Τεχνικές Βελτιστοποίησης

Νικόλαος Τουλκερίδης

AEM: 10718

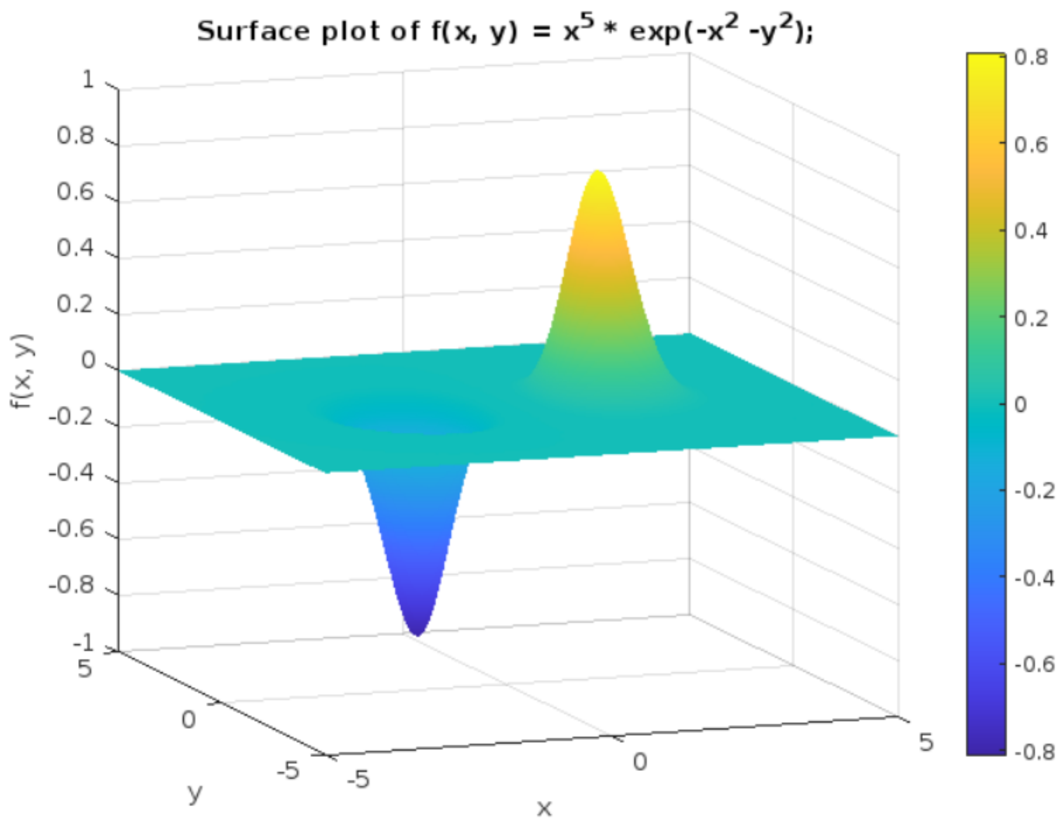
1 Εισαγωγή

Η εργασία αυτή εξετάζει τρεις αλγόριθμους αναζήτησης οι οποίοι ελαχιστοποιούν μια συνάρτηση πολλών μεταβλητών χωρίς περιορισμούς. Οι αλγόριθμοι που θα μελετηθούν είναι:
1. Η μέθοδος Μέγιστης καθόδου (Steepest Descent)
2. Η μέθοδος Newton
3. Η μέθοδος Levenberg-Marquardt

Η υπό μελέτη συνάρτηση είναι η f(x,y)=$x^5 * exp(-x^2 - y^2)$.

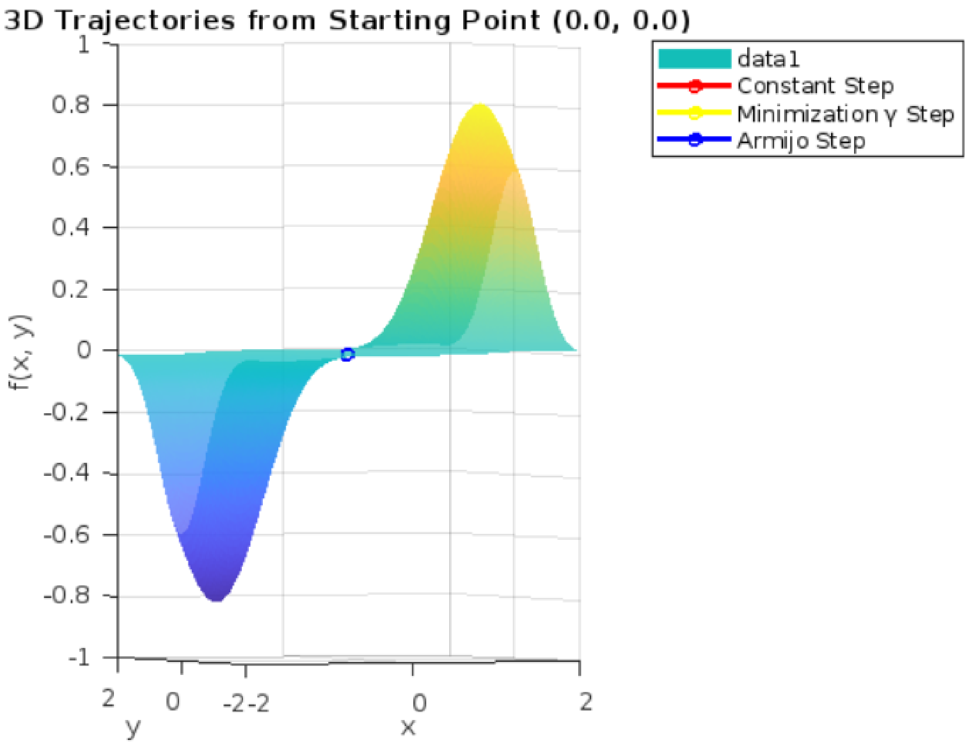
2 Θέμα 1

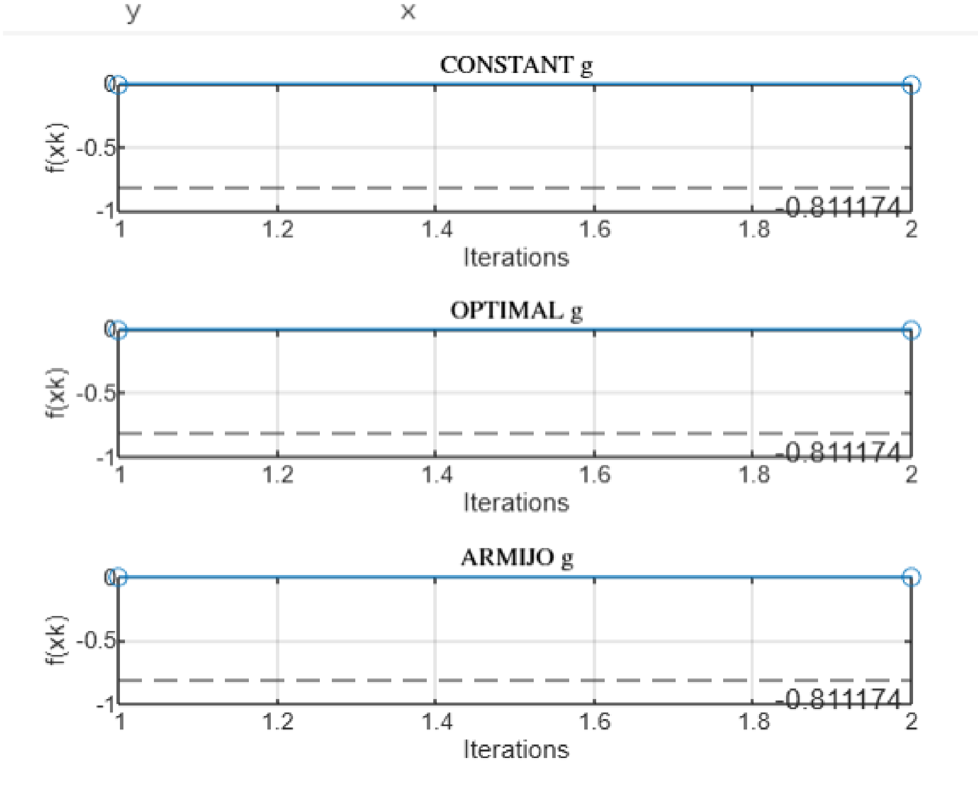
Εδώ παρουσιάζεται μία γενική εικόνα της μορφής της συνάρτησης.



3 Θέμα 2

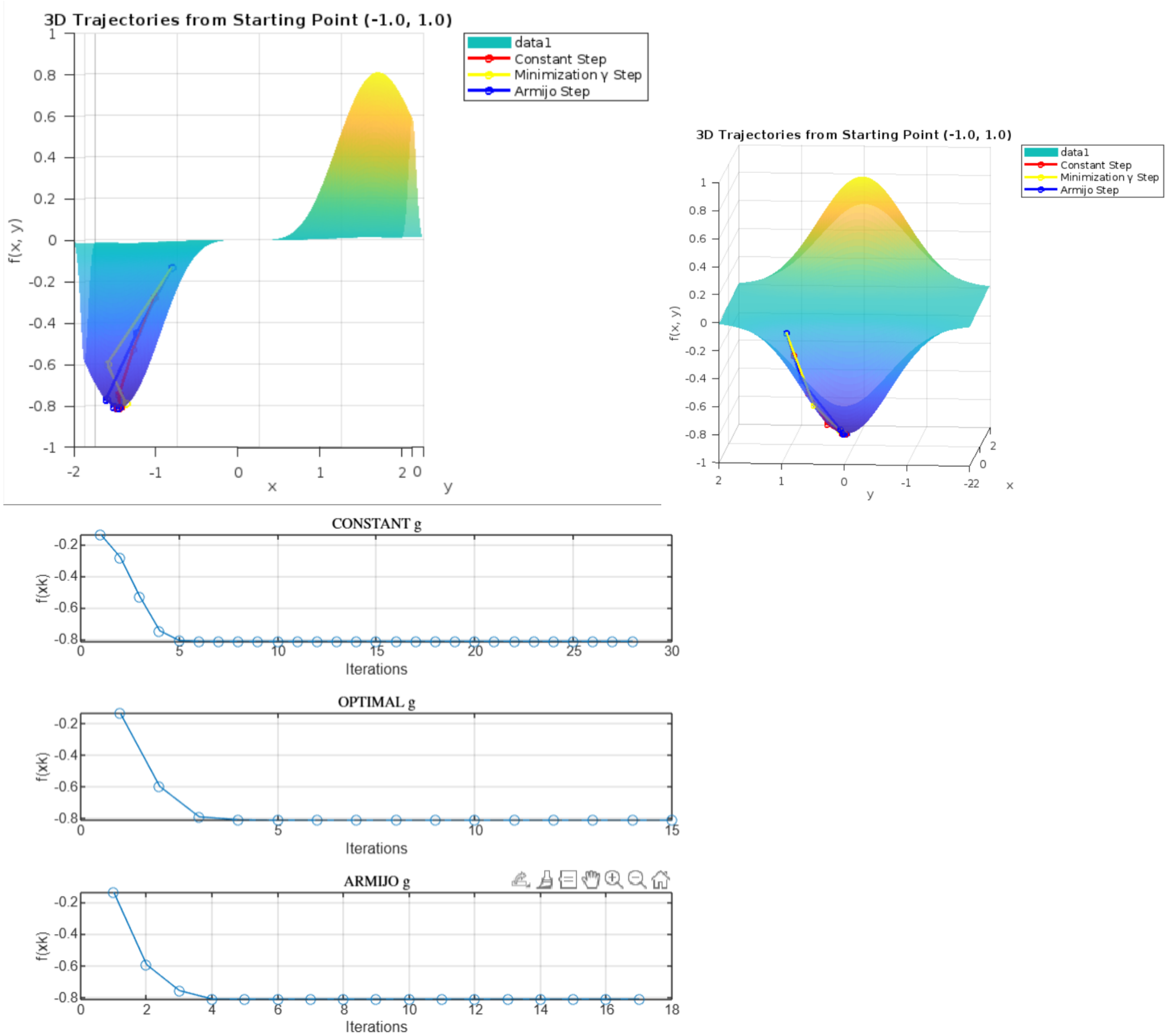
3.1 Αρχικό σημείο: (0,0)





Όπως φαίνεται κανένας αλγόριθμος δε συγκλίνει σε ελάχιστο διότι στο αρχικό σημείο υπάρχει τοπικό ελάχιστο και κατά συνέπεια το $\text{grad}(f) = 0$. Αυτό δεν επιτρέπει τον αλγόριθμο να ¨κινηθεί' προς καμία κατεύθυνση κι έτσι δεν υπάρχει πρόοδος. Αυτή η περίπτωση είναι ένα κλασικό παράδειγμα της εξάρτησης του αποτελέσματος από την τιμή εκκίνησης του αλγορίθμου.

3.2 Αρχικό σημείο: (-1,1)



Για αυτό το σημείο εκκίνησης συγκλίνουν και οι τρεις αλγόριθμοι στο ελάχιστο της συνάρτησης, ωστόσο δεν έχουν την ίδια αποδοτικότητα.

Περίπτωση α)

Για σταθερό βήμα γκ ο αλγόριθμος συγκλίνει σε 27 επαναλήψεις. Αυτό συμβαίνει γιατί δεν λαμβάνει υπ΄όψη του χαρακτηριστικά της συνάρτησης, όπως το πόσο απότομη είναι σε ορισμένα σημεία. Έτσι, κάνει κάθε φορά το ίδιο, σταθερό βήμα προς το ελάχιστο ¨χάνοντας’ ίσως ευκαιρίες να ¨κατάβει ταχύτερα την πλαγιά¨.

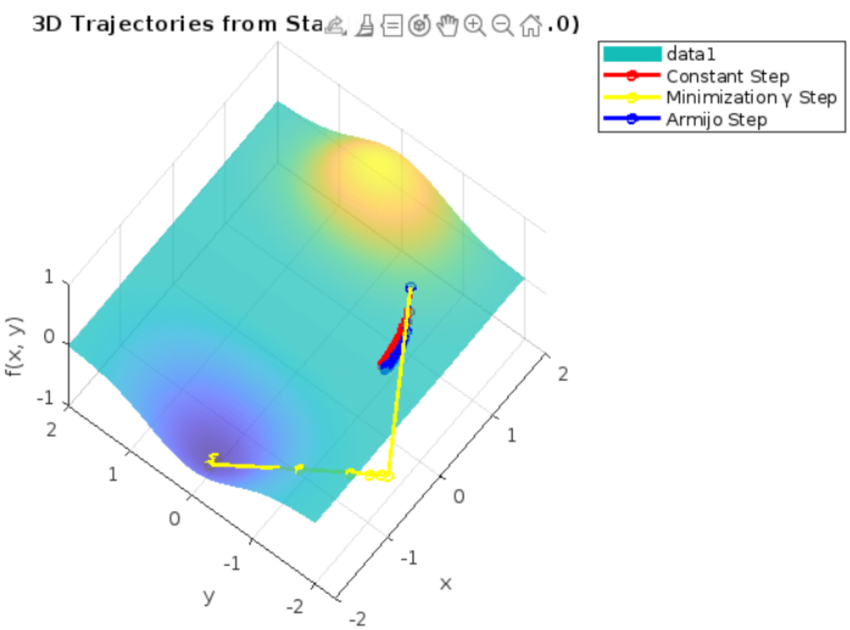
Περίπτωση β)

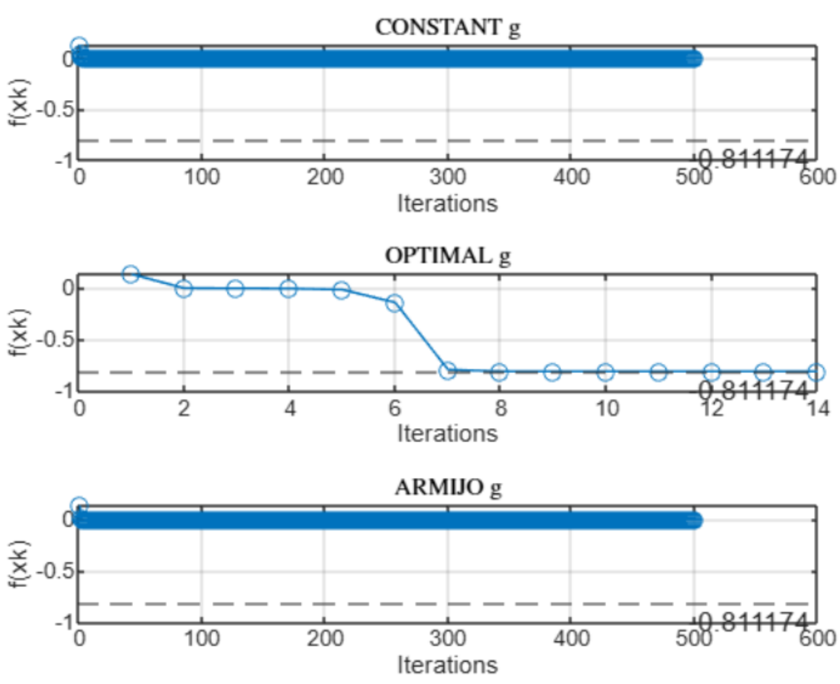
Εδώ το βήμα γκ υπολογίζεται τέτοιο ώστε να ελαχιστοποιεί την f(xk + gk*dk). Ο αλγόριθμος συγκλίνει σε 14 επαναλήψεις. Αυτό συμβαίνει επειδή το βήμα γκ προσαρμόζεται δυναμικά έτσι ώστε σε κάθε επανάληψη το βήμα να είναι κατάλληλο έτσι ώστε να ¨κατέβουμε ταχύτερα την πλαγιά¨, δηλαδή να κάνει μεγαλύτερα βήματα όταν η παράγωγος δείχνει καθαρά την κατεύθυνση, είτε να ¨κόβουμε την ταχύτητα’ όταν φθάνουμε κοντά στο ελάχιστο, αποφεύγοντας και το γεγονός να το ξεπεράσουμε. Κάτι τέτοιο δε συμβαίνει στην παραπάνω περίπτωση που το βημα γκ είναι σταθερό σε κάθε επανάληψη.

Περίπτωση γ)

Σε αυτή την περίπτωση χρησιμοποιείται ο κανόνας Armijo για την επιλογή του βήματος γ. Η λογική αυτού είναι να βρεθεί το κατάλληλο βήμα έτσι ώστε μην είναι πολύ μεγάλο και φύγουμε πολύ μακρυά αλλά ούτε και πολύ μικρό αυξάνοντας κατά πολύ την σύγκλιση. Η μέθοδος έχει ως προτεραιότητα την σταθερότητα από την πολύ γρήγορη σύγκλιση, κάτι που αποδεικνύεται πολύ χρήσιμο σε υψηλότερης διάστασης προβλήματα. Εάν το μονοδιάστατο πρόβλημα βελτιστοποίησης που καλείται να λύσει η παραπάνω περίπτωση λύνεται εύκολα τότε το να υπολογίζουμε σε κάθε επανάληψη το βέλτιστο βήμα γ θα μας δώσει καλύτερη σύγκλιση όπως φαίνεται να γίνεται και εδώ, καθώς ο αλγόριθμος Armijo χρειάζεται 33 επαναλήψεις να συγκλίνει, κάτι εμφανώς παραπάνω από ότι θα περιμέναμε. Συνοπτικά αυτή η μέθοδος βεβαιώνει ένα ¨αρκετά καλό’ βήμα που συχνά είναι χειρότερο από την πιο ακριβή ελαχιστοποιήση που επιχειρεί η παραπάνω.

3.3 Αρχικό σημείο: (1,-1)



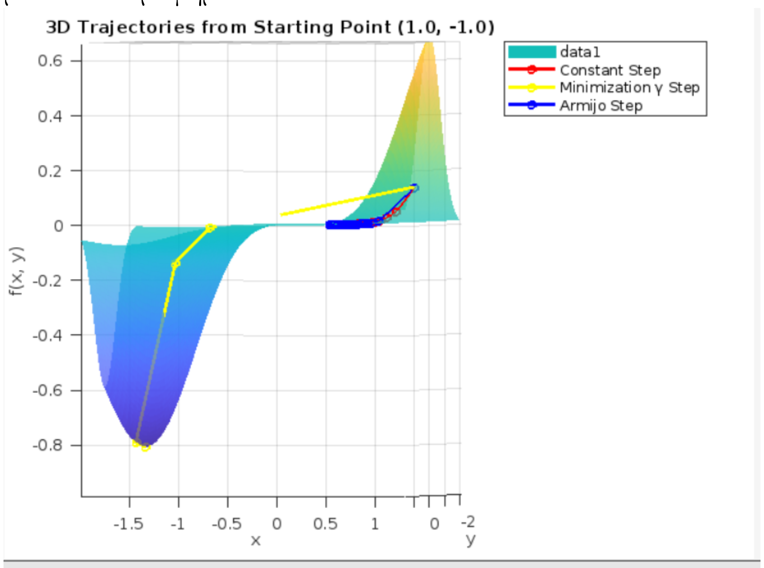


Περίπτωση α)

Για αυτό το αρχικό σημείο η μέθοδος με σταθερό βήμα γ δεν καταφέρνει να συγκλίνει. Φαίνεται να ¨παγιδεύεται’ γύρω από ενα τοπικό ελάχιστο κοντά στο $\chi = 0$ και παραμένει εκεί. Ακόμη μία φορά φαίνεται η εξάρτηση του αποτελέσματος από το σημείο εκκίνησης και το πρόβλημα του σταθερού βήματος γ.

Περίπτωση β)

Αυτή η μέθοδος επιφέρει πολύ καλά αποτελέσματα. Καταφέρνει να ξεφύγει από το τοπικό ελάχιστο που παγιδεύεται η μέθοδος με το σταθερό βήμα.



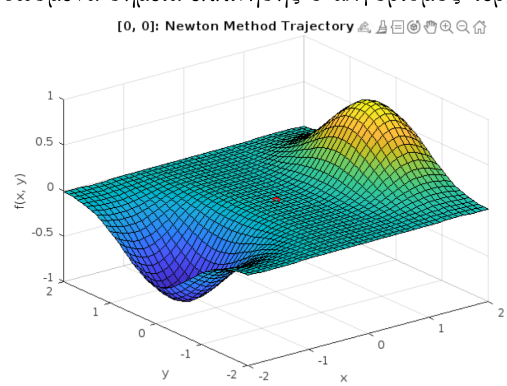
Όπως φαίνεται από το σχήμα ο αλγόριθμος κάνει ένα πολύ μεγάλο βήμα στην 1η επανάληψη εξαιτίας της κλίσης στης συνάρτησης στο αρχικό σημείο.

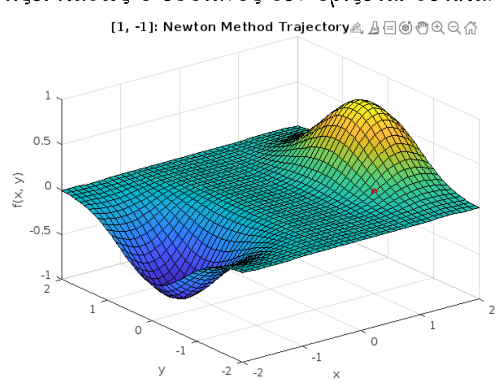
Περίπτωση γ)

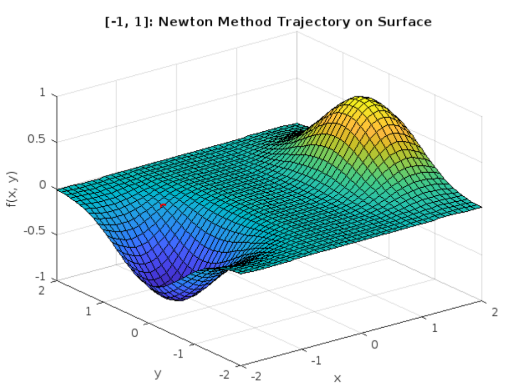
Η μέθοδος που χρησιμοποιεί τον κανόνα Armijo δεν συγκλίνει γιατί εγκλωβίζεται και αυτή στο τοπικό ελάχιστο που παγιδεύεται η μέθοδος σταθερού βήματος. Ακριβώς επειδή αυτή η μέθοδος λαμβάνει πιο επιφυλακτικές αποφάσεις όσον αφορά το βήμα γ δεν κάνει αρκετά μεγάλο βήμα ώστε να ξεπεράσει το τοπικό ελάχιστο και καταλήγει να εγκλωβίζεται σε αυτό. Φαίνεται ξανά η αναποτελεσματικότητα αυτής της μεθόδου για το συγκεκριμένο πρόβλημα.

4 Θέμα 3

Η μέθοδος Newton στηρίζεται στην ύπαρξη του εσσιανού πίνακα ο οποίος πρέπει να είναι θετικά ορισμένος. Ωστόσο για τα δωσμένα σημεία εκκίνησης ο αλγόριθμος τερματίζει καθώς ο εσσιανός δεν ορίζεται θετικά.





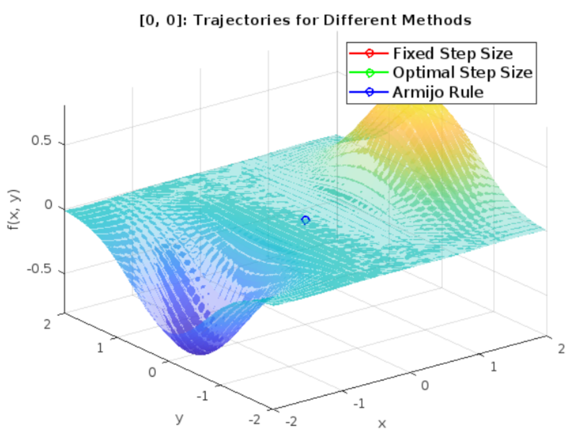


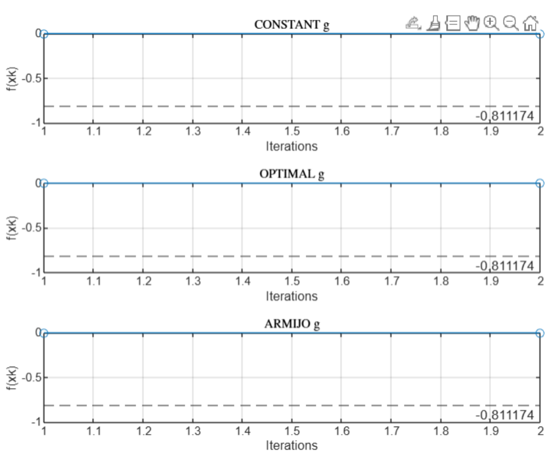

```
Converged to minimum at iteration 0.
Minimum is: [0.000000, 0.000000] = 0.000000
Eigen values of hess are: [-0.5413, 1.0827]
Warning: Hessian is not positive definite. Terminating.
> In Thema3_new>test (line 70)
In Thema3_new (line 132)
Minimum is: [1.000000, -1.000000] = 0.135335
Eigen values of hess are: [-1.0827, 0.5413]
Warning: Hessian is not positive definite. Terminating.
> In Thema3_new>test (line 70)
In Thema3_new (line 133)
Minimum is: [-1.000000, 1.000000] = -0.135335
```

5 Θέμα 4

Η συγκεκριμένη μέθοδος αποτελεί τροποποίηση της μεθόδου Newton. Όπως παρατηρήθηκε στο προηγούμενο θέμα η μέθοδος Newton έχει τον περιορισμό του εσσιανού πίνακα που πρέπει να είναι θετικά ορισμένος. Με αυτήν την μέθοδο χρησιμοποιούμε έναν πίνακα όπου έχουμε προσθέσει μία νέα παράμετρο μκ στα διαγώνια στοιχεία του εσσιανού, τέτοιο ώστε να είναι μεγαλύτερο από την απόλυτη τιμή της μικρότερης ιδιοτιμής του εσσιανού. Έτσι, ο νέος πίνακας προκύπτει θετικά ορισμένος και λύνουμε αυτό το πρόβλημα.

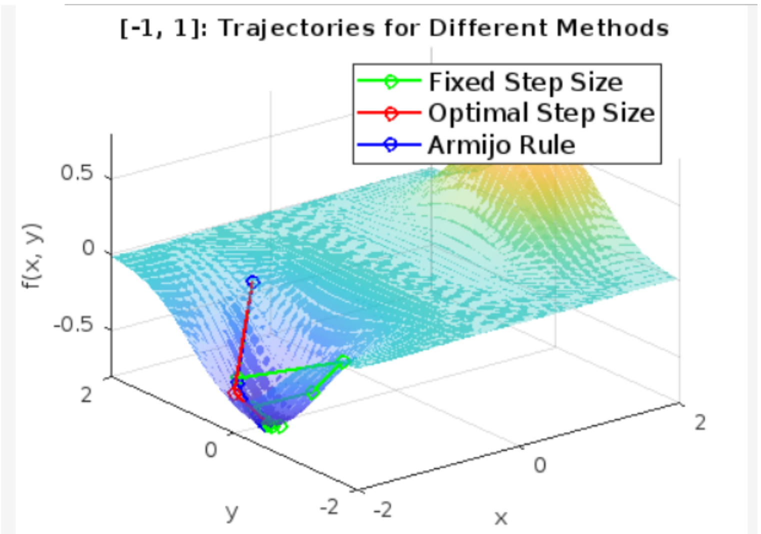
5.1 Αρχικό σημείο: (0,0)



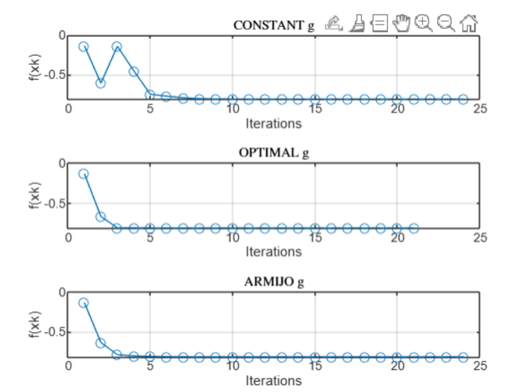


Στην περίπτωση σταθερού βήματος γ, ο αλγόριθμος δεν συγκλίνει για καμία από τις περιπτώσεις και πάλι καθώς βρίσκεται σε τοπικό ακρότατο.

5.2 Αρχικό σημείο: (-1,1)

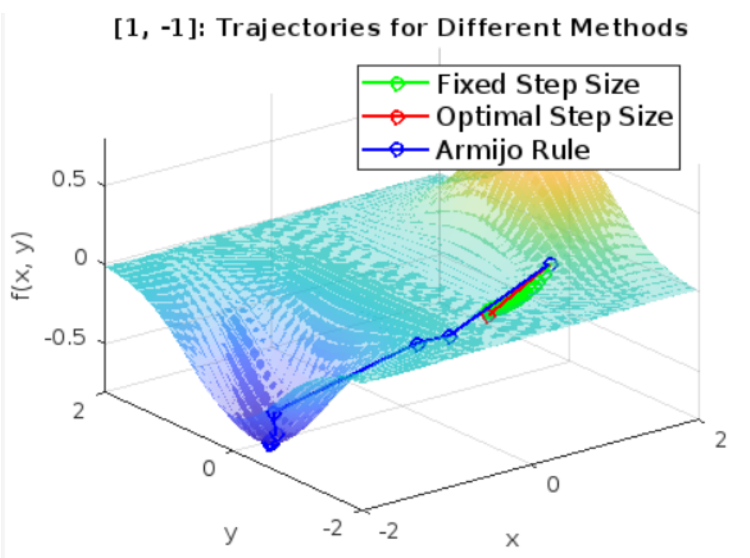


Βλέπουμε πως η σύγκλιση είναι εφικτή και για τις τρεις περιπτώσεις.

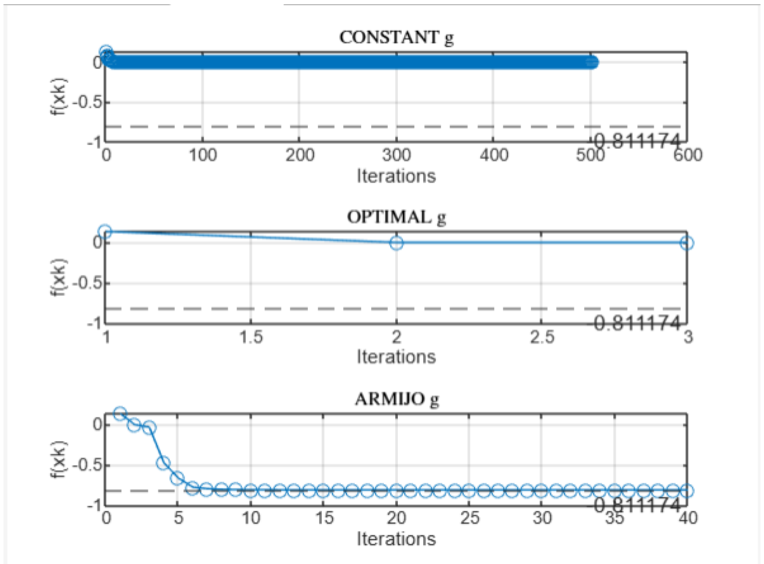


Φαίνεται ότι η μέθοδος Armijo συγκλίνει ταχύτερα μαζί με την μέθοδο Newton, ενώ αμέσως μετά είναι η μέθοδος σταθερού βήματος γ.

5.3 Αρχικό σημείο: (1,-1)



Βλέπουμε πως η σύγκλιση είναι εφικτή μόνο για την δεύτερη περίπτωση.



Περίπτωση α)

Ο αλγόριθμος δεν καταφέρνει να συγκλίνει γιατί πάλι πέφτει στο τοπικό ακρότατο κοντά για χ=0. Μόλις στην πρώτη επανάληψη συγκλίνει στο τοπικό ελάχιστο.

Περίπτωση β)

Ο αλγόριθμος δε συγκλίνει καθώς πέφτει πάνω στο τοπικό ελάχιστο.

Περίπτωση γ)

Η μέθοδος που χρησιμοποιεί τον κανόνα armijo καταφέρνει να συγκλίνει εκεί που η παραπάνω μέθοδος δε το καταφέρνει.Φαίνεται πως η σταθερότητα ως προς την σύγκλιση που παρέχει αυτός ο αλγόριθμος παίζει σημαντικό ρόλο εδώ.

6 Συμπεράσματα

Η μέθοδος της μέγιστης καθόδου απαιτεί περισσότερες επαναλήψεις απο την μέθοδο L-M. Αυτο οφείλεται στο γεγονός ότι γεωμετρικά τα βήματα που χρησιμοποιούνται είναι κάθετα μεταξύ τους και αυτό κάνει την σύγκλιση πιο αργή. Ωστόσο, μπορούμε φυσικά και πάλι να οδηγηθούμε στο ελάχιστο με κατάλληλη επιλογή αρχικού σημείο εκκίνησης και βήματος.

Η μέθοδος Newton, όπως φάνηκε δεν ήταν κατάλληλη για το συγκεκριμένο πρόβλημα δοσμένων των αρχικών σημείων, καθώς ο εσσιανός πίνακας δεν ήταν θετικά ορισμένος. Άρα, η επιτυχία του συγκεκριμένου αλγορίθμου εξαρτάται σημαντικά από το σημείο εκκίνησης.

Η μέθοδος L-M φαίνεται πως είναι η βέλτιση καθώς φτάνουμε στο ελάχιστο με λιγότερες επαναλήψεις. Αποτελεί μια ¨βελτιστοποιημένη’ μέθοδο Newton.

Η επιλογή του σημείου εκκίνησης κρίθηκε πολύ σημαντική, καθώς όπως είδαμε τα αποτελέσματα ήταν αρκετά διαφορετικά συναρτήσει των σημείων εκκίνησης. Το σημείο [-1,1] ήταν πιο κόντα στο ελάχιστο σε σχέση με τα υπόλοιπα σημεία και γιάυτό είχαμε σύγκλιση όταν ξεκινούσαμε από εκεί. Στα άλλα 2 σημεία υπήρχε το πρόβλημα ότι είτε βρισκόμασταν πάνω σε τοπικό ακρότατο είτε φθάναμε σε ένα κατά την διαδρομή προς το ολικό ελάχιστο.

Αποδοτικότερη λύση συνολικά ήταν η επιλογή γκ μέσω του κανόνα Armijo, αμέσως επόμενη η επιλογή γκ με κανόνα της ελαχιστοποίησης της F(xk+1), και τέλος χειρότερη φάνηκε η λύση χρησιμοποιώντας σταθερό γ, κάτι που μπορεί να οδηγήσει τον αλγόριθμο στο να καταλήξει σε τοπικό ακρότατο. Θα χρησιμοποιούσαμε Armijo γενικά, προκειμένου να εγγυηθούμε την σύγκλιση ή ίσως για μεγαλύτερης διάστασης προβλήματα, ενώ αν μας ενδιέφερε η ταχύτητα και η ακρίβεια θα επιλέγαμε μάλλον την μέθοδο της ελαχιστοποίησης του γκ.

Για όλες τις μεθόδους χρησιμοποιήθηκε συνάρτηση ελαχιστοποίησης ως προς γκ αυτή της διχοτόμου με χρήση παραγώγων.