ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

Τομέας Σημάτων, Ελέγχου και Ρομποτικής

Επίλυση συστημάτων εξισώσεων κυρτών συναρτήσεων με χρήση αλγορίθμου ολικής βελτιστοποίησης

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

Σάββα Π. Κιναλή

Επιβλέπων: Νικόλαος Μαράτος
Καθηγητής

Αθήνα, Ιούλιος 2013

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ
ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
Τομέας Σημάτων, Ελέγχου και Ρομποτικής

Επίλυση συστημάτων εξισώσεων κυρτών συναρτήσεων με χρήση αλγορίθμου ολικής βελτιστοποίησης

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

Σάββα Π. Κιναλή

Επιβλέπων: Νικόλαος Μαράτος
Καθηγητής

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 16/07/2013.

.....	.....	.....
Νικόλαος Μαράτος	Τρύφων Κουσιουρής	Κωνσταντίνος Τζαφέστας
Καθηγητής	Καθηγητής	Επίκουρος Καθηγητής

Αθήνα, Ιούλιος 2013.

.....

Σάββας Π. Κιναλής
Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών.

© 2013 Εθνικό Μετσόβιο Πολυτεχνείο. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα. Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Ο σκοπός της διπλωματικής εργασίας ήταν η επίλυση συστημάτων εξισώσεων κυρτών συναρτήσεων. Αυτό αντιμετωπίστηκε ανάγοντας το πρόβλημα σε ένα πρόβλημα ελαχιστοποίησης κοίλων συναρτήσεων με κυρτούς περιορισμούς. Επιλέχθηκε ένας γνωστός από τη βιβλιογραφία ντετερμινιστικός αλγόριθμος διαδοχικών εξωτερικών προσεγγίσεων για να λυθεί αυτό το πρόβλημα ελαχιστοποίησης. Με αυτό τον τρόπο επίλυσης αποκτάμε γνώση τόσο για τη λύση του συστήματος, όσο και για την ύπαρξη αυτής. Στα πλαίσια της διπλωματικής υλοποιήθηκε η μέθοδος αυτή στο προγραμματιστικό περιβάλλον του Matlab© και ελέγχθηκε η ορθή λειτουργία της σε διάφορα παραδείγματα. Τέλος, είδαμε ότι η μέθοδος αυτή, με χρήση ιδιοτήτων των συναρτήσεων κυρτών διαφορών, μπορεί να επεκταθεί και στη λύση συστημάτων εξισώσεων δύο φορές συνεχώς διαφορίσιμων στο $\mathbb{R}^n$ συναρτήσεων.

Λέξεις Κλειδιά

μη γραμμικές εξισώσεις, κοίλη ελαχιστοποίηση, κυρτές συναρτήσεις, κοίλες συναρτήσεις, αριθμητική ανάλυση, μέθοδοι διαδοχικής εξωτερικής προσέγγισης, συναρτήσεις κυρτών διαφορών

(this page is left intentionally blanc)

Abstract

The goal of this thesis was to solve systems of equations of convex functions. Our approach was to handle this as a problem of concave minimization. A known deterministic algorithm was selected, which is using the successive outer approximation method. That way, knowledge of the existence of the solution, as well as the solution itself is achieved. As part of the thesis we implemented the method using the software package Matlab© and its proper use was checked on several examples. Finally, we came to the conclusion that the method, by using properties of difference convex functions, can also be used to solve systems of equations of two times continuously differentiable on $\mathbb{R}^n$ functions.

Key Words

nonlinear equations, concave minimization, convex functions, concave functions, numerical analysis, successive outer approximation methods, d.c. functions

(this page is left intentionally blanc)

Ευχαριστίες

Καταρχήν θα ήθελα να ευχαριστήσω τον κύριο καθ. Τ. Κουσιουρή που μου αναγέννησε το ενδιαφέρον για τη σχολή, στα μαθήματά του στο 5ο εξάμηνο και με ώθησε προς τη ροή Σ. Ακόμη θα ήθελα να ευχαριστήσω και τον κύριο καθ. Γ. Παπαβασιλόπουλο καθώς και τους υποψήφιους διδάκτορες του Α. Χαραλαμπίδη και Γ. Κορδώνη. Τα μαθήματά που προσέφεραν στο 7ο, 8ο και 9ο εξάμηνο, μαζί με το μάθημα Τεχνικές Βελτιστοποίησης αποτέλεσαν το αποκορύφωμα της προπτυχιακής ακαδημαϊκής μου μόρφωσης. Τελευταίος αλλά και σημαντικότερος ο καθηγητής μου ο κύριος Νικόλαος Μαράτος, τον οποίο και ευχαριστώ θερμά που με εμπιστεύθηκε με αυτή τη διπλωματική και που με καθοδήγησε καθόλη τη διάρκειά της. Η προθυμία του να βοηθήσει με κάθε τρόπο τον φοιτητή του, καθώς και το γνήσιο ενδιαφέρον του γι΄ αυτόν είναι απαράμιλλα. Θέλω επίσης να ευχαριστώ την οικογένεια μου για τη στήριξή τους όλα αυτά τα χρόνια των σπουδών μου.

(this page is left intentionally blanc)

Περιεχόμενα

1 Εισαγωγή **1**
1.1 Μη γραμμικές εξισώσεις 1
1.2 Ελαχιστοποίηση κοίλων συναρτήσεων 4
1.2.1 Μέθοδοι διαδοχικής προσέγγισης 6

2 Διατύπωση προβλήματος και αλγόριθμος επίλυσης **11**
2.1 Διατύπωση προβλήματος 11
2.2 Αλγόριθμος 12
2.2.1 Υπολογιστικές διεργασίες 13

3 Παραδείγματα **15**
3.1 Τετραγωνικές συναρτήσεις δύο διαστάσεων 15
3.1.1 Μέθοδος Newton 23
3.1.2 Μη ύπαρξη λύσης 23
3.1.3 Ειδική περίπτωση 24
3.2 Άθροισμα εκθετικών και τετραγωνικών συναρτήσεων 26
3.2.1 Παράδειγμα σε δύο διαστάσεις 26
3.2.2 Παράδειγμα σε μεγαλύτερες διαστάσεις 28

4 Συμπεράσματα - Επεκτάσεις **33**

A Κώδικας Matlab **35**

Βιβλιογραφία **45**

(this page is left intentionally blanc)

Κεφάλαιο 1

Εισαγωγή

Σε αυτή τη διπλωματική ασχοληθήκαμε με τη λύση συστημάτων εξισώσεων κυρτών συναρτήσεων της μορφής:

$$f_i(x) = 0, \quad i = 1, 2, \ldots, n$$

,όπου $x \in \mathbb{R}^n$ και οι $f_i \colon \mathbb{R}^n \to \mathbb{R}$ κυρτές συναρτήσεις. Ανάγαμε το πρόβλημα στη λύση ενός προβλήματος ελαχιστοποίησης κοίλων συναρτήσεων, οπότε σε αυτή την εισαγωγή θα δούμε κάποια βασικά αποτελέσματα στο πεδίο των μη γραμμικών εξισώσεων καθώς και της ελαχιστοποίησης κοίλων συναρτήσεων.

1.1 Μη γραμμικές εξισώσεις

Το πιο βασικό πρόβλημα που μελετάται είναι η λύση του συστήματος μη γραμμικών εξισώσεων:

$$\begin{aligned} &\text{δεδομένης } F \colon \mathbb{R}^n \to \mathbb{R}^n \\ &\qquad \text{βρες } x_* \in \mathbb{R}^n \text{ τέτοιο ώστε } F(x_*) = 0 \end{aligned} \tag{1.1}$$

όπου η F θεωρείται ότι είναι συνεχώς διαφορίσιμη. Αγνοώντας ερωτήματα σχετικά με την ύπαρξη και τη μοναδικότητα των λύσεων του συστήματος, καθώς θεωρούμε ότι στην πράξη σπάνια αυτό θα αποτελεί πρόβλημα, θα παρουσιάσουμε αρχικά τη πλέον διαδεδομένη μέθοδο Newton (βλέπε [1]).

Η μέθοδος Newton παράγεται βρίσκοντας τη ρίζα μιας γραμμικής προσέγγισης της F στην παρούσα επανάληψη x_c:

$$F(x_c + p) = F(x_c) + \int_0^1 J(x_c + \lambda p) d\lambda.$$

Προσεγγίζουμε το ολοκλήρωμα στο $\mathbb{R}^n$ με το γραμμικό όρο $J(x_x)\cdot p$ για να πάρουμε τη γραμμική προσέγγιση της F με διαταράξεις p του x_c,

$$M_c(x_c + p) = F(x_c) + J(x_c)p.$$

Έπειτα λύνουμε για βήμα s^N το οποίο κάνει το $M(x_c + s^N) = 0$, δίνοντας την επανάληψη Newton για το αρχικό πρόβλημα. Λύσε

$$\begin{aligned} J(x_c)s^N &= -F(x_c), \\ x_+ &= x_c + s^N. \end{aligned} \tag{1.2}$$

Ένας ισοδύναμος τρόπος να δούμε αυτή τη διαδικασία είναι ότι βρίσκουμε ένα ταυτόχρονο μηδενικό του γραμμικού μοντέλου των n συνιστωσών της F που δίνεται από

$$(M_c)_i(x_c + s^N) = f_i(x_c) + \nabla f_i(x_c)^T s^N, \;\; i = 1, 2, \ldots, n.$$

Εφόσον το x_+ δεν προσδοκούμε να είναι ίσο με το x_*, αλλά μόνο μία καλύτερη εκτίμηση από το x_c, κάνουμε την επανάληψη Newton έναν αλγόριθμο, υπολογίζοντάς την επαναληπτικά ξεκινώντας από μια αρχική πρόβλεψη x_0.

Αλγόριθμος 1

Δεδομένων $F\colon \mathbb{R}^n \to \mathbb{R}^n$ συνεχώς διαφορίσιμης συνάρτησης και $x_0 \in \mathbb{R}^n$: σε κάθε επανάληψη k, λύσε

$$\begin{aligned} J(x_k)s_k &= -F(x_k), \\ x_{k+1} &= x_k + s_k. \end{aligned} \tag{1.3}$$

Θα δούμε τώρα πως μπορεί να συνδυαστεί η μέθοδος Newton με μεθόδους βελτιστοποίησης χωρίς περιορισμούς για το πρόβλημα:

δεδομένης $F\colon \mathbb{R}^n \to \mathbb{R}^n$

βρες $x_* \in \mathbb{R}^n$ τέτοιο ώστε $F(x_*) = 0$

Το βήμα της μεθόδου Newton για το πρόβλημα είναι

$$x_+ = x_c - J(x_c)^{-1}F(x_c), \tag{1.4}$$

,όπου η $J(x_c)$ είναι η Ιακωβιανή μήτρα της F στο x_c. Γνωρίζουμε ότι η παραπάνω ακολουθία είναι τοπικά τετραγωνικά συγκλίνουσα στο x_*, αλλά όχι απαραίτητα ολικά συγκλίνουσα.Τώρα θεωρούμε ότι το x_c δεν είναι κοντά σε καμία λύση x_* του αρχικού προβλήματος. Πώς θα αποφάσιζε κάποιος αν θα δεχόταν το x_+ για την επόμενη επανάληψη; Μία λογική απάντηση θα ήταν το $\|F(x_+)\|$ να είναι μικρότερο από το $\|F(x_c)\|$ για κάποια νόρμα $\|\cdot\|$, με βολική επιλογή την l_2 νόρμα $\|F(x)\|_2^2 = F(x)^T F(x)$.

Απαιτώντας το επόμενο αποτέλεσμα του βήματος μας να οδηγήσει σε μείωση της $\|F(x)\|_2$ είναι το ίδιο πράγμα με το να προσπαθούμε να βρούμε ένα ελάχιστο της συνάρτησης $\|F(x)\|_2$. Συνεπώς, έχουμε στρέψει την προσοχή μας στο αντίστοιχο πρόβλημα ελαχιστοποίησης:

$$\min_{x\in\mathbb{R}^n} f(x) = \frac{1}{2}F(x)^T F(x), \tag{1.5}$$

όπου το 1/2 χρησιμοποιείται για διευκόλυνση της μετέπειτα άλγεβρας. Αξίζει να σημειωθεί ότι κάθε λύση του αρχικού προβλήματος είναι λύση της ελαχιστοποίησης, αλλά μπορεί να υπάρχουν πολλά τοπικά ελάχιστα της (1.5) που δεν είναι λύσεις της (1.1). Επομένως, παρόλο που θα μπορούσαμε να προσπαθήσουμε να λύσουμε το πρόβλημα (1.1) απλά χρησιμοποιώντας μια ρουτίνα ελαχιστοποίησης για την (1.5), είναι προτιμότερο να χρησιμοποιήσουμε τη δομή του αρχικού προβλήματος όποτε είναι εφικτό και συγκεκριμένα για να υπολογίζουμε το βήμα Newton (1.4). Ωστόσο η γενική μας στρατηγική θα βασιστεί στην προσπάθεια για ολική λύση του προβλήματος ελαχιστοποίησης (1.5).

Μια σημαντική ερώτηση είναι: "Ποιά είναι η κατεύθυνση κατάβασης για το πρόβλημα ελαχιστοποίησης (1.5);" Είναι οποιαδήποτε κατεύθυνση p για την οποία $\nabla f(x_c)^T p < 0$, όπου

$$\nabla f(x_c) = \frac{d}{dx}\sum_{i=1}^{n}\frac{1}{2}(f_i(x_c))^2 = \sum_{i=1}^{n}\nabla f_i(x_c)\cdot f(x_c) = J(x_c)^T F(x_c).$$

Επομένως η κατεύθυνση της απότομης καθόδου για την (1.5) είναι κατά μήκος της $-J(x_c)^T F(x_c)$. Επίσης, η καινούργια κατεύθυνση της μεθόδου Newton κατά μήκος της $s^N = -J(x_c)^{-1}F(x_c)$ είναι κατεύθυνση κατάβασης, αφού

$$\nabla f(x_c)^T s^N = -F(x_c)^T J(x_c) J(x_c)^{-1} F(x_c) = -F(x_c)^T F(x_c) < 0$$

όσο $F(x_c) \neq 0$. Αυτό είναι λογικό από γεωμετρικής άποψης. Αφού το βήμα Newton αποτελεί ρίζα της

$$M_c(x_c + s) = F(x_c) + J(x_c)s,$$

πάει επίσης σε ένα ελάχιστο της τετραγωνικής συνάρτησης

$$\begin{aligned}\hat{m}_c(x_c+s) &\triangleq \frac{1}{2}M_c(x_c+s)^T M_c(x_c+s)\\ &= \frac{1}{2}F(x_c)^T F(x_c) + (J(x_c)^T F(x_c))^T s + \frac{1}{2}s^T(J(x_c)^T J(x_c))s,\end{aligned} \tag{1.6}$$

επειδή $\hat{m}_c(x_c+s) \geq 0$ για όλα τα s και $\hat{m}_c(x_c+s^N) = 0$. Επομένως, το s^N είναι μια κατεύθυνση κατάβασης για το $\hat{m}_c$, και καθώς η κλίση στο x_c της $\hat{m}_c$ και της f είναι ίδια, είναι και κατεύθυνση κατάβασης για την f.

Η παραπάνω εξέλιξη δίνει την κατεύθυνση ως προς τη δημιουργία μεθόδων για την (1.1). Θα βασιστούν στη χρησιμοποίηση αλγορίθμων γραμμικής αναζήτησης με οπισθοπορεία για το τετραγωνικό μοντέλο $\hat{m}_c(x)$ της (1.6). Αφού $\nabla^2 \hat{m}_c(x_c) = J(x_c)^T J(x_c)$, αυτό το μοντέλο είναι θετικά ορισμένο όσο η $J(x_c)$ έχει μη μηδενική ορίζουσα, το οποίο είναι συνεπές με το γεγονός ότι η $x_c + s^N$ είναι η μοναδική ρίζα της $M_c(x)$ και επομένως μοναδικός ελαχιστοποιητής της $\hat{m}_c(x)$ σε αυτή την περίπτωση. Συνεπώς το μοντέλο έχει τις ελκυστικές ιδιότητες ότι ο ελαχιστοποιητής του είναι το σημείο Newton του αρχικού προβλήματος, και όλες οι κατευθύνσεις καθόδου του είναι και κατευθύνσεις καθόδου για την $f(x)$ επειδή $\nabla \hat{m}_c(x_c) = \nabla f(x_c)$. Οπότε οι μέθοδοι που βασίζονται σε αυτό το μοντέλο, πηγαίνοντας προς τα κάτω και προσπαθώντας να ελαχιστοποιήσουν το $m_c(x)$, θα συνδυάζουν τη μέθοδο Newton για μη γραμμικές εξισώσεις με μεθόδους για ένα σχετιζόμενο πρόβλημα ελαχιστοποίησης.
Σημειώνουμε ότι το $\hat{m}_c(x)$ δεν είναι το ίδιο με το τετραγωνικό μοντέλο της $f(x) = \frac{1}{2}F(x)^T F(x)$ γύρω από το x_c,

$$m_c(x_c + s) = f(x_c) + \nabla f(x_c)^T s + \frac{1}{2} s^T \nabla^2 f(x_c) s,$$

διότι $\nabla^2 f(x_c) \neq J(x_c)^T J(x_c)$.

Τώρα μπορούμε να χρησιμοποιήσουμε ευθέως την ολική μέθοδο της γραμμικής αναζήτησης στο πρόβλημά μας.

Η ιδέα της γραμμικής αναζήτησης είναι απλή: δεδομένης μιας κατεύθυνσης καθόδου p_k, κάνουμε ένα βήμα προς αυτή την κατεύθυνση το οποίο θα μας δώσει ένα δεκτό x_{k+1}.

Στην επανάληψη k:

• υπολόγισε μια κατεύθυνση καθόδου p_k,

• θέσε $x_{k+1} := x_k + \lambda_k p_k$ για κάποιο $\lambda_k > 0$ το οποίο κάνει την x_{k+1} αποδεκτή επανάληψη.

Οπότε όσο η $J(x_c)^T J(x_c)$ είναι ασφαλώς θετικά ορισμένη, ο αλγόριθμος χρησιμοποιείται χωρίς αλλαγές, αν ορίσουμε την αντικειμενική συνάρτηση ως την $\frac{1}{2}\|F(x)\|_2^2$, τη κατεύθυνση Newton ως $-J(x)^{-1}F(x)$, και το θετικά ορισμένο τετραγωνικό μοντέλο από την (1.6). Αυτό σημαίνει ότι ψάχνουμε στην κατεύθυνση Newton, κοιτώντας για μία ικανοποιητική μείωση στην $\|F(x)\|_2$.

1.2 Ελαχιστοποίηση κοίλων συναρτήσεων

Η ελαχιστοποίηση κοίλων συναρτήσεων με κυρτούς περιορισμούς αποτελεί μία από τις πιο θεμελιώδεις και εντατικά μελετημένες κλάσεις προβλημάτων στην ολική

βελτιστοποίηση. Θα παρουσιαστούν κάποιες βασικές ιδέες από τους ντετερμινιστικούς αλγορίθμους, οι οποίοι μέχρι σήμερα έχουν δείξει ότι εκμεταλλεύονται τη δομή της ολικής ελαχιστοποίησης κοίλων συναρτήσεων αποτελεσματικά.

Το πρόβλημα που θα συζητηθεί είναι το ακόλουθο:

$$\text{glob min } f(x), \text{ subject to } x \in D, \tag{P}$$

όπου το $D \subseteq \mathbb{R}^n$ είναι ένα μη κενό, κλειστό κυρτό σύνολο, και η f είναι μία κοίλη συνάρτηση που παίρνει πραγματικές τιμές, ορισμένη σε ένα κατάλληλο κυρτό σύνολο $A \subseteq \mathbb{R}^n$ το οποίο περιέχει το D. Ο στόχος στο πρόβλημα (P) είναι να βρεθεί η (ολική) ελάχιστη τιμή, την οποία μπορεί να πάρει η f στο D, και αν αυτή δεν είναι ίση με $-\infty$, να βρεθεί, αν υπάρχει, τουλάχιστον ένα σημείο στο D που να επιτυγχάνει αυτή την τιμή. Στις περισσότερες εφαρμογές, το A είτε είναι ίσο με το $\mathbb{R}^n$ είτε με κάποιο ανοιχτό κυρτό σύνολο που περιέχει το D, και το D είναι συμπαγές. Υπό αυτές τις συνθήκες, όπως θα δούμε, η f είναι συνεχής στο A, έτσι ώστε από το γνωστό θεώρημα του Weierstrass, η ολική ελάχιστη τιμή είναι πεπερασμένη, και υπάρχει τουλάχιστον ένα σημείο στο D που παίρνει αυτή την τιμή.

Τυπικά, το D ορίζεται από ένα σύστημα γραμμικών και/ή μη γραμμικών ανισωτικών και ισωτικών περιορισμών. Σε αυτές τις περιπτώσεις μπορούμε να θεωρήσουμε ότι το D έχει τη μορφή

$$D = \{x \in A | g_i(x) \leq 0, \;\; i = 1, 2, \ldots, m\},$$

όπου για κάθε $i = 1, 2, \ldots, m, \;\; g_i \colon A \to \mathbb{R}$.

Το πρόβλημα της ελαχιστοποίησης κοίλων συναρτήσεων γενικά θα έχει πολλές λύσεις x_* που είναι τοπικά, αλλά όχι ολικά ελάχιστα. Αυτές οι λύσεις ανήκουν στο D και παίρνουν τιμές $f(x_*)$ οι οποίες είναι οι ελάχιστες της f σε κάποιες υποπεριοχές $(N_{x_*} \cap D)$ του D, όπου το N_{x_*} είναι ένα ανοιχτό σύνολο που περιέχει το x_*, αλλά οι τιμές $f(x_*)$ δεν είναι οι μικρότερες στο D. Γι' αυτό το λόγο η εφαρμογή των κλασσικών αλγορίθμων που είναι σχεδιασμένοι για τη λύση προβλημάτων κυρτού προγραμματισμού με περιορισμούς, στη γενική περίπτωση, θα αποτύχουν να λύσουν αυτής της μορφής πρόβλημα με πολλαπλά ακρότατα.

Σε αυτό το κεφάλαιο, το ολικό ελάχιστο του προβλήματος (P) θα καλείται η ολική ελάχιστη τιμή του προβλήματος (P) και θα συμβολίζεται με v. Ένα σημείο x^* στο D για το οποίο $f(x^*) = v$ θα καλείται ολική βέλτιστη λύση ή ολικό ελάχιστο για το πρόβλημα (P). Το σύνολο όλων των ολικών βέλτιστων λύσεων του προβλήματος (P) θα συμβολίζεται με X^*. Όταν $v = -\infty$, το πρόβλημα (P) θα καλείται μη φραγμένο. Σημειώστε ότι από τη στιγμή που μια κυρτή συνάρτηση g μπορεί να μεγιστοποιηθεί σε ένα κυρτό σύνολο Y ελαχιστοποιώντας την $-g$ στο Y, η μεγιστοποίηση κυρτών συναρτήσεων μπορεί να θεωρηθεί ως ελαχιστοποίηση κοίλων συναρτήσεων. Το πρόβλημα (P) δηλαδή της ελαχιστοποίησης μιας κοίλης συνάρτησης με κυρτούς περιορισμούς θα αναφέρεται και ως πρόβλημα κοίλης ελαχιστοποίησης.

Υπάρχουν τουλάχιστον τέσσερα χαρακτηριστικά αυτών των προβλημάτων που εξηγούν γιατί έχουν τραβήξει τόση προσοχή αυτά τα προβλήματα:

(i) Εμφανίζουν ενδιαφέρουσες μαθηματικές ιδιότητες που τα καθιστούν πιο εύκολα διαχειρίσιμα από ένα πιο γενικό πρόβλημα βελτιστοποίησης.

(ii) Έχουν πολύ ευρύ πεδίο άμεσων εφαρμογών.

(iii) Πολλά μοντέλα που δεν είναι αρχικά τέτοιου τύπου μπορούνε εύκολα να μετατραπούν σε προβλήματα τα οποία είναι ή λύνονται σαν κοίλης ελαχιστοποίησης.

(iv) Πολλές αλγοριθμικές τεχνικές οι οποίες χρησιμοποιούνται για να λύσουν προβλήματα κοίλης βελτιστοποίησης παίζουν σημαντικό ρόλο και στη λύση άλλων τύπων προβλημάτων γενικής βελτιστοποίησης.

Τα προβλήματα κοίλης βελτιστοποίησης παρουσιάζουν μία εντυπωσιακή ποικιλία πρακτικών εφαρμογών. Η βιβλιογραφία αφθονεί με προβλήματα από τόσο διαφορετικούς κλάδους, όπως τα οικονομικά, η διαχείριση παραγωγής, οι τηλεπικοινωνίες, θέματα σχετικά με τους ηλεκτρολόγους μηχανικούς, τις μεταφορές και τη σχεδίαση υπολογιστών, τα οποία μπορούν να διατυπωθούν ως προβλήματα κοίλης βελτιστοποίησης.

Η ποικιλία των δηλωμένων εφαρμογών του κοίλου προγραμματισμού έχει αυξηθεί συνεχόμενα, κυρίως τα τελευταία χρόνια. Αυτό πιθανώς οφείλεται τουλάχιστον μερικώς στο γεγονός ότι όσο οι αλγόριθμοι που λύνουν προβλήματα κοίλης ελαχιστοποίησης βελτιώνονται σε αποτελεσματικότητα, οι ερευνητές παρακινούνται να διατυπώσουν διάφορα πρακτικά προβλήματα σαν κοίλης ελαχιστοποίησης.

Στην κοίλη ελαχιστοποίηση, τρεις βασικές αλγοριθμικές προσεγγίσεις ξεχωρίζουν. Η πρώτη προσέγγιση είναι η χρησιμοποίηση μιας μεθόδου απαρίθμησης για την αναζήτηση μιας ολικής βέλτιστης λύσης. Αυτή η προσέγγιση εφαρμόζεται μόνο όταν το D είναι πολύεδρο, Οι άλλες δύο προσεγγίσεις μπορούν να χρησιμοποιηθούν είτε όταν το D είναι πολύεδρο είτε όταν είναι ένα γενικό κλειστό κυρτό σύνολο. Αυτές οι δύο προσεγγίσεις είναι η διαδοχική προσεγγιστική μέθοδος και η διαδοχική μέθοδος διαχωρισμού (branch and bound).

1.2.1 Μέθοδοι διαδοχικής προσέγγισης

Μια από τις συνηθισμένες προσεγγίσεις για τη λύση κυρτών και ακεραίων προβλημάτων προγραμματισμού με περιορισμούς είναι με χρήση πιο απλών προσεγγίσεων πολύπλοκων συνόλων, συναρτήσεων και προβλημάτων βελτιστοποίησης, Σε πολλές περιπτώσεις, διαδοχικές βελτιώσεις προσεγγίσεων αυτών των τύπων χρησιμοποιούνται για τη δημιουργία ακολουθιών προβλημάτων βελτιστοποίησης που είναι πιο απλά στη λύση από το αρχικό πρόβλημα. Αυτά τα προβλήματα λύνονται διαδοχικά μέχρι να βρεθεί μια ακριβής ή προσεγγιστική λύση στο αρχικό πρόβλημα. Τέτοιου είδους προσεγγίσεις στα προγραμματιστικά προβλήματα με περιορισμούς ονομάζονται μέθοδοι διαδοχικής προσέγγισης. Από τη

δεκαετία του 1970 τέτοιες μέθοδοι έχουν επίσης χρησιμοποιηθεί στην ολική βελτιστοποίηση και, πιο συγκεκριμένα, στην κοίλη ελαχιστοποίηση.

Θα παρουσιάσουμε τώρα δύο βασικές ιδιότητες αυτών των προβλημάτων (βλέπε [2]).

Ιδιότητα 1. *Αν το $H = \mathbb{R}^n$ ή το H είναι ένα ανοιχτό κυρτό υποσύνολο του $\mathbb{R}^n$, τότε η h είναι μία συνεχής συνάρτηση στο H.*

Ιδιότητα 2. *Υποθέστε ότι το $D \subseteq \mathbb{R}^n$ είναι ένα μη κενό,κλειστό κυρτό σύνολο, και ότι η f είναι μια πραγματική, κοίλη συνάρτηση που ορίζεται σε ένα κατάλληλο κυρτό σύνολο $A \subseteq \mathbb{R}^n$ που περιέχει το D. Ακόμη, υποθέστε ότι το A είτε είναι ίσο με $\mathbb{R}^n$ είτε είναι ένα ανοιχτό κυρτό σύνολο στο $\mathbb{R}^n$, και ότι το D είναι φραγμένο. Τότε το πρόβλημα (P) πρέπει να έχει μία ολική βέλτιστη λύση, η οποία είναι ακραίο σημείο του D.*

Για απλότητα θα υποθέσουμε ότι στο πρόβλημα (P), το D είναι ένα μη κενό, συμπαγές κυρτό σύνολο στο $\mathbb{R}^n$, και ότι το $A \supseteq D$ είναι ένα ανοιχτό κυρτό σύνολο. Υπό αυτές τις υποθέσεις και από τις ιδιότητες 1 και 2 η f είναι συνεχής στο A και το (P) έχει τουλάχιστον μία ολική βέλτιστη λύση η οποία είναι ακραίο σημείο του D.

Από τις πιο δημοφιλείς μεθόδους για να λύνουμε το (P) με διαδοχικές προσεγγίσεις είναι η εξωτερική προσέγγιση. Σε αυτή τη μέθοδο, πιο χαλαρά σύνολα $D_k \supseteq D, \;\; k = 1, 2, \ldots,$ της εφικτής περιοχής D κατασκευάζονται διαδοχικά. Αυτά τα χαλαρά σύνολα καλούνται εξωτερικές προσεγγίσεις του D. Για κάθε εξωτερική προσέγγιση του D_k του D η οποία κατασκευάζεται, το πρόβλημα της προσέγγισης που δίνεται από

$$\min f(x), \text{ subject to } x \in D_k, \tag{P_k}$$

λύνεται για μια βέλτιστη λύση x^k. Τυπικά, το D_k είναι ένα μη κενό, συμπαγές πολύεδρο στο $\mathbb{R}^n$, αλλά πιο γενικές εξωτερικές προσεγγίσεις είναι επιτρεπτές. Αν το $x^k \in D$, τότε το x^k είναι μία βέλτιστη λύση του (P) και η διαδικασία τερματίζει. Διαφορετικά, χρησιμοποιώντας το D_k και το x^k, μία καινούργια εξωτερική προσέγγιση D_{k+1} του D κατασκευάζεται τέτοια ώστε $x^k \notin D_{k+1}$ και $D_{k+1} \subset D_k$, και η διαδικασία επαναλαμβάνεται. Αφού $x^k \notin D_{k+1}$, το x^k λέγεται ότι κόβεται από το D_{k+1}, αλλά αφού $D \subseteq D_{k+1}$, κανένα επιτρεπόμενο σημείο του (P) δεν κόβεται από το D_{k+1}.

Οι κύριες προϋποθέσεις σε καθέναν από τους αλγορίθμους εξωτερικής προσέγγισης είναι για κάθε $k = 1, 2, \ldots,$ δεδομένων των x^k και $D_k,$ το σύνολο D_{k+1} πρέπει να είναι κατασκευασμένο έτσι ώστε $D_{k+1} \subset D_k, \;\; x^k \notin D_{k+1}$ και $D_{k+1} \supseteq D$, και, αν ο αλγόριθμος δε σταματάει, πρέπει να συγκλίνει με κάποια έννοια σε μία ολική βέλτιστη λύση του (P). Υπάρχουν τουλάχιστον δύο τύποι μεθόδων που έχουν προταθεί για την επίτευξη αυτού. Εμείς θα ασχοληθούμε με τη γενική μέθοδο των ανισωτικών "κοψιμάτων". Για κάθε k, προστίθεται μία ανισότητα στον ορισμό του D_k για να οριστεί το D_{k+1}. Αυτές οι ανισότητες ονομάζονται κοψίματα.

Στα τέλη της δεκαετίας του 1980, ο Horst και οι συνεργάτες του δημοσίευσαν δύο σημαντικά papers (βλέπε [3],[4]) με ένα γενικό σχήμα και θεωρία σύγκλισης για τη λύση του (P) μέσω εξωτερικών προσεγγίσεων με χρήση ανισωτικών κοψιμάτων.

Το γενικό αυτό σχήμα μπορεί να παρουσιαστεί απλά ως εξής:

Βήμα 0. Διάλεξε ένα συμπαγές σύνολο $D_1 \subseteq A$ το οποίο περιέχει το D. Θέσε $k = 1$ και πήγαινε στο βήμα k.

Βήμα $k(k = 1, 2, \dots)$. Λύσε το "χαλαρότερο" πρόβλημα (P_k) που δίνεται από

$$\min f(x), \text{ subject to } x \in D_k$$

για μια βέλτιστη λύση x^k. Αν το $x^k \in D$, σταμάτα: το x^k είναι μια ολική βέλτιστη λύση του (P). Διαφορετικά, κατασκεύασε μία συνάρτηση $q_k \colon A \to \mathbb{R}$ τέτοια ώστε

(i) $q_k(x) \leq 0$ for all $x \in D$

και

(ii) $q_k(x^k) > 0$,

θέσε $D_{k+1} = D_k \cap \{x \in A | q_k(x) \leq 0\}$, θέσε $k = k + 1$, και πήγαινε στο βήμα k.

Είναι αμέσως εμφανές στο σχήμα ότι για κάθε
$k = 1, 2, \dots, \;\; D_{k+1} \subset D_k, \;\; x^k \notin D_{k+1},$ και $D_{k+1} \supseteq D$ όπως απαιτείται. Επίσης, αν οι συναρτήσεις $q_k, \;\; ka = 1, 2, \dots,$ είναι κατασκευασμένες κατάλληλα, τότε η σύγκλιση εξασφαλίζεται από το παρακάτω θεώρημα (βλέπε [5]).

Θεώρημα 1
Αν το παραπάνω σχήμα γενικής εξωτερικής προσέγγισης τερματίζει στο βήμα k για κάποιο $k = 1, 2, \dots,$ τότε το x^k είναι μια ολική βέλτιστη λύση για το πρόβλημα (P). Διαφορετικά, αν η q_k είναι χαμηλότερη ημι-συνεχής στο για κάθε $k = 1, 2, \dots,$ και για κάθε συγκλίνουσα υποακολουθία $\left\{x^{k_w}\right\}$ των $\left\{x^k\right\}$ που ικανοποιούν τη σχέση

$$\lim_i q_{k_{w_i}}(x^{k_w}) = \lim_i q_{k_{w_i}}(\bar{x}),$$

όπου το $\bar{x} \in D$ όποτε $\lim_i q_{k_{w_i}}(\bar{x}) = 0$, τότε οποιοδήποτε σημείο συσσώρευσης της ακολουθίας $\left\{x^k\right\}$ είναι μια ολική βέλτιστη λύση για το πρόβλημα (P).

Οι Horst, Thoai και Tuy (βλέπε [4]) έδειξαν ότι οι περισσότεροι αλγόριθμοι εξωτερικής προσέγγισης που προτείνονται για το πρόβλημα (P) και χρησιμοποιούν ανισωτικά κοψίματα μπορούν να θεωρηθούν ειδικές περιπτώσεις του γενικού σχήματος που πρότειναν, και μπορεί να δειχθεί ότι συγκλίνουν μέσω του θεωρήματος 1. Στην πράξη σε όλους αυτούς τους αλγορίθμους, οι εξωτερικές προσεγγίσεις $D_k, \;\; k = 1, 2, \dots,$ είναι κυρτά πολύεδρα, και οι συναρτήσεις $q_k, \;\; k = 1, 2, \dots,$ είναι γραμμικές ανισότητες. Συνεπώς οι εξισώσεις $q_k = 0, \;\; k = 1, 2, \dots,$ σε αυτές τις μεθόδους ορίζουν υπερεπίπεδα τα οποία χωρίζουν αυστηρά το x_k από το D. Η εφικτή περιοχή D σε αυτές τις μεθόδους θεωρείται ότι έχει τη μορφή

$$D = \{x \in A | g(x) \leq 0\},$$

όπου η $g\colon A \to \mathbb{R}$ είναι μια κυρτή συνάρτηση. Παρατηρούμε ότι οι εφικτές περιοχές της μορφής $\{x \in A | g_i(x) \leq 0, \ \ i = 1, 2, \ldots, m\}$, όπου για κάθε $i = 1, 2, \ldots, m,$ η $g_i\colon A \to \mathbb{R}$ είναι μια κυρτή συνάρτηση, μπορούν να αναδιατυπωθούν στη μορφή της D θέτοντας $g(x) = \min\{g_i(x) | i = 1, 2, \ldots, m\}$ για κάθε $x \in A$.

Για να δείξουν ότι το γενικό σχήμα και το αποτέλεσμα σύγκλισης υπάγονται στην πληθώρα των αλγορίθμων εξωτερικής προσέγγισης με κοψίματα για το πρόβλημα (P), οι Horst, Thoai και Tuy (βλέπε [4]) έδωσαν μία γενική φόρμουλα για την επιλογή καταλλήλων γραμμικών συναρτήσεων για να ορίσουν τις $q_k, \ \ k = 1, 2, \ldots$. Για να παρουσιάσουμε αυτή τη φόρμουλα, θεωρούμε ότι το D έχει την παραπάνω μορφή και ότι το σύνολο D^I που δίνεται από

$$D^I = \{x \in A | g(x) < 0\}$$

είναι μη κενό. Επίσης, έστω ότι το K είναι ένα οποιοδήποτε κυρτό σύνολο του D^I. Παρατηρούμε ότι εφόσον η g είναι κυρτή στο $A, \ \ \partial g(x) \neq$ σε κάθε $x \in A$. Η φόρμουλα για κάθε $q_k\colon A \to \mathbb{R}$ μπορεί να γραφεί ως

$$q_k(x) = \langle p(y^k), x - y^k \rangle + g(y^k), \tag{1.7}$$

όπου y^k είναι οποιοσδήποτε κυρτός συνδυασμός των σημείων στο $K \cup \{x^k\}$ ο οποίος δε βρίσκεται στο D^I, και $p(y^k) \in \partial g(y^k)$. Θυμηθείτε ότι για κάθε $k,$ το x^k ελαχιστοποιεί την f στο D_k στο γενικό σχήμα της εξωτερικής προσέγγισης.

Έστω $k \in \{1, 2, \ldots\}$. Παρατηρούμε ότι το $p(y^k) \in \partial g(y^k)$ θα είναι διάφορο του μηδενός με την ακόλουθη λογική. Για κάθε $w \in D^I$ και $p(y^k) \in \partial g(y^k)$, γνωρίζουμε ότι εφόσον η g είναι κυρτή συνάρτηση στο A

$$0 > g(w) \geq \langle p(y^k), w - y^k \rangle + g(y^k),$$

όπου οι δύο ανισώσεις προκύπτουν από τον ορισμό της D^I και $\partial g(y^k)$ αντίστοιχα. Αφού $y^k \notin D^I, \ \ g(y^k) \geq 0$. Μαζί με την (1) προκύπτει ότι $p(y^k) \neq 0$.

Συνεπάγεται ότι αν χρησιμοποιηθεί η (1) για την παραγωγή των γραμμικών περιορισμών q_k, τότε η $q_k(x) = 0$ ορίζει ένα μη τετριμμένο υπερεπίπεδο. Επιπροσθέτως, είναι εύκολο να δείξουμε ότι αυτό το υπερεπίπεδο χωρίζει αυστηρά το x^k από το D όπως απαιτείται στα βήματα k(i) και k(ii) του γενικού σχήματος που έχει δοθεί παραπάνω. Τελικά, με χρήση του θεωρήματος 1, έδειξαν ότι αν για κάθε $k = 1, 2, \ldots,$ η φόρμουλα (25) χρησιμοποιείται για τον ορισμό των q_k στο σχήμα, τότε το σχήμα είτε λύνει το πρόβλημα (P) μετά από πεπερεασμένο αριθμό βημάτων, είτε, αν όχι, κάθε σημείο συσσώρευσης της ακολουθίας των $\{x^k\}$ που παράγεται είναι μία ολική βέλτιστη λύση.

Κεφάλαιο 2

Διατύπωση προβλήματος και αλγόριθμος επίλυσης

2.1 Διατύπωση προβλήματος

Έστω ότι έχουμε το σύστημα εξισώσεων:

$$f_i(x) = 0, \;\; i = 1, \ldots, n, \;\; x \in \mathbb{R}^n, \;\; f_i \text{ κυρτές συναρτήσεις.}$$

Θέλουμε να βρούμε τα $x \in \mathbb{R}^n$ τα οποία ικανοποιούν το σύστημα. Μπορούμε να προσεγγίσουμε τη λύση αυτού του προβλήματος ως ένα πρόβλημα κοίλης ελαχιστοποίησης, δηλαδή ελαχιστοποίησης μιας κοίλης συνάρτησης με κυρτούς περιορισμούς. Αυτό θα πραγματοποιηθεί ως εξής: θεωρούμε τη συνάρτηση $h(x) = \sum_{i=1}^{n} f_i(x)$, καθώς και το σύνολο $G = \{x \in \mathbb{R}^n \colon f_i(x) \leq 0, \;\; i = 1, \ldots, n\}$. Τα $x \in \mathbb{R}^n$ που ικανοποιούν το σύστημα εξισώσεων, ικανοποιούν και το πρόβλημα μεγιστοποίησης της h στο G,

$$\max_{x \in \mathbb{R}^n} \{h(x) \colon f_i(x) \leq 0, \;\; i = 1, \ldots, n\}.$$

καθώς εύκολα φαίνεται ότι $h(x) \leq 0, \;\; \forall x \in G$.

Άρα θα ικανοποιούν και το πρόβλημα ελαχιστοποίησης της $-h$ στο G:

$$\min_{x \in \mathbb{R}^n} \{-h(x) \colon f_i(x) \leq 0, \;\; i = 1, \ldots, n\}.$$

Η h είναι κυρτή συνάρτηση ως άθροισμα κυρτών συναρτήσεων, συνεπώς η $-h$ θα είναι κοίλη συνάρτηση. Συνεπώς έχουμε να λύσουμε ένα πρόβλημα κοίλης ελαχιστοποίησης το οποίο εφόσον έχει λύση και αυτή είναι ίση με το 0 θα έχουμε και τη λύση του αρχικού προβλήματος του συστήματος εξισώσεων.

2.2 Αλγόριθμος

Ο αλγόριθμος της Karla Leigh Hoffman (βλέπε [6]) που θα παρουσιαστεί λύνει το πρόβλημα της κοίλης ελαχιστοποίησης.Έστω το πρόβλημα:

$$P: \qquad \text{minimize } f(x), \quad \text{subject to } g_i(x) \leq 0,\ i = 1, \ldots, m$$

όπου η f είναι μια κοίλη συνάρτηση που ορίζεται στο $\mathbb{R}^n$ και g_i $(i = 1, \ldots, m)$ είναι κυρτές συναρτήσεις με συνεχείς παραγώγους. Θεωρούμε ότι υπάρχει ένα σημείο p αυστηρά στο εσωτερικό της εφικτής περιοχής, δηλαδή
$p \in \{x\colon g_i < 0,\ \ i = 1, \ldots, m\}$. Έστω $S = \{x\colon g_i \leq 0,\ \ i = 1, \ldots, m\}$. Θα θεωρήσουμε ότι το S είναι μη κενό και συμπαγές σύνολο.

Διάλεξε $\epsilon > 0$ αρκετά μικρό.
Βήμα 0: Βρες ένα σημείο p στο εσωτερικό του S.
Θέσε $u^0 = f(p)$ (το καλύτερο άνω άκρο μέχρι στιγμής).
Βρες ένα γραμμικό πολύεδρο S^0 που να περιέχει το σύνολο S, τέτοιο ώστε όλες οι κορυφές του S^0 να είναι γνωστές, και να μπορούν να περιγραφούν ως ένα σύνολο από περιορισμούς $Ax \leq b$.
Έστω V^0 =το σύνολο των κορυφών του S^0.
Θέσε $k = 1$.
Πήγαινε στο βήμα 1.
Βήμα k: (Αυτό το βήμα παίρνει ως είσοδο ένα πολύεδρο S^{k-1} που περικλείει το S, το σύνολο V^{k-1}, τα ακραία σημεία του S^{k-1}, και το τωρινό άνω άκρο u^{k-1} για το πρόβλημα P.)
Λύσε το πρόβλημα

$$P^k: \qquad \text{minimize f(x)}, \quad \text{subject to } x \in S^{k-1}$$

επιλέγοντας $min\left\{f(v)\colon v \in V^{k-1}\right\}$.
Έστω ότι το x^k λύνει το πρόβλημα P^k.
Τώρα έστω ότι το λ_k λύνει το μονοδιάστατο πρόβλημα βελτιστοποίησης

$$Q^k: \qquad \text{minimize } \lambda, \quad \text{subject to } 0 \leq \lambda \leq 1,\ \ x^k + \lambda(p - x^k) \in S$$

Αν $\lambda_k = 0$, τότε $x^k \in S$ και επομένως είναι λύση του προβλήματος P; σταμάτα.

Αν $\lambda_k > 0$, θέσε $z^k = x^k + \lambda_k(p - x^k)$. Αν $f(z^k) < u^{k-1}$, θέσε $u^k = f(z^k)$; αλλιώς $u^k = u^{k-1}$. Αν $z^k - x^k < \epsilon$, (ή παρομοίως $u^k - f(x^k) < \epsilon$) σταμάτα; το x^k είναι η

λύση. Αν όχι, βρες όλους τους περιορισμούς του S που ικανοποιούνται από το z^k και θέσε

$$J^k = \left\{ i\colon \quad g_i(z^k) = 0, \quad i = 1, \ldots, m \right\},$$

και διάλεξε κάποιο $j^k \in J^k$.
Στους περιορισμούς του S^{k-1}, πρόσθεσε τον περιορισμό

$$G^k_{j^k}(x) = g_{j^k}(z^k) + \left[\nabla g_{j^k}(z^k) \right]^T (x - z^k) \leq 0.$$

Θέσε

$$S^k = S^{k-1} \cap \left\{ x\colon G^k_{j^k}(x) \leq 0 \right\}, \qquad V^k = \left\{ \text{ακραία σημεία του } S^k \right\}.$$

Θέσε $k = k + 1$.
Επίστρεψε στο βήμα k.
Προσαρμόζοντας τον αλγόριθμο της κοίλης ελαχιστοποίησηςστο αρχικό πρόβλημα της λύσης του συστήματος:

$$f_i(x) = 0, \quad i = 1, \ldots, n, \quad x \in \mathbb{R}^n, \quad f_i \text{ κυρτές συναρτήσεις.}$$

θέτουμε $g_i = f_i$ και $f(x) = -\sum_{i=0}^{n} f_i(x)$.

Αν υπάρχει p στο εσωτερικό της εφικτής περιοχής, και ο αλγόριθμος σταματάει για κάποιο k^*, τότε αν $f(x^{k^*}) = 0$ το σύστημα έχει λύση το x^{k^*}. Διαφορετικά το σύστημα δεν έχει λύση.

Αν δεν υπάρχει p αυστηρά στο εσωτερικό της εφικτής περιοχής, τότε αν το p που θα βρούμε είναι στα όρια όλων των περιορισμών, δηλαδή τους ικανοποιεί όλους είναι λύση του συστήματος. Διαφορετικά το σύστημα δεν έχει λύση.

Σε μια προσπάθεια να βρούμε όλες τις λύσεις της ελαχιστοποίησης αφήνουμε τον αλγόριθμο να συνεχίσει την αναζήτησή και έπειτα από την εύρεση ενός ολικού ελαχίστου. Σε κάθε βήμα αφαιρούμε (κατά τη διάρκεια του βήματος της αναζήτησης της ελάχιστης κορυφής) από το σύνολο των V^k τις κορυφές εκείνες που γνωρίζουμε ήδη ότι λύνουν το πρόβλημα της ελαχιστοποίησης, ή βρίσκονται σε πολύ κοντινή ευκλείδεια απόσταση. Μετά την εύρεση του πρώτου ολικού ελαχίστου κάνουμε έναν αριθμό επιπλέον βημάτων και μετά ο αλγόριθμος τερματίζει.

2.2.1 Υπολογιστικές διεργασίες

Για να βρούμε ένα πολύεδρο που να περικλείει την περιοχή S των εφικτών λύσεων, η πιο απλή προσέγγιση θα ήταν να ζητάμε ανώτερα και κατώτερα όρια για κάθε μεταβλητή. Όμως επειδή η αποδοτικότητα αυξάνει όσο πιο "σφιχτό" είναι το

πολύεδρο, θα εξασφαλίσουμε ένα "σφιχτό" πολύεδρο S^0 λύνοντας τα $2n$ κυρτά προβλήματα:

$$\begin{aligned}
&\min x_j, \\
&\text{s.t. } g_i(x) \le 0; \quad i = 1, \dots, m; \\
&\qquad\qquad\qquad\qquad\qquad\qquad (j = 1, \dots, n)\,. \\
&\max x_j, \\
&\text{s.t. } g_i(x) \le 0; \quad i = 1, \dots, m;
\end{aligned}$$

,όπου αντί για μεγιστοποίηση για κάθε x_j θα λύσουμε το πρόβλημα της ελαχιστοποίησης των $-x_j$.

Για να βρούμε το εσωτερικό σημείο p που απαιτείται για να εκκινήσει ο αλγόριθμός μας θα λύσουμε το κυρτό πρόβλημα:

$$\min_{x,\varepsilon} \{\varepsilon \colon f_i(x) \le \varepsilon, \quad i = 1, \dots, n\}$$

Παίρνουμε το p ίσο με το x που προέκυψε από αυτή την ελαχιστοποίηση.

Το πιο απαιτητικό υπολογιστικά κομμάτι του αλγορίθμου περιλαμβάνει την εύρεση των καινούργιων ακμών του πολυέδρου που δημιουργείται με την προσθήκη του επιπλέον υπερεπιπέδου (περιορισμού της μορφής $Ax \le b$). Αυτό πραγματοποιείται με τον ακόλουθο τρόπο: Διάλεξε ένα j^k από το J^k.

(i) αν v^t είναι μια κορυφή του S^{k-1} η οποία ικανοποιεί τους περιορισμούς $G^k_{j^k}(x) \le 0$, τότε η v^t είναι μία κορυφή του S^k, δηλαδή $v^t \in V^k$;

(ii) αν v^t είναι μια κορυφή του S^{k-1} η οποία ικανοποιεί την $G^k_{j^k}(x) > 0$ για κάποιο $j^k \in J^k$, τότε βρες όλους τους γείτονες της v^t στο πολύεδρο $\bar{S}^k = \left\{x \colon G^k_{j^k}(x) \ge 0\right\} \cap S^k$; αν ένας τέτοιος γείτονας ικανοποιεί την $G^k_{j^k}(x) = 0$, τότε είναι στοιχείο του V^k, διαφορετικά δεν είναι.

Συνοψίζοντας, οι κορυφές του S^{k+1} είτε είναι κορυφές του S^k είτε είναι γείτονες στο $\bar{S}^{k+1}$ των κορυφών v^t του S^k.

Κεφάλαιο 3

Παραδείγματα

3.1 Τετραγωνικές συναρτήσεις δύο διαστάσεων

Σαν πρώτο παράδειγμα θα δούμε κάτι πολύ απλό, τη λύση συστήματος τετραγωνικών συναρτήσεων στις δύο διαστάσεις. Με αυτό το απλό παράδειγμα όμως μας δίνεται η δυνατότητα να αυξήσουμε τη διαισθητική κατανόηση της ορθότητας του αλγορίθμου που χρησιμοποιούμε και να επιβεβαιώσουμε την ολική σύγκλισή του, καθώς το σύστημα που έχουμε επιλέξει έχει δύο ολικά και ένα τοπικό ελάχιστο. Το σύστημα των εξισώσεων είναι το ακόλουθο:

$$\frac{1}{2}x^T Q_i x + b_i^T x + a_i = 0, x \in \mathbb{R}^n, \;\; i = 1, 2.$$

$$\text{όπου } Q_1 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \; Q_2 = \begin{pmatrix} 2 & 0 \\ 0 & 40 \end{pmatrix}, \; b_1 = \begin{pmatrix} 0 \\ -8 \end{pmatrix},$$

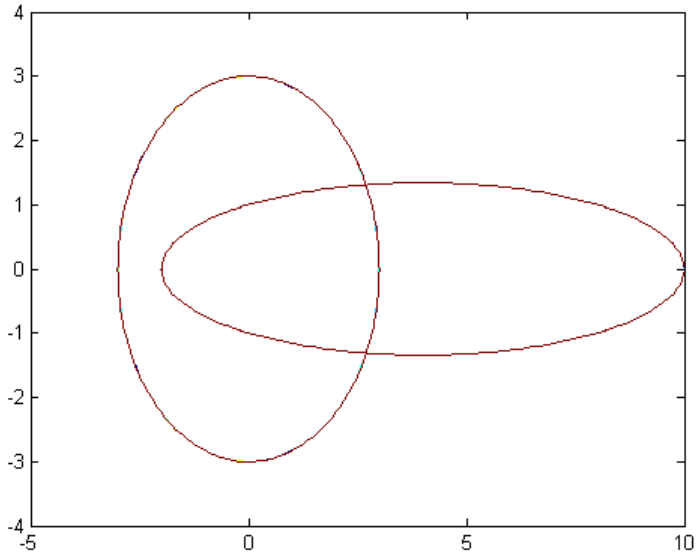
$$b_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \;\; a_1 = -9 \text{ και } a_2 = -20.$$

Ορίζουμε τη συνάρτηση $f(x) = x^T \hat{Q} x + \hat{b}^T x + \hat{a}$,
όπου $\hat{Q} = Q_1 + Q_2, \;\; \hat{b} = b_1 + b_2, \;\; \hat{a} = a_1 + a_2$. Ορίζουμε και το σύνολο
$F = \left\{ x \in \mathbb{R}^n \colon \frac{1}{2}x^T Q_i x + b_i^T x + a_i \leq 0, \;\; i = 1, 2 \right\}$

Η f κυρτή ως άθροισμα κυρτών συναρτήσεων (είναι προφανές ότι οι Q_i είναι θετικά ορισμένες) και $f(x) \leq 0, \;\; \forall x \in F$. Άρα η $-f$ κοίλη.

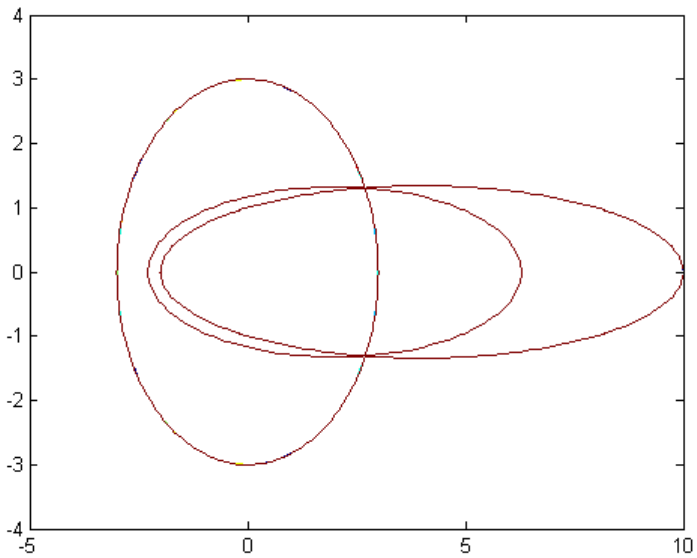
Οπότε λύνουμε το πρόβλημα κοίλης ελαχιστοποίησης:

$$min \left\{ f(x) \colon x \in F, \;\; x \in \mathbb{R}^2 \right\}$$

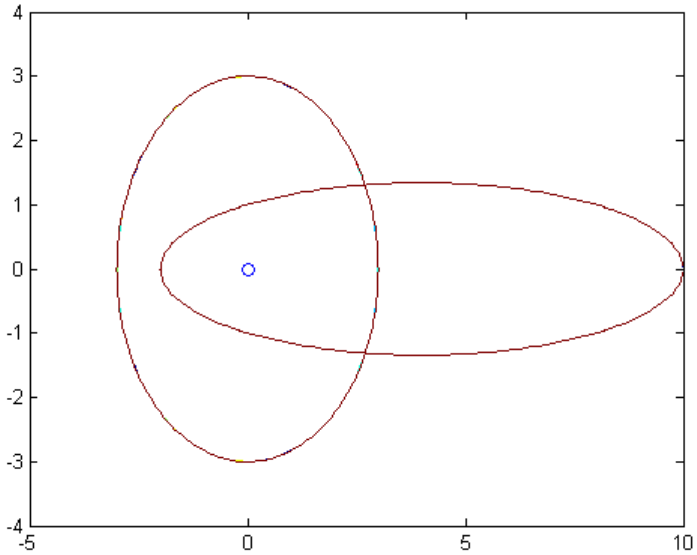


Σχήμα 3.1: Οι τετραγωνικές εξισώσεις $\frac{1}{2}x^T Q_i x + b_i^T x + a_i = 0, \;\; i = 1, 2.$

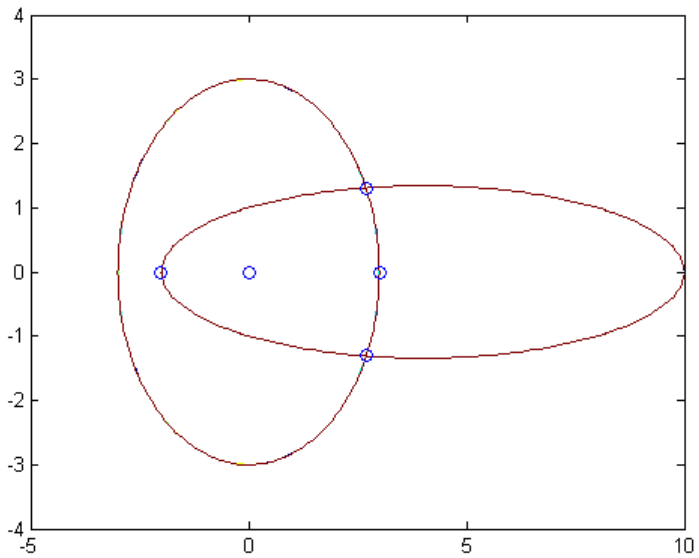
Στα σημεία τομής των δύο ελλείψεων ικανοποιείται το σύστημα εξισώσεων. Αυτό επιβεβαιώνεται και στο παρακάτω σχήμα.



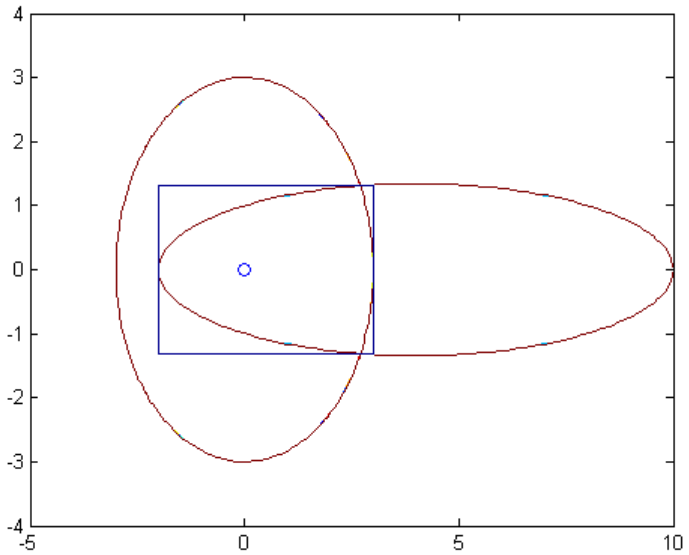
Σχήμα 3.2: Οι τετραγωνικές εξισώσεις και η $f(x) = 0$.



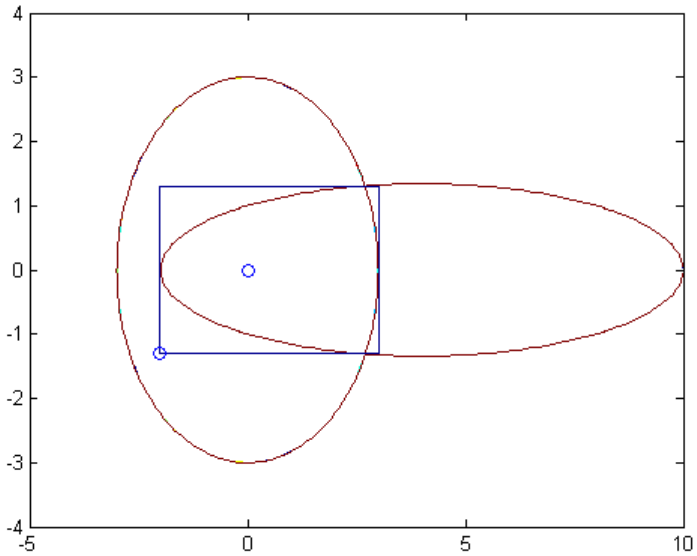
Σχήμα 3.3: Το εσωτερικό σημείο p που προκύπτει από την ελαχιστοποίηση $\min_{x,\varepsilon} \left\{ \varepsilon \colon \frac{1}{2} x^T Q_i x + b_i^T x + a_i \leq \varepsilon, \;\; i = 1, 2 \right\}$.



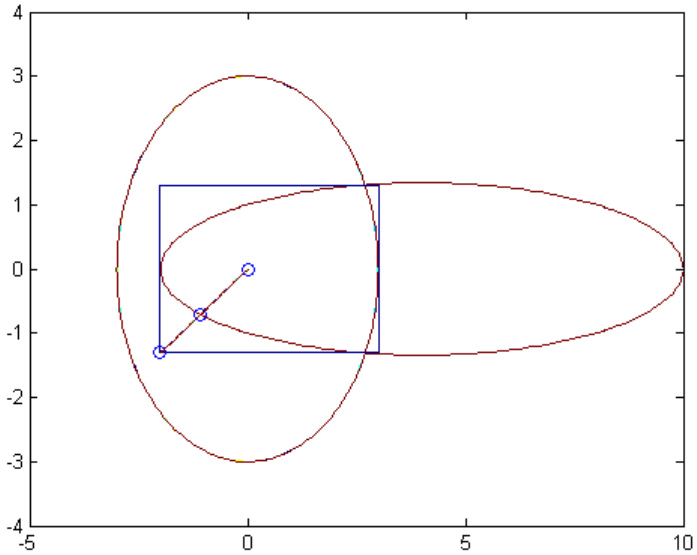
Σχήμα 3.4: Η ελάχιστη και η μέγιστη τιμή της κάθε διάστασης του αρχικού πολυέδρου S_0 που απαιτείται για να ξεκινήσει ο αλγόριθμος.



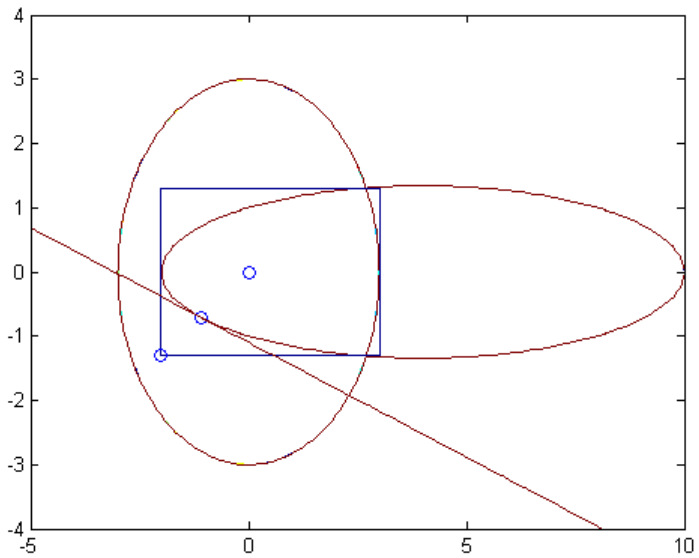
Σχήμα 3.5: Το αρχικό πολύεδρο S_0 που επιλέξαμε να περικλείει το σύνολο F των περιορισμών.



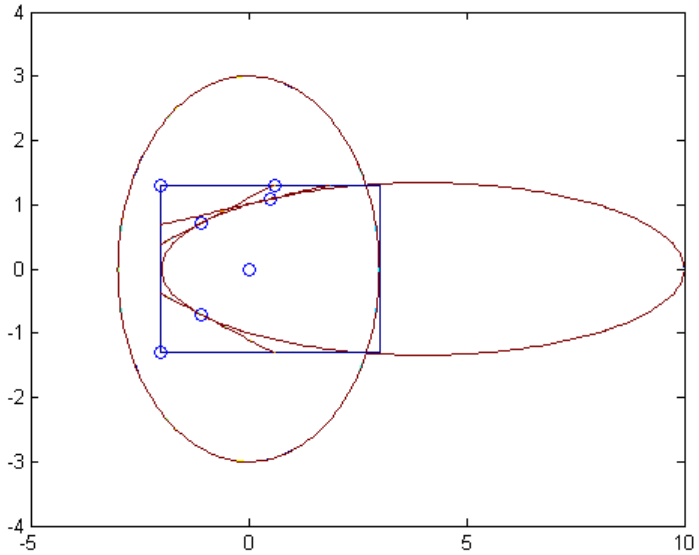
Σχήμα 3.6: Εξετάζονται όλες οι κορυφές του πολυέδρου και βρίσκουμε τη x_1 με την ελάχιστη τιμή της αντικειμενικής συνάρτησης.



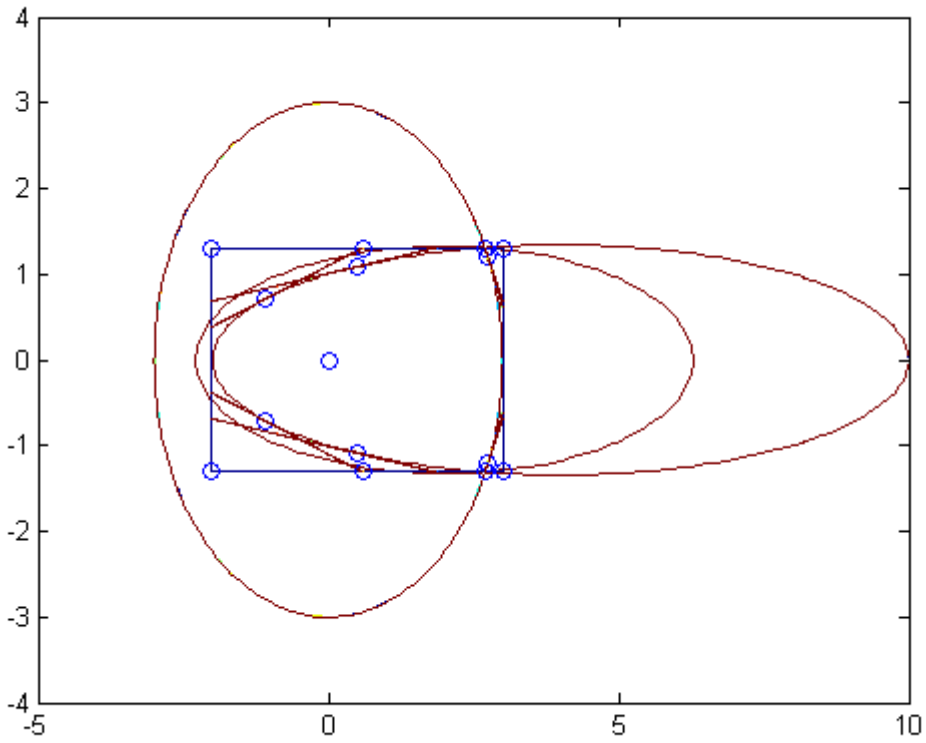
Σχήμα 3.7: Με γραμμική αναζήτηση βρίσκουμε που τέμνει το κυρτό σύνολο των περιορισμών, στο z_1.



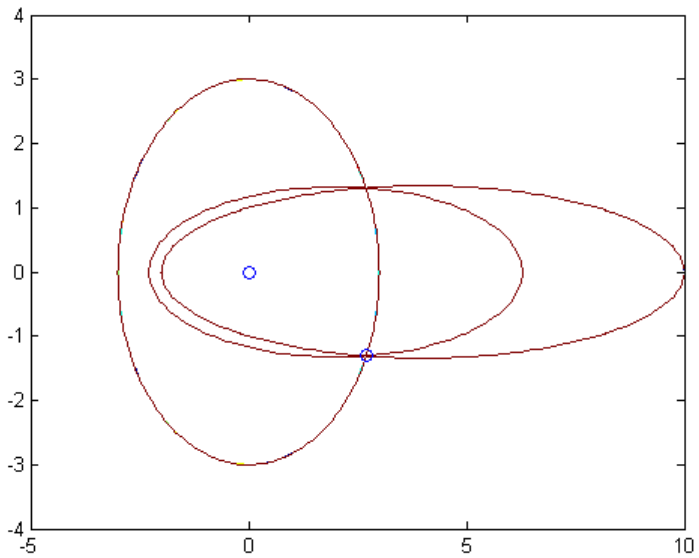
Σχήμα 3.8: Φέρνουμε την κάθετη στο σύνολο στο z_1. Έχουμε τώρα το νέο πολύεδρο S_1 που περικλείει το σύνολο F.



Σχήμα 3.9: Μετά την προσθήκη 3 περιορισμών έχουμε το S_3.



Σχήμα 3.10: Όταν τερματίζει ο αλγόριθμός μας έπειτα από 11 επαναλήψεις λαμβάνουμε το παραπάνω σχήμα.



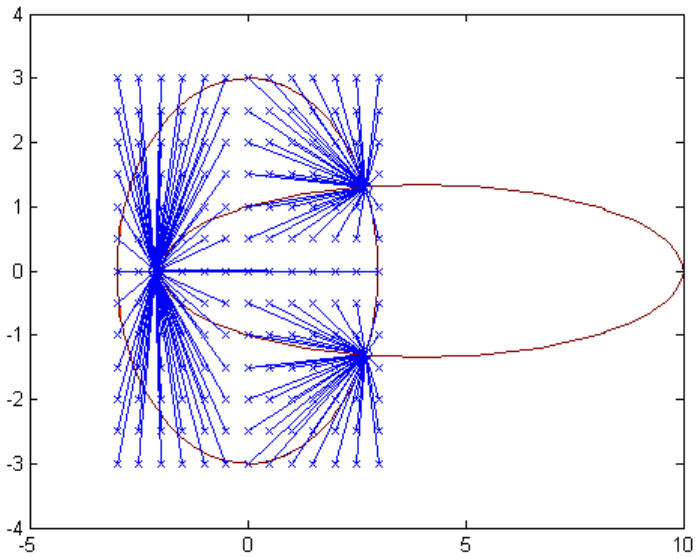
Σχήμα 3.11: Συγκλίνων σημείο του αλγορίθμου.

Πίνακας 3.1: Οι κορυφές που επιλέχθηκαν και οι αντίστοιχες τιμές της αντικειμενικής συνάρτησης φαίνονται στον παρακάτω πίνακα:

Επανάληψη	$f(x)$	$x(1)$	$x(2)$
1	-31,0228	-2	-1,30972
2	-31,0228	-2	1,309722
3	-3,01631	0,586931	1,309722
4	-3,01631	0,586931	-1,30972
5	-1,02278	3	1,309722
6	-1,02278	3	-1,30972
7	-0,00739	2,701644	1,309722
8	-0,00739	2,701644	-1,30972
9	-7,15E-07	2,699006	1,309722
10	-7,15E-07	2,699006	-1,30972
11	-2,87E-08	2,699005	-1,30972

3.1.1 Μέθοδος Newton

Για το ίδιο σύστημα με πριν, εφαρμόζουμε τη μέθοδο Newton όπως την περιγράψαμε στην εισαγωγή. Για τη λύση του χρησιμοποιήσαμε διάφορες αρχικές συνθήκες. Φαίνεται ξεκάθαρα ότι χρειάζονται πολλές αρχικές συνθήκες για την εύρεση της λύσης, καθώς η μέθοδος πολλές φορές συγκλίνει στο τοπικό ελάχιστο. Ξεκάθαρο πλεονέκτημα επίσης της μεθόδου που χρησιμοποιούμε είναι ότι σε κάθε περίπτωση όταν σταματήσει ο αλγόριθμος γνωρίζουμε την ύπαρξη ή μη, λύσης, και την τιμή της.



Σχήμα 3.12: Με το x φαίνονται οι αρχικές συνθήκες και η γραμμή δείχνει προς τα που συγκλίνει ο αλγόριθμος.

3.1.2 Μη ύπαρξη λύσης

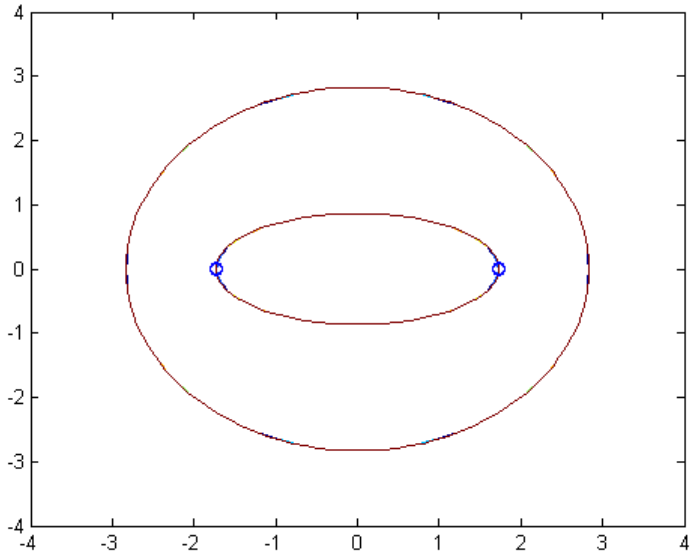
Στην περίπτωση μη ύπαρξης λύσης δεν υπάρχει τομή των άκρων των συνόλων των περιορισμών στο ίδιο σημείο. Όσον αφορά το πως φαίνεται αυτό στον αλγόριθμο έχουμε δύο πιθανές περιπτώσεις. Είτε δεν υπάρχει κάποιο εσωτερικό σημείο στο σύνολο των περιορισμών, οπότε ο αλγόριθμος τερματίζει κατευθείαν και γνωρίζουμε ότι δεν υπάρχει λύση στο πρόβλημα, είτε υπάρχουν πλεονάζοντες περιορισμοί, οπότε ο αλγόριθμος μας θα τερματίσει κανονικά αλλά το ελάχιστο που θα βρεθεί θα είναι διάφορο του μηδενός (θετικό).

Ένα απλό παράδειγμα για τη δεύτερη περίπτωση είναι το ακόλουθο:

$$\frac{1}{2}x^T Q_i x + b_i^T x + a_i = 0, x \in \mathbb{R}^2, \quad i = 1, 2.$$

$$\text{όπου } Q_1 = \begin{pmatrix} 2 & 0 \\ 0 & 8 \end{pmatrix}, \quad Q_2 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix},$$

$$b_1 = b_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad a_1 = -4 \text{ και } a_2 = -9.$$



Σχήμα 3.13: Οι λύσεις που βρίσκει ο αλγόριθμος είναι τα ολικά ελάχιστα αλλά δεν ικανοποιούν το σύστημά μας καθώς $f(solution) = 5.00038 > 0$.

3.1.3 Ειδική περίπτωση

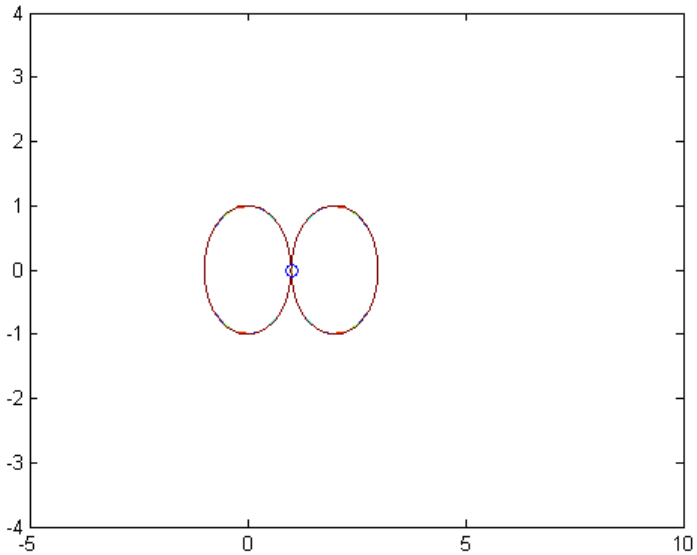
Στην ειδική περίπτωση που η τομή των συνόλων είναι ένα σημείο, η λύση του συστήματος τελειώνει άμεσα με την εύρεση του εσωτερικού σημείου. Για να επιτευχθεί αυτό ο αλγόριθμός μας έχει τροποποιηθεί κατάλληλα ώστε να ελέγχει αν το εσωτερικό σημείο που βρέθηκε ικανοποιεί το σύστημα. Το σύστημά μας σε αυτή

την περίπτωση είναι το παρακάτω:

$$\frac{1}{2}x^T Q_i x + b_i^T x + a_i = 0, x \in \mathbb{R}^2, \quad i = 1, 2.$$

όπου $Q_1 = Q_2 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \quad b_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$

$b_2 = \begin{pmatrix} -4 \\ 0 \end{pmatrix}, \quad a_1 = -1$ και $a_2 = 3$.



Σχήμα 3.14: Το εσωτερικό σημείο που βρίσκει ο αλγόριθμος αποτελεί και λύση του συστήματος!

3.2 Άθροισμα εκθετικών και τετραγωνικών συναρτήσεων

3.2.1 Παράδειγμα σε δύο διαστάσεις

Θα δούμε τώρα ένα παράδειγμα με συναρτήσεις της μορφής

$$e^{\alpha_i^T x} + \frac{1}{2}x^T Q_i x + b_i^T x + c_i = 0, x \in \mathbb{R}^2, \quad i = 1, 2$$

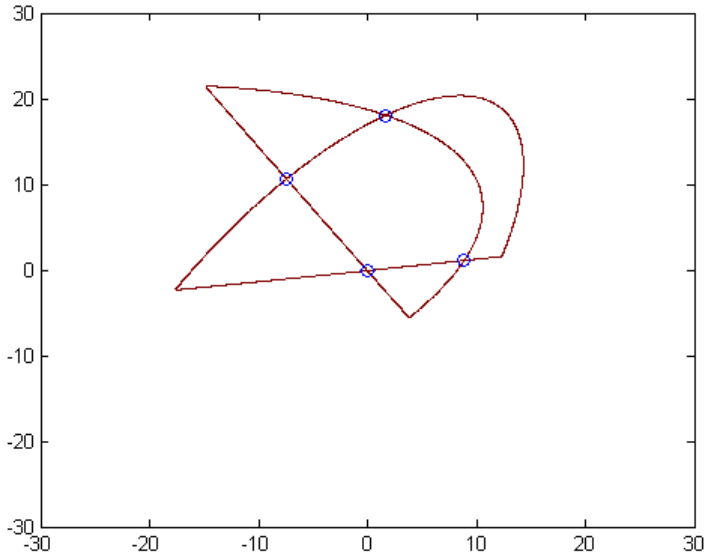
όπου $a_i, b_i \in \mathbb{R}^2, \quad c_i \in (-\infty, 0),$ και Q_i θετικά ορισμένη 2x2 μήτρα. Οι εξισώσεις μας είναι κυρτές ως άθροισμα κυρτών συναρτήσεων.

$$\text{με } Q_1 = \begin{pmatrix} 7.9702 & -4.1622 \\ -4.1622 & 4.0641 \end{pmatrix}, \; Q_2 = \begin{pmatrix} 1.9021 & -2.0598 \\ -2.0598 & 8.8718 \end{pmatrix}, \; b_1 = \begin{pmatrix} 17.0528 \\ 9.9447 \end{pmatrix},$$

$$b_2 = \begin{pmatrix} 83.4387 \\ -42.8322 \end{pmatrix}, \; c_1 = -757.2002, \; c_2 = -753.7291,$$

$$a_1 = \begin{pmatrix} 9.4431 \\ -72.2751 \end{pmatrix} \text{ και } a_2 = \begin{pmatrix} -70.1412 \\ -48.4983 \end{pmatrix}$$

Εδώ παρουσιάζει ιδιαίτερο ενδιαφέρον, το γεγονός ότι το σύστημα έχει 4 λύσεις, και έχουμε την ευκαιρία να εξετάσουμε αν δουλεύει η παραλλαγή του αλγορίθμου μας. Στη συγκεκριμένη περίπτωση ο αλγόριθμος δουλεύει ικανοποιητικά.



Σχήμα 3.15: Φαίνονται καθαρά οι 4 λύσεις του συστήματος.

Πίνακας 3.2: Οι κορυφές που επιλέχθηκαν και οι αντίστοιχες τιμές της αντικειμενικής συνάρτησης φαίνονται στον παρακάτω πίνακα:

Επανάληψη	$f(x)$	$x(1)$	$x(2)$
1	-4,34e+228	-7,44094	-0,09547
2	-2,40e+46	10,57844	-0,09547
3	-1762,94	-7,44094	10,59729
4	-474,305	10,57844	1,296508
5	-432,427	10,57844	18,06711
6	-368,172	-7,44094	18,06711
7	-204,999	8,86158	1,071158
8	-42,0632	4,162357	18,06711
9	-22,9329	-7,44094	10,61646
10	-1,88974	0,437957	18,06711
11	-1,58537	1,935971	18,06711
12	-0,88275	-0,02843	-0,09547
13	-0,45954	8,865414	1,078502
14	-0,01627	1,668765	18,06711
15	-0,01595	-7,44094	10,61688
16	-0,00725	8,865425	1,078523
17	-3,00e-06	8,865345	1,078512
18	-2,88e-06	1,66518	18,06711
19	-1,26e-07	-0,02841	-0,09547
20	-7,87e-09	-7,44094	10,61688

3.2.2 Παράδειγμα σε μεγαλύτερες διαστάσεις

Πάλι θα έχουμε κυρτές συναρτήσεις της μορφής:

$$e^{\alpha_i^T x} + \frac{1}{2}x^T Q_i x + b_i^T x + c_i = 0, x \in \mathbb{R}^n, \quad i = 1, 2, \ldots, n.$$

Εδώ επιλέγουμε $n = 5$, με

$$Q_1 = \begin{pmatrix} 0.8830 & -11.893 & 0.9687 & 11.002 & -24.357 \\ -11.893 & 69.120 & 10.159 & 32.799 & 11.300 \\ 0.9687 & 10.159 & 49.116 & 43.199 & -47.556 \\ 11.002 & 3,28E+04 & 43.199 & 67.941 & -55.865 \\ -24.357 & 1,13E+04 & -47.556 & -55.865 & 99.362 \end{pmatrix},$$

$$Q_2 = \begin{pmatrix} 27.120 & -30.458 & 23.559 & -25.409 & 41.245 \\ -30.458 & 80.642 & -66.712 & 0.1062 & -99.188 \\ 23.559 & -66.712 & 114.201 & -0.9064 & 102.680 \\ -25.409 & 0.1062 & -0.9064 & 60.969 & -35.247 \\ 41.245 & -9,92E+04 & 102.680 & -35.247 & 207.366 \end{pmatrix},$$

$$Q_3 = \begin{pmatrix} 22.845 & 31.009 & 30.191 & -22.088 & 10.289 \\ 31.009 & 68.364 & 76.959 & -52.857 & -0.3727 \\ 30.191 & 76.959 & 109.178 & -62.260 & 0.1194 \\ -22.088 & -5,29E+04 & -62.260 & 99.447 & 0.7678 \\ 10.289 & -0.3727 & 0.1194 & 0.7678 & 66.557 \end{pmatrix},$$

$$Q_4 = \begin{pmatrix} 54.796 & 60.795 & 66.397 & -43.255 & -50.720 \\ 60.795 & 89.555 & 75.932 & -45.283 & -24.057 \\ 66.397 & 75.932 & 129.463 & -20.166 & -55.499 \\ -43.255 & -4,53E+04 & -20.166 & 106.512 & -16.051 \\ -50.720 & -2,41E+04 & -55.499 & -16.051 & 194.538 \end{pmatrix},$$

$$Q_5 = \begin{pmatrix} 31.900 & -18.126 & -0.8111 & -24.619 & -32.465 \\ -18.126 & 83.778 & 67.677 & -0.3710 & 62.216 \\ -0.8111 & 67.677 & 95.006 & -23.423 & 12.255 \\ -24.619 & -0.3710 & -23.423 & 36.539 & 23.782 \\ -32.465 & 6,22E+04 & 12.255 & 23.782 & 99.492 \end{pmatrix}$$

$$b_1 = \begin{pmatrix} 6,813 \\ -82,010 \\ -77,659 \\ -72,741 \\ 35,730 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 6,813 \\ -82,010 \\ -77,659 \\ -72,741 \\ 35,730 \end{pmatrix}, \quad b_3 = \begin{pmatrix} -0,965 \\ -62,058 \\ -0,999 \\ -70,478 \\ -89,005 \end{pmatrix},$$

$$b_4 = \begin{pmatrix} 70,143 \\ 12,112 \\ 85,922 \\ 39,333 \\ 16,558 \end{pmatrix}, \quad b_5 = \begin{pmatrix} 63,079 \\ 75,803 \\ 97,782 \\ -99,896 \\ 73,088 \end{pmatrix},$$

$$c_1 = -612,566, \quad c_2 = -989,950, \quad c_3 = -527,680,$$

$$c_4 = -479,523, \quad c_5 = -801,348,$$

$$a_1 = \begin{pmatrix} 6,173 \\ 30,889 \\ -18,476 \\ 63,996 \\ 43,672 \end{pmatrix}, \quad a_2 = \begin{pmatrix} 93,730 \\ 6,267 \\ -34,971 \\ -78,874 \\ 22,192 \end{pmatrix}, \quad a_3 = \begin{pmatrix} 55,760 \\ -15,309 \\ -81,835 \\ -46,706 \\ -69,269 \end{pmatrix}$$

$$a_4 = \begin{pmatrix} -43,799 \\ -11,983 \\ 5,429 \\ -8,515 \\ 75,074 \end{pmatrix} \text{ και } a_5 = \begin{pmatrix} 3,610 \\ 88,725 \\ 27,542 \\ 91,539 \\ -51,859 \end{pmatrix}$$

Πίνακας 3.3: Οι κορυφές που επιλέχθηκαν και οι αντίστοιχες τιμές της αντικειμενικής συνάρτησης φαίνονται στον παρακάτω πίνακα:

Επανάληψη	$f(x)$	$x(1)$	$x(2)$	$x(3)$	$x(4)$	$x(5)$
1	$-\infty$	3,175739	3,378209	-1,6857	4,52688	-3,52312
2	$-\infty$	3,175739	3,041079	-1,6857	-4,49407	-3,52312
3	$-\infty$	3,175739	3,378209	3,225582	-4,49407	0,937756
4	$-\infty$	-5,80758	-8,20608	6,30844	-4,49407	0,937756
5	$-\infty$	1,617392	-3,58052	-1,6857	4,52688	0,937756
6	-3863,44	-5,67883	-5,87885	-1,6857	4,52688	-3,52312
7	-1951,31	-1,19046	2,054892	6,30844	-4,49407	-1,25605
8	-1714,53	-1,18779	2,025566	6,30844	-4,49407	-1,25917
9	-1,66e+03	-1,17882	1,927206	6,30844	-4,49407	-1,26962
10	-1,45e+03	3,175739	-3,7218	5,905054	-0,22179	-3,52312
11	-1,06e+03	-0,25458	-8,20608	6,30844	-4,49407	-2,34718
12	-847,837	-5,67614	-5,88867	-1,6857	4,52688	-3,52312
13	-478,873	-0,27672	-1,41397	5,781083	-3,40754	-1,11641
14	-442,364	-2,51281	-7,95334	6,26673	4,52688	-2,6043
15	-348,709	3,175739	-3,47131	1,866602	2,414423	-0,59726
16	-313,766	-3,02312	3,378209	1,239619	-4,44909	-1,72871
17	-296,607	3,175739	-3,47185	1,866475	2,414498	-0,59704
18	-262,042	3,08088	-3,36382	1,80733	2,341815	-0,57849
19	-208,539	-0,27638	-1,37566	5,760578	-3,40314	-1,1452
20	-164,557	3,080062	-3,36302	1,808869	2,339851	-0,57981
21	-135,253	3,079802	-3,36254	1,809308	2,339498	-0,57941
22	-45,633	-3,02007	3,378209	1,270471	-4,45965	-1,73042
23	-13,835	-0,13737	0,300619	-0,19944	-0,12397	0,051222
24	-8,17533	-0,13732	0,300485	-0,19942	-0,12391	0,051239
25	-5,19693	-0,13719	0,300345	-0,19934	-0,12382	0,051215
26	-2,13232	-0,13714	0,300303	-0,19926	-0,12379	0,051234
27	-0,68292	-0,13717	0,300327	-0,19925	-0,12383	0,05122
28	-0,13645	3,079237	-3,36376	1,809081	2,340452	-0,57996
29	-0,10277	3,079237	-3,36376	1,809081	2,340452	-0,57996
30	-0,00969	3,079236	-3,36376	1,809083	2,34045	-0,57997
31	-0,00734	3,079223	-3,36374	1,809075	2,34044	-0,57996
32	-0,00128	3,079223	-3,36374	1,809075	2,34044	-0,57996
33	-0,00042	-0,13717	0,300322	-0,19925	-0,12382	0,051217
34	-8,27e-06	-0,13717	0,300322	-0,19925	-0,12382	0,051217
35	-3,18e-06	-0,13717	0,300322	-0,19925	-0,12382	0,051217

Πίνακας 3.4: Παρουσιάζεται ένα πλήθος επαναλήψεων για το πρόβλημα στις 10 διαστάσεις. Παρατηρούμε αναμενόμενη αύξηση στον αριθμό των επαναλήψεων αλλά και στον αριθμό των κορυφών. Χαρακτηριστικό είναι ότι το $V_0 = 1024$ και το V_105 όπου και τερματίζει ο αλγόριθμος είναι 272151!

Επανάληψη	f(x)	x(1)	x(2)	x(3)	x(4)	x(5)	x(6)	x(7)	x(8)	x(9)	x(10)
1	-782,479	-0,383	0,124	-0,030	-1,149	0,359	0,412	0,480	0,448	-0,743	-1,113
2	-724,355	-0,383	-0,132	-1,168	-1,149	0,359	0,412	0,480	-1,074	-0,743	-1,113
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
103	-2,87E-05	0,119	-0,487	-0,603	-0,770	0,463	0,079	0,215	-0,268	0,243	-0,571
104	-2,29E-05	-0,091	-0,223	-0,319	-0,456	0,557	0,065	0,471	-0,630	0,006	-0,118
105	-1,50E-05	-0,091	-0,223	-0,319	-0,456	0,557	0,065	0,471	-0,630	0,006	-0,118

Κεφάλαιο 4

Συμπεράσματα - Επεκτάσεις

Είδαμε ότι ο αλγόριθμος της Hoffman δουλεύει αποτελεσματικά για τη λύση συστημάτων εξισώσεων κυρτών συναρτήσεων. Έχει το ξεκάθαρο πλεονέκτημα της εγγυημένης ολικής σύγκλισης έναντι των περισσοτέρων άλλων γνωστών αλγορίθμων. Βέβαια το υπολογιστικό του κόστος είναι μεγάλο. Σε δέκα διαστάσεις για παράδειγμα το αρχικό πολύεδρο S_0 που απαιτείται για να εκκινήσει ο αλγόριθμος, έτσι όπως έχουμε επιλέξει εμείς να το υπολογίζουμε, έχει 1024 κορυφές. Στο παράδειγμά μας με τις δέκα διαστάσεις το S_{105} που δίνει τελικά και το συγκλίνων σημείο έχει 272151 κορυφές! Οπότε υπάρχουν έντονες δυσκολίες στην υλοποίηση του σε πραγματικά προβλήματα πολλών διαστάσεων. Παρόλες αυτές τις δυσκολίες όμως μας δίνει τη δυνατότητα να λύσουμε συστήματα εξισώσεων πολύ γενικής μορφής (των συναρτήσεων κυρτών διαφορών) και όχι μόνο κυρτών συναρτήσεων.

Καταρχήν μία συνάρτηση g ορισμένη στο κυρτό σύνολο $G \subseteq \mathbb{R}^n$ ονομάζεται συνάρτηση κυρτών διαφορών στο G, όταν για κάθε $x \in G,$ η $g(x)$μπορεί να εκφραστεί στη μορφή

$$g(x) = p(x) - q(x),$$

όπου οι p και q είναι κυρτές συναρτήσεις στο G.

Το παρακάτω θεώρημα με το λήμμα του (βλέπε [5]) επεκτείνει πολύ την κλάση των συναρτήσεων κυρτών διαφορών.

Θεώρημα: *Κάθε τοπικά κυρτών διαφορών συνάρτηση, είναι κυρτών διαφορών.*

Λήμμα: *Κάθε συνάρτηση $f \in \mathbb{C}^2$ είναι συνάρτηση κυρτών διαφορών.*

Όπου $\mathbb{C}$ είναι η κλάση των συναρτήσεων $\mathbb{R}^n \to \mathbb{R}$ των οποίων οι δεύτερες μερικές παράγωγοι είναι συνεχείς παντού.

Έστω λοιπόν ότι έχουμε το πρόβλημα:

$$f_j(x) = 0, \quad i = 1, \ldots, n \tag{4.1}$$

,με τις συναρτήσεις $f \in \mathbb{C}^2$, άρα και κυρτών διαφορών:

$$f_j(x) = g_j(x) - h_j(x), \quad j = 1, \ldots, n$$

,όπου οι g, h κυρτές συναρτήσεις ορισμένες στο $\mathbb{R}^n$. Άρα έχουμε:

$$\begin{aligned}(4.1) \Leftrightarrow & g_j(x) = h_j(x) = y_j, \quad j = 1, \ldots, n \\ \Leftrightarrow & \begin{cases} g_j(x) - y_j = & 0 \\ h_j(x) - y_j = & 0 \end{cases}, \quad j = 1, \ldots, n\end{aligned}$$

,με y_j σταθερές συναρτήσεις.

Μπορούμε ισοδύναμα να εκφράσουμε το πρόβλημα ως εξής:

$$F_i(x, y) = 0, \quad i = 1, \ldots, 2n.$$

$$\text{όπου} \begin{cases} F_i(x, y) = g_i(x) - y_i, & i = 1, \ldots, n \\ F_{n+i}(x, y) = h_i(x) - y_i, & 1, \ldots, n \end{cases}$$

Η F είναι προφανώς κυρτή στο $\mathbb{R}^n$, άρα το αρχικό σύστημα αντιμετωπίζεται όπως είδαμε στην παρούσα διπλωματική.

Ένα πρόβλημα έγκειται στον προσδιορισμό των συναρτήσεων g και h που είναι απαραίτητες για την αναγωγή του προβλήματος σε ένα πρόβλημα κοίλης ελαχιστοποίησης. Αυτό θα μπορούσε να γίνει θεωρώντας ότι $h(x) = \mu\|x\|_2^2$ και $g(x) = f(x) - h(x)$. Η h είναι προφανώς κυρτή. Όσον αφορά τη g έχουμε:

$$\frac{\partial^2 g(x)}{\partial x^2} = \frac{\partial^2 f(x)}{\partial x^2} + 2\mu I$$

Ελέγχουμε την κυρτότητα:

$$y^T \frac{\partial^2 g(x)}{\partial x^2} y = y^T \frac{\partial^2 f(x)}{\partial x^2} y + 2\mu\|y\|^2 \geq (\lambda_{min}(x) + 2\mu)\|y\|^2$$

,η ανίσωση προκύπτει από τη γνωστή σχέση της τετραγωνικής μορφής μιας μήτρας: $x^T A x \geq \lambda_{min}(A)\|x\|^2$. Οπότε παίρνοντας $\mu \geq \frac{1}{2}\lambda_{min}(x)$ εξασφαλίζουμε την κυρτότητα της συνάρτησης $g(x)$. Για να προχωρήσουμε από δω ίσως χρειαστεί να φράξουμε το x σε ένα κλειστό σύνολο ώστε να υπολογιστεί αυτή η τιμή.

Συνεπώς θεωρούμε σαν επέκταση ότι θα μπορούσε να είναι η λύση συστημάτων εξισώσεων τέτοιας μορφής. Να εξεταστεί η ορθή λειτουργία της μεθόδου μας και να συγκριθεί με άλλες γνωστές μεθόδους της βιβλιογραφίας ως προς τη χρονική κυρίως πολυπλοκότητά τους. Πιθανώς αυτές οι μέθοδοι να μην είναι ντετερμινιστικές αλλά τυχαιοκρατικές.

Ένθετο Α

Κώδικας Matlab

```
function x = diplwmatikh(n)
%Η συνάρτησή μας παίρνει ως είσοδο το n και επιστρέψει
%την ακολουθία των κορυφών που επιλέχθηκαν.

clc;
%Δημιουργούμε n τυχαία διανύσματα μεγέθους n από το
%-100 έως το 100.
a=construct_b(n,-100,100);
%Δημιουργούμε έναν θετικά ορισμένο πίνακα με γνωστή
%μέθοδο.
Q=construct_Q(n,-3,3);
%Ομοίως παίρνουμε n τυχαία διανύσματα μεγέθους n από
%το -100 εως το 100.
b=construct_b(n,-100,100);
%Δημιουργούμε n τυχαίους αρνητικούς αριθμούς στο
%[-1000,0].
c=-1000*rand(n,1);
%Φτιάχνουμε τις συναρτήσεις f_i των εξισώσεων.
f=construct_functions(a,b,c,Q,n);
%Φτιάχνουμε την αντικειμενική συνάρτηση της κοίλης
%ελαχιστοποίησης.
min_func = @(x) construct_objective_function(x,f);
%Προφανώς στο πρόβλημά μας η διάσταση του
%προβλήματος είναι ίδια με το πλήθος των περιορισμών.
m=n;
x_min=zeros(n,1);
x_max=zeros(n,1);

%Έχουμε τώρα όλα τα δεδομένα που απαιτούνται για τη
```

```
%λύση του προβλήματος.
%Βρίσκουμε το εσωτερικό σημείο της τομής των συνόλων
%που απαιτείται για να ξεκινήσει ο αλγόριθμος.
options = optimset('Algorithm','active-set',
'GradObj','on','GradConstr','off');
y1 = fmincon(@(x) interior_point_function(x,n),
zeros((n+1),1),[],[],[],[],[],[],
@(x) interior_point_constraints(x,f,m),options);
p=y1(2:end);
[g,~]=general_constraints(p,f,m);
%Έλεγχος αν υπήρχε τελικά εσωτερικό σημείο.
for i=1:length(g)
    if g(i)>0
        error('There is no interior point')
    end
end
%Έλεγχος αν το σημείο βρίσκεται στα σύνορα των συνόλων
%και αποτελεί λύση του προβλήματος
if abs(min_func(p))<1e-6
    x=p;
else
%Κατασκευή του πολυέδρου που περικλείει το
%επιτρεπόμενο σύνολο.
%Αρχικά βρίσκουμε τα ελάχιστα και τα μέγιστα της κάθε
%διάστασης ικανοποιώντας τους περιορισμούς.
options = optimset('Algorithm','active-set',
    'GradObj','on','GradConstr','off',
    'MaxFunEvals',100000,'MaxIter',100000,
    'TolFun',1e-10);
for i=1:n
    y2 = fmincon(@(x) id_i(x,i,n),zeros(n,1),[],[],[],
        [],[],[],@(x) general_constraints(x,f,m),
        options);
    x_min(i)=y2(i);
end;
for i=1:n
    y2 = fmincon(@(x) id_i_minus(x,i,n),zeros(n,1),[],
        [],[],[],[],[], @(x) general_constraints(x,f,m),
        options);
    x_max(i)=y2(i);

end;
%Με δεδομένα τα ελάχιστα και τα μέγιστα της κάθε
%διάστασης εκφράζουμε το πολύεδρό μας στη μορφή
```

```
%LQ*x<=Lb.
[LQ,Lb]=enclosing_linear_polyhedron(x_max,x_min,10000,
-10000,n);
%Βρίσκουμε τις κορυφές του πολυέδρου.
V = (con2vert(LQ,Lb))';
%Αρχικοποίηση τιμών για εκκίνηση του αλγορίθμου.
u(1) = min_func(p);
x(:,1)=zeros(n,1);
z(:,1)=zeros(n,1);
num=length(Lb);
J{1}=0;
k=2;
stop=0;
j=1;
jj=0;
solution(:,1)=p;

while stop==0
    clearvars Vtemp y Jtemp nr;
    Vtemp=V;
    check=1;
    %Εύρεση του ελαχίστου από τις υπάρχουσες κορυφές
    %της συγκεκριμένης επανάληψης.
    while check==1
        check=0;
        clearvars y;
        for i=1:size(Vtemp,2)
                y(i)=min_func(Vtemp(:,i));
        end
        [~,I]=min(y);
        x(:,k)=Vtemp(:,I);
        %Αν θέλουμε να αναζητήσουμε πιθανώς πολλαπλά
        %ολικά ελάχιστα αφαιρούμε τα ελάχιστα που έχουμε
        %ήδη βρει από το σύνολο των κορυφών.
        if jj>0
            for ii=1:size(solution,2)
                if abs(sum((x(:,k)-solution(:,ii)).^2))<1e-9
                    Vtemp(:,I)=[];

                    check=1;
                end
            end
        end
    end
```

```
    %Κάνουμε γραμμική αναζήτηση για να βρούμε το σημείο z
    %ανάμεσα στο p και στην κορυφή που έχουμε επιλέξει
    %που τέμνει τους περιορισμούς.
    lamvda = fmincon(@(l) fun(l),0,[],[],[],[],0,1,
     @(l) general_constraints((x(:,k)+l*(p-x(:,k))),f,m),
     options);
    %Αν το λ=0 τερματίζει ο αλγόριθμος, αν το λ>0 κάνουμε
    %την τομή και συνεχίζουμε.
    if lamvda > 0
        z(:,k) = x(:,k)+lamvda*(p-x(:,k));
        if min_func(z(:,k))<u(k-1)
            u(k)=min_func(z(:,k));
        else
            u(k)=u(k-1);
        end
        %Αν η κορυφή που βρήκαμε είναι πολύ κοντά στους
        %περιορισμούς, σταματάμε.
        if abs(u(k)-min_func(x(:,k)))<1e-6
            stop=1;
            %Αν θέλουμε να αναζητήσουμε πιθανώς και άλλες
            %λύσεις, δεν κάνουμε το stop->1 και συνεχίζουμε
            %τον αλγόριθμο.
%             jj=jj+1;
%             j=j+1;
%             solution(:,j)=x(:,k);
%             Vtemp(:,I)=[];
%             V=Vtemp;
%             u(k)=u(k-1);
        else
           %Βρίσκουμε τους περιορισμούς που ικανοποιεί το z.
            [g,~]=general_constraints(z(:,k),f,m);
            J{k}=[];
            for i=1:m
                if abs(g(i))<1e-6
                    J{k}=[J{k};i];
                end

            end
            Jtemp=J{k};
            %Επιλέγουμε τον πρώτο χωρίς βλάβη της
            %γενικότητας.
            ind=Jtemp(1);
            num=num+1;
            %Γράφουμε το κόψιμο σε μορφή A*x<=b και το
```

```
            %προσθέτουμε στους υπάρχοντες περιορισμούς.
            LQ(num,:)=(exp(a(:,ind)'*z(:,k))*a(:,ind)
             +Q(:,:,ind)*z(:,k) +b(:,ind))';
            Lb(num,:)=-g(ind)+(exp(a(:,ind)'*z(:,k))*
              a(:,ind)+Q(:,:,ind)*z(:,k)+b(:,ind))'*z(:,k);
            %Βρίσκουμε το νέο σύνολο κορυφών.
            V=(con2vert(LQ,Lb))';
        end
    else
        stop=1;

%         jj=jj+1;
%         j=j+1;
%         solution(:,j)=x(:,k);
%         Vtemp(:,I)=[];
%         V=Vtemp;
%         u(k)=u(k-1);
    end
    if jj>0
        jj=jj+1;
    end
    k=k+1;
    %Σταματάμε αν φτάσουμε τις 1000 επαναλήψεις, ή τις 50
    %επαναληψεις μετά το πρωτο ελάχιστο που βρίσκουμε.
    if ((k==1000) || (jj==50))
        stop=1;
    end
end
end
%Τυπώνουμε τα αποτελέσματα μας σε μορφή πίνακα.
output_display(x,min_func);
end


function b = construct_b (n,min,max)
    for k=1:n
        b(:,k)=min+(max-min).*rand(n,1);
    end
end
%Για την κατασκευή μιας θετικά ορισμένης μήτρας nxn
%παίρνουμε μια κάτω τριγωνική μήτρα με θετικά στοιχεία
%στη διαγώνιο και την πολλαπλασιάζουμε εξ αριστερών με
%την ανάστροφή της.
```

```
function Q = construct_Q (n,min,max)
    for k=1:n
        L=min+(max-min).*rand(n);
        for i=1:n
            for j=n:-1:(i+1)
                L(i,j)=0;
            end
        end
        for i=1:n
            L(i,i)=.5+(max-.5).*rand(1);
        end

        Q(:,:,k)=L*L';

    end
end

function f=construct_functions(a,b,c,Q,n)
    for i=1:n
        f{i}=@(x) (exp(a(:,i)'*x)+1/2*x'*Q(:,:,i)*x+
         b(:,i)'*x+c(i));
    end
end

function fun = construct_objective_function(x,f)
    fun=0;
    for i=1:length(f)
        fun = fun - feval(f{i},x);
    end
end


function [f,fgrad] = interior_point_function(x,n)
         f = x(1);
         fgrad = [1;zeros(n,1)];
end

function [c,ceq] = interior_point_constraints(x,f,m)
        for i=1:m
            c(i) = feval(f{i},x(2:end))-x(1);
        end
        ceq = [];
end
```

```
function [c,ceq] = general_constraints(x,f,m)
        for i=1:m
            c(i) = feval(f{i},x);
        end
        ceq = [];
end

function [A,b] = enclosing_linear_polyhedron(x_max,x_min,
max,min,n)
        A=[eye(n);eye(n);-eye(n);-eye(n)];
        b=[x_max;max*ones(n,1);-min*ones(n,1);-x_min];
end

function [ f,fgrad ] = id_i( x,i,n )
    f = x(i);
    fgrad = zeros(n,1);
    fgrad(i) = 1;
end

function [ f,fgrad ] = id_i_minus( x,i,n )
    f = -x(i);
    fgrad = zeros(n,1);
    fgrad(i) = -1;
end

function [ f,fgrad ] = id( x )
    f = x;
    fgrad = 1;
end

function output_display(x,min_func)
    A{1}='f(x) ';
    for j=1:size(x,1)
        A{j+1}=sprintf('x(%d) ',j);
    end
    B=strjoin(A,'\t');
    for j=1:(size(x,2)-1)
        A{j}=sprintf('Επανάληψη%d ',j);
    end
    A=strjoin(A,'\t');
    C=zeros(size(x,2)-1,size(x,1)+1);
    for i=1:(size(x,2)-1)
        C(i,1)=min_func(x(:,i+1));
    end
```

```
    for i=2:size(x,2)
        C(i-1,2:end)=x(:,i)';
    end
    printmat(C, 'Output',A,B);
end

%Προσθέτω και μια απλή περιγραφή της fmincon από το
%documentation της matlab.
% FMINCON finds a constrained minimum of a function of
% several variables.
%     FMINCON attempts to solve problems of the form:
%      min F(X)  subject to:  A*X <= B, Aeq*X  = Beq (linear
% constraints)
%       X                    C(X)<= 0, Ceq(X) = 0 (nonlinear
% constraints)
%                                LB <= X <= UB      (bounds)
%
%     X = FMINCON(FUN,X0,A,B) starts at X0 and finds a
% minimum X to the function FUN, subject to the linear
% inequalities A*X <= B. FUN accepts input X and
%       returns a scalar function value F evaluated at X. X0
%       may be a scalar, vector, or matrix.
%
%     X = FMINCON(FUN,X0,A,B,Aeq,Beq) minimizes FUN
% subject to the linear equalities Aeq*X = Beq as well
%       as A*X <= B. (Set A=[] and B=[] if no inequalities
%       exist.)

%
%     X = FMINCON(FUN,X0,A,B,Aeq,Beq,LB,UB) defines a
% set of lower and upper bounds on the design variables,
% X, so that a solution is found in the range
% LB <= X <= UB. Use empty matrices for LB and UB if
%       no bounds exist.
% Set LB(i) = -Inf if X(i) is unbounded below;
%       set UB(i) = Inf if X(i) is unbounded above.
%
%     X = FMINCON(FUN,X0,A,B,Aeq,Beq,LB,UB,NONLCON)
% subjects the minimization to the constraints defined
%       in NONLCON. The function NONLCON accepts X and
% returns the vectors C and Ceq, representing the
% nonlinear inequalities and equalities respectively.
% FMINCON minimizes FUN such that C(X) <= 0 and
% Ceq(X) = 0. (Set LB = [] and/or UB = [] if
```

```
%       no bounds exist.)
%
%     X = FMINCON(FUN,X0,A,B,Aeq,Beq,LB,UB,
% NONLCON,OPTIONS) minimizes with the default
% optimization parameters replaced by values in the
% structure OPTIONS, an argument created with the
% OPTIMSET function. See OPTIMSET for details. For a
% list of options accepted by FMINCON refer to the
% documentation.

function [V,nr] = con2vert(A,b)
% CON2VERT - convert a convex set of constraint inequalities
%           into the set of vertices at the intersections of
%            those inequalities;i.e., solve the "vertex
%            enumeration" problem. Additionally, identify
%            redundant entries in the list of inequalities.
%
% V = con2vert(A,b)
% [V,nr] = con2vert(A,b)
%
% Converts the polytope (convex polygon, polyhedron, etc.)
% defined by the system of inequalities A*x <= b into a list
% of vertices V. Each ROW of V is a vertex. For n variables:
% A = m x n matrix, where m >= n (m constraints, n variables)
% b = m x 1 vector (m constraints)

% V = p x n matrix (p vertices, n variables)
% nr = list of the rows in A which are NOT redundant
% constraints

c = A\b;
if ~all(A*c < b);
    [c,f,ef] = fminsearch(@obj,c,'params',{A,b});
    if ef ~= 1
        error('Unable to locate a point within the interior
            of a feasible region.')
    end
end
b = b - A*c;
D = A ./ repmat(b,[1 size(A,2)]);
[k,v2] = convhulln([D;zeros(1,size(D,2))]);
[k,v1] = convhulln(D);
if v2 > v1
    error('Non-bounding constraints detected. (Consider
```

```
     box constraints on variables.)')
end
nr = unique(k(:));
G  = zeros(size(k,1),size(D,2));
for ix = 1:size(k,1)
    F = D(k(ix,:),:);
    G(ix,:)=F\ones(size(F,1),1);
end
V = G + repmat(c',[size(G,1),1]);
[null,I]=unique(num2str(V,6),'rows');
V=V(I,:);
return
function d = obj(c,params)
A=params{1};
b=params{2};
d = A*c-b;
k=(d>=-1e-15);
d(k)=d(k)+1;
d = max([0;d]);
return
```

Βιβλιογραφία

[1] J. E. Dennis, Robert B. Schnabel. "Numerical methods for unconstrained optimization and nonlinear equations". 1983 by Prentice-Hall, Inc., Englewood Cliffs, New Jersey 07632.

[2] Harold P. Benson. "Handbook of Global Optimization (Nonconvex Optimization and Its Applications)", chapter "Concave minimization: theory, applications and algorithms". Edited by R. Horst, Panos M. Pardalos. Kluwer Academic Publishers 1995.

[3] R. Horst, N. V. Thoai and H. Tuy. "Outer Approximation by Polyhedral Convex Sets". Operations Research Spektrum 1987.

[4] R. Horst, N. V. Thoai and H. Tuy. "On an Outer Approximation Concept in Global Optimization". Optimization 1989.

[5] R. Horst and H. Tuy. "Global Optimization: Deterministic Approaches". 2nd Edn. Springer-Verlag. Berlin 1993

[6] Karla Leigh Hoffman. "A method for globally minimizing concave functions over convex sets". Mathematical Programming 20 (1981) 22-32. North Holland Publishing Company.