

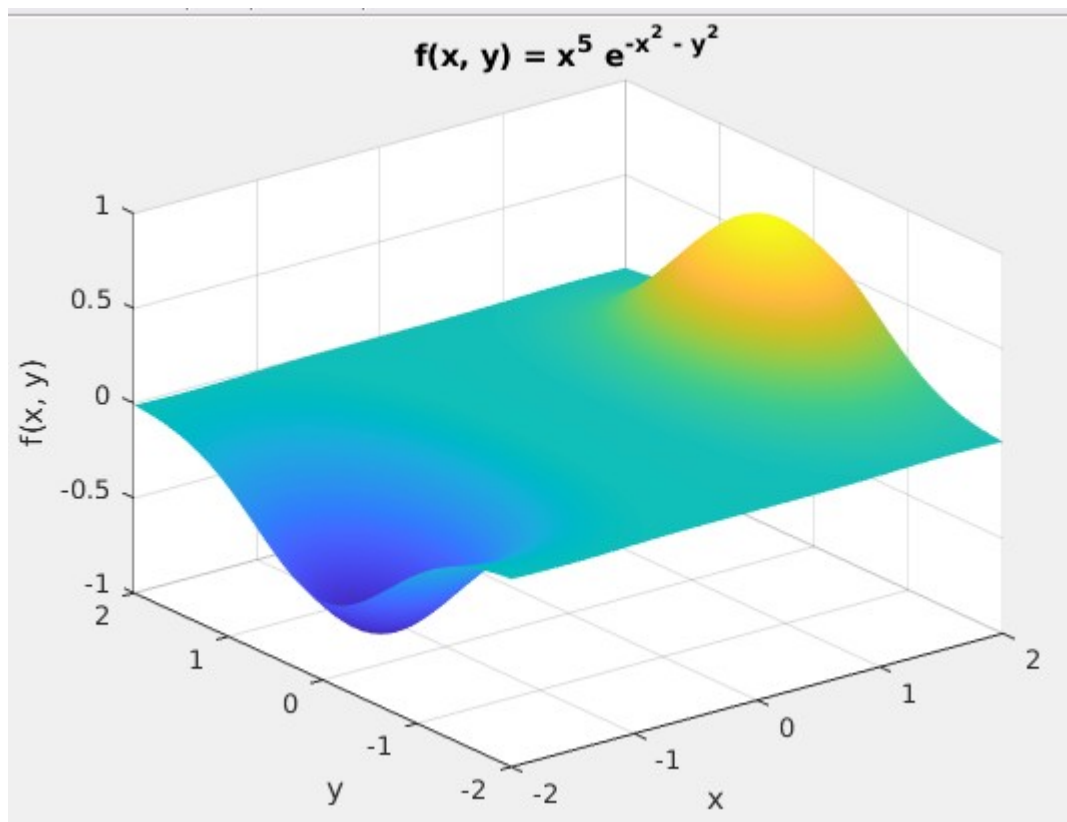
# ΤΕΧΝΙΚΕΣ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ

## 2η Εργαστηριακή Άσκηση

Καράτης Δημήτριος 10775

### Θέμα 1:

Σχεδιάζοντας την γραφική παράσταση της  $f(x,y) = x^5 e^{-x^2-y^2}$  έχουμε:

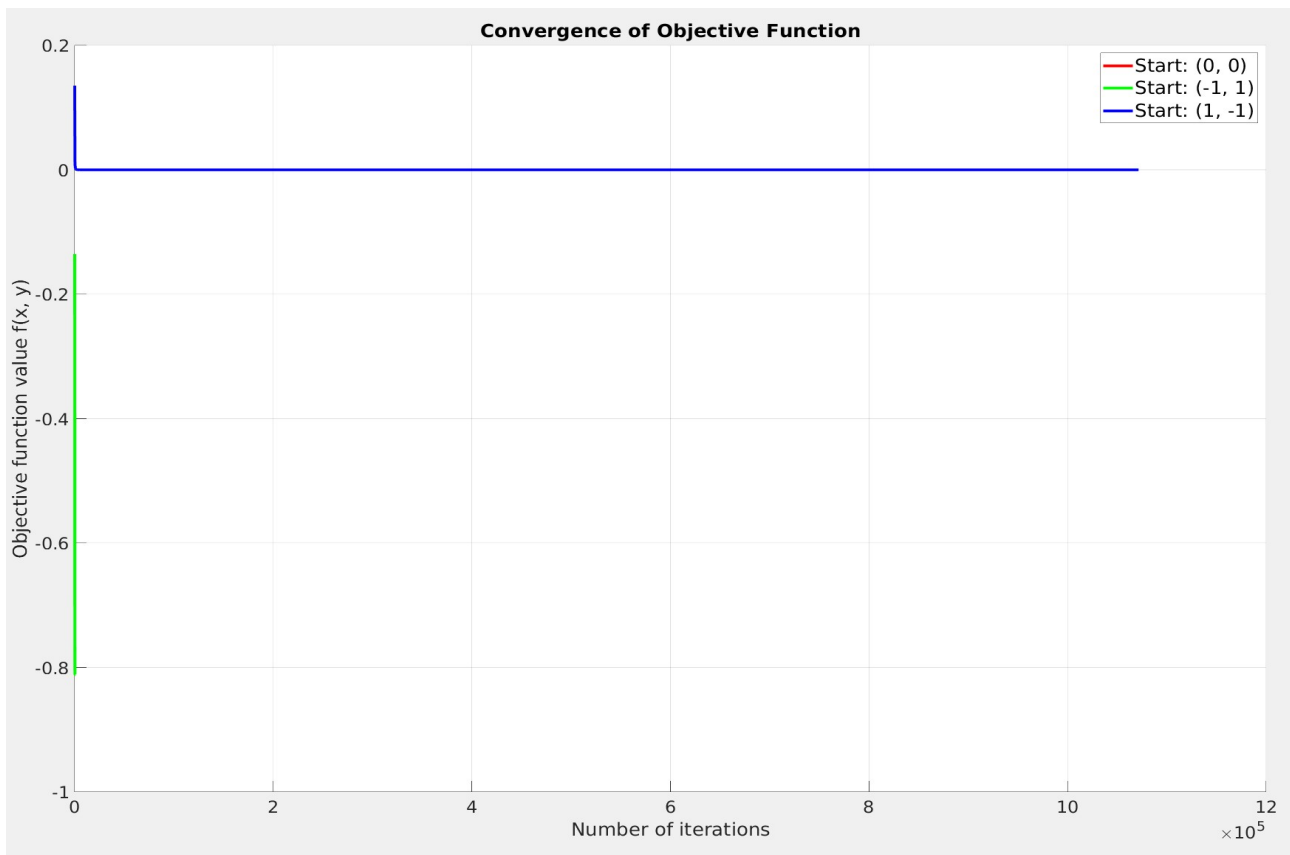


Επίσης, η ελάχιστη τιμή της  $f$  είναι  $-0.811174$  για  $(x, y)$ :  $(-1.5811, 0.0000)$ .

Αναλυτικότερη επεξήγηση και κώδικας βρίσκεται στο αρχείο *work2\_ex1.m*

## Θέμα 2:

α) Για  $\gamma_k = 0.01$ , προκύπτει η εξής γραφική παράσταση της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων που απαιτούνται μέχρι να τερματίσει ο αλγόριθμος:



ενώ, για τα διάφορα σημεία έναρξης του αλγορίθμου ισχύει:

```
Initial point: (0.00, 0.00)
Minimum found at: (0.0000, 0.0000)
Final value f(x, y) = 0.000000
Number of iterations: 1

Initial point: (-1.00, 1.00)
Minimum found at: (-1.5811, 0.0000)
Final value f(x, y) = -0.811174
Number of iterations: 943

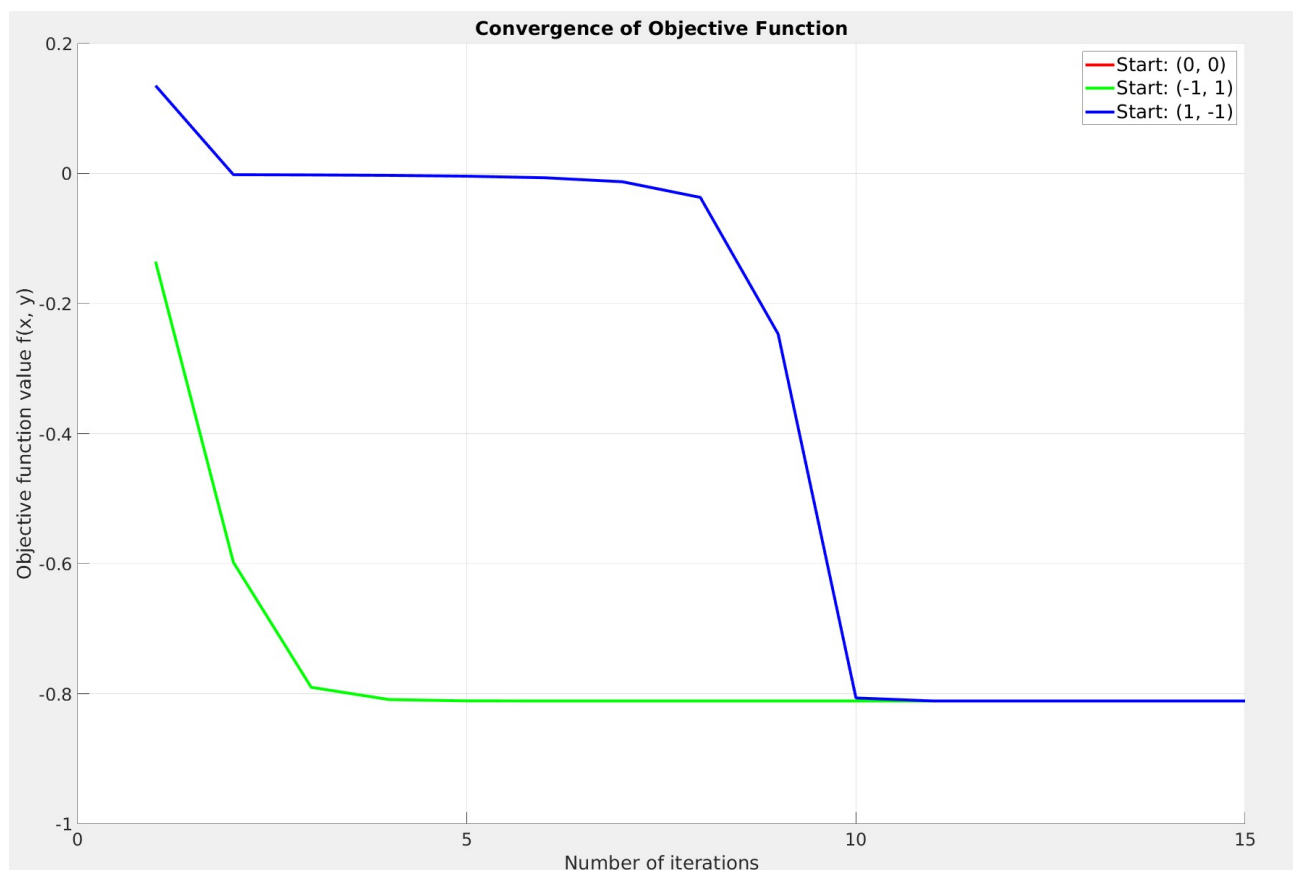
Initial point: (1.00, -1.00)
Minimum found at: (0.0321, -1.2913)
Final value f(x, y) = 0.000000
Number of iterations: 1071508
```

Με βάσει αυτά είναι προφανές ότι η μέθοδος της **Μέγιστης Καθόδου για σταθερό  $\gamma_k$** , έδωσε σωστό αποτέλεσμα (-0.8111174) μόνο στην περίπτωση όπου το σημείο έναρξης ήταν το (-1.0, 1.0). Στην περίπτωση που ήταν το (0.0, 0.0) παρατηρούμε ότι “εγκλωβίζεται” σε τοπικό ελάχιστο και δεν μπορεί να διαφύγει, ενώ, αν είναι το (1.0, -1.0) παρατηρούμε ότι συμβαίνει το ίδιο.

Η διαδικασία της μεθόδου ξεκινά από ένα αρχικό σημείο, όπου υπολογίζεται η κλίση (gradient) της συνάρτησης, η οποία υποδεικνύει την κατεύθυνση της μέγιστης αύξησής της. Στη συνέχεια, κινείται προς την αντίθετη κατεύθυνση (μέγιστης καθόδου) για να μειώσει την τιμή της συνάρτησης. Η μέθοδος υπολογίζει το κατάλληλο βήμα (step size) κατά μήκος αυτής της κατεύθυνσης, ώστε να εξασφαλίσει τη μεγαλύτερη δυνατή μείωση σε κάθε επανάληψη. Τα πλεονεκτήματα της μεθόδου περιλαμβάνουν την απλότητά της και την ευκολία εφαρμογής σε διάφορα προβλήματα. Ωστόσο, παρουσιάζει σημαντικά μειονεκτήματα, όπως η αργή σύγκλιση (μεγάλος αριθμός επαναλήψεων) σε προβλήματα με απότομες και ακανόνιστες καμπύλες, καθώς και η εξάρτηση από το βήμα, που μπορεί να οδηγήσει σε αναποτελεσματική πορεία προς το ελάχιστο ή ακόμα και σε στασιμότητα.

Αναλυτικότερη επεξήγηση και κώδικας βρίσκεται στο αρχείο *work2\_ex2\_a.m*

β) Επιλέγοντας  $\gamma_k$  τέτοιο ώστε να ελαχιστοποιεί την  $f(x_k + \gamma_k d_k)$ : Για τον σκοπό αυτόν, θα χρησιμοποιήσουμε την μέθοδο του **Χρυσού Τομέα**, γνωστή από την προηγούμενη εργασία. Έτσι, έχουμε τις εξής γραφικές παραστάσεις.



ενώ, για τα διάφορα σημεία έναρξης του αλγορίθμου ισχύει:

```
Initial point: (0.00, 0.00)
Minimum found at: (0.0000, 0.0000)
Number of iterations: 1
Final value f(x, y) = 0.000000

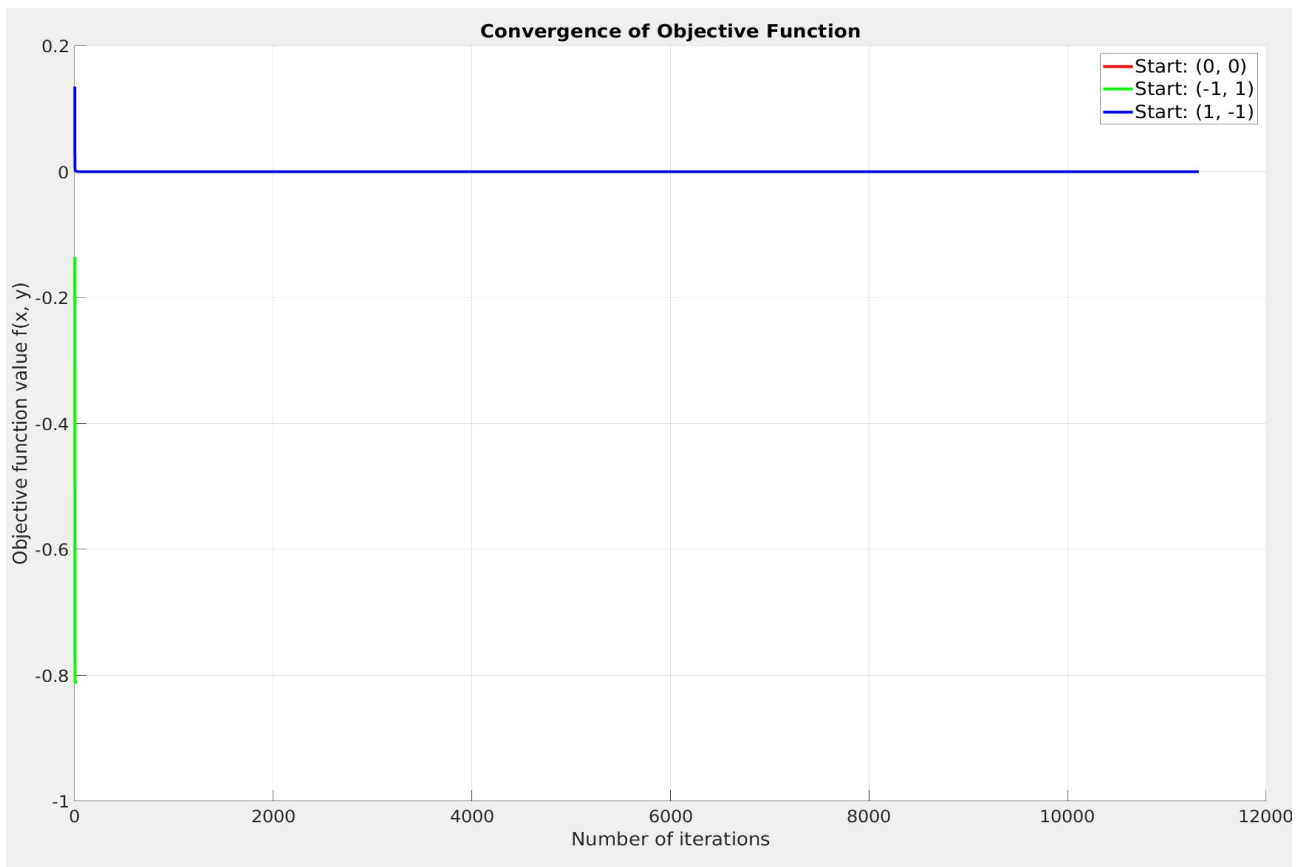
Initial point: (-1.00, 1.00)
Minimum found at: (-1.5811, 0.0000)
Number of iterations: 14
Final value f(x, y) = -0.811174

Initial point: (1.00, -1.00)
Minimum found at: (-1.5811, -0.0000)
Number of iterations: 15
Final value f(x, y) = -0.811174
```

Με βάσει αυτά, είναι προφανές ότι η μέθοδος της **Μέγιστης Καθόδου για  $\gamma_k$  που ελαχιστοποιεί την  $f(\mathbf{x}_k + \gamma_k \mathbf{d}_k)$**  έδωσε σωστό αποτέλεσμα (-0.811174) για σημεία έναρξης τα (1.0, -1.0) και (-1.0, 1.0), (έχουμε σημαντικά λιγότερες επαναλήψεις σε σχέση με πριν). Στην περίπτωση που ήταν το (0.0, 0.0) παρατηρούμε ότι και πάλι “εγκλωβίζεται” στο τοπικό ελάχιστο 0.0 και δεν μπορεί να διαφύγει.

*Αναλυτικότερη επεξήγηση και κώδικας βρίσκεται στο αρχείο `work2_ex2_b.m`*

γ) Επιλέγοντας το  $\gamma_k$  με βάσει τον κανόνα του Armijo: Εφαρμόζοντας τον κανόνα στην σελίδα 140 του βιβλίου προκύπτουν οι παρακάτω γραφικές παραστάσεις:



ενώ, για τα διάφορα σημεία έναρξης του αλγορίθμου ισχύει:

```
Initial point: (0.00, 0.00)
Minimum found at: (0.0000, 0.0000)
Final value  $f(x, y) = 0.000000$ 
Number of iterations: 1

Initial point: (-1.00, 1.00)
Minimum found at: (-1.5811, 0.0000)
Final value  $f(x, y) = -0.811174$ 
Number of iterations: 25

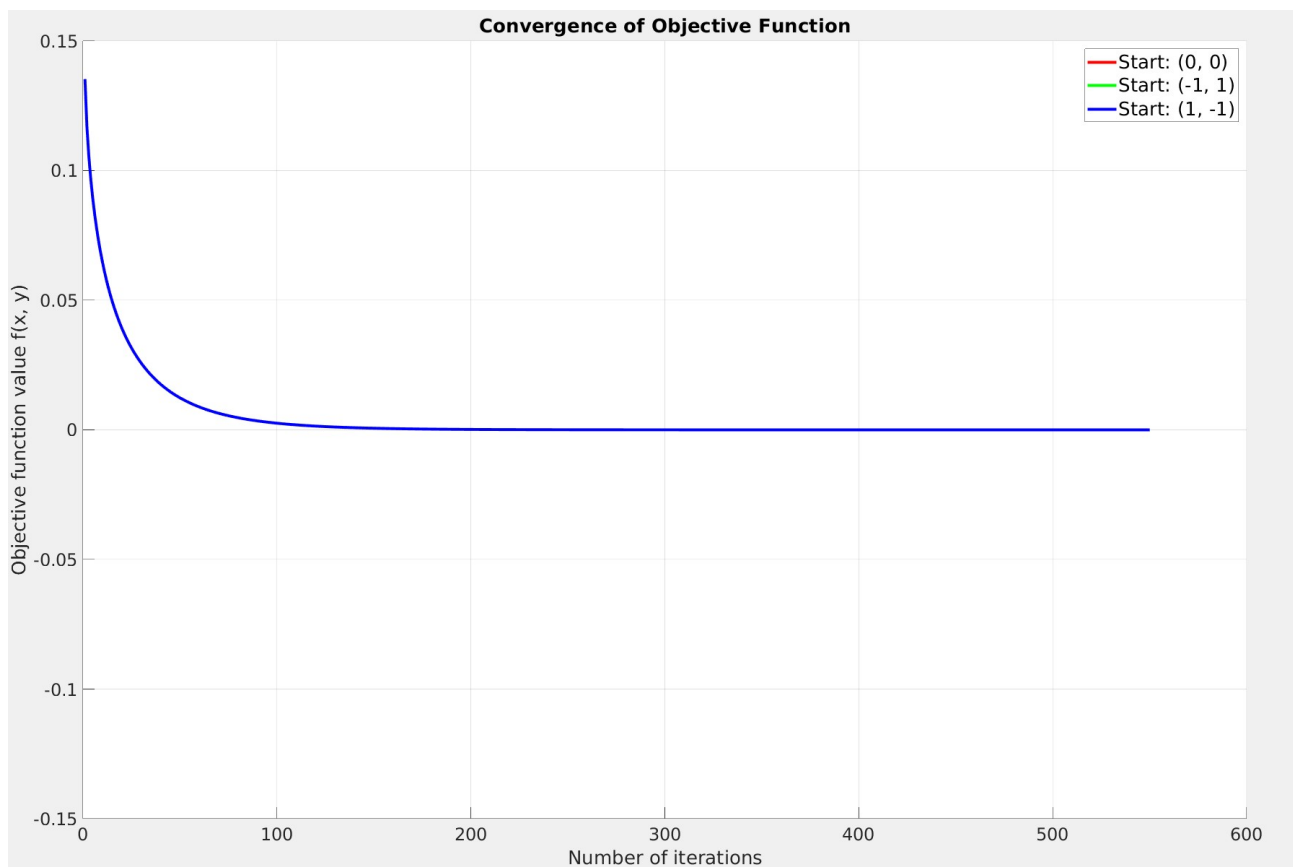
Initial point: (1.00, -1.00)
Minimum found at: (0.0339, -1.3746)
Final value  $f(x, y) = 0.000000$ 
Number of iterations: 11322
```

Με βάσει αυτά, είναι προφανές ότι η μέθοδος της **Μέγιστης Καθόδου για  $\gamma_k$  με βάσει τον κανόνα Armijo** έδωσε σωστό αποτέλεσμα (-0.8111174) μόνο στην περίπτωση όπου το σημείο έναρξης ήταν το (-1.0, 1.0), ωστόσο χρειάστηκαν περισσότερες επαναλήψεις σε σχέση με πριν. Στην περίπτωση που ήταν το (0.0, 0.0) παρατηρούμε ότι “εγκλωβίζεται” στο τοπικό ελάχιστο και δεν μπορεί να διαφύγει, ενώ, αν είναι το (1.0, -1.0) παρατηρούμε ότι συμβαίνει και πάλι το ίδιο.

Αναλυτικότερη επεξήγηση και κώδικας βρίσκεται στο αρχείο *work2\_ex2\_c.m*

### Θέμα 3:

α) Για  $\gamma_k = 0.01$ , προκύπτει η εξής γραφική παράσταση της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων που απαιτούνται μέχρι να τερματίσει ο αλγόριθμος:



ενώ, για τα διάφορα σημεία έναρξης του αλγορίθμου ισχύει:

```
Initial point: (0.00, 0.00)
Minimum found at: (0.0000, 0.0000)
Final value f(x, y) = 0.000000
Number of iterations: 1

Warning: Hessian is not positive definite!
> In work2_ex3_a (line 63)
Initial point: (-1.00, 1.00)
Minimum found at: (-1.0000, 1.0000)
Final value f(x, y) = -0.135335
Number of iterations: 1

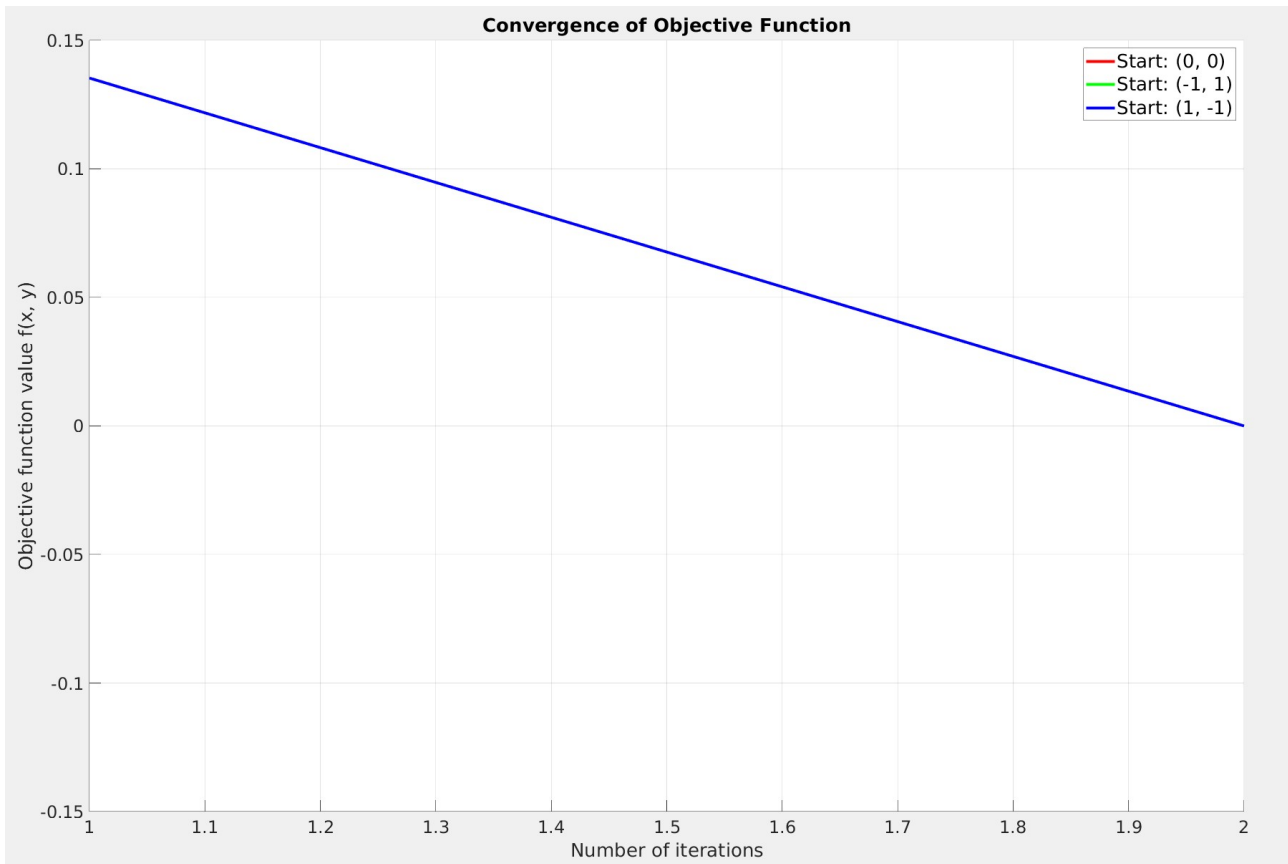
Initial point: (1.00, -1.00)
Minimum found at: (0.1796, -2.9240)
Final value f(x, y) = 0.000000
Number of iterations: 550
```

Με βάσει αυτά είναι προφανές ότι η μέθοδος **Newton για σταθερό  $\gamma_k$**  δεν κατάφερε να δώσει σωστό αποτέλεσμα (-0.8111174) σε καμία περίπτωση. Στην περίπτωση που το σημείο έναρξης ήταν το (0.0, 0.0) παρατηρούμε ότι “εγκλωβίζεται” στο τοπικό ελάχιστο 0.0 και δεν μπορεί να διαφύγει, ενώ, αν είναι το (1.0, -1.0) παρατηρούμε συμβαίνει το ίδιο. Τέλος στην περίπτωση που το αρχικό σημείο είναι το (-1.0, 1.0) η μέθοδος δεν μπορεί να εφαρμοστεί διότι ο Εσσιανός πίνακας δεν είναι θετικά ορισμένος.

Η μέθοδος Newton είναι μια αριθμητική τεχνική βελτιστοποίησης που χρησιμοποιείται για την εύρεση των στατικών σημείων (π.χ. ελαχίστων ή μεγίστων) μιας συνάρτησης. Βασίζεται στη χρήση της δεύτερης παραγώγου (ή του Εσσιανού, Hessian matrix, για πολυδιάστατες συναρτήσεις) για να προσαρμόσει την κατεύθυνση και το μέγεθος του βήματος σε κάθε επανάληψη. Σε κάθε βήμα, η μέθοδος υπολογίζει τη βαθμίδα της συνάρτησης για να προσδιορίσει την κατεύθυνση και χρησιμοποιεί τον Εσσιανό για να υπολογίσει τη βέλτιστη προσαρμογή της κίνησης. Ένα από τα μεγαλύτερα πλεονεκτήματά της είναι η ταχεία σύγκλιση, ειδικά όταν η αρχική προσέγγιση είναι κοντά στο ελάχιστο. Επιπλέον, προσαρμόζει δυναμικά το βήμα ανάλογα με την καμπυλότητα της συνάρτησης, καθιστώντας την πολύ αποτελεσματική για καλά συμπεριφερόμενες συναρτήσεις. Ωστόσο, παρουσιάζει μειονεκτήματα, όπως το υψηλό υπολογιστικό κόστος, καθώς απαιτεί τον υπολογισμό και την αντιστροφή του Εσσιανού, κάτι που είναι απαγορευτικό για προβλήματα μεγάλης διάστασης. Επιπλέον, σε συναρτήσεις με περίπλοκη καμπυλότητα, μπορεί να οδηγήσει σε αναποτελεσματικά ή λανθασμένα αποτελέσματα.

*Αναλυτικότερη επεξήγηση και κώδικας βρίσκεται στο αρχείο work2\_ex3\_a.m*

β) Επιλέγοντας  $\gamma_k$  τέτοιο ώστε να ελαχιστοποιεί την  $f(x_k + \gamma_k d_k)$ : Για τον σκοπό αυτόν, θα χρησιμοποιήσουμε την μέθοδο του **Χρυσού Τομέα**, γνωστή από την προηγούμενη εργασία. Έτσι, έχουμε τις εξής γραφικές παραστάσεις:



ενώ, για τα διάφορα σημεία έναρξης του αλγορίθμου ισχύει:

```
Initial point: (0.00, 0.00)
Minimum found at: (0.0000, 0.0000)
Final value f(x, y) = 0.000000
Number of iterations: 1
```

**Warning: Hessian is not positive definite!**

> In work2\_ex3\_b (line 63)

```
Initial point: (-1.00, 1.00)
Minimum found at: (-1.0000, 1.0000)
Final value f(x, y) = -0.135335
Number of iterations: 1
```

```
Initial point: (1.00, -1.00)
Minimum found at: (-0.2311, -4.1657)
Final value f(x, y) = -0.000000
Number of iterations: 2
```

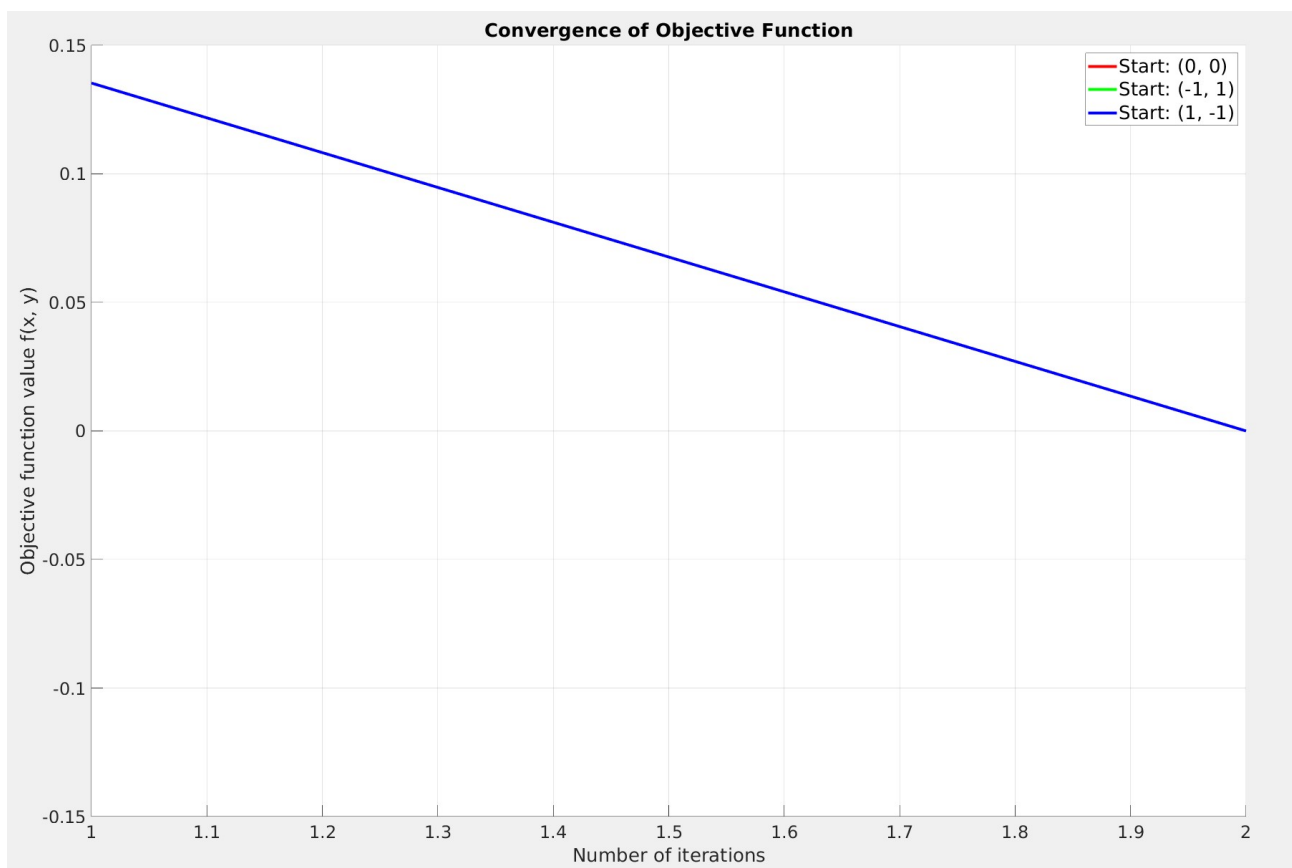
fx



Με βάσει αυτά, είναι προφανές ότι η μέθοδος **Newton για  $\mathbf{g}_k$  που ελαχιστοποιεί την  $f(\mathbf{x}_k + \mathbf{g}_k \mathbf{d}_k)$**  δεν δίνει σωστό αποτέλεσμα (-0.8111174) για κανένα σημείο έναρξης. Μάλιστα για σημείο έναρξης το (-1.0, 1.0), ο αλγόριθμος δεν μπορεί να εφαρμοστεί καθώς ο Εσσιανός πίνακας δεν είναι θετικά ορισμένος. Στις άλλες περιπτώσεις παρατηρούμε ότι “εγκλωβίζεται” σε τοπικό ελάχιστο και δεν μπορεί να διαφύγει.

Αναλυτικότερη επεξήγηση και κώδικας βρίσκεται στο αρχείο *work2\_ex3\_b.m*

γ) Επιλέγοντας το  $\mathbf{g}_k$  με βάσει τον κανόνα του Armijo: Εφαρμόζοντας τον κανόνα στην σελίδα 140 του βιβλίου προκύπτουν οι παρακάτω γραφικές παραστάσεις:



ενώ, για τα διάφορα σημεία έναρξης του αλγορίθμου ισχύει:

```
Initial point: (0.00, 0.00)
Minimum found at: (0.0000, 0.0000)
Final value f(x, y) = 0.000000
Number of iterations: 1

Warning: Hessian is not positive definite!
> In work2_ex3_c (line 66)
Initial point: (-1.00, 1.00)
Minimum found at: (-1.0000, 1.0000)
Final value f(x, y) = -0.135335
Number of iterations: 1

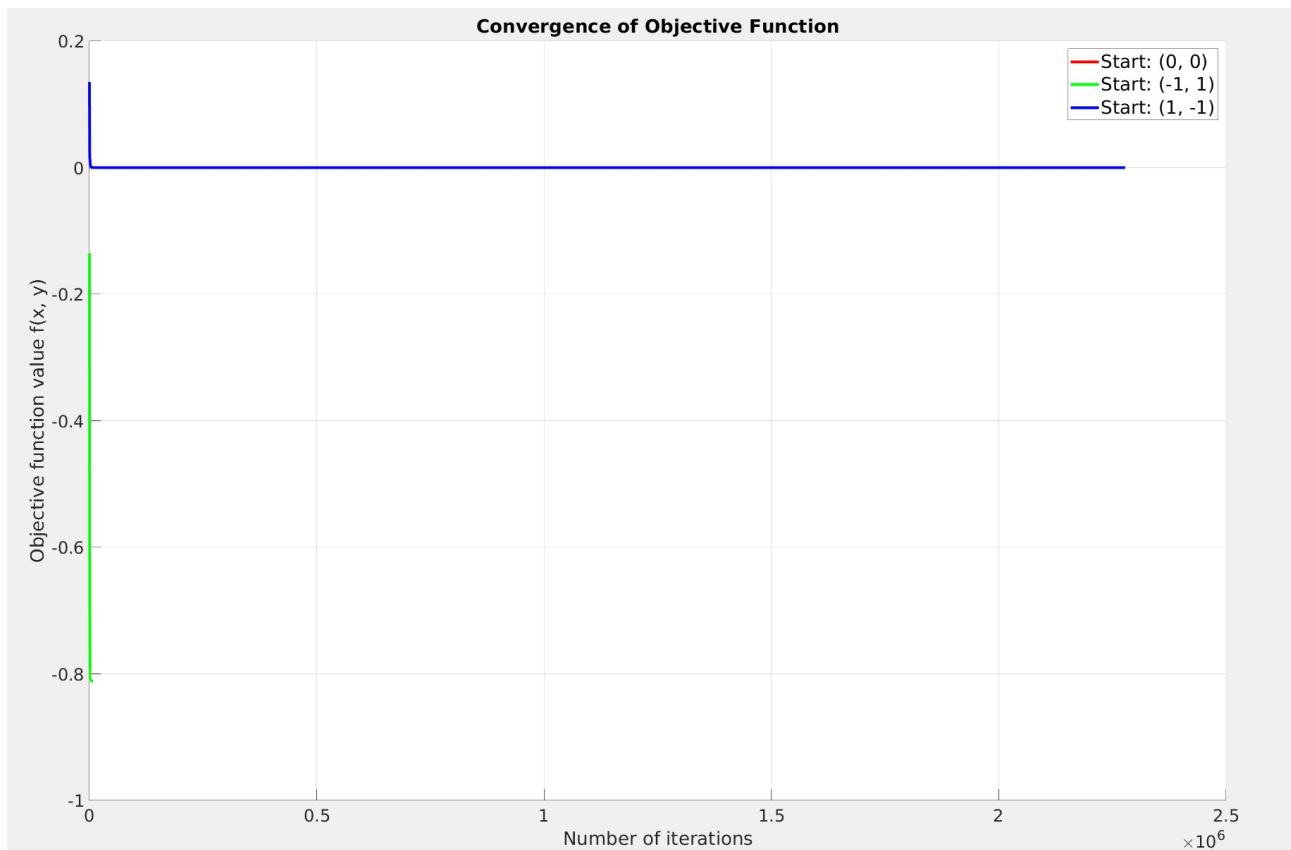
Initial point: (1.00, -1.00)
Minimum found at: (-0.7500, -5.5000)
Final value f(x, y) = -0.000000
Number of iterations: 2
```

Με βάσει αυτά, είναι προφανές ότι η μέθοδος **Newton** για  $\gamma_k$  με βάσει τον κανόνα **Armijo** δεν δίνει σωστό αποτέλεσμα (-0.8111174) σε καμία περίπτωση. Μάλιστα για σημείο έναρξης το (-1.0, 1.0), ο αλγόριθμος δεν μπορεί να εφαρμοστεί καθώς ο Εσσιανός πίνακας δεν είναι θετικά ορισμένος. Στις άλλες περιπτώσεις παρατηρούμε ότι “εγκλωβίζεται” σε τοπικό ελάχιστο και δεν μπορεί να διαφύγει.

*Αναλυτικότερη επεξήγηση και κώδικας βρίσκεται στο αρχείο work2\_ex3\_c.m*

#### **Θέμα 4:**

α) Για  $\gamma_k = 0.01$ , προκύπτει η εξής γραφική παράσταση της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων που απαιτούνται μέχρι να τερματίσει ο αλγόριθμος:



ενώ, για τα διάφορα σημεία έναρξης του αλγορίθμου ισχύει:

```
Initial point: (0.00, 0.00)
Minimum found at: (0.0000, 0.0000)
Final value  $f(x, y) = 0.000000$ 
Number of iterations: 1

Initial point: (-1.00, 1.00)
Minimum found at: (-1.5811, 0.0000)
Final value  $f(x, y) = -0.811174$ 
Number of iterations: 7746

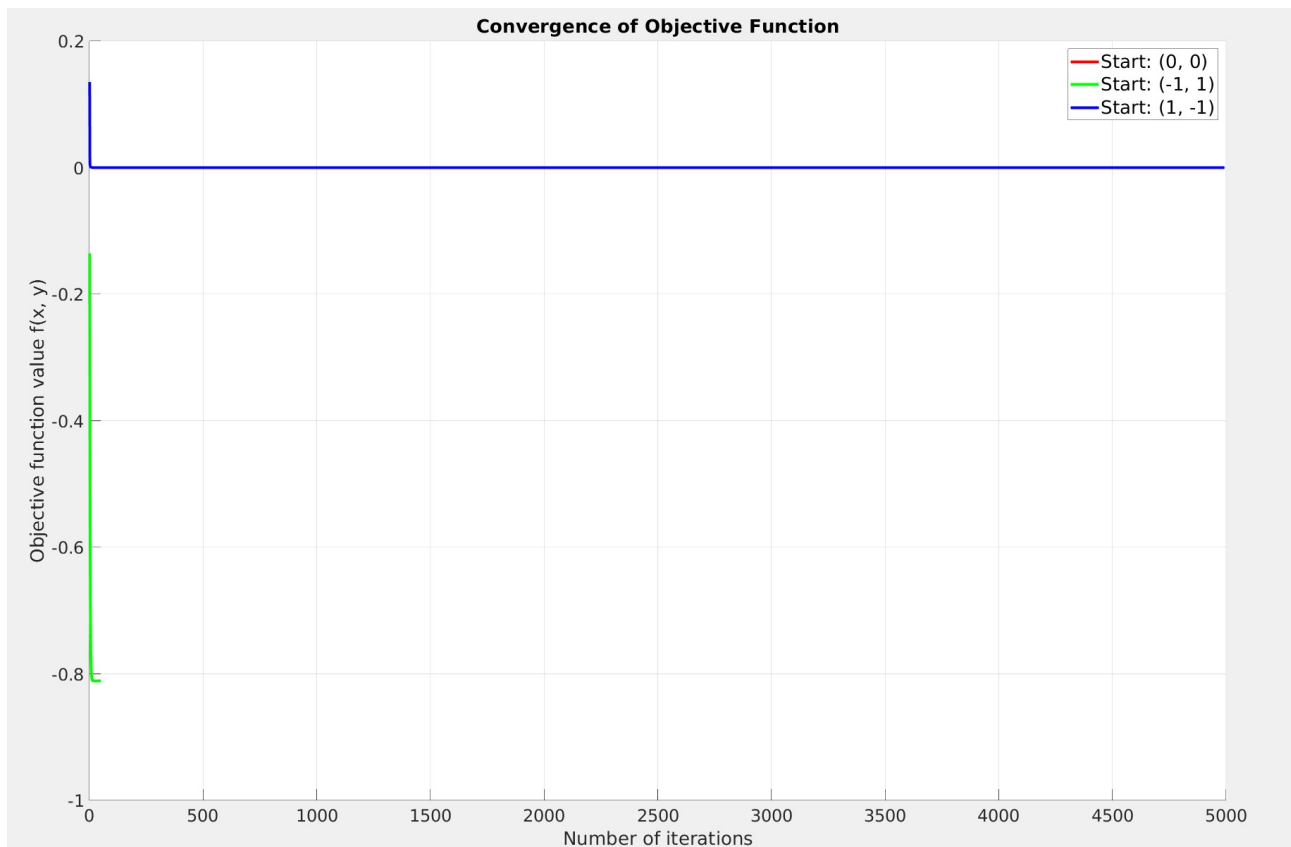
Initial point: (1.00, -1.00)
Minimum found at: (0.0341, -1.3811)
Final value  $f(x, y) = 0.000000$ 
Number of iterations: 2278176
```

Με βάσει αυτά είναι προφανές ότι η μέθοδος **Levenberg-Marquardt για σταθερό  $\gamma_k$**  έδωσε σωστό αποτέλεσμα (-0.8111174) μόνο όταν το αρχικό σημείο ήταν το (-1.0, 1.0). Στην περίπτωση που το σημείο έναρξης ήταν το (0.0, 0.0) παρατηρούμε ότι “εγκλωβίζεται” στο τοπικό ελάχιστο 0.0 και δεν μπορεί να διαφύγει, ενώ, αν είναι το (1.0, -1.0) παρατηρούμε ότι συμβαίνει το ίδιο.

Η μέθοδος Levenberg-Marquardt είναι ένας υβριδικός αλγόριθμος βελτιστοποίησης που συνδυάζει στοιχεία της μεθόδου Newton και της μεθόδου Μέγιστης Καθόδου για την επίλυση μη γραμμικών προβλημάτων ελαχίστων τετραγώνων. Η μέθοδος λειτουργεί προσαρμόζοντας δυναμικά τη στρατηγική της: όταν η τρέχουσα προσέγγιση είναι μακριά από τη λύση, εφαρμόζει μια πιο σταθερή αλλά αργή προσέγγιση, παρόμοια με τη Μέγιστη Κάθοδος. Αντίθετα, όταν βρίσκεται κοντά στη λύση, μεταβαίνει σε μια πιο ταχεία και ακριβή μέθοδο, παρόμοια με τη μέθοδο Newton. Αυτό επιτυγχάνεται μέσω ενός παραμέτρου απόσβεσης, που ελέγχει τη μετάβαση μεταξύ των δύο μεθόδων. Πλεονεκτήματα της μεθόδου περιλαμβάνουν την ταχεία σύγκλιση σε προβλήματα με καλά ορισμένες συναρτήσεις και την ικανότητά της να αποφεύγει προβλήματα που σχετίζονται με την αντιστροφή μη θετικά ορισμένων Εσσιανών πινάκων. Επίσης, είναι ιδιαίτερα αποτελεσματική για μη γραμμικές συναρτήσεις με μικρό αριθμό παραμέτρων. Από την άλλη, τα μειονεκτήματα της περιλαμβάνουν την ανάγκη προσεκτικής ρύθμισης των παραμέτρων απόσβεσης και το αυξημένο υπολογιστικό κόστος για μεγάλης κλίμακας προβλήματα, καθώς απαιτεί τον επαναληπτικό υπολογισμό της ιακωβιανής και των παραγώγων.

*Αναλυτικότερη επεξήγηση και κώδικας βρίσκεται στο αρχείο `work2_ex4_a.m`*

β) Επιλέγοντας  $\gamma_k$  τέτοιο ώστε να ελαχιστοποιεί την  $f(x_k + \gamma_k d_k)$ : Για τον σκοπό αυτόν, θα χρησιμοποιήσουμε την μέθοδο του **Χρυσού Τομέα**, γνωστή από την προηγούμενη εργασία. Έτσι, έχουμε τις εξής γραφικές παραστάσεις:



ενώ, για τα διάφορα σημεία έναρξης του αλγορίθμου ισχύει:

```
Initial point: (0.00, 0.00)
Minimum found at: (0.0000, 0.0000)
Final value  $f(x, y) = 0.000000$ 
Number of iterations: 1

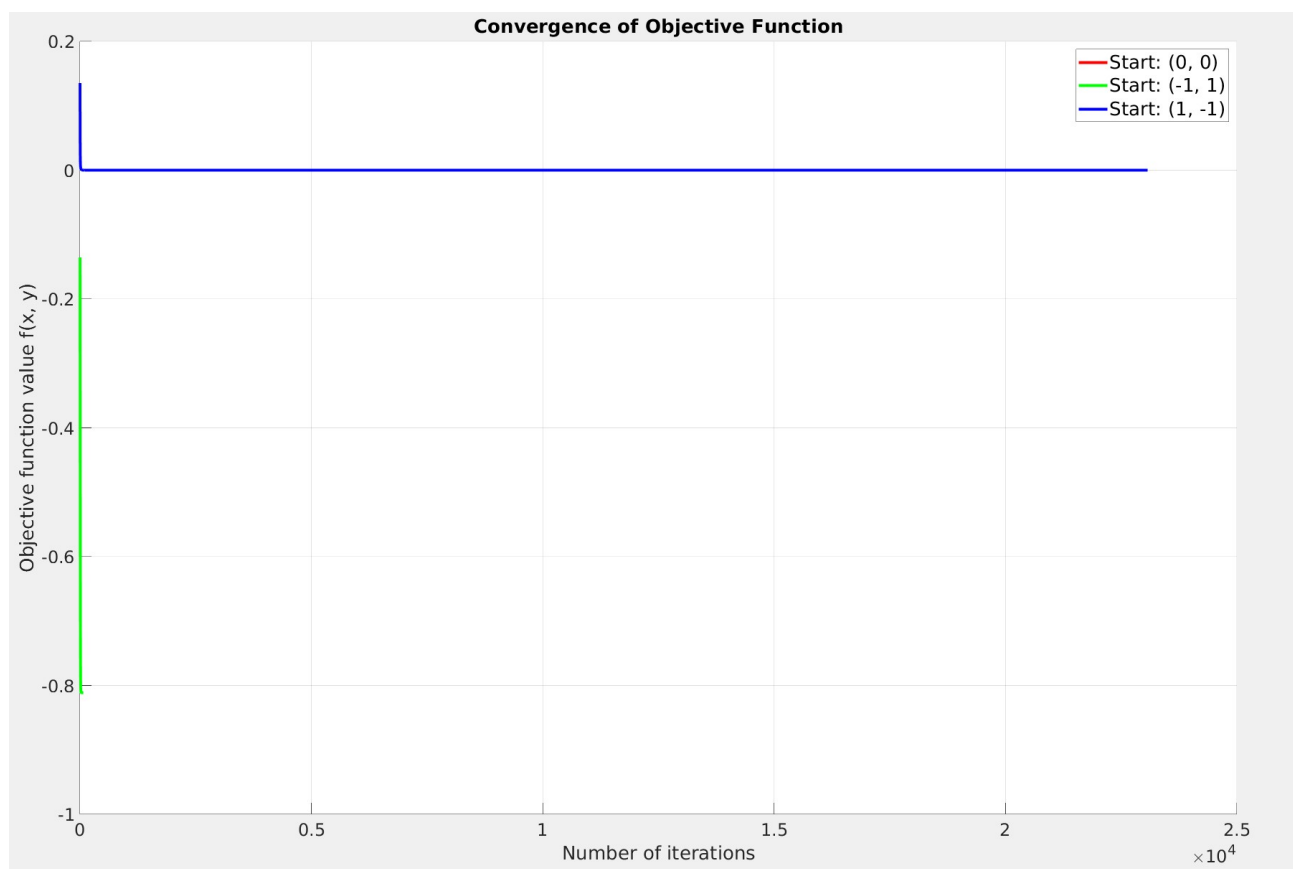
Initial point: (-1.00, 1.00)
Minimum found at: (-1.5811, 0.0000)
Final value  $f(x, y) = -0.811174$ 
Number of iterations: 50

Initial point: (1.00, -1.00)
Minimum found at: (0.0374, -1.5083)
Final value  $f(x, y) = 0.000000$ 
Number of iterations: 4993
```

Με βάσει αυτά είναι προφανές ότι η μέθοδος **Levenberg-Marquardt για  $\gamma_k$  που ελαχιστοποιεί την  $f(\mathbf{x}_k + \gamma_k \mathbf{d}_k)$**  κατάφερε να δώσει σωστό αποτέλεσμα (-0.8111174) μόνο όταν το αρχικό σημείο είναι το (-1.0, 1.0). Στην περίπτωση που το σημείο έναρξης ήταν το (0.0, 0.0) παρατηρούμε ότι “εγκλωβίζεται” σε τοπικό ελάχιστο και δεν μπορεί να διαφύγει, ενώ, αν είναι το (1.0, -1.0) παρατηρούμε ότι συμβαίνει το ίδιο.

Αναλυτικότερη επεξήγηση και κώδικας βρίσκεται στο αρχείο *work2\_ex4\_b.m*

γ) Επιλέγοντας το  $\gamma_k$  με βάσει τον κανόνα του Armijo: Εφαρμόζοντας τον κανόνα στην σελίδα 140 του βιβλίου προκύπτουν οι παρακάτω γραφικές παραστάσεις:



ενώ, για τα διάφορα σημεία έναρξης του αλγορίθμου ισχύει:

```
Initial point: (0.00, 0.00)
Minimum found at: (0.0000, 0.0000)
Final value f(x, y) = 0.000000
Number of iterations: 1

Initial point: (-1.00, 1.00)
Minimum found at: (-1.5811, 0.0000)
Final value f(x, y) = -0.811174
Number of iterations: 72

Initial point: (1.00, -1.00)
Minimum found at: (0.0345, -1.3995)
Final value f(x, y) = 0.000000
Number of iterations: 23070
```

Με βάσει αυτά, είναι προφανές ότι η μέθοδος **Levenberg-Marquardt** για  $\gamma_k$  με βάσει τον **κανόνα Armijo** δίνει σωστό αποτέλεσμα (-0.811174) μόνο όταν το αρχικό σημείο είναι το (-1.0, 1.0). Στις άλλες περιπτώσεις παρατηρούμε ότι “εγκλωβίζεται” σε τοπικό ελάχιστο και δεν μπορεί να διαφύγει.

*Αναλυτικότερη επεξήγηση και κώδικας βρίσκεται στο αρχείο `work2_ex4_c.m`*

*Από την ανάλυση όλων των γραφημάτων, φαίνεται ότι:*

1. Η επιλογή του αρχικού σημείου παίζει σημαντικό ρόλο στο αν τελικά οι μέθοδοι θα δώσουν τα επιθυμητά αποτελέσματα ή όχι.
2. Η μέθοδος Μέγιστης Καθόδου υπερέχει σε ταχύτητα σύγκλισης, ειδικά σε περιοχές κοντά στο ελάχιστο δεδομένου αυτής της ανάλυσης.