ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

Προβλήματα Βελτιστοποίησης με Περιορισμούς και Δεσμεύσεις

Θεωρία και Υπολογισμοί

Διπλωματική Εργασία

Γεωργία Κουρμουλή

Επιβλέπων: Κωνσταντίνος Χρυσαφίνος, Αν. Καθηγητής ΕΜΠ

Αθήνα, Οκτώβριος 2020

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

Προβλήματα Βελτιστοποίησης με Περιορισμούς και Δεσμεύσεις

Θεωρία και Υπολογισμοί

Διπλωματική Εργασία

Γεωργία Κουρμουλή

Επιβλέπων: Κωνσταντίνος Χρυσαφίνος, Αν. Καθηγητής ΕΜΠ

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή:

....	....	....
Κωνσταντίνος Χρυσαφίνος Αναπλ. Καθηγητής ΕΜΠ	Εμμανουήλ Γεωργούλης Αναπλ. Καθηγητής ΕΜΠ	Ιωάννης Κολέτσος Επίκουρος Καθηγητής ΕΜΠ

Αθήνα, Οκτώβριος 2020

Πρόλογος

Με την ολοκλήρωση της παρούσας διπλωματικής εργασίας, θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή μου κύριο Κωνσταντίνο Χρυσαφίνο για την καθοδήγηση και τις εύστοχες παρατηρήσεις του, καθώς και για την στήριξή του καθ΄ όλη την διάρκεια συγγραφής της εργασίας.

Επιπλέον, θα ήθελα να ευχαριστήσω την οικογένεια και τους φίλους μου, για την ψυχολογική τους στήριξη, που με βοήθησε να ξεπεράσω όλες τις δυσκολίες που εμφανίστηκαν στον δρόμο μου, αλλά και για τις γνώσεις που μου προσέφεραν, καθιστώντας την ολοκλήρωση αυτής της προσπάθειας δυνατή.

Περίληψη

Η διπλωματική αυτή εργασία πραγματεύεται το θεωρητικό υπόβαθρο των προβλημάτων βελτιστοποίησης με περιορισμούς, στο οποίο στηρίζονται οι τεχνικές επίλυσής τους. Στην συνέχεια, παρουσιάζονται κάποιοι υπολογισμοί που γίνονται με την βοήθεια του μαθηματικού πακέτου Matlab και των μεθόδων Active Set και Interior-point.

Στην εργασία παρουσιάζονται αρχικά, όλες οι βασικές μαθηματικές έννοιες που είναι χρήσιμες για την κατανόηση της θεωρίας αλλά και για την επίλυση των προβλημάτων. Στην συνέχεια, αναπτύσσονται οι τεχνικές επίλυσης αυτού του τύπου προβλημάτων, οι οποίες βασίζονται στις συνθήκες του θεωρήματος των Karush - Kuhn - Tucker. Έπειτα, εξετάζονται οι προϋποθέσεις που πρέπει να καλύπτουν οι περιορισμοί ενός προβλήματος βελτιστοποίησης, προκειμένου ένα σημείο να αποτελεί λύση του, αλλά και η προσέγγιση της λύσης αυτής να γίνεται γρήγορα μέσα στην εφικτή περιοχή. Τέλος, παρουσιάζεται ο αλγόριθμος της μεθόδου Active Set και η υλοποίησή του στο Matlab, ενώ μέσα από την επίλυση παραδειγμάτων επιβεβαιώνεται η σωστή λειτουργία του και γίνεται σύγκριση των αποτελεσμάτων του με αυτά του αλγορίθμου Interior-point.

Abstract

This diploma thesis is related to the theoretical setting of the constrained optmization, in which problem solving techniques are based. Also some calculations are presented, with the help of the mathematical package Matlab and the methods Active Set and Interior-point.

Initially, in this diploma thesis, the mathematical background is presented, which is necessary for the comprehension of the theory and for the problem solving. Later on, solving techniques for this type of problems are discussed, which are based on the theorem of Karush - Kuhn - Tucker. Moreover, the constraint qualifications are examined, which ensure that a point is the solution of a problem and that the approximation of this solution is made by taking big steps inside the feasible region. Finally, the algorithm of the Active Set method and its corresponding code in Matlab are presented, while the optimization algorithm is assessed through some examples and compared to the Interior-point one.

Περιεχόμενα

Πίνακας Συντομογραφιών

KKT	Karush-Kuhn-Tucker
LICQ	Linear Independence Constraint Qualification
MFCQ	Mangasarian-Fromovitz Constraint Qualification

Πίνακας Συμβόλων

$\mathbb{R}$, $\mathbb{R}^n$, $\mathbb{R}^{m\times n}$	Τα σύνολα των πραγματικών αριθμών, διανυσμάτων και $m \times n$ πινάκων
x^*	Βέλτιστη λύση του προβλήματος
$f(x)$	Αντικειμενική συνάρτηση
$c_i(x)$	Περιορισμοί
$\mathcal{E}$	Σύνολο των δεικτών των ισοτικών περιορισμών
$\mathcal{I}$	Σύνολο των δεικτών των ανισοτικών περιορισμών
$\mathcal{U}$	Εφικτό σύνολο
$\nabla f(x)$	Κλίση της συνάρτησης f
$\nabla^2 f(x)$	Εσσιανός πίνακας της συνάρτησης f
λ	Διάνυσμα πολλαπλασιαστών Lagrange
$\mathcal{L}(x,\lambda)$	Συνάρτηση Lagrange
$\mathcal{A}(x^*)$	Ενεργό σύνολο στο σημείο x^*
$\{z_k\}, k \in \mathbb{N}$	Εφικτή ακολουθία
$T_{\mathcal{U}}(x)$	Εφαπτόμενος κώνος
$\mathcal{F}(x)$	Σύνολο των εφικτών διευθύνσεων πρώτης τάξης
$\mathcal{A}(x^*)^T$	Πίνακας με γραμμές τις κλίσεις των ενεργών περιορισμών στο σημείο x^*
$\mathcal{C}(x^*,\lambda^*)$	Κρίσιμος κώνος
$rank(A)$	Βαθμός του πίνακα A
I_n	Μοναδιαίος πίνακας τάξης n
$Null(A)$ ή $Ker(A)$	Μηδενοχώρος ή πυρήνας του πίνακα A
$N_{\mathcal{U}}(x)$	Κανονικός κώνος
$v \in N_{\mathcal{U}}(x)$	Κανονικό διάνυσμα
$q(\lambda)$	Δυϊκή αντικειμενική συνάρτηση
$\mathcal{W}$	Σύνολο εργασίας

Κεφάλαιο 1

Βασικές Έννοιες

Η *βελτιστοποίηση προβλημάτων* είναι μία διαδικασία, την οποία για να υλοποιήσουμε θα πρέπει αρχικά να εντοπίσουμε το *αντικείμενο* που πρέπει να μελετηθεί. Αυτό το αντικείμενο μπορεί να είναι το κέρδος κάποιας επιχείρησης, ο χρόνος παραγωγής κάποιου προϊόντος ή κάποιο μέγεθος ή και συνδυασμός μεγεθών, τα οποία μπορούν να αναπαρασταθούν με ένα νούμερο ή να ποσοτικοποιηθούν. Ο σκοπός της μελέτης, εξαρτάται από κάποια χαρακτηριστικά του συστήματος στο οποίο ανήκει, τα οποία καλούνται *μεταβλητές* ή *παράμετροι* του συστήματος. Στόχος της βελτιστοποίησης είναι να βρεθούν οι τιμές των μεταβλητών, για τις οποίες το αντικείμενο της μελέτης βελτιστοποιείται. Πολλές φορές, οι μεταβλητές του συστήματος πρέπει να καλύπτουν κάποιες προϋποθέσεις, τους *περιορισμούς* του συστήματος. Τέτοιοι περιορισμοί μπορεί να δημιουργηθούν λόγω της φύσης των μεταβλητών, για παράδειγμα το φορτίο ενός ηλεκτρονίου πρέπει να είναι αρνητικό, ή λόγω της φύσης του συστήματος, για παράδειγμα η ημερομηνία παράδοσης κάποιου έργου.

Η διαδικασία εύρεσης του αντικειμένου μελέτης, των μεταβλητών και των περιορισμών του συστήματος καλείται *μαθηματική μοντελοποίηση του προβλήματος*. Η δημιουργία ενός κατάλληλου μοντέλου είναι το πρώτο βήμα και το πιο σημαντικό στην διαδικασία της βελτιστοποίησης. Εάν το μοντέλο είναι πολύ απλό, δηλαδή κάνει χρήση μερικών μόνο παραμέτρων του συστήματος, τότε είναι πιθανό να μην δώσει χρήσιμα αποτελέσματα για τον σκοπό της μελέτης. Από την άλλη, βέβαια, εάν το μοντέλο είναι ιδιαίτερα πολύπλοκο, είναι πολύ πιθανό να μην μπορεί να λυθεί.

Από την στιγμή που το μοντέλο έχει διαμορφωθεί, μπορεί να χρησιμοποιηθεί ένας *αλγόριθμος βελτιστοποίησης*, μέσω ενός υπολογιστή, ώστε να βρεθεί η λύση του. Ο αλγόριθμος βελτιστοποίησης δεν είναι ενιαίος για όλα τα προβλήματα, αλλά συνήθως γίνεται χρήση ενός συνδυασμού αλγορίθμων, οι οποίοι έχουν σχεδιαστεί για να επιλύουν συγκεκριμένα στάδια του προβλήματος. Η επιλογή του καταλληλότερου αλγορίθμου είναι στα χέρια αυτού που κάνει την εκάστοτε μελέτη και είναι πολύ σημαντική, καθώς καθορίζει αρχικά εάν μπορεί να βρεθεί η λύση και έπειτα εάν το πρόβλημα μπορεί να λυθεί γρήγορα ή αργα.

Αφού ένας αλγόριθμος βελτιστοποίησης εφαρμοστεί στο πρόβλημα, μπορούμε να διαπιστώσουμε εάν πέτυχε τον στόχο του, δηλαδή αν βρήκε λύση στο πρόβλημα. Η λύση αυτή που μας επέδειξε ο αλγόριθμος, ισχύει για τις παραμέτρους του συστηματος που ορίσαμε κατά την μοντελοποίηση του προβλήματος. Ωστόσο, συχνά ενδιαφερόμαστε να γνωρίζουμε τι επιπτώσεις θα είχε στην λύση του προβλήματος μία τυχόν αλλαγή στο σύστημα ή στα δεδομένα. Αυτή η μέθοδος, κατά την οποία προσδιορίζεται η ευαισθησία της λύσης σε μεταβολές των παραμέτρων του συστήματος, ονομάζεται *ανάλυση ευαισθησίας (sensitivity analysis).*

1.1 Μαθηματική μοντελοποίηση

Από μαθηματικής άποψης, βελτιστοποίηση προβλημάτων θεωρείται η διαδικασία εύρεσης της βέλτιστης λύσης ενός προβλήματος. Το πρόβλημα αυτό είναι η ελαχιστοποίηση ή η μεγιστοποίηση της αντικειμενικής συνάρτησης, δηλαδή της μαθηματικής έκφρασης του σκοπού της μελέτης, μέσα σε ένα σύνολο περιορισμών.

Οι βασικοί συμβολισμοί που θα χρησιμοποιηθούν είναι οι εξής:

- x το διάνυσμα των μεταβλητών ή αλλιώς παραμέτρων (variables)
- x^* η λύση του προβλήματος, δηλαδή το διάνυσμα των μεταβλητών που επιλύει το πρόβλημα
- f η αντικειμενική συνάρτηση (objective function), δηλαδή η συνάρτηση του διανύσματος x, της οποίας αναζητούμε το ελάχιστο ή το μεγιστο
- c_i οι περιορισμοί (constraints), δηλαδή συναρτήσεις του διανύσματος x, οι οποίες δημιουργούνται από τις ισότητες (equalities) ή τις ανισότητες (inequalities) που το άγνωστο διάνυσμα x πρέπει να ικανοποιεί
- $\mathcal{E}$ το σύνολο που περιέχει τους δείκτες των ισοτικών περιορισμών
- $\mathcal{I}$ το σύνολο που περιέχει τους δείκτες των ανισοτικών περιορισμών
- $\mathcal{U}$ το εφικτό σύνολο (feasible set), δηλαδή το σύνολο των x που ικανοποιούν τους περιορισμούς, τα στοιχεία του οποίου ονομάζονται εφικτά σημεία (feasible points).

Σύμφωνα με τους παραπάνω συμβολισμούς, το *γενικό πρόβλημα ελαχιστοποίησης με περιορισμούς* για συναρτήσεις $f, c_i : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, $S \subset \Omega$ γράφεται ως εξής:

$$\text{Να βρεθεί } x^* \in \mathcal{U} \text{ τέτοιο ώστε } f(x^*) = \min_{x\in\mathcal{U}} f(x), \tag{1.1}$$
$$\text{όπου } \mathcal{U} = \{x \in S \mid c_i(x) = 0,\, i \in \mathcal{E}, c_i(x) \geq 0,\, i \in \mathcal{I}\}.$$

Παράδειγμα 1.1.1. *Έστω η συνάρτηση $f(x) = (x_1 - 2)^2 + (x_2 - 1)^2$ την οποία θέλουμε να ελαχιστοποιήσουμε υπό τους περιορισμούς $x_1^2 - x_2 \leq 0$ και $x_1 + x_2 \leq 2$, τότε το πρόβλημα γράφεται στην μορφή (1.1) ως εξής:*

$$\min_{x\in\mathcal{U}} (x_1 - 2)^2 + (x_2 - 1)^2, \ \textit{με} \ \mathcal{E} = \emptyset, \ \mathcal{I} = \{1, 2\}$$

$$\textit{και} \ \mathcal{U} = \{x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 : -x_1^2 + x_2 \geq 0, -x_1 - x_2 + 2 \geq 0\}.$$

Παράδειγμα 1.1.2. *Έστω η συνάρτηση $f(x) = x_1^2 + x_2^2 + x_3^2$ την οποία θέλουμε να ελαχιστοποιήσουμε υπό τους περιορισμούς $x_1 + x_2 + x_3 \geq 1$, $x_1 \geq 0$, $x_2 \geq 0$ και $x_3 \geq 0$, τότε το πρόβλημα γράφεται στην μορφή (1.1) ως εξής:*

$$\min_{x\in\mathcal{U}} x_1^2 + x_2^2 + x_3^2, \ \text{με}\ \mathcal{E} = \emptyset,\ \mathcal{I} = \{1, 2, 3, 4\}$$

$$\text{και}\ \ \mathcal{U} = \{x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 : x_1 + x_2 + x_3 - 1 \geq 0, x_1 \geq 0, x_2 \geq 0, x_3 \geq 0\},$$

όπου το εφικτό σύνολο του προβλήματος σχηματίζει ένα πολύεδρο.

Παράδειγμα 1.1.3. *Έστω η συνάρτηση $f(x) = -x_1 - x_2 - 2x_3 + 5$ την οποία θέλουμε να ελαχιστοποιήσουμε υπό τους περιορισμούς $x_1^2 + x_2 = 1$, $x_1 + x_2 + x_3 = 1$ και $x_2 \geq 0$, τότε το πρόβλημα γράφεται στην μορφή (1.1) ως εξής:*

$$\min_{x\in\mathcal{U}} \ - x_1 - x_2 - 2x_3 + 5,\ \text{με}\ \mathcal{E} = \{1, 2\},\ \mathcal{I} = \{1\}$$

$$\text{και}\ \ \mathcal{U} = \{x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 : x_1^2 + x_2 - 1 = 0, x_1 + x_2 + x_3 - 1 = 0, x_2 \geq 0\}.$$

Ορισμός 1.1. *Έστω συνάρτηση $f : S \subset \mathbb{R}^n \longrightarrow \mathbb{R}$. Το x^* είναι σημείο ολικού ελαχίστου ή σημείο ελαχίστου της f στο S εάν:*

$$f(x^*) \leq f(x), \forall x \in S.$$

Αντίστοιχα, το x^* είναι σημείο ολικού μεγίστου ή σημείο μεγίστου της f στο S εάν:

$$f(x^*) \geq f(x), \forall x \in S.$$

Ορισμός 1.2. *Έστω συνάρτηση $f : S \subset \mathbb{R}^n \longrightarrow \mathbb{R}$. Το x^* είναι σημείο τοπικού ελαχίστου της f στο S εάν:*

$$\exists\ \delta > 0\ \text{τέτοιο ώστε}\ f(x^*) \leq f(x), \forall x \in S\ \text{με}\ \|x - x^*\| < \delta.$$

Αντίστοιχα, το x^* είναι σημείο τοπικού μεγίστου της f στο S εάν:

$$\exists\ \delta > 0\ \text{τέτοιο ώστε}\ f(x^*) \geq f(x), \forall x \in S\ \text{με}\ \|x - x^*\| < \delta.$$

Παρατηρήσεις

1. Αν το x^* είναι σημείο ολικού ελαχίστου (μεγίστου) της f, τότε είναι και σημείο τοπικού ελαχίστου (μεγίστου).

2. Ένα σημείο x^* είναι σημείο μεγίστου (ολικό ή τοπικό) της f αν και μόνο αν είναι σημείο ελαχίστου (ολικό ή τοπικό) της $-f$.

1.2 Βελτιστοποίηση προβλημάτων με περιορισμούς και χωρίς περιορισμούς

Τα προβλήματα της μορφής (1.1) χωρίζονται σε διάφορες κατηγορίες, ανάλογα με την μορφή των αντικειμένων που τα απαρτίζουν. Ένα βασικό στοιχείο για την κατηγοριοποίησή τους είναι η ύπαρξη ή όχι περιορισμών.

Προβλήματα χωρίς περιορισμούς

Σε αυτή την κατηγορία προβλημάτων ψάχνουμε το ελάχιστο ή το μέγιστο της συνάρτησης f, με τα σύνολα $\mathcal{E} = \mathcal{I} = \emptyset$. Σε μερικές περιπτώσεις, ακόμη και στα προβλήματα με περιορισμούς, οι περιορισμοί μπορούν να απαλειφθούν εάν δεν επηρεάζουν την λύση του προβλήματος. Ωστόσο, ένας ακόμη τρόπος για την απαλοιφή τους, είναι μέσω της συνάρτησης ποινής και της μεθόδου των ποινών, όπου η λύση ενός προβλήματος βελτιστοποίησης με περιορισμούς ανάγεται σε λύση μιας ακολουθίας προβλημάτων χωρίς περιορισμούς.

Προβλήματα με περιορισμούς

Τα προβλήματα με περιορισμούς δημιουργούνται από μοντέλα στα οποία οι περιορισμοί παίζουν πολύ σημαντικό ρόλο στην λυση του προβλήματος, για παράδειγμα οι οικονομικοί περιορισμοί μιας επιχείρησης σε ένα πρόβλημα ελαχιστοποίησης του κόστους. Αυτοί οι περιορισμοί μπορεί να είναι από μία απλή ανίσωση, μία γραμμική εξίσωση ή ακόμη και πιο περίπλοκες μαθηματικές εκφράσεις των μεταβλητών.

Όταν η αντικειμενική συνάρτηση f και οι περιορισμοί c_i είναι γραμμικές συναρτήσεις ως προς x, τότε είναι πρόβλημα *γραμμικού προγραμματισμού*. Τέτοιου είδους προβλήματα πραγματεύονται αρκετά συχνά οι επιχειρήσεις. Ωστόσο, υπάρχουν και τα προβλήματα *μη γραμμικού προγραμματισμού*, στα οποία τουλάχιστον ένας από τους περιορισμούς ή η αντικειμενική συνάρτηση, είναι μη γραμμικές συναρτήσεις του x.

Στα πλαίσια αυτής της διπλωματικής θα ασχοληθούμε κατά κύριο λόγο με προβλήματα ελαχιστοποίησης με περιορισμούς και συγκεκριμένα μη γραμμικού προγραμματισμού.

1.3 Πίνακες

Στα πλαίσια αυτής της διπλωματικής ασχολούμαστε μόνο με διανύσματα και πίνακες, των οποίων τα στοιχεία είναι πραγματικοί αριθμοί. Τον χώρο των διανυσμάτων μήκους n τον συμβολίζουμε με $\mathbb{R}^n$ και τον χώρο των πινάκων διαστάσεων $m \times n$ τον συμβολίζουμε με $\mathbb{R}^{m\times n}$. Τα *στοιχεία ενός διανύσματος* $x \in \mathbb{R}^n$ τα συμβολίζουμε με $x_i, i = 1, 2, \ldots, n$. Αντίστοιχα, τα *στοιχεία ενός πίνακα* $A \in \mathbb{R}^{m\times n}$ τα συμβολίζουμε με $a_{ij}, i = 1, 2, \ldots, m, j = 1, 2, \ldots, n$.

Ορισμός 1.3. *Ένας πίνακας $A \in \mathbb{R}^{m\times n}$ του οποίου ο αριθμός των γραμμών ισούται με τον αριθμό των στηλών, δηλαδή $m = n$, λέγεται τετραγωνικός πίνακας.*

Ορισμός 1.4. *Έστω ένας τετραγωνικός πίνακας $A \in \mathbb{R}^{n\times n}$. Τα στοιχεία a_{ij}, με $i = j$, ορίζουν την κύρια διαγώνιο του πίνακα.*

Ορισμός 1.5. *Ένας τετραγωνικός πίνακας $A \in \mathbb{R}^{n \times n}$ λέγεται άνω τριγωνικός εάν όλα τα στοιχεία του που βρίσκονται κάτω από την κύρια διαγώνιο είναι 0, δηλαδή $a_{ij} = 0$, με $i > j$.*

Αντίστοιχα, ο πίνακας A λέγεται κάτω τριγωνικός εάν όλα τα στοιχεία του που βρίσκονται πάνω από την κύρια διαγώνιο είναι 0, δηλαδή $a_{ij} = 0$, με $i < j$.

Ορισμός 1.6. *Έστω τετραγωνικός πίνακας $A \in \mathbb{R}^{n \times n}$. Ο πίνακας που προκύπτει με εναλλαγή μεταξύ των γραμμών και των στηλών του λέγεται ανάστροφος πίνακας του A και συμβολίζεται με A^T, δηλαδή $A^T = (a_{ji}) \in \mathbb{R}^{n \times n}$.*

Ορισμός 1.7. *Ένας τετραγωνικός πίνακας $A \in \mathbb{R}^{n \times n}$ λέγεται συμμετρικός εάν $A^T = A$, δηλαδή εάν $a_{ij} = a_{ji}$ για κάθε $i, j = 1, 2, \ldots, n$.*

Αντίστοιχα, ο πίνακας A λέγεται αντισυμμετρικός εάν $A^T = -A$, δηλαδή εάν $a_{ij} = -a_{ji}$ για κάθε $i, j = 1, 2, \ldots, n$.

Ορισμός 1.8. *Ο πίνακας $I_n \in \mathbb{R}^{n \times n}$, με:*

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

λέγεται μοναδιαίος πίνακας.

Ορισμός 1.9. *Ένας τετραγωνικός πίνακας $A \in \mathbb{R}^{n \times n}$ λέμε ότι είναι αντιστρέψιμος, εάν υπάρχει πίνακας B τέτοιος ώστε:*

$$AB = BA = I_n.$$

Τότε ο πίνακας B καλείται αντίστροφος του A και γράφουμε $B = A^{-1}$.

Ορισμός 1.10 (βλ. [1], Κεφάλαιο 14, Ορισμός 14.3.1 (β), σελ. 426)**.** *Ένας τετραγωνικός πίνακας $A \in \mathbb{R}^{n \times n}$ λέγεται ορθογώνιος εάν ικανοποιεί τις ισότητες:*

$$AA^T = A^T A = I_n \quad \text{ή ισοδύναμα} \quad A^T = A^{-1}.$$

Ορισμός 1.11 (βλ. [1], Κεφάλαιο 15, Ορισμός 15.5.1 (i), σελ. 468)**.** *Ένας τετραγωνικός πίνακας $A \in \mathbb{R}^{n \times n}$ λέγεται θετικά ημιορισμένος ή θετικός, αν $x^T A x \geq 0$, για κάθε $x \in \mathbb{R}^n$.*

Αντίστοιχα, ο πίνακας $A \in \mathbb{R}^{n \times n}$ λέγεται θετικά ορισμένος, αν $x^T A x > 0$, για κάθε $x \in \mathbb{R}^n, x \neq 0$.

Ορισμός 1.12 (βλ. [1], Κεφάλαιο 11, Ορισμός 11.1.1 (β), σελ. 300)**.** *Έστω ένας τετραγωνικός πίνακας $A \in \mathbb{K}^{n \times n}$, όπου $\mathbb{K} = \mathbb{R}$ ή $\mathbb{C}$. Ένα διάνυσμα $q \in \mathbb{K}^n - \{0\}$ λέγεται ιδιοδιάνυσμα του πίνακα A, αν ικανοποιεί την εξίσωση:*

$$Aq = \lambda q,$$

για κάποιο $\lambda \in \mathbb{K}$. Τότε το λ αυτό, λέγεται ιδιοτιμή του πίνακα A.

Ορισμός 1.13 (βλ. [2], Κεφάλαιο 3, Ορισμός, σελ. 111). *Έστω ένας τετραγωνικός πίνακας $A \in \mathbb{R}^{n\times n}$. Ορίζουμε την φασματική ακτίνα $\rho(A)$ του A ως το μέγιστο των απολύτων τιμών των ιδιοτιμών του πίνακα A.*

Παρατηρήσεις

1. Ένας πίνακας $A_{n\times n}$ είναι αντιστρέψιμος εάν $\lambda_i \neq 0$, $\forall i = 1, \ldots, n$.
2. Ένας πίνακας $A_{n\times n}$ είναι θετικά ημιορισμένος αν και μόνο αν οι ιδιοτιμές του είναι $\lambda_i \geq 0$, $\forall i = 1, \ldots, n$.
3. Ένας πίνακας $A_{n\times n}$ είναι θετικά ορισμένος αν και μόνο αν οι ιδιοτιμές του είναι $\lambda_i > 0$, $\forall i = 1, \ldots, n$.

1.4 Υπόχωροι

Ορισμός 1.14 (βλ. [3], Κεφάλαιο 1, Ορισμός 1.2, σελ. 8). *Έστω ο Ευκλείδειος χώρος $\mathbb{R}^n$. Ένα μη κενό $S \subset \mathbb{R}^n$ είναι υπόχωρος του $\mathbb{R}^n$ αν για κάθε $x, y \in S$ και $\lambda \in \mathbb{R}$ ισχύει $x + y \in S$ και $\lambda x \in S$.*

Παραδείγματα υποχώρων

1. Για κάθε χώρο $\mathbb{R}^n$, οι $\mathbb{R}^n$ και $\{0\}$ είναι προφανώς υπόχωροί του.
2. Οι υπόχωροι του $\mathbb{R}^2$ είναι:
   (i) Ο μηδενικός υπόχωρος $\{0\}$.
   (ii) Οι ευθείες που περνάνε από το 0.
   (iii) Ο ίδιος ο $\mathbb{R}^2$.
3. Για κάθε σύνολο διανυσμάτων $a_i \in \mathbb{R}^n, i = 1, 2, \ldots, m$, το σύνολο:
$$S = \{w \in \mathbb{R}^n \mid a_i^T w = 0, i = 1, 2, \ldots, m\}$$
   είναι υπόχωρος του $\mathbb{R}^n$.

Ορισμός 1.15 (βλ.[3], Κεφάλαιο 1, Ορισμός 1.6, σελ. 10). *Ένα πεπερασμένο υποσύνολο $\{x_1, \ldots, x_n\}$ του $\mathbb{R}^n$ λέγεται γραμμικά ανεξάρτητο αν για κάθε $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ τέτοια ώστε $\lambda_1 x_1 + \ldots + \lambda_n x_n = 0$ έχουμε ότι $\lambda_1 = \ldots = \lambda_n = 0$.*

Αντίστοιχα, ένα πεπερασμένο υποσύνολο $\{x_1, \ldots, x_n\}$ του $\mathbb{R}^n$ λέγεται γραμμικά εξαρτημένο αν υπάρχουν $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ όχι όλα ίσα με μηδέν, τέτοια ώστε $\lambda_1 x_1 + \ldots + \lambda_n x_n = 0$.

Ορισμός 1.16 (βλ.[3], Κεφάλαιο 1, Ορισμός 1.6, σελ. 10). *Ένα $S \subset \mathbb{R}^n$ λέγεται γραμμικά ανεξάρτητο αν κάθε πεπερασμένο υποσύνολό του είναι γραμμικά ανεξάρτητο.*

Αντίστοιχα, το S λέγεται γραμμικά εξαρτημένο αν έχει τουλάχιστον ένα γραμμικά εξαρτημένο πεπερασμένο υποσύνολο.

Ορισμός 1.17 (βλ. [1], Κεφάλαιο 8, Ορισμός, σελ. 197)**.** *Έστω ένας διανυσματικός χώρος V και ένα $S \subset V$. Το σύνολο των γραμμικών συνδυασμών μεταξύ των στοιχείων του S είναι υπόχωρος του V και λέγεται γραμμική θήκη του S. Δηλαδή το σύνολο:*

$$< S >= \left\{ \sum_{i=1}^{n} \lambda_i x_i \mid \lambda_i \in \mathbb{R}, x_i \in S, i = 1, \dots, n, n \in \mathbb{N} \right\}.$$

Ορισμός 1.18 (βλ. [1], Κεφάλαιο 8, Ορισμός, σελ. 197)**.** *Έστω ένας διανυσματικός χώρος V και ένα $S \subset V$. Αν $< S >= V$, δηλαδή εάν:*

$$v = \lambda_1 x_1 + \ldots + \lambda_n x_n, \quad \text{για κάθε} \ \ v \in V, x_i \in S, \lambda_i \in \mathbb{R}, i = 1, \dots, n$$

λέμε ότι το σύνολο S παράγει τον χώρο V.

Ορισμός 1.19 (βλ. [1], Κεφάλαιο 8, Ορισμός 8.4.1, σελ. 203)**.** *Έστω ένας διανυσματικός χώρος V και ένα $S \subset V$. Αν τα στοιχεία του συνόλου S είναι γραμμικά ανεξάρτητα μεταξύ τους και το S παράγει τον χώρο V, τότε λέμε ότι το S είναι βάση του V.*

Ορισμός 1.20 (βλ. [1], Κεφάλαιο 13, Παράδειγμα 1, σελ. 386)**.** *Έστω διανύσματα $x, y \in \mathbb{R}^n$. Το εσωτερικό γινόμενο των διανυσμάτων είναι:*

$$x^T y = \sum_{i=1}^{n} x_i y_i.$$

Ορισμός 1.21 (βλ. [1], Κεφάλαιο 10, Ορισμός, σελ. 279-280)**.** *Έστω πίνακας $A \in \mathbb{R}^{m \times n}$. Το μέγιστο πλήθος των γραμμικά ανεξάρτητων στηλών του λέγεται βαθμός στηλών του πίνακα A.*

Αντίστοιχα, το μέγιστο πλήθος των γραμμικά ανεξάρτητων γραμμών του πίνακα A λέγεται βαθμός γραμμών του πίνακα A.

Ορισμός 1.22 (βλ. [1], Κεφάλαιο 10, Ορισμός 10.2.1 (α), σελ. 280)**.** *Έστω πίνακας $A \in \mathbb{R}^{m \times n}$. Βαθμός ή τάξη του πίνακα A λέγεται το μέγιστο πλήθος των γραμμικά ανεξάρτητων στηλών του ή ισοδύναμα το μέγιστο πλήθος των γραμμικά ανεξάρτητων γραμμών του και συμβολίζεται με $rank(A)$.*

Ορισμός 1.23 (βλ. [2], Κεφάλαιο 3, Ορισμός, σελ. 77)**.** *Έστω πίνακας $A \in \mathbb{R}^{m \times n}$. Ο πίνακας A λέγεται πλήρους τάξης όταν ισχύει $rank(A) = \min\{m, n\}$.*

Ορισμός 1.24 (βλ. [2], Κεφάλαιο 3, Ορισμός, σελ. 77)**.** *Έστω πίνακας $A \in \mathbb{R}^{m \times n}$. Ο πίνακας A λέγεται πλήρους τάξης ως προς τις γραμμές του όταν ισχύει $rank(A) = m$.*

Αντίστοιχα, ο A λέγεται πλήρους τάξης ως προς τις στήλες του όταν ισχύει $rank(A) = n$.

Ορισμός 1.25 (βλ. [1], Κεφάλαιο 10, Ορισμός 10.4.1 (iv), σελ. 285)**.** *Έστω πίνακας $A \in \mathbb{R}^{m \times n}$. Το σύνολο των λύσεων του ομογενούς γραμμικού συστηματος $AX = 0$ που συμβολίζεται $Null(A)$ ή $\ker(A)$, είναι διανυσματικός υπόχωρος του $\mathbb{R}^n$ και λέγεται μηδενοχώρος ή πυρήνας του A. Δηλαδή το σύνολο:*

$$Null(A) = \{X \in \mathbb{R}^n \mid AX = 0\}.$$

1.5 Νόρμες

Ορισμός 1.26. *Έστω X διανυσματικός χώρος. Νόρμα λέγεται μία απεικόνιση $\|\cdot\| : X \longrightarrow \mathbb{R}$ αν ικανοποιεί τις εξής ιδιότητες:*

(i) $\|x\| \geq 0$, *για κάθε* $x \in X$

(ii) $\|x\| = 0 \iff x = 0$

(iii) $\|\lambda x\| = |\lambda|\|x\|$, *για κάθε* $x \in X$ *και* $\lambda \in \mathbb{R}$

(iv) $\|x + y\| \leq \|x\| + \|y\|$, *για κάθε* $x, y \in X$

Παραδείγματα νορμών διανυσμάτων

Για κάθε διάνυσμα $x \in \mathbb{R}^n$, ορίζουμε τις παρακάτω νόρμες:

1. Η 1-νόρμα: $\|x\|_1 = \sum_{i=1}^{n} |x_i|$

2. Η Ευκλείδεια νόρμα: $\|x\|_2 = \left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}} = (x^T x)^{\frac{1}{2}}$

3. Η $p-$νόρμα: $\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} = (x^T x)^{\frac{1}{p}}$

4. Η νόρμα άπειρο: $\|x\|_\infty = \max\limits_{i=1,\ldots,n} |x_i|$

Οι παραπάνω νόρμες μετράνε κατά μία έννοια το μήκος ενός διανύσματος και είναι ισοδύναμες καθώς ισχύει:

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty, \ \|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty.$$

Ορισμός 1.27. *Έστω $X \subset \mathbb{R}^n$, $x \in X$ και $M \in \mathbb{R}, M > 0$. Ορίζουμε τα σύνολα:*

(i) Ανοιχτή μπάλα κέντρου x και ακτίνας M

$$B(x, M) = \{y \in X \mid \|x - y\| < M\}.$$

(ii) Κλειστή μπάλα κέντρου x και ακτίνας M

$$B[x, M] = \{y \in X \mid \|x - y\| \leq M\}.$$

(iii) Σφαίρα κέντρου x και ακτίνας M

$$S(x, M) = \{y \in X \mid \|x - y\| = M\}.$$

Ορισμός 1.28 (βλ. [4], Κεφάλαιο 1, Ορισμός, σελ. 4)**.** *Η ανισότητα Cauchy-Schwarz για κάθε $x, y \in \mathbb{R}^n$ είναι:*

$$|x^T y| \leq \|x\|_2 \|y\|_2$$

Ορισμός 1.29 (βλ. [2], Κεφάλαιο 3, Ορισμός, σελ. 129)**.** *Έστω μία νόρμα $\|\cdot\|$ στον $\mathbb{R}^n$, πίνακας $A \in \mathbb{R}^{n\times n}$ και $x \in \mathbb{R}^n$. Ορίζουμε την φυσική νόρμα του πίνακα A ως την απεικόνιση:*

$$\|\cdot\| \ : \ \mathbb{R}^{n\times n} \longrightarrow \mathbb{R}, \ \|A\| = \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|}.$$

Παραδείγματα νορμών πινάκων

Για κάθε πίνακα $A \in \mathbb{R}^{n\times n}$, ορίζουμε τις παρακάτω νόρμες:

1. Η 1-νόρμα: $\|A\|_1 = \max_{j=1,\dots,n} \sum_{i=1}^n |a_{ij}|$.

2. Η Ευκλείδεια νόρμα: $\|A\|_2 = [\rho(A^T A)]^{\frac{1}{2}}$.

   Για την οποία ισχύει η ιδιότητα: $\|AB\|_2 \leq \|A\|_2 \ \|B\|_2$.

   Εάν ο πίνακας A είναι συμμετρικός και θετικά ορισμένος, τότε ισχύει: $\|A\|_2 = \rho(A)$.

3. Η $p-$νόρμα: $\|A\|_p = [\rho(A^T A)]^{\frac{1}{p}}$.

4. Η νόρμα άπειρο: $\|A\|_\infty = \max_{i=1,\dots,n} \sum_{j=1}^n |a_{ij}|$.

1.6 Κυρτότητα

Η έννοια της κυρτότητας εφαρμόζεται και σε σύνολα και σε συναρτήσεις και είναι πολύ σημαντική για την βελτιστοποίηση προβλημάτων.

Ορισμός 1.30. *Έστω σύνολο $S \subset \mathbb{R}^n$. Το S είναι κυρτό εάν:*

$$ax + (1-a)y \in S, \forall \ x, \ y \in S, a \in [0,1].$$

Παραδείγματα κυρτών συνόλων

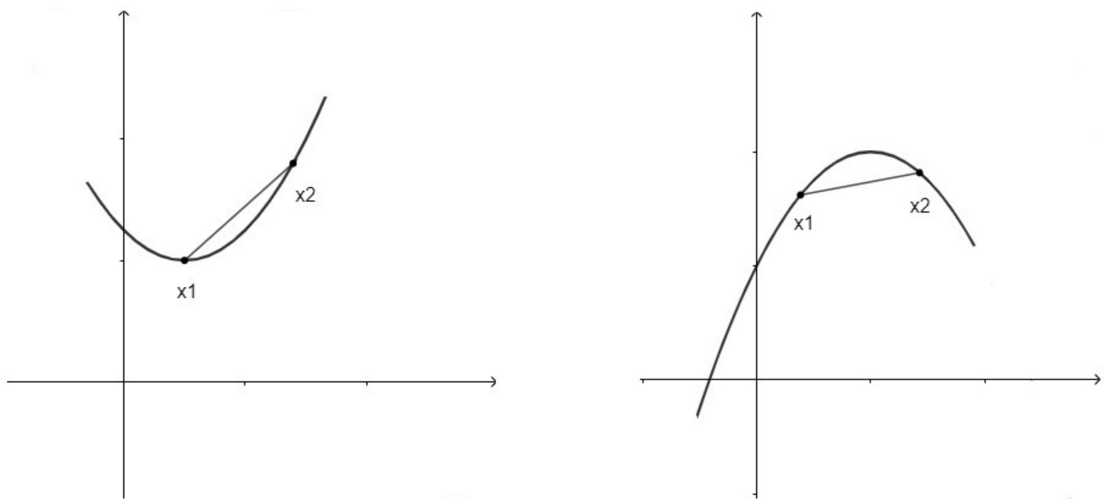
1. Για κάθε $x \in S$, το μονοσύνολο $\{x\}$ είναι κυρτό.

2. Ένα $A \subset \mathbb{R}$ είναι κυρτό αν και μόνο αν είναι διάστημα ανοιχτό ή κλειστό.

3. Αν $x, \ y \in \mathbb{R}^2$ με $x \neq y$, είναι γνωστό ότι το παρακάτω σύνολο ταυτίζεται με το ευθύγραμμο τμήμα με άκρα τα σημεία $x, \ y$:

$$[x, y\ ] = \{ax + (1-a)y : a \in [0,1]\}.$$

4. Το σύνολο $\{x \in \mathbb{R}^n : \|x\|_2 \leq 1\}$, δηλαδή η κλειστή μοναδιαία μπάλα του $\mathbb{R}^n$, είναι κυρτό υποσύνολο του $\mathbb{R}^n$.



Σχήμα 1.1: Κυρτή συνάρτηση (αριστερά) και κοίλη συνάρτηση (δεξιά)

Ορισμός 1.31. *Έστω συνάρτηση $f : S \subset \mathbb{R}^n \longrightarrow \mathbb{R}$. Η f είναι κυρτή συνάρτηση εάν το S είναι κυρτό και:*

$$f[ax + (1-a)y] \leq af(x) + (1-a)f(y), \forall\ x,\ y \in S, a \in [0,1].$$

Αντίστοιχα, η f είναι κοίλη συνάρτηση εάν η $-f$ είναι κυρτή ή ισοδύναμα εάν το S είναι κυρτό και:

$$f[ax + (1-a)y] \geq af(x) + (1-a)f(y), \forall\ x,\ y \in S, a \in [0,1].$$

Παραδείγματα κυρτών συναρτήσεων

1. Η γραμμική συνάρτηση $f(x) = c^T x + a$, για κάθε διάνυσμα $c \in \mathbb{R}^n$ και $a \in \mathbb{R}$.

2. Η τετραγωνική συνάρτηση $f(x) = x^T Hx$, όπου H είναι ένας συμμετρικός θετικά ημιορισμένος πίνακας.

Ορισμός 1.32. *Η συνάρτηση f είναι αυστηρά κυρτή συνάρτηση εάν:*

$$f[ax + (1-a)y] < af(x) + (1-a)f(y), \forall\ x,\ y \in S, \textit{με } x \neq y \textit{ και } a \in (0,1).$$

Ορισμός 1.33 (βλ. [5], Κεφάλαιο 6, Ορισμός, σελ. 129-130)**.** *Έστω συνάρτηση $f : S \subset \mathbb{R}^n \longrightarrow \mathbb{R}$. Η f είναι σχεδόν κυρτή συνάρτηση εάν:*

$$f[ax + (1-a)y] \leq \max\{f(x), f(y)\}, \forall\ x,\ y \in S, a \in [0,1].$$

Αντίστοιχα, η f είναι αυστηρά σχεδόν κυρτή συνάρτηση εάν:

$$f[ax + (1-a)y] < \max\{f(x), f(y)\}, \forall\ x,\ y \in S, \textit{με } f(x) \neq f(y) \textit{ και } a \in (0,1).$$

Παρατηρήσεις

1. Αν η f είναι κυρτή, τότε είναι προφανώς σχεδόν κυρτή.

2. Αν η f είναι αύξουσα ή φθίνουσα, τότε είναι σχεδόν κυρτή.

3. Αν η f είναι φθίνουσα για $x \leq x_0$ και αύξουσα για $x \geq x_0$, για κάποιο $x_0 \in S$, τότε είναι σχεδόν κυρτή.

4. Στην πράξη, μία συναρτηση f είναι κυρτή και σχεδόν κυρτή κατά τμήματα.

1.7 Πιεστικές συναρτήσεις, συνέχεια, συμπάγεια

Ορισμός 1.34. *Έστω $\|\cdot\|$ μία νόρμα στο $\mathbb{R}^n$. Ένα συνολο $S \subset \mathbb{R}^n$ λέγεται φραγμένο αν υπάρχει $M > 0$ τέτοιο ώστε:*

$$\|x\| \leq M, \text{ για } x \in S.$$

Ορισμός 1.35 (βλ. [6], Κεφάλαιο 1, Ορισμός, σελ. 26)**.** *Μία συνεχής συνάρτηση $f : S \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, με S μη φραγμένο, λέγεται πιεστική (στο S) αν:*

$$f(x) \to +\infty, \text{ όταν } \|x\| \to +\infty, \text{ για } x \in S.$$

Παράδειγμα 1.7.1. *Η συνάρτηση $f(x,y) = x^2 + y^2$ είναι πιεστική καθώς:*

$$f(z) = \|z\|^2 \to +\infty, \text{ όταν } \|z\| \to +\infty.$$

Ορισμός 1.36. *Ένα συνολο $S \subset \mathbb{R}^n$ λέγεται ανοιχτό, αν για κάθε $x \in S$ υπάρχει $\delta > 0$ τέτοιο ώστε:*

$$y \in S, \text{ για } \|y - x\| < \delta.$$

Ορισμός 1.37. *Ένα συνολο $S \subset \mathbb{R}^n$ λέγεται κλειστό, αν το συμπλήρωμά του S^c είναι ανοιχτό.*

Υπενθύμιση

Ένα συνολο $S \subset \mathbb{R}^n$ είναι κλειστό αν και μόνο αν περιέχει το όριο κάθε συγκλίνουσας ακολουθίας στοιχείων του S.

Ορισμός 1.38 (βλ. [7], Κεφάλαιο 2, Ορισμός 2.5-1, σελ. 77)**.** *Ένα συνολο $S \subset \mathbb{R}^n$ λέγεται συμπαγές, αν κάθε ακολουθία στοιχείων του S περιέχει μία υπακολουθία που συγκλίνει σε ένα στοιχείο του S.*

Υπενθύμιση

Σε χώρους πεπερασμένων διαστάσεων, ένα συνολο $S \subset \mathbb{R}^n$ είναι συμπαγές αν και μόνο αν είναι κλειστό και φραγμένο.

Ορισμός 1.39. *Μία συνάρτηση $f : S \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ είναι συνεχής στο $x_0 \in S$ αν:*

$$\lim_{x \to x_0} f(x) = f(x_0), \text{ για } x \in S.$$

Επίσης, μία συνάρτηση $f : S \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ είναι συνεχής στο S αν είναι συνεχής σε κάθε σημείο του S.

1.8 Παράγωγοι συναρτήσεων

Ορισμός 1.40 (βλ. [8], Κεφάλαιο 5, Ορισμός, σελ. 198). *Η συνάρτηση $f : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ λέγεται παραγωγίσιμη στο σημείο $x \in \Omega$ αν ισχύει:*

$$f(x+h) - f(x) = v(x)^T h + \varepsilon(h)\|h\|$$

όπου $v(x) \in \mathbb{R}^n$ και $\varepsilon(h) \to 0$ όταν $\|h\| \to 0$.

Ορισμός 1.41. *Αν η συνάρτηση f είναι παραγωγίσιμη σε κάθε σημείο $x \in S \subset \Omega$, τότε λέμε ότι η f είναι παραγωγίσιμη στο S.*

Παρατηρήσεις

1. Αν η f είναι παραγωγίσιμη στο $x \in \Omega$, τότε είναι συνεχής στο $x \in \Omega$.
2. Η ύπαρξη μερικών παραγώγων της f συνεχών στο σημείο x μόνο, δεν εξασφαλίζει την παραγωγισιμότητα της f στο x.
3. Αν η $f : \Omega \longrightarrow \mathbb{R}$ έχει μερικές παραγώγους μέχρι και την τάξη m συνεχείς στο Ω, τότε γράφουμε $f \in C^m(\Omega)$.
4. Έστω $V \subset \Omega$ ανοιχτό σύνολο. Αν $f \in C^1(V)$, τότε η f είναι παραγωγίσιμη στο V.

Ορισμός 1.42. *Έστω $\Omega \subset \mathbb{R}^n$ ένα ανοιχτό σύνολο. Αν η συνάρτηση $f : \Omega \longrightarrow \mathbb{R}$ έχει μερικές παραγώγους πρώτης τάξης στο σημείο $x \in \Omega$, τότε ορίζουμε ως κλίση (gradient) της f στο x να είναι:*

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x) \\ \frac{\partial f}{\partial x_2}(x) \\ \cdots \\ \cdots \\ \cdots \\ \frac{\partial f}{\partial x_n}(x) \end{bmatrix}.$$

Ορισμός 1.43 (βλ. [4], Κεφάλαιο 4, Ορισμός 3.1, σελ. 116-117). *Έστω $f : S \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ μία συνάρτηση και $x \in S$, $h \in \mathbb{R}^n$. Αν $x + ah \in S, \forall a \neq 0$ αρκετά μικρό και το όριο:*

$$\delta f(x,h) = \lim_{a \to 0} \frac{f(x+ah) - f(x)}{a} \text{ υπάρχει,}$$

τότε καλείται παράγωγος της f στο σημείο x κατά την κατεύθυνση του h.

Ορισμός 1.44. *Έστω $f : S \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ μία συνάρτηση και $x \in S$, $h \in \mathbb{R}^n$. Αν $x + ah \in S, \forall a \neq 0$ αρκετά μικρό και το όριο:*

$$\delta_+ f(x,h) = \lim_{\substack{a \to 0 \\ a > 0}} \frac{f(x+ah) - f(x)}{a} \text{ υπάρχει,}$$

τότε καλείται παράγωγος της f στο σημείο x θετικά κατά την κατεύθυνση του h.

Παρατηρήσεις

1. Αν η συνάρτηση $f : \Omega \longrightarrow \mathbb{R}$ είναι παραγωγίσιμη στο σημείο $x \in \Omega$ και Ω ανοιχτό, τότε οι παράγωγοι $\delta f(x,h)$ και $\delta_+ f(x,h)$ υπάρχουν στο σημείο x για κάθε $h \in \mathbb{R}^n$ και ισχύει:
$$\delta_+ f(x,h) = \delta f(x,h) = \nabla f(x)^T h.$$

2. Έστω f παραγωγίσιμη στο $x \in \Omega \subset \mathbb{R}^n$ με Ω ανοιχτό και $h \in \mathbb{R}^n$ με $\|h\|_2 = 1$, τότε από την ανισότητα Cauchy-Schwarz έχουμε ότι:
$$|\delta f(x,h)| = |\nabla f(x)^T h| \leq \|\nabla f(x)\|_2 \|h\|_2 = \|\nabla f(x)\|_2.$$
Έστω το μοναδιαίο διάνυσμα $u := \dfrac{\nabla f(x)}{\|\nabla f(x)\|_2}$, τότε:
$$\delta f(x,u) = \nabla f(x)^T u = \nabla f(x)^T \frac{\nabla f(x)}{\|\nabla f(x)\|_2} = \nabla f(x)^T \nabla f(x) \frac{1}{\|\nabla f(x)\|_2}$$
$$= \frac{\|\nabla f(x)\|_2^2}{\|\nabla f(x)\|_2} = \|\nabla f(x)\|_2.$$
Επομένως, για κάθε μοναδιαίο διάνυσμα $h \in \mathbb{R}^n$ έχουμε:
$$\delta f(x,h) \leq \delta f(x,u) = \|\nabla f(x)\|_2.$$
Δηλαδή η κατεύθυνση u του $\nabla f(x)$ είναι αυτή κατά την οποία η f έχει τον μεγαλύτερο ρυθμό αύξησης $\delta f(x,h)$ κοντά στο σημείο x.

Ορισμός 1.45 (βλ. [4], Κεφάλαιο 4, Ορισμός, σελ. 144)**.** *Αν η συνάρτηση $f : \Omega \longrightarrow \mathbb{R}$ έχει μερικές παραγώγους δεύτερης τάξης:*
$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x), i, j = 1, \ldots, n$$
στο σημείο $x \in \Omega$, ορίζουμε τον Εσσιανό πίνακα της f στο x ως:
$$\nabla^2 f(x) = f_{xx}(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(x) \\ \vdots & & \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n}(x) \end{bmatrix}.$$

Παρατηρήσεις

1. Αν $f \in C^2$, τότε $f \in C^1$.

2. Αν $f \in C^2$, τότε ο πίνακας $f_{xx}(x)$ είναι συμμετρικός για κάθε $x \in \Omega$.

1.9 Συνθήκες για την ύπαρξη ελαχίστου

Θεώρημα 1.1 (βλ. [8], Κεφάλαιο 5, Θεώρημα 3, σελ. 197). *Έστω συνάρτηση $f : S \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, με το σύνολο S κυρτό και την f κυρτή. Το $x^* \in S$ είναι σημείο ολικού ελαχίστου της f στο S, αν και μόνο αν είναι σημείο τοπικού ελαχίστου της f στο S.*

Απόδειξη. Έστω $x^* \in S$ σημείο τοπικού ελαχίστου της f, $\forall\ x \in S$. Αφού το S είναι κυρτό, από τον ορισμό έχουμε:

$$ax + (1-a)x^* = x^* + a(x - x^*) \in S, \forall\ a \in (0,1). \tag{1}$$

Αφού η f είναι κυρτή, από τον ορισμό έχουμε:

$$f[ax + (1-a)x^*] \leq af(x) + (1-a)f(x^*) = f(x^*) + a[f(x) - f(x^*)]. \tag{2}$$

Επειδή $a \neq 0$ και x^* τοπικό ελάχιστο, από τον ορισμό έχουμε:

$$\exists\ \delta > 0 \text{ τέτοιο ώστε } f(x^*) \leq f(y), \forall\ y \in S \text{ με } \|y - x^*\| < \delta. \tag{3}$$

Επομένως, για κατάλληλο a από την (1): $y = x^* + a(x - x^*) \in S$.
Άρα απο την (3):

$$\|x^* - [x^* + a(x - x^*)]\| \leq |a|\ \|x - x^*\| < \delta$$

και

$$f(y) = f[x^* + a(x - x^*)] \geq f(x^*).$$

Επομένως:

$$f(x^*) \leq f[x^* + a(x - x^*)] = f[ax + (1-a)x^*] \overset{(2)}{\leq} f(x^*) + a[f(x) - f(x^*)]$$

$$\Longrightarrow 0 \leq a[f(x) - f(x^*)] \overset{a>0}{\Longrightarrow} 0 \leq f(x) - f(x^*) \Longrightarrow f(x^*) \leq f(x), \forall x \in S.$$

Επομένως, το x^* είναι σημείο ολικού ελαχίστου της f στο S. □

Θεώρημα 1.2 (βλ. [8], Κεφάλαιο 5, Θεώρημα 4, σελ. 198). *Έστω συνάρτηση $f : S \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, με το σύνολο S κυρτό και την f αυστηρά κυρτή. Τότε η f έχει το πολύ ένα σημείο ελαχίστου.*

Απόδειξη. Έστω x_1^*, $x_2^* \in S$ σημεία ελαχίστου της f, $\forall\ x \in S$, με $x_1^* \neq x_2^*$. Αφού το S είναι κυρτό, για $a = \frac{1}{2}$ από τον ορισμό έχουμε:

$$\frac{1}{2}\ x_1^* + (1 - \frac{1}{2})\ x_2^* = \frac{1}{2}\ x_1^* + \frac{1}{2}\ x_2^* \in S.$$

Αφού η f είναι αυστηρά κυρτή, για $a = \frac{1}{2}$ από τον ορισμό έχουμε:

$$f(\frac{1}{2}\ x_1^* + \frac{1}{2}\ x_2^*) < \frac{1}{2}f(x_1^*) + \frac{1}{2}f(x_2^*) \leq \frac{1}{2}f(x_1^*) + \frac{1}{2}f(x_1^*) = f(x_1^*)$$

$$\Longrightarrow f(\frac{1}{2}\ x_1^* + \frac{1}{2}\ x_2^*) < f(x_1^*).$$

Αυτό όμως είναι άτοπο, αφού το x_1^* είναι σημείο ελαχίστου της f στο S. Άρα, η f έχει το πολύ ένα σημείο ελαχίστου. □

Θεώρημα 1.3 (βλ. [8], Κεφάλαιο 5, Θεώρημα 1, σελ. 197). *Έστω $f : S \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ μία συνάρτηση συνεχής στο μη κενό, συμπαγές σύνολο S. Τότε η f έχει τουλάχιστον ένα σημείο ελαχίστου στο S.*

Παράδειγμα 1.9.1. *Έστω η γραμμική συνάρτηση $f : S \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, με $f(x) = c^T x + a$, $c \in \mathbb{R}^n, a \in \mathbb{R}$ και $S = \{y \in \mathbb{R}^n : \|x - y\| \leq r\}$, δηλαδή η κλειστή μπάλα κέντρου x και ακτίνας r. Θα δείξουμε ότι η f έχει τουλάχιστον ένα σημείο ελαχίστου στο S.*

Παρατηρούμε ότι το σύνολο S είναι προφανώς κλειστό. Θα δείξουμε ότι το S είναι φραγμένο. Πράγματι, για κάθε $y \in S$ υπάρχει $x \in S \subset \mathbb{R}^n$ και $r > 0$ τέτοιο ώστε να ισχύει $\|y - x\| \leq r$. Άρα, από τον ορισμό του φραγμένου συνόλου το S είναι φραγμένο. Επομένως, αφού το S είναι κλειστό και φραγμένο, τότε είναι συμπαγές.

Η συνάρτηση f είναι συνεχής στο S ως γραμμική. Επομένως, από το Θεώρημα 1.3 η f έχει τουλάχιστον ένα σημείο ελαχίστου στο S.

Θεώρημα 1.4 (βλ. [8], Κεφάλαιο 5, Θεώρημα 2, σελ. 197). *Έστω $S \subset \mathbb{R}^n$ ένα μη κενό, κλειστό και μη φραγμένο σύνολο, και $f : S \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ μία συνάρτηση συνεχής και πιεστική. Τότε η f έχει τουλάχιστον ένα σημείο ελαχίστου στο S.*

Απόδειξη. Έστω $x_0 \in S$. Επειδή η f είναι πιεστική, $\exists\ r > 0$ τέτοιο ώστε:

$$f(x) > f(x_0), \text{ για } \|x\| > r, x \in S.$$

Έστω $S_0 = \{x \in S : \|x\| \leq r\} = S \cap \{x \in \mathbb{R}^n : \|x\| \leq r\}$. Αφού το S και το $\{x \in \mathbb{R}^n : \|x\| \leq r\}$ είναι κλειστά σύνολα, τότε και το S_0 θα είναι κλειστό και φραγμένο ως υποσύνολο του $\mathbb{R}^n$. Επομένως, το S_0 είναι συμπαγές σύνολο.

Η ελαχιστοποίηση της f στο S ισοδυναμεί, τότε, με την ελαχιστοποίηση της f στο S_0.

Επομένως, από το Θεώρημα 1.3 έχουμε ότι η f θα έχει τουλάχιστον ένα σημείο ελαχίστου στο S_0. □

Παράδειγμα 1.9.2. *Έστω η τετραγωνική συνάρτηση $f : S \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, με $f(x) = x^T H x$, όπου H είναι ένας συμμετρικός θετικά ορισμένος πίνακας και $S = \{x = [x_1, x_2, \ldots, x_n]^T \in \mathbb{R}^n : x_1, x_2, \ldots, x_n \geq 0\}$. Θα δείξουμε ότι η f έχει τουλάχιστον ένα σημείο ελαχίστου στο S.*

Το σύνολο S είναι κλειστό. Πράγματι, το συμπληρωματικό σύνολο του S, $S^c = \mathbb{R}^n \backslash S = \{x = [x_1, x_2, \ldots, x_n]^T \in \mathbb{R}^n : x_1, x_2, \ldots, x_n < 0\}$ είναι ανοιχτό, καθώς για κάθε $x \in S^c$ υπάρχει $\delta > 0$ τέτοιο ώστε $\|x - 0\| < \delta$. Επομένως, το σύνολο S είναι κλειστό.

Το σύνολο S είναι μη φραγμένο. Πράγματι, θα δείξουμε ότι δεν υπάρχει ανοιχτή μπάλα που να το περιέχει. Έστω ότι υπάρχει $r > 0$ τέτοιο ώστε $S \subseteq B(0, r)$. Έστω, επίσης, $z \in S$ με $z = [z_1, \ldots, z_n]^T = [r+1, \ldots, r+1]^T$. Τότε έχουμε:

$$\begin{aligned}
\|z - 0\| &= \sqrt{(r+1)^2 + \ldots + (r+1)^2} \\
&= \sqrt{n(r+1)^2} \\
&= \sqrt{n}|r+1| > r, \forall r > 0 \\
&\implies z \notin B(0, r), \forall r > 0.
\end{aligned}$$

Όμως, το $z \in S, \forall r > 0$, *άρα* $S \nsubseteq B(0,r), \forall r > 0$. *Επομένως, το* S *δεν περιέχεται σε καμία ανοιχτή μπάλα και άρα είναι μη φραγμένο.*

Η συνάρτηση f *είναι πιεστική στο* S. *Πράγματι, έστω* $S_0 = \{x \in \mathbb{R}^n : \|x\| = 1\} \subset \mathbb{R}^n$. *Το σύνολο* S_0 *είναι κλειστό και φραγμένο και καθώς η* f *είναι συνεχής, τότε από το Θεώρημα 1.3 η* f *θα έχει τουλάχιστον ένα σημείο ελαχίστου στο* S_0. *Έστω* $x^* \in S_0$.

Θέτουμε $a = f(x^*) = (x^*)^T H x^* > 0$, *αφού ο πίνακας* H *είναι θετικά ορισμένος. Τότε, για κάθε* $x \in \mathbb{R}^n$ *για τα οποία ισχύει* $\|x\| = 1$, *έχουμε ότι* $f(x) \geq a$.

Έστω τυχαίο $x \in \mathbb{R}^n, x \neq 0$. *Τότε:*

$$f(x) = x^T H x = \left(\frac{x}{\|x\|}\right)^T H \left(\frac{x}{\|x\|}\right) \|x\|^2 \geq a\|x\|^2 \implies f(x) \geq a\|x\|^2.$$

Θα δείξουμε ότι η συνάρτηση $a\|x\|^2$ *είναι πιεστική στο* $\mathbb{R}^n$. *Πράγματι:*

$$\lim a\|x\|^2 = +\infty, \quad \text{καθώς} \ \ \|x\| \to +\infty.$$

Άρα, από την παραπάνω ανισότητα και η συνάρτηση f *είναι πιεστική στο* $\mathbb{R}^n$. *Επομένως, από το Θεώρημα 1.4 η* f *έχει τουλάχιστον ένα σημείο ελαχίστου στο* S.

Θεώρημα 1.5 (Μέσης Τιμής, βλ. [8], Κεφάλαιο 5, Θεώρημα 5, σελ. 200)**.** *Έστω συνάρτηση* $f : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$. *Αν* $x + ah \in \Omega$, *για κάθε* $a \in [0,1]$, *η* f *είναι συνεχής σε κάθε σημείο* $x + ah, a \in [0,1]$ *και παράγωγίσιμη σε κάθε σημείο* $x + ah, a \in (0,1)$, *τότε υπάρχει* $t \in (0,1)$ *τέτοιο ώστε:*

$$f(x+h) = f(x) + \nabla f(x+th)^T h.$$

Θεώρημα 1.6 (Θεώρημα Taylor, βλ. [8], Κεφάλαιο 5, Θεώρημα 6, σελ. 200)**.** *Έστω συνάρτηση* $f : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, $f \in C^2$. *Αν* $x + ah \in \Omega$, *για κάθε* $a \in [0,1]$, *τότε υπάρχει* $t \in (0,1)$ *τέτοιο ώστε:*

$$f(x+h) = f(x) + \nabla f(x)^T h + \frac{1}{2} h^T f_{xx}(x+th) h.$$

Επιπλέον, για $\|h\|$ *αρκετά μικρό, έχουμε ότι* $x + h \in \Omega$ *και:*

$$f(x+h) = f(x) + \nabla f(x)^T h + \frac{1}{2} h^T f_{xx}(x) h + \varepsilon(h)\|h\|^2,$$

όπου $\varepsilon(h) \to 0$, *όταν* $\|h\| \to 0$.

Θεώρημα 1.7 (Αναγκαία Συνθήκη Πρώτης Τάξης, βλ. [8], Κεφάλαιο 5, Θεώρημα 7, σελ. 200)**.** *Έστω συνάρτηση* $f : S \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, *με* S *κυρτό και* $x^* \in S$. *Αν η παράγωγος* $\delta_+ f(x^*, x - x^*)$ *υπάρχει για κάθε* $x \in S$ *και το* x^* *είναι σημείο τοπικού ή ολικού ελαχίστου της* f *στο* S, *τότε ισχύει η συνθήκη:*

$$\delta_+ f(x^*, x - x^*) \geq 0, \ \text{για κάθε} \ x \in S. \tag{1.2}$$

Αν επιπλέον $S \subset \Omega$, Ω *ανοιχτό,* $f : \Omega \longrightarrow \mathbb{R}$ *και η* f *είναι παραγωγίσιμη στο* x^*, *τότε η συνθήκη (1.2) γράφεται:*

$$\nabla f(x^*)^T (x - x^*) \geq 0, \ \text{για κάθε} \ x \in S. \tag{1.3}$$

Απόδειξη. Αν το x^* είναι σημείο τοπικού ελαχίστου της f στο S, επειδή το S είναι κυρτό τότε: $y = ax + (1-a)x^* \in S, \forall x \in S, a \in [0,1]$. Επομένως, αν $a \in (0,1)$ αρκετά μικρό, από τον ορισμό του τοπικού ελαχίστου υπάρχει $\delta > 0$ τέτοιο ώστε:

$$f(y) = f[ax + (1-a)x^*] = f[x^* + a(x - x^*)] \geq f(x^*)$$

με $\|y - x^*\| = \|x^* + a(x-x^*) - x^*\| = |a|\|x - x^*\| = a\|x - x^*\| < \delta$.
Άρα

$$\lim_{\substack{a \to 0 \\ a > 0}} \frac{f[x^* + a(x-x^*)] - f(x^*)}{a} \geq 0$$

δηλαδή

$$\delta_+ f(x^*, x - x^*) \geq 0, \text{ για κάθε } x \in S.$$

Αν το x^* είναι ολικό ελάχιστο, τότε είναι και τοπικό ελάχιστο επομένως η απόδειξη ισχύει.

Με τις πρόσθετες υποθέσεις του θεωρήματος η συνθήκη αυτή γράφεται:

$$\nabla f(x^*)^T (x - x^*) \geq 0, \text{ για κάθε } x \in S.$$

□

Παρατηρήσεις

1. Αν ο χώρος S είναι κατάλληλος ή πιο γενικά αν υπάρχει $\delta > 0$ τέτοιο ώστε:

   $$x \in S \text{ και } \|x - x^*\| < \delta \Rightarrow x' = x^* + (x^* - x) \in S,$$

   δηλαδή αν το S είναι συμμετρικό ως προς το x^*, κοντά στο x^*, τότε για $x = x'$ η συνθήκη (1.3) από το Θεώρημα 1.7 ισοδυναμεί με την:

   $$\nabla f(x^*)^T (x - x^*) = 0, \text{ για κάθε } x \in S.$$

2. Αν $S = \mathbb{R}^n$ ή πιο γενικά αν υπάρχει $\delta > 0$ τέτοιο ώστε:

   $$\|x - x^*\| < \delta \Rightarrow x \in S,$$

   δηλαδή αν $x \in S$ για x κοντά στο x^*, τότε η συνθήκη (1.3) από το Θεώρημα 1.7 ισοδυναμεί με την:

   $$\nabla f(x^*) = 0. \tag{1.4}$$

Θεώρημα 1.8 (Ικανή Συνθήκη Πρώτης Τάξης, βλ. [8], Κεφάλαιο 5, Θεώρημα 8, σελ. 202)**.** *Έστω συνάρτηση $f : S \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, με S κυρτό και $x^* \in S$. Αν η παράγωγος $\delta_+ f(x^*, x - x^*)$ υπάρχει για κάθε $x \in S$ και αν η f είναι κυρτή στο S, τότε η αναγκαία συνθήκη (1.2) από το Θεώρημα 1.7 είναι ικανή και αναγκαία για να είναι το x^* σημείο ελαχίστου της f στο S.*

Απόδειξη. Αν η f είναι κυρτή στο S, τότε για $x \in S$ και $a \in (0,1)$ αφού το S κυρτό έχουμε:

$$ax + (1-a)x^* = x^* + a(x - x^*) \in S.$$

Επομένως

$$\begin{aligned} f[ax+(1-a)x^*] = f[x^*+a(x-x^*)] &\le f(x^*)+a[f(x)-f(x^*)] \\ \Longrightarrow f[x^*+a(x-x^*)]-f(x^*) &\le a[f(x)-f(x^*)] \\ \Longrightarrow \frac{f[x^*+a(x-x^*)]-f(x^*)}{a} &\le f(x)-f(x^*), \end{aligned}$$

άρα

$$\lim_{\substack{a\to 0\\ a>0}} \frac{f[x^*+a(x-x^*)]-f(x^*)}{a} = \delta_+ f(x^*, x-x^*) \le f(x)-f(x^*),$$

που δείχνει ότι η αναγκαία συνθήκη (1.2) είναι και ικανή για να είναι το x^* σημείο ελαχίστου της f στο S. □

Ορισμός 1.46. *Ένα σημείο x^* καλείται στάσιμο σημείο εάν $\nabla f(x^*) = 0$.*

Παρατήρηση

Αν $S = \mathbb{R}^n$ ή πιο γενικά αν ένα σύνολο ικανοποιεί την συνθήκη (1.4), τότε από την (1.4) και το Θεώρημα 1.7 κάθε τοπικό ελάχιστο πρέπει να είναι στάσιμο σημείο.

Θεώρημα 1.9 (Αναγκαία Συνθήκη Δεύτερης Τάξης, βλ. [9], Κεφάλαιο 2, Θεώρημα 2.3, σελ. 15). *Εαν το σημείο x^* είναι τοπικό ελάχιστο της συνάρτησης $f : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ και ο Εσσιανός πίνακας $\nabla^2 f$ υπάρχει και είναι συνεχής κοντά στο x^*, τότε $\nabla f(x^*) = 0$ και ο $\nabla^2 f(x^*)$ είναι θετικά ημιορισμένος.*

Απόδειξη. Από την συνθήκη (1.4) και το Θεώρημα 1.7 ξέρουμε ότι $\nabla f(x^*) = 0$. Υποθέτουμε ότι ο $\nabla^2 f(x^*)$ δεν είναι θετικά ημιορισμένος. Διαλέγουμε ένα διάνυσμα h για το οποίο ισχύει: $h^T\nabla^2 f(x^*)h < 0$ και επειδή $\nabla^2 f$ είναι συνεχής κοντά στο x^*, υπάρχει $T > 0$ τέτοιο ώστε:

$$h^T\nabla^2 f(x^*+th)h < 0, \forall t \in [0,T].$$

Από το Θεώρημα Taylor υπάρχει $\bar{t} \in (0,T]$ και $t \in (0,\bar{t})$ τέτοια ωστε:

$$f(x^*+\bar{t}h) = f(x^*) + \bar{t}h^T\nabla f(x^*) + \frac{1}{2}\bar{t}^2 h^T\nabla^2 f(x^*+th)h < f(x^*).$$

Άρα το x^* δεν είναι τοπικό ελάχιστο. Άτοπο, επομένως ο $\nabla^2 f(x^*)$ είναι θετικά ημιορισμένος. □

Θεώρημα 1.10 (Ικανή Συνθήκη Δεύτερης Τάξης, βλ. [8], Κεφάλαιο 5, Θεώρημα 9, σελ. 203). *Έστω συνάρτηση $f : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, $x^* \in \Omega$, $f \in C^2$ κοντά στο x^*. Αν $\nabla f(x^*) = 0$ και ο Εσσιανός πίνακας $\nabla^2 f(x^*)$ είναι θετικά ορισμένος, τότε το x^* είναι σημείο τοπικού ελαχίστου της f στο Ω.*

Απόδειξη. Από το Θεώρημα Taylor, για $\|h\|$ αρκετά μικρό και $h \neq 0$, έχουμε ότι $x+h \in \Omega$ και:

$$f(x^*+h) - f(x^*) = \frac{1}{2}\|h\|^2[u^T\nabla^2 f(x^*)u + \varepsilon(h)],$$

όπου $u = \dfrac{h}{\|h\|}$ και $\varepsilon(h) \to 0$, όταν $\|h\| \to 0$.

Η συνάρτηση $u \longrightarrow u^T \nabla^2 f(x^*) u$ είναι συνεχής στο συμπαγές σύνολο $S = \{u : \|u\| = 1\}$.

Επομένως, από το Θεώρημα 1.3 υπάρχει $\bar{u} \in S$ με $\bar{u}^T \nabla^2 f(x^*) \bar{u} = c > 0$, για $u \in S$, αφού ο πίνακας $\nabla^2 f(x^*)$ είναι θετικά ορισμένος. Άρα:

$$x^* + h \in \Omega \text{ και } f(x^* + h) - f(x^*) > 0,$$

για $\|h\|$ αρκετά μικρό, $h \neq 0$, που δείχνει ότι το x^* είναι σημείο τοπικού ελαχίστου της f στο Ω. □

Σημείωση

Δοσμένων δύο μη αρνητικών ακολουθιών $\{t_k\}$ και $\{n_k\}$, θα χρησιμοποιούμε τους συμβολισμούς $o(\cdot)$ και $O(\cdot)$ όταν:

- Η ακολουθία $\left\{\dfrac{n_k}{t_k}\right\} \to 0$, για $k \to 0$, τότε γράφουμε $n_k = o(t_k)$.
- Υπάρχει $C > 0$ τέτοιο ώστε να ισχύει $|n_k| \leq C|t_k|$, τότε γράφουμε $n_k = O(t_k)$.

1.10 Αλγόριθμοι βελτιστοποίησης

Στην πράξη πολλές φορές δεν είναι δυνατή η εύρεση της λύσης ενός προβλήματος υπολογιστικά, καθώς οι συναρτήσεις από τις οποίες αποτελείται μπορεί να είναι ιδιαίτερα πολύπλοκες. Για τον λόγο αυτό χρησιμοποιούμε προσεγγιστικές μεθόδους, δηλαδή επαναληπτικούς αλγορίθμους που προσεγγίζουν την λύση. Η επίλυση ξεκινάει με μία τιμή της μεταβλητής x, όπου μέσα από αναδρομικούς τύπους, παράγονται προσεγγίσεις της λύσης x^* σε κάθε επανάληψη, μέχρι ο αλγόριθμος να τερματίσει με την καλύτερη προσέγγιση της λύσης του προβλήματος. Ο κάθε αλγόριθμος χρησιμοποιεί διαφορετικές τεχνικές για να πηγαίνει από την μία επανάληψη στην άλλη και τελικά για την προσέγγιση της λύσης. Κάποιες από αυτές τις τεχνικές, κάνουν χρήση των τιμών της αντικειμενικής συνάρτησης f, των περιορισμών c_i καθώς και των παραγώγων των συναρτήσεων αυτών. Κάποιοι αλγόριθμοι χρησιμοποιούν τα δεδομένα από όλες τις επαναλήψεις, ενώ κάποιοι άλλοι μόνο της τελευταίας. Σε κάθε περίπτωση, όμως, οι αλγόριθμοι πρέπει να έχουν τα εξής χαρακτηριστικά:

- **Ευρωστία (Robustness).** Να μπορούν να εφαρμόζονται σε ένα μεγάλο εύρος προβλημάτων της τάξης τους, για διάφορες αρχικές τιμές.
- **Αποδοτικότητα (Efficiency).** Οι απαιτήσεις τους σε υπολογιστικούς πόρους, δηλαδή ο απαιτούμενος χρόνος εκτέλεσης και ο απαιτούμενος χώρος αποθήκευσης, να είναι μικρές.
- **Ακρίβεια (Accuracy).** Η προσέγγιση της λύσης να είναι όσο το δυνατόν πιο ακριβής, χωρίς να είναι υπερβολικά ευαίσθητοι στα σφάλματα των δεδομένων και στα σφάλματα στρογγύλευσης που προκύπτουν από την εισαγωγή του αλγορίθμου στον υπολογιστή.

Μερικές φορές δεν είναι εφικτό ένας αλγόριθμος να έχει και τα τρία χαρακτηριστικά ταυτόχρονα, καθώς είναι πιθανό να έρχονται σε σύγκρουση. Για παράδειγμα, μπορεί ένας αλγόριθμος να έχει ένα μεγάλο εύρος εφαρμογής, αλλά για αυτόν τον λόγο να απαιτεί αρκετό χρόνο εκτέλεσης. Στην βελτιστοποίηση, τα πλεονεκτήματα και τα μειονεκτήματα του κάθε αλγορίθμου εξετάζονται πολύ προσεκτικά ώστε να γίνει και η βέλτιστη επιλογή για το εκάστοτε πρόβλημα.

Βιβλιογραφικές αναφορές

Η εισαγωγή του Κεφαλαίου 1 βασίστηκε στην εισαγωγή του Κεφαλαίου 1 του βιβλίου Numerical Optimization [βλ. [9], Κεφάλαιο 1, σελ. 2].

Ενότητα 1.1: Η ενότητα αυτή βασίστηκε στο Κεφάλαιο Μέθοδοι Βελτιστοποίησης του βιβλίου Εισαγωγή στην Αριθμητική Ανάλυση [βλ. [8], Κεφάλαιο 5, σελ. 197, 203], στην ενότητα Mathematical Formulation του βιβλίου Numerical Optimization [βλ. [9], Κεφάλαιο 1, σελ. 2-4], καθώς επίσης και στην ενότητα Εισαγωγή - Ορισμοί, του Κεφαλαίου Ακρότατα, του βιβλίου Ανάλυση ΙΙ [βλ. [4], Κεφάλαιο 6, σελ. 226].

Ενότητα 1.2: Η ενότητα αυτή βασίστηκε στην ενότητα Constrained and Unconstrained Optimization του βιβλίου Numerical Optimization [βλ. [9], Κεφάλαιο 1, σελ. 6].

Ενότητα 1.3: Η ενότητα αυτή βασίστηκε στις υποενότητες The Spectrum και Positive Definite and Nonnegative Definite Matrices του βιβλίου Matrix Algebra: Theory, Computations and Applications in Statistics [βλ. [2], Κεφάλαιο 3, σελ. 111, 124], καθώς επίσης και στις ενότητες Βασικοί Ορισμοί, Οι Πράξεις στο Σύνολο των $\mu \times \nu$ Πινάκων, Ιδιοτιμές και Ιδιοδιανύσματα, Ορθομοναδιαίοι και Ορθογώνιοι Πίνακες, και Θετικά Ορισμένες Τετραγωνικές Μορφές του βιβλίου Γραμμική Άλγεβρα και Αναλυτική Γεωμετρία [βλ. [1], Κεφάλαια 2, 11, 14, 15, σελ. 18, 22-25, 300, 426, 468].

Ενότητα 1.4: Η ενότητα αυτή βασίστηκε στην ενότητα Ορισμός Διανυσματικού Χώρου από τις Σημειώσεις Παραδόσεων Συναρτησιακής Ανάλυσης [βλ. [3], Κεφάλαιο 1, σελ. 8-10], στην ενότητα Matrix Rank and the Inverse of a Full Rank Matrix του βιβλίου Matrix Algebra: Theory, Computations and Applications in Statistics [βλ. [2], Κεφάλαιο 3, σελ. 76-78], καθώς επίσης και στα Κεφάλαια Διανυσματικοί Χώροι, Βαθμός Γραμμικής Απεικόνισης και Πίνακα, και Χώροι Εσωτερικού Γινομένου του βιβλίου Γραμμική Άλγεβρα και Αναλυτική Γεωμετρία [βλ. [1], Κεφάλαια 8, 10, 13, σελ. 195, 197, 203, 279-280, 285, 386].

Ενότητα 1.5: Η ενότητα αυτή βασίστηκε στην ενότητα Matrix Norms του βιβλίου Matrix Algebra: Theory, Computations and Applications in Statistics [βλ. [2], Κεφάλαιο 3, σελ. 128-134], στις ενότητες Ορισμός και Παραδείγματα, και Μπάλες και Σφαίρες σε Χώρους με Νόρμα από τις Σημειώσεις Παραδόσεων Συναρτησιακής Ανάλυσης [βλ. [3], Κεφάλαιο 2, σελ. 19-20], καθώς επίσης και στις ενότητες Βασικές Έννοιες και η Τοπολογία του $\mathbb{R}^n$ του βιβλίου Ανάλυση ΙΙ [βλ. [4], Κεφάλαιο 1, σελ. 3-5, 8-9].

Ενότητα 1.6: Η ενότητα αυτή βασίστηκε στην ενότητα Convex Functions του βιβλίου The Mathematics of Nonlinear Programming [βλ. [6], Κεφάλαιο 2, σελ. 45-57], στην ενότητα Κυρτά Σύνολα από τις Σημειώσεις Παραδόσεων Συναρτησιακής Ανάλυσης [βλ. [3], Κεφάλαιο 1, σελ. 17], καθώς επίσης και στις ενότητες Κυρτές Συναρτήσεις και Συναρτήσεις Χρησιμότητας του βιβλίου Θέματα Ανάλυσης και Θεωρία Γενικής Ισορροπίας στην Οικονομία [βλ. [5], Κεφάλαια 2,6, σελ. 37-38, 129-130].

Ενότητα 1.7: Η ενότητα αυτή βασίστηκε στο Κεφάλαιο Μέθοδοι Βελτιστοποίησης του βιβλίου Εισαγωγή στην Αριθμητική Ανάλυση [βλ. [8], Κεφάλαιο 5, σελ. 195-203], στις ενότητες Η Τοπολογία του $\mathbb{R}^n$, Συμπαγή Σύνολα και Συνέχεια του βιβλίου Ανάλυση ΙΙ [βλ. [4], Κεφάλαια 1,2, σελ. 9, 13-14, 18, 41], στην ενότητα Coercive Functions and Global Minimizers του βιβλίου The Mathematics of Nonlinear Programming [βλ. [6], Κεφάλαιο 1, σελ. 25-28], στο Παράρτημα Topology of the Euclidean Space $\mathbb{R}^n$ του βιβλίου Numerical Optimization [βλ. [9], Παράρτημα Α, σελ. 620-621], καθώς επίσης και στις ενότητες Open Set, Closed Set, Neighborhood, Completeness και Compactness and Finite Dimension του βιβλίου Introductory Functional Analysis with Applications [βλ. [7], Κεφάλαια 1,2, σελ. 18, 30, 77].

Ενότητα 1.8: Η ενότητα αυτή βασίστηκε στο Κεφάλαιο Μέθοδοι Βελτιστοποίησης του βιβλίου Εισαγωγή στην Αριθμητική Ανάλυση [βλ. [8], Κεφάλαιο 5, σελ. 195-203], στο Κεφάλαιο Διαφορίσιμες Συναρτήσεις του βιβλίου Ανάλυση ΙΙ [βλ. [4], Κεφάλαιο 4, σελ. 105-121, 144, 158-159], στο Κεφάλαιο Διαφορισιμότητα Συνάρτησης Πολλών Μεταβλητών του βιβλίου Μαθηματική Ανάλυση ΙΙ [βλ. [10], Κεφάλαιο 4, σελ. 111-121], στο Κεφάλαιο Παράγωγος Συνάρτησης του βιβλίου Μαθηματική Ανάλυση Ι [βλ. [11], Κεφάλαιο 6, σελ. 314-321], καθώς επίσης και στο Παράρτημα Derivatives και Directional Derivatives του βιβλίου Numerical Optimization [βλ. [9], Παράρτημα Α, σελ. 625-629].

Ενότητα 1.9: Η ενότητα αυτή βασίστηκε στο Κεφάλαιο Μέθοδοι Βελτιστοποίησης του βιβλίου Εισαγωγή στην Αριθμητική Ανάλυση [βλ. [8], Κεφάλαιο 5, σελ. 195-203], στην ενότητα Ακρότατα Συνάρτησης του βιβλίου Μαθηματική Ανάλυση ΙΙ [βλ. [10], Κεφάλαιο 7, σελ. 402-403], καθώς επίσης και στην ενότητα What Is a Solution? του βιβλίου Numerical Optimization [βλ. [9], Κεφάλαιο 2, σελ. 12-17].

Ενότητα 1.10: Η ενότητα αυτή βασίστηκε στην ενότητα Optimization Algorithms του βιβλίου Numerical Optimization [βλ. [9], Κεφάλαιο 1, σελ. 8-9].

Κεφάλαιο 2

Βελτιστοποίηση Προβλημάτων με Περιορισμούς

Σε αυτό το κεφάλαιο θα ασχοληθούμε με την ελαχιστοποίηση συναρτήσεων υπό κάποιους περιορισμούς των μεταβλητών τους. Όπως αναφέραμε και στο Κεφάλαιο 1, το *γενικό πρόβλημα ελαχιστοποίησης με περιορισμούς*, το οποίο θα εξετάζουμε από εδώ και πέρα είναι το εξής:

$$\text{Να βρεθεί } x^* \in \mathcal{U} \text{ τέτοιο ώστε } f(x^*) = \min_{x \in \mathcal{U}} f(x) \tag{2.1}$$
$$\text{όπου } f, c_i : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}, S \subset \Omega$$
$$\text{και } \mathcal{U} = \{x \in S : c_i(x) = 0, i \in \mathcal{E}, c_i(x) \geq 0, i \in \mathcal{I}\}$$

Στο γενικό πρόβλημα (2.1), όπως ορίστηκαν στο Κεφάλαιο 1, καλούμε f την *αντικειμενική συνάρτηση*, $c_i, i \in \mathcal{E}$ τους *περιορισμούς ισότητες (equality constraints)*, $c_i, i \in \mathcal{I}$ τους *περιορισμούς ανισότητες (inequality constraints)* και $\mathcal{U}$ το *εφικτό σύνολο (feasible set)*, το οποίο περιέχει όλα τα *εφικτά σημεία (feasible points)*. Επίσης, καλούμε *εφικτή περιοχή (feasible region)* του προβλήματος το χωρίο του γραφήματος που σχηματίζεται από τους περιορισμούς και είναι το σύνολο όλων των εφικτών σημείων.

Σε αυτό το κεφάλαιο θα εξάγουμε μαθηματικούς τύπους για τις λύσεις του προβλήματος (2.1) και τις συνθήκες που πρέπει να ικανοποιούν. Όπως στο Κεφάλαιο 1, θα παρουσιάσουμε δύο ειδών συνθήκες, που αφορούν στην εξασφάλιση της λύσης του προβλήματος. Τις *αναγκαίες* συνθήκες, τις οποίες πρέπει να ικανοποιεί κάθε λύση του προβλήματος και τις *ικανές* συνθήκες, οι οποίες αν ικανοποιούνται σε ένα σημείο x^* τότε αυτό το σημείο είναι λύση του προβλήματος.

Υπενθυμίζουμε ότι στο *γενικό προβλημα ελαχιστοποίησης χωρίς περιορισμούς* οι συνθήκες ήταν οι εξής:

- Οι *αναγκαίες* συνθήκες: Τα τοπικά ελάχιστα έχουν $\nabla f(x^*) = 0$ και ο $\nabla^2 f(x^*)$ είναι θετικά ημιορισμένος.
- Οι *ικανές* συνθήκες: Κάθε σημείο x^*, για το οποίο ισχύει $\nabla f(x^*) = 0$ και ο $\nabla^2 f(x^*)$ είναι θετικά ορισμένος, είναι τοπικό ελάχιστο της f.

2.1 Τοπική και ολική λύση

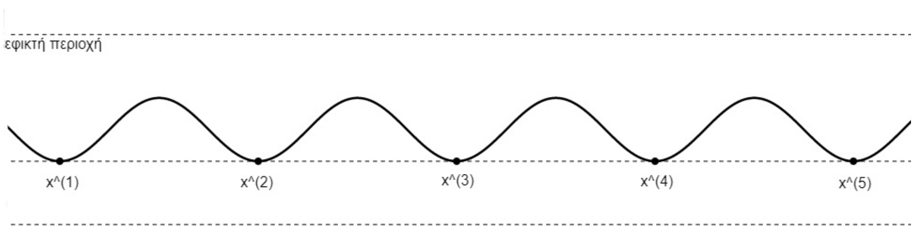
Η εύρεση ολικής λύσης σε ένα πρόβλημα, είναι μία δύσκολη διαδικασία και στα προβλήματα χωρίς περιορισμούς και σε εκείνα με περιορισμούς. Ωστόσο, η ύπαρξη περιορισμών διευκολύνει μερικές φορές την κατάσταση, καθώς μέσω του εφικτού συνόλου, μπορούν να εξαιρεθούν σημεία τοπικού ελαχίστου, εφόσον αυτά δεν ικανοποιούν τους περιορισμούς, αλλά και να βρεθεί το ολικό ελάχιστο, από τα εναπομείναντα σημεία. Από την άλλη, μερικές φορές οι περιορισμοί μπορεί και να δυσκολέψουν την επίλυση του προβλήματος.

Παράδειγμα 2.1.1. *Έχουμε το εξής πρόβλημα (Σχήμα 2.1):*

$$\min\,(x_2+100)^2 + 0.01x_1^2, \text{ υπό τον περιορισμό } x_2 - \cos x_1 \geq 0.$$

Χωρίς τον περιορισμό, το πρόβλημα έχει μοναδική λύση την $(0,-100)^T$. *Με τον περιορισμό, υπάρχουν τοπικά ελάχιστα της μορφής:*

$$x^{(k)} = (k\pi, -1)^T, \text{ για } k = \pm 1, \pm 3, \pm 5, \ldots$$



Σχήμα 2.1: Πρόβλημα με πολλά τοπικά ελάχιστα

Ορισμός 2.1. *Ένα διάνυσμα* x^* *είναι τοπική λύση του προβλήματος (2.1) εάν* $x^* \in \mathcal{U}$ *και υπάρχει περιοχή* $\mathcal{N}$ *του* x^* *τέτοια ώστε* $f(x) \geq f(x^*)$, *για κάθε* $x \in \mathcal{N} \cap \mathcal{U}$.

Ορισμός 2.2 (βλ. [9], Κεφάλαιο 12, Ορισμός, σελ. 306). *Ένα διάνυσμα* x^* *είναι αυστηρά τοπική λύση εάν* $x^* \in \mathcal{U}$ *και υπάρχει περιοχή* $\mathcal{N}$ *του* x^* *τέτοια ώστε* $f(x) > f(x^*)$, *για κάθε* $x \in \mathcal{N} \cap \mathcal{U}$ *με* $x \neq x^*$.

Ορισμός 2.3 (βλ. [9], Κεφάλαιο 12, Ορισμός, σελ. 306). *Ένα σημείο* x^* *είναι απομονωμένη τοπική λύση εάν* $x^* \in \mathcal{U}$ *και υπάρχει περιοχή* $\mathcal{N}$ *του* x^* *τέτοια ώστε το* x^* *να είναι η μόνη τοπική λύση στο σύνολο* $\mathcal{N} \cap \mathcal{U}$.

Παρατήρηση

Κάθε απομονωμένη τοπική λύση είναι προφανώς αυστηρά τοπική λύση, όμως το αντίστροφο δεν ισχύει.

Ορισμός 2.4 (βλ. [9], Κεφάλαιο 12, Ορισμός, σελ. 320). *Η συνάρτηση Lagrange του προβλήματος (2.1) σε κάποιο σημείο* $x \in \mathcal{U}$ *ορίζεται ως:*

$$\mathcal{L}(x,\lambda) = f(x) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i c_i(x),$$

όπου το διάνυσμα $\lambda \in \mathbb{R}^n$ *είναι το διάνυσμα των πολλαπλασιαστών Lagrange, με την συνιστώσα* λ_i *του διανύσματος να είναι ο πολλαπλασιαστής Lagrange που αντιστοιχεί στον περιορισμό* $c_i(x)$.

Ορισμός 2.5 (βλ. [9], Κεφάλαιο 12, Ορισμός, σελ. 321)**.** *Ορίζουμε ως κλίση της συνάρτησης Lagrange να είναι:*

$$\nabla_x \mathcal{L}(x, \lambda) = \nabla f(x) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i \nabla c_i(x).$$

Ορισμός 2.6 (βλ. [9], Κεφάλαιο 12, Ορισμός, σελ. 308)**.** *Σε ένα εφικτό σημείο x, ο ανισοτικός περιορισμός για $i \in \mathcal{I}$ ονομάζεται ενεργός εάν ικανοποιείται ως ισότητα, δηλαδή εάν ισχύει $c_i(x) = 0$.*

Αντίστοιχα, ονομάζεται ανενεργός εάν ικανοποιείται ως καθαρή ανισότητα, δηλαδή εάν ισχύει $c_i(x) > 0$.

Ορισμός 2.7 (βλ. [9], Κεφάλαιο 12, Ορισμός 12.1, σελ. 308)**.** *Το σύνολο $\mathcal{A}(x)$, σε κάθε εφικτό σημείο x, ονομάζεται ενεργό σύνολο (active set), εάν αποτελείται από τους δείκτες των ισοτικών περιορισμών του συνόλου $\mathcal{E}$ και από αυτούς των ανισοτικών περιορισμών του συνόλου $\mathcal{I}$ για τους οποίους ισχύει $c_i(x) = 0$. Δηλαδή το σύνολο:*

$$\mathcal{A}(x) = \mathcal{E} \cup \{i \in \mathcal{I} : c_i(x) = 0\}.$$

2.2 Λείες συναρτήσεις

Ορισμός 2.8 (βλ. [4], Κεφάλαιο 3, Ορισμός, σελ. 73)**.** *Μία συνάρτηση $f : (a, b) \longrightarrow \mathbb{R}$ ονομάζεται λεία συνάρτηση εάν είναι συνεχώς παραγωγίσιμη και $f^{'}(t) \neq 0$ για κάθε $t \in (a, b)$.*

Μία διανυσματική συνάρτηση $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ είναι λεία συνάρτηση εάν κάθε συνιστώσα της είναι συνεχώς παραγωγίσιμη και όλες οι παράγωγοι $\frac{df_i}{dt} \neq 0$, $i = 1, \ldots, n$.

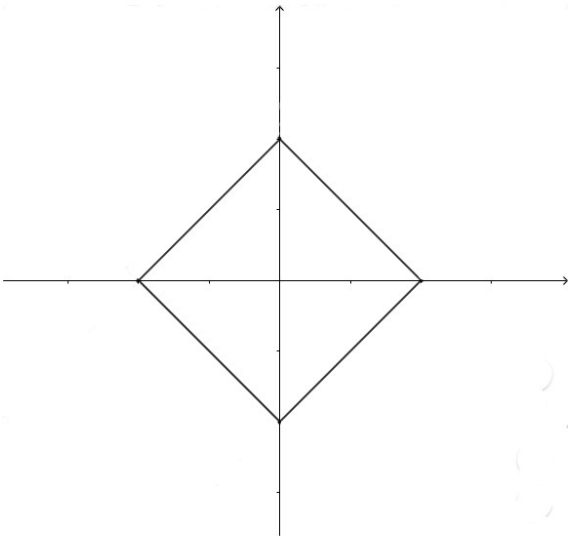
Η αντικειμενική συνάρτηση καθώς και οι περιορισμοί είναι σημαντικό να είναι λείες συναρτήσεις για τον χαρακτηρισμό των λύσεων. Αυτό το χαρακτηριστικό εξασφαλίζει ότι αυτές οι συναρτήσεις συμπεριφέρονται ομαλά και λόγω αυτού, ο αλγόριθμος μπορεί να εξάγει αποτελέσματα προς την σωστή κατεύθυνση.

Οι συναρτήσεις που δεν είναι λείες, εμφανίζουν στα γραφήματά τους σημεία ασυνέχειας ή γωνίες στην καμπύλη τους, όπου η παράγωγος δεν υπάρχει. Πολλές φορές παρατηρούμε ότι το γράφημα της εφικτής περιοχής ενός προβλήματος με περιορισμούς, περιέχει πολλά τέτοια σημεία, δηλαδή σημεία που δημιουργούν γωνίες. Ωστόσο, αυτό δεν σημαίνει ότι οι συναρτήσεις των περιορισμών, που δημιουργούν αυτή την περιοχή, δεν είναι λείες συναρτήσεις. Επομένως, οι καμπύλες που δημιουργούν τα σύνορα ενός προβλήματος και οι οποίες δεν είναι λείες, δημιουργούνται συνήθως από συναρτήσεις περιορισμών οι οποίες είναι λείες.

Παράδειγμα 2.2.1. *Έχουμε τον εξής περιορισμό:*

$$\|x\|_1 = |x_1| + |x_2| \leq 1 \implies \begin{cases} x_1 + x_2 & \leq 1, \ \text{όταν} \ \ x_1, x_2 \geq 0 \\ x_1 - x_2 & \leq 1, \ \text{όταν} \ \ x_1 \geq 0, x_2 \leq 0 \\ -x_1 + x_2 & \leq 1, \ \text{όταν} \ \ x_1 \leq 0, x_2 \geq 0 \\ -x_1 - x_2 & \leq 1, \ \text{όταν} \ \ x_1, x_2 \leq 0 \end{cases} \tag{2.2}$$

Η εφικτή περιοχή των παραπάνω περιορισμών μπορεί να απεικονιστεί με έναν ρόμβο στον $\mathbb{R}^2$ (Σχήμα 2.2), όπου καθένας από τους τέσσερις περιορισμούς αναπαριστά μία ακμή του.



Σχήμα 2.2: Μία εφικτή περιοχή με μη λεία σύνορα μπορεί να αναπαρασταθεί από λείες συναρτήσεις

Παρατήρηση

Γενικά, οι περιορισμοί επιλέγονται έτσι ώστε ο κάθενας τους να αναπαριστά μία λεία πλευρά του συνόρου $\mathcal{U}$.

Ωστόσο, μερικές φορές προβλήματα χωρίς περιορισμούς των οποίων η αντικειμενική συνάρτηση δεν είναι λεία, μπορούν να επαναδιατυπωθούν ως προβλήματα λείων συναρτήσεων με περιορισμούς.

Παράδειγμα 2.2.2. *Έστω το πρόβλημα χωρίς περιορισμούς:*

$$\min f(x), \ \text{όπου } f(x) = \max(x^2, x). \tag{2.3}$$

Η παραπάνω συνάρτηση στα σημεία $x = 0$, $x = 1$ δεν είναι παραγωγίσιμη και το πρόβλημα έχει λύση την $x^* = 0$.

Με μερικούς μετασχηματισμούς το πρόβλημα (2.3) γίνεται ένα πρόβλημα λείων συναρτήσεων με περιορισμούς:

$$\min t, \ \text{υπό τους περιορισμούς} \begin{cases} t \geq x \\ t \geq x^2 \end{cases}. \tag{2.4}$$

Παρατηρήσεις

1. Μετασχηματισμοί της μορφής (2.2) και (2.4) χρησιμοποιούνται συχνά σε προβλήματα όπου η f είναι το max κάποιων συναρτήσεων και όταν είναι ίση με την $\|\cdot\|_1$ ή την $\|\cdot\|_\infty$ μιας διανυσματικής συνάρτησης.

2. Στα Παραδείγματα 2.2.1 και 2.2.2 οι περιορισμοί των προβλημάτων βρίσκονται σε διαφορετική μορφή από αυτή που παρουσιάστηκε στο γενικό πρόβλημα (2.1). Ωστόσο, είναι προφανές ότι οποιασδήποτε μορφής περιορισμοί μπορούν να εκφραστούν στην επιθυμητή.

2.3 Παραδείγματα

Για να μπορέσουμε να παρουσιάσουμε τις βασικές αρχές για τον χαρακτηρισμό των λύσεων, θα δουλέψουμε πάνω σε μερικά απλά παραδείγματα.

2.3.1 Πρόβλημα με έναν περιορισμό ισότητα

Παράδειγμα 2.3.1. *Έστω το πρόβλημα:*

$$\min x_1 + x_2, \text{ υπό τον περιορισμό } x_1^2 + x_2^2 - 2 = 0. \tag{2.5}$$

Σύμφωνα με το πρόβλημα (2.1) έχουμε:

$$f(x) = x_1 + x_2, \mathcal{E} = \{1\}, \mathcal{I} = \emptyset \text{ και } c_1(x) = x_1^2 + x_2^2 - 2.$$

Το εφικτό σύνολο αυτού του προβλήματος είναι:

$$\mathcal{U} = \{x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 : x_1^2 + x_2^2 - 2 = 0\},$$

δηλαδή ο κύκλος κέντρου O(0,0) και ακτίνας $\sqrt{2}$ (Σχήμα 2.3 το σύνορο και όχι το εσωτερικό του).

Από το Σχήμα 2.3 παρατηρούμε ότι η λύση του είναι $x^* = (-1, -1)^T$, καθώς για αυτή την τιμή η αντικειμενική συνάρτηση παίρνει την ελάχιστη τιμή της με $f(x^*) = -2$.

Παρατηρήσεις

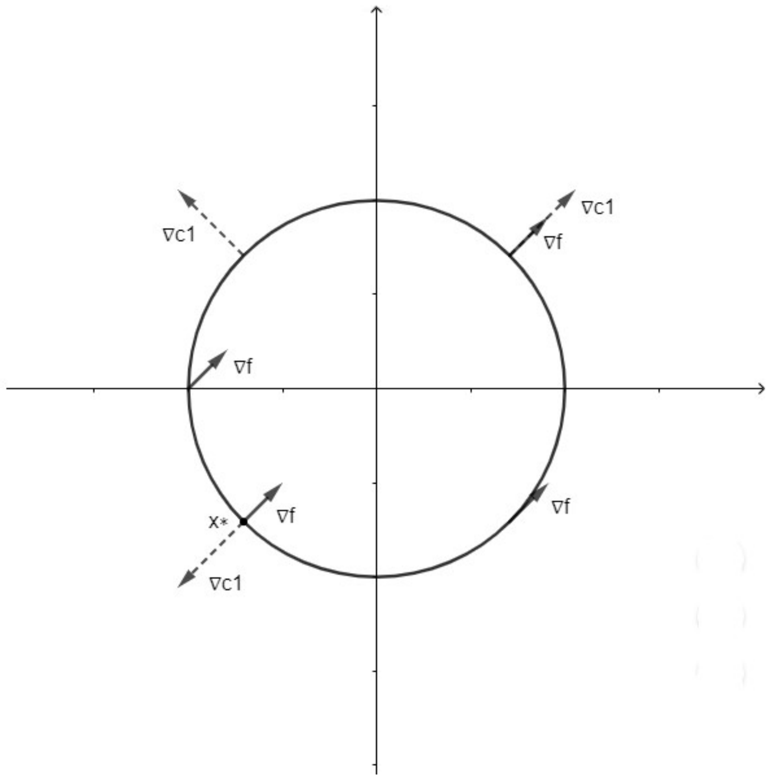
1. Από κάθε σημείο του κύκλου (Σχήμα 2.3), μπορούμε να βρούμε μία διεύθυνση μετακίνησης έτσι ώστε να παραμένουμε πάνω στην εφικτή περιοχή, δηλαδή πάνω στον κύκλο, ενώ ταυτόχρονα μειώνεται η συνάρτηση f. Για παράδειγμα, από το σημείο $x = (\sqrt{2}, 0)^T$, οποιαδήποτε μετακίνηση δεξιόστροφα και πάνω στον κύκλο, θα επιφέρει το επιθυμητό αποτέλεσμα.

2. Από το Σχήμα 2.3 παρατηρούμε ότι στην λύση x^*, οι κλίσεις:

$$\nabla f(x^*) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ και } \nabla c_1(x^*) = \begin{bmatrix} -2 \\ -2 \end{bmatrix}$$

είναι συγγραμμικά διανύσματα. Αυτό σημαίνει ότι υπάρχει $\lambda_1^* \in \mathbb{R}$, στην περίπτωση αυτή $\lambda_1^* = -\frac{1}{2}$, τέτοιο ώστε:

$$\nabla f(x^*) = \lambda_1^* \nabla c_1(x^*). \tag{2.6}$$



Σχήμα 2.3: Το πρόβλημα (2.5), όπου φαίνεται η κλίση της $f(x)$ και του $c_1(x)$ σε διάφορα εφικτά σημεία

3. Ο παραπάνω τύπος (2.6) προκύπτει και από το Θεώρημα Taylor αν εφαρμοστεί στην αντικειμενική συνάρτηση $f(x)$ και στον περιορισμό $c_1(x)$. Σύμφωνα με αυτό, για να παραμείνουμε στην εφικτή περιοχή του προβλήματος θεωρούμε βήμα $s > 0$, κατάλληλα μικρό, έτσι ώστε να ισχύει: $c_1(x+s) = 0$. Τότε από τον τύπο Taylor έχουμε:

$$0 = c_1(x+s) \approx c_1(x) + \nabla c_1(x)^T s \overset{c_1(x)=0}{=} \nabla c_1(x)^T s.$$

Επομένως, για να παραμείνουμε στην εφικτή περιοχή πρέπει το βήμα s να ικανοποιεί την:

$$\nabla c_1(x)^T s = 0. \tag{2.7}$$

Αντίστοιχα, εάν θέλουμε το βήμα s να προκαλέσει μείωση στην f, θα πρέπει $f(x+s) - f(x) < 0$. Τότε από τον τύπο Taylor έχουμε:

$$0 > f(x+s) - f(x) \approx f(x) + \nabla f(x)^T s - f(x).$$

Άρα:

$$\nabla f(x)^T s < 0. \tag{2.8}$$

Για να υπάρχει τέτοιο βήμα s το οποίο να ικανοποιεί τις συνθήκες (2.7) και (2.8), θα πρέπει να υπάρχει κάποια διεύθυνση d, θα μπορούσαμε να επιλέξουμε $d = \dfrac{s}{\|s\|}$ ώστε η νόρμα του d να είναι 1, τέτοια ώστε:

$$\nabla c_1(x)^T d = 0 \;\; \text{και} \;\; \nabla f(x)^T d < 0. \tag{2.9}$$

Εάν δεν υπάρχει τέτοια διεύθυνση, που να ικανοποιεί τις συνθήκες (2.9), τότε είναι πιθανό να μην μπορεί να βρεθεί βήμα s, που να ικανοποιεί τις συνθήκες (2.7) και (2.8). Στην περίπτωση αυτή, η λύση x^* θα είναι τοπική λύση του προβλήματος.

Παρατηρούμε ότι η μόνη περίπτωση όπου κάποια διεύθυνση d μπορεί να ικανοποιεί τις συνθήκες (2.9), είναι μόνο όταν τα διανύσματα $\nabla f(x)$ και $\nabla c_1(x)$ δεν είναι συγγραμμικά. Αυτό, σύμφωνα με την Παρατήρηση 2 σημαίνει ότι δεν πρέπει να ισχύει η συνθήκη $\nabla f(x) = \lambda_1 \nabla c_1(x)$, για κάποιο $\lambda_1 \in \mathbb{R}$. Επομένως, εάν τα διανύσματα δεν είναι συγγραμμικά, τότε θέτοντας:

$$d = \frac{\bar{d}}{\|\bar{d}\|}, \quad \text{όπου} \quad \bar{d} = -\left(I - \frac{\nabla c_1(x) \nabla c_1(x)^T}{\|\nabla c_1(x)\|^2} \right) \nabla f(x),$$

εύκολα διαπιστώνουμε ότι αυτο το d ικανοποιεί τις συνθήκες (2.9).

Πράγματι:

$$\begin{aligned}
\nabla c_1(x)^T d &= \nabla c_1(x)^T \frac{\bar{d}}{\|\bar{d}\|} \\
&= \nabla c_1(x)^T \frac{-\left(I - \frac{\nabla c_1(x) \nabla c_1(x)^T}{\|\nabla c_1(x)\|^2} \right) \nabla f(x)}{\|\bar{d}\|} \\
&= \frac{-\nabla c_1(x)^T \nabla f(x) + \nabla c_1(x)^T \frac{\nabla c_1(x) \nabla c_1(x)^T}{\|\nabla c_1(x)\|^2} \nabla f(x)}{\|\bar{d}\|} \\
&= \frac{-\nabla c_1(x)^T \nabla f(x) + \nabla c_1(x)^T \nabla f(x)}{\|\bar{d}\|} \\
&= 0.
\end{aligned}$$

Επίσης:

$$\begin{aligned}
\nabla f(x)^T d &= \nabla f(x)^T \frac{\bar{d}}{\|\bar{d}\|} \\
&= \nabla f(x)^T \frac{-\left(I - \frac{\nabla c_1(x) \nabla c_1(x)^T}{\|\nabla c_1(x)\|^2} \right) \nabla f(x)}{\|\bar{d}\|} \\
&= \frac{-\nabla f(x)^T \nabla f(x) + \nabla f(x)^T \frac{\nabla c_1(x) \nabla c_1(x)^T}{\|\nabla c_1(x)\|^2} \nabla f(x)}{\|\bar{d}\|} \\
&= \frac{-\|\nabla f(x)\|^2 \|\nabla c_1(x)\|^2 + \nabla f(x)^T \nabla c_1(x) \nabla c_1(x)^T \nabla f(x)}{\|\bar{d}\| \|\nabla c_1(x)\|^2} \\
&= \frac{-\|\nabla f(x)\|^2 \|\nabla c_1(x)\|^2 + (\nabla f(x)^T \nabla c_1(x))^2}{\|\bar{d}\| \|\nabla c_1(x)\|^2} \\
&< 0.
\end{aligned}$$

Η τελευταία ανισότητα προκύπτει από την ανισότητα Cauchy-Schwarz ως εξής:

$$\nabla f(x)^T \nabla c_1(x) \leq \|\nabla f(x)\| \|\nabla c_1(x)\|$$

$$\implies (\nabla f(x)^T \nabla c_1(x))^2 \leq \|\nabla f(x)\|^2 \|\nabla c_1(x)\|^2.$$

Καθώς, όμως, τα διανύσματα δεν είναι συγγραμμικά, δηλαδή η σχέση $\nabla f(x) = \lambda_1 \nabla c_1(x)$, δεν ισχύει για κάποιο $\lambda_1 \in \mathbb{R}$, έχουμε ότι:

$$(\nabla f(x)^T \nabla c_1(x))^2 < \|\nabla f(x)\|^2 \|\nabla c_1(x)\|^2.$$

4. Από τον Ορισμό 2.5 της κλίσης της συνάρτησης Lagrange, παρατηρούμε ότι η συνθήκη (2.6) μπορεί ισοδύναμα να γραφτεί ως:

$$\nabla_x \mathcal{L}(x^*, \lambda_1^*) = 0, \tag{2.10}$$

όπου x^* η λύση του προβλήματος και $\lambda_1^* \in \mathbb{R}$ ο πολλαπλασιαστής Lagrange του περιορισμού $c_1(x)$.

Από το παραπάνω συμπεραίνουμε ότι για να βρούμε την λύση του προβλήματος (2.5) πρέπει να εξετάσουμε τα στάσιμα σημεία της συνάρτησης Lagrange.

5. Παρόλο που η συνθήκη (2.6) ή ισοδύναμα η (2.10) φαίνεται να είναι αναγκαία συνθήκη για την εύρεση της βέλτιστης λύσης του προβλήματος (2.5), είναι εμφανές ότι δεν είναι και ικανή. Για παράδειγμα, η συνθήκη (2.6) ικανοποιείται και στο σημείο $x = (1,1)^T$ με $\lambda_1 = \frac{1}{2}$. Ωστόσο, το σημείο αυτό δεν αποτελεί λύση του προβλήματος, καθώς μεγιστοποιεί την αντικειμενική συνάρτηση.

2.3.2 Πρόβλημα με έναν περιορισμό ανισότητα

Παράδειγμα 2.3.2. *Έστω το πρόβλημα:*

$$\min x_1 + x_2, \text{ υπό τον περιορισμό } 2 - x_1^2 - x_2^2 \geq 0. \tag{2.11}$$

Σύμφωνα με το πρόβλημα (2.1) έχουμε:

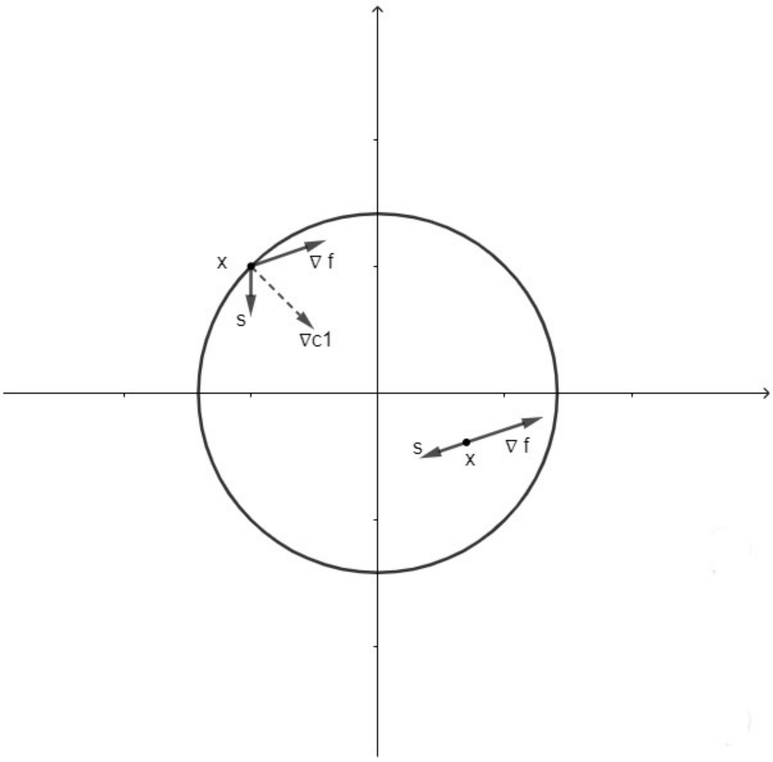
$$f(x) = x_1 + x_2, \mathcal{E} = \emptyset, \mathcal{I} = \{1\} \text{ και } c_1(x) = 2 - x_1^2 - x_2^2.$$

Το εφικτό σύνολο αυτού του προβλήματος είναι:

$$\mathcal{U} = \{x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 : 2 - x_1^2 - x_2^2 \geq 0\},$$

δηλαδή ο κύκλος κέντρου O(0,0) και ακτίνας $\sqrt{2}$, το σύνορο και το εσωτερικό του (Σχήμα 2.4).

Η λύση του είναι: $x^* = (-1,-1)^T$ και η συνθήκη (2.6) ισχύει για $\lambda_1^* = \frac{1}{2}$.



Σχήμα 2.4: Το πρόβλημα (2.11), όπου φαίνονται οι διευθύνσεις του βήματος s από δύο εφικτά σημεία x, στα οποία ο περιορισμός είναι ενεργός και ανενεργός αντίστοιχα

Παρατηρήσεις

1. Από το Σχήμα 2.4 παρατηρούμε ότι το διάνυσμα της κλίσης του περιορισμού $c_1(x)$ από οποιοδήποτε σημείο πάνω στο σύνορο, έχει κατεύθυνση προς το εσωτερικό της εφικτής περιοχής, δηλαδή του κύκλου.

2. Όπως και στο προηγούμενο παράδειγμα, από το Θεώρημα Taylor υποθέτουμε ότι ένα εφικτό σημείο x δεν είναι η βέλτιστη λύση εάν μπορούμε να βρούμε ένα βήμα $s > 0$, κατάλληλα μικρό, με το οποίο να παραμένουμε στην εφικτή περιοχή αλλά ταυτόχρονα να προκαλεί μείωση στην f.

   Δηλαδή εάν θέλουμε το βήμα s να προκαλέσει μείωση στην f, θα πρέπει, όπως και προηγουμένως, να ισχύει:

   $$0 > f(x+s) - f(x) \approx \nabla f(x)^T s \implies \nabla f(x)^T s < 0. \tag{2.12}$$

   Αντίστοιχα, για να παραμείνουμε στην εφικτή περιοχή πρέπει να ισχύει:

   $$0 \leq c_1(x+s) \approx c_1(x) + \nabla c_1(x)^T s.$$

   Επομένως, για να παραμείνουμε στην εφικτή περιοχή πρέπει το βήμα s να ικανοποιεί την:

   $$c_1(x) + \nabla c_1(x)^T s \geq 0. \tag{2.13}$$

   Για να βρούμε εάν υπάρχει τέτοιο βήμα s το οποίο να ικανοποιεί τις συνθήκες (2.12) και (2.13), θεωρούμε δύο περιπτώσεις:

- **Περίπτωση 1.** Έστω ότι το εφικτό σημείο x βρίσκεται εντός του κύκλου (Σχήμα 2.4), έτσι ώστε να ικανοποιείται η αυστηρή ανισότητα του περιορισμού $c_1(x) > 0$. Τότε, κάθε διάνυσμα του βήματος s ικανοποιεί την συνθήκη (2.13), υπό την προϋπόθεση να είναι κατάλληλα μικρό. Πράγματι, εφόσον $\nabla f(x) \neq 0$, μπορούμε να βρούμε βήμα s που να ικανοποιεί τις συνθήκες (2.12) και (2.13) εάν θέσουμε:

$$s = -a\nabla f(x), \quad \text{όπου} \quad a \in \mathbb{R}, a > 0, \quad \text{κατάλληλα μικρό.}$$

Πράγματι, για την συνθήκη (2.12) έχουμε:

$$\begin{aligned} \nabla f(x)^T s &= \nabla f(x)^T(-a\nabla f(x)) \\ &= -a\nabla f(x)^T \nabla f(x) \\ &= -a\|\nabla f(x)\|_2^2 \\ &< 0, \text{ καθώς } a > 0. \end{aligned}$$

Αντίστοιχα, για την συνθήκη (2.13) έχουμε:

$$\begin{aligned} c_1(x) + \nabla c_1(x)^T s &= c_1(x) + \nabla c_1(x)^T(-a\nabla f(x)) \\ &= c_1(x) - a\nabla c_1(x)^T \nabla f(x) \\ &\overset{C-S}{\geq} c_1(x) - a\|\nabla f(x)\|\|\nabla c_1(x)\|. \end{aligned}$$

Αν $\nabla c_1(x) = 0$, τότε:

$$\begin{aligned} c_1(x) + \nabla c_1(x)^T s &\overset{C-S}{\geq} c_1(x) - a\|\nabla f(x)\|\|\nabla c_1(x)\| \\ &= c_1(x) - 0 \\ &= c_1(x) \\ &> 0. \end{aligned}$$

Αν $\nabla c_1(x) \neq 0$, τότε θέτουμε $a = \dfrac{c_1(x)}{\|\nabla f(x)\|\|\nabla c_1(x)\|} > 0$, αφού $c_1(x) > 0$ και $\nabla f(x), \nabla c_1(x) \neq 0$. Τότε η παραπάνω σχέση γίνεται:

$$\begin{aligned} c_1(x) + \nabla c_1(x)^T s &\overset{C-S}{\geq} c_1(x) - a\|\nabla f(x)\|\|\nabla c_1(x)\| \\ &\geq c_1(x) - \frac{c_1(x)}{\|\nabla f(x)\|\|\nabla c_1(x)\|}\|\nabla f(x)\|\|\nabla c_1(x)\| \\ &= 0 \\ &\implies c_1(x) + \nabla c_1(x)^T s \geq 0. \end{aligned}$$

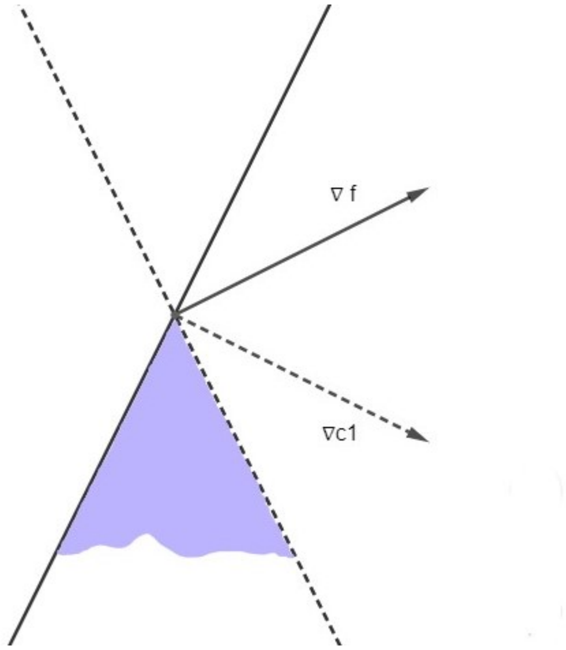
Επομένως, και στις δύο περιπτώσεις παραμένουμε στην εφικτή περιοχή. Ωστόσο, το παραπάνω βήμα s δεν ικανοποιεί τις απαραίτητες συνθήκες όταν $\nabla f(x) = 0$.

- **Περίπτωση 2.** Έστω ότι το εφικτό σημείο x βρίσκεται πάνω στον κύκλο (Σχήμα 2.4), έτσι ώστε να ικανοποιείται η ισότητα του περιορισμού $c_1(x) = 0$. Τότε οι συνθήκες (2.12) και (2.13) γίνονται:

$$\nabla f(x)^T s < 0 \ \text{ και } \ \nabla c_1(x)^T s \geq 0.$$

Η πρώτη συνθήκη ορίζει ένα ανοιχτό ημιεπίπεδο και η δεύτερη ένα κλειστό ημιεπίπεδο (Σχήμα 2.5). Από το σχήμα παρατηρούμε ότι ο χώρος ανάμεσα στα δύο ημιεπίπεδα είναι κενός μόνο όταν τα διανύσματα $\nabla f(x)$ και $\nabla c_1(x)$ είναι συγγραμμικά και ομόρροπα, δηλαδή όταν:

$$\nabla f(x) = \lambda_1 \nabla c_1(x), \quad \text{για κάποιο} \quad \lambda_1 \geq 0.$$



Σχήμα 2.5: Η διεύθυνση d που ικανοποιεί και τις δύο συνθήκες (2.12) και (2.13), βρίσκεται ανάμεσα από ένα ανοιχτό και ένα κλειστό ημιεπίπεδο, δηλαδή οποιοδήποτε d μέσα στον κώνο έχει το επιθυμητό αποτέλεσμα

3. Όταν υπάρχει ο περιορισμός ανισότητας το πρόσημο του αντίστοιχου πολλαπλασιαστή Lagrange έχει σημασία, καθώς αν $\lambda_1 < 0$ τότε τα δύο διανύσματα θα είναι αντίρροπα και έτσι δεν θα ικανοποιούνται οι συνθήκες (2.12) και (2.13).

4. Όπως και στο προηγούμενο παράδειγμα, Από τον Ορισμό 2.5 της κλίσης της συνάρτησης Lagrange, παρατηρούμε ότι οι συνθήκες των περιπτώσεων 1. και 2. μπορούν ισοδύναμα να γραφτούν ως:

$$\nabla_x \mathcal{L}(x^*, \lambda_1^*) = 0, \quad \text{για κάποιο} \quad \lambda_1^* \geq 0, \tag{2.14}$$

όπου x^* στάσιμο σημείο της συνάρτησης Lagrange και $\lambda_1^* \in \mathbb{R}$ ο πολλαπλασιαστής Lagrange του περιορισμού $c_1(x)$.

Καθώς επίσης:

$$\lambda_1^* c_1(x^*) = 0. \tag{2.15}$$

5. Η συνθήκη (2.15) ονομάζεται συνθήκη συμπληρωματικότητας (complementarity condition) και εξασφαλίζει ότι ο πολλαπλασιαστής Lagrange λ_1 είναι αυστηρά θετικός μόνο όταν ο αντίστοιχος περιορισμός $c_1(x)$ είναι ενεργός και μηδέν όταν ο περιορισμός είναι ανενεργός.

2.3.3 Πρόβλημα με δύο περιορισμούς ανισότητες

Παράδειγμα 2.3.3. *Έστω το πρόβλημα:*

$$\min x_1 + x_2, \text{ υπό τους περιορισμούς } \begin{aligned} 2 - x_1^2 - x_2^2 &\geq 0, \\ x_2 &\geq 0. \end{aligned} \tag{2.16}$$

Σύμφωνα με το πρόβλημα (2.1) έχουμε:

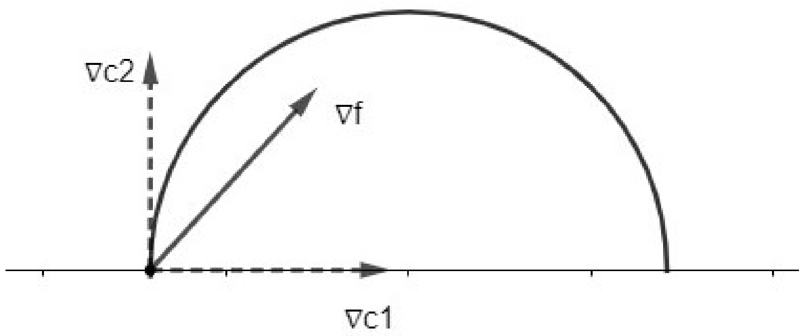
$$f(x) = x_1 + x_2, \mathcal{E} = \emptyset, \mathcal{I} = \{1,2\} \text{ και } c_1(x) = 2 - x_1^2 - x_2^2, c_2(x) = x_2.$$

Το εφικτό σύνολο αυτού του προβλήματος είναι:

$$\mathcal{U} = \{x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 : 2 - x_1^2 - x_2^2 \geq 0, x_2 \geq 0\},$$

δηλαδή ο μισός δίσκος κέντρου O(0,0) και ακτίνας $\sqrt{2}$ (Σχήμα 2.6).

Η λύση του είναι: $x^* = (-\sqrt{2}, 0)^T$ στην οποία και οι δύο περιορισμοί είναι ενεργοί.



Σχήμα 2.6: Το πρόβλημα (2.16), όπου φαίνεται η κλίση των ενεργών περιορισμών και της αντικειμενικής συνάρτησης στην λύση του προβλήματος

Παρατηρήσεις

1. Όπως και στα παραδείγματα 2.3.1 και 2.3.2, θεωρούμε διεύθυνση d η οποία να ικανοποιεί τις συνθήκες:

$$\nabla c_i(x)^T d \geq 0, \quad \text{για } i \in \mathcal{I} = \{1,2\} \text{ και } \nabla f(x)^T d < 0. \tag{2.17}$$

Ωστόσο, από το Σχήμα 2.6 παρατηρούμε ότι δεν μπορεί να υπάρξει τέτοια διεύθυνση όταν $x = (-\sqrt{2}, 0)^T$. Πράγματι, οι συνθήκες $\nabla c_i(x)^T d \geq 0$, για $i = 1, 2$, ικανοποιούνται μόνο όταν η διεύθυνση d βρίσκεται μέσα στο τεταρτημόριο που δημιουργείται από τα διανύσματα $\nabla c_1(x)$ και $\nabla c_2(x)$ και είναι εμφανές ότι κάθε διεύθυνση d στο τεταρτημόριο αυτό ικανοποιεί την συνθήκη $\nabla f(x)^T d \geq 0$.

2. Η συνάρτηση Lagrange για το πρόβλημα (2.16) είναι:

$$\mathcal{L}(x,\lambda) = f(x) - \lambda_1 c_1(x) - \lambda_2 c_2(x),$$

όπου $\lambda = (\lambda_1, \lambda_2)^T$ είναι το διάνυσμα των πολλαπλασιαστών Lagrange των αντίστοιχων περιορισμών.

Η αντίστοιχη συνθήκη (2.14) του προβλήματος (2.16) είναι:

$$\nabla_x \mathcal{L}(x^*, \lambda^*) = 0, \quad \text{για κάποιο} \quad \lambda^* \geq 0, \tag{2.18}$$

όπου η ανισότητα $\lambda^* \geq 0$ σημαίνει ότι και οι δύο συνιστώσες του διανύσματος λ^* πρέπει να είναι μη αρνητικές.

Η αντίστοιχη συνθήκη συμπληρωματικότητας (2.15) του προβλήματος (2.16) είναι:

$$\lambda_1^* c_1(x^*) = 0 \quad \text{και} \quad \lambda_2^* c_2(x^*) = 0. \tag{2.19}$$

Επομένως, για $x^* = (-\sqrt{2}, 0)^T$ έχουμε:

$$\nabla f(x^*) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \nabla c_1(x^*) = \begin{bmatrix} 2\sqrt{2} \\ 0 \end{bmatrix}, \nabla c_2(x^*) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Άρα εύκολα διαπιστώνουμε ότι η συνθήκη (2.18) ισχύει όταν $\lambda^* = \begin{bmatrix} \frac{1}{2\sqrt{2}} \\ 1 \end{bmatrix}$.

Παρατηρούμε ότι και οι δύο συνιστώσες του διανύσματος λ^* είναι θετικές για να ικανοποιείται η συνθήκη (2.18).

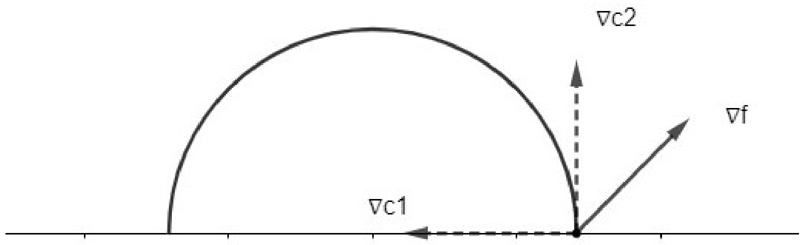
3. Θεωρούμε εφικτό σημείο $x = (\sqrt{2}, 0)^T$ το οποίο δεν αποτελεί λύση του προβλήματος (2.16). Σε αυτό το σημείο και οι δύο περιορισμοί είναι ενεργοί (Σχήμα 2.7). Είναι εύκολο να βρούμε διανύσματα d τα οποία να ικανοποιούν την συνθήκη (2.17). Ένα τέτοιο διάνυσμα είναι το $d = (-1, 0)^T$. Για αυτό το σημείο x η συνθήκη $\nabla_x \mathcal{L}(x,\lambda) = 0$ ικανοποιείται μόνο όταν $\lambda = (-\frac{1}{2\sqrt{2}}, 1)^T$. Ωστόσο, η μία συνιστώσα του διανύσματος λ είναι αρνητική, επομένως η συνθήκη (2.18) δεν ικανοποιείται στο σημείο αυτό.

4. Θεωρούμε ένα άλλο εφικτό σημείο $x = (1, 0)^T$ το οποίο επίσης δεν αποτελεί λύση του προβλήματος (2.16). Σε αυτό το σημείο μόνο ο περιορισμός c_2 είναι ενεργός. Εφόσον κάθε μικρό βήμα s από αυτό το σημείο θα ικανοποιεί την συνθήκη $c_1(x+s) > 0$, θα εξετάσουμε μόνο την συμπεριφορά του περιορισμού c_2 και της αντικειμενικής συνάρτησης για να κρίνουμε εάν το s αποτελεί ένα εφικτό βήμα. Όπως και στα προηγούμενα παραδείγματα, θεωρούμε διεύθυνση d η οποία να ικανοποιεί τις συνθήκες:

$$\nabla c_2(x)^T d \geq 0 \quad \text{και} \quad \nabla f(x)^T d < 0. \tag{2.20}$$

Αφού:

$$\nabla f(x) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \nabla c_2(x) = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$



Σχήμα 2.7: Το πρόβλημα (2.16), όπου φαίνεται η κλίση των ενεργών περιορισμών και της αντικειμενικής συνάρτησης σε ένα εφικτό σημείο

εύκολα διαπιστώνουμε ότι το διάνυσμα $d = \left(-\frac{1}{2}, \frac{1}{4}\right)^T$ ικανοποιεί τις συνθήκες (2.20).

Ωστόσο, αφού $c_1(x) > 0$ για αυτό το σημείο x, τότε από την συνθήκη (2.19) έχουμε ότι $\lambda_1 = 0$. Επόμενως, για να ικανοποιείται η συνθήκη $\nabla_x \mathcal{L}(x, \lambda) = 0$, ψάχνουμε λ_2 τέτοιο ώστε: $\nabla f(x) - \lambda_2 \nabla c_2(x) = 0$. Όμως, δεν υπάρχει τέτοιο λ_2, και άρα η συνθήκη δεν ικανοποιείται.

2.4 Εφαπτόμενος κώνος και προϋποθέσεις περιορισμών

Στην προηγούμενη ενότητα αναλύσαμε το κατά πόσο είναι εφικτό με ένα βήμα s να απομακρυνθούμε από ένα εφικτό σημείο παραμένοντας, παράλληλα, στην εφικτή περιοχή, εξετάζοντας την πρώτη παράγωγο της αντικειμενικής συνάρτησης και των περιορισμών. Για την ανάλυση αυτή, χρησιμοποιήθηκε ο τύπος Taylor πρώτης τάξης σε κάθε μία από αυτές τις συναρτήσεις, με σκοπό την εύρεση συνθηκών που πρέπει να ικανοποιεί το βήμα s. Αυτή η προσέγγιση έχει νόημα μόνο όταν οι παραγώμενες συνθήκες ορίζουν γεωμετρικά μία εφικτή περιοχή κοντά στο σημείο x που εξετάζεται. Εάν η εφικτή περιοχή που βρέθηκε κοντά στο x διαφέρει από την εφικτή περιοχή του προβλήματος, για παράδειγμα εάν ορίζει ένα επίπεδο σε σχέση με το μεμονωμένο σημείο που ορίζει το εφικτό σύνολο του προβλήματος, τότε οι παραγώμενες συνθήκες είναι πιθανό να μην δώσουν χρήσιμα αποτελέσματα για την λύση του προβλήματος. Σε αυτή την περίπτωση, θα χρειαστεί να γίνουν υποθέσεις για τους περιορισμούς που είναι ενεργοί στο σημείο x, έτσι ώστε οι δύο εφικτές περιοχές να είναι παρόμοιες κόντα στο x. Τέτοιου είδους υποθέσεις είναι οι προϋποθέσεις που πρέπει να ικανοποιούν οι περιορισμοί, προκειμένου να εξασφαλιστεί η σχέση του συνόλου των περιορισμών $\mathcal{U}$ με την εφικτή περιοχή κοντά στην λύση του προβλήματος x^*.

Ορισμός 2.9 (βλ. [9], Κεφάλαιο 12, Ορισμός, σελ. 316). *Έστω ένα εφικτό σημείο x. Ορίζουμε την ακολουθία $\{z_k\}, k \in \mathbb{N}$ να είναι η εφικτή ακολουθία (feasible sequence) που συγκλίνει στο x, όταν:*

$$z_k \in \mathcal{U}, \forall k \in \mathbb{N} \;\; \textit{και} \;\; z_k \to x.$$

Ορισμός 2.10 (βλ. [9], Κεφάλαιο 12, Ορισμός, σελ. 316). *Ένα σημείο x^* είναι τοπική λύση του προβλήματος (2.1) εάν για κάθε εφικτή ακολουθία $\{z_k\}$ που συγκλίνει στο x^* ισχύει:*

$$f(z_k) \geq f(x^*), \quad \textit{για κάθε} \;\; k \in \mathbb{N}.$$

Ορισμός 2.11 (βλ. [9], Κεφάλαιο 12, Ορισμός, σελ. 316). *Ορίζουμε ως εφαπτομένη, την περιοριστική διεύθυνση (limiting direction) μίας εφικτής ακολουθίας.*

Ορισμός 2.12 (βλ. [9], Κεφάλαιο 12, Ορισμός 12.2, σελ. 316). *Ορίζουμε ως εφαπτομένη (ή εφαπτόμενο διάνυσμα) του συνόλου $\mathcal{U}$ στο σημείο x, ένα διάνυσμα d για το οποίο ισχύει:*

$$\lim_{k \to \infty} \frac{z_k - x}{t_k} = d,$$

όπου $\{z_k\}$ μία εφικτή ακολουθία που συγκλίνει στο x και $\{t_k\}$ μία ακολουθία θετικών αριθμών με $t_k \to 0$.

Το σύνολο όλων των εφαπτομένων του $\mathcal{U}$ στο σημείο x^* καλείται εφαπτόμενος κώνος και συμβολίζεται με $T_{\mathcal{U}}(x^*)$.

Παρατήρηση

Ο εφαπτόμενος κώνος είναι ένας κώνος σύμφωνα με τον ορισμό:

Ορίζουμε ως κώνο ένα σύνολο $\mathcal{F}$ εάν για κάθε $x \in \mathcal{F}$ ισχύει:

$$x \in \mathcal{F} \implies ax \in \mathcal{F}, \forall a > 0.$$

Πράγματι, αν d είναι ένα εφαπτόμενο διάνυσμα με τις αντίστοιχες ακολουθίες να είναι $\{z_k\}$ και $\{t_k\}$, τότε αντικαθιστώντας κάθε t_k με $a^{-1}t_k$, για κάποιο $a > 0$, έχουμε ότι $ad \in T_{\mathcal{U}}(x^*)$. Επίσης, $0 \in T_{\mathcal{U}}(x)$, αν θέσουμε $z_k \equiv x$ στον ορισμό της εφικτής ακολουθίας.

Ορισμός 2.13 (βλ. [9], Κεφάλαιο 12, Ορισμός 12.3, σελ. 316). *Έστω ένα εφικτό σημείο x και το ενεργό σύνολο $\mathcal{A}(x)$. Ορίζουμε το σύνολο των γραμμικοποιημένων εφικτών διευθύνσεων (linearized feasible directions) $\mathcal{F}(x)$ να είναι:*

$$\mathcal{F}(x) = \{d \mid d^T \nabla c_i(x) = 0, \forall i \in \mathcal{E}, d^T \nabla c_i(x) \geq 0, \forall i \in \mathcal{A}(x) \cap \mathcal{I}\}.$$

Παρατηρήσεις

1. Όπως και στον εφαπτόμενο κώνο είναι εύκολο να αποδείξουμε ότι το σύνολο $\mathcal{F}(x)$ είναι πράγματι ένας κώνος, σύμφωνα με τον ορισμό του κώνου που δώσαμε προηγουμένως.

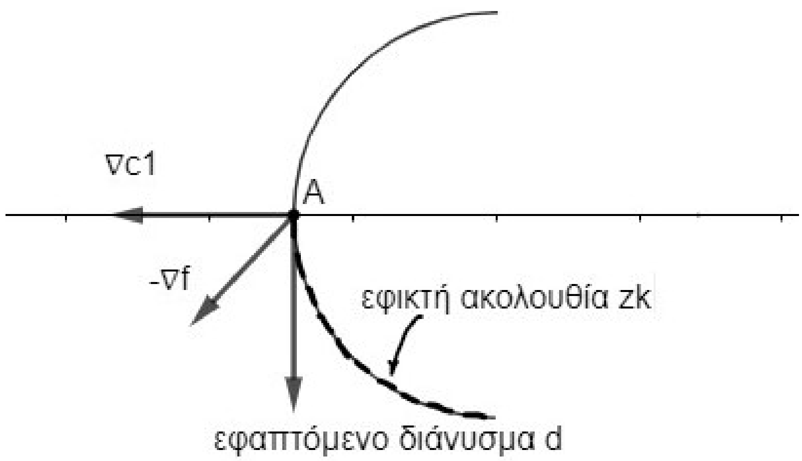
2. Ο ορισμός του εφαπτόμενου κώνου δεν βασίζεται στην αλγεβρική έκφραση του συνόλου $\mathcal{U}$, παρά μόνο στην γεωμετρική του ερμηνεία.

3. Το σύνολο των γραμμικοποιημένων εφικτών διευθύνσεων $\mathcal{F}(x)$ εξαρτάται από τον ορισμό των περιορισμών $c_i, i \in \mathcal{E} \cup \mathcal{I}$.

Παράδειγμα 2.4.1 (Παράδειγμα 2.3.1. αναθεωρημένο). *Έστω το πρόβλημα (2.5), όπου το εφικτό του σύνολο είναι ο κύκλος κέντρου* $O(0,0)$ *και ακτίνας* $\sqrt{2}$. *Στο Σχήμα 2.8 φαίνεται κοντά στο εφικτό σημείο* $x = (-\sqrt{2}, 0)^T$, *μία εφικτή ακολουθία η οποία συγκλίνει στο* x. *Η ακολουθία αυτή είναι:*

$$z_k = \begin{bmatrix} -\sqrt{2 - \dfrac{1}{k^2}} \\ -\dfrac{1}{k} \end{bmatrix}.$$

Επιλέγοντας $t_k = \|z_k - x\|$, *βρίσκουμε ότι το διάνυσμα* $d = (0, -1)^T$ *είναι εφαπτόμενο.*



Σχήμα 2.8: Το πρόβλημα (2.5), όπου φαίνεται η κλίση του περιορισμού και της αντικειμενικής συνάρτησης και η εφικτή ακολουθία

Η αντικειμενική συνάρτηση $f(x) = x_1 + x_2$ *αυξάνεται καθώς κινούμαστε πάνω στην ακολουθία* z_k. *Πράγματι:* $f(z_{k+1}) > f(z_k), \forall k = 2, 3, \ldots$ *και* $f(z_k) < f(x), \forall k = 2, 3, \ldots$. *Επομένως, το* x *δεν μπορεί να είναι λύση του προβλήματος.*

Μία ακόμη εφικτή ακολουθία είναι εκείνη που συγκλίνει στο $x = (-\sqrt{2}, 0)^T$ *από την αντίθετη φορά. Η ακολουθία αυτή είναι:*

$$z_k = \begin{bmatrix} -\sqrt{2 - \dfrac{1}{k^2}} \\ \dfrac{1}{k} \end{bmatrix}.$$

Η εφαπτομένη αυτής της ακολουθίας είναι η $d = (0, a)^T$ *και η αντικειμενική συνάρτηση μειώνεται καθώς κινούμαστε πάνω στην ακολουθία αυτή. Επομένως, ο εφαπτόμενος* *κώνος στο σημείο* $x = (-\sqrt{2}, 0)^T$ *είναι* $\{(0, d_2)^T | d_2 \in \mathbb{R}\}$.

Από τον Ορισμό 2.13. έχουμε ότι ένα $d = (d_1, d_2)^T \in \mathcal{F}(x)$ *εάν:*

$$0 = \nabla c_1(x)^T d = \begin{bmatrix} 2x_1 & 2x_2 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = -2\sqrt{2} d_1.$$

Επομένως, $\mathcal{F}(x) = \{(0, d_2)^T | d_2 \in \mathbb{R}\}$. *Παρατηρούμε ότι σε αυτή την περίπτωση ισχύει* $T_{\mathcal{U}}(x) = \mathcal{F}(x)$.

Παρατήρηση

Τα δύο σύνολα $T_{\mathcal{U}}(x)$ και $\mathcal{F}(x)$ δεν είναι πάντα ίσα.

Παράδειγμα 2.4.2. *Έστω ότι το εφικτό σύνολο του προηγούμενου προβλήματος ήταν διαφορετικό. Δηλαδή:*

$$\mathcal{U} = \{x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 : c_1(x) = 0\}, \quad \text{όπου} \;\; c_1(x) = (x_1^2 + x_2^2 - 2)^2.$$

Τότε ένα διάνυσμα $d = (d_1, d_2)^T \in \mathcal{F}(x)$ *εάν:*

$$0 = \nabla c_1(x)^T d = \begin{bmatrix} 4(x_1^2 + x_2^2 - 2)x_1 & 4(x_1^2 + x_2^2 - 2)x_2 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}.$$

Το οποίο ισχύει για κάθε $(d_1, d_2)^T \in \mathbb{R}^2$. *Επομένως, θα είναι* $\mathcal{F}(x) = \mathbb{R}^2$ *και έτσι, για αυτό το εφικτό σύνολο, ο εφαπτόμενος κώνος και το σύνολο των γραμμικοποιημένων εφικτών διευθύνσεων είναι διαφορετικά μεταξύ τους.*

Παράδειγμα 2.4.3 (Παράδειγμα 2.3.2. αναθεωρημένο)**.** *Έστω το πρόβλημα (2.11) με αντικειμενική συνάρτηση την* $f(x) = x_1 + x_2$ *και περιορισμό τον* $c_1(x) = 2 - x_1^2 - x_2^2$, *όπου το εφικτό του σύνολο είναι ο κύκλος κέντρου* $O(0, 0)$ *και ακτίνας* $\sqrt{2}$, *το σύνορο και το εσωτερικό του, και η λύση του προβλήματος είναι* $x^* = (-1, -1)^T$.

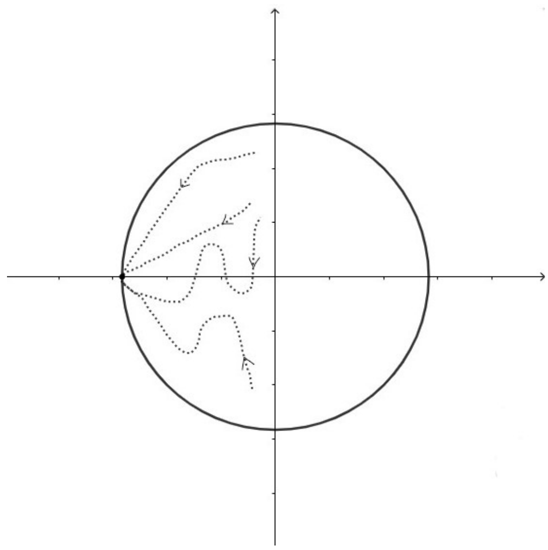
Στο Σχήμα 2.9 φαίνονται κοντά στο εφικτό σημείο $x = (-\sqrt{2}, 0)^T$, *οι διάφορες εφικτές ακολουθίες που συγκλίνουν στο* x, *οι οποίες είναι άπειρες και είναι δύο ειδών, αυτές που είναι κατά μήκος μιας ευθείας και εκείνες που είναι κατά μήκος μιας καμπύλης στο εσωτερικό του κύκλου.*

Οι ακολουθίες του προηγούμενου παραδείγματος:

$$z_k = \begin{bmatrix} -\sqrt{2 - \frac{1}{k^2}} \\ -\frac{1}{k} \end{bmatrix} \quad \text{και} \;\; z_k = \begin{bmatrix} -\sqrt{2 - \frac{1}{k^2}} \\ \frac{1}{k} \end{bmatrix},$$

είναι εφικτές και σε αυτό το πρόβλημα. Επομένως, οι εφικτές ακολουθίες στο σημείο $x = (-\sqrt{2}, 0)^T$, *που βρίσκονται κατά μήκος μίας ευθείας στο εσωτερικό του, έχουν την εξής μορφή:*

$$z_k = \begin{bmatrix} -\sqrt{2} \\ 0 \end{bmatrix} + \frac{1}{k} w, \quad \text{όπου} \;\; w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \quad \text{είναι κάθε διάνυσμα με} \;\; w_1 > 0.$$



Σχήμα 2.9: Το πρόβλημα (2.11), όπου φαίνονται εφικτές ακολουθίες σε ένα εφικτό σημείο

Το σημείο z_k είναι εφικτό εάν:

$$\|z_k\| \leq \sqrt{2} \implies \Big(-\sqrt{2} + \frac{w_1}{k} \Big)^2 + \Big(\frac{w_2}{k} \Big)^2 \leq 2.$$

Η παραπάνω συνθήκη ισχύει όταν $k \geq \dfrac{w_1^2 + w_2^2}{2\sqrt{2}w_1}$.

Επομένως, ο εφαπτόμενος κώνος στο σημείο $x = (-\sqrt{2}, 0)^T$ *είναι:*

$$\{(w_1, w_2)^T | w_1 \geq 0\}.$$

Για το πρόβλημα (2.11), από τον Ορισμό 2.13. έχουμε ότι ένα $d = (d_1, d_2)^T \in \mathcal{F}(x)$ *εάν:*

$$0 \leq \nabla c_1(x)^T d = \begin{bmatrix} -2x_1 & -2x_2 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = 2\sqrt{2}d_1.$$

Επομένως, $\mathcal{F}(x) = \{(d_1, d_2)^T | d_1 \geq 0\}$. *Παρατηρούμε ότι σε αυτή την περίπτωση ισχύει* $T_{\mathcal{U}}(x) = \mathcal{F}(x)$.

Παρατήρηση

Οι προϋποθέσεις που πρέπει να ικανοποιούν οι περιορισμοί είναι συνθήκες οι οποίες εξασφαλίζουν ότι τα δύο σύνολα, $T_{\mathcal{U}}(x)$ και $\mathcal{F}(x)$, είναι παρόμοια. Ειδικότερα, στις περισσότερες περιπτώσεις, αυτά τα σύνολα θα ταυτίζονται. Όπως προαναφέραμε, το σύνολο $\mathcal{F}(x)$ κατασκευάζεται μέσω της γραμμικοποίησης του συνόλου $\mathcal{U}$ στο εφικτό σημείο x, ενώ το σύνολο $T_{\mathcal{U}}(x)$ αποτελεί την γεωμετρική του έκφραση.

Παράδειγμα 2.4.4. *Έστω ένα πρόβλημα με τους εξής περιορισμούς:*

$$c_1(x) = 1 - x_1^2 - (x_2 - 1)^2 \geq 0 \quad \text{και} \quad c_2(x) = -x_2 \geq 0. \tag{2.21}$$

Από το Σχήμα 2.10 παρατηρούμε ότι το εφικτό σύνολο των παραπάνω περιορισμών είναι το σημείο $\mathcal{U} = \{(0,0)^T\}$. *Για το σημείο* $x = (0,0)^T$ *είναι προφανές ότι ο εφαπτόμενος* *κώνος είναι* $T_{\mathcal{U}}(x) = \{(0,0)^T\}$, *καθώς όλες οι εφικτές ακολουθίες που συγκλίνουν στο* x *πρέπει να είναι* $z_k = x = (0,0)^T$, *για κάθε* $k \in \mathbb{N}$. *Αντίστοιχα, το σύνολο των γραμμικοποιημένων εφικτών διευθύνσεων είναι το* $\mathcal{F}(x) = \{(d_1, 0)^T | d_1 \in \mathbb{R}\}$, *δηλαδή όλος ο οριζόντιος άξονας.*

Παρατήρηση

Στο παραπάνω πρόβλημα παρατηρούμε ότι το σύνολο $\mathcal{F}(x)$ δεν εκφράζει την γεωμετρία του εφικτού συνόλου $\mathcal{U}$. Επομένως, τα δύο σύνολα $T_{\mathcal{U}}(x)$ και $\mathcal{F}(x)$ δεν ταυτίζονται και άρα οι προϋποθέσεις των περιορισμών δεν ικανοποιούνται.

Ορισμός 2.14 (LICQ, βλ. [9], Κεφάλαιο 12, Ορισμός 12.4, σελ. 320)**.** *Έστω ένα σημείο* x *και το ενεργό σύνολο* $\mathcal{A}(x)$. *Λέμε ότι η προϋπόθεση της γραμμικής ανεξαρτησίας των περιορισμών (Linear Independence Constraint Qualification (LICQ)) ικανοποιείται, εάν το σύνολο που περιέχει τις κλίσεις των ενεργών περιορισμών, δηλαδή το σύνολο:*

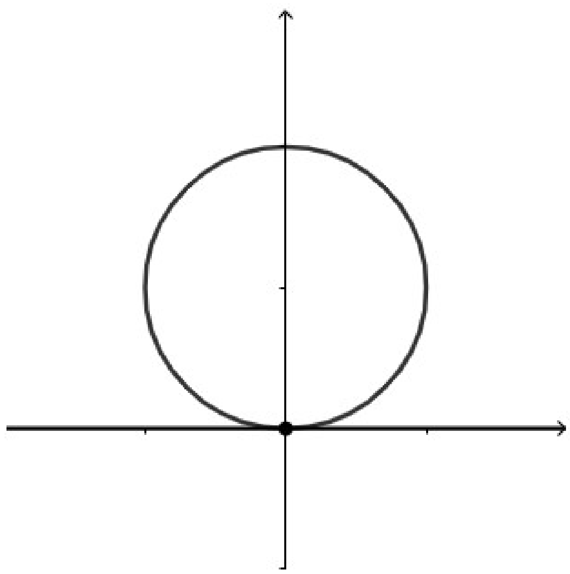
$$\{\nabla c_i(x), i \in \mathcal{A}(x)\},$$

είναι γραμμικά ανεξάρτητο.

Παρατηρήσεις

1. Παρατηρούμε ότι η παραπάνω προϋπόθεση δεν ικανοποιείται στα Παραδείγματα 2.4.2. και 2.4.4.

2. Γενικά εάν η προϋπόθεση LICQ ικανοποιείται, τότε οι κλίσεις όλων των ενεργών περιορισμών είναι διάφορες του μηδενός.



Σχήμα 2.10: Το πρόβλημα (2.21), όπου φαίνεται το εφικτό σύνολο να είναι το σημείο τομής ενός κύκλου και μιας ευθείας

2.5 Αναγκαίες συνθήκες πρώτης τάξης

Οι αναγκαίες συνθήκες που θα παρουσιαστούν στο επόμενο θεώρημα καλούνται πρώτης τάξης, καθώς περιλαμβάνουν τις κλίσεις της αντικειμενικής συνάρτησης και των περιορισμών. Επομένως, μέσω των παραγώγων πρώτης τάξης των συναρτήσεων αυτών, ορίζουμε τις αναγκαίες συνθήκες προκειμένου ένα σημείο να είναι τοπικό ελάχιστο της αντικειμενικής συνάρτησης.

Θεώρημα 2.1 (Αναγκαίες Συνθήκες Πρώτης Τάξης, βλ. [9], Κεφάλαιο 12, Θεώρημα 12.1, σελ. 321)**.** *Έστω ότι το x^* είναι τοπική λύση του γενικού προβλήματος (2.1), οι συναρτήσεις f και c_i είναι συνεχώς παραγωγίσιμες συναρτήσεις και η προϋπόθεση LICQ ικανοποιείται στο x^*. Τότε υπάρχει ένα διάνυσμα πολλαπλασιαστών Lagrange λ^*, με συνιστώσες λ_i^*, $i \in \mathcal{E} \cup \mathcal{I}$, τέτοιο ώστε οι παρακάτω συνθήκες να ικανοποιούνται στο (x^*, λ^*):*

$$\begin{aligned}
\nabla_x \mathcal{L}(x^*, \lambda^*) &= 0 && (2.22)\\
c_i(x^*) &= 0, \quad \text{για κάθε } i \in \mathcal{E} && (2.23)\\
c_i(x^*) &\geq 0, \quad \text{για κάθε } i \in \mathcal{I} && (2.24)\\
\lambda_i^* &\geq 0, \quad \text{για κάθε } i \in \mathcal{I} && (2.25)\\
\lambda_i^* c_i(x^*) &= 0, \quad \text{για κάθε } i \in \mathcal{E} \cup \mathcal{I}. && (2.26)
\end{aligned}$$

Οι συνθήκες (2.22) έως και (2.26) είναι γνωστές ως *συνθήκες Karush-Kuhn-Tucker* ή *KKT*. Οι συνθήκες (2.26) λέγονται *συνθήκες συμπληρωματικότητας (complementarity conditions)*.

Παρατηρήσεις

1. Οι συνθήκες συμπληρωματικότητας (2.26) εξασφαλίζουν ότι είτε ο περιορισμός $c_i(x^*)$ είναι ενεργός, είτε $\lambda_i^* = 0$, είτε και τα δύο.

2. Οι πολλαπλασιαστές Lagrange, που αντιστοιχούν σε ανενεργούς περιορισμούς, είναι ίσοι με μηδέν. Επομένως, εάν αφαιρέσουμε τους όρους για αυτούς τους δείκτες $i \notin \mathcal{A}(x^*)$, τότε η συνθήκη (2.22) γίνεται:

$$0 = \nabla_x \mathcal{L}(x^*, \lambda^*) = \nabla f(x^*) - \sum_{i \in \mathcal{A}(x^*)} \lambda_i^* \nabla c_i(x^*). \tag{2.27}$$

Ορισμός 2.15 (Αυστηρή Συμπληρωματικότητα, βλ. [9], Κεφάλαιο 12, Ορισμός 12.5, σελ. 321)**.** *Έστω ότι το x^* είναι τοπική λύση του γενικού προβλήματος (2.1) και ένα διάνυσμα λ^* ικανοποιεί τις συνθήκες KKT. Λέμε ότι η συνθήκη αυστηρής συμπληρωματικότητας (strict complementarity condition) ικανοποιείται εάν ακριβώς ένα από τα λ_i^* και $c_i(x^*)$ είναι μηδέν, για κάθε $i \in \mathcal{I}$. Δηλαδή εάν ισχύει:*

$$\lambda_i^* > 0, \ \forall\, i \in \mathcal{I} \cap \mathcal{A}(x^*).$$

Παρατηρήσεις

1. Όταν η συνθήκη αυστηρής συμπληρωματικότητας ικανοποιείται, τότε συνήθως είναι πιο εύκολο για τους αλγορίθμους να βρουν το ενεργό σύνολο $\mathcal{A}(x^*)$ και να συγκλίνουν πιο γρήγορα στην λύση x^*.

2. Για ένα πρόβλημα της μορφής (2.1) με λύση x^*, μπορεί να υπάρχουν πολλά διανύσματα λ^* που να ικανοποιούν τις συνθήκες KKT. Ωστόσο, όταν η προϋπόθεση LICQ ικανοποιείται το βέλτιστο διάνυσμα λ^* είναι μοναδικό.

Παράδειγμα 2.5.1. *Έστω το πρόβλημα:*

$$\min_x \left(x_1 - \frac{3}{2}\right)^2 + \left(x_2 - \frac{1}{2}\right)^4,$$

$$\text{υπό τους περιορισμούς} \quad \begin{cases} -x_1 - x_2 + 1 \geq 0 \\ -x_1 + x_2 + 1 \geq 0 \\ \ \ x_1 - x_2 + 1 \geq 0 \\ \ \ x_1 + x_2 + 1 \geq 0 \end{cases}. \tag{2.28}$$

Από το Σχήμα 2.11, παρατηρούμε ότι η λύση του προβλήματος είναι η $x^* = (1,0)^T$, στην οποία οι δύο πρώτοι περιορισμοί του προβλήματος είναι ενεργοί. Επομένως έχουμε:

$$\nabla f(x) = \begin{bmatrix} 2\left(x_1 - \frac{3}{2}\right) \\ 4\left(x_2 - \frac{1}{2}\right)^3 \end{bmatrix}, \nabla c_1(x) = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \nabla c_2(x) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Άρα:

$$\nabla f(x^*) = \begin{bmatrix} -1 \\ -\frac{1}{2} \end{bmatrix}, \nabla c_1(x^*) = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \nabla c_2(x^*) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Τότε από την πρώτη συνθήκη KKT έχουμε:

$$0 = \nabla_x \mathcal{L}(x^*, \lambda^*) = \nabla f(x^*) - \lambda_1^* \nabla c_1(x^*) - \lambda_2^* \nabla c_2(x^*)$$

$$\implies 0 = \begin{bmatrix} -1 \\ -\frac{1}{2} \end{bmatrix} - \lambda_1^* \begin{bmatrix} -1 \\ -1 \end{bmatrix} - \lambda_2^* \begin{bmatrix} -1 \\ 1 \end{bmatrix} \implies \begin{cases} \lambda_1^* + \lambda_2^* - 1 = 0 \\ \lambda_1^* - \lambda_2^* - \frac{1}{2} = 0 \end{cases}$$

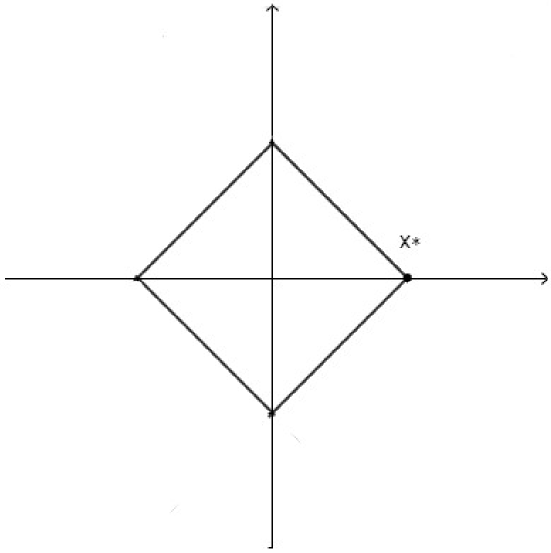
$$\implies \lambda_1^* = \frac{3}{4} \ \text{ και } \ \lambda_2^* = \frac{1}{4}.$$

Επομένως, οι συνθήκες KKT ικανοποιούνται όταν:

$$\lambda^* = \left(\frac{3}{4}, \frac{1}{4}, 0, 0\right)^T.$$

2.6 Αναγκαίες συνθήκες πρώτης τάξης: Η απόδειξη

Σε αυτή την ενότητα θα παρουσιάσουμε την απόδειξη του Θεωρήματος 2.1. ή αλλιώς Karush-Kuhn-Tucker. Για να μπορέσουμε να αποδείξουμε το θεώρημα, θα πρέπει πρώτα να αναλύσουμε κάποια χρήσιμα αποτελέσματα τα οποία είναι βασικά για την ολοκλήρωση της απόδειξης αλλά και για την βελτιστοποίηση γενικότερα.



Σχήμα 2.11: Το πρόβλημα (2.28), όπου φαίνονται οι περιορισμοί και η λύση του προβλήματος στο $(1,0)^T$

2.6.1 Σχέση εφαπτόμενου κώνου και συνόλου εφικτών διευθύνσεων πρώτης τάξης

Κάνοντας χρήση της προϋπόθεσης LICQ, θα συσχετίσουμε τον εφαπτόμενο κώνο $T_{\mathcal{U}}(x)$ με το σύνολο των εφικτών διευθύνσεων πρώτης τάξης $\mathcal{F}(x)$. Στην παρακάτω απόδειξη καθώς και στα επόμενα αποτελέσματα, θα συμβολίζουμε με $\mathcal{A}(x^*)^T$ τον πίνακα που έχει ως γραμμές τις κλίσεις των ενεργών περιορισμών στην βέλτιστη λύση. Δηλαδή το σύνολο:

$$\mathcal{A}(x^*)^T = \{\nabla c_i(x^*)\}, \quad \text{όπου} \;\; i \in \mathcal{A}(x^*) \;\; \text{και} \;\; \mathcal{A}(x^*) \;\; \text{το ενεργό σύνολο.}$$

Λήμμα 2.1 (βλ. [9], Κεφάλαιο 12, Λήμμα 12.2, σελ. 323)**.** *Έστω ένα εφικτό σημείο x^*. Τότε ισχύουν:*

(i) $T_{\mathcal{U}}(x^*) \subset \mathcal{F}(x^*)$.

(ii) Εάν η προϋπόθεση LICQ ικανοποιείται στο x^*, τότε $T_{\mathcal{U}}(x^*) = \mathcal{F}(x^*)$.

Απόδειξη. Χωρίς βλάβη της γενικότητας, υποθέτουμε ότι όλοι οι περιορισμοί $c_i(x), i = 1, 2, \ldots, m$ είναι ενεργοί στο x^*.

(i) $\Longrightarrow$ Έστω $\{z_k\}$ και $\{t_k\}$ ακολουθίες που ικανοποιούν την ισότητα:

$$\lim_{k\to\infty} \frac{z_k - x}{t_k} = d, \quad \text{με} \;\; t_k > 0 \;\; \text{για κάθε} \;\; k \in \mathbb{N}. \tag{2.29}$$

Τότε από τον ορισμό έχουμε:

$$z_k = x^* + t_k d + o(t_k).$$

Αφού $c_i(z_k) = 0$, για $i \in \mathcal{E}$, από τον τύπο Taylor έχουμε:

$$\begin{aligned} 0 &= \frac{1}{t_k} c_i(z_k) \\ &= \frac{1}{t_k}[c_i(x^*) + t_k \nabla c_i(x^*)^T d + o(t_k)] \\ &= \nabla c_i(x^*)^T d + \frac{o(t_k)}{t_k}. \end{aligned}$$

Στο παραπάνω, παίρνοντας όριο με $k \to \infty$, ο τελευταίος όρος εξαφανίζεται και έτσι έχουμε το ζητούμενο:

$$\nabla c_i(x^*)^T d = 0.$$

Αντίστοιχα για τους ενεργούς ανισοτικούς περιορισμούς, όπου $i \in \mathcal{A}(x^*) \cap \mathcal{I}$, έχουμε:

$$\begin{aligned} 0 &\leq \frac{1}{t_k} c_i(z_k) \\ &= \frac{1}{t_k}[c_i(x^*) + t_k \nabla c_i(x^*)^T d + o(t_k)] \\ &= \nabla c_i(x^*)^T d + \frac{o(t_k)}{t_k}. \end{aligned}$$

Παίρνοντας όριο με $k \to \infty$, ο τελευταίος όρος εξαφανίζεται και έτσι έχουμε το ζητούμενο:

$$\nabla c_i(x^*)^T d \geq 0.$$

(ii) $\Longrightarrow$ Για την απόδειξη αυτού, κάνουμε χρήση του Θεωρήματος Α΄.1 (βλ. Παράρτημα Α΄.1). Αρχικά, εφόσον η προϋπόθεση LICQ ικανοποιείται, από τον ορισμό της έχουμε ότι ο πίνακας $\mathcal{A}_{m\times n}(x^*)^T$ των κλίσεων των ενεργών περιορισμών είναι πλήρους τάξης m. Έστω Z ένας πίνακας του οποίου οι στήλες είναι βάση του μηδενοχώρου του πίνακα $\mathcal{A}(x^*)^T$. Τότε:

$$Z \in \mathbb{R}^{n\times(n-m)}, Z \text{ είναι πλήρους τάξης ως προς τις στήλες του }, \mathcal{A}(x^*)^T Z = 0.$$

Επιλέγουμε $d \in \mathcal{F}(x^*)$ και υποθέτουμε ότι η $\{t_k\}_{k=0}^{\infty}$ είναι μια ακολουθία θετικών αριθμών τέτοια ώστε $\lim_{k\to\infty} t_k = 0$. Ορίζουμε το παραμετρικό σύστημα εξισώσεων $R : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ ως εξής:

$$R(z,t) = \begin{bmatrix} c(z) - t\mathcal{A}(x^*)d \\ Z^T(z - x^* - td) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{2.30}$$

Απαιτούμε η λύση $z = z_k$ του παραπάνω συστήματος, για μικρό $t = t_k > 0$, να δίνει μία εφικτή ακολουθία η οποια να συγκλίνει στο x^* και να ικανοποιεί την ισότητα (2.29).

Αφού $\mathcal{A}(x^*) = \{\nabla c_i(x^*)\}_{i\in\mathcal{A}(x^*)}$, τότε για $t = 0$ και $z = x^*$ ο Ιακωβιανός πίνακας του συστήματος R είναι:

$$\nabla_z R(x^*, 0) = \begin{bmatrix} \mathcal{A}(x^*) \\ Z^T \end{bmatrix},$$

ο οποίος είναι αντιστρέψιμος λόγω της κατασκευής του Z. Έτσι, σύμφωνα με το Θεώρημα Α′.1 (Θεώρημα Πεπλεγμένων Συναρτήσεων), το σύστημα (2.30) έχει μοναδική λύση z_k για κάθε t_k αρκετά μικρό. Επίσης, από το σύστημα (2.30) και από τον ορισμό του συνόλου των εφικτών διευθύνσεων πρώτης τάξης έχουμε:

$$i \in \mathcal{E} \implies c_i(z_k) = t_k \nabla c_i(x^*)^T d = 0,$$

$$i \in \mathcal{A}(x^*) \cap \mathcal{I} \implies c_i(z_k) = t_k \nabla c_i(x^*)^T d \geq 0,$$

έτσι ώστε η z_k να είναι εφικτή.

Θα αποδείξουμε ότι η ισότητα (2.29) ικανοποιείται για αυτη την $\{z_k\}$. Αφού $R(z_k, t_k) = 0$, για κάθε $k \in \mathbb{N}$, από το Θεώρημα Taylor έχουμε:

$$\begin{aligned} 0 = R(z_k, t_k) &= \begin{bmatrix} c(z_k) - t_k \mathcal{A}(x^*) d \\ Z^T(z_k - x^* - t_k d) \end{bmatrix} \\ &= \begin{bmatrix} \mathcal{A}(x^*)(z_k - x^*) + o(\|z_k - x^*\|) - t_k \mathcal{A}(x^*) d \\ Z^T(z_k - x^* - t_k d) \end{bmatrix} \\ &= \begin{bmatrix} \mathcal{A}(x^*) \\ Z^T \end{bmatrix} (z_k - x^* - t_k d) + o(\|z_k - x^*\|). \end{aligned}$$

Διαιρώντας το παραπάνω με t_k και χρησιμοποιώντας την αντιστρεψιμότητα του πίνακα, έχουμε:

$$\frac{z_k - x^*}{t_k} = d + o\left(\frac{\|z_k - x^*\|}{t_k}\right).$$

Επομένως, η (2.29) ικανοποιείται για $x = x^*$, οπότε $d \in T_{\mathcal{U}}(x^*)$ για ένα αυθαίρετο $d \in \mathcal{F}(x^*)$ και έτσι έχουμε το ζητούμενο. □

2.6.2 Μία θεμελιώδης αναγκαία συνθήκη

Από τον Ορισμό 2.10. είδαμε ότι ένα x^* είναι τοπική λύση του γενικού προβλήματος (2.1), εάν για κάθε εφικτή ακολουθία που συγκλίνει στο σημείο αυτο ισχύει ότι $f(z_k) \geq f(x^*), \forall k \in \mathbb{N}$. Στο παρακάτω θεώρημα θα δούμε ότι εάν μια τέτοια ακολουθία υπάρχει, τότε το γινόμενο των περιοριστικών της διευθύνσεων d με την κλίση της αντικειμενικής συνάρτησης είναι μη αρνητικό.

Θεώρημα 2.2 (βλ. [9], Κεφάλαιο 12, Θεώρημα 12.3, σελ. 325)**.** *Έστω x^* μία τοπική λύση του γενικού προβλήματος (2.1). Τότε:*

$$\nabla f(x^*)^T d \geq 0, \quad \textit{για κάθε} \;\; d \in T_{\mathcal{U}}(x^*).$$

Απόδειξη. Έστω ότι υπάρχει εφαπτομένη d τέτοια ώστε να ισχύει $\nabla f(x^*)^T d < 0$. Έστω επίσης $\{z_k\}$ και $\{t_k\}$ ακολουθίες που να ικανοποιούν τον Ορισμό 2.12 για αυτό το d, δηλαδή:

$$\lim_{k \to \infty} \frac{z_k - x}{t_k} = d, \quad \text{με} \;\; t_k \to 0 \;\; \text{για κάθε} \;\; k \in \mathbb{N}.$$

Τότε αφού
$$z_k = x^* + t_k d + o(t_k),$$
έχουμε ότι:
$$\begin{aligned} f(z_k) &= f(x^*) + (z_k - x^*)^T \nabla f(x^*) + o(\|z_k - x^*\|) \\ &= f(x^*) + t_k d^T \nabla f(x^*) + o(t_k). \end{aligned}$$
Καθώς ισχύει $\nabla f(x^*)^T d = d^T \nabla f(x^*) < 0$, ο τελευταίος όρος τελικά εξαφανίζεται και έχουμε:
$$f(z_k) < f(x^*) + \frac{1}{2} t_k d^T \nabla f(x^*), \quad \text{για κάθε } \; k \; \text{ αρκετά μεγάλο.}$$
Έστω ανοιχτή περιοχή κοντά στο x^*. Τότε διαλέγοντας ένα k αρκετά μεγάλο έτσι ώστε η $\{z_k\}$ να βρίσκεται μέσα σε αυτή την περιοχή, έχουμε μία μικρότερη τιμή της αντικειμενικής συνάρτησης f. Τότε το x^* δεν είναι τοπική λύση του προβλήματος. Αυτό όμως είναι άτοπο, άρα έχουμε το ζητούμενο. □

Παρατήρηση

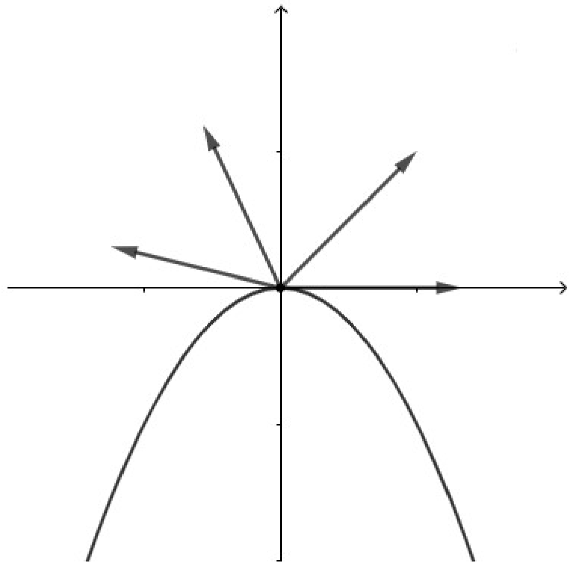
Το αντίστροφο του παραπάνω θεωρήματος δεν ισχύει πάντα. Δηλαδή εάν ισχύει $\nabla f(x^*)^T d \geq 0, \;$ για κάθε $\; d \in T_{\mathcal{U}}(x^*)$, το x^* δεν είναι απαραίτητα τοπική λύση του προβλήματος (2.1).

Παράδειγμα 2.6.1. *Έστω το πρόβλημα:*
$$\min x_2, \quad \text{υπό τον περιορισμό } \; x_2 \geq -x_1^2. \tag{2.31}$$
Από το Σχήμα 2.12, παρατηρούμε ότι στο εφικτό σημείο $x = (0,0)^T$, ενώ προφανώς δεν είναι λύση του προβλήματος, οι περιοριστικές διευθύνσεις d των εφικτών ακολουθιών στο σημείο αυτό πρέπει να έχουν $d_2 \geq 0$, έτσι ώστε $\nabla f(x^*)^T d = d_2 \geq 0$. Το σημείο $(a, -a^2)^T$, με $a > 0$, δίνει μικρότερη τιμή στην αντικειμενική συνάρτηση από το $x = (0,0)^T$ και μπορεί να φτάσει οσοδήποτε κοντά του, για a κατάλληλα μικρό.

2.6.3 Το λήμμα του Farkas

Ένα σημαντικό βήμα για την απόδειξη του Θεωρηματος KKT είναι το Λήμμα του Farkas. Στο Λήμμα αυτό αναφέρεται ένας κώνος K με τον εξής ορισμό:
$$K = \{By + Cw \mid y \geq 0\}, \tag{2.32}$$
όπου B, C είναι πίνακες διάστασης $n \times m$ και $n \times p$ αντίστοιχα, y, w είναι διανύσματα κατάλληλων διαστάσεων και ο συμβολισμός $y \geq 0$ σημαίνει ότι $y_i \geq 0$, για κάθε $i = 1, \ldots, m$.

Δεδομένου ενός διανύσματος $g \in \mathbb{R}^n$, το Λήμμα του Farkas αποδεικνύει ότι μία από τις δύο περιπτώσεις μπορεί να ισχύει. Είτε $g \in K$, είτε υπάρχει διάνυσμα $d \in \mathbb{R}^n$ το οποίο ορίζει ένα υπερεπίπεδο που τα διαχωρίζει. Οι δύο αυτές περιπτώσεις απεικονίζονται στο Σχήμα 2.13 με τον πίνακα B να έχει τρεις στήλες, τον C κενό, $n = 2$ και το d να είναι ένα επίπεδο στον $\mathbb{R}^n$ που διαχωρίζει το διάνυσμα g από τον κώνο K.



Σχήμα 2.12: Το πρόβλημα (2.31), όπου φαίνονται οι διάφορες περιοριστικές διευθύνσεις των εφικτών ακολουθιών στο σημείο $(0,0)^T$

Λήμμα 2.2 (Λήμμα του Farkas, βλ. [9], Κεφάλαιο 12, Λήμμα 12.4, σελ. 327)**.** *Έστω κώνος K της μορφής (2.32). Δεδομένου ενός διανύσματος $g \in \mathbb{R}^n$, μπορεί να ισχύει ένα από τα παρακάτω:*

(i) $g \in K$.

(ii) *Υπάρχει $d \in \mathbb{R}^n$ τέτοιο ώστε:*

$$g^T d < 0, \ B^T d \geq 0, \ C^T d = 0,$$

όπου $B^T d \geq 0$ με κάθε συνιστώσα του παραγόμενου διανύσματος $B^T d$ να είναι μη αρνητική.

Απόδειξη. Αρχικά θα αποδείξουμε ότι τα (i), (ii) δεν μπορούν να ισχύουν ταυτόχρονα.

Εάν $g \in K$, τότε υπάρχουν διανύσματα $y \geq 0$ και w τέτοια ώστε:

$$g = By + Cw.$$

Εάν επίσης υπάρχει ένα διάνυσμα d για το οποίο να ισχύουν:

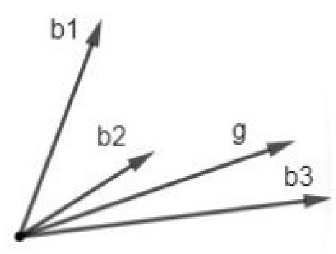
$$g^T d < 0, \ B^T d \geq 0, \ C^T d = 0,$$

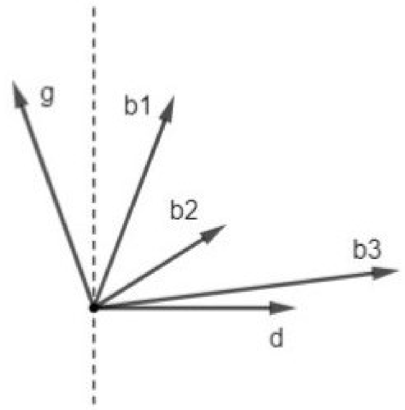
τότε έχουμε:

$$0 > d^T g = d^T By + d^T Cw = (B^T d)^T y + (C^T d)^T w \geq 0,$$

όπου η τελευταία ανισότητα προκύπτει από τα $C^T d = 0, \ B^T d \geq 0$ και $y \geq 0$. Επομένως, τα (i), (ii) δεν μπορούν να ισχύουν ταυτόχρονα.

Στην συνέχεια θα αποδείξουμε ότι ισχύει ένα από τα (i), (ii). Πιο συγκεκριμένα, πως κατασκευάζουμε ένα διάνυσμα d με τις ιδιότητες που περιγράφονται στο (ii) καθώς





Σχήμα 2.13: Το Λήμμα του Farkas, όπου φαίνονται οι δύο περιπτώσεις $g \in K$ (αριστερά) ή υπάρχει ένα υπερεπίπεδο που τα διαχωρίζει (δεξιά)

$g \notin K$. Για αυτή την απόδειξη θα χρησιμοποιήσουμε το γεγονός ότι το K είναι κλειστό σύνολο (βλ. Λήμμα Α′.1 - Παράρτημα Α′.2).

Έστω διάνυσμα $\hat{s} \in K$ του οποίου η απόσταση από το g είναι μικρή. Αφού το K είναι κλειστό, το $\hat{s}$ είναι καλά ορισμένο και είναι η λύση του παρακάτω προβλήματος βελτιστοποίησης:

$$\min \|s - g\|_2^2, \quad \text{υπό τον περιορισμό} \quad s \in K.$$

Αφού $\hat{s} \in K$ και K κώνος, τότε $a\hat{s} \in K$ για κάποιο $a \in \mathbb{R}, a \geq 0$. Καθώς το $\|a\hat{s} - g\|_2^2$ ελαχιστοποιείται για $a = 1$, έχουμε:

$$\frac{d}{da}\|a\hat{s} - g\|_2^2\bigg|_{a=1} = 0 \implies (-2\hat{s}^T g + 2t\hat{s}^T\hat{s})\bigg|_{a=1} = 0 \tag{2.33}$$

$$\implies \hat{s}^T(\hat{s} - g) = 0. \tag{2.34}$$

Έστω s τυχαίο διάνυσμα που ανήκει στο K. Αφού K είναι κυρτό σύνολο, έχουμε:

$$\|\hat{s} + \theta(s - \hat{s}) - g\|_2^2 \geq \|\hat{s} - g\|_2^2, \text{ για κάθε } \theta \in [0, 1].$$

Επομένως:

$$\begin{aligned}
& (\hat{s} + \theta(s - \hat{s}) - g)^T(\hat{s} + \theta(s - \hat{s}) - g) - (\hat{s} - g)^T(\hat{s} - g) & \geq 0 \\
\implies & ((\hat{s} - g)^T + \theta(s - \hat{s})^T)((\hat{s} - g) + \theta(s - \hat{s})) - (\hat{s} - g)^T(\hat{s} - g) & \geq 0 \\
\implies & \theta(s - \hat{s})^T(\hat{s} - g) + \theta(\hat{s} - g)^T(s - \hat{s}) + \theta^2(s - \hat{s})^T(s - \hat{s}) & \geq 0 \\
\implies & \theta(s - \hat{s})^T(\hat{s} - g) + \theta(s - \hat{s})^T(\hat{s} - g) + \theta^2(s - \hat{s})^T(s - \hat{s}) & \geq 0.
\end{aligned}$$

Τότε:

$$2\theta(s - \hat{s})^T(\hat{s} - g) + \theta^2\|s - \hat{s}\|_2^2 \geq 0.$$

Διαιρώντας το παραπάνω με θ και παίρνοντας όριο για $\theta \to 0$ έχουμε ότι $(s - \hat{s})^T(\hat{s} - g) \geq 0$. Τότε από την (2.34) έχουμε:

$$s^T(\hat{s} - g) \geq 0, \quad \text{για κάθε} \quad s \in K. \tag{2.35}$$

Απαιτούμε το διάνυσμα $d = \hat{s} - g$ να ικανοποιεί τις συνθήκες:

$$g^T d < 0,\ B^T d \geq 0,\ C^T d = 0.$$

Το $d \neq 0$ αφού $g \notin K$. Τότε από την (2.34) έχουμε:

$$d^T g = d^T(\hat{s} - d) = (\hat{s} - g)^T \hat{s} - d^T d = -\|d\|_2^2 < 0,$$

έτσι ώστε το d να ικανοποιεί την $g^T d < 0$.

Από την (2.35) έχουμε ότι $d^T s \geq 0$, για κάθε $s \in K$, έτσι ώστε:

$$d^T(By + Cw) \geq 0, \quad \text{για κάθε} \ \ y \geq 0 \ \ \text{και} \ \ w.$$

Θέτοντας $y = 0$ έχουμε ότι $(C^T d)^T w \geq 0$, για κάθε w, το οποίο ισχύει μόνο όταν $C^T d = 0$. Θέτοντας $w = 0$ έχουμε ότι $(B^T d)^T y \geq 0$, για κάθε $y \geq 0$, το οποίο ισχύει μόνο όταν $B^T d \geq 0$. Επομένως, το d ικανοποιεί όλες τις συνθήκες. □

Παρατήρηση

Έστω κώνος N που ορίζεται ως εξής:

$$N = \left\{ \sum_{i \in \mathcal{A}(x^*)} \lambda_i \nabla c_i(x^*), \lambda_i \geq 0,\ \forall i \in \mathcal{A}(x^*) \cap \mathcal{I} \right\}.$$

Εφαρμόζοντας το Λήμμα 2.2 (Λήμμα του Farkas) στον κώνο N και θέτοντας $g = \nabla f(x^*)$, ένα από τα δύο μπορεί να ισχύει:

(i)

$$\nabla f(x^*) = \sum_{i \in \mathcal{A}(x^*)} \lambda_i \nabla c_i(x^*) = \mathcal{A}(x^*)^T \lambda^*, \lambda_i \geq 0,\ \forall i \in \mathcal{A}(x^*) \cap \mathcal{I} \qquad (2.36)$$

(ii) Υπάρχει διεύθυνση d τέτοια ώστε:

$$d^T \nabla f(x^*) < 0 \ \ \text{και} \ \ d \in \mathcal{F}(x^*),$$

όπου $\mathcal{F}(x^*) = \{d \mid d^T \nabla c_i(x^*) = 0, \forall i \in \mathcal{E}, d^T \nabla c_i(x^*) \geq 0, \forall i \in \mathcal{A}(x^*) \cap \mathcal{I}\}$ το σύνολο των γραμμικοποιημένων εφικτών διευθύνσεων.

2.6.4 Απόδειξη Θεωρήματος Karush-Kuhn-Tucker

Συνδυάζοντας τα Λήμματα 2.1 και 2.2 έχουμε τις συνθήκες KKT που περιγράφονται στο Θεώρημα 2.1 (Θεώρημα Karush-Kuhn-Tucker). Έτσι καταλήγουμε στο τελευταίο κομμάτι της απόδειξης.

Έστω εφικτό σημείο $x^* \in \mathbb{R}^n$ στο οποίο ικανοποιείται η προϋπόθεση LICQ. Από το Θεώρημα 2.1 έχουμε ότι αν το σημείο x^* είναι τοπική λύση του γενικού προβλήματος (2.1), τότε υπάρχει διάνυσμα $\lambda^* \in \mathbb{R}^m$, με $\lambda_i^* \in \mathbb{R}, \forall i = 1, \ldots, m$, το οποίο ικανοποιεί τις συνθήκες (2.22) έως και (2.26) (συνθήκες KKT). Αρχικά θα δείξουμε ότι υπάρχουν πολλαπλασιαστές $\lambda_i, i \in \mathcal{A}(x^*)$ τέτοιοι ώστε η (2.36) να ικανοποιείται.

Από το Θεώρημα 2.2 έχουμε ότι $d^T \nabla f(x^*) \geq 0$, για κάθε διάνυσμα εφαπτομένων $d \in T_{\mathcal{U}}(x^*)$. Από το Λήμμα 2.1, αφού η προϋπόθεση LICQ ικανοποιείται, έχουμε ότι $T_{\mathcal{U}}(x^*) = \mathcal{F}(x^*)$. Επομένως:

$$d^T \nabla f(x^*) \geq 0,\ \forall d \in \mathcal{F}(x^*).$$

Τότε από το Λήμμα 2.2 υπάρχει διάνυσμα λ που να ικανοποιεί την (2.36). Ορίζουμε διάνυσμα λ^*, ως εξής:

$$\lambda_i^* = \begin{cases} \lambda_i,\ \forall i \in \mathcal{A}(x^*) \\ 0,\ \forall i \in \mathcal{I} \backslash \mathcal{A}(x^*) \end{cases}. \tag{2.37}$$

Θα αποδείξουμε ότι αυτό το λ^* με την τοπική λύση x^* ικανοποιούν τις συνθήκες KKT:

- Η συνθήκη (2.22) προκύπτει από την συνθήκη (2.36) και τους Ορισμούς 2.4 της συνάρτησης Lagrange και (2.37) του διανύσματος λ^*.
- Καθώς το x^* είναι εφικτό σημείο, οι συνθήκες (2.23) και (2.24) ικανοποιούνται.
- Από την συνθήκη (2.36) έχουμε ότι $\lambda_i^* \geq 0,\ \forall i \in \mathcal{A}(x^*) \cap \mathcal{I}$. Από τον Ορισμό (2.37) του διανύσματος λ^* έχουμε ότι $\lambda_i^* = 0,\ \forall i \in \mathcal{I} \backslash \mathcal{A}(x^*)$. Επομένως, $\lambda_i^* \geq 0, \forall i \in \mathcal{I}$ και έτσι η συνθήκη (2.25) ικανοποιείται.
- Έχουμε ότι $c_i(x^*) = 0,\ \forall i \in \mathcal{A}(x^*) \cap \mathcal{I}$ και $\lambda_i^* = 0,\ \forall i \in \mathcal{I} \backslash \mathcal{A}(x^*)$. Επομένως, $\lambda_i^* c_i(x^*) = 0,\ \forall i \in \mathcal{I}$ και έτσι η συνθήκη (2.26) ικανοποιείται.

2.7 Συνθήκες δεύτερης τάξης

Στην προηγούμενη ενότητα αναλύσαμε τις αναγκαίες συνθήκες πρώτης τάξης, τις συνθήκες KKT, όπου σύμφωνα με αυτές είδαμε πως συνδέονται οι κλισεις της αντικειμενικής συνάρτησης και των περιορισμών σε μία λύση x^*. Όταν αυτές ικανοποιούνται, η μετακίνηση πάνω σε οποιοδήποτε διάνυσμα $w \in \mathcal{F}(x^*)$ είτε αυξάνει την τιμή της πρώτης τάξεως προσέγγισης, δηλαδή $w^T \nabla f(x^*) > 0$, είτε κρατάει σταθερή την τιμή της, δηλαδή $w^T \nabla f(x^*) = 0$.

Σε αυτή την ενότητα θα δούμε πως επηρεάζουν οι δευτέρας τάξεως παράγωγοι της αντικειμενικής συνάρτησης και των περιορισμών, τις αναγκαίες συνθήκες. Για τις διευθύνσεις $w \in \mathcal{F}(x^*)$, για τις οποίες ισχύει $w^T \nabla f(x^*) = 0$, δεν μπορούμε να αποφανθούμε εάν μία μετακίνηση πάνω σε αυτές θα προκαλέσει αύξηση ή μείωση στην τιμή της αντικειμενικής συνάρτησης χρησιμοποιώντας μόνο τις παραγώγους πρώτης τάξης, και για τον λόγο αυτό θα εξετάσουμε και τις παραγώγους δεύτερης τάξης. Έτσι, οι συνθήκες δεύτερης τάξης εξετάζουν τους όρους των παραγώγων δεύτερης τάξης στις σειρές Taylor της αντικειμενικής συνάρτησης και των περιορισμών, για να λύσουν το πρόβλημα της αύξησης ή της μείωσης της f. Ειδικότερα, οι συνθήκες δεύτερης τάξης εξετάζουν την καμπυλότητα της συνάρτησης Lagrange στις κρίσιμες διευθύνσεις, δηλαδή στις διευθύνσεις $w \in \mathcal{F}(x^*)$, για τις οποίες ισχύει $w^T \nabla f(x^*) = 0$.

Σε αυτη την ενότητα καθώς γίνεται χρήση των παραγώγων δεύτερης τάξης, χρειαζόμαστε κάποιες επιπλέον προϋποθέσεις για τις συναρτήσεις f και c_i. Επομένως, για αυτή την ενότητα υποθέτουμε ότι οι συναρτήσεις f και c_i, $i \in \mathcal{E} \cup \mathcal{I}$, είναι δύο φορές συνεχώς παραγωγίσιμες.

Ορισμός 2.16 (βλ. [9], Κεφάλαιο 12, Ορισμός, σελ. 330)**.** *Έστω ένα διάνυσμα πολλαπλασιαστών Lagrange λ^* το οποίο ικανοποιεί τις συνθήκες KKT. Ορίζουμε τον κρίσιμο κώνο $\mathcal{C}(x^*, \lambda^*)$ να είναι το σύνολο:*

$$\mathcal{C}(x^*, \lambda^*) = \{w \in \mathcal{F}(x^*) \mid \nabla c_i(x^*)^T w = 0,\ \forall i \in \mathcal{A}(x^*) \cap \mathcal{I} \ \ \mu\epsilon\ \ \lambda_i^* > 0\},$$

όπου $\mathcal{F}(x^*) = \{w \mid w^T \nabla c_i(x^*) = 0, \forall i \in \mathcal{E}, w^T \nabla c_i(x^*) \geq 0, \forall i \in \mathcal{A}(x^*) \cap \mathcal{I}\}$ το σύνολο των γραμμικοποιημένων εφικτών διευθύνσεων και $\mathcal{A}(x^*) = \mathcal{E} \cup \{i \in \mathcal{I} : c_i(x^*) = 0\}$ το ενεργό σύνολο στο x^*.

Ορισμός 2.17 (βλ. [9], Κεφάλαιο 12, Ορισμός, σελ. 330)**.** *Ένα διάνυσμα w ανήκει στον κώνο $\mathcal{C}(x^*, \lambda^*)$ εάν:*

$$w \in \mathcal{C}(x^*, \lambda^*) \Leftrightarrow \begin{cases} \nabla c_i(x^*)^T w = 0, & \forall i \in \mathcal{E} \\ \nabla c_i(x^*)^T w = 0, & \forall i \in \mathcal{A}(x^*) \cap \mathcal{I}\ \mu\epsilon\ \lambda_i^* > 0 \\ \nabla c_i(x^*)^T w \geq 0, & \forall i \in \mathcal{A}(x^*) \cap \mathcal{I}\ \mu\epsilon\ \lambda_i^* = 0. \end{cases}$$

Παρατήρηση

Από τον παραπάνω Ορισμό και από το γεγονός ότι $\lambda_i^* = 0$, για όλους τους ανενεργούς περιορισμούς με δείκτες $i \in \mathcal{I} \backslash \mathcal{A}(x^*)$, έχουμε ότι:

$$w \in \mathcal{C}(x^*, \lambda^*) \implies \lambda_i^* \nabla c_i(x^*)^T w = 0,\ \forall i \in \mathcal{E} \cup \mathcal{I}.$$

Επομένως, από την συνθήκη KKT (2.22) και τον Ορισμό 2.4 της συνάρτησης Lagrange, έχουμε ότι:

$$w \in \mathcal{C}(x^*, \lambda^*) \implies w^T \nabla f(x^*) = \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i^* w^T \nabla c_i(x^*) = 0.$$

Επομένως, ο κρίσιμος κώνος $\mathcal{C}(x^*, \lambda^*)$ αποτελείται από τις διευθύνσεις του συνόλου $\mathcal{F}(x^*)$, για τις οποίες δεν μπορούμε να αποφανθούμε από τις παραγώγους πρώτης τάξης εάν προκαλούν αύξηση ή μείωση στην τιμή της αντικειμενικής συνάρτησης.

Παράδειγμα 2.7.1. *Έστω το πρόβλημα:*

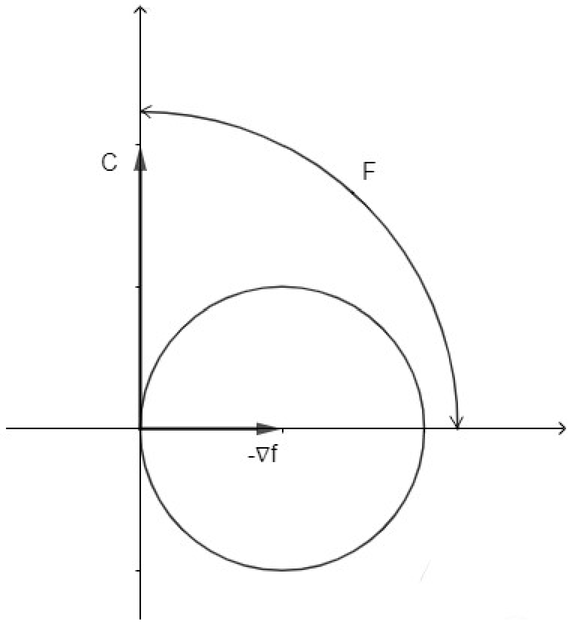
$$\min x_1, \quad \text{υπό τους περιορισμούς}\ \ x_2 \geq 0,\ 1 - (x_1 - 1)^2 - x_2^2 \geq 0. \tag{2.38}$$

Από το Σχήμα 2.14, παρατηρούμε ότι η λύση του προβλήματος είναι $x^* = (0, 0)^T$ και το ενεργό σύνολο στο σημείο αυτό είναι $\mathcal{A}(x^*) = \{1, 2\}$. Το βέλτιστο διάνυσμα των πολλαπλασιαστών Lagrange για την λύση αυτή είναι $\lambda^* = (0, 0.5)^T$. Η προϋπόθεση LICQ ικανοποιείται, καθώς στο σημείο x^* οι κλίσεις των ενεργών περιορισμών είναι $\nabla c_1(x^*) = (0, 1)^T$ και $\nabla c_2(x^*) = (2, 0)^T$, επομένως το διάνυσμα λ^* είναι μοναδικό. Τότε το σύνολο των γραμμικοποιημένων εφικτών διευθύνσεων είναι $\mathcal{F}(x^*) = \{d \mid d \geq$

$0\}$. *Αφού* $\lambda_1^* = 0$ *και ο* $c_1(x^*)$ *είναι ενεργός ανισοτικός περιορισμός, τότε από τον Ορισμό 2.17 έχουμε ότι:*

$$w \in \mathcal{C}(x^*, \lambda^*) \Leftrightarrow \nabla c_1(x^*)^T w \geq 0 \implies \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \geq 0 \implies w_2 \geq 0.$$

Άρα ο κρίσιμος κώνος είναι: $\mathcal{C}(x^*, \lambda^*) = \left\{ \begin{bmatrix} 0 & w_2 \end{bmatrix}^T | w_2 \geq 0 \right\}$.



Σχήμα 2.14: Το πρόβλημα (2.38), όπου φαίνονται τα σύνολα $\mathcal{F}(x^*)$ και $\mathcal{C}(x^*, \lambda^*)$

Στην συνέχεια θα αποδείξουμε δύο θεωρήματα τα οποία παρουσιάζουν τις ικανές και αναγκαίες συνθήκες, έτσι ώστε κάποιο x^* να είναι τοπική λύση του γενικού προβλήματος. Αυτά τα θεωρήματα εξετάζουν την καμπυλότητα της συνάρτησης Lagrange, μέσω του Εσσιανού της πίνακα, στις κρίσιμες διευθύνσεις, δηλαδή σε αυτές που ανήκουν στον κρίσιμο κώνο.

Θεώρημα 2.3 (Αναγκαίες Συνθήκες Δεύτερης Τάξης, βλ. [9], Κεφάλαιο 12, Θεώρημα 12.5, σελ. 332)**.** *Έστω* x^* *η τοπική λύση του γενικού προβλήματος (2.1) και ότι η προϋπόθεση LICQ ικανοποιείται. Έστω επίσης* λ^* *το διάνυσμα των πολλαπλασιαστών Lagrange που ικανοποιεί τις συνθήκες* KKT*. Τότε:*

$$w^T \nabla_{xx}^2 \mathcal{L}(x^*, \lambda^*) w \geq 0, \quad \text{για κάθε} \;\; w \in \mathcal{C}(x^*, \lambda^*). \tag{2.39}$$

Απόδειξη. Αφού το x^* είναι τοπική λύση του προβλήματος (2.1), τότε για κάθε εφικτή ακολουθία $\{z_k\}$ που συγκλίνει στο x^*, πρέπει να ισχύει $f(z_k) \geq f(x^*)$, για κάθε k αρκετά μεγάλο. Θα κατασκευάσουμε μία εφικτή ακολουθία της οποίας η περιοριστική διεύθυνση είναι η w, και θα δείξουμε ότι η συνθήκη $f(z_k) \geq f(x^*)$ οδηγεί στην συνθήκη (2.39).

Χρησιμοποιώντας την τεχνική με την οποία αποδείξαμε το Λήμμα 2.1, καθώς και το γεγονός ότι $w \in \mathcal{C}(x^*, \lambda^*) \subset \mathcal{F}(x^*)$, επιλέγουμε ακολουθία $\{t_k\}$, με $t_k \in \mathbb{R}, \forall k \in \mathbb{N}$

για να κατασκευάσουμε μία εφικτή ακολουθία $\{z_k\}$ που συγκλίνει στο x^* τέτοια ώστε:

$$\lim_{k\to\infty} \frac{z_k - x^*}{t_k} = w.$$

Το παραπάνω μπορεί να γραφτεί και ως:

$$z_k - x^* = t_k w + o(t_k). \tag{2.40}$$

Λόγω της κατασκευής της ακολουθίας $\{z_k\}$ και από τις συνθήκες της απόδειξης του Λήμματος 2.1:

$$i \in \mathcal{E} \implies c_i(z_k) = t_k \nabla c_i(x^*)^T d = 0,$$
$$i \in \mathcal{A}(x^*) \cap \mathcal{I} \implies c_i(z_k) = t_k \nabla c_i(x^*)^T d \geq 0,$$

έχουμε ότι:

$$c_i(z_k) = t_k \nabla c_i(x^*)^T w, \ \forall i \in \mathcal{A}(x^*). \tag{2.41}$$

Από τον Ορισμό 2.17 είδαμε ότι:

$$w \in \mathcal{C}(x^*, \lambda^*) \implies \lambda_i^* \nabla c_i(x^*)^T w = 0, \ \forall i \in \mathcal{E} \cup \mathcal{I}.$$

Από το παραπάνω και από την (2.41), η συνάρτηση Lagrange γίνεται:

$$\begin{aligned}
\mathcal{L}(z_k, \lambda^*) &= f(z_k) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i^* c_i(z_k) \\
&= f(z_k) - t_k \sum_{i \in \mathcal{A}(x^*)} \lambda_i^* \nabla c_i(x^*)^T w \\
&= f(z_k).
\end{aligned}$$

Χρησιμοποιώντας τον τύπο Taylor και την συνέχεια των Εσσιανών πινάκων $\nabla^2 f$ και $\nabla^2 c_i, i \in \mathcal{E} \cup \mathcal{I}$, μπορούμε να βρουμε μία προσέγγιση της συνάρτησης $\mathcal{L}(z_k, \lambda^*)$ κοντά στο x^*. Έτσι έχουμε:

$$\begin{aligned}
\mathcal{L}(z_k, \lambda^*) = \mathcal{L}(x^*, \lambda^*) &+ (z_k - x^*)^T \nabla_x \mathcal{L}(x^*, \lambda^*) \\
&+ \frac{1}{2}(z_k - x^*)^T \nabla_{xx}^2 \mathcal{L}(x^*, \lambda^*)(z_k - x^*) + o(\|z_k - x^*\|^2).
\end{aligned} \tag{2.42}$$

Από την συνθήκη συμπληρωματικότητας (2.26) έχουμε ότι:

$$\mathcal{L}(x^*, \lambda^*) = f(x^*). \tag{2.43}$$

Από τις συνθήκες (2.22) και (2.40) η συνθήκη (2.42) γράφεται ως εξής:

$$\mathcal{L}(z_k, \lambda^*) = f(x^*) + \frac{1}{2} t_k^2 w^T \nabla_{xx}^2 \mathcal{L}(x^*, \lambda^*) + o(t_k^2). \tag{2.44}$$

Λόγω της (2.43) αντικαθιστώντας στην (2.44) έχουμε:

$$f(z_k) = f(x^*) + \frac{1}{2} t_k^2 w^T \nabla_{xx}^2 \mathcal{L}(x^*, \lambda^*) w + o(t_k^2). \tag{2.45}$$

Εάν $w^T \nabla_{xx}^2 \mathcal{L}(x^*, \lambda^*) w < 0$, τότε από την (2.45) θα είχαμε ότι $f(z_k) < f(x^*)$, για κάθε k αρκετά μεγάλο, το οποίο είναι άτοπο, καθώς τότε το x^* δεν θα ήταν τοπική λύση του προβλήματος (2.1). Επομένως, η συνθήκη (2.39) ισχύει. □

Θεώρημα 2.4 (Ικανές Συνθήκες Δεύτερης Τάξης, βλ. [9], Κεφάλαιο 12, Θεώρημα 12.6, σελ. 333)**.** *Έστω ότι για κάποιο εφικτό σημείο $x^* \in \mathbb{R}^n$ υπάρχει ένα διάνυσμα πολλαπλασιαστών Lagrange λ^* που ικανοποιεί τις συνθήκες KKT. Έστω επίσης:*

$$w^T \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*) w > 0, \quad \text{για κάθε} \;\; w \in \mathcal{C}(x^*, \lambda^*), w \neq 0. \tag{2.46}$$

Τότε το x^* είναι αυστηρά τοπική λύση του γενικού προβλήματος (2.1).

Απόδειξη. Ορίζουμε το σύνολο $\bar{\mathcal{C}} = \{d \in \mathcal{C}(x^*, \lambda^*) \mid \|d\| = 1\}$, το οποίο είναι ενα συμπαγές υποσύνολο του $\mathcal{C}(x^*, \lambda^*)$, τότε από την συνθήκη (2.46) το ελάχιστο της $d^T \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*) d$, σε αυτό το σύνολο, είναι ένας θετικός αριθμός έστω σ. Αφού το σύνολο $\mathcal{C}(x^*, \lambda^*)$ είναι κώνος, έχουμε ότι το $\frac{w}{\|w\|} \in \bar{\mathcal{C}}$ αν και μόνο αν $w \in \mathcal{C}(x^*, \lambda^*), w \neq 0$. Επομένως, από την συνθήκη (2.46):

$$w^T \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*) w \geq \sigma \|w\|^2, \quad \text{για κάθε} \;\; w \in \mathcal{C}(x^*, \lambda^*), \sigma > 0.$$

Θα δείξουμε ότι το x^* είναι αυστηρά τοπική λύση του γενικού προβλήματος (2.1), δείχνοντας ότι για κάθε εφικτή ακολουθία $\{z_k\}$ που συγκλίνει στο x^*, ισχύει $f(z_k) \geq f(x^*) + \frac{\sigma}{4}\|z_k - x^*\|^2$, για κάθε k αρκετά μεγάλο. Έστω ότι δεν ισχύει. Τότε υπάρχει ακολουθία $\{z_k\}$ που συγκλίνει στο x^* για την οποία ισχύει:

$$f(z_k) < f(x^*) + \frac{\sigma}{4}\|z_k - x^*\|^2, \quad \text{για κάθε} \;\; k \;\; \text{αρκετά μεγάλο}. \tag{2.47}$$

Τότε, μπορούμε να βρούμε μία περιοριστική διεύθυνση d τέτοια ώστε:

$$\lim_{k \to \infty} \frac{z_k - x^*}{\|z_k - x^*\|} = d. \tag{2.48}$$

Από το Λήμμα 2.1 (i) και τον ορισμό του συνόλου των εφικτών διευθύνσεων $\mathcal{F}(x^*)$ έχουμε ότι $d \in \mathcal{F}(x^*)$. Από τον ορισμό της συνάρτησης Lagrange και το γεγονός ότι $\lambda_i^* \geq 0$, $c_i(z_k) \geq 0, \forall i \in \mathcal{I}$ και $c_i(z_k) = 0, \forall i \in \mathcal{E}$, έχουμε ότι:

$$\mathcal{L}(z_k, \lambda^*) = f(z_k) - \sum_{i \in \mathcal{A}(x^*)} \lambda_i^* c_i(z_k) \leq f(z_k), \tag{2.49}$$

καθώς η προσέγγιση της συνάρτησης Lagrange, μέσω του τύπου Taylor, από την απόδειξη του Θεωρήματος 2.3:

$$\mathcal{L}(z_k, \lambda^*) = f(x^*) + \frac{1}{2} t_k^2 w^T \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*) + o(t_k^2)$$

συνεχίζει να ισχύει.

Εάν $d \notin \mathcal{C}(x^*, \lambda^*)$, θα υπήρχε ένας δείκτης $j \in \mathcal{A}(x^*) \cap \mathcal{I}$ τέτοιος ώστε η αυστηρή ανισότητα:

$$\lambda_j^* \nabla c_j(x^*)^T d > 0 \tag{2.50}$$

να ικανοποιείται, καθώς για τους υπόλοιπους δείκτες $i \in \mathcal{A}(x^*)$ να ισχύει:

$$\lambda_i^* \nabla c_i(x^*)^T d \geq 0.$$

Από το Θεώρημα Taylor και την (2.48) για αυτό τον δείκτη j έχουμε ότι:

$$\begin{aligned}\lambda_j^* c_j(z_k) &= \lambda_j^* c_j(x^*) + \lambda_j^* \nabla c_j(x^*)^T (z_k - x^*) + o(\|z_k - x^*\|) \\ &= \|z_k - x^*\| \lambda_j^* \nabla c_j(x^*)^T d + o(\|z_k - x^*\|).\end{aligned}$$

Επομένως, από την (2.49) έχουμε ότι:

$$\begin{aligned}\mathcal{L}(z_k, \lambda^*) &= f(z_k) - \sum_{i \in \mathcal{A}(x^*)} \lambda_i^* c_i(z_k) \\ &\leq f(z_k) - \lambda_j^* c_j(z_k) \\ &\leq f(z_k) - \|z_k - x^*\| \lambda_j^* \nabla c_j(x^*)^T d + o(\|z_k - x^*\|).\end{aligned}$$

Από την προσέγγιση της συνάρτησης Lagrange μέσω του τύπου Taylor έχουμε ότι:

$$\mathcal{L}(z_k, \lambda^*) = f(x^*) + O(\|z_k - x^*\|^2).$$

Συνδυάζοντας τις δύο τελευταίες συνθήκες έχουμε:

$$f(z_k) \geq f(x^*) + \|z_k - x^*\| \lambda_j^* \nabla c_j(x^*)^T d + o(\|z_k - x^*\|).$$

Από την (2.50), η παραπάνω ανισότητα με την (2.47) δεν συμφωνούν. Επομένως, καταλήγουμε ότι $d \in \mathcal{C}(x^*, \lambda^*)$ και έτσι $d^T \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*) d \geq \sigma$.

Συνδυάζοντας την προσέγγιση της συνάρτησης Lagrange μέσω του τύπου Taylor με την (2.49) και την (2.48) έχουμε:

$$\begin{aligned}f(z_k) &\geq f(x^*) + \tfrac{1}{2}(z_k - x^*)^T \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*)(z_k - x^*) + o(\|z_k - x^*\|^2) \\ &= f(x^*) + \tfrac{1}{2} d^T \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*) d \|z_k - x^*\|^2 + o(\|z_k - x^*\|^2) \\ &\geq f(x^*) + \tfrac{\sigma}{2} \|z_k - x^*\|^2 + o(\|z_k - x^*\|^2).\end{aligned}$$

Το οποίο είναι άτοπο λόγω της (2.47). Επομένως, κάθε εφικτή ακολουθία $\{z_k\}$ που συγκλίνει στο x^* πρέπει να ικανοποιεί την $f(z_k) \geq f(x^*) + \frac{\sigma}{4}\|z_k - x^*\|^2$, για κάθε k αρκετά μεγάλο, έτσι ώστε το x^* να είναι αυστηρά τοπική λύση. □

Παράδειγμα 2.7.2. *Έστω το πρόβλημα (2.11) του Παραδείγματος 2.3.2:*

$$\min x_1 + x_2, \quad \textit{υπό τον περιορισμό} \;\; 2 - x_1^2 - x_2^2 \geq 0,$$

δηλαδή $f(x) = x_1 + x_2, c_1(x) = 2 - x_1^2 - x_2^2, \mathcal{E} = \emptyset, \mathcal{I} = \{1\}$.

Η συνάρτηση Lagrange του προβλήματος είναι:

$$\mathcal{L}(x, \lambda) = (x_1 + x_2) - \lambda_1(2 - x_1^2 - x_2^2).$$

Είναι εύκολο να δείξουμε ότι οι συνθήκες KKT ικανοποιούνται για $\lambda_1^* = \frac{1}{2}$ *και* $x^* = (-1, -1)^T$.

Ο Εσσιανός πίνακας της συνάρτησης Lagrange σε αυτό το σημείο είναι:

$$\nabla_{xx}\mathcal{L}(x^*, \lambda^*) = \begin{bmatrix} 2\lambda_1^* & 0 \\ 0 & 2\lambda_1^* \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Ο παραπάνω πίνακας είναι θετικά ορισμένος, επομένως, ικανοποιεί τις συνθήκες του Θεωρήματος 2.4, δηλαδή τις Ικανές Συνθήκες Δεύτερης Τάξης. Έτσι, το σημείο $x^* = (-1, -1)^T$ *είναι αυστηρά τοπική λύση του προβλήματος (2.11).*

Παράδειγμα 2.7.3. *Έστω το πρόβλημα:*

$$\min -0.1(x_1 - 4)^2 + x_2^2, \quad \text{υπό τον περιορισμό } \; x_1^2 + x_2^2 - 1 \geq 0, \tag{2.51}$$

όπου το εφικτό σύνολο είναι το εξωτερικό του μοναδιαίου κύκλου. Η αντικειμενική συνάρτηση του προβλήματος δεν είναι κυρτή και δεν είναι κάτω φραγμένη μέσα στην εφικτή περιοχή, καθώς παίρνοντας την εφικτή ακολουθία:

$$\begin{bmatrix} 10 \\ 0 \end{bmatrix}, \begin{bmatrix} 20 \\ 0 \end{bmatrix}, \begin{bmatrix} 30 \\ 0 \end{bmatrix}, \begin{bmatrix} 40 \\ 0 \end{bmatrix},$$

η $f(x)$ *συγκλίνει στο* $-\infty$. *Επομένως, το πρόβλημα δεν έχει ολική λύση, αν και μπορούμε να βρούμε μία αυστηρά τοπική λύση στο σύνορο της εφικτής περιοχής. Θα βρούμε αυτή την λύση μέσω των συνθηκών KKT και των Ικανών Συνθηκών Δεύτερης Τάξης του Θεωρήματος 2.4.*

Η συνάρτηση Lagrange του προβλήματος (2.51) είναι:

$$\mathcal{L}(x, \lambda) = (-0.1(x_1 - 4)^2 + x_2^2) - \lambda_1(x_1^2 + x_2^2 - 1).$$

Τότε έχουμε:

$$\nabla_x\mathcal{L}(x, \lambda) = \begin{bmatrix} -0.2(x_1 - 4) - 2\lambda_1 x_1 \\ 2x_2 - 2\lambda_1 x_2 \end{bmatrix},$$

$$\nabla^2_{xx}\mathcal{L}(x, \lambda) = \begin{bmatrix} -0.2 - 2\lambda_1 & 0 \\ 0 & 2 - 2\lambda_1 \end{bmatrix}. \tag{2.52}$$

Το σημείο $x^* = (1, 0)^T$ *ικανοποιεί τις συνθήκες KKT με* $\lambda_1^* = 0.3$ *και ενεργό σύνολο το* $\mathcal{A}(x^*) = \{1\}$. *Για να δουμε εάν οι Ικανές Συνθήκες Δεύτερης Τάξης ικανοποιούνται σε αυτό το σημείο έχουμε:*

$$\nabla c_1(x^*) = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

Επομένως:

$$w \in \mathcal{C}(x^*, \lambda^*) \Leftrightarrow \nabla c_1(x^*)^T w = 0 \implies \begin{bmatrix} 2 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = 0 \implies \begin{cases} w_1 = 0 \\ w_2 \in \mathbb{R} \end{cases}.$$

Άρα ο κρίσιμος κώνος είναι:

$$\mathcal{C}(x^*, \lambda^*) = \{(0, w_2)^T \mid w_2 \in \mathbb{R}\}.$$

Έτσι, από την (2.52) για κάθε $w \in \mathcal{C}(x^*, \lambda^*)$ με $w \neq 0$ και $w_2 \neq 0$ έχουμε:

$$w^T \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*) w = \begin{bmatrix} 0 & w_2 \end{bmatrix} \begin{bmatrix} -0.4 & 0 \\ 0 & 1.4 \end{bmatrix} \begin{bmatrix} 0 \\ w_2 \end{bmatrix} = 1.4 w_2^2 > 0.$$

Επομένως, οι Ικανές Συνθήκες Δεύτερης Τάξης ικανοποιούνται και έτσι το σημείο $x^* = (1, 0)^T$ είναι αυστηρά τοπική λύση του προβλήματος (2.51).

2.8 Άλλες προϋποθέσεις περιορισμών

Σε αυτή την ενότητα θα δούμε κάποιες ακόμα προϋποθέσεις που πρέπει να ικανοποιούν οι περιορισμοί εκτός από αυτές που είδαμε στις ενότητες 2.4 και 2.5, που ήταν η προϋπόθεση LICQ και η συνθήκη αυστηρής συμπληρωματικότητας. Οι συνθηκες αυτές εξασφάλιζαν ότι η γραμμική προσέγγιση του εφικτού συνόλου $\mathcal{U}$ αποτυπώνει το σχήμα της εφικτής περιοχής κόντα στο x^*.

Μία περίπτωση κατά την οποία το σύνολο των γραμμικοποιημένων εφικτών διευθύνσεων $\mathcal{F}(x^*)$ είναι μία επαρκής αναπαράσταση του εφικτού συνόλου, είναι όταν όλοι οι ενεργοί περιορισμοί είναι γραμμικές συναρτήσεις. Δηλαδή:

$$c_i(x) = a_i^T x + b_i, \quad \text{για κάποιο} \;\; a_i \in \mathbb{R}^n \;\; \text{και} \;\; b_i \in \mathbb{R}. \tag{2.53}$$

Λήμμα 2.3 (βλ. [9], Κεφάλαιο 12, Λήμμα 12.7, σελ. 338)**.** *Έστω εφικτό σημείο $x^* \in \mathcal{U}$ και ότι όλοι οι ενεργοί περιορισμοί $c_i(x)$, με $i \in \mathcal{A}(x^*)$, είναι γραμμικές συναρτήσεις. Τότε $\mathcal{F}(x^*) = T_{\mathcal{U}}(x^*)$.*

Απόδειξη. Από το Λήμμα 2.1 (i) έχουμε ότι $T_{\mathcal{U}}(x^*) \subset \mathcal{F}(x^*)$. Για να δείξουμε ότι $\mathcal{F}(x^*) \subset T_{\mathcal{U}}(x^*)$, επιλέγουμε αυθαίρετο $w \in \mathcal{F}(x^*)$ και θα δείξουμε ότι $w \in T_{\mathcal{U}}(x^*)$.

Από τον ορισμό του συνόλου των εφικτών διευθύνσεων $\mathcal{F}(x^*)$ και την μορφή των περιορισμών (2.53) έχουμε:

$$\mathcal{F}(x^*) = \left\{ d \mid \begin{array}{l} a_i^T d = 0, \forall i \in \mathcal{E} \\ a_i^T d \geq 0, \forall i \in \mathcal{A}(x) \cap \mathcal{I} \end{array} \right\}.$$

Υπάρχει $\bar{t} \in \mathbb{R}, \bar{t} > 0$, τέτοιο ώστε οι ανενεργοί περιορισμοί να παραμένουν ανενεργοί στο σημείο $x^* + tw$, για κάθε $t \in [0, \bar{t}]$. Δηλαδή:

$$c_i(x^* + tw) > 0, \; \forall i \in \mathcal{I} \backslash \mathcal{A}(x^*) \;\; \text{και} \;\; t \in [0, \bar{t}].$$

Ορίζουμε ακολουθία $\{z_k\}$ ως εξής:

$$z_k = x^* + \frac{\bar{t}}{k} w, \; \forall k = 1, 2, \ldots.$$

Αφού $a_i^T w \geq 0$ για κάθε $i \in \mathcal{I} \cap \mathcal{A}(x^*)$ έχουμε:

$$c_i(z_k) = c_i(z_k) - c_i(x^*) = a_i^T (z_k - x^*) = \frac{\bar{t}}{k} a_i^T w \geq 0, \; \forall i \in \mathcal{I} \cap \mathcal{A}(x^*),$$

έτσι ώστε η $\{z_k\}$ να είναι εφικτή ως προς τους ενεργούς περιορισμούς c_i, με $i \in \mathcal{I} \cap \mathcal{A}(x^*)$. Με αυτό το $\bar{t}$, η $\{z_k\}$ είναι επίσης εφικτή ως προς τους ανενεργούς περιορισμούς c_i, με $i \in \mathcal{I} \backslash \mathcal{A}(x^*)$ και είναι εύκολο να δείξουμε ότι $c_i(z_k) = 0$ για τους περιορισμούς c_i, με $i \in \mathcal{E}$. Επομένως, η $\{z_k\}$ είναι εφικτή για κάθε $k = 1, 2 \ldots$. Επίσης έχουμε:

$$\frac{z_k - x^*}{\frac{\bar{t}}{k}}) = \frac{\frac{\bar{t}}{k} w}{\frac{\bar{t}}{k}} = w,$$

έτσι ώστε το w να είναι περιοριστική διεύθυνση της $\{z_k\}$. Επομένως, $w \in T_{\mathcal{U}}(x^*)$ και άρα $\mathcal{F}(x^*) = T_{\mathcal{U}}(x^*)$. □

Παρατηρήσεις

1. Το Λήμμα 2.3 είναι μία εκδοχή του Λήμματος 2.1, για την περίπτωση όπου οι περιορισμοί είναι γραμμικές συναρτήσεις.

2. Από το παραπάνω, συμπεραίνουμε ότι η περίπτωση όπου οι περιορισμοί είναι γραμμικές συναρτήσεις, είναι μία πιθανή προϋπόθεση που θα μπορούσαν να ικανοποιούν οι περιορισμοί. Η προϋπόθεση αυτή, δεν είναι πιο ασθενής από την προϋπόθεση LICQ, ωστόσο υπάρχουν περιπτώσεις όπου η μία ικανοποιείται ενώ η άλλη όχι.

Ορισμός 2.18 (MFCQ, βλ. [9], Κεφάλαιο 12, Ορισμός 12.6, σελ. 339)**.** *Λέμε ότι η προϋπόθεση για τους περιορισμούς Mangasarian-Fromovitz (Mangasarian - Fromovitz Constraint Qualification (MFCQ)) ικανοποιείται εάν υπάρχει διάνυσμα $w \in \mathbb{R}^n$ τέτοιο ώστε:*

$$\nabla c_i(x^*)^T w = 0, \forall i \in \mathcal{E},$$
$$\nabla c_i(x^*)^T w > 0, \forall i \in \mathcal{A}(x^*) \cap \mathcal{I},$$

και το σύνολο των κλίσεων των ισοτικών περιορισμών $\{\nabla c_i(x^*), i \in \mathcal{E}\}$ είναι γραμμικά ανεξάρτητο.

Παρατηρήσεις

1. Η παραπάνω προϋπόθεση είναι μία γενίκευση της προϋπόθεσης LICQ.

2. Η προϋπόθεση MFCQ είναι πιο ασθενής από την προϋπόθεση LICQ. Εάν η LICQ ικανοποιείται, τότε το σύστημα:

$$\nabla c_i(x^*)^T w = 0, \forall i \in \mathcal{E},$$
$$\nabla c_i(x^*)^T w = 1, \forall i \in \mathcal{A}(x^*) \cap \mathcal{I},$$

   έχει λύση w και οι κλίσεις των ενεργών περιορισμών είναι πίνακες πλήρους τάξης. Επομένως, μπορούμε να επιλέξουμε το w του Ορισμού 2.18 να είναι αυτή η λύση. Από την άλλη, είναι εύκολο να κατασκευάσουμε παραδείγματα στα οποία να ικανοποιείται η προϋπόθεση MFCQ και όχι η LICQ.

3. Μία μορφή των Αναγκαίων Συνθηκών Πρώτης Τάξης, από το Θεώρημα 2.1, στην οποία η προϋπόθεση LICQ της υπόθεσης αντικαθίσταται από την MFCQ, μπορεί να αποδειχθεί. Η προϋπόθεση MFCQ δίνει μια ιδιότητα, η οποία είναι ισοδύναμη με το ότι το σύνολο των διανυσμάτων των πολλαπλασιαστών Lagrange λ^*, που ικανοποιούν τις συνθήκες KKT, είναι φραγμένο. Στην περίπτωση όπου η LICQ ικανοποιείται, το σύνολο αυτό είναι εξ' ορισμού φραγμένο, καθώς το διάνυσμα λ^* που ικανοποιεί τις KKT είναι μοναδικό.

4. Οι προϋποθέσεις των περιορισμών είναι ικανές συνθήκες προκειμένου η γραμμική προσέγγιση να είναι επαρκής, και όχι αναγκαίες. Για παράδειγμα, εάν έχουμε τους περιορισμούς $x_2 \geq -x_1^2$ και $x_2 \leq x_1^2$ και το εφικτό σημείο $x^* = (0,0)^T$, καμία προϋπόθεση δεν ικανοποιείται. Ωστόσο, η γραμμική προσέγγιση $\mathcal{F}(x^*) = \{(w_1, 0)^T | w_1 \in \mathbb{R}\}$ απεικονίζει την γεωμετρία του εφικτού συνόλου κοντά στο x^*.

2.9 Δυϊκότητα

Σε αυτή την ενότητα παρουσιάζονται κάποια κομμάτια της δυϊκης θεωρίας για τον μη γραμμικό προγραμματισμό. Η θεωρία αυτή έχει χρησιμοποιηθεί για την κατασκευή κάποιων πολύ σημαντικών αλγορίθμων βελτιστοποίησης και εφαρμόζεται σε διάφορα είδη βελτιστοποίησης όπως στην γραμμική (linear otimization), στην συνδυαστική (discrete otimization), στην οποία οι μεταβλητές μπορούν να πάρουν μόνο διακριτές τιμές, και στην κυρτή μη λεία (convex nonsmooth otimization).

Η δυϊκή θεωρία μας δείχνει πως κατασκευάζουμε ένα δεύτερο πρόβλημα βελτιστοποίησης, ισοδύναμο με το πρώτο, από τις συναρτήσεις και τα δεδομένα του πρώτου προβλήματος. Το δεύτερο αυτό πρόβλημα ονομάζεται *δυϊκό πρόβλημα* και το αρχικό ονομάζεται *πρωτεύον πρόβλημα*. Μέσω του δυϊκού προβλήματος, έχουν κατασκευαστεί αλγόριθμοι για την επίλυση του πρωτεύοντος και σε μερικές περιπτώσεις, το δυϊκό πρόβλημα μπορεί να είναι πιο εύκολα επιλύσιμο από το πρωτεύον ή να δίνει χαμηλότερα φράγματα για την τιμή της αντικειμενικής συνάρτησης για το πρωτεύον πρόβλημα.

Τα κομμάτια που παρουσιάζονται σε αυτή την ενότητα επικεντρώνονται μόνο στην ειδική περίπτωση του γενικού προβλήματος (2.1), όπου δεν υπάρχουν ισοτικοί περιορισμοί, και η αντικειμενική συνάρτηση f μαζί με τους ανισοτικούς περιορισμούς με αρνητικό πρόσημο $-c_i$ είναι κυρτές συναρτήσεις. Έστω ότι υπάρχουν m στο πλήθος ανισοτικοί περιορισμοί. Θεωρούμε το *πρωτεύον πρόβλημα* ως το γενικό πρόβλημα (2.1) που γράφεται ως εξής:

$$\min_{x \in \mathbb{R}^n} f(x), \quad \text{υπό τον περιορισμό} \quad c(x) \geq 0, \tag{2.54}$$

όπου $c(x) = (c_1(x), c_2(x), \ldots, c_m(x))^T$ και $c_i(x) \geq 0, i = 1, 2, \ldots, m$ οι ανισοτικοί περιορισμοί. Η συνάρτηση Lagrange του πρωτεύοντος προβλήματος (2.54) είναι:

$$\mathcal{L}(x, \lambda) = f(x) - \lambda^T c(x),$$

όπου $\lambda \in \mathbb{R}^m$ το διάνυσμα πολλαπλασιαστών Lagrange.

Ορισμός 2.19 (βλ. [9], Κεφάλαιο 12, Ορισμός, σελ. 344)**.** *Ορίζουμε την δυϊκή αντικειμενική συνάρτηση, δηλαδή την αντικειμενική συνάρτηση του δυϊκού προβλήματος, να είναι $q : \mathbb{R}^n \longrightarrow \mathbb{R}$:*

$$q(\lambda) = \inf_x \mathcal{L}(x, \lambda). \tag{2.55}$$

Παρατηρήσεις

1. Σε πολλά προβλήματα, για κάποιες τιμές του λ, το μέγιστο κάτω φράγμα της συνάρτησης Lagrange είναι $-\infty$.

2. Ο υπολογισμός του μέγιστου κάτω φράγματος στην (2.55) απαιτεί την εύρεση του ολικού ελαχίστου της συνάρτησης $\mathcal{L}(\cdot, \lambda)$ για το δοσμένο λ, το οποίο μπορεί να είναι πολύ δύσκολο. Ωστόσο, όταν οι συναρτήσεις f και $-c_i$ είναι κυρτές και $\lambda \geq 0$, η συνάρτηση $\mathcal{L}(\cdot, \lambda)$ είναι και αυτή κυρτή. Σε αυτή την περίπτωση, έχουμε δείξει ότι κάθε τοπικό ελάχιστο είναι και ολικό ελάχιστο, οπότε ο υπολογισμός του $q(\lambda)$ είναι πιο εύκολος.

Ορισμός 2.20. *Ορίζουμε το πεδίο ορισμού της αντικειμενικής συνάρτησης του δυϊκού προβλήματος να είναι οι τιμές του λ για τις οποίες η αντικειμενική συνάρτηση είναι πεπερασμένη, δηλαδή το σύνολο:*

$$\mathcal{D} = \{\lambda \mid q(\lambda) > -\infty\}. \tag{2.56}$$

Ορισμός 2.21 (Δυϊκό Πρόβλημα)**.** *Το δυϊκό πρόβλημα του πρωτεύοντος προβλήματος (2.54) είναι το εξής:*

$$\max_{\lambda \in \mathbb{R}^n} q(\lambda), \quad \textit{υπό τον περιορισμό} \;\; \lambda \geq 0. \tag{2.57}$$

Παράδειγμα 2.9.1. *Έστω το πρόβλημα:*

$$\min_{(x_1,x_2)} 0.5(x_1^2 + x_2^2), \quad \textit{υπό τον περιορισμό} \;\; x_1 - 1 \geq 0. \tag{2.58}$$

Η συνάρτηση Lagrange του προβλήματος είναι:

$$\mathcal{L}(x_1, x_2, \lambda_1) = 0.5(x_1^2 + x_2^2) - \lambda_1(x_1 - 1).$$

Εάν κρατήσουμε σταθερό το λ_1 η συνάρτηση Lagrange είναι μία κυρτή συνάρτηση των $(x_1, x_2)^T$. Επομένως, το μέγιστο κάτω φράγμα της επιτυγχάνεται όταν οι μερικές παράγωγοι ως προς x_1 και x_2 είναι μηδέν. Δηλαδή όταν:

$$\nabla \mathcal{L}(x_1, x_2, \lambda_1) = \begin{bmatrix} x_1 - \lambda_1 \\ x_2 \end{bmatrix} = 0 \implies \begin{cases} x_1 - \lambda_1 & = 0 \\ x_2 & = 0 \end{cases}.$$

Αντικαθιστώντας τις τιμές των x_1 και x_2 στην συνάρτηση Lagrange έχουμε ότι η δυϊκή αντικειμενική συνάρτηση του προβλήματος είναι:

$$q(\lambda_1) = 0.5(\lambda_1^2 + 0) - \lambda_1(\lambda_1 - 1) = -0.5\lambda_1^2 + \lambda_1.$$

Επομένως, το δυϊκό πρόβλημα του (2.58) είναι:

$$\max_{\lambda_1 \geq 0} \; -0.5\lambda_1^2 + \lambda_1,$$

του οποίου προφανώς η λύση είναι $\lambda_1 = 1$.

Παρατηρήσεις

1. Στο παράδειγμα αυτό παρατηρούμε ότι το $\lambda_1 = 1$ είναι βέλτιστος πολλαπλασιαστής Lagrange για το πρωτεύον πρόβλημα (2.58) αλλά και λύση για το δυϊκό του. Επίσης, η βέλτιστη τιμή και των δύο προβλημάτων είναι 0.5.

2. Στην δυϊκή λύση $\lambda_1 = 1$, το μέγιστο κάτω φράγμα $\inf \mathcal{L}(x_1, x_2, \lambda_1)$ επιτυγχάνεται στο σημείο $(x_1, x_2) = (1, 0)^T$, το οποίο είναι η λύση του πρωτεύοντος προβλήματος (2.58).

Θεώρημα 2.5 (βλ. [9], Κεφάλαιο 12, Θεώρημα 12.10, σελ. 345). *Η δυϊκή αντικειμενική συνάρτηση q (2.55) είναι κοίλη και το πεδίο ορισμού της $\mathcal{D}$ (2.56) είναι κυρτό σύνολο.*

Απόδειξη. Έστω $\lambda^0, \lambda^1 \in \mathbb{R}^m$, $x \in \mathbb{R}^n$ και $a \in [0,1]$. Τότε:

$$\mathcal{L}(x, (1-a)\lambda^0 + a\lambda^1) = (1-a)\mathcal{L}(x, \lambda^0) + a\mathcal{L}(x, \lambda^1).$$

Παίρνοντας το infimum και στα δύο μέλη της ισότητας έχουμε ότι:

$$\begin{aligned} \inf\ \mathcal{L}(x, (1-a)\lambda^0 + a\lambda^1) &= \inf[(1-a)\mathcal{L}(x, \lambda^0) + a\mathcal{L}(x, \lambda^1)] \\ &\geq (1-a)\inf\ \mathcal{L}(x, \lambda^0) + a\inf\ \mathcal{L}(x, \lambda^1). \end{aligned}$$

Από τον ορισμό της δυϊκής αντικειμενικής συνάρτησης έχουμε:

$$q((1-a)\lambda^0 + a\lambda^1) \geq (1-a)q(\lambda^0) + aq(\lambda^1).$$

Επομένως, η q είναι κοίλη.

Εάν $\lambda^0, \lambda^1 \in \mathcal{D}$ τότε από την παραπάνω και τον ορισμό του συνόλου $\mathcal{D}$ έχουμε ότι:

$$q((1-a)\lambda^0 + a\lambda^1) \geq -\infty \implies (1-a)\lambda^0 + a\lambda^1 \in \mathcal{D}.$$

Επομένως, το $\mathcal{D}$ είναι κυρτό σύνολο. □

Θεώρημα 2.6 (Ασθενής Δυϊκότητα, βλ. [9], Κεφάλαιο 12, Θεώρημα 12.11, σελ. 345). *Για κάθε εφικτό σημείο $\bar{x}$ του πρωτεύοντος προβλήματος (2.54) και για κάθε $\bar{\lambda} \geq 0$, έχουμε ότι:*

$$q(\bar{\lambda}) \leq f(\bar{x}).$$

Απόδειξη. Αφού $\bar{\lambda} \geq 0$ και $c(\bar{x}) \geq 0$ έχουμε:

$$q(\bar{\lambda}) = \inf_x f(x) - \bar{\lambda}^T c(x) \leq f(\bar{x}) - \bar{\lambda}^T c(\bar{x}) \leq f(\bar{x}).$$

□

Παρατήρηση

Από το Θεώρημα 2.6 παρατηρούμε ότι η βέλτιστη τιμή του δυϊκού προβλήματος (2.57) δίνει ένα χαμηλότερο φράγμα για την βέλτιστη τιμή της αντικειμενικής συνάρτησης του πρωτεύοντος προβλήματος (2.54).

Θεώρημα 2.7 (Αναγκαίες Συνθήκες Πρώτης Τάξης). *Οι συνθήκες KKT για το πρωτεύον πρόβλημα (2.54) είναι:*

$$\begin{aligned} \nabla_x \mathcal{L}(\bar{x}, \bar{\lambda}) = \nabla f(\bar{x}) - \nabla c(\bar{x})\bar{\lambda} &= 0, && (2.59) \\ c(\bar{x}) &\geq 0, && (2.60) \\ \bar{\lambda} &\geq 0, && (2.61) \\ \bar{\lambda}_i c_i(\bar{x}) &= 0, \forall i = 1, 2, \ldots, m, && (2.62) \end{aligned}$$

όπου $\nabla c(x) = [\nabla c_1(x), \nabla c_2(x), \ldots, \nabla c_m(x)]$ είναι ένας $n \times m$ πίνακας.

Θεώρημα 2.8 (βλ. [9], Κεφάλαιο 12, Θεώρημα 12.12, σελ. 346). *Έστω $\bar{x}$ λύση του πρωτεύοντος προβλήματος (2.54) και η f μαζί με τους περιορισμούς $-c_i, i = 1, 2, \ldots, m$ είναι κυρτές συναρτήσεις στο $\mathbb{R}^n$ και παραγωγίσιμες στο $\bar{x}$. Τότε κάθε $\bar{\lambda}$ για το οποίο το ζεύγος $(\bar{x}, \bar{\lambda})$ ικανοποιεί τις συνθήκες KKT (2.59) έως (2.62), είναι λύση του δυϊκού προβλήματος (2.57).*

Απόδειξη. Έστω ζεύγος $(\bar{x}, \bar{\lambda})$ που ικανοποιεί τις συνθήκες KKT (2.59) έως (2.62). Αφού $\bar{\lambda} \geq 0$, η συνάρτηση $\mathcal{L}(\cdot, \bar{\lambda})$ είναι κυρτή και παραγωγίσιμη. Επομένως, από την πρώτη συνθήκη KKT (2.59), για κάθε x έχουμε:

$$\mathcal{L}(x, \bar{\lambda}) \geq \mathcal{L}(\bar{x}, \bar{\lambda}) + \nabla_x \mathcal{L}(\bar{x}, \bar{\lambda})^T (x - \bar{x}) = \mathcal{L}(\bar{x}, \bar{\lambda}).$$

Επομενως, από την συνθήκη KKT (2.62), η δυϊκή αντικειμενική συνάρτηση είναι:

$$q(\bar{\lambda}) = \inf_x \mathcal{L}(x, \bar{\lambda}) = \mathcal{L}(\bar{x}, \bar{\lambda}) = f(\bar{x}) - \bar{\lambda}^T c(\bar{x}) = f(\bar{x}).$$

Όμως, από το Θεώρημα 2.6 έχουμε ότι $q(\lambda) \leq f(\bar{x})$ για κάθε $\lambda \geq 0$. Επομένως, αφού $q(\bar{\lambda}) = f(\bar{x})$ το $\bar{\lambda}$ θα είναι λύση του δυϊκού προβλήματος (2.57). □

Παρατήρηση

Εάν οι συναρτήσεις είναι συνεχώς παραγωγίσιμες και μία προϋπόθεση όπως η LICQ ικανοποιείται στο $\bar{x}$, τότε από το Θεώρημα 2.1 των Αναγκαίων Συνθηκών Πρωτης Τάξης είναι σίγουρη η ύπαρξη ενός βέλτιστου πολλαπλασιαστή Lagrange.

Θεώρημα 2.9 (βλ. [9], Κεφάλαιο 12, Θεώρημα 12.13, σελ. 347). *Έστω f και $-c_i, i = 1, 2, \ldots, m$ κυρτές και συνεχώς παραγωγίσμες συναρτήσεις στο $\mathbb{R}^n$. Έστω $\bar{x}$ λύση του πρωτεύοντος προβλήματος (2.54), στο οποίο η προϋπόθεση LICQ ικανοποιείται. Έστω ότι το $\hat{\lambda}$ είναι λύση του δυϊκού προβλήματος (2.57) και ότι το μέγιστο κάτω φράγμα $\inf_x \mathcal{L}(x, \hat{\lambda})$ επιτυγχάνεται στο $\hat{x}$. Έστω επίσης ότι η $\mathcal{L}(\cdot, \hat{\lambda})$ είναι αυστηρά κυρτή συνάρτηση. Τότε $\bar{x} = \hat{x}$, δηλαδή το $\hat{x}$ είναι η μοναδική λύση του πρωτεύοντος προβλήματος (2.54), και $f(\bar{x}) = \mathcal{L}(\hat{x}, \hat{\lambda})$.*

Απόδειξη. Έστω ότι ισχύει $\bar{x} \neq \hat{x}$. Από το Θεώρημα 2.1 (Αναγκαίες Συνθήκες Πρώτης Τάξης), καθώς η προϋπόθεση LICQ ικανοποιείται, υπάρχει $\bar{\lambda}$ που να ικανοποιεί τις συνθήκες KKT (2.59) έως (2.62). Επομένως, από το Θεώρημα 2.8, το $\bar{\lambda}$ αυτό είναι και λύση του δυϊκού προβλήματος (2.57), έτσι ώστε:

$$\mathcal{L}(\bar{x}, \bar{\lambda}) = q(\bar{\lambda}) = q(\hat{\lambda}) = \mathcal{L}(\hat{x}, \hat{\lambda}).$$

Επειδή το $\hat{x}$ είναι σημείο τοπικού ελαχίστου της $\mathcal{L}(x, \hat{\lambda})$, από το Θεώρημα 1.7 (Αναγκαία Συνθήκη Πρώτης Τάξης) και την Παρατήρηση 2. (συνθήκη (1.4)) γνωρίζουμε ότι $\nabla_x \mathcal{L}(\hat{x}, \hat{\lambda}) = 0$. Επίσης, αφού η $\mathcal{L}(\cdot, \hat{\lambda})$ είναι αυστηρά κυρτή συνάρτηση έχουμε ότι:

$$\mathcal{L}(\bar{x}, \hat{\lambda}) - \mathcal{L}(\hat{x}, \hat{\lambda}) > \nabla_x \mathcal{L}(\hat{x}, \hat{\lambda})^T (\bar{x} - \hat{x}) = 0.$$

Επομένως, από τα προηγούμενα και την συνθήκη KKT (2.62) έχουμε:

$$\begin{aligned}
\mathcal{L}(\bar{x}, \hat{\lambda}) > \quad & \mathcal{L}(\hat{x}, \hat{\lambda}) & = \mathcal{L}(\bar{x}, \bar{\lambda}) \\
\implies f(\bar{x}) - \hat{\lambda}^T c(\bar{x}) > \quad & f(\hat{x}) - \hat{\lambda}^T c(\hat{x}) & = f(\bar{x}) - \bar{\lambda}^T c(\bar{x}) \\
\implies -\hat{\lambda}^T c(\bar{x}) > \quad & -\bar{\lambda}^T c(\bar{x}) & = 0.
\end{aligned}$$

Όμως $\hat{\lambda} \geq 0$ ως λύση του δυϊκού προβλήματος και $c(\bar{x}) \geq 0$ από την συνθήκη KKT (2.60), άρα καταλήγουμε σε άτοπο και επομένως $\bar{x} = \hat{x}$. □

Παρατήρηση

Η συνθήκη του Θεωρήματος 2.9 ότι η $\mathcal{L}(\cdot, \hat{\lambda})$ πρέπει να είναι αυστηρά κυρτή, ισχύει είτε αν η συνάρτηση f είναι αυστηρά κυρτή, είτε αν οι περιορισμοί c_i είναι αυστηρά κυρτοί για κάποιο $i = 1, 2, \ldots, m$ με $\hat{\lambda}_i > 0$.

Ορισμός 2.22 (Wolfe Δυϊκό Πρόβλημα, βλ. [9], Κεφάλαιο 12, Ορισμός, σελ. 347). *Το Wolfe δυϊκό πρόβλημα του πρωτεύοντος προβλήματος (2.54) είναι το εξής:*

$$\max_{x,\lambda} \mathcal{L}(x, \lambda), \quad \textit{υπό τους περιορισμούς} \;\; \nabla_x \mathcal{L}(x, \lambda) = 0, \lambda \geq 0. \tag{2.63}$$

Θεώρημα 2.10 (βλ. [9], Κεφάλαιο 12, Θεώρημα 12.14, σελ. 347). *Έστω f και $-c_i, i = 1, 2, \ldots, m$ κυρτές και συνεχώς παραγωγίσιμες συναρτήσεις στο $\mathbb{R}^n$. Έστω, επίσης, το ζεύγος $(\bar{x}, \bar{\lambda})$ λύση του πρωτεύοντος προβλήματος (2.54) στο οποίο η προϋπόθεση LICQ ικανοποιείται. Τότε το $(\bar{x}, \bar{\lambda})$ είναι λύση του Wolfe δυϊκού προβλήματος (2.63).*

Απόδειξη. Από τις συνθήκες KKT (2.59) έως (2.62), έχουμε ότι το $(\bar{x}, \bar{\lambda})$ ικανοποιεί τους περιορισμούς του Wolfe δυϊκού προβλήματος και ότι $\mathcal{L}(\bar{x}, \bar{\lambda}) = f(\bar{x})$. Αφού $\mathcal{L}(\cdot, \lambda)$ κυρτή, τότε για κάθε ζεύγος (x, λ) που ικανοποιεί τους περιορισμούς του Wolfe δυϊκού προβλήματος, έχουμε ότι:

$$\begin{aligned}
\mathcal{L}(\bar{x}, \bar{\lambda}) &= f(\bar{x}) \\
&\geq f(\bar{x}) - \lambda^T c(\bar{x}) \\
&= \mathcal{L}(\bar{x}, \lambda) \\
&\geq \mathcal{L}(x, \lambda) + \nabla_x \mathcal{L}(x, \lambda)^T (\bar{x} - x) \\
&= \mathcal{L}(x, \lambda).
\end{aligned}$$

Επομένως, το $(\bar{x}, \bar{\lambda})$ μεγιστοποιεί την συνάρτηση $\mathcal{L}$ υπό τους περιορισμούς του Wolfe δυϊκού προβλήματος και άρα είναι λύση του (2.63). □

Παράδειγμα 2.9.2 (Γραμμικός Προγραμματισμός). *Μία ειδική περίπτωση του Wolfe δυϊκού προβλήματος (2.63) είναι το πρόβλημα γραμμικού προγραμματισμού:*

$$\min c^T x, \quad \textit{υπό τον περιορισμό} \;\; Ax - b \geq 0. \tag{2.64}$$

Η δυϊκή αντικειμενική συνάρτηση του (2.64) είναι:

$$q(\lambda) = \inf_x [c^T x - \lambda^T (Ax - b)] = \inf_x [(c - A^T \lambda)^T x + b^T \lambda].$$

Εάν $c - A^T \lambda \neq 0$, τότε το μέγιστο κάτω φράγμα είναι προφανώς $-\infty$ (θέτοντας το x ως ένα μεγάλο αρνητικό πολλαπλάσιο του $-(c - A^T \lambda)$, το q γίνεται αυθαίρετα μεγάλο και αρνητικό).

Εάν $c - A^T \lambda = 0$, τότε $q(\lambda) = \inf_x [b^T \lambda] = b^T \lambda$.

Έπομένως, για την μεγιστοποίηση της q, μπορούμε να απορρίψουμε τα λ για την περίπτωση όπου $c - A^T \lambda \neq 0$, καθώς είναι προφανές ότι το μέγιστο δεν μπορεί να *επιτευχθεί σε ένα σημείο* λ *για το οποίο* $q(\lambda) = -\infty$. *Επομένως, το δυϊκό πρόβλημα (2.57) μπορεί να γραφτεί ως:*

$$\max_{\lambda} b^T\lambda, \quad \text{υπό τους περιορισμούς } \; A^T\lambda = c, \lambda \geq 0. \tag{2.65}$$

Το Wolfe δυϊκό πρόβλημα του (2.64) είναι:

$$\max_{\lambda} c^T x - \lambda^T(Ax - b), \quad \text{υπό τους περιορισμούς } \; A^T\lambda = c, \lambda \geq 0.$$

Εάν στην δυϊκή αντικειμενική συνάρτηση του (2.64) αντικαταστήσουμε το $A^T\lambda - c = 0$, *τότε καταλήγουμε στο πρόβλημα (2.65) πάλι.*

Παρατήρηση

Για κάποιους πίνακες A, το δυϊκό πρόβλημα (2.65) μπορεί να είναι πιο εύκολα επιλύσιμο από ότι το πρωτεύον πρόβλημα (2.64).

Παράδειγμα 2.9.3 (Κυρτός Τετραγωνικός Προγραμματισμός)**.** *Έστω το πρόβλημα:*

$$\min \frac{1}{2}x^T Gx + c^T x, \quad \text{υπό τον περιορισμό } \; Ax - b \geq 0, \tag{2.66}$$

όπου G *ένας συμμετρικός θετικά ορισμένος πίνακας. Η δυϊκή αντικειμενική συνάρτηση του προβλήματος (2.66) είναι:*

$$q(\lambda) = \inf_x \mathcal{L}(x, \lambda) = \inf_x \frac{1}{2}x^T Gx + c^T x - \lambda^T(Ax - b).$$

Αφού ο G *είναι θετικά ορισμένος πίνακας και η* $\mathcal{L}(\cdot, \lambda)$ *είναι αυστηρά κυρτή τετραγωνική συνάρτηση, το μέγιστο κάτω φράγμα επιτυγχάνεται όταν* $\nabla_x \mathcal{L}(x, \lambda) = 0$, *δηλαδή όταν:*

$$Gx + c - A^T\lambda = 0. \tag{2.67}$$

Αντικαθιστώντας το x *στην δυϊκή αντικειμενική συνάρτηση έχουμε:*

$$q(\lambda) = -\frac{1}{2}(A^T\lambda - c)^T G^{-1}(A^T\lambda - c)^T + b^T\lambda.$$

Επομένως, το Wolfe δυϊκό πρόβλημα (2.63) με τον περιορισμό (2.67) γράφεται ως εξής:

$$\max_{(x,\lambda)} \frac{1}{2}x^T Gx + c^T x - \lambda^T(Ax - b), \tag{2.68}$$

$$\text{υπό τους περιορισμούς } \; Gx + c - A^T\lambda = 0, \lambda \geq 0.$$

Για να είναι πιο εμφανές ότι η αντικειμενική συνάρτηση είναι κοίλη, μπορούμε χρησιμοποιώντας τον περιορισμό (2.67) να αντικαταστήσουμε στην αντικειμενική το $(c - A^T\lambda)^T x = -x^T Gx$, *και έτσι έχουμε το δυϊκό πρόβλημα:*

$$\max_{(x,\lambda)} \; -\frac{1}{2}x^T Gx + \lambda^T b, \tag{2.69}$$

$$\text{υπό τους περιορισμούς } \; Gx + c - A^T\lambda = 0, \lambda \geq 0.$$

Παρατήρηση

Το Wolfe δυϊκό πρόβλημα προϋποθέτει μόνο ότι ο πίνακας G είναι θετικά ημιορισμένος.

Βιβλιογραφικές αναφορές

Το Κεφάλαιο 2 βασίστηκε στο Κεφάλαιο 12 του βιβλίου Numerical Optimization [βλ. [9], Κεφάλαιο 12, σελ. 304-349], καθώς επίσης και στο Λεξικό Μαθηματικών Όρων [βλ. [12]].

Ενότητα 2.2: Για την ενότητα αυτή, επιπλέον, αντλήθηκαν στοιχεία από το βιβλίο Ανάλυση ΙΙ [βλ. [4], Κεφάλαιο 3, σελ. 73].

Ενότητα 2.9: Για την ενότητα αυτή και συγκεκριμένα για το Wolfe δυϊκό πρόβλημα, αντλήθηκαν στοιχεία και από το άρθρο A Duality Theorem for Nonlinear Programming [βλ. [13]].

Κεφάλαιο 3

Εφαρμογές Βελτιστοποίησης σε Μαθηματικά Προβλήματα

Σε αυτό το κεφάλαιο θα αναλύσουμε τις διαφορές που υπάρχουν ανάμεσα σε διαφορετικές μεθόδους που χρησιμοποιούνται στην επίλυση του γενικού προβλήματος βελτιστοποίησης με περιορισμούς, όπως αυτό ορίστηκε στο Κεφάλαιο 2. Δηλαδή, για λείες συναρτήσεις $f, c_i : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}, S \subset \Omega$ και πεπερασμένα σύνολα $\mathcal{E}$ και $\mathcal{I}$ των δεικτών των ισοτικών και ανισοτικών περιορισμών αντίστοιχα, έχουμε το εξής πρόβλημα:

$$\text{Να βρεθεί } x^* \in \mathcal{U} \text{ τέτοιο ώστε } f(x^*) = \min_{x \in \mathcal{U}} f(x), \tag{3.1}$$
$$\text{όπου } \mathcal{U} = \{x \in S : c_i(x) = 0, i \in \mathcal{E}, c_i(x) \geq 0, i \in \mathcal{I}\}.$$

Στο Κεφάλαιο 2, αναλύσαμε την θεωρία των συνθηκών που χαρακτηρίζουν τις λύσεις του παραπάνω γενικού προβλήματος καθώς και τις προϋποθέσεις που πρέπει να ικανοποιούν οι περιορισμοί του. Σε αυτό το κεφάλαιο, θα δούμε μέσα από παραδείγματα πως αυτή η θεωρία υλοποιείται μέσα από συγκεκριμένους αλγορίθμους. Η βασική λειτουργία των αλγορίθμων είναι η παραγωγή μίας ακολουθίας προσεγγίσεων της λύσης x^*, μέσω της επανάληψης κάποιων βημάτων, η οποία ιδανικά συγκλίνει στην λύση. Ωστόσο, υπάρχουν και αλγόριθμοι οι οποίοι εκτός από την ακολουθία των προσεγγίσεων αυτών, παράγουν και μία ακολουθία των πολλαπλασιαστών Lagrange που σχετίζονται με τους περιορισμούς του δοσμένου προβλήματος. Οι αλγόριθμοι που θα εξετάσουμε σε αυτό το κεφάλαιο, βρίσκουν μόνο τοπικές λύσεις του προβλήματος (3.1).

3.1 Μέθοδος Ενεργού Συνόλου (Active Set)

Μία από τις βασικές δυσκολίες στην επίλυση των προβλημάτων μη γραμμικού προγραμματισμού είναι ο προσδιορισμός εκείνων των περιορισμών που είναι ενεργοί στην λύση του προβλήματος. Ένας τρόπος για να γίνει αυτό, που είναι και η βασική αρχή των μεθόδων του ενεργού συνόλου, είναι η επιλογή τυχαίων περιορισμών που υποθέτουμε ότι ικανοποιούνται ως ισότητες στην λύση. Αυτή η τεχνική βασίζεται στο γεγονός ότι ένα πρόβλημα μη γραμμικού προγραμματισμού με περιορισμούς ισότητες είναι πιο εύκολα επιλύσιμο από ότι ένα πρόβλημα με ανισότητες.

Ορισμός 3.1 (βλ. [9], Κεφάλαιο 15, Ορισμός, σελ. 424). *Η υπόθεση του βέλτιστου ενεργού συνόλου καλείται σύνολο εργασίας και συμβολίζεται με $\mathcal{W}$.*

Αφού ορίσουμε το σύνολο εργασίας $\mathcal{W}$, επιλύουμε το πρόβλημα στο οποίο οι περιορισμοί που ανήκουν στο σύνολο $\mathcal{W}$ επιβάλλονται ως ισότητες, ενώ εκείνοι που δεν ανήκουν σε αυτό αγνοούνται. Στην συνέχεια, ψάχνουμε να βρούμε εάν υπάρχει διάνυσμα πολλαπλασιαστών Lagrange λ το οποίο να ικανοποιεί τις συνθήκες KKT, με την λύση x^* που βρήκαμε για αυτό το $\mathcal{W}$. Εάν υπάρχει τέτοιο λ, τότε το x^* που βρήκαμε αποτελεί τοπική λύση του προβλήματος (3.1). Εάν δεν υπάρχει τέτοιο λ, τότε επιλέγουμε ένα διαφορετικό σύνολο $\mathcal{W}$ και επαναλαμβάνουμε την διαδικασία.

Αλγόριθμος 3.1. (Μέθοδος Ενεργού Συνόλου, βλ. [14], Κεφάλαιο 10, Αλγόριθμος, σελ. 241)

Επίλεξε αρχικό σημείο x^0
$W_0 = \{i \in \mathcal{I} : A_i x^0 = b_i\}$
$k \leftarrow 0$
Όσο ένα κριτήριο σύγκλισης δεν ικανοποιείται **επανάλαβε**
    Υπολόγισε την βέλτιστη λύση y^k του προβλήματος

$$\min \ \frac{1}{2}x^T Q x + c^T x + d,$$

$$\text{υπό τους περιορισμούς} \ \ A_i x = b_i, \forall i \in W_k$$

    Αν $y^k \neq x^k$ **τότε**
        Αν y^k είναι εφικτό σημείο **τότε**
            $t_k = 1$
        αλλιώς

$$t_k = \min\left\{\frac{b_i - A_i x^k}{A_i(y^k - x^k)} \ : i \notin W_k, \ \ - A_i(y^k - x^k) > 0\right\}$$

        Τέλος αν
        $x^{k+1} = x^k + t_k(y^k - x^k)$
        $W_{k+1} = W_k \cup \{i \notin W_k : A_i x^{k+1} = b_i\}$
        $k \leftarrow k+1$
    αλλιώς
        Υπολόγισε τους πολλαπλασιαστές Lagrange λ^k που αντιστοιχούν στην λύση y^k
        Αν $\lambda^k \geq 0$ **τότε**
            Σταμάτησε και επίστρεψε x^k
        αλλιώς
            $x^{k+1} = x^k$
            $\lambda_j^k = \min_{i \in W_k} \lambda_i^k$
            $W_{k+1} = W_k \setminus \{j\}$
            $k \leftarrow k+1$
        Τέλος αν
    Τέλος αν
Τέλος επανάληψης

Ο παραπάνω αλγόριθμος είναι σχεδιασμένος ειδικά για τα προβλήματα με τετραγωνικές συναρτήσεις, δηλαδή συναρτήσεις της μορφής $f(x) = \frac{1}{2}x^TQx + c^Tx + d$, όπου Q είναι ένας θετικά ορισμένος πίνακας, c ένα διάνυσμα και d ένας σταθερός όρος, και οι περιορισμοί είναι γραμμικές συναρτήσεις. Το αρχικό σύνολο εργασίας που χρησιμοποιείται αποτελείται από εκείνους τους περιορισμούς που είναι ενεργοί στο αρχικό σημείο x^0. Για τον κώδικα Matlab της μεθόδου βλ. Παράρτημα Β΄.1.

3.2 Παραδείγματα

3.2.1 Πρόβλημα με τρεις περιορισμούς ανισότητες

Για την επιβεβαίωση της σωστής λειτουργίας του κώδικα θα συγκρίνουμε την θεωρητική λύση με εκείνη που προκύπτει από τον αλγόριθμο, μέσω ενός απλού παραδείγματος Τετραγωνικού Προγραμματισμού. Αρχικά, με τις συνθήκες KKT του Θεωρήματος 2.1 θα βρούμε το σημείο στο οποίο η συνάρτηση λαμβάνει την ελάχιστη τιμή της σύμφωνα με τους περιορισμούς, καθώς και τους αντίστοιχους πολλαπλασιαστές Lagrange. Έπειτα με την μέθοδο του Ενεργού Συνόλου θα επιβεβαιώσουμε αυτά τα αποτελέσματα με την χρήση του μαθηματικού πακέτου Matlab.

Παράδειγμα 3.2.1 (βλ. [9], Κεφάλαιο 16, Άσκηση 16.1, σελ. 492)**.** *Έστω το πρόβλημα:*

$$\min\ 2x_1 + 3x_2 + 4x_1^2 + 2x_1x_2 + x_2^2, \quad \textit{υπό τους περιορισμούς} \quad \begin{array}{rl} x_1 - x_2 & \geq 0, \\ x_1 + x_2 & \leq 4, \\ x_1 & \leq 3. \end{array} \tag{3.2}$$

Επομένως, για το πρόβλημα (3.2) έχουμε ότι η αντικειμενική συνάρτηση είναι:

$$f(x) = 2x_1 + 3x_2 + 4x_1^2 + 2x_1x_2 + x_2^2, \quad \textit{με} \ \ \mathcal{E} = \emptyset, \ \mathcal{I} = \{1, 2, 3\}$$

και οι περιορισμοί: $c_1(x) = x_1 - x_2,\ c_2(x) = 4 - x_1 - x_2,\ c_3(x) = 3 - x_1.$

Άρα, οι κλίσεις των συναρτήσεων αυτών είναι:

$$\nabla f(x) = \begin{bmatrix} 2 + 8x_1 + 2x_2 \\ 3 + 2x_1 + 2x_2 \end{bmatrix}, \nabla c_1(x) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \nabla c_2(x) = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \nabla c_3(x) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}.$$

Σύμφωνα με τις συνθήκες KKT έχουμε να λύσουμε το εξής σύστημα:

$$\begin{aligned} \nabla f(x) - \lambda_1 \nabla c_1(x) - \lambda_2 \nabla c_2(x) - \lambda_3 \nabla c_3(x) &= 0, \\ \lambda_i &\geq 0, \quad \textit{για κάθε} \ \ i \in \mathcal{I} \\ \lambda_i^* c_i(x^*) &= 0, \quad \textit{για κάθε} \ \ i \in \mathcal{I}. \end{aligned}$$

Επομένως το σύστημα:

$$\begin{aligned}
\begin{bmatrix} 2+8x_1+2x_2 \\ 3+2x_1+2x_2 \end{bmatrix} - \lambda_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} - \lambda_2 \begin{bmatrix} -1 \\ -1 \end{bmatrix} - \lambda_3 \begin{bmatrix} -1 \\ 0 \end{bmatrix} &= 0, & (3.3)\\
\lambda_1, \lambda_2, \lambda_3 &\geq 0, & (3.4)\\
\lambda_1(x_1 - x_2) &= 0, & (3.5)\\
\lambda_2(4 - x_1 - x_2) &= 0, & (3.6)\\
\lambda_3(3 - x_1) &= 0. & (3.7)
\end{aligned}$$

Επιλύοντας το παραπάνω σύστημα, διαπιστώνουμε ότι για $\lambda_1 = \lambda_2 = \lambda_3 = 0$ *η αντικειμενική συνάρτηση λαμβάνει την ελάχιστη τιμή της. Έτσι, από την συνθήκη (3.3) για αυτούς τους πολλαπλασιαστές Lagrange έχουμε ότι η βέλτιστη λύση του προβλήματος είναι:*

$$\begin{cases} 8x_1 + 2x_2 + 2 = 0 \\ 2x_1 + 2x_2 + 3 = 0 \end{cases} \Longrightarrow \begin{cases} x_1 = \dfrac{1}{6} \\ x_2 = -\dfrac{5}{3} \end{cases} \Longrightarrow x^* = \begin{bmatrix} \dfrac{1}{6} \\ -\dfrac{5}{3} \end{bmatrix},$$

με την αντίστοιχη τιμή της αντικειμενικής συνάρτησης $f(x^*) = -\dfrac{7}{3}$.

Θα επιλύσουμε το πρόβλημα (3.2) με την μέθοδο active set που παρουσιάστηκε στην Ενότητα 3.1. Αρχικά ξαναγράφουμε το πρόβλημα έτσι ώστε όλοι οι περιορισμοί να είναι της μορφής $Ax \geq b$. *Τότε εχουμε:*

$$\min\ 2x_1 + 3x_2 + 4x_1^2 + 2x_1x_2 + x_2^2, \quad \text{υπό τους περιορισμούς} \quad \begin{aligned} x_1 - x_2 &\geq 0, \\ -x_1 - x_2 &\geq -4, \\ -x_1 &\geq -3. \end{aligned}$$

Ο πίνακας Q *και* c *της συνάρτησης* f *είναι:*

$$Q = \begin{bmatrix} 8 & 2 \\ 2 & 2 \end{bmatrix}, c = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

Ο πίνακας των συντελεστών των γραμμικων περιορισμών και των σταθερών όρων είναι:

$$A = \begin{bmatrix} 1 & -1 \\ -1 & -1 \\ -1 & 0 \end{bmatrix}, b = \begin{bmatrix} 0 \\ -4 \\ -3 \end{bmatrix}.$$

Επιλέγουμε ως αρχικό σημείο το $x^0 = (0,0)^T$, *το οποίο είναι εύκολο να διαπιστώσουμε ότι ανήκει στο εφικτό σύνολο.*

Από το Σχήμα 3.1, παρατηρούμε ότι ο αλγόριθμος active set μετά από 3 επαναλήψεις συγκλίνει σε ένα σημείο. Το τοπικό ελάχιστο που υπολογίζει το Matlab είναι το σημείο $x^* = (0.1667, -1,6667)^T = \left(\dfrac{1}{6}, -\dfrac{5}{3}\right)^T$, *με την τιμή της αντικειμενικής συνάρτησης* *στο σημείο αυτό να είναι* $f(x^*) = -2.3333 = -\dfrac{7}{3}$ *και το βέλτιστο διάνυσμα πολλαπλασιαστών Lagrange να είναι το* $\lambda^* = (0, 0, 0)$.

Έτσι, λοιπόν, διαπιστώνουμε ότι τα αποτελέσματα του αλγορίθμου είναι ίδια με εκείνα της θεωρητικής επίλυσης, επιβεβαιώνοντας την σωστή λειτουργία του κώδικα.

```
>> activeset

  Iter        W          x          f(x)
     0        1          0             0
                         0
     1        1    -0.3571       -0.8929
                   -0.3571
     2        0    -0.3571       -0.8929
                   -0.3571
     3        0     0.1667       -2.3333
                   -1.6667

 xk =               fvalue =

    0.1667             -2.3333
   -1.6667
```

Σχήμα 3.1: Το πρόβλημα (3.2), όπου φαίνονται τα αποτελέσματα της μεθόδου active set

3.2.2 Πρόβλημα με πέντε περιορισμούς ανισότητες

Για την αξιολόγηση του αλγορίθμου που παρουσιάστηκε στην Ενότητα 3.1 θα επιλύσουμε ένα παράδειγμα Τετραγωνικού Προγραμματισμού με την μέθοδο του Ενεργού Συνόλου και στην συνέχεια θα συγκρίνουμε τα αποτελέσματα που θα προκύψουν με εκείνα του solver fmincon στο Matlab.

Στην Ενότητα 1.10 είδαμε τα χαρακτηριστικά που πρέπει να έχουν οι αλγόριθμοι που χρησιμοποιούμε για την επίλυση προβλημάτων βελτιστοποίησης. Ωστόσο, θα δούμε ότι δεν έχουν όλοι αλγόριθμοι τα ίδια χαρακτηριστικά και δεν χρησιμοποιούνται όλοι στα ίδια προβλήματα.

Τα προβλήματα χωρίζονται σε διάφορες κατηγορίες, ανάλογα τις συναρτήσεις από τις οποίες αποτελούνται. Έτσι, λοιπόν, το κάθε πρόβλημα επιλύεται με έναν διαφορετικό solver, χρησιμοποιώντας ο καθένας τους διαφορετικούς αλγορίθμους του Matlab.

Παράδειγμα 3.2.2 (βλ. [9], Κεφάλαιο 16, Παράδειγμα 16.4, σελ. 475)**.** *Έστω το πρόβλημα:*

$$\min\ (x_1-1)^2+(x_2-2.5)^2, \quad \text{υπό τους περιορισμούς} \quad \begin{aligned} x_1-2x_2 &\geq -2,\\ -x_1-2x_2 &\geq -6,\\ -x_1+2x_2 &\geq -2,\\ x_1 &\geq 0,\\ x_2 &\geq 0. \end{aligned} \tag{3.8}$$

Επομένως, για το πρόβλημα (3.8) έχουμε ότι η αντικειμενική συνάρτηση είναι:

$$f(x)=(x_1-1)^2+(x_2-2.5)^2, \quad \text{με } \mathcal{E}=\emptyset,\ \mathcal{I}=\{1,2,3,4,5\}$$

και οι περιορισμοί:

$$c_1(x)=x_1-2x_2+2,\ c_2(x)=-x_1-2x_2+6,\ c_3(x)=-x_1+2x_2+2,$$

$$c_4(x)=x_1,\ c_5(x)=x_2.$$

Θα επιλύσουμε το πρόβλημα (3.8) με τον solver fmincon στο Matlab. Αρχικά, ξαναγράφουμε το πρόβλημα έτσι ώστε όλοι οι περιορισμοί να είναι της μορφής $Ax \leq b$. Τότε έχουμε:

$$\min\ (x_1-1)^2+(x_2-2.5)^2, \quad \text{υπό τους περιορισμούς} \quad \begin{aligned} -x_1+2x_2 &\leq 2,\\ x_1+2x_2 &\leq 6,\\ x_1-2x_2 &\leq 2,\\ -x_1 &\leq 0,\\ -x_2 &\leq 0. \end{aligned}$$

Ο πίνακας των συντελεστών των γραμμικων περιορισμών και των σταθερών όρων είναι:

$$A=\begin{bmatrix} -1 & 2\\ 1 & 2\\ 1 & -2\\ -1 & 0\\ 0 & -1 \end{bmatrix}, b=\begin{bmatrix} 2\\ 6\\ 2\\ 0\\ 0 \end{bmatrix}.$$

Επιλέγουμε ως αρχικό σημείο το $x^0=(2,0)^T$, το οποίο είναι εύκολο να διαπιστώσουμε ότι ανήκει στο εφικτό σύνολο.

Από το Σχήμα 3.2, παρατηρούμε ότι ο αλγόριθμος interior-point μετά από 9 επαναλήψεις συγκλίνει σε ένα σημείο. Το τοπικό ελάχιστο που υπολογίζει το Matlab είναι το σημείο $x^*=(1.4,1.7)^T$, με την τιμή της αντικειμενικής συνάρτησης στο σημείο αυτό να είναι $f(x^*)=0.8$ και το βέλτιστο διάνυσμα πολλαπλασιαστών Lagrange να είναι το $\lambda^*=(0.8,0,0,0,0)$.

Θα επιλύσουμε το πρόβλημα (3.8) με την μέθοδο active set στο Matlab. Αρχικά, ξαναγράφουμε την συνάρτηση f έτσι ώστε να είναι στην μορφή $f(x) = \frac{1}{2}x^TQx + c^Tx + d$, καθώς οι περιορισμοί του προβλήματος βρίσκονται ήδη στην μορφή $Ax \geq b$. Τότε έχουμε:

$$f(x) = (x_1-1)^2+(x_2-2.5)^2 = x_1^2-2x_1+1+x_2^2-5x_2+6.25 = x_1^2+x_2^2-2x_1-5x_2+7.25$$

Ο πίνακας Q και c της συνάρτησης f είναι:

$$Q = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, c = \begin{bmatrix} -2 \\ -5 \end{bmatrix}.$$

Ο πίνακας των συντελεστών των γραμμικων περιορισμών και των σταθερών όρων είναι:

$$A = \begin{bmatrix} 1 & -2 \\ -1 & -2 \\ -1 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, b = \begin{bmatrix} -2 \\ -6 \\ -2 \\ 0 \\ 0 \end{bmatrix}.$$

Επιλέγουμε πάλι ως αρχικό σημείο το $x^0 = (2,0)^T$.

Από το Σχήμα 3.3, παρατηρούμε ότι ο αλγόριθμος active set μετά από 5 επαναλήψεις συγκλίνει σε ένα σημείο. Το τοπικό ελάχιστο που υπολογίζει το Matlab είναι το σημείο $x^* = (1.4, 1.7)^T$, με την τιμή της αντικειμενικής συνάρτησης στο σημείο αυτό να είναι $f(x^*) = 0.8$ και το βέλτιστο διάνυσμα πολλαπλασιαστών Lagrange να είναι το $\lambda^* = (0.8, 0, 0, 0, 0)$. Επομένως, οι δύο αλγόριθμοι δίνουν τα ίδια αποτελέσματα.

Από το Σχήμα 3.2 και 3.3 παρατηρούμε ότι ο αλγόριθμος interior-point έχει μικρότερη ταχύτητα σύγκλισης στην βέλτιστη λύση σε σχέση με τον active set, καθώς ο πρώτος εκτελεί 9 επαναλήψεις ενώ ο δεύτερος 5 για να βρεί το x^*. Αυτό συμβαίνει διότι ο δεύτερος αλγόριθμος χρησιμοποιεί πιο μεγάλο βήμα στην προσέγγιση της λύσης το οποίο αυξάνει την ταχύτητά του.

Ωστόσο, ο interior-point σε σχέση με τον active set έχει πιο μεγάλη αποδοτικότητα. Αυτό συμβαίνει διότι η πρώτη μέθοδος απαιτεί λιγότερους υπολογιστικούς πόρους από τον υπολογιστή, καθώς εκτελεί λιγότερες πράξεις μέσα σε μία επανάληψη.

```
function f=objective(x)
f=(x(1)-1)^2+(x(2)-2.5)^2
```

```
>> [x,fval,exitflag,output,lambda]=fmincon(@objective,x0,A,b)

                                                    First-order      Norm of
 Iter F-count            f(x)  Feasibility    optimality         step
    0       3    7.250000e+00    0.000e+00     4.995e+00
    1       6    6.260936e+00    0.000e+00     4.496e+00    2.665e-01
    2       9    1.839587e+00    0.000e+00     2.265e+00    1.878e+00
    3      12    1.678729e+00    0.000e+00     8.969e-01    1.070e+00
    4      15    1.051286e+00    0.000e+00     2.682e-01    2.704e-01
    5      18    8.447435e-01    0.000e+00     4.474e-02    1.064e-01
    6      21    8.009155e-01    0.000e+00     1.177e-03    2.599e-02
    7      24    8.001998e-01    0.000e+00     1.998e-04    8.632e-04
    8      27    8.000020e-01    0.000e+00     2.016e-06    1.249e-04
    9      30    8.000000e-01    0.000e+00     2.000e-08    1.303e-06

x =              fval =           exitflag =

    1.4000           0.8000             1
    1.7000

output =

  struct with fields:

         iterations: 9
          funcCount: 30
     constrviolation: 0
           stepsize: 1.3028e-06
          algorithm: 'interior-point'
       firstorderopt: 2.0002e-08
        cgiterations: 0
             message: 'Local minimum found that satisfies the constra-
                       ints.↵↵Optimization completed because the
                       objective function is non-decreasing in ↵
                       feasible directions, to within the default
                       value of the optimality tolerance,↵and
                       constraints are satisfied to within the
                       default value of the constraint tolerance.↵↵
```

Σχήμα 3.2: Το πρόβλημα (3.8), όπου φαίνονται τα αποτελέσματα του solver fmincon με την χρήση του αλγορίθμου interior point

```
>> activeset

   Iter         W          x          f(x)
     0    3     5          2        7.2500
                           0
     1          5          2        7.2500
                           0
     2          5          1        6.2500
                           0
     3          0          1        6.2500
                           0
     4          1     1.0000        1.0000
                      1.5000
     5          1     1.4000        0.8000
                      1.7000

  xk =           fvalue =

     1.4000          0.8000
     1.7000
```

Σχήμα 3.3: Το πρόβλημα (3.8), όπου φαίνονται τα αποτελέσματα της μεθόδου active set

Βιβλιογραφικές αναφορές

Το Κεφάλαιο 3 βασίστηκε στο Κεφάλαιο Quadratic Programming του βιβλίου Numerical Optimization [βλ. [9], Κεφάλαιο 16, σελ. 467-478, 480-485, 492], καθώς επίσης και στην ενότητα Active Set Methods του βιβλίου Practical Methods of Optimization [βλ. [14], Κεφάλαιο 10, σελ. 240-245].

Ενότητα 3.1: Για την ενότητα αυτή, επιπλέον, αντλήθηκαν στοιχεία από το Κεφάλαιο Fundamentals of Constrained Algorithms του βιβλίου Numerical Optimization [βλ. [9], Κεφάλαιο 15, σελ. 421-426].

Βιβλιογραφία

[1] Ανάργυρος Γ. Φελλούρης. *Γραμμική Άλγεβρα και Αναλυτική Γεωμετρία*. Τσότρας, 2017.

[2] James E. Gentle. *Matrix Algebra: Theory, Computations, and Applications in Statistics*. Springer Publishing Company, Incorporated, 2007.

[3] Σπύρος Αργυρός. Σημειώσεις Παραδόσεων Συναρτησιακής Ανάλυσης. ΣΕΜΦΕ, ΕΜΠ, 05 2004.

[4] Ν. Καδιανάκης, Σ. Καρανάσιος και Α. Φελλούρης. *Ανάλυση ΙΙ Συναρτήσεις Πολλών Μεταβλητών*. Ιδιωτική Έκδοση, Έκδοση Όγδοη, 2009.

[5] Ιωάννης Πολυράκης. *Θέματα Ανάλυσης και Θεωρία Γενικής Ισορροπίας στην Οικονομία*. Ιδιωτική Έκδοση, 2016.

[6] Anthony L. Peressini, Francis E. Sullivan and J.J. Jr. Uhl. *The Mathematics of Nonlinear Programming*. Springer-Verlag New York, 1988.

[7] Erwin Kreyszig. *Introductory Functional Analysis with Applications*. Wiley, 1989.

[8] Αλέξανδρος Μπακόπουλος και Ίων Χρυσοβέργης. *Εισαγωγή στην Αριθμητική Ανάλυση*. Συμεών, 2009.

[9] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization*. Springer, New York, NY, Second Edition, 2006.

[10] Θεμιστοκλής Μ. Ρασσιάς. *Μαθηματική Ανάλυση ΙΙ Συναρτήσεις πολλών μεταβλητών*. Τεύχος Α'. Συμεών, 2007.

[11] Θεμιστοκλής Μ. Ρασσιάς. *Μαθηματική Ανάλυση Ι*. Συμεών, 2011.

[12] Α. Καλογεροπούλου, Μ. Γκίκας, Δ. Καραγιαννάκης και Μ. Λάμπρου. *Αγγλο-Ελληνικό Λεξικό Μαθηματικών Όρων*. Τροχαλία, 1992.

[13] Philip Wolfe. A duality theorem for non-linear programming. *Quarterly of applied mathematics*, 19(3):239–244, 1961.

[14] Roger Fletcher. *Practical Methods of Optimization*. John Wiley and Sons, Second Edition, 2000.

[15] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[16] Λάζαρος Μωυσής, Μαρία Τσικαλοπούλου, Δημήτρης Διαμαντίδης, Ροδή Λύκου, Γεωργία Βουγιούκα, Μαρία Τσαουσίδου, Χρήστος Τσολάκης, Θάνος Τσάπαρης και Δανάη Παπαθεοδώρου. Εισαγωγή στη Latex για φοιτητές. (An Introduction to Latex in Greek). 09 2014.

[17] Cesar Lopez. *MATLAB Optimization Techniques.* Apress, 2014.

Παράρτημα Α΄

Α΄.1 Πεπλεγμένες συναρτήσεις (Implicit functions)

Ορισμός Α΄.1 (Πραγματική Πεπλεγμένη Συνάρτηση, βλ. [4], Κεφάλαιο 5, Ορισμός 4.1, σελ. 197)**.** *Θεωρούμε την πραγματική συνάρτηση* $h : A \subseteq \mathbb{R}^n \times \mathbb{R} \longrightarrow \mathbb{R}$, *$n+1$ μεταβλητών. Η συνάρτηση:*

$$z : U \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}, U \text{ ανοιχτό},$$

θα λέμε ότι είναι υπό πεπλεγμένη μορφή ή ότι είναι πεπλεγμένη συνάρτηση στην εξίσωση $h(x_1, x_2, \ldots, x_n, y) = 0$, *αν για κάθε* $(x_1, x_2, \ldots, x_n) \in U$ *ισχύουν:*

i) $(x_1, x_2, \ldots, x_n, z(x_1, x_2, \ldots, x_n)) \in A$,

ii) $h(x_1, x_2, \ldots, x_n, z(x_1, x_2, \ldots, x_n)) = 0$.

Ορισμός Α΄.2 (Διανυσματική Πεπλεγμένη Συνάρτηση, βλ. [4], Κεφάλαιο 5, Ορισμός, σελ. 205)**.** *Θεωρούμε την διανυσματική συνάρτηση* $h : A \subseteq \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^m$. *Η συνάρτηση:*

$$z : U \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m, U \text{ ανοιχτό},$$

θα λέμε ότι είναι υπό πεπλεγμένη μορφή ή ότι είναι πεπλεγμένη συνάρτηση στην εξίσωση $h(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m) = 0$, *αν για κάθε* $(x_1, x_2, \ldots, x_n) \in U$ *ισχύουν:*

i) $(x_1, x_2, \ldots, x_n, z_1(x_1, x_2, \ldots, x_n), \ldots, z_m(x_1, x_2, \ldots, x_n)) \in A$,

ii) $h(x_1, x_2, \ldots, x_n, z_1(x_1, x_2, \ldots, x_n), \ldots, z_m(x_1, x_2, \ldots, x_n)) = 0$.

Θεώρημα Α΄.1 (Θεώρημα Πεπλεγμένων Συναρτήσεων, βλ. [9], Παράρτημα Α, Θεώρημα Α.2, σελ. 631)**.** *Έστω συνάρτηση* $h : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^n$ *τέτοια ώστε να ικανοποιούνται οι συνθήκες:*

(i) $h(z^*, 0) = 0$, *για κάποιο* $z^* \in \mathbb{R}^n$,

(ii) η $h(\cdot, \cdot)$ *είναι συνεχώς παραγωγίσιμη σε μια περιοχή του* $(z^*, 0)$,

(iii) ο πίνακας $\nabla_z h(z, t)$ *είναι αντιστρέψιμος στο σημείο* $(z, t) = (z^*, 0)$, *δηλαδή* $\nabla_z h(z, t) \neq 0$ *στο* $(z, t) = (z^*, 0)$.

Τότε:

1. *υπάρχουν ανοιχτές περιοχές* $\mathcal{N}_z \subset \mathbb{R}^n$ *και* $\mathcal{N}_t \subset \mathbb{R}^m$, *των σημείων* z^* *και* 0 *αντίστοιχα, έτσι ώστε να ορίζεται μονοσήμαντα η συνεχής συνάρτηση:*

$$z(t) : \mathcal{N}_t \longrightarrow \mathcal{N}_z,$$

με τις ιδιότητες:

$$z^* = z(0) \quad \text{και} \quad h(z(t), t) = 0, \quad \text{για κάθε} \;\; t \in \mathcal{N}_t,$$

2. *η συνάρτηση* h *είναι* p *φορές συνεχώς παραγωγίσιμη ως προς* $z(t)$ *και* t, *για κάποιο* $p > 0$,

3. *η συνάρτηση* $z(t)$ *είναι* p *φορές συνεχώς παραγωγίσιμη ως προς* t *και ισχύει η σχεση*

$$\nabla z(t) = -\nabla_t h(z(t), t)[\nabla_z h(z(t), t)]^{-1}, \quad \text{για κάθε} \;\; t \in \mathcal{N}_t.$$

Παρατήρηση

Το Θεώρημα Α΄.1 χρησιμοποιείται συχνά σε παραμετρικά συστήματα γραμμικών εξισώσεων, όπου το z είναι η λύση του:

$$M(t)z = g(t),$$

όπου $M(\cdot) \in \mathbb{R}^{n \times n}$ με $M(0)$ αντιστρέψιμος πίνακας και $g(\cdot) \in \mathbb{R}^n$. Για να εφαρμόσουμε το Θεώρημα ορίζουμε:

$$h(z, t) = M(t)z - g(t).$$

Εάν $M(\cdot)$ και $g(\cdot)$ είναι συνεχώς παραγωγίσιμες σε κάποια περιοχή του 0, τότε από το θεώρημα έχουμε ότι η $z(t) = M(t)^{-1}g(t)$ είναι συνεχής συνάρτηση του t σε κάποια περιοχή του 0.

Α΄.2 Κλειστά σύνολα

Ορισμός Α΄.3 (Πληθικότητα). *Η πληθικότητα ή πληθάριθμος ενός πεπερασμένου συνόλου είναι ένας φυσικός αριθμός, που αντιπροσωπεύει τον αριθμό των στοιχείων του.*

Το παρακάτω Λήμμα χρησιμοποιείται στην απόδειξη του Λήμματος 2.2 (Λήμμα του Φάρκα) προκειμένου να διασφαλιστεί ότι η λύση του προβλήματος βελτιστοποίησης:

$$\min \|s - g\|_2^2, \quad \text{υπό τον περιορισμό} \;\; s \in K,$$

είναι καλά ορισμένη.

Λήμμα Α΄.1 (βλ. [9], Κεφάλαιο 12, Λήμμα 12.15, σελ. 350). *Ορίζουμε το σύνολο* N *ως εξής:*

$$N = \{By + Ct \mid y \geq 0\},$$

όπου B *και* C *πίνακες διαστάσεων* $n \times m$ *και* $n \times p$ *αντίστοιχα και* y, t *διανύσματα κατάλληλων διαστάσεων. Το σύνολο αυτό είναι κλειστό.*

Απόδειξη. Διαχωρίζοντας το διάνυσμα t σε θετικά και αρνητικά στοιχεία, έχουμε:

$$N = \left\{ \begin{bmatrix} B & C & -C \end{bmatrix} \begin{bmatrix} y \\ t^+ \\ t^- \end{bmatrix} \;\middle|\; \begin{bmatrix} y \\ t^+ \\ t^- \end{bmatrix} \geq 0 \right\}.$$

Επομένως, χωρίς βλάβη της γενικότητας υποθέτουμε ότι το σύνολο N γράφεται ως εξής:

$$N = \{By \mid y \geq 0\}.$$

Έστω ότι ο B είναι $n \times m$ διαστάσεων. Αρχικά θα δείξουμε ότι για κάθε $s \in N$, μπορούμε να γράψουμε $s = B_I y_I$ με $y_I \geq 0$, όπου $I \subset \{1, 2, \dots, m\}$, B_I είναι ο υποπίνακας-I στήλη του B, ο οποίος είναι πλήρους τάξης ως προς τις στήλες του, και ο I έχει την ελάχιστη δυνατή πληθικότητα.

Έστω ότι $K \subset \{1, 2, \dots, m\}$ είναι ένα σύνολο δεικτών με την ελάχιστη δυνατή πληθικότητα, τέτοιο ώστε:

$$s = B_K y_K, \;\; \text{με} \;\; y_K \geq 0,$$

με τις στήλες του B_K να είναι γραμμικά ανεξάρτητες. Εφόσον το K είναι το ελάχιστο, το y_K δεν έχει μηδενικά στοιχεία. Τότε υπάρχει διάνυσμα $w \neq 0$ τέτοιο ώστε $B_K w = 0$. Αφού $s = B_K(y_K + \tau w)$ για κάθε τ, μπορούμε να αυξήσουμε ή να μειώσουμε το τ από το 0, μέχρι ένα ή περισσότερα στοιχεία του $y_K + \tau w$ να γίνουν μηδέν, ενώ τα υπόλοιπα να παραμείνουν θετικά. Ορίζουμε το $\bar{K}$ ως το σύνολο που περιέχει τους δείκτες του K που αντιστοιχούν σε μη μηδενικά στοιχεία του $y_K + \tau w$ και το $\bar{y}_{\bar{K}}$ ως το διάνυσμα των αυστηρά θετικών στοιχείων του $y_K + \tau w$. Τότε έχουμε:

$$s = B_{\bar{K}} \bar{y}_{\bar{K}} \;\; \text{και} \;\; \bar{y}_{\bar{K}} \geq 0,$$

το οποίο είναι άτοπο καθώς το K είναι το σύνολο με την ελάχιστη δυνατή πληθικότητα που να έχει αυτή την ιδιότητα.

Έστω ακολουθία $\{s^k\}$ με $s^k \in N$, για κάθε k, και $s^k \to s$. Θα αποδείξουμε το Λήμμα δείχνοντας ότι $s \in N$. Από την προηγούμενη παράγραφο, για κάθε k μπορούμε να γράψουμε:

$$s^k = B_{I_k} y_{I_k}^k, \;\; \text{με} \;\; y_{I_k}^k \geq 0,$$

όπου I_k είναι το σύνολο με την ελάχιστη δυνατή πληθικότητα και οι στήλες του B_{I_k} είναι γραμμικά ανεξάρτητες. Αφού οι πιθανές επιλογές του συνόλου δεικτών I_k είναι πεπερασμένες, τουλάχιστον ένα σύνολο δεικτών εμφανίζεται στην ακολουθία. Διαλέγοντας κατάλληλο σύνολο I, μπορούμε να πάρουμε μία υπακολουθία I_k και χωρίς βλάβη της γενικότητας να υποθέσουμε ότι $I_k \equiv I$, για κάθε k. Τότε:

$$s^k = A_I y_I^k, \;\; \text{με} \;\; y_I^k \geq 0 \;\; \text{και} \;\; A_I \;\; \text{πλήρους τάξης ως προς τις στήλες του.}$$

Από το τελευταίο έχουμε ότι το $A_I^T A_I$ είναι αντιστρέψιμο, έτσι ώστε το y_I^k να ορίζεται μονοσήμαντα ως:

$$y_I^k = (A_I^T A_I)^{-1} A_I^T s^k, \;\; k = 0, 1, 2, \dots$$

Παίρνοντας όρια και χρησιμοποιώντας το γεγονός ότι $s^k \to s$, έχουμε:

$$y_I^k \to y_I \overset{\text{ορισ.}}{=} (A_I^T A_I)^{-1} A_I^T s,$$

και επίσης $y_I \geq 0$, αφού $y_I^k \geq 0$ για κάθε k. Επομένως, $s = B_I y_I$ με $y_I \geq 0$ και άρα $s \in N$. □

Α'.3 Παραγοντοποίηση QR

Η παραγοντοποίηση πινάκων είναι μία σημαντική διαδικασία τόσο στον σχεδιασμό αλγορίθμων όσο και στην ανάλυση προβλημάτων. Μία τέτοια παραγοντοποίηση είναι η QR. Για την παραγοντοποιήση ενός πραγματικού $m \times n, m \geq n$ πίνακα A, χρησιμοποιούμε τους μετασχηματισμούς Householder για να βρούμε πίνακες Q και R, τέτοιους ώστε $A = QR$. Ο πίνακας $Q \in \mathbb{R}^{m\times m}$ είναι ορθογώνιος και ο πίνακας $R \in \mathbb{R}^{m\times n}$ είναι άνω τριγωνικός.

Μετασχηματισμοί Householder

Ο αλγόριθμος είναι ο εξής:

1. Έστω a_1 η πρώτη στήλη του A. Υπολογίζουμε διάνυσμα $u = a_1 \pm ae_1$, όπου $a = \|a_1\|_2$. Το πρόσημο στο διάνυσμα u επιλέγεται να είναι το αντίθετο του πρώτου στοιχείου του a_1.

2. Θέτουμε $v = \dfrac{u}{\|u\|_2}$ και υπολογίζουμε τον πίνακα $Q_1 = I_m - 2vv^T$, όπου ο I_m είναι ο μοναδιαίος πίνακας m τάξης. Ο Q_1 έχει την ιδιότητα:
$$Q_1 A = \begin{bmatrix} a & * & \cdots & * \\ 0 & & & \\ \vdots & & \tilde{A} & \\ 0 & & & \end{bmatrix}.$$

3. Επαναλαμβάνουμε τα βήματα (1) και (2) στον πίνακα $\tilde{A} \in \mathbb{R}^{(m-1)\times(n-1)}$ και έτσι έχουμε έναν ορθογώνιο πίνακα $\tilde{Q}_2 \in \mathbb{R}^{(m-1)\times(m-1)}$. Τότε ο Q_2 ορίζεται ως εξής:
$$Q_2 = \begin{bmatrix} I_{m-1} & 0 \\ 0 & \tilde{Q}_2 \end{bmatrix} \in \mathbb{R}^{m\times m}.$$

4. Μετά από $t = \min\{m-1, n\}$ βήματα έχουμε τον πίνακα $R = Q_t \cdots Q_2 Q_1 A$ και τον πίνακα $Q = Q_1^T Q_2^T \cdots Q_t^T$.

Βιβλιογραφικές αναφορές

Παράρτημα Α'.1: Η ενότητα αυτή βασίστηκε στην ενότητα Τα Θεωρήματα Πεπλεγμένων Συναρτήσεων του βιβλίου Ανάλυση ΙΙ [βλ. [4], Κεφάλαιο 5, σελ. 196-210], καθώς επίσης και στο Παράρτημα Implicit Function Theorem του βιβλίου Numerical Optimization [βλ. [9], Παράρτημα Α, σελ. 630-631].

Παράρτημα Α'.2: Η ενότητα αυτή βασίστηκε στο Λήμμα 12.15 του βιβλίου Numerical Optimization [βλ. [9], Κεφάλαιο 12, σελ. 350-351].

Παράρτημα Α'.3: Η ενότητα αυτη βασίστηκε στις ενότητες Householder Transformations και Householder Reflections to Form the QR Factorization του βιβλίου

Matrix Algebra: Theory, Computations, and Applications in Statistics [βλ. [2], Κεφάλαιο 5, σελ. 180-181, 188-191].

Παράρτημα Β΄

Β΄.1 Μέθοδος Ενεργού Συνόλου στο Matlab

```
%Inputs by user:
%maxiter (μεγιστος αριθμος επαναληψεων)
%tol (σφαλμα)
%x0 (αρχικο σημειο διανυσμα n x 1)
%A (συντελεστες περιορισμων A*x>=b πινακας m x n)
%b (σταθεροι οροι περιορισμων A*x>=b πινακας m x 1)
%f (συναρτηση προς ελαχιστοποιηση)
%Q (πινακας της f)
%c (πινακας της f)
%d (σταθερος ορος της f)

%QP standard problem:
%min 1/2*x^T*Q*x+c^T*x+d
%υπο τους περιορισμους A*x>=b, i in I

%Αναγκαιες συνθηκες πρωτης ταξης για το παραπανω συστημα:
%[Q,-A^T;A,0] [x*;λ*] = [-c;b]
%οπου [Q,A^T;A,0] ο KKT πινακας του συστηματος
%οπου [x*;λ*] ο X πινακας του συστηματος
%οπου [-c;b] ο B πινακας του συστηματος

%εκκινηση αλγοριθμου
xk=x0;
k=0
W=[];
maxiter=100;  %default maxiterations
tol=1e-8;  %default tolerance

%βημα 0 αλγοριθμου
j=1;  %j μετρητης των στοιχειων του W
for i=1:size(A,1)  %i μετρητης του πληθους των περιορισμων
    if A{i}*x0==b(i)  %δημιουργια συνολου εργασιας με βαση το x0
        W(j)=i;
        j=j+1;
    end
end
W

%ελεγχος εαν το συνολο εργασιας ειναι κενο
if size(W)==0
    fprintf("x0 not feasible")  %τοτε το x0 δεν ειναι εφικτο σημειο
end
```

Σχήμα Β΄.1: Κώδικας Matlab Αλγορίθμου 3.1

```
while k<=maxiter
    %βημα 1 αλγοριθμου

    %ελεγχος εαν το W ειναι κενο
    Wempty = isempty(W);
    if Wempty==1  %αν το W ειναι κενο
        KKT=Q;
        B=-c;

    else  %αν το W δεν ειναι κενο

        %δημιουργια KKT πινακα του υποπροβληματος QP
        %δημιουργια πινακα [Q;A] με τους περιορισμους που εχει το W
        kkt_at = Q;
        for i=1:size(W,2)
            kkt_at = cat(1,kkt_at,A{W(i)});
        end

        %δημιουργια πινακα transpose(A) με τους περιορισμους που εχει το W
        kkt_bt=transpose(-A{W(1)});
        for i=2:size(W,2)
            kkt_bt=cat(2,kkt_bt,transpose(-A{W(i)}));
        end

        %δημιουργια πινακα [-A^T;0]
        kkt_bt = cat(1,kkt_bt,zeros(size(kkt_bt,2),size(kkt_bt,2)));
        KKT = cat(2,kkt_at,kkt_bt);

        %δημιουργια B πινακα του υποπροβληματος QP
        B = -c;
        for i=1:size(W,2)
            B=cat(1,B,b(W(i)));
        end

    end  %end if W κενο/μη κενο

    X = linsolve (KKT, B);  %X, ο πινακας των y,lambda
                            %KKT, ο KKT πινακας του συστηματος
                            %B, ο πινακας των -c,newb

    %διασπαση του πινακα X σε y και lambda
    y=[];
    lambda=[];
    for m=1:size(Q,1)  %m μετρητης των μεταβλητων xi
        y(m)=X(m);
    end
    for i=(m+1):(size(W,2)+m)  %i μετρητης των υπολοιπων στοιχειων του X
        lambda(i-m)=X(i);
    end
    y=transpose(y);
    lambda=transpose(lambda);
    %βημα 2 αλγοριθμου
    y
    xk
    fvalue=1/2*transpose(xk)*Q*xk+transpose(c)*xk+d
    dd = round(y,4)==round(xk,4);
    %ελεγχος αν το y~=xk
    if sum(dd) ~= size(xk,1)  %αν y~=xk
```

Σχήμα Β΄.2: Κώδικας Matlab Αλγορίθμου 3.1

```
        %βημα 3 αλγοριθμου αρχικοι υπολογισμοι
        %υπολογισμοι αν το y ειναι εφικτο σημειο με ολους τους περιορισμους
        samt=0;
        for i=1:size(A,1)
            if A{i}*y>=b(i)
                samt=samt+1;
            else
                samt=samt+0;
            end
        end

        %δημιουργια στοιχειων που θα συγκριθουν στο min
        fullW=1:size(A,1);
        antiW=setdiff(fullW,W);
        tchoices=[];
        j=1;
        for i=1:size(antiW,2)
            if -A{antiW(i)}*(y-xk)>0
                tchoices(j)=(b(antiW(i))-A{antiW(i)}*xk)/(A{antiW(i)}*(y-xk));
                j=j+1;
            end
        end

        %βημα 3 αλγοριθμου
        %ελεγχος αν το y ειναι εφικτο σημειο με ολους τους περιορισμους
        if samt==size(A,1)
            t=1;
        else
            t=min(tchoices);

        end

        xk=xk+t*(y-xk);
        for i=1:size(antiW,2)
            if A{antiW(i)}*xk==b(antiW(i))
                W=union(W,i);
            end
        end
        W
        k=k+1

    else  %αν y=xk

        %βημα 4 αλγοριθμου
        L = lambda>=0;
        if sum(L) == size(lambda,1)
            break
        else
            [lambdaj,j]=min(lambda);
            W=setdiff(W,W(j));
            W
            k=k+1
        end
    end  %end if y~=/=xk
end
```

Σχήμα Β΄.3: Κώδικας Matlab Αλγορίθμου 3.1

Βιβλιογραφικές αναφορές

Παράρτημα Β΄.1: Η ενότητα αυτή βασίστηκε στο βιβλίο MATLAB Optimization Techniques [βλ. [17]] και στο βιβλίο Numerical Optimization [βλ. [9], Κεφάλαιο 16, σελ. 451].

Κατάλογος Σχημάτων