



ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ
ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ



Μαριλένα Σπυράκη

ΜΕΘΟΔΟΙ ΥΠΟΧΩΡΩΝ ΚΡΥΛΟΝ ΓΙΑ ΤΗΝ ΕΠΙΛΥΣΗ
ΓΡΑΜΜΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΑΤΡΙΒΗ

Ιωάννινα, 2016

Στους γονείς μου

Η παρούσα Μεταπτυχιακή Διατριβή εκπονήθηκε στο πλαίσιο των σπουδών για την απόκτηση του Μεταπτυχιακού Διπλώματος Ειδίκευσης στα Εφαρμοσμένα Μαθηματικά και Πληροφορική που απονέμει το Τμήμα Μαθηματικών του Πανεπιστημίου Ιωαννίνων.

Εγκρίθηκε την 14/12/2016 από την εξεταστική επιτροπή:

| | |
|----------------------|-----------------------|
| Ονοματεπώνυμο | Βαθμίδα |
| Νούτσος Δημήτριος | Καθηγητής |
| Αχρίβης Γεώργιος | Καθηγητής |
| Χωρίκης Θεόδωρος | Αναπληρωτής Καθηγητής |

ΥΠΕΥΘΥΝΗ ΔΗΛΩΣΗ

“Δηλώνω υπεύθυνα ότι η παρούσα διατριβή εκπονήθηκε κάτω από τους διεθνείς ηθικούς και ακαδημαϊκούς κανόνες δεοντολογίας και προστασίας της πνευματικής ιδιοκτησίας. Σύμφωνα με τους κανόνες αυτούς, δεν έχω προβεί σε ιδιοποίηση ξένου επιστημονικού έργου και έχω πλήρως αναφέρει τις πηγές που χρησιμοποίησα στην εργασία αυτή.”

Μαριλένα Σπυράκη

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω πρώτα από όλους τον επιβλέποντα καθηγητή μου Κύριο Δημήτριο Νούτσο για την εμπιστοσύνη που μου έδειξε, τη στήριξη και την καθοδήγησή του στην προσπάθειά μου. Ακόμη, θα ήθελα να ευχαριστήσω θερμά τον Κύριο Γεώργιο Ακρίβη για τις πολύ εύστοχες παρατηρήσεις του και τη συμβολή του στην τελική διαμόρφωση του κειμένου της εργασίας μου. Ευχαριστώ επίσης τον Κύριο Θεόδωρο Χωρίκη για τις παρατηρήσεις του. Ευχαριστώ όμως και όλους ανεξαιρέτως τους καθηγητές του τμήματος Μαθηματικών του Πανεπιστημίου Ιωαννίνων, τους οποίους είχα την τιμή να γνωρίσω φοιτώντας σε αυτό. Τέλος, τους γονείς μου, που μου συμπαραστάθηκαν και με βοήθησαν με κάθε τρόπο σε κάθε μου βήμα μέχρι σήμερα.

Η παρούσα διατριβή αναφέρεται σε αριθμητικές μεθόδους επίλυσης γραμμικών αλγεβρικών συστημάτων. Αρχικά, γίνεται μια σύντομη παρουσίαση εισαγωγικών εννοιών και κάποιων χρήσιμων στοιχείων Γραμμικής Άλγεβρας, καθώς και Αριθμητικής Γραμμικής Άλγεβρας. Δίνονται ορισμοί εσωτερικών γινομένων, νορμών, πινάκων ειδικής μορφής, ακολουθιών διανυσμάτων και πινάκων και συνθηκών σύγκλισής τους. Στη συνέχεια καταγράφονται, για την πληρότητα του θέματος, οι Απλές Επαναληπτικές Μέθοδοι, που βασίζονται στην προσέγγιση της λύσης από μια ακολουθία διανυσμάτων που παράγεται από μια αναδρομική επαναληπτική διαδικασία. Οι πιο γνωστές είναι, γνωστές ως κλασικές επαναληπτικές μέθοδοι, οι μέθοδοι Jacobi, Gauss - Seidel, SOR και SSOR. Στο κύριο μέρος της εργασίας μελετώνται και παρουσιάζονται οι επαναληπτικές μέθοδοι ελαχιστοποίησης ή μέθοδοι που βασίζονται σε υποχώρους Krylov. Οι μέθοδοι αυτές αναφέρονται με τη χρονολογική σειρά με την οποία εισήχθησαν. Αρχίζουμε με τις μεθόδους Orlthomin(1) και Απότομης Καθόδου - Steepest Descent. Γι' αυτές παρουσιάζεται η βασική μαθηματική θεωρία και δίνονται εκτιμήσεις για τις νόρμες διανυσμάτων υπόλοιπου και σφάλματος. Στη συνέχεια, παρουσιάζονται μέθοδοι πιο εξελιγμένες ως προς την πολυπλοκότητα, τις απαιτήσεις μνήμης και την ευστάθεια. Αυτές είναι μέθοδοι όπως η Μέθοδος Συζυγών Κλίσεων - Conjugate Gradient και η Γενικευμένη Μέθοδος Ελαχίστου Υπολοίπου - Generalized Minimal Residual. Βασίζονται στην ορθογωνοποίηση ενός υποχώρου Krylov. Σε κάθε μέθοδο τονίζονται τα πλεονεκτήματα και τα μειονεκτήματα, καθώς και η ανάγκη για περαιτέρω εξέλιξη. Τελικά, παρουσιάζονται παραλλαγές των μεθόδων αυτών που βασίζονται σε ένα είδος διορθογωνοποίησης του υποχώρου Krylov (κατασκευή δύο συνόλων διανυσμάτων των οποίων τα στοιχεία του ενός είναι ορθογώνια ως προς τα στοιχεία του άλλου). Αυτές είναι η μέθοδος Ημιαχίστου Υπολοίπου - QMR, η Τετραγωνική Μέθοδος Συζυγών Κλίσεων - CGS και η Ευσταθειοποιημένη Μέθοδος Δισυζυγών Κλίσεων - BiCGSTAB.

ABSTRACT

This thesis is concerned with numerical methods for the solution of linear algebraic systems. First, we present introductory concepts and some useful results from Linear Algebra and Numerical Linear Algebra. Definitions of inner products, norms, some special kinds of matrices, vector sequences and sequences of matrices, as well as their convergence properties. Next, for completeness, we present the Iterative Methods that yield approximations to the solution, i.e. a sequence of vectors generated from a recursive iterative process. The most well known methods, referred to as classical iterative methods, are the Jacobi, Gauss - Seidel, SOR and SSOR methods. In the main part of the thesis, we study the iterative minimization methods or methods based on Krylov subspaces. The methods we study, are listed in chronological order in which they were introduced. We begin with the Orthomin(1) and Steepest Descent methods. For these methods we present the basic mathematical theory and provide estimates for the norms of the residual - vector and the error - vector. Moreover, we present improved methods as regards complexity, memory requirements, as well as stability. These are the Conjugate Gradient and Generalized Minimal Residual methods. The latter methods are based on the orthogonalization of a Krylov subspace. In each method we describe, we point out the advantages and disadvantages, as well as the necessity of further development. Finally, we present modifications of these methods relying on a kind of biorthogonalization of a Krylov subspace (construction of two sets of vectors, where the elements of the one set are orthogonal to the elements of the other). These are the Quasi Minimal Residual (QMR), Conjugate Gradient Squared (CGS) and BiConjugate Gradient Stabilized (BiCGSTAB) methods.

ΠΕΡΙΕΧΟΜΕΝΑ

| | |
|---|-----------|
| Περίληψη | i |
| Abstract | ii |
| 1 Εισαγωγή | 1 |
| 1.1 Εισαγωγικές Έννοιες | 1 |
| 1.2 Κλασικές Επαναληπτικές Μέθοδοι | 14 |
| 1.3 Σύγκριση της Απλής Επαναληπτικής Μεθόδου | 16 |
| 2 Μέθοδοι Ορθογωνοποίησης | 19 |
| 2.1 Μέθοδος Orthomin(1) - Μέθοδος Απότομης Καθόδου (Steepest Descent - SD) | 19 |
| 2.2 Μέθοδος Γενικών Διευθύνσεων (Generalized Direction - GD) | 24 |
| 2.3 Γενικευμένη Μέθοδος Συζυγών Υπολοίπων (Generalized Conjugate Residual - GCR) - Μέθοδος Συζυγών Διευθύνσεων (Conjugate Direction - CD) Μέθοδος Συζυγών Κλίσεων (Conjugate Gradient - CG) | 26 |
| 2.4 Γενικευμένη Μέθοδος Ελαχίστου Υπολοίπου (Generalized Minimal Residual Method - GMRES) | 38 |
| 3 Μέθοδοι Διορθογωνοποίησης | 53 |
| 3.1 Μέθοδος Δισυζυγών Κλίσεων (Biconjugate Gradient - BiCG) | 53 |
| 3.1.1 Μέθοδος Δισυζυγών Κλίσεων - BiConjugate Gradient | 57 |
| 3.2 Μέθοδος Ημιελαχίστου Υπολοίπου (Quasi Minimal Residual - QMR) | 59 |
| 3.3 Τετραγωνική Μέθοδος Συζυγών Κλίσεων (Conjugate Gradient Squared - CGS) | 61 |

| | | |
|-----|--|-----------|
| 3.4 | Ευσταθειοποιημένη Μέθοδος Δισυζυγών Κλίσεων (BiConjugate Gradient Stabilized - BiCGSTAB) | 64 |
| | Επίλογος | 69 |
| | Βιβλιογραφία | 71 |

ΚΕΦΑΛΑΙΟ 1

Εισαγωγή

Στόχος μας είναι η επίλυση αλγεβρικών γραμμικών συστημάτων $Ax = b$, με $A \in \mathbb{C}^{n,n}$, $\det A \neq 0$ και $b \in \mathbb{C}^n$. Η λύση του συστήματος είναι ένα διάνυσμα $x \in \mathbb{C}^n$. Για την επίλυση του συστήματος αυτού προτείνονται διάφορες μέθοδοι. Όλες οι μέθοδοι που παρουσιάζονται στη συνέχεια χαρακτηρίζονται ως επαναληπτικές μέθοδοι ελαχιστοποίησης. Επαναληπτικές, γιατί η ζητούμενη λύση $x \in \mathbb{C}^n$ προκύπτει από μια διαδικασία υπολογισμού διαδοχικών προσεγγίσεων $\{x^{(m)}\} = \{x^{(0)}, x^{(1)}, x^{(2)}, \dots, x^{(m)}\}$, $m \in \mathbb{N}$. Ως αρχική προσέγγιση επιλέγεται συνήθως η $x^{(0)} = 0$. Ως επόμενες προσεγγίσεις οι $x^{(1)} \in \text{span}\{b\}$, $x^{(2)} \in \text{span}\{b, Ab\}$, $x^{(3)} \in \text{span}\{b, Ab, A^2b\}$ κ.ο.κ. Τέλος, ως τελική προσέγγιση επιλέγεται η $x = x^{(m)} \in \text{span}\{b, Ab, A^2b, \dots, A^{m-1}b\}$. Ένας υπόχωρος του \mathbb{C}^n που περιέχει αυτές τις διαδοχικές προσεγγίσεις ονομάζεται χώρος Krylov διάστασης m .

1.1 Εισαγωγικές Έννοιες

Ορισμός 1.1.1 (Υπόχωρος Krylov). *Υπόχωρος Krylov διάστασης m του \mathbb{C}^n είναι ο χώρος της μορφής $K_m(A, b) = \text{span}\{b, Ab, A^2b, \dots, A^{m-1}b\}$. Συμβολίζουμε συχνά από εδώ και στο εξής τον $K_m(A, b)$ απλά ως K_m και έχουμε κατά νου ότι αναφερόμαστε στον υπόχωρο Krylov που αντιστοιχεί στο αλγεβρικό γραμμικό σύστημα $Ax = b$.*

Οι μέθοδοι που θα παρουσιαστούν χαρακτηρίζονται ακόμη ως μέθοδοι ελαχιστοποίησης, γιατί κατά την υλοποίησή τους σε κάθε βήμα m προσπαθούν να ελαχιστοποιήσουν την Ευκλείδεια νόρμα του διανύσματος - υπόλοιπο ($r^{(m)} = b - Ax^{(m)}$) ή την νόρμα ενέργειας του διανύσματος - σφάλμα ($\varepsilon^{(m)} = A^{-1}b - x^{(m)}$) πάνω σε ένα χώρο Krylov της μορφής

$$K_m(A, r^{(0)}) = \text{span}\{r^{(0)}, Ar^{(0)}, A^2r^{(0)}, \dots, A^{m-1}r^{(0)}\},$$

όπου $r^{(0)} = b - Ax^{(0)}$ το διάνυσμα - υπόλοιπο της αρχικής προσέγγισης $x^{(0)}$.

Ορισμός 1.1.2 (Μέθοδος Υποχώρου Krylov). Η μέθοδος υπόχωρου Krylov είναι μια επαναληπτική μέθοδος ελαχιστοποίησης με K_m να είναι ο υπόχωρος Krylov $K_m(A, r^{(0)}) = \text{span}\{r^{(0)}, Ar^{(0)}, A^2r^{(0)}, \dots, A^{m-1}r^{(0)}\}$, όπου $r^{(0)} = b - Ax^{(0)}$. Συμβολίζουμε σε αντιστοιχία με τον προηγούμενο ορισμό τον $K_m(A, r^{(0)})$ απλά ως K_m .

Η βασική ιδιότητα που ισχύει σε έναν υπόχωρο Krylov K_m του \mathbb{C}^n είναι ότι κάθε διάνυσμα $x \in \mathbb{C}^n$ του χώρου αυτού μπορεί να γραφεί στη μορφή $x = p(A)b$, όπου με p συμβολίζουμε ένα πολυώνυμο βαθμού το πολύ $m - 1$.

Ορισμός 1.1.3 (Ελάχιστο πολυώνυμο). Ελάχιστο πολυώνυμο (*minimal polynomial*) p ενός διανύσματος $b \in \mathbb{C}^n$ που αντιστοιχεί στον πίνακα $A \in \mathbb{C}^{n,n}$ είναι το μη μηδενικό μονικό (δηλαδή με συντελεστή μεγιστοβαθμίου όρου 1) πολυώνυμο ελάχιστου βαθμού για το οποίο ισχύει $p(A)b = 0$. Ο βαθμός (*degree*) του ελάχιστου πολυωνύμου του διανύσματος b ως προς τον πίνακα A , από εδώ και στο εξής θα καλείται απλά βαθμός του b και θα συμβολίζεται με μ .

Έχει νόημα να αναζητήσουμε την ακριβή λύση $x = A^{-1}b$ του συστήματος (ή έστω μια καλή προσεγγιστική λύση) σε έναν υπόχωρο Krylov K_m , αφού προκύπτει ότι πάντα ένας τέτοιος χώρος θα την περιέχει. Πράγματι, έστω το σύστημα $Ax = b$, $b \in \mathbb{C}^n$ οποιοδήποτε και $x^{(0)} = 0$ η αρχική προσέγγιση. Τότε, αρκεί να υπάρχει πολυώνυμο βαθμού το πολύ $m - 1$ με $p_{m-1}(A)b = A^{-1}b$, δηλαδή ο $A^{-1}b$ να γράφεται ως γραμμικός συνδυασμός των διανυσμάτων που ορίζουν τον $K_m(A, b)$. Το πολυώνυμο r βαθμού $t \leq n$ για το οποίο $r(A) = (\alpha_0 I + \alpha_1 A + \alpha_2 A^2 + \dots + \alpha_t A^t)b = 0$, $\alpha_0 \neq 0$, είναι το ελάχιστο πολυώνυμο του b ως προς τον πίνακα A . Τότε αν πολλαπλασιάσουμε τη σχέση αυτή επί A^{-1} παίρνουμε $x = A^{-1}b = \frac{-1}{\alpha_0} \sum_{i=0}^{t-1} \alpha_{i+1} A^i b$. Παρακάτω δίνεται και το σχετικό Θεώρημα.

Θεώρημα 1.1.4. Αν το ελάχιστο πολυώνυμο ενός αντιστρέψιμου πίνακα $A \in \mathbb{C}^{n,n}$ είναι βαθμού $t \leq n$, τότε η λύση $x = A^{-1}b$ του συστήματος $Ax = b$ περιέχεται σε έναν υπόχωρο Krylov $K_t(A, r^{(0)})$ (ή απλά K_t).

Πρόταση 1.1.5 (Διάσταση ενός υπόχωρου Krylov). Ο υπόχωρος Krylov K_m είναι διάστασης m , αν και μόνο αν ο βαθμός μ του ελάχιστου πολυωνύμου p είναι μεγαλύτερος ή ίσος του m , δηλαδή $\dim(K_m) = m \iff \mu \geq m$, συνεπώς $\dim(K_m) = \min\{m, \mu\}$.

Απόδειξη. Έστω ο m - διάστατος υπόχωρος Krylov

$$K_m = \{b, Ab, A^2b, \dots, A^{m-1}b\}.$$

Τα διανύσματα $b, Ab, A^2b, \dots, A^{m-1}b$ αποτελούν βάση του K_m , άρα είναι γραμμικά ανεξάρτητα. Δηλαδή για κάποιο σύνολο m πραγματικών αριθμών $\alpha_i, i = 0, 1, \dots, m-1$, με τουλάχιστον ένα από τα $\alpha_i \neq 0$, έπεται ότι ο γραμμικός συνδυασμός $\sum_{i=0}^{m-1} \alpha_i A^i b = p_{m-1}(A)b$ είναι διάφορος του μηδενικού διανύσματος. Ισοδύναμα, το μοναδικό πολυώνυμο βαθμού $\leq m-1$ με $p(A)b = 0$ είναι το μηδενικό πολυώνυμο. Επομένως, $\mu \geq m$.

Έστω $\mu < m$ και $p_\mu(A)$ το ελάχιστο πολυώνυμο του b . Τότε, για $y \in K_m$ υπάρχει πολυώνυμο q_{m-1} το πολύ $m-1$ βαθμού τέτοιο ώστε $y = q_{m-1}(A)b$. Συνεπώς,

$$\begin{aligned} q_{m-1}(A)b &= (s_{m-\mu-1}(A)p_\mu(A) + r_{\mu-1}(A))b \\ &= s_{m-\mu-1}(A)p_\mu(A)b + r_{\mu-1}(A)b \\ &= r_{\mu-1}(A)b \in K_\mu, \end{aligned}$$

όπου s είναι το πολυώνυμο - πηλίκο της διαίρεσης $q_m(A)/p_m(A)$ βαθμού το πολύ $m-\mu-1$ και r το αντίστοιχο υπόλοιπο βαθμού το πολύ $\mu-1$ αντίστοιχα. Αυτό σημαίνει ότι $\dim(K_m) = \mu$. Άρα τελικά, $\dim(K_m) = \min\{m, \mu\}$. \square

Είναι σκόπιμο, για καλύτερη κατανόηση, να δοθούν στη συνέχεια κάποιοι βασικοί ορισμοί. Θα δοθούν ορισμοί διαφόρων τύπων πινάκων, εσωτερικού γινομένου, νόρμας διανύσματος και νόρμας πίνακα, φασματικής ακτίνας και δείκτη κατάστασης πίνακα, σύγκλισης ακολουθίας διανυσμάτων και πινάκων, καθώς και κάποιες χρήσιμες ιδιότητες πινάκων που θα χρησιμοποιηθούν στη συνέχεια. Οι βασικές παραγοντοποιήσεις πινάκων που χρησιμοποιούνται συχνά στην Αριθμητική Γραμμική Άλγεβρα και χρειάζονται σε κάποιους από τους αλγορίθμους που θα παρουσιαστούν, θα αναφέρονται όπου κρίνεται αναγκαίο.

Ορισμός 1.1.6. Για κάθε $A \in \mathbb{C}^{n,n}$ υπάρχουν αριθμοί $\lambda_i \in \mathbb{C}$ και διανύσματα $x_i \in \mathbb{C}^n \setminus \{0\}, i = 1(1)n$, τέτοια ώστε $Ax_i = \lambda_i x_i, i = 1(1)n$. Οι λ_i λέγονται ιδιοτιμές και τα x_i αντίστοιχα ιδιοδιανύσματα.

Συμπέρασμα. Εξ ορισμού ένας πίνακας $A \in \mathbb{C}^{n,n}$ έχει ως ιδιοτιμή την τιμή $\lambda \in \mathbb{C}$ και ως ιδιοδιάνυσμα το μη μηδενικό διάνυσμα $x \in \mathbb{C}^n \neq 0$ αν και μόνο αν $Ax = \lambda x$. Έτσι,

$$Ax = \lambda x \iff Ax - \lambda x = 0 \iff (A - \lambda I)x = 0,$$

δηλαδή ο $A - \lambda I$ έχει μηδενική ορίζουσα ($\det(A - \lambda I) = 0$, όπου $\chi(A) := A - \lambda I$, $\chi \in \mathbb{P}_n$ είναι το χαρακτηριστικό πολυώνυμο του πίνακα A). Οπότε, πρακτικά, για τον προσδιορισμό των ιδιοτιμών ενός πίνακα $A \in \mathbb{C}^{n,n}$ αρκεί να λύσουμε την εξίσωση $\det(A - \lambda I) = 0$. Παρατηρούμε ότι οι ρίζες αυτής της εξίσωσης είναι γενικά μιγαδικές για οποιονδήποτε μιγαδικό πίνακα A , μάλιστα μπορεί να υπάρχουν κάποιες συζυγείς μιγαδικές ρίζες και στην περίπτωση πραγματικού πίνακα A .

Παρατήρηση. Το φάσμα των ιδιοτιμών ενός πίνακα $A \in \mathbb{C}^{n,n}$, δηλαδή το σύνολο $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ θα το συμβολίζουμε με $\sigma(A)$ και τη φασματική ακτίνα, δηλαδή το $\max_{i=1(1)n} |\lambda_i|$ θα τη συμβολίζουμε με $\rho(A)$.

Ορισμός 1.1.7. Ευκλείδειο εσωτερικό γινόμενο δύο διανυσμάτων $x, y \in \mathbb{C}^n$ ονομάζεται το

$$(x, y)_2 = \sum_{i=1}^n \bar{x}_i y_i.$$

και συμβολίζεται με $(x, y)_2$ ή απλά ως (x, y) .

Προφανώς ισχύει εξ' ορισμού ότι $(x, x) \geq 0$. Η ποσότητα (x, x) ισούται με $\|x\|_2^2$, όπου $\|x\|_2$ (ή απλά $\|x\|$) είναι η Ευκλείδεια νόρμα ή στάθμη του διανύσματος $x \in \mathbb{C}^n$. Ακόμη, είναι εύκολο να παρατηρήσει κανείς ότι $\overline{(x, y)_2} = \sum_{i=1}^n \bar{\bar{x}_i y_i} = (y, x)_2$.

Ορισμός 1.1.8. Ένα διάνυσμα $x \in \mathbb{C}^n$ λέγεται κανονικοποιημένο ως προς μία νόρμα αν και μόνο αν έχει νόρμα ίση με 1.

Ορισμός 1.1.9. Ένας πίνακας $A \in \mathbb{C}^{n,n}$ λέγεται αντιστρέψιμος, αν υπάρχει ο αντίστροφός του A^{-1} . Ισοδύναμα, ένας πίνακας $A \in \mathbb{C}^{n,n}$ είναι αντιστρέψιμος αν και μόνο αν $\det A \neq 0$. Τότε, ισχύει η σχέση $A^{-1}A = I_n = AA^{-1}$, όπου με I_n συμβολίζουμε τον μοναδιαίο πίνακα διάστασης n .

Σχόλιο. Άλλα κριτήρια που χαρακτηρίζουν έναν πίνακα $A \in \mathbb{C}^{n,n}$ αντιστρέψιμο, είναι το να αποτελείται από n γραμμικά ανεξάρτητα διανύσματα στήλες ή n γραμμικά ανεξάρτητα διανύσματα γραμμές. Ακόμη, το αλγεβρικό γραμμικό σύστημα $Ax = b$ να έχει μοναδική λύση, την $x = A^{-1}b$, ή το αντίστοιχο ομογενές σύστημα $Ax = 0$ να έχει μόνο τη μηδενική λύση, δηλαδή τη $x = 0$. Γι αυτό, για να έχει νόημα να αναζητήσουμε τη λύση x του συστήματος $Ax = b$ υποθέτουμε εξ' αρχής ότι ο πίνακας των συντελεστών A είναι αντιστρέψιμος.

Ορισμός 1.1.10. Ένας πίνακας $A \in \mathbb{C}^{n,n}$ λέγεται ερμιτιανός, αν $a_{ij} = \bar{a}_{ji}$, όπου με a_{ij} συμβολίζουμε το στοιχείο που βρίσκεται στη θέση (i, j) του πίνακα A και με \bar{a}_{ji} το μιγαδικό συζυγές του στοιχείου που βρίσκεται στη θέση (j, i) του πίνακα A .

Ορισμός 1.1.11. Ένας πίνακας $A \in \mathbb{R}^{n,n}$ λέγεται *συμμετρικός*, αν $a_{ij} = a_{ji}$.

Παρατήρηση. Ένας πίνακας $A \in \mathbb{C}^{n,n}$ είναι *ερμιτιανός*, αν και μόνο αν ισχύει $A = A^H$, όπου $A^H = \overline{A}^T$. Ο $A \in \mathbb{R}^{n,n}$ είναι *συμμετρικός*, αν και μόνο αν $A = A^T$. Δηλαδή, το αντίστοιχο του ερμιτιανού πίνακα στην πραγματική περίπτωση είναι ο *συμμετρικός* πίνακας.

Σχόλιο. Τα στοιχεία της κύριας διαγωνίου του a_{ii} , $i = 1(1)n$ ενός ερμιτιανού πίνακα $A \in \mathbb{C}^{n,n}$, λόγω του τρόπου ορισμού του, μπορούμε να δούμε ότι έχουν την ιδιότητα να είναι πραγματικά, αφού κάθε ένα από αυτά είναι ίσο με το συζυγές του. Ακόμη, αποδεικνύεται ότι ένας ερμιτιανός πίνακας έχει n το πλήθος πραγματικές ιδιοτιμές και n γραμμικά ανεξάρτητα ιδιοδιανύσματα τα οποία μπορούν να ληφθούν ορθοκανονικά, δηλαδή: $(x_j, x_k) = \delta_{jk}$, όπου δ_{jk} το σύμβολο του Kronecker με

$$\delta_{jk} = \begin{cases} 1, & j = k \\ 0, & j \neq k \end{cases}$$

Ορισμός 1.1.12. Έστω $A \in \mathbb{C}^{n,n}$ με $A = A^H$. Ο A λέγεται *θετικά ορισμένος*, αν και μόνο αν $\forall x \in \mathbb{C}^n \setminus \{0\}$ ισχύει $(x, Ax)_2 = (Ax, x)_2 > 0$.

Ορισμός 1.1.13. Έστω $A \in \mathbb{C}^{n,n}$ με $A = A^H$. Ο A λέγεται *θετικά ημιορισμένος*, αν και μόνο αν $\forall x \in \mathbb{C}^n \setminus \{0\}$ ισχύει $(x, Ax)_2 = (Ax, x)_2 \geq 0$.

Αντίστοιχοι είναι οι ορισμοί για θετικά ορισμένο και θετικά ημιορισμένο πίνακα στην περίπτωση πίνακα $A \in \mathbb{R}^{n,n}$.

Ορισμός 1.1.14. Έστω $A \in \mathbb{C}^{n,n}$. Ο A λέγεται *αυστηρά διαγώνια υπέρτερος* κατά γραμμές, αν και μόνο αν $|\alpha_{ii}| > \sum_{j=1, j \neq i}^n |\alpha_{ij}|$, $i = 1(1)n$. (Αντίστοιχα, ο A λέγεται *αυστηρά διαγώνια υπέρτερος* κατά στήλες, αν και μόνο αν $|\alpha_{ii}| > \sum_{i=1, i \neq j}^n |\alpha_{ij}|$, $j = 1(1)n$.)

Σχόλιο. Αποδεικνύεται ότι ένας αυστηρά διαγώνια υπέρτερος πίνακας κατά γραμμές ή κατά στήλες είναι αναγκαστικά αντιστρέψιμος.

Ορισμός 1.1.15. Ένας πίνακας $A \in \mathbb{C}^{n,n}$ λέγεται *θετικός*, αν και μόνο αν $a_{ij} > 0$, $\forall i, j = 1(1)n$.

Σχόλιο. Αντίστοιχοι είναι οι ορισμοί για μηδενικό και αρνητικό πίνακα, αρκεί $a_{ij} = 0$, $\forall i, j = 1(1)n$ ή αντίστοιχα $a_{ij} < 0$, $\forall i, j = 1(1)n$.

Ορισμός 1.1.16. Ένας πίνακας $Q \in \mathbb{C}^{n,n}$ λέγεται *ορθομοναδιαίος*, αν και μόνο αν $Q^H Q = I = Q Q^H$.

Ορισμός 1.1.17. Ένας πίνακας $Q \in \mathbb{R}^{n,n}$ λέγεται ορθογώνιος, αν και μόνο αν $Q^T Q = I = Q Q^T$.

Ορισμός 1.1.18. Δύο διανύσματα $x, y \in \mathbb{C}^n \setminus \{0\}$ λέγονται ορθογώνια, αν και μόνο αν $(x, y) = 0$.

Σχόλιο. Αν στον προηγούμενο ορισμό ισχύει επιπλέον ότι $\|x\| = \|y\| = 1$, τα διανύσματα λέγονται ορθοκανονικά.

Αντίστοιχος είναι ο ορισμός για διανύσματα $x, y \in \mathbb{R}^n$.

Ορισμός 1.1.19. Έστω $A \in \mathbb{C}^{n,n} \setminus \{0\}$ ένας ερμιτιανός πίνακας. Δύο διανύσματα $x, y \in \mathbb{C}^n$ λέγονται *A-ορθογώνια* ή *A-συζυγή* (*A - conjugate*), αν και μόνο αν $(Ax, y) = 0$.

Βάσει των δύο παραπάνω ορισμών παρατηρούμε ότι η *A* - ορθογωνιότητα για τον πίνακα $A = I$ δεν είναι άλλη από την γνωστή ορθογωνιότητα δύο διανυμάτων.

Ορισμός 1.1.20. Ο $A \in \mathbb{C}^{n,n}$ λέγεται διαγωνιοποιήσιμος, αν ο $P^{-1}AP$ είναι διαγώνιος πίνακας, για κάποιον αντιστρέψιμο πίνακα P .

Θεώρημα 1.1.21. Έστω $A \in \mathbb{C}^{n,n}$. Τα επόμενα είναι ισοδύναμα:

- α) Ο A είναι διαγωνιοποιήσιμος.
- β) Ο A έχει n γραμμικά ανεξάρτητα ιδιοδιανύσματα.

Απόδειξη. ((α) \implies (β)) Έστω A διαγωνιοποιήσιμος. Τότε, υπάρχει αντιστρέψιμος πίνακας P τέτοιος ώστε: $P^{-1}AP = D$, όπου $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$. Πολλαπλασιάζοντας από τα αριστερά με τον P έχουμε $AP = PD$. Αν με p_1, \dots, p_n συμβολίσουμε τα διανύσματα στήλες του πίνακα P τότε έχουμε ότι

$$Ap_1 = \lambda_1 p_1, \dots, Ap_n = \lambda_n p_n.$$

Αφού ο P είναι αντιστρέψιμος τα p_1, \dots, p_n είναι μη μηδενικά και γραμμικά ανεξάρτητα, άρα οι $\lambda_1, \lambda_2, \dots, \lambda_n$ είναι οι ιδιοτιμές και τα p_1, \dots, p_n τα αντίστοιχα ιδιοδιανύσματα του πίνακα A . Έτσι ο A έχει n γραμμικά ανεξάρτητα ιδιοδιανύσματα.

((β) \implies (α)) Έστω ότι ο A έχει n γραμμικά ανεξάρτητα ιδιοδιανύσματα p_1, \dots, p_n που αντιστοιχούν στις ιδιοτιμές $\lambda_1, \lambda_2, \dots, \lambda_n$. Έστω P ο πίνακας με διανύσματα στήλες αυτά τα ιδιοδιανύσματα. Αν $A \in \mathbb{C}^{n,n}$, τότε ο AP έχει ως διανύσματα στήλες τα Ap_1, \dots, Ap_n . Άρα

$$AP = [Ap_1, \dots, Ap_n] = [\lambda_1 p_1, \dots, \lambda_n p_n] = PD,$$

όπου $D = \text{diag}(\lambda_1, \dots, \lambda_n)$. Αφού έχουμε υποθέσει ότι τα p_1, \dots, p_n είναι γραμμικά ανεξάρτητα, ο πίνακας P είναι αντιστρέψιμος. Έτσι,

$$AP = PD \implies P^{-1}AP = D,$$

δηλαδή ο A είναι διαγωνιοποιήσιμος. \square

Σχόλιο. Αν ένας πίνακας A είναι διαγωνιοποιήσιμος, τότε μπορούμε να υπολογίσουμε διάφορες δυνάμεις του πίνακα (δηλαδή τον πίνακα A^k) με τον εξής απλό τρόπο. Παρατηρούμε ότι

$$(P^{-1}AP)^2 = (P^{-1}AP)(P^{-1}AP) = P^{-1}AIP = P^{-1}A^2P.$$

Οπότε, για κάποιο ακέραιο $k > 0$ παρατηρούμε ότι με απλή επαγωγή έχουμε

$$(P^{-1}AP)^k = (P^{-1}AP) \dots (P^{-1}AP) = P^{-1}A^kP.$$

Επομένως,

$$P^{-1}A^kP = (P^{-1}AP)^k = D^k \iff A^kP = PD^k$$

(πολλαπλασιάζοντας και τα δύο μέλη από αριστερά με τον πίνακα P) $\iff A^k = PD^kP^{-1}$ (πολλαπλασιάζοντας και τα δύο μέλη από δεξιά με τον πίνακα P^{-1}). Ο υπολογισμός του D^k είναι πολύ απλός λόγω της μορφής του D . Στην περίπτωση αυτή, η πιο δαπανηρή διαδικασία είναι ο προσδιορισμός των πινάκων D και P , ο προσδιορισμός δηλαδή όλων των ιδιοτιμών και των ιδιοδιανυσμάτων του πίνακα A .

Ορισμός 1.1.22. Δύο πίνακες $A \in \mathbb{C}^{n,n}$ και $B \in \mathbb{C}^{n,n}$ λέγονται όμοιοι, αν και μόνο αν υπάρχει αντιστρέψιμος πίνακας $P \in \mathbb{C}^{n,n}$ τέτοιος ώστε $P^{-1}AP = B$.

Για δύο όμοιους πίνακες ισχύουν οι ακόλουθες ιδιότητες:

Λήμμα 1.1.23. Αν οι πίνακες $A, B \in \mathbb{C}^{n,n}$ είναι όμοιοι, τότε ισχύει $\sigma(A) = \sigma(B)$.

Απόδειξη. Η απόδειξη είναι προφανής, αν παρατηρήσει κανείς ότι βάσει του ορισμού των όμοιων πινάκων ισχύει $\chi_A(\lambda) = \chi_{P^{-1}AP}(\lambda)$, αφού

$$\begin{aligned} \chi_{P^{-1}AP}(\lambda) &= \det(\lambda I - P^{-1}AP) \\ &= \det(P^{-1}(\lambda I - A)P) \\ &= \det(P^{-1}) \det(\lambda I - A) \det(P) \\ &= \frac{1}{\det(P)} \det(\lambda I - A) \det(P) \\ &= \det(\lambda I - A) = \chi_A(\lambda). \end{aligned}$$

Άρα δύο όμοιοι πίνακες θα έχουν τις ίδιες ιδιοτιμές (οι οποίες προκύπτουν ως λύσεις δύο χαρακτηριστικών εξισώσεων που αντιστοιχούν στο ίδιο χαρακτηριστικό πολυώνυμο). \square

Ακόμη, ισχύει το επόμενο θεώρημα.

Θεώρημα 1.1.24. *Αν οι πίνακες $A, B \in \mathbb{C}^{n,n}$ είναι όμοιοι, τότε ισχύει $\det(A) = \det(B)$.*

Απόδειξη. Είναι $A, B \in \mathbb{C}^{n,n}$ όμοιοι. Άρα υπάρχει $P \in \mathbb{C}^{n,n}$ με

$$\begin{aligned} P^{-1}AP = B &\implies \det(P^{-1}AP) = \det(B) \\ &\iff \det(P^{-1}) \det(A) \det(P) = \det(B) \\ &\iff \frac{1}{\det(P)} \det(A) \det(P) = \det(B) \\ &\iff \det(A) = \det(B). \end{aligned} \quad \square$$

Μπορούμε να πούμε έτσι, ότι γενικά δύο όμοιοι πίνακες έχουν τα ίδια χαρακτηριστικά μεγέθη (ορίζουσα, ιδιοτιμές). Το αντίστροφο δεν ισχύει, δηλαδή αν λόγου χάριν για δύο τετραγωνικούς πίνακες $A, B \in \mathbb{C}^{n,n}$ ισχύει $\det(A) = \det(B)$, οι A και B δεν είναι αναγκαστικά όμοιοι.

Ορισμός 1.1.25. Ένας πίνακας $A \in \mathbb{C}^{n,n}$ λέγεται *άνω τριγωνικός*, αν και μόνο αν $\alpha_{ij} = 0, \forall i > j$. Δηλαδή, τα στοιχεία που βρίσκονται κάτω από την κύρια διαγώνιο είναι ίσα με 0.

Ορισμός 1.1.26. Ένας πίνακας $A \in \mathbb{C}^{n,n}$ λέγεται *κάτω τριγωνικός*, αν και μόνο αν $\alpha_{ij} = 0, \forall i < j$. Δηλαδή, τα στοιχεία που βρίσκονται πάνω από την κύρια διαγώνιο είναι ίσα με 0.

Ορισμός 1.1.27. Ένας πίνακας $A \in \mathbb{C}^{n,n}$ λέγεται *άνω Hessenberg*, αν και μόνο αν $\alpha_{ij} = 0, \forall i > j + 1$.

Ορισμός 1.1.28. Ένας πίνακας $A \in \mathbb{C}^{n,n}$ λέγεται *κάτω Hessenberg* αν και μόνο αν $\alpha_{ij} = 0, \forall i < j + 1$.

Ορισμός 1.1.29. Νόρμα $\|\cdot\| : \mathbb{C}^n \rightarrow \mathbb{R}^{0,+}$ είναι μια απεικόνιση η οποία ικανοποιεί τις επόμενες τρεις ιδιότητες:

- i) $\|x\| \geq 0$ και $\|x\| = 0$ αν και μόνο αν $x = 0, \forall x \in \mathbb{C}^n$
- ii) $\|cx\| = |c|\|x\|, \forall c \in \mathbb{C}$ και $\forall x \in \mathbb{C}^n$

$$\text{iii) } \|x + y\| \leq \|x\| + \|y\|, \forall x, y \in \mathbb{C}^n.$$

Ορισμός 1.1.30. Βαθμίδα (rank) ενός πίνακα $A \in \mathbb{C}^{n,n}$ ονομάζεται το μέγιστο πλήθος των γραμμικά ανεξάρτητων γραμμών ή στηλών του.

Παρατήρηση. Μια διανυσματική νόρμα ορίζει προφανώς μια απόσταση $d(x, y) = \|x - y\|$ στον \mathbb{C}^n , αφού προφανώς ικανοποιεί τις εξής ιδιότητες:

- $d(x, x) = 0, \forall x \in \mathbb{C}^n$ και $d(x, y) = 0$ αν και μόνο αν $x = y$.
- $d(x, y) = d(y, x), \forall x, y \in \mathbb{C}^n$
- $d(x, y) \leq d(x, z) + d(z, y), \forall x, y, z \in \mathbb{C}^n$.

Επομένως, ο $(\mathbb{C}^n, d(x, y) = \|x - y\|)$ είναι μετρικός χώρος.

Ορισμός 1.1.31. Νόρμα - p (p - norm) ενός διανύσματος $x \in \mathbb{C}^n$ με $p \geq 1$ είναι η

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}, \quad p \geq 1.$$

Για $p = \infty$ η νόρμα - p ορίζεται ως $\|x\|_\infty = \lim_{x \rightarrow \infty} \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} = \max_{i=1(1)n} \|x_i\|$.

Αντικαθιστώντας στον ορισμό της νόρμας - p όπου $p = 1, 2, \infty$ λαμβάνουμε τις τρεις νόρμες που χρησιμοποιούνται συνήθως στην Αριθμητική Ανάλυση. Έτσι έχουμε τις (ℓ_1 - norm) $\|x\|_1 = \sum_{i=1}^n \|x_i\|$, (ℓ_2 - norm ή Ευκλείδεια norm) $\|x\|_2 = \left(\sum_{i=1}^n \|x_i\|^2 \right)^{\frac{1}{2}}$ και (ℓ_∞ - norm ή norm μεγίστου) $\|x\|_\infty = \max_{i=1(1)n} \|x_i\|$.

Ορισμός 1.1.32. Δύο νόρμες $\mathbb{C}^n \rightarrow \mathbb{R}^{0,+}$ με $\|\cdot\|_\alpha$ και $\|\cdot\|_\beta$ λέγονται ισοδύναμες (συγκρίσιμες), αν και μόνο αν υπάρχουν σταθερές $c_1, c_2 > 0$ τέτοιες ώστε:

$$c_1 \|\cdot\|_\alpha \leq \|\cdot\|_\beta \leq c_2 \|\cdot\|_\alpha \quad (\text{ή } c_2^{-1} \|x\|_\beta \leq \|x\|_\alpha \leq c_1^{-1} \|x\|_\beta)$$

Θεώρημα 1.1.33. Δύο οποιεσδήποτε νόρμες διανυσμάτων στον \mathbb{C}^n είναι ισοδύναμες (συγκρίσιμες).

Με τρόπο παρόμοιο με τη νόρμα διανύσματος μπορούμε να ορίσουμε τη νόρμα πίνακα.

Ορισμός 1.1.34. Η απεικόνιση $\|\cdot\| : \mathbb{C}^{n,n} \rightarrow \mathbb{R}^{0,+}$ ορίζει μια νόρμα πίνακα, αν και μόνο αν:

- i) $\|A\| \geq 0$ και $\|A\| = 0$ αν και μόνο αν $A = 0$, $\forall A \in \mathbb{C}^{n,n}$
- ii) $\|cA\| = |c|\|A\|$, $\forall c \in \mathbb{C}$ και $\forall A \in \mathbb{C}^{n,n}$
- iii) $\|A + B\| \leq \|A\| + \|B\|$, $\forall A, B \in \mathbb{C}^{n,n}$
- iv) $\|AB\| \leq \|A\|\|B\|$, $\forall A, B \in \mathbb{C}^{n,n}$.

Θεώρημα 1.1.35. Η απεικόνιση $\|A\| := \sup_{x \in \mathbb{C}^n \setminus \{0\}} \frac{\|Ax\|}{\|x\|}$, $\forall A \in \mathbb{C}^{n,n}$ ορίζει μία νόρμα πίνακα.

Απόδειξη. Για να αποδείξουμε ότι η απεικόνιση αποτελεί νόρμα αρκεί να αποδείξουμε ότι ικανοποιεί τις 4 ιδιότητες της νόρμας.

Αρχικά, για την απόδειξη της πρώτης ιδιότητας μπορούμε να δούμε από τον τρόπο ορισμού της ότι $\|A\| \geq 0$ και $\|A\| = 0$ αν και μόνο αν $A = 0$. Για την απόδειξη της δεύτερης ιδιότητας,

$$\|cA\| = \sup_{x \in \mathbb{C}^n \setminus \{0\}} \frac{\|cAx\|}{\|x\|} = |c| \sup_{x \in \mathbb{C}^n \setminus \{0\}} \frac{\|Ax\|}{\|x\|} = |c|\|A\|.$$

Για την τρίτη ιδιότητα,

$$\begin{aligned} \|A + B\| &= \sup_{x \in \mathbb{C}^n \setminus \{0\}} \frac{\|(A + B)x\|}{\|x\|} \\ &= \sup_{x \in \mathbb{C}^n \setminus \{0\}} \frac{\|Ax + Bx\|}{\|x\|} \\ &\leq \sup_{x \in \mathbb{C}^n \setminus \{0\}} \frac{\|Ax\|}{\|x\|} + \sup_{x \in \mathbb{C}^n \setminus \{0\}} \frac{\|Bx\|}{\|x\|} \\ &= \|A\| + \|B\|. \end{aligned}$$

Τέλος, για την τέταρτη ιδιότητα από τον ορισμό της νόρμας μπορούμε να παρατηρήσουμε ότι $\forall x \in \mathbb{C}^n \setminus \{0\}$, $\frac{\|Ax\|}{\|x\|} \leq \|A\| \iff \|Ax\| \leq \|A\|\|x\|$. Εφαρμόζοντας αυτή την ανισότητα για το διάνυσμα Bx δύο φορές έχουμε

$$\frac{\|ABx\|}{\|x\|} \leq \frac{\|A\|\|Bx\|}{\|x\|} \leq \frac{\|A\|\|B\|\|x\|}{\|x\|} = \|A\|\|B\|. \quad \square$$

Οι νόρμες που παράγονται από αντίστοιχες διανυσματικές λέγονται φυσικές νόρμες. Οι φυσικές νόρμες πινάκων που χρησιμοποιούνται συχνότερα είναι εκείνες που παράγονται από τις αντίστοιχες νόρμες διανυσμάτων ℓ_1 - norm, ℓ_2 - norm και ℓ_∞ - norm. Έστω $A \in \mathbb{C}^{n,n}$. Τότε, για τον πίνακα A αποδεικνύεται ότι:

$$\|A\|_1 = \max_{j=1(1)n} \sum_{i=1}^n |a_{ij}|, \quad \|A\|_2 = \rho^{\frac{1}{2}}(A^H A) \quad \text{και} \quad \|A\|_\infty = \max_{i=1(1)m} \sum_{j=1}^m |a_{ij}|.$$

Για τις φυσικές νόρμες πινάκων ισχύουν οι επόμενες προτάσεις.

Θεώρημα 1.1.36. Έστω $\|\cdot\|$ κάποια φυσική νόρμα. Τότε $\|I\| = 1$.

Απόδειξη. Από τον ορισμό της φυσικής νόρμας

$$\|A\| := \sup_{x \in \mathbb{C}^n \setminus \{0\}} \frac{\|Ax\|}{\|x\|},$$

για $A = I$ έχουμε $\|I\| = \sup_{x \in \mathbb{C}^n \setminus \{0\}} \frac{\|x\|}{\|x\|} = 1$. □

Θεώρημα 1.1.37. Έστω $\|\cdot\|$ κάποια φυσική νόρμα. Τότε, $\forall A \in \mathbb{C}^{n,n}$ ισχύει: $\rho(A) \leq \|A\|$.

Απόδειξη. Έστω $\lambda \in \mathbb{C}$ ιδιοτιμή του πίνακα A και $x \in \mathbb{C}^n \setminus \{0\}$ αντίστοιχο ιδιοδιάνυσμα. Τότε ισχύει, $Ax = \lambda x$. Οπότε, παίρνοντας νόρμες και στα δύο μέλη έχουμε $\|Ax\| = \|\lambda x\|$. Εφαρμόζοντας την ιδιότητα (ii) της διανυσματικής νόρμας λαμβάνουμε: $\|Ax\| = \|\lambda x\| = |\lambda| \|x\|$. Αλλά, από τον ορισμό της φυσικής νόρμας ισχύει ότι

$$\forall x \in \mathbb{C}^n \setminus \{0\}, \quad \frac{\|Ax\|}{\|x\|} \leq \|A\| \iff \|Ax\| \leq \|A\| \|x\|.$$

Άρα τελικά, $|\lambda| \|x\| \leq \|A\| \|x\|$. Διαιρώντας και τα δύο μέλη με $\|x\|$ (που είναι θετική ποσότητα γιατί $x \neq 0$) έχουμε $|\lambda| \leq \|A\|$, όπου λ τυχαία ιδιοτιμή του πίνακα A . Άρα, $\rho(A) \leq \|A\|$. □

Σχόλιο. Η φασματική ακτίνα φράσσεται από όλες τις φυσικές νόρμες. Εύκολα μπορεί να παρατηρήσει κανείς ότι δεν είναι νόρμα, μπορεί όμως να προσεγγιστεί, όσο καλά θέλουμε, από μια φυσική νόρμα. Αυτό το εξασφαλίζει η εξής Πρόταση που δίνεται χωρίς απόδειξη: Για κάθε $A \in \mathbb{C}^{n,n}$ και $\forall \varepsilon > 0$ υπάρχει φυσική νόρμα $\|\cdot\|$ τέτοια ώστε $\|A\| \leq \rho(A) + \varepsilon$.

Θεώρημα 1.1.38 (Neumann). Έστω $\|\cdot\|$ κάποια φυσική νόρμα. Αν $A \in \mathbb{C}^{n,n}$ με $\|A\| < 1$, τότε:

i) Ο $I - A$ είναι αντιστρέψιμος

$$\text{ii) } \frac{1}{1 + \|A\|} \leq \|(I - A)^{-1}\| \leq \frac{1}{1 - \|A\|}$$

Απόδειξη. i) Έστω ότι ο $I - A$ είναι μη αντιστρέψιμος. Τότε, ο $I - A$ έχει ως ιδιοτιμή τη $\lambda = 0$ με αντίστοιχο ιδιοδιάνυσμα $x \in \mathbb{C}^n \setminus \{0\}$, δηλαδή ισχύει ότι:

$$(I - A)x = 0.$$

Άρα,

$$x - Ax = 0 \iff x = Ax.$$

Παίρνοντας νόρμες και στα δύο μέλη έχουμε: $\|x\| = \|Ax\|$. Από τον ορισμό της φυσικής νόρμας

$$\forall x \in \mathbb{C}^n \setminus \{0\}, \|Ax\| \leq \|A\|\|x\|.$$

Άρα τελικά,

$$\|x\| \leq \|A\|\|x\|$$

και διαιρώντας και τα δύο μέλη με $\|x\| (> 0)$ προκύπτει ότι $\|A\| \geq 1$, πράγμα άτοπο από υπόθεση.

ii) Αφού από (i) ο $I - A$ είναι αντιστρέψιμος ισχύει ότι: $I = (I - A)^{-1}(I - A)$.

$$1 = \|I\| = \|(I - A)^{-1}(I - A)\| \leq \|(I - A)^{-1}\| \|(I - A)\|,$$

από την ιδιότητα (iv) της νόρμας πινάκων. Ακόμη, από την ιδιότητα (iii) της νόρμας πινάκων

$$\|(I - A)\| \leq \|I\| + \|A\| = 1 + \|A\|.$$

Οπότε τελικά προκύπτει ότι: $\frac{1}{1 + \|A\|} \leq \|(I - A)^{-1}\|$. Αντίστοιχα,

$$\begin{aligned} 1 &= \|I\| = \|(I - A)^{-1}(I - A)\| \\ &= \|(I - A)^{-1} - (I - A)^{-1}A\| \\ &\geq \|(I - A)^{-1}\| - \|(I - A)^{-1}A\| \\ &\geq \|(I - A)^{-1}\| - \|(I - A)^{-1}\|\|A\| \\ &= \|(I - A)^{-1}\|(1 - \|A\|), \end{aligned}$$

αφού $\|(I - A)^{-1}A\| \leq \|(I - A)^{-1}\|\|A\|$ από την ιδιότητα (iv) της νόρμας πινάκων.

Οπότε τελικά: $\|(I - A)^{-1}\| \leq \frac{1}{1 - \|A\|}$. □

Σχόλιο. Το παραπάνω θεώρημα, το οποίο είναι γνωστό ως θεώρημα Neumann, ισχύει και για τον πίνακα $I + A$ αρκεί να θέσουμε όπου A τον $-A$ τόσο κατά τη διατύπωση όσο και κατά την απόδειξή του.

Ορισμός 1.1.39. Μια ακολουθία διανυσμάτων $\{x^{(k)}\}$, $k = 0(1)\infty$, με όρους $x^{(0)}, x^{(1)}, x^{(2)}, \dots \in \mathbb{C}^n$ συγκλίνει (converges) στο διάνυσμα $x \in \mathbb{C}^n$ ή έχει όριο το διάνυσμα $x \in \mathbb{C}^n$, αν και μόνο αν $x_i^{(k)} \rightarrow x_i, \forall i = 1(1)n$. Συμβολίζουμε με $x_{k \rightarrow \infty}^{(k)} \rightarrow x$ ή $\lim_{k \rightarrow \infty} x^{(k)} = x$ ή ισοδύναμα $(x^{(k)} - x)_{k \rightarrow \infty} \rightarrow 0$ και $\lim_{k \rightarrow \infty} (x^{(k)} - x) = 0$.

Οπότε, αν $\{x^{(k)}\}$, $k = 0(1)\infty$, είναι μια συγκλίνουσα ακολουθία διανυσμάτων με όριο το διάνυσμα $x \in \mathbb{C}^n$, τότε για οποιαδήποτε νόρμα $\|\cdot\|_\alpha$ είναι

$$\lim_{k \rightarrow \infty} \|x^{(k)} - x\|_\alpha = 0.$$

Απόδειξη. Η απόδειξη είναι προφανής από τον ορισμό της σύγκλισης ακολουθίας διανυσμάτων. \square

Ορισμός 1.1.40. Δείκτης κατάστασης ενός αντιστρέψιμου πίνακα $A \in \mathbb{C}^{n,n}$ ως προς μια φυσική νόρμα $\|\cdot\|$ καλείται ο αριθμός $\kappa(A) = \|A\| \|A^{-1}\|$.

Σχόλιο. Για τον δείκτη κατάστασης ενός αντιστρέψιμου πίνακα $A \in \mathbb{C}^{n,n}$ ισχύει προφανώς λόγω του ορισμού του ότι $\kappa(A) \geq 1$, αφού

$$1 = \|I\| = \|AA^{-1}\| \leq \|A\| \|A^{-1}\| = \kappa(A).$$

Ως δείκτης κατάστασης ενός μη αντιστρέψιμου πίνακα θεωρείται το ∞ .

Πρόταση 1.1.41. Έστω $A \in \mathbb{C}^{n,n}$ ένας ερμιτιανός και αντιστρέψιμος πίνακας. Ο δείκτης κατάστασης $\kappa_2(A)$ ως προς την Ευκλείδεια νόρμα $\|\cdot\|_2$ δίνεται από τον τύπο

$$\kappa_2(A) = \frac{\max_i |\lambda_i|}{\min_i |\lambda_i|},$$

όπου λ_i είναι οι ιδιοτιμές του πίνακα A .

Απόδειξη.

$$\begin{aligned} \kappa_2(A) &= \|A\|_2 \|A^{-1}\|_2 = [\rho(A^H A)]^{\frac{1}{2}} \left[\rho\left((A^{-1})^H A^{-1}\right) \right]^{\frac{1}{2}} \\ &= [\rho(A^2)]^{\frac{1}{2}} \left[\rho\left((A^H)^{-1} A^{-1}\right) \right]^{\frac{1}{2}} = [\rho(A^2)]^{\frac{1}{2}} \left[\rho\left((A^{-1})^2\right) \right]^{\frac{1}{2}} \end{aligned}$$

$$\begin{aligned}
&= \sqrt{\max_i (\lambda_i)^2} \sqrt{\max_i \frac{1}{(\lambda_i)^2}} = \max_i |\lambda_i| \max_i \left| \frac{1}{\lambda_i} \right| \\
&= \max_i |\lambda_i| \frac{1}{\min_i |\lambda_i|} = \frac{\max_i |\lambda_i|}{\min_i |\lambda_i|}. \quad \square
\end{aligned}$$

Σχόλιο. Ο δείκτης κατάστασης χαρακτηρίζει την κατάσταση του προβλήματος. Αν είναι πολύ μεγάλος, τότε το πρόβλημα έχει κακή κατάσταση, που σημαίνει ότι μικρές μεταβολές των στοιχείων του πίνακα A ή και του διανύσματος b επιφέρουν μεγάλες μεταβολές στη λύση x . Αν είναι σχετικά μικρός, τότε λέμε ότι έχει καλή κατάσταση. Προβλήματα με κακή κατάσταση αντιμετωπίζονται με την τεχνική της προρρύθμισης, όπου το σύστημα μετασχηματίζεται σε ένα ισοδύναμό του με καλή κατάσταση.

Ορισμός 1.1.42. Έστω $A \in \mathbb{C}^{n,n}$ ένας ερμιτιανός και θετικά ορισμένος πίνακας. Τότε υπάρχει μοναδικός ερμιτιανός και θετικά ορισμένος πίνακας $B \in \mathbb{C}^{n,n}$ τέτοιος ώστε $A = BB$. Ο πίνακας B λέγεται θετική τετραγωνική ρίζα του A και συμβολίζεται με $A^{1/2}$.

Ορισμός 1.1.43. Έστω $A \in \mathbb{C}^{n,n}$ ένας ερμιτιανός και θετικά ορισμένος πίνακας. Τότε, για κάθε $x \in \mathbb{C}^n$ η απεικόνιση $\mathbb{C}^n \rightarrow \mathbb{R}^{0,+}$ με τύπο $\|x\|_{A^{1/2}} := (Ax, x)^{1/2}$ ορίζει μια διανυσματική νόρμα η οποία είναι γνωστή ως A - νόρμα (A - norm) ή νόρμα ενέργειας (energy norm).

Απόδειξη. Είναι εύκολο να αποδείξουμε ότι η απεικόνιση ορίζει μια διανυσματική νόρμα, αρκεί να την μετασχηματίσουμε σε

$$\|x\|_{A^{1/2}} = (Ax, x)^{1/2} = \left(A^{1/2}x, A^{1/2}x \right)^{1/2} = \|A^{1/2}x\|.$$

Το ότι ικανοποιούνται οι τρεις ιδιότητες της διανυσματικής νόρμας αποδεικνύεται εύκολα. □

1.2 Κλασικές Επαναληπτικές Μέθοδοι

Για την επίλυση του γραμμικού αλγεβρικού συστήματος

$$Ax = b, \quad A \in \mathbb{C}^{n,n}, b \in \mathbb{C}^n,$$

έχουν προταθεί διάφορες μέθοδοι. Αυτές χαρακτηρίζονται ως άμεσες, όταν χρησιμοποιώντας ακριβή αριθμητική, βρίσκουν ακριβώς τη λύση του συστήματος μετά

από πεπερασμένο πλήθος πράξεων. Σε αυτήν την κατηγορία κατατάσσονται η μέθοδος απαλοιφής του Gauss με ολική ή μερική οδήγηση, η μέθοδος της LU παραγοντοποίησης και η παραγοντοποίηση Cholesky όταν ο A είναι συμμετρικός και θετικά ορισμένος. Στο πλαίσιο της παρούσης εργασίας δεν θα γίνει αναφορά σε αυτές και ο αναγνώστης παραπέμπεται στο βιβλίο [4]. Οι μέθοδοι που βρίσκουν μια προσέγγιση της λύσης, γιατί αυτή δεν μπορεί να υπολογιστεί ακριβώς λέγονται έμμεσες μέθοδοι. Σε αυτήν την κατηγορία ανήκουν οι κλασικές επαναληπτικές μέθοδοι καθώς και οι επαναληπτικές μέθοδοι ελαχιστοποίησης (ή επαναληπτικές μέθοδοι βελτιστοποίησης) που θα αποτελέσουν το κύριο θέμα αυτής της διατριβής. Πριν όμως αρχίσουμε να αναπτύξουμε μία προς μία αυτές τις μεθόδους, θα γίνει μια αναφορά στις κλασικές επαναληπτικές μεθόδους. Αφού δοθεί η βασική ιδέα για τον τρόπο κατασκευής τους, θα αναφερθούν οι κυριότερες από αυτές.

Έστω το γραμμικό αλγεβρικό σύστημα

$$Ax = b, \quad A \in \mathbb{C}^{n,n}, \det(A) \neq 0, b \in \mathbb{C}^n.$$

Θα κατασκευάσουμε έναν επαναληπτικό αλγόριθμο, στον οποίο η παραγόμενη ακολουθία διανυσμάτων θα συγκλίνει στην ακριβή λύση του συστήματος, εφόσον βέβαια είναι συγκλίνουσα. Αρχικά, παρατηρούμε ότι ο πίνακας των συντελεστών μπορεί να γραφεί ως $A = M - N$, όπου ο πίνακας M , ο οποίος ονομάζεται προρρυθμιστής (preconditioner) πίνακας, είναι αντιστρέψιμος (ισοδύναμα $\det(M) \neq 0$) και το γραμμικό αλγεβρικό σύστημα

$$Mx = b, \quad M \in \mathbb{C}^{n,n}, b \in \mathbb{C}^n,$$

λύνεται πολύ ευκολότερα από το αρχικό σύστημα. Τότε

$$Ax = b \iff (M - N)x = b \iff Mx = Nx + b.$$

Πολλαπλασιάζοντας από αριστερά και τα δύο μέλη αυτής της σχέσης με τον πίνακα M^{-1} , έχουμε:

$$x = M^{-1}Nx + M^{-1}b$$

Έτσι, $x = Tx + c$, όπου $T := M^{-1}N$, $c := M^{-1}b$. Ένας προφανής αλγόριθμος που μπορεί να κατασκευαστεί από αυτή τη σχέση για την προσέγγιση της λύσης x του συστήματος είναι:

$$x^{(k+1)} = Tx^{(k)} + c, \quad k = 0, 1, 2, \dots, \quad x^{(0)} \in \mathbb{C}^n \text{ τυχόν.}$$

Μπορούμε να παρατηρήσουμε ακόμη ότι αν η ακολουθία που παράγεται από τον παραπάνω αλγόριθμο είναι συγκλίνουσα, αυτή θα συγκλίνει στη λύση $x = A^{-1}b$

του συστήματος. Πράγματι, υπάρχει $y \in \mathbb{C}^n$ τέτοιο ώστε: $\lim_{k \rightarrow \infty} x^{(k)} = y$. Θα αποδείξουμε ότι ισχύει $y = x$. Αφού

$$x^{(k+1)} = Tx^{(k)} + c \implies \lim_{k \rightarrow \infty} x^{(k+1)} = \lim_{k \rightarrow \infty} (Tx^{(k)} + c),$$

αντικαθιστώντας τις ποσότητες για τα T και c έχουμε

$$\begin{aligned} y &= M^{-1}Ny + M^{-1}b \implies My = Ny + b \\ &\implies My - Ny = b \\ &\implies (M - N)y = b \\ &\implies Ay = b. \end{aligned}$$

Οπότε, $y = A^{-1}b$, δηλαδή έχουμε σύγκλιση στη λύση του συστήματος.

Όλες οι κλασικές επαναληπτικές μέθοδοι βασίζονται στο ότι ο αρχικός πίνακας A του συστήματος μπορεί να γραφεί ως $A = D - L - U$, όπου $D = \text{diag}(A)$, δηλαδή διαγώνιος πίνακας με στοιχεία κύριας διαγωνίου τα στοιχεία του A , L αυστηρά κάτω τριγωνικός και U αυστηρά άνω τριγωνικός. Μπορούμε να παρατηρήσουμε δε, από τον τρόπο με τον οποίο ορίζεται, ότι μια τέτοια διάσπαση του A είναι μοναδική. Ανάλογα με την επιλογή του προρρυθμιστή πίνακα M που θα κάνουμε, έχουμε τις δύο βασικές επαναληπτικές μεθόδους Jacobi και Gauss - Seidel και ως γενίκευση της δεύτερης την μέθοδο SOR.

Στη συνέχεια, παρατίθενται τα θεωρήματα που εξασφαλίζουν τη σύγκλιση των κλασικών επαναληπτικών μεθόδων. Άλλες επαναληπτικές μέθοδοι είναι η μέθοδος παρεκβολής της Jacobi (extrapolated Jacobi), η μέθοδος παρεκβολής της Gauss-Seidel (extrapolated Gauss-Seidel) και η συμμετρική μέθοδος SOR (SSOR). Η τεχνική της παρεκβολής αφορά στη βελτίωση της (ασυμπτωτικής) ταχύτητας σύγκλισης της μεθόδου πάνω στην οποία εφαρμόζεται ή ακόμη και στην μετατροπή μιας αποκλίνουσας μεθόδου σε συγκλίνουσα σε ορισμένες περιπτώσεις. Για περισσότερες λεπτομέρειες βλέπε το [4].

1.3 Σύγκλιση της Απλής Επαναληπτικής Μεθόδου

Η πιο απλή και γενική ικανή και αναγκαία συνθήκη σύγκλισης της απλής επαναληπτικής μεθόδου προκύπτει βάσει του ορισμού της σύγκλισης μιας ακολουθίας διαδοχικών προσεγγίσεων στην ακριβή λύση $x = A^{-1}b$ του συστήματος. Παρακάτω παρατίθεται ένα προκαταρκτικό λήμμα το οποίο αποδεικνύεται χρήσιμο και στη συνέχεια το θεώρημα το οποίο καθορίζει τη σύγκλιση η μή της μεθόδου.

Λήμμα 1.3.1. Έστω $T \in \mathbb{C}^{n,n}$. Τότε $\lim_{k \rightarrow \infty} T^k = 0 \iff \rho(T) < 1$.

Απόδειξη. (\implies) Έστω ότι $\lim_{k \rightarrow \infty} T^k = 0$. Τότε $\lim_{k \rightarrow \infty} \|T^k\| = 0$ για κάποια φυσική νόρμα. Άρα υπάρχει $k_0 \in \mathbb{N}$ τέτοιο ώστε $\forall k \geq k_0 \ \|T^k\| < 1$. Επίσης, αφού

$$\rho^k(T) = \rho(T^k) \leq \|T^k\|,$$

προκύπτει τελικά ότι $\rho^k(T) < 1 \implies \rho(T) < 1$.

(\impliedby) Έστω ότι $\rho(T) < 1$. Τότε, $\forall \varepsilon > 0$ υπάρχει $\|\cdot\|$ τέτοια ώστε $\|T\| \leq \rho(T) + \varepsilon$, όπου $\|\cdot\|$ φυσική νόρμα. Οπότε, για κάποιο ε με $0 < \varepsilon < 1 - \rho(T)$ θα είναι

$$\|T\| < 1 \implies \|T^k\| \leq \|T\|^k < 1, \quad \forall k \in \mathbb{N}.$$

Έτσι,

$$\lim_{k \rightarrow \infty} \|T^k\| = 0 \implies \lim_{k \rightarrow \infty} T^k = 0.$$

□

Συμπέρασμα. Έστω $A \in \mathbb{C}^{n,n}$ και $\|\cdot\|$ κάποια φυσική νόρμα. Ισχύει ότι $\rho(A) = \lim_{k \rightarrow \infty} \|A^k\|^{1/k}$

Το παρακάτω Θεώρημα δίνει μια ικανή και αναγκαία συνθήκη για τη σύγκλιση της (διανυσματικής) ακολουθίας των διαδοχικών προσεγγίσεων που παράγονται μέσω του επαναληπτικού αλγορίθμου στην ακριβή λύση $x = A^{-1}b$ του συστήματος.

Θεώρημα 1.3.2. Ικανή και αναγκαία συνθήκη για τη σύγκλιση της ακολουθίας των διαδοχικών προσεγγίσεων $\{x^{(k)}\}$, $k = 0, 1, 2, \dots$ στην ακριβή λύση του συστήματος $x = A^{-1}b$ είναι η $\rho(T) < 1$.

Απόδειξη. Ισχύει

$$x = Tx + c, \text{ με } T := M^{-1}N, \text{ και } c := M^{-1}b$$

για τον προσδιορισμό της λύσης $x = A^{-1}b$ και

$$x^{(k+1)} = Tx^{(k)} + c, \quad k = 0, 1, 2, \dots, \quad x^{(0)} \in \mathbb{C}^n \text{ οποιοδήποτε.}$$

Αφαιρώντας κατά μέλη παίρνουμε διαδοχικά

$$x - x^{(k+1)} = T(x - x^{(k)}) \implies \varepsilon^{(k+1)} = T\varepsilon^{(k)}, \quad \forall k = 0, 1, 2, \dots,$$

όπου με $\varepsilon^{(k)}$, $\forall k = 0, 1, 2, \dots$ συμβολίζουμε το διάνυσμα - σφάλμα στην k - επανάληψη. Από την τελευταία σχέση για $k = 0$ προκύπτει ότι: $\varepsilon^{(1)} = T\varepsilon^{(0)}$. Για $k = 1$ ότι $\varepsilon^{(2)} = T\varepsilon^{(1)} = T^2\varepsilon^{(0)}$. Έτσι, με επαγωγή στο k προκύπτει τελικά ότι: $\varepsilon^{(k)} = T^k\varepsilon^{(0)}$, $\varepsilon^{(0)}$ τυχόν. Για να έχουμε σύγκλιση στη λύση του συστήματος απαιτούμε να ισχύει $\lim_{k \rightarrow \infty} x^{(k)} = A^{-1}b$, $x^{(0)}$ τυχόν. Ισοδύναμα,

$$\lim_{k \rightarrow \infty} (x - x^{(k)}) = x - A^{-1}b \Rightarrow \lim_{k \rightarrow \infty} \varepsilon^{(k)} = 0, \varepsilon^{(0)} \text{ τυχόν.}$$

Αρκεί αυτό να ισχύει για όλα τα διανύσματα μιας βάσης του C^n στη θέση του $\varepsilon^{(0)}$. Ως τέτοια παίρνουμε τις στήλες του μοναδιαίου πίνακα e^j , $j = 1(1)n$. Τότε, $\varepsilon^{(k)} = T^k\varepsilon^{(0)} = T^ke^j$, είναι η j στήλη του T^k . Δηλαδή,

$$\lim_{k \rightarrow \infty} \varepsilon^{(k)} = \lim_{k \rightarrow \infty} T^ke^j = 0, j = 1(1)n \iff \lim_{k \rightarrow \infty} T^k = 0 \iff \rho(T) < 1. \quad \square$$

Ορισμός 1.3.3. Έστω δύο πίνακες $A, B \in C^{n,n}$. Αν για κάποιο θετικό ακέραιο k υπάρχει φυσική νόρμα $\|\cdot\|_\alpha$ με $\|A^k\|_\alpha < 1$, τότε η ποσότητα

$$\mathfrak{R}(A^k) = -\log\left(\|A^k\|_\alpha^{\frac{1}{k}}\right) = \frac{-\log\|A^k\|_\alpha}{k}$$

καλείται μέση ταχύτητα σύγκλισης για k επαναλήψεις. Αν ισχύει $\|B^k\| < 1$ και $\mathfrak{R}(A^k) < \mathfrak{R}(B^k)$, τότε ο B είναι επαναληπτικά ταχύτερος του A για k επαναλήψεις.

Παρατήρηση. Αν κάνοντας χρήση του ορισμού της μέσης ταχύτητας σύγκλισης υπολογίσουμε την ταχύτητα σύγκλισης για $k \rightarrow \infty$ (πρακτικά για μεγάλες τιμές του k), δηλαδή την ασυμπτωτική ταχύτητα σύγκλισης, έχουμε $\mathfrak{R}_\infty(A) = \lim_{k \rightarrow \infty} \mathfrak{R}(A^k) = -\log\rho(A)$. Παρατηρούμε ότι σε κάποια επαναληπτική μέθοδο η ταχύτητα σύγκλισης εξαρτάται μόνο από τον δείκτη κατάστασης του επαναληπτικού πίνακα της μεθόδου. Συγκεκριμένα, όσο μικρότερη είναι η φασματική ακτίνα, τόσο ταχύτερη σύγκλιση στη λύση του συστήματος έχουμε και όσο αυτή μεγαλώνει, τόσο πιο αργή γίνεται η σύγκλιση της μεθόδου.

ΚΕΦΑΛΑΙΟ 2

Μέθοδοι Ορθογωνοποίησης

2.1 Μέθοδος Orthomin(1) - Μέθοδος Απότομης Καθόδου (Steepest Descent - SD)

Οι πρώτες μέθοδοι αυτής της κατηγορίας που αναπτύχθηκαν σε χρονολογική σειρά είναι η μέθοδος Orthomin(1) και ως ειδικότερη περίπτωση της, αυτή της Απότομης Καθόδου. Αρχικά, για τη μέθοδο Orthomin(1) θα ασχοληθούμε με τον τρόπο κατασκευής της και τη βασική ιδέα πίσω από αυτή. Αν $x^{(k)}$ είναι μια προσέγγιση της λύσης στο k βήμα της μεθόδου, τότε κάθε επόμενη επανάληψη δίνεται από τον τύπο

$$x^{(k+1)} = x^{(k)} + \alpha_k r^{(k)},$$

όπου με $r^{(k)}$ συμβολίζουμε το διάνυσμα - υπόλοιπο (residual - vector) το οποίο ισούται με $r^{(k)} = b - Ax^{(k)}$.

Σε κάθε βήμα το ζητούμενο είναι η ελαχιστοποίηση της Ευκλείδειας νόρμας (ℓ_2 - norm) του διανύσματος - υπόλοιπο $r^{(k)}$. Ας θεωρήσουμε λοιπόν την Ευκλείδεια νόρμα του διανύσματος - υπόλοιπο στο $k+1$ βήμα της μεθόδου. Είναι κατ' αρχάς

$$\begin{aligned} r^{(k+1)} &= b - Ax^{(k+1)} = b - A(x^{(k)} + \alpha_k r^{(k)}) \\ &= b - Ax^{(k)} - A\alpha_k r^{(k)} = r^{(k)} - \alpha_k Ar^{(k)}. \end{aligned}$$

Οπότε,

$$\begin{aligned} \|r^{(k+1)}\|^2 &= (r^{(k+1)}, r^{(k+1)}) = (r^{(k)} - \alpha_k Ar^{(k)}, r^{(k)} - \alpha_k Ar^{(k)}) \\ &= (r^{(k)}, r^{(k)}) - \alpha_k (r^{(k)}, Ar^{(k)}) + |\alpha_k|^2 (Ar^{(k)}, Ar^{(k)}) \\ &\quad - \bar{\alpha}_k (Ar^{(k)}, r^{(k)}) \\ &= (r^{(k)}, r^{(k)}) - 2\operatorname{Re}(\alpha_k (r^{(k)}, Ar^{(k)})) + |\alpha_k|^2 (Ar^{(k)}, Ar^{(k)}). \end{aligned}$$

Γράφοντας τους (γενικά) μιγαδικούς αριθμούς α_k και $(r^{(k)}, Ar^{(k)})$ στη μορφή $\alpha_k = \alpha + i\beta$, $\alpha, \beta \in \mathbb{R}$, και $(r^{(k)}, Ar^{(k)}) = \gamma + i\delta$, $\gamma, \delta \in \mathbb{R}$ έχουμε

$$F(\alpha, \beta) := \|r^{(k+1)}\|^2 = \|r^{(k)}\|^2 - 2(\gamma\alpha - \beta\delta) + (\alpha^2 + \beta^2) \|Ar^{(k)}\|^2.$$

Για να βρούμε το σημείο που ελαχιστοποιεί αυτήν την ποσότητα αρκεί να προσδιορίσουμε τα κρίσιμα σημεία της και έπειτα από αυτά εκείνο που ελαχιστοποιεί την τιμή της. Έτσι,

$$\begin{aligned} \nabla F(\alpha, \beta) = 0 &\iff \frac{\partial F}{\partial \alpha} = 0 \\ &\iff -2\gamma + 2\alpha \|Ar^{(k)}\|^2 = 0 \\ &\iff \alpha = \frac{\gamma}{\|Ar^{(k)}\|^2} \end{aligned}$$

και

$$\frac{\partial F}{\partial \beta} = 0 \iff \beta = -\frac{\delta}{\|Ar^{(k)}\|^2}.$$

Θεωρούμε τον Εσσιανό πίνακα H_F της F , ο οποίος είναι

$$H_F = \begin{bmatrix} 2\|Ar^{(k)}\|^2 & 0 \\ 0 & 2\|Ar^{(k)}\|^2 \end{bmatrix}.$$

Ο H_F είναι συμμετρικός, αφού προφανώς ισχύει $H_F^T = H_F$, και θετικά ορισμένος, άρα το κρίσιμο σημείο που βρέθηκε είναι το σημείο ελαχίστου της F . Αντικαθιστώντας τις τιμές για τα α, β που βρήκαμε έχουμε ότι

$$\alpha_k = \frac{\gamma}{\|Ar^{(k)}\|^2} - i \frac{\delta}{\|Ar^{(k)}\|^2} = \frac{\gamma - i\delta}{(Ar^{(k)}, Ar^{(k)})} = \frac{(Ar^{(k)}, r^{(k)})}{(Ar^{(k)}, Ar^{(k)})}.$$

Στην περίπτωση που ο πίνακας A είναι συμμετρικός και θετικά ορισμένος, εκτός από τη διαδικασία που περιγράφηκε παραπάνω, μπορούμε με αρκετά παρόμοιο τρόπο να οδηγηθούμε στη μέθοδο Απότομης Καθόδου. Η μέθοδος Απότομης Καθόδου κατασκευάζεται αν σε κάθε βήμα της επιδιώκουμε την ελαχιστοποίηση της A - νόρμας (A - norm) ή νόρμας ενέργειας (energy norm) του διάνυσματος - σφάλμα. Ως διάνυσμα - σφάλμα ορίζεται η διαφορά της προσεγγιστικής (υπολογιστικής) λύσης από την ακριβή λύση, δηλαδή ισχύει: $\varepsilon^{(k+1)} = x - x^{(k+1)}$. Αφού

$$x^{(k+1)} = x^{(k)} + \alpha_k r^{(k)}$$

έπεται ότι

$$\varepsilon^{(k+1)} = \varepsilon^{(k)} + \alpha_k r^{(k)}.$$

Είναι:

$$\begin{aligned} \|\varepsilon^{(k+1)}\|_{A^{\frac{1}{2}}}^2 &= (\varepsilon^{(k+1)}, A\varepsilon^{(k+1)}) = (\varepsilon^{(k)} - \alpha_k r^{(k)}, A(\varepsilon^{(k)} - \alpha_k r^{(k)})) \\ &= (\varepsilon^{(k)} - \alpha_k r^{(k)}, A\varepsilon^{(k)} - \alpha_k Ar^{(k)}) \\ &= (\varepsilon^{(k)}, A\varepsilon^{(k)}) - \alpha_k (\varepsilon^{(k)}, Ar^{(k)}) - \alpha_k (r^{(k)}, A\varepsilon^{(k)}) \\ &\quad + \alpha_k^2 (Ar^{(k)}, r^{(k)}) \\ &= \|\varepsilon^{(k)}\|_{A^{\frac{1}{2}}}^2 - 2\alpha_k (r^{(k)}, r^{(k)}) + \alpha_k^2 \|r^{(k)}\|^2, \end{aligned}$$

γιατί

$$(\varepsilon^{(k)}, Ar^{(k)}) = (r^{(k)}, A\varepsilon^{(k)}) = (r^{(k)}, r^{(k)}),$$

αφού ισχύει

$$A\varepsilon^{(k)} = A(x - x^{(k)}) = b - Ax^{(k)} = r^{(k)}.$$

Παρατηρούμε ότι η τελευταία ποσότητα είναι ένα τριώνυμο ως προς α_k (για πραγματικό α_k) με συντελεστές Ευκλείδεια εσωτερικά γινόμενα οι οποίοι προφανώς είναι θετικοί. Η τιμή του ελαχιστοποιείται για $\alpha_k = \frac{(r^{(k)}, r^{(k)})}{(Ar^{(k)}, r^{(k)})}$.

Ένας άλλος τρόπος κατασκευής της Μεθόδου Απότομης Καθόδου για πραγματικό, συμμετρικό και θετικά ορισμένο πίνακα βασίζεται στην ελαχιστοποίηση της τιμής ενός συναρτησιακού και παρουσιάζεται παρακάτω. Ως συναρτησιακό προς ελαχιστοποίηση επιλέγεται το

$$f(x) := \frac{1}{2}(Ax, x) - (b, x), \quad x \in \mathbb{C}^n.$$

Η f λαμβάνει ελάχιστο στο σημείο $x = A^{-1}b$. Για να βρούμε ποια είναι η τιμή του α_k που ελαχιστοποιεί τη συνάρτηση f έχουμε:

$$\begin{aligned} f(x^{(k)} + \alpha_k r^{(k)}) &= \frac{1}{2}(A(x^{(k)} + \alpha_k r^{(k)}), x^{(k)} + \alpha_k r^{(k)}) - (b, x^{(k)} + \alpha_k r^{(k)}) \\ &= \frac{1}{2}(Ax^{(k)}, x^{(k)}) + \frac{1}{2}\alpha_k (Ax^{(k)}, r^{(k)}) + \frac{1}{2}\alpha_k (Ar^{(k)}, x^{(k)}) \\ &\quad + \frac{1}{2}\alpha_k^2 (Ar^{(k)}, r^{(k)}) - (b, x^{(k)}) - \alpha_k (b, r^{(k)}) \\ &= f(x^{(k)}) + \frac{1}{2}\alpha_k^2 (Ar^{(k)}, r^{(k)}) + \alpha_k (Ax^{(k)}, r^{(k)}) - \alpha_k (b, r^{(k)}) \\ &= f(x^{(k)}) + \frac{1}{2}\alpha_k^2 (Ar^{(k)}, r^{(k)}) - \alpha_k (r^{(k)}, r^{(k)}). \end{aligned}$$

Η παραπάνω ποσότητα είναι ένα τριώνυμο δευτέρου βαθμού ως προς α_k . Ο συντελεστής του μεγιστοβάθμιου όρου είναι θετικός αριθμός, αφού ο A είναι συμμετρικός και θετικά ορισμένος και το $r^{(k)} \neq 0$, οπότε το τριώνυμο ελαχιστοποιείται για $\alpha_k = \frac{(r^{(k)}, r^{(k)})}{(Ar^{(k)}, r^{(k)})}$. Παρακάτω παρατίθεται ο αλγόριθμος της Μεθόδου Απότομης Καθόδου.

Αλγόριθμος 2.1 Μέθοδος Απότομης Καθόδου

Δεδομένα: $A \in \mathbb{C}^{n,n}$, $A^H = A$, A θετικά ορισμένος, $b \in \mathbb{C}^n$, $\varepsilon \in \mathbb{R}^+$ η μέγιστη επιτρεπτή ανοχή σφάλματος.

Αποτέλεσμα: $x = x^{(k)}$ η προσέγγιση της λύσης.

- 1: $x^{(0)} = 0$
- 2: $r^{(0)} = b$
- 3: $k = 0$
- 4: **Εφόσον** $\|r^{(k)}\| > \varepsilon$
- 5: $\alpha_k = \frac{(r^{(k)}, r^{(k)})}{(Ar^{(k)}, r^{(k)})}$
- 6: $x^{(k+1)} = x^{(k)} + \alpha_k r^{(k)}$
- 7: $r^{(k+1)} = b - Ax^{(k+1)}$
- 8: $k = k + 1$
- 9: **Τέλος 'Εφόσον'**

Το κόστος του αλγορίθμου είναι $O(n^2)$ πράξεις ανά βήμα, αφού υπάρχουν δύο πολλαπλασιασμοί πίνακα επί διάνυσμα (ο $Ax^{(k)}$ και ο $Ar^{(k)}$). Αυτοί, δύναται να αναχθούν σε έναν, αν ισοδύναμα για τον υπολογισμό του $r^{(k)}$ υπολογίσουμε το $r^{(k+1)} = b - Ax^{(k+1)} = b - A(x^{(k)} + \alpha_k r^{(k)}) = r^{(k)} - \alpha_k Ar^{(k)}$. Έτσι υπολογίζουμε μόνο το $Ar^{(k)}$ κατά τον υπολογισμό του α_k και στη συνέχεια το επαναχρησιμοποιούμε για να υπολογίσουμε το $r^{(k+1)}$. Στην περίπτωση ερμιτιανού και θετικά ορισμένου πίνακα A , έχει αποδειχθεί ότι ισχύουν κάποιες σχέσεις που συνδέουν τις Ευκλείδειες νόρμες των διανυσμάτων - υπόλοιπο που παράγονται από τη Μέθοδο Orthomin(1), καθώς και τις νόρμες ενέργειας των διανυσμάτων - σφάλμα που παράγονται από τη Μέθοδο Απότομης Καθόδου. Αυτές οι σχέσεις δίνονται στη συνέχεια.

Θεώρημα 2.1.1. Έστω $A \in \mathbb{C}^{n,n}$ ένας ερμιτιανός και θετικά ορισμένος πίνακας. Για την Ευκλείδεια νόρμα του διανύσματος - υπόλοιπο της Μεθόδου Ortho-

min(1) ισχύει: $\|r^{(k)}\| \leq \left(\frac{\kappa - 1}{\kappa + 1}\right)^k \|r^{(0)}\|$.

Απόδειξη. Από τον τρόπο ορισμού της μεθόδου έχουμε

$$\begin{aligned} r^{(k+1)} &= b - Ax^{(k+1)} = b - A(x^{(k)} + \alpha_k r^{(k)}) \\ &= r^{(k)} - \alpha_k A r^{(k)} = (I - \alpha_k A) r^{(k)} \end{aligned}$$

Από αυτή παίρνουμε

$$\|r^{(k+1)}\| = \|(I - \alpha_k A) r^{(k)}\| \leq \|I - \alpha_k A\| \|r^{(k)}\|.$$

Αρκεί να υπολογίσουμε το α έτσι ώστε να ελαχιστοποιείται η $\|I - \alpha A\|$, δηλαδή να βρεθεί το $\min_{\alpha \in \mathbb{R}} \|I - \alpha A\|$. Έστω λ_i , $i = 1(1)n$, οι n το πλήθος ιδιοτιμές του πίνακα A και λ_{\max} , λ_{\min} η μέγιστη και η ελάχιστη ιδιοτιμή αντίστοιχα. Ζητάμε το

$$\min_{\alpha \in \mathbb{R}} \|I - \alpha A\| = \min_{\alpha \in \mathbb{R}} \max_{\lambda_i} |1 - \alpha \lambda_i| \leq \min_{\alpha \in \mathbb{R}} \max_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} |1 - \alpha \lambda|$$

Αυτό επιτυγχάνεται όταν

$$|1 - \alpha \lambda_{\min}| = |1 - \alpha \lambda_{\max}|$$

ή

$$1 - \alpha \lambda_{\min} = \alpha \lambda_{\max} - 1 \iff \alpha = \frac{2}{\lambda_{\min} + \lambda_{\max}}.$$

Επομένως,

$$\min_{\alpha \in \mathbb{R}} |1 - \alpha A| \leq 1 - \frac{2\lambda_{\min}}{\lambda_{\max} + \lambda_{\min}} = \frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}}.$$

Διαιρώντας αριθμητή και παρονομαστή με το $\lambda_{\min} \neq 0$, έχουμε

$$\|r^{(k+1)}\| \leq \left(\frac{\kappa - 1}{\kappa + 1} \right) \|r^{(k)}\|,$$

όπου $\kappa = \kappa_2 = \frac{\lambda_{\max}}{\lambda_{\min}}$ ο δείκτης κατάστασης του συστήματος. □

Επαγωγικά, έχουμε ότι $\|r^{(k)}\| \leq \left(\frac{\kappa - 1}{\kappa + 1} \right)^k \|r^{(0)}\|$. Ακριβώς η ίδια εκτίμηση ισχύει και στην περίπτωση ερμιτιανού και θετικά ορισμένου πίνακα για την A - νόρμα του διανύσματος - σφάλμα της μεθόδου Απότομης Καθόδου. Δηλαδή ισχύει το εξής:

Θεώρημα 2.1.2. Έστω $A \in \mathbb{C}^{n,n}$ ένας ερμιτιανός και θετικά ορισμένος πίνακας. Για την νόρμα ενέργειας του διανύσματος - υπόλοιπο που παράγεται από τη Μέθοδο Απότομης Καθόδου ισχύει: $\|\varepsilon^{(k+1)}\|_{A^{\frac{1}{2}}} \leq \left(\frac{\kappa - 1}{\kappa + 1}\right)^k \|\varepsilon^{(0)}\|_{A^{\frac{1}{2}}}$, όπου $\kappa = \kappa_2 = \frac{\lambda_{\max}}{\lambda_{\min}}$ ο δείκτης κατάστασης του συστήματος.

Θεώρημα 2.1.3. Για δύο διαδοχικά διανύσματα υπόλοιπο της μεθόδου Απότομης Καθόδου ισχύει ότι $(r^{(k+1)}, r^{(k)}) = 0$, δηλαδή τα $r^{(k+1)}$ και $r^{(k)}$ είναι ορθογώνια μεταξύ τους.

Απόδειξη. Ισχύει ότι

$$r^{(k+1)} = r^{(k)} - \alpha_k Ar^{(k)} = r^{(k)} - \frac{(r^{(k)}, r^{(k)})}{(Ar^{(k)}, r^{(k)})} Ar^{(k)}$$

Οπότε,

$$\begin{aligned} (r^{(k+1)}, r^{(k)}) &= \left(r^{(k)} - \frac{(r^{(k)}, r^{(k)})}{(Ar^{(k)}, r^{(k)})} Ar^{(k)}, r^{(k)} \right) \\ &= (r^{(k)}, r^{(k)}) - \frac{(r^{(k)}, r^{(k)})}{(Ar^{(k)}, r^{(k)})} (Ar^{(k)}, r^{(k)}) = 0. \quad \square \end{aligned}$$

Ο αλγόριθμος εκτελείται όσο ικανοποιείται μία συνθήκη τερματισμού η οποία έχει τεθεί από τον χρήστη εκ των προτέρων, όπως και όλοι οι αλγόριθμοι που θα παρουσιαστούν. Στην περίπτωση της μεθόδου Απότομης Καθόδου αυτή είναι η $\|r^{(k)}\| \leq \varepsilon$, όπου ε η επιθυμητή ανοχή σφάλματος.

2.2 Μέθοδος Γενικών Διευθύνσεων (Generalized Direction - GD)

Μετά τη μέθοδο Απότομης Καθόδου αναπτύχθηκε η μέθοδος Γενικών Διευθύνσεων. Στην μέθοδο Απότομης Καθόδου κατά την $k + 1$ επανάληψη η διεύθυνση που επιλέγεται κατά την οποία ελαχιστοποιείται η τιμή του συναρτησιακού $f(x^{(k+1)})$ είναι η

$$r^{(k)} = -\nabla f(x^{(k)}), \quad k = 0, 1, \dots$$

Παρόμοια είναι η λογική ελαχιστοποίησης για τη Μέθοδο Γενικών Διευθύνσεων, με τη διαφορά ότι οι διευθύνσεις είναι γενικά τυχαίες. Αν $x^{(k)}$ είναι η

προσέγγιση της λύσης στο βήμα k , τότε η επόμενη προσέγγιση ορίζεται ως

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)},$$

όπου τα διανύσματα $p^{(k)} \in \mathbb{C}^n$ είναι γενικά τυχαίες διευθύνσεις. Επίσης, ισχύουν ανάλογοι τύποι με τη μέθοδο Απότομης Καθόδου για το διάνυσμα - υπόλοιπο και το διάνυσμα - σφάλμα. Έχουμε

$$\begin{aligned} r^{(k+1)} &= b - Ax^{(k+1)} \\ &= b - A(x^{(k)} + \alpha_k p^{(k)}) \\ &= b - Ax^{(k)} - \alpha_k Ap^{(k)} \\ &= r^{(k)} - \alpha_k Ap^{(k)} \end{aligned}$$

και

$$\varepsilon^{(k+1)} = x - x^{(k+1)} = x - x^{(k)} - \alpha_k p^{(k)} = \varepsilon^{(k)} - \alpha_k p^{(k)}.$$

Σχόλιο. Αξίζει να σημειωθεί ότι η Μέθοδος Γενικών Διευθύνσεων είναι μια πάρα πολύ γενική μέθοδος, αλλά για συγκεκριμένη επιλογή των διευθύνσεων $p^{(k)}$ μπορούν να κατασκευαστούν χρήσιμες μέθοδοι που παρουσιάζονται στη συνέχεια.

Ένας τρόπος κατασκευής της Μεθόδου Γενικών Διευθύνσεων, όπως είδαμε και για τη Μέθοδο Απότομης Καθόδου, είναι μέσω ελαχιστοποίησης ενός συναρτησιακού. Θεωρώντας το συναρτησιακό $f(x) := \frac{1}{2}(Ax, x) - (b, x)$ λαμβάνουμε ότι

$$\begin{aligned} f(x^{(k+1)}) &= f(x^{(k)} + \alpha_k p^{(k)}) \\ &= \frac{1}{2}(A(x^{(k)} + \alpha_k p^{(k)}), x^{(k)} + \alpha_k p^{(k)}) - (b, x^{(k)} + \alpha_k p^{(k)}) \\ &= \frac{1}{2}(Ax^{(k)}, x^{(k)}) + \alpha_k (Ax^{(k)}, p^{(k)}) + \frac{1}{2}\alpha_k^2 (Ap^{(k)}, p^{(k)}) \\ &\quad - (b, x^{(k)}) - \alpha_k (b, p^{(k)}) \\ &= f(x^{(k)}) + \frac{1}{2}\alpha_k^2 (Ap^{(k)}, p^{(k)}) - \alpha_k (p^{(k)}, r^{(k)}). \end{aligned}$$

Καταλήγουμε σε ένα τριώνυμο ως προς α_k . Η ελάχιστη τιμή του συναρτησιακού δίνεται για $\alpha = \alpha_k = \frac{(p^{(k)}, r^{(k)})}{(Ap^{(k)}, p^{(k)})}$ και είναι η

$$f(x^{(k+1)}) = f(x^{(k)}) - \frac{(p^{(k)}, r^{(k)})^2}{2(Ap^{(k)}, p^{(k)})} < f(x^{(k)}),$$

αν και μόνο αν $(p^{(k)}, r^{(k)}) \neq 0$.

Παρακάτω παρατίθεται ο αλγόριθμος της Μεθόδου Γενικών Διευθύνσεων.

Αλγόριθμος 2.2 Μέθοδος Γενικών Διευθύνσεων

Δεδομένα: $A \in \mathbb{C}^{n,n}$, $A^H = A$, A θετικά ορισμένος, $b \in \mathbb{C}^n$, $\varepsilon \in \mathbb{R}^+$ η μέγιστη επιτρεπτή ανοχή σφάλματος.

Αποτέλεσμα: $x = x^{(k)}$ η προσέγγιση της λύσης.

- 1: $x^{(0)} = 0$
- 2: $r^{(0)} = b$
- 3: $k = 0$
- 4: **Εφόσον** $\|r^{(k)}\| > \varepsilon$
- 5: *Επέλεξε* διεύθυνση $p^{(k)}$: $(p^{(k)}, r^{(k)}) \neq 0$
- 6: $\alpha_k = \frac{(p^{(k)}, r^{(k)})}{(Ap^{(k)}, p^{(k)})}$
- 7: $x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}$
- 8: $r^{(k+1)} = b - Ax^{(k+1)}$
- 9: $k = k + 1$
- 10: **Τέλος** 'Εφόσον'

Το κόστος του αλγορίθμου είναι $O(n^2)$ πράξεις ανά βήμα. Όπως και στον αλγόριθμο της Μεθόδου Απότομης Καθόδου υπάρχουν δύο πολλαπλασιασμοί πίνακα επί διάνυσμα, ο $Ax^{(k+1)}$ και ο $Ap^{(k)}$. Αυτοί οι δύο μπορούν να αναχθούν σε έναν, αν παρατηρήσουμε ότι $r^{(k+1)} = b - Ax^{(k+1)} = b - A(x^{(k)} + \alpha_k p^{(k)}) = r^{(k)} - \alpha_k Ap^{(k)}$, οπότε και για τον υπολογισμό του $r^{(k+1)}$ δεν χρειάζεται να κάνουμε τον πολλαπλασιασμό $Ax^{(k+1)}$, αλλά χρησιμοποιούμε το $Ap^{(k)}$ που έχουμε ήδη υπολογίσει.

2.3 Γενικευμένη Μέθοδος Συζυγών Υπολοίπων (Generalized Conjugate Residual - GCR) - Μέθοδος Συζυγών Διευθύνσεων (Conjugate Direction - CD) Μέθοδος Συζυγών Κλίσεων (Conjugate Gradient - CG)

Οι μέθοδοι ελαχιστοποίησης που παρουσιάστηκαν έχουν κάτι κοινό. Σε κάθε βήμα της μεθόδου η προσεγγιστική λύση μπορεί να γραφεί ως γραμμικός συνδυασμός της αρχικής προσέγγισης $x^{(0)}$ και των διευθύνσεων που έχουν χρησιμοποιηθεί κάθε φορά μέχρι το συγκεκριμένο βήμα της μεθόδου. Έτσι, για τις

Μεθόδους Orthomin(1) και Απότομης Καθόδου ισχύει

$$\begin{aligned} x^{(k)} &= x^{(k-1)} + \alpha_{k-1}r^{(k-1)} = x^{(k-2)} + \alpha_{k-2}r^{(k-2)} + \alpha_{k-1}r^{(k-1)} \\ &= x^{(k-3)} + \alpha_{k-3}r^{(k-3)} + \alpha_{k-2}r^{(k-2)} + \alpha_{k-1}r^{(k-1)} \\ &= \dots = x^{(0)} + \sum_{i=1}^k \alpha_{i-1}r^{(i-1)}. \end{aligned}$$

Αντίστοιχα, για τη Μέθοδο Γενικών Διευθύνσεων ισχύει

$$x^{(k)} = x^{(0)} + \sum_{i=1}^k \alpha_{i-1}p^{(i-1)}.$$

Οπότε και στις δύο περιπτώσεις η διαφορά $x^{(k)} - x^{(0)}$ μπορεί να γραφεί ως γραμμικός συνδυασμός των διευθύνσεων. Για $k \leq n$ τα διανύσματα αυτά είναι γραμμικά ανεξάρτητα. Όμως για $k > n$ τα διανύσματα είναι προφανώς γραμμικά εξαρτημένα, αφού η διάσταση του \mathbb{C}^n είναι n .

Για $k \leq n$ είναι δυνατόν να βρεθεί ένα σύνολο k γραμμικά ανεξάρτητων διανυσμάτων, τα οποία είναι ορθογώνια μεταξύ τους. Γι' αυτόν τον σκοπό χρησιμοποιείται συνήθως ο αλγόριθμος ορθογωνοποίησης Gram-Schmidt. Οι μέθοδοι που προκύπτουν με αυτή την επιλογή των διανυσμάτων παρουσιάζονται στη συνέχεια. Σ' αυτές (όπως άλλωστε και στη Μέθοδο Γενικών Διευθύνσεων) αν $x^{(k)}$ είναι η προσέγγιση της λύσης στην k επανάληψη, η επόμενη προσέγγιση ορίζεται να είναι η

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}.$$

Για το διάνυσμα υπόλοιπο της $k+1$ επανάληψης ισχύει

$$\begin{aligned} r^{(k+1)} &= b - Ax^{(k+1)} \\ &= b - A(x^{(k)} + \alpha_k p^{(k)}) \\ &= b - Ax^{(k)} - \alpha_k Ap^{(k)} \\ &= r^{(k)} - \alpha_k Ap^{(k)}. \end{aligned}$$

Τέλος, το διάνυσμα - σφάλμα είναι:

$$\varepsilon^{(k+1)} = x - x^{(k+1)} = x - x^{(k)} - \alpha_k p^{(k)} = \varepsilon^{(k)} - \alpha_k p^{(k)}.$$

Η Γενικευμένη Μέθοδος Συζυγών Υπολοίπων παράγεται αν απαιτήσουμε το $r^{(k+1)}$ να είναι ορθογώνιο στο $Ap^{(k)}$ ή ισοδύναμα, τα $Ap^{(k)}$ να είναι ορθογώνια

μεταξύ των $((Ap^{(k)}, Ap^{(j)}) = 0, k \neq j)$. Τότε, τα α_k δίνονται ως $\alpha_k = \frac{(r^{(k)}, Ap^{(k)})}{(Ap^{(k)}, Ap^{(k)})}$. Παρακάτω παρατίθεται ο αλγόριθμος της Μεθόδου Συζυγών Υπολοίπων.

Αλγόριθμος 2.3 Γενικευμένη Μέθοδος Συζυγών Υπολοίπων

Δεδομένα: $A \in \mathbb{C}^{n,n}$, $A^H = A$, A θετικά ορισμένος, $b \in \mathbb{C}^n$, $\varepsilon \in \mathbb{R}^+$ η μέγιστη επιτρεπτή ανοχή σφάλματος

Αποτέλεσμα: $x = x^{(k)}$ η προσέγγιση της λύσης

- 1: $x^{(0)} = 0$
- 2: $r^{(0)} = b$
- 3: $p^{(0)} = r^{(0)}$
- 4: $k = 0$
- 5: **Εφόσον** $\|r^{(k)}\| > \varepsilon$
- 6: $\alpha_k = \frac{(r^{(k)}, Ap^{(k)})}{(Ap^{(k)}, Ap^{(k)})}$
- 7: $x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}$
- 8: $r^{(k+1)} = b - Ax^{(k+1)}$
- 9: **Για** $i = 0, \dots, k$
- 10: $\beta_{ik} = -\frac{(Ar^{(k+1)}, Ap^{(i)})}{(Ap^{(i)}, Ap^{(i)})}$
- 11: **Τέλος ‘Για’**
- 12: $p^{(k+1)} = r^{(k+1)} + \sum_{i=0}^k \beta_{ik} p^{(i)}$
- 13: $k = k + 1$
- 14: **Τέλος ‘Εφόσον’**

Η Μέθοδος Συζυγών Διευθύνσεων προκύπτει αν απαιτήσουμε το $\varepsilon^{(k+1)}$ να είναι A - ορθογώνιο στο $p^{(k)}$ ή ισοδύναμα, οι διευθύνσεις $p^{(k)}$ να είναι A - ορθογώνιες μεταξύ των $((Ap^{(k)}, p^{(j)}) = 0, k \neq j)$. Αντικαθιστώντας το $\varepsilon^{(k+1)} = \varepsilon^{(k)} - \alpha_k p^{(k)}$ στην απαίτηση $(A\varepsilon^{(k+1)}, p^{(k)}) = 0$ παίρνουμε το α_k ως

$$\begin{aligned} \alpha_k &= \frac{(A\varepsilon^{(k)}, p^{(k)})}{(Ap^{(k)}, p^{(k)})} = \frac{(A(x - x^{(k)}), p^{(k)})}{(Ap^{(k)}, p^{(k)})} \\ &= \frac{(b - Ax^{(k)}, p^{(k)})}{(Ap^{(k)}, p^{(k)})} = \frac{(r^{(k)}, p^{(k)})}{(Ap^{(k)}, p^{(k)})}. \end{aligned}$$

Θεώρημα 2.3.1 (Gram-Schmidt). Έστω $A \in \mathbb{C}^{n,n}$, $A^H = A$ και A θετικά ορισμένος. Έστω ακόμη πίνακας $U \in \mathbb{C}^{n,k}$, $k \leq n$, τα διανύσματα - στήλες w^j ,

$j = 1(1)k$, του οποίου είναι γραμμικά ανεξάρτητα. Υπάρχουν μοναδικοί πίνακες $P \in \mathbb{C}^{n,k}$, με $P^T A P = D = \text{diag}(d_1, d_2, \dots, d_k)$, με $d_j > 0$, $j = 0(1)k - 1$, και άνω τριγωνικός $R \in \mathbb{C}^{k,k}$ με $r_{j,j} = 1$, $j = 0(1)k - 1$, τ.ω.

$$U = PR.$$

Σημείωση. Αν $p^{(j)}$, $j = 1(1)k$, είναι τα διανύσματα - στήλες του P , τότε η σχέση $P^T A P = D = \text{diag}(d_1, d_2, \dots, d_k)$, με $d_j > 0$, $j = 1(1)k$, δεν αποτελεί παρά την απαίτησή μας τα $p^{(j)}$ να είναι μή μηδενικά και ανά δύο A - ορθογώνια (ή A - συζυγή), οπότε θα είναι και γραμμικά ανεξάρτητα. Για την απόδειξη αυτού του ισχυρισμού, έστω ένας γραμμικός συνδυασμός k διανυσμάτων

$$\sum_{i=0}^{k-1} c_i p^{(i)} = 0, \text{ όπου } c_i \in \mathbb{C}, i = 0(1)k - 1.$$

Θεωρώντας εσωτερικό γινόμενο της σχέσης αυτής με $Ap^{(j)}$ έχουμε ότι

$$\sum_{i=0}^{k-1} c_i (Ap^{(j)}, p^{(i)}) = 0.$$

Οπότε, λόγω της A - ορθογωνιότητας των $p^{(i)}$ και $p^{(j)}$ για $i \neq j$ αναπτύσσοντας το άθροισμα έχουμε: $c_j (Ap^{(j)}, p^{(j)}) = 0$, αφού οι υπόλοιποι όροι του αθροίσματος είναι 0. Τελικά, αφού ο A είναι θετικά ορισμένος, δηλαδή $(Ap^{(j)}, p^{(j)}) > 0$, προκύπτει ότι $c_j = 0$, $j = 0(1)k - 1$, δηλαδή ότι τα $p^{(i)}$, $i = 0(1)k - 1$, είναι γραμμικά ανεξάρτητα.

Από τη σχέση $U = PR$ μπορούμε να προσδιορίσουμε τα διανύσματα στήλες $p^{(j)}$, $j = 0(1)k - 1$ καθώς και τα στοιχεία r_{ij} , $i = 0(1)k - 2$, $j = 0(1)k - 1$. Η σχέση $U = PR$ γράφεται ως:

$$[u^0 \ u^1 \ u^2 \ \dots \ u^{k-1}] = [p^{(0)} \ p^{(1)} \ p^{(2)} \ \dots \ p^{(k-1)}] \begin{bmatrix} 1 & r_{0,1} & r_{0,2} & \dots & r_{0,k-1} \\ & 1 & r_{1,2} & \dots & r_{1,k-1} \\ & & 1 & \dots & r_{2,k-1} \\ & & & \ddots & \vdots \\ & & & & 1 \end{bmatrix}$$

Μετά από στοιχειώδεις πράξεις προκύπτει ότι:

$$p^{(j)} = u^j - \sum_{i=0}^{j-1} r_{i,j} p^{(i)} \text{ και } r_{i+1,j+1} = \frac{(Ap^{(i)}, u^j)}{(Ap^{(i)}, p^{(i)})}, \text{ με } i = 0(1)j-1, j = 0(1)k-1.$$

Παρακάτω παρατίθεται ο αλγόριθμος Gram - Schmidt A - Ορθογωνοποίησης καθώς και μια τροποποίησή του. Στη συνέχεια δίνεται ο πλήρης αλγόριθμος της Μεθόδου Συζυγών Διευθύνσεων.

Αλγόριθμος 2.4 QR Ανάλυση με Gram - Schmidt A - Ορθογωνοποίηση

Δεδομένα: $U \in \mathbb{C}^{n,m}$ με διανύσματα στήλες u^i , $i = 0(1)m - 1$, γραμμικά ανεξάρτητα.

- 1: **Για** $i = 0(1)m - 1$
- 2: $p^{(i)} = u^i$
- 3: **Για** $j = 0(1)i - 1$
- 4: $r_{ji} = \frac{(Ap^{(j)}, u^i)}{(Ap^{(j)}, p^{(j)})}$
- 5: $p^{(i)} = p^{(i)} - r_{ji}p^{(j)}$
- 6: **Τέλος 'Για'**
- 7: **Τέλος 'Για'**

Πολλές φορές για να επιτύχουμε την QR ανάλυση του πίνακα A ή να κατασκευάσουμε τον πίνακα P με A - ορθογώνια διανύσματα, αντί του κλασικού αλγορίθμου Gram - Schmidt ορθογωνοποίησης χρησιμοποιούμε του τροποποιημένου αλγορίθμου Gram - Schmidt ορθογωνοποίησης (Modified Gram - Schmidt) ή A - ορθογωνοποίησης. Οι δύο αλγόριθμοι είναι ισοδύναμοι μαθηματικά και χρησιμοποιούνται για τον ίδιο ακριβώς σκοπό. Στον κλασικό αλγόριθμο που δώσαμε, για τον υπολογισμό των r_{ji} υπολογίζουμε το $(Ap^{(j)}, u^i)$. Στο u^i όμως περιέχονται στον γραμμικό συνδυασμό και τα $p^{(k)}$, $k = 0(1)j - 1$, που είναι A - ορθογώνια με το $p^{(j)}$. Αν στη θέση του u^i θέσουμε την τρέχουσα τιμή $p^{(i)}$, τότε $(Ap^{(j)}, p^{(i)}) = (Ap^{(j)}, u^i)$. Το $p^{(i)}$ αυτό είναι απαλλαγμένο από τις συνιστώσες $p^{(0)}, p^{(1)}, \dots, p^{(j-1)}$. Για τον λόγο αυτό ο τροποποιημένος αλγόριθμος είναι πιο ευстаθής από τον κλασικό, διότι στον κλασικό ουσιαστικά υπολογίζονται και τα $(Ap^{(j)}, p^{(k)})$, $k = 0(1)j - 1$, που λόγω σφαλμάτων στρογγύλευσης δεν είναι ακριβώς μηδέν. Η διαφορά του τροποποιημένου αλγορίθμου από τον κλασικό βρίσκεται μόνο στον υπολογισμό των r_{ji} . Στον τροποποιημένο θα γίνει $\frac{(Ap^{(j)}, p^{(i)})}{(Ap^{(j)}, p^{(j)})}$.

Αλγόριθμος 2.5 Μέθοδος Συζυγών Διευθύνσεων

Δεδομένα: $A \in \mathbb{C}^{n,n}$, $A^H = A$, A θετικά ορισμένος, $b \in \mathbb{C}^n$, $U \in \mathbb{C}^{n,n}$, με διανύσματα στήλες u^j , $j = 0(1)n - 1$, γραμμικά ανεξάρτητα, $\varepsilon \in \mathbb{R}^+$ η μέγιστη επιτρεπτή ανοχή σφάλματος.

Αποτέλεσμα: $x = x^{(k)}$ η προσέγγιση της λύσης.

- 1: $x^{(0)} = 0$
- 2: $r^{(0)} = b$
- 3: $k = 0$
- 4: **Εφόσον** $\|r^{(k)}\| > \varepsilon$
- 5: $p^{(k)} = u^k$
- 6: **Για** $j = 0(1)k - 1$
- 7: $r_{jk} = \frac{(Ap^{(j)}, p^{(k)})}{(Ap^{(j)}, p^{(j)})}$
- 8: $p^{(k)} = p^{(k)} - r_{jk}p^{(j)}$
- 9: **Τέλος 'Για'**
- 10: $\alpha_k = \frac{(p^{(k)}, r^{(k)})}{(Ap^{(k)}, p^{(k)})}$
- 11: $x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}$
- 12: $r^{(k+1)} = b - Ax^{(k)} (= r^{(k)} - \alpha_k Ap^{(k)})$
- 13: $k = k + 1$
- 14: **Τέλος 'Εφόσον'**

Στην περίπτωση που ο πίνακας A είναι ερμιτιανός και θετικά ορισμένος η Μέθοδος Συζυγών Διευθύνσεων επεκτείνεται και λαμβάνουμε τη Μέθοδο Συζυγών Κλίσεων. Η Μέθοδος Συζυγών Κλίσεων είναι μια μέθοδος συζυγών διευθύνσεων με τη διαφορά ότι ως στήλες του πίνακα U στον αλγόριθμο Gram - Schmidt λαμβάνονται τα διαδοχικά διανύσματα - υπόλοιπο $u^k = r^{(k)}$, $k = 0(1)n - 1$. Επίσης, για τον προσδιορισμό της επόμενης διεύθυνσης $p^{(k+1)}$ αποδεικνύεται ότι χρειάζεται μόνο η αμέσως προηγούμενη διεύθυνση $p^{(k)}$ και όχι όλες οι προηγούμενες διευθύνσεις, όπως συμβαίνει στη Μέθοδο Συζυγών Διευθύνσεων.

Στην $k + 1$ επανάληψη της Μεθόδου Συζυγών Κλίσεων ισχύει

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}$$

και για το αντίστοιχο διάνυσμα υπόλοιπο

$$r^{(k+1)} = r^{(k)} - \alpha_k Ap^{(k)}.$$

Αφού δύο διαδοχικά διανύσματα - σφάλμα $r^{(k)}$ είναι ορθογώνια μεταξύ τους, δηλαδή $(r^{(k+1)}, r^{(k)}) = 0$, αντικαθιστώντας το $r^{(k+1)}$ με $r^{(k)} - \alpha_k Ap^{(k)}$ και λύνοντας ως προς α_k έχουμε ότι

$$\alpha_k = \frac{(r^{(k)}, r^{(k)})}{(Ap^{(k)}, r^{(k)})}.$$

Ο ίδιος τύπος προκύπτει αν απαιτήσουμε οι διευθύνσεις $p^{(k)}$ να είναι A - ορθογώνιες μεταξύ τους.

Για τον προσδιορισμό των β_k εργαζόμαστε με παρόμοιο τρόπο. Αφού η διεύθυνση $p^{(k+1)}$ είναι ορθογώνια στο $Ap^{(k)}$, δηλαδή $(p^{(k+1)}, Ap^{(k)}) = 0$, αντικαθιστώντας το $p^{(k+1)}$ με $r^{(k+1)} + \beta_k p^{(k)}$ έχουμε ότι

$$\begin{aligned} (r^{(k+1)} + \beta_k p^{(k)}, Ap^{(k)}) = 0 &\Leftrightarrow \beta_k = \frac{(r^{(k+1)}, Ap^{(k)})}{(p^{(k)}, Ap^{(k)})} \\ &= \frac{\left(r^{(k+1)}, \frac{1}{\alpha_k} (r^{(k)} - r^{(k+1)}) \right)}{(Ap^{(k)}, p^{(k)})}, \end{aligned}$$

αφού το $Ap^{(k)}$ δίνεται συναρτήσει των $r^{(k)}$ και $r^{(k+1)}$ ως

$$Ap^{(k)} = \frac{1}{\alpha_k} (r^{(k)} - r^{(k+1)}).$$

Οπότε τελικά προκύπτει ότι

$$\beta_k = \frac{(r^{(k+1)}, r^{(k+1)})}{(r^{(k)}, r^{(k)})}.$$

Κατά τη Μέθοδο Συζυγών Κλίσεων, κάθε νέα διεύθυνση, δηλαδή η $p^{(k+1)}$ αποτελεί την ορθή προβολή του $r^{(k+1)}$ στον υπόχωρο $\text{span}\{Ap^{(1)}, \dots, Ap^{(k)}\}$, $k = 1(1)n - 1$. Επίσης, το διάνυσμα σφάλμα $\varepsilon^{(k)}$ είναι A -ορθογώνιο προς την αμέσως προηγούμενη διεύθυνση $p^{(k-1)}$. Δηλαδή στην $k + 1$ επανάληψη ισχύει $(\varepsilon^{(k+1)}, Ap^{(k)}) = 0$. Επιτυγχάνεται ελαχιστοποίηση της A -νόρμας του διανύσματος - σφάλμα στον χώρο $\varepsilon^{(k)} + \text{span}\{r^{(k)}, p^{(k-1)}\}$. Παρακάτω δίνεται ο αλγόριθμος της Μεθόδου Συζυγών Κλίσεων. Συμπερασματικά, μπορούμε να πούμε ότι είναι πανομοιότυπος με αυτόν της Μεθόδου Συζυγών Διευθύνσεων, με μόνη διαφορά ότι στο σημείο όπου βρίσκουμε την επόμενη διεύθυνση $p^{(k)}$ αυτή βρίσκεται συναρτήσει μόνο της προηγούμενης διεύθυνσης $p^{(k-1)}$ και όχι συναρτήσει όλων των προηγούμενων διευθύνσεων. Το γεγονός αυτό καθιστά τη μέθοδο πιο ευσταθή και ταχύτερη από εκείνη των Συζυγών Διευθύνσεων.

Αλγόριθμος 2.6 Μέθοδος Συζυγών Κλίσεων

Δεδομένα: $A \in \mathbb{C}^{n,n}$, $A^H = A$, θετικά ορισμένος, $b \in \mathbb{C}^n$, $\varepsilon \in \mathbb{R}^+$ η μέγιστη επιτρεπτή ανοχή σφάλματος.

- 1: $x^{(0)} = 0$
- 2: $r^{(0)} = b$
- 3: $p^{(0)} = r^{(0)}$
- 4: $\alpha_0 = \frac{(r^{(0)}, r^{(0)})}{(Ap^{(0)}, p^{(0)})}$

- 5: $x^{(1)} = x^{(0)} + \alpha_0 p^{(0)}$
- 6: $r^{(1)} = b - Ax^{(1)} (= r^{(0)} - \alpha_0 Ap^{(0)})$
- 7: $k = 1$
- 8: **Εφόσον** $\|r^{(k)}\| > \varepsilon$
- 9: $\beta_{k-1} = \frac{(r^{(k)}, r^{(k)})}{(r^{(k-1)}, r^{(k-1)})}$
- 10: $p^{(k)} = r^{(k)} + \beta_{k-1} p^{(k-1)}$
- 11: $\alpha_k = \frac{(r^{(k)}, r^{(k)})}{(Ap^{(k)}, p^{(k)})}$
- 12: $x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}$
- 13: $r^{(k+1)} = b - Ax^{(k+1)} (= r^{(k)} - \alpha_k Ap^{(k)})$
- 14: $k = k + 1$
- 15: **Τέλος ‘Εφόσον’**

Είναι γνωστό ότι η Μέθοδος Συζυγών Κλίσεων δίνει (θεωρητικά) ακριβώς τη λύση σε n επαναλήψεις. Στην πράξη μας ενδιαφέρει το κριτήριο σύγκλισης ($\|r^{(k)}\| \leq \varepsilon$), που αποδεικνύεται ότι επιτυγχάνεται πολύ πιο πριν το k γίνει ίσο με το n σε προβλήματα με καλή κατάσταση. Σε προβλήματα με κακή κατάσταση προτιμάται η τεχνική της προρρύθμισης που μετατρέπει το πρόβλημα σε ισοδύναμο, με καλή κατάσταση. Στην τεχνική της προρρύθμισης χρησιμοποιείται ένας προρρυθμιστής πίνακας $M \in \mathbb{C}^{n,n}$ συμμετρικός και θετικά ορισμένος, οικονομικά αντιστρέψιμος σε σχέση με τον A και τέτοιος ώστε $\frac{\lambda_{\max}(M^{-1}A)}{\lambda_{\min}(M^{-1}A)} \ll \kappa_2(A)$, όπου με $\lambda_{\max}, \lambda_{\min}$ αντίστοιχα συμβολίζεται η μικρότερη και η μεγαλύτερη ιδιοτιμή του πίνακα $M^{-1}A$ που αποδεικνύεται ότι είναι θετικές. Πολλαπλασιάζοντας το αρχικό σύστημα με M^{-1} έχουμε:

$$M^{-1}Ax = M^{-1}b$$

Ο $M^{-1}A$ δεν είναι συμμετρικός και άρα δεν μπορεί να εφαρμοστεί η Μέθοδος Συζυγών Κλίσεων. Η ύπαρξη όμως ενός συμμετρικού και θετικά ορισμένου πίνακα C τέτοιου ώστε $C^2 = M$ μετατρέπει το παραπάνω σύστημα σε ισοδύναμό του της μορφής:

$$\overline{A}\overline{x} = \overline{b}, \text{ όπου } \overline{A} = C^{-1}AC^{-1}, \overline{x} = Cx, \overline{b} = C^{-1}b.$$

Σε αυτό το σύστημα ο πίνακας \overline{A} είναι συμμετρικός. Οπότε, για το νέο αυτό σύστημα είναι εύκολο να εφαρμοστεί ο αλγόριθμος της Μεθόδου Συζυγών Κλίσεων, αφού ικανοποιούνται οι προϋποθέσεις του. Ως προρρυθμιστές, λόγω της φύσης του πίνακα A , επιλέγονται συμμετρικοί και θετικά ορισμένοι πίνα-

κες. Αυτοί οι προρρυθμιστές εμφανίζονται και στις κλασικές επαναληπτικές μεθόδους, δεδομένου ότι είναι συμμετρικοί και θετικά ορισμένοι πίνακες. Μπορούμε λόγω χάριν, να επιλέξουμε τον $M = \text{diag}(\alpha_{11}, \alpha_{22}, \dots, \alpha_{nn})$ και τότε η μέθοδος που αντιστοιχεί είναι η Jacobi - Συζυγών Κλίσεων (Jacobi - CG) ή τον $M = \frac{1}{\omega(2-\omega)}(D - \omega L)D^{-1}(D - \omega L^T)$ και τότε η μέθοδος που αντιστοιχεί είναι η Συμμετρική SOR - Συζυγών Κλίσεων (SSOR - CG) και χρησιμοποιείται συχνότερα με παράμετρο $\omega = 1$.

Στη συνέχεια θα ασχοληθούμε με την εκτίμηση σφάλματος της Μεθόδου Συζυγών Κλίσεων. Θα προσπαθήσουμε δηλαδή να εκτιμήσουμε τη διαφορά της προσεγγιστικής λύσης, $x^{(m)}$, από την ακριβή λύση, x , του συστήματος συναρτήσει της διαφοράς της αρχικής προσέγγισης, $x^{(0)}$, από τη λύση x . Συγκεκριμένα, θα δοθεί μια εκ των προτέρων (a priori) εκτίμηση (ή ισοδύναμα ένα άνω φράγμα) του σφάλματος. Το διάνυσμα - σφάλμα εκφράζεται ως γραμμικός συνδυασμός των διανυσμάτων $\{\varepsilon^{(0)}, A\varepsilon^{(0)}, \dots, A^k\varepsilon^{(0)}\}$ Δηλαδή έχουμε ότι

$$\varepsilon^{(k)} = \sum_{i=1}^n A^i \varepsilon^{(0)} = p_m(A)\varepsilon^{(0)},$$

όπου με p_m συμβολίζουμε ένα πολυώνυμο βαθμού το πολύ m . Ζητούμενο είναι η ελαχιστοποίηση της νόρμας ενέργειας του διανύσματος - σφάλμα. Το πρόβλημά μας ανάγεται σε ένα διακριτό min - max πρόβλημα της μορφής

$$\min_{\substack{p_m \in P_m \\ p_m(0)=1}} \max_{i=1(1)n} |p_m(\lambda_i)| \leq \min_{\substack{p_m \in P_m \\ p_m(0)=1}} \max_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} |p_m(\lambda)|$$

στον χώρο P_m των πολυωνύμων βαθμού το πολύ m . Απάντηση σε τέτοια προβλήματα δίνει η θεωρία προσέγγισης και συγκεκριμένα τα πολυώνυμα του Chebyshev.

Τα πολυώνυμα Chebyshev πρώτου είδους, βαθμού m , ορίζονται ως

$$T_m(t) = \cos[m \cdot \arccos(t)], \quad t \in [-1, 1].$$

Αν υποθέσουμε ότι αναφερόμαστε σε πραγματικά τόξα και όχι σε (γενικά) μιγαδικά, θέτουμε

$$\arccos(t) = \theta \ (\in \mathbb{R}) \iff t = \cos(\theta)$$

και τότε ισχύει προφανώς $-1 \leq \cos(\theta) \leq 1 \iff \theta \in [0, \pi]$. Οπότε, ισοδύναμα, τα πολυώνυμα Chebyshev ορίζονται από τη σχέση

$$T_m(t) = \cos(m \cdot \theta), \quad \theta \in [0, \pi].$$

Ισχύει προφανώς ότι $|T_m(t)| = |\cos(m \cdot \theta)| \leq 1$, δηλαδή τα πολυώνυμα φράσσονται κατά απόλυτη τιμή από τη μονάδα.

Στην περίπτωση μιγαδικών τόξων ($\theta \in \mathbb{C}$) ο ορισμός των πραγματικών πολυωνύμων Chebyshev επεκτείνεται εύκολα ως εξής, δίνοντας τα μιγαδικά πολυώνυμα Chebyshev

$$T_m(t) = \cos h [m \cdot \operatorname{arccosh}(t)], \quad t \geq 1.$$

Θέτοντας

$$\operatorname{arccosh}(t) = \theta \Leftrightarrow t = \cosh(\theta),$$

έχουμε $t \geq 1 \Leftrightarrow \cosh(\theta) \geq 1 \Leftrightarrow \theta \geq 0$. Αντίστοιχα, σε αυτή την περίπτωση ισχύει ότι $|T_m(t)| > 1$.

Τα πολυώνυμα του Chebyshev μπορούν να προκύψουν εναλλακτικά μέσω μιας αναδρομικής σχέσης.

Για $m = 0$ έχουμε $T_0(t) = 1$.

Για $m = 1$ έχουμε $T_1(t) = t$.

Τότε, κάθε επόμενο πολυώνυμο ικανοποιεί τη σχέση

$$T_{m+1}(t) - 2t \cdot T_m(t) + T_{m-1}(t) = 0, \quad m = 1, 2, 3, 4, \dots$$

Η απόδειξη αυτού του ισχυρισμού δίνεται μέσω της τριγωνομετρικής ταυτότητας

$$\cos(m+1)\theta + \cos(m-1)\theta = 2\cos(\theta)\cos(m\theta)$$

Θέτοντας $\theta = \arccos(t) \iff t = \cos(\theta)$ έχουμε

$$\begin{aligned} & \cos(m+1)\arccos(t) + \cos(m-1)\arccos(t) \\ &= 2\cos(\arccos(t))\cos(m\arccos(t)) \\ &\iff T_{m+1}(\theta) - 2mT_m(\theta) + T_{m-1}(\theta) = 0 \end{aligned}$$

Για την νόρμα ενέργειας του διανύσματος - σφάλμα στην m επανάληψη της Μεθόδου Συζυγών Κλίσεων ισχύει η ακόλουθη εκτίμηση.

Θεώρημα 2.3.2. Αν $x^{(0)} \in \mathbb{C}^n$ είναι η αρχική προσέγγιση της μεθόδου συζυγών κλίσεων και $x^{(m)} \in \mathbb{C}^n$ είναι η προσεγγιστική λύση στο m βήμα της μεθόδου για τη λύση x του συστήματος $Ax = b$, $A \in \mathbb{C}^{n,n}$, με A συμμετρικό και θετικά ορισμένο και $\kappa = \frac{\lambda_{\max}}{\lambda_{\min}}$ ο δείκτης κατάστασης του συστήματος, όπου $0 < \lambda_1 \equiv \lambda_{\min} \leq \lambda_2 \leq \dots \leq \lambda_n \equiv \lambda_{\max}$ οι n το πλήθος ιδιοτιμές του A , τότε για το διάνυσμα σφάλμα στην m επανάληψη της μεθόδου ισχύει:

$$\|\varepsilon^{(m)}\|_{A^{\frac{1}{2}}} \leq 2 \left[\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1} \right)^m + \left(\frac{\sqrt{\kappa}+1}{\sqrt{\kappa}-1} \right)^m \right]^{-1} \|\varepsilon^{(0)}\|_{A^{\frac{1}{2}}}, \quad m = 1(1)n-1.$$

Απόδειξη. Έστω $\varepsilon^{(m)} = x - x^{(m)}$ το διάνυσμα - σφάλμα στο βήμα m της μεθόδου. Για το αρχικό διάνυσμα - σφάλμα της Μεθόδου Συζυγών Κλίσεων ισχύει $\varepsilon^{(0)} = x - x^{(0)} = \sum_{i=1}^n \xi_i x^i$, δηλαδή αυτό γράφεται ως γραμμικός συνδυασμός των ιδιοδιανυσμάτων του A . Ζητούμενο είναι, όπως έχει ήδη αναφερθεί, η ελαχιστοποίηση της νόρμας ενέργειας του διανύσματος - σφάλμα, δηλαδή ζητάμε το

$$\min_{\substack{p_m \in P_m \\ p(0)=1}} \|p_m(A)\varepsilon^{(0)}\|_{A^{\frac{1}{2}}}.$$

Οπότε, η ελαχιστοποίηση της νόρμας ενέργειας γίνεται στον χώρο των πολυωνύμων βαθμού το πολύ m . Από εδώ και στο εξής το πολυώνυμο p_m θα συμβολίζεται απλά ως p . Ισχύει ότι

$$\begin{aligned} \|p(A)\varepsilon^{(0)}\|_{A^{\frac{1}{2}}}^2 &= \left(Ap(A)\varepsilon^{(0)}, p(A)\varepsilon^{(0)} \right) \\ &= \left(\sum_{i=1}^n \xi_i Ap(A)x^i, \sum_{i=1}^n \xi_i p(A)x^i \right) \\ &= \sum_{i=1}^n \xi_i^2 (\lambda_i p(\lambda_i)x^i, p(\lambda_i)x^i) \\ &= \sum_{i=1}^n \xi_i^2 \lambda_i p(\lambda_i)^2 \leq \max_i p(\lambda_i)^2 \sum_{i=1}^n \xi_i^2 \lambda_i \\ &= \max_i p(\lambda_i)^2 \left(A \sum_{i=1}^n \xi_i x^i, \sum_{i=1}^n \xi_i x^{(i)} \right) \\ &= \max_i p(\lambda_i)^2 \|\varepsilon^{(0)}\|_{A^{\frac{1}{2}}}^2 \\ &= \max_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} (p(\lambda))^2 \|\varepsilon^{(0)}\|_{A^{\frac{1}{2}}}^2 \end{aligned}$$

Οπότε, ζητάμε να λύσουμε ένα πρόβλημα της μορφής

$$\|\varepsilon^{(m)}\|_{A^{\frac{1}{2}}} \leq \min_{\substack{p \in P_m \\ p(0)=1}} \max_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} |p(\lambda)| \|\varepsilon^{(0)}\|_{A^{\frac{1}{2}}}.$$

Το τελευταίο αποτελεί ένα συνεχές min-max πρόβλημα. Τη λύση στο πρόβλημα αυτό θα μας δώσουν τα πολυώνυμα Chebyshev πρώτου είδους στο διάστημα $[\lambda_{\min}, \lambda_{\max}]$. Αποδεικνύεται ότι το πολυώνυμο που δίνει απάντηση στο πρόβλημά

μας είναι το

$$p = p_m(t) = \frac{T_m\left(\frac{2t - \lambda_{\max} - \lambda_{\min}}{\lambda_{\max} - \lambda_{\min}}\right)}{T_m\left(\frac{-\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} - \lambda_{\min}}\right)}.$$

Ο αριθμητής φράσσεται από τη μονάδα, αφού

$$-1 < \frac{2t - \lambda_{\max} - \lambda_{\min}}{\lambda_{\max} - \lambda_{\min}} < 1, \quad \forall t \in [\lambda_{\min}, \lambda_{\max}].$$

Για τον παρονομαστή, αφού $\lambda_{\max} - \lambda_{\min} > 0 \iff \frac{-\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} - \lambda_{\min}} < -1$, σε αυτήν την περίπτωση το πολυώνυμο Chebyshev ορίζεται από τον τύπο $T_m(t) = \cosh[\operatorname{arccosh}(t)]$, $|t| < 1$, και έχουμε τελικά ότι

$$\begin{aligned} \|\varepsilon^{(m)}\|_{A^{\frac{1}{2}}}^2 &\leq \left[T_m\left(\frac{\lambda_{\max} + \lambda_{\min}}{\lambda_{\max} - \lambda_{\min}}\right) \right]^{-1} \|\varepsilon^{(0)}\|_{A^{\frac{1}{2}}}^2 \\ &= \left[T_m\left(\frac{\kappa + 1}{\kappa - 1}\right) \right]^{-1} \|\varepsilon^{(0)}\|_{A^{\frac{1}{2}}}^2 \\ &= 2 \left[\left(\frac{\sqrt{\kappa} + 1}{\sqrt{\kappa} - 1}\right)^m + \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}\right)^m \right]^{-1} \|\varepsilon^{(0)}\|_{A^{\frac{1}{2}}}^2 \end{aligned}$$

και αφού $\left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}\right)^m \rightarrow 0$ προκύπτει τελικά ότι

$$\|\varepsilon^{(m)}\|_{A^{\frac{1}{2}}} \leq 2 \left[\left(\frac{\sqrt{\kappa} + 1}{\sqrt{\kappa} - 1}\right) \right]^m \|\varepsilon^{(0)}\|_{A^{\frac{1}{2}}}.$$

Εναλλακτικά, χρησιμοποιώντας τη σχέση

$$\begin{aligned} T_m(t) &= \frac{1}{2} \left[(t + \sqrt{t^2 - 1})^m + (t - \sqrt{t^2 - 1})^m \right] \\ &\geq \frac{1}{2} (t + \sqrt{t^2 - 1})^m, \end{aligned}$$

παρατηρούμε ότι για $\eta = \frac{\lambda_{\min}}{\lambda_{\max} - \lambda_{\min}}$ το μέγιστο του πολυωνύμου Chebyshev δίνεται από τον τύπο

$$T_m(1 + 2\eta) \geq \frac{1}{2} \left(1 + 2\eta + \sqrt{(1 + 2\eta)^2 - 1} \right)^m$$

$$\begin{aligned}
&\geq \frac{1}{2} \left(1 + 2\eta + 2\sqrt{\eta(\eta+1)} \right)^m \\
&= \frac{1}{2} \left(\sqrt{\eta} + \sqrt{\eta+1} \right)^{2m} \\
&= \frac{1}{2} \left[\frac{(\sqrt{\lambda_{\min}} + \sqrt{\lambda_{\max}})^2}{\lambda_{\max} - \lambda_{\min}} \right]^m \\
&= \frac{1}{2} \left(\frac{\sqrt{\lambda_{\max}} + \sqrt{\lambda_{\min}}}{\sqrt{\lambda_{\max}} - \sqrt{\lambda_{\min}}} \right)^m \\
&= \frac{1}{2} \left(\frac{\sqrt{\kappa} + 1}{\sqrt{\kappa} - 1} \right)^m. \quad \square
\end{aligned}$$

Σχόλιο. Για την απόδειξη της σχέσης

$$T_m(t) = \frac{1}{2} \left[\left(t + \sqrt{t^2 - 1} \right)^m + \left(t + \sqrt{t^2 - 1} \right)^{-m} \right]$$

που χρησιμοποιήθηκε στο τέλος του προηγούμενου θεωρήματος, χρησιμοποιούμε την τριγωνομετρική ταυτότητα

$$\cos(m\theta) + i \sin(m\theta) = (\cos(\theta) + i \sin(\theta))^m,$$

με $\theta = \arccos(t)$ η οποία δίνει $(t + \sqrt{t^2 - 1})^m$. Αντίστοιχα η

$$\cos(m\theta) - i \sin(m\theta) = (\cos(\theta) + i \sin(\theta))^{-m}$$

δίνει $(t + \sqrt{t^2 - 1})^{-m}$. Οπότε η ζητούμενη σχέση προκύπτει εύκολα ως ημίαθροισμα αυτών των όρων.

2.4 Γενικευμένη Μέθοδος Ελαχίστου Υπολοίπου (Generalized Minimal Residual Method - GMRES)

Η Γενικευμένη Μέθοδος Ελαχίστου Υπολοίπου (GMRES) είναι μια μέθοδος προσέγγισης της λύσης του συστήματος $Ax = b$ για γενικό αντιστρέψιμο πίνακα $A \in \mathbb{C}^{n,n}$, όχι κατ' ανάγκη συμμετρικό και θετικά ορισμένο. Ζητούμενο είναι να βρούμε μια ορθοκανονική βάση $\{v_1, v_2, \dots, v_k\}$ του υποχώρου Krylov $\text{span}\{r^{(0)}, Ar^{(0)}, \dots, A^{m-1}r^{(0)}\}$, όπου για την προσεγγιστική λύση $x^{(k)}$ η οποία έχει βρεθεί στο βήμα k του αλγορίθμου ως $x^{(k)} = x^{(k-1)} + y^{(k)}v_k$, ισχύει $b - Ax^{(k)} \perp \text{span}\{r^{(0)}, Ar^{(0)}, \dots, A^k r^{(0)}\}$. Η μέθοδος επιτυγχάνει την ελαχιστοποίηση της νόρμας του διανύσματος - υπόλοιπο $\|r^{(k)}\|$ ως προς όλα τα διανύσματα

του χώρου $x_0 + K_k$. Για τον προσδιορισμό της βάσης διανυσμάτων v_1, v_2, \dots, v_k με αρχικό διάνυσμα το $v_1 = r^{(0)}/\beta$, με $\beta = \|r^{(0)}\|$ χρησιμοποιείται ο αλγόριθμος του Arnoldi, ο οποίος είναι ουσιαστικά μία παραλλαγή του αλγορίθμου Gram-Schmidt. Παρακάτω παρατίθεται ο αλγόριθμος Arnoldi.

Αλγόριθμος 2.7 Arnoldi

Δεδομένα: $A \in \mathbb{C}^{n,n}$, K_m .

Αποτέλεσμα: Ορθοκανονική βάση διανυσμάτων v_i του K_m .

- 1: *Επέλεξε ένα διάνυσμα $v_1 \in \mathbb{C}^n \setminus \{0\}$ νόρμας 1*
- 2: **Για** $j = 1, 2, \dots, m$
- 3: **Για** $i = 1, 2, \dots, j$
- 4: $h_{ij} = (Av_j, v_i)$
- 5: $w_j := Av_j - h_{ij}v_i$
- 6: **Τέλος ‘Για’**
- 7: $h_{j+1,j} = \|w_j\|$
- 8: **Αν** $h_{j+1,j} = 0$
- 9: *Σταμάτα*
- 10: **Τέλος ‘Αν’**
- 11: $v_{j+1} = \frac{w_j}{h_{j+1,j}}$
- 12: **Τέλος ‘Για’**

Συχνά, αντί αυτού του αλγορίθμου χρησιμοποιείται ο τροποποιημένος αλγόριθμος του Arnoldi, που δίνεται στη συνέχεια, ο οποίος προτιμάται λόγω καλύτερης ευστάθειας.

Αλγόριθμος 2.8 Τροποποιημένος Arnoldi

Δεδομένα: $A \in \mathbb{C}^{n,n}$, K_k .

Αποτέλεσμα: Ορθοκανονική βάση διανυσμάτων v_i του K_k .

- 1: *Επέλεξε ένα διάνυσμα $v_1 \in \mathbb{C}^n \setminus \{0\}$ νόρμας 1*
- 2: **Για** $j = 1, 2, \dots, k$
- 3: $w_j := Av_j$
- 4: **Για** $i = 1, 2, \dots, j$
- 5: $h_{ij} = (w_j, v_i)$
- 6: $w_j := w_j - h_{ij}v_i$
- 7: **Τέλος ‘Για’**
- 8: $h_{j+1,j} = \|w_j\|$
- 9: **Αν** $h_{j+1,j} = 0$
- 10: *Σταμάτα*
- 11: **Τέλος ‘Αν’**

$$12: \quad v_{j+1} = \frac{w_j}{h_{j+1,j}}$$

13: **Τέλος ‘Για’**

Αν παρατηρήσουμε προσεκτικά τον αλγόριθμο του Arnoldi, διαπιστώνουμε ότι ενδέχεται σε κάποιο βήμα j να σταματήσει να εκτελείται. Αυτό συμβαίνει στην περίπτωση που το $h_{j+1,j} = \|w_j\| = 0$, πράγμα που πρακτικά σημαίνει ότι δεν μπορεί να υπολογιστεί το αμέσως επόμενο ορθοκανονικό διάνυσμα v_{j+1} . Το παρακάτω Θεώρημα επιβεβαιώνει αυτό τον ισχυρισμό.

Θεώρημα 2.4.1. *Ο αλγόριθμος του Arnoldi καταρρέει (breakdown) στο βήμα j , δηλ. $h_{j+1,j} = 0$, αν και μόνο αν το ελάχιστο πολυώνυμο του διανύσματος v_1 είναι βαθμού j , δηλαδή ισχύει $\text{grade}(v_1) = j$. Ακόμη, σε αυτή την περίπτωση ο υπόχωρος K_j είναι αναλλοίωτος στην επίδραση του A (invariant under A).*

Απόδειξη. (\Leftarrow) Έστω ότι ο βαθμός του ελαχίστου πολυωνύμου είναι j , δηλαδή ισχύει $\text{grade}(v_1) = j$. Τότε συνεπάγεται ότι $w_j = 0$. Αν υποθέσουμε ότι $w_j \neq 0$, τότε από τον αλγόριθμο του Arnoldi θα μπορούσε να οριστεί το διάνυσμα $v_{j+1} = \frac{w_j}{h_{j+1,j}}$. Αυτό είναι αδύνατο, γιατί σε αυτή την περίπτωση θα οριζόταν ένας υπόχωρος Krylov K_{j+1} με διάσταση $j + 1$, μεγαλύτερη του j που είναι εξ υποθέσεως η μέγιστη δυνατή διάσταση που μπορεί να έχει ο χώρος Krylov. (\Rightarrow) Για να αποδειχθεί το αντίστροφο, έστω $w_j = 0$. Τότε, αφού από τον αλγόριθμο

$$\begin{aligned} w_j &= h_{j+1,j}v_{j+1} = 0 \\ \implies Aq_{j-1}(A)v_1 - \sum_{i=1}^j h_{ij}q_{i-1}(A)v_1 &= 0 \\ \implies \text{grade}(v_1) = \mu &\leq j \end{aligned}$$

Αν ίσχυε ότι $\text{grade}(v_1) = \mu < j$, ο αλγόριθμος του Arnoldi θα είχε σταματήσει να εκτελείται στο βήμα μ , βάσει του άλλου μέρους της απόδειξης. Άρα τελικά προκύπτει ότι $\text{grade}(v_1) = \mu = j$. Επίσης, προκύπτει εύκολα ότι ο K_j είναι αναλλοίωτος στην επίδραση του A , δηλαδή ότι $K_m = K_j$, για όλα τα $m \geq j$. Για την απόδειξη αυτού, αρκεί να δείχθει ότι το διάνυσμα Ax μπορεί να γραφεί ως γραμμικός συνδυασμός των διανυσμάτων του χώρου K_j , άρα $Ax \in K_j$. \square

Πρόταση 2.4.2. *Αν ο αλγόριθμος του Arnoldi περατώνεται στα πρώτα m βήματα, τα διανύσματα v_1, v_2, \dots, v_m αποτελούν ορθοκανονική βάση του υπόχωρου Krylov $K_m = \text{span}\{v_1, Av_1, \dots, A^{m-1}v_1\}$.*

Απόδειξη. Θα αποδείξουμε ότι τα v_1, v_2, \dots, v_k αποτελούν ορθοκανονική βάση

του K_m , άρα αρκεί να δείξουμε ότι είναι ορθοκανονικά και ότι παράγουν τον χώρο. Το ότι είναι ορθοκανονικά ισχύει προφανώς εκ κατασκευής τους. Για να δείξουμε ότι παράγουν τον K_m , αρκεί να αποδείξουμε ότι κάθε ένα από τα v_j μπορεί να γραφεί στη μορφή $q_{j-1}(A)v_1$, όπου q_{j-1} είναι ένα πολυώνυμο βαθμού $j-1$. Η απόδειξη ολοκληρώνεται με επαγωγή.

Για $j=1$ είναι $v_1 = q_0(A)v_1$, συνεπώς αρκεί να επιλέγει $q_0(t) \equiv 1$.

Έστω ότι η υπόθεση ισχύει για κάθε ακέραιο μικρότερο ή ίσο του j , δηλαδή $v_j = q_{j-1}(A)v_1$.

Θα δείξουμε ότι ισχύει για $j+1$. Από τον αλγόριθμο προκύπτει ότι $h_{j+1,j}v_{j+1} = Av_j - \sum_{i=1}^j h_{ij}v_i = Aq_{j-1}(A)v_1 - \sum_{i=1}^j h_{ij}q_{i-1}(A)v_1 = q_j(A)v_1$, όπου q_j είναι πολυώνυμο βαθμού j . □

Σε μορφή πινάκων, ο αλγόριθμος του Arnoldi μας λέει ότι

$$A[v_1, v_2, \dots, v_m] = [v_1, v_2, \dots, v_m, v_{m+1}] \begin{bmatrix} h_{1,1} & h_{1,2} & \dots & h_{1,m} \\ h_{2,1} & h_{2,2} & \dots & h_{2,m} \\ & h_{3,2} & \dots & h_{3,m} \\ & & \ddots & h_{m,m} \\ & & & h_{m+1,m} \end{bmatrix}.$$

Ισοδύναμα,

$$AV_m = V_m H_m + v_{m+1} h_{m+1,m} = V_m H_m + w_m e^{(m)T} = V_{m+1} \bar{H}_m,$$

όπου το διάνυσμα $e^{(m)}$ είναι το μοναδιαίο διάνυσμα με μηδενικά παντού και 1 στην m συνιστώσα, ο V_m είναι ο $n \times m$ πίνακας με διανύσματα στήλες v_1, \dots, v_m , ο H_m είναι ο $m \times m$ Hessenberg πίνακας με τα μη μηδενικά στοιχεία h_{ij} να δίνονται από τον αλγόριθμο του Arnoldi, ο \bar{H}_m είναι ο $(m+1) \times m$ επαυξημένος κατά μία γραμμή πίνακας H_m με μόνο μη μηδενικό στοιχείο το $h_{m+1,m}$, και το διάνυσμα $w_m = v_{m+1} h_{m+1,m}$. Συγκεκριμένα, ο αλγόριθμος του Arnoldi συνοψίζεται στην επόμενη πρόταση.

Πρόταση 2.4.3. Αν V_m είναι ο $m \times n$ πίνακας με διανύσματα στήλες v_1, \dots, v_m , \bar{H}_m είναι ο $(m+1) \times m$ πίνακας Hessenberg με μη μηδενικά στοιχεία h_{ij} να δίνονται από τον αλγόριθμο του Arnoldi και H_m είναι ο πίνακας που προκύπτει από τον \bar{H}_m διαγράφοντας την τελευταία του γραμμή, τότε ισχύουν οι σχέσεις

$$AV_m = V_m H_m + w_m e^{(m)T} = V_{m+1} \bar{H}_m, \text{ όπου } V_m^T AV_m = H_m.$$

Στη συνέχεια, θα ασχοληθούμε με την κατασκευή της Γενικευμένης Μεθόδου Ελαχίστου Υπολοίπου. Για κάθε διάνυσμα

$$x \in x^{(0)} + K_m = \text{span}\{v_1, Av_1, \dots, A^m v_1\}$$

ισχύει $x = x^{(0)} + V_m y$, όπου $y \in \mathbb{C}^m$. Τότε, $\|b - Ax\| = \|b - A(x^{(0)} + V_m y)\|$. Στόχος μας, όπως έχει ήδη αναφερθεί, είναι η ελαχιστοποίηση αυτής της νόρμας. Οπότε έχουμε διαδοχικά

$$\begin{aligned} \|b - Ax\| &= \|b - A(x^{(0)} + V_m y)\| = \|b - Ax^{(0)} - AV_m y\| \\ &= \|r_0 - AV_m y\| = \|\beta v_1 - V_{m+1} \bar{H}_m y\| \\ &= \|V_{m+1}(\beta e^{(1)} - \bar{H}_m y)\|. \end{aligned}$$

Αφού τα διανύσματα στήλες του πίνακα V_{m+1} είναι ορθοκανονικά και κάθε ορθοκανονικός μετασχηματισμός διατηρεί αναλλοίωτη την Ευκλείδεια νόρμα, συμπεραίνουμε ότι $\|b - Ax\| = \|\beta e^{(1)} - \bar{H}_m y\|$. Για να ελαχιστοποιήσουμε τη ζητούμενη νόρμα, αρκεί να λύσουμε ένα $(m+1) \times m$ πρόβλημα ελαχίστων τετραγώνων όπου το ζητούμενο είναι να βρεθεί το ελάχιστο $y^{(m)} = \arg\min_y \|\beta e^{(1)} - \bar{H}_m y\|$ το οποίο ελαχιστοποιεί τη $\|\beta e^{(1)} - \bar{H}_m y\|$. Για την επίλυση του προβλήματος ελαχίστων τετραγώνων αρκεί να επιτύχουμε την QR ανάλυση του άνω Hessenberg πίνακα \bar{H}_m . Ο αλγόριθμος που χρησιμοποιείται για να λάβουμε την QR ανάλυση του πίνακα A είναι ο αλγόριθμος εφαρμογής διαδοχικών στροφών Givens. Το δε y δίνεται από την επίλυση ενός άνω τριγωνικού συστήματος.

Παρακάτω περιγράφεται η βασική ιδέα πίσω από τη μέθοδο των στροφών Givens, οι οποίες είναι ορθογώνιοι μετασχηματισμοί. Η ιδέα βασίζεται στο γεγονός ότι ο απλούστερος ορθογώνιος μετασχηματισμός είναι ένας 2×2 πίνακας της μορφής $\Omega_\theta = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$, ο οποίος είναι προφανώς ορθογώνιος και αν εφαρμοστεί σε κάποιο διάνυσμα $\alpha = (x, y)$ του επιπέδου το στρέφει κατά μια γωνία θ . Γενικεύοντας αυτό τον ισχυρισμό στη διάσταση $n \times n$, μπορούμε να

παρατηρήσουμε ότι ο Ω_k είναι ένας $n \times n$ πίνακας της μορφής

$$\begin{array}{c}
 \begin{array}{c} k \text{ στήλη} \\ \downarrow \end{array} \qquad \qquad \begin{array}{c} j \text{ στήλη} \\ \downarrow \end{array} \\
 \begin{array}{c} 1 \\ \vdots \\ 1 \\ \cos \theta \\ \qquad \qquad 1 \\ \qquad \qquad \qquad \vdots \\ \sin \theta \\ \qquad \qquad \qquad \qquad 1 \\ \qquad \qquad \qquad \qquad \qquad 1 \\ \qquad \qquad \qquad \qquad \qquad \qquad \vdots \\ \qquad \qquad \qquad \qquad \qquad \qquad \qquad 1 \end{array}
 \end{array}$$

που έχει μονάδες στη διαγώνιο, στην k γραμμή τα στοιχεία $\cos \theta$, $-\sin \theta$ και στην j γραμμή τα στοιχεία $\sin \theta$, $\cos \theta$ και παντού αλλού μηδέν. Αν εφαρμοστεί σε ένα n - διάστατο διάνυσμα το στρέφει ως προς τις k , j συνιστώσες του κατά γωνία θ κατά τη θετική φορά. Σημειώνεται ότι η γωνία θ θα είναι (γενικά) μιγαδικός αριθμός.

Πρακτικά, για τη μέθοδο GMRES ορίζουμε ως πίνακα στροφή τον πίνακα

$$\Omega_i = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & c_i & s_i & & \\ & & -s_i & c_i & & \\ & & & & \ddots & \\ & & & & & 1 \end{bmatrix},$$

ο οποίος έχει στην i γραμμή τα στοιχεία c_i , s_i και στην $i + 1$ γραμμή τα $-s_i$, c_i με $c_i^2 + s_i^2 = 1$, παντού αλλού 0 και στη διαγώνιο 1. Αν υποθέσουμε ότι έχουν εφαρμοστεί τα πρώτα k βήματα του αλγορίθμου GMRES, οι πίνακες στροφής Ω_i θα έχουν διάσταση $(k + 1) \times (k + 1)$. Κατά την εφαρμογή διαδοχικών στροφών Givens στόχος μας είναι ο μηδενισμός όλων των στοιχείων κάτω από την κύρια διαγώνιο. Έτσι, σε κάθε βήμα τα c_i , s_i επιλέγονται έτσι ώστε να μηδενίζουν κάθε φορά το στοιχείο $h_{i+1,i}$. Γενικά, για την i εφαρμογή μιας στροφής ισχύει ότι

$$s_i = \frac{h_{i+1,i}}{\sqrt{(h_{ii}^{(i-1)})^2 + h_{i+1,i}^2}}, \quad c_i = \frac{h_{ii}^{(i-1)}}{\sqrt{(h_{ii}^{(i-1)})^2 + h_{i+1,i}^2}},$$

όπου $h_{ii}^{(i-1)}$ είναι το στοιχείο στην (i, i) θέση, όπως αυτό έχει τροποποιηθεί κατά την $i-1$ εφαρμογή. Αν εφαρμόσουμε αυτή τη διαδικασία σε όλα τα στοιχεία του πίνακα $A \in \mathbb{C}^{n,n}$ που βρίσκονται κάτω από την κύρια διαγώνιο λαμβάνουμε ότι

$$\bar{R}_m = \Omega_m \Omega_{m-1} \dots \Omega_1 \bar{H}_m \iff \Omega_1 \Omega_2 \dots \Omega_m \bar{R}_m = \bar{H}_m \iff Q_m \bar{R}_m = \bar{H}_m,$$

αρκεί με Q_m να συμβολίσουμε το γινόμενο των m πρώτων πινάκων στροφής. Τότε έχουμε την ανάλυση QR του πίνακα \bar{H}_m . Το πρόβλημα ελαχίστων τετραγώνων γίνεται

$$\begin{aligned} \min_{y \in \mathbb{C}^n} \|\beta e^{(1)} - \bar{H}_m y\| &= \min_y \|Q_m^T \beta e^{(1)} - Q_m^T \bar{H}_m y\| \\ &= \min_y \|\bar{g}_m - \bar{R}_m y\| = \\ &= \left\| \frac{g_m - R_m y}{\gamma_{m+1}} \right\|. \end{aligned}$$

Το ελάχιστο δίνεται από τη λύση του συστήματος $R_m y = g_m = (\gamma_1, \gamma_2, \dots, \gamma_m)^T$, όπου R_m είναι ο $m \times m$ πίνακας που προκύπτει από τον \bar{R}_m διαγράφοντας την τελευταία γραμμή και g_m το m -διάστατο διάνυσμα που προκύπτει από το \bar{g}_m διαγράφοντας το τελευταίο του στοιχείο. Τα προαναφερθέντα συνοψίζονται στο ακόλουθο θεώρημα.

Θεώρημα 2.4.4. Έστω $\Omega_i, i = 1(1)m$ οι πίνακες στροφής που χρησιμοποιούνται για να μετατρέψουν τον \bar{H}_m σε άνω τριγωνικό πίνακα, \bar{R}_m ο πίνακας που προκύπτει και $\bar{g}_m = Q_m^H \beta e^{(1)} = (\gamma_1, \dots, \gamma_{m+1})^T$ το διάνυσμα που προκύπτει στο δεξιό μέλος. Επίσης, έστω R_m, g_m ο πίνακας και το διάνυσμα που προκύπτουν αν από τους \bar{R}_m και \bar{g}_m αν διαγράψουμε την τελευταία γραμμή. Τότε ισχύει:

1. Το διάνυσμα $y^{(m)}$ που ελαχιστοποιεί τη νόρμα $\|\beta e^{(1)} - \bar{H}_m y\|$ δίνεται από τη σχέση $y^{(m)} = R_m^{-1} g_m$.
2. Το διάνυσμα σφάλμα στο βήμα m του αλγορίθμου ικανοποιεί τη σχέση $r^{(m)} = b - Ax^{(m)} = V_{m+1}(\beta e^{(1)} - \bar{H}_m y^{(m)}) = V_{m+1} Q_m^T (\gamma_{m+1} e^{(m+1)})$, άρα $\|b - Ax^{(m)}\| = |\gamma_{m+1}|$.

Απόδειξη. Για την απόδειξη του πρώτου ισχυρισμού, παρατηρούμε ότι:

$$\begin{aligned}\|\beta e^{(1)} - \bar{H}_m y\|^2 &= \|Q_m^T(\beta e^{(1)} - \bar{H}_m y)\|^2 \\ &= \|Q_m^T \beta e^{(1)} - Q_m^T \bar{H}_m y\|^2 \\ &= \|\bar{g}_m - \bar{R}_m y\|^2 \\ &= |\gamma_{m+1}|^2 + \|g_m - R_m y\|^2,\end{aligned}$$

αφού το γ_{m+1} είναι το τελευταίο στοιχείο του διανύσματος \bar{g}_m . Το ελάχιστο (ως προς y) της ποσότητας $|\gamma_{m+1}|^2 + \|g_m - R_m y\|^2$ δίνεται όταν $y = R_m^{-1} g_m$.

Για την απόδειξη του δεύτερου ισχυρισμού, αν x είναι ένα διάνυσμα που γράφεται ως $x = x^{(0)} + V_m y$, τότε

$$\begin{aligned}b - Ax &= V_{m+1}(\beta e^{(1)} - \bar{H}_m y) \\ &= V_{m+1} Q_m Q_m^T (\beta e^{(1)} - \bar{H}_m y) \\ &= V_{m+1} Q_m (\bar{g}_m - \bar{R}_m y).\end{aligned}$$

Από την προηγούμενη απόδειξη παρατηρούμε ότι η $\|\bar{g}_m - \bar{R}_m y\|$ ελαχιστοποιείται όταν μείνει μόνο το τελευταίο στοιχείο από το διάνυσμα \bar{g}_m . Τότε ισχύει,

$$b - Ax = V_{m+1} Q_m (\gamma_{m+1} e^{(m+1)})$$

και άρα $\|b - Ax^{(m)}\| = |\gamma_{m+1}|$, αφού τα διανύσματα στήλες του $V_{m+1} Q_m^T$ είναι ορθομοναδιαία. \square

Παραθέτουμε τώρα τον πλήρη αλγόριθμο της Μεθόδου GMRES.

Αλγόριθμος 2.9 Γενικευμένη Μέθοδος Ελαχίστου Υπολοίπου GMRES

- 1: Υπολόγισε τα $r^{(0)} = b - Ax^{(0)}$, $\beta := \|r^{(0)}\|$ και $v_1 := r^{(0)}/\beta$
- 2: Όρισε τον $(m+1) \times m$ πίνακα $\bar{H}_m = h_{ij}$, $1 \leq i \leq m+1$, $1 \leq j \leq m$
- 3: Θέσε $\bar{H}_m = 0$
- 4: **Για** $j = 1, 2, \dots, m$
- 5: $w_j = Av_j$
- 6: **Για** $i = 1, \dots, j$
- 7: $h_{ij} := (w_j, v_i)$
- 8: $w_j := w_j - h_{ij} v_j$
- 9: **Τέλος 'Για'**
- 10: $h_{j+1,j} = \|w_j\|$.
- 11: **Αν** $h_{j+1,j} = 0$
- 12: Θέσε $m := j$ και πήγαινε στο βήμα 15

13: **Τέλος ‘Αν’**

14: **Τέλος ‘Για’**

15: Υπολόγισε την $y^{(m)}$ που ελαχιστοποιεί την $\|\beta e^{(1)} - \bar{H}_m y\|$ και την $x^{(m)} = x^{(0)} + V_m y^{(m)}$

Η μέθοδος GMRES είναι δυνατόν να σταματήσει να εκτελείται χωρίς να έχει καταφέρει να δώσει λύση. Αυτό ονομάζεται κατάρρευση (breakdown) της μεθόδου. Αν παρατηρήσουμε προσεκτικά τον αλγόριθμο, βλέπουμε ότι αυτό μπορεί να συμβεί αν κατά την κατασκευή των ορθοκανονικών διανυσμάτων που παράγονται από τον αλγόριθμο του Arnoldi για το διάνυσμα που κατασκευάζεται στο βήμα j ισχύει $h_{j+1,j} = 0$. Τότε, προφανώς και δεν μπορεί να βρεθεί το αμέσως επόμενο ορθοκανονικοποιημένο διάνυσμα v_{j+1} και ο αλγόριθμος της μεθόδου σταματάει να εκτελείται στο βήμα j . Παρ’ όλα αυτά, για το συγκεκριμένο βήμα, δύναται να βρεθεί λύση, η οποία είναι ακριβής. Αυτό περιγράφεται στην επόμενη πρόταση.

Πρόταση 2.4.5. Έστω $A \in \mathbb{C}^{n,n}$ ένας αντιστρέψιμος πίνακας. Τότε, ο αλγόριθμος της μεθόδου GMRES καταρρέει (breakdown) στο βήμα j , δηλ. $h_{j+1,j} = 0$, αν και μόνο αν η προσεγγιστική λύση $x^{(j)}$ είναι ακριβής.

Απόδειξη. (\Rightarrow) Έστω ότι ο αλγόριθμος της μεθόδου GMRES καταρρέει στο βήμα j , δηλ. $h_{j+1,j} = 0$. Επειδή ισχύει ότι $h_{j+1,j} = 0$, κατά την j εφαρμογή του πίνακα στροφής Ω_j θα ισχύει $s_j = 0$, αφού $s_j = h_{j+1,j} / \sqrt{(h_{jj}^{(j-1)})^2 + h_{j+1,j}^2}$. Τότε, για το συγκεκριμένο βήμα j , η τελευταία γραμμή του πίνακα \bar{R}_j , η οποία είναι η μηδενική γραμμή, καθώς και το τελευταίο στοιχείο του διανύσματος \bar{g}_j το οποίο είναι το $\gamma_{j+1} = -s_j \gamma_j = 0$ αποτελούν ταυτοτική εξίσωση και το $y^{(j)}$ υπολογίζεται από τις πρώτες j εξισώσεις. Οπότε τελικά προκύπτει ότι $\gamma_{j+1} = 0 \Leftrightarrow \|b - Ax^{(j)}\| = 0 \Rightarrow \|r^{(j)}\| = 0$, δηλαδή ότι η λύση είναι ακριβής στο βήμα j . (\Leftarrow) Έστω ότι η λύση $x^{(j)}$ είναι ακριβής στο βήμα j και όχι στο προηγούμενο βήμα $j - 1$. Τότε, αφού ισχύει $\gamma_{j+1} = -s_j \gamma_j$, έπεται ότι $s_j = 0$. Αφού $s_j = h_{j+1,j} / \sqrt{(h_{jj}^{(j)})^2 + h_{j+1,j}^2} \Rightarrow h_{j+1,j} = 0$, δηλαδή ο αλγόριθμος της GMRES σταματάει να εκτελείται στο βήμα j . \square

Αν ο αλγόριθμος της μεθόδου συνεχίζει να εκτελείται μέχρι το βήμα m , τότε μπορούμε να έχουμε για την Ευκλείδεια νόρμα του διανύσματος - υπόλοιπο μία εκτίμηση ανάλογη με αυτή που δόθηκε για τη Μέθοδο Συζυγών Κλίσεων. Υπενθυμίζεται ότι κατά τη μέθοδο GMRES το ζητούμενο είναι η ελαχιστοποίηση της

Ευκλείδεια νόρμας του διανύσματος - υπόλοιπο

$$\|r^{(m)}\| = \min_{\substack{p_m \in P_m \\ p_m(0)=1}} \|p_m(A)r^{(0)}\|,$$

υποθέτοντας ότι ο πίνακας A είναι διαγωνοποιήσιμος. Το παρακάτω θεώρημα δίνει μια εκτίμηση για την $\|r^{(m)}\|$ συναρτήσει της $\|r^{(0)}\|$, ανεξάρτητη όμως από το $r^{(0)}$.

Θεώρημα 2.4.6. Έστω $A \in \mathbb{C}^{n,n}$, $\det(A) \neq 0$ διαγωνοποιήσιμος και $A = V\Lambda V^{-1}$ η ανάλυση Jordan του πίνακα A , όπου $\Lambda = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ ο διαγώνιος πίνακας που περιέχει τις n το πλήθος ιδιοτιμές του πίνακα A και V ο πίνακας με τα αντίστοιχα ιδιοδιανύσματα. Ορίζουμε ως:

$$\varepsilon^{(m)} = \min_{\substack{p \in P_m \\ p(0)=1}} \max_{i=1(1)n} |p(\lambda_i)|.$$

Τότε η Ευκλείδεια νόρμα του διανύσματος - υπόλοιπο στο m βήμα του αλγορίθμου GMRES ικανοποιεί την εκτίμηση:

$$\|r^{(m)}\| \leq \kappa(V) \cdot \varepsilon^{(m)} \|r^{(0)}\|, \quad \kappa(V) = \|V\| \|V^{-1}\|.$$

Απόδειξη. Έχουμε $r^{(m)} = b - Ax^{(m)} = p(A)r^{(0)}$, με $p \in P_m$, $p(0) = 1$. Έστω ένα τέτοιο p και το διάνυσμα $x \in K_m$ τέτοια ώστε $b - Ax = p(A)r^{(0)}$. Τότε:

$$\begin{aligned} \|b - Ax\| &= \|Vp(\Lambda)V^{-1}r^{(0)}\| \\ &= \|V\| \|V^{-1}\| \|r^{(0)}\| \|p(\lambda)\|. \end{aligned}$$

Αφού ο Λ είναι διαγώνιος, έπεται ότι $\|p(\lambda)\| = \max_{i=1(1)n} |p(\lambda_i)|$.

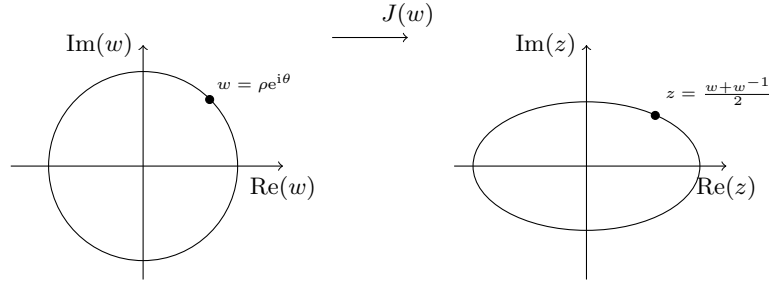
Οπότε, για την προσέγγιση $x^{(m)}$ ισχύει

$$\begin{aligned} \|b - Ax^{(m)}\| &\leq \|b - Ax\| \\ &\leq \|V\| \|V^{-1}\| \|r^{(0)}\| \max_{i=1(1)n} |p(\lambda)|. \end{aligned}$$

Ελαχιστοποιώντας το δεύτερο μέλος παίρνουμε

$$\begin{aligned} \|b - Ax^{(m)}\| &\leq \min_{\substack{p \in P_m \\ p(0)=1}} \|V\| \|V^{-1}\| \|r^{(0)}\| \max_{i=1(1)n} |p(\lambda_i)| \\ \iff \|b - Ax^{(m)}\| &\leq \|V\| \|V^{-1}\| \|r^{(0)}\| \varepsilon^{(m)}. \end{aligned}$$

Η απόδειξη ολοκληρώνεται με τον υπολογισμό ενός άνω φράγματος για το $\varepsilon^{(m)}$. □



Σχήμα 2.1: Η απεικόνιση Joukowski

Ένα άνω φράγμα δίνεται, όπως και για την αντίστοιχη εκτίμηση σφάλματος της Μεθόδου Συζυγών Κλίσεων, από τη Θεωρία Προσέγγισης. Συγκεκριμένα, για τον προσδιορισμό του φράγματος χρησιμοποιούνται τα πολυώνυμα Chebyshev. Αυτά, σχετίζονται με ελλείψεις. Αυτό τεκμηριώνεται από τον ακόλουθο ισχυρισμό. Έστω T_m ο κύκλος με ακτίνα m και κέντρο $O(0, 0)$. Η απεικόνιση $J(w) = \frac{1}{2}[w + w^{-1}]$ (γνωστή ως απεικόνιση Joukowski) μετατρέπει αυτό τον κύκλο σε έλλειψη, με εστίες στο $[-1, 1]$, μεγάλο ημιάξονα τον $\frac{1}{2}[m + m^{-1}]$ και μικρό ημιάξονα τον $\frac{1}{2}[m - m^{-1}]$. Σχηματικά:

Παρακάτω παρατίθεται η απόδειξη αυτού του ισχυρισμού. Ο κύκλος $\gamma_\rho: \theta \rightarrow \rho e^{i\theta}$, $0 \leq \theta \leq 2\pi$, μετατρέπεται σε μία έλλειψη μέσω της απεικόνισης Joukowski ως εξής:

$$\begin{aligned}
 J(\gamma_\rho) : w(\theta) &= \frac{1}{2} \left(\rho e^{i\theta} + \frac{1}{\rho} e^{i\theta} \right) \\
 &= \frac{1}{2} \left[\rho(\cos \theta + i \sin \theta) + \frac{1}{\rho}(\cos \theta - i \sin \theta) \right] \\
 &= \frac{1}{2} \left(\rho \cos \theta + \rho i \sin \theta + \frac{1}{\rho} \cos \theta - \frac{1}{\rho} i \sin \theta \right) \\
 &= \frac{1}{2} \left(\rho + \frac{1}{\rho} \right) \cos \theta + \frac{1}{2} \left(\rho - \frac{1}{\rho} \right) i \sin \theta \\
 &= \alpha \cos \theta + \beta i \sin \theta,
 \end{aligned}$$

όπου $0 \leq \theta \leq 2\pi$, μεγάλος άξονας $\alpha = \frac{1}{2} \left(\rho + \frac{1}{\rho} \right)$, μικρός άξονας $\beta =$

$\frac{1}{2} \left(\rho - \frac{1}{\rho} \right)$. Αυτός περιέχει το διάστημα $[-1, 1]$ για κίνηση όμοια με αυτή των δεικτών του ρολογιού όταν $\rho > 1$, και για κίνηση αντίθετη από αυτή των δεικτών του ρολογιού όταν $\rho < 1$. Στην περίπτωση όπου $\rho = 1$, η έλλειψη που προκύπτει είναι η $E(0, 1, -1)$ και εκφυλίζεται σε ευθύγραμμο τμήμα το οποίο διαγράφεται δύο φορές.

Αν υποθέσουμε ότι όλες οι ιδιοτιμές του πίνακα A περιέχονται στην έλλειψη $E(c, d, a)$ χωρίς τους άξονες συντεταγμένων, όπου c είναι το κέντρο, d η εστιακή απόσταση και a το μήκος του μικρού ημιάξονα, τότε αποδεικνύεται ότι το πηλίκο $\frac{T_m \left(\frac{a}{d} \right)}{T_m \left(\frac{c}{d} \right)} \approx \left(\frac{a + \sqrt{a^2 - d^2}}{c + \sqrt{c^2 - d^2}} \right)^k$ είναι ένα άνω φράγμα για το ε^m . Για περισσότερα βλέπε [9].

Η μέθοδος GMRES είναι στην πράξη από τις πιο ασύμφορες υπολογιστικά μεθόδους. Αυτό συμβαίνει κατ' αρχάς διότι ο αλγόριθμος Arnoldi που χρησιμοποιείται είναι της τάξης $O(m^2n)$, πράγμα που αυξάνει αρκετά το υπολογιστικό κόστος. Επιπροσθέτως, τα διανύσματα που κατασκευάζονται στα m πρώτα βήματά του, χρειάζονται αρκετό χώρο αποθήκευσης στη μνήμη, της τάξης $O(mn)$. Το γεγονός αυτό, για πολύ μεγάλα m και πόσω μάλλον όταν το μέγεθος του πίνακα A αυξάνει, είναι μη επιθυμητό. Θα ήταν ευχής έργο να μειώσουμε το υπολογιστικό κόστος του αλγορίθμου, καθώς και τις απαιτήσεις στη μνήμη ταυτόχρονα. Για να γίνει αυτό, αρκεί να μπορέσουμε να διατηρήσουμε το m όσο πιο μικρό γίνεται. Έτσι, είναι δυνατόν, αντί του αλγορίθμου GMRES, να εκτελέσουμε τον αλγόριθμο της GMRES με επανεκκίνηση (Restarted GMRES). Αυτό σημαίνει ότι στην πράξη μετά από λίγες επαναλήψεις, εκτελούμε ξανά τη μέθοδο GMRES παίρνοντας ως εκάστοτε αρχική προσέγγιση την τελευταία προσέγγιση που μας έδωσε η αμέσως προηγούμενη εκτέλεση. Συμβολίζουμε με GMRES(m) τη μέθοδο με επανεκκίνηση μετά τις πρώτες m επαναλήψεις. Ως εσωτερική επανάληψη ορίζεται κάθε επανάληψη μεταξύ δύο επανεκκινήσεων της μεθόδου. Οι m εσωτερικές επαναλήψεις κατασκευάζουν την ορθοκανονική βάση K_m της μεθόδου. Ως εξωτερική επανάληψη ορίζεται το σύνολο των m εσωτερικών επαναλήψεων της μεθόδου μεταξύ δύο επανεκκινήσεων. Οπότε είναι φανερό ότι η GMRES(m) εμπεριέχει δύο διαφορετικά είδη επαναλήψεων.

Αλγόριθμος 2.10 Γενικευμένη Μέθοδος Ελαχίστου Υπολοίπου GMRES(m)

- 1: Υπολόγισε τα $r^{(0)} = b - Ax^{(0)}$, $\beta = \|r^{(0)}\|$ και $v_1 = r^{(0)}/\beta$
- 2: Κατασκεύασε με αρχικό διάνυσμα το v_1 μια ορθοκανονική βάση και τον πίνακα \overline{H}_m
- 3: Υπολόγισε το $y^{(m)}$ που ελαχιστοποιεί την Ευκλείδεια νόρμα $\|\beta e^{(1)} - \overline{H}_m y\|$

και το $x^{(m)} = x^{(0)} + V_m y^{(m)}$

4: Αν ικανοποιείται κάποιο κριτήριο τερματισμού, τότε Σταμάτα

5: Αλλιώς, θέσε $x^{(0)} = x^{(m)}$ και επανεκτέλεσε

Ως κριτήριο τερματισμού για τη μέθοδο GMRES(m) μπορεί να επιλεγεί το $\|b - Ax^{(m)}\| = |\gamma_{m+1}| < \varepsilon\beta$, όπου ε πολύ μικρός αριθμός (π.χ. 10^{-5}).

Υπενθυμίζεται ότι αν ο πίνακας A δεν είναι συμμετρικός, τότε αυτός λέγεται θετικά ορισμένος αν το συμμετρικό του τμήμα $\frac{A + A^T}{2}$ είναι θετικά ορισμένο. Αποδεικνύεται ότι η GMRES(m) συγκλίνει, αν ο πίνακας συντελεστών είναι συμμετρικός και θετικά ορισμένος.

Θεώρημα 2.4.7. Αν ο πίνακας A είναι συμμετρικός και θετικά ορισμένος, η μέθοδος GMRES(m) συγκλίνει στην λύση x για κάθε $m \geq 1$.

Μια σημαντική δυσκολία της GMRES(m) που συναντάται συχνά στην πράξη είναι ότι μπορεί να τερματίσει χωρίς να δώσει λύση (stagnate) για πίνακες μη θετικά ορισμένους. Οπότε, αν δεν γνωρίζουμε εκ των προτέρων ότι ο πίνακας μας είναι θετικά ορισμένος, ούτε η ιδέα της επανεκκίνησης εγγυάται ταχεία σύγκλιση στη λύση. Ο πλήρης αλγόριθμος της μεθόδου GMRES παρόλα αυτά εγγυάται σύγκλιση σε το πολύ n βήματα. Για πρακτικούς λόγους, μπορούμε να χρησιμοποιήσουμε την τεχνική της προρρυθμίσσης. Στόχος μας είναι να μειώσουμε όσο γίνεται το πλήθος των βημάτων για να επιτύχουμε σύγκλιση. Το πλήθος αυτό είναι πολύ μεγάλο, για μεγάλα n , ειδικά σε προβλήματα με κακή κατάσταση (ο δείκτης κατάστασης $\kappa(A)$ του πίνακα A είναι πολύ μεγάλος). Ακόμη, όπως είναι φανερό, για ακόμη ταχύτερη σύγκλιση μπορούμε να εφαρμόσουμε έναν καλύτερο προρρυθμιστή σε ένα ήδη προρρυθμισμένο σύστημα. Έστω το αρχικό σύστημα $Ax = b$, με $A \in \mathbb{C}^{n,n}$, $\det(A) \neq 0$ και $b \in \mathbb{C}^n$. Έστω ακόμη ο προρρυθμιστής πίνακας $M \in \mathbb{C}^{n,n}$ που επιλέγεται έτσι ώστε να είναι αντιστρέψιμος και να επιλύει το σύστημα σε λιγότερο χρόνο σε σχέση με το χρόνο που θα έκανε έχοντας ως πίνακα συντελεστών τον A . Σημειώνεται ότι η προρρυθμισμένη μέθοδος GMRES(m) επιβαρύνεται με μια επίλυση του συστήματος με τον πίνακα M , ανά επανάληψη. Επομένως ο χρόνος για τη λύση του συστήματος με πίνακα τον M θα πρέπει να είναι κατά τάξη μεγέθους μικρότερος από εκείνον για τη λύση του συστήματος με πίνακα τον A και η μείωση των επαναλήψεων να είναι τόση, ώστε συνολικά να έχουμε όφελος στο χρόνο. Τότε, πολλαπλασιάζοντας από αριστερά το σύστημα $Ax = b$ με τον αντίστροφο του M , έχουμε: $M^{-1}Ax = M^{-1}b$, σύστημα ισοδύναμο με το αρχικό με πίνακα συντελεστών τον $M^{-1}A$ και σταθερό διάνυσμα το $M^{-1}b$. Η διαδικασία που περιγράφηκε αποτελεί την αριστερή προρρυθμίσση (left preconditioning) του συστήματος $Ax = b$. Τότε η νόρμα ενέργειας του διανύσματος - υπόλοιπο ελαχιστοποιείται στον υπόχωρο

Krylov $K_m = \text{span}\{r^{(0)}, M^{-1}Ar^{(0)}, \dots, (M^{-1}A)^{m-1}r^{(0)}\}$. Με παρόμοιο τρόπο λαμβάνουμε τη σχέση: $AM^{-1}c = b$ και $c = Mx$, όπου $c \in \mathbb{C}^n$. Η διαδικασία αυτή αποτελεί τη δεξιά προρρυθμηση του συστήματος. Σε αυτήν την περίπτωση η νόρμα ενέργειας του διανύσματος - υπόλοιπο ελαχιστοποιείται στον υπόχωρο Krylov $K_m = \text{span}\{r^{(0)}, AM^{-1}r^{(0)}, \dots, (AM^{-1})^{m-1}r^{(0)}\}$. Παρακάτω παρατίθενται οι αλγόριθμοι της αριστερής και δεξιάς προρρυθμισμένης μεθόδου GMRES(m) (GMRES(m) with left/right preconditioning) που εφαρμόζονται σε ένα αριστερά και ένα δεξιά προρρυθμισμένο σύστημα, αντίστοιχα.

Αλγόριθμος 2.11 Αριστερά Προρρυθμισμένη GMRES(m)

- 1: Υπολόγισε τα $r^{(0)} = M^{-1}(b - Ax^{(0)})$, $\beta := \|r^{(0)}\|$ και $v_1 = \frac{r^{(0)}}{\beta}$
- 2: Όρισε τον $(m+1) \times m$ πίνακα $\bar{H}_m = h_{ij}$, $1 \leq i \leq m+1$, $1 \leq j \leq m$
- 3: Θέσε $\bar{H}_m = 0$
- 4: **Για** $j = 1, 2, \dots, m$
- 5: $w_j := M^{-1}Av_j$
- 6: **Για** $i = 1, \dots, j$
- 7: $h_{ij} := (w_j, v_i)$
- 8: $w_j := w_j - h_{ij}v_j$
- 9: **Τέλος 'Για'**
- 10: $h_{j+1,j} = \|w_j\|$
- 11: Αν $h_{j+1,j} = 0$ θέσε $m := j$ και πήγαινε στο βήμα 13
- 12: **Τέλος 'Για'**
- 13: Υπολόγισε την $y^{(m)}$ που ελαχιστοποιεί την $\|\beta e^{(1)} - \bar{H}_m y\|$ και το $x^{(m)} = x^{(0)} + V_m y^{(m)}$
- 14: Αν ικανοποιείται κάποιο κριτήριο τερματισμού, τότε Σταμάτα
- 15: Αλλιώς, θέσε $x^{(0)} = x^{(m)}$ και επανεκτέλεσε

Αλγόριθμος 2.12 Δεξιά Προρρυθμισμένη GMRES(m)

- 1: Υπολόγισε τα $r^{(0)} = b - Ax^{(0)}$, $\beta := \|r^{(0)}\|$ και $v_1 = \frac{r^{(0)}}{\beta}$
- 2: Όρισε τον $(m+1) \times m$ πίνακα $\bar{H}_m = h_{ij}$, $1 \leq i \leq m+1$, $1 \leq j \leq m$
- 3: Θέσε $\bar{H}_m = 0$
- 4: **Για** $j = 1, 2, \dots, m$
- 5: $w_j := AM^{-1}v_j$
- 6: **Για** $i = 1, \dots, j$
- 7: $h_{ij} := (w_j, v_i)$
- 8: $w_j := w_j - h_{ij}v_j$
- 9: **Τέλος 'Για'**
- 10: $h_{j+1,j} = \|w_j\|$

- 11: Αν $h_{j+1,j} = 0$ θέσε $m := j$ και πήγαινε στο βήμα 13
- 12: **Τέλος 'Για'**
- 13: Υπολόγισε την $y^{(m)}$ που ελαχιστοποιεί την $\|\beta e^{(1)} - \bar{H}_m y\|$ και το $x^{(m)} = x^{(0)} + V_m y^{(m)}$
- 14: Αν ικανοποιείται κάποιο κριτήριο τερματισμού, τότε Σταμάτα
- 15: Αλλιώς, θέσε $x^{(0)} = x^{(m)}$ και επανεκτέλεσε

Στον αλγόριθμο δεξιά προρρυθμισμένης GMRES(m) το διάνυσμα c που προκύπτει μετά την προρύθμιση δεν χρειάζεται απαραίτητα να οριστεί ρητά. Με το που αρχίζει να εκτελείται ο αλγόριθμος υπολογίζεται το αρχικό διάνυσμα υπόλοιπο $r^{(0)} = b - Ax^{(0)}$ που ισούται στην πράξη με $r^{(0)} = b - AM^{-1}c^{(0)}$, αφού υποθέτουμε ότι το σύστημα είναι δεξιά προρρυθμισμένο. Τα επόμενα βήματα του αλγορίθμου συνεχίζουν να εκτελούνται κανονικά χωρίς να χρησιμοποιηθεί κάπου το διάνυσμα c . Στο τέλος του αλγορίθμου η προσεγγιστική λύση $x^{(m)}$ προκύπτει από το διάνυσμα c από τη σχέση:

$$c^{(m)} = c^{(0)} + V_m y^{(m)}, \quad c^{(0)} = Mx^{(0)}.$$

Πράγματι, αν πολλαπλασιάσουμε αυτή τη σχέση από αριστερά με M^{-1} έχουμε:

$$\begin{aligned} M^{-1}c^{(m)} &= M^{-1}c^{(0)} + M^{-1}V_m y^{(m)} \\ \iff x^{(m)} &= x^{(0)} + M^{-1}V_m y^{(m)}. \end{aligned}$$

ΚΕΦΑΛΑΙΟ 3

Μέθοδοι Διορθογωνοποίησης

3.1 Μέθοδος Δισυζυγών Κλίσεων (Biconjugate Gradient - BiCG)

Στο παρόν κεφάλαιο θα ασχοληθούμε με τον αλγόριθμο διορθογωνοποίησης (biorthogonalization) του Lanczos. Οι μέθοδοι που παρουσιάστηκαν αφορούσαν κυρίως συμμετρικούς πίνακες. Στο παρόν κεφάλαιο θα γίνει επέκταση και στη μη συμμετρική περίπτωση. Τονίζεται ότι η Γενικευμένη Μέθοδος Ελαχίστου Υπολοίπου (GMRES) είναι αυτή που δεν απαιτεί ο πίνακας A να έχει συγκεκριμένη μορφή. Είναι όμως από τις πιο δαπανηρές υπολογιστικά μεθόδους. Αυτό συμβαίνει διότι ανά επανάληψη απαιτεί μεγάλο πλήθος πράξεων, αλλά και μεγάλο χώρο μνήμης. Έτσι επινοήθηκε η ιδέα της επανεκκίνησης, δηλαδή η μέθοδος της GMRES με επανεκκίνηση GMRES(m). Και πάλι όμως υπάρχουν περιπτώσεις αποτυχίας όταν ο πίνακας δεν είναι συμμετρικός.

Οι μέθοδοι που θα παρουσιαστούν στη συνέχεια έχουν ως στόχο να ελαχιστοποιήσουν όσο γίνεται αυτή την αδυναμία και έχουν κάποιες παραπάνω ιδιότητες από αυτές που έχουν ήδη παρουσιαστεί. Υστερούν βέβαια στο γεγονός ότι δεν έχουν ακόμη αποδειχθεί γι' αυτές κάποιες εκ των προτέρων θεωρητικές εκτιμήσεις σφαλμάτων. Οι προσπάθειες μάλιστα που έχουν γίνει προς αυτή την κατεύθυνση είναι μάλλον απαισιόδοξες.

Αρχικά, θα παρουσιαστεί ο αλγόριθμος διορθογωνοποίησης του Lanczos. Έστω το γραμμικό αλγεβρικό σύστημα $Ax = b$, A μη συμμετρικός, $A \in \mathbb{C}^{n,n}$, $b \in \mathbb{C}^n$, $\det(A) \neq 0$. Σε αντιδιαστολή με τον αλγόριθμο του Arnoldi ο οποίος κατασκευάζει μία ορθογώνια ακολουθία διανυσμάτων, ο αλγόριθμος του Lanczos παράγει δύο διορθογώνιες ακολουθίες διανυσμάτων.

Κατασκευάζουμε δύο διορθογώνια σύνολα διανυσμάτων $\{v_1, v_2, \dots, v_{m-1}\}$ και $\{w_1, w_2, \dots, w_{m-1}\}$, τέτοια ώστε να αποτελούν βάσεις των υποχώρων Krylov

$$K_m(A, v_1) = \text{span}\{v_1, Av_1, \dots, A^{m-1}v_1\}$$

$$K_m(A^T, w_1) = \text{span}\{w_1, A^T w_1, \dots, (A^T)^{m-1}w_1\},$$

αντίστοιχα. Η διορθογωνιότητα έγκειται στο γεγονός ότι $(v_j, w_i) = \delta_{ij}$. Στην πράξη επιλύονται συγχρόνως τα συστήματα $Ax = b$, $A^T x = b'$, με πίνακες συντελεστών τους A και A^T αντίστοιχα. Παραθέτουμε τώρα τον αλγόριθμο του Lanczos.

Αλγόριθμος 3.1 Lanczos

- 1: *Επέλεξε* δύο διανύσματα v_1, w_1 με $(v_1, w_1) = 1$
- 2: *Θέσε* $\beta_1 \equiv \delta_1 \equiv 0$, $w_0 \equiv v_0 \equiv 0$
- 3: **Για** $j = 1, 2, \dots, m$
- 4: $\alpha_j = (Av_j, w_j)$
- 5: $\hat{v}_{j+1} = Av_j - \alpha_j v_j - \beta_j v_{j-1}$
- 6: $\hat{w}_{j+1} = A^T w_j - \alpha_j w_j - \delta_j w_{j-1}$
- 7: $\delta_{j+1} = |(\hat{v}_{j+1}, \hat{w}_{j+1})|^{1/2}$
- 8: **Αν** $\delta_{j+1} = 0$
- 9: *Σταμάτα*
- 10: **Τέλος 'Αν'**
- 11: $\beta_{j+1} = \frac{(\hat{v}_{j+1}, \hat{w}_{j+1})}{\delta_{j+1}}$
- 12: $w_{j+1} = \frac{\hat{w}_{j+1}}{\beta_{j+1}}$
- 13: $v_{j+1} = \frac{\hat{v}_{j+1}}{\delta_{j+1}}$
- 14: **Τέλος 'Για'**

Υπάρχουν διάφοροι τρόποι για την επιλογή των δ_{j+1} , β_{j+1} . Σε κάθε περίπτωση απαιτούμε για τα διανύσματα που παράγονται στο βήμα $j + 1$ του αλγορίθμου να ισχύει: $(v_{j+1}, w_{j+1}) = 1$. Από τον αλγόριθμο έχουμε ότι: $w_{j+1} = \frac{\hat{w}_{j+1}}{\beta_{j+1}}$ και $v_{j+1} = \frac{\hat{v}_{j+1}}{\delta_{j+1}}$. Έτσι

$$(v_{j+1}, w_{j+1}) = 1 \iff \frac{1}{\delta_{j+1}\beta_{j+1}}(\hat{v}_{j+1}, \hat{w}_{j+1}) = 1 \iff \delta_{j+1}\beta_{j+1} = (\hat{v}_{j+1}, \hat{w}_{j+1}),$$

οπότε επιλέγονται έτσι ώστε $\delta_{j+1}\beta_{j+1} = (\hat{v}_{j+1}, \hat{w}_{j+1})$. Με τη συγκεκριμένη επιλογή που έγινε τα δ_j είναι θετικά και τέτοια ώστε: $\beta_j = \pm\delta_j$. Ορίζουμε ως

T_m τον τριδιαγώνιο πίνακα

$$T_m = \begin{bmatrix} \alpha_1 & \beta_2 & & & & \\ \delta_2 & \alpha_2 & \beta_3 & & & \\ \cdot & \cdot & \cdot & & & \\ & & \delta_{m-1} & \alpha_{m-1} & \beta_m & \\ & & & \delta_m & \alpha_m & \end{bmatrix}.$$

Πρόταση 3.1.1. Αν ο αλγόριθμος του Lanczos δεν καταρρέει (breakdown) στα πρώτα m βήματα, τα διανύσματα v_i , $i = 1(1)m$, και w_j , $j = 1(1)m$, αποτελούν ένα διορθογώνιο σύστημα διανυσμάτων τέτοιο ώστε

$$(v_j, w_i) = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}, \quad i \geq 1, j \leq m.$$

Επιπλέον, τα $\{v_i\}$ και $\{w_i\}$ αποτελούν βάσεις των $K_m(A, v_1)$ και $K_m(A^T, w_1)$, αντίστοιχα.

Απόδειξη. Θα αποδείξουμε επαγωγικά ότι τα διανύσματα είναι διορθογώνια.

Αρχικά, ισχύει προφανώς $(v_1, w_1) = 1$, αφού έτσι επιλέγονται τα v_1 και w_1 στο πρώτο βήμα του αλγορίθμου.

Υποθέτουμε ότι στο βήμα j ισχύει:

$$(v_j, w_i) = \delta_{ij}, \quad i \geq 1, j \leq m.$$

Θα αποδείξουμε ότι τα v_1, \dots, v_{j+1} και w_1, \dots, w_{j+1} είναι διορθογώνια.

Αρκεί να δειχθεί ότι $(v_{j+1}, w_i) = 0$, για $i \leq j$. Με τον ίδιο τρόπο αποδεικνύεται ότι $(v_i, w_{j+1}) = 0$, για $i \leq j$.

Για $i = j$ έχουμε

$$\begin{aligned} (v_{j+1}, w_j) &= \left(\frac{\hat{v}_{j+1}}{\delta_{j+1}}, w_j \right) = \delta_{j+1}^{-1} [(Av_j - \alpha_j v_j - \beta_j v_{j-1}, w_j)] \\ &= \delta_{j+1}^{-1} [(Av_j, w_j) - \alpha_j (v_j, w_j) - \beta_j (v_{j-1}, w_j)] \\ &= \delta_{j+1}^{-1} [(Av_j, w_j) - (Av_j, w_j)] = 0, \end{aligned}$$

όπου θέσαμε $\alpha_j = (Av_j, w_j)$, είναι $(v_j, w_j) = 1$ εκ κατασκευής και $(v_{j-1}, w_j) = 0$ από την επαγωγική υπόθεση.

Για $i < j$ έχουμε

$$\begin{aligned}
(v_{j+1}, w_i) &= \delta_{j+1}^{-1} [(Av_j, w_i) - \alpha_j(v_j, w_i) - \beta_j(v_{j-1}, w_i)] = \\
&= \delta_{j+1}^{-1} [(v_j, A^T w_i) - \beta_j(v_{j-1}, w_i)] \\
&= \delta_{j+1}^{-1} [(v_j, \beta_{j+1} w_{i+1} + \alpha_i w_i + \delta_i w_{i-1}) - \beta_j(v_{j-1}, w_i)] \\
&= \delta_{j+1}^{-1} [(v_j, \beta_{i+1} w_{i+1}) + (v_j, \alpha_i w_i) + (v_j, \delta_i w_{i-1}) - \beta_j(v_{j-1}, w_i)],
\end{aligned}$$

όπου για $i < j - 1$ από την επαγωγική υπόθεση το δεύτερο μέλος ισούται με 0. Ακόμη, για $i = j - 1$ αυτή γίνεται

$$\begin{aligned}
(v_{j+1}, w_i) &= (v_{j+1}, w_{j-1}) = \delta_{j+1}^{-1} [(v_j, \beta_j w_j + \alpha_{j-1} w_{j-1} + \delta_{j-1} w_{j-2}) \\
&\quad - \beta_j(v_{j-1}, w_{j-1})] \\
&= \delta_{j+1}^{-1} [\beta_j(v_j, w_j) + \alpha_{j-1}(v_j, w_{j-1}) \\
&\quad + \delta_{j+1}(v_j, w_{j-2}) - \beta_j(v_{j-1}, w_{j-1})] = 0
\end{aligned}$$

Επίσης, ισχύει εκ κατασκευής ότι $(v_{j+1}, w_{j+1}) = 1$. Άρα τα v_i, w_i που κατασκευάζονται από τον αλγόριθμο του Lanczos είναι όντως διορθογώνια. \square

Σε μορφή πινάκων ο αλγόριθμος του Lanczos μας λέει ότι

- $AV_m = V_m T_m + \delta_{m+1} v_{m+1} (e^{(m)})^T$
- $A^T W_m = W_m T_m^T + \beta_{m+1} w_{m+1} (e^{(m)})^T$
- $W_m^T AV_m = T_m$

Συμβολίζουμε με V_m τον πίνακα με στήλες τα διανύσματα v_1, \dots, v_m και με W_m τον πίνακα με στήλες τα διανύσματα w_1, \dots, w_m .

Ο αλγόριθμος του Lanczos καταρρέει όταν σε κάποιο βήμα προκύπτει $\delta_{j+1} = 0$. Τότε δεν μπορεί να υπολογιστεί το β_{j+1} και κατά συνέπεια τα v_{j+1}, w_{j+1} . Σε αυτήν την περίπτωση, τα \hat{v}_{j+1} και \hat{w}_{j+1} έχουν εσωτερικό γινόμενο ίσο με το μηδέν. Για να συμβεί αυτό υπάρχουν δύο δυνατοί λόγοι:

1. Κανονικός τερματισμός

Όταν $(\hat{v}_{j+1}, \hat{w}_{j+1}) = 0$ και ένα από τα $\hat{v}_{j+1}, \hat{w}_{j+1}$ είναι μηδέν. Τότε, τα διανύσματα v_1, v_2, \dots, v_m παράγουν έναν A - αναλλοίωτο υπόχωρο ή αντίστοιχα τα διανύσματα w_1, w_2, \dots, w_m παράγουν έναν A^T - αναλλοίωτο υπόχωρο.

2. Κατάρρευση (breakdown)

Όταν $(\hat{v}_{j+1}, \hat{w}_{j+1}) = 0$ και κανένα από τα δύο διανύσματα δεν είναι ίσο με το μηδέν. Τότε, δεν μπορούν να υπολογιστούν από τον αλγόριθμο τα v_{j+1} και w_{j+1} . Σε αυτήν την περίπτωση μπορεί σε επόμενο βήμα του αλγορίθμου να υπάρχουν τέτοια $\hat{v}_{j+1}, \hat{w}_{j+1}$, κ.ο.κ. Τότε, συχνά προτείνεται αντί του αλγορίθμου του Lanczos να χρησιμοποιηθεί κάποια τροποποίησή του, π.χ. Lanczos lookahead. Υπάρχουν όμως και περιπτώσεις όπου αυτό δεν μπορεί να αντιμετωπιστεί.

3.1.1 Μέθοδος Δισυζυγών Κλίσεων - BiConjugate Gradient

Η Μέθοδος Δισυζυγών Κλίσεων είναι μια διαδικασία ορθογώνιας προβολής του υποχώρου

$$K_m(A, v_1) = \text{span}\{v_1, Av_1, \dots, A^{m-1}v_1\}$$

στον

$$K_m(A^T, w_1) = \text{span}\{w_1, A^T w_1, \dots, (A^T)^{m-1}w_1\},$$

με αρχικό διάνυσμα $v_1 = \frac{r^{(0)}}{\|r^{(0)}\|}$ και το w_1 επιλέγεται τέτοιο ώστε $(v_1, w_1) = 1$.

Παρακάτω θα ασχοληθούμε με την κατασκευή της μεθόδου.

Δεδομένων των αρχικών διανυσμάτων v_1 και w_1 κατασκευάζονται από τον αλγόριθμο του Lanczos δύο ακολουθίες διανυσμάτων v_1, \dots, v_m και w_1, \dots, w_m καθώς και ο τριδιαγώνιος πίνακας T_m . Τότε η προσεγγιστική λύση $x^{(m)}$ στο βήμα m του αλγορίθμου δίνεται από τον τύπο

$$x^{(m)} = x^{(0)} + V_m y^{(m)}, \text{ όπου } y^{(m)} = T_m^{-1}(\beta e^{(1)}).$$

Αυτή μπορεί να εκφραστεί ως

$$\begin{aligned} x^{(m)} &= x^{(0)} + V_m y^{(m)} = x^{(0)} + V_m T_m^{-1}(\beta e^{(1)}) \\ &= x^{(0)} + V_m U_m^{-1} L_m^{-1}(\beta e^{(1)}) = x^{(0)} + P_m L_m^{-1}(\beta e^{(1)}), \end{aligned}$$

όπου $T_m = L_m U_m$ η LU παραγοντοποίηση του T_m και $P_m = V_m U_m^{-1}$.

Όπως ακριβώς ορίστηκε ο πίνακας $P_m = V_m U_m^{-1}$ με διανύσματα στήλες τις διευθύνσεις $p^{(i)}$, μπορούμε να ορίσουμε τον πίνακα $P_m^* = W_m L_m^{-T}$ με διανύσματα στήλες τις $p^{*(i)}$. Οι στήλες των δύο αυτών πινάκων είναι A - ορθογώνιες. Αυτό προκύπτει εύκολα ως εξής

$$(P_m^*)^T A P_m = (W_m L_m^{-T})^T A V_m U_m^{-1} = L_m^{-1} W_m^T A V_m U_m^{-1} = L_m^{-1} T_m U_m^{-1} = I.$$

Αλγόριθμος 3.2 Μέθοδος Διςυζυγών Κλίσεων BiCG

- 1: Υπολόγισε $r^{(0)} = b - Ax^{(0)}$
- 2: Επέλεξε $\hat{r}^{(0)}$ τέτοιο ώστε $(r^{(0)}, \hat{r}^{(0)}) \neq 0$
- 3: Θέσε $p^{(0)} := r^{(0)}, p^{*(0)} := \hat{r}^{(0)}$
- 4: **Για** $j = 0, 1, \dots$ υπολόγισε
- 5: $\alpha_j := \frac{(r^{(j)}, r^{*(j)})}{(Ap^{(j)}, p^{*(j)})}$
- 6: $x^{(j+1)} := x^{(j)} + \alpha_j p^{(j)}$
- 7: $r^{(j+1)} := r^{(j)} - \alpha_j Ap^{(j)}$
- 8: $r^{*(j+1)} := r^{*(j)} - \alpha_j A^T p^{*(j)}$
- 9: $\beta_j := \frac{(r^{(j+1)}, r^{*(j+1)})}{(r^{(j)}, r^{*(j)})}$
- 10: $p^{(j+1)} := r^{(j+1)} + \beta_j p^{(j)}$
- 11: $p^{*(j+1)} := r^{*(j+1)} + \beta_j p^{*(j)}$
- 12: **Τέλος 'Για'**

Στον αλγόριθμο της Μεθόδου Διςυζυγών Κλίσεων η εντολή 'Για $j = 0, 1, \dots$ υπολόγισε' εκτελείται μέχρι να υπάρξει σύγκλιση. Μπορεί να επιλεγεί οποιοδήποτε κριτήριο σύγκλισης.

Σχόλιο. Για τον υπολογισμό της νόρμας του $r^{(j)}$ ισχύει $\|r^{(j)}\| = \|b - Ax^{(j)}\| = \|\delta_{j+1}(e^{(j)})^T y^{(j)}\| \|v^{(j+1)}\|$. Αυτό αποδεικνύεται ως εξής

$$\begin{aligned} r^{(j)} &= b - Ax^{(j)} \\ &= b - Ax^{(0)} - AV_j T_j^{-1}(\beta e^{(1)}) \\ &= r^{(0)} - (V_j T_j + \delta_{j+1} v_{j+1} e^{(j)T}) T_j^{-1}(\beta e^{(1)}) \\ &= r^{(0)} - V_j(\beta e^{(1)}) - \delta_{j+1} v_{j+1} e^{(j)T} T_j^{-1} \beta e^{(1)} \\ &= r^{(0)} - \beta v_1 - \delta_{j+1} v_{j+1} e^{(j)T} y^{(j)} \\ &= -\delta_{j+1} e^{(j)T} y^{(j)} v_{j+1}. \end{aligned}$$

Οπότε η νόρμα του διανύσματος - υπόλοιπο μπορεί να υπολογιστεί οικονομικότερα χωρίς απαραίτητα στο j βήμα να υπολογιστεί το $x^{(j)}$. Αυτό μας δίνει τη δυνατότητα να μπορούμε να ελέγχουμε αν ικανοποιείται το δεδομένο κριτήριο σύγκλισης σε κάθε επανάληψη.

3.2 Μέθοδος Ημιελαχίστου Υπολοίπου (Quasi Minimal Residual - QMR)

Κατά την Μέθοδο Ημιελαχίστου Υπολοίπου (QMR) το ζητούμενο είναι η ελαχιστοποίηση της Ευκλείδειας νόρμας του διανύσματος - υπόλοιπο. Αν $x^{(0)}$ είναι η αρχική προσέγγιση και V_m είναι ο πίνακας που περιέχει τα διανύσματα που παράγονται από τον αμφίπλευρο αλγόριθμο του Lanczos, τότε η προσεγγιστική λύση $x^{(m)}$ στο βήμα m του αλγορίθμου δίνεται από τη σχέση:

$$x^{(m)} = x^{(0)} + V_m y^{(m)},$$

όπου το $y^{(m)}$ υπολογίζεται έτσι ώστε να ελαχιστοποιεί την Ευκλείδεια νόρμα του διανύσματος - υπόλοιπο. Ισχύει:

$$\begin{aligned} b - Ax^{(m)} &= b - A(x^{(0)} + V_m y^{(m)}) \\ &= b - Ax^{(0)} - AV_m y^{(m)} \\ &= r^{(0)} - V_{m+1} \bar{T}_m y^{(m)} \\ &= V_{m+1} (\beta e^{(1)} - \bar{T}_m y^{(m)}), \end{aligned}$$

όπου με \bar{T}_m συμβολίζεται ο $(m+1) \times m$ πίνακας που προκύπτει αν επεκτείνουμε τον T_m με την $m+1$ γραμμή $\delta_{m+1}(e^{(m)})^T$.

Οπότε, η Ευκλείδεια νόρμα του διανύσματος - υπόλοιπο δίνεται από τον τύπο $\|r^{(m)}\| = \|V_{m+1}(\beta e^{(1)} - \bar{T}_m y^{(m)})\|$. Παρατηρούμε ότι αν τα διανύσματα v_i κατασκευαστούν έτσι ώστε να είναι ορθοκανονικά (δηλαδή αν V_m ορθομοναδιαίος), για την εύρεση του $y^{(m)}$ αρκεί να λυθεί ένα πρόβλημα ελαχίστων τετραγώνων της μορφής $\operatorname{argmin}_y \|(\beta e^{(1)} - \bar{T}_m y^{(m)})\|$. Στην περίπτωση μας όπου τα v_i , $i = 1(1)m$, δεν είναι (γενικά) ορθογώνια, μπορούμε και πάλι να υπολογίσουμε το $y^{(m)}$ ως το y που ελαχιστοποιεί την ποσότητα $\|\beta e^{(1)} - \bar{T}_m y^{(m)}\|$ και τότε να λάβουμε τη λύση $x^{(m)} = x^{(0)} + V_m y^{(m)}$. Τότε, αν υποθέσουμε ότι τα v_i είναι κανονικοποιημένα ως προς την Ευκλείδεια νόρμα, για τη νόρμα του διανύσματος - σφάλμα ισχύει ότι $\|r^{(m)}\| \leq \sqrt{m+1} \|(\beta e^{(1)} - \bar{T}_m y^{(m)})\|$, αφού $\|V_{m+1}\| \leq \sqrt{m+1}$. Η ελαχιστοποίηση αυτή επιτυγχάνεται με τη σταδιακή παραγοντοποίηση QR του \bar{T}_m χρησιμοποιώντας στροφές Givens. Δηλαδή $R_m = Q_m \bar{T}_m$, όπου $Q_m = \Omega_m \Omega_{m-1} \dots \Omega_1$. Τότε το y που ελαχιστοποιεί την Ευκλείδεια νόρμα $\|\beta e^{(1)} - \bar{T}_m y\|$ δίνεται ως

$$y^{(m)} = R_m^{-1} Q_m (\beta e^{(1)}) = R_m^{-1} g_m,$$

όπου g_m το διάνυσμα \bar{g}_m χωρίς την m συνιστώσα. Το διάνυσμα $x^{(m)}$ δίνεται ως

$$x^{(m)} = x^{(0)} + V_m y^{(m)} = x^{(0)} + V_m R_m^{-1} g_m.$$

Ορίζουμε τώρα τον πίνακα P_m ως $P_m = V_m R_m^{-1}$ ή ισοδύναμα $V_m = P_m R_m$ και η λύση $x^{(m)}$ θα είναι

$$\begin{aligned} x^{(m)} &= x^{(0)} + P_m R_m y^{(m)} \\ &= x^{(0)} + P_m R_m R_m^{-1} g_m \\ &= x^{(0)} + P_m g_m \\ &= x^{(0)} + P_{m-1} g_{m-1} + \gamma_m p^{(m)} \\ &= x^{(m-1)} + \gamma_m p^{(m)}. \end{aligned}$$

Από τη σχέση $V_m = P_m R_m$ δίνεται η στήλη m του P_m , δηλαδή ισχύει $v_m = \sum_{i=m-2}^m r_{im} p_i$, όπου μπορεί να υπολογιστεί το εκάστοτε $p^{(m)}$. Λύνοντας ως προς $p^{(m)}$ έχουμε μια αναδρομική σχέση για τον υπολογισμό του πίνακα P_m και κατά συνέπεια για τη λύση $x^{(m)}$.

Παραθέτουμε τον αλγόριθμο της Μεθόδου Ημιελαχίστου Υπολοίπου.

Αλγόριθμος 3.3 Μέθοδος Ημιελαχίστου Υπολοίπου QMR

- 1: Υπολόγισε $r^{(0)} = b - Ax^{(0)}$ και $\gamma_1 := \|r^{(0)}\|$, $w_1 := v_1 := \frac{r^{(0)}}{\gamma_1}$
- 2: **Για** $m = 1, 2, \dots$ υπολόγισε
- 3: $\alpha_m = (Av_m, w_m)$
- 4: $\hat{v}_{m+1} = Av_m - \alpha_m v_m - \beta_m v_{m-1}$
- 5: $\hat{w}_{m+1} = A^T w_m - \alpha_m w_m - \delta_m w_{m-1}$
- 6: $\delta_{m+1} = |(\hat{v}_{m+1}, \hat{w}_{m+1})|^{1/2}$
- 7: **Αν** $\delta_{m+1} = 0$
- 8: **Σταμάτα**
- 9: **Τέλος 'Αν'**
- 10: $\beta_{m+1} = \frac{(\hat{v}_{m+1}, \hat{w}_{m+1})}{\delta_{m+1}}$
- 11: $w_{m+1} = \frac{\hat{w}_{m+1}}{\beta_{m+1}}$
- 12: $v_{m+1} = \frac{\hat{v}_{m+1}}{\delta_{m+1}}$
- 13: **Κατασκεύασε τον τριδιαγώνιο πίνακα \bar{T}_m , δηλ. θέσε $t_{mm} = \alpha_m$, $t_{m,m+1} = \beta_{m+1}$, $t_{m+1,m} = \delta_{m+1}$**
- 14: **Ανακάλεσε και εφάρμοσε τις στροφές Givens $\Omega_{m-2}, \Omega_{m-1}$ στην m στήλη του \bar{T}_m**
- 15: **Υπολόγισε τους συντελεστές c_m και s_m της στροφής Givens Ω_m**
- 16: **Εφάρμοσε τη στροφή Givens Ω_m στον \bar{T}_m και \bar{g}_m , δηλ. υπολόγισε**

- 17: $\gamma_{m+1} = -s_m \gamma_m$
 18: $\gamma_m = c_m \gamma_m$
 19: $\alpha_m = c_m \alpha_m + s_m \delta_{m+1} = \sqrt{\delta_{m+1}^2 + \alpha_m^2}$
 20: $p^{(m)} = (v_m - \sum_{i=m-2}^{m-1} t_{im} p^{(i)}) / t_{mm}$
 21: $x^{(m)} = x^{(m-1)} + \gamma_m p^{(m)}$
 22: Αν $|\gamma_{m+1}|$ είναι αρκετά μικρή ποσότητα, τότε Σταματά
 23: **Τέλος 'Για'**

Θεώρημα 3.2.1. Αν υποθέσουμε ότι ο αλγόριθμος του Lanczos δεν καταρρέει μέχρι το βήμα m και συμβολίσουμε με $r^{Q(m)}$, $r^{G(m)}$ τα διανύσματα - υπόλοιπο που αντιστοιχούν στις μεθόδους QMR και GMRES, αντίστοιχα, τότε για τις Ευκλείδειες νόρμες τους ισχύει:

$$\|r^{Q(m)}\| \leq \kappa_2(V_{m+1}) \|r^{G(m)}\|,$$

όπου V_{m+1} ο πίνακας V της QMR.

Σχόλιο. Είναι δύσκολο να βρεθεί ένα φράγμα για τον δείκτη κατάστασης του πίνακα V_{m+1} , ο οποίος σε κάποιες περιπτώσεις ενδέχεται να μεγαλώσει αρκετά. Επομένως, για τη μέθοδο QMR δεν μπορεί να δοθεί κάποια εκ των προτέρων εκτίμηση για τη νόρμα του διανύσματος - υπόλοιπο. Από την άποψη αυτή η QMR υστερεί της GMRES. Από την άποψη όμως της πολυπλοκότητας ανά επανάληψη η QMR υπερτερεί σημαντικά της GMRES.

3.3 Τετραγωνική Μέθοδος Συζυγών Κλίσεων (Conjugate Gradient Squared - CGS)

Οι δύο τελευταίες μέθοδοι που αναπτύχθηκαν, δηλαδή οι μέθοδοι BiCG και QMR, απαιτούν σε κάθε επανάληψη πολλαπλασιασμούς των πινάκων A και A^T επί κάποια διανύσματα. Είναι δυνατόν να κατασκευαστεί μια νέα μέθοδος που χρησιμοποιεί σε όλους αυτούς τους πολλαπλασιασμούς μόνο τον πίνακα A και είναι απαλλαγμένη από τον A^T . Στόχος της είναι η ελαχιστοποίηση της Ευκλείδειας νόρμας του διανύσματος - υπόλοιπο σε έναν υπόχωρο Krylov της μορφής $K_m = \{v_1, Av_1, \dots, A^{m-1}v_1\}$. Η ιδέα κατασκευής της μεθόδου βασίζεται στην εξής απλή παρατήρηση:

Στην επανάληψη j , για το διάνυσμα - υπόλοιπο της Μεθόδου Δισυζυγών Κλίσεων ισχύει

$$r^{(j)} = \phi_j(A)r^{(0)},$$

όπου το ϕ_j είναι ένα συγκεκριμένο πολυώνυμο βαθμού το πολύ j με $\phi_j(0) = 1$. Από αυτή τη σχέση προκύπτει ότι $\|r^{(j)}\| = \|\phi_j(A)r^{(0)}\|$. Αν τροποποιήσουμε τον αλγόριθμο ώστε το διάνυσμα υπόλοιπο κατά την επανάληψη j να δίνεται ως $r^{(j)} = \phi_j^2(A)r^{(0)}$, κατά τεκμήριο θα έχουμε ταχύτερη σύγκλιση και στην ιδανική περίπτωση διπλάσια ταχύτητα.

Η μέθοδος που θα κατασκευαστεί στη συνέχεια, έχει αφενός περίπου το ίδιο υπολογιστικό κόστος με τη Μέθοδο Διςζυγών Κλίσεων (πράγμα που είναι φανερό από την κατασκευή της), πετυχαίνει δε ταχύτερη σύγκλιση συγκριτικά με τις μεθόδους BiCG και QMR που βασίζονται στον αμφίπλευρο αλγόριθμο του Lanczos. Συμπερασματικά, θα λέγαμε ότι η μέθοδος CGS είναι μια μέθοδος με στόχο τη βελτίωση των μεθόδων BiCG και QMR. Παρακάτω θα ασχοληθούμε με την κατασκευή της μεθόδου:

Για το διάνυσμα - υπόλοιπο της Μεθόδου Διςζυγών Κλίσεων στο βήμα j ισχύει

$$r^{(j)} = \phi_j(A)r^{(0)},$$

όπου το ϕ_j είναι ένα συγκεκριμένο πολυώνυμο βαθμού το πολύ j με $\phi_j(0) = 1$. Ομοίως, για τη συζυγή διεύθυνση $p^{(j)}$ ισχύει

$$p^{(j)} = \psi_j(A)r^{(0)},$$

όπου ψ_j είναι πολυώνυμο βαθμού το πολύ j .

Από τον αλγόριθμο της Μεθόδου Διςζυγών Κλίσεων είναι φανερό ότι οι ίδιες σχέσεις ισχύουν για τον πίνακα A^T , δηλαδή

$$\begin{aligned}\widehat{r}^{(j)} &= \phi_j(A^T)\widehat{r}^{(0)} \\ \widehat{p}^{(j)} &= \psi_j(A^T)\widehat{r}^{(0)}.\end{aligned}$$

Επίσης, από τον ίδιο αλγόριθμο έχουμε ότι

$$\alpha_j = \frac{(p^j, p^{*(j)})}{(Ap^{(j)}, p^{(j)})} = \frac{(\phi_j(A)r^{(0)}, \phi_j(A^T)\widehat{r}^{(0)})}{(A\psi_j(A)r^{(0)}, \psi_j(A^T)\widehat{r}^{(0)})} = \frac{(\phi_j^2(A)r^{(0)}, \widehat{r}^{(0)})}{(A\psi_j^2(A)r^{(0)}, \widehat{r}^{(0)})},$$

από γνωστές ιδιότητες εσωτερικού γινομένου. Για τον υπολογισμό των α_j, β_j αρκεί να βρεθεί μια αναδρομική σχέση (recursion) που περιέχει τα διανύσματα $\phi_j^2(A)r^{(0)}, \psi_j^2(A)r^{(0)}$. Τότε, για τα διανύσματα - υπόλοιπο της μεθόδου θα ισχύει

$$r'^{(j)} = \phi_j^2(A)r^{(0)}.$$

Από τον αλγόριθμο Διςζυγών Κλίσεων έχουμε ότι

$$\begin{aligned}\phi_{j+1}(t) &= \phi_j(t) - \alpha_j t \psi_j(t) \\ \psi_{j+1}(t) &= \phi_{j+1}(t) + \beta_j \psi_j(t).\end{aligned}$$

Υψώνοντας στο τετράγωνο τις δύο παραπάνω σχέσεις παίρνουμε

$$\begin{aligned}\phi_{j+1}^2(t) &= (\phi_j(t) - \alpha_j t \psi_j(t))^2 = \phi_j^2(t) - 2\alpha_j \psi_j(t) \phi_j(t) + \alpha_j^2 t^2 \psi_j^2(t) \\ \psi_{j+1}^2(t) &= (\phi_{j+1}(t) + \beta_j \psi_j(t))^2 = \phi_{j+1}^2(t) + 2\beta_j \phi_{j+1}(t) \psi_j(t) + \beta_j^2 \psi_j^2(t).\end{aligned}$$

Είναι φανερό ότι αν έλειπαν οι όροι $-2\alpha_j \psi_j(t) \phi_j(t)$ και $2\beta_j \phi_{j+1}(t) \psi_j(t)$, θα προέκυπτε ένα σύστημα δύο αναδρομικών σχέσεων για τον υπολογισμό των ϕ_j , ψ_j . Στην περίπτωσή μας, αρκεί να θέσουμε κάποιο από τα δύο γινόμενα (έστω το $\phi_{j+1}(t) \psi_j(t)$) ως τρίτη σχέση του αναδρομικού συστήματος. Τότε, για το $\phi_j(t) \psi_j(t)$ έχουμε

$$\begin{aligned}\phi_j(t) \psi_j(t) &= \phi_j(t) [\phi_j(t) + \beta_{j-1} \psi_{j-1}(t)] \\ &= \phi_j^2(t) + \beta_{j-1} \phi_j(t) \psi_{j-1}(t).\end{aligned}$$

Τελικά, έχουμε το αναδρομικό σύστημα τριών σχέσεων

$$\begin{aligned}\phi_{j+1}^2(t) &= \phi_j^2(t) - \alpha_j t (2\phi_j^2(t) + 2\beta_{j-1} \phi_j(t) \psi_{j-1}(t) - \alpha_j^2 t^2 \psi_j^2(t)) \\ \phi_{j+1} \psi_j(t) &= \phi_j^2(t) + \beta_{j-1} \phi_j(t) \psi_{j-1}(t) - \alpha_j t \psi_j^2(t) \\ \psi_{j+1}^2 &= \phi_{j+1}^2(t) + 2\beta_j \phi_{j+1}(t) \psi_j(t) + \beta_j^2 \psi_j^2(t).\end{aligned}$$

Θέτοντας

$$\begin{aligned}r^{(j)} &\equiv \phi_j^2(A) r^{(0)} \\ p^{(j)} &\equiv \psi_j^2(A) r^{(0)} \\ q^{(j)} &\equiv \phi_{j+1}(A) \psi_j(A) r^{(0)},\end{aligned}$$

το αναδρομικό σύστημα γίνεται:

$$\begin{aligned}r^{(j+1)} &= r^{(j)} - \alpha_j A (2r^{(j)} + 2\beta_{j-1} q^{(j-1)} - \alpha_j A p^{(j)}) \\ q^{(j+1)} &= r^{(j)} + \beta_{j-1} q^{(j-1)} - \alpha_j A p^{(j)} \\ p^{(j+1)} &= r^{(j+1)} + 2\beta_j q^{(j)} + \beta_j^2 p^{(j)}.\end{aligned}$$

Οι τρεις παραπάνω σχέσεις αποτελούν τον πυρήνα της Τετραγωνικής Μεθόδου Συζυγών Κλίσεων (CGS), ο οποίος προκύπτει με την ανάλυση που έχουμε κάνει από αυτόν της Μεθόδου Δισυζυγών Κλίσεων.

Παρατηρούμε ότι ο αλγόριθμος είναι απαλλαγμένος από τον A^T , ενώ περιέχει δύο πολλαπλασιασμούς του πίνακα A επί κάποιο διάνυσμα ανά επανάληψη. Αν γνωρίζουμε ότι η BiCG συγκλίνει, θα ήταν αναμενόμενο για τη σύγκλιση της CGS, όπου δεν υπάρχουν πολλαπλασιασμοί του πίνακα A^T επί διάνυσμα, να απαιτούνται οι μισές επαναλήψεις. Μια σημαντική δυσκολία της μεθόδου CGS οφείλεται στο

γεγονός ότι τα πολυώνυμα υψώνονται στο τετράγωνο. Αυτό μπορεί να επιφέρει σφάλματα στρογγύλευσης σε σχέση με την μέθοδο BiCG. Αυτό, σε περίπτωση μη σύγκλισης της μεθόδου BiCG, όπου υπάρχουν μεγάλες αυξομειώσεις στην νόρμα του διανύσματος - υπόλοιπο από την μια επανάληψη στην άλλη, θα επιφέρει ακόμη μεγαλύτερες αυξομειώσεις στη νόρμα της μεθόδου CGS.

Αλγόριθμος 3.4 Τετραγωνική Μέθοδος Συζυγών Κλίσεων

- 1: *Επέλεξε* $\hat{r}^{(0)}$
- 2: *Υπολόγισε* $r^{(0)} = b - Ax^{(0)}$
- 3: *Θέσε* $p^{(0)} = u^{(0)} = r^{(0)}$
- 4: **Για** $j = 0, 1, 2, \dots$
- 5: $\alpha_j = \frac{(r^{(j)}, r^{(0)})}{(Ap^{(j)}, r^{(0)})}$
- 6: $q^{(j)} = u^{(j)} - \alpha_j Ap^{(j)}$
- 7: $x^{(j+1)} = x^{(j)} + \alpha_j(u^{(j)} + q^{(j)})$
- 8: $r^{(j+1)} = r^{(j)} - \alpha_j A(u^{(j)} + q^{(j)})$
- 9: $\beta_j = \frac{(r^{(j+1)}, \hat{r}^{(0)})}{(r^{(j)}, \hat{r}^{(0)})}$
- 10: $u^{(j+1)} = r^{(j+1)} + \beta_j q^{(j)}$
- 11: $p^{(j+1)} = u^{(j+1)} + \beta_j(q^{(j)} + \beta_j p^{(j)})$
- 12: **Τέλος 'Για'**

Στον αλγόριθμο τετραγωνικής μεθόδου συζυγών κλίσεων η εντολή Για $j = 0, 1, 2, \dots$ εκτελείται μέχρις ότου να ικανοποιηθεί κάποιο κριτήριο σύγκλισης.

3.4 Ευσταθειοποιημένη Μέθοδος Δισυζυγών Κλίσεων (BiConjugate Gradient Stabilized - BiCGSTAB)

Όπως είδαμε, η Τετραγωνική Μέθοδος Συζυγών Κλίσεων βασίζεται στον τετραγωνισμό του διανύσματος - υπόλοιπο. Οπότε, όπως είναι αναμενόμενο, η σύγκλιση της μεθόδου σχετίζεται με το τετράγωνο του διανύσματος - υπόλοιπο. Έτσι, σε περιπτώσεις μη ομαλής σύγκλισης είναι αναμενόμενο να έχει επηρεαστεί αρκετά από σφάλματα στρογγύλευσης. Η Ευσταθειοποιημένη Μέθοδος Δισυζυγών Κλίσεων BiCGSTAB προσπαθεί να αντιμετωπίσει αυτή την αδυναμία, όντας πιο ευσταθής στην παρουσία σφαλμάτων συγκριτικά με την Τετραγωνική Μέθοδο Δισυζυγών Κλίσεων. Έχει ως στόχο την ελαχιστοποίηση της Ευκλείδειας νόρμας του διανύσματος - υπόλοιπο σε έναν υπόχωρο Krylov της μορ-

φής $K_m = \{v_1, Av_1, \dots, A^{(m-1)}v_1\}$. Ως διάνυσμα - υπόλοιπο όμως επιλέγεται μια τροποποιημένη μορφή. Συγκεκριμένα, το διάνυσμα - υπόλοιπο της μεθόδου BiCGSTAB είναι

$$r^{(j)} = \chi^{(j)}(t)\phi^{(j)}(t)r^{(0)},$$

όπου $\phi^{(j)}(t)$ είναι το πολυώνυμο που παράγεται μέσω του αλγόριθμου της Μεθόδου Δισυζυγών Κλίσεων και το $\chi^{(j)}(t)$ είναι ένα νέο πολυώνυμο που ορίζεται από μια αναδρομική σχέση και έχει στόχο να «ευσταθειοποιήσει» τη σύγκλιση του αλγορίθμου. Αυτό γίνεται, αν επιλεγεί τέτοιο ώστε σε κάθε βήμα του αλγορίθμου να διατηρεί τη νόρμα του διανύσματος - υπόλοιπο μικρή, διατηρώντας την ταχεία σύγκλιση της Τετραγωνικής Μεθόδου Συζυγών Κλίσεων. Το $\chi(t)$ ορίζεται από την αναδρομική σχέση

$$\chi^{(j+1)}(t) = (1 - \omega_j t)\chi^{(j)}(t),$$

όπου ω_j προσδιοριστέο.

Το διάνυσμα $p^{(j)}$ ορίζεται ως

$$p^{(j)}(t) = \phi^{(j)}(t)\chi^{(j)}(t).$$

Η διαδικασία κατασκευής της μεθόδου είναι παραπλήσια με αυτή της CGS. Αρχικά, αναζητούμε μία αναδρομική σχέση για τα $r^{(j)}$, $p^{(j)}$.

Έχουμε

$$\begin{aligned} r^{(j+1)}(t) &= \chi^{(j+1)}(t)\phi^{(j+1)}(t) \\ &= (1 - \omega_j t)\chi^{(j)}(t)(\phi^{(j)}(t) - \alpha_j t\psi^{(j)}(t)), \end{aligned}$$

αφού από CGS ισχύει $\phi^{(j+1)}(t) = \phi^{(j)}(t) - \alpha_j t\psi^{(j)}(t)$.

Οπότε

$$r^{(j+1)}(t) = (1 - \omega_j t)(r^{(j)}(t) - \alpha_j t p^{(j)}(t)).$$

Τελικά $r^{(j+1)}(t) = (1 - \omega_j A)(r^{(j)} - \alpha_j A p^{(j)})$.

Θέτουμε $s^{(j)} = r^{(j)} - \alpha_j A p^{(j)}$, οπότε $r^{(j+1)} = (1 - \omega_j A) s^{(j)}$. Η παράμετρος ω , δίνεται από αυτή τη σχέση ελαχιστοποιώντας την $\|r^{(j+1)}\|$ και είναι η εξής:

$$\omega_j = \frac{(As^{(j)}, s^{(j)})}{(As^{(j)}, As^{(j)})}.$$

Ομοίως, για τον υπολογισμό των $p^{(j)}$ έχουμε

$$\begin{aligned} p^{(j+1)}(t) &= \chi^{(j+1)}(t)\psi^{(j+1)}(t) \\ &= \chi^{(j+1)}(t)(\phi^{(j)}(t) + \beta_j \psi^{(j)}(t)), \end{aligned}$$

αφού από CGS ισχύει $\phi^{(j+1)}(t) = \phi^{(j)}(t) + \beta_j \psi^{(j)}(t)$.

Έτσι

$$p^{(j+1)}(t) = r^{(j)}(t) + \beta_j(1 - \omega_j t)p^{(j)}(t).$$

Τελικά έχουμε $p^{(j+1)} = r^{(j)} + \beta_j(1 - \omega_j A)p^{(j)}$.

Μένει να εκφράσουμε τους συντελεστές α_j, β_j της Μεθόδου Διςυζυγών Κλίσεων συναρτήσει των νέων διανυσμάτων $r^{(j)}, p^{(j)}$. Στον αλγόριθμο BiCG ισχύει η σχέση διορθωσιμότητας $W_m^T A V_m = T_m$.

Χρησιμοποιώντας τις εκφράσεις των ϕ_j, ψ_j , καθώς και την παραπάνω σχέση διορθωσιμότητας έχουμε ότι τα α_j της Ευσταθειοποιημένης Μεθόδου Διςυζυγών Κλίσεων είναι

$$\begin{aligned} \alpha_j &= \frac{(\phi_j(A)r^{(0)}, \phi_j(A^T)\hat{r}^{(0)})}{(A\psi_j(A)r^{(0)}, \phi_j(A^T)\hat{r}^{(0)})} \\ &= \frac{(\phi_j(A)r^{(0)}, c_j(A^T)^j \hat{r}^{(0)})}{(A\psi_j(A)r^{(0)}, c_j(A^T)^j \hat{r}^{(0)})} \\ &= \frac{\left(\phi_j(A)r^{(0)}, \frac{c_j}{c_j} \chi_j(A^T)\hat{r}^{(0)} \right)}{\left(A\psi_j(A)r^{(0)}, \frac{c_j}{c_j} \chi_j(A^T)\hat{r}^{(0)} \right)} \\ &= \frac{(r^{(k-1)}, \hat{r}^{(0)})}{(Ap^{(k-1)}, \hat{r}^{(0)})}, \end{aligned}$$

όπου με c_j συμβολίζουμε τον συντελεστή του μεγιστοβαθμίου όρου του πολυωνύμου που αντιστοιχεί στον πίνακα A και με c_j' αυτόν που αντιστοιχεί στον A^T .

Τελικά προκύπτει ότι $\alpha_{k-1} = \frac{(r^{(k-1)}, \hat{r}^{(0)})}{(Ap^{(k-1)}, \hat{r}^{(0)})}$ και με παρόμοιο τρόπο $\beta_k = \frac{\alpha_{k-1}}{\omega_k} \frac{(r^{(k)}, \hat{r}^{(0)})}{(r^{(k-1)}, \hat{r}^{(0)})}$, $k = 1, 2, \dots$

Παραθέτουμε τον αλγόριθμο της Ευσταθειοποιημένης Μεθόδου Διςυζυγών Κλίσεων που προτάθηκε από τον van der Vorst.

Αλγόριθμος 3.5 BiCGSTAB

- 1: *Επέλεξε* $\hat{r}^{(0)}$
- 2: *Υπολόγισε* $r^{(0)} = b - Ax^{(0)}$
- 3: $p^{(0)} = r^{(0)}$
- 4: **Για** $j = 0, 1, \dots$

- 5: $\alpha_j = \frac{(r^{(j)}, \hat{r}^{(0)})}{(Ap^{(j)}, \hat{r}^{(0)})}$
- 6: $s^{(j)} = r^{(j)} - \alpha_j Ap^{(j)}$
- 7: $\omega_j = \frac{(As^{(j)}, s^{(j)})}{(As^{(j)}, As^{(j)})}$
- 8: $x^{(j+1)} = x^{(j)} + \alpha_j p^{(j)} + \omega_j s^{(j)}$
- 9: $r^{(j+1)} = s^{(j)} - \omega_j As^{(j)}$
- 10: $\beta_j = \frac{(r^{(j+1)}, \hat{r}^{(0)})}{(r^{(j)}, \hat{r}^{(0)})} \times \frac{\alpha_j}{\omega_j}$
- 11: $p^{(j+1)} = r^{(j+1)} + \beta_j (p^{(j)} - \omega_j Ap^{(j)})$
- 12: **Τέλος 'Για'**

Επίλογος

Έγινε μια συνοπτική παρουσίαση των κυριότερων μεθόδων ελαχιστοποίησης ή μεθόδων υποχώρων Krylov. Η προσοχή μας εστιάστηκε στο να τονίσουμε τα μειονεκτήματα της κάθε μεθόδου, επιδιώκοντας κάθε φορά τη βελτίωση τους. Σε κάθε περίπτωση, δύναται σε μια μέθοδο να χρησιμοποιηθεί κατάλληλος προρρυθμιστής πίνακας για την επίτευξη καλύτερων αριθμητικών αποτελεσμάτων. Αξίζει να τονιστεί ότι η χρήση προρρυθμιστή δεν αλλάζει τη μέθοδο, απλώς επιφέρει αλλαγές στον αλγόριθμο της μεθόδου. Είδαμε ότι για ερμιτιανούς πίνακες η εξέλιξη σταματά στην ουσία στη Μέθοδο Συζυγών Κλίσεων (CG). Κάθε άλλη περαιτέρω προσπάθεια γίνεται προς την κατεύθυνση της προρρύθμισης. Στην περίπτωση όπου οι πίνακες δεν απαιτείται να είναι ερμιτιανοί η εξέλιξη των μεθόδων σταματά στη Γενικευμένη Μέθοδο Ελαχίστου Υπολοίπου (GMRES). Τότε κάθε άλλη προσπάθεια αφορά ειδικές μορφές της, όπως η Γενικευμένη Μέθοδος Ελαχίστου Υπολοίπου με Επανεκκίνηση (GMRES(m)) καθώς και προρρυθμισμένη GMRES. Σε πίνακες που έχουν κάποια (γνωστή) ειδική μορφή, όλες οι μέθοδοι που μελετήθηκαν θα μπορούσαν να τροποποιηθούν κατάλληλα, έτσι ώστε να αξιοποιήσουμε κάθε φορά τη δομή των πινάκων αυτών. Προς αυτή την κατεύθυνση δε, γίνεται πρόταση για περαιτέρω έρευνα. Αναφορικά με τα σφάλματα των μεθόδων, παρατίθενται οι εκτιμήσεις των περισσότερων από αυτές, για τις οποίες χρησιμοποιούνται εργαλεία της Ανάλυσης, της Αριθμητικής Γραμμικής Άλγεβρας και της Θεωρίας Προσέγγισης. Στις μεθόδους που βασίζονται στον αμφίπλευρο αλγόριθμο του Lanczos επισημαίνεται ότι δεν έχουν ακόμη αποδειχθεί εκτιμήσεις σφαλμάτων, ενώ οι προοπτικές μοιάζουν απαισιόδοξες.

Βιβλιογραφία

- [1] Axelsson, O. *Iterative Solution Methods*. Cambridge University Press, London, 1994.
- [2] Brezinski, C., Zaglia, M Redivo, and Sadok, H. “Avoiding Breakdown and Near-Breakdown in Lanczos Type Algorithms”. In: *Numer. Algorithms* 1.2 (1991), pp. 261–284.
- [3] Demmel, J.W. *Applied Numerical Linear Algebra*. SIAM, Philadelphia, 1997.
- [4] Δουγαλής, Β., Νούτσος, Δ. και Χατζηδήμος, Α. *Αριθμητική γραμμική άλγεβρα*. Πανεπιστημιακό Τυπογραφείο Ιωαννίνων, 2009.
- [5] Faber, V. and Manteuffel, T.A. “Orthogonal Error Methods”. In: *SIAM J. Numer. Anal.* 24.1 (1987), pp. 170–187.
- [6] Gallopoulos, Ef., Philippe, B., and Sameh, A.H. *Parallelism in Matrix Computations*. Springer, 2016.
- [7] Golub, G. and Van Loan, C. *Matrix computations*. Vol. 3. JHU Press, 2012.
- [8] Parlett, B.N., Taylor, D.R., and Liu, Z.A. “A Look-Ahead Lanczos Algorithm for Unsymmetric Matrices”. In: *Math. Comp.* 44.169 (1985), pp. 105–124.
- [9] Saad, Y. *Iterative Methods for Sparse Linear Systems*. 2nd ed. SIAM, 2003.
- [10] Saad, Y. “Krylov Subspace Methods for Solving Large Unsymmetric Linear Systems”. In: *Math. Comp.* 37.155 (1981), pp. 105–126.

- [11] Saad, Y. and Schultz, M.H. “GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems”. In: *SIAM J. Sci. Statist. Comput.* 7.3 (1986), pp. 856–869.