



ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ

# Τεχνικές Βελτιστοποίησης Αναφορά

## Εργασία 2

Διακολουκάς Δημήτριος  
ΑΕΜ 10642

*Email: [ddiakolou@ece.auth.gr](mailto:ddiakolou@ece.auth.gr)*

# Περιεχόμενα

1	Ορισμός και σχεδιασμός συνάρτησης	2
2	Μέθοδος Μέγιστης Καθόδου (Steepest Descent)	4
3	Μέθοδος Newton	9
4	Μέθοδος Levenberg-Marquardt	14
5	Αποτελέσματα και Συμπεράσματα	19

# Κεφάλαιο 1

## Ορισμός και σχεδιασμός συνάρτησης

Ο στόχος αυτής της ανάλυσης είναι να παρουσιάσει και να σχολιάσει τα αποτελέσματα της ελαχιστοποίησης της συνάρτησης:

$$f(x, y) = x^5 e^{-x^2 - y^2}.$$

Για την επίλυση του προβλήματος χρησιμοποιήθηκαν τρεις μέθοδοι βελτιστοποίησης: η μέθοδος της μέγιστης καθόδου (Steepest Descent), η μέθοδος του Νεύτωνα (Newton), και η μέθοδος Levenberg-Marquardt.

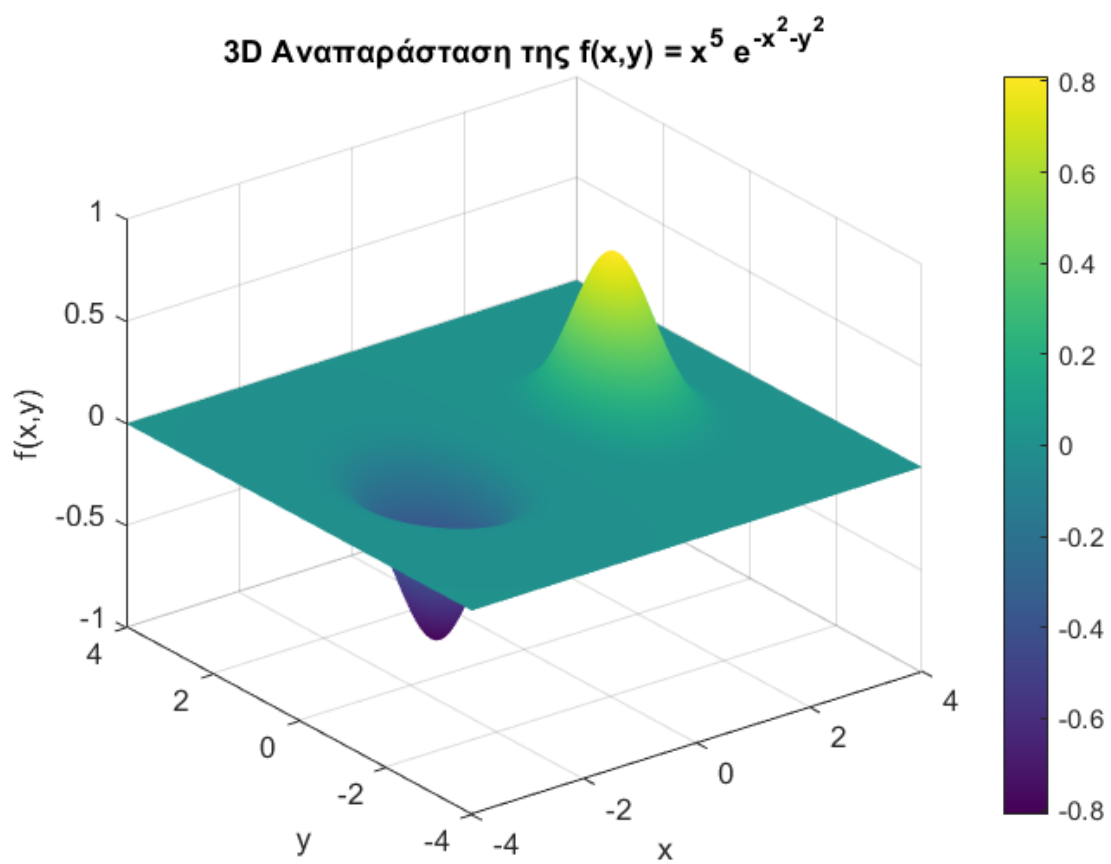
Η παρούσα εργασία επικεντρώνεται στην επίλυση προβλημάτων ελαχιστοποίησης συναρτήσεων πολλών μεταβλητών χωρίς περιορισμούς, χρησιμοποιώντας τις παραπάνω τρεις μεθόδους.

Η ανάλυση αποσκοπεί στη σύγκριση της απόδοσής τους ως προς την ταχύτητα σύγκλισης, την ακρίβεια, και τη σταθερότητα, λαμβάνοντας υπόψη διαφορετικές αρχικές συνθήκες και στρατηγικές προσαρμογής βημάτων.

Η υλοποίηση των αλγορίθμων πραγματοποιήθηκε σε περιβάλλον *Matlab*, και τα αποτελέσματα συνοδεύονται από σχολιασμό και εξαγωγή χρήσιμων συμπερασμάτων. Η εργασία απαιτεί τόσο τη θεωρητική κατανόηση των μεθόδων όσο και την πρακτική εφαρμογή τους, παρέχοντας μια ολοκληρωμένη εικόνα των προσεγγίσεων βελτιστοποίησης. Έτσι λοιπόν, όπως φαίνεται και στο Σχήμα 1.1 έχουμε την 3D αναπαράσταση της αντικειμενικής συνάρτησης.

Όπως αλλωστε φαίνεται και στο σχήμα, από την ανάλυση του τρισδιάστατου γραφήματος της συνάρτησης  $f$ , διαπιστώνεται ότι η ελάχιστη τιμή της εντοπίζεται για αρνητικές τιμές του  $x$ , κοντά στο σημείο όπου ο άξονας  $x$  τέμνει τον άξονα  $y$ . Από την άλλη πλευρά, ένα μέγιστο εμφανίζεται στις θετικές τιμές του  $x$  στην περιοχή κοντά στο σημείο όπου ο άξονας  $x$  τέμνει τον  $y$ .

Ο αντίστοιχος κώδικας βρίσκεται στο αρχείο **Part1.m**.



Σχήμα 1.1: Γενική εικόνα της μορφής της  $f$ .

## Κεφάλαιο 2

### Μέθοδος Μέγιστης Καθόδου (Steepest Descent)

Στο συγκεκριμένο ερώτημα, εξετάζεται η εφαρμογή της μεθόδου Μεγίστης Καθόδου για την ελαχιστοποίηση της συνάρτησης  $f(x, y)$ , λαμβάνοντας υπόψη διαφορετικές αρχικές συνθήκες και στρατηγικές επιλογής του βήματος  $\gamma_k$ . Στόχος είναι να αναλυθεί πώς οι παράμετροι αυτές επηρεάζουν την απόδοση του αλγορίθμου, τόσο ως προς τη σύγκλιση όσο και ως προς την ακρίβεια του τελικού αποτελέσματος.

Ειδικότερα, η μέθοδος εφαρμόζεται ξεκινώντας από τα αρχικά σημεία  $(x_0, y_0) = (0, 0)$ ,  $(-1, 1)$ , και  $(1, -1)$ , με τρεις διαφορετικές προσεγγίσεις για την επιλογή του  $\gamma_k$ :

1. Σταθερή τιμή, που ορίζεται από τον χρήστη (την όρισα 0.1),
2. Υπολογισμός για την ελαχιστοποίηση της  $f(x_k + \gamma_k d_k)$ ,
3. Χρήση του κανόνα Armijo.

Η ανάλυση αποσκοπεί στη σύγκριση των αποτελεσμάτων, εντοπίζοντας πιθανές διαφορές λόγω της επιλογής του αρχικού σημείου ή του τρόπου καθορισμού του βήματος. Επιπλέον, εξετάζεται αν ο αλγόριθμος οδηγεί πάντοτε σε σωστά αποτελέσματα και, σε περίπτωση αποκλίσεων, διερευνώνται τα αίτια αυτών.

#### Ιδιαιτερότητα Σημείου Αρχικοποίησης

Κατά την αρχικοποίηση του αλγορίθμου στο σημείο  $(0, 0)$ , παρατηρείται ότι όλοι οι αλγόριθμοι τερματίζουν αμέσως. Αυτό είναι λογικό, καθώς το συγκεκριμένο σημείο ικανοποιεί τη συνθήκη τερματισμού

$$|\nabla f| < \epsilon, \quad \epsilon > 0.$$

Η συνάρτηση κλίσης, δεδομένη από τη σχέση:

$$\nabla f = \begin{bmatrix} 5x^4 e^{-x^2-y^2} - 2x^6 e^{-x^2-y^2} \\ -2x^5 y e^{-x^2-y^2} \end{bmatrix},$$

μηδενίζεται στο  $(0, 0)$ . Συνεπώς, το σημείο αυτό αναγνωρίζεται ως σημείο εγκλωβισμού, καθιστώντας το μη κατάλληλο για την έναρξη της διαδικασίας βελτιστοποίησης.

## Υλοποίηση των Μεθόδων Βελτιστοποίησης

Στην παρούσα ανάλυση, η μέθοδος Μεγίστης Καθόδου (*Steepest Descent*) υλοποιείται με τέσσερις διαφορετικές στρατηγικές επιλογής του βήματος  $\gamma_k$ , για την ελαχιστοποίηση της συνάρτησης  $f(x, y)$ . Η διαδικασία περιγράφεται γενικά ως εξής:

### Γενική Διαδικασία

Η μέθοδος Μεγίστης Καθόδου ξεκινά από ένα αρχικό σημείο  $\mathbf{x}_0 = [x_0, y_0]^T$  και επαναλαμβάνει τα εξής βήματα:

1. Υπολογισμός της κατεύθυνσης καθόδου:

$$\mathbf{d}_k = -\nabla f(\mathbf{x}_k),$$

όπου  $\nabla f(\mathbf{x}_k)$  είναι η κλίση της συνάρτησης  $f$  στο σημείο  $\mathbf{x}_k$ .

2. Υπολογισμός του βήματος  $\gamma_k$  σύμφωνα με μία από τις ακόλουθες στρατηγικές.
3. Ενημέρωση της θέσης:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \gamma_k \mathbf{d}_k.$$

4. Έλεγχος σύγκλισης: Η διαδικασία σταματά όταν:

$$\|\nabla f(\mathbf{x}_k)\| < \epsilon,$$

όπου  $\epsilon > 0$  είναι μια μικρή προκαθορισμένη τιμή.

### 1. Σταθερό Βήμα (*Fixed Step Size*)

Το βήμα  $\gamma_k$  λαμβάνει μία σταθερή τιμή  $\gamma > 0$  που ορίζεται εκ των προτέρων:

$$\gamma_k = \gamma.$$

### 2. Βέλτιστο Βήμα μέσω Ελαχιστοποίησης (*Optimal Step Size*)

Το βέλτιστο  $\gamma_k$  υπολογίζεται ως εξής:

$$\gamma_k = \arg \min_{\gamma > 0} f(\mathbf{x}_k + \gamma \mathbf{d}_k).$$

Η παραπάνω ελαχιστοποίηση υλοποιείται αριθμητικά, για παράδειγμα με τη μέθοδο *fminbnd* του *MATLAB* ή τη μέθοδο Χρυσής Τομής.

### 3. Μέθοδος Χρυσής Τομής (*Golden Section Search*)

Η μέθοδος Χρυσού Τομέα χρησιμοποιείται για την αριθμητική ελαχιστοποίηση της  $f(\mathbf{x}_k + \gamma \mathbf{d}_k)$ . Στη μέθοδο αυτή:

i ) Ορίζεται αρχικό διάστημα  $[a, b]$  (όπου το όρισα  $[0, 10]$ ).

ii ) Υπολογίζονται τα σημεία:

$$x_1 = a + (1 - \gamma)(b - a), \quad x_2 = a + \gamma(b - a),$$

$$\text{όπου } \gamma = \frac{\sqrt{5}-1}{2}.$$

- iii ) Επαναλαμβάνεται η διαδικασία σύγκρισης  $f(x_1)$  και  $f(x_2)$ , έως ότου το διάστημα  $[a, b]$  να ικανοποιεί:

$$b - a < \ell,$$

όπου  $\ell$  είναι μια μικρή προκαθορισμένη ανοχή που ορίστηκε στο 0.0001.

#### 4. Κανόνας Armijo

Το  $\gamma_k$  επιλέγεται έτσι ώστε να ικανοποιείται η συνθήκη Armijo:

$$f(\mathbf{x}_k) - f(\mathbf{x}_k + \gamma_k \mathbf{d}_k) \geq -\alpha \gamma_k \nabla f(\mathbf{x}_k)^\top \mathbf{d}_k,$$

όπου  $0 < \alpha < 1$  είναι σταθερά. Η τιμή  $\gamma_k$  προσαρμόζεται δυναμικά μειώνοντας την αρχική τιμή  $\gamma$  έως ότου η παραπάνω ανισότητα να ισχύει.

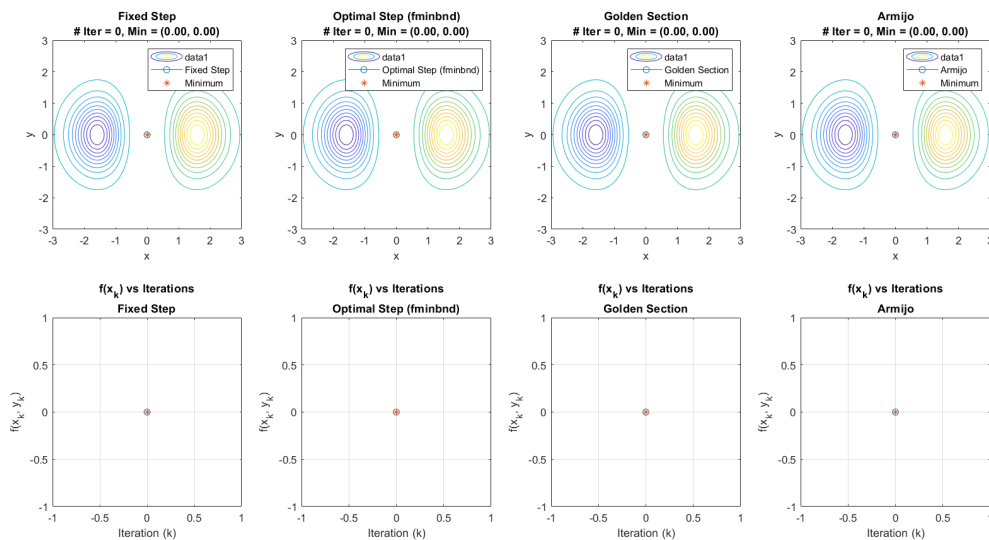
#### Σύγκλιση

Η διαδικασία επαναλαμβάνεται έως ότου το μέγεθος της κλίσης γίνει μικρότερο από ένα προκαθορισμένο όριο:

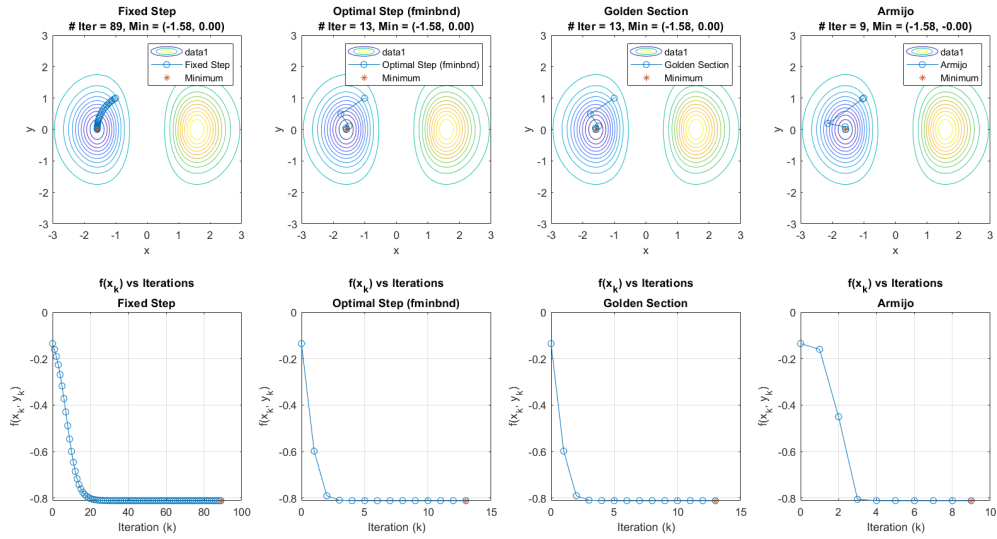
$$\|\nabla f(\mathbf{x}_k)\| < \epsilon.$$

Επιπλέον, για λόγους πρακτικούς θεωρήσα ότι αν υπερβαίνει το όριο επαναλήψεων των 40000 να διακόπτει τον αλγόριθμο γιατί δεν έχουμε σύγκλιση.

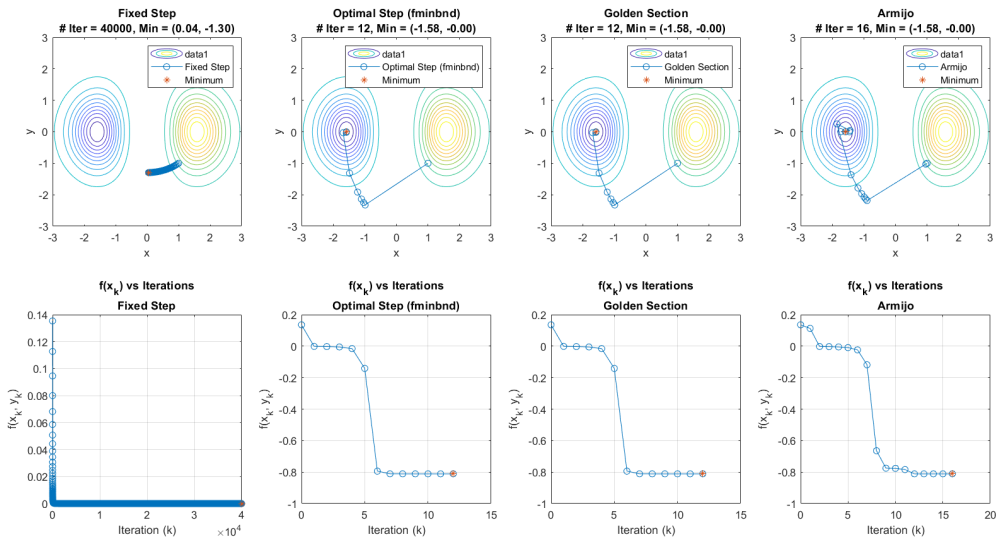
Στην συνέχεια δημιούργησα και τη γραφική παράσταση της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων που απαιτούνται μέχρι να τερματίσει ο αλγόριθμος αλλά και την αναπαράσταση τυχών συγκλίσεων στην αναπαράσταση της συνάρτησης  $f$ . Έτσι παρατηρήθηκαν τα παρακάτω αποτελέσματα όπως φαίνονται στα Σχήματα 2.1, 2.2 και 2.3 για αρχικά σημεία  $(x_0, y_0) = (0, 0)$ ,  $(-1, 1)$ , και  $(1, -1)$  αντίστοιχα.



Σχήμα 2.1: Αποτελέσματα για  $(0, 0)$   $f$ .



Σχήμα 2.2: Αποτελέσματα για  $(-1, 1)$   $f$ .



Σχήμα 2.3: Αποτελέσματα για  $(1, -1)$   $f$ .

	Fixed Step	Optimal Step	Armijo
$(0, 0)$	X	X	X
$(-1, 1)$	-0.811174	-0.811174	-0.811174
$(1, -1)$	X	-0.811174	-0.811174

Πίνακας 2.1: Πίνακας για την καταγραφή της σύγκλισης.



	Fixed Step	Optimal Step	Armijo
(0, 0)	0	0	0
(-1, 1)	89	13	9
(1, -1)	40000	12	16

Πίνακας 2.2: Πίνακας για την καταγραφή των επαναλήψεων.

Από τα σχήματα πολύ εύκολα παρατηρούμε ότι η μέθοδος που υπολογίζει το βήμα με τρόπο που ελαχιστοποιεί τη συνάρτηση απαιτεί λιγότερες επαναλήψεις για να συγκλίνει και έχει ίδια αποτελέσματα είτε υλοποιηθεί με `fminbnd` είτε με μέθοδο του χρυσού τομέα, σε σύγκριση με τη μέθοδο που χρησιμοποιεί σταθερό βήμα. Επιπλέον, η μέθοδος που καθορίζει το βήμα σύμφωνα με τον κανόνα του Armijo παρουσιάζει ακόμα καλύτερη απόδοση, καθώς συγκλίνει με ακόμη λιγότερες επαναλήψεις στην περίπτωση (-1, 1), ενώ στο (1, -1) απαιτούνται λιγότερες επαναλήψεις στην μέθοδο εύρεσης βέλτιστου γ άρα έχει καλύτερη απόδοση. Στις περιπτώσεις που στον Πίνακα 2.1 έχουμε ένδειξη  $X$  παρατηρείται εκλωβισμός σε τοπικό ελάχιστο και δεν προσεγγίζεται σωστά το ελάχιστο ή μέγιστο αντίστοιχα.

Ο αντίστοιχος κώδικας βρίσκεται στο αρχείο **Part2.m**.

## Κεφάλαιο 3

### Μέθοδος Newton

Ακολουθώντας ακριβώς την ίδια διαδικασία με την μέθοδο Steepest Descent θα κληθούμε να υλοποιήσουμε και την μέθοδο Newton .

Στο συγκεκριμένο ερώτημα, εξετάζεται η εφαρμογή της μεθόδου Νεύτωνα για την ελαχιστοποίηση της συνάρτησης  $f(x, y)$ , λαμβάνοντας υπόψη διαφορετικές αρχικές συνθήκες και στρατηγικές επιλογής του βήματος  $\gamma_k$ . Η μέθοδος Νεύτωνα λαμβάνει υπόψη τόσο την κλίση όσο και τη δεύτερη παράγωγο (Hessian) της συνάρτησης, για να επιτύχει γρηγορότερη σύγκλιση. Στόχος είναι να αναλυθεί πώς οι παράμετροι αυτές επηρεάζουν την απόδοση του αλγορίθμου, τόσο ως προς τη σύγκλιση όσο και ως προς την ακρίβεια του τελικού αποτελέσματος.

#### Υλοποίηση της Μεθόδου Νεύτωνα

Η μέθοδος Νεύτωνα υλοποιείται με τέσσερις διαφορετικές στρατηγικές επιλογής του βήματος  $\gamma_k$ , για την ελαχιστοποίηση της συνάρτησης  $f(x, y)$ . Η γενική διαδικασία περιγράφεται ως εξής:

#### Γενική Διαδικασία

Η μέθοδος Νεύτωνα ξεκινά από ένα αρχικό σημείο  $\mathbf{x}_0 = [x_0, y_0]^T$  και επαναλαμβάνει τα εξής βήματα:

1. Υπολογισμός της κατεύθυνσης καθόδου χρησιμοποιώντας την αντίστροφη του πίνακα Hessian :

$$\mathbf{d}_k = -[\nabla^2 f(\mathbf{x}_k)]^{-1} \nabla f(\mathbf{x}_k),$$

όπου  $\nabla f(\mathbf{x}_k)$  είναι η κλίση και  $\nabla^2 f(\mathbf{x}_k)$  ο πίνακας Hessian της συνάρτησης  $f$  στο σημείο  $\mathbf{x}_k$ .

2. Υπολογισμός του βήματος  $\gamma_k$  σύμφωνα με μία από τις ακόλουθες στρατηγικές.
3. Ενημέρωση της θέσης:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \gamma_k \mathbf{d}_k.$$

4. Έλεγχος σύγκλισης: Η διαδικασία σταματά όταν:

$$\|\nabla f(\mathbf{x}_k)\| < \epsilon,$$

όπου  $\epsilon > 0$  είναι μια μικρή προκαθορισμένη τιμή.

### 1. Σταθερό Βήμα (*Fixed Step Size*)

Το βήμα  $\gamma_k$  λαμβάνει μία σταθερή τιμή  $\gamma > 0$  που ορίζεται εκ των προτέρων:

$$\gamma_k = \gamma.$$

### 2. Κανόνας Armijo

Το  $\gamma_k$  επιλέγεται έτσι ώστε να ικανοποιείται η συνθήκη Armijo:

$$f(\mathbf{x}_k) - f(\mathbf{x}_k + \gamma_k \mathbf{d}_k) \geq -\alpha \gamma_k \nabla f(\mathbf{x}_k)^\top \mathbf{d}_k,$$

όπου  $0 < \alpha < 1$  είναι σταθερά. Η τιμή  $\gamma_k$  προσαρμόζεται δυναμικά μειώνοντας την αρχική τιμή  $\gamma$  έως ότου η παραπάνω ανισότητα να ισχύει.

### 3. Βέλτιστο Βήμα μέσω Ελαχιστοποίησης (*Optimal Step Size*)

Το βέλτιστο  $\gamma_k$  υπολογίζεται ως εξής:

$$\gamma_k = \arg \min_{\gamma > 0} f(\mathbf{x}_k + \gamma \mathbf{d}_k).$$

Η παραπάνω ελαχιστοποίηση υλοποιείται αριθμητικά, για παράδειγμα με τη μέθοδο *fminbnd* του *MATLAB*.

### 4. Μέθοδος Χρυσής Τομής (*Golden Section Search*)

Η μέθοδος Χρυσής Τομής χρησιμοποιείται για την αριθμητική ελαχιστοποίηση της  $f(\mathbf{x}_k + \gamma \mathbf{d}_k)$  όπως ακριβώς φαίνεται παραπάνω στην μέθοδο μεγίστης καθόδου.

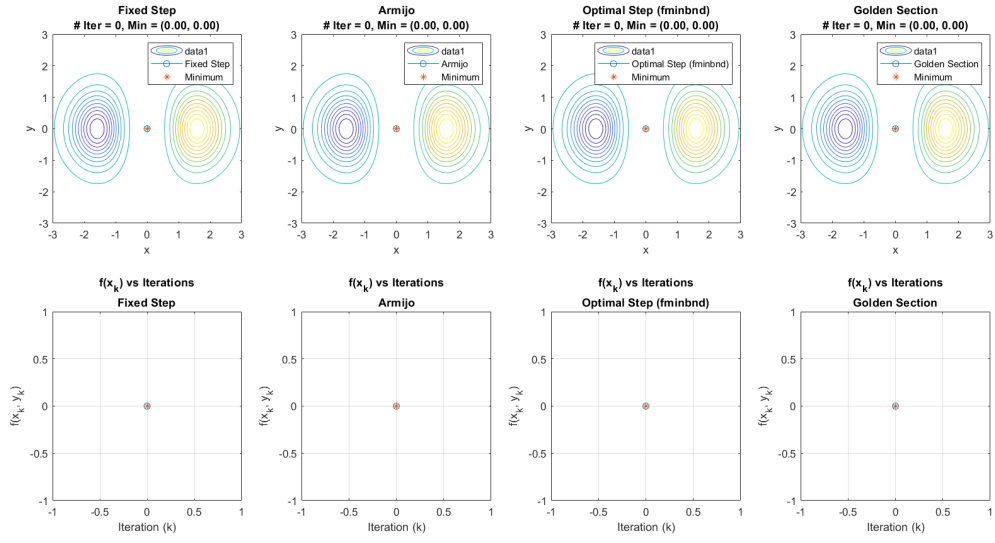
### Σύγκλιση

Η διαδικασία επαναλαμβάνεται έως ότου το μέγεθος της κλίσης γίνει μικρότερο από ένα προκαθορισμένο όριο:

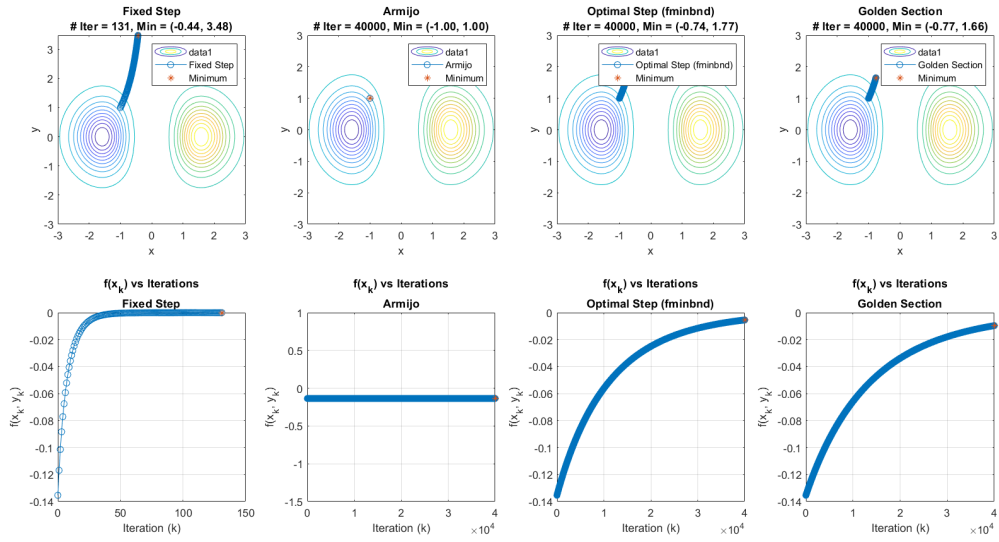
$$\|\nabla f(\mathbf{x}_k)\| < \epsilon.$$

Επιπλέον, για λόγους πρακτικούς, θεωρήθηκε ότι αν υπερβεί το όριο επαναλήψεων των 40000, ο αλγόριθμος διακόπτεται, καθώς θεωρείται ότι δεν υπάρχει σύγκλιση όπως και στην μέθοδο μεγίστης καθόδου.

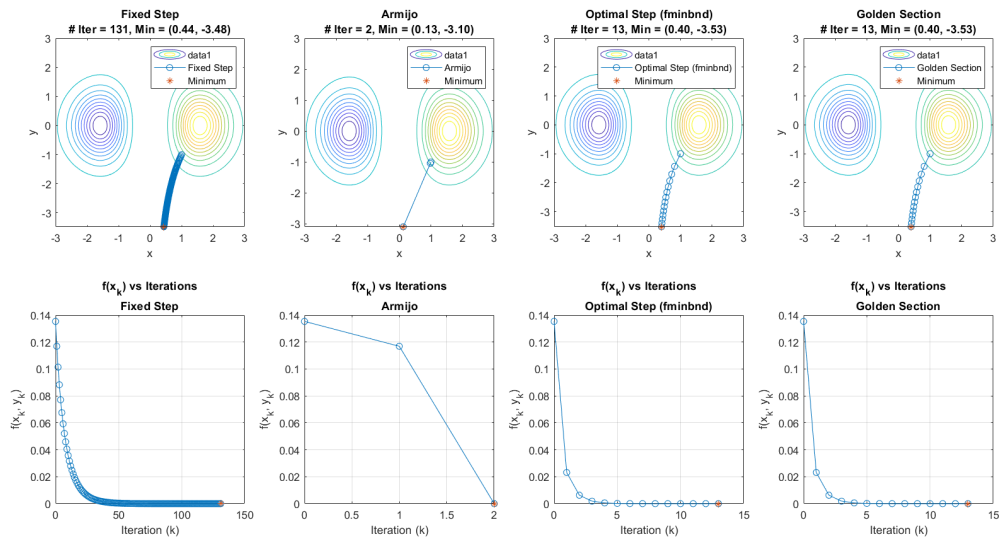
Στην συνέχεια όπως και πριν αναπαριστώ και τη γραφική παράσταση της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων που απαιτούνται μέχρι να τερματίσει ο αλγόριθμος αλλά και την αναπαράσταση τυχών συγκλίσεων στην αναπαράσταση της συνάρτησης  $f$ . Έτσι παρατηρήθηκαν τα παρακάτω αποτελέσματα όπως φαίνονται στα Σχήματα 3.1, 3.2 και 3.3 για αρχικά σημεία  $(x_0, y_0) = (0, 0)$ ,  $(-1, 1)$ , και  $(1, -1)$  αντίστοιχα.



Σχήμα 3.1: Αποτελέσματα για  $(0, 0)$   $f$ .



Σχήμα 3.2: Αποτελέσματα για  $(-1, 1)$   $f$ .



Σχήμα 3.3: Αποτελέσματα για  $(1, -1)$   $f$ .

	Fixed Step	Optimal Step	Armijo
(0, 0)	X	X	X
(-1, 1)	X	X	X
(1, -1)	X	X	X

Πίνακας 3.1: Πίνακας για την καταγραφή της σύγκλισης.

	Fixed Step	Optimal Step	Armijo
(0, 0)	0	0	0
(-1, 1)	131	40000	40000
(1, -1)	131	13	2

Πίνακας 3.2: Πίνακας για την καταγραφή των επαναλήψεων.

Η μέθοδος του Νεύτωνα, όπως ορίζεται, δεν μπορεί να εφαρμοστεί για τη συγκεκριμένη συνάρτηση. Αυτό συμβαίνει διότι, για να έχει η μέθοδος την ιδιότητα της επαναληπτικής καθόδου και να συγκλίνει σωστά, ο Εσσιανός πίνακας της συνάρτησης  $f$ , δηλαδή το  $\nabla^2 f$ , πρέπει να είναι θετικά ορισμένος.

Ωστόσο, για καθένα από τα τρία αρχικά σημεία  $(0, 0)$ ,  $(-1, 1)$  και  $(1, -1)$ , ο πίνακας δεν είναι θετικά ορισμένος. Στην πρώτη περίπτωση έχουμε τον μηδενικό πίνακα, ενώ στις άλλες δύο περιπτώσεις οι ιδιοτιμές του πίνακα είναι οι  $[4e^{-2}, -8e^{-2}]$  και  $[-4e^{-2}, 8e^{-2}]$ , οπότε προφανώς ο πίνακας δεν είναι θετικά ορισμένος και η μέθοδος ξεκινάει λανθασμένα από τα πρώτα κιόλας σημεία. Έτσι δεν έχουμε καλά συμπεράσματα με την μέθοδο Νεωτον.

Ο αντίστοιχος κώδικας βρίσκεται στο αρχείο **Part3.m**.

## Κεφάλαιο 4

### Μέθοδος Levenberg-Marquardt

Σε αυτό το ερώτημα για μία ακόμα φορά εξετάζεται η εφαρμογή της μεθόδου Levenberg-Marquardt για την ελαχιστοποίηση της συνάρτησης  $f(x, y)$ , λαμβάνοντας υπόψη διαφορετικές αρχικές συνθήκες και στρατηγικές επιλογής του βήματος  $\gamma_k$ . Η μέθοδος Levenberg-Marquardt αποτελεί συνδυασμό της μεθόδου Νεύτωνα και της μεθόδου Μεγίστης Καθόδου, προσφέροντας αυξημένη σταθερότητα μέσω της εισαγωγής μιας παραμέτρου ρύθμισης  $\mu$ , η οποία διασφαλίζει ότι ο πίνακας Hessian παραμένει θετικά ορισμένος. Στόχος είναι να αναλυθεί πώς οι παράμετροι αυτές επηρεάζουν την απόδοση του αλγορίθμου.

#### Υλοποίηση της Μεθόδου Levenberg-Marquardt

Η μέθοδος Levenberg-Marquardt υλοποιείται με τέσσερις διαφορετικές στρατηγικές επιλογής του βήματος  $\gamma_k$ , για την ελαχιστοποίηση της συνάρτησης  $f(x, y)$ . Η γενική διαδικασία περιγράφεται ως εξής:

##### Γενική Διαδικασία

Η μέθοδος Levenberg-Marquardt ξεκινά από ένα αρχικό σημείο  $\mathbf{x}_0 = [x_0, y_0]^T$  και επαναλαμβάνει τα εξής βήματα:

1. Υπολογισμός της κατεύθυνσης καθόδου με ρύθμιση μέσω της παραμέτρου  $\mu$ :

$$\mathbf{d}_k = -[\nabla^2 f(\mathbf{x}_k) + \mu I]^{-1} \nabla f(\mathbf{x}_k),$$

όπου  $\nabla f(\mathbf{x}_k)$  είναι η κλίση και  $\nabla^2 f(\mathbf{x}_k)$  ο πίνακας Hessian της συνάρτησης  $f$  στο σημείο  $\mathbf{x}_k$ , ενώ  $I$  είναι ο μοναδιαίος πίνακας.

2. Υπολογισμός του βήματος  $\gamma_k$  σύμφωνα με μία από τις ακόλουθες στρατηγικές.
3. Ενημέρωση της θέσης:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \gamma_k \mathbf{d}_k.$$

4. Έλεγχος σύγκλισης: Η διαδικασία σταματά όταν:

$$\|\nabla f(\mathbf{x}_k)\| < \epsilon,$$

όπου  $\epsilon > 0$  είναι μια μικρή προκαθορισμένη τιμή.

### 1. Σταθερό Βήμα (*Fixed Step Size*)

Το βήμα  $\gamma_k$  λαμβάνει μία σταθερή τιμή  $\gamma > 0$  που ορίζεται εκ των προτέρων:

$$\gamma_k = \gamma.$$

### 2. Κανόνας Armijo

Το  $\gamma_k$  επιλέγεται έτσι ώστε να ικανοποιείται η συνθήκη Armijo:

$$f(\mathbf{x}_k) - f(\mathbf{x}_k + \gamma_k \mathbf{d}_k) \geq -\alpha \gamma_k \nabla f(\mathbf{x}_k)^\top \mathbf{d}_k,$$

όπου  $0 < \alpha < 1$  είναι σταθερά. Η τιμή  $\gamma_k$  προσαρμόζεται δυναμικά μειώνοντας την αρχική τιμή  $\gamma$  έως ότου η παραπάνω ανισότητα να ισχύει.

### 3. Βέλτιστο Βήμα μέσω Ελαχιστοποίησης (*Optimal Step Size*)

Το βέλτιστο  $\gamma_k$  υπολογίζεται ως εξής:

$$\gamma_k = \arg \min_{\gamma > 0} f(\mathbf{x}_k + \gamma \mathbf{d}_k).$$

Η παραπάνω ελαχιστοποίηση υλοποιείται αριθμητικά, για παράδειγμα με τη μέθοδο *fminbnd* του *MATLAB*.

### 4. Μέθοδος Χρυσής Τομής (*Golden Section Search*)

Η μέθοδος Χρυσής Τομής χρησιμοποιείται και υλοποιείται όπως δείχτηκε στο Κεφάλαιο 2.

### Σύγκλιση

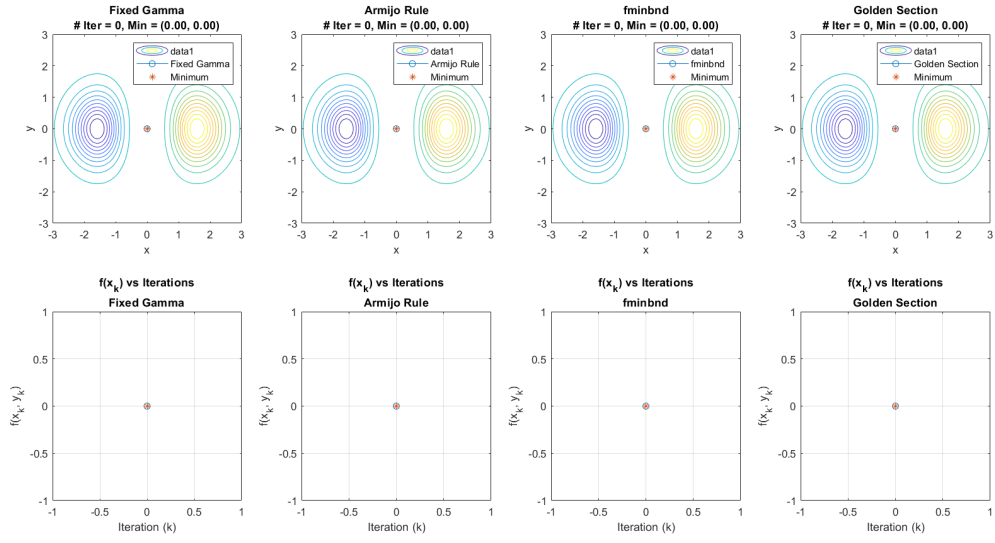
Η διαδικασία επαναλαμβάνεται έως ότου το μέγεθος της κλίσης γίνει μικρότερο από ένα προκαθορισμένο όριο:

$$\|\nabla f(\mathbf{x}_k)\| < \epsilon.$$

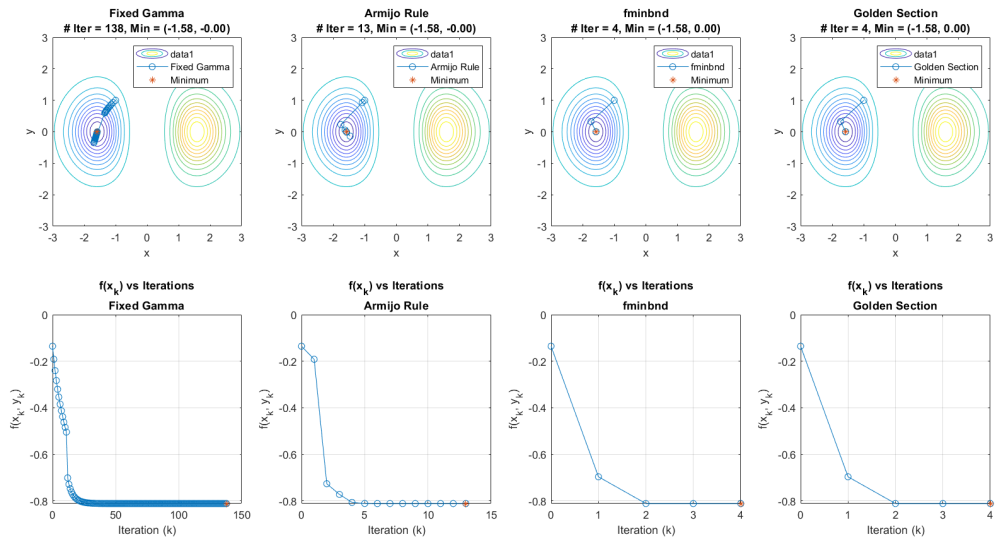
Ξανά, ο αλγόριθμος διακόπτεται στις 40000 επαναλήψεις.

Όπως και στα προηγούμενα 2 ερωτήματα θα δείξω την γραφική παράσταση της σύγκλισης της συνάρτησης ως προς τον αριθμό των επαναλήψεων που απαιτούνται μέχρι να τερματίσει ο αλγόριθμος αλλά και την αναπαράσταση τυχών συγκλίσεων στην αναπαράσταση της συνάρτησης  $f$ . Έτσι παρατηρήθηκαν τα παρακάτω αποτελέσματα όπως φαίνονται στα Σχήματα 4.1, 4.2 και 4.3 για αρχικά σημεία  $(x_0, y_0) = (0, 0)$ ,  $(-1, 1)$ , και  $(1, -1)$  αντίστοιχα.

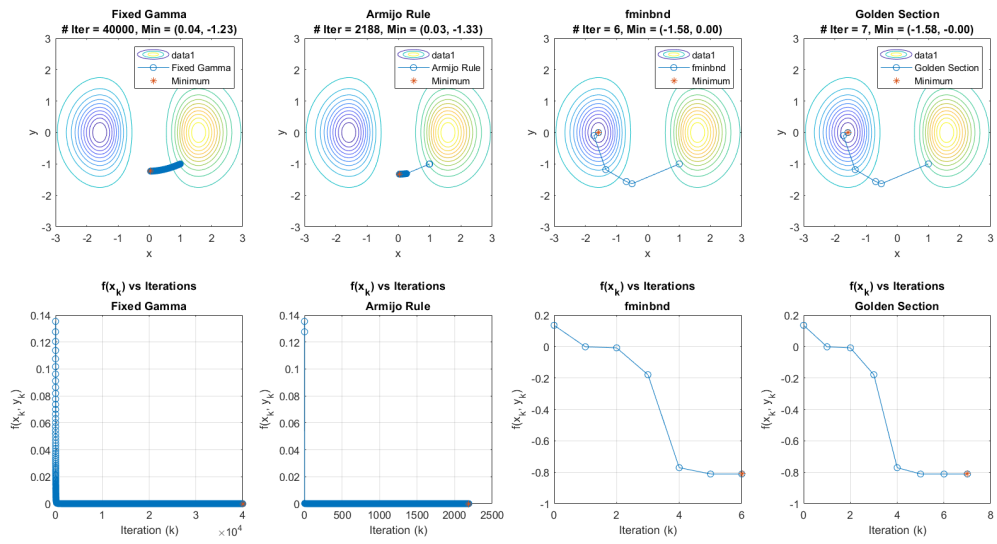




Σχήμα 4.1: Αποτελέσματα για  $(0, 0)$   $f$ .



Σχήμα 4.2: Αποτελέσματα για  $(-1, 1)$   $f$ .



Σχήμα 4.3: Αποτελέσματα για  $(1, -1)$   $f$ .

	Fixed Step	Optimal Step	Armijo
(0, 0)	X	X	X
(-1, 1)	-811174	-0.811174	-0.811174
(1, -1)	X	-0.811174	X

Πίνακας 4.1: Πίνακας για την καταγραφή της σύγκλισης.

	Fixed Step	Optimal Step	Armijo
(0, 0)	0	0	0
(-1, 1)	138	4	13
(1, -1)	40000	6 (7 Golden Section )	2188

Πίνακας 4.2: Πίνακας για την καταγραφή των επαναλήψεων.

Από τα αποτελέσματα παρατηρούμε ότι η μέθοδος Levenberg-Marquardt παρουσιάζει αρκετά βελτιωμένη συμπεριφορά σε σύγκριση με την μέθοδο Newton , ειδικά όταν συνδυάζεται με δυναμικές στρατηγικές επιλογής βήματος, όπως η εύρεση βέλτιστου  $\gamma$  (είτε με τη μέθοδο fminbnd είτε με τη μέθοδο Χρυσής Τομής).

Συγκεκριμένα:

- Στην περίπτωση  $(-1, 1)$ , η μέθοδος Levenberg-Marquardt με την εύρεση βέλτιστου  $\gamma$  συγκλίνει σε μόλις λίγες επαναλήψεις (6 ή 7, αναλόγως της στρατηγικής), ενώ η μέθοδος με σταθερό βήμα χρειάζεται πολύ περισσότερες επαναλήψεις (138). Ο κανόνας Armijo επίσης επιτυγχάνει αποδοτική σύγκλιση με 13 επαναλήψεις.
- Στην περίπτωση  $(1, -1)$ , οι μέθοδοι εύρεσης βέλτιστου  $\gamma$  (π.χ. Χρυσή Τομή) είναι και πάλι εξαιρετικά αποτελεσματικές (6 επαναλήψεις). Ωστόσο, η μέθοδος με σταθερό βήμα αποτυγχάνει να συγκλίνει εντός του ορίου επαναλήψεων (40000 επαναλήψεις), γεγονός που υποδεικνύει ότι η συγκεκριμένη επιλογή βήματος δεν είναι κατάλληλη για αυτή την περίπτωση.

Επιπλέον, στις περιπτώσεις που εμφανίζεται ένδειξη  $X$  στον πίνακα, η μέθοδος Levenberg-Marquardt ενδέχεται να αντιμετωπίζει δυσκολίες λόγω μη θετικά ορισμένου πίνακα:

$$\nabla^2 f(\mathbf{x}_k) + \mu I$$

Έτσι οδηγούμαστε σε "εγκλωβισμό" σε τοπικά ακρότατα (τοπικό ελάχιστο ή μέγιστο) ή αποτυχία προσέγγισης του ελάχιστου. Αυτό ενισχύει τη σημασία της δυναμικής ρύθμισης της παραμέτρου  $\mu$ , ώστε να εξασφαλιστεί η σταθερότητα της μεθόδου. Η αποδοτικότερη επιλογή στην περίπτωση της εφαρμογής μεθόδου Levenberg-Marquardt φαίνεται από το Πίνακα 4.2 να είναι η εύρεση βέλτιστου  $\gamma$  ανεξαρτήτως μεθόδου ελαχιστοποίησης που επιλέξαμε.

Ο αντίστοιχος κώδικας βρίσκεται στο αρχείο **Part4.m**.

## Κεφάλαιο 5

### Αποτελέσματα και Συμπεράσματα

Κατά την εφαρμογή των μεθόδων, παρατηρούμε ότι για αρχικά σημεία όπως το  $(1, -1)$  αλλά και το  $(-1, 1)$ , η σύγκλιση μπορεί να παρεμποδιστεί από τοπικά ακρότατα, οδηγώντας σε "παγίδευση" του αλγορίθμου. Για αρχικό σημείο  $(0, 0)$  οι μέθοδοι μένουν στάσιμοι χωρίς επανάληψη. Επιπλέον, στρατηγικές όπως η ελαχιστοποίηση του βήματος ή ο κανόνας του Armijo μειώνουν σημαντικά τις απαιτούμενες επαναλήψεις για τη σύγκλιση του αλγορίθμου. Επιπρόσθετα, η μέθοδος Newton, δεν είναι βοηθητική για την συγκεκριμένη διερεύνηση καθώς δεν οδηγεί σε σωστά συμπεράσματα λόγω του μη θετικά ορισμένου πίνακα  $\nabla^2 f(\mathbf{x}_k)$  σε μερικές περιπτώσεις. Τέλος, σχετικά με την επιλογή του  $\gamma_k$ , η αποδοτικότερη επιλογή φαίνεται σε πολλές περιπτώσεις να είναι αυτή που οδηγεί στην ελαχιστοποίηση της  $f(\mathbf{x}_k + \gamma_k \mathbf{d}_k)$ .

## Βιβλιογραφία

- [1] Γεώργιος Α. Ροβιθάκης, *Τεχνικές Βελτιστοποίησης*. Εκδόσεις ΤΖΙΟΛΑ.