ΕΠΕΚΤΕΙΝΟΝΤΑΣ ΤΟΝ ΑΛΓΟΡΙΘΜΟ “DOG-LEG” ΜΕ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ ΕΝΤΟΣ ΥΠΟΧΩΡΩΝ

Η
ΜΕΤΑΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ ΕΞΕΙΔΙΚΕΥΣΗΣ

Υποβάλλεται στην

ορισθείσα από την Γενική Συνέλευση Ειδικής Σύνθεσης
του Τμήματος Πληροφορικής
Εξεταστική Επιτροπή

από την

Αρσινόη Ειρήνη Ρίζου

ως μέρος των Υποχρεώσεων

για τη λήψη

του

ΜΕΤΑΠΤΥΧΙΑΚΟΥ ΔΙΠΛΩΜΑΤΟΣ ΣΤΗΝ ΠΛΗΡΟΦΟΡΙΚΗ

ΜΕ ΕΞΕΙΔΙΚΕΥΣΗ ΣΤΟΥΣ ΕΠΙΣΤΗΜΟΝΙΚΟΥΣ ΥΠΟΛΟΓΙΣΜΟΥΣ

Δεκέμβριος 2018

Εξεταστική Επιτροπή:

- Ισαάκ Η. Λαγαρής, Καθηγητής, Τμήμα Μηχανικών Η/Υ & Πληροφορικής, Πανεπιστήμιο Ιωαννίνων
- Κωνσταντίνος Παρσόπουλος, Αναπληρωτής Καθηγητής, Τμήμα Μηχανικών Η/Υ & Πληροφορικής, Πανεπιστήμιο Ιωαννίνων
- Χριστόφορος Νίκου, Καθηγητής, Τμήμα Μηχανικών Η/Υ & Πληροφορικής, Πανεπιστήμιο Ιωαννίνων

ΑΦΙΕΡΩΣΗ

Στους γονείς μου Σταύρο και Ευαγγελία

ΕΥΧΑΡΙΣΤΙΕΣ

Χωρίς την παρουσία, την ειλικρινή υποστήριξη και ενθάρρυνση κάποιων ανθρώπων δεν θα ήταν δυνατή η υλοποίηση αυτής της μεταπτυχιακής διατριβής.

Πρώτα από όλα, θα ήθελα από καρδιάς να ευχαριστήσω τον επιβλέποντα μου, κύριο Ισαάκ Λαγαρή, Καθηγητή του τμήματος Μηχανικών Η/Υ & Πληροφορικής για την επιστημονική, πνευματική και ηθική υποστήριξη που μου παρείχε σε όλη την διάρκεια της εργασίας μου.

Ιδιαίτερη μνεία, αξίζει ο Λάμπρος, το αδερφάκι μου, για την απύθμενη αγάπη και την στήριξη του, μα και για το γεγονός ότι αποτελεί πρότυπο και επιρροή στην ζωή μου.

Ως πράξη ευγνωμοσύνης, μιας και είναι το λιγότερο που μπορώ να κάνω για αυτούς τους δύο ανθρώπους, ευχαριστώ μέσα από την ψυχή μου την αδερφική μου φίλη Έλενα και το «Παιδί Χρυσάφι», τον θετικό μου άνθρωπο, που χωρίς τον γλυκό τους ψυχαναγκασμό για την περάτωση της εργασίας μου, το αμείωτο ενδιαφέρον και την ατέρμονη στήριξη τους δεν θα ήταν τίποτα το ίδιο. Με τέτοιους ανθρώπους πηγαίνεις πάντα ένα βήμα παραπέρα.

Για τέλος, δεν γινόταν να παραλείψω να αναφέρω έναν πολύ ιδιαίτερο άνθρωπο για εμένα, την φίλη μου Κατερίνα που με γνωρίζει τόσο καλά όσο και τον εαυτό της. Αν και μακριά μου πάντα βρίσκει ένα μοναδικό τρόπο να είναι τόσο κοντά μου.

ΠΕΡΙΕΧΟΜΕΝΑ

	Σελ
ΑΦΙΕΡΩΣΗ	iii
ΕΥΧΑΡΙΣΤΙΕΣ	iv
ΠΕΡΙΕΧΟΜΕΝΑ	v
ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ	vi
ΕΥΡΕΤΗΡΙΟ ΣΧΗΜΑΤΩΝ	vii
ΕΠΕΞΗΓΗΣΕΙΣ ΣΥΜΒΟΛΙΣΜΩΝ	viii
ΠΕΡΙΛΗΨΗ	ix
EXTENDED ABSTRACT IN ENGLISH	xi
ΚΕΦΑΛΑΙΟ 1. ΕΙΣΑΓΩΓΗ	1
1.1. Στόχοι	1
1.2. Δομή της Διατριβής	2
ΚΕΦΑΛΑΙΟ 2. ΕΙΣΑΓΩΓΙΚΕΣ ΈΝΝΟΙΕΣ ΕΠΙ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ	4
2.1. Μαθηματικός Συμβολισμός	4
2.2. Κατηγορίες Βελτιστοποίησης	5
2.2.1. Υπό και Δίχως περιορισμούς	5
2.2.2. Διακριτή και Συνεχής	6
2.2.3. Καθολική και Τοπική	7
2.2.4. Αιτιοκρατική και Στοχαστική	8
2.3. Τι ορίζεται ως λύση	8
2.4. Αναγωνρίζοντας ένα τοπικό ελάχιστο	10
2.5. ΑλγόριθμοιΒελτιστοποίησης	13
ΚΕΦΑΛΑΙΟ 3. ΘΕΜΕΛΙΩΔΗΣ ΑΝΑΛΥΣΗ ΠΡΟΑΠΑΙΤΟΥΜΕΝΩΝ	16
3.1. Πολυδιάστατες Τετραγωνικές Συναρτήσεις	16
3.1.1. Τετραγωνικές Συναρτήσεις	16
3.1.2. Τετραγωνικά Μοντέλα	18
3.2. Μέθοδοι Αριθμητικής Ανάλυσης και Άλγεβρας	19
3.2.1. Μέθοδος Cholesky	19
3.2.2. Μέθοδοι Εύρεσης Ριζών	21
3.3. Στρατηγική "Περιοχής Εμπιστοσύνης"	22
3.3.1. Περιγραφή Περιοχής Εμπιστοσύνης	23
3.4. Αλγόριθμοι που βασίζονται στο σημείο Cauchy	26
3.4.1. Σημείο Dogleg	29
3.4.2. Υποπροβήματα Ελαχιστοποίησης Δύο Διαστάσεων	31
3.5. Μέθοδος Newton	31
3.5.1. Βασική Μορφή Αλγορίθμου	31
3.5.2. Σύγκλιση Μεθόδου	33
3.5.3. Τοπική σύγκλιση μεθόδων Newton Περιοχή Εμπιστοσύνης	34

ΚΕΦΑΛΑΙΟ 4. Η ΕΠΕΚΤΑΣΗ ΤΟΥ ΑΛΓΟΡΙΘΜΟΥ DOGLEG 35
4.1. Ανάλυση και παρουσίαση 35
4.1.1. Γενικό Πρόβλημα 35
4.1.2. Δομή της Περιοχής-Εμπιστοσύνης 37
4.1.3. Dogleg κατά Powell 38
4.2. Ακριβές Dogleg στον δισδιάστατο χώρο 39
4.3. Διαχείριση Εσσιανού 41
4.4. Επέκταση λύσης στον τρισδιάστατο χώρο 45
ΚΕΦΑΛΑΙΟ 5. ΑΠΟΤΕΛΕΣΜΑΤΑ 48
5.1. Συναρτήσεις Πειραμάτων 48
5.2. Αριθμητικά Αποτελέσματα 50
5.3. Μελλοντική Έρευνα 53
ΑΝΑΦΟΡΕΣ 55
ΣΥΝΤΟΜΟ ΒΙΟΓΡΑΦΙΚΟ 57

ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ

Πίνακας	Σελ
Πίνακας 3.1 Περιπτώσεις Ύπαρξη Ελαχίστου	17
Πίνακας 5.1 Συναρτήσεις n=5	51
Πίνακας 5.2 Συναρτήσεις n=50	52
Πίνακας 5.3 Συναρτήσεις n=100	52
Πίνακας 5.4 Συναρτήσεις διαφόρων διαστάσεων	49

ΕΥΡΕΤΗΡΙΟ ΣΧΗΜΑΤΩΝ

Σχήμα	Σελ
Σχήμα 2.1	5
Σχήμα 2.2	9
Σχήμα 2.3	15
Σχήμα 3.1	23
Σχήμα 3.2	28
Σχήμα 3.3	30
Σχήμα 4.1	39
Σχήμα 4.2	40
Σχήμα 4.3	41
Σχήμα 4.4	43
Σχήμα 4.5	46

ΕΠΕΞΗΓΗΣΕΙΣ ΣΥΜΒΟΛΙΣΜΩΝ

X	διάνυσμα
A^{-1}	αντίστροφος πίνακας του A
A^{T}	ανάστροφος πίνακας του A
$\mathbb{R}^{n,n}$	nXn-διάστατος πραγματικός χώρος
x^{*}	τοπικός ελαχιστοποιητής
$f(x)$	αντικειμενική συνάρτηση βελτιστοποίησης
$g(x)=\boldsymbol{\nabla} f(x)$	διάνυσμα πρώτων παραγώγων (κλίση)
$G(x)= \boldsymbol{\nabla}^{2} f(x)$	πίνακας δευτέρων παραγώγων
$f_k, g_k, \ldots$	$f(x_k), g(x_k), \ldots$

ΠΕΡΙΛΗΨΗ

Αρσινόη Ειρήνη Ρίζου του Σταύρου και της Ευαγγελίας, Μ.Δ.Ε. Τμήμα Μηχανικών Η/Υ & Πληροφορικής, Πανεπιστήμιο Ιωαννίνων, Δεκέμβριος, 2018.
Επέκταση του αλγορίθμου «Dog-Leg» με βελτιστοποίηση εντός υποχώρων.
Επιβλέπων Μεταπτυχιακής Διατριβής: Ισαάκ Η. Λαγαρής, Καθηγητής.

Ο στόχος της παρούσας διατριβής, είναι να υλοποιηθεί μία βελτιωμένη εκδοχή του αλγορίθμου Powell's Dogleg υπό δύο πτυχές. Αρχικά αυτή η εκδοχή λύνει το προκύπτον, δύο διαστάσεων πρόβλημα βελτιστοποίησης, ακριβώς και χειρίζεται θετικά και αρνητικά ορισμένους αλλά και απροσδιόριστους Εσσιανούς πίνακες στο ίδιο σημείο. Το σχήμα είναι αποτελεσματικό, υπολογιστικά αποδοτικό και απλό στην εφαρμογή.
Ανάμεσα στις πολλές μεθοδολογίες βελτιστοποίησης, οι μέθοδοι Newton και Quasi-Newton είναι εκείνες που κατέχουν τα σκήπτρα της επιτυχίας για λείες συναρτήσεις με συνεχείς πρώτες και δεύτερες παραγώγους. Οι μέθοδοι Quasi-Newton δεν υπολογίζουν επακριβώς τον Εσσιανό Πίνακα αλλά χρησιμοποιούν ένα σχήμα ενημέρωσης για να χτίσουν μία προσέγγιση για αυτόν και χρησιμοποιούν μόνο τις πρώτες παραγώγους. Από την άλλη μεριά οι μέθοδοι Newton απαιτούν τον υπολογισμό της πρώτης και της δεύτερης παραγώγου της συνάρτησης. Υπάρχουν δύο προσεγγίσεις η «Ευθύγραμμη Αναζήτηση» και η «Περιοχή Εμπιστοσύνης». .
Η παρούσα εργασία εστιάζει στις μεθόδους Newton υπό την δομή της τεχνικής Περιοχή Εμπιστοσύνης. Παρουσιάζονται ενδελεχώς και αναλυτικά οι παραπάνω μέθοδοι καθώς και τα μαθηματικά προαπαιτούμενα για την κατανόηση του περιεχομένου του τομέα της Βελτιστοποίησης.

EXTENDED ABSTRACT IN ENGLISH

Rizou Arsinoi Eirini, M. Sc. Department of Computer Science and Engineering, University of Ioannina, Greece, December 2018.
Subspace optimization extending the Dog-Leg Algorithm.
M. Sc. Thesis Supervisor: Isaac E. Lagaris, Professor.

The field of « Nonlinear Optimization » has undergone a tremendous development in the last decades with the advent powerful computers. Among the several optimization methodologies, Newton and Quasi-Newton are regarded as the most successful for smooth objective functions with continuous first and second derivatives.
Quasi Newton (QN) methods do not need to explicitly calculate the Hessian matrix but use an updating scheme to build the estimation for it over several iterations and need only to calculate the gradient of the objective function.
Newton methods on the other hand, require the calculation of both the gradient and the second derivative matrix (Hessian). There are two main approaches currently followed by Newton and QN methods, namely the “Line-search” and the “Trust-region” (also known as Restricted-step). The focus of the present thesis is on “Trust-region” Newton Methods. In particular, the approximate 2-d subspace minimization effected by Powell’s “Dog-Leg”, is replaced by an exact technique that can also treat higher dimensional subspaces.
This improves Powell’s Dogleg in two aspects: a) Solves the arising two dimensional optimization subproblem exactly, and b) handles positive-definite, indefinite or negative-definite Hessian matrices on the same footing. The scheme is effective, computationally efficient, and rather simple to implement

ΚΕΦΑΛΑΙΟ 1. ΕΙΣΑΓΩΓΗ

1.1 Στόχοι
1.2 Δομή της Διατριβής

1.1 Στόχοι

Ο συνολικός στόχος της διατριβής είναι να παρουσιαστεί μία βελτιωμένη έκδοση της τεχνικής Powell's Dogleg. Στην ουσία πρόκειται για υλοποίηση δύο λύσεων, η μία στον δισδιάστατο χώρο και η άλλη στον τρισδιάστατο με τα ακόλουθα χαρακτηριστικά: το σχήμα να είναι αποτελεσματικό, υπολογιστικά αποδοτικό και απλό στην εφαρμογή.

Οι επί μέρους στόχοι :

- Να αναλυθούν ενδελεχώς, από θεωρητικής και όχι μόνο απόψεως, γνώσεις και μαθηματικό υπόβαθρο στον τομέα της Βελτιστοποίησης .

- Η ανάπτυξη (υλοποίηση αντίστοιχου κώδικα σε γλώσσα προγραμματισμού) των δύο σχετικών αλγορίθμων για να εξάγουμε τα πειραματικά-συγκριτικά αποτελέσματα και αποφανθούμε.

- Να συγκριθούν οι μέθοδοι τόσο ποιοτικά όσο και χρονικά ώστε να έχουμε τους απαραίτητους συγκριτικούς πίνακες για να αποφανθούμε για την απόδοσή τους.

1.2 Δομή της Διατριβής

Η διατριβή εκτός από το εισαγωγικό Κεφάλαιο 1, περιέχει τα τρία επόμενα κεφάλαια: Το Κεφάλαιο 2, αποτελείται από πέντε μέρη. Το πρώτο μέρος είναι ο «Μαθηματικός Συμβολισμός» που περιέχει την γενική μαθηματική διατύπωση ενός προβλήματος βελτιστοποίησης. Στην συνέχεια ακολουθούν οι «Κατηγορίες Βελτιστοποίησης» με βάση τα χαρακτηριστικά. Ακολουθεί ο ορισμός της λύσης ενός προβλήματος βελτιστοποίησης σε κάθε περίπτωση. Εν συνεχεία στο «Αναγνωρίζοντας ένα τοπικό ελάχιστο» παρουσιάζονται μαθηματικά υπόβαθρα για την κατανόηση εύρεσης ενός τοπικού ελαχίστου. Τέλος ακολουθεί η παράγραφος «Αλγόριθμοι Βελτιστοποίησης» με βασική αναφορά στις δύο πολύ γνωστές τεχνικές που εφαρμόζονται για εύρεση βέλτιστου σημείου.

Το Κεφάλαιο 3 περιέχει πέντε μέρη, ξεκινώντας με την αναφορά στις τετραγωνικές συναρτήσεις και τα τετραγωνικά μοντέλα, όπου αναφέρεται η μορφή τους και η σημαντικότητα της χρήσης τους. Μετέπειτα, αναφερόμαστε σε μεθόδους αριθμητικής γραμμικής άλγεβρας και εύρεσης ρίζας που υλοποιήθηκαν στην μεταπτυχιακή εργασία όπως η παραγοντοποίηση Cholesky και ο αλγόριθμός της μεθόδου Regula-Falsi. Στο τέταρτο μέρος αυτού του κεφαλαίου αναφερόμαστε ενδελεχώς στην τεχνική της Περιοχής Εμπιστοσύνης αναλύοντας το σημείο Cauchy και το σημείο dogleg. Τέλος, αναλύουμε τις μεθόδους Newton παρουσιάζοντας τον βασικό αλγόριθμο και επισημαίνοντας θέματα σύγκλισης.

Το Κεφάλαιο 4, περιέχει, με την σειρά του έξι τμήματα. Στο πρώτο εμπεριέχεται η περιγραφή της σκέψης που αποτέλεσε το εναρκτήριο λάκτισμα της υλοποίησης της βελτιωμένης εκδοχής της τεχνικής dogleg. Στην ουσία παρουσιάζεται η επέκταση με δύο τεχνικές και τις λεπτομέρειες τους. Η μία αποτελεί βελτιστοποίηση της τεχνικής του Powell’s Dogleg στις δύο διαστάσεις και η επόμενη στις τρεις διαστάσεις. Παρουσιάζονται τα βήματα των αλγορίθμων καθώς και λεπτομέρειες τους. Αναφέρεται ο λόγος που ο αλγόριθμος μας βρίσκει προσεγγιστικά την ακριβή λύση έναντι της γνωστής προσεγγιστικής λύση..

Το Κεφάλαιο 5, είναι το τελευταίο της μεταπτυχιακής διατριβής κι αποτελείται από τρία επιμέρους τμήματα που έχουν σχέση με τα πειράματα και τα αποτελέσματα από

την υλοποίηση του κώδικα. Παρουσιάζονται οι συναρτήσεις, που μας απασχόλησαν, με την γραφική τους παράσταση, τις παραμέτρους τους και το ολικό τους ελάχιστο το οποίο είναι ξεκάθαρο και από την γραφική τους παράσταση. Στη συνέχεια ακολουθεί ο πίνακας σύγκρισης των δύο τεχνικών-επεκτάσεων. Τέλος, υπάρχουν τα συμπεράσματα από την παρούσα εργασία καθώς οι προτάσεις για μελλοντική εργασία επί του θέματος.

ΚΕΦΑΛΑΙΟ 2. ΕΙΣΑΓΩΓΙΚΕΣ ΕΝΝΟΙΕΣ ΕΠΙ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ

2.1 Μαθηματικός Συμβολισμός
2.2 Κατηγορίες Βελτιστοποίησης
2.3 Τι ορίζεται ως λύση
2.4 Αναγνωρίζοντας ένα τοπικό ελάχιστο
2.5 Αλγόριθμοι Βελτιστοποίησης

2.1 Μαθηματικός Συμβολισμός

Μαθηματικά μιλώντας, η βελτιστοποίηση είναι η ελαχιστοποίηση ή η μεγιστοποίηση μίας συνάρτησης και ίσως υπό περιορισμούς αρκετές φορές. Χρησιμοποιείται η ακόλουθη σημειογραφία:

x είναι ένα διάνυσμα από μεταβλητές που καλείται επίσης παράμετροι

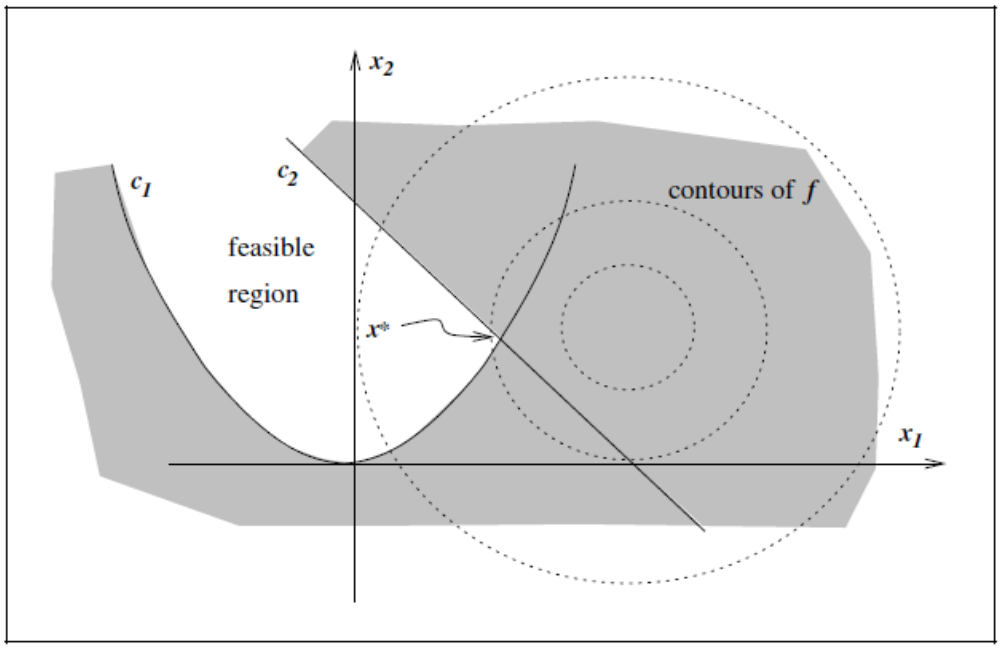
f είναι η αντικειμενική συνάρτηση, μία βαθμωτή συνάρτηση που θέλουμε να μεγιστοποιήσουμε ή να ελαχιστοποιήσουμε

c_i είναι οι περιορισμοί, ένα σύνολο από συναρτήσεις που ορίζουν εξισώσεις ή ανισώσεις που θα πρέπει να ικανοποιούνται από τον άγνωστο x

Χρησιμοποιώντας τον παραπάνω συμβολισμό ,το πρόβλημα της βελτιστοποίησης μπορεί να αναπαρασταθεί ως εξής:

$$\min_{x \epsilon IR^n} f(x) \ \ \upsilon\pi\acute{o}\ \tau o\nu\ \pi\varepsilon\rho\iota o\rho\iota\sigma\mu\acute{o} \quad \begin{matrix} c_i(x) = 0\ , i\ \epsilon\ \mathcal{E} \\ c_i(x) \geq 0\ , i\ \epsilon\ \mathfrak{I} \end{matrix} \qquad \text{Εξ.2.1}$$

Όπου $\mathcal{E}$, $\mathfrak{I}$ είναι σύνολα δεικτών για τις ισότητες και ανισότητες αντίστοιχα.



Σχήμα 2.1

Στο Σχήμα 2.1 παρουσιάζονται οι ισοϋψείς καμπύλες, που είναι ένα σύνολο σημείων που η συνάρτηση έχει σταθερή τιμή. Και τέλος το εφικτό σύνολο, δηλαδή το σύνολο των σημείων που ικανοποιούν τους περιορισμούς

2.2 Κατηγορίες Βελτιστοποίησης

Σε αυτό το σημείο αναφέρονται ενδελεχώς τα είδη της βελτιστοποίησης ανάλογα με το αν έχουν περιορισμούς ή δεν έχουν, είναι στον διακριτό ή στον συνεχή χώρο, αν ψάχνουν μια γενική λύση ή όχι και τέλος αν είναι στοχαστική ή ντετερμινιστική βελτιστοποίηση.

2.2.1 Υπό Περιορισμούς και Δίχως Περιορισμούς Βελτιστοποίηση

Προβλήματα με γενική μορφή Εξίσωση (2.1) μπορούν να ταξινομηθούν ανάλογα με τη φύση της αντικειμενικής συνάρτησης και τον περιορισμό (γραμμικός, μη γραμμικός, κυρτός), τον αριθμό των μεταβλητών (μεγάλος ή μικρός), την ομαλότητα των συναρτήσεων (διαφορίσιμη ή μη διαφορίσιμη) και ούτω καθεξής. Μια σημαντική

διάκριση είναι μεταξύ των προβλημάτων είναι σε αυτά που έχουν περιορισμούς στις μεταβλητές και εκείνα που δεν έχουν.

Προβλήματα άνευ περιορισμών, για τα οποία έχουμε $\mathcal{E} = \mathfrak{I} = \emptyset$ προκύπτουν άμεσα σε πολλές εφαρμογές. Ακόμη και για ορισμένα προβλήματα με φυσικούς περιορισμούς στις μεταβλητές, μπορεί να είναι προτιμότερο να αγνοηθούν, καθώς δεν επηρεάζουν τη λύση και δεν παρεμβάλλονται στους αλγόριθμους επίλυσης του προβλήματος. Άνευ περιορισμού προβλήματα προκύπτουν επίσης, ως αναδιατυπώσεις προβλημάτων βελτιστοποίησης με περιορισμούς, στα οποία οι περιορισμοί αντικαθίστανται από την επιβολή «κύρωσης-τιμωρίας» που προστίθενται στην αντικειμενική συνάρτηση που έχει ως αποτέλεσμα την αποθάρρυνση παραβιάσεων περιορισμού.

Προβλήματα βελτιστοποίησης με περιορισμούς προκύπτουν από μοντέλα στα οποία οι περιορισμοί διαδραματίζουν ουσιαστικό ρόλο, για παράδειγμα στην επιβολή δημοσιονομικών περιορισμών σε ένα οικονομικό πρόβλημα ή τον περιορισμό σχήματος σε ένα πρόβλημα σχεδίασης. Αυτοί οι περιορισμοί μπορεί να είναι απλοί όπως $0 \leq x_i \leq 1$, πιο γενικοί $\sum_{i=1}^{n} x_i \geq 0$ ή περισσότερο πολύπλοκοι ώστε να ανταποκρίνονται στην πολυπλοκότητα του προβλήματος.

2.2.2 Διακριτή και Συνεχής Βελτιστοποίηση

Σε κάποια προβλήματα βελτιστοποίησης είναι αναγκαίο οι μεταβλητές να λαμβάνουν τιμές ακέραιες. Ο μαθηματικός συμβολισμός τέτοιων προβλημάτων περιλαμβάνουν περιορισμούς οι οποίοι έχουν την μορφή $x_i \in Z$, όπου Z είναι το σύνολο των ακεραίων, ή δυαδικοί περιορισμούς οι οποίοι έχουν την μορφή $x_i \in \{0,1\}$ επιπρόσθετες στους αλγεβρικούς περιορισμούς της Εξίσωση (2.1). Προβλήματα σαν κι αυτό ονομάζονται ακέραια προγραμματιστικά προβλήματα. Αυτά είναι και μία συγκεκριμένη κατηγορία βελτιστοποίησης, διακριτή. Το καθοριστικό χαρακτηριστικό ενός διακριτού προβλήματος βελτιστοποίησης είναι ότι το άγνωστο x προέρχεται από ένα πεπερασμένο σύνολο.

Αντίθετα, το εφικτό σύνολο για προβλήματα συνεχούς βελτιστοποίησης είναι φανερά άπειρο, όπως όταν οι συνιστώσες του x επιτρέπεται να είναι πραγματικοί αριθμοί. Τα προβλήματα συνεχούς βελτιστοποίησης, είναι συνήθως πιο εύκολα στην επίλυση επειδή η ομαλότητα των συναρτήσεων καθιστά δυνατή τη χρήση αντικειμενικών συναρτήσεων και πληροφορίες περιορισμού σε ένα συγκεκριμένο σημείο x για την εξαγωγή πληροφοριών σχετικά με την συμπεριφορά των συναρτήσεων σε όλα τα σημεία κοντά στο x . Σε διακριτά προβλήματα, υπό περιορισμό, η συμπεριφορά της συνάρτησης και οι περιορισμοί μπορεί να αλλάξουν σημαντικά καθώς μεταβαίνουμε από ένα εφικτό σημείο σε ένα άλλο, ακόμη και αν τα δύο σημεία είναι "κοντά" με κάποιο μέτρο. Τα εφικτά σύνολα, για διακριτά προβλήματα βελτιστοποίησης, μπορούν να θεωρηθούν ότι παρουσιάζουν μια ακραία μορφή μη κυρτότητας, όπως επίσης και ένας κυρτός συνδυασμός δύο εφικτών σημείων δεν είναι γενικά εφικτός.

Πρέπει, τέλος, να σημειωθεί πως αρκετές φορές παρατηρούμε, ωστόσο, ότι οι τεχνικές συνεχούς βελτιστοποίησης συχνά διαδραματίζουν σημαντικό ρόλο στην επίλυση διακριτών προβλημάτων βελτιστοποίησης. προβλήματα βελτιστοποίησης, μπορούν να θεωρηθούν ότι παρουσιάζουν μια ακραία μορφή μη κυρτότητας, όπως επίσης και ένας κυρτός συνδυασμός δύο εφικτών σημείων δεν είναι γενικά εφικτός.

2.2.3 Καθολική και Τοπική Βελτιστοποίηση

Πολλοί αλγόριθμοι για προβλήματα μη γραμμικής βελτιστοποίησης ψάχνουν μόνο μία τοπική λύση, ένα σημείο όπου η αντικειμενική συνάρτηση είναι μικρότερη από όλα τα άλλα εφικτά γειτονικά σημεία. Δεν βρίσκουν πάντα ολική λύση, η οποία είναι το σημείο με την μικρότερη τιμή που εμφανίζει η συνάρτηση γενικότερα, ανάμεσα σε όλα τα σημεία. Οι ολικές λύσεις χρειάζονται σε κάποιες εφαρμογές αλλά για πολλά προβλήματα είναι δύσκολο να αναγνωριστούν και πόσο μάλλον να βρεθούν αυτές οι τιμές. Για τα κυρτά προβλήματα προγραμματισμού και κυρίως για τα γραμμικά, οι τοπικές λύσεις είναι επίσης ολικές. Γενικά, τα μη γραμμικά προβλήματα και αυτά που έχουν περιορισμούς και αυτά δίχως περιορισμούς, ίσως παρουσιάζουν τοπικές λύσεις που δεν είναι καθολικές. Θα πρέπει να τονιστεί πως αρκετοί αλγόριθμοι καθολικής βελτιστοποίησης χρησιμοποιούν τα όλες τις τοπικές λύσεις για να αποφανθούν.

2.2.4 Αιτιοκρατική και Στοχαστική Βελτιστοποίηση

Σε ορισμένα προβλήματα βελτιστοποίησης, το μοντέλο δεν μπορεί να καθοριστεί πλήρως, επειδή εξαρτάται από ποσότητες που είναι άγνωστες κατά τη στιγμή της διαμόρφωσης. Αυτό το χαρακτηριστικό μοιράζονται πολλά μοντέλα οικονομικού και χρηματοοικονομικού σχεδιασμού, τα οποία μπορεί να εξαρτώνται π.χ. από τα μελλοντικά επιτόκια, ή τις μελλοντικές τιμές των βασικών εμπορευμάτων, αλλά η αβεβαιότητα μπορεί να προκύψει φυσικά σε σχεδόν οποιοδήποτε τύπο εφαρμογής. Αντί να χρησιμοποιηθεί μόνο μια "καλύτερη εικασία" για τις αβέβαιες ποσότητες, οι «μοντελιστές» μπορούν να αποκτήσουν πιο χρήσιμες λύσεις ενσωματώνοντας πρόσθετες γνώσεις, σχετικές με αυτές, στο μοντέλο. Για παράδειγμα, μπορεί να είναι γνωστά ορισμένα πιθανά σενάρια για την αβέβαιη ζήτηση, μαζί με εκτιμήσεις των πιθανοτήτων κάθε σεναρίου. Τα προβλήματα στοχαστικής βελτιστοποίησης χρησιμοποιούν αυτές τις ποσοτικοποιήσεις της αβεβαιότητας για να παράγουν λύσεις που βελτιστοποιούν την αναμενόμενη απόδοση του μοντέλου.

Αντίθετα, ως προβλήματα αιτιοκρατικής βελτιστοποίησης αναφέρονται τα προβλήματα στα οποία το μοντέλο είναι ολοκληρωτικά άγνωστο. Ωστόσο, πολλοί αλγόριθμοι για στοχαστική βελτιστοποίηση προχωρούν με τη διατύπωση ενός ή περισσότερων αλγορίθμων για αιτιοκρατικά υποπροβλήματα .

2.3 Τι ορίζεται ως λύση

Γενικά, αυτό που θέλουμε να πετύχουμε είναι έναν ολικό ελαχιστοποιητή της συνάρτησης f , ένα σημείο όπου η συνάρτηση επιτυγχάνει την ελάχιστη τιμή της.

Μαθηματικά ορίζεται ως εξής:

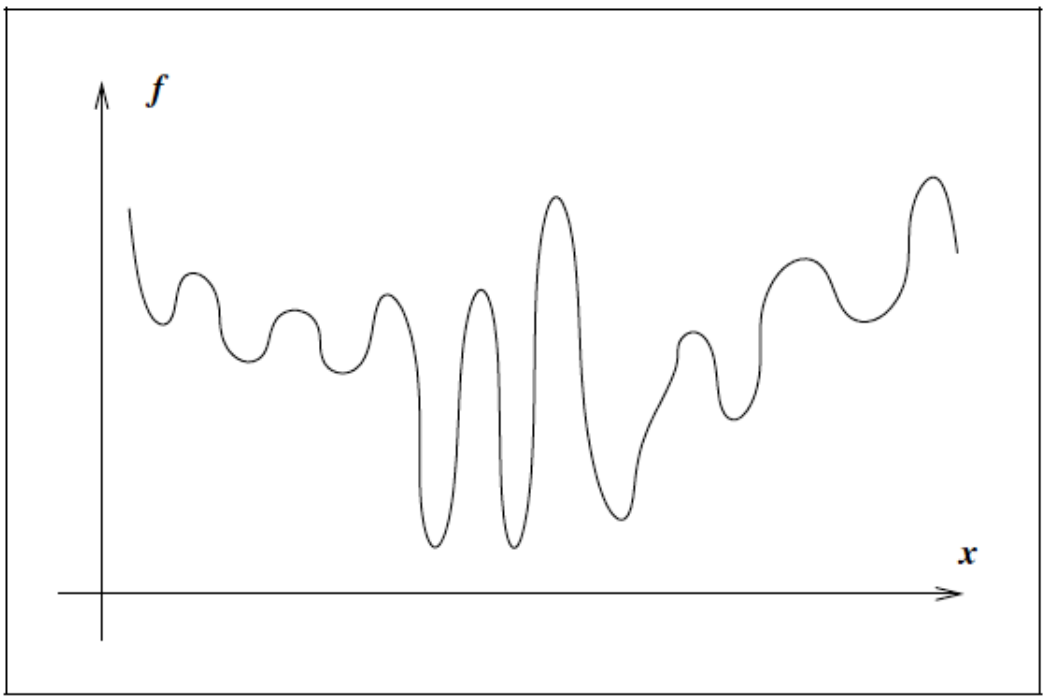
Ένα σημείο x* είναι **ολικός ελαχιστοποιητής** αν $f(x^*) \leq f(x), \forall x \in \mathbb{R}^n$. Το να βρούμε έναν ελαχιστοποιητή ίσως μας δυσκολέψει γιατί συνήθως η γνώση μας είναι συχνά τοπικού χαρακτήρα. Οι περισσότεροι αλγόριθμοι είναι ικανοί να βρίσκουν μόνο ένα τοπικό ελαχιστοποιητή ,το οποίο είναι μια μικρή τιμή σε μία γειτονιά του. Μπορούμε να το διατυπώσουμε ως εξής: Ένα σημείο x^* είναι ένας **τοπικός ελαχιστοποιητής** αν υπάρχει γειτονιά N του x^* τέτοια ώστε $f(x^*) \leq f(x)$ για όλα τα $x \in N$. Ένα σημείο που ικανοποιεί αυτό τον ορισμό μερικές φορές

καλείται ένας αδύναμος τοπικός ελαχιστοποιητής. Ένα σημείο καλείται **ακριβής τοπικός ελαχιστοποιητής** ,ο οποίος είναι ο «νικητής» στην γειτονιά του. Μπορούμε να το διατυπώσουμε ως εξής: Ένα σημείο x^* είναι ένας ακριβής τοπικός ελαχιστοποιητής αν υπάρχει γειτονιά N του x^*τέτοια ώστε $f(x^*) < f(x)$ για όλα τα $x \in N$ με $x \neq x^*$

Παράδειγμα

Για την συνεχή συνάρτηση $f(x) = 2$,κάθε σημείο x είναι ένας αδύναμος τοπικός ελαχιστοποιητής ενώ για την συνάρτηση $f(x) = (x-2)^4$ έχει έναν ακριβή τοπικό ελαχιστοποιητή στο σημείο $x = 2$.

Ένα σημείο καλείται **απομονωμένος τοπικός ελαχιστοποιητής** αν υπάρχει γειτονιά N του x^*τέτοια ώστε το x^* είναι ο μοναδικός τοπικός ελαχιστοποιητής στο N. Μερικοί ακριβείς τοπικοί ελαχιστοποιητές δεν είναι απομονωμένοι , όπως στην συνάρτηση ακολούθως $f(x)$ =x^4cos(1/x) +2x^4 , $f(0) = 0$. Η συνάρτηση είναι δύο φορές παραγωγίσιμη και έχει έναν ακριβή τοπικό ελαχιστοποιητή στο σημείο x^* =0.



Σχήμα 2.2

Ενώ, οι ακριβείς τοπικοί ελαχιστοποιητές δεν είναι πάντα απομονωμένοι ,όλοι οι απομονωμένοι τοπικοί ελαχιστοποιητές είναι ακριβείς. Το Σχήμα 2.2 δείχνει μία συνάρτηση με πολλούς τοπικούς ελαχιστοποιητές. Είναι δύσκολο να βρούμε έναν ολικό ελαχιστοποιητή για τέτοιες συναρτήσεις, επειδή οι αλγόριθμοι τείνουν να παγιδεύονται στους τοπικούς ελαχιστοποιητές. Σε αρκετές περιπτώσεις βλέπουμε συναρτήσεις με εκατομμύρια τοπικά ελάχιστα. Μία σημαντική ειδική περίπτωση είναι, αυτή των κυρτών συναρτήσεων, για τις οποίες κάθε τοπικός ελαχιστοποιητής είναι επίσης ένας ολικός ελαχιστοποιητής.

2.4 Αναγνωρίζοντας ένα τοπικό ελάχιστο

Από τον ορισμό που δώσαμε παραπάνω, ίσως φαίνεται ότι ο μόνος τρόπος για να βρούμε αν ένα σημείο x^* είναι ένα τοπικό ελάχιστο είναι να εξετάσουμε όλα τα σημεία στην άμεση γειτονιά τους για να σιγουρευτούμε ότι κανένα από αυτά δεν έχει μικρότερη τιμή συνάρτησης. Όταν μια συνάρτηση f είναι λεία, υπάρχουν πιο αποδοτικοί και πρακτικοί τρόποι για να προσδιορίσουμε ένα τοπικό ελάχιστο. Αν η f είναι δύο φορές συνεχώς παραγωγίσιμη, είμαστε ικανοί να πούμε ότι το x^* είναι ένας τοπικός ελαχιστοποιητής ελέγχοντας το $\nabla f(x^*)$ και τον Εσσιανό. Το μαθηματικό εργαλείο που χρησιμοποιείται για να μελετήσουμε λείες συναρτήσεις είναι το Θεώρημα Taylor .

Θεώρημα Taylor

Έστω συνάρτηση $f: \mathbb{R} \to \mathbb{R}$ είναι συνεχώς παραγωγίσιμη και $p \in R^n$. Τότε έχουμε:

$f(x+p) = f(x) + \nabla f(x+tp)^T p$, για $t \in (0,1)$.

Επιπλέον, αν η συνάρτηση f είναι δύο φορές συνεχώς παραγωγίσιμη έχουμε ότι:

$$f(x+p) = f(x) + \nabla f(x)^T p + \frac{1}{2} pT\nabla^2 f(x+tp)^T p \text{ , για } t \in (0,1) .$$

Απαραίτητες προϋποθέσεις για την εύρεση είναι τα δύο ακόλουθα θεωρήματα.

Θεώρημα 1ης παραγώγου

Αν x^* είναι τοπικός ελαχιστοποιητής και η f είναι συνεχώς παραγωγίσιμη σε μία ανοιχτή γειτονιά του x^* τότε $\nabla f(x^*) = 0$.

Θεώρημα 2ης παραγώγου

Αν x^*είναι τοπικός ελαχιστοποιητής και υπάρχει $\nabla^2 f$ συνεχής σε μία ανοιχτή γειτονιά του x^* τότε $\nabla f(x^*)$ και $\nabla^2 f(x^*)$ είναι θετικά ημιορισμένος πίνακας.

Στην συνέχεια βλέπουμε μια ικανή προϋπόθεση.

Θεώρημα

Έστω ότι $\nabla^2 f$ είναι συνεχής σε μία ανοιχτή γειτονιά του x^*και ακόμη $\nabla f(x^*)$ 0 και ∇^2 είναι θετικά ορισμένος. Τότε x^* είναι ένας ακριβής τοπικός ελαχιστοποιητής της f.

Αυτό το θεώρημα εγγυάται κάτι πιο «δυνατό» που δεν συμβαίνει με αντίστοιχα με τις απαραίτητες προϋποθέσεις που αναφέρονται ως θεωρήματα παραπάνω, κυρίως ότι ο ελαχιστοποιητής είναι ένας ακριβής τοπικός ελαχιστοποιητής. Στο επόμενο παράδειγμα δίνεται μία συνάρτηση $f(x)=x^4$ για την οποία το σημείο $x^*=0$ είναι ένας ακριβής τοπικός ελαχιστοποιητής στο οποίο ο Εσσιανός πίνακας εκλείπει και δεν είναι θετικά ορισμένος.

2.5 Αλγόριθμοι Βελτιστοποίησης

Οι αλγόριθμοι βελτιστοποίησης είναι επαναληπτικοί. Αρχίζουν με μια αρχική εικασία της μεταβλητής x και παράγουν μια ακολουθία βελτιωμένων εκτιμήσεων (που ονομάζονται "iterates") έως ότου τερματιστούν, ευελπιστώντας να βρεθεί λύση. Η στρατηγική που χρησιμοποιείται για τη μετάβαση από τη μια επανάληψη στην άλλη διακρίνει έναν αλγόριθμο από τον άλλο. Οι περισσότερες στρατηγικές κάνουν χρήση των τιμών της αντικειμενικής συνάρτησης f, οι συναρτήσεις περιορισμού c_i και πιθανώς την πρώτη και το δεύτερη παράγωγο αυτών των συναρτήσεων. Μερικοί αλγόριθμοι συσσωρεύουν πληροφορίες που συγκεντρώθηκαν σε προηγούμενες επαναλήψεις, ενώ άλλοι χρησιμοποιούν μόνο τοπικές πληροφορίες που αποκτήθηκαν

στο τρέχον σημείο. Για να είναι ένας αλγόριθμος καλός πρέπει να διαθέτει τις ακόλουθες ιδιότητες:

- Ευστάθεια. Θα πρέπει να αποδίδει καλά σε μια ποικιλία προβλημάτων, για τις διάφορες τιμές του σημείου εκκίνησης.
- Αποδοτικότητα. Δεν πρέπει να απαιτεί υπερβολικό χρόνο από τον υπολογιστή ή για αποθήκευση.
- Ακρίβεια. Θα πρέπει να είναι σε θέση να προσδιορίσει μια λύση με ακρίβεια, χωρίς να είναι υπερβολικά ευαίσθητος σε σφάλματα των δεδομένων ή σε αριθμητικά σφάλματα στρογγυλοποίησης που εμφανίζονται όταν ο αλγόριθμος υλοποιείται σε έναν υπολογιστή.

Αυτοί οι στόχοι μπορεί να είναι αλληλοσυγκρουόμενοι. Για παράδειγμα, μια ταχέως συγκλίνουσα μέθοδος για ένα μεγάλο χωρίς περιορισμούς μη γραμμικό πρόβλημα ενδέχεται να απαιτεί υπερβολική αποθήκευση στον υπολογιστή. Αφ' ετέρου, μια ισχυρή μέθοδος μπορεί επίσης να είναι η πιο αργή. Εναλλαγές μεταξύ ρυθμού σύγκλισης και απαιτήσεις αποθήκευσης , καθώς και μεταξύ της ευρωστίας και της ταχύτητας και ούτω καθεξής, είναι κεντρικά ζητήματα για την αριθμητική βελτιστοποίηση.

Υπάρχουν δύο θεμελιώδεις στρατηγικές για τη μετάβαση από το τρέχον σημείο x_k σε ένα νέο σημείο $x_k + 1$. Ακολουθεί βασική περιγραφή τους.

2.5.1 Ευθύγραμμη Αναζήτηση

Στην τεχνική της ευθύγραμμης αναζήτησης, ο αλγόριθμος επιλέγει μία διεύθυνση p_kκαι ψάχνει κατά μήκος αυτής της διεύθυνσης από το τρέχον x_k προς ένα νέο με μικρότερη τιμή συνάρτησης. Η απόσταση για να κινηθούμε κατά μήκος του p_k μπορεί να υπολογιστεί προσεγγιστικά λύνοντας ένα μονοδιάστατο πρόβλημα βελτιστοποίησης ώστε να υπολογίσουμε το α

$$\min_{a>0} f(x_{k+} a p_k) \qquad \text{Εξ. 2.2}$$

Λύνοντας το παραπάνω πρόβλημα μπορούμε να εξάγουμε το μεγαλύτερο όφελος από την κατεύθυνση του p_k, αλλά μία ακριβής ελαχιστοποίηση ίσως είναι κοστοβόρα και αρκετές φορές αχρείαστη. Εν αντιθέσει ο αλγόριθμος ευθύγραμμης αναζήτησης παράγει έναν περιορισμένο αριθμό από δοκιμαστικά μήκη βήματος μέχρι να βρει εκείνο που προσεγγίζει το ελάχιστο της Εξίσωση (2.2). Στο νέο σημείο, θα υπολογιστούν μία νέα κατεύθυνση αναζήτησης και σαφώς ένα νέο βήμα και η διαδικασία θα συνεχίζεται.

2.5.2 Περιοχή Εμπιστοσύνης

Στην δεύτερη αλγοριθμική τεχνική, γνωστή και ως ***περιοχή εμπιστοσύνης***, η πληροφορία που συλλέγεται για την f χρησιμοποιείται για την κατασκευή ενός μοντέλου συνάρτησης m_k του οποίου η συμπεριφορά κοντά στο σημείο x_k είναι όμοια με αυτή που έχει η πραγματική αντικειμενική συνάρτηση f. Επειδή το μοντέλο m_k ίσως να μην είναι καλή προσέγγιση της συνάρτησης f όταν το x είναι αρκετά μακριά από το x_k , περιορίζουμε την αναζήτηση του ελαχιστοποιητή του m_k σε μια περιοχή γύρω από το x_k. Με άλλα λόγια βρίσκουμε το υποψήφιο βήμα p λύνοντας προσεγγιστικά το παρακάτω υποπρόβλημα:

$$\min_p m_k(x_k + p)\ ,$$

όταν το $x_k + p$ βρίσκεται εντός της περιοχής εμπιστοσύνης. Εξ. 2.3

Το μοντέλο m_k ορίζεται ως τετραγωνικό έχοντας την ακόλουθη μορφή:

$$m_k(x_k + p) = f_k + p^T \nabla f_k + \frac{1}{2} p^T B_k p \qquad \text{Εξ. 2.4}$$

Όπου f_k , ∇f_k και B_k είναι βαθμωτό μέγεθος διάνυσμα και πίνακας αντίστοιχα.
f_k είναι η τιμή της συνάρτησης στο τρέχον x_k
∇f_k είναι η τιμή της παραγώγου στο τρέχον x_k
B_k είναι ο Εσσιανός πίνακας $\nabla^2 f_k$ ή μια προσέγγισή του στο τρέχον x_k

Αν η υποψήφια λύση δεν παράγει μία ικανοποιητική μείωση στις τιμές της f, καταλήγουμε ότι η περιοχή δεν είναι τόσο μεγάλη. Τότε μικραίνουμε την περιοχή εμπιστοσύνης και ξαναλύνουμε το πρόβλημα Εξίσωση (2.3). Συνήθως, η περιοχή

εμπιστοσύνης είναι μία σφαίρα όπου $\|p\|_2 \leq \Delta$, όπου $\Delta > 0$ είναι η ακτίνα της περιοχής εμπιστοσύνης. Το μέγεθος της περιοχής εμπιστοσύνης είναι κρίσιμο για την αποτελεσματικότητα κάθε βήματος. Εάν η περιοχή είναι πολύ μικρή, ο αλγόριθμος χάνει την ευκαιρία να κάνει ένα σημαντικό βήμα που θα τον μετακινήσει πολύ πιο κοντά στον ελαχιστοποιητή της αντικειμενικής συνάρτησης. Εάν είναι πολύ μεγάλο, ο ελαχιστοποιητής του μοντέλου μπορεί να απέχει πολύ από τον ελαχιστοποιητή της αντικειμενικής συνάρτησης στην περιοχή, έτσι μπορεί να χρειάζεται να μειώσουμε το μέγεθος της περιοχής και να προσπαθήσουμε ξανά. Σε πρακτικούς αλγόριθμους, επιλέγουμε το μέγεθος της περιοχής σύμφωνα με την απόδοση του αλγορίθμου κατά τις προηγούμενες επαναλήψεις.

Επιπλέον χρησιμοποιούνται περιοχές εμπιστοσύνης σχημάτων έλλειψης και κουτιού.

Παράδειγμα

Αν η αντικειμενική συνάρτηση είναι η $f(x) = 10(x_2 - {x_1}^2)^2 + (1 - x_1)^2$.

Στο σημείο $x_k = (0,1)$ το gradient και ο Hessian είναι:

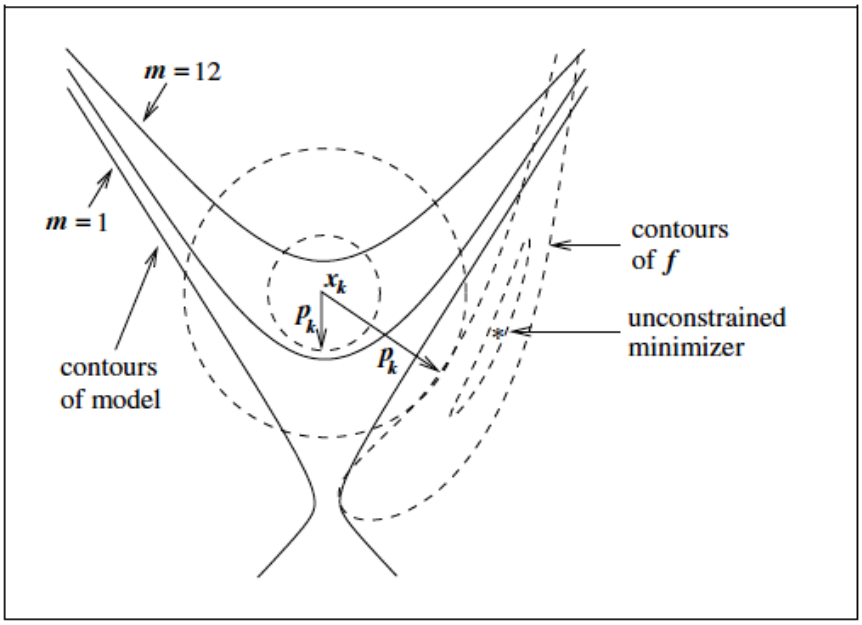
$$\nabla f_k = \begin{bmatrix} -2 \\ 20 \end{bmatrix} \quad \text{και} \quad \nabla^2 f_k = \begin{bmatrix} -38 & 0 \\ 0 & 20 \end{bmatrix}$$

Οι παραπάνω πίνακες εφαρμόζονται στο μοντέλο της Εξίσωση (2.4) .

Στην εικόνα Σχήμα 2.3 παρουσιάζονται οι ισοϋψείς καμπύλες του τετραγωνικού μοντέλου Εξίσωση (2.4) καθώς επίσης και της αντικειμενικής συνάρτησης αλλά και της περιοχής εμπιστοσύνης. Όπως μπορούμε να παρατηρήσουμε και στην εικόνα, κάθε φορά που μειώνουμε το μέγεθος της περιοχής εμπιστοσύνης μετά από «αποτυχία» μιας υποψήφιας επανάληψης, το βήμα από το x_k στο νέο υποψήφιο θα είναι μικρότερη και συνήθως δείχνει σε διαφορετική κατεύθυνση από το προηγούμενο.

Η στρατηγική για την περιοχή εμπιστοσύνης διαφέρει από αυτή την άποψη από την ευθύγραμμη αναζήτηση, η οποία παραμένει με μία μόνο κατεύθυνση αναζήτησης. Εν κατακλείδι, κατά μία έννοια, οι μέθοδοι της ευθύγραμμης αναζήτησης και της

περιοχής εμπιστοσύνης διαφέρουν κατά τη σειρά με την οποία επιλέγουν την κατεύθυνση και την απόσταση της μετάβασης στην επόμενη επανάληψη.



Σχήμα 2.3

Η βασική διαφορά των δύο παραπάνω στρατηγικών είναι ότι :στην πρώτη περίπτωση έχουμε καθορισμό κατεύθυνσης και μετά μήκους βήματος, ενώ στην δεύτερη ορίζουμε αρχικά μία μέγιστη τιμή για την ακτίνα και στην συνέχεια αναζητούμε βήμα και κατεύθυνση που επιτυγχάνουν την καλύτερη δυνατή βελτίωση. Διεξοδικά η τεχνική της περιοχής εμπιστοσύνης αναλύεται στο Κεφάλαιο 3 μιας και είναι η τεχνική που χρησιμοποιήθηκε στο προγραμματιστικό κομμάτι του κώδικα.

ΚΕΦΑΛΑΙΟ 3. ΘΕΜΕΛΙΩΔΗΣ ΑΝΑΛΥΣΗ ΠΡΟΑΠΑΙΤΟΥΜΕΝΩΝ

3.1 Πολυδιάστατες Τετραγωνικές Συναρτήσεις
3.2 Μέθοδοι Αριθμητικής Ανάλυσης και Άλγεβρας
3.3 Στρατηγική «Περιοχή Εμπιστοσύνης»
3.4 Ανάλυση Μεθόδου Dog-Leg
3.5 Μέθοδοι Newton

3.1 Πολυδιάστατες Τετραγωνικές Συναρτήσεις

Στην παρούσα διπλωματική εργασία γίνεται πολύς λόγος για τα τετραγωνικά μοντέλα ή τις τετραγωνικές συναρτήσεις. Εξ'αυτού δεν μπορούμε να παραλείψουμε τον ορισμό τους και κάποια βασικά τους στοιχεία.

3.1.1 Τετραγωνικές Συναρτήσεις

Η γενική μορφή μιας τετραγωνικής συνάρτησης αποτελείται από τρεις όρους:

$$q(x) = \frac{1}{2} x^T Q x + p^T x + c$$

Όπου c μία πραγματική σταθερά, x, p $\in \mathbb{R}^n$ και $Q \in \mathbb{R}^{(nXn)}$. Κάθε τετραγωνικός πίνακας μπορεί να γραφτεί ως: $Q = \frac{Q+Q^T}{2} + \frac{Q-Q^T}{2}$

όπου ο πρώτος όρος είναι ένας συμμετρικός πίνακας και ο δεύτερος ένας αντισυμμετρικός. Η συνεισφορά του αντισυμμετρικού πίνακα όρου στην q(x) είναι μηδενική διότι: $x^T(Q-Q^T)x = x^TQx - x^TQ^Tx = x^T(Qx) - (Qx)^Tx = 0$ και άρα στο εξής θα θεωρούμε ότι ο πίνακας Q είναι συμμετρικός. Η κλίση της τετραγωνικής συνάρτησης και ο Εσσιανός πίνακας δίνονται αντίστοιχα από τις παρακάτω εκφράσεις:

$$\nabla q(x) = Qx + p \qquad \text{Εξ. 3.1}$$

$$\nabla^2 q(x) = Q$$

Εάν ο Q είναι θετικά ορισμένος τότε το ακρότατο $x^* = -Q^{-1}p$ υπάρχει και είναι ελάχιστο. Από την Εξίσωση (3.1) συνεπάγεται ότι: $\nabla q(x_2) - \nabla q(x_1) = Q(x_2 - x_1)$

Η παραπάνω σχέση όπου ο πίνακας Q συνδέει την διαφορά των θέσεων με την διαφορά των κλίσεων, έχει παίξει σημαντικό ρόλο στην ανάπτυξη των μεθόδων «Μεταβλητή Μετρικής» ή γνωστές ως «Quasi-Newton».

Το ενδιαφέρον με τις τετραγωνικές συναρτήσεις είναι ότι μπορούμε να ορίσουμε τη συμπεριφορά τους και την ύπαρξη ή όχι ελάχιστων με τη μελέτη της ορισιμότητας του πίνακα Q, δηλαδή ανάλογα με το αν ο πίνακας είναι θετικά ή αρνητικά είναι (ημι)ορισμένος ή ακόμα αν είναι αόριστος μπορούμε να έχουμε συνολική εικόνα σχετικά με τη μορφή της τετραγωνικής συνάρτησης. Στον πίνακα, Πίνακας 3.1 παρουσιάζονται οι διαφορετικές περιπτώσεις για τετραγωνικές συναρτήσεις δύο διαστάσεων .

Πίνακας 3.1 Περιπτώσεις ύπαρξης ελαχίστων

x^TQx	Ορισιμότητα	Φύση x^*
> 0	Θετικά Ορισμένος	Ελάχιστο
≥ 0	Θετικά ημιορισμένος	Κοιλάδα
$\neq 0$	Αόριστος	Σαγματικό Σημείο
< 0	Αρνητικά Ορισμένος	Μέγιστο
≤ 0	Αρνητικά ημιορισμένος	Ράχη

3.1.2 Τετραγωνικά Μοντέλα

Οι περισσότεροι αλγόριθμοι ελαχιστοποίησης που θα παρουσιάσουμε είναι σχεδιασμένοι να λειτουργούν με τον καλύτερο τρόπο για τετραγωνικές συναρτήσεις. Η έννοια της τετραγωνικής συνάρτησης είναι πολύ σημαντική στην ανάπτυξη επιτυχημένων αλγορίθμων ελαχιστοποίησης για τους εξής λόγους:

1. Οι τετραγωνικές συναρτήσεις είναι από τις πιο απλές ομαλές συναρτήσεις, με καλά καθορισμένο ελάχιστο και συνεπώς εύκολο στο χειρισμό.
2. Σε μία γενική συνάρτηση, περιοχή γύρω από έναν ελαχιστοποιητή x μπορεί να προσεγγιστεί από μία κατάλληλα κατασκευασμένη τετραγωνική συνάρτηση. Για το λόγο αυτό οι αλγόριθμοι που βασίζονται σε τετραγωνικά μοντέλα έχουν γρήγορη σύγκλιση κοντά στον ελαχιστοποιητή.
3. Ακόμα και σε περιοχές μακριά από τον ελαχιστοποιητή είναι προτιμότερο να χρησιμοποιούμε τετραγωνικές συναρτήσεις γιατί η πληροφορία που εσωκλείουν αποδεικνύεται πολύ χρήσιμη και σαφώς καλύτερη από γραμμικές συναρτήσεις. Ο λόγος της προτίμησης αυτής είναι το γεγονός ότι το ανάπτυγμα Taylor μιας συνάρτησης f γύρω από ένα τυχαίο σημείο x_k αν αποτιμηθεί μέχρι και τον τετραγωνικό όρο

$$f(x_k + h) = f(x_k) + h^T \nabla f(x_k) + \frac{1}{2} h^T \nabla^2 f(x_k) h + O(\| h \|^3)$$

   παρέχει μεγαλύτερη ακρίβεια προσέγγισης σε μία γειτονιά του x_k από τη γραμμική προσέγγιση.
4. Μέθοδοι ελαχιστοποίησης που βασίζονται σε τετραγωνικά μοντέλα μπορούν να είναι αμετάβλητες σε γραμμικούς μετασχηματισμούς των μεταβλητών.

Εάν οι πρώτες και δεύτερες παράγωγοι είναι διαθέσιμες είναι προφανές ότι μπορούν να χρησιμοποιηθούν στον ορισμό ενός κατάλληλου τετραγωνικού μοντέλου. Σε περίπτωση που ο πίνακας δεύτερων παραγωγών δεν είναι διαθέσιμος μπορούμε να δημιουργήσουμε ένα τετραγωνικό μοντέλο προσεγγίζοντας τον πίνακα δεύτερον παραγώγων χρησιμοποιώντας πρώτες παραγωγούς ή/και τις τιμές συνάρτησης.

Μία άλλη περίπτωση που μπορούμε να χρησιμοποιήσουμε ένα τετραγωνικό μοντέλο είναι για να δείξουμε ότι ένας αλγόριθμος έχει τετραγωνική σύγκλιση. Σύμφωνα με

τον ορισμό ένας αλγόριθμος ελαχιστοποίησης έχει τετραγωνική σύγκλιση εάν καταφέρνει να προσεγγίσει το ελάχιστο x^* μιας τετραγωνικής συνάρτησης μετά από γνωστό και πεπερασμένο αριθμό βημάτων. Οι αλγόριθμοι αυτοί θα μπορούν να εφαρμοστούν επαναληπτικά και σε μη τετραγωνικές συναρτήσεις και να διατηρήσουν καλές ιδιότητες σύγκλισης.

3.2 Μέθοδοι Αριθμητικής Ανάλυσης και Άλγεβρας

Σε αυτές τις παραγράφους αναλύεται η μέθοδος επίλυσης ενός γραμμικού συστήματος με την μέθοδο «Παραγοντοποίηση Cholesky» καθώς και κάποιοι χρήσιμοι αλγόριθμοι μονοδιάστατης ελαχιστοποίησης οι οποίοι υλοποιήθηκαν στην παρούσα διατριβή. Συγκεκριμένα η μέθοδος Cholesky υλοποιήθηκε στα πλαίσια υπολογισμού του αντίστροφου Εσσιανού πίνακα.

3.2.1 Μέθοδος Cholesky

Ας υποθέσουμε, ότι έχουμε προς επίλυση ένα πραγματικό γραμμικό σύστημα εξισώσεων με n αγνώστους του οποίου ο πίνακας των συντελεστών των άγνωστων A είναι συμμετρικός δηλαδή: $A^T=A$. Είναι προφανές ότι λόγω της συμμετρίας δεν χρειάζεται να αποθηκεύσουμε όλα τα στοιχεία του A, αρκούν τα στοιχεία του κάτω ή άνω τριγωνικού μέρους του ,που είναι όλα πλήθους $n(n+1)/2$ αντί n^2 που είναι όλα, δηλαδή περίπου τα μισά. Το ερώτημα που γεννιέται τότε είναι μήπως μπορούμε να εκμεταλλευτούμε τη συμμετρία του A και να τον παραγοντοποιήσουμε κατά τέτοιο τρόπο ώστε να έχουμε ένα γινόμενο ενός κάτω και ενός άνω τριγωνικού πίνακα όπου ο ένας θα αποτελεί τον ανάστροφο του άλλου. Στην περίπτωση αυτή θα αρκεί μόνο η αποθήκευση του ενός από τους δύο παράγοντες. Αυτό, φυσικά υπό την προϋπόθεση ότι οι παράγοντες θα έχουν στοιχεία πραγματικά. Η απάντηση είναι καταφατική αν ο A έχει μία επιπλέον ιδιότητα αυτή του θετικά ορισμένου πίνακα. Θα παρατεθούν αναγκαίος ορισμός και θεώρημα.

Ορισμός

Έστω πίνακας $A \in \mathbb{R}^{(nXn)}$ με $A^T=A$. Ο A είναι ένας θετικά ορισμένος πίνακας αν και μόνο αν $\forall\, x \in \mathbb{R}^n \backslash \{0\}$ ισχύει $(x,Ax)_2 = (Ax,x)_2 > 0$.

Θεώρημα

Έστω πίνακας $A \in \mathbb{R}^{(nXn)}$, συμμετρικός και θετικά ορισμένος. Υπάρχει μοναδικός κάτω τριγωνικός πίνακας $L \in \mathbb{R}^{(nXn)}$ με $l_{ii}>0$,$i=0,1,2\ldots,n$ τ.ω. $LL^T=A$.

Για την εύρεση αλγορίθμου που θα υπολογίζει τα στοιχεία του πίνακα L στην παραγοντοποίηση Cholesky χρησιμοποιούμε την ισότητα $A^T=A$ με A πραγματικό συμμετρικό και θετικά ορισμένο. Εκτελούμε τον πολλαπλασιασμό στο αριστερό μέλος και εξισώνουμε ένα προς ένα τα στοιχεία του κάτω τριγωνικού πίνακα του γινομένου, που προκύπτει με τα αντίστοιχα στοιχεία του A. Είναι δυνατόν να βρίσκουμε τα στοιχεία του L είτε προχωρώντας από την πρώτη γραμμή προς τη n-ιοστή είτε από την πρώτη στήλη προς τη n-ιοστή. Ακολουθούν τα αναλυτικά ο βήματα του αλγόριθμου, προχωρώντας κατά γραμμές.

Αλγόριθμος Cholesky κατά γραμμές

Δεδομένα: $A \in IR^{(n,n)}$, $A^T=A$, A:θετικάορισμένος

Για i=1(1) n

Για j=1(1)i − 1

$$l_{ij} = \left(a_{ij} - \sum_{k=1}^{j-1} l_{ik} l_{jk} \right) / l_{jj}$$

Τελος 'Για'

$$l_{ii} = \left(a_{ij} - \sum_{j=1}^{j-1} l_{ij} \right)^2$$

Τέλος 'Για'

Δηλαδή μπορούμε να γράψουμε για τον A την παρακάτω ισότητα γινομένων πινάκων:

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{12} & \alpha_{22} & \alpha_{23} \\ \alpha_{13} & \alpha_{23} & \alpha_{33} \end{pmatrix} = \begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix} \begin{pmatrix} l_{11} & l_{21} & l_{31} \\ 0 & l_{22} & l_{32} \\ 0 & 0 & l_{33} \end{pmatrix}$$
$$= LL^T$$

3.2.2 Μέθοδοι Εύρεσης Ριζών

Στην συνέχεια περιγράφονται σε μορφή ψευδοκώδικα βήματα αλγορίθμων. Οι δύο παρακάτω αλγόριθμοι είναι πολύ γνωστοί για την εφαρμογή τους σε προβλήματα μίας διάστασης για την εύρεση ρίζας μίας συνάρτησης.

Αλγόριθμος Regula-Falsi:

Δεδομένα:Συνάρτηση f , ε μία πολύ μικρή θετική ποσότητα
Αρχή
Βήμα1: Πρόβλεψη δύο σημείων x_0,x_1
Βήμα2: Υπολογίζουμε τις τιμές $f(x_0)$ και $f(x_1)$
Βήμα3: Ελέγχουμε το πρόσημο του γινόμενου: $f(x_0)f(x_1)$
Αν $f(x_0)f(x_1) > 0$ πάμε στο Βήμα1
Αν $f(x_0)f(x_1) < 0$ πάμε στο Βήμα5
Βήμα4: Υπολογίζουμε την ποσότητα: $x = \dfrac{x_0 f(x_1) - x_1 f(x_0)}{f(x_0) - f(x_0)}$
Βήμα5: Ελέγχουμε το πρόσημο του γινόμενου: $f(x_0)f(x_1)$
Αν $f(x_0)f(x_1) > 0$ θέτουμε $x_1 = x$.
Αν $f(x_0)f(x_1) < 0$ θέτουμε $x_0 = x$.
Βήμα6: Ελέγχουμε αν $f(x) > \varepsilon$
Αν ισχύει πάμε στο Βήμα4
Αν δεν ισχύει πάμε στο Βήμα7
Βήμα7: Ρίζα ο αριθμός x
Τέλος

Αλγόριθμος Διχοτόμησης:

Δεδομένα: Συνάρτηση f ,ε μία πολύ μικρή θετική ποσότητα,
Αρχικό διάστημα αβεβαιότητας [α,β]

Βήμα1 : Όσο β - α > 0

Βήμα2 : $x_1 = \frac{a+\beta}{2} - \frac{\varepsilon}{2}$

$x_2 = \frac{a+\beta}{2} + \frac{\varepsilon}{2}$

Υπολογίζουμε τις $f(x_1)$,$f(x_2)$

Βήμα3 : Αν $f_1 < f_2$ τότε

$\beta = x_2$

Αλλιώς

$\alpha = x_1$

Τέλος 'Αν'

Τέλος ' Όσο'

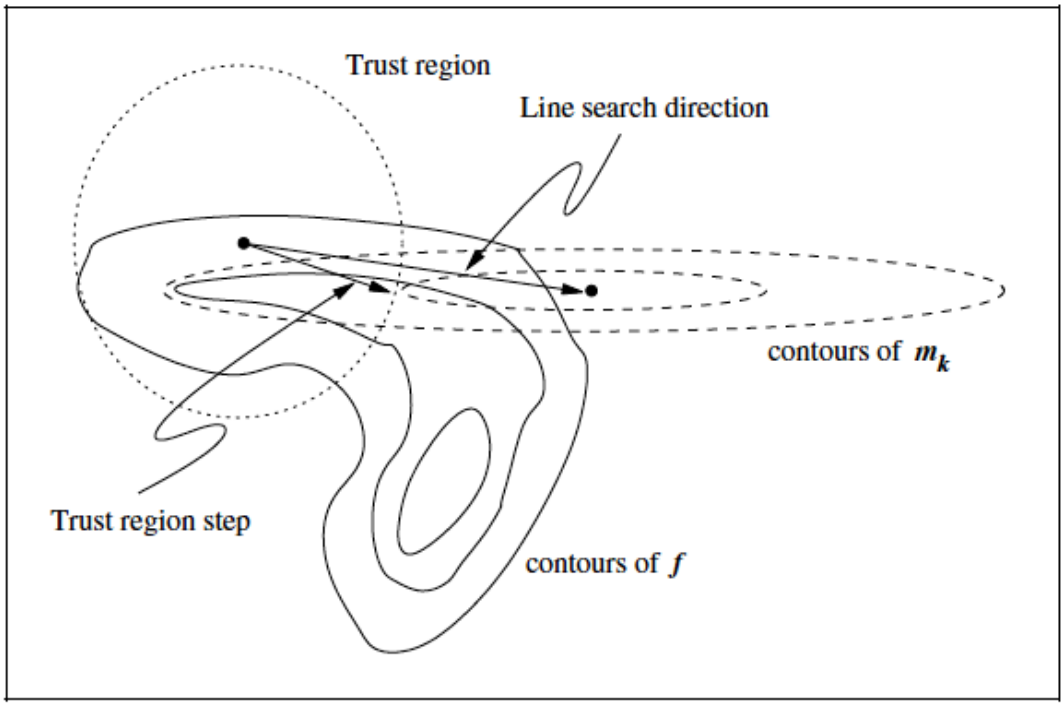
$x^* = \frac{\alpha+\beta}{2}$

$f^* = f(x^*)$

3.3 Στρατηγική «Περιοχής Εμπιστοσύνης»

Όπως αναφέρθηκε εν συντομία στο Κεφάλαιο 2, οι μέθοδοι περιοχής εμπιστοσύνης ορίζουν μια περιοχή γύρω από την τρέχουσα επανάληψη στην οποία εμπιστεύονται το μοντέλο ώστε να είναι μια επαρκής αναπαράσταση της αντικειμενικής συνάρτησης και στη συνέχεια επιλέγουν το βήμα, να είναι ο κατά προσέγγιση ελαχιστοποιητής μοντέλο στην περιοχή αυτή. Στην πραγματικότητα, επιλέγουν ταυτόχρονα την κατεύθυνση και τη διάρκεια του βήματος. Αν ένα βήμα δεν είναι αποδεκτό, μειώνουν το μέγεθος της περιοχής και βρίσκουν ένα νέο ελαχιστοποιητή. Σε γενικές γραμμές, η κατεύθυνση του βήματος αλλάζει όποτε το μέγεθος της περιοχής εμπιστοσύνης τροποποιείται. Το μέγεθος της περιοχής εμπιστοσύνης είναι κρίσιμο για την αποτελεσματικότητα κάθε βήματος. Ακολουθεί διεξοδική ανάλυση για αυτή την στρατηγική. Στην εικόνα Σχήμα 3.1 φαίνονται οι ισοϋψείς καμπύλες της συνάρτησης, οι ισοϋψείς καμπύλες του τετραγωνικού μοντέλου m_k, το βήμα εμπιστοσύνης της περιοχής και το βήμα ευθύγραμμης αναζήτησης.

Αυτό που θα πρέπει να τονισθεί είναι πως καθ' όλη την εργασία, έχουμε ότι η συνάρτηση μοντέλου m_k που χρησιμοποιείται σε κάθε μία επανάληψη x_k είναι τετραγωνική. Επιπλέον, m_k βασίζεται στο ανάπτυγμα της σειράς Taylor της f γύρω από το x_k ,όπως περιγράφεται στο προηγούμενο κεφάλαιο.



Σχήμα 3.1

Το πρόβλημα βελτιστοποίησης υπό την χρήση αυτής της στρατηγικής διατυπώνεται ως εξής:

$$\min_{p \epsilon IR} \; m_k(p) = f_k + g_k^T p + \frac{1}{2} p^T B_k p \quad \text{με } ||p|| \leq \Delta_k \qquad \text{Εξ. 3.1}$$

όπου Δ_k η ακτίνα της περιοχής εμπιστοσύνης και $||\cdot||$ η Ευκλείδια Νόρμα.

3.3.1 Περιγραφή Περιοχής Εμπιστοσύνης

Ένα από τα βασικά κλειδιά ενός αλγόριθμου περιοχής εμπιστοσύνης είναι η στρατηγική για την επιλογή της ακτίνα εμπιστοσύνης k σε κάθε επανάληψη. Βάσει

αυτής της επιλογής στη συμφωνία μεταξύ της συνάρτηση μοντέλου p_k και της αντικειμενικής συνάρτησης f σε προηγούμενες επαναλήψεις. Με δεδομένο ένα βήμα p_k εμείς καθορίζουμε την ακόλουθη αναλογία:

$$\rho_k = \frac{f(x_k) - f(x_k + p_k)}{m_k(0) - m_k(p_k)} \qquad \text{Εξ. 3.2}$$

Όπου ο αριθμητής ονομάζεται πραγματική μείωση και ο παρονομαστής είναι η προβλεπόμενη μείωση (δηλαδή, η μείωση της f που προβλέπεται από τη συνάρτηση του μοντέλου). Σημειώνουμε ότι μιας και το βήμα p_k επιτυγχάνεται με την ελαχιστοποίηση του μοντέλου m_k σε μια περιοχή που περιλαμβάνει το $p = 0$, η προβλεπόμενη μείωση θα είναι πάντοτε μη αρνητική. Επομένως, αν το p_k είναι αρνητικό, η νέα αντικειμενική τιμή $f\ (x_k + p_k)$ είναι μεγαλύτερη από την τρέχουσα τιμή $f\ (x_k)$, οπότε το βήμα πρέπει να απορριφθεί.

Από την άλλη πλευρά, αν το p_k είναι κοντά στο 1, υπάρχει καλή συμφωνία μεταξύ του μοντέλου m_k και της συνάρτησης f πάνω από αυτό το βήμα, επομένως είναι ασφαλές να επεκταθεί η περιοχή εμπιστοσύνης για την επόμενη επανάληψη. Αν το p_k είναι θετικό αλλά σημαντικά μικρότερο από 1, δεν αλλάζουμε την περιοχή εμπιστοσύνης, αλλά αν είναι κοντά στο μηδέν ή αρνητικό, συρρικνώνουμε την περιοχή εμπιστοσύνης μειώνοντας το k στην επόμενη επανάληψη.

Ο σχετικός αλγόριθμος και το θεώρημα επ'αυτού

Αλγόριθμος 3.1 (Περιοχή Εμπιστοσύνης)

Δοθέντος $R > 0 \in (0, R)$ και $\eta \in [0, 1/4)$:

Για k=0,1,2,…

1. Λαμβάνουμε p_k προσεγγιστικά λύνοντας την Εξίσωση (3.1)
2. Υπολογίζουμε την ποσότητα ρ_k από την Εξίσωση (3.2)

3. **Αν** $\rho_k < \frac{1}{4}$

$\Delta_{k+1} = \frac{1}{4}\Delta_k$

Αλλιώς

4. **Αν** $\rho_k < \frac{3}{4}$ και $||\rho_k|| = \Delta_k$

$\Delta_{k+1} = \min\,(2\Delta_k, R)$

Αλλιώς

$\Delta_{k+1} = \Delta_k$

Τέλος_αν

Τέλος_αν

5. **Αν** $p_k > \eta$

$x_{k+1} = x_k + \rho_k$

Αλλιώς

$x_{k+1} = x_k$

Τέλος_αν

6.Τέλος Επανάληψης

Εδώ το R είναι μια συνολική δέσμευση για τα μήκη βημάτων. Σημειώστε ότι η ακτίνα αυξάνεται μόνο αν το $||p_k||$ φθάνει πραγματικά στο όριο της περιοχής εμπιστοσύνης. Εάν το βήμα παραμένει αυστηρά εντός της περιοχής, συμπεραίνουμε ότι η τρέχουσα τιμή του k δεν παρεμβαίνει στην εξέλιξη του αλγορίθμου, οπότε αφήνουμε την τιμή του αμετάβλητη για την επόμενη επανάληψη.

Για να μετατρέψουμε τον Αλγόριθμο 3.1 σε έναν πρακτικό αλγόριθμο, πρέπει να επικεντρωθούμε στην επίλυση του υποπροβλήματος εμπιστοσύνης Εξίσωση (3.1). Όταν συζητούμε αυτό το θέμα, πετάμε μερικές φορές τον δείκτη επανάληψης k και επαναδιατυπώνουμε το πρόβλημα Εξίσωση (3.1) ως εξής:

$$\min_{p\epsilon IR} m(p) = \mathrm{f} + \mathrm{g}^{\mathrm{T}}p + \frac{1}{2}p^T \mathrm{B}p \quad \text{με } ||p|| \leq \Delta \qquad \text{Εξ. 3.3}$$

Ένα πρώτο βήμα για τον χαρακτηρισμό των ακριβών λύσεων της Εξίσωσης (3.3) δίνεται από το ακόλουθο θεώρημα που δείχνει ότι η λύση p^* του ικανοποιεί την

$$(B + \lambda I)p^* = -g \ \ , \gamma\iota\alpha\ \kappa\acute{\alpha}\pi o\iota o\ \lambda \geq 0 \qquad \text{Εξ. 3.4}$$

Θεώρημα:

Το διάνυσμα p^* είναι μια ολική λύση του προβλήματος της περιοχής εμπιστοσύνης ,δηλαδή του: $\min_{p\epsilon IR} m(p) = \mathrm{f} + \mathrm{g}^{\mathrm{T}}p + \frac{1}{2}p^T \mathrm{B}p$ με $||p|| \leq \Delta$,αν και μόνο το p^* είναι εφικτό και υπάρχει ένα βαθμωτό λ ≥ 0 έτσι ώστε να πληρούνται οι ακόλουθες συνθήκες : $(B + \lambda I)p^* = -g$ Εξ. 3.5

$\lambda(\Delta - ||p^*||) = 0$ και Εξ. 3.6

$(B + \lambda I)$ είναι θετικά ημιορισμένος Εξ. 3.7

Θα πρέπει να σημειωθεί πως εξαιτίας της Εξίσωση (3.5) ισχύει η ακόλουθη ισότητα:

$$\lambda p^* = -Bp^* - g = -\nabla m(p^*)$$

Στην επόμενη παράγραφο περιγράφονται δύο στρατηγικές για την εύρεση προσεγγιστικών λύσεων του υποπροβλήματος Εξίσωση (3.1), που επιτυγχάνουν τουλάχιστον τόσο μείωση του m_k όσο και η μείωση που επιτυγχάνεται από το λεγόμενο *σημείο Cauchy*. Αυτό το σημείο είναι απλά ο ελαχιστοποιητής του m_k κατά μήκος της κατεύθυνσης των πιο απότομων καθόδου -m_k. υπό την επιφύλαξη του ορίου της περιοχής εμπιστοσύνης. Η πρώτη προσεγγιστική στρατηγική είναι η μέθοδος dogleg, η οποία είναι κατάλληλη όταν το μοντέλο Hessian B_k είναι θετικά ορισμένο ενώ η δεύτερη στρατηγική, γνωστή ως δισδιάστατη ελαχιστοποίηση υποχώρου, μπορεί να εφαρμοστεί όταν B_k είναι αόριστος, αν και απαιτεί μια εκτίμηση της πιο αρνητικής ιδιοτιμής αυτής του πίνακα.

3.4 Αλγόριθμοι που βασίζονται στο σημείο Cauchy

Όπως γνωρίζουμε από το Κεφάλαιο 1, οι μέθοδοι ευθύγραμμης αναζήτησης μπορούν να συγκλίνουν σε καθολικό επίπεδο ακόμη και όταν δεν χρησιμοποιείται το βέλτιστο

μήκος βημάτων σε κάθε επανάληψη. Στην πραγματικότητα, το μήκος βήματος α_k πρέπει να ικανοποιεί μόνο αρκετά κι όχι τόσο αυστηρά κριτήρια. Ανάλογη κατάσταση ισχύει στις μεθόδους περιοχής εμπιστοσύνης. Παρόλο που επιδιώκουμε την βέλτιστη λύση του υποπροβλήματος Εξίσωση (3.1) αρκεί για τους σκοπούς της καθολικής σύγκλισης να βρεθεί ,μια κατά προσέγγιση, λύση p_k που να βρίσκεται μέσα στην περιοχή εμπιστοσύνης και να δίνει μια επαρκή μείωση του μοντέλου. Η επαρκής μείωση μπορεί να προσδιοριστεί ποσοτικά με βάση το σημείο Cauchy, το οποίο υποδηλώνουμε με p_k^C και ορίζεται με βάση την ακόλουθη απλή διαδικασία:

Αλγόριθμος 3.2 (Υπολογισμός σημείου Cauchy)

Βρίσκουμε το διάνυσμα p_k^S το οποίο είναι λύση της γραμμικής εκδοχής της Εξίσωσης (3.1) το οποίο είναι :

$$p_k^S = arg\min_{p\epsilon IR^n} f_k + g_k^T p \quad \mu\varepsilon \ \ ||p|| \leq \Delta_k \qquad \text{Εξ. 3.8}$$

Υπολογίζουμε το βαθμωτό $\tau_k > 0$ το οποίο ελαχιστοποιεί το $m_k(\tau p_k^S)$ υπό τον περιορισμό να ικανοποιείται το όριο της περιοχής εμπιστοσύνης το οποίο διατυπώνεται ως ακολούθως:

$$\tau_k = argarg\min_{\tau\geq 0} m_k(\tau p_k^S) \ \ \text{με} \ \left|\left|\tau p_k^S\right|\right| \leq \Delta_k \qquad \text{Εξ.3.9}$$

Θέτοντας $p_k^C = \tau_k p_k^S$, είναι εύκολο να διαμορφώσουμε έναν ορισμό κλειστής μορφής του σημείου Cauchy. Για αρχή, η λύση του προβλήματος Εξίσωση (3.8) είναι απλή και είναι η ακόλουθη :

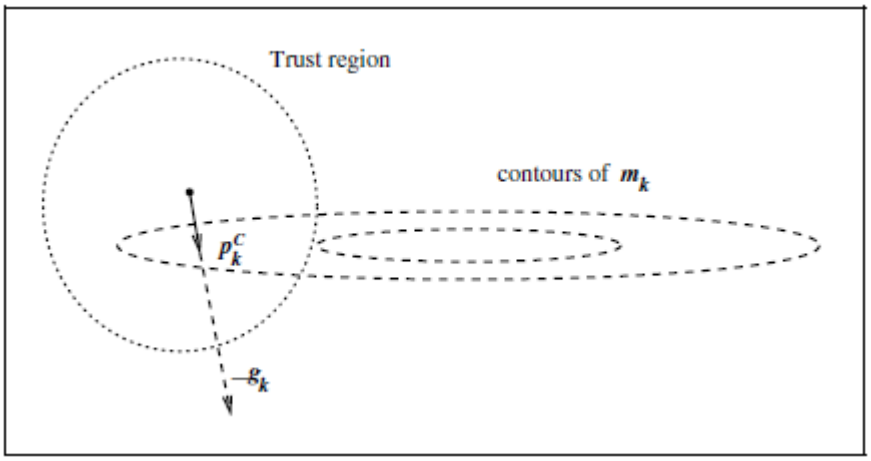
$$p_k^S = \frac{\Delta_k}{||g_k||} g_k$$

Για να λάβουμε ακριβώς την τιμή του βήματος Cauchy θα πρέπει να λάβουμε υπ όψιν δύο περιπτώσεις ξεχωριστά αν η ποσότητα $g_k^T B_\kappa g_k$ είναι θετική ή αρνητική. Τυπικά, η συνάρτηση $m_k \tau p_k^S$ μειώνει μονότονα ,λόγω του τ. Έτσι το $\tau\alpha_\kappa$ είναι η μεγαλύτερη τιμή που ικανοποιεί το όριο της περιοχής εμπιστοσύνης ,κυρίως για τα=1. Για την περίπτωση τώρα που η ποσότητα $g_k^T B_\kappa g_k > 0$ το $m_k \tau p_k^S$ είναι ένα τετραγωνικό κυρτό μοντέλο συναρτήσει του τ. Τότε είναι ενας δίχως περιορισμούς ελαχιστοποιητής του $\frac{||g_k||^3}{\Delta_\kappa g_k^T B_\kappa g_k}$ ή η τιμή όριο που ισούται με 1.

Γενικεύοντας, μπορούμε να εξάγουμε το εξής συμπέρασμα:

$$p_k^c = -\tau_\kappa \frac{\Delta_k}{||g_k||} g_k$$

$$\text{Όπου: } \tau_\kappa = \begin{cases} 1 & , \quad g_k^T B_\kappa g_k \leq 0 \\ \min\left(\frac{||g_k||^3}{\Delta_\kappa g_k^T B_\kappa g_k}, 1\right), & \delta\iota\alpha\varphi o\rho\varepsilon\tau\iota\kappa\acute{\alpha} \end{cases}$$



Σχήμα 3.2 Σημείο Cauchy

Το σημείο Cauchy είναι φτηνό υπολογιστικά δεν υπάρχουν παραγοντοποιήσεις πίνακα - και είναι κρίσιμης σημασίας για να αποφασιστεί εάν είναι αποδεκτή μια κατά προσέγγιση λύση του υποπροβλήματος της εμπιστοσύνης. Συγκεκριμένα, μια μέθοδος περιοχής εμπιστοσύνης θα έχει ολική σύγκλιση εάν τα βήματά του p_k δίνουν μια μείωση στο μοντέλο m_k που είναι τουλάχιστον κάποια σταθερή θετική πολλαπλάσια τιμή της μείωσης που επιτεύχθηκε από το βήμα Cauchy. Στην εικόνα Σχήμα 3.2 παρατηρούμε πως το σημείο Cauchy είναι αυστηρά εντός της περιοχής εμπιστοσύνης πράγμα το οποίο δεν αποτελεί σταθερά.

Μιας και το σημείο Cauchy παρέχει επαρκή μείωση της συνάρτησης του τετραγωνικού μοντέλου, για να αποφέρει ολική σύγκλιση και επειδή το κόστος υπολογισμού είναι τόσο μικρό, γιατί να αναζητήσουμε ένα ακόμη σημείο για μια

καλύτερη λύση κατά προσέγγιση για την Εξίσωση (3.1); Ο λόγος είναι ότι παίρνοντας πάντα το σημείο Cauchy ως βήμα μας, απλά εφαρμόζουμε την πιο απότομη καθοδική μέθοδο με μια συγκεκριμένη επιλογή του βήματος. Όπως γνωρίζουμε, η απότομη κάθοδος εκτελείται με κάποιου είδους ανεπάρκεια, ακόμη και αν χρησιμοποιείται ένα βέλτιστο μήκος βημάτων σε κάθε επανάληψη.

Το σημείο Cauchy δεν εξαρτάται κατά πολύ από τη μήτρα B_k, η οποία χρησιμοποιείται μόνο στον υπολογισμό του μήκους βήματος. Η ταχεία σύγκλιση μπορεί να αναμένεται μόνο εάν ο B_k παίζει ρόλο στον καθορισμό της κατεύθυνσης του βήματος καθώς και του μήκους του, και εάν ο B_k περιέχει έγκυρες πληροφορίες καμπυλότητας σχετικά με τη συνάρτηση. Ένα πλήθος αλγορίθμων περιοχής εμπιστοσύνης υπολογίζουν το σημείο Cauchy και στη συνέχεια προσπαθούν να το βελτιώσουν αυτό. Η στρατηγική βελτίωσης συχνά σχεδιάζεται έτσι ώστε το πλήρες βήμα ${p_k}^B = -B_k^{-1} g_k$ επιλέγεται όταν ο B_k είναι θετικά ορισμένος και $||p_k^B|| \leq \Delta_k$.Όταν ο B_k είναι ακριβώς ο Εσσιανός ή μια προσέγγιση του,αυτή η στρατηγική αναμένεται να οδηγήσει σε υπεργραμμική σύγκλιση.

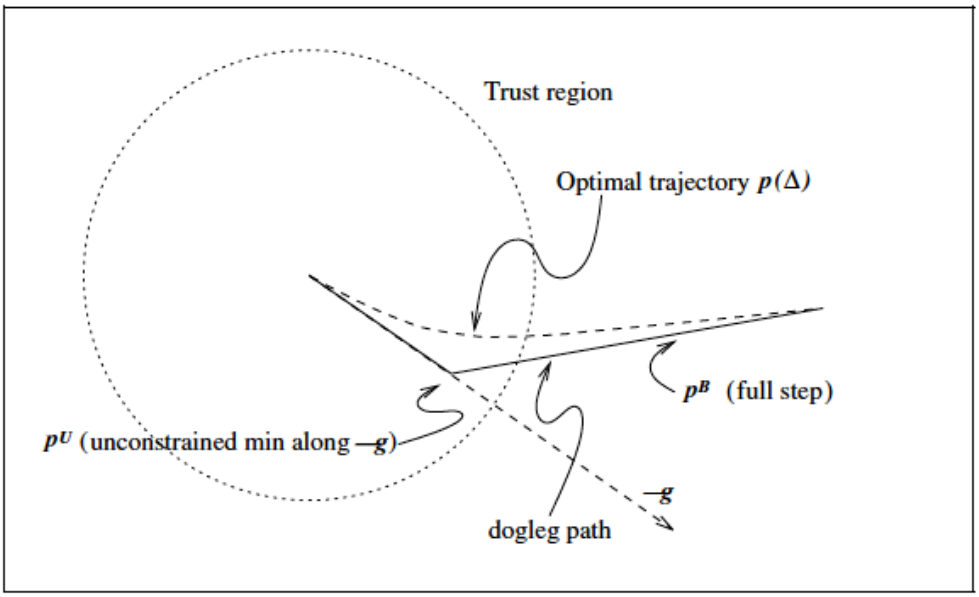
Στις επόμενες παραγράφους θα αναφερθούμε σε μεθόδους που ικανοποιούν αυτά τα χαρακτηριστικά.

3.4.1 Σημείο dogleg

Η προσέγγιση που συζητάμε πηγαίνει μπορεί να χρησιμοποιηθεί όταν ο B είναι θετικό. Ξεκινάμε εξετάζοντας την επίδραση της ακτίνας Δ της περιοχής εμπιστοσύνης για την λύση *p*(Δ)* του υποπροβλήματος Εξίσωση (3.1). Όταν ο B είναι θετικά ορισμένος, έχουμε ήδη επισημάνει ότι ο δίχως περιορισμούς ελαχιστοποιητής του μοντέλου m είναι ο $p = -B^{-1}g$. Όταν αυτό το σημείο είναι εφικτό για την Εξίσωση (3.1) είναι φανερά μία λύση οπότε εμείς έχουμε:

$$p^*(\Delta) = p^B \ , \quad \text{όταν} \quad \Delta \geq ||p^B||$$

Όταν η Δ είναι μικρή σε σχέση με το p^B, ο περιορισμός $||p|| \leq \Delta$ εξασφαλίζει ότι ο τετραγωνικός όρος στο m έχει μικρή επίδραση στην λύση της Εξίσωση (3.1).



Σχήμα 3.3

Για τέτοια Δ, μπορούμε να πάρουμε μια προσέγγιση στο p (Δ) απλά παραλείποντας τον τετραγωνικό όρο από την Εξίσωση (3.1) και γράφοντας:

$$p^*(\Delta) \approx -\Delta \frac{g}{|g|}\ , \quad \text{όταν η } \Delta \text{ είναι πολύ μικρή}$$

Για τις ενδιάμεσες τιμές του Δ, η λύση $p^*(\Delta)$ ακολουθεί τυπικά μια καμπύλη τροχιά όπως αυτό στο Σχήμα 3.3.

Η μέθοδος dogleg βρίσκει μια κατά προσέγγιση λύση αντικαθιστώντας την καμπύλη τροχιά για $p^*(\Delta)$ με διαδρομή που αποτελείται από δύο ευθύγραμμα τμήματα. Το πρώτο ευθύγραμμο τμήμα τρέχει από την προέλευση στον ελαχιστοποιητή του m κατά μήκος της πιο απότομης κατεύθυνσης, η οποία είναι:

$$p^U = -\frac{g^T g}{g^T B g} g$$

ενώ το δεύτερο ευθύγραμμο τμήμα τρέχει από p^U σε p^B όπως βλέπουμε και στο Σχήμα 3.3. . Αυτή την τροχιά την ορίζουμε από το $\tilde{p}(\tau)$ για τ ∈ [0, 2], όπου:

$$\tilde{p}(\tau) = \begin{cases} \tau p^U, & 0 \leq \tau \leq 1 \\ p^U + (\tau - 1)(p^B - p^U), & 1 \leq \tau \leq 2 \end{cases}$$

Η μέθοδος dogleg επιλέγει το p για να ελαχιστοποιήσει το μοντέλο m κατά μήκος αυτής της διαδρομής, με την διατήρηση του ορίου στην ακτίνα της περιοχής εμπιστοσύνης. Το παρακάτω, ευκόλως αποδεικνυόμενο, λήμμα δείχνει ότι το ελάχιστο κατά μήκος του μονοπατιού dogleg μπορεί να βρεθεί πολύ εύκολα.

Λήμμα

Έστω ότι ο πίνακας B είναι θετικά ορισμένος τότε:

1. $||\tilde{p}(\tau)||$ είναι μία αύξουσα συνάρτηση ως προς τ
2. $m(\tilde{p}(\tau))$ είναι μία φθίνουσα συνάρτηση ως προς τ

Εν κατακλείδι, η στρατηγική dogleg μπορεί να προσαρμοστεί για να χειριστεί αόριστες μήτρες B, αλλά δεν υπάρχει σημαντικότητα για να γίνει αυτό επειδή το πλήρες βήμα p^B δεν είναι ο μη περιορισμένος ελαχιστοποιητής του m σε αυτήν την περίπτωση. Αντ 'αυτού, περιγράφουμε τώρα μια άλλη στρατηγική, η οποία στοχεύει να συμπεριλάβει κατευθύνσεις αρνητικής καμπυλότητας (δηλαδή, κατευθύνσεις d για τις οποίες $d^T Bd < 0$) στο χώρο των υποψηφίων βημάτων της περιοχής εμπιστοσύνης.

3.4.2 Υποπρόβλημα Ελαχιστοποίησης Δύο Διαστάσεων

Όταν ο B είναι θετικός, η στρατηγική της μεθόδου dogleg μπορεί να γίνει ελαφρώς πιο εξελιγμένη διευρύνοντας την αναζήτηση του p σε ολόκληρο την δισδιάστατη υποπεριοχή που καλύπτεται από p^U και p^B. Το υποπρόβλημα Εξίσωση (3.1) αντικαθίσταται από το:

$\min_p m(p) = f + g^T p + \frac{1}{2} p^T B,$ με $||p|| \leq \Delta$ και $p \in$ στον χώρο $[g, B^{-1}g]$

Αυτό είναι ένα πρόβλημα δύο μεταβλητών που είναι υπολογιστικά φθηνό για να λυθεί. (Μετά από κάποιες αλγεβρικές παρεμβάσεις μπορεί να μειωθεί για την εύρεση των ριζών ενός πολυωνύμου τέταρτου βαθμού.) Είναι προφανές ότι το σημείο Cauchy είναι εφικτό για το προαναφερθέν πρόβλημα έτσι ώστε η βέλτιστη λύση αυτού του υποπροβλήματος να αποφέρει τουλάχιστον μια μείωση στο m όπως το σημείο Cauchy, με αποτέλεσμα την καθολική σύγκλιση του αλγορίθμου. Η στρατηγική δισδιάστατης ελαχιστοποίησης του υποχώρου είναι προφανώς μια επέκταση της μεθόδου dogleg, καθώς ολόκληρη η διαδρομή dogleg βρίσκεται στο διάστημα $[g, B^{-1}g]$. Αυτή η στρατηγική μπορεί να τροποποιηθεί για να χειριστεί την περίπτωση του απροσδιόριστου B με τρόπο που είναι διαισθητικός, πρακτικός και θεωρητικά καλός.

Στην ουσία αν ο πίνακας B έχει αρνητικές ιδιοτιμές τότε το υποπρόβλημα δύο διαστάσεων παράγεται στον χώρο των διανυσμάτων $[g, (B + \alpha I)^{-1}]$ για κάποιο $\alpha \in (-\lambda_1, -2\lambda_1)$, όπου λ_1 είναι η μικρότερη αρνητική τιμή του πίνακα.

3.5 Μέθοδος Newton

3.5.1 Βασική Μορφή Αλγορίθμου

Σε αυτήν την παράγραφο θα παρουσιάσουμε τη μέθοδο newton για πολυδιάστατη βελτιστοποίηση. Πρόκειται για τον πιο σημαντικό αλγόριθμος βελτιστοποίησης δεύτερης τάξης (που χρησιμοποιεί δηλαδή και τον πίνακα δευτέρων παραγωγών). Η μέθοδος προκύπτει από το ανάπτυγμα taylor της Εξίσωση (3.1) κρατώντας και τους τρεις όρους. Παραγωγίζοντας το τετραγωνικό μοντέλο m_k ως προς p παίρνουμε:

$$\nabla m_k(p) = 0 \Rightarrow g_k + G_k p = 0 \Rightarrow G_k p = -g_k \qquad \text{Εξ. 3.10}$$

και λύνοντας το σύστημα της Εξίσωσης (3.10) παίρνουμε τη μετατόπιση p_k την οποία προσθέτουμε στην τρέχουσα επανάληψη και παίρνουμε την επομένη: $x_{k+1} = x_k + p_k$.

Στην ουσία η μέθοδος Newton προσεγγίζει την αντικειμενική συνάρτηση με το τετραγωνικό μοντέλο της Εξίσωσης (3.1) και σε κάθε επανάληψη προσθέτει το ελάχιστο του συγκεκριμένου μοντέλου .

Παρακάτω ακολουθούν τα βήματα του αλγόριθμου της μεθόδου Newton για βελτιστοποίηση σε βασική μορφή ψευδοκώδικα.

Αλγόριθμος Newton

Δεδομένα: Συνάρτηση f:IR^n $\rightarrow$ IR ,αρχικό σημείο x(0)
Βήμα1: $k = 0$
Βήμα2: Όσο ικανοποιούνται τα κριτήρια σύγκλισης
Βήμα3: $p_k = -G_k^{-1} g_k$
Βήμα4: $x_k = x_{k+1} + p_k$
Βήμα5: $k = k + 1$
Βήμα6: Τέλος 'Οσο'

3.5.1.Σύγκλιση μεθόδου Newton

Η βασική μορφή της μεθόδου που περιγράφεται σε αυτόν τον αλγόριθμο δεν είναι σίγουρο ότι θα συγκλίνει πάντα στο ελάχιστο. Μάλιστα υπάρχουν και περιπτώσεις που μπορεί να αποκλίνει ή ακόμα σε περίπτωση που ο G δεν αντιστρέφεται και να καταρρεύσει. Στο ακόλουθο θεώρημα παραθέτουμε τις συνθήκες, οι οποίες αν πληρούνται, η βασική μέθοδος του Newton να συγκλίνει στον ελαχιστοποιητή με τετραγωνικό αριθμό σύγκλισης.

Θεώρημα

Έστω f μία συνεχής ,δύο φορές παραγωγίσιμη συνάρτηση ορισμένη σε ένα ανοιχτό κυρτό σύνολο $S \subset \mathbb{R}^n$. Έστω ότι, ακόμη, η $\nabla^2 f(x)$ είναι συνεχής κατά Lipschitz στο

S, δηλαδή: $\|\nabla^2 f(x) - \nabla^2 f(y)\| \leq \|x - y\|$ για όλα τα x,y $\in S$ και για κάποια σταθερά $L < \infty$. Αν υπάρχει ελαχιστοποιητής x^* τέτοιος ώστε να είναι θετικά ορισμένος τότε οι επαναλήψεις του Αλγόριθμου Newton συγκλίνουν με τετραγωνό ρυθμό στο x^*.

3.5.2 Τοπική Σύγκλιση Μεθόδων Netwon Trust Region

Σε αυτό το σημείο θα αναφερθούμε σε σύγκλιση τοπικού χαρακτήρα. Το κλειδί για την επίτευξη γρήγορου ρυθμού σύγκλισης, που συσχετίζεται συνήθως με τη μέθοδο του Newton είναι να δείξει ότι τα όρια απ' την περιοχή εμπιστοσύνης τελικά δεν παρεμβαίνουν καθώς πλησιάζουμε σε μια λύση. Συγκεκριμένα, ελπίζουμε ότι κοντά στην λύση, η (κατά προσέγγιση) λύση του υποπροβλήματος της εμπιστοσύνης είναι καλά ορισμένο μέσα στην περιοχή εμπιστοσύνης και μετατρέπεται όλο και πιο κοντά στο πραγματικό βήμα Newton. Τα βήματα που ικανοποιούν την τελευταία ιδιότητα λέγεται ότι είναι παρόμοια με τα βήματα Newton .

ΚΕΦΑΛΑΙΟ 4. Η ΕΠΕΚΤΑΣΗ ΤΟΥ ΑΛΓΟΡΙΘΜΟΥ “DOG-LEG”

4.1 Ανάλυση και Παρουσίαση
4.2 Ακριβές Dogleg
4.3 Διαχείριση Εσσιανού
4.4 Επέκταση Λύσης στον Τρισδιάστατο Χώρο

4.1 Ανάλυση και Παρουσίαση

Στην παρούσα εργασία όπως αναφέρθηκε και στην εισαγωγή παρουσιάζεται μία τεχνική η οποία βελτιώνει την μέθοδο “Dogleg” κατά Powell σε δύο πτυχές. Αρχικά επιλύει, το προκύπτον υποπρόβλημα βελτιστοποίησης δύο-διαστάσεων, ακριβώς και χειρίζεται στο ίδιο σημείο και τον θετικά ορισμένο και ημιορισμένο ή αρνητικά ορισμένους Εσσιανούς πίνακες. Το σχήμα είναι αποτελεσματικό, υπολογιστικά, υπό το πρίσμα της απόδοσης και απλό στην εφαρμογή.

4.1.1 Γενικό Πρόβλημα

Το γενικό πρόβλημα μπορεί να παρουσιαστεί ως εξής:

Να βρεθεί το τοπικό ελάχιστο της συνάρτησης $f(x)$, $x\epsilon\, \mathbb{R}^n$.

Η συνάρτηση είναι δύο φορές παραγωγίσιμη και η παράγωγος (κλίση) της υπολογίζεται ως $g = \nabla f(x)$ ενώ ο συσχετισμένος Εσσιανός πίνακας ως $G = \nabla^2 f(x)$.

Στις μεθόδους Newton χρησιμοποιείται ένα τετραγωνικό μοντέλο $Q(h)$ για να προσεγγίσει τοπικά την συνάρτηση ως ακολούθως :

$$f(x+h) - f(x) \approx Q(h) \equiv h^T g + \frac{1}{2} h^T B h \qquad \text{Εξ. 4.1}$$

Αν ο Εσσιανός πίνακας είναι θετικά ορισμένος τότε $B = G$, διαφορετικά ως B ορίζεται ένας τροποποιημένος (πίνακας) G, τέτοιος ώστε να υπάρχει ελαχιστοποιητής του $Q(h)$.Υπάρχουν διάφορες δυνατές επιλογές για να έχουμε έναν θετικά ορισμένο πίνακα B τροποποιώντας έναν απροσδιόριστο Εσσιανό πίνακα. Συνήθως αυτό που κάνουμε είναι να τον ορίζουμε ως εξής $B = G + vI$, όπου $v > |\lambda_1|$, όπου λ_1 είναι η πιο αρνητική ιδιοτιμή του G.

Μία αρκετά επιτυχημένη στρατηγική είναι να κατασκευάσουμε τον πίνακα B μέσω της παραγοντοποίησης Cholesky για τον G όπως προτείνεται

Το βήμα h υπολογίζεται από την επίλυση του προβλήματος ελαχιστοποίησης του $Q(h)$ επιλύοντας κυρίως το $\nabla Q(h) = 0$. Εξ. 4.2

Το οποίο ισοδύναμα είναι

$$Bh = -g \Rightarrow h = -B^{-1} g. \qquad \text{Εξ.4.3}$$

Η τεχνική Περιοχής Εμπιστοσύνης περιορίζει το $|h|$ έτσι ώστε το $Q(h)$ να προσομοιώνει-προσεγγίζει την διαφορά $f(x+h) - f(x)$ με ικανοποιητική ακρίβεια. Μία κοινή επιλογή είναι να προσδιορίσουμε μία ακτίνα R (ακτίνα περιοχής εμπιστοσύνης) και να απαιτηθεί η ακόλουθη ανίσωση $|h| \leq R$. Συνεπώς το αρχικό πρόβλημα μπορεί να παρουσιαστεί ως εξής:

$\min Q(h)$, υπό τον περιορισμό $|h| \leq R$.

Η τακτική για να επιλύσουμε το παραπάνω πρόβλημα υπό τον περιορισμό ορίζεται ως ακολούθως:

Ορίζουμε το Newton Step ως: $\boldsymbol{h_N = -B^{-1} g}$ Εξ. 4.4

Αν $|h| \leq R$, Τότε: $h = h_N$.

Διαφορετικά λύνουμε το $\min Q(h)$, με $|h| = R$. Εξ. 4.5

Χρησιμοποιώντας τους πολλαπλασιαστές Lagrange το πρόβλημα στην Εξίσωση (4.5) είναι ισοδύναμο με το : $(B + \alpha I)h(a) = -g$, $|h(a)| = R$, $a \geq 0$.

4.1.2 Δομή Τεχνικής: Περιοχή Εμπιστοσύνης

Στην πορεία παραθέτουμε την δομή του αλγορίθμου την οποία εφαρμόσαμε στην πορεία ελαφρώς τροποποιημένη. Σαφώς, ο ακόλουθος αλγόριθμος βασίζεται στον Αλγορίθμου 3.1 του Κεφαλαίου 3.

Δοθέντος ενός αρχικού σημείου $x_0 \in \mathbb{R}^n$ επιλέγουμε τιμή για το $R_0 \in (0, R_{max})$.
Η ποσότητα R_{max} είναι ένα άνω όριο για την ακτίνα της περιοχής εμπιστοσύνης και ως R_0 ορίζεται μία αρχική εκτίμηση. Θέτω $k = 0$ και η $\in [0, 1/4)$ και ε μία πολύ μικρή θετική σταθερά.

Βήματα

1.Υπολογίζουμε την κλίση $g_k = \nabla f(x_k)$.
Αν $|g_k| \leq \varepsilon$, τότε $x^* = x_k$ και σταματάμε εδώ.

2.Υπολογίζουμε ή «ανανεώνουμε» τον πίνακα B_k . (Έναν τροποποιημένο και θετικά ορισμένο Εσσιανό πίνακα)

3.Βρίσκουμε το h_k που ελαχιστοποιεί το $Q_k(h) \equiv h^T g_k(x) + \frac{1}{2} h^T B_k h$ υπό τον περιορισμό $|h| \leq R$.

4.Υπολογίζουμε την πραγματική προς την αναλογία μείωσης που προβλέφθηκε

$$\rho_k = \frac{f(x_k) - f(x_k + h_k)}{Q_k(0) - Q_k(h_k)}$$

5.Ενημερώνουμε κατάλληλα την ακτίνα της περιοχής εμπιστοσύνης ως εξής:

Αν $\rho_k < 0.25$

Τότε $R_{k+1} = 0.25|h_k|$

Αλλιώς

Αν $\rho_k > 0.75$ και $|h_k| = R_k$

Τότε $R_{k+1} = \min(2R_k, R_{max})$

Αλλιώς

$R_{k+1} = R_k$

Τέλος αν

Τέλος αν

6. Αποδεχόμαστε ή απορρίπτουμε το νέο σημείο ως εξής:

   Αν $\rho_k > \eta$

   Τότε $x_{k+1} = x_k + h_k$ και θέτουμε J=1.

   Αλλιώς

   $x_{k+1} = x_k$ και θέτουμε J=0

   Τέλος αν

7. k=k+1
8. Αν J=0 πήγαινε στο βήμα *3*.
9. Αν J=1 πήγαινε στο βήμα *1*.

4.1.3 DogLeg κατά Powell

Για να λυθεί ακριβώς η Εξίσωση (4.9) κάποιος θα πρέπει επανειλημμένα να λύσει ένα σχετικό γραμμικό σύστημα για τις διάφορες τιμές του α μέχρις ότου ικανοποιηθεί η ισότητα $|h(a)| = R$. Αν η διάσταση του προβλήματος είναι μεγάλη τότε υπολογιστικά είναι κοστοβόρο. Επομένως μία προσεγγιστική μα «φτηνή» λύση θα προτιμηθεί. Ο Powell, στην τεχνική του, χρησιμοποίησε ένα επίπεδο που αποτελείται από τα διανύσματα h_N και $-g$ για να λύσει το πρόβλημα προσεγγιστικά. Επομένως αναφερόμαστε σε αυτό το επίπεδο ως επίπεδο DLP (DogLeg Plane). Ας θεωρήσουμε $H = B^{-1}$ και το σημείο Cauchy h_C κατά μήκος της κλίσης. Οπότε:

$$h_C = -\frac{g^{\mathrm{T}} g}{g^{\mathrm{T}} \mathrm{B} g} g$$

Συνεπώς το h υπολογίζεται από την ακόλουθη διαδικασία:

Αν $|h_C| \geq R$

Τότε $h_C = -R \frac{g}{|g|}$

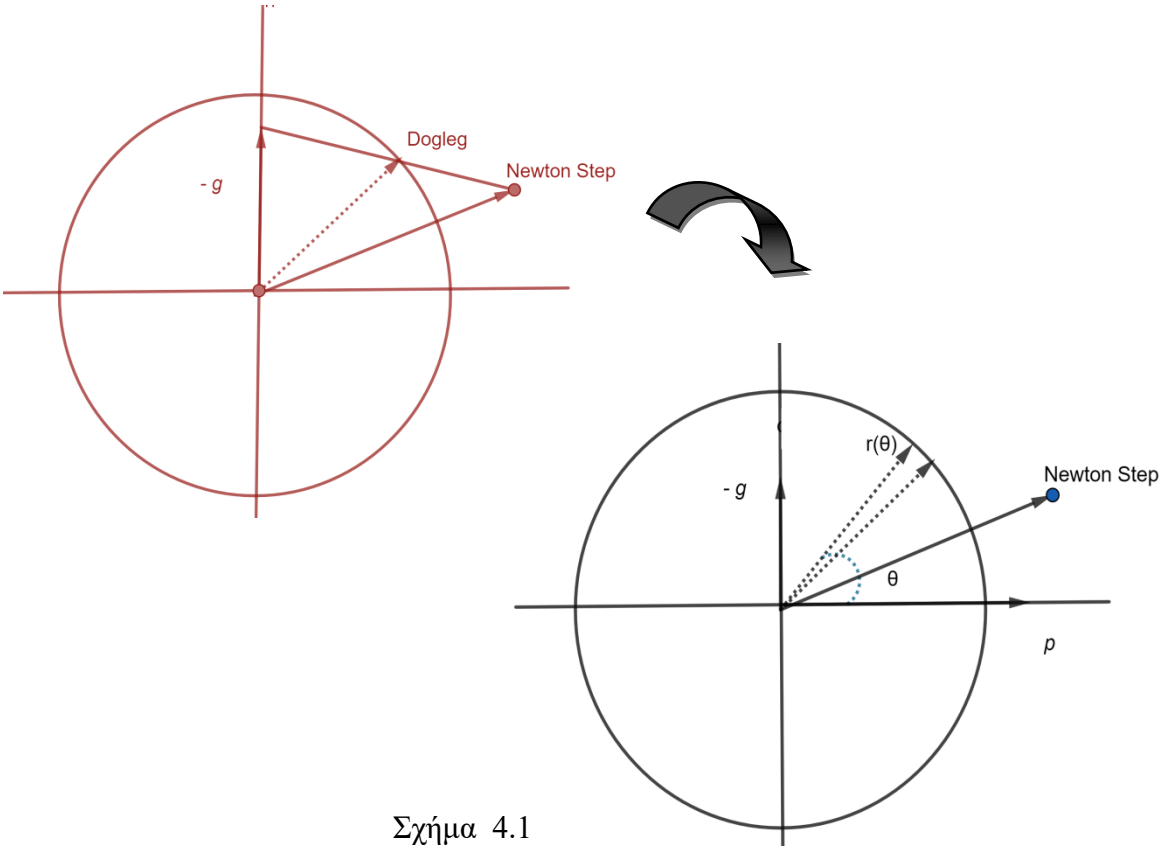
Αλλιώς

$h = h_C + \beta(h_N - h_C)$, όπου β ∈ (0,1) είναι τέτοιο ώστε $|h_C + \beta(h_N - h_C)| = R$

Η τελευταία εξίσωση είναι μια απλή τετραγωνική εξίσωση στο β. Οι Dennis και Mei επεσήμαναν πως το dogleg h έχει μία «προτίμηση» προς την απότομη καθοδική κατεύθυνση και κατά συνέπεια ανέπτυξαν μία βελτιωμένη έκδοση με το όνομα «double dogleg» και βελτίωσαν τον παράγοντα Newton για να πάρουν γρηγορότερη σύγκλιση. Και οι δύο μέθοδοι βρίσκουν ένα h στο DLP ου ικανοποιεί την ισότητα $|h| = R$ και ελαχιστοποιούν το τετραγωνικό μοντέλο $Q(h)$ προσεγγιστικά.

4.2 Ακριβές Dog-Leg

Στην ουσία αυτό που περιγράφεται σε αυτό το τμήμα της εργασίας γίνεται ευκόλως κατανοητό με την εικόνα Σχήμα 4.1. Αποφεύγουμε πλέον την τυχαία προσέγγιση και δουλεύουμε μία μία πιο ακριβή «προσέγγιση». Στο πάνω μέρος της εικόνας το διάνυσμα που οδηγείται από το O→ Cauchy Point → Newton Point είναι κάθε φορά εκείνο που δημιουργείται από την τομή πανω στον κύκλο του ευθύγραμμου τμήματος των δύο points. Στο κάτω μέρος βλέπουμε πως αυτή ακριβώς η διαδικασία δεν είναι τυχαία.Καθώς όπως θα γίνει κατανοητό και στην πορεία το διάνυσμα Dogleg δεν είναι απλά ένα τυχαίο σημείο τομής αλλά θα εξαρτάται από ένα ακόμη διάνυσμα.



Σχήμα 4.1

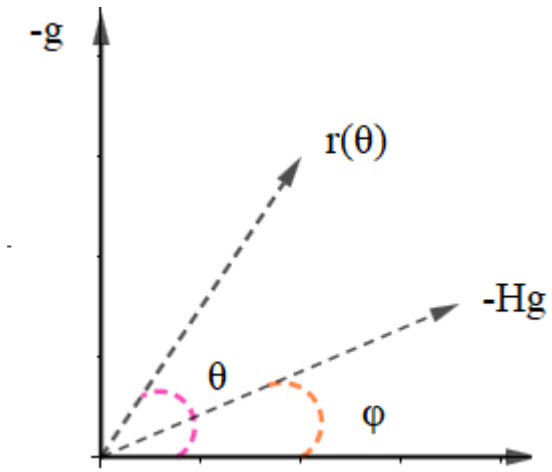
Ένα διάνυσμα p, στο βασικό DLP , και στο επίπεδο του g ,μπορεί να κατασκευαστεί ως εξής:

$$\mathrm{p} = \left(\mathrm{I} - \frac{\mathrm{H}}{\langle \mathrm{H} \rangle}\right) g\,, \quad \text{όπου} \quad \langle \mathrm{H} \rangle = \left(\mathrm{I} - \frac{g^{\mathrm{T}} H g}{g^{\mathrm{T}} g}\right)$$

Η γεωμετρία του DLP φαίνεται στο Σχήμα 4.2. Ως ω ορίζεται η γωνία μεταξύ των $-Hg$ και p, ενώ $r(\theta)$ είναι ένα μοναδιαίο διάνυσμα στο DLP που σχηματίζει με το διάνυσμα p γωνία θ και συνεπώς ορίζεται ως εξής:

$$r(\theta) = \frac{p}{|p|}\cos(\theta) - \frac{g}{|g|}\sin(\theta)$$

Μιας και $g^T r(\theta) = -|g| sin(\theta)$, $r(\theta)$ είναι καθοδικά $\forall\ \theta \in (0,\pi)$. Η λύση για το h στο DLP με |h| = R θα μπορούσε να γραφτεί δηλαδή έτσι : $h = Rr(\theta),\ \forall\ \theta \in (0,\pi)$.

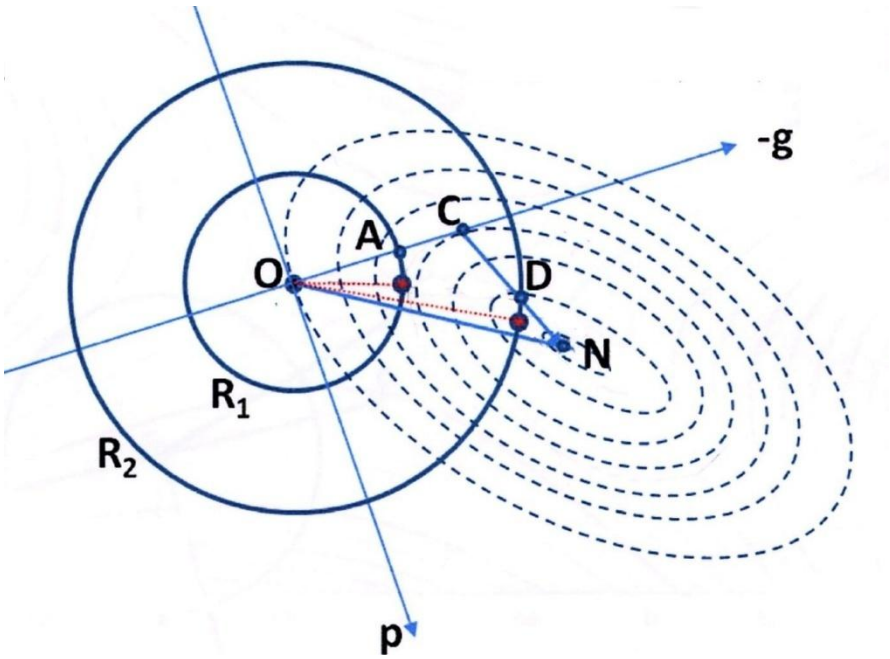


Σχήμα 4.2. Η γεωμετρία του επιπέδου του dogleg

Έτσι $h = h(\theta)$ υπολογίζεται συναρτήσει μίας γωνίας θ. Επομένως λαμβάνοντας υπ' όψιν την τελευταία εξίσωση που γράψαμε καθώς και την Εξίσωση (4.1) λαμβάνουμε τις ακόλουθες σχέσεις:

$$Q(\Theta) = \frac{1}{2}R^2\left[\frac{1}{2}\left(\frac{p^TBp}{p^Tp} - \frac{g^TBg}{g^Tg}\right)\cos(2\theta) - \frac{p^TBg}{p^Tg}\sin(2\theta)\right] - R|g|\sin(\theta) + \frac{1}{4}R^2\left(\frac{p^TBp}{p^Tp} + \frac{g^TBg}{g^Tg}\right) \qquad \text{Εξ. 4.18}$$

$$\frac{dQ(\theta)}{d\theta} = -\frac{1}{2}R^2\left(\frac{p^TBp}{p^Tp} - \frac{g^TBg}{g^Tg}\right)sin(2\theta) + \frac{p^TBg}{p^Tg}\cos(2\theta) - |g|Rcos(\theta) \qquad \text{Εξ. 4.19}$$



Σχήμα 4.3. Σημεία Cauchy (C) , Newton(N) και Dogleg(D)

Έχοντας την αναλυτική έκφραση, μπορεί εύκολα κάποιος να υπολογίσει την γωνία έχοντας κανείς την αναλυτική έκφραση για το μοντέλο και την παράγωγο, μπορεί εύκολα να υπολογίσει γωνία $\theta^* \in (0,\pi)$ που ελαχιστοποιεί το μοντέλο. Στο Σχήμα 4.3 βλέπουμε τις ισοϋψείς καμπύλες από ένα κυρτό τετραγωνικό μοντέλο και δύο περιοχές εμπιστοσύνης με ακτίνες εμπιστοσύνης $R1$ και $R2$ αντίστοιχα. Το σημείο Cauchy $\boldsymbol{h_C} = \overrightarrow{OC}$ είναι έξω από τον μικρό κύκλο και μέσα στον μεγάλο κύκλο, παράγοντας τα βήματα dogleg $\overrightarrow{OA}$ και $\overrightarrow{OD}$ αντίστοιχα .Τα ακριβή σημεία dogleg είναι με τις συμπαγείς τελείες στους δύο κύκλους και συνδέονται με την περιοχή με τις διακεκομμένες κόκκινες γραμμές.

4.3 Διαχείριση απροσδιοριστίας Εσσιανού Πίνακα

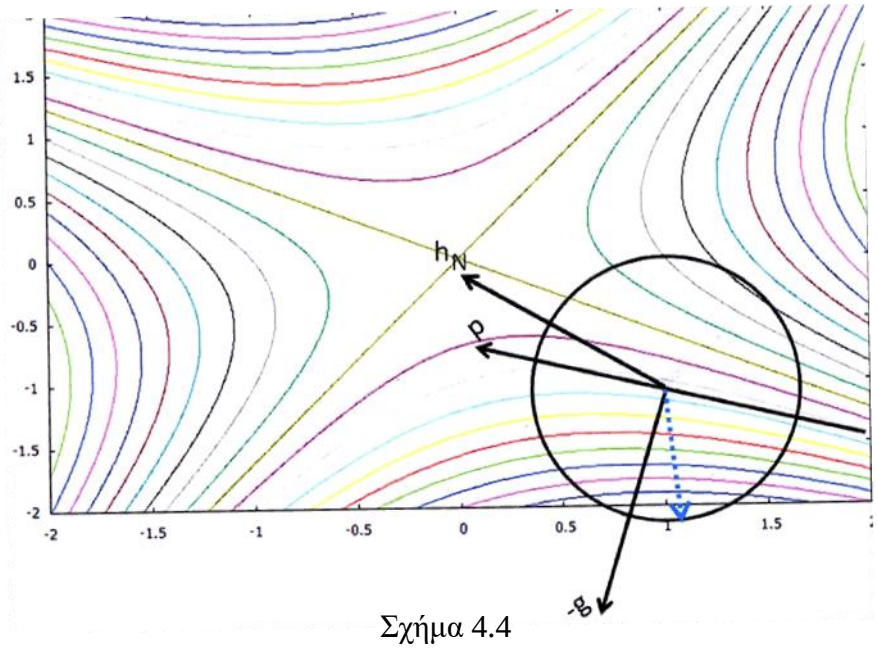
Σε αυτή την παράγραφο παρουσιάζεται ο εναλλακτικός τρόπος χρήσης του Εσσιανού πίνακα χωρίς την μετατροπή του σε θετικά ορισμένο πίνακα. Όταν ο G δεν είναι θετικά ορισμένος η τετραγωνική του προσέγγιση στην Εξίσωση (4.1) χρησιμοποιεί μία αντικατάστασή του με τον B, η οποία αναπόφευκτα είναι χαμηλής ποιότητας.

Η διαδικασία «ακριβές dogleg» όπως περιγράφηκε στην προηγούμενη παράγραφο δεν βασίζεται στον θετικά ορισμένο πίνακα B και πάντα θα βρίσκει ένα ελάχιστο στην περιφέρεια του κύκλου για $\theta \in (0,\pi)$.Υπάρχουν αρκετά σημεία στα οποία πρέπει να δοθεί ιδιαίτερη έμφαση:

- Αν ο G δεν είναι θετικά ορισμένος, το τετραγωνικό μοντέλο ίσως να έχει παραπάνω από ένα ακρότατο. Συνεπώς το να μηδενίσω απλά την παράγωγο δεν αρκεί. Μιας και ο υπολογισμός του $Q(\theta)$ σε διάφορα θ (γωνίες δλδ.) είναι φτηνός υπολογιστικά, για την εύρεση του ελαχίστου μπορεί να εφαρμοστεί μία τεχνική «ομαδοποίησης» μέσω αναζήτησης λεπτού πλέγματος (fine grid search). Όταν παίρνουμε μία «ομάδα», ένας μηδενισμός της παραγώγου μπορεί να τοποθετηθεί μέσα στην ομάδα με επιτυχή γραμμική παρεμβολή και περαιτέρω μείωση της ομάδας.
- Ένας άλλος προβληματισμός, είναι ότι οι παράγοντες της παραγοντοποίησης Cholesky ενός αόριστου πίνακα ίσως να μην υπάρχουν .Όμως για να

επιτευχθεί μία λύση στο $Gh_N = -g$ ίσως χρειαστεί να πάρουμε και να λύσουμε το ισοδύναμο σύστημα $G^2 h_N = -Gg$. Αν ο G είναι τετραγωνικός πίνακας και έχει αντίστροφο πίνακα (είναι non-singular δλδ.) τότε ο G^2 είναι θετικά ορισμένος και μπορεί να αναλυθεί σε παραγοντοποίηση Cholesky,έτσι ώστε με ασφάλεια να έχουμε υπολογίσει την ακόλουθη ποσότητα για το βήμα: $h_N = -G^{-2}(Gg) = -G^{-1}\mathrm{g}$.

- Όταν ο πίνακας G είναι απροσδιόριστος (αόριστος) ίσως το h_n να είναι καθοδικό. Όμως να μην είναι απαραίτητα ελαχιστοποιητής για το τετραγωνικό μοντέλο Q(θ). Για να το δούμε αυτό θεωρούμε ένα βήμα $h_n + \delta$ αντί αυτού. Επομένως $Q(h_N + \delta) - Q(h_N) = \frac{1}{2}\delta^T G\delta$, το οποίο μπορεί να είναι αρνητικό για κάποια κατάλληλη τιμή του δ, για παράδειγμα κατά μήκος ενός ιδιοδιανύσματος του G που αντιστοιχεί σε μία αρνητική ιδιοτιμή. Συνεπώς θα έχουμε επιπλέον όφελος εάν ελαχιστοποιήσουμε το τετραγωνικό μοντέλο σε ένα κύκλο ακτίνας $|h_N|$.



Σχήμα 4.4

Στο Σχήμα 4.4 παρουσιάζονται οι ισοϋψείς καμπύλες ενός μη κυρτού τετραγωνικού μοντέλου. Σε αυτή την περίπτωση το Newton βήμα είναι ανοδικής κατεύθυνσης μιας

και $g^T h_N > 0$. Η τεχνική Exact-Dogleg προσδιορίζει (υπολογίζει) το σημείο ελαχιστοποίησης στον κύκλο, με τη διακεκομμένη γαλάζια γραμμή. Αφότου επισημάναμε τα πιο σημαντικά σημεία, μπορούμε να αναπλάσουμε την Δομή της Τεχνικής Περιοχής Εμπιστοσύνης, αφήνοντας ανέπαφο το τμήμα της αρχικοποίησης, που αναφέρθηκε στην παράγραφο *4.1.2*, ως ακολούθως:

Βήματα

1.Υπολογίζουμε την κλίση $g_k = \nabla f(x_k)$.

Αν $|g_k| \leq \varepsilon$, τότε $x^* = x_k$ και σταμάταμε εδώ.

2.Υπολογίζουμε $G_k = \nabla^2 f(x_k)$.

3.Βρίσκουμε το h_k που ελαχιστοποιεί το $Q_k(h) \equiv h^T g_k(x) + \frac{1}{2} h^T B_k h$ υπό τον περιορισμό $|h| \leq R_k$.

4.Υπολογίζουμε την πραγματική προς την προβλεπόμενη αναλογία μείωσης ρ_k

5. Ανανέωση ακτίνας περιοχής εμπιστοσύνης (αναλύεται παρακάτω)

6. k = k+1

7. Αν $\rho_k > \eta$

Τότε αποδεχόμαστε το νέο σημείο και πηγαίνουμε στο βήμα 1.

Αλλιώς

Απορρίπτουμε το σημείο και επανατακτοποιώ: $R_k = \min\,(R_k, a|h_k|)$ και το α∈ $(0,1)$ και πηγαίνουμε στο βήμα 3.

Τέλος αν

Περαιτέρω , αναγκαία , ανάλυση για τα βήματα **3.** και **5.** .

Βήμα **3**.

Υπολογίζουμε $h_N = -G_k^{-1} g_k$

Θέτουμε $R_C = min\,(R, |h_N|)$ και $h = R_C r(\theta)$ και ελαχιστοποιούμε το $Q_k(\theta)$ για να πάρουμε θ_k και $h_k = R_C r(\theta_k)$

Βήμα **5**.

$$R = R * m(\rho)$$

$$\text{όπου} \qquad m(\rho) = \begin{cases} 0.2\,, & \rho < 0 \\ 0.2 + 2.4p\,, & \rho \in [0, 0.75] \\ 0.2\,, & \rho \in (0, 1{,}25) \\ 7 - 4p\,, & \rho \in [1.25\,, 1.5] \\ 1\,, & \rho > 1.5 \end{cases}$$

4.4 Επέκταση λύσης στον τρισδιάστατο χώρο

Στο τελευταίο μέρος του συγκεκριμένου κεφαλαίου θα περιγραφεί η δεύτερη εκδοχή της λύσης που περιγράψαμε στις παραπάνω παραγράφους. Πρόκειται ουσιαστικά για μία επέκταση λύσης στον τρισδιάστατο χώρο όπως φαίνεται και στο Σχήμα 4.5 .

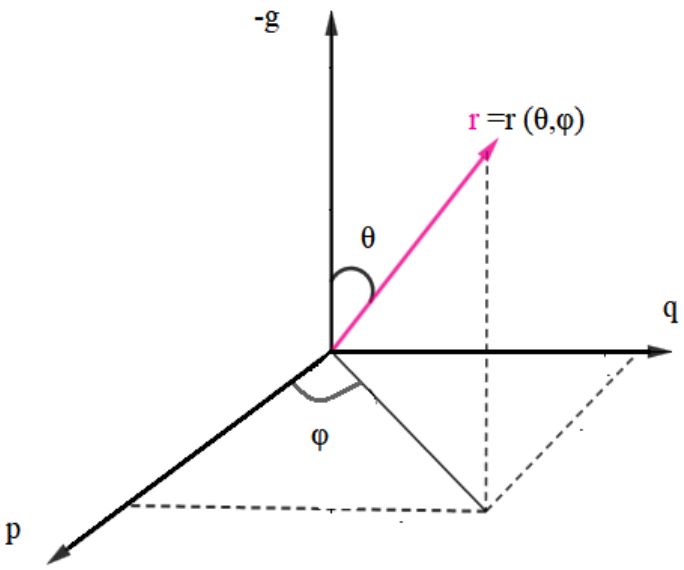
Έστω τρία μη συνεπίπεδα διανύσματα g, Hg και ένα ακόμη d μία ποσότητα ακόμη που μπορεί κανείς να την ορίσει με την ίδια λογική όπως προηγουμένως στα g και p και να κατασκευάσει ένα διάνυσμα q κάθετο στο p και στο q ,ως ακολούθως:

$$q = g + \frac{(p^T d)(g^T g)}{g^T d(p^T p)} p - \frac{g^T g}{g^T d} d$$

Έπειτα ένα μοναδιαίο διάνυσμα r (τέτοιο ώστε $r^T r = 1$) ,μπορεί να γραφτεί ως εξής

$$r(\varphi, \theta) = -\frac{g}{|g|}\cos(\theta) + \left(\frac{p}{|p|}\cos(\varphi) \quad + \frac{q}{|q|}\sin(\varphi)\right)\sin(\theta)$$

Ως ορθοκανονικό σύστημα συντεταγμένων : x-άξονας κατά μήκος του p , y- άξονας κατά μήκος του και z-άξονας κατά μήκος του $-g$. Με τις γωνίες φ$\in [0,2\pi]$ και $\theta \in \left[0, \frac{\pi}{2}\right]$.



Σχήμα 4.5

Μια σφαίρα ακτίνας R περιγράφεται από ένα διάνυσμα h(φ,θ)= Rr(φ,θ). Το τετραγωνικό μοντέλο τότε γίνεται:

$$Q(\varphi,\theta) = \frac{1}{2}R^2 r^T(\varphi,\theta)Br(\varphi,\theta) + Bg^T r(\varphi,\theta)$$

Παίρνοντας την μερική παράγωγο του τετραγωνικού μοντέλου ως προς θ και ως προς φ αντίστοιχα και εξισώνοντας τες με το μηδέν έχουμε:

- $$\frac{dQ(\varphi,\theta)}{d\theta} = 0 \Rightarrow \quad tan(\theta) = \frac{B_{gp}\sin(\varphi) - B_{qq}cos(\varphi)}{\left(B_{pp} - B_{qq}\right)\sin(2\varphi) - B_{pq}\cos(2\varphi)}$$

- $$\frac{dQ(\varphi,\theta)}{d\varphi} = 0 \Rightarrow \frac{|g|}{Rsin(\theta)} = B_{gg} - B_{pp}cos^2(\varphi) - B_{qq}\sin^2(\varphi) + \frac{2}{\tan(2\theta)}\left(B_{gp}\cos(\varphi) + B_{gq}\sin(\varphi)\right)$$

όπου:

$$B_{uv} = B_{vu} = \frac{u^T B u}{|u||v|}$$

ΚΕΦΑΛΑΙΟ 5. ΑΠΟΤΕΛΕΣΜΑΤΑ

5.1 Συναρτήσεις Πειραμάτων
5.2 Αριθμητικά Αποτελέσματα
5.3 Μελλοντική Έρευνα

Σε αυτό το κεφάλαιο αναφέρονται τα πειράματα και οι συγκρίσεις. Θα πρέπει να αναφερθεί πως στα πλαίσια της υλοποίησης, το προγραμματιστικό μέρος υλοποιήθηκε με την γλώσσα Fortran 77. Πρόκειται για μία από τις πρώτες γλώσσες υψηλού προγραμματισμού που έχει αποτελεσματικότητα και απλή εφαρμογή σε διάφορα προβλήματα κυρίως μαθηματικού περιεχομένου.

5.1 Συναρτήσεις Πειραμάτων

Στην συνέχεια παρουσιάζονται αναλυτικά οι συναρτήσεις που αποτέλεσαν το πειραματικό τμήμα αυτής της εργασίας. Αναφέρονται :το όνομα της συνάρτησης, ο τύπος της καθώς και οι παράμετροί της (αν έχει), οι διαστάσεις που μελετήθηκε.

Problem 1. Rosenbrock: $f(x) = \sum_{i=1}^{n-1}[(100(x_{i+1} - {x_i}^2)^2) - (x_i - 1)^2]$

Problem 2. Sphere: $f(x) = \sum_{i=1}^{n} x_i\, f(x)$

Problem 3.Sumsquare: $f(x) = \sum_{i=1}^{n} i{x_i}^2$

Problem 4.Rotated Ellipse: $f(x) = \sum_{1}^{d}\sum_{1}^{j} x_j^2$

Problem 5. Rastrigin: $f(x) = 10n + \sum_{i=1}^{n}\left(x_i^2 - 10\cos(2\pi x_i)\right)$

Problem 6. Qing: $f(x) = \sum_1^n (x_i^2 - i)^2$

Problem 7. Schumer Steiglitz: $f(x) = \sum_{i=1}^n x_i^4$

Problem 8. Schwefel: $f(x) = \sum_{i=1}^n [(x_i - 1)^2 + (x_1 - x_i{}^2)]$

Problem 9. Zakharov: $f(x) = \sum_{i=1}^{n} x_i^2 + \left(\frac{1}{2}\sum_{i=1}^{n} i x_i\right)^2 + \left(\frac{1}{2}\sum_{i=1}^{n} i x_i\right)^4$

Problem 10. Cosine Mixture model: $f(x) = -0.1\sum_{i=1}^{n} \cos(5\pi x_i) - \sum_{i=1}^{n} x_i^2$

Problem 11. Park: $f(x) = \frac{2}{3} e^{(x_1+x_2)} - x_4 \sin(x_3) + x_3$

Problem 12. Mattyas: $f(x) = 0.26(x_1^2 + x_2^2) - 0.48 x_1 x_2$

Problem 13. Wood:
$$f(x) = 100(x_2 - x_1{}^2)^2 + (1 - x_1)^2 + 90(x_4 - x_3{}^2)^2 + (1 - x_3)^2 + 10.1\left[(x_2 - 1) + (x_4 - 1)^2\right] + 19.8(x_2 - 1)(x_4 - 1)$$

Problem 14. Moon: $f(x) = x_1 + x_2 + 3x_1 x_3$

Στον κάθε πίνακα αναφέρεται η διάσταση, το πρόβλημα (δηλαδή η συνάρτηση) πόσες φορές, κατά μέσο όρο, κλήθηκε το πρόγραμμα να υπολογίσει την τιμή της συνάρτησης σε ένα x_κ στην πρώτη εκδοχή του dog-leg στις δύο διαστάσεις (V1) αλλά και στις τρεις διαστάσεις (V2), πόσες φορές κλήθηκε να υπολογίσει την παράγωγο την συνάρτησης σε διάφορα x_k στην πρώτη και την δεύτερη εκδοχή και τέλος πόσες φορές υπολόγισε τον Εσσιανό πίνακα. Στην τελευταία στήλη υπάρχει ο χρόνος επεξεργασίας κάθε εκδοχής με μονάδα μέτρησης τα δευτερόλεπτα.

5.2 Αριθμητικά Αποτελέσματα

Πίνακας 5.1 Συναρτήσεις με διάσταση n=5

Problem	F calls		G calls		CPU time	
	V1	V2	V1	V2	V1	V2
1.	40	34	35	31	4.0002E-004	3.8592E-004
2.	29	25	24	22	4.0001E-004	3.8592E-004
3.	47	44	43	40	4.0002E-004	3.8592E-004
4.	42	40	37	33	4.0002E-004	3.8592E-004
5.	27	32	24	29	4.0001E-004	3.8592E-004
6.	23	17	20	12	4.0001E-004	3.8592E-004
7.	29	26	24	22	4.0001E-004	3.8592E-004
8.	43	39	37	34	4.0002E-004	3.8592E-004
9.	34	32	29	27	4.0001E-004	3.8592E-004
10.	45	49	44	54	3.8592E-004	4.0001E-004

Πίνακας 5.2 Συναρτήσεις με διάσταση n=50

Problem	F calls		G calls		CPU time	
	V1	V2	V1	V2	V1	V2
1.	129	112	121	103	0.192	0.1599
2.	154	139	142	132	0.192	0.16
3.	98	93	91	84	0.191	0.1499
4.	87	80	71	68	0.192	0.1499
5.	79	76	78	73	0.192	0.1499
6.	138	127	129	129	0.194	0.1597
7.	107	96	102	102	0.192	0.1399
8.	125	120	119	119	0.192	0.1399
9.	97	85	88	80	0.191	0.138
10.	122	135	138	142	0.139	0.191

Πίνακας 5.3 Συναρτήσεις με διάσταση n=100

Problem	F calls		G calls		CPU time	
	V1	V2	V1	V2	V1	V2
1.	201	194	167	164	1.7	1.678
2.	210	207	204	195	1.7	1.67
3.	318	312	310	297	1.73	1.7
4.	184	179	171	163	1.7	1.67
5.	195	149	183	142	1.69	1.6
6.	245	237	240	233	1.7	1.67
7.	328	317	324	314	1.75	1.7
8.	289	282	275	272	1.7	1.67
9.	512	508	502	497	1.7	1.67
10.	259	263	254	260	1.6	1.69

Πίνακας 5.4 Συναρτήσεις διαφορετικών διαστάσεων

Συνάρτηση	F calls		G calls		CPU time	
	V1	V2	V1	V2	V1	V2
Wood n=4	14	11	13	9	3.0041E-04	3.0002E-04
Moon n=3	12	10	11	8	3.0045E-04	3.0002E-04
Park n=3	16	14	13	12	3.0019E-04	3.0018E-04
Mattyas n=2	7	7	7	7	1.0001E-08	1.0001E-08

Όπως παρατηρούμε από τους παραπάνω συγκριτικούς πίνακες η v2 τις περισσότερες των περιπτώσεων εμφανίζει ελαφρώς καλύτερες τιμές για το ίδιο ελάχιστο και με κοινό αρχικό σημείο προφανώς. Στην ουσία λιγότερες κλήσεις της συνάρτησης και της παραγώγου. Και όπως είναι αναμενόμενο λιγότερος χρόνος. Και τα δύο σχήματα –εκδοχές είναι αποτελεσματικά ,απλά στην υλοποίηση και καθόλου κοστοβόρα ως φαίνεται ακόμη και για μεγάλες διαστάσεις.

5.3 Μελλοντική Έρευνα

Σε αυτό το κεφάλαιο παρουσιάζουμε κάποιες μελλοντικές προς υλοποίηση και έρευνα σκέψεις, πάνω στον αλγόριθμο που περιγράφηκε και υλοποιήθηκε όπως περιγράφεται στο Κεφάλαιο 4.

Όπως είναι ευρέως γνωστό, αυτό που αποτελεί τάση πλέον στην επιστήμη της πληροφορικής, είναι ο παράλληλος προγραμματισμός. Παρράλληλος προγραμματισμός λέγεται η ανάπτυξη εφαρμογών οι οποίες εκμεταλλεύονται την ύπαρξη πολλαπλών επεξεργαστικών μονάδων σε έναν πολυεπεξεργαστή ή πολυυπολογιστη για να επιτύχουν αύξηση των υπολογιστικών επιδόσεων και μείωση του απαιτούμενου χρόνου εκτέλεσης της εφαρμογής. Επομένως ο παράλληλος προγραμματισμός μπορεί να ιδωθεί ως ειδική περίπτωση ταυτόχρονου προγραμματισμού, όπου η εκτέλεση γίνεται πραγματικά παράλληλα και όχι ψευδοπαράλληλα.

Συνήθως ένας αλγόριθμος παραλληλοποιείται με διάσπασή του σε πολλαπλά τμήματα τα οποία ανατίθενται σε ξεχωριστά νήματα ή διεργασίες και έτσι εκτελούνται παράλληλα σε διαφορετικές επεξεργαστικές μονάδες. Ωστόσο δεν είναι βέβαιο ότι ένας αλγόριθμος, υλοποιημένος σε κάποιο πρόγραμμα, μπορεί να παραλληλοποιηθεί πάντα.

Ωστόσο ακόμα και στην περίπτωση μη παραλληλοποιήσιμου προγράμματος η χρήση διαφορετικών νημάτων ή διεργασιών μπορεί να είναι ευεγερτική αν ο προγραμματιστής επιθυμεί να επικαλύψει υπολογισμούς με επικοινωνίες (π.χ. ένα νήμα να συνεχίζει τους υπολογισμούς όσο το άλλο αναμένει είσοδο από το δίκτυο ή

από τον χρήστη· αυτό έχει νόημα ακόμα και αν το πρόγραμμα εκτελείται ψευδοπαράλληλα σε μονοεπεξεργαστικό, σειριακό υπολογιστή).

Η παραλληλοποίηση ενός σειριακού αλγορίθμου, προκειμένου να εκμεταλλευτούμε την αύξηση των υπολογιστικών επιδόσεων που προσφέρει ο παραλληλισμός, διακρίνεται σε τέσσερα βήματα που εκτελούνται διαδοχικά:

1. Τη διάσπαση του ολικού υπολογισμού σε επιμέρους εργασίες.
2. Την ανάθεση εργασιών σε οντότητες εκτέλεσης
3. Την ενορχήστρωση των οντοτήτων.
4. Την αντιστοίχηση οντοτήτων σε επεξεργαστές.

Στην παρούσα εργασία θα μπορούσε μελλοντικά να υλοποιηθεί και ο παράλληλος προγραμματισμός στο σημείο υπολογισμού της κλίσης (άνυσμα πρώτης παραγώγου) και του Εσσιανού πίνακα.

Δηλαδή θα μπορούσαμε να έχουμε n+1 πυρήνες στην περίπτωση διάστασης προβλήματος n. Στο οποίο θα ίσχυε ο ακόλουθος συλλογισμός:

Ο πρώτος πυρήνας θα υπολογίζει το $gradient(x_1,\ldots,x_n)$

Και κάθε μία γραμμή του $Hessian = \begin{pmatrix} \frac{\partial^2 f}{\partial^2 x_1} & \cdots\cdots\cdots & \cdots\cdots \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ \cdots\cdots & \cdots\cdot\cdot & \frac{\partial^2 f}{\partial^2 x_n} \end{pmatrix}$ να υπολογίζεται από έναν διαφορετικό πυρήνα.

Μία ακόμη, σε πρώιμο πάντα στάδιο σκέψη, είναι η ακόλουθη: να επεκτείνω την ακριβή λύση στις τέσσερις διαστάσεις. Στην ουσία προσεγγίζω το τετραγωνικό μοντέλο με διανύσματα στον χώρο [g, Bg ,Hg ,H^2g] τα οποία ανά δύο να είναι κάθετα μεταξύ τους. Ακολουθούμε την συλλογιστική πορεία που ακολουθήσαμε στις τρεις και στις δύο διαστάσεις προφανώς.

Τέλος, μία ακόμη πρόταση για μελλοντική εργασία επί του θέματος είναι ο τρόπος που επιλέγουμε να υπολογιστεί ο πίνακας B. Δηλαδή αντί να υπολογίζουμε με διάφορες μεθόδους τον ακριβή αντίστροφο να χρησιμοποιήσουμε την γνωστή Μέθοδο SR1 με περιοχή εμπιστοσύνης η οποία χρησιμοποιεί ενημερωτικό σχήμα

(updating) για τον πίνακα. Οι πίνακες αυτοί, που παράγονται, αποτελούν πολύ καλές προσεγγίσεις των πραγματικών πινάκων δευτέρων παραγώγων. Αρκεί να προσεχθεί ιδιαίτερα το σημείο που ανανεώνεται ο πίνακας καθώς υπάρχει περίπτωση να προσεγγίζει ο παρανομαστής το μηδέν οπότε να πρέπει να παρακαμφθεί αυτή η ενημέρωση. Αυτό δεν επιφέρει ζημιά στα αποτελέσματα του αλγόριθμου.

ΑΝΑΦΟΡΕΣ

[1] J. E. Jr Dennis, H. H. W. Mei, “Two new unconstrained optimization algorithms which use function and gradient values”, Journal of Optimization Theory and Applications, 1979.

[2] J. Nocedal, S. Wright, “Numerical Optimization Theoritial and PracticalAspects”, Springer, 1999.

[3] M. J. D. Powell, “A hybrid method for nonlinear equations”, Gordon and Breach Science, London, 1970.

[4] P. E. Gill, W. Murray, M.H. Wright, “Practical Optimization”, Academic Press, May 1981.

[5] Α. Κουράκος, “Fortran 77/90/95”, Κλειδάριθμος, Αθήνα 2004.

[6] Γ. Ροβιθάκης, “Τεχνικές Βελτιστοποίησης”, Εκδόσεις Τζιόλα, Αθήνα 2007.

[7] Δ. Νούτσος, Β. Δούγαλης, Α. Χατζηδήμος, “Αριθμητική Γραμμική Άλγεβρα”, Πανεπιστημιακό Τυπογραφείο Ιωαννίνων, Ιωάννινα 2012.

[8] Η. Λυπιτάκης, “Προγραμματισμός με Fortran 77”, Εκδόσεις Οικονομικό Πανεπιστήμιο Αθηνών, 1987.

ΣΥΝΤΟΜΟ ΒΙΟΓΡΑΦΙΚΟ

Γεννήθηκα στα Ιωάννινα 04/05/1991 και κατάγομαι από την Ηγουμενίτσα. Είμαι απόφοιτη του τμήματος Μαθηματικών της Σχολής Θετικών Επιστημών του Πανεπιστημίου Ιωαννίνων. Έχω εργαστεί ως μαθηματικός σε ιδιωτικούς εκπαιδευτικούς φορείς. Εκπαιδεύομαι διαρκώς στα αντικείμενα Διδακτικής Μαθηματικών και Ειδικής Αγωγής με παρουσία δημοσιεύσεων σε σχετικά συνέδρια. Ερευνητικά ενδιαφέροντα Εφαρμοσμένα Μαθηματικά, Διδακτική Μαθηματικών, Ψυχολογία Παίδων.