ΑΝΩΤΑΤΟ ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ

ΠΑΤΡΩΝ

ΣΧΟΛΗ ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΟΙΚΟΝΟΜΙΑΣ

ΤΜΗΜΑ ΕΠΙΧΕΙΡΗΜΑΤΙΚΟΥ ΣΧΕΔΙΑΣΜΟΥ ΚΑΙ ΠΛΗΡΟΦΟΡΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

ΘΕΜΑ:

ΜΑΘΗΜΑΤΙΚΕΣ ΜΕΘΟΔΟΙ ΚΑΙ ΕΜΠΕΙΡΙΚΕΣ ΕΦΑΡΜΟΓΕΣ ΣΤΗ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗ ΘΕΩΡΙΑ ΚΑΙ ΤΗ ΒΙΟΜΗΧΑΝΙΚΗ ΟΡΓΑΝΩΣΗ

ΦΟΙΤΗΤΕΣ:

ΑΝΤΩΝΙΑΔΗΣ ΑΝΤΩΝΗΣ

ΚΟΥΝΤΟΥΡΗΣ ΙΩΑΝΝΗΣ

ΕΠΟΠΤΕΥΩΝ ΚΑΘΗΓΗΤΗΣ:

ΤΟΥΛΙΑΣ ΘΩΜΑΣ

ΠΕΡΙΕΧΟΜΕΝΑ

Σελίδα

ΠΕΡΙΛΗΨΗ

Στην ακόλουθη εργασία γίνεται μία ανάλυση της διαδικασίας επίλυσης προβλημάτων μη γραμμικής βελτιστοποίησης της οικονομικής δραστηριότητας. Για να γίνει κατανοητή η διαδικασία αυτή, πρέπει να γίνει αναφορά στα μέσα και τις διαδικασίες που είναι απαραίτητες στην επίλυση αυτού του είδους των προβλημάτων.

Μία βασική έννοια στη διαδικασία επίλυσης είναι η εύρεση των ακρότατων σημείων των συναρτήσεων. Για τον καθορισμό των ακρότατων χρησιμοποιούνται κάποια κριτήρια παραγώγων μέσω των οποίων προκύπτουν οι μέγιστες και ελάχιστες τιμές των συναρτήσεων, οι οποίες αποτελούν βασικό εργαλείο στην εύρεση των βέλτιστων τιμών. Μία σημαντική μέθοδος της διαδικασίας επίλυσης είναι η μέθοδος πολλαπλασιαστών Lagrange, για την οποία ακολουθεί εκτενέστερη ανάλυση στη συνέχεια της εργασίας.

Εφόσον γίνουν κατανοητές οι βασικές μέθοδοι που χρησιμοποιούνται στην επίλυση τέτοιου είδους προβλημάτων, γίνεται μια εισαγωγή στα μη γραμμικά προβλήματα και συγκεκριμένα στην γενική μορφή μη γραμμικού προβλήματος. Για τον σκοπό αυτό γίνεται χρήση αλγορίθμων οι οποίοι συντελούν στην ταχύτερη και με μεγαλύτερη ακρίβεια επίλυσής τους. Οι αλγόριθμοι αυτοί είναι η Γενική Μορφή Αλγορίθμου Βελτιστοποίησης καθώς και η Μέθοδος της Μεγαλύτερης Αλλαγής του Cauchy.

ΕΙΣΑΓΩΓΗ

Οι μαθηματικές μέθοδοι αποτελούν βασικό εργαλείο της οικονομικής δραστηριότητας καθώς και την βάση της οικονομικής θεωρίας. Ένα από τα βασικότερα προβλήματα στις Οικονομικές Επιστήμες αποτελεί η προσπάθεια για την μεγιστοποίηση των κερδών και την ελαχιστοποίηση του κόστους. Η μαθηματική μέθοδος που βοηθάει στην επίλυση τέτοιου είδους οικονομικών προβλημάτων είναι η εύρεση των ακρότατων στην συνάρτηση που προκύπτει από την εκάστοτε οικονομική δραστηριότητα. Μέσα από την εύρεση των ακρότατων γίνεται δυνατός ο εντοπισμός των μέγιστων και ελάχιστων τιμών των συναρτήσεων που προκύπτουν από τις οικονομικές δραστηριότητες, οι οποίες βοηθούν στην επιλογή των βέλτιστων λύσεων για τις συναρτήσεις αυτές.

Η βελτιστοποίηση αποτελεί την μαθηματική εκείνη διαδικασία κατά την οποία υπολογίζεται η καλύτερη λύση ενός μαθηματικού προβλήματος. Απαραίτητη προϋπόθεση είναι η ύπαρξη εναλλακτικών προτάσεων για την λύση του προβλήματος οι οποίες συνδέονται με κάποιο κόστος. Η έννοια "καλύτερη" αλληλεπιδρά με την έννοια του κόστους καθώς η καλύτερη επιλογή προϋποθέτει κάποιο κόστος . Εν αντιθέσει, το κόστος δεν συνεπάγεται χρηματική απόδοση. Για τη μέθοδο της βελτιστοποίησης σημαντικό παράγοντα αποτελεί η δημιουργία αλγορίθμων οι οποίοι βοηθούν στην ταχύτερη και πιο ακριβής επίλυση των προβλημάτων που προκύπτουν σε κάθε οικονομική δραστηριότητα.

1. ΚΕΦΑΛΑΙΟ 1

1.1 Ακρότατα και Σαγματικά σημεία

Η εύρεση των μέγιστων και ελάχιστων τιμών μιας συνάρτησης πολλών μεταβλητών, καθώς και των σημείων όπου αυτές προκύπτουν, είναι μία σημαντική εφαρμογή του διαφορικού λογισμού πολλών μεταβλητών. Σε αυτή την ενότητα θα εξεταστούν τέτοιου είδους εφαρμογές και θα βρίσκονται οι απαντήσεις με την μελέτη των μερικών παραγώγων κάποιας κατάλληλης συνάρτησης.

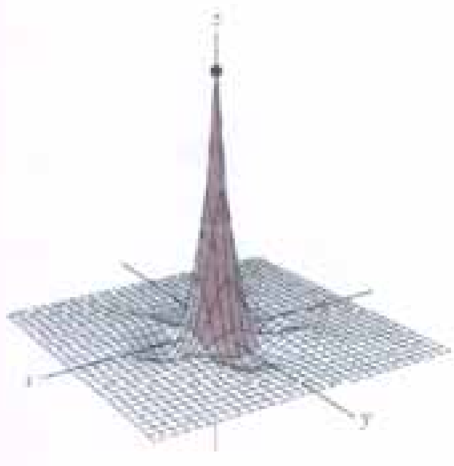
1.1.1. Συμπεριφορά συναρτήσεως σε κλειστές φραγμένες περιοχές

Όπως γίνεται γνωστό από την μελέτη συναρτήσεων μίας μεταβλητής, οι διαφορίσιμες συναρτήσεις είναι πολύ χρήσιμες στη μαθηματική περιγραφή προβλημάτων βελτιστοποίησης. Επειδή οι συναρτήσεις αυτές είναι συνεχείς, είναι γνωστό ότι σε κλειστά διαστήματα θα έπαιρναν τόσο την ελάχιστη όσο και την μέγιστη τιμή τους. Επειδή είναι διαφορίσιμες, προκύπτει ακόμη ότι θα έπαιρναν τις τιμές αυτές μονάχα σε συνοριακά σημεία ή σε εσωτερικά σημεία του πεδίου ορισμού όπου μηδενίζεται η πρώτη παράγωγος. Μερικές φορές, βρίσκονται συναρτήσεις που δεν είναι διαφορίσιμες σε ένα ή σε περισσότερα σημεία του πεδίου ορισμού, τα οποία και πρέπει να τα εξετάζονται χωριστά για την ύπαρξη ακρότατων.

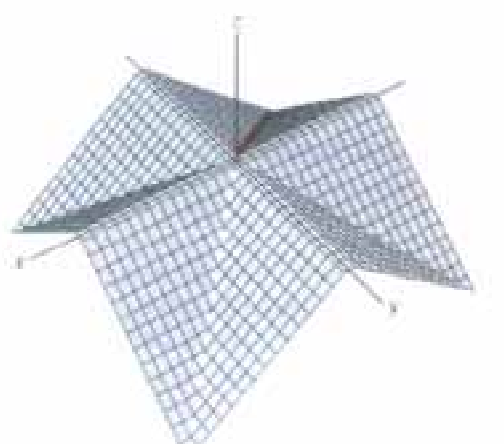
Είναι επίσης γνωστό ότι η συνθήκη $f'(c) = 0$, δεν σημαίνει πάντα την ύπαρξη ακρότατου. Σε τέτοια σημεία c, η γραφική παράσταση ενδέχεται να έχει σημείο καμπής αντί για τοπικό μέγιστο ή ελάχιστο. Δηλαδή η καμπύλη ενδέχεται να ανέρχεται καθώς πλησιάζει στο c, και κατόπιν πάλι να ανέρχεται καθώς αφήνει το c. Ή ακόμη μπορεί να κατέρχεται καθώς πλησιάζει το c, να οριζοντιώνεται στο c, και να συνεχίζει την κάθοδο της αποκρινόμενη. Με άλλα λόγια, η καμπύλη μπορεί να τέμνει την εφαπτόμενη της στο $x = c$.

Οι συναρτήσεις δύο μεταβλητών παρουσιάζουν παρόμοια συμπεριφορά. Επίσης, οι συνεχείς συναρτήσεις δύο μεταβλητών παίρνουν μέγιστες και ελάχιστες τιμές σε κλειστά, φραγμένα πεδία ορισμού, όπως παρατηρείται και στα Σχήματα 1 και 2. Είναι εφικτό να αποκλειστούν μερικά πιθανά σημεία ύπαρξης ακρότατου, αν διερευνηθούν οι πρώτες παράγωγοι των συναρτήσεων αυτών. Μία συνάρτηση δύο μεταβλητών μπορεί να εμφανίζει ακρότατα μονάχα σε συνοριακά σημεία του πεδίου ορισμού όπου και οι δύο πρώτες μερικές παράγωγοι μηδενίζονται ή όπου μία τουλάχιστον εκ των παραγώγων αυτών δεν υπάρχει.

Για μία ακόμη φορά, ο μηδενισμός των παραγώγων σε ένα εσωτερικό σημείο (a,b) δεν συνεπάγεται αναγκαστικά την ύπαρξη ακρότατου. Η επιφάνεια που αποτελεί τη γραφική παράσταση της συνάρτησης μπορεί να έχει το σχήμα σαμαριού πάνω από το (a,b) και τέμνει το εκεί εφαπτόμενο επίπεδο.



Σχήμα 1



Σχήμα 2

1.1.2. Κριτήρια παραγώγων για τοπικά ακρότατα.

Για να βρεθούν τα τοπικά ακρότατα συνάρτησης μίας μεταβλητής γίνεται έρευνα για σημεία όπου η γραφική παράσταση έχει οριζόντια εφαπτόμενη. Στη συνέχεια εξετάζεται αν καθένα από τα σημεία αυτά είναι σημείο τοπικού μεγίστου, τοπικού ελαχίστου, ή σημείο καμπής. Αν τώρα η συνάρτηση $f(x,y)$ έχει δύο μεταβλητές, γίνεται έρευνα για σημεία όπου η επιφάνεια $z=f(x,y)$ έχει οριζόντιο εφαπτόμενο επίπεδο. Κατόπιν εξετάζεται αν στα σημεία αυτά η συνάρτηση παρουσιάζει τοπικό μέγιστο, τοπικό ελάχιστο, ή σαγματική συμπεριφορά.

ΟΡΙΣΜΟΣ

Τοπικό μέγιστο και ελάχιστο

Έστω $f(x,y)$ ορισμένη σε περιοχή R που περιέχει το σημείο (a,b). Στην περίπτωση αυτή

1. Η $f(a,b)$ είναι τοπική μέγιστη τιμή της f αν $f(a,b) \geq f(x,y)$ για όλα τα σημεία (x,y) του πεδίου ορισμού που ανήκουν σε έναν ανοιχτό δίσκο με κέντρο το (a,b).

2. Η $f(a,b)$ είναι τοπική ελάχιστη τιμή της f αν $f(a,b) \leq f(x,y)$ για όλα τα σημεία (x,y) του πεδίου ορισμού που ανήκουν σε έναν ανοιχτό δίσκο με κέντρο το (a,b).

Τα τοπικά μέγιστα αντιστοιχούν στις κορυφές των « βουνών » της επιφάνειας $z=f(x,y)$, ενώ τα τοπικά ελάχιστα αντιστοιχούν στα χαμηλότερα σημεία των « κοιλάδων » όπως γίνεται αντιληπτό και στο Σχήμα

3. Στα σημεία αυτά, τα εφαπτόμενα επίπεδα, όταν υπάρχουν, είναι οριζόντια.

Τα τοπικά ακρότατα αποκαλούνται και **σχετικά ακρότατα**.

Σχήμα 3

Όπως ισχύει για τις συναρτήσεις μίας μεταβλητής, έτσι και εδώ το « κλειδί » για τον εντοπισμό τοπικών ακρότατων είναι ένα κριτήριο πρώτης παραγώγου.

1.1.2.1. Κριτήριο πρώτης παραγώγου για τοπικά ακρότατα

Αν η $f(x,y)$ εμφανίζει τοπικό μέγιστο ή ελάχιστο σε κάποιο εσωτερικό σημείο (a,b) του πεδίου ορισμού της όπου υπάρχουν οι πρώτες μερικές της παράγωγοι, τότε $f_x(a,b)=0$ και $f_y(a,b)=0$

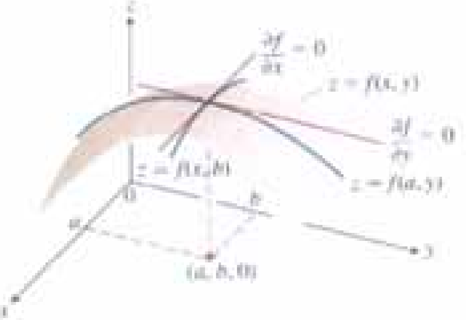
Απόδειξη : Έστω ότι η f έχει τοπικό μέγιστο σε ένα εσωτερικό σημείο (a,b) του πεδίου ορισμού της. Τότε

Το $x=a$ είναι εσωτερικό σημείο του πεδίου ορισμού της καμπύλης $z=f(x,b)$ που αποτελεί τομή του επιπέδου $y=b$ με την επιφάνεια $z=f(x,y)$

2. Η συνάρτηση $z=f(x,b)$ είναι διαφορίσιμη συνάρτηση του x στο $x=a$ (η παράγωγος είναι $f_x(a,b)$).

Η συνάρτηση $z=f(x,b)$ έχει τοπικό μέγιστο στο $x=a$.

4. Η τιμή της παραγώγου της $z = f(x,b)$ στο x = a είναι συνεπώς μηδέν. Εφόσον η παράγωγος είναι $f_x(a,b)$ συμπεραίνουμε ότι $f_x(a,b) = 0$.



Για $x = a$, $y = b$, προκύπτει τοπικό μέγιστο της f

Σχήμα 4

Επιχειρηματολογώντας κατά παρόμοιο τρόπο για τη συνάρτηση $z = f(a,y)$ βρίσκουμε ότι $f_y(a,b) = 0$

Αν αντικαταστήσουμε $f_x(a,b) = 0$ και $f_y(a,b) = 0$ στην εξίσωση $f_x(a,b)(x-a) + f_y(a,b)(y-b) - (z - f(a,b)) = 0$ επιπέδου που εφάπτεται στην επιφάνεια $z = f(x,y)$ στο σημείο (a,b), η εξίσωση παίρνει τη μορφή

$$0 * (x-a) + 0 * (y-b) - z + f(a,b) = 0 \text{ δηλαδή } z = f(a,b)$$

Έτσι το κριτήριο πρώτης παραγώγου για τοπικά ακρότατα αποδεικνύει πως αν υπάρχει εφαπτόμενο επίπεδο της επιφάνειας σε τοπικό ακρότατο, το επίπεδο αυτό θα είναι οριζόντιο.

Όπως ισχύει και για τις συναρτήσεις μίας μεταβλητής το θεώρημα αποδεικνύει ότι τα μόνα σημεία όπου μία συνάρτηση $f(x,y)$ μπορεί να εμφανίζει ακρότατα είναι:

1. Εσωτερικά σημεία όπου $f_x = f_y = 0$
2. Εσωτερικά σημεία όπου μία τουλάχιστον των f_x και f_y δεν υπάρχει
3. Συνοριακά σημεία του πεδίου ορισμού της συναρτήσεως.

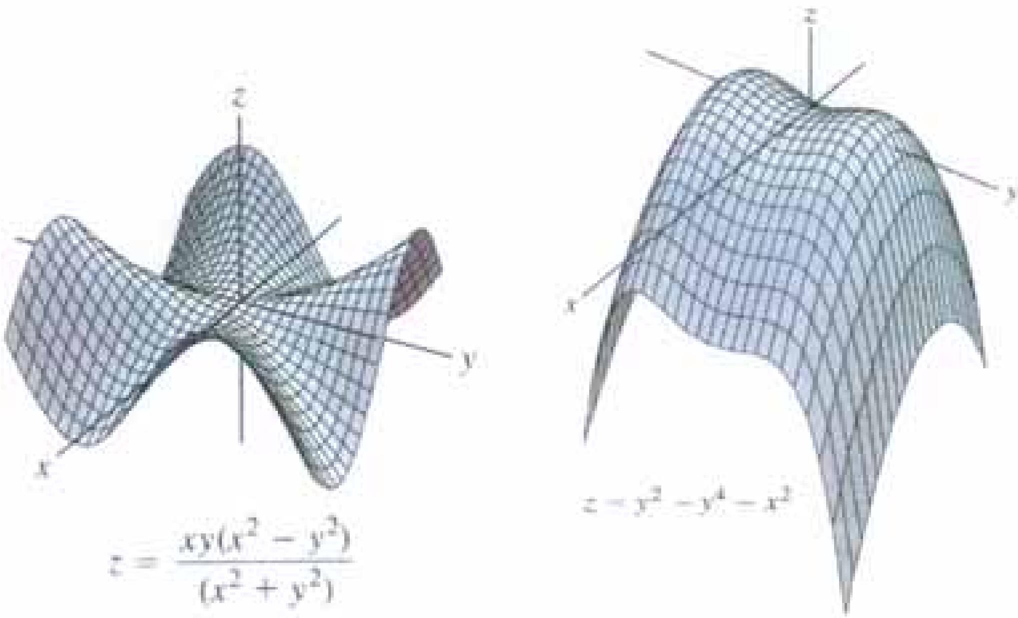
ΟΡΙΣΜΟΣ

Κρίσιμο σημείο

Ένα εσωτερικό σημείο του πεδίου ορισμού της συνάρτησης $f(x,y)$ όπου τόσο η f_x όσο και η f_y μηδενίζεται, ή όπου τουλάχιστον η μία εκ των f_x και f_y δεν υπάρχει, καλείται κρίσιμο σημείο της.

Δηλαδή, τα μόνα σημεία όπου μια συνάρτηση $f(x,y)$ μπορεί να εμφανίζει ακρότατα είναι τα κρίσιμα και τα συνοριακά (ακραία) σημεία. Όπως συμβαίνει και για τις διαφορίσιμες συναρτήσεις μίας μεταβλητής, ένα κρίσιμο σημείο δεν είναι απαραίτητα σημείο τοπικού ακρότατου. Μία διαφορίσιμη συνάρτηση δύο μεταβλητών μπορεί να διαθέτει ένα **σαγματικό σημείο.**

Δηλαδή μία διαφορίσιμη συνάρτηση $f(x,y)$ έχει σαγματικό σημείο σε ένα κρίσιμο σημείο (a,b) αν σε κάθε ανοιχτό κυκλικό δίσκο με κέντρο το (a,b) υπάρχουν σημεία (x,y) του πεδίου ορισμού όπου $f(x,y) > f(a,b)$ και σημεία (x,y) του πεδίου ορισμού όπου $f(x,y) < f(a,b)$. Το σημείο $(a,b,f(a,b))$ πάνω στην επιφάνεια $z = f(x,y)$ καλείται **σαγματικό σημείο.** Αυτό γίνεται καλύτερα αντιληπτό στο Σχήμα 5 που παρουσιάζονται δύο περιπτώσεις σαγματικών σημείων στην αρχή των αξόνων.



Σχήμα 5

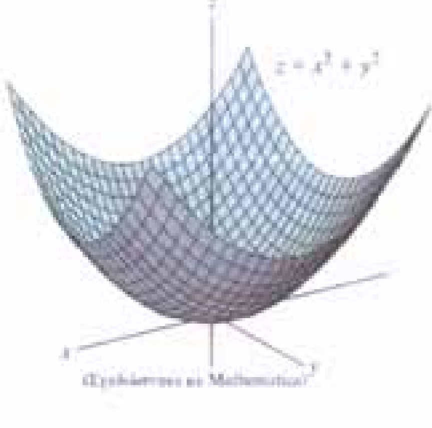
<u>*ΠΑΡΑΔΕΙΓΜΑ:*</u>

Εύρεση τοπικών ακρότατων

Να βρεθούν τα τοπικά ακρότατα της $f(x,y) = x^2 + y^2$.

ΛΥΣΗ: Πεδίο ορισμού της f είναι όλο το επίπεδο (άρα δεν υπάρχουν συνοριακά σημεία), ενώ οι μερικές παράγωγοι $f_x = 2x$ και $f_y = 2y$ υπάρχουν παντού. Συνεπώς, τοπικά ακρότατα μπορούν μονάχα να προκύψουν σε σημεία όπου

$$f_x = 2x = 0 \text{ και } f_y = 2y = 0$$



Σχήμα 6

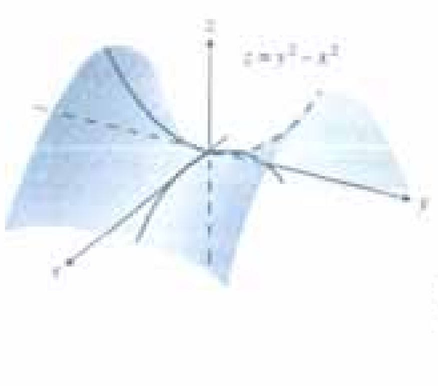
Σχήμα 6: Η γραφική παράσταση της συνάρτησης $f(x,y)=x^2+y^2$ είναι το παραβολοειδές $z=x^2+y^2$. Η συνάρτηση έχει μόνο ένα κρίσιμο σημείο, στην αρχή των αξόνων. Πρόκειται για σημείο τοπικού ελαχίστου όπου η συνάρτηση παίρνει εκεί την τιμή 0.

Έτσι, μόνο πιθανό σημείο ύπαρξης ακρότατου είναι η αρχή, όπου η συνάρτηση f μηδενίζεται. Εφόσον η f δεν γίνεται ποτέ αρνητική, βλέπουμε ότι η αρχή όντως αποτελεί σημείο τοπικού ελαχίστου.

Εντοπισμός και ταυτοποίηση σαγματικού σημείου

Βρείτε τα τοπικά ακρότατα (αν υπάρχουν) της $f(x,y)=y^2-x^2$.

ΛΥΣΗ: Πεδίο ορισμού της f είναι όλο το επίπεδο (συνεπώς δεν υπάρχουν συνοριακά σημεία), ενώ οι μερικές παράγωγοι $f_x=-2x$ και $f_y=2y$ υπάρχουν παντού. Συνεπώς, τοπικά ακρότατα μπορούν μονάχα να προκύψουν στην αρχή (0, 0). Κατά μήκος του θετικού ημιάξονα x η f παίρνει την τιμή $f(x,0)=-x^2<0$. Από την άλλη, κατά μήκος του θετικού ημιάξονα y, η f παίρνει την τιμή $f(0,y)=y^2>0$. Συνεπώς, κάθε ανοιχτός κυκλικός δίσκος του επιπέδου xy με κέντρο το (0, 0) θα περιέχει σημεία όπου η συνάρτηση είναι θετική και σημεία όπου η συνάρτηση είναι αρνητική. Η συνάρτηση έχει λοιπόν σαγματικό σημείο στην αρχή αντί για τοπικό ακρότατο(Σχήμα 7).



Σχήμα 7

Το γεγονός ότι $f_x = f_y = 0$ σε ένα εσωτερικό σημείο (a,b) του R δεν εγγυάται ότι η f θα έχει τοπικό ακρότατο εκεί. Αν ωστόσο η f και οι μερικές παράγωγοι πρώτης και δεύτερης τάξης είναι συνεχείς στο R, τότε το ακόλουθο θεώρημα μπορεί να διαφωτίσει περισσότερο την κατάσταση.

1.1.2.2. Κριτήριο δεύτερης παραγώγου για τοπικά ακρότατα

Έστω ότι η $f(x,y)$ και οι μερικές της παράγωγοι πρώτης και δεύτερης τάξης είναι παντού συνεχείς σε έναν κυκλικό δίσκο με κέντρο το (a,b) και ότι $f_x(a,b)=0$. Τότε :

1. Η f έχει <u>τοπικό μέγιστο</u> στο (a,b) αν $f_{xx}<0$ και $f_{xx}f_{yy}-f_{xy}^2>0$ στο (a,b).
2. Η f έχει <u>τοπικό ελάχιστο</u> στο (a,b) αν $f_{xx}>0$ και $f_{xx}f_{yy}-f_{xy}^2>0$ στο (a,b).
3. Η f έχει σαγματικό σημείο στο (a,b) αν $f_{xx}f_{yy}-f_{xy}^2<0$ στο (a,b)
4. Δεν μπορούμε να αποφανθούμε για τη φύση του σημείου (a,b) αν $f_{xx}f_{yy}-f_{xy}^2=0$ στο (a,b). Στην περίπτωση αυτή, θα πρέπει να βρούμε άλλον τρόπο μελέτης της συμπεριφοράς της f στο (a,b).

Η έκφραση $f_{xx}f_{yy}-f_{xy}^2$ καλείται <u>*διακρίνουσα*</u> ή <u>*Εσσιαννή*</u> της f. Η απομνημόνευσή της είναι μάλλον ευκολότερη στη μορφή ορίζουσας

$$f_{xx}f_{yy}-f_{xy}^2=\begin{vmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{vmatrix}$$

Σύμφωνα με το ανωτέρω θεώρημα, αν η διακρίνουσα είναι θετική στο σημείο (a,b), τότε η επιφάνεια θα καμπυλώνεται κατά τον ίδιο τρόπο σε όλες τις κατευθύνσεις: προς τα κάτω αν $f_{xx}<0$, οπότε θα υπάρχει τοπικό μέγιστο, και προς τα πάνω αν $f_{xx}>0$, οπότε θα υπάρχει τοπικό ελάχιστο. Από την άλλη, αν η διακρίνουσα είναι αρνητική στο (a,b), τότε η επιφάνεια θα καμπυλώνεται προς τα πάνω σε μερικές κατευθύνσεις και προς τα κάτω σε άλλες άρα θα έχουμε <u>*σαγματικό σημείο*</u>.

Εύρεση τοπικών ακρότατων

Βρείτε τα τοπικά ακρότατα της συνάρτησης $f(x,y)=xy-x^2-y^2-2x-2y+4$

ΛΥΣΗ: Η συνάρτηση ορίζεται και είναι διαφορίσιμη για κάθε x και y το πεδίο ορισμού της δεν έχει συνοριακά σημεία. Η συνάρτηση έχει συνεπώς ακρότατα μόνο σε σημεία όπου οι fx και fy μηδενίζονται ταυτόχρονα. Δηλαδή σε σημεία όπου

$$f_x = y-2x-2=0\ ,$$

δηλαδή

$$x=y=-2$$

Συνεπώς, το $(-2,-2)$ είναι το μόνο σημείο όπου η f ενδέχεται να έχει ακρότατο. Για να δούμε τι τελικά συμβαίνει, υπολογίζουμε τις

$$f_{xx}=-2,\quad f_{yy}=-2,\quad f_{xy}=1$$

Η διακρίνουσα της f στο $(a,b)=(-2,-2)$ είναι

$$f_{xx}f_{yy}-f_{xy}^2=(-2)(-2)-(1)^2=4-1=3$$

Το γεγονός ότι $f_{xx}<0$ και $f_{xx}f_{yy}-f_{xy}^2>0$ σημαίνει ότι η f εμφανίζει τοπικό μέγιστο στο $(-2,-2)$. Η τιμή της f στο σημείο αυτό είναι $f(-2,-2)=8$.

Αναζήτηση τοπικών ακρότατων

Να βρεθούν τα τοπικά ακρότατα της $f(x,y)=xy$

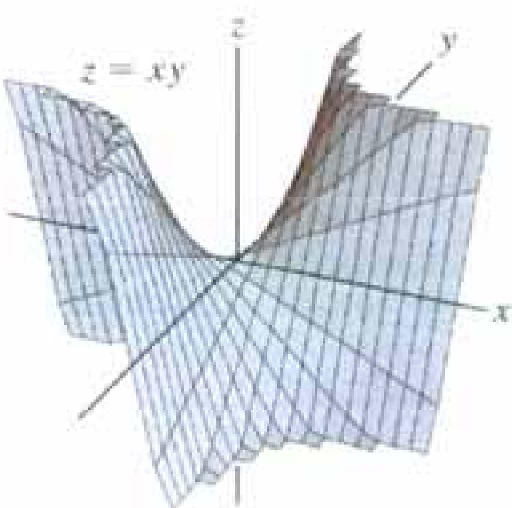
ΛΥΣΗ: Εφόσον η f είναι διαφορίσιμη παντού(Σχήμα 8), θα εμφανίζει ακρότατα μόνο σε σημεία όπου $f_x=y=0$ και $f_y=x=0$.

Με άλλα λόγια , η αρχή των αξόνων είναι το μόνο πιθανό σημείο ύπαρξης ακρότατων της f. Για να δούμε τι όντως συμβαίνει εκεί, υπολογίζουμε τις παραγώγους

$$f_{xx}=0 \quad , \quad f_{yy}=0 \quad , \quad f_{xy}=1 .$$

Η διακρίνουσα $f_{xx}f_{yy}-f_{xy}^2=-1$, είναι αρνητική. Συνεπώς, η συνάρτηση έχει σαγματικό σημείο στο $(0,0)$. Συμπεραίνουμε λοιπόν ότι η $f(x,y)=xy$ δεν έχει τοπικά ακρότατα.

Σχήμα 8



1.1.3. Ολικά (απόλυτα) μέγιστα και ελάχιστα σε κλειστές φραγμένες περιοχές

Η διαδικασία διερεύνησης για ολικά (ή απόλυτα) ακρότατα συνεχούς συναρτήσεως $f(x,y)$ σε ένα κλειστό και φραγμένο χωρίο R μπορεί να παρουσιαστεί σε τρία βήματα.

<u>*Βήμα 1^o*</u>: Καταγράφουμε τα εσωτερικά σημεία του R όπου η f ενδέχεται να έχει τοπικά μέγιστα και ελάχιστα και υπολογίζουμε την f στα σημεία αυτά.

Πρόκειται για τα σημεία όπου $f_x = f_y = 0$ ή όπου τουλάχιστον μία εκ των f_x και f_y δεν υπάρχει (κρίσιμα σημεία της f).

Βήμα 2°: Καταγράφουμε τα συνοριακά σημεία του R όπου η f έχει τοπικά μέγιστα και ελάχιστα και υπολογίζουμε την f στα σημεία αυτά.

Βήμα 3°: Από τα σημεία που καταγράψαμε, απομονώνουμε εκείνα που αντιστοιχούν στη μεγαλύτερη και στη μικρότερη τιμή της f. Στα σημεία αυτά η f εμφανίζει ολικά ακρότατα στο R. Εφόσον ένα ολικό ακρότατο είναι και τοπικό ακρότατο, το μόνο που έχουμε να κάνουμε είναι να εντοπίσουμε από τον κατάλογο των σημείων που κάναμε στα βήματα 1 και 2, τη μέγιστη και την ελάχιστη τιμή της f.

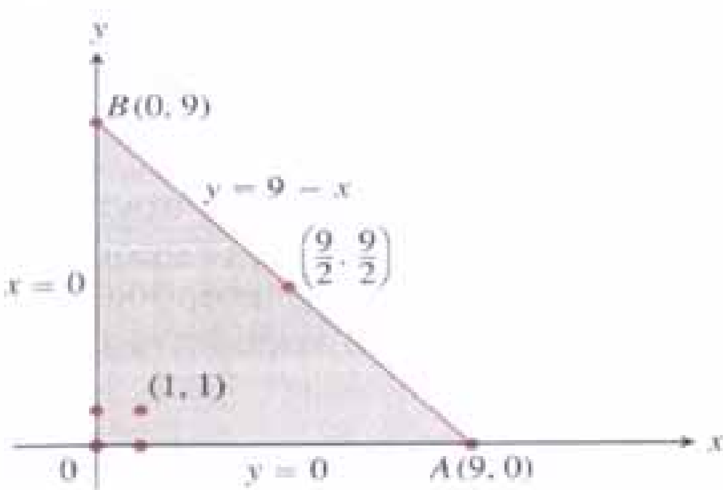
Εύρεση ολικών ακρότατων

Βρείτε το ολικό μέγιστο και το ολικό ελάχιστο της

$f(x,y) = 2 + 2x + 2y - x^2 - y^2$

στο τριγωνικό χωρίο του πρώτου τεταρτημορίου που περικλείεται από τις ευθείες $x = 0$, $y = 0$ και $y = 9 - x$

ΛΥΣΗ: Εφόσον η f είναι διαφορίσιμη, τα μόνα υποψήφια σημεία είναι εκείνα στο εσωτερικό του τριγώνου στο Σχήμα 9 όπου $f_x = f_y = 0$, καθώς και τα συνοριακά σημεία.



Σχήμα 9

Εσωτερικά σημεία

Έχουμε

$$f_x = 2 - 2x = 0 \quad , \quad f_y = 2 - 2y = 0$$

οπότε προκύπτει ως υποψήφιο το σημείο $(x,y)=(1,1)$. Η τιμή της f εκεί είναι $f(1,1)=4$.

Συνοριακά σημεία

Ερευνούμε μία προς μία τις πλευρές του τριγώνου:

Ευθύγραμμο τμήμα OA, $y=0$. Η συνάρτηση

$f(x,y)=f(x,0)=2+2x-x^2$

Μπορεί τώρα να θεωρηθεί ως συνάρτηση του x ορισμένη στο κλειστό διάστημα $0 \le x \le 9$. Τα ακρότατα της συνάρτησης αυτής μπορεί να προκύψουν στα συνοριακά σημεία

$x=0$ όπου $f(0,0)=2$

$x=9$ όπου $f(9,0)=2+18-81=-61$

καθώς και στα εσωτερικά σημεία όπου $f'(x,0)=2-2x=0$. Το μόνο εσωτερικό σημείο όπου $f'(x,0)=0$ είναι το $x=1$, όπου

$$f(x,0)=f(1,0)=3$$

Ευθύγραμμο τμήμα OB, $x=0$. Η συνάρτηση παίρνει τη μορφή

$f(x,y)=f(0,y)=2=2y-y^2$

Λόγο της συμμετρίας της f στην εναλλαγή των x και y και λόγο της ανάλυσης που προηγήθηκε, γνωρίζουμε ότι υποψήφια σημεία ολικού ακρότατου είναι τα σημεία όπου

$$f(0,0)=2, \qquad f(0,9)=-61, \qquad f(0,1)=3$$

Εφόσον έχουμε ήδη ασχοληθεί με τα άκρα του τμήματος AB, το μόνο που μας μένει είναι να εξετάσουμε εσωτερικά σημεία του AB. Για $y=9-x$, παίρνουμε $f(x,y)=2+2x+2(9-x)-x^2-(9-x)^2=-61+18x-2x^2$.

Θέτοντας $f'(x,9-x)=18-4x=0$, παίρνουμε

$$x=\frac{18}{4}=\frac{9}{2}$$

Για την τιμή αυτή του x,

$$y=9-\frac{9}{2}=\frac{9}{2} \qquad \text{και} \qquad f(x,y)=f(\frac{9}{2},\frac{9}{2})=-\frac{41}{2}.$$

<u>Περίληψη</u>:

Παραθέτουμε όλες τις υποψήφιες τιμές ολικού ακρότατου: 4, 2, -61, 3, -(41/2). Η μεγαλύτερη τιμή από αυτές είναι η 4, την οποία η f παίρνει στο σημείο (1,1). Η ελάχιστη τιμή είναι -61, την οποία η f παίρνει στα σημεία (0,9) και (9,0).

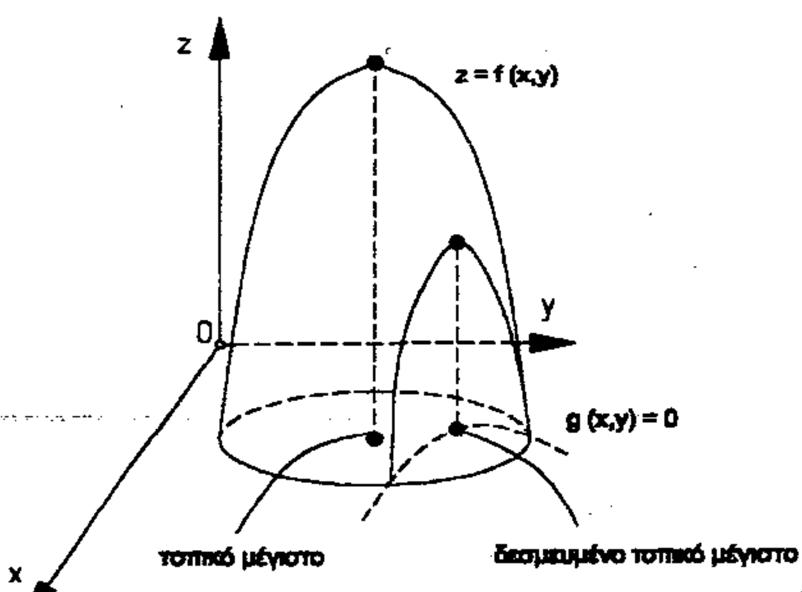
Η επίλυση προβλημάτων ακρότατων υπό συνθήκη απαιτεί συνήθως την εφαρμογή της μεθόδου των πολλαπλασιαστών Lagrange, μέθοδος η οποία παρατίθεται στη συνέχεια.

1.2. Πολλαπλασιαστές του Lagrange

Έστω Α ανοικτό υποσύνολο του R^n και $f: A \to R$. Έστω ότι ζητούνται τα ακρότατα σημεία της συνάρτησης f όταν οι μεταβλητές επαληθεύουν $m \le n-1$ εξισώσεις

$$
\begin{gathered}
g_1(x_1,\dots,x_n)=0, \\
g_2(x_1,\dots,x_n)=0, \\
\dots \\
\dots \\
\dots \\
g_m(x_1,\dots,x_n)=0
\end{gathered}
$$

Στην περίπτωση αυτή τα ακρότατα (αν υπάρχουν) καλούνται δεσμευμένα ακρότατα. Αυτά απεικονίζονται στο Σχήμα 10.



Σχήμα 10

1.2.1. Περιγραφή της μεθόδου των πολλαπλασιαστών Lagrange

Έστω Α ανοικτό υποσύνολο του R^n και έστω $f: A \to R$ συνάρτηση για την οποία υπάρχουν οι μερικές παράγωγοι

$$\frac{\partial f}{\partial x_i}, i = 1,2,...,n$$

και είναι συνεχείς. Έστω για τις συναρτήσεις $g_k : A \to R, k = 1,2,...,m$, όπου $m \leq n-1$ ότι υπάρχουν οι μερικές παράγωγοι

$$\frac{\partial g_k}{\partial x_i}, k = 1,2,...,m, \quad i = 1,2,...,n$$

και είναι συνεχείς συναρτήσεις.

Θεωρούμε τη συνάρτηση

$$G(x_1,...,x_n,\lambda_1,\lambda_2,...,\lambda_m) = f(x_1,...,x_n) + \sum_{k=1}^{m} \lambda_k g_k(x_1,...,x_n)$$

όπου $\lambda_1, \lambda_2,...,\lambda_m$ πραγματικοί αριθμοί. Η συνάρτηση G καλείται συνάρτηση του *Lagrange* ενώ οι πραγματικοί αριθμοί λ_k, $k = 1,2,...,m$ πολλαπλασιαστές του *Lagrange*.

Στη συνέχεια θεωρούμε το σύστημα:

$$
\begin{cases}
\dfrac{\partial G}{\partial x_1} = \dfrac{\partial f}{\partial x_1} + \sum_{k=1}^{m} \lambda_k \dfrac{\partial g_k}{\partial x_1} = 0 \\
\dfrac{\partial G}{\partial x_2} = \dfrac{\partial f}{\partial x_2} + \sum_{k=1}^{m} \lambda_k \dfrac{\partial g_k}{\partial x_2} = 0 \\
\dots \\
\dots \\
\dots \\
\dfrac{\partial G}{\partial x_n} = \dfrac{\partial f}{\partial x_n} + \sum_{k=1}^{m} \lambda_k \dfrac{\partial g_k}{\partial x_n} = 0 \\
g_1(x_1, \dots, x_n) = 0 \\
g_2(x_1, \dots, x_n) = 0 \\
\dots \\
\dots \\
\dots \\
g_m(x_1, \dots, x_n) = 0
\end{cases}
$$

Εάν το σημείο $\overline{\xi} = (\xi_1, \xi_2, \dots, \xi_\nu) \in A$ είναι το τοπικό ακρότατο για την συνάρτηση f που ικανοποιεί τις εξισώσεις:

$$g_\kappa(x_1, x_2, \dots, x_n) = 0, k = 1, 2, \dots, m$$

Τότε αποδεικνύεται ότι υπάρχουν πραγματικοί αριθμοί $\lambda_1, \lambda_2, \dots, \lambda_m$ έτσι ώστε το σημείο $\overline{k} = (\xi_1, \xi_2, \dots, \xi_n, \lambda_1, \lambda_2, \dots, \lambda_m)$ να είναι λύση του παραπάνω συστήματος. Έτσι στην πράξη προσπαθούμε να λύσουμε το παραπάνω σύστημα. Εάν η $\overline{k} = (\xi_1, \xi_2, \dots, \xi_n, \lambda_1, \lambda_2, \dots, \lambda_m)$ είναι μία λύση του συστήματος, τότε το σημείο $(\xi_1, \xi_2, \dots, \xi_n)$ ενδέχεται να είναι τοπικό ακρότατο της f που ικανοποιεί τις εξισώσεις

$$g_\kappa(x_1, x_2, \dots, x_n) = 0, k = 1, 2, \dots, m$$

Τώρα έστω:

$$\Delta_1(\bar{k})=\begin{vmatrix} \frac{\partial^2 G(\bar{k})}{\partial x_1^2} & \frac{\partial^2 G(\bar{k})}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 G(\bar{k})}{\partial x_1 \partial x_n} & \frac{\partial g_1(\bar{\xi})}{\partial x_1} & \cdots & \frac{\partial g_m(\bar{\xi})}{\partial x_1} \\ \frac{\partial^2 G(\bar{k})}{\partial x_2 \partial x_1} & \frac{\partial^2 G(\bar{k})}{\partial x_2^2} & \cdots & \frac{\partial^2 G(\bar{k})}{\partial x_2 \partial x_n} & \frac{\partial g_1(\bar{\xi})}{\partial x_2} & \cdots & \frac{\partial g_m(\bar{\xi})}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 G(\bar{k})}{\partial x_n \partial x_1} & \frac{\partial^2 G(\bar{k})}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 G(\bar{k})}{\partial x_n^2} & \frac{\partial g_1(\bar{\xi})}{\partial x_n} & \cdots & \frac{\partial g_m(\bar{\xi})}{\partial x_n} \\ \frac{\partial g_1(\bar{\xi})}{\partial x_1} & \cdots & \cdots & \frac{\partial g_1(\bar{\xi})}{\partial x_n} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_m(\bar{\xi})}{\partial x_1} & \cdots & \cdots & \frac{\partial g_m(\bar{\xi})}{\partial x_n} & 0 & \cdots & 0 \end{vmatrix}$$

και $\Delta_i(\bar{k}), i=0,1,2,\ldots,n-m$ η ελλάσονα ορίζουσα του στοιχείου α_{11} της ορίζουσας $\Delta_{i-1}(\bar{k})$.

Τότε έχουμε:

Εάν $\Delta_i(\bar{k})<0$, για κάθε $i=0,1,2,\ldots,n-m$ και m περιττός, τότε το σημείο $(\xi_1,\ldots,\xi_n)$ είναι σημείο δεσμευμένου τοπικού ελαχίστου.

Εάν $\Delta_i(\bar{k})>0$, για κάθε $i=0,1,2,\ldots,n-m$ και m άρτιος, τότε το σημείο $(\xi_1,\ldots,\xi_n)$ είναι σημείο δεσμευμένου τοπικού ελαχίστου.

Εάν $(-1)^i\Delta_i(\bar{k})>0$, για κάθε $i=0,1,2,\ldots,n-m$ και n περιττός, τότε το σημείο $(\xi_1,\ldots,\xi_n)$ είναι σημείο δεσμευμένου τοπικού μεγίστου.

Εάν $(-1)^i\Delta_i(\bar{k})<0$, για κάθε $i=0,1,2,\ldots,n-m$ και n άρτιος, τότε το σημείο $(\xi_1,\ldots,\xi_n)$ είναι σημείο δεσμευμένου τοπικού μεγίστου.

Η εύρεση τοπικών ακρότατων με την μέθοδο των πολλαπλασιαστών του *Lagrange* είναι αρκετά χρονοβόρα και επίπονη. Για το λόγο αυτό θα παρατεθούν μερικές ειδικές περιπτώσεις που συνήθως χρησιμοποιούνται σε πάρα πολλά προβλήματα και εφαρμογές.Οι παρακάτω περιπτώσεις αποτελούν απλή εφαρμογή της γενικής θεωρίας που αναφέρθηκε παραπάνω για τους πολλαπλασιαστές *Lagrange*.

Περίπτωση $1^η$:

Έστω A ανοικτό υποσύνολο του R^2 και $f : A \to R$ συνάρτηση για την οποία υπάρχουν οι μερικές παράγωγοι:

$$\frac{\partial f}{\partial x_i}, i = 1,2$$

και είναι συνεχείς συναρτήσεις. Έστω η συνάρτηση $g : A \to R$ για την οποία υπάρχουν οι μερικές παράγωγοι:

$$\frac{\partial g}{\partial x_i}, i = 1,2$$

και είναι συνεχείς συναρτήσεις. Ζητούνται τα ακρότατα σημεία της συνάρτησης f με τον περιορισμό $g(x_1, x_2) = 0$.

Θεωρούμε τη συνάρτηση

$$G(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2)$$

και λύνουμε το σύστημα

$$\begin{cases} \dfrac{\partial G(x_1, x_2, \lambda)}{\partial x_1} = \dfrac{\partial f(x_1, x_2)}{\partial x_1} + \lambda \dfrac{\partial g(x_1, x_2)}{\partial x_1} = 0 \\ \dfrac{\partial G(x_1, x_2, \lambda)}{\partial x_2} = \dfrac{\partial f(x_1, x_2)}{\partial x_2} + \lambda \dfrac{\partial g(x_1, x_2)}{\partial x_2} = 0 \\ g(x_1, x_2) = 0 \end{cases}$$

Έστω $\bar{k} = (\xi_1, \xi_2, \lambda)$ λύση του συστήματος και έστω

$$\Delta_1(\bar{k}) = \begin{vmatrix} \dfrac{\partial^2 G(\bar{k})}{\partial x_1^2} & \dfrac{\partial^2 G(\bar{k})}{\partial x_1 \partial x_2} & \dfrac{\partial g(\bar{\xi})}{\partial x_1} \\ \dfrac{\partial^2 G(\bar{k})}{\partial x_2 \partial x_1} & \dfrac{\partial^2 G(\bar{k})}{\partial x_2^2} & \dfrac{\partial g(\bar{\xi})}{\partial x_2} \\ \dfrac{\partial g(\bar{\xi})}{\partial x_1} & \dfrac{\partial g(\bar{\xi})}{\partial x_2} & 0 \end{vmatrix}$$

Τότε έχουμε:

Εάν $\Delta_1(\bar{k})<0$, τότε το σημείο $\bar{\xi}=(\xi_1,\xi_2)$ είναι σημείο δεσμευμένου τοπικού ελαχίστου.

Εάν $(-1)^1\Delta_1(\bar{k})<0$ ή $\Delta_1(\bar{k})>0$, τότε το σημείο $\bar{\xi}=(\xi_1,\xi_2)$ είναι σημείο δεσμευμένου τοπικού μεγίστου.

Περίπτωση 2η :

Έστω A ανοικτό υποσύνολο του R^3 και $f:A\to R$ συνάρτηση για την οποία υπάρχουν οι μερικές παράγωγοι:

$$\frac{\partial f}{\partial x_i},\ i=1,2,3$$

και είναι συνεχείς συναρτήσεις. Έστω η συνάρτηση $g:A\to R$ για την οποία υπάρχουν οι μερικές παράγωγοι

$$\frac{\partial g}{\partial x_i},\ i=1,2,3$$

και είναι συνεχείς συναρτήσεις. Ζητούνται τα ακρότατα σημεία της συνάρτησης f με τον περιορισμό $g(x_1,x_2,x_3)=0$.

Θεωρούμε τη συνάρτηση:

$$G(x_1,x_2,x_3,\lambda)=f(x_1,x_2,x_3)+\lambda g(x_1,x_2,x_3)$$

και λύνουμε το σύστημα:

$$\begin{cases}\dfrac{\partial G(x_1,x_2,x_3,\lambda)}{\partial x_1}=\dfrac{\partial f(x_1,x_2,x_3)}{\partial x_1}+\lambda\dfrac{\partial g(x_1,x_2,x_3)}{\partial x_1}=0\\ \dfrac{\partial G(x_1,x_2,x_3,\lambda)}{\partial x_2}=\dfrac{\partial f(x_1,x_2,x_3)}{\partial x_2}+\lambda\dfrac{\partial g(x_1,x_2,x_3)}{\partial x_2}=0\\ \dfrac{\partial G(x_1,x_2,x_3,\lambda)}{\partial x_3}=\dfrac{\partial f(x_1,x_2,x_3)}{\partial x_3}+\lambda\dfrac{\partial g(x_1,x_2,x_3)}{\partial x_3}=0\\ g(x_1,x_2,x_3)=0\end{cases}$$

Έστω $\bar{k}=(\xi_1,\xi_2,\xi_3,\lambda)$ λύση του συστήματος και έστω:

$$\Delta_1\left(\overline{k}\right)=\begin{vmatrix} \dfrac{\partial^2 G\left(\overline{k}\right)}{\partial x_1^2} & \dfrac{\partial^2 G\left(\overline{k}\right)}{\partial x_1 \partial x_2} & \dfrac{\partial^2 G\left(\overline{k}\right)}{\partial x_1 \partial x_3} & \dfrac{\partial g\left(\overline{\xi}\right)}{\partial x_1} \\ \dfrac{\partial^2 G\left(\overline{k}\right)}{\partial x_2 \partial x_1} & \dfrac{\partial^2 G\left(\overline{k}\right)}{\partial x_2^2} & \dfrac{\partial^2 G\left(\overline{k}\right)}{\partial x_2 \partial x_3} & \dfrac{\partial g\left(\overline{\xi}\right)}{\partial x_2} \\ \dfrac{\partial^2 G\left(\overline{k}\right)}{\partial x_3 \partial x_1} & \dfrac{\partial^2 G\left(\overline{k}\right)}{\partial x_3 \partial x_2} & \dfrac{\partial^2 G\left(\overline{k}\right)}{\partial x_3^2} & \dfrac{\partial g\left(\overline{\xi}\right)}{\partial x_3} \\ \dfrac{\partial g\left(\overline{\xi}\right)}{\partial x_1} & \dfrac{\partial g\left(\overline{\xi}\right)}{\partial x_2} & \dfrac{\partial g\left(\overline{\xi}\right)}{\partial x_3} & 0 \end{vmatrix}$$

$$\Delta_1\left(\overline{k}\right)=\begin{vmatrix} \dfrac{\partial^2 G\left(\overline{k}\right)}{\partial x_2^2} & \dfrac{\partial^2 G\left(\overline{k}\right)}{\partial x_2 \partial x_3} & \dfrac{\partial g\left(\overline{\xi}\right)}{\partial x_2} \\ \dfrac{\partial^2 G\left(\overline{k}\right)}{\partial x_3 \partial x_2} & \dfrac{\partial^2 G\left(\overline{k}\right)}{\partial x_3^2} & \dfrac{\partial g\left(\overline{\xi}\right)}{\partial x_3} \\ \dfrac{\partial g\left(\overline{\xi}\right)}{\partial x_2} & \dfrac{\partial g\left(\overline{\xi}\right)}{\partial x_3} & 0 \end{vmatrix}$$

Τότε έχουμε:

Εάν $\Delta_i\left(\overline{k}\right)<0,\ i=1,2$, τότε το σημείο (ξ_1,ξ_2,ξ_3) είναι σημείο δεσμευμένου τοπικού ελαχίστου.

Εάν $(-1)^i\Delta_i\left(\overline{k}\right)>0$, τότε το σημείο (ξ_1,ξ_2,ξ_3) είναι σημείο δεσμευμένου τοπικού μεγίστου.

<u>Περίπτωση 3^η :</u>

Έστω A ανοικτό υποσύνολο του R^3 και $f:A\rightarrow R$ συνάρτηση για την οποία υπάρχουν οι μερικές παράγωγοι:

$$\frac{\partial f}{\partial x_i}, i=1,2,3$$

και είναι συνεχείς συναρτήσεις. Έστω οι συναρτήσεις $g_j:A\rightarrow R,\ j=1,2$ για την οποία υπάρχουν οι μερικές παράγωγοι

$$\frac{\partial g_j}{\partial x_i}, i=1,2,3 \text{ και } j=1,2$$

και είναι συνεχείς συναρτήσεις. Ζητούνται τα ακρότατα σημεία της συνάρτησης f με τους περιορισμούς:

$$\begin{cases} g_1(x_1,x_2,x_3)=0 \\ g_2(x_1,x_2,x_3)=0 \end{cases}$$

Θεωρούμε τη συνάρτηση:

$$G(x_1,x_2,x_3,\lambda_1,\lambda_2)=f(x_1,x_2,x_3)+\lambda_1 g_1(x_1,x_2,x_3)+\lambda_2 g_2(x_1,x_2,x_3)$$

και λύνουμε το σύστημα:

$$\begin{cases} \dfrac{\partial G(x_1,x_2,x_3,\lambda_1,\lambda_2)}{\partial x_1}=\dfrac{\partial f(x_1,x_2,x_3)}{\partial x_1}+\lambda_1\dfrac{\partial g_1(x_1,x_2,x_3)}{\partial x_1}+\lambda_2\dfrac{\partial g_2(x_1,x_2,x_3)}{\partial x_1}=0 \\ \dfrac{\partial G(x_1,x_2,x_3,\lambda_1,\lambda_2)}{\partial x_2}=\dfrac{\partial f(x_1,x_2,x_3)}{\partial x_2}+\lambda_1\dfrac{\partial g_1(x_1,x_2,x_3)}{\partial x_2}+\lambda_2\dfrac{\partial g_2(x_1,x_2,x_3)}{\partial x_2}=0 \\ \dfrac{\partial G(x_1,x_2,x_3,\lambda_1,\lambda_2)}{\partial x_3}=\dfrac{\partial f(x_1,x_2,x_3)}{\partial x_3}+\lambda_1\dfrac{\partial g_1(x_1,x_2,x_3)}{\partial x_3}+\lambda_2\dfrac{\partial g_2(x_1,x_2,x_3)}{\partial x_3}=0 \\ g(x_1,x_2,x_3)=0 \end{cases}$$

Έστω $\overline{k}=(\xi_1,\xi_2,\xi_3,\lambda_1,\lambda_2)$ λύση του συστήματος και έστω:

$$\Delta_1(\overline{k})=\begin{vmatrix} \dfrac{\partial^2 G(\overline{k})}{\partial x_1^2} & \dfrac{\partial^2 G(\overline{k})}{\partial x_1 \partial x_2} & \dfrac{\partial^2 G(\overline{k})}{\partial x_1 \partial x_3} & \dfrac{\partial g_1(\overline{\xi})}{\partial x_1} & \dfrac{\partial g_2(\overline{\xi})}{\partial x_1} \\ \dfrac{\partial^2 G(\overline{k})}{\partial x_2 \partial x_1} & \dfrac{\partial^2 G(\overline{k})}{\partial x_2^2} & \dfrac{\partial^2 G(\overline{k})}{\partial x_2 \partial x_3} & \dfrac{\partial g_1(\overline{\xi})}{\partial x_2} & \dfrac{\partial g_2(\overline{\xi})}{\partial x_2} \\ \dfrac{\partial^2 G(\overline{k})}{\partial x_3 \partial x_1} & \dfrac{\partial^2 G(\overline{k})}{\partial x_3 \partial x_2} & \dfrac{\partial^2 G(\overline{k})}{\partial x_3^2} & \dfrac{\partial g_1(\overline{\xi})}{\partial x_3} & \dfrac{\partial g_2(\overline{\xi})}{\partial x_3} \\ \dfrac{\partial g_1(\overline{\xi})}{\partial x_1} & \dfrac{\partial g_1(\overline{\xi})}{\partial x_2} & \dfrac{\partial g_1(\overline{\xi})}{\partial x_3} & 0 & 0 \\ \dfrac{\partial g_2(\overline{\xi})}{\partial x_1} & \dfrac{\partial g_2(\overline{\xi})}{\partial x_2} & \dfrac{\partial g_2(\overline{\xi})}{\partial x_3} & 0 & 0 \end{vmatrix}$$

Τότε έχουμε:

Εάν $\Delta_1(\overline{k})>0$, τότε το σημείο (ξ_1,ξ_2,ξ_3) είναι σημείο δεσμευμένου τοπικού ελαχίστου.

Εάν $\Delta_1(\overline{k})<0$, τότε το σημείο (ξ_1,ξ_2,ξ_3) είναι σημείο δεσμευμένου τοπικού μεγίστου.

Στο σημείο αυτό παρουσιάζονται δύο παραδείγματα της χρήσης των πολλαπλασιαστών *Lagrange* για την επίλυση προβλημάτων.

<u>ΠΑΡΑΔΕΙΓΜΑ 1°</u>:

Να υπολογιστούν τα τοπικά ακρότατα της συνάρτησης $f(x,y)=x^2-3y^2$ με τον περιορισμό $x+2y-1=0$.

ΛΥΣΗ:

Θεωρούμε τη συνάρτηση:

$$G(x,y)=x^2-3y^2+\lambda(x+2y-1)$$

και σχηματίζουμε το παρακάτω σύστημα:

$$\begin{cases} \dfrac{\partial G(x,y)}{\partial x}=2x+\lambda=0 \\ \dfrac{\partial G(x,y)}{\partial y}=-6y+2\lambda=0 \\ x+2y-1=0 \end{cases} \quad \text{ή} \quad \begin{cases} x=-3 \\ y=2 \\ \lambda=6 \end{cases}$$

Οπότε πιθανό δεσμευμένο ακρότατο είναι το σημείο $(-3,2)$. Επίσης, έχουμε:

$$\frac{\partial^2 G(x,y)}{\partial x^2}=2,\quad \frac{\partial^2 G(x,y)}{\partial y^2}=-6,$$

$$\frac{\partial^2 G(x,y)}{\partial x\partial y}=\frac{\partial^2 G(x,y)}{\partial y\partial x}=0.$$

Τώρα, έστω $g(x,y)=x+2y-1$. Τότε επιπλέων έχουμε:

$$\frac{\partial g(x,y)}{\partial x}=1,\quad \frac{\partial g(x,y)}{\partial y}=2.$$

Συνεπώς, έχουμε:

$$\Delta_1(-3,2,6)=\begin{vmatrix}2 & 0 & 1\\ 0 & -6 & 2\\ 1 & 2 & 0\end{vmatrix}=-2<0 \quad .$$

Οπότε, το σημείο $(-3,2)$ είναι σημείο δεσμευμένου τοπικού ελαχίστου με τιμή $f(-3,2)=-3$.

ΠΑΡΑΔΕΙΓΜΑ 2°:

Να αποδειχθεί ότι η συνάρτηση $f(x_1,x_2,x_3)=x_1x_2x_3$ παρουσιάζει στο σημείο $\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)$ δεσμευμένο τοπικό μέγιστο που ικανοποιεί την εξίσωση $x_1+x_2+x_3-1=0$.

ΛΥΣΗ:

Θεωρούμε τη συνάρτηση

$$G(x_1,x_2,x_3,\lambda)=x_1x_2x_3+\lambda(x_1+x_2+x_3-1)$$

και σχηματίζουμε το παρακάτω σύστημα:

$$\begin{cases}\dfrac{\partial G(x_1,x_2,x_3,\lambda)}{\partial x_1}=x_2x_3+\lambda=0\\[2mm] \dfrac{\partial G(x_1,x_2,x_3,\lambda)}{\partial x_2}=x_1x_3+\lambda=0\\[2mm] \dfrac{\partial G(x_1,x_2,x_3,\lambda)}{\partial x_3}=x_2x_1+\lambda=0\\[2mm] x_1+x_2+x_3-1=0\end{cases}$$

Επειδή το σημείο $\left(\frac{1}{3},\frac{1}{3},\frac{1}{3},-\frac{1}{9}\right)$ είναι λύση του συστήματος, το σημείο $\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)$ είναι πιθανό δεσμευμένο ακρότατο.

Επίσης, έχουμε:

$$\frac{\partial^2 G(x_1,x_2,x_3)}{\partial x_1^2}=\frac{\partial^2 G(x_1,x_2,x_3)}{\partial x_2^2}=\frac{\partial^2 G(x_1,x_2,x_3)}{\partial x_3^2}=0,$$

$$\frac{\partial^2 G(x_1,x_2,x_3)}{\partial x_1\partial x_2}=\frac{\partial^2 G(x_1,x_2,x_3)}{\partial x_2\partial x_1}=x_3,$$

$$\frac{\partial^2 G(x_1,x_2,x_3)}{\partial x_1 \partial x_3} = \frac{\partial^2 G(x_1,x_2,x_3)}{\partial x_3 \partial x_1} = x_2,$$

$$\frac{\partial^2 G(x_1,x_2,x_3)}{\partial x_2 \partial x_3} = \frac{\partial^2 G(x_1,x_2,x_3)}{\partial x_3 \partial x_2} = x_1.$$

Έστω $g(x_1,x_2,x_3) = x_1 + x_2 + x_3 - 1$. Τότε επιπλέον έχουμε:

$$\frac{\partial g(x_1,x_2,x_3)}{\partial x_1} = \frac{\partial g(x_1,x_2,x_3)}{\partial x_2} = \frac{\partial g(x_1,x_2,x_3)}{\partial x_3} = 1.$$

Συνεπώς:

$$\Delta_1\left(\frac{1}{3},\frac{1}{3},\frac{1}{3},-\frac{1}{9}\right) = \begin{vmatrix} 0 & \frac{1}{3} & \frac{1}{3} & 1 \\ \frac{1}{3} & 0 & \frac{1}{3} & 1 \\ \frac{1}{3} & \frac{1}{3} & 0 & 1 \\ 1 & 1 & 1 & 1 \end{vmatrix} = -\frac{1}{3} < 0,$$

$$\Delta_2\left(\frac{1}{3},\frac{1}{3},\frac{1}{3},-\frac{1}{9}\right) = \begin{vmatrix} 0 & \frac{1}{3} & 1 \\ \frac{1}{3} & 0 & 1 \\ 1 & 1 & 0 \end{vmatrix} = \frac{2}{3} > 0,$$

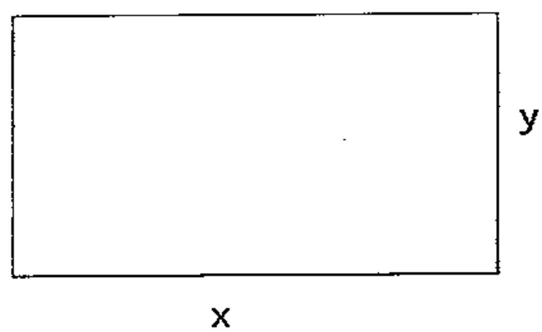
Οπότε η συνάρτηση f παρουσιάζει στο σημείο $\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)$ δεσμευμένο τοπικό μέγιστο που είναι $f\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right) = \frac{1}{27}$.

Ύστερα από την περιγραφή της μεθόδου των πολλαπλασιαστών Lagrange και των παραδειγμάτων που παρουσιάστηκαν για την καλύτερη κατανόησή της, θα παρουσιαστούν άλλα δύο παραδείγματα που θα εμφανίζουν την εφαρμογή της στην οικονομική θεωρία και πως η μέθοδος αυτή βοηθάει στην επίλυση προβλημάτων οικονομικής φύσεως.

ΠΑΡΑΔΕΙΓΜΑ $1^{\text{ο}}$:

Να βρεθεί το ορθογώνιο παραλληλόγραμμο με περίμετρο 4 μονάδες μήκος και μέγιστο εμβαδόν.

ΛΥΣΗ :



Αν x, y είναι το μήκος και το πλάτος του ορθογωνίου τότε το πρόβλημα γράφεται ως :

$$\begin{cases} \max\limits_{(x,y)} Q = xy \\ x + y = 4 \Leftrightarrow g(x,y) = 0 \ \text{,όπου } g(x,y) = 4 - 2x - 2y \end{cases}$$

Βήμα $1^{\text{ο}}$: Δημιουργούμε τη συνάρτηση Lagrange

$$G(x,y,\lambda) = xy - \lambda(4 - 2x - 2y)$$

όπου λ ο πολλαπλασιαστής Lagrange.

Βήμα $2^{\text{ο}}$: Ορίζουμε τις συνθήκες $1^{\text{ης}}$ τάξης

Είναι

$$\begin{cases} \dfrac{\partial G}{\partial x} = y - 2\lambda = 0 \\ \dfrac{\partial G}{\partial y} = x - 2\lambda = 0 \\ g(x,y) = 4 - 2x - 2y = 0 \end{cases}$$

Η λύση του συστήματος είναι:

$$x=1,\ y=1,\ \lambda=\frac{1}{2}$$

Η λύση βρίσκεται ως εξής :

$$\begin{cases} y-2\lambda=0 \\ x-2\lambda=0 \\ 4-2x-2y=0 \end{cases} \Leftrightarrow \begin{cases} y=2\lambda \\ x=2\lambda \\ 4-2x-2y=0 \end{cases} \Leftrightarrow \begin{cases} y=2\lambda \\ x=2\lambda \\ 4-4\lambda-4\lambda=0 \end{cases} \Leftrightarrow \begin{cases} y=2\lambda \\ x=2\lambda \\ \lambda=\frac{1}{2} \end{cases} \Leftrightarrow \begin{cases} y=1 \\ x=1 \\ \lambda=\frac{1}{2} \end{cases}$$

<u>Βήμα 3^{o}</u>: Ελέγχουμε τις συνθήκες $2^{ης}$ τάξης. Επειδή η συνάρτηση που θέλουμε να μεγιστοποιήσουμε είναι $n=2$ μεταβλητών και έχουμε $m=1$ περιορισμό ελέγχουμε μόνο την $\Delta\left(1,1,\frac{1}{2}\right)$.

Είναι

$$\Delta\left(1,1,\frac{1}{2}\right)=\begin{vmatrix} \frac{\partial^2 G}{\partial x^2} & \frac{\partial^2 G}{\partial x\partial y} & \frac{\partial g}{\partial x} \\ \frac{\partial^2 G}{\partial y\partial x} & \frac{\partial^2 G}{\partial y^2} & \frac{\partial g}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} & 0 \end{vmatrix}=\begin{vmatrix} 0 & 1 & -2 \\ 1 & 0 & -2 \\ -2 & -2 & 0 \end{vmatrix}=-\begin{vmatrix} 1 & -2 \\ -2 & 0 \end{vmatrix}+(-2)\begin{vmatrix} 1 & 0 \\ -2 & -2 \end{vmatrix}=-(-4)+(-2)(-2)=4+4=8$$

Επειδή $(-1)\Delta_1\left(1,1,\frac{1}{2}\right)<0$ και $n=2$ άρτιος, το σημείο $(1,1)^T$ είναι σημείο <u>μέγιστο</u>.

<u>Σημείωση:</u> Είναι γνωστό ότι η λύση του παραπάνω προβλήματος είναι το τετράγωνο με $x=y=1$. Με "στοιχειώδη" άλγεβρα αυτό δείχνεται ως εξής. Αν το μέγιστο εμβαδό είναι $E=xy$, τότε τα x,y είναι οι ρίζες του τριωνύμου $\omega^2-2\omega+E=0$ (άθροισμα ριζών: $x+y=2$, γινόμενο ριζών: $xy=E$). Για να έχει λύσεις πρέπει:

$0\le\Delta=4-4E\Leftrightarrow E\le 1$

Άρα $xy=1$ και $x=1,y=1$

ΠΑΡΑΔΕΙΓΜΑ 2^o :

Έστω ότι υπάρχει ένας καταναλωτής που έχει στη διάθεσή του $M = 24$ χρηματικές μονάδες τις οποίες σκοπεύει να καταναλώσει για την προμήθεια τριών αγαθών των οποίων οι τιμές είναι $P_x = 2$, $P_y = 1$ και $P_z = 4$.

Έστω επίσης ότι εκτιμήθηκε η συνάρτηση ωφέλειάς του που είναι $U(x, y, z) = xyz$. Ζητείται να υπολογιστούν τα x, y και z στη θέση όπου μεγιστοποιεί την ωφέλειά του. Να υπολογιστεί η μέγιστη αυτή ωφέλεια.

ΛΥΣΗ :

Πρέπει να λυθεί το πρόβλημα :

$$\begin{cases} \max\limits_{(x,y,z)} U = xyz \\ 2x + y + 4z = 24 \Leftrightarrow g(x,y,z) = 0 \text{ όπου } g(x,y,z) = 24 - 2x - y - 4z \end{cases}$$

Βήμα 1^o : Δημιουργούμε τη συνάρτηση Lagrange

$$L = xyz - \lambda(24 - 2x - y - 4z)$$

όπου λ ο πολλαπλασιαστής Lagrange.

Βήμα 2^o : Ορίζουμε τις συνθήκες $1^{ης}$ τάξης

Είναι

$$\begin{cases} \dfrac{\partial L}{\partial x} = yz - 2\lambda = 0 \\ \dfrac{\partial L}{\partial y} = xz - \lambda = 0 \\ \dfrac{\partial L}{\partial z} = xy - 4\lambda = 0 \\ 2x + y + 4z = 24 \end{cases}$$

Η λύση του συστήματος είναι:

$$\lambda = \frac{xy}{4} = xz = \frac{yz}{2}$$

Άρα $\frac{yz}{2} = xz$.

Αν $\underline{z \neq 0}$ τότε $y = 2x$.

Άρα η $xz = \frac{xy}{4}$ γίνεται $xz = \frac{2x^2}{4}$.

Αν $\underline{x \neq 0}$ τότε έχουμε από την τελευταία $2z = x$ και από την

$2x + y + 4z = 24$ έχουμε

$$4z + 4z + 4z = 24$$

Άρα

$$x = 4,\ y = 8,\ z = 2,\ \lambda = 8$$

Αν $\underline{x = 0}$ τότε και $y = 0$

Άρα

$$x = 0,\ y = 0,\ z = 6,\ \lambda = 0$$

Αν $\underline{z = 0}$ τότε $\lambda = 0$ και $\frac{xy}{4} = \lambda = 0$

Άρα αν $x = 0$ τότε $y = 24$

αν $y = 0$ τότε $x = 12$

Τελικά οι λύσεις του συστήματος είναι:

$$(x, y, z, \lambda) = (4, 8, 2, 8)$$

$$\text{ή } (x, y, z, \lambda) = (0, 0, 6, 0)$$

$$\text{ή } (x, y, z, \lambda) = (0, 24, 0, 0)$$

$$\text{ή } (x, y, z, \lambda) = (12, 0, 0, 0).$$

Οι λύσεις όπου κάποια από τις τιμές των x, y, z είναι 0 προφανώς δεν μεγιστοποιούν την ωφέλεια. Συνεπώς θα γίνει έλεγχος στις συνθήκες $2^{ης}$ τάξης μόνο για τη λύση $(x, y, z, \lambda) = (4, 8, 2, 8)$.

<u>Βήμα 3^{ο}</u>: Έλεγχος συνθηκών $2^{ης}$ τάξης :

Επειδή η συνάρτηση που θέλουμε να μεγιστοποιήσουμε είναι $n=3$ μεταβλητών και έχουμε $m=1$ περιορισμό, θα πρέπει να γίνει έλεγχος της

$$\Delta_1(4,8,2,8) \text{ και } \Delta_2(4,8,2,8)$$

$$\Delta_1(4,8,2,8)=\begin{vmatrix} \frac{\partial^2 L}{\partial x^2} & \frac{\partial^2 L}{\partial x\partial y} & \frac{\partial^2 L}{\partial x\partial z} & \frac{\partial g}{\partial x} \\ \frac{\partial^2 L}{\partial x\partial y} & \frac{\partial^2 L}{\partial y^2} & \frac{\partial^2 L}{\partial y\partial z} & \frac{\partial g}{\partial y} \\ \frac{\partial^2 L}{\partial z\partial x} & \frac{\partial^2 L}{\partial z\partial y} & \frac{\partial^2 L}{\partial z^2} & \frac{\partial g}{\partial z} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} & \frac{\partial g}{\partial z} & 0 \end{vmatrix} = \begin{vmatrix} 0 & 2 & 8 & -16 \\ 2 & 0 & 4 & -8 \\ 8 & 4 & 0 & -32 \\ -16 & -8 & -32 & 0 \end{vmatrix} =$$

(Από την $3^η$ στήλη αφαιρούμε το τετραπλάσιο της $2^{ης}$ και από την $4^η$ στήλη αφαιρούμε το οκταπλάσιο της $2^{ης}$)

$$= \begin{vmatrix} 0 & 2 & 0 & 0 \\ 2 & 0 & 4 & -8 \\ 8 & 4 & -16 & 0 \\ -16 & -8 & 0 & -64 \end{vmatrix} = -2\begin{vmatrix} 2 & 4 & -8 \\ 8 & -16 & 0 \\ -16 & 0 & -64 \end{vmatrix} =$$

(Από τη $2^η$ στήλη αφαιρούμε το διπλάσιο της $1^{ης}$ και στην $3^η$ στήλη προσθέτουμε το τετραπλάσιο της $1^{ης}$)

$$= -2\begin{vmatrix} 2 & 0 & 0 \\ 8 & -32 & 32 \\ -16 & 32 & -128 \end{vmatrix} = -4\begin{vmatrix} -32 & 32 \\ 32 & -128 \end{vmatrix} = -4(4096-1024) = -4*3072 = -12288$$

Είναι

$$\Delta_2(4,8,2,8) = \begin{vmatrix} 0 & 4 & -8 \\ 4 & 0 & -32 \\ -8 & -32 & 0 \end{vmatrix} =$$

(Στην $3^η$ γραμμή προσθέτουμε το διπλάσιο της $2^{ης}$)

$$= \begin{vmatrix} 0 & 4 & -8 \\ 4 & 0 & -32 \\ -8 & -32 & -64 \end{vmatrix} = -4\begin{vmatrix} 4 & -8 \\ -32 & -64 \end{vmatrix} = -4(-256-256) = 4*512 = 2048$$

Επειδή

$$(-1)\Delta_1(4,8,2,8) > 0$$
$$\Delta_2(4,8,2,8) > 0$$

και $n=3$ περιττός, το σημείο $(4,8,2,8)$ είναι σημείο μεγίστου.

Άρα

$$x=4, y=8, z=2 \text{ και } U=64$$

2. ΚΕΦΑΛΑΙΟ 2

2.1. Εισαγωγή Στα Μη Γραμμικά Προβλήματα

Πραγματοποιώντας μια ιστορική αναζήτηση της ανάπτυξης των μεθόδων βελτιστοποίησης συναρτήσεων πολλών μεταβλητών, παρατηρείται η σχεδόν ολική απουσία της, ως την εποχή του Β' Παγκοσμίου Πολέμου. Καταλυτικός παράγοντας στην ανάπτυξη αλγορίθμων, που παρείχαν προσεγγιστικές λύσεις ταχέως και με ακρίβεια, αποτέλεσε η δημιουργία και η χρήση ηλεκτρονικών υπολογιστών. Κατά συνέπεια, στις δεκαετίες του 1940 και 1950, παρατηρείται μια ραγδαίου ρυθμού ανάπτυξη των μεθόδων Μαθηματικού Προγραμματισμού, που οδήγησε στην μεταβολή και άλλων κλάδων μεθόδων βελτιστοποίησης όπως ο Δυναμικός Προγραμματισμός, ο Ακέραιος Προγραμματισμός, οι Μη γραμμικές Μέθοδοι κλπ. Απαραίτητος είναι ο διαχωρισμός του όρου Προγραμματισμός από την επικρατούσα ταύτισή του με τον προγραμματισμό των ηλεκτρονικών υπολογιστών, και η παράλληλη σύνδεσή του με την έννοια Βελτιστοποίηση.

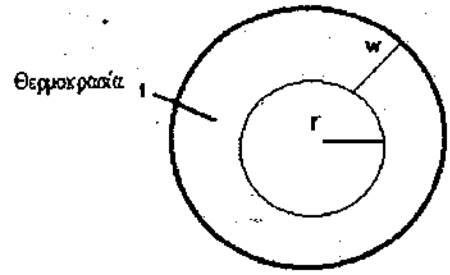
Στην πορεία θα τεθεί το γενικό πρόβλημα του Μη Γραμμικού Προγραμματισμού, δηλαδή το γενικό πρόβλημα βελτιστοποίησης μη γραμμικών προβλημάτων.

2.1.1. Δομή των προβλημάτων

Η δομή των προβλημάτων γίνεται πιο κατανοητή με τη βοήθεια των ακόλουθων παραδειγμάτων:

ΠΑΡΑΔΕΙΓΜΑ 1°:

Το ακόλουθο παράδειγμα είναι μια απλουστευμένη μορφή ενός προβλήματος που παρουσιάζεται στη μελέτη αντιδραστήρων πυρηνικής σύντηξης (Dennis & Schnabel 1983). Ένας αντιδραστήρας μπορεί να παρασταθεί όπως στο Σχήμα 11, δηλαδή από δύο ομόκεντρους κύκλους. Το πρόβλημα είναι να βρεθεί ο ιδανικός συνδυασμός εσωτερικής ακτίνας r, του πλάτους w και της θερμοκρασίας t, έτσι ώστε να οδηγηθούμε σε παραγωγή ενέργειας.



Σχήμα 11

Μετά από διάφορες εργαστηριακές μελέτες καθορίστηκε ότι το κόστος ανά μονάδα ενέργειας περιγράφεται από το ακόλουθο μοντέλο:

$$f(r,w,t)=10^6 t^2\left[c_1\left(r^2+\frac{2rw}{3}+\frac{w^2}{6}\right)+c_2\left(c_4+\frac{c_4^2}{r^2-c_4^2}\right)\right]+c_3\left(\frac{1}{2}+\frac{r}{w}\right)\left(10^3 t\right)^{-1.46}$$

στο οποίο οι c_1,c_2,c_3,c_4 είναι σταθερές. Οι μεταβλητές του προβλήματος είναι οι r,w,t ενώ οι c_1,c_2,c_3,c_4 είναι παράμετροι. Η λύση του πραγματικού φυσικού μη γραμμικού προβλήματος ανάγεται στον προσδιορισμό ενός ελαχίστου της πραγματικής συνάρτησης f η οποία εξαρτάται από την ακτίνα r, από το πλάτος w και από τη θερμοκρασία t. Αν το πρόβλημα έπρεπε να μελετηθεί πιο αυστηρά, (όπως γίνεται και στην πραγματικότητα), θα έπρεπε να θέτονται και κάποιοι περιορισμοί στις μεταβλητές, όπως για παράδειγμα ότι η ακτίνα και το πλάτος πρέπει να είναι θετικά. Το πρόβλημα στην πραγματική του μορφή είχε και πέντε περιορισμούς. Έτσι μπορεί να θεωρηθεί σαν ένα *πρόβλημα μη γραμμικού προγραμματισμού με περιορισμούς.*

ΠΑΡΑΔΕΙΓΜΑ 2°:

Ανάμεσα σε όλα τα τρίγωνα με βάση 8 εκατοστά και εμβαδόν τουλάχιστον 12 τετραγωνικά εκατοστά να βρεθεί αυτό με την ελάχιστη περίμετρο (McCormick 1982).

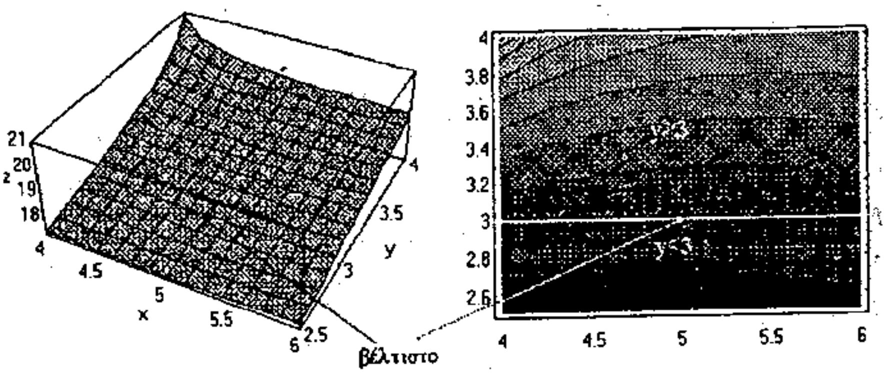
Η βάση είναι σταθερή 8 εκατοστά οπότε θα θέλαμε να γνωρίζουμε για παράδειγμα, τις τιμές από μία πλευρά και το ύψος. Έτσι αν η μία πλευρά είναι η x και το ύψος είναι το y, αυτές είναι οι μεταβλητές του προβλήματος και η βάση είναι η παράμετρος. Από τα άπειρα x και y με τα οποία παράγονται τρίγωνα με εμβαδόν μεγαλύτερο η ίσο των 12 τετραγωνικών εκατοστών, αναζητούμε τις τιμές x^* και y^* για τις οποίες η περίμετρος είναι ελάχιστη. Αφού $E = 8y/2 \geq 12$ έχουμε ότι $y \geq 3$. Είναι επίσης προφανές ότι $x \geq y$. Από το Πυθαγόρειο Θεώρημα (το τετράγωνο της υποτείνουσας είναι ίσο με το άθροισμα των τετραγώνων των δύο απέναντι πλευρών), προκύπτει το παρακάτω πρόβλημα:

$$\underset{(x,y)}{Minimize}\ f(x,y) = 8 + x + \left\{ y^2 + \left[8 - \left(x^2 - y^2\right)^{1/2} \right]^2 \right\}^{1/2}$$

με περιορισμούς:

$$y \geq 3 \text{ και } x \geq y.$$

Η λύση του προβλήματος είναι το σημείο $(5,3)$ που δίνει τιμή για τη συνάρτηση ίση με 18 (Σχήμα 12 την ανωτέρω συνάρτηση στον R^3)



Σχήμα 12

Όπως είναι φανερό από τα παραπάνω παραδείγματα, η μελέτη των μη γραμμικών προβλημάτων και η κατασκευή ενός αντίστοιχου μαθηματικού μοντέλου βελτιστοποίησης, εξαρτάται άμεσα από την αναγκαιότητα μεταφοράς συναρτησιακών ή ποιοτικών σχέσεων που ενυπάρχουν μεταξύ μετρήσιμων ή ποιοτικών χαρακτηριστικών, σε μαθηματικά σύμβολα. Αυτή η διαδικασία ονομάζεται μοντελοποίηση. Κάθε μαθηματικό μοντέλο σαν μεταφορά πραγματικών προβλημάτων σε μαθηματικούς τύπους, υπόκειται σε σφάλματα τα οποία προκύπτουν τουλάχιστον από τις παραδοχές που πρέπει να γίνουν. Μία από τις πιο βασικές ικανότητες που πρέπει να αναπτύξει ένας επιχειρησιακός ερευνητής είναι η εύστοχη απόρριψη των ήσσονος σημασίας στοιχείων ενός προβλήματος και ταυτοχρόνως η ανάδειξη των σημαντικών στοιχείων και η προσκόλλησή τους στο μοντέλο.

Στο σημείο αυτό παρατίθεται η γενική μορφή ενός μοντέλου μη γραμμικού προγραμματισμού το οποίο διατυπώνεται με τον παρακάτω τρόπο

Γενική Μορφή Μη Γραμμικού Προβλήματος

Έστω $x \in R^n$ *για κάποια διάσταση* n, $x = (x_1, x_2, ..., x_n)^T$ *και έστω* $f(x)$ *η αντικειμενική συνάρτηση η οποία υποθέτουμε ότι είναι συνεχείς. Επίσης, συμβολίζουμε με* $h_1(x), h_2(x), h_3(x), ..., h_m(x)$ *τους περιορισμούς που εκφράζονται με εξισώσεις και με* $a_{m+1}(x), a_{m+2}(x), a_{m+3}(x), ..., a_k(x)$ *τους περιορισμούς που εκφράζονται με ανισώσεις για τις μεταβλητές* $x_1, x_2, ..., x_n$. *Τότε, το γενικό πρόβλημα του μη γραμμικού προγραμματισμού μπορεί να γραφτεί στη ακόλουθη μορφή:*

$$Minimize\ f(x), x \in R^n$$

με τρόπο ώστε να ικανοποιούνται οι m *γραμμικές ή μη γραμμικές εξισώσεις*

$$h_1(x) = 0, h_2(x) = 0, ..., h_m(x) = 0$$

και οι $k - m$ *γραμμικές ή μη γραμμικές ανισώσεις*

$$a_{m+1}(x) \geq 0, a_{m+2}(x) \geq 0, ..., a_k(x) \geq 0$$

Το διάνυσμα $x^* = (x_1^*, x_2^*, ..., x_n^*)^T$ που ικανοποιεί τις σχέσεις της Γενικής Μορφής Μη Γραμμικού Προβλήματος καλείται **βέλτιστο σημείο** (optimal point) και η αντίστοιχη τιμή της συνάρτησης δηλαδή το ελάχιστο (ή το μέγιστο) καλείται **βέλτιστη τιμή** (optimal value) της συνάρτησης. Και τα δύο μαζί αποτελούν μία **βέλτιστη λύση** (optimal solution).

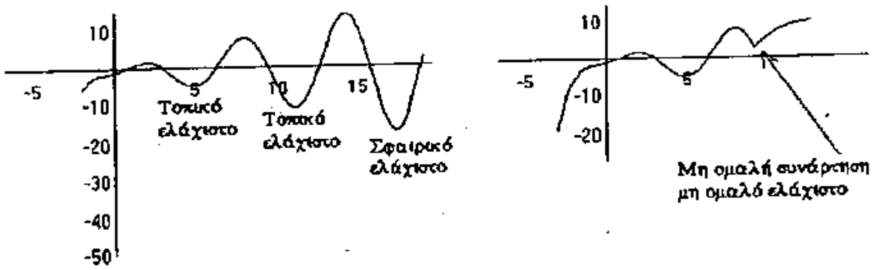
ΠΑΡΑΔΕΙΓΜΑ :

$$Minimize\ f(x) = x_1^3 + x_2^2 + 2x_1x_2 - 3x_2$$

με τρόπο ώστε να ικανοποιούνται οι ακόλουθοι περιορισμοί:

$$h_1(x) = x_1^2 - 2x_2^2 + 7 = 0$$
$$a_2(x) = 0,7x_1^2 + 4x_2^3 - x_1 + 10 \geq 0$$
$$a_3(x) = x_1 \geq 0$$

Γενικά το πρόβλημα δεν είναι δυνατόν να λυθεί με τις κλασικές μεθόδους της ανάλυσης, (ανεξαρτήτως αν τα παραδείγματα που θα χρησιμοποιηθούν, σε αρκετές περιπτώσεις είναι φανερό ότι λύνονται εύκολα με τις μεθόδους αυτές), γιατί το σημείο x^* στο οποίο επιτυγχάνεται το μέγιστο ή ελάχιστο της συνάρτησης $f(x)$, μπορεί να βρίσκεται στο εσωτερικό ή στην τομή των επιφανειών που εκφράζουν οι περιορισμοί. Βεβαίως, υπάρχει από την αρχή και το πρόβλημα της πιθανής μη ύπαρξης βέλτιστου σημείου όταν π.χ. η αντικειμενική συνάρτηση είναι μη φραγμένη κάτω ή αυστηρά φθίνουσα. Αλλά και φραγμένη να είναι μπορεί πάλι να μην έχει ένα ελάχιστο (π.χ. $f(x) = e^{-x}$). Άλλες φορές υπάρχουν μη φραγμένες συναρτήσεις με τοπικά ελάχιστα τα οποία πιθανόν να ενδιαφέρουν (π.χ. $f(x) = x^3 - 3x$). Ακόμα, ένα μόνιμο μειονέκτημα που εμφανίζεται σε όλους τους αλγορίθμους που προτείνονται, είναι ότι δεν εγγυώνται την εύρεση του μέγιστου ή ελάχιστου της συνάρτησης $f(x)$ που αντιστοιχεί σε όλο το χώρο μεταβλητότητας του x. Αυτό το ακρότατο (μέγιστο ή ελάχιστο) σημείο ονομάζεται **σφαιρικό ή ολικό** (global minimum or maximum). Αντίθετα, είναι πιο πιθανή η εύρεση ενός ακρότατου που αντιστοιχεί σε μια περιοχή μεταβλητότητας του x και το οποίο ονομάζεται **τοπικό** (local minimum or maximum). Μια απλή πρακτική τεχνική προτείνει την επίλυση του προβλήματος από πολλά διαφορετικά σημεία εκκίνησης και την επιλογή της καλύτερης ανάμεσά τους λύσης. Ακόμα και η εύρεση ενός τοπικού ελαχίστου κρύβει δυσκολίες. Είναι αποδεκτή η εύρεση ενός σημείου x^* για το οποίο ισχύει $f(x) \geq f(x^*)$ για όλα τα x που βρίσκονται αρκετά κοντά του. Επίσης, προβλήματα προκύπτουν όταν η συνάρτηση δεν είναι **ομαλή** (smooth), δεδομένου ότι τα ελάχιστα σημεία αυτών των συναρτήσεων δεν έχουν τις ίδιες ιδιότητες με τα κλασικά ελάχιστα σημεία (Σχήμα 13).



Σχήμα 13

2.1.2. Διαχωρισμός των προβλημάτων

Όπως σημειώθηκε παραπάνω, η γενική μορφή ενός μη γραμμικού προβλήματος δίνεται από τις εξισώσεις της Γενικής Μορφής Μη Γραμμικού Προβλήματος. Σε γενικές γραμμές, τα προβλήματα μπορούν να χωριστούν σε δύο μεγάλες κατηγορίες.

- Προβλήματα χωρίς περιορισμούς (unconstrained problems)
- Προβλήματα με περιορισμούς (constrained problems)

Στην πρώτη κατηγορία πρέπει να βρεθεί το ελάχιστο κάποιας αντικειμενικής συνάρτησης και ταυτόχρονα θα πρέπει να ικανοποιούνται κάποιοι περιορισμοί οι οποίοι θα παρουσιάζονται με τη μορφή εξισώσεων ή ανισώσεων. Στην δεύτερη κατηγορία θα πρέπει να βρεθεί το ελάχιστο κάποιας αντικειμενικής συνάρτησης και ταυτόχρονα θα πρέπει να ικανοποιούνται κάποιοι περιορισμοί που τίθενται μεταξύ των μεταβλητών, όπως φαίνεται και στις εξισώσεις της Γενικής Μορφής Μη Γραμμικού Προβλήματος.

Συνήθως, η μελέτη για τις μεθόδους αυτές ξεκινάει από τα προβλήματα χωρίς περιορισμούς. Η σειρά αυτή ακολουθείται για το λόγο ότι πολλά από τα προβλήματα βελτιστοποίησης υπό περιορισμούς μπορούν να λυθούν με τεχνικές που τα μετατρέπουν σε προβλήματα χωρίς περιορισμούς με μικρότερη διάσταση.

Ένας ακόμη περιορισμός που γίνεται, είναι η μελέτη και στις δύο κατηγορίες, αρχικά των μονοδιάστατων προβλημάτων, δηλαδή των προβλημάτων με μία μεταβλητή. Η διαπραγμάτευση πρώτα των μονοδιάστατων προβλημάτων γίνεται για δύο κυρίως λόγους. Πρώτον, γιατί σε πάρα πολλές πολυδιάστατες μεθόδους ένα υποπρόβλημα που εμφανίζεται είναι η βελτιστοποίηση συναρτήσεων μίας μεταβλητής, οι οποίες προκύπτουν μέσα από την αλγοριθμική διαδικασία και πρέπει να αντιμετωπιστούν. Δεύτερον, με τον τρόπο αυτό, είναι πιο εύκολο να αποκτηθεί οικειότητα με τον τρόπο σκέψης που ακολουθείται στον συγκεκριμένο χώρο. Άλλωστε, θα πρέπει να σημειωθεί ότι μερικές μέθοδοι μεταφέρονται αμέσως από το μονοδιάστατο πρόβλημα στο πολυδιάστατο με απλή γενίκευση.

Προβλήματα βελτιστοποίησης χωρίς περιορισμούς εμφανίζονται σε πολλές γνωστές στατιστικές μεθόδους όπως στην παλινδρόμηση, στην ανάλυση κατά παράγοντες, στην εκτιμητική, στη θεωρία προσεγγίσεων, στην επίλυση συστημάτων μη γραμμικών εξισώσεων. Για παράδειγμα ο γνωστός εκτιμητής μέγιστης πιθανοφάνειας προκύπτει από την επίλυση ενός εν γένει πολυδιάστατου μη γραμμικού προβλήματος χωρίς περιορισμούς. Η διάσταση του προβλήματος είναι ίση με το πλήθος των παραμέτρων που πρέπει να εκτιμηθούν.

Άλλη εφαρμογή στην οποία προκύπτει μη γραμμικό πρόβλημα βελτιστοποίησης είναι π.χ. (Hillier and Lieberman 1989), η περίπτωση μεγιστοποίησης κέρδους μίας εταιρείας από την πώληση ενός προϊόντος όταν η τιμή πώλησης εξαρτάται από την ποσότητα που πωλείται. Άλλη περίπτωση προκύπτει σε προβλήματα μεταφοράς αγαθών από διάφορες πηγές σε πολλούς προορισμούς όταν το κόστος μεταφοράς μίας μονάδας εξαρτάται από το συνολικό αριθμό μονάδων που μεταφέρονται, δεδομένων πάντα κάποιων περιορισμών που αφορούν στην προσφορά και τη ζήτηση. Ακόμα μη γραμμικές μέθοδοι βελτιστοποίησης χρησιμοποιούνται ευρέως σε προβλήματα επενδύσεων όπου κάποιος πρέπει να αποφασίσει για παράδειγμα πόσες μετοχές να αγοράσει από n διαφορετικές επενδύσεις έτσι ώστε δεδομένης κάποιας μέσης τιμής και διασποράς απόδοσης της κάθε επένδυσης να μεγιστοποιήσει το κέρδος του.

2.1.3. Χρήσιμοι ορισμοί

Από τις βασικότερες ιδιότητες μίας συνάρτησης στα μη γραμμικά προβλήματα είναι μεταξύ άλλων, η συνέχεια, η ομαλότητα, η καμπυλότητα και η κυρτότητα.

Έστω ότι υπάρχει μία συνάρτηση $f(x)$ ορισμένη σε κάποιο σύνολο, και έστω E ένα υποσύνολο του πεδίου ορισμού της.

2.1.3.1. *Ορισμός Συνέχειας Συνάρτησης*

Αν σε ένα σημείο $x_0 \in E$, $\lim_{x \to x_0} f(x) = f(x_0)$ *τότε η συνάρτηση* $f(x)$ *είναι* ***συνεχής*** *στο σημείο* x_0. *Αν η* $f(x)$ *είναι συνεχής για κάθε* $x \in E$ *τότε θεωρείται ότι η συνάρτηση είναι συνεχής στο* E *και χρησιμοποιείται ο συμβολισμός* $f \in C$ *στο* E.

Ο συμβολισμός $f \in C^k$ στο E χρησιμοποιείται όταν υπάρχει η παράγωγος k τάξης της συνάρτησης $f(x)$ και είναι συνεχής. Μπορεί εύκολα να δειχθεί ότι αν μία συνάρτηση είναι $f \in C^{k+1}$ στο $E \subset R^n$ τότε είναι και $f \in C^k$ στο E.

2.1.3.2. *Ορισμός Ομαλότητας Συνάρτησης*

*Μία συνάρτηση $f : R^n \to R$ είναι **ομαλή** σε ένα σημείο του πεδίου ορισμού της, αν υπάρχουν οι παράγωγοί της στο σημείο αυτό και είναι συνεχής.*

Για παράδειγμα, η συνάρτηση $f(x)=\begin{cases}-x+1, & x>0\\ x+1, & x<0\end{cases}$ δεν είναι ομαλή στο σημείο $x=0$. Η έννοια της ομαλότητας είναι πολύ σημαντική στη θεωρία βελτιστοποίησης. Αυτό συμβαίνει διότι οι διάφορες συνθήκες βελτιστοποίησης ισχύουν κυρίως για ομαλές αντικειμενικές συναρτήσεις. Με τις συνθήκες αυτές διερευνώνται οι δυνατότητες ύπαρξης υποψηφίων βέλτιστων σημείων. Έτσι, κατασκευάζονται αλγόριθμοι από την προσπάθεια εύρεσης βέλτιστων σημείων που να ικανοποιούν τις λεγόμενες συνθήκες βελτιστοποίησης. Στην περίπτωση που οι αντικειμενικές συναρτήσεις δεν είναι ομαλές είναι αναγκαία η κατασκευή ειδικών αλγορίθμων πράγμα που συνήθως δεν είναι εύκολο. Η τρίτη πολύ σημαντική ιδιότητα στα προβλήματα μη γραμμικού προγραμματισμού είναι η ιδιότητα της συνεχούς καμπυλότητας. Η καμπυλότητα μίας συνάρτησης μελετάται από τη δεύτερη της παράγωγο εφόσον πρόκειται για συνάρτηση μίας μεταβλητής. Όταν υπάρχει μία συνάρτηση n μεταβλητών τότε χρησιμοποιείται ο πίνακας διάστασης $n \times n$ που περιέχει σε κάθε θέση τις δεύτερες μερικές παραγώγους. Η συνέχεια των δεύτερων μερικών παραγώγων είναι πολύ σημαντική όταν απασχολεί ο ρυθμός σύγκλισης ενός αλγορίθμου που επιλύει ένα αντίστοιχο πρόβλημα. Επίσης είναι σημαντική στην ανάλυση ευαισθησίας δηλαδή στις διαφορές που προκύπτουν στην βέλτιστη λύση όταν γίνονται αλλαγές στις παραμέτρους του προβλήματος.

2.1.3.3. *Ορισμός Κλίσης Συνάρτησης*

Αν μία συνάρτηση $f \in C^1$ στο E, τότε το ανάστροφο διάνυσμα της $f(x)$ ονομάζεται ***κλίση*** *(gradient) και συμβολίζεται με $\nabla f(x)$, ή απλούστερα με ∇f εφόσον δεν αναφέρεται σε συγκεκριμένο σημείο. Άρα είναι:*

$$\nabla f(x) = \begin{bmatrix} \partial f(x)/\partial x_1 \\ \partial f(x)/\partial x_2 \\ \vdots \\ \partial f(x)/\partial x_n \end{bmatrix}$$

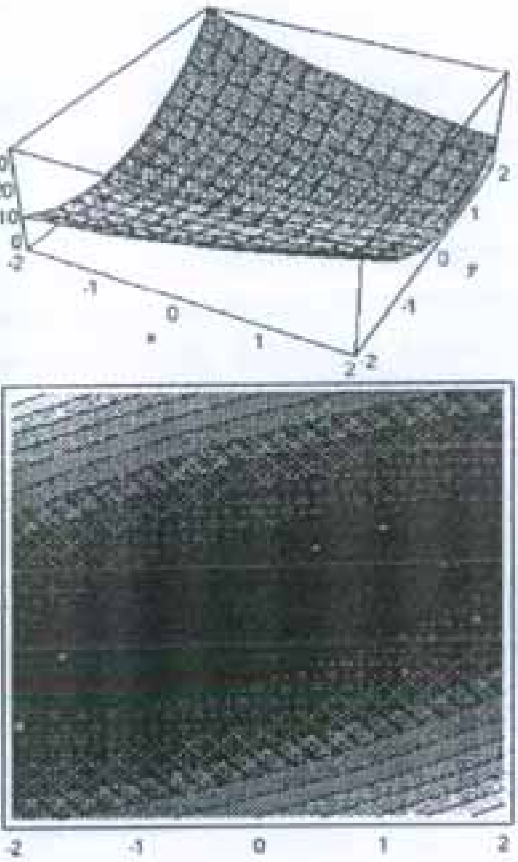
Στα προβλήματα που θα προκύψουν, η κλίση (ή η πρώτη παράγωγος στα μονοδιάστατα) είναι βαρύνουσα παράμετρος, διότι στην πράξη διερευνείται η δυνατότητα ύπαρξης σημείων για τα οποία η κλίση της αντικειμενικής συνάρτησης μηδενίζεται, οπότε οδηγεί σε τοπικά ακρότατα, υποψήφια ελάχιστα ή μέγιστα που είναι και τα ζητούμενα βέλτιστα σημεία.

2.1.3.4. *Ορισμός Καμπυλότητας Συνάρτησης*

Το σύνολο των σημείων για τα οποία η αντικειμενική συνάρτηση $f(x)$ που παριστάνει μία επιφάνεια (εν γένει), έχει μία σταθερή τιμή, ονομάζεται μία ***καμπύλη*** *(contour) της συνάρτησης.*

ΠΑΡΑΔΕΙΓΜΑ :

Στο Σχήμα 14 παρατηρείται η γραφική παράσταση της συνάρτησης $f(x)=x_1^2+x_2^2-3x_1x_2$, όπου $-2<x_1<2$ και $-2<x_2<2$ και στη συνέχεια η προβολή των ισοϋψών καμπυλών της στο επίπεδο όπου όμοιος τόνος γκρίζου χρώματος σημαίνει σταθερή τιμή.



Σχήμα 14

Η κλίση μίας συνάρτησης σε ένα σημείο x_0 όπως ορίστηκε παραπάνω, είναι ένα διάνυσμα κάθετο στη καμπύλη της επιφάνειας $f(x)$ που περνάει από το σημείο x_0, και ορίζει την κατεύθυνση της μέγιστης αύξησης-ανόδου της $f(x)$ στο σημείο αυτό (steepest ascent direction). Η αντίθετη της κλίσης διεύθυνση ορίζει τη κατεύθυνση της μέγιστης μείωσης-πτώσης (steepest descent direction). Όλα αυτά είναι αληθή, εφόσον για την έννοια της απόστασης χρησιμοποιείται η Ευκλείδεια norm. Σ' αυτό το σημείο θα πρέπει να σημειωθεί ότι σαν Ευκλείδεια norm (μέτρο) ενός διανύσματος $x \in E \subseteq R^n$ ορίζεται η ακόλουθη:

$$\|x\| = \sqrt{\sum_{j=1}^{n} (x_j)^2}$$

Έτσι, κάθε διάνυσμα έστω u, κάθετο πάνω στη κλίση $\nabla f(x_0)$, όπως είναι για παράδειγμα η εφαπτόμενη επιφάνεια στο $f(x_0)$, επαληθεύει τη σχέση $u^T \nabla f(x_0) = 0$.

Πρέπει επίσης να σημειωθεί ότι μία ακολουθία διανυσμάτων, (σημείων στο $E \subseteq R^n$), $\{x_k\}_{k=0}^{\infty}$ έχει όριο το διάνυσμα x^* αν για κάθε $\varepsilon > 0$ υπάρχει ένας ακέραιος $k(\varepsilon)$ τέτοιος ώστε για κάθε $k > k(\varepsilon)$ ισχύει

$$\|x_k - x^*\| < \varepsilon$$

2.1.3.5. *Ορισμός Πίνακα Hessian*

Όταν η συνάρτηση $f \in C^2$ στο E, ο πίνακας που περιέχει τις δεύτερες μερικές παραγώγους στο σημείο x_k για παράδειγμα, είναι τετραγωνικός διάστασης n και ονομάζεται Hessian πίνακας. Στις αντίστοιχες θέσεις i και j περιέχει το στοιχείο $\partial f(x) / \partial x_i \partial x_j$. Συμβολίζεται με $f''(x)$ ή $\nabla^2 f(x)$ ή $H(x_k)$ ή απλά με $\nabla^2 f$ ή H αν δεν αναφέρεται σε συγκεκριμένο σημείο.

Ο Hessian πίνακας έχει την ακόλουθη μορφή:

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial f(x_k)}{\partial x_1^2} & \cdots & \frac{\partial f(x_k)}{\partial x_1 \partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial f(x_k)}{\partial x_n \partial x_1} & \cdots & \frac{\partial f(x_k)}{\partial x_n^2} \end{bmatrix}$$

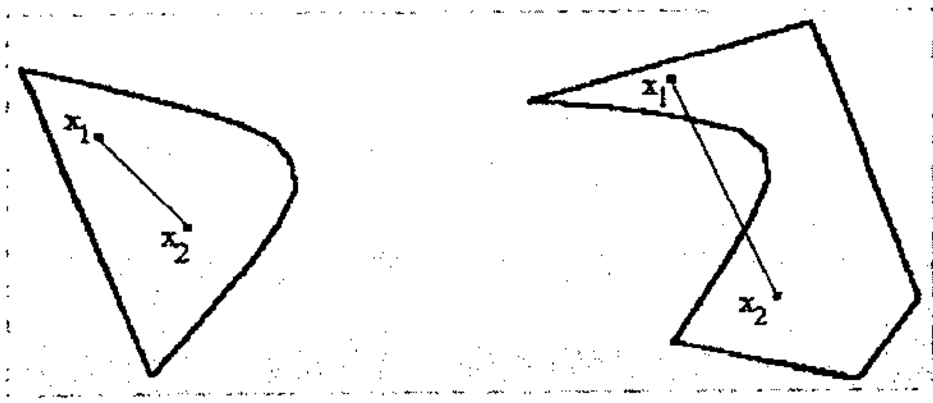
Είναι γνωστό ότι αν $f \in C^2$ στο $E \subseteq R^n$, τότε $\partial f(x)/\partial x_i \partial x_j = \partial f(x)/\partial x_j \partial x_i$. Δηλαδή, ο πίνακας στην περίπτωση αυτή είναι συμμετρικός (προφανώς είναι πάντα τετραγωνικός).

Η άλλη ιδιότητα η οποία αναφέρθηκε στην αρχή της παραγράφου, είναι η κυρτότητα. Η κυρτότητα είναι σημαντική διότι καθιστά τοπικές (local) βέλτιστες λύσεις, (δηλαδή τοπικά ακρότατα σημεία), σφαιρικές (global) όταν αφορούν κυρτές συναρτήσεις ορισμένες σε κυρτά σύνολα, γεγονός που επαυξάνει δραστικά την αποτελεσματικότητα μίας λύσης που βρέθηκε. Όταν αναφέρεται η έννοια κυρτότητα συνήθως οι έννοιες που συναντούνται είναι η κυρτή συνάρτηση, η κοίλη συνάρτηση και το κυρτό σύνολο.

2.1.3.6. *Ορισμός Κυρτότητας Συνόλου*

*Ένα σύνολο $E \subseteq R^n$ είναι **κυρτό** αν το ευθύγραμμο τμήμα που ενώνει οποιαδήποτε δύο σημεία του, περιέχεται στο E. Δηλαδή, αν x και y ανήκουν στο E, τότε $\lambda x + (1-\lambda) y$ ανήκει στο E για κάθε $0 \le \lambda \le 1$.*

Τα κυρτά σύνολα μπορούν να είναι ανοιχτά, κλειστά ή τίποτε από τα δύο. Στο Σχήμα 15 παρατηρείται ένα κυρτό και ένα μη κυρτό σύνολο στον R^2. Επίσης, παρατηρείται ότι στο πρώτο σύνολο κάθε γραμμή που συνδέει δύο σημεία περιέχεται σ' αυτό. Στο δεύτερο σύνολο αυτό δε συμβαίνει πάντα.



Σχήμα 15

2.1.3.7. *Ορισμός Κυρτότητας Συνάρτησης*

*Μία συνάρτηση $f(x)$ καλείται **κυρτή** στο κυρτό σύνολο $E \subseteq R^n$ αν για οποιαδήποτε σημεία x_1 και $x_2 \in E$ ισχύει*

$$f(\lambda x_1 + (1-\lambda x_2)) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$

όπου $0 \leq \lambda \leq 1$.

Η συνάρτηση καλείται αυστηρά κυρτή αν γίνεται να αντικατασταθεί το $\leq$ με το $<$.

2.1.3.8. *Ορισμός Κοιλότητας Συνάρτησης*

*Μία συνάρτηση $f(x)$ καλείται **κοίλη** στο σύνολο $E \subseteq R^n$ αν για οποιαδήποτε σημεία x_1 και $x_2 \in E$ ισχύει*

$$f(\lambda x_1 + (1-\lambda x_2)) \geq \lambda f(x_1) + (1-\lambda) f(x_2)$$

όπου $0 \leq \lambda \leq 1$.

Η συνάρτηση καλείται αυστηρά κοίλη αν γίνεται να αντικατασταθεί το $\geq$ με το $>$. Ακόμα, είναι εύκολα αντιληπτό ότι μία συνάρτηση $f(x)$ είναι κοίλη αν η συνάρτηση $-f(x)$ είναι κυρτή.

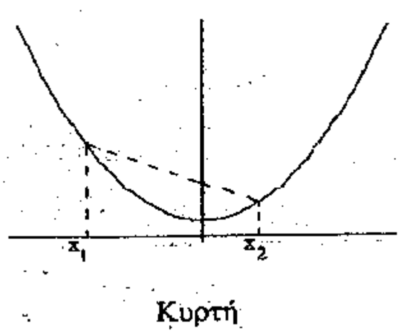
Μία διαφορίσιμη, $(f \in C^2)$, κυρτή συνάρτηση $f : E \subseteq R^n \to R, E$ κυρτό, έχει τις ακόλουθες ιδιότητες:

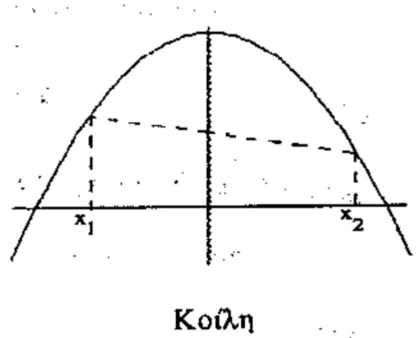
- $f(x_2) - f(x_1) \geq \nabla^T f(x_1)(x_2 - x_1) \forall x_1, x_2 \in E \subseteq R^n$,
- Ο Hessian πίνακας είναι θετικά ορισμένος (ή ημιορισμένος), για κάθε x, αν η $f(x)$ είναι αυστηρά κυρτή (ή κυρτή)

δηλαδή
$$s^T \nabla^2 f(x) s > 0 \;\; \forall\, s \neq 0, \in R^n, x \in E,$$
$$\left(s^T \nabla^2 f(x) s > 0 \;\; \forall\, s \in R^n, x \in E\right)$$

- Στο πεδίο τιμών της έχει ένα μόνο ελάχιστο.

Οι παραπάνω ιδιότητες καθιστούν πράγματι την κυρτότητα πολύ σημαντική έννοια στα προβλήματα τα οποία πρέπει να λυθούν. Στο Σχήμα 16 υπάρχει μία κυρτή και μία κοίλη συνάρτηση στο R .γίνεται αντιληπτό πως κάθε ευθεία που ενώνει δύο σημεία της κυρτής συνάρτησης περνά "πάνω" απο την καμπύλη της, δηλαδή δεν μπορεί να πάρει καμία τιμή μεγαλύτερη από αυτές που δίνει η γραμμική παρεμβολή χρησιμοποιώντας τα σημεία αυτά. Το αντίθετο συμβαίνει στην άλλη περίπτωση.



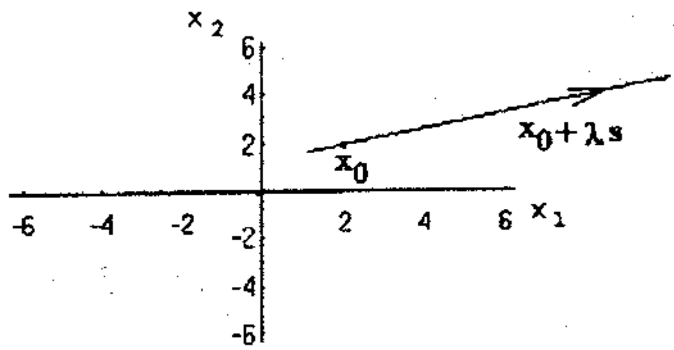


Σχήμα 16

Μία άλλη σημαντική έννοια στα μη γραμμικά προβλήματα είναι η ευθεία μέσα σε ένα n – διάστατο χώρο. Ουσιαστικά αφορά το σύνολο των σημείων

$$x\left(=x(\lambda)\right)=x_0+\lambda s$$

όπου $x, x_0 \in E \subseteq R^n$ για κάποια διάσταση n και για όλα τα $\lambda \geq 0$. Είναι φανερό ότι ξεκινάει από ένα σταθερό σημείο x_0, για $\lambda = 0$, πάνω στην ευθεία και s είναι η κατεύθυνση της ευθείας. Στο Σχήμα 17 (Fletcher 1987), βλέπουμε την ευθεία $x = x_0 + \lambda s \Rightarrow \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} 3 \\ 1 \end{pmatrix}$



Σχήμα 17

Διακρίνεται το διάνυσμα s με το βέλος. Στη συνέχεια, το διάνυσμα της κατεύθυνσης κανονικοποιείται διαιρώντας με το μέτρο του, ώστε να ισχύει: $s^T s = 1$, δηλαδή $\sum_i s_i^2 = 1$, ισοδύναμα $\|s\| = \sqrt{s^T s} = 1$

Αν υπολογιστεί η παράγωγος μίας συνάρτησης f στη κατεύθυνση μίας ευθείας $x(\lambda)$, τότε θεωρείται ότι

$$\frac{df}{d\lambda} = \sum_i \frac{d}{d\lambda} x_i(\lambda) \frac{\partial}{\partial x_i} = \sum_i s_i \frac{\partial}{\partial x_i} = s^T \nabla$$

Άρα, η παράγωγος (slope), της συνάρτησης $f\left(=f\left(x(\lambda)\right)\right)$ κατά μήκος της ευθείας σε κάθε σημείο της $x(\lambda)$, είναι:

$$\frac{df}{d\lambda} = s^T \nabla f = \nabla^T f s$$

Όμοια, η καμπυλότητα (curvature) κατά μήκος της γραμμής θα είναι:

$$\frac{d^2 f}{d\lambda^2} = \frac{d}{d\lambda}\frac{df}{d\lambda} s^T \nabla\left(\nabla^T fs\right) f = s^T \nabla^2 fs$$

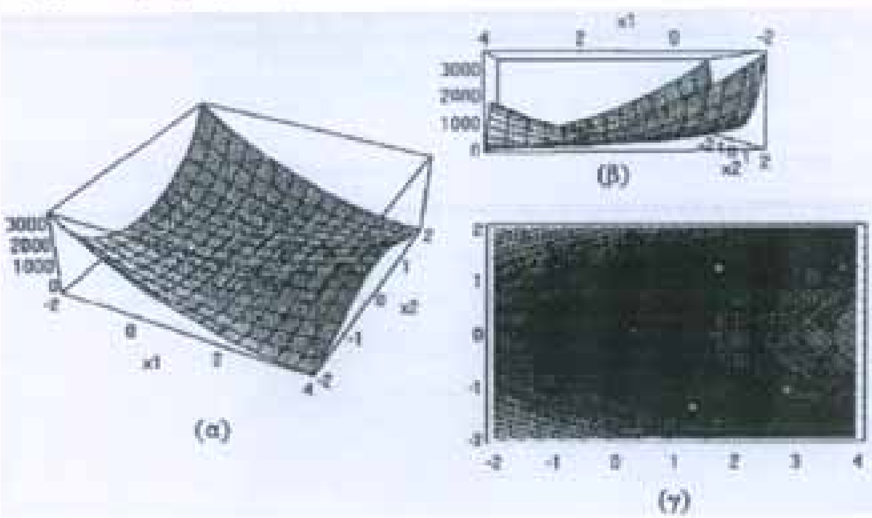
Όταν ένα σημείο x^* είναι τοπικό ελάχιστο μίας αντικειμενικής συνάρτησης $f(x)$, τότε κατά μήκος κάθε ευθείας $x(=x(\lambda)) = x^* + \lambda s$ που διέρχεται από το x^*, η συνάρτηση έχει μηδενική παράγωγο (slope) και μη αρνητική καμπυλότητα (non-negative curvature) στο x^*.

<u>ΠΑΡΑΔΕΙΓΜΑ</u> :

Μία κλασική δισδιάστατη αντικειμενική συνάρτηση με την οποία μπορεί να δοκιμαστεί η απόδοση των μεθόδων βελτιστοποίησης, είναι η συνάρτηση του Rosenbrock που δίνεται παρακάτω:

$$f(x)100\left(x_2 - x_1^2\right)^2 + \left(1 - x_1\right)^2$$

Τα σχήματα της συνάρτησης αυτής παρατίθενται στη συνέχεια και είναι τα Σχήματα 18 α,β,γ, δηλαδή το γράφημά της από δύο οπτικές γωνίες καθώς επίσης και η προβολή των ισοβαρών καμπυλών της.



Σχήματα 18 α,β,γ

Για τη συνάρτηση αυτή, αν υπολογιστεί η κλίση (gradient) και ο Πίνακας Hessian θα βρεθούν οι παρακάτω εκφράσεις:

$$\nabla f(x)=\begin{pmatrix}-400x_1\left(x_2-x_1^2\right)-2\left(1-x_1\right)\\ 200\left(x_2-x_1^2\right)\end{pmatrix}$$

και

$$\nabla^2 f(x)=\begin{bmatrix}1200x_1^2-400x_2+2 & -400x_1\\ -400x_1 & 200\end{bmatrix}$$

Έτσι, παρατηρείται ότι γενικά η κλίση και η καμπυλότητα εξαρτώνται από το σημείο x. Για $x=\begin{pmatrix}0\\0\end{pmatrix}$, $\nabla f(x)=\begin{pmatrix}-2\\0\end{pmatrix}$ και $\nabla^2 f(x)=\begin{bmatrix}2 & 0\\ 0 & 200\end{bmatrix}$.

Άρα, από τους τύπους $\frac{df}{d\lambda}=s^T\nabla f=\nabla^T fs$ και

$\frac{d^2 f}{d\lambda^2}=\frac{d}{d\lambda}\frac{df}{d\lambda}s^T\nabla\left(\nabla^T fs\right)f=s^T\nabla^2 fs$ (που αναφέρθηκαν προηγουμένως)

αντίστοιχα, στο σημείο $x=\begin{pmatrix}0\\0\end{pmatrix}$, η παράγωγος (slope) κατά μήκος της γραμμής που παράγεται από τη διεύθυνση $s=\begin{pmatrix}1\\0\end{pmatrix}$ δηλαδή από τον άξονα των συντεταγμένων x_1, είναι $s^T\nabla f=-2$ (η προβολή της κλίσης στον άξονα των x) και η καμπυλότητα είναι $s^T\nabla^2 fs=2$.

Από το προηγούμενο παράδειγμα φαίνεται και η χρησιμότητα της κανονικοποίησης των διευθύνσεων, όπως δίνεται στην εξίσωση $\|s\|=\sqrt{s^T s}=1$, δεδομένου ότι τα παραπάνω αποτελέσματα εξαρτώνται από το μέγεθος του s. Ενώ αν μετατραπούν όλες οι διευθύνσεις σε μοναδιαία διανύσματα θα υπάρχουν ίδια αποτελέσματα με διαφοροποίηση στο συντελεστή λ. Με το ίδιο σκεπτικό, το διάνυσμα $\nabla f(x)/\|\nabla f(x)\|$ υποδηλώνει τη μοναδιαία διεύθυνση της κλίσης. Ενώ η παραπάνω διεύθυνση είναι η διεύθυνση της μεγαλύτερης αύξησης στο σημείο x, η $-\nabla f(x)/\|\nabla f(x)\|$ είναι η διεύθυνση της μεγαλύτερης μείωσης της συνάρτησης στο σημείο αυτό, για όλες τις δυνατές μοναδιαίες διευθύνσεις. Υπενθυμίζοντας, η κλίση (gradient) είναι κάθετη στη καμπύλη $f(x)$ και στο εφαπτόμενο σ'αυτήν επίπεδο, στο σημείο x.

2.1.3.9. Ορισμός Διεύθυνσης Μείωσης-Πτώσης

*Έστω μία συνάρτηση $f: R^n \to R$. Ένα διάνυσμα s_k ονομάζεται **διεύθυνση μείωσης-πτώσης** (nonascent direction) στο σημείο $x_k \in R^n$ αν ισχύει $s_k^T \nabla f(x_k) \leq 0$. Αν επιπλέον αντικατασταθεί το $\leq$ με το $<$ τότε υπάρχει διεύθυνση αυστηρής μείωσης-πτώσης (descent direction) στο σημείο $x_k \in R^n$.*

2.1.3.10. ***Ορισμός Διεύθυνσης Αύξησης***

*Έστω μία συνάρτηση $f: R^n \to R$. Ένα διάνυσμα s_k ονομάζεται **διεύθυνση αύξησης** (nondescent direction) στο σημείο $x_k \in R^n$ αν ισχύει $s_k^T \nabla f(x_k) \geq 0$. Αν επιπλέον αντικατασταθεί το $\geq$ με το $>$ τότε υπάρχει διεύθυνση αυστηρής αύξησης (ascent direction) στο σημείο $x_k \in R^n$.*

Οι δύο ανωτέρω ορισμοί αντιστοιχούν στην ιδιότητα της αρνητικής ή θετικής πρώτης παραγώγου σε μία μονοδιάστατη συνάρτηση και κατ'επέκταση στην φθίνουσα και αύξουσα συνάρτηση σε κάποιο διάστημα.

2.1.3.11. Ορισμός Διεύθυνσης Μη Θετικής Καμπυλότητας

*Έστω μία συνάρτηση $f : R^n \to R$. Ένα διάνυσμα s_k ονομάζεται διεύθυνση **μη θετικής καμπυλότητας** (nonpositive curvature) στο σημείο $x_k \in R^n$ αν ισχύει $s_k^T \nabla f(x_k) s_k \leq 0$. Αν αντικατασταθεί το $\leq$ με το $<$ τότε υπάρχει διεύθυνση αρνητικής καμπυλότητας (negative curvature) στο σημείο $x_k \in R^n$.*

2.1.3.12. *Ορισμός Διεύθυνσης Μη Αρνητικής Καμπυλότητας*

*Έστω μία συνάρτηση $f : R^n \to R$. Ένα διάνυσμα s_k ονομάζεται διεύθυνση **μη αρνητικής καμπυλότητας** (nonnegative curvature) στο σημείο $x_k \in R^n$ αν ισχύει $s_k^T \nabla f(x_k) s_k \geq 0$. Αν αντικατασταθεί το $\geq$ με το $>$ τότε υπάρχει διεύθυνση θετικής καμπυλότητας (positive curvature) στο σημείο $x_k \in R^n$.*

Ο Ορισμός Διεύθυνσης Μη Θετικής Καμπυλότητας είναι ισοδύναμος με τη μη θετική ή αρνητική τιμή της δεύτερης παραγώγου σε ένα σημείο x σε μία μονοδιάστατη συνάρτηση $f : R \to R$, ενώ ο Ορισμός Διεύθυνσης Μη Αρνητικής Καμπυλότητας ισοδυναμεί με μη αρνητική ή θετική δεύτερη παράγωγο και κατ'επέκταση στην ιδιότητα της κοιλότητας ή κυρτότητας.

Αν στον Ορισμό Διεύθυνσης Μη Αρνητικής Καμπυλότητας η ιδιότητα $s_k^T \nabla f(x_k) s_k \geq 0$ της συνάρτησης $f : R^n \to R$, επεκταθεί για κάθε διάνυσμα $s_k \in R^n$ και για κάθε σημείο $x_k \in R^n$, μέσα σε ένα υποσύνολο του πεδίου ορισμού της, η συνάρτηση είναι κυρτή (ή αυστηρά κυρτή αν αντικατασταθεί το $\geq$ με $>$) στο υποσύνολο αυτό. Ανάλογη ιδιότητα σε μονοδιάστατη συνάρτηση είναι να είναι η δεύτερη παράγωγος μη αρνητική (ή θετική), σε ένα διάστημα του πεδίου ορισμού.

2.2. Πολυδιάστατα Προβλήματα Χωρίς Περιορισμούς

Στο σημείο αυτό, πρέπει να αναπτυχθούν οι μέθοδοι με τις οποίες μπορεί να βρεθεί το μέγιστο ή το ελάχιστο μίας αντικειμενικής συνάρτησης πολλών μεταβλητών. Δηλαδή το πρόβλημα στη γενική του μορφή, λαμβάνοντας υπόψη τις σχέσεις της Γενικής Μορφής Μη Γραμμικού Προβλήματος, είναι το ακόλουθο:

$$Minimize\left(Maximize\right)\ f\left(x\right),\ x \in E \subseteq R^n$$

Οι ιδέες και οι τεχνικές που χρησιμοποιούν οι υπάρχοντες αλγόριθμοι διαφοροποιούνται ως προς το μέγεθος της πληροφορίας που χρειάζονται από τη συνάρτηση $f(x)$ για να λύσουν κάποιο πρόβλημα. Στα μονοδιάστατα προβλήματα, οι αλγόριθμοι εντοπίζουν ένα τοπικό ακρότατο για το οποίο πρέπει να βρεθούν συνθήκες για να είναι κατ'αρχήν ελάχιστο (ή μέγιστο) και αν γίνεται να αποδειχθεί ότι είναι σφαιρικό.

Στη συνέχεια δίνονται κάποιες ικανές ή αναγκαίες συνθήκες σε μορφή θεωρήματος και οι οποίες αφορούν την ύπαρξη ενός ελαχίστου μίας πολυδιάστατης αντικειμενικής συνάρτησης. Οι συνθήκες αυτές είναι φυσική γενίκευση των αποτελεσμάτων για τις μονοδιάστατες συναρτήσεις. Όπως αναφέρθηκε παραπάνω, όταν ένα σημείο x^* είναι τοπικό ελάχιστο μίας αντικειμενικής συνάρτησης $f(x)$, τότε κατά μήκος κάθε ευθείας $x(\lambda)=x^*+\lambda s$ που διέρχεται από το x^*, η συνάρτηση έχει μηδενική παράγωγο (slope) και μη αρνητική καμπυλότητα (non-negative curvature) στο x^*. Από τις σχέσεις $\frac{df}{d\lambda}=s^T\nabla f=\nabla^T fs$ και $\frac{d^2 f}{d\lambda^2}=\frac{d}{d\lambda}\frac{df}{d\lambda}s^T\nabla\left(\nabla^T fs\right)f=s^T\nabla^2 fs$ (που αναφέρθηκε παραπάνω) είναι εύκολο να βγει το συμπέρασμα ότι για κάθε διεύθυνση s θα υπάρχει μηδενική παράγωγος και μη αρνητική καμπυλότητα στο x^*.

Άρα ισχύει:

$s^T \nabla f(x^*) = 0$ και $s^T \nabla^2 f(x^*) s \geq 0$

Αφού οι παραπάνω σχέσεις ισχύουν για κάθε διεύθυνση s, από την πρώτη σχέση προκύπτει:

$\nabla f(x^*) = 0$

ενώ η δεύτερη είναι

$s^T \nabla^2 f(x^*) s \geq 0, \quad \forall s$.

Δεδομένου ότι οι δύο παραπάνω σχέσεις προκύπτουν από το γεγονός ότι το x^* είναι βέλτιστο σημείο για τη συνάρτηση $f(x)$, είναι αναγκαίες συνθήκες για την ύπαρξη ενός τοπικού ελαχίστου. Η συνθήκη $\nabla f(x^*) = 0$ ονομάζεται αναγκαία συνθήκη πρώτης τάξης ενώ η συνθήκη $s^T \nabla^2 f(x^*) s \geq 0, \quad \forall s$ ονομάζεται αναγκαία συνθήκη δεύτερης τάξης.

2.2.1. *Αναγκαία Συνθήκη Πρώτης Τάξης*

Έστω μία συνάρτηση $f: E \subseteq R^n \to R$ η οποία είναι $f \in C^1$ στο E. Αν ένα σημείο x^* δίνει τοπικό ελάχιστο στο E, τότε μία αναγκαία συνθήκη είναι $\nabla f(x^*) = 0$.

2.2.2. *Αναγκαία Συνθήκη Δεύτερης Τάξης*

Έστω μία συνάρτηση $f: E \subseteq R^n \to R$ η οποία είναι $f \in C^2$ στο E. Αν ένα σημείο x^* δίνει τοπικό ελάχιστο στο E, τότε μία αναγκαία συνθήκη είναι ο Πίνακας Hessian να είναι θετικά ορισμένος.

Η τάξη που αναφέρεται στις συνθήκες έχει να κάνει με τις πληροφορίες που χρησιμοποιούνται. Στην πρώτη περίπτωση υπάρχει κλίση (πρώτη παράγωγος), ενώ στη δεύτερη υπάρχει ο Πίνακας Hessian (δεύτερη παράγωγος). Ένα κενό που υπάρχει στις αναγκαίες συνθήκες είναι η περίπτωση της μηδενικής καμπυλότητας (ανάλογο του σημείου καμπής). Είναι όμως δυνατόν να προκύψουν ικανές συνθήκες στηριζόμενες στα αντίστοιχα αποτελέσματα που αφορούν μονοδιάστατα προβλήματα αρκεί να μετατραπεί η συνθήκη δεύτερης τάξης. Έτσι προκύπτει:

2.2.3. ***Ικανές Συνθήκες Δεύτερης Τάξης***

Έστω μία συνάρτηση $f : E \subseteq R^n \to R$ *η οποία είναι* $f \in C^2$ *στο* E. *Ικανές συνθήκες έτσι ώστε ένα σημείο* x^* *δίνει τοπικό ελάχιστο στο* E, *είναι να ισχύει η αναγκαία συνθήκη πρώτης τάξης και ο Πίνακας Hessian να είναι θετικά ορισμένος, δηλαδή να ισχύει*

$$\nabla f\left(x^*\right)=0 \ \textit{και} \ s^T \nabla^2 f\left(x^*\right) s>0 \quad \forall s \neq 0$$

Όταν η συνάρτηση $f(x)$ είναι κυρτή, ορισμένη σε ένα κυρτό σύνολο $E \subseteq R^n$, οι αναγκαίες συνθήκες πρώτης τάξης είναι ικανές για να είναι ένα σημείο x^* βέλτιστο, δεδομένου ότι ο Πίνακας Hessian είναι πάντα θετικά ορισμένος. Το ελάχιστο που θα βρεθεί θα είναι ταυτοχρόνως και σφαιρικό ελάχιστο.

2.2.4. *Ικανή Συνθήκη Κυρτότητας*

Έστω μία κυρτή συνάρτηση $f : E \subseteq R^n \to R$ *η οποία είναι* $f \in C^1$ *στο κυρτό* E. *Αν σε ένα σημείο* $x^* \in E, \nabla f(x^*) = 0$, τότε το x^* *δίνει σφαιρικό ελάχιστο της συνάρτησης στο* E.

Κάθε αλγόριθμος αρχίζει από μία ή περισσότερες αρχικές προσεγγίσεις $x_0 \in E \subseteq R^n$ του ελαχίστου της συνάρτησης $f(x)$, το οποίο επιτυγχάνεται στο σημείο $x^* \in E$. Στη συνέχεια παράγεται μια ακολουθία σημείων $\{x_k\}, x_k \in E, k = 1,2,...$ η οποία καθώς το $k \to \infty$ συγκλίνει προς το x^* και κατ'επέκταση $x_k \to x^*$. Η γενική ιδέα κάθε αλγόριθμου από αυτούς που θα παρατεθούν στη συνέχεια είναι πολύ απλή και θα μπορούσε να γίνει πιο κατανοητή με το ακόλουθο παράδειγμα. Έστω ότι ψάχνουμε κατά μήκος ενός μονοπατιού, ξεκινώντας από το τρέχον σημείο, δηλαδή κατά μήκος κάποιας ευθείας μέσα στο χώρο μεταβλητότητας της αντικειμενικής συνάρτησης. Ο τρόπος με τον οποίο ορίζεται η διεύθυνση αυτής της ευθείας (Line Search problem), το μήκος του βήματος που θα κάνουμε ως προς κάθε άξονα για να εντοπίσουμε την επόμενη προσέγγιση του βέλτιστου σημείου, οι πληροφορίες που παρέχονται από τη συνάρτηση και τις παραγώγους της, διαφοροποιούν κατά ένα μεγάλο ποσοστό και τις μεθόδους. Η γενική εικόνα ενός αλγορίθμου βελτιστοποίησης μπορεί συνοπτικά να πάρει την ακόλουθη μορφή.

2.2.5. *Γενική Μορφή Αλγορίθμου Βελτιστοποίησης*

Βήμα 1°: Εκκίνηση με μία αρχική προσέγγιση x_0.

Βήμα 2°: Στην k *επανάληψη το τρέχον σημείο είναι το* $x_k \in E$.

Βήμα 3°: Κατασκευάζουμε μία καμπύλη $Q_k(\lambda_k)$, *με μεταβλητή τη* λ_k *και με τις εξής ιδιότητες:*

Το x_k *προκύπτει σαν τιμή της* $Q_k(\lambda_k)$ *για* $\lambda_k = 0$.

Για λ_k *θετικό και αρκετά μικρό, να ισχύει* $f(Q_k(\lambda_k)) \leq f(x_k)$.

Συνήθως η καμπύλη $Q_k(\lambda_k)$ *είναι μία ευθεία* $x_k + \lambda_k s_k$ *όπου το* s_k *είναι ένα διάνυσμα* nx1 *το οποίο καθορίζει τη διεύθυνση έρευνας για την εύρεση ενός ελαχίστου (direction of search) το οποίο μπορεί να είναι και μοναδιαίο.*

Βήμα 4°: Χρησιμοποιούμε κατάλληλη διαδικασία επιλογής του βήματος (Step Size Procedure, Line Search), καθορισμού δηλαδή, της κατάλληλης τιμής για το λ_k *επάνω στην ευθεία μετακίνησης (Line Search problem).*

Βήμα 5°: Θέτουμε $x_{k+1} = x_k + \lambda_k s_k$ *ή γενικά* $x_{k+1} = Q_k(\lambda_k)$

Βήμα 6°: Ελέγχουμε κάποιο κριτήριο σύγκλισης και ανακυκλώνουμε αν δεν είναι αληθές.

Όπως παρατηρείται στον παραπάνω αλγόριθμο, το πρώτο πρόβλημα είναι η εύρεση και πάλι αρχικής προσέγγισης. Στη συνέχεια για κάθε ανακύκλωση χρειάζεται να αποφασιστεί από ποιό μονοπάτι θα γίνει η μετάβαση, δηλαδή ποιά θα είναι η ευθεία που αναφέρεται στο 3° βήμα. Επιλέγουμε διάφορους τρόπους μετακίνησης οι οποίοι να δείχνουν προς κάποια διεύθυνση στην οποία η αντικειμενική συνάρτηση να μειώνεται, ή να αυξάνεται αν θέλουμε να βρούμε το μέγιστο. Το επόμενο πρόβλημα που προκύπτει είναι ο καθορισμός του βήματος μετακίνησης ως προς κάθε άξονα επάνω στην ευθεία όπως αναφέρεται στο 4° βήμα του αλγορίθμου. Ο εντοπισμός αυτός γίνεται με την επιλογή της κατάλληλης διαδικασίας που μπορεί να χρησιμοποιηθεί σε κάθε πρόβλημα, δηλαδή το κατάλληλο SSP. Εντοπίζεται η επόμενη προσέγγιση του βέλτιστου σημείου, ελέγχεται κριτήριο σύγκλισης και επαναλαμβάνεται. Πριν γίνει αναλυτική αναφορά στην ανάλυση των μεθόδων επιλογής βήματος, ακολουθεί μία συνοπτική παρουσία τους. Όλες επιλύουν το ακόλουθο μερικό πρόβλημα (Line Search Problem), που παρουσιάζεται στον Αλγόριθμο Βελτιστοποίησης στο 4° Βήμα.

2.2.6. *Μερικό Πρόβλημα Εύρεσης του Μήκους του Βήματος* λ_k

Δίνεται ένα σημείο $x_k \in E$ *και μία διεύθυνση* s_k. *Το κατάλληλο* λ_k *που πρέπει να χρησιμοποιηθεί για να γίνει μετάβαση στο επόμενο σημείο* x_{k+1}, *προκύπτει από τη λύση του*

$$\underset{\lambda_k \geq 0}{Minimize} \ f\left(x_k + \lambda_k s_k\right)$$

έτσι ώστε να ισχύει ότι

$$\lambda_k \in \left\{\lambda_k \geq 0 / x_k + \lambda_k s_k \in E\right\}$$

Δηλαδή, το κατάλληλο λ_k είναι αυτό που ελαχιστοποιεί τη συνάρτηση $f\left(x_k + \lambda_k s_k\right)$ του λ_k, για τα συγκεκριμένα x_k και s_k. Οι διαδικασίες που μπορούν να χρησιμοποιηθούν είναι οι ακόλουθες:

- SSP1: Θέτει ως λ_k ένα τοπικό ελάχιστο για το πρόβλημα (δηλαδή το $\underset{\lambda_k \geq 0}{Minimize} \ f\left(x_k + \lambda_k s_k\right)$)
- SSP2: Θέτει ως λ_k το σφαιρικό ελάχιστο για το πρόβλημα
- SSP3: Θέτει ως λ_k οποιοδήποτε τοπικό ελάχιστο του προβλήματος, το οποίο διασφαλίζει ταυτοχρόνως ότι $f\left(x_k + \lambda_k s_k\right) \leq f\left(x_k\right)$.

Προφανώς, όταν η συνάρτηση $f(x)$ είναι αυστηρά κυρτή οι τρεις αυτές διαδικασίες συμπίπτουν. Μάλιστα, η περίπτωση αυτή είναι και η μοναδική όπου υπάρχει εγγύηση εύρεσης λύσης στο μερικό πρόβλημα και ομαλής εφαρμογής μετά του γενικού αλγορίθμου.

Ο εντοπισμός του βήματος λ_k που ελαχιστοποιεί την $f(x_k+\lambda_k s_k)$ μπορεί να γίνει με αναλυτική μέθοδο αν αυτό είναι δυνατόν, δηλαδή με παραγώγιση της $f(x_k+\lambda_k s_k)$ ως προς λ_k και εύρεση των ριζών της πρώτης παραγώγου για να βρεθούν τα ακρότατα. Μετά, από τα ακρότατα και εφόσον είναι δυνατόν, βρίσκονται αυτά που δίνουν ελάχιστο. Αυτό είναι πολύ εύκολο να γίνει όταν η συνάρτηση είναι τετραγωνική μορφή ή προσεγγίζεται από μία τετραγωνική μορφή όπως θα παρατηρηθεί στη συνέχεια.

Υπάρχουν αλγόριθμοι των οποίων ο ρυθμός σύγκλισης εξαρτάται άμεσα από τον τρόπο επιλογής του βήματος ενώ για άλλους αυτό δεν ισχύει. Οι τρείς αυτές διαδικασίες επίλυσης του μερικού προβλήματος $\underset{\lambda_k\geq 0}{Minimize}\ f(x_k+\lambda_k s_k)$, ονομάζονται βέλτιστες διαδικασίες επιλογής βήματος (Optimal Step Size Procedures, OSSP) και είναι πολύ σημαντικές για ότι αναφερθεί στη συνέχεια. Όταν θα αναφέρεται ότι θα βρεθεί το βήμα λ_k με OSSP θα σημαίνει ότι θα επιλύεται το πρόβλημα $\underset{\lambda_k\geq 0}{Minimize}\ f(x_k+\lambda_k s_k)$ χρησιμοποιώντας μία από τις SSP1, SSP2, SSP3.

Στο σημείο αυτό θα παρατεθούν ορισμένες ενδιαφέρουσες ιδιότητες που προκύπτουν αν υπάρχει μία αντικειμενική συνάρτηση η οποία είναι τετραγωνική μορφή ή προσεγγίζεται από μία τετραγωνική μορφή. Τότε η ανάπτυξη της συνάρτησης γύρω από ένα τρέχον σημείο x_k, σύμφωνα με τον τύπο του Taylor είναι:

$$f(x)\approx f(x_k)+\nabla^T f(x_k)(x-x_k)+\frac{1}{2}(x-x_k)^T\nabla^2 f(x_k)(x-x_k)$$

Αν στη συνάρτηση αυτή γίνεται ελαχιστοποίηση κατά μήκος μίας διεύθυνσης s_k τότε σύμφωνα με τη Γενική Μορφή Αλγορίθμου Βελτιστοποίησης στο 5° Βήμα , είναι $x_{k+1} = x_k + \lambda_k s_k$. Άρα, $\lambda_k s_k = x_{k+1} - x_k$.

Αντικαθιστώντας στην

$f(x) \approx f(x_k) + \nabla^T f(x_k)(x - x_k) + \frac{1}{2}(x - x_k)^T \nabla^2 f(x_k)(x - x_k)$ για $x = x_{k+1}$, θα γίνει:

$$f(x) \approx f(x_k) + \nabla^T f(x_k)\lambda_k s_k + \frac{1}{2}\lambda_k s_k^T \nabla^2 f(x_k)\lambda_k s_k$$

Σύμφωνα με το μερικό πρόβλημα εύρεσης του λ_k της

$\underset{\lambda_k \geq 0}{Minimize}\ f(x_k + \lambda_k s_k)$ και τα SSP1, SSP2, SSP3 θα πρέπει να βρεθεί ένα ελάχιστο της παραπάνω συνάρτησης με μεταβλητή το λ_k δηλαδή με διαδικασία OSSP. Όλα τα υπόλοιπα στοιχεία της συνάρτησης είναι γνωστά διανύσματα. Στο σημείο αυτό, επειδή η συνάρτηση είναι τετραγωνική μορφή, με αναλυτικά βήματα, γίνεται παραγώγιση της συνάρτησης η οποία εξισώνεται με το μηδέν για να βρεθούν τα ακρότατα σημεία ως προς λ_k. Έτσι η συνάρτηση μετά την παραγώγιση γίνεται:

$$\frac{df(x_k + \lambda_k s_k)}{d\lambda_k} = \nabla^T f(x_k)s_k + s_k^T H \lambda_k s_k = 0$$

Αν λυθεί ως προς λ_k προκύπτει

$$\lambda_k = -\frac{\nabla^T f(x_k)s_k}{s_k^T H s_k}$$

Η σχέση που προέκυψε είναι καθοριστική για ότι ακολουθήσει, διότι ουσιαστικά παρέχει το μήκος βήματος για την εύρεση του επόμενου σημείου στο γενικό αλγόριθμο, όταν η συνάρτηση είναι τετραγωνική μορφή ή προσεγγίζεται από μία τετραγωνική μορφή. Δηλαδή, καλύπτει τη λογική των SSP1, SSP2, SSP3 και επομένως τον εντοπισμό του μήκους βήματος με μία βέλτιστη διαδικασία επιλογής, OSSP. Στις περιπτώσεις όπου η αντικειμενική συνάρτηση δεν είναι τετραγωνική μορφή, τότε ακολουθείται κανονικά η

$\underset{\lambda_k \geq 0}{Minimize}\ f(x_k + \lambda_k s_k)$ ελαχιστοποιώντας, είτε αναλυτικά είτε με αριθμητική μέθοδο.

Επίσης η τετραγωνική μορφή γράφεται

$$f(x) = a + x^T b + \frac{1}{2} x^T H x$$

και παραγωγίζοντας δίνει για την κλίση

$$\nabla f(x) = b + Hx$$

και για κάθε σημείο x_k είναι

$$\nabla f(x_k) = b + Hx_k$$

Αν γίνει αντικατάσταση της $\nabla f(x_k) = b + Hx_k$ στην

$\dfrac{df(x_k + \lambda_k s_k)}{d\lambda_k} = \nabla^T f(x_k) s_k + s_k^T H \lambda_k s_k = 0$ θα προκύψει ότι

$$(b + Hx_k)^T s_k + s_k^T H \lambda_k s_k = 0$$

Στη συνέχεια γίνεται αντικατάσταση και το $\lambda_k s_k = x_{k+1} - x_k$ και προκύπτει

$$s_k^T (b + Hx_k) + s_k^T H (x_{k+1} - x_k) = 0$$

που δίνει

$$s_k^T b + s_k^T Hx_k + s_k^T Hx_{k+1} - s_k^T Hx_k = 0$$

οπότε

$$s_k^T (b + Hx_{k+1}) = s_k^T \nabla f(x_{k+1}) = 0$$

Άρα αποδείχτηκε μία άλλη ιδιότητα της διαδικασίας βελτιστοποίησης σε μία τετραγωνική μορφή. Η κλίση στην επόμενη προσέγγιση x_{k+1} είναι κάθετη στην προηγούμενη διεύθυνση ελαχιστοποίησης s_k με την οποία εντοπίστηκε.

Στο σημείο αυτό θα παρουσιαστούν δύο παραδείγματα που κάνουν κατανοητή την εφαρμογή της γενικής μορφής του αλγορίθμου βελτιστοποίησης σε προβλήματα οικονομικής φύσεως.

<u>ΠΑΡΑΔΕΙΓΜΑ 1°</u>:

Να βρεθεί το ελάχιστο της συνάρτησης $f(x,y)=x^2+y^2$ με αρχικό σημείο το $x_0=\begin{pmatrix}2\\1\end{pmatrix}$ και διεύθυνση έρευνας $\begin{pmatrix}-1\\-1\end{pmatrix}$.

<u>ΛΥΣΗ</u>:

Είναι $x_0=\begin{pmatrix}2\\1\end{pmatrix}$, $s_0=\begin{pmatrix}-1\\-1\end{pmatrix}$. Άρα $x_1=x_0+\lambda_0 s_0=\begin{pmatrix}2\\1\end{pmatrix}+\lambda_0\begin{pmatrix}-1\\-1\end{pmatrix}=\begin{pmatrix}2-\lambda_0\\1-\lambda_0\end{pmatrix}$

Για την εύρεση του λ_0 πρέπει να λυθεί το πρόβλημα :

$$\min_{\lambda_0\geq 0} imize\, f\left(x_0+\lambda_0 s_0\right)$$

Έχουμε

$$f\left(x_0+\lambda_0 s_0\right)=f\left(2-\lambda_0,1-\lambda_0\right)=\left(2-\lambda_0\right)^2+\left(1-\lambda_0\right)^2=4-4\lambda_0+\lambda_0^2+1-2\lambda_0+\lambda_0^2=2\lambda_0^2-6\lambda_0+5$$

Το ελάχιστο δίνεται για $\lambda_0=-\frac{-6}{2*2}=\frac{6}{4}=1,5$

Άρα

$$x_1=\begin{pmatrix}0,5\\-0,5\end{pmatrix} \text{ και } f\left(x_1\right)=0,5\leq f\left(x_0\right)=5$$

Αν $s_1=\begin{pmatrix}-1\\-1\end{pmatrix}$ τότε :

$$x_2=x_1+\lambda_1 s_1=\begin{pmatrix}0,5-\lambda_1\\-0,5-\lambda_1\end{pmatrix}$$

Για να βρεθεί το λ_1 πρέπει να λυθεί το

$$\min_{\lambda_0\geq 0} imize\, f\left(x_1+\lambda_1 s_1\right)$$

Έχουμε

$$f\left(x_1+\lambda_1 s_1\right)=\left(0,5-\lambda_1\right)^2+\left(-0,5-\lambda_1\right)^2=0,25-\lambda_1+\lambda_1^2+0,25+\lambda_1+\lambda_1^2=0,5+2\lambda_1^2$$

Άρα υπάρχει ελάχιστο για $\lambda_1 = 0$ και συνεπώς $x_2 = x_1$ και στο σημείο αυτό σταματάμε τον αλγόριθμο.

ΠΑΡΑΔΕΙΓΜΑ 2^o:

Έστω ότι υπάρχει ένας καταναλωτής που έχει στη διάθεσή του $M = 24$ χρηματικές μονάδες τις οποίες σκοπεύει να καταναλώσει για την προμήθεια τριών αγαθών των οποίων οι τιμές είναι $P_x = 2$, $P_y = 1$ και $P_z = 4$.
Έστω επίσης ότι εκτιμήθηκε η συνάρτηση ωφέλειάς του που είναι $U(x, y, z) = xyz$. Ζητείται να υπολογιστούν τα x, y και z στη θέση όπου μεγιστοποιεί την ωφέλειά του. Να υπολογιστεί η μέγιστη αυτή ωφέλεια.

ΛΥΣΗ :

Παρατηρείται ότι η συνάρτηση ωφέλειας του καταναλωτή μεγαλώνει όσο αυξάνουν οι τιμές των x, y και z. Συνεπώς αυτή θα λαμβάνει τη μέγιστη τιμή (δεδομένων των περιορισμών) όταν

$$2x + y + 4z = 24 \Leftrightarrow y = 24 - 2x - 4z$$

Άρα ζητείται το μέγιστο της

$$U = xyz = x(24 - 2x - 4z)z = 24xz - 2x^2z - 4xz^2$$

ή αντίστοιχα το ελάχιστο της συνάρτησης

$$f = -U = 2x^2z + 4xz^2 - 24xz$$

Αν επιλεγεί για αρχικό σημείο το $x_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, δηλαδή $x = 1, z = 0$ και για διεύθυνση έρευνας το $s_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ τότε προκύπτει

$$x_1 = x_0 + \lambda_0 s_0 = \begin{pmatrix} 1 \\ \lambda_0 \end{pmatrix}$$

Για την εύρεση του λ_0 πρέπει να λυθεί το

$$\min_{\lambda_0 \geq 0} imize\, f\left(x_0 + \lambda_0 s_0\right)$$

Είναι

$$f\left(x_0 + \lambda_0 s_0\right) = 4\lambda^2 - 22\lambda$$

Το ελάχιστο βρίσκεται για

$$\lambda_0 = \frac{-22}{2*4} = 2,75$$

Συνεπώς

$$x_1 = \begin{pmatrix} 1 \\ 2,75 \end{pmatrix} \text{ και } f\left(x_1\right) = -30,75 \leq f\left(x_0\right) = 0$$

2.3. Εμπειρικές Μέθοδοι

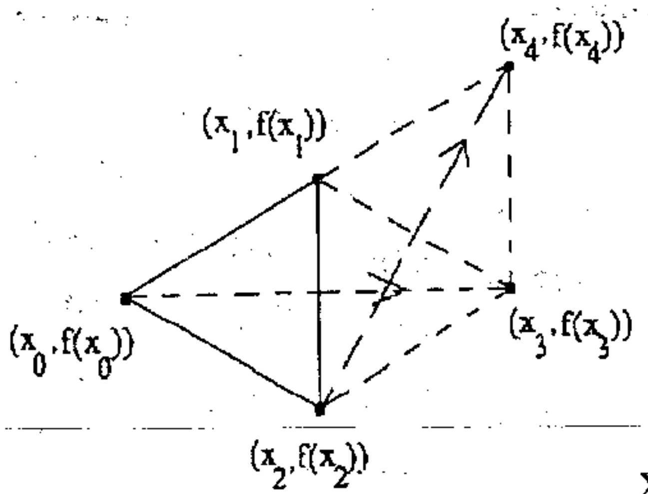
Στο σημείο αυτό θα πρέπει να γίνει αναφορά σε ελάχιστα στοιχεία που αφορούν πολυδιάστατες μεθόδους που αναπτύχθηκαν παλαιότερα σε εμπειρική βάση, με ελλιπές θεωρητικό υπόβαθρο και χρησιμοποιούσαν κυρίως μόνο τις τιμές της συνάρτησης. Τέτοιες μέθοδοι είναι:

- Παραλλαγές Εντοπισμού διαστήματος / Διαχωρισμού.
- Παραγωγή σημείων με τυχαίο τρόπο.
- Μέθοδος κανονικού simplex.
- Μέθοδος εναλλαγής μεταβλητών.

Στην πρώτη περίπτωση και ειδικά σε μικρή διάσταση $n=2,3$, υπάρχει προσπάθεια εφαρμογής μεθόδου όπως είναι για παράδειγμα η διχοτόμηση σε κάθε άξονα - μεταβλητή. Έτσι γίνεται προσπάθεια να βρεθούν περιοχές, (αντί διαστήματα), στις οποίες πιθανόν να ανήκει το βέλτιστο και οι οποίες πρέπει να μικραίνουν όσο είναι δυνατόν.

Στη δεύτερη περίπτωση παράγονται με τυχαίο τρόπο σημεία μέσα σε μία περιοχή βελτιστοποίησης της συνάρτησης και επιλέγεται το σημείο με την καλύτερη τιμή. Μοιάζει δηλαδή με τη γενική αρχή δοκιμασίας και λάθους. Οι μέθοδοι αυτές είναι οι λιγότερο ικανές αλλά αποδίδουν καλά σε ψηφιακούς και αναλογικούς υπολογιστές.

Η μέθοδος του κανονικού simplex ανήκει στις μεθόδους που δεν χρησιμοποιούν παραγώγους και οφείλεται στους Spendley, Hext & Himsworth (1962). Κατασκευάζεται κανονικά simplex με κάποια διαδικασία. Στο σημείο αυτό πρέπει να αναφερθεί ότι το κανονικό simplex είναι ένα σύνολο από $n+1$ σημεία τα οποία ισαπέχουν μεταξύ τους μέσα στον R^n. Για παράδειγμα για $n=2$ το κανονικό simplex είναι ένα ισόπλευρο τρίγωνο. Για $n=3$ έχουμε ένα κανονικό τετράεδρο κλπ. Έστω λοιπόν ότι εφαρμόζεται η μέθοδος σε ένα δισδιάστατο πρόβλημα, τότε ξεκινάει με τρία ισαπέχοντα σημεία x_0, x_1, x_2 τα οποία σχηματίζουν ένα ισόπλευρο τρίγωνο στο επίπεδο. Από τα τρία σημεία επιλέγεται αυτό που δίνει τη μεγαλύτερη τιμή και χρησιμοποιώντας σαν άξονα συμμετρίας το ευθύγραμμο τμήμα που ορίζουν τα άλλα δύο, βρίσκεται το νέο σημείο στη διαδικασία που είναι το συμμετρικό του και ονομάζεται αντανάκλαση. Συνεχίζοντας με αυτό το τρόπο, επιλέγεται το επόμενο ισόπλευρο τρίγωνο που σχηματίζει ο προηγούμενος άξονας συμμετρίας και το σημείο που προκύπτει, όπως παρατηρείται και στο Σχήμα 19



Σχήμα 19

Αν όμως στη συνέχεια, το τελευταίο σημείο έχει τη μεγαλύτερη τιμή της συνάρτησης, θα πρέπει να ξαναγυρίσει πίσω στο προηγούμενο τρίγωνο. Αυτό αποφεύγεται με διάφορες αλλαγές στον αλγόριθμο. Για παράδειγμα, (Fletcher, 1987), αν ένα σημείο εμφανιστεί μέσα στο τρέχον simplex για κάποιο αριθμό ανακυκλώσεων, τότε στην επόμενη επανάληψη επιλέγεται ως άξονας συμμετρίας ένας άξονας στη μισή απόσταση από το σημείο προς τα δύο γειτονικά του, αντί αυτόν που ορίζεται ακριβώς από τα δύο γειτονικά του. Μία βελτίωση της μεθόδου αυτής είναι η μέθοδος του ευέλικτου πολύεδρου (flexible polyhedron) των Nelder & Mead (1965) οι οποίοι αφήνουν το πολύεδρο να παίρνει διάφορα σχήματα δηλαδή παύει να είναι αναγκαστικά simplex. Η μέθοδος ανήκει στις βελτιωμένες μεθόδους που δεν χρησιμοποιούν παραγώγους για πολυδιάστατα προβλήματα.

Η τελευταία περίπτωση είναι η μέθοδος εναλλαγής μεταβλητών. Εδώ σε ένα n-διάστατο πρόβλημα, στην κάθε επανάληψη k όπου $k = 1,2,...,n$, μόνο η συντεταγμένη x_k αλλάζει ενώ οι άλλες παραμένουν σταθερές. Μετά από n βήματα όλες οι συντεταγμένες έχουν αλλάξει από μία φορά και ο κύκλος επαναλαμβάνεται. Το βασικό μειονέκτημα της μεθόδου αυτής είναι ότι δεν λαμβάνονται υπόψη τυχόν αλληλεπιδράσεις μεταξύ των συντεταγμένων οπότε μία αλλαγή ως προς ένα άξονα αναπόφευκτα επηρεάζει το προηγούμενο βέλτιστο κ.ο.κ. άρα και τη τιμή της συνάρτησης. Για παράδειγμα, έστω ότι στο προηγούμενο βήμα, γινόταν αναζήτηση κατά μήκος του άξονα των x_1 διατηρώντας τα άλλα σταθερά και προέκυπτε μία λύση x_2 η οποία είναι βέλτιστη ως προς τις λύσεις παράλληλα του άξονα x_1. Στη συνέχεια γίνεται αναζήτηση κατά μήκος ή παράλληλα του άξονα x_2 για να βρεθεί το επόμενο x_3. Τυχόν συσχέτιση των τιμών μεταξύ x_1 και x_2 θα καταστρέψει τη σημασία της προηγούμενης προσέγγισης μια και το καινούργιο x_2 επηρεάζει τη τιμή του x_1.

Στις μεθόδους που ενδεικτικά αναφέρθηκαν στην παράγραφο αυτή, έχουν γίνει πολλές παραλλαγές και βελτιώσεις μερικές μάλιστα από αυτές τις παραλλαγές οδήγησαν σε επιτυχείς μεθόδους κυρίως στην κατηγορία αυτών που δεν χρησιμοποιούν παραγώγους, το βασικό τους όμως μειονέκτημα είναι η έλλειψη ικανοποιητικού μαθηματικού υπόβαθρου που να στηρίζει τα αποτελέσματά τους. Τα τελευταία χρόνια η χρήση τους μειώνεται συνεχώς.

2.4. Η μέθοδος της μεγαλύτερης αλλαγής

Όπως αναφέρθηκε παραπάνω, η κλίση (gradient) μίας αντικειμενικής συνάρτησης σε ένα σημείο x_0 (όπως ορίστηκε παραπάνω), είναι ένα διάνυσμα κάθετο στη καμπύλη (contour) της επιφάνειας $f(x)$ που περνάει από το σημείο x_0, και ορίζει την κατεύθυνση της μέγιστης τοπικής αύξησης της $f(x)$ στο σημείο αυτό (steepest ascent direction). Έτσι, η αντίθετη διεύθυνση ορίζει τη διεύθυνση της μέγιστης τοπικής μείωσης (steepest descent direction). Η μέθοδος αυτή εκμεταλλεύεται την ιδιότητα αυτή, για να βρει το ελάχιστο της συνάρτησης σε ένα σύνολο και μπορεί κανείς να αναζητήσει τις ρίζες της από την εποχή του Γάλλου Μαθηματικού Cauchy (1847) και των μελετών του. Γι' αυτό η μέθοδος της μεγαλύτερης αλλαγής ή μέθοδος της μέγιστης πτώσης ονομάζεται και μέθοδος του Cauchy. Τα προβλήματα που έχουν αναπτυχθεί για τις μεθόδους βελτιστοποίησης στις μονοδιάστατες αντικειμενικές συναρτήσεις, παραμένουν και εδώ σε σχέση με τα τοπικά-σφαιρικά ελάχιστα, την εύστοχη επιλογή αρχικής προσέγγισης, τη σύγκλιση κλπ. Αν όμως υπάρχει καλή εμπειρική γνώση για τη συμπεριφορά της συνάρτησης, τα πιθανά ανώμαλα ή ακρότατα σημεία της και την περιοχή που πέφτει το ελάχιστο, τότε είναι δυνατόν με μεγάλη ακρίβεια να βρεθεί ένα τοπικό ελάχιστο. Ακολούθως αναφέρεται ποια είναι ακριβώς η λογική της μεθόδου.

Έστω μία συνάρτηση $f : E \subseteq R^n \to R$ και μία τρέχουσα προσέγγιση $x_k \in E$ του βέλτιστου σημείου που είναι το x^* κατά τα γνωστά. Μία καλύτερη προσέγγιση του ελαχίστου πρέπει να αναζητηθεί προς τη κατεύθυνση που δείχνει η αντίθετη της κλίσης στο x_k δηλαδή στη κατεύθυνση $-\nabla f(x_k)$. Αν γίνει κανονικοποίηση, το μοναδιαίο διάνυσμα στην κατεύθυνση αυτή είναι προφανώς το

$$-\frac{\nabla f(x_k)}{\left\|\nabla f(x_k)\right\|}$$

Έτσι, αν οριστεί σαν διεύθυνση μετάβασης στο επόμενο σημείο η

$$s_k = -\frac{\nabla f(x_k)}{\left\|\nabla f(x_k)\right\|}$$

σύμφωνα με τη Γενική Μορφή Αλγορίθμου Βελτιστοποίησης στο 5^{o} Βήμα, είναι

$$x_{k+1} = x_k + \lambda_k s_k$$

άρα προκύπτει ότι

$$x_{k+1} = x_k + \lambda_k\left(-\frac{\nabla f(x_k)}{\left\|\nabla f(x_k)\right\|}\right) = x_k - \lambda_k \frac{\nabla f(x_k)}{\left\|\nabla f(x_k)\right\|}$$

Ο τρόπος με τον οποίο επιλέγεται το βήμα λ_k δηλαδή το πώς λύνεται το μερικό πρόβλημα $\underset{\lambda_k \geq 0}{Minimize}\ f(x_k + \lambda_k s_k)$, ο τρόπος ορισμού της norm, όπως επίσης και το κριτήριο τερματισμού δίνουν και διαφορετικές παραλλαγές στη μέθοδο.

Η norm είναι η Ευκλείδεια εξίσωση $\|x\| = \sqrt{\sum_{j=1}^{n}(x_j)^2}$, όπως έχει σημειωθεί παραπάνω, και δεν θα μπορούσε να είναι διαφορετική σ' αυτό το σημείο για να υπάρχει η διεύθυνση της μεγαλύτερης αλλαγής.

Το βήμα λ_k επιλέγεται με τη μέθοδο SSP1 που είναι βέλτιστη μέθοδος επιλογής βήματος (OSSP). Αυτό σημαίνει ότι πρέπει να βρεθεί μία λύση για το μερικό πρόβλημα $\underset{\lambda_k \geq 0}{Minimize}\ f(x_k + \lambda_k s_k)$ (ελαχιστοποίηση ως προς λ_k), παραγωγίζοντας την $f(x_k + \lambda_k s_k)$ και βρίσκοντας τις ρίζες της παραγώγου ως προς λ_k δηλαδή

$$\frac{df(x_k + \lambda_k s_k)}{d\lambda_k} = \frac{df\left(x_k - \lambda_k \frac{\nabla f(x_k)}{\left\|\nabla f(x_k)\right\|}\right)}{d\lambda_k} = 0$$

Σύμφωνα με τον Goldstein (1962) η μέθοδος της μέγιστης αλλαγής με την SSP1 όπως στην παραπάνω εξίσωση για την επιλογή του βήματος, συγκλίνει καθώς το $k \to \infty$, στο ελάχιστο, αν είναι κυρτή και υπάρχει η παράγωγος της μέχρι τρίτης τάξης. Εξίσου ενδιαφέρουσα είναι η περίπτωση όπου η εφαρμογή της παραπάνω SSP1 δεν είναι εφικτή αναλυτικά, οπότε εφαρμόζεται μία μονοδιάστατη μέθοδος για τον εντοπισμό ενός βέλτιστου λ_k στο πρόβλημα $\underset{\lambda_k \geq 0}{Minimize} \; f(x_k + \lambda_k s_k)$, για παράδειγμα τη μέθοδο των χρυσών τομών.

Επίσης, είναι δυνατόν στον αλγόριθμο, να χρησιμοποιηθεί σταθερό ή προσαρμοζόμενο από επανάληψη σε επανάληψη βήμα λ_k. Για παράδειγμα αν τεθεί σταθερό ίσο με τη μονάδα ως προς κάθε διάσταση και για κάθε επανάληψη, ή να αλλάζει η τιμή του από βήμα σε βήμα ανάλογα με την εμπειρία που υπάρχει για το πρόβλημα. Στην περίπτωση αυτή, η τιμή του πρέπει να επιλέγεται με προσοχή, έτσι ώστε να αποφεύγονται καταστάσεις όπως αύξηση της τιμής της συνάρτησης από μία τρέχουσα προσέγγιση στην επόμενή της, (μεγαλύτερο λ_k απ' όσο χρειάζεται), ή επίσης μεγάλο αριθμό επαναλήψεων του αλγορίθμου (μικρότερο λ_k απ' όσο πρέπει).

Τέλος, όπως έχει αναφερθεί και παραπάνω, σε ότι αφορά το λ_k, αν η συνάρτηση είναι τετραγωνικής μορφής ή προσεγγίζεται από μία τέτοια, τότε το μήκος του βήματος μπορεί να δοθεί κατευθείαν από τον τύπο

$$\lambda_k = -\frac{\nabla^T f(x_k) s_k}{s_k^T H s_k}$$

Αν αντικατασταθεί η $s_k = -\frac{\nabla f(x_k)}{\|\nabla f(x_k)\|}$ στην $\lambda_k = -\frac{\nabla^T f(x_k) s_k}{s_k^T H s_k}$ προκύπτει

$$\lambda_k = -\frac{\nabla^T f(x_k)\left(-\frac{\nabla f(x_k)}{\|\nabla f(x_k)\|}\right)}{\left(-\frac{\nabla f(x_k)}{\|\nabla f(x_k)\|}\right)^T H \left(-\frac{\nabla f(x_k)}{\|\nabla f(x_k)\|}\right)}$$

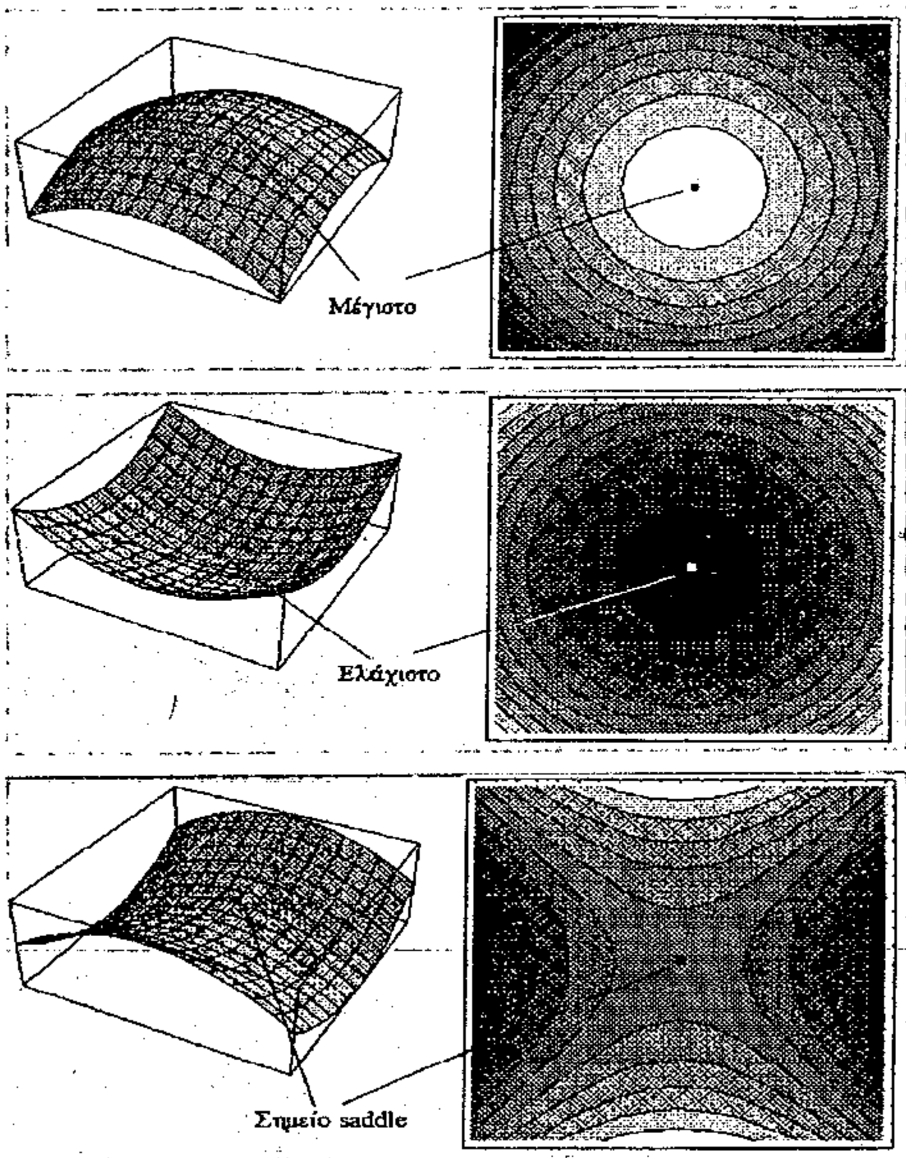
η οποία ύστερα από κάποιες πράξεις γίνεται

$$\lambda_k = \frac{\frac{\|\nabla f(x_k)\|^2}{\|\nabla f(x_k)\|}}{\frac{\nabla^T f(x_k) H \nabla f(x_k)}{\|\nabla f(x_k)\|^2}} = \frac{\|\nabla f(x_k)\|^3}{\nabla^T f(x_k) H \nabla f(x_k)}$$

Άρα στην περίπτωση της τετραγωνικής μορφής η εφαρμογή του SSP1 είναι ισοδύναμη με τη χρήση του παραπάνω τύπου και δεν χρειάζεται καμία άλλη ενέργεια για την εύρεση του λ_k.

Όσον αφορά τον τερματισμό, η ανακύκλωση του αλγορίθμου για τον εντοπισμό ενός ελαχίστου σταματάει με βάση διάφορους κανόνες άλλοτε βασισμένους στο x, άλλοτε στο $\nabla f(x)$ και άλλοτε στο $f(x)$. Γενικά υπάρχει δυνατότητα να σταματήσει όταν η $\nabla f(x)$ πλησιάζει αρκετά στο μηδέν μια και αυτή είναι η χαρακτηριστική ιδιότητα του ελαχίστου στην περιοχή αυτή.

Η δυνατότητα μίας αλγοριθμικής διαδικασίας να συγκλίνει αποτελεσματικά και με αποδεκτό ρυθμό εξαρτάται άμεσα από την αρχική προσέγγιση, τα επόμενα σημεία, τα κριτήρια σύγκλισης, τη μορφή της συνάρτησης κλπ. όμως η κλίση είναι θεωρητικά μηδέν και σε άλλα σημεία, για παράδειγμα τα σημεία ανωμαλίας της επιφάνειας $f(x)$ (saddle points). Στο Σχήμα 20 παρουσιάζονται τρείς διαφορετικές περιπτώσεις σημείων.



Σχήμα 20

Αν κάτι τέτοιο προκύψει στη συνέχεια, θα πρέπει να εφαρμοστεί μία διαδικασία μέσω της οποίας θα ξεφεύγει από το σημείο αυτό. Ένας τρόπος για να επιτευχθεί αυτό είναι ο έλεγχος στη συνέχεια του Πίνακα Hessian ο οποίος θα πρέπει να είναι θετικά ορισμένος. Αν το σημείο είναι saddle η συνάρτηση έχει μηδενική κλίση άλλα ο πίνακας δεν είναι θετικά ορισμένος.

Στο σημείο αυτό παρατίθεται συνοπτικά ο αλγόριθμος με τα βήματά του:

2.4.1. *Μέθοδος της Μεγαλύτερης Αλλαγής, Cauchy*

Έστω μία πραγματική συνάρτηση $f \in C^1$ *στο* $E \subseteq R^n$. *Για την εύρεση ενός σημείου* x^* *το οποίο δίνει τοπικό ελάχιστο μέσα στο* E *επιλέγεται μία αρχική προσέγγιση* $x_0 \in E$ *και δημιουργείται μία ακολουθία σημείων* $\{x_k\}$ *η οποία συγκλίνει στο βέλτιστο σημείο. Για να προχωρήσει από το σημείο* x_k *στο* x_{k+1} *ακολουθείται η εξής διαδικασία:*

<u>Βήμα 1°</u>: Υπολογίζεται η κλίση $\nabla f(x_k)$

<u>Βήμα 2°</u>: Υπολογίζεται η διεύθυνση μετάβασης $s_k = -\dfrac{\nabla f(x_k)}{\|\nabla f(x_k)\|}$

<u>Βήμα 3°</u>: Λύνεται το πρόβλημα $\underset{\lambda_k \geq 0}{\text{Minimize}}\ f(x_k + \lambda_k s_k)$ *για να βρεθεί το βήμα* λ_k.

Η λύση μπορεί να βρεθεί

Από την $\dfrac{df\left(x_k - \lambda_k \dfrac{\nabla f(x_1)}{\|\nabla f(x_1)\|}\right)}{d\lambda_k} = 0$

Από την εφαρμογή κάποιας μονοδιάστατης μεθόδου

Αν η συνάρτηση είναι τετραγωνική μορφή τότε $\lambda_k = -\dfrac{\nabla^T f(x_k)s_k}{s_k^T H s_k}$ *δηλαδή*

$$\lambda_k = \frac{\|\nabla f(x_1)\|^3}{\nabla^T f(x_k) H \nabla f(x_k)}$$

<u>Βήμα 4°</u>: Εντοπίζεται το επόμενο σημείο $x_{k+1} = x_k + \lambda_k s_k$

<u>Βήμα 5°</u>: Ελέγχεται κριτήριο σύγκλισης. Αν είναι αληθές τέλος, διαφορετικά γίνεται ανακύκλωση.

<u>ΠΑΡΑΔΕΙΓΜΑ 1°</u>:

Στο πρόβλημα αυτό πρέπει να βρεθεί το ελάχιστο της συνάρτησης $f(x) = x_1^2 + 5x_2^2 + x_3^2 - 4$ με τη μέθοδο Cauchy. Επιλέγεται σαν αρχική προσέγγιση το σημείο $x_0 = (2,2,2)^T$. Η κλίση σε κάθε σημείο x θα είναι $\nabla f(x) = \begin{pmatrix} 2x_1 \\ 10x_2 \\ 2x_3 \end{pmatrix}$ αφού

$\frac{\partial f(x)}{\partial x_1} = 2x_1$, $\frac{\partial f(x)}{\partial x_2} = 10x_2$ και $\frac{\partial f(x)}{\partial x_3} = 2x_3$. Άρα

$\left\|\nabla f(x)\right\| = \sqrt{\nabla^T f(x) \nabla f(x)} = 2\sqrt{x_1^2 + 25x_2^2 + x_3^2}$ που για $x_0 = (2,2,2)^T$ δίνει

$\left\|\nabla f(x_0)\right\| = 2\sqrt{108}$. Συνεπώς η διεύθυνση μετάβασης στο επόμενο σημείο είναι η

$-\nabla f(x_0) / \left\|\nabla f(x_0)\right\|$ που δίνει $s_0 = -\frac{1}{2\sqrt{108}}\begin{pmatrix} 4 \\ 20 \\ 4 \end{pmatrix} = -\frac{1}{\sqrt{108}}\begin{pmatrix} 2 \\ 10 \\ 2 \end{pmatrix}$. Οπότε από το 4^{o}

Βήμα του παραπάνω αλγορίθμου προκύπτει ότι $x_1 = x_0 + \lambda_0 s_0$. Μετά από την

αντικατάσταση η σχέση αυτή δίνει $x_1 = (2,2,2)^T - \frac{1}{\sqrt{108}}\lambda_0 (2,10,2)^T$ οπότε

$$x_1 = x_0 + \lambda_0 s_0 = \left(2 - \frac{2}{\sqrt{108}}\lambda_0, 2 - \frac{10}{\sqrt{108}}\lambda_0, 2 - \frac{2}{\sqrt{108}}\lambda_0\right)^T.$$

Στο σημείο αυτό γίνεται αντιληπτό ότι το πρόβλημα έχει φτάσει στο 3^{o} Βήμα του αλγορίθμου του Cauchy και από την $f(x_0 + \lambda_0 s_0)$ με αντικατάσταση της τελευταίας εξισώσεις προκύπτει ότι

$$f(x_0 + \lambda_0 s_0) = \left(2 - \frac{2}{\sqrt{108}}\lambda_0\right)^2 + 5\left(2 - \frac{10}{\sqrt{108}}\lambda_0\right)^2 + \left(2 - \frac{2}{\sqrt{108}}\lambda_0\right)^2 - 4$$

που δίνει

$$f(x_0 + \lambda_0 s_0) = 2\left(2 - \frac{2}{\sqrt{108}}\lambda_0\right)^2 + 5\left(2 - \frac{10}{\sqrt{108}}\lambda_0\right)^2 - 4$$

Εφαρμόζοντας το i. από το 3^{o} Βήμα του αλγορίθμου, πρέπει να γίνει παραγώγισης της προηγούμενης συνάρτησης ως προς λ_0 έτσι ώστε να βρεθούν οι ρίζες της παραγώγου και το ελάχιστο της $f(x_0 + \lambda_0 s_0)$. Παραγωγίζοντας προκύπτει:

$$\frac{df\left(x_0-\lambda_0\frac{\nabla f(x_0)}{\|\nabla f(x_0)\|}\right)}{d\lambda_0}=4\left(2-\frac{2}{\sqrt{108}}\lambda_0\right)\left(-\frac{2}{\sqrt{108}}\right)+10\left(2-\frac{10}{\sqrt{108}}\lambda_0\right)\left(-\frac{10}{\sqrt{108}}\right)=0$$

Άρα $2\left(2-\frac{2}{\sqrt{108}}\lambda_0\right)+25\left(2-\frac{10}{\sqrt{108}}\lambda_0\right)=0$ που δίνει

$4-\frac{4}{\sqrt{108}}\lambda_0+50-\frac{250}{\sqrt{108}}\lambda_0=0$ δηλαδή $\frac{254}{\sqrt{108}}\lambda_0=54$ και λύνοντας ως προς το ζητούμενο προκύπτει $\lambda_0\approx 2.21$. Έτσι για να γίνει μετάβαση στην επόμενη προσέγγιση, γίνεται αντικατάσταση το βήμα $\lambda_0\approx 2.21$ στην εξίσωση

$x_1=x_0+\lambda_0 s_0=\left(2-\frac{2}{\sqrt{108}}\lambda_0, 2-\frac{10}{\sqrt{108}}\lambda_0, 2-\frac{2}{\sqrt{108}}\lambda_0\right)^T$ και προκύπτει ότι

$x_1=(1.575, 0.127, 1.575)^T$. Στο σημείο αυτό τελειώνει μία ανακύκλωση (αφού ελεγχθεί και κάποιο κριτήριο σύγκλισης). Η διαδικασία συνεχίζεται με τρέχον σημείο το x_1 για να εντοπιστεί το x_2. Στον Πίνακα 1 που ακολουθεί δίνονται μερικές επαναλήψεις του αλγορίθμου.

k	x_k^T	λ_k	$\nabla^T f(x_k)$	$f(x_k)$
0	(2, 2, 2)	2.2100	(4, 20, 4)	24
1	(1.575, -0.127, 1.575)	1.7807	(3.15, -1.27, 3.15)	1.042
2	(0.354, 0.361, 0.364)	0.5100	(0.728, 3.61, 0.728)	-3.08
3	(0.265, -0.130, 0.265)	0.1824	(0.53, -1.3, 0.53)	-3.775
4	(0.190, 0.032, 0.190)	0.1550	(0.38, 0.32, 0.38)	-3.923
5	(0.090, -0.040, 0.090)	0.0610	(0.18, -0.4, 0.18)	-3.976
6	(0.060, 0.010, 0.060)		(0.12, 0.1, 0.12)	-3.999

Πίνακας 1

Συνεχίζοντας τις επαναλήψεις θα βρεθεί το βέλτιστο σημείο $x^* = (0,0,0)^T$ με βέλτιστη τιμή $f(x^*) = -4$.

Σ'αυτό το πρόβλημα θα γινόταν εναλλακτικά, για να βρεθεί το λ_k, στο 3^o Βήμα του αλγορίθμου του Cauchy να χρησιμοποιηθεί η περίπτωση iii. Επειδή είναι τετραγωνική μορφή η αντικειμενική συνάρτηση. Αν παραγωγιστεί για δεύτερη φορά η συνάρτηση, ο Πίνακας Hessian της συνάρτησης για κάθε σημείο x είναι ο εξής:

$H = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 2 \end{bmatrix}$. Από τον τύπο $\lambda_k = -\dfrac{\nabla^T f(x_k) s_k}{s_k^T H s_k} = \dfrac{\|\nabla f(x_k)\|^3}{\nabla^T f(x_k) H \nabla f(x_k)}$ αν γίνει αντικατάσταση τα αντίστοιχα διανύσματα για το x_0 που βρέθηκε προηγουμένως θα βρεθεί και πάλι το ίδιο $\lambda_0 = \dfrac{(27*2^2)^{3/2}}{127*2^2} = \dfrac{108^{3/2}}{508} \approx 2.21$.

Στο παράδειγμα αυτό η σύγκλιση γίνεται αρκετά αργά. Σε αυτό συμβάλλει αρνητικά η διαφορετική κλίμακα των συντελεστών των μεταβλητών δηλαδή το 5 που υπάρχει για το x_2 ενώ στα άλλα είναι μονάδες.

<u>ΠΑΡΑΔΕΙΓΜΑ 2^o</u>:

Έστω ότι πρέπει να βρεθεί ένα ελάχιστο της συνάρτησης $f(x) = x_1^2 + x_2^2 + x_3^2 - 4$. Η συνάρτηση αυτή διαφέρει από την συνάρτηση του προηγούμενου παραδείγματος μόνο στον τελεστή του x_2 ο οποίος είναι μονάδα. Έτσι μπορεί να βρεθεί αμέσως ότι $\nabla f(x) = (2x_1, 2x_2, 2x_3)^T$ ενώ ο Πίνακας Hessian είναι:

$$H = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

και για το ίδιο αρχικό σημείο $x_0 = (2,2,2)^T$ προκύπτει $\nabla f(x_0) = (4,4,4)^T$ ενώ

$\|\nabla f(x)\| = 2\sqrt{x_1^2 + x_2^2 + x_3^2}$ που δίνει $\|\nabla f(x_0)\| = 4\sqrt{3}$.

Στη συνέχεια προκύπτει ότι

$$x_1 = x_0 + \lambda_0 s_0 = x_0 - \lambda_0 \frac{\nabla f(x_0)}{\|\nabla f(x_0)\|} = (2,2,2)^T - \lambda_0 \frac{(4,4,4)^T}{4\sqrt{3}} = \left(2 - \frac{\sqrt{3}}{3}\lambda_0, 2 - \frac{\sqrt{3}}{3}\lambda_0, 2 - \frac{\sqrt{3}}{3}\lambda_0\right)^T$$

Άρα $f(x_0 + \lambda_0 s_0) = 3\left(2 - \frac{\sqrt{3}}{3}\lambda_0\right)^2 - 4$, οπότε

$$\frac{df\left(x_0 - \lambda_0 \frac{\nabla f(x_0)}{\|\nabla f(x_0)\|}\right)}{d\lambda_0} = 6\left(2 - \frac{\sqrt{3}}{3}\lambda_0\right)\left(-\frac{\sqrt{3}}{3}\right) = 0$$

το οποίο δίνει $\lambda_0 = 2\sqrt{3}$.

Θα γινόταν βέβαια να βρεθεί αμέσως το λ_0 από τον τύπο iii του 3^{ου} Βήματος στον αλγόριθμο του Cauchy. Αν γίνει αντικατάσταση στον τύπο

$$\lambda_k = -\frac{\nabla^T f(x_k) s_k}{s_k^T H s_k} = \frac{\|\nabla f(x_k)\|^3}{\nabla^T f(x_k) H \nabla f(x_k)}$$

την κλίση και τον Πίνακα Hessian θα προκύψει ότι

$$\lambda_k = \frac{\left(\sqrt{x_1^2 + x_2^2 + x_3^2}\right)^3}{(2x_1, 2x_2, 2x_3)\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}\begin{pmatrix} 2x_1 \\ 2x_2 \\ 2x_3 \end{pmatrix}} = \sqrt{x_1^2 + x_2^2 + x_3^2}$$

Ο τύπος αυτός δίνει, για $k = 0$ δηλαδή για το αρχικό σημείο x_0, και πάλι το $\lambda_0 = 2\sqrt{3}$. Αντικαθιστώντας το λ_0 στην

$$x_1 = x_0 + \lambda_0 s_0 = x_0 - \lambda_0 \frac{\nabla f(x_0)}{\|\nabla f(x_0)\|} = (2,2,2)^T - \lambda_0 \frac{(4,4,4)^T}{4\sqrt{3}} = \left(2 - \frac{\sqrt{3}}{3}\lambda_0, 2 - \frac{\sqrt{3}}{3}\lambda_0, 2 - \frac{\sqrt{3}}{3}\lambda_0\right)^T$$

παίρνεται η επόμενη προσέγγιση $x_1 = (0,0,0)^T = x^*$, που είναι η βέλτιστη.

Όπως παρατηρήθηκε στο Παράδειγμα 2 βρέθηκε σε ένα βήμα το βέλτιστο σημείο. Η διαφορά του από το Παράδειγμα 1 είναι η ανυπαρξία προβλημάτων κλίμακας που αποτελούν ένα από τα βασικά μειονεκτήματα της μεθόδου της μεγαλύτερης αλλαγής. Ότι δηλαδή επηρεάζεται πολύ το πλήθος των επαναλήψεων από τη διαφορά κλίμακας των τιμών ανά συντεταγμένη.

Στο σημείο αυτό θα παρουσιαστούν τρία παραδείγματα που κάνουν κατανοητή την εφαρμογή της μεθόδου της μεγαλύτερης αλλαγής, Caushy σε προβλήματα οικονομικής φύσεως.

ΠΑΡΑΔΕΙΓΜΑ 1° :

Να βρεθεί το ελάχιστο της συνάρτησης $f(x,y)=x^2+y^2$ με τη μέθοδο της μεγαλύτερης αλλαγής του Cauchy και αρχικό σημείο το $x_0=\begin{pmatrix}2\\1\end{pmatrix}$

ΛΥΣΗ :

Βήμα 1° : Υπολογίζουμε την κλίση

$$\nabla f(x_0)=\begin{pmatrix}2*2\\2*1\end{pmatrix}=\begin{pmatrix}4\\1\end{pmatrix}$$

Βήμα 2° : Υπολογίζουμε τη διεύθυνση μετάβασης

$$s_0=-\frac{\nabla f(x_0)}{\left\|\nabla f(x_0)\right\|}=-\frac{1}{\sqrt{5}}\begin{pmatrix}4\\1\end{pmatrix}$$

Άρα

$$x_1=x_0+\lambda_0 s_0=\begin{pmatrix}2-\frac{4}{\sqrt{5}}\lambda_0\\1-\frac{1}{\sqrt{5}}\lambda_0\end{pmatrix}$$

Βήμα 3° : Λύνουμε το πρόβλημα

$$\min_{\lambda_0 \geq 0} imize\ f\left(x_0 + \lambda_0 s_0\right)$$

Είναι

$$f\left(x_0+\lambda_0 s_0\right)=\left(2-\frac{4}{\sqrt{5}}\lambda_0\right)^2+\left(1-\frac{1}{\sqrt{5}}\lambda_0\right)^2=4-\frac{16}{\sqrt{5}}\lambda_0+\frac{16}{5}\lambda_0^2+1-\frac{2}{\sqrt{5}}\lambda_0+\frac{1}{5}\lambda_0^2=\frac{17}{5}\lambda_0^2-\frac{18}{\sqrt{5}}\lambda_0+5$$

Το ελάχιστο δίνεται για $\lambda_0=-\dfrac{-\dfrac{18}{\sqrt{5}}}{2\dfrac{17}{5}}=\dfrac{90}{34\sqrt{5}}\simeq 1,18$

Βήμα 4^o:

Άρα $x_1=\begin{pmatrix} 2-\dfrac{4}{\sqrt{5}}1,18 \\ 1-\dfrac{1}{\sqrt{5}}1,18 \end{pmatrix}=\begin{pmatrix} -1,22 \\ 0,47 \end{pmatrix}$ και $f\left(x_1\right)=1,7093 \leq f\left(x_0\right)=5$

Στο σημείο αυτό τελειώνει μία ανακύκλωση (αφού ελεγχθεί και κάποιο κριτήριο σύγκλισης). Μπορούμε να συνεχίσουμε για να βρούμε μία καλύτερη προσέγγιση x_2.

Βήμα 1^o: Υπολογίζουμε την κλίση

$$\nabla f\left(x_0\right)=\begin{pmatrix} -2,44 \\ 0,94 \end{pmatrix}$$

Βήμα 2^o: Υπολογίζουμε τη διεύθυνση μετάβασης

$$s_0=-\frac{\nabla f\left(x_0\right)}{\left\|\nabla f\left(x_0\right)\right\|}=-\frac{1}{\sqrt{6,83}}\begin{pmatrix} -2,44 \\ 0,94 \end{pmatrix}$$

Άρα

$$x_2=x_1+\lambda_1 s_1=\begin{pmatrix} -1,22+\dfrac{2,44}{\sqrt{6,38}}\lambda_0 \\ 0,47-\dfrac{0,94}{\sqrt{6,38}}\lambda_0 \end{pmatrix}$$

Βήμα 3^o: Λύνουμε το πρόβλημα

$$\min_{\lambda_1 \geq 0} imize\ f\left(x_1+\lambda_1 s_1\right)$$

Είναι

$$f\left(x_1+\lambda_1 s_1\right)=\left(\frac{2,44}{\sqrt{6,38}}\lambda_1-1,22\right)^2+\left(0,47-\frac{0,94}{\sqrt{6,38}}\lambda_1\right)^2=\frac{5,95}{6,38}\lambda_1^2-\frac{5,95}{\sqrt{6,38}}\lambda_1+1,48+$$

$$+0,22-\frac{0,88}{\sqrt{6,38}}\lambda_1+\frac{0,88}{6,38}\lambda_1^2=\frac{6,82}{6,38}\lambda_1^2-\frac{6,83}{\sqrt{6,38}}\lambda_1+1,7$$

Το ελάχιστο δίνεται για

$$\lambda_1=\frac{-\frac{6,83}{\sqrt{6,38}}}{2\frac{6,82}{6,38}}\simeq 1,26$$

Βήμα 4°:

Άρα $x_2=\begin{pmatrix}-0,002\\0,001\end{pmatrix}$ και $f\left(x_2\right)=0,0005\le f\left(x_1\right)=1,7093$

ΠΑΡΑΔΕΙΓΜΑ 2°:

Να βρεθεί το ελάχιστο της συνάρτησης $f(x,y,z)=2x^2+y^2+z^2-2xy-2x+1$ με τη μέθοδο της μεγαλύτερης αλλαγής του Cauchy με αρχική προσέγγιση το σημείο $x_0=\begin{pmatrix}0\\8\end{pmatrix}$.

ΛΥΣΗ:

Βήμα 1°: Υπολογίζουμε την κλίση

$$\nabla f\left(x_0\right)=\begin{pmatrix}-2\\0\\0\end{pmatrix}$$

Βήμα 2°: Υπολογίζουμε τη διεύθυνση μετάβασης

$$s_0=-\frac{\nabla f\left(x_1\right)}{\left\|\nabla f\left(x_1\right)\right\|}=\begin{pmatrix}1\\0\\0\end{pmatrix}$$

Άρα

$$x_1 = x_0 + \lambda_0 s_0 = \begin{pmatrix} \lambda \\ 0 \\ 0 \end{pmatrix}$$

<u>Βήμα 3°</u>: Λύνουμε το πρόβλημα

$$\min_{\lambda_0 \geq 0} imize\, f\left(x_0 + \lambda_0 s_0\right)$$

Είναι

$$f\left(x_0 + \lambda_0 s_0\right) = 2\lambda^2 - 2\lambda + 1$$

Το ελάχιστο δίνεται για

$$\lambda_0 = -\frac{-2}{2*2} = \frac{1}{2}$$

<u>Βήμα 4°</u>:

Άρα $x_1 = \begin{pmatrix} 1/2 \\ 0 \\ 0 \end{pmatrix}$ και $f\left(x_1\right) = \frac{1}{2} \leq f\left(x_0\right) = 1$

Στο σημείο αυτό τελειώνει μία ανακύκλωση (αφού ελεγχθεί και κάποιο κριτήριο σύγκλισης). Μπορούμε να συνεχίσουμε για να βρούμε μία καλύτερη προσέγγιση.

<u>Βήμα 1°</u>: Υπολογίζουμε την κλίση

$$\nabla f\left(x_1\right) = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$$

<u>Βήμα 2°</u>: Υπολογίζουμε τη διεύθυνση μετάβασης

$$s_1 = -\frac{\nabla f\left(x_1\right)}{\left\|\nabla f\left(x_1\right)\right\|} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Άρα

$$x_2 = x_1 + \lambda_1 s_1 = \begin{pmatrix} 1/2 \\ \lambda_1 \\ 0 \end{pmatrix}$$

<u>Βήμα 3°</u>: Λύνουμε το πρόβλημα

$$\underset{\lambda_1 \geq 0}{\min imize}\, f\left(x_1 + \lambda_1 s_1\right)$$

Είναι

$$f\left(x_1 + \lambda_1 s_1\right) = \lambda_1^2 - \lambda_1 + \frac{1}{2}$$

Το ελάχιστο δίνεται για

$$\lambda_1 = \frac{1}{2}$$

<u>Βήμα 4°</u>:

Άρα $x_2 = \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \end{pmatrix}$ και $f\left(x_2\right) = \frac{1}{4} \leq f\left(x_1\right) = \frac{1}{2}$

Στο σημείο αυτό τελειώνει άλλη μία ανακύκλωση. Στον παρακάτω πίνακα δίνονται μερικές επαναλήψεις του αλγορίθμου.

k	x_k^T	λ_k	$\nabla^T f\left(x_k\right)$	$f\left(x_k\right)$
0	$(0,0,0)$	$1/2$	$(-2,0,0)$	1
1	$\left(1/2,0,0\right)$	$1/2$	$(0,-1,0)$	$1/2$
2	$\left(1/2,1/2,0\right)$	$1/4$	$(-1,0,0)$	$1/4$
3	$\left(3/4,1/2,0\right)$	$1/4$	$\left(0,-1/2,0\right)$	$1/8$
4	$\left(3/4,3/4,0\right)$	$1/8$	$\left(-1/2,0,0\right)$	$1/16$
5	$\left(7/8,3/4,0\right)$		$\left(0,-1/4,0\right)$	$1/32$

Συνεχίζοντας τις επαναλήψεις θα προσεγγίζεται όλο και περισσότερο το βέλτιστο σημείο

$$x^{*}(x,y,z)^T = (1,1,0)^T$$

Αυτό είναι το βέλτιστο σημείο γιατί

$$2x^2 + y^2 + z^2 - 2xy - 2x + 1 = (x-1)^2 + (x-y)^2 + z^2$$

ΠΑΡΑΔΕΙΓΜΑ 3°:

Ένας καταναλωτής με $M = 10$ χρηματικές μονάδες θέλει να αγοράσει 3 αγαθά με τιμές $P_x = 2, P_y = P_z = 1$. Αν η συνάρτηση ωφέλειας του είναι $U = 2xy + xz$ να βρεθούν τα x, y, z ώστε αυτή να μεγιστοποιηθεί.

ΛΥΣΗ:

Το μέγιστο (δεδομένων των περιορισμών) θα βρίσκεται όταν ο καταναλωτής ξοδέψει όλα τα χρήματά του. Όταν δηλαδή ισχύει

$$2x + y + z = 10 \Leftrightarrow z = 10 - 2x - y$$

πρέπει να μεγιστοποιηθεί η

$$U = 2xy + xz = 2xy + x(10 - 2x - y) = xy - 2x^2 + 10x$$

ή αντίστοιχα να ελαχιστοποιηθεί η

$$f = -U = 2x^2 - 10x - xy$$

Έστω $x_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ μια αρχική προσέγγιση.

Βήμα 1°: Υπολογίζουμε την κλίση

$$\nabla f(x_0) = \begin{pmatrix} -10 \\ 0 \end{pmatrix}$$

Βήμα 2°: Υπολογίζουμε τη διεύθυνση μετάβασης

$$s_0 = -\frac{\nabla f(x_0)}{\|\nabla f(x_0)\|} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Άρα

$$x_1 = \begin{pmatrix} \lambda \\ 0 \end{pmatrix}$$

<u>Βήμα 3°</u> : Λύνουμε το πρόβλημα

$$\underset{\lambda_0 \geq 0}{\min imize}\, f(x_0 + \lambda_0 s_0)$$

Είναι

$$f(x_0 + \lambda_0 s_0) = 2\lambda_0^2 - 10\lambda_0$$

Άρα το ελάχιστο δίνεται για

$$\lambda_0 = \frac{-10}{2*2} = 2,5$$

<u>Βήμα 4°</u> : Το νέο σημείο είναι το

$$x_1 = \begin{pmatrix} 2,5 \\ 0 \end{pmatrix} \text{ με } f(x_1) = -12,5 \leq f(x_0) = 0$$

Στο σημείο αυτό τελειώνει μια ανακύκλωση. Για τη νέα προσέγγιση θα πρέπει να ελεγχθεί κάποιο κριτήριο σύγκλισης. Θα πρέπει επίσης να ελεγχθεί αν ικανοποιεί τους περιορισμούς που θέτει το πρόβλημα. Σε αυτό το παράδειγμα πρέπει

$$x \geq 0, y \geq 0 \text{ και } z \geq 0 \text{ ή ισοδύναμα } 10 - 2x - y \geq 0$$

Για το x_1 προκύπτει ότι $x_1 = \begin{pmatrix} 2,5 \\ 0 \end{pmatrix}$

Άρα

$$x = 2,5, y = 0 \text{ και } z = 5$$

-Οπότε ικανοποιούνται οι περιορισμοί του προβλήματος.

Σε αντίθετη περίπτωση θα έπρεπε να επιλεγεί άλλο σημείο αρχικής προσέγγισης ή κάποιον άλλο αλγόριθμο (όπως για παράδειγμα τη γενική μορφή του αλγορίθμου βελτιστοποίησης όπου για παράδειγμα η διεύθυνση έρευνας δεν θα προερχόταν από την κλίση $\nabla f(x_0)$).

ΒΙΒΛΙΟΓΡΑΦΙΑ

ΞΕΝΟΓΛΩΣΣΗ ΒΙΒΛΙΟΓΡΑΦΙΑ

- Απειροστικός Λογισμός, Finney, Ross L., Thomas, George B., Πανεπιστημιακές Εκδόσεις Κρήτης 2004.
- Μαθηματική Θεωρία της Βελτιστοποίησης, Du, DingZhu, Pardalos, Panos M., Wu, Weili, Εκδόσεις Νέων Τεχνολογιών 2005.

ΕΛΛΗΝΙΚΗ ΒΙΒΛΙΟΓΡΑΦΙΑ

- Εφαρμοσμένα Μαθηματικά για Οικονομικές Επιστήμες, Πέκος, Γ.Δ., Θεσσαλονίκη 2003.
- Μη Γραμμικές Μέθοδοι Βελτιστοποίησης, Α. Γεωργίου, Π. Βασιλείου, Θεσσαλονίκη 1993.
- Εισαγωγή στον Διαφορικό Λογισμό Συναρτήσεων Πολλών Μεταβλητών, Καρανικόλας, Ν.Δ., 2004.
- Συναρτήσεις Πολλών Μεταβλητών-Μαθηματικά ΙΙ Τόμος Α', Αναστασάτος, Αναστασίου, Γάγαλης, Κομισόπουλος, Θεσσαλονίκη 2001.
- Λογισμός Συναρτήσεων Πολλών Μεταβλητών, Κρόκος, Ι.Π., Άρνος 2002
- Μαθηματικά ΙΙ (Διδακτικές Σημειώσεις), Δ. Γεωργίου, Πάτρα 2000.