ΑΠΟΣΤΟΛΟΥ Κ. ΓΕΓΙΟΥ

ΕΠΙΚ. ΚΑΘΗΓΗΤΟΥ
ΠΑΝΕΠΙΣΤΗΜΙΟΥ ΙΩΑΝΝΙΝΩΝ

ΑΡΙΘΜΗΤΙΚΕΣ ΜΕΘΟΔΟΙ

ΤΟΜΟΣ Ι

ΙΩΑΝΝΙΝΑ

ΒΙΒΛΙΟΘΗΚΗ
ΠΑΝΕΠΙΣΤΗΜΙΟΥ ΙΩΑΝΝΙΝΩΝ
026000021611

318/86

ΜΑΘΗΜΑΤΙΚΟ ΤΜΗΜΑ
ΒΙΒΛΙΟΘΗΚΗ
ΙΩΑΝΝΙΝΩΝ

515.1
Γ.2

ΑΠΟΣΤΟΛΟΥ Κ. ΓΕΓΙΟΥ

ΕΠΙΚ. ΚΑΘΗΓΗΤΟΥ
ΠΑΝΕΠΙΣΤΗΜΙΟΥ ΙΩΑΝΝΙΝΩΝ

ΑΡΙΘΜΗΤΙΚΕΣ ΜΕΘΟΔΟΙ

ΤΟΜΟΣ Ι

ΙΩΑΝΝΙΝΑ
1986

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

318/86

ΤΜΗΜΑ ΒΙΒΛΙΟΘΗΚΗ ΕΛΛΗΝΙΚΩΝ

Απαγορεύεται η μετάφραση ή αναπαραγωγή σε οποιαδήποτε μορφή, οποιουδήποτε τμήματος του παρόντος βιβλίου, χωρίς την γραπτή άδεια του συγγραφέα.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ

ΠΡΟΛΟΓΟΣ

Με την ονομασία "Υπολογιστικά (ή Αριθμητικά) Μαθηματικά" ή "Αριθμητική Ανάλυση" είναι γνωστός ο κλάδος των εφαρμοσμένων Μαθηματικών, που έχει σαν βασικό σκοπό την παραγωγή, ανάλυση και χρήση αποτελεσματικών,στην πράξη,αριθμητικών (υπολογιστικών) μεθόδων (αλγορίθμων),για την λύση μαθηματικών προβλημάτων. Πρέπει να σημειωθεί ότι τα υπολογιστικά Μαθηματικά είναι στενά συνδεδεμένα με τους υπολογιστές. Μία αριθμητική μέθοδος είναι μία διαδικασία με την οποία, μέσω ενός υπολογιστού (ο οποίος αναλαμβάνει την εκτέλεση των αριθμητικών υπολογισμών), αναζητούμε την αριθμητική και επομένως κατασκευαστική λύση ενός μαθηματικού προβλήματος, δηλαδή, παίρνουμε αποτελέσματα υπό μορφή αριθμών, χρησιμοποιώντας αριθμητικά δεδομένα. Μία τέτοια μέθοδος θα είναι πρακτικά αποτελεσματική και επομένως χρήσιμη,εάν δίνει αποτελέσματα με κάποια συγκεκριμένη ακρίβεια και εντός λογικού χρόνου.

Έργο της Αριθμητικής Αναλύσεως είναι επίσης η μελέτη και εκτίμηση των σφαλμάτων,που υπεισέρχονται στους υπολογισμούς. Οι ρίζες αυτής βρίσκονται στην αρχαιότητα, αλλά ως εποχή της συγχρόνου Αριθμητικής Αναλύσεως θεωρείται η περίοδος από το 1950 μέχρι σήμερα, κατά τη διάρκεια της οποίας η αλματώδης ανάπτυξη των αυτόματων ηλεκτρονικών υπολογιστών επηρέασε σημαντικά την πορεία αυτής. Είναι γεγονός αναμ-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

φισβήτητο ότι τα τελευταία χρόνια έχει αυξηθεί σημαντικά η χρησιμότητα των αριθμητικών μεθόδων σε πολλές επιστήμες,όπου παρουσιάζονται πρακτικά προβλήματα (φυσικές, κοινωνικές επιστήμες, κ.λ.π.).

Το παρόν βιβλίο εξετάζει τα βασικά θέματα που περιέχει συνήθως ένα εισαγωγικό βιβλίο Αριθμητικής Αναλύσεως και δεν απαιτεί προηγούμενη γνώση του αντικειμένου του. Γράφτηκε με σκοπό να χρησιμοποιηθεί για τη διδασκαλία των σχετικών μαθημάτων του προγράμματος σπουδών στους φοιτητές (κυρίως, των δύο τελευταίων ετών) του τμήματος Μαθηματικών του Πανεπιστημίου Ιωαννίνων, αλλά και να χρησιμεύσει σαν βασικό βοήθημα σε εκείνους που θα ασχοληθούν ειδικώτερα με τις αριθμητικές μεθόδους.

Θεωρώ καθήκον μου να ευχαριστήσω την σύζυγό μου Λουκρητία Παπαγεωργίου – Γέγιου, μέλος του Ε.Δ.Τ.Π. του Δ' Τομέα του τμήματος Μαθηματικών του Πανεπιστημίου Ιωαννίνων, για την επιμελημένη και προσεκτική δακτυλογράφηση των χειρογράφων και γενικά για την συνολική προσπάθειά της για την καλύτερη εμφάνιση του παρόντος.

Ιωάννινα, Μάϊος 1986 Α. Κ. Γέγιος

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Στους γονείς μου
και στην οικογένειά μου

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΡΟΛΟΓΟΣ

ΠΙΝΑΚΑΣ ΟΡΙΣΜΕΝΩΝ ΣΥΜΒΟΛΩΝ

1. ΣΦΑΛΜΑΤΑ

1.1 Σφάλμα — Απόλυτο και σχετικό σφάλμα 1

1.2 Πηγές και είδη σφαλμάτων 4

1.3 Παράσταση των αριθμών — Σημαντικά ψηφία — Στρογγύλευση 7

1.4 Διάδοση και δημιουργία σφαλμάτων κατά τους υπολογισμούς 12

1.5 Ολικό σφάλμα στην αριθμητική λύση ενός μαθηματικού προβλήματος 20

ΑΣΚΗΣΕΙΣ 21

2. ΠΑΡΕΜΒΟΛΗ 24

2.1 Εισαγωγή 24

2.2 Πολυωνυμική παρεμβολή 25

2.2.1 Παρεμβολή κατά Lagrange 27

2.2.2 Εκτίμηση του σφάλματος στην πολυωνυμική παρεμβολή 30

2.2.3 Η μέθοδος ή αλγόριθμος του Aitken 32

2.2.4 Αντίστροφη παρεμβολή 36

2.3 Διηρημένες διαφορές 38

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

2.3.1 Τύποι παρεμβολής του Newton με διηρημένες διαφορές 41
2.4 Πεπερασμένες διαφορές — Πίνακες πεπερασμένων διαφορών 44
2.4.1 Διάδοση σφαλμάτων σε πίνακες διαφορών 52
2.5 Τελεστές διαφορών 56
2.5.1 Παραγοντικές δυνάμεις 62
2.6 Τύποι πολυωνυμικής παρεμβολής για ισαπέχοντα σημεία παρεμβολής 64
2.6.1 Τύποι των προς τα εμπρός διαφορών των Newton — Gregory 65
2.6.2 Τύποι των προς τα πίσω διαφορών των Newton — Gregory 69
2.6.3 Τύποι παρεμβολής των Gauss, Stirling και Bessel 71
2.7 Παρεμβολή κατά Hermite 75
ΑΣΚΗΣΕΙΣ 76

3. ΑΡΙΘΜΗΤΙΚΗ ΠΑΡΑΓΩΓΙΣΗ 83

3.1 Εισαγωγή 83
3.2 Τύποι αριθμητικής παραγωγίσεως με την μέθοδο των πολυωνύμων παρεμβολής 84
3.2.1 Τύποι αριθμητικής παραγωγίσεως για ισαπέχοντα σημεία παρεμβολής 86
3.3 Τύποι αριθμητικής παραγωγίσεως με την μέθοδο των προσδιοριστέων συντελεστών 93
3.4 Σφάλματα στην αριθμητική παραγώγιση 96
ΑΣΚΗΣΕΙΣ 98

4. ΣΥΝΟΛΑ ΟΡΘΟΓΩΝΙΩΝ ΠΟΛΥΩΝΥΜΩΝ 102

4.1 Εισαγωγή 102
4.2 Ιδιότητες συνόλων ορθογωνίων πολυωνύμων 105
4.3 Τα πολυώνυμα Legendre 107
4.4 Τα πολυώνυμα Chebyshev 113
ΑΣΚΗΣΕΙΣ 119

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ

5 ΑΡΙΘΜΗΤΙΚΗ ΟΛΟΚΛΗΡΩΣΗ 124

5.1 Εισαγωγή 124

5.2 Κλειστοί τύποι αριθμητικής ολοκληρώσεως των Newton — Cotes 125

5.3 Σφάλματα αποκοπής των κλειστών τύπων των Newton — Cotes 131

5.4 Ανοικτοί — ημιανοικτοί τύποι αριθμητικής ολοκληρώσεως των Newton — Cotes 136

5.5 Extrapolation — Αριθμητική ολοκλήρωση του Romberg 141

5.6 Μέθοδοι αριθμητικής ολοκληρώσεως του Gauss 146

5.6.1 Εισαγωγή 146

5.6.2 Εύρεση των τύπων ολοκληρώσεως του Gauss 148

5.6.3 Τύποι ολοκληρώσεως των Gauss — Legendre 153

5.6.4 Τύποι ολοκληρώσεως των Gauss — Chebyshev 155

5.6.5 Σφάλμα αποκοπής των τύπων ολοκληρώσεως του Gauss 159

5.7 Αριθμητική ολοκλήρωση του Chebyshev 162

5.8 Η μέθοδος των προσδιοριστέων συντελεστών 167

ΑΣΚΗΣΕΙΣ 169

6 ΑΡΙΘΜΗΤΙΚΗ ΕΠΙΛΥΣΗ ΕΞΙΣΩΣΕΩΝ 180

6.1 Εισαγωγή 180

6.2 Εντοπισμός των ριζών 181

6.3 Η μέθοδος της διχοτομήσεως 183

6.4 Η επαναληπτική μέθοδος 186

6.4.1 Σύγκλιση της επαναληπτικής μεθόδου 188

6.4.2 Γεωμετρική ερμηνεία της επαναληπτικής μεθόδου 199

6.5 Η μέθοδος Newton — Raphson 202

6.5.1 Γεωμετρική ερμηνεία της μεθόδου Newton — Raphson 203

6.5.2 Σύγκλιση και τάξη της μεθόδου Newton — Raphson 204

6.5.3 Εφαρμογές — Παραδείγματα 209

6.5.4 Η μέθοδος Newton — Raphson για την εύρεση μιγαδικών ριζών 211

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

X Θεωρημα 6.11.

6.6 Μέθοδοι παρεμβολής για τον προσδιορισμό ριζών **212**

6.6.1 Η μέθοδος της εσφαλμένης θέσεως (Regula falsi) **212**

6.6.2 Τάξη συγκλίσεως της μεθόδου Regula falsi **216**

6.6.3 Γεωμετρική ερμηνεία της μεθόδου Regula falsi **217**

6.6.4 Η μέθοδος της τέμνουσας **219**

6.6.5 Σύγκλιση της μεθόδου της τέμνουσας **220**

6.7 Επιτάχυνση της συγκλίσεως — Η Δ^2 — μέθοδος του Aitken — Η μέθοδος του Steffensen **225**

ΑΣΚΗΣΕΙΣ **230**

7. ΠΟΛΥΩΝΥΜΙΚΕΣ ΕΞΙΣΩΣΕΙΣ 239

7.1 Εισαγωγή **239**

7.2 Γενικές ιδιότητες των πολυωνύμων — Φράγματα ριζών **240**

7.3 Η μέθοδος της συνθετικής διαιρέσεως (Μέθοδος του Horner) **248**

7.4 Η μέθοδος Newton — Bairstow **257**

7.5 Άλλες μέθοδοι για πολυωνυμικές εξισώσεις **262**

7.6 Γενικές οδηγίες και παρατηρήσεις για την εύρεση των ριζών πολυωνυμικών εξισώσεων **263**

ΑΣΚΗΣΕΙΣ **266**

8. ΣΥΣΤΗΜΑΤΑ ΜΗ ΓΡΑΜΜΙΚΩΝ ΕΞΙΣΩΣΕΩΝ 269

8.1 Εισαγωγή **269**

8.2 Γραφική παράσταση του (8.1) **270**

8.3 Η επαναληπτική μέθοδος **270**

8.4 Η γενικευμένη μέθοδος Newton — Raphson ή μέθοδος Newton **278**

8.5 Η μέθοδος της κλίσεως (Gradient, Steepest descent) **282**

ΑΣΚΗΣΕΙΣ **286**

9. ΣΤΟΙΧΕΙΑ ΘΕΩΡΙΑΣ ΠΙΝΑΚΩΝ ΚΑΙ ΓΡΑΜΜΙΚΗΣ ΑΛΓΕΒΡΑΣ. 289

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

9.1 Εισαγωγή 289
9.2 Διανύσματα και πίνακες 289
9.3 Ορίζουσες 294
9.4 Συστήματα γραμμικών εξισώσεων 295
9.5 Ιδιοτιμές και ιδιοδιανύσματα 297
9.6 Ορισμοί και θεωρήματα για ειδικούς και μη πίνακες 305
9.7 Norms διανύσματος και norms πίνακα 313
ΑΣΚΗΣΕΙΣ 326

10. ΑΡΙΘΜΗΤΙΚΗ ΕΠΙΛΥΣΗ ΓΡΑΜΜΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ 336

10.1 Εισαγωγή 336
10.2 Άμεσες μέθοδοι 338
10.2.1 Η μέθοδος απαλοιφής του Gauss 338
10.2.2 Υπολογισμός ορίζουσας 349
10.2.3 Επίλυση πολλών γραμμικών συστημάτων με τον ίδιο πίνακα συντελεστών 350
10.2.4 Εύρεση του αντιστρόφου A^{-1} ενός πίνακα A ($\det(A) \neq 0$) 351
10.2.5 Η μέθοδος απαλοιφής του Jordan 353
10.2.6 Υπολογισμός του πλήθους των αριθμητικών πράξεων στις μεθόδους απαλοιφής Gauss και Jordan 357
10.2.7 Μέθοδοι τριγωνικοποιήσεως (μέθοδοι τριγωνικής παραγοντοποιήσεως) 361
10.2.8 Η μέθοδος Cholesky για θετικά ορισμένους πίνακες 365
10.2.9 Εκτίμηση σφάλματος στη λύση γραμμικού συστήματος 366
10.2.10 Ασταθή γραμμικά συστήματα 368
10.2.11 Επαναληπτική βελτίωση μιας προσεγγιστικής λύσεως 372
10.3 Επαναληπτικές μέθοδοι 373
10.3.1 Βασικές επαναληπτικές μέθοδοι 374
10.3.2 Η μέθοδος Jacobi 380
10.3.3 Η μέθοδος Gauss — Seidel 384
10.3.4 Η μέθοδος SOR (Successive OverRelaxation) 389

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

10.3.5 Άλλες επαναληπτικές μέθοδοι **392**
10.3.6 Επιτάχυνση της συγκλίσεως **392**
ΑΣΚΗΣΕΙΣ **399**

11. ΥΠΟΛΟΓΙΣΜΟΣ ΙΔΙΟΤΙΜΩΝ ΚΑΙ ΙΔΙΟΔΙΑΝΥΣΜΑΤΩΝ 409

11.1 Εισαγωγή **409**
11.2 Η μέθοδος της δυνάμεως **410**
11.3 Η μέθοδος της δυνάμεως στον πίνακα A − qI **415**
11.4 Υπολογισμός της απολύτως μικρότερης ιδιοτιμής ενός πίνακα A **416**
11.5 Η αντίστροφη μέθοδος της δυνάμεως **417**
11.6 Επέκταση της μεθόδου της δυνάμεως. Η διαδικασία deflation **419**
11.7 Η LR μέθοδος **425**
11.8 Η QR μέθοδος **427**
11.9 Ευστάθεια προβλημάτων ιδιοτιμών **428**
ΑΣΚΗΣΕΙΣ **431**

12. ΠΟΛΥΩΝΥΜΙΚΗ ΠΡΟΣΕΓΓΙΣΗ 437

12.1 Εισαγωγή **437**
12.2 Πολυωνυμική προσέγγιση **439**
12.3 Άριστη ομοιόμορφη προσέγγιση **441**
12.4 Πολυώνυμο ελαχίστων τετραγώνων **453**
ΑΣΚΗΣΕΙΣ **463**

ΒΙΒΛΙΟΓΡΑΦΙΑ 468

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ

ΠΙΝΑΚΑΣ ΟΡΙΣΜΕΝΩΝ ΣΥΜΒΟΛΩΝ

Σύμβολο	Έννοια
$\approx$	κατά προσέγγιση ίσον, περίπου ίσον
$\ll$	πολύ μικρότερο από
$\lesssim$	μικρότερο ή περίπου ίσον
▲	τέλος θεωρήματος, παραδείγματος, πορίσματος, ορισμού
$i = 0(1)n$	$i = 0,1,2,\ldots,n$
$\ln x$	φυσικός λογάριθμος του x
$\in$	ανήκει
$\mathbb{R}$	το σύνολο των πραγματικών αριθμών
$\mathbb{C}$	το σύνολο των μιγαδικών αριθμών
$C[a,b]$	το σύνολο των συνεχών συναρτήσεων στο διάστημα $[a,b]$
$\forall$	για κάθε
$\Rightarrow$	έπεται
$(2n)!!$	$2 \cdot 4 \cdot 6 \cdot \ldots \cdot (2n)$
$(2n+1)!!$	$1 \cdot 2 \cdot 3 \cdot \ldots \cdot (2n+1)$
π	$3.14159\ldots$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

1

ΣΦΑΛΜΑΤΑ

1.1 Σφάλμα – Απόλυτο και σχετικό σφάλμα.

Έστω x κάποιος πραγματικός αριθμός, δηλαδή, x είναι η ακριβής ή αληθής τιμή του. Γενικά, στους υπολογισμούς, στη θέση του x χρησιμοποιείται ένας προσεγγιστικός αριθμός $\bar{x}$, ο οποίος διαφέρει "ελαφρώς" από τον x. Λέμε ότι ο $\bar{x}$ αποτελεί μία προσέγγιση του x και γράφουμε $x \approx \bar{x}$. Για παράδειγμα, αν $x = \sqrt{3}$, τότε ο $\bar{x} = 1.732$ είναι μία προσέγγιση του $\sqrt{3}$, $\sqrt{3} \approx 1.732$.

Καλούμε σφάλμα ε του προσεγγιστικού αριθμού $\bar{x}$ τη διαφορά

$$\varepsilon = \bar{x} - x . \tag{1.1}$$

Τη διαφορά $x - \bar{x} = -\varepsilon$ καλούμε διόρθωση. Αν το σφάλμα ε είναι γνωστό, τότε από την (1.1) έχουμε $x = \bar{x} - \varepsilon$. Επειδή σε πολλές περιπτώσεις το πρόσημο του σφάλματος δεν είναι γνωστό, χρησιμοποιείται ο όρος απόλυτο σφάλμα του $\bar{x}$ που είναι η απόλυτη τιμή του ε, δηλαδή,

$$|\varepsilon| = |\bar{x} - x| = \text{απόλυτο σφάλμα του } \bar{x} . \tag{1.2}$$

Στην πράξη, συνήθως, ενδιαφερόμαστε για την εύρεση άνω φραγμά-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

των του απολύτου σφάλματος. Αν M είναι ένα άνω φράγμα για το $|\varepsilon|$, δηλαδή, $|\varepsilon| \leqq M$, τότε λόγω της (1.2) προκύπτει ότι

$$\bar{x} - M \leqq x \leqq \bar{x} + M . \tag{1.3}$$

Μερικές φορές, αντί της $|\varepsilon| = |\bar{x} - x| \leqq M$ χρησιμοποιείται ο συμβολισμός $x = \bar{x} \pm M$.

<u>Παράδειγμα 1.1</u> Αν $\sqrt{2} \approx 1.41$, να βρεθεί ένα άνω φράγμα για το απόλυτο σφάλμα του 1.41 αν είναι γνωστό ότι $1.41 < \sqrt{2} < 1.42$.

<u>Λύση</u> Είναι $|\varepsilon| = |\bar{x} - x| = |1.41 - \sqrt{2}| = \sqrt{2} - 1.41 < 1.42 - 1.41 = 0.01$. Άρα, ένα άνω φράγμα του $|\varepsilon|$ είναι ο αριθμός 0.01. ▲

Το απόλυτο σφάλμα δεν είναι αρκετό για να μας περιγράψει την ακρίβεια μιας μέτρησης ή ενός υπολογισμού. Πράγματι, ας υποθέσουμε ότι μετρήσαμε την απόσταση x δύο σημείων και την απόσταση y δύο άλλων σημείων και βρήκαμε αντίστοιχα $\bar{x} = 1005.1$ m και $\bar{y} = 7.9$ m με το ίδιο μέγιστο απόλυτο σφάλμα 0.1 m. Ισχύουν δηλαδή $|\bar{x} - x| \leqq 0.1$ m και $|\bar{y} - y| \leqq 0.1$ m. Επομένως, λόγω της (1.3) έχουμε $1005.0 \text{ m} \leqq x \leqq 1005.2$ m, $7.8 \text{ m} \leqq y \leqq 8.0$ m και προφανώς η μέτρηση για την απόσταση x είναι πολύ καλύτερη (πιο ακριβής), αφού το μέγιστο απόλυτο σφάλμα 0.1 m είναι πολύ μικρό σε σύγκριση με την απόσταση x. Για το λόγο αυτό εισάγεται η έννοια του σχετικού σφάλματος.

Καλούμε σχετικό σφάλμα του προσεγγιστικού αριθμού $\bar{x}$ τον λόγο δ του απόλυτου σφάλματος του $\bar{x}$ δια της απολύτου τιμής του ακριβούς αριθμού x $(x \neq 0)$, δηλαδή,

$$\delta = \frac{|\varepsilon|}{|x|} . \tag{1.4}$$

Το σχετικό σφάλμα είναι ανεξάρτητο των μονάδων μετρήσεως ενώ το απόλυτο σφάλμα εξαρτάται από τις μονάδες που χρησιμοποιήθηκαν για τη μέτρηση. Μερικές φορές πολλαπλασιάζουμε το σχετικό σφάλμα επί 100% για να το έχουμε σε αναλογία επί τοις εκατό (%). Βάσει του ορισμού (1.4) μπορούμε να θέσουμε $\bar{x} = x(1 + v)$, όπου $|v| = \delta =$ σχετικό σφάλμα του

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$\bar{x}$. Ακόμη, είναι $|\bar{x}| = |x + \bar{x} - x| \leqq |x| + |\varepsilon|$ και επομένως $|x| \geqq$ $\geqq |\bar{x}| - |\varepsilon|$. Εάν $|\varepsilon| \leqq M << |\bar{x}|$, πράγμα που ισχύει συνήθως, τότε από την (1.4) παίρνουμε

$$\delta = \frac{|\varepsilon|}{|x|} \leqq \frac{|\varepsilon|}{|\bar{x}| - |\varepsilon|} \approx \frac{|\varepsilon|}{|\bar{x}|} \quad \text{ή}$$

$$\delta \lessapprox \frac{|\varepsilon|}{|\bar{x}|} \quad , \quad \bar{x} \neq 0 \ . \tag{1.5}$$

Αντί της (1.5) χρησιμοποιούμε συνήθως την

$$\delta \approx \frac{|\varepsilon|}{|\bar{x}|} \ , \quad \bar{x} \neq 0 \tag{1.6}$$

για την εκτίμηση του σχετικού σφάλματος, αφού στην (1.4) ο x είναι γενικά άγνωστος.

<u>Θεώρημα 1.1</u> Αν $\delta \leqq N << 1$ τότε ισχύουν:

i) $|\varepsilon| \leqq \dfrac{|\bar{x}|N}{1 - N} \approx |\bar{x}|N$

ii) $\min\{\bar{x}(1 - N), \bar{x}(1 + N)\} \lessapprox x \lessapprox \max\{\bar{x}(1 - N), \bar{x}(1 + N)\}$.

<u>Απόδειξη</u> Από την (1.1) παίρνουμε $|x| = |\bar{x} - \varepsilon| \leqq |\bar{x}| + |\varepsilon|$, οπότε από την (1.4) προκύπτει $|\varepsilon| = \delta|x| \leqq |x|N \leqq (|\bar{x}| + |\varepsilon|)N$. Άρα, $(1 - N)|\varepsilon| \leqq |\bar{x}|N$ ή

$$|\varepsilon| \leqq \frac{|\bar{x}|N}{1 - N} \approx |\bar{x}|N \quad , \tag{1.7}$$

δηλαδή, $|\varepsilon| \lessapprox M = |\bar{x}|N$. Λόγω της (1.3) θα είναι

$$\bar{x} - |\bar{x}|N \lessapprox x \lessapprox \bar{x} + |\bar{x}|N \quad \text{ή ισοδύναμα}$$

$$\min\{\bar{x}(1 - N), \bar{x}(1 + N)\} \lessapprox x \lessapprox \max\{\bar{x}(1 - N), \bar{x}(1 + N)\}. \tag{1.8}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Αντί της (1.8) μπορούμε να γράψουμε $x \approx \bar{x}(1 + \mu)$, $|\mu| \leqq N$. ▲

<u>Παράδειγμα 1.2</u> Σαν συνέχεια του παραδείγματος 1.1, να βρεθεί ένα άνω φράγμα για το σχετικό σφάλμα του $\bar{x} = 1.41$.

<u>Λύση</u> Βρέθηκε ότι $|\varepsilon| < 0.01 = M$. Έτσι έχουμε

$$\delta \leqq \frac{|\varepsilon|}{|\bar{x}| - |\varepsilon|} < \frac{M}{|\bar{x}| - M} = \frac{0.01}{1.41 - 0.01} = \frac{0.01}{1.40} \approx 0.007 = 0.7\%. \blacktriangle$$

<u>Σημείωση</u> Οι ορισμοί (1.1), (1.2), (1.4) επεκτείνονται και στην περίπτωση που αντί για αριθμούς τα x , $\bar{x}$ παριστούν διανύσματα ή πίνακες με την διαφορά ότι οι απόλυτες τιμές αντικαθίστανται από κάποια norm, δηλαδή, $\|\varepsilon\| = \|\bar{x} - x\|$, $\delta = \dfrac{\|\varepsilon\|}{\|x\|}$ (βλέπε κεφάλαιο 9).

1.2 Πηγές και είδη σφαλμάτων

Οι λόγοι για τους οποίους η λύση, που υπολογίζεται τελικά για ένα πραγματικό (π.χ. φυσικό) πρόβλημα με τη βοήθεια αριθμητικών μεθόδων, θα διαφέρει πάντοτε από την ακριβή λύση, εξηγούνται από την παρουσία των παρακάτω πηγών σφαλμάτων. Μία πρώτη πηγή σφαλμάτων βρίσκεται στη μετατροπή του αρχικού πραγματικού προβλήματος σε ένα μαθηματικό πρόβλημα, δηλαδή, στην κατασκευή του μαθηματικού μοντέλου μιας δοθείσης πραγματικής κατάστασης. Η μετατροπή αυτή περιέχει συνήθως απλοποιήσεις και παραλείψεις. Τα μαθηματικά μοντέλα είναι κατά το πλείστον εξιδανικευμένα. Τα σφάλματα της πηγής αυτής καλούνται σφάλματα του μαθηματικού προβλήματος.

Μία δεύτερη κατηγορία σφαλμάτων είναι τα σφάλματα δεδομένων ή αρχικά σφάλματα. Αυτά παρουσιάζονται επειδή τα μαθηματικά προβλήματα περιέχουν συνήθως παραμέτρους, όπως μήκος, χρόνος, μάζα, θερμοκρασία κ.λ.π., των οποίων οι τιμές μπορούν να βρεθούν μόνο από μετρήσεις ή εκτιμήσεις. Τα δεδομένα από μετρήσεις και παρατηρήσεις περιέχουν πολύ συχνά ανακρίβειες (σφάλματα) που περιορίζουν ασφαλώς την ακρίβεια των

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

αποτελεσμάτων του προβλήματος. Σε πολλά προβλήματα η μαθηματική (ακριβής) λύση είναι πολύ "ευαίσθητη" σε μικρομεταβολές των τιμών ορισμένων ή όλων των παραμέτρων - δεδομένων. Δηλαδή, μικρές μεταβολές των δεδομένων δημιουργούν μεγάλη μεταβολή στη μαθηματική λύση. Τα προβλήματα αυτά καλούνται ασταθή (ill conditioned, unstable). Στην αντίθετη περίπτωση καλούνται ευσταθή (well conditioned, stable). Ένα παράδειγμα ασταθούς προβλήματος είναι το παρακάτω γραμμικό σύστημα

$$\begin{cases} x + 2y = 3 \\ 0.499x + 1.001y = 1.5 \end{cases},$$

του οποίου η ακριβής λύση είναι $(x,y) = (1,1)$. Εάν ο συντελεστής του x στην δεύτερη εξίσωση μεταβληθεί από 0.499 σε 0.5 τότε η ακριβής λύση του προκύπτοντος συστήματος είναι $(x,y) = (3,0)$.

Μία άλλη πηγή σφαλμάτων αποτελεί το γεγονός ότι η παράσταση των αριθμών είναι πεπερασμένη στους διάφορους ηλεκτρονικούς υπολογιστές. Εφ' όσον υπάρχουν αριθμοί (είτε σαν δεδομένα είτε σαν αποτελέσματα αριθμητικών πράξεων) που για την παράστασή τους απαιτείται ακόμη και άπειρο πλήθος ψηφίων (π.χ. στο δεκαδικό σύστημα είναι

$$\pi = 3.14159265358979323846264 3\ldots ,$$

$$\frac{1}{3} = 0.333333\ldots),$$

είμαστε αναγκασμένοι στην πράξη για την εκτέλεση των υπολογισμών να προσεγγίζουμε τους πολυψήφιους (με πεπερασμένο ή άπειρο πλήθος ψηφίων) αριθμούς, δηλαδή, να τους αντικαθιστούμε με άλλους οι οποίοι έχουν ένα περιορισμένο πλήθος ψηφίων που εξαρτάται από το μέσον που χρησιμοποιούμε. Τα σφάλματα αυτής της προσέγγισης που καλείται στρογγύλευση είναι γνωστά σαν σφάλματα στρογγυλεύσεως και έχουν την ιδιότητα να συσσωρεύονται κατά την διάρκεια των διαδοχικών υπολογισμών, με αποτέλεσμα πολλές φορές να επηρεάζουν σημαντικά τη λύση του προβλήματος που μας ενδιαφέρει. Βασικό ρόλο στον περιορισμό και έλεγχο των σφαλμάτων στρογγυλεύσεως παίζει η αριθμητική μέθοδος που χρησιμοποιούμε. Ο τρόπος με τον οποίο γίνεται η στρογγύλευση των αριθμών

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

θα εξηγηθεί στην επόμενη παράγραφο. Πρέπει να σημειωθεί ότι τα σφάλματα των δεδομένων ενός προβλήματος μπορεί να είναι αποκλειστικά σφάλματα στρογγυλεύσεως.

Ένα τελευταίο σημαντικό είδος σφαλμάτων είναι αποτέλεσμα του γεγονότος ότι διαδικασίες που στην Μαθηματική Ανάλυση απαιτούν ένα άπειρο πλήθος βημάτων πρέπει στην πράξη, δηλαδή, κατά την εφαρμογή μιας αριθμητικής μεθόδου να έχουν πεπερασμένη παράσταση. Τα σφάλματα του είδους αυτού καλούνται σφάλματα αποκοπής ή ακόμη σφάλματα προσεγγίσεως ή διακεκριμενοποίησης ή σφάλματα της μεθόδου. Τέτοια σφάλματα συμβαίνουν π.χ. όταν το πρόβλημα του αθροίσματος μιας απειροσειράς αντικατασταθεί με το απλούστερο πρόβλημα του αθροίσματος μόνον ενός πεπερασμένου πλήθους όρων της σειράς, τα ορισμένα ολοκληρώματα προσεγγιστούν με πεπερασμένα αθροίσματα, τα διαφορικά πηλίκα με πηλίκα διαφορών κ.λ.π. . Σαν ενδεικτικό παράδειγμα ας θεωρήσουμε το ανάπτυγμα σε σειρά Taylor της συναρτήσεως $f(x)$ γύρω από το σημείο x_0

$$f(x_0+h) = f(x_0) + hf'(x_0) + \frac{h^2}{2!} f''(x_0) + \ldots + \frac{h^m}{m!} f^{(m)}(x_0) + \ldots \quad . \tag{1.9}$$

Από την (1.9) είναι φανερό ότι για κάποιο $h > 0$ η τιμή $f(x_0+h)$ προσδιορίζεται ακριβώς γνωρίζοντας τις τιμές $f^{(k)}(x_0)$, $k = 0,1,2,\ldots$ (υποθέτουμε ότι όλες οι παράγωγοι υπάρχουν) με την εύρεση ενός αθροίσματος απείρων όρων. Εάν για την εύρεση της $f(x_0+h)$ χρησιμοποιήσουμε μόνον τους $m+1$ πρώτους όρους του αναπτύγματος τότε η (1.9) γράφεται

$$f(x_0+h) = f(x_0) + hf'(x_0) + \ldots + \frac{h^m}{m!} f^{(m)}(x_0) - \varepsilon(h), \tag{1.10}$$

όπου $\varepsilon(h)$ το σφάλμα αποκοπής. Για το σφάλμα αποκοπής λόγω των (1.9) και (1.10) έχουμε

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\varepsilon(h) = -\frac{h^{m+1}}{(m+1)!}f^{(m+1)}(x_0) - \frac{h^{m+2}}{(m+2)!}f^{(m+2)}(x_0) - \dots \quad (1.11)$$

Επομένως, για μικρό αρκετά h $(h \to 0)$ το σφάλμα αποκοπής είναι ανάλογο του h^{m+1} (με την προυπόθεση ότι $f^{(m+1)}(x_0) \neq 0$), δηλαδή

$$\varepsilon(h) \approx \alpha h^{m+1} \quad .$$

Στην περίπτωση αυτή λέμε ότι το σφάλμα αποκοπής είναι της τάξεως του h^{m+1} και γράφουμε $\varepsilon(h) = O(h^{m+1})$.

Σε πολλές περιπτώσεις προσπαθώντας να ελαττώσουμε το σφάλμα αποκοπής προκαλούμε αύξηση του σφάλματος στρογγυλεύσεως. Αυτό συμβαίνει γιατί γενικά το ποσό του σφάλματος στρογγυλεύσεως είναι ευθέως ανάλογο της πολυπλοκότητας μιας αριθμητικής μεθόδου ενώ το ποσό του σφάλματος αποκοπής είναι αντιστρόφως ανάλογο.

Από την προηγούμενη ανάλυση των σφαλμάτων είναι φανερό ότι τα αριθμητικά αποτελέσματα, δηλαδή,η αριθμητική λύση ενός συγκεκριμένου μαθηματικού προβλήματος επηρεάζεται λιγώτερο ή περισσότερο από τα τρία τελευταία είδη σφαλμάτων. Για τον λόγο αυτό ένας από τους βασικώτερους σκοπούς της Αριθμητικής Αναλύσεως είναι και η εκτίμηση της ακρίβειας των αποτελεσμάτων των υπολογισμών, δηλαδή η εκτίμηση του συνολικού σφάλματος που περιέχουν οι απαντήσεις και επί πλέον η ελαχιστοποίηση αυτού. Τέλος, υποθέτουμε ότι τυχόν ανθρώπινα λάθη που οφείλονται π.χ. σε αμέλεια ή απροσεξία μπορούν να βρεθούν και να διορθωθούν χωρίς κανένα επακόλουθο.

1.3 Παράσταση των αριθμών – Σημαντικά ψηφία – Στρογγύλευση

Όπως είναι γνωστό, εάν $q > 1$ είναι κάποιος θετικός ακέραιος, τότε κάθε πραγματικός αριθμός x μπορεί να γραφεί υπό την μορφή

$$x = (\text{πρόσημο του } x)(\alpha_n q^n + \alpha_{n-1} q^{n-1} + \dots + \alpha_0 q^0 + \alpha_{-1} q^{-1} + \alpha_{-2} q^{-2} + \dots), \quad (1.12)$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

όπου τα α_i είναι τα ψηφία του x $(0 \leqq \alpha_i \leqq q-1)$ στο σύστημα αριθμήσεως με βάση q. Η (1.12) γράφεται συντόμως ως εξής:

$$x = (\text{πρόσημο του } x)(\alpha_n \alpha_{n-1} \ldots \alpha_0 \bullet \alpha_{-1} \alpha_{-2} \ldots)_q \,. \qquad (1.13)$$

Αν $q = 10$ έχουμε το γνωστό δεκαδικό σύστημα αριθμήσεως που χρησιμοποιείται από τους ανθρώπους. Για τεχνικούς λόγους οι περισσότεροι συγχρονοι ηλεκτρονικοί υπολογιστές παριστούν τους αριθμούς εσωτερικά σε συστήματα με βάση $q = 2$ (δυαδικό) ή $q = 8$ (οκταδικό) ή $q = 16$ (δεκαεξαδικό). Για παράδειγμα, ο αριθμός $18.5 = (18.5)_{10}$ γράφεται

$$18.5 = 1\cdot 2^4 + 0\cdot 2^3 + 0\cdot 2^2 + 1\cdot 2^1 + 0\cdot 2^0 + 1\cdot 2^{-1}$$

και επομένως είναι $18.5 = (10010.1)_2$.

Αν υποθέσουμε ότι ο υπολογιστής που διαθέτουμε χρησιμοποιεί π.χ. το δυαδικό σύστημα, τότε οι αριθμοί (δεδομένα) που στέλνουμε σ' αυτόν και που φυσιολογικά έχουν δεκαδική παράσταση $(q = 10)$ πρέπει να μετατραπούν στην δυαδική μορφή. Επίσης, τα τελικά αποτελέσματα του υπολογιστού μετατρέπονται πάλι στο δεκαδικό σύστημα προτού γίνουν διαθέσιμα σε μας. Πρέπει να σημειωθεί ότι ο αριθμός των ψηφίων ενός αριθμου εξαρτάται από την βάση q του συστήματος στο οποίο γράφεται . Πράγματι, μπορεί ένας αριθμός στο δεκαδικό σύστημα να έχει λίγα ψηφία ενώ στο δυαδικό σύστημα να έχει άπειρο πλήθος ψηφίων. Είναι π.χ.

$$0.1 = (0.0001100110011\ldots)_2 \,.$$

Αφ' ετέρου μπορεί να αποδειχθεί ότι εάν ένας αριθμός έχει ένα πεπερασμένο αριθμό ℓ ψηφίων στο δυαδικό σύστημα τότε θα έχει το πολύ ℓ ψηφία και στο δεκαδικό σύστημα. Από τα παραπάνω προκύπτει το συμπέρασμα ότι αφού οι υπολογιστές χρησιμοποιούν αριθμούς που έχουν περιορισμένο πλήθος ψηφίων η πιθανότητα για σφάλματα στρογγυλεύσεως είναι μεγάλη.

Στη συνέχεια, για καλύτερη κατανόηση, αναφερόμαστε μόνον στο δεκαδικό σύστημα, δηλαδή, υποθέτουμε ότι ο υπολογιστής μας χρησιμοποιεί

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

το δεκαδικό σύστημα (δεκαδικός υπολογιστής) παρ' όλο που αυτό δεν συμβαίνει στην πράξη. Η αντιστοιχία με άλλα συστήματα μπορεί να γίνει εύκολα.

Κάθε πραγματικός αριθμός x στο δεκαδικό σύστημα μπορεί να γραφεί υπό την μορφή

$$x = \alpha \cdot 10^{b}, \tag{1.14}$$

όπου α = (πρόσημο του x) $0.\alpha_1\alpha_2\alpha_3\ldots$, b ακέραιος, α_i ακέραιοι αριθμοί (ψηφία) με $0 \leq \alpha_i \leq 9$, $i = 1,2,\ldots$ και $\alpha_1 \neq 0$ (εκτός εάν $x = 0$), π.χ. $30.421 = 0.30421 \cdot 10^2$, $-0.0000603 = -0.603 \cdot 10^{-4}$.

Η μορφή (1.14) καλείται κανονικοποιημένη παράσταση κινητής δεκ. τελείας (normalized floating-point representation) του αριθμού x ή εκθετική μορφή ή επιστημονικός συμβολισμός και είναι ειδικά χρήσιμη για επιστημονικούς υπολογισμούς. Προφανώς ισχύει $10^{-1} \leq |\alpha| < 1$ αν $x \neq 0$, ενώ αν $x = 0$ τότε $\alpha = 0$. Ο αριθμός $0.\alpha_1\alpha_2\alpha_3\ldots$ καλείται κλασματικό μέρος (mantissa) ενώ ο b εκθέτης ή χαρακτηριστικό.

Σημαντικά ψηφία (σ.ψ.) του αριθμού x καλούνται όλα τα ψηφία της mantissa μετά την δεκαδική τελεία. Επειδή στην πράξη η παράσταση των αριθμών πρέπει να είναι πεπερασμένη ας υποθέσουμε ότι ο υπολογιστής μας μπορεί να εκτελεί υπολογισμούς με αριθμούς που έχουν το πολύ κ σημαντικά ψηφία (κ θετικός ακέραιος) και ο εκθέτης b των οποίων βρίσκεται μέσα σε κάποια όρια, δηλαδή, $m \leq b \leq M$, όπου m είναι αρνητικός ακέραιος και M θετικός ακέραιος. Επομένως κάθε αριθμός που έχει περισσότερα από κ σημαντικά ψηφία (υποθέτουμε ότι για τον εκθέτη b ικανοποιείται η $m \leq b \leq M$) πρέπει να στρογγυλευθεί ώστε να έχει ακριβώς κ σημαντικά ψηφία. Για τον σκοπό αυτό αρκεί να στρογγυλεύσουμε την mantissa του αριθμού σε κ δεκαδικές θέσεις (δ.θ.). Μία συνήθης διαδικασία στρογγυλεύσεως είναι η εξής. Έστω ότι θέλουμε να στρογγυλεύσουμε τον αριθμό $x = \alpha \cdot 10^{b}$ με $|\alpha| = 0.\alpha_1\alpha_2\alpha_3\ldots\alpha_\kappa\alpha_{\kappa+1}\alpha_{\kappa+2}\ldots$ σε κ σημαντικά ψηφία. Αν καλέσουμε $\bar{x}$ τον στρογγυλευμένο αριθμό, τότε είναι

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\bar{x} = (\text{πρόσημο του } x)\tilde{\alpha}\cdot 10^{b} ,$$

όπου

$$\tilde{\alpha} = \begin{cases} 0.\alpha_1\alpha_2\ldots\alpha_K , & \text{αν } 0 \leqq \alpha_{K+1} \leqq 4 \\ 0.\alpha_1\alpha_2\ldots\alpha_K + 10^{-K} , & \text{αν } \alpha_{K+1} \geqq 5 . \end{cases}$$

Ένας άλλος τρόπος στρογγυλεύσεως τον οποίο και θα ακολουθούμε στη συνέχεια είναι ο εξής

$$\bar{x} = (\text{πρόσημο του } x)\tilde{\alpha}\cdot 10^{b} ,$$

όπου

$$\tilde{\alpha} = \begin{cases} 0.\alpha_1\alpha_2\ldots\alpha_K , & \text{αν } c < 0.5 \text{ ή αν } c = 0.5 \text{ και } \alpha_K \text{ άρτιος} \\ 0.\alpha_1\alpha_2\ldots\alpha_K + 10^{-K}, & \text{αν } c > 0.5 \text{ ή αν } c = 0.5 \text{ και } \alpha_K \text{ περιττός} \end{cases}$$

και $c = 0.\alpha_{K+1}\alpha_{K+2}\ldots$. Από τον τρόπο που γίνεται η στρογγύλευση του x σε K σημαντικά ψηφία και στις δύο περιπτώσεις για το απόλυτο σφάλμα στρογγυλεύσεως του $\bar{x}$ ισχύει

$$|\varepsilon| = |\bar{x} - x| \leqq \frac{1}{2}\, 10^{-K}\cdot 10^{b} = \frac{1}{2}\, 10^{b-K} \tag{1.15}$$

ενώ για το σχετικό σφάλμα στρογγυλεύσεως του $\bar{x}$ έχουμε $(x \neq 0)$

$$\delta = \frac{|\varepsilon|}{|x|} \leqq \frac{\frac{1}{2}\, 10^{-K}\cdot 10^{b}}{|\alpha|\cdot 10^{b}} = \frac{\frac{1}{2}\, 10^{-K}}{|\alpha|} \leqq \frac{\frac{1}{2}\, 10^{-K}}{10^{-1}} = \frac{1}{2}\, 10^{1-K} = 5\cdot 10^{-K} . \tag{1.16}$$

Αξίζει να σημειωθεί ότι στο πεπερασμένο σύνολο (σύστημα) $F(q,K,m,M)$ των πραγματικών αριθμών της μορφής (1.14), το οποίο καθορίζεται από την βάση q του συστήματος αριθμήσεως (εδώ έχει ληφθεί $q = 10$), τον πεπερασμένο αριθμό K (π.χ. $K = 10$) των σημαντικών ψηφίων και τα όρια m και M (π.χ. $m = -98$, $M = 100$) του εκθέτη b , και στο οποίο υποθέσαμε ότι εργαζόμαστε, λόγω της διαδικασίας της στρογγυλεύσεως δεν ισχύουν ο προσεταιριστικός και ο επιμεριστι-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

κός νόμος της Άλγεβρας. Μπορεί εύκολα να δειχθεί ότι στο $F(q,\kappa,m,M)$ υπάρχουν ακριβώς $1+2(q-1)q^{\kappa-1}(M-m+1)$ διαφορετικοί αριθμοί. Για $q = 10$, $\kappa = 10$, $m = -98$ και $M = 100$, ο αριθμός αυτός ισούται με $1+3582\cdot 10^9 \approx 3.6\cdot 10^{12}$.

Αν θεωρήσουμε τον πραγματικό αριθμό x στην συνηθισμένη του μορφή, που αποτελείται γενικά από το ακέραιο τμήμα, την δεκαδική τελεία και το δεκαδικό (κλασματικό) τμήμα, τότε τα σημαντικά του ψηφία είναι όλα τα ψηφία αρχίζοντας από αριστερά προς τα δεξιά με το πρώτο μη μηδενικό ψηφίο. Για παράδειγμα, στους αριθμούς $0.0\underline{3045}$ και $0.0\underline{304500}$ τα σημαντικά ψηφία είναι υπογραμμισμένα. Όταν ο αριθμός είναι ακέραιος και έχει ορισμένα μηδενικά στο τέλος, τότε πρέπει να δηλώνονται πόσα από τα μηδενικά αυτά είναι σημαντικά ψηφία ή να γράφεται ο αριθμός υπό την εκθετική μορφή για να μην υπάρχει αμφιβολία. Είναι φανερό ότι η στρογγύλευση ενός αριθμού σε ένα δοθέντα αριθμό σημαντικών ψηφίων δεν δίνει γενικά το ίδιο αποτέλεσμα με την στρογγύλευση αυτού στον ίδιο αριθμό δεκαδικών θέσεων. Όταν ο αριθμός x στρογγυλεύεται σε κ δεκαδικές θέσεις τότε για το απόλυτο σφάλμα (στρογγυλεύσεως) του στρογγυλευμένου αριθμού $\bar{x}$ ισχύει προφανώς

$$|\varepsilon| = |\bar{x} - x| \leq \frac{1}{2}10^{-\kappa} \quad . \tag{1.17}$$

Ένα σημαντικό ψηφίο ενός προσεγγιστικού αριθμού $\bar{x}$ λέγεται ότι είναι ακριβές εάν το απόλυτο σφάλμα του $\bar{x}$ δεν υπερβαίνει την μισή μονάδα της τάξεως που αντιστοιχεί σ' αυτό το ψηφίο. Άν κ σημαντικά ψηφία του $\bar{x}$ είναι ακριβή τότε λέμε ότι ο $\bar{x}$ είναι ακριβής σε κ σημαντικά ψηφία ή ακόμη ακριβής στον αριθμό των δεκαδικών θέσεων που προσδιορίζεται από το τελευταίο ακριβές σημαντικό ψηφίο. Εάν όλα τα σημαντικά ψηφία του $\bar{x}$ είναι ακριβή, τότε λέμε ότι ο $\bar{x}$ είναι ακριβής στον δοθέντα αριθμό ψηφίων. Για παράδειγμα, αν ο $\bar{x} = 0.03045$ έχει απόλυτο σφάλμα $|\varepsilon| = 0.0004 = 0.4\cdot 10^{-3}$, τότε, αφού $|\varepsilon| < 0.5\cdot 10^{-3}$, ο $\bar{x}$ είναι ακριβής σε δύο σημαντικά ψηφία ή ακριβής σε τρείς δεκαδικές θέσεις. Αν για τον $\bar{x} = 36.00$ εί-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ναι $|\varepsilon| = 0.03$ τότε αυτός είναι ακριβής σε 3 σημαντικά ψηφία ή σε μία δεκαδική θέση, αφού $0.03 < 0.05 = 0.5 \cdot 10^{-1}$.

Τονίζεται εδώ ότι όταν ένας ακριβής αριθμός x στρογγυλεύεται σε κ σημαντικά ψηφία ή σε κ δεκαδικές θέσεις, τότε ο στρογγυλευμένος αριθμός $\bar{x}$ είναι ακριβής σε κ σημαντικά ψηφία ή σε κ δεκαδικές θέσεις αντίστοιχα. Ακόμη, είναι φανερό ότι αν ο προσεγγιστικός αριθμός $\bar{x}$ είναι ακριβής σε κ δεκαδικές θέσεις, τότε για το απόλυτο σφάλμα του ισχύει $|\varepsilon| \leqq \frac{1}{2} 10^{-\kappa}$, ενώ αν είναι ακριβής σε κ σημαντικά ψηφία τότε για το σχετικό σφάλμα του ισχύει $\delta \leqq 5 \cdot 10^{-\kappa}$.

Λεμε ότι δύο αριθμοί συμφωνούν σε κ σημαντικά ψηφία εάν μετά την στρογγύλευσή τους σε κ σημαντικά ψηφία οι αριθμοί ταυτίζονται.

Λέμε ότι δύο αριθμοί συμφωνούν σε κ δεκαδικές θέσεις εάν η απόλυτη τιμή της διαφοράς των είναι $\leqq \frac{1}{2} 10^{-\kappa}$. Π.χ. οι αριθμοί 0.125 και 0.130 συμφωνούν σε 1 σημαντικό ψηφίο ενώ συμφωνούν σε 2 δεκαδικές θέσεις. Οι αριθμοί 9.995 και 10.000 συμφωνούν σε δύο δεκαδικές θέσεις ενώ συμφωνούν σε τρία σημαντικά ψηφία.

Η ακρίβεια ενός προσεγγιστικού αριθμού εξαρτάται από τον αριθμό των ακριβών σημαντικών ψηφίων του και όχι από τον συνολικό αριθμό των σημαντικών ψηφίων του. Για τον λόγο αυτό ο ακόλουθος πρακτικός κανόνας μπορεί να χρησιμοποιηθεί σαν οδηγός: Όταν εκτελούμε υπολογισμούς με προσεγγιστικούς αριθμούς, ο αριθμός των σημαντικών ψηφίων στα ενδιάμεσα αποτελέσματα δεν πρέπει να υπερβαίνει τον αριθμό των ακριβών σημαντικών ψηφίων περισσότερο από 1 ή 2 μονάδες (η στρογγύλευση εφαρμόζεται για τον σκοπό αυτό). Το τελικό αποτέλεσμα στρογγυλεύεται επίσης σε τόσα σημαντικά ψηφία όσος είναι ο αριθμός των ακριβών σημαντικών ψηφίων του.

Ακολουθώντας τον παραπάνω κανόνα μειώνουμε τον χρόνο υπολογισμού χωρίς να μειώνεται η ακρίβεια του τελικού αποτελέσματος.

1.4 Διάδοση και δημιουργία σφαλμάτων κατά τους υπολογισμούς

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Ας υποθέσουμε ότι $f(\bar{x}_1,\bar{x}_2,\ldots,\bar{x}_n)$ είναι μία πραγματική συνάρτηση των πραγματικών ποσοτήτων $\bar{x}_1,\bar{x}_2,\ldots,\bar{x}_n$ που αποτελούν τα δεδομένα του προβλήματος. Τα δεδομένα στους πρακτικούς υπολογισμούς είναι συνήθως μόνον προσεγγιστικές τιμές αφού προέρχονται από μετρήσεις, εκτιμήσεις ή προηγούμενους υπολογισμούς. Έτσι οι τιμές $\bar{x}_1,\bar{x}_2,\ldots,\bar{x}_n$ είναι προσεγγίσεις (π.χ. μπορεί να προέρχονται από στρογγύλευση) κάποιων ακριβών τιμών $x_1,x_2,\ldots,x_n$ και έστωσαν $|\varepsilon_i|$, $i = 1(1)n$ τα απόλυτα σφάλματα των $\bar{x}_i$, $i = 1(1)n$ (συνήθως είναι γνωστά μόνον άνω φράγματα για τα $|\varepsilon_i|$, $i = 1(1)n$) . Σκοπός μας είναι να προσδιορίσουμε (εκτιμήσουμε) το απόλυτο και σχετικό σφάλμα της προσεγγιστικής τιμής $f(\bar{x}_1,\bar{x}_2,\ldots,\bar{x}_n)$. Υποθέτουμε εδώ ότι οι απαιτούμενες πράξεις (υπολογισμοί) για την εύρεση της $f(\bar{x}_1,\bar{x}_2,\ldots,\bar{x}_n)$ εκτελούνται ακριβώς, δηλαδή, δεν δημιουργούνται νέα σφάλματα (στρογγυλεύσεως) και επομένως το σφάλμα στο αποτέλεσμα, δηλαδή, στην τιμή $f(\bar{x}_1,\bar{x}_2,\ldots,\bar{x}_n)$ προέρχεται αποκλειστικά από την διάδοση των σφαλμάτων των αρχικών προσεγγιστικών δεδομένων $\bar{x}_i$, $i = 1(1)n$ δια μέσου των υπολογισμών (διαδοθέν σφάλμα).

Με την προϋπόθεση ότι η συνάρτηση $f(x_1,x_2,\ldots,x_n)$ αναπτύσσεται σε σειρά Taylor γύρω από το σημείο $\bar{x} = (\bar{x}_1,\bar{x}_2,\ldots,\bar{x}_n)$ έχουμε

$$f(x_1,x_2,\ldots,x_n) =$$

$$= \sum_{\kappa=0}^{\infty} \frac{1}{\kappa!}\left(-\varepsilon_1\frac{\partial}{\partial x_1} - \varepsilon_2\frac{\partial}{\partial x_2} - \varepsilon_3\frac{\partial}{\partial x_3} - \ldots - \varepsilon_n\frac{\partial}{\partial x_n}\right)^{\kappa} f \ , \qquad (1.18)$$

όπου $\varepsilon_i = \bar{x}_i - x_i$, $i = 1(1)n$ και μετά την παραγώγιση οι μερικές παράγωγοι της f υπολογίζονται στο σημείο $\bar{x}$. Όταν τα ε_i , $i = 1(1)n$ είναι μικρά (όπως συμβαίνει συνήθως) χωρίς όλες οι μερικές παράγωγοι $\frac{\partial f(\bar{x})}{\partial x_i}$ να είναι επίσης μικρές ποσότητες, από την (1.18) παίρνουμε παραλείποντας τους υπόλοιπους όρους

$$f(x_1,x_2,\ldots,x_n) \approx f(\bar{x}_1,\bar{x}_2,\ldots,\bar{x}_n) - \Big(\varepsilon_1 f_{x_1}(\bar{x}_1,\bar{x}_2,\ldots,\bar{x}_n) +$$

$$+ \varepsilon_2 f_{x_2}(\bar{x}_1,\bar{x}_2,\ldots,\bar{x}_n) + \ldots + \varepsilon_n f_{x_n}(\bar{x}_1,\bar{x}_2,\ldots,\bar{x}_n)\Big) \ ,$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

όπου $f_{x_i} \equiv \frac{\partial f}{\partial x_i}$. Αν καλέσουμε $|\varepsilon|$ το απόλυτο (διαδοθέν) σφάλμα της $f(\bar{x}_1,\bar{x}_2,\ldots,\bar{x}_n)$ τότε έχουμε

$$|\varepsilon| = |f(\bar{x}_1,\bar{x}_2,\ldots,\bar{x}_n) - f(x_1,x_2,\ldots,x_n)| \approx$$

$$\approx |\varepsilon_1 f_{x_1}(\bar{x}) + \varepsilon_2 f_{x_2}(\bar{x}) + \ldots + \varepsilon_n f_{x_n}(\bar{x})| =$$

$$= \left|\sum_{i=1}^{n} \varepsilon_i f_{x_i}(\bar{x})\right| \leqq \sum_{i=1}^{n} |\varepsilon_i|\,|f_{x_i}(\bar{x})| \ , \qquad (1.19)$$

δηλαδή,

$$|\varepsilon| \lessapprox \sum_{i=1}^{n} |\varepsilon_i| \left|f_{x_i}(\bar{x})\right| \ . \qquad (1.20)$$

Αν $|\varepsilon_i| \leqq M_i$, $i = 1(1)n$ είναι τα φράγματα των απολύτων σφαλμάτων των $\bar{x}_i$, $i = 1(1)n$, τότε από την (1.20) προκύπτει

$$|\varepsilon| \lessapprox \sum_{i=1}^{n} M_i \left|f_{x_i}(\bar{x})\right| \ . \qquad (1.21)$$

Από την (1.20) διαιρώντας τα δύο μέλη με $|f(\bar{x}_1,\bar{x}_2,\ldots,\bar{x}_n)|$ παίρνουμε μία εκτίμηση για το σχετικό σφάλμα δ της $f(\bar{x})$

$$\delta \approx \frac{|\varepsilon|}{|f(\bar{x})|} \lessapprox \sum_{i=1}^{n} \left|\frac{f_{x_i}(\bar{x})}{f(\bar{x})}\right| |\varepsilon_i| = \sum_{i=1}^{n} \left|\frac{\partial}{\partial x_i} \ln f(\bar{x})\right| |\varepsilon_i| \leqq$$

$$\leqq \sum_{i=1}^{n} M_i \left|\frac{\partial}{\partial x_i} \ln f(\bar{x})\right| \ . \qquad (1.22)$$

<u>Παράδειγμα 1.3</u> Δίνεται η συνάρτηση $f(t) = 6t^2 + 8t - 30$. Έστωσαν $x_1 = 3.01$ και $\bar{x}_1 = 3$ η ακριβής και η προσεγγιστική τιμή αντίστοιχα ενός αριθμού. Να βρεθεί το απόλυτο σφάλμα της τιμής $f(\bar{x}_1)$ α-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

κριβώς και ύστερα με την βοήθεια του τύπου (1.20) (κατά προσέγγιση).

Λύση Σύμφωνα με τον ορισμό είναι $|\varepsilon| = |f(\bar{x}_1) - f(x_1)| = |48 - 48.4406| = 0.4406$. Βάσει των σχέσεων (1.19) και (1.20) έχουμε

$$|\varepsilon| \approx |\varepsilon_1||f'(\bar{x}_1)| = |\bar{x}_1 - x_1||f'(\bar{x}_1)| =$$
$$= |3 - 3.01||f'(3)| = (0.01)44 = 0.44 \quad . \qquad \blacktriangle$$

Τα θεωρήματα που ακολουθούν βασίζονται στη θεωρία που αναπτύχθηκε προηγουμένως.

Θεώρημα 1.3 Το απόλυτο σφάλμα του αλγεβρικού αθροίσματος n προσεγγιστικών αριθμών $\bar{x}_i$, $i = 1(1)n$ δεν υπερβαίνει το άθροισμα των απολύτων σφαλμάτων των αριθμών αυτών.

Απόδειξη Το αλγεβρικό άθροισμα των $\bar{x}_i$, $i = 1(1)n$ είναι η τιμή

$$f(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n) = c_1\bar{x}_1 + c_2\bar{x}_2 + \ldots + c_n\bar{x}_n \quad , \qquad (1.23)$$

όπου $c_i = 1$ ή -1 , $i = 1(1)n$. Αν $\varepsilon_i = \bar{x}_i - x_i$, $i = 1(1)n$ τότε είναι

$$\varepsilon = f(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n) - f(x_1, x_2, \ldots, x_n) = c_1\varepsilon_1 + c_2\varepsilon_2 + \ldots + c_n\varepsilon_n$$

και επομένως, αφού $|c_i| = 1$, $i = 1(1)n$,

$$|\varepsilon| \leqq |\varepsilon_1| + |\varepsilon_2| + \ldots + |\varepsilon_n| \quad . \qquad \blacktriangle \qquad (1.24)$$

Παρατήρηση Αν οι $\bar{x}_i$ προέρχονται από στρογγύλευση των x_i σε κ δεκαδικές θέσεις, τότε προφανώς $|\varepsilon_i| \leqq \frac{1}{2}10^{-\kappa}$, $i = 1(1)n$ και επομένως για το απόλυτο σφάλμα στρογγυλεύσεως του αλγεβρικού αθροίσματος των $\bar{x}_i$, λόγω της (1.24) ισχύει

$$|\varepsilon| \leqq n\,\frac{1}{2}10^{-\kappa} \qquad (1.25)$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

(Σημειώνεται ότι η στατιστική θεωρία, για μεγάλο n δίνει μια πιο ρεαλιστική εκτίμηση και συγκεκριμένα $|\varepsilon| \leqq \frac{1}{2}\sqrt{n}\,10^{-K}$) .

<u>Πόρισμα</u> Το απόλυτο σφάλμα της διαφοράς $\bar{x}_1 - \bar{x}_2$ δεν υπερβαίνει το άθροισμα των απολύτων σφαλμάτων των $\bar{x}_1$ και $\bar{x}_2$.

<u>Σημείωση</u> Στην πράξη πρέπει να αποφεύγεται εάν είναι δυνατόν η αφαίρεση δύο σχεδόν ίσων προσεγγιστικών αριθμών διότι αυτή οδηγεί σε απώλεια σημαντικών ψηφίων (μείωση της ακρίβειας του αποτελέσματος). Για παράδειγμα, αν οι $\bar{x}_1 = 15.243$ και $\bar{x}_2 = 15.222$ είναι ακριβείς στον δοθέντα αριθμό ψηφίων (5 σ.ψ.), τότε για τα απόλυτα σφάλματα $|\varepsilon_i|$, $i = 1,2$ αυτών ισχύει $|\varepsilon_i| \leqq \frac{1}{2}10^{-3}$, $i = 1,2$ και επομένως για το απόλυτο σφάλμα $|\varepsilon|$ της διαφοράς $\bar{x}_1 - \bar{x}_2 = 0.021$ έχουμε $|\varepsilon| \leqq 2\,\frac{1}{2}\,10^{-3} = 10^{-3} = 0.001 < 0.5\cdot 10^{-2}$. Δηλαδή, η τιμή 0.021 είναι ακριβής σε 1 σ.ψ. . Η αύξηση της ακρίβειας μπορεί μερικές φορές να επιτεύχθεί με απλούς αλγεβρικούς μετασχηματισμούς. Έτσι, για τον υπολογισμό π.χ. της διαφοράς $70-\sqrt{4899}$, όπου οι αριθμοί 70 και 4899 είναι ακριβείς και δίνεται ότι $\sqrt{4899} \approx 69.99$ (με ακρίβεια 4 σ.ψ.) έχουμε τους ακόλουθους δύο τρόπους:

i) $70 - \sqrt{4899} \approx 70 - 69.99 = 0.01 = \bar{x}_1$ και

ii) $$70 - \sqrt{4899} = \frac{1}{70+\sqrt{4899}} \approx \frac{1}{70+69.99} = \frac{1}{139.99} \approx$$

$$\approx 0.00714336 = \bar{x}_2 .$$

Μπορεί να δειχθεί (βλέπε άσκηση 12) ότι ο $\bar{x}_1$ είναι ακριβής σε 1 σ.ψ. ενώ ο $\bar{x}_2$ είναι ακριβής σε 4 σ.ψ. .

Στο σημείο αυτό πρέπει να αναφερθεί ότι, όπως συμβαίνει με τα προβλήματα έτσι και στις αριθμητικές μεθόδους (υπολογιστικές διαδικασίες – αλγόριθμοι) που χρησιμοποιούνται για την εύρεση των αριθμητικών λύσεων των προβλημάτων, υπάρχει αντίστοιχος διαχωρισμός σε ευσταθείς και ασταθείς μεθόδους. Μία ασταθής μέθοδος είναι δυνατόν μερικές φορές να δώσει αποτελέσματα με μειωμένη ακρίβεια (δεν υπάρχει

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

σταθερότης στην ακρίβεια) ενώ άλλες φορές δεν δημιουργεί δυσκολίες. Έτσι στο παραπάνω πρόβλημα η μέθοδος (i) χαρακτηρίζεται ασταθής ένώ η (ii) ευσταθής.

Θεώρημα 1.4 Το σχετικό σφάλμα του γινομένου n προσεγγιστικών αριθμών $\bar{x}_i$, $i = 1(1)n$ είναι μικρότερο από,ή περίπου ίσον με το άθροισμα των σχετικών σφαλμάτων των αριθμών αυτών.

Απόδειξη Θεωρούμε την συνάρτηση $f(\bar{x}_1,\bar{x}_2,\ldots\bar{x}_n) = \bar{x}_1\bar{x}_2\bar{x}_3\ldots\bar{x}_n$. Σύμφωνα με την (1.20) έχουμε για το απόλυτο σφάλμα $|\varepsilon|$ αυτής

$$|\varepsilon| \lessapprox |\varepsilon_1||\bar{x}_2\bar{x}_3\ldots\bar{x}_n| + |\varepsilon_2||\bar{x}_1\bar{x}_3\ldots\bar{x}_n| + \ldots +$$

$$+ |\varepsilon_n||\bar{x}_1\bar{x}_2\ldots\bar{x}_{n-1}| , \qquad (1.26)$$

όπου $\varepsilon_i = \bar{x}_i - x_i$, $i = 1(1)n$. Για το σχετικό σφάλμα δ προκύπτει ότι

$$\delta = \frac{|\varepsilon|}{|x_1x_2\ldots x_n|} \approx \frac{|\varepsilon|}{|\bar{x}_1\bar{x}_2\ldots\bar{x}_n|} \lessapprox \frac{|\varepsilon_1|}{|\bar{x}_1|} + \frac{|\varepsilon_2|}{|\bar{x}_2|} + \ldots + \frac{|\varepsilon_n|}{|\bar{x}_n|} \approx$$

$\approx \delta_1 + \delta_2 + \ldots + \delta_n$, λόγω της (1.26) και των $\delta_i = \dfrac{|\varepsilon_i|}{|x_i|} \approx \dfrac{|\varepsilon_i|}{|\bar{x}_i|}$,

$i = 1(1)n$. Άρα τελικά $\delta \lessapprox \delta_1 + \delta_2 + \ldots + \delta_n$ και επί πλέον

$$|\varepsilon| \lessapprox |\bar{x}_1\bar{x}_2\ldots\bar{x}_n| \, (\delta_1 + \delta_2 + \ldots + \delta_n) . \qquad \blacktriangle \qquad (1.27)$$

Θεώρημα 1.5 Το σχετικό σφάλμα του πηλίκου δύο προσεγγιστικών αριθμών είναι μικρότερο από,ή περίπου ίσον με το άθροισμα των σχετικών σφαλμάτων των αριθμών αυτών.

Απόδειξη Θεωρούμε την συνάρτηση $f(\bar{x}_1,\bar{x}_2) = \dfrac{\bar{x}_1}{\bar{x}_2}$, $\bar{x}_2 \neq 0$. Σύμφωνα με την (1.20) για το απόλυτο σφάλμα $|\varepsilon|$ αυτής έχουμε

$$|\varepsilon| \lessapprox |\varepsilon_1|\left|\frac{1}{\bar{x}_2}\right| + |\varepsilon_2|\left|-\frac{\bar{x}_1}{\bar{x}_2^2}\right| = \frac{|\varepsilon_1|}{|\bar{x}_2|} + \frac{|\varepsilon_2|}{|\bar{x}_2|}\,\frac{|\bar{x}_1|}{|\bar{x}_2|} , \qquad (1.28)$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

όπου $\varepsilon_i = \bar{x}_i - x_i$, $i = 1,2$. Λόγω της (1.28) για το σχετικό σφάλμα δ (υποτίθεται $x_1, x_2, \bar{x}_1 \neq 0$) προκύπτει

$$\delta = \frac{|\varepsilon|}{\left|\frac{x_1}{x_2}\right|} \approx \frac{|\varepsilon|}{\left|\frac{\bar{x}_1}{\bar{x}_2}\right|} \lessapprox \frac{\frac{|\varepsilon_1|}{|\bar{x}_2|} + \frac{|\varepsilon_2|}{|\bar{x}_2|} \frac{|\bar{x}_1|}{|\bar{x}_2|}}{\left|\frac{\bar{x}_1}{\bar{x}_2}\right|} = \frac{|\varepsilon_1|}{|\bar{x}_1|} + \frac{|\varepsilon_2|}{|\bar{x}_2|} \approx \delta_1 + \delta_2 \quad \text{ή}$$

$\delta \lessapprox \delta_1 + \delta_2$. Επίσης ισχύει

$$|\varepsilon| \lessapprox \left|\frac{\bar{x}_1}{\bar{x}_2}\right| (\delta_1 + \delta_2) \quad . \qquad \blacktriangle \qquad (1.29)$$

<u>Παράδειγμα 1.4</u> Αν οι αριθμοί 23.86 , 0.01762 , 37.26 , 0.02146 προέρχονται από στρογγύλευση σε 4 σ.ψ. αντίστοιχων ακριβών αριθμών, να βρεθούν τα αποτελέσματα των πράξεων i) 23.86 + 0.01762 και ii) (37.26) (0.02146) στρογγυλευμένα στο πλήθος των ακριβών σημαντικών ψηφίων.

<u>Λύση</u> i) Είναι 23.86 + 0.01762 = 23.87762. Το απόλυτο σφάλμα του αποτελέσματος είναι $|\varepsilon| \leqq |\varepsilon_1| + |\varepsilon_2|$, όπου $|\varepsilon_i|$, $i = 1,2$ είναι τα απόλυτα σφάλματα των προσθετέων. Επειδή $|\varepsilon_1| \leqq \frac{1}{2} 10^{-2}$ και $|\varepsilon_2| \leqq \frac{1}{2} 10^{-5}$ έχουμε

$$|\varepsilon| \leqq 0.5 \cdot 10^{-2} + 0.5 \cdot 10^{-5} = 0.005005 < 0.5 \cdot 10^{-1} = 0.05 \quad .$$

Άρα το αποτέλεσμα 23.87762 είναι ακριβές σε 3 σ.ψ. και με την στρογγύλευση γίνεται 23.9 . ii) (37.26) (0.02146) = 0.7995996 . Βάσει της (1.26) έχουμε για το απόλυτο σφάλμα $|\varepsilon|$ του γινομένου $|\varepsilon| \lessapprox |\varepsilon_1||\bar{x}_2| + |\varepsilon_2||\bar{x}_1| \leqq \frac{1}{2} 10^{-2}(0.02146) + \frac{1}{2} 10^{-5}(37.26) =$ $= 0.0002936 < 0.5 \cdot 10^{-3}$. Άρα το αποτέλεσμα 0.7995996 είναι ακριβές σε 3 σ.ψ. και στρογγυλεύεται σε 0.800 . ▲

<u>Παράδειγμα 1.5</u> Να βρεθούν άνω φράγματα για το απόλυτο και σχε-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

τικό σφάλμα του όγκου $V = \frac{1}{6}\pi d^3$ μιας σφαίρας εάν η διάμετρος d στρογγυλευμένη σε 2 σ.ψ. είναι 3.7 cm και $\pi \approx 3.14$.

<u>Λύση</u> Στην περίπτωσή μας η συνάρτηση είναι $f(\bar{x}_1,\bar{x}_2) = \frac{1}{6}\bar{x}_1\bar{x}_2^3$ όπου $\bar{x}_1 = 3.14$ και $\bar{x}_2 = 3.7$. Αν $|\varepsilon_1| = |\bar{x}_1 - x_1|$, $|\varepsilon_2| = |\bar{x}_2 - x_2|$, τότε είναι $|\varepsilon_1| \approx 0.0016$ και $|\varepsilon_2| \leq \frac{1}{2}10^{-1} = 0.05$ και σύμφωνα με τις (1.20) και (1.21) έχουμε για το απόλυτο σφάλμα $|\varepsilon|$ του όγκου

$$|\varepsilon| \lesssim |\varepsilon_1|\left|\frac{1}{6}\bar{x}_2^3\right| + |\varepsilon_2|\left|\frac{1}{2}\bar{x}_1\bar{x}_2^2\right| \lesssim 0.0016\left|\frac{1}{6}(3.7)^3\right| +$$

$$+ 0.05\left|\frac{1}{2}(3.14)(3.7)^2\right| \approx 0.0016(8.44) + 0.05(21.5) \approx$$

$$\approx 0.014 + 1.075 = 1.089 \approx 1.1 \text{ cm}^3 \quad .$$

Εξ άλλου για το σχετικό σφάλμα ισχύει

$$\delta \approx \frac{|\varepsilon|}{\left|\frac{1}{6}\bar{x}_1\bar{x}_2^3\right|} \lesssim \frac{1.089}{\frac{1}{6}(3.14)(3.7)^3} \approx \frac{1.089}{26.5} \approx 0.04 = 4\% \quad . \qquad \blacktriangle$$

Κατά την ανάπτυξη της θεωρίας της παρούσας παραγράφου αγνοήσαμε τελείως τα παραγόμενα (δημιουργούμενα) σφάλματα στρογγυλεύσεως κατά την διάρκεια των υπολογισμών μέσω ενός υπολογιστού. Ο υπολογιστής για την εύρεση της τιμής $f(\bar{x}_1,\bar{x}_2,\ldots,\bar{x}_n)$ εκτελεί ένα σύνολο αριθμητικών πράξεων, δηλαδή, προσθέσεων, αφαιρέσεων, πολλαπλασιασμών και διαιρέσεων. Επειδή οι πράξεις στον υπολογιστή δεν εκτελούνται ακριβώς αλλά αντικαθίστανται κατά κάποιο τρόπο από άλλες προσεγγιστικές (εξ αιτίας του περιορισμένου αριθμού σημαντικών ψηφίων που χρησιμοποιεί), κάθε πράξη δημιουργεί γενικά ένα σφάλμα στρογγυλεύσεως και επομένως η τελική τιμή που υπολογίζεται για την $f(\bar{x}_1,\bar{x}_2,\ldots,\bar{x}_n)$ θα περιέχει ένα σφάλμα που καλείται υπολογιστικό σφάλμα και που οφείλεται αφ' ενός στο διαδοθέν σφάλμα και αφ' ετέρου στα δημιουργούμενα σφάλματα των επί μέρους προσεγγιστικών πράξεων, δηλαδή, στο καλούμενο παραχθέν σφάλμα. Το μέγεθος του παραχθέντος σφάλματος εξαρτάται από την σειρά με

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

την οποία εκτελούνται οι πράξεις και είναι δύσκολο να εκτιμηθεί αυτό στο τελικό αποτέλεσμα ενός πολύπλοκου υπολογισμού. Σαν παράδειγμα θεωρούμε την περίπτωση όπου μία μόνον αριθμητική πράξη εκτελείται μεταξύ δύο προσεγγιστικών αριθμών. Ας υποθέσουμε χάριν απλότητος ότι $f(\bar{x}_1,\bar{x}_2) = \bar{x}_1 + \bar{x}_2$, $\bar{x}_1 = 0.962$, $\bar{x}_2 = 0.6732$, $|\varepsilon_1|,|\varepsilon_2| \leqq \frac{1}{2}10^{-4}$ και ότι ο υπολογιστής μας χρησιμοποιεί μόνον 4 σ.ψ.Αν συμβολίσουμε με E_υ το υπολογιστικό σφάλμα στην τιμή $\bar{x}_1 + \bar{x}_2$ που τελικά υπολογίζεται, με ε το διαδοθέν σφάλμα και με ε_π το παραχθέν σφάλμα (στρογγυλεύσεως) λόγω της προσεγγιστικής προσθέσεως που συμβολίζουμε με $\oplus$, τότε έχουμε

$$E_\upsilon = (\bar{x}_1 \oplus \bar{x}_2) - (x_1 + x_2) = \left[(\bar{x}_1 \oplus \bar{x}_2) - (\bar{x}_1 + \bar{x}_2)\right] +$$

$$+ \left[(\bar{x}_1 + \bar{x}_2) - (x_1 + x_2)\right] = \varepsilon_\pi + \varepsilon ,$$

αφού

$$\varepsilon_\pi = (\bar{x}_1 \oplus \bar{x}_2) - (\bar{x}_1 + \bar{x}_2) \quad \text{και} \quad \varepsilon = (\bar{x}_1 + \bar{x}_2) - (x_1 + x_2) .$$

Άρα,

$$|E_\upsilon| = |\varepsilon_\pi + \varepsilon| \leqq |\varepsilon_\pi| + |\varepsilon| . \tag{1.30}$$

Επειδή $\bar{x}_1 + \bar{x}_2 = 1.6353$, $\bar{x}_1 \oplus \bar{x}_2 = 1.635$, $\varepsilon_\pi = -0.0003$ και $|\varepsilon| \leqq |\varepsilon_1| + |\varepsilon_2| \leqq 2\frac{1}{2}10^{-4} = 10^{-4}$ (σύμφωνα με την (1.24)), από την (1.30) προκύπτει ότι

$$|E_\upsilon| \leqq 0.0003 + 0.0001 = 0.0004 .$$

1.5 Ολικό σφάλμα στην αριθμητική λύση ενός μαθηματικού προβλήματος

Χάριν απλότητος ας υποθέσουμε ότι το μαθηματικό πρόβλημα συνίσταται στην εύρεση της τιμής της συναρτήσεως $f(x)$ για κάποιο x . Συνήθως η f προσεγγίζεται (αντικαθίσταται) από κάποια απλούστερη συνάρ-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

τηση $f_1(x)$ και η διαφορά $E_A = f_1(x) - f(x)$ είναι τότε το καλούμενο σφάλμα αποκοπής. Στη θέση του x χρησιμοποιείται συνήθως μία προσεγγιστική τιμή $\bar{x}$ και η διαφορά $\varepsilon = f_1(\bar{x}) - f_1(x)$ είναι το καλούμενο διαδοθέν σφάλμα. Τελικά, λόγω των προσεγγιστικών πράξεων δεν βρίσκουμε την τιμή $f_1(\bar{x})$ αλλά την $\bar{f}_1(\bar{x})$ και η διαφορά $\varepsilon_\pi = \bar{f}_1(\bar{x}) - f_1(\bar{x})$ είναι το καλούμενο παραχθέν σφάλμα. Έτσι, για το ολικό σφάλμα $E_{o\lambda}$ στην τιμή $\bar{f}_1(\bar{x})$ έχουμε

$$E_{o\lambda} = \bar{f}_1(\bar{x}) - f(x) = \left[\bar{f}_1(\bar{x}) - f_1(\bar{x})\right] + \left[f_1(\bar{x}) - f_1(x)\right] +$$

$$+ \left[f_1(x) - f(x)\right] = \varepsilon_\pi + \varepsilon + E_A \quad . \tag{1.31}$$

Ας σημειωθεί ότι το υπολογιστικό σφάλμα είναι $E_\upsilon = \bar{f}_1(\bar{x}) - f_1(x) = \varepsilon_\pi + \varepsilon$, οπότε

$$E_{o\lambda} = E_\upsilon + E_A \quad . \tag{1.32}$$

ΑΣΚΗΣΕΙΣ

1. Σύμφωνα με την (1.6) είναι $\delta \approx \bar{\delta} = \dfrac{|\varepsilon|}{|\bar{x}|}$, $\bar{x} \neq 0$. Να δειχθεί ότι $|\bar{\delta} - \delta| \lesssim (\bar{\delta})^2$.

2. Να βρεθεί μία απλή προσέγγιση για την παράγωγο $f'(x_0)$ με την βοήθεια της σχέσης (1.10) και να γραφεί η εκφραση του σφάλματος αποκοπής.

3. Να στρογγυλευθούν οι αριθμοί: -617.950 , 6890150 , -15.0032 , $\frac{5}{3}$, $\frac{1}{17}$, $-\frac{1}{19}$ σε 4 σ.ψ. .

4. Να στρογγυλευθούν οι αριθμοί: -75684.3195 , 0.02345 , -0.04795 , 0.086251 , -0.6485003 , 0.0005985 , 689000.0095 i) σε 3 σ.ψ. και ii) σε 3 δ.θ. .

5. Με την προυπόθεση ότι οι παρακάτω αριθμοί είναι ακριβείς και ότι

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

μπορούμε να εκτελούμε πράξεις με αριθμούς που έχουν το πολύ 4 σ. ψ., να βρεθεί το αποτέλεσμα του υπολογισμού $(0.3789)(0.2614) - 0.0990$ και το σφάλμα της υπολογισμένης τιμής.

6. Να βρεθεί η λύση του συστήματος $\begin{cases} 0.0995x + 0.9949y = 1.0944 \\ 0.0985x + 0.9951y = 1.0936 \end{cases}$ και στη συνέχεια η λύση του προκύπτοντος συστήματος μετά την στρογγύλευση όλων των αριθμών σε 3 δ.θ. . Τι παρατηρείτε;

7. Αν οι αριθμοί 3.45 , 4.87 , 5.16 , 3.55 , 2.73 προέρχονται από στρογγύλευση σε 3 σ.ψ. αντίστοιχων ακριβών αριθμών, να βρεθούν τα αποτελέσματα των πράξεων: i) $3.45 + 4.87 - 5.16$ ii) $(3.55)(2.73)$ στρογγυλευμένα στο πλήθος των ακριβών σημαντικών ψηφίων.

8. Οι ακριβείς αριθμοί $x_i > 0$, $i = 1(1)n$ προσεγγιζονται από τους $\bar{x}_i > 0$, $i = 1(1)n$. Να δειχθεί ότι το σχετικό σφάλμα του αθροίσματος των $\bar{x}_i$ δεν υπερβαίνει το μέγιστο των σχετικών σφαλμάτων των $\bar{x}_i$.

9. Με ακρίβεια πόσων τουλάχιστον δ.θ. πρέπει να βρεθεί η $\sqrt{11.2}$ έτσι ώστε το απόλυτο σφάλμα της παράστασης $2340(3 + \sqrt{11.2})$ να είναι ≤ 0.2 ; (Οι αριθμοί 2340, 11.2 και 3 είναι ακριβείς).

10. Ο ακριβής αριθμός $x = \sqrt{4.5}$ προσεγγίζεται από τον $\bar{x} = \sqrt{4} = 2$. Να βρεθεί κατά προσέγγιση το απόλυτο σφάλμα του $\bar{x}$ και στη συνέχεια να εκτιμηθεί η $\sqrt{4.5}$.

11. Οι ακριβείς αριθμοί x_1, x_2, x_3 όταν στρογγυλευθούν σε 3 δ.θ. είναι $\bar{x}_1 = 4.70$, $\bar{x}_2 = 0.832$, $\bar{x}_3 = 2.413$ αντίστοιχα. Να βρεθούν τα αποτελέσματα i) $\frac{\bar{x}_1\bar{x}_2}{\bar{x}_3}$ ii) $\frac{1}{\bar{x}_1\bar{x}_2}$ στρογγυλευμένα στο πλήθος των ακριβών σημαντικών ψηφίων τους.

12. Να υπολογιστεί η διαφορά $70 - \sqrt{4899}$ με την μεγαλύτερη ακρίβεια, αν είναι γνωστό ότι $\sqrt{4899} \approx 69.99$ (με ακρίβεια 4 σ.ψ.) .

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

13. Δίνεται η εξίσωση $f(x) = 0$. Αν x_1 είναι μία ρίζα αυτής και $x_1 \approx \bar{x}_1$ να εκτιμηθεί το απόλυτο σφάλμα της $\bar{x}_1$.

14. Να εκτιμηθεί το απόλυτο (διαδοθέν) σφάλμα της εκφράσεως $6.80x^2 + 3.25x + 8.24$, όπου $x = 10.00$, με την προυπόθεση ότι όλοι οι δοθέντες αριθμοί είναι στρογγυλευμένοι σε 2 δ.θ. .

15. i) Ο προσεγγιστικός αριθμός $\bar{x} = 45.1267$ είναι ακριβής σε 4 σ. ψ. . Να βρεθεί ένα άνω φράγμα για το σχετικό σφάλμα του. ii) Για το σχετικό σφάλμα δ του αριθμού $\bar{x} = 39.257587$ ισχύει $\delta \leq 10^{-5}$. Πόσα από τα σ.ψ. του $\bar{x}$ είναι ακριβή ;

16. Αν οι προσεγγιστικοί αριθμοί $\bar{x}_i$, $\bar{y}_i$, $i = 1(1)n$ έχουν το ίδιο άνω φράγμα M για τα απόλυτα σφάλματά τους, να βρεθεί ένα άνω φράγμα για το απόλυτο σφάλμα του $\sum_{i=1}^{n} \bar{x}_i \bar{y}_i$.

17. Να εκτιμηθεί το απόλυτο (διαδοθέν) σφάλμα για τις δύο ισοδύναμες εκφράσεις $\ln(x - \sqrt{x^2 - 1}) = -\ln(x + \sqrt{x^2 - 1})$ αν $x = 30$ (ακριβής τιμή) και η τετραγωνική ρίζα δίνεται με ακρίβεια 6 δ.θ. ($\sqrt{899} \approx 29.983329$).

18. Έστω $f(x) = \frac{\pi x}{1+x^2}$. Ας υποθέσουμε ότι ζητούμε την τιμή $f(0.16037)$ αλλά μπορούμε να χρησιμοποιούμε το πολύ 3 σ.ψ. στους υπολογισμούς. Να βρεθεί το διαδοθέν σφάλμα, το παραχθέν σφάλμα και το ολικό σφάλμα.

19. Αν οι αριθμοί $\bar{x}_1 = 2$, $\bar{x}_2 = 3$, $\bar{x}_3 = 4$ προσεγγίζουν αντίστοιχα τις ακριβείς τιμές x_1, x_2, x_3 για τις οποίες ισχύει $1.5 \leq x_1 \leq 2.5$, $2.9 \leq x_2 \leq 3.1$ και $3 \leq x_3 \leq 5$, να βρεθεί ακριβώς το μέγιστο απόλυτο διαδοθέν σφάλμα στην τιμή $\frac{\bar{x}_1 \bar{x}_2}{\bar{x}_3}$. Επίσης να εκτιμηθεί αυτό σύμφωνα με την θεωρία.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

2

ΠΑΡΕΜΒΟΛΗ

2.1 Εισαγωγή

Συχνά μια άγνωστη ή γνωστή συνάρτηση $y = f(x)$ ορίζεται από ένα πίνακα τιμών, δηλαδή,ένα πίνακα (αριθμών) που περιέχει τις τιμές της $f(x)$ για ένα σύνολο τιμών του x. Οι πινακοποιημένες τιμές μπορεί να είναι είτε εμπειρικά δεδομένα που προέρχονται από μετρήσεις ή παρατηρήσεις στην περίπτωση που είναι άγνωστη η μορφή της $f(x)$, είτε να έχουν προκύψει με κάποια ορισμένη ακρίβεια από μια γνωστή και σχετικά πολύπλοκη συνάρτηση $f(x)$, όπως συμβαίνει συνήθως με τους μαθηματικούς πίνακες. Ένα πρόβλημα που προκύπτει είναι ο υπολογισμός της $f(x)$ για κάποιες τιμές του x που δεν ανήκουν στον πίνακα (π. χ. ενδιάμεσες τιμές). Για το σκοπό αυτό ζητούμε να προσδιορίσουμε μια άλλη απλούστερη συνάρτηση $F(x)$ (π.χ. ένα πολυώνυμο) τέτοια ώστε για μια οποιαδήποτε συγκεκριμένη τιμή $\hat{x}$ του x η τιμή $F(\hat{x})$ να είναι αρκετά κοντά στην τιμή $f(\hat{x})$.

Ας υποθέσουμε ότι σε ένα πίνακα δίνονται οι $n+1$ τιμές $f_i \equiv f(x_i)$, $i = 0(1)n$ της $f(x)$ για τις αντίστοιχες διακεκριμένες τιμές x_i, $i = 0(1)n$ του x. Ο προσδιορισμός μιας κατάλληλης απλής συναρτήσεως $F(x)$, η οποία θα βρίσκεται πιο "κοντά", υπό κάποια

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

έννοια, στη συνάρτηση $f(x)$ είναι ένα γενικώτερο πρόβλημα της θεωρίας προσεγγίσεως (βλέπε κεφάλαιο 12). Η παρεμβολή που εξετάζεται στο παρόν κεφάλαιο αποτελεί μια ειδική περίπτωση του προβλήματος αυτού. Το απλούστερο πρόβλημα της παρεμβολής συνίσταται στην κατασκευή μιας "απλής" συναρτήσεως $F(x)$ (συνάρτηση παρεμβολής) η οποία παίρνει τις ίδιες τιμές με την $f(x)$ στα σημεία x_i , $i = 0(1)n$ του πίνακα , δηλαδή, απαιτούμε να ισχύει

$$F(x_i) = f_i \ , \quad i = 0(1)n \ . \tag{2.1}$$

Τα σημεία x_i , $i = 0(1)n$ καλούνται σημεία παρεμβολής. Γεωμετρικά, αυτό σημαίνει ότι η καμπύλη $y = F(x)$ πρέπει να διέρχεται από όλα τα δοθέντα σημεία $M_i(x_i,f_i)$, $i = 0(1)n$. Στη συνέχεια θα περιοριστούμε μόνον στην περίπτωση που η $F(x)$ είναι ένα πολυώνυμο $P_n(x)$ βαθμού το πολύ n , δηλαδή, στην καλούμενη πολυωνυμική παρεμβολή. Έτσι ζητούμε ένα πολυώνυμο $P_n(x)$ τέτοιο ώστε

$$P_n(x_i) = f_i \ , \quad i = 0(1)n \ . \tag{2.2}$$

Το πολυώνυμο παρεμβολής $P_n(x)$ χρησιμοποιείται πρωταρχικά για να προσεγγίσει τις τιμές της συναρτήσεως $f(x)$ για τιμές του x διαφορετικές από τα σημεία παρεμβολής $(f(x) \approx P_n(x))$. Εκείνο όμως που έχει μεγαλύτερη σημασία είναι το γεγονός ότι πολλές αριθμητικές μέθοδοι που χρησιμοποιούνται για τη λύση άλλων προβλημάτων (π.χ. αριθμητική παραγώγιση, αριθμητική ολοκλήρωση, αριθμητική επίλυση διαφορικών εξισώσεων κ.λ.π.) βασίζονται στην ιδέα της πολυωνυμικής παρεμβολής.

2.2 Πολυωνυμική παρεμβολή

Έστω

$$P_n(x) = \alpha_0 x^n + \alpha_1 x^{n-1} + \ldots + \alpha_{n-1} x + \alpha_n \tag{2.3}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

το ζητούμενο πολυώνυμο παρεμβολής βαθμού το πολύ n . Ζητούμε να προσδιορίσουμε τους $n+1$ συντελεστές α_i , $i = 0(1)n$ έτσι ώστε να ισχύει η (2.2). Προφανώς τα α_i , $i = 0(1)n$ πρέπει να ικανοποιούν το ακόλουθο σύστημα των $n+1$ γραμμικών αλγεβρικών εξισώσεων

$$\begin{cases} x_0^n\alpha_0 + x_0^{n-1}\alpha_1 + \ldots + x_0\alpha_{n-1} + \alpha_n = f_0 \\ x_1^n\alpha_0 + x_1^{n-1}\alpha_1 + \ldots + x_1\alpha_{n-1} + \alpha_n = f_1 \\ \vdots \qquad\qquad\qquad\qquad \vdots \\ x_n^n\alpha_0 + x_n^{n-1}\alpha_1 + \ldots + x_n\alpha_{n-1} + \alpha_n = f_n \end{cases} \tag{2.4}$$

Το σύστημα (2.4) έχει μία μοναδική λύση, αφού η ορίζουσα των συντελεστών των αγνώστων α_i, $i = 0(1)n$ είναι η καλούμενη ορίζουσα Vandermonde

$$D = \begin{vmatrix} x_0^n & x_0^{n-1} & \ldots & x_0 & 1 \\ x_1^n & x_1^n & \ldots & x_1 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ x_n^n & x_n^{n-1} & \ldots & x_n & 1 \end{vmatrix} \tag{2.5}$$

της οποίας η τιμή είναι $D = \prod_{\substack{i,j=0 \\ i<j}}^{n} (x_i - x_j) \neq 0$ επειδή τα x_i, $i = 0(1)n$ είναι διακεκριμένα εξ υποθέσεως. Επομένως, λύνοντας το σύστημα (2.4) π.χ. με την μέθοδο απαλοιφής του Gauss (βλέπε κεφάλαιο 10) προσδιορίζεται το $P_n(x)$. Έτσι απεδείχθη ότι το πολυώνυμο παρεμβολής υπάρχει πάντοτε και είναι μοναδικό. Ένας καλύτερος τρόπος προσδιορισμού του $P_n(x)$ που αποφεύγει την λύση του συστήματος (2.4) δίνεται στη συνέχεια.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

2.2.1 Παρεμβολή κατα Lagrange

Εάν είναι δυνατόν να κατασκευασθούν πολυώνυμα $L_i(x)$, $i = 0(1)n$ βαθμού το πολύ n (θα αποδειχθεί ότι όλα είναι βαθμού ακριβώς n) τέτοια ώστε

$$L_i(x_j)=\begin{cases} 1 & \text{εάν} \quad j = i \\ 0 & \text{εάν} \quad j \neq i \end{cases} \tag{2.6}$$

όπου $i,j = 0(1)n$, τότε το πολυώνυμο

$$P_n(x) = \sum_{j=0}^{n} L_j(x)f_j \quad , \tag{2.7}$$

που είναι βαθμού το πολύ n θα είναι το ζητούμενο πολυώνυμο παρεμβολής. Πράγματι θα έχουμε

$$P_n(x_i)= \sum_{j=0}^{n} L_j(x_i)f_j = L_i(x_i)f_i = f_i \quad , \quad i = 0(1)n \quad .$$

Αφού $L_i(x_j) = 0$ για $j \neq i$, προκύπτει ότι το $L_i(x)$ διαιρείται με το $\prod_{\substack{j=0 \\ j\neq i}}^{n} (x - x_j)$ και επειδή το $L_i(x)$ είναι βαθμού το πολύ n θα ισχύει

$$L_i(x) = c_i \prod_{\substack{j=0 \\ j\neq i}}^{n} (x - x_j) \quad ,$$

όπου c_i είναι σταθερά. Από την συνθήκη $L_i(x_i) = 1$ προσδιορίζεται η c_i , οπότε παίρνουμε τελικά

$$L_i(x) = \prod_{\substack{j=0 \\ j\neq i}}^{n} \left(\frac{x - x_j}{x_i - x_j}\right) \quad , \quad i = 0(1)n \tag{2.8}$$

και επομένως κάθε $L_i(x)$ είναι προφανώς πολυώνυμο βαθμού n και προσδιορίζεται μονοσήμαντα. Βάσει της (2.8) το πολυώνυμο παρεμβολής (2.7) γράφεται

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$P_n(x) = \sum_{i=0}^{n} f_i \prod_{\substack{j=0 \\ j\neq i}}^{n} \left(\frac{x - x_j}{x_i - x_j}\right) \tag{2.9}$$

και καλείται πολυώνυμο παρεμβολής του Lagrange. Τα πολυώνυμα (2.8) καλούνται συντελεστές παρεμβολής του Lagrange, η δε (2.9) καλείται και τύπος παρεμβολής του Lagrange. Από τον τύπο (2.9) είναι φανερό ότι οι συντελεστές του πολυωνύμου $P_n(x)$ είναι γραμμικές εκφράσεις των τιμών f_i , $i = 0(1)n$. Εάν θεωρήσουμε την συνάρτηση

$$\omega(x) = \prod_{j=0}^{n} (x - x_j) \tag{2.10}$$

τότε ισχύει $\omega'(x) = \sum_{\kappa=0}^{n} \left[\prod_{\substack{j=0 \\ j\neq\kappa}}^{n} (x - x_j)\right]$. Για $x = x_i$, $\kappa \neq i$ η παράσταση που βρίσκεται μέσα στις αγκύλες μηδενίζεται. Επομένως έχουμε

$$\omega'(x_i) = \prod_{\substack{j=0 \\ j\neq i}}^{n} (x_i - x_j) \ , \ i = 0(1)n \ . \tag{2.11}$$

Λόγω των (2.10) και (2.11), οι (2.8) και (2.9) γίνονται

$$L_i(x) = \frac{\omega(x)}{(x - x_i)\omega'(x_i)} \ , \quad i = 0(1)n \tag{2.12}$$

και

$$P_n(x) = \omega(x) \sum_{i=0}^{n} \frac{f_i}{(x - x_i)\omega'(x_i)} \ . \tag{2.13}$$

θα δειχθεί ότι για τους συντελεστές παρεμβολής του Lagrange (2.8) ισχύει

$$\sum_{i=0}^{n} L_i(x) = 1 \ . \tag{2.14}$$

Πράγματι, το πολυώνυμο $R(x) \equiv \sum_{i=0}^{n} L_i(x) - 1$ είναι βαθμού το πολύ n και ισχύει $R(x_i) = 0$, $i = 0(1)n$ λόγω των (2.6). Άρα $R(x) \equiv 0$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

και επομένως ισχύει η (2.14). Η ταυτότητα (2.14) είναι πολύ χρήσιμη για τον έλεγχο και μερικές φορές για την μείωση των υπολογισμών.

Για $n = 1$ έχουμε δύο σημεία παρεμβολής x_0, x_1 με f_0, f_1 τις αντίστοιχες τιμές της συναρτήσεως. Τότε, το πολυώνυμο παρεμβολής του Lagrange είναι το

$$P_1(x) = \frac{x - x_1}{x_0 - x_1} f_0 + \frac{x - x_0}{x_1 - x_0} f_1 = f_0 + \frac{f_1 - f_0}{x_1 - x_0}(x - x_0), \quad (2.15)$$

δηλαδή, η ευθεία που διέρχεται από τα δύο σημεία $(x_0, f_0), (x_1, f_1)$. Η παρεμβολή στην περίπτωση αυτή καλείται γραμμική.

Για $n = 2$ ο τύπος παρεμβολής του Lagrange είναι η εξίσωση της παραβολής που διέρχεται από τα τρία σημεία $(x_0, f_0), (x_1, f_1), (x_2, f_2)$ (τετραγωνική παρεμβολή)

$$P_2(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} f_0 + \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} f_1 + \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} f_2 . \quad (2.16)$$

<u>Παράδειγμα 2.1</u> Δίνεται ο πίνακας τιμών της $f(x)$

x_i	-1	0	1	2
f_i	2	2	2	-1

Να βρεθεί το πολυώνυμο παρεμβολής με σημεία παρεμβολής τα x_i, $i = 0(1)3$ και στη συνέχεια να προσεγγιστεί η τιμή $f(\frac{1}{2})$.

<u>Λύση</u> Το ζητούμενο πολυώνυμο παρεμβολής βαθμού το πολύ 3 θα είναι λόγω της (2.7), $P_3(x) = L_0(x)f_0 + L_1(x)f_1 + L_2(x)f_2 + L_3(x)f_3$, όπου τα $L_i(x)$, $i = 0(1)3$ δίνονται από την (2.8). Άρα $P_3(x) = 2\left[L_0(x) + L_1(x) + L_2(x)\right] - L_3(x)$ ή, λόγω της (2.14), $P_3(x) = 2\left[1 - L_3(x)\right] - L_3(x) = 2 - 3L_3(x)$, όπου

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$L_3(x) = \frac{(x-x_0)(x-x_1)(x-x_2)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)} = \frac{(x+1)(x-0)(x-1)}{(2+1)(2-0)(2-1)} =$$

$$= \frac{1}{6}(x^2-1)x \ .$$

Έτσι έχουμε $P_3(x) = 2 - 3\,\frac{1}{6}(x^2-1)x = -\frac{1}{2}(x^3-x-4)$ και $f(\frac{1}{2}) \approx P_3(\frac{1}{2}) = \frac{35}{16} = 2.1875$. ▲

Πρέπει να σημειωθεί ότι η παρεμβολή κατά Lagrange χρησιμοποιείται γενικά για μη ισαπέχοντα σημεία παρεμβολής. Μπορεί βεβαίως να χρησιμοποιηθεί και για ισαπέχοντα σημεία παρεμβολής,σαν ειδική περίπτωση, αλλά στην περίπτωση αυτή το πολυώνυμο παρεμβολής μπορεί να προσδιοριστεί ευκολώτερα με άλλους τύπους όπως θα δειχθεί σε επόμενες παραγράφους. Ο τύπος παρεμβολής του Lagrange έχει περισσότερο θεωρητική αξία παρά πρακτική. Στις πρακτικές εφαρμογές και ειδικά όταν ο αριθμός των σημείων παρεμβολής είναι μεγάλος,πρέπει να χρησιμοποιείται με μεγάλη προσοχή,επειδή το πολυώνυμο παρεμβολής μεγάλου βαθμού μπορεί να μην προσεγγίζει αρκετά καλά την (άγνωστη) ομαλή συνάρτηση $f(x)$ μεταξύ των σημείων παρεμβολής. Η παρεμβολή κατά Lagrange είναι οπωσδήποτε χρήσιμη σε περιπτώσεις που πολλά προβλήματα παρεμβολής πρόκειται να λυθούν με τα ίδια σημεία παρεμβολής x_i , $i = 0(1)n$ αλλά με διαφορετικά σύνολα τιμών f_i , $i = 0(1)n$ του πίνακα.

2.2.2 Εκτίμηση του σφάλματος στην πολυωνυμική παρεμβολή

Με την εύρεση του πολυωνύμου παρεμβολής $P_n(x)$ από τα σημεία παρεμβολής x_i , $i = 0(1)n$ και τις τιμές $f_i \equiv f(x_i)$, $i = 0(1)n$ της συναρτήσεως $f(x)$, η τιμή της $f(x)$ σε ένα σημείο $\hat{x} \neq x_i$, $i = 0(1)n$ προσεγγίζεται από την τιμή $P_n(\hat{x})$, δηλαδή, $f(\hat{x}) \approx P_n(\hat{x})$. Η προσέγγιση αυτή δημιουργεί ένα σφάλμα αποκοπής το οποίο δίνεται από την

$$\varepsilon(\hat{x}) = P_n(\hat{x}) - f(\hat{x}) \qquad (2.17)$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Το σφάλμα αυτό μπορεί μερικές φορές να εκτιμηθεί με την βοήθεια του ακολούθου θεωρήματος.

<u>Θεώρημα 2.1</u> Εάν η συνάρτηση $f(x)$ είναι παραγωγίσιμη $n+1$ φορές στο $I \equiv [\alpha,b]$, όπου $\alpha = \min\{x_0,x_1,\ldots,x_n,\hat{x}\}$, $b = \max\{x_0,x_1,\ldots,x_n,\hat{x}\}$, τότε για κάθε $\hat{x}$ υπάρχει ένας αριθμός $\xi \in I$ τέτοιος ώστε

$$\varepsilon(\hat{x}) = -\frac{\omega(\hat{x})f^{(n+1)}(\xi)}{(n+1)!} ,$$

όπου η συνάρτηση $\omega(x)$ δίνεται από την (2.10).

<u>Απόδειξη</u> Θεωρούμε την βοηθητική συνάρτηση

$$\Phi(x) \equiv f(x) - P_n(x) - c\omega(x) , \tag{2.18}$$

όπου c είναι μία σταθερά η οποία εκλέγεται έτσι ώστε $\Phi(\hat{x}) = 0$. Έτσι έχουμε

$$c = \frac{f(\hat{x}) - P_n(\hat{x})}{\omega(\hat{x})} , \tag{2.19}$$

όπου $\omega(\hat{x}) \neq 0$, αφού $\hat{x} \neq x_i$, $i = 0(1)n$. Έτσι η συνάρτηση $\Phi(x)$ με το c να δίνεται από την (2.19) μηδενίζεται στα $n+2$ σημεία: x_0, $x_1,\ldots,x_n$, $\hat{x}$ του διαστήματος I. Σύμφωνα με το θεώρημα του Rolle, εάν αυτό εφαρμοσθεί επαναληπτικά, συμπεραίνουμε ότι η $\Phi'(x)$ έχει τουλάχιστον $n+1$ ρίζες στο I, η $\Phi''(x)$ έχει τουλάχιστον n ρίζες και τέλος η $\Phi^{(n+1)}(x)$ έχει τουλάχιστον μία ρίζα $\xi \in I$. Αφού $P_n^{(n+1)}(x) \equiv 0$ και $\omega^{(n+1)}(x) = (n+1)!$, από την (2.18) παίρνουμε

$$\Phi^{(n+1)}(x) = f^{(n+1)}(x) - c(n+1)! \tag{2.20}$$

και επομένως $\Phi^{(n+1)}(\xi) = 0 = f^{(n+1)}(\xi) - c(n+1)!$ ή

$$c = \frac{f^{(n+1)}(\xi)}{(n+1)!} . \tag{2.21}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Από τις (2.19) και (2.20) προκύπτει ότι

$$f(\hat{x}) - P_n(\hat{x}) = \frac{\omega(\hat{x}) f^{(n+1)}(\xi)}{(n+1)!} \tag{2.22}$$

ή

$$\varepsilon(\hat{x}) = -\frac{\omega(\hat{x}) f^{(n+1)}(\xi)}{(n+1)!} \quad . \qquad \blacktriangle \tag{2.23}$$

<u>Παρατηρήσεις</u> i) Αφού το $\hat{x}$ είναι αυθαίρετο μπορούμε αντί της (2.23) να γράψουμε

$$\varepsilon(x) = -\frac{\omega(x) f^{(n+1)}(\xi)}{(n+1)!} \quad , \tag{2.24}$$

όπου το ξ εξαρτάται από το x , δηλαδή, $\xi = \xi(x)$ και $\xi \in [\min\{x_0, x_1, \ldots, x_n, x\}, \max\{x_0, x_1, \ldots, x_n, x\}]$.

ii) Πρέπει να σημειωθεί ότι ο τύπος (2.23) ισχύει και όταν το $\hat{x}$ ισούται με κάποιο από τα σημεία παρεμβολής. Πράγματι, αφού $\omega(x_i) = 0$, $i = 0(1)n$ προκύπτει ότι $\varepsilon(x_i) = 0$, $i = 0(1)n$ όπως ανεμένετο λόγω των $f(x_i) = P_n(x_i)$, $i = 0(1)n$.

iii) Αν $M_{n+1} = \max\limits_{x \in I} \left| f^{(n+1)}(x) \right|$, τότε για το απόλυτο σφάλμα αποκοπής ισχύει

$$|\varepsilon(\hat{x})| \leq \frac{M_{n+1}}{(n+1)!} |\omega(\hat{x})| \quad . \tag{2.25}$$

iv) Για να εφαρμοσθεί ο τύπος (2.23) του σφάλματος πρέπει απαραιτήτως να είναι γνωστή η συνάρτηση $f(x)$ (οπότε θα είναι και η $f^{(n+1)}(x)$) , πράγμα που συνήθως δεν συμβαίνει στην πράξη. Αλλά και αν ακόμη η $f(x)$ είναι γνωστή, η (2.23) δεν μπορεί να δώσει μία ακριβή εκτίμηση του σφάλματος, επειδή η ακριβής θέση του σημείου ξ δεν είναι γνωστή.

2.2.3 Η μέθοδος ή αλγόριθμος του Aitken

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Με την μέθοδο του Aitken είναι δυνατόν να υπολογίσουμε το πολυώνυμο παρεμβολής $P_n(x)$ ή την τιμή του σε ένα σημείο $\hat{x}$ διαφορετικό από τα σημεία παρεμβολής x_i , $i = 0(1)n$, χρησιμοποιώντας μία ακολουθία από γραμμικές παρεμβολές. Χάριν απλότητος υποθέτουμε ότι $n = 4$, δηλαδή, δίνονται σε ένα πίνακα τα σημεία x_i , $i = 0(1)4$ και οι αντίστοιχες τιμές f_i , $i = 0(1)4$. Καλούμε $I_{0,j}(x) \equiv I_{0,j}$ το πολυώνυμο παρεμβολής το πολύ πρώτου βαθμού που διέρχεται από τα δύο σημεία (x_0,f_0) και (x_j,f_j) , όπου $j = 1,2,3,4$. Σύμφωνα με την (2.15) είναι

$$I_{0,j} = \frac{(x_j - x)f_0 + (x - x_0)f_j}{x_j - x_0} \quad , \quad j = 1(1)4 \ . \tag{2.26}$$

Καλούμε $I_{0,1,j}(x) \equiv I_{0,1,j}$ το πολυώνυμο που κατασκευάζεται σαν να είναι το πολυώνυμο παρεμβολής που διέρχεται από τα δύο σημεία $(x_1, I_{0,1})$ και $(x_j, I_{0,j})$, όπου $j = 2,3,4$. Είναι

$$I_{0,1,j} = \frac{(x_j - x)I_{0,1} + (x - x_1)I_{0,j}}{x_j - x_1} \quad , \quad j = 2,3,4 \ . \tag{2.27}$$

Κατ' αναλογίαν ορίζουμε τα πολυώνυμα

$$I_{0,1,2,j}(x) \equiv I_{0,1,2,j} = \frac{(x_j - x)I_{0,1,2} + (x - x_2)I_{0,1,j}}{x_j - x_2} \ , \tag{2.28}$$

$$j = 3,4$$

και

$$I_{0,1,2,3,4}(x) \equiv I_{0,1,2,3,4} = \frac{(x_4 - x)I_{0,1,2,3} + (x - x_3)I_{0,1,2,4}}{x_4 - x_3} \ . \tag{2.29}$$

Με βάση τις (2.26) – (2.29) σχηματίζουμε τον ακόλουθο πίνακα

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

x_i	f_i				
x_0	f_0				
		$I_{0,1}$			
x_1	f_1		$I_{0,1,2}$		
		$I_{0,2}$		$I_{0,1,2,3}$	
x_2	f_2		$I_{0,1,3}$		$I_{0,1,2,3,4}$
		$I_{0,3}$		$I_{0,1,2,4}$	
x_3	f_3		$I_{0,1,4}$		
		$I_{0,4}$			
x_4	f_4				

Πίνακας 2.1

Προφανώς είναι

$$I_{0,j}(x_0) = f_0 \quad \text{και} \quad I_{0,j}(x_j) = f_j \ , \quad j = 1,2,3,4 \ . \tag{2.30}$$

Επίσης, παίρνουμε λόγω των (2.30)

$$\left.\begin{aligned}
I_{0,1,j}(x_0) &= \frac{(x_j - x_0)I_{0,1}(x_0) + (x_0 - x_1)I_{0,j}(x_0)}{x_j - x_1} = \\
&= \frac{(x_j - x_0)f_0 + (x_0 - x_1)f_0}{x_j - x_1} = f_0 \ , \\
I_{0,1,j}(x_1) &= \frac{(x_j - x_1)I_{0,1}(x_1) + (x_1 - x_1)I_{0,j}(x_1)}{x_j - x_1} = f_1 \ , \\
I_{0,1,j}(x_j) &= \frac{(x_j - x_j)I_{0,1}(x_j) + (x_j - x_1)I_{0,j}(x_j)}{x_j - x_1} = \\
&= I_{0,j}(x_j) = f_j \ , \qquad j = 2,3,4 \ .
\end{aligned}\right\} \tag{2.31}$$

Συνεχίζοντας την ίδια διαδικασία βρίσκουμε ότι

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$I_{0,1,2,j}(x_K) = f_K \quad , \kappa = 0,1,2,j \quad , \quad j = 3,4 \quad ,$$
$$I_{0,1,2,3,4}(x_K) = f_K \quad , \kappa = 0,1,2,3,4 \quad . \tag{2.32}$$

Συνεπώς το πολυώνυμο $I_{0,1,2,3,4}(x)$, που είναι προφανώς το πολύ τετάρτου βαθμού, ταυτίζεται με το πολυώνυμο παρεμβολής $P_4(x)$ του Lagrange. Στην περίπτωση που το x έχει την συγκεκριμένη τιμή $\hat{x}$ τότε ο πίνακας 2.1 είναι ένας πίνακας αριθμών και είναι $P_4(\hat{x}) = I_{0,1,2,3,4}$. Ένα βασικό πλεονέκτημα της μεθόδου του Aitken είναι ότι μπορούμε να προσθέτουμε και άλλα σημεία παρεμβολής χωρίς να χρειάζεται να εκτελέσουμε όλους τους υπολογισμούς από την αρχή, πράγμα που δεν συμβαίνει στην παρεμβολή κατά Lagrange.

<u>Παράδειγμα 2.2</u> Δίνεται ο πίνακας τιμών

x_i	0	1	3
f_i	1	3	2

. Με την μέθοδο του Aitken να βρεθεί μία προσέγγιση για την $f(2)$.

<u>Λύση</u> Είναι $x_0 = 0$, $x_1 = 1$, $x_2 = 3$, $f_0 = 1$, $f_1 = 3$, $f_2 = 2$; $\hat{x} = 2$ και επομένως

$$I_{0,1} = \frac{(x_1 - \hat{x})f_0 + (\hat{x} - x_0)f_1}{x_1 - x_0} = \frac{(1-2)1 + (2-0)3}{1-0} = 5 \quad ,$$

$$I_{0,2} = \frac{(x_2 - \hat{x})f_0 + (\hat{x} - x_0)f_2}{x_2 - x_0} = \frac{(3-2)1 + (2-0)2}{3-0} = \frac{5}{3} \quad ,$$

$$I_{0,1,2} = \frac{(x_2 - \hat{x})5 + (\hat{x} - x_1)\frac{5}{3}}{x_2 - x_1} = \frac{(3-2)5 + (2-1)\frac{5}{3}}{3-1} = \frac{10}{3} \quad .$$

Άρα, $f(2) \approx I_{0,1,2} = \frac{10}{3}$. Ο αντίστοιχος του πίνακα 2.1 είναι

x_i	f_i		
0	1		
1	3	$I_{0,1} = 5$ $I_{0,2} = \frac{5}{3}$	$I_{0,1,2} = \frac{10}{3}$
3	2		

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

2.2.4 Αντίστροφη παρεμβολή

Ας υποθέσουμε ότι σ' ένα πίνακα δίνονται οι τιμές f_i , $i = 0(1)n$ της συναρτήσεως $y = f(x)$ στα αντίστοιχα σημεία x_i , $i = 0(1)n$. Το πρόβλημα της αντίστροφης παρεμβολής συνίσταται στον προσδιορισμό μιάς προσεγγιστικής τιμής $\bar{x}$ του σημείου x στο οποίο η $f(x)$ παίρνει την δοσμένη τιμή $\hat{y}$, δηλαδή,

$$f(x) = \hat{y} \ . \tag{2.33}$$

Ένας τρόπος για την λύση του προβλήματος αυτού είναι να κατασκευάσουμε το πολυώνυμο παρεμβολής $P_n(x)$, οπότε $f(x) \approx P_n(x)$ και στη συνέχεια να λύσουμε την πολυωνυμική εξίσωση

$$P_n(x) = \hat{y} \ . \tag{2.34}$$

Για παράδειγμα, άν $n = 1$ τότε η (2.34) γίνεται

$$P_1(x) = \frac{x - x_1}{x_0 - x_1} f_0 + \frac{x - x_0}{x_1 - x_0} f_1 = \hat{y}$$

και επομένως η λύση αυτής είναι

$$\bar{x} = \frac{\hat{y} - f_1}{f_0 - f_1} x_0 + \frac{\hat{y} - f_0}{f_1 - f_0} x_1 \ . \tag{2.35}$$

Ένας άλλος απλούστερος τρόπος για την λύση του προβλήματος της αντίστροφης παρεμβολής προκύπτει αν υποθέσουμε ότι υπάρχει η αντίστροφη συνάρτηση της $y = f(x)$. Αυτό συμβαίνει στα διαστήματα που η $f(x)$ είναι γνησίως μονότονος. Έτσι από την (2.33) παίρνουμε $x = f^{-1}(\hat{y})$ που σημαίνει ότι αρκεί να προσεγγίσουμε την τιμή της συναρτήσεως $x = f^{-1}(y) \equiv g(y)$ στο σημείο $\hat{y}$. Αλλά αυτό είναι το γνωστό πρόβλημα της παρεμβολής και λύνεται αφού κατασκευασθεί το πολυώνυμο παρεμβολής $P_n(y)$ που διέρχεται από τα σημεία (f_i, x_i) , $i = 0(1)n$, δηλαδή, αντιστρέφοντας τους ρόλους των x_i, f_i , $i = 0(1)n$. Επομένως, αφού $g(y) \approx P_n(y)$, θα είναι

$$\bar{x} = P_n(\hat{y}) \ .$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Πρέπει να σημειωθεί ότι οι τιμές της ανεξάρτητης μεταβλητής y της συναρτήσεως $x = g(y)$, δηλαδή, οι τιμές f_i , $i = 0(1)n$ που δίνονται στον πίνακα είναι σχεδόν πάντοτε μη ισαπέχουσες και επομένως η παρεμβολή κατά Lagrange είναι κατάλληλη. Για παράδειγμα ας θεωρήσουμε πάλι την περίπτωση $n = 1$. Τότε, το πολυώνυμο παρεμβολής $P_1(y)$ είναι

$$P_1(y) = \frac{y - f_1}{f_0 - f_1} x_0 + \frac{y - f_0}{f_1 - f_0} x_1$$

και επομένως

$$\bar{x} = P_1(\hat{y}) = \frac{\hat{y} - f_1}{f_0 - f_1} x_0 + \frac{\hat{y} - f_0}{f_1 - f_0} x_1 \quad ,$$

που συμπίπτει με την (2.35). Για $n > 1$ οι δύο τρόποι που αναφέρθηκαν δεν δίνουν το ίδιο αποτέλεσμα. Μία σημαντική εφαρμογή της αντίστροφης παρεμβολής είναι η εύρεση προσεγγιστικών τιμών για τις ρίζες μιας συναρτήσεως $f(x)$ (περίπτωση $\hat{y} = 0$).

<u>Παράδειγμα 2.3</u> Αν είναι γνωστό ότι $f(-4) = 4$, $f(-2) = 1$ και $f(-1) = -3$, να εκτιμηθεί μία ρίζα της $f(x)$ με την αντίστροφη παρεμβολή.

<u>Λύση</u> Είναι $\hat{y} = 0$ και $x_0 = -4$, $x_1 = -2$, $x_2 = -1$, $f_0 = 4$, $f_1 = 1$, $f_2 = -3$.

<u>1ος τρόπος</u>: Είναι $P_2(x) = \dfrac{(x+2)(x+1)}{(-4+2)(-4+1)} 4 + \dfrac{(x+4)(x+1)}{(-2+4)(-2+1)} 1 +$

$$+ \frac{(x+4)(x+2)}{(-1+4)(-1+2)} (-3) = -\frac{1}{6} (5x^2+39x+52).$$

Οι ρίζες της $P_2(x) = 0$ είναι $\xi_1 \approx -1.70683$ και $\xi_2 \approx -6.09317$. Άρα $\bar{x} = -1.70683$ είναι η ζητούμενη προσεγγιστική τιμή της ρίζας.

<u>2ος τρόπος</u>: Το πολυώνυμο παρεμβολής $P_2(y)$ που διέρχεται από τα σημεία $(4, -4), (1, -2)$ και $(-3, -1)$ είναι

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$P_2(y) = \frac{(y-1)(y+3)}{(4-1)(4+3)}(-4) + \frac{(y-4)(y+3)}{(1-4)(1+3)}(-2) +$$

$$+ \frac{(y-4)(y-1)}{(-3-4)(-3-1)}(-1) .$$

Αρα,

$$\bar{x} = P_2(\hat{y}) = P_2(0) = -\frac{11}{7} \approx -1.57143 . \qquad \blacktriangle$$

2.3 Διηρημένες διαφορές

Υποθέτουμε ότι η συνάρτηση $y = f(x)$ είναι πινακοποιημένη στα διακεκριμένα (γενικά μη ισαπέχοντα) σημεία x_i, $i = 0(1)n$ και $f_i \equiv f(x_i)$, $i = 0(1)n$ είναι οι αντίστοιχες τιμές της. Οι διηρημένες διαφορές ορίζονται ως εξής: Η διηρημένη διαφορά μηδενικής τάξεως στο σημείο x_i συμβολίζεται $f[x_i]$ και είναι

$$f[x_i] \equiv f_i \quad , \quad i = 0(1)n .$$

Η διηρημένη διαφορά πρώτης τάξεως των σημείων x_i, x_j ορίζεται από την σχέση

$$f[x_i,x_j] = \frac{f[x_j] - f[x_i]}{x_j - x_i} = \frac{f_j - f_i}{x_j - x_i} . \tag{2.36}$$

Γεωμετρικά η διηρημένη διαφορά πρώτης τάξεως είναι η κλίση μιας χορδής στην γραφική παράσταση της $y = f(x)$. Η διηρημένη διαφορά δευτέρας τάξεως των σημείων x_i,x_j,x_ℓ ορίζεται από την σχέση

$$f[x_i,x_j,x_\ell] = \frac{f[x_j,x_\ell] - f[x_i,x_j]}{x_\ell - x_i} \tag{2.37}$$

και γενικά η διηρημένη διαφορά κ – τάξεως $(1 \leq \kappa \leq n)$, $f[x_i,x_{i+1}, x_{i+2},\ldots,x_{i+\kappa}]$ ορίζεται συναρτήσει των διηρημένων διαφορών $(\kappa - 1)$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

– τάξεως από τον τύπο:

$$f[x_i,x_{i+1},\ldots,x_{i+K}] = \frac{f[x_{i+1},x_{i+2},\ldots x_{i+K}] - f[x_i,x_{i+1},\ldots,x_{i+K-1}]}{x_{i+K} - x_i} \quad . \tag{2.38}$$

Είναι φανερό ότι από τα δεδομένα x_i, f_i, $i = 0(1)n$ του πίνακα της $f(x)$ μπορούμε να κατασκευάσουμε ένα πίνακα διηρημένων διαφορών της ίδιας μορφής με τον ακόλουθο που κατασκευάστηκε για $n = 4$.

		ΔΙΗΡΗΜΕΝΕΣ ΔΙΑΦΟΡΕΣ			
x_i	f_i	1ης τάξεως	2ας τάξεως	3ης τάξεως	4ης τάξεως
x_0	f_0				
		$f[x_0,x_1]$			
x_1	f_1		$f[x_0,x_1,x_2]$		
		$f[x_1,x_2]$		$f[x_0,x_1,x_2,x_3]$	
x_2	f_2		$f[x_1,x_2,x_3]$		$f[x_0,x_1,x_2,x_3,x_4]$
		$f[x_2,x_3]$		$f[x_1,x_2,x_3,x_4]$	
x_3	f_3		$f[x_2,x_3,x_4]$		
		$f[x_3,x_4]$			
x_4	f_4				

Πίνακας 2.2

<u>Θεώρημα 2.2</u> Για την διηρημένη διαφορά κ – τάξεως που ορίζεται από τα σημεία $x_0,x_1,\ldots,x_K$ ισχύει

$$f[x_0,x_1,\ldots,x_K] = \sum_{i=0}^{K} \frac{f_i}{\prod\limits_{\substack{j=0\\ j\neq i}}^{K} (x_i - x_j)} \quad . \tag{2.39}$$

<u>Απόδειξη</u> Με την βοήθεια της επαγωγικής μεθόδου έχουμε: Για $\kappa = 1$ το πρώτο μέλος της (2.39) είναι $f[x_0,x_1] = \dfrac{f_1 - f_0}{x_1 - x_0}$ (λόγω

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

της (2.36)), ενώ από το δεύτερο μέλος της (2.39) προκύπτει

$$\sum_{i=0}^{1} \frac{f_i}{\prod\limits_{\substack{j=0\\ j\neq i}}^{1} (x_i - x_j)} = \frac{f_0}{x_0 - x_1} + \frac{f_1}{x_1 - x_0} = \frac{f_1 - f_0}{x_1 - x_0} .$$

Υποθέτουμε ότι η (2.39) ισχύει για $\kappa \leq \ell$. Τότε είναι

$$f[x_0, x_1, \ldots, x_\ell, x_{\ell+1}] = \frac{f[x_1, x_2, \ldots, x_{\ell+1}] - f[x_0, x_1, \ldots, x_\ell]}{x_{\ell+1} - x_0} =$$

$$= \frac{1}{x_{\ell+1} - x_0} \left[\sum_{i=1}^{\ell+1} \frac{f_i}{\prod\limits_{\substack{j=1\\ j\neq i}}^{\ell+1} (x_i - x_j)} - \sum_{i=0}^{\ell} \frac{f_i}{\prod\limits_{\substack{j=0\\ j\neq i}}^{\ell} (x_i - x_j)} \right] =$$

$$= \alpha_0 f_0 + \alpha_1 f_1 + \ldots + \alpha_\ell f_\ell + \alpha_{\ell+1} f_{\ell+1} = \sum_{i=0}^{\ell+1} \alpha_i f_i . \qquad (2.40)$$

Αρκεί να δειχθεί ότι

$$\alpha_i = \frac{1}{\prod\limits_{\substack{j=0\\ j\neq i}}^{\ell+1} (x_i - x_j)} , \quad i = 0(1)\ell+1 . \qquad (2.41)$$

Αλλά από την (2.40), για $i \neq 0, \ell+1$ είναι

$$\alpha_i = \frac{1}{(x_{\ell+1} - x_0) \prod\limits_{\substack{j=1\\ j\neq i}}^{\ell+1} (x_i - x_j)} - \frac{1}{(x_{\ell+1} - x_0) \prod\limits_{\substack{j=0\\ j\neq i}}^{\ell} (x_i - x_j)} =$$

$$= \frac{(x_i - x_0)}{(x_{\ell+1} - x_0) \prod\limits_{\substack{j=0\\ j\neq i}}^{\ell+1} (x_i - x_j)} - \frac{(x_i - x_{\ell+1})}{(x_{\ell+1} - x_0) \prod\limits_{\substack{j=0\\ j\neq i}}^{\ell+1} (x_i - x_j)} = \frac{1}{\prod\limits_{\substack{j=0\\ j\neq i}}^{\ell+1} (x_i - x_j)} .$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Επίσης, για $i = 0$, $\ell+1$ από την (2.40) παίρνουμε

$$\alpha_0 = \frac{1}{x_{\ell+1} - x_0}\left[-\frac{1}{\prod\limits_{\substack{j=0\\ j\neq 0}}^{\ell}(x_0 - x_j)}\right] = \frac{1}{\prod\limits_{\substack{j=0\\ j\neq 0}}^{\ell+1}(x_0 - x_j)} \quad \text{και}$$

$$\alpha_{\ell+1} = \frac{1}{x_{\ell+1} - x_0}\left[\frac{1}{\prod\limits_{\substack{j=1\\ j\neq \ell+1}}^{\ell+1}(x_{\ell+1} - x_j)}\right] = \frac{1}{\prod\limits_{\substack{j=0\\ j\neq \ell+1}}^{\ell+1}(x_{\ell+1} - x_j)} \ .$$

Δηλαδή, απεδείχθη η (2.41) για $i = 0(1)\ell+1$ και επομένως η (2.39).▲

<u>Παρατηρήσεις</u> Από την (2.39) προκύπτει ότι:

i) Η $f[x_0,x_1,\ldots,x_K]$ είναι μία συμμετρική συνάρτηση των σημείων $x_0,x_1,\ldots,x_K$ που σημαίνει ότι αυτή δεν μεταβάλλεται με οποιαδήποτε μετάθεση αυτών.

ii) Ισχύει $(c_1 f + c_2 g)[x_0,x_1,\ldots,x_K] = c_1 f[x_0,x_1,\ldots,x_K] + c_2 g[x_0,x_1,\ldots,x_K]$ όπου c_1,c_2 σταθερές και f,g συναρτήσεις.

Σημειώνεται τέλος ότι εάν η $f(x)$ είναι επαρκώς παραγωγίσιμη, τότε η διηρημένη διαφορά $f[x_0,x_1,\ldots,x_K]$ μπορεί να οριστεί και στην περίπτωση που κάποια από τα x_i , $i = 0(1)\kappa$ συμπίπτουν. Για παράδειγμα, αν υπάρχει η παράγωγος της $f(x)$ στο σημείο x_0 , τότε μπορούμε να ορίσουμε $f[x_0,x_0] = f'(x_0)$.

2.3.1 Τύπος παρεμβολής του Newton με διηρημένες διαφορές

Το πολυώνυμο παρεμβολής βαθμού το πολύ n που διέρχεται από τα $n+1$ σημεία (x_i,f_i) , $i = 0(1)n$ μπορεί να πάρει μία άλλη διαφορετική μορφή αν χρησιμοποιηθούν οι διηρημένες διαφορές. Αν $P_n(x)$ είναι το πολυώνυμο παρεμβολής του Lagrange τότε ισχύει

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$f(x) - P_n(x) = f(x) - \sum_{i=0}^{n} f_i \prod_{\substack{j=0 \\ j\neq i}}^{n} \left(\frac{x - x_j}{x_i - x_j}\right) =$$

$$= \prod_{i=0}^{n} (x - x_i)\left[\frac{f(x)}{\prod_{i=0}^{n} (x - x_i)} + \sum_{i=0}^{n} \frac{f_i}{(x_i - x) \prod_{\substack{j=0 \\ j\neq i}}^{n} (x_i - x_j)}\right] =$$

$$= \prod_{i=0}^{n} (x - x_i)\, f[x, x_0, x_1, \ldots, x_n] \,, \qquad (2.42)$$

όπου η έκφραση για την $f[x,x_0,x_1,\ldots,x_n]$ προκύπτει βάσει της (2.39) (βλέπε άσκηση 16). Από τους ορισμούς των διηρημένων διαφορών έχουμε

$$\left.\begin{aligned}
f(x) &= f_0 + (x - x_0)\, f[x,x_0] \\
f[x,x_0] &= f[x_0,x_1] + (x - x_1) f[x,x_0,x_1] \\
f[x,x_0,x_1] &= f[x_0,x_1,x_2] + (x - x_2) f[x,x_0,x_1,x_2] \\
&\;\;\vdots \\
f[x,x_0,x_1,\ldots,x_{n-1}] &= f[x_0,x_1,\ldots,x_n] + (x - x_n)\, f[x,x_0,x_1,\ldots,x_n].
\end{aligned}\right\} (2.43)$$

Πολλαπλασιάζοντας την δεύτερη από τις (2.43) επί $(x - x_0)$, την τρίτη επί $(x - x_0)(x - x_1)$ κ.ο.κ. και τέλος την τελευταία από τις (2.43) επί $\prod_{i=0}^{n-1} (x - x_i)$ και προσθέτοντας κατά μέλη παίρνουμε

$$f(x) = f_0 + (x - x_0)\, f[x_0,x_1] + (x - x_0)(x - x_1)\, f[x_0,x_1,x_2] + \ldots +$$

$$+ (x - x_0)(x - x_1)\ldots(x - x_{n-1}) f[x_0,x_1,\ldots,x_n] +$$

$$+ \prod_{i=0}^{n} (x - x_i)\, f[x,x_0,x_1,\ldots,x_n] \,. \qquad (2.44)$$

Από τις (2.42) και (2.44) συνεπάγεται ότι για το πολυώνυμο παρεμβολής $P_n(x)$ ισχύει

$$P_n(x) = f_0 + (x - x_0)\, f[x_0,x_1] + (x - x_0)(x - x_1)\, f[x_0,x_1,x_2] + \ldots +$$

$$+ (x - x_0)(x - x_1)\ldots(x - x_{n-1})\, f[x_0,x_1,\ldots,x_n] \,. \qquad (2.45)$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Το πολυώνυμο παρεμβολής στην μορφή (2.45) καλείται πολυώνυμο παρεμβολής του Newton με διηρημένες διαφορές. Η μορφή (2.45) σε σχέση με την (2.9) του Lagrange έχει το πλεονέκτημα που αναφέρθηκε και για την μέθοδο του Aitken. Συγκρίνοντας τις (2.24) και (2.42) προκύπτει ότι

$$f[x,x_0,x_1,\ldots,x_n] = \frac{f^{(n+1)}(\xi)}{(n+1)!}, \tag{2.46}$$

$$\xi \in [\min\{x_0,x_1,\ldots,x_n,x\},\ \max\{x_0,x_1,\ldots,x_n,x\}] .$$

Από την (2.46) έχουμε προφανώς

$$f[x_0,x_1,\ldots,x_n] = \frac{f^{(n)}(\xi)}{n!} \tag{2.47}$$

για κάποιο $\xi \in [a,b]$, όπου $a = \min_i\{x_i\}$, $b = \max_i\{x_i\}$, $i = 0(1)n$.
Η σχέση (2.47) συνδέει τις διηρημένες διαφορές με τις παραγώγους (υποτίθεται ότι η $f(x)$ είναι παραγωγίσιμη n φορές στο $[a,b]$).

Από την (2.47) συμπεραίνουμε ότι αν η $f(x)$ είναι ένα πολυώνυμο $Q_K(x)$ βαθμού $\kappa \leq n$ με συντελεστή μεγιστοβαθμίου όρου q_0 τότε

$$Q_K[x_0,x_1,x_2,\ldots,x_n] = \begin{cases} q_0 , & \text{αν } \kappa = n \\ 0 , & \text{αν } \kappa < n , \end{cases}$$

όπου x_i , $i = 0(1)n$ είναι αυθαίρετα διακεκριμένα σημεία.

Πρέπει να σημειωθεί ότι οι διηρημένες διαφορές στον τύπο παρεμβολής του Newton (2.45) βρίσκονται εύκολα από τον πίνακα διηρημένων διαφορών. Εάν θέσουμε $\alpha_0 = f_0$,

$$\alpha_1 = f[x_0,x_1],\ \alpha_2 = f[x_0,x_1,x_2],\ldots,\ \alpha_n = f[x_0,x_1,\ldots,x_n],$$

τότε ο υπολογισμός της τιμής του $P_n(x)$ για $x = \hat{x} \neq x_i$, $i = 0(1)n$ από την (2.45) μπορεί να γίνει με την βοήθεια της παρακάτω εκφράσεως

$$P_n(\hat{x}) = \Big(\ldots(\alpha_n(\hat{x} - x_{n-1}) + \alpha_{n-1})(\hat{x} - x_{n-2}) + \ldots\Big)(\hat{x} - x_0) + \alpha_0$$

η οποία είναι ανάλογη με το σχήμα του Horner (βλέπε κεφάλαιο 7).

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Παράδειγμα 2.4 Δίνεται ο πίνακας τιμών της $f(x)$:

x_i	-1	0	1	2
f_i	2	2	2	-1

. Να βρεθεί το πολυώνυμο παρεμβολής του Newton με διηρημένες διαφορές με σημεία παρεμβολής τα x_i , $i = 0(1)3$.

Λύση Το ζητούμενο πολυώνυμο παρεμβολής βαθμού το πολύ 3 θα είναι λόγω της (2.45)

$$P_3(x) = f_0 + (x - x_0)\, f[x_0, x_1] + (x - x_0)(x - x_1)\, f[x_0, x_1, x_2] +$$
$$+ (x - x_0)(x - x_1)(x - x_2)\, f[x_0, x_1, x_2, x_3] \ .$$

Κατασκευάζουμε τον ακόλουθο πίνακα διηρημένων διαφορών:

x_i	f_i	1ης τάξεως	2ας τάξεως	3ης τάξεως
$x_0 = -1$	2			
		0		
$x_1 = 0$	2		0	
		0		$-\frac{1}{2}$
$x_2 = 1$	2		$-\frac{3}{2}$	
		-3		
$x_3 = 2$	-1			

Έτσι έχουμε $P_3(x) = 2 + (x+1)\cdot 0 + (x+1)(x-0)\cdot 0 +$
$+ (x+1)(x-0)(x-1)(-\frac{1}{2}) = -\frac{1}{2}(x^3 - x - 4)$. ▲

2.4 Πεπερασμένες διαφορές—Πίνακες πεπερασμένων διαφορών

Ας υποθέσουμε ότι δίνονται σ'ένα πίνακα οι τιμές $f_i \equiv f(x_i)$, $i = 0,1,2,\ldots$ της συναρτήσεως $f(x)$, στα ισαπέχοντα σημεία $x_i = x_0 + ih$, όπου $h(>0)$ είναι ένας σταθερός αριθμός, που καλείται βήμα ή διάστημα πινακοποιήσεως. Από τις πινακοποιημένες τιμές της $f(x)$ μπορούμε να κατασκευάσουμε ένα πίνακα αριθμών, ο οποίος καλείται πίνακας των πεπερασμένων διαφορών, με τον ακόλουθο τρόπο: Οι δια-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

φορές πρώτης τάξεως (ο όρος πεπερασμένες θα παραλείπεται στη συνέχεια) βρίσκονται δι' αφαιρέσεως κάθε τιμής f_i από την αμέσως επόμενή της τιμή f_{i+1} στον πίνακα τιμών της $f(x)$. Οι διαφορές δευτέρας τάξεως βρίσκονται επαναλαμβάνοντας την παραπάνω διαδικασία στις διαφορές πρώτης τάξεως κ.ο.κ. για τις διαφορές ανωτέρας τάξεως. Ο τρόπος με τον οποίο γράφονται συνήθως οι διαφορές των διαφόρων τάξεων για να σχηματίσουν τον μοναδικό πίνακα των πεπερασμένων διαφορών γίνεται φανερός από το ακόλουθο παράδειγμα:

		Δ Ι Α Φ Ο Ρ Ε Σ			
x_i	f_i	1ης τάξεως	2ας τάξεως	3ης τάξεως	4ης τάξεως
− 2	34				
		− 29			
− 1	5		26		
		− 3		− 24	
0	2		2		24
		− 1		0	
1	1		2		
		1			
2	2				

Υπάρχουν τρείς διαφορετικοί συμβολισμοί που χρησιμοποιούνται συνήθως για να παραστήσουν τις διαφορές. Οι συμβολισμοί αυτοί χρησιμοποιούν αντίστοιχα τα σύμβολα Δ, ∇, δ που καλούνται τελεστές διαφορών και ικανοποιούν εξ ορισμού τις ακόλουθες σχέσεις:

$$\Delta f_i = f_{i+1} - f_i \ , \tag{2.48}$$

$$\nabla f_i = f_i - f_{i-1} \ , \tag{2.49}$$

$$\delta f_i = f_{i+\frac{1}{2}} - f_{i-\frac{1}{2}} \ . \tag{2.50}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Οι σχέσεις (2.48),(2.49) και (2.50) ορίζουν αντίστοιχα τις προς τα εμπρός, τις προς τα πίσω και τις κεντρικές διαφορές πρώτης τάξεως στο σημείο x_i . Οι προς τα εμπρός, οι προς τα πίσω και οι κεντρικές διαφορές ανωτέρας τάξεως στο σημείο x_i ορίζονται αντίστοιχα με την βοήθεια των παρακάτω αναδρομικών σχέσεων:

$$\Delta^m f_i = \Delta(\Delta^{m-1} f_i) = \Delta^{m-1} f_{i+1} - \Delta^{m-1} f_i \quad ,$$

$$\nabla^m f_i = \nabla(\nabla^{m-1} f_i) = \nabla^{m-1} f_i - \nabla^{m-1} f_{i-1} \quad , \qquad (2.51)$$

$$\delta^m f_i = \delta(\delta^{m-1} f_i) = \delta^{m-1} f_{i+\frac{1}{2}} - \delta^{m-1} f_{i-\frac{1}{2}} \quad ,$$

όπου m ακέραιος > 1 (σημειώνεται ότι $\Delta^0 f_i = \nabla^0 f_i = \delta^0 f_i = f_i$) .
Για παράδειγμα, συμφωνα με τις (2.48) – (2.51) έχουμε

$$\Delta^2 f_i = \Delta f_{i+1} - \Delta f_i = (f_{i+2} - f_{i+1}) - (f_{i+1} - f_i) = f_{i+2} - 2f_{i+1} + f_i ,$$

$$\nabla^2 f_i = \nabla f_i - \nabla f_{i-1} = f_i - 2f_{i-1} + f_{i-2} \quad ,$$

$$\delta^2 f_i = \delta f_{i+\frac{1}{2}} - \delta f_{i-\frac{1}{2}} = f_{i+1} - 2f_i + f_{i-1} \quad .$$

Είναι φανερό από τις (2.48) – (2.50) ότι η ποσότητα $f_{i+1} - f_i$ μπορεί να συμβολιστεί ως εξής:

$$f_{i+1} - f_i = \Delta f_i = \nabla f_{i+1} = \delta f_{i+\frac{1}{2}} \quad . \qquad (2.52)$$

Χρησιμοποιώντας και τις (2.51), είναι εύκολο να δειχθεί ότι ο μοναδικός πίνακας των πεπερασμένων διαφορών μπορεί να πάρει από την άποψη του συμβολισμού τις ακόλουθες τρείς μορφές:

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

i) Συμβολισμός που χρησιμοποιεί προς τα εμπρός διαφορές

x_i	f_i	1ης τάξεως	2ας τάξεως	3ης τάξεως	4ης τάξεως	...
x_0	f_0					
		Δf_0				
x_1	f_1		$\Delta^2 f_0$			
		Δf_1		$\Delta^3 f_0$		
x_2	f_2		$\Delta^2 f_1$		$\Delta^4 f_0$	
		Δf_2		$\Delta^3 f_1$		
x_3	f_3		$\Delta^2 f_2$			
		Δf_3				
x_4	f_4					
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Πίνακας 2.3

ii) Συμβολισμός με προς τα πίσω διαφορές

x_i	f_i	1ης τάξεως	2ας τάξεως	3ης τάξεως	4ης τάξεως	...
x_0	f_0					
		∇f_1				
x_1	f_1		$\nabla^2 f_2$			
		∇f_2		$\nabla^3 f_3$		
x_2	f_2		$\nabla^2 f_3$		$\nabla^4 f_4$	
		∇f_3		$\nabla^3 f_4$		
x_3	f_3		$\nabla^2 f_4$			
		∇f_4				
x_4	f_4					
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Πίνακας 2.4

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

iii) Συμβολισμός με κεντρικές διαφορές

x_i	f_i	1ης τάξεως	2ας τάξεως	3ης τάξεως	4ης τάξεως	...
x_0	f_0					
		$\delta f_{\frac{1}{2}}$				
x_1	f_1		$\delta^2 f_1$			
		$\delta f_{\frac{3}{2}}$		$\delta^3 f_{\frac{3}{2}}$		
x_2	f_2		$\delta^2 f_2$		$\delta^4 f_2$	
		$\delta f_{\frac{5}{2}}$		$\delta^3 f_{\frac{5}{2}}$		
x_3	f_3		$\delta^2 f_3$			
		$\delta f_{\frac{7}{2}}$				
x_4	f_4					
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Πίνακας 2.5

Παρατηρήσεις: i) Δεν είναι απαραίτητο να συμβολίζουμε με x_0 εκείνο από τα ισαπέχοντα σημεία x_i που βρίσκεται στην κορυφή του πίνακα.

ii) Στους παραπάνω πίνακες 2.3 – 2.5 σημειώνονται οι ευθείες κατά μήκος των οποίων οι διαφορές έχουν τους ίδιους δείκτες σε σχέση με τον αντίστοιχο συμβολισμό.

iii) Είναι προφανές ότι εάν $n+1$ είναι το πλήθος των ισαπεχόντων σημείων x_i, τότε ο πίνακας διαφορών θα περιέχει μέχρι και n–τάξεως διαφορές.

iv) Συνηθίζεται, χάριν συντομίας, όταν γράφουμε τις διαφορές στον πίνακα διαφορών να παραλείπεται η δεκαδική τελεία και να χρησιμοποιούνται μόνον τα σημαντικά ψηφία τους. Όταν όμως ζητούνται συγκεκριμένες διαφορές του πίνακα πρέπει αυτές να δίνονται με την κανονική τους μορφή.

Θεώρημα 2.3 Ισχύουν οι σχέσεις:

$$\Delta^m f_i = \nabla^m f_{i+m} = \delta^m f_{i+\frac{m}{2}}\ , \quad m = 0,1,2,\ldots \ . \tag{2.53}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Απόδειξη Μπορεί να γίνει εύκολα με την επαγωγή (βλέπε άσκηση 44).▲

Θεώρημα 2.4 Η m – τάξεως προς τα εμπρός διαφορά στο σημείο x_i εκφράζεται συναρτήσει των τιμών της συναρτήσεως $f(x)$ από τον τύπο:

$$\Delta^m f_i = \sum_{j=0}^{m} (-1)^j \binom{m}{j} f_{i+m-j} = \sum_{j=0}^{m} (-1)^{m-j} \binom{m}{j} f_{i+j} \quad . \qquad (2.54)$$

Απόδειξη Προφανώς για $m = 1$ η (2.54) ισχύει λόγω της (2.48). Υποθέτουμε ότι αυτή ισχύει για $m \leq \kappa$. Τότε έχουμε:

$$\Delta^{\kappa+1} f_i = \Delta(\Delta^\kappa f_i) =$$

$$= \Delta^\kappa f_{i+1} - \Delta^\kappa f_i = \sum_{j=0}^{\kappa} (-1)^j \binom{\kappa}{j} f_{i+1+\kappa-j} - \sum_{j=0}^{\kappa} (-1)^j \binom{\kappa}{j} f_{i+\kappa-j} =$$

$$= \binom{\kappa}{0}(f_{i+\kappa+1} - f_{i+\kappa}) - \binom{\kappa}{1}(f_{i+\kappa} - f_{i+\kappa-1}) + \binom{\kappa}{2}(f_{i+\kappa-1} - f_{i+\kappa-2}) - \dots +$$

$$+ (-1)^{\kappa-1}\binom{\kappa}{\kappa-1}(f_{i+2} - f_{i+1}) + (-1)^\kappa \binom{\kappa}{\kappa}(f_{i+1} - f_i) =$$

$$= \binom{\kappa+1}{0} f_{i+\kappa+1} - \left[\binom{\kappa}{0} + \binom{\kappa}{1}\right] f_{i+\kappa} + \left[\binom{\kappa}{1} + \binom{\kappa}{2}\right] f_{i+\kappa-1} - \dots +$$

$$+ (-1)^\kappa \left[\binom{\kappa}{\kappa-1} + \binom{\kappa}{\kappa}\right] f_{i+1} + (-1)^{\kappa+1}\binom{\kappa+1}{\kappa+1} f_i = \sum_{j=0}^{\kappa+1} (-1)^j \binom{\kappa+1}{j} f_{i+\kappa+1-j} ,$$

αφού ισχύει $\binom{\kappa}{j} + \binom{\kappa}{j-1} = \binom{\kappa+1}{j}$. Άρα η (2.54) ισχύει για κάθε $m = 0,1,2,\dots$. ▲

Από την (2.54) είναι φανερό ότι για τον υπολογισμό της $\Delta^m f_i$ πρέπει να είναι γνωστές οι $m+1$ τιμές $f_i, f_{i+1}, \dots, f_{i+m}$.

Θεώρημα 2.5 Αν $x_i = x_0 + ih$, τότε ισχύει

$$f[x_i, x_{i+1}, \dots, x_{i+m}] = \frac{\Delta^m f_i}{h^m m!} \quad . \qquad (2.55)$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

<u>Απόδειξη</u> Η απόδειξη γίνεται επαγωγικά. Για $m = 1$ είναι

$f[x_i, x_{i+1}] = \dfrac{f_{i+1} - f_i}{x_{i+1} - x_i} = \dfrac{\Delta f_i}{h}$. Υποθέτοντας ότι η (2.55) ισχύει για κάθε $m \leq \kappa$ παίρνουμε:

$$f[x_i x_{i+1}, \ldots, x_{i+\kappa+1}] =$$

$$= \frac{f[x_{i+1}, x_{i+2}, \ldots, x_{i+\kappa+1}] - f[x_i, x_{i+1}, \ldots, x_{i+\kappa}]}{x_{i+\kappa+1} - x_i} =$$

$$= \frac{\Delta^\kappa f_{i+1} - \Delta^\kappa f_i}{(h^\kappa \kappa!)(\kappa+1)h} = \frac{\Delta^{\kappa+1} f_i}{h^{\kappa+1}(\kappa+1)!} .$$

Δηλαδή, η (2.55) ισχύει και για $m = \kappa+1$ και επομένως το θεώρημα απεδείχθη. ▲

Εχοντας υπ' όψιν την (2.47) και τις (2.53), (2.55) συμπεραίνουμε ότι

$$\Delta^m f_i = \nabla^m f_{i+m} = \delta^m f_{i+\frac{m}{2}} = h^m f^{(m)}(\xi) \quad , \qquad (2.56)$$

όπου $x_i \leq \xi \leq x_{i+m}$.

<u>Πόρισμα 2.1</u> Οι πεπερασμένες διαφορές τάξεως n ενός πολυωνύμου βαθμού n είναι σταθερές ενώ οι διαφορές τάξεως $> n$ είναι ίσες με μηδέν. Συγκεκριμένα, οι διαφορές n τάξεως είναι όλες ίσες με $\alpha_0 n! h^n$, όπου α_0 είναι ο συντελεστής του μεγιστοβαθμίου όρου του πολυωνύμου και h το βήμα πινακοποιήσεως.

<u>Απόδειξη</u> Εάν $P_n(x) = \alpha_0 x^n + \alpha_1 x^{n-1} + \ldots + \alpha_{n-1} x + \alpha_n$ είναι ένα πολυώνυμο n βαθμού ($\alpha_0 \neq 0$) , τότε για το τυχόν σημείο $x \equiv x_i$ του πίνακα, από την (2.56) παίρνουμε

$$\Delta^n P_n(x) = h^n P_n^{(n)}(\xi) = h^n \alpha_0 n! = \text{σταθερά}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

και επομένως $\Delta^{\kappa}P_n(x) = 0$ για $\kappa > n$. ▲

Σημείωση Αληθεύει επίσης και το αντίστροφο του πορίσματος 2.1. Δηλαδή, εάν οι n – τάξεως διαφορές μιας συναρτήσεως $f(x)$ (που ορίζεται από ένα πίνακα τιμών) ισούνται όλες με την ίδια σταθερά ($\neq 0$), τότε η συνάρτηση $f(x)$ είναι ένα πολυώνυμο ακριβώς n βαθμού. Πρέπει να τονιστεί ακόμη ότι το πόρισμα 2.1 ισχύει μόνον όταν οι πινακοποιημένες τιμές του πολυωνύμου είναι ακριβείς.

Παράδειγμα 2.5 Να κατασκευασθεί ο πίνακας διαφορών για την συνάρτηση $f(x) = x^3 - 2x^2 + 3x - 4$ όταν δίνονται οι τιμές της για $x = 0(1)5$.

Λύση

x_i	f_i	Δ	Δ^2	Δ^3	Δ^4	Δ^5
0	−4					
		2				
1	−2		2			
		4		6		
2	2		8		0	
		12		6		0
3	14		14		0	
		26		6		
4	40		20			
		46				
5	86					

Το παράδειγμα αυτό επαληθεύει την ισχύ του πορίσματος 2.1. Πράγματι είναι

$$\Delta^3 f(x) \doteq 1^3 \cdot \alpha_0 \cdot 3! = 6\alpha_0 = 6 \ . \qquad ▲$$

Παράδειγμα 2.6 Να επεκταθεί ο πίνακας τιμών του πολυωνύμου του

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

προηγουμένου παραδείγματος για $x = 6,7,8$ χωρίς να χρησιμοποιηθεί η $f(x)$.

<u>Λύση</u> Από τον πίνακα διαφορών στο παράδειγμα 2.5 φαίνεται ότι όλες οι διαφορές τρίτης τάξεως θα είναι ίσες με 6 . Έτσι, προσθέτοντας τρείς ακόμη τιμές ίσες με 6 στην στήλη των τρίτων διαφορών και προχωρώντας προς τα αριστερά με προσθέσεις, όπως δείχνουν τα βέλη στον παρακάτω πίνακα διαφορών βρίσκουμε διαδοχικά τις τιμές $f(6)$, $f(7)$ και $f(8)$.

x_i	f_i	Δ	Δ^2	Δ^3	Δ^4	Δ^5
0	− 4					
		2				
1	− 2		2			
		4		6		
2	2		8		0	
		12		6		0
3	14		14		0	
		16		6		0
4	40		20		0	
		46		6		0
5	86		26		0	
		72		6		0
6	158		32		0	
		104		6		
7	262		38			
		142				
8	404					

▲

2.4.1 Διάδοση σφαλμάτων σε πίνακες διαφορών

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Θα θεωρήσουμε τις ακόλουθες δύο περιπτώσεις:

i) Υποθέτουμε ότι αντί της ακριβούς τιμής f_i στον πίνακα τιμών της $f(x)$ έχουμε μία προσεγγιστική της τιμή $\bar{f}_i = f_i + \varepsilon$, όπου ε είναι το σφάλμα της $\bar{f}_i$, ενώ όλες οι άλλες τιμές της $f(x)$ στον πίνακα είναι ακριβείς. Θα εξετάσουμε το αποτέλεσμα του απομονωμένου σφάλματος ε στις διαφορές των διαφόρων τάξεων. Όπως φαίνεται από τον παρακάτω πίνακα διαφορών

f	Δ	Δ^2	Δ^3	Δ^4	Δ^5	...
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$		
f_{i-3}						
	Δf_{i-3}					
f_{i-2}		$\Delta^2 f_{i-3}$		$\Delta^4 f_{i-4} + \varepsilon$		
	Δf_{i-2}		$\Delta^3 f_{i-3} + \varepsilon$			
f_{i-1}		$\Delta^2 f_{i-2} + \varepsilon$		$\Delta^4 f_{i-3} - 4\varepsilon$		
	$\Delta f_{i-1} + \varepsilon$		$\Delta^3 f_{i-2} - 3\varepsilon$			
$f_i + \varepsilon$		$\Delta^2 f_{i-1} - 2\varepsilon$		$\Delta^4 f_{i-2} + 6\varepsilon$	$\vdots$	$\vdots$
	$\Delta f_i - \varepsilon$		$\Delta^3 f_{i-1} + 3\varepsilon$			
f_{i+1}		$\Delta^2 f_i + \varepsilon$		$\Delta^4 f_{i-1} - 4\varepsilon$		
	Δf_{i+1}		$\Delta^3 f_i - \varepsilon$			
f_{i+2}		$\Delta^2 f_{i+1}$		$\Delta^4 f_i + \varepsilon$		
	Δf_{i+2}		$\vdots$			
f_{i+3}		$\vdots$		$\vdots$		
$\vdots$	$\vdots$					

Πίνακας 2.6

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

η διάδοση του σφάλματος ε στις διαφορές των διαφόρων τάξεων ακολουθεί μία τριγωνική διάταξη. Ακόμη, παρατηρούμε ότι τα σφάλματα που εμφανίζονται στις διαφορές m – τάξεως έχουν αντίστοιχα τις τιμές:

$$\binom{m}{0}\varepsilon,\ -\binom{m}{1}\varepsilon,\ \binom{m}{2}\varepsilon,\ldots,\ (-1)^m\binom{m}{m}\varepsilon\ ,$$

όπου γίνεται φανερή η παρουσία των διωνυμικών συντελεστών με εναλλασσόμενα πρόσημα (λόγω του θεωρήματος 2.4). Σημειώνεται ότι η τιμή της συναρτήσεως f(x) η οποία έχει το σφάλμα ε βρίσκεται στην ίδια οριζόντια γραμμή με την διαφορά m – τάξεως που έχει το μεγαλύτερο απόλυτο σφάλμα, αν αυτή είναι μία (m = άρτιος),ή βρίσκεται στην οριζόντια γραμμή που διέρχεται ανάμεσα από τις διαφορές m – τάξεως που έχουν το μέγιστο απόλυτο σφάλμα, αν αυτές είναι δύο (m = περιττός). Έτσι είναι δυνατόν σε ορισμένες περιπτώσεις να ανακαλύπτουμε και να προσδιορίζουμε το απομονωμένο σφάλμα ε της τιμής $\bar{f}_i$ και στη συνέχεια να βρίσκουμε την f_i .

Με ανάλογο τρόπο μπορεί να εξεταστεί και η περίπτωση όπου περισσότερες από μία τιμές της f(x) έχουν σφάλματα.

<u>Παράδειγμα 2.7</u> Να βρεθεί και να διορθωθεί το απομονωμένο σφάλμα που υπάρχει σε μία από τις τιμές της f(x) στον ακόλουθο πίνακα, αν είναι γνωστό ότι η f(x) είναι πολυώνυμο τρίτου βαθμού

x	5.3	5.4	5.5	5.6	5.7	5.8
f(x)	148.877	157.464	166.357	175.616	185.193	195.112

x	5.9	6.0
f(x)	205.379	216.000

.

<u>Λύση</u> Κατασκευάζουμε τον πίνακα διαφορών (βλέπε επόμενη σελίδα). Σύμφωνα με το πόρισμα 2.1, εάν δεν υπήρχε το απομονωμένο σφάλμα ε σε κάποια τιμή της f(x) οι διαφορές τετάρτης τάξεως θα ήταν όλες ίσες με 0 . Αφού η ακολουθία των σφαλμάτων στις διαφορές 4ης τάξεως είναι {ε, –4ε, 6ε, 4ε, ε} , είναι φανερό από τον πίνακα διαφορών ότι θα έ-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

χουμε $6\varepsilon = -108$, $-4\varepsilon = 72$, $\varepsilon = -18$ και επομένως $\varepsilon = -0.018$. Επιπλέον, η τιμή που περιέχει το σφάλμα είναι η $\bar{f}(5.5) = 166.357$, οπότε η ακριβής τιμή είναι $f(5.5) = 166.357 - \varepsilon = 166.375$.

x	f(x)	Δ	Δ^2	Δ^3	Δ^4	
5.3	148.877		.		.	ε
		8587		.		
5.4	157.464		306		.	−4ε
		8893		60		
5.5	166.357		366		−108 →	6ε
		9259		− 48		
5.6	175.616		318		72 →	−4ε
		9577		24		
5.7	185.193		342		−18 →	ε
		9919		6		
5.8	195.112		348		0	
		10267		6		
5.9	205.379		354			
		10621				
6.0	216.000				.	

▲

ii) Υποθέτουμε τώρα ότι όλες οι πινακοποιημένες τιμές της f(x) είναι στρογγυλευμένες σε κ δεκαδικές θέσεις. Αν δηλαδή ε_i είναι το σφάλμα στρογγυλεύσεως στην τιμή $\bar{f}_i$ $(f_i \approx \bar{f}_i)$, τότε θα ισχύει προφανώς $|\varepsilon_i| \leq \frac{1}{2} 10^{-K}$. Επομένως το μέγιστο απόλυτο σφάλμα (στρογγυλεύσεως) στις τιμές της f(x) είναι $0.5 \cdot 10^{-K}$. Θα εξετάσουμε τώρα ποιό είναι το μέγιστο απόλυτο σφάλμα στις διαφορές m τάξεως $(m \geq 1)$ που οφείλεται στην στρογγύλευση των τιμών της f(x) . Σύμφωνα με την (2.54) και το κεφάλαιο 1 η τιμή $\Delta^m \bar{f}_i$ θα περιέχει ένα

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

σφάλμα $\varepsilon_i^{(m)}$ τέτοιο ώστε

$$\varepsilon_i^{(m)} = \sum_{j=0}^{m} (-1)^j \binom{m}{j} \varepsilon_{i+m-j} \quad . \tag{2.57}$$

Από την (2.57) συμπεραίνουμε ότι

$$|\varepsilon_i^{(m)}| \leq \sum_{j=0}^{m} \binom{m}{j} |\varepsilon_{i+m-j}| \leq 0.5 \cdot 10^{-K} \left(\sum_{j=0}^{m} \binom{m}{j} \right) =$$

$$= 0.5 \cdot 10^{-K} \cdot 2^m = 2^{m-1} \cdot 10^{-K} \ .$$

Άρα το μέγιστο απόλυτο σφάλμα στις διαφορές m – τάξεως είναι ίσο με $2^{m-1} 10^{-K}$.

<u>Παρατήρηση</u> Ο πίνακας διαφορών που κατασκευάζεται από ένα σύνολο τιμών της συναρτήσεως $f(x)$ (που αντιστοιχούν σ' ένα σύνολο ισαπεχόντων τιμών του x) μπορεί να χρησιμοποιηθεί για να αποφανθούμε εάν η $f(x)$ μπορεί να προσεγγιστεί καλά από ένα πολυώνυμο και ποιός είναι ο κατάλληλος βαθμός για το προσεγγιστικό πολυώνυμο. Για τον σκοπό αυτό εξετάζουμε αν οι διαφορές κάποιας τάξεως είναι μικρές (κατά προσέγγιση ίσες με μηδέν), λαμβανομένων υπ' όψιν και των ορίων του σφάλματος στρογγυλεύσεως (εάν υπάρχει) που αντιστοιχεί στις διαφορές αυτές. Εάν αυτό συμβαίνει για τις διαφορές $m+1$ τάξεως, τότε ο κατάλληλος βαθμός για το προσεγγιστικό πολυώνυμο είναι m .

2.5 Τ ε λ ε σ τ έ ς δ ι α φ ο ρ ώ ν

Είναι χρήσιμο για πολλούς λόγους, όπως θα φανεί και από τα επόμενα κεφάλαια, να θεωρήσουμε τα σύμβολα Δ , ∇ και δ , που ορίστηκαν στην παράγραφο 2.4 για να συμβολίσουμε τις πεπερασμένες διαφορές, σαν τελεστές οι οποίοι μετασχηματίζουν μία δοσμένη συνάρτηση $f(x)$ σύμφωνα με τις ακόλουθες σχέσεις:

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\Delta f(x) = f(x+h) - f(x) \ ,$$
$$\nabla f(x) = f(x) - f(x-h) \ , \qquad (2.58)$$
$$\delta f(x) = f(x+\frac{h}{2}) - f(x-\frac{h}{2}) \ ,$$

όπου h είναι το σταθερό βήμα (ή διάστημα) μεταβολής του x. Γενικά, καλείται τελεστής θ η παράσταση μιας απεικονίσεως ή ενός μετασχηματισμού από ένα σύνολο S_1 σ' ένα σύνολο S_2 ($\theta: S_1 \rightarrow S_2$) έτσι ώστε σε κάθε στοιχείο του S_1 αντιστοιχεί ένα μοναδικό στοιχείο του S_2 ($f(x) \in S_1 \xrightarrow{\theta} g(x) \in S_2$ ή $\theta f(x) = g(x)$). Οι τελεστές Δ, ∇ και δ καλούνται αντίστοιχα τελεστής των προς τα εμπρός, των προς τα πίσω και των κεντρικών διαφορών και είναι εύκολο να δειχθεί ότι είναι γραμμικοί τελεστές, με την έννοια ότι, αν θ είναι ένας οποιοσδήποτε από τους Δ, ∇ και δ και α,β αυθαίρετες σταθερές, τότε ισχύει

$$\theta(\alpha f(x) + \beta g(x)) = \alpha\theta f(x) + \beta\theta g(x) \ , \qquad (2.59)$$

για όλες τις συναρτήσεις $f(x)$, $g(x)$ που ανήκουν στο πεδίο ορισμού.

Είναι σκόπιμο να ορίσουμε και τους ακόλουθους τελεστές οι οποίοι συνδέονται με τους τελεστές διαφορών. Ο ταυτοτικός ή μοναδιαίος τελεστής I ή 1 ορίζεται από την

$$If(x) = f(x) \quad \text{για κάθε} \quad f(x) \ , \qquad (2.60)$$

ο τελεστής μετατοπίσεως E από την

$$Ef(x) = f(x+h) \ , \qquad (2.61)$$

οι τελεστές D και J της παραγώγου και του ορισμένου ολοκληρώματος αντίστοιχα, σύμφωνα με τις σχέσεις

$$Df(x) = f'(x) \ ,$$
$$Jf(x) = \int_x^{x+h} f(t)dt \ , \qquad (2.62)$$

και τέλος ο τελεστής μέσης τιμής μ είναι τέτοιος ώστε

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\mu f(x) = \frac{1}{2}\left[f(x + \frac{h}{2}) + f(x - \frac{h}{2})\right] . \qquad (2.63)$$

Οι τελεστές E, D, J, μ είναι επίσης γραμμικοί.

Δύο τελεστές θ_1 και θ_2 θα λέμε ότι είναι ίσοι εάν $\theta_1 f(x) = \theta_2 f(x)$ για κάθε συνάρτηση $f(x)$ που ανήκει στο κοινό πεδίο ορισμού των. Το άθροισμα και το γινόμενο τελεστών ορίζονται ως εξής:

$(\theta_1 \pm \theta_2)f(x) = \theta_1 f(x) \pm \theta_2 f(x)$, για κάθε $f(x)$ που ανήκει στο κοινό πεδίο ορισμού,

$(\theta_1\theta_2)f(x) = \theta_1(\theta_2 f(x))$ (γενικά, $\theta_1\theta_2 \neq \theta_2\theta_1$)

$(\alpha\theta)f(x) = \alpha(\theta f(x))$, α σταθερά .

Αν θ είναι ένας οποιοσδήποτε από τους μέχρι τώρα ορισθέντες τελεστές, οι δυνάμεις του θ ορίζονται από την $\theta^n = \theta \cdot \theta^{n-1}$, n ακέραιος >1 και ισχύει η ιδιότητα $\theta^m(\theta^n f(x)) = \theta^{m+n} f(x)$, όπου m , n είναι μη αρνητικοί ακέραιοι. Εξ ορισμού είναι $\theta^0 = I$. Για τον τελεστή E η δύναμη E^m ορίζεται για κάθε πραγματικό αριθμό m έτσι ώστε

$$E^m f(x) = f(x + mh) , \qquad (2.64)$$

με την προυπόθεση ότι ορίζεται η $f(x + mh)$. Ειδικά έχουμε

$$E^{-1} f(x) = f(x - h) . \qquad (2.65)$$

Μπορεί εύκολα να δειχθεί ότι για όλους τους τελεστές που ορίστηκαν ισχύει η αντιμεταθετική, η επιμεριστική και η προσεταιριστική ιδιότητα των πραγματικών αριθμών.

Εάν θ είναι ένας τελεστής, $\theta: X \to Y$, ένα προς ένα και επί, τότε υπάρχει ένας τελεστής που συμβολίζεται θ^{-1} και καλείται ο αντίστροφος του τελεστή θ . Είναι $\theta^{-1}: Y \to X$ και ισχύει $\theta^{-1}\theta = \theta\,\theta^{-1} = I$.

Ο τελεστής θ είναι ένα προς ένα εάν

$$\theta f_1(x) = \theta f_2(x) \Rightarrow f_1(x) = f_2(x)$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ή ισοδύναμα

$$f_1(x) \neq f_2(x) \Rightarrow \theta f_1(x) \neq \theta f_2(x) \ . \qquad (2.66)$$

Θα δειχθεί τώρα ότι μόνον ο E έχει αντίστροφο E^{-1} (για τον οποίο ισχύει $E^{-1}f(x) = f(x-h)$ (βλέπε (2.65)) και $E^{-1}E = E\,E^{-1} = I$) . Πράγματι, έστω θ ένας τελεστής και $\theta f(x) = g(x)$. Αν $\varphi(x)$ είναι μία συνάρτηση $\neq 0$ τέτοια ώστε $\theta\varphi(x) \equiv 0$ θα έχουμε $\theta(f(x)+\varphi(x)) = \theta f(x) + \theta\varphi(x) = \theta f(x) = g(x)$ με $f(x)+\varphi(x) \neq f(x)$. Άρα δεν ισχύει η (2.66). Για τον E δεν υπάρχει $\varphi(x) \neq 0$ τέτοια ώστε $E\varphi(x) = \varphi(x+h) \equiv 0$ δηλαδή, η $E\varphi(x) \equiv 0$ ισχύει εάν και μόνον εάν $\varphi(x) \equiv 0$, ενώ όλοι οι υπόλοιποι τελεστές μηδενίζουν ορισμένες συναρτήσεις (π.χ. οι Δ , ∇ , δ μηδενίζουν κάθε συνάρτηση περιόδου h , ο J μηδενίζει την παράγωγο κάθε τέτοιας συναρτήσεως και ο D κάθε σταθερά. Για τους Δ , ∇ , δ , D μπορεί να ληφθεί $\varphi(x) \equiv c =$ σταθερά). Ας υποθέσουμε τώρα ότι $f_1(x) \neq f_2(x)$ και $Ef_1(x) = Ef_2(x)$. Μπορούμε να γράψουμε $f_2(x) = f_1(x) + \psi(x)$, $\psi(x) \neq 0$ οπότε παίρνουμε $Ef_1(x) = Ef_2(x) = E(f_1(x)+\psi(x)) = Ef_1(x) + E\psi(x)$ ή $E\psi(x) \equiv 0$ ή $\psi(x) \equiv 0$, πράγμα άτοπον. Άρα για τον E ισχύει η (2.66).

Για τους άλλους τελεστές (εκτός του E) το σύμβολο θ^{-1} δεν έχει έννοια τελεστή και ορίζεται έτσι ώστε

$$\theta^{-1}g(x) = f(x) \ , \qquad (2.67)$$

όπου $f(x)$ είναι μία οποιαδήποτε συνάρτηση που ικανοποιεί την

$$\theta f(x) = g(x) \ . \qquad (2.68)$$

Για παράδειγμα, $\Delta^{-1}g(x) =$ το σύνολο των συναρτήσεων $f(x)$ τέτοιων ώστε $\Delta f(x) = g(x)$. Από την (2.67) έχουμε $\theta\,\theta^{-1}g(x) = \theta f(x) = g(x)$, δηλαδή,

$$\theta\,\theta^{-1} = I \ . \qquad (2.69)$$

Έτσι, ο θ^{-1} θα καλείται δεξιός αντίστροφος του θ . Αν $\varphi(x)$ είναι συνάρτηση τέτοια ώστε $\theta\varphi(x) \equiv 0$, $\varphi(x) \not\equiv 0$ και ισχύει η (2.68) τότε έχουμε $\theta(f(x)+\varphi(x)) = g(x)$. Άρα, θα έχουμε επίσης

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\theta^{-1}g(x) = f(x) + \varphi(x) \ . \tag{2.70}$$

Η (2.70) λόγω της (2.68) γίνεται

$$\theta^{-1}\theta f(x) = f(x) + \varphi(x) \tag{2.71}$$

(για $\theta = \Delta, \nabla, \delta, D$ μπορεί να ληφθεί $\varphi(x) \equiv c =$ σταθερά). Αντιστρόφως, αν ισχύει η (2.71) τότε λόγω της (2.69) έχουμε $\theta(\theta^{-1}\theta f(x)) = \theta(f(x) + \varphi(x))$ ή $\theta f(x) = \theta f(x) + \theta\varphi(x)$, δηλαδή, $\theta\varphi(x) \equiv 0$.

Από τα παραπάνω είναι φανερό ότι πρέπει να δίνεται ιδιαίτερη προσοχή σε πράξεις που περιέχουν το σύμβολο θ^{-1}.

Από τους ορισμούς των τελεστών που δόθηκαν προκύπτουν εύκολα οι ακόλουθες σχέσεις μεταξύ αυτών

$$\begin{aligned} &\Delta = E - 1 \ , \quad \nabla = 1 - E^{-1} \ , \quad \delta = E^{\frac{1}{2}} - E^{-\frac{1}{2}} \ , \\ &\mu = \frac{1}{2}\left(E^{\frac{1}{2}} + E^{-\frac{1}{2}}\right) \ , \quad DJ = JD = \Delta \ . \end{aligned} \tag{2.72}$$

Ακόμη, χρησιμοποιώντας τον τύπο του Taylor (υποθέτοντας ότι $f(x)$ είναι μία αναλυτική συνάρτηση)

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2!} f''(x) + \ldots$$

και τους ορισμούς των τελεστών E και D παίρνουμε

$$Ef(x) = \left(1 + hD + \frac{h^2D^2}{2!} + \ldots\right)f(x) = e^{hD}f(x) \ .$$

Επομένως, συμβολικά έχουμε

$$E = e^{hD} \ , \tag{2.73}$$

όπου ο τελεστής e^{hD} ορίζεται από το ανάπτυγμά του σε δυναμοσειρά. Στο σημείο αυτό πρέπει να σημειωθεί ότι κάθε ένας από τους τελεστές Δ, ∇, δ και D όταν εφαρμόζεται σε ένα πολυώνυμο ελαττώνει τον βαθμό του, ενώ δεν ισχύει το ίδιο για τους E, J και μ . Έτσι, από την (2.73) προκύπτει ότι ο τελεστής E είναι ισοδύναμος με το άθροισμα των $n+1$ πρώτων όρων του αναπτύγματος σε σειρά του δεύτερου μέλους της

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

(2.73) όταν εφαρμόζεται σε ένα πολυώνυμο n βαθμού.

Γενικά, αν θ είναι ένας τελεστής, τότε με την σειρά

$$\alpha_0 I + \alpha_1\theta + \alpha_2\theta^2 + \ldots + \alpha_n\theta^n + \ldots$$

εννοούμε ένα τελεστή θ_1 τέτοιο ώστε

$$\theta_1 f(x) = \lim_{n\to\infty} \sum_{i=0}^{n} \alpha_i\theta^i f(x) .$$

Είναι φανερό ότι όλες οι σειρές των τελεστών Δ, ∇, δ και D όταν εφαρμόζονται σε πολυώνυμο n βαθμού ισοδυναμούν με πεπερασμένα αθροίσματα. Στην πράξη, όταν οι τύποι των τελεστών εφαρμόζονται σε συναρτήσεις που δεν είναι πολυώνυμα, τα αναπτύγματα των σειρών πρέπει να αποκόπτονται ύστερα από ένα ορισμένο αριθμό όρων και να προσδιορίζεται το σφάλμα αποκοπής.

Από τις (2.72), (2.73) μπορεί ακόμη να δειχθεί εύκολα ότι

$$D = \frac{1}{h}\ln(1+\Delta) = \frac{1}{h}\left(\Delta - \frac{\Delta^2}{2} + \frac{\Delta^3}{3} - \frac{\Delta^4}{4} + \ldots + (-1)^{n+1}\frac{\Delta^n}{n} + \ldots\right) \tag{2.74}$$

και

$$E = (1-\nabla)^{-1} = 1 + \nabla + \nabla^2 + \nabla^3 + \ldots \quad . \tag{2.75}$$

Οι μέθοδοι της Αριθμητικής Αναλύσεως που χρησιμοποιούν την θεωρία των τελεστών καλούνται συμβολικές μέθοδοι και είναι ιδιαίτερα χρήσιμες για την παραγωγή τύπων αριθμητικής παραγωγίσεως και ολοκληρώσεως.

<u>Παράδειγμα 2.8</u> Να υπολογισθεί το άθροισμα $S = 1^2 + 2^2 + 3^2 + \ldots + \kappa^2$ χρησιμοποιώντας συμβολικές μεθόδους.

<u>Λύση</u> Θεωρούμε το πολυώνυμο $P(x) = x^2$ και $h = 1$. Τότε είναι

$$S = E^0P_1(x) + EP_1(x) + E^2P_1(x) + \ldots + E^{\kappa-1}P_1(x) ,$$

όπου $E^mP_1(x) \equiv \left[E^mP(x)\right]_{x=1}$, $m = 0(1)\kappa-1$. Άρα ,

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$S = (1 + E + E^2 + \ldots + E^{K-1})P_1(x) = \left(\frac{E^K - 1}{E - 1}\right)P_1(x) = \left[\frac{(1+\Delta)^K - 1}{\Delta}\right]P_1(x)$$

$$= \left[\frac{\binom{K}{1}\Delta + \binom{K}{2}\Delta^2 + \ldots + \binom{K}{K}\Delta^K}{\Delta}\right]P_1(x) = \left[K + \frac{K(K-1)}{2}\Delta + \right.$$

$$\left. + \frac{K(K-1)(K-2)}{6}\Delta^2\right]P_1(x) = KP_1(x) + \frac{K(K-1)}{2}\Delta P_1(x) +$$

$$+ \frac{K(K-1)(K-2)}{6}\Delta^2 P_1(x)$$. Αλλά είναι $P_1(x) = \left[P(x)\right]_{x=1} = 1$,

$\Delta P_1(x) = \left[\Delta P(x)\right]_{x=1} = \left[(x+1)^2 - x^2\right]_{x=1} = \left[2x+1\right]_{x=1} = 3$ και

$\Delta^2 P_1(x) = \left[2(x+1) + 1 - 2x - 1\right]_{x=1} = 2$. Επομένως, έχουμε τελικά

$$S = K + \frac{3}{2}K(K-1) + \frac{K(K-1)(K-2)}{6}\,2 = \frac{1}{6}K(K+1)(2K+1) \;. \quad \blacktriangle$$

2.5.1 Παραγοντικές δυνάμεις

Αν x είναι ένας αριθμός και h είναι ένα ορισμένο σταθερό βήμα, τότε η K-στή παραγοντική δύναμη του x που συμβολίζεται με $x^{[K]}$ ορίζεται ως εξής:

$$x^{[K]} = x(x-h)(x-2h)\ \ldots\ [x-(K-1)h] \;, \qquad (2.76)$$

όπου K είναι μη αρνητικός ακέραιος. Εξ ορισμού έχουμε $x^{[0]} = 1$. Αν $h = 0$ τότε προφανώς $x^{[K]} = x^K$. Αν $x = n =$ θετικός ακέραιος και $h = 1$ τότε $n^{[n]} = n!$.

<u>Θεώρημα 2.6</u> Ισχύει

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\Delta^m x^{[\kappa]} = \begin{cases} \kappa(\kappa-1)(\kappa-2)\ \ldots\ [\kappa-(m-1)]h^m x^{[\kappa-m]} & , \text{για } m = 1(1)\kappa \\ 0 \quad , \quad \text{για } m > \kappa \ . & \end{cases} \tag{2.77}$$

<u>Απόδειξη</u> Από την (2.76) είναι φανερό ότι το $x^{[\kappa]}$ είναι ένα πολυώνυμο βαθμού κ και επομένως σύμφωνα με το πόρισμα 2.1 είναι $\Delta^m x^{[\kappa]} = 0$ για $m > \kappa$. Εφαρμόζοντας τον τελεστή Δ στην (2.76) παίρνουμε

$$\Delta x^{[\kappa]} = (x+h)^{[\kappa]} - x^{[\kappa]} =$$

$$= (x+h)x(x-h)\ldots[x-(\kappa-3)h][x-(\kappa-2)h] - x(x-h)(x-2h)\ldots$$

$$\ldots[x-(\kappa-1)h] = \Big((x+h) - [x-(\kappa-1)h]\Big)x(x-h)\ldots$$

$$\ldots[x-(\kappa-2)h] = \kappa h x^{[\kappa-1]} \ , \quad \text{δηλαδή} \ ,$$

$$\Delta x^{[\kappa]} = \kappa h x^{[\kappa-1]} \ . \tag{2.78}$$

Επαγωγικά τώρα μπορεί εύκολα να δειχθεί ο γενικός τύπος (2.77). ▲

<u>Παρατήρηση</u> Από την (2.78) θέτοντας $h = \Delta x$ έχουμε

$$\frac{\Delta x^{[\kappa]}}{\Delta x} = \kappa x^{[\kappa-1]} \ . \tag{2.79}$$

Η (2.79) είναι το διακεκριμένο ανάλογο του τύπου της παραγωγίσεως $\frac{dx^\kappa}{dx} = \kappa x^{\kappa-1}$.

Μπορεί να δειχθεί ότι η (2.76) γράφεται επίσης (για $\kappa \geqq 1$)

$$x^{[\kappa]} = \sum_{j=1}^{\kappa} \alpha_j^{(\kappa)} h^{\kappa-j} x^j \ ,$$

όπου οι αριθμοί $\alpha_j^{(\kappa)}$ καλούνται αριθμοί του Stirling πρώτου είδους και βρίσκονται χρησιμοποιώντας την αναδρομική σχέση

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\alpha_j^{(\kappa+1)} = \alpha_{j-1}^{(\kappa)} - \kappa\alpha_j^{(\kappa)}, \quad \text{με} \quad \alpha_\kappa^{(\kappa)} = 1 \quad \text{και} \quad \alpha_0^{(\kappa)} = 0 .$$

Για παράδειγμα, είναι $x^{[1]} = \alpha_1^{(1)}x = x$,

$$x^{[2]} = \alpha_1^{(2)}hx + \alpha_2^{(2)}x^2 = \left[\alpha_0^{(1)} - \alpha_1^{(1)}\right]hx + x^2 = -hx + x^2 ,$$

$$x^{[3]} = \alpha_1^{(3)}h^2x + \alpha_2^{(3)}hx^2 + \alpha_3^{(3)}x^3 = \left[\alpha_0^{(2)} - 2\alpha_1^{(2)}\right]h^2x +$$

$$+ \left[\alpha_1^{(2)} - 2\alpha_2^{(2)}\right]hx^2 + x^3 = 2h^2x - 3hx^2 + x^3 .$$

Επίσης, οι συνήθεις δυνάμεις του x μπορούν να εκφραστούν συναρτήσει των παραγοντικών δυνάμεων αυτού σύμφωνα με τον ακόλουθο τύπο $(\kappa \geq 1)$

$$x^\kappa = \sum_{j=1}^{\kappa} \beta_j^{(\kappa)} h^{\kappa-j} x^{[j]} ,$$

όπου οι αριθμοί $\beta_j^{(\kappa)}$ καλούνται αριθμοί του Stirling δευτέρου είδους και υπολογίζονται με την βοήθεια του τύπου

$$\beta_j^{(\kappa+1)} = \beta_{j-1}^{(\kappa)} + j\beta_j^{(\kappa)} , \quad \text{όπου} \quad \beta_\kappa^{(\kappa)} = 1 \quad \text{και} \quad \beta_0^{(\kappa)} = 0 .$$

Για παράδειγμα, έχουμε:

$$x = \beta_1^{(1)}x^{[1]} = x^{[1]} ,$$

$$x^2 = \beta_1^{(2)}hx^{[1]} + \beta_2^{(2)}x^{[2]} = \left[\beta_0^{(1)} + \beta_1^{(1)}\right]hx^{[1]} + x^{[2]} = hx^{[1]} + x^{[2]},$$

$$x^3 = \beta_1^{(3)}h^2x^{[1]} + \beta_2^{(3)}hx^{[2]} + \beta_3^{(3)}x^{[3]} = \left[\beta_0^{(2)} + \beta_1^{(2)}\right]h^2x^{[1]} +$$

$$+ \left[\beta_1^{(2)} + 2\beta_2^{(2)}\right]hx^{[2]} + x^{[3]} = h^2x^{[1]} + 3hx^{[2]} + x^{[3]} .$$

2.6 Τύποι πολυωνυμικής παρεμβολής για ισαπέχοντα σημεία παρεμβολής

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

2.6.1 Τύπος των προς τα εμπρός διαφορών των Newton-Gregory

Υποθέτουμε ότι δίνονται οι τιμές $f_i \equiv f(x_i)$, $i = 0(1)n$ της συναρτήσεως $f(x)$ στα ισαπέχοντα σημεία παρεμβολής $x_i = x_0 + ih$, $i = 0(1)n$, όπου h είναι το σταθερό βήμα. Για την περίπτωση των ισαπεχόντων σημείων είναι χρήσιμο να ορίσουμε μία κανονικοποιημένη μεταβλητή t από τη σχέση

$$t = \frac{x - x_0}{h} \quad . \tag{2.80}$$

Έτσι, μια μεταβολή στο x κατά h μονάδες ισοδυναμεί με μεταβολή της t κατά μία μονάδα και αντιστρόφως. Από την (2.80) είναι $x = x_0 + ht$ και θέτοντας $f(x) = f(x_0 + ht) \equiv F(t)$ προφανώς έχουμε $f_i = F(i)$, $i = 0(1)n$. Πρέπει να σημειωθεί ότι η t είναι μία συνεχής μεταβλητή και δεν περιορίζεται σε ακέραιες τιμές.

Σκοπός μας είναι η κατασκευή του πολυωνύμου παρεμβολής βαθμού το πολύ n με σημεία παρεμβολής τα x_i , $i = 0(1)n$ χρησιμοποιώντας τιμές του πίνακα διαφορών. Με την βοήθεια των τελεστών έχουμε

$$F(t) \equiv f(x_0 + ht) = E^t f(x_0) = (1 + \Delta)^t f_0 = \sum_{\kappa=0}^{\infty} \binom{t}{\kappa} \Delta^\kappa f_0 =$$

$$= f_0 + \binom{t}{1} \Delta f_0 + \binom{t}{2} \Delta^2 f_0 + \ldots + \binom{t}{n} \Delta^n f_0 + \ldots \ , \tag{2.81}$$

όπου $\binom{t}{\kappa} = \dfrac{t(t-1)\ldots(t-\kappa+1)}{\kappa!}$ για κάθε θετικό ακέραιο κ . Είναι φανερό ότι από τον πίνακα διαφορών για την $f(x)$ είναι διαθέσιμες μόνον οι τιμές Δf_0 , $\Delta^2 f_0, \ldots, \Delta^n f_0$. Εάν θεωρήσουμε τώρα την συνάρτηση

$$P_n(x) = \sum_{\kappa=0}^{n} \binom{t}{\kappa} \Delta^\kappa f_0 \ , \tag{2.82}$$

όπου το t δίνεται από την (2.80), παρατηρούμε ότι η $P_n(x)$ είναι ένα πολυώνυμο ως προς x βαθμού το πολύ n και ισχύει

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$P_n(x_i) = \sum_{\kappa=0}^{n}\binom{i}{\kappa}\Delta^{\kappa}f_0 = \sum_{\kappa=0}^{i}\binom{i}{\kappa}\Delta^{\kappa}f_0 = (1+\Delta)^i f_0 = E^i f_0 = f_i,\ i = 0(1)n$$

Επομένως, το πολυώνυμο $P_n(x)$ που δίνεται από την (2.82) είναι το ζητούμενο πολυώνυμο παρεμβολής με σημεία παρεμβολής τα x_i , $i = 0(1)n$ και καλείται πολυώνυμο παρεμβολής των προς τα εμπρός διαφορών των Newton – Gregory (Δ – NG) ο δε τύπος (2.81) καλείται τύπος των προς τα εμπρός διαφορών των Newton – Gregory (τύπος Δ – NG) . Ο τύπος (2.82) είναι περισσότερο κατάλληλος για να χρησιμοποιηθεί για την παρεμβολή της $y = f(x)$ στην αρχή του πίνακα, δηλαδή, στην περιοχή του αρχικού σημείου x_0 όπου το t είναι μικρό απολύτως ($|t| < 1$) .

Στην πράξη όταν δίνεται σ' ένα πίνακα ένα σύνολο τιμών της συναρτήσεως $y = f(x)$ για ισαπέχουσες τιμές του x και χρειάζεται να προσεγγίσουμε όσο γίνεται καλύτερα την $f(x)$ σε κάποιο σημείο $\hat{x}$ χρησιμοποιώντας πολυωνυμική παρεμβολή, πρέπει να αποφασίσουμε τι βαθμού πολυώνυμο παρεμβολής πρέπει να χρησιμοποιήσουμε, δηλαδή, πόσες και ποιές από τις τιμές του x πρέπει να χρησιμοποιηθούν σαν τα ισαπέχοντα σημεία παρεμβολής. Πρέπει να σημειωθεί ότι δεν βρίσκουμε αναγκαστικά καλύτερες προσεγγίσεις με τη χρησιμοποίηση πολυωνύμου παρεμβολής όσον το δυνατόν μεγαλύτερου βαθμού. Ένας τρόπος για να έχουμε κάποια ένδειξη για τον βαθμό του πολυωνύμου παρεμβολής που πρέπει να χρησιμοποιηθεί σε μια συγκεκριμένη περίπτωση είναι να παρατηρήσουμε την συμπεριφορά των διαφορών στον πίνακα διαφορών που κατασκευάζεται από τις δοθείσες τιμές της $f(x)$ (βλέπε παρατηρηση σελ. 56). Σαν κανόνας, όταν χρησιμοποιείται το πολυώνυμο παρεμβολής Δ – NG , καλούμε x_0 εκείνο το σημείο του πίνακα που είναι αμέσως προηγούμενο (μικρότερο) του σημείου $\hat{x}$ στο οποίο εφαρμόζεται η παρεμβολή.

Αξίζει να σημειωθεί ότι ο τύπος (2.82) μπορεί να προκύψει και από τον τύπο (2.45) με την χρησιμοποίηση των (2.55) και (2.80). Έτσι, αν χρησιμοποιήσουμε και τον ορισμό (2.76) η (2.45) γράφεται

$$P_n(x) = f_0 + \frac{\Delta f_0}{1!h}(x-x_0)^{[1]} + \frac{\Delta^2 f_0}{2!h^2}(x-x_0)^{[2]} + \ldots + \frac{\Delta^n f_0}{n!h^n}(x-x_0)^{[n]} . \tag{2.83}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Μπορεί εύκολα να δειχθεί ότι $\dfrac{(x-x_0)^{[\kappa]}}{\kappa!h^\kappa} = \binom{t}{\kappa}$, $\kappa = 1(1)n$ και επομένως η (2.83) γίνεται $P_n(x) = \sum_{\kappa=0}^{n} \binom{t}{\kappa} \Delta^\kappa f_0$ με $t = \dfrac{x-x_0}{h}$, δηλαδή, συμπίπτει με την (2.82). Όταν το $h \to 0$ ο τύπος (2.82) γίνεται ένα πολυώνυμο του Taylor για την $f(x)$. Πράγματι, έχουμε

$$\lim_{h\to 0} (x-x_0)^{[\kappa]} = (x-x_0)^\kappa$$

και λόγω της (2.56) ισχύει $\lim\limits_{h\to 0} \dfrac{\Delta^\kappa f_0}{h^\kappa} = f^{(\kappa)}(x_0) = f_0^{(\kappa)}$, οπότε η (2.83) για $h \to 0$ γίνεται

$$P_n(x) = f_0 + (x-x_0)f_0' + \frac{(x-x_0)^2}{2!} f_0'' + \ldots + \frac{(x-x_0)^n}{n!} f_0^{(n)} .$$

Μπορούμε δηλαδή να πούμε ότι ο τύπος (2.81) είναι το διακεκριμένο ανάλογο του τύπου του Taylor. Τονίζεται, τέλος, ότι στην περίπτωση των ισαπεχόντων σημείων παρεμβολής συμφέρει να χρησιμοποιούμε τον τύπο (2.82), παρά το πολυώνυμο παρεμβολής του Lagrange ή την μέθοδο του Aitken, από την άποψη του πλήθους των υπολογισμών. Ο τύπος (2.82) έχει επίσης το πλεονέκτημα που αναφέρθηκε και για την μέθοδο του Aitken.

<u>Παράδειγμα 2.9</u> Δίνεται ο πίνακας τιμών της $f(x)$

x_i	-1	0	1	2
f_i	2	2	2	-2

.

Να βρεθεί το πολυώνυμο παρεμβολής $\Delta - NG$ με σημεία παρεμβολής τα x_i , $i = 0(1)3$.

<u>Λύση</u> Το ζητούμενο πολυώνυμο παρεμβολής βαθμού το πολύ 3 βάσει του τύπου (2.82) είναι

$$P_3(x) = f_0 + \binom{t}{1}\Delta f_0 + \binom{t}{2}\Delta^2 f_0 + \binom{t}{3}\Delta^3 f_0 ,$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

όπου $t = \dfrac{x - x_0}{h}$, $h = 1$, $x_0 = -1$ και $x_i = x_0 + ih$, $i = 0(1)3$.
Κατασκευάζουμε τον πίνακα διαφορών:

x_i	f_i	Δ	Δ^2	Δ^3
-1	2			
		0		
0	2		0	
		0		-3
1	2		-3	
		-3		
2	-1			

Έτσι, έχουμε $P_3(x) = 2 + 0 + 0 + \dfrac{t(t-1)(t-2)}{6}(-3)$ με $t = x + 1$ ή

$$P_3(x) = 2 - \frac{1}{2}\left[(x+1)x(x-1)\right] = -\frac{1}{2}(x^3 - x - 4) \ . \qquad \blacktriangle$$

<u>Παράδειγμα 2.10</u> Δίνεται ο ακόλουθος πίνακας τιμών της $f(x)$:

x	1.0	1.2	1.4	1.6	1.8	2.0
$f(x)$	2.98	3.55	4.45	5.63	7.16	9.01

.

Χρησιμοποιώντας πολυωνυμική παρεμβολή (πολυώνυμο παρεμβολής Δ – NG) να βρεθεί μία καλή προσέγγιση για την τιμή της $f(x)$ στο σημείο $x = 1.5$.

<u>Λύση</u> Από τα δεδομένα κατασκευάζουμε τον ακόλουθο πίνακα διαφορών:

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

x	f(x)	Δ	Δ^2	Δ^3	Δ^4
1.0	2.98				
		57			
1.2	3.55		33		
		90		−5	
1.4	4.45		28		12
		118		7	
1.6	5.63		35		−10
		153		−3	
1.8	7.16		32		
		185			
2.0	9.01				

Παρατηρούμε ότι οι διαφορές τρίτης τάξεως είναι κατά προσέγγιση ίσες με 0 (οι διαφορές δευτέρας τάξεως είναι πρακτικά σταθερές ενώ οι διαφορές τετάρτης τάξεως αυξάνουν) και επομένως η συνάρτηση $f(x)$ προσεγγίζεται καλύτερα από ένα πολυώνυμο παρεμβολής δευτέρου βαθμού. Χρειαζόμαστε μόνον τρία σημεία για να προσδιορίσουμε το πολυώνυμο αυτό. Επειδή ζητούμε να προσεγγίσουμε την τιμή $f(1.5)$ και πρόκειται να χρησιμοποιήσουμε το πολυώνυμο παρεμβολής Δ−NG καλούμε $x_0 = 1.4$, οπότε τα τρία ισαπέχοντα σημεία παρεμβολής είναι $x_0 = 1.4$, $x_1 = 1.6$, $x_2 = 1.8$. Έτσι, το πολυώνυμο παρεμβολής $P_2(x)$ είναι $P_2(x) = f_0 + \binom{t}{1}\Delta f_0 + \binom{t}{2}\Delta^2 f_0$ με $t = \frac{x - x_0}{h} = \frac{x - 1.4}{0.2}$ και επομένως $f(1.5) \approx$

$$\approx P(1.5) = 4.45 + (0.5)(1.18) + \frac{(0.5)(0.5-1)}{2}(0.35) = 4.99625 \approx 5.00 .$$

(Για $x = 1.5$ είναι $t = 0.5$). ▲

2.6.2 Τύπος των προς τα πίσω διαφορών των Newton−Gregory

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Εάν καλέσουμε x_0 το τελευταίο από τα $n+1$ ισαπέχοντα σημεία παρεμβολής στον πίνακα της συναρτήσεως $f(x)$, τότε θα είναι $x_i = x_0 + ih$, $i = -n(1)0$, δηλαδή, τα σημεία παρεμβολής είναι x_{-n}, $x_{-(n-1)}, \ldots, x_{-2}, x_{-1}, x_0$ και f_i, $i = -n(1)0$ είναι οι αντίστοιχες τιμές της $f(x)$. Θεωρώντας πάλι την (2.80) και θέτοντας $f(x_0 + ht) \equiv F(t)$ προκύπτει ότι

$$F(t) \equiv f(x_0 + ht) = E^t f_0 = (1-\nabla)^{-t} f_0 = \sum_{\kappa=0}^{\infty} (-1)^\kappa \binom{-t}{\kappa} \nabla^\kappa f_0 =$$

$$= f_0 - \binom{-t}{1}\nabla f_0 + \binom{-t}{2}\nabla^2 f_0 - \ldots + (-1)^n \binom{-t}{n}\nabla^n f_0 + \ldots =$$

$$= f_0 + t\nabla f_0 + \frac{t(t+1)}{2!}\nabla^2 f_0 + \frac{t(t+1)(t+2)}{3!}\nabla^3 f_0 + \ldots \quad . (2.84)$$

Το πολυώνυμο παρεμβολής στην περίπτωση αυτή, με σημεία παρεμβολής τα x_i, $i = -n(1)0$, δίνεται από τον τύπο

$$P_n(x) = \sum_{\kappa=0}^{n} (-1)^\kappa \binom{-t}{\kappa} \nabla^\kappa f_0 , \qquad (2.85)$$

όπου το t δίνεται από την (2.80) και καλείται πολυώνυμο παρεμβολής των προς τα πίσω διαφορών των Newton – Gregory (∇ – NG). Αντίστοιχα ο τύπος (2.84) καλείται τύπος ∇ –NG. Σε αντίθεση με τον τύπο (2.82) ο (2.85) είναι κατάλληλος για παρεμβολή κοντά στο τέλος του πίνακα. Στην πράξη όταν πρόκειται να χρησιμοποιηθεί ο τύπος (2.85) καλούμε γενικά x_0 εκείνο το σημείο του πίνακα που είναι αμέσως επόμενο (μεγαλύτερο) του σημείου $\hat{x}$ στο οποίο ζητούμε να προσεγγίσουμε την $f(x)$.

<u>Παράδειγμα 2.11</u> Δίνεται ο πίνακας τιμών της $f(x)$ του παραδείγματος 2.10. Να προσεγγιστούν οι τιμές $f(1.5)$ και $f(2.2)$ χρησιμοποιώντας τον τύπο (2.85).

<u>Λύση</u> Έχοντας υπ' όψιν τον πίνακα διαφορών του παραδείγματος 2.10 και την παρατήρηση σχετικά με τον κατάλληλο βαθμό του πολυωνύμου

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

παρεμβολής, εργαζόμαστε ως εξής: i) Για να προσεγγίσουμε την $f(1.5)$ χρησιμοποιούμε το πολυώνυμο παρεμβολής $\nabla - NG$ με σημεία παρεμβολής τα $x_0 = 1.6$, $x_{-1} = 1.4$, $x_{-2} = 1.2$. Έτσι, για $x = 1.5$ είναι $t = \frac{x - x_0}{h} = \frac{1.5 - 1.6}{0.2} = -\frac{1}{2}$ και από την (2.85) για $n = 2$ έχουμε

$$f(1.5) \approx P_2(1.5) = 5.63 - 0.5(1.18) + \frac{(-0.5)(-0.5+1)}{2}(0.28) =$$

$= 5.005 \approx 5.00$. ii) Για την προσέγγιση της $f(2.2)$ καλούμε $x_0 = 2.0$, $x_{-1} = 1.8$, $x_{-2} = 1.6$ και επομένως το πολυώνυμο παρεμβολής $\nabla - NG$ με σημεία παρεμβολής τα x_0 , x_{-1} , x_{-2} είναι $P_2(x) = f_0 + t\nabla f_0 + \frac{t(t+1)}{2!}\nabla^2 f_0$, όπου $t = \frac{x - x_0}{h} = \frac{x - 2.0}{0.2}$. Άρα, $f(2.2) \approx P_2(2.2) = 9.01 + 1.85 + 0.32 = 11.18$. ▲

2.6.3 Τύποι παρεμβολής των Gauss, Stirling και Bessel

Ας υποθέσουμε ότι έχουμε ένα σύνολο $2m+1$ $(n = 2m)$ ισαπεχόντων σημείων παρεμβολής, τα οποία συμβολίζουμε ως εξής:

$$x_{\pm i} = x_0 \pm ih \ , \quad i = 0(1)m \ , \tag{2.86}$$

δηλαδή, x_{-m} , $x_{-(m-1)}, \ldots, x_{-1}$, x_0 , $x_1, \ldots, x_m$. Αν $f_{\pm i}$, $i = 0(1)m$ είναι οι αντίστοιχες γνωστές τιμές της συναρτήσεως $f(x)$ και θέσουμε τα σημεία παρεμβολής με την σειρά

$$x_0 \ , \ x_1 \ , \ x_{-1} \ , \ x_2 \ , \ x_{-2} \ , \ldots , \ x_m \ , x_{-m} \ , \tag{2.87}$$

τότε το πολυώνυμο παρεμβολής του Newton με διηρημένες διαφορές βαθμού το πολύ $n = 2m$ είναι

$$\begin{aligned} P_n(x) = f_0 &+ (x - x_0) f[x_0, x_1] + (x - x_0)(x - x_1) f[x_0, x_1, x_{-1}] + \\ &+ (x - x_0)(x - x_1)(x - x_{-1}) f[x_0, x_1, x_{-1}, x_2] + \ldots + \\ &+ (x - x_0)(x - x_1)(x - x_{-1}) \ldots (x - x_m) f[x_0, x_1, x_{-1}, \ldots, x_{-m}] \ . \end{aligned} \tag{2.88}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Αφού οι διηρημένες διαφορές είναι συμμετρικές συναρτήσεις ισχύουν οι σχέσεις

$$f[x_0,x_1,x_{-1}] = f[x_{-1},x_0,x_1] \ , \ f[x_0,x_1,x_{-1},x_2] = f[x_{-1},x_0,x_1,x_2],$$

$$\ldots, \ f[x_0,x_1,x_{-1},\ldots, x_m,x_{-m}] = f[x_{-m},\ldots,x_{-1},x_0,x_1,\ldots,x_m] \ . \qquad (2.89)$$

Χρησιμοποιώντας κατάλληλα τις (2.55) και (2.53) προκύπτει ότι

$$f[x_{-\kappa},\ldots,x_0,\ldots,x_\kappa] = \frac{\Delta^{2\kappa} f_{-\kappa}}{h^{2\kappa}(2\kappa)!} = \frac{\delta^{2\kappa} f_0}{h^{2\kappa}(2\kappa)!} \quad \text{και}$$

$$f[x_{-\kappa+1},\ldots,x_0,\ldots,x_\kappa] = \frac{\Delta^{2\kappa-1} f_{-\kappa+1}}{h^{2\kappa-1}(2\kappa-1)!} = \frac{\delta^{2\kappa-1} f_{\frac{1}{2}}}{h^{2\kappa-1}(2\kappa-1)!} \ , \ \kappa = 1(1)m. \qquad (2.90)$$

Έτσι, λόγω των (2.89) και (2.90) η (2.88) γίνεται

$$P_n(x) = P_{2m}(x) = f_0 + \binom{t}{1}\delta f_{\frac{1}{2}} + \binom{t}{2}\delta^2 f_0 + \binom{t+1}{3}\delta^3 f_{\frac{1}{2}} +$$

$$+ \binom{t+1}{4}\delta^4 f_0 + \ldots + \binom{t+m-1}{2m-1}\delta^{2m-1} f_{\frac{1}{2}} + \binom{t+m-1}{2m}\delta^{2m} f_0, \qquad (2.91)$$

όπου $t = \dfrac{x - x_0}{h}$. Ο τύπος (2.91) καλείται πρώτος (προς τα εμπρός) τύπος του Gauss για το πολυώνυμο παρεμβολής βαθμού το πολύ $n = 2m$.

Εάν θέσουμε τα σημεία παρεμβολής με την σειρά

$$x_0 \ , \ x_{-1} \ , \ x_1 \ , \ x_{-2} \ , \ x_2 \ , \ldots, x_{-m} \ , \ x_m \ , \qquad (2.92)$$

τότε με ανάλογο τρόπο προκύπτει ο δεύτερος (προς τα πίσω) τύπος του Gauss για το $P_n(x)$ που είναι

$$P_n(x) = P_{2m}(x) = f_0 + \binom{t}{1}\delta f_{-\frac{1}{2}} + \binom{t+1}{2}\delta^2 f_0 + \binom{t+1}{3}\delta^3 f_{-\frac{1}{2}} + \ldots +$$

$$+ \binom{t+m-1}{2m-1}\delta^{2m-1} f_{-\frac{1}{2}} + \binom{t+m}{2m}\delta^{2m} f_0 \ , \qquad (2.93)$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

όπου $t = \frac{x - x_0}{h}$. Οι τύποι (2.91),(2.93) του Gauss έχουν κυρίως θεωρητικό ενδιαφέρον και χρησιμοποιούνται για την εύρεση άλλων τύπων όπως θα δούμε στη συνέχεια.

Εάν σχηματίσουμε το ημιάθροισμα των τύπων (2.91) και (2.93) παίρνουμε τον τύπο του Stirling:

$$P_n(x) = P_{2m}(x) = f_0 + \binom{t}{1}\mu\delta f_0 + \frac{t^2}{2!}\delta^2 f_0 + \binom{t+1}{3}\delta^3 f_0 + \frac{t^2(t^2-1)}{4!}\delta^4 f_0 +$$

$$+ \binom{t+2}{5}\mu\delta^5 f_0 + \ldots + \binom{t+m-1}{2m-1}\mu\delta^{2m-1} f_0 +$$

$$+ \frac{t^2(t^2-1^2)(t^2-2^2)\ldots[t^2-(m-1)^2]}{(2m)!}\delta^{2m} f_0 , \qquad (2.94)$$

όπου μ είναι ο τελεστής μέσης τιμής. Ο τύπος του Stirling είναι κατάλληλος να χρησιμοποιηθεί για μικρό t και συγκεκριμένα για $|t| \leq \frac{1}{4}$.

Αν τώρα υποθέσουμε ότι έχουμε $2m+2$ $(n = 2m+1)$ ισαπέχοντα σημεία παρεμβολής, τα οποία συμβολίζουμε x_{-m}, $x_{-(m-1)}$, ..., x_{-1}, x_0, x_1, ..., x_m, x_{m+1}, όπου ισχύει η (2.86) και $x_{m+1} = x_m + h$, τότε θέτοντας αυτά με την σειρά

$$x_0, x_1, x_{-1}, x_2, x_{-2}, \ldots, x_m, x_{-m}, x_{m+1},$$

ο πρώτος τύπος του Gauss για το πολυώνυμο παρεμβολής βαθμού το πολύ $n = 2m+1$ γίνεται

$$P_n(x) = P_{2m+1}(x) = f_0 + \binom{t}{1}\delta f_{\frac{1}{2}} + \binom{t}{2}\delta^2 f_0 + \binom{t+1}{3}\delta^3 f_{\frac{1}{2}} + \ldots +$$

$$+ \binom{t+m-1}{2m}\delta^{2m} f_0 + \binom{t+m}{2m+1}\delta^{2m+1} f_{\frac{1}{2}} , \qquad (2.95)$$

όπου $t = \frac{x - x_0}{h}$. Επίσης μπορεί να βρεθεί ότι ο δεύτερος τύπος του Gauss για το πολυώνυμο παρεμβολής $P_n(x)$ $(n = 2m+1)$, σε σχέση με την ακόλουθη σειρά των $2m+2$ σημείων παρεμβολής

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$x_1,\ x_0,\ x_2,\ x_{-1},\ x_3,\ x_{-2},\ldots,\ x_{m+1},\ x_{-m}$$

είναι

$$P_n(x) = P_{2m+1}(x) = f_1 + \binom{t-1}{1}\delta f_{\frac{1}{2}} + \binom{t}{2}\delta^2 f_1 + \binom{t}{3}\delta^3 f_{\frac{1}{2}} + \ldots +$$

$$+ \binom{t-1+m}{2m}\delta^{2m} f_1 + \binom{t-1+m}{2m+1}\delta^{2m+1} f_{\frac{1}{2}} \quad . \qquad (2.96)$$

Εάν σχηματίσουμε το ημιάθροισμα των τύπων (2.95) και (2.96) προκύπτει ο ακόλουθος τύπος του Bessel:

$$P_n(x) = P_{2m+1}(x) = \mu f_{\frac{1}{2}} + \left(t - \frac{1}{2}\right)\delta f_{\frac{1}{2}} + \binom{t}{2}\mu\delta^2 f_{\frac{1}{2}} + \frac{t(t-1)(t-\frac{1}{2})}{3!}\delta^3 f_{\frac{1}{2}} +$$

$$+ \binom{t+1}{4}\mu\delta^4 f_{\frac{1}{2}} + \ldots + \binom{t-1+m}{2m}\mu\delta^{2m} f_{\frac{1}{2}} +$$

$$+ \frac{t(t^2-1^2)(t^2-2^2)\ldots[t^2-(m-1)^2](t-m)(t-\frac{1}{2})}{(2m+1)!}\delta^{2m+1} f_{\frac{1}{2}}, \qquad (2.97)$$

όπου $t = \dfrac{x - x_0}{h}$. Ο τύπος του Bessel είναι κατάλληλος για εφαρμογή για $\frac{1}{4} \leq |t| \leq \frac{3}{4}$. Για $t = \frac{1}{2}$ όλοι οι όροι με διαφορές περιττής τάξεως μηδενίζονται. Είναι φανερό ότι οι τύποι των Stirling και Bessel είναι κατάλληλοι γαι παρεμβολή στο κέντρο ενός πίνακα τιμών.

<u>Παράδειγμα 2.12</u> Δίνεται ο πίνακας τιμών του παραδείγματος 2.10. Να προσεγγιστεί η $f(1.5)$ με τον τύπο του Stirling $(m=1)$ και με τον τύπο του Bessel $(m=1)$.

<u>Λύση</u> i) Ο τύπος του Stirling, αν ληφθεί $x_0 = 1.4$ οπότε $x_{-1} = 1.2$, $x_1 = 1.6$, $t = \dfrac{1.5-1.4}{0.2} = 0.5$, δίνει (έχοντας υπ' όψιν τον πίνακα διαφορών του παραδείγματος 2.10)

$$f(1.5) \approx P_2(1.5) = 4.45 + \frac{1}{2}\left(\frac{0.90+1.18}{2}\right) + \frac{\left(\frac{1}{2}\right)^2}{2}(0.28) = 4.45 +$$

$$+ 0.52 + 0.035 = 5.005 \approx 5.00 \ .$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ii) Για τον τύπο του Bessel παίρνουμε $x_0 = 1.4$ οπότε $x_{-1} = 1.2$, $x_1 = 1.6$, $x_2 = 1.8$ $(m = 1,\ n = 3)$. Έτσι έχουμε $t = \frac{x - x_0}{h} = \frac{1.5 - 1.4}{0.2} = \frac{1}{2}$. Λαμβάνοντας υπ' όψιν τον πίνακα διαφορών του παραδείγματος 2.10, από τον τύπο (2.97) παίρνουμε

$$f(1.5) \approx P_3(1.5) = \frac{f_0 + f_1}{2} + \binom{\frac{1}{2}}{2}\left(\frac{\delta^2 f_0 + \delta^2 f_1}{2}\right) = \frac{4.45 + 5.63}{2} +$$

$$+ \frac{\frac{1}{2}(\frac{1}{2} - 1)}{2}\left(\frac{0.28 + 0.35}{2}\right) = 5.04 - \frac{1}{8}(0.315) = 5.000625 \approx 5.00 \ . \blacktriangle$$

2.7 Παρεμβολή κατα Hermite

Ένα άλλο, κάπως διαφορετικό και γενικώτερο πρόβλημα πολυωνυμικής παρεμβολής είναι το ακόλουθο: Εάν δίνονται οι τιμές της συναρτήσεως $f(x)$ καθώς και οι τιμές της πρώτης παραγώγου $f'(x)$ αυτής σ' ένα σύνολο $n+1$ διακεκριμένων σημείων παρεμβολής x_i , $i = 0(1)n$, δηλαδή, είναι γνωστές οι τιμές f_i , f'_i , $i = 0(1)n$, να βρεθεί ένα πολυώνυμο $P(x)$ βαθμού το πολύ $2n+1$ τέτοιο ώστε

$$P(x_i) = f_i \ , \quad P'(x_i) = f'_i \ , \quad i = 0\ (1)n \ . \tag{2.98}$$

Η απάντηση στο πρόβλημα αυτό είναι ότι υπάρχει ένα μοναδικό πολυώνυμο $P(x)$ βαθμού το πολύ $2n+1$ που ικανοποιεί τις (2.98). Αυτό καλείται πολυώνυμο παρεμβολής του Hermite και έχει τη μορφή

$$P(x) = \sum_{i=0}^{n} \left[1 - 2L'_i(x_i)(x - x_i)\right]\left[L_i(x)\right]^2 f_i + \sum_{i=0}^{n}(x - x_i)\left[L_i(x)\right]^2 f'_i \ , \tag{2.99}$$

όπου τα $L_i(x)$, $i = 0(1)n$ ορίζονται από την (2.8), είναι δηλαδή οι καλούμενοι συντελεστές παρεμβολής του Lagrange. Μπορεί να δειχθεί, με τρόπο ανάλογο προς το θεωρημα 2.1, ότι στην περίπτωση της παρεμβολής του Hermite που αναφέρθηκε παραπάνω το σφάλμα αποκοπής δίνεται από την έκφραση

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\varepsilon(x) = P(x) - f(x) = -\frac{f^{(2n+2)}(\xi)}{(2n+2)!} \prod_{i=0}^{n} (x - x_i)^2, \qquad (2.100)$$

όπου ξ ανήκει στο μικρότερο διάστημα που περιέχει τα $x_0, x_1, \ldots, x_n, x$ και με την προυπόθεση ότι υπάρχει η $f^{(2n+2)}(x)$.

Κλείνοντας το κεφάλαιο της παρεμβολής πρέπει να σημειωθεί ότι εξετάστηκε σ' αυτό μόνον η απλούστερη περίπτωση κατά την οποία η συνάρτηση παρεμβολής $F(x)$ (βλέπε παράγραφο 2.1) είναι ένα πολυώνυμο που ικανοποιεί ορισμένες απλές συνθήκες ((2.2) ή (2.98)). Εκτός από την πολυωνυμική παρεμβολή κατά Hermite η οποία μπορεί να γενικευθεί, άλλα είδη παρεμβολής που χρησιμοποιούνται στην πράξη, σε περιπτώσεις που η πολυωνυμική παρεμβολή δεν είναι ικανοποιητική, είναι: η ρητή παρεμβολή όπου η συνάρτηση παρεμβολής είναι μία ρητή συνάρτηση (λόγος δύο πολυωνύμων), η τριγωνομετρική παρεμβολή, η εκθετική παρεμβολή και η spline παρεμβολή (παρεμβολή με spline συναρτήσεις).

Α Σ Κ Η Σ Ε Ι Σ

1. Να δειχθεί με την μέθοδο της εις άτοπον απαγωγής ότι το πολυώνυμο παρεμβολής $P_n(x)$ βαθμού το πολύ n με σημεία παρεμβολής τα x_i , $i = 0(1)n$, που προσεγγίζει την συνάρτηση $f(x)$ είναι μοναδικό.
2. Χρησιμοποιώντας γραμμική παρεμβολή με σημεία παρεμβολής $x_0 = 8$, $x_1 = 27$, να εκτιμηθεί η $\sqrt[3]{20}$.
3. Να κατασκευασθεί το πολυώνυμο παρεμβολής του Lagrange για την συνάρτηση $y = \sin \pi x$, με σημεία παρεμβολής $x_0 = 0$, $x_1 = \frac{1}{6}$, $x_2 = \frac{1}{2}$ και στη συνέχεια να εκτιμηθεί η τιμή $\sin \frac{\pi}{3}$.
4. Να δειχθεί ότι στην περίπτωση που τα σημεία παρεμβολής x_i , $i = 0(1)n$ είναι ισαπέχοντα, $x_i = x_0 + ih$, $i = 0(1)n$, το πολυώνυμο παρεμβολής του Lagrange παίρνει τη μορφή

$$P_n(x) = \frac{1}{n!} \prod_{j=0}^{n} (t - j) \sum_{i=0}^{n} (-1)^{n-i} \binom{n}{i} \frac{f_i}{t - i} , \text{ όπου } t = \frac{x - x_0}{h} .$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

5. Να κατασκευασθεί ένα πολυώνυμο βαθμού ακριβώς 2 που διέρχεται από τα σημεία (x_0, f_0) και (x_1, f_1) όπου $x_0 \neq x_1$.

6. Δίνεται ο πίνακας τιμών

x_i	x_0	x_1	x_2	x_3
f_i	f_0	f_1	f_2	f_3

. Ποιά συνθήκη πρέπει να πληρούν τα δεδομένα του πίνακα ώστε το πολυώνυμο παρεμβολής με σημεία παρεμβολής τα x_i , $i = 0(1)3$ να είναι βαθμού μικρότερου από 3 .

7. Αν υποθέσουμε ότι η $f(x)$ έχει μια ρίζα ξ στο διάστημα $[-1,1]$ και είναι $f(-1) = 9$, $f(0) = 1$ και $f(1) = 1$ να εκτιμηθεί η ρίζα ξ .

8. Να βρεθεί ο συντελεστής του μεγιστοβαθμίου όρου του πολυωνύμου παρεμβολής του Lagrange με σημεία παρεμβολής τα x_i , $i = 0(1)n$ και αντίστοιχες τιμές για την $f(x)$ τις $f_i \equiv f(x_i)$, $i = 0(1)n$.

9. Έχοντας υπ' όψιν τις (2.10), (2.12) να δειχθεί ότι $L_i'(x_i) = \dfrac{\omega''(x_i)}{2\omega'(x_i)}$, $i = 0(1)n$.

0. Αν x_i , $i = 0(1)n$ είναι τα σημεία παρεμβολής και $L_i(x)$, $i = 0(1)n$ οι συντελεστές παρεμβολής του Lagrange, να δειχθεί ότι:

i) $\sum_{i=0}^{n} x_i^m L_i(x) = x^m$, $m = 0(1)n$

ii) $\sum_{i=0}^{n} L_i(0) x_i^m = \begin{cases} 1 & \text{για } m = 0 \\ 0 & \text{για } m = 1(1)n \\ (-1)^n x_0 x_1 \ldots x_n & \text{για } m = n+1 . \end{cases}$

1. Να βρεθεί το πολυώνυμο παρεμβολής για την συνάρτηση $\sqrt{x}$ με σημεία παρεμβολής τα $x_i = i^2$, $i = 0(1)3$.

2. Δίνονται τα ισαπέχοντα σημεία παρεμβολής $x_i = x_0 + ih$, $i = 0(1)3$ $(h > 0)$. Πως πρέπει να εκλεγεί το βήμα h ώστε το απόλυτο σφάλμα αποκοπής κατά την παρεμβολή για την προσέγγιση της συναρτήσεως

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$f(x) = \sin x$ στο διάστημα (x_0, x_3) να είναι $\leq 5 \cdot 10^{-5}$.

13. Να δειχθεί ότι για ισαπέχοντα σημεία παρεμβολής $x_i = x_0 + ih$, $i = 0(1)n$ $(h > 0)$ ισχύει $\max\limits_{x_0 < x < x_n} |\omega(x)| \equiv \max\limits_{x_0 < x < x_n} |(x - x_0) \cdot$

$\cdot (x - x_1) \ldots (x - x_n)| < n!\, h^{n+1}$. Στη συνέχεια με βάση την (2.24) να δειχθεί ότι για $x_0 < x < x_n$ είναι $|\varepsilon(x)| < \frac{h^{n+1}}{n+1} M_{n+1}$, όπου $M_{n+1} = \max\limits_{x_0 < x < x_n} |f^{(n+1)}(x)|$.

14. Δοθέντων των σημείων (x_i, f_i): $(0,1)$, $(1,-1)$, $(2,1)$, $(3,-1)$ και $(4,1)$, να βρεθεί μία προσέγγιση για την $f(\frac{1}{2})$ με την μέθοδο του Aitken.

15. Αφού δειχθεί ότι η εξίσωση $x^3 + x^2 + x - 1 = 0$ έχει ακριβώς μία πραγματική ρίζα $\xi \in (0,1)$, να εκτιμηθεί αυτή χρησιμοποιώντας αντίστροφη παρεμβολή στα σημεία $x = 0, \frac{1}{2}, 1$.

16. Εάν η $f[x, x_0, x_1, \ldots, x_K]$ είναι ένα πολυώνυμο (ως προς x) βαθμού $m > 0$, να δειχθεί ότι η $f[x, x_0, x_1, \ldots, x_K, x_{K+1}]$ είναι ένα πολυώνυμο βαθμού $m - 1$.

17. Εάν $x, x_0, x_1, \ldots, x_n$ είναι $n+2$ διακεκριμένα σημεία, να δειχθεί ότι

$$f[x, x_0, x_1, \ldots, x_n] = \frac{f(x)}{\prod\limits_{i=0}^{n}(x - x_i)} + \sum_{i=0}^{n} \frac{f_i}{(x_i - x)\prod\limits_{\substack{j=0 \\ j \neq i}}^{n}(x_i - x_j)} =$$

$$= \sum_{i=0}^{n} \frac{f[x, x_i]}{\prod\limits_{\substack{j=0 \\ j \neq i}}^{n}(x_i - x_j)} .$$

18. Αν $f(x) = \prod\limits_{i=0}^{n}(x - x_i)$ να δειχθεί ότι $f[x_0, x_1, \ldots, x_p] = 0$ για $p \leq n$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

19. Να εκτιμηθεί η τιμή $f(0)$ με την βοήθεια πίνακα διαφορών αν είναι γνωστό ότι $f(-2) = f(2) = 0$ και $f(-1) = f(1) = 1$.

20. Σε δύο διαδοχικές τιμές της $f(x)$ στον παρακάτω πίνακα υπάρχουν αντίστοιχα τα σφάλματα ε_1 και ε_2 . Να βρεθούν αυτά και να διορθωθούν οι εσφαλμένες τιμές της $f(x)$ με την προυπόθεση ότι η $f(x)$ είναι πολυώνυμο δεύτερου βαθμού.

x	−4	−3	−2	−1	0	1	2	3	4
f(x)	19	11	5	1	0	1	1	5	11

.

21. Να κατασκευασθεί ο πίνακας διαφορών για την συνάρτηση $f(x) = x^3$, για $x = 0(0.1)0.7$ όταν: i) οι τιμές της $f(x)$ είναι ακριβείς ii) οι τιμές της $f(x)$ είναι στρογγυλευμένες σε 2 δ.θ. . Να συγκριθεί το μέγιστο απόλυτο σφάλμα στρογγυλεύσεως στις διαφορές τρίτης τάξεως με το αντίστοιχο που εκτιμάται σύμφωνα με την θεωρία της παραγράφου 2.4.1.

22. Να βρεθούν οι δύο επόμενοι όροι της ακολουθίας $\alpha_1 = 0$, $\alpha_2 = 4$, $\alpha_3 = 6$, $\alpha_4 = 3$, $\alpha_5 = 2, \ldots,$ δοθέντος ότι ο γενικός όρος αυτής παρίσταται από ένα πολυώνυμο τετάρτου βαθμού.

23. Να δειχθούν οι ακόλουθες σχέσεις μεταξύ των τελεστών:

1) $\delta^2 = \nabla\Delta$ 2) $\mu^2 = 1 + \frac{\delta^2}{4}$

3) $2\mu\delta = (E+1)\nabla = \Delta + \nabla$ 4) $\delta^{2m} = \Delta^{2m}E^{-m}$

5) $J = h\left(1 + \frac{1}{2}\Delta - \frac{1}{12}\Delta^2 + \frac{1}{24}\Delta^3 - \frac{19}{720}\Delta^4 + \ldots\right)$

6) $D^2 = \frac{1}{h^2}\left(\Delta^2 - \Delta^3 + \frac{11}{12}\Delta^4 - \frac{5}{6}\Delta^5 + \ldots\right)$

7) $\Delta^2 = D^2 + D^3 + \frac{7}{12}D^4 + \ldots,$ εάν $h = 1$, 8) $E = 1 + \frac{1}{2}\delta^2 + \delta\left(1 + \frac{\delta^2}{4}\right)^{\frac{1}{2}}$

9) $\nabla = -\frac{1}{2}\delta^2 + \delta\left(1 + \frac{\delta^2}{4}\right)^{\frac{1}{2}}$ 10) $\Delta = \nabla(1-\nabla)^{-1}$

11) $E^{\frac{1}{2}} = \frac{1}{2}\delta + \left(1 + \frac{\delta^2}{4}\right)^{\frac{1}{2}}$ 12) $D = \frac{2}{h}(\sinh)^{-1}\frac{\delta}{2}$.

24. Χρησιμοποιώντας την σχέση 7) της ασκήσεως 23 να βρεθεί η $\Delta^2 f(3)$ εάν $f(x) = x^4 - 3x + 1$ και $h = 1$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

25. Να δειχθούν οι ακόλουθες σχέσεις:

1) $D^{-1}f(x) = \int f(x)dx + c$, c = σταθερά .

2) $\Delta(f_K g_K) = f_K \Delta g_K + g_K \Delta f_K + (\Delta f_K)(\Delta g_K) =$

$= f_{K+1}\Delta g_K + g_K \Delta f_K = g_{K+1}\Delta f_K + f_K \Delta g_K$.

3) $\Delta\left(\frac{f_K}{g_K}\right) = \frac{g_K \Delta f_K - f_K \Delta g_K}{g_K g_{K+1}}$.

4) $\Delta \ln(\alpha x) = \ln\left(1 + \frac{h}{x}\right)$, α σταθερά .

5) $\Delta f(x) = f(x)$, όπου $f(x) = 2^{\frac{x}{h}}$.

6) $\Delta f(x) = -2f(x)$, όπου $f(x) = \sin\left(\frac{\pi x}{h}\right)$.

7) $Ef(x) = e^{\alpha h} f(x)$, όπου $f(x) = e^{\alpha x + b}$.

8) $\Delta^2 \sum_{j=0}^{K-1} f_j = \Delta f_K - \Delta f_0$.

9) $\nabla x^{[K]} = hK(x-h)^{[K-1]}$.

10) $\Delta\binom{x}{K} = \binom{x}{K-1}$, αν $h = 1$.

11) $\Delta^n \alpha^x = (\alpha^h - 1)^n \alpha^x$.

26. Αν $h = 1$ να βρεθούν οι εκφράσεις 1) $\Delta\left[x(x-1)(x-2)(x-3)\right]$ 2) $\frac{\Delta^3 x!}{x!}$ για $x = 9$ 3) $\Delta^6\left[(1+x)(1+ax)^2(1+bx)^3\right]$.

27. Αν $h = 1$ να βρεθεί συνάρτηση $f(x)$ για την οποία ισχύει $\Delta f(x) = 2x^2 - 3x + 4$.

28. Να προσδιοριστεί ο βαθμός του πολυωνύμου $P(x)$ εάν $P_K = P_{K-1} + P_{K-3} - P_{K-4}$, όπου $P_K \equiv P(K)$, K ακέραιος .

29. Να εκφραστεί η παράσταση $\Delta^3 E^{-2} \nabla^2 f_4$ συναρτήσει των τιμών της $f(x)$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

30. Να μετασχηματισθεί το πολυώνυμο $P(x) = x^4 + 6x^2 - 7x + 8$ σε παραγοντικές δυνάμεις του x $(h=1)$ και κατόπιν να βρεθεί η $\Delta P(x)$.

31. Να βρεθεί το άθροισμα $S_n = 1^2 + 2^2 + \ldots + n^2$ με την βοήθεια του τύπου $\Delta - NG$.

32. Να δειχθεί ότι αν τα σημεία παρεμβολής x_i , $i = 0(1)n$ είναι ισαπέχοντα, $x_i = x_0 + ih$, $i = 0(1)n$, το πολυώνυμο παρεμβολής $P_n(x)$ λαμβάνει την μορφή

$$P_n(x) = \sum_{i=0}^{n} \sum_{\kappa=0}^{i} (-1)^{i-\kappa} \binom{t}{i} \binom{i}{\kappa} f_\kappa \quad , \quad \text{όπου} \quad t = \frac{x - x_0}{h} \quad .$$

33. Να βρεθεί το πολυώνυμο ελαχίστου βαθμού που διέρχεται από τα σημεία (0,3),(1,2),(2,7),(3,24),(4,59),(5,118) χρησιμοποιώντας τον τύπο του Bessel.

34. Να προσδιορισθεί το πολυώνυμο παρεμβολής, χρησιμοποιώντας τους τύπους των πολυωνύμων παρεμβολής $\Delta - NG$, $\nabla - NG$, τους δύο τύπους του Gauss και τον τύπο του Stirling, από τα ακόλουθα δεδομένα

x	− 2	− 1	0	1	2
y	11	0	1	2	15

.

35. Να βρεθεί ο πρώτος τύπος του Gauss για $n = 2m-1$ και για τα σημεία παρεμβολής

$$x_{-m+1} \ , \ x_{-m+2} \ , \ldots , \ x_0 \ , \ x_1 \ , \ldots , \ x_m \ .$$

36. Να δειχθεί ότι το πολυώνυμο παρεμβολής του Hermite (2.99) ικανοποιεί τις (2.98) και επιπλέον είναι μοναδικό.

37. Να δειχθεί ότι για $n = 1$ το πολυώνυμο παρεμβολής του Hermite βαθμού το πολύ 3 λαμβάνει την μορφή:

$$P_n(x) = \left(\frac{x - x_1}{x_0 - x_1}\right)^2 \left[\left\{1 - 2\left(\frac{x - x_0}{x_0 - x_1}\right)\right\} f_0 + (x - x_0) f_0'\right] +$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$+\left(\frac{x-x_0}{x_1-x_0}\right)^2\left[\left\{1-2\left(\frac{x-x_1}{x_1-x_0}\right)\right\}f_1+(x-x_1)f'_1\right] \quad .$$

38. Αν δίνονται οι τιμές $f_0, f'_0, f''_0, \ldots, f_0^{(n)}$ να βρεθεί ένα πολυώνυμο $P_n(x)$ βαθμού n τέτοιο ώστε $P_n^{(\kappa)}(x_0) = f^{(\kappa)}(x_0) \equiv f_0^{(\kappa)}$, $\kappa = 0(1)n$.

39. Εάν η $f(x)$ είναι ένα πολυώνυμο βαθμού $\leq n$ και $P_n(x)$ είναι το πολυώνυμο παρεμβολής για την $f(x)$ με σημεία παρεμβολής τα x_i, $i = 0(1)n$, να δειχθεί ότι $P_n(x) \equiv f(x)$.

40. Αν x_i, $i = 0(1)n$ είναι σημεία παρεμβολής και $L_i(x)$, $i = 0(1)n$ οι συντελεστές παρεμβολής του Lagrange, να δειχθεί ότι:

$$\sum_{i=0}^{n}(x_i - x)^{\kappa}L_i(x) = 0 \ , \quad \kappa = 1(1)n \ .$$

41. Δίνονται οι τιμές της $f(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt$ στον παρακάτω πίνακα

x	0.0	0.4	0.8	1.2	1.6	2.0
f(x)	0	0.42839	0.74210	0.91031	0.97635	0.99532

Χρησιμοποιώντας τις τιμές του x του πίνακα σαν σημεία παρεμβολής να βρεθεί το πολυώνυμο παρεμβολής και να προσεγγιστεί η $f(x)$ για $x = 0.2(0.4)1.8$.

42. Από τον πίνακα διαφορών ενός πολυωνύμου $P(x)$ τετάρτου βαθμού με $h = 1$ είναι $P(5) = 446$, $\nabla P(5) = 273$, $\delta^2P(4) = 152$, $\Delta^3P(2) = 72$ και $\nabla^4P(5) = 24$. Να βρεθούν οι τιμές $P(1)$ και $P(6)$.

43. Να βρεθεί ένα άνω φράγμα για το απόλυτο σφάλμα αποκοπής, όταν η συνάρτηση $\sqrt{x}$ προσεγγίζεται στο σημείο $x = 115$ από το πολυώνυμο παρεμβολής με σημεία παρεμβολής τα $x_0 = 100$, $x_1 = 121$, $x_2 = 144$.

44. Να αποδειχθούν οι σχέσεις (2.53).

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

3

ΑΡΙΘΜΗΤΙΚΗ ΠΑΡΑΓΩΓΙΣΗ

3.1 Εισαγωγή

Σε πολλά πρακτικά προβλήματα χρειάζεται να υπολογίσουμε τις τιμές των παραγώγων διαφόρων τάξεων σε ορισμένα σημεία, κάποιας συναρτήσεως $y = f(x)$, η οποία δίνεται είτε από μία πολύ πολύπλοκη αναλυτική έκφραση ή από ένα πίνακα τιμών. Στις περιπτώσεις αυτές καταφεύγουμε, αναγκαστικά ίσως, στην αριθμητική παραγώγιση, δηλαδή, στον προσεγγιστικό υπολογισμό των διαφόρων παραγώγων. Η βασική ιδέα της αριθμητικής παραγωγίσεως συνίσταται στην αντικατάσταση της $y = f(x)$ από μία άλλη απλούστερη συνάρτηση $F(x)$, για την οποία η ζητούμενη παράγωγος μπορεί να βρεθεί εύκολα αναλυτικά. Αν δηλαδή η $f(x)$ αντικατασταθεί (προσεγγιστεί) σ' ένα διάστημα $[a,b]$ που μας ενδιαφέρει, από την συνάρτηση παρεμβολής $F(x)$ (συνήθως από ένα πολυώνυμο παρεμβολής), τότε θα έχουμε π.χ. για την πρώτη παράγωγο $f'(x) \approx F'(x)$, $x \in [a,b]$. Εδώ πρέπει να τονισθεί ότι η αριθμητική παραγώγιση είναι γενικά λιγώτερο καλή προσεγγιστική διαδικασία από ότι η παρεμβολή. Αυτό προκύπτει από το γεγονός ότι, ενώ η $F(x)$ μπορεί να προσεγγίζει αρκετά καλά την $f(x)$ στο $[a,b]$, αυτό δεν σημαίνει ότι θα ισχύει το ίδιο και για την $F'(x)$ σε σχέση με την $f'(x)$, πολύ δε περισ-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

σότερο για παραγώγους ανωτέρας τάξεως. Σαν γενικός κανόνας, μπορούμε να πούμε ότι η αριθμητική παραγώγιση πρέπει να χρησιμοποιείται με ιδιαίτερη προσοχή και μόνον όταν δεν είναι δυνατόν να αποφευχθεί.

3.2 Τύποι αριθμητικής παραγωγίσεως με την μέθοδο των πολυωνύμων παρεμβολής

Οι απλούστεροι τύποι αριθμητικής παραγωγίσεως προκύπτουν δια παραγωγίσεως των τύπων παρεμβολής που αναφέρθηκαν στο προηγούμενο κεφάλαιο.

Ας υποθέσουμε ότι δίνονται οι τιμές f_i , $i = 0(1)n$ της συναρτήσεως $y = f(x)$ στα σημεία x_i , $i = 0(1)n$ (γενικά μη ισαπέχοντα) και ζητούμε να υπολογίσουμε την παράγωγο $f^{(\kappa)}(x)$. Αν $P_n(x)$ είναι το πολυώνυμο παρεμβολής βαθμού το πολύ n με σημεία παρεμβολής τα x_i , $i = 0(1)n$, τότε θέτουμε

$$f^{(\kappa)}(x) \approx P_n^{(\kappa)}(x) \ , \quad \kappa = 0,1,2,\ldots \ . \tag{3.1}$$

Αν $\varepsilon(x)$ είναι το σφάλμα αποκοπής του πολυωνύμου παρεμβολής $P_n(x)$, δηλαδή, $\varepsilon(x) = P_n(x) - f(x)$ (βλέπε θεώρημα 2.1), τότε το σφάλμα αποκοπής $E_\kappa(x)$ της παραγώγου $P_n^{(\kappa)}(x)$ δίνεται από τον τύπο

$$E_\kappa(x) = P_n^{(\kappa)}(x) - f^{(\kappa)}(x) = \varepsilon^{(\kappa)}(x) \ , \quad \kappa = 0,1,2,\ldots \ . \tag{3.2}$$

Στην περίπτωση που τα σημεία x_i , $i = 0(1)n$ είναι μη ισαπέχοντα, τότε χρησιμοποιώντας το πολυώνυμο παρεμβολής του Lagrange (2.9) στην (3.1) παίρνουμε για $\kappa = 1$

$$f'(x) \approx P_n'(x) = \sum_{i=0}^{n} L_i'(x) f_i \ , \tag{3.3}$$

όπου τα $L_i(x)$ δίνονται από την (2.8) ή την (2.12). Για $n = 1$ μπορεί εύκολα να βρεθεί (λόγω της (2.15)) ότι

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$f'(x) \approx P_1'(x) = \frac{f_1 - f_0}{x_1 - x_0} \quad . \tag{3.4}$$

Επίσης για $n = 2$, λαμβάνοντας υπ' όψιν την (2.16) έχουμε

$$f'(x) \approx P_2(x) = \frac{(x-x_1)+(x-x_2)}{(x_0-x_1)(x_0-x_2)} f_0 + \frac{(x-x_0)+(x-x_2)}{(x_1-x_0)(x_1-x_2)} f_1 +$$

$$+ \frac{(x-x_0)+(x-x_1)}{(x_2-x_0)(x_2-x_1)} f_2 \quad . \tag{3.5}$$

Εάν ζητούμε την τιμή της $f'(x)$ σε κάποιο από τα σημεία παρεμβολής, έστω το x_m , τότε ο τύπος (3.3) λαμβάνει την μορφή

$$f'(x_m) \approx P_n'(x_m) = \sum_{i=0}^{n} L_i'(x_m) f_i \quad , \tag{3.6}$$

όπου

$$L_i'(x_m) = \begin{cases} \displaystyle\sum_{\substack{j=0 \\ j\neq m}}^{n} \frac{1}{x_m - x_j} \quad , & \text{αν} \quad i = m \\[2ex] \dfrac{\displaystyle\prod_{\substack{j=0 \\ j\neq m,i}}^{n} (x_m - x_j)}{\displaystyle\prod_{\substack{j=0 \\ j\neq i}}^{n} (x_i - x_j)} \quad , & \text{αν} \quad i \neq m \quad . \end{cases} \tag{3.7}$$

Για το σφάλμα αποκοπής του $P_n'(x)$ λόγω της (3.2) και του τύπου (2.24) παίρνουμε

$$E_1(x) = P_n'(x) - f'(x) = \varepsilon'(x) = -\frac{1}{(n+1)!}\Big[\omega'(x) f^{(n+1)}(\xi(x)) +$$

$$+ \omega(x)\frac{d}{dx} f^{(n+1)}(\xi(x))\Big] \quad . \tag{3.8}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Στον τύπο (3.8) η $\frac{d}{dx} f^{(n+1)}(\xi(x))$ δεν μπορεί να υπολογιστεί αφού η συνάρτηση $\xi(x)$ είναι άγνωστη. Αν όμως το x είναι ένα από τα σημεία παρεμβολής x_i , $i = 0(1)n$ (και η $\frac{d}{dx} f^{(n+1)}(\xi(x))$ είναι φραγμένη), τότε ο τύπος (3.8), επειδή $\omega(x_i) = 0$, γίνεται

$$E_1(x_i) = -\frac{\omega'(x_i)}{(n+1)!} f^{(n+1)}(\xi_i) \ , \ i = 0(1)n \ , \tag{3.9}$$

όπου $\xi_i \equiv \xi(x_i)$ και $\xi_i \in [\min\{x_0,x_1,\ldots,x_n\}, \max\{x_0,x_1,\ldots,x_n\}]$. Λόγω της (2.11) η (3.9) γράφεται

$$E_1(x_i) = -\frac{f^{(n+1)}(\xi_i)}{(n+1)!} \prod_{\substack{j=0 \\ j\neq i}}^{n} (x_i - x_j) \ , \ i = 0(1)n \ . \tag{3.10}$$

Από την (3.10) για $n = 1$ και $i = 0$ προκύπτει ότι το σφάλμα αποκοπής στην (3.4) (για $x = x_0$) είναι

$$E_1(x_0) = -\frac{f''(\xi_0)}{2} (x_0 - x_1) \ . \tag{3.11}$$

Στη συνέχεια θα εξεταστεί η συνηθέστερη περίπτωση όπου τα σημεία x_i , $i = 0(1)n$ στα οποία δίνονται οι τιμές της $f(x)$ είναι ισαπέχοντα.

3.2.1 Τύποι αριθμητικής παραγωγίσεως για ισαπέχοντα σημεία παρεμβολής

Αν $x_i = x_0 + ih$, $i = 0(1)n$, $(h > 0)$ και $f_i \equiv f(x_i)$, $i = 0(1)n$ είναι οι γνωστές τιμές της $f(x)$, τότε το πολυώνυμο παρεμβολής του Lagrange παίρνει την μορφή (βλέπε άσκηση 2.4) .

$$P_n(x) = \frac{1}{n!} \prod_{i=0}^{n} (t-i) \sum_{i=0}^{n} (-1)^{n-i} \binom{n}{i} \frac{f_i}{t-i} \ , \text{ όπου } t = \frac{x-x_0}{h} \ .$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Επομένως, λόγω και της $\frac{dt}{dx} = \frac{1}{h}$, παίρνουμε

$$f'(x) \approx P_n'(x) = \frac{1}{h} \sum_{i=0}^{n} \frac{(-1)^{n-i} f_i}{i!(n-i)!} \frac{d}{dt} \left\{ \frac{\prod_{i=0}^{n} (t-i)}{t-i} \right\} , \qquad (3.12)$$

ο δε τύπος (3.10) του σφάλματος αποκοπής γίνεται

$$E_1(x_i) = (-1)^{n-i+1} h^n i!(n-i)! \frac{f^{(n+1)}(\xi_i)}{(n+1)!} , \qquad (3.13)$$

όπου ξ_i είναι κάποια τιμή μεταξύ των x_0 και x_n . Ο τύπος (3.13) για $i = 0$ απλουστεύεται στον ακόλουθο .

$$E_1(x_0) = (-n)^{n+1} h^n \frac{f^{(n+1)}(\xi_0)}{n+1} . \qquad (3.14)$$

Από τον τύπο (3.12) για $n = 1$ έχουμε

$$f'(x) \approx \frac{f_1 - f_0}{h} = \frac{\Delta f_0}{h} , \qquad (3.15)$$

με αντίστοιχο σφάλμα αποκοπής για $x = x_0$ (προκύπτει από την (3.13) για $n = 1$, $i = 0$)

$$E_1(x_0) = \frac{h}{2} f''(\xi_0) . \qquad (3.16)$$

Επίσης, για $n = 2$ και $x = x_1$ από την (3.12) παίρνουμε

$$f'(x_1) \approx \frac{1}{2h} (f_2 - f_0) , \qquad (3.17)$$

με αντίστοιχο σφάλμα αποκοπής

$$E_1(x_1) = \frac{h^2}{6} f'''(\xi_1) . \qquad (3.18)$$

Παραγωγίζοντας την (3.12) και για $n = 2$ παίρνουμε

$$f''(x) \approx \frac{1}{h^2} (f_0 - 2f_1 + f_2) = \frac{\Delta^2 f_0}{h^2} .$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Μέχρι τώρα χρησιμοποιήσαμε το πολυώνυμο παρεμβολής του Lagrange για την εύρεση των διαφόρων τύπων. Αν χρησιμοποιήσουμε το πολυώνυμο παρεμβολής Δ – NG που δίνεται από την (2.82) θα έχουμε

$$f'(x) \approx P_n'(x) = \frac{dP_n(x)}{dx} = \frac{dP_n(x)}{dt}\,\frac{dt}{dx} = \frac{1}{h}\,\frac{dP_n(x)}{dx} =$$

$$= \frac{1}{h}\,\frac{d}{dt}\left\{\sum_{i=0}^{n}\binom{t}{\kappa}\Delta^{\kappa}f_0\right\} = \frac{1}{h}\sum_{i=0}^{n}\frac{d}{dt}\binom{t}{\kappa}\Delta^{\kappa}f_0 = \frac{1}{h}\left[\Delta f_0 + \frac{2t-1}{2}\Delta^2 f_0 + \right.$$

$$\left. + \frac{3t^2-6t+2}{6}\Delta^3 f_0 + \frac{2t^3-9t^2+11t-3}{12}\Delta^4 f_0 + \ldots + \frac{d}{dt}\binom{t}{n}\Delta^n f_0\right]. \tag{3.19}$$

Για $x = x_0$ δηλαδή $t = 0$ και λαμβάνοντας υπ' όψιν ότι για $\kappa \geqq 1$ ισχύει

$$\left[\frac{d}{dt}\binom{t}{\kappa}\right]_{t=0} = \frac{1}{\kappa!}\left\{\frac{d}{dt}\left[t(t-1)(t-2)\ldots(t-\kappa+1)\right]\right\}_{t=0} =$$

$$= \frac{1}{\kappa!}(-1)^{\kappa-1}(\kappa-1)! = \frac{(-1)^{\kappa-1}}{\kappa},$$

ο τύπος (3.19) απλοποιείται στον ακόλουθο

$$f'(x_0) \approx \frac{1}{h}\sum_{\kappa=1}^{n}\frac{(-1)^{\kappa-1}}{\kappa}\Delta^{\kappa}f_0 = \frac{1}{h}\left[\Delta - \frac{\Delta^2}{2} + \frac{\Delta^3}{3} - \frac{\Delta^4}{4} + \ldots + \right.$$

$$\left. + \frac{(-1)^{n-1}}{n}\Delta^n\right]f_0, \tag{3.20}$$

με αντίστοιχο σφάλμα αποκοπής που δίνεται από τον τύπο (3.14). Επειδή σε πολλές περιπτώσεις είναι δύσκολο να εκτιμηθεί η $f^{(n+1)}(\xi_0)$ στον τύπο (3.14), θέτουμε για μικρό h (βλέπε (2.56))

$$f^{(n+1)}(\xi_0) \approx \frac{\Delta^{n+1}f_0}{h^{n+1}},$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

οπότε ο (3.14) δίνει

$$E_1(x_0) \approx \frac{(-1)^{n+1}}{(n+1)h} \Delta^{n+1} f_0 \quad . \tag{3.21}$$

Από τον (3.20) για $n = 1,2$ παίρνουμε αντίστοιχα

$$f'(x_0) \approx \frac{\Delta f_0}{h} = \frac{f_1 - f_0}{h} \quad , \text{(βλέπε και (3.15))}$$

$$f'(x_0) \approx \frac{1}{2h}(-3f_0 + 4f_1 - f_2) \ . \tag{3.22}$$

Για την εύρεση προσεγγιστικών τύπων για την $f^{(\kappa)}(x)$, $1 < \kappa \leq n$ παίρνουμε

$$f^{(\kappa)}(x) \approx P_n^{(\kappa)}(x) = \frac{1}{h^\kappa} \sum_{i=0}^{n} \frac{d^\kappa}{dt^\kappa} \binom{t}{i} \Delta^i f_0 = \frac{1}{h^\kappa} \sum_{i=\kappa}^{n} \frac{d^\kappa}{dt^\kappa} \binom{t}{i} \Delta^i f_0 \ .$$

Για $\kappa = n$ ο παραπάνω τύπος δίνει

$$f^{(n)}(x) \approx \frac{\Delta^n f_0}{h^n} \quad , \quad \text{για κάθε } x \in [x_0, x_n] \tag{3.23}$$

αφού $\frac{d^n}{dt^n}\binom{t}{n} = \frac{1}{n!} \frac{d^n}{dt^n}(t^n) = 1$.

Αν η $f(x)$ είναι πολυώνυμο n βαθμού τότε ο τύπος (3.23) είναι ακριβής (βλέπε άσκηση 6).

Ειδικά για την $f''(x)$ έχουμε

$$f''(x) \approx P_n''(x) = \frac{d}{dx} P_n'(x) = \frac{1}{h} \frac{d}{dt} P_n'(x) = \frac{1}{h^2} \sum_{\kappa=2}^{n} \frac{d^2}{dt^2} \binom{t}{\kappa} \Delta^\kappa f_0 =$$

$$= \frac{1}{h^2} \left[\Delta^2 f_0 + (t-1)\Delta^3 f_0 + \frac{6t^2 - 18t + 11}{12} \Delta^4 f_0 + \ldots + \frac{d^2}{dt^2}\binom{t}{n} \Delta^n f_0 \right] . \tag{3.24}$$

Από τον (3.24) για $x = x_0$, $(t = 0)$ προκύπτει ότι

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$f''(x_0) \approx \frac{1}{h^2}\left[\Delta^2 f_0 - \Delta^3 f_0 + \frac{11}{12}\Delta^4 f_0 - \frac{5}{6}\Delta^5 f_0 + \ldots + \right.$$

$$\left. + \left\{\frac{d^2}{dt^2}\binom{t}{n}\right\}_{t=0} \Delta^n f_0\right] . \qquad (3.25)$$

Από την (3.25) για $n = 2$ προκύπτει ο απλούστερος τύπος για την $f''(x_0)$ (βλέπε και (3.23))

$$f''(x_0) \approx \frac{\Delta^2 f_0}{h^2} = \frac{1}{h^2}(f_2 - 2f_1 + f_0) . \qquad (3.26)$$

Αν χρησιμοποιήσουμε το πολυώνυμο ∇ – NG (τα σημεία παρεμβολής τώρα είναι $x_{-n}, x_{-n+1}, \ldots, x_{-1}, x_0$) που δίνεται από την (2.85), τότε ο αντίστοιχος του τύπου (3.20) είναι

$$f'(x_0) \approx \frac{1}{h}\sum_{\kappa=1}^{n} \frac{\nabla^\kappa f_0}{\kappa} = \frac{1}{h}\left[\nabla + \frac{\nabla^2}{2} + \frac{\nabla^3}{3} + \ldots + \frac{\nabla^n}{n}\right] f_0 , \qquad (3.27)$$

με αντίστοιχο σφάλμα αποκοπής

$$E_1(x_0) = -\frac{h^n}{n+1} f^{(n+1)}(\xi_0) . \qquad (3.28)$$

Ο (3.28) προκύπτει από τον (3.10) με κατάλληλη τροποποίηση ως προς τα σημεία παρεμβολής.

Επίσης, αν χρησιμοποιηθεί ο τύπος του πολυωνύμου παρεμβολής του Stirling (2.94) π.χ. με $m = 3$ ($n = 6$) παίρνουμε τους τύπους

$$f'(x_0) \approx \frac{1}{h}\left[\mu\delta f_0 - \frac{1}{6}\mu\delta^3 f_0 + \frac{1}{30}\mu\delta^5 f_0\right] , \qquad (3.29)$$

$$f''(x_0) \approx \frac{1}{h^2}\left[\delta^2 f_0 - \frac{1}{12}\delta^4 f_0 + \frac{1}{90}\delta^6 f_0\right] . \qquad (3.30)$$

Πρέπει να σημειωθεί ότι πολλοί από τους τύπους της παραγράφου αυτής μπορούν να προκύψουν ευκολώτερα με την χρησιμοποίηση συμβολικών μεθόδων. Πράγματι, έχουμε

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$f^{(\kappa)}(x_0) = D^{\kappa} f_0 \ , \qquad (3.31)$$

όπου D είναι ο τελεστής παραγωγίσεως. Λόγω της (2.74) και για $\kappa = 1$ από την (3.31) προκύπτει ο τύπος (3.20) με την προυπόθεση ότι έχουμε τα $n+1$ σημεία παρεμβολής $x_i = x_0 + ih$, $i = 0(1)n$. Για $\kappa = 2$ προκύπτει ο τύπος (3.24). Αν χρησιμοποιηθεί η σχέση $D = -\frac{1}{h}\ln(1-\nabla)$ στην (3.31), τότε για $\kappa = 1$ προκύπτει ο τύπος (3.27). Με ανάλογο τρόπο μπορούν να προκύψουν και οι (3.29), (3.30).

Κλείνουμε την παρούσα παράγραφο με την παρατήρηση ότι η αριθμητική παραγώγιση μπορεί να χρησιμοποιηθεί για την εκτίμηση τοπικών ακροτάτων μιας συναρτήσεως $f(x)$ η οποία δίνεται από ένα πίνακα τιμών, δηλαδή, για την αριθμητική λύση της $f'(x) = 0$.

<u>Παράδειγμα 3.1</u> Στον παρακάτω πίνακα δίνονται οι τιμές της $f(x) = \sqrt{x}$ για $x = 1.00(0.05)1.30$ στρογγυλευμένες σε 5 δ.θ. . Να εκτιμηθούν οι $f'(1)$, $f''(1)$, $f'(1.30)$, $f'(1.15)$, $f''(1.15)$, χρησιμοποιώντας μέχρι και τρίτης τάξεως διαφορές, καθώς και το σφάλμα αποκοπής για την προσέγγιση της $f'(1)$.

x	$f(x) = \sqrt{x}$
1.00	1.00000
1.05	1.02470
1.10	1.04881
1.15	1.07238
1.20	1.09545
1.25	1.11803
1.30	1.14018

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

<u>Λύση</u> Σχηματίζουμε τον πίνακα διαφορών:

x	f(x)	Δ	Δ^2	Δ^3	Δ^4
1.00	1.00000				
		2470			
1.05	1.02470		− 59		
		2411		5	
1.10	1.04881		− 54		− 1
		2357		4	
1.15	1.07238		− 50		− 3
		2307		1	
1.20	1.09545		− 49		5
		2258		6	
1.25	1.11803		− 43		
		2215			
1.30	1.14018				

Σύμφωνα με τον τύπο (3.20) για $n = 3$ έχουμε $(h = 0.05)$

$$f'(1) \approx \frac{1}{0.05}\left[\Delta - \frac{1}{2}\Delta^2 + \frac{\Delta^3}{3}\right] f(1) = 20\left[0.02470 - \frac{1}{2}(-0.00059) + \right.$$

$$\left. + \frac{1}{3}(0.00005)\right] \approx 0.50024 \ .$$

Αν χρησιμοποιηθεί ο τύπος (3.21) για το σφάλμα αποκοπής έχουμε

$$E_1(1) \approx \frac{1}{0.05}\ \frac{\Delta^4 f(1)}{4} = 5(-0.00001) = -0.5 \cdot 10^{-4} \ .$$

Από τον τύπο (3.25) για $n = 3$ παίρνουμε $f''(1) \approx$

$\approx \dfrac{1}{(0.05)^2}\left[(-59)\cdot 10^{-5} - 5\cdot 10^{-5}\right] = -0.256$. Χρησιμοποιώντας τον

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

τύπο (3.27) για $n = 3$ προκύπτει ότι $f'(1.30) \approx \frac{1}{0.05}\left[2215 \cdot 10^{-5} + \frac{(-43) \cdot 10^{-5}}{2} + \frac{6 \cdot 10^{-5}}{3}\right] = 0.43910$. Τέλος με τη βοήθεια των τύπων (3.29), (3.30) έχουμε

$$f'(1.15) \approx \frac{1}{0.05}\left[\frac{2307 + 2357}{2} \cdot 10^{-5} - \frac{1}{6}\,\frac{4+1}{2} \cdot 10^{-5}\right] \approx 0.46632 \quad \text{και}$$

$$f''(1.15) \approx \frac{1}{(0.05)^2}\left[(-50) \cdot 10^{-5}\right] = -0.2$$. Αξίζει να σημειωθεί ότι, αφού $f'(x) = \frac{1}{2\sqrt{x}}$ και $f''(x) = -\frac{1}{4x^{3/2}}$, οι ακριβείς τιμές είναι $f'(1) = 0.5$, $f''(1) = -0.25$, $f'(1.30) \approx 0.43853$, $f'(1.15) \approx 0.46625$ και $f''(1.15) \approx -0.20272$. ▲

3.3 Τύποι αριθμητικής παραγωγίσεως με την μέθοδο των προσδιοριστέων συντελεστών

Στην παράγραφο 3.2 βρέθηκαν τύποι αριθμητικής παραγωγίσεως δια παραγωγίσεως του πολυωνύμου παρεμβολής. Ένας άλλος τρόπος για τον προσδιορισμό τύπων αριθμητικής παραγωγίσεως που σε πολλές περιπτώσεις απαιτεί λιγώτερη προσπάθεια είναι η καλούμενη μέθοδος των προσδιοριστέων συντελεστών. Σκοπός μας με τη μέθοδο αυτή είναι να προσδιοριστούν οι συντελεστές σ' ένα προκαθορισμένο γραμμικό συνδυασμό τιμών της $f(x)$ (και παραγώγων της $f(x)$), έτσι ώστε ο γραμμικός αυτός συνδυασμός να αποτελεί καλή προσέγγιση για την ζητούμενη παράγωγο. Ο προσδιορισμός των συντελεστών γίνεται έτσι ώστε ο προσεγγιστικός τύπος που θα προκύψει να είναι ακριβής για πολυώνυμα βαθμού $\leq \ell$ όπου το ℓ είναι όσο γίνεται μεγαλύτερο. Το μεγαλύτερο ℓ , για το οποίο ο τύπος αριθμητικής παραγωγίσεως είναι ακριβής για πολυώνυμα βαθμού $\leq \ell$ καλείται τάξη του τύπου. Πρέπει να σημειωθεί ότι οι τύποι που προκύπτουν με την μέθοδο των προσδιοριστέων συντελεστών συμπίπτουν με τους

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

αντίστοιχους που προκύπτουν με την βοήθεια του πολυωνύμου παρεμβολής. Το πλεονέκτημα της μεθόδου των προσδιοριστέων συντελεστών είναι ότι μας επιτρέπει να εκλέγουμε εκ των προτέρων την μορφή του τύπου που θα χρησιμοποιήσουμε.

Σαν παράδειγμα, ας υποθέσουμε ότι δίνονται τα τρία ισαπέχοντα σημεία $x_{\pm i} = x_0 \pm ih$, $i = 0,1$ και οι αντίστοιχες τιμές $f_{\pm i}$, $i = 0,1$ της $f(x)$ και θέλουμε να εκφράσουμε την $f'(x_0)$ από τον ακόλουθο προσεγγιστικό τύπο

$$f'(x_0) \approx \alpha_{-1} f_{-1} + \alpha_0 f_0 + \alpha_1 f_1 \quad . \tag{3.32}$$

Οι συντελεστές α_{-1}, α_0, α_1 στον τύπο (3.32) θα προσδιοριστούν έτσι ώστε αυτός να είναι ακριβής (δηλαδή τα δύο μέλη του να ταυτίζονται) για πολυώνυμα όσον το δυνατόν μεγαλύτερου βαθμού. Είναι εύκολο να δειχθεί, λόγω της γραμμικότητος του τελεστή της παραγώγου, ότι αν ο τύπος (3.32) είναι ακριβής όταν εφαρμόζεται στα ειδικά πολυώνυμα 1, $x, x^2, \ldots, x^p$, τότε θα είναι ακριβής και όταν εφαρμόζεται σε οποιοδήποτε πολυώνυμο βαθμού $\leq p$. Είναι φανερό επίσης ότι αν m είναι το πλήθος των προσδιοριστέων συντελεστών (στην περίπτωση του (3.32) είναι $m = 3$) τότε η τάξη του ζητούμενου τύπου είναι τουλάχιστον $m-1$. Έτσι από τον τύπο (3.32), για $f(x) = 1, x, x^2$ εξισώνοντας τα δύο μέλη του, παίρνουμε αντίστοιχα τις παρακάτω εξισώσεις

$$\begin{cases} \alpha_{-1} + \alpha_0 + \alpha_1 = 0 \\ \alpha_{-1} x_{-1} + \alpha_0 x_0 + \alpha_1 x_1 = 0 \\ \alpha_{-1} x_{-1}^2 + \alpha_0 x_0^2 + \alpha_1 x_1^2 = 2x_0 \quad . \end{cases} \tag{3.33}$$

Το σύστημα (3.33) είναι ισοδύναμο με το ακόλουθο

$$\begin{cases} \alpha_{-1} + \alpha_0 + \alpha_1 = 0 \\ -\alpha_{-1} \qquad + \alpha_1 = \dfrac{1}{h} \\ \alpha_{-1} \qquad + \alpha_1 = 0 \end{cases}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

από το οποίο προκύπτει ότι $\alpha_1 = \alpha_{-1} = \frac{1}{2h}$, $\alpha_0 = 0$ και επομένως ο τύπος (3.32) γίνεται

$$f'(x_0) \approx \frac{1}{2h}(f_1 - f_{-1}) \ . \tag{3.34}$$

Ο τύπος (3.34) είναι ακριβής για πολυώνυμα βαθμού ≤ 2 και είναι εύκολο να δειχθεί ότι για $f(x) = x^3$ τα δύο μέλη του είναι διαφορετικά οπότε η τάξη του είναι $\ell = 2$. Αξίζει να σημειωθεί ότι ο τύπος (3.34) συμπίπτει με εκείνον που προκύπτει αν στον τύπο (3.29) διατηρήσουμε μόνον ένα όρο στο δεύτερο μέλος $(m = 1,\ n = 2)$.

Μία άλλη μέθοδος παραγωγής τύπων αριθμητικής παραγωγίσεως ειδικά για την περίπτωση ισαπεχόντων σημείων $x_{\pm i} = x_0 \pm ih$, $i = 0,1,2,\ldots$ είναι η μέθοδος της σειράς Taylor. Αφού $f_{\pm i} \equiv f(x_{\pm i}) = f(x_0 \pm ih)$, $i = 0,1,2,\ldots$ από το ανάπτυγμα της σειράς Taylor για τις $f_{\pm i}$ γύρω από το σημείο x_0 έχουμε

$$f_{\pm i} = f(x_0 \pm ih) = f_0 \pm ihf_0' + \frac{(ih)^2}{2!}f_0'' \pm \frac{(ih)^3}{3!}f_0''' +$$

$$+ \frac{(ih)^4}{4!}f_0^{(4)} \pm \frac{(ih)^5}{5!}f_0^{(5)} + \ldots \quad , \ i = 0,1,2,\ldots \ . \tag{3.35}$$

Για την εύρεση τύπου π.χ. για την $f_0' \equiv f'(x_0)$ σχηματίζουμε την διαφορά $f_1 - f_{-1}$. Λόγω της (3.35) παίρνουμε

$$f_1 - f_{-1} = 2hf_0' + \frac{h^3}{3}f_0''' + \frac{h^5}{60}f_0^{(5)} + \ldots \ . \tag{3.36}$$

Εάν αποκόψουμε την σειρά στο δεύτερο μέλος της (3.36) υπολογίζοντας την f''' σε κάποιο σημείο ξ μεταξύ των x_{-1} και x_1 θα έχουμε

$$f_1 - f_{-1} = 2hf_0' + \frac{h^3}{3}f'''(\xi) \quad \text{ή}$$

$$f'(x_0) = f_0' = \frac{f_1 - f_{-1}}{2h} - \frac{h^2}{6}f'''(\xi) \ . \tag{3.37}$$

Επομένως, ο τύπος που βρέθηκε είναι $f'(x_0) \approx \frac{f_1 - f_{-1}}{2h}$ (βλέπε και (3.34)) με αντίστοιχο σφάλμα αποκοπής $\frac{h^2}{6}f'''(\xi)$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Για την εύρεση τύπου για την $f''(x_0)$ έχουμε

$$f_1 + f_{-1} = 2f_0 + h^2 f_0'' + \frac{h^4}{12} f_0^{(4)} + \dots ,$$

από την οποία προκύπτει ότι

$$f_0'' = f''(x_0) \approx \frac{f_1 - 2f_0 + f_{-1}}{h^2} \tag{3.38}$$

με αντίστοιχο σφάλμα αποκοπής $\frac{h^2}{12} f^{(4)}(\xi)$, όπου ξ μεταξύ των x_{-1} και x_1 . Ο τύπος (3.38) συμπίπτει με τον (3.30) αν διατηρήσουμε μόνον ένα όρο στο δέυτερο μέλος του (3.30) $(m = 1,\ n = 2)$.

Θεωρώντας κατάλληλους γραμμικούς συνδυασμούς των εξισώσεων που προκύπτουν από τις (3.35) για τις διάφορες τιμές του i , μπορούμε να βρούμε διάφορους τύπους για οποιαδήποτε παράγωγο της $f(x)$ στο σημείο x_0 καθώς και τα αντίστοιχα σφάλματα αποκοπής.

3.4 Σφάλματα στην αριθμητική παραγώγιση

Από τις προηγούμενες παραγράφους γίνεται φανερό ότι στην περίπτωση των ισαπεχόντων σημείων παρεμβολής $x_i = x_0 + ih$, $i = 0(1)n$ κάθε τύπος αριθμητικής παραγωγίσεως για την εύρεση της $f^{(\kappa)}(x)$ $(\kappa \geq 1)$ σε κάποιο σημείο παρεμβολής έχει τη γενική μορφή

$$f^{(\kappa)}(x_j) \approx P_n^{(\kappa)}(x_j) = \frac{1}{h^\kappa} \sum_{i=0}^{n} A_i f_i , \tag{3.39}$$

$j = 0(1)n$, όπου A_i , $i = 0(1)n$ είναι σταθερές. Το αντίστοιχο σφάλμα αποκοπής έχει την γενική μορφή

$$E_\kappa(x_j) = \frac{h^p f^{(m)}(\xi_j)}{\alpha} , \tag{3.40}$$

όπου $m \geq n+1$, α σταθερά και p θετικός ακέραιος με $p \geq n+1-\kappa$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

(Για $\kappa = 1$ είναι $p = n$ και $m = n+1$ (βλέπε (3.13))).

Έστω $\left| \dfrac{f^{(m)}(\xi_j)}{\alpha} \right| \leqq A$. Τότε έχουμε

$$\left| E_\kappa(x_j) \right| \leqq Ah^p \ , \quad A \text{ σταθερά} \ . \tag{3.41}$$

Αφού το σφάλμα αποκοπής που δίνεται από την (3.40) μηδενίζεται για πολυώνυμα βαθμού το πολύ $m-1$ είναι φανερό ότι η τάξη του τύπου (3.39) θα είναι $m-1$, θα ισχύει δε $\sum_{i=0}^{n} A_i = 0$.

Στο δεύτερο μέλος του τύπου (3.39) η f_i συμβολίζει την ακριβή τιμή της $f(x)$ στο σημείο x_i . Αν υποθέσουμε ότι το δεύτερο μέλος του (3.39) υπολογίζεται με την αντικατάσταση της f_i από την προσεγγιστική της τιμή $\bar{f}_i$, πράγμα που συμβαίνει συνήθως στην πράξη, όπου $\bar{f}_i - f_i = \varepsilon_i$, $i = 0(1)n$ και $|\varepsilon_i| \leqq M$, $i = 0(1)n$, ενώ οι τιμές των h , A_i θεωρούνται ακριβείς και επιπλέον αγνοήσουμε χάριν απλότητος το παραχθέν σφάλμα κατά τους υπολογισμούς (ως αμελητέο σε σχέση με το διαδοθέν), τότε για το διαδοθέν σφάλμα ε , σύμφωνα με το κεφάλαιο 1 θα ισχύει

$$|\varepsilon| \leqq \frac{1}{h^\kappa} \sum_{i=0}^{n} |A_i||\varepsilon_i| \leqq \frac{M}{h^\kappa} \sum_{i=0}^{n} |A_i| \equiv \frac{B}{h^\kappa} \ . \tag{3.42}$$

Επομένως, για το ολικό σφάλμα $E_{o\lambda}$ κατά την χρησιμοποίηση του τύπου (3.39) ισχύει

$$|E_{o\lambda}| = |\varepsilon + E_\kappa(x_j)| \leqq |\varepsilon| + |E_\kappa(x_j)| \leqq \frac{B}{h^\kappa} + Ah^p \equiv g(h) \ . \tag{3.43}$$

Από την (3.42) προκύπτει ότι εάν θέλουμε το απόλυτο υπολογιστικό σφάλμα να είναι μικρό πρέπει να χρησιμοποιήσουμε μία μεγάλη σχετικά τιμή για το h . Εξ άλλου από την (3.41) είναι φανερό ότι για να είναι το απόλυτο σφάλμα αποκοπής μικρό πρέπει το h να είναι αρκετά μικρό. Ε-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

τσι για την ελαχιστοποίηση της $g(h)$ στην (3.43) πρέπει να εκλέξουμε κάποια βέλτιστη τιμή h_{opt} για το h. Από την εξίσωση $g'(h) = 0$ προκύπτει ότι $h_{opt} = \left(\frac{\kappa B}{pA}\right)^{\frac{1}{p+\kappa}}$. Πρέπει να σημειωθεί ότι το παραπάνω συμπέρασμα προέκυψε θεωρώντας ένα άνω φράγμα για το $|E_{o\lambda}|$, δηλαδή, την $g(h)$ στην (3.43). Στην πράξη βέβαια τα πράγματα είναι διαφορετικά και δυστυχώς είναι σχεδόν αδύνατον να εκτιμηθεί η καλύτερη τιμή για το h. Συνήθως προτιμούμε να χρησιμοποιήσουμε ένα περισσότερο ακριβή τύπο με σχετικά μεγάλο h παρά ένα λιγώτερο ακριβή τύπο με πολύ μικρό h. Πρέπει να σημειωθεί επίσης ότι τύποι της γενικής μορφής των (3.39) και (3.40) ισχύουν για το τυχόν σημείο x και όχι μόνον για τα σημεία παρεμβολής.

Α Σ Κ Η Σ Ε Ι Σ

1. Να δειχθούν οι σχέσεις (3.7).
2. Να βρεθεί το σφάλμα αποκοπής του τύπου (3.5) για $x = x_1$.
3. Να βρεθούν οι τύποι αριθμητικής παραγωγίσεως και τα αντίστοιχα σφάλματα αποκοπής, που προκύπτουν από τις (3.12) και (3.13) για: i) $n = 2$, $x = x_0$ ii) $n = 2$, $x = x_2$.
4. Ένα αυτοκίνητο κινούμενο ευθυγράμμως με σταθερή επιτάχυνση διήνυσε τις ακόλουθες αποστάσεις στον αντίστοιχο χρόνο

t(sec)	1	2	2.2
S(m)	35	70	78.2

. Να εκτιμηθεί η ταχύτητά του για $t = 1.5$ sec και $t = 2.2$ sec.

5. Δίνεται η συνάρτηση $f(x) = 3x^2 + 2x + 1$. Να εκτιμηθεί η $f'(2)$ με τους τύπους (3.15), (3.27) (για $n = 1$) και (3.34), με $h = 1$. Ποιός είναι πιο ακριβής;

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

6. Να δειχθεί ότι ο τύπος (3.23) είναι ακριβής όταν η $f(x)$ είναι πολυώνυμο n βαθμού.

7. Να εκτιμηθεί η $f'(0)$ με την βοήθεια του τύπου (3.22) από τα ακόλουθα δεδομένα:

x	-2	-1	0	1	2	3
$f(x)$	0.7	0.8	3.0	1.5	0.3	0

.

8. Να βρεθεί τύπος ανάλογος των (3.20), (3.25) για την $f'''(x_0)$ για $n = 5$.

9. Να βρεθεί ο τύπος αριθμητικής παραγωγίσεως $f'(x_0) \approx \dfrac{f_1 - f_0}{h}$ και το αντίστοιχο σφάλμα αποκοπής με την βοήθεια της σειράς Taylor.

10. Να βρεθεί ο τύπος (3.26) με την βοήθεια του τύπου $f'(x_0) \approx \approx \dfrac{f_1 - f_0}{h}$.

11. Αν $x_i = x_0 + ih$, $i = 0,1,2$ είναι τα σημεία παρεμβολής ($n = 2$), να προσεγγιστούν οι τιμές $f'(x_1)$, $f'(x_2)$, $f''(x_1)$, $f''(x_2)$ με την βοήθεια συμβολικών μεθόδων.

12. Με την βοήθεια της σειράς Taylor να βρεθούν τα σφάλματα αποκοπής για τους παρακάτω τύπους αριθμητικής παραγωγίσεως:

i) $f'(x_0 + \frac{h}{2}) \approx \dfrac{f(x_0 + h) - f(x_0)}{h}$

ii) τύπος (3.26)

iii) $f''(x_0 + h) \approx \dfrac{1}{h^2}\left[f(x_0 + 2h) - 2f(x_0 + h) + f(x_0)\right]$

iv) $f'(x_0) \approx \dfrac{1}{2h}\left[f(x_0 - 2h) - 8f(x_0 - h) + 8f(x_0 + h) - \right.$
$\left. - f(x_0 + 2h)\right]$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

13. Να προσεγγιστεί η f_0''' εφαρμόζοντας τον τύπο (3.34) στην f'' και χρησιμοποιώντας τον τύπο (3.38).

14. Χρησιμοποιώντας το πολυώνυμο παρεμβολής του Stirling (2.94) να δειχθεί ότι

$$f''(x_0) \approx \frac{1}{h^2} \sum_{i=1}^{m} \frac{2(-1)^{i-1}[(i-1)!]^2}{(2i)!} \delta^{2i} f_0 \ ,$$

με σφάλμα αποκοπής $E_2(x_0) = \dfrac{2(-1)^{m+1}(m!)^2}{(2m+2)!} h^{2m} f^{(2m+2)}(\xi_0)$.

15. Πως πρέπει να εκλεγεί το h στον τύπο (3.34) ώστε το απόλυτο σφάλμα να είναι $\leq \frac{1}{2} 10^{-6}$; (Υποθέτουμε ότι δεν υπάρχουν σφάλματα στρογγυλεύσεως). Δίνεται ότι $f(x) = \frac{1}{2} x^3$.

16. Για την προσέγγιση της $f'(x_0)$ χρησιμοποιείται ο τύπος $f'(x_0) \approx \approx \dfrac{f_1 - f_0}{h}$, όπου οι τιμές f_0, f_1 είναι στρογγυλευμένες σε 2 δ.θ. Τι βήμα h πρέπει να χρησιμοποιηθεί ώστε το απόλυτο ολικό σφάλμα να είναι ελάχιστο; Δίνεται ότι $f''(x_0) = 1$.

17. Να προσδιορισθούν οι συντελεστές A, B, C, D και E έτσι ώστε ο τύπος

$$hf'(x_0) \approx Af_{-2} + Bf_{-1} + Cf_0 + Df_1 + Ef_2$$

να είναι όσον το δυνατόν περισσότερο ακριβής. Δίνεται ότι

$$f_{\pm i} \equiv f(x_{\pm i}) \equiv f(x_0 \pm ih) \ , \quad i = 0,1,2 \ .$$

18. Να γίνει το ίδιο με την άσκηση 17 για τους τύπους

i) $h^2 f_0'' \approx Af_{-2} + Bf_{-1} + Cf_0 + Df_1 + Ef_2$,

ii) $h^3 f_0''' \approx Af_{-2} + Bf_{-1} + Cf_0 + Df_1 + Ef_2$.

19. Αν $x_{\pm i} = x_0 \pm ih$, $i = 0(1)3$, να προσδιορισθούν οι συντελεστές $\alpha_{\pm i}$, $i = 0(1)3$, έτσι ώστε οι παρακάτω τύποι να είναι όσον το δυ-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

νατόν περισσότερο ακριβείς

i) $f'(x_2) \approx \alpha_{-1} f_{-1} + \alpha_1 f_1 + \alpha_2 f_2$

ii) $f'(x_0 + \frac{h}{2}) \approx \alpha_0 f_0 + \alpha_1 f_1 + \alpha_2 f_2$

iii) $f'(x_2) \approx \alpha_{-1} f_{-1} + \alpha_0 f_0 + \alpha_1 f_1 + \alpha_2 f_2$

iv) $hf'_0 \approx \alpha_{-3} f_{-3} + \alpha_0 f_0 + \alpha_2 f_2$.

20. Με την μέθοδο των προσδιοριστέων συντελεστών να βρεθούν τα α, b, A στον παρακάτω τύπο

$$hf'_0 \approx f(x_0 + \alpha h) + Af(x_0 + bh) .$$

21. Αν η $f(x)$ είναι πολυώνυμο τρίτου βαθμού και δίνεται ο πίνακας τιμών:

x	-2	-1	1
$f'(x)$	1	0	2

,

να βρεθεί η $f'(0)$ με τη βοήθεια πίνακα διαφορών. Στη συνέχεια να βρεθούν τα τοπικά ακρότατα της $f(x)$, δοθέντος ότι $f(0) = f'(0)$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

4

ΣΥΝΟΛΑ ΟΡΘΟΓΩΝΙΩΝ ΠΟΛΥΩΝΥΜΩΝ

4.1 Εισαγωγή

Στο κεφάλαιο αυτό περιέχονται ορισμένα βασικά συμπεράσματα από την θεωρία των ορθογωνίων πολυωνύμων, τα οποία είναι απαραίτητα για την ανάπτυξη της θεωρίας επομένων κεφαλαίων και ειδικά της αριθμητικής ολοκληρώσεως και της πολυωνυμικής προσεγγίσεως.

Έστω $[a,b]$ ένα οποιοδήποτε πεπερασμένο ή απέραντο διάστημα και $w(x)$ μία δοσμένη συνάρτηση, η οποία θα καλείται συνάρτηση βάρους, τέτοια ώστε: i) $w(x) \geqq 0$ στο $[a,b]$ και ii) τα ολοκληρώματα $\int_a^b w(x)x^K dx$, $\kappa = 0,1,2,\ldots$ υπάρχουν και είναι πεπερασμένα με $\int_a^b w(x)dx > 0$. Έστω $L_2[a,b]$ το σύνολο όλων των πραγματικών συναρτήσεων $f(x)$ για τις οποίες το ολοκλήρωμα $\int_a^b w(x)[f(x)]^2 dx$ υπάρχει και είναι πεπερασμένο. Στον γραμμικό χώρο $L_2[a,b]$ ορίζουμε το εσωτερικό γινόμενο (f,g) δύο συναρτήσεων $f(x)$ και $g(x)$, σε σχέση με την συνάρτηση βάρους $w(x)$ ως εξής:

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$(f,g) = \int_a^b w(x)f(x)g(x)dx \quad . \tag{4.1}$$

Λέμε ότι οι συναρτήσεις $f(x)$ και $g(x)$ $(f(x),g(x) \not\equiv 0)$ είναι ορθογώνιες σε σχέση με την συνάρτηση βάρους $w(x)$ στο διάστημα $[a,b]$ εάν $(f,g) = 0$. Ειδικά, αν $w(x) \equiv 1$, τότε απλώς λέμε ότι οι $f(x)$ και $g(x)$ είναι ορθογώνιες στο $[a,b]$.

<u>Ορισμός 4.1</u> Ένα σύνολο πολυωνύμων $\{\varphi_0(x), \varphi_1(x), \varphi_2(x),...\}$, όπου το $\varphi_\kappa(x)$ είναι ένα πολυώνυμο βαθμού κ , καλείται σύνολο ή σύστημα ορθογωνίων πολυωνύμων σε σχέση με την συνάρτηση βάρους $w(x)$ στο διάστημα $[a,b]$ εάν

$$(\varphi_\kappa,\varphi_j) = \int_a^b w(x)\varphi_\kappa(x)\varphi_j(x)dx = 0 \qquad \text{για} \quad \kappa \neq j \quad . \tag{4.2}$$

Το παραπάνω σύνολο πολυωνύμων λέγεται ότι είναι ορθοκανονικό εάν

$$(\varphi_\kappa,\varphi_j) = \delta_{\kappa j} = \begin{cases} 0 \ , & \kappa \neq j \\ 1 \ , & \kappa = j \ . \end{cases} \tag{4.3}$$

▲

Εάν $\{\varphi_\kappa(x)\}$, $\kappa = 0,1,2,...$ είναι ένα σύνολο ορθογωνίων πολυωνύμων για την $w(x)$ στο $[a,b]$, τότε το σύνολο $\{\varphi^*_\kappa(x)\}$, $\kappa = 0,1,2,...$, όπου

$$\varphi^*_\kappa(x) = \frac{\varphi_\kappa(x)}{(\varphi_\kappa,\varphi_\kappa)^{\frac{1}{2}}}$$

είναι ορθοκανονικό. Σημειώνεται ότι $(\varphi_\kappa,\varphi_\kappa) > 0$, $\kappa = 0,1,2,...$, αφού $\varphi_\kappa(x) \not\equiv 0$, $\kappa = 0,1,2,...$.

Για την κατασκευή συνόλων ορθογωνίων πολυωνύμων (κάθε σύνολο σχετίζεται με κάποιο διάστημα $[a,b]$ και κάποια $w(x)$) χρησιμοποιείται η ακόλουθη αναδρομική σχέση σύμφωνα με την Gram – Schmidt διαδικασία ορθογωνιοποιήσεως:

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\begin{cases} \varphi_0(x) \equiv 1 \ , \\ \varphi_1(x) \equiv x - \dfrac{(x\varphi_0,\varphi_0)}{(\varphi_0,\varphi_0)}\varphi_0(x) = x - \dfrac{(x,1)}{(1,1)} \ , \\ \varphi_{K+1}(x) \equiv x\varphi_K(x) - \alpha_K\varphi_K(x) - \beta_K\varphi_{K-1}(x) \ , \ K = 1,2,3.... \ , \end{cases} \tag{4.4}$$

όπου

$$\begin{cases} \alpha_K = \dfrac{(x\varphi_K,\varphi_K)}{(\varphi_K,\varphi_K)} \ , \\ \beta_K = \dfrac{(\varphi_K,\varphi_K)}{(\varphi_{K-1},\varphi_{K-1})} \ . \end{cases} \tag{4.5}$$

Είναι φανερό ότι τα ορθογώνια πολυώνυμα που προκύπτουν με την χρησιμοποίηση των (4.4) και (4.5) για κάποιο διάστημα $[a,b]$ και την συνάρτηση βάρους $w(x)$ έχουν συντελεστή μεγιστοβαθμίου όρου την μονάδα. Σημειώνεται ακόμη, χωρίς απόδειξη, ότι αυτά είναι μονοσήμαντα ορισμένα. Γενικώτερα, μπορεί να αποδειχθεί ότι, εάν $[a,b]$ είναι ένα οποιοδήποτε δοσμένο διάστημα και $w(x)$ μία συνάρτηση βάρους, τότε, εκτός από σταθερούς παράγοντες, υπάρχει ένα μοναδικό σύνολο ορθογωνίων πολυωνύμων σε σχέση με την $w(x)$ στο $[a,b]$.

Για παράδειγμα, με $w(x) \equiv 1$ και $[a,b] \equiv [-1,1]$ από τις (4.4) και (4.5) έχουμε

$$\varphi_0(x) = 1 \ , \ \varphi_1(x) = x - \frac{\int_{-1}^{1} x dx}{\int_{-1}^{1} dx} = x - 0 = x \ ,$$

$$\varphi_2(x) = x\varphi_1(x) - \alpha_1\varphi_1(x) - \beta_1\varphi_0(x) = x^2 - \alpha_1 x - \beta_1 \ , \text{ όπου}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\alpha_1 = \frac{\int_{-1}^{1} x^3 dx}{\int_{-1}^{1} x^2 dx} = 0 \quad \text{και} \quad \beta_1 = \frac{\int_{-1}^{1} x^2 dx}{\int_{-1}^{1} dx} = \frac{\frac{2}{3}}{2} = \frac{1}{3} \quad .$$

Άρα $\varphi_2(x) = x^2 - \frac{1}{3}$. Ομοίως βρίσκουμε $\varphi_3(x) = x^3 - \frac{3}{5}x$, κ.ο.κ. .

Πολλές φορές, αντί της ακολουθίας ορθογωνίων πολυωνύμων $\{\varphi_\kappa(x)\}$, $\kappa = 0,1,2,\ldots$, που προκύπτει με βάση τις (4.4), (4.5) και στην οποία το $\varphi_\kappa(x)$ έχει συντελεστή του x^κ την μόνάδα, χρησιμοποιείται η $\{c_\kappa\varphi_\kappa(x)\}$, $\kappa = 0,1,2,\ldots$, όπου c_κ, $\kappa = 0,1,2,\ldots$ είναι σταθερές. Με την έννοια αυτή από τα ορθογώνια πολυώνυμα που βρέθηκαν προηγουμένως για $w(x) \equiv 1$ και $[a,b] \equiv [-1,1]$ προκύπτουν τα πολυώνυμα Legendre. Άλλες ακολουθίες γνωστών ορθογωνίων πολυωνύμων προκύπτουν από τις περιπτώσεις: i) $w(x) \equiv (1-x^2)^{-\frac{1}{2}}$ και $[a,b] \equiv [-1,1]$ (πολυώνυμα Chebyshev), ii) $w(x) \equiv e^{-x}$ και $[a,b] \equiv [0,+\infty]$ (πολυώνυμα Laguerre), iii) $w(x) = e^{-x^2}$ και $[a,b] \equiv [-\infty,+\infty]$ (πολυώνυμα Hermite) κ.λ.π. .

4.2 Ιδιότητες συνόλων ορθογωνίων πολυωνύμων

Έστω $\{\varphi_\kappa(x)\}$, $\kappa = 0,1,2,\ldots$ ένα σύνολο ορθογωνίων πολυωνύμων (στο εξής υποθέτουμε ότι ο συντελεστής του x^κ στο $\varphi_\kappa(x)$ δεν είναι αναγκαστικά η μονάδα) σε σχέση με την συνάρτηση βάρους $w(x)$ στο διάστημα $[a,b]$ και έστω P_n το σύνολο όλων των πολυωνύμων βαθμού $\leq n$. Τότε ισχύουν τα ακόλουθα:

<u>Θεώρημα 4.1</u> Κάθε πολυώνυμο $p = p(x) \in P_n$ μπορεί να εκφραστεί μονοσήμαντα σαν γραμμικός συνδυασμός των ορθογωνίων πολυωνύμων

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$\varphi_j(x)$, $j = 0(1)n$.

Απόδειξη Επειδή το σύνολο $\{\varphi_\kappa(x)\}$, $\kappa = 0(1)n$ είναι γραμμικώς ανεξάρτητο (βλέπε άσκηση 2), αποτελεί μία βάση για τον γραμμικό χώρο P_n . Έτσι, κάθε $p(x) \in P_n$ μπορεί να εκφραστεί σαν γραμμικός συνδυασμός των $\varphi_j(x)$, $j = 0(1)n$, δηλαδή ,

$$p(x) = \sum_{j=0}^{n} d_j \varphi_j(x) \ , \tag{4.6}$$

όπου οι συντελεστές d_j προσδιορίζονται μονοσήμαντα από τις

$$d_j = \frac{(p\ ,\varphi_j)}{(\varphi_j,\varphi_j)} \quad , \ j = 0(1)n \ . \qquad \blacktriangle \tag{4.7}$$

Πόρισμα 4.1 Για κάθε $p \in P_{n-1}$ ισχύει $(p,\varphi_n) = 0$ $(n \geq 1)$.

Πόρισμα 4.2 Αν $R_n(x)$ είναι ένα πολυώνυμο n βαθμού και για κάθε $p \in P_{n-1}$ ισχύει $(p,R_n) \equiv \int_a^b w(x)p(x)R_n(x)dx = 0$, τότε $R_n(x) \equiv$ $\equiv c_n\varphi_n(x)$, όπου c_n σταθερός παράγων.

Απόδειξη Σύμφωνα με το θεώρημα 4.1 είναι $R_n(x) = \sum_{i=0}^{n} c_i\varphi_i(x)$.

Επειδή η $(p,R_n) = 0$ ισχύει για κάθε $p \in P_{n-1}$ θα ισχύει και για $p = \varphi_0(x),\varphi_1(x),\ldots,\varphi_{n-1}(x)$. Έτσι θα έχουμε $(\varphi_j,R_n) = 0$, $j = 0(1)n-1$ ή

$$\int_a^b w(x)\varphi_j(x)\left(\sum_{i=0}^{n} c_i\varphi_i(x)\right)dx = c_j\int_a^b w(x)\left[\varphi_j(x)\right]^2 dx = 0 \ , \ j = 0(1)n-1.$$

Επειδή $\int_a^b w(x)\left[\varphi_j(x)\right]^2 dx \neq 0$ (αφού $\int_a^b w(x)dx > 0$), προκύπτει ότι $c_j = 0$, $j = 0(1)n-1$ και επομένως $R_n(x) \equiv c_n\varphi_n(x)$. ▲

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

<u>Θεώρημα 4.2</u> Το πολυώνυμο $\varphi_n(x)$ έχει ακριβώς n διακεκριμένες ρίζες στο διάστημα (a,b) .

<u>Απόδειξη</u> Επειδή $\int_a^b w(x)\varphi_n(x)dx = 0$, $n \geq 1$ (βλέπε πόρισμα 4.1) συμπεραίνουμε ότι το πολυώνυμο $\varphi_n(x)$ αλλάζει πρόσημο στο $[a,b]$ τουλάχιστον μία φορά. Υποθέτουμε ότι το $\varphi_n(x)$ έχει στο (a,b) μόνον $\ell < n$ ρίζες περιττής πολλαπλότητος τις $x_1, x_2, \ldots, x_\ell$, δηλαδή, ρίζες στις οποίες το $\varphi_n(x)$ αλλάζει πρόσημο. Τότε το πολυώνυμο $\varphi_n(x)\prod_{j=1}^{\ell}(x-x_j)$ δεν αλλάζει πρόσημο στο $[a,b]$ και επομένως έχουμε

$$\int_a^b w(x)\varphi_n(x)\prod_{j=1}^{\ell}(x-x_j)dx \neq 0 ,$$

πράγμα που είναι άτοπο αφού το $\varphi_n(x)$ είναι ορθογώνιο προς κάθε πολυώνυμο μικροτέρου βαθμού, άρα και προς το $\prod_{j=1}^{\ell}(x-x_j)$. Επομένως, είναι $\ell = n$ που σημαίνει ότι το $\varphi_n(x)$ έχει n διακεκριμένες ρίζες στο (a,b) . ▲

4.3 Τα πολυώνυμα Legendre

Τα πολυώνυμα Legendre ορίζονται από τον τύπο

$$\Phi_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n , \quad n = 0,1,2,\ldots \tag{4.8}$$

και όπως θα δειχθεί (θεώρημα 4.4) αποτελούν ένα σύνολο ορθογωνίων πολυωνύμων σε σχέση με την συνάρτηση βάρους $w(x) \equiv 1$ στο διάστημα $[-1,1]$. Στη συνέχεια αναφέρονται υπό μορφή θεωρημάτων οι πιο βασικές ιδιότητες των πολυωνύμων Legendre.

<u>Θεώρημα 4.3</u> Αν $g(x)$ είναι οποιαδήποτε n-φορές συνεχώς παραγω-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

γίσιμη συνάρτηση στο $[-1,1]$ τότε

$$\int_{-1}^{1}\Phi_n(x)g(x)dx = (-1)^n \frac{1}{2^n n!}\int_{-1}^{1}(x^2-1)^n g^{(n)}(x)dx \ . \qquad (4.9)$$

<u>Απόδειξη</u> Θέτουμε $\Phi_n(x) = \frac{1}{2^n n!} R^{(n)}(x)$, όπου $R(x) = (x^2-1)^n$. Προφανώς ισχύει $R^{(\kappa)}(\pm 1) = 0$, $\kappa = 0(1)n-1$. Ολοκληρώνοντας κατά παράγοντας έχουμε

$$\int_{-1}^{1}\Phi_n(x)g(x)dx = \frac{1}{2^n n!}\int_{-1}^{1}R^{(n)}(x)g(x)dx =$$

$$= \frac{1}{2^n n!}\left[g(x)R^{(n-1)}(x)\right]_{-1}^{1} - \frac{1}{2^n n!}\int_{-1}^{1}g'(x)R^{(n-1)}(x)dx =$$

$$= -\frac{1}{2^n n!}\int_{-1}^{1}g'(x)R^{(n-1)}(x)dx = \ldots = (-1)^n \frac{1}{2^n n!}\int_{-1}^{1}g^{(n)}(x)R(x)dx \ . \ \blacktriangle$$

<u>Θεώρημα 4.4</u>

$$\int_{-1}^{1}\Phi_\kappa(x)\Phi_j(x)dx = 0 \quad \text{για} \quad \kappa \neq j \ .$$

<u>Απόδειξη</u> Μπορούμε να υποθέσουμε $\kappa > j$ οπότε η απόδειξη προκύπτει από το θεώρημα 4.3 αν ληφθεί $g(x) = \Phi_j(x)$. ▲

<u>Θεώρημα 4.5</u> Για κάθε πολυώνυμο $p(x)$ βαθμού $< n$ ισχύει

$$\int_{-1}^{1}p(x)\Phi_n(x)dx = 0 \ , \quad n \geq 1 \ .$$

<u>Απόδειξη</u> Προκύπτει από το θεώρημα 4.3 αν $g(x) = p(x)$ (βλέπε επίσης πόρισμα 4.1) . ▲

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Πόρισμα 4.3

$$\int_{-1}^{1} \Phi_n(x)dx = 0 \ , \quad n \geq 1 \quad .$$

Θεώρημα 4.6 Το πολυώνυμο $\Phi_n(x)$ έχει n διακεκριμένες ρίζες στο διάστημα $(-1,1)$.

Θεώρημα 4.7 Ο συντελεστής του μεγιστοβαθμίου όρου του πολυωνύμου $\Phi_n(x)$ του Legendre ισούται με $\dfrac{(2n)!}{2^n(n!)^2} = \dfrac{(2n-1)!!}{n!}$ (όπου $(2n-1)!! = (2n-1)(2n-3)(2n-5)\ldots 5\cdot 3\cdot 1$, για $n \geq 1$) .

Απόδειξη Έχουμε

$$\Phi_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n}(x^2-1)^n = \frac{1}{2^n n!} \frac{d^n}{dx^n}\left\{ \sum_{r=0}^{n} (-1)^r \binom{n}{r} x^{2n-2r} \right\} =$$

$$= \frac{1}{2^n n!} \sum_{r=0}^{\left[\frac{n}{2}\right]} (-1)^r \binom{n}{r} \frac{(2n-2r)!}{(n-2r)!} x^{n-2r} =$$

$$= \sum_{r=0}^{\left[\frac{n}{2}\right]} (-1)^r \frac{(2n-2r)!}{2^n r!(n-r)!(n-2r)!} x^{n-2r} \ , \qquad (4.10)$$

όπου

$$\left[\frac{n}{2}\right] = \begin{cases} \dfrac{n}{2} & \text{αν} \quad n = 2\lambda \quad \text{(άρτιος)} \\ \dfrac{n-1}{2} & \text{αν} \quad n = 2\lambda+1 \quad \text{(περιττός)} \end{cases} \quad .$$

Από την παραπάνω έκφραση για το $\Phi_n(x)$ προκύπτει ότι

$$\Phi_n(x) = (-1)^0 \frac{(2n)!}{2^n 0!n!n!} x^n + S_{n-2}(x) = \frac{(2n)!}{2^n(n!)^2} x^n + S_{n-2}(x) \ ,$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

όπου $S_{n-2}(x)$ πολυώνυμο βαθμού $n-2$. ▲

<u>Πόρισμα 4.4</u> Τα πολυώνυμα που προκύπτουν από τις (4.4) και (4.5) για $[a,b] = [-1,1]$ και $w(x) \equiv 1$ δίνονται από τον τύπο

$$\varphi_n(x) = \frac{n!}{(2n)!} \frac{d^n}{dx^n}(x^2-1)^n \ , \quad n = 0,1,2,\ldots \ . \tag{4.11}$$

<u>Θεώρημα 4.8</u>

$$\int_{-1}^{1} \left[\Phi_n(x)\right]^2 dx = \frac{2}{2n+1} \ .$$

<u>Απόδειξη</u> Από το θεώρημα 4.3 για $g(x) = \Phi_n(x)$ και επειδή

$$\Phi_n^{(n)}(x) = \frac{(2n)!}{2^n(n!)^2} n! = \frac{(2n)!}{2^n n!} = (2n-1)!!$$ (λόγω του θεωρήματος 4.7), έχουμε:

$$\int_{-1}^{1} \left[\Phi_n(x)\right]^2 dx = (-1)^n \frac{1}{2^n n!} \frac{(2n)!}{2^n n!} \int_{-1}^{1} (x^2-1)^n dx =$$

$$= 2 \frac{(2n)!}{2^{2n}(n!)^2} \int_0^1 (1-x^2)^n dx$$. Μπορεί να δειχθεί ότι

$$\int_0^1 (1-x^2)^n dx = \frac{2^{2n}(n!)^2}{(2n+1)!} = \frac{(2n)!!}{(2n+1)!!}$$, οπότε προκύπτει

$$\int_{-1}^{1} \left[\Phi_n(x)\right]^2 dx = 2 \frac{(2n)!}{2^{2n}(n!)^2} \frac{2^{2n}(n!)^2}{(2n+1)!} = \frac{2}{2n+1} \ .$$ ▲

<u>Θεώρημα 4.9</u>

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\int_{-1}^{1} x^n \Phi_n(x)dx = \frac{2^{n+1}(n!)^2}{(2n+1)!} \quad .$$

Απόδειξη Προκύπτει από το θεώρημα 4.3 με $g(x) = x^n$ (βλέπε και απόδειξη του θεωρήματος 4.8).

Θεώρημα 4.10 Όλα τα πολυώνυμα $\Phi_{2K}(x)$ είναι άρτιες συναρτήσεις και τα $\Phi_{2K+1}(x)$ είναι περιττές, δηλαδή ισχύει

$$\Phi_j(-x) = (-1)^j \Phi_j(x) \quad , \; j = 0,1,2,\ldots \quad . \tag{4.12}$$

Απόδειξη Προκύπτει από την έκφραση (4.10) αφού τα $\Phi_{2K}(x)$ περιέχουν μόνον άρτιες δυνάμεις του x ενώ τα $\Phi_{2K+1}(x)$ περιέχουν μόνον περιττές δυνάμεις του x . ▲

Θεώρημα 4.11 Ισχύει η αναδρομική σχέση

$$(n+1)\Phi_{n+1}(x) - (2n+1)x\Phi_n(x) + n\Phi_{n-1}(x) = 0 \quad , \quad n = 1,2,\ldots \quad .$$

Απόδειξη Σύμφωνα με το θεώρημα 4.1 έχουμε για το πολυώνυμο $x\Phi_n(x)$

$$x\Phi_n(x) = \sum_{i=0}^{n+1} c_i \Phi_i(x) \quad , \tag{4.13}$$

από την οποία για $K \leq n+1$ παίρνουμε

$$\int_{-1}^{1} x\Phi_n(x)\Phi_K(x)dx = \int_{-1}^{1} \sum_{i=0}^{n+1} c_i \Phi_i(x)\Phi_K(x)dx = \sum_{i=0}^{n+1} c_i \int_{-1}^{1} \Phi_i(x)\Phi_K(x)dx =$$

$$= c_K \int_{-1}^{1} \Phi_K^2(x)dx \quad \text{ή} \quad c_K = \frac{\int_{-1}^{1} x\Phi_n(x)\Phi_K(x)dx}{\int_{-1}^{1} \Phi_K^2(x)dx} \quad , \; K = 0(1)n+1 . \tag{4.14}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Είναι φανερό ότι για $\kappa \leq n-2$ το πολυώνυμο $x\Phi_\kappa(x)$ είναι βαθμού $\kappa+1 \leq n-1$ και λόγω του θεωρήματος 4.5 από την (4.14) προκύπτει ότι

$$c_\kappa = 0 \ , \quad \kappa = 0(1)n-2 \ . \tag{4.15}$$

Για $\kappa = n$ από την (4.14) έχουμε

$$c_n = \frac{\int_{-1}^{1} x\Phi_n^2(x)dx}{\int_{-1}^{1} \Phi_n^2(x)dx} = 0 \ , \tag{4.16}$$

αφού η συνάρτηση $x\Phi_n^2(x)$ είναι περιττή λόγω του θεωρήματος 4.10 . Για $\kappa = n-1$ η (4.14) δίνει

$$c_{n-1} = \frac{\int_{-1}^{1} x\Phi_n(x)\Phi_{n-1}(x)dx}{\int_{-1}^{1} \Phi_{n-1}^2(x)dx} \ . \tag{4.17}$$

Αν d_{n-1} είναι ο συντελεστής του μεγιστοβαθμίου όρου του $\Phi_{n-1}(x)$, τότε είναι $x\Phi_{n-1}(x) = d_{n-1}x^n + S_{n-2}(x)$, όπου $S_{n-2}(x)$ είναι πολυώνυμο βαθμού $n-2$. Έτσι έχοντας υπ' όψιν τα θεωρήματα 4.5, 4.7, 4.8 και (4.9), ή (4.17) γίνεται

$$c_{n-1} = \frac{d_{n-1}\int_{-1}^{1} x^n\Phi_n(x)dx}{\int_{-1}^{1} \Phi_{n-1}^2(x)dx} = \frac{\dfrac{[2(n-1)]!}{2^{n-1}[(n-1)!]^2}\,\dfrac{2^{n+1}(n!)^2}{(2n+1)!}}{\dfrac{2}{2(n-1)+1}} = \frac{n}{2n+1} \ . \tag{4.18}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Τέλος, για $\kappa = n+1$ είναι

$$c_{n+1} = \frac{\int_{-1}^{1} x\Phi_n(x)\Phi_{n+1}(x)dx}{\int_{-1}^{1} \Phi^2_{n+1}(x)dx} \quad . \tag{4.19}$$

Ακολουθώντας την ίδια διαδικασία που χρησιμοποιήθηκε για τον υπολογισμό του c_{n-1} βρίσκουμε ότι

$$c_{n+1} = \frac{n+1}{2n+1} \quad . \tag{4.20}$$

Λόγω των (4.15), (4.16), (4.18) και (4.20) από την (4.13) προκύπτει η ζητούμενη αναδρομική σχέση. ▲

Με την βοήθεια της αναδρομικής σχέσεως του θεωρήματος 4.11 μπορούν να βρεθούν εύκολα τα πολυώνυμα Legendre έχοντας υπ' όψιν ότι $\Phi_0(x) \equiv 1$, $\Phi_1(x) \equiv x$. Έτσι τα πρώτα 6 πολυώνυμα Legendre είναι:

$$\begin{aligned}
\Phi_0(x) &= 1 \\
\Phi_1(x) &= x \\
\Phi_2(x) &= \frac{1}{2}(3x^2-1) \\
\Phi_3(x) &= \frac{1}{2}(5x^3-3x) \\
\Phi_4(x) &= \frac{1}{8}(35x^4-30x^2+3) \\
\Phi_5(x) &= \frac{1}{8}(63x^5-70x^3+15x) \quad .
\end{aligned} \tag{4.21}$$

4.4 Τα πολυώνυμα Chebyshev

Ένα σύνολο ορθογωνίων πολυωνύμων σε σχέση με την συνάρτηση βά-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ρους $w(x) = (1-x^2)^{-\frac{1}{2}}$ στο διάστημα $[-1,1]$ αποτελούν τα πολυώνυμα Chebyshev που ορίζονται από την σχέση

$$T_n(x) = \cos(n \arccos x) \quad , \quad n = 0,1,2,\ldots \quad . \tag{4.22}$$

Προφανώς είναι $T_0(x) = 1$, $T_1(x) = x$ και $|T_n(x)| \leq 1$ για $|x| \leq 1$.

Για τα πολυώνυμα Chebyshev ισχύουν τα ακόλουθα θεωρήματα:

<u>Θεώρημα 4.12</u>

$$\int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} T_\kappa(x) T_j(x) dx = 0 \quad \text{για} \quad \kappa \neq j \quad .$$

<u>Απόδειξη</u> Θέτοντας $\arccos x = \theta$, $0 \leq \theta \leq \pi$ έχουμε $x = \cos\theta$, $dx = -\sin\theta d\theta$ και $(1-x^2)^{-\frac{1}{2}} = \frac{1}{\sin\theta}$, οπότε

$$\int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} T_\kappa(x) T_j(x) dx = \int_0^\pi \cos\kappa\theta \cos j\theta d\theta = 0 \ , \ \text{για} \quad \kappa \neq j \ .$$

<u>Θεώρημα 4.13</u> Για κάθε πολυώνυμο $p(x)$ βαθμού $\leq n-1$ ισχύει

$$\int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} p(x) T_n(x) dx = 0 \quad , \quad n \geq 1 \quad .$$

<u>Θεώρημα 4.14</u> Το πολυώνυμο $T_n(x)$ έχει ακριβώς n διακεκριμένες ρίζες στο διάστημα $(-1,1)$ οι οποίες δίνονται από τον τύπο

$$x_\kappa = \cos\left(\frac{(2\kappa+1)\pi}{2n}\right) \quad , \quad \kappa = 0(1)n-1 \ , \quad n \geq 1 \ . \tag{4.23}$$

<u>Απόδειξη</u> Από την εξίσωση $T_n(x) = \cos(n \arccos x) = 0$, θέτοντας $\arccos x = \theta$, $0 \leq \theta \leq \pi$ προκύπτει $\theta = \frac{(2\kappa+1)\pi}{2n}$, κ ακέραιος με

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$0 \leq \dfrac{(2\kappa+1)\pi}{2n} \leq \pi$, δηλαδή, $\kappa = 0(1)n-1$. Έτσι οι ρίζες του $T_n(x)$ είναι

$$x_\kappa = \cos\left(\frac{(2\kappa+1)\pi}{2n}\right) , \quad \kappa = 0(1)n-1 , \quad n \geq 1 . \qquad (4.24)$$

Αυτές είναι διακεκριμένες και ανήκουν στο $(-1,1)$. ▲

<u>Θεώρημα 4.15</u> Για τα πολυώνυμα Chebyshev ισχύει η αναδρομική σχέση:

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x) \quad , \quad n = 1,2,\ldots \ .$$

<u>Απόδειξη</u> Από την τριγωνομετρική ταυτότητα

$$\cos\left((n+1)\theta\right) + \cos\left((n-1)\theta\right) = 2\cos\theta \cdot \cos n\theta \quad \text{με} \quad \theta = \arccos x$$

προκύπτει $T_{n+1}(x) + T_{n-1}(x) = 2xT_n(x)$, δηλαδή η ζητούμενη σχέση. ▲

Με την αναδρομική σχέση του θεωρήματος 4.15 και γνωρίζοντας ότι $T_0(x) = 1$, $T_1(x) = x$, μπορούμε να βρούμε διαδοχικά τα πολυώνυμα Chebyshev. Τα πρώτα δέκα από αυτά είναι

$$\begin{aligned}
T_0(x) &= 1 \quad , \quad T_1(x) = x \\
T_2(x) &= 2x^2 - 1 \quad , \quad T_3(x) = 4x^3 - 3x \\
T_4(x) &= 8x^4 - 8x^2 + 1 \\
T_5(x) &= 16x^5 - 20x^3 + 5x \\
T_6(x) &= 32x^6 - 48x^4 + 18x^2 - 1 \\
T_7(x) &= 64x^7 - 112x^5 + 56x^3 - 7x \\
T_8(x) &= 128x^8 - 256x^6 + 160x^4 - 32x^2 + 1 \\
T_9(x) &= 256x^9 - 576x^7 + 432x^5 - 120x^3 + 9x \ .
\end{aligned} \qquad (4.25)$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Θεώρημα 4.16 Ο συντελεστής του μεγιστοβαθμίου όρου του πολυωνύμου $T_n(x)$ του Chebyshev ισούται με 2^{n-1} $(n \geq 1)$.

Απόδειξη Από το θεώρημα 4.15 προκύπτει ότι ο μεγιστοβάθμιος όρος του $T_{n+1}(x)$ βρίσκεται αν πολλαπλασιάσουμε τον μεγιστοβάθμιο όρο του $T_n(x)$ επί $2x$. Επειδή $T_1(x) = x = 2^0x$, ο μεγιστοβάθμιος όρος του $T_2(x)$ θα είναι 2^1x^2, του $T_3(x)$ θα είναι 2^2x^3 κ.ο.κ. και του $T_n(x)$ θα είναι $2^{n-1}x^n$, $n \geq 1$.

Θεώρημα 4.17 Όλα τα πολυώνυμα $T_{2K}(x)$ είναι άρτιες συναρτήσεις και όλα τα $T_{2K+1}(x)$ είναι περιττές συναρτήσεις, δηλαδή ισχύει

$$T_j(-x) = (-1)^j T_j(x) \ , \quad j = 0,1,2,\ldots \quad . \tag{4.26}$$

Απόδειξη Για $j = 0,1$ η (4.26) είναι προφανής. Υποθέτουμε ότι η (4.26) ισχύει για $j \leq n$ $(n \geq 1)$. Λόγω του θεωρήματος 4.15 έχουμε

$$T_{n+1}(-x) = -2xT_n(-x) - T_{n-1}(-x) = (-1)^{n+1}2xT_n(x) -$$
$$-(-1)^{n-1}T_{n-1}(x) = (-1)^{n+1}\left[2xT_n(x) - T_{n-1}(x)\right] =$$
$$= (-1)^{n+1}T_{n+1}(x) \ .$$

Έτσι, απεδείχθη η (4.26) για $j = n+1$ και επομένως για $j = 0,1,2,\ldots$.

Θεώρημα 4.18

$$\int_{-1}^{1}(1-x^2)^{-\frac{1}{2}}\, T_n^2(x)dx = \begin{cases} \pi & \text{για } n = 0 \\ \frac{\pi}{2} & \text{για } n > 0 \end{cases} \ .$$

Απόδειξη Ακολουθώντας την απόδειξη του θεωρήματος 4.12 βρίσκουμε ότι

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\int_{-1}^{1}(1-x^2)^{-\frac{1}{2}}\,T_n^2(x)dx = \int_0^{\pi}\cos^2 n\theta d\theta = \begin{cases} \pi & \text{για } n = 0 \\ \frac{\pi}{2} & \text{για } n > 0 \end{cases} .$$

Θεώρημα 4.19 Οι ακρότατες τιμές του $T_n(x)$ $(n \geq 1)$ στο διάστημα $[-1,1]$ λαμβάνονται στα σημεία

$$x_\kappa = \cos\left(\frac{\kappa\pi}{n}\right) , \quad \kappa = 0(1)n \tag{4.27}$$

για τα οποία ισχύει $T_n(x_\kappa) = (-1)^\kappa$.

Απόδειξη Τα σημεία του $[-1,1]$ στα οποία το $T_n(x)$ παίρνει ακρότατη τιμή προκύπτουν από την λύση της εξισώσεως $|T_n(x)| = 1$, δηλαδή των

$$\cos(n \arccos x) = \pm 1 . \tag{4.28}$$

Θέτοντας $\arccos x = \theta$, $0 \leq \theta \leq \pi$, οπότε $x = \cos\theta$, η (4.28) γίνεται $\cos n\theta = \pm 1$ με λύσεις $\theta_\kappa = \frac{\kappa\pi}{n}$, κ ακέραιος και $0 \leq \frac{\kappa\pi}{n} \leq \pi$. Άρα οι λύσεις των (4.28) είναι

$$x_\kappa = \cos\theta_\kappa = \cos\left(\frac{\kappa\pi}{n}\right) , \quad \kappa = 0(1)n .$$

Επί πλέον έχουμε $T_n(x_\kappa) = \cos(n\theta_\kappa) = \cos(\kappa\pi) = (-1)^\kappa$, $\kappa = 0(1)n$.

Θεώρημα 4.20 Εάν $P_n(x)$ είναι ένα οποιοδήποτε πολυώνυμο βαθμού $n \geq 1$ με συντελεστή του x^n την μονάδα, τότε ισχύει

$$\max_{x \in [-1,1]} \left|P_n(x)\right| \geq \max_{x \in [-1,1]} \left|2^{1-n}T_n(x)\right| = 2^{1-n} . \tag{4.29}$$

Απόδειξη Ας υποθέσουμε ότι ισχύει η αντίθετη της (4.29), δηλαδή

$$\max_{x \in [-1,1]} \left|P_n(x)\right| < 2^{1-n} \tag{4.30}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

και επομένως $\left|P_n(x)\right| < 2^{1-n}$ για κάθε $x \in [-1,1]$. Επειδή το πολυώνυμο $2^{1-n}T_n(x)$ έχει συντελεστή του x^n την μονάδα προκύπτει ότι το πολυώνυμο $R(x) = 2^{1-n}T_n(x) - P_n(x)$ είναι βαθμού $\leq n-1$ και για τις τιμές (4.27) ισχύει

$$\operatorname{sign} R(x_\kappa) = \operatorname{sign}\left[2^{1-n}T_n(x_\kappa) - P_n(x_\kappa)\right] =$$

$$= \operatorname{sign}\left[2^{1-n}(-1)^\kappa - P_n(x_\kappa)\right] , \quad \kappa = 0(1)n . \tag{4.31}$$

Επειδή $\left|P_n(x_\kappa)\right| < 2^{1-n}$, δηλαδή, $-2^{1-n} < P_n(x_\kappa) < 2^{1-n}$ από την (4.31) συμπεραίνουμε ότι

$$\operatorname{sign} R(x_\kappa) = (-1)^\kappa , \quad \kappa = 0(1)n . \tag{4.32}$$

Η (4.32) σημαίνει ότι το πολυώνυμο $R(x)$, βαθμού το πολύ $n-1$, με $R(x) \not\equiv 0$ αφού στα σημεία x_κ , $\kappa = 0(1)n$ παίρνει τιμές $\neq 0$, αλλάζει πρόσημο σε κάθε ένα από τα διαστήματα $(x_{\kappa+1}, x_\kappa)$, $\kappa = 0(1)n-1$ και επομένως έχει τουλάχιστον n διακεκριμένες ρίζες, πράγμα που είναι άτοπον. Άρα ισχύει η (4.29). Μπορεί να αποδειχθεί (βλέπε άσκηση) ότι η ισότητα στην (4.29) ισχύει εάν και μόνον εάν $P_n(x) \equiv \equiv \pm 2^{1-n}T_n(x)$.

<u>Πόρισμα 4.5</u> Εάν $P_m(x)$ είναι ένα οποιοδήποτε πολυώνυμο βαθμού m , όπου $1 \leq m < n$ με συντελεστή του x^m την μονάδα, τότε

$$\max_{x \in [-1,1]} \left|P_m(x)\right| > \max_{x \in [-1,1]} \left|2^{1-n}T_n(x)\right| = 2^{1-n} .$$

<u>Απόδειξη</u> Σύμφωνα με το θεώρημα 4.20 έχουμε

$$\max_{x \in [-1,1]} \left|P_m(x)\right| \geq 2^{1-m} > 2^{1-n} = \max_{x \in [-1,1]} \left|2^{1-n}T_n(x)\right| .$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ΑΣΚΗΣΕΙΣ

1. Να βρεθούν τα ορθογώνια πολυώνυμα $P_1(x)$, $P_2(x)$, $P_3(x)$, $P_4(x)$, $P_5(x)$ σε σχέση με την συνάρτηση βάρους $w(x) \equiv x^2$ στο $[-1,1]$. Στη συνέχεια να βρεθούν οι ρίζες αυτών.

2. Να δειχθεί ότι κάθε πεπερασμένο σύνολο ορθογωνίων πολυωνύμων είναι γραμμικώς ανεξάρτητο.

3. Να δειχθεί ότι τα πολυώνυμα Legendre $\Phi_n(x)$, $n = 0,1,2,\ldots$ δίνονται από τις ακόλουθες ισοδύναμες εκφράσεις:

   i) $$\Phi_n(x) = \frac{1}{2^n} \sum_{\kappa=0}^{n} \binom{n}{\kappa}^2 (x-1)^{n-\kappa}(x+1)^{\kappa}$$

   ii) $$\Phi_n(x) = \frac{1}{\pi} \int_0^{\pi} (x + \sqrt{x^2-1}\cos\theta)^n d\theta$$

   iii) $$\Phi_n(x) = \frac{1}{2^n} \sum_{i=0}^{\left[\frac{n}{2}\right]} (-1)^i \binom{n}{i} \binom{2n-2i}{n} x^{n-2i} .$$

4. Να βρεθεί η σχέση που συνδέει τους συντελεστές των μεγιστοβαθμίων όρων δύο διαδοχικών πολυωνύμων $\Phi_n(x)$, $\Phi_{n+1}(x)$ του Legendre.

5. Να δειχθούν οι ακόλουθες ιδιότητες των πολυωνύμων Legendre $\Phi_n(x)$, $n = 0,1,2,\ldots$

   i) $\Phi_n(1) = 1$, $\Phi_n(-1) = (-1)^n$.

   ii) $$\Phi_n(0) = \begin{cases} 0 , & n = 2\kappa+1 \\ (-1)^{\kappa} \dfrac{(2\kappa-1)!!}{(2\kappa)!!} = (-1)^{\kappa} \dfrac{1\cdot 3\cdot 5 \ldots (2\kappa-1)}{2\cdot 4\cdot 6 \ldots (2\kappa)} , & n = 2\kappa . \end{cases}$$

   iii) $(1-x^2)\Phi_n''(x) - 2x\Phi_n'(x) + n(n+1)\Phi_n(x) = 0$.

   iv) $|\Phi_n(x)| \leq 1$ για $|x| \leq 1$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

6. Να βρεθεί το γινόμενο των ριζών του $\Phi_n(x)$.

7. Να δειχθεί ότι για τους συντελεστές α_i , $i = 0(1)n$ του $\Phi_n(x$ ισχύει $\sum_{i=0}^{n} \alpha_i = 1$.

8. Να δειχθούν οι ταυτότητες

i) $\Phi'_n(x) = x\Phi'_{n-1}(x) + n\Phi_{n-1}(x)$

ii) $\Phi'_{n+1}(x) - \Phi'_{n-1}(x) = (2n+1)\Phi_n(x)$

iii) $\int \Phi_n(x)dx = \dfrac{\Phi_{n+1}(x) - \Phi_{n-1}(x)}{2n+1}$

iv) $x\Phi'_n(x) - \Phi'_{n-1}(x) = n\Phi_n(x)$.

9. Να δειχθεί ότι δύο διαδοχικά πολυώνυμα του Legendre δεν μπορούν να έχουν κοινή ρίζα στο $[-1, 1]$.

10. Για τα πολυώνυμα Chebyshev να δειχθούν οι ακόλουθες σχέσεις

i) $T_n(1) = 1$, $T_n(-1) = (-1)^n$

ii) $T_{2K}(0) = (-1)^K$, $T_{2K+1}(0) = 0$

iii) $T_{m+n}(x) + T_{m-n}(x) = 2T_m(x)T_n(x)$

iv) $T_{2m}(x) = 2T_m^2(x) - 1$

v) $T_K(T_m(x)) = T_{Km}(x)$.

11. Να βρεθεί το γινόμενο των ριζών του $T_n(x)$.

12. Να δειχθεί ότι η ελάχιστη τιμή της εκφράσεως $\max\limits_{x\in[-1,1]} |(x-x_0)(x-x_1)\ldots(x-x_n)|$ επιτυγχάνεται όταν τα x_i , $i = 0(1)n$ εκλεγούν να είναι οι ρίζες του πολυωνύμου $T_{n+1}(x)$ του Chebyshev.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

13. Να δειχθεί ότι

i) $(1-x^2)T_n''(x) - xT_n'(x) + n^2T_n(x) = 0$.

ii) $$\int T_n(x)dx = \begin{cases} T_1(x) + c\ , & n = 0 \\ \frac{1}{4}T_2(x) + c\ , & n = 1 \\ \frac{1}{2}\left(\frac{T_{n+1}(x)}{n+1} - \frac{T_{n-1}(x)}{n-1}\right) + c\ , & n > 1\ . \end{cases}$$

14. Αν $P_n(x)$ είναι ένα οποιοδήποτε πολυώνυμο βαθμού n με συντελεστή του μεγιστοβαθμίου όρου τον α_n , τέτοιο ώστε $\max\limits_{x\in[-1,1]} |P_n(x)| \leq 1$, να δειχθεί ότι $|\alpha_n| \leq 2^{n-1}$.

15. Να δειχθεί ότι $T_n'(x) = \dfrac{n\sin(n\arccos x)}{\sin(\arccos x)}$.

16. Να δειχθούν οι ακόλουθες ισοδύναμες εκφράσεις για τα πολυώνυμα Chebyshev:

i) $$T_n(x) = \frac{1}{2}\left[\left(x+\sqrt{x^2-1}\right)^n + \left(x-\sqrt{x^2-1}\right)^n\right]$$

ii) $$T_n(x) = \frac{n}{2}\sum_{j=0}^{\left[\frac{n}{2}\right]} \frac{(-1)^j(n-j-1)!}{j!(n-2j)!}(2x)^{n-2j} =$$

$$= \sum_{j=0}^{\left[\frac{n}{2}\right]} (-1)^j \frac{n}{n-j}\binom{n-j}{j} 2^{n-2j-1}x^{n-2j}\ , \quad n = 1,2,\ldots$$

iii) $$T_n(x) = \sum_{i=0}^{\left[\frac{n}{2}\right]} (-1)^i \binom{n}{2i} x^{n-2i}(1-x^2)^i =$$

$$= \sum_{i=0}^{\left[\frac{n}{2}\right]} \sum_{\kappa=0}^{i} (-1)^{i+\kappa}\binom{n}{2i}\binom{i}{\kappa} x^{n+2\kappa-2i}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

iv) $$T_n(x) = \sum_{\kappa=0}^{\left[\frac{n}{2}\right]} \sum_{j=\kappa}^{\left[\frac{n}{2}\right]} (-1)^{\kappa} \binom{n}{2j} \binom{j}{\kappa} x^{n-2\kappa} \quad , \quad n = 1,2,\ldots$$

v) $$T_n(x) = \frac{(-1)^n}{2^n n!} (1-x^2)^{\frac{1}{2}} \frac{d^n}{dx^n} \left[(1-x^2)^{n-\frac{1}{2}}\right] .$$

17. Να δειχθεί ότι τα πολυώνυμα Chebyshev και τα πολυώνυμα Legendre συνδέονται με τις σχέσεις

$$T_0(x) = \Phi_0(x) \quad , \quad T_1(x) = \Phi_1(x) ,$$

$$T_n(x) = \sum_{j=0}^{n} \Phi_j(x)\Phi_{n-j}(x) - x \sum_{j=0}^{n-1} \Phi_j(x)\Phi_{n-j-1}(x) \quad , \quad n > 1 .$$

18. Να δειχθεί ότι

$$\int_{-1}^{1} T_n(x)dx = \begin{cases} 0 & , \text{ για } n = 2\kappa+1 \\ \dfrac{2}{1-n^2} & , \text{ για } n = 2\kappa . \end{cases}$$

19. Αν $P_n(t)$ είναι ένα οποιοδήποτε πολυώνυμο βαθμού $n \geqq 1$ με συντελεστή του t^n την μονάδα, τότε ισχύει

$$\max_{t\in[a,b]} |P_n(t)| \geqq \max_{t\in[a,b]} \left| (b-a)^n 2^{1-2n} T_n\left(\frac{2t-b-a}{b-a}\right)\right| = 2^{1-2n}(b-a)^n.$$

20. Να δειχθεί ότι η ισότητα στην (4.29) ισχύει εάν και μόνον εάν $P_n(x) \equiv 2^{1-n}T_n(x)$.

21. Να δειχθεί ότι i)

$$\int_{-1}^{1} (1-x^2)^{-\frac{1}{2}} T_n(x)dx = \begin{cases} \pi & \text{για } n = 0 \\ 0 & \text{για } n \geqq 1 . \end{cases}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ii) $\int_{1}^{1}(1-x^2)^{-\frac{1}{2}}\, P_n(x)T_n(x)dx = \frac{\alpha_0 \pi}{2^n}$,

όπου $P_n(x)$ πολυώνυμο n βαθμού με συντελεστή του x^n τον α_0 .

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

5

ΑΡΙΘΜΗΤΙΚΗ ΟΛΟΚΛΗΡΩΣΗ

5.1 Εισαγωγή

Πολλές φορές, σε διάφορα προβλήματα, απαιτείται ο υπολογισμός της τιμής του ορισμένου ολοκληρώματος

$$\int_a^b f(x)dx \ , \tag{5.1}$$

όπου $f(x)$ είναι κάποια γνωστή πραγματική συνάρτηση, συνεχής στο πεπερασμένο διάστημα $[a,b]$. Εάν η αντιπαράγωγος της $f(x)$, δηλαδή , μία συνάρτηση $F(x)$ για την οποία ισχύει $F'(x) = f(x)$, είναι γνωστή,τότε το ολοκλήρωμα (5.1) μπορεί να υπολογισθεί από τον γνωστό τύπο

$$\int_a^b f(x)dx = F(b) - F(a) \ . \tag{5.2}$$

Συχνά όμως, η $F(x)$ δεν μπορεί με αναλυτικές μεθόδους να εκφρασθεί σε "κλειστή" μορφή, σαν ένας πεπερασμένος συνδυασμός στοιχειωδών συναρτήσεων, ή όταν αυτό είναι δυνατό είναι αρκετά πολύπλοκης

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

μορφής, με αποτέλεσμα η χρησιμοποίηση του τύπου (5.2) να είναι δύσκολη ή πρακτικά αδύνατη. Επίσης, στην περίπτωση που η $f(x)$ δεν είναι γνωστή αναλυτικά, αλλά δίνεται ένα σύνολο τιμών της (π.χ. πειραματικά δεδομένα) για τιμές του x στο διάστημα $[a,b]$, δεν μπορούν να εφαρμοστούν αναλυτικές μέθοδοι για τον υπολογισμό του (5.1). Για τους λόγους αυτούς χρησιμοποιείται η αριθμητική ολοκλήρωση, δηλαδή μέθοδοι προσδιορισμού μιας προσεγγιστικής τιμής του (5.1). Οι τύποι που προκύπτουν για τον προσεγγιστικό υπολογισμό των ορισμένων ολοκληρωμάτων έχουν μεγάλη πρακτική αξία .

Η βασική ιδέα των μεθόδων αριθμητικής ολοκληρώσεως στηρίζεται στην προσέγγιση της $f(x)$ στο $[a,b]$, συνήθως, από ένα πολυώνυμο $P(x)$ (π.χ. πολυώνυμο παρεμβολής). Σύμφωνα με το θεώρημα προσεγγίσεως του Weierstrass, κάθε συνεχής συνάρτηση μπορεί να προσεγγιστεί σε ένα κλειστό διάστημα όσο θέλουμε καλύτερα από ένα πολυώνυμο. Έτσι, αφού $f(x) \approx P(x)$ στο $[a,b]$, θέτουμε

$$\int_a^b f(x)dx \approx \int_a^b P(x)dx \,, \tag{5.3}$$

όπου το δεύτερο μέλος της (5.3) υπολογίζεται αναλυτικά εύκολα αφού το $P(x)$ είναι πολυώνυμο.

5.2 Κλειστοί τύποι αριθμητικής ολοκληρώσεως των Newton–Cotes

Ας υποθέσουμε ότι θέλουμε να υπολογίσουμε το ολοκλήρωμα

$$\int_a^b f(x)dx$$

και δίνονται ή μπορούν να βρεθούν οι τιμές της $f(x)$ στα $n+1$ $(n > 0)$ ισαπέχοντα σημεία $x_i = a + ih$, $i = 0(1)n$, όπου $h = \frac{b-a}{n}$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Προφανώς $x_0 = \alpha$ και $x_n = \alpha + nh = b$. Έστω $P_n(x)$ το πολυώνυμο παρεμβολής βαθμού το πολύ n με $P_n(x_i) = f_i = f(x_i) = f(\alpha + ih)$, $i = 0(1)n$. Από τις (2.8) και (2.9) (τύπος παρεμβολής του Lagrange) έχουμε $P_n(x) = \sum_{i=0}^{n} L_i(x) f_i$, όπου

$$L_i(x) = \prod_{\substack{j=0 \\ j\neq i}}^{n} \left(\frac{x - x_j}{x_i - x_j} \right) .$$

Εισάγοντας την μεταβλητή t από την σχέση $x = x_0 + ht = \alpha + ht$ (βλέπε (2.80)), οπότε είναι

$$L_i(x) = L_i(x(t)) \equiv \Lambda_i(t) = \prod_{\substack{j=0 \\ j\neq i}}^{n} \left(\frac{t - j}{i - j} \right) \tag{5.4}$$

και επειδή $f(x) \approx P_n(x)$ στο $[\alpha, b]$, προκύπτει ότι

$$\int_\alpha^b f(x)dx \approx \int_\alpha^b P_n(x)dx = \sum_{i=0}^{n} f_i \int_\alpha^b L_i(x)dx = h \sum_{i=0}^{n} f_i \int_0^n \Lambda_i(t)dt \ , \quad \text{ή}$$

$$\int_\alpha^b f(x)dx \approx h \sum_{i=0}^{n} A_i f_i \ , \tag{5.5}$$

όπου

$$A_i = \int_0^n \Lambda_i(t)dt \ .$$

Λόγω της (5.4) οι συντελεστές A_i , $i = 0(1)n$ στον προσεγγιστικό τύπο (5.5) εξαρτώνται μόνον από το n (δηλαδή, δεν εξαρτώνται από την $f(x)$ και τα α, b) και μπορούν να βρεθούν σε πίνακες για τις διάφορες τιμές του n . Μπορεί ακόμη να δειχθεί ότι $A_i = A_{n-i}$, $i = 0(1)n$.

Από τον τύπο (5.5) για τις διάφορες τιμές του n προκύπτουν οι καλούμενοι κλειστοί τύποι αριθμητικής ολοκληρώσεως των Newton – Cotes.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Καλούνται κλειστοί επειδή και τα δύο όρια a,b της ολοκληρώσεως είναι σημεία παρεμβολής.

Για παράδειγμα, αν $n = 1$ από την (5.5) έχουμε

$$\int_a^b f(x)dx \approx h(A_0 f_0 + A_1 f_1) \ , \tag{5.6}$$

όπου

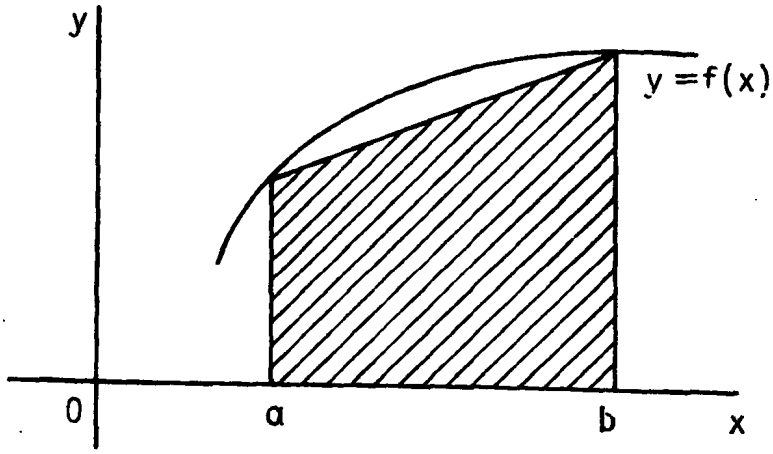
$$A_0' = \int_0^1 \Lambda_0(t)dt = \int_0^1 \frac{t-1}{0-1}\,dt = -\int_0^1 (t-1)dt = \frac{1}{2} \ ,$$

$$A_1 = \int_0^1 \Lambda_1(t)dt = \int_0^1 \frac{t-0}{1-0}\,dt = \int_0^1 t dt = \frac{1}{2} \ .$$

Άρα ο (5.6) γίνεται

$$\int_a^b f(x)dx \approx \frac{h}{2}(f_0 + f_1) = \frac{b-a}{2}\Big[f(a) + f(b)\Big] \tag{5.7}$$

και καλείται απλός κανόνας (ή τύπος) του τραπεζίου. Η γεωμετρική ερμηνεία αυτού φαίνεται στο σχήμα 5.1 . (Το εμβαδόν του γραμμοσκιασμένου τραπεζίου δίνεται από το δεύτερο μέλος του (5.7)).



σχ. 5.1

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Για $n = 2$ προκύπτει ο καλούμενος απλός κανόνας (του $\frac{1}{3}$) του Simpson:

$$\int_a^b f(x)dx \approx \frac{h}{3}(f_0 + 4f_1 + f_2) , \tag{5.8}$$

όπου

$$f_i = f(a + ih) , \quad h = \frac{b-a}{2} , \quad i = 0,1,2 .$$

Για $n = 3$ προκύπτει ο καλούμενος απλός κανόνας των $\frac{3}{8}$

$$\int_a^b f(x)dx \approx \frac{3h}{8}(f_0 + 3f_1 + 3f_2 + f_3) , \tag{5.9}$$

όπου

$$f_i = f(a + ih) , \quad h = \frac{b-a}{3} , \quad i = 0(1)3 .$$

Πρέπει να σημειωθεί ότι ο τύπος αριθμητικής ολοκληρώσεως των Newton – Cotes (5.5) είναι ακριβής για όλα τα πολυώνυμα βαθμού $\leqq n$, αφού, όπως είναι γνωστό από την παρεμβολή, εάν η $f(x)$ είναι ένα πολυώνυμο βαθμού $\leqq n$, τότε $P_n(x) \equiv f(x)$. Έτσι, θεωρώντας την περίπτωση $f(x) \equiv 1$ προκύπτει ότι

$$\sum_{i=0}^{n} A_i = n . \tag{5.10}$$

Θα λέμε ότι η τάξη του τύπου (5.5) είναι ℓ εάν αυτός είναι ακριβής για όλα τα πολυώνυμα βαθμού $\leqq \ell$ και εάν δεν είναι ακριβής για όλα τα πολυώνυμα βαθμού $\ell + 1$. (Σημειώνεται ότι εάν υπάρχει τουλάχιστον ένα πολυώνυμο βαθμού $\ell + 1$ για το οποίο δεν είναι ακριβής, τότε για όλα τα πολυώνυμα βαθμού $\ell + 1$ δεν είναι ακριβής).

Στην πράξη, συνήθως, για να επιτύχουμε μεγαλύτερη ακρίβεια δεν χρησιμοποιούμε τύπους των Newton – Cotes με μεγάλο n, αφού όπως είναι γνωστό τα πολυώνυμα παρεμβολής μεγάλου βαθμού (σε ισαπέχοντα ση-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

μεία παρεμβολής) μπορεί να μην προσεγγίζουν καλά μία ομαλή συνάρτηση $f(x)$. Επί πλέον, έχει αποδειχθεί ότι η ακολουθία των δεύτερων μελών του τύπου (5.5) των Newton – Cotes για $n = 1,2,3,...$ δεν συγκλίνει αναγκαστικά στην ακριβή τιμή του $\int_a^b f(x)dx$ καθώς το $n \to \infty$. Για τους λόγους αυτούς η ακόλουθη πρακτική διαδικασία εφαρμόζεται. Το διάστημα ολοκληρώσεως $[a,b]$ διαιρείται σε ένα αριθμό υποδιαστημάτων και κατόπιν στα υποδιαστήματα αυτά εφαρμόζονται οι τύποι των Newton – Cotes με μικρό n . Το ζητούμενο ολοκλήρωμα προσεγγίζεται επομένως από το άθροισμα των προσεγγίσεων για όλα τα υποδιαστήματα. Από την διαδικασία αυτή η οποία θα γίνει πιο κατανοητή στη συνέχεια προκύπτουν οι καλούμενοι γενικευμένοι (σύνθετοι) κανόνες (τύποι) αριθμητικής ολοκληρώσεως των Newton – Cotes που χρησιμοποιούνται πολύ στην πράξη.

Για τον υπολογισμό του ολοκληρώματος (5.1), πρώτα διαιρούμε το διάστημα ολοκληρώσεως $[a,b]$ σε $N \geq 1$ ισομήκη υποδιαστήματα $[x_i, x_{i+1}]$ $i = 0(1)N-1$. Προφανώς, το μήκος κάθε υποδιαστήματος θα είναι

$$h = \frac{b-a}{N} \tag{5.11}$$

και ισχύει $x_i = a + ih$, $i = 0(1)N$. Έτσι έχουμε

$$\int_a^b f(x)dx = \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} f(x)dx \ . \tag{5.12}$$

Εάν εφαρμοστεί τώρα ο απλός κανόνας του τραπεζίου (5.7) στο διάστημα $[x_i, x_{i+1}]$ παίρνουμε

$$\int_{x_i}^{x_{i+1}} f(x)dx \approx \frac{h}{2}\left(f_i + f_{i+1}\right) \tag{5.13}$$

και επομένως η (5.12) δίνει

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\int_a^b f(x)dx \approx \frac{h}{2}\sum_{i=0}^{N-1}\left(f_i + f_{i+1}\right) = h\left(\frac{f_0}{2} + f_1 + f_2 + \ldots + f_{N-1} + \frac{f_N}{2}\right). \quad (5.14)$$

Ο τύπος (5.14) καλείται γενικευμένος κανόνας του τραπεζίου.

Εάν ο θετικός ακέραιος N εκλεγεί άρτιος, τότε αντί της (5.12) θεωρούμε την

$$\int_a^b f(x)dx = \int_{x_0}^{x_2} f(x)dx + \int_{x_2}^{x_4} f(x)dx + \ldots + \int_{x_{N-2}}^{x_N} f(x)dx = \sum_{i=0}^{\frac{N}{2}-1} \int_{x_{2i}}^{x_{2i+2}} f(x)dx \,. \quad (5.15)$$

Στο κάθε διπλό υποδιάστημα $\left[x_{2i}, x_{2i+2}\right]$, $i = 0(1)\frac{N}{2}-1$ εφαρμοζόμενος ο απλός κανόνας του Simpson (5.8) δίνει

$$\int_{x_{2i}}^{x_{2i+2}} f(x)dx \approx \frac{h}{3}\left(f_{2i} + 4f_{2i+1} + f_{2i+2}\right) . \quad (5.16$$

Λόγω της (5.16), από την (5.15) προκύπτει τώρα ότι

$$\int_a^b f(x)dx \approx \sum_{i=0}^{\frac{N}{2}-1} \frac{h}{3}\left(f_{2i} + 4f_{2i+1} + f_{2i+2}\right) = \frac{h}{3}\left(f_0 + 4f_1 + 2f_2 + 4f_3 + \ldots\right.$$

$$\left. + 2f_{N-2} + 4f_{N-1} + f_N\right) . \quad (5.17)$$

Ο τύπος (5.17) καλείται γενικευμένος κανόνας του Simpson και είναι σως ο περισσότερο διαδεδομένος τύπος αριθμητικής ολοκληρώσεως.

Κατά τον ίδιο τρόπο, αν υποθέσουμε ότι ο N διαιρείται ακριβώς με το 3 και εφαρμοστεί ο απλός κανόνας των $\frac{3}{8}$ σε κάθε τριπλό υπο διάστημα $\left[x_{3i}, x_{3i+3}\right]$, $i = 0(1)\frac{N}{3}-1$, προκύπτει ο ακόλουθος γενικ μένος κανόνας των $\frac{3}{8}$:

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\int_a^b f(x)dx \approx \frac{3h}{8}\left(f_0 + 3f_1 + 3f_2 + 2f_3 + 3f_4 + 3f_5 + 2f_6 + \ldots + 3f_{N-1} + f_N\right). \tag{5.18}$$

Πρέπει να σημειωθεί ότι όταν ζητούμε να υπολογίσουμε το ολοκλήρωμα (5.1) με κάποια δεδομένη ακρίβεια, π.χ. με ακρίβεια κ δ.θ., χρησιμοποιώντας κάποιο συγκεκριμένο γενικευμένο κανόνα των Newton – Cotes, αρχίζουμε με την μικρότερη δυνατή τιμή του N (π.χ. N = 1 για τον (5.14), N = 2 για τον (5.17), N = 3 για τον (5.18) κ.λ.π.) και διπλασιάζοντας διαδοχικά την τιμή του N βρίσκουμε διαδοχικές προσεγγίσεις για το ζητούμενο ολοκλήρωμα. Η επαναληπτική αυτή υπολογιστική διαδικασία σταματά όταν η διαφορά δύο διαδοχικών προσεγγίσεων είναι απολύτως $\leq \frac{1}{2}10^{-\kappa}$.

Αξίζει να αναφερθεί τέλος, ότι η ακολουθία των προσεγγίσεων που προκύπτει από ένα γενικευμένο κανόνα των Newton – Cotes καθώς το $N \to \infty$ συγκλίνει πάντοτε στην ακριβή τιμή του $\int_a^b f(x)dx$.

5.3 Σφάλματα αποκοπής των κλειστών τύπων των Newton–Cotes

Το σφάλμα αποκοπής $E_n(f)$ του τύπου αριθμητικής ολοκληρώσεως (5.5) δίνεται ως γνωστόν από την διαφορά

$$E_n(f) = h\sum_{i=0}^{n} A_i f_i - \int_a^b f(x)dx = \int_a^b P_n(x)dx - \int_a^b f(x)dx =$$

$$= \int_a^b \left(P_n(x) - f(x)\right)dx . \tag{5.19}$$

Εάν υποθέσουμε ότι η f(x) είναι συνεχώς παραγωγίσιμη n + 1 φορές στο [a,b] , τότε σύμφωνα με το θεώρημα 2.1 (βλέπε και (2.24)) η

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

(5.19) γίνεται

$$E_n(f) = -\frac{1}{(n+1)!}\int_a^b \prod_{i=0}^{n}(x - x_i)f^{(n+1)}(\xi(x))dx \ , \qquad (5.20)$$

όπου $\xi \in [x_0, x_n] \equiv [a,b]$ ή ακόμη, βάσει των $x = x_0 + ht$, $x_i = a + ih$, $i = 0(1)n$, $h = \frac{b-a}{n}$,

$$E_n(f) = -\frac{h^{n+2}}{(n+1)!}\int_0^n \prod_{i=0}^{n}(t-i)f^{(n+1)}(\xi(t))dt \ . \qquad (5.21)$$

Από την (5.21) για $n = 1$ (κανόνας του τραπεζίου) έχουμε $E_1(f) = -\frac{h^3}{2}\int_0^1 t(t-1)f''(\xi(t))dt$. Επειδή $t(t-1) \leq 0$ στο $[0,1]$, σύμφωνα με το γενικό θεώρημα μέσης τιμής για ολοκληρώματα προκύπτει ότι

$$E_1(f) = -\frac{h^3 f''(\xi(\theta))}{2}\int_0^1 t(t-1)dt \quad \text{για κάποιο} \quad \theta \in [0,1] \ .$$

Έτσι, τελικά το σφάλμα αποκοπής του απλού κανόνα του τραπεζίου είναι

$$E_1(f) = \frac{1}{12}h^3 f''(\xi) = \frac{1}{12}(b-a)^3 f''(\xi) \ , \qquad (5.22)$$

όπου $\xi \in [x_0, x_1] \equiv [a,b]$, αφού $\int_0^1 t(t-1)dt = -\frac{1}{6}$. Έστω $|f''(x)| \leq M_2$, $\forall x \in [a,b]$. Τότε,από την (5.22) έχουμε $|E_1(f)| \leq \frac{(b-a)^3 M_2}{12}$.

Για το σφάλμα αποκοπής του απλού κανόνα του Simpson $(n = 2)$, είναι εύκολο να δειχθεί ότι δεν μπορεί να εφαρμοσθεί η ίδια διαδικασία με την περίπτωση $n = 1$. Μπορεί πάντως να αποδειχθεί (βλέπε ά-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

σκηση 9) ότι ισχύει

$$E_2(f) = \frac{1}{90} h^5 f^{(4)}(\xi) \ , \tag{5.23}$$

όπου $\xi \in [x_0, x_2] \equiv [a,b]$ και $h = \frac{b-a}{2}$. (Υποτίθεται επί πλέον ότι η $f^{(4)}(x)$, δηλαδή, η $n+2$ τάξεως παράγωγος της $f(x)$, υπάρχει και είναι συνεχής στο $[a,b]$).

Επίσης, το σφάλμα αποκοπής του απλού κανόνα των $\frac{3}{8}$ (περίπτωση $n = 3$) δίνεται από τον τύπο

$$E_3(f) = \frac{3}{80} h^5 f^{(4)}(\xi) \ , \tag{5.24}$$

όπου $\xi \in [x_0, x_3] \equiv [a,b]$ και $h = \frac{b-a}{3}$.

Από τους τύπους των σφαλμάτων αποκοπής (5.22), (5.23) και (5.24) είναι φανερό ότι η τάξη των κανόνων του τραπεζίου, του Simpson και των $\frac{3}{8}$ είναι αντίστοιχα 1,3 και 3 .

Είναι δυνατόν να δειχθεί δια γενικωτέρων μεθόδων ότι,εάν η $f(x)$ είναι συνεχώς παραγωγίσιμη $n+1$ φορές εάν ο n είναι περιττός και $n+2$ φορές εάν ο n είναι άρτιος, τότε ο τύπος του σφάλματος αποκοπής (5.20) ή (5.21) λαμβάνει τις ακόλουθες ισοδύναμες μορφές:

$$\text{i)} \quad E_n(f) = \begin{cases} -\dfrac{h^{n+2} f^{(n+1)}(\xi)}{(n+1)!} \displaystyle\int_0^n \prod_{i=0}^{n} (t-i)\,dt \ , & \text{εάν } n \text{ περιττός} \\[2ex] -\dfrac{h^{n+3} f^{(n+2)}(\xi)}{(n+2)!} \displaystyle\int_0^n t \prod_{i=0}^{n} (t-i)\,dt \ , & \text{εάν } n \text{ άρτιος} \end{cases} \tag{5.25}$$

$$\text{ii)} \quad E_n(f) = \begin{cases} \dfrac{E_n(x^{n+1})}{(n+1)!} f^{(n+1)}(\xi) \ , & \text{εάν } n \text{ περιττός} \\[2ex] \dfrac{E_n(x^{n+2})}{(n+2)!} f^{(n+2)}(\xi) \ , & \text{εάν } n \text{ άρτιος} \end{cases} \tag{5.26}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

όπου $\xi \in [a,b]$ και $E_n(x^m)$ είναι το σφάλμα αποκοπής όταν $f(x) \equiv x^m$.

Για παράδειγμα, έστω $[a,b] \equiv [-1,1]$ και $n = 2$. Τότε, από την (5.26) έχουμε $E_2(f) = \dfrac{E_2(x^4)}{4!} f^{(4)}(\xi)$, $\xi \in [-1,1]$, όπου λόγω της (5.8) είναι

$$E_2(x^4) = \frac{1}{3}\left[(-1)^4 + 4(0)^4 + 1^4\right] - \int_0^1 x^4 dx = \frac{2}{3} - \frac{2}{5} = \frac{4}{15}.$$ Άρα,

$$E_2(f) = \frac{4}{15 \cdot 4!} f^{(4)}(\xi) = \frac{1}{90} f^{(4)}(\xi) \quad (\text{βλέπε (5.23) για } h = 1).$$

Από τις (5.25), (5.26) προκύπτει ότι η τάξη του τύπου αριθμητικής ολοκληρώσεως (5.5) των Newton – Cotes είναι n αν ο n είναι περιττός, ενώ είναι $n+1$ αν ο n είναι άρτιος.

Αφού βρέθηκαν τύποι για τα σφάλματα αποκοπής των απλών κανόνων των Newton – Cotes, είναι εύκολο να βρεθούν τα σφάλματα αποκοπής των αντίστοιχων γενικευμένων κανόνων των Newton – Cotes. Έτσι, στη συνέχεια βρίσκουμε το σφάλμα αποκοπής του γενικευμένου κανόνα του τραπεζίου (5.14). Ανάλογη διαδικασία εφαρμόζεται για οποιονδήποτε άλλο γενικευμένο κανόνα.

Αν καλέσουμε $E_{1,N}(f)$ το σφάλμα αποκοπής του (5.14), τότε λόγω των (5.12), (5.13), (5.14) και (5.22) είναι

$$E_{1,N}(f) = \frac{h}{2}\sum_{i=0}^{N-1}\left(f_i + f_{i+1}\right) - \sum_{i=0}^{N-1}\int_{x_i}^{x_{i+1}} f(x)dx = \sum_{i=0}^{N-1}\left[\frac{h}{2}\left(f_i + f_{i+1}\right) - \int_{x_i}^{x_{i+1}} f(x)dx\right] =$$

$$= \sum_{i=0}^{N-1}\left[\frac{h^3}{12} f''(\xi_i)\right] = \frac{h^3}{12}\sum_{i=0}^{N-1} f''(\xi_i), \qquad (5.27)$$

όπου $\xi_i \in [x_i, x_{i+1}]$ και το h δίνεται από την (5.11). Αφού η $f''(x)$ είναι συνεχής, σύμφωνα με το θεώρημα μέσης τιμής για αθροίσματα, έχουμε

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\sum_{i=0}^{N-1} f''(\xi_i) = f''(\xi)\sum_{i=0}^{N-1} 1 = Nf''(\xi) \quad \text{με} \quad \xi \in \left[\min_i\{\xi_i\},\ \max_i\{\xi_i\}\right] \subseteq [a,b].$$

Άρα η (5.27) γίνεται

$$E_{1,N}(f) = \frac{h^3N}{12} f''(\xi) = \frac{b-a}{12} h^2 f''(\xi)\ , \quad \xi \in [a,b]\ , \tag{5.28}$$

αφού $hN = b - a$.

Για τον γενικευμένο κανόνα (5.17) του Simpson (N άρτιος) μπορεί να δειχθεί ότι

$$E_{2,N}(f) = \frac{h^5N}{180} f^{(4)}(\xi) = \frac{b-a}{180} h^4 f^{(4)}(\xi)\ , \quad \xi \in [a,b]\ , \tag{5.29}$$

ενώ για τον γενικευμένο κανόνα (5.18) των $\frac{3}{8}$ (N=πολλ. του 3) είναι

$$E_{3,N}(f) = \frac{Nh^5}{80} f^{(4)}(\xi) = \frac{b-a}{80} h^4 f^{(4)}(\xi)\ , \quad \xi \in [a,b]\ . \tag{5.30}$$

Οι τύποι για το σφάλμα αποκοπής των γενικευμένων κανόνων είναι χρήσιμοι και για την εκλογή του h $\left(N = \frac{b-a}{h}\right)$. Έτσι, εάν, για παράδειγμα, θέλουμε να υπολογίσουμε το ολοκλήρωμα $\int_a^b f(x)dx$ με τον γενικευμένο κανόνα του Simpson, έτσι ώστε το μέγιστο απόλυτο σφάλμα αποκοπής να είναι $< \varepsilon$, τότε, λόγω της (5.29), το βήμα h πρέπει να εκλεγεί έτσι ώστε

$$\left|E_{2,N}(f)\right| \leq \frac{b-a}{180} h^4 \max_{x\in[a,b]} \left|f^{(4)}(x)\right| < \varepsilon\ , \quad \text{δηλαδή,}$$

$$h < \left[\frac{180\,\varepsilon}{(b-a)M_4}\right]^{\frac{1}{4}}\ , \quad \text{όπου} \quad M_4 = \max_{x\in[a,b]} \left|f^{(4)}(x)\right|\ .$$

Πρέπει να σημειωθεί ότι κατά την χρησιμοποίηση κάποιου γενικευμένου κανόνα αριθμητικής ολοκληρώσεως εκτός από το σφάλμα αποκοπής υπεισέρχεται και το υπολογιστικό σφάλμα (συνήθως σφάλμα στρογγυλεύσεως), το οποίο αυξάνει όταν ο N αυξάνει, δηλαδή, όταν το βήμα h

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ελαττώνεται. Πάντως, το σφάλμα στρογγυλεύσεως δεν είναι τόσο σοβαρό πρόβλημα στην αριθμητική ολοκλήρωση, όπως αντίθετα συμβαίνει με την αριθμητική παραγώγιση και μπορούμε να πούμε ότι η αριθμητική ολοκλήρωση είναι ευσταθής διαδικασία.

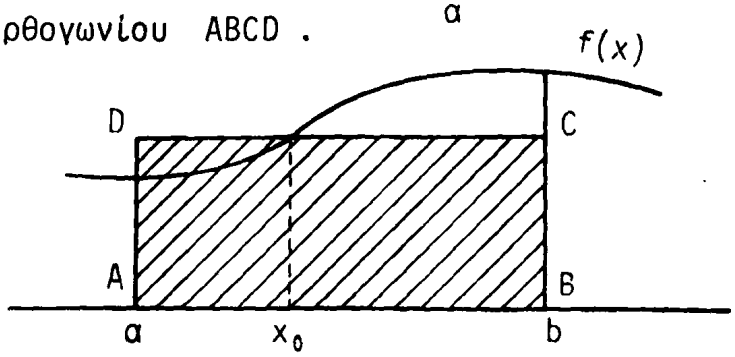
5.4 Ανοικτοί – ημιανοικτοί τύποι αριθμητικής ολοκληρώσεως των Newton – Cotes

Όπως αναφέρθηκε στην παράγραφο 5.2, οι τύποι που προκύπτουν από τον (5.5) για $n = 1,2,3,...$ καλούνται κλειστοί. Όπως θα δούμε στη συνέχεια μπορούν να προκύψουν τύποι της μορφής (5.5) όπου κανένα από τα άκρα a και b του διαστήματος ολοκληρώσεως δεν είναι σημείο παρεμβολής. Οι τύποι αυτοί καλούνται ανοικτοί τύποι των Newton – Cotes. Στους ημιανοικτούς τύπους ένα μόνον από τα άκρα a,b είναι σημείο παρεμβολής.

Αν υποθέσουμε τώρα ότι έχουμε ένα σημείο παρεμβολής $x_0 \in [a,b]$, τότε το πολυώνυμο παρεμβολής είναι $P_0(x) \equiv f(x_0)$ και επομένως προκύπτει ότι

$$\int_a^b f(x)dx \approx \int_a^b P_0(x)dx = \int_a^b f(x_0)dx = f(x_0)\int_a^b dx = (b-a)f(x_0). \quad (5.31)$$

Ο τύπος (5.31) καλείται κανόνας του ορθογωνίου. Η γεωμετρική ερμηνεία αυτού (βλέπε σχ. 5.2) είναι ότι το $\int_a^b f(x)dx$ προσεγγίζεται από το εμβαδόν του ορθογωνίου ABCD .



σχ. 5.2

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Αν $x_0 = a, b, \frac{a+b}{2}$ προκύπτουν αντίστοιχα οι ακόλουθοι τύποι

$$\int_a^b f(x)dx \approx (b-a)\ f(a), \tag{5.32}$$

$$\int_a^b f(x)dx \approx (b-a)f(b), \tag{5.33}$$

$$\int_a^b f(x)dx \approx (b-a)f\left(\frac{a+b}{2}\right) \quad . \tag{5.34}$$

Οι τύποι (5.32), (5.33), (5.34) είναι αντίστοιχα γνωστοί σαν, προς τα εμπρός κανόνας του ορθογωνίου, προς τα πίσω κανόνας του ορθογωνίου και κανόνας του μέσου. Οι δύο πρώτοι είναι ημιανοικτοί ο δε τρίτος ανοικτός.

Για την εύρεση του σφάλματος αποκοπής του τύπου (5.32), από την (5.20) παίρνουμε για $n = 0$

$$E_0(f) = -\int_a^b (x - x_0)f'(\xi(x))dx = -\int_a^b (x-a)f'(\xi(x))dx \ , \tag{5.35}$$

όπου $\xi \in [a,x]$. Επειδή $x - a \geqq 0$ στο $[a,b]$, σύμφωνα με το γενικό θεώρημα μέσης τιμής για ολοκληρώματα έχουμε

$$E_0(f) = -f'(\xi(\theta))\int_a^b (x-a)dx \ , \ \theta \in [a,b] \ , \ \text{ή τελικά}$$

$$E_0(f) = -\frac{(b-a)^2}{2}f'(\xi) \ , \ \xi \in [a,b] \quad . \tag{5.36}$$

Ομοίως,μπορεί να βρεθεί ότι το σφάλμα αποκοπής του τύπου (5.33) είναι

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$E_0^{(1)}(f) = \frac{(b-a)^2}{2} f'(\xi), \quad \xi \in [a,b] \quad . \tag{5.37}$$

Ακόμη, μπορεί να δειχθεί (βλέπε άσκηση 47) ότι το σφάλμα αποκοπής του τύπου (5.34) είναι

$$E_0^{(2)}(f) = -\frac{(b-a)^3}{24} f''(\xi) \ , \quad \xi \in [a,b] \ . \tag{5.38}$$

Η γενική μορφή των ανοικτών τύπων αριθμητικής ολοκληρώσεως των Newton – Cotes προκύπτει ως εξής: θεωρούμε τα $n+1$ ισαπέχοντα σημεία x_i , $i = 0(1)n$, τέτοια ώστε

$$x_i = a + (i+1)h \ , \quad i = 0(1)n \ , \text{ όπου } h = \frac{b-a}{n+2} \ . \tag{5.39}$$

Προφανώς $x_{i+1} - x_i = h$, $i = 0(1)n-1$ και $a = x_0 - h$, $b = x_n + h$. Αν $P_n(x)$ είναι το πολυώνυμο παρεμβολής των σημείων $(x_i, f(x_i)) \equiv$ $\equiv (x_i, f_i)$, $i = 0(1)n$, τότε κατ' αναλογίαν προς τον τύπο (5.5) προκύπτει ότι

$$\int_a^b f(x)dx \approx \int_a^b P_n(x)dx = h \sum_{i=0}^{n} B_i f_i \ , \tag{5.40}$$

όπου

$$B_i = \int_{-1}^{n+1} \prod_{\substack{j=0 \\ j \neq i}}^{n} \left(\frac{t-j}{i-j} \right) dt \ , \quad f_i \equiv f(x_i) = f(a+(i+1)h) \ , \ i = 0(1)n$$

και $h = \dfrac{b-a}{n+2}$.

Για $n = 0$ από τον (5.40) προκύπτει ο τύπος (5.34) του μέσου. Για $n = 1$ προκύπτει ο τύπος

$$\int_a^b f(x)dx \approx \frac{3h}{2} \left(f_0 + f_1 \right) \ , \tag{5.41}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

όπου $h = \frac{b-a}{3}$ και $f_i = f(a+(i+1)h)$, $i = 0,1$ με αντίστοιχο σφάλμα αποκοπής

$$E_1^A(f) = -\frac{3h^3}{4} f''(\xi) \ , \ \xi \in [a,b] \ . \tag{5.42}$$

Για $n = 2$ (βλέπε άσκηση 47) προκύπτει ο ακόλουθος τύπος

$$\int_a^b f(x)dx \approx \frac{4h}{3}(2f_0 - f_1 + 2f_2) \ , \ . \tag{5.43}$$

όπου $h = \frac{b-a}{4}$ και $f_i = f(a+(i+1)h)$, $i = 0,1,2,$. Το σφάλμα αποκοπής του (5.43) είναι

$$E_2^A(f) = -\frac{14h^5}{45} f^{(4)}(\xi) \ , \ \xi \in [a,b] \ . \tag{5.44}$$

Γενικά οι ανοικτοί τύποι είναι κατώτεροι (λιγώτερο ακριβείς) από τους αντίστοιχους κλειστούς. Ο κανόνας του μέσου που είναι καλύτερος από τον κανόνα του τραπεζίου (όπως προκύπτει δια συγκρίσεως των απολύτων σφαλμάτων αποκοπής των κανόνων αυτών) είναι η πιο σπουδαία εξαίρεση. Οι ανοικτοί τύποι χρησιμοποιούνται κυρίως στην αριθμητική επίλυση συνήθων διαφορικών εξισώσεων και σε άλλες ειδικές περιπτώσεις.

Είναι φανερό ότι οι ανοικτοί και ημιανοικτοί τύποι μπορούν να γενικευθούν όπως και οι κλειστοί. Έτσι, ακολουθώντας την ίδια διαδικασία, ο γενικευμένος προς τα εμπρός κανόνας του ορθογωνίου προκύπτει ως εξής:

$$\int_a^b f(x)dx = \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} f(x)dx \approx \sum_{i=0}^{N-1} hf_i = h\left(f_0+f_1+\ldots+f_{N-2}+f_{N-1}\right) \ , \tag{5.45}$$

όπου $h = \frac{b-a}{N}$ και $f_i = f(a+ih)$, $i = 0(1)N-1$. Το σφάλμα αποκοπής του (5.45) είναι

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$E_{0,N}(f) = -\frac{b-a}{2} hf'(\xi) \ , \quad \xi \in [a,b] \ . \tag{5.46}$$

Επίσης, ο γενικευμένος προς τα πίσω κανόνας του ορθογωνίου είναι

$$\int_a^b f(x)dx \approx h\left[f_1 + f_2 + \ldots + f_{N-1} + f_N\right] \ , \tag{5.47}$$

όπου $h = \frac{b-a}{N}$, $f_i = f(a+ih)$, $i = 1(1)N$ με σφάλμα αποκοπής

$$E^{(1)}_{0,N}(f) = \frac{b-a}{2} hf'(\xi) \ , \quad \xi \in [a,b] \tag{5.48}$$

και ο γενικευμένος κανόνας του μέσου είναι.

$$\int_a^b f(x)dx \approx h\sum_{i=0}^{N-1} f\left[a + (i+\frac{1}{2})h\right] \ , \tag{5.49}$$

$h = \frac{b-a}{N}$ με σφάλμα αποκοπής

$$E^{(2)}_{0,N}(f) = -\frac{b-a}{24} h^2 f''(\xi) \ , \quad \xi \in [a,b] \ . \tag{5.50}$$

<u>Παράδειγμα 5.1</u> Να προσεγγιστεί το ολοκλήρωμα $\int_0^1 f(x)dx$, όπου $f(x) = \frac{1}{1+x^4}$ με την βοήθεια των γενικευμένων κανόνων του τραπεζίου, του Simpson και του μέσου. Να χρησιμοποιηθεί $N = 4$ και να διατηρούνται 6 δ.θ. στους υπολογισμούς. (Η τιμή του ολοκληρώματος με ακρίβεια 8 δ.θ. είναι 0.86697299) .

<u>Λύση</u> Είναι $h = \frac{b-a}{N} = \frac{1}{4} = 0.25$.

i) Γενικευμένος κανόνας του τραπεζίου (βλέπε (5.14)):

$$\int_0^1 \frac{dx}{1+x^4} \approx 0.25\left[\frac{1}{2(1+0^4)} + \frac{1}{1+(0.25)^4} + \frac{1}{1+(0.5)^4} + \frac{1}{1+(0.75)^4} + \frac{1}{2(1+1^4)}\right] \approx$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$\approx 0.25(0.5 + 0.996109 + 0.941176 + 0.759644 + 0.25) =$

$= 0.25(3.446929) \approx 0.861732$.

ii) Γενικευμένος κανόνας του Simpson (βλέπε (5.17)):

$$\int_0^1 \frac{dx}{1+x^4} \approx \frac{0.25}{3}\Big[1 + 4(0.996109) + 2(0.941176) + 4(0.759644) +$$

$$0.5\Big] = \frac{0.25}{3}(10.405364) \approx 0.867114 .$$

iii) Γενικευμένος κανόνας του μέσου (βλέπε (5.49)):

$$\int_0^1 \frac{dx}{1+x^4} \approx 0.25\left(\frac{1}{1+(0.125)^4} + \frac{1}{1+(0.375)^4} + \frac{1}{1+(0.625)^4} + \frac{1}{1+(0.875)^4}\right) \approx$$

$\approx 0.25(0.999756 + 0.980608 + 0.867613 + 0.630445) =$

$= 0.25(3.478422) \approx 0.869606$.

5.5 Extrapolation—Αριθμητική ολοκλήρωση του Romberg

Ας υποθέσουμε ότι η ακριβής τιμή $I = \int_a^b f(x)dx$ προσεγγίζεται με την βοήθεια κάποιου γενικευμένου κανόνα των Newton — Cotes από την $I(h)$, όπου $h = \frac{b-a}{N}$, με αντίστοιχο σφάλμα αποκοπής $E(h)$ της τάξεως του $h^m (m \geq 1)$ $(E(h) = O(h^m))$, δηλαδή, $E(h) \approx Mh^m$, όπου η ποσότης M θεωρείται σταθερά στο $[a,b]$ για την δοσμένη συνάρτηση $f(x)$. Έτσι έχουμε

$$E(h) = I(h) - I \tag{5.51}$$

και

$$I(h) - I \approx Mh^m . \tag{5.52}$$

Εκλέγοντας τώρα βήμα $h_1 = \lambda h$, $0 < \lambda < 1$ παίρνουμε μία νέα καλύτερη

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

προσέγγιση $I(h_1) = I(\lambda h)$ της I , για την οποία θα ισχύει λόγω της (5.52)

$$I(\lambda h) - I \approx M(\lambda h)^m \quad . \tag{5.53}$$

Οι σχέσεις (5.52) και (5.53) αποτελούν ένα σύστημα δύο εξισώσεων με τους δύο αγνώστους I και M . Πολλαπλασιάζοντας την (5.52) επί $-\lambda^m$ και προσθέτοντας την προκύπτουσα στην (5.53) παίρνουμε

$$I \approx \frac{I(\lambda h) - \lambda^m I(h)}{1 - \lambda^m} = I(\lambda h) + \frac{I(\lambda h) - I(h)}{\left(\frac{1}{\lambda}\right)^m - 1} =$$

$$= I(\lambda h) + \omega\Big[I(\lambda h) - I(h)\Big] \equiv I^{(1)}(\lambda h) \ , \tag{5.54}$$

όπου $\omega = \dfrac{1}{\left(\frac{1}{\lambda}\right)^m - 1}$. Η τιμή $I^{(1)}(\lambda h)$ αποτελεί μία βελτιωμένη προσέγγιση για την I (γενικά μπορεί να αποδειχθεί ότι το σφάλμα αποκοπής για την $I^{(1)}(\lambda h)$ είναι τουλάχιστον της τάξεως του h^{m+1}) και καλείται "extrapolated" προσέγγιση της I που βασίζεται στις τιμές $I(h)$ και $I(\lambda h)$.

Η παραπάνω διαδικασία που καλείται extrapolation του Richardson μπορεί να χρησιμοποιηθεί επαναληπτικά ("extrapolation στο όριο"), καθώς το h πλησιάζει στο 0 , και είναι η βάση για την αριθμητική ολοκλήρωση του Romberg.

Συνήθως εκλέγεται $\lambda = \frac{1}{2}$, οπότε η (5.54) γίνεται

$$I^{(1)}\left(\frac{h}{2}\right) = I\left(\frac{h}{2}\right) + \frac{I\left(\frac{h}{2}\right) - I(h)}{2^m - 1} \quad . \tag{5.55}$$

Με την extrapolation διαδικασία αποφεύγουμε την χρησιμοποίηση πολύ μικρού h , που θα είχε σαν αποτέλεσμα την αύξηση των υπολογιστικών πράξεων και επομένως του σφάλματος στρογγυλεύσεως, προκειμένου

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

να επιτύχουμε προσεγγίσεις ορισμένης ακρίβειας.

Η ολοκλήρωση του Romberg συνίσταται στην επαναληπτική εφαρμογή της διαδικασίας extrapolation με $\lambda = \frac{1}{2}$ στους γενικευμένους κανόνες των Newton – Cotes και συνήθως στον γενικευμένο κανόνα του τραπεζίου. Έτσι, εάν καλέσουμε $T(h) \equiv T_N$ την προσέγγιση που προκύπτει από τον γενικευμένο κανόνα του τραπεζίου $(h = \frac{b-a}{N})$, οπότε $T(\frac{h}{2}) \equiv T_{2N}$ και επειδή από την (5.28) είναι $m = 2$, η (5.55) δίνει

$$T_{2N}^{(1)} = T_{2N} + \frac{T_{2N} - T_N}{3} . \tag{5.56}$$

Μπορεί να δειχθεί (βλέπε άσκηση 26) ότι η $T_{2N}^{(1)}$ συμπίπτει με την προσέγγιση $S(\frac{h}{2}) \equiv S_{2N}$ που προκύπτει από τον γενικευμένο κανόνα του Simpson και επομένως το σφάλμα αποκοπής της είναι $O(h^4)$. Με βάση την (5.56) προκύπτει ότι

$$T_{4N}^{(1)} = T_{4N} + \frac{T_{4N} - T_{2N}}{3} . \tag{5.57}$$

Για την εύρεση της $T_{4N}^{(1)}$ απαιτείται και η T_{4N} . Επίσης θα ισχύει $T_{4N}^{(1)} \equiv S_{4N} \equiv S(\frac{h}{4})$. Έτσι, από την (5.55) για $m = 4$ η νέα προσέγγιση που βασίζεται στις $T_{2N}^{(1)}$ και $T_{4N}^{(1)}$ θα είναι

$$T_{4N}^{(2)} = T_{4N}^{(1)} + \frac{T_{4N}^{(1)} - T_{2N}^{(1)}}{15} . \tag{5.58}$$

Μπορεί να δειχθεί ότι η $T_{4N}^{(2)}$ συμπίπτει με την τιμή που προκύπτει από τον τέταρτο στη σειρά (επόμενο εκείνου των $\frac{3}{8}$) γενικευμένο κλειστό κανόνα των Newton – Cotes με βήμα $\frac{h}{4}$ (βλέπε άσκηση 28). Επομένως η επόμενη προσέγγιση που βασίζεται στις $T_{4N}^{(2)}$ και $T_{8N}^{(2)}$ είναι (τώρα είναι $m = 6$) .

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$T_{8N}^{(3)} = T_{8N}^{(2)} + \frac{T_{8N}^{(2)} - T_{4N}^{(2)}}{63} . \qquad (5.59)$$

Η διαδικασία συνεχίζεται μέχρις ότου επιτευχθεί η επιζητούμενη ακρίβεια. Είναι φανερό ότι η $T_{8N}^{(3)}$ απαιτεί την εύρεση των $T_{8N}^{(2)}$, $T_{8N}^{(1)}$, T_{8N} . Η συστηματική εύρεση των διαδοχικών επαναλήψεων γίνεται με την βοήθεια του παρακάτω πίνακα του Romberg

h	N	T_N			
$\frac{h}{2}$	$2N$	T_{2N}	$T_{2N}^{(1)}$		
$\frac{h}{4}$	$4N$	T_{4N}	$T_{4N}^{(1)}$	$T_{4N}^{(2)}$	
$\frac{h}{8}$	$8N$	T_{8N}	$T_{8N}^{(1)}$	$T_{8N}^{(2)}$	$T_{8N}^{(3)}$
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.

Πίνακας 5.1

Πρέπει να σημειωθεί ότι η $T_{8N}^{(3)}$ και οι επόμενες προσεγγίσεις δεν συμπίπτουν πλέον με τιμές που προκύπτουν από τους γενικευμένους κανόνες των Newton – Cotes .

<u>Παράδειγμα 5.2</u> Να υπολογισθεί το $\int_{-3}^{5} x^7 dx$ με ολοκλήρωση Romberg αρχίζοντας με βήμα $h = 8$ και φθάνοντας μέχρι $h = 1$ (Ακριβής τιμή = 48008) .

<u>Λύση</u> Έχουμε $N = \frac{b-a}{h} = \frac{5-(-3)}{h} = \frac{8}{h}$, οπότε από τον γενικευμένο κανόνα του τραπεζίου παίρνουμε

$$T(8) = T_1 = \frac{8}{2}\left[(-3)^7 + 5^7\right] = 4(-2187 + 78125) = 303752 ,$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$T(4) = T_2 = 4\left[\frac{(-3)^7}{2} + 1^7 + \frac{5^7}{2}\right] = 4(-1093.5 + 1 + 39062.5) = 151880,$$

$$T(2) = T_4 = 2\left[\frac{(-3)^7}{2} + (-1)^7 + 1^7 + 3^7 + \frac{5^7}{2}\right] = 80312\ ,$$

$$T(1) = T_8 = \left[\frac{(-3)^7}{2} + (-2)^7 + (-1)^7 + 0^7 + 1^7 + 2^7 + 3^7 + \frac{5^7}{2}\right] = 56540\ .$$

'Εχοντας υπ' όψιν τον παρακάτω πίνακα βρίσκουμε διαδοχικά

h	N				
8	1	T_1			
4	2	T_2	$T_2^{(1)}$		
2	4	T_4	$T_4^{(1)}$	$T_4^{(2)}$	
1	8	T_8	$T_8^{(1)}$	$T_8^{(2)}$	$T_8^{(3)}$

$$T_2^{(1)} = \frac{4T_2 - T_1}{3} = 101256\ ,$$

$$T_4^{(1)} = \frac{4T_4 - T_2}{3} = 56456\ ,$$

$$T_8^{(1)} = \frac{4T_8 - T_4}{3} = 48616\ ,$$

$$T_4^{(2)} = \frac{16T_4^{(1)} - T_2^{(1)}}{15} = \frac{160408}{3} \approx 53469.3333,$$

$$T_8^{(2)} = \frac{16T_8^{(1)} - T_4^{(1)}}{15} = \frac{144280}{3} \approx 48093.3333,$$

$$T_8^{(3)} = \frac{64T_8^{(2)} - T_4^{(2)}}{63} = 48008\ .$$

Η $T_8^{(3)}$ συμπίπτει με την ακριβή τιμή του ολοκληρώματος.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

5.6 Μέθοδοι αριθμητικής ολοκληρώσεως του Gauss

5.6.1 Εισαγωγή

Θα ασχοληθούμε τώρα με το γενικώτερο πρόβλημα της αριθμητικής ολοκληρώσεως, δηλαδή, με τον προσεγγιστικό υπολογισμό ολοκληρωμάτων της μορφής

$$\int_a^b w(x)f(x)dx \ . \tag{5.60}$$

Στην (5.60) $w(x)$ είναι η καλούμενη συνάρτηση βάρους, δηλαδή μία καθορισμένη συνάρτηση για την οποία υποθέτουμε ότι ισχύουν οι ιδιότητες που αναφέρονται στην παράγραφο 4.1 (σε σχέση με το διάστημα $[a,b]$, το οποίο δεν είναι αναγκαστικά πεπερασμένο). Υποθέτουμε επίσης ότι η $f(x)$ είναι συνεχής και υπάρχει το ολοκλήρωμα (5.60). Είναι εύκολο να δειχθεί ότι ισχύει το παρακάτω θεώρημα .

<u>Θεώρημα 5.1</u> Εάν δοθούν οποιαδήποτε $n+1$ διακεκριμένα σημεία x_i, $i = 0(1)n$, μπορούμε να βρούμε σταθερές (συντελεστές) A_i, $i = 0(1)n$ έτσι ώστε ο προσεγγιστικός τύπος αριθμητικής ολοκληρώσεως

$$\int_a^b w(x)f(x)dx \approx \sum_{i=0}^{n} A_i f(x_i) \tag{5.61}$$

να είναι ακριβής στην περίπτωση που η $f(x)$ είναι πολυώνυμο βαθμού $\leq n$. (Υποτίθεται ότι η $f(x)$ ορίζεται στα σημεία x_i, $i = 0(1)n$, τα οποία δεν ανήκουν αναγκαστικά στο $[a,b]$).

<u>Απόδειξη</u> Έστω $P_n(x)$ το πολυώνυμο παρεμβολής του Lagrange , με σημεία παρεμβολής τα x_i, $i = 0(1)n$, τέτοιο ώστε $P_n(x_i) = f(x_i)$, $i = 0(1)n$. Προφανώς

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$P_n(x) = \sum_{i=0}^{n} L_i(x) f(x_i) \qquad \text{(βλέπε (2.7), (2.8))} .$$

Τότε έχουμε

$$\int_a^b w(x)f(x)dx \approx \int_a^b w(x)P_n(x)dx = \sum_{i=0}^{n} \left[\int_a^b w(x)L_i(x)dx\right] f(x_i) .$$

θέτοντας $A_i = \int_a^b w(x)L_i(x)dx$, $i = 0(1)n$, προκύπτει ότι

$$\int_a^b w(x)f(x)dx \approx \sum_{i=0}^{n} A_i f(x_i) . \tag{5.62}$$

Εάν η $f(x)$ είναι ένα πολυώνυμο $R(x)$ βαθμού $\leq n$, τότε από την παρεμβολή είναι γνωστό ότι $P_n(x) \equiv R(x)$ και επομένως ο τύπος (5.62) θα είναι ακριβής για πολυώνυμα βαθμού $\leq n$. ▲

Πρέπει να σημειωθεί ότι οι τύποι των Newton – Cotes (5.5) και (5.40) προκύπτουν από τον γενικώτερο τύπο (5.61) στην περίπτωση που $w(x) \equiv 1$ και τα σημεία x_i, $i = 0(1)n$ είναι ισαπέχοντα και ανήκουν στο $[a,b]$. Επίσης, είναι δυνατόν ο τύπος (5.61) σε μερικές περιπτώσεις να είναι ακριβής για πολυώνυμα βαθμού μεγαλύτερου του n (π.χ. ο τύπος (5.5) των Newton – Cotes με n άρτιο).

Η αριθμητική ολοκλήρωση του Gauss συνίσταται στην εύρεση ενός προσεγγιστικού τύπου της μορφής (5.61), χωρίς να υποθέσουμε ότι τα σημεία x_i, $i = 0(1)n$ είναι γνωστά (δεδομένα) εκ των προτέρων. Δηλαδή, τα σημεία x_i, $i = 0(1)n$ και οι συντελεστές A_i, $i = 0(1)n$ είναι προσδιοριστέα και εκλέγονται έτσι ώστε ο τύπος (5.61) να είναι ακριβής για πολυώνυμα όσον το δυνατόν μεγαλύτερου βαθμού. (Υποτίθεται φυσικά ότι μπορούμε να υπολογίζουμε την συνάρτηση $f(x)$ σε οποιοδήποτε σημείο x).

Επειδή υπάρχουν $2n+2$ προσδιοριστέες παράμετροι στον τύπο (5.61),

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

είναι λογικό να αναμένουμε ότι με κατάλληλη εκλογή αυτών μπορεί να προκύψει ένας τύπος αριθμητικής ολοκληρώσεως ακριβής για πολυώνυμα βαθμού $\leq 2n+1$, δηλαδή, για όλους τους γραμμικούς συνδυασμούς των $2n+2$ δυνάμεων $1, x, x^2, \ldots, x^{2n+1}$. Ένας τέτοιος τύπος θα καλείται τύπος ή κανόνας αριθμητικής ολοκληρώσεως του Gauss .

Στον τύπο (5.61) μόνον η συνάρτηση $f(x)$ υπεισέρχεται στον υπολογισμό του ολοκληρώματος ενώ η συνάρτηση βάρους $w(x)$ χρησιμοποιείται μόνον για τον προσδιορισμό των A_i και x_i, $i = 0(1)n$. Όταν τα A_i και x_i έχουν προσδιορισθεί,τότε ο τύπος (5.61) μπορεί να εφαρμοσθεί για τον υπολογισμό ολοκληρωμάτων της μορφής (5.60) με διαφορετικές συναρτήσεις $f(x)$ αλλά με την ίδια συνάρτηση βάρους $w(x)$. Θα λέμε ότι η τάξη (ακρίβειας) του τύπου (5.61) είναι ℓ εάν αυτός είναι ακριβής στην περίπτωση που η $f(x)$ είναι πολυώνυμο βαθμού $\leq \ell$ ενώ δεν είναι ακριβής για $f(x) = x^{\ell+1}$.

5.6.2 Εύρεση των τύπων ολοκληρώσεως του Gauss

Σύμφωνα με το κεφάλαιο 4 υπάρχει ένα μονοσήμαντα ορισμένο (εκτός από σταθερούς παράγοντες) σύνολο ορθογωνίων πολυωνύμων $\{\varphi_\kappa(x)\}$, $\kappa = 0,1,2,\ldots$ σε σχέση με την συνάρτηση βάρους $w(x)$ στο διάστημα $[a,b]$. Ακόμη, εδείχθη ότι το πολυώνυμο $\varphi_n(x)$ βαθμού n έχει n πραγματικές διακεκριμένες ρίζες στο (a,b) . Θα αποδειχθεί τώρα το ακόλουθο θεώρημα στο οποίο βασίζεται η εύρεση των τύπων ολοκληρώσεως του Gauss.

<u>Θεώρημα 5.2</u> Έστωσαν x_i , $i = 0(1)n$ οι ρίζες του ορθογωνίου πολυωνύμου $\varphi_{n+1}(x)$ (βαθμού $n+1$) σε σχέση με την συνάρτηση βάρους $w(x)$ στο $[a,b]$. Εάν οι συντελεστές A_i , $i = 0(1)n$ στον προσεγγιστικό τύπο

$$\int_a^b w(x)f(x)dx \approx \sum_{i=0}^{n} A_i f(x_i) \tag{5.63}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

εκλεγούν σύμφωνα με το θεώρημα 5.1, τότε ο τύπος (5.63) θα είναι ακριβής για όλα τα πολυώνυμα βαθμού $\leq 2n+1$. Επί πλέον η τάξη του είναι $2n+1$ και $A_i > 0$, $i = 0(1)n$. Αντιστρόφως, εάν τα x_i και A_i, $i = 0(1)n$ στον τύπο (5.63) είναι τέτοια ώστε αυτός να είναι ακριβής για όλα τα πολυώνυμα βαθμού $\leq 2n+1$, τότε τα x_i, $i = 0(1)n$ πρέπει να είναι οι ρίζες του ορθογωνίου πολυωνύμου $\varphi_{n+1}(x)$ για την $w(x)$ στο $[a,b]$.

<u>Απόδειξη</u> Αφού οι συντελεστές A_i, $i = 0(1)n$ προσδιορίζονται σύμφωνα με το θεώρημα 5.1, τότε ο τύπος (5.63) θα είναι ακριβής για όλα τα πολυώνυμα βαθμού $\leq n$. Ας υποθέσουμε τώρα ότι η $f(x)$ είναι ένα πολυώνυμο βαθμού m όπου, $n+1 \leq m \leq 2n+1$. Τότε θα ισχύει

$$f(x) = S(x)\varphi_{n+1}(x) + R(x) , \tag{5.64}$$

όπου το πηλίκο $S(x)$ είναι ένα πολυώνυμο βαθμού $m-(n+1)$ και το υπόλοιπο $R(x)$ είναι ένα πολυώνυμο βαθμού $\leq n$. Προφανώς $0 \leq m-(n+1) \leq n$. Λόγω της (5.64) έχουμε

$$\int_a^b w(x)f(x)dx = \int_a^b w(x)\left[S(x)\varphi_{n+1}(x) + R(x)\right]dx =$$

$$= \int_a^b w(x)S(x)\varphi_{n+1}(x)dx + \int_a^b w(x)R(x)dx = \int_a^b w(x)R(x)dx , \tag{5.65}$$

αφού $\int_a^b w(x)S(x)\varphi_{n+1}(x)dx = 0$ λόγω του πορίσματος 4.1.

Από την (5.64) προκύπτει ότι

$$f(x_i) = S(x_i)\varphi_{n+1}(x_i) + R(x_i) = R(x_i) , \quad i = 0(1)n \tag{5.66}$$

Ακόμη, αφού το $R(x)$ είναι πολυώνυμο βαθμού $\leq n$, ο τύπος (5.63) θα είναι ακριβής για το $R(x)$, δηλαδή,

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\int_a^b w(x)R(x)dx = \sum_{i=0}^{n} A_i R(x_i) \ . \tag{5.67}$$

Επομένως η (5.65) λόγω των (5.66) και (5.67) γίνεται

$$\int_a^b w(x)f(x)dx = \sum_{i=0}^{n} A_i R(x_i) = \sum_{i=0}^{n} A_i f(x_i) \ , \tag{5.68}$$

που σημαίνει ότι ο τύπος (5.63) είναι ακριβής για πολυώνυμα βαθμού $\leq 2n+1$.

Θα αποδειχθεί τώρα ότι για οποιαδήποτε εκλογή των $n+1$ διακεκριμένων σημείων x_i, $i = 0(1)n$ (παραβλέπουμε προς το παρόν το γεγονός ότι τα x_i, $i = 0(1)n$ είναι οι ρίζες του $\varphi_{n+1}(x)$) και των συντελεστών A_i, $i = 0(1)n$ είναι αδύνατον να βρεθεί τύπος της μορφής (5.63) που να είναι ακριβής για πολυώνυμα βαθμού $2n+2$. Πράγματι, αν θεωρήσουμε το πολυώνυμο $2n+2$ βαθμού

$$P(x) = \prod_{i=0}^{n}(x - x_i)^2 \ , \tag{5.69}$$

παρατηρούμε ότι ο τύπος (5.63) δεν είναι ακριβής για το $P(x)$, αφού το πρώτο μέλος του τύπου (5.63) είναι

$$\int_a^b w(x)P(x)dx > 0 \ , \tag{5.70}$$

ενώ το δεύτερο μέλος είναι

$$\sum_{i=0}^{n} A_i P(x_i) = 0 \ , \tag{5.71}$$

για κάθε εκλογή των A_i . Η (5.70) ισχύει διότι λόγω των υποθέσεων για την $w(x)$ και λόγω της (5.69) η συνάρτηση $w(x)P(x)$ είναι ≥ 0 στο $[a,b]$ και μηδενίζεται σ' αυτό το διάστημα το πολύ σε ένα πεπε-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ρασμένο αριθμό σημείων. Αφού λοιπόν υπάρχει τουλάχιστον ένα πολυώνυμο βαθμού $2n+2$ για το οποίο ο (5.63) δεν είναι ακριβής, η τάξη του θα είναι $2n+1$. ▲

Θα αποδειχθεί τώρα ότι $A_i > 0$, $i = 0(1)n$. Θεωρούμε τα ακόλουθα πολυώνυμα βαθμού $2n$

$$P_j(x) = \prod_{\substack{\kappa=0 \\ \kappa\neq j}}^{n}(x - x_\kappa)^2, \; j = 0(1)n \;, \tag{5.72}$$

για τα οποία βεβαίως ο τύπος (5.63) θα είναι ακριβής. Έτσι έχουμε

$$\int_a^b w(x)P_j(x)dx = \sum_{i=0}^{n} A_i P_j(x_i) = A_j P_j(x_j) = A_j \prod_{\substack{\kappa=0 \\ \kappa\neq j}}^{n}(x_j - x_\kappa)^2, \quad j = 0(1)n. \tag{5.73}$$

Επειδή $\int_a^b w(x)P_j(x)dx > 0$ και $\prod_{\substack{\kappa=0 \\ \kappa\neq j}}^{n}(x_j - x_\kappa)^2 > 0$, από την (5.73) συνεπάγεται ότι $A_j > 0$, $j = 0(1)n$.

Ας υποθέσουμε τώρα ότι στον ακόλουθο τύπο

$$\int_a^b w(x)f(x)dx \approx \sum_{i=0}^{n} \hat{A}_i f(\hat{x}_i) \tag{5.74}$$

τα $\hat{A}_i$, $\hat{x}_i$, $i = 0(1)n$ είναι τέτοια ώστε αυτός να είναι ακριβής για πολυώνυμα βαθμού $\leq 2n+1$. Θεωρούμε το πολυώνυμο

$$V(x) = \prod_{i=0}^{n}(x - \hat{x}_i) \;, \tag{5.75}$$

το οποίο είναι βαθμού $n+1$. Αν $Z(x)$ είναι οποιοδήποτε πολυώνυμο βαθμού n, τότε ο τύπος (5.74) θα είναι ακριβής για το πολυώνυμο $V(x)Z(x)$ βαθμού $2n+1$. Επομένως θα ισχύει

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\int_a^b w(x)V(x)Z(x)dx = \sum_{i=0}^{n} \hat{A}_i V(\hat{x}_i)Z(\hat{x}_i) = 0 \ , \tag{5.76}$$

αφού $V(\hat{x}_i) = 0$, $i = 0(1)n$. Από την (5.76), λόγω του πορίσματος 4.2, προκύπτει ότι $V(x) \equiv c_{n+1}\varphi_{n+1}(x)$, όπου c_{n+1} σταθερός παράγων και $\varphi_{n+1}(x)$ το ορθογώνιο πολυώνυμο βαθμού $n+1$ σε σχέση με την συνάρτηση βάρους $w(x)$ στο $[a,b]$. Άρα τα $\hat{x}_i$, $i = 0(1)n$ θα είναι οι ρίζες του $\varphi_{n+1}(x)$. ▲

Τύποι της μορφής (5.63) που κατασκευάζονται σύμφωνα με το θεώρημα 5.2 καλούνται τύποι του Gauss. Τα x_i, $i = 0(1)n$ είναι οι ρίζες του πολυωνύμου $\varphi_{n+1}(x)$ (βαθμού $n+1$) του ορθογωνίου συνόλου πολυωνύμων $\{\varphi_k(x)\}$, $k = 0,1,2,\ldots$ σε σχέση με την συνάρτηση βάρους $w(x)$ στο διάστημα $[a,b]$, οι δε συντελεστές A_i δίνονται από τις

$$A_i = \int_a^b w(x)L_i(x)dx \ , \quad i = 0(1)n \ , \tag{5.77}$$

όπου $L_i(x)$ είναι οι συντελεστές παρεμβολής του Lagrange για την $f(x)$ με σημεία παρεμβολής τα x_i , $i = 0(1)n$. Λαμβάνοντας υπ' όψιν την (2.12), η (5.77) μπορεί να γραφεί

$$A_i = \int_a^b w(x) \frac{\varphi_{n+1}(x)}{(x - x_i)\varphi'_{n+1}(x_i)} dx \ , \ i = 0(1)n \ . \tag{5.78}$$

Σημειώνεται ότι ο συντελεστής του μεγιστοβαθμίου όρου του $\varphi_{n+1}(x)$ δεν είναι αναγκαστικά η μονάδα. Αν υποθέσουμε ότι c_n είναι ο συντελεστής του x^n στο $\varphi_n(x)$, $n = 0,1,2,\ldots$, μπορεί να αποδειχθεί ότι ο τύπος (5.78) μετασχηματίζεται στον ακόλουθο

$$A_i = \frac{c_{n+1}}{c_n} \frac{\int_a^b w(x)\left[\varphi_n(x)\right]^2 dx}{\varphi'_{n+1}(x_i)\varphi_n(x_i)} \ , \ i = 0(1)n \ . \tag{5.79}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Στη συνέχεια θα εξετασθούν οι τύποι του Gauss που προκύπτουν αν χρησιμοποιηθούν τα γνωστά από το κεφάλαιο 4 ορθογώνια πολυώνυμα Legendre και Chebyshev. Αυτοί καλούνται αντίστοιχα τύποι αριθμητικής ολοκληρώσεως των Gauss – Legendre και Gauss – Chebyshev .

5.6.3 Τύποι ολοκληρώσεως των Gauss-Legendre

Όπως είναι γνωστό από το κεφάλαιο 4 τα πολυώνυμα Legendre $\Phi_\kappa(x)$, $\kappa = 0,1,2,\ldots$ είναι ορθογώνια στο διάστημα $[-1,1]$ σε σχέση με την συνάρτηση βάρους $w(x) \equiv 1$. Έτσι από την (5.63) για $w(x) \equiv 1$ και $[a,b] \equiv [-1,1]$ προκύπτουν οι τύποι

$$\int_{-1}^{1} f(x)dx \approx \sum_{i=0}^{n} A_i f(x_i) \ , \tag{5.80}$$

όπου τα x_i, $i = 0(1)n$ είναι οι ρίζες του $\Phi_{n+1}(x)$ και τα A_i μπορούν να προσδιορισθούν από την (5.79) για $\varphi_n(x) \equiv \Phi_n(x)$, $w(x) \equiv 1$ και $[a,b] \equiv [-1,1]$. Έχοντας υπ' όψιν τις ιδιότητες των πολυωνύμων Legendre παίρνουμε

$$A_i = \frac{2}{(n+1)\Phi_n(x_i)\Phi'_{n+1}(x_i)} \ , \ i = 0(1)n \ . \tag{5.81}$$

Μπορεί να δειχθεί με κατάλληλους μετασχηματισμούς και χρησιμοποιώντας τις ιδιότητες των πολυωνύμων Legendre ότι ένας πιο κατάλληλος τύπος υπολογισμού των A_i είναι ο ακόλουθος

$$A_i = \frac{2}{(1-x_i)^2\left(\Phi'_{n+1}(x_i)\right)^2} \ , \ i = 0(1)n \ . \tag{5.82}$$

Για $n = 0$ είναι $\Phi_1(x) = x$, $\Phi'_1(x) = 1$, $x_0 = 0$, $A_0 = 2$ και επομένως από την (5.80) έχουμε

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\int_{-1}^{1} f(x)dx \approx 2f(0) \ . \tag{5.83}$$

Ο τύπος (5.83) είναι ο γνωστός κανόνας του μέσου (βλέπε (5.34)) .

Για $n=1$ είναι $\Phi_2(x) = \frac{1}{2}(3x^2-1)$, $\Phi_2'(x) = 3x$, $x_0 = -\frac{1}{\sqrt{3}}$, $x_1 = \frac{1}{\sqrt{3}}$, $A_0 = \frac{2}{(1-x_0^2)(3x_0)^2} = 1$, $A_1 = \frac{2}{(1-x_1^2)(3x_1)^2} = 1$. Συνεπώς για $n=1$ ο κανόνας των Gauss – Legendre είναι

$$\int_{-1}^{1} f(x)dx \approx f\left(-\frac{1}{\sqrt{3}}\right) + f\left(\frac{1}{\sqrt{3}}\right) \tag{5.84}$$

και είναι ακριβής για πολυώνυμα βαθμού $\leqq 3$.

Για $n=2$ είναι $\Phi_3(x) = \frac{1}{2}(5x^3-3x)$ και μπορούμε να βρούμε ανάλογα ότι

$$x_0 = -\sqrt{0.6} \ , \ x_1 = 0 \ , \ x_2 = \sqrt{0.6} \ , \ A_0 = \frac{5}{9} \ , \ A_1 = \frac{8}{9} \ , \ A_2 = \frac{5}{9} \ ,$$

$$\int_{-1}^{1} f(x)dx \approx \frac{1}{9}\left(5f(-\sqrt{0.6}) + 8f(0) + 5f(\sqrt{0.6})\right) \ . \tag{5.85}$$

Η τάξη του τύπου (5.85) είναι 5 . Τα σημεία (ρίζες) x_i, $i = 0(1)n$ και οι αντίστοιχοι συντελεστές A_i, $i = 0(1)n$ για τις διάφορες τιμές του n μπορούν να βρεθούν σε πίνακες.

Ο τύπος (5.80) μπορεί να εφαρμοστεί για τον υπολογισμό ολοκληρωμάτων και όταν το διάστημα ολοκληρώσεως είναι διαφορετικό από το $[-1,1]$. Έτσι εάν $[c,d]$ είναι ένα αυθαίρετο διάστημα και ζητούμε να υπολογίσουμε το ολοκλήρωμα

$$\int_{c}^{d} F(t)dt \ , \tag{5.86}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

εκτελούμε τον μετασχηματισμό $t = \frac{d-c}{2}x + \frac{c+d}{2}$, οπότε η (5.86) γίνεται

$$\int_c^d F(t)dt = \frac{d-c}{2}\int_{-1}^{1} F\left(\frac{d-c}{2}x + \frac{c+d}{2}\right)dx = \frac{d-c}{2}\int_{-1}^{1} f(x)dx \ , \quad (5.87)$$

όπου $F\left(\frac{d-c}{2}x + \frac{c+d}{2}\right) \equiv f(x)$. Εφαρμόζοντας τον τύπο (5.80) στην (5.87) παίρνουμε

$$\int_c^d F(t)dt \approx \frac{d-c}{2}\sum_{i=0}^{n} A_i f(x_i) = \frac{d-c}{2}\sum_{i=0}^{n} A_i F(t_i) \ , \quad (5.88)$$

όπου $t_i = \frac{d-c}{2}x_i + \frac{c+d}{2}$, $i = 0(1)n$.

Για παράδειγμα, αν $n=1$ θα έχουμε

$$\int_c^d F(t)dt \approx \frac{d-c}{2}\left(F(t_0) + F(t_1)\right) \ ,$$

όπου $t_0 = \frac{d-c}{2}\left(-\frac{1}{\sqrt{3}}\right) + \frac{c+d}{2}$, $t_1 = \frac{d-c}{2}\left(\frac{1}{\sqrt{3}}\right) + \frac{c+d}{2}$.

5.6.4 Τύποι ολοκληρώσεως των Gauss–Chebyshev

Είναι γνωστό ότι τα πολυώνυμα Chebyshev $T_\kappa(x)$, $\kappa = 0,1,2,\ldots$ είναι ορθογώνια στο διάστημα $[-1,1]$ σε σχέση με την συνάρτηση βάρους $w(x) \equiv (1-x^2)^{-\frac{1}{2}}$. Επομένως οι τύποι των Gauss – Chebyshev θα έχουν την μόρφη

$$\int_{-1}^{1}(1-x^2)^{-\frac{1}{2}}f(x)dx \approx \sum_{i=0}^{n} A_i f(x_i) \ , \quad (5.89)$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

όπου τα x_i , $i = 0(1)n$ είναι οι ρίζες του $T_{n+1}(x)$,οι οποίες σύμφωνα με το θεώρημα 4.14 δίνονται από τον τύπο

$$x_i = \cos\left(\frac{(2i+1)\pi}{2(n+1)}\right) \quad , \; i = 0(1)n \tag{5.90}$$

και τα A_i , $i = 0(1)n$ προκύπτουν από τον (5.79) για $\varphi_n(x) \equiv T_n(x)$, $w(x) \equiv (1-x^2)^{-\frac{1}{2}}$ και $[a,b] \equiv [-1,1]$. Είναι δυνατόν να δειχθεί ότι

$$A_i = \frac{\pi}{n+1} \; , \; i = 0(1)n \quad , \tag{5.91}$$

δηλαδή όλα τα A_i είναι ίσα μεταξύ τους. Έτσι η (5.89) λαμβάνει την μορφή

$$\int_{-1}^{1}(1-x^2)^{-\frac{1}{2}} f(x)dx \approx \frac{\pi}{n+1} \sum_{i=0}^{n} f\left(\cos\frac{(2i+1)\pi}{2(n+1)}\right) \quad . \tag{5.92}$$

Για $n = 0,1,2,$ από τον τύπο (5.92) προκύπτουν αντίστοιχα οι τύποι

$$\int_{-1}^{1}(1-x^2)^{-\frac{1}{2}}f(x)dx \approx \pi f(0) \; , \tag{5.93}$$

$$\int_{-1}^{1}(1-x^2)^{-\frac{1}{2}}f(x)dx \approx \frac{\pi}{2}\left[f\left(-\frac{1}{\sqrt{2}}\right) + f\left(\frac{1}{\sqrt{2}}\right)\right] \; , \tag{5.94}$$

$$\int_{-1}^{1}(1-x^2)^{-\frac{1}{2}}f(x)dx \approx \frac{\pi}{3}\left[f\left(-\frac{\sqrt{3}}{2}\right) + f(0) + f\left(\frac{\sqrt{3}}{2}\right)\right] \quad . \tag{5.95}$$

Πρέπει να σημειωθεί ότι ο τύπος (5.92) μπορεί να εφαρμοστεί σε ένα ολοκλήρωμα της μορφής $\int_{-1}^{1}F(x)dx$, αν γράψουμε

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$F(x) = (1-x^2)^{-\frac{1}{2}}\left[(1-x^2)^{\frac{1}{2}}F(x)\right] \equiv (1-x^2)^{-\frac{1}{2}}f(x) \ .$$

Οι τύποι ολοκληρώσεως του Gauss παρά την απλή μορφή τους είναι δύσχρηστοι για υπολογισμούς με το χέρι, επειδή γενικά τα x_i και A_i είναι άρρητοι αριθμοί και πρέπει να υπολογίζονται με κατάλληλη ακρίβεια. Η εφαρμογή τους όμως με την βοήθεια Ηλεκτρονικού υπολογιστή είναι απλή και γρήγορη. Το πλεονέκτημα των τύπων του Gauss είναι ότι με την χρησιμοποίηση μικρού σχετικά αριθμού τιμών της συναρτήσεως $f(x)$ επιτυγχάνεται μεγάλη τάξη ακρίβειας. (Συγκρίνατε π.χ. τον κανόνα του Simpson με τον (5.85)). Σημειώνεται ότι μπορούν και αυτοί, όπως οι τύποι των Newton – Cotes, να γενικευθούν (γενικευμένοι κανόνες του Gauss). Ακόμη, μπορεί να αποδειχθεί ότι αν το $[a,b]$ είναι πεπερασμένο και η $f(x)$ συνεχής στο $[a,b]$ τότε η ακολουθία των τύπων του Gauss για $n \to \infty$ συγκλίνει πάντοτε στην ακριβή τιμή του

$$\int_a^b w(x)f(x)dx \ .$$

5.

<u>Παράδειγμα 5.3</u> Να προσεγγιστεί το ολοκλήρωμα $\int_0^1 \frac{dx}{1+x^4}$ (βλέπε και παράδειγμα 5.1) με τους τύπους Gauss – Legendre για $n = 1,2$.

<u>Λύση</u> Έχουμε $I = \int_0^1 \frac{dt}{1+t^4}$. Θέτουμε $t = \frac{1}{2}x + \frac{1}{2} = \frac{1}{2}(x+1)$,

οπότε

$$\int_0^1 \frac{dt}{1+t^4} = \frac{1}{2}\int_{-1}^1 \frac{dx}{1+\left(\frac{1}{2}(x+1)\right)^4} \ .$$

i) Για $n = 1$ (βλέπε (5.84)) έχουμε

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\frac{1}{2}\int_{-1}^{1}\frac{1}{1+\left(\frac{1}{2}(x+1)\right)^4}dx \approx \frac{1}{2}\left[\frac{1}{1+\left[\frac{1}{2}\left(-\frac{1}{\sqrt{3}}+1\right)\right]^4}+\frac{1}{1+\left[\frac{1}{2}\left(\frac{1}{\sqrt{3}}+1\right)\right]^4}\right]\approx$$

≈ 0.859522 .

ii) Για $n = 2$ (βλέπε (5.85)) παίρνουμε

$$\frac{1}{2}\int_{-1}^{1}\frac{1}{1+\left(\frac{1}{2}(x+1)\right)^4}dx \approx \frac{1}{2}\,\frac{1}{9}\left[5\,\frac{1}{1+\left(\frac{1}{2}(-\sqrt{0.6}+1)\right)^4}+8\,\frac{1}{1+\left(\frac{1}{2}\right)^4}+\right.$$

$$\left.+\,5\,\frac{1}{1+\left(\frac{1}{2}(\sqrt{0.6}+1)\right)^4}\right]\approx 0.867518 \;. \qquad \blacktriangle$$

<u>Παράδειγμα 5.4</u> Να προσεγγιστεί το ολοκλήρωμα $\int_0^1 \frac{t^2}{\sqrt{t(1-t)}}dt$ με τον τύπο των Gauss – Chebyshev για $n = 2$.

<u>Λύση</u> Θέτουμε $t = \frac{1}{2}x + \frac{1}{2} = \frac{1}{2}(x+1)$, οπότε

$$\int_0^1 \frac{t^2}{\sqrt{t(1-t)}}dt = \frac{1}{4}\int_{-1}^{1}\frac{(x+1)^2}{\sqrt{1-x^2}}dx = \frac{1}{4}\int_{-1}^{1}(1-x^2)^{-\frac{1}{2}}f(x)dx \;, \quad \text{όπου}$$

$f(x) = (x+1)^2$. Άρα, λόγω του τύπου (5.95), παίρνουμε

$$\frac{1}{4}\int_{-1}^{1}(1-x^2)^{-\frac{1}{2}}f(x)dx \approx \frac{1}{4}\;\frac{\pi}{3}\left[\left(-\frac{\sqrt{3}}{2}+1\right)^2+1+\left(\frac{\sqrt{3}}{2}+1\right)^2\right] = \frac{3\pi}{8}\approx$$

≈ 1.178097 . Επειδή η $f(x)$ είναι πολυώνυμο δεύτερου βαθμού ο τύ-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

πος Gauss – Chebyshev για $n = 2$ είναι ακριβής. Πρέπει να σημειωθεί ότι και ο τύπος (5.94) $(n = 1)$ δίνει επίσης την ακριβή τιμή. Πράγματι, είναι

$$\frac{1}{4}\int_{-1}^{1}(1-x^2)^{-\frac{1}{2}}f(x)dx \approx \frac{1}{4}\frac{\pi}{2}\left[\left(-\frac{1}{\sqrt{2}}+1\right)^2+\left(\frac{1}{\sqrt{2}}+1\right)^2\right] =$$

$$= \frac{\pi}{8}\left[2\left(\frac{1}{2}+1\right)\right] = \frac{3\pi}{8} \quad . \qquad \blacktriangle$$

5.6.5 Σφάλμα αποκοπής των τύπων ολοκληρώσεως του Gauss

Το σφάλμα αποκοπής του τύπου (5.63) δίνεται ως γνωστόν από την διαφορά

$$E_n(f) = \sum_{i=0}^{n} A_i f(x_i) - \int_a^b w(x)f(x)dx \ . \tag{5.96}$$

θα αποδειχθεί τώρα ένα θεώρημα που δίνει μια έκφραση για το σφάλμα αποκοπής (5.96).

<u>Θεώρημα 5.3</u> Εάν η συνάρτηση $f(x)$ έχει συνεχή παράγωγο τάξεως $2n+2$ στο $[a,b]$, τότε

$$E_n(f) = -\frac{f^{(2n+2)}(\xi)}{(2n+2)!}\int_a^b w(x)\varphi_{n+1}^2(x)dx \ , \quad \xi \in [a,b] \ , \tag{5.97}$$

όπου $\varphi_{n+1}(x)$ είναι το πολυώνυμο βαθμού $n+1$, με συντελεστή μεγιστοβαθμίου όρου την μονάδα, ορθογώνιο στο διάστημα $[a,b]$ σε σχέση με την συνάρτηση βάρους $w(x)$.

<u>Απόδειξη</u> Θεωρούμε το πολυώνυμο παρεμβολής $P(x)$ του Hermite

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

βαθμού το πολύ $2n+1$ τέτοιο ώστε

$$P(x_i) = f(x_i) \ , \quad P'(x_i) = f'(x_i) \ , \quad i = 0(1)n \ . \tag{5.98}$$

Αφού το $P(x)$ είναι βαθμού $\leq 2n+1$ ο τύπος (5.63) θα είναι ακριβής για το $P(x)$ και επομένως

$$\int_a^b w(x)P(x)dx = \sum_{i=0}^{n} A_i P(x_i) = \sum_{i=0}^{n} A_i f(x_i) \ . \tag{5.99}$$

Από τις (5.96) και (5.99) προκύπτει ότι

$$E_n(f) = \int_a^b w(x)P(x)dx - \int_a^b w(x)f(x)dx = \int_a^b w(x)\Big[P(x) - f(x)\Big]dx \tag{5.100}$$

και από την (2.100) έχουμε

$$P(x) - f(x) = -\frac{f^{(2n+2)}(\theta)}{(2n+2)!} \prod_{i=0}^{n}(x - x_i)^2 \ , \tag{5.101}$$

όπου $\theta = \theta(x)$ ανήκει στο μικρότερο διάστημα που περιέχει τα x_0, $x_1, \ldots, x_n, x$. Αφού τα x_i , $i = 0(1)n$ είναι οι ρίζες του $\varphi_{n+1}(x)$, θα έχουμε $x_i \in (a,b)$, $i = 0(1)n$ και

$$\prod_{i=0}^{n}(x - x_i)^2 = \varphi_{n+1}^2(x) \ . \tag{5.102}$$

Έτσι, η (5.100) λόγω των (5.101) και (5.102) γίνεται

$$E_n(f) = -\int_a^b w(x) \frac{f^{(2n+2)}(\theta)}{(2n+2)!} \varphi_{n+1}^2(x)dx \ , \tag{5.103}$$

όπου $\theta \in [a,b]$. Αφού $w(x)\varphi_{n+1}^2(x) \geq 0$ και $f^{(2n+2)}(\theta)$ συνεχής στο $[a,b]$, σύμφωνα με το γενικό θεώρημα μέσης τιμής, από την (5.103) προκύπτει ότι

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$E_n(f) = -\frac{f^{(2n+2)}(\xi)}{(2n+2)!}\int_a^b w(x)\varphi_{n+1}^2(x)dx \ ,$$

για κάποιο $\xi \in [a,b]$. ▲

Για την εύρεση τώρα των σφαλμάτων αποκοπής των τύπων Gauss – Legendre και Gauss – Chebyshev που δίνονται από τις (5.80) και (5.89) αντίστοιχα, εργαζόμεθα με βάση τον γενικό τύπο (5.97). Επειδή το πολυώνυμο Legendre $\Phi_{n+1}(x)$ έχει συντελεστή μεγιστοβαθμίου όρου $\frac{(2n+1)!!}{(n+1)!}$ (βλέπε θεώρημα 4.7), από την (5.97) για $w(x) \equiv 1$ και $[a,b] = [-1,1]$ παίρνουμε

$$E_n(f) = -\frac{f^{(2n+2)}(\xi)}{(2n+2)!}\int_{-1}^{1}\left[\frac{(n+1)!}{(2n+1)!!}\Phi_{n+1}(x)\right]^2 dx =$$

$$= -\left[\frac{(n+1)!}{(2n+1)!!}\right]^2 \frac{f^{(2n+2)}(\xi)}{(2n+2)!}\int_{-1}^{1}\Phi_{n+1}^2(x)dx =$$

$$= -\frac{2}{(2n+3)(2n+2)!}\left[\frac{(n+1)!}{(2n+1)!!}\right]^2 f^{(2n+2)}(\xi) \ , \quad \xi \in [-1,1], \quad (5.104)$$

όπου χρησιμοποιήθηκε το θεώρημα 4.8 . Δηλαδή, το σφάλμα αποκοπής των τύπων Gauss – Legendre (5.80) είναι τελικά

$$E_n(f) = -2\left[\frac{(n+1)!}{(2n+1)!!}\right]^2 \frac{f^{(2n+2)}(\xi)}{(2n+3)!} \ , \quad \xi \in [-1,1] \ . \qquad (5.105)$$

Είναι εύκολο να δειχθεί (βλέπε άσκηση 35) ότι το σφάλμα αποκοπής, όταν ο τύπος (5.80) εφαρμόζεται για τον υπολογισμό του ολοκληρώματος (5.86), λαμβάνει την μορφή

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$E_n(F) = -\left(\frac{d-c}{2}\right)^{2n+3}\left(\frac{(n+1)!}{(2n+1)!!}\right)^2 \frac{F^{(2n+2)}(\xi)}{(2n+3)!}, \xi \in [c,d]. \quad (5.106)$$

Όσον αφορά το σφάλμα αποκοπής των τύπων Gauss – Chebyshev (5.89), με ανάλογο τρόπο και στηριζόμενοι στα θεωρήματα 4.16 και 4.18 βρίσκουμε από τον (5.97)

$$E_n(f) = -\frac{f^{(2n+2)}(\xi)}{(2n+2)!}\int_{-1}^{1}(1-x^2)^{-\frac{1}{2}}\left[2^{-n}T_{n+1}(x)\right]^2 dx =$$

$$= -\frac{\pi}{2^{2n+1}(2n+2)!} f^{(2n+2)}(\xi), \quad \xi \in [-1,1]. \quad (5.107)$$

5.7 Αριθμητική ολοκλήρωση του Chebyshev

Θεωρούμε πάλι το πρόβλημα του προσεγγιστικού υπολογισμού του ολοκληρώματος (5.60) με τις ίδιες προυποθέσεις για την συνάρτηση βάρους $w(x)$. Ο Chebyshev εισηγήθηκε την χρησιμοποίηση ενός τύπου αριθμητικής ολοκληρώσεως της μορφής (5.61) με τις ακόλουθες προυποθέσεις: i) Όλοι οι συντελεστές είναι ίσοι μεταξύ τους, δηλαδή, $A_i = A$, $i = 0(1)n$ και ii) Τα x_i, $i = 0(1)n$ και A είναι προσδιοριστέα και εκλέγονται έτσι ώστε ο υπ' όψιν τύπος να είναι ακριβής για πολυώνυμα όσον το δυνατόν μεγαλύτερου βαθμού.

Ο τύπος του Chebyshev έχει την μορφή

$$\int_a^b w(x)f(x)dx \approx A\sum_{i=0}^{n} f(x_i). \quad (5.108)$$

Στον τύπο (5.108) υπάρχουν $n+2$ προσδιοριστέες παράμετροι και επομένως η τάξη του θα είναι τουλάχιστον $n+1$. Για την εύρεση των x_i,

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$i = 0(1)n$ και A απαιτούμε ο τύπος (5.108) να είναι ακριβής για $f(x) = x^K$, $\kappa = 0(1)n+1$ (οπότε θα είναι ακριβής για όλους τους γραμμικούς συνδυασμούς των $n+2$ δυνάμεων $1, x, x^2, \ldots, x^{n+1}$). Πρέπει να σημειωθεί ότι οι τύποι των Gauss – Chebyshev που βρέθηκαν στην παράγραφο 5.6.4 αποτελούν ειδική περίπτωση του τύπου (5.108). Συγκεκριμένα, για $w(x) = (1-x)^{-\frac{1}{2}}$ και $[a,b] = [-1,1]$ το πρόβλημα του προσδιορισμού των παραμέτρων στον τύπο (5.108) λύνεται με τους τύπους των Gauss – Chebyshev (τα x_i είναι οι ρίζες του $T_{n+1}(x)$ και $A = \frac{\pi}{n+1}$) .

Θα εξετάσουμε τώρα την περίπτωση $w(x) = 1$ και $[a,b] = [-1,1]$, οπότε ο (5.108) γίνεται

$$\int_{-1}^{1} f(x)dx \approx A \sum_{i=0}^{n} f(x_i) \ . \tag{5.109}$$

θέτοντας στην (5.109) $f(x) = x^K$, $\kappa = 0(1)n+1$ και απαιτώντας να ισχύει η ισότητα παίρνουμε το ακόλουθο σύστημα εξισώσεων

$$\begin{cases} A(n+1) = 2 \\ A(x_0 + x_1 + \ldots + x_n) = 0 \\ A(x_0^2 + x_1^2 + \ldots + x_n^2) = \frac{2}{3} \\ A(x_0^3 + x_1^3 + \ldots + x_n^3) = 0 \\ \vdots \qquad\qquad \vdots \\ A(x_0^{n+1} + x_1^{n+1} + \ldots + x_n^{n+1}) = \dfrac{1-(-1)^{n+2}}{n+2} \ . \end{cases} \tag{5.110}$$

Από την πρώτη εξίσωση του (5.110) προκύπτει ότι

$$A = \frac{2}{n+1} \ , \tag{5.111}$$

το δε σύστημα των υπολοίπων $n+1$ εξισώσεων με αγνώστους τα x_i ,

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$i = 0(1)n$ μπορεί να γραφεί συνοπτικά, αφού χρησιμοποιηθεί η (5.111), ως εξής:

$$\begin{cases} \sum_{i=0}^{n} x_i^{\kappa} = 0 \quad , \quad \kappa = 1(1)n+1 \ , \ \kappa \text{ περιττός} \\ \sum_{i=0}^{n} x_i^{\kappa} = \frac{n+1}{\kappa+1} \quad , \quad \kappa = 1(1)n+1 \ , \ \kappa \text{ άρτιος} . \end{cases} \tag{5.112}$$

Για τον υπολογισμό των x_i , $i = 0(1)n$, αντί να λύσουμε το σύστημα (5.112), υποθέτουμε ότι

$$P(x) = x^{n+1} + \alpha_1 x^n + \alpha_2 x^{n-1} + \ldots + \alpha_n x + \alpha_{n+1} \ , \tag{5.113}$$

είναι το πολυώνυμο που έχει ρίζες τα x_i , $i = 0(1)n$, δηλαδή $P(x) \equiv \equiv \prod_{i=0}^{n}(x - x_i)$. Εάν βρεθούν τα α_i , $i = 1(1)n+1$ στην (5.113), τότε τα x_i , $i = 0(1)n$ προσδιορίζονται από την λύση της πολυωνυμικής εξισώσεως $P(x) = 0$. Όπως είναι γνωστό από την Άλγεβρα τα α_i , $i = 1(1)n+1$ συνδέονται με τους αριθμούς

$$S_{\kappa} = \sum_{i=0}^{n} x_i^{\kappa} \quad , \quad \kappa = 1(1)n+1 \quad , \tag{5.114}$$

οι οποίοι δίνονται από τις (5.112), με τις σχέσεις (τύποι του Newton)

$$\begin{cases} S_1 + 1\alpha_1 = 0 \\ S_2 + \alpha_1 S_1 + 2\alpha_2 = 0 \\ S_3 + \alpha_1 S_2 + \alpha_2 S_1 + 3\alpha_3 = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ S_{n+1} + \alpha_1 S_n + \alpha_2 S_{n-1} + \ldots + (n+1)\alpha_{n+1} = 0 \ . \end{cases} \tag{5.115}$$

Έτσι, λύνοντας το σύστημα (5.115) προσδιορίζονται τα α_i , $i =$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$= 1(1)n+1$ και επομένως το πολυώνυμο (5.113). Για να μπορεί να κατασκευασθεί ο τύπος (5.109) πρέπει τα x_i , $i = 0(1)n$ που θα προσδιορίζονται σαν οι ρίζες του πολυωνύμου (5.113) να είναι πραγματικά και διακεκριμένα. Έχει αποδειχθεί ότι για $n = 0,1,2,\ldots,6$ και $n = 8$ οι ρίζες του πολυωνύμου (5.113) είναι πραγματικές, διακεκριμένες και ανήκουν στο $[-1,1]$. Για $n = 7$ και $n \geq 9$ απεδείχθη ότι μεταξύ των ριζών του (5.113) υπάρχουν και μιγαδικές ρίζες και επομένως για τις τιμές αυτές του n δεν υπάρχουν τύποι του Chebyshev της μορφής (5.109). Το γεγονός αυτό αποτελεί και το βασικό μειονέκτημα της αριθμητικής ολοκληρώσεως του Chebyshev.

Για $n = 0$ είναι $A = 2$ και η (5.114) λόγω των (5.112) δίνει $S_1 = 0$. Από το σύστημα (5.115) που για $n = 0$ περιέχει μόνον την πρώτη εξίσωση προκύπτει ότι $\alpha_1 = 0$. Έτσι, από την (5.113) είναι $P(x) = x$ και επομένως $x_0 = 0$. Άρα, για $n = 0$ ο τύπος του Chebyshev είναι

$$\int_{-1}^{1} f(x)dx \approx 2f(0) \tag{5.116}$$

και συμπίπτει με τον τύπο (5.83) των Gauss – Legendre $(n = 0)$.

Για $n = 1$ είναι $A = 1$, $S_1 = 0$, $S_2 = \frac{2}{3}$, $\alpha_1 = 0$, $\alpha_2 = -\frac{S_2}{2} = -\frac{1}{3}$ και τα x_0, x_1 είναι οι ρίζες της $x^2 - \frac{1}{3} = 0$, δηλαδή, $x_0 = -\frac{1}{\sqrt{3}}$, $x_1 = \frac{1}{\sqrt{3}}$. Έτσι, ο τύπος (5.109) για $n = 1$ γίνεται

$$\int_{-1}^{1} f(x)dx \approx f\left(-\frac{1}{\sqrt{3}}\right) + f\left(\frac{1}{\sqrt{3}}\right) \tag{5.117}$$

και συμπίπτει με τον τύπο (5.84) των Gauss – Legendre για $n = 1$.

Για $n = 2$ μπορεί ανάλογα να βρεθεί ότι

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\int_{-1}^{1} f(x)dx \approx \frac{2}{3}\left[f\left(-\frac{1}{\sqrt{2}}\right)+f(0)+f\left(\frac{1}{\sqrt{2}}\right)\right] . \tag{5.118}$$

Είναι φανερό ότι η τάξη του τύπου (5.109) είναι τουλάχιστον $n+1$. Μπορεί να αποδειχθεί ότι το σφάλμα αποκοπής του τύπου (5.109) δίνεται από την έκφραση

$$E_n(f) = -\frac{2}{(m+1)!}\left[\frac{1}{m+2} - \frac{1}{n+1}\sum_{i=0}^{n} x_i^{m+1}\right] f^{(m+1)}(\xi) , \tag{5.119}$$

όπου $\xi \in [-1,1]$ και

$$m = \begin{cases} n+1 , & \text{αν } n \text{ άρτιος} \\ n+2 , & \text{αν } n \text{ περιττός} , \end{cases}$$

οπότε η τάξη του τύπου (5.109) είναι $n+1$ αν ο n είναι άρτιος, ενώ είναι $n+2$ αν ο n είναι περιττός. (Το n παίρνει μόνον τις τιμές $0,1,2,\ldots,6$ και 8) .

Για να εφαρμόσουμε τον τύπο ολοκληρώσεως του Chebyshev (5.109) σε ένα ολοκλήρωμα της μορφής (5.86), όπου $[c,d] \neq [-1,1]$, λαμβάνουμε υπ' όψιν την (5.87). Έτσι έχουμε

$$\int_{c}^{d} f(t)dt = \frac{d-c}{2}\int_{1}^{1} F\left(\frac{d-c}{2}x + \frac{c+d}{2}\right)dx \approx$$

$$\approx \frac{d-c}{2}\,\frac{2}{n+1}\sum_{i=0}^{n} F\left(\frac{d-c}{2}x_i + \frac{c+d}{2}\right)dx ,$$

όπου x_i , $i = 0(1)n$ είναι οι ρίζες της (5.113). Δηλαδή, τελικά έχουμε

$$\int_{c}^{d} F(t)dt \approx \frac{d-c}{n+1}\sum_{i=0}^{n} F(t_i) , \tag{5.120}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

όπου $t_i = \frac{d-c}{2} x_i + \frac{c+d}{2}$, $i = 0(1)n$.

<u>Παράδειγμα 5.5</u> Να προσεγγιστεί το ολοκλήρωμα $\int_0^1 \sqrt{1-t^2}dt$ (ακριβής τιμή $= \frac{\pi}{4} \approx 0.785398$) με την βοήθεια του τύπου ολοκληρώσεως του Chebyshev για $n = 2$.

<u>Λύση</u> θέτουμε $t = \frac{1}{2}x + \frac{1}{2} = \frac{1}{2}(x+1)$. Για $n = 2$ λόγω της (5.120) έχουμε

$$\int_0^1 \sqrt{1-t^2}dt \approx \frac{1}{3}\left(\sqrt{1-t_0^2} + \sqrt{1-t_1^2} + \sqrt{1-t_2^2}\right) \text{ , όπου } t_i = \frac{1}{2}(x_i+1),$$

$i = 0,1,2$, $x_0 = -\frac{1}{\sqrt{2}}$, $x_1 = 0$, $x_2 = \frac{1}{\sqrt{2}}$. Άρα $t_0 = \frac{1}{2}\left(-\frac{1}{\sqrt{2}}+1\right)$, $t_1 = \frac{1}{2}$, $t_2 = \frac{1}{2}\left(\frac{1}{\sqrt{2}}+1\right)$ και επομένως

$$\int_0^1 \sqrt{1-t^2}dt \approx \frac{1}{3}\left[\left(1-\frac{1}{4}\left(-\frac{1}{\sqrt{2}}+1\right)^2\right)^{\frac{1}{2}} + \left(1-\frac{1}{4}\right)^{\frac{1}{2}} + \left(1-\frac{1}{4}\left(\frac{1}{\sqrt{2}}+1\right)^2\right)^{\frac{1}{2}}\right] \approx$$

≈ 0.792083 . ▲

5.8 Η μέθοδος των προσδιοριστέων συντελεστών

Η μέθοδος των προσδιοριστέων συντελεστών, που αναπτύχθηκε στην παράγραφο 3.3 για την εύρεση τύπων αριθμητικής παραγωγίσεως, μπορεί να χρησιμοποιηθεί ανάλογα και για την εύρεση τύπων αριθμητικής ολοκληρώσεως. Η μέθοδος αυτή έχει ήδη χρησιμοποιηθεί για την εύρεση των τύπων ολοκληρώσεως (5.109) του Chebyshev. Πρέπει να σημειωθεί ότι οποιοσδήποτε συγκεκριμένος τύπος απ' αυτούς που βρέθηκαν στις προηγούμενες παραγράφους, μπορεί να προκύψει, πολλές φορές ίσως ευκολώτερα, με την μέθοδο των προσδιοριστέων συντελεστών. Στη συνέχεια θα δοθεί ένα παράδειγμα, αφού η διαδικασία είναι γνωστή.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Παράδειγμα 5.6 Να προσδιορισθούν τα x_0, x_1, A_0, A_1 στον ακόλουθο προσεγγιστικό τύπο

$$\int_{-1}^{1}(1-x^2)f(x)dx \approx A_0 f(x_0) + A_1 f(x_1) \ , \tag{5.121}$$

έτσι ώστε να είναι όσον το δυνατόν περισσότερο ακριβής (ακριβής για πολυώνυμα όσον το δυνατόν μεγαλύτερου βαθμού).

Λύση Από τον (5.121) για $f(x) \equiv 1$, x, x^2, x^3 εξισώνοντας τα δύο μέλη του παίρνουμε το ακόλουθο σύστημα

$$\begin{cases} A_0 + A_1 = \frac{4}{3} \\ A_0 x_0 + A_1 x_1 = 0 \\ A_0 x_0^2 + A_1 x_1^2 = \frac{4}{15} \\ A_0 x_0^3 + A_1 x_1^3 = 0 \ . \end{cases} \tag{5.122}$$

Για την λύση του (5.122) θα ακολουθήσουμε μία διαδικασία που μπορεί να γενικευθεί. Έστω $P(x) = (x - x_0)(x - x_1) \equiv x^2 + \alpha_1 x + \alpha_2$ το πολυώνυμο με ρίζες τα x_0, x_1. Βρίσκοντας τα α_1, α_2 υπολογίζουμε στη συνέχεια τα x_0, x_1 και από τις δύο πρώτες εξισώσεις του (5.122) βρίσκονται τα A_0, A_1. Εάν πολλαπλασιάσουμε την τρίτη των (5.122) με το 1, την δεύτερη με α_1 και την πρώτη με α_2 και προσθέσουμε τις προκύπτουσες, έχοντας υπ' όψιν ότι $P(x_0) = P(x_1) = 0$, παίρνουμε

$$\frac{4}{3}\alpha_2 + \frac{4}{15} = 0 \quad \text{ή} \quad \alpha_2 = -\frac{1}{5} \ .$$

Εφαρμόζοντας την ίδια διαδικασία στις τετάρτη, τρίτη και δεύτερη των (5.122) λαμβάνουμε $\frac{4}{15}\alpha_1 = 0$, δηλαδή, $\alpha_1 = 0$, οπότε $P(x) = x^2 - \frac{1}{5}$ και επομένως

$$x_0 = -\frac{1}{\sqrt{5}} \ , \quad x_1 = \frac{1}{\sqrt{5}} \ , \quad A_0 = A_1 = \frac{2}{3} \ .$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Έτσι ο (5.121) γίνεται

$$\int_{1}^{1}(1-x^2)f(x)dx \approx \frac{2}{3}\left[f\left(-\frac{1}{\sqrt{5}}\right)+f\left(\frac{1}{\sqrt{5}}\right)\right] \quad . \qquad (5.123)$$

Είναι εύκολο να δειχθεί ότι για $f(x) = x^4$ ο (5.123) δεν είναι ακριβής. Άρα η τάξη του είναι 3 . Αξίζει να σημειωθεί ότι ο τύπος (5.121) είναι τύπος του Gauss με $w(x) = 1-x^2 \geq 0$ στο $[a,b] = [-1,1]$ και $n = 1$.

ΑΣΚΗΣΕΙΣ

1. Να υπολογισθεί το ολοκλήρωμα $\int_0^1 e^{-x^2}dx$, χρησιμοποιώντας το ανάπτυγμα της σειράς MacLaurin για την συνάρτηση e^{-x^2} .

2. Να προσεγγιστεί το $\int_0^1 (1+x^2)^{\frac{1}{2}}dx$, χρησιμοποιώντας τους 4 πρώτους όρους από το ανάπτυγμα του διωνύμου $(1+x^2)^{\frac{1}{2}}$.

3. Υποθέτοντας ότι η $f(x)$ αναπτύσσεται σε σειρά Taylor γύρω από το σημείο $x = c$ (όπου $c \in [a,b]$) , η οποία συγκλίνει ομοιόμορφα στο διάστημα $[a,b]$, να βρεθεί μία έκφραση σε σειρά για το ολοκλήρωμα $\int_a^b f(x)dx$. Τί συμβαίνει εάν $c = \frac{a+b}{2}$;

4. Να δειχθεί ότι i) $\int_{-1}^{1}\frac{e^x-1}{x}dx = \sum_{\substack{k=1\\ k=\text{περιττός}}}^{\infty}\frac{2}{k(k!)} \approx 2.11450175$

   ii) $\frac{1}{\sqrt{2\pi}}\int_{-1}^{1}e^{-\frac{x^2}{2}}dx = \frac{2}{\sqrt{2\pi}}\sum_{n=0}^{\infty}\frac{(-1)^n}{2^n n!(2n+1)}$.

5. Να προσδιορισθούν με διαφορετικό τρόπο οι συντελεστές A_i , $i = 0(1)n$ στον τύπο (5.5), λαμβάνοντας υπ' όψιν το γεγονός ότι ο

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

τύπος (5.5) είναι ακριβής για κάθε πολυώνυμο βαθμού $\leq n$. (Υπόδειξη: θέτοντας διαδοχικά $f(x) \equiv 1, x, x^2, \ldots, x^n$ στον (5.5) προκύπτουν οι ακόλουθες $n+1$ εξισώσεις με τους $n+1$ αγνώστους A_i, $i = 0(1)n$

$$\int_a^b x^\kappa dx = \frac{b^{\kappa+1} - a^{\kappa+1}}{\kappa + 1} = h \sum_{i=0}^{n} A_i x_i^\kappa \quad , \quad \kappa = 0(1)n \quad .$$

Να δειχθεί ότι το γραμμικό αυτό σύστημα έχει πάντοτε μία μοναδική λύση).

6. Να δειχθεί ότι ο κανόνας του Simpson για το ολοκλήρωμα $\int_a^b f(x)dx$ προκύπτει λαμβάνοντας το $\frac{1}{3}$ του κανόνα του τραπεζίου και τα $\frac{2}{3}$ του κανόνα του μέσου.

7. Να δειχθεί ότι ο γενικευμένος κανόνας του τραπεζίου είναι το ημιάθροισμα των γενικευμένων κανόνων του ορθογωνίου (5.45) και (5.47).

8. Να βρεθεί ο τέταρτος στη σειρά γενικευμένος κανόνας των Newton – Cotes (N = πολλ. του 4) και το αντίστοιχο σφάλμα αποκοπής.

9. Να δειχθεί ο τύπος (5.23) του σφάλματος αποκοπής του κανόνα του Simpson.

10. Χρησιμοποιώντας την (5.20) ή την (5.21), να βρεθεί ένα άνω φράγμα για το απόλυτο σφάλμα αποκοπής των κλειστών τύπων των Newton – Cotes, εάν $|f^{(n+1)}(x)| \leq M_{n+1}$ για κάθε $x \in [a,b]$.

11. Η γενική μορφή ενός γενικευμένου κλειστού τύπου των Newton – Cotes είναι:

$$\int_a^b f(x)dx \approx h \sum_{i=0}^{N} A_i f_i \quad , \qquad (1)$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

όπου $h = \frac{b-a}{N}$. Να δειχθεί ότι $\sum_{i=0}^{N} A_i = N$. Εάν, αντί των ακριβών τιμών f_i , $i = 0(1)N$, στο δεύτερο μέλος του (1) χρησιμοποιούνται οι προσεγγιστικές τιμές $\bar{f}_i$, $i = 0(1)N$ με αντίστοιχα σφάλματα $\varepsilon_i = \bar{f}_i - f_i$, τέτοια ώστε $|\varepsilon_i| \leq M$, $i = 0(1)N$, να βρεθεί ένα άνω φράγμα για το απόλυτο διαδοθέν σφάλμα κατά την εφαρμογή του τύπου (1). Να γίνει εφαρμογή στους γενικευμένους κανόνες του τραπεζίου,του Simpson και των $\frac{3}{8}$.

12. Εάν ο n είναι άρτιος $(n = 2\kappa , \kappa \geq 1)$ να δειχθεί ότι ο τύπος (5.5) είναι ακριβής για πολυώνυμα βαθμού $n+1$ (χωρίς την χρησιμοποίηση των (5.25), (5.26)).

13. Να υπολογισθεί το ολοκλήρωμα $\int_2^4 \sqrt{1+x}\, dx$, εφαρμόζοντας τον γενικευμένο κανόνα του τραπεζίου με $N = 4$. Να χρησιμοποιηθούν 4 δ.θ. στους υπολογισμούς. Στη συνέχεια να βρεθεί το μέγιστο απόλυτο σφάλμα αποκοπής καθώς και το μέγιστο απόλυτο διαδοθέν σφάλμα στρογγυλεύσεως.

14. Τα ίδια με την προηγούμενη άσκηση,χρησιμοποιώντας τον γενικευμένο κανόνα του Simpson (με $N = 4$) .

15. Να υπολογισθεί το $\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{dx}{1+\sin^2 x}$, χρησιμοποιώντας τους γενικευμένους κανόνες του τραπεζίου και του Simpson με $N = 6$.

16. Να υπολογισθεί το $\int_1^{2.2} f(x)dx$ με τον γενικευμένο κανόνα του Simpson από τα ακόλουθα δεδομένα:

x_i	1.0	1.2	1.4	1.6	1.8	2.0	2.2
f_i	0	0.03	0.09	0.19	0.33	0.51	0.71

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

17. Δοθέντος ότι $\int_0^1 4(1+x^2)^{-1}dx = \pi$, να βρεθεί μία προσέγγιση για τον π χρησιμοποιώντας τον γενικευμένο κανόνα του Simpson με $N = 4$.

18. Εάν $\int_0^{0.8} f(x)dx = 2$ και διαθέτουμε μόνον τα ακόλουθα δεδομένα

x_i	0	0.1	0.2	0.3	0.4	0.5	0.6	0.8
f_i	5	8	6	3	0	−3	−3	5

,

να εκτιμηθεί η $f(0.7)$ χρησιμοποιώντας κατάλληλα τον κανόνα του Simpson.

19. Να προσδιορισθεί ο μικρότερος N, έτσι ώστε το απόλυτο σφάλμα αποκοπής κατά τον υπολογισμό του $\int_1^{2.5} \ln x\, dx$, με τον γενικευμένο κανόνα του Simpson, να είναι $< 0.5 \cdot 10^{-5}$.

20. Να προσδιορισθεί ο μικρότερος N έτσι ώστε, εφαρμόζοντας τον γενικευμένο κανόνα του τραπεζίου στο ολοκλήρωμα $\int_{-1}^{3} \frac{dx}{x^2+1}$, το απόλυτο σφάλμα αποκοπής να είναι $< 0.5 \cdot 10^{-6}$.

21. Το πλάτος ενός ποταμού είναι 20 m . Να προσεγγιστεί το εμβαδόν S της καθέτου τομής αυτού αν δίνεται ο ακόλουθος πίνακας τιμών:

x	0	2	4	6	8	10	12	14	16	18	20
y	0.2	0.5	0.9	1.1	1.3	1.7	2.1	1.5	1.1	0.6	0.2

όπου x είναι η απόσταση (σε μέτρα) από την μία όχθη κατά μήκος της καθέτου τομής και y το αντίστοιχο βάθος (σε μέτρα).

22. Υποθέτουμε ότι δίνονται οι τιμές της $f(x)$ στα ισαπέχοντα σημεία

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$x_i = x_0 + ih$, $i = 0,1,2,\ldots$. Να δειχθεί δια συμβολικών μεθόδων ότι

$$\int_{x_0}^{x_1} f(x)dx = \int_{x_0}^{x_0+h} f(x)dx = h(1 + \frac{1}{2}\Delta - \frac{1}{12}\Delta^2 + \frac{1}{24}\Delta^3 - \frac{19}{720}\Delta^4 + \ldots)f_0 .$$

Στη συνέχεια, να δειχθεί πώς είναι δυνατόν να προκύψουν οι κανόνες του τραπεζίου, του Simpson κ.λ.π. .

23. Να βρεθεί η τάξη (ακρίβειας) των ακολούθων τύπων αριθμητικής ολοκληρώσεως:

i) $$\int_a^b f(x)dx \approx h(f_0 + f_1 + \ldots + f_N) - \frac{5h}{8}(f_0 + f_N) + \frac{h}{6}(f_1 + f_{N-1}) - \frac{h}{24}(f_2 + f_{N-2}) ,$$

ii) $$\int_a^b f(x)dx \approx h\left(\frac{f_0}{2} + f_1 + f_2 + \ldots + f_{N-1} + \frac{f_N}{2}\right) + \frac{h^2}{12}\left[f'(a) - f'(b)\right],$$

όπου $h = \frac{b-a}{N}$, $x_i = a + ih$, $i = 0(1)N$, $f_i = f(x_i)$ και $N \geq 2$ για τον τύπο (i) .

24. Να δειχθεί ότι ο τύπος αριθμητικής ολοκληρώσεως

$$\int_{-1}^{1} f(x)dx \approx \frac{1}{15}\left[7f(-1) + 16f(0) + 7f(1) + f'(-1) - f'(1)\right]$$

είναι ακριβής για πολυώνυμα βαθμού ≤ 5 .

25. Έχοντας υπ' όψιν την (5.54) να δειχθεί ότι, αν $I(h) \neq I(\lambda h)$, τότε $I^{(1)}(\lambda h) \notin [\min\{I(h), I(\lambda h)\} , \max\{I(h), I(\lambda h)\}]$.

26. Να δειχθεί ότι η $T_{2N}^{(1)}$ (βλέπε (5.56)) συμπίπτει με την προσέγγιση $S_{2N} \equiv S(\frac{h}{2})$ που προκύπτει από τον γενικευμένο κανόνα του Simpson.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

27. Για το ολοκλήρωμα $\int_0^1 \frac{dx}{1+x^4}$ να βρεθούν οι τιμές T_1, $T_2^{(1)}$, $T_4^{(2)}$ (βλέπε παράγραφο 5.5 και παράδειγμα 5.1) .

28. Να δειχθεί ότι η $T_{4N}^{(2)}$ (βλέπε (5.58)) συμπίπτει με την τιμή που προκύπτει από τον τέταρτο στη σειρά (επόμενο εκείνου των $\frac{3}{8}$) γενικευμένο κλειστό κανόνα των Newton – Cotes με βήμα $\frac{h}{4}$.

29. Από τα δεδομένα του πίνακα

x_i	-4	-3	-2	-1	0	1	2	3	4
f_i	0	0	1	0	0	1	0	0	0

,

να εκτιμηθεί το ολοκλήρωμα $\int_{-4}^{4} f(x)dx$ χρησιμοποιώντας ολοκλήρωση του Romberg.

30. Να δειχθεί ότι στους τύπους ολοκληρώσεως του Gauss ισχύουν

i) $\sum_{i=0}^{n} A_i = \sum_{i=0}^{n} |A_i| = \int_a^b w(x)dx$,

ii) $A_i = \int_a^b w(x)L_i(x)dx = \int_a^b w(x)L_i^2(x)dx$.

31. Αν $f(x)$, $g(x)$ είναι πολυώνυμα βαθμού $\leq n$ και x_i, $i = 0(1)n$ οι ρίζες του πολυωνύμου $\Phi_{n+1}(x)$ του Legendre, να δειχθεί ότι

$$\int_{-1}^{1} f(x)g(x)dx = \sum_{i=0}^{n} A_i f(x_i)g(x_i) ,$$

όπου τα A_i δίνονται από την (5.82) .

32. Αν τα πολυώνυμα $\varphi_0(x)$, $\varphi_1(x),\ldots,\varphi_n(x)$, $\varphi_{n+1}(x)$ είναι ορθογώνια σε σχέση με την συνάρτηση βάρους $w(x)$ στο διάστημα $[a,b]$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

και x_i , $i = 0(1)n$ είναι οι ρίζες του $\varphi_{n+1}(x)$, να δειχθεί ότι

$$\sum_{i=0}^{n} A_i \varphi_j(x_i)\varphi_\kappa(x_i) = 0 \ , \quad j,\kappa \leq n \ , \ j \neq \kappa \ ,$$

όπου τα A_i δίνονται από την (5.77) .

33. Να βρεθεί ο γενικευμένος κανόνας των Gauss – Legendre για $n = 1$.

34. Να βρεθεί ο τύπος ολοκληρώσεως των Gauss – Legendre για $n = 3$.

35. Να δειχθεί ο τύπος του σφάλματος αποκοπής (5.106).

36. Να βρεθούν τα σφάλματα αποκοπής για $n = 0,1,2$ από τον τύπο (5.119) .

37. Να υπολογισθεί το $\int_{-1}^{1} \frac{x^3 - 1}{\sqrt{1 - x^2}} dx$ χρησιμοποιώντας τον τύπο των Gauss – Chebyshev για $n = 1$.

38. Να προσεγγιστεί το $\int_{-1}^{1} \frac{x \sin x}{\sqrt{1 - x^2}} dx$ χρησιμοποιώντας τους τύπους Gauss – Legendre για $n = 1$ και Gauss – Chebyshev για $n = 1$.

39. Να προσεγγιστεί το $\int_{0}^{1} \sqrt{1+2t}\, dt$ με τους τύπους: του Simpson, των Gauss – Legendre για $n = 2$ και του Chebyshev για $n = 2$, χρησιμοποιώντας 6 σ.ψ. στους υπολογισμούς. Ποιός είναι περισσότερο ακριβής;

40. Να υπολογισθεί το $\int_{-1}^{1} \sqrt{1 - x^2}\, dx$ με τους τύπους (5.94) και (5.123).

41. Να προσεγγιστεί το $\int_{-1}^{1} \frac{dx}{\sqrt{x^4 - 17x^2 + 16}}$ με τους τύπους των Gauss – Chebyshev για $n = 1,2$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

42. Να προσδιορισθούν οι συντελεστές A,B,C,D έτσι ώστε οι ακόλουθοι τύποι αριθμητικής ολοκληρώσεως να είναι όσον το δυνατόν περισσότερο ακριβείς.

i) $\int_a^b f(x)dx \approx (b-a)\left[Af(a)+Bf(b)\right] + (b-a)^2\left[Cf'(a)+Df'(b)\right]$

ii) $\int_{-\pi}^{\pi} f(x)\sin mx\, dx \approx Af(-\pi) + Bf(0) + Cf(\pi)$

iii) $\int_{-\pi}^{\pi} f(x)\cos mx\, dx \approx Af(-\pi) + Bf(0) + Cf(\pi)$,

όπου m θετικός ακέραιος.

43. Όπως και στην προηγούμενη άσκηση, να χρησιμοποιηθεί η μέθοδος των προσδιοριστέων συντελεστών στους ακόλουθους τύπους για την εύρεση των αντίστοιχων αγνώστων.

i) $\int_a^b f(x)dx \approx Af(x_1)+Bf'(x_2)$ (άγνωστοι: A,B,x_1,x_2)

ii) $\int_0^{\infty} e^{-x}f(x)dx \approx Af(0)+Bf(1)+Cf(2)$ (άγνωστοι: A,B,C)

iii) $\int_0^h f(x)dx \approx h\left[Af(0)+Bf(h)\right]+h^2Cf'(Dh)$ (άγνωστοι: A,B,C,D)

iv) $\int_0^h f(x)dx \approx h\left[Af(Bh)+hCf'(Dh)\right]$ (άγνωστοι: A,B,C,D)

v) $\int_{-1}^{1}(1-x^2)f(x)dx \approx A\left[f(x_0)+f(x_1)+f(x_2)\right]$ (άγνωστοι: A,x_0,x_1,x_2)

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

vi) $\int_{-1}^{1}(1-x^2)^{-\frac{1}{2}} f(x)dx \approx Af(-1)+Bf(0)+Cf(1)$ (άγνωστοι: A,B,C)

vii) $\int_{-1}^{1} f(x)dx \approx Af(-2)+Bf(-1)+Cf(0)+Df(1)+Ef(2)$ (άγνωστοι: A,B,C,D,E)

viii) $\int_{a}^{b} f(x)dx \approx Af(x_1)+Bf''(x_2)$ (άγνωστοι: A,B,x_1,x_2)

44. Να βρεθεί η τάξη του τύπου:

$$\int_0^1 xf(x)dx \approx \frac{9-\sqrt{6}}{36} f\left(\frac{6-\sqrt{6}}{10}\right) + \frac{9+\sqrt{6}}{36} f\left(\frac{6+\sqrt{6}}{10}\right).$$

45. Με την βοήθεια υπολογιστού και χρησιμοποιώντας τους γενικευμένους κανόνες του τραπεζίου, του Simpson και του μέσου, να υπολογισθούν τα παρακάτω ολοκληρώματα με ακρίβεια 6 δ.θ.

i) $\int_0^1 \ln(x+1)\ln(x+2)dx$ ii) $\int_0^1 \sqrt{1+10x^3}\,dx$

iii) $\int_0^{\pi} \frac{1-\cos x}{x}dx$ iv) $\int_0^1 e^{-x^2}dx$ v) $\int_0^1 \frac{dx}{1+10x^2}$

vi) $\int_0^2 \frac{e^{-x}}{1+x^2}dx$ vii) $\int_0^1 \sqrt{x}\sin x\,dx$ viii) $\int_0^1 \frac{dx}{\sqrt{x}(1+x)}$.

46. Να υπολογισθεί ο ln2 με ακρίβεια 6 δ.θ. με την βοήθεια του γενικευμένου κανόνα του Simpson.

47. Να δειχθεί ο τύπος (5.38) και ο (5.43).

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

48. Να βρεθεί ο τύπος ολοκληρώσεως (5.109) του Chebyshev για $n = 3$.

49. Να προσεγγισθεί το $\int_{-1}^{1} f(x)dx$,όπου $f(x) = -25x^4 + 45x^2 - 8$, εφαρμόζοντας τον γενικευμένο κανόνα του τραπεζίου με $h = 1$ και τον κανόνα του Simpson. Ποιά προσέγγιση είναι περισσότερο ακριβής; Να βρεθούν τα αντίστοιχα σφάλματα αποκοπής και να προσδιορισθεί ο αριθμός ξ στον τύπο του σφάλματος αποκοπής του γενικευμένου κανόνα του τραπεζίου (βλέπε (5.28)).

50. Δίνεται ο ακόλουθος πίνακας τιμών της συναρτήσεως $f(x)$:

x	-1	0	1	3
$f(x)$	-1	1	3	79

Να βρεθεί η ακριβής τιμή του ολοκληρώματος $\int_{-1}^{3} P(x)dx$, όπου $P(x)$ είναι το πολυώνυμο παρεμβολής που προσεγγίζει την $f(x)$, με σημεία παρεμβολής τις τιμές του x του πίνακα, χωρίς να βρεθεί το $P(x)$.

51. Να προσεγγισθεί το ολοκλήρωμα $\int_{0}^{1} 4\sqrt{x(1-x)}\,dx$ i) με τον κανόνα του Simpson ii) με τον τύπο του Chebyshev για $n = 2$ iii) με τον τύπο Gauss – Legendre για $n = 1$ και iv) με τον τύπο Gauss – – Chebyshev για $n = 1$. Ποιά είναι η ακριβής τιμή του ολοκληρώματος;

52. Δίνεται το ολοκλήρωμα $\int_{-1}^{1} (1 + x + x^2 + x^3 + x^4)dx$. Να βρεθεί πρώτα η ακριβής τιμή του αναλυτικά. Στη συνέχεια, να προσεγγισθεί η τιμή του χρησιμοποιώντας i) τον κανόνα του Simpson ii) τον κανόνα των 3/8 iii) τον τύπο του Chebyshev για $n = 2$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

και iv) τον τύπο των Gauss — Legendre για $n = 1$. Ποιός από τους παραπάνω 4 τύπους αριθμητικής ολοκληρώσεως είναι περισσότερο ακριβής για το δοσμένο ολοκλήρωμα;

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

6

ΑΡΙΘΜΗΤΙΚΗ ΕΠΙΛΥΣΗ ΕΞΙΣΩΣΕΩΝ

6.1 Εισαγωγή

Εάν μία εξίσωση είναι αρκετά πολύπλοκης μορφής, συνήθως,δεν είναι δυνατόν να βρούμε τις ακριβείς τιμές των ριζών της. Είναι γνωστό επίσης,ότι δεν υπάρχει αναλυτική έκφραση για τις ρίζες ενός γενικού πολυωνύμου βαθμού μεγαλύτερου από τέσσερα. Για τις ρίζες πολυωνύμων τρίτου και τέταρτου βαθμού υπάρχουν διαθέσιμοι τύποι σε κλειστή μορφή, αλλά αυτοί είναι σχετικά πολύπλοκοι και σπανίως χρησιμοποιούνται στην πράξη. Επομένως, εκτός από ειδικές περιπτώσεις, προτιμούμε ή μάλλον υποχρεωνόμαστε στην πράξη να χρησιμοποιήσουμε κάποια αριθμητική μέθοδο για να προσεγγίσουμε τις ρίζες μιας εξισώσεως.

Ας υποθέσουμε ότι έχουμε την εξίσωση

$$f(x) = 0 \ , \tag{6.1}$$

όπου $f(x)$ είναι μία γνωστή συνάρτηση και θέλουμε να υπολογίσουμε μερικές ή όλες τις ρίζες αυτής (καλείται ρίζα ή λύση της (6.1) ή μηδενική τιμή για την $f(x)$ κάθε τιμή ξ τέτοια ώστε $f(\xi) = 0$) . Η εξίσωση (6.1) καλείται αλγεβρική ή πολυωνυμική εάν η $f(x)$ είναι ένα πολυώνυμο με συντελεστές πραγματικούς ή μιγαδικούς αριθμούς, καλείται

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

δε μη αλγεβρική ή υπερβατική εάν η $f(x)$ περιέχει π.χ. τριγωνομετρικές ή εκθετικές συναρτήσεις (π.χ. $f(x) = x + e^x = 0$). Στη συνέχεια του κεφαλαίου αυτού, εκτός κι αν διαφορετικά καθορίζεται, υποθέτουμε ότι η $f(x)$ είναι μία πραγματική συνάρτηση, ορισμένη και συνεχής σε κάποιο πεπερασμένο ή απέραντο διάστημα $[a,b]$ και ζητούμε να υπολογίσουμε τις πραγματικές ρίζες της.

Πριν χρησιμοποιήσουμε μία αριθμητική μέθοδο, είναι χρήσιμο να έχουμε κάποια πληροφορία όσον αφορά τον αριθμό, την φύση και την προσεγγιστική θέση των ριζών. Η συνήθης τακτική περιλαμβάνει την κατασκευή της γραφικής παράστασης ή πίνακα τιμών της συναρτήσεως $f(x)$, βάσει των οποίων λαμβάνονται οι απαιτούμενες πληροφορίες. Βασικός σκοπός του αρχικού αυτού σταδίου είναι ο εντοπισμός των πραγματικών ριζών της (6.1), δηλαδή, η εύρεση των μικρότερων δυνατών διαστημάτων καθένα από τα οποία περιέχει τουλάχιστον μία ρίζα της (6.1). Σε μία ιδανική περίπτωση κάθε διάστημα θα περιέχει ακριβώς μία ρίζα.

Το δεύτερο στάδιο επιλύσεως της (6.1) περιλαμβάνει την χρησιμοποίηση μιας αριθμητικής μεθόδου, η οποία είναι συνήθως μία επαναληπτική διαδικασία, με την οποία βελτιώνουμε τις ήδη γνωστές (από τον εντοπισμό) προσεγγίσεις των ριζών μέχρι να επιτευχθεί η επιζητούμενη ακρίβεια.

6.2 Εντοπισμός των ριζών

Το ακόλουθο γνωστό θεώρημα από την Μαθηματική Ανάλυση είναι πολύ χρήσιμο στην διαδικασία εντοπισμού των ριζών της $f(x) = 0$.

<u>Θεώρημα 6.1</u> Εάν η συνεχής συνάρτηση $f(x)$ παίρνει τιμές αντιθέτου προσήμου στα άκρα ενός διαστήματος $[a,b]$, δηλαδή, $f(a)f(b) < 0$, τότε το διάστημα $[a,b]$ περιέχει τουλάχιστον μία ρίζα της εξισώσεως $f(x) = 0$, δηλαδή, υπάρχει τουλάχιστον ένας αριθμός $\xi \in (a,b)$ τέτοιος ώστε $f(\xi) = 0$. Επιπλέον, το διάστημα $[a,b]$ θα περιέχει μία μοναδική ρίζα ξ εάν η $f'(x)$ υπάρχει και ισχύει $f'(x) > 0$ (ή

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$f'(x) < 0$) για κάθε $x \in (a,b)$. ▲

Εάν η $f'(x)$ είναι συνεχής και οι ρίζες της $f'(x) = 0$ βρίσκονται εύκολα, είναι αρκετό, για τον εντοπισμό των ριζών της (6.1), να βρούμε το πρόσημο της $f(x)$ στα σημεία που είναι ρίζες της $f'(x) = 0$ καθώς και στα άκρα του διαστήματος που ορίζεται η $f(x)$.

Παράδειγμα 6.1 Να εντοπισθούν οι ρίζες της $f(x) = 2x^3 - 6x + 2 = 0$.

Λύση Έχουμε $f'(x) = 6(x^2 - 1)$. Οι ρίζες της $f'(x) = 0$ είναι $x = \pm 1$. Επίσης είναι $f(-\infty) < 0$, $f(-1) > 0$, $f(1) < 0$, $f(+\infty) > 0$. Επομένως η $f(x) = 0$ έχει τρείς πραγματικές ρίζες ξ_1, ξ_2, ξ_3 με $\xi_1 \in (-\infty,1)$, $\xi_2 \in (-1,1)$, $\xi_3 \in (1,+\infty)$. Πιο συγκεκριμένα $\xi_1 \in (-2,-1)$, $\xi_2 \in (0,1)$, $\xi_3 \in (1,2)$ αφού $f(-2) < 0$, $f(2) > 0$, $f(0) > 0$. ▲

Ένας άλλος τρόπος εντοπισμού και προσδιορισμού (αρχικών) προσεγγιστικών τιμών των ριζών της $f(x) = 0$, είναι η χρησιμοποίηση γραφικών μεθόδων. Τα σημεία τομής της καμπύλης, που παριστά γραφικά την συνάρτηση $y = f(x)$, με τον άξονα των x αποτελούν τις ζητούμενες προσεγγίσεις των ριζών. Οι προσεγγιστικές αυτές τιμές των ριζών χρησιμοποιούνται συνήθως σαν αρχικές τιμές από τις οποίες μπορούμε να βρούμε πιο ακριβείς προσεγγίσεις των ριζών με την βοήθεια επαναληπτικών μεθόδων,όπως θα δούμε σε επόμενες παραγράφους.

Συχνά, είναι προτιμώτερο να αντικαταστήσουμε την εξίσωση $f(x) = 0$ με μια ισοδύναμη εξίσωση $g(x) = h(x)$, όπου οι συναρτήσεις $g(x)$ και $h(x)$ είναι απλούστερες από την $f(x)$.Κατόπιν, γίνεται η γραφική παράσταση των $y = g(x)$ και $y = h(x)$ και οι τετμημένες των σημείων τομής των δύο καμπύλων αποτελούν τις προσεγγίσεις των ριζών της $f(x) = 0$.

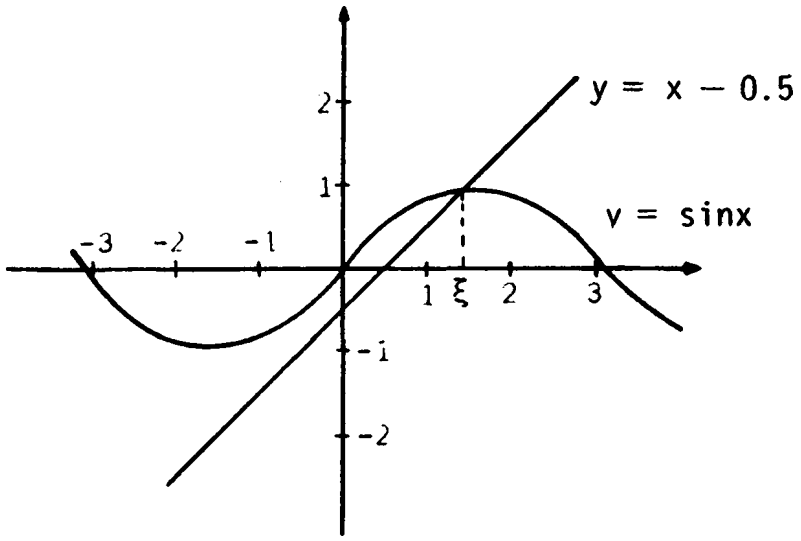
Παράδειγμα 6.2 Να λυθεί γραφικά η εξίσωση

$$f(x) = \sin x - x + \frac{1}{2} = 0 . \qquad (6.2)$$

Λύση Η (6.2) γράφεται $\sin x = x - \frac{1}{2}$. Κατασκευάζουμε την καμπύλη $y = \sin x$ και την ευθεία $y = x - 0.5$ (βλέπε σχ. 6.1). Από το σχή-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

μα 6.1 προκύπτει ότι η εξίσωση (6.2) έχει μία μόνον πραγματική ρίζα $\xi \approx 1.5$.



σχ. 6.1

6.3 Η μέθοδος της διχοτομήσεως

Υποθέτουμε ότι έχουμε την εξίσωση

$$f(x) = 0 \ , \tag{6.3}$$

όπου η συνάρτηση $f(x)$ είναι συνεχής στο $[a,b]$ και $f(a)f(b) < 0$. Σύμφωνα με το θεώρημα 6.1 στο $[a,b]$ υπάρχει τουλάχιστον μία ρίζα της (6.3). Υπολογίζουμε την τιμή $f(\frac{a+b}{2})$, δηλαδή, την τιμή της $f(x)$ στο μέσον του διαστήματος $[a,b]$. Εάν $f(\frac{a+b}{2}) = 0$, τότε η $\xi = \frac{a+b}{2}$ είναι μία ρίζα της (6.3). Εάν $f(\frac{a+b}{2}) \neq 0$, τότε εκλέγουμε εκείνο το μισό διάστημα $\left[a, \frac{a+b}{2}\right]$ ή $\left[\frac{a+b}{2}, b\right]$ στα άκρα του οποίου η $f(x)$ έχει αντίθετα πρόσημα. Στο νέο αυτό διάστημα $[a_1, b_1]$, η διαδικασία της διχοτομήσεως (που εφαρμόστηκε στο διάστημα $[a,b]$) επαναλαμβάνεται κ.ο.κ., οπότε τελικά βρίσκουμε είτε την ακριβή τιμή μιας ρίζας της (6.3) ή μία άπειρο ακολουθία διαστημάτων $[a_1,b_1]$, $[a_2,b_2]$,...,

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$[a_n,b_n]$,..., τέτοιων ώστε

$$f(a_n)f(b_n) < 0 \ , \quad n = 1,2,... \tag{6.4}$$

και

$$b_n - a_n = \frac{1}{2^n}(b-a) \ . \tag{6.5}$$

Αφού τα αριστερά άκρα $a_1,a_2,...,a_n,...$ των διαστημάτων σχηματίζουν μία αύξουσα και φραγμένη ακολουθία και τα δεξιά άκρα $b_1,b_2,...,b_n,...$ σχηματίζουν μία φθίνουσα και φραγμένη ακολουθία, τότε λόγω της (6.5) θα υπάρχει ένα κοινό όριο

$$\xi = \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n \ .$$

Από την (6.4), παίρνοντας όρια και λόγω της συνεχείας της $f(x)$, προκύπτει ότι $\left(f(\xi)\right)^2 \leqq 0$, δηλαδή $f(\xi) = 0$ και επομένως ξ είναι μία ρίζα της (6.3) .

Προφανώς, επειδή $\xi \in [a_n,b_n]$, $n = 1,2,...,$ από την (6.5) έχουμε

$$0 \leqq \xi - a_n \leqq \frac{1}{2^n}(b-a) \ , \tag{6.6}$$

οπότε, απαιτώντας να ισχύει $\frac{1}{2^n}(b-a) \leqq \varepsilon = \frac{1}{2}10^{-\kappa}$, βρίσκουμε τον ελάχιστο αριθμό επαναλήψεων n που απαιτείται ώστε η ρίζα ξ να υπολογίζεται με ακρίβεια κ δ.θ. (βλέπε άσκηση 5) .

Έτσι εφαρμόζοντας την μέθοδο της διχοτομήσεως, όπως αυτή περιγράφηκε προηγουμένως, βρίσκουμε πάντοτε μία ρίζα της (6.3) από αυτές που υπάρχουν μέσα στο $[a,b]$, δηλαδή, η μέθοδος της διχοτομήσεως παρέχει εγγυημένη σύγκλιση σε μια ρίζα $\xi \in [a,b]$ και είναι κατάλληλη για να χρησιμοποιηθεί από ένα ηλεκτρονικό υπολογιστή. Η ακρίβεια με την οποία η ρίζα ξ μπορεί να βρεθεί περιορίζεται μόνον από την ακρίβεια με την οποία η $f(x)$ μπορεί να υπολογισθεί. Στα μειονεκτήματα της μεθόδου της διχοτομήσεως συγκαταλέγεται το γεγονός ότι δεν μπορεί να ε-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

φαρμοσθεί αυτή για την εύρεση ριζών πέριξ των οποίων η $f(x)$ δεν αλλάζει πρόσημο (ρίζες αρτίας πολλαπλότητος), καθώς και η άργή σύγκλισή της. Γι' αυτό και δεν συνιστάται η χρησιμοποίησή της σε μεγάλο βαθμό για πρακτικές εφαρμογές.

Παράδειγμα 6.3 Με την μέθοδο της διχοτομήσεως να βρεθεί μία ρίζα της $f(x) = 2x^3 - x - 0.5 = 0$ στο διάστημα $[0,1]$ με ακρίβεια 2 δ.θ. .

Λύση Είναι $[a,b] = [0,1]$ και $f(0) = -0.5 < 0$, $f(1) = 0.5 > 0$. Έχοντας υπ' όψιν την μέθοδο της διχοτομήσεως κατασκευάζουμε τον ακόλουθο πίνακα, όπου τέθηκε $a_0 = a$, $b_0 = b$, διατηρώντας 3 δ.θ. στους υπολογισμούς. Ο ελάχιστος αριθμός επαναλήψεων που απαιτείται για να υπολογίσουμε την ρίζα με ακρίβεια 2 δ.θ. είναι $n = 8$.

n	a_n	b_n	$x_n = \frac{a_n + b_n}{2}$	$f(x_n)$
0	0	1	0.5	− 0.75
1	0.5	1	0.75	− 0.406
2	0.75	1	0.875	− 0.035
3	0.875	1	0.938	0.212
4	0.875	0.938	0.906	0.082
5	0.875	0.906	0.890	0.02
6	0.875	0.890	0.882	− 0.01
7	0.882	0.890	0.886	0.006
8	0.882	0.886	0.884	− 0.002

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Επειδή $a_8 = 0.882 \approx 0.88$ και $b_8 = 0.886 \approx 0.89$, για την ζητούμενη ρίζα ξ παίρνουμε $\xi \approx \frac{a_8 + b_8}{2} = 0.884$ ή $\xi \approx 0.88$.

6.4 Η Επαναληπτική μέθοδος

Μία από τις πιό σημαντικές μεθόδους για την αριθμητική επίλυση εξισώσεων είναι η επαναληπτική μέθοδος ή μέθοδος των διαδοχικών προσεγγίσεων ή μέθοδος της διαδοχικής αντικαταστάσεως. Γενικά, με την μέθοδο αυτή αρχίζουμε με μία αρχική (πρόχειρη) προσέγγιση x_0 της ζητούμενης ρίζας και από αυτή βρίσκουμε μία νέα προσέγγιση x_1 , από την οποία πάλι βρίκουμε μία νέα προσέγγιση x_2 κ.ο.κ., μέχρις ότου (με την προϋπόθεση ότι η ακολουθία των διαδοχικών προσεγγίσεων $x_0, x_1, x_2, \ldots,$ που καλούνται και επαναλήψεις, συγκλίνει) βρούμε μία προσέγγιση x_n για την ζητούμενη ρίζα με την επιζητούμενη (προκαθορισμένη) ακρίβεια.

Η επαναληπτική μέθοδος προκύπτει ως εξής: Υποθέτουμε πάλι ότι έχουμε την εξίσωση (6.1). Εάν αντικαταστήσουμε την (6.1) με μία ισοδύναμη εξίσωση (δύο εξισώσεις λέγονται ισοδύναμες εάν έχουν ακριβώς τις ίδιες ρίζες)

$$x = g(x) \ , \tag{6.7}$$

όπου $g(x)$ είναι συνεχής συνάρτηση, τότε οι διαδοχικές προσεγγίσεις x_n, $n = 1,2,\ldots$ για την ζητούμενη ρίζα υπολογίζονται με την βοήθεια της συναρτήσεως $g(x)$ που καλείται επαναληπτική συνάρτηση από την σχέση

$$x_{n+1} = g(x_n) \ , \quad n = 0,1,2,\ldots \ . \tag{6.8}$$

Εάν η ακολουθία των διαδοχικών προσεγγίσεων συγκλίνει, δηλαδή, $\lim_{n \to \infty} x_n = \xi$, τότε παίρνοντας όρια στην (6.8) έχουμε

$$\xi = g(\xi) \ , \tag{6.9}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

δηλαδή, το όριο ξ είναι ένα σταθερό σημείο της $g(x)$ και επομένως μία ρίζα της $f(x) = 0$. Βεβαίως, στην πράξη, μόνον ένας πεπερασμένος αριθμός επαναλήψεων $x_1, x_2, \ldots, x_n$ υπολογίζεται για την ρίζα ξ. Ένα σημαντικό πρόβλημα που προκύπτει είναι ο προσδιορισμός του μικρότερου αριθμού επαναλήψεων, που απαιτούνται με σκοπό να βρούμε μία καλή προσέγγιση x_n της ρίζας ξ της $f(x) = 0$. Ο ισχυρισμός ότι η x_n είναι καλή προσέγγιση της ρίζας ξ έχει δύο πιθανές ερμηνείες. Η πρώτη είναι ότι ο αριθμός $|x_n - \xi|$ είναι αρκετά μικρός και η δεύτερη ότι ο αριθμός $|f(x_n)|$ είναι αρκετά μικρός. Αν και σε πολλές περιπτώσεις είναι αληθές ότι αν ισχύει η πρώτη συνθήκη, τότε θα ισχύει και η δεύτερη, εν τούτοις, αυτές οι δύο συνθήκες είναι συχνά εντελώς ανεξάρτητες μεταξύ τους και η εκλογή τους πρέπει να γίνεται κατάλληλα.

Εάν δεχθούμε σαν πιο κατάλληλη την δεύτερη συνθήκη, τότε θα σταματήσουμε να υπολογίζουμε νέες επαναλήψεις εφ' όσον ικανοποιηθεί η σχέση

$$|f(x_n)| \leqq \varepsilon_1 , \tag{6.10}$$

όπου $\varepsilon_1 > 0$ είναι ένας αρκετά μικρός προκαθορισμένος αριθμός.

Εάν ενδείκνυται η πρώτη συνθήκη, τότε, επειδή η ρίζα ξ είναι άγνωστη και δεν μπορούμε να εξετάζουμε την ποσότητα $|x_n - \xi|$, προσπαθούμε με αναλυτικές μεθόδους να σχετίσουμε την ποσότητα $|x_n - \xi|$ με την $|x_n - x_{n-1}|$, η οποία εξετάζεται εύκολα. Συνήθως, στην περίπτωση αυτή, σταματάμε να υπολογίζουμε νέες επαναλήψεις όταν ικανοποιηθεί η συνθήκη

$$|x_n - x_{n-1}| \leqq \varepsilon_2 , \tag{6.11}$$

όπου $\varepsilon_2 > 0$ είναι ένας αρκετά μικρός προκαθορισμένος αριθμός, αν και η ικανοποίηση της συνθήκης (6.11), γενικά, δεν σημαίνει ότι και $|x_n - \xi| \leqq \varepsilon_2$. Αν ληφθεί $\varepsilon_2 = \frac{1}{2} 10^{-\kappa}$ (κ ακέραιος $\geqq 0$), τότε η επαναληπτική διαδικασία σταματά όταν δύο διαδοχικές επαναλήψεις συμφωνούν σε κ δ.θ. και λέμε ότι η τελευταία τιμή x_n που βρέθηκε

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

προσεγγίζει την ρίζα ξ με ακρίβεια κ δ.θ. .

Οι σχέσεις (6.10), (6.11) καλούνται συνθήκες ή κριτήρια τερματισμού της επαναληπτικής μεθόδου. Πολλές φορές επίσης, ένα άλλο κριτήριο τερματισμού των επαναλήψεων χρησιμοποιείται σε συνδυασμό με τα προηγουμένως αναφερθέντα. Συγκεκριμένα, επειδή μπορεί η ακολουθία των επαναλήψεων να μην συγκλίνει ή να συγκλίνει πολύ αργά, για να εμποδίσουμε τον υπολογισμό απεριορίστου αριθμού επαναλήψεων, είναι απαραίτητο να μετράμε τον αριθμό των επαναλήψεων που υπολογίζουμε και να σταματάμε όταν ο αριθμός αυτός υπερβεί ένα προκαθορισμένο μέγιστο αριθμό επαναλήψεων.

Από την περιγραφή της επαναληπτικής μεθόδου, που έγινε προηγουμένως, προκύπτουν τα ακόλουθα ερωτήματα: i) Πώς βρίσκεται μία κατάλληλη συνάρτηση $g(x)$; ii) Κάτω από ποιές προυποθέσεις συγκλίνει η ακολουθία του αλγορίθμου (6.8); και iii) Πόσο γρήγορα θα συγκλίνει η ακολουθία αυτή; Απαντήσεις στα ερωτήματα αυτά δίνονται στη συνέχεια.

6.4.1 Σύγκλιση της επαναληπτικής μεθόδου

Η επαναληπτική μέθοδος ορίζεται πλήρως από την επαναληπτική συνάρτηση $g(x)$ και την αρχική προσέγγιση x_0 . Η σύγκλιση ή μή της ακολουθίας του αλγορίθμου (6.8) σε κάποια ρίζα ξ (άρα και η ταχύτητα συγκλίσεως) εξαρτάται προφανώς από την συνάρτηση $g(x)$ και από την αρχική προσέγγιση x_0 .

Έστω ξ μία ρίζα της (6.1). Τότε, λόγω της (6.7) έχουμε $\xi = g(\xi)$. Η επαναληπτική μέθοδος που ορίζεται από την (6.8) λέγεται ότι είναι τοπικώς συγκλίνουσα στην ρίζα ξ εάν υπάρχει μία περιοχή $\Pi(\xi)$ της ρίζας ξ , τέτοια ώστε, για όλες τις αρχικές προσεγγίσεις $x_0 \in \Pi(\xi)$, η ακολουθία που παράγεται από την (6.8) να συγκλίνει στη ρίζα ξ . Εάν σαν $\Pi(\xi)$ μπορεί να ληφθεί το σύνολο $\mathbb{R}$ των πραγματικών αριθμών, τότε η μέθοδος καλείται ολικώς συγκλίνουσα.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

<u>Ορισμός 6.1</u> Έστω $\{x_n\}$, $n = 0,1,2,\ldots$ μία ακολουθία τιμών που παράγεται από την επαναληπτική μέθοδο

$$x_{n+1} = g(x_n) \ , \ n = 0,1,2,\ldots \tag{6.12}$$

και $\lim\limits_{n\to\infty} x_n = \xi$. Εάν υπάρχει ένας αριθμός $p \geqq 1$ και μία σταθερά $C > 0$ (εάν $p = 1$ τότε $0 < C < 1$) , τέτοια ώστε

$$\lim_{n\to\infty} \frac{|\varepsilon_{n+1}|}{|\varepsilon_n|^p} = C \ , \tag{6.13}$$

όπου $\varepsilon_n = x_n - \xi$, τότε ο p καλείται η τάξη συγκλίσεως της ακολουθίας (ή λέμε ότι η ταχύτητα συγκλίσεως ή απλώς η σύγκλιση της ακολουθίας είναι τάξεως p) και η C σταθερά του ασυμπτωτικού σφάλματός της (η C καλείται και σταθερά συγκλίσεως). Ειδικά,για $p = 1,2$ και 3 η σύγκλιση λέγεται ότι είναι γραμμική, τετραγωνική και κυβική αντίστοιχα. Η αντίστοιχη μέθοδος (6.12) λέγεται ότι είναι τάξεως p για τον προσδιορισμό της ρίζας ξ . ▲

Από την (6.13) προκύπτει ότι για πολύ μεγάλες τιμές του n έχουμε $|\varepsilon_{n+1}| \approx C|\varepsilon_n|^p$. Αυτό σημαίνει ότι όσο πιό μεγάλος είναι ο αριθμός p και όσο πιό μικρή η σταθερά C , τόσο μεγαλύτερη είναι η βελτίωση που επιτυγχάνεται στην ακρίβεια από την μία επαναληψη στην άλλη και επομένως απαιτούνται λιγώτερες επαναλήψεις για την εύρεση της τελικής προσεγγίσεως με την επιζητούμενη ακρίβεια.

Τα ακόλουθα θεωρήματα δίνουν ικανές συνθήκες για την σύγκλιση της ακολουθίας που παράγεται από την επαναληπτική μέθοδο (6.8) .

<u>Θεώρημα 6.2</u> Εάν ξ είναι ένα σταθερό σημείο της συναρτήσεως $g(x)$ της επαναληπτικής μεθόδου (6.8) ($\xi = g(\xi)$) και για κάθε $x,y \in I \equiv [\xi - \varepsilon, \xi + \varepsilon]$, $\varepsilon > 0$ έχουμε

$$|g(x) - g(y)| \leqq L\ |x - y| \ , \quad 0 \leqq L < 1 \ , \tag{6.14}$$

τότε για κάθε $x_0 \in I$ η ακολουθία που παράγεται από τον αλγόριθμο $x_{n+1} = g(x_n)$, $n = 0,1,2,\ldots$ συγκλίνει στο ξ που είναι και η μο-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ναδική ρίζα της $x = g(x)$ στο I . (Η συνθήκη (6.14) καλείται συνθήκη του Lipschitz και η ισχύς της συνεπάγεται την συνέχεια της $g(x)$ στο I).

<u>Απόδειξη</u> Θα δειχθεί ότι

$$x_n \in I \ , \quad n = 0,1,2,\ldots \ . \tag{6.15}$$

Για $n = 0$ είναι $x_0 \in I$ εξ υποθέσεως. Εάν υποθέσουμε ότι η (6.15) ισχύει για $n = 0(1)\kappa$, $\kappa \geqq 1$, τότε προκύπτει αμέσως, λόγω της (6.14), ότι

$$|x_{\kappa+1} - \xi| = |g(x_\kappa) - g(\xi)| \leqq L|x_\kappa - \xi| \leqq L^2|x_{\kappa-1} - \xi| \leqq \cdots \leqq$$

$$\leqq L^{\kappa+1}|x_0 - \xi| < |x_0 - \xi| \ . \tag{6.16}$$

Άρα $x_{\kappa+1} \in I$ και επομένως απεδείχθη η (6.15) .
Επίσης, από την (6.16) παίρνοντας όρια έχουμε

$$\lim_{\kappa \to \infty} |x_{\kappa+1} - \xi| \leqq \lim_{\kappa \to \infty} L^{\kappa+1}|x_0 - \xi| = 0 \text{ ,δηλαδή, } \lim_{n \to \infty} x_n = \xi \ .$$

Θα δειχθεί τώρα ότι στο I δεν μπορεί να υπάρχει και άλλο σταθερό σημείο $\hat{\xi}$ της $g(x)$. Πράγματι, αν $\hat{\xi} \in I$ με $\hat{\xi} \neq \xi$ και $\hat{\xi} = g(\hat{\xi})$ τότε θα έχουμε

$$|\xi - \hat{\xi}| = |g(\xi) - g(\hat{\xi})| \leqq L|\xi - \hat{\xi}| < |\xi - \hat{\xi}| \ ,$$

πράγμα άτοπον. Άρα $\hat{\xi} = \xi$. ▲

<u>Παρατήρηση</u> Από την (6.16), θέτοντας $\varepsilon_\kappa = x_\kappa - \xi$, $\kappa = 0,1,2,\ldots$, έχουμε $|\varepsilon_{\kappa+1}| \leqq L|\varepsilon_\kappa|$, $\kappa = 0,1,2,\ldots$. Έτσι η ποσότης L μπορεί να χρησιμοποιηθεί σαν ένα μέτρο της ταχύτητας με την οποία η επαναληπτική μέθοδος συγκλίνει. Όσο μικρότερο είναι το L τόσο ταχύτερη θα είναι η σύγκλιση.

<u>Θεώρημα 6.3</u> Εάν ξ είναι ένα σταθερό σημείο της συναρτήσεως $g(x)$ της μεθόδου (6.8), η $g(x)$ είναι ορισμένη και παραγωγίσιμη

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

στο $I \equiv [\xi - \varepsilon, \xi + \varepsilon]$, $\varepsilon > 0$ και για κάθε $x \in (\xi - \varepsilon, \xi + \varepsilon)$ έχουμε $|g'(x)| \leq L < 1$, τότε για κάθε $x_0 \in I$ η ακολουθία που παράγεται από τον αλγόριθμο $x_{n+1} = g(x_n)$, $n = 0,1,2,\ldots$ συγκλίνει στο ξ που είναι και η μοναδική ρίζα της $x = g(x)$ στο I .

<u>Απόδειξη</u> Αυτή γίνεται όπως και στο θεώρημα 6.2,με την διαφορά ότι για την απόδειξη της (6.16) εφαρμόζουμε τώρα το θεώρημα μέσης τιμής. Έτσι, έχουμε λόγω των υποθέσεων

$$|x_{K+1} - \xi| = |g(x_K) - g(\xi)| = |(x_K - \xi)g'(\hat{x}_K)| =$$

$$= |g'(\hat{x}_K)||x_K - \xi| \leq L|x_K - \xi| \leq \ldots \leq L^{K+1}|x_0 - \xi| < |x_0 - \xi|,$$

αφού το $\hat{x}_K$ βρίσκεται μεταξύ των x_K και ξ ,οπότε, $\hat{x}_K \in (\xi - \varepsilon, \xi + \varepsilon)$.▲

<u>Θεώρημα 6.4</u> Δίνεται η επαναληπτική μέθοδος

$$x_{n+1} = g(x_n) \ , \ n = 0,1,2,\ldots \tag{6.17}$$

και έστω ξ μία ρίζα της $x = g(x)$. Εάν: i) η $g(x)$ είναι επαρκώς παραγωγίσιμη με συνεχείς παραγώγους σε μία περιοχή της ξ, ii) $g^{(K)}(\xi) = 0$, για $K = 1,2,\ldots, p-1$ και iii) $g^{(p)}(\xi) \neq 0$, όπου $p \geq 1$ και για $p = 1$, εκτός από την $g'(\xi) \neq 0$, υποθέτουμε επιπλέον ότι ισχύει $|g'(\xi)| < 1$, τότε η επαναληπτική μέθοδος (6.17) είναι τοπικώς συγκλίνουσα στην ρίζα ξ και είναι τάξεως p για τον προσδιορισμό της ξ .

<u>Απόδειξη</u> Έστω $p = 1$. Τότε ισχύει $0 < |g'(\xi)| < 1$. Επειδή η $g'(x)$ είναι συνεχής πλησίον της ξ συνεπάγεται ότι θα υπάρχει μία περιοχή $\Pi(\xi) \equiv [\xi - \varepsilon, \xi + \varepsilon]$, $\varepsilon > 0$ της ξ, τέτοια ώστε $|g'(x)| \leq L < 1$, $\forall\, x \in \Pi(\xi)$. Τότε, σύμφωνα με το θεώρημα 6.3,για κάθε $x_0 \in \Pi(\xi)$ η ακολουθία που παράγεται από την (6.17) συγκλίνει στη ρίζα ξ . Επίσης, στην $\Pi(\xi)$, δηλαδή, $\forall\, x_0 \in \Pi(\xi)$ θα έχουμε

$$x_{n+1} = g(x_n) = g(\xi) + (x_n - \xi)g'(\xi_n) = \xi + (x_n - \xi)g'(\xi_n),$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

για κάποιο ξ_n που βρίσκεται μεταξύ των x_n και ξ, οπότε

$$|x_{n+1} - \xi| = |\varepsilon_{n+1}| = |\varepsilon_n| \, |g'(\xi_n)| \, . \tag{6.18}$$

Από την (6.18), λόγω της $\lim\limits_{n\to\infty} x_n = \xi$, προκύπτει ότι

$$\lim_{n\to\infty} \frac{|\varepsilon_{n+1}|}{|\varepsilon_n|} = |g'(\xi)| \equiv C \quad \text{με} \quad 0 < C < 1 \, .$$

Άρα, η σύγκλιση (βλέπε ορισμό 6.1) είναι γραμμική και η μέθοδος (6.17) είναι πρώτης τάξεως για τον προσδιορισμό της ρίζας ξ.

Για $p > 1$, η τοπική σύγκλιση αποδεικνύεται όπως και στην περίπτωση $p = 1$, αφού τώρα είναι $g'(\xi) = 0$ και επομένως $|g'(\xi)| < 1$. Επί πλέον, χρησιμοποιώντας τον τύπο του Taylor και τις υποθέσεις του θεωρήματος, $\forall\, x_0 \in \Pi(\xi)$ = περιοχή συγκλίσεως παίρνουμε

$$x_{n+1} = g(x_n) = g(\xi) + (x_n - \xi)g'(\xi) + (x_n - \xi)^2 \frac{g''(\xi)}{2!} + \ldots +$$

$$+ (x_n - \xi)^{p-1} \frac{g^{(p-1)}(\xi)}{(p-1)!} + (x_n - \xi)^p \frac{g^{(p)}(\xi_n)}{p!} = \xi +$$

$$+ (x_n - \xi)^p \frac{g^{(p)}(\xi_n)}{p!} \, ,$$ για κάποιο ξ_n μεταξύ των x_n και ξ,

οπότε

$$|\varepsilon_{n+1}| = |\varepsilon_n|^p \frac{|g^{(p)}(\xi_n)|}{p!} \, . \tag{6.19}$$

Από την (6.19) προκύπτει τώρα, λόγω της $\lim\limits_{n\to\infty} x_n = \xi$, ότι

$$\lim_{n\to\infty} \frac{|\varepsilon_{n+1}|}{|\varepsilon_n|^p} = \frac{1}{p!} |g^{(p)}(\xi)| \equiv C > 0 \, ,$$

που σημαίνει ότι η μέθοδος (6.17) είναι τάξεως p $(p > 1)$ για τον προσδιορισμό της ρίζας ξ. ▲

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Σημείωση Αν παραλείψουμε από τις υποθέσεις του θεωρήματος 6.4 την συνθήκη (iii), τότε η μέθοδος (6.17) είναι τουλάχιστον p τάξεως για τον προσδιορισμό της ξ. (Για $p = 1$ απαιτείται μόνον η $|g'(\xi)| < 1$, ενώ για $p = 2$ η $g'(\xi) = 0$).

Βασιζόμενοι στο θεώρημα 6.4, προσπαθούμε στην πράξη να κατασκευάσουμε επαναληπτικές μεθόδους όπου $p \geq 2$, δηλαδή, για τις οποίες ισχύει $g'(\xi) = 0$, διότι, αφ' ενός εξασφαλίζεται τοπική σύγκλιση, αφ' ετέρου δε η σύγκλιση είναι πολύ γρήγορη. Γενικά όμως, όσο πιό μεγάλη είναι η τάξη p, τόσο πιό πολύπλοκη είναι η μορφή της επαναληπτικής μεθόδου, με αποτέλεσμα περισσότεροι υπολογισμοί να απαιτούνται για κάθε επανάληψη. Έτσι, το κέρδος από την μεγαλύτερη τάξη συγκλίσεως δεν πρέπει να μετριέται μόνον με βάση τον αριθμό των επαναλήψεων, αλλά και από την συνολική προσπάθεια, δηλαδή, τον ολικό χρόνο υπολογισμού για την εύρεση του τελικού αποτελέσματος. Για τον λόγο αυτό, περισσότερο χρησιμοποιούνται στην πράξη μέθοδοι πρώτης και δευτέρας τάξεως.

Θεώρημα 6.5 Εάν η συνάρτηση $g(x)$ της μεθόδου (6.8) είναι ορισμένη και παραγωγίσιμη σε ένα διάστημα $[a,b]$, έχει ένα μοναδικό σταθερό σημείο ξ στο $[a,b]$ και για κάθε $x \in (a,b)$ ισχύει $|g'(x)| \leq L < 1$, τότε για κάθε x_0 που ανήκει στο μικρότερο από τα δύο διαστήματα $[a,\xi]$, $[\xi,b]$, η ακολουθία που παράγεται από τον αλγόριθμο $x_{n+1} = g(x_n)$, $n = 0,1,2,\ldots$ συγκλίνει (τουλάχιστον γραμμικά) στο ξ.

Απόδειξη Εάν $[a,\xi]$ είναι το μικρότερο διάστημα (δηλαδή, $\xi < \frac{a+b}{2}$), τότε θεωρούμε το διάστημα $I \equiv [\xi-\varepsilon, \xi+\varepsilon]$ με $\varepsilon = \xi - a$, δηλαδή, $I \equiv [a, 2\xi - a]$, για το οποίο ισχύει $[a,\xi] \subset I \subset [a,b]$, αφού $2\xi - a < b$. Έτσι, σύμφωνα με το θεώρημα 6.3, για κάθε $x_0 \in I$ και επομένως για κάθε $x_0 \in [a,\xi]$, η ακολουθία που παράγεται από τον αλγόριθμο $x_{n+1} = g(x_n)$, $n = 0,1,2,\ldots$ συγκλίνει στο ξ. Ανάλογη είναι η απόδειξη όταν το $[\xi,b]$ είναι το μικρότερο

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

διάστημα.

▲

Ορισμός 6.2 Μία συνάρτηση $g(x)$ λέγεται ότι είναι μία απεικόνιση συστολής (contraction or contractive mapping) σε ένα διάστημα $[a,b]$ εάν

i) $\forall\, x \in [a,b] \Rightarrow g(x) \in [a,b]$

ii) $\forall\, x,y \in [a,b] \Rightarrow |g(x) - g(y)| \leqq L|x - y|\ ,\quad 0 \leqq L < 1\ .$

Θεώρημα 6.6 Εάν η επαναληπτική συνάρτηση $g(x)$ της μεθόδου (6.8) είναι ορισμένη σε ένα διάστημα $[a,b]$ και είναι μία απεικόνιση συστολής στο $[a,b]$, τότε ισχύουν τα ακόλουθα:

i) Η εξίσωση $x = g(x)$ έχει μία μοναδική ρίζα ξ στο $[a,b]$.

ii) Για κάθε $x_0 \in [a,b]$ η ακολουθία $x_{n+1} = g(x_n)$, $n = 0,1,2,\ldots$ συγκλίνει στη ρίζα ξ ,

$$|x_n - \xi| \leqq \frac{L^n}{1-L}|x_1 - x_0|\ ,\quad |x_{n+1} - \xi| < |x_n - \xi| \quad \text{και}$$

$$|x_{n+1} - \xi| \leqq \frac{L}{1-L}|x_{n+1} - x_n|\ ,\quad n = 0,1,2,\ldots\ .$$

Απόδειξη Προφανώς, ισχύει $x_n \in [a,b]$ $n = 1,2,\ldots$ για κάθε $x_0 \in [a,b]$. Για δύο διαδοχικές τιμές x_n, x_{n+1} έχουμε

$$|x_{n+1} - x_n| = |g(x_n) - g(x_{n-1})| \leqq L\ |x_n - x_{n-1}| \leqq \ldots \leqq$$
$$\leqq L^n |x_1 - x_0|\ . \tag{6.20}$$

Θεωρούμε την σειρά $\sum_{n=0}^{\infty} c_n$, όπου $c_0 = x_0$, $c_{n+1} = x_{n+1} - x_n$, $n = 0,1,2,\ldots$, δηλαδή,

$$\sum_{n=0}^{\infty} c_n = x_0 + (x_1 - x_0) + (x_2 - x_1) + \ldots + (x_n - x_{n-1}) + \ldots\ , \tag{6.21}$$

της οποίας τα μερικά αθροίσματα είναι $S_{n+1} = x_n$, $n = 0,1,2,\ldots$,

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

καθώς επίσης και τη σειρά $\sum_{n=0}^{\infty} d_n$, όπου $d_0 = |x_0|$, $d_{n+1} =$

$= L^n |x_1 - x_0|$, $n = 0,1,2,\ldots$, η οποία προφανώς συγκλίνει διότι οι όροι της εκτός του πρώτου αποτελούν γεωμετρική πρόοδο με λόγο $L < 1$. Επειδή, λόγω της (6.20), ισχύει $|c_n| \leqq |d_n|$, $n = 0,1,2,\ldots$, έπεται ότι η σειρά $\sum_{n=0}^{\infty} |c_n|$ συγκλίνει, οπότε θα συγκλίνει και η (6.21), αφού αυτή συγκλίνει απολύτως.

Έστω, $\lim_{n\to\infty} S_{n+1} \equiv \lim_{n\to\infty} x_n = \xi$. Προφανώς, $\xi \in [a,b]$ και από την $x_{n+1} = g(x_n)$, $n = 0,1,2,\ldots$, παίρνοντας όρια, έχουμε $\xi = g(\xi)$, δηλαδή, ξ είναι μία ρίζα της $x = g(x)$. Αν $\hat{\xi}$ είναι μία άλλη ρίζα, $\hat{\xi} = g(\hat{\xi})$, $\hat{\xi} \in [a,b]$, τότε έχουμε

$$|\hat{\xi} - \xi| = |g(\hat{\xi}) - g(\xi)| \leqq L|\hat{\xi} - \xi| \quad \text{ή} \quad (1 - L)|\hat{\xi} - \xi| \leqq 0 \quad \text{ή}$$

$$\hat{\xi} = \xi .$$

Λόγω της (6.20) έχουμε

$$|x_{n+K} - x_n| \leqq |x_{n+K} - x_{n+K-1}| + |x_{n+K-1} - x_{n+K-2}| + \ldots +$$

$$|x_{n+1} - x_n| \leqq L^{n+K-1}|x_1 - x_0| + L^{n+K-2}|x_1 - x_0| + \ldots + L^n|x_1 - x_0| =$$

$$= L^n(1 + L + L^2 + \ldots + L^{K-1})|x_1 - x_0| = L^n \left(\frac{1 - L^K}{1 - L}\right)|x_1 - x_0| \leqq$$

$$\leqq \frac{L^n}{1 - L}|x_1 - x_0| . \qquad (6.22)$$

Από την (6.22), παίρνοντας όρια για $K \to \infty$ και έχοντας υπ' όψιν ότι $\lim_{K\to\infty} x_{n+K} = \xi$, προκύπτει τελικά ότι

$$|\xi - x_n| = |x_n - \xi| \leqq \frac{L^n}{1 - L}|x_1 - x_0| , \; n = 0,1,2,\ldots . \qquad (6.23)$$

Επίσης, έχουμε

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$|x_{n+1} - \xi| = |g(x_n) - g(\xi)| \leqq L|x_n - \xi| < |x_n - \xi| \ ,$$
$$n = 0,1,2,\ldots, \quad (6.24)$$

δηλαδή, $|\varepsilon_{n+1}| < |\varepsilon_n|$, $n = 0,1,2,\ldots$, που σημαίνει ότι η ακολουθία των απολύτων σφαλμάτων $|\varepsilon_n| = |x_n - \xi|$, $n = 0,1,2,\ldots$ είναι γνησίως φθίνουσα. Έτσι, κάθε νέα επανάληψη είναι πιο ακριβής από την προηγούμενη.

Ακόμη, αφού $\xi = g(\xi)$ παίρνουμε

$$|x_{n+1} - g(x_{n+1})| = |x_{n+1} - \xi + g(\xi) - g(x_{n+1})| \geqq |x_{n+1} - \xi| -$$
$$- |g(\xi) - g(x_{n+1})| \geqq |x_{n+1} - \xi| - L|x_{n+1} - \xi| = (1 - L)|x_{n+1} - \xi|$$

ή

$$|x_{n+1} - \xi| \leqq \frac{|x_{n+1} - g(x_{n+1})|}{1 - L} = \frac{|x_{n+2} - x_{n+1}|}{1 - L} \ . \quad (6.25)$$

Η (6.25), λόγω της (6.20) γίνεται

$$|x_{n+1} - \xi| \leqq \frac{L}{1 - L} |x_{n+1} - x_n| \ , \quad n = 0,1,2,\ldots \ . \quad \blacktriangle \quad (6.26)$$

<u>Θεώρημα 6.7</u> Εάν η επαναληπτική συνάρτηση $g(x)$ της μεθόδου (6.8) είναι ορισμένη και παραγωγίσιμη σε ένα διάστημα $[a,b]$ και

i) $\forall\, x \in [a,b] \Rightarrow g(x) \in [a,b]$,

ii) $\forall\, x \in (a,b) \Rightarrow |g'(x)| \leqq L < 1$,

τότε ισχύουν τα συμπεράσματα του θεωρήματος 6.6 .

<u>Απόδειξη</u> Είναι εύκολο να δειχθεί, εφαρμόζοντας το θεώρημα μέσης τιμής, ότι η $g(x)$ είναι μία απεικόνιση συστολής στο $[a,b]$ και επομένως η απόδειξη είναι όμοια με εκείνη του θεωρήματος 6.6 . ▲

<u>Παρατηρήσεις</u>: i) Τα θεωρήματα 6.6 και 6.7 ισχύουν και στην περίπτωση που το διάστημα $[a,b]$ είναι το $(-\infty , +\infty)$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ii) Κατά την ανάπτυξη της θεωρίας συγκλίσεως της επαναληπτικής μεθόδου αγνοήσαμε τελείως την ύπαρξη των αναπόφευκτων στην πράξη σφαλμάτων στρογγυλεύσεως, δηλαδή, υποθέσαμε ότι οι διαδοχικές επαναλήψεις (προσεγγίσεις) υπολογίζονται ακριβώς.

iii) Εάν είναι $L \leqq \frac{1}{2}$, τότε από την (6.26) έχουμε $|x_{n+1} - \xi| \leqq$ $\leqq |x_{n+1} - x_n|$. Έτσι, όταν ικανοποιηθεί η ανισότητα $|x_{n+1} - x_n| \leqq \varepsilon$, τότε θα είναι και $|\varepsilon_{n+1}| \leqq \varepsilon$. Ακόμη, από την (6.26) είναι φανερό ότι για να βρούμε μία επανάληψη x_n τέτοια ώστε $|x_n - \xi| \leqq \varepsilon$, αρκεί να ικανοποιηθεί η ανισότητα $|x_n - x_{n-1}| \leqq \frac{1-L}{L}\varepsilon$ $(0 \leqq L < 1)$.

iv) Το x_0 στα θεωρήματα 6.6 και 6.7 μπορεί να εκλεγεί αυθαίρετα από το $[a,b]$. Εάν όμως ληφθεί $x_0 = \frac{a+b}{2}$ και επιπλέον ισχύει $|x_1 - x_0| \leqq (1-L)\left(\frac{b-a}{2}\right)$, τότε, μπορεί να αποδειχθεί (βλέπε άσκηση 6) ότι η συνθήκη (i) του ορισμού 6.2 μπορεί να παραλειφθεί από τις υποθέσεις των θεωρημάτων 6.6 και 6.7 .

v) Εάν η εξίσωση $x = g(x)$ έχει μία μοναδική ρίζα ξ στο $[a,b]$ και $\forall\, x \in [a,b]$ ισχύει $|g'(x)| \geqq L > 1$, τότε, εκτός από ορισμένες εξαιρέσεις (βλέπε π.χ. άσκηση 7), $\forall\, x_0 \in [a,b]$, $x_0 \neq \xi$, η επαναληπτική μέθοδος (6.8) δεν θα συγκλίνει.

vi) Επειδή στην ακολουθία των επαναλήψεων $x_0, x_1, x_2, \ldots,$ που προκύπτει από την επαναληπτική μέθοδο, κάθε όρος x_n μπορεί να θεωρηθεί σαν η αρχική προσέγγιση x_0, μικρά σφάλματα κατά τον υπολογισμό των διαφόρων επαναλήψεων δεν θεωρούνται άξια προσοχής, με την έννοια ότι, εάν η ακολουθία συγκλίνει, αυτά δεν μπορούν, φυσιολογικά, να εμποδίσουν την σύγκλιση της ακολουθίας στην ζητούμενη ρίζα. Για τον λόγο αυτό η επαναληπτική μέθοδος λέγεται ότι είναι αυτοδιορθωτική.

Σχετικά με το ερώτημα πώς κατασκευάζεται ή πώς εκλέγεται μία κατάλληλη επαναληπτική συνάρτηση $g(x)$, πρέπει να σημειωθούν τα εξής: Εάν δοθεί μία εξίσωση $f(x) = 0$, τότε είναι δυνατόν, με κατάλληλους μετασχηματισμούς, να γράψουμε αυτή με την ισοδύναμη μορφή $x = g(x)$, για διαφορετικές συναρτήσεις $g(x)$ και επομένως να δημιουργήσουμε

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

αρκετές διαφορετικές επαναληπτικές μεθόδους. Μερικές απ' αυτές είναι δυνατόν να χρησιμοποιηθούν για την εύρεση όλων των πραγματικών ριζών της $f(x) = 0$, ορισμένες δεν μπορούν να χρησιμοποιηθούν για καμμία ρίζα και άλλες μπορούν να χρησιμοποιηθούν για την εύρεση ορισμένων ριζών. Έτσι, στην πράξη, πολλές φορές χρειαζόμαστε να χρησιμοποιήσουμε αρκετές διαφορετικές συναρτήσεις $g(x)$ για να προσδιορίσουμε όλες τις ρίζες της $f(x)$. Μεταξύ των άλλων εκλογών για την $g(x)$ είναι και οι ακόλουθες δύο:

i) $g(x) = x - cf(x)$, όπου c σταθερά $\neq 0$, (6.27)

ii) $g(x) = x - h(x)f^m(x)$, (6.28)

όπου $h(x) \neq 0$ και m θετική σταθερά.

Πρέπει να αναφερθεί ακόμη, ότι πολλές φορές η $x = g(x)$ δεν είναι ισοδύναμη με την $f(x) = 0$, με την έννοια ότι το σύνολο των σταθερών σημείων της $g(x)$ είναι υποσύνολο των ριζών της $f(x) = 0$ (σε ειδικές περιπτώσεις η $g(x)$ μπορεί να περιέχει και μερικά σταθερά σημεία τα οποία δεν είναι ρίζες της $f(x) = 0$) . Γενικά, η $g(x)$ πρέπει να εκλεγεί έτσι ώστε, όχι μόνον τα σταθερά της σημεία να είναι ρίζες της $f(x) = 0$, αλλά και η ακολουθία $x_{n+1} = g(x_n)$, $n = 0,1,2,...$ να συγκλίνει στα σταθερά αυτά σημεία. Επειδή, σύμφωνα με την θεωρία συγκλίσεως, σε μια κατάλληλη $g(x)$ η $|g'(x)|$ είναι μικρός αριθμός σε ένα διάστημα που περιέχει μια ρίζα ξ , συνιστάται κατά την μετατροπή της εξισώσεως $f(x) = 0$ στην $x = g(x)$ η χρησιμοποίηση τετραγωνικών ριζών, κυβικών ριζών ή λογαρίθμων, επειδή οι συναρτήσεις αυτές είναι σχετικά επίπεδες σε ένα μεγάλο διάστημα. Ακόμη, μερικές φορές, είναι προτιμώτερο να λύνουμε ως προς x από τον όρο της $f(x)$ που έχει τον συντελεστή με το μεγαλύτερο μέτρο (π.χ. όταν η $f(x)$ είναι ένα πολυώνυμο).

Όπως τονίστηκε και προηγουμένως, προσπαθούμε στην πράξη να κατασκευάσουμε μεθόδους που να είναι τουλάχιστον δευτέρας τάξεως για τον προσδιορισμό κάποιας ρίζας ξ . Έτσι, αν θεωρήσουμε την επανα-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ληπτική μέθοδο $x_{n+1} = g(x_n)$, $n = 0,1,2,\ldots$, όπου $g(x) = x - h(x)f(x)$ (βλέπε (6.28) για $m = 1$) , προσπαθούμε να προσδιορίσουμε μία συνάρτηση $h(x)$, τέτοια ώστε $g'(\xi) = 0$, όπου ξ η ζητούμενη ρίζα της $f(x) = 0$. Παραγωγίζοντας, έχουμε $g'(x) = 1 - h'(x)f(x) - h(x)f'(x)$. θέτοντας $g'(\xi) = 0$ και εάν $f'(\xi) \neq 0$ (ξ απλή ρίζα), έχουμε $h(\xi) = \frac{1}{f'(\xi)}$. Άρα, μπορούμε να εκλέξουμε $h(x) \equiv \frac{1}{f'(\xi)}$. Επειδή όμως η ρίζα ξ είναι άγνωστη, μία καλύτερη εκλογή είναι η $h(x) \equiv \frac{1}{f'(x)}$, $f'(x) \neq 0$, οπότε $g(x) = x - \frac{f(x)}{f'(x)}$. Έτσι, προκύπτει η ακόλουθη σημαντική επαναληπτική μέθοδος:

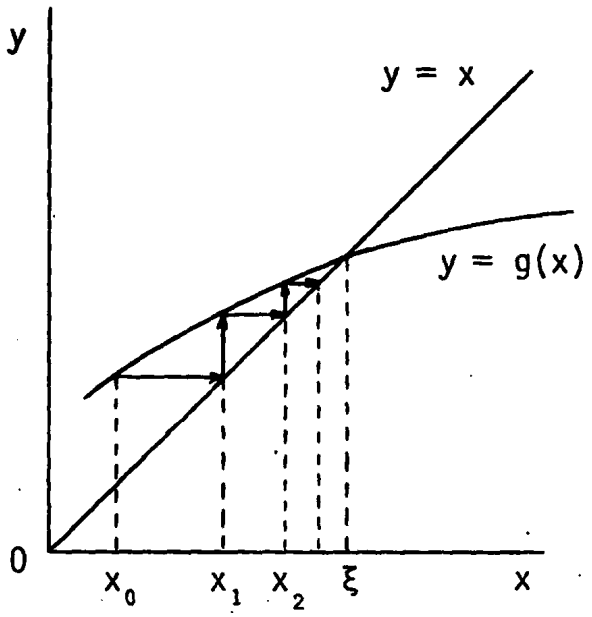
$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0,1,2,\ldots, \tag{6.29}$$

που είναι γνωστή σαν μέθοδος των Newton – Raphson. Η μέθοδος (6.29) θα μελετηθεί αναλυτικώτερα σε επόμενη παράγραφο.

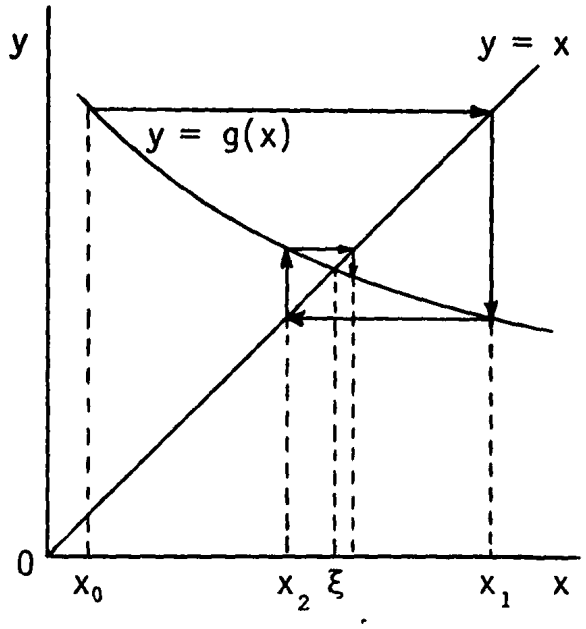
6.4.2 Γεωμετρική ερμηνεία της επαναληπτικής μεθόδου

Η γεωμετρική ερμηνεία της επαναληπτικής μεθόδου είναι απλή. Σχεδιάζουμε στο επίπεδο με συντεταγμένες x,y τις συναρτήσεις $y = x$ και $y = g(x)$. Οι τετμημένες των σημείων τομής αυτών των καμπύλων είναι οι πραγματικές ρίζες της εξισώσεως $x = g(x)$. Η εύρεση των διαδοχικών επαναλήψεων γίνεται γεωμετρικά ως εξής: Αρχίζοντας από ένα σημείο $(x_0,x_1) \equiv (x_0,g(x_0))$ φέρνουμε την ευθεία $y = x_1$ μέχρι το σημείο τομής της με την ευθεία $y = x$ και έπειτα από το σημείο τομής φέρνουμε την ευθεία $x = x_1$ μέχρι το σημείο τομής της με την $y = g(x)$. Έτσι, βρίσκουμε το σημείο (x_1,x_2) κ.ο.κ. . Τα σχήματα 6.2, 6.3, 6.4, 6.5 δείχνουν την συμπεριφορά των διαδοχικών προσεγγίσεων στις περιπτώσεις i) $0 < g'(x) < 1$, ii) $-1 < g'(x) < 0$, iii) $1 < g'(x)$ και iv) $g'(x) < -1$. Για $g'(x) > 0$ η συμπεριφορά είναι μονότονος (μονοπλευρική σύγκλιση ή απόκλιση), ενώ για $g'(x) < 0$ είναι ταλαντωτική (διπλευρική σύγκλιση ή απόκλιση).

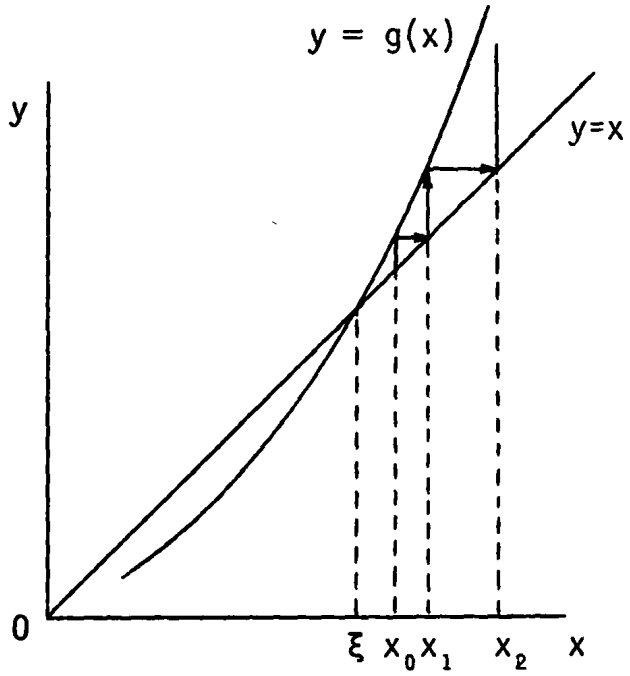
ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ



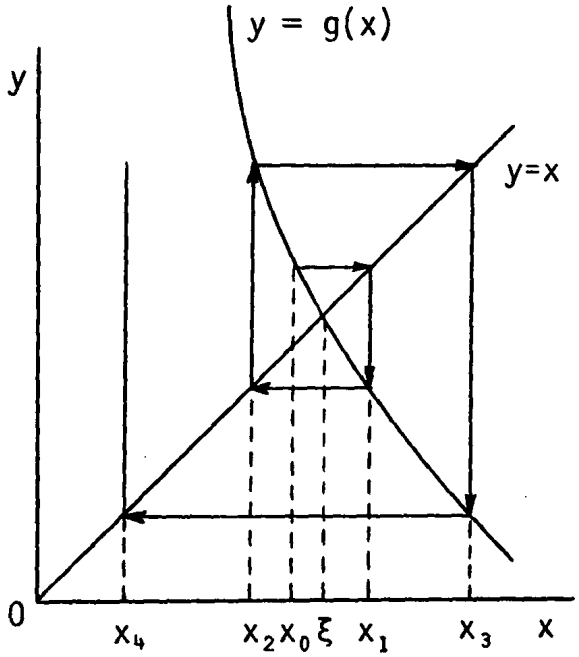
σχ. 6.2 $(0 < g'(x) < 1)$



σχ. 6.3 $(-1 < g'(x) < 0)$



σχ. 6.4 $(g'(x) > 1)$



σχ. 6.5 $(g'(x) < -1)$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Παράδειγμα 6.4 Δίνεται η εξίσωση $f(x) = x^2 - a = 0$, όπου a πραγματικός αριθμός > 1 . Να δειχθεί ότι η επαναληπτική μέθοδος $x_{n+1} = g(x_n)$, $n = 0,1,2,\ldots$, όπου $g(x) = \frac{a(x+1)}{x+a}$ συγκλίνει $\forall\, x_0 \geqq 0$ στην ρίζα $\xi = \sqrt{a}$ της $f(x) = 0$.

Λύση Η εξίσωση $x = g(x)$ είναι ισοδύναμη με την $f(x) = 0$. Θεωρούμε το διάστημα $I = [0, +\infty)$. Η $g(x)$ είναι ορισμένη και παραγωγίσιμη στο I και επίσης $\forall\, x \in I \Rightarrow g(x) \in I$. Παραγωγίζοντας έχουμε $g'(x) = \frac{a(a-1)}{(x+a)^2}$. Προφανώς, $\forall\, x \in I$ είναι $0 < g'(x) < 1$. Επομένως, σύμφωνα με το θεώρημα 6.7, $\forall\, x_0 \in I$ η ακολουθία $x_{n+1} = g(x_n)$, $n = 0,1,2,\ldots$ συγκλίνει στην ρίζα $\xi = \sqrt{a} \in I$. Για $a = 2$ και $x_0 = 1$ οι 4 πρώτες επαναλήψεις είναι: $x_1 = \frac{4}{3}$, $x_2 = 1.4$, $x_3 = \frac{24}{17}$, $x_4 = \frac{41}{29} \approx 1.414$. ▲

Παράδειγμα 6.5 Αφού δειχθεί ότι η εξίσωση $f(x) = x2^x - 1 = 0$ έχει μία μοναδική ρίζα ξ στο $[0, +\infty)$ και ότι ο αλγόριθμος $x_{n+1} = \frac{1}{2^{x_n}}$, $n = 0,1,2,\ldots$ συγκλίνει στη ρίζα ξ $\forall\, x_0 \geqq 0$, να βρεθεί η ρίζα ξ με ακρίβεια 6 δ.θ. .

Λύση Η $f(x) = 0$ είναι ισοδύναμη με την $x = \frac{1}{2^x} \equiv g(x)$. Επίσης, έχουμε $g(x) > 0$, $\forall\, x \geqq 0$ και $g'(x) = -\frac{\ln 2}{2^x}$, οπότε $|g'(x)| = \frac{\ln 2}{2^x} < 1$, $\forall\, x \geqq 0$, αφού $\ln 2 \approx 0.69314718$. Έτσι, σύμφωνα με το θεώρημα 6.7, η $f(x) = 0$ έχει μία μοναδική ρίζα ξ στο $[0, +\infty)$ (πιο συγκεκριμένα $\xi \in [0,1]$, αφού $f(0) = -1$ και $f(1) = 1$) και ο αλγόριθμος $x_{n+1} = \frac{1}{2^{x_n}}$, $n = 0,1,2,\ldots$ συγκλίνει στη ρίζα ξ $\forall\, x_0 \geqq 0$. Λαμβάνοντας π.χ. $x_0 = 1$, μπορεί να βρεθεί ότι $\xi \approx 0.641185$. ▲

Παράδειγμα 6.6 Για την εύρεση της ρίζας $\xi = \sqrt[3]{a}$, $a > 0$ της εξισώσεως $f(x) = x(x^3 - a) = 0$, προτείνεται η επαναληπτική μέθοδος $x_{n+1} = \frac{x_n(2a + x_n^3)}{2x_n^3 + a}$, $n = 0,1,2,\ldots$. Να δειχθεί ότι αυτή είναι τρίτης

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

τάξεως για τον προσδιορισμό της ξ , για κατάλληλο x_0 .

<u>Λύση</u> Η εξίσωση $x = g(x)$, όπου $g(x) = \dfrac{x(2a + x^3)}{2x^3 + a}$ είναι ισοδύναμη με την $f(x) = 0$. Είναι εύκολο να δειχθεί ότι $g'(x) = \dfrac{2(x^3 - a)^2}{(2x^3 + a)^2}$, οπότε $g'(\xi) = 0$. Ακόμη, μπορεί να δειχθεί ότι $g''(\xi) = 0$ και $g'''(\xi) = \dfrac{4}{a^{2/3}} \neq 0$. Έτσι, σύμφωνα με το θεώρημα 6.4 η δοθείσα επαναληπτική μέθοδος είναι τοπικώς συγκλίνουσα στην ρίζα $\xi = \sqrt[3]{a}$ και τάξεως 3 για τον προσδιορισμό αυτής. ▲

6.5 Η μέθοδος Newton–Raphson

Όπως είδαμε στο τέλος της παραγράφου 6.4.1, η μέθοδος Newton–Raphson μπορεί να θεωρηθεί σαν μία ειδική περίπτωση της επαναληπτικής μεθόδου, όπου η επαναληπτική συνάρτηση έχει την μορφή

$$g(x) = x - \frac{f(x)}{f'(x)} \tag{6.30}$$

και επομένως οι διαδοχικές επαναλήψεις δίνονται από τον αλγόριθμο

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} , \quad n = 0,1,2,\dots , \tag{6.31}$$

όπου x_0 η αρχική προσέγγιση της ζητούμενης ρίζας ξ της $f(x) = 0$.

Ένας άλλος τρόπος παραγωγής της μεθόδου Newton – Raphson είναι ο ακόλουθος, που βασίζεται στο θεώρημα του Taylor:

Εάν ξ είναι μία ρίζα της $f(x) = 0$ και x_n μία προσέγγιση αυτής, τότε έχουμε

$$f(\xi) = f(x_n) + (\xi - x_n)f'(x_n) + \dots \approx f(x_n) + (\xi - x_n)f'(x_n) .$$

Άρα, $f(x_n) + (\xi - x_n)f'(x_n) \approx 0$ και λύνοντας ως προς ξ (με την προυπόθεση ότι $f'(x_n) \neq 0$) παίρνουμε

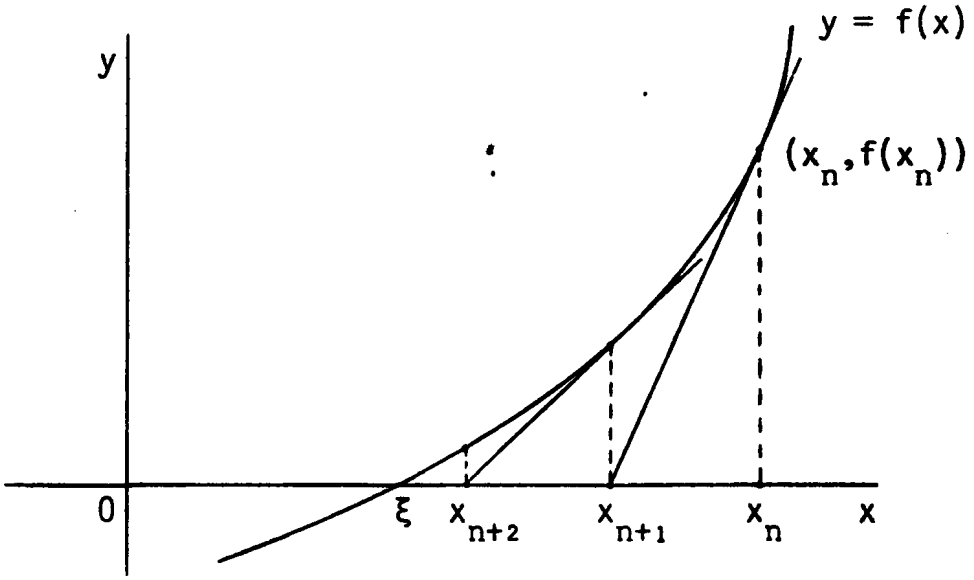
$$\xi \approx x_n - \frac{f(x_n)}{f'(x_n)} ,$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

που σημαίνει ότι η $x_{n+1} = x_n - \dfrac{f(x_n)}{f'(x_n)}$, θα πρέπει να είναι μία καλύτερη προσέγγιση για την ρίζα ξ απ' ότι η x_n . Έτσι, προκύπτει πάλι η (6.31).

Από την (6.31) είναι φανερό ότι για τον υπολογισμό της τιμής x_{n+1} απαιτούνται οι τιμές $f(x_n)$ και $f'(x_n)$. Πρέπει να σημειωθεί ότι, εάν η $f(x)$ είναι ένα πολυώνυμο, τότε οι $f(x_n)$ και $f'(x_n)$ υπολογίζονται χρησιμοποιώντας το σχήμα του Horner (βλέπε κεφ. 7).

6.5.1 Γεωμετρική ερμηνεία της μεθόδου Newton-Raphson



σχ. 6.6

Το σχήμα 6.6 δείχνει την γραφική παράσταση μιας συναρτήσεως $f(x)$ και ξ είναι η ζητούμενη ρίζα της $f(x) = 0$. Στο σημείο $(x_n, f(x_n))$ φέρνουμε την εφαπτόμενη της καμπύλης. Η εξίσωση αυτής είναι

$$y - f(x_n) = f'(x_n)(x - x_n) \ . \qquad (6.32)$$

Αν $(x_{n+1}, 0)$ είναι το σημείο τομής της εφαπτομένης με τον άξονα των x , τότε από την (6.32) έχουμε $0 - f(x_n) = f'(x_n)(x_{n+1} - x_n)$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ή $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$, δηλαδή, παίρνουμε πάλι την (6.31) .

6.5.2 Σύγκλιση και τάξη της μεθόδου Newton–Raphson

Έστω ξ μία ρίζα της $f(x) = 0$ και $[a,b]$ ένα διάστημα, το οποίο περιέχει την ξ και στο οποίο οι $f'(x)$, $f''(x)$ υπάρχουν και είναι συνεχείς. Επί πλέον, υποθέτουμε ότι $f'(\xi) \neq 0$. Από την (6.30) παίρνουμε

$$g'(x) = \frac{f(x)f''(x)}{[f'(x)]^2} , \tag{6.33}$$

για κάθε x τέτοιο ώστε $f'(x) \neq 0$. Επομένως, αφού $f(\xi) = 0$ και $f'(\xi) \neq 0$ προκύπτει ότι

$$g'(\xi) = 0 . \tag{6.34}$$

Ακόμη, υποθέτοντας ότι η $f'''(x)$ υπάρχει και είναι συνεχής στο $[a,b]$, μπορεί να δειχθεί παραγωγίζοντας την (6.33) ότι

$$g''(\xi) = \frac{f''(\xi)}{f'(\xi)} . \tag{6.35}$$

Από τις (6.34), (6.35) συμπεραίνουμε (βλέπε θεώρημα 6.4) ότι η μέθοδος Newton – Raphson, κάτω από τις προυποθέσεις που θεωρήσαμε, είναι τουλάχιστον δευτέρας τάξεως για τον προσδιορισμό της ρίζας ξ . (Η $g''(\xi)$ θα είναι ίση με 0 αν $f''(\xi) = 0$). Έτσι, απεδείχθη το ακόλουθο θεώρημα.

<u>Θεώρημα 6.8</u> Εάν η συνάρτηση $f(x)$ έχει συνεχείς παραγώγους $f'(x)$, $f''(x)$ σε μια περιοχή μιας απλής ρίζας ξ της $f(x) = 0$, τότε για κάθε x_0 που είναι πολύ κοντά στη ρίζα ξ οι επαναλήψεις που παράγονται από τον αλγόριθμο Newton – Raphson συγκλίνουν στην ρίζα ξ και μάλιστα η σύγκλιση είναι τουλάχιστον τετραγωνική. ▲

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Παρατηρήσεις i) Το θεώρημα 6.8 έχει την έννοια ότι κάτω από τις προϋποθέσεις του υπάρχει πάντοτε μία περιοχή $I \equiv [\xi - \varepsilon, \xi + \varepsilon]$, $\varepsilon > 0$ της ρίζας ξ τέτοια ώστε $|g'(x)| \leq L < 1$, $\forall x \in I$, όπου η $g'(x)$ δίνεται από την (6.33) και επομένως εφαρμόζεται το θεώρημα 6.3 . Στην πράξη όμως, το θεώρημα 6.8 δεν είναι πολύ βοηθητικό γιατί συχνά είναι δύσκολο να προσδιορισθεί μια περιοχή της ρίζας ξ στην οποία να εφαρμόζεται το θεώρημα 6.3 και έτσι η εκλογή ενός κατάλληλου x_0 δεν είναι εύκολη. Επειδή όμως για κατάλληλο x_0 η ταχύτητά συγκλίσεως της μεθόδου Newton – Raphson είναι μεγάλη (τουλάχιστον τετραγωνική), στην πράξη, συνήθως, ακολουθούμε την εξής διαδικασία: Χρησιμοποιούμε πρώτα μία άλλη μέθοδο που σίγουρα συγκλίνει (π.χ. την μέθοδο της διχοτομήσεως) και βρίσκουμε μία προσέγγιση της ρίζας ξ που είναι αρκετά κοντά στη ρίζα ξ . Μετά, χρησιμοποιώντας την προσέγγιση αυτή σαν x_0 , εφαρμόζουμε την μέθοδο Newton – Raphson για να επιτύχουμε την επιζητούμενη ακρίβεια.

ii) Κάτω από τις προϋποθέσεις του θεωρήματος 6.8 και έχοντας υπ' όψιν την απόδειξη του θεωρήματος 6.4 και την (6.35) για τα διαδοχικά σφάλματα ε_n και ε_{n+1} ισχύει

$$\lim_{n \to \infty} \frac{\varepsilon_{n+1}}{\varepsilon_n^2} = \frac{1}{2} \frac{f''(\xi)}{f'(\xi)} \quad . \tag{6.36}$$

Θεώρημα 6.9 Εάν ξ είναι μία ρίζα πολλαπλότητος m $(m \geq 2)$ της $f(x) = 0$, όπου η $f(x)$ είναι επαρκώς παραγωγίσιμη με συνεχείς παραγώγους σε μία περιοχή της ρίζας ξ , τότε ισχύει πάλι το συμπέρασμα του θεωρήματος 6.8 με την διαφορά ότι η σύγκλιση τώρα είναι γραμμική.

Απόδειξη Προφανώς, ισχύει $f^{(\kappa)}(\xi) = 0$ για $\kappa = 0(1)m - 1$ και η $f(x)$ θα είναι της μορφής

$$f(x) = (x - \xi)^m h(x) \quad , \tag{6.37}$$

όπου $h(x)$ παραγωγίσιμη και $h(\xi) \neq 0$, οπότε

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$f'(x) = m(x-\xi)^{m-1}h(x) + (x-\xi)^m h'(x) \ . \tag{6.38}$$

Από την (6.30) παίρνουμε

$$g(x) = x - \frac{f(x)}{f'(x)} = x - \frac{(x-\xi)^m h(x)}{m(x-\xi)^{m-1}h(x) + (x-\xi)^m h'(x)} =$$

$$x - \frac{(x-\xi)h(x)}{mh(x) + (x-\xi)h'(x)} \tag{6.39}$$

και επομένως

$$g'(x) = 1 - \frac{[h(x) + (x-\xi)h'(x)]G(x) - (x-\xi)h(x)G'(x)}{G^2(x)} \ , \text{ όπου}$$

$G(x) = mh(x) + (x-\xi)h'(x)$. Άρα,

$$g'(\xi) = 1 - \frac{mh^2(\xi)}{m^2h^2(\xi)} = 1 - \frac{1}{m} \ , \text{ και προφανώς } 0 < g'(\xi) = 1 - \frac{1}{m} < 1.$$

Έτσι, η μέθοδος των Newton – Raphson, σύμφωνα με το θεώρημα 6.4, είναι πρώτης τάξεως για τον προσδιορισμό της ρίζας ξ , για κατάλληλο x_0 . ▲

<u>Παρατήρηση</u> Στην περίπτωση του θεωρήματος 6.9 για τα διαδοχικά σφάλματα ισχύει

$$\lim_{n\to\infty} \frac{\varepsilon_{n+1}}{\varepsilon_n} = 1 - \frac{1}{m} \ . \tag{6.40}$$

<u>Θεώρημα 6.10</u> Εάν υπάρχει ένα διάστημα $[a,b]$ τέτοιο ώστε: i) $f(a)f(b) < 0$ και ii) οι $f'(x)$, $f''(x)$ είναι συνεχείς και δεν μηδενίζονται στο $[a,b]$, τότε $\forall\, x_0 \in [a,b]$ τέτοιο ώστε $f(x_0)f''(x_0) \geqq 0$, η μέθοδος Newton – Raphson συγκλίνει στην μοναδική ρίζα ξ της $f(x) = 0$ που υπάρχει στο $[a,b]$.

<u>Απόδειξη</u> Χωρίς βλάβη της γενικότητος υποθέτουμε ότι $f(a) < 0$,

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$f(b) > 0$, $f'(x) > 0$ και $f''(x) > 0$ $\forall\, x \in [a,b]$ (οι άλλες περιπτώσεις εξετάζονται ανάλογα). Σύμφωνα με το θεώρημα 6.1 στο $[a,b]$ υπάρχει μία μοναδική ρίζα ξ της $f(x) = 0$. Επειδή $f(x_0)f''(x_0) \geqq 0$, $x_0 \in [a,b]$ και $f''(x_0) > 0$ έπεται ότι $f(x_0) \geqq 0$. Αν εξαιρέσουμε την περίπτωση $f(x_0) = 0$, η οποία συνεπάγεται προφανώς ότι $\xi = x_0 = x_1 = \ldots$, θα είναι $f(x_0) > 0$, οπότε $x_0 > \xi$. Θα δειχθεί τώρα ότι $x_n > \xi$, $n = 1,2,\ldots$ και $x_{n+1} < x_n$, $n = 0,1,2,\ldots$. Επειδή $f'(x_0) > 0$ έχουμε

$$x_1 = g(x_0) = x_0 - \frac{f(x_0)}{f'(x_0)} < x_0 \quad , \tag{6.41}$$

όπου η $g(x)$ δίνεται από την (6.30). Από το θεώρημα μέσης τιμής παίρνουμε

$$x_1 - \xi = g(x_0) - g(\xi) = g'(\lambda)(x_0 - \xi) \quad , \tag{6.42}$$

όπου λ είναι κάποια τιμή μεταξύ των ξ και x_0 . Από την (6.33), επειδή $f''(x) > 0$, $\forall\, x \in [a,b]$ και $f(\lambda) > 0$, προκύπτει ότι $g'(\lambda) > 0$, οπότε από την (6.42) είναι $x_1 - \xi > 0$. Έτσι, λόγω και της (6.41), είναι $\xi < x_1 < x_0$. Επαναλαμβάνοντας την παραπάνω διαδικασία, σχηματίζεται μία φραγμένη, γνησίως φθίνουσα ακολουθία επαναλήψεων $x_0, x_1, \ldots, x_n, \ldots$ και , έστω $\xi^* = \lim\limits_{n \to \infty} x_n$. Παίρνοντας όρια στον αλγόριθμο Newton – Raphson (6.31) έχουμε

$$\xi^* = \xi^* - \frac{f(\xi^*)}{f'(\xi^*)} \quad \text{ή} \quad f(\xi^*) = 0 \ , \quad \text{δηλαδή,} \quad \xi = \xi^* \ . \qquad \blacktriangle$$

<u>Παρατηρήσεις</u> i) Η εκλογή του x_0 στο θεώρημα 6.10 μπορεί να γίνει ως εξής: Παίρνουμε σαν x_0 εκείνο από τα άκρα a,b του $[a,b]$ στο οποίο η $f(x)$ έχει το ίδιο πρόσημο με την $f''(x)$ (Η $f''(x)$ διατηρεί πρόσημο στο $[a,b]$) .

ii) Η σύγκλιση της μεθόδου Newton – Raphson στο θεώρημα 6.10

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

καλείται μονότονος και είναι προφανώς τετραγωνική.

<u>Πόρισμα 6.1</u> Εάν: i) η $f(x)$ είναι ορισμένη και συνεχής στο $(-\infty, +\infty)$, ii) υπάρχει ένα διάστημα $[a,b]$ τέτοιο ώστε $f(a)f(b) < 0$, iii) η $f'(x)$ είναι συνεχής και δεν μηδενίζεται στο $[a,b]$ και iv) η $f''(x)$ είναι συνεχής και δεν μηδενίζεται στο $(-\infty, +\infty)$, τότε, $\forall\, x_0 \in [a,b]$ η μέθοδος Newton – Raphson συγκλίνει στην μοναδική ρίζα ξ της $f(x) = 0$ που υπάρχει στο $[a,b]$.

<u>Απόδειξη</u> Υποθέτουμε πάλι ότι $f(a) < 0$, $f(b) > 0$, $f'(x) > 0$, $\forall\, x \in [a,b]$, $f''(x) > 0$, $\forall\, x \in (-\infty, +\infty)$ και θεωρούμε κάποιο $x_0 \in [a,b]$. Εάν $f(x_0) = 0$ τότε $\xi = x_0$. Αν $f(x_0) > 0$, δηλαδή, $\xi < x_0 \leqq b$, τότε $f(x_0)f''(x_0) > 0$ και η απόδειξη προκύπτει από το θεώρημα 6.10. Έστω τώρα $f(x_0) < 0$, δηλαδή, $a \leqq x_0 < \xi$. Τότε, είναι

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} > x_0 \quad .$$

Από το θεώρημα μέσης τιμής έχουμε $f(x_0) = f(x_0) - f(\xi) = (x_0 - \xi)f'(\lambda)$ όπου $\lambda \in (x_0, \xi)$. Επειδή $f''(x) > 0$, συνεπάγεται ότι η $f'(x)$ είναι γνησίως αύξουσα στο $[a,b]$. Άρα,

$$f(x_0) = (x_0 - \xi)f'(\lambda) < (x_0 - \xi)f'(x_0) \quad \text{ή} \quad \frac{f(x_0)}{f'(x_0)} < x_0 - \xi \quad \text{ή}$$

$$-\frac{f(x_0)}{f'(x_0)} > \xi - x_0 \quad .$$

Επομένως, έχουμε $x_1 = x_0 - \dfrac{f(x_0)}{f'(x_0)} > x_0 + \xi - x_0 = \xi$ ή $\xi < x_1$. Επειδή $f''(x) > 0$, έχουμε ότι $f'(x) > f'(a) > 0$, $\forall\, x > a$, οπότε στο διάστημα $[a, +\infty)$ υπάρχει μόνον η ρίζα ξ. Έτσι, έχουμε $f(x_1) > 0$ και $f(x_1)f''(x_1) > 0$, οπότε η απόδειξη προκύπτει εφαρμό-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ζοντας το θεώρημα 6.10 στο διάστημα $[a, +\infty)$ με $x_0 = x_1$. ▲

Το ακόλουθο θεώρημα, που δίνεται χωρίς απόδειξη, έχει σχέση με το θεώρημα 6.10 και παρέχει επίσης ικανές συνθήκες που εξασφαλίζουν την μονότονη σύγκλιση της μεθόδου Newton – Raphson για κάθε x_0 σε ένα διάστημα $[a,b]$. Η απόδειξη αυτού μπορεί να γίνει και γεωμετρικά.

Θεώρημα 6.11 Εάν $[a,b]$ είναι ένα διάστημα τέτοιο ώστε: i) $f(a)f(b) < 0$, ii) $f'(x) \neq 0$ στο $[a,b]$, iii) $f''(x) \geq 0$, $\forall\, x \in [a,b]$ ή $f''(x) \leq 0$, $\forall\, x \in [a,b]$ και iv) Η εφαπτομένη της $y = f(x)$ σε εκείνο από τα σημεία $(a,f(a))$, $(b,f(b))$, για το οποίο η $|f'(x)|$ είναι μικρότερη, τέμνει τον άξονα των x εντός του $[a,b]$, τότε η $f(x) = 0$ έχει μία μοναδική ρίζα ξ στο $[a,b]$ και η μέθοδος Newton – Raphson συγκλίνει στην ρίζα ξ , $\forall\, x_0 \in [a,b]$. (Η συνθήκη (iv) ικανοποιείται εάν $\left|\dfrac{f(c)}{f'(c)}\right| \leq b - a$, όπου $c = a$ αν $|f'(a)| < |f'(b)|$ και $c = b$ αν $|f'(b)| < |f'(a)|$) .

6.5.3. Εφαρμογές – Παραδείγματα

Η μέθοδος Newton – Raphson παρουσιάζει μεγάλο ενδιαφέρον στην πράξη επειδή είναι γενικά τουλάχιστον δευτέρας τάξεως, έχει απλή μορφή και είναι κατάλληλη για να χρησιμοποιηθεί σε ένα ηλεκτρονικό υπολογιστή. Αυτή είναι κατάλληλη για τον υπολογισμό της $N^{1/p} = \sqrt[p]{N}$, όπου N πραγματικός > 0 και p ακέραιος. Το πρόβλημα αυτό είναι ισοδύναμο με το να βρούμε την θετική ρίζα της εξισώσεως

$$f(x) = x^p - N = 0 \ . \tag{6.43}$$

Ο αλγόριθμος (6.31) για την (6.43) γίνεται

$$x_{n+1} = \frac{p-1}{p}\, x_n + \frac{N}{p x_n^{p-1}} \ , \quad n = 0,1,2,\ldots \ . \tag{6.44}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Αν $p > 1$, τότε μπορεί να αποδειχθεί (βλέπε άσκηση 20) ότι $\forall\, x_0 > 0$ ο αλγόριθμος (6.44) συγκλίνει στην θετική ρίζα της (6.43). Σαν αρχική προσέγγιση x_0 της $\sqrt[p]{N}$, στην (6.44), λαμβάνεται, συνήθως, η τιμή ενός πολυωνύμου που προσεγγίζει την συνάρτηση $y = \sqrt[p]{x}$.

Για $p = 2$ από την (6.44) έχουμε

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{N}{x_n}\right) , \quad n = 0,1,2,\ldots, \quad \text{και} \quad x_0 > 0 . \tag{6.45}$$

Ο αλγόριθμος (6.45), γνωστός σαν αλγόριθμος του Ήρωνος, ήταν γνωστός στους αρχαίους Έλληνες (100 π.Χ.) και χρησιμοποιείται σήμερα για την εύρεση της τετραγωνικής ρίζας ενός θετικού αριθμού N από πολλούς ηλεκτρονικούς υπολογιστές.

Για $p = -1$ ο (6.44) γίνεται

$$x_{n+1} = x_n(2 - Nx_n) , \quad n = 0,1,2,\ldots . \tag{6.46}$$

Ο αλγόριθμος (6.46) (με κατάλληλο x_0) υπολογίζει τον αντίστροφο του N χωρίς να υπεισέρχεται η πράξη της διαιρέσεως (βλέπε και άσκηση 21).

<u>Παράδειγμα 6.7</u> Να υπολογισθούν οι τρείς πραγματικές ρίζες της εξισώσεως $f(x) = x^3 - 10x + 4 = 0$ με ακρίβεια 3 δ.θ., χρησιμοποιώντας την μέθοδο Newton – Raphson.

<u>Λύση</u> Η δοθείσα εξίσωση έχει τρείς πραγματικές ρίζες ξ_1, ξ_2, ξ_3 τέτοιες ώστε $\xi_1 \in [-4, -3]$, $\xi_2 \in [0,1]$ και $\xi_3 \in [2,3]$. Ο αλγόριθμος της μεθόδου Newton – Raphson είναι

$$x_{n+1} = x_n - \frac{x_n^3 - 10x_n + 4}{3x_n^2 - 10} \quad \text{ή} \quad x_{n+1} = \frac{2x_n^3 - 4}{3x_n^2 - 10} , \quad n = 0,1,2,\ldots .$$

Για την ρίζα ξ_1 λαμβάνουμε $x_0 = -3$ και διατηρώντας 4.δ.θ. στους υπολογισμούς βρίσκουμε διαδοχικά $x_1 \approx -3.4118$, $x_2 \approx -3.3477$, $x_3 \approx -3.3460$, $x_4 \approx -3.3460$. Επειδή $|x_4 - x_3| \leq \frac{1}{2}\, 10^{-3}$, έχουμε $\xi_1 \approx -3.346$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Για την ρίζα ξ_2 παίρνουμε $x_0 = 0$, οπότε βρίσκουμε $x_1 = 0.4$, $x_2 \approx 0.4067$, $x_3 \approx 0.4067$. Άρα $\xi_2 \approx 0.407$.

Για την ρίζα ξ_3 , χρησιμοποιώντας $x_0 = 3$, βρίσκουμε $x_1 \approx$ ≈ 2.9412 , $x_2 \approx 2.9392$, $x_3 \approx 2.9392$. Άρα $\xi_3 \approx 2.939$. (Πρέπει να σημειωθεί ότι η εκλογή του x_0 έγινε έχοντας υπ' όψιν τα θεωρήματα 6.10 και 6.11). ▲

6.5.4 Η μέθοδος Newton–Raphson για την εύρεση μιγαδικών ριζών

Πολλές φορές,στην πράξη,είναι απαραίτητο να βρούμε καλές προσεγγίσεις για τις μιγαδικές ρίζες μιας εξισώσεως

$$f(z) = 0 \quad , \tag{6.47}$$

όπου $f(z)$ είναι μία πραγματική ή μιγαδική συνάρτηση, κυρίως,πολυωνυμική.

Γενικά, έστω $f(z)$ μία μιγαδική συνάρτηση μιας μιγαδικής μεταβλητής z , όπου $z = x + iy$ συμβολίζει ένα σημείο του πεδίου ορισμού της (x,y είναι πραγματικά και $i = \sqrt{-1}$.) και $f(z) = u + iv$ ένα σημείο του πεδίου τιμών της , όπου $u = u(x,y)$ και $v = v(x,y)$ είναι πραγματικές συναρτήσεις.

Η εύρεση μιας ρίζας της (6.47) ισοδυναμεί με το να βρούμε τέτοια x και y που μηδενίζουν συγχρόνως τις u και v . Έτσι, το πρόβλημα της ευρέσεως των μιγαδικών ριζών της (6.47) ισοδυναμεί με την εύρεση των πραγματικών ριζών του, γενικά, μη γραμμικού συστήματος, των δύο πραγματικών εξισώσεων

$$\begin{cases} u(x,y) = 0 \\ v(x,y) = 0 \quad . \end{cases} \tag{6.48}$$

Τό πρόβλημα (6.48) θα μας απασχολήσει ειδικώτερα στο κεφάλαιο 8 . Προς το παρόν, θα δούμε με ένα παράδειγμα πώς η μέθοδος Newton – Raphson μπορεί να χρησιμοποιηθεί για την εύρεση των μιγαδικών ριζών

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

της (6.47). Ας θεωρήσουμε την εξίσωση

$$f(x) = x^2 + 1 = 0 \ , \tag{6.49}$$

της οποίας οι ρίζες είναι $\pm i$. Η μέθοδος Newton – Raphson είναι

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \ \text{ή} \ x_{n+1} = x_n - \frac{x_n^2 + 1}{2x_n} = \frac{x_n^2 - 1}{2x_n} \ , \tag{6.50}$$

$$n = 0,1,2,\ldots \ .$$

Όπως και για την εύρεση πραγματικών ριζών, είναι απαραίτητο και εδώ η αρχική προσέγγιση x_0 να εκλεγεί, σε μία περιοχή του μιγαδικού επιπέδου, πολύ κοντά στην ζητούμενη μιγαδική ρίζα ξ . Επί πλέον, αφού η $f(x) = x^2 + 1$ παίρνει πραγματικές τιμές για x πραγματικό, με σκοπό να βρούμε μιγαδική ρίζα πρέπει το x_0 να είναι μιγαδικό, διότι, διαφορετικά, όλες οι επαναλήψεις θα είναι πραγματικές. Αν πάρουμε π.χ. $x_0 = 0.5\,i$, τότε από την (6.50) βρίσκουμε $x_1 = 1.25i$, $x_2 = 1.025i$, $x_3 \approx 1.0003i$ κ.ο.κ. . Προφανώς, η ακολουθία που παράγεται συγκλίνει στην ρίζα $\xi = i$.

Το μειονέκτημα στην περίπτωση των μιγαδικών ριζών είναι ότι χρειάζονται πολύ περισσότεροι υπολογισμοί σε κάθε επανάληψη, λόγω των περιεχομένων πράξεων μεταξύ μιγαδικών αριθμών.

6.6 Μέθοδοι παρεμβολής για τον προσδιορισμό ριζών

Πολύ χρήσιμες για τον προσδιορισμό των ριζών της εξισώσεως $f(x) = 0$ είναι και οι μέθοδοι παρεμβολής, που εξετάζονται στη συνέχεια. Αυτές, σε σύγκριση με την μέθοδο Newton – Raphson, έχουν το πλεονέκτημα ότι δεν απαιτούν τον υπολογισμό της $f'(x)$, ακόμη δε, σε πολλές περιπτώσεις συγκλίνουν ταχύτερα.

6.6.1 Η μέθοδος της εσφαλμένης θέσεως (Regula falsi)

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Η απλούστερη και παλαιότερη από τις μεθόδους παρεμβολής είναι γνωστή σαν μέθοδος της εσφαλμένης θέσεως. Όπως και στην μέθοδο της διχοτομήσεως, υποθέτουμε και εδώ ότι η $f(x)$ είναι συνεχής σε ένα διάστημα $[a_0,b_0]$ και $f(a_0)f(b_0) < 0$. Επομένως, το διάστημα $[a_0,b_0]$ περιέχει τουλάχιστον μία ρίζα της $f(x) = 0$. Έστω x_0 η ρίζα της γραμμικής συναρτήσεως παρεμβολής $P_0(x)$ με $P_0(a_0) = f(a_0)$ και $P_0(b_0) = f(b_0)$. Τότε, έχουμε (βλέπε (2.15))

$$P_0(x) = f(a_0) + (x - a_0)\left[\frac{f(b_0) - f(a_0)}{b_0 - a_0}\right] \tag{6.51}$$

και

$$x_0 = a_0 - \frac{(b_0 - a_0)f(a_0)}{f(b_0) - f(a_0)} \quad . \tag{6.52}$$

Αφού η $f(a_0)f(b_0) < 0$ συνεπάγεται ότι $f(a_0) \neq f(b_0)$, το x_0 είναι πάντοτε ορισμένο και είναι εύκολο να δειχθεί ότι $a_0 < x_0 < b_0$. Το x_0 , που δίνεται από την (6.52), αποτελεί την αρχική προσέγγιση για μια ρίζα της $f(x) = 0$ στο $[a_0,b_0]$. Αν $f(x_0) = 0$, τότε η $\xi = x_0$ είναι μία ρίζα της $f(x) = 0$. Διαφορετικά, καλούμε $[a_1,b_1]$ εκείνο από τα δύο διαστήματα $[a_0,x_0]$, $[x_0,b_0]$ στα άκρα του οποίου η συνάρτηση $f(x)$ έχει αντίθετα πρόσημα και στη συνέχεια από το $[a_1,b_1]$ βρίσκεται η προσέγγιση x_1 με την ίδια διαδικασία, με την οποία βρέθηκε και η x_0 από το $[a_0,b_0]$. Κατόπιν, βρίσκεται κατά τον ίδιο τρόπο το διάστημα $[a_2,b_2]$ και η x_2 κ.ο.κ.. Γενικά, η προσέγγιση x_n δίνεται από τον τύπο

$$x_n = a_n - (b_n - a_n)\frac{f(a_n)}{f(b_n) - f(a_n)} , \quad n = 0,1,2,\ldots , \tag{6.53}$$

όπου $f(a_n)f(b_n) < 0$. Μπορεί να αποδειχθεί ότι η ακολουθία που παράγεται από τον αλγόριθμο (6.53) συγκλίνει πάντοτε σε μία ρίζα της

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$f(x) = 0$ που βρίσκεται στο $[a_0, b_0]$. Στη συνέχεια, χάριν απλότητος, θα αποδειχθεί η σύγκλιση της ακολουθίας $\{x_n\}$, $n = 0,1,2,\ldots$, υποθέτοντας ότι για κάποιο $n \geqq 0$ ισχύουν:

$$\begin{aligned} &\text{i)}\quad f(a_n) < 0 \ , \quad f(b_n) > 0 \quad \text{και} \\ &\text{ii)}\quad \text{ή } f''(x) \text{ υπάρχει και } f''(x) \geqq 0 \ , \quad \forall\, x \in [a_n, b_n] \ . \end{aligned} \tag{6.54}$$

Κάτω από τις προυποθέσεις (6.54), θα δειχθεί ότι είτε $f(x_n) = 0$ ή $f(x_n)f(b_n) < 0$, και επομένως $a_n < a_{n+1} = x_n < b_{n+1} = b_n$.

Από τον ορισμό (6.53) του x_n έχουμε αμέσως ότι $a_n < x_n < b_n$. Το σφάλμα κατά την παρεμβολή, για την εύρεση του x_n, δίνεται από τον τύπο

$$P_n(x) - f(x) = -(x - a_n)(x - b_n)\,\frac{f''(\lambda)}{2}\ , \tag{6.55}$$

όπου $P_n(x)$ είναι το πρωτοβάθμιο πολυώνυμο παρεμβολής με σημεία παρεμβολής τα a_n, b_n, $x \in [a_n, b_n]$ και λ μία κατάλληλη τιμή στο $[a_n, b_n]$. Λόγω της (ii) των (6.54), έχουμε $f(x) - P_n(x) \leqq 0$, $\forall\, x \in [a_n, b_n]$ και επομένως

$$f(x_n) - P_n(x_n) = f(x_n) \leqq 0\ , \tag{6.56}$$

αφού $P_n(x_n) = 0$. Έστω $f(x_n) \neq 0$. Τότε, προφανώς, είναι $f(x_n)f(b_n) < 0$. Στο νέο διάστημα τώρα $[a_{n+1}, b_{n+1}] \equiv [x_n, b_n]$ ισχύουν $f(a_{n+1}) < 0$, $f(b_{n+1}) > 0$ και $f''(x) \geqq 0$, $\forall\, x \in [a_{n+1}, b_{n+1}]$, δηλαδή, αντίστοιχες των (6.54). Έτσι, είναι φανερό ότι αν οι (6.54) ισχύουν για $n = n_0$, τότε θα ισχύουν $\forall\, n \geqq n_0$. Επομένως, για $n \geqq n_0$ θα έχουμε $b_n = b$ και οι τιμές $x_n = a_{n+1}$ σχηματίζουν μία γνησίως αύξουσα ακολουθία, φραγμένη από το b. Έστω $\xi = \lim\limits_{n\to\infty} x_n$. Αφού η $f(x)$ είναι συνεχής, λόγω της (6.56), έπεται $f(\xi) \leqq 0$ και από την (6.53) παίρνουμε

$$\xi = \xi - (b - \xi)\,\frac{f(\xi)}{f(b) - f(\xi)} \quad \text{ή} \quad (\xi - b)f(\xi) = 0\ .$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Αλλά $\xi \neq b$, αφού $f(b) > 0$ ενώ $f(\xi) \leqq 0$. Άρα, $f(\xi) = 0$ και επομένως το όριο ξ της ακολουθίας είναι μία ρίζα της $f(x) = 0$. Μπορεί να αποδειχθεί (βλέπε άσκηση 26) ότι κάτω από τις προυποθέσεις (6.54), υπάρχει μία μοναδική ρίζα της $f(x) = 0$ στο $[a_n, b_n]$.

Αν οι σχέσεις (6.54) ισχύουν για $n = 0$, δηλαδή, έχουμε

$$f(a_0) < 0 \ , \ f(b_0) > 0 \ , \ f''(x) \geqq 0 \ , \ \forall \ x \in [a_0, b_0] \ , \qquad (6.57)$$

τότε το άκρο $b_0 = b$ του διαστήματος $[a_0, b_0]$ παραμένει σταθερό και ο αλγόριθμος (6.53) μετασχηματίζεται στον

$$x_n = x_{n-1} - (b - x_{n-1}) \frac{f(x_{n-1})}{f(b) - f(x_{n-1})} \ , \ n = 0,1,2,\ldots \ , \qquad (6.58)$$

όπου $x_{-1} = a_0$.

Εάν αντί των (6.57) ισχύουν οι σχέσεις

$$f(a_0) > 0 \ , \ f(b_0) < 0 \ , \ f''(x) \geqq 0 \ , \ \forall \ x \in [a_0, b_0] \ , \qquad (6.59)$$

τότε μπορεί εύκολα να δειχθεί (βλέπε άσκηση 27) ότι το άκρο $a_0 = a$ του διαστήματος $[a_0, b_0]$ παραμένει σταθερό και ο αλγόριθμος (6.53) γίνεται

$$x_n = a - (x_{n-1} - a) \frac{f(a)}{f(x_{n-1}) - f(a)} \quad \text{ή}$$

$$x_n = x_{n-1} - (x_{n-1} - a) \frac{f(x_{n-1})}{f(x_{n-1}) - f(a)} \ , \ n = 0,1,2,\ldots \ , \qquad (6.60)$$

όπου $x_{-1} = b_0$. Σαν συμπέρασμα, από τις (6.57) και (6.59) προκύπτει ότι το σταθερό άκρο στο $[a_0, b_0]$ είναι εκείνο στο οποίο το πρόσημο της $f(x)$ συμπίπτει με το πρόσημο της $f''(x)$. Στην περίπτωση που αντί της $f''(x) \geqq 0$, $\forall \ x \in [a_0, b_0]$ ισχύει η $f''(x) \leqq 0$, $\forall \ x \in [a_0, b_0]$, τότε θεωρούμε την εξίσωση $-f(x) = 0$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

6.6.2 Τάξη συγκλίσεως της μεθόδου Regula falsi

Όπως τονίστηκε στην παράγραφο 6.3, η μέθοδος της διχοτομήσεως έχει αργή σύγκλιση (συγκεκριμένα, μπορεί να δειχθεί ότι η σύγκλιση είναι γραμμική). Αυτό εξηγείται κατά κάποιο τρόπο από το γεγονός ότι χρησιμοποιεί πολύ λίγο την συνάρτηση $f(x)$. Πράγματι, μόνον το πρόσημο της $f(x)$ σε διάφορα σημεία απαιτείται. Αντιθέτως, η μέθοδος Regula falsi χρησιμοποιεί και τιμές της $f(x)$ για την εύρεση των διαδοχικών προσεγγίσεων. Για τον λόγο αυτό, γενικά, αναμένουμε η μέθοδος Regula falsi να συγκλίνει ταχύτερα από την μέθοδο της διχοτομήσεως.

Υποθέτουμε ότι ισχύουν οι προυποθέσεις (6.57),οπότε ο αλγόριθμος της μεθόδου Regula falsi δίνεται από την (6.58), η οποία μπορεί να γραφεί

$$x_{n+1} = x_n - \frac{(b - x_n)f(x_n)}{f(b) - f(x_n)}, \quad n = -1,0,1,\ldots, \tag{6.61}$$

όπου $x_{-1} = a_0$ και $b = b_0$.

Ο αλγόριθμος (6.61) έχει την μορφή $x_{n+1} = g(x_n)$, όπου η επαναληπτική συνάρτηση $g(x)$ δίνεται από την

$$g(x) = x - \frac{(b - x)f(x)}{f(b) - f(x)} . \tag{6.62}$$

Αν ξ είναι η ρίζα της $f(x) = 0$, στην οποία συγκλίνει ο (6.61), τότε λόγω της (6.62) έχουμε

$$g'(x) = 1 - \frac{[-f(x) + (b - x)f'(x)][f(b) - f(x)] + (b - x)f(x)f'(x)}{[f(b) - f(x)]^2},$$

οπότε

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$g'(\xi) = 1 - \frac{(b-\xi)f'(\xi)f(b)}{f^2(b)} = 1 - f'(\xi)\,\frac{b-\xi}{f(b)} =$$

$$= 1 - f'(\xi)\,\frac{\xi - b}{f(\xi) - f(b)}\,. \tag{6.63}$$

Από το θεώρημα μέσης τιμής έχουμε

$$f(\xi) - f(b) = (\xi - b)f'(\lambda_1)\ ,\ \text{όπου}\ \ \xi < \lambda_1 < b \quad \text{και} \tag{6.64}$$

$$f(x_n) - f(\xi) = (x_n - \xi)f'(\lambda_2) \quad \text{με}\ \ x_n < \lambda_2 < \xi\ . \tag{6.65}$$

Από την (6.56) έχουμε $f(x_n) < 0$ και επομένως $x_n < \xi$ (εκτός εάν $x_n = \xi$) . Άρα, από την (6.65) είναι

$$\frac{f(x_n) - f(\xi)}{x_n - \xi} = f'(\lambda_2) > 0\ .$$

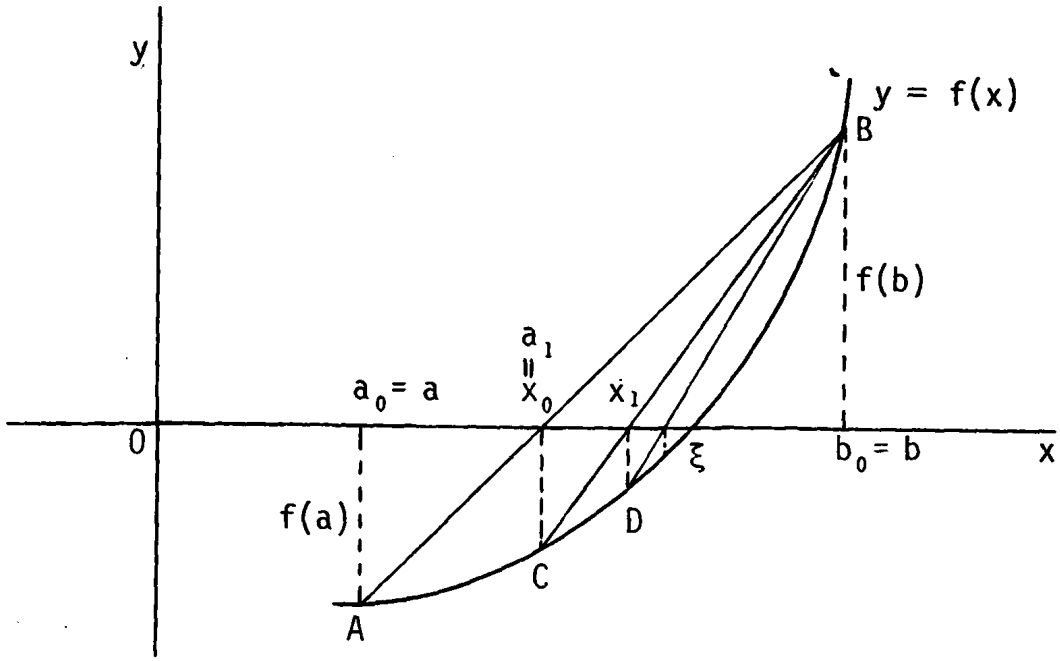
Επειδή $f''(x) \geqq 0$, $\forall\, x \in [a_0, b_0]$, η $f'(x)$ θα είναι αύξουσα στο $[a_0, b_0]$ και επομένως $0 < f'(\lambda_2) \leqq f'(\xi) \leqq f'(\lambda_1)$, οπότε από τις (6.63) και (6.64) παίρνουμε

$$0 \leqq g'(\xi) = 1 - \frac{f'(\xi)}{f'(\lambda_1)} < 1\ . \tag{6.66}$$

Η (6.66) σημαίνει ότι,κάτω από τις συνθήκες (6.57),η σύγκλιση της μεθόδου Regula falsi είναι τουλάχιστον γραμμική. Γενικά, η μέθοδος Regula falsi έχει τα ίδια μειονεκτήματα με την μέθοδο της διχοτομήσεως, πράγμα που σημαίνει ότι πολλές φορές είναι ασύμφορη η χρησιμοποίησή της στην πράξη. Σημειώνεται τέλος, ότι οι μέθοδοι διχοτομήσεως και Regula falsi καλούνται και μέθοδοι διαστημάτων.

6.6.3 Γεωμετρική ερμηνεία της μεθόδου Regula falsi

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ



σχ. 6.7

Η γεωμετρική ερμηνεία της μεθόδου Regula falsi, στην περίπτωση που αυτή δίνεται από τον αλγόριθμο (6.58) φαίνεται στο σχήμα 6.7. Στο διάστημα $[a_0,b_0]$ η καμπύλη $y = f(x)$ "αντικαθίσταται" από την ευθεία AB με $A(a_0,f(a_0))$ και $B(b_0,f(b_0))$, και η τομή x_0 αυτής με τον άξονα των x ,μας δίνει την "εσφαλμένη θέση" της ρίζας ξ (μέθοδος της εσφαλμένης θέσεως). Μετά βρίσκουμε τα $x_1,x_2,\ldots$ επαναλαμβάνοντας την διαδικασία. Όπως φαίνεται στο σχήμα, το άκρο $b_0 = b$ παραμένει σταθερό και οι διαδοχικές προσεγγίσεις συγκλίνουν στην ρίζα ξ.

<u>Παράδειγμα 6.8</u> Δίνεται η εξίσωση $f(x) = x^2 - N = 0$, όπου N πραγματικός > 1 και θεωρούμε το διάστημα $[a_0,b_0] \equiv [0,N]$, για το οποίο ισχύει $f(0)f(N) < 0$. Να βρεθεί ο αλγόριθμος της μεθόδου Regula falsi και το όριο της ακολουθίας που παράγεται.

<u>Λύση</u> Έχουμε $f(0) < 0$, $f(N) > 0$ και $f''(x) = 2 > 0$, $\forall\, x \in [0,N]$, οπότε ικανοποιούνται οι (6.57). Έτσι, από την (6.58) παίρνουμε

$$x_n = x_{n-1} - (N - x_{n-1})\,\frac{f(x_{n-1})}{f(N) - f(x_{n-1})} \quad \text{ή}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$x_n = x_{n-1} - (N - x_{n-1}) \frac{x_{n-1}^2 - N}{N(N-1) - x_{n-1}^2 + N} \quad \text{ή}$$

$$x_n = \frac{N(x_{n-1} + 1)}{N + x_{n-1}} \quad , \; n = 0,1,2,\ldots \quad \text{και} \quad x_{-1} = a_0 = 0 \; .$$

Προφανώς, η ακολουθία $\{x_n\}$, $n = 0,1,2,\ldots$ θα συγκλίνει στην μοναδική ρίζα $\xi = \sqrt{N}$ της $f(x) = 0$, που υπάρχει στο $[0,N]$ (βλέπε και παράδειγμα 6.4). ▲

6.6.4 Η μέθοδος της τέμνουσας

Μία σημαντική παραλλαγή της μεθόδου Regula falsi είναι η καλούμενη μέθοδος της τέμνουσας ή μέθοδος της χορδής, η οποία όμως δεν έχει το πλεονέκτημα της εγγυημένης συγκλίσεως. Στην μέθοδο αυτή αρχίζουμε με ένα αρχικό διάστημα $[a,b]$, στο οποίο υποθέτουμε ότι η συνάρτηση $f(x)$ είναι συνεχής και στο οποίο συνήθως ισχύει $f(a)f(b) < 0$ (η συνθήκη $f(a)f(b) < 0$ δεν είναι απαραίτητη) . Καλούμε $x_0 = a$ (ή $x_0 = b$) και $x_1 = b$ (ή $x_1 = a$) . Σαν μια νέα προσέγγιση x_2 , παίρνουμε την τετμημένη του σημείου τομής με τον άξονα των x , της ευθείας που διέρχεται από τα σημεία $(x_0, f(x_0))$ και $(x_1, f(x_1))$ (βλέπε σχήμα 6.8), δηλαδή, x_2 είναι η ρίζα του πολυωνύμου παρεμβολής πρώτου βαθμού, με σημεία παρεμβολής x_0, x_1 , που προσεγγίζει την $f(x)$. Η προσέγγιση τώρα x_3 βρίσκεται από τα x_1, x_2 κατά τον ίδιο τρόπο που βρέθηκε η x_2 από τα x_0 και x_1 , κ.ο.κ.. Γενικά, αν x_{n-1} και x_n είναι δύο διαδοχικές προσεγγίσεις, τότε η x_{n+1} δίνεται από τον αλγόριθμο (σύμφωνα με την (6.52))

$$x_{n+1} = x_n - \frac{(x_n - x_{n-1}) f(x_n)}{f(x_n) - f(x_{n-1})} \quad , \; n = 1,2,\ldots \; , \tag{6.67}$$

όπου $f(x_{n-1}) \neq f(x_n)$. Εάν σε κάποια φάση της διαδικασίας συμβεί

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$f(x_{n-1}) = f(x_n)$, τότε η μέθοδος δεν είναι αποτελεσματική. Στην μέθοδο της τέμνουσας δεν είμαστε βέβαιοι ότι ισχύει $f(x_{n-1})f(x_n) < 0$ για κάθε n , ακόμη και αν $f(x_0)f(x_1) < 0$. Ο αλγόριθμος (6.67) είναι απλούστερος (αφού το πρόσημο της $f(x_{n+1})$ δεν ελέγχεται) από τον αντίστοιχο της μεθόδου Regula falsi (βλέπε (6.53)) και συχνά συγκλίνει ταχύτερα.

Ένας άλλος τρόπος παραγωγής της μεθόδου της τέμνουσας είναι ο ακόλουθος:

Εάν στην μέθοδο Newton – Raphson (6.31) αντικαταστήσουμε την παράγωγο $f'(x_n)$ (πολλές φορές ο υπολογισμός της παραγώγου $f'(x_n)$ είναι πολύ δαπανηρός) με την διηρημένη διαφορά

$$f[x_{n-1},x_n] = \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}} ,$$

παίρνουμε $x_{n+1} = x_n - \dfrac{f(x_n)}{\dfrac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}}$, δηλαδή, την (6.67) .

Μία σημαντική παρατήρηση για την μέθοδο της τέμνουσας είναι ότι αυτή δεν έχει την μορφή της επαναληπτικής μεθόδου (6.8) αλλά την μορφή

$$x_{n+1} = g(x_n,x_{n-1}) , \quad n = 1,2,\ldots , \tag{6.68}$$

δηλαδή, η επανάληψη x_{n+1} προσδιορίζεται και από τις δύο προηγούμενες επαναλήψεις x_n και x_{n-1} . Μία μέθοδος της μορφής (6.68) καλείται επαναληπτική μέθοδος δύο βημάτων σε αντίθεση με την (6.8) που είναι μέθοδος του ενός βήματος.

6.6.5 Σύγκλιση της μεθόδου της τέμνουσας

Υποθέτουμε ότι ξ είναι μία απλή ρίζα της $f(x) = 0$ και ότι

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

υπάρχουν οι $f'(x)$, $f''(x)$ και είναι συνεχείς σε μία περιοχή $\Pi(\xi)$ της ρίζας ξ .

Αν x_{n-1}, x_n είναι δύο διαδοχικές προσεγγίσεις της μεθόδου της τέμνουσας και $P_n(x)$ είναι το πολυώνυμο παρεμβολής, με σημεία παρεμβολής τα x_{n-1} και x_n , που προσεγγίζει την $f(x)$, τότε θα ισχύει

$$f(x) = P_n(x) + \frac{1}{2}(x - x_{n-1})(x - x_n)f''(\lambda_n) , \tag{6.69}$$

όπου λ_n ανήκει στο μικρότερο διάστημα που περιέχει τα x , x_{n-1} και x_n . Το $P_n(x)$ δίνεται από τον τύπο

$$P_n(x) = f(x_n) + (x - x_n)\frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}} \tag{6.70}$$

και επειδή $P_n(x_{n+1}) = 0$,από την (6.70) προκύπτει ότι

$$0 = f(x_n) + (x_{n+1} - x_n)\frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}} \quad \text{ή}$$

$$f(x_n) = (x_n - x_{n+1})\frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}} . \tag{6.71}$$

Η (6.70), λόγω της (6.71),γίνεται

$$P_n(x) = (x - x_{n+1})\frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}} . \tag{6.72}$$

Από την (6.69),για $x = \xi$ και λόγω της (6.72), έχουμε

$$0 = f(\xi) = (\xi - x_{n+1})\frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}} +$$

$$+\frac{1}{2}(\xi - x_{n-1})(\xi - x_n)f''(\lambda_n) , \tag{6.73}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

όπου λ_n ανήκει στο μικρότερο διάστημα που περιέχει τα x_{n-1}, x_n και ξ. Από το θεώρημα μέσης τιμής είναι

$$f(x_n) - f(x_{n-1}) = (x_n - x_{n-1})f'(\mu_n) \ , \tag{6.74}$$

όπου μ_n είναι κάποιο σημείο μεταξύ των x_n και x_{n-1}. Η (6.73) λόγω της (6.74) γίνεται

$$0 = (\xi - x_{n+1})f'(\mu_n) + \frac{1}{2}(\xi - x_{n-1})(\xi - x_n)f''(\lambda_n) \qquad \text{ή}$$

$$x_{n+1} - \xi = (x_{n-1} - \xi)(x_n - \xi)\frac{f''(\lambda_n)}{2f'(\mu_n)} \qquad \text{ή}$$

$$\varepsilon_{n+1} = M\varepsilon_n\varepsilon_{n-1} \ , \tag{6.75}$$

όπου $\varepsilon_n = x_n - \xi$ είναι γενικά το σφάλμα στην n επανάληψη και $M = \frac{1}{2}\frac{f''(\lambda_n)}{f'(\mu_n)}$.

Από την (6.75) προκύπτει το συμπέρασμα ότι η μέθοδος της τέμνουσας θα συγκλίνει εάν οι αρχικές προσεγγίσεις x_0, x_1 βρίσκονται επαρκώς κοντά στην ρίζα ξ, δηλαδή, σε μία κατάλληλη περιοχή της ρίζας ξ. Με άλλα λόγια, η μέθοδος της τέμνουσας είναι τοπικώς συγκλίνουσα στην ρίζα ξ.

Υποθέτουμε τώρα ότι η μέθοδος συγκλίνει και $f''(\xi) \neq 0$, οπότε από την (6.75) παίρνουμε

$$\lim_{n\to\infty} \frac{\varepsilon_{n+1}}{\varepsilon_n\varepsilon_{n-1}} = \frac{1}{2}\frac{f''(\xi)}{f'(\xi)} \equiv C \ , \tag{6.76}$$

αφού $\lambda_n \underset{n\to\infty}{\to} \xi$ και $\mu_n \underset{n\to\infty}{\to} \xi$. Δηλαδή, για πολύ μεγάλες τιμές του n θα έχουμε

$$\varepsilon_{n+1} \approx C\varepsilon_n\varepsilon_{n-1} \ . \tag{6.77}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Με σκοπό να βρούμε την τάξη της μεθόδου της τέμνουσας για τον προσδιορισμό της ρίζας ξ (εδώ πρέπει να σημειωθεί ότι χρησιμοποιείται ο ορισμός 6.1, όπου η (6.12) αντικαθίσταται από την (6.68)), από την (6.77), θέτοντας $|C\varepsilon_n| = R_n$, παίρνουμε

$$R_{n+1} \approx R_n R_{n-1} \ . \tag{6.78}$$

Από την (6.78) προκύπτει ότι $\ln R_{n+1} \approx \ln R_n + \ln R_{n-1}$. Είναι λογικό να υποθέσουμε ότι για μεγάλο n η ποσότης $\ln R_n$ συμπεριφέρεται όπως η λύση της εξισώσεως διαφορών

$$y_{n+1} = y_n + y_{n-1} \ . \tag{6.79}$$

Η λύση της (6.79) είναι $y_n = c_1 p_1^n + c_2 p_2^n$, όπου $p_1 = \frac{1+\sqrt{5}}{2}$, $p_2 = \frac{1-\sqrt{5}}{2}$ και c_1, c_2 αυθαίρετες σταθερές, οπότε για μεγάλο n $y_n \approx c_1 p_1^n$, αφού $|p_2| < 1$. Δηλαδή, έχουμε $\ln R_n \approx y_n \approx c_1 p_1^n$, οπότε $\ln R_{n+1} \approx c_1 p_1^{n+1}$, ή $\ln|C\varepsilon_{n+1}| \approx c_1 p_1^{n+1}$, ή $|C\varepsilon_{n+1}| \approx e^{c_1 p_1^{n+1}}$,

ή $|\varepsilon_{n+1}| \approx \frac{1}{|C|} e^{c_1 p_1^{n+1}} = |C|^{p_1 - 1} \left(\frac{1}{|C|} e^{c_1 p_1^n} \right)^{p_1} =$

$= |C|^{\frac{1}{p_1}} \left(\frac{1}{|C|} e^{c_1 p_1^n} \right)^{p_1} \approx |C|^{\frac{1}{p_1}} |\varepsilon_n|^{p_1}$. Έτσι, για μεγάλο n ισχύει

$$|\varepsilon_{n+1}| \approx |C|^{\frac{1}{p_1}} |\varepsilon_n|^{p_1} \ , \tag{6.80}$$

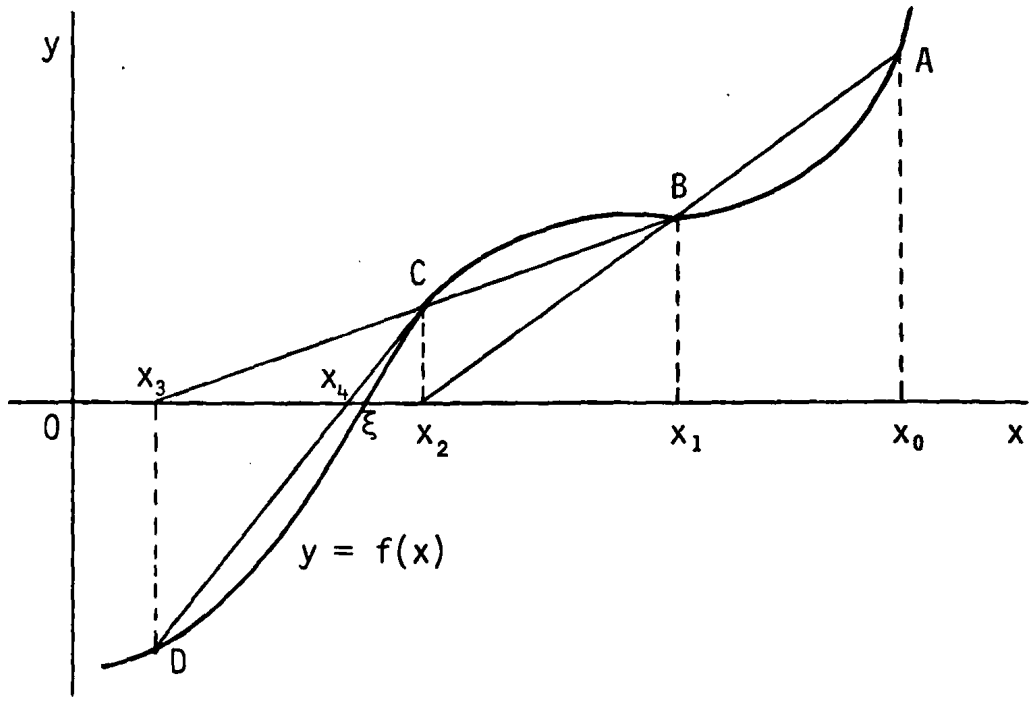
όπου $p_1 = \frac{1+\sqrt{5}}{2} \approx 1.618$ και η C δίνεται από την (6.76). Η (6.80) σημαίνει ότι, κάτω από τις προυποθέσεις που θεωρήσαμε, η μέθοδος της τέμνουσας είναι τάξεως περίπου $p_1 \approx 1.618$ για τον προσδιορισμό της ρίζας ξ. Στην περίπτωση αυτή η σύγκλιση καλείται υπεργραμμική $(1 < p_1 < 2)$.

Πρέπει να σημειωθεί ότι αν ξ είναι ρίζα πολλαπλότητος $m > 1$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

της $f(x) = 0$, τότε είναι δυνατόν η μέθοδος της τέμνουσας να μην συγκλίνει στη ρίζα ξ ακόμη κι' αν τα x_0, x_1 εκλεγούν πολύ κοντά στη ρίζα ξ . Εάν όμως αυτή συγκλίνει η σύγκλιση θα είναι γραμμική.

Κλείνοντας την παρούσα παράγραφο πρέπει να αναφερθεί ότι διάφορες τροποποιήσεις και συνδυασμοί των τριών μεθόδων: διχοτομήσεως, Regula falsi και τέμνουσας έχουν προταθεί κατά καιρούς με τα πλεονεκτήματα της εγγυημένης και ταχύτερης συγκλίσεως (π.χ. αλγόριθμος Dekker – Brent) . Ακόμη, αναφέρουμε την μέθοδο Muller, στην οποία μία παραβολή χρησιμοποιείται για να προσεγγίσει την $f(x)$, δοθέντων τριών διαφορετικών προσεγγίσεων x_{n-2}, x_{n-1} και x_n μιας ρίζας ξ της $f(x) = 0$ (η παραβολή διέρχεται από τα σημεία: $(x_{n-2}, f(x_{n-2}))$, $(x_{n-1}, f(x_{n-1}))$ και $(x_n, f(x_n)))$. Σαν νέα προσέγγιση x_{n+1} για την ρίζα ξ εκλέγεται η ρίζα της παραβολής που βρίσκεται πιο κοντά στην x_n .



σχ. 6.8

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Παράδειγμα 6.9 Δίνεται η εξίσωση $f(x) = x^2 - 2 = 0$ και θεωρούμε το διάστημα $[a,b] \equiv [1,2]$. Να δοθεί ο αλγόριθμος της μεθόδου της τέμνουσας και να βρεθούν οι επαναλήψεις x_2, x_3, x_4 και x_5 λαμβάνοντας $x_0 = 1$ και $x_1 = 2$.

Λύση Σύμφωνα με την (6.67) έχουμε

$$x_{n+1} = x_n - \frac{(x_n - x_{n-1})(x_n^2 - 2)}{(x_n^2 - 2) - (x_{n-1}^2 - 2)} \quad , \text{ ή } x_{n+1} = \frac{x_n x_{n-1} + 2}{x_n + x_{n-1}} \quad ,$$

$n = 1,2,\ldots,$ όπου $x_0 = 1$, $x_1 = 2$ και $x_n + x_{n-1} \neq 0$. Έτσι, παίρνουμε $x_2 = \frac{4}{3}$, $x_3 = 1.4$, $x_4 = \frac{58}{41}$ και $x_5 \approx 1.41421$. ▲

6.7 Επιτάχυνση της συγκλίσεως – Η Δ^2 – μέθοδος του Aitken – Η μέθοδος του Steffensen

Μία επαναληπτική μέθοδος πρώτης τάξεως για τον προσδιορισμό μιας ρίζας ξ της $f(x) = 0$ μπορεί πολλές φορές να συγκλίνει πάρα πολύ αργά, με αποτέλεσμα να μην έχει πρακτική αξία. Μπορούμε όμως, συχνά να επιταχύνουμε την σύγκλιση της ακολουθίας των επαναλήψεων $\{x_n\}$, που παράγεται από την θεωρούμενη μέθοδο πρώτης τάξεως, δημιουργώντας μία νέα ακολουθία επαναλήψεων $\{x'_n\}$ η οποία, γενικά, συγκλίνει ταχύτερα από την $\{x_n\}$. Μία τέτοια διαδικασία είναι η καλούμενη Δ^2 – μέθοδος του Aitken, η οποία εφαρμόζεται όταν είναι γνωστό ότι η ακολουθία $\{x_n\}$ συγκλίνει και η σύγκλιση είναι γραμμική.

Υποθέτουμε κατ' αρχάς ότι έχουμε μία ακολουθία $\{x_n\}$ με $\lim_{n\to\infty} x_n = \xi$ και ισχύει

$$x_{n+1} - \xi = \kappa(x_n - \xi) \ , \ n = 0,1,2,\ldots, \tag{6.81}$$

όπου $0 < |\kappa| < 1$, δηλαδή, τα διαδοχικά σφάλματα είναι όροι γεωμετρικής προόδου (γεωμετρική ακολουθία). Τότε, τα κ και ξ μπορούν να προσδιορισθούν από τρείς διαδοχικούς όρους x_n, x_{n+1}, x_{n+2} της

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ακολουθίας, βάσει των εξισώσεων

$$x_{n+1} - \xi = \kappa(x_n - \xi) \quad \text{και} \quad x_{n+2} - \xi = \kappa(x_{n+1} - \xi) \ . \tag{6.82}$$

Πράγματι, από τις (6.82) προκύπτει

$$\kappa = \frac{x_{n+2} - x_{n+1}}{x_{n+1} - x_n} \quad \text{και}$$

$$\xi = \frac{x_n x_{n+2} - x_{n+1}^2}{x_{n+2} - 2x_{n+1} + x_n} = x_n - \frac{(x_{n+1} - x_n)^2}{x_{n+2} - 2x_{n+1} + x_n} \ . \tag{6.83}$$

Εάν χρησιμοποιήσουμε τον τελεστή Δ των προς τα εμπρός διαφορών, τότε αφού $\Delta x_n = x_{n+1} - x_n$ και $\Delta^2 x_n = x_{n+2} - 2x_{n+1} + x_n$ η (6.83) γράφεται

$$\xi = x_n - \frac{(\Delta x_n)^2}{\Delta^2 x_n} \ . \tag{6.84}$$

Όπως είδαμε προηγουμένως, αν ισχύει η (6.81) τότε η (6.84) μας δίνει το όριο ξ της ακολουθίας $\{x_n\}$. Αλλά, και όταν δεν ισχύει η (6.81) μπορούμε συνήθως να επιταχύνουμε την γραμμική σύγκλιση της ακολουθίας $\{x_n\}$ δημιουργώντας μία νέα ακολουθία $\{x'_n\}$ με την βοήθεια της (6.84), δηλαδή,

$$x'_n = x_n - \frac{(\Delta x_n)^2}{\Delta^2 x_n} \ , \ n = 0,1,2,\ldots \ . \tag{6.85}$$

Η (6.85) καλείται Δ^2–μέθοδος του Aitken. Μπορεί να αποδειχθεί ότι η ακολουθία $\{x'_n\}$ συγκλίνει ταχύτερα από την $\{x_n\}$ στο ξ για $n \to \infty$, με την προϋπόθεση ότι η (6.81) ισχύει ασυμπτωτικά, δηλαδή, για $n \to \infty$. Γενικά, η σύγκλιση της $\{x'_n\}$ είναι επίσης γραμμική. Εάν η ακολουθία $\{x_n\}$ δεν συγκλίνει, τότε και η $\{x'_n\}$ δεν συγκλίνει.

Όταν η ακολουθία $\{x_n\}$ προέρχεται από την επαναληπτική μέθοδο

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$x_{n+1} = g(x_n) \ , \ n = 0,1,2,\dots \ , \tag{6.86}$$

για την οποία ισχύει $0 < |g'(\hat{\xi})| < 1$, όπου $\hat{\xi}$ είναι μία ρίζα της $x = g(x)$, τότε ως γνωστόν η (6.86) για κατάλληλο x_0 είναι πρώτης τάξεως για τον προσδιορισμό της ρίζας $\hat{\xi}$ και ισχύει ασυμπτωτικά η (6.81) με $\kappa = g'(\hat{\xi})$ και $\xi = \hat{\xi}$ (βλέπε θεώρημα 6.4). Επομένως, από τον τύπο (6.85), χρησιμοποιώντας τριάδες διαδοχικών επαναλήψεων x_n, x_{n+1}, x_{n+2} της ακολουθίας (6.86), σχηματίζουμε την ακολουθία $\{x'_n\}$, η οποία θα συγκλίνει ταχύτερα από την $\{x_n\}$ στην ρίζα $\hat{\xi}$.

Ας υποθέσουμε τώρα ότι δίνεται η (6.86). Εάν τεθεί $y_n = g(x_n)$ και $z_n = g(y_n) = g(g(x_n))$, τότε η (6.85) γράφεται

$$x'_n = x_n - \frac{(y_n - x_n)^2}{z_n - 2y_n + x_n} = x_n - \frac{(g(x_n) - x_n)^2}{g(g(x_n)) - 2g(x_n) + x_n} \ . \tag{6.87}$$

Λαμβάνοντας το x'_n που δίνεται από την (6.87) σαν την νέα επανάληψη x_{n+1} έχουμε

$$x_{n+1} = x_n - \frac{(g(x_n) - x_n)^2}{g(g(x_n)) - 2g(x_n) + x_n} \ , \ n = 0,1,2,\dots,$$

δηλαδή, προκύπτει η επαναληπτική μέθοδος:

$$x_{n+1} = \Phi(x_n) \ , \ n = 0,1,2,\dots, \tag{6.88}$$

όπου

$$\Phi(x) = x - \frac{(g(x) - x)^2}{g(g(x)) - 2g(x) + x} = \frac{xg(g(x)) - g^2(x)}{g(g(x)) - 2g(x) + x} \ . \tag{6.89}$$

Η μέθοδος (6.88) καλείται μέθοδος του Steffensen και ισχύουν τα ακόλουθα θεωρήματα:

<u>Θεώρημα 6.12</u> Αν $\Phi(\xi) = \xi$ τότε και $g(\xi) = \xi$. Αντιστρόφως, αν $g(\xi) = \xi$ και $g'(\xi) \neq 1$, τότε $\Phi(\xi) = \xi$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

<u>Απόδειξη</u> Αν $\Phi(\xi) = \xi$, τότε από την (6.89) προκύπτει ότι

$$\frac{(g(\xi) - \xi)^2}{g(g(\xi)) - 2g(\xi) + \xi} = 0 \quad \text{και επομένως} \quad g(\xi) = \xi .$$

Υποθέτουμε τώρα ότι $g(\xi) = \xi$ και υπάρχει η $g'(\xi)$ με $g'(\xi) \neq 1$. Τότε, από την (6.89) και εφαρμόζοντας τον κανόνα του Hospital έχουμε

$$\Phi(\xi) = \frac{\xi g(g(\xi)) - g^2(\xi)}{g(g(\xi)) - 2g(\xi) + \xi} = \frac{0}{0} = \frac{g(g(\xi)) + \xi g'(g(\xi))g'(\xi) - 2g(\xi)g'(\xi)}{g'(g(\xi))g'(\xi) - 2g'(\xi) + 1}$$

$$= \frac{\xi + \xi[g'(\xi)]^2 - 2\xi g'(\xi)}{1 + [g'(\xi)]^2 - 2g'(\xi)} = \frac{\xi[1 - g'(\xi)]^2}{[1 - g'(\xi)]^2} = \xi . \qquad \blacktriangle$$

<u>Θεώρημα 6.13</u> Εάν ξ είναι μία ρίζα της $x = g(x)$ και $g'(\xi) \neq 1$, τότε η μέθοδος Steffensen (6.88) είναι τοπικώς συγκλίνουσα στην ρίζα ξ και είναι τουλάχιστον δευτέρας τάξεως για τον προσδιορισμό της ρίζας ξ.

<u>Απόδειξη</u> Σύμφωνα με το θεώρημα 6.12 είναι $\Phi(\xi) = \xi$. Θα δειχθεί ότι $\Phi'(\xi) = 0$, οπότε έχει εφαρμογή το θεώρημα 6.4. Θέτοντας $h = x - \xi$, όπου το x ανήκει σε μια περιοχή της ξ, από τον τύπο του Taylor παίρνουμε

$$g(x) = g(\xi) + hg'(\xi) + O(h^2) = \xi + hg'(\xi) + O(h^2) , \qquad (6.90)$$

οπότε είναι

$$g(g(x)) = \xi + [\xi + hg'(\xi) + O(h^2) - \xi]g'(\xi) +$$
$$+ O((\xi + hg'(\xi) + O(h^2) - \xi)^2) = \xi + h[g'(\xi)]^2 + O(h^2). \qquad (6.91)$$

Έτσι, από την (6.89) παίρνουμε

$$\Phi(x) = \xi + h - \frac{[\xi + hg'(\xi) + O(h^2) - \xi - h]^2}{\xi + h[g'(\xi)]^2 + O(h^2) - 2[\xi + hg'(\xi) + O(h^2)] + \xi + h} =$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$= \xi + h - \frac{h^2[g'(\xi)]^2 + h^2 - 2h^2 g'(\xi) + O(h^3)}{h[g'(\xi)]^2 + h - 2hg'(\xi) + O(h^2)} =$$

$$= \xi + h - \frac{h^2[g'(\xi) - 1]^2 + O(h^3)}{h[g'(\xi) - 1]^2 + O(h^2)} =$$

$$= \xi + \frac{h^2[g'(\xi) - 1]^2 - h^2[g'(\xi) - 1]^2 + O(h^3)}{h[g'(\xi) - 1]^2 + O(h^2)} =$$

$$= \xi + \frac{O(h^3)}{h[g'(\xi) - 1]^2 + O(h^2)} = \xi + O(h^2) , \qquad (6.92)$$

αφού $g'(\xi) \neq 1$. Η (6.92) συνεπάγεται ότι $\Phi'(\xi) = 0$. ▲

Παρατηρήσεις i) Από την υπόθεση του θεωρήματος 6.13 είναι φανερό ότι η μέθοδος του Steffensen είναι τουλάχιστον δευτέρας τάξεως για τον προσδιορισμό της ρίζας ξ , ανεξάρτητα από την συμπεριφορά της μεθόδου $x_{n+1} = g(x_n)$, $n = 0,1,2,\ldots$, δηλαδή, και όταν αυτή έχει τοπική σύγκλιση στη ρίζα ξ (περίπτωση $|g'(\xi)| < 1$), αλλά και όταν δεν έχει (περίπτωση $|g'(\xi)| > 1$) . Το γεγονός αυτό καθιστά την μέθοδο Steffensen πολύ χρήσιμη στην πράξη.

ii) Κάτω από τις προϋποθέσεις του θεωρήματος 6.13, μπορεί να αποδειχθεί, ότι για τα διαδοχικά σφάλματα στην μέθοδο Steffensen ισχύει

$$\lim_{n\to\infty} \frac{\varepsilon_{n+1}}{\varepsilon_n^2} = \frac{g'(\xi)}{g'(\xi) - 1} \; \frac{g''(\xi)}{2} . \qquad (6.93)$$

Παράδειγμα 6.10 Σύμφωνα με το παράδειγμα 6.5, για την εύρεση της ρίζας $\xi \in [0,1]$ της $f(x) = x2^x - 1 = 0$ προτείνεται ο αλγόριθμος $x_{n+1} = g(x_n)$, $n = 0,1,2,\ldots$, όπου $g(x) = \frac{1}{2^x} = 2^{-x}$. Να βρεθεί

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ο αλγόριθμος της μεθόδου του Steffensen και οι δυο πρώτες επαναλήψεις x_1, x_2 αυτού, χρησιμοποιώντας $x_0 = 1$.

Λύση Η ξ είναι ρίζα της $x = g(x)$ τέτοια ώστε $g'(\xi) \neq 1$. Σύμφωνα με τις (6.88), (6.89), η μέθοδος Steffensen είναι $x_{n+1} = \Phi(x_n)$, $n = 0,1,2,\ldots$, όπου

$$\Phi(x) = \frac{x2^{-(2^{-x})} - (2^{-x})^2}{2^{-(2^{-x})} - 2(2^{-x}) + x} = \frac{x2^{-2^{-x}} - 2^{-2x}}{2^{-2^{-x}} - 2^{1-x} + x} .$$

Για $x_0 = 1$ βρίσκουμε $x_1 = \Phi(x_0) = \Phi(1) = 1 - \frac{\sqrt{2}}{4} \approx 0.646447$, $x_2 \approx 0.641189$. ▲

ΑΣΚΗΣΕΙΣ

1. Η εξίσωση $f(x) = 0$ έχει μία ρίζα $\xi \in [a,b]$. Αν $\bar{x} \in [a,b]$ είναι μία προσεγγιστική τιμή της ρίζας ξ και $f'(x) \geqq M > 0$ $\forall\, x \in [a,b]$, να δειχθεί ότι $|\bar{x} - \xi| \leqq \dfrac{|f(\bar{x})|}{M}$.

2. Να δειχθεί ότι ο συλλογισμός: ο $\bar{x}$ είναι μία καλή προσέγγιση μιας ρίζας ξ της $f(x) = 0$ εάν $|f(\bar{x})|$ είναι πολύ μικρός αριθμός, ενώ είναι κακή προσέγγιση εάν $|f(\bar{x})|$ είναι μεγάλος αριθμός, δεν είναι πάντοτε σωστός.

3. Να προσδιορισθεί το πλήθος των πραγματικών ριζών της $f(x) = x + e^x = 0$.

   Έστωσαν $[a_n,b_n]$, $[a_{n+1},b_{n+1}]$ δύο διαδοχικά διαστήματα στην μέθοδο της διχοτομήσεως. Τότε, ως γνωστόν, μία από τις παρακάτω δύο περιπτώσεις ισχύει: i) $a_{n+1} = \dfrac{a_n + b_n}{2}$, $b_{n+1} = b_n$ ii) $a_{n+1} = a_n$, $b_{n+1} = \dfrac{a_n + b_n}{2}$. Εάν ισχύει η (i), να δειχθεί ότι

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$|a_{n+1} - \xi| \leqq \frac{1}{2} |a_n - \xi|$, ενώ αν ισχύει η (ii) , να δειχθεί ότι $|b_{n+1} - \xi| \leqq \frac{1}{2} |b_n - \xi|$, όπου ξ η ρίζα στην οποία συγκλίνει η μέθοδος.

5. Να προσδιορισθεί ο ελάχιστος αριθμός επαναλήψεων n που απαιτείται με την μέθοδο της διχοτομήσεως ώστε μία ρίζα $\xi \in [a,b]$ της $f(x) = 0$ $(f(a)f(b) < 0)$ να υπολογίζεται με ακρίβεια κ δ.θ. .

6. Εάν η συνάρτηση $g(x)$ της μεθόδου (6.8) είναι ορισμένη στο $[a,b]$ και i) $\forall\, x,y \in [a,b] \Rightarrow |g(x) - g(y)| \leqq L|x - y|$, $0 \leqq L < 1$, ii) $x_0 = \frac{a+b}{2}$ και iii) $|x_1 - x_0| \leqq (1 - L)\left(\frac{b-a}{2}\right)$, τότε η εξίσωση $x = g(x)$ έχει μία μοναδική ρίζα ξ στο $[a,b]$ και η ακολουθία $x_{n+1} = g(x_n)$, $n = 0,1,2,\ldots$ συγκλίνει στην ρίζα ξ . Επιπλέον, είναι $|x_n - \xi| \leqq L^n\left(\frac{b-a}{2}\right)$, $|x_{n+1} - \xi| < |x_n - \xi|$ και $|x_{n+1} - \xi| \leqq \frac{L}{1-L} |x_{n+1} - x_n|$, $n = 0,1,2,\ldots$.

7. Η εξίσωση $x = g(x)$, όπου $g(x) = x^2 - 4x + 4$ έχει μία μοναδική ρίζα $\xi \in [0\,,1.4]$. Να εφαρμοσθεί η επαναληπτική μέθοδος $x_{n+1} = g(x_n)$, $n = 0,1,2,\ldots$ με $x_0 = 0$. Τί παρατηρείτε;

8. Να δειχθεί ότι η επαναληπτική μέθοδος $x_{n+1} = g(x_n)$, $n = 0,1,2,\ldots$ όπου $g(x) = x + (x-2)^2$ συγκλίνει $\forall\, x_0 \in [1,2]$ στην ρίζα $\xi = 2$ της $x = g(x)$.

9. Εάν η $f'(x)$ είναι συνεχής στο $[a,b]$ με $0 < m \leqq f'(x) \leqq M$, $\forall\, x \in [a,b]$, να δειχθεί ότι η $g(x) = x - cf(x)$ ικανοποιεί την $|g'(x)| < 1$, $\forall\, x \in [a,b]$ για κάποια σταθερά c κατάλληλα εκλεγμένη. Τι συμβαίνει αν $m \leqq f'(x) \leqq M < 0$, $\forall\, x \in [a,b]$;

10. Να προσδιορισθεί το μεγαλύτερο διάστημα $I = [a,b]$, τέτοιο ώστε $\forall\, x_0 \in I$ η μέθοδος $x_{n+1} = g(x_n)$, $n = 0,1,2,\ldots$, όπου $g(x) = x - \frac{x^2}{3} + \frac{1}{2}$ να συγκλίνει. Ποιο θα είναι το όριο της ακολουθίας

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$\{x_n\}$, $n = 0,1,2,\ldots$;

11. Ποιες από τις ακόλουθες επαναληπτικές συναρτήσεις $g(x)$ μπορούν να χρησιμοποιηθούν για την εύρεση της ρίζας $\xi \in (\frac{1}{4}, \frac{1}{2})$ της εξισώσεως $x^3 + 2x - 1 = 0$; i) $\frac{1}{2}(1 - x^3)$ ii) $\frac{1 - 2x}{x^2}$ iii) $1 - x - x^3$ iv) $(1 - 2x)^{1/3}$ v) $x - 0.2(x^3 + 2x - 1)$ vi) $x - \frac{x^3 + 2x - 1}{3x^2 + 2}$.

12. Η εξίσωση $x^3 - x^2 + 20x + 10 = 0$ έχει μία ρίζα $\xi \in [-1,0]$. Να δειχθεί ότι η μέθοδος $x_{n+1} = g(x_n)$, $n = 0,1,2,\ldots$, όπου $g(x) = \frac{x^2 - x^3 - 10}{20}$ συγκλίνει στη ρίζα ξ για κατάλληλο x_0 . Να ληφθεί $x_0 = 0$ ή $x_0 = -1$ και να βρεθούν οι 4 πρώτες επαναλήψεις της μεθόδου.

13. Να προσδιορισθούν οι συναρτήσεις $h(x)$ και $H(x)$, έτσι ώστε η μέθοδος $x_{n+1} = g(x_n)$, $n = 0,1,2,\ldots$, όπου $g(x) = x - h(x)f(x) - H(x)f^2(x)$ να είναι τρίτης τάξεως για τον προσδιορισμό μιας απλής ρίζας ξ της $f(x) = 0$, για κατάλληλο x_0 .

14. Να δειχθεί ότι η εξίσωση $xe^{-x} = e^{-3}$ έχει ακριβώς δύο πραγματικές ρίζες ξ_1 , ξ_2 $(\xi_1 < \xi_2)$. Να δειχθεί ότι η επαναληπτική μέθοδος $x_{n+1} = 3 + \ln x_n$, $n = 0,1,2,\ldots$, συγκλίνει στην ξ_2 εάν $x_0 > \xi_1$ και αποκλίνει εάν $x_0 < \xi_1$. Να δειχθεί επίσης ότι η μέθοδος $x_{n+1} = e^{x_n - 3}$, $n = 0,1,2,\ldots$, συγκλίνει στην ξ_1 εάν $x_0 < \xi_2$ και αποκλίνει εάν $x_0 > \xi_2$. Να βρεθούν οι δύο ρίζες ξ_1 , ξ_2 με ακρίβεια 6 δ.θ. χρησιμοποιώντας $x_0 = 2$ και για τις δύο μεθόδους.

15. Η εξίσωση $x^3 - 10x + 4 = 0$ έχει από μία ρίζα σε καθένα από τα διαστήματα $[-4, -3]$, $[0,1]$ και $[2,3]$. i) Να δειχθεί ότι ο

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

αλγόριθμος $x_{n+1} = \frac{1}{10}(x_n^3 + 4)$, $n = 0,1,2,\ldots$, συγκλίνει στην ρίζα που ανήκει στο $[0,1]$, $\forall\, x_0 \in [0,1]$ και είναι πρώτης τάξεως για τον προσδιορισμό αυτής. ii) Είναι ο παραπάνω αλγόριθμος κατάλληλος για τον προσδιορισμό μιας οποιασδήποτε από τις άλλες δύο ρίζες; iii) Να δειχθεί ότι ο αλγόριθμος $x_{n+1} = (10x_n - 4)^{1/3}$, $n = 0,1,2,\ldots$,συγκλίνει $\forall\, x_0 \in [-4, -3]$ στην ρίζα που ανήκει στο $[-4, -3]$, ενώ $\forall\, x_0 \in [2,3]$ συγκλίνει στην ρίζα που ανήκει στο $[2,3]$. iv) Να προσδιορισθούν οι τρείς ρίζες με ακρίβεια 3 δ.θ. .

16. Δίνεται η εξίσωση $x^3 - 3x^2 + 2x = 0$ και η επαναληπτική μέθοδος $x_{n+1} = \frac{3x_n^2 - x_n^3}{2}$, $n = 0,1,2,\ldots$ i) Να δειχθεί ότι $\forall\, x_0 \in [-\frac{1}{4}, \frac{1}{4}]$ αυτή συγκλίνει στη ρίζα $\xi = 0$. ii) Να βρεθεί ένα διάστημα I τέτοιο ώστε $\forall\, x_0 \in I$ να συγκλίνει στην ρίζα $\xi = 2$. iii) Να δειχθεί ότι δεν υπάρχει διάστημα I τέτοιο ώστε $\forall\, x_0 \in I$ να συγκλίνει στη ρίζα $\xi = 1$.

17. Να δειχθεί ότι ο αλγόριθμος $x_{n+1} = \frac{1}{x_n^2 + x_n + 2}$, $n = 0,1,2,\ldots,$ συγκλίνει σε μία πραγματική ρίζα της $x^3 + x^2 + 2x - 1 = 0$, $\forall\, x_0 > 0$.

18. Να εξετασθεί εάν οι αλγόριθμοι: i) $x_{n+1} = \frac{1}{4}\left(x_n^2 + \frac{6}{x_n}\right)$, ii) $x_{n+1} = 4 - \frac{6}{x_n^2}$, $n = 0,1,2,\ldots$, είναι κατάλληλοι για τον υπολογισμό μιας ρίζας της $x^3 - 4x^2 + 6 = 0$ που ανήκει στο $[3,4]$.

19. Να δειχθεί ότι ο αλγόριθμος $x_{n+1} = x_n^3 - 9x_n^2 + 27x_n - 24$, $n = 0,1,2,\ldots,$ είναι τρίτης τάξεως για τον προσδιορισμό μιας ρίζας της $f(x) = x^3 - 9x^2 + 26x - 24 = 0$ για κατάλληλο x_0 .

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

20. Να δειχθεί ότι ο αλγόριθμος (6.44), όπου $p > 1$, συγκλίνει στην ρίζα $\sqrt[p]{N}$, $\forall\, x_0 > 0$.

21. Να δειχθεί ότι ο αλγόριθμος (6.46) συγκλίνει εάν και μόνον εάν $0 < x_0 < \frac{2}{N}$. Εάν $0 < N < 2$ και εκλεγεί $x_0 = 1$, να δειχθεί ότι $x_n = \frac{1-(1-N)^{2^n}}{N}$, $n = 0,1,2,\ldots$.

22. Να δειχθεί ότι ο αλγόριθμος $x_{n+1} = x_n - m\,\frac{f(x_n)}{f'(x_n)}$, $n = 0,1,2,\ldots$, είναι τουλάχιστον δευτέρας τάξεως για τον προσδιορισμό μιας ρίζας ξ της $f(x) = 0$ πολλαπλότητος $m \geqq 1$, για κατάλληλο x_0.

23. Να δειχθεί ότι η μέθοδος Newton – Raphson συγκλίνει $\forall\, x_0 \in [2,3]$ στην ρίζα $\xi \in [2,3]$ της εξισώσεως $f(x) = x^3 - x^2 - x - 3 = 0$. Να βρεθεί η ρίζα ξ με ακρίβεια 4 δ.θ. .

24. Να υπολογισθεί το $S = \sqrt[3]{6 + \sqrt[3]{6 + \sqrt[3]{6 + \ldots}}}$.

25. Να προσδιορισθούν οι σταθερές p,q και r, έτσι ώστε η τάξη της μεθόδου $x_{n+1} = g(x_n)$, $n = 0,1,2,\ldots$, όπου $g(x) = px + \frac{qN}{x^2} + \frac{rN^2}{x^5}$, για τον προσδιορισμό της $\sqrt[3]{N}$ να είναι όσον το δυνατόν μεγαλύτερη (για κατάλληλο x_0).

26. Να δειχθεί ότι κάτω από τις προυποθέσεις (6.54) υπάρχει μία μοναδική απλή ρίζα της $f(x) = 0$ στο $[a_n, b_n]$.

27. Κάτω από τις προυποθέσεις (6.59), να δειχθεί ότι ο αλγόριθμος (6.53) μετασχηματίζεται στον (6.60).

28. Να βρεθούν οι τρείς πρώτες επαναλήψεις της μεθόδου Newton – Raphson για την εξίσωση $f(x) = x^2 + 1 = 0$, λαμβάνοντας $x_0 = \frac{1+i}{2}$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

29. Με την μέθοδο Regula falsi να βρεθούν οι θετικές ρίζες $\xi_1 \in [0,1]$, $\xi_2 \in [2,3]$ της εξισώσεως $x^3 - 10x + 4 = 0$ με ακρίβεια 3 δ.θ..

30. Να βρεθεί η ρίζα $\xi \in [1,2]$ της εξισώσεως $x^3 - x - 3 = 0$ με την μέθοδο της τέμνουσας, όπου $x_0 = 1$, $x_1 = 2$ και με ακρίβεια 6 δ.θ..

31. Οι τρείς ρίζες της εξισώσεως $x^3 + 6x^2 + 11x + 6 = 0$ είναι διαδοχικοί ακέραιοι αριθμοί. Είναι δυνατόν να βρεθεί μία ισοδύναμη μορφή $x = g(x)$ αυτής, όπου $g(x) = \alpha x^3 + \beta x^2 + \gamma x + \delta$, έτσι ώστε ο αλγόριθμος $x_{n+1} = g(x_n)$, $n = 0,1,2,\ldots$ να είναι τρίτης τάξεως για τον προσδιορισμό κάποιας ρίζας, για κατάλληλο x_0 ;

32. Να δειχθεί ότι στην μέθοδο Newton – Raphson, για την εύρεση της απλής ρίζας ξ της $f(x) = 0$, ισχύει η ακόλουθη σχέση μεταξύ των διαδοχικών σφαλμάτων:

$$\varepsilon_{n+1} = \frac{1}{2}\frac{f''(\xi)}{f'(\xi)}\varepsilon_n^2 + \left(\frac{1}{3}\frac{f'''(\xi)}{f'(\xi)} - \frac{1}{2}\left[\frac{f''(\xi)}{f'(\xi)}\right]^2\right)\varepsilon_n^3 + \ldots ,$$

όπου $\varepsilon_n = x_n - \xi$, $n = 0,1,2,\ldots$.

33. Να δειχθεί ότι ο αλγόριθμος $x_{n+1} = x_n\left[3(1 - Nx_n) + (Nx_n)^2\right]$, $n = 0,1,2,\ldots$, είναι τρίτης τάξεως για τον προσδιορισμό του $\frac{1}{N}$, $N > 0$, $\forall\, x_0 \in \left(0, \frac{2}{N}\right)$. Να δειχθεί ότι ισχύει $x_{n+1} - N^{-1} = N^2(x_n - N^{-1})^3$.

34. Να δειχθεί ότι ο αλγόριθμος (Halley)

$$x_{n+1} = x_n - \frac{2f(x_n)f'(x_n)}{2\left[f'(x_n)\right]^2 - f(x_n)f''(x_n)}, \quad n = 0,1,2,\ldots ,$$

είναι τρίτης τάξεως για τον προσδιορισμό μιας απλής ρίζας ξ της

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$f(x) = 0$ για κατάλληλο x_0 . Να βρεθούν οι συγκεκριμένοι αλγόριθμοι αν $f(x) = x^p - N$ για $p = 2,3,$ όπου $N > 0$.

35. Να δειχθεί ότι οι ακόλουθοι αλγόριθμοι: i) $x_{n+1} = \dfrac{x_n p N}{(p-1)N + x_n^p}$,

$n = 0,1,2,\ldots$, ii) $x_{n+1} = x_n - \dfrac{x_n}{Np}(x_n^p - N)$, $n = 0,1,2,\ldots$,

iii) $x_{n+1} = \dfrac{p x_n^{p+1}}{(p+1)x_n^p - N}$, $n = 0,1,2,\ldots$, είναι δευτέρας τάξεως για τον προσδιορισμό της $\sqrt[p]{N}$, $N > 0$, για κατάλληλο x_0 . Να δειχθεί επίσης ότι ο (i) για $p = 2$, συγκλίνει στην $\sqrt{N}$ $\forall\, x_0 > 0$ και επιπλέον το απόλυτο σφάλμα της επαναλήψεως x_n αυτού είναι $\leq$ από το απόλυτο σφάλμα της επαναλήψεως x_n του αλγορίθμου (6.45) , με την προυπόθεση ότι το x_0 είναι το ίδιο για τους δύο αλγορίθμους.

36. Να δειχθεί ότι ο αλγόριθμος

$$x_{n+1} = \frac{1}{4}\left(x_n + \frac{N}{x_n}\right) + \frac{N}{x_n + \dfrac{N}{x_n}} , \quad n = 0,1,2,\ldots ,$$

είναι τετάρτης τάξεως για τον προσδιορισμό της $\sqrt{N}$, $N > 0$, $\forall\, x_0 > 0$.

37. Να δειχθεί ότι ο αλγόριθμος $x_{n+1} = \dfrac{x_n^5 + 10x_n^3 N + 5x_n N^2}{5x_n^4 + 10x_n^2 N + N^2}$,

$n = 0,1,2,\ldots$, είναι πέμπτης τάξεως για τον προσδιορισμό της $\sqrt{N}$, $N > 0$, για κατάλληλο x_0 .

38. Να δειχθεί ότι ο αλγόριθμος $x_{n+1} = \dfrac{x_n}{p}\left[(p+1) - N x_n^p\right]$, $n = 0,1,2,\ldots$, είναι δευτέρας τάξεως για τον προσδιορισμό της

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$\frac{1}{\sqrt[p]{N}}$, $N > 0$, για κατάλληλο x_0 .

39. Να δειχθεί ότι ο αλγόριθμος $x_{n+1} = x_n - \dfrac{f(x_n)f'(x_n)}{\left[f'(x_n)\right]^2 - f(x_n)f''(x_n)}$,

$n = 0,1,2,\ldots$ είναι δευτέρας τάξεως για τον προσδιορισμό μιας ρίζας ξ , οποιασδήποτε πολλαπλότητος, της $f(x) = 0$, για κατάλληλο x_0 .

40. Να δειχθεί ότι ο αλγόριθμος $x_{n+1} = x_n - \dfrac{f^2(x_n)}{f(x_n + f(x_n)) - f(x_n)}$,

$n = 0,1,2,\ldots$, είναι τουλάχιστον δευτέρας τάξεως για τον προσδιορισμό μιας απλής ρίζας ξ της $f(x) = 0$ για κατάλληλο x_0 . Να δειχθεί επίσης ότι αυτός είναι ισοδύναμος με τον αλγόριθμο του Steffensen (6.88), εάν τεθεί $f(x) = g(x) - x$.

41. Για τον αλγόριθμο (6.45) να δειχθεί ότι, αν $0 \leqq x_{n-1} - x_n < \varepsilon$ $(n \geqq 2)$, τότε $0 \leqq x_n - \sqrt{N} < \varepsilon$.

42. Να προσδιορισθούν οι τιμές των a,b,c,d , έτσι ώστε η τάξη του αλγορίθμου

$$x_{n+1} = \frac{x_n(ax_n^2 + bN)}{cx_n^2 + dN} , \quad n = 0,1,2,\ldots ,$$

για τον προσδιορισμό της $\sqrt{N}$, $N > 0$ να είναι όσον το δυνατόν μεγαλύτερη (για κατάλληλο x_0) .

43. Να δειχθεί ότι ο αλγόριθμος του Steffensen (6.88) προκύπτει από τον αλγόριθμο της μεθόδου της τέμνουσας (6.67) εάν τεθεί $f(x) = x - g(x)$ και $x_{n+1} = g(x_n)$, $n = 0,1,2,\ldots$.

44. Για την εύρεση της ρίζας ξ της εξισώσεως $x = g(x)$, όπου $g'(\xi) > 1$ προτείνεται ο αλγόριθμος $x_{n+1} = (1 - \lambda)x_n + \lambda g(x_n)$,

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$n = 0,1,2,\ldots$, όπου $\lambda \neq 0$. Πώς πρέπει να εκλεγεί η σταθερά λ, ώστε ο αλγόριθμος να συγκλίνει ταχύτερα στην ρίζα ξ για κατάλληλο x_0 ;

45. Αν x_n είναι μία προσέγγιση της ρίζας ξ της $f(x) = 0$,τότε σαν μία νέα προσέγγιση x_{n+1} της ρίζας ξ θεωρούμε την ρίζα του πρωτοβαθμίου πολυωνύμου $P_1(x)$, τέτοιου ώστε $P_1(x_n) = f(x_n)$ και $P_1'(x_n) = f'(x_n)$. Να βρεθεί ο αλγόριθμος που δίνει τις διαδοχικές προσεγγίσεις.

46. Η μέθοδος Newton – Raphson εφαρμόζεται στην εξίσωση $f(x) = x^{-\frac{1}{p}}(x^p - N) = 0$, $N > 0$. Για ποιές τιμές του p είναι αυτή τουλάχιστον τρίτης τάξεως για τον προσδιορισμό της ρίζας $\sqrt[p]{N}$.

47. Να εξετασθούν ως προς την σύγκλιση οι κάτωθι αλγόριθμοι για τον προσδιορισμό της $\sqrt[p]{N}$, $N > 0$

i) $$x_{n+1} = \frac{x_n[(p-1)x_n^p + N]}{x_n^p + (p-1)N}, \quad n = 0,1,2,\ldots , \; p \neq 2 ,$$

ii) $$x_{n+1} = \frac{N}{x_n^{p-1}}, \quad n = 0,1,2,\ldots .$$

48. Να υπολογισθούν οι πραγματικές ρίζες των εξισώσεων: i) $e^x - 2x^2 = 0$ ii) $\sin x - \frac{1}{x} + 1 = 0$ iii) $x^2 - \log x - 2 = 0$ iv) $xe^x - 1 = 0$ v) $e^{2x} - 3 = 0$ vi) $x - 0.2 \sin x - 0.5 = 0$ vii) $e^x - 3x = 0$ με ακρίβεια 6 δ.θ. .

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

7

ΠΟΛΥΩΝΥΜΙΚΕΣ ΕΞΙΣΩΣΕΙΣ

7.1 Εισαγωγή

Στο κεφάλαιο αυτό θεωρούμε μία ειδική περίπτωση του προβλήματος που μελετήθηκε στο προηγούμενο κεφάλαιο. Συγκεκριμένα, θα ασχοληθούμε με την αριθμητική επίλυση της αλγεβρικής ή πολυωνυμικής εξισώσεως

$$P(x) = 0 \ , \tag{7.1}$$

όπου

$$P(x) \equiv a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n = \sum_{i=0}^{n} a_i x^{n-i} \tag{7.2}$$

είναι ένα πολυώνυμο βαθμού n με συντελεστές a_i , $i = 0(1)n$ πραγματικούς ή μιγαδικούς αριθμούς $(a_0 \neq 0)$. Η μεταβλητή x θα θεωρείται μιγαδική στην γενική περίπτωση.

Οποιαδήποτε από τις μεθόδους που αναπτύχθηκαν στο προηγούμενο κεφάλαιο μπορεί ασφαλώς να χρησιμοποιηθεί και για την επίλυση της (7.1). Ο λόγος για τον οποίο η εξίσωση (7.1) εξετάζεται ξεχωριστά είναι η μεγάλη σημασία που έχουν τα πολυώνυμα,γενικά,στην Αριθμητική Ανάλυση. Είναι π.χ. γνωστό ότι κάθε συνεχής συνάρτηση μπορεί να προ-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

σεγγιστεί από ένα πολυώνυμο πριν αρχίσουμε να αναζητούμε τις ρίζες της. Λόγω της ειδικής φύσεως των πολυωνύμων (ειδικές ιδιότητες), ο χειρισμός του προβλήματος (7.1) γίνεται απλούστερος και αποτελεσματικώτερος, ακόμη δε, το πρόβλημα (7.1) μπορεί να αντιμετωπισθεί επιτυχώς και με άλλες μεθόδους ειδικά προσαρμοσμένες για πολυωνυμικές εξισώσεις.

Στη συνέχεια, θεωρούμε μόνον την περίπτωση όπου το $P(x)$ στην (7.2) είναι πραγματικό πολυώνυμο, δηλαδή, οι a_i, $i = 0(1)n$ είναι πραγματικοί αριθμοί. Αυτό γίνεται αφ' ένός μεν διότι η περίπτωση αυτή είναι η πιο ενδιαφέρουσα στην πράξη, αφ'ετέρου δε διότι, αν μερικοί ή όλοι οι συντελεστές a_i, $i = 0(1)n$ είναι μιγαδικοί αριθμοί, τότε σχηματίζουμε το πολυώνυμο

$$R(x) \equiv P(x)\left(\sum_{i=0}^{n} a_i^* x^{n-i}\right) = \left(\sum_{i=0}^{n} a_i x^{n-i}\right)\left(\sum_{i=0}^{n} a_i^* x^{n-i}\right), \qquad (7.3)$$

όπου a_i^* είναι ο συζυγής μιγαδικός του a_i. Μπορεί να δειχθεί ότι το $R(x)$, που είναι βαθμού $2n$, έχει πραγματικούς συντελεστές. Ακόμη, είναι φανερό, ότι κάθε ρίζα της (7.1) είναι και ρίζα της

$$R(x) = 0 . \qquad (7.4)$$

Πιο συγκεκριμένα, μπορεί να αποδειχθεί ότι οι πραγματικές ρίζες (αν υπάρχουν) της (7.4) θα είναι αρτίας πολλαπλότητος και κάθε μία απ' αυτές θα είναι ρίζα της (7.1) με την μισή πολλαπλότητα. Οι μιγαδικές ρίζες της (7.4) πρέπει να δοκιμασθούν με αντικατάσταση στην (7.1), για να βρούμε ποιές απ'αυτές είναι και ρίζες της (7.1).

7.2 Γενικές ιδιότητες των πολυωνύμων – Φράγματα ριζών

Στην παράγραφο αυτή παραθέτουμε χωρίς αποδείξεις ορισμένα βασικά και χρήσιμα θεωρήματα που αφορούν τις ρίζες των πολυωνύμων. Με την βοήθεια αυτών αποκτούμε πληροφορίες για το είδος των ριζών και την

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

θέση τους στο μιγαδικό επίπεδο, δηλαδή, βρίσκουμε περιοχές που περιέχουν μία ή περισσότερες ρίζες.

Θεώρημα 7.1 Μία αλγεβρική εξίσωση βαθμού n

$$P(x) = 0 \ , \tag{7.5}$$

έχει ακριβώς n ρίζες, πραγματικές ή μιγαδικές, με την προϋπόθεση ότι μία ρίζα πολλαπλότητος m υπολογίζεται σαν m ρίζες. (Μία ρίζα ξ της (7.5) έχει πολλαπλότητα m εάν $P(\xi) = P'(\xi) = \ldots = P^{(m-1)}(\xi) = 0$ και $P^{(m)}(\xi) \neq 0$) .

Θεώρημα 7.2 Εάν η πολυωνυμική εξίσωση $P(x) = 0$ με πραγματικούς συντελεστές έχει την ρίζα $\xi = a + ib$ (a,b πραγματικά) με πολλαπλότητα κ , τότε και ο αριθμός $\xi^* = a - ib$ είναι επίσης ρίζα αυτής με την ίδια πολλαπλότητα κ .

Πόρισμα 7.1 Κάθε πολυωνυμική εξίσωση περιττού βαθμού με πραγματικούς συντελεστές έχει τουλάχιστον μία πραγματική ρίζα.

Θεώρημα 7.3 Έστωσαν $P(x)$ και $R(x)$ πολυώνυμα βαθμού n και m αντίστοιχα, όπου $1 \leq m \leq n$. Τότε, υπάρχει ένα μοναδικό πολυώνυμο $S(x)$ βαθμού $n - m$ και ένα μοναδικό πολυώνυμο $V(x)$ βαθμού $\leq m - 1$, τέτοια ώστε

$$P(x) = R(x)S(x) + V(x) \ .$$

Θεώρημα 7.4 Εάν ξ_i , $i = 1(1)n$ είναι οι n ρίζες της (7.1) (δεν χρειάζεται να είναι διακεκριμένες), τότε η (7.2) γράφεται

$$P(x) = a_0(x - \xi_1)(x - \xi_2)\ldots(x - \xi_n) \equiv a_0\prod_{i=0}^{n}(x - \xi_i) \ . \tag{7.6}$$

Η (7.6) γράφεται επίσης ως εξής

$$P(x) = a_0\left[x^n - S_1x^{n-1} + S_2x^{n-2} + \ldots + (-1)^nS_n\right] , \tag{7.7}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

όπου $S_1 = \sum_{i=1}^{n} \xi_i$, $S_2 = \sum_{\substack{i,j=1 \\ j<i}}^{n} \xi_i \xi_j$,..., $S_n = \prod_{i=1}^{n} \xi_i$.

Από τις (7.7) και (7.2) προκύπτει ότι

$$\frac{a_\kappa}{a_0} = (-1)^\kappa S_\kappa \ , \ \kappa = 1(1)n \ . \tag{7.8}$$

Ειδικά, για $\kappa = 1,n$ έχουμε

$$\sum_{i=1}^{n} \xi_i = -\frac{a_1}{a_0} \ , \ \prod_{i=1}^{n} \xi_i = (-1)^n \frac{a_n}{a_0} \ . \tag{7.9}$$

<u>Θεώρημα 7.5</u> Έστωσαν ξ_i , $i = 1(1)n$ οι ρίζες της (7.1). Τότε:
i) Οι ρίζες της εξισώσεως

$$x^n P(\tfrac{1}{x}) \equiv a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 = 0 \tag{7.10}$$

(για $a_n \neq 0$) είναι οι $\lambda_i = \frac{1}{\xi_i}$, $i = 1(1)n$.
ii) Οι ρίζες της εξισώσεως

$$P(cx) \equiv \sum_{i=0}^{n} a_i (cx)^{n-i} = 0 \quad (\text{για } c \neq 0)$$

είναι οι $\lambda_i = \frac{\xi_i}{c}$, $i = 1(1)n$. (Αν θέλουμε να ισχύει $\prod_{i=1}^{n} \lambda_i = 1$, τότε σύμφωνα με την (7.9) προκύπτει

$$1 = \prod_{i=1}^{n} \left(\frac{\xi_i}{c} \right) = \frac{\prod_{i=1}^{n} (\xi_i)}{c^n} = \frac{(-1)^n \frac{a_n}{a_0}}{c^n} \ . \text{ Άρα ,}$$

$$c^n = (-1)^n \frac{a_n}{a_0} \ , \text{ οπότε } \quad c = \sqrt[n]{(-1)^n \frac{a_n}{a_0}} \ .$$

Είναι φανερό ότι εάν $a_0 a_n < 0$ και n άρτιος τότε το c δεν είναι

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

πραγματικός αριθμός. Εάν θέλουμε όμως να ισχύει $\left|\prod_{i=1}^{n} \lambda_i\right| = 1$, τότε το c εκλέγεται πάντοτε πραγματικό και δίνεται από την $c = \left|\frac{a_n}{a_0}\right|^{\frac{1}{n}} = \left|\prod_{i=1}^{n} \xi_i\right|^{\frac{1}{n}}$).

iii) Οι ρίζες της εξισώσεως

$$(-1)^n P(-x) \equiv a_0 x^n - a_1 x^{n-1} + a_2 x^{n-2} - \ldots + (-1)^n a_n = 0 \quad (7.11)$$

είναι οι $\lambda_i = -\xi_i$, $i = 1(1)n$.

iv) Οι ρίζες της εξισώσεως

$$P(x + c) \equiv P(c) + \frac{x}{1!} P'(c) + \frac{x^2}{2!} P''(c) + \ldots + \frac{x^n}{n!} P^{(n)}(c) = 0 \quad (7.12)$$

είναι οι $\lambda_i = \xi_i - c$, $i = 1(1)n$. (Αν θέλουμε να ισχύει $\sum_{i=1}^{n} \lambda_i = 0$, τότε, σύμφωνα με την (7.9), πρέπει να ληφθεί $c = -\frac{a_1}{na_0}$) .

<u>Πόρισμα 7.2</u> i) Αν οι ρίζες ξ_i, $i = 1(1)n$ της (7.1) πληρούν την σχέση $|\xi_i| < \rho$, τότε οι ρίζες λ_i της $P(\rho x) = 0$ πληρούν την $|\lambda_i| < 1$.

ii) Αν η (7.1) έχει μία ρίζα εντός του κύκλου $|x - c| \leq \rho$, τότε η εξίσωση $P(\rho x + c) = 0$ έχει μία ρίζα εντός του μοναδιαίου κύκλου ($|x| \leq 1$) .

Τα ακόλουθα θεωρήματα δίνουν φράγματα για τα μέτρα των ριζών και τις πραγματικές ρίζες της πολυωνυμικής εξισώσεως (7.1), όπου το $P(x)$ δίνεται από την (7.2) και έχει πραγματικούς συντελεστές και είναι από τα πιο χρήσιμα μεταξύ των πολλών που υπάρχουν.

<u>Θεώρημα 7.6</u> Για όλες τις ρίζες ξ_i, $i = 1(1)n$ της (7.1) ισχύουν:

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

i) $r < |\xi_i| < R$, όπου

$$r = \frac{1}{1 + \max\limits_{0 \leq j \leq n-1} \left|\frac{a_j}{a_n}\right|} , \ (a_n \neq 0) \quad \text{και} \quad R = 1 + \max_{1 \leq j \leq n} \left|\frac{a_j}{a_0}\right| . \tag{7.14}$$

<u>Σημείωση</u> Οι αριθμοί r και R είναι τα κάτω και άνω φράγματα αντίστοιχα για τις θετικές ρίζες της (7.1). Ομοίως οι αριθμοί $-R$ και $-r$ είναι τα κάτω και άνω φράγματα αντίστοιχα των αρνητικών ριζών της (7.1) .

ii) $$|\xi_i| \leq \max\left\{ 1 , \sum_{j=1}^{n} \left|\frac{a_j}{a_0}\right| \right\} \tag{7.15}$$

iii) $$|\xi_i| \leq \max\left\{ \left|\frac{a_n}{a_0}\right| , 1 + \max_{1 \leq j \leq n-1} \left|\frac{a_j}{a_0}\right| \right\} \tag{7.16}$$

iv) $$|\xi_i| \leq \sum_{j=0}^{n-1} \left|\frac{a_{j+1}}{a_j}\right| \tag{7.17}$$

v) $$|\xi_i| \leq 2 \max_{1 \leq j \leq n} \left|\frac{a_j}{a_0}\right|^{\frac{1}{j}} \tag{7.18}$$

vi) $$|\xi_i| \leq \max\left\{ \left|\frac{a_n}{a_{n-1}}\right| , 2 \max_{0 \leq j \leq n-2} \left|\frac{a_{j+1}}{a_j}\right| \right\} \tag{7.19}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

vii) $$\frac{|a_n|}{\sqrt{\mu}} \leq |\xi_i| \leq \max\left\{1, \sqrt{\frac{\mu}{|a_0|^2} - \frac{|a_n|^2}{\mu}}\right\}, \qquad (7.20)$$

όπου $\mu = \sum_{j=0}^{n} |a_j|^2$.

Παρατήρηση Κάτω φράγματα για τις $|\xi_i|$ στις περιπτώσεις (ii) – – (vi) του θεωρήματος 7.6 μπορούν να προσδιορισθούν αφού βρεθούν τα άνω φράγματα για τις $|\lambda_i|$, όπου λ_i είναι οι ρίζες της (7.10) .

Θεώρημα 7.7 Αν στην (7.1) είναι $a_0 = 1$ και c είναι ένας αριθμός (γενικά μιγαδικός) τέτοιος ώστε $P'(c) \neq 0$, τότε ο κύκλος $|x - c| \leq n \left|\frac{P(c)}{P'(c)}\right|$ περιέχει τουλάχιστον μία ρίζα της (7.1) .

Θεώρημα 7.8 Αν στην (7.1) είναι $a_0 > 0$, τότε κάθε πραγματική ρίζα ξ της (7.1) ικανοποιεί τις ακόλουθες ανισότητες:

i) $$\xi < 1 + \frac{|a_K|}{a_0}, \qquad (7.21)$$

όπου a_K είναι ο αρνητικός συντελεστής του (7.2) με το μεγαλύτερο μέτρο (εάν δεν υπάρχει αρνητικός συντελεστής τότε $\xi \leq 0$) .

ii) $$\xi < 1 + \left(\frac{|a_K|}{a_0}\right)^{\frac{1}{r}}, \qquad (7.22)$$

όπου a_K ορίζεται όπως και στην περίπτωση (i) και r είναι ο δείκτης του πρώτου αρνητικού συντελεστού στην ακολουθία $a_0, a_1, \ldots, a_n$ $(r \geq 1)$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\text{iii)} \quad \xi < 1 + \left(\frac{|a_K|}{a_S} \right)^{\frac{1}{r-s}} , \tag{7.23}$$

όπου a_K και r ορίζονται όπως στις (i) και (ii) και a_S είναι ο μεγαλύτερος από τους πρώτους r θετικούς συντελεστές του (7.2) .

Παρατήρηση Για την εύρεση αντίστοιχων κάτω φραγμάτων των πραγματικών ριζών της (7.1) εφαρμόζουμε το θεώρημα 7.8 στην εξίσωση (7.12).

Τα ακόλουθα θεωρήματα δίνουν απάντηση στο ερώτημα που αφορά το πλήθος των πραγματικών ριζών της (7.1) σε ένα γνωστό διάστημα (a,b) .

Θεώρημα 7.9 (Descarte's) Ο αριθμός κ των θετικών ριζών της (7.1) ικανοποιεί τις σχέσεις: i) $\kappa \leqq \upsilon$ και ii) $\upsilon - \kappa$ είναι άρτιος ακέραιος $\geqq 0$, όπου υ είναι ο αριθμός των αλλαγών προσήμου στην ακολουθία $a_0, a_1, \ldots, a_n$ (οι μηδενικοί συντελεστές παραλείπονται.Μία αλλαγή προσήμου συμβαίνει όταν μετά από ένα θετικό συντελεστή ακολουθεί ένας αρνητικός ή αντιστρόφως. Μία ρίζα πολλαπλότητος m υπολογίζεται σαν m ρίζες).

Παράδειγμα 7.1 Να βρεθεί ο αριθμός των θετικών ριζών της

$$P(x) = x^5 - x^3 - x^2 + x - 1 = 0 . \tag{7.24}$$

Λύση Η ακολουθία των προσήμων είναι $+ - - + -$, δηλαδή, υπάρχουν τρείς αλλαγές προσήμου και επομένως για την (7.24) είναι $\kappa = 1$ ή $\kappa = 3$. ▲

Παρατήρηση Στηριζόμενοι στο θεώρημα 7.5 (iii) μπορούμε να βρούμε το πλήθος των αρνητικών ριζών της (7.1) .

Παράδειγμα 7.2 Να βρεθεί ο αριθμός των αρνητικών ριζών της (7.24).

Λύση Θεωρούμε την εξίσωση $(-1)^5 P(-x) = (-1)^5(-x^5 + x^3 - x^2 - x - 1) = x^5 - x^3 + x^2 + x + 1 = 0$. Ο αριθμός των θετικών ρι-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ζών αυτής σύμφωνα με το θεώρημα 7.9 είναι 0 ή 2 . Επομένως η (7.24) έχει 0 ή 2 αρνητικές ρίζες. ▲

Θεώρημα 7.10 Εάν για το $P(x)$ στην (7.1) και για το διάστημα $[a,b]$ ισχύει $P(a)P(b) < 0$, τότε στο (a,b) υπάρχει ένας περιττός αριθμός ριζών της (7.1), υπολογίζοντας και τις πολλαπλότητες των ριζών (δηλαδή, μία ρίζα πολλαπλότητος m υπολογίζεται σαν m ρίζες).

Εάν ισχύει η $P(a)P(b) > 0$, τότε στο (a,b) είτε δεν υπάρχουν καθόλου ρίζες της (7.1) ή υπάρχει ένας άρτιος αριθμός ριζών (υπολογίζοντας και τις πολλαπλότητες των ριζών).

Θεώρημα 7.11 (Hua's) Εάν όλες οι ρίζες της (7.1) είναι πραγματικές, τότε ισχύει $a_K^2 > a_{K-1}a_{K+1}$, $K = 1(1)n-1$.

Πόρισμα 7.3 Εάν στην (7.1) για κάποιο $K = 1(1)n-1$ ισχύει $a_K^2 \leq a_{K-1}a_{K+1}$, τότε η (7.1) έχει τουλάχιστον ένα ζεύγος μιγαδικών ριζών.

Θέωρημα 7.12 (Budan – Fourier) Αν $P(x)$ είναι το πολυώνυμο (7.2) με $a_0 > 0$, $P(a)P(b) \neq 0$, όπου $a < b$ και $w(x)$ συμβολίζει τον αριθμό των αλλαγών προσήμου στην ακολουθία $P(x)$, $P'(x)$, $P''(x)$, ..., $P^{(n)}(x)$, τότε ο αριθμός των πραγματικών ριζών (υπολογίζοντας και τις πολλαπλότητες των ριζών) της (7.1) στο διάστημα $[a,b]$ είναι είτε ίσος με $w(a) - w(b)$ ή μικρότερος κατά ένα άρτιο αριθμό.

Θεώρημα 7.13 (Sturm) Έστω ότι η εξίσωση (7.1) έχει όλες τις ρίζες απλές. Εάν $a < b$ και $P(a)P(b) \neq 0$, τότε υπάρχουν ακριβώς $V(a) - V(b)$ πραγματικές ρίζες της (7.1) στο διάστημα (a,b) , όπου $V(x)$ συμβολίζει τον αριθμό των αλλαγών προσήμου στην ακολουθία:

$$P(x),\ P'(x),\ r_1(x),\ldots,\ r_K(x) \quad ,$$

όπου $-r_1(x)$ είναι το υπόλοιπο της διαιρέσεως $\dfrac{P(x)}{P'(x)}$, $-r_2(x)$ το

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

υπόλοιπο της διαιρέσεως $\frac{P'(x)}{r_1(x)}$ κ.ο.κ. και $-r_K(x)$ είναι το υπόλοιπο της διαιρέσεως $\frac{r_{K-2}(x)}{r_{K-1}(x)}$, όπου $r_K(x)$ είναι μία σταθερά $\neq 0$.

7.3 Η μέθοδος της συνθετικής διαιρέσεως (Μέθοδος του Horner)

Εάν χρησιμοποιήσουμε την μέθοδο Newton – Raphson του προηγουμένου κεφαλαίου για την επίλυση της (7.1), τότε, ως γνωστόν, οι διαδοχικές προσεγγίσεις για κάποια ρίζα υπολογίζονται από τον αλγόριθμο

$$x_{n+1} = x_n - \frac{P(x_n)}{P'(x_n)} \quad , \; n = 0,1,2,\ldots \; . \tag{7.25}$$

Δηλαδή, για την εύρεση της x_{n+1} απαιτείται ο υπολογισμός της τιμής του $P(x)$ καθώς επίσης και της τιμής της παραγώγου $P'(x)$ στο σημείο $x = x_n$. Γενικά, το πρόβλημα του υπολογισμού της τιμής ενός πραγματικού πολυωνύμου για μια πραγματική ή μιγαδική τιμή της μεταβλητής του είναι από τα πιο συνηθισμένα στους αριθμητικούς υπολογισμούς.

Ας υποθέσουμε ότι δίνεται το πολυώνυμο

$$P(x) \equiv a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n \; , \tag{7.26}$$

όπου a_i , $i = 0(1)n$ είναι πραγματικοί αριθμοί με $a_0 \neq 0$ και θέλουμε να υπολογίσουμε την τιμή $P(\xi)$, όπου ξ είναι ένας πραγματικός ή μιγαδικός αριθμός. Προφανώς, είναι

$$P(\xi) = a_0\xi^n + a_1\xi^{n-1} + \ldots + a_{n-1}\xi + a_n \; . \tag{7.27}$$

Η εύρεση της $P(\xi)$ γίνεται πιο εύκολα, δηλαδή, με λιγώτερες πράξεις, εάν η (7.27) γραφεί ως εξής:

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$P(\xi) = (\ldots((a_0\xi + a_1)\xi + a_2)\xi + \ldots)\xi + a_n \quad . \tag{7.28}$$

Σύμφωνα με την (7.28) διαδοχικά υπολογίζουμε τους αριθμούς

$$\begin{aligned} b_0 &= a_0 \\ b_1 &= b_0\xi + a_1 \\ b_2 &= b_1\xi + a_2 \\ &\vdots \\ b_n &= b_{n-1}\xi + a_n \quad , \end{aligned} \tag{7.29}$$

οπότε $P(\xi) = b_n$. Οι σχέσεις (7.29) γράφονται επίσης ως εξής

$$b_i = b_{i-1}\xi + a_i \ , \quad i = 0(1)n \quad \text{με} \quad b_{-1} \equiv 0 \ . \tag{7.30}$$

Ο αλγόριθμος που χρησιμοποιεί τις (7.30) για τον υπολογισμό της $P(\xi)$ είναι γνωστός σαν μέθοδος του Horner ή μέθοδος της συνθετικής διαιρέσεως.

Θα δειχθεί τώρα ότι οι αριθμοί b_i , $i = 0(1)n-1$ στην (7.30) είναι οι συντελεστές του πολυωνύμου $Z(x)$ που προκύπτει σαν πηλίκο κατά την διαίρεση του πολυωνύμου $P(x)$ δια του διωνύμου $x - \xi$.Πράγματι, έστω

$$Z(x) = c_0x^{n-1} + c_1x^{n-2} + \ldots + c_{n-2}x + c_{n-1} \tag{7.31}$$

και

$$P(x) = (x - \xi)Z(x) + c_n \quad . \tag{7.32}$$

Από την (7.32) προκύπτει ότι $c_n = P(\xi)$ και

$$\begin{aligned} P(x) &= (x - \xi)(c_0x^{n-1} + c_1x^{n-2} + \ldots + c_{n-2}x + c_{n-1}) + c_n = \\ &= c_0x^n + (c_1 - c_0\xi)x^{n-1} + (c_2 - c_1\xi)x^{n-2} + \ldots + \\ &+ (c_{n-1} - c_{n-2}\xi)x + (c_n - c_{n-1}\xi) \quad . \end{aligned} \tag{7.33}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Από τις (7.26) και (7.33) παίρνουμε

$$c_0 = a_0 \ , \quad c_i - c_{i-1}\xi = a_i \ , \ i = 1(1)n$$

ή, λόγω και των (7.29),

$$c_0 = a_0 = b_0 \ , \quad c_i = c_{i-1}\xi + a_i = b_i \ , \ i = 1(1)n \ ,$$

πράγμα που θέλαμε να αποδείξουμε. Άρα, από τις (7.30) βρίσκουμε τους συντελεστές b_i , $i = 0(1)n-1$ του πηλίκου $Z(x)$ και το υπόλοιπο $P(\xi) = b_n$.

Στην πράξη,όταν οι υπολογισμοί γίνονται με το χέρι,χρησιμοποιούμε τον ακόλουθο σχηματικό τρόπο (σχήμα του Horner) για την εύρεση των b_i , $i = 0(1)n$

$$\begin{array}{c|ccccc} & a_0 & a_1 & a_2 & \dots & a_n \\ \xi & & b_0\xi & b_1\xi & \dots & b_{n-1}\xi \\ \hline & b_0 & b_1 & b_2 & \dots & b_n = P(\xi) \end{array} \ ,$$

του οποίου η εξήγηση είναι προφανής.

<u>Παρατήρηση</u> Αν είναι $a_0 > 0$, $\xi > 0$ και ισχύει $b_i \geqq 0$, $i = 1(1)n$ στην (7.30), τότε από την (7.33) προκύπτει ότι για $x > \xi$ είναι $P(x) > 0$, οπότε όλες οι πραγματικές ρίζες ξ_i , $i = 1(1)n$ του $P(x)$ θα ικανοποιούν την σχέση $\xi_i \leqq \xi$, $i = 1(1)n$.

Πρέπει να σημειωθεί ότι η παραπάνω μέθοδος του Horner είναι περισσότερο κατάλληλη για την περίπτωση που το ξ είναι πραγματικός αριθμός, γιατί, αν το ξ είναι μιγαδικός αριθμός, τότε αυξάνει σημαντικά το πλήθος των πράξεων,επειδή θα πρέπει να εκτελούνται πράξεις με μιγαδικούς αριθμούς. Για την περίπτωση που το ξ είναι μιγαδικό μια πιο αποτελεσματική μέθοδος θα αναπτυχθεί στη συνέχεια, αφού όμως πρώτα εξετάσουμε πώς είναι δυνατόν η μέθοδος του Horner να χρησιμοποιηθεί για τον υπολογισμό των τιμών των παραγώγων του $P(x)$ στο πραγ-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ματικό σημείο ξ .

Κατ' αρχάς από την (7.32) παραγωγίζοντας έχουμε

$$P'(x) = Z(x) + (x - \xi)Z'(x) \quad . \tag{7.34}$$

Θέτοντας στην (7.34) $x = \xi$ προκύπτει ότι $P'(\xi) = Z(\xi)$, δηλαδή, η τιμή της πρώτης παραγώγου του $P(x)$ για $x = \xi$ είναι η τιμή $Z(\xi)$, η οποία βρίσκεται εφαρμόζοντας πάλι την μέθοδο του Horner στο πολυώνυμο $Z(x)$, που είναι ήδη γνωστό.

Καλούμε τώρα $Z_1(x) \equiv Z(x)$ και θεωρούμε την ταυτότητα της διαιρέσεως πρώτα του $P(x)$ δια $x - \xi$ και στη συνέχεια των διαδοχικών πηλίκων δια $x - \xi$. Έτσι έχουμε

$$\begin{aligned} P(x) &= (x - \xi)Z_1(x) + P(\xi) \quad . \\ Z_1(x) &= (x - \xi)Z_2(x) + Z_1(\xi) \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ Z_{n-1}(x) &= (x - \xi)Z_n(x) + Z_{n-1}(\xi) \\ Z_n(x) &= (x - \xi) \cdot 0 + Z_n(\xi) \quad . \end{aligned} \tag{7.35}$$

Εξ άλλου,αν το $P(x)$ αναπτυχθεί σε σειρά Taylor γύρω από το σημείο ξ θα έχουμε

$$P(x) = P(\xi) + \frac{(x - \xi)}{1!} P'(\xi) + \frac{(x - \xi)^2}{2!} P''(\xi) + \ldots +$$

$$+ \frac{(x - \xi)^n}{n!} P^{(n)}(\xi) = P(\xi) + (x - \xi) \left[\frac{P'(\xi)}{1!} + \frac{(x - \xi)}{2!} P''(\xi) + \right.$$

$$\left. + \ldots + \frac{(x - \xi)^{n-1}}{n!} P^{(n)}(\xi) \right] \quad . \tag{7.36}$$

Από την πρώτη των (7.35) και την (7.36) προκύπτει ότι

$$Z_1(x) = \frac{P'(\xi)}{1!} + \frac{(x - \xi)}{2!} P''(\xi) + \ldots + \frac{(x - \xi)^{n-1}}{n!} P^{(n)}(\xi) \ , \tag{7.37}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

οπότε

$$Z_1(\xi) = \frac{P'(\xi)}{1!} \quad . \tag{7.38}$$

Η (7.37), λόγω της (7.38), γίνεται

$$Z_1(x) = Z_1(\xi) + (x-\xi)\left[\frac{P''(\xi)}{2!} + \frac{(x-\xi)}{3!}P'''(\xi) + \ldots + \right.$$

$$\left. + \frac{(x-\xi)^{n-2}}{n!}P^{(n)}(\xi)\right] \quad . \tag{7.39}$$

Συγκρίνοντας την (7.39) με την δεύτερη των (7.35) συνεπάγεται ότι

$$Z_2(x) = \frac{P''(\xi)}{2!} + \frac{(x-\xi)}{3!}P'''(\xi) + \ldots + \frac{(x-\xi)^{n-2}}{n!}P^{(n)}(\xi)$$

και επομένως

$$Z_2(\xi) = \frac{P''(\xi)}{2!} \quad .$$

Συνεχίζοντας την ίδια διαδικασία, τελικά καταλήγουμε στον γενικό τύπο

$$P^{(\kappa)}(\xi) = \kappa! Z_\kappa(\xi) \ , \quad \kappa = 1(1)n \ , \tag{7.40}$$

δηλαδή, η τιμή μιας παραγώγου του $P(x)$ οποιασδήποτε τάξεως στο σημείο ξ μπορεί να βρεθεί επαναλαμβάνοντας την διαδικασία του σχήματος του Horner στα διαδοχικά πηλίκα $Z_i(x)$, $i = 1(1)n$ όσες φορές χρειάζεται. Σημειώνεται ακόμη ότι η (7.36), λόγω της (7.40), γράφεται

$$P(x) = P(\xi) + (x-\xi)Z_1(\xi) + (x-\xi)^2 Z_2(\xi) + \ldots + (x-\xi)^n Z_n(\xi) \ . \tag{7.41}$$

Θέτοντας $y = x - \xi$ στην (7.41) έχουμε αμέσως την μετασχηματισμένη

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

μορφή του πολυωνύμου $P(x)$ βάσει της αντικαταστάσεως $x = y + \xi$. Οι τιμές $P(\xi)$ και $Z_i(\xi)$, $i = 1(1)n$ μπορούν να βρεθούν, όπως τονίστηκε προηγουμένως, με διαδοχικές εφαρμογές του σχήματος του Horner.

<u>Παράδειγμα 7.3</u> Για το πολυώνυμο $P(x) = x^4 + 3x^3 - 5x - 10$, να βρεθούν: i) Οι τιμές $P(-2)$, $P'''(-2)$, ii) Το πηλίκο $Z(x)$ της διαιρέσεως του $P(x)$ δια $x + 2$ και iii) Το νέο πολυώνυμο ως προς y αν τεθεί $x = y - 2$ στο $P(x)$.

<u>Λύση</u>

	1	3	0	−5	−10
−2		−2	−2	4	2
	1	1	−2	−1	$-8 = P(-2)$
−2		−2	2	0	
	1	−1	0	$-1 = \frac{P'(-2)}{1!}$	
−2		−2	6		
	1	−3	$6 = \frac{P''(-2)}{2!}$		
−2		−2			
	1	$-5 = \frac{P'''(-2)}{3!}$			
−2					
	$1 = \frac{P^{(4)}(-2)}{4!}$				

Άρα είναι $P(-2) = -8$, $P'''(-2) = 3!(-5) = -30$, $Z(x) = x^3 + x^2 - 2x - 1$, $P(y-2) = -8 - y + 6y^2 - 5y^3 + y^4$. ▲

Με σκοπό να βρούμε τις μιγαδικές ρίζες ενός πολυωνύμου $P(x)$ είναι απαραίτητο μερικές φορές να υπολογίζουμε το $P(x)$ για διάφορες μιγαδικές τιμές του x (βλέπε παράγραφο 6.5.4). Ο υπολογισμός του

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$P(x)$ που δίνεται από την (7.26) για μια μιγαδική τιμή του x γίνεται με την ακόλουθη αποτελεσματική μέθοδο, η οποία αποτελεί γενίκευση της μεθόδου του Horner.

Αν $\xi = u + iv$ είναι μία μιγαδική τιμή, όπου u,v είναι πραγματικά και $i = \sqrt{-1}$, τότε θέτουμε $\xi^* = u - iv$. Από την ταυτότητα της διαιρέσεως του $P(x)$ δια $(x - \xi)(x - \xi^*)$ έχουμε

$$P(x) = (x - \xi)(x - \xi^*)Z(x) + \kappa x + \lambda \tag{7.42}$$

ή

$$P(x) = (x^2 - 2ux + u^2 + v^2)Z(x) + \kappa x + \lambda , \tag{7.43}$$

όπου, όπως θα δειχθεί παρακάτω, τα κ,λ και οι συντελεστές του $Z(x)$ είναι πραγματικοί αριθμοί. Θέτοντας $p = 2u$ και $q = -(u^2 + v^2)$ η (7.43) γράφεται

$$P(x) = (x^2 - px - q)Z(x) + \kappa x + \lambda . \tag{7.44}$$

Από την (7.42) προκύπτει αμέσως ότι

$$P(\xi) = \kappa\xi + \lambda = \kappa(u + iv) + \lambda = (\kappa u + \lambda) + i\kappa v . \tag{7.45}$$

Άρα, για την εύρεση της $P(\xi)$ μας χρειάζονται οι τιμές των κ και λ. Αν υποθέσουμε ότι

$$Z(x) = b_0x^{n-2} + b_1x^{n-3} +\ldots+ b_{n-3}x + b_{n-2} , \tag{7.46}$$

τότε από την (7.44) προκύπτει ότι

$$P(x) = (x^2 - px - q)(b_0x^{n-2} + b_1x^{n-3} +\ldots+ b_{n-3}x + b_{n-2}) + \kappa x + \lambda. \tag{7.47}$$

Εξισώνοντας τους συντελεστές των ίδιων δυνάμεων του x στα δύο μέλη της (7.47) παίρνουμε

$$a_0 = b_0 \tag{7.48}$$

$$a_1 = b_1 - pb_0$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$
\begin{aligned}
a_2 &= b_2 - pb_1 - qb_0 \\
&\ldots\ldots\ldots\ldots\ldots\ldots \\
a_{n-2} &= b_{n-2} - pb_{n-3} - qb_{n-4} \\
a_{n-1} &= \kappa - pb_{n-2} - qb_{n-3} \\
a_n &= \lambda - qb_{n-2} \quad .
\end{aligned}
\tag{7.48}
$$

Εάν οι σχέσεις (7.48) λυθούν ως προς b_i , $i = 0(1)n-2$ βρίσκουμε

$$b_i = a_i + pb_{i-1} + qb_{i-2} \ , \ i = 0(1)n-2 \ , \tag{7.49}$$

όπου $b_{-1} \equiv 0$, $b_{-2} \equiv 0$ και οι δύο τελευταίες σχέσεις των (7.48) δίνουν

$$
\begin{aligned}
\kappa &= a_{n-1} + pb_{n-2} + qb_{n-3} \\
\lambda &= a_n + qb_{n-2} \quad ,
\end{aligned}
\tag{7.50}
$$

οπότε από την (7.45) υπολογίζεται η $P(\xi)$.

Είναι φανερό από τις (7.49) και (7.50) ότι, αφού τα a_i και p,q είναι πραγματικά, τα b_i καθώς και τα κ,λ θα είναι πραγματικά.

Στην πράξη, όταν οι υπολογισμοί γίνονται με το χέρι, ακολουθείται η παρακάτω διάταξη, που είναι γνωστή σαν γενίκευση του σχήματος του Horner, για την εύρεση των b_i , $i = 0(1)n-2$ και κ,λ

	a_0	a_1	a_2	a_3	$\ldots$	a_{n-2}	a_{n-1}	a_n
p		pb_0	pb_1	pb_2	$\ldots$	pb_{n-3}	pb_{n-2}	
q			qb_0	qb_1	$\ldots$	qb_{n-4}	qb_{n-3}	qb_{n-2}
	b_0	b_1	b_2	b_3	$\ldots$	b_{n-2}	κ	λ

Εάν τώρα παραγωγίσουμε την (7.42) ως προς x παίρνουμε

$$P'(x) = (x-\xi^*)Z(x) + (x-\xi)Z(x) + (x-\xi)(x-\xi^*)Z'(x) + \kappa \ , \tag{7.51}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

οπότε

$$P'(\xi) = (\xi - \xi^*)Z(\xi) + \kappa = i2vZ(\xi) + \kappa \ . \qquad (7.52)$$

Η (7.52) μας δίνει την τιμή της πρώτης παραγώγου του $P(x)$ για $x = \xi$. Η τιμή $Z(\xi)$ βρίσκεται ομοίως με την γενικευμένη μέθοδο του Horner ,αφού το $Z(x)$ έχει βρεθεί κατά την εύρεση της $P(\xi)$.

<u>Παρατήρηση</u> Η γενικευμένη μέθοδος του Horner ή της συνθετικής διαιρέσεως για την εύρεση του πηλίκου $Z(x)$ και του υπολοίπου $\kappa x + \lambda$ της διαιρέσεως του $P(x)$ δια του τριωνύμου $x^2 - px - q$, ισχύει και όταν το $x^2 - px - q$ έχει πραγματικές ρίζες .

<u>Παράδειγμα 7.4</u> Δίνεται το πολυώνυμο $P(x) = x^5 - x^3 + x^2 - 1$. Να βρεθούν οι τιμές $P(\xi)$, $P'(\xi)$, όπου $\xi = -1 + 2i$ και το πηλίκο $Z(x)$ της διαιρέσεως του $P(x)$ δια $(x - \xi)(x - \xi^*)$, όπου $\xi^* = -1 - 2i$.

<u>Λύση</u> Στην περίπτωσή μας είναι $u = -1$, $v = 2$, οπότε $p = -2$, $q = -5$. Σύμφωνα με την γενικευμένη μέθοδο του Horner παίρνουμε

	1	0	−1	1	0	−1
−2		−2	4	4	−30	
−5			−5	10	10	−75
	1	−2	−2	15	−20	−76

Άρα, το πηλίκο είναι $Z(x) = x^3 - 2x^2 - 2x + 15$ και σύμφωνα με την (7.45) είναι

$$P(\xi) = (-20)(-1) + (-76) + i(-20)(2) = -56 - 40i =$$
$$= -8(7 + 5i) \ .$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Για την εύρεση της $P'(\xi)$ βρίσκουμε τώρα την $Z(\xi)$

	1	−2	−2	15
−2		−2	8	
−5			−5	20
	1	−4	1	35

Έτσι, έχουμε $Z(\xi) = -1 + 35 + 2i = 34 + 2i$ και από την (7.52) παίρνουμε

$$P'(\xi) = i2(2)(34 + 2i) + (-20) = -4(7 - 34i) \quad . \qquad \blacktriangle$$

7.4 Η μέθοδος Newton – Bairstow

Όπως τονίστηκε στην παράγραφο 6.5.4, η μέθοδος Newton – Raphson μπορεί να χρησιμοποιηθεί για την εύρεση των μιγαδικών ριζών μιας πολυωνυμικής εξισώσεως, με την προυπόθεση ότι η αρχική προσέγγιση θα είναι μιγαδικός αριθμός. Έτσι, είναι φανερό ότι στους υπολογισμούς υπεισέρχονται πράξεις μεταξύ μιγαδικών αριθμών. Η μέθοδος Newton – Bairstow, που αναπτύσσεται στη συνέχεια, χρησιμοποιείται κυρίως για την εύρεση ζευγών συζυγών μιγαδικών ριζών της $P(x) = 0$ και αποφεύγει πράξεις μεταξύ μιγαδικών αριθμών. Πιο συγκεκριμένα, η μέθοδος Newton – Bairstow χρησιμοποιείται για την εύρεση πραγματικών δευτεροβαθμίων παραγόντων του πολυωνύμου $P(x)$. Είναι φανερό ότι όλες οι μιγαδικές ρίζες του $P(x)$ θα βρίσκονται μεταξύ των ριζών των δευτεροβαθμίων αυτών παραγόντων του $P(x)$.

Η μέθοδος βασίζεται στο γεγονός ότι οι ρίζες ενός πραγματικού δευτεροβαθμίου τριωνύμου

$$x^2 - px - q \qquad (7.53)$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

είναι επίσης ρίζες της δοθείσης πολυωνυμικής εξισώσεως

$$P(x) = a_0 x^n + a_1 x_1^{n-1} + \ldots + a_n = 0 \tag{7.54}$$

εάν και μόνον εάν στην ταυτότητα της διαιρέσεως

$$P(x) = (x^2 - px - q)Z(x) + Kx + \Lambda \tag{7.55}$$

ισχύει

$$K = \Lambda = 0 \quad . \tag{7.56}$$

Γενικά στην (7.55), όλοι οι συντελεστές του $Z(x)$ και τα K , Λ μπορούν να εκφραστούν σαν συναρτήσεις των p και q , δηλαδή, έχουμε

$$K = K(p,q) \quad \text{και} \quad \Lambda = \Lambda(p,q) \ . \tag{7.57}$$

Επομένως, οι (7.56) ισχύουν όταν τα p και q ικανοποιούν το ακόλουθο σύστημα:

$$\begin{cases} K(p,q) = 0 \\ \Lambda(p,q) = 0 \quad . \end{cases} \tag{7.58}$$

Το σύστημα (7.58) είναι ένα μη γραμμικό σύστημα δύο εξισώσεων με δύο αγνώστους και μπορεί να λυθεί αριθμητικά με μεθόδους που εξετάζονται στο επόμενο κεφάλαιο. Μία από αυτές είναι και η γενικευμένη μέθοδος Newton – Paphson για συστήματα, που για την περίπτωση του συστήματος (7.58) καλείται μέθοδος Newton – Bairstow. Ο αλγόριθμος αυτής που τον δεχόμαστε προς το παρόν χωρίς απόδειξη (για την απόδειξη αυτού βλέπε κεφάλαιο 8) είναι:

$$\begin{cases} p_{n+1} = p_n - \dfrac{1}{J(p_n,q_n)} \left[K(p_n,q_n) \dfrac{\partial \Lambda}{\partial q}(p_n,q_n) - \Lambda(p_n,q_n) \dfrac{\partial K}{\partial q}(p_n,q_n) \right], \\ q_{n+1} = q_n - \dfrac{1}{J(p_n,q_n)} \left[\Lambda(p_n,q_n) \dfrac{\partial K}{\partial p}(p_n,q_n) - K(p_n,q_n) \dfrac{\partial \Lambda}{\partial p}(p_n,q_n) \right], \end{cases} \tag{7.59}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

όπου

$$J(p_n,q_n) = \frac{\partial K}{\partial p}(p_n,q_n)\frac{\partial \Lambda}{\partial q}(p_n,q_n) - \frac{\partial K}{\partial q}(p_n,q_n)\frac{\partial \Lambda}{\partial p}(p_n,q_n) \neq 0 ,$$

$$n = 0,1,2,\ldots$$

και (p_0,p_0) η αρχική προσέγγιση μιας λύσεως του (7.58) . Για την εύρεση των μερικών παραγώγων στις (7.51) εργαζόμαστε ως εξής: Παραγωγίζοντες την (7.55) ως προς p και q παίρνουμε (το $P(x)$ είναι ανεξάρτητο των p και q):

$$\frac{\partial P(x)}{\partial p} \equiv 0 = (x^2 - px - q)\frac{\partial Z(x)}{\partial p} - xZ(x) + \frac{\partial K}{\partial p}x + \frac{\partial \Lambda}{\partial p} ,$$

$$\frac{\partial P(x)}{\partial q} \equiv 0 = (x^2 - px - q)\frac{\partial Z(x)}{\partial q} - Z(x) + \frac{\partial K}{\partial q}x + \frac{\partial \Lambda}{\partial q} . \tag{7.60}$$

Από την ταυτότητα της διαιρέσεως του $Z(x)$ δια του $x^2 - px - q$ έχουμε

$$Z(x) = (x^2 - px - q)Z_1(x) + K_1x + \Lambda_1 . \tag{7.61}$$

Υποθέτοντας ότι το $x^2 - px - q$ έχει δύο διακεκριμένες ρίζες (πράγμα που συμβαίνει όταν αυτές είναι συζυγείς μιγαδικές) ξ_0 και ξ_1 , από την (7.61) έπεται

$$Z(\xi_i) = K_1\xi_i + \Lambda_1 \quad , \; i = 0,1 . \tag{7.62}$$

Λόγω των (7.62) ,από τις (7.60) προκύπτουν οι εξισώσεις:

$$-\xi_i(K_1\xi_i + \Lambda_1) + \frac{\partial K}{\partial p}\xi_i + \frac{\partial \Lambda}{\partial p} = 0 ,$$

$$-(K_1\xi_i + \Lambda_1) + \frac{\partial K}{\partial q}\xi_i + \frac{\partial \Lambda}{\partial q} = 0 \quad , \quad i = 0,1 . \tag{7.63}$$

Από την δεύτερη των (7.63) έχουμε, αφού $\xi_0 \neq \xi_1$,

$$\frac{\partial K}{\partial q} = K_1 \quad \text{και} \quad \frac{\partial \Lambda}{\partial q} = \Lambda_1 . \tag{7.64}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Λόγω των (7.64), η πρώτη των (7.63) γίνεται

$$-\xi_i\left(\frac{\partial K}{\partial q}\xi_i+\frac{\partial \Lambda}{\partial q}\right)+\frac{\partial K}{\partial p}\xi_i+\frac{\partial \Lambda}{\partial p}=0 \quad \text{ή}$$

$$-\xi_i^2\frac{\partial K}{\partial q}+\xi_i\left(\frac{\partial K}{\partial p}-\frac{\partial \Lambda}{\partial q}\right)+\frac{\partial \Lambda}{\partial p}=0\ ,\ i=0,1\ . \tag{7.65}$$

Επειδή $\xi_i^2-p\xi_i-q=0$, $i=0,1$, προκύπτει ότι $\xi_i^2=p\xi_i+q$ και επομένως η (7.65) γράφεται

$$-(p\xi_i+q)\frac{\partial K}{\partial q}+\xi_i\left(\frac{\partial K}{\partial p}-\frac{\partial \Lambda}{\partial q}\right)+\frac{\partial \Lambda}{\partial p}=0 \quad \text{ή}$$

$$\xi_i\left(\frac{\partial K}{\partial p}-\frac{\partial \Lambda}{\partial q}-\frac{\partial K}{\partial q}\right)+\frac{\partial \Lambda}{\partial p}-q\frac{\partial K}{\partial q}=0\ ,\ i=0,1\ . \tag{7.66}$$

Έτσι, λόγω των (7.64), από την (7.66) προκύπτει ότι

$$\frac{\partial K}{\partial p}=\Lambda_1+pK_1\ ,\quad \frac{\partial \Lambda}{\partial p}=qK_1\ . \tag{7.67}$$

Από την παραπάνω ανάλυση συμπεραίνουμε τα εξής: Εάν έχουμε μία αρχική προσέγγιση $u+iv$ μιας μιγαδικής ρίζας της (7.54), τότε, θεωρώντας και την συζυγή της $u-iv$, κατασκευάζουμε το δευτεροβάθμιο τριώνυμο που έχει αυτές σαν ρίζες, δηλαδή, το $x^2-p_0x-q_0$, όπου $p_0=2u$ και $q_0=-(u^2+v^2)$. Το ζεύγος (p_0,q_0) αποτελεί την αρχική προσέγγιση μιας λύσεως του (7.58). Από τον αλγόριθμο (7.59), αρχίζοντας με το (p_0,q_0) βρίσκουμε ένα ζεύγος (p_n,q_n) που προσεγγίζει μια λύση του (7.58) με την επιζητούμενη ακρίβεια και λύνοντας κατόπιν την εξίσωση $x^2-p_nx-q_n=0$ προσδιορίζουμε τις ζητούμενες προσεγγίσεις για δύο συζυγείς μιγαδικές ρίζες της (7.54).

Πρέπει να σημειωθεί ότι ο αλγόριθμος (7.59), λόγω των (7.64) και (7.67), γίνεται

$$p_{n+1}=p_n-\frac{1}{J(p_n,q_n)}\left[K(p_n,q_n)\Lambda_1(p_n,q_n)-\Lambda(p_n,q_n)K_1(p_n,q_n)\right]\ , \tag{7.68}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$q_{n+1} = q_n - \frac{1}{J(p_n,q_n)} \Big[\Lambda(p_n,q_n)\Big(\Lambda_1(p_n,q_n) + p_n K_1(p_n,q_n)\Big) - $$
$$- q_n K(p_n,q_n) K_1(p_n,q_n)\Big] \ , \qquad (7.68)$$

όπου $J(p_n,q_n) = \Big[\Lambda_1(p_n,q_n) + p_n K_1(p_n,q_n)\Big]\Lambda_1(p_n,q_n) - q_n\Big[K_1(p_n,q_n)\Big]^2$,

$n = 0,1,2,\ldots$ και (p_0,q_0) η αρχική προσέγγιση, και ότι οι τιμές των K, Λ και K_1, Λ_1 στο σημείο (p_n,q_n) υπολογίζονται με την γενικευμένη μέθοδο του Horner (βλέπε παράγραφο 7.3), έχοντας υπ' όψιν τις (7.55) και (7.61). Συχνά οι $p_0 = q_0 = 0$ χρησιμοποιούνται σαν αρχικές τιμές, εάν δεν υπάρχουν καθόλου πληροφορίες για τις μιγαδικές ρίζες. Η μέθοδος Newton – Bairstow συγκλίνει γρήγορα σε μερικές περιπτώσεις (συνήθως όταν έχουμε καλές αρχικές τιμές p_0,q_0), ενώ σε άλλες περιπτώσεις μπορεί να μην συγκλίνει.

<u>Παράδειγμα 7.5</u> Δοθέντος ότι $-\frac{1}{2} + i$ είναι η αρχική προσέγγιση μιας ρίζας της εξισώσεως $x^3 - x^2 - x - 2 = 0$, να εφαρμοσθεί η μέθοδος Newton – Bairstow μια φορά για την εύρεση μιας καλύτερης προσεγγίσεως.

<u>Λύση</u> Έχουμε $u + iv = -0.5 + i$, οπότε $u - iv = -0.5 - i$, $p_0 = -1$, $q_0 = -\frac{5}{4} = -1.25$ και $x^2 - p_0 x - q_0 = x^2 + x + \frac{5}{4}$. Η εύρεση των τιμών των K, Λ, K_1, Λ_1 στο σημείο $(p_0,q_0) = (-1, -\frac{5}{4})$ γίνεται με την μέθοδο του Horner ως ακολούθως:

	1	−1	−1	−2
−1		−1	2	
−1.25			−1.25	2.5
	1	−2	−0.25 = K	0.5 = Λ
−1				
−1.25		0		
	1 = K_1	−2 = Λ_1		

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Έτσι, έχουμε $K = -0.25$, $\Lambda = 0.5$, $K_1 = 1$, $\Lambda_1 = -2$, οπότε είναι
$J(p_0,q_0) = \{-2 + (-1)1\}(-2) - (-1.25)\cdot 1^2 = 6 + 1.25 = 7.25$
και από την (7.68) παίρνουμε

$$p_1 = -1 - \frac{1}{7.25}\left[(-0.25)(-2) - 0.5\cdot 1\right] = -1 ,$$

$$q_1 = -1.25 - \frac{1}{7.25}\left[0.5\{(-2) + (-1)\cdot 1\} + 1.25(-0.25)\cdot 1\right] =$$

$$= -1.25 + \frac{1.8125}{7.25} = -1 .$$

Η ζητούμενη προσέγγιση προκύπτει λύνοντας την εξίσωση $x^2 - p_1x - q_1 \equiv$ $\equiv x^2 + x + 1 = 0$. Οι ρίζες αυτής είναι $\frac{1}{2}(-1 \pm i\sqrt{3})$, που είναι και ρίζες της δοθείσης εξισώσεως, αφού $(x^3 - x^2 - x - 2) \equiv (x^2 + x + 1)(x - 2)$.

▲

7.5 Άλλες μέθοδοι για πολυωνυμικές εξισώσεις

Υπάρχει ένας μεγάλος αριθμός μεθόδων που μπορούν να χρησιμοποιηθούν για την αριθμητική επίλυση πολυωνυμικών εξισώσεων, από τις οποίες άλλες είναι περισσότερο και άλλες λιγώτερο αποτελεσματικές. Οι πιο απλούστερες απ' αυτές, ως προς την ανάλυση και τον προγραμματισμό, έχουν περιορισμένες δυνατότητες. Από τις πιο γνωστές είναι και οι ακόλουθες μέθοδοι: i) Η μέθοδος Muller, σε συνδυασμό με την μέθοδο Cauchy. Οι δύο αυτές μέθοδοι σχετίζονται μεταξύ των όπως οι μέθοδοι τέμνουσας και Newton – Raphson. ii) Η μέθοδος Lehmer – Schur. iii) Η μέθοδος Bernoulli, η οποία βασίζεται σε μεθόδους για την εύρεση ιδιοτιμών πινάκων (βλέπε κεφάλαιο 11) . iv) Η μέθοδος Laguerre, που χρησιμοποιείται για την εύρεση των ριζών ενός πολυωνύμου $P(x)$ βαθμού n που έχει μόνον πραγματικές ρίζες. Ο αλγόριθμος αυτής είναι:

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$x_{K+1} = x_K - \frac{nP(x_K)}{P'(x_K) \pm \sqrt{(n-1)\left[(n-1)(P'(x_K))^2 - nP(x_K)P''(x_K)\right]}},$$

$$\kappa = 0,1,2,\ldots . \qquad (7.69)$$

Το πρόσημο της τετραγωνικής ρίζας στην (7.69) εκλέγεται το ίδιο με το πρόσημο της $P'(x_K)$. Η μέθοδος συγκλίνει για κάθε x_0 πραγματικό.
v) Η μέθοδος Root – squaring (Graffe). vi) Η μέθοδος Quotient – Difference που ενδείκνυται κυρίως για την εύρεση καλών αρχικών προσεγγίσεων των ριζών. Αυτή έχει το πλεονέκτημα ότι δεν χρησιμοποιεί καμμία άλλη πληροφορία παρά μόνον τους συντελεστές του πολυωνύμου.

Η συστηματική μελέτη και ανάλυση των παραπάνω μεθόδων είναι εκτός του σκοπού του παρόντος βιβλίου.

7.6 Γενικές οδηγίες και παρατηρήσεις για την εύρεση των ριζών πολυωνυμικών εξισώσεων

Στην πράξη, για την εύρεση όλων των ριζών της πολυωνυμικής εξισώσεως (7.1), ακολουθείται συνήθως η εξής τακτική. Πρώτα υπολογίζονται οι πραγματικές ρίζες (εάν υπάρχουν) και κατόπιν τα ζεύγη των συζυγών μιγαδικών ριζών (εάν υπάρχουν). Για την εύρεση των πραγματικών ριζών ένα σημαντικό βήμα είναι η απομόνωση αυτών, που επιτυγχάνεται με την εύρεση των φραγμάτων των ριζών και στη συνέχεια με την εύρεση του πλήθους των ριζών στα διαστήματα που περιέχουν ρίζες (βλέπε παράγραφο 7.2). Έχοντας υπ' όψιν το θεώρημα 7.5 (i), όταν ψάχνουμε για τις πραγματικές ρίζες της (7.1), είναι αρκετό να βρούμε τις πραγματικές ρίζες των $P(x) = 0$ και $x^n P(\frac{1}{x}) = 0$ στο διάστημα $[-1, 1]$.

Ας υποθέσουμε τώρα ότι, εφαρμόζοντας μία από τις μεθόδους του κεφαλαίου 6 , π.χ. την μέθοδο Newton – Raphson, βρήκαμε την προσεγγιστική τιμή ξ (με κάποια ακρίβεια) για μία πραγματική ρίζα της (7.1) . Τότε, με την βοήθεια της μεθόδου του Horner βρίσκουμε το πολυώνυμο

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$Z(x)$ βαθμού $n-1$ που πληροί την

$$P(x) = (x-\xi)Z(x) + P(\xi) \ . \tag{7.70}$$

Εάν η τιμή $P(\xi)$ είναι μικρή απολύτως (γενικά είναι $P(\xi) \neq 0$), τότε οι ρίζες του $Z(x)$ θα είναι γενικά κοντά στις υπόλοιπες ρίζες του $P(x)$. Έτσι, για τον υπολογισμό των υπολοίπων ριζών του $P(x)$ εργαζόμαστε με το πολυώνυμο $Z(x)$ βαθμού $n-1$. Η διαδικασία αυτή της διαιρέσεως του αρχικού πολυωνύμου δια $x-\xi$ καλείται deflation και είναι χρήσιμη, διότι απαλλάσσει το πολυώνυμο $P(x)$ από την ρίζα ξ που υπολογίσθηκε (απομακρύνει τον "παράγοντα" $x-\xi$) και επομένως η ίδια ρίζα μπορεί να βρεθεί πάλι, εάν πρόκειται περί πολλαπλής ρίζας. Στη συνέχεια, μία πραγματική ρίζα ξ_1 του $Z(x)$ υπολογίζεται και η διαδικασία deflation επαναλαμβάνεται στο $Z(x)$ για την ρίζα ξ_1 κ.ο.κ., μέχρις ότου τελικά βρεθούν (προσεγγιστικά) όλες οι πραγματικές ρίζες του $P(x)$.

Με την εύρεση και απαλοιφή όλων των πραγματικών ριζών, έχει απομείνει τελικά ένα πολυώνυμο $R(x)$ αρτίου βαθμού m που θα περιέχει τα ζεύγη των συζυγών μιγαδικών ριζών της (7.1). Η επαναληπτική μέθοδος, η μέθοδος Newton – Raphson και η μέθοδος της τέμνουσας του κεφαλαίου 6, μπορούν να χρησιμοποιηθούν (όπως και η μέθοδος Newton – Bairstow) για την εύρεση των μιγαδικών ριζών του $R(x)$, με την προυπόθεση βέβαια ότι οι αρχικές προσεγγίσεις αυτών θα είναι μιγαδικές. Γενικά, το πρόβλημα της εκλογής των αρχικών τιμών, για την εύρεση μιγαδικών ριζών με μία επαναληπτική μέθοδο, δεν είναι τόσο εύκολο στην πράξη (όσο στην περίπτωση των πραγματικών ριζών), αφού συχνά είναι απαραίτητο η αρχική προσέγγιση να είναι πολύ κοντά σε μία ρίζα για να συγκλίνει η μέθοδος (π.χ. μέθοδος Newton – Raphson) .Τώρα, αντί για διαστήματα ψάχνουμε για περιοχές (κύκλους) στο μιγαδικό επίπεδο που περιέχουν τις ρίζες. Τα θεωρήματα 7.6 και 7.7 δίνουν τέτοιες πληροφορίες. Ακόμη, στηριζόμενοι πάλι στο θεώρημα 7.5(i), αρκεί να βρούμε τις μιγαδικές ρίζες των πολυωνυμικών εξισώσεων $R(x) = 0$ και $x^m R(\frac{1}{x}) = 0$ στην περιοχή $|x| \leq 1$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Έτσι, βρίσκουμε, με κάποια από τις μεθόδους που αναφέρθηκαν προηγουμένως, ένα ζεύγος συζυγών μιγαδικών ριζών του $R(x)$ μέ κάποια ακρίβεια. Στη συνέχεια,εφαρμόζεται η διαδικασία deflation για μιγαδικές ρίζες, δηλαδή, με την βοήθεια της γενικευμένης μεθόδου του Horner απομακρύνουμε τον δευτεροβάθμιο πραγματικό "παράγοντα" $x^2 - px - q$, που έχει ρίζες αυτές που βρέθηκαν, από το πολυώνυμο $R(x)$. Στο πολυώνυμο που απομένει εφαρμόζεται η ίδια εργασία κ.ο.κ., μέχρις ότου υπολογισθούν όλες οι μιγαδικές ρίζες του $R(x)$.

Πρέπει να σημειωθεί ότι κάθε φορά που ένας "παράγων", πρωτοβάθμιος (για πραγματική ρίζα) ή δευτεροβάθμιος (για συζυγείς μιγαδικές ρίζες), απαλείφεται από το αρχικό πολυώνυμο $P(x)$ ή τα διαδοχικά πηλίκα, απομένει ένα πολυώνυμο του οποίου οι ρίζες δεν είναι ακριβώς ρίζες του αρχικού πολυωνύμου $P(x)$,λόγω της μεταδόσεως των σφαλμάτων που υπεισέρχονται στα διαδοχικά βήματα. Για τον λόγο αυτό είναι συνήθως καλύτερο από την άποψη της ακρίβειας των ριζών να υπολογίζονται και να απαλείφονται οι ρίζες (πραγματικές ή μιγαδικές) κατά αύξουσα τάξη των μέτρων τους. Επίσης, οι προσεγγιστικές τιμές που βρίσκονται τελικά για όλες τις ρίζες πρέπει να ελεγχθούν με αντικατάσταση στο αρχικό πολυώνυμο $P(x)$ και εάν είναι απαραίτητο πρέπει να βελτιωθούν χρησιμοποιώντας μία επαναληπτική μέθοδο στο αρχικό πολυώνυμο $P(x)$.

Ένα πρόβλημα που παρουσιάζεται συχνά στις εφαρμογές είναι η αστάθεια των ριζών του πολυωνύμου $P(x)$. Δηλαδή, σχετικά μικρές μεταβολές στους συντελεστές του μπορούν να επιφέρουν σχετικά μεγάλες μεταβολές στις ρίζες του. Για παράδειγμα, η εξίσωση $x^2 - 1.498x + 0.561 = 0$ έχει τις ακριβείς ρίζες $\xi_1 = 0.75$ και $\xi_2 = 0.748$, ενώ η εξίσωση $x^2 - 1.5x + 0.561 = 0$ έχει ρίζες τις $p_1 \approx 0.789$ και $p_2 \approx 0.711$ (με ακρίβεια 3 δ.θ.). Για τον λόγο αυτό,πολλά προβλήματα που θα μπορούσαν να λυθούν λύνοντας μία πολυωνυμική εξίσωση,συνήθως,λύνονται με άλλες μεθόδους. Σαν συμπέρασμα, μπορούμε να πούμε ότι το γενικό πρόβλημα της ευρέσεως όλων των ριζών ενός πραγματικού αυθαίρετου πολυωνύμου είναι αρκετά δύσκολο στην πράξη και πρέπει να αποφεύγεται εάν είναι δυνατόν.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ΑΣΚΗΣΕΙΣ

1. Να βρεθούν φράγματα για τις ρίζες των εξισώσεων: i) $x^3 - 2x^2 - x + 2 = 0$ ii) $19x^5 - 7x^3 + 2x^2 - 8x + 38 = 0$ iii) $3x^9 + 6x^7 - 2x^6 + x^5 - 2 = 0$ iv) $6x^8 + 20x^7 + 5x^6 + x^4 + x^3 + x^2 + 17x - 144 = 0$. Να δειχθεί ότι η (iv) έχει τουλάχιστον δύο πραγματικές ρίζες.

2. Να δειχθεί ότι η εξίσωση $x^4 - 2x^3 - 2x + 1 = 0$ έχει δύο θετικές ρίζες και δύο (συζυγείς) μιγαδικές.

3. Να προσδιορισθεί ο αριθμός των πραγματικών ριζών της $x^3 - x^2 + 2x - 3 = 0$ στο διάστημα $(0,2)$ βάσει του θεωρήματος 7.12 .

4. Να δειχθεί ότι η εξίσωση $x^3 + ax^2 + bx + c = 0$ έχει τουλάχιστον μία θετική ρίζα αν $c < 0$, ενώ αν $c > 0$ έχει τουλάχιστον μία αρνητική ρίζα.

5. Εάν η εξίσωση (7.1) έχει n πραγματικές ρίζες ξ_i, $i = 1(1)n$, να δειχθεί ότι

$$|\xi_i| \leq \frac{1}{|a_0|}(a_1^2 - 2a_0a_2)^{\frac{1}{2}} \; ; \; i = 1(1)n \; .$$

6. Να δειχθεί ότι για τον υπολογισμό της τιμής του πολυωνύμου $P(x) = \sum_{\kappa=0}^{n} a_\kappa x^{n-\kappa}$ στο πραγματικό σημείο ξ με άμεση αντικατάσταση (υπολογισμός όρου προς όρο) απαιτούνται $2n - 1$ πολλαπλασιασμοί και n προσθαφαιρέσεις, ενώ με την μέθοδο Horner απαιτούνται μόνον n πολλαπλασιασμοί και n προσθαφαιρέσεις.

7. Να προσδιορισθεί η πολλαπλότης της ρίζας $\xi = 1$ του πολυωνύμου $P(x) = x^5 - 2x^4 + 4x^3 - x^2 - 7x + 5$, με την βοήθεια του σχήματος του Horner. Να βρεθούν κατόπιν οι τιμές $P'(2)$ και $P''(2)$.

8. Να προσδιορισθούν τα a,b αν είναι γνωστό ότι το $P(x) = 4x^4 + ax^3 + 2x^2 - 3x + b$ διαιρείται ακριβώς με το $x^2 - 1$ (μέθοδος Horner)

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

9. Με την βοήθεια της μεθόδου Horner να δειχθεί ότι

$$\frac{x^2 + x + 1}{(x + 2)^4} = \frac{1}{(x + 2)^2} - \frac{3}{(x + 2)^3} + \frac{3}{(x + 2)^4} .$$

10. Να βρεθεί το πηλίκο και το υπόλοιπο της διαιρέσεως του πολυωνύμου $x^4 - 2x^3 - 11x^2 + 12x + 36$ δια του $2x + 1$ (μέθοδος Horner).

11. Να βρεθεί το πηλίκο και το υπόλοιπο της διαιρέσεως του $P(x) = nx^{n+1} - (n + 1)x^n + 1$ δια του $(x - 1)^2$ (μέθοδος Horner).

12. Να υπολογισθούν οι τιμές $P(1 \pm i)$, $P'(1 \pm i)$ όταν: i) $P(x) = x^4 - 7x^3 + 2x^2 - x - 5$ ii) $P(x) = x^6 - 5x^4 - 2x^2 + 5$.

13. Να υπολογισθεί η τιμή $P(2 + 3i)$, όπου $P(x) = x^3 - (2 + i)x + 1$.

14. Να βρεθεί το πηλίκο και το υπόλοιπο της διαιρέσεως i) του $P(x) = 2x^6 - 4x^5 - 9x^4 - 10x^3 - 7x^2 + 13x - 6$ δια του $2x^2 + 6x + 10$, ii) του $P(x) = x^4 + 2x^3 + 2x + 15$ δια του $x^2 + x - 1$.

15. Για την εξίσωση $x^3 - 2 = 0$ και με αρχική προσέγγιση $-\frac{2}{3} + \frac{2}{3} i$ για μία ρίζα αυτής να εφαρμοσθεί η μέθοδος Newton – Bairstow μία φορά.

16. Για το πολυώνυμο $x^3 - 3x^2 - x + 9$ να βρεθούν οι συναρτήσεις $K = K(p,q)$ και $\Lambda = \Lambda(p,q)$ (βλέπε (7.57)).

17. Να δειχθεί ότι εάν η μέθοδος Newton – Bairstow εφαρμοσθεί στην εξίσωση $a_0x^2 + a_1x + a_2 = 0$, $a_0 \neq 0$ με p_0 και q_0 οποιαδήποτε, τότε προκύπτει $p_1 = -\frac{a_1}{a_0}$ και $q_2 = -\frac{a_2}{a_0}$.

18. Να εφαρμοσθεί μία φορά η μέθοδος Newton – Bairstow στις ακόλουθες εξισώσεις: i) $2x^5 + 4x^3 - 1 = 0$, $(p_0,q_0) = (2, -2)$ ii) $x^5 - 2x^4 + 3x^3 - 4x^2 + 5x - 6$, $(p_0,q_0) = (-1, -2)$ iii) $2x^6 + 3x^5 - 5x^4 + 3x^3 + 7x^2 + 8x + 1$, $(p_0,q_0) = (2, -2)$ iv) $5x^4 + 7x^2 + x + 3 = 0$, $(p_0,q_0) = (0.5, -1)$ v) $x^3 - 4x^2 + 6x - 4$, $(p_0,q_0) = (0,0)$ vi) $x^3 + 2x^2 + 3x + 2$, $(p_0,q_0) = (-2, -3)$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

19. Να υπολογισθούν οι ρίζες των εξισώσεων: 1) $x^4 - 1.1x^3 + 2.3x^2 + 0.5x + 3.3 = 0$ 2) $x^3 - 3x^2 - x + 9 = 0$ 3) $x^3 + 2x^2 - x + 1 = 0$ 4) $x^3 - x - 1 = 0$ 5) $x^5 - 3x - 1 = 0$ 6) $x^3 + x - 1 = 0$ 7) $x^3 - x^2 - x - 1 = 0$ 8) $4x^4 - 9x^3 + 1 = 0$ 9) $x^4 - 8x^3 + 23x^2 + 16x - 50 = 0$ 10) $x^8 - 15x^4 - 16 = 0$ 11) $x^{10} + 2x^8 - x^2 - 2 = 0$ 12) $x^6 - 2x^5 + 5x^4 - 6x^3 + 2x^2 + 8x - 8 = 0$ 13) $x^3 + (1 - 3i)x^2 - (8 + 5i)x - 6 + 12i = 0$ 14) $x^3 + (2 + i)x^2 - 2ix + 2 - 4i = 0$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

8

ΣΥΣΤΗΜΑΤΑ ΜΗ ΓΡΑΜΜΙΚΩΝ ΕΞΙΣΩΣΕΩΝ

8.1 Εισαγωγή

Όπως είδαμε στην παράγραφο 6.5.4, το πρόβλημα της ευρέσεως των μιγαδικών ριζών μιας πραγματικής ή μιγαδικής συναρτήσεως $f(x)$ μπορεί να μετατραπεί στο πρόβλημα της ευρέσεως των πραγματικών λύσεων ενός πραγματικού μη γραμμικού συστήματος δύο εξισώσεων με δύο αγνώστους Μία μέθοδος για την επίλυση ενός τέτοιου συστήματος, η οποία θα μελετηθεί στη συνέχεια, χρησιμοποιήθηκε για να προκύψει η μέθοδος Newton – – Bairstow.

Σκοπός του κεφαλαίου αυτού είναι να δειχθεί πώς μερικές μέθοδοι του κεφαλαίου 6 μπορούν να γενικευθούν ώστε να εφαρμόζονται σε συστήματα n μη γραμμικών εξισώσεων με n αγνώστους. Χάριν απλότητος, θα περιοριστούμε μόνον στην περίπτωση ενός συστήματος δύο πραγματικών εξισώσεων με δύο αγνώστους $(n = 2)$, δηλαδή, του

$$\begin{cases} f_1(x,y) = 0 \\ f_2(x,y) = 0 , \end{cases} \tag{8.1}$$

όπου υποτίθεται ότι η απαλοιφή ενός αγνώστου είναι αδύνατη ή δύσκολη.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Η επέκταση για οποιοδήποτε n γίνεται εύκολα.

8.2 Γραφική παράσταση του (8.1)

Για την εύρεση προσεγγιστικών αρχικών τιμών των πραγματικών λύσεων του (8.1) μπορούμε να κατασκευάσουμε τις γραφικές παραστάσεις των συναρτήσεων $f_1(x,y) = 0$ και $f_2(x,y) = 0$. Οι συντεταγμένες (x_i, y_i) των σημείων τομής (εάν υπάρχουν) αυτών είναι οι πραγματικές λύσεις του (8.1). (Πρέπει να σημειωθεί ότι η εξίσωση $f_1(x,y) = 0$ ορίζει γενικά ένα σύνολο καμπύλων στο επίπεδο xy και η $f_2(x,y) = 0$ ορίζει ένα άλλο σύνολο καμπύλων. Οι πραγματικές λύσεις του συστήματος (8.1) βρίσκονται στις τομές των δύο αυτών συνόλων καμπύλων). Αφού οι συναρτήσεις $f_1(x,y) = 0$ και $f_2(x,y) = 0$ είναι πεπλεγμένες, για την εύρεση της γραφικής απεικονίσεως αυτών εργαζόμεθα ως εξής: Για την $f_1(x,y) = 0$, για παράδειγμα, δίνουμε στο x διαδοχικά ένα σύνολο κατάλληλων, λογικών τιμών $\{x_K\}$, λύνοντας κάθε φορά (δηλαδή, για κάθε συγκεκριμένη τιμή x_K) την εξίσωση

$$f_1(x_K, y) = 0 \ , \tag{8.2}$$

με άγνωστο το y , με κάποια από τις μεθόδους του κεφαλαίου 6 . Έτσι, βρίσκουμε τελικά το σύνολο $\{y_K\}$ των λύσεων της (8.2) για όλες τις τιμές $\{x_K\}$. (Πρέπει να σημειωθεί ότι η (8.2) μπορεί για πολλές από τις τιμές $\{x_K\}$ να έχει περισσότερες από μία λύσεις. Ακόμη, πολλές φορές, οι ρόλοι των x και y εναλλάσσονται). Στη συνέχεια, με την βοήθεια των σημείων (x_K, y_K) κατασκευάζεται η γραφική παράσταση της $f_1(x,y) = 0$. Η ίδια εργασία γίνεται και για την συνάρτηση $f_2(x,y) = 0$.

8.3 Η επαναληπτική μέθοδος

Υποθέτουμε ότι ζητούμε να βρούμε τις πραγματικές λύσεις του (8.1) και ότι (x_0, y_0) είναι μία προσεγγιστική αρχική τιμή (που βρέθηκε γραφικώς) μιας λύσεως του (8.1).

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Κατ' αναλογίαν προς την περίπτωση της μιας εξισώσεως (βλέπε παράγραφο 6.4) γράφουμε το σύστημα (8.1) υπό την ισοδύναμη μορφή:

$$\begin{cases} x = g_1(x,y) \\ y = g_2(x,y) \end{cases} , \tag{8.3}$$

οπότε η επαναληπτική μέθοδος συνίσταται στην εφαρμογή του αλγορίθμου:

$$\begin{aligned} x_{n+1} &= g_1(x_n,y_n) \\ y_{n+1} &= g_2(x_n,y_n) \end{aligned} , \tag{8.4}$$

$n = 0,1,2,\ldots,$ και (x_0 , y_0) η αρχική προσέγγιση. Επειδή είναι καλύτερο να χρησιμοποιούμε τις νέες τιμές μόλις αυτές είναι διαθέσιμες, συχνά αντί του (8.4) χρησιμοποιείται ο αλγόριθμος (Seidel) :

$$\begin{aligned} x_{n+1} &= g_1(x_n,y_n) \\ y_{n+1} &= g_2(x_{n+1} , y_n) \end{aligned} , \quad n = 0,1,2,\ldots . \tag{8.5}$$

Από τον (8.4) (ή (8.5)) υπολογίζονται οι διαδοχικές προσεγγίσεις (επαναλήψεις) (x_n , y_n) $n = 1,2,\ldots$. Εάν ο αλγόριθμος (8.4) συγκλίνει, δηλαδή, εάν υπάρχουν τα όρια

$$\xi = \lim_{n\to\infty} x_n \quad \text{και} \quad \theta = \lim_{n\to\infty} y_n , \tag{8.6}$$

τότε υποθέτοντας ότι οι συναρτήσεις $g_1(x , y)$, $g_2(x , y)$ στην (8.3) είναι συνεχείς, από τις (8.4) έχουμε $\lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} g_1(x_n , y_n)$ και $\lim_{n\to\infty} y_{n+1} = \lim_{n\to\infty} g_2(x_n , y_n)$. Επομένως, $\xi = g_1(\xi , \theta)$ και $\theta = g_2(\xi , \theta)$, δηλαδή, το (ξ , θ) είναι λύση του συστήματος (8.3), άρα και του (8.1).

Η σύγκλιση του αλγορίθμου (8.4) εξαρτάται προφανώς από την εκλογή των $g_1(x , y)$, $g_2(x , y)$ και την αρχική προσέγγιση (x_0 , y_0) . Γενικά, η αρχική προσέγγιση (x_0, y_0) πρέπει να είναι αρκετά καλή, δηλαδή, αρκετά κοντά στην ζητούμενη λύση (πρέπει να σημειωθεί ότι σε πολ-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

λά πρακτικά προβλήματα η προσεγγιστική θέση μιας ζητούμενης λύσεως είναι γνωστή).

Το ακόλουθο θεώρημα παρέχει ικανές συνθήκες για την σύγκλιση του αλγορίθμου (8.4).

<u>Θεώρημα 8.1</u> Υποθέτουμε ότι σε μία κλειστή περιοχή $R\{a \leqq x \leqq b\ ,\ c \leqq y \leqq d\}$ του επιπέδου $x\,y$ υπάρχει μία και μόνον λύση $(\xi\ ,\theta)$ του συστήματος (8.3). Εάν: i) Οι συναρτήσεις $g_1(x\ ,y)$ και $g_2(x\ ,y)$ έχουν συνεχείς μερικές παραγώγους πρώτης τάξεως στην περιοχή R , που ικανοποιούν τις ανισότητες

$$\left|\frac{\partial g_1}{\partial x}\right| + \left|\frac{\partial g_1}{\partial y}\right| \leqq q_1 < 1 \quad ,$$
$$\left|\frac{\partial g_2}{\partial x}\right| + \left|\frac{\partial g_2}{\partial y}\right| \leqq q_2 < 1 \quad , \tag{8.7}$$

στην περιοχή R και ii) Η αρχική προσέγγιση $(x_0\ ,y_0)$ και όλες οι διαδοχικές προσεγγίσεις $(x_n\ ,y_n)$, $n = 1,2,\ldots$, που παράγονται από τον αλγόριθμο (8.4) ανήκουν στην περιοχή R , τότε ο αλγόριθμος (8.4) θα συγκλίνει στην λύση $(\xi\ ,\theta)$ του συστήματος (8.3), δηλαδή, $\lim\limits_{n\to\infty} x_n = \xi$ και $\lim\limits_{n\to\infty} y_n = \theta$.

<u>Απόδειξη</u> Αφού η $(\xi\ ,\theta)$ είναι λύση του (8.3), προφανώς, ισχύουν $\xi = g_1(\xi\ ,\theta)$ και $\theta = g_2(\xi\ ,\theta)$. Αν καλέσουμε

$$\varepsilon_\kappa = x_\kappa - \xi \quad \text{και} \quad E_\kappa = y_\kappa - \theta \quad , \quad \kappa = 0,1,2,\ldots \ , \tag{8.8}$$

τότε χρησιμοποιώντας το θεώρημα Taylor παίρνουμε

$$\begin{aligned}\varepsilon_{\kappa+1} &= x_{\kappa+1} - \xi = g_1(x_\kappa\ ,y_\kappa) - g_1(\xi\ ,\theta) = \\ &= g_1(\xi\ ,\theta) + (x_\kappa - \xi)\frac{\partial g_1}{\partial x} + (y_\kappa - \theta)\frac{\partial g_1}{\partial y} - g_1(\xi\ ,\theta) = \\ &= \varepsilon_\kappa \frac{\partial g_1}{\partial x} + E_\kappa \frac{\partial g_1}{\partial y}\end{aligned} \tag{8.9}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Κατά τον ίδιο τρόπο βρίσκουμε ότι

$$E_{K+1} = \varepsilon_K \frac{\partial g_2}{\partial x} + E_K \frac{\partial g_2}{\partial y} \quad . \tag{8.10}$$

Στις (8.9) και (8.10) οι μερικές παράγωγοι υπολογίζονται σε ένα σημείο $(\hat{\xi}, \hat{\theta})$, όπου το $\hat{\xi}$ βρίσκεται μεταξύ των x_K και ξ και το $\hat{\theta}$ μεταξύ των y_K και θ. Προφανώς, το $(\hat{\xi}, \hat{\theta})$ ανήκει στην περιοχή R. Από τις (8.9) και (8.10), παίρνοντας μέτρα έχουμε

$$\begin{aligned} |\varepsilon_{K+1}| &\leqq |\varepsilon_K| \left|\frac{\partial g_1}{\partial x}\right| + |E_K| \left|\frac{\partial g_1}{\partial y}\right| \\ |E_{K+1}| &\leqq |\varepsilon_K| \left|\frac{\partial g_2}{\partial x}\right| + |E_K| \left|\frac{\partial g_2}{\partial y}\right| \quad . \end{aligned} \tag{8.11}$$

Έστω $S_K = \max\{|\varepsilon_K|, |E_K|\}$ και $Z = \max\{q_1, q_2\}$. Τότε, λόγω των (8.7), από τις (8.11) προκύπτει ότι

$$\begin{aligned} |\varepsilon_{K+1}| &\leqq S_K \left(\left|\frac{\partial g_1}{\partial x}\right| + \left|\frac{\partial g_1}{\partial y}\right| \right) \leqq S_K Z \quad , \\ |E_{K+1}| &\leqq S_K \left(\left|\frac{\partial g_2}{\partial x}\right| + \left|\frac{\partial g_2}{\partial y}\right| \right) \leqq S_K Z \quad . \end{aligned} \tag{8.12}$$

Οι (8.12) συνεπάγονται ότι $S_{K+1} \leqq ZS_K$ και επομένως, αφού όλες οι επαναλήψεις (x_K, y_K) εξ υποθέσεως ανήκουν στην περιοχή R έχουμε

$$\begin{aligned} &S_{K+1} \leqq ZS_K \leqq Z^2 S_{K-1} \leqq \dots \leqq Z^{K+1} S_0 \quad \text{ή} \\ &S_K \leqq Z^K S_0 \quad . \end{aligned} \tag{8.13}$$

Αφού $0 \leqq Z < 1$, έπεται ότι $\lim\limits_{K \to \infty} S_K \leqq \lim\limits_{K \to \infty} Z^K S_0 = 0$, οπότε

$$S_K \underset{K \to \infty}{\longrightarrow} 0 \quad . \tag{8.14}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Η σχέση (8.14) σημαίνει ότι ο αλγόριθμος (8.4) θα συγκλίνει στην λύση (ξ, θ) . ▲

<u>Σημείωση</u> Το θεώρημα 8.1 ισχύει επίσης και στην περίπτωση που οι συνθήκες (8.7) αντικατασταθούν από τις ακόλουθες:

$$\left|\frac{\partial g_1}{\partial x}\right| + \left|\frac{\partial g_2}{\partial x}\right| \leqq q_1 < 1 \quad ,$$
$$\left|\frac{\partial g_1}{\partial y}\right| + \left|\frac{\partial g_2}{\partial y}\right| \leqq q_2 < 1 \quad . \tag{8.15}$$

Το σύστημα (8.1) μπορεί να γραφεί συνοπτικά ως εξής:

$$F(X) = 0 \quad , \tag{8.16}$$

όπου $X = \begin{pmatrix} x \\ y \end{pmatrix}$ και $F(X) = \begin{pmatrix} f_1(X) \\ f_2(X) \end{pmatrix}$ είναι διδιάστατα διανύσματα. Κατ' αναλογίαν, το σύστημα (8.3) γράφεται

$$X = G(X) \quad , \tag{8.17}$$

όπου $G(X) = \begin{pmatrix} g_1(X) \\ g_2(X) \end{pmatrix}$. Η διανυσματική εξίσωση (8.17) είναι επέκταση της εξισώσεως $x = g(x)$, με μία μεταβλητή (βλέπε παράγραφο 6.4). Από την (8.17) προκύπτει ο αλγόριθμος της επαναληπτικής μεθόδου (8.4) σε συνοπτική μορφή:

$$X_{n+1} = G(X_n) \ , \ n = 0,1,2,\ldots \ , \tag{8.18}$$

όπου $X_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ κάποιο αρχικό διάνυσμα (αρχική προσέγγιση μιας λύσεως του (8.16)). Αν $Z = \begin{pmatrix} \xi \\ \theta \end{pmatrix}$ είναι μία λύση του (8.16), τότε, προφανώς,

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$F(Z) = 0 \quad \text{και} \quad Z = G(Z) \quad . \tag{8.19}$$

Αν δοθεί η εξίσωση (8.16), τότε μία εκλογή για την $G(X)$ στην (8.17) μπορεί να είναι η

$$G(X) = X + AF(X) \ ,$$

όπου A είναι ένας ομαλός 2×2 πίνακας. Το παρακάτω θεώρημα που παρατίθεται χωρίς απόδειξη αποτελεί γενίκευση του θεωρήματος 6.6 και είναι γενικώτερο του θεωρήματος 8.1 .

<u>Θεώρημα 8.2</u> Εάν υπάρχει μία κλειστή ορθογώνιος περιοχή R του επιπέδου $x\,y$ στην οποία η $G(X)$ είναι μία απεικόνιση συστολής (contraction mapping) σε σχέση με κάποια από τις γνωστές norms $\|\cdot\|_p$, $p = 1,2,\infty$ (βλέπε κεφάλαιο 9), δηλαδή ισχύουν:

i) $X \in R \Rightarrow G(X) \in R$ και

ii) $\|G(X) - G(Y)\|_p \leqq L\,\|X - Y\|_p$, $\forall\, X, Y \in R$ με $0 \leqq L < 1$,

τότε η εξίσωση (8.17) έχει μία μοναδική λύση (έστω Z) που ανήκει στην περιοχή R και $\forall\, X_0 \in R$ η ακολουθία των διανυσμάτων που παράγεται από τον αλγόριθμο (8.18) συγκλίνει στη λύση Z .

<u>Σημείωση</u> Αν η συνεχής διανυσματική συνάρτηση $G(X)$ έχει συνεχή παράγωγο $G'(X)$ στην περιοχή R , τότε από αντίστοιχο θεώρημα μέσης τιμής για διανυσματικές συναρτήσεις παίρνουμε

$$\|G(X) - G(Y)\|_p \leqq \|G'(\hat{X})\|_p \|X - Y\|_p \ , \quad X, Y \in R \ , \tag{8.20}$$

όπου $\hat{X}$ είναι ένα σημείο που βρίσκεται "μεταξύ" των X και Y ,δηλαδή, $\hat{X} = (1 - \lambda)X + \lambda Y$, για κάποιο $\lambda \in (0,1)$. Σημειώνεται ότι ο υπολογισμός της $\|\cdot\|_2$ δεν είναι εύκολος και γι' αυτό στην πράξη χρησιμοποιούμε τις $\|\cdot\|_p$, $p = 1, \infty$. Στην (8.20) η παράγωγος, γενικά, $G'(X)$ είναι ο γνωστός πίνακας Jacobi των συναρτήσεων g_1, g_2 σε σχέση με τις μεταβλητές x, y, δηλαδή,

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$G'(X) = W(X) = \begin{bmatrix} \dfrac{\partial g_1}{\partial x} & \dfrac{\partial g_1}{\partial y} \\ \\ \dfrac{\partial g_2}{\partial x} & \dfrac{\partial g_2}{\partial y} \end{bmatrix} . \tag{8.21}$$

Από τον πίνακα $G'(X)$ κατασκευάζουμε τον πίνακα:

$$\tilde{G} = \begin{bmatrix} \max\limits_{X \in R} \left|\dfrac{\partial g_1}{\partial x}\right| & \max\limits_{X \in R} \left|\dfrac{\partial g_1}{\partial y}\right| \\ \\ \max\limits_{X \in R} \left|\dfrac{\partial g_2}{\partial x}\right| & \max\limits_{X \in R} \left|\dfrac{\partial g_2}{\partial y}\right| \end{bmatrix} . \tag{8.22}$$

Προφανώς, ισχύει $\|G'(X)\|_p \leqq \|\tilde{G}\|_p$, $p = 1, \infty$, $X \in R$. Επομένως, από την (8.20) έχουμε

$$\|G(X) - G(Y)\|_p \leqq \|\tilde{G}\|_p \|X - Y\|_p , \quad X, Y \in R , \quad p = 1, \infty . \tag{8.23}$$

Από τις (8.20) και (8.23) γίνεται φανερό ότι η πρόταση (ii) στις υποθέσεις του θεωρήματος 8.2 μπορεί να αντικατασταθεί με την πρόταση:

$$\|\tilde{G}\|_p \leqq L < 1 , \quad p = 1, \infty , \tag{8.24}$$

ή με την πρόταση

$$\|G'(X)\|_p \leqq L < 1 , \quad \forall X \in R , \quad p = 1, \infty . \tag{8.25}$$

Για $p = 1$ η (8.25) γίνεται:

$$\max \left\{ \left|\frac{\partial g_1}{\partial x}\right| + \left|\frac{\partial g_2}{\partial x}\right| , \left|\frac{\partial g_1}{\partial y}\right| + \left|\frac{\partial g_2}{\partial y}\right| \right\} \leqq L < 1 , \quad \forall X \in R , \tag{8.26}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ενώ για $p = \infty$ γίνεται

$$\max\left\{\left|\frac{\partial g_1}{\partial x}\right| + \left|\frac{\partial g_1}{\partial y}\right| , \left|\frac{\partial g_2}{\partial x}\right| + \left|\frac{\partial g_2}{\partial y}\right|\right\} \leqq L < 1 , \ \forall X \in R . \qquad (8.27)$$

Η (8.26) είναι ισοδύναμη με τις συνθήκες (8.15), ενώ η (8.27) είναι ισοδύναμη με τις συνθήκες (8.7). Δηλαδή, το θεώρημα 8.1 προκύπτει σαν πόρισμα του γενικώτερου θεωρήματος 8.2 και της σημειώσεως γι' αυτό.

<u>Παράδειγμα 8.1</u> Το σύστημα

$$\begin{cases} x - \sqrt{y} = 7 \\ \sqrt{x} - y = 7 \end{cases} \qquad (8.28)$$

έχει μία μοναδική λύση (ξ , θ) στην περιοχή $R\{1 \leqq x \leqq 10 , 1 \leqq y \leqq 6\}$. Να δειχθεί ότι $\forall (x_0 , y_0) \in R$ ο αλγόριθμος

$$\begin{aligned} x_{n+1} &= 7 + \sqrt{y_n} \\ y_{n+1} &= 7 - \sqrt{x_n} \quad , n = 0,1,2,\ldots \end{aligned} \qquad (8.29)$$

συγκλίνει στην λύση (ξ , θ) . Να ληφθεί $(x_0 , y_0) = (1 , 1)$ και να βρεθεί η λύση (ξ , θ).

<u>Λύση</u> Έχουμε $g_1(x , y) = 7 + \sqrt{y}$, $g_2(x , y) = 7 - \sqrt{x}$, $\frac{\partial g_1}{\partial x} = 0$, $\frac{\partial g_1}{\partial y} = \frac{1}{2\sqrt{y}}$, $\frac{\partial g_2}{\partial x} = -\frac{1}{2\sqrt{x}}$, $\frac{\partial g_2}{\partial y} = 0$. Είναι εύκολο να δειχθεί ότι $\forall$ $(x_0 , y_0) \in R$ όλες οι επαναλήψεις (x_n , y_n) , $n = 1,2,\ldots$, που παράγονται από τον αλγόριθμο (8.29) ανήκουν στην περιοχή R . Επί πλέον, $\forall$ $(x , y) \in R$ είναι $\left|\frac{\partial g_1}{\partial x}\right| + \left|\frac{\partial g_1}{\partial y}\right| \leqq \frac{1}{2} < 1$ και $\left|\frac{\partial g_2}{\partial x}\right| + \left|\frac{\partial g_2}{\partial y}\right| \leqq \frac{1}{2} < 1$.

Έτσι, σύμφωνα με το θεώρημα 8.1 $\forall$ $(x_0,y_0) \in R$ ο αλγόριθμος (8.29) συγκλίνει στην λύση (ξ , θ) του (8.28). Λαμβάνοντας $(x_0,y_0) = (1 , 1)$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

από τις (8.29) έχουμε $x_1 = 8$, $y_1 = 6$, $x_2 \approx 9.4495$, $y_2 \approx 4.17157$, $x_3 \approx 9.04244$, $y_3 \approx 3.92600$, $x_4 \approx 8.98141$, $y_4 \approx 3.99293$, $x_5 \approx$ ≈ 8.99823 , $y_5 \approx 4.00301$, $x_6 \approx 9.00075$, $y_6 \approx 4.00030$ κ.λ.π. . Είναι φανερό ότι $(\xi , \theta) = (9 , 4)$.

8.4 Η γενικευμένη μέθοδος Newton-Raphson ή μέθοδος Newton

Έστω (x_n , y_n) μία προσεγγιστική τιμή μιας λύσεως του συστήματος (8.1), όπου οι συναρτήσεις f_1, f_2 έχουν συνεχείς παραγώγους. Για την εύρεση της επόμενης προσεγγίσεως (x_{n+1} , y_{n+1}) της λύσεως θέτουμε

$$x_{n+1} = x_n + h_n , \quad y_{n+1} = y_n + k_n \tag{8.30}$$

και χρησιμοποιώντας τον τύπο του Taylor παίρνουμε

$$f_1(x_{n+1},y_{n+1}) = f_1(x_n,y_n) + h_n \frac{\partial f_1}{\partial x}(x_n,y_n) + k_n \frac{\partial f_1}{\partial y}(x_n,y_n) + \dots ,$$

$$f_2(x_{n+1},y_{n+1}) = f_2(x_n,y_n) + h_n \frac{\partial f_2}{\partial x}(x_n,y_n) + k_n \frac{\partial f_2}{\partial y}(x_n,y_n) + \dots .$$

Απαιτούμε τώρα να ισχύουν

$$\begin{cases} f_1(x_{n+1},y_{n+1}) \approx f_1(x_n,y_n) + h_n \dfrac{\partial f_1}{\partial x}(x_n,y_n) + k_n \dfrac{\partial f_1}{\partial y}(x_n,y_n) = 0 \\ f_2(x_{n+1},y_{n+1}) \approx f_2(x_n,y_n) + h_n \dfrac{\partial f_2}{\partial x}(x_n,y_n) + k_n \dfrac{\partial f_2}{\partial y}(x_n,y_n) = 0. \end{cases} \tag{8.31}$$

Το σύστημα των (8.31) είναι γραμμικό ως προς h_n και k_n . Εάν η Jacobian των f_1 και f_2 , δηλαδή, η ορίζουσα

$$J(x_n , y_n) = \begin{vmatrix} \dfrac{\partial f_1}{\partial x}(x_n , y_n) & \dfrac{\partial f_1}{\partial y}(x_n , y_n) \\ \dfrac{\partial f_2}{\partial x}(x_n , y_n) & \dfrac{\partial f_2}{\partial y}(x_n , y_n) \end{vmatrix}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

είναι $\neq 0$, τότε από το σύστημα (8.31) λύνοντας ως προς h_n και k_n έχουμε

$$h_n = -\frac{1}{J(x_n,y_n)} \begin{vmatrix} f_1(x_n, y_n) & \frac{\partial f_1}{\partial y}(x_n, y_n) \\ f_2(x_n, y_n) & \frac{\partial f_2}{\partial y}(x_n, y_n) \end{vmatrix} \tag{8.32}$$

και

$$k_n = -\frac{1}{J(x_n,y_n)} \begin{vmatrix} \frac{\partial f_1}{\partial x}(x_n, y_n) & f_1(x_n, y_n) \\ \frac{\partial f_2}{\partial x}(x_n, y_n) & f_2(x_n, y_n) \end{vmatrix} . \tag{8.33}$$

Επομένως, από τις (8.30) έχουμε

$$x_{n+1} = x_n - \frac{f_1 \frac{\partial f_2}{\partial y} - f_2 \frac{\partial f_1}{\partial y}}{\frac{\partial f_1}{\partial x}\frac{\partial f_2}{\partial y} - \frac{\partial f_1}{\partial y}\frac{\partial f_2}{\partial x}} , \tag{8.34}$$

$$y_{n+1} = y_n - \frac{f_2 \frac{\partial f_1}{\partial x} - f_1 \frac{\partial f_2}{\partial x}}{\frac{\partial f_1}{\partial x}\frac{\partial f_2}{\partial y} - \frac{\partial f_1}{\partial y}\frac{\partial f_2}{\partial x}} , \tag{8.35}$$

όπου οι τιμές των f_1 , f_2 και των μερικών παραγώγων τους υπολογίζονται στο σημείο (x_n, y_n) . Οι σχέσεις (8.34), (8.35) , όπου $n = 0,1,2,...$ και (x_0, y_0) η αρχική προσέγγιση, αποτελούν τον αλγόριθμο της μεθόδου Newton για την επίλυση του (8.1) .

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Αν χρησιμοποιήσουμε τους συμβολισμούς με διανύσματα (βλέπε (8.16)) και έχοντας υπ' όψιν αντίστοιχη της (8.21), τότε η μέθοδος Newton υπό διανυσματική μορφή είναι

$$X_{n+1} = X_n - \left[F'(X_n)\right]^{-1} F(X_n) \ , \ n = 0,1,2,\ldots \tag{8.36}$$

και X_0 η αρχική προσέγγιση, με την προυπόθεση ότι η ορίζουσα του πίνακα $F'(X_n)$ είναι $\neq 0$. Στην (8.36) είναι

$$X_n = \begin{pmatrix} x_n \\ y_n \end{pmatrix} \ , \ F(X_n) = \begin{pmatrix} f_1(X_n) \\ f_2(X_n) \end{pmatrix} \ , \ \text{και}$$

$$F'(X_n) = \begin{bmatrix} \dfrac{\partial f_1}{\partial x}(X_n) & \dfrac{\partial f_1}{\partial y}(X_n) \\ \dfrac{\partial f_2}{\partial x}(X_n) & \dfrac{\partial f_2}{\partial y}(X_n) \end{bmatrix} . \tag{8.37}$$

Η μέθοδος Newton συγκλίνει και μάλιστα γρήγορα εάν η αρχική προσέγγιση είναι πολύ καλή. Συγκεκριμένα, μπορεί να δειχθεί ότι, εάν η $F(X)$ έχει συνεχείς μερικές παραγώγους δευτέρας τάξεως, ο πίνακας Jacobi $F'(X)$ είναι ομαλός στην ζητούμενη λύση Z και η μέθοδος συγκλίνει, τότε η σύγκλιση είναι τετραγωνική, όπως και στην μονοδιάστατη περίπτωση. Στην πράξη, σε πολλές περιπτώσεις αρχίζουμε με κάποια πρόχειρη,λογική αρχική προσέγγιση και εάν μετά την εκτέλεση ορισμένων επαναλήψεων δεν παρατηρηθεί σύγκλιση, τότε χρησιμοποιούμε κάποια άλλη αρχική προσέγγιση κ.ο.κ., μέχρις ότου επιτευχθεί σύγκλιση.

Πρέπει να σημειωθεί ότι στον αλγόριθμο (8.36), για την εύρεση των διαδοχικών προσεγγίσεων, χρειάζεται να υπολογίζεται σε κάθε βήμα ο αντίστροφος ενός πίνακα. Επειδή η αντιστροφή πίνακα απαιτεί περισσότερες πράξεις από την επίλυση ενός γραμμικού συστήματος (βλέπε κεφ. 10), γράφουμε τον αλγόριθμο (8.36) υπό την μορφή

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$F'(X_n)\,(X_n - X_{n+1}) = F(X_n)\ ,\ n = 0,1,2,\dots\ .$$

θέτοντας

$$X_n - X_{n+1} \equiv Z_{n+1}\ ,\ n = 0,1,2,\dots\ , \tag{8.38}$$

παίρνουμε

$$F'(X_n)Z_{n+1} = F(X_n)\ ,\ n = 0,1,2,\dots\ . \tag{8.39}$$

Έτσι, χρησιμοποιώντας τον αλγόριθμο (8.39), λύνουμε κάθε φορά ένα γραμμικό σύστημα με άγνωστο το διάνυσμα Z_{n+1}, με κάποια από τις μεθόδους του κεφαλαίου 10 (π.χ. απαλοιφή Gauss) και από την (8.38) βρίσκουμε το $X_{n+1} = X_n - Z_{n+1}$. Η εύρεση των επαναλήψεων σταματά όταν για δύο διαδοχικά διανύσματα X_n, X_{n+1} και για κάποια norm π.χ. την $\|\cdot\|_\infty$ ικανοποιηθεί η συνθήκη

$$\|Z_{n+1}\|_\infty = \|X_{n+1} - X_n\|_\infty \leq \varepsilon\ ,$$

όπου ε είναι ένας μικρός προκαθορισμένος αριθμός.

Μία τροποποίηση της μεθόδου Newton, η οποία μειώνει σημαντικά τον συνολικό αριθμό των υπολογισμών, συνίσταται στην χρησιμοποίηση σταθερού πίνακα στα διαδοχικά γραμμικά συστήματα (8.39) και συγκεκριμένα του πίνακα $F'(X_0)$, εφ' όσον βέβαια το X_0 είναι αρκετά κοντά σε μία λύση, οπότε $F'(X_n) \approx F'(X_0)$. Στην περίπτωση αυτή συνιστάται η εύρεση του αντιστρόφου του $F'(X_0)$ και η χρησιμοποίηση του τροποποιημένου αλγορίθμου υπό την μορφή (8.36). Τονίζεται ακόμη, ότι η σύγκλιση (εάν υπάρχει) της τροποποιημένης μεθόδου Newton είναι πιο αργή.

<u>Παράδειγμα 8.2</u> Να εφαρμοσθεί η μέθοδος Newton στο σύστημα

$$\begin{cases} x^3 - 3xy^2 + 1 = 0 \\ 3x^2y - y^3 = 0 \end{cases}\ ,\ \text{με}\ (x_0,y_0) = (1,1)\ \text{για την εύρεση της}$$

(x_1,y_1).

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

<u>Λύση</u> Έχουμε $f_1(x,y) = x^3 - 3xy^2 + 1$, $f_2(x,y) = 3x^2y - y^3$, οπότε $\frac{\partial f_1}{\partial x} = 3x^2 - 3y^2$, $\frac{\partial f_1}{\partial y} = -6xy$, $\frac{\partial f_2}{\partial x} = 6xy$, $\frac{\partial f_2}{\partial y} = 3x^2 - 3y^2$. Έτσι, από τις (8.34), (8.35) παίρνουμε για $(x_0,y_0) = (1,1)$, $x_1 = \frac{2}{3}$ και $y_1 = \frac{5}{6}$, δηλαδή $(x_1,y_1) \approx (0.667 , 0.833)$. Σημειώνεται ότι οι ακριβείς λύσεις του δοθέντος συστήματος είναι οι $(-1,0)$, $(\frac{1}{2} , \pm \frac{\sqrt{3}}{2}) \approx$ $\approx(0.5 , \pm 0.86602)$. ▲

8.5 Η μέθοδος της κλίσεως (Gradient Steepest descent)

Θεωρούμε πάλι το σύστημα (8.1) υπό την μορφή (8.16):

$$F(X) = 0 \quad . \tag{8.40}$$

Το πρόβλημα της ευρέσεως των λύσεων του συστήματος (8.40) μπορεί να διατυπωθεί σαν το πρόβλημα της ελαχιστοποιήσεως της συναρτήσεως

$$f(X) = \frac{1}{2} \|F(X)\|_2^2 = \frac{1}{2} \left[f_1(X)\right]^2 + \frac{1}{2} \left[f_2(X)\right]^2 \quad . \tag{8.41}$$

Η ελαχιστοποίηση της $f(X)$ ανήκει στην κατηγορία των μή γραμμικών προβλημάτων βελτιστοποιήσεως πολλών (γενικά) μεταβλητών (μη γραμμικό πρόβλημα ελαχίστων τετραγώνων) και οι επαναληπτικές μέθοδοι που χρησιμοποιούνται για την λύση αυτών καλούνται μέθοδοι μη γραμμικού προγραμματισμού.

Η συνάρτηση $f(X) = f(x,y)$ είναι μη αρνητική . Είναι προφανές ότι κάθε λύση του (8.40) μηδενίζει την $f(X)$. Αντιστρόφως, κάθε σημείο στο οποίο η $f(X)$ μηδενίζεται (επιτυγχάνεται δηλαδή το ολικό ελάχιστο της $f(X)$) είναι λύση του (8.40) .

Υποθέτοντας ότι το σύστημα (8.40) έχει μία λύση σε μία περιοχή R, προσπαθούμε να ελαχιστοποιήσουμε την $f(X)$ στην περιοχή R . Η διαδικασία που εφαρμόζεται είναι η εξής: Έστω (x_n,y_n) ένα σημείο στην περιοχή R . Ζητούμε τώρα να βρούμε ένα νέο σημείο (x_{n+1},y_{n+1}) , τέτοιο

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ώστε

$$f(x_{n+1},y_{n+1}) < f(x_n,y_n) \ . \tag{8.42}$$

Η εξίσωση

$$f(x,y) = f(x_n,y_n) = \text{σταθ.} \tag{8.43}$$

παριστά στο επίπεδο xy μία καμπύλη, η οποία διέρχεται από το σημείο (x_n,y_n) . Είναι γνωστό από την Μαθηματική Ανάλυση ότι στο οποιοδήποτε σημείο (x_n,y_n) η κατεύθυνση στην οποία η $f(x,y)$ ελαττώνεται πιο γρήγορα είναι η $-\nabla f(x_n,y_n)$, όπου, γενικά, το διάνυσμα

$$\nabla f(x,y) = \begin{pmatrix} \frac{\partial f}{\partial x}(x,y) \\ \\ \frac{\partial f}{\partial y}(x,y) \end{pmatrix} \tag{8.44}$$

είναι η καλούμενη κλίσις της $f(x,y)$ (συμβολίζεται επίσης σαν grad $f(x,y)$. Το σύμβολο ∇ καλείται nabla ή del). Το διάνυσμα – κλίσις $\nabla f(x_n,y_n)$ (και επομένως και το $-\nabla f(x_n,y_n)$) έχει επίσης την ιδιότητα ότι στο σημείο (x_n,y_n) είναι κάθετο στην καμπύλη (8.43) . Υποθέτοντας ότι η $f(x,y)$ έχει μερικές παραγώγους πρώτης τάξεως από την (8.41) έχουμε

$$\begin{aligned} \frac{\partial f}{\partial x} &= f_1 \frac{\partial f_1}{\partial x} + f_2 \frac{\partial f_2}{\partial x} \ , \\ \frac{\partial f}{\partial y} &= f_1 \frac{\partial f_1}{\partial y} + f_2 \frac{\partial f_2}{\partial y} \ . \end{aligned} \tag{8.45}$$

Με την προυπόθεση, επομένως, ότι $\nabla f(x_n,y_n) \neq 0$, είναι δυνατόν να βρούμε θετικό αριθμό t_n , τέτοιο ώστε για το σημείο

$$(x_{n+1},y_{n+1}) = (x_n - t_n \frac{\partial f}{\partial x}(x_n,y_n),\ y_n - t_n \frac{\partial f}{\partial y}(x_n,y_n)\ )$$

να ισχύει η (8.42). Στην μέθοδο της κλίσεως το t_n υπολο-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

γίζεται έτσι ώστε να ελαχιστοποιεί την συνάρτηση ως προς t:

$$f(x_n - t\,\frac{\partial f}{\partial x}(x_n,y_n)\ ,\ \ y_n - t\,\frac{\partial f}{\partial x}(x_n,y_n)\ )\ , \tag{8.46}$$

δηλαδή, υπολογίζεται σαν η μικρότερη θετική ρίζα της εξισώσεως ως προς t:

$$\frac{d}{dt}\,f(x_n - t\,\frac{\partial f}{\partial x}(x_n,y_n)\ ,\ \ y_n - t\,\frac{\partial f}{\partial x}(x_n,y_n)\ ) = 0\ . \tag{8.47}$$

Επομένως, ο αλγόριθμος της μεθόδου της κλίσεως είναι

$$\begin{cases} x_{n+1} = x_n - t_n\dfrac{\partial f}{\partial x}(x_n,y_n) \\[2ex] y_{n+1} = y_n - t_n\dfrac{\partial f}{\partial y}(x_n,y_n)\ , \end{cases} \tag{8.48}$$

όπου $n = 0,1,2,\ldots$ και (x_0,y_0) η αρχική προσέγγιση. Η εύρεση των t_n , $n = 0,1,2,\ldots$ γίνεται με τον τρόπο που περιγράφηκε προηγουμένως.

Ο αλγόριθμος (8.48) υπό διανυσματική μορφή γράφεται

$$X_{n+1} = X_n - t_n\nabla f(X_n)\quad ,\ n = 0,1,2,\ldots\ . \tag{8.49}$$

Λόγω των (8.45) είναι

$$\nabla f(X_n) = \left[F'(X_n)\right]^T F(X_n)\ , \tag{8.50}$$

όπου $F(X_n)$, $F'(X_n)$ δίνονται από την (8.37) και $\left[F'(X_n)\right]^T$ ο ανάστροφος του πίνακα $F'(X_n)$. Άρα,η (8.49) γίνεται

$$X_{n+1} = X_n - t_n\left[F'(X_n)\right]^T F(X_n)\ ,\ n = 0,1,2,\ldots\ . \tag{8.51}$$

Είναι δυνατόν, η μέθοδος της κλίσεως να συγκλίνει σε σημεία στάσεως (π.χ. τοπικό ελάχιστο) στα οποία έχουμε $\nabla f(X) = 0$ ενώ $F(X) \neq 0$, δηλαδή σε σημεία όπου ο πίνακας $F'(X)$ είναι μη ομαλός (λόγω της (8.50)) Για τον λόγο αυτό,πρέπει να υπολογίζεται η τιμή της $f(x,y)$ στο σημείο που συγκλίνουν τελικά οι διαδοχικές επαναλήψεις (x_n^*,y_n) του αλγορίθμου

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

(8.48), με σκοπό να διαπιστώσουμε αν βρήκαμε ή όχι μία λύση του (8.40). Πρέπει να σημειωθεί ότι, συνήθως, στην πράξη αποφεύγεται η λύση της εξισώσεως (8.47) με την χρησιμοποίηση προσεγγιστικών τιμών για το t_n. Μπορεί να δειχθεί ότι μια τέτοια προσεγγιστική τιμή είναι η

$$t_n = \frac{f(x_n,y_n)}{\left[\frac{\partial f}{\partial x}(x_n,y_n)\right]^2 + \left[\frac{\partial f}{\partial y}(x_n,y_n)\right]^2} = \frac{f(x_n,y_n)}{\|\nabla f(x_n,y_n)\|_2^2} \quad . \qquad (8.52)$$

Επίσης, πολλές φορές η gradient μέθοδος χρησιμοποιείται για την εύρεση μιας καλής αρχικής προσεγγίσεως, έτσι ώστε η μέθοδος Newton της προηγούμενης παραγράφου να μπορεί να εφαρμοσθεί αποτελεσματικά στη συνέχεια.

<u>Παράδειγμα 8.3</u> Να εφαρμοσθεί η μέθοδος της κλίσεως στο σύστημα

$$\begin{cases} x^2 - y - 2 = 0 \\ x - y = 0 \end{cases} \quad \text{με} \quad (x_0,y_0) = (0,0) \quad .$$

<u>Λύση</u> Είναι $f_1(x,y) = x^2 - y - 2$, $f_2(x,y) = x - y$, $f(x,y) = \frac{1}{2}\left[(x^2 - y - 2)^2 + (x - y)^2\right]$, $\frac{\partial f}{\partial x} = (x^2 - y - 2)2x + x - y$, $\frac{\partial f}{\partial y} = -(x^2 - y - 2) - (x - y) = -x^2 + 2y - x + 2$ και

$$\frac{\partial f}{\partial x}(x_0,y_0) = 0 \ , \quad \frac{\partial f}{\partial y}(x_0,y_0) = 2 \ , \text{ οπότε } \nabla f(x_0,y_0) \neq 0 \ .$$

Για την εύρεση του t_0 σχηματίζουμε την συνάρτηση (8.46), δηλαδή, την

$$h_0(t) = f(x_0 - t\cdot 0, y_0 - 2t) = f(0, -2t) =$$
$$= \frac{1}{2}\left[(2t - 2)^2 + (2t)^2\right] = 2(2t^2 - 2t + 1) \ .$$

Από την $\frac{d}{dt}h_0(t) = 8t - 4 = 0$ προκύπτει $t = \frac{1}{2}$. Άρα, είναι $t_0 = \frac{1}{2}$. Από τις (8.48) παίρνουμε $x_1 = 0$, $y_1 = -1$, δηλαδή, $(x_1,y_1) = (0,-1)$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Επαναλαμβάνοντας την διαδικασία, έχουμε $\frac{\partial f}{\partial x}(x_1,y_1) = 1$, $\frac{\partial f}{\partial y}(x_1,y_1) = 0$, $\nabla f(x_1,y_1) \neq 0$, $h_1(t) = f(-t,-1) = \frac{1}{2}\left[(t^2+1-2)^2 + (1-t)^2\right] = \frac{1}{2}\left[t^4 - t^2 - 2t + 2\right]$. Από την $h_1'(t) = (t-1)(2t^2+2t+1) = 0$ προκύπτει $t = 1$, οπότε $t_1 = 1$. Έτσι, βρίσκουμε $x_2 = -1$, $y_2 = -1$ ή $(x_2,y_2) = (-1,-1)$. Επίσης, είναι $f(x_0,y_0) = f(0,0) = 2$, $f(x_1,y_1) = f(0,-1) = 1$, $f(x_2,y_2) = f(-1,-1) = 0$. Άρα μία λύση του συστήματος είναι η $(-1,-1)$ (η άλλη λύση είναι η $(2,2)$). ▲

ΑΣΚΗΣΕΙΣ

1. Το σύστημα $\begin{cases} x^2 + y^2 = 1 \\ 10xy = 1 \end{cases}$ έχει μία μοναδική λύση (ξ,θ) στην περιοχή $R\{0.4 \leq x \leq 1,\ 0 \leq y \leq 0.7\}$. Να εξετασθεί ως προς την σύγκλιση ο αλγόριθμος $\begin{cases} x_{n+1} = \sqrt{1 - y_n^2} \\ y_{n+1} = \frac{1}{10x_n} \end{cases}$, $n = 0,1,2,\ldots,$ με $(x_0,y_0) \in R$.

2. Να βρεθεί μία λύση του συστήματος $\begin{cases} 2x^2 + y^2 - 5 = 0 \\ x + 2y - 3 = 0 \end{cases}$, εφαρμόζοντας την επαναληπτική μέθοδο: $\begin{cases} x_{n+1} = \left(\frac{5 - y_n^2}{2}\right)^{\frac{1}{2}} \\ y_{n+1} = \frac{1}{2}(3 - x_n) \end{cases}$, $n = 0,1,2,\ldots,$ με $(x_0,y_0) = (1.5, 1)$. Να εξετασθεί η σύγκλιση του αλγοοίθμου στην περιοχή $R\{1 \leq x \leq 2,\ \frac{1}{2} \leq y \leq 1.5\}$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

3. Να βρεθεί μία λύση του συστήματος $\begin{cases} x + \frac{1}{4} y^2 = \frac{1}{16} \\ y + \frac{1}{3} \sin x = \frac{1}{2} \end{cases}$, εφαρμόζοντας

την επαναληπτική μέθοδο $\begin{cases} x_{n+1} = \frac{1}{16} - \frac{1}{4} y_n^2 \\ y_{n+1} = \frac{1}{2} - \frac{1}{3} \sin x_n \end{cases}$, $n = 0,1,2,\ldots$ με

$(x_0, y_0) = (0,0)$.

4. Να εφαρμοσθεί η μέθοδος Newton στο γραμμικό σύστημα $Ax = b$, όπου $\det(A) \neq 0$.

5. Να βρεθεί ο αλγόριθμος της μεθόδου Newton για το σύστημα

$$\begin{cases} xy - N = 0 \\ x - y = 0 \end{cases}, \quad N > 0 .$$

6. Να εφαρμοσθεί η μέθοδος Newton στα ακόλουθα συστήματα:

i) $\begin{cases} 2x^3 - y^3 - 1 = 0 \\ xy^3 - y - 4 = 0 \end{cases}$, με $(x_0, y_0) = (1.2,\ 1.7)$

ii) $\begin{cases} x^2y^2 - 2x^3 - 5y^3 + 10 = 0 \\ x^4 - 8y + 1 = 0 \end{cases}$, με $(x_0, y_0) = (-2,\ 2)$

iii) $\begin{cases} x^2 - xy + 20 = 0 \\ y^2 - 2xy + 10 = 0 \end{cases}$, με $(x_0, y_0) = (5,9)$

iv) $\begin{cases} x^2 + y^2 - 6x + 2y + 3 = 0 \\ x^3 + y^3 - 1 = 0 \end{cases}$, με $(x_0, y_0) = (1,1)$.

7. Υποθέτουμε ότι ζητούμε να βρούμε τις ρίζες μιας αναλυτικής συναρτήσεως $f(z) = u(x,y) + iv(x,y)$, $z = x + iy$, λύνοντας το σύστημα (βλέπε (6.48))

$$\begin{cases} u(x,y) = 0 \\ v(x,y) = 0 \end{cases} . \qquad (1)$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Χρησιμοποιώντας τις καλούμενες Cauchy – Riemann εξισώσεις $\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}$, $\frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$, να δειχθεί ότι ο αλγόριθμος της μεθόδου Newton για το σύστημα (1) γράφεται

$$x_{n+1} = g_1(x_n,y_n) \ , \quad y_{n+1} = g_2(x_n,y_n) \ , \ n = 0,1,2,\dots \ ,$$

όπου $g_1(x,y) = x - \dfrac{\frac{\partial |f|^2}{\partial x}}{2\|\nabla u\|_2^2}$, $g_2(x,y) = y - \dfrac{\frac{\partial |f|^2}{\partial y}}{2\|\nabla u\|_2^2}$.

8. Να εφαρμοσθεί μία φορά η μέθοδος της κλίσεως στο σύστημα

$$\begin{cases} 3x + 2y = 0 \\ 5x + 4y = -2 \end{cases} , \text{ με } (x_0,y_0) = (0,0) \ .$$

9. Να εφαρμοσθεί η μέθοδος της κλίσεως στο σύστημα

$$\begin{cases} e^{xy} - 3x - 2y + 1 = 0 \\ \ln(5 + x + y) - xy = 0 \end{cases} , \text{ με } (x_0,y_0) = (1,1) \text{ και } t_0 = \frac{1}{2} ,$$

για την εύρεση της (x_1,y_1) .

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

9

ΣΤΟΙΧΕΙΑ ΘΕΩΡΙΑΣ ΠΙΝΑΚΩΝ ΚΑΙ ΓΡΑΜΜΙΚΗΣ ΑΛΓΕΒΡΑΣ

9.1 Εισαγωγή

Σκοπός του κεφαλαίου αυτού είναι να δοθούν οι βασικές έννοιες και ορισμένα χρήσιμα συμπεράσματα από την θεωρία πινάκων και την γραμμική Άλγεβρα, που είναι απαραίτητα για την μελέτη των διαφόρων αριθμητικών μεθόδων, που παρουσιάζονται σε επόμενα κεφάλαια. Συγκεκριμένα, το περιεχόμενο του παρόντος κεφαλαίου είναι ιδιαίτερα χρήσιμο στην αριθμητική επίλυση γραμμικών συστημάτων και στην αριθμητική εύρεση των ιδιοτιμών και ιδιοδιανυσμάτων ενός πίνακα.

Σε πολλά από τα θεωρήματα του κεφαλαίου αυτού οι αποδείξεις παραλείπονται. Αυτές μπορούν να βρεθούν σε πολλά βασικά βιβλία γραμμικής Άλγεβρας.

9.2 Διανύσματα και πίνακες

Συμβολίζουμε με $\mathbb{C}^{n,n}$ το σύνολο όλων των τετραγωνικών πινάκων τάξεως n με στοιχεία πραγματικά ή μιγαδικά. Εάν

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

είναι ένα στοιχείο του $\mathbb{C}^{n,n}$, τότε γράφουμε

$$A = (a_{ij}) \ , \quad i,j = 1(1)n \ . \tag{9.1}$$

Πρέπει να σημειωθεί ότι θα περιοριστούμε στην χρησιμοποίηση πινάκων που ανήκουν στο σύνολο $\mathbb{C}^{n,n}$, δηλαδή, $n \times n$ πινάκων, εκτός εάν διαφορετικά καθορίζεται.

Καλούμε διάνυσμα στήλη ή απλώς διάνυσμα ένα $n \times 1$ πίνακα. Θα συμβολίζουμε με $\mathbb{C}^n$ το σύνολο όλων των $n \times 1$ πινάκων με στοιχεία πραγματικούς ή μιγαδικούς αριθμούς. Εάν

$$v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

είναι ένα στοιχείο του $\mathbb{C}^n$, τότε γράφουμε εν συντομία

$$v = (v_i) \ , \quad i = 1(1)n \ . \tag{9.2}$$

Μία βάση του $\mathbb{C}^n$ είναι ένα σύνολο n γραμμικώς ανεξαρτήτων διανυσμάτων του . Εάν τα διανύσματα $v^{(i)}$, $i = 1(1)n$ αποτελούν μία βάση του $\mathbb{C}^n$, τότε κάθε διάνυσμα $v \in \mathbb{C}^n$ μπορεί να εκφραστεί μονοσήμαντα σαν γραμμικός συνδυασμός των διανυσμάτων της βάσεως, δηλαδή, υπάρχουν μονοσήμαντα ορισμένοι αριθμοί c_i , $i = 1(1)n$, έτσι ώ-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ώστε

$$v = \sum_{i=1}^{n} c_i v^{(i)} \quad .$$

Θα συμβολίζουμε με v^T το ανάστροφο του διανύσματος v , δηλαδή,

$$v^T = [v_1, v_2, \ldots, v_n] \tag{9.3}$$

με v^* το συζυγές (μιγαδικό) του v και με v^H το συζυγές ανάστροφο του v . Ομοίως, με A^T , A^* , A^H συμβολίζουμε αντίστοιχα τον ανάστροφο, τον συζυγή (μιγαδικό) και τον συζυγή ανάστροφο του A . Ο ταυτοτικός πίνακας θα συμβολίζεται με I , δηλαδή, $I = (\delta_{ij})$, $i,j = 1(1)n$, όπου $\delta_{ij} = \begin{cases} 0 , & i \neq j \\ 1 , & i = j \end{cases}$. Ο μηδενικός πίνακας έχει όλα τα στοιχεία του ίσα με 0 και θα συμβολίζεται με το ίδιο σύμβολο 0 . Από το κείμενο θα γίνεται φανερό κάθε φορά αν το 0 παριστά τον μηδενικό πίνακα ή το μηδενικό διάνυσμα ή τον αριθμό μηδέν.

Ένας πίνακας $D = (d_{ij})$, $i,j = 1(1)n$ καλείται διαγώνιος εάν $d_{ij} = 0$ για $i \neq j$.

Το γινόμενο AB των πινάκων $A = (a_{ij})$, $i,j = 1(1)n$ και $B = (b_{ij})$, $i,j = 1(1)n$ είναι ένας πίνακας $C = (c_{ij})$, $i,j = 1(1)n$, όπου

$$c_{ij} = \sum_{\kappa=1}^{n} a_{i\kappa} b_{\kappa j} , \quad i,j = 1(1)n . \tag{9.4}$$

Το γινόμενο διαγωνίων πινάκων είναι διαγώνιος πίνακας.

<u>Ορισμός 9.1</u> Ο $A = (a_{ij})$, $i,j = 1(1)n$ καλείται κάτω τριγωνικός (αυστηρά κάτω τριγωνικός) εάν $a_{ij} = 0$ για $i < j$ $(i \leqq j)$. Εάν $a_{ij} = 0$ για $i > j$ $(i \geqq j)$, τότε ο A καλείται άνω τριγωνικός (αυστηρά άνω τριγωνικός).

<u>Θεώρημα 9.1</u> Το γινόμενο κάτω (άνω) τριγωνικών πινάκων είναι κά-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

τω (άνω) τριγωνικός πίνακας.

<u>Θεώρημα 9.2</u> Εάν $A, B \in \mathbb{C}^{n,n}$, τότε ισχύουν τα εξής:

i) $(A^H)^H = (A^T)^T = (A^*)^* = A$, $A^H = (A^*)^T = (A^T)^*$

ii) $(A + B)^T = A^T + B^T$, $(AB)^T = B^T A^T$

iii) $(A + B)^* = A^* + B^*$, $(A + B)^H = A^H + B^H$, $(AB)^* = A^* B^*$,

$(AB)^H = B^H A^H$, $(\kappa A)^H = \kappa^* A^H$, $\kappa \in \mathbb{C}$. ▲

Εάν ο πίνακας A είναι πραγματικός (έχει πραγματικά στοιχεία) τότε $A^* = A$ και $A^H = A^T$.

Ο πίνακας A καλείται συμμετρικός (symmetric) εάν $A^T = A$.

Ο πίνακας A καλείται Ερμιτιανός (Hermitian) εάν $A^H = A$.

Ο πίνακας A καλείται αντισυμμετρικός (skew – symmetric) εάν $A^T = -A$.

Ο πίνακας A καλείται αντι-Ερμιτιανός (skew – Hermitian) εάν $A^H = -A$.

Ο πίνακας A καλείται κανονικός (normal) εάν $AA^H = A^H A$.

Ο πίνακας A καλείται ομαλός (nonsingular) εάν υπάρχει ένας πίνακας J τέτοιος ώστε $AJ = JA = I$. Εάν υπάρχει ένας τέτοιος J, τότε είναι μοναδικός και καλείται αντίστροφος του A , συμβολίζεται δε με A^{-1} .

Ο πίνακας A καλείται ορθογώνιος (orthogonal) εάν $A^{-1} = A^T$.

Ο πίνακας A καλείται unitary εάν $A^{-1} = A^H$.

<u>Θεώρημα 9.3</u> Εάν A, B είναι ομαλοί πίνακες, τότε ισχύουν

i) $(A^T)^{-1} = (A^{-1})^T$, $(A^H)^{-1} = (A^{-1})^H$, $(A^*)^{-1} = (A^{-1})^*$, $(A^{-1})^{-1} = A$.

ii) $A^{-K} = (A^{-1})^K$, κ θετικός ακέραιος, $(\lambda A)^{-1} = \frac{1}{\lambda} A^{-1}$, $\lambda =$ σταθ. $\neq 0$.

iii) ο AB είναι ομαλός και $(AB)^{-1} = B^{-1} A^{-1}$.

<u>Θεώρημα 9.4</u> i) Οι διαγώνιοι, οι Ερμιτιανοί, οι αντι-Ερμιτιανοί και οι unitary πίνακες είναι normal πίνακες.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ii) Οι συμμετρικοί, οι αντισυμμετρικοί, οι ορθογώνιοι, οι normal και οι κάτω (άνω) τριγωνικοί πίνακες παραμένουν ομοίως κατά την αντιστροφή (εφ' όσον είναι ομαλοί). ▲

Το ίχνος (trace) του $A = (a_{ij})$, $i,j = 1(1)n$ ορίζεται από την σχέση

$$tr(A) = \sum_{i=1}^{n} a_{ii} \ . \tag{9.5}$$

<u>Θεώρημα 9.5</u> Εάν $A = (a_{ij})$, $B = (b_{ij})$, $i,j = 1(1)n$, τότε ισχύουν

i) $tr(AB) = tr(BA)$

ii) $tr(A^H A) = tr(AA^H) = \sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2$. ▲

Αν $v = (v_i)$, $w = (w_i)$, $i = 1(1)n$ είναι δύο διανύσματα, το εσωτερικό γινόμενο (v,w) των v και w ορίζεται από την σχέση

$$(v,w) \equiv v^H w = \sum_{i=1}^{n} v_i^* w_i \ . \tag{9.6}$$

<u>Θεώρημα 9.6</u> Για διανύσματα v,w και πίνακα A ισχύουν

i) $(v,v) \equiv v^H v = \sum_{i=1}^{n} |v_i|^2$

ii) $v^H w = (w^H v)^*$, $v^T w = w^T v$

iii) $(v^H A^H w)^* = w^H A v$ (9.7)

iv) $|v^H w| \leq (v^H v)^{\frac{1}{2}} (w^H w)^{\frac{1}{2}}$ (Ανισότης Cauchy – Schwarz, το = ισχύει εάν και μόνον εάν τα v,w είναι γραμμικώς εξηρτημένα).

<u>Ορισμός 9.2</u> Ένα σύνολο διανυσμάτων $v^{(i)}$, $i = 1(1)\kappa$ καλείται ορθογώνιο εάν $(v^{(i)}, v^{(j)}) = 0$ για $i \neq j$. Το σύνολο καλείται ορθοκανονικό εάν

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$(v^{(i)}, v^{(j)}) = \begin{cases} 0 & \text{για } i \neq j \\ 1 & \text{για } i = j \end{cases} . \tag{9.8}$$

9.3 Ο ρ ί ζ ο υ σ ε ς

Η ορίζουσα ενός πίνακα $A = (a_{ij})$ του $\mathbb{C}^{n,n}$ συμβολίζεται $\det(A)$ και ορίζεται επαγωγικά ως εξής:

Εάν $n = 1$, τότε $\det(A) = a_{11}$.

Εάν $n = 2$, τότε $\det(A) = a_{11}a_{22} - a_{21}a_{12}$.

Για $n > 2$ είναι:

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}) \ , \ i \in \{1,2,\ldots,n\} \ ,$$

ή (9.9)

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}) \ , \ j \in \{1,2,\ldots,n\} \ ,$$

όπου A_{ij} είναι ο $(n-1)\times(n-1)$ πίνακας που προκύπτει από τον A αν παραλείψουμε την i γραμμή και j στήλη του. Κάθε $\det(A_{ij})$ καλείται ελάσσων ορίζουσα του A . Ο αριθμός $(-1)^{i+j}\det(A_{ij})$ καλείται αλγεβρικό συμπλήρωμα του a_{ij} ή συμπαράγων του a_{ij} .

<u>Θεώρημα 9.7</u> Ο πίνακας $A \in \mathbb{C}^{n,n}$ είναι ομαλός, δηλαδή, υπάρχει ο αντίστροφος A^{-1} του A εάν και μόνον εάν $\det(A) \neq 0$.

Ένας πίνακας A είναι μη ομαλός εάν $\det(A) = 0$.

<u>Θεώρημα 9.8</u> i) Εάν όλα τα στοιχεία κάποιας γραμμής (ή στήλης) του A είναι 0 , τότε $\det(A) = 0$.

ii) Εάν κάθε στοιχείο κάποιας γραμμής (ή στήλης) του A πολλαπλασιασθεί με την σταθερά c , τότε για την ορίζουσα του πίνακα B που προκύπτει ισχύει $\det(B) = c\det(A)$.

iii) Ανταλλάσσοντας δύο γραμμές (ή στήλες) του A προκύπτει ένας πίνακας B με $\det(B) = -\det(A)$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

iv) Αν δύο γραμμές (ή στήλες) του A ταυτίζονται, τότε $\det(A) = 0$.

v) Εάν το πολλαπλάσιο μιας γραμμής (ή στήλης) του A προστεθεί σε κάποια άλλη γραμμή (στήλη) αυτού, τότε προκύπτει ένας πίνακας B με $\det(B) = \det(A)$.

<u>Θεώρημα 9.9</u> Αν $A,B \in \mathbb{C}^{n,n}$, τότε ισχύουν

i) $\det(A^T) = \det(A)$, $\det(\lambda A) = \lambda^n \det(A)$, όπου λ σταθερά

ii) $\det(AB) = \det(A)\det(B) = \det(BA)$

iii) $\det(A^{-1}) = [\det(A)]^{-1}$ (υποτίθεται ότι υπάρχει ο A^{-1}) .

<u>Θεώρημα 9.10</u> Η ορίζουσα ενός (άνω ή κάτω) τριγωνικού ή διαγώνιου πίνακα ισούται με το γινόμενο των στοιχείων της κυρίας διαγωνίου του.

<u>Θεώρημα 9.11</u> Έστω $A = (a_{ij}) \in \mathbb{C}^{n,n}$ ένας ομαλός πίνακας και m_{ij} το αλγεβρικό συμπλήρωμα του a_{ij} . Αν $M = (m_{ij})$, $i,j = 1(1)n$ είναι ο πίνακας των αλγεβρικών συμπληρωμάτων των στοιχείων του A , τότε

$$A^{-1} = \frac{M^T}{\det(A)} \quad . \tag{9.10}$$

(Ο πίνακας M^T καλείται προσηρτημένος (adjoint) πίνακας του A).

9.4 Συστήματα γραμμικών εξισώσεων

Ένα σύστημα n γραμμικών εξισώσεων με n αγνώστους έχει την μορφή

$$Ax = b \ , \tag{9.11}$$

όπου $A \in \mathbb{C}^{n,n}$, $x,b \in \mathbb{C}^n$. Τα A,b είναι γνωστά και το διάνυσμα – λύση $x = (x_i)$, $i = 1(1)n$ προσδιοριστέο.

<u>Θεώρημα 9.12</u> Το σύστημα (9.11) έχει μία μοναδική λύση εάν και

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

μόνον εάν ο A είναι ομαλός. Εάν ο A είναι ομαλός, τότε η λύση του (9.11) είναι

$$x = A^{-1}b \ . \tag{9.12}$$

Εάν ο A είναι μη ομαλός, τότε το σύστημα (9.11) ή δεν έχει λύση ή έχει άπειρες λύσεις.

Θεώρημα 9.13 (κανών του Cramer) Εάν $\det(A) \neq 0$, τότε η μοναδική λύση $x = (x_i)$, $i = 1(1)n$ του (9.11) δίνεται από τις

$$x_i = \frac{\det(A_i)}{\det(A)} \ , \quad i = 1(1)n \ , \tag{9.13}$$

όπου A_i είναι ο πίνακας που προκύπτει από τον A εάν αντικατασταθεί η i στήλη του από το διάνυσμα b .

Θεώρημα 9.14 Το σύστημα $Ax = 0$ έχει λύσεις $\neq 0$ εάν και μόνον εάν $\det(A) = 0$. ▲

Έστω A ένας $m \times n$ πίνακας. Η rank του A (συμβολίζεται rank(A)) ορίζεται ως η τάξη του μεγαλύτερου ομαλού τετραγωνικού υποπίνακα του A (ένας υποπίνακας του A είναι ένας πίνακας που προκύπτει από τον A παραλείποντας κάποιες από τις γραμμές του και/ ή κάποιες από τις στήλες του). Με άλλα λόγια, η rank(A) είναι ο μεγαλύτερος ακέραιος r , τέτοιος ώστε ο A έχει τουλάχιστον ένα $r \times r$ υποπίνακα A_r με $\det(A_r) \neq 0$. Η rank(A) ισούται με τον μέγιστο αριθμό των γραμμικώς ανεξαρτήτων γραμμών του A , που είναι ίσος με τον μέγιστο αριθμό των γραμμικώς ανεξαρτήτων στηλών αυτού. Προφανώς, ισχύει $\text{rank}(A) \leq \min\{m,n\}$.

Το σύνολο των λύσεων $x \in \mathbb{C}^n$ του $Ax = 0$ (ο A είναι $m \times n$ πίνακας) καλείται μηδενικός χώρος (null space) του A . Ο μηδενικός χώρος του A είναι ένας διανυσματικός χώρος του οποίου η διάσταση καλείται nullity του A (διάσταση ενός διανυσματικού χώρου (πεπερασμένης βάσεως) είναι ο αριθμός των διανυσμάτων σε κάθε βάση αυτού). Η nullity

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

του A ισούται με $n - \text{rank}(A)$, δηλαδή, το σύστημα $Ax = 0$ έχει $n - \text{rank}(A)$ γραμμικώς ανεξάρτητες λύσεις και κάθε λύση αυτού είναι ένας γραμμικός συνδυασμός των $n - \text{rank}(A)$ αυτών λύσεων.

Το σύνολο όλων των $y \in \mathbb{C}^m$, τέτοιων ώστε για κάποιο $x \in \mathbb{C}^n$ να έχουμε $Ax = y$ καλείται χώρος τιμών (range space) του A . Η $\text{rank}(A)$ είναι η διάσταση του χώρου τιμών του A .

<u>Θεώρημα 9.15</u> Εάν ο A είναι $m \times n$ πίνακας, τότε το σύστημα $Ax = b$ έχει τουλάχιστον μία λύση εάν και μόνον εάν $\text{rank}(A) = \text{rank}(\hat{A})$, όπου $\hat{A}$ είναι ο καλούμενος αυξημένος (augmented) , $m \times (n + 1)$ πίνακας του συστήματος, που προκύπτει θέτοντας δεξιά του A σαν μία νέα στήλη το διάνυσμα b . Εάν $\hat{x}$ είναι μία λύση του $Ax = b$, τότε για οποιαδήποτε άλλη λύση x αυτού έχουμε $x = \hat{x} + y$, όπου y ανήκει στον μηδενικό χώρο του A .

<u>Θεώρημα 9.16</u> Το σύστημα $Ax = b$, όπου ο A είναι $m \times n$ πίνακας έχει μία μοναδική λύση εάν και μόνον εάν $\text{rank}(A) = \text{rank}(\hat{A}) = n$ (η ισότητα αυτή προυποθέτει ότι $m \geqq n$).

<u>Θεώρημα 9.17</u> Το σύστημα $Ax = b$, όπου ο A είναι $m \times n$ πίνακας έχει άπειρες λύσεις εάν και μόνον εάν $\text{rank}(A) = \text{rank}(\hat{A}) < n$ (εδώ όλες οι περιπτώσεις: $m > n$, $m = n$, $m < n$ μπορούν να ισχύουν).

<u>Θεώρημα 9.18</u> Αν ο A είναι $m \times n$ με $m < n$, τότε το σύστημα $Ax = 0$ έχει λύση $x \neq 0$.

9.5 Ιδιοτιμές και ιδιοδιανύσματα

Ο πραγματικός ή μιγαδικός αριθμός λ είναι μία ιδιοτιμή του $A \in \mathbb{C}^{n,n}$, εάν για κάποιο διάνυσμα $v \neq 0$ έχουμε

$$Av = \lambda v \ . \tag{9.14}$$

Το οποιοδήποτε διάνυσμα $v \neq 0$, που ικανοποιεί την (9.14) καλείται ιδιοδιάνυσμα του A που αντιστοιχεί στην ιδιοτιμή λ . Εάν v και w

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

είναι ιδιοδιανύσματα του A που αντιστοιχούν στην ιδιοτιμή λ, τότε και κάθε γραμμικός συνδυασμός $c_1 v + c_2 w \neq 0$ είναι επίσης ιδιοδιάνυσμα του A που αντιστοιχεί στην ίδια ιδιοτιμή λ.

<u>Θεώρημα 9.19</u> Ο αριθμός λ είναι μία ιδιοτιμή του A εάν και μόνον εάν

$$P(\lambda) \equiv \det(A - \lambda I) = 0 . \qquad \blacktriangle \quad (9.15)$$

Η εξίσωση (9.15) καλείται χαρακτηριστική εξίσωση του A και το πρώτο μέλος της είναι ένα πολυώνυμο (το χαρακτηριστικό πολυώνυμο του A) $P(\lambda) = (-1)^n(\lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \ldots + a_{n-1}\lambda + a_n)$ βαθμού n ως προς λ. Οι ιδιοτιμές του A είναι οι n ρίζες της πολυωνυμικής εξισώσεως (9.15). Το σύνολο όλων των ιδιοτιμών του A καλείται φάσμα του A.

<u>Θεώρημα 9.20</u> Αν λ_i, $i = 1(1)p$ είναι οι διακεκριμένες ιδιοτιμές του A, τότε το v είναι ένα ιδιοδιάνυσμα του A που αντιστοιχεί στην ιδιοτιμή λ_j εάν και μόνον εάν το v είναι μη μηδενικό διάνυσμα του μηδενικού χώρου του $A - \lambda_j I$.

<u>Θεώρημα 9.21</u> Εάν λ_i, $i = 1(1)n$ είναι οι ιδιοτιμές του $A \in \mathbb{C}^{n,n}$, τότε

$$\det(A) = \prod_{i=1}^{n} \lambda_i \quad \text{και} \quad tr(A) = \sum_{i=1}^{n} \lambda_i .$$

<u>Ορισμός 9.3</u> Εάν λ_i, $i = 1(1)n$ είναι οι ιδιοτιμές του A, τότε η φασματική ακτίνα $\rho(A)$ αυτού ορίζεται ως εξής:

$$\rho(A) \equiv \max_i |\lambda_i| . \qquad (9.16)$$

<u>Θεώρημα 9.22</u> Εάν λ_i, $i = 1(1)\kappa$ είναι διακεκριμένες ιδιοτιμές του A και $v^{(i)}$, $i = 1(1)\kappa$ αντίστοιχα ιδιοδιανύσματα, τότε τα $v^{(i)}$, $i = 1(1)\kappa$ είναι γραμμικώς ανεξάρτητα.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Θεώρημα 9.23 i) Οι ιδιοτιμές ενός (άνω ή κάτω) τριγωνικού ή διαγώνιου πίνακα είναι τα στοιχεία της κυρίας διαγωνίου του.

ii) Ο πίνακας A^T έχει τις ίδιες ιδιοτιμές με τον A (οι A, A^T έχουν το ίδιο χαρακτηριστικό πολυώνυμο).

iii) Οι ιδιοτιμές του A^H είναι οι συζυγείς (μιγαδικές) των ιδιοτιμών του A.

iv) Αν $A, B \in \mathbb{C}^{n,n}$, τότε οι ιδιοτιμές του AB είναι οι ίδιες, με αυτές του BA (οι AB, BA έχουν το ίδιο χαρακτηριστικό πολυώνυμο).

Ορισμός 9.4 Ο πίνακας B λέγεται ότι είναι όμοιος με τον πίνακα A εάν υπάρχει ένας ομαλός πίνακας P, τέτοιος ώστε $B = PAP^{-1}$. Γράφουμε $B \sim A$.

Θεώρημα 9.24 Αν A, B είναι όμοιοι ($B = PAP^{-1}$, $\det(P) \neq 0$), τότε ισχύουν: i) $\det(A) = \det(B)$ ii) $tr(A) = tr(B)$ iii) οι A, B έχουν ακριβώς τις ίδιες ιδιοτιμές (το ίδιο χαρακτηριστικό πολυώνυμο) iv) Εάν v είναι ένα ιδιοδιάνυσμα του A που αντιστοιχεί στην ιδιοτιμή λ, το $w = Pv$ είναι ιδιοδιάνυσμα του B που αντιστοιχεί στην ιδιοτιμή λ.

Θεώρημα 9.25 Εάν ο A έχει ένα σύνολο n γραμμικώς ανεξαρτήτων ιδιοδιανυσμάτων, τότε είναι όμοιος με ένα διαγώνιο πίνακα του οποίου τα διαγώνια στοιχεία είναι ιδιοτιμές του A (βλέπε άσκηση 17).

Ορισμός 9.5 Ένας πίνακας A καλείται μη ελλιπής (non - defective) εάν ο αριθμός κ των γραμμικώς ανεξαρτήτων ιδιοδιανυσμάτων του ισούται με n. Εάν $\kappa < n$, τότε ο A καλείται ελλιπής (defective).

Θεώρημα 9.26 Αν ο πίνακας A έχει n διακεκριμένες ιδιοτιμές λ_i, $i = 1(1)n$, τότε ο A είναι μή ελλιπής.

Ορισμός 9.6 Ένας πίνακας A καλείται διαγωνίσιμος εάν είναι όμοιος με ένα διαγώνιο πίνακα.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Θεώρημα 9.27 Ένας πίνακας A είναι διαγωνίσιμος εάν και μόνον εάν υπάρχει ένας θετικά ορισμένος (βλέπε ορισμό 9.15) Ερμιτιανός πίνακας P, τέτοιος ώστε ο $M = PAP^{-1}$ είναι normal πίνακας. ▲

Κάθε normal πίνακας (π.χ. πραγματικός συμμετρικός ή Ερμιτιανός) είναι διαγωνίσιμος.

Θεώρημα 9.28 Ένας πίνακας είναι διαγωνίσιμος εάν και μόνον εάν είναι μή ελλιπής.

Θεώρημα 9.29 Εάν λ_i είναι μία ιδιοτιμή του A πολλαπλότητος m_i, τότε η διάσταση d_i του μηδενικού χώρου του $A - \lambda_i I$ ικανοποιεί την $1 \leq d_i \leq m_i$. Η d_i καλείται γεωμετρική πολλαπλότης της λ_i, ενώ η m_i καλείται αλγεβρική πολλαπλότης της λ_i. Σύμφωνα με την παράγραφο 9.4 είναι

$$d_i \equiv (\text{nullity του } A - \lambda_i I) = n - \text{rank}(A - \lambda_i I) \quad . \qquad (9.17)$$

Θεώρημα 9.30 (Η κανονική μορφή Jordan) Έστω $A \in \mathbb{C}^{n,n}$ και λ_1, $\lambda_2, \ldots \lambda_p$ οι διακεκριμένες ιδιοτιμές του με αλγεβρικές και γεωμετρικές πολλαπλότητες m_i, $i = 1(1)p$ και d_i, $i = 1(1)p$ αντίστοιχα. Τότε, υπάρχει ένας ομαλός πίνακας T, τέτοιος ώστε

$$J \equiv T^{-1}AT = \begin{bmatrix} J_1 & 0 & \ldots & 0 \\ 0 & J_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & J_p \end{bmatrix} , \qquad (9.18)$$

όπου κάθε J_κ είναι ένας τετραγωνικός πίνακας τάξεως m_κ, $\kappa = 1(1)p$, που έχει την μορφή

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$J_\kappa = \begin{bmatrix} \lambda_\kappa & \delta_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_\kappa & \delta_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \lambda_\kappa & \delta_3 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \lambda_\kappa & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \lambda_\kappa & \delta_{m_\kappa - 1} \\ 0 & 0 & 0 & 0 & \cdots & 0 & \lambda_\kappa \end{bmatrix} . \qquad (9.19)$$

Στον J_κ , τα $d_\kappa - 1$ από τα δ_i , $i = 1(1)m_\kappa - 1$ είναι ίσα με 0 και τα άλλα $m_\kappa - d_\kappa$ από τα δ_i είναι ίσα με 1 . Αν συμβολίσουμε με $J_{\nu_j(\kappa)}$ τον παρακάτω $\nu_j(\kappa) \times \nu_j(\kappa)$ πίνακα $(\nu_j(\kappa) \geqq 1)$

$$J_{\nu_j(\kappa)} = \begin{bmatrix} \lambda_\kappa & 1 & & & \\ & \lambda_\kappa & 1 & & \bigcirc \\ & & \ddots & \ddots & \\ & \bigcirc & & & 1 \\ & & & & \lambda_\kappa \end{bmatrix} , \qquad (9.20)$$

τότε για κάθε $\kappa = 1(1)p$ υπάρχουν d_κ φυσικοί αριθμοί $\nu_j(\kappa)$, $j = 1(1)d_\kappa$ με $\sum_{j=1}^{d_\kappa} \nu_j(\kappa) = m_\kappa$ και ο J_κ που δίνεται από την (9.19) παίρνει την μορφή

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$J_\kappa = \begin{bmatrix} J_{\nu_1(\kappa)} & & & \bigcirc \\ & J_{\nu_2(\kappa)} & & \\ & & \ddots & \\ \bigcirc & & & J_{\nu_{d_\kappa}(\kappa)} \end{bmatrix} . \tag{9.21}$$

Ο block διαγώνιος πίνακας J στην (9.18) καλείται η κανονική μορφή Jordan του A και είναι μοναδικός εκτός από πιθανές μεταθέσεις των $J_{\nu_j(\kappa)}$, $j = 1(1)d_\kappa$, $\kappa = 1(1)p$, που καλούνται Jordan blocks. ▲

Είναι φανερό ότι ο J είναι άνω τριγωνικός και $A \sim J$. Ακόμη, ο αριθμός των Jordan blocks ισούται με τον αριθμό $\sum_{\kappa=1}^{p} d_\kappa$ των γραμμικώς ανεξαρτήτων ιδιοδιανυσμάτων του πίνακα A . Ο J είναι διαγώνιος πίνακας εάν και μόνον εάν ο A είναι μη ελλιπής.

<u>Θεώρημα 9.31</u> Δύο πίνακες είναι όμοιοι εάν και μόνον εάν έχουν την ίδια κανονική μορφή Jordan (εκτός από πιθανές μεταθέσεις των Jordan blocks). ▲

Οι στήλες του πίνακα T στην (9.18) αποτελούν μία βάση του $\mathbb{C}^n$, η οποία αποτελείται από τα γραμμικώς ανεξάρτητα ιδιοδιανύσματα του A και από τα καλούμενα κύρια διανύσματα του A (principal vectors) ή γενικευμένα ιδιοδιανύσματα του A (βλέπε ορισμό 9.8). Σε κάθε Jordan block $J_{\nu_j(\kappa)}$, $j = 1(1)d_\kappa$, $\kappa = 1(1)p$ αντιστοιχούν 1 ιδιοδιάνυσμα του A και $\nu_j(\kappa) - 1$ γενικευμένα ιδιοδιανύσματα του A .

Τα χαρακτηριστικά πολυώνυμα

$$(\lambda_\kappa - \lambda)^{\nu_j(\kappa)} = \det(J_{\nu_j(\kappa)} - \lambda I) \tag{9.22}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

όλων των Jordan blocks $J_{\nu_j(\kappa)}$, $j = 1(1)d_\kappa$, $\kappa = 1(1)p$ καλούνται οι στοιχειώδεις διαιρέτες (elementary divisors) του A .

Θεώρημα 9.32 Ο $A \in \mathbb{C}^{n,n}$ είναι διαγωνίσιμος εάν και μόνον εάν όλοι οι στοιχειώδεις διαιρέτες του είναι γραμμικοί ($\nu_j(\kappa) = 1$, $\forall$ j και κ).

Ορισμός 9.7 Ο πίνακας $A \in \mathbb{C}^{n,n}$ καλείται μη ακυρωτικός (non derogatory) εάν ισχύει $d_\kappa = 1$, $\kappa = 1(1)p$ στην κανονική μορφή Jordan του A (βλέπε θεώρημα (9.30)), δηλαδή, κάθε διακεκριμένη ιδιοτιμή λ_κ ($\kappa = 1(1)p$) εμφανίζεται μόνον σε ένα Jordan block. Διαφορετικά, ο A καλείται ακυρωτικός (derogatory).

Ορισμός 9.8 Ένα διάνυσμα $v \in \mathbb{C}^n$ είναι ένα κύριο διάνυσμα (γενικευμένο ιδιοδιάνυσμα) βαθμού κ του πίνακα $A \in \mathbb{C}^{n,n}$, που αντιστοιχεί στην ιδιοτιμή λ αυτού, εάν

$$(A - \lambda I)^\kappa v = 0 , \quad \text{αλλά} \quad (A - \lambda I)^{\kappa-1} v \neq 0 \quad , \kappa \geq 1 . \qquad (9.23)$$

Θεώρημα 9.33 Κάθε πίνακας $A \in \mathbb{C}^{n,n}$ έχει n γραμμικώς ανεξάρτητα κύρια διανύσματα. Κύρια διανύσματα που αντιστοιχούν σε διακεκριμένες ιδιοτιμές είναι γραμμικώς ανεξάρτητα. Εάν λ είναι μία ιδιοτιμή του A πολλαπλότητος m , τότε ο A έχει m γραμμικώς ανεξάρτητα κύρια διανύσματα που αντιστοιχούν στην λ .

Θεώρημα 9.34 (Schur) Για κάθε $A \in \mathbb{C}^{n,n}$ υπάρχει ένας unitary πίνακας U τέτοιος ώστε $UAU^H = T$, όπου T είναι τριγωνικός και επομένως οι ιδιοτιμές του A είναι τα διαγώνια στοιχεία του T .

Ορισμός 9.9 Εάν λ είναι μία ιδιοτιμή του $A^H A$, $A \in \mathbb{C}^{n,n}$, τότε ο $\lambda^{\frac{1}{2}}$ καλείται μία μη ομαλή (singular) τιμή του A .

Ορισμός 9.10 Καλείται ελάχιστο (minimal) πολυώνυμο του $A \in \mathbb{C}^{n,n}$ το πολυώνυμο

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$P(\lambda) = \lambda^m + a_1\lambda^{m-1} + a_2\lambda^{m-2} + \ldots + a_{m-1}\lambda + a_m$$

του μικρότερου βαθμού, που έχει την ιδιότητα $P(A) = 0$.

<u>Θεώρημα 9.35</u> Έχοντας υπ' όψιν το θεώρημα (9.30), έστω $\tau_\kappa = \max_j\{\nu_j(\kappa)\}$, $j = 1(1)d_\kappa$, $\kappa = 1(1)p$. Τότε, το

$$P(\lambda) = (\lambda - \lambda_1)^{\tau_1}(\lambda - \lambda_2)^{\tau_2} \ldots (\lambda - \lambda_p)^{\tau_p} \tag{9.24}$$

είναι το ελάχιστο πολυώνυμο του A . Το ελάχιστο πολυώνυμο του A διαιρεί κάθε πολυώνυμο $R(\lambda)$ με $R(A) = 0$.

<u>Θεώρημα 9.36</u> (Cayley – Hamilton) Αν $P(\lambda) \equiv \det(A - \lambda I)$ είναι το χαρακτηριστικό πολυώνυμο του A , τότε $P(A) = 0$.

<u>Θεώρημα 9.37</u> Έστω πίνακας $A \in \mathbb{C}^{n,n}$ και

$$\det(\lambda I - A) \equiv P(\lambda) = \lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \ldots + a_{n-1}\lambda + a_n.$$

Τότε, ο πίνακας $F \in \mathbb{C}^{n,n}$ (ή ο F^T) , όπου

$$F = \begin{bmatrix} 0 & 1 & 0 & & & \\ 0 & 0 & 1 & & \bigcirc & \\ & & \ddots & \ddots & & \\ & \bigcirc & & & 0 & 1 \\ -a_n & -a_{n-1} & \ldots & & -a_2 & -a_1 \end{bmatrix} \tag{9.25}$$

καλείται συνοδεύων πίνακας (του Frobenius) του $P(\lambda)$ και ισχύει

$$\det(\lambda I - F) = P(\lambda) \quad .$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

<u>Θεώρημα 9.38</u> (Gerschgorin) Εάν $A = (a_{ij})$, $i,j = 1(1)n$, τότε όλες οι ιδιοτιμές του βρίσκονται στην τομή των συνόλων S και $\hat{S}$, όπου $S = \bigcup_{\kappa=1}^{n} D_\kappa$ και $\hat{S} = \bigcup_{\kappa=1}^{n} \hat{D}_\kappa$ με

$$D_\kappa = \left\{ z \in \mathbb{C} : |z - a_{\kappa\kappa}| \leqq \sum_{\substack{j=1 \\ j \neq \kappa}}^{n} |a_{\kappa j}| \right\} \quad \text{και}$$

$$\hat{D}_\kappa = \left\{ z \in \mathbb{C} : |z - a_{\kappa\kappa}| \leqq \sum_{\substack{j=1 \\ j \neq \kappa}}^{n} |a_{j\kappa}| \right\} . \qquad \blacktriangle$$

Για $\kappa = 1(1)n$ τα σύνολα D_κ και $\hat{D}_\kappa$ είναι δίσκοι (δίσκος = ο κύκλος και το εσωτερικό του) στο μιγαδικό επίπεδο.

<u>Θεώρημα 9.39</u> Εάν $A \in \mathbb{C}^{n,n}$, $A = (a_{ij})$, τότε

$$\rho(A) \leqq \min\left[\max_i \left\{ \sum_{j=1}^{n} |a_{ij}| \right\} , \max_j \left\{ \sum_{i=1}^{n} |a_{ij}| \right\} \right] , \quad i,j = 1(1)n .$$

<u>Θεώρημα 9.40</u> Γράφοντας τον πίνακα $A \in \mathbb{C}^{n,n}$ στην μορφή $A = H_1 + iH_2$, όπου H_1, H_2 είναι Ερμιτιανοί πίνακες $(H_1 = \frac{1}{2}(A + A^H)$, $H_2 = \frac{1}{2i}(A - A^H))$, τότε για κάθε ιδιοτιμή λ του A $(\lambda = \operatorname{Re}\lambda + i \operatorname{Im}\lambda)$ ισχύει

$$\lambda_m(H_1) \leqq \operatorname{Re}\lambda \leqq \lambda_M(H_1) ,$$

$$\lambda_m(H_2) \leqq \operatorname{Im}\lambda \leqq \lambda_M(H_2) ,$$

όπου $\lambda_m(H_i)$, $\lambda_M(H_i)$ είναι αντίστοιχα η μικρότερη και η μεγαλύτερη ιδιοτιμή του H_i , $i = 1,2$.

9.6 Ορισμοί και θεωρήματα για ειδικούς και μη πίνακες

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Ορισμός 9.11 Ο πίνακας $P \in \mathbb{R}^{n,n}$ ($\mathbb{R}^{n,n}$ είναι το σύνολο όλων των $n \times n$ πραγματικών πινάκων) καλείται πίνακας μεταθέσεως (permutation) εάν στήλες του είναι οι στήλες του ταυτοτικού πίνακα I με οποιαδήποτε σειρά. Δηλαδή, ο P έχει σε κάθε γραμμή και σε κάθε στήλη ένα μόνο στοιχείο διάφορο του μηδενός που ισούται με 1.

Θεώρημα 9.41 i) Κάθε πίνακας μεταθέσεως είναι ορθογώνιος
ii) Αν $A \in \mathbb{C}^{n,n}$ και P πίνακας μεταθέσεως, τότε το γινόμενο PA έχει σαν αποτέλεσμα μετάθεση των γραμμών του A, ενώ το AP μετάθεση των στηλών του A.

Ορισμός 9.12 Ο πίνακας $A \in \mathbb{R}^{n,n}$, $A = (a_{ij})$, καλείται μη αρνητικός (θετικός) και γράφουμε $A \geqq 0$ $(A > 0)$ εάν $a_{ij} \geqq 0$ $(a_{ij} > 0)$, για κάθε $i,j = 1(1)n$. Εάν $A \geqq 0$, $B \geqq 0$ και $A - B \geqq 0$, τότε λέμε ότι $A \geqq B$. ▲

Εάν $A \in \mathbb{C}^{n,n}$, $A = (a_{ij})$, τότε με $|A|$ συμβολίζουμε τον πίνακα του οποίου στοιχεία είναι τα μέτρα των αντίστοιχων στοιχείων του A, δηλαδή, $|A| = (|a_{ij}|)$. Προφανώς, $|A| \geqq 0$.

Θεώρημα 9.42 Εάν $A > 0$, τότε ο A έχει μία απλή θετική ιδιοτιμή $\rho = \rho(A)$ με $|\lambda_i| < \rho$ για όλες τις άλλες ιδιοτιμές λ_i του A. Στην ιδιοτιμή ρ αντιστοιχεί ένα ιδιοδιάνυσμα $u > 0$.

Θεώρημα 9.43 Εάν $A \in \mathbb{R}^{n,n}$, $B \in \mathbb{C}^{n,n}$ και $A \geqq |B|$, τότε $\rho(A) \geqq \rho(B)$.

Ορισμός 9.13 i) Ένας πίνακας $A \in \mathbb{C}^{n,n}$, $A = (a_{ij})$ καλείται αυστηρά διαγώνια υπέρτερος εάν

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| , \tag{9.26}$$

για όλα τα $i = 1(1)n$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ii) Ο $A = (a_{ij})$ καλείται ασθενώς διαγώνια υπέρτερος εάν

$$|a_{ii}| \geqq \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| \ , \ i = 1(1)n \tag{9.27}$$

και για ένα τουλάχιστον i η (9.27) είναι της μορφής (9.26).

Ορισμός 9.14 Οι πίνακες $A,B \in \mathbb{C}^{n,n}$ καλούνται ισοδύναμοι εάν $A = PBQ$, όπου $P,Q \in \mathbb{C}^{n,n}$ είναι ομαλοί πίνακες.

Θεώρημα 9.44 Ο πίνακας $A \in \mathbb{C}^{n,n}$ είναι normal εάν και μόνον εάν υπάρχει ένας unitary πίνακας U , τέτοιος ώστε $U^{-1}AU = \Lambda$, όπου Λ διαγώνιος (ο U είναι ο modal πίνακας του A).

Θεώρημα 9.45 Εάν $A \in \mathbb{C}^{n,n}$, τότε ο $x^H Ax$ είναι πραγματικός αριθμός για κάθε $x \in \mathbb{C}^n$ εάν και μόνον εάν ο A είναι Ερμιτιανός.

Ορισμός 9.15 Ο Ερμιτιανός πίνακας A καλείται θετικά ορισμένος (positive definite) εάν $x^H Ax > 0$ για κάθε $x \in \mathbb{C}^n$, $x \neq 0$. Ο A καλείται μη αρνητικά ορισμένος (nonnegative definite ή semipositive definite) εάν $x^H Ax \geqq 0$ για κάθε $x \in \mathbb{C}^n$.

Ορισμός 9.16 Ο πραγματικός συμμετρικός πίνακας A καλείται θετικά ορισμένος εάν $x^T Ax > 0$ για κάθε $x \in \mathbb{R}^n$, $x \neq 0$.

Παρατήρηση Ο ορισμός 9.15 στην περίπτωση που ο A είναι πραγματικός πίνακας συμπίπτει με τον ορισμό 9.16, αφού ισχύει το ακόλουθο θεώρημα.

Θεώρημα 9.46 Εάν $A \in \mathbb{R}^{n,n}$, συμμετρικός και $x^T Ax > 0$, $\forall\, x \in \mathbb{R}^n$, $x \neq 0$, τότε $x^H Ax > 0$ $\forall\, x \in \mathbb{C}^n$, $x \neq 0$.

Ορισμός 9.17 Εάν $A \in \mathbb{C}^{n,n}$ και $x^T Ax > 0$ $\forall\, x \in \mathbb{R}^n$, $x \neq 0$, τότε ο A καλείται θετικά πραγματικός πίνακας (positive real).

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Σημείωση Οι θετικά ορισμένοι πίνακες είναι και θετικά πραγματικοί. Το αντίστροφο δεν ισχύει αναγκαστικά.

Θεώρημα 9.47 Ο $A \in \mathbb{C}^{n,n}$ είναι θετικά πραγματικός εάν και μόνον εάν ο

$A^T + A$ είναι πραγματικός και θετικά ορισμένος.

Θεώρημα 9.48 Εάν $A \in \mathbb{C}^{n,n}$, τότε ο $A^H A$ (και ο AA^H) είναι Ερμιτιανός και μη αρνητικά ορισμένος. Ο $A^H A$ είναι θετικά ορισμένος εάν και μόνον εάν $\det(A) \neq 0$.

Θεώρημα 9.49 Ο Ερμιτιανός πίνακας A είναι θετικά ορισμένος (μή αρνητικά ορισμένος) εάν και μόνον εάν όλες οι ιδιοτιμές του είναι θετικές (μή αρνητικές).

Θεώρημα 9.50 Εάν A είναι Ερμιτιανός και θετικά ορισμένος, τότε το ίδιο ισχύει για τον A^{-1} και όλους τους κύριους υποπίνακες του A. Επίσης, όλες οι κύριες υποορίζουσες του A είναι θετικές και το μεγαλύτερο απολύτως στοιχείο του βρίσκεται επί της κυρίας διαγωνίου του.

Θεώρημα 9.51 Ο Ερμιτιανός πίνακας A είναι θετικά ορισμένος εάν και μόνον εάν $\det(A_\kappa) > 0$, $\kappa = 1(1)n$, όπου A_κ είναι ο $\kappa \times \kappa$ πίνακας που σχηματίζεται από τις πρώτες κ γραμμές και κ στήλες του A, $\kappa = 1(1)n$.

Θεώρημα 9.52 Ένας Ερμιτιανός πίνακας A είναι θετικά ορισμένος (μη αρνητικά ορισμένος) εάν και μόνον εάν κάθε υποορίζουσα που σχηματίζεται παραλείποντας m οποιεσδήποτε γραμμές του A και τις αντίστοιχες στήλες είναι θετική (μή αρνητική), όπου $0 \leq m < n$. Ειδικά έχουμε $\det(A) > 0$ (≥ 0) και $a_{ii} > 0$ (≥ 0), $i = 1(1)n$.

Θεώρημα 9.53 Εάν ο A είναι Ερμιτιανός και θετικά ορισμένος, τότε για κάθε ομαλό πίνακα L ο πίνακας LAL^H είναι Ερμιτιανός και

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

θετικά ορισμένος.

Θεώρημα 9.54 Για κάθε Ερμιτιανό και θετικά ορισμένο πίνακα A υπάρχει ένας μοναδικός κάτω τριγωνικός πίνακας L με θετικά διαγώνια στοιχεία, τέτοιος ώστε $A = LL^H$. Εάν ο $A \in \mathbb{R}^{n,n}$ τότε και $L \in \mathbb{R}^{n,n}$.

Θεώρημα 9.55 Το πραγματικό μέρος ενός Ερμιτιανού και θετικά ορισμένου (μη αρνητικά ορισμένου) πίνακα είναι ένας πραγματικός συμμετρικός θετικά ορισμένος (μή αρνητικά ορισμένος) πίνακας.

Θεώρημα 9.56 Εάν ο A είναι Ερμιτιανός και θετικά ορισμένος (μή αρνητικά ορισμένος), τότε υπάρχει ένας μοναδικός θετικά (μη αρνητικά) ορισμένος πίνακας B, τέτοιος ώστε $B^2 = A$. Αν $A \in \mathbb{R}^{n,n}$, τότε $B \in \mathbb{R}^{n,n}$. Γράφουμε $B = A^{\frac{1}{2}}$.

Θεώρημα 9.57 Εάν ο $A = (a_{ij})$ είναι ένας Ερμιτιανός και αυστηρά διαγώνια υπέρτερος πίνακας με $a_{ii} > 0$, τότε ο A είναι θετικά ορισμένος.

Θεώρημα 9.58 Έστω $A \in \mathbb{C}^{n,n}$. Τότε, ο A γράφεται σαν το γινόμενο $A = LR$, όπου ο L είναι unit κάτω τριγωνικός και ο R άνω τριγωνικός με μη μηδενικά διαγώνια στοιχεία εάν και μόνον εάν όλοι οι άνω αριστερά κύριοι υποπίνακες του A είναι ομαλοί (οπότε $\det(A) \neq 0$). Επί πλέον η παραγοντοποίηση $A = LR$, όταν είναι δυνατή, είναι μοναδική.

Θεώρημα 9.59 Εάν ο $A \in \mathbb{C}^{n,n}$ είναι ομαλός, τότε υπάρχει μία ανάλυση $A = QR$, όπου Q είναι unitary και R είναι άνω τριγωνικός. Επί πλέον, εάν τα διαγώνια στοιχεία του R είναι θετικοί αριθμοί τότε η ανάλυση είναι μοναδική.

Θεώρημα 9.60 Κάθε $A \in \mathbb{C}^{n,n}$ μπορεί πάντοτε να γραφεί σαν $A = QR$, όπου Q είναι unitary και R άνω τριγωνικός ("QR" διάσπαση του A).

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Θεώρημα 9.61 Για κάθε $A \in \mathbb{C}^{n,n}$ υπάρχουν unitary πίνακες U και V, τέτοιοι ώστε UAV είναι διαγώνιος πίνακας με διαγώνια στοιχεία τις μή ομαλές τιμές του A (βλέπε ορισμό 9.9).

Θεώρημα 9.62 Εάν $A \in \mathbb{C}^{n,n}$ είναι ομαλός πίνακας, τότε υπάρχει ένας πίνακας μεταθέσεως P, τέτοιος ώστε ο πίνακας PA έχει όλους τους άνω αριστερά κύριους υποπίνακες του ομαλούς.

Θεώρημα 9.63 Εάν όλοι οι άνω αριστερά κύριοι υποπίνακες του A είναι ομαλοί, τότε ο A μπορεί να παρασταθεί μονοσήμαντα σαν $A = LDU$, όπου L είναι unit κάτω τριγωνικός, D διαγώνιος με διαγώνια στοιχεία $\neq 0$ και U unit άνω τριγωνικός.

Θεώρημα 9.64 Έστω $A \in \mathbb{C}^{n,n}$. Τότε, ο A γράφεται κατά μοναδικό τρόπο ως $A = LDU$, όπου L unit κάτω τριγωνικός, D διαγώνιος και U unit άνω τριγωνικός, εάν και μόνον εάν οι πρώτοι $n-1$ άνω αριστερά κύριοι υποπίνακες του A είναι ομαλοί.

Θεώρημα 9.65 Ένας normal πίνακας έχει ένα πλήρες ορθοκανονικό σύνολο ιδιοδιανυσμάτων (n γραμμικώς ανεξάρτητα ιδιοδιανύσματα).

Θεώρημα 9.66 Έστω $A \in \mathbb{C}^{n,n}$ ένας Ερμιτιανός πίνακας $(A^H = A)$. Τότε, ισχύουν:

1) Οι ιδιοτιμές του είναι πραγματικές.
2) Ιδιοδιανύσματα που αντιστοιχούν σε διακεκριμένες ιδιοτιμές είναι ορθογώνια.
3) Έχει ένα πλήρες ορθοκανονικό σύνολο ιδιοδιανυσμάτων (n γραμμικώς ανεξάρτητα ιδιοδιανύσματα).
4) Υπάρχει ένας unitary πίνακας U τέτοιος ώστε $U^{-1}AU = \Lambda$, όπου Λ διαγώνιος (ο U είναι ο modal πίνακας του A).
5) $\lambda_m \leq \dfrac{x^H A x}{x^H x} \leq \lambda_M$, $\forall\, x \in \mathbb{C}^n$, $x \neq 0$, όπου $\lambda_m = \min\limits_i \lambda_i$, $\lambda_M = \max\limits_i \lambda_i$, $i = 1(1)n$ και λ_i, $i = 1(1)n$ οι ιδιοτιμές του

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

A. Επί πλέον, υπάρχουν διανύσματα w και z διάφορα του μηδενός, τέτοια ώστε:

$$\frac{w^H A w}{w^H w} = \lambda_m \ , \ \frac{z^H A z}{z^H z} = \lambda_M \quad \text{και} \quad Aw = \lambda_m w \ , \ Az = \lambda_M z \ ,$$

δηλαδή, $\lambda_m = \min\limits_{x \neq 0} \dfrac{x^H A x}{x^H x}$ και $\lambda_M = \max\limits_{x \neq 0} \dfrac{x^H A x}{x^H x}$.

6) Εάν λ_i είναι μία ιδιοτιμή του, αλγεβρικής πολλαπλότητος m_i, τότε η γεωμετρική πολλαπλότης d_i της λ_i ισούται με m_i , δηλαδή, $m_i = n - \text{rank}(A - \lambda_i I)$.

7) ο iA είναι αντι – Ερμιτιανός πίνακας.

8) Τα ιδιοδιανύσματα του A^T είναι τα συζυγή των ιδιοδιανυσμάτων του A .

9) Αν λ_i , $i = 1(1)n$ είναι οι ιδιοτιμές του και $x^{(i)}$ τα αντίστοιχα ιδιοδιανύσματα, τότε

$$A = \lambda_1 x^{(1)} x^{(1)^H} + \lambda_2 x^{(2)} x^{(2)^H} + \ldots + \lambda_n x^{(n)} x^{(n)^H} \ .$$

10) Εάν όλοι οι άνω αριστερά κύριοι υποπίνακές του είναι ομαλοί, τότε ο A μπορεί να γραφεί μονοσήμαντα σαν $A = LDL^H$, όπου L είναι unit κάτω τριγωνικός και D διαγώνιος πραγματικός με διαγώνια στοιχεία $\neq 0$. Ισχύει $D > 0$ εάν και μόνον εάν ο A είναι θετικά ορισμένος.

11) Έστω $n > 1$ και B ο $(n-1) \times (n-1)$ Ερμιτιανός πίνακας, που σχηματίζεται παραλείποντας την τελευταία γραμμή και τελευταία στήλη του A . Αν $\lambda_1 \geqq \lambda_2 \geqq \ldots \geqq \lambda_{n-1} \geqq \lambda_n$ είναι οι ιδιοτιμές του A και $\mu_1 \geqq \mu_2 \geqq \ldots \geqq \mu_{n-2} \geqq \mu_{n-1}$ οι ιδιοτιμές του B , τότε $\lambda_1 \geqq \mu_1 \geqq$ $\geqq \lambda_2 \geqq \mu_2 \geqq \ldots \geqq \lambda_{n-1} \geqq \mu_{n-1} \geqq \lambda_n$.

<u>Θεώρημα 9.67</u> Έστω $A \in \mathbb{R}^{n,n}$ ένας συμμετρικός πίνακας. Τότε ισχύουν:

1) ο A είναι Ερμιτιανός .

2) Βλέπε Θεώρημα 9.66 , (1), (2) και (3) .

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

3) Τα ιδιοδιανύσματά του μπορούν πάντοτε να εκλεγούν πραγματικά.

4) Υπάρχει ένας πραγματικός ορθογώνιος πίνακας Q, τέτοιος ώστε $Q^{-1}AQ = \Lambda$, όπου Λ διαγώνιος (ο Q είναι ο modal πίνακας του A).

5) Βλέπε θεώρημα 9.66, (5).

6) Εάν όλοι οι άνω αριστερά κύριοι υποπίνακες του A είναι ομαλοί, τότε ο A μπορεί να γραφεί μονοσήμαντα σαν $A = LDL^T$ όπου ο L είναι πραγματικός unit κάτω τριγωνικός και D διαγώνιος πραγματικός με διαγώνια στοιχεία $\neq 0$. Ισχύει $D > 0$ εάν και μόνον εάν ο A είναι θετικά ορισμένος.

7) Εάν όλοι οι άνω αριστερά κύριοι υποπίνακες του A είναι ομαλοί, τότε ο A μπορεί να γραφεί σαν $A = MM^T$, όπου M είναι γενικά μιγαδικός, κάτω τριγωνικός με διαγώνια στοιχεία $\neq 0$. Ο M είναι πραγματικός εάν και μόνον εάν ο A είναι θετικά ορισμένος.

<u>Θεώρημα 9.68</u> Έστω $A \in \mathbb{C}^{n,n}$ ένας αντι – Ερμιτιανός πίνακας. Τότε ισχύουν:

1) Οι ιδιοτιμές του είναι καθαροί φανταστικοί αριθμοί ή 0.

2) Τα διαγώνια στοιχεία του είναι πολλαπλάσια του i ή 0.

3) Ο iA είναι Ερμιτιανός.

4) $\forall\, x \in \mathbb{C}^n$ ο $x^H Ax$ είναι καθαρός φανταστικός αριθμός.

<u>Θεώρημα 9.69</u> Έστω $A \in \mathbb{R}^{n,n}$ ένας αντισυμμετρικός πίνακας. Τότε ισχύουν:

1) Οι ιδιοτιμές του είναι είτε 0 ή καθαροί φανταστικοί αριθμοί (συζυγείς). Εάν n περιττός, τότε τουλάχιστον μία ιδιοτιμή του είναι ίση με 0.

2) Η $\operatorname{rank}(A)$ είναι πάντοτε άρτιος αριθμός.

3) $\forall\, x \in \mathbb{R}^n$, $x^T Ax = 0$.

4) Τα διαγώνια στοιχεία του $-a_{ii} = 0$, $i = 1(1)n$.

<u>Θεώρημα 9.70</u> Έστω $A \in \mathbb{C}^{n,n}$ ένας unitary πίνακας. Τότε ισχύουν:

1) Οι γραμμές και οι στήλες του αποτελούν ορθοκανονικά σύνολα διανυσμάτων.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

2) $|\det(A)| = 1$ (Αν $A \in \mathbb{R}^{n,n}$ τότε $\det(A) = \pm 1$).

3) $|\lambda_i| = 1$, $i = 1(1)n$, όπου λ_i , $i = 1(1)n$ οι ιδιοτιμές του A (Αν $A \in \mathbb{R}^{n,n}$, τότε αν λ είναι μία ιδιοτιμή του A και η $\frac{1}{\lambda}$ είναι ιδιοτιμή του A).

9.7 Norms διανύσματος και norms πίνακα

Ορισμός 9.18 Η ακολουθία διανυσμάτων $\{x^{(\kappa)}\} = x^{(1)}, x^{(2)},\ldots, x^{(\kappa)},\ldots$, όπου $x^{(\kappa)} = (x_i^{(\kappa)}) \in \mathbb{C}^n$, $\kappa = 1,2,\ldots$ συγκλίνει σε ένα διάνυσμα $x = (x_i)$ καθώς το $\kappa \to \infty$ και γράφουμε $\lim\limits_{\kappa\to\infty} x^{(\kappa)} = x$ ή $x^{(\kappa)} \underset{\kappa\to\infty}{\to} x$ εάν και μόνον εάν $\lim\limits_{\kappa\to\infty} x_i^{(\kappa)} = x_i$, $i = 1(1)n$.

Ορισμός 9.19 Η απειροσειρά διανυσμάτων $\sum\limits_{\kappa=1}^{\infty} x^{(\kappa)}$ συγκλίνει στο διάνυσμα y και γράφουμε $\sum\limits_{\kappa=1}^{\infty} x^{(\kappa)} = y$ εάν και μόνον εάν η ακολουθία των μερικών αθροισμάτων $\{y^{(\kappa)}\}$, όπου $y^{(\kappa)} = x^{(1)} + x^{(2)} +\ldots+ x^{(\kappa)}$, συγκλίνει στο y καθώς το $\kappa \to \infty$.

Ορισμός 9.20 Η ακολουθία πινάκων $\{A^{(\kappa)}\} = A^{(1)}, A^{(2)},\ldots, A^{(\kappa)}, \ldots$, όπου $A^{(\kappa)} = (a_{ij}^{(\kappa)}) \in \mathbb{C}^{n,n}$, $\kappa = 1,2,\ldots$ συγκλίνει σε ένα πίνακα $A = (a_{ij}) \in \mathbb{C}^{n,n}$, καθώς το $\kappa \to \infty$ και γράφουμε $\lim\limits_{\kappa\to\infty} A^{(\kappa)} = A$ ή $A^{(\kappa)} \underset{\kappa\to\infty}{\to} A$ εάν και μόνον εάν $\lim\limits_{\kappa\to\infty} a_{ij}^{(\kappa)} = a_{ij}$, $i,j = 1(1)n$.

Ορισμός 9.21 Η απειροσειρά πινάκων $\sum\limits_{\kappa=1}^{\infty} A^{(\kappa)}$ συγκλίνει στον πίνακα B και γράφουμε $\sum\limits_{\kappa=1}^{\infty} A^{(\kappa)} = B$ εάν και μόνον εάν η ακολουθία των μερικών αθροισμάτων $\{B^{(\kappa)}\}$, όπου $B^{(\kappa)} = A^{(1)} + A^{(2)} +\ldots+ A^{(\kappa)}$, συγκλίνει στον B καθώς το $\kappa \to \infty$. ▲

Με σκοπό να μελετηθούν οι διάφορες υπολογιστικές διαδικασίες,στις

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

οποίες υπεισέρχονται διανύσματα και πίνακες είναι χρήσιμο να θεωρήσουμε συνδεδεμένο με κάθε διάνυσμα ή πίνακα ένα μη αρνητικό αριθμό, που κατά κάποια έννοια θα μας δίνει ένα μέτρο του "μεγέθους" του . Αυτό γίνεται με την εισαγωγή της έννοιας norm (στάθμη) διανύσματος και πίνακα αντίστοιχα.

Ορισμός 9.22 Μία norm διανύσματος είναι μία πραγματική συνάρτηση $\|\cdot\|$, η οποία ορίζεται για κάθε $x \in \mathbb{C}^n$ και έχει τις ακόλουθες ιδιότητες:

i) $\|x\| \geqq 0$ και $\|x\| = 0$ εάν και μόνον εάν $x = 0$

ii) $\|cx\| = |c|\,\|x\|$, $\forall\, c \in \mathbb{C}$ και κάθε $x \in \mathbb{C}^n$ (9.28)

iii) $\|x + y\| \leqq \|x\| + \|y\|$, $\forall\, x,y \in \mathbb{C}^n$. ▲

Είναι φανερό ότι δια μέσου μιας norm διανύσματος σε κάθε $x \in \mathbb{C}^n$ αντιστοιχεί μία πραγματική μη αρνητική τιμή $\|x\|$ που καλείται norm του x . Υπάρχουν άπειρες norms διανύσματος. Πράγματι, αν θεωρήσουμε $\forall\, x = (x_i) \in \mathbb{C}^n$ τίς συναρτήσεις

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}, \quad p = 1,2,\ldots, \tag{9.29}$$

είναι δυνατόν να δειχθεί ότι αυτές ικανοποιούν τις ιδιότητες (9.28). Πρέπει να σημειωθεί ότι για $p \to \infty$ από την (9.29) μπορεί να προκύψει ότι

$$\|x\|_\infty = \lim_{p \to \infty} \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} = \max_i |x_i| .$$

Θεώρημα 9.71 Για κάθε norm διανύσματος ισχύει

$$\big|\,\|x\| - \|y\|\,\big| \leqq \|x \pm y\| , \quad \forall\, x,y \in \mathbb{C}^n . \tag{9.30}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

<u>Ορισμός 9.23</u> Μία norm πίνακα είναι μία πραγματική συνάρτηση $\|\cdot\|$, η οποία ορίζεται για κάθε $A \in \mathbb{C}^{n,n}$ και έχει τις ακόλουθες ιδιότητες:

$$\begin{aligned}
&\text{i)} \quad \|A\| \geqq 0 \quad \text{και} \quad \|A\| = 0 \quad \text{εάν και μόνον εάν} \quad A = 0 \\
&\text{ii)} \quad \|cA\| = |c|\,\|A\| \;,\; \forall\, c \in \mathbb{C} \quad \text{και} \quad A \in \mathbb{C}^{n,n} \\
&\text{iii)} \quad \|A + B\| \leqq \|A\| + \|B\| \;,\; \forall\, A,B \in \mathbb{C}^{n,n} \\
&\text{iv)} \quad \|AB\| \leqq \|A\|\,\|B\| \;,\; \forall\, A,B \in \mathbb{C}^{n,n} .
\end{aligned} \tag{9.31}$$

Όπως θα φανεί στη συνέχεια, υπάρχουν επίσης άπειρες norms πίνακα.

<u>Θεώρημα 9.72</u> Για κάθε norm πίνακα ισχύει

$$\big|\,\|A\| - \|B\|\,\big| \leqq \|A \pm B\| \;,\; \forall\, A,B \in \mathbb{C}^{n,n} . \qquad \blacktriangle$$

Οι πιο γνωστές norms διανύσματος, που παρουσιάζουν πρακτικό ενδιαφέρον προκύπτουν από την (9.29) για $p = 1,2$ και ∞ και είναι

$$\begin{aligned}
&\|x\|_1 = \sum_{i=1}^{n} |x_i| \,, \\
&\|x\|_2 = \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}} = (x^H x)^{\frac{1}{2}} \,,\quad \|x\|_\infty = \max_i |x_i| \,.
\end{aligned} \tag{9.32}$$

Η $\|x\|_2$ καλείται και Ευκλείδια norm ή μήκος του x, η δε $\|x\|_\infty$ καλείται και μεγίστη norm.

<u>Θεώρημα 9.73</u> Εάν $\|x\|_\alpha$ είναι μία norm διανύσματος, τότε για κάθε $A \in \mathbb{C}^{n,n}$ η συνάρτηση $\sup\limits_{x \neq 0} \dfrac{\|Ax\|_\alpha}{\|x\|_\alpha}$ είναι μία norm πίνακα, που συμβολίζεται $\|A\|_\alpha$ και καλείται η norm πίνακα η εξαρτώμενη (subordinate) από την $\|x\|_\alpha$. $\blacktriangle$

Για την απόδειξη του θεωρήματος 9.73, που παραλείπεται, τονίζεται

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

απλώς ότι κατ' αρχήν με βάση την ιδιότητα (9.31), (ii) μπορούμε να πάρουμε

$$\sup_{x\neq 0} \frac{\|Ax\|_\alpha}{\|x\|_\alpha} = \sup_{\|x\|_\alpha=1} \|Ax\|_\alpha \ . \tag{9.33}$$

Ακόμη, μπορεί να δειχθεί ότι κάθε norm διανύσματος είναι συνεχής συνάρτηση και το σύνολο των διανυσμάτων x για τα οποία $\|x\|_\alpha = 1$ είναι κλειστό και φραγμένο, οπότε η (9.33) γίνεται

$$\sup_{\|x\|_\alpha=1} \|Ax\|_\alpha = \max_{\|x\|_\alpha=1} \|Ax\|_\alpha \ . \tag{9.34}$$

Τέλος, μπορεί να δειχθεί ότι η συνάρτηση $\max_{\|x\|_\alpha=1} \|Ax\|_\alpha$ είναι μία norm πίνακα (ικανοποιεί τις ιδιότητες (9.31)). Έτσι, σε κάθε norm διανύσματος $\|x\|_\alpha$ αντιστοιχεί μία norm πίνακα που ορίζεται από την

$$\|A\|_\alpha = \sup_{x\neq 0} \frac{\|Ax\|_\alpha}{\|x\|_\alpha} = \max_{\|x\|_\alpha=1} \|Ax\|_\alpha \ . \tag{9.35}$$

<u>Ορισμός 9.24</u> Η norm πίνακα $\|A\|_\beta$ και η norm διανύσματος $\|x\|_\alpha$ καλούνται συμβιβαστές εάν

$$\|Ax\|_\alpha \leqq \|A\|_\beta \|x\|_\alpha \ , \tag{9.36}$$

για κάθε $x \in \mathbb{C}^n$ και $A \in \mathbb{C}^{n,n}$. ▲

Είναι προφανές (λόγω της (9.35)) ότι μία norm διανύσματος $\|x\|_\alpha$ και η εξαρτώμενη από αυτή norm πίνακα $\|A\|_\alpha$ είναι πάντοτε συμβιβαστές

Έτσι, εάν δοθεί μία norm διανύσματος, μπορούμε πάντοτε να βρούμε μία norm πίνακα που να είναι συμβιβαστή με αυτή. Ισχύει όμως και το αντίστροφο (βλέπε άσκηση 43) .

Προσδιορίζουμε τώρα τις τρείς norms πίνακα που εξαρτώνται από τις τρείς norms διανύσματος $\|x\|_1, \|x\|_2$ και $\|x\|_\infty$ αντίστοιχα. Θα δειχθεί ότι, αν $A = (a_{ij})$, $i,j = 1(1)n$, τότε

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\|A\|_1 = \max_{\|x\|_1=1} \|Ax\|_1 = \max_j \sum_{i=1}^{n} |a_{ij}| \ , \tag{9.37}$$

$$\|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2 = \left(\rho(A^H A)\right)^{\frac{1}{2}} \ , \tag{9.38}$$

$$\|A\|_\infty = \max_{\|x\|_\infty=1} \|Ax\|_\infty = \max_i \sum_{j=1}^{n} |a_{ij}| \ . \tag{9.39}$$

Για την $\|A\|_1$ έχουμε διαδοχικά,χρησιμοποιώντας τον ορισμό της norm διανύσματος $\|x\|_1$:

$$\|A\|_1 = \max_{\|x\|_1=1} \|Ax\|_1 = \max_{\|x\|_1=1} \sum_{i=1}^{n} \left| \sum_{j=1}^{n} a_{ij} x_j \right| \leq$$

$$\leq \max_{\|x\|_1=1} \sum_{i=1}^{n} \left(\sum_{j=1}^{n} |a_{ij}||x_j| \right) = \max_{\|x\|_1=1} \sum_{j=1}^{n} \left(\sum_{i=1}^{n} |a_{ij}||x_j| \right) =$$

$$= \max_{\|x\|_1=1} \sum_{j=1}^{n} \left(|x_j| \left\{ \sum_{i=1} |a_{ij}| \right\} \right) \leq \max_{\|x\|_1=1} \sum_{j=1}^{n} \left(|x_j| \left\{ \max_j \sum_{i=1}^{n} |a_{ij}| \right\} \right) =$$

$$= \left(\max_j \sum_{i=1}^{n} |a_{ij}| \right) \left(\max_{\|x\|_1=1} \sum_{j=1}^{n} |x_j| \right) = \max_j \sum_{i=1}^{n} |a_{ij}| \ , \quad \text{αφού}$$

$$\|x\|_1 = \sum_{j=1}^{n} |x_j| = 1 \ .$$

Εδείχθη λοιπόν μέχρι τώρα ότι

$$\|A\|_1 \leq \max_j \sum_{i=1}^{n} |a_{ij}| \ . \tag{9.40}$$

Ας υποθέσουμε ότι

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\max_j \sum_{i=1}^{n} |a_{ij}| = \sum_{i=1}^{n} |a_{i\kappa}| \ . \tag{9.41}$$

Εάν εκλέξουμε το διάνυσμα $x = e^{(\kappa)}$, όπου $e_i^{(\kappa)} = 0$ για $i \neq \kappa$ και $e_\kappa^{(\kappa)} = 1$, τότε είναι $\|x\|_1 = \|e^{(\kappa)}\|_1 = 1$ και

$$\|A\|_1 = \max_{\|x\|_1=1} \|Ax\|_1 \geqq \|Ae^{(\kappa)}\|_1 = \sum_{i=1}^{n} |a_{i\kappa}| = \max \sum_{i=1}^{n} |a_{ij}| \ . \tag{9.42}$$

Από τις (9.40) και (9.42) προκύπτει η (9.37).

Για την απόδειξη της (9.39), έχοντας υπ'όψιν ότι $\|x\|_\infty = \max_i |x_i|$, παίρνουμε

$$\|A\|_\infty = \max_{\|x\|_\infty=1} \|Ax\|_\infty = \max_{\|x\|_\infty=1} \left\{ \max_i \left| \sum_{j=1}^{n} a_{ij} x_j \right| \right\} \leqq$$

$$\leqq \max_{\|x\|_\infty=1} \left\{ \max_i \left(\sum_{j=1}^{n} |a_{ij}| \, |x_j| \right) \right\} \leqq \max_{\|x\|_\infty=1} \left\{ \max_i \left(\left[\max_i |x_i| \right] \sum_{j=1}^{n} |a_{ij}| \right) \right\} =$$

$$= \max_i \sum_{j=1}^{n} |a_{ij}| \ , \text{ αφού } \|x\|_\infty = \max_i |x_i| = 1 \ .$$

Άρα,

$$\|A\|_\infty \leqq \max_i \sum_{j=1}^{n} |a_{ij}| \ . \tag{9.43}$$

Ας υποθέσουμε ότι

$$\max_i \sum_{j=1}^{n} |a_{ij}| = \sum_{j=1}^{n} |a_{\kappa j}| \ . \tag{9.44}$$

Εκλέγουμε διάνυσμα $y = (y_j)$, $j = 1(1)n$, τέτοιο ώστε

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$y_j = \begin{cases} \dfrac{|a_{\kappa j}|}{a_{\kappa j}} \ , & \text{αν} \quad a_{\kappa j} \neq 0 \\ 1 \ , & \text{αν} \quad a_{\kappa j} = 0 \ . \end{cases} \tag{9.45}$$

Για το διάνυσμα y που ορίζεται από την (9.45) ισχύει $|y_j| = 1$, άρα, και $\|y\|_\infty = \max_i |y_i| = 1$ και επομένως

$$\|A\|_\infty = \max_{\|x\|_\infty = 1} \|Ax\|_\infty \geqq \|Ay\|_\infty = \max_i \left| \sum_{j=1}^n a_{ij} y_j \right| =^*$$

$$=^* \left| \sum_{j=1}^n a_{\kappa j} y_j \right| = \sum_{j=1}^n |a_{\kappa j}| = \max_i \sum_{j=1}^n |a_{ij}| \ . \tag{9.46}$$

* Το = αυτό ισχύει διότι $\left| \sum_{j=1}^n a_{ij} y_j \right| \leqq \sum_{j=1}^n |a_{ij}||y_j| = \sum_{j=1}^n |a_{ij}| \leqq$

$\leqq \sum_{j=1}^n |a_{\kappa j}| = \left| \sum_{j=1}^n a_{\kappa j} y_j \right|$. Από τις (9.44) και (9.46) προκύπτει η (9.39).

Για την απόδειξη της (9.38) έχουμε

$$\|A\|_2 = \max_{\|x\|_2 = 1} \|Ax\|_2 = \max_{\|x\|_2 = 1} \left[(Ax)^H Ax \right]^{\frac{1}{2}} = \max_{\|x\|_2 = 1} (x^H A^H A x)^{\frac{1}{2}} =$$

$$= \left[\max_{\|x\|_2 = 1} (x^H A^H A x) \right]^{\frac{1}{2}} \ . \tag{9.47}$$

Σύμφωνα με τα θεωρήματα 9.48 και 9.49 ο πίνακας $A^H A$ είναι Ερμιτιανός και όλες οι ιδιοτιμές του είναι μη αρνητικές. Έστωσαν μ_m, μ_M η μικρότερη και η μεγαλύτερη αντίστοιχα, από τις ιδιοτιμές του. Σύμφωνα με το θεώρημα 9.66, (5) για κάθε $x \in \mathbb{C}^n$ με $\|x\|_2 = 1$ θα ισχύει $0 \leqq \mu_m \leqq x^H A^H A x \leqq$ $\leqq \mu_M$ και $\max_{\|x\|_2 = 1} (x^H A^H A x) = \mu_M$, οπότε η (9.47) γίνεται $\|A\|_2 = \mu_M^{\frac{1}{2}} =$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$= \left[\rho(A^H A)\right]^{\frac{1}{2}}$, αφού η μ_M είναι προφανώς και η φασματική ακτίνα του $A^H A$. Σύμφωνα με τον ορισμό 9.9, η $\|A\|_2$ ισούται με την μεγαλύτερη μη ομαλή τιμή του A . Η $\|A\|_2$ καλείται επίσης και φασματική norm του A .

Μία άλλη γνωστή norm πίνακα είναι η καλούμενη Ευκλείδια norm, που ορίζεται από την

$$\|A\|_E = \left[\sum_{j=1}^{n} \sum_{i=1}^{n} |a_{ij}|^2\right]^{\frac{1}{2}} = \left[tr(A^H A)\right]^{\frac{1}{2}} .$$

Μπορεί να αποδειχθεί ότι η $\|A\|_E$ είναι συμβιβαστή με την $\|x\|_2$, δηλαδή, ισχύει

$$\|Ax\|_2 \leqq \|A\|_E \|x\|_2 \ , \ \forall\, x \in \mathbb{C}^n \quad \text{και} \quad A \in \mathbb{C}^{n,n} \tag{9.48}$$

και ότι η $\|A\|_E$ δεν είναι εξαρτώμενη από καμμία norm διανύσματος.

<u>Θεώρημα 9.74</u> Για κάθε $A \in \mathbb{C}^{n,n}$ και για οποιαδήποτε norm πίνακα ισχύει $\rho(A) \leqq \|A\|$.

<u>Απόδειξη</u> Έστω λ_i μία οποιαδήποτε ιδιοτιμή του A και $x^{(i)}$ ένα αντίστοιχο ιδιοδιάνυσμα $(x^{(i)} \neq 0)$. Τότε, έχουμε

$$\lambda_i x^{(i)} = A x^{(i)} . \tag{9.49}$$

Εκλέγοντας τώρα μία οποιαδήποτε norm πίνακα $\|\cdot\|$ και μία συμβιβαστή μ' αυτή norm διανύσματος $\|\cdot\|$, από την (9.49) παίρνουμε

$$\|\lambda_i x^{(i)}\| = |\lambda_i| \|x^{(i)}\| = \|Ax^{(i)}\| \leqq \|A\| \|x^{(i)}\| \ , \text{ οπότε}$$

$$|\lambda_i| \leqq \|A\| . \tag{9.50}$$

Η (9.50) ισχύει για κάθε ιδιοτιμή λ_i του A , που σημαίνει ότι $\rho(A) \leqq \|A\|$. ▲

<u>Θεώρημα 9.75</u> Έστω $\{x^{(K)}\}$ μία ακολουθία διανυσμάτων. Τότε ,

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$x^{(\kappa)} \underset{\kappa\to\infty}{\to} x$ εάν και μόνον εάν $\|x^{(\kappa)} - x\| \underset{\kappa\to\infty}{\to} 0$ για οποιαδήποτε norm διανύσματος.

<u>Θεώρημα 9.76</u> Έστω $\{A^{(\kappa)}\}$ μία ακολουθία πινάκων. Τότε, $A^{(\kappa)} \underset{\kappa\to\infty}{\to} A$ εάν και μόνον εάν $\|A^{(\kappa)} - A\| \underset{\kappa\to\infty}{\to} 0$ για οποιαδήποτε norm πίνακα.

<u>Θεώρημα 9.77</u> Για κάθε norm πίνακα ισχύει

$$\|A^{\kappa}\| \leqq \|A\|^{\kappa}, \quad \forall A \in \mathbb{C}^{n,n}, \quad \kappa = 1,2,\ldots .$$

<u>Απόδειξη</u> Για $\kappa = 1$ ισχύει η ισότητα. Για $\kappa > 1$ έχουμε

$$\|A^{\kappa}\| = \|AA^{\kappa-1}\| \leqq \|A\|\|A^{\kappa-1}\| \leqq \|A\|^2\|A^{\kappa-2}\| \leqq \ldots \leqq \|A\|^{\kappa} . \qquad \blacktriangle$$

<u>Θεώρημα 9.78</u> Εάν για κάποια norm πίνακα είναι $\|A\| < 1$, τότε $A^{\kappa} \underset{\kappa\to\infty}{\to} 0$, δηλαδή, η ακολουθία των διαδοχικών δυνάμεων $\{A^{\kappa}\}$ του πίνακα $A \in \mathbb{C}^{n,n}$ συγκλίνει στον μηδενικό πίνακα.

<u>Απόδειξη</u> Εάν $\|A\| < 1$, τότε από την $\|A^{\kappa}\| \leqq \|A\|^{\kappa}$ προκύπτει ότι $\|A\|^{\kappa} \underset{\kappa\to\infty}{\to} 0$ και $\|A^{\kappa}\| \underset{\kappa\to\infty}{\to} 0$, οπότε από το θεώρημα 9.76 προκύπτει ότι $A^{\kappa} \underset{\kappa\to\infty}{\to} 0$. $\blacktriangle$

<u>Θεώρημα 9.79</u> Για κάθε πίνακα $A \in \mathbb{C}^{n,n}$ και κάθε $\varepsilon > 0$ υπάρχει μία norm διανύσματος $\|x\|_{\alpha}$, τέτοια ώστε για την εξαρτώμενη απ' αυτή norm πίνακα $\|A\|_{\alpha}$ να ισχύει

$$\|A\|_{\alpha} \leqq \rho(A) + \varepsilon .$$

<u>Απόδειξη</u> Σύμφωνα με το θεώρημα 9.30 υπάρχει ένας ομαλός πίνακας T, τέτοιος ώστε $T^{-1}AT = J$, όπου J είναι η κανονική μορφή Jordan του A. Κάθε Jordan block έχει την μορφή (βλέπε (9.20))

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$J_{\nu_j}(\kappa) = \begin{bmatrix} \lambda_\kappa & 1 & & & \\ & \lambda_\kappa & 1 & & \bigcirc \\ & & \ddots & \ddots & \\ & \bigcirc & & & 1 \\ & & & & \lambda_\kappa \end{bmatrix} . \qquad (9.51)$$

Για κάθε $\varepsilon > 0$ ορίζουμε τον ακόλουθο διαγώνιο πίνακα $D \in \mathbb{C}^{n,n}$

$$D = \begin{bmatrix} 1 & & & & \\ & \varepsilon & & & \bigcirc \\ & & \varepsilon^2 & & \\ & \bigcirc & & \ddots & \\ & & & & \varepsilon^{n-2} \\ & & & & & \varepsilon^{n-1} \end{bmatrix}$$

και σχηματίζουμε τον πίνακα $\hat{J} = D^{-1}JD = D^{-1}(T^{-1}AT)D = (TD)^{-1}A(TD)$. Ο $\hat{J}$ προκύπτει από τον J αν το Jordan block (9.51) αντικατασταθεί από το

$$\begin{bmatrix} \lambda_\kappa & \varepsilon & & & \\ & \lambda_\kappa & \varepsilon & & \bigcirc \\ & & \ddots & \ddots & \\ & \bigcirc & & & \varepsilon \\ & & & & \lambda_\kappa \end{bmatrix} \qquad (9.53)$$

Έτσι, έχουμε $\|\hat{J}\|_\infty = \|D^{-1}JD\|_\infty = \|(TD)^{-1}A(TD)\|_\infty \leq \rho(A) + \varepsilon$. Θέτοντας $(TD)^{-1}=L$, ορίζουμε την norm διανύσματος $\|x\|_{\infty,L}$ (βλέπε και άσκηση 45)

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\|x\|_{\infty,L} = \|Lx\|_{\infty} \quad . \tag{9.54}$$

Για την εξαρτώμενη από την (9.54) norm πίνακα (βλέπε άσκηση 45) ισχύει επομένως,

$$\|A\|_{\infty,L} = \|LAL^{-1}\|_{\infty} = \|(TD)^{-1}A(TD)\|_{\infty} \leqq \rho(A) + \varepsilon \ . \ \blacktriangle \tag{9.55}$$

<u>Θεώρημα 9.80</u> $A^{K} \underset{K\to\infty}{\to} 0$ εάν και μόνον εάν $\rho(A) < 1$.

<u>Απόδειξη</u> i) Έστω ότι $A^{K} \underset{K\to\infty}{\to} 0$. Εάν υποθέσουμε ότι $\rho(A) \geqq 1$, τότε $\forall\, \kappa$ είναι $\rho(A^{K}) \geqq 1$ και σύμφωνα με το θεώρημα 9.74 ισχύει $\|A^{K}\| \geqq \rho(A^{K}) \geqq 1$, για οποιαδήποτε norm πίνακα. Αυτό όμως είναι άτοπο, διότι σύμφωνα με το θεώρημα 9.76 η $A^{K} \underset{K\to\infty}{\to} 0$ συνεπάγεται ότι $\|A^{K}\| \underset{K\to\infty}{\to} 0$.

ii) Έστω $\rho(A) < 1$. Εφαρμόζοντας το θεώρημα 9.79 με $\varepsilon = \dfrac{1-\rho(A)}{2}$, έχουμε ότι (βλέπε (9.55))

$$\|A\|_{\infty,L} \leqq \rho(A) + \frac{1-\rho(A)}{2} = \frac{1+\rho(A)}{2} < 1 \ .$$

Σύμφωνα τώρα με το θεώρημα 9.78, προκύπτει ότι $A^{K} \underset{K\to\infty}{\to} 0$. $\blacktriangle$

<u>Θεώρημα 9.81</u> Η απειροσειρά $\sum_{\kappa=0}^{\infty} A^{K} = I + A + A^{2} + \ldots + A^{K} + \ldots,$ συγκλίνει εάν και μόνον εάν $\rho(A) < 1$. Επί πλέον, όταν αυτή συγκλίνει έχουμε

$$\sum_{\kappa=0}^{\infty} A^{K} = (I - A)^{-1} \ . \tag{9.56}$$

<u>Απόδειξη</u> i) Έστω $\rho(A) < 1$. Τότε, η $\lambda = 1$ δεν μπορεί να είναι ιδιοτιμή του A και επομένως ο πίνακας $I - A$ είναι ομαλός. Θεωρούμε την ταυτότητα

$$(I - A)(I + A^{2} + A + \ldots + A^{K}) \equiv 1 - A^{K+1} \ , \ \kappa = 0,1,2,\ldots, \tag{9.57}$$

από την οποία προκύπτει ότι

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$I + A + A^2 + \dots + A^K \equiv (I - A)^{-1} - (I - A)^{-1}A^{K+1}. \quad (9.58)$$

Αφού $\rho(A) < 1$, σύμφωνα με το θεώρημα 9.80 , είναι $A^K \underset{K\to\infty}{\to} 0$, οπότε από την (9.58) παίρνουμε

$$(I + A + A^2 + \dots + A^K) \underset{K\to\infty}{\to} (I - A)^{-1}.$$

Άρα,

$$\sum_{K=0}^{\infty} A^K = (I - A)^{-1}. \quad \text{(βλέπε ορισμό 9.21)}$$

ii) Έστω ότι η απειροσειρά συγκλίνει στον πίνακα B . Τότε, σύμφωνα με τον ορισμό 9.21 η ακολουθία $\{B^{(K)}\}$, όπου $B^{(K)} = I + A + A^2 + \dots + A^K$, συγκλίνει στον B , δηλαδή, $B^{(K)} \underset{K\to\infty}{\to} B$.
Έχουμε, $A^K = B^{(K)} - B^{(K-1)}$ και επομένως

$$\lim_{K\to\infty} A^K = \lim_{K\to\infty} B^{(K)} - \lim_{K\to\infty} B^{(K-1)} = B - B = 0 .$$ Άρα,

$A^K \underset{K\to\infty}{\to} 0$, οπότε σύμφωνα με το θεώρημα 9.80 είναι $\rho(A) < 1$ και συνεπώς

$$B = (I - A)^{-1},$$

σύμφωνα με την περίπτωση (i) . ▲

<u>Θεώρημα 9.82</u> Εάν για κάποια norm πίνακα ισχύει $\|A\|_\beta < 1$, τότε

i) $\sum_{K=0}^{\infty} A^K = (I - A)^{-1}$

ii) $\|(I - A)^{-1} - (I + A + A^2 + \dots + A^K)\|_\beta \leq \dfrac{\|A\|_\beta^{K+1}}{1 - \|A\|_\beta}$.

<u>Απόδειξη</u> i) Προκύπτει από τα θεωρήματα 9.74 και 9.81 .
ii) Είναι

$$(I - A)^{-1} - (I + A + A^2 + \dots + A^K) = A^{K+1} + A^{K+2} + \dots$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

και επομένως

$$\|(I-A)^{-1}-(I+A+A^2+\dots+A^K)\|_\beta = \|A^{K+1}+A^{K+2}+\dots\|_\beta \leq$$

$$\leq \|A\|_\beta^{K+1}+\|A\|_\beta^{K+2}+\dots = \|A\|_\beta^{K+1}(1+\|A\|_\beta+\|A\|_\beta^2+\dots) =$$

$$= \frac{\|A\|_\beta^{K+1}}{1-\|A\|_\beta}. \qquad \blacktriangle$$

<u>Θεώρημα 9.83</u> $A^K v \underset{K\to\infty}{\to} 0$ για κάθε $v \in \mathbb{C}^n$ εάν και μόνον εάν $\rho(A) < 1$.

<u>Απόδειξη</u> i) Εάν $\rho(A) < 1$, τότε από το θεώρημα 9.80 είναι $A^K \underset{K\to\infty}{\to} 0$, οπότε και

$$A^K v \underset{K\to\infty}{\to} 0 \quad \text{για κάθε } v \in \mathbb{C}^n$$

ii) Έστω ότι $A^K v \underset{K\to\infty}{\to} 0$ για κάθε $v \in \mathbb{C}^n$. Εάν το v εκλεγεί έτσι ώστε να είναι ιδιοδιάνυσμα του A, που αντιστοιχεί στην ιδιοτιμή λ με $|\lambda| = \rho(A)$, τότε $A^K v = \lambda^K v$ και $\lambda^K v \underset{K\to\infty}{\to} 0$, οπότε $|\lambda| < 1$ και συνεπώς $\rho(A) < 1$.

<u>Θεώρημα 9.84</u> Εάν για την norm πίνακα $\|A\|_\alpha$, που είναι εξαρτώμενη κάποιας norm διανύσματος, ισχύει $\|A\|_\alpha < 1$, τότε

$$\frac{1}{1+\|A\|_\alpha} \leq \|(I \pm A)^{-1}\|_\alpha \leq \frac{1}{1-\|A\|_\alpha}. \tag{9.59}$$

<u>Απόδειξη</u> Αφού η norm πίνακα $\|A\|_\alpha$ είναι εξαρτώμενη από κάποια norm διανύσματος θα ισχύει $\|I\|_\alpha = 1$. Εάν $\|A\|_\alpha < 1$, τότε σύμφωνα με το θεώρημα 9.74 είναι $\rho(A) < 1$ και προφανώς οι πίνακες $I \pm A$ είναι ομαλοί. Θα συνεχίσουμε την απόδειξη με τον πίνακα $I - A$, αφού πάντοτε μπορούμε να αντικαταστήσουμε τον A με $-A$. Από την ισότητα

$$I = (I-A)(I-A)^{-1},$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

προκύπτει ότι

$$1 = \|I\|_\alpha = \|(I-A)(I-A)^{-1}\|_\alpha \leqq \|I-A\|_\alpha \|(I-A)^{-1}\|_\alpha \leqq$$

$$\leqq \left(\|I\|_\alpha + \|-A\|_\alpha\right)\|(I-A)^{-1}\|_\alpha = \left(1 + \|A\|_\alpha\right)\|(I-A)^{-1}\|_\alpha$$

και διαιρώντας με τον παράγοντα $1 + \|A\|_\alpha$ προκύπτει η αριστερή ανισότητα (9.59) . Εξ άλλου ,με βάση το θεώρημα 9.82 έχουμε

$$\|(I-A)^{-1}\|_\alpha = \|\sum_{\kappa=0}^{\infty} A^\kappa\|_\alpha \leqq \|I\|_\alpha + \|A\|_\alpha + \|A^2\|_\alpha + \dots +$$

$$+ \|A^\kappa\|_\alpha + \dots \leqq 1 + \|A\|_\alpha + \|A\|_\alpha^2 + \dots + \|A\|_\alpha^\kappa + \dots = \frac{1}{1-\|A\|_\alpha} .$$

<u>Θεώρημα 9.85</u> Εάν $A \in \mathbb{C}^{n,n}$, τότε

$$\rho(A) = \lim_{\kappa \to \infty} (\|A^\kappa\|_\alpha)^{1/\kappa} . \qquad \blacktriangle$$

Για τα επόμενα κεφάλαια υποθέτουμε ότι $\|x\|$, $x \in \mathbb{C}^n$ συμβολίζει κάποια από τις norms διανύσματος (9.32) και $\|A\|$ την αντίστοιχη εξαρτώμενη norm του πίνακα $A \in \mathbb{C}^{n,n}$.

Α Σ Κ Η Σ Ε Ι Σ

<u>Σημείωση</u> Στις ακόλουθες ασκήσεις υποθέτουμε ότι οι πίνακες ανήκουν στο $\mathbb{C}^{n,n}$ και τα διανύσματα στο $\mathbb{C}^n$, εκτός εάν διαφορετικά καθορίζεται.

1. Εάν οι A,B είναι συμμετρικοί (ή Ερμιτιανοί) πίνακες, να δειχθεί ότι ο AB είναι συμμετρικός (Ερμιτιανός) εάν και μόνον εάν $AB = BA$.

2. Να δειχθεί ότι κάθε πίνακας A μπορεί να γραφεί: i) σαν άθροισμα ενός συμμετρικού και ενός αντισυμμετρικού πίνακα ii) υπό την μορφή

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$A = B + iC$, όπου B,C είναι Ερμιτιανοί πίνακες.

3. Εάν L είναι ένας $n \times n$ αυστηρά κάτω (ή άνω) τριγωνικός πίνακας, να δειχθεί ότι $L^n = 0$.

4. Ένας πίνακας $A = (a_{ij})$, $i,j = 1(1)n$ καλείται unit τριγωνικός, εάν είναι τριγωνικός και $a_{ii} = 1$, $i = 1(1)n$. Να δειχθεί ότι ο αντίστροφος ενός τέτοιου πίνακα είναι επίσης unit τριγωνικός.

5. Να δειχθεί ότι κάθε ορθογώνιο σύνολο διανυσμάτων $v^{(i)}$, $i = 1(1)\kappa$, $v^{(i)} \neq 0$, είναι γραμμικώς ανεξάρτητο.

6. Να βρεθεί η γενική μορφή ενός 2×2 πίνακα A όταν: i) $A^H = -A$ ii) $A^{-1} = A^H$.

7. Να δειχθεί ότι ο πίνακας A είναι Ερμιτιανός εάν και μόνον εάν $A = R + iS$, όπου ο R είναι πραγματικός συμμετρικός και ο S πραγματικός αντισυμμετρικός.

8. Εάν $A,B \neq 0$ και $AB = 0$, να δειχθεί ότι $\det(A) = \det(B) = 0$.

9. Εάν A είναι ομαλός πίνακας, να δειχθεί ότι $Ax \neq 0$, για κάθε διάνυσμα $x \neq 0$.

10. Αν $A \neq 0$, να δειχθεί ότι υπάρχει $x \neq 0$ τέτοιο ώστε $Ax \neq 0$.

11. Εάν λ είναι μία ιδιοτιμή του πίνακα A και v ένα ιδιοδιάνυσμα αυτού που αντιστοιχεί στην λ , να δειχθεί ότι: i) $(cA)v = (c\lambda)v$, $c \in \mathbb{C}$ ii) $(A - \mu I)v = (\lambda - \mu)v$, $\mu \in \mathbb{C}$ iii) $A^{-1}v = \frac{1}{\lambda} v$ (με την προυπόθεση ότι ο A είναι ομαλός) iv) $A^m v = \lambda^m v$, m θετικός ακέραιος.

12. Έστω $A \in \mathbb{R}^{n,n}$. Να δειχθεί ότι: i) Εάν λ είναι μία πραγματική ιδιοτιμή του, τότε ο A έχει ένα πραγματικό ιδιοδιάνυσμα που αντιστοιχεί στην λ . ii) Εάν λ είναι μία μιγαδική ιδιοτιμή του, τότε ο A έχει ένα μιγαδικό ιδιοδιάνυσμα x που αντιστοιχεί στην λ και η λ^* είναι επίσης ιδιοτιμή του A , με αντίστοιχο ιδιοδιάνυσμα το x^* .

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

13. Εάν λ είναι μία ιδιοτιμή του A , να δειχθεί ότι η λ^* είναι ιδιοτιμή του A^H .

14. Εάν $u^{(i)}$ είναι ιδιοδιάνυσμα του A που αντιστοιχεί στην ιδιοτιμή λ_i αυτού και $v^{(\kappa)}$ ιδιοδιάνυσμα του A^T που αντιστοιχεί στην ιδιοτιμή λ_κ αυτού με $\lambda_i \neq \lambda_\kappa$, να δειχθεί ότι $u^{(i)T}v^{(\kappa)} = 0$.

15. Εάν λ είναι μία ιδιοτιμή του A , y ιδιοδιάνυσμα του A^T που αντιστοιχεί στην ίδια ιδιοτιμή λ αυτού και z ιδιοδιάνυσμα του A^H που αντιστοιχεί στην λ^* , να δειχθεί ότι $y^T = z^H$ και $z^H A = \lambda z^H$.

16. Υποθέτουμε ότι ο πίνακας A έχει διακεκριμένες ιδιοτιμές λ_i , $i = 1(1)n$. Εάν $u^{(i)}$ είναι ένα ιδιοδιάνυσμα του A που αντιστοιχεί στην ιδιοτιμή λ_i αυτού και $v^{(j)}$ ένα ιδιοδιάνυσμα του A^H που αντιστοιχεί στην ιδιοτιμή λ_j^* αυτού, να δειχθεί ότι $u^{(i)H}v^{(j)} = 0$, $i \neq j$. Να δειχθεί επίσης ότι τα $v^{(i)}$, $i = 1(1)n$ μπορούν να εκλεγούν έτσι ώστε $u^{(i)H}v^{(i)} = 1$.

17. Έστω A ένας πίνακας και λ_i , $i = 1(1)n$ οι ιδιοτιμές του με αντίστοιχα ιδιοδιανύσματα $x^{(i)}$, $i = 1(1)n$. Αν X είναι ο πίνακας με στήλες τα $x^{(i)}$, $i = 1(1)n$, να δειχθεί ότι $AX = X\Lambda$,

όπου $\Lambda = \begin{bmatrix} \lambda_1 & & \bigcirc \\ & \lambda_2 & \\ \bigcirc & & \ddots \\ & & & \lambda_n \end{bmatrix}$. Ο πίνακας X καλείται modal πίνακας

του A . Αν τα $x^{(i)}$, $i = 1(1)n$ είναι γραμμικώς ανεξάρτητα, τότε $A = X\Lambda X^{-1}$.

18. Εάν στην i γραμμή του $A = (a_{ij})$ το μόνο μη μηδενικό στοιχείο είναι το a_{ii} , να δειχθεί ότι το a_{ii} είναι μία ιδιοτιμή του A.

19. Εάν A,B είναι Ερμιτιανοί, να δειχθεί ότι $AB = BA$ εάν και μόνον εάν οι A,B έχουν τους ίδιους modal πίνακες.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

20. Αν μ_i είναι οι ιδιοτιμές του A , ρ_i οι ιδιοτιμές του B , λ_i οι ιδιοτιμές του $A + B$ και ξ_i οι ιδιοτιμές του AB , $i = 1(1)n$, να δειχθεί ότι

$$\text{i)} \quad \sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} (\mu_i + \rho_i)$$

$$\text{ii)} \quad \prod_{i=1}^{n} \xi_i = \prod_{i=1}^{n} (\mu_i \rho_i) \ .$$

21. Εάν ο πίνακας A έχει ιδιοτιμές λ_i , $i = 1(1)n$ με αντίστοιχα ιδιοδιανύσματα $x^{(i)}$, $i = 1(1)n$, να δειχθεί ότι ο πίνακας $[Q(A)][R(A)]^{-1}$, $\det(R(A)) \neq 0$, όπου Q,R είναι πολυώνυμα,έχει ιδιοτιμές τις $\dfrac{Q(\lambda_i)}{R(\lambda_i)}$, $i = 1(1)n$, με αντίστοιχα ιδιοδιανύσματα $x^{(i)}$, $i = 1(1)n$.

22. Δίνεται ο $n \times n$ πραγματικός τριδιαγώνιος πίνακας

$$A = \begin{bmatrix} a_1 & b_2 & & & \\ c_2 & a_2 & b_3 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & b_n \\ & & & c_n & a_n \end{bmatrix} \text{, με } b_i c_i > 0 \ , \ i = 2(1)n \ .$$

Να δειχθεί ότι i) ο A έχει πραγματικές ιδιοτιμές ii) Αν $a_i = a$, $b_i = b$, $c_i = c$, τότε οι ιδιοτιμές του A είναι $\lambda_j = a + 2\sqrt{bc} \cos \dfrac{j\pi}{n+1}$, $j = 1(1)n$.

23. Να δειχθεί ότι κάθε ιδιοτιμή του Jordan block $J_{\nu_j(\kappa)}$ (βλέπε (9.20)), που αντιστοιχεί στην ιδιοτιμή λ_κ αυτού είναι της μορφής $\alpha e^{(1)}$ όπου $e^{(1)} = (1, 0, 0, \ldots, 0)^T$ και $\alpha \in \mathbb{C}$, $\alpha \neq 0$, δηλαδή, η γεωμετρική πολλαπλότης της λ_κ είναι 1 .

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

24. Να δειχθεί ότι κάθε πραγματικός πίνακας που είναι όμοιος με ένα πραγματικό συμμετρικό πίνακα έχει n γραμμικώς ανεξάρτητα πραγματικά ιδιοδιανύσματα.

25. Δίνεται ο πραγματικός πίνακας $A = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$. Πότε ισχύει: i) $\rho(A^2) = \rho(A^T A)$; ii) $A^T A = AA^T$;

26. Αν $A = \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix}$, A πραγματικός, να δειχθεί ότι $\rho(A) \leqq \sqrt{\alpha^2 + 2\beta^2 + \gamma^2}$. Πότε ισχύει το = ;

27. Εάν λ_i , $i = 1(1)n$ είναι οι ιδιοτιμές του A και μ_m, μ_M η μικρότερη και η μεγαλύτερη αντίστοιχα ιδιοτιμή του $A^H A$, να δειχθεί ότι $\mu_m \leqq |\lambda_i|^2 \leqq \mu_M$.

28. Αν A είναι normal πίνακας και λ μία ιδιοτιμή του, να δειχθεί ότι η $\lambda\lambda^* = |\lambda|^2$ είναι ιδιοτιμή του AA^H .

29. Αν A είναι ομαλός, να δειχθεί ότι υπάρχει ένας πίνακας μεταθέσεως P , τέτοιος ώστε τα διαγώνια στοιχεία του PA να είναι $\neq 0$.

30. Εάν ο $A = (a_{ij})$, $i,j = 1(1)n$ είναι αυστηρά διαγώνια υπέρτερος, να δειχθεί ότι $\det(A) \neq 0$ και $a_{ii} \neq 0$, $i = 1(1)n$.

31. Αν A,B είναι Ερμιτιανοί και θετικά ορισμένοι και $AB = BA$, να δειχθεί ότι οι πίνακες $A^{-1}B$ και AB είναι Ερμιτιανοί και θετικά ορισμένοι.

32. Αν λ_i , $i = 1(1)n$ είναι οι ιδιοτιμές του A και ο $A + A^H$ είναι θετικά ορισμένος, να δειχθεί ότι $\mathrm{Re}\lambda_i > 0$, $i = 1(1)n$.

33. Εάν ο A είναι πραγματικός συμμετρικός και θετικά ορισμένος $n \times n$ πίνακας και $b \in \mathbb{R}^n$, να δειχθεί ότι $\min\limits_{x \in \mathbb{R}^n} \left[\frac{1}{2} x^T A x - x^T b\right] = -\frac{1}{2} \hat{x}^T b$, όπου $\hat{x}$ τέτοιο ώστε $A\hat{x} = b$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

34. Αν ο $A = (a_{ij})$, $i,j = 1(1)n$ είναι πραγματικός συμμετρικός, να δειχθεί ότι $\lambda_m \leqq \min_i a_{ii}$ και $\lambda_M \geqq \max_i a_{ii}$, όπου λ_m, λ_M η μικρότερη και η μεγαλύτερη αντίστοιχα ιδιοτιμή του A .

35. Αν $A,B \in \mathbb{C}^{n,n}$, να δειχθεί ότι i) $|A + B| \leqq |A| + |B|$ ii) $|AB| \leqq |A||B|$ iii) $|\lambda A| = |\lambda||A|$, $\lambda \in \mathbb{C}$ iv) $|A^K| \leqq |A|^K$, κ φυσικός v) $\rho(AB) \leqq \rho(|A||B|)$ vi) $\rho(A + B) \leqq \rho(|A| + |B|)$.

36. Για κάθε $n \times n$ πίνακα A , να δειχθεί ότι i) $\rho(A) \geqq \dfrac{|tr(A)|}{n}$ ii) $\rho(A^K) = [\rho(A)]^K$, κ φυσικός.

37. Για τις norms διανύσματος (9.32) να δειχθεί ότι

i) $\|x\|_\infty \leqq \|x\|_2 \leqq \|x\|_1 \leqq \sqrt{n}\,\|x\|_2 \leqq n\,\|x\|_\infty$ ii) $\|x\|_2^2 \leqq \|x\|_1\|x\|_\infty$.

38. Αν $x \in \mathbb{R}^2$, να γίνει η γραφική παράσταση των: i) $\|x\|_1 \leqq 1$ ii) $\|x\|_2 \leqq 1$ iii) $\|x\|_\infty \leqq 1$.

39. Δίνεται η ακολουθία διανυσμάτων $\{x^{(K)}\}$. Εάν $x^{(K)} \underset{K\to\infty}{\to} x$, να δειχθεί ότι $\|x^{(K)}\| \underset{K\to\infty}{\to} \|x\|$, για οποιαδήποτε norm διανύσματος.

40. Για την $\|\cdot\|_2$ να δειχθεί η (9.28) (iii). Να δειχθεί ότι το = ισχύει εάν και μόνον εάν $y = \alpha x$, όπου α σταθερά $\geqq 0$.

41. Εάν x,y είναι ορθογώνια διανύσματα, να δειχθεί ότι $\|x + y\|_2^2 = \|x\|_2^2 + \|y\|_2^2$.

42. Έστω $\|x\|_\alpha$ μία norm διανύσματος. Εάν $\|A\|_\alpha$ είναι η norm πίνακα η εξαρτώμενη από την $\|x\|_\alpha$ και $\|A\|_\beta$ μία norm συμβιβαστή με την $\|x\|_\alpha$, να δειχθεί ότι $\|A\|_\alpha \leqq \|A\|_\beta$.

43. Έστω $\|A\|_\beta$ μία norm πίνακα. Τότε, υπάρχει μία norm διανύσματος $\|x\|_\alpha$ που είναι συμβιβαστή με την $\|A\|_\beta$, δηλαδή, $\|Ax\|_\alpha \leqq \|A\|_\beta\|x\|_\alpha$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

44. Μία norm πίνακα $\|A\|$, $A = (a_{ij})$, $i,j = 1(1)n$ καλείται κανονική εάν i) $|a_{ij}| \leqq \|A\|$ και για $n = 1$ $\|A\| = |a_{11}|$, ii) $|A| \leqq$ $\leqq |B| \Rightarrow \|A\| \leqq \|B\|$ και iii) $\|A\| = \| |A| \|$. Να δειχθεί ότι οι $\|A\|_1$, $\|A\|_\infty$, $\|A\|_E$ είναι κανονικές.

45. Αν $\|x\|_\alpha$ είναι μία norm διανύσματος και L ένας οποιοσδήποτε ομαλός πίνακας, τότε η $\|x\|_{\alpha,L} = \|Lx\|_\alpha$ είναι μία norm διανύσματος (καλείται α,L – norm). Να δειχθεί ότι για την εξαρτώμενη norm πίνακα $\|A\|_{\alpha,L}$ ισχύει $\|A\|_{\alpha,L} = \|LAL^{-1}\|_\alpha$.

46. Αν $A,B \in \mathbb{C}^{n,n}$, $x,y \in \mathbb{C}^n$, να δειχθούν οι σχέσεις i) $\|A\|_2 \leqq$ $\leqq \| |A| \|_2 \leqq \|A\|_E \leqq \sqrt{n}\, \|A\|_2$ ii) $\|AB\|_E \leqq \|A\|_E \|B\|_2$ iii) $\|AB\|_E \leqq$ $\leqq \|A\|_2 \|B\|_E$ iv) $\|A^H\|_\infty = \|A\|_1$, $\|A^H\|_2 = \|A\|_2$, $\|A^H\|_E = \|A\|_E$ v) $\|A\|_2^2 \leqq \|A\|_1 \|A\|_\infty$ vi) $|x^H Ay| \leqq \|A\|_2 \|x\|_2 \|y\|_2$ vii) $\|A^H A\|_2 = \|A\|_2^2$.

47. Αν A είναι normal πίνακας, να δειχθεί ότι $\|A\|_2 = \rho(A)$.

48. Αν $D = (d_{ij})$, είναι διαγώνιος πίνακας, να δειχθεί ότι i)

$$\|D\|_1 = \|D\|_2 = \|D\|_\infty \quad \text{ii)} \quad \min_{x \neq 0} \frac{\|Dx\|_2}{\|x\|_2} = \min_i |d_{ii}| \ , \ x \in \mathbb{C}^n .$$

49. Αν $A = \begin{bmatrix} \alpha & 1 & 1 \\ 1 & \alpha & 1 \\ 1 & 1 & \alpha \end{bmatrix}$, $\alpha \geqq 0$, να δειχθεί ότι $\|A\|_1 = \|A\|_2 = \|A\|_\infty$.

50. Εάν $A = (a_{ij})$, $i,j = 1(1)n$ είναι αυστηρά διαγώνια υπέρτερος, να δειχθεί ότι

$$\|A^{-1}\|_\infty \leqq \frac{1}{\min_i \left(|a_{ii}| - \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \right)} .$$

51. Να δειχθεί ότι $\|A^{-1} - B^{-1}\| \leqq \|A^{-1}\| \|B^{-1}\| \|A - B\|$, όπου A,B ομα-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

λοί πίνακες.

52. Αν για κάποια norm πίνακα ισχύει $\|A^m\| < 1$, $m \geqq 1$, να δειχθεί ότι $\rho(A) < 1$.

53. Να δειχθεί ότι i) Για οποιαδήποτε norm πίνακα $\|\cdot\|$ ισχύει $\|I\| \geqq 1$. ii) Για κάθε norm πίνακα $\|\cdot\|_\alpha$, που είναι εξαρτώμενη κάποιας norm διανύσματος, ισχύει $\|I\|_\alpha = 1$.

54. Αν $A,B \in \mathbb{C}^{n,n}$, $A^H = A^{-1}$, να δειχθεί ότι i) $\|AB\|_2 = \|BA\|_2 = \|B\|_2 = \|ABA^H\|_2$ ii) $\|AB\|_E = \|ABA^H\|_E = \|B\|_E$ iii) $\|A\|_1 \geqq 1$, $\|A\|_\infty \geqq 1$.

55. Να βρεθούν οι $\|A\|_1 \|A\|_2 \|A\|_\infty$, αν $A = \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}$.

56. Αν $A = \begin{bmatrix} 1 & \alpha \\ \alpha & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & \beta \\ \beta & 1 \end{bmatrix}$, $\alpha,\beta \geqq 0$, να δειχθεί ότι $\|AB\|_2 = \|A\|_2 \|B\|_2$.

57. Αν A είναι Ερμιτιανός, να δειχθεί ότι $\|A\|_E^2 = \sum_{i=1}^{n} \lambda_i^2$, όπου λ_i , $i = 1(1)n$ οι ιδιοτιμές του A .

58. Αν $\|x\|_\alpha$ είναι μία norm διανύσματος και $\|A\|_\alpha$ η εξαρτώμενη από αυτή norm πίνακα, τότε υπάρχει διάνυσμα $x \neq 0$ τέτοιο ώστε $\|Ax\|_\alpha = \|A\|_\alpha \|x\|_\alpha$.

59. Αν A είναι ομαλός πίνακας, να δειχθεί ότι $\min_{\|x\|=1} \|Ax\| = \|A^{-1}\|^{-1}$, $x \in \mathbb{C}^n$.

60. Αν A,B είναι ομαλοί πίνακες και λ_i , $i = 1(1)n$ οι ιδιοτιμές του A, να δειχθεί ότι

$$\text{i)} \quad \kappa(A) \equiv \|A^{-1}\| \|A\| \geqq \frac{\max_i |\lambda_i|}{\min_i |\lambda_i|} \geqq 1 .$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ii) $\kappa(A^2) \leqq [\kappa(A)]^2$ iii) $\kappa(AB) \leqq \kappa(A)\kappa(B)$

iv) $\kappa(cA) = \kappa(A)$, $c \in \mathbb{C}$ v) $\kappa_2(A) \equiv \|A^{-1}\|_2\|A\|_2 = \sqrt{\frac{\mu_M}{\mu_m}}$,

όπου μ_M, μ_m η μεγαλύτερη και η μικρότερη αντίστοιχα ιδιοτιμή του $A^H A$.

61. Εάν A είναι ομαλός Ερμιτιανός πίνακας, να δειχθεί ότι

i) $\kappa_2(A) \equiv \|A^{-1}\|_2\|A\|_2 = \frac{\max_i|\lambda_i|}{\min_i|\lambda_i|}$, όπου λ_i , $i = 1(1)n$ οι ιδιοτιμές του A .

ii) $\kappa_2(A^2) = [\kappa_2(A)]^2 \geqq \kappa_2(A)$.

62. Αν $A,B \in \mathbb{C}^{n,n}$, $A^H = A^{-1}$, $\|B\|_2 < 1$, να δειχθεί ότι

$$\kappa_2(A + B) \leqq \frac{1 + \|B\|_2}{1 - \|B\|_2} .$$

63. Αν $A,U \in \mathbb{C}^{n,n}$, $U^H = U^{-1}$, να δειχθεί ότι $\kappa_2(UA) = \kappa_2(AU) = \kappa_2(A)$.

64. Εάν $A,E \in \mathbb{C}^{n,n}$, A ομαλός και $\|A^{-1}\|\|E\| < \frac{1}{2}$, να δειχθεί ότι $\|I - (A + E)^{-1}A\| < 1$. Με την προυπόθεση ότι και ο E είναι ομαλός, να δειχθεί επίσης ότι $\|(I \pm E^{-1}A)\| < 1$.

65. Να δειχθεί ότι ο πίνακας $B = \begin{bmatrix} 1 & 0.4 & 0.7 \\ 0.1 & 1 & 0.2 \\ 0.3 & 0.2 & 1 \end{bmatrix}$ είναι ομαλός και να βρεθεί ένα άνω σφάλμα για την $\|B^{-1}\|_1$, χωρίς να υπολογισθεί ο B^{-1} .

66. Εάν $A,B \in \mathbb{C}^{n,n}$, $\det(A) = 0$, $\|A^{-1}\|\|B\| < 1$, να δειχθεί ότι

$$\|(A - B)^{-1}\| \leqq \frac{1}{\|A^{-1}\|^{-1} - \|B\|} .$$

67. Για κάθε $A \in \mathbb{C}^{n,n}$, να δειχθεί ότι $\frac{1}{\kappa!} A^{\kappa} \underset{\kappa \to \infty}{\to} 0$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

68. Αν $A = I - 2xx^H$ με $\|x\|_2 = 1$, $x \in \mathbb{C}^n$, να δειχθεί ότι $A^H = A^{-1}$ και $A^2 = I$.

69. Αν $A \in \mathbb{C}^{n,n}$, να δειχθεί ότι $\max_{i,j}|a_{ij}| \leqq \|A\| \leqq n \max_{i,j}|a_{ij}|$.

70. Εάν $A,E \in \mathbb{C}^{n,n}$, $\det(A) \neq 0$ και $\|A^{-1}E\| < 1$, να δειχθεί ότι

i) $\|I - (A + E)^{-1}A\| \leqq \dfrac{\|A^{-1}E\|}{1 - \|A^{-1}E\|}$

ii) $\|A^{-1} - (A + E)^{-1}\| \leqq \dfrac{\|A^{-1}\|\|A^{-1}E\|}{1 - \|A^{-1}E\|}$.

71. Εάν A είναι ομαλός και ο $A + B$ είναι μη ομαλός, να δειχθεί ότι $\|B\|\|A^{-1}\| \geqq 1$.

72. Αν $A,C \in \mathbb{C}^{n,n}$, $\det(A) \neq 0$, $B = A - C$ και $\|C\|\|A^{-1}\| < 1$, να δειχθεί ότι i) ο B είναι ομαλός ii) $\|B^{-1}\| \leqq \dfrac{\|A^{-1}\|}{1 - \|A^{-1}\|\|C\|}$ και iii) $\|B^{-1} - A^{-1}\| \leqq \dfrac{\|A^{-1}\|^2\|C\|}{1 - \|A^{-1}\|\|C\|}$.

73. Έστω $R = I - AC$, $\|R\| < 1$. Να δειχθεί ότι i) οι A,C είναι ομαλοί ii) $\|A^{-1}\| \leqq \dfrac{\|C\|}{1 - \|R\|}$ iii) $\|C - A^{-1}\| \leqq \dfrac{\|C\|\|R\|}{1 - \|R\|}$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

10

ΑΡΙΘΜΗΤΙΚΗ ΕΠΙΛΥΣΗ ΓΡΑΜΜΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

10.1 Εισαγωγή

Στο κεφάλαιο αυτό θα ασχοληθούμε με αριθμητικές μεθόδους επιλύσεως συστημάτων n γραμμικών εξισώσεων με n αγνώστους. Η γενική μορφή ενός τέτοιου συστήματος είναι:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= b_n \,, \end{aligned} \tag{10.1}$$

όπου οι συντελεστές a_{ij} , $i,j = 1(1)n$ και τα δεύτερα μέλη b_i , $i = 1(1)n$ είναι δοσμένοι αριθμοί και x_i , $i = 1(1)n$ είναι οι άγνωστοι. Το σύστημα (10.1) γράφεται υπό μορφή πινάκων ως εξής

$$Ax = b \,, \tag{10.2}$$

όπου $A = (a_{ij})$, $x = (x_i)$ και $b = (b_i)$, $i,j = 1(1)n$. Ο πίνακας A καλείται πίνακας των συντελεστών του συστήματος (10.1) ή (10.2) και το διάνυσμα b καλείται διάνυσμα των σταθερών όρων του συστήματος.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Υποθέτουμε επίσης ότι $A \in \mathbb{R}^{n,n}$ και $b \in \mathbb{R}^n$ (πρέπει να σημειωθεί ότι αν $A \in \mathbb{C}^{n,n}$ και $b \in \mathbb{C}^n$, τότε μπορεί να δειχθεί (βλέπε π.χ.άσκηση 1) ότι η επίλυση του (10.2) ανάγεται στην επίλυση ενός συστήματος γραμμικών εξισώσεων $By = c$, όπου $B \in \mathbb{R}^{2n,2n}$ και $c \in \mathbb{R}^{2n}$, δηλαδή, ενός πραγματικού συστήματος διπλάσιας τάξεως.

Σύμφωνα με το θεώρημα 9.12 το σύστημα (10.2) έχει μία μοναδική λύση (την $x = A^{-1}b$) εάν και μόνον εάν $\det(A) \neq 0$. Η μοναδική αυτή λύση δίνεται επίσης από τις (9.13) (κανών του Cramer). Εάν $\det(A) = 0$, τότε το σύστημα (10.2) ή δεν έχει λύση ή έχει άπειρες λύσεις, πράγμα που εξαρτάται από το διάνυσμα b. Στη συνέχεια, θα υποθέτουμε ότι $\det(A) \neq 0$. (Τονίζεται εδώ, ότι στους πρακτικούς υπολογισμούς, λόγω των αναπόφευκτων σφαλμάτων στρογγυλεύσεως, που υπεισέρχονται, το πρόβλημα του προσδιορισμού εάν ο A είναι ομαλός ή όχι παρουσιάζει δυσκολίες. Έτσι, ένας "θεωρητικά" μη ομαλός πίνακας μπορεί να εμφανίζεται σαν ομαλός και αντιστρόφως).

Υπάρχουν δύο βασικές κατηγορίες μεθόδων για την αριθμητική επίλυση του συστήματος (10.2). Αυτές είναι: i) οι άμεσες μέθοδοι, οι οποίες μας δίνουν την ακριβή λύση του (10.2) σε ένα πεπερασμένο αριθμό βημάτων, εάν υποθέσουμε ότι οι υπολογισμοί γίνονται χωρίς σφάλματα στρογγυλεύσεως. Στην πράξη, λόγω των σφαλμάτων στρογγυλεύσεως, μία προσεγγιστική λύση βρίσκεται, για την οποία ο αριθμός των πράξεων που απαιτούνται είναι ανεξάρτητος από την επιζητούμενη ακρίβεια και ii) οι έμμεσες ή επαναληπτικές μέθοδοι, οι οποίες αρχίζουν με μία (αυθαίρετη) αρχική προσέγγιση της λύσεως και παράγουν με την βοήθεια ενός αναδρομικού τύπου μία ακολουθία διαδοχικών προσεγγίσεων, η οποία (κάτω από κατάλληλες συνθήκες) θα συγκλίνει στην ακριβή λύση του συστήματος, μετά από άπειρες επαναλήψεις και υποθέτοντας την απουσία των σφαλμάτων στρογγυλεύσεως. Στην πράξη, ο αριθμός των επαναλήψεων εξαρτάται από την επιζητούμενη ακρίβεια της λύσεως.

Το πρόβλημα της επιλύσεως συστημάτων γραμμικών εξισώσεων παρουσιάζεται συχνά σε πολλούς κλάδους των εφαρμοσμένων Μαθηματικών και γενικώτερα των θετικών επιστημών. Η τάξη του πίνακα A μπορεί να κυ-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

μαίνεται από μερικές δεκάδες μέχρι αρκετές εκατοντάδες χιλιάδες. Ακόμη, στην Αριθμητική Ανάλυση, εμφανίζονται γραμμικά συστήματα σε συνδυασμό και με άλλα προβλήματα, όπως είναι η αριθμητική επίλυση διαφορικών εξισώσεων με μερικές παραγώγους, με την μέθοδο των πεπερασμένων διαφορών, η επίλυση προβλημάτων ιδιοτιμών κ.λ.π. . Οι μέθοδοι που θα εξεταστούν στη συνέχεια είναι από τις πιο γνωστές που χρησιμοποιούνται για πρακτικές εφαρμογές.

10.2 Άμεσες Μέθοδοι

Είναι γνωστό ότι η λύση του συστήματος 10.2 μπορεί να βρεθεί θεωρητικά με τον κανόνα του Cramer (βλέπε (9.13)), ο οποίος απαιτεί τον υπολογισμό $n + 1$ οριζουσών τάξεως n . Μπορεί να δειχθεί ότι για την πλήρη επίλυση του συστήματος, αναπτύσσοντας πλήρως κάθε ορίζουσα, απαιτούνται αριθμητικές πράξεις ισοδύναμες με περίπου $2(n + 1)!$ πολλαπλασιασμούς. Έτσι, ακόμη και για σχετικά μικρό n , π.χ. $n = 100$, ($100! = 10^{157.9\cdots}$),ο αριθμός των αριθμητικών πράξεων, που απαιτούνται, είναι πέρα από τις δυνατότητες των συγχρόνων αλλά και μελλοντικών υπολογιστικών μηχανών. Για τον λόγο αυτό, είναι πρακτικά αδύνατη η χρησιμοποίηση της μεθόδου του Cramer. Πρέπει να σημειωθεί ακόμη,ότι ο αριθμός των πράξεων μιας μεθόδου είναι σημαντικός παράγων και για την ακρίβεια των αποτελεσμάτων, λόγω της πιθανής συσσωρεύσεως των σφαλμάτων στρογγυλεύσεως κατά τους υπολογισμούς. Γι' αυτό, είναι απαραίτητο οι μέθοδοι που χρησιμοποιούνται στην πράξη να είναι ταχείς και ακριβείς. Η απλούστερη και περισσότερο γνωστή άμεση μέθοδος είναι η καλούμενη μέθοδος απαλοιφής του Gauss .

10.2.1 Η Μέθοδος απαλοιφής του Gauss

Με την μέθοδο αυτή το σύστημα (10.2) μετασχηματίζεται σε ένα ισοδύναμο σύστημα (δηλαδή, ένα σύστημα με την ίδια λύση)

$$Rx = c \ , \qquad (10.3)$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

όπου ο πίνακας R είναι άνω τριγωνικός. Έτσι, το σύστημα (10.3) μπορεί εύκολα να λυθεί με την γνωστή μέθοδο της προς τα πίσω αντικατάστασης, όπως θα δειχθεί στη συνέχεια. Αλλά, πρώτα ας δούμε πώς προκύπτει το (10.3) από το (10.2) .

Είναι γνωστό ότι οι ακόλουθες στοιχειώδεις διαδικασίες εφαρμοζόμενες στο σύστημα (10.1) μετατρέπουν αυτό σε ισοδύναμα συστήματα:

i) Αντιμετάθεση δύο οποιωνδήποτε εξισώσεων (αλλαγή της σειράς των εξισώσεων).

ii) Πολλαπλασιασμός μιας οποιασδήποτε εξισώσεως με μία σταθερά $\neq 0$.

iii) Πρόσθεση του γινομένου μιας εξισώσεως επί μία σταθερά $\neq 0$ σε κάποια άλλη εξίσωση.

Αντί να εργαζόμαστε με τις εξισώσεις του συστήματος, συνήθως, οι παραπάνω διαδικασίες εκτελούνται στις γραμμές του αυξημένου πίνακα του συστήματος (ο οποίος προκύπτει θέτοντας δεξιά του πίνακα των συντελεστών A σαν νέα στήλη το διάνυσμα των σταθερών όρων b), που θα συμβολίζεται στη συνέχεια με (A,b) , δηλαδή,

$$(A,b) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{bmatrix} . \qquad (10.4)$$

Για να γίνει η περιγραφή της μεθόδου πιο συστηματική θα συμβολίσουμε με

$$A^{(1)}x = b^{(1)} \qquad (10.5)$$

τό σύστημα (10.2) , οπότε

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$(A^{(1)},b^{(1)}) = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \dots & a_{1n}^{(1)} & b_1^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & \dots & a_{2n}^{(1)} & b_2^{(1)} \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1}^{(1)} & a_{n2}^{(1)} & \dots & a_{nn}^{(1)} & b_n^{(1)} \end{bmatrix} . \qquad (10.6)$$

Υποθέτουμε ότι $a_{11}^{(1)} \neq 0$. Τότε, η πρώτη γραμμή του $(A^{(1)},b^{(1)})$ καλείται οδηγός γραμμή και το $a_{11}^{(1)}$ οδηγό στοιχείο. (Εάν $a_{11}^{(1)} = 0$, τότε, επειδή $\det(A^{(1)}) \neq 0$, μπορούμε να βρούμε ένα στοιχείο $a_{j1}^{(1)} \neq$ $\neq 0$ $(j > 1)$. Ανταλλάσσουμε τότε τις γραμμές 1 και j του (10.6) και εργαζόμαστε στη συνέχεια με τον προκύπτοντα πίνακα που θα δίνεται από το γινόμενο $P^{(1)}(A^{(1)},b^{(1)})$, όπου $P^{(1)}$ είναι ένας κατάλληλος πίνακας μεταθέσεως). Στο πρώτο βήμα της απαλοιφής, για $i = 2(1)n$, προσθέτουμε το γινόμενο της οδηγού γραμμής επί

$$m_{i1} = -\frac{a_{i1}^{(1)}}{a_{11}^{(1)}} \qquad (10.7)$$

στην i γραμμή του πίνακα (10.6). Τα m_{i1} , $i = 2(1)n$, που δίνονται από την (10.7) καλούνται πολλαπλασιαστές. Σαν αποτέλεσμα παίρνουμε τον πίνακα

$$(A^{(2)},b^{(2)}) = \left[\begin{array}{c|cccc} a_{11}^{(1)} & a_{12}^{(1)} & \dots & a_{1n}^{(1)} & b_1^{(1)} \\ \hline 0 & a_{22}^{(2)} & \dots & a_{2n}^{(2)} & b_2^{(2)} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & a_{n2}^{(2)} & \dots & a_{nn}^{(2)} & b_n^{(2)} \end{array}\right] . \qquad (10.8)$$

Με άλλα λόγια, το σύστημα $A^{(2)}x = b^{(2)}$ είναι ισοδύναμο με το (10.5)

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

αλλά, ο άγνωστος x_1 έχει απαλειφθεί από όλες τις εξισώσεις εκτός της πρώτης. Με την προυπόθεση ότι δεν υπάρχει ανταλλαγή γραμμών στον $A^{(1)}$, είναι εύκολο να δειχθεί ότι $A^{(2)} = M^{(1)}A^{(1)}$ και $b^{(2)} = M^{(1)}b^{(1)}$, όπου $M^{(1)}$ είναι ο $n \times n$ πίνακας:

$$M^{(1)} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ m_{21} & 1 & 0 & \dots & 0 \\ m_{31} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ m_{n1} & 0 & 0 & \dots & 1 \end{bmatrix} . \qquad (10.9)$$

Υποθέτουμε τώρα ότι $a_{22}^{(2)} \neq 0$. (Εάν $a_{22}^{(2)} = 0$, τότε η δεύτερη γραμμή στον (10.8) ανταλλάσσεται με κάποια από τις επόμενες γραμμές,έστω την q $(q > 2)$ με $a_{q2}^{(2)} \neq 0$. Αυτό είναι πάντοτε δυνατό, αφού $\det(A^{(1)}) \neq 0$). Η δεύτερη γραμμή του (10.8) είναι τώρα οδηγός γραμμή και το $a_{22}^{(2)}$ οδηγό στοιχείο. Στο δεύτερο βήμα της απαλοιφής η ίδια διαδικασία, που περιγράφηκε στο πρώτο βήμα, εφαρμόζεται τώρα στον πίνακα που προκύπτει εάν παραλείψουμε την πρώτη γραμμή και πρώτη στήλη του (10.8). Οι νέοι πολλαπλασιαστές είναι:

$$m_{i2} = -\frac{a_{i2}^{(2)}}{a_{22}^{(2)}} \quad , \; i = 3(1)n \; . \qquad (10.10)$$

Έτσι, προκύπτει ο πίνακας

$$(A^{(3)},b^{(3)}) = \left[\begin{array}{cc|ccc} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \dots & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \dots & a_{2n}^{(2)} & b_2^{(2)} \\ \hline 0 & 0 & a_{33}^{(3)} & \dots & a_{3n}^{(3)} & b_3^{(3)} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & a_{n3}^{(3)} & \dots & a_{nn}^{(3)} & b_n^{(3)} \end{array}\right] , \qquad (10.11)$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

που αντιστοιχεί στο σύστημα

$$A^{(3)}x = b^{(3)} \tag{10.12}$$

(ισοδύναμο με το $A^{(2)}x = b^{(2)}$) , όπου ο άγνωστος x_2 έχει απαλειφθεί από όλες εκτός από τις δύο πρώτες εξισώσεις. Επίσης, (υποθέτοντας ότι δεν υπάρχει ανταλλαγή γραμμών στον $A^{(2)}$) , ισχύει $A^{(3)} = M^{(2)}A^{(2)}$, $b^{(3)} = M^{(2)}b^{(2)}$, όπου

$$M^{(2)} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & m_{32} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & m_{n2} & 0 & \dots & 1 \end{bmatrix} . \tag{10.13}$$

Μετά από $\kappa - 1$ $(1 < \kappa < n)$ βήματα απαλοιφής, φθάνουμε στον πίνακα

$$(A^{(\kappa)}, b^{(\kappa)}) = \left[\begin{array}{cccc|cccc} a_{11}^{(1)} & a_{12}^{(1)} & \dots & a_{1,\kappa-1}^{(1)} & a_{1\kappa}^{(1)} & \dots & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(2)} & \dots & a_{2,\kappa-1}^{(2)} & a_{2\kappa}^{(2)} & \dots & a_{2n}^{(2)} & b_2^{(2)} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & a_{\kappa-1,\kappa-1}^{(\kappa-1)} & a_{\kappa-1,\kappa}^{(\kappa-1)} & \dots & a_{\kappa-1,n}^{(\kappa-1)} & b_{\kappa-1}^{(\kappa-1)} \\ \hline 0 & 0 & \dots & 0 & a_{\kappa\kappa}^{(\kappa)} & \dots & a_{\kappa n}^{(\kappa)} & b_\kappa^{(\kappa)} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & a_{n\kappa}^{(\kappa)} & \dots & a_{nn}^{(\kappa)} & b_n^{(\kappa)} \end{array}\right] \tag{10.14}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

που αντιστοιχεί στο σύστημα

$$A^{(\kappa)}x = b^{(\kappa)} , \tag{10.15}$$

το οποίο είναι ισοδύναμο με το αρχικό. Για την εφαρμογή του επόμενου βήματος απαλοιφής (κ βήμα), υποθέτουμε ότι $a_{\kappa\kappa}^{(\kappa)} \neq 0$. Τότε, το $a_{\kappa\kappa}^{(\kappa)}$ είναι οδηγό στοιχείο και η κ γραμμή του (10.14) οδηγός γραμμή. Για $i = \kappa + 1 , \kappa + 2,\ldots,n,$ προσθέτουμε το γινόμενο της οδηγού γραμμής επί

$$m_{i\kappa} = -\frac{a_{i\kappa}^{(\kappa)}}{a_{\kappa\kappa}^{(\kappa)}} \tag{10.16}$$

στην i γραμμή του (10.14) . Έτσι, προκύπτει ο πίνακας $(A^{(\kappa+1)} , b^{(\kappa+1)})$, που αντιστοιχεί στο σύστημα $A^{(\kappa+1)}x = b^{(\kappa+1)}$ και ισχύει $A^{(\kappa+1)} = M^{(\kappa)}A^{(\kappa)}$, $b^{(\kappa+1)} = M^{(\kappa)}b^{(\kappa)}$, όπου

$$M^{(\kappa)} = \left[\begin{array}{cccc|cccc}
1 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & \cdots & 1 & 0 & 0 & 0 & \cdots & 0 \\
\hline
0 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 & m_{\kappa+1,\kappa} & 1 & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 & m_{\kappa+2,\kappa} & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & \cdots & 0 & m_{n\kappa} & 0 & 0 & \cdots & 1
\end{array}\right] \tag{10.17}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Μία ενδιαφέρουσα ιδιότητα του πίνακα (10.17) είναι ότι ο αντίστροφός του προκύπτει αλλάζοντας απλώς το πρόσημο των m_{ik}, $i = k + 1, k + 2, \ldots, n$. (Εάν $a_{kk}^{(k)} = 0$, τότε η k γραμμή του (10.14) ανταλλάσσεται με κάποια γραμμή q , $q > k$, με $a_{qk}^{(k)} \neq 0$. Η νέα k γραμμή γίνεται τότε οδηγός γραμμή και το νέο $a_{kk}^{(k)}$ οδηγό στοιχείο. Αυτό είναι πάντοτε δυνατό αφού η $\det(A^{(1)}) \neq 0$ συνεπάγεται ότι $\det(A^{(k)}) \neq 0$. Έτσι, προκύπτει ο πίνακας $P^{(k)}(A^{(k)}, b^{(k)})$, με τον οποίο εργαζόμαστε στη συνέχεια, όπου $P^{(k)}$ είναι ένας κατάλληλος πίνακας μεταθέσεως. Συγκεκριμένα, ο $P^{(k)}$ προκύπτει από τον I εάν γίνει ανταλλαγή των γραμμών k και q και ισχύει $(P^{(k)})^{-1} = P^{(k)}$) .

Τελικά, μετά από $n - 1$ βήματα απαλοιφής ο πίνακας (10.6) μετασχηματίζεται στον

$$(A^{(n)}, b^{(n)}) = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1,n-1}^{(1)} & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2,n-1}^{(2)} & a_{2n}^{(2)} & b_2^{(2)} \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{n-1,n-1}^{(n-1)} & a_{n-1,n}^{(n-1)} & b_{n-1}^{(n-1)} \\ 0 & 0 & \cdots & 0 & a_{nn}^{(n)} & b_n^{(n)} \end{bmatrix}, \quad (10.18)$$

δηλαδή, το σύστημα (10.5) μετασχηματίζεται στο ισοδύναμο σύστημα

$$A^{(n)}x = b^{(n)} , \qquad (10.19)$$

όπου ο $A^{(n)}$ είναι άνω τριγωνικός. Σημειώνεται ότι στον πίνακα $A^{(n)}$ θα είναι $a_{nn}^{(n)} \neq 0$. Το $a_{nn}^{(n)}$ θεωρείται σαν το n – στό οδηγό στοιχείο. Θέτοντας $A^{(n)} = R = (r_{ij})$ και $b^{(n)} = c = (c_i)$, λύνουμε το σύστημα

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$Rx = c$ με την μέθοδο της προς τα πίσω αντικατάστασης ως εξής: Η τελευταία εξίσωση αυτού δίνει

$$x_n = \frac{c_n}{r_{nn}} .$$

Η υπολογισθείσα αυτή τιμή του x_n αντικαθίσταται στην προτελευταία $(n-1)$ εξίσωση, από την οποία παίρνουμε

$$x_{n-1} = \frac{c_{n-1} - r_{n-1,n}x_n}{r_{n-1,n-1}} .$$

Οι τιμές τώρα των x_n, x_{n-1} που βρέθηκαν αντικαθίστανται στην $n-2$ εξίσωση, από την οποία βρίσκουμε τον άγνωστο x_{n-2} και η διαδικασία συνεχίζεται μέχρις ότου υπολογισθούν όλοι οι άγνωστοι x_i, $i = 1(1)n$. Ο γενικός τύπος που μας δίνει τις τιμές των αγνώστων είναι

$$x_i = \frac{c_i - \sum_{\kappa=i+1}^{n} r_{i\kappa}x_\kappa}{r_{ii}} , \quad i = n, n-1, n-2, \ldots, 1 . \quad (10.20)$$

Σημειώνεται ότι $r_{ii} \neq 0$, αφού στον (10.18) είναι $a_{\kappa\kappa}^{(\kappa)} \neq 0$, $\kappa = 1(1)n$.

Από την παραπάνω περιγραφή της μεθόδου απαλοιφής του Gauss είναι φανερό ότι, εάν δεν υπάρχουν ανταλλαγές γραμμών, τότε

$$\begin{aligned} R &= A^{(n)} = MA^{(1)} , \\ c &= b^{(n)} = Mb^{(1)} , \end{aligned} \qquad (10.21)$$

όπου

$$M = M^{(n-1)}M^{(n-2)} \ldots M^{(2)}M^{(1)} .$$

Ο M είναι προφανώς κάτω τριγωνικός και μπορεί να δειχθεί ότι

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$M^{-1} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -m_{21} & 1 & 0 & \cdots & 0 \\ -m_{31} & -m_{32} & 1 & \cdots & 0 \\ -m_{41} & -m_{42} & -m_{43} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -m_{n1} & -m_{n2} & -m_{n3} & \cdots & 1 \end{bmatrix} . \qquad (10.22)$$

Από την (10.21) προκύπτει ότι

$$A^{(1)} = M^{-1}R = LR ,$$

όπου ο $L = M^{-1}$ είναι κάτω τριγωνικός και ο R άνω τριγωνικός. Δηλαδή, ο $A^{(1)}$ έχει γραφεί σαν το γινόμενο δύο τριγωνικών πινάκων. (Είναι φανερό ότι ο M^{-1} στην (10.22) μπορεί να κατασκευασθεί εύκολα, γνωρίζοντας τους πολλαπλασιαστές m_{ik} , $i > k$ (βλέπε (10.16)). Μπορεί να αποδειχθεί ότι εάν ο $A^{(1)}$ είναι ένας πραγματικός συμμετρικός θετικά ορισμένος πίνακας ή ένας αυστηρά διαγώνια υπέρτερος πίνακας (βλέπε ορισμούς 9.13 και 9.16), τότε η απαλοιφή του Gauss ολοκληρώνεται χωρίς ανταλλαγές γραμμών. (Σημειώνεται ότι αν λάβουμε υπ' όψιν και την πιθανή αντιμετάθεση δύο γραμμών του πίνακα $(A^{(k)}, b^{(k)})$ για να εφαρμοσθεί το k βήμα απαλοιφής, $k = 1(1)n-1$ (όταν συμβεί $a_{kk}^{(k)} = 0$) , τότε $(A^{(k+1)}, b^{(k+1)}) = M^{(k)}P^{(k)}(A^{(k)}, b^{(k)})$ και επομένως οι (10.21) πρέπει να αντικατασταθούν από τις

$$R = A^{(n)} = M^{(n-1)}P^{(n-1)}M^{(n-2)}P^{(n-2)} \ldots M^{(2)}P^{(2)}M^{(1)}P^{(1)}A^{(1)} = GA^{(1)},$$
$$c = b^{(n)} = Gb^{(1)} . \qquad (10.23)$$

Το σύστημα $GA^{(1)}x = Gb^{(1)}$ είναι ισοδύναμο με το αρχικό $A^{(1)}x = b^{(1)}$. Από την (10.23) έχουμε $A^{(1)} = G^{-1}R = \hat{L}R$, όπου ο $\hat{L} = G^{-1}$ δεν είναι γενικά κάτω τριγωνικός. Μπορεί όμως να δειχθεί στην περίπτωση

αυτή ότι

$$PA^{(1)} = LR ,$$

όπου $P = P^{(n-1)}P^{(n-2)} \dots P^{(2)}P^{(1)}$ είναι ένας πίνακας μεταθέσεως και $L = P\hat{L} = PG^{-1}$
είναι ένας κάτω τριγωνικός πίνακας. Δηλαδή, για κάθε ομαλό πίνακα $A^{(1)}$ υπάρχει ένας πίνακας μεταθέσεως P , έτσι ώστε ο $PA^{(1)}$ να γράφεται σαν γινόμενο δύο τριγωνικών πινάκων).

Όπως τονίστηκε στην περιγραφή της μεθόδου απαλοιφής του Gauss, ο μόνος περιορισμός για το οδηγό στοιχείο (στο κ βήμα της απαλοιφής) $a_{\kappa\kappa}^{(\kappa)}$ ήταν να είναι διάφορο του μηδενός, $\kappa = 1(1)n-1$. Εάν το οδηγό στοιχείο $a_{\kappa\kappa}^{(\kappa)}$ είναι μικρό απολύτως, σε σύγκριση με τα $a_{i\kappa}^{(\kappa)}$, $i > \kappa$, τότε οι αντίστοιχοι πολλαπλασιαστές $m_{i\kappa}$ που δίνονται από την (10.16) θα είναι απολύτως μεγαλύτεροι της μονάδος, με αποτέλεσμα να προκαλείται μεγέθυνση των σφαλμάτων στρογγυλεύσεως στα διάφορα βήματα της απαλοιφής και της προς τα πίσω αντικατάστασης. Για τον λόγο αυτό,με σκοπό να ελαχιστοποιήσουμε την επίδραση των σφαλμάτων στρογγυλεύσεως, εκλέγουμε, συνήθως, το οδηγό στοιχείο $a_{\kappa\kappa}^{(\kappa)}$ έτσι ώστε

$$|a_{\kappa\kappa}^{(\kappa)}| \geqq \max_{i>\kappa}|a_{i\kappa}^{(\kappa)}| . \tag{10.24}$$

Αυτό πετυχαίνεται με την καλούμενη τεχνική της "οδήγησης" . Μιά μορφή αυτής, που χρησιμοποιείται περισσότερο στην πράξη, είναι η "μερική οδήγηση" και εφαρμόζεται ως εξής: Υποθέτουμε ότι έχει βρεθεί ο πίνακας (10.14) και έστω

$$|a_{r\kappa}^{(\kappa)}| = \max_{i}|a_{i\kappa}^{(\kappa)}| , \quad i = \kappa(1)n .$$

Τότε, ανταλλάσσουμε τις γραμμές κ και r του (10.14) και το νέο $a_{\kappa\kappa}^{(\kappa)}$ θα είναι το οδηγό στοιχείο στο κ βήμα της απαλοιφής. Έτσι, η (10.24) , ικανοποιείται. Με την τροποποίηση αυτή η μέθοδος απαλοιφής του Gauss καλείται μέθοδος απαλοιφής του Gauss με μερική οδήγηση και

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

αποτελεί μία πιο ευσταθή διαδικασία για την λύση γραμμικών συστημάτων. Σημειώνεται τέλος, ότι η μέθοδος απαλοιφής του Gauss με μερική οδήγηση, υπό μορφή πινάκων, παρίσταται από σχέσεις της μορφής (10.23) .

Παράδειγμα 10.1 Χρησιμοποιώντας την μέθοδο απαλοιφής του Gauss με μερική οδήγηση, να βρεθεί η λύση του συστήματος:

$$\begin{cases} 3x_3 + x_4 = 2 \\ 5x_1 + x_2 + x_3 = 7 \\ 4x_2 + 2x_3 + x_4 = 5 \\ x_1 + x_2 - 3x_4 = -5 \end{cases} .$$

Λύση Η εργασία για την λύση του συστήματος μπορεί να παρασταθεί με τον ακόλουθο πρακτικό και οικονομικό τρόπο:

Πολλαπλασιαστές	Αυξημένος πίνακας					Διαδικασία
	0	0	3	1	2	Ανταλλαγή των δύο πρώτων γραμμών
	5	1	1	0	7	
	0	4	2	1	5	
	1	1	0	−3	−5	
	5	1	1	0	7	1ο βήμα απαλοιφής
0	0	0	3	1	2	
0	0	4	2	1	5	
− 0.2	1	1	0	−3	−5	
	5	1	1	0	7	
	0	0	3	1	2	

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$x_i = \frac{\quad}{V_{n-1,n-1}}$

	0	4	2	1	5	Ανταλλαγή των γραμμών 2 και 3
	0	0.8	−0.2	−3	−6.4	
	$\underline{5}$	1	1	0	7	
	0	$\underline{4}$	2	1	5	$2^{\underline{o}}$ βήμα απαλοιφής
0	0	0	3	1	2	
−0.2	0	0.8	−0.2	−3	−6.4	
	$\underline{5}$	1	1	0	7	
	0	$\underline{4}$	2	1	5	$3^{\underline{o}}$ βήμα απαλοιφής
	0	0	$\underline{3}$	1	2	
0.2	0	0	−0.6	−3.2	−7.4	
	$\underline{5}$	1	1	0	7	
	0	$\underline{4}$	2	1	5	προς τα πίσω αντικα-
	0	0	$\underline{3}$	1	2	τάσταση
	0	0	0	$-\underline{3}$	−7	
Λύση	$\frac{23}{18}$	$\frac{13}{18}$	$-\frac{1}{9}$	$\frac{7}{3}$		

Δηλαδή, η λύση του συστήματος είναι: $x_1 = \frac{23}{18} \approx 1.278$, $x_2 = \frac{13}{18} \approx 0.722$, $x_3 = -\frac{1}{9} \approx -0.111$ και $x_4 = \frac{7}{3} \approx 2.333$.

<u>Σημείωση</u> Τα οδηγά στοιχεία έχουν υπογραμμισθεί. ▲

10.2.2 Υπολογισμός ορίζουσας

Η μέθοδος απαλοιφής του Gauss (χωρίς ή με μερική οδήγηση) μπορεί να χρησιμοποιηθεί αποτελεσματικά για τον υπολογισμό της ορίζουσας του

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

πίνακα $A^{(1)}$ (εδώ δεν υποτίθεται ότι $\det(A^{(1)}) \neq 0$). Πρέπει να σημειωθεί ότι η διαδικασία εφαρμόζεται μόνον στον πίνακα $A^{(1)}$, αφού δεν υπάρχει αυξημένος πίνακας. Εάν σε κάποιο πίνακα $A^{(\kappa)}$, $\kappa = 1(1)n-1$ συμβεί να ισχύει $a_{i\kappa}^{(\kappa)} = 0$ για κάθε $i \geqq \kappa$, οπότε δεν μπορεί να εφαρμοσθεί το κ βήμα της απαλοιφής, τότε $\det(A^{(1)}) = 0$. Εάν με την εφαρμογή των $n-1$ βημάτων απαλοιφής φθάσουμε στον άνω τριγωνικό πίνακα $A^{(n)}$, τότε είναι φανερό ότι

$$\det(A^{(1)}) = (-1)^{\ell} a_{11}^{(1)} a_{22}^{(2)} \ldots a_{nn}^{(n)} , \qquad (10.25)$$

όπου ℓ είναι ο αριθμός των απαιτούμενων αντιμεταθέσεων μεταξύ δύο γραμμών, στους πίνακες $A^{(\kappa)}$, $\kappa = 1(1)n-1$, ώστε να ολοκληρωθεί η διαδικασία της απαλοιφής. Σημειώνεται ότι το $a_{nn}^{(n)}$ μπορεί να είναι 0 , οπότε $\det(A^{(1)}) = 0$. (Για την εύρεση του τύπου (10.25) χρησιμοποιήθηκαν τα θεωρήματα 9.8 και 9.10. Τα $a_{\kappa\kappa}^{(\kappa)}$, $\kappa = 1(1)n$ είναι τα οδηγά στοιχεία).

Παράδειγμα 10.2 Να υπολογισθεί η ορίζουσα του πίνακα των συντελεστών A ,του συστήματος του παραδείγματος 10.1 .

Λύση Για τη μέθοδο απαλοιφής του Gauss με μερική οδήγηση, που χρησιμοποιήθηκε στο παράδειγμα 10.1 , είναι $\ell = 2$, αφού έγιναν δύο ανταλλαγές γραμμών και $a_{11}^{(1)} = 5$, $a_{22}^{(2)} = 4$, $a_{33}^{(3)} = 3$, $a_{44}^{(4)} = -3$. Άρα ,$\det(A) = (-1)^2(5)(4)(3)(-3) = -180$. ▲

10.2.3 Επίλυση πολλών γραμμικών συστημάτων με τον ίδιο πίνακα συντελεστών

Εάν έχουμε να λύσουμε πολλά γραμμικά συστήματα,με τον ίδιο πίνακα συντελεστών A ($\det(A) \neq 0$) , τότε μπορούμε να εφαρμόσουμε την μέθοδο απαλοιφής του Gauss, χρησιμοποιώντας ένα αυξημένο πίνακα,που περιέχει τον A και όλα τα διανύσματα των σταθερών όρων των συστημάτων. Έτσι, η επίλυση των συστημάτων γίνεται συγχρόνως,αποφεύγοντας επανάληψη της διαδικασίας.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

10.2.4 Εύρεση του αντιστρόφου A^{-1} ενός πίνακα A $(\det(A) \neq 0)$

Το πρόβλημα του προσδιορισμού του A^{-1}, όπου A ομαλός πίνακας, ανάγεται σ' αυτό της προηγουμένης παραγράφου. Πράγματι, αν καλέσουμε X τον αντίστροφο πίνακα του A, τότε $AX = I$. Εάν $x^{(j)}$ είναι η j στήλη του X και $e^{(j)}$ η j στήλη του I, τότε προφανώς $Ax^{(j)} = e^{(j)}$, $j = 1(1)n$, δηλαδή, η j στήλη του X είναι η λύση του γραμμικού συστήματος $Ax = e^{(j)}$, $j = 1(1)n$. Έτσι, λύνοντας n γραμμικά συστήματα προσδιορίζουμε τον πίνακα $X = A^{-1}$.

<u>Παράδειγμα 10.3</u> Χρησιμοποιώντας την μέθοδο απαλοιφής του Gauss, να βρεθεί ο αντίστροφος του πίνακα:

$$A = \begin{bmatrix} 1 & -2 & 1 & 0 \\ 1 & -2 & 2 & -3 \\ 0 & 1 & -1 & 1 \\ -2 & 3 & -2 & 3 \end{bmatrix}.$$

<u>Λύση</u>

Πολ/στές	Αυξημένος πίνακας								Διαδικασία
	$\underline{1}$	-2	1	0	1	0	0	0	
-1	1	-2	2	-3	0	1	0	0	$1^{\underline{o}}$ βήμα
0	0	1	-1	1	0	0	1	0	απαλοιφής
2	-2	3	-2	3	0	0	0	1	
	$\underline{1}$	-2	1	0	1	0	0	0	
	0	0	1	-3	-1	1	0	0	

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

	0	1	−1	1	0	0	1	0	Ανταλλαγή των γραμμών 2 και 3
	0	−1	0	3	2	0	0	1	
	1	−2	1	0	1	0	0	0	
	0	1	−1	1	0	0	1	0	2ο βήμα απαλοιφής
0	0	0	1	−3	−1	1	0	0	
1	0	−1	0	3	2	0	0	1	
	1	−2	1	0	1	0	0	0	
	0	1	−1	1	0	0	1	0	3ο βήμα απαλοιφής
	0	0	1	−3	−1	1	0	0	
1	0	0	−1	4	2	0	1	1	
	1	−2	1	0	1	0	0	0	
	0	1	−1	1	0	0	1	0	Προς τα πίσω αντικατάσταση
	0	0	1	−3	−1	1	0	0	
	0	0	0	1	1	1	1	1	
Λύση 1ου	1	1	2	1					
Λύση 2ου	2	3	4	1					
Λύση 3ου	3	3	3	1					
Λύση 4ου	1	2	3	1					

Άρα, η πρώτη στήλη του A^{-1} είναι $(1,1,2,1)^T$, η δεύτερη (2,3,4,1 η τρίτη $(3,3,3,1)^T$ και η τέταρτη $(1,2,3,1)^T$, δηλαδή,

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$A^{-1} = \begin{bmatrix} 1 & 2 & 3 & 1 \\ 1 & 3 & 3 & 2 \\ 2 & 4 & 3 & 3 \\ 1 & 1 & 1 & 1 \end{bmatrix} .$$

10.2.5 Η μέθοδος απαλοιφής του Jordan

Με την μέθοδο αυτή το αρχικό σύστημα $A^{(1)}x = b^{(1)}$ $(\det(A^{(1)}) \neq 0)$ μετασχηματίζεται σε ένα ισοδύναμο $Dx = c$, όπου ο πίνακας D είναι διαγώνιος. Έτσι, η λύση του τελευταίου προκύπτει εύκολα κάνοντας n διαιρέσεις (αποφεύγοντας, δηλαδή, την προς τα πίσω αντικατάσταση).

Η μέθοδος απαλοιφής του Jordan αποτελεί μία παραλλαγή της μεθόδου απαλοιφής του Gauss. Έχοντας υπ' όψιν την θεωρία της παραγράφου 10.2.1, η μέθοδος απαλοιφής του Jordan χρειάζεται n βήματα απαλοιφής αντί για $n-1$ που χρειάζεται η μέθοδος του Gauss. Στο κ βήμα της απαλοιφής $(\kappa = 1(1)n)$ και με την προυπόθεση ότι $a_{\kappa\kappa}^{(\kappa)} \neq 0$ μηδενίζουμε τα στοιχεία της κ στήλης του $(A^{(\kappa)}, b^{(\kappa)})$ όχι μόνον κάτω από την κυρία διαγώνιο αλλά και πάνω απ' αυτή. Αυτό γίνεται αν κατάλληλα πολλαπλάσια της (οδηγού) κ γραμμής προστεθούν σ' όλες τις άλλες γραμμές. Πρέπει να σημειωθεί ότι ο αντίστοιχος του πίνακα (10.14), για την περίπτωση της μεθόδου του Jordan, θα είναι ο

$$(A^{(\kappa)}, b^{(\kappa)}) \doteq \text{(βλέπε αρχή επόμενης σελίδας)} , \qquad (10.26)$$

όπου για λόγους ομοιομορφίας τέθηκε $a_{\kappa-1,j}^{(\kappa)}$ αντί $a_{\kappa-1,j}^{(\kappa-1)}$, $j = \kappa(1)n$ και $b_{\kappa-1}^{(\kappa)}$ αντί $b_{\kappa-1}^{(\kappa-1)}$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$(A^{(\kappa)},b^{(\kappa)})=\left[\begin{array}{cccc|ccc}
a_{11}^{(1)} & 0 & \cdots & 0 & a_{1\kappa}^{(\kappa)} & \cdots & a_{1n}^{(\kappa)} & b_1^{(\kappa)} \\
0 & a_{22}^{(2)} & \cdots & 0 & a_{2\kappa}^{(\kappa)} & \cdots & a_{2n}^{(\kappa)} & b_2^{(\kappa)} \\
\vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\
0 & 0 & \cdots & a_{\kappa-1,\kappa-1}^{(\kappa-1)} & a_{\kappa-1,\kappa}^{(\kappa)} & \cdots & a_{\kappa-1,n}^{(\kappa)} & b_{\kappa-1}^{(\kappa)} \\
\hline
0 & 0 & \cdots & 0 & a_{\kappa\kappa}^{(\kappa)} & \cdots & a_{\kappa n}^{(\kappa)} & b_\kappa^{(\kappa)} \\
\vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\
0 & 0 & \cdots & 0 & a_{n\kappa}^{(\kappa)} & \cdots & a_{nn}^{(\kappa)} & b_n^{(\kappa)}
\end{array}\right] \tag{10.26}$$

Έτσι, στο κ βήμα απαλοιφής της μεθόδου του Jordan, για $i = 1(1)n$, $i \neq \kappa$, προσθέτουμε το γινόμενο της κ γραμμής του (10.26) επί

$$m_{i\kappa} = -\frac{a_{i\kappa}^{(\kappa)}}{a_{\kappa\kappa}^{(\kappa)}} \tag{10.27}$$

στην i γραμμή αυτού. Τελικά, μετά από n βήματα απαλοιφής, ο πίνακας $(A^{(1)},b^{(1)})$ μετασχηματίζεται στον

$$(A^{(n+1)},b^{(n+1)})=\begin{bmatrix}
a_{11}^{(1)} & 0 & 0 & \cdots & 0 & b_1^{(n+1)} \\
0 & a_{22}^{(2)} & 0 & \cdots & 0 & b_2^{(n+1)} \\
0 & 0 & a_{33}^{(3)} & \cdots & 0 & b_3^{(n+1)} \\
\vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & 0 & 0 & \cdots & a_{nn}^{(n)} & b_n^{(n+1)}
\end{bmatrix}, \tag{10.28}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

όπου ο $A^{(n+1)}$ είναι διαγώνιος. Το σύστημα $A^{(n+1)}x = b^{(n+1)}$ είναι ισοδύναμο με το αρχικό και επομένως η λύση του θα δίνεται από τις σχέσεις:

$$x_i = \frac{b_i^{(n+1)}}{a_{ii}^{(i)}}, \quad i = 1(1)n . \tag{10.29}$$

Τονίζεται ότι η τεχνική της μερικής οδήγησης μπορεί να χρησιμοποιηθεί και με την μέθοδο του Jordan. (Πρέπει να σημειωθεί όμως,ότι οι οριζόμενοι πολλαπλασιαστές δεν είναι αναγκαστικά όλοι απολύτως ≤ 1 ,όπως συμβαίνει στην μέθοδο του Gauss) . Με μια πρώτη ματιά θα νόμιζε κανείς ότι η μέθοδος απαλοιφής του Jordan απαιτεί λιγώτερες αριθμητικές πράξεις απ' ότι η μέθοδος του Gauss,για την λύση ενός γραμμικού συστήματος. Στην πραγματικότητα όμως,το αντίθετο ισχύει,όπως θα δούμε σε επόμενη παράγραφο.

<u>Παράδειγμα 10.4</u> Να λυθεί το σύστημα

$$\begin{cases} 3x_1 + x_2 - 2x_3 - x_4 = 3 \\ 2x_1 - 2x_2 + 2x_3 + 3x_4 = -8 \\ x_1 + 5x_2 - 4x_3 - x_4 = 3 \\ 3x_1 + x_2 + 2x_3 + 3x_4 = -1 , \end{cases}$$

με την μέθοδο απαλοιφής του Jordan με μερική οδήγηση .

<u>Λύση</u>

Πολλαπλασιαστές	Αυξημένος πίνακας					Διαδικασία
	3	1	−2	−1	3	
$-\frac{2}{3}$	2	−2	2	3	−8	1ο βήμα απαλοιφής
$-\frac{1}{3}$	1	5	−4	−1	3	
−1	3	1	2	3	−1	

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

	3	1	−2	−1	3	
	0	$-\frac{8}{3}$	$\frac{10}{3}$	$\frac{11}{3}$	−10	Ανταλλαγή των γραμμών
	0	$\frac{14}{3}$	$-\frac{10}{3}$	$-\frac{2}{3}$	2	2 και 3
	0	0	4	4	−4	
$-\frac{3}{14}$	3	1	−2	−1	3	
	0	$\frac{14}{3}$	$-\frac{10}{3}$	$-\frac{2}{3}$	2	
$\frac{4}{7}$	0	$-\frac{8}{3}$	$\frac{10}{3}$	$\frac{11}{3}$	−10	$2^{\underline{o}}$ βήμα απαλοιφής
0	0	0	4	4	−4	
	3	0	$-\frac{9}{7}$	$-\frac{6}{7}$	$\frac{18}{7}$	
	0	$\frac{14}{3}$	$-\frac{10}{3}$	$-\frac{2}{3}$	2	Ανταλλαγή των γραμμών
	0	0	$-\frac{10}{7}$	$\frac{23}{7}$	$-\frac{62}{7}$	3 και 4
	0	0	4	4	−4	
$\frac{9}{28}$	3	0	$-\frac{9}{7}$	$-\frac{6}{7}$	$\frac{18}{7}$	
$\frac{5}{6}$	0	$\frac{14}{3}$	$-\frac{10}{3}$	$-\frac{2}{3}$	2	$3^{\underline{o}}$ βήμα απαλοιφής
	0	0	4	4	−4	
$-\frac{5}{14}$	0	0	$\frac{10}{7}$	$\frac{23}{7}$	$-\frac{62}{7}$	
$-\frac{3}{13}$	3	0	0	$\frac{3}{7}$	$\frac{9}{7}$	
$-\frac{56}{39}$	0	$\frac{14}{3}$	0	$\frac{8}{3}$	$-\frac{4}{3}$	$4^{\underline{o}}$ βήμα απαλοιφής
$-\frac{28}{13}$	0	0	4	4	−4	
	0	0	0	$\frac{13}{7}$	$-\frac{52}{7}$	
	3	0	0	0	3	
	0	$\frac{14}{3}$	0	0	$\frac{28}{3}$	

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

	0	0	4	0	12	Εύρεση των αγνώστων με διαιρέσεις
	0	0	0	$\frac{13}{7}$	$-\frac{52}{7}$	
Λύση	1	2	3	−4		

Σημείωση: Τα στοιχεία μέσα στους κύκλους είναι τα οδηγά στοιχεία. Άρα, $x_1 = 1$, $x_2 = 2$, $x_3 = 3$, $x_4 = -4$. ▲

10.2.6 Υπολογισμός του πλήθους των αριθμητικών πράξεων στις μεθόδους απαλοιφής Gauss και Jordan

Θεωρούμε το πρόβλημα της επιλύσεως m γραμμικών συστημάτων με τον ίδιο πίνακα συντελεστών A , όπου $A \in \mathbb{R}^{n,n}$, με $\det(A) \neq 0$. Θα υπολογίσουμε το πλήθος των αριθμητικών πράξεων, που απαιτούνται, χρησιμοποιώντας για την λύση του πρώτα την μέθοδο απαλοιφής του Gauss και κατόπιν την μέθοδο απαλοιφής του Jordan. Δίνοντας μετά στο m τις τιμές $0, 1, n$, θα είμαστε σε θέση να υπολογίσουμε το πλήθος των αριθμητικών πράξεων που απαιτούνται για τον υπολογισμό της $\det(A)$, ενός γραμμικού συστήματος και του A^{-1} αντίστοιχα.

Ι. Θεωρούμε, χάριν απλότητος, την περίπτωση που η μέθοδος απαλοιφής του Gauss μπορεί να ολοκληρωθεί χωρίς να υπάρξουν ανταλλαγές γραμμών (δηλαδή, $A = LR$) . Στο κ βήμα απαλοιφής $(\kappa = 1(1)n-1)$, είναι φανερό, ότι εκτελούνται $n-1$ διαιρέσεις για την εύρεση των πολλαπλασιαστών (10.16), $(n-\kappa)(n-\kappa+m)$ πολλαπλασιασμοί καθώς και $(n-\kappa)(n-\kappa+m)$ προσθέσεις (οι αφαιρέσεις υπολογίζονται σαν προσθέσεις). Σημειώνεται ότι κατά την εφαρμογή του κ βήματος απαλοιφής, τα κάτωθεν του οδηγού στοιχείου $a_{\kappa\kappa}^{(\kappa)}$ στοιχεία του αυξημένου πίνακα τίθενται ίσα με 0, υποθέτοντας ότι δεν απαιτούνται αριθμητικές πράξεις για το σκοπό αυτό. Έτσι, γιά την ολοκλήρωση της διαδικασίας της απα-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

λοιφής απαιτούνται

$$\sum_{\kappa=1}^{n-1}(n-\kappa) = 1 + 2 + \ldots + n - 1 = \frac{n(n-1)}{2} \text{ διαιρέσεις,}$$

$$\sum_{\kappa=1}^{n-1}(n-\kappa)(n-\kappa+m) = \sum_{\kappa=1}^{n-1}(n-\kappa)^2 + m\sum_{\kappa=1}^{n-1}(n-\kappa) =$$

$$= \frac{n^3}{3} - \frac{n^2}{2} + \frac{n}{6} + \frac{mn(n-1)}{2} \text{ πολλαπλασιασμοί} \tag{10.30}$$

$$\text{και } \frac{n^3}{3} - \frac{n^2}{2} + \frac{n}{6} + \frac{mn(n-1)}{2} \text{ προσθέσεις .}$$

Έχοντας υπ' όψιν τον τύπο (10.20) , η προς τα πίσω αντικατάσταση για την λύση των m γραμμικών συστημάτων απαιτεί

$$n\,m \text{ διαιρέσεις}$$

$$(1 + 2 + 3 + \ldots + n - 1)m = \frac{mn(n-1)}{2} \text{ πολλαπλασιασμούς} \tag{10.31}$$

$$\text{και } \frac{mn(n-1)}{2} \text{ προσθέσεις .}$$

Από τις (10.30) και (10.31) προκύπτει ότι απαιτούνται συνολικά

$$\frac{n(n-1)}{2} + n\,m \text{ διαιρέσεις ,}$$

$$\frac{n^3}{3} - \frac{n^2}{2} + \frac{n}{6} + mn(n-1) \text{ πολλαπλασιασμοί} \tag{10.32}$$

$$\text{και } \frac{n^3}{3} - \frac{n^2}{3} + \frac{n}{6} + mn(n-1) \text{ προσθέσεις .}$$

Είναι φανερό, ότι για τον υπολογισμό της $\det(A)$ απαιτούνται οι πράξεις (10.32) με $m = 0$ και επί πλέον $n - 1$ πολλαπλασιασμοί, λόγω του τύπου (10.25), δηλαδή, συνολικά

$$\frac{n(n-1)}{2} = \frac{n^2}{2} - \frac{n}{2} \text{ διαιρέσεις ,} \tag{10.33}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\frac{n^3}{3} - \frac{n^2}{2} + \frac{n}{6} + n - 1 = \frac{n^3}{3} - \frac{n^2}{2} + \frac{7n}{6} - 1 \quad \text{πολλαπλασιασμοί}$$

$$\text{και} \quad \frac{n^3}{3} - \frac{n^2}{2} + \frac{n}{6} \quad \text{προσθέσεις .} \tag{10.33}$$

Για την λύση ενός γραμμικού συστήματος $(m = 1)$, από τις (10.31), προκύπτει ότι απαιτούνται

$$\frac{n(n-1)}{2} + n = \frac{n^2}{2} + \frac{n}{2} \quad \text{διαιρέσεις ,}$$

$$\frac{n^3}{3} - \frac{n^2}{2} + \frac{n}{6} + n(n-1) = \frac{n^3}{3} + \frac{n^2}{2} - \frac{5n}{6} \quad \text{πολλαπλασιασμοί} \tag{10.34}$$

$$\text{και} \quad \frac{n^3}{3} + \frac{n^2}{2} - \frac{5n}{6} \quad \text{προσθέσεις.}$$

Για τον υπολογισμό του A^{-1} απαιτούνται $(m = n$ στις (10.32))

$$\frac{n(n-1)}{2} + n^2 = \frac{3n^2}{2} - \frac{n}{2} \quad \text{διαιρέσεις ,}$$

$$\frac{n^3}{3} - \frac{n^2}{2} + \frac{n}{6} + n^2(n-1) = \frac{4n^3}{3} - \frac{3n^2}{2} + \frac{n}{6} \quad \text{πολλαπλασιασμοί} \tag{10.35}$$

$$\text{και} \quad \frac{4n^3}{3} - \frac{3n^2}{2} + \frac{n}{6} \quad \text{προσθέσεις .}$$

Πρέπει να σημειωθεί ότι οι πράξεις (10.35) απαιτούνται για την λύση n οποιωνδήποτε γραμμικών συστημάτων με τον ίδιο πίνακα συντελεστών A . Για την εύρεση του A^{-1} οι πράξεις (10.35) μπορούν να ελαττωθούν αν επωφεληθούμε από την παρουσία του ταυτοτικού πίνακα I στον αυξημένο πίνακα με τον οποίο εργαζόμαστε. Συγκεκριμένα, μπορεί να δειχθεί ότι απαιτούνται n^3 πολλαπλασιασμοί και διαιρέσεις και $n^3 - 2n^2 + n$ προσθέσεις.

Αν όλες οι πράξεις (10.34) θεωρηθούν πολλαπλασιασμοί, τότε έχουμε ένα σύνολο πολλαπλασιασμών

$$\frac{2n^3}{3} + \frac{3n^2}{2} - \frac{7n}{6} < n^3 \ , \quad \text{για} \quad n \geqq 4 \ .$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Από την σύγκριση με τον κανόνα του Cramer βλέπουμε ότι $n^3 << 2(n+1)!$.

Είναι εύκολο να δειχθεί ότι η μέθοδος απαλοιφής του Gauss για την λύση ενός γραμμικού συστήματος $(m = 1)$ είναι προτιμώτερη από τον κανόνα του Cramer, όπου όμως οι $n + 1$ ορίζουσες υπολογίζονται με την μέθοδο απαλοιφής του Gauss (δηλαδή, για κάθε ορίζουσα απαιτούνται οι πράξεις (10.33)). Ακόμη, προφανώς, είναι προτιμώτερο για την λύση των m γραμμικών συστημάτων $Ax = b^{(j)}$, $j = 1(1)m$, $m \geqq 1$, να χρησιμοποιήσουμε την μέθοδο απαλοιφής του Gauss (πράξεις (10.32)), παρά να υπολογίσουμε πρώτα τον A^{-1} με την μέθοδο απαλοιφής του Gauss και να σχηματίσουμε κατόπιν τα γινόμενα $A^{-1}b^{(j)}$, $j = 1(1)m$.

II. Θεωρούμε τώρα την μέθοδο απαλοιφής του Jordan και υποθέτουμε πάλι ότι δεν υπάρχουν ανταλλαγές γραμμών. Στο κ βήμα της απαλοιφής $(\kappa = 1(1)n)$, σύμφωνα με τη θεωρία της παραγράφου 10.2.4, είναι φανερό ότι εκτελούνται $n - 1$ διαιρέσεις για την εύρεση των πολλαπλασιαστών (10.27), $(n-1)(n-\kappa+m)$ πολλαπλασιασμοί καθώς και $(n-1)(n-\kappa+m)$ προσθέσεις. Επομένως, η απαλοιφή ολοκληρώνεται απαιτώντας συνολικά

$$\sum_{\kappa=1}^{n}(n-1) = n(n-1) \quad \text{διαιρέσεις,}$$

$$\sum_{\kappa=1}^{n}(n-1)(n-\kappa+m) = (n-1)\sum_{\kappa=1}^{n}(n-\kappa) + \sum_{\kappa=1}^{n}(n-1)m =$$

$$= \frac{n(n-1)^2}{2} + m\,n(n-1) \quad \text{πολλαπλασιασμούς και} \qquad (10.36)$$

$$\frac{n(n-1)^2}{2} + m\,n(n-1) \quad \text{προσθέσεις .}$$

Οι λύσεις τώρα των m συστημάτων προκύπτουν κάνοντας nm διαιρέσεις. Άρα, οι πράξεις που απαιτούνται συνολικά είναι

$$n(n-1) + nm \quad \text{διαιρέσεις ,}$$

$$\frac{n(n-1)^2}{2} + mn(n-1) \quad \text{πολλαπλασιασμοί και} \qquad (10.37)$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\frac{n(n-1)^2}{2} + mn(n-1) \quad \text{προσθέσεις} . \tag{10.37}$$

Συγκρίνοντας τις (10.32) και (10.37), εύκολα προκύπτει το συμπέρασμα ότι, για κάθε $m \geqq 0$ και $n \geqq 2$, η μέθοδος απαλοιφής του Gauss είναι προτιμώτερη από την μέθοδο απαλοιφής του Jordan. Μπορεί να δειχθεί ότι η απαλοιφή του Jordan είναι εξίσου αποτελεσματική, όπως η απαλοιφή Gauss, για την εύρεση του A^{-1} (δηλαδή, απαιτεί περίπου n^3 πολλαπλασιασμούς και διαιρέσεις και $n^3 - 2n^2 + n$ προσθέσεις) και γι' αυτό εφαρμόζεται σε πολλές περιπτώσεις.

10.2.7 Μέθοδοι τριγωνικοποιήσεως (μέθοδοι τριγωνικής παραγοντοποιήσεως)

Όπως τονίστηκε στην παράγραφο 10.2.1, η απαλοιφή του Gauss, όταν εφαρμοσθεί για την λύση του γραμμικού συστήματος

$$Ax = b , \tag{10.38}$$

όπου $\det(A) \neq 0$, μας παρέχει συγχρόνως και τους πίνακες P, L, R, έτσι ώστε

$$PA = LR , \tag{10.39}$$

όπου ο P είναι ένας πίνακας μεταθέσεως (μπορεί να είναι $P = I$), ο L κάτω τριγωνικός με μονάδες στη διαγώνιο και ο R άνω τριγωνικός. Η (10.39) αποτελεί μία τριγωνική ανάλυση του PA. Εάν, δοθέντος του συστήματος (10.38), η ανάλυση (10.39) είναι γνωστή (δηλαδή, οι P, L, R είναι γνωστοί), τότε μπορούμε να λύσουμε αμέσως εύκολα το σύστημα (10.38) ως εξής: Από την (10.38) προκύπτει το ισοδύναμο σύστημα

$$PAx = LRx = Pb . \tag{10.40}$$

Το διάνυσμα λύση x μπορεί να βρεθεί τώρα λύνοντας διαδοχικά τα δύο τριγωνικά συστήματα

$$Ly = Pb \quad \text{και} \quad Rx = y , \tag{10.41}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

με τη βοήθεια των μεθόδων της προς τα εμπρός και προς τα πίσω αντικατάστασης αντίστοιχα. Πράγματι, έχουμε

$$x = R^{-1}y = R^{-1}(L^{-1}Pb) = R^{-1}L^{-1}(Pb) = A^{-1}P^{-1}Pb = A^{-1}b \ .$$

Η τριγωνική ανάλυση (10.39) μπορεί να βρεθεί απ' ευθείας με την ακόλουθη διαδικασία (μέθοδος τριγωνικοποιήσεως). Για την απλούστευση της διαδικασίας, υποθέτουμε ότι ο A έχει όλους τους άνω αριστερά κύριους υποπίνακες ομαλούς (βλέπε θεωρήματα 9.58 και 9.62), οπότε είναι δυνατή (και μοναδική) η παρακάτω ανάλυση του A (δηλαδή, $P = I$) .

$$A = LR \ , \tag{10.42}$$

όπου $A = (a_{ij})$, ο $L = (\ell_{ij})$ είναι κάτω τριγωνικός με $\ell_{ii} = 1$ και ο $R = (r_{ij})$ άνω τριγωνικός με $r_{ii} \neq 0$, $i,j = 1(1)n$. Για τον προσδιορισμό των n^2 αγνώστων ℓ_{ij} , $i > j$ και r_{ij} , $i \leq j$, από την εξίσωση (10.42) παίρνουμε τις n^2 μη γραμμικές εξισώσεις

$$a_{ij} = \sum_{\kappa=1}^{\min(i,j)} \ell_{i\kappa} r_{\kappa j} \ , \ i,j = 1(1)n \ , \ (\ell_{ii} = 1) \ . \tag{10.43}$$

Ο υπολογισμός των αγνώστων στοιχείων των πινάκων L και R μπορεί να γίνει από τις εξισώσεις (10.43) με την ακόλουθη σειρά (μέθοδος Crout). Για $i = 1(1)n$ είναι

$$\begin{aligned} r_{i\kappa} &= a_{i\kappa} - \sum_{j=1}^{i-1} \ell_{ij} r_{j\kappa} \qquad \kappa = i, i+1, \ldots, n \\ \ell_{\kappa i} &= \frac{a_{\kappa i} - \sum_{j=1}^{i-1} \ell_{\kappa j} r_{ji}}{r_{ii}} \ , \quad \kappa = i+1, i+2, \ldots, n \ . \end{aligned} \tag{10.44}$$

Για παράδειγμα, αν $n = 3$, έχουμε

$$r_{1\kappa} = a_{1\kappa} - \sum_{j=1}^{0} \ell_{1j} r_{j\kappa} = a_{1\kappa} \ , \ \kappa = 1,2,3, \tag{10.45}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\ell_{K1} = \frac{a_{K1} - \sum_{j=1}^{0} \ell_{Kj} r_{j1}}{r_{11}} = \frac{a_{K1}}{r_{11}}, \quad K = 2,3,$$

$$r_{2K} = a_{2K} - \sum_{j=1}^{1} \ell_{2j} r_{jK} = a_{2K} - \ell_{21} r_{1K}, \quad K = 2,3$$

$$\ell_{K2} = \frac{a_{K2} - \sum_{j=1}^{1} \ell_{Kj} r_{j2}}{r_{22}} = \frac{a_{K2} - \ell_{K1} r_{12}}{r_{22}}, \quad K = 3, \tag{10.45}$$

$$r_{3K} = a_{3K} - \sum_{j=1}^{2} \ell_{3j} r_{jK} = a_{3K} - \ell_{31} r_{1K} - \ell_{32} r_{2K}, \quad K = 3.$$

Η διαδικασία τριγωνικοποιήσεως, που περιγράφηκε προηγουμένως, μπορεί να δειχθεί ότι είναι ισοδύναμη θεωρητικά και αριθμητικά με την απαλοιφή του Gauss, που δεν απαιτεί ανταλλαγές γραμμών. Αυτό σημαίνει ότι και οι δύο μέθοδοι απαιτούν τις ίδιες ακριβώς πράξεις για την λύση του (10.38). Οι μέθοδοι τριγωνικοποιήσεως εξυπηρετούν καλύτερα όταν εργαζόμαστε με το χέρι, αφού, σε αντίθεση με την απαλοιφή του Gauss, δεν χρειάζεται να καταγράφουμε τα ενδιάμεσα βήματα, καθώς και σε περιπτώσεις που έχουμε να λύσουμε πολλά συστήματα με τον ίδιο πίνακα A αλλά τα δεύτερα μέλη δεν είναι όλα διαθέσιμα από την αρχή.

<u>Παράδειγμα 10.5</u> Δίνεται το γραμμικό σύστημα $Ax = b$, $\det(A) \neq 0$, όπου για τον πίνακα A είναι δυνατή η ανάλυση (10.42). Να γραφούν οι σχέσεις με τις οποίες υπολογίζονται οι συνιστώσες του διανύσματος – λύση x του συστήματος. Να γίνει εφαρμογή για

$$A = \begin{bmatrix} 2 & 2 & -1 \\ -3 & 0 & 2 \\ 4 & -5 & -1 \end{bmatrix}, \quad b = \begin{bmatrix} 5 \\ -5 \\ 0 \end{bmatrix}.$$

<u>Λύση</u> Σύμφωνα με την θεωρία της παρούσας παραγράφου, εάν $A = LR$,

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

τότε για την λύση του $Ax = b$ αρκεί να λύσουμε τα δύο τριγωνικά συστήματα $Ly = b$ και $Rx = y$. Οι πίνακες L και R υπολογίζονται με βάση τις (10.44), οπότε για την λύση του $Ly = b$ με τη μέθοδο της προς τα εμπρός αντικατάστασης έχουμε

$$y_i = b_i - \sum_{\kappa=1}^{i-1} \ell_{i\kappa} y_\kappa \ , \ i = 1(1)n \ , \tag{10.46}$$

ενώ για την λύση του $Rx = y$ ο αντίστοιχος του τύπου (10.20) είναι

$$x_i = \frac{y_i - \sum_{\kappa=i+1}^{n} r_{i\kappa} x_\kappa}{r_{ii}} \ , \ i = n, n-1, \ldots, 1 \ . \tag{10.47}$$

Για τον πίνακα $A = \begin{bmatrix} 2 & 2 & -1 \\ -3 & 0 & 2 \\ 4 & -5 & -1 \end{bmatrix}$, υπολογίζουμε τους L και R από τις (10.45). Έτσι, έχουμε

$$r_{11} = 2 \ , \quad r_{12} = 2 \ , \quad r_{13} = -1 ,$$

$$\ell_{21} = \frac{-3}{2} \ , \quad \ell_{31} = \frac{4}{2} = 2 \ ,$$

$$r_{22} = 0 - \left(-\frac{3}{2}\right)2 = 3 \ , \ r_{23} = 2 - \left(-\frac{3}{2}\right)(-1) = 2 - \frac{3}{2} = \frac{1}{2} \ ,$$

$$\ell_{32} = \frac{-5 - 2 \cdot 2}{3} = -3 \ ,$$

$$r_{33} = -1 - 2(-1) - (-3)\left(\frac{1}{2}\right) = -1 + 2 + \frac{3}{2} = \frac{5}{2} \ .$$

Άρα, $L = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{3}{2} & 1 & 0 \\ 2 & -3 & 1 \end{bmatrix}$, $R = \begin{bmatrix} 2 & 2 & -1 \\ 0 & 3 & \frac{1}{2} \\ 0 & 0 & \frac{5}{2} \end{bmatrix}$, οπότε .

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$y_1 = b_1 = 5 \ ,$$
$$y_2 = b_2 - \ell_{21}y_1 = -5 - \left(-\frac{3}{2}\right)5 = -5 + \frac{15}{2} = \frac{5}{2} \ ,$$
$$y_3 = b_3 - \ell_{31}y_1 - \ell_{32}y_2 = 0 - 2(5) - (-3)\,\frac{5}{2} = -10 + \frac{15}{2} = \frac{-5}{2} \ ,$$
$$x_3 = \frac{y_3}{r_{33}} = -1 \ ,$$
$$x_2 = \frac{y_2 - r_{23}x_3}{r_{22}} = \frac{\frac{5}{2} - \frac{1}{2}(-1)}{3} = 1,$$
$$x_1 = \frac{y_1 - r_{12}x_2 - r_{13}x_3}{r_{11}} = \frac{5 - 2(1) - (-1)(-1)}{2} = \frac{5 - 2 - 1}{2} = 1.$$

Δηλαδή, οι συνιστώσες του διανύσματος λύση είναι: $x_1 = 1$, $x_2 = 2$, $x_3 = -1$.

10.2.8 Η μέθοδος Cholesky για θετικά ορισμένους πίνακες

Υποθέτουμε ότι έχουμε να λύσουμε το γραμμικό σύστημα

$$Ax = b \ , \tag{10.48}$$

όπου ο $A = (a_{ij})$ είναι ένας πραγματικός συμμετρικός θετικά ορισμένος πίνακας. Τότε, σύμφωνα με το θεώρημα 9.54 , υπάρχει ένας μοναδικός πραγματικός κάτω τριγωνικός πίνακας $L = (\ell_{ij})$ με $\ell_{ii} > 0$, τέτοιος ώστε:

$$A = LL^T \ . \tag{10.49}$$

Εάν βρεθεί ο L , τότε το σύστημα (10.48) είναι ισοδύναμο με τα δύο τριγωνικά συστήματα

$$Ly = b \quad \text{και} \quad L^Tx = y \ , \tag{10.50}$$

τα οποία λύνονται εύκολα κατά τα γνωστά. Για τον υπολογισμό των στοιχείων του πίνακα L , από την εξίσωση (10.49) μπορούμε να πάρουμε τις ακόλουθες εξισώσεις για $i = 1(1)n$:

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\ell_{ii} = \left[a_{ii} - \sum_{\kappa=1}^{i-1} \ell_{i\kappa}^2\right]^{\frac{1}{2}} ,$$

$$\ell_{ji} = \frac{1}{\ell_{ii}} \left[a_{ij} - \sum_{\kappa=1}^{i-1} \ell_{i\kappa}\ell_{j\kappa}\right] , \; j > i \; . \tag{10.51}$$

Για παράδειγμα, αν $n = 3$ έχουμε

$$\ell_{11} = \sqrt{a_{11}} \quad ,$$

$$\ell_{21} = \frac{a_{12}}{\ell_{11}} , \; \ell_{31} = \frac{a_{13}}{\ell_{11}} ,$$

$$\ell_{22} = \sqrt{a_{22} - \ell_{21}^2} \quad , \tag{10.52}$$

$$\ell_{32} = \frac{1}{\ell_{22}} \; (a_{23} - \ell_{21}\ell_{31}) \; ,$$

$$\ell_{33} = \sqrt{a_{33} - \ell_{31}^2 - \ell_{3.2}^2} \quad .$$

Η μέθοδος που αναπτύχθηκε προηγουμένως καλείται μέθοδος τριγωνικής αναλύσεως του Cholesky και απαιτεί εκτός των άλλων πράξεων και τον υπολογισμό n τετραγωνικών ριζών. Λόγω της πρώτης των (10.51) θα ισχύει $|\ell_{i\kappa}| \leqq \sqrt{a_{ii}}$, $\kappa = 1(1)i$, $i = 1(1)n$.

10.2.9 Εκτίμηση σφάλματος στην λύση γραμμικού συστήματος

Είναι φανερό ότι, εάν μία οποιαδήποτε από τις άμεσες μεθόδους, που περιγράφηκαν στις προηγούμενες παραγράφους, χρησιμοποιηθεί για την λύση του γραμμικού συστήματος

$$Ax = b \quad (\det(A) \neq 0 \; , \quad b \neq 0) \; , \tag{10.53}$$

δεν βρίσκουμε γενικά την ακριβή λύση $\hat{x} = A^{-1}b$ αυτού, αλλά μία προ-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

σέγγιση $\bar{x}$ αυτής, λόγω των σφαλμάτων στρογγυλεύσεως. Η υπολογισθείσα λύση $\bar{x}$ θα περιέχει ένα σφάλμα ε, που δίνεται από την διαφορά

$$\varepsilon = \bar{x} - \hat{x} \ . \tag{10.54}$$

Για να μετρήσουμε το μέγεθος του σφάλματος ε, δηλαδή, για να προσδιορίσουμε πόσο καλή είναι η προσέγγιση $\bar{x}$, χρησιμοποιούμε μία norm διανύσματος (βλέπε παράγραφο 9.7). Έτσι, ο αριθμός $\|\varepsilon\|$ καλείται απόλυτο σφάλμα του $\bar{x}$ και ο $\dfrac{\|\varepsilon\|}{\|\hat{x}\|}$ σχετικό σφάλμα του $\bar{x}$.

Το διάνυσμα

$$r = A\bar{x} - b \tag{10.55}$$

θα καλείται υπόλοιπο και ισχύει

$$A\varepsilon = A(\bar{x} - \hat{x}) = A\bar{x} - A\hat{x} = A\bar{x} - b = r \quad \text{ή}$$

$$\varepsilon = A^{-1}r \ . \tag{10.56}$$

Από την (10.56), χρησιμοποιώντας κάποια norm, παίρνουμε για το απόλυτο σφάλμα του $\bar{x}$

$$\|\varepsilon\| = \|A^{-1}r\| \leqq \|A^{-1}\|\|r\| \ . \tag{10.57}$$

Από την $A\hat{x} = b$ εξ' άλλου, προκύπτει ότι $\|b\| = \|A\hat{x}\| \leqq \|A\|\|\hat{x}\|$ ή

$$\frac{1}{\|\hat{x}\|} \leqq \frac{\|A\|}{\|b\|} \ . \tag{10.58}$$

Από τις (10.57) και (10.58), για το σχετικό σφάλμα του $\bar{x}$ έχουμε

$$\frac{\|\varepsilon\|}{\|\hat{x}\|} \leqq \frac{\|A\|\|A^{-1}\|\|r\|}{\|b\|} \ . \tag{10.59}$$

Από τις (10.57) και (10.59) γίνεται φανερό ότι ένα "μικρό" υπόλοιπο r (δηλαδή, με μικρή $\|r\|$) δεν συνεπάγεται αναγκαστικά μικρά $\|\varepsilon\|$ και $\dfrac{\|\varepsilon\|}{\|\hat{x}\|}$, αφού π.χ. στην (10.57) ο παράγων $\|A^{-1}\|$ μπορεί να είναι πολύ μεγάλος.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Θα βρεθεί τώρα ένα κάτω φράγμα για το σχετικό σφάλμα του $\bar{x}$. Έχουμε

$$\|r\| = \|A\bar{x} - A\hat{x}\| \leqq \|A\|\|\bar{x} - \hat{x}\| \quad , \text{ οπότε}$$

$$\|\bar{x} - \hat{x}\| \geqq \frac{\|r\|}{\|A\|} \tag{10.60}$$

και $\|\hat{x}\| = \|A^{-1}b\| \leqq \|A^{-1}\|\|b\|$ ή

$$\frac{1}{\|\hat{x}\|} \geqq \frac{1}{\|A^{-1}\|\|b\|} \quad . \tag{10.61}$$

Από τις (10.60) και (10.61) προκύπτει ότι

$$\frac{\|\varepsilon\|}{\|\hat{x}\|} \geqq \frac{1}{\|A\|\|A^{-1}\|\|b\|} \quad . \tag{10.62}$$

10.2.10 Ασταθή γραμμικά συστήματα

Το σύστημα (10.53) θα καλείται ασταθές εάν η λύση του $\hat{x}$ είναι πολύ ευαίσθητη σε μικρές μεταβολές των στοιχείων του αυξημένου πίνακα (A,b), δηλαδή, σχετικά μικρές μεταβολές στα δεδομένα A,b προκαλούν, γενικά,σχετικά μεγάλη μεταβολή στην ακριβή του λύση. Εάν δεχθούμε την άποψη ότι η προσεγγιστική (λόγω των σφαλμάτων στρογγυλεύσεως) λύση $\bar{x}$, που υπολογίζουμε για το $Ax = b$ με μία οποιαδήποτε αριθμητική μέθοδο, είναι η ακριβής λύση ενός συστήματος που προκύπτει από το $Ax = b$ προκαλώντας μικρές μεταβολές στα δεδομένα A,b, είναι φανερό ότι για ένα ασταθές σύστημα η λύση που υπολογίζεται μπορεί να είναι πολύ διαφορετική από την ακριβή λύση και επομένως χωρίς αξία. Όπως θα δειχθεί στη συνέχεια, η αστάθεια ή μη ενός συστήματος εξαρτάται μόνον από τον πίνακα A . Ένα μη ασταθές σύστημα θα καλείται ευσταθές. Ευτυχώς, τα περισσότερα συστήματα που προκύπτουν από πρακτικές εφαρμογές είναι ευσταθή.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Θα εξετάσουμε τώρα ποια επίδραση στην λύση $\hat{x}$ του (10.53) έχει η εισαγωγή μικρών μεταβολών (διαταράξεων) στα δεδομένα A και b. Μαζί με το σύστημα (10.53), του οποίου η ακριβής λύση είναι $\hat{x} = A^{-1}b$, θεωρούμε και το διαταραγμένο σύστημα

$$(A + \delta A)x = b + \delta b \ , \tag{10.63}$$

όπου δA, δb είναι οι διαταράξεις των A και b αντίστοιχα. Υποθέτουμε επίσης ότι

$$\|A^{-1}\|\|\delta A\| < 1 \ , \tag{10.64}$$

οπότε ο πίνακας $A + \delta A$ είναι ομαλός. Πράγματι, έχουμε

$$A + \delta A = A(I + A^{-1}\delta A) \ . \tag{10.65}$$

Επειδή

$$\|A^{-1}\delta A\| \leq \|A^{-1}\|\|\delta A\| < 1 \ , \tag{10.66}$$

σύμφωνα με το θεώρημα 9.84, ο πίνακας $I + A^{-1}\delta A$ είναι ομαλός, οπότε από την (10.65) προκύπτει ότι ο $A + \delta A$ είναι ομαλός, αφού και ο A είναι ομαλός.

Αν $\bar{x}$ είναι η ακριβής λύση του (10.63), τότε η μεταβολή (διατάραξη) της $\hat{x}$, δηλαδή, το σφάλμα της $\bar{x}$ θα είναι

$$\delta\hat{x} = \bar{x} - \hat{x} \ . \tag{10.67}$$

Αναζητούμε τώρα μια σχέση μεταξύ των σχετικών διαταράξεων $\dfrac{\|\delta\hat{x}\|}{\|\hat{x}\|}$, $\dfrac{\|\delta A\|}{\|A\|}$ και $\dfrac{\|\delta b\|}{\|b\|}$. Από την (10.67) είναι $\bar{x} = \hat{x} + \delta\hat{x}$, οπότε από την (10.63) παίρνουμε

$$(A + \delta A)(\hat{x} + \delta\hat{x}) = b + \delta b \ . \tag{10.68}$$

Αφαιρώντας την $A\hat{x} = b$ από την (10.68) προκύπτει ότι

$$A\delta\hat{x} + (\delta A)\hat{x} + \delta A\delta\hat{x} = \delta b \quad \text{ή} \quad \delta\hat{x} = A^{-1}(\delta b - (\delta A)\hat{x} - \delta A\delta\hat{x}) \ , \tag{10.69}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

οπότε

$$\|\delta\hat{x}\| = \|A^{-1}(\delta b - (\delta A)\hat{x} - \delta A\delta\hat{x})\| \leqq \|A^{-1}\|\|\delta b\| +$$

$$+ \|A^{-1}\|\|\delta A\|\|\hat{x}\| + \|A^{-1}\|\|\delta A\|\|\delta\hat{x}\|$$

ή

$$\|\delta\hat{x}\| \leqq \frac{\|A^{-1}\|(\|\delta b\| + \|\delta A\|\|\hat{x}\|)}{1 - \|A^{-1}\|\|\delta A\|}, \tag{10.70}$$

αφού, λόγω της (10.64), είναι $1 - \|A^{-1}\|\|\delta A\| > 0$. Από την $A\hat{x} = b$ προκύπτει

$$\frac{1}{\|\hat{x}\|} \leqq \frac{\|A\|}{\|b\|} . \tag{10.71}$$

Από την (10.70) και λόγω της (10.71) παίρνουμε

$$\frac{\|\delta\hat{x}\|}{\|\hat{x}\|} \leqq \frac{\|A^{-1}\|}{1 - \|A^{-1}\|\|\delta A\|}\left(\frac{\|\delta b\|}{\|\hat{x}\|} + \|\delta A\|\right) \leqq \frac{\|A^{-1}\|}{1 - \|A^{-1}\|\|\delta A\|}\left(\frac{\|\delta b\|\|A\|}{\|b\|} +\right.$$

$$\left. + \|\delta A\|\right) = \frac{\|A^{-1}\|\|A\|}{1 - \|A^{-1}\|\|\delta A\|}\left(\frac{\|\delta b\|}{\|b\|} + \frac{\|\delta A\|}{\|A\|}\right) \quad \text{ή}$$

$$\frac{\|\delta\hat{x}\|}{\|\hat{x}\|} \leqq \frac{\kappa}{1 - \kappa\dfrac{\|\delta A\|}{\|A\|}}\left(\frac{\|\delta b\|}{\|b\|} + \frac{\|\delta A\|}{\|A\|}\right), \tag{10.72}$$

όπου

$$\kappa = \|A\|\|A^{-1}\| . \tag{10.73}$$

Έτσι, βρήκαμε ένα άνω φράγμα για την σχετική διατάραξη της λύσης $\hat{x}$ του (10.53).

Αν $\delta A = 0$, τότε από τις (10.69) και (10.72) προκύπτει ότι

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\|\delta\hat{x}\| \leqq \|A^{-1}\|\|\delta b\| ,$$
$$\frac{\|\delta\hat{x}\|}{\|\hat{x}\|} \leqq \kappa \frac{\|\delta b\|}{\|b\|} . \tag{10.74}$$

Αν $\delta b = 0$, τότε από την (10.72) έχουμε (υποθέτοντας ότι ισχύει η (10.64))

$$\frac{\|\delta\hat{x}\|}{\|\hat{x}\|} \leqq \frac{\kappa}{1 - \kappa \dfrac{\|\delta A\|}{\|A\|}} \frac{\|\delta A\|}{\|A\|} . \tag{10.75}$$

Από την (10.72) παρατηρούμε ότι, αν η ποσότητα $\|A^{-1}\|\|\delta A\| = \kappa \dfrac{\|\delta A\|}{\|A\|}$ είναι επαρκώς μικρή, τότε μπορούμε να αντικαταστήσουμε με 1 την παράσταση $1 - \kappa \dfrac{\|\delta A\|}{\|A\|}$ στην (10.72), οπότε η σταθερά $\kappa = \|A\|\|A^{-1}\|$ είναι ένα μέτρο της ευαισθησίας της λύσης $\hat{x}$, λόγω των διαταράξεων δA , δb . Εάν ο κ είναι μικρός, τότε μικρές σχετικές διαταράξεις των A , b δημιουργούν μικρή σχετική διατάραξη στη λύση $\hat{x}$ και επομένως το σύστημα $Ax = b$ είναι ευσταθές. Εάν ο κ είναι μεγάλος, τότε μικρές σχετικές διαταράξεις στα δεδομένα A , b είναι δυνατόν να προκαλέσουν μεγάλη σχετική διατάραξη στη λύση $\hat{x}$, δηλαδή, το σύστημα $Ax = b$ μπορεί να είναι ασταθές. Ο αριθμός κ καλείται αριθμός καταστάσεως του A σε σχέση με το σύστημα $Ax = b$ και ισχύει

$$1 = \|I\| = \|A^{-1}A\| \leqq \|A^{-1}\|\|A\| = \kappa ,$$

δηλαδή, $\kappa \geqq 1$. Λέμε, συνήθως, ότι ο A είναι ευσταθής (ασταθής) σε σχέση με το $Ax = b$ εάν ο κ είναι μικρός (μεγάλος, δηλαδή, $\kappa \gg 1$). Λόγω του A^{-1} δεν είναι εύκολη η εκτίμηση του κ στην πράξη. Ο κ, προφανώς, εξαρτάται από την norm που χρησιμοποιείται. Έτσι, έχουμε

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\kappa_1 = \|A^{-1}\|_1 \|A\|_1 \ ,$$
$$\kappa_2 = \|A^{-1}\|_2 \|A\|_2 \ , \qquad (10.76)$$
$$\kappa_\infty = \|A^{-1}\|_\infty \|A\|_\infty \ .$$

Ο κ_2 καλείται φασματικός αριθμός καταστάσεως του A σε σχέση με το $Ax = b$.

Πρέπει να σημειωθεί ότι αν ο A είναι σχεδόν μη ομαλός (η $\det(A)$ πολύ κοντά στο 0), τότε, γενικά, περιμένουμε το σύστημα $Ax = b$ να είναι ασταθές. Πιο συγκεκριμένα, εάν ο κανονικοποιημένος πίνακας,που προκύπτει διαιρώντας την i γραμμή $a^{(i)}$ του A με το μέτρο της $\|a^{(i)}\|_2 = \left(\sum_{j=1}^{n} a_{ij}^2\right)^{\frac{1}{2}}$, $i = 1(1)n$, είναι σχεδόν μη ομαλός, τότε το σύστημα $Ax = b$ είναι ασταθές.

Ένα παράδειγμα ασταθούς συστήματος είναι το εξής:

$$\begin{cases} x_1 + 0.99x_2 = 1.99 \\ 0.99x_1 + 0.98x_2 = 1.97 \ , \end{cases}$$

του οποίου η ακριβής λύση είναι $x_1 = x_2 = 1$. Το διαταραγμένο σύστημα

$$\begin{cases} x_1 + 0.99x_2 = 2.00 \\ 0.99x_1 + 0.98x_2 = 1.97 \end{cases}$$

έχει την ακριβή λύση $x_1 = -97$, $x_2 = 100$.

10.2.11 Επαναληπτική βελτίωση μιας προσεγγιστικής λύσεως

Ας υποθέσουμε ότι $x^{(0)}$ είναι μία προσέγγιση της ακριβούς λύσεως $\hat{x}$ του (10.53), που βρέθηκε με κάποια αριθμητική μέθοδο. Το σφάλμα ε στην $x^{(0)}$ θα είναι

$$\varepsilon = x^{(0)} - \hat{x}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

και λόγω των (10.55) και (10.56) θα έχουμε

$$r^{(0)} = Ax^{(0)} - b \ , \ A\varepsilon = r^{(0)} \ .$$

Δηλαδή, το ε είναι η λύση του συστήματος $Ax = r^{(0)}$, στο οποίο τα $A, r^{(0)}$ είναι γνωστά. Αν το $r^{(0)}$ υπολογιζόταν ακριβώς και μπορούσαμε να βρούμε την ακριβή λύση ε του $Ax = r^{(0)}$, τότε θα είχαμε

$$\hat{x} = x^{(0)} - \varepsilon \ .$$

Επειδή όμως, στην πράξη, λύνοντας το σύστημα $Ax = r^{(0)}$, βρίσκουμε μία προσέγγιση $\varepsilon^{(0)}$ του ε, είναι λογικό να υποθέσουμε ότι η

$$x^{(1)} = x^{(0)} - \varepsilon^{(0)}$$

είναι μία βελτιωμένη προσέγγιση της $\hat{x}$. Βρίσκουμε τώρα τα $r^{(1)}$ και $\varepsilon^{(1)}$, οπότε $x^{(2)} = x^{(1)} - \varepsilon^{(1)}$ και η διαδικασία μπορεί να εφαρμοσθεί επαναληπτικά. Γενικά, έχουμε

$$x^{(i+1)} = x^{(i)} - \varepsilon^{(i)} \ , \ i = 0,1,2,\ldots, \tag{10.77}$$

όπου το $\varepsilon^{(i)}$ είναι η λύση που υπολογίζεται για το σύστημα

$$Ax = r^{(i)} = Ax^{(i)} - b \ , \ i = 0,1,2,\ldots \ . \tag{10.78}$$

Πρέπει να σημειωθεί ότι το $r^{(i)}$ πρέπει να υπολογίζεται με μεγάλη ακρίβεια. Η παραπάνω διαδικασία είναι ιδιαίτερα χρήσιμη για ασταθή συστήματα και συγκλίνει γενικά πολύ γρήγορα (δηλαδή, $\varepsilon^{(i)} \to 0$ καθώς το i αυξάνει).

10.3 Επαναληπτικές μέθοδοι

Θεωρούμε πάλι το πρόβλημα της επιλύσεως του συστήματος (10.1) ή ισοδύναμα του (10.2), όπου $A \in \mathbb{R}^{n,n}$, $b \in \mathbb{R}^n$ και $\det(A) \neq 0$. Πολλές φορές, στην πράξη, έχουμε $n > 10^3$ ή $n > 10^4$ και επί πλέον ο πίνακας A είναι "αραιός", δηλαδή, τα περισσότερα από τα στοιχεία του είναι ίσα με μηδέν. Τέτοια συστήματα προκύπτουν π.χ. όταν οι μέθοδοι των

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

πεπερασμένων διαφορών ή του πεπερασμένου στοιχείου εφαρμοσθούν για την αριθμητική επίλυση προβλημάτων συνοριακών τιμών στις διαφορικές εξισώσεις με μερικές παραγώγους. Είναι φανερό ότι, στις περιπτώσεις αυτές, οι άμεσες μέθοδοι, που εξετάστηκαν προηγουμένως, απαιτούν πολύ μεγάλο αριθμό πράξεων (όπως προκύπτει από τις (10.34)) και επομένως είναι πρακτικά δύσκολη ή αδύνατη η χρησιμοποίησή τους. Οι επαναληπτικές μέθοδοι εργάζονται τελείως διαφορετικά και λόγω της απλότητας της λειτουργίας των προτιμούνται συνήθως για την λύση μεγάλων συστημάτων με αραιούς πίνακες συντελεστών, απαιτώντας χρόνους υπολογισμού μέσα σε λογικά όρια για την επίτευξη ικανοποιητικής ακρίβειας. Για συστήματα μετρίου μεγέθους ($n < 10^3$) προτιμούνται συνήθως οι άμεσες μέθοδοι. Μία επαναληπτική μέθοδος για την βελτίωση μιας προσεγγιστικής λύσεως έχει ήδη περιγραφεί στην παράγραφο 10.2.10. .

10.3.1 Βασικές επαναληπτικές μέθοδοι

Η κατασκευή μιας επαναληπτικής μεθόδου, που θα χρησιμοποιηθεί για την επίλυση του συστήματος

$$Ax = b \ , \tag{10.79}$$

γίνεται ώς εξής: Εάν M και N είναι πραγματικοί πίνακες, τέτοιοι ώστε ο M να είναι ομαλός και

$$A = M - N \ , \tag{10.80}$$

τότε το σύστημα (10.79) γράφεται ισοδύναμα

$$Mx = Nx + b \quad \text{ή}$$

$$x = M^{-1}Nx + M^{-1}b \ . \tag{10.81}$$

θέτοντας

$$G = M^{-1}N \ , \quad \kappa = M^{-1}b \tag{10.82}$$

η (10.81) γράφεται

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$x = Gx + \kappa \tag{10.83}$$

ή

$$(I - G)x = \kappa \ . \tag{10.84}$$

Αν $x^{(0)}$ είναι μία δοσμένη αρχική προσέγγιση της ακριβούς λύσεως $\hat{x} = A^{-1}b$ του (10.79), τότε μία ακολουθία $x^{(1)}, x^{(2)}, x^{(3)}, \ldots$ διανυσμάτων μπορεί να παραχθεί με βάση την (10.83), αν γράψουμε

$$x^{(m+1)} = Gx^{(m)} + \kappa \ , \quad m = 0,1,2,\ldots \ . \tag{10.85}$$

Η (10.85) αποτελεί την γενική μορφή μιας βασικής επαναληπτικής μεθόδου. Ο G καλείται επαναληπτικός πίνακας της μεθόδου (10.85) και λόγω της (10.80) έχουμε

$$G = M^{-1}N = M^{-1}(M - A) = I - M^{-1}A \ . \tag{10.86}$$

Έτσι, για μια συγκεκριμένη εκλογή του M τα G, κ είναι καθορισμένα. Πρέπει να σημειωθεί ότι αν και γενικά ο M μπορεί να εκλεγεί αυθαίρετα, εν τούτοις, για να είναι χρήσιμη πρακτικά η μέθοδος (10.85), που μπορεί να γραφεί επίσης

$$Mx^{(m+1)} = (M - A)x^{(m)} + b \ , \ m = 0,1,2,\ldots \ , \tag{10.87}$$

πρέπει το σύστημα (10.87) να λύνεται εύκολα ως προς $x^{(m+1)}$. Αυτό συμβαίνει π.χ. εάν ο M εκλεγεί έτσι ώστε να είναι τριγωνικός ή διαγώνιος. Όσον αφορά την ακολουθία $\{x^{(m)}\}$, $m = 1,2,\ldots$, που παράγει ο τύπος (αλγόριθμος) (10.85), ισχύει το εξής θεώρημα.

<u>Θεώρημα 10.1</u> Εάν η ακολουθία $x^{(1)}, x^{(2)}, \ldots$, που παράγεται από τον (10.85) συγκλίνει, τότε αυτή θα συγκλίνει στην ακριβή λύση $\hat{x}$ του (10.79).

<u>Απόδειξη</u> Εάν $y = \lim\limits_{m \to \infty} x^{(m)}$, τότε παίρνοντας όρια στα δύο μέλη της (10.85) έχουμε

$$y = Gy + \kappa \quad \text{ή} \quad (I - G)y = \kappa \ ,$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

δηλαδή, το y είναι λύση του (10.84). Αλλά το σύστημα (10.84) είναι ισοδύναμο προφανώς με το (10.79) και επομένως $y = \hat{x}$. ▲

Ορισμός 10.1 Η επαναληπτική μέθοδος (10.85) λέγεται ότι συγκλίνει (είναι συγκλίνουσα), εάν για κάθε αρχικό διάνυσμα $x^{(0)}$ η ακολουθία $x^{(1)}, x^{(2)}, \ldots,$ που παράγεται απ' αυτή, συγκλίνει στην ακριβή λύση $\hat{x}$ του (10.79) .

Θεώρημα 10.2 Η επαναληπτική μέθοδος (10.85) συγκλίνει εάν και μόνον εάν $\rho(G) < 1$.

Απόδειξη Το σφάλμα $\varepsilon^{(m)}$ στην επανάληψη $x^{(m)}$ είναι

$$\varepsilon^{(m)} = x^{(m)} - \hat{x} \quad . \tag{10.88}$$

Επίσης, αφού το σύστημα (10.79) είναι ισοδύναμο με το (10.83), ισχύει

$$\hat{x} = G\hat{x} + \kappa \quad . \tag{10.89}$$

Αφαιρώντας κατά μέλη τις (10.85), (10.89) παίρνουμε, λόγω της (10.88),

$$\varepsilon^{(m+1)} = G\varepsilon^{(m)} \quad , \; m = 0,1,2,\ldots \; . \tag{10.90}$$

Εάν η (10.90) εφαρμοσθεί επαναληπτικά προκύπτει εύκολα ότι

$$\varepsilon^{(m)} = G^m\varepsilon^{(0)} \quad , \; m = 0,1,2,\ldots \quad . \tag{10.91}$$

Υποθέτουμε τώρα ότι η μέθοδος (10.85) συγκλίνει, δηλαδή, για κάθε $x^{(0)}$ και επομένως για κάθε $\varepsilon^{(0)} = x^{(0)} - \hat{x}$, ισχύει $\lim\limits_{m\to\infty} \varepsilon^{(m)} = 0$. Έτσι, από την (10.91) προκύπτει ότι

$$\lim_{m\to\infty} \{G^m\varepsilon^{(0)}\} = 0 \; , \; \forall\, \varepsilon^{(0)} \; . \tag{10.92}$$

Η (10.92), σύμφωνα με το θεώρημα 9.83, συνεπάγεται ότι $\rho(G) < 1$.

Αντιστρόφως, εάν $\rho(G) < 1$, τότε σύμφωνα με το θεώρημα 9.83 θα ισχύει η (10.92), οπότε από την (10.91) έχουμε $\lim\limits_{m\to\infty} \varepsilon^{(m)} = 0$, $\forall\, \varepsilon^{(0)}$, δηλαδή, η (10.85) συγκλίνει. ▲

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

<u>Θεώρημα 10.3</u> Εάν $\|G\| \leq \lambda < 1$, τότε η μέθοδος (10.85) συγκλίνει και ισχύουν

i) $\|\varepsilon^{(m)}\| \leq \lambda^m \|\varepsilon^{(0)}\|$

ii) $\|\varepsilon^{(m)}\| \leq \dfrac{\lambda}{1-\lambda} \|x^{(m)} - x^{(m-1)}\|$, $m = 1,2,\ldots$

iii) $\|\varepsilon^{(m)}\| \leq \dfrac{\lambda^m}{1-\lambda} \|x^{(1)} - x^{(0)}\|$, $m = 1,2,\ldots$.

<u>Απόδειξη</u> Επειδή $\|G\| < 1$, από τα θεωρήματα 9.74 και 10.2 προκύπτει ότι η μέθοδος (10.85) συγκλίνει. Εξ' άλλου, από την (10.91) παίρνουμε

$$\|\varepsilon^{(m)}\| = \|G^m \varepsilon^{(0)}\| \leq \|G^m\| \|\varepsilon^{(0)}\| \leq \|G\|^m \|\varepsilon^{(0)}\| \leq \lambda^m \|\varepsilon^{(0)}\| .$$

Για την απόδειξη των (ii) και (iii), από την (10.85) παίρνουμε

$$x^{(m+1)} = Gx^{(m)} + \kappa$$
$$x^{(m)} = Gx^{(m-1)} + \kappa ,$$

οπότε

$$x^{(m+1)} - x^{(m)} = G(x^{(m)} - x^{(m-1)}) , \quad m = 1,2,\ldots . \qquad (10.93)$$

Λόγω της (10.90) θα είναι

$$\varepsilon^{(m)} = G\varepsilon^{(m-1)} = G(x^{(m-1)} - \hat{x}) = G(x^{(m-1)} - x^{(m)} + x^{(m)} - \hat{x}) = G(x^{(m-1)} - x^{(m)}) + G\varepsilon^{(m)}$$

και επομένως

$$\|\varepsilon^{(m)}\| = \|G(x^{(m-1)} - x^{(m)}) + G\varepsilon^{(m)}\| \leq \|G\| \|x^{(m-1)} - x^{(m)}\| + \|G\| \|\varepsilon^{(m)}\| \quad \text{ή}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\|\varepsilon^{(m)}\| \leqq \frac{\|G\|}{1-\|G\|}\|x^{(m-1)}-x^{(m)}\| \leqq \frac{\lambda}{1-\lambda}\|x^{(m)}-x^{(m-1)}\|. \quad (10.94)$$

Από την (10.93) προκύπτει με επαναληπτική εφαρμογή ότι

$$x^{(m+1)}-x^{(m)} = G^m(x^{(1)}-x^{(0)}) \quad \text{ή ισοδύναμα}$$

$$x^{(m)}-x^{(m-1)} = G^{m-1}(x^{(1)}-x^{(0)}) . \quad (10.95)$$

Έτσι, από την (10.95) έχουμε

$$\|x^{(m)}-x^{(m-1)}\| \leqq \|G\|^{m-1}\|x^{(1)}-x^{(0)}\| \leqq \lambda^{m-1}\|x^{(1)}-x^{(0)}\|$$

και η (10.94) γίνεται

$$\|\varepsilon^{(m)}\| \leqq \frac{\lambda^m}{1-\lambda}\|x^{(1)}-x^{(0)}\| . \quad (10.96)$$

Αν εκλέξουμε $x^{(0)} = \kappa$, τότε είναι $x^{(1)} = G\kappa + \kappa$ και επομένως

$$\|x^{(1)}-x^{(0)}\| = \|G\kappa\| \leqq \|G\|\|\kappa\| \leqq \lambda\|\kappa\| .$$

Έτσι, η (10.96) δίνει

$$\|\varepsilon^{(m)}\| \leqq \frac{\lambda^{m+1}}{1-\lambda}\|\kappa\| . \qquad \blacktriangle$$

Πρέπει να σημειωθεί ότι το θεώρημα 10.3 έχει μεγαλύτερη πρακτική αξία από το 10.2 , γιατί οι ιδιοτιμές του πίνακα G δεν είναι γνωστές και ο προσδιορισμός της $\rho(G)$ απαιτεί γενικά μεγαλύτερη προσπάθεια από το να λύσουμε το σύστημα (10.79) που είναι ο σκοπός μας. Εάν δεν έχουμε καμμιά πληροφορία για την λύση $\hat{x}$, τότε συνήθως εκλέγουμε $x^{(0)} = 0$.

Έχοντας στη διάθεσή μας μία συγκλίνουσα επαναληπτική μέθοδο (10.85), εκείνο που ζητάμε συνήθως στην πράξη είναι να υπολογίσουμε μία προσέγγιση $\bar{x}$ της $\hat{x}$ με κάποια δεδομένη ακρίβεια, π.χ. ℓ δεκαδικών θέσεων. (Η ακριβής λύση $\hat{x}$ δεν μπορεί γενικά να βρεθεί και αν ακόμη γίνουν άπειρες επαναλήψεις, λόγω των σφαλμάτων στρογγυλεύσεως).

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Για τον σκοπό αυτό ο υπολογισμός όρων (επαναλήψεων) της ακολουθίας που παράγεται από την (10.85) συνεχίζεται μέχρις ότου βρεθούν δύο διαδοχικές επαναλήψεις, τέτοιες ώστε

$$\|x^{(m+1)} - x^{(m)}\| \leqq \varepsilon = \frac{1}{2}\, 10^{-\ell} . \tag{10.97}$$

Τότε, δεχόμαστε ότι $\bar{x} = x^{(m+1)}$. Αν $x^{(m)} = (x_i^{(m)})$, $i = 1(1)n$, $m = 0,1,2,\ldots$, τότε η (10.97) συνεπάγεται τις

$$|x_i^{(m+1)} - x_i^{(m)}| \leqq \varepsilon = \frac{1}{2}\, 10^{-\ell} , \quad i = 1(1)n .$$

Ας υποθέσουμε ότι έχουμε μία συγκλίνουσα επαναληπτική μέθοδο (10.85), οπότε $\rho(G) < 1$. Για να έχει η μέθοδος αυτή πρακτική αξία πρέπει να συγκλίνει γρήγορα, δηλαδή, η ταχύτητα με την οποία η $\|\varepsilon^{(m)}\|$ τείνει στο 0 καθώς το $m \to \infty$ να είναι μεγάλη. Με σκοπό να ορίσουμε την ταχύτητα συγκλίσεως της μεθόδου, θα εξετάσουμε τον αριθμό των επαναλήψεων, που απαιτούνται για να επιτύχουμε κάποια δεδομένη ακρίβεια στην προσεγγιστική μας λύση. Η εκτίμηση του αριθμού αυτού των επαναλήψεων γίνεται με τη βοήθεια της (10.91) από την οποία προκύπτει ότι

$$\|\varepsilon^{(m)}\| = \|G^m \varepsilon^{(0)}\| \leqq \|G^m\| \|\varepsilon^{(0)}\| . \tag{10.98}$$

Λόγω των θεωρημάτων 9.76 και 9.80, $\|G^m\| \underset{m \to \infty}{\to} 0$ και επομένως, για επαρκώς μεγάλα m, ώστε $\|G^m\| < 1$, το αρχικό σφάλμα $\|\varepsilon^{(0)}\|$ ελαττώνεται τουλάχιστον κατά τον παράγοντα $\|G^m\|$ μετά από m επαναλήψεις. Εάν απαιτήσουμε να ισχύει

$$\|\varepsilon^{(m)}\| \leqq \rho \|\varepsilon^{(0)}\| , \tag{10.99}$$

όπου ρ ένας προκαθορισμένος παράγων $(0 < \rho < 1)$, τότε ο μικρότερος αριθμός των επαναλήψεων m που απαιτούνται για να ισχύει η (10.99) μπορεί να εκτιμηθεί από την

$$\|G^m\| = \rho . \tag{10.100}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Επειδή για επαρκώς μεγάλα m, λόγω του θεωρήματος 9.85, θα ισχύει

$$\|G^m\| \approx [\rho(G)]^m ,$$

η (10.100) αντικαθίσταται από την

$$[\rho(G)]^m \approx \rho ,$$

η οποία δίνει

$$m \approx \frac{-\ln \rho}{-\ln \rho(G)} . \qquad (10.101)$$

Ορίζουμε την ασυμπτωτική ταχύτητα συγκλίσεως ή απλώς ταχύτητα συγκλίσεως της μεθόδου (10.85) από την

$$R(G) \equiv -\ln \rho(G) . \qquad (10.102)$$

Είναι φανερό από την (10.102) ότι όσο μικρότερη είναι η $\rho(G)$, τόσο μεγαλύτερη είναι η ταχύτητα συγκλίσεως της μεθόδου.

10.3.2 Η μέθοδος Jacobi

Ο πίνακας $A = (a_{ij})$ του συστήματος (10.79) μπορεί πάντοτε να γραφεί ως

$$A = D - C_L - C_U , \qquad (10.103)$$

όπου

$$D = \begin{bmatrix} a_{11} & 0 & 0 \dots & 0 \\ 0 & a_{22} & 0 \dots & 0 \\ 0 & 0 & a_{33} \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 \dots & a_{nn} \end{bmatrix} , \quad C_L = -\begin{bmatrix} 0 & 0 & 0 \dots & 0 \\ a_{21} & 0 & 0 \dots & 0 \\ a_{31} & a_{32} & 0 \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} \dots & 0 \end{bmatrix} ,$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$C_U = -\begin{bmatrix} 0 & a_{12} & a_{13} \dots & a_n \\ 0 & 0 & a_{23} \dots & a_{2n} \\ 0 & 0 & 0 \dots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 \dots & 0 \end{bmatrix} .$$

Η απλούστερη βασική επαναληπτική μέθοδος προκύπτει εάν στην (10.80) εκλεγεί $M = D$, με την προυπόθεση ότι ο D είναι ομαλός, δηλαδή, $a_{ii} \neq \neq 0$, $i = 1(1)n$. (Εάν $\det(D) = 0$, τότε, επειδή $\det(A) \neq 0$, μπορούμε πάντοτε να αναδιατάξουμε τις εξισώσεις του (10.79) έτσι ώστε τα διαγώνια στοιχεία του πίνακα των συντελεστών του νέου ισοδύναμου συστήματος που προκύπτει να είναι διάφορα του μηδενός). Έτσι, έχοντας υπ' όψιν τις (10.80), (10.82), (10.86) και (10.103) βρίσκουμε

$$N = C_L + C_U \ , \quad G = D^{-1}(C_L + C_U) = I - D^{-1}A \ , \ \kappa = D^{-1}b$$

και η (10.85) γίνεται

$$x^{(m+1)} = D^{-1}(C_L + C_U)x^{(m)} + D^{-1}b \ , \quad m = 0,1,2,\dots \ . \quad (10.104)$$

Η (10.104) είναι γνωστή σαν η μέθοδος Jacobi. Ο επαναληπτικός πίνακας αυτής θα συμβολίζεται με J και είναι

$$J = D^{-1}(C_L + C_U) = L + U = I - D^{-1}A \ , \quad (10.105)$$

όπου

$$L \equiv D^{-1}C_L \ , \quad U \equiv D^{-1}C_U \ . \quad (10.106)$$

Εάν γράψουμε την (10.104) ως εξής

$$Dx^{(m+1)} = (C_L + C_U)x^{(m)} + b \ ,$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

τότε για την i συνιστώσα του πρώτου μέλους θα έχουμε

$$a_{ii}x_i^{(m+1)} = -\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}x_j^{(m)} + b_i \quad , \quad i = 1(1)n$$

ή

$$x_i^{(m+1)} = \frac{1}{a_{ii}}\left[b_i - \sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}x_j^{(m)}\right] \quad , \; i = 1(1)n \; , \; m = 0,1,2,\ldots \quad (10.10$$

Στην πράξη, χρησιμοποιείται συνήθως η (10.107) για την εύρεση των διαδοχικών επαναλήψεων από τις συνιστώσες τους.

<u>Θεώρημα 10.4</u> Εάν ο πίνακας A του συστήματος (10.79) είναι αυστηρά διαγώνια υπέρτερος (βλέπε ορισμό 9.13i), τότε η μέθοδος Jacobi συγκλίνει.

<u>Απόδειξη</u> Έχουμε

$$\|J\|_\infty = \|D^{-1}(C_L + C_U)\|_\infty = \max_i \sum_{\substack{j=1\\ j\neq i}}^{n} \left|\frac{a_{ij}}{a_{ii}}\right| = \max_i \left\{\frac{1}{|a_{ii}|}\sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}|\right\} .$$

Αφού ο A είναι αυστηρά διαγώνια υπέρτερος θα ισχύει

$$|a_{ii}| > \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}| \; , \; \forall\, i = 1(1)n \; , \qquad (10.108)$$

οπότε προκύπτει ότι $\|J\|_\infty < 1$ και σύμφωνα με το θεώρημα 10.3 η μέθοδος Jacobi συγκλίνει. ▲

<u>Σημείωση</u> Μπορεί να αποδειχθεί ότι ένας αυστηρά διαγώνια υπέρτερος πίνακας A είναι ομαλός και $a_{ii} \neq 0$, $i = 1(1)n$. Συχνά, είναι δυνατόν ένας πίνακας να μετατραπεί σε ένα αυστηρώς διαγώνια υπέρτερο πίνακα με κατάλληλη αναδιάταξη των γραμμών του.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Παράδειγμα 10.6 Δίνεται το σύστημα

$$\begin{aligned} x_1 + 2x_2 - 2x_3 &= 1 \\ x_1 + x_2 + x_3 &= 1 \\ 2x_1 + 2x_2 + x_3 &= 1 \end{aligned}$$

Να δειχθεί ότι η μέθοδος Jacobi συγκλίνει και να βρεθούν οι 3 πρώτες επαναλήψεις, αν $x^{(0)} = (1,1,1)^T$.

Λύση Ο επαναληπτικός πίνακας της μεθόδου Jacobi, σύμφωνα με την (10.105), είναι

$$J = I - D^{-1}A = I - A = \begin{bmatrix} 0 & -2 & 2 \\ -1 & 0 & -1 \\ -2 & -2 & 0 \end{bmatrix} .$$

Είναι $\|J\|_1 = \|J\|_\infty = 4 > 1$ και γι' αυτό θα βρούμε την $\rho(J)$. Από την $\det(J - \lambda I) = 0$ έχουμε ότι οι ιδιοτιμές του J είναι $\lambda_1 = \lambda_2 = \lambda_3 = 0$ και επομένως $\rho(J) = 0 < 1$, δηλαδή, η μέθοδος Jacobi συγκλίνει. Από την (10.107) έχουμε

$$\begin{aligned} x_1^{(m+1)} &= 1 - 2x_2^{(m)} + 2x_3^{(m)} \\ x_2^{(m+1)} &= 1 - x_1^{(m)} - x_3^{(m)} \quad , \quad m = 0,1,2,\ldots . \qquad (10.109) \\ x_3^{(m+1)} &= 1 - 2x_1^{(m)} - 2x_2^{(m)} \end{aligned}$$

Για $x^{(0)} = (1,1,1)^T$, δηλαδή, $x_1^{(0)} = x_2^{(0)} = x_3^{(0)} = 1$, από τις (10.109) βρίσκουμε

$$\begin{aligned} x_1^{(1)} &= 1 , \; x_2^{(1)} = -1 , \; x_3^{(1)} = -3 \Rightarrow x^{(1)} = (1, -1, -3)^T \\ x_1^{(2)} &= -3 , \; x_2^{(2)} = 3 , \; x_3^{(2)} = 1 \Rightarrow x^{(2)} = (-3, 3, 1)^T \\ x_1^{(3)} &= -3 , \; x_2^{(3)} = 3 , \; x_3^{(3)} = 1 \Rightarrow x^{(3)} = (-3, 3, 1)^T . \end{aligned}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Αφού $x^{(3)} = x^{(2)}$ συμπεραίνουμε ότι η λύση του συστήματος είναι η $\hat{x} = (-3, 3, 1)^T$. Στο παράδειγμα αυτό, στην ουσία, βρήκαμε την ακριβή λύση με την δεύτερη επανάληψη. Η πολύ ταχεία αυτή σύγκλιση εξηγείται από το γεγονός ότι $\rho(J) = 0$. ▲

10.3.3 Η μέθοδος Gauss – Seidel

Έχοντας υπ' όψιν τις (10.80) και (10.103), μία άλλη βασική επαναληπτική μέθοδος προκύπτει εάν εκλέξουμε $M = D - C_L$, οπότε $N = C_U$. Υποθέτουμε ότι ο $D - C_L$ είναι ομαλός, δηλαδή, $a_{ii} \neq 0$, $i = 1(1)n$. Διαφορετικά, ισχύει η ίδια παρατήρηση που έγινε και στην μέθοδο Jacobi. Έτσι, από την (10.85) προκύπτει η μέθοδος

$$x^{(m+1)} = (D - C_L)^{-1}C_U x^{(m)} + (D - C_L)^{-1}b \ , \ m = 0,1,2,\ldots, \qquad (10.110)$$

που είναι γνωστή σαν μέθοδος των Gauss – Seidel. Πρέπει να σημειωθεί ότι ο $D - C_L$ είναι κάτω τριγωνικός και ο επαναληπτικός πίνακας της μεθόδου (10.110), που θα συμβολίζεται με S , με βάση τις (10.106) μπορεί να γραφεί

$$S = (D - C_L)^{-1}C_U = (I - L)^{-1}U = I - (I - L)^{-1}D^{-1}A \ . \qquad (10.111)$$

Εάν η (10.110) γραφεί με την μορφή .

$$(D - C_L)x^{(m+1)} = C_U x^{(m)} + b \ , \quad m = 0,1,2,\ldots \ , \qquad (10.112)$$

τότε το $x^{(m+1)}$ υπολογίζεται εύκολα με τη μέθοδο της προς τα εμπρός αντικατάστασης. Για την εύρεση σχέσεως αντίστοιχης της (10.107), από την (10.112) έχουμε

$$\sum_{j=1}^{i} a_{ij}x_j^{(m+1)} = -\sum_{j=i+1}^{n} a_{ij}x_j^{(m)} + b_i \ , \ i = 1(1)n$$

και λύνοντας ως προς $x_i^{(m+1)}$ παίρνουμε

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$x_i^{(m+1)} = \frac{1}{a_{ii}} \left[b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(m+1)} - \sum_{j=i+1}^{n} a_{ij} x_i^{(m)} \right] ,$$

$$i = 1(1)n \ , \quad m = 0,1,2,\ldots \ . \tag{10.113}$$

Συγκρίνοντας τις (10.107) και (10.113) βλέπουμε ότι η διαφορά τους έγκειται στο ότι με την μέθοδο Gauss – Seidel για τον υπολογισμό της $x_i^{(m+1)}$ χρησιμοποιούμε τις $x_{i-1}^{(m+1)}$, $x_{i-2}^{(m+1)}$,..., $x_1^{(m+1)}$ αντί των $x_{i-1}^{(m)}$, $x_{i-2}^{(m)}$,..., $x_1^{(m)}$ αντίστοιχα, που χρησιμοποιούμε στην μέθοδο Jacobi. Με άλλα λόγια, η μέθοδος Gauss – Seidel χρησιμοποιεί τις πιο πρόσφατα υπολογισμένες προσεγγίσεις των συνιστωσών, μόλις αυτές γίνονται διαθέσιμες. Το γεγονός αυτό δεν σημαίνει αναγκαστικά για ένα οποιοδήποτε δεδομένο σύστημα,για το οποίο η μέθοδος Jacobi συγκλίνει, ότι η μέθοδος Gauss – Seidel θα συγκλίνει ταχύτερα. Πράγματι, υπάρχουν περιπτώσεις όπου η Jacobi συγκλίνει,ενώ η Gauss – Seidel δεν συγκλίνει αλλά και αντιστρόφως. Γενικά, δεν μπορούμε να πούμε ποια από τις δύο θα είναι ταχύτερη.

<u>Θεώρημα 10.5</u> Εάν στο σύστημα (10.79) ο A είναι αυστηρά διαγώνια υπέρτερος, τότε η μέθοδος Gauss – Seidel συγκλίνει και επιπλέον ισχύει

$$\|S\|_\infty \leqq \|J\|_\infty < 1 \ ,$$

όπου S και J ορίζονται στις (10.111) και (10.105) αντίστοιχα.

<u>Απόδειξη</u> Σύμφωνα με το θεώρημα 10.4 θα είναι $\|J\|_\infty < 1$. Εάν $e = (1,1,\ldots,1)^T$, τότε για τον πίνακα $|J| = |L| + |U|$ έχουμε

$$|J|e \leqq \|J\|_\infty e < e \qquad \text{ή}$$

$$(|L| + |U|)e \leqq \|J\|_\infty e \qquad \text{ή}$$

$$|U|e \leqq (\|J\|_\infty \ I - |L|)e \ . \tag{10.114}$$

Επειδή οι L , $|L|$ είναι αυστηρά κάτω τριγωνικοί θα ισχύει $L^n =$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$L^n = |L|^n = 0$, θα υπάρχουν οι $(I - L)^{-1}$ και $(I - |L|)^{-1}$ και επί πλέον

$$0 \leqq |(I - L)^{-1}| = |I + L + L^2 + \ldots + L^{n-1}| \leqq$$
$$\leqq I + |L| + |L|^2 + \ldots + |L|^{n-1} = (I - |L|)^{-1} .$$

Αφού $S = (I - L)^{-1}U$, λόγω της (10.114), παίρνουμε

$$|S|e = |(I - L)^{-1}U|e \leqq |(I - L)^{-1}||U|e \leqq (I - |L|)^{-1}|U|e \leqq$$
$$\leqq (I - |L|)^{-1}(\|J\|_\infty I - |L|)e = (I - |L|)^{-1}\Big[I - |L| + (\|J\|_\infty - 1)I\Big]e =$$
$$= \Big[I + (\|J\|_\infty - 1)(I - |L|)^{-1}\Big]e \leqq \Big[I + (\|J\|_\infty - 1)I\Big]e = \|J\|_\infty e .$$

Η τελευταία ανισότητα ισχύει επειδή $(I - |L|)^{-1} \geqq I$ και $\|J\|_\infty < 1$. Άρα, $|S|e \leqq \|J\|_\infty e$, που σημαίνει ότι

$$\|S\|_\infty \leqq \|J\|_\infty < 1$$

και σύμφωνα με το θεώρημα 10.3 η μέθοδος Gauss – Seidel συγκλίνει. ▲

<u>Θεώρημα 10.6</u> Εάν στο σύστημα (10.79) ο A είναι ένας πραγματικός συμμετρικός θετικά ορισμένος πίνακας (βλέπε ορισμό 9.16), τότε η μέθοδος Gauss – Seidel συγκλίνει.

<u>Απόδειξη</u> Αφού ο A είναι συμμετρικός, είναι $A = D - C_L - C_U$ με $C_U = C_L^T$ και $S = (D - C_L)^{-1}C_U = (D - C_L)^{-1}C_L^T$. Αρκεί να δείξουμε ότι $\rho(S) < 1$. Εάν λ είναι μία ιδιοτιμή του S και u ένα αντίστοιχο ιδιοδιάνυσμα $(\lambda \in \mathbb{C}$, $u \in \mathbb{C}^n)$, τότε $Su = \lambda u$ ή $(D - C_L)^{-1}C_L^T u = \lambda u$ ή

$$C_L^T = \lambda(D - C_L)^{-1}u \quad . \tag{10.115}$$

Πολλαπλασιάζοντας τα δύο μέλη της (10.115) από αριστερά με u^H παίρνουμε

$$u^H C_L^T u = \lambda u^H (D - C_L)u \ . \tag{10.116}$$

Αφού ο A είναι θετικά ορισμένος θα έχουμε $u^H A u > 0$ και $a_{ii} > 0$,

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

οπότε ο D είναι και αυτός θετικά ορισμένος. Θέτουμε $u^H Du = \gamma > 0$, $u^H C_L u = \alpha + i\beta$, οπότε $u^H C_L^T u = (u^H C_L u)^H = \alpha - i\beta$ και $u^H Au = \gamma - 2\alpha > 0$. Στην (10.116) είναι $u^H(D - C_L)u \neq 0$, διότι αν ήταν ίσο με 0, τότε θα είχαμε $\gamma - (\alpha + i\beta) = 0$, δηλαδή, $\gamma = \alpha > 0$ και $\beta = 0$, ενώ από την $u^H Au > 0$ προκύπτει ότι $-\alpha > 0$, πράγμα που είναι άτοπο. Επομένως, από την (10.116) έχουμε

$$\lambda = \frac{u^H C_L^T u}{u^H Du - u^H C_L u} = \frac{\alpha - i\beta}{\gamma - (\alpha + i\beta)} = \frac{\alpha - i\beta}{(\gamma - \alpha) - i\beta}$$

και $|\lambda|^2 = \dfrac{\alpha^2 + \beta^2}{(\gamma - \alpha)^2 + \beta^2} < 1$, αφού $\gamma(\gamma - 2\alpha) > 0$. Έτσι, ισχύει $\rho(S) < 1$ και επομένως η μέθοδος Gauss – Seidel συγκλίνει. ▲

Αξίζει να σημειωθεί ότι εάν τα δύο μέλη ενός γραμμικού συστήματος $Ax = b$, όπου $\det(A) \neq 0$, πολλαπλασιαστούν από αριστερά με A^T, τότε, στο ισοδύναμο σύστημα $A^T Ax = A^T b$, που προκύπτει, ο $A^T A$ είναι θετικά ορισμένος και επομένως η μέθοδος Gauss – Seidel θα συγκλίνει.

Το παρακάτω θεώρημα δίνεται χωρίς απόδειξη.

<u>Θεώρημα 10.7</u> (Stein, Rosenberg) Εάν για τον επαναληπτικό πίνακα της μεθόδου Jacobi ισχύει $J \geqq 0$, τότε για τους J και S ακριβώς μία από τις παρακάτω σχέσεις ισχύει:

i) $\rho(S) = \rho(J) = 0$

ii) $0 < \rho(S) < \rho(J) < 1$

iii) $\rho(S) = \rho(J) = 1$

iv) $1 < \rho(J) < \rho(S)$. ▲

<u>Σημείωση</u>: Εάν ο πίνακας A στο σύστημα (10.79) είναι τέτοιος ώστε $a_{ii} > 0$, $a_{ij} \leqq 0$, $i,j = 1(1)n$, $i \neq j$, τότε η προϋπόθεση $J \geqq 0$ του θεωρήματος ικανοποιείται.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

<u>Παράδειγμα 10.7</u> Να βρεθούν οι τρείς πρώτες επαναλήψεις της μεθόδου Gauss – Seidel για το σύστημα

$$\begin{cases} 4x_1 - x_2 = 1 \\ -x_1 + 4x_2 - x_3 = 0 \\ -x_2 + 4x_3 = 1 \end{cases}, \quad \text{αν} \quad x^{(0)} = 0 .$$

<u>Λύση</u> 0 $A = \begin{bmatrix} 4 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 4 \end{bmatrix}$ είναι αυστηρά διαγώνια υπέρτερος

και επομένως η μέθοδος Gauss – Seidel συγκλίνει. Σύμφωνα με την (10.113) έχουμε

$$x_1^{(m+1)} = \frac{1}{4}(1 + x_2^{(m)})$$

$$x_2^{(m+1)} = \frac{1}{4}(x_1^{(m+1)} + x_3^{(m)})$$

$$x_3^{(m+1)} = \frac{1}{4}(1 + x_2^{(m+1)}), \qquad m = 0,1,2,\ldots$$

και επομένως με $x_1^{(0)} = x_2^{(0)} = x_3^{(0)} = 0$ βρίσκουμε

$$x_1^{(1)} = \frac{1}{4}(1 + 0) = 0.25$$

$$x_2^{(1)} = \frac{1}{4}(0.25 + 0) = 0.125$$

$$x_3^{(1)} = \frac{1}{4}(1 + 0.125) = 0.28125 ,$$

$$x_1^{(2)} = \frac{1}{4}(1 + 0.125) = 0.28125$$

$$x_2^{(2)} = \frac{1}{4}(0.28125 + 0.28125) = 0.140625$$

$$x_3^{(2)} = \frac{1}{4}(1 + 0.140625) \approx 0.28516 ,$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$x_1^{(3)} \approx 0.28516$$

$$x_2^{(3)} \approx 0.14258$$

$$x_3^{(3)} \approx 0.28564 \quad .$$

Άρα, $\hat{x} \approx (0.28516,\ 0.14258,\ 0.28564)^T$. Σημειώνεται ότι η ακριβής λύση $\hat{x}$ του συστήματος είναι $\hat{x} = (\frac{2}{7}, \frac{1}{7}, \frac{2}{7})^T \approx (0{,}28571,\ 0.14286,\ 0{,}28571)^T$.

10.3.4 Η μέθοδος SOR (Successive OverRelaxation)

Μία άλλη σημαντική μέθοδος, η οποία αποτελεί τροποποίηση της μεθόδου Gauss - Seidel με σκοπό να βελτιωθεί η ταχύτητα συγκλίσεως, είναι η γνωστή ως SOR (successive overrelaxation) μέθοδος. Για την δημιουργία αυτής, με βάση τις (10.80) και (10.103), θεωρούμε την εκλογή

$$M = M(\omega) = \frac{1}{\omega} D - C_L \ , \tag{10.117}$$

όπου ω είναι μία πραγματική παράμετρος $\neq 0$. Έτσι, κατά τα γνωστά η SOR μέθοδος είναι

$$x^{(m+1)} = (\frac{1}{\omega} D - C_L)^{-1} \left[C_U - (1 - \frac{1}{\omega})D\right] x^{(m)} + (\frac{1}{\omega} D - C_L)^{-1} b \ , \tag{10.118}$$

$m = 0,1,2,\ldots$, με την προυπόθεση ότι ο $M(\omega)$ είναι ομαλός, δηλαδή, $a_{ii} \neq 0$, $i = 1(1)n$ (ο $M(\omega)$ είναι κάτω τριγωνικός με διαγώνια στοιχεία $\frac{a_{ii}}{\omega}$, $i = 1(1)n$).

Ο επαναληπτικός πίνακας της (10.118) θα συμβολίζεται με $S(\omega)$ και παίρνει την μορφή

$$S(\omega) = \omega(D - \omega C_L)^{-1} \frac{1}{\omega}\left[(1 - \omega)D + \omega C_U\right] =$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$= (D - \omega C_L)^{-1}\Big[(1-\omega)D + \omega C_U\Big] = (I - \omega L)^{-1}\Big[(1-\omega)I + \omega U\Big] =$$

$$= I - \omega(I - \omega L)^{-1}D^{-1}A \ . \qquad (10.119)$$

Έτσι, η (10.118) μπορεί να γραφεί επίσης

$$x^{(m+1)} = (I - \omega L)^{-1}\Big[(1-\omega)I + \omega U\Big]x^{(m)} + \omega(I - \omega L)^{-1}D^{-1}b \ , \qquad (10.120)$$

όπου τα L , U ορίζονται στην (10.106). Επειδή ο $S(\omega)$ εξαρτάται από την παράμετρο ω (relaxation παράμετρος), βασικός σκοπός είναι να εκλεγεί το ω έτσι ώστε η μέθοδος να συγκλίνει όσο γίνεται πιο γρήγορα, δηλαδή, η $\rho(S(\omega))$ να ελαχιστοποιείται ως προς ω . Παρατηρούμε ότι για $\omega = 1$ η SOR συμπίπτει με την μέθοδο Gauss – Seidel. Για την εύρεση του επαναληπτικού αλγορίθμου της SOR αναλυτικά με τις συνιστώσες, από την (10.118) εύκολα προκύπτει ότι

$$a_{ii}x_i^{(m+1)} + \omega\sum_{j=1}^{i-1} a_{ij}x_j^{(m+1)} = (1-\omega)a_{ii}x_i^{(m)} - \omega\sum_{j=i+1}^{n} a_{ij}x_j^{(m)} + \omega b_i \ ,$$

$i = 1(1)n$, οπότε

$$x_i^{(m+1)} = (1-\omega)x_i^{(m)} + \frac{\omega}{a_{ii}}\left[b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(m+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(m)}\right] ,$$

$$i = 1(1)n \ , \quad m = 0,1,2,\ldots \ . \qquad (10.121)$$

Για την $\rho(S(\omega))$ ισχύει το ακόλουθο θεώρημα .

<u>Θεώρημα 10.8</u> (Kahan) Είναι $\rho(S(\omega)) \geq |\omega - 1|$. Επί πλέον, εάν η SOR συγκλίνει, τότε $0 < \omega < 2$.

<u>Απόδειξη</u> Αν λ_i , $i = 1(1)n$ είναι οι ιδιοτιμές του $S(\omega)$, τότε, σύμφωνα με το θεώρημα 9.17, έχουμε $\det(S(\omega)) = \prod_{i=1}^{n} \lambda_i$. Λόγω της (10.119), είναι

$$\det(S(\omega)) = \det\Big((I - \omega L)^{-1}\Big[(1-\omega)I + \omega U\Big]\Big) = \qquad (10.122)$$

$$= \det\Big((I - \omega L)^{-1}\Big)\det\Big((1-\omega)I + \omega U\Big) =$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$= \left[\det(I - \omega L)\right]^{-1} \det\left[(1-\omega)I + \omega U\right] = 1 \cdot (1-\omega)^n = (1-\omega)^n ,$$

αφού ο $I - \omega L$ είναι κάτω τριγωνικός με μονάδες στη διαγώνιο και ο $(1-\omega)I + \omega U$ άνω τριγωνικός με διαγώνια στοιχεία είναι ίσα με $1-\omega$. Άρα,

$$\left[\rho(S(\omega))\right]^n \geqq \prod_{i=1}^{n} |\lambda_i| = \left| \prod_{i=1}^{n} \lambda_i \right| = |(1-\omega)^n| = |1-\omega|^n , \text{ ή}$$

$$\rho(S(\omega)) \geqq |1-\omega| . \qquad (10.123)$$

Εάν η SOR συγκλίνει, τότε δεν μπορεί να είναι $|1-\omega| \geqq 1$, διότι από την (10.123) θα είχαμε $\rho(S(\omega)) \geqq 1$, πράγμα που είναι άτοπον. Άρα $|1-\omega| < 1$ ή $0 < \omega < 2$. ▲

Για την απόδειξη (η οποία παραλείπεται) του ακόλουθου θεωρήματος μπορεί να χρησιμοποιηθεί τρόπος ανάλογος εκείνου του θεωρήματος 10.6.

<u>Θεώρημα 10.9</u> (Ostrowski, Reich) Εάν στο σύστημα (10.79) ο A είναι ένας πραγματικός συμμετρικός θετικά ορισμένος πίνακας, τότε για κάθε $0 < \omega < 2$ η SOR μέθοδος συγκλίνει. ▲

Είναι φανερό ότι εάν η μέθοδος Gauss – Seidel $(\omega = 1)$ συγκλίνει για ένα δεδομένο σύστημα, τότε, λόγω της συνέχειας ως προς ω , η μέθοδος SOR θα συγκλίνει για ένα σύνολο τιμών του ω , που περιέχει την τιμή $\omega = 1$, και είναι υποσύνολο του $(0,2)$. Για να συγκλίνει η SOR μέθοδος όσο γίνεται πιο γρήγορα πρέπει να προσδιορίσουμε την "άριστη" τιμή ω_{opt} της relaxation παραμέτρου ω , η οποία χαρακτηρίζεται από την

$$\rho(S(\omega_{opt})) = \min_{\omega \in \mathbb{R}} \rho(S(\omega)) = \min_{0<\omega<2} \rho(S(\omega)) . \qquad (10.124)$$

Δυστυχώς, ο αναλυτικός προσδιορισμός της ω_{opt} πετυχαίνεται μόνον όταν ο πίνακας A στο σύστημα $Ax = b$ ανήκει σε ορισμένες ειδικές κατηγορίες πινάκων που παρουσιάζουν πρακτικό ενδιαφέρον, στις οποίες όμως δεν θα αναφερθούμε.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

10.3.5 Άλλες επαναληπτικές μέθοδοι

Μία άλλη απλή μέθοδος προκύπτει εάν στην (10.80) εκλεγεί $M = I$, οπότε η (10.85) γίνεται

$$x^{(m+1)} = (I - A)x^{(m)} + b \ , \quad m = 0,1,2,\dots \ . \tag{10.125}$$

Η μέθοδος (10.125) είναι γνωστή σαν μέθοδος του Richardson. Είναι φανερό ότι η (10.125) συγκλίνει εάν και μόνον εάν $|1 - \lambda_i| < 1$, $i = 1(1)n$, όπου λ_i, $i = 1(1)n$ είναι οι ιδιοτιμές του πίνακα A.

Εάν αντί της (10.117) θεωρήσουμε την εκλογή

$$M = M(r) = D - rC_L \ , \tag{10.126}$$

όπου r είναι μία πραγματική παράμετρος (relaxation παράμετρος) και με την προυπόθεση ότι ο $M(r)$ είναι ομαλός, δηλαδή, $a_{ii} \neq 0$, $i = 1(1)n$, προκύπτει η ακόλουθη επαναληπτική μέθοδος

$$x^{(m+1)} = (D - rC_L)^{-1}\Big[(1 - r)C_L + C_U\Big]x^{(m)} + (D - rC_L)^{-1}b \ , \tag{10.127}$$

$m = 0,1,2,\dots$. Ο επαναληπτικός πίνακας της (10.127) θα συμβολίζεται με $R(r)$ και είναι

$$R(r) = (D - rC_L)^{-1}\Big[(1 - r)C_L + C_U\Big] = (I - rL)^{-1}\Big[(1 - r)L + U\Big] =$$
$$= I - (I - rL)^{-1}D^{-1}A \ . \tag{10.128}$$

Για $r = 0$ η (10.127) συμπίπτει με την μέθοδο Jacobi, ενώ για $r = 1$ συμπίπτει με την μέθοδο Gauss – Seidel .

10.3.6 Επιτάχυνση της συγκλίσεως

Θεωρούμε πάλι την γενική μορφή (10.85) μιας βασικής επαναληπτικής μεθόδου

$$x^{(m+1)} = Gx^{(m)} + \kappa \ , \quad m = 0,1,2,\dots \ , \tag{10.129}$$

για την λύση του (10.79), η οποία προέκυψε με βάση την συγκεκριμένη

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

εκλογή $M = M_1$ στην (10.80) ($N_1 = M_1 - A$, $G = M_1^{-1}N_1 = I - M_1^{-1}A$, $\kappa = M_1^{-1}b$) . Εάν $\rho(G) \geqq 1$, οπότε η (10.129) δεν συγκλίνει, ή εάν $\rho(G)<1$ αλλά η σύγκλιση είναι αργή, τότε με σκοπό να βρούμε μέθοδο συγκλίνουσα ή να επιταχύνουμε την σύγκλιση αντίστοιχα, είναι δυνατόν, κάτω από ορισμένες προυποθέσεις, να χρησιμοποιήσουμε μία νέα μέθοδο που κατασκευάζεται με βάση την (10.129) και καλείται extrapolation της (10.129). Για την δημιουργία της extrapolation μεθόδου της (10.129) θεωρούμε στην (10.80) την εκλογή $M = M(\omega) = \frac{1}{\omega} M_1$, όπου ω είναι μία πραγματική παράμετρος $\neq 0$, η οποία καλείται extrapolation παράμετρος. Έτσι, η μέθοδος που προκύπτει είναι

$$x^{(m+1)} = \left(\frac{1}{\omega} M_1\right)^{-1} \left(\frac{1}{\omega} M_1 - A\right) x^{(m)} + \left(\frac{1}{\omega} M_1\right)^{-1} b \quad \text{ή}$$

$$x^{(m+1)} = (I - \omega M_1^{-1} A) x^{(m)} + \omega M_1^{-1} b \quad \text{ή}$$

$$x^{(m+1)} = G_\omega x^{(m)} + \omega\kappa \ , \ m = 0,1,2,\ldots \ , \tag{10.130}$$

όπου για τον επαναληπτικό πίνακα G_ω αυτής έχουμε, λόγω της $G = I - M_1^{-1}A$,

$$G_\omega = I - \omega M_1^{-1} A = I - \omega(I - G) = (1 - \omega) I + \omega G \ . \tag{10.131}$$

Για $\omega = 1$ η (10.130) συμπίπτει με την (10.129). Εάν λ_j , $j = 1(1)n$ είναι οι ιδιοτιμές του G , τότε από την (10.131) προκύπτει ότι οι ιδιοτιμές του G_ω είναι

$$\mu_j = 1 - \omega + \omega\lambda_j \ , \ j = 1(1)n \tag{10.132}$$

και επομένως η (10.130) συγκλίνει εάν και μόνον εάν

$$\rho(G_\omega) = \max_j |1 - \omega + \omega\lambda_j| < 1 \ .$$

Εάν $\lambda_j = a_j + ib_j$, $i = \sqrt{-1}$, $a_j, b_j \in \mathbb{R}$, $j = 1(1)n$, τότε έχουμε

$$|\mu_j|^2 = (1 - \omega + \omega a_j)^2 + \omega^2 b_j^2 \ . \tag{10.133}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

<u>Θεώρημα 10.10</u> Η extrapolation μέθοδος (10.130) συγκλίνει εάν και μόνον εάν μία από τις παρακάτω προτάσεις ισχύει:

i) $a_j < 1$ για κάθε $j = 1(1)n$ και

$$0 < \omega < \min_j \left\{ \frac{2(1-a_j)}{(1-a_j)^2 + b_j^2} \right\}$$

ii) $a_j > 1$ για κάθε $j = 1(1)n$ και $\max_j \left\{ \frac{2(1-a_j)}{(1-a_j)^2 + b_j^2} \right\} < \omega < 0$,

όπου $\lambda_j = a_j + ib_j$, $j = 1(1)n$ είναι οι ιδιοτιμές του πίνακα G της μεθόδου (10.129).

<u>Απόδειξη</u> Λόγω της (10.133), η μέθοδος (10.130) συγκλίνει εάν και μόνον εάν

$$|\mu_j|^2 = (1 - \omega + \omega a_j)^2 + \omega^2 b_j^2 = \left[1 + \omega(a_j - 1)\right]^2 + \omega^2 b_j^2 < 1, \quad (10.134)$$

για κάθε $j = 1(1)n$. Από την (10.133) είναι φανερό ότι, αν για κάποιο j είναι $a_j = 1$, τότε η (10.134) δεν μπορεί να ισχύει, αφού $|\mu_j|^2 \geqq 1$ $\forall\, \omega$. Ακόμη, παρατηρούμε ότι αν για κάποιο $j = j_1$ είναι $a_{j_1} < 1$ τότε $\forall\, \omega < 0$ είναι $|\mu_{j_1}|^2 > 1$, ενώ αν για κάποιο $j = j_2$ είναι $a_{j_2} > 1$, τότε $\forall\, \omega > 0$ είναι $|\mu_{j_2}|^2 > 1$. Από τα παραπάνω προκύπτει το συμπέρασμα ότι εάν η (10.134) ισχύει για κάποιες τιμές του ω , τότε πρέπει $a_j < 1$ για κάθε $j = 1(1)n$ ή $a_j > 1$ για κάθε $j = 1(1)n$. Υποθέτουμε τώρα ότι $a_j < 1$, $j = 1(1)n$. Θα δειχθεί ότι η (10.134) ισχύει (δηλαδή, η μέθοδος (10.130) συγκλίνει) εάν και μόνον εάν $0 < \omega < \min_j \left\{ \frac{2(1-a_j)}{(1-a_j)^2 + b_j^2} \right\}$. Πράγματι, η (10.134) είναι ισοδύναμη με την

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\omega\left(\omega\left[(1-a_j)^2+b_j^2\right]+2(a_j-1)\right)<0\ ,\ j=1(1)n \qquad (10.135)$$

ή

$$0<\omega<\frac{2(1-a_j)}{(1-a_j)^2+b_j^2}\ ,\ j=1(1)n$$

ή

$$0<\omega<\min_j\left\{\frac{2(1-a_j)}{(1-a_j)^2+b_j^2}\right\}\ .$$

Αν υποθέσουμε ότι $a_j>1$, $j=1(1)n$, τότε η (10.134) ή η (10.135) ισχύει εάν και μόνον εάν

$$\max_j\left\{\frac{2(1-a_j)}{(1-a_j)^2+b_j^2}\right\}<\omega<0\ . \qquad \blacktriangle$$

<u>Θεώρημα 10.11</u> Εάν οι ιδιοτιμές λ_i , $i=1(1)n$ του επαναληπτικού πίνακα G της μεθόδου (10.129) είναι πραγματικές και $\lambda_m \leqq \lambda_i \leqq \lambda_M$, $i=1(1)n$, τότε η extrapolation μέθοδος (10.130) συγκλίνει εάν και μόνον εάν μία από τις παρακάτω προτάσεις ισχύει:

i) $\lambda_M<1$ και $0<\omega<\dfrac{2}{1-\lambda_m}$

ii) $\lambda_m>1$ και $\dfrac{2}{1-\lambda_M}<\omega<0$.

Επί πλέον, η άριστη τιμή ω_{opt} της extrapolation παραμέτρου ω , για την οποία ισχύει $\rho(G_{\omega_{opt}})=\min_\omega \rho(G_\omega)$, δίνεται από την

$$\omega_{opt}=\frac{2}{2-\lambda_m-\lambda_M}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

και για τις δύο περιπτώσεις (i) και (ii) του θεωρήματος και ακόμη είναι

$$\rho(G_{\omega_{opt}}) = \frac{\lambda_M - \lambda_m}{|2 - \lambda_m - \lambda_M|} \quad .$$

<u>Απόδειξη</u> Το πρώτο μέρος του θεωρήματος, που αφορά τις προτάσεις (i) και (ii) μπορεί να προκύψει εύκολα σαν πόρισμα του θεωρήματος 10.10. Θα προσδιορίσουμε τώρα το $\min_{\omega} \rho(G_\omega)$, με την προυπόθεση ότι ισχύει η πρόταση (i) . Ανάλογη είναι η εργασία όταν ισχύει η πρόταση (ii) . Έτσι, έχουμε

$$\min_{\omega} \rho(G_\omega) = \min_{\omega \in W} \max_{\lambda \in \Lambda} |1 - \omega + \omega\lambda| \ , \tag{10.136}$$

όπου $W = (0, \frac{2}{1 - \lambda_m})$, $\Lambda = [\lambda_m , \lambda_M]$, $\lambda_M < 1$. Επειδή $\frac{\partial}{\partial\lambda}(1 - \omega + \omega\lambda) = \omega$ (ανεξάρτητο του λ) , θα ισχύει (σημειώνεται ότι $\lambda_m + \lambda_M < 2$)

$$\max_{\lambda \in \Lambda} |1 - \omega + \omega\lambda| = \max \left\{ |1 - \omega + \omega\lambda_m| \ , \ |1 - \omega + \omega\lambda_M| \right\} =$$

$$= \begin{cases} |1 - \omega + \omega\lambda_M| \ , & \text{εάν} \quad 0 < \omega \leqq \omega^* \\ |1 - \omega + \omega\lambda_m| \ , & \text{εάν} \quad \omega^* \leqq \omega < \frac{2}{1 - \lambda_m} \ , \end{cases} \tag{10.137}$$

όπου

$$\omega^* = \frac{2}{2 - \lambda_m - \lambda_M} \quad . \tag{10.138}$$

Είναι εύκολο να δειχθεί ότι $1 - \omega + \omega\lambda_M \geqq 0$ εάν $0 < \omega \leqq \omega^*$, ενώ $1 - \omega + \omega\lambda_m \leqq 0$ εάν $\omega^* \leqq \omega < \frac{2}{1 - \lambda_m}$, οπότε η (10.137) γίνεται

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\max_{\lambda\in\Lambda} |1-\omega+\omega\lambda| = \begin{cases} 1-\omega+\omega\lambda_M \ , & \text{εάν } 0<\omega\leqq\omega^* \\ \omega-1-\omega\lambda_m \ , & \text{εάν } \omega^*\leqq\omega<\dfrac{2}{1-\lambda_m} \ . \end{cases} \tag{10.139}$$

Επειδή $\frac{d}{d\omega}(1-\omega+\omega\lambda_M) = \lambda_M - 1 < 0$ και $\frac{d}{d\omega}(\omega-1-\omega\lambda_m) = 1 - \lambda_m > 0$, προκύπτει ότι

$$\min_{0<\omega\leqq\omega^*} (1-\omega+\omega\lambda_M) = 1-\omega^*+\omega^*\lambda_M \quad \text{και}$$

$$\min_{\omega^*\leqq\omega<\frac{2}{1-\lambda_m}} (\omega-1-\omega\lambda_m) = \omega^*-1-\omega^*\lambda_m \ .$$

Αλλά, είναι $1-\omega^*+\omega^*\lambda_M = \omega^*-1-\omega^*\lambda_m = \dfrac{\lambda_M-\lambda_m}{2-\lambda_m-\lambda_M}$. Από τα παραπάνω και λόγω της (10.139) η (10.136) τελικά γίνεται

$$\min_{\omega} \rho(G_\omega) = \rho(G_{\omega_{opt}}) = \frac{\lambda_M-\lambda_m}{|2-\lambda_m-\lambda_M|} \ , \text{ όπου } \omega_{opt} = \omega^* \ . \qquad \blacktriangle$$

<u>Θεώρημα 10.12</u> Εάν η μέθοδος (10.129) συγκλίνει, τότε η extrapolation αυτής (10.130) συγκλίνει για κάθε $0<\omega<\dfrac{2}{1+\rho(G)}$.

<u>Απόδειξη</u> Αφού η (1.129) συγκλίνει είναι $\rho(G)<1$. Θα δειχθεί ότι για κάθε $0<\omega<\dfrac{2}{1+\rho(G)}$ είναι $\rho(G_\omega)<1$, όπου $G_\omega = (1-\omega)I+\omega G$. Αν λ_j , $j=1(1)n$ είναι οι ιδιοτιμές του G , τότε,λόγω της (10.132),είναι

$$\rho(G_\omega) = \max_j|1-\omega+\omega\lambda_j| \leqq \max_j\left\{|1-\omega|+|\omega||\lambda_j|\right\} =$$

$$= |1-\omega|+|\omega|\max_j|\lambda_j| = |1-\omega|+|\omega|\rho(G) \ . \tag{10.140}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Επειδή $\dfrac{2}{1+\rho(G)} > 1$, διακρίνουμε τις περιπτώσεις:

i) $0 < \omega \leqq 1$. Τότε, από την (10.140) έχουμε $\rho(G_\omega) \leqq$
$\leqq 1 - \omega + \omega\rho(G) < 1 - \omega + \omega = 1$.

ii) $1 < \omega < \dfrac{2}{1+\rho(G)}$. Τότε, η (1.140) δίνει $\rho(G_\omega) \leqq$
$\leqq \omega - 1 + \omega\rho(G) = \omega(1+\rho(G)) - 1 < 1$. ▲

Έχοντας υπ' όψιν τον τρόπο κατασκευής της extrapolation μεθόδου μιας βασικής επαναληπτικής μεθόδου, είναι εύκολο να σχηματίσουμε τις αντίστοιχες extrapolation μεθόδους (με extrapolation παράμετρο ω) των βασικών μεθόδων, που μελετήθηκαν στις προηγούμενες παραγράφους. Έτσι, η extrapolation της μεθόδου Jacobi, που καλείται επίσης και JOR (Jacobi OverRelaxation) είναι:

$$x^{(m+1)} = J_\omega x^{(m)} + \omega D^{-1} b \ , \quad m = 0,1,2,\ldots, \tag{10.141}$$

όπου $J_\omega = I - \omega D^{-1} A = (1-\omega)I + \omega J$.

Η extrapolation της μεθόδου Gauss – Seidel είναι:

$$x^{(m+1)} = S_\omega x^{(m)} + \omega(D - C_L)^{-1} b \ , \quad m = 0,1,2,\ldots, \tag{10.142}$$

όπου $S_\omega = (1-\omega)I + \omega S = I - \omega(I-L)^{-1}D^{-1}A$.

Η extrapolation της μεθόδου του Richardson είναι

$$x^{(m+1)} = (I - \omega A)x^{(m)} + \omega b \ , \quad m = 0,1,2,\ldots \ . \tag{10.143}$$

Τέλος, η extrapolation της μεθόδου (10.127) δίνεται από την

$$x^{(m+1)} = R_{r,\omega} x^{(m)} + \omega(D - rC_L)^{-1} b \ , \quad m = 0,1,2,\ldots \ , \tag{10.144}$$

όπου $R_{r,\omega} = (1-\omega)I + \omega R(r) = I - \omega(I - rL)^{-1}D^{-1}A$.

Η μέθοδος (10.144) είναι γνωστή με το όνομα AOR (Accelerated OverRelaxation) μέθοδος και περιέχει τις δύο παραμέτρους r και ω , οι ο-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ποίες πρέπει να εκλέγονται έτσι ώστε αυτή να συγκλίνει όσο γίνεται πιο γρήγορα. Είναι φανερό ότι για $r = 0$ η AOR συμπίπτει με την extrapolation της μεθόδου Jacobi, δηλαδή την (10.141), ενώ για $r = 1$ συμπίπτει με την extrapolation της μεθόδου Gauss – Seidel (βλέπε (10.142)). Ακόμη, είναι εύκολο να δειχθεί ότι για $r = \omega$ αυτή συμπίπτει με την μέθοδο SOR . Με σκοπό την εύρεση του επαναληπτικού αλγορίθμου της AOR αναλυτικά με τις συνιστώσες, η (10.144) μπορεί να γραφεί ως εξής:

$$(D - rC_L)x^{(m+1)} = \left[(1-\omega)D + (\omega - r)C_L + \omega U\right]x^{(m)} + \omega b ,$$

από την οποία προκύπτει ότι

$$a_{ii}x_i^{(m+1)} + r\sum_{j=1}^{i-1} a_{ij}x_j^{(m+1)} = (1-\omega)a_{ii}x_i^{(m)} - (\omega - r)\sum_{j=1}^{i-1} a_{ij}x_j^{(m)} -$$

$$- \omega\sum_{j=i+1}^{n} a_{ij}x_j^{(m)} + \omega b_i , \quad i = 1(1)n ,$$ ή, λύνοντας ως προς $x_i^{(m+1)}$,

$$x_i^{(m+1)} = (1-\omega)x_i^{(m)} + \frac{1}{a_{ii}}\left[\omega\left\{b_i - \sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}x_j^{(m)}\right\} + \right.$$

$$\left. + r\sum_{j=1}^{i-1} a_{ij}(x_j^{(m)} - x_j^{(m+1)})\right] , \quad i = 1(1)n , \quad m = 0,1,2,\ldots . \quad (10.145)$$

ΑΣΚΗΣΕΙΣ

1. Να δειχθεί ότι η επίλυση του μιγαδικού συστήματος $(i = \sqrt{-1})$

$$\begin{cases} (1+i)z_1 - 2iz_2 = -5 + 3i \\ (3-2i)z_1 + (2+i)z_2 = 18 - 5i , \end{cases}$$

μπορεί να αναχθεί στην επίλυση ενός γραμμικού συστήματος $Ax = b$, όπου $A \in \mathbb{R}^{4,4}$ και $b \in \mathbb{R}^4$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

2. Να λυθεί το σύστημα

$$\begin{cases} 3x_1 - 12x_2 + 8x_3 = 7 \\ 4x_1 + 7x_2 - 2x_3 = 4 \\ 7x_1 + 9x_2 + 5x_3 = 11 \end{cases},$$

με την μέθοδο απαλοιφής του Gauss . Να χρησιμοποιηθούν 3 δ.θ. στους υπολογισμούς.

3. Να λυθεί το σύστημα

$$\begin{cases} 10^{-5}x_1 + x_2 = 0.6 \\ x_1 + x_2 = 1 \end{cases},$$

i) με την μέθοδο απαλοιφής του Gauss και ii) με την μέθοδο απαλοιφής του Gauss με μερική οδήγηση. Να χρησιμοποιηθούν 4 σ.ψ. στους υπολογισμούς.

4. Να λυθεί το σύστημα

$$\begin{cases} 0.7x_1 + 46x_2 - 16x_3 = 13.3 \\ 19x_1 - 13x_2 + 23x_3 = 14 \\ 31x_1 + 26x_2 + 11x_3 = 17 \end{cases},$$

i) με την μέθοδο απαλοιφής του Gauss και ii) με την μέθοδο απαλοιφής του Gauss με μερική οδήγηση, χρησιμοποιώντας 3 σ.ψ. στους υπολογισμούς. (Ακριβής λύση: $x_1 = -1$, $x_2 = 1$, $x_3 = 2$) .

5. Να λυθεί το σύστημα

$$\begin{cases} x_1 + \frac{1}{2}x_2 + \frac{1}{3}x_3 = 8 \\ \frac{1}{2}x_1 + \frac{1}{3}x_2 + \frac{1}{4}x_3 = 8 \\ \frac{1}{3}x_1 + \frac{1}{4}x_2 + \frac{1}{5}x_3 = 7 \end{cases}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\begin{cases} 6x_1 + 13x_2 - 17x_3 = 1 \\ 13x_1 + 29x_2 - 38x_3 = 2 \\ -17x_1 - 38x_2 + 50x_3 = -3 \end{cases}, \quad \begin{cases} 20x + 99y = 218 \\ 101x + 500y = 1101 \end{cases}.$$

16. Αν στην (10.63) είναι $\delta b = 0$, να δειχθεί ότι

$$\frac{\|\delta\hat{x}\|}{\|\bar{x}\|} \leqq \kappa \frac{\|\delta A\|}{\|A\|}, \text{ όπου } \kappa = \|A^{-1}\|\|A\| .$$

17. Δίνεται το σύστημα $Ax = b$ και $\bar{x} \in \mathbb{C}^n$ με $r = b - A\bar{x}$. Εάν $B \in \mathbb{C}^{n,n}$ τέτοιος ώστε $\|I - AB\| < 1$, να δειχθεί ότι ο A είναι ομαλός και

$$\|A^{-1}b - \bar{x}\| \leqq \frac{\|r\|\|B\|}{1 - \|I - AB\|} .$$

18. Δίνεται το σύστημα $Ax = b$, $\det(A) \neq 0$, $\hat{x} = A^{-1}b$. Αν ο πίνακας B είναι τέτοιος ώστε $\|A - B\| \leqq \|A^{-1}\|^{-1}$, να δειχθεί ότι το σύστημα $Bx = b$ έχει μία μοναδική λύση $\hat{y}$ που ικανοποιεί την

$$\|\hat{x} - \hat{y}\| \leqq \frac{\kappa(A)\left(\|A - B\|\|\hat{x}\| + \|b - d\|\right)}{\|A\| - \kappa(A)\|A - B\|}, \text{ όπου } \kappa(A) = \|A\|\|A^{-1}\|.$$

19. Έστω A ένας ομαλός πίνακας και X_0 ένας πίνακας τέτοιος ώστε $\|I - AX_0\| < 1$. Αν $X_\kappa = X_{\kappa-1}(2I - AX_{\kappa-1})$, $\kappa = 1,2,\ldots$, να δειχθεί ότι $X_\kappa \underset{\kappa\to\infty}{\to} A^{-1}$.

20. Έχοντας υπ' όψιν τις (10.79), (10.85) και θέτοντας $r^{(m)} = Ax^{(m)} - b$ $m = 0,1,2,\ldots$, να δειχθεί ότι $x^{(m)} - x^{(m+1)} = (I - G)A^{-1}r^{(m)}$, $m = 0,1,2,\ldots$. Αν $\rho(G) < 1$, να δειχθεί ότι $\varepsilon^{(m)} = (I - G)^{-1}G^m(x^{(0)} - x^{(1)})$, όπου $\varepsilon^{(m)} = x^{(m)} - \hat{x}$, $\hat{x} = A^{-1}b$.

21. Δίνεται το σύστημα $Ax = b$, όπου A είναι πραγματικός συμμετρι-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

με την μέθοδο απαλοιφής του Gauss χρησιμοποιώντας i) κλάσματα στους υπολογισμούς ii) 2 σ.ψ. και iii) 3 σ.ψ. στους υπολογισμούς.

6. Να βρεθούν οι λύσεις των συστημάτων:

$$Ax = b^{(1)}, \quad Ax = b^{(2)}, \text{ όπου } A = \begin{bmatrix} 6 & 4 & 3 \\ 20 & 15 & 12 \\ 15 & 12 & 10 \end{bmatrix}$$

$b^{(1)} = (13.1,\ 46.9,\ 37.1)^T$, $b^{(2)} = (13,\ 47,\ 37)^T$.

7. Για τον πίνακα A του συστήματος του παραδείγματος 10.1 να βρεθούν οι πίνακες P, L, R έτσι ώστε $PA = LR$ (ο P είναι πίνακας μεταθέσεως και οι L, R κάτω και άνω αντίστοιχα τριγωνικοί πίνακες).

8. Για ποιές τιμές του κ ο πίνακας $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & \kappa \end{bmatrix}$ είναι μη ομαλός;

9. Να υπολογισθεί η $\det(A)$ με την μέθοδο απαλοιφής του Gauss με μερική οδήγηση, όπου

$$A = \begin{bmatrix} 3 & 2 & -1 & 2 \\ -6 & -2 & 4 & -3 \\ -3 & -6 & -3 & -1 \\ 3 & 4 & 1 & 1 \end{bmatrix} .$$

10. Να βρεθούν οι αντίστροφοι των πινάκων

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$A = \begin{bmatrix} 5 & 4 & 3 & 2 & 1 \\ 4 & 4 & 3 & 2 & 1 \\ 3 & 3 & 3 & 2 & 1 \\ 2 & 2 & 2 & 2 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 4 & 5 \\ 4 & 5 & 5 & 7 \end{bmatrix}, \quad C = \begin{bmatrix} 6 & 4 & 3 \\ 20 & 15 & 12 \\ 15 & 12 & 10 \end{bmatrix}.$$

11. Να λυθεί το σύστημα του παραδείγματος 10.1 με την μέθοδο απαλοιφής του Jordan με μερική οδήγηση.

12. Να γίνει η τριγωνική ανάλυση του θετικά ορισμένου πίνακα

$$\begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}.$$

13. Δίνεται το σύστημα $\begin{bmatrix} 1 & 2 \\ 2 & 4.0001 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1.9999 \end{bmatrix}$. Αν $\hat{x}$ είναι η ακριβής λύση του και $\bar{x} = (1,0)^T$ μία προσέγγιση αυτής, να βρεθεί το υπόλοιπο $r = A\bar{x} - b$ και φράγματα για το απόλυτο σφάλμα $\|\varepsilon\|_\infty$ και το απόλυτο σχετικό σφάλμα $\dfrac{\|\varepsilon\|_\infty}{\|\hat{x}\|_\infty}$ της $\bar{x}$.

14. Για ποιές τιμές του λ ο αριθμός καταστάσεως $\kappa_\infty(A)$ όπου

$$A = \begin{bmatrix} 2\lambda & \lambda \\ 1 & 1 \end{bmatrix},$$ γίνεται ελάχιστος;

15. Να εξετασθούν ως προς την ευστάθεια τα ακόλουθα συστήματα:

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

κός και θετικά ορισμένος $(A = D - C_L - C_L^T)$. Εάν ο πίνακας $2D - A$ είναι θετικά ορισμένος, να δειχθεί ότι η μέθοδος Jacobi συγκλίνει.

22. Αν στο σύστημα $Ax = b$ ο A είναι αυστηρά διαγώνια υπέρτερος ως προς τις στήλες του, δηλαδή ισχύει

$$|a_{jj}| > \sum_{\substack{i=1 \\ i \neq j}}^{n} |a_{ij}| \ , \ \forall\, j = 1(1)n \ ,$$

να δειχθεί ότι η μέθοδος Jacobi συγκλίνει.

23. Με ποιά σειρά πρέπει να γραφούν οι εξισώσεις του συστήματος

$$\begin{cases} 3x - 8y + 2z = -4 \\ -4x + y - 6z = 3 \\ 7x - y + 4z = 8 \ , \end{cases}$$

για να είμαστε βέβαιοι ότι η μέθοδος Jacobi θα συγκλίνει;

24. Εάν $A^H = A$, $A = M - N$, $\det(M) \neq 0$ και οι A , $M + M^H - A$ είναι θετικά ορισμένοι, να δειχθεί ότι $\rho(M^{-1}N) < 1$.

25. Να δειχθεί ότι η χαρακτηριστική εξίσωση $\det(J - \lambda I) = 0$ του επαναληπτικού πίνακα J της μεθόδου Jacobi για το σύστημα $Ax = b$, όπου $A = (a_{ij})$, $i,j = 1(1)n$ είναι ισοδύναμη με την

$$\begin{vmatrix} a_{11}\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}\lambda & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}\lambda \end{vmatrix} = 0 \ .$$

26. Αν $A = \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix}$, $A \in \mathbb{R}^{2,2}$, $\alpha > 0$, $\det(A) > 0$, $b \in \mathbb{R}^2$, να

δειχθεί ότι η μέθοδος Jacobi για το σύστημα $Ax = b$ συγκλίνει.

27. Αν $G = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$, $G \in \mathbb{R}^{2,2}$, να δειχθεί ότι $\rho(G) < 1$ εάν και μόνον εάν $|\alpha + \delta| < 1 + \det(G)$ και $|\det(G)| < 1$.

28. Δίνεται ο αλγόριθμος

$$\begin{cases} x_1^{(m+1)} = -\frac{1}{2} x_2^{(m)} \\ x_2^{(m+1)} = -\frac{1}{2} x_1^{(m)} - \frac{1}{2} x_3^{(m)} \\ x_3^{(m+1)} = -\frac{1}{2} x_2^{(m)} \end{cases}, \quad m = 0,1,2,\ldots \quad \text{και}$$

$x_i^{(0)}$, $i = 1(1)3$ αυθαίρετοι πραγματικοί αριθμοί. Να εξεταστεί ως προς την σύγκλιση.

29. Να βρεθούν οι τρείς πρώτες επαναλήψεις της μεθόδου Jacobi για το σύστημα

$$\begin{cases} 4x_1 - x_2 = 1 \\ -x_1 + 4x_2 - x_3 = 0 \\ -x_2 + 4x_3 = 1 \end{cases}, \quad \text{αν} \quad x^{(0)} = (0, 0, 0)^T .$$

30. Να δειχθεί ότι η χαρακτηριστική εξίσωση $\det(S - \lambda I) = 0$ του επαναληπτικού πίνακα S της μεθόδου Gauss – Seidel για το σύστημα $Ax = b$, όπου $A = (a_{ij})$, $i,j = 1(1)n$ είναι ισοδύναμη με την

$$\begin{vmatrix} a_{11}\lambda & a_{12} & \ldots & a_{1n} \\ a_{21}\lambda & a_{22}\lambda & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1}\lambda & a_{n2}\lambda & \ldots & a_{nn}\lambda \end{vmatrix} = 0 .$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

31. Να βρεθούν όλες οι τιμές του α , έτσι ώστε για τον πίνακα

$$A = \begin{bmatrix} 4 & -\alpha \\ -\alpha & 1 \end{bmatrix}$$, η μέθοδος Gauss – Seidel να συγκλίνει.

32. Αν $A = \begin{bmatrix} 1 & \alpha & \alpha \\ \alpha & 1 & \alpha \\ \alpha & \alpha & 1 \end{bmatrix}$, να δειχθεί ότι η μέθοδος Gauss – Seidel συγκλίνει $\forall\, 0 < \alpha < 1$. Για ποιές τιμές του α συγκλίνει η μέθοδος Jacobi;

33. Δίνεται το σύστημα $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$, όπου $a_{11}a_{22} \neq 0$ και $a_{11}a_{22} - a_{12}a_{21} \neq 0$. Να δειχθεί ότι η μέθοδος Gauss – Seidel συγκλίνει εάν και μόνον εάν $\left|\dfrac{a_{12}a_{21}}{a_{11}a_{22}}\right| < 1$. Να δειχθεί επίσης ότι $\rho(S) = \rho^2(J)$, όπου S, J οι επαναληπτικοί πίνακες των μεθόδων Gauss – Seidel και Jacobi αντίστοιχα.

34. Δίνεται ο πίνακας $A = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}$. Να δειχθεί ότι $\rho(S) = \rho^2(J)$.

35. Να εξετασθούν ως προς την σύγκλιση οι μέθοδοι Jacobi και Gauss – – Seidel για τα συστήματα

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\begin{cases} 2x_1 - x_2 + x_3 = 1 \\ x_1 + x_2 + x_3 = 6 \\ x_1 + x_2 - 2x_3 = 0 \end{cases} \qquad \begin{bmatrix} 1 & \frac{1}{2} & -\frac{1}{2} \\ -1 & 1 & -1 \\ \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & -2 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} .$$

36. Δίνεται το σύστημα

$$\begin{cases} 2x_1 + 2x_2 + x_3 = 1 \\ x_1 + x_2 + x_3 = 1 \\ x_1 + 2x_2 - 2x_3 = 1 \end{cases} .$$

Να δειχθεί ότι η μέθοδος Gauss – Seidel συγκλίνει και να βρεθούν οι τρείς πρώτες επαναλήψεις αυτής.

37. Να δειχθεί ότι η χαρακτηριστική εξίσωση του επαναληπτικού πίνακα $S(\omega)$ της SOR μεθόδου είναι $\det(S(\omega) - \lambda I) = 0$ ή ισοδύναμα $\det(\lambda L + U - \frac{\lambda + \omega - 1}{\omega} I) = 0$. Να δειχθεί ότι η $\lambda = 0$ είναι μία ιδιοτιμή του $S(\omega)$ εάν και μόνον εάν $\omega = 1$.

38. Αν $A = \begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix}$, για ποιές τιμές του ω συγκλίνει η SOR μέθοδος; Ποιό είναι το ω_{opt};

39. Θεωρούμε το σύστημα $Ax = b$, όπου $A = \begin{bmatrix} \frac{1}{2} & \alpha \\ -\alpha & \frac{1}{2} \end{bmatrix}$, $\alpha \in \mathbb{R}$. Για

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ποιές τιμές του α συγκλίνει η μέθοδος Richardson;

40. Αν $A = \begin{bmatrix} 1 & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & 1 & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & 1 \end{bmatrix}$, να βρεθεί για ποιά ω συγκλίνει η JOR μέθοδος και τα ω_{opt}, $\rho(G_{\omega_{opt}})$.

41. Δίνεται η μέθοδος $x^{(m+1)} = Gx^{(m)} + \kappa$, $m = 0,1,2,\ldots$ για την επίλυση του $Ax = b$. Εάν για τις ιδιοτιμές λ_i, $i = 1(1)n$ του G ισχύει $|\lambda_i| = 1$, $\lambda_i \neq 1$, $i = 1(1)n$, να δειχθεί ότι $\rho(G_\omega) < 1$ εάν και μόνον εάν $0 < \omega < 1$, όπου $G_\omega = (1-\omega)I + \omega G$. Επί πλέον, να δειχθεί ότι $\omega_{opt} = \frac{1}{2}$.

42. Να λυθούν με την μέθοδο Gauss – Seidel τα συστήματα

$$\begin{cases} 2x_1 + x_2 + x_3 - x_4 = -3 \\ x_1 + 9x_2 + 8x_3 + 4x_4 = 15 \\ -x_1 + 3x_2 + 5x_3 + 2x_4 = 10 \\ x_2 + x_4 = 2 \end{cases}$$

$$\begin{cases} 10x_1 + 2x_2 + 3x_3 - x_4 = 80 \\ x_1 - 20x_2 - x_3 + 3x_4 = 40 \\ x_1 + x_2 - 10x_3 + 2x_4 = 40 \\ 2x_1 - x_2 - x_3 + 30x_4 = 120 \end{cases}.$$

43. Να βρεθεί για ποιές τιμές του ω συγκλίνει η extrapolation μέθοδος της μεθόδου του Richardson, καθώς και η "άριστη" τιμή ω_{opt}, για το σύστημα

$$\begin{bmatrix} 2-i & 0 & 0 \\ -1 & 2+i & 0 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

11

ΥΠΟΛΟΓΙΣΜΟΣ ΙΔΙΟΤΙΜΩΝ ΚΑΙ ΙΔΙΟΔΙΑΝΥΣΜΑΤΩΝ

11.1 Εισαγωγή

Έχοντας υπ' όψιν την βασική θεωρία για τις ιδιοτιμές και τα ιδιοδιανύσματα ενός πίνακα $A \in \mathbb{C}^{n,n}$, που αναφέρθηκε στην παράγραφο 9.5, στο παρόν κεφάλαιο θα εξετάσουμε ορισμένες αριθμητικές μεθόδους για τον υπολογισμό των ιδιοτιμών και ιδιοδιανυσμάτων του A.

Είναι γνωστό ότι οι ιδιοτιμές ενός πίνακα $A \in \mathbb{C}^{n,n}$ είναι οι n ρίζες της χαρακτηριστικής εξισώσεως του A

$$P(\lambda) \equiv \det(A - \lambda I) =$$

$$= (-1)^n(\lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \ldots + a_{n-1}\lambda + a_n) = 0. \quad (11.1)$$

Επί πλέον, εάν οι ιδιοτιμές είναι γνωστές, τότε αντίστοιχα ιδιοδιανύσματα μπορούν να βρεθούν σαν μη μηδενικές λύσεις συγκεκριμένων ομογενών γραμμικών συστημάτων (βλέπε θεωρήματα 9.20, 9.29). Είναι φανερό ότι, αν το n είναι μεγάλο, τότε η διαδικασία υπολογισμού των συντελεστών του χαρακτηριστικού πολυωνύμου $P(\lambda)$, αναπτύσσοντας την $\det(A - \lambda I)$, δεν είναι πρακτικά αποτελεσματική. Επί πλέον, και αν ακόμη είναι εύκολο να βρεθεί (προσεγγισθεί) το πολυώνυμο $P(\lambda)$, ο προσ-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

διορισμός των ριζών του είναι ένα δύσκολο πρόβλημα που γίνεται δυσκολώτερο όταν αυτό είναι ασταθές. Για τους παραπάνω λόγους, στην πράξη, σχεδόν ποτέ δεν χρησιμοποιούνται μέθοδοι υπολογισμού του χαρακτηριστικού πολυωνύμου με σκοπό τον υπολογισμό των ιδιοτιμών. Σημειώνουμε επ' ευκαιρία, έχοντας υπ'όψιν το θεώρημα 9.37, ότι ο προσδιορισμός των ριζών ενός πολυωνύμου (βλέπε κεφάλαιο 7) μπορεί να αναχθεί στον προσδιορισμότων ιδιοτιμών του συνοδεύοντα πίνακα του πολυωνύμου.

Το πρόβλημα του υπολογισμού των ιδιοτιμών και ιδιοδιανυσμάτων ένός πίνακα A είναι πιο απλό στην περίπτωση που ο A είναι Ερμιτιανός, λόγω των ιδιοτήτων που ισχύουν για ένα τέτοιο πίνακα (βλέπε θεωρήματα 9.66, 9.67). Μπορεί να δειχθεί επίσης ότι ένας Ερμιτιανός πίνακας είναι ευσταθής σε σχέση με το πρόβλημα ευρέσεως των ιδιοτιμών του, δηλαδή, μικρές μεταβολές στα στοιχεία του, προκαλούν μικρές μεταβολές στις ιδιοτιμές του (βλέπε παράγραφο 11.9).

Για τον υπολογισμό φραγμάτων για τις ιδιοτιμές,που απαιτούνται σε ορισμένες εφαρμογές,χρήσιμα είναι τα θεωρήματα 9.38, 9.39, 9.40 και 9.74 .

11.2 Η μέθοδος της δυνάμεως

Η απλούστερη και πιο διαδεδομένη ίσως επαναληπτική μέθοδος για τον υπολογισμό ιδιοτιμών και ιδιοδιανυσμάτων ενός πίνακα $A \in \mathbb{C}^{n,n}$ είναι η μέθοδος της δυνάμεως . Αυτή μπορεί να θεωρηθεί σαν πρόδρομος άλλων επαναληπτικών μεθόδων με μεγαλύτερη πρακτική αξία και είναι ειδικά κατάλληλη για μεγάλους αραιούς πίνακες. Για την εφαρμογή αυτής της μεθόδου ο πίνακας A δεν θεωρείται τελείως αυθαίρετος, αλλά υποθέτουμε ότι είναι διαγωνίσιμος, δηλαδή, έχει n γραμμικώς ανεξάρτητα ιδιοδιανύσματα. Υποθέτουμε επίσης ότι για τις ιδιοτιμές λ_i , $i = 1(1)n$ του A ισχύει

$$|\lambda_1| \geqq |\lambda_2| \geqq \dots \geqq |\lambda_n| \tag{11.2}$$

και ότι δεν υπάρχει ιδιοτιμή $\lambda_j \neq \lambda_1$ με $|\lambda_j| = |\lambda_1|$. Δηλαδή, υπάρ-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

χει ένας ακέραιος $m > 0$ τέτοιος ώστε

$$\lambda_1 = \lambda_2 = \ldots = \lambda_m \ , \qquad (11.3)$$
$$|\lambda_1| = \ldots = |\lambda_m| > |\lambda_{m+1}| \geqq \ldots \geqq |\lambda_n| \ .$$

Έστωσαν $x^{(i)}$, $i = 1(1)n$ τα n γραμμικώς ανεξάρτητα ιδιοδιανύσματα του A , οπότε

$$Ax^{(i)} = \lambda_i x^{(i)} \ , \ i = 1(1)n \ . \qquad (11.4)$$

Τα $x^{(i)}$, $i = 1(1)n$ αποτελούν μία βάση του $\mathbb{C}^n$ και επομένως αν $y^{(0)} \in \mathbb{C}^n$ είναι ένα αυθαίρετο αρχικό διάνυσμα, τότε μπορούμε να γράψουμε

$$y^{(0)} = a_1 x^{(1)} + a_2 x^{(2)} + \ldots + a_n x^{(n)} \ , \qquad (11.5)$$

όπου, γενικά, $a_i \in \mathbb{C}$, $i = 1(1)n$. Σχηματίζουμε τώρα την ακολουθία διανυσμάτων $\{y^{(\kappa)}\}$ από την αναδρομική σχέση (υποθέτοντας ότι $y^{(0)} \neq 0$)

$$y^{(\kappa+1)} = Ay^{(\kappa)} \ , \quad \kappa = 0,1,2,\ldots \ . \qquad (11.6)$$

Χρησιμοποιώντας επαναληπτικά την (11.6) παίρνουμε

$$y^{(\kappa)} = A^{\kappa} y^{(0)} \ , \quad \kappa = 0,1,2,\ldots \ . \qquad (11.7)$$

Η (11.7), λόγω των (11.5) και (11.4), γίνεται διαδοχικά

$$y^{(\kappa)} = A^{\kappa}\left(\sum_{i=1}^{n} a_i x^{(i)}\right) = \sum_{i=1}^{n} a_i A^{\kappa} x^{(i)} = \sum_{i=1}^{n} a_i \lambda_i^{\kappa} x^{(i)} =$$

$$= a_1 \lambda_1^{\kappa} x^{(1)} + a_2 \lambda_2^{\kappa} x^{(2)} + \ldots + a_n \lambda_n^{\kappa} x^{(n)} \ . \qquad (11.8)$$

Υποθέτοντας τώρα ότι το $y^{(0)}$ στην (11.5) ικανοποιεί την

$$a_1 x^{(1)} + a_2 x^{(2)} + \ldots + a_m x^{(m)} \neq 0 \qquad (11.9)$$

(η (11.9) εξασφαλίζεται με κατάλληλη αυθαίρετη εκλογή του $y^{(0)}$) και έχοντας υπ' όψιν, λόγω της (11.3), ότι

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\left|\frac{\lambda_j}{\lambda_1}\right| < 1 \ , \ j \geqq m+1 \ ,$$

από την (11.8) προκύπτει ότι

$$y^{(\kappa)} = \lambda_1^{\kappa}\left[a_1x^{(1)} + a_2x^{(2)}+\ldots+ a_mx^{(m)} + a_{m+1}\left(\frac{\lambda_{m+1}}{\lambda_1}\right)^{\kappa} x^{(m+1)}+\ldots+\right.$$

$$\left. + a_n\left(\frac{\lambda_n}{\lambda_1}\right)^{\kappa} x^{(n)}\right] \qquad (11.10)$$

και

$$\lim_{\kappa\to\infty}\frac{y^{(\kappa)}}{\lambda_1^{\kappa}} = a_1x^{(1)} + a_2x^{(2)} +\ldots+ a_mx^{(m)} \ . \qquad (11.11)$$

Χρησιμοποιώντας την έκφραση για το $y^{(\kappa+1)}$, όπως προκύπτει από την (11.10), και υποθέτοντας γενικά ότι $y^{(\kappa)} = (y_j^{(\kappa)})$, $j = 1(1)n$, παίρνουμε

$$\lim_{\kappa\to\infty}\frac{y_j^{(\kappa+1)}}{y_j^{(\kappa)}} = \lambda_1 \quad , \ j = 1(1)n \ , \qquad (11.12)$$

με την προυπόθεση ότι

$$a_1x_j^{(1)} + a_2x_j^{(2)} +\ldots+ a_mx_j^{(m)} \neq 0 \ . \qquad (11.13)$$

Λόγω της (11.9), θα υπάρχει τουλάχιστον ένα j $(j = 1(1)n)$, για το οποίο ισχύει η (11.13). Έτσι, με τις παραπάνω προυποθέσεις η μέθοδος της δυνάμεως, που βασίζεται στην (11.6), δια μέσου της (11.12) μας παρέχει την ιδιοτιμή λ_1 η οποία έχει το μεγαλύτερο μέτρο και δια μέσου της (11.11) ένα ιδιοδιάνυσμα, το $a_1x^{(1)} + a_2x^{(2)} +\ldots+ a_mx^{(m)}$, που αντιστοιχεί στην λ_1 . Η ταχύτητα συγκλίσεως της μεθόδου της δυνάμεως εξαρτάται προφανώς από τον λόγο $\left|\frac{\lambda_{m+1}}{\lambda_1}\right|$ (βλέπε (11.3) και (11.10)).

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Όσο πιο μικρός είναι ο λόγος αυτός, τόσο ταχύτερη θα είναι η σύγκλιση.

Για επαρκώς μεγάλο κ από την (11.10) έχουμε

$$y^{(\kappa)} \approx \lambda_1^{\kappa}(a_1 x^{(1)} + a_2 x^{(2)} + \ldots + a_m x^{(m)}) \qquad (11.14)$$

και από την (11.12)

$$\lambda_1 \approx \frac{y_j^{(\kappa+1)}}{y_j^{(\kappa)}}, \quad j = 1(1)n .$$

Η (11.14) σημαίνει πως το $y^{(\kappa)}$ είναι κατά προσέγγιση ένα ιδιοδιάνυσμα του A, που αντιστοιχεί στην ιδιοτιμή λ_1. Για την εύρεση άλλων γραμμικά ανεξάρτητων ιδιοδιανυσμάτων, που αντιστοιχούν στην λ_1 (υπάρχουν m τέτοια σύμφωνα με τις υποθέσεις), εφαρμόζουμε την μέθοδο της δυνάμεως αλλάζοντας διαδοχικά το αρχικό διάνυσμα $y^{(0)}$. Τονίζεται ότι δεν είμαστε βέβαιοι ότι με τον τρόπο αυτό θα βρούμε όλα τα γραμμικώς ανεξάρτητα ιδιοδιανύσματα του A που αντιστοιχούν στην λ_1.

Πρέπει να σημειωθεί ότι το όριο στην (11.12) υπολογίζεται για κάθε j για το οποίο $y_j^{(\kappa)} \neq 0$. Έτσι, είναι φανερό, λόγω της (11.14), ότι για επαρκώς μεγάλο κ η $y_j^{(\kappa)} \neq 0$ συνεπάγεται την (11.13). Επίσης, από την (11.8) ή την (11.10) παρατηρούμε ότι καθώς το κ αυξάνει $(\kappa \to \infty)$ οι συνιστώσες του $y^{(\kappa)}$ αυξάνουν απεριόριστα απολύτως εάν $|\lambda_1| > 1$, ενώ ελαττούνται απεριόριστα απολύτως εάν $|\lambda_1| < 1$. Για τον λόγο αυτό με σκοπό να περιορίσουμε την επίδραση των σφαλμάτων στρογγυλεύσεως στους υπολογισμούς, η ακόλουθη διαδικασία κανονικοποιήσεως του $y^{(\kappa)}$ σε κάθε βήμα εφαρμόζεται.

Έστω ότι $y^{(\kappa)} = (y_j^{(\kappa)})$, $j = 1(1)n$, $\kappa = 0,1,2,\ldots$. θέτουμε

$$z^{(\kappa)} = \frac{y^{(\kappa)}}{y_{j_\kappa}^{(\kappa)}}, \text{ όπου } \left|y_{j_\kappa}^{(\kappa)}\right| = \max_i \left|y_i^{(\kappa)}\right| = \|y^{(\kappa)}\|_\infty , \qquad (11.15)$$

δηλαδή, διαιρούμε το διάνυσμα $y^{(\kappa)}$ με την συνιστώσα του $y_{j_\kappa}^{(\kappa)}$, που

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

έχει το μεγαλύτερο μέτρο. Έτσι, η (11.6) τροποποιείται στην

$$y^{(\kappa+1)} = Az^{(\kappa)} \ , \quad \kappa = 0,1,2,\ldots \tag{11.16}$$

και η ιδιοτιμή λ_1 και ένα ιδιοδιάνυσμα που αντιστοιχεί σ' αυτή υπολογίζονται από τις

$$\lim_{\kappa\to\infty} y_{j_\kappa}^{(\kappa+1)} = \lambda_1 \quad \text{και} \quad \lim_{\kappa\to\infty} z^{(\kappa)} = \beta(a_1x^{(1)} + a_2x^{(2)} + \ldots + a_mx^{(m)}) \ , \tag{11.17}$$

όπου β σταθερά $\neq 0$ (βλέπε άσκηση 10) . Στην πράξη, η διαδικασία συνήθως σταματά όταν π.χ. $\|z^{(\kappa+1)} - z^{(\kappa)}\|_\infty \leqq \varepsilon$, όπου ε μικρός προκαθορισμένος θετικός αριθμός.

Πρέπει να σημειωθεί ότι η παραπάνω μελέτη της μεθόδου της δυνάμεως έγινε, χάριν απλότητος, υπό την προυπόθεση της (11.3). Με κατάλληλες τροποποιήσεις μπορεί αυτή να εφαρμοσθεί και σε άλλες περιπτώσεις (βλέπε π.χ. άσκηση 11) , καθώς και σε μη διαγωνίσιμους πίνακες.

<u>Παράδειγμα 11.1</u> Να προσεγγισθεί η μεγαλύτερη απολύτως ιδιοτιμή του πίνακα

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 1 \end{bmatrix}$$

και το αντίστοιχο ιδιοδιάνυσμα αυτής.

<u>Λύση</u> Ο πίνακας A είναι πραγματικός συμμετρικός και επομένως έχει τρία γραμμικώς ανεξάρτητα ιδιοδιανύσματα. Επίσης, αφού $A > 0$, σύμφωνα με το θεώρημα 9.42, η συνθήκη (11.3) ικανοποιείται για $m = 1$, δηλαδή, αν $\lambda_1, \lambda_2, \lambda_3$ είναι οι ιδιοτιμές του A , θα ισχύει

$$|\lambda_1| > |\lambda_2| \geqq |\lambda_3| \ .$$

Εκλέγουμε το αρχικό διάνυσμα $y^{(0)} = (1, 0, 0)^T$ και εφαρμόζουμε τις (11.15), (11.16) και (11.17), οπότε παίρνουμε διαδοχικά (κριτήριο τερ-

ματισμού θα είναι η ικανοποίηση της συνθήκης $\|z^{(K+1)} - z^{(K)}\|_\infty \leqq$ $\leqq \frac{1}{2}\, 10^{-1} = 0.05$)

$z^{(0)} = (1, 0, 0)^T$, $y^{(1)} = (1, 1, 2)^T$, $\lambda_1^{(1)} = 1$,

$z^{(1)} = (0.5, 0.5, 1)^T$, $y^{(2)} = (3, 2.5, 2.5)^T$, $\lambda_1^{(2)} = 2.5$,

$z^{(2)} = (1, 0.83, 0.83)^T$, $y^{(3)} = (3.49, 3.49, 3.66)^T$, $\lambda_1^{(3)} = 3.49$,

$z^{(3)} = (0.95, 0.95, 1)^T$, $y^{(4)} = (3.9, 3.85, 3.85)^T$, $\lambda_1^{(4)} = 3.85$,

$z^{(4)} = (1, 0.99, 0.99)^T$, $y^{(5)} = (3.97, 3.97. 3.98)^T$, $\lambda_1^{(5)} = 3.97$.

Επειδή $\|z^{(4)} - z^{(3)}\|_\infty = \frac{1}{2}\, 10^{-1} = 0.05$, προκύπτει ότι $\lambda_1 \approx \lambda_1^{(5)} \approx 4.0$ και το αντίστοιχο ιδιοδιάνυσμα αυτής $x^{(1)} \approx z^{(4)} \approx (1.0, 1.0, 1.0)^T$. Μπορεί να δειχθεί ότι οι τρείς ιδιοτιμές του A είναι: $\lambda_1 = 4$, $\lambda_2 =$ $= 1$, $\lambda_3 = -1$.

11.3 Η μέθοδος της δυνάμεως στον πίνακα $A - qI$

Έστω $A \in \mathbb{C}^{n,n}$ με ιδιοτιμές λ_i , $i = 1(1)n$ και αντίστοιχα ιδιοδιανύσματα $x^{(i)}$, $i = 1(1)n$. Είναι εύκολο να δειχθεί ότι αν $q \in \mathbb{C}$, τότε ο πίνακας $A - qI$ έχει ιδιοτιμές $\lambda_i - q$, $i = 1(1)n$ και αντίστοιχα ιδιοδιανύσματα τα $x^{(i)}$, $i = 1(1)n$, δηλαδή, τα ίδια με τον A . Με άλλα λόγια, η αφαίρεση του αριθμού q από τα διαγώνια στοιχεία του A έχει σαν αποτέλεσμα την αφαίρεση του q από τις ιδιοτιμές του. Η παραπάνω τροποποίηση του πίνακα A είναι πολλές φορές χρήσιμη , με σκοπό να εφαρμοσθεί η μέθοδος της δυνάμεως. Για παράδειγμα, ας υποθέσουμε ότι ο πίνακας A είναι διαγωνίσιμος και έχει πραγματικές ιδιοτιμές λ_i , $i = 1(1)n$, τέτοιες ώστε

$$\lambda_1 \geqq \lambda_2 \geqq \dots \geqq \lambda_m > \lambda_{m+1} \geqq \dots \geqq \lambda_n \,, \tag{11.18}$$

όπου $m > 0$ και θέλουμε να υπολογίσουμε την μεγαλύτερη και μικρότερη από τις ιδιοτιμές του A , δηλαδή, τις λ_1 και λ_n και τα αντίστοιχα

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

ιδιοδιανύσματα. (Σημειώνεται ότι, λόγω της (11.18), η λ_1 μπορεί να είναι πολλαπλή καθώς και η λ_n) . Σύμφωνα με το θεώρημα 9.74 θα έχουμε

$$|\lambda_i| \leqq \|A\| \quad \text{ή}$$

$$-\|A\| \leqq \lambda_i \leqq \|A\| \ , \quad i = 1(1)n \ , \tag{11.19}$$

δηλαδή,

$$\lambda_i + \|A\| \geqq 0 \quad \text{και} \quad \lambda_i - \|A\| \leqq 0 \ , \quad i = 1(1)n \ . \tag{11.20}$$

Για την εύρεση της λ_1 εκλέγουμε $q = -\|A\|$ και σχηματίζουμε τον πίνακα $A - qI = A + \|A\|I$. Οι ιδιοτιμές του πίνακα $A + \|A\|I$ είναι $\lambda_i + \|A\| \geqq 0$, $i = 1(1)n$ και προφανώς, λόγω της (11.18), η μεγαλύτερη απολύτως ιδιοτιμή του $A + \|A\|I$ είναι η $\lambda_1 + \|A\| > 0$, πολλαπλότητος όσο και η λ_1 . Έτσι, εφαρμόζοντας την μέθοδο της δυνάμεως στον πίνακα $A + \|A\|I$, μπορούμε να βρούμε την $\lambda_1 + \|A\|$ και το αντίστοιχο ιδιοδιάνυσμα $cx^{(1)}$, οπότε αφαιρώντας την $\|A\|$ από την $\lambda_1 + \|A\|$ βρίσκουμε την λ_1.

Για την εύρεση της λ_n εκλέγουμε $q = \|A\|$, οπότε $A - qI = A - \|A\|I$. Οι ιδιοτιμές του πίνακα $A - \|A\|I$ είναι $\lambda_i - \|A\| \leqq 0$, $i = 1(1)n$ και λόγω της (11.18) η μεγαλύτερη απολύτως ιδιοτιμή του $A - \|A\|I$ είναι η $\lambda_n - \|A\| < 0$, πολλαπλότητος όσο και η λ_n . Με την εφαρμογή επομένως της μεθόδου της δυνάμεως στον πίνακα $A - \|A\|I$ μπορούμε να βρούμε την $\lambda_n - \|A\|$ και κατόπιν την λ_n καθώς και το ιδιοδιάνυσμα $cx^{(n)}$. Σημειώνεται ότι η μεγαλύτερη απολύτως από τις λ_1, λ_n είναι και η μεγαλύτερη απολύτως ιδιοτιμή του πίνακα A .

11.4 Υπολογισμός της απολύτως μικρότερης ιδιοτιμής ενός πίνακα A.

Έστω $A \in \mathbb{C}^{n,n}$ ένας διαγωνίσιμος πίνακας, για τις ιδιοτιμές λ_i , $i = 1(1)n$ του οποίου ισχύουν

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\lambda_{n-m+1} = \lambda_{n-m+2} = \ldots = \lambda_n \neq 0 \ , \quad m > 0 \ ,$$
$$|\lambda_1| \geqq |\lambda_2| \geqq \ldots \geqq |\lambda_{n-m}| > |\lambda_{n-m+1}| = \ldots = |\lambda_n| \ , \tag{11.21}$$

δηλαδή, η απολύτως μικρότερη ιδιοτιμή $\lambda_n \neq 0$ είναι πολλαπλότητος m. Υποθέτουμε επίσης ότι $x^{(i)}$, $i = 1(1)n$ είναι τα αντίστοιχα γραμμικώς ανεξάρτητα ιδιοδιανύσματα του A. Προφανώς, ο A θα είναι ομαλός και επομένως οι ιδιοτιμές του A^{-1} θα είναι οι $\frac{1}{\lambda_i}$, $i = 1(1)n$, με αντίστοιχα ιδιοδιανύσματα τα $x^{(i)}$, $i = 1(1)n$.
Λόγω των (11.21) θα έχουμε

$$\frac{1}{\lambda_{n-m+1}} = \frac{1}{\lambda_{n-m+2}} = \ldots = \frac{1}{\lambda_n} \ ,$$
$$\left|\frac{1}{\lambda_n}\right| = \left|\frac{1}{\lambda_{n-1}}\right| = \ldots = \left|\frac{1}{\lambda_{n-m+1}}\right| > \left|\frac{1}{\lambda_{n-m}}\right| \geqq \ldots \geqq \left|\frac{1}{\lambda_1}\right| \ , \tag{11.22}$$

οπότε η $\frac{1}{\lambda_n}$ (και στη συνέχεια η λ_n) μπορεί να βρεθεί εφαρμόζοντας την μέθοδο της δυνάμεως, σύμφωνα με την παράγραφο 11.2, στον πίνακα A^{-1}. Συγχρόνως, βρίσκουμε και το $x^{(n)}$. Η αντίστοιχη της (11.16) θα είναι τώρα

$$y^{(\kappa+1)} = A^{-1} z^{(\kappa)} \ , \quad \kappa = 0,1,2,\ldots \ . \tag{11.23}$$

Στην πράξη, για την εύρεση των διανυσμάτων $y^{(\kappa+1)}$, $\kappa = 0,1,2,\ldots$, δεν απαιτείται η εύρεση του A^{-1}, αφού η (11.23) μπορεί να γραφεί

$$Ay^{(\kappa+1)} = z^{(\kappa)} \ , \quad \kappa = 0,1,2,\ldots \ , \tag{11.24}$$

και επομένως για $\kappa = 0,1,2,\ldots$ το $y^{(\kappa+1)}$ προκύπτει λύνοντας το γραμμικό σύστημα (11.24). Για την λύση όλων αυτών των γραμμικών συστημάτων, που έχουν τον ίδιο πίνακα συντελεστών A, κατάλληλη είναι μία μέθοδος τριγωνικοποιήσεως (βλέπε παράγραφο 10.2.6).

11.5 Η αντίστροφη μέθοδος της δυνάμεως

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Με την μέθοδο της δυνάμεως, όπως είδαμε, μπορούμε κάτω από ορισμένες προυποθέσεις,να υπολογίσουμε την μεγαλύτερη απολύτως και την μικρότερη απολύτως ιδιοτιμή ενός πίνακα A καθώς και τα αντίστοιχα ιδιοδιανύσματα. Με την αντίστροφη μέθοδο της δυνάμεως μπορούμε να υπολογίσουμε μία οποιαδήποτε ιδιοτιμή. Για την εφαρμογή αυτής υποθέτουμε πάλι ότι ο A είναι διαγωνίσιμος και ότι μία καλή εκτίμηση λ είναι ήδη γνωστή για κάποια από τις ιδιοτιμές του $\lambda_1, \lambda_2, \ldots, \lambda_n$, π.χ. την λ_j την οποία θέλουμε να υπολογίσουμε, δηλαδή, ισχύει

$$|\lambda_j - \lambda| < |\lambda_m - \lambda| \quad \text{για κάθε} \quad \lambda_m \neq \lambda_j \,. \tag{11.25}$$

Έστωσαν ακόμη $x^{(i)}$, $i = 1(1)n$ τα αντίστοιχα ιδιοδιανύσματα του A. Εάν $\lambda \neq \lambda_i$, $i = 1(1)n$, τότε ο πίνακας $A - \lambda I$ είναι ομαλός και ο αντίστροφός του $(A - \lambda I)^{-1}$ έχει ιδιοτιμές $\frac{1}{\lambda_i - \lambda}$, $i = 1(1)n$, με αντίστοιχα ιδιοδιανύσματα τα $x^{(i)}$, $i = 1(1)n$. Λόγω της (11.25), θα έχουμε

$$\left|\frac{1}{\lambda_j - \lambda}\right| > \left|\frac{1}{\lambda_m - \lambda}\right| \,, \quad \text{για} \quad \lambda_m \neq \lambda_j \,. \tag{11.26}$$

Η αντίστροφη μέθοδος της δυνάμεως, για την εύρεση της λ_j του πίνακα A, ισοδυναμεί με την μέθοδο της δυνάμεως για την εύρεση της μεγαλύτερης απολύτως ιδιοτιμής (της $\frac{1}{\lambda_j - \lambda}$) του πίνακα $(A - \lambda I)^{-1}$. Έτσι, αρχίζοντας με ένα αυθαίρετο αρχικό διάνυσμα $y^{(0)} \in \mathbb{C}^n$, η αντίστοιχη της (11.16) θα είναι τώρα

$$y^{(\kappa+1)} = (A - \lambda I)^{-1} z^{(\kappa)} \,, \quad \kappa = 0,1,2,\ldots \,. \tag{11.27}$$

Για την (11.27),η οποία γράφεται

$$(A - \lambda I) y^{(\kappa+1)} = z^{(\kappa)} \,, \quad \kappa = 0,1,2,\ldots \,, \tag{11.28}$$

ισχύουν τα ίδια, που αναφέρθηκαν για την (11.23), όσον αφορά την εύρεση των $y^{(\kappa+1)}$, $\kappa = 0,1,2,\ldots$. Αφού βρεθεί η $\frac{1}{\lambda_j - \lambda}$, η λ_j βρίσκεται εύκολα, αφού το λ είναι γνωστό. Έπίσης, το αντίστοιχο ιδιο-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

διάνυσμα $x^{(j)}$ μπορεί να βρεθεί συγχρόνως με την $\frac{1}{\lambda_j - \lambda}$. Όσο πιο καλή είναι η προσέγγιση λ για την λ_j, τόσο πιο γρήγορα θα συγκλίνει η αντίστροφη μέθοδος της δυνάμεως. Σε περιπτώσεις, όπου η μέθοδος της δυνάμεως για την εύρεση της μεγαλύτερης απολύτως (ή της μικρότερης απολύτως) ιδιοτιμής του A συγκλίνει αργά, πράγμα που συμβαίνει όταν τα μέτρα των ιδιοτιμών του A δεν χωρίζονται επαρκώς, τότε, αφού εφαρμοσθούν μερικά βήματα αυτής, ώστε να βρεθεί κάποια εκτίμηση της ζητούμενης ιδιοτιμής, στη συνέχεια εφαρμόζεται η αντίστροφη μέθοδος της δυνάμεως για να έχουμε γρήγορη σύγκλιση. Στην πράξη, η αντίστροφη μέθοδος της δυνάμεως εφαρμόζεται κυρίως σε πίνακες A, για τους οποίους το σύστημα (11.28) λύνεται εύκολα, π.χ. σε τριδιαγώνιο πίνακα A.

11.6 Επέκταση της μεθόδου της δυνάμεως. Η διαδικασία deflation

Ας υποθέσουμε ότι οι ιδιοτιμές λ_i, $i = 1(1)n$ ενός διαγωνίσιμου πίνακα $A \in \mathbb{R}^{n,n}$ (ο A χάριν απλότητος θεωρείται πραγματικός) είναι τέτοιες ώστε

$$|\lambda_1| > |\lambda_2| > |\lambda_3| \geqq \dots \geqq |\lambda_n| \tag{11.29}$$

και ότι εφαρμόζοντας την μέθοδο της δυνάμεως έχουμε υπολογίσει την λ_1 και το αντίστοιχο ιδιοδιάνυσμα $x^{(1)}$. (Λόγω της (11.29) οι λ_1, λ_2 θα είναι απλές πραγματικές ιδιοτιμές $\neq 0$). Όπως θα δούμε στη συνέχεια, είναι δυνατόν, χρησιμοποιώντας μία διαδικασία που καλείται deflation, να υπολογίσουμε με την μέθοδο της δυνάμεως την ιδιοτιμή λ_2 και άλλες ιδιοτιμές καθώς και τα αντίστοιχα ιδιοδιανύσματα.

Η βασική ιδέα μιας διαδικασίας deflation είναι να μετασχηματισθεί ο πίνακας A σε ένα άλλο πίνακα A_1, του οποίου οι ιδιοτιμές είναι οι $\lambda_2, \lambda_3, \dots, \lambda_n, 0$. Επομένως, η μεγαλύτερη απολύτως ιδιοτιμή του A_1 θα είναι η λ_2 και θα μπορεί να βρεθεί αυτή, εφαρμόζοντας την μέθοδο της δυνάμεως στον πίνακα A_1. Εάν η πρώτη συνιστώσα του $x^{(1)}$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

είναι $\neq 0$, μπορούμε πάντοτε να υποθέσουμε ότι το $x^{(1)}$ έχει κανονικοποιηθεί ώστε $x_1^{(1)} = 1$. Εάν το $x^{(1)}$ που υπολογίστηκε έχει $x_1^{(1)} = 0$, τότε, αφού $x^{(1)} \neq 0$, θα υπάρχει τουλάχιστον μία συνιστώσα αυτού έστω η $x_m^{(1)} \neq 0$. Εάν P είναι ο πίνακας μεταθέσεως που προκύπτει από τον I αν ανταλλάξουμε την πρώτη με την m γραμμή του, τότε έχουμε $P^T = P = P^{-1}$ και ο πίνακας PAP είναι όμοιος με τον A . Επί πλέον, ισχύει προφανώς

$$(PAP)Px^{(1)} = \lambda_1 Px^{(1)}. \tag{11.30}$$

Η (11.30) σημαίνει ότι στην ιδιοτιμή λ_1 του PAP αντιστοιχεί το ιδιοδιάνυσμα $Px^{(1)}$ που η πρώτη του συνιστώσα είναι η $x_m^{(1)} \neq 0$. Από τα παραπάνω προκύπτει το συμπέρασμα ότι, χωρίς βλάβη της γενικότητος, μπορούμε να υποθέσουμε ότι η πρώτη συνιστώσα του ιδιοδιανύσματος $x^{(1)}$, που υπολογίστηκε με την λ_1 , είναι ίση με 1 . Εάν $A = (a_{ij})$, $i,j, = 1(1)n$, τότε μπορούμε να γράψουμε

$$A = \begin{bmatrix} a^{(1)^T} \\ a^{(2)^T} \\ \vdots \\ a^{(n)^T} \end{bmatrix} , \tag{11.31}$$

όπου $a^{(i)^T} = (a_{i1}, a_{i2}, \ldots, a_{in})$ είναι η i γραμμή του A , $i = 1(1)n$. Έτσι, αφού $Ax^{(1)} = \lambda_1 x^{(1)}$, θα έχουμε

$$a^{(1)^T} x^{(1)} = \lambda_1 x_1^{(1)} = \lambda_1 \ . \tag{11.32}$$

Σχηματίζουμε τώρα τον πίνακα

$$A_1 = A - x^{(1)} a^{(1)^T} . \tag{11.33}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Λόγω της (11.33), είναι φανερό ότι όλα τα στοιχεία της πρώτης γραμμής του A_1 είναι 0. Προφανώς, ο A_1 είναι μη ομαλός και επομένως το 0 είναι ιδιοτιμή του. Είναι εύκολο να δειχθεί ότι $A_1x^{(1)} = 0 = 0 \cdot x^{(1)}$, δηλαδή, το $x^{(1)}$ είναι ιδιοδιάνυσμα του A_1 που αντιστοιχεί στην ιδιοτιμή 0. Θα δειχθεί τώρα ότι οι άλλες ιδιοτιμές του A_1 είναι οι $\lambda_2, \lambda_3, \ldots, \lambda_n$. Για ένα οποιοδήποτε ιδιοδιάνυσμα $x^{(j)}$, $j \neq 1$ του A, που αντιστοιχεί στην ιδιοτιμή λ_j μπορούμε να υποθέσουμε (όπως και για το $x^{(1)}$) ότι η πρώτη του συνιστώσα ή είναι ίση με 1 ή είναι ίση με 0. Εάν $x_1^{(j)} = 1$, τότε λόγω της $Ax^{(j)} = \lambda_j x^{(j)}$, θα έχουμε

$$a^{(1)^T} x^{(j)} = \lambda_j \tag{11.34}$$

και

$$\begin{aligned} A_1(x^{(1)} - x^{(j)}) &= (A - x^{(1)}a^{(1)^T})(x^{(1)} - x^{(j)}) = \\ &= A(x^{(1)} - x^{(j)}) - x^{(1)}a^{(1)^T}(x^{(1)} - x^{(j)}) = \\ &= \lambda_1 x^{(1)} - \lambda_j x^{(j)} - x^{(1)}\lambda_1 + x^{(1)}\lambda_j = \lambda_j(x^{(1)} - x^{(j)}) . \end{aligned} \tag{11.35}$$

Η (11.35) αποδεικνύει ότι η λ_j είναι ιδιοτιμή του A_1 με αντίστοιχο ιδιοδιάνυσμα $x^{(1)} - x^{(j)}$, του οποίου η πρώτη συνιστώσα είναι 0. Εάν $x_1^{(j)} = 0$, τότε αντί της (11.34) θα έχουμε $a^{(1)^T}x^{(j)} = 0$ και επομένως

$$\begin{aligned} A_1x^{(j)} &= (A - x^{(1)}a^{(1)^T})x^{(j)} = \lambda_j x^{(j)} - x^{(1)}a^{(1)^T}x^{(j)} = \\ &= \lambda_j x^{(j)} , \end{aligned} \tag{11.36}$$

δηλαδή, πάλι η λ_j είναι ιδιοτιμή του A_1 με αντίστοιχο ιδιοδιάνυσμα $x^{(j)}$, του οποίου η πρώτη συνιστώσα είναι 0. Έτσι, έχουμε αποδείξει ότι οι $\lambda_2, \lambda_3, \ldots, \lambda_n$ είναι ιδιοτιμές του A_1 με αντίστοιχα ιδιοδιανύσματα που έχουν την πρώτη συνιστώσα ίση με 0. (Μπορεί να

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

δειχθεί ότι ο A_1 είναι επίσης διαγωνίσιμος).

Εάν υποθέσουμε ότι

$$A_1 = \left[\begin{array}{c|ccc} 0 & 0 & \dots & 0 \\ \hline b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{array}\right] \equiv \left[\begin{array}{c|c} 0 & \bigcirc \\ \hline b & B \end{array}\right]$$

και ότι $\begin{bmatrix} 0 \\ z^{(j)} \end{bmatrix}$, $j = 2(1)n$ είναι ιδιοδιανύσματά του που αντιστοιχούν στις λ_j, $j = 2(1)n$, όπου $z^{(j)}$ είναι $(n-1)$ – διάστατο διάνυσμα, τότε είναι

$$\left[\begin{array}{c|c} 0 & \bigcirc \\ \hline b & B \end{array}\right] \left[\begin{array}{c} 0 \\ \hline z^{(j)} \end{array}\right] = \lambda_j \left[\begin{array}{c} 0 \\ \hline z^{(j)} \end{array}\right] . \qquad (11.37)$$

Από την (11.37) έχουμε $Bz^{(j)} = \lambda_j z^{(j)}$, $j = 2(1)n$, δηλαδή, ο $(n-1)\times(n-1)$ πίνακας B, που προκύπτει από τον A_1 παραλείποντας την πρώτη γραμμή και την πρώτη στήλη του έχει ιδιοτιμές τις λ_j, $j = 2(1)n$, με αντίστοιχα ιδιοδιανύσματα τα $z^{(j)}$, $j = 2(1)n$. Έτσι, εφαρμόζοντας την μέθοδο της δυνάμεως στον πίνακα B (αντί για τον A_1), μπορούμε να βρούμε την μεγαλύτερη απολύτως ιδιοτιμή του που είναι η λ_2 και ένα αντίστοιχο ιδιοδιάνυσμα $z^{(2)}$. (Μπορεί να δειχθεί ότι ο B είναι διαγωνίσιμος).

Για την εύρεση του ιδιοδιανύσματος $x^{(2)}$ του A, που αντιστοιχεί στην λ_2, παρατηρούμε τα εξής: Εάν η πρώτη συνιστώσα του $x^{(2)}$ είναι 0, τότε λόγω της (11.36), θα είναι

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$w^{(2)} \equiv \begin{bmatrix} 0 \\ z^{(2)} \end{bmatrix} = cx^{(2)} \quad , \tag{11.38}$$

όπου c σταθερά $\neq 0$. Εάν $x_1^{(2)} = 1$, τότε, λόγω της (11.35), θα είναι

$$w^{(2)} = \hat{c}(x^{(1)} - x^{(2)}) \quad , \tag{11.39}$$

όπου $\hat{c}$ σταθερά $\neq 0$. Πολλαπλασιάζοντες τα δύο μέλη της (11.38) από αριστερά με $a^{(1)^T}$ παίρνουμε

$$a^{(1)^T} w^{(2)} = \hat{c}(a^{(1)^T} x^{(1)} - a^{(1)^T} x^{(2)}) = \hat{c}(\lambda_1 - \lambda_2). \text{ Άρα,}$$

$$\hat{c} = \frac{a^{(1)^T} w^{(2)}}{\lambda_1 - \lambda_2}$$

και από την (11.39) προκύπτει

$$x^{(2)} = x^{(1)} + \frac{\lambda_2 - \lambda_1}{a^{(1)^T} w^{(2)}} w^{(2)} \quad . \tag{11.40}$$

Επειδή στην πράξη το $x^{(2)}$ είναι άγνωστο, μόλις βρεθεί το $w^{(2)}$, σχηματίζουμε το $a^{(1)^T} w^{(2)}$. Εάν $a^{(1)^T} w^{(2)} = 0$, τότε ισχύει η (11.38). Εάν $a^{(1)^T} w^{(2)} \neq 0$ τότε ισχύει η (11.40) (βλέπε άσκηση 21) .

Η διαδικασία deflation με την οποία βρέθηκε ο πίνακας B μπορεί να επαναληφθεί στον B με σκοπό την εύρεση και των άλλων ιδιοτιμών και ιδιοδιανυσμάτων του A . Στην πράξη, λόγω των σφαλμάτων που υπεισέρχονται σε κάθε βήμα, εφαρμόζεται συνήθως για τον υπολογισμό ολίγων μόνον ιδιοτιμών,που έχουν κατά σειρά τα μεγαλύτερα μέτρα.

<u>Παράδειγμα 11.2</u> Να εφαρμοσθεί η διαδικασία deflation και η μέθοδος της δυνάμεως στον πίνακα A του παραδείγματος 11.1, για την εύρεση και των άλλων ιδιοτιμών και των αντίστοιχων ιδιοδιανυσμάτων του

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

(βρέθηκε ότι $\lambda_1 = 4$ και $x^{(1)} = (1, 1, 1)^T$) .

<u>Λύση</u> Είναι $A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 1 \end{bmatrix}$, $a^{(1)^T} = (1, 1, 2)$,

$$A_1 = A - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} [1, 1, 2] = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ 1 & 1 & 2 \end{bmatrix} = \left[\begin{array}{c|cc} 0 & 0 & 0 \\ \hline 0 & 1 & -1 \\ 1 & 0 & -1 \end{array}\right],$$

$B = \begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix}$. Οι ιδιοτιμές του B είναι προφανώς οι $\lambda_2 = 1$, $\lambda_3 = -1$.

Ας υποθέσουμε ότι βρέθηκε $z^{(2)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, οπότε $w^{(2)} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. Είναι

$$a^{(1)^T} w^{(2)} = (1, 1, 2) \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = 1 \neq 0 \text{ . Άρα, } x^{(2)} = x^{(1)} + \frac{\lambda_2 - \lambda_1}{a^{(1)^T} w^{(2)}} w^{(2)} =$$

$$= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \frac{1-4}{1} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} .$$

Αν ληφθεί σαν λ_2 η ιδιοτιμή -1 του B , τότε βρίσκουμε $z^{(2)} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$,

οπότε $w^{(2)} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$, $a^{(1)^T} w^{(2)} = (1, 1, 2) \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} = 5$,

$$x^{(2)} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \frac{-1-4}{5} \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} .$$

Δηλαδή, οι ιδιοτιμές του A είναι $\lambda_1 = 4$, $\lambda_2 = 1$, $\lambda_3 = -1$, με

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

αντίστοιχα ιδιοδιανύσματα $x^{(1)} = (1, 1, 1)^T$, $x^{(2)} = (1, -2, 1)^T$, $x^{(3)} = (1, 0, -1)^T$. ▲

11.7 Η LR μέθοδος

Έστω $A \in \mathbb{C}^{n,n}$ ένας ομαλός πίνακας, για τον οποίο είναι δυνατή η παραγοντοποίηση $A = LR$, όπου ο L είναι unit κάτω τριγωνικός και ο R άνω τριγωνικός (βλέπε θεώρημα 9.58). Η παραγοντοποίηση αυτή μπορεί να γίνει π.χ. με την βοήθεια της μεθόδου απαλοιφής του Gauss.

Θέτοντας $A_1 = A$, δημιουργούμε μία ακολουθία πινάκων A_κ, $\kappa = 1,2,\ldots$ ως εξής: Με την προυπόθεση ότι είναι δυνατή η παραγοντοποίηση του A_κ σε $A_\kappa = L_\kappa R_\kappa$ για κάθε κ, όπου L_κ είναι unit κάτω τριγωνικός και R_κ άνω τριγωνικός, τότε

$$A_{\kappa+1} = R_\kappa L_\kappa = L_{\kappa+1} R_{\kappa+1} \quad , \kappa = 1,2,\ldots . \tag{11.41}$$

Θα δειχθεί τώρα ότι

$$A_{\kappa+1} = L_\kappa^{-1} A_\kappa L_\kappa \quad , \kappa = 1,2,\ldots, \tag{11.42}$$

δηλαδή, ο $A_{\kappa+1}$ είναι όμοιος με τον A_κ και

$$A_{\kappa+1} = (L_1 L_2 \ldots L_\kappa)^{-1} A_1 (L_1 L_2 \ldots L_\kappa) \ , \kappa = 1,2,\ldots . \tag{11.43}$$

Πράγματι, από την $A_\kappa = L_\kappa R_\kappa$ παίρνουμε

$$L_\kappa^{-1} A_\kappa L_\kappa = R_\kappa L_\kappa = A_{\kappa+1} \ ,$$

οπότε

$$A_{\kappa+1} = L_\kappa^{-1} (L_{\kappa-1}^{-1} A_{\kappa-1} L_{\kappa-1}) L_\kappa = \ldots = L_\kappa^{-1} L_{\kappa-1}^{-1} \ldots L_1^{-1} A_1 L_1 L_2 \ldots$$
$$\ldots L_{\kappa-1} L_\kappa = (L_1 L_2 \ldots L_\kappa)^{-1} A_1 (L_1 L_2 \ldots L_\kappa) \ . \tag{11.44}$$

Αν θέσουμε $B_\kappa = L_1 L_2 \ldots L_\kappa$ και $U_\kappa = R_\kappa R_{\kappa-1} \ldots R_1$, $\kappa = 1,2,\ldots$, τότε οι B_κ και U_κ είναι αντίστοιχα κάτω τριγωνικός και άνω τριγω-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

νικός και λόγω της (11.44) έχουμε

$$B_K U_K = L_1 L_2 \dots L_K R_K R_{K-1} \dots R_1 = L_1 L_2 \dots L_{K-1} A_K R_{K-1} \dots R_1 =$$
$$= A_1 L_1 L_2 \dots L_{K-1} R_{K-1} \dots R_1 = A_1 B_{K-1} U_{K-1} = A_1^2 B_{K-2} U_{K-2} = \dots =$$
$$= A_1^{K-1} B_1 U_1 = A_1^{K-1} L_1 R_1 = A_1^K \text{ , δηλαδή,}$$

$$A^K = A_1^K = B_K U_K \quad , \kappa = 1,2,\dots \ . \tag{11.45}$$

Πρέπει να σημειωθεί ότι τα διαγώνια στοιχεία του B_K είναι ίσα με την μονάδα. Μπορεί να δειχθεί, κάτω από ορισμένες προυποθέσεις που αφορούν τον πίνακα A , ότι

$$\lim_{K\to\infty} A_K = \lim_{K\to\infty} R_K = U \ , \quad \lim_{K\to\infty} L_K = I \ , \tag{11.46}$$

όπου U είναι ένας άνω τριγωνικός πίνακας του οποίου τα διαγώνια στοιχεία είναι οι ιδιοτιμές του πίνακα A .

Στην περίπτωση που ο A είναι Ερμιτιανός θετικά ορισμένος πίνακας, τότε η μέθοδος LR τροποποιείται ως εξής: θέτουμε $\hat{A}_1 = A$ και ορίζουμε την ακολουθία $\hat{A}_K$, $\kappa = 1,2,\dots$ από τις

$$\begin{aligned} \hat{A}_K &= L_K L_K^H \\ \hat{A}_{K+1} &= L_K^H L_K = L_{K+1} L_{K+1}^H \quad , \kappa = 1,2,\dots, \end{aligned} \tag{11.47}$$

δηλαδή, για την τριγωνικοποίηση του $\hat{A}_K$ χρησιμοποιείται η μέθοδος τριγωνικής αναλύσεως του Cholesky (βλέπε παράγραφο 10.2.7 και θεώρημα 9.54) . Είναι φανερό ότι οι πίνακες $\hat{A}_K$, $\kappa = 1,2,\dots$ είναι Ερμιτιανοί θετικα ορισμένοι και ισχύει

$$\hat{A}_{K+1} = L_K^H L_K = L_K^H (\hat{A}_K (L_K^H)^{-1}) = L_K^H \hat{A}_K (L_K^H)^{-1} = L_K^{-1} \hat{A}_K L_K \ .$$

Στην περίπτωση αυτή η ακολουθία $\hat{A}_K$, $\kappa = 1,2,\dots$ συγκλίνει σε διαγώνιο πίνακα.

Στην πράξη ,η εύρεση όρων της ακολουθίας (11.41) A_K , $\kappa = 1,2,\dots$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

συνεχίζεται μέχρις ότου βρεθεί ένας πίνακας A_m στον οποίο τα κάτω της διαγωνίου στοιχεία του είναι επαρκώς κοντά στο 0 , δηλαδή, ο A_m είναι ουσιαστικά άνω τριγωνικός. Πρέπει να σημειωθεί ότι σε κάθε βήμα της LR μεθόδου (για τον υπολογισμό του A_{K+1} από τον A_K) απαιτείται μία πλήρης εφαρμογή της μεθόδου απαλοιφής του Gauss και επομένως συνολικά απαιτείται ένας μεγάλος αριθμός πράξεων . Για τον λόγο αυτό η LR μέθοδος εφαρμόζεται, αφού πρώτα ο αρχικός πίνακας, του οποίου θέλουμε να προσδιορίσουμε τις ιδιοτιμές, μετασχηματισθεί με τη βοήθεια ενός πεπερασμένου αριθμού μετασχηματισμών ομοιότητος σε ένα πίνακα A απλούστερης μορφής και συγκεκριμένα τριδιαγώνιο ή μορφής Hessenberg. Ο $A = (a_{ij})$ είναι άνω Hessenberg πίνακας εάν $a_{ij} = 0$ για $j \leq i - 2$, $i,j = 1(1)n$. (Ο $A = (a_{ij})$ είναι κάτω Hessenberg πίνακας εάν $a_{ij} = 0$ για $j \geq i + 2$, $i,j = 1(1)n$). Σημειώνεται ότι ένας Ερμιτιανός πίνακας μπορεί να μετασχηματισθεί σε ένα όμοιο Ερμιτιανό τριδιαγώνιο πίνακα (με τις μεθόδους των Householder, και Givens στις οποίες δεν θα αναφερθούμε), ενώ ένας γενικός πίνακας μπορεί να μετασχηματισθεί σε ένα όμοιο άνω Hessenberg πίνακα με τη μέθοδο Householder ή με μία παραλλαγή της απαλοιφής Gauss (πιο αποτελεσματική μέθοδος). Μπορεί να δειχθεί ότι, εφαρμόζοντας την LR μέθοδο, εάν ο A_K είναι τριδιαγώνιος ή Hessenberg πίνακας, τότε και ο A_{K+1} θα είναι της ίδιας μορφής.

11.8 Η QR μέθοδος

Επειδή η LR μέθοδος παρουσιάζει ορισμένες δυσκολίες και αδυναμίες,μία παρόμοια μέθοδος γνωστή σαν QR μέθοδος έχει προταθεί. Με βάση το θεώρημα 9.60 η QR μέθοδος αρχίζει με τον πίνακα $A_1 = A$ και δημιουργεί μία ακολουθία πινάκων A_K , $K = 1,2,\ldots,$ με τον ακόλουθο τρόπο:

$$A_K = Q_K R_K \quad , \qquad (11.48)$$
$$A_{K+1} = R_K Q_K = Q_{K+1} R_{K+1} \quad , \quad K = 1,2,\ldots,$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

όπου $Q_K^H = Q_K^{-1}$ και R_K είναι άνω τριγωνικός πίνακας, $\kappa = 1,2,\ldots$. Δηλαδή, ο πίνακας A_K αναλύεται σε γινόμενο ενός unitary πίνακα Q_K και ενός άνω τριγωνικού πίνακα R_K . Είναι εύκολο να δειχθεί από τις (11.48) ότι

$$A_{K+1} = Q_K^H A_K Q_K =$$

$$= B_K^H A_1 B_K \ , \quad \text{όπου} \quad B_K = Q_1 Q_2 \ldots Q_K \tag{11.49}$$

$$\text{και} \quad A^K = B_K U_K \ , \ \text{όπου} \quad U_K = R_K R_{K-1} \ldots R_1 \ . \tag{11.50}$$

Όπως και στην LR μέθοδο, κάτω από ορισμένες προυποθέσεις, είναι δυνατόν να αποδειχθεί ότι η ακολουθία A_K , $\kappa = 1,2,\ldots$ συγκλίνει σε ένα άνω τριγωνικό πίνακα του οποίου τα διαγώνια στοιχεία είναι οι ιδιοτιμές του A . Για την εφαρμογή της QR μεθόδου πρώτα ο αρχικός πίνακας μετασχηματίζεται σε ένα όμοιο απλούστερο πίνακα (ισχύουν εδώ όσα αναφέρθηκαν στην LR μέθοδο), δηλαδή, τριδιαγώνιο ή Hessenberg . Επειδή η QR μέθοδος χρησιμοποιεί unitary μετασχηματισμούς ομοιότητος (βλέπε 11.49) είναι φανερό ότι αν A_1 είναι Hermitian, τότε οι A_K , $\kappa = 1,2,\ldots$ είναι όλοι Ερμιτιανοί πίνακες και η ακολουθία A_K , $\kappa = 1,2,\ldots$ σε περίπτωση συγκλίσεως θα συγκλίνει σε διαγώνιο πίνακα. Πρέπει να σημειωθεί ότι αφού προσεγγισθούν οι ιδιοτιμές ενός τριδιαγωνίου ή Hessenberg πίνακα A , εφαρμόζοντας κάποια απ' τις μεθόδους LR και QR , στη συνέχεια, ο υπολογισμός των αντίστοιχων ιδιοδιανυσμάτων γίνεται με τη βοήθεια της αντίστροφης μεθόδου της δυνάμεως.

11.9 Ευστάθεια προβλημάτων ιδιοτιμών

Θα εξετάσουμε τώρα πως μεταβάλλονται οι ιδιοτιμές ενός πίνακα A εξαιτίας μικρών μεταβολών (διαταράξεων) στα στοιχεία αυτού. Θα περιοριστούμε στην κλάση των διαγωνισίμων πινάκων, δηλαδή, αυτών που έχουν n γραμμικώς ανεξάρτητα ιδιοδιανύσματα.

Έστω $A \in \mathbb{C}^{n,n}$ ένας διαγωνίσιμος πίνακας με ιδιοτιμές λ_i , $i =$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ

$= 1(1)n$ και E μία διατάραξη αυτού, δηλαδή, E είναι ένας πίνακας του οποίου τα στοιχεία είναι κατά κάποια έννοια μικρά απολύτως, συγκρινόμενα με τα στοιχεία του A. Για τις ιδιοτιμές του διαταραγμένου πίνακα $A + E$ ισχύει το ακόλουθο θεώρημα.

<u>Θεώρημα 11.1</u> Έστω $A \in \mathbb{C}^{n,n}$ διαγωνίσιμος με $X^{-1}AX = D$, όπου D είναι ο διαγώνιος πίνακας των ιδιοτιμών λ_i, $i = 1(1)n$ του A. Αν $E \in \mathbb{C}^{n,n}$ είναι μία διατάραξη του A, τότε για κάθε ιδιοτιμή μ του πίνακα $A + E$ ισχύει

$$\min_i |\mu - \lambda_i| \leqq \|X^{-1}\| \|X\| \|E\| . \tag{11.51}$$

<u>Απόδειξη</u> Εάν τεθεί $F = X^{-1}EX$, τότε ο πίνακας $X^{-1}(A + E)X = D + F$ είναι όμοιος με τον $A + E$. Έστω μ μία ιδιοτιμή του $D + F$. Τότε, ο πίνακας $\mu I - D - F$ είναι μή ομαλός. Εάν $\mu = \lambda_i$ για κάποιο i, τότε $\min_i |\mu - \lambda_i| = 0$ και επομένως ισχύει η (11.51). Υποθέτουμε τώρα ότι ο μ δεν είναι μία ιδιοτιμή του A. Έτσι, ο πίνακας $\mu I - D$ είναι ομαλός, ενώ ο πίνακας $(\mu I - D)^{-1}(\mu I - D - F) = I - (\mu I - D)^{-1}F$ είναι μή ομαλός. Αυτό σημαίνει ότι $\|(\mu I - D)^{-1}F\| \geqq 1$, διότι, αν ήταν $\|(\mu I - D)^{-1}F\| < 1$, τότε ο πίνακας $I - (\mu I - D)^{-1}F$ θα ήταν ομαλός (βλέπε θεώρημα 9.84), πράγμα που είναι άτοπο. Επομένως, έχουμε

$$1 \leqq \|(\mu I - D)^{-1}F\| \leqq \|(\mu I - D)^{-1}\| \|F\| = \max_i \{|\mu - \lambda_i|^{-1}\} \|F\| =$$

$$= \frac{1}{\min_i |\mu - \lambda_i|} \|X^{-1}EX\| \leqq \frac{1}{\min_i |\mu - \lambda_i|} \|X^{-1}\| \|X\| \|E\|$$

ή

$$\min_i |\mu - \lambda_i| \leqq \|X^{-1}\| \|X\| \|E\| . \qquad \blacktriangle$$

(Σημειώνεται ότι στην (11.51) η norm πίνακα είναι μία από τις $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$). Η σχέση (11.51) σημαίνει ότι κάθε ιδιοτιμή του πίνακα $A + E$ κείται σε ένα τουλάχιστον από τους κυκλικούς δίσκους με κέντρα

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

λ_i , $i = 1(1)n$ και ακτίνα $\|X^{-1}\|\|X\|\|E\|$. Δηλαδή, οι ιδιοτιμές του πίνακα $A + E$ κείνται στην ένωση των δίσκων αυτών. Είναι φανερό από την (11.51) ότι η ποσότης $\|X^{-1}\|\|X\|$ αποτελεί ένα μέτρο της ευαισθησίας των ιδιοτιμών του A , λόγω της διαταράξεως E . Επειδή ο πίνακας X δεν είναι μοναδικός (διότι αν T είναι ένας οποιοσδήποτε ομαλός διαγώνιος πίνακας, τότε $(XT)^{-1}A(XT) = T^{-1}(X^{-1}AX)T = T^{-1}DT = D$) , η (11.51) μπορεί να γραφεί

$$\min_i|\mu - \lambda_i| \leqq \left(\min_X\|X^{-1}\|\|X\|\right)\|E\| \ . \tag{11.52}$$

Ο αριθμός

$$K = \min_X\|X^{-1}\|\|X\| \ , \tag{11.53}$$

καλείται αριθμός καταστάσεως του A σε σχέση με το πρόβλημα υπολογισμού των ιδιοτιμών του και ισχύει $K \geqq 1$ αφού $\|X\|\|X^{-1}\| \geqq 1$. Αν χρησιμοποιηθεί η $\|\cdot\|_2$, τότε ο K_2 καλείται φασματικός αριθμός καταστάσεως. Γενικά, όσο πιο μεγάλος είναι ο αριθμός K , τόσο πιο ασταθές είναι το πρόβλημα ιδιοτιμών για τον A .

<u>Πόρισμα 11.1</u> Έστω $A \in \mathbb{C}^{n,n}$ Ερμιτιανός και λ_i , $i = 1(1)n$ οι ιδιοτιμές του. Αν $E \in \mathbb{C}^{n,n}$ είναι μία διατάραξη του A , τότε για κάθε ιδιοτιμή μ του $A + E$ ισχύει

$$\min_i|\mu - \lambda_i| \leqq \|E\|_2 \ . \tag{11.54}$$

<u>Απόδειξη</u> Αφού ο A είναι Ερμιτιανός, τότε ο πίνακας X στο θεώρημα 11.1 μπορεί να ληφθεί unitary, δηλαδή, $X^{-1} = X^H$, οπότε ισχύει $\|X\|_2 = \|X^{-1}\|_2 = 1$ και επομένως η (11.51) γίνεται

$$\min_i|\mu - \lambda_i| \leqq \|E\|_2 \ . \qquad \blacktriangle$$

Είναι φανερό από τα παραπάνω ότι για ένα Ερμιτιανό πίνακα A είναι $K_2 = 1$, δηλαδή ο K_2 παίρνει την μικρότερη δυνατή τιμή. Έτσι,από την (11.54) προκύπτει το συμπέρασμα ότι ο προσδιορισμός των ιδιοτιμών κάθε

Ερμιτιανού πίνακα είναι ένα τελείως ευσταθές πρόβλημα.

Παρατήρηση: Το πόρισμα 11.1 ισχύει γενικώτερα για κάθε normal πίνακα A (βλέπε θεώρημα 9.40).

ΑΣΚΗΣΕΙΣ

1. Να δειχθεί ότι το πρόβλημα ιδιοτιμών $Ax = \lambda x$, όπου $A \in \mathbb{C}^{n,n}$, $A^H = A$, μπορεί να αναχθεί σε ένα πρόβλημα ιδιοτιμών $Zy = \lambda y$, όπου $Z \in \mathbb{R}^{2n,2n}$, $Z^T = Z$.

2. Να δειχθεί ότι οι ρίζες της πολυωνυμικής εξισώσεως $P(\lambda) = \lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-1} + \ldots + a_{n-1}\lambda + a_n = 0$ είναι οι ιδιοτιμές του πίνακα F , που δίνεται από την (9.25). Ποια είναι τα αντίστοιχα ιδιοδιανύσματα του πίνακα F . Να δειχθεί ότι $\rho(F) \leqq 1 + \max_i |a_i|$.

3. Έστω $A \in \mathbb{C}^{n,n}$, $x \in \mathbb{C}^n$, $x \neq 0$, $\lambda \in \mathbb{C}$, $r = Ax - \lambda x$. Αν το x είναι σταθερό, να δειχθεί ότι η $\|r\|_2$ ελαχιστοποιείται ως προς λ για $\lambda = \dfrac{x^H A x}{x^H x}$.

4. Αν $x,y \in \mathbb{R}^n$, $x,y \neq 0$, να βρεθούν οι ιδιοτιμές του πίνακα xy^T .

5. i) Εάν ο $A = (a_{ij})$ είναι Ερμιτιανός και θετικά ορισμένος, να δειχθεί ότι $\rho(A) \geqq \max_i a_{ii}$ ii) Εάν $B = (b_{ij})$, $B > 0$, να δειχθεί, ότι $\rho(B) > \max_i b_{ii}$, $i = 1(1)n$, $n > 1$.

6. Έστω A ομαλός και $\det(\lambda I - A) = \lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \ldots + a_{n-1}\lambda + a_n$, $a_n \neq 0$. Να δειχθεί ότι $A^{-1} = -\dfrac{1}{a_n}(A^{n-1} + a_1 A^{n-2} + \ldots + a_{n-1} I)$.

7. Αν $\bar{\lambda}$ είναι μία προσέγγιση της ιδιοτιμής λ του Ερμιτιανού πίνακα A και $\bar{x}$ μία προσέγγιση του ιδιοδιανύσματος x που αντιστοιχεί στην λ , να δειχθεί ότι μία καλύτερη προσέγγιση της λ είναι η $\bar{\lambda} + \dfrac{c^H \bar{x}}{\bar{x}^H \bar{x}}$, όπου $c = A\bar{x} - \lambda \bar{x}$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

8. Αν $A \in \mathbb{R}^{n,n}$, $a + ib$ είναι μία ιδιοτιμή του και $x + iy$ ένα αντίστοιχο ιδιοδιάνυσμα, να δειχθεί ότι το x είναι ιδιοδιάνυσμα του πίνακα $B = A(A - 2aI)$ και ότι $y = \frac{1}{b}(ax - Ax)$. Ποιά είναι η αντίστοιχη ιδιοτιμή του B ;

9. Δίνεται το ακόλουθο γενικευμένο πρόβλημα ιδιοτιμών: Να βρεθεί αριθμός λ τέτοιος ώστε για κάποιο διάνυσμα $x \neq 0$ να ισχύει $Ax = \lambda Bx$, όπου $A,B \in \mathbb{C}^{n,n}$ είναι δοσμένοι πίνακες. Εάν είτε ο A ή ο B είναι ομαλός πίνακας, να δειχθεί ότι το πρόβλημα αυτό ανάγεται σε ένα πρόβλημα ιδιοτιμών $Cy = \mu y$ (κανονική μορφή). Εάν οι A,B είναι πραγματικοί συμμετρικοί και ένας απ' αυτούς θετικά ορισμένος, τότε, να δειχθεί ότι το γενικευμένο πρόβλημα ανάγεται στην κανονική μορφή $Cy = \mu y$, όπου ο C είναι πραγματικός συμμετρικός.

10. Να αποδειχθούν οι σχέσεις (11.17). Να δειχθεί ότι αντί της πρώτης των (11.17) μπορεί να χρησιμοποιηθεί η

$$\lim_{\kappa \to \infty} \frac{y_j^{(\kappa+1)}}{z_j^{(\kappa)}} = \lambda_1 \ , \ j = 1(1)n \quad \text{με} \quad z_j^{(\kappa)} \neq 0 \ .$$

11. Έστω $A \in \mathbb{R}^{n,n}$, διαγωνίσιμος, με ιδιοτιμές λ_i , $i = 1(1)n$, τέτοιες ώστε

$$|\lambda_1| = |\lambda_2| > |\lambda_3| \geqq \dots \geqq |\lambda_n| \ ,$$

όπου $\lambda_2 = -\lambda_1$ και αντίστοιχα ιδιοδιανύσματα $x^{(i)}$, $i = 1(1)n$.

Με βάση την (11.6) να δειχθεί ότι $\lim_{\kappa \to \infty} \frac{y_j^{(\kappa+2)}}{y_j^{(\kappa)}} = \lambda_1^2$ και

$\lim_{\kappa \to \infty} \left\{ y^{(\kappa+1)} + \lambda_j y^{(\kappa)} \right\} = c_j x^{(j)}$, $j = 1,2$, όπου c_1, c_2 σταθερές

$\neq 0$. Να γίνει εφαρμογή στον πίνακα $A = \begin{bmatrix} 5 & -10 & -10 \\ 3 & -1 & -3 \\ -3 & -4 & -2 \end{bmatrix}$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

12. Έστω $A \in \mathbb{R}^{n,n}$, διαγωνίσιμος, με ιδιοτιμές λ_i , $i = 1(1)n$ τέτοιες ώστε $\lambda_1 \in \mathbb{C}$, $\lambda_2 = \lambda_1^*$, $|\lambda_1| = |\lambda_2| > |\lambda_3| \geqq \ldots \geqq |\lambda_n|$ και αντίστοιχα ιδιοδιανύσματα $x^{(i)}$, $i = 1(1)n$. Να χρησιμοποιηθεί κατάλληλα η μέθοδος της δυνάμεως για την εύρεση των λ_1 , $x^{(1)}$.

13. Για την λύση του συστήματος $Ax = b$, $\det(A) \neq 0$, χρησιμοποιείται η επαναληπτική μέθοδος $x^{(m+1)} = Gx^{(m)} + \kappa$, $m = 0,1,2,\ldots$ και $x^{(0)}$ αυθαίρετο διάνυσμα. Αν ο G είναι διαγωνίσιμος και για τις ιδιοτιμές του λ_i , $i = 1(1)n$ ισχύει $|\lambda_1| > |\lambda_2| \geqq \ldots$ $\ldots \geqq |\lambda_n|$, να δειχθεί ότι $\lim\limits_{m \to \infty} \left\{ \dfrac{x_j^{(m)} - x_j^{(m-1)}}{x_j^{(m-1)} - x_j^{(m-2)}} \right\} = \lambda_1$,

$j = 1(1)n$ ..

14. Εάν $A^H = A$ και για τις ιδιοτιμές λ_i , $i = 1(1)n$ του A ισχύει $|\lambda_1| > |\lambda_2| \geqq \ldots \geqq |\lambda_n|$, να δειχθεί ότι

$$\lim_{\kappa \to \infty} \frac{y^{(\kappa)^H} A y^{(\kappa)}}{y^{(\kappa)^H} y^{(\kappa)}} = \lambda_1 ,$$

όπου $y^{(\kappa+1)} = Ay^{(\kappa)}$, $\kappa = 0,1,2,\ldots$ και $y^{(0)}$ αυθαίρετο διάνυσμα $\neq 0$. Να γίνει εφαρμογή στον πίνακα A του παραδείγματος 11.1 .

15. Έστω $A \in \mathbb{R}^{n,n}$ με ιδιοτιμές λ_i , $i = 1(1)n$ τέτοιες ώστε

$$|\lambda_1| > |\lambda_2| > |\lambda_3| \geqq \ldots \geqq |\lambda_n| , \lambda_1\lambda_2 < 0$$

και αντίστοιχα ιδιοδιανύσματα $x^{(i)}$, $i = 1(1)n$. Εάν η λ_1 και το $x^{(1)}$ είναι γνωστά, τότε για την εύρεση της λ_2 και του $x^{(2)}$ θεωρούμε τον πίνακα $A - \lambda_1 I$. Να δειχθεί ότι η $\lambda_2 - \lambda_1$ είναι η μεγαλύτερη απολύτως ιδιοτιμή του $A - \lambda_1 I$, δηλαδή, ισχύει

$|\lambda_2 - \lambda_1| > |\lambda_j - \lambda_1|$, $j = 1(1)n, j \neq 2$. Αν η συνθήκη $\lambda_1\lambda_2 < 0$ αντικατασταθεί από την $\lambda_1\lambda_2 > 0$, να δειχθεί ότι $0 < |\lambda_2 - \lambda_1| < |\lambda_i - \lambda_1|$, $i = 3(1)n$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

16. Έστω A Ερμιτιανός, με ιδιοτιμές λ_i , $i = 1(1)n$ και αντίστοιχα ιδιοδιανύσματα $x^{(i)}$, $i = 1(1)n$, όπου

$$|\lambda_1| > |\lambda_2| > |\lambda_3| \geqq \dots \geqq |\lambda_n| \quad \text{και} \quad \lambda_1 \ , \ x^{(1)} \quad \text{γνωστά.}$$

Να δειχθεί ότι ο πίνακας $A_1 = A - \lambda_1 x^{(1)} x^{(1)^H}$ έχει ιδιοτιμές τις $0, \lambda_2, \lambda_3, \dots, \lambda_n$ με αντίστοιχα ιδιοδιανύσματα τα $x^{(i)}$, $i = 1(1)n$.

17. Αν $A = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$ και $y = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, να υπολογιστεί το $A^{10}y$.

18. Να βρεθεί η μικρότερη και η μεγαλύτερη ιδιοτιμή του πίνακα.

$$A = \begin{bmatrix} 9 & 10 & 8 \\ 10 & 5 & -1 \\ 8 & -1 & 3 \end{bmatrix} ,$$

με ακρίβεια 2 δ.θ., χρησιμοποιώντας την θεωρία της παραγράφου 11.3 .

19. Για τις ιδιοτιμές του πίνακα $A = \begin{bmatrix} 1 & -1 & 2 \\ -2 & 0 & 5 \\ 6 & -3 & 6 \end{bmatrix}$ ισχύει

$|\lambda_1| < |\lambda_2| < |\lambda_3|$. Αν $\lambda_1 = 5$ και $x^{(1)} = (5, 16, 18)^T$, να βρεθούν οι λ_2, λ_3 και τα $x^{(2)}$, $x^{(3)}$ με την μέθοδο της δυνάμεως και την διαδικασία deflation.

20. Να βρεθούν οι ιδιοτιμές και τα ιδιοδιανύσματα των πινάκων

$$\begin{bmatrix} -1 & 1 & 0 \\ 1 & 5 & -1 \\ 0 & -1 & 2 \end{bmatrix} , \begin{bmatrix} 0 & 1 & 1 \\ 2 & 2 & -2 \\ -4 & 1 & 5 \end{bmatrix} , \begin{bmatrix} 2 & 8 & -6 \\ 1 & 3 & 0 \\ 5 & -3 & 2 \end{bmatrix} ,$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ

με την μέθοδο της δυνάμεως και την διαδικασία deflation.

21. Να δειχθεί ότι αν $a^{(1)^T} w^{(2)} = 0$ $(\lambda_2 \neq 0)$, τότε $x_1^{(2)} = 0$ (βλέπε παράγραφο 11.6) .

22. Να χρησιμοποιηθεί η LR μέθοδος στον πίνακα $A = \begin{bmatrix} 7 & 6 \\ 3 & 4 \end{bmatrix}$, για την εύρεση των A_2, A_3, A_4 . Να χρησιμοποιηθούν 4 δ.θ. στους υπολογισμούς. Ποιές είναι οι ιδιοτιμές του A ;

23. Στην QR μέθοδο, να δειχθεί ότι $\|A_{K+1}\|_2 \leq \|A_K\|_2$, $\kappa = 1,2,\ldots$.

24. Εάν $A \in \mathbb{C}^{n,n}$ είναι normal πίνακας, λ_i , $i = 1(1)n$ οι ιδιοτιμές του και $x^{(i)}$, $i = 1(1)n$ αντίστοιχα ιδιοδιανύσματα και για κάποια $\mu \in \mathbb{C}$ και $x \in \mathbb{C}^n$ με $\|x\|_2 = 1$ ισχύει $Ax - \mu x = c$, τότε να δειχθεί ότι $\min_i |\lambda_i - \mu| \leq \|c\|_2$. Χρησιμοποιώντας το παραπάνω συμπέρασμα, να αποδειχθεί το πόρισμα 11.1 .

25. Αν ο $A = (a_{ij})$, $i,j = 1(1)n$ είναι Ερμιτιανός πίνακας με ιδιοτιμές λ_i , $i = 1(1)n$, να δειχθεί ότι

$$\min_i |\lambda_i - a_{jj}| \leq \left(\sum_{\substack{\kappa=1 \\ \kappa \neq j}}^{n} |a_{j\kappa}|^2 \right)^{\frac{1}{2}} , \quad j = 1(1)n .$$

26. Εάν $A = Q^H B Q$, όπου $Q^H = Q^{-1}$ και A διαγωνίσιμος, να δειχθεί ότι οι A,B έχουν τον ίδιο φασματικό αριθμό καταστάσεως, σε σχέση με το πρόβλημα των ιδιοτιμών τους.

27. Να βρεθούν οι ιδιοτιμές και τα ιδιοδιανύσματα των ακολούθων πινάκων

$$\begin{bmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \\ -2 & -2 & 1 \end{bmatrix}, \begin{bmatrix} 6 & 4 & 4 & 1 \\ 4 & 6 & 1 & 4 \\ 4 & 1 & 6 & 4 \\ 1 & 4 & 4 & 6 \end{bmatrix}, \begin{bmatrix} 5 & 4 & 1 & 1 \\ 4 & 5 & 1 & 1 \\ 1 & 1 & 4 & 2 \\ 1 & 1 & 2 & 4 \end{bmatrix},$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\begin{bmatrix} 4 & -5 & 0 & 3 \\ 0 & 4 & -3 & -5 \\ 5 & -3 & 4 & 0 \\ 3 & 0 & 5 & 4 \end{bmatrix}, \quad \begin{bmatrix} 120 & 80 & 40 & -16 \\ 80 & 120 & 16 & -40 \\ 40 & 16 & 120 & -80 \\ -16 & -40 & -80 & 120 \end{bmatrix},$$

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 \\ -1 & 5 & 3 & 4 & 0 \\ 2 & 6 & 1 & 2 & 2 \\ 7 & 1 & 3 & 1 & 4 \end{bmatrix}.$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ

12

ΠΟΛΥΩΝΥΜΙΚΗ ΠΡΟΣΕΓΓΙΣΗ

12.1 Εισαγωγή

Ένα από τα βασικώτερα θέματα των Μαθηματικών και πρωταρχικής σημασίας για την ανάπτυξη των αριθμητικών μεθόδων είναι η θεωρία προσεγγίσεως των συναρτήσεων. Το γενικό πρόβλημα που παρουσιάζεται συχνά στην πράξη είναι το ακόλουθο. Υποθέτουμε ότι $f(x)$ είναι μία πραγματική συνάρτηση μιας πραγματικής μεταβλητής, ορισμένη στο πεπερασμένο διάστημα $[a,b]$, η οποία δίνεται είτε από μία πολύπλοκη αναλυτική εκφραση είτε από ένα πίνακα τιμών (π.χ. πειραματικά δεδομένα) και θέλουμε να προσδιορίσουμε μία άλλη απλούστερη συνάρτηση $F(x)$, η οποία να προσεγγίζει την $f(x)$, δηλαδή, να είναι,υπό κάποια έννοια,αρκετά "κοντά" στην $f(x)$ στο διάστημα $[a,b]$. Έτσι, για διάφορους υπολογιστικούς σκοπούς η $F(x)$ θα μπορεί να χρησιμοποιηθεί αντί της $f(x)$ στο διάστημα ορισμού της. Υπάρχουν βεβαίως πολλοί τρόποι για να εκλέξουμε την προσεγγιστική συνάρτηση $F(x)$, η οποία πρέπει να έχει πεπερασμένη παράσταση. Η εκλογή της κατάλληλης κλάσης συναρτήσεων στην οποία θα ανήκει η $F(x)$ εξαρτάται συνήθως από την υπολογιστική διαδικασία στην οποία πρόκειται να χρησιμοποιηθεί η $F(x)$. Η πιο συνηθισμένη κλάση συναρτήσεων $F(x)$ είναι τα πολυώνυμα, που είναι συνεχείς και ομαλές

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

συναρτήσεις που παραγωγίζονται και ολοκληρώνονται εύκολα. Στην περίπτωση αυτή, η προσέγγιση καλείται πολυωνυμική και σ' αυτή θα περιοριστούμε στις επόμενες παραγράφους. Μία μορφή πολυωνυμικής προσεγγίσεως είναι η πολυωνυμική παρεμβολή που αναπτύχθηκε στο κεφάλαιο 2. Εκτός από τα πολυώνυμα, άλλες συνήθεις κλάσεις προσεγγιστικών συναρτήσεων $F(x)$ είναι οι ρητές, οι τριγωνομετρικές και οι εκθετικές συναρτήσεις.

Είναι φανερό από την περιγραφή της $F(x)$ ότι πρέπει να έχουμε κάποιο τρόπο για να μπορούμε να μετρήσουμε την "απόσταση" της $F(x)$ από την $f(x)$ στο $[a,b]$, δηλαδή, πόσο "κοντά" είναι η $F(x)$ στην $f(x)$. Ο τρόπος αυτός είναι να μετρήσουμε το μέγεθος του σφάλματος, δηλαδή, το μέγεθος της συναρτήσεως $F(x) - f(x)$ στο $[a,b]$, με τη βοήθεια κάποιας norm συναρτήσεως. Οι πιο γνωστές norms συναρτήσεως αποτελούν φυσικές επεκτάσεις των γνωστών norms $\|\cdot\|_p$, $p = 1,2,\infty$ διανύσματος (βλέπε (9.32)). Έτσι, για την συνάρτηση $F(x) - f(x)$ έχουμε

$$\|F(x) - f(x)\|_1 = \int_a^b |F(x) - f(x)|dx , \tag{12.1}$$

$$\|F(x) - f(x)\|_2 = \left\{ \int_a^b |F(x) - f(x)|^2 dx \right\}^{\frac{1}{2}} , \tag{12.2}$$

$$\|F(x) - f(x)\|_\infty = \sup_{x\in[a,b]} |F(x) - f(x)| , \tag{12.3}$$

με την προυπόθεση ότι οι συναρτήσεις $F(x)$ και $f(x)$ στις (12.1) και (12.2) είναι ολοκληρώσιμες. Εάν για κάποια norm $\|\cdot\|_p$, η ποσότητα $\|F(x) - f(x)\|_p$ είναι μικρός αριθμός, τότε λέμε ότι η $F(x)$ είναι καλή προσέγγιση της $f(x)$ σε σχέση με την norm αυτή. Πρέπει να σημειωθεί ότι μπορεί η $F(x)$ να είναι καλή προσέγγιση της $f(x)$ σε σχέση με κάποια norm, αλλά να μην είναι καλή προσέγγιση σε σχέση με κάποια άλλη norm. Η εκλογή της norm εξαρτάται συνήθως από το πρόβλημα στο οποίο θα χρησιμοποιηθεί η $F(x)$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Εάν η συνάρτηση $f(x)$ δίνεται από ένα πίνακα τιμών, με τις τιμές της $f(x_i)$ στα $m+1$ σημεία x_i, $i = 0(1)m$ και $F(x_i)$, $i = 0(1)m$ είναι οι τιμές της $F(x)$ στα σημεία αυτά, τότε χρησιμοποιούμε τα διακεκριμένα ανάλογα των (12.1), (12.2) και (12.3), δηλαδή,

$$\|F(x) - f(x)\|_1 = \sum_{i=0}^{m} |F(x_i) - f(x_i)| \ , \tag{12.4}$$

$$\|F(x) - f(x)\|_2 = \left(\sum_{i=0}^{m} |F(x_i) - f(x_i)|^2 \right)^{\frac{1}{2}} \tag{12.5}$$

$$\|F(x) - f(x)\|_\infty = \max_{0 \leqq i \leqq m} |F(x_i) - f(x_i)| \ . \tag{12.6}$$

Η $\|\cdot\|_\infty$ που ορίζεται από την (12.3) ή (12.6) καλείται και "ομοιόμορφη" norm ή μεγίστη norm, ενώ η $\|\cdot\|_2$ καλείται Ευκλείδια norm.

12.2 Πολυωνυμική προσέγγιση

Έστω $f(x)$ μία συνάρτηση συνεχής σ' ένα (πεπερασμένο) διάστημα $[a,b]$. Εάν $\|\cdot\|_p$ είναι κάποια norm συναρτήσεως ($p = 1,2,\infty$, βλέπε (12.1) – (12.3)), τότε το πρόβλημά μας είναι να βρούμε ένα πολυώνυμο $\tilde{P}_n(x)$ βαθμού $\leqq n$ τέτοιο ώστε

$$\|\tilde{P}_n(x) - f(x)\|_p \leqq \|P_n(x) - f(x)\|_p \ , \tag{12.7}$$

για κάθε πολυώνυμο $P_n(x)$ βαθμού $\leqq n$. Εάν ένα τέτοιο πολυώνυμο $\tilde{P}_n(x)$ υπάρχει, τότε αυτό καλείται το "άριστο" προσεγγιστικό πολυώνυμο βαθμού $\leqq n$ για την $f(x)$ σε σχέση με την $\|\cdot\|_p$ στο $[a,b]$. Η απάντηση στο ερώτημα αυτό βρίσκεται στο ακόλουθο θεώρημα, που δίνεται χωρίς απόδειξη.

<u>Θεώρημα 12.1</u> Εάν η $f(x)$ είναι συνεχής στο $[a,b]$ και $\|\cdot\|_p$ είναι μία οποιαδήποτε norm συναρτήσεως ($p = 1,2,\infty$), τότε για κάθε ακέραιο $n \geqq 0$ υπάρχει ένα μοναδικό πολυώνυμο $\tilde{P}_n(x)$ βαθμού $\leqq n$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

τέτοιο ώστε

$$\|\tilde{P}_n(x) - f(x)\|_p \leqq \|P_n(x) - f(x)\|_p \quad ,$$

για κάθε πολυώνυμο $P_n(x)$ βαθμού $\leqq n$. ▲

Πρέπει να τονισθεί ότι, γενικά, το "άριστο" πολυώνυμο ως προς μία norm δεν είναι το ίδιο με το "άριστο" πολυώνυμο σε σχέση με κάποια άλλη norm. Το "άριστο" προσεγγιστικό πολυώνυμο για την $f(x)$ σε σχέση με την $\|\cdot\|_2$ καλείται "πολυώνυμο των ελαχίστων τετραγώνων" , ενώ το "άριστο" προσεγγιστικό πολυώνυμο για την $f(x)$ ως προς την $\|\cdot\|_\infty$ καλείται "άριστη ομοιόμορφη προσέγγιση" ή "προσέγγιση του Chebyshev" ή ακόμη "min – max πολυώνυμο". Το πολυώνυμο των ελαχίστων τετραγώνων για την $f(x)$ στο $[a,b]$ μπορεί να κατασκευαστεί αναλυτικά, δηλαδή, να υπολογιστεί σε ένα πεπερασμένο αριθμό βημάτων με μία ακολουθία τύπων, ενώ δεν ισχύει συνήθως το ίδιο για το $\tilde{P}_n(x)$ ως προς τις $\|\cdot\|_1$ και $\|\cdot\|_\infty$. Στις περιπτώσεις αυτές, το $\tilde{P}_n(x)$ μπορεί να προσεγγισθεί με διάφορες επαναληπτικές μεθόδους.

Στην περίπτωση που η συνάρτηση $f(x)$ ορίζεται σε ένα πεπερασμένο σύνολο σημείων $S = \{x_0, x_1, \ldots, x_m\}$, δηλαδή, είναι γνωστές οι τιμές της $f(x_i)$, $i = 0(1)m$, τότε το πρόβλημα της πολυωνυμικής προσεγγίσεως συνίσταται στην εύρεση ενός πολυωνύμου $\tilde{P}_n(x)$ βαθμού $\leqq n$, έτσι ώστε, αν $\|\cdot\|_p$ είναι κάποια norm ($p = 1,2,\infty$, βλέπε (12.4) – (12.6)), να ισχύει

$$\|\tilde{P}_n(x) - f(x)\|_p \leqq \|P_n(x) - f(x)\|_p \quad ,$$

για κάθε πολυώνυμο $P_n(x)$ βαθμού $\leqq n$. Είναι φανερό ότι αν $n = m$, τότε το $\tilde{P}_n(x)$ είναι το πολυώνυμο παρεμβολής των σημείων x_i , $i = 0(1)m$. Για $n > m$ υπάρχουν άπειρες λύσεις του προβλήματος, αφού υπάρχουν άπειρα πολυώνυμα τα οποία διέρχονται από τα σημεία $(x_i, f(x_i))$, $i = 0(1)m$. Η περίπτωση που έχει μεγαλύτερο πρακτικό ενδιαφέρον είναι η $n < m$. Στην περίπτωση αυτή για τις $\|\cdot\|_2$ και $\|\cdot\|_\infty$ ισχύει ανάλογο του θεωρήματος 12.1 και το $\tilde{P}_n(x)$ καλείται το "άριστο" προσεγ-

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

γιστικό πολυώνυμο βαθμού $\leq n$ για την $f(x)$ σε σχέση με την $\|\cdot\|_p$ ($p = 2$ ή ∞), που βασίζεται στα σημεία x_i , $i = 0(1)m$. Επίσης, οι ίδιες, όπως και προηγουμένως (στην συνεχή περίπτωση), ονομασίες χρησιμοποιούνται για το $\tilde{P}_n(x)$.

Στις επόμενες παραγράφους θα εξετάσουμε το πρόβλημα της ευρέσεως (κατασκευής) της "άριστης ομοιόμορφης προσεγγίσεως" και του "πολυωνύμου ελαχίστων τετραγώνων" στην συνεχή και διακεκριμένη περίπτωση. Κλείνοντας την παράγραφο αυτή αξίζει να αναφέρουμε (χωρίς απόδειξη) το ακόλουθο θεώρημα που έχει κυρίως θεωρητική αξία και είναι βασικό για την ανάπτυξη της θεωρίας της πολυωνυμικής προσεγγίσεως.

<u>Θεώρημα 12.2</u> (Weierstrass) Εάν η $f(x)$ είναι συνεχής στο κλειστό και φραγμένο διάστημα $[a,b]$, τότε για κάθε $\varepsilon > 0$ υπάρχει ένα πολυώνυμο $P_n(x)$ βαθμού $n = n(\varepsilon)$, τέτοιο ώστε $\|P_n(x) - f(x)\|_\infty = \max_{x\in[a,b]} |P_n(x) - f(x)| \leq \varepsilon$. ▲

12.3 Άριστη ομοιόμορφη προσέγγιση

Θεωρούμε πρώτα την διακεκριμένη περίπτωση και αποδεικνύουμε το ακόλουθο θεώρημα

<u>Θεώρημα 12.3</u> Εάν η συνάρτηση $f(x)$ ορίζεται στο σύνολο S των $n + 2$ σημείων x_i , $i = 0(1)n + 1$, όπου $-\infty < a \leq x_0 < x_1 < \ldots < x_{n+1} \leq b < \infty$, τότε υπάρχει ένα μοναδικό min max πολυώνυμο $\tilde{P}_n(x)$ βαθμού $\leq n$ για την $f(x)$ στο S . Εάν θέσουμε $\varepsilon(x) = \tilde{P}_n(x) - f(x)$, $E_n(S) = \max_{x\in S}|\varepsilon(x)| \equiv \|\varepsilon(x)\|_\infty$, τότε ισχύουν οι σχέσεις

$$\begin{aligned} &|\varepsilon(x_i)| = E_n(S) \quad , \; i = 0(1)n + 1 \quad \text{και} \\ &\varepsilon(x_i) = -\varepsilon(x_{i+1}) \; , \; i = 0(1)n \; . \end{aligned} \tag{12. 8}$$

<u>Απόδειξη</u> Έστω $P(x) = a_n x^n + \ldots + a_1 x + a_0$ ένα αυθαίρετο πολυώνυμο βαθμού $\leq n$ και $L = \max_{x\in S}|P(x) - f(x)|$. Γενικά , είναι $L \neq 0$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Αν $L = 0$, τότε το $P(x)$ είναι το ζητούμενο min max πολυώνυμο $\tilde{P}_n(x)$ και είναι μοναδικό διότι ταυτίζεται με το πολυώνυμο παρεμβολής με σημεία παρεμβολής οποιαδήποτε $n + 1$ σημεία του S . Ορίζουμε τις ποσότητες u_i από τις εξισώσεις

$$P(x_i) - f(x_i) = (-1)^i u_i L \ , \quad i = 0(1)n + 1 \ . \tag{12.9}$$

Προφανώς, είναι $|u_i| \leq 1$. Υποθέτουμε προς στιγμήν ότι τα u_i είναι γνωστά και θεωρούμε το ακόλουθο σύστημα των $n + 2$ εξισώσεων που προκύπτει από τις (12.9)

$$\begin{cases} -u_0 L + a_0 + a_1 x_0 + \ldots + a_n x_0^n = f(x_0) \\ u_1 L + a_0 + a_1 x_1 + \ldots + a_n x_1^n = f(x_1) \\ \vdots \qquad\qquad\qquad\qquad \vdots \quad \vdots \\ -(-1)^{n+1} u_{n+1} L + a_0 + a_1 x_{n+1} + \ldots + a_n x_{n+1}^n = f(x_{n+1}) \ , \end{cases} \tag{12.10}$$

σαν ένα γραμμικό σύστημα με αγνώστους τους a_i , $i = 0(1)n$ και L . Το σύστημα (12.10) γράφεται υπό μορφή πινάκων ως εξής:

$$\begin{bmatrix} -u_0 & 1 & x_0 & x_0^2 & \ldots & x_0^n \\ u_1 & 1 & x_1 & x_1^2 & \ldots & x_1^n \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ -(-1)^{n+1} u_{n+1} & 1 & x_{n+1} & x_{n+1}^2 & & x_{n+1}^n \end{bmatrix} \begin{bmatrix} L \\ a_0 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ f(x_{n+1}) \end{bmatrix} . \tag{12.11}$$

Η ορίζουσα D του πίνακα των συντελεστών του συστήματος (12.11) δίνεται από την

$$D = -u_0 D_0 - u_1 D_1 - \ldots - u_{n+1} D_{n+1} \ ,$$

όπου D_i , $i = 0(1)n + 1$ είναι οι ελάσσονες ορίζουσες των στοιχείων

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

της πρώτης στήλης του πίνακα. Είναι $D_i > 0$, $i = 0(1)n + 1$, αφού κάθε D_i είναι μία ορίζουσα του Vandermonde και $x_i - x_j > 0$ για $i > j$. Χρησιμοποιώντας τον κανόνα του Cramer, βρίσκουμε για τον άγνωστο L , συναρτήσει των u_i

$$L = \frac{D_0 f(x_0) - D_1 f(x_1) + \ldots + (-1)^{n+1} D_{n+1} f(x_{n+1})}{-u_0 D_0 - u_1 D_1 - \ldots - u_{n+1} D_{n+1}} . \qquad (12.12)$$

Ο αριθμητής στην (12.12) είναι ανεξάρτητος του $P(x)$. Εάν αυτός είναι μηδέν, τότε παίρνοντας $u_i = 1$, $i = 0(1)n + 1$ έχουμε $D \neq 0$, οπότε από την (12.12) είναι $L = 0$ και από την λύση του συστήματος (12.11) προκύπτει ένα μοναδικό πολυώνυμο $P(x)$ βαθμού $\leq n$, τέτοιο ώστε $P(x_i) = f(x_i)$, $i = 0(1)n + 1$, που είναι το ζητούμενο min max πολυώνυμο $\tilde{P}_n(x)$ (Αυτό συμπίπτει με το πολυώνυμο παρεμβολής με σημεία παρεμβολής οποιαδήποτε $n + 1$ σημεία του S). Είναι, δηλαδή, $E_n(S) = 0$ και προφανώς ισχύουν οι (12.8) . Εάν ο αριθμητής στην (12.12) είναι $\neq 0$, τότε για να βρούμε το $\tilde{P}_n(x)$ πρέπει και αρκεί να προσδιορίσουμε τα u_i , $i = 0(1)n + 1$, έτσι ώστε το $L(>0)$ να γίνεται ελάχιστο. Για την ελαχιστοποίηση του L πρέπει και αρκεί να μεγιστοποιήσουμε την $|D|$. Αφού $|u_i| \leq 1$ και $D_i > 0$, $i = 0(1)n + 1$, έχουμε

$$|D| = \left| \sum_{i=0}^{n+1} u_i D_i \right| \leq \sum_{i=0}^{n+1} |u_i| |D_i| \leq \sum_{i=0}^{n+1} D_i .$$

Επομένως, η $|D|$ γίνεται μεγίστη είτε λαμβάνοντας $u_i = 1$, $i = 0(1)r$ ή $u_i = -1$, $i = 0(1)n + 1$. Εάν ο αριθμητής στην (12.12) είναι > 0 τότε λαμβάνουμε $u_i = -1$, ενώ αν είναι < 0 τότε $u_i = 1$. Και στις δύο περιπτώσεις είναι $D \neq 0$ και επομένως το σύστημα (12.11) έχει μια μοναδική λύση για τους συντελεστές a_i , $i = 0(1)n$, βάσει των οποίων προσδιορίζεται το ζητούμενο $\tilde{P}_n(x)$. Είναι εύκολο να δειχθεί ότι ανεξάρτητα από το πρόσημο του αριθμητή στην (12.12) οι λύσεις των δύο γραμμικών συστημάτων, που προκύπτουν από το (12.11) για τις δύο διαφορετ

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

εκλογές των u_i , $i = 0(1)n + 1$, διαφέρουν μόνον ως προς το πρόσημο του L . Για το $\tilde{P}_n(x)$ θα ισχύει

$$E_n(S) = \frac{|D_0 f(x_0) - D_1 f(x_1) + \ldots + (-1)^{n+1} D_{n+1} f(x_{n+1})|}{D_0 + D_1 + \ldots + D_{n+1}} . \quad (12.13)$$

Επί πλέον, λόγω της (12.9), θα έχουμε είτε $\varepsilon(x_i) = \tilde{P}_n(x_i) - f(x_i) = (-1)^i E_n(S)$ ή $\varepsilon(x_i) = (-1)^i(-1)E_n(S)$, $i = 0(1)n + 1$ και επομένως οι (12.8) ικανοποιούνται. ▲

<u>Σημείωση</u> Από την απόδειξη του προηγουμένου θεωρήματος συνάγεται ότι για τον υπολογισμό του min max πολυωνύμου $\tilde{P}_n(x)$ βαθμού $\leq n$ για την $f(x)$ στο S αρκεί να λύσουμε το γραμμικό σύστημα (12.11) με $u_i = 1$, $i = 0(1)n + 1$. Αφού βρεθούν τα a_i , $i = 0(1)n$ και L τότε $\tilde{P}_n(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0$ και $E_n(S) = |L|$.

<u>Παράδειγμα 12.1</u> Αν $S = \{x_0 = 0 , x_1 = 1 , x_2 = 4\}$ και $f(x_0) = 0$, $f(x_1) = 1$, $f(x_2) = 2$, να βρεθεί το min max πολυώνυμο $\tilde{P}_1(x)$ βαθμού ≤ 1 για την $f(x)$ στο S .

<u>Λύση</u> Σχηματίζουμε το γραμμικό σύστημα

$$\begin{bmatrix} -1 & 1 & 0 \\ 1 & 1 & 1 \\ -1 & 1 & 4 \end{bmatrix} \begin{bmatrix} L \\ a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} ,$$

από το οποίο προκύπτει ότι $L = \frac{1}{4}$, $a_0 = \frac{1}{4}$, $a_1 = \frac{1}{2}$ και επομένως

$$\tilde{P}_1(x) = \frac{1}{2} x + \frac{1}{4} = \frac{1}{4} (2x + 1) \quad \text{και} \quad E_1(S) = \frac{1}{4} . \qquad ▲$$

Το ακόλουθο θεώρημα που δίνεται χωρίς απόδειξη χρησιμεύει για την εύρεση του min max πολυωνύμου $\tilde{P}_n(x)$ βαθμού $\leq n$ για την $f(x)$ σε ένα οποιοδήποτε πεπερασμένο σύνολο G από N σημεία όπου $N > n + 2$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

<u>Θεώρημα 12.4</u> Έστω $f(x)$ συνάρτηση ορισμένη σε ένα πεπερασμένο σύνολο G από N σημεία, όπου $N > n + 2$. Τότε, υπάρχει ένα μοναδικό min max πολυώνυμο $\tilde{P}_n(x)$ βαθμού $\leq n$ για την $f(x)$ στο G. Επί πλέον, εάν $S \subset G$, όπου το S αποτελείται από τα $n + 2$ σημεία $x_0 < x_1 < \ldots < x_{n+1}$ και είναι τέτοιο ώστε $E_n(S) \geq E_n(S')$ για κάθε υποσύνολο S' του G που αποτελείται από $n + 2$ σημεία, τότε $E_n(S) = E_n(G)$ και το min max πολυώνυμο για την $f(x)$ στο S είναι το ζητούμενο min max πολυώνυμο $\tilde{P}_n(x)$ για την $f(x)$ στο G. ▲

Σύμφωνα με το θεώρημα 12.4, για τον υπολογισμό του $\tilde{P}_n(x)$ για την $f(x)$ στο G, θεωρούμε όλα τα $\binom{N}{n+2}$ υποσύνολα του G που αποτελούνται από $n + 2$ σημεία. Για κάθε τέτοιο υποσύνολο S υπολογίζουμε το $E_n(S)$ από τον τύπο (12.13) και κατόπιν εκλέγουμε ένα σύνολο S για το οποίο το $E_n(S)$ είναι μέγιστο. Για αυτό το S είναι $E_n(S) = E_n(G)$ και το min max πολυώνυμο $\tilde{P}_n(x)$ για την $f(x)$ στο S αυτό, που υπολογίζεται σύμφωνα με το θεώρημα 12.3, είναι το ζητούμενο min max πολυώνυμο για την $f(x)$ στο G. Πρέπει να τονισθεί ότι ο τρόπος που περιγράφηκε με βάση το θεώρημα 12.4 για την εύρεση του $\tilde{P}_n(x)$ για την $f(x)$ στο G, όταν το N είναι μεγάλο δεν έχει πρακτική αξία. Για τον λόγο αυτό άλλοι αποτελεσματικώτεροι αλγόριθμοι έχουν προταθεί. Ένας απ' αυτούς, γνωστός σαν πρώτος αλγόριθμος του Remes ή αλγόριθμος της απλής ανταλλαγής, δίνεται στη συνέχεια χωρίς απόδειξη.

<u>Αλγόριθμος της απλής ανταλλαγής</u>

1. Εκλέγουμε ένα οποιοδήποτε υποσύνολο S του G από $n + 2$ σημεία τέτοια ώστε $x_0 < x_1 < \ldots < x_{n+1}$.

2. Υπολογίζουμε το min max πολυώνυμο $\tilde{P}_n(x)$ για την $f(x)$ στο S και το $E_n(S)$, σύμφωνα με το θεώρημα 12.3.

3. Βρίσκουμε ένα $\hat{x} \in G$ τέτοιο ώστε

$$|\tilde{P}_n(\hat{x}) - f(\hat{x})| = \max_{x \in G} |\tilde{P}_n(x) - f(x)| \equiv \Lambda .$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Εάν $\Lambda = E_n(S)$, τότε η διαδικασία τελειώνει και το $\tilde{P}_n(x)$ για την $f(x)$ στο S είναι το ζητούμενο min max πολυώνυμο για την $f(x)$ στο G . Διαφορετικά, προχωρούμε στο επόμενο βήμα.

4. Εάν το $\hat{x}$ κείται σε κάποιο διάστημα $[x_i, x_{i+1}]$, $i = 0(1)n$, τότε αντικαθιστούμε το σημείο x_i του S με το $\hat{x}$ εάν η $\tilde{P}_n(\hat{x}) - f(\hat{x})$ έχει το ίδιο πρόσημο με την $\tilde{P}_n(x_i) - f(x_i)$. Διαφορετικά, αντικαθιστούμε το x_{i+1} με το $\hat{x}$.

Εάν $\hat{x} < x_0$, τότε αντικαθιστούμε το x_0 με το $\hat{x}$ εάν η $\tilde{P}_n(x_0) - f(x_0)$ έχει το ίδιο πρόσημο με την $\tilde{P}_n(\hat{x}) - f(\hat{x})$. Διαφορετικά, αντικαθιστούμε το x_{n+1} με το $\hat{x}$.

Εάν $\hat{x} > x_{n+1}$ και τα πρόσημα των $\tilde{P}_n(x_{n+1}) - f(x_{n+1})$, $\tilde{P}_n(\hat{x}) - f(\hat{x})$ συμφωνούν, τότε αντικαθιστούμε το x_{n+1} με το $\hat{x}$. Διαφορετικά, αντικαθιστούμε το x_0 με το $\hat{x}$.

5. Έχουμε βρεί τώρα ένα νέο υποσύνολο S του G από $n + 2$ σημεία και επιστρέφουμε στο βήμα 2.

<u>Παράδειγμα 12.2</u> Δίνεται το σύνολο $G = \{0, 1, 4, 9\}$ και $f(0) = 0$, $f(1) = 1$, $f(4) = 2$, $f(9) = 3$. Να βρεθεί το min max πολυώνυμο $\tilde{P}_1(x)$ βαθμού ≤ 1 για την $f(x)$ στο G , πρώτα με βάση το θεώρημα 12.4 και κατόπιν με τον αλγόριθμο της απλής ανταλλαγής .

<u>Λύση</u> Είναι $N = 4$ και $n = 1$. Υπάρχουν $\binom{4}{3} = 4$ υποσύνολα του G από 3 σημεία. Αυτά είναι $S_1 = \{0, 1, 4\}$, $S_2 = \{0, 1, 9\}$, $S_3 = \{0, 4, 9\}$ και $S_4 = \{1, 4, 9\}$. Για το S_1 ο πίνακας του συστήματος (12.11) με $u_i = 1$ είναι

$$\begin{bmatrix} -1 & 1 & 0 \\ 1 & 1 & 1 \\ -1 & 1 & 4 \end{bmatrix}$$, οπότε $D_0 = 3$, $D_1 = 4$, $D_2 = 1$ και

$$E_1(S_1) = \frac{|3(0) - 4(1) + 1(2)|}{8} = \frac{1}{4}$$. Ομοίως για τα S_2, S_3 και

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

S_4 βρίσκουμε

$$E_1(S_2) = \frac{|8(0) - 9(1) + 1(3)|}{18} = \frac{1}{3}\ ,\ E_1(S_3) = \frac{|5(0) - 9(2) + 4(3)|}{18} =$$

$$= \frac{1}{3}\ ,\ E_1(S_4) = \frac{|5(1) - 8(2) + 3(3)|}{16} = \frac{1}{8}\ .$$ Επειδή $E_1(S_2) = \frac{1}{3} \geqq$ $\geqq E_1(S_i)$, $i = 0,1,3$, βρίσκουμε το min max πολυώνυμο $\tilde{P}_1(x)$ για την $f(x)$ στο S_2 κατά τα γνωστά, λύνοντας το σύστημα

$$\begin{bmatrix} -1 & 1 & 0 \\ 1 & 1 & 1 \\ -1 & 1 & 9 \end{bmatrix} \begin{bmatrix} L \\ a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} .$$

Έτσι, προκύπτει $\tilde{P}_1(x) = \frac{1}{3}(x + 1)$, που είναι το ζητούμενο min max πολυώνυμο.

Θα χρησιμοποιήσουμε τώρα τον αλγόριθμο της απλής ανταλλαγής και εκλέγουμε $S = S_1$. Σύμφωνα με το παράδειγμα 12.1 το $\tilde{P}_1(x)$ για την $f(x)$ στο S_1 είναι $\tilde{P}_1(x) = \frac{1}{4}(2x + 1)$ και $E_1(S_1) = \frac{1}{4}$.Βρίσκουμε $|\tilde{P}_1(9) - f(9)| = |\frac{1}{4}(18 + 1) - 3| = \frac{7}{4} > \frac{1}{4}$. Άρα $\hat{x} = 9$. Επειδή $\hat{x} > 4$ και $\tilde{P}_1(4) - f(4) = \frac{9}{4} - 2 = \frac{1}{4} > 0$ και $\tilde{P}_1(9) - f(9) = \frac{7}{4} > 0$, αντικαθιστούμε το σημείο 4 με το $\hat{x} = 9$ και έτσι έχουμε το νέο υποσύνολο $S_2 = \{0, 1, 9\}$. Το $\tilde{P}_1(x)$ για την $f(x)$ στο S_2 είναι $\tilde{P}_1(x) =$ $= \frac{1}{3}(x + 1)$ και $E_1(S_2) = \frac{1}{3}$. Επειδή $|\tilde{P}_1(4) - f(4)| = \left|\frac{5}{3} - 2\right| = \frac{1}{3} =$ $= E_1(S_2)$, συμπεραίνουμε ότι το ζητούμενο min max πολυώνυμο για την $f(x)$ στο G είναι το $\tilde{P}_1(x) = \frac{1}{3}(x + 1)$. ▲

Έστω τώρα $f(x)$ μία δοσμένη συνάρτηση, συνεχής στο διάστημα $[a,b]$. Σύμφωνα με το θεώρημα 12.1 υπάρχει ένα μοναδικό min max πολυώνυμο $\tilde{P}_n(x)$ βαθμού $\leqq n$ για την $f(x)$ στο $[a,b]$. Εκείνο που γίνεται συνήθως στην πράξη, με σκοπό την κατασκευή του $\tilde{P}_n(x)$, είναι η

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

αντικάτασταση του συνεχούς αυτού προβλήματος με το αντίστοιχο διακεκριμένο. Δηλαδή, εκλέγουμε ένα σύνολο S_N από N σημεία του διαστήματος $[a,b]$, όπου $N \geqq n+2$ και βρίκουμε το min max πολυώνυμο $\tilde{P}_{n,N}(x)$ βαθμού $\leqq n$ για την $f(x)$ στο S_N, σύμφωνα με όσα αναπτύχθηκαν στην παρούσα παράγραφο. Ένας τρόπος εκλογής του S_N είναι να αποτελείται από τα N σημεία $x_i = a + i\left(\frac{b-a}{N-1}\right)$, $i = 0(1)N-1$. Εάν $N \to \infty$, μπορεί να αποδειχθεί ότι η ακολουθία των min max πολυωνύμων $\tilde{P}_{n,N}(x)$, $N = n+2, n+3, \ldots$, συγκλίνει ομοιόμορφα στο ζητούμενο $\tilde{P}_n(x)$ για την $f(x)$ στο $[a,b]$, δηλαδή,

$$\tilde{P}_{n,N}(x) \underset{N\to\infty}{\to} \tilde{P}_n(x) .$$

Έτσι, είμαστε σε θέση να βρίσκουμε προσεγγίσεις για το $\tilde{P}_n(x)$, αφού εκτός από ορισμένες ειδικές περιπτώσεις δεν υπάρχει αποτελεσματική μέθοδος για τον αναλυτικό προσδιορισμό του. Τα ακόλουθα δύο βασικά θεωρήματα, που δίνονται χωρίς απόδειξη, περιέχουν τις απαιτούμενες πληροφορίες για τον προσδιορισμό του $\tilde{P}_n(x)$.

<u>Θεώρημα 12.5</u> (Chebyshev) Έστω $f(x)$ συνεχής συνάρτηση στο $[a,b]$ και $P_n(x)$ πολυώνυμο βαθμού $\leqq n$. Θέτουμε

$$\begin{aligned} \varepsilon(x) &= P_n(x) - f(x) , \\ E_n &= \max_{x\in[a,b]} |\varepsilon(x)| = \|\varepsilon(x)\|_\infty . \end{aligned} \tag{12.14}$$

Τότε, το $P_n(x)$ είναι το min max πολυώνυμο $\tilde{P}_n(x)$ για την $f(x)$ στο $[a,b]$ εάν και μόνον εάν υπάρχουν τουλάχιστον $n+2$ διακεκριμένα σημεία x_i, $i = 0(1)n+1$, τέτοια ώστε

$$a \leqq x_0 < x_1 < \ldots < x_{n+1} \leqq b \quad \text{και}$$

$$|\varepsilon(x_i)| = E_n , \quad i = 0(1)n+1 , \tag{12.15}$$

$$\varepsilon(x_i) = -\varepsilon(x_{i+1}) , \quad i = 0(1)n . \quad ▲ \tag{12.16}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

<u>Θεώρημα 12.6</u> Εάν η $f(x)$ είναι συνεχής στο $[a,b]$, τότε το μοναδικό min max πολυώνυμο $\tilde{P}_n(x)$ βαθμού $\leqq n$ για την $f(x)$ στο $[a,b]$ συμπίπτει με το min max πολυώνυμο για την $f(x)$ σε ένα σύνολο S από $n+2$ σημεία του $[a,b]$, το οποίο εκλέγεται έτσι ώστε $E_n(S) \geqq E_n(S')$, για οποιοδήποτε σύνολο S' από $n+2$ σημεία του $[a,b]$. ▲

Μία ειδική περίπτωση που μπορούμε να υπολογίσουμε εύκολα το min max πολυώνυμο $\tilde{P}_n(x)$ περιέχεται στο ακόλουθο θεώρημα.

<u>Θεώρημα 12.7</u> Το min max πολυώνυμο $\tilde{P}_n(x)$ βαθμού $\leqq n$ για την $f(x) \equiv Q_{n+1}(x)$, όπου $Q_{n+1}(x)$ είναι ένα πολυώνυμο $n+1$ βαθμού με συντελεστή μεγιστοβάθμιου όρου a_{n+1}, στο $[-1,1]$, σύμπίπτει με το πολυώνυμο παρεμβολής για την $f(x)$ με σημεία παρεμβολής τις ρίζες του πολυωνύμου του Chebyshev $T_{n+1}(x)$, $n+1$ βαθμού. Επί πλέον, αν

$$Q_{n+1}(x) = \sum_{i=0}^{n+1} c_i T_i(x) \text{ , τότε είναι } \tilde{P}_n(x) = \sum_{i=0}^{n} c_i T_i(x) .$$

<u>Απόδειξη</u> Αν x_i, $i = 0(1)n$ είναι οι ρίζες του $T_{n+1}(x)$ (βλέπε θεώρημα 4.14) και $P_n(x)$ είναι το πολυώνυμο παρεμβολής βαθμού $\leqq n$ για την $f(x) \equiv Q_{n+1}(x)$ με σημεία παρεμβολής τα x_i, $i = 0(1)n$, τότε έχουμε

$$P_n(x) - Q_{n+1}(x) = -\prod_{i=0}^{n}(x - x_i)\,\frac{Q_{n+1}^{(n+1)}(\xi)}{(n+1)!} = -\frac{T_{n+1}(x)}{2^n}\,\frac{a_{n+1}(n+1)!}{(n+1)!} =$$

$$= -\frac{a_{n+1}}{2^n}\,T_{n+1}(x) . \qquad (12.17)$$

Έτσι, είναι

$$|P_n(x) - Q_{n+1}(x)| = |a_{n+1}|\left|\frac{P_n(x) - Q_{n+1}(x)}{a_{n+1}}\right| = |a_{n+1}|\left|\frac{1}{2^n}\,T_{n+1}(x)\right| =$$

$$= \frac{|a_{n+1}|}{2^n}\,|T_{n+1}(x)| \leqq \frac{|a_{n+1}|}{2^n} , \quad x \in [-1,1] , \qquad (12.18)$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

δηλαδή,

$$\max_{x\in[-1,1]} |P_n(x) - Q_{n+1}(x)| = \frac{|a_{n+1}|}{2^n} \quad . \tag{12.19}$$

Αν τώρα $R_n(x)$ είναι ένα οποιοδήποτε πολυώνυμο βαθμού $\leq n$, τότε, σύμφωνα με το θεώρημα 4.20 , έχουμε

$$\max_{x\in[-1,1]} |R_n(x) - Q_{n+1}(x)| = \max_{x\in[-1,1]} |Q_{n+1}(x) - R_n(x)| =$$

$$= |a_{n+1}| \max_{x\in[-1,1]} \left| \frac{Q_{n+1}(x) - R_n(x)}{a_{n+1}} \right| \geq |a_{n+1}| \frac{1}{2^n} , \tag{12.20}$$

αφού το πολυώνυμο $\dfrac{Q_{n+1}(x) - R_n(x)}{a_{n+1}}$ είναι βαθμού $n + 1$ και έχει συντελεστή του x^{n+1} την μονάδα.

Από τις (12.19) και (12.20) προκύπτει ότι

$$\min_{R_n(x)} \max_{x\in[-1,1]} |R_n(x) - Q_{n+1}(x)| = \max_{x\in[-1,1]} |P_n(x) - Q_{n+1}(x)|$$

και επομένως το $P_n(x)$ είναι το ζητούμενο minmax πολυώνυμο $\tilde{P}_n(x)$ για το $Q_{n+1}(x)$ στο $[-1,1]$. Εξ άλλου, από την (12.17) έχουμε

$$\tilde{P}_n(x) \equiv P_n(x) = Q_{n+1}(x) - \frac{a_{n+1}}{2^n} T_{n+1}(x) \quad . \tag{12.21}$$

Αν $Q_{n+1}(x) = \sum_{i=0}^{n+1} c_i T_i(x)$, τότε εξισώνοντας τους συντελεστές του x^{n+1} στα δύο μέλη παίρνουμε $a_{n+1} = c_{n+1} 2^n$, οπότε η (12.21) δίνει

$$\tilde{P}_n(x) = \sum_{i=0}^{n+1} c_i T_i(x) - \frac{a_{n+1}}{2^n} T_{n+1}(x) = \sum_{i=0}^{n} c_i T_i(x) \quad . \qquad \blacktriangle \tag{12.22}$$

Τα ακόλουθα δύο θεωρήματα (το δεύτερο δίνεται χωρίς απόδειξη) αφορούν το min max πολυώνυμο $\tilde{P}_n(x)$ για μια συνεχή συνάρτηση στο $[a,b]$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ

<u>Θεώρημα 12.8</u> Έστω $f(x)$ συνάρτηση συνεχώς παραγωγίσιμη $n+1$ φορές στο $[a,b] \equiv [-1,1]$ και έστω $|f^{(n+1)}(x)| \leq M_{n+1}$, $\forall x \in [-1,1]$. Αν $\tilde{P}_n(x)$ είναι το min max πολυώνυμο βαθμού $\leq n$ για την $f(x)$ στο $[-1,1]$, τότε

$$\max_{x\in[-1,1]} |\tilde{P}_n(x) - f(x)| \leq \frac{M_{n+1}}{2^n(n+1)!} .$$

<u>Απόδειξη</u>

Έστω $P_n(x)$ το πολυώνυμο παρεμβολής βαθμού $\leq n$ για την $f(x)$ με σημεία παρεμβολής τις ρίζες x_i, $i = 0(1)n$ του πολυωνύμου Chebyshev $T_{n+1}(x)$. Τότε, ισχύει

$$P_n(x) - f(x) = -\prod_{i=0}^{n}(x - x_i)\,\frac{f^{(n+1)}(\xi)}{(n+1)!} = -\frac{1}{2^n}\,T_{n+1}(x)\,\frac{f^{(n+1)}(\xi)}{(n+1)!} ,$$

όπου $\xi \in [-1,1]$. Επομένως, έχουμε

$$\max_{x\in[-1,1]} |\tilde{P}_n(x) - f(x)| \leq \max_{x\in[-1,1]} |P_n(x) - f(x)| =$$

$$= \max_{x\in[-1,1]} \left|\frac{1}{2^n}\,T_{n+1}(x)\,\frac{f^{(n+1)}(\xi)}{(n+1)!}\right| \leq \frac{M_{n+1}}{2^n(n+1)!} . \qquad \blacktriangle$$

<u>Θεώρημα 12.9</u> Έστω $f(x)$ συνεχής συνάρτηση στο $[a,b]$ και $\tilde{P}_n(x)$, $n = 0,1,2,\ldots$ η ακολουθία των min max πολυωνύμων για την $f(x)$ στο $[a,b]$. Τότε,

$$\lim_{n\to\infty}\left\{\max_{x\in[a,b]} |\tilde{P}_n(x) - f(x)|\right\} = 0 ,$$

δηλαδή, η ακολουθία $\tilde{P}_n(x)$, $n = 0,1,2,\ldots$ συγκλίνει ομοιόμορφα στην $f(x)$. $\blacktriangle$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

<u>Παράδειγμα 12.3</u> Να βρεθεί το min max πολυώνυμο βαθμού ≤ 1 για την $f(x) = \sqrt{x}$ στο $[a,b]$, όπου $a \geq 0$.

<u>Λύση</u> Έστω $P_1(x) = a_1x + a_0$ πολυώνυμο βαθμού ≤ 1. Είναι $\varepsilon(x) = P_1(x) - f(x) = a_1x + a_0 - \sqrt{x}$. Εάν $\hat{x}$ είναι ένα σημείο του $[a,b]$, τέτοιο ώστε $\varepsilon(\hat{x}) = \max\limits_{x\in[a,b]} |\varepsilon(x)|$, τότε το $\hat{x}$ είναι σημείο ακροτάτου της $\varepsilon(x)$ και επομένως θα είναι είτε ένα από τα άκρα του $[a,b]$ ή ρίζα της $\varepsilon'(x) = 0$. Η μόνη ρίζα της $\varepsilon'(x) = 0$ είναι η $x = \dfrac{1}{4a_1^2}$, $(a_1 \neq 0)$. Επειδή σύμφωνα με το θεώρημα 12.5 πρέπει να ικανοποιούνται οι συνθήκες (12.15) και (12.16) για $n = 1$, ζητούμε να προσδιορίσουμε τα a_0, a_1 έτσι ώστε

$$\begin{cases} \varepsilon(a) = a_1a + a_0 - \sqrt{a} = E \\ \varepsilon\left(\dfrac{1}{4a_1^2}\right) = a_1\dfrac{1}{4a_1^2} + a_0 - \sqrt{\dfrac{1}{4a_1^2}} = \dfrac{1}{4a_1} + a_0 - \dfrac{1}{2a_1} = -E \\ \varepsilon(b) = a_1b + a_0 - \sqrt{b} = E. \end{cases} \tag{12.23}$$

Οι (12.23) αποτελούν ένα μή γραμμικό σύστημα με αγνώστους τα a_0, a_1 και E. Από την λύση του συστήματος αυτού προκύπτει ότι

$$\begin{cases} a_1 = \dfrac{1}{\sqrt{a} + \sqrt{b}}, \\ a_0 = \dfrac{1}{2}\left[\sqrt{b} - \dfrac{b}{\sqrt{a} + \sqrt{b}} + \dfrac{\sqrt{a} + \sqrt{b}}{4}\right] \end{cases} \tag{12.24}$$

και $E = a_1a + a_0 - \sqrt{a}$. Άρα, το ζητούμενο min max πολυώνυμο είναι $\tilde{P}_1(x) = a_1x + a_0$, όπου τα a_0, a_1 δίνονται από τις (12.24) και $\max\limits_{x\in[a,b]} |\tilde{P}_1(x) - \sqrt{x}| = |E|$. Για $[a,b] \equiv [0,9]$ βρίκουμε ότι $\tilde{P}_1(x) = \frac{1}{3}x + \frac{3}{8}$. Σημειώνεται ότι στο παράδειγμα 12.2 έχουμε βρεί σαν μία προσέγγιση αυτού, χρησιμοποιώντας τα σημεία 0,1,4,9 του $[0,9]$,

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

το $\frac{1}{3}x + \frac{1}{3}$.

12.4 Πολυώνυμο ελαχίστων τετραγώνων

Θα εξετάσουμε πρώτα την συνεχή περίπτωση, δηλαδή, το πρόβλημα του προσδιορισμού του πολυωνύμου ελαχίστων τετραγώνων (Π.Ε.Τ.) $\tilde{P}_n(x)$ βαθμού $\leq n$ για την συνεχή συνάρτηση $f(x)$ στο διάστημα $[a,b]$. Το $\tilde{P}_n(x)$ ως γνωστόν πρέπει να είναι τέτοιο ώστε

$$\|\tilde{P}_n(x) - f(x)\|_2^2 = \int_a^b \left(\tilde{P}_n(x) - f(x)\right)^2 dx \leq \int_a^b \left(P_n(x) - f(x)\right)^2 dx =$$

$$= \|P_n(x) - f(x)\|_2^2 , \qquad (12.25)$$

για κάθε πολυώνυμο $P_n(x)$ βαθμού $\leq n$.

Στον γραμμικό χώρο $C[a,b]$ ορίζουμε το εσωτερικό γινόμενο (f,g) δύο συναρτήσεων $f(x)$ και $g(x)$ από την

$$(f,g) = \int_a^b f(x)g(x)dx . \qquad (12.26)$$

Πρέπει να σημειωθεί ότι η (12.26) προκύπτει από την (4.1) αν

$w(x) \equiv 1$ ($w(x)$ είναι η συνάρτηση βάρους) .

Σύμφωνα με το κεφάλαιο 4 υπάρχει ένα σύνολο ορθογωνίων πολυωνύμων $\varphi_0(x)$, $\varphi_1(x)$, $\varphi_2(x)$, ... (σε σχέση με την συνάρτηση βάρους $w(x) \equiv 1$) στο $[a,b]$. Για παράδειγμα, αν $[a,b] \equiv [-1,1]$, τότε ένα τέτοιο είναι το σύνολο των πολυωνύμων Legendre. Λόγω της (12.26) είναι $\|g(x)\|_2^2 = (g,g)$ και επομένως

$$\|P_n(x) - f(x)\|_2^2 = (P_n - f , P_n - f) . \qquad (12.27)$$

Αφού το $P_n(x)$ είναι ένα οποιοδήποτε πολυώνυμο βαθμού $\leq n$, σύμφωνα

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

με το θεώρημα 4.1 , θα έχουμε κατά μοναδικό τρόπο

$$P_n(x) = \sum_{i=0}^{n} c_i\varphi_i(x) \ . \tag{12.28}$$

Ζητούμε τώρα να εκλέξουμε τα c_i , $i = 0(1)n$ έτσι ώστε η ποσότητα (12.27) να ελαχιστοποιείται. Λόγω των $(\varphi_i, \varphi_j) = 0$, $i \neq j$ και της (12.28), η (12.27) γίνεται

$$(P_n - f \ , \ P_n - f) = (P_n, P_n) - 2(P_n, f) + (f,f) =$$

$$= \left(\sum_{i=0}^{n} c_i\varphi_i \ , \ \sum_{i=0}^{n} c_i\varphi_i \right) - 2 \left(\sum_{i=0}^{n} c_i\varphi_i \ , \ f \right) + (f,f) =$$

$$= \sum_{i=0}^{n} c_i^2(\varphi_i,\varphi_i) - 2\sum_{i=0}^{n} c_i(\varphi_i,f) + (f,f) \equiv R(c_0,c_1,\ldots,c_n). \tag{12.29}$$

Αναγκαία συνθήκη για την ελαχιστοποίηση της συναρτήσεως $R(c_0,c_1,\ldots,c_n)$ είναι ο μηδενισμός όλων των μερικών παραγώγων πρώτης τάξεως αυτής, δηλαδή,

$$\frac{\partial R}{\partial c_j} = 2c_j(\varphi_j,\varphi_j) - 2(\varphi_j,f) = 0 \ , \ j = 0(1)n \ . \tag{12.30}$$

Από τις (12.30) προκύπτει ότι

$$c_j = \frac{(\varphi_j,f)}{(\varphi_j,\varphi_j)} \ , \ j = 0(1)n \ . \tag{12.31}$$

Πρέπει να σημειωθεί ότι $(\varphi_j,\varphi_j) \neq 0$, αφού $\varphi_j(x) \not\equiv 0$, $j = 0(1)n$. Το γεγονός ότι οι τιμές (12.31) πράγματι ελαχιστοποιούν την συνάρτηση R μπορεί εύκολα να δειχθεί λόγω των

$$\frac{\partial^2 R}{\partial c_j \partial c_\kappa} = \begin{cases} 2(\varphi_j,\varphi_j) > 0 & \text{εάν} \quad j = \kappa \\ 0 & \text{εάν} \quad j \neq \kappa \ . \end{cases} \tag{12.32}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Από τα παραπάνω προκύπτει ότι το ζητούμενο Π.Ε.Τ. $\tilde{P}_n(x)$ βαθμού $\leqq n$ για την $f(x)$ στο $[a,b]$ δίνεται από την

$$\tilde{P}_n(x) = \sum_{i=0}^{n} c_i \varphi_i(x) \ , \tag{12.33}$$

όπου τα c_i , $i = 0(1)n$ δίνονται από την (12.31). Επίσης, λόγω των (12.27), (12.29) και (12.33), είναι

$$\|\tilde{P}_n(x) - f(x)\|_2^2 = (f,f) - \sum_{i=0}^{n} \frac{(\varphi_i,f)^2}{(\varphi_i,\varphi_i)} \ . \tag{12.34}$$

Από την (12.34) προκύπτει η γνωστή ανισότητα του Bessel

$$\|f(x)\|_2^2 \geqq \sum_{i=0}^{n} \frac{(\varphi_i,f)^2}{(\varphi_i,\varphi_i)} \ . \tag{12.35}$$

Για την απόδειξη της μοναδικότητας του Π.Ε.Τ. $\tilde{P}_n(x)$ θα δείξουμε πρώτα ότι $(\tilde{P}_n - f, P_n) = 0$ για κάθε πολυώνυμο $P_n(x)$ βαθμού $\leqq n$.

Επειδή $P_n(x) = \sum_{i=0}^{n} \beta_i \varphi_i(x)$, αρκεί να δείξουμε ότι $(\tilde{P}_n - f, \varphi_i) = 0$ για $i = 0(1)n$. Πράγματι, έχουμε

$$(\tilde{P}_n - f \ , \ \varphi_i) = (\tilde{P}_n,\varphi_i) - (f,\varphi_i) = \left(\sum_{j=0}^{n} \frac{(\varphi_j,f)}{(\varphi_j,\varphi_j)} \varphi_j , \varphi_i \right) - (f,\varphi_i) =$$

$$= \frac{(\varphi_i,f)}{(\varphi_i,\varphi_i)} (\varphi_i,\varphi_i) - (f,\varphi_i) = 0 \ , \ i = 0(1)n \ .$$

Από τα παραπάνω συμπεραίνουμε ότι αν $P_n(x)$ είναι ένα οποιοδήποτε πολυώνυμο βαθμού $\leqq n$, $\tilde{P}_n(x) \neq P_n(x)$, τότε $(\tilde{P}_n - f, P_n - \tilde{P}_n) = 0$ και επομένως

$$(P_n - f, P_n - f) = ([P_n - \tilde{P}_n] + [\tilde{P}_n - f] \ , \ [P_n - \tilde{P}_n] + [\tilde{P}_n - f]) =$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$= (P_n - \tilde{P}_n, P_n - \tilde{P}_n) + (\tilde{P}_n - f, \tilde{P}_n - f) \quad \text{ή} \quad \|P_n - f\|_2^2 =$$

$$= \|P_n - \tilde{P}_n\|_2^2 + \|\tilde{P}_n - f\|_2^2 > \|\tilde{P}_n - f\|_2^2 \quad (\text{αφού } \|P_n - \tilde{P}_n\|_2^2 > 0) ,$$

πράγμα που αποδεικνύει την μοναδικότητα του $\tilde{P}_n(x)$.

<u>Παράδειγμα 12.4</u> Να βρεθεί το Π.Ε.Τ. $\tilde{P}_2(x)$ βαθμού ≤ 2 για την $f(x) = e^x$ στο $[-1,1]$.

<u>Λύση</u> Τα πολυώνυμα Legendre $\Phi_i(x)$, $i = 0,1,2,\ldots$ (βλέπε κεφάλαιο 4) αποτελούν ένα σύνολο ορθογωνίων πολυωνύμων σε σχέση με τη συνάρτηση βάρους $w(x) \equiv 1$ στο $[-1,1]$. Έτσι, θα είναι

$$\tilde{P}_2(x) = c_0\Phi_0(x) + c_1\Phi_1(x) + c_2\Phi_2(x) ,$$

όπου

$$c_i = \frac{(\Phi_i, e^x)}{(\Phi_i, \Phi_i)} = \frac{\int_{-1}^{1} \Phi_i(x)e^x dx}{\int_{-1}^{1} \Phi_i^2(x)\, dx} = \frac{2i+1}{2}\int_{-1}^{1} \Phi_i(x)e^x dx , \quad i = 0,1,2 .$$

Άρα,

$$c_0 = \frac{1}{2}\int_{-1}^{1} e^x dx = \frac{1}{2}\left(e - \frac{1}{e}\right) , \quad c_1 = \frac{3}{2}\int_{-1}^{1} xe^x dx = \frac{3}{2}\,\frac{2}{e} = \frac{3}{e} ,$$

$$c_2 = \frac{5}{2}\int_{-1}^{1} \frac{1}{2}(3x^2 - 1)e^x dx = \frac{15}{4}\int_{-1}^{1} x^2 e^x dx - \frac{5}{4}\int_{-1}^{1} e^x dx = \frac{15}{4}\left(e - \frac{5}{e}\right) -$$

$$- \frac{5}{4}\left(e - \frac{1}{e}\right) = \frac{10}{4}e - \frac{70}{4}\frac{1}{e} = \frac{5}{2}\left(e - \frac{7}{e}\right) \quad \text{και}$$

$$\tilde{P}_2(x) = \frac{1}{2}\left(e - \frac{1}{e}\right) + \frac{3}{e}x + \frac{5}{2}\left(e - \frac{7}{e}\right)\frac{1}{2}(3x^2 - 1) = \frac{15(e^2 - 7)}{4e}x^2 +$$

$$+ \frac{3}{e}x + \frac{3(11 - e^2)}{4e} \approx 0.53672x^2 + 1.10364x + 0.99629 .$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Λόγω της (12.34) έχουμε $\|\tilde{P}_2(x) - e^x\|_2^2 = (f,f) - \sum_{i=0}^{2} c_i^2(\varphi_i,\varphi_i) =$

$$= (f,f) - \sum_{i=0}^{2} \frac{2}{2i+1} c_i^2 \quad \text{ή}$$

$$\|\tilde{P}_2(x) - e^x\|_2^2 = \int_1^1 \left(e^x\right)^2 dx - 2\left[\frac{1}{2}\left(e - \frac{1}{e}\right)\right]^2 - \frac{2}{3}\left(\frac{3}{e}\right)^2 -$$

$$- \frac{2}{5}\left[\frac{5}{2}\left(e - \frac{7}{e}\right)\right]^2 = \frac{1}{2}\left(e^2 - \frac{1}{e^2}\right) - \frac{1}{2}\left(e - \frac{1}{e}\right)^2 - \frac{6}{e^2} - \frac{5}{2}\left(e - \frac{7}{e}\right)^2 =$$

$$= 36 - \frac{5}{2}e^2 - \frac{259}{2e^2} \approx 0.00144 \quad \text{και} \quad \|\tilde{P}_2(x) - e^x\|_2 \approx 0.03795 \ . \quad \blacktriangle$$

Πρέπει να σημειωθεί ακόμη ότι, αν $\tilde{P}_n(x)$, $n = 0,1,2,\ldots$ είναι η ακολουθία των πολυωνύμων ελαχίστων τετραγώνων για την $f(x) \in C[a,b]$, στο $[a,b]$, τότε μπορεί να αποδειχθεί ότι

$$\lim_{n\to\infty} \|\tilde{P}_n(x) - f(x)\|_2 = 0 \ , \tag{12.36}$$

δηλαδή, η ακολουθία $\tilde{P}_n(x)$, $n = 0,1,2,\ldots$ συγκλίνει στην $f(x)$ σε σχέση με την $\|\cdot\|_2$ και γράφουμε

$$f(x) = \sum_{i=0}^{\infty} \frac{(\varphi_i,f)}{(\varphi_i,\varphi_i)} \varphi_i(x) \ . \tag{12.37}$$

Η απειροσειρά στη (12.37) καλείται μια ορθογώνια σειρά ή σειρά Fourier για την $f(x)$. Λόγω της (12.36), από την (12.34) προκύπτει ότι

$$\|f(x)\|_2^2 = \sum_{i=0}^{\infty} \frac{(\varphi_i,f)^2}{(\varphi_i,\varphi_i)} \ . \tag{12.38}$$

Η σχέση (12.38) καλείται ισότητα του Parseval .

<u>Παρατηρήσεις:</u> i) Το πρόβλημα της ευρέσεως του Π.Ε.Τ. $\tilde{P}_n(x)$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

για την $f(x)$ στο $[a,b] \neq [-1,1]$, με την αλλαγή μεταβλητής: $x = \frac{b-a}{2}t + \frac{b+a}{2}$ μετατρέπεται στην εύρεση του Π.Ε.Τ. $\tilde{P}_n(t)$ για την $f(x(t)) \equiv F(t)$ στο $[-1,1]$ και έτσι μπορούμε να χρησιμοποιούμε τα πολυώνυμα Legendre.

ii) Εάν έχει βρεθεί το Π.Ε.Τ. $\tilde{P}_n(x)$ για την $f(x)$ στο $[a,b]$ τότε το Π.Ε.Τ. $\tilde{P}_{n+1}(x) = \tilde{P}_n(x) + c_{n+1}\varphi_{n+1}(x)$, όπου $c_{n+1} = \frac{(\varphi_{n+1}, f)}{(\varphi_{n+1}, \varphi_{n+1})}$ δηλαδή, για την εύρεση του $\tilde{P}_{n+1}(x)$ χρειάζεται να υπολογίσουμε μόνον τα $\varphi_{n+1}(x)$, c_{n+1} .

iii) Αν αντί της (12.26) χρησιμοποιηθεί η (4.1) , οπότε η $\|\cdot\|_2$ ορίζεται από την

$$\|f(x)\|_2^2 = (f,f) = \int_a^b w(x)\Big(f(x)\Big)^2 dx \ , \tag{12.39}$$

τότε η θεωρία της παρούσας παραγράφου μπορεί εύκολα να γενικευθεί. Για $w(x) = (1-x^2)^{-\frac{1}{2}}$ και $[a,b] \equiv [-1,1]$ χρησιμοποιούνται τα πολυώνυμα Chebyshev (βλέπε κεφάλαιο 4).

Στη συνέχεια θα αναζητήσουμε το Π.Ε.Τ. $\tilde{P}_n(x)$ βαθμού $\leq n$ για μιά συνάρτηση $f(x)$, που ορίζεται σε ένα πεπερασμένο σύνολο σημείων $S = \{x_0, x_1, \ldots, x_m\}$, όπου $m > n$. Για τον σκοπό αυτό αρκεί να βρούμε το

$$\min_{P_n(x)} \|P_n(x) - f(x)\|_2^2 = \min_{P_n(x)} \left\{ \sum_{i=0}^{m} \Big[P_n(x_i) - f(x_i)\Big]^2 \right\} \ ,$$

όπου το $P_n(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ διατρέχει το σύνολο όλων των πολυωνύμων βαθμού $\leq n$. Για την ελαχιστοποίηση της συναρτήσεως

$$R(a_0, a_1, \ldots, a_n) = \sum_{i=0}^{m} \left(\sum_{j=0}^{n} a_j x_i^j - f(x_i) \right)^2 \tag{12.40}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

εργαζόμαστε ως εξής: θέτουμε $f_i \equiv f(x_i)$, $i = 0(1)m$,

$$\hat{f} = (f_0, f_1, \ldots, f_m)^T , \quad \hat{a} = (a_0, a_1, \ldots, a_n)^T ,$$

$$X = \begin{bmatrix} 1 & x_0 & x_0^2 & \ldots & x_0^n \\ 1 & x_1 & x_1^2 & \ldots & x_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_m & x_m^2 & \ldots & x_m^n \end{bmatrix} ,$$

$$A = X^T X \quad \text{και} \quad \hat{g} = X^T \hat{f} . \tag{12.41}$$

Ο X είναι ένας $(m + 1) \times (n + 1)$ πίνακας. Η rank του X είναι $n + 1$, αφού ο υποπίνακας του X που σχηματίζεται από τις πρώτες $n + 1$ γραμμές αυτού και τις $n + 1$ στήλες του είναι ομαλός (ορίζουσα Vandermonde). Θα δειχθεί τώρα ότι ο A είναι θετικά ορισμένος και επομένως ομαλός. Πράγματι, ο A είναι $(n + 1) \times (n + 1)$ πραγματικός συμμετρικός πίνακας και για κάθε διάνυσμα $y \in \mathbb{R}^{n+1}$ είναι

$$y^T A y = y^T X^T X y = (Xy)^T (Xy) = \|Xy\|_2^2 \geqq 0 . \tag{12.42}$$

Από την (12.42) προκύπτει ότι $y^T A y = 0$ εάν και μόνον εάν $Xy = 0$. Επειδή $\text{rank}(X) = n + 1$, σύμφωνα με το θεώρημα 9.16 , η $Xy = 0$ ισχύει εάν και μόνον εάν $y = 0$, πράγμα που αποδεικνύει ότι $y^T A y > 0$, $\forall\, y \in \mathbb{R}^{n+1}$, $y \neq 0$. Είναι εύκολο να δειχθεί τώρα, έχοντας υπ' όψιν τις (12.40) και (12.41) , ότι

$$R(\hat{a}) = \|X\hat{a} - \hat{f}\|_2^2 = (X\hat{a} - f)^T (X\hat{a} - \hat{f}) = \hat{a}^T A \hat{a} - \hat{a}^T \hat{g} - \\ - \hat{g}^T \hat{a} + \hat{f}^T \hat{f} + \hat{g}^T A^{-1} \hat{g} - \hat{g}^T A^{-1} \hat{g} , \tag{12.43}$$

όπου προσθέσαμε και αφαιρέσαμε την ποσότητα $\hat{g}^T A^{-1} \hat{g}$. Η (12.43) μπορεί ακόμη να γραφεί

$$R(\hat{a}) = \hat{f}^T \hat{f} - \hat{g}^T A^{-1} \hat{g} + (\hat{a} - A^{-1} \hat{g})^T A (\hat{a} - A^{-1} \hat{g}) . \tag{12.44}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Στο δεξιό μέλος της (12.44) μόνον ο τρίτος όρος, που είναι ≥ 0 λόγω της (12.42), εξαρτάται από το άγνωστο διάνυσμα $\hat{a}$. Ο όρος αυτός μηδενίζεται εάν και μόνον εάν $\hat{a} - A^{-1}\hat{g} = 0$ και επομένως η ελάχιστη τιμή για την $R(\hat{a})$ συμβαίνει για

$$\hat{a} = A^{-1}\hat{g} \quad ,$$

δηλαδή, για την μοναδική λύση του γραμμικού συστήματος

$$Ax = \hat{g} \quad . \tag{12.45}$$

Αφού προσδιοριστεί το $\hat{a}$ σαν η μοναδική λύση του (12.45), ο προσδιορισμός στη συνέχεια του Π.Ε.Τ. $\tilde{P}_n(x)$ είναι προφανής. Επίσης, από την (12.44) προκύπτει ότι

$$\|\tilde{P}_n(x) - f(x)\|_2^2 = \hat{f}^T\hat{f} - \hat{g}^T\hat{a} \ . \tag{12.46}$$

Έτσι, απεδείχθη ότι υπάρχει ένα μοναδικό Π.Ε.Τ. $\tilde{P}_n(x)$, βαθμού $\leq n$ για την $f(x)$ στο σύνολο S . Οι εξισώσεις του γραμμικού συστήματος (12.45) καλούνται κανονικές εξισώσεις. Εκτός και αν το n είναι μικρό ($n \leq 6$) , μπορεί να δειχθεί ότι το σύστημα των κανονικών εξισώσεων είναι συνήθως πολύ ασταθές. Για τον λόγο αυτό στη συνέχεια, για την εύρεση του Π.Ε.Τ. $\tilde{P}_n(x)$ για την $f(x)$ στο σύνολο S θα χρησιμοποιηθεί μία διαφορετική μέθοδος η οποία βασίζεται στα ορθογώνια πολυώνυμα (όπως και στην συνεχή περίπτωση).

Στην διακεκριμένη περίπτωση του συνόλου $S = \{x_0, x_1, \ldots, x_m\}$, αν $f(x)$, $g(x)$ είναι δύο συναρτήσεις, τότε η έκφραση

$$(f,g) = \sum_{i=0}^{m} f(x_i)g(x_i) \tag{12.47}$$

ικανοποιεί τις ιδιότητες του εσωτερικού γινομένου με την μόνη εξαίρεση ότι η $(g,g) = 0$ δεν συνεπάγεται κατ' ανάγκην ότι $g(x) \equiv 0$. Αν όμως η $g(x)$ είναι πολυώνυμο βαθμού $\leq m$ τότε $(g,g) = 0$ εάν και μόνον εάν $g(x) \equiv 0$, αφού εάν ένα πολυώνυμο βαθμού $\leq m$ μηδενίζεται σε

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$m + 1$ διακεκριμένα σημεία πρέπει να είναι το μηδενικό πολυώνυμο. Δηλαδή, η έκφραση (12.47) είναι πράγματι ένα εσωτερικό γινόμενο στο σύνολο των πολυωνύμων βαθμού $\leq m$. Η (12.47) ορίζει το καλούμενο διακεκριμένο εσωτερικό γινόμενο και είναι φανερό ότι $\|g(x)\|_2^2 = (g,g)$ (βλέπε (12.5)).

<u>Ορισμός 12.1</u> Ένα σύνολο πολυωνύμων $\{\varphi_0(x),\varphi_1(x),\ldots,\varphi_n(x)\}$, όπου το $\varphi_\kappa(x)$ είναι ένα πολυώνυμο βαθμού κ καλείται σύνολο ορθογωνίων πολυωνύμων στο σύνολο S των σημείων x_i , $i = 0(1)m$, όπου $n \leq m$ εάν

$$(\varphi_\kappa,\varphi_j) = \sum_{i=0}^{m} \varphi_\kappa(x_i)\varphi_j(x_i) = 0 \quad \text{για} \quad \kappa \neq j \ , \tag{12.48}$$

όπου $\kappa, j = 0(1)n$. ▲

Είναι φανερό ότι $(\varphi_j,\varphi_j) = \|\varphi_j(x)\|_2^2 > 0$, $j = 0(1)n$, αφού $\varphi_j(x) \not\equiv 0$, $j = 0(1)n$. Η κατασκευή των πολυωνύμων $\varphi_0(x),\varphi_1(x),\ldots,\varphi_n(x)$ επιτυγχάνεται με τη βοήθεια των (4.4) και (4.5), με τη διαφορά ότι τα ολοκληρώματα αντικαθίστανται από αθροίσματα και $w(x) \equiv 1$. Συγκεκριμένα, θα είναι

$$\varphi_0(x) \equiv 1 \ , \quad \varphi_1(x) \equiv x - \frac{1}{m+1}\sum_{i=0}^{m} x_i \tag{12.49}$$

και

$$\varphi_{j+1}(x) = x\varphi_j(x) - \alpha_j\varphi_j(x) - \beta_j\varphi_{j-1}(x) \ , \ j = 1(1)n-1 \ ,$$

όπου

$$\alpha_j = \frac{\sum_{i=0}^{m} x_i\left(\varphi_j(x_i)\right)^2}{\sum_{i=0}^{m}\left(\varphi_j(x_i)\right)^2} \ ,$$

$$\beta_j = \frac{\sum_{i=0}^{m}\left(\varphi_j(x_i)\right)^2}{\sum_{i=0}^{m}\left(\varphi_{j-1}(x_i)\right)^2} \ . \tag{12.50}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

Για την εύρεση του Π.Ε.Τ. $\tilde{P}_n(x)$ βαθμού $\leq n$, $n < m$, για την $f(x)$ στο σύνολο S αρκεί να ελαχιστοποιήσουμε την έκφραση

$$\|P_n(x) - f(x)\|_2^2 = (P_n - f, P_n - f) = \sum_{i=0}^{m} \left[P_n(x_i) - f(x_i)\right]^2$$

ως προς όλα τα πολυώνυμα $P_n(x)$ βαθμού $\leq n$. Η διαδικασία που εφαρμόζεται στη συνέχεια είναι ίδια με εκείνη για την συνεχή περίπτωση (βλέπε (12.28) – (12.31)) και γι' αυτό παραλείπεται. Επομένως, θα είναι

$$\tilde{P}_n(x) = \sum_{j=0}^{n} c_j \varphi_j(x) \quad \text{με}$$

$$c_j = \frac{\sum_{i=0}^{m} f(x_i)\varphi_j(x_i)}{\sum_{i=0}^{m} \left[\varphi_j(x_i)\right]^2} \; ; \quad j = 0(1)n \; . \tag{12.51}$$

Η μέθοδος που περιγράφηκε είναι υπολογιστικά πιο αποτελεσματική αφού αποφεύγει τα μειονεκτήματα της μεθόδου που απαιτεί την λύση του συστήματος των κανονικών εξισώσεων.

<u>Παράδειγμα 12.5</u> Δίνεται το σύνολο $S = \{0, 1, 4, 9\}$ και οι τιμές $f(0) = 0$, $f(1) = 1$, $f(4) = 2$, $f(9) = 3$. Να βρεθεί το Π.Ε.Τ. $\tilde{P}_1(x)$ βαθμού ≤ 1 για την $f(x)$ στο S .

<u>Λύση</u> Είναι $x_0 = 0$, $x_1 = 1$, $x_2 = 4$, $x_3 = 9$,
$f_0 = 0$, $f_1 = 1$, $f_2 = 2$, $f_3 = 3$. Για την εύρεση του $\tilde{P}_1(x) = a_0 + a_1 x$ σχηματίζουμε το σύστημα των κανονικών εξισώσεων

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ x_0 & x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 1 & x_0 \\ 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ x_0 & x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ f_2 \\ f_3 \end{bmatrix} \quad \text{ή}$$

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$\begin{bmatrix} 4 & 14 \\ 14 & 98 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 6 \\ 36 \end{bmatrix} . \tag{12.52}$$

Λύνοντας το σύστημα (12.52) προκύπτει $a_0 = \frac{3}{7}$, $a_1 = \frac{15}{49}$. Άρα $\tilde{P}_1(x) =$ $= \frac{3}{7}(1 + \frac{5}{7}x)$. Λόγω της (12.46) θα είναι $\|\tilde{P}_1(x) - f(x)\|_2^2 =$

$$= (0,\ 1,\ 2,\ 3)\begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix} - (6,\ 36)\begin{pmatrix} 3/7 \\ 15/49 \end{pmatrix} = 14 - \frac{666}{49} = \frac{20}{49} \approx 0.40816 \quad \text{και}$$

$\|\tilde{P}_1(x) - f(x)\|_2 \approx 0.63888$.

Εάν χρησιμοποιήσουμε την μέθοδο με τα ορθογώνια πολυώνυμα, τότε θα είναι $\tilde{P}_1(x) = c_0\varphi_0(x) + c_1\varphi_1(x)$, όπου $\varphi_0(x) = 1$, $\varphi_1(x) =$ $= x - \frac{1}{4}\ 14 = x - \frac{7}{2}$ και

$$c_0 = \frac{6}{4} = \frac{3}{2}\ ,\ c_1 = \frac{1\left(1 - \frac{7}{2}\right) + 2\left(4 - \frac{7}{2}\right) + 3\left(9 - \frac{7}{2}\right)}{\left(0 - \frac{7}{2}\right)^2 + \left(1 - \frac{7}{2}\right)^2 + \left(4 - \frac{7}{2}\right)^2 + \left(9 - \frac{7}{2}\right)^2} =$$

$$= \frac{15}{49}\ .\ \text{Άρα,}\ \tilde{P}_1(x) = \frac{3}{2}\ 1 + \frac{15}{49}\left(x - \frac{7}{2}\right) = \frac{3}{7} + \frac{15}{49}x\ .$$

ΑΣΚΗΣΕΙΣ

1. Αν $f(x) = e^x$, $F(x) = 1 + x$ και $[a,b] = [-1,1]$, να υπολογιστούν οι $\|F(x) - f(x)\|_p$, $p = 1,2,\infty$.

2. Με βάση τις (12.1) – (12.3) , να δειχθεί ότι $\|F - f\|_1 \leq$ $\leq (b-a)\|F - f\|_\infty$ και $\|F - f\|_2 \leq (b-a)^{\frac{1}{2}}\|F - f\|_\infty$.

3. Να δειχθεί ότι το min max πολυώνυμο $\tilde{P}_1(x)$ για την $f(x)$ στο σύνολο $S = \{x_0, x_1, x_2\}$, όπου $x_0 < x_1 < x_2$, ταυτίζεται με την ευθεία

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

που διέρχεται από τα σημεία $\left(\frac{x_0 + x_1}{2}, \frac{f_0 + f_1}{2}\right)$, $\left(\frac{x_1 + x_2}{2}, \frac{f_1 + f_2}{2}\right)$. Ποια είναι η ικανή και αναγκαία συνθήκη ώστε $\tilde{P}_1(x) \equiv c =$ σταθερά.

4. Να βρεθεί το min max πολυώνυμο $\tilde{P}_3(x)$ για την $f(x) = \frac{1}{1 + x^2}$, στο σύνολο των σημείων $x_0 = -1$, $x_1 = -\frac{\sqrt{2}}{2}$, $x_2 = 0$, $x_3 = \frac{\sqrt{2}}{2}$, $x_4 = 1$.

5. Να βρεθεί το min max πολυώνυμο $\tilde{P}_2(x)$ για την $f(x) = 5x^3 + 3x^2 - 2x + 8$, πρώτα στο $[-1,1]$ και κατόπιν στο $[0,1]$.

6. Να βρεθεί το min max πολυώνυμο $\tilde{P}_1(x)$ για την $f(x) = e^x$ στο $[0,1]$. Πόσοι όροι του αναπτύγματος της σειράς Taylor για την e^x γύρω από το $x = 0$ χρειάζονται, για να έχουμε την ίδια ακρίβεια στο $[0,1]$ με εκείνη που δίνει το $\tilde{P}_1(x)$;

7. Έστω $f(x)$ συνεχής συνάρτηση στο $[a,b]$. Να βρεθεί το min max πολυώνυμο βαθμού 0 για την $f(x)$ στο $[a,b]$.

8. Να γίνει γενίκευση του θεωρήματος 12.7 για το διάστημα $[a,b]$.

9. Να δειχθεί ότι το πολυώνυμο $P(x) \equiv 0$ είναι το min max πολυώνυμο $\tilde{P}_n(x)$ βαθμού $\leq n$, $n = 0,1,2,\ldots,6$ για την $f(x) = \sin 4x$ στο $[0,2\pi]$.

10. Αν $m > n$, να δειχθεί ότι $\|\tilde{P}_m(x) - f(x)\|_\infty \leq \|\tilde{P}_n(x) - f(x)\|_\infty$ και επομένως $\lim_{n\to\infty} \|\tilde{P}_n(x) - f(x)\|_\infty = 0$, όπου $\tilde{P}_m(x)$, $\tilde{P}_n(x)$ είναι min max πολυώνυμα για την συνεχή συνάρτηση $f(x)$ στο $[a,b]$.

11. Να βρεθεί το min max πολυώνυμο $\tilde{P}_1(x)$ για την $f(x) = \frac{1}{1 + x}$ στο $[0,1]$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

12. Να βρεθεί το min max πολυώνυμο $\tilde{P}_2(x)$ για την $f(x) = \dfrac{1}{1 + x^2}$ στο $[-1,1]$.

13. Να βρεθεί το min max πολυώνυμο $\tilde{P}_2(x)$ για την $f(x) = |x|$ στο $[-1,1]$.

14. Να βρεθεί το πολυώνυμο ελαχίστων τετραγώνων για την συνεχή συνάρτηση $f(x)$ στο $[a,b]$.

15. Να βρεθεί το πολυώνυμο ελαχίστων τετραγώνων (Π.Ε.Τ.) $\tilde{P}_1(x)$ για την $f(x) = e^x$ στο $[-1,1]$. Να βρεθούν επίσης τα Π.Ε.Τ. $\tilde{P}_1(x)$ και $\tilde{P}_2(x)$ για την e^x στο $[0,1]$.

16. Για τις ακόλουθες συναρτήσεις να βρεθούν τα αντίστοιχα Π.Ε.Τ. στα αντίστοιχα διαστήματα. i) $f(x) = |x|$, $\tilde{P}_2(x)$ στο $[-1,1]$ ii) $f(x) = \cos x$, $\tilde{P}_1(x)$ στο $\left[0, \frac{\pi}{2}\right]$ iii) $f(x) = 1 + x^3$, $\tilde{P}_2(x)$ στο $[-1,1]$ iv) $f(x) = \sin x$, $\tilde{P}_3(x)$ στο $[-\pi,\pi]$ v) $f(x) = \cos \pi x$, $\tilde{P}_3(x)$ στο $[-1,1]$ vi) $f(x) = \sqrt{x}$, $\tilde{P}_1(x)$ στα $[0,9]$, $[0,1]$ και $\left[\frac{1}{4}, 1\right]$.

17. Αν $\tilde{P}_n(x)$ είναι το Π.Ε.Τ. για την $f(x)$ στο $[a,b]$, να δειχθεί ότι

$$\|\tilde{P}_n - f\|_2^2 = \|f\|_2^2 - \|\tilde{P}_n\|_2^2 .$$

18. Έστω $f(x) \in C[a,b]$. Να δειχθεί ότι το πολυώνυμο $P_n(x)$ βαθμού $\leq n$ είναι το Π.Ε.Τ. για την $f(x)$ στο $[a,b]$ εάν και μόνον εάν $(\hat{P}_n - f , P_n) = 0$, για κάθε πολυώνυμο $P_n(x)$ βαθμού $\leq n$, όπου $(\hat{P}_n - f , P_n) = \int_a^b \left(\hat{P}_n(x) - f(x)\right) P_n(x)dx$.

19. Να βρεθούν τα Π.Ε.Τ. $\tilde{P}_1(x)$, $\tilde{P}_2(x)$ για τα ακόλουθα δεδομένα:

x_i	-1	0	1	3	5
f_i	6	3	4	18	48

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

20. Να βρεθεί το Π.Ε.Τ. $\tilde{P}_1(x)$ για την $f(x)$ με δεδομένα:

x_i	-2	-1	0	1	2
f_i	$\frac{1}{2}$	$\frac{1}{2}$	2	3.5	3.5

.

21. Να κατασκευασθούν τα Π.Ε.Τ. $\tilde{P}_1(x)$, $\tilde{P}_2(x)$, $\tilde{P}_3(x)$ για τα ακόλουθα δεδομένα

x_i	-2	-1	0	1	2
f_i	-1	-1	0	1	1

.

22. Να βρεθεί το Π.Ε.Τ. $\tilde{P}_2(x)$ για τα δεδομένα

x_i	0	2	3	5
f_i	-1	0	2	3

.

23. Να προσεγγισθούν τα δεδομένα

x_i	-1	0	1	2
f_i	5	1	0.3	0.06

με μιά συνάρτηση Ae^{Bx} , ως προς την $\|\cdot\|_2$.

24. Δίνονται τα σημεία $(x_i, f(x_i))$, $i = 0(1)m$, $m > 1$. Να βρεθεί η ευθεία που ελαχιστοποιεί το άθροισμα των τετραγώνων των αποστάσεων των δοθέντων σημείων απ' αυτή.

25. Να προσαρμοσθεί μία καμπύλη της μορφής $y = \dfrac{1}{a + bx}$ στα ακόλουθα δεδομένα:

x_i	1	2	3	4	5
f_i	0.42	0.30	0.30	0.24	0.20

, με την βοήθεια Π.Ε.Τ.

26. Θεωρώντας τα σημεία $x_i = \frac{1}{4} + \frac{3}{4}\frac{i}{n}$, $i = 0(1)n$ $(n \to \infty)$, να δειχθεί ότι το Π.Ε.Τ. $\tilde{P}_1(x)$ για την $f(x) = \sqrt{x}$, στο $\left[\frac{1}{4}, 1\right]$ είναι $\tilde{P}_1(x) = \frac{88}{135}x + \frac{10}{27}$.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ

27. Να δειχθεί ότι το Π.Ε.Τ. $\tilde{P}_1(x)$ για τα δεδομένα (x_i, f_i), $i = 0(1)m$ διέρχεται από το σημείο $\left(\frac{1}{m+1}\sum_{i=0}^{m} x_i, \frac{1}{m+1}\sum_{i=0}^{m} f_i\right)$.

28. Αν $\tilde{P}_1(x)$ είναι το Π.Ε.Τ. βαθμού ≤ 1 για τα δεδομένα (x_i, f_i), $i = 0(1)m$, να δειχθεί ότι $\sum_{i=0}^{m}(f_i - \tilde{P}_1(x_i)) = 0$.

29. Για την $R(\hat{a})$ που δίνεται από την (12.43) να υπολογισθεί η $\frac{\partial R}{\partial a_i}$ και η $\frac{\partial^2 R}{\partial a_j \partial a_k}$. Να δειχθεί ότι $\frac{\partial R}{\partial a_i} = 0$, $i = 0(1)n$ εάν και μόνον εάν $X^T X \hat{a} = X^T \hat{f}$.

30. Να κατασκευασθεί ένα σύνολο ορθογωνίων πολυωνύμων $\{\varphi_0(x), \varphi_1(x), \varphi_2(x), \varphi_3(x)\}$ i) στο σύνολο των σημείων $x_i = 1 + i$, $i = 0(1)6$. ii) στο σύνολο των σημείων $x_0 = -1$, $x_1 = -\frac{1}{2}$, $x_2 = 0$, $x_3 = \frac{1}{2}$, $x_4 = 1$. iii) στο σύνολο $S = \{x_0 = -2, x_1 = -1, x_2 = 1, x_3 = 2\}$.

31. Να βρεθεί η λύση ελαχίστων τετραγώνων του συστήματος

$$\begin{cases} x_1 + x_2 = 1 \\ 2x_1 + x_2 = 0 \\ x_1 - x_2 = 0 . \end{cases}$$

32. Να δειχθεί ότι το σύστημα

$$\begin{cases} x_1 + 3x_2 + 7x_3 = 1 \\ 2x_1 + x_2 - x_3 = 1 \\ x_1 + 2x_2 + 4x_3 = 1 \end{cases}$$

δεν έχει λύση (είναι αδύνατο) και να βρεθεί μια λύση ελαχίστων τετραγώνων αυτού.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

$$x_i = \frac{c_i - \sum_{k=1, k \neq i}^{n} c_k x_k}{c_{ii}}$$

ΒΙΒΛΙΟΓΡΑΦΙΑ

1. ACTON F.S., "Numerical Methods that work", Harper & Row, 1970.
2. ALLEN SMITH, W., "Elementary Numerical Analysis", Harper & Row, 1979.
3. ARDEN B.W. and ASTILL K.N., "Numerical algorithms: origins and applications", Addison — Wesley, 1970.
4. BAKHVALOV N.S., "Numerical Methods" (Translated from the Russian), Mir publishers, 1977.
5. BEREZIN I.S. and ZHIDKOV N.P., "Computing Methods", Vol.1-2, Pergamon Press, 1965.
6. BLUM E.K., "Numerical Analysis and Computation: Theory and Practice", Addison — Wesley, 1972.
7. COHEN A.M., et al. "Numerical Analysis", McGraw — Hill, 1973.
8. CONTE S.D. and C. de BOOR, "Elementary Numerical Analysis", 2d ed., McGraw — Hill, 1972.
9. DODES I.A., "Numerical Analysis for Computer Science", Elsevier, North-Holland, Inc., 1978.
10. DAHLQUIST G. and BJÖRCK A., "Numerical Methods", Prentice — Hall, Inc., 1974.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΠΑΤΡΩΝ

11. DANIELS R.W., "An Introduction to Numerical Methods and Optimization Techniques", Elsevier North — Holland, 1978.

12. DAVIS P.J., RABINOWITZ P., "Methods of Numerical Integration", Academic Press, 1975.

13. DAVIS P.J., "Interpolation and Approximation", Blaisdell Publishing Company, 1963.

14. DIXON C., "Numerical Analysis", Blackie — Chambers, 1974.

15. DEMIDOVICH B.P. and MARON I.A., "Computational Mathematics", Mir Publishers, 1973.

16. ENGELS H., "Numerical Quadrature and Cubature", Academic Press, 1980.

17. FORSYTHE G., MALCOLM M. and MOLER C., "Computer Methods for Mathematical Computations", Prentice — Hall, 1977.

18. FRANKLIN J.N., "Matrix theory", Prentice — Hall, Inc., 1968.

19. FRÖBERG C.E., "Introduction to Numerical Analysis", Second edition, Addison — Wesley, 1969.

20. GASTINEL N., "Linear Numerical Analysis", Hermann, 1970.

21. GEORGE A. and LIU J., "Computer Solution of Large Sparse Positive Definite Systems", Prentice — Hall, 1981.

22. GERALD C.F., "Applied Numerical analysis", Addison — Wesley, 1970.

23. GOLUB G.H. and VAN LOAN C.F., "Matrix Computations", North Oxford Academic, 1983.

24. GOULT R.J., HOSKINS R.F., MILNER J.A., PRATT M.J., "Computational Methods in Linear Algebra", Stanley Thornes, 1974.

25. GREENSPAN D., "Introduction to Numerical analysis and Applications", Markham, 1971.

26. HAGEMAN L.A. and YOUNG D.M., "Applied Iterative Methods", Academic Press, 1981.

27. HAMMING R.W., "Introduction to applied numerical analysis", Mc Graw — Hill, 1971.

28. HENRICI P., "Applied and Computational Complex Analysis", Vol. 1, Wiley, 1974.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

29. HILDEBRAND F.B., "Introduction to Numerical Analysis",2d ed., Mc Graw — Hill, 1974.

30. HORNBECK R.W. "Numerical Methods", Quantum Publishers, Inc., 1975.

31. HOSKING R.J., JOYCE D.C., TURNER J.C., "First steps in Numerical Analysis", Hodder and Stoughton, 1978.

32. HOUSEHOLDER A.S.,"The theory of Matrices in Numerical Analysis", Blaisdell Publ. Company, 1964.

33. HOUSEHOLDER A.S., "The Numerical Treatment of a Single Nonlinear Equation", McGraw — Hill, 1970.

34. ISAACSON E. and KELLER H.B., "Analysis of Numerical Methods", Wiley, 1966.

35. JOHNSON L.W. and RIESS R.D., "Numerical Analysis", Addison — Wesley, 1977.

36. KELLY L.G., "Handbook of Numerical Methods and Applications", Addison — Wesley, 1967.

37. KRONSJÖ L.I., "Algorithms, their complexity and efficiency", John Wiley & Sons, 1979.

38. KUO S.S., "Computer Applications of Numerical Methods", Addison — Wesley, 1972.

39. LAW A., and SCHNEY B., "Theory of Approximation with Applications", Academic Press, 1976.

40. LAWSON C.L. and HANSON R.J., "Solving Least Squares Problems", Prentice — Hall, 1974.

41. LINZ P., "Theoretical Numerical Analysis" John Wiley & Sons, 1979.

MYSOVSKIH I.P., "Lectures on Numerical Methods", Wolters — Noordhoff Publishing Groningen, 1969.

NOBLE B., "Applied Linear Algebra", Prentice — Hall, Inc. 1969.

NOBLE B. and DANIEL J.N., "Applied Linear Algebra, "Prentice — Hall, 1977.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

45. ORTEGA J.M., "Numerical Analysis: A second course", Academic Press, 1972.

46. ORTEGA J.M. and POOLE W.G., Jr., "An Introduction to Numerical Methods for Differential Equations", Pitman Publ. Inc., 1981.

47. PARLETT B., "The Symmetric Eigenvalue Problem", Prentice — Hall, 1980.

48. PENNINGTON R.H. "Introductory Computer Methods and Numerical Analysis",Macmillan Company (second edition) 1970.

49. PETTOFREZZO A.J., "Introductory Numerical Analysis", D.C. Heath and Company, 1967.

50. PHILLIPS G.M., and TAYLOR P.J., "Theory and applications of Numerical Analysis", Academic Press, 1973.

51. RALSTON A., "A first course in Numerical Analysis", McGraw — Hill, 1965.

52. RALSTON A., RABINOWITZ P., "A first course in Numerical Analysis", McGraw — Hill, 1978.

53. RECTORYS K., (ed.) "Survey of Applicable Mathematics", Iliffe Books Ltd, 1969.

54. RICE J.R., "The Approximation of Functions", Vol. 1, Addison — Wesley, 1964.

55. RICE J.R., "Numerical Methods, Software, and Analysis", McGraw — Hill, 1983.

56. RIVLIN T.J., "The Chebyshev polynomials", Wiley, 1976.

57. SAWYER W.W., "A first look at Numerical Functional Analysis", Clarendon Press, 1978.

58. SQUIRE W., "Integration for engineers and Scientists", Elsevier, 1970.

59. STARK P.A. "Introduction to Numerical Methods", Macmillan Company, 1970.

60. STEINBERG D.I., "Computational Matrix Algebra", McGraw — Hill, 1974.

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ

61. STEWART G.W., "Introduction to matrix computations", Academic Press, 1973.

62. STOER J. and BULIRSCH R., "Introduction to Numerical Analysis", Springer — Verlag, 1980.

63. STRANG G., "Linear Algebra and its Applications", Academic Press, (second edition), 1980.

64. STROUD A.H., "Numerical Quadrature and Solution of Ordinary Differential Equations", Springer — Verlag, 1974.

65. SZIDAROVSZKY F. and YAKOWITZ S., "Principles and Procedures of Numerical Analysis", Plenum Press, 1978.

66. TODD J., "Basic Numerical Mathematics", Vol. 2 (Numerical Algebra), 1977, Vol. 1 (Numerical Analysis), 1979, Birkhäuser Verlag.

67. TRAUB J.F., "Iterative Methods for the Solution of Equations", Prentice — Hall, inc., 1964.

68. VANDERGRAFT J.S., "Introduction to Numerical Computations", Academic Press, 1978. Second edition 1983.

69. VARGA R.S., "Matrix Iterative Analysis", Prentice — Hall, 1962.

70. WAIT R., "The Numerical Solution of Algebraic Equations", Wiley, 1979.

71. WENDROFF B., "First principles of Numerical Analysis", Addison — Wesley, 1969.

72. WESTLAKE J.R., "A Handbook of Numerical Matrix inversion and solution of Linear equations", John Wiley & Sons, 1968.

73. WILKINSON J.H., "The Algebraic eigenvalue Problem", Clarendon Press, 1965.

74. ΧΑΤΖΗΔΗΜΟΣ Α., "Εισαγωγή στην Αριθμητική Ανάλυση", Ιωάννινα, 1977.

75. ΧΑΤΖΗΔΗΜΟΣ Α., " Αριθμητική Ανάλυση Ι, ΙΙ", Ιωάννινα, 1978,1979.

76. YOUNG D.M., "Iterative solution of Large Linear Systems", Academic Press, 1971.

77. YOUNG D.M. and GREGORY R.T., "A survey of Numerical Mathematics". Volumes I — II , Addison — Wesley, 1972, 1973.

78. ZAGUSKIN V.L., "Handbook of Numerical Methods for the Solution of Algebraic and Transcendental equations", Pergamon Press, 1961.

ΕΚΔΟΣΕΙΣ

Φωτοστοιχειοθεσία Foto Ofsset Γ. ΤΣΟΛΗΣ — Γιάννενα Τηλ. 22.877

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ — ΔΑΠΑΝΗ ΤΗΣ ΠΟΛΙΤΕΙΑΣ

ΠΑΝΕΠΙΣΤΗΜΙΑΚΗ ΒΙΒΛΙΟΘΗΚΗ ΙΩΑΝΝΙΝΩΝ