Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης

Πολυτεχνική σχολή

Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

ΤΕΧΝΙΚΕΣ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ

ΑΝΑΦΟΡΑ ΔΕΥΤΕΡΗΣ ΕΡΓΑΣΙΑΣ:

Ημερομηνία 25/11/2024

Λιαροπούλου Κλεοπάτρα (ΑΕΜ : 10066, Email : liaropou@ece.auth.gr)

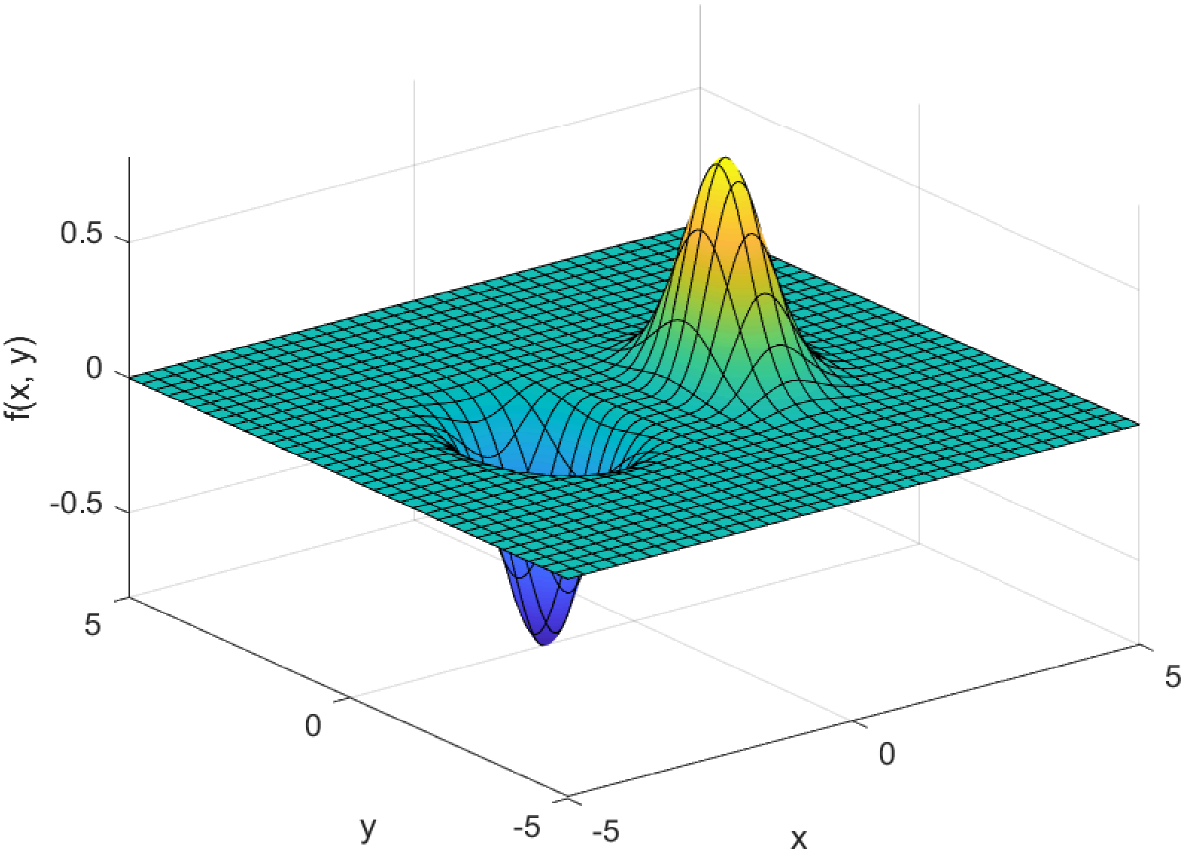
ΠΕΡΙΕΧΟΜΕΝΑ :

1.Εισαγωγή:

Στην προκειμένη εργασία παρουσιάζονται τα αποτελέσματα των εξής αλγορίθμων ελαχιστοποίησης:

- Μέθοδος Steepest descent
- Μέθοδος Newton
- Μέθοδος Levenberg-marquardt

Η γραφική παράσταση που θα μελετηθεί είναι η $f(x) = x^5 e^{-x^2-y^2}$ της οποίας η γραφική παράσταση απεικονίζεται στο *σχημα 1*:

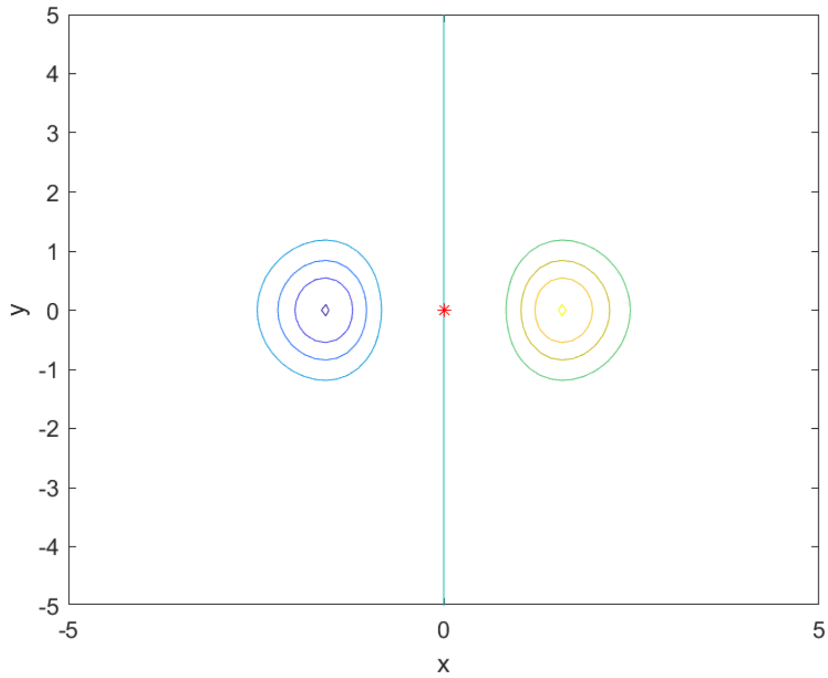


Σχήμα 1: Γραφική παράσταση της f(x,y)

2.Μέθοδος steepest descent

Στην μέθοδο της μέγιστης καθόδου η επαναληπτική διαδικασία ορίζεται μέσω του τύπου: $x_k + 1 = x_k - \gamma_k \nabla f(x_k)$

<u>Σημείο εκκίνησης (0,0):</u>
Για σημείο εκκίνησης του αλγορίθμου (0,0) παρατηρούμε ότι ο αλγόριθμος παραμένει στο ίδιο σημείο και δεν πραγματοποιούνται άλλες επαναλήψεις του αλγορίθμου καθώς αυτός τερματίζεται εφόσον η συνθήκη τερματισμού ∇f < ε ικανοποιείται. Δηλαδη, για το σημείο (0,0) το διάνυσμα ∇f είναι μηδενικό το οποίο σημαίνει ότι στο (0,0) υπάρχει τοπικό ελάχιστο και άρα το συγκεκριμένο σημείο είναι σημείο εγκλωβισμού. Αυτό φαίνεται και στο σχήμα όπου δεν σχηματίζεται κάποια γραμμή που να συμβολίζει την πορεία που ακολουθούν τα σημεία με την πάροδο των επαναλήψεων.



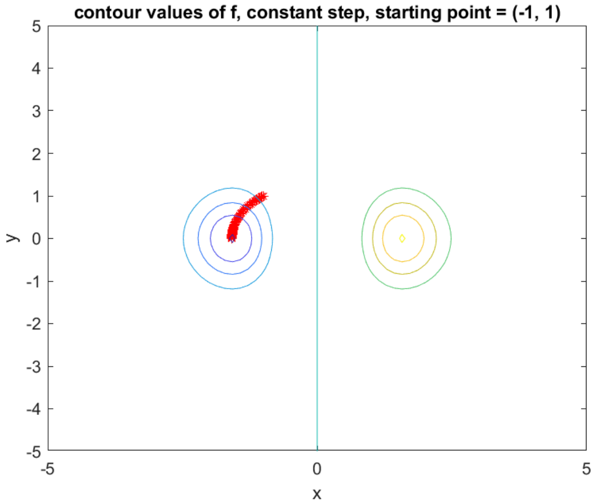
Αυτό ισχύει και για τους τρεις τρόπους προσέγγισης (Σταθερού βήματος, ελαχιστοποίηση και Armijo) σε όλες τις μεθόδους (Steepest Descent, Newton, Levenberg Marquardt) καθώς αυτή η ιδιότητα σχετίζεται με την μορφή της συναρτήσης και όχι τις μεθόδους ελαχιστοποίησης. Συνεπώς, θεωρείται δεδομένο αυτό το φαινόμενο για το υπόλοιπο της εργασίας.

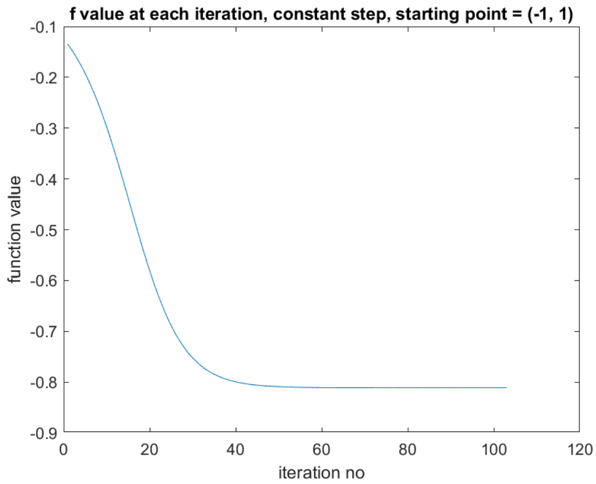
Στο Θέμα 2, η ελαχιστοποίηση της συνάρτηση f πραγματοποιείται με τη χρήση της μεθόδου Μέγιστης Καθόδου. Κάθε φορά διαφέρει η επιλογή του βήματος γ_k:

- Σταθερό βήμα = 0.05
- Βήμα που ελαχιστοποιεί τη συνάρτηση $f(x_k + \gamma_k d_k)$
- Βήμα που καθορίζεται βάσει του κανόνα του Armijo

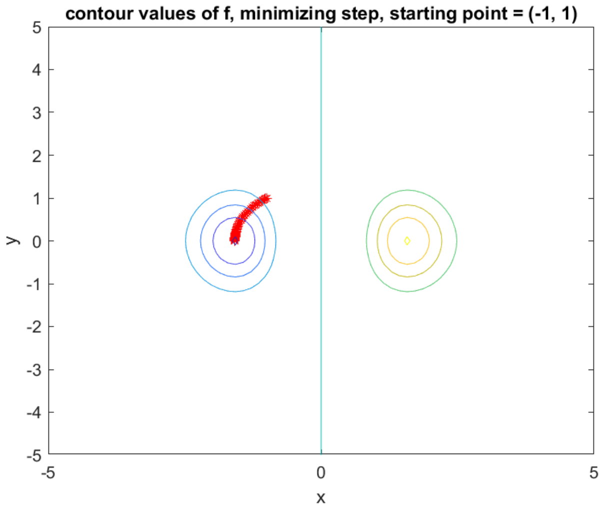
Σημείο εκκίνησης (-1,1):

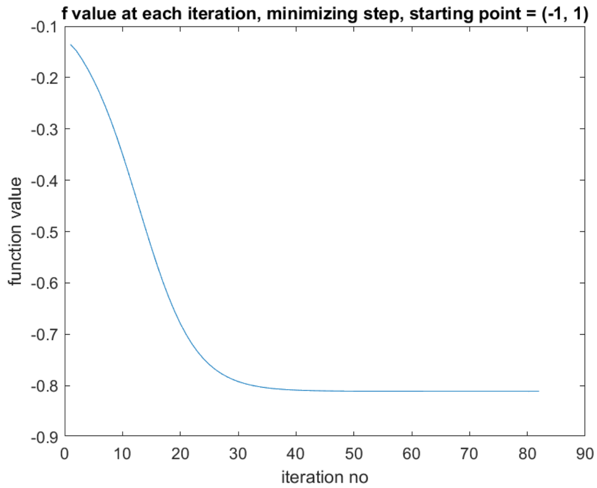
- Σταθερό βήμα = 0.05



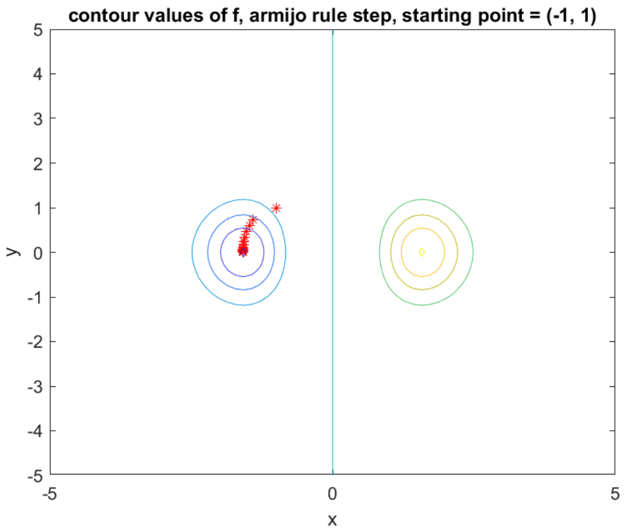


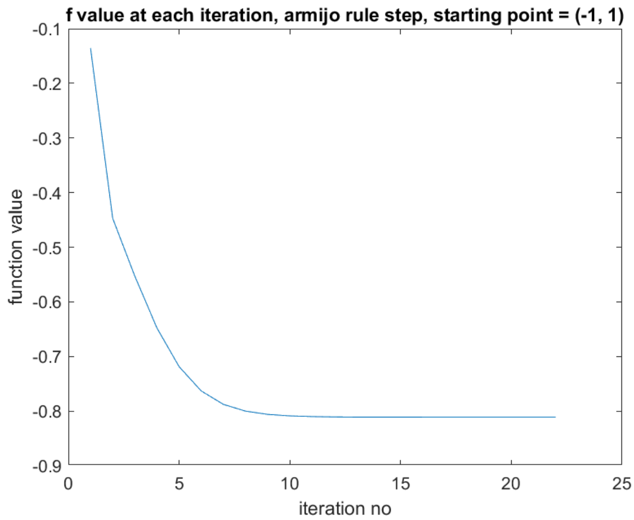
- Βήμα που ελαχιστοποιεί τη συνάρτηση $f\,(x_k + \gamma_k d_k)$





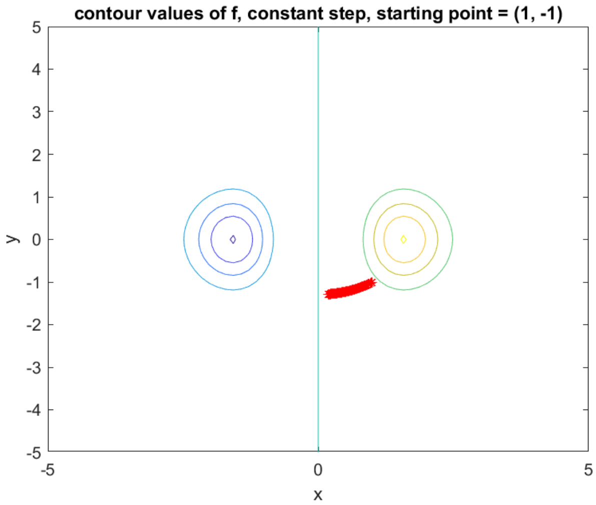
- Βήμα που καθορίζεται βάσει του κανόνα του Armijo

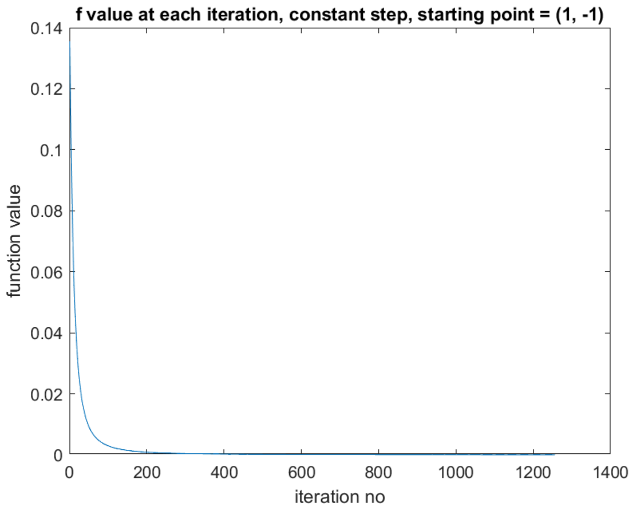




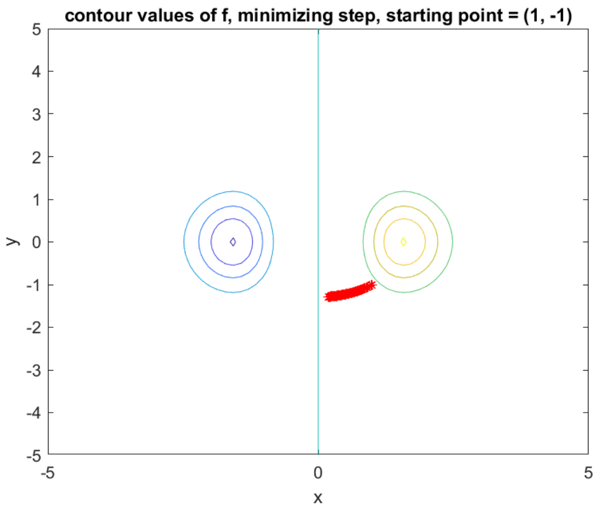
Σημείο εκκίνησης (1,-1):

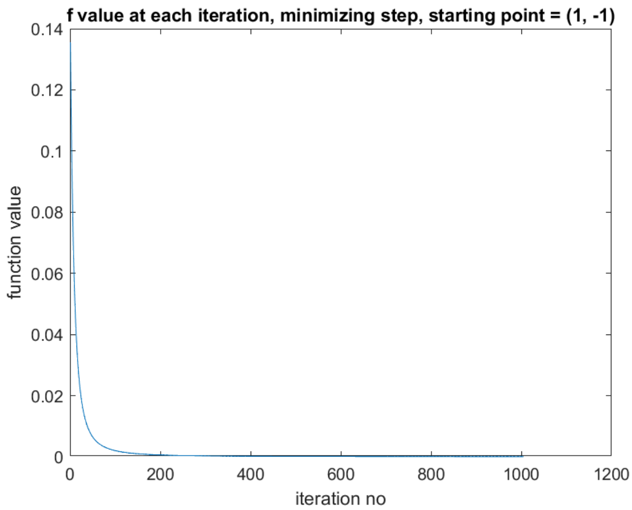
- Σταθερό βήμα = 0.05



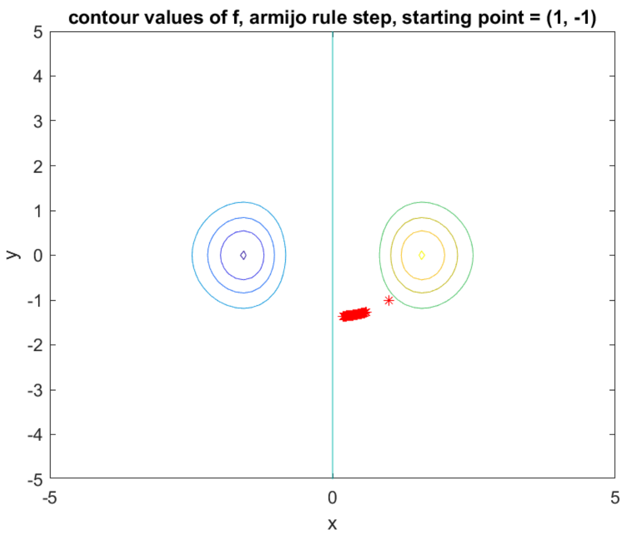


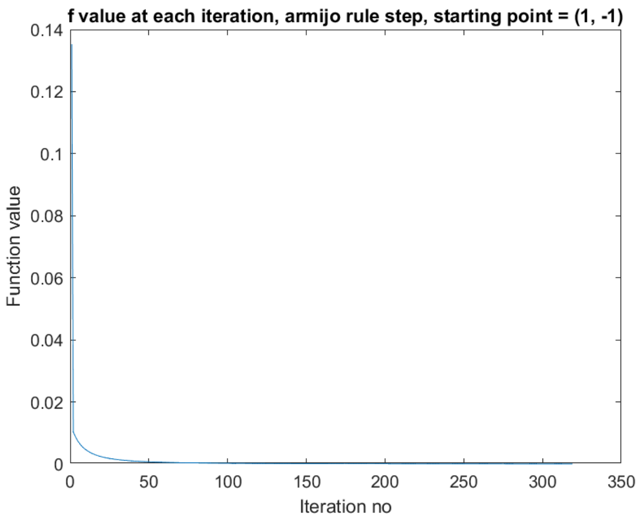
- Βήμα που ελαχιστοποιεί τη συνάρτηση $f\,(x_k + \gamma_k d_k)$





- Βήμα που καθορίζεται βάσει του κανόνα του Armijo





Συμπεράσματα μεθόδου steepest descent:

Γενικά το (-1, 1) φαίνεται να είναι ένα ικανοποιητικό σημείο εκκίνησης των αλγορίθμων καθώς προσεγγίζει ακριβώς το ελάχιστο της συνάρτησης. Όσον αφορά την σύγκριση μεταξύ του τρόπου επιλογής του βήματος για την μέθοδο Steepest Decent, καλύτερη επιλογή βήματος προσφέρει η Armijo, έπειτα η ελαχιστοποίηση της $f\,(x_k + \gamma_k d_k)$ και τέλος η επιλογή σταθερού βήματος. Αυτό είναι προφανές και απο τα διαγράμματα που δείχνουν την τιμή της f με την πάροδο των επαναλήψεων όπου με την σειρά που υλοποιήθηκαν ισχύει ότι:

- Η σταθερου βήματος προσέγγισε το ελάχιστο σε 100 επαναλήψεις
- Η ελαχιστοποίηση σε 80
- Η Armijo σε μόλις 22

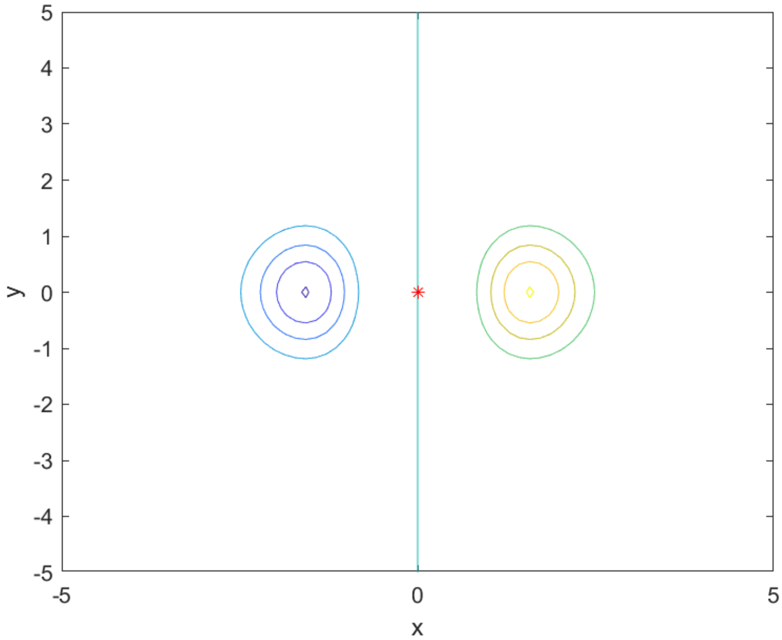
Αυτό φαίνεται άλλωστε στο διάγραμμα της τιμής της f που παρουσιάζει και πιό απότομη πτώση. Αντίστοιχη ιεραρχία ισχύει και για το σημείο εκκίνησης (1,-1) όπου αξίζει να σημειωθεί οτι σε αυτή την περίπτωση δεν προσεγγίζεται το ολικό ελάχιστο αλλά κάποιο τοπικό ελάχιστο. Αυτό συμβαίνει επειδή παρόλο που πλησιάζει όντως το ελάχιστο εγκλωβίζεται για τιμές κοντά στο χ = 0 με αποτέλεσμα να τερματιζει ο αλγόριθμος. Αυτό πιθανώς συμβαίνει γιατί στο σημείο που τερματίζει ισχύει η συνθήκη ∇f < ε.

3.Μέθοδος Newton

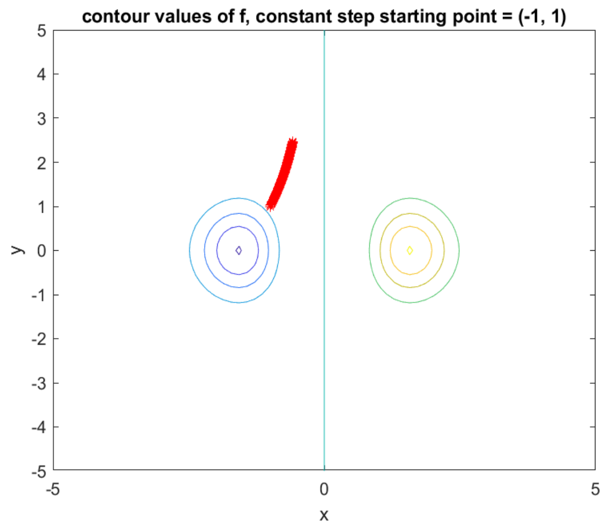
Η μέθοδος Newton αναζητεί σημείο ελαχίστου χρησιμοποιώντας την κατεύθυνση $d_k = -[\nabla^2 f(x_k)]^{-1}\nabla f(x_k)$. Αυτή η μέθοδος λαμβάνει υπόψη και τον εσσιανο πίνακα, ο οποίος πρέπει να είναι θετικά ορισμένος. Για την συγκεκριμένη συνάρτηση ο εσσιανός είναι ο εξής:

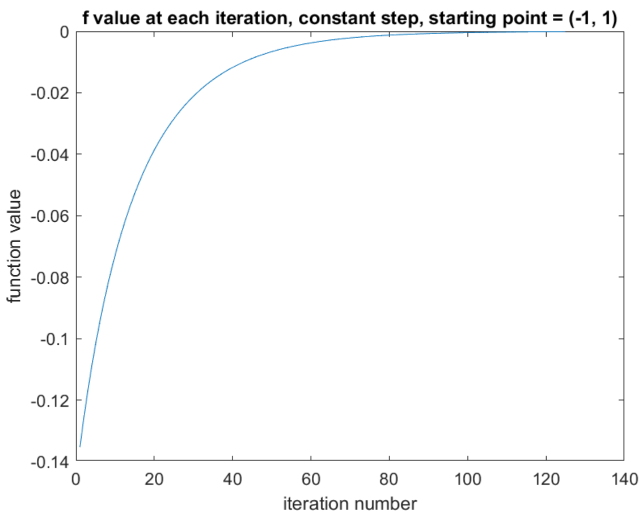
$$H = \begin{bmatrix} 4x^7e^{-x^2-y^2} - 22x^5e^{-x^2-y^2} + 20x^3e^{-x^2-y^2} & 4x^6ye^{-x^2-y^2} - 10x^4ye^{-x^2-y^2} \\ 4x^6ye^{-x^2-y^2} - 10x^4ye^{-x^2-y^2} & 4x^5y^2e^{-x^2-y^2} - 2x^5e^{-x^2-y^2} \end{bmatrix}$$

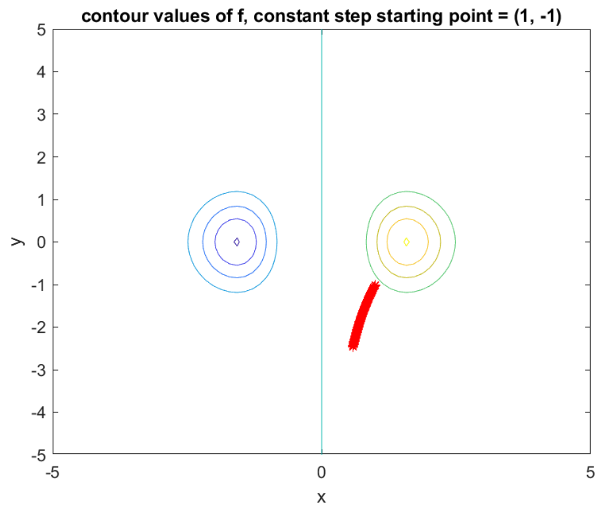
Για τα σημεία (0,0), (-1,1), (1,-1) ο εσσιανός δεν είναι θετικά ορισμένος, συνεπώς δεν πληρούνται οι προϋποθέσεις για να εφαρμοστεί μεθοδος Newton. Όπως προηγουμένως για σημείο εκκίνησης (0,0) ο αλγόριθμος εγκλωβίζεται και τερματίζεται από την πρώτη επανάληψη.

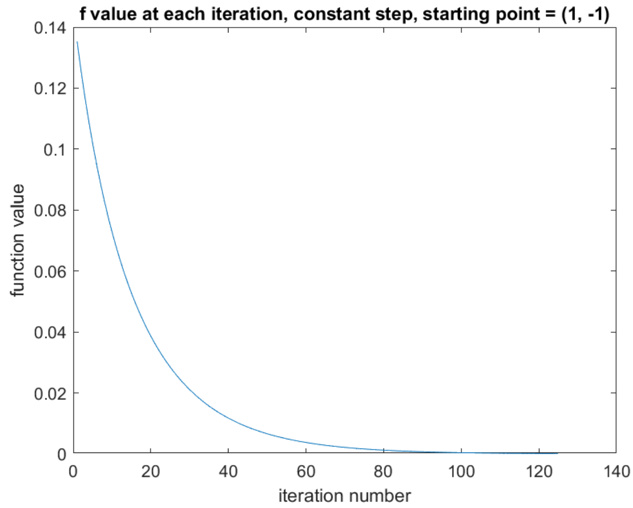


Ωστόσο για σημεία εκκίνησης (-1,1) όσο και για (1,-1) ο εσσιανός πίνακας (∇^2f) όπως τονίστηκε φαίνεται να είναι **αρνητικά ορισμένος** και για αυτόν τον λόγο ο αλγόριθμος δεν δίνει αποτέλεσμα ελαχίστου της συνάρτησης. Άλλωστε αυτό είναι εμφανές και από τα διαγράμματα παρακάτω στα οποία είναι καρφιτσωμένο το σημείο εκκίνησης και είναι προφανές ότι απομακρυνόμαστε από το ελάχιστο. Συμπερασματικά, η προκειμένη συνάρτηση δεν είναι ιδανική για την μέθοδο Newton καθώς και οι περιορισμοί για τον εσσιανο πινακα την καθιστούν μη εφαρμόσιμη. Τα αποτελέσματα μη συγκλισης είναι κοινά για όλες τις μεθόδους επιλογής βήματος δηλαδή δεν εξάγεται αποτέλεσμα.









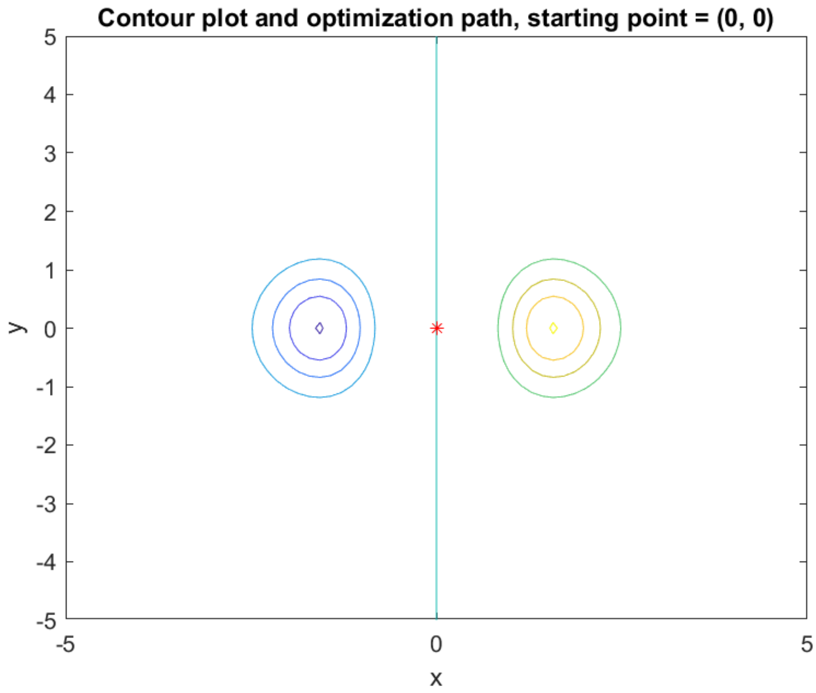
Για το σημείο (1, -1) όπως βλέπουμε και στα διαγράμματα της τιμής συνάρτησης με τις επαναλήψεις φαίνεται να ελαχιστοποιείται η συνάρτηση. Ωστόσο, όπως και πριν εγκλωβίζεται σε ένα τοπικό ελάχιστο. Από τα παραπάνω γραφήματα και από το γεγονός ότι δεν πληρουνται οι προϋποθεσεις βγαίνει το συμπέρασμα ότι η Newton δεν είναι η κατάλληλη μέθοδος ελαχιστοποίησης για την συγκεκριμένη συνάρτηση.

4.Μέθοδος Levenberg Marquardt

Η μέθοδος αυτή αποτελεί μια βελτιωμένη τροποποίηση της μεθόδου Newton. Συγκεκριμένα, μπορεί να παρακαμφθεί το ζήτημα του αρνητικά ορισμένου εσσιανού που παρουσιαστηκε νωρίτερα. Η διαφοροποίηση είναι η εξής: Ορίζεται συντελεστής μ_k τέτοιος ώστε ο πίνακας $D = \nabla^2 f(x_k) + \mu_k I$ να είναι **πάντα θετικά ορισμένος**. Ως αποτέλεσμα μπορεί να γίνει αντιστροφή του D.

<u>Σημείο εκκίνησης (0,0):</u>

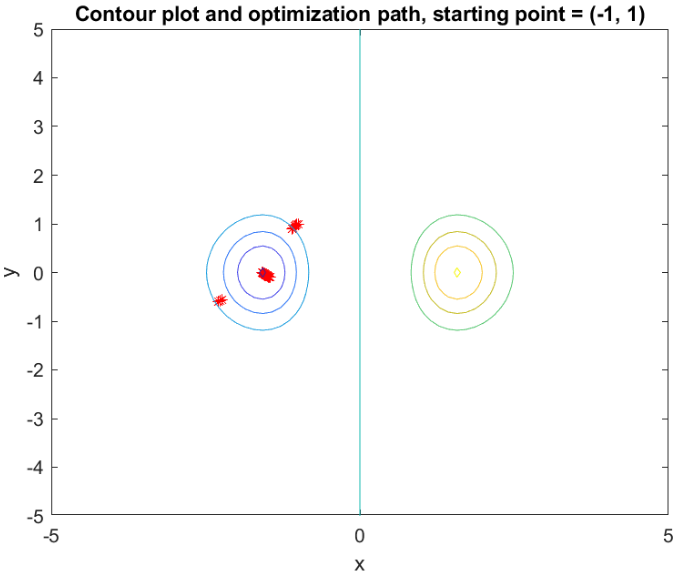
Όπως και πριν ο αλγόριθμος τερματίζεται κατευθείαν καθώς εγκλωβίζεται ($\nabla f(x_k) = 0$)

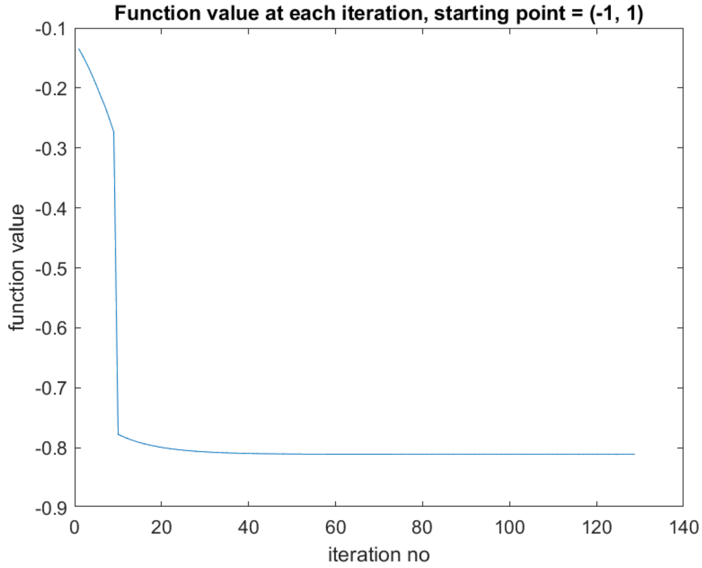


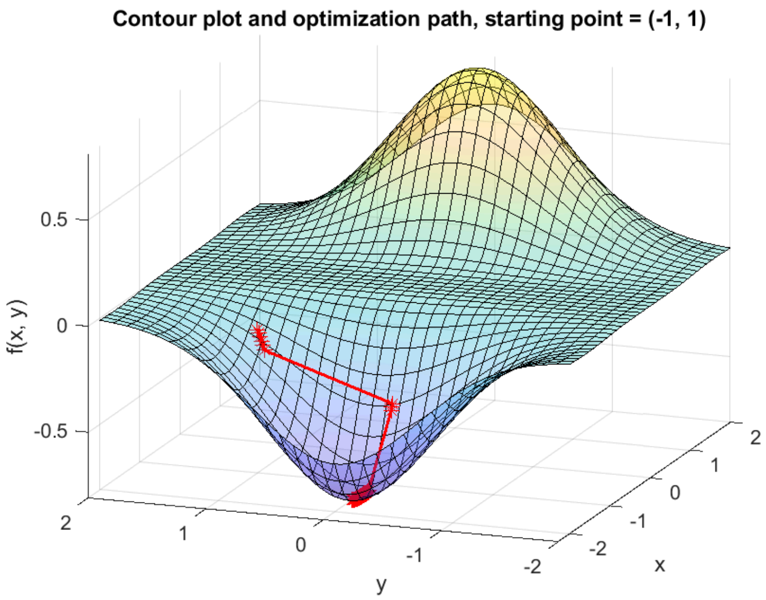
Σημείο εκκίνησης (-1, 1):

Παρατηρούμε ότι σε αυτήν την περίπτωση ο αλγόριθμος λειτουργεί ορθα και δεν υπάρχει ζήτημα να επηρεάσει ο αρνητικά ορισμενος εσσιανος πίνακας. Η σύγκλιση γίνεται κανονικά:

- Σταθερό βήμα = 0.05

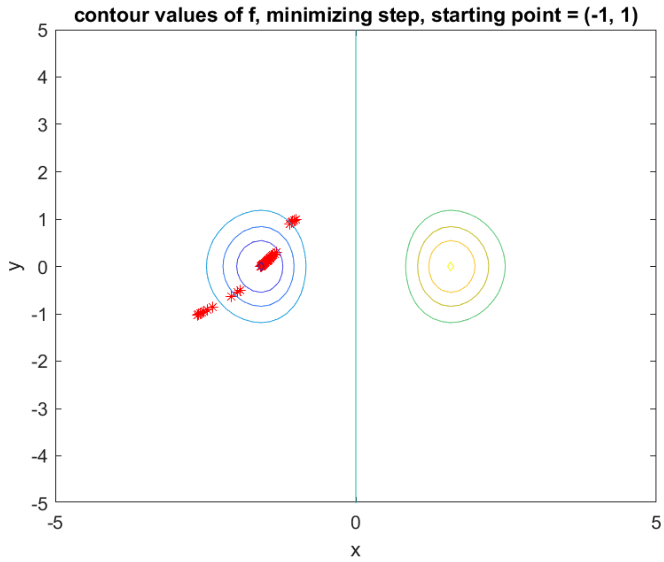


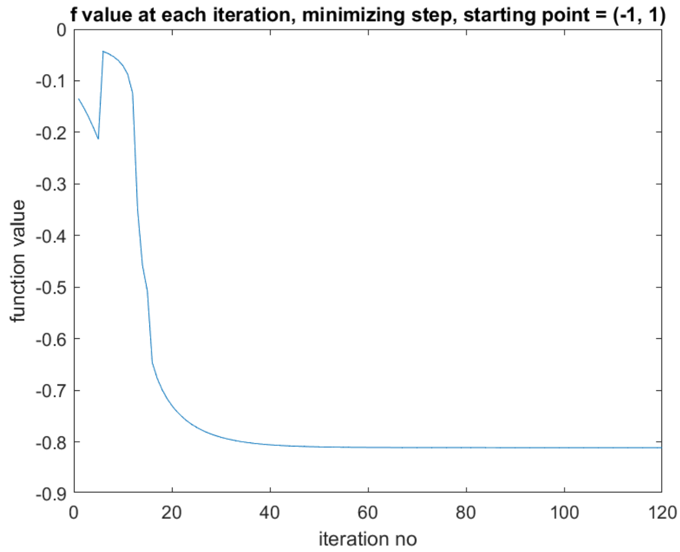


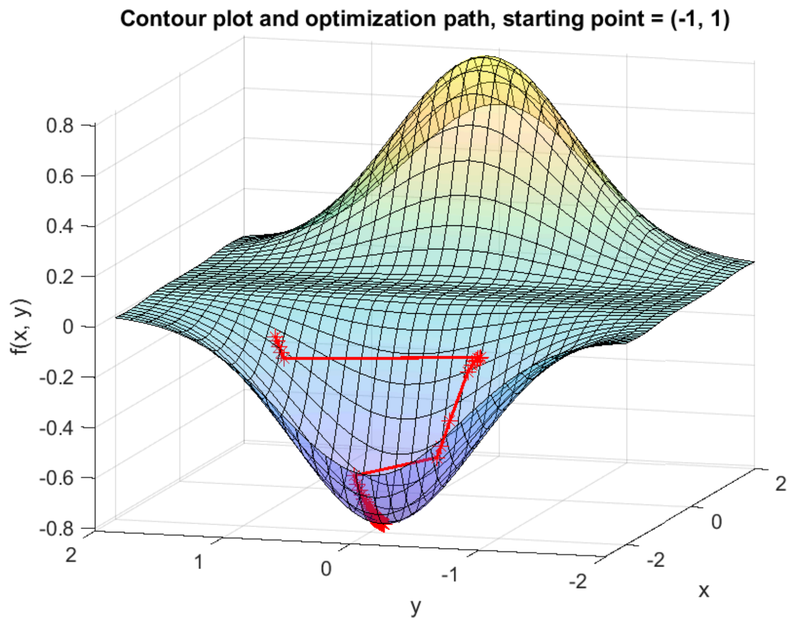


Η παραπάνω γραφική παράσταση οπτικοποιεί την επιτυχημένη εύρεση του ελαχίστου σε τρισδιάστατο σχήμα καθώς στο δισδιάστατο διάγραμμα δεν ήταν κατανοητή η σειρά των σημείων με την τεθλασμένη γραμμή.

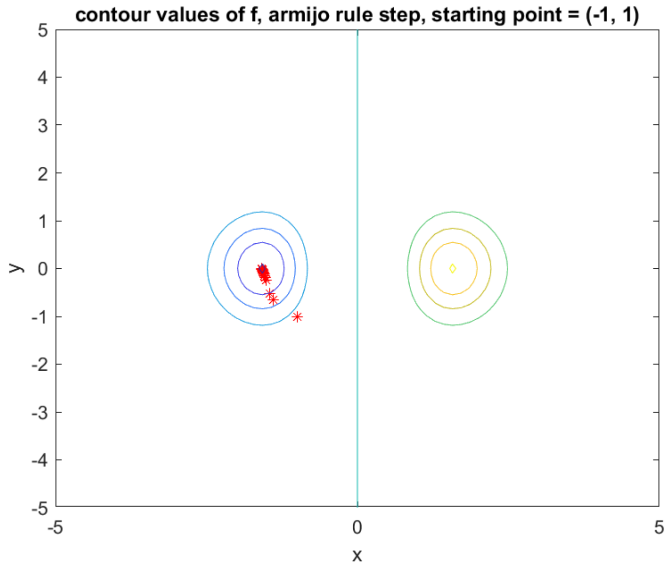
- Βήμα που ελαχιστοποιεί τη συνάρτηση $f\,(x_k + \gamma_k d_k)$

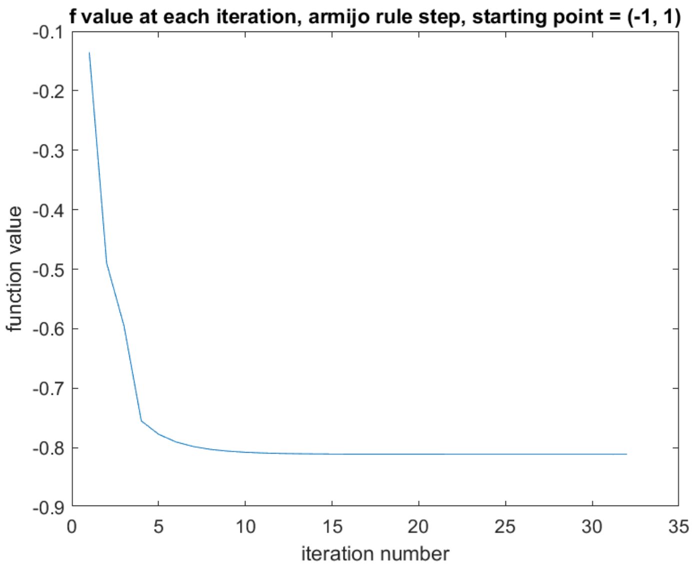


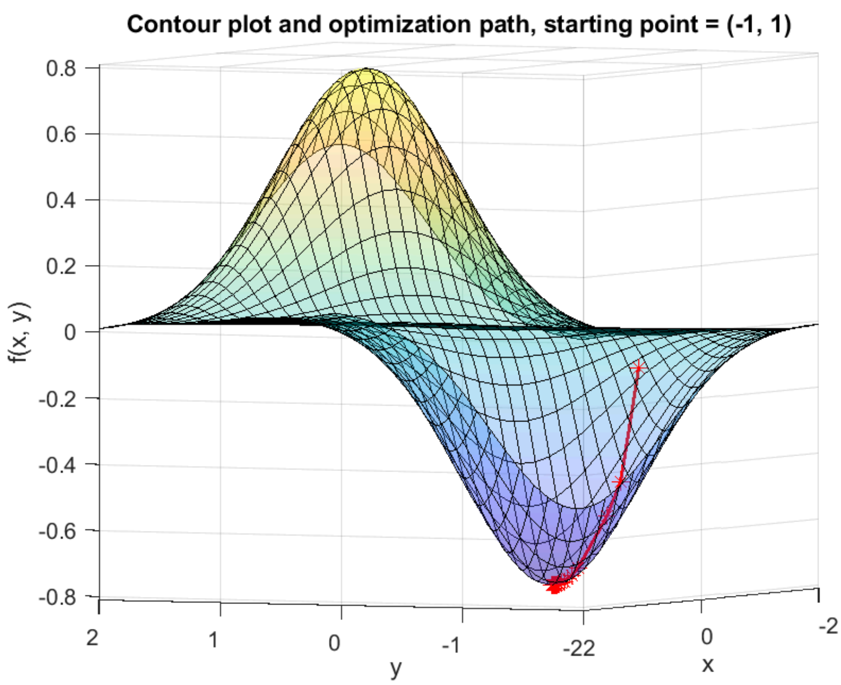




- Βήμα που καθορίζεται βάσει του κανόνα του Armijo





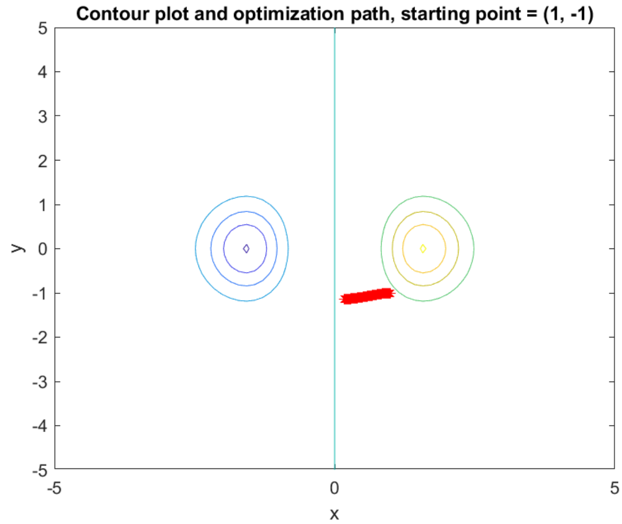


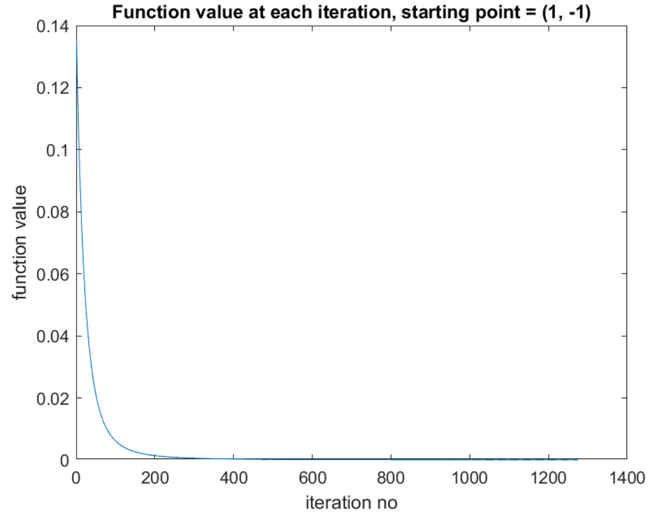
<u>Συμπεράσματα για το σημείο εκκίνησης (-1, 1):</u>

Γενικά το (-1, 1) φαίνεται να είναι ένα πολύ καλό σημείο εκκίνησης των αλγορίθμων καθώς όπως φαίνεται στα τρισδιάστατα διαγράμματα προσεγγίζει ακριβώς το ελάχιστο της συνάρτησης. Όσον αφορά την σύγκριση μεταξύ του τρόπου επιλογής του βήματος για την μέθοδο Levenberg Marquardt καλύτερη επιλογή βήματος προσφέρει η Armijo, έπειτα η ελαχιστοποίηση της $f\,(x_k\;+\;\gamma_k d_k)$ και τέλος η επιλογή σταθερού βήματος. Αυτό είναι προφανές και απο τα διαγράμματα που δείχνουν την τιμή της f με την πάροδο των επαναλήψεων όπου η σταθερου βήματος προσέγγισε το ελάχιστο σε 130 επαναλήψεις, η ελαχιστοποίηση σε 120 ενώ η Armijo σε μόλις 33.

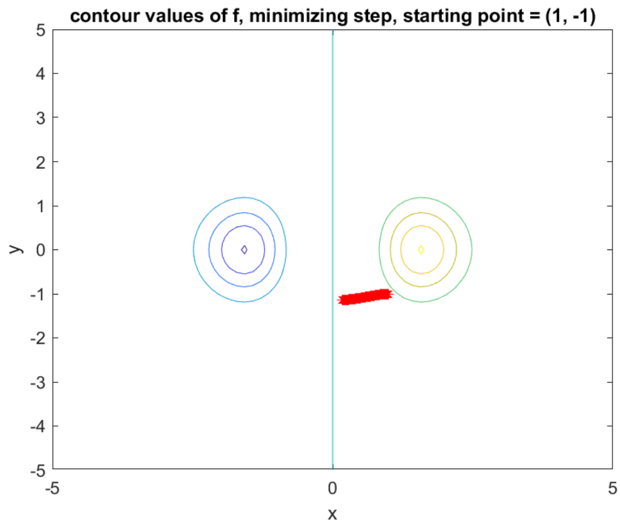
Σημείο εκκίνησης (1, -1):

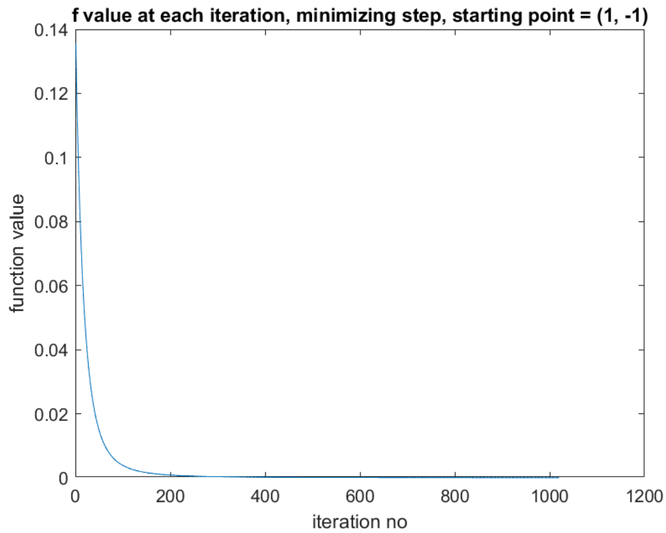
- Σταθερό βήμα = 0.05



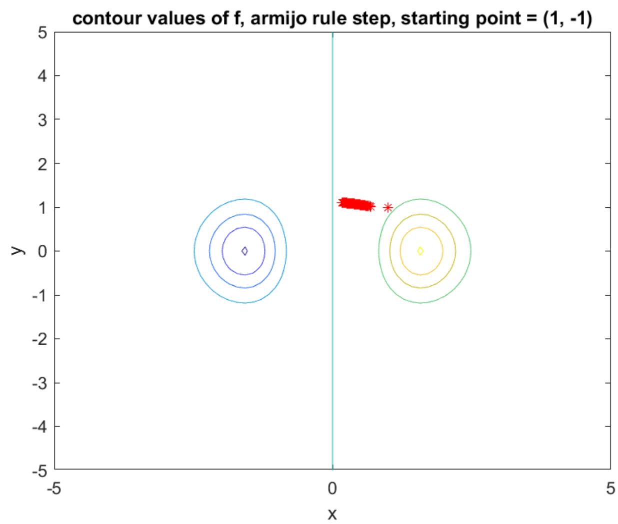


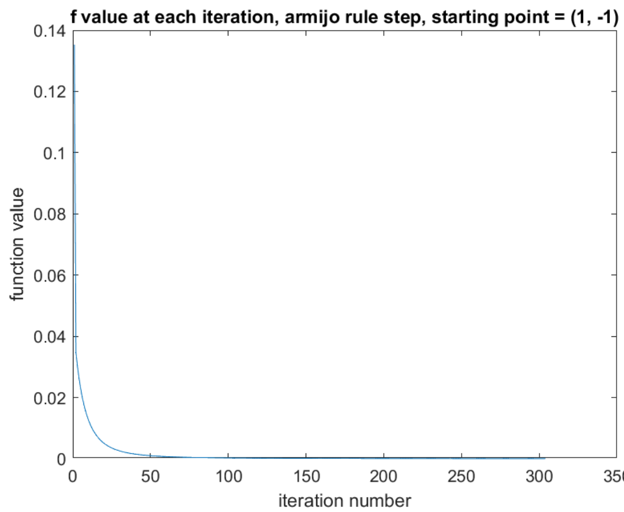
- Βήμα που ελαχιστοποιεί τη συνάρτηση $f\ (x_k + \gamma_k d_k)$





- Βήμα που καθορίζεται βάσει του κανόνα του Armijo





Συμπεράσματα για το σημείο εκκίνησης (1, -1):

Γενικά το σημείο (1, -1) δεν είναι αρκετά ικανοποιητικό σημείο εκκίνησης των αλγορίθμων καθώς όπως φαίνεται στα διαγράμματα δεν προσεγγίζει το ολικό ελάχιστο της συνάρτησης αλλά κάποιο τοπικό ελαχιστο. Αυτο συμβαίνει καθώς φτάνει σε κάποιο σημείο εγκλωβισμού όπου το $\nabla f(x_k)$ έχει μικρή τιμή και εκεί ο αλγόριθμος τερματίζει. Όσον αφορά την σύγκριση μεταξύ του τρόπου επιλογής του βήματος για την μέθοδο Levenberg Marquardt καλύτερη επιλογή βήματος προσφέρει η Armijo, έπειτα η ελαχιστοποίηση της $f(x_k + \gamma_k d_k)$ και τέλος η επιλογή σταθερού βήματος. Έστω και αν προσεγγίζεται κάποιο τοπικό ελάχιστο, απο τα διαγράμματα που δείχνουν την τιμή της f με την πάροδο των επαναλήψεων η μέθοδος σταθερου βήματος προσέγγισε το ελάχιστο σε 1300 επαναλήψεις, η ελαχιστοποίηση σε 1000 ενώ η Armijo σε μόλις 300.

5.Συμπεράσματα

Η επιλογή του αρχικού σημείου είναι ιδιαιτέρως καθοριστική στην αποτελεσματικότητα του αλγορίθμου. Αυτό έγινε αισθητό σε ολόκληρη την αναφορά καθώς ιδανικό σημείο εκκίνησης του αλγορίθμου κρίθηκε το (-1, 1) ωστέ να μπορούν να εξαχθούν ασφαλή συμπεράσματα κάθε φορά. Έπειτα, το σημείο (0,0) ως σημείο εγκλωβισμού με $\nabla f = 0$ δεν ενδείκνυται καθώς τερματίζει κατευθείαν τον αλγόριθμο. Τέλος, το (1, -1) παρουσιάζεται ώς ενδιάμεση λύση εφόσον δεν τερματίζεται ακαριαία ο αλγόριθμος αλλα ωστόσο προσεγγίζεται ένα τοπικό ελάχιστο κόντα στον άξονα του χ= 0 και όχι το ολικό μέγιστο της συνάρτησης. Συγκεκριμένα, γινοταν εγκλωβισμός κοντά στο 0 λόγω της οριζόντιας κλίσης της συνάρτησης. Άρα, ως βελτιστη επιλογή είναι κάποιο αρχικό σημείο που να είναι ουτωσιαλλως σχετικά κοντά στο ελάχιστο της συνάρτησης που στην προκειμένη περίπτωση είναι το (-1,1).

Ως προς τις μεθόδους ελαχιστοποίησης που χρησιμοποιήθηκαν, η Newton αποδείχθηκε ιδιαίτερα μη λειτουργική για την προκειμένη περίπτωση λόγω των περιορισμών του εσσιανου πινακα να είναι θετικά ορισμένος. Ωστόσο η μέθοδος Levenberg Marquardt που ξεπερνά αυτό ακριβώς το πρόβλημα αποδεικνύεται περισσότερο λειτουργική δεδομένου ότι τερμάτιζε φτάνοντας στο ελάχιστο με τον μικρότερο αριθμό επαναλήψεων σε σχέση με τις υπόλοιπες μεθόδους.

Όσον αφορά την μέθοδο επιλογής βήματος, η Armijo φαίνεται να είναι η βέλτιστη καθώς από τα διαγράμματα της τιμής της f με την πάροδο των επαναλήψεων, ο αλγόριθμος έβρισκε το ελάχιστο με τις λιγότερες επαναλήψεις. Τέλος, ακολουθουν με την σειρά ώς προς την αποδοτικότητα η επιλογή βήματος που ελαχιστοποιεί την $f\,(x_k + \gamma_k d_k)$ και η επιλογή σταθερού βήματος.

6.Βιβλιογραφικές πηγές:

[1] Γεώργιος Α. Ροβιθάκης. Τεχνικές Βελτιστοποίησης. Εκδόσεις ΤΖΙΟΛΑ.