**ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ**
**ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ**
**ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ**


# ΤΕΧΝΙΚΕΣ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ

ΕΡΓΑΣΙΑ ΔΕΥΤΕΡΗ


**Πηγή Παπανικολάου**


Νοέμβριος 2023

# Περιεχόμενα


1 Περιγραφή Προβλήματος ..... 2

2 Επίλυση ..... 2

Θέμα 1ο ..... 2

Θέμα 2ο: Μέθοδος Μέγιστης Καθόδου (Steepest Descent) ..... 4

Θέμα 3ο: Μέθοδος Newton ..... 10

Θέμα 3ο: Μέθοδος Lavenberg-Marquardt ..... 10

3 Παρατηρήσεις ..... 17

# 1 Περιγραφή Προβλήματος

Στην εργασία αυτή θα ασχοληθούμε με το πρόβλημα ελαχιστοποίησης μιας δοσμένης συνάρτησης πολλών μεταβλητών $f: \mathbb{R}^n \to \mathbb{R}$ χωρίς περιορισμούς με χρήση παραγώγων. Οι αλγόριθμοι που θα χρησιμοποιήσουμε βασίζονται στην ιδέα της επαναληπτικής καθόδου, βάσει της οποίας ξεκινάμε από κάποιο σημείο $x_0 \in \mathbb{R}^n$ και παράγουμε διαδοχικά τα διανύσματα $x_1, x_2, \ldots$ έτσι ώστε $f(x_{k+1}) < f(x_k), k = 1,2, \ldots$ .

Οι αλγόριθμοι αναζήτησης που θα μελετηθούν είναι:

- Μέθοδος Μέγιστης Καθόδου (Steepest Descent)
- Μέθοδος Newton
- Μέθοδος Levenberg-Marquardt

Η αντικειμενική συνάρτηση που θα μελετήσουμε είναι η:

$$f(x, y) = x^3 e^{-x^2 - y^4} \tag{1}$$

Τα αρχικά σημεία που επιλέχθηκαν για την ανάλυση είναι:

- $(x_0, y_0) = (0,0)$
- $(x_0, y_0) = (1,1)$
- $(x_0, y_0) = (-1,-1)$

Σε κάθε περίπτωση το βήμα $\gamma_k$ θα επιλέγεται:

- Σταθερό
- Τέτοιο ώστε να ελαχιστοποιεί την $f(x_k + \gamma_k d_k)$
- Βάσει του κανόνα Armijo

Οι μέθοδοι που μας απασχολούν αποτελούν αλγορίθμους κλίσης, άρα είναι της μορφής:

$$x_{k+1} = x_k - \gamma_k \Delta_k \nabla f(x_k) \tag{2}$$

# 2 Επίλυση

### Θέμα 1ο

Η τρισδιάστατη γραφική παράσταση της $f$ όπως προκύπτει από το matlab φαίνεται στο Figure 1. Παρατηρούμε ότι η συνάρτηση στο ολικό ελάχιστό της παίρνει τιμή κοντά στο $-0.4$ για $x \approx -1.22$ και $y \approx 0$. Αυτή η διαπίστωση θα μας βοηθήσει στην επαλήθευση των μεθόδων που θα εξετάσουμε.

Ακόμα δίνεται στο Figure 2 και το διάγραμμα των ισοβαρών καμπυλών της $f$ για καλύτερη εποπτεία.

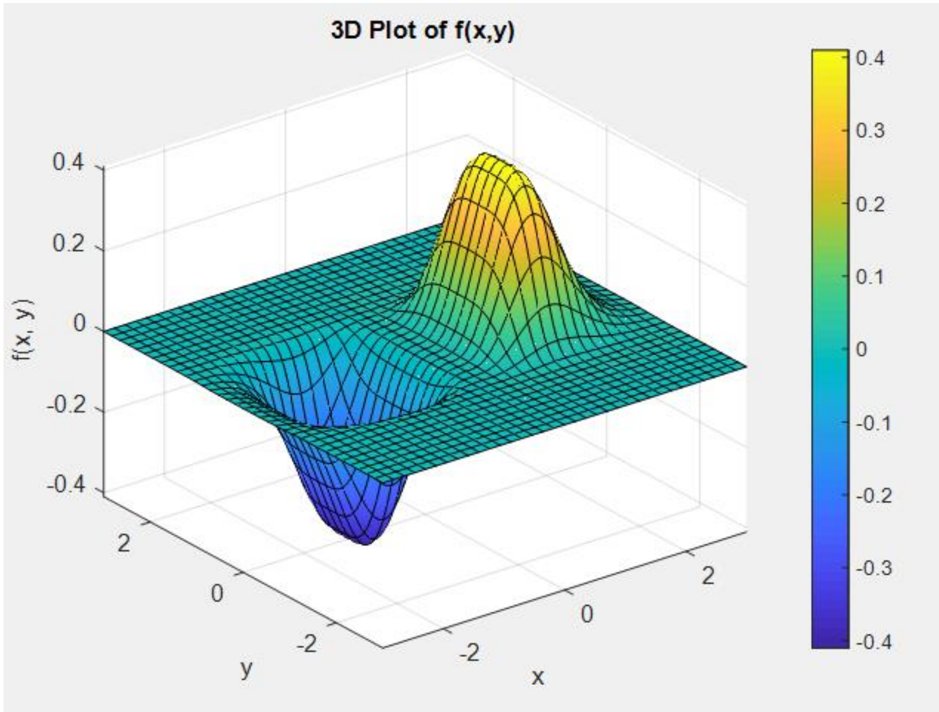



*Figure 1. Τρισδιάστατη απεικόνιση της f*

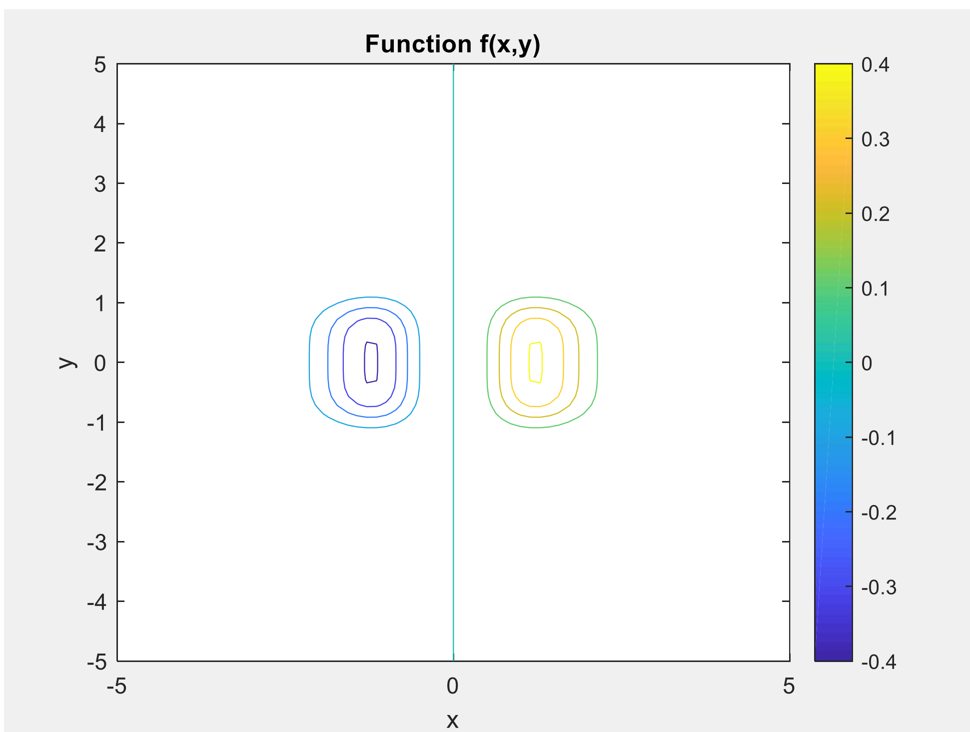



*Figure 2. Ισοβαρείς καμπύλες της f*

## Θέμα 2ο: Μέθοδος Μέγιστης Καθόδου (Steepest Descent)

Η μέθοδος μέγιστης καθόδου (steepest descent) προκύπτει επιλέγοντας $\Delta_k = I, k = 1,2,3, \dots$ . η (2) λοιπόν γίνεται

$$x_{k+1} = x_k - \gamma_k \nabla f(x_k) \tag{3}$$

Το διάνυσμα κατεύθυνσης $d_k = -\nabla f(x_k)$ ισούται με την αρνητική κλίση της $f$ υπολογισμένες στο σημείο $x_k$.

Ως σταθερά τερματισμού $\varepsilon$ επιλέγουμε $\varepsilon = 10^{-3}$. Όσο μικρότερο επιλέγετε το $\varepsilon$ τόσες περισσότερες θα είναι οι επαναλήψεις του αλγορίθμου για σύγκλιση ώστε να ικανοποιείται η μεγαλύτερη ακρίβεια που ζητείται.

Μελετούμε, λοιπόν, τη συμπεριφορά του αλγορίθμου με αφετηρία το κάθε σημείο που μας ζητείται:

### a) Με αρχικό σημείο $(\mathbf{x_0}, \mathbf{y_0}) = (\mathbf{0}, \mathbf{0})$

Όταν ξεκινάμε την αναζήτηση του ελαχίστου από το σημείο $(0,0)$, ο αλγόριθμος τερματίζει από την πρώτη επανάληψη και το αποτέλεσμα που δίνει είναι το $(0,0)$. Αυτό συμβαίνει καθώς $\nabla f(0,0) = 0$ και άρα η συνθήκη τερματισμού $|\nabla f(x_k, y_k)| \leq \varepsilon$ πληρείται εξ αρχής.

Το αποτέλεσμα αυτό είναι αναμενόμενο καθώς όπως βλέπουμε και στην τρισδιάστατη αναπαράσταση το σημείο $(0,0)$ είναι σημείο του επιπέδου $xy$ και αποτελεί τοπικό ελάχιστο της συνάρτησης. Άρα ο αλγόριθμος εγκλωβίζεται σε αυτό το τοπικό ελάχιστο και αδυνατεί να βρει το ολικό.

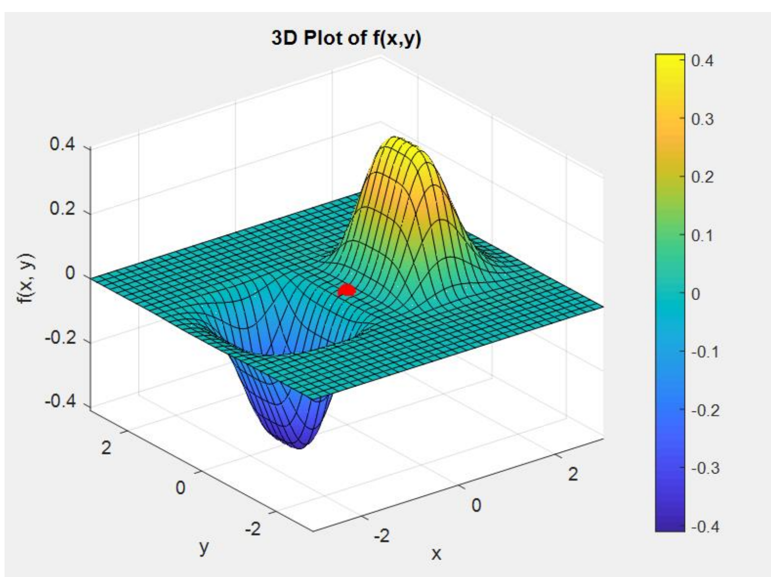



*Figure 3. Σημείο f(0,0)*

Το παραπάνω συμπέρασμα ισχύει για όλες τις μεθόδους που καλούμαστε να ελέγξουμε και δεν μεταβάλλεται ανάλογα με την επιλογή του βήματος $\gamma_k$.

Συνεπώς, το σημείο $(0,0)$ δεν αποτελεί καλό σημείο εκκίνησης.

### b) Με αρχικό σημείο $(\mathbf{x_0}, \mathbf{y_0}) = (-\mathbf{1}, -\mathbf{1})$

Όπως παρατηρούμε από τα Figure 4, Figure 5, Figure 6, για όλες τις μεθόδους καθορισμού του $\gamma_k$ ο αλγόριθμος συγκλίνει στο σωστό σημείο, δηλαδή το ελάχιστο της $f$. Η διαφορά που υπάρχει ανάμεσα τους είναι ο αριθμός των επαναλήψεων της μεθόδου ώστε να συγκλίνουμε στο ελάχιστο.

- Στην περίπτωση του σταθερού βήματος, επιλέγουμε αρχικά βήμα $\gamma_k = 1$ για το οποίο προέκυψαν και οι γραφικές παραστάσεις στο Figure 4. Για να ελέγξουμε αν η επιλογή αυτή είναι καλή, τρέχουμε τον αλγόριθμο για ακόμα τέσσερις (4) τιμές $\gamma_k$. Συγκεκριμένα για $\gamma_k = 0.2, 0.5, 1.2, 1.5$. Το αποτέλεσμα είναι το εξής:

| $\gamma_k$ | $k$ | $x^*$ | $y^*$ |
|---|---|---|---|
| **0.2** | 211 | -1.22 | -0.0847 |
| **0.5** | 83 | -1.22 | -0.0847 |
| **1** | 41 | -1.22 | -0.084 |
| **1.2** | 136 | -1.22 | -0.0427 |
| **1.5** | ΔΕΝ ΣΥΓΚΛΙΝΕΙ | | |

  Γίνεται κατανοητό λοιπόν ότι ο αλγόριθμος δουλεύει αποτελεσματικότερα για $\gamma_k = 1$ και για τιμές που αποκλίνουν λίγο από το 1 οι επαναλήψεις αυξάνονται κατά πολύ. Δηλαδή χρειάζεται μεγάλη εμπειρία ώστε να επιλεγεί ένα $\gamma_k$ που να ελαχιστοποιεί τις συνολικές επαναλήψεις.

- Στην περίπτωση που το $\gamma_k$ επιλέγεται σε κάθε επανάληψη με σκοπό να ελαχιστοποιεί την $f(x_k + \gamma_k d_k)$, με τη μέθοδο της διχοτόμου, προκύπτουν οι γραφικές παραστάσεις Figure 5. Βλέπουμε ότι χρειάζονται μόλις 13 επαναλήψεις, ώσπου να οδηγηθούμε στο ελάχιστο της συνάρτησης.
- Στην περίπτωση που το $\gamma_k$ επιλέγεται σύμφωνα με τον κανόνα Armijo, δηλαδή μειώνεται διαδοχικά σύμφωνα με τη σχέση

$$\gamma_k = s\beta^{m_k}, 0 < \beta < 1 \tag{4}$$

  και $m_k$ να είναι ο μικρότερος μη-αρνητικός ακέραιος που ικανοποιεί το κριτήριο:

$$f(x_k) - f(x_{k+1}) \geq -\alpha\beta^{m_k} s {d_k}^T \nabla f(x_k) \tag{5}$$

  προκύπτουν οι Figure 6.

Ο κανόνας Armijo εφαρμόστηκε χρησιμοποιώντας ως τιμές παραμέτρων $\beta = 0.5, \alpha = 10^{-5}$ και αρχικό βήμα $s = 0.5$. Από το Figure 6 διαπιστώνουμε ότι χρειάζεται σχετικά μεγάλος αριθμός επαναλήψεων.

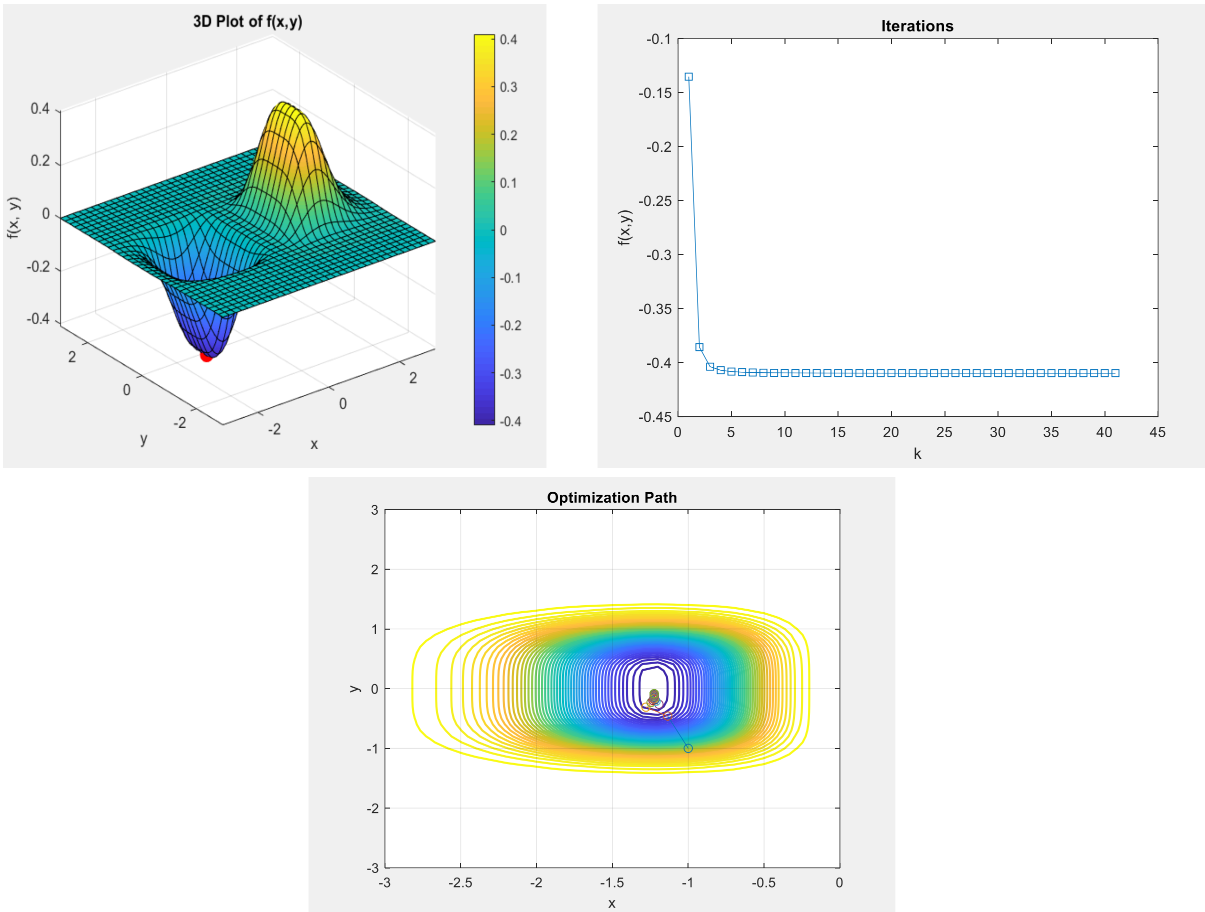
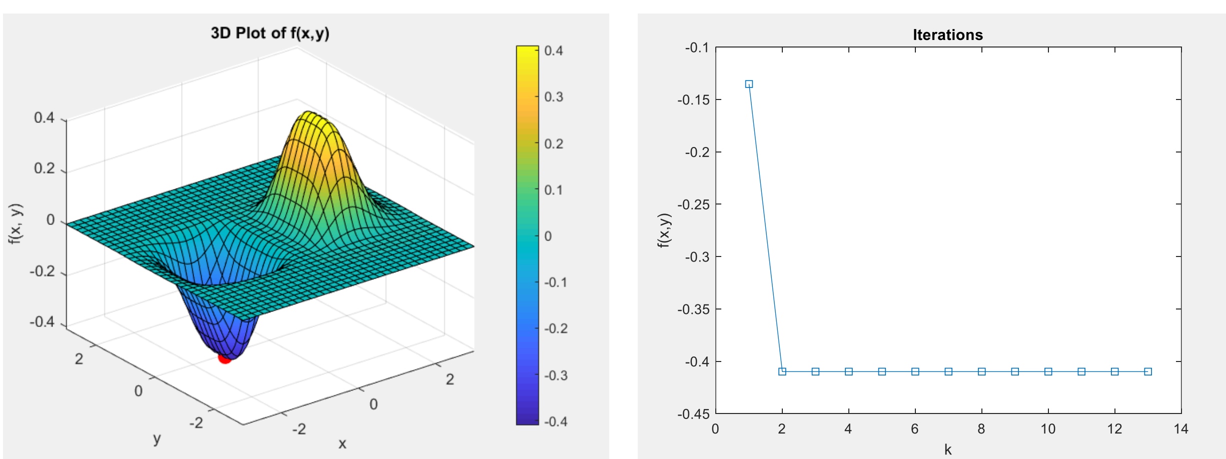



*Figure 4. Σταθερό $\gamma_k$ για αρχικό σημείο (-1,-1)*



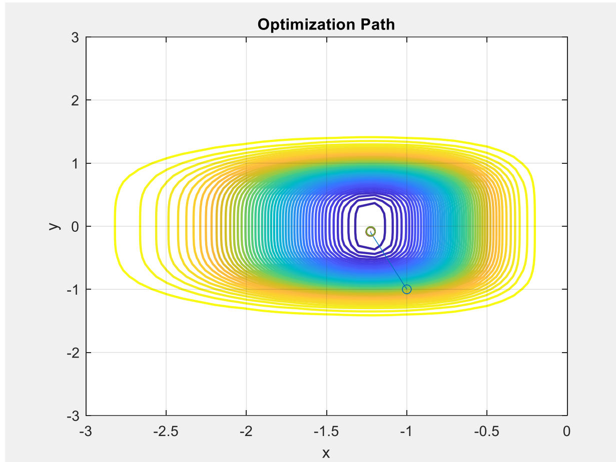



*Figure 5. Το $\gamma_k$ προκύπτει από την ελαχιστοποίησης της $f(x_{k+1}+\gamma_k d_k)$ για αρχικό σημείο (-1, -1)*

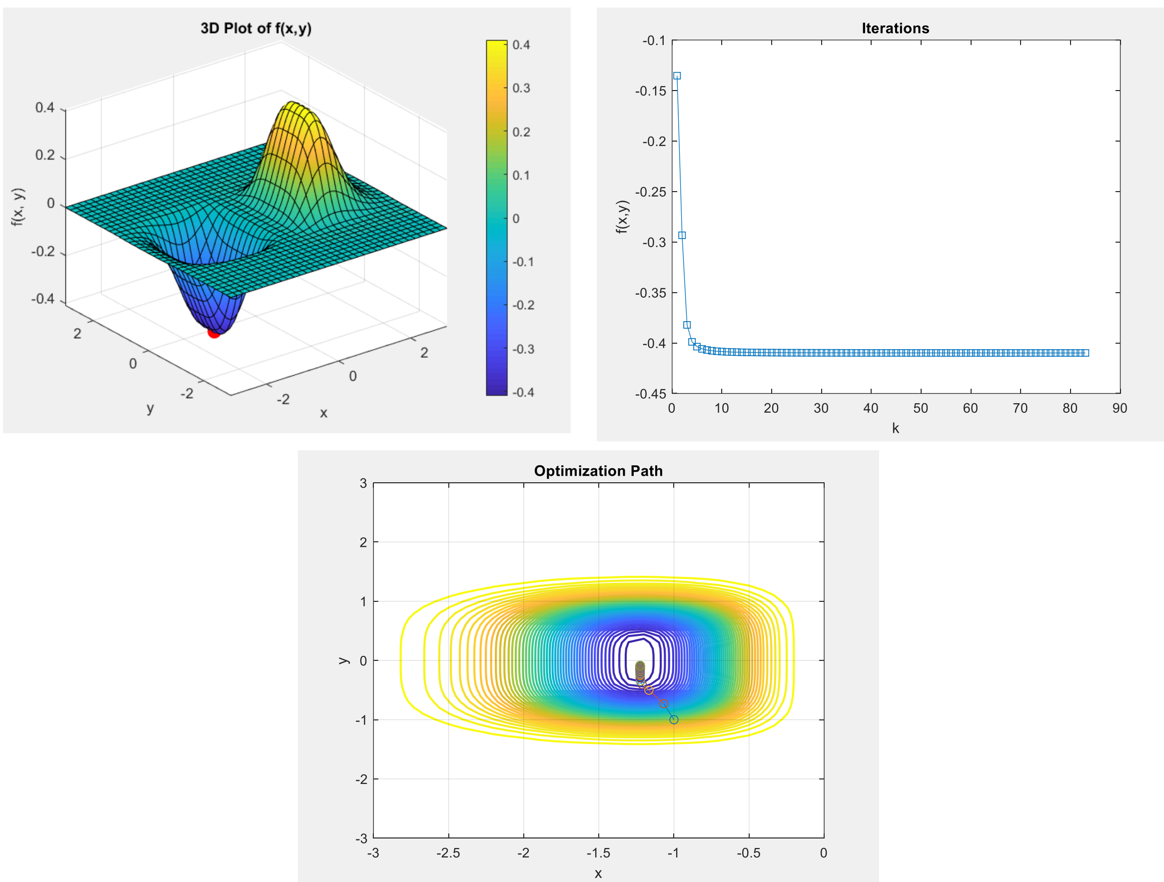



*Figure 6. Το γk υπολογίζεται βάσει του κανόνα Armijo για το σημείο (-1,-1)*

### c) Με αρχικό σημείο $(\mathbf{x_0}, \mathbf{y_0}) = (\mathbf{1}, \mathbf{1})$

Όπως παρατηρούμε από τα Figure 7, Figure 8, Figure 9, για όλες τις μεθόδους καθορισμού του $\gamma_k$ ο αλγόριθμος αποτυγχάνει να πετύχει σύγκλιση στο σωστό σημείο, δηλαδή το ελάχιστο της $f$. Μετά από κάποιο αριθμό επαναλήψεων, ο αλγόριθμος οδηγείται στο επίπεδο $xy$ όπου η κλίση είναι 0 και έτσι εγκλωβιζόμαστε πάλι σε τοπικό ελάχιστο.

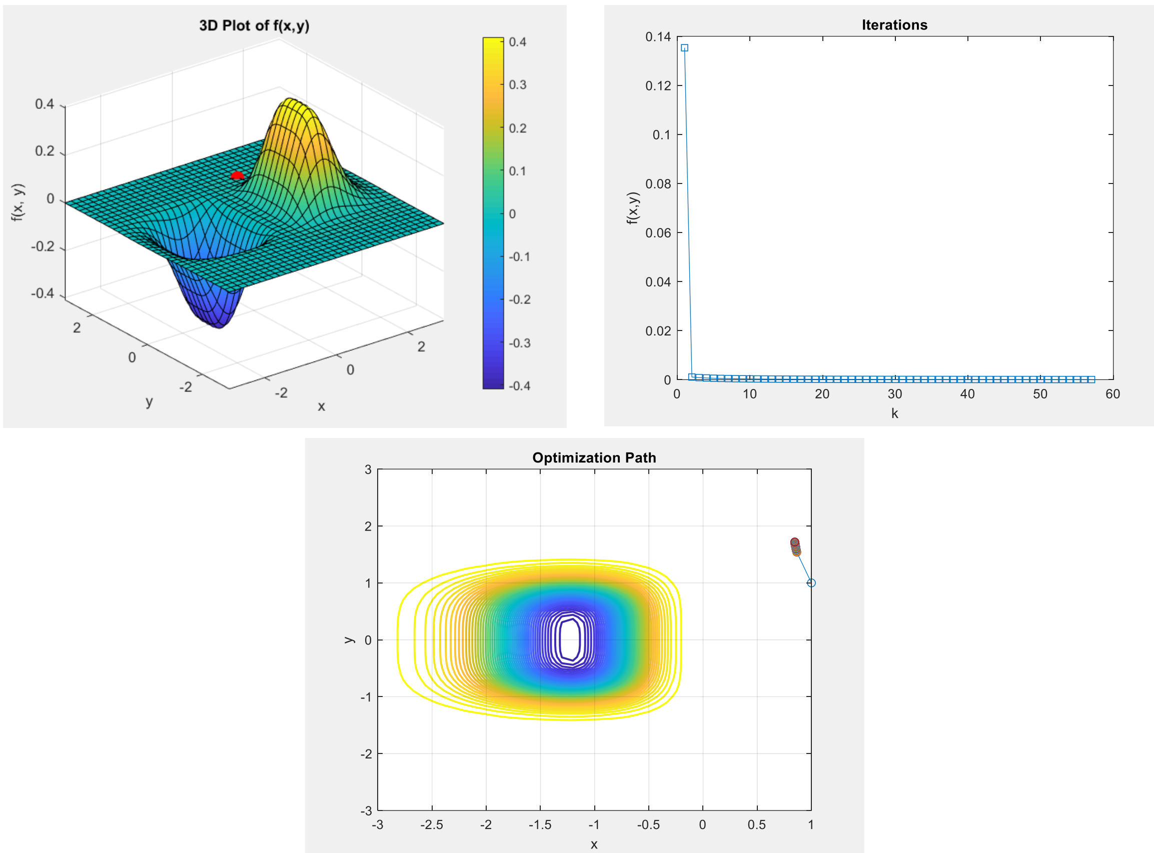
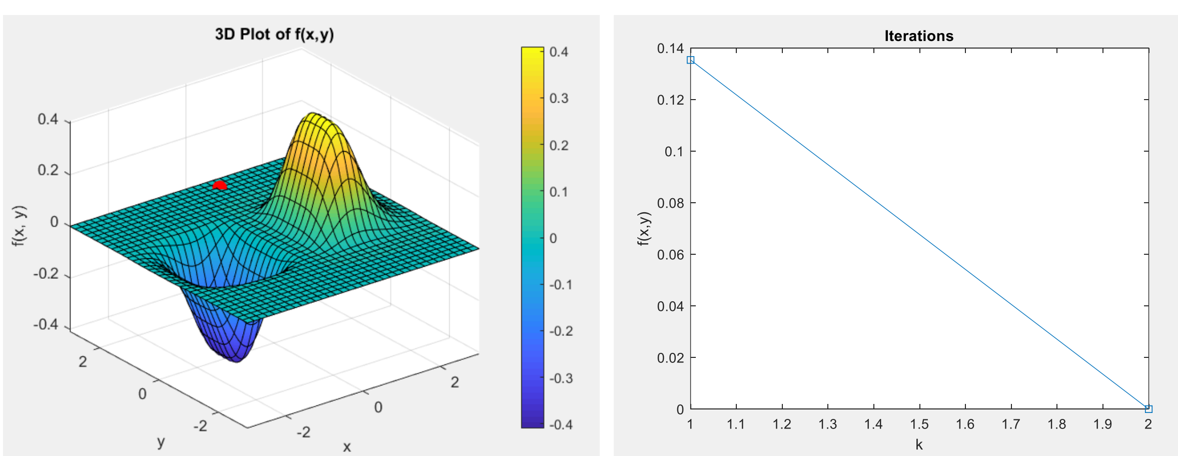



*Figure 7. Σταθερό $\gamma_k$ για αρχικό σημείο (1,1)*



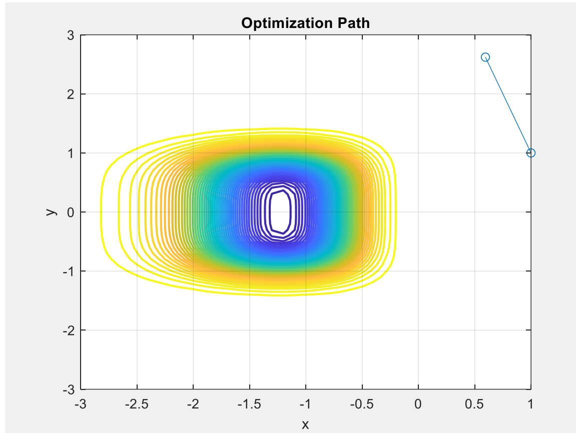



*Figure 8. Το $\gamma_k$ προκύπτει από την ελαχιστοποίησης της $f(x_{k+1}+\gamma_k d_k)$ για αρχικό σημείο (1, 1)*

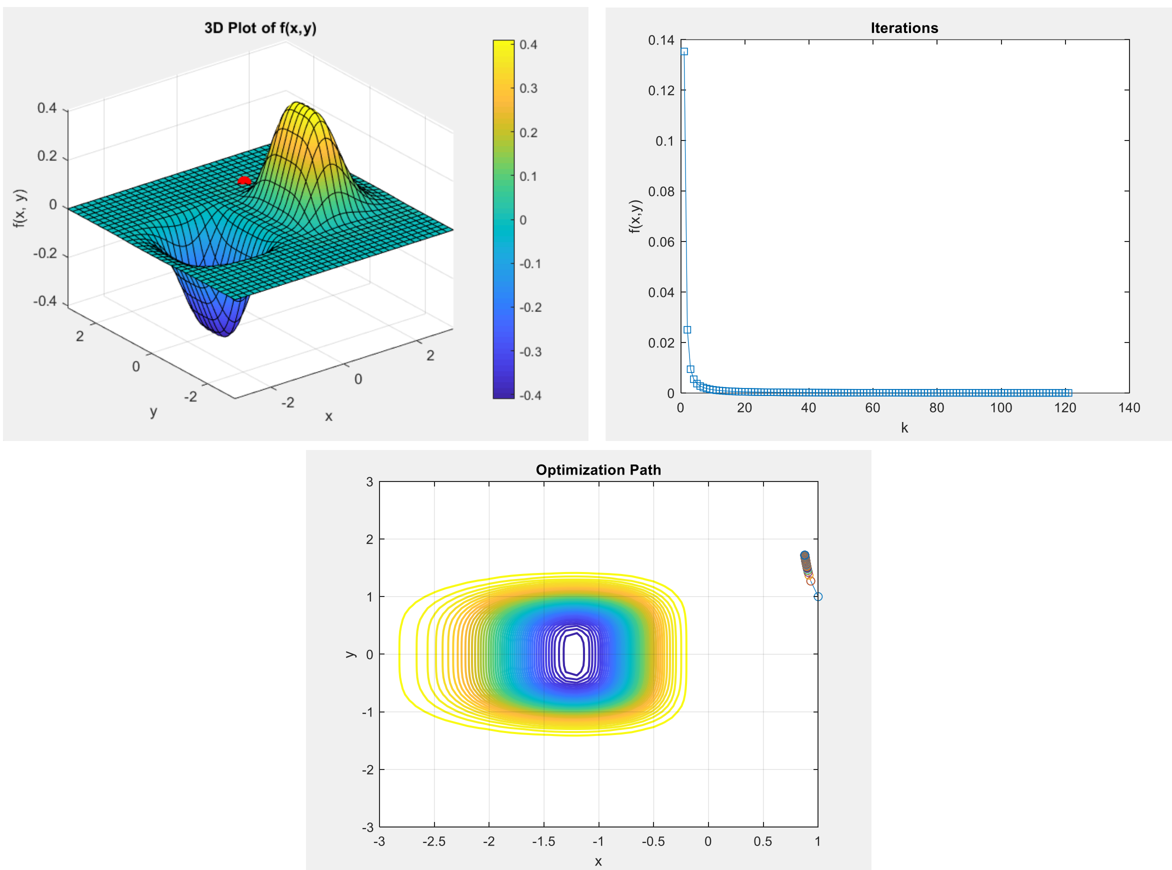



*Figure 9. Το γk υπολογίζεται βάσει του κανόνα Armijo για το σημείο (1,1)*

## Θέμα 3ο: Μέθοδος Newton

Η μέθοδος Newton προκύπτει αν στην (2) επιλέξουμε:

$$\Delta_k = [\nabla^2 f(x_k)]^{-1}, k = 0,1,2, \dots \tag{6}$$

με την προϋπόθεση ότι ο $\nabla^2 f(x_k)$ είναι θετικά ορισμένος. Αντικαθιστώντας την (6) στην (2) βρίσκουμε:

$$x_{k+1} = x_k - \gamma_k [\nabla^2 f(x_k)]^{-1} \nabla f(x_k) \tag{7}$$

Η μέθοδος του Newton Λοιπόν αναζητά το σημείο ελαχίστου στην κατεύθυνση:

$$d_k = -[\nabla^2 f(x_k)]^{-1} \nabla f(x_k) \tag{8}$$

Ως σταθερά τερματισμού $\varepsilon$ επιλέγουμε πάλι $\varepsilon = 10^{-3}$. Στον κώδικα που υλοποιήθηκε σε matlab για την μέθοδο Newton υπάρχει και έλεγχος που δείχνει εάν ο εσσιανός πίνακας της συνάρτησης είναι θετικά ορισμένος σε συγκεκριμένο σημείο.

Όπως παρατηρούμε, λοιπόν, αφού τρέξουμε τον αλγόριθμο ότι αδυνατεί να εφαρμοστεί για τα συγκεκριμένα σημεία αφετηρίας που μας έχουν δοθεί. Ο πίνακας $\nabla f(x_k)$ δεν είναι θετικά ορισμένος και έτσι ο αλγόριθμος τερματίζει αμέσως μετά τον εντοπισμό του μη θετικά ορισμένου πίνακα.

## Θέμα 3ο: Μέθοδος Lavenberg-Marquardt

Αν ο $\nabla^2 f(x_k)$ δεν είναι θετικά ορισμένος $\forall x \in \mathbb{R}^n$ τότε μπορούμε να χρησιμοποιήσουμε τον τροποποιημένο αλγόριθμο Newton ή μέθοδο Lavenberg-Marquardt.

Σύμφωνα με τον αλγόριθμο εφαρμόζουμε διαγώνια φόρτιση $(\nabla^2 f(x_k) + \mu_k I)$ στον εσσιανό πίνακα της συνάρτησης έως ότου να γίνει θετικά ορισμένος. Ο αριθμός $\mu_k$ που προστίθεται στη διαγώνιο του εσσιανού διαφέρει ανάλογα την επιλογή μεθόδου εύρεσης του $\gamma_k$. Η επιλογή του $\mu_k$ και του $\gamma_k$ γίνεται έτσι ώστε να πληρούνται τα εξής κριτήρια:

**Κριτήριο 1:** Υπάρχει $\beta$ με $0 < \beta < 1$ τέτοιο ώστε

$$d_k{}^T \nabla f(x_{k+1}) > \beta d_k{}^T \nabla f(x_k) \tag{9}$$

**Κριτήριο 2:** Υπάρχει $\alpha$ με $0 < \alpha < 1$ τέτοιο ώστε

$$f(x_{k+1}) \leq f(x_k) + \alpha \gamma_k d_k{}^T \nabla f(x_k) \tag{10}$$

Αυτά διασφαλίζουν πως η τιμή του βήματος $\gamma_k$ δεν θα πάρει πολύ μικρές ή πολύ μεγάλες τιμές, ώστε να αποφευχθούν ατέρμονες επαναλήψεις του αλγορίθμου και αποκλίσεις.

Ως σταθερά τερματισμού $\varepsilon$ επιλέγουμε $\varepsilon = 10^{-3}$. Όσο μικρότερο επιλέγετε το $\varepsilon$ τόσες περισσότερες θα είναι οι επαναλήψεις του αλγορίθμου για σύγκλιση ώστε να ικανοποιείται η μεγαλύτερη ακρίβεια που ζητείται.

Μελετούμε, λοιπόν, τη συμπεριφορά του αλγορίθμου με αφετηρία το κάθε σημείο που μας ζητείται:

### a) Με αρχικό σημείο $(\mathbf{x_0}, \mathbf{y_0}) = (\mathbf{0}, \mathbf{0})$

Όταν ξεκινάμε την αναζήτηση του ελαχίστου από το σημείο $(0{,}0)$, ο αλγόριθμος τερματίζει από την πρώτη επανάληψη και το αποτέλεσμα που δίνει είναι το $(0{,}0)$. Αυτό συμβαίνει καθώς $\nabla f(0{,}0) = 0$ και άρα η συνθήκη τερματισμού $|\nabla f(x_k, y_k)| \leq \varepsilon$ πληρείται εξ αρχής.

Το παραπάνω συμπέρασμα ισχύει για όλες τις μεθόδους που καλούμαστε να ελέγξουμε και δεν μεταβάλλεται ανάλογα με την επιλογή του βήματος $\gamma_k$.

Συνεπώς, το σημείο $(0{,}0)$ δεν αποτελεί καλό σημείο εκκίνησης.

### b) Με αρχικό σημείο $(\mathbf{x_0}, \mathbf{y_0}) = (-\mathbf{1}, -\mathbf{1})$

Όπως παρατηρούμε από τα Figure 10, Figure 11, Figure 12, για όλες τις μεθόδους καθορισμού του $\gamma_k$ ξανά ο αλγόριθμος συγκλίνει στο σωστό σημείο, δηλαδή το ελάχιστο της $f$. Η διαφορά που υπάρχει ανάμεσα τους είναι ο αριθμός των επαναλήψεων της μεθόδου ώστε να συγκλίνουμε στο ελάχιστο.

- Στην περίπτωση του σταθερού βήματος, επιλέγουμε αρχικά βήμα $\gamma_k = 0.3$ για το οποίο προέκυψαν και οι γραφικές παραστάσεις στο Figure 10. Αξίζει να σημειώσουμε, ότι απαιτείται ένας σχετικά μεγάλος αριθμός επαναλήψεων για να συγκλίνουμε και να τερματίσει ο αλγόριθμος. Παρατηρούμε όμως ότι πλησιάζει πολύ γρήγορα το ελάχιστο της συνάρτησης. Οποιαδήποτε μικρή απόκλιση από αυτή την τιμή οδηγεί στην μη εκπλήρωση των κριτηρίων 1 και 2 είτε στην μεγάλη αύξηση του αριθμού επαναλήψεων είτε σε μικρή μεταβολή της συντεταγμένης $y$. Αυτό το βλέπουμε πιο καθαρά στον παρακάτω πίνακα:

| $\gamma_k$ | $k$ | $x^*$ | $y^*$ |
|---|---|---|---|
| 0.055 | 530 | -1.22 | -0.0847 |
| 0.065 | 128 | -1.22 | -0.0352 |
| 0.07 | 122 | -1.22 | 0.0372 |
| 0.08 | 367 | -1.22 | 0.0846 |
| 0.09 | 398 | -1.22 | 0.0847 |
| 0.1 | 379 | -1.22 | 0.0847 |
| 0.11 | 355 | -1.22 | 0.0846 |

| | |
|---|---|
| 0.15 | ΣΥΓΚΡΟΥΣΗ ΜΕ ΚΡΙΤΗΡΙΑ |
| 0.2 | ΣΥΓΚΡΟΥΣΗ ΜΕ ΚΡΙΤΗΡΙΑ |

Στα Figure 10, Figure 11 φαίνονται τα διαγράμματα για $\gamma_k = 0.065$.

- Όπως και στην πρώτη μέθοδο, ο αλγόριθμος στον οποίο το $\gamma_k$ επιλέγεται τέτοιο ώστε να ελαχιστοποιεί την συνάρτηση $f(x_k + \gamma_k d_k)$ συγκλίνει σε πολύ λιγότερες επαναλήψεις σε σχέση με την επιλογή σταθερού $\gamma_k$. Η προσέγγιση το ελάχιστου σημείου επίσης, γίνεται πολύ γρήγορα όπως φαίνεται στο Figure 11.
- Στην μέθοδο Armijo επιλέγονται οι παράμετροι $\alpha = 10^{-3}, \beta = 0.2$. Το αρχικό βήμα επιλέγεται $\gamma_k = 0.25$, μετά από δοκιμές, ώστε να πετύχουμε όσο λιγότερες επαναλήψεις γίνεται. Ο αριθμός των επαναλήψεων βρίσκεται μεταξύ των δύο προηγούμενων μεθόδων όπως είναι φανερό στο Figure 12.

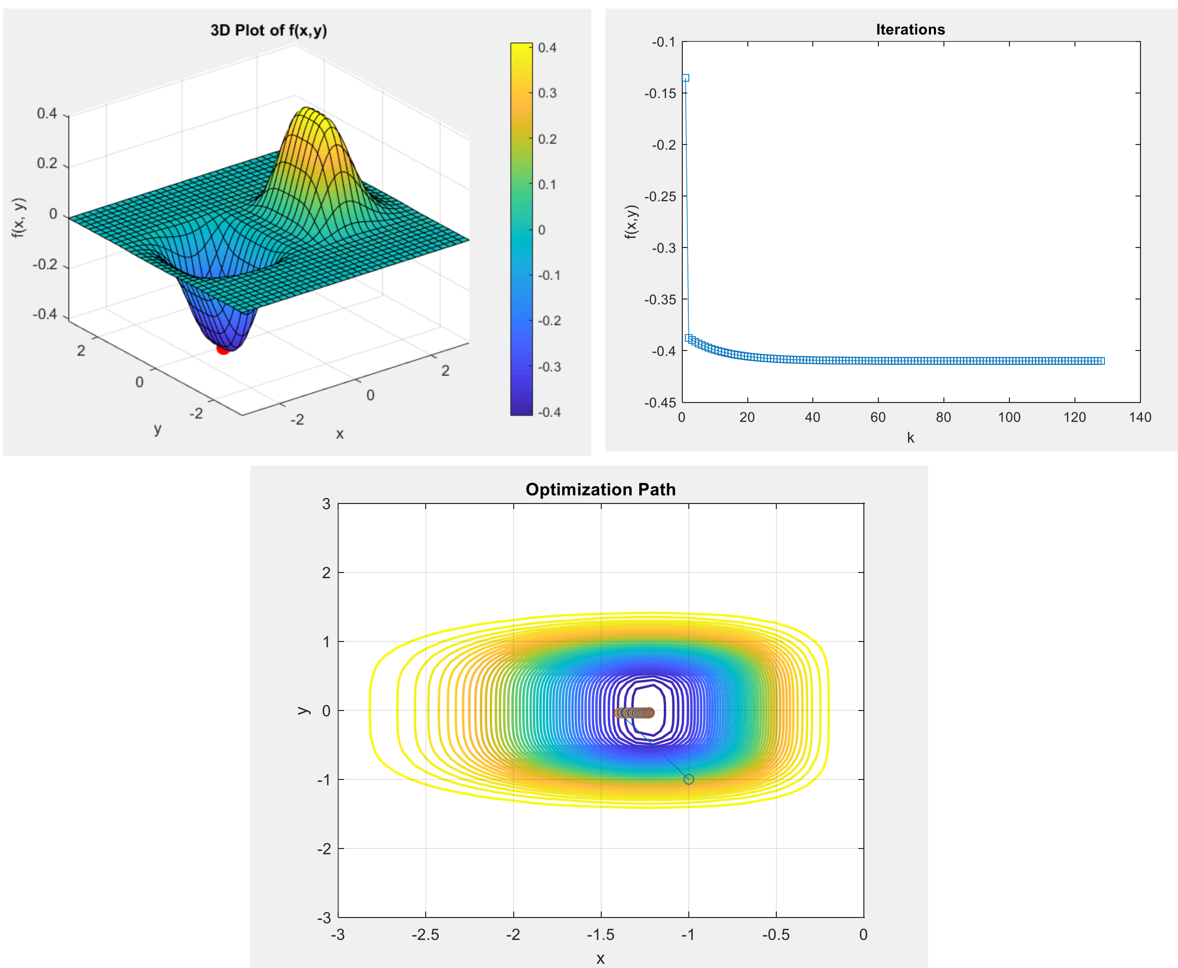



*Figure 10. Σταθερό $\gamma_k$ για αρχικό σημείο (-1,-1)*

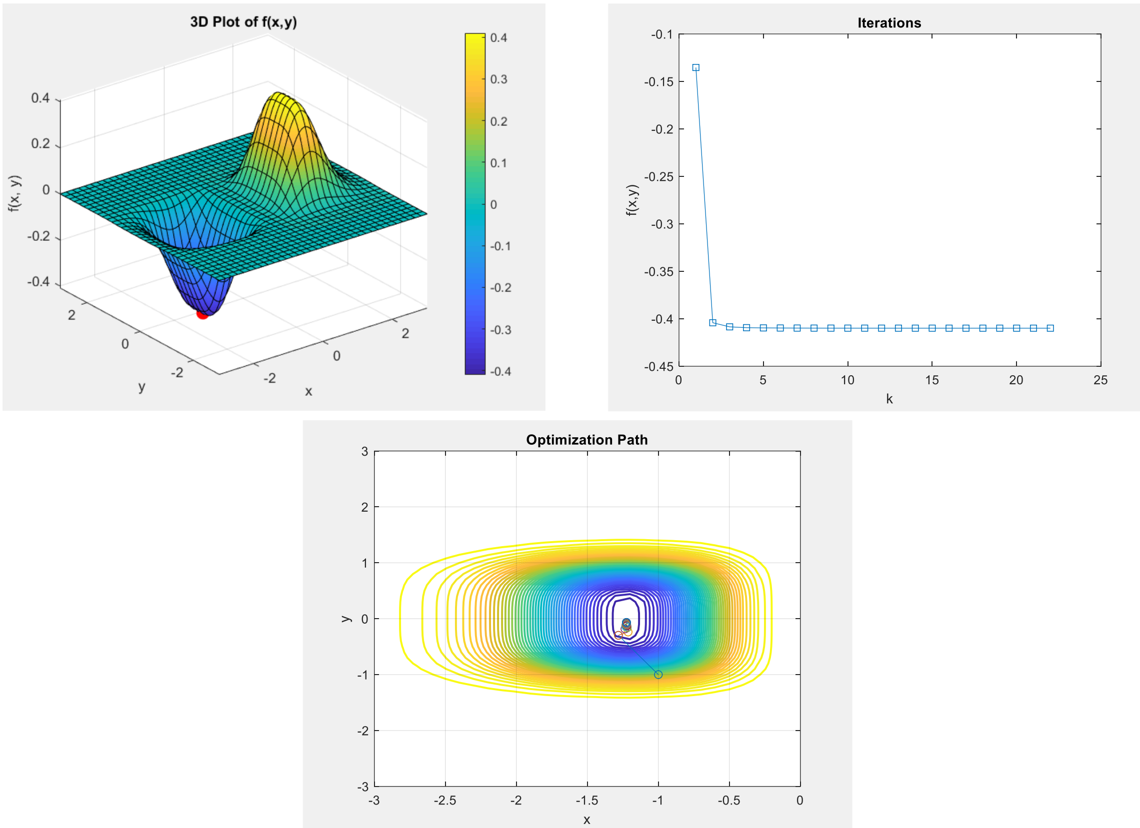
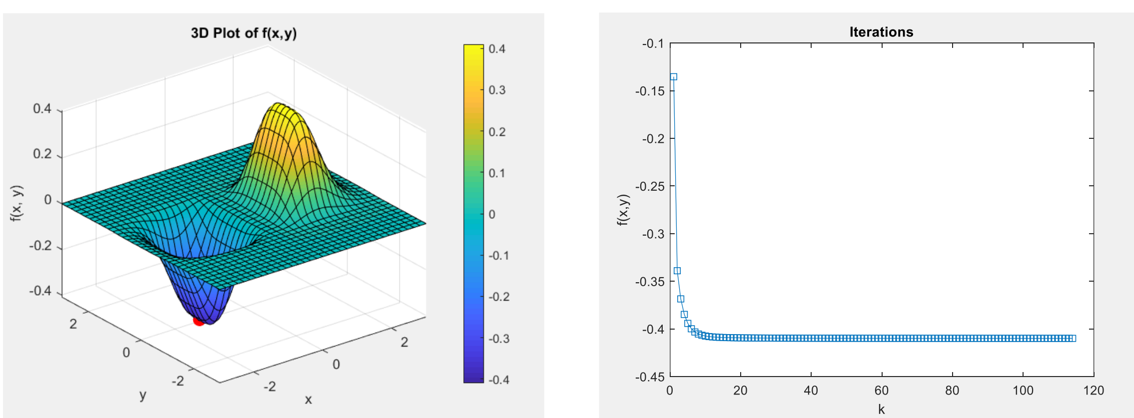



*Figure 11. Το $\gamma_k$ προκύπτει από την ελαχιστοποίησης της $f(x_{k+1}+\gamma_k d_k)$ για αρχικό σημείο (-1, -1)*



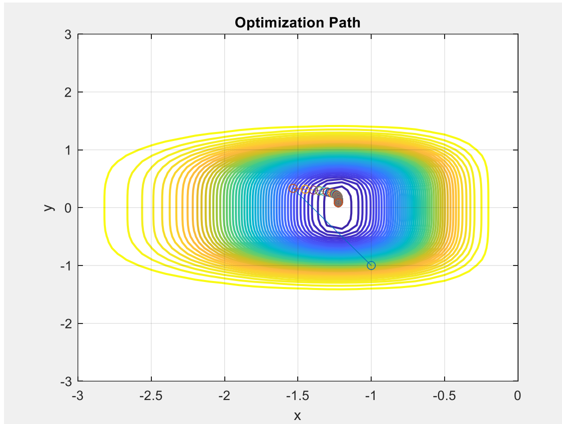



*Figure 12. Το $\gamma_k$ υπολογίζεται βάσει του κανόνα Armijo για το σημείο (-1,-1)*

### c) Με αρχικό σημείο $(\mathbf{x_0}, \mathbf{y_0}) = (\mathbf{1}, \mathbf{1})$

Όπως παρατηρούμε από τα Figure 13, Figure 14, Figure 15, παρόμοια με την μέθοδο μέγιστης καθόδου, για όλες τις μεθόδους καθορισμού του $\gamma_k$ ο αλγόριθμος αποτυγχάνει να πετύχει σύγκλιση στο σωστό σημείο, δηλαδή το ελάχιστο της $f$. Μετά από κάποιο αριθμό επαναλήψεων, ο αλγόριθμος οδηγείται στο επίπεδο $xy$ όπου η κλίση είναι 0 και έτσι εγκλωβιζόμαστε πάλι σε τοπικό ελάχιστο.

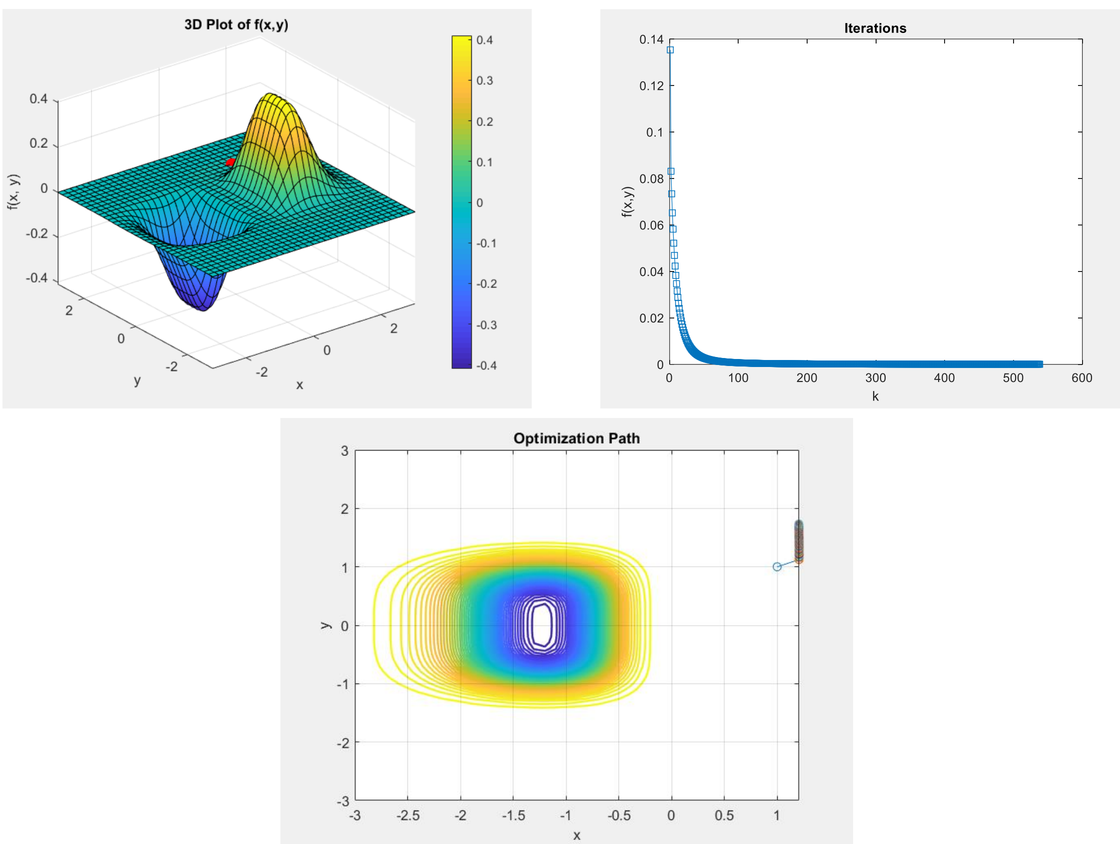



*Figure 13. Σταθερό $\gamma_k$ για αρχικό σημείο (1,1)*

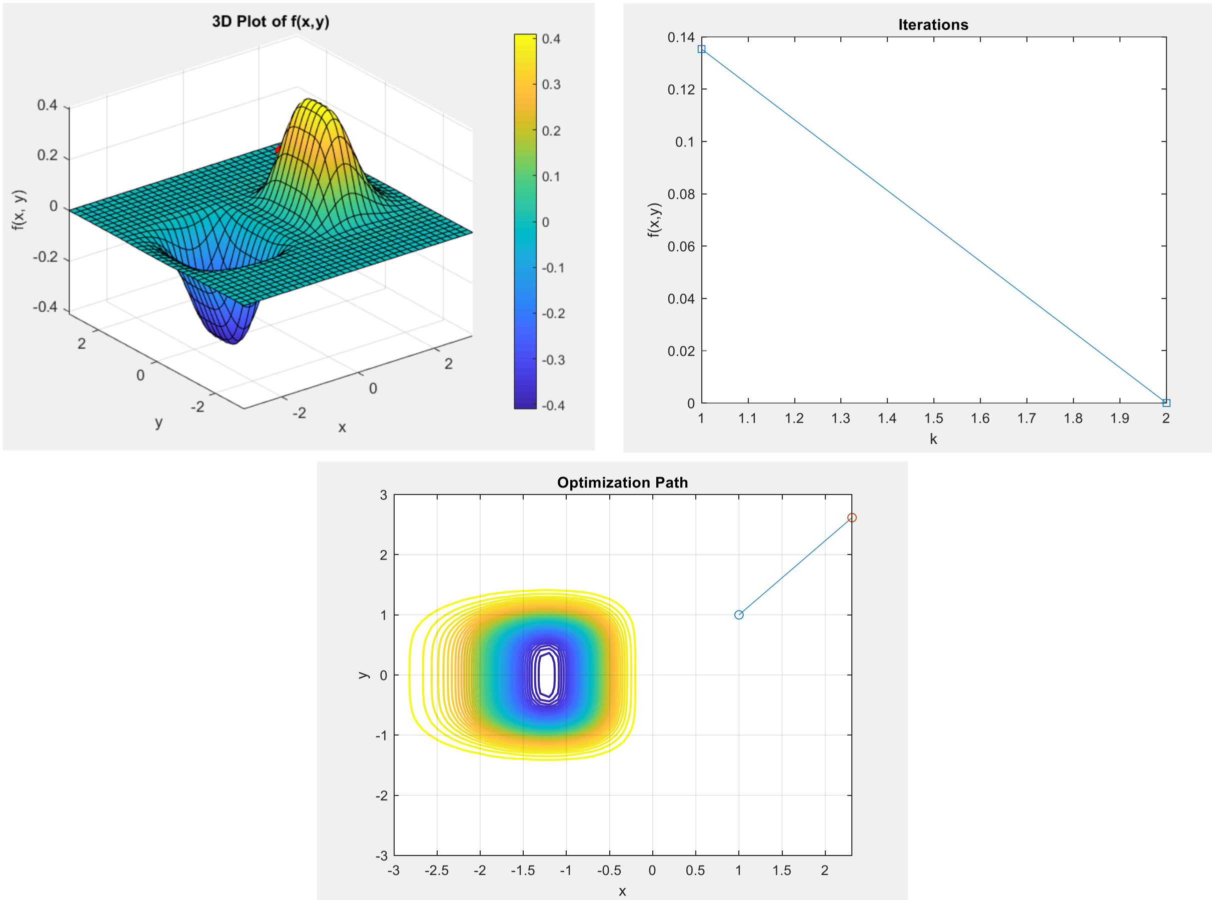
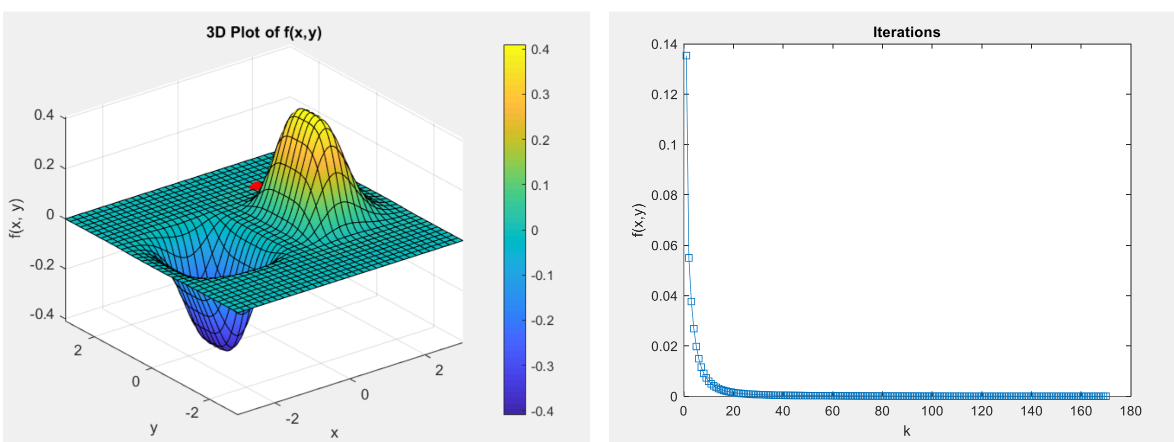



*Figure 14. Το $\gamma_k$ προκύπτει από την ελαχιστοποίησης της $f(x_{k+1}+\gamma_k d_k)$ για αρχικό σημείο (1, 1)*



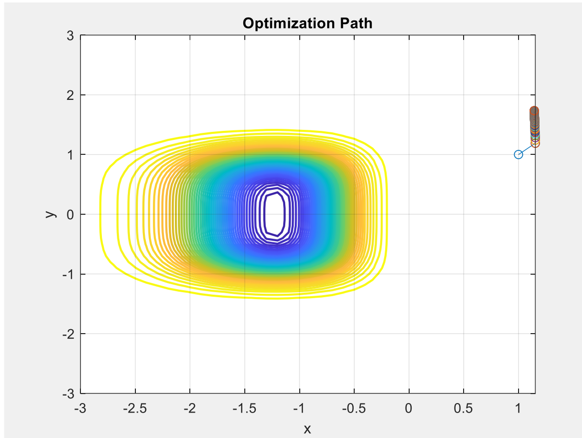



*Figure 15. Το $\gamma_k$ υπολογίζεται βάσει του κανόνα Armijo για το σημείο (1,1)*

# 3 Παρατηρήσεις

Στην παρούσα εργασία μελετήθηκαν τρείς μέθοδοι εύρεσης ελαχίστου μίας συνάρτησης πολλών μεταβλητών χωρίς περιορισμούς με χρήση παραγώγων, επιλέγοντας το βήμα $\gamma_k$ με τρείς διαφορετικούς τρόπους.

Και οι τρείς μέθοδοι έχουν θετικά και αρνητικά. Η ποιο αποτελεσματική από όλες σε θέμα ταχύτητας και ακρίβειας φαίνεται να είναι η Μέθοδος Μέγιστης Καθόδου. Είναι και η λιγότερη πολύπλοκη μέθοδος. Στην μέθοδο Newton από την άλλη χρειάζεται να πληρούνται περιορισμοί, καθώς εφαρμόζεται μόνο σε περιπτώσεις που ο εσσιανός πίνακας της αντικειμενικής συνάρτησης είναι θετικά ορισμένος. Η μέθοδος Levenberg-Marquardt έχει μεγάλη υπολογιστική πολυπλοκότητα σε σχέση με τις άλλες μεθόδους, ως μια τροποποιημένη έκδοση του αλγορίθμου Newton. Η μέθοδος αυτή είναι πολύ ευαίσθητη στην επιλογή της τιμής $\mu_k$ και πρέπει να ικανοποιεί πολλούς περιορισμούς.

Θα μπορούσε εύκολα κανείς να βγάλει το συμπέρασμα ότι η μέθοδος μέγιστης καθόδου έχει καλύτερα αποτελέσματα από τις άλλες μεθόδους. Πιο συγκεκριμένα, εάν αναφερθούμε στην αποτελεσματικότητα των μεθόδων επιλογής του $\gamma_k$ που εφαρμόζουμε σε κάθε αλγόριθμο, αδιαμφησβήτητα υπερτερεί η μέθοδος εκλογής του $\gamma_k$ με σκοπό να ελαχιστοποιηθεί η συνάρτηση $f(x_k + \gamma_k d_k)$. Χρησιμοποιώντας αυτή την μέθοδο οδηγούμαστε με μεγάλη ταχύτητα και σχετικά μικρό αριθμό επαναλήψεων στο ελάχιστο της αντικειμενικής συνάρτησης. Η μέθοδος σταθερού $\gamma_k$ μπορεί να φανεί πολύ αποτελεσματική με την κατάλληλη επιλογή του βήματος $\gamma_k$, όμως χρειάζεται μεγάλη εμπειρία για τον προσδιορισμό του. Για την μέθοδο Armijo μπορούμε να πούμε, πως παρά τους αρκετούς περιορισμούς που μπορούμε να μεταβάλουμε, υπάρχει το θετικό ότι εξασφαλίζει πως το $\gamma_k$ δεν θα πάρει μεγάλες τιμές.

Για να λειτουργήσουν αποτελεσματικά οι αλγόριθμοι χρειάζεται η κατάλληλη επιλογή του σημείου εκκίνησης. Στη μέθοδο Newton αυτό είναι πολύ πιο εμφανές, καθώς μόνο εάν επιλέξουμε σημείο εκκίνησης αρκετά κοντά στο ελάχιστο ο αλγόριθμος το εντοπίζει σωστά. Αλλιώς, όπως σε όλες τις μεθόδους, ο αλγόριθμός είναι πιθανό να εγκλωβιστεί σε τοπικά ακρότατα και να μην οδηγηθεί στο ολικό ελάχιστο. Τα σημεία που ζητήθηκε να μελετήσουμε δεν φέρουν στην πλειοψηφία τους ένα σωστό αποτέλεσμα, με εξαίρεση το $(-1, -1)$ το οποίο βρίσκεται πιο κοντά στο ολικό ελάχιστο σε σχέση με τα άλλα δύο, αλλά και πάλι αδυνατεί να σταθεί αρκετό για την σωστή διεκπεραίωση της μεθόδου Newton.