

2^η Εργασία Τεχνικών Βελτιστοποίησης

Αλεξόπουλος Δημήτριος AEM 10091

11 Δεκεμβρίου 2022

Εισαγωγή

Στην ακόλουθη εργασία ζητούμενο είναι η ελαχιστοποίηση μιας δοσμένης συνάρτησης πολλών μεταβλητών $f : R^n \rightarrow R$ χωρίς περιορισμούς. Οι αλγόριθμοι που θα χρησιμοποιήσουμε βασίζονται στην ιδέα της επαναληπτικής καθόδου, βάσει της οποίας ξεκινάμε από κάποιο σημείο $x_0 \in R^n$ και παράγουμε διαδοχικά τα διανύσματα x_1, x_2, \dots έτσι ώστε $f(x_{k+1}) < f(x_k), k = 1, 2, 3, \dots$. Οι αλγόριθμοι αναζήτησης που θα μελετήσουμε είναι:

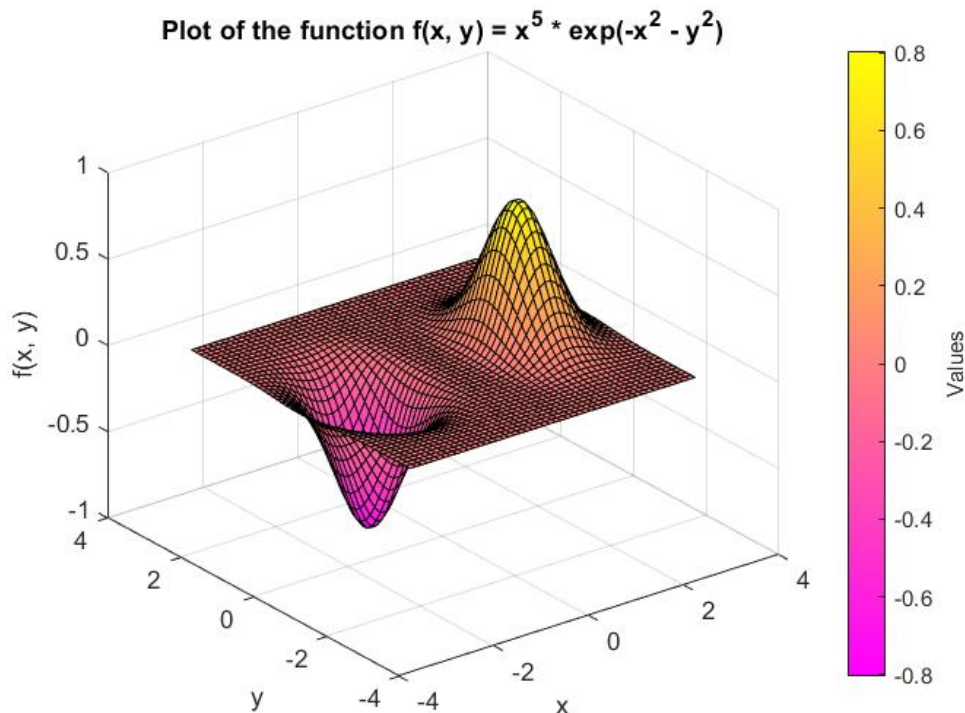
1. Μέθοδος Μέγιστης Καθόδου (Steepest Descent)
2. Μέθοδος Newton
3. Μέθοδος Levenberg-Marquardt

Η αντικειμενική συνάρτηση που θα μελετήσουμε είναι η:

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

1 Γραφική Παράσταση της Αντικειμενικής Συνάρτησης

Παρακάτω βλέπουμε την εικόνα της συνάρτησης στις τρεις διαστάσεις. Παρατηρούμε την μορφή της και τις θέσεις, διαισθητικά, των μεγίστων και ελαχίστων.



Σχήμα 1: Γραφική παράσταση της f

2 Μέθοδος Μέγιστης Καθόδου

Η μέθοδος της μέγιστης καθόδου (Steepest Descent) προκύπτει επιλέγοντας $\Delta_k = I, k = 1, 2, 3, \dots$ στην σχέση $x_{k+1} = x_k - \gamma_k \Delta_k \nabla f(x_k)$, όπου I ο μοναδιαίος $n \times n$ πίνακας.

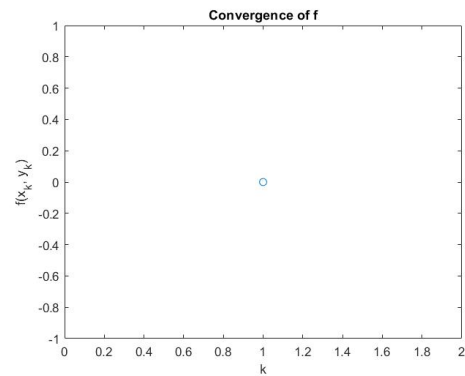
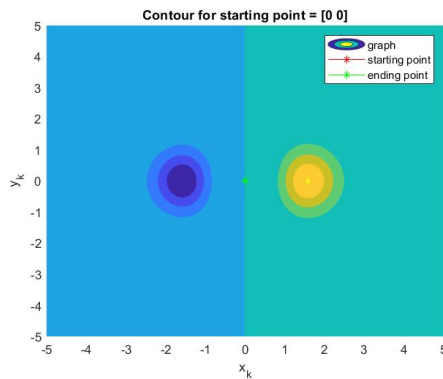
Χρησιμοποιήσαμε τρία διαφορετικά αρχικά σημεία για τους αλγορίθμους μας, ενώ για την επιλογή του βήματος γ_k χρησιμοποιήθηκαν τρεις μέθοδοι. Αυτοί παρατίθενται παρακάτω.

2.1 Σταθερό γ_k (της επιλογής μας)

Για κάθε ένα από τα τρία σημεία εκκίνησης του αλγορίθμου έχουμε:

2.1.1 Σημείο $(0, 0)$

Η τιμή της συνάρτησης παραμένει στο $(0, 0)$, καθώς το διάνυσμα κλίσης είναι ίσο με το μηδέν.

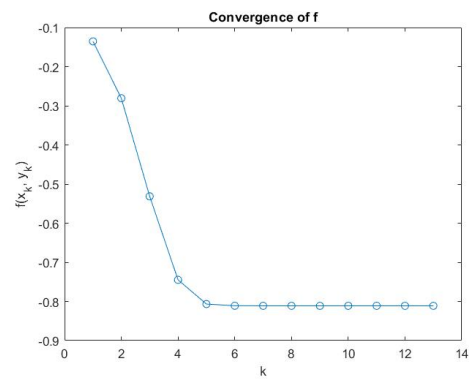
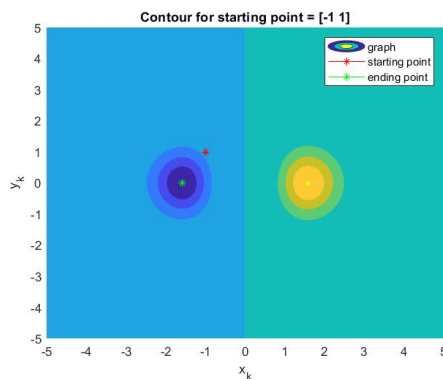


Σχήμα 2: Οπτικοποίηση της εύρεσης ελαχίστου για σταθερό γ_k

Σχήμα 3: Σύγκλιση της αντικειμενικής για σταθερό γ_k

2.1.2 Σημείο $(-1, 1)$

Η τιμή της συνάρτησης συγκλίνει στο ελάχιστο για $(x_k, y_k) = (-1.580981, 0.000000)$ με τιμή ελαχίστου -0.811174 . Ο αριθμός των επαναλήψεων k φαίνεται στο παρακάτω διάγραμμα.

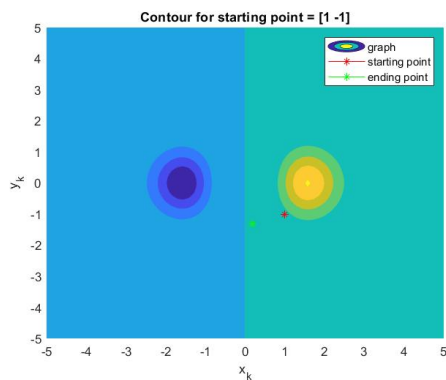


Σχήμα 4: Οπτικοποίηση της εύρεσης ελαχίστου για σταθερό γ_k

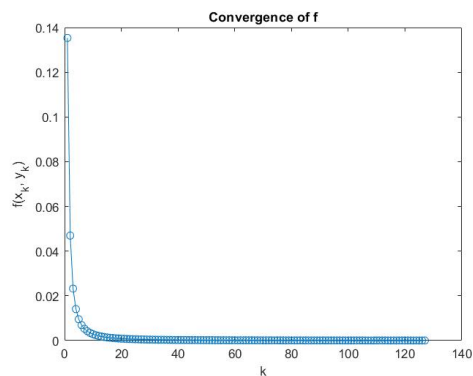
Σχήμα 5: Σύγκλιση της αντικειμενικής για σταθερό γ_k

2.1.3 Σημείο $(1, -1)$

Η τιμή της συνάρτησης εγκλωβίζεται στο τοπικό ελάχιστο για $(x_k, y_k) = (0.184688, -1.314906)$ με τιμή 0.000037 . Ο αριθμός των επαναλήψεων k φαίνεται στο παρακάτω διάγραμμα.



Σχήμα 6: Οπτικοποίηση της εύρεσης ελαχίστου για σταθερό γ_k



Σχήμα 7: Σύγκλιση της αντικειμενικής για σταθερό γ_k

2.2 Τέτοιο γ_k ώστε να ελαχιστοποιεί την $f(x_k + \gamma_k d_k)$

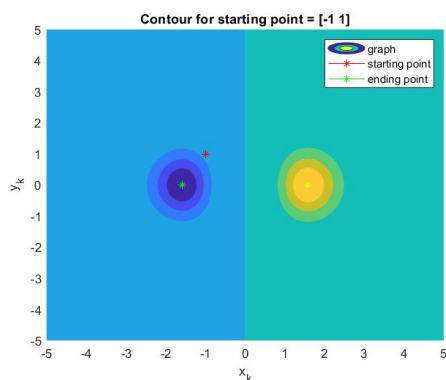
Για κάθε ένα από τα τρία σημεία εκκίνησης του αλγορίθμου έχουμε:

2.2.1 Σημείο (0,0)

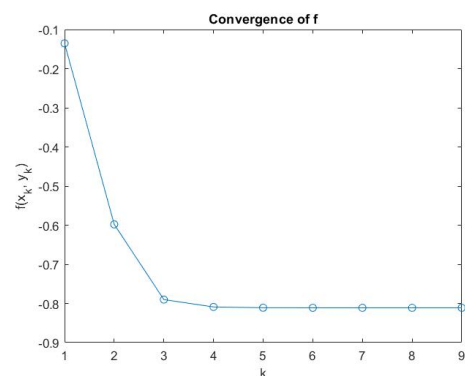
Ακριβώς όπως και πριν, η τιμή της συνάρτησης παραμένει στο (0,0), καθώς το διάνυσμα κλίσης είναι ίσο με το μηδέν.

2.2.2 Σημείο (-1,1)

Η τιμή της συνάρτησης συγκλίνει στο ελάχιστο για $(x_k, y_k) = (-1.581062, 0.000102)$ με τιμή ελαχίστου -0.811174 . Ο αριθμός των επαναλήψεων k φαίνεται στο παρακάτω διάγραμμα.



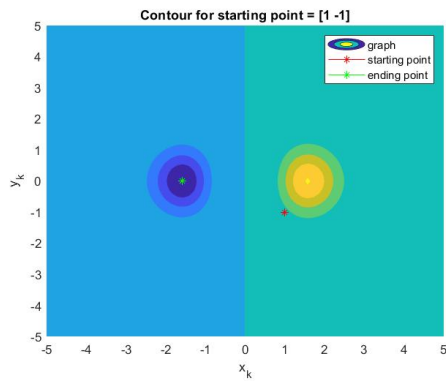
Σχήμα 8: Οπτικοποίηση της εύρεσης ελαχίστου για τέτοιο γ_k ώστε να ελαχιστοποιεί την $f(x_k + \gamma_k d_k)$



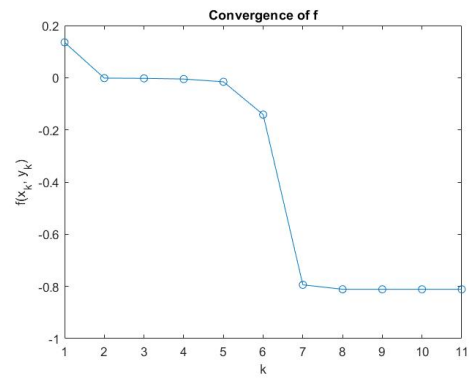
Σχήμα 9: Σύγκλιση της αντικειμενικής για τέτοιο γ_k ώστε να ελαχιστοποιεί την $f(x_k + \gamma_k d_k)$

2.2.3 Σημείο (1,-1)

Η τιμή της συνάρτησης συγκλίνει στο ελάχιστο για $(x_k, y_k) = (-1.581152, -0.000004)$ με τιμή ελαχίστου -0.811174 . Ο αριθμός των επαναλήψεων k φαίνεται στο παρακάτω διάγραμμα.



Σχήμα 10: Οπτικοποίηση της εύρεσης ελαχίστου για τέτοιο γ_k ώστε να ελαχιστοποιεί την $f(x_k + \gamma_k d_k)$



Σχήμα 11: Σύγκλιση της αντικειμενικής για τέτοιο γ_k ώστε να ελαχιστοποιεί την $f(x_k + \gamma_k d_k)$

2.3 Το γ_k βάσει του κανόνα Armijo

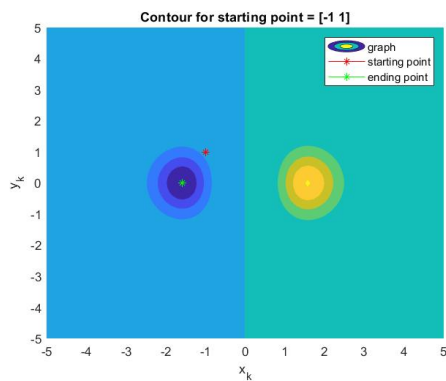
Για κάθε ένα από τα τρία σημεία εκκίνησης του αλγορίθμου έχουμε:

2.3.1 Σημείο (0,0)

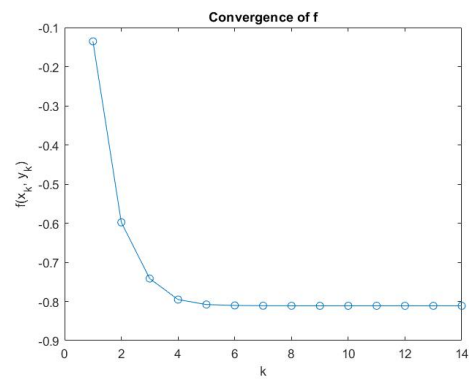
Ακριβώς όπως και πριν, η τιμή της συνάρτησης παραμένει στο (0,0), καθώς το διάνυσμα κλίσης είναι ίσο με το μηδέν.

2.3.2 Σημείο (-1,1)

Η τιμή της συνάρτησης συγκλίνει στο ελάχιστο για $(x_k, y_k) = (-1.581266, -0.000000)$ με τιμή ελαχίστου -0.811174 . Ο αριθμός των επαναλήψεων k φαίνεται στο παρακάτω διάγραμμα.



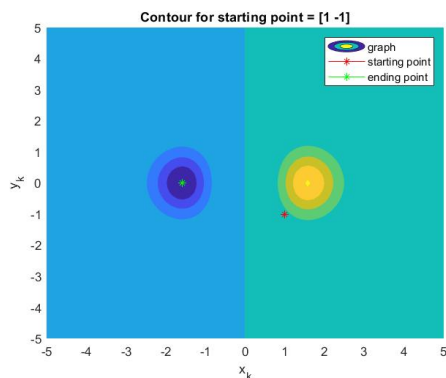
Σχήμα 12: Οπτικοποίηση της εύρεσης ελαχίστου για γ_k βάσει του κανόνα Armijo



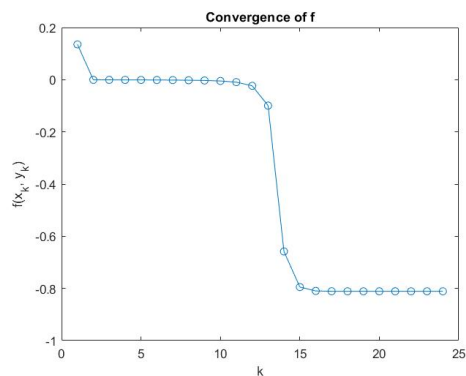
Σχήμα 13: Σύγκλιση της αντικειμενικής για γ_k βάσει του κανόνα Armijo

2.3.3 Σημείο (1,-1)

Η τιμή της συνάρτησης συγκλίνει στο ελάχιστο για $(x_k, y_k) = (-1.581231, -0.000000)$ με τιμή ελαχίστου -0.811174 . Ο αριθμός των επαναλήψεων k φαίνεται στο παρακάτω διάγραμμα.



Σχήμα 14: Οπτικοποίηση της εύρεσης ελαχίστου για γ_k βάσει του κανόνα Armijo



Σχήμα 15: Σύγκλιση της αντικειμενικής για γ_k βάσει του κανόνα Armijo

3 Μέθοδος Newton

Η Μέθοδος Newton προκύπτει εάν αυτή τη φορά επιλέξουμε $\Delta_k = \text{Hess}(f(x_k))^{-1}$, $k = 0, 1, 2, \dots$ με την προϋπόθεση ότι ο εσσιανός πίνακας της αντικειμενικής συνάρτησης είναι **θετικά ορισμένος**.

Παρατηρούμε, ωστόσο, και επιβεβαιώνουμε προγραμματιστικά ότι ο εσσιανός πίνακας στην συγκεκριμένη περίπτωση και για κάθε σημείο εκκίνησης **ΔΕΝ** είναι θετικά ορισμένος. Επομένως, δεν μπορεί να χρησιμοποιηθεί αυτή η μέθοδος για την εύρεση ελαχίστου και γι' αυτό θα χρησιμοποιήσουμε τον τροποποιημένο αλγόριθμο Newton ή, όπως αλλιώς λέγεται, την μέθοδο Levenberg-Marquardt.

4 Μέθοδος Levenberg-Marquardt

Η μέθοδος αυτή εξασφαλίζει ότι θα έχουμε έναν πίνακα $A = \text{Hess}(f(x_k)) + \mu_k I$, ο οποίος θα είναι πάντα **θετικά ορισμένος**.

Χρησιμοποιήσαμε και πάλι τρία διαφορετικά αρχικά σημεία, ενώ για την επιλογή του βήματος γ_k χρησιμοποιήθηκαν οι δυο μέθοδοι μεταβαλλόμενου γ_k , καθώς αυτή η μέθοδος δεν επιτρέπει σταθερο γ_k . Αυτά παρατίθενται παρακάτω.

4.1 Τέτοιο γ_k ώστε να ελαχιστοποιεί την $f(x_k + \gamma_k d_k)$

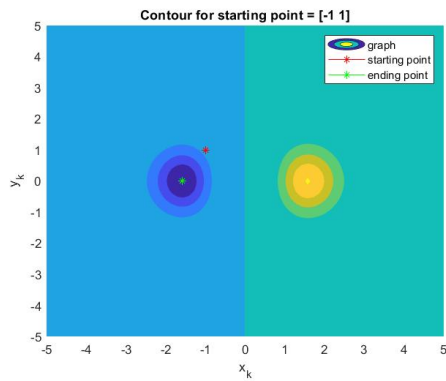
Για κάθε ένα από τα τρία σημεία εκκίνησης του αλγορίθμου έχουμε:

4.1.1 Σημείο (0,0)

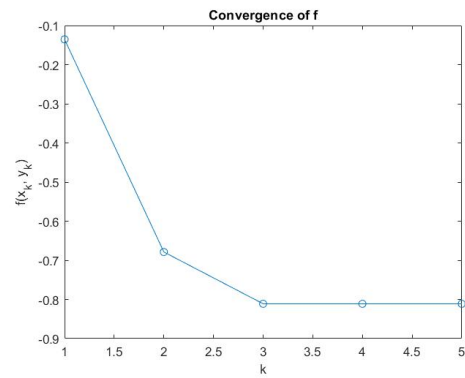
Ακριβώς όπως και πριν, η τιμή της συνάρτησης παραμένει στο (0,0), καθώς το διάνυσμα κλίσης είναι ίσο με το μηδέν.

4.1.2 Σημείο (-1,1)

Η τιμή της συνάρτησης συγκλίνει στο ελάχιστο για $(x_k, y_k) = (-1.581139, -0.000000)$ με τιμή ελαχίστου -0.811174 . Ο αριθμός των επαναλήψεων k φαίνεται στο παρακάτω διάγραμμα.



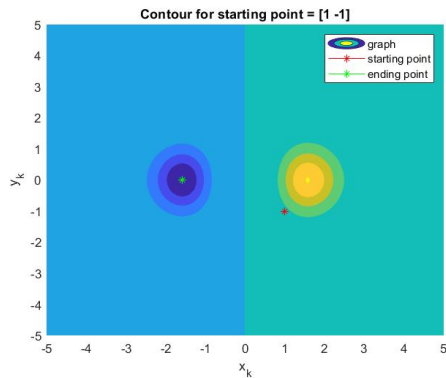
Σχήμα 16: Οπτικοποίηση της εύρεσης ελαχίστου για τέτοιο γ_k ώστε να ελαχιστοποιεί την $f(x_k + \gamma_k d_k)$



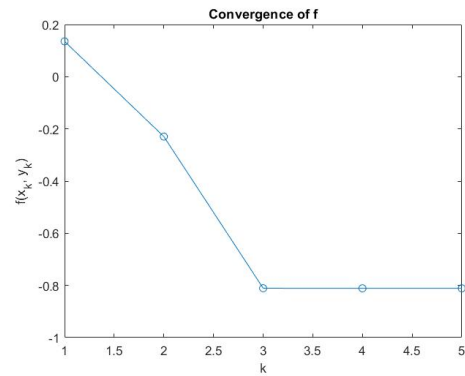
Σχήμα 17: Σύγκλιση της αντικειμενικής για τέτοιο γ_k ώστε να ελαχιστοποιεί την $f(x_k + \gamma_k d_k)$

4.1.3 Σημείο (1, -1)

Η τιμή της συνάρτησης συγκλίνει στο ελάχιστο για $(x_k, y_k) = (-1.581139, -0.000000)$ με τιμή ελαχίστου -0.811174 . Ο αριθμός των επαναλήψεων k φαίνεται στο παρακάτω διάγραμμα.



Σχήμα 18: Οπτικοποίηση της εύρεσης ελαχίστου για τέτοιο γ_k ώστε να ελαχιστοποιεί την $f(x_k + \gamma_k d_k)$



Σχήμα 19: Σύγκλιση της αντικειμενικής για τέτοιο γ_k ώστε να ελαχιστοποιεί την $f(x_k + \gamma_k d_k)$

4.2 Το γ_k βάσει του κανόνα Armijo

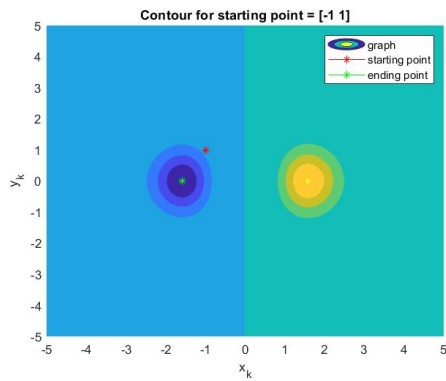
Για κάθε ένα από τα τρία σημεία εκκίνησης του αλγορίθμου έχουμε:

4.2.1 Σημείο (0, 0)

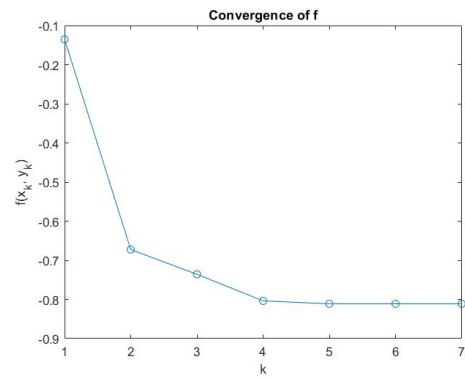
Ακριβώς όπως και πριν, η τιμή της συνάρτησης παραμένει στο $(0, 0)$, καθώς το διάνυσμα κλίσης είναι ίσο με το μηδέν.

4.2.2 Σημείο (-1, 1)

Η τιμή της συνάρτησης συγκλίνει στο ελάχιστο για $(x_k, y_k) = (-1.581062, -0.000142)$ με τιμή ελαχίστου -0.811174 . Ο αριθμός των επαναλήψεων k φαίνεται στο παρακάτω διάγραμμα.



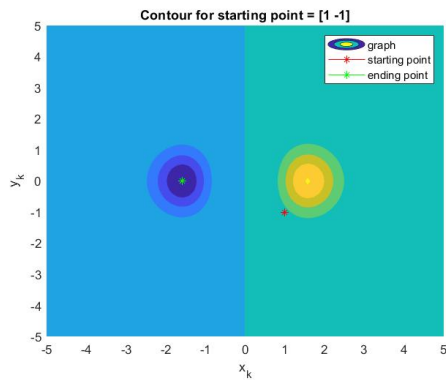
Σχήμα 20: Οπτικοποίηση της εύρεσης ελαχίστου για γ_k βάσει του κανόνα Armijo



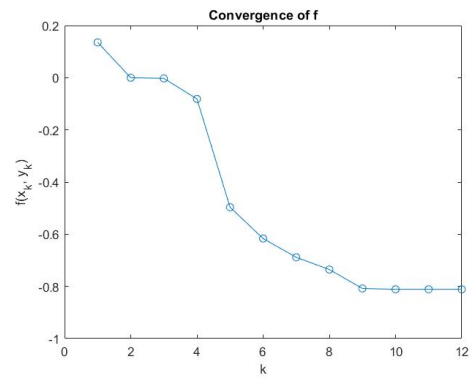
Σχήμα 21: Σύγκλιση της αντικειμενικής για γ_k βάσει του κανόνα Armijo

4.2.3 Σημείο (1, -1)

Η τιμή της συνάρτησης συγκλίνει στο ελάχιστο για $(x_k, y_k) = (-1.581232, 0.000182)$ με τιμή ελαχίστου -0.811174 . Ο αριθμός των επαναλήψεων k φαίνεται στο παρακάτω διάγραμμα.



Σχήμα 22: Οπτικοποίηση της εύρεσης ελαχίστου για γ_k βάσει του κανόνα Armijo



Σχήμα 23: Σύγκλιση της αντικειμενικής για γ_k βάσει του κανόνα Armijo

5 Σύγκριση Μεθόδων - Σχόλια

Από την παραπάνω ανάλυση μπορούμε να βγάλουμε συμπεράσματα για την αποδοτικότητα των αλγορίθμων και να συγκρίνουμε τις επιμέρους μεθόδους, τις επιλογές του γ_k , αλλά και τις επιλογές των αρχικών σημείων αναζήτησης.

Αρχικά, είναι προφανής η σημασία του σημείου εκκίνησης, καθώς υπάρχει περίπτωση ο αλγόριθμος να εγκλωβιστεί σε αυτό ή και να μην ξεκινήσει καν - όπως έγινε με το σημείο $(0, 0)$ - λόγω μηδενικού διανύσματος κλίσης.

Από την άλλη, παρατηρώντας τον παρακάτω πίνακα με την σύγκριση του αριθμού των επαναλήψεων για κάθε μέθοδο και κάθε επιλογή γ_k για το σημείο εκκίνησης $(-1, 1)$:

Επιλογή γ_k	Σταθερό	Ελαχ. την $f(x_k + \gamma_k d_k)$	Κανόνας Armijo
Μέθοδος Μέγιστης Καθόδου	13	9	14
Μέθοδος Levenberg-Marquardt	-	5	7

Πίνακας 1: Σύγκριση επαναλήψεων μεθόδων για το σημείο εκκίνησης $(-1, 1)$.

και για το σημείο εκκίνησης $(1, -1)$:

Επιλογή γ_k	Σταθερό	Ελαχ. την $f(x_k + \gamma_k d_k)$	Κανόνας Armijo
Μέθοδος Μέγιστης Καθόδου	127	11	24
Μέθοδος Levenberg-Marquardt	-	5	12

Πίνακας 2: Σύγκριση επαναλήψεων μεθόδων για το σημείο εκκίνησης $(1, -1)$.

συμπεραίνουμε πως η αποδοτικότερη επιλογή του γ_k είναι αυτή που ελαχιστοποιεί την $f(x_k + \gamma_k d_k)$, ενώ ο αποδοτικότερος αλγόριθμος φαίνεται να είναι η μέθοδος Levenberg-Marquardt, στην οποία το ελάχιστο εντοπίζεται σωστά για τα δύο σημεία εκκίνησης σε πολύ μικρό αριθμό επαναλήψεων.

Παρ' όλ' αυτά, και οι άλλες δύο επιλογές του γ_k είναι πολύ αποδοτικές, όπως και η μέθοδος μέγιστης καθόδου είναι πολύ γρήγορη. Μοναδική εξαίρεση αποτελεί η σταθερή επιλογή του γ_k , η οποία απαίτησε παραπάνω επαναλήψεις και στην περίπτωση του σημείου εκκίνησης $(1, -1)$ δεν εντόπισε το μικρότερο δυνατό ελάχιστο, αλλά εγκλωβίστηκε σε ένα τοπικό ελάχιστο.