Δοϊνάκης Μιχαήλ
AEM:9292
e-mail: doinakis@ece.auth.gr

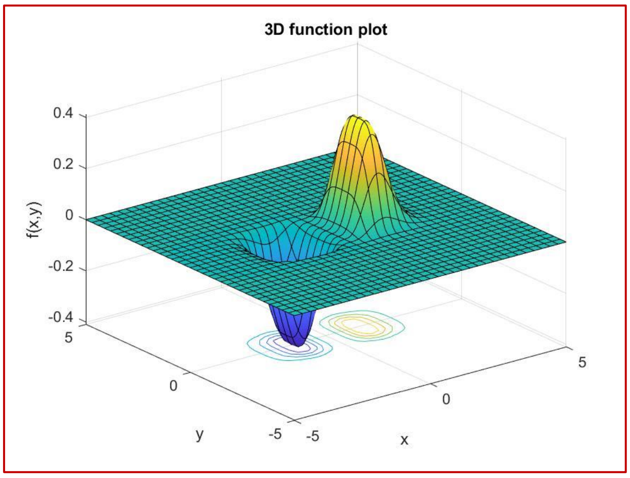
ΕΙΣΑΓΩΓΗ

Στόχος της δεύτερης εργασίας είναι η ελαχιστοποίηση δοσμένης συνάρτησης πολλών μεταβλητών, στη προκειμένη περίπτωση στο δισδιάστατο χώρο, χωρίς περιορισμούς. Υλοποιήθηκαν τρεις μέθοδοι ελαχιστοποίησης, η μέθοδος της μέγιστης καθόδου (Steepest Descent), η μέθοδος Newton και η μέθοδος Levenberg-Marquardt, όπου για κάθε μία μελετήθηκαν διάφοροι τρόποι επιλογής της παραμέτρου γ_κ αλλά και διαφορετικά σημεία εκκίνησης. Πιο συγκεκριμένα υλοποιήθηκαν τρεις συναρτήσεις με τα αντίστοιχα ονόματα οι οποίες επιστρέφουν τον πίνακα με τα σημεία χ_κ που υπολογίζονται σε κάθε επανάληψη (το τελευταίο σημείο είναι το ελάχιστο) και τον αριθμό των επαναλήψεων για κάθε υλοποίηση. Τα κατώφλι τερματισμού (epsilon = η τιμή για την οποία το gradient της συνάρτησης στο σημείο κ θεωρείται μηδέν) τέθηκε στο 0.01 και η αρχική τιμή για το γ_κ είναι το 0.5. Τα σημεία για τα οποία θα πραγματοποιηθεί η μελέτη είναι τα $(0,0), (-1,-1), (1,1)$ και οι τιμές του γ_κ επιλέγεται ως σταθερή(μέθοδος α), έτσι ώστε να ελαχιστοποιεί την $f(\chi_\kappa + \gamma_\kappa * d_k)$ (μέθοδος β) και με τη μέθοδο του Armijo(μέθοδος γ)(επιλέχθηκαν οι τιμές $\alpha = 0.001 \; \kappa\alpha\iota \; \beta = \frac{1}{5}$). Για την ελαχιστοποίηση της συνάρτησης και την εύρεση του γ_κ χρησιμοποιήθηκε η συνάρτηση bisection_with_derivatives που είχε υλοποιηθεί στην πρώτη εργασία.

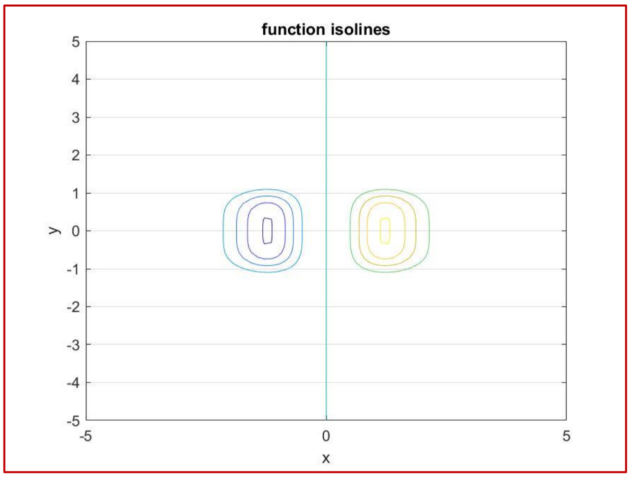
Επίσης υλοποιήθηκαν και τέσσερα matlab scripts **Subject1st**, **Subject2nd**, **Subject3rd**, **Subject4th**, από τα οποία το πρώτο είναι αυτό που δίνει την αναπαράσταση της συνάρτησης και τα υπόλοιπα τρία εφαρμόζουν τις μεθόδους της μέγιστης καθόδου (Steepest Descent), η μέθοδος Newton και η μέθοδος Levenberg-Marquardt αντίστοιχα.

ΘΕΜΑ 1: ΑΝΑΠΑΡΑΣΤΑΣΗ ΤΗΣ ΣΥΝΑΡΤΗΣΗΣ

Παρακάτω παρουσιάζεται η μορφή της συνάρτησης στον τρισδιάστατο χώρο αλλά και οι ισοβαρείς καμπύλες της. Από τη γραφική παράσταση της συνάρτησης μπορεί να εξαχθεί το συμπέρασμα ότι η ελάχιστη τιμή βρίσκεται σε αρνητικά x και y. Επιπλέον παρουσιάζονται οι ισοβαρείς καμπύλες οι οποίες θα χρησιμοποιηθούν παρακάτω για να παρουσιαστεί η πορεία που ακολουθεί ο κάθε αλγόριθμος έως ότου τερματιστεί. Η συνάρτηση η οποία παρουσιάζεται είναι η εξής: $f(x,y) = x^3 e^{(-x^2-y^4)}$.



1. Γραφική παράσταση της f (με χρήση fsurf).

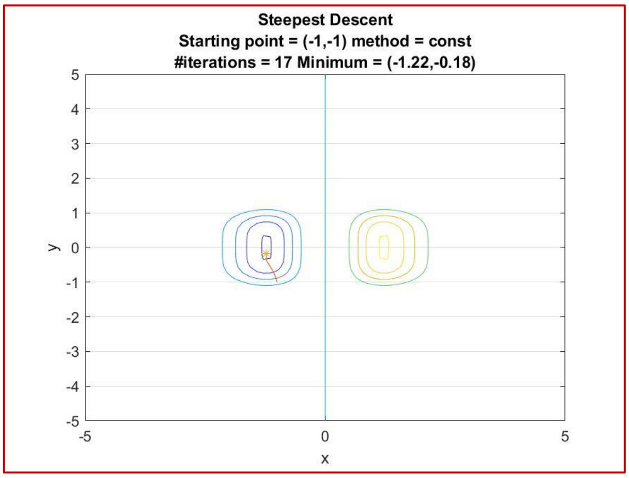


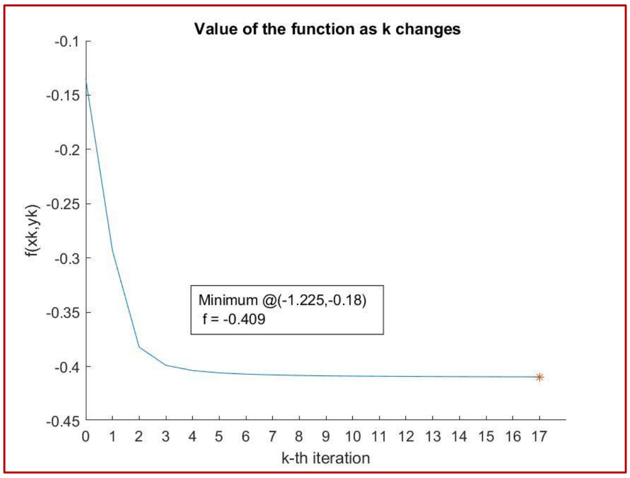
2. Ισοβαρείς καμπύλες της f (με χρήση της fcontour).

ΘΕΜΑ 2: ΜΕΘΟΔΟΣ ΜΕΓΙΣΤΗΣ ΚΑΘΟΔΟΥ

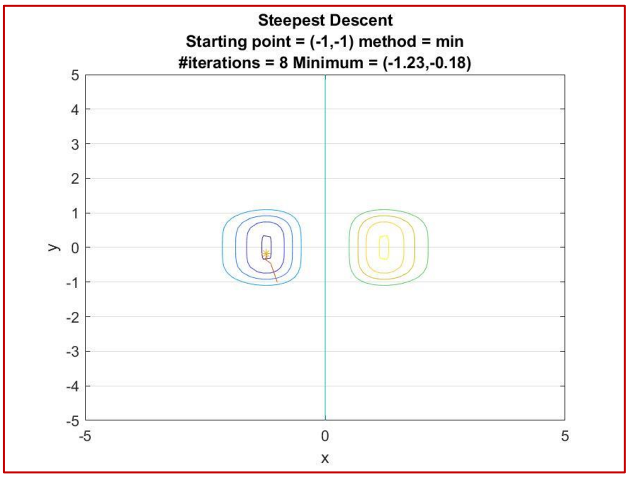
Στα παρακάτω διαγράμματα φαίνεται η πορεία που ακολουθείται για την εύρεση του ελαχίστου της συνάρτησης σε σχέση με τα διάφορα σημεία εκκίνησης αλλά και το πώς συγκλίνει η συνάρτηση στο ελάχιστο με τη πάροδο της κάθε επανάληψης. Ως μηδενική επανάληψη θεωρείται αυτή για το σημείο εκκίνησης. Όταν, λοιπόν, ως αρχικό σημείο τεθεί το $(0,0)$, η μέθοδος της μέγιστης καθόδου εγκλωβίζεται σε αυτό, διότι είναι τοπικό ελάχιστο της συνάρτησης σε εκείνη την περιοχή και δεν εκτελεί έτσι καμία επανάληψη (για το λόγο αυτό δεν παρουσιάζονται και διαγράμματα για το σημείο αυτό).

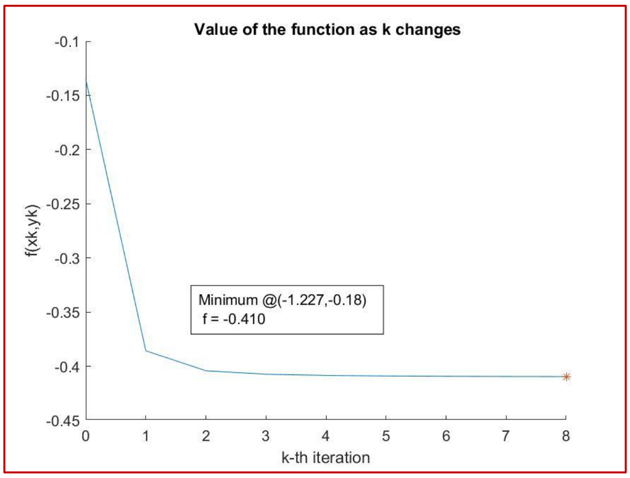
Συνεχίζοντας τη μελέτη όταν ως αρχικό σημείο τεθεί το $(-1,-1)$, παρατηρείται ότι ο αλγόριθμος φτάνει στο ελάχιστο σημείο (πολύ κοντά σε αυτό ανάλογα με το epsilon που έχει επιλεχτεί) και με τις τρεις μεθόδους επιλογής του γ_κ. Από τα διαγράμματα της σύγκλισης της συνάρτησης συναρτήσει του αριθμού της επανάληψης φαίνεται ότι η μέθοδος β είναι αρκετά γρηγορότερη σε σχέση με αυτή του σταθερού γ_κ, γεγονός που είναι αναμενόμενο καθώς η πρώτη κινείται σε κάθετα μεταξύ τους ευθύγραμμα τμήματα. Ήταν επιπλέον γρηγορότερη και από την μέθοδο Armijo.



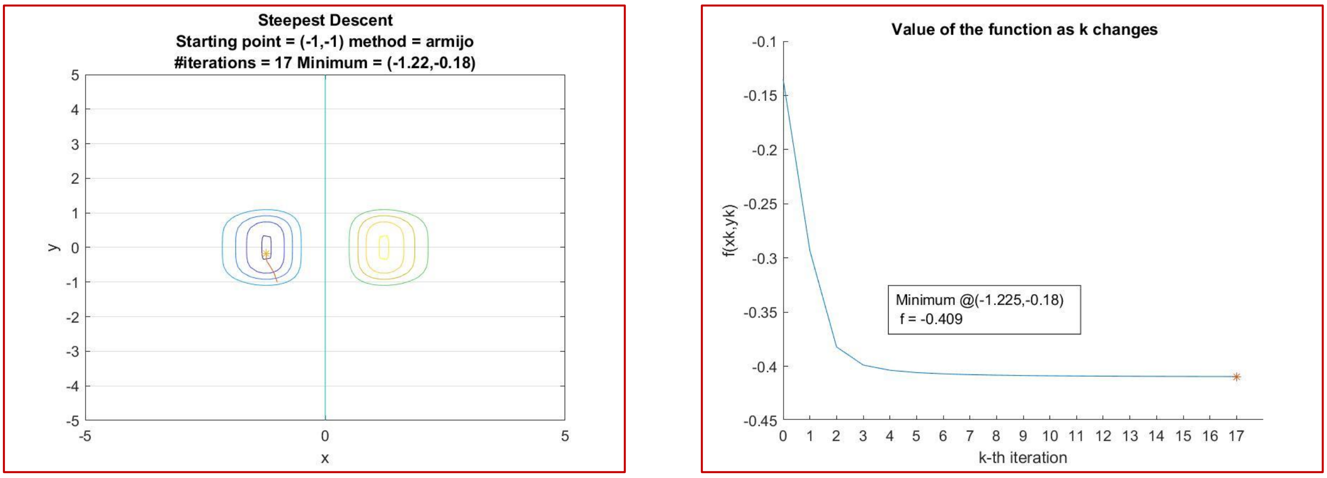


3. Σταθερό γ_κ και σημείο εκκίνησης (-1,-1).



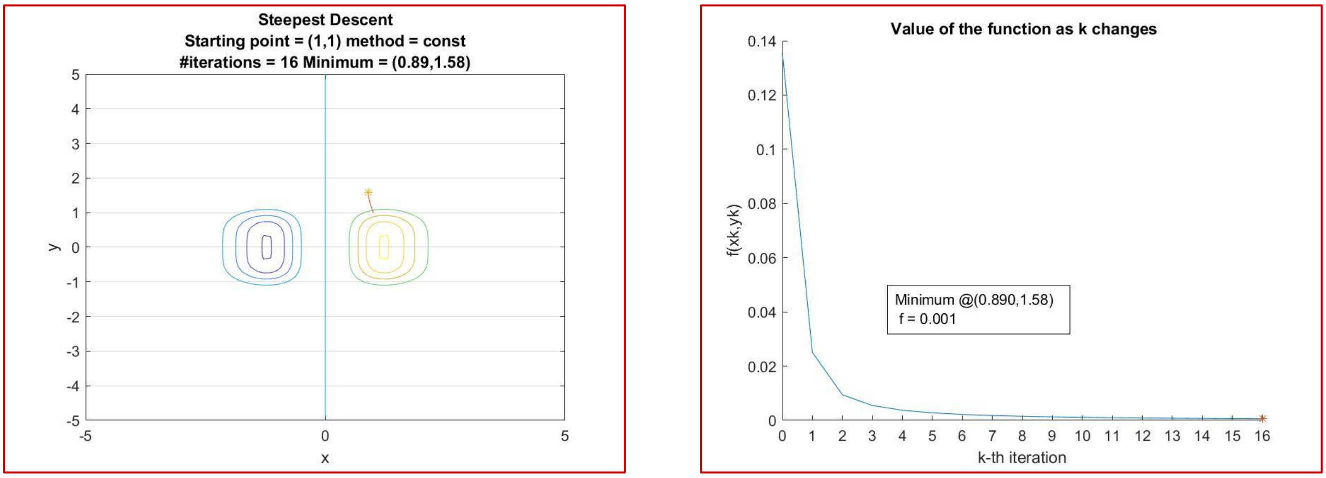


4. γ_κ που ελαχιστοποιεί τη συνάρτηση $f(\chi_\kappa + \gamma_\kappa * d_k)$ και σημείο εκκίνησης το (-1,-1).

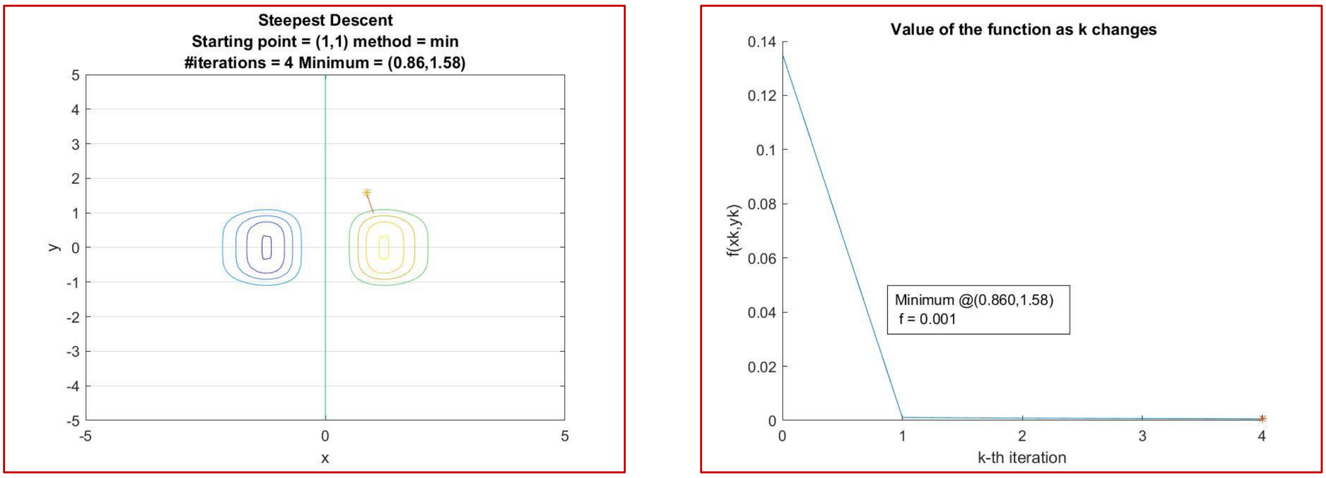


5. γ_κ υπολογισμένο με τη μέθοδο armijo και σημείο εκκίνησης (-1,-1).

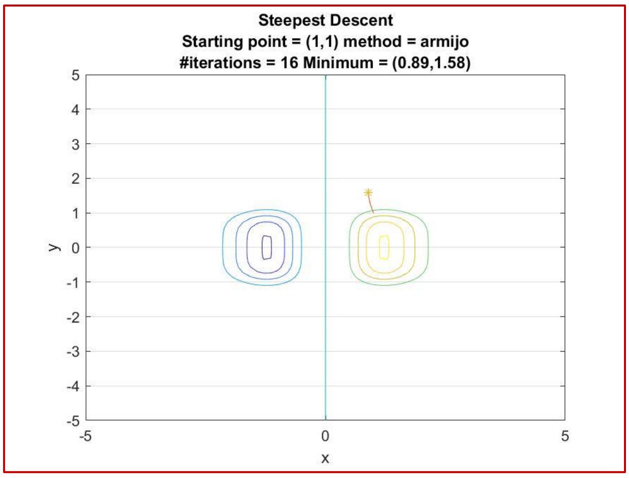
Με αρχικό σημείο το $(1,1)$, η μέθοδος εγκλωβίζεται πάλι όπως και με το $(0,0)$, σε τοπικό ελάχιστο και αδυνατεί να βρει το πραγματικό. Εξάγεται, λοιπόν το συμπέρασμα πως η επιλογή του σημείου έναρξης του αλγορίθμου παίζει καθοριστικό ρόλο στο αν τελικά ο αλγόριθμος θα καταλήξει να βρει το ολικό ελάχιστο της συνάρτησης. Σε ότι αφορά την ταχύτητα της σύγκλισης ανάλογα με τη μέθοδο επιλογής του γ_κ ισχύει ότι ειπώθηκε και για το σημείο $(-1,-1)$.

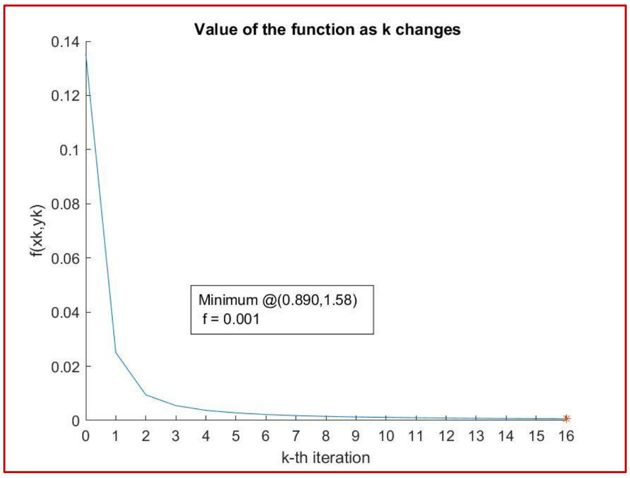


6. Σταθερό γ_κ και σημείο εκκίνησης (1,1).



7. γ_κ που ελαχιστοποιεί τη συνάρτηση $f(\chi_\kappa + \gamma_\kappa * d_k)$ και σημείο εκκίνησης το (1,1).



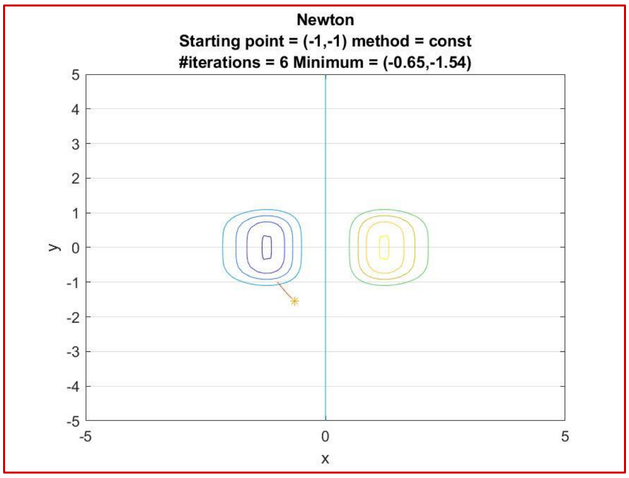


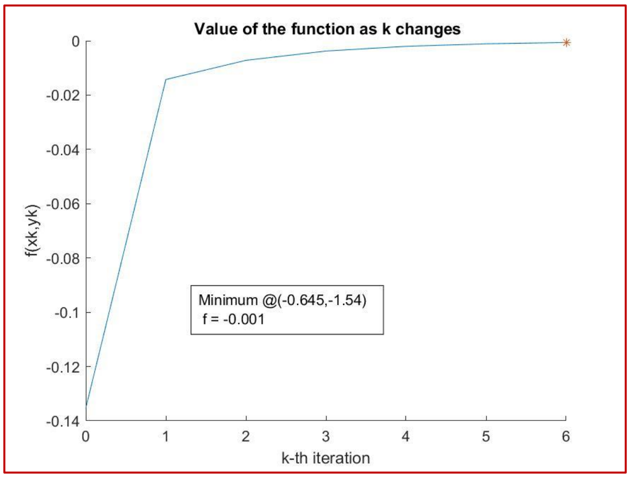
8. γ_κ υπολογισμένο με τη μέθοδο armijo και σημείο εκκίνησης (1,1).

ΘΕΜΑ 3: ΜΕΘΟΔΟΣ ΤΟΥ NEWTON

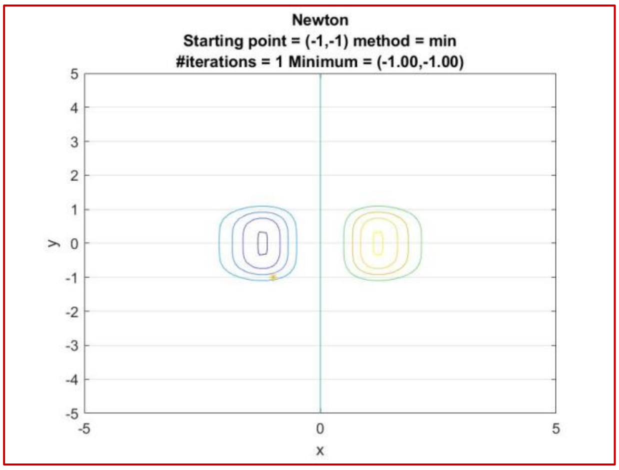
Η υλοποίηση της μεθόδου του Newton παρουσιάζεται στα παρακάτω διαγράμματα, με την ίδια λογική όπως και στην προηγούμενη ενότητα. Στο σημείο $(0,0)$ η συνάρτηση εγκλωβίζεται ξανά σε τοπικό ελάχιστο (αφού και εδώ η κλίση της συνάρτησης στο σημείο είναι 0).

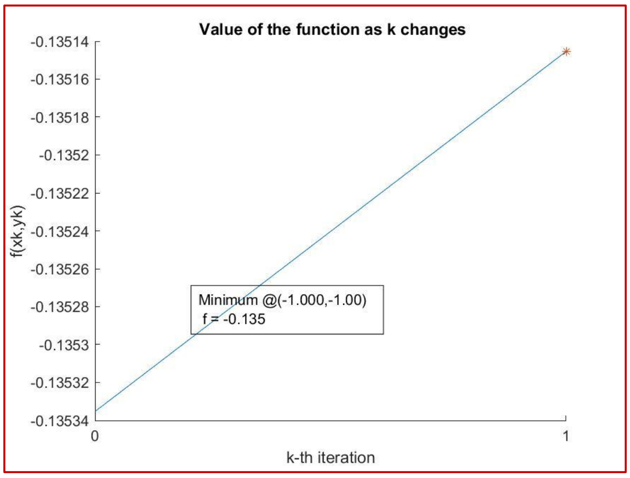
Περισσότερο ενδιαφέρον παρουσιάζουν τα διαγράμματα για σημείο εκκίνησης το $(-1,-1)$. Η μέθοδος του Newton εμφανίζει σημαντικό πρόβλημα όταν ο εσσιανός πίνακας της f(x) δεν είναι θετικά ορισμένος ή ακόμα και αντιστρέψιμος. Για της μεθόδους β και γ τέθηκε ως κριτήριο ο αλγόριθμος να σταματάει όταν η τιμή της συνάρτησης για χ_κ που υπολογίστηκε είναι μεγαλύτερη ή ίση με την τιμή της συνάρτησης στο $\chi_{\kappa-1}$. Σε διαφορετική περίπτωση, ο αλγόριθμος δεν συγκλίνει ποτέ σε σημείο ελαχίστου. Επιπλέον, στη μέθοδο α όπου το γ_κ παραμένει σταθερό αντί να η συνάρτηση να ελαχιστοποιείται, η τιμή της στο τελικό σημείο έχει μεγαλύτερη τιμή από την τιμή στο σημείο έναρξης.



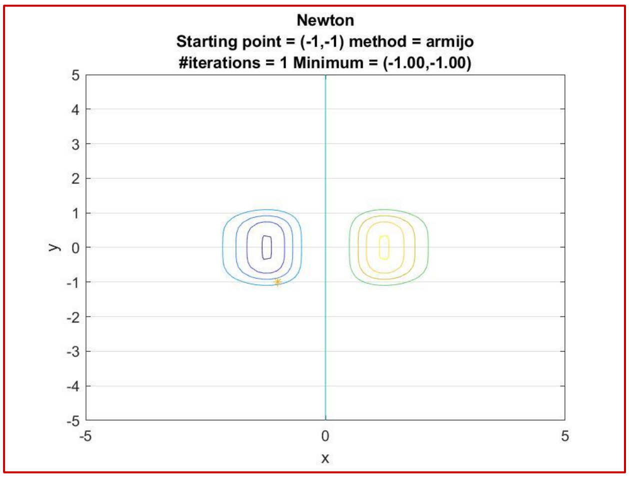


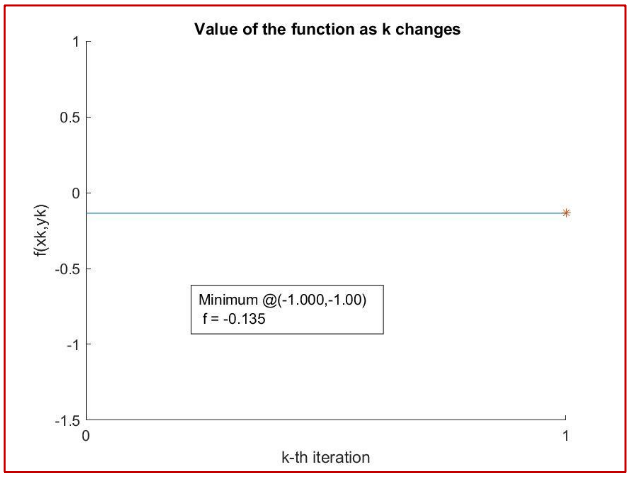
9. Σταθερό γ_κ και σημείο εκκίνησης (-1,-1).





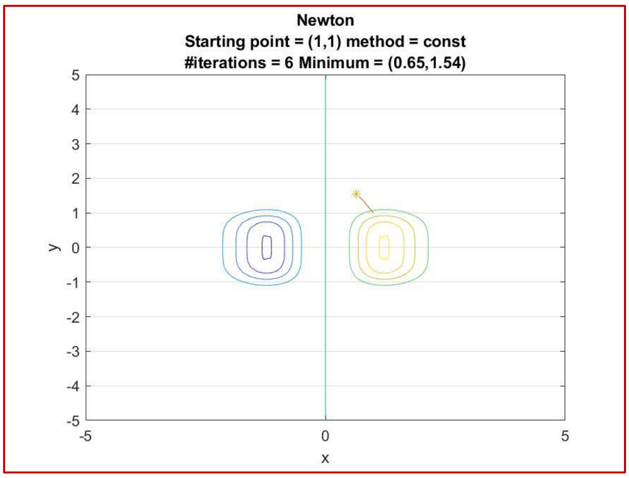
10. γ_κ που ελαχιστοποιεί τη συνάρτηση $f(\chi_\kappa + \gamma_\kappa * d_k)$ και σημείο εκκίνησης το (-1,-1).

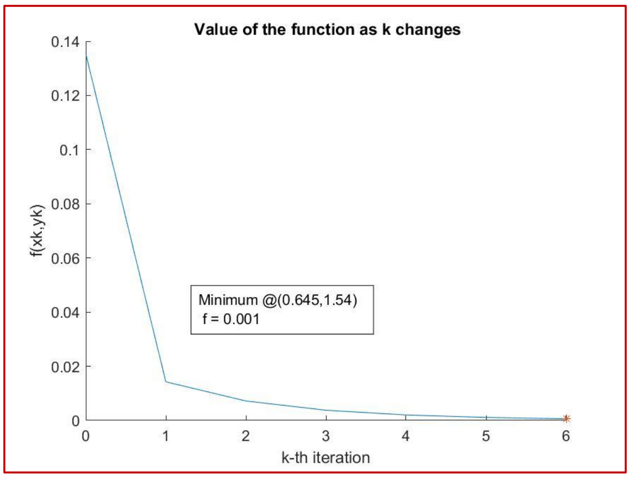




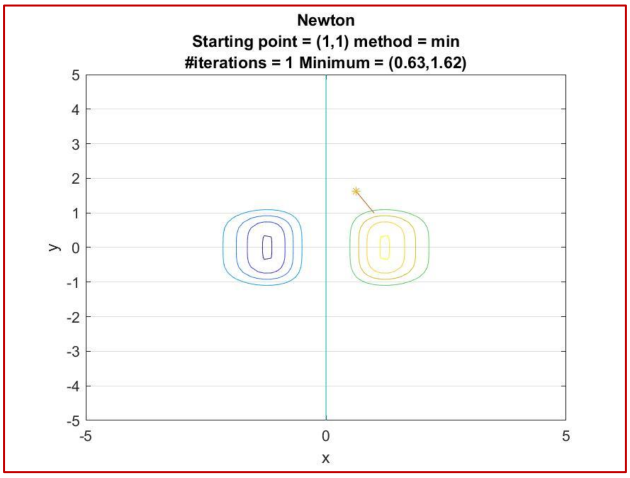
11. γ_κ υπολογισμένο με τη μέθοδο armijo και σημείο εκκίνησης (-1,-1).

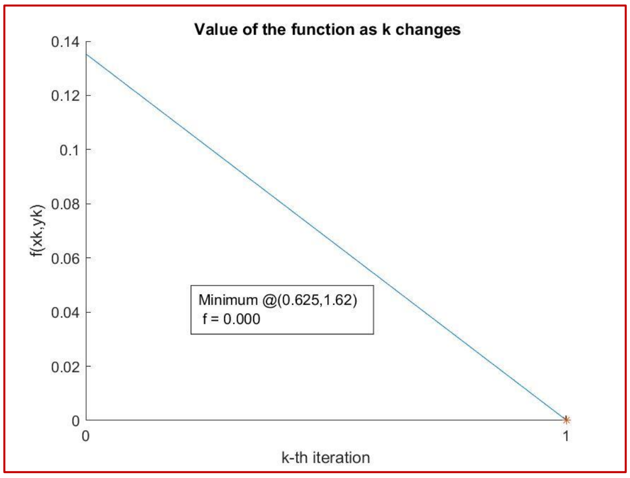
Για το σημείο (1,1) ο αλγόριθμος έχει παρόμοια αποτελέσματα με αυτόν τη μέγιστης καθόδου, δηλαδή δεν βρίσκει το ολικό ελάχιστο αλλά εγκλωβίζεται σε τοπικό.



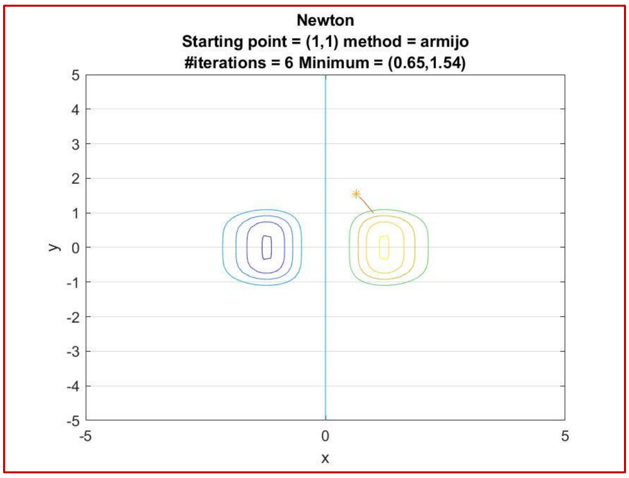


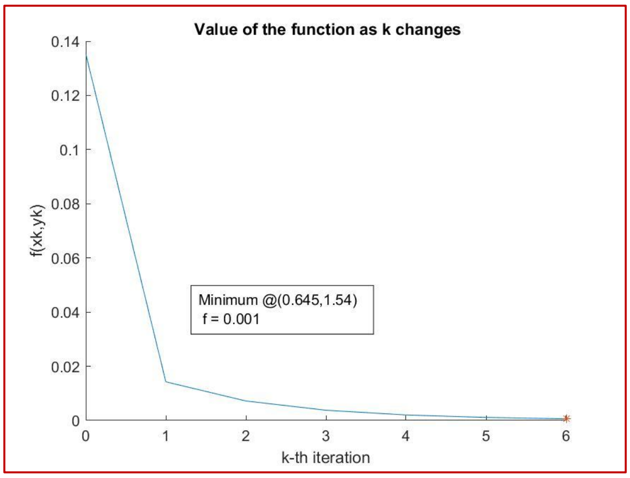
12. Σταθερό γ_κ και σημείο εκκίνησης (1,1).





13. γ_κ που ελαχιστοποιεί τη συνάρτηση $f(\chi_\kappa + \gamma_\kappa * d_k)$ και σημείο εκκίνησης το $(1,1)$.





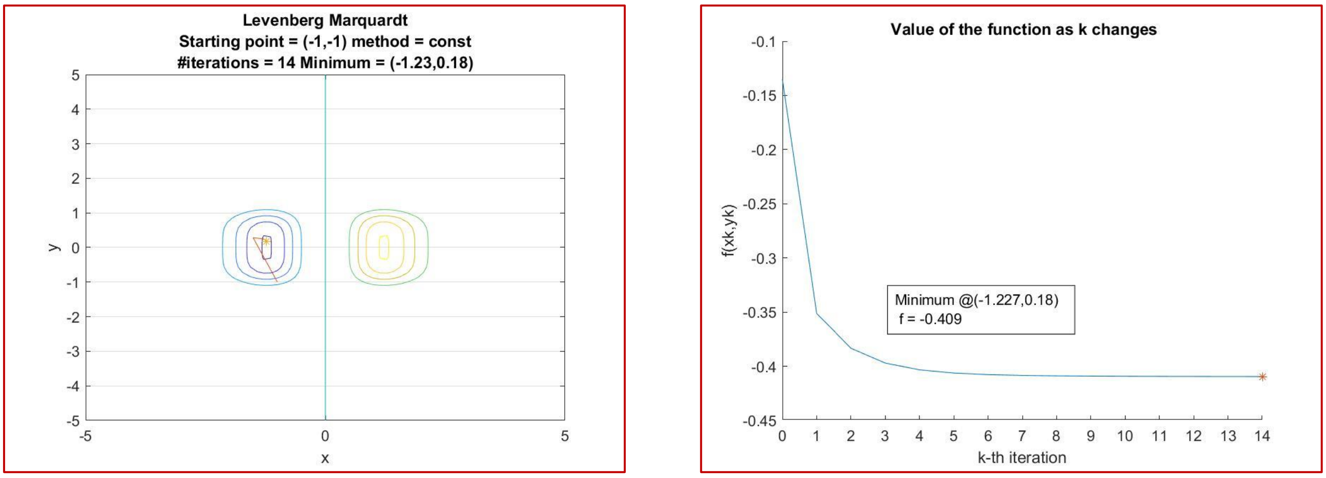
14. γ_κ υπολογισμένο με τη μέθοδο armijo και σημείο εκκίνησης $(1,1)$.

ΘΕΜΑ 4: ΜΕΘΟΔΟΣ LEVENBERG MARQUARDT

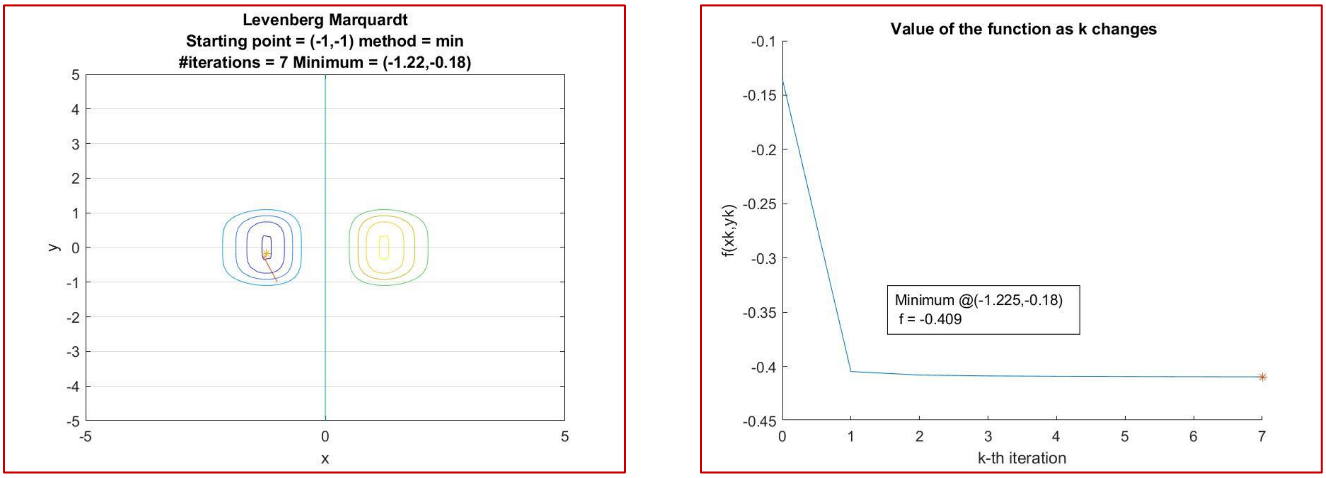
Τέλος, παρουσιάζεται η μέθοδος Levenberg Marquardt η οποία λύνει και το πρόβλημα της μεθόδου του Newton όταν ο πίνακας δεν είναι θετικά ορισμένος. Υπολογίζεται το μ_κ ώστε ο πίνακας $[\nabla^2 f(x) + \mu_\kappa I]$ να είναι θετικά ορισμένος (το $\mu_\kappa = \textbf{μεγαλύτερη απόλυτη ιδιοτιμή} + 0.2$). Με αυτόν τον τρόπο ο αλγόριθμος δεν αποκλίνει και όταν το μ_κ είναι αρκετά μεγάλο η μέθοδος συμπεριφέρεται σαν τη μέθοδο της μέγιστης καθόδου, αφού ο όρος που επικρατεί είναι ο $\mu_\kappa I$, ενώ όταν το μ_κ είναι μικρό η μέθοδος συμπεριφέρεται όπως η μέθοδος Newton.

Για το σημείο $(0,0)$ παρουσιάζεται, όπως και στις προηγούμενες μεθόδους, τοπικό ελάχιστο και ο αλγόριθμος δεν μπορεί να προχωρήσει.

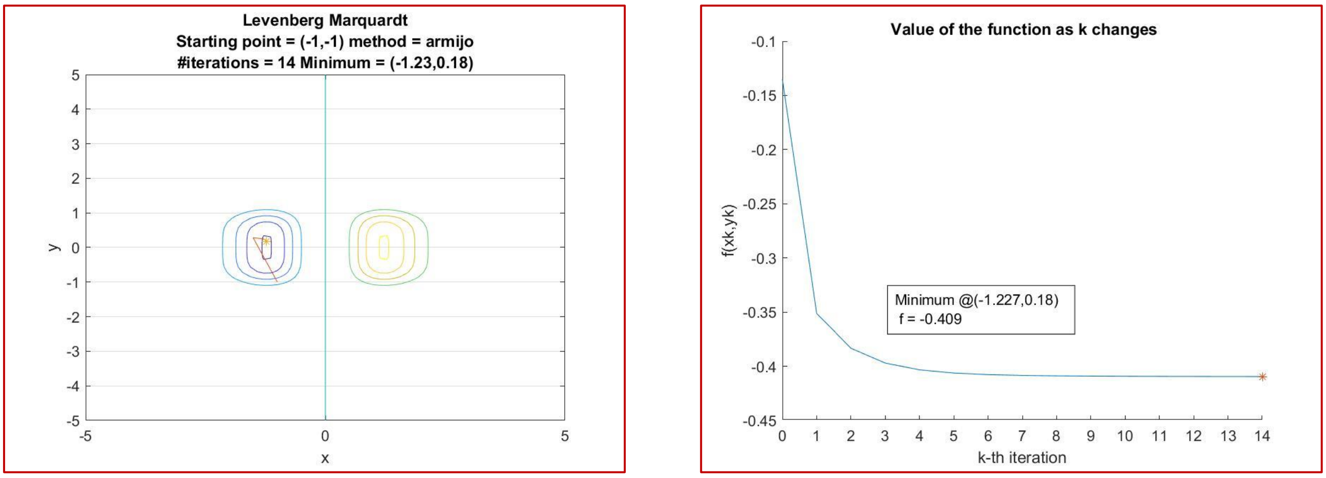
Όταν το αρχικό σημείο είναι το $(-1,-1)$, ο αλγόριθμος συγκλίνει προς το ολικό ελάχιστο της συνάρτησης και μάλιστα με μικρότερο αριθμό επαναλήψεων από ότι η steepest descent.



15. Σταθερό γ_κ και σημείο εκκίνησης (-1,-1).

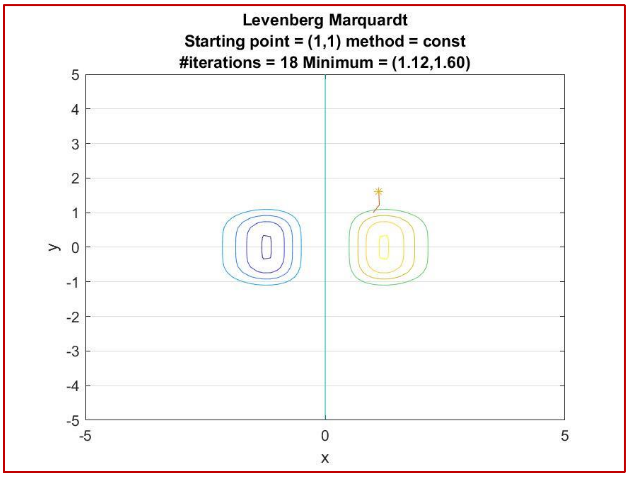


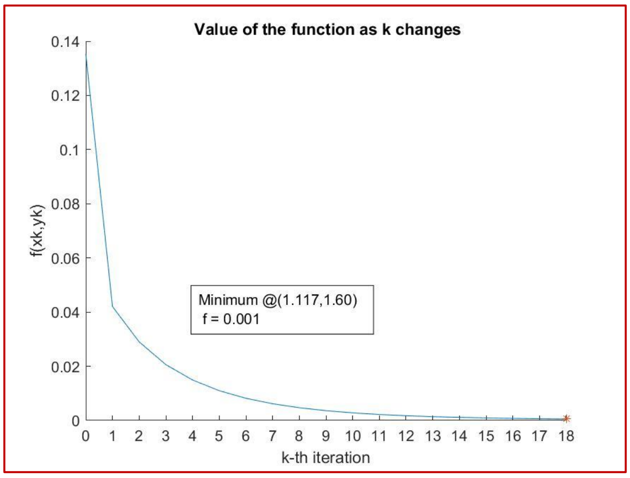
16. γ_κ που ελαχιστοποιεί τη συνάρτηση $f(\chi_\kappa + \gamma_\kappa * d_k)$ και σημείο εκκίνησης το (-1,-1).



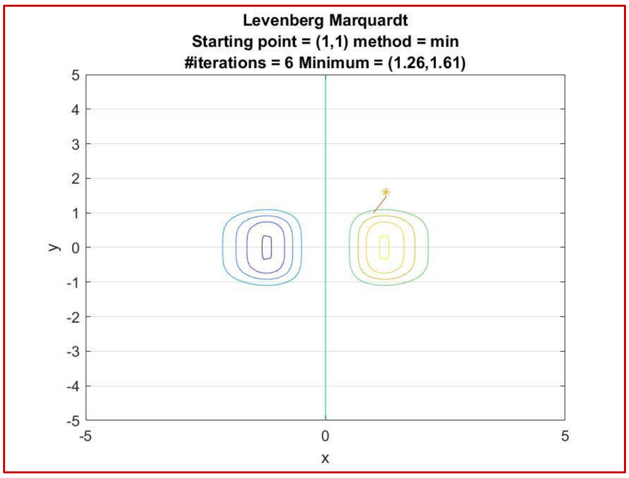
17. γ_κ υπολογισμένο με τη μέθοδο armijo και σημείο εκκίνησης (-1,-1).

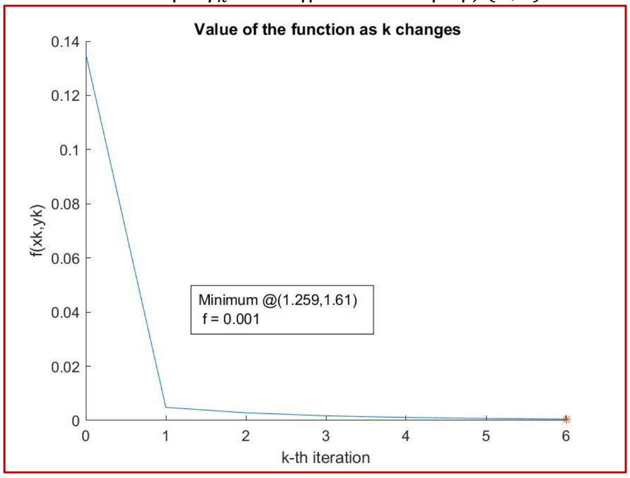
Τέλος, για το σημείο (1,1) και η μέθοδος Levenberg Marquardt βρίσκει τοπικό ελάχιστο της συνάρτησης και εγκλωβίζεται σε αυτό.



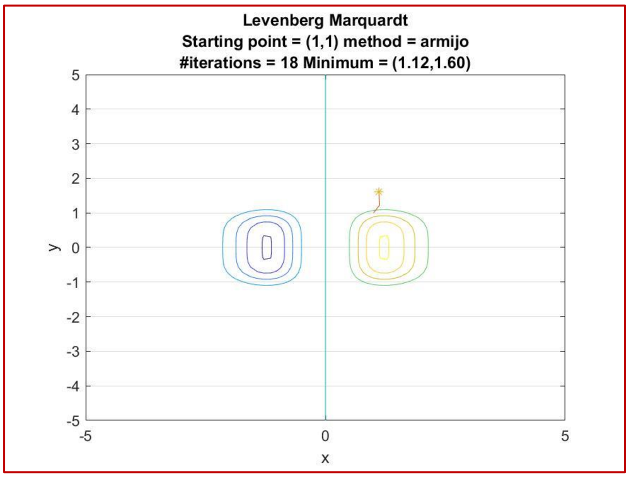


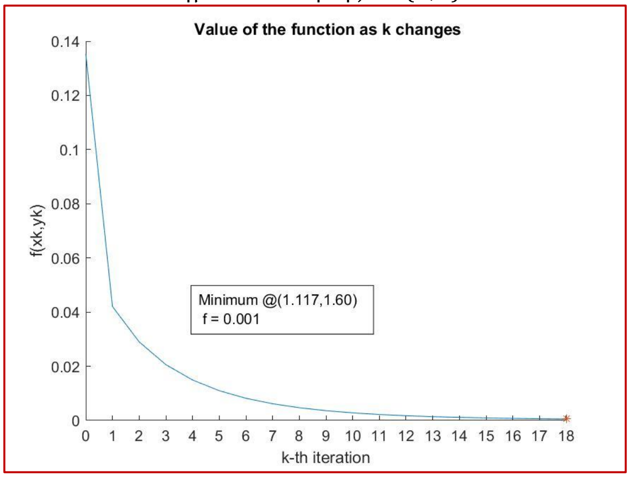
4 Σταθερό γ_κ και σημείο εκκίνησης (1,1).





54. γ_κ που ελαχιστοποιεί τη συνάρτηση $f(\chi_\kappa + \gamma_\kappa * d_k)$ και σημείο εκκίνησης το (1,1).





6γ_κ υπολογισμένο με τη μέθοδο armijo και σημείο εκκίνησης (1,1).

ΣΥΜΠΕΡΑΣΜΑΤΑ/ΣΧΟΛΙΑΣΜΟΣ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

Συγκρίνοντας τις μεθόδους επιλογής του γ_κ παρατηρείται ότι η μέθοδος α και γ χρειάζονταν περισσότερο αριθμό επαναλήψεων για την εύρεση του ελαχίστου ή τοπικού ελαχίστου σε σχέση με τη μέθοδο β.

Όσον αφορά τη σύγκριση των κύριων μεθόδων ελαχιστοποίησης, είναι προφανές ότι αυτός με το περισσότερο υπολογιστικό φόρτο είναι αυτός της μέγιστης καθόδου καθώς απαιτεί τα περισσότερα βήματα (συνήθως) για την εύρεση του ελαχίστου. Ο αλγόριθμος του Newton, αν και υπολογιστικά γρηγορότερος έχει σοβαρό πρόβλημα όταν ο εσσιανός της f δεν είναι θετικά ορισμένος. Από την άλλη πλευρά, ο αλγόριθμος Levenberg Marquardt συνδυάζει τους δύο αυτούς αλγορίθμους και καταλήγει στο ελάχιστον σε ικανοποιητικό αριθμό βημάτων, ξεκινώντας στην αρχή κάνοντας κάθετα βήματα όπως ο αλγόριθμος της μέγιστης καθόδου και έπειτα ακολουθεί τον αλγόριθμο του Newton, καθώς πλησιάζει στο ελάχιστο.

Παράλληλα βγαίνει ότι η επιλογή του σημείου έναρξης του αλγορίθμου παίζει καθοριστικό ρόλο στην εύρεση του ολικού ελαχίστου της συνάρτησης καθώς οι μέθοδοι εγκλωβίζονται εύκολα σε τοπικά ακρότατα. Τέλος, αυξάνοντας την τιμή του $\boldsymbol{\gamma_k}$ ο αριθμός των επαναλήψεων τον αλγορίθμων μειώνεται, καθώς αυξάνεται το «βήμα» με το οποίο προχωράμε στο επόμενο σημείο.