ΝΙΟΒΗ ΠΕΡΑΤΙΚΟΥ
ΑΕΜ:10494

Τεχνικές Βελτιστοποίησης

Άσκηση 2: Εύρεση ελαχίστου σε συνάρτηση 2 μεταβλητών

Σε αυτήν την άσκηση μας ζητήθηκε η μελέτη τριών μεθόδων:

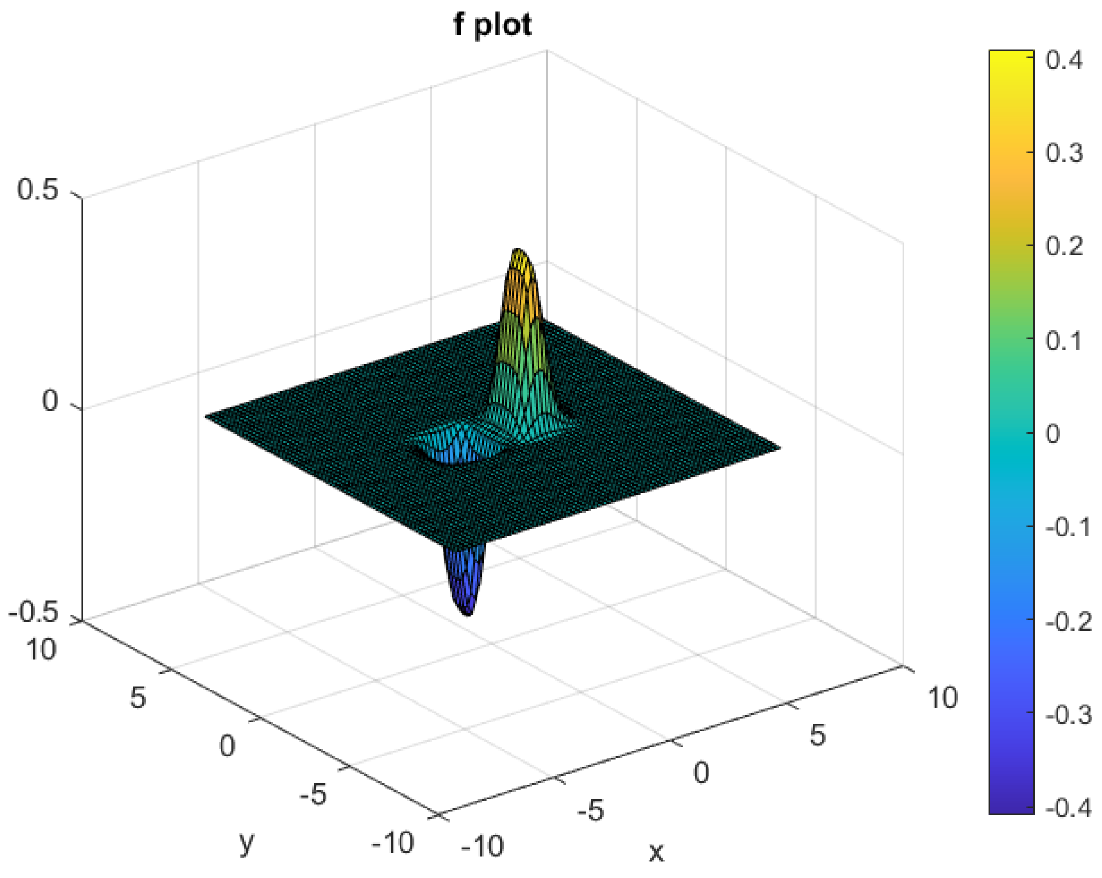
1. Μέθοδος Μέγιστης Καθόδου(Steepest Descent)
2. Μέθοδος Newton
3. Μέθοδος Levenberg-Marquardt

σε συνάρτηση f, για δοσμένη αντικειμενική συνάρτηση:

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

Οι τρεις αυτοί μέθοδοι, ανήκουν στις μεθόδους κλίσης.

Σχεδιάζω την f χρησιμοποιώντας την plot.m και παίρνω μια γενική εικόνα της μορφής της:



ΝΙΟΒΗ ΠΕΡΑΤΙΚΟΥ
ΑΕΜ:10494

Ζητούμενα:

Μας ζητήθηκε να υλοποιήσουμε και να εφαρμόσουμε τη μέθοδο Μέγιστης Καθόδου, την μέθοδο Newton και την μέθοδο Levenberg – Marquardt για να ελαχιστοποιήσουμε την συνάρτηση παίρνοντας τα αρχικά σημεία i) (0,0), ii) (-1,-1), iii) (1,1).

Το βήμα $γ_κ$ θα επιλεγεί:

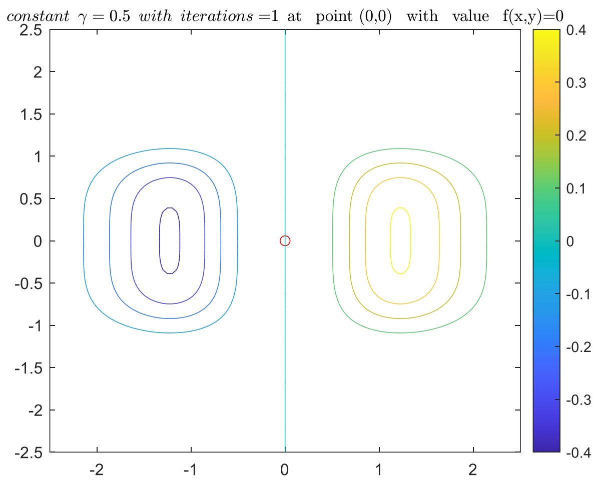
- σταθερό της επιλογής μας
- μεταβλητό τέτοιο ώστε σε κάθε επανάληψη να ελαχιστοποιείται η f ($x_κ$ +$γ_κ$ · dk)
- βάσει του κανόνα Armijo

Μέθοδος της Μέγιστης Καθόδου:

Η γενική ιδέα της μεθόδου είναι, ότι μεγαλύτερη αύξηση, || dx ||, της μεταβλητής x, θα υπάρξει όταν το διάνυσμα dx είναι συγγραμικό με το διάνυσμα κλίσης, ∇f, της f.

Συγκεκριμένα αν στην $xκ+1=xκ−γκΔκ\nabla f(xκ)$, όπου $Δ_κ$ θέσουμε τον μοναδιαίο πίνακα I, θα έχουμε $xκ+1=xκ−γκ\nabla f(xκ)$.

1. Αρχικό σημείο (0,0) για τη συνάρτηση f
   a. Με σταθερό γ=0.5

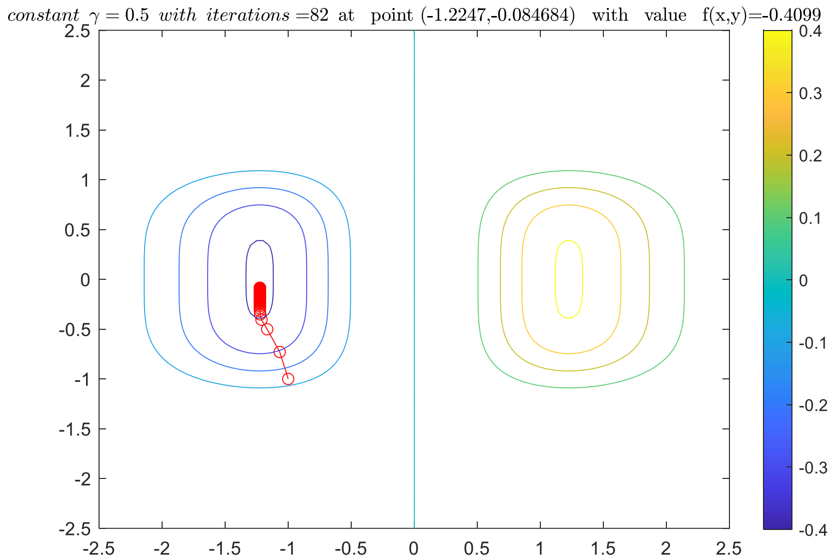


Για αρχικό σημείο (0,0) η παράγωγος της f είναι μηδέν οπότε ο αλγόριθμος τερματίζεται πρόωρα και συνεπώς εγκλωβιζόμαστε στο σημείο f(x,y) = 0. Όπου αυτό θα συμβεί και στην περίπτωση ελαχιστοποίησης της f($x_κ$ + $γ_κ$ * $d_κ$), και βάσει του κανόνα Armijo.

Άρα δεν θα ελεγχτεί το (0,0) για τις ακόλουθες μεθόδους.

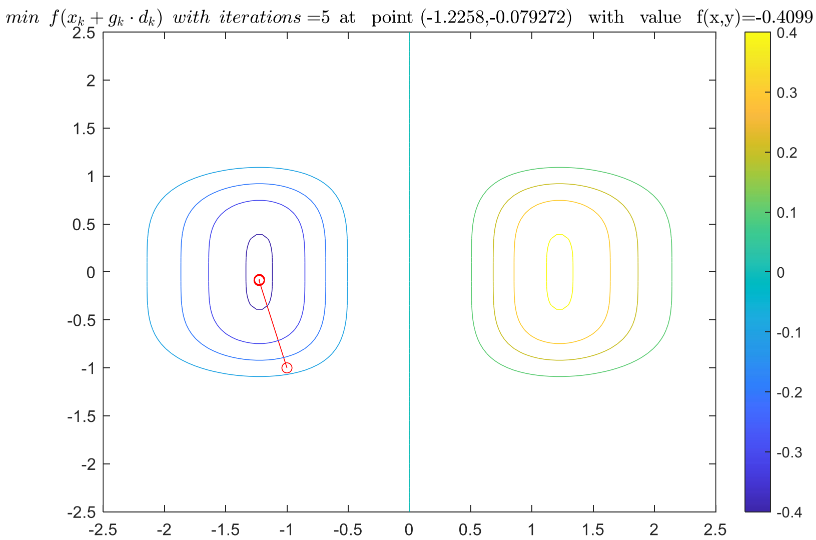
ΝΙΟΒΗ ΠΕΡΑΤΙΚΟΥ
AEM:10494

2. Αρχικό σημείο (-1,-1) για την συνάρτηση f
   a. Με σταθερό γ = 0.5



Για αρχικό σημείο (-1,-1) μετά από 82 επαναλήψεις καταλήγουμε στο σημείο (-1.2247,-0.084684) για δοσμένη ακρίβεια e=10^-3 με τιμή f(x, y)= -0.4099 όπου ταλαντώνεται γύρω από το τοπικό ελάχιστο.

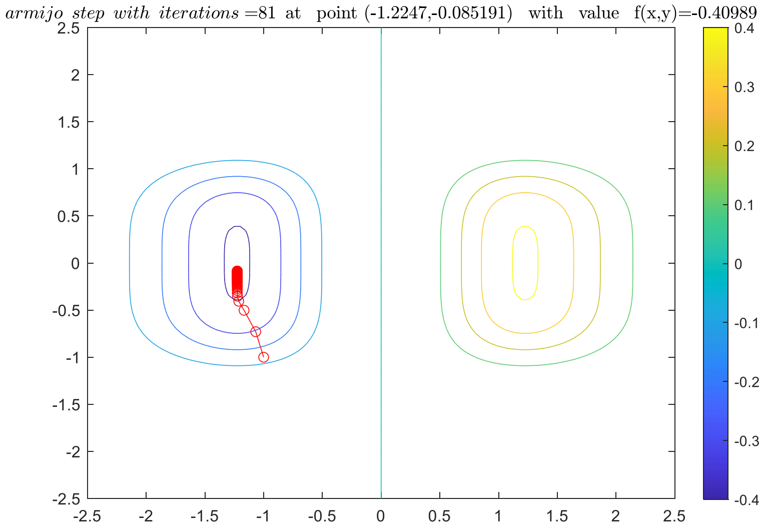
b) Για να ελαχιστοποιηθεί η συνάρτηση $h(\gamma)=f(xk+\gamma \kappa dk)$, αφού είναι συνάρτηση μιας μεταβλητής μετασχηματισμού χρησιμοποιώ την golden_section.m.



Για αρχικό σημείο (-1,-1) μετά από 5 επαναλήψεις καταλήγουμε στο σημείο (-1.2258,-0.079272) για δοσμένη ακρίβεια e=10^-3 με τιμή f(x, y)=-0.4099 που είναι το ολικό ελάχιστο.

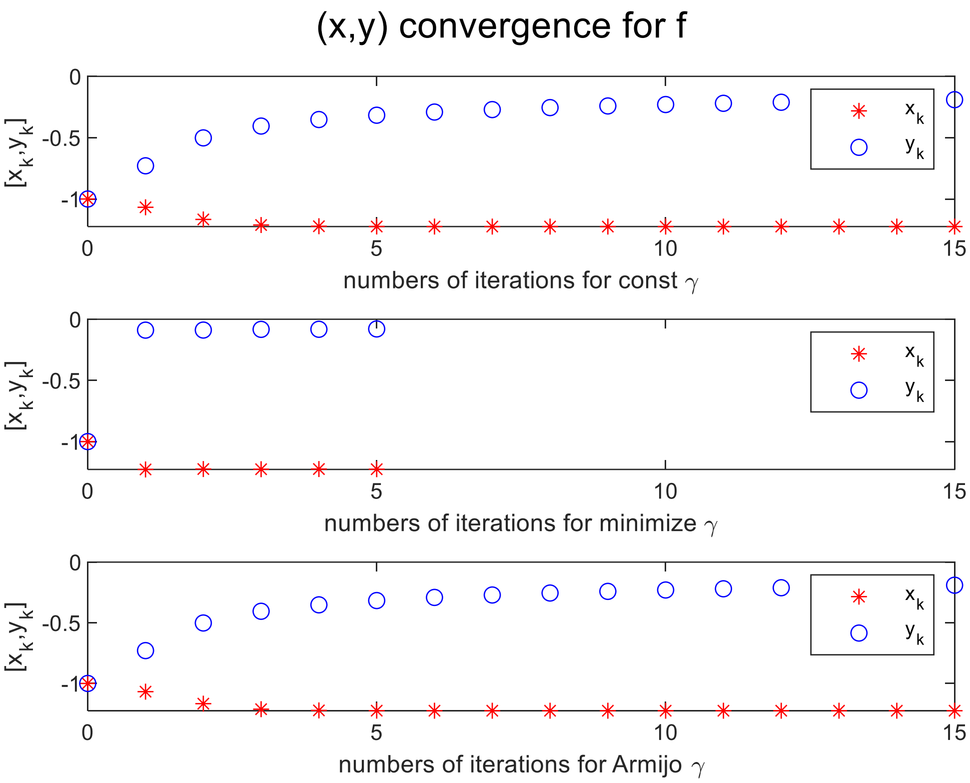
ΝΙΟΒΗ ΠΕΡΑΤΙΚΟΥ
ΑΕΜ:10494

c) Κανόνας Armijo. Είναι μια μέθοδος διαδοχικής μείωσης του γ_k.



Για αρχικό σημείο (-1,-1) μετά από 81 επαναλήψεις καταλήγουμε στο σημείο (-1.2247,-0.085191) για δοσμένη ακρίβεια e=10^-3 με τιμή f(x, y)=-0.40989 που είναι το ολικό ελάχιστο.

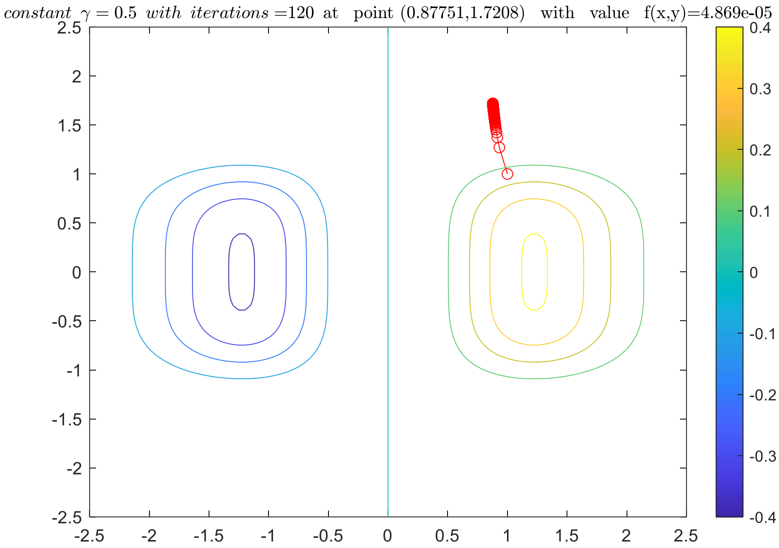
Για να γίνει πιο σαφής η σύγκριση μεταξύ των διαφορετικών <<μορφών>> της Μέγιστης Καθόδου, χρησιμοποιώ τη steepest_descent_x_y_compare.m, όπου το ζεύγος (x , y), για κάθε μια από τις παραπάνω μορφές, σχεδιάζεται συναρτήσει των αριθμών επαναλήψεων.
Για το σημείο (-1,-1), έχουμε:



ΝΙΟΒΗ ΠΕΡΑΤΙΚΟΥ
ΑΕΜ:10494

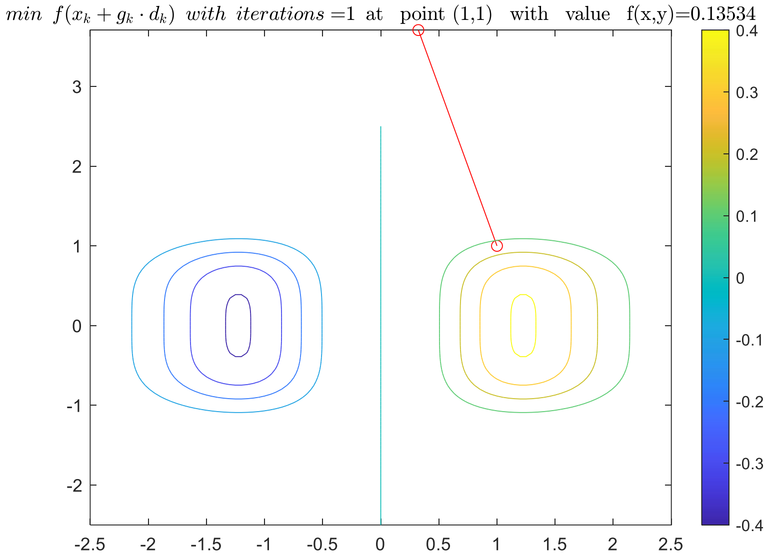
3. Αρχικό σημείο (1,1) για την συνάρτηση f

   a) Με σταθερό γ = 0.5



Για αρχικό σημείο (1,1) μετά από 120 επαναλήψεις καταλήγουμε στο σημείο (0.87751,1.7208) για δοσμένη ακρίβεια e=10^-3 με τιμή f(x, y)=4.869*e-05 που είναι το ολικό ελάχιστο.

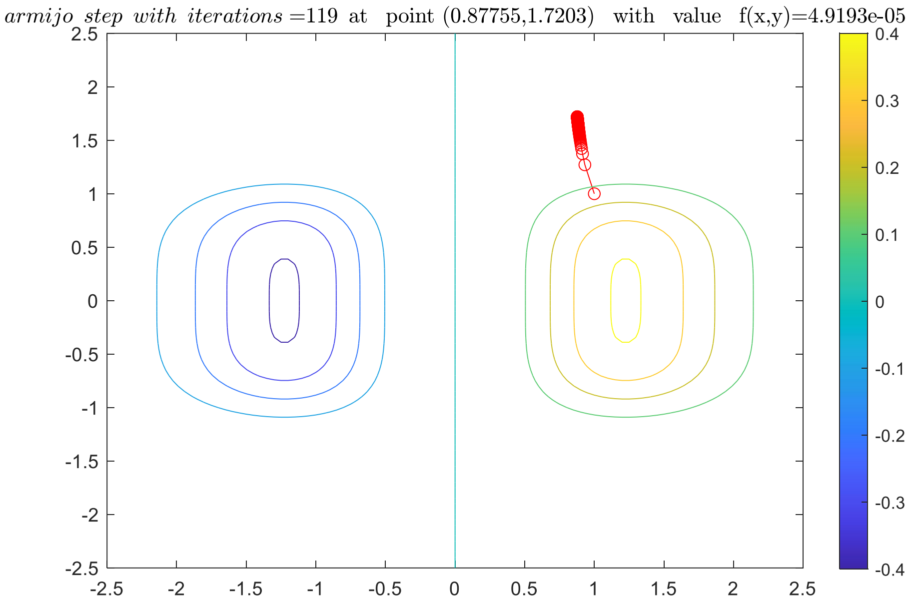
b) Για να ελαχιστοποιηθεί η συνάρτηση $h(\gamma)=f(xk+\gamma\kappa dk)$



Για αρχικό σημείο (1,1) μετά από 1 επανάληψη καταλήγουμε στο σημείο (1,1) για δοσμένη ακρίβεια e=10^-3 με τιμή f(x, y)=-0.13534 που είναι το ολικό ελάχιστο.

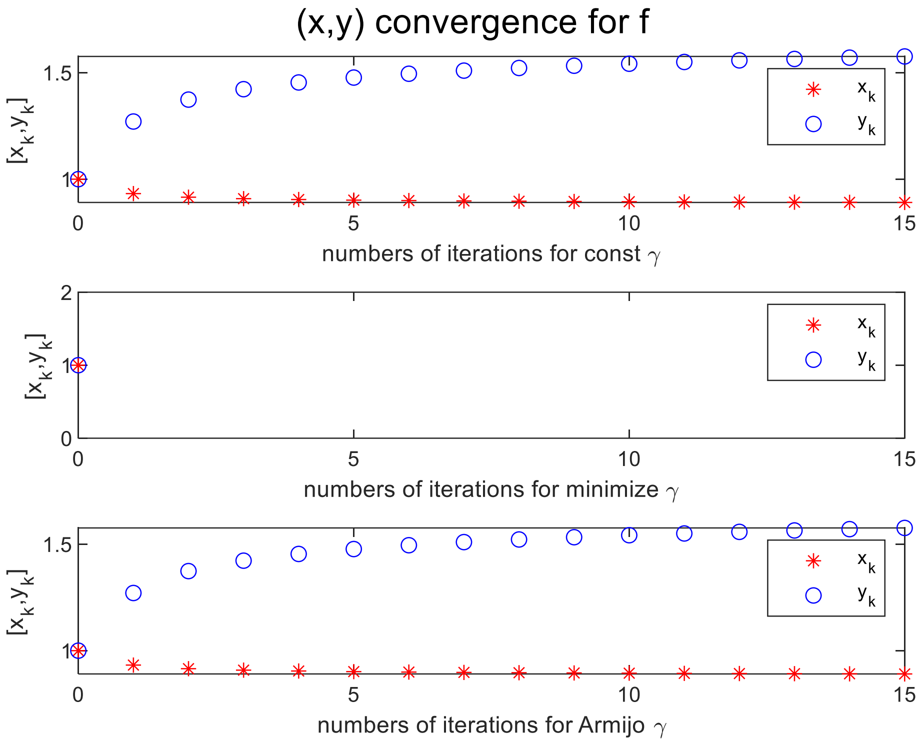
ΝΙΟΒΗ ΠΕΡΑΤΙΚΟΥ
ΑΕΜ:10494

c) Κανόνας Armijo



Για αρχικό σημείο (1,1) μετά από 119 επαναλήψεις καταλήγουμε στο σημείο (0.87755,1.7203) για δοσμένη ακρίβεια e=10^-3 με τιμή f(x, y)=4.9193*e^-05 που είναι το ολικό ελάχιστο.

Για να γίνει πιο σαφής η σύγκριση μεταξύ των διαφορετικών <<μορφών>> της Μέγιστης Καθόδου, χρησιμοποιώ τη steepest_descent_x_y_compare.m, όπου το ζεύγος (x , y), για κάθε μια από τις παραπάνω μορφές, σχεδιάζεται συναρτήσει των αριθμών επαναλήψεων.
Για το σημείο (1,1), έχουμε:



<u>Μέθοδος Newton:</u> Σε αυτή τη μέθοδο, προσεγγίζουμε την αντικειμενική συνάρτηση, με την τετραγωνική της μορφή, με σκοπό την σύγκλιση στο σημείο ελαχίστου σε ένα και μόνο βήμα.

Συγκεκριμένα, αν στην σχέση $x_{\kappa+1} = x_\kappa - \gamma_\kappa \Delta_\kappa \nabla f(x_\kappa)$, όπου Δ_κ θέσουμε τον αντίστροφο εσσιανό πίνακα

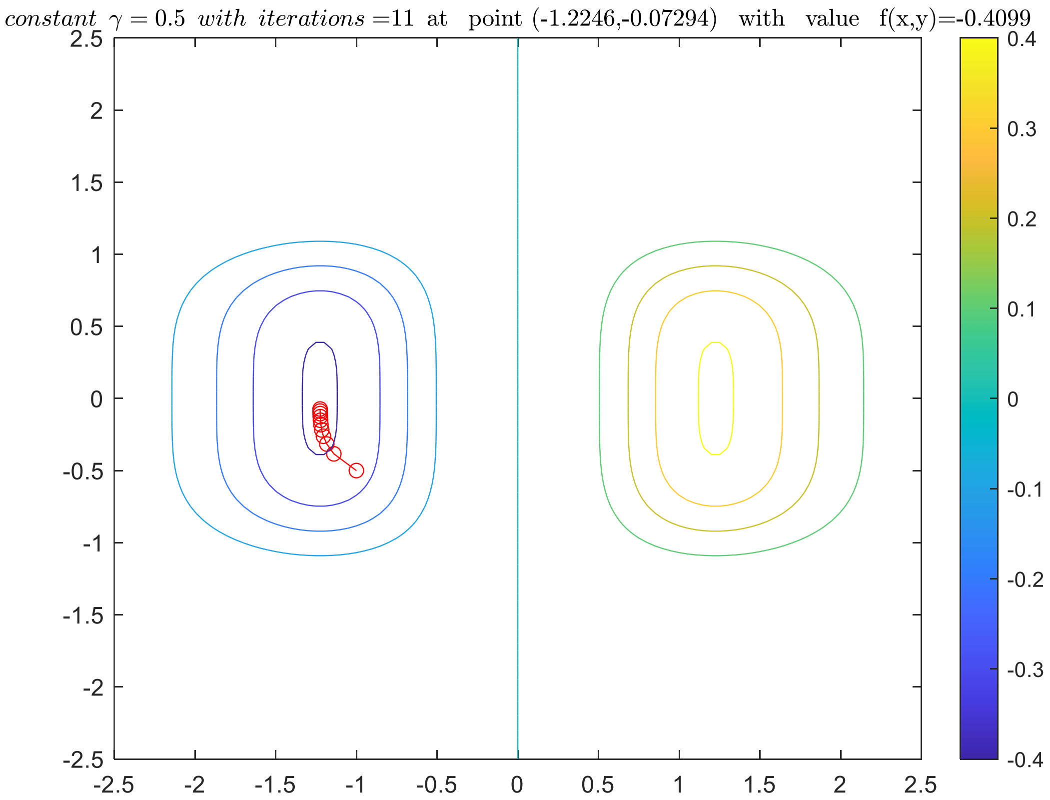
$$[\nabla^2 f(x_k)]^{-1} \text{ , θα έχουμε } x_{\kappa+1} = x_\kappa - \gamma_\kappa [\nabla^2 f(x_k)]^{-1} \nabla f(x_\kappa).$$

Η μέθοδος Newton, για να λειτουργήσει απαιτεί ο $\nabla^2 f(x_k)$ να είναι θετικά ορισμένος και να αντιστρέφεται.

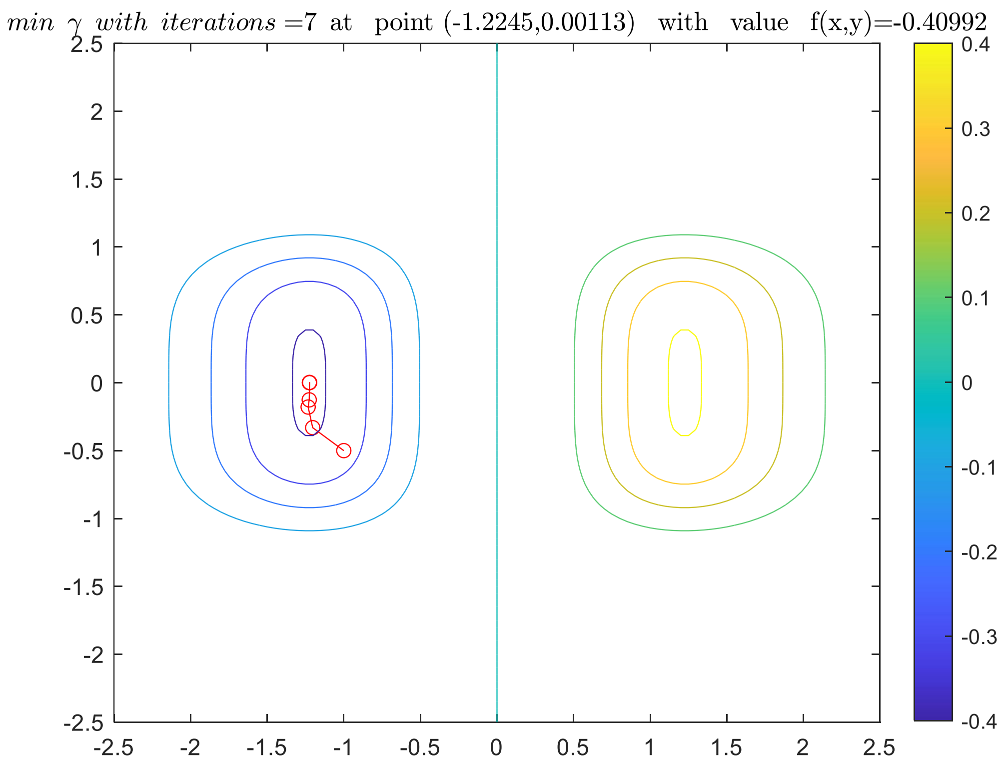
Για αρχικό σημείο (0,0) ο $\nabla^2 f(x_k)$ είναι αρνητικός ορισμένος και δεν τηρείται το κριτήριο, επίσης για αρχικό σημείο (-1,-1) ο $\nabla^2 f(x_k)$ είναι αρνητικός ορισμένος και δεν τηρείται το κριτήριο. Τέλος για αρχικό σημείο (1,1) ο $\nabla^2 f(x_k)$ είναι αρνητικός ορισμένος και δεν τηρείται ούτε εδώ το κριτήριο. Άρα δεν μπορεί να υλοποιηθεί η μέθοδος Newton.

Έτσι για να μπορούμε να έχουμε κάποιο αποτέλεσμα στον κώδικά μας επιλέγουμε αυθαίρετα ως αρχικό σημείο το (-1,-0.5) με το οποίο βλέπουμε ότι τηρείται το κριτήριο και έχουμε τα εξής αποτελέσματα:

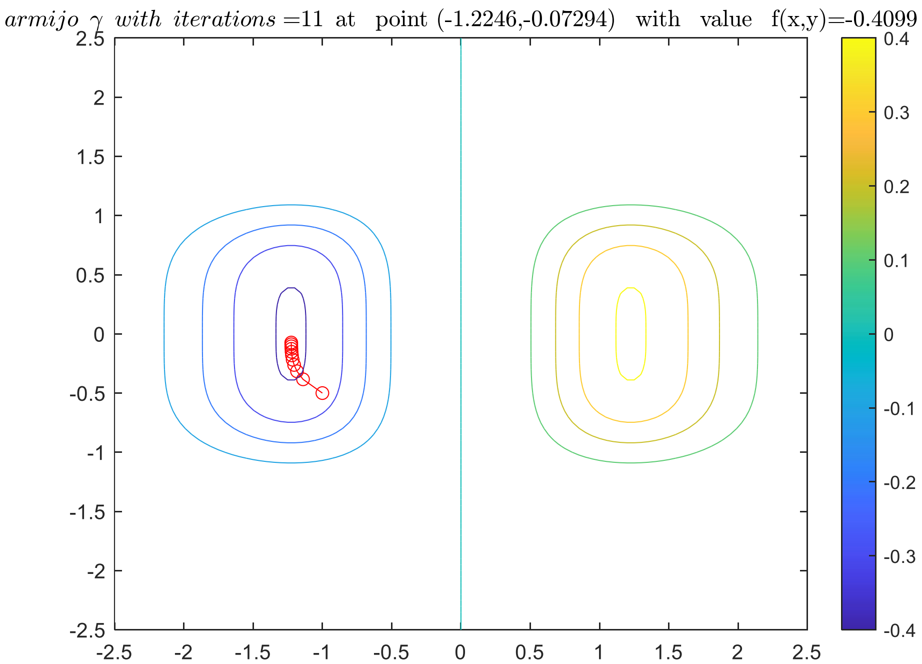
I.Κρατώντας σταθερό το ε=0.001 και επιλέγοντας το ίδιο γ με πριν, έχω:



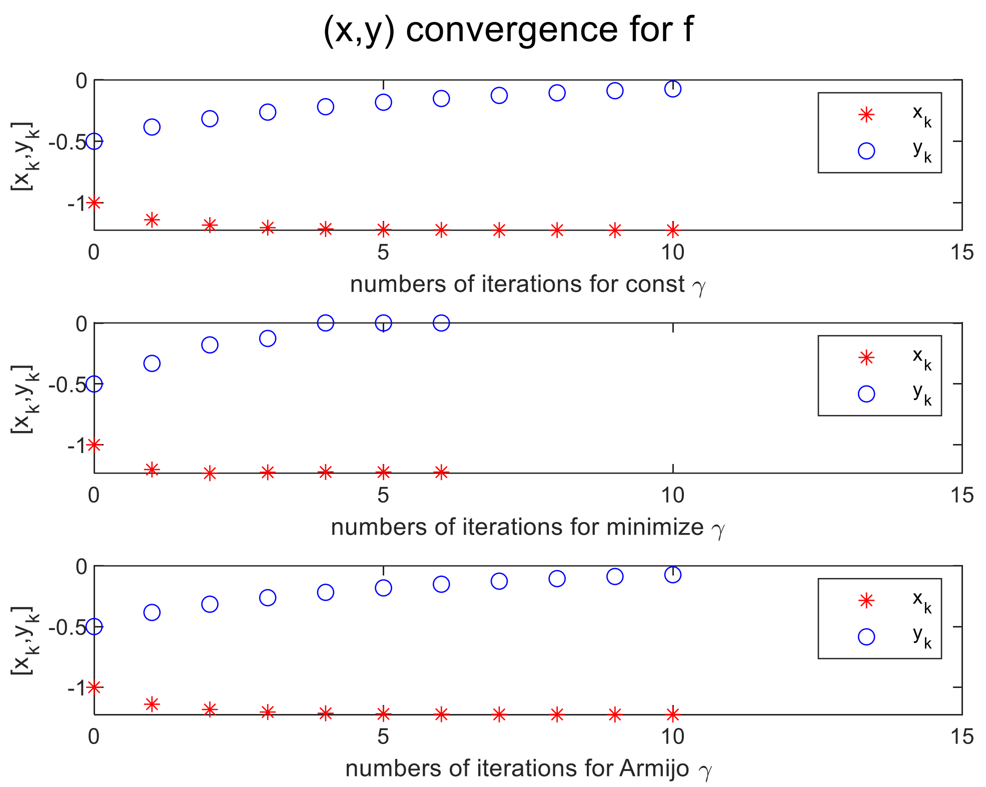
II. Εφαρμόζοντας την ελαχιστοποίηση για το γ, πετυχαίνουμε τον στόχο μας. Συγκλίνουμε στο ελάχιστο, σε μία και μόνο επανάληψη.



III. Και τέλος ο Armijo κανόνας.



Για να γίνει πιο σαφής η σύγκριση μεταξύ των διαφορετικών <<μορφών>> της μεθόδου Newton, χρησιμοποιώ το newton_plot_x_y_compare.m όπου το ζεύγος (x,y), για κάθε μια από τις παραπάνω μορφές, σχεδιάζεται συναρτήσει των αριθμών επαναλήψεων.

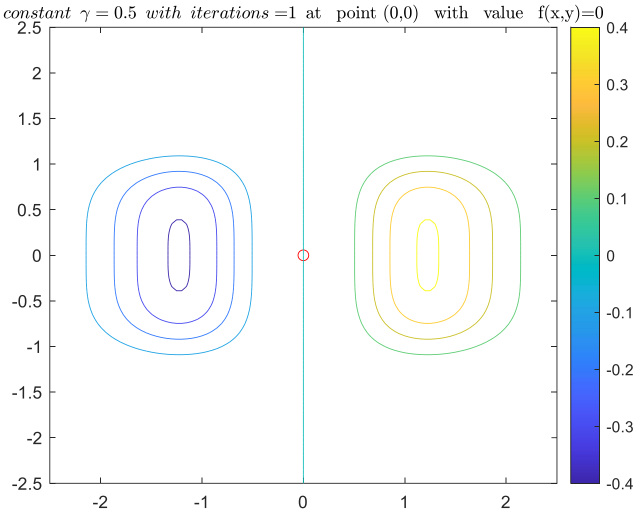


ΝΙΟΒΗ ΠΕΡΑΤΙΚΟΥ
ΑΕΜ:10494

Μέθοδος Levenberg-Marquardt: Σε αυτή τη μέθοδο αν ο $\nabla 2f(xk)$ δεν είναι θετικά ορισμένος τότε μπορούμε να τροποποιήσουμε τον αλγόριθμο Newton.

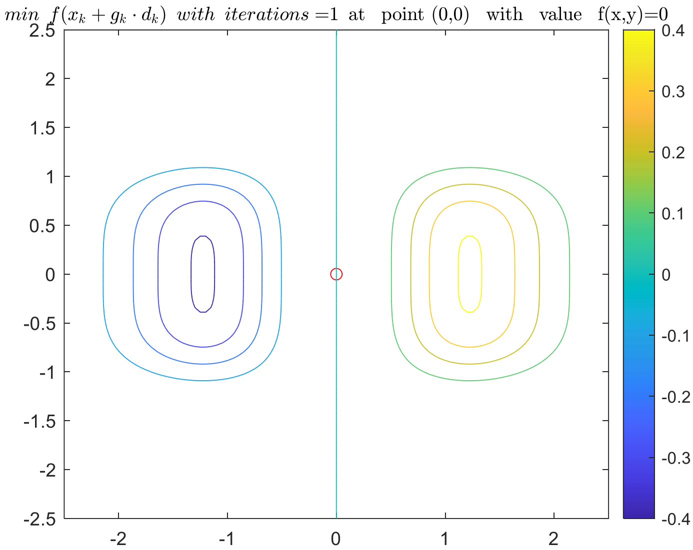
Η μέθοδος Levenberg-Marquardt, για να λειτουργήσει δεν απαιτεί ο $\nabla 2f(xk)$ να είναι θετικά ορισμένος.

1. Αρχικό σημείο (0,0) για την συνάρτηση f
   a) Με σταθερό γ = 0.5



Για αρχικό σημείο (0,0) δεν ικανοποιούνται οι προδιαγραφές των Κριτηρίων 3 και 4.

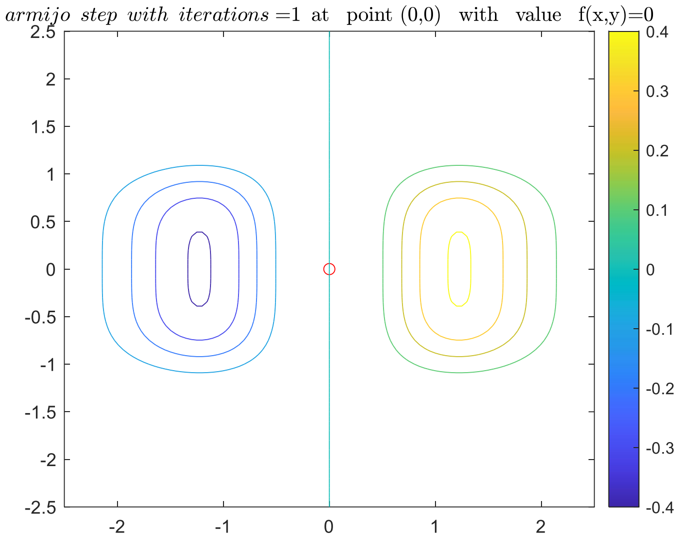
   b) Για να ελαχιστοποιηθεί η συνάρτηση $h(\gamma)=f(xk+\gamma\kappa dk)$



Για αρχικό σημείο (0,0) δεν ικανοποιούνται οι προδιαγραφές των Κριτηρίων 3 και 4.

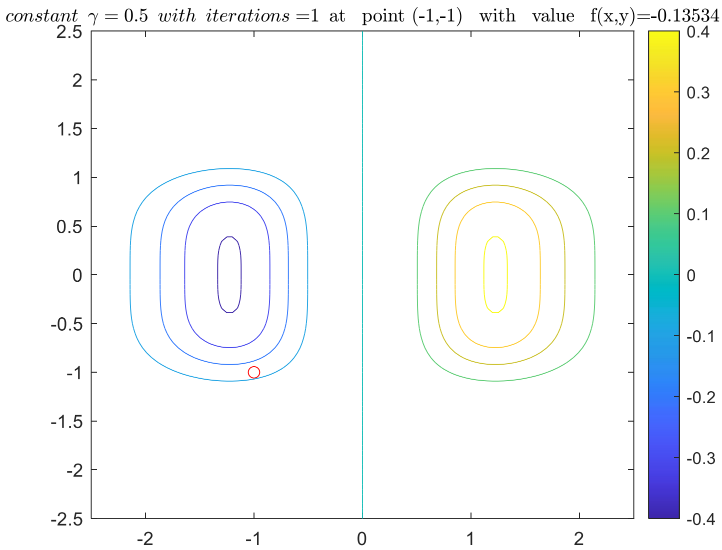
ΝΙΟΒΗ ΠΕΡΑΤΙΚΟΥ
ΑΕΜ:10494

c) Κανόνας Armijo



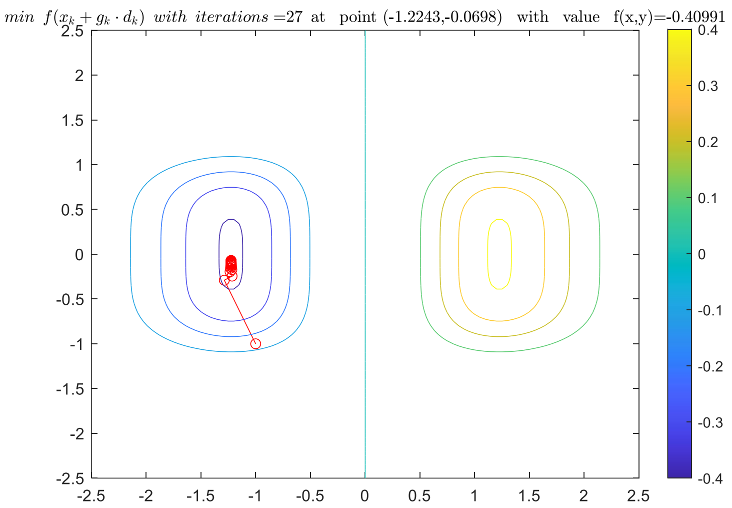
Για αρχικό σημείο (0,0) δεν ικανοποιούνται οι προδιαγραφές των Κριτηρίων 3 και 4.

2. Αρχικό σημείο (-1,-1) για την συνάρτηση f
   a) Με σταθερό γ=0.5



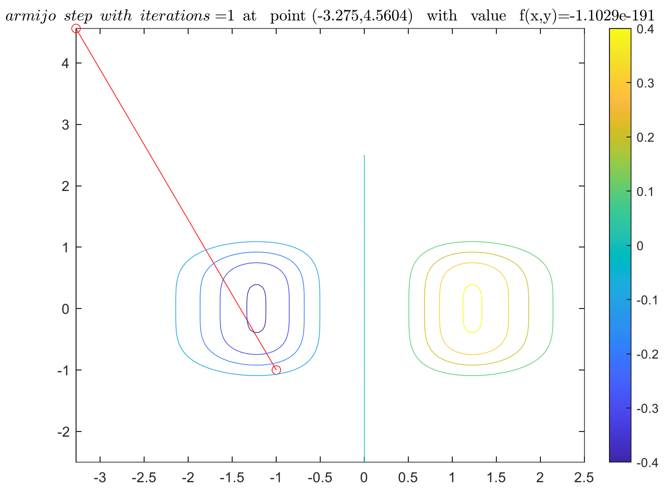
Για αρχικό σημείο (-1,-1) δεν ικανοποιούνται οι προδιαγραφές των Κριτηρίων 3 και 4.

b) Για να ελαχιστοποιηθεί η συνάρτηση $h(\gamma)=f(xk+\gamma\kappa dk)$



Για αρχικό σημείο (-1,-1) μετά από 27 επαναλήψεις καταλήγουμε σε τοπικό ελάχιστο (-1.2243,-0.0698) για δοσμένη ακρίβεια e=10^-3 με τιμή f(x,y)=-0.40991. Αποτέλεσμα λογικό αφού με τόσο μικρό βήμα είναι αδύνατον να βρεθούμε στην απέναντι περιοχή του ολικού ελαχίστου χωρίς να εγκλωβιστούμε.

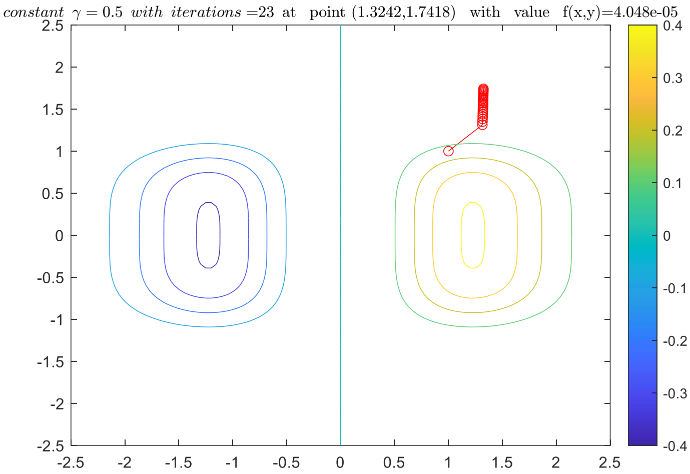
c) Κανόνας Armijo



Για αρχικό σημείο (-1,-1) μετά από 1 επανάληψη καταλήγουμε στο σημείο (-3.275,4.5604) για δοσμένη ακρίβεια e=10^-3 με τιμή f(x, y)=-1.1029*e-191 ο αλγόριθμος αποκλίνει.

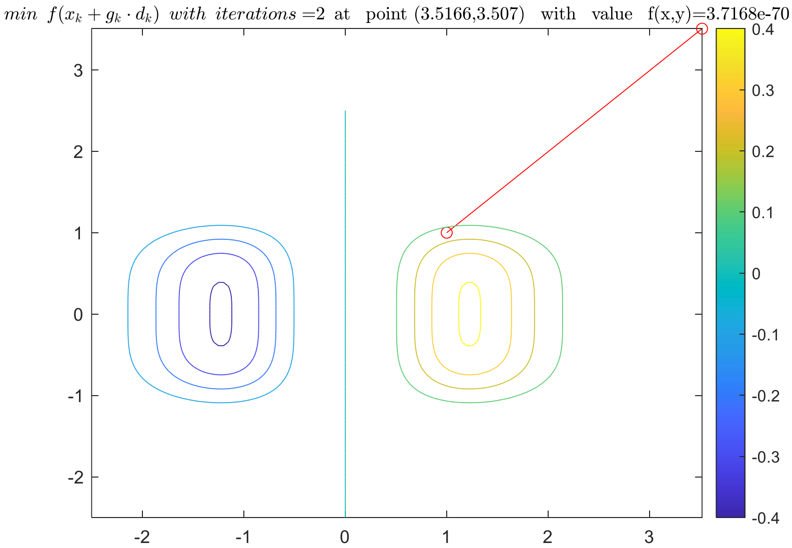
ΝΙΟΒΗ ΠΕΡΑΤΙΚΟΥ
AEM:10494

3. Αρχικό σημείο (1,1) για την συνάρτηση f
   a) Σταθερό γ=0.5



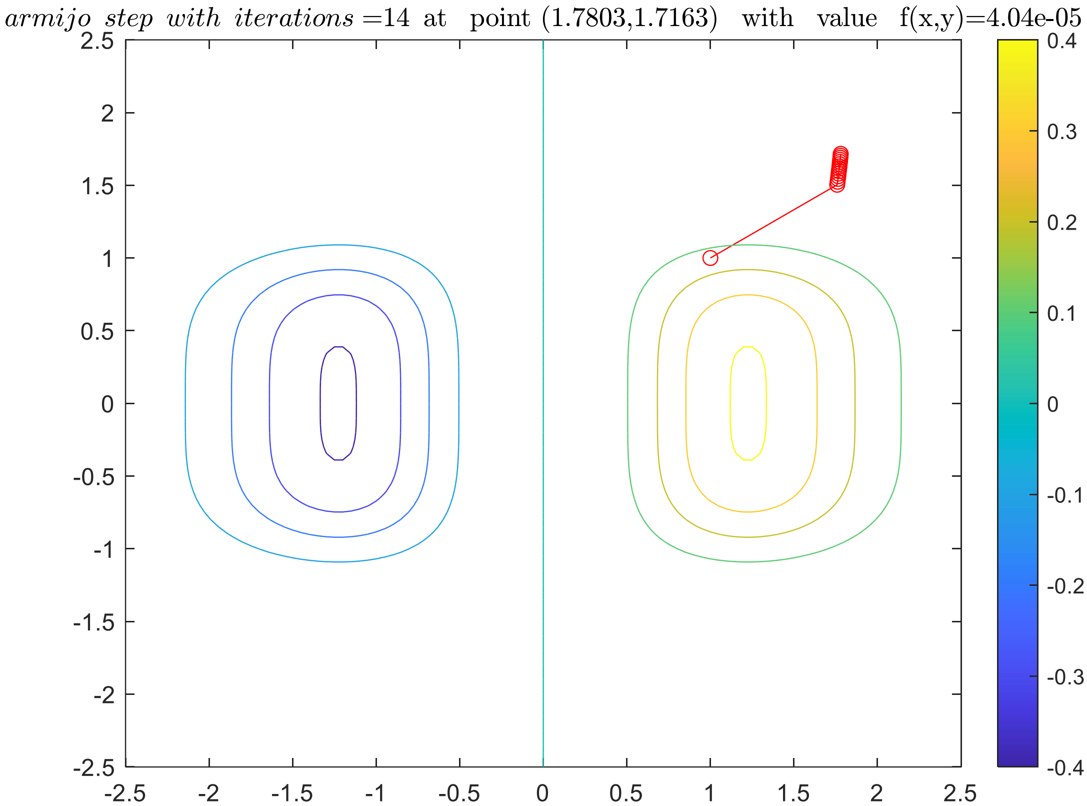
Για αρχικό σημείο (1,1) μετά από 23 επαναλήψεις καταλήγουμε σε τοπικό ελάχιστο (1.3242,1.7418) για δοσμένη ακρίβεια e=10^-3 με τιμή f(x, y)=4,048·10−5. Αποτέλεσμα λογικό αφού με τόσο μικρό βήμα είναι αδύνατον να βρεθούμε στην απέναντι περιοχή του ολικού ελαχίστου χωρίς να εγκλωβιστούμε.

   b) Ελαχιστοποίηση της $h(\gamma)=f(xk+\gamma\kappa dk)$



Για αρχικό σημείο (1,1) μετά από 2 επαναλήψεις καταλήγουμε σε τοπικό ελάχιστο (3.5166,3.507) για δοσμένη ακρίβεια e=10^-3 με τιμή f(x, y)=3,7168·10−70. Αποτέλεσμα λογικό αφού με τόσο μικρό βήμα είναι αδύνατον να βρεθούμε στην απέναντι περιοχή του ολικού ελαχίστου χωρίς να εγκλωβιστούμε.

c) Κανόνας Armijo.



Για αρχικό σημείο (1,1) μετά από 14 επαναλήψεις καταλήγουμε σε τοπικό ελάχιστο (1.7803,1.7163) για δοσμένη ακρίβεια e=10^-3 με τιμή f(x, y)=4.04·e^−05. Αποτέλεσμα λογικό αφού με τόσο μικρό βήμα είναι αδύνατον να βρεθούμε στην απέναντι περιοχή του ολικού ελαχίστου χωρίς να εγκλωβιστούμε.

Συμπεράσματα:

Παρατηρώντας, τα παραπάνω διαγράμματα η μέθοδος Newton είναι σε γενικές γραμμές καλύτερη από τη μέθοδο της μέγιστης καθόδου. Ωστόσο, και οι δύο μέθοδοι έχουν σημαντικές αδυναμίες. Από την μία στην μέγιστη κάθοδο, ο αλγόριθμος καθυστερεί στην σύγκλισή του και από την άλλη η Newton, χρειάζεται θετικά ορισμένο τον εσσιανό πίνακα, κάτι που δεν μπορεί να εγγυηθεί πως θα ισχύει σε κάθε επανάληψη, όπου αυτό το πρόβλημα προσπαθεί να επιλυθεί εν μέρος από την μέθοδο Levenberg-Marguardt όπου όμως βάζει κάποιους άλλους περιορισμούς. Τέλος και οι τρεις μέθοδοι έχουν τοπικό χαρακτήρα, εγκλωβίζονται στην γειτονιά του ελαχίστου και έτσι δεν υπάρχει εγγύηση ότι το ελάχιστο που βρέθηκε είναι το ολικό.