2022-2023

ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ ΑΠΘ - Τεχνικές Βελτιστοποίησης

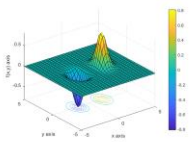
Φίλιππος Γερμανόπουλος

[2Η ΕΡΓΑΣΤΗΡΙΑΚΗ ΑΣΚΗΣΗ]

Ελαχιστοποίηση με χρήση Παραγώγων

ΘΕΜΑ 1

Η εργασία πραγματεύεται το πρόβλημα ελαχιστοποίησης μίας δοσμένης συνάρτησης πολλών μεταβλητών χωρίς περιορισμούς και χρησιμοποιούνται αλγόριθμοι που βασίζονται στην ιδέα της επαναληπτικής μεθόδου, έτσι ώστε $f(\chi\kappa+1) < f(\chi\kappa)$. Το πρώτο ζητούμενο της εργασίας ήταν η γραφική αναπαράσταση της $f(\chi,y) = x^5 e^{(-x^2-y^2)}$ και έχει την εξής μορφή:



ΘΕΜΑ 2

Σε αυτό το σημείο αξίζει να αναφερθεί πως και οι 3 παρακάτω αλγόριθμοι που χρησιμοποιούνται, αποτελούν αλγόριθμους κλίσης. Πολλοί από τους αλγόριθμους κλίσης είναι της μορφής :

$$X(\kappa+1)=\chi\kappa - \Delta\kappa \cdot \text{διάνυσμα κλίσης της } f(\chi\kappa)$$

Στο θέμα 2 υλοποιείται η μέθοδος της Μέγιστης Καθόδου που προκύπτει εάν Δκ=I ,και χρησιμοποιούνται ως αρχικά σημεία τα (x0,y0)= (0,0) , (-1,1) , (1,-1).
Το βήμα γκ επιλέγεται :

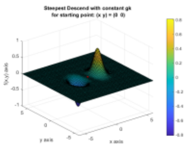
1. Αρχικά ως μία σταθερή τιμή.
2. Κατάλληλο ώστε να ελαχιστοποιεί την f(xk+γk*dk), δια μέσω της χρήσης αλγορίθμου της προηγούμενης εργαστηριακής άσκησης, και μάλιστα της χρυσής τομής.

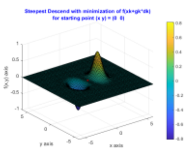
3. Βάσει του κανόνα Armijo.

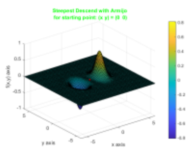
Σε όλες τις περιπτώσεις, ως σταθερά τερματισμού του αλγορίθμου, επιλέχθηκε η τιμή ε=0.01

I. **<u>Για αρχικό σημείο (x0,y0)= (0,0):</u>**

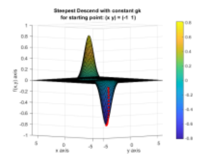
Ο αλγόριθμος σταματά στη πρώτη κιόλας επανάληψη, εφόσον ικανοποιείται εξ αρχής η συνθήκη τερματισμού του αλγορίθμου. Αυτό συμβαίνει φυσικά και για τις τρεις διαφορετικές περιπτώσεις επιλογής του γκ. Αυτό φυσικά συμβαίνει καθώς το σημείο (0,0) βρίσκεται πάνω στο επίπεδο και έχει μηδενική κλίση. Συνεπώς ο αλγόριθμος εγκλωβίζεται σε ένα τοπικό ελάχιστο και αδυνατεί να βρει το ολικό ελάχιστο.

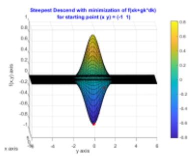

Steepest Descend with constant gk
for starting point: (x y) = (0 0)
f(x,y) axis
y axis
x axis

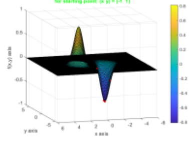

Steepest Descend with minimization of f(xk+gk*dk)
for starting point (x y) = (0 0)
f(x,y) axis
y axis
x axis


Steepest Descend with Armijo
for starting point: (x y) = (0 0)
f(x,y) axis
y axis
x axis
1
0.5
0
-0.5
-1
5
0
-5
-5
0
5
0.8
0.6
0.4
0.2
0
-0.2
-0.4

II. Για αρχικό σημείο (x0,y0)= (-1,1):







Η εύρεση του ολικού ελαχίστου είναι επιτυχής, ανεξαρτήτου μεθόδου επιλογής του γκ που χρησιμοποιείται. Ωστόσο φαίνεται καθαρά πως η ειδοποιός διαφορά τους είναι ο αριθμός των επαναλήψεων που απαιτούνται. **Ο αριθμός των επαναλήψεων (k) και η τελική ελάχιστη τιμή της f (minimum_point), όπως επίσης και όλα τα σημεία που προκύπτουν από τον αλγόριθμο κατά τη διάρκεια της υλοποίησής του (x_points, y_points,f_points) αποθηκεύονται και είναι διαθέσιμες στις αντίστοιχες μεταβλητές στον κώδικα MATLAB.** Ο αριθμός επαναλήψεων και η ελάχιστη τιμή της f που εντοπίζει ο αλγόριθμος τυπώνονται επίσης στο παράθυρο εντολών (command window) για κάθε σημείο.

Αναλυτικότερα, στην μέθοδο του σταθερού βήματος επιλέχθηκε το γκ= 0,1 ως μία σχετικά καλή επιλογή(αλλά όχι βέλτιστη). Δοκιμάστηκαν και επιλογές αρκετά μικρότερες, όπως 0.001, και αρκετά μεγαλύτερες, όπως γκ=10.Αυτές οι τιμές όμως είναι ανεπιθύμητες εφόσον δεν ικανοποιούν τη γεωμετρία της κίνησης και είτε αυξάνεται κατά πολύ ο αριθμός των επαναλήψεων, είτε η εύρεση ελαχίστου καθίσταται αδύνατη. Η ορθή επιλογή του γκ λοιπόν είναι κρίσιμης αξίας και απαιτεί εμπειρία και αρκετή εξοικείωση με την μέθοδο.

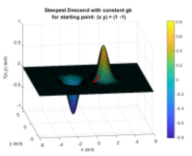
Ακόμη και με κατάλληλη επιλογή του γκ ωστόσο, η μέθοδος του σταθερού βήματος παραμένει μία αναποτελεσματική και αργή μέθοδος, που απαιτεί πολλές επαναλήψεις. Αντιθέτως, στην περίπτωση όπου η επιλογή του γκ γίνεται έτσι ώστε να ελαχιστοποιείται η f(χκ+γκ*dκ), ο αλγόριθμος τερματίζει αποτελεσματικότατα και με ελάχιστες επαναλήψεις. Αυτό συμβαδίζει και με τις θεωρητικές μας γνώσεις, όπου το διάνυσμα που συνδέει το χκ με το χ άστρο είναι συγγραμικό με το διάνυσμα κλίσης της f(χκ). Σε αυτό το σημείο χρησιμοποιήθηκε η μέθοδος της χρυσής τομής της προηγούμενης εργασίας, η οποία ελαχιστοποιεί μία κυρτή συνάρτηση και μπορεί να αποτελέσει βάση των αλγορίθμων εύρεσης ελαχίστου συναρτήσεων με περισσότερες μεταβλητές, όπως στην περίπτωσή μας. Η συνάρτηση MATLAB της χρυσής τομής τροποποιήθηκε ελαφρώς, ούτως ώστε να επιστρέφει μόνο την τιμή της ελαχιστοποίησης του γκ και όχι όλες τις τιμές που επέστρεφε στο πρώτο παραδοτέο.

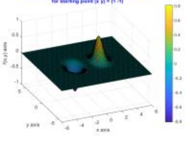
Όσον αφορά τον κανόνα Armijo, θυμόμαστε πως το βήμα μειώνεται διαδοχικά σύμφωνα με τον τύπο : γκ = s*β^(mκ) και το β ανήκει 0<β<1.

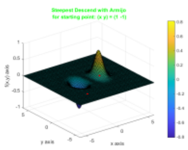
Το mκ είναι ο μικρότερος θετικός ακέραιος που ικανοποιεί το κριτήριο:
f(χκ)-f(χκ+1) >= - α*β*^(mκ)*s*transpose(dκ)*διάνυσμα κλίσης f(χκ).

Συνήθως το α ανήκει [10^(-5),10^(-1)] και το β στο [0,1 , 0,5], και για αυτό επιλέχθησαν οι τιμές **α=10^(-3), β=0,3 , s=10.** Το ελάχιστο βρέθηκε με σχετικά λίγες επαναλήψεις.

III. Για αρχικό σημείο (x0,y0)= (1,-1):


Steepest Descend with constant gk
for starting point: (x y) = (1 -1)
f(x,y) axis
y axis
x axis


Steepest Descend with minimization of f(xk+gk*dk)
for starting point (x y) = (1 -1)
f(x,y) axis
y axis
x axis



Για τούτο το σημείο εκκίνησης ο αλγόριθμος αποτυγχάνει να βρει το ολικό ελάχιστο, ανεξάρτητα από την μέθοδο με την οποία επιλέγεται το γκ. Και αυτό διότι μετά από λίγες επαναλήψεις, οδηγείται στο μηδενικό επίπεδο , όπου αποκτά κλίση 0 και εγκλωβίζεται ξανά στο τοπικό ελάχιστο, όπως στην περίπτωση όπου το αρχικό μας σημείο ήταν το (0,0).

ΘΕΜΑ 3

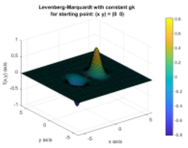
Στο θέμα 3 ζητούνται τα ίδια ερωτήματα αλλά με χρήση μεθόδου Newton. Η Newton , όμοια με την μέθοδο της μέγιστης καθόδου, προκύπτει εάν επιλέξουμε όπου Δκ τον αντίστροφο του εσσιανού της f(χκ). **Οφείλει να αναφερθεί η σημαντικότατη παρατήρηση ότι η εφαρμογή της μεθόδου Newton προϋποθέτει ο εσσιανός της f(χκ) να είναι θετικά ορισμένος.** Έτσι λοιπόν ο κώδικας MATLAB ελέγχει εάν ισχύει αυτή η προϋπόθεση, προτού εφαρμόσει τον αλγόριθμο. Ωστόσο, και για τα τρία σημεία εκκίνησης που επιλέχθηκαν , η μέθοδος Newton είναι αδύνατον να εφαρμοσθεί επειδή δεν ικανοποιείται η προκείμενη προϋπόθεση. Έτσι, ο αλγόριθμος τερματίζει αμέσως μετά τον έλεγχο αυτόν.

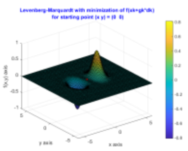
ΘΕΜΑ 4

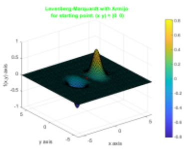
Στο θέμα 3 ζητούνται τα ίδια ερωτήματα αλλά με χρήση της μεθόδου Levenberg-Marquardt. Η μέθοδος αυτή είναι μία τροποποιημένη εκδοχή της μεθόδου Newton, όπου εφαρμόζεται και σε περιπτώσεις όπου ο εσσιανός της f(χκ) δεν είναι θετικά ορισμένος. Αυτό επιτυγχάνεται με τον ορισμό μίας παραμέτρου μκ που υπολογίζεται κατάλληλα σε κάθε επανάληψη, έτσι ώστε το άθροισμα του εσσιανού της f(χκ) με το γινόμενο της παραμέτρου μκ με τον μοναδιαίο πίνακα , να δίνουν έναν θετικά ορισμένο πίνακα.

- **Για αρχικό σημείο (x0,y0) = (0,0)**

Για το συγκεκριμένο σημείο ο αλγόριθμος τερματίζει λόγω του προαναφερθέντος εγκλωβισμού σε κρίσιμο σημείο, όπως και στις 2 προηγούμενες μεθόδους.

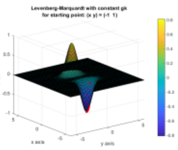

Levenberg-Marquardt with constant gk
for starting point: (x y) = (0 0)
f(x,y) axis
y axis
x axis
1
0.5
0
-0.5
-1
5
0
-5
-5
0
5
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8

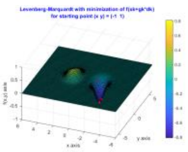

Levenberg-Marquardt with minimization of f(xk+gk*dk)
for starting point (x y) = (0 0)
f(x,y) axis
y axis
x axis
1
0.5
0
-0.5
-1
5
0
-5
-5
0
5
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8

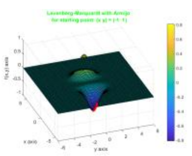


- **<u>Για αρχικό σημείο (x0,y0) = (-1,1)</u>**

Ομοίως με νωρίτερα, το σημείο αυτό μας οδηγεί επιτυχώς στην εύρεση του ολικού ελαχίστου για όλους τους τρόπους επιλογής γκ.


Levenberg-Marquardt with constant gk
for starting point: (x y) = (-1 1)
f(x,y) axis
x axis
y axis


Levenberg-Marquardt with minimization of f(xk+gk*dk)
for starting point (x y) = (-1 1)
f(x,y) axis
x axis
y axis

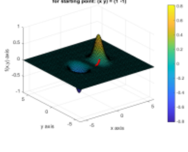


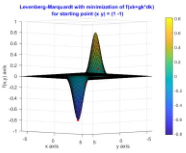
Στην περίπτωση διατήρησης σταθερού βήματος, επιλέχθηκε το γκ=0,1. Εδώ παρατηρούμε ότι φτάνουμε γρήγορα στην γειτονιά του ελαχίστου, αλλά μετά απαιτούνται πολλές επαναλήψεις για την πραγματική εύρεση αυτού. Αντίστοιχα, η υλοποίηση όπου το γκ επιλέγεται έτσι ώστε να ελαχιστοποιείται η f(χκ+γκ*dκ) είναι αποτελεσματικότερη και γρηγορότερη, όπως ακριβώς αναμέναμε. Από την δεύτερη κιόλας επανάληψη οδηγούμαστε κοντά στην περιοχή του ολικού ελαχίστου.

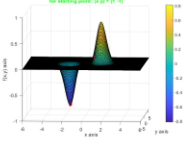
Για την μέθοδο Armijo, επιλέχθηκαν ακριβώς ίδιες τιμές για τις αναγκαίες παραμέτρους όπως και πριν : **α=10^(-3), β=0,3 , s=10**. Υλοποιήθηκε και μία συνάρτηση η οποία ελέγχει την ικανοποίηση της καλής λειτουργίας του αλγορίθμου και διασφαλίζουν πως το γκ δεν θα πάρει πολύ μεγάλες ή πολύ μικρές τιμές, προκειμένου να αποφευχθούν αποκλίσεις. Ωστόσο είναι πολύ σημαντική η κατάλληλη αρχικοποίηση του αρχικού βήματος s, καθώς για πολλές τιμές δεν ικανοποιούνται τα κριτήρια και τερματίζεται ο αλγόριθμος. Η επιθυμητή τιμή δεν εντοπίσθηκε, για αυτό και περιλαμβάνεται το κομμάτι του ελέγχου των κριτηρίων σε μορφή σχολίων στον κώδικα MATLAB.

- **<u>Για αρχικό σημείο (x0,y0) = (1,-1)</u>**

Για το συγκεκριμένο σημείο, για τους λόγους που ισχύουν στη μέθοδο Μέγιστης Καθόδου και στη μέθοδο Newton, ο αλγόριθμος εγκλωβίζεται σε τοπικό ελάχιστο και αποτυγχάνει στην προσπάθεια εύρεσης του ολικού ελαχίστου. Αυτήν την φορά όμως, αυτό συμβαίνει μόνο για την σταθερή επιλογή του γκ, ενώ για τις άλλες 2 περιπτώσεις το ελάχιστο βρίσκεται επιτυχώς.




Levenberg-Marquardt with minimization of f(xk+gk*dk)
for starting point (x y) = (1 -1)
f(x,y) axis
x axis
y axis


Levenberg-Marquardt with Armijo
for starting point: (x y) = (1 -1)
f(x,y) axis
x axis
y axis

<u>ΠΑΡΑΤΗΡΗΣΕΙΣ</u>

Η εργασία αυτή μελετά διαφορετικούς τρόπους εύρεσης ελαχίστου μίας συνάρτησης πολλών μεταβλητών για διαφορετικούς τρόπους καθορισμού του βήματος γκ και για πολλαπλά διαφορετικά σημεία εκκίνησης της αναζήτησης. Εξάγονται συμπεράσματα με αυτόν τον τρόπο για την αξία της θεωρητικής ανάλυσης, στην οποία πρέπει να βασιζόμαστε για την αντίληψη ,κατανόηση και επεξήγηση των πρακτικών εφαρμογών και αποτελεσμάτων.

Κάθε μέθοδος έχει θετικά στοιχεία και αδυναμίες, αναλόγως την εκάστοτε περίπτωση. Η μέθοδος της Μέγιστης Καθόδου είναι μία απλή μέθοδος που δεν απαιτεί μεγάλη υπολογιστική πολυπλοκότητα, αλλά οι ευθείες που ενώνουν τα διαδοχικά σημεία επιλογής του αλγορίθμου σχηματίζουν ορθή γωνία μεταξύ τους και οδηγούν σε πολλές επαναλήψεις , κάνοντάς την έτσι μία αργή μέθοδο. Η μέθοδος Newton είναι αρκετά περιοριστική, εφόσον εφαρμόζεται μόνο όταν ο Εσσιανός πίνακας της αντικειμενικής συνάρτησης είναι θετικά ορισμένος. Η μέθοδος Levenberg-Marquardt απαιτεί μεγαλύτερη υπολογιστική πολυπλοκότητα και αναλόγως το μέγεθος της χαρακτηριστικής τιμής μκ, προσεγγίζει την τόσο τη μέθοδο της Μέγιστης Καθόδου, όσο και τη μέθοδο Newton.

Η κάθε μία αναδεικνύει τα χαρίσματά της σε διαφορετικές περιπτώσεις και έτσι δεν υπάρχει κάποιο αντικειμενικό κριτήριο που να αποδεικνύουν την ανωτερότητα κάποιας μεθόδου έναντι των άλλων. Παρόλα αυτά, συμπεραίνουμε πως ο βέλτιστος τρόπος καθορισμού του βήματος γκ είναι αυτός που ελαχιστοποιεί την συνάρτηση f(χκ+γκ*dκ) καθώς προσδίδει στον αλγόριθμο μεγάλη ταχύτητα , μικρό αριθμό επαναλήψεων και ακρίβεια εντοπισμού του ολικού ελαχίστου. Η μέθοδος σταθερής τιμής γκ δύναται να είναι αποτελεσματική, αλλά απαιτεί μεγάλη εμπειρία για την κατάλληλη επιλογή του βήματος γκ. Κάτι τέτοιο είναι εξαιρετικά δύσκολο να επιτευχθεί από συνάρτηση σε συνάρτηση. Η μέθοδος Armijo επηρεάζεται από πολλές παραμέτρους και εξασφαλίζει πως η τιμή του βήματος γκ θα παραμείνει σε ένα επιθυμητό πεδίο τιμών, χωρίς να ξεφύγει και να γίνει πολύ μεγάλο.

Τέλος, είναι πολύ σημαντική η επιρροή του αρχικού σημείο εκκίνησης του αλγορίθμου στην σύγκλιση των αλγορίθμων. Ένα σημείο εκκίνησης μπορεί να οδηγήσει σε εσφαλμένες μετρήσεις, σε εγκλωβισμό του αλγορίθμου σε τοπικό ελάχιστο και έτσι ελλοχεύει ο κίνδυνος λήψης λανθασμένων συμπερασμάτων. Η υλοποίηση των αλγορίθμων όμως, με τη βοήθεια εργαλείων όπως αυτό της MATLAB, μας βοηθά να οπτικοποιήσουμε τη διαδικασία και με κριτική σκέψη, σε συνδυασμό με τα πορίσματα της θεωρίας, να ερμηνεύσουμε ορθώς τα αποτελέσματα.