Τεχνικές Βελτιστοποίησης

Ιωάννης-Παναγιώτης Μπουντουρίδης

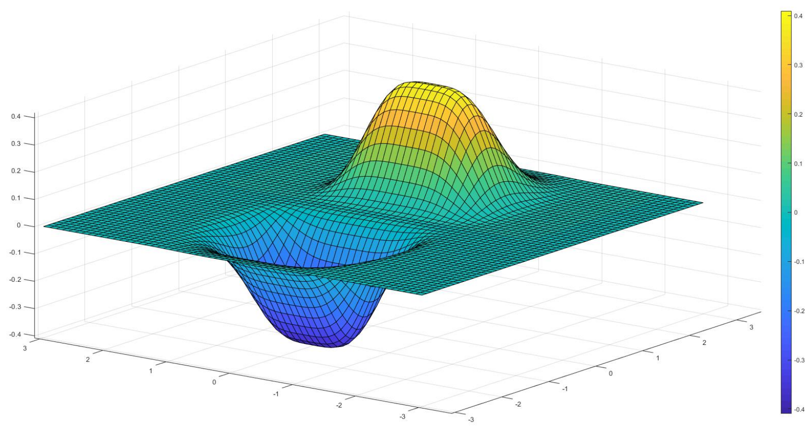
ΑΕΜ: 8872

work 2

Θέμα 1

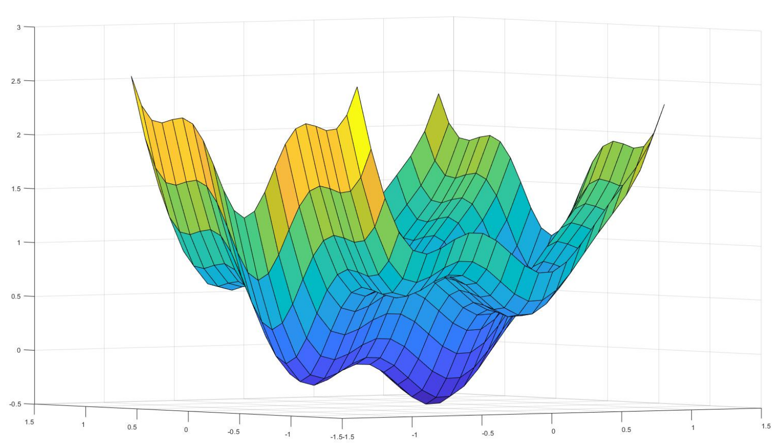
Γραφική παράσταση της συνάρτησης

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



Γραφική παράσταση της συνάρτησης

$$g(x, y) = x^4 + y^2 - 0.2sin(2\pi x) - 0.3cos(2\pi y)$$



ο κώδικας των γραφικών παραστάσεων υπάρχει στο αρχείο ***plot_functions.m***

Θέμα 2

Ζητούμενα

Σε αυτό το θέμα μας ζητείται να υλοποιήσουμε και να εφαρμόσουμε τη μέθοδο μέγιστης καθόδου (steepest descent) για να ελαχιστοποιήσουμε την συναρτηαη f 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 $f(x_k + g_k \cdot d_k)$
- βάσει του κανόνα Armijo

Περιγραφή αλγορίθμου

Θεωρούμε το πρόβλημα ελαχιστοποίησης μιας συνάρτησης τουλάχιστον δυο φορές παραγωγίσιμης f, στην ιδέα της επαναληπτικής διαδικασίας η οποία έχει ως εξής: Ξεκινάμε από το σημείο x_0 και παράγουμε διαδοχικά τα διανύσματα $x_1, x_2, ..$ ώστε

$$f(x_{k+1}) < f(x_k) \;\; k = 0, 1, 2, ...$$

Ο αλγόριθμος υλοποιεί την ιδέα της επαναληπτικής καθόδου που μας οδηγεί σε ολοένα και βελτιωμένες τιμές της f, προς την ελαχιστοποίηση της.

Αρχικοποίηση Ορίστε $\varepsilon > 0$ τη σταθερά τερματισμού. Επιλέξτε το αρχικό σημείο x_1 και θέστε $k = 1$.

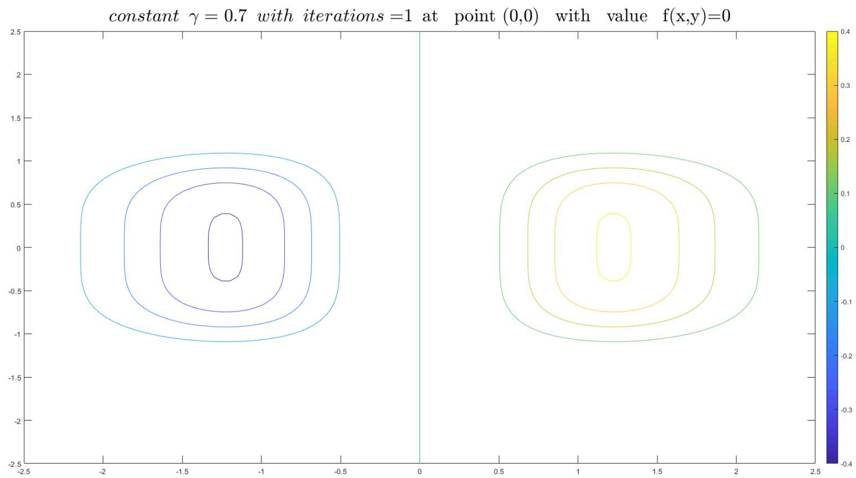
Κύριο Βήμα Αν $|\nabla f(x_k)| < \varepsilon$ ο αλγόριθμος τερματίζει. Διαφορετικά, θέστε $d_k = -\nabla f(x_k)$ και ορίστε ως γ_k την τιμή που ελαχιστοποιεί την $f(x_k - \gamma_k \nabla f(x_k))$ ως προς γ_k με $\gamma_k \geq 0$. Θέστε $x_{k+1} = x_k - \gamma_k \nabla f(x_k)$, $k = k + 1$ και επαναλάβετε το κυρίως βήμα.

Σταθερό γάμμα

Θέτουμε ένα σταθερό $\boxed{\gamma_f = 0.7}$ της επιλογής μας για την συνάρτηση f. Εντός του αλγορίθμου κάνουμε τους απαραίτητους ελέγχους για να δούμε αν η επιλογή μας αυτή θα συγκλίνει σε κάποιο αποτέλεσμα. Όταν η τιμή του γ είναι πολύ μικρή τότε τα βήματα των επαναλήψεων για την εύρεση ελαχίστου αυξάνουν σημαντικά. Απο την άλλη η επιλογή ενος μεγάλου γ για το σύστημα προκαλεί αστάθεια καθώς με μεγάλο βήμα ο αλγόριθμος αδυνατεί να βρει τον ελάχιστο. Επίσης για την ακρίβεια e, δηλαδή πόσο κοντά θα είμαστε στο ελάχιστο επιλέξαμε αρκούντος μικρή τιμή $\boxed{e = 10^{-4}}$

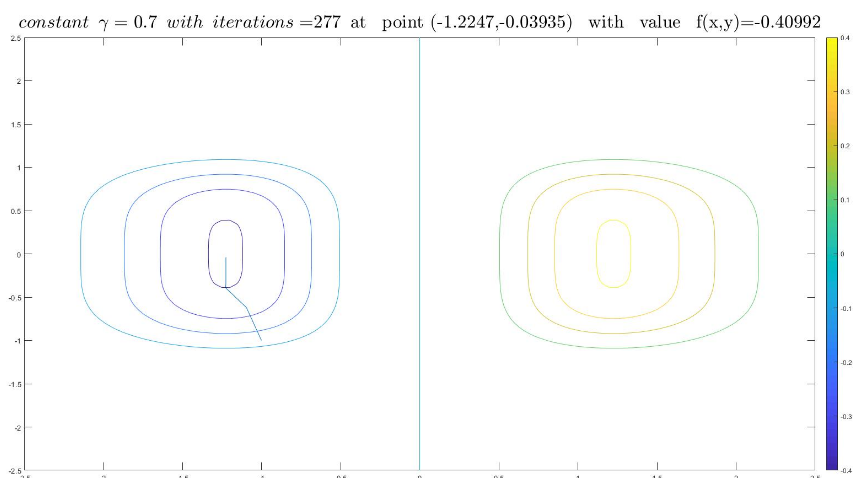
Πιο αναλυτικά, στο ελάχιστο η παράγωγος είναι μηδέν οπότε για να προσεγγίσουμε το ελάχιστο και να βρισκόμαστε κοντά του πρέπει η παράγωγος να είναι πολύ μικρή-σχεδόν μηδέν. Οπότε ξεκινώντας για γάμα 0.7 έχουμε τις εξής γραφικές για την μέθοδο μέγιστης καθόδου.

Αρχικό σημείο (0,0) για την συνάρτηση f



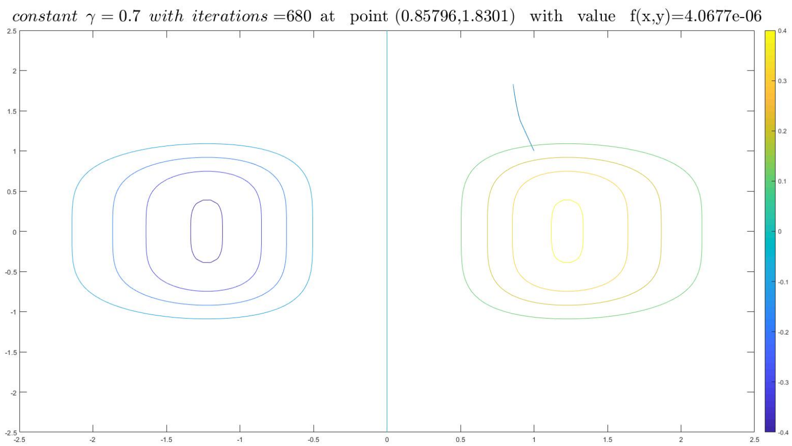
Για αρχικό σημείο (0,0) η παράγωγος της f είναι μηδέν οπότε ο αλγόριθμος τερματίζεται πρόωρα και συνεπώς εγκλωβιζόμαστε στο σημείο f(x,y)=0.

Αρχικό σημείο (-1,-1) για την συνάρτηση f



Για αρχικό σημείο (-1,-1) μετά απο 277 επαναλήψεις καταλήγουμε στο σημείο (-1.22,-0.03) για δοσμένη ακρίβεια e=10^{-4} με τιμή f(x,y)=-0.409 που είναι το **ολικό ελάχιστο**

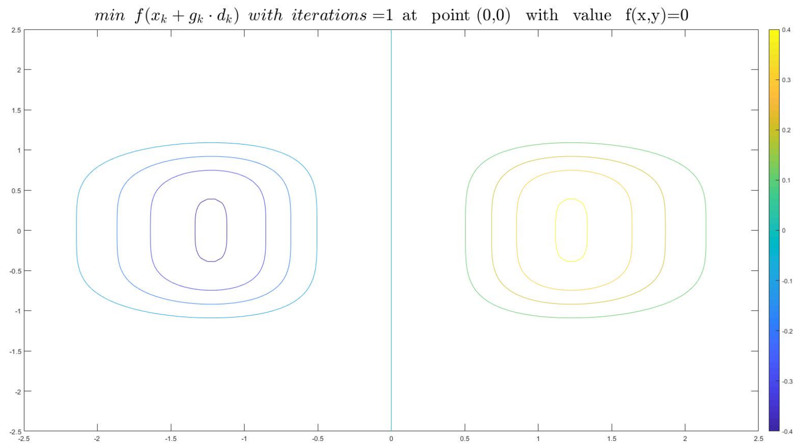
Αρχικό σημείο (1,1) για την συνάρτηση f



Για αρχικό σημείο (1,1) μετά απο 680 επαναλήψεις καταλήγουμε σε τοπικό ελάχιστο (0.85,1.83) για δοσμένη ακρίβεια e=10^{-4} με τιμή f(x,y)=$4{\cdot}10^{-6}$. Αποτέλεσμα λογικό αφού με τόσο μικρό βήμα είναι αδύνατον να βρεθούμε στην απέναντι περιοχή του ολικού ελάχιστου χωρίς να εγκλωβιστούμε.

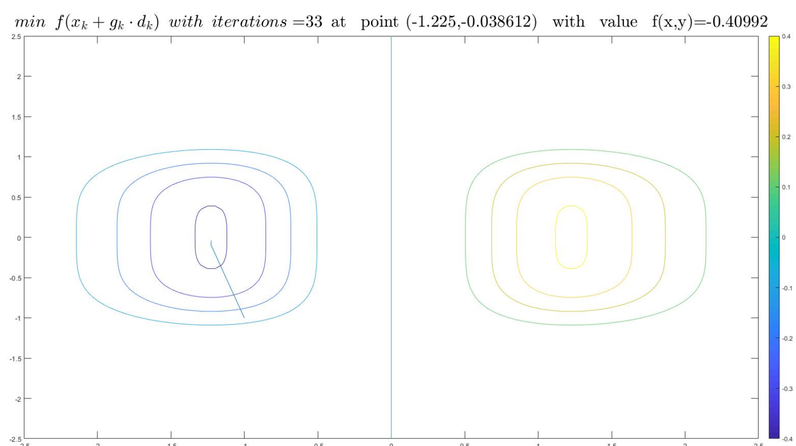
Ελαχιστοποίηση $f(x_k + g_k \cdot d_k)$

Αρχικό σημείο (0,0) για την συνάρτηση f



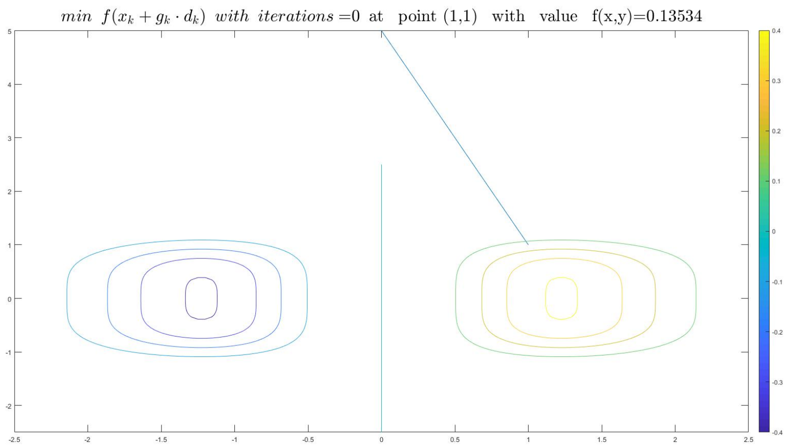
Για αρχικό σημείο (0,0) η παράγωγος της f είναι μηδέν οπότε ο αλγόριθμος τερματίζεται πρόωρα και συνεπώς εγκλωβιζόμαστε στο σημείο f(x,y)=0.

Αρχικό σημείο (-1,-1) για την συνάρτηση f



Για αρχικό σημείο (-1,-1) μετά απο 33 επαναλήψεις καταλήγουμε στο σημείο (-1.22,-0.03) για δοσμένη ακρίβεια e=10^{-4} με τιμή f(x,y)=-0.409 που είναι το **ολικό ελάχιστο**

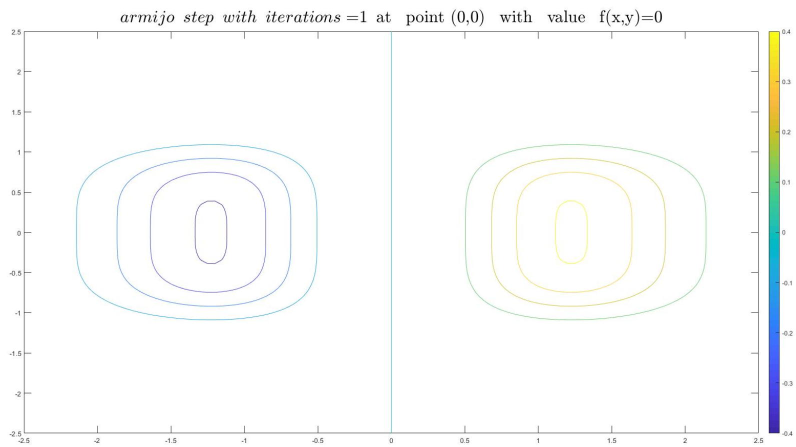
Αρχικό σημείο (1,1) για την συνάρτηση f



Για αρχικό σημείο (1,1) βλέπουμε και από την παραπάνω γραφική ο αλγόριθμος δεν συγκλίνει στο ελάχιστο. Μπορεί το γάμμα να μην είναι σταθερό και να ακολουθεί έναν κανόνα αύξησης(όταν αυτό κρίνεται απαραίτητο) ωστόσο ένας απλός κανόνας δεν φτάνει ώστε να καταφέρουμε την σύγκλιση του αλγορίθμου. Ο εκάστοτε κανόνας,συνήθως για να είναι αποδοτικός, πρέπει να είναι καλά σχεδιασμένος και να βασίζεται σε μαθηματικές σχέσεις και περιορισμούς.

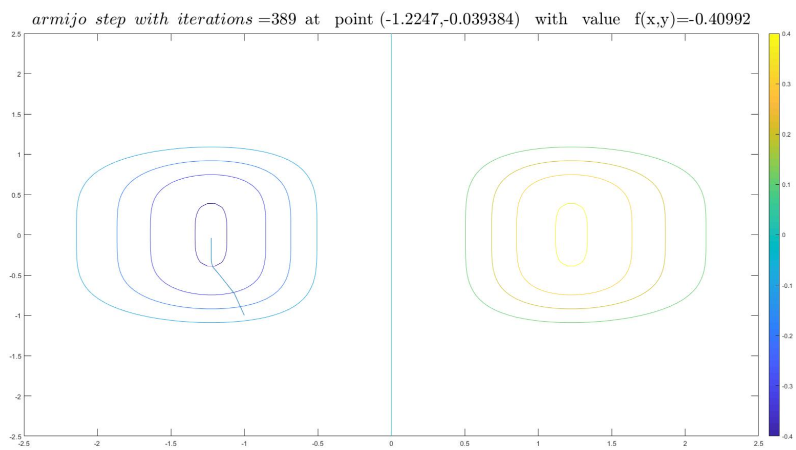
Κανόνας Armijo

Αρχικό σημείο (0,0) για την συνάρτηση f



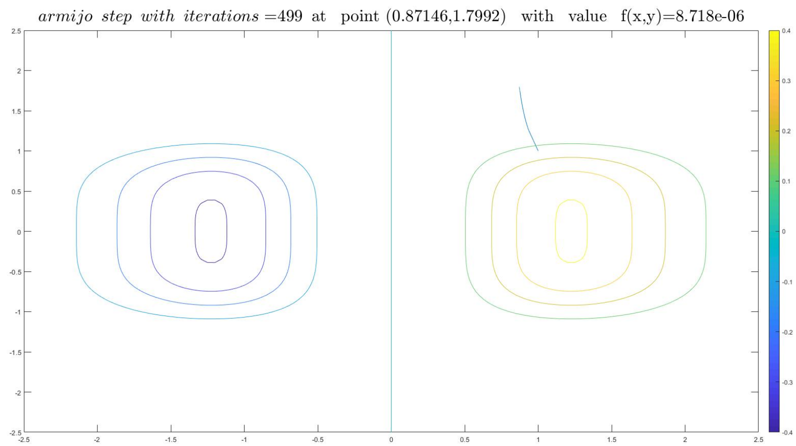
Για αρχικό σημείο (0,0) η παράγωγος της f είναι μηδέν οπότε ο αλγόριθμος τερματίζεται πρόωρα και συνεπώς εγκλωβιζόμαστε στο σημείο f(x,y)=0.

Αρχικό σημείο (-1,-1) για την συνάρτηση f



Για αρχικό σημείο (-1,-1) μετά απο 389 επαναλήψεις καταλήγουμε στο σημείο (-1.22,-0.03) για δοσμένη ακρίβεια e=10^{-4} με τιμή f(x,y)=-0.409 που είναι το **ολικό ελάχιστο**

Αρχικό σημείο (1,1) για την συνάρτηση f



Για αρχικό σημείο (1,1) μετά απο 499 επαναλήψεις καταλήγουμε σε τοπικό ελάχιστο (0.87,1.79) για δοσμένη ακρίβεια e=10^{-4} με τιμή f(x,y)=$8.7\cdot10^{-6}$.

Θέμα 3

Ζητούμενα

Σε αυτό το θέμα μας ζητείται να υλοποιήσουμε και να εφαρμόσουμε τη μέθοδο newton για να ελαχιστοποιήσουμε την συναρτηση f 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 $f(x_k + g_k \cdot d_k)$
- βάσει του κανόνα Armijo

Περιγραφή αλγορίθμου

Αρχικοποίηση Ορίστε το σημείο εκκίνησης x_0, τη σταθερά τερματισμού $\varepsilon \geq 0$ και θέστε $k = 1$.

Βήμα 1 Υπολογίστε την $\nabla f(x_k)$. Αν $|\nabla f(x_k)| < \varepsilon$ ο αλγόριθμος τερματίζει. Διαφορετικά, πηγαίνετε στο Βήμα 2.

Βήμα 2 Υπολογίστε τον $\nabla^2 f(x) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}\right] \quad i, j = 1, 2, \ldots, n.$

Βήμα 3 Υπολογίστε την κατεύθυνση αναζήτησης $d_k = -[\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$.

Βήμα 4 Θέστε $x_{k+1} = x_k + \gamma_k d_k$ όπου το $\gamma_k \geq 0$ υπολογίζεται έτσι ώστε να ελαχιστοποιεί την $f(x_k + \gamma_k d_k)$ ως προς γ_k.

Βήμα 5 Θέστε $k = k + 1$ και πηγαίνετε στο Βήμα 1.

Σταθερό γάμμα

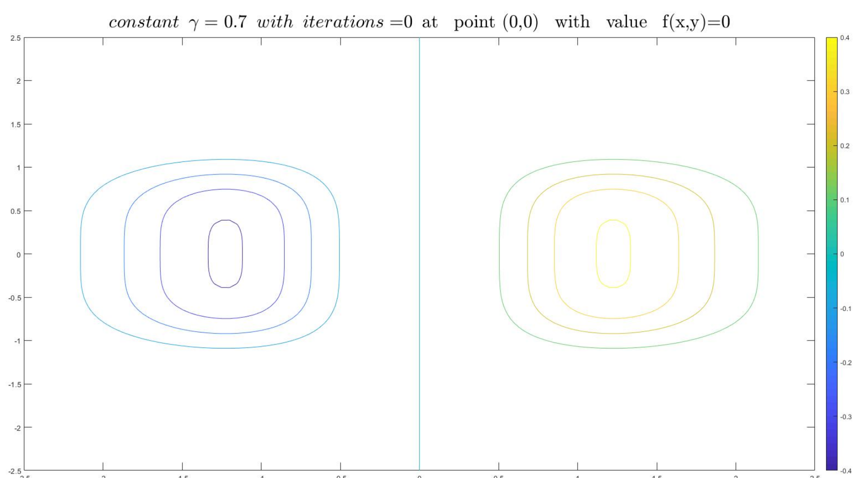
Θέτουμε ένα σταθερό $\boxed{\gamma_f = 0.7}$ της επιλογής μας για την συνάρτηση f. Όταν η τιμή του γ είναι πολύ μικρή τότε τα βήματα των επαναλήψεων για την εύρεση ελαχίστου αυξάνουν σημαντικά. Απο την άλλη η επιλογή ενος μεγάλου γ για το σύστημα προκαλεί αστάθεια καθώς με μεγάλο βήμα ο αλγόριθμος αδυνατεί να βρει τον ελάχιστο. Επίσης για την ακρίβεια e, δηλαδή πόσο κοντά θα είμαστε στο ελάχιστο επιλέξαμε αρκούντος μικρή τιμή $\boxed{e = 10^{-4}}$

Σχόλια

Αφού υλοποιήσαμε τον αλγόριθμο της μεθόδο Newton, δοκιμάσαμε τα σημεία που μας ζητήθηκε να εξετάσουμε. Στη συνέχεια, κατασκευάσαμε τη συνάρτηση function x = checkPositiveDefinite(matrix) η οποία επιστρέφει 1 αν πίνακας πίνακας matrix είναι θετικά ορισμένος και αλλιώς μηδέν.Στα σημεία (-1,-1) και (1,1) όπου ο αλγόριθμος τερματίζεται πριν την πρώτη επανάληψη στα ο εσσιανός πίνακας για κάθε σημείο είναι αρνητικά ορισμένος. Οπότε όπως είναι γνωστό όταν ο εσσιανός της αντικειμενικής συναρτήσεις είναι αρνητικά ορισμένος η μέθοδος Newton είτε δεν ορίζεται είτε δίνει λύσεις πολύ μακρυά από το σημείο του ελαχίστου για οποιοδήποτε γάμμα. Για αυτό το λόγο θα παρουσιάσουμε ενδεικτικά την πορεία εύρεσης ελαχίστου στο σημείο (-1,-0.5) με γάμμα σταθερό και ίσο με 0.7 και ακρίβεια .

Σταθερό γαμμα

Αρχικό σημείο (0,0) για την συνάρτηση f



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.

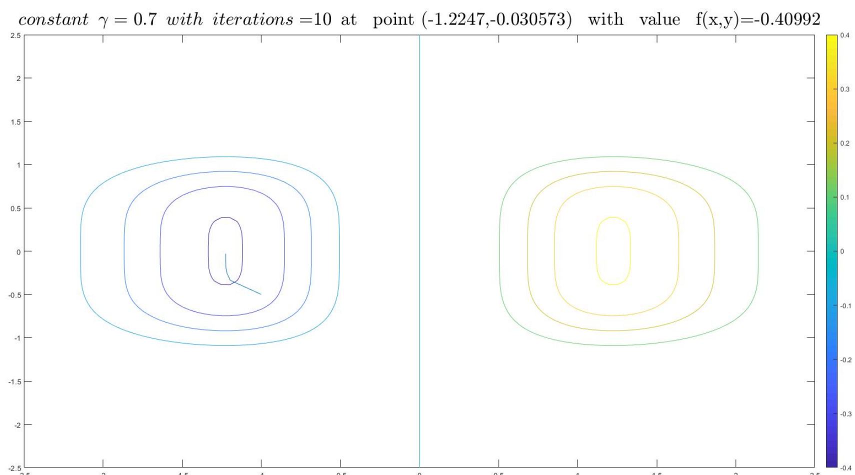
Αρχικό σημείο (-1,-1) για την συνάρτηση f

Για αρχικό σημείο (-1,-1) ο αλγόριθμος τερματίζει αφού έχουμε αρνητικά ορισμένο εσσιανό πίνακα.

Αρχικό σημείο (1,1) για την συνάρτηση f

Για αρχικό σημείο (1,1) ο αλγόριθμος τερματίζει αφού έχουμε αρνητικά ορισμένο εσσιανό πίνακα.

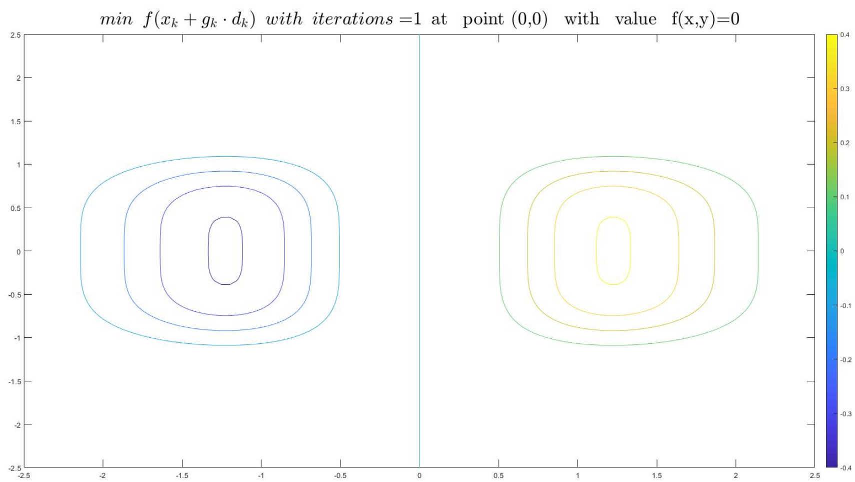
Αρχικό σημείο (-1,-0.5) για την συνάρτηση f



Για αρχικό σημείο (-1,-0.5) μετά απο 10 επαναλήψεις καταλήγουμε στο σημείο (-1.22,-0.03) για δοσμένη ακρίβεια e=10^{-4} με τιμή f(x,y)=-0.409 που είναι το **ολικό ελάχιστο**

Ελαχιστοποίηση $f(x_k + g_k \cdot d_k)$

Αρχικό σημείο (0,0) για την συνάρτηση f



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.

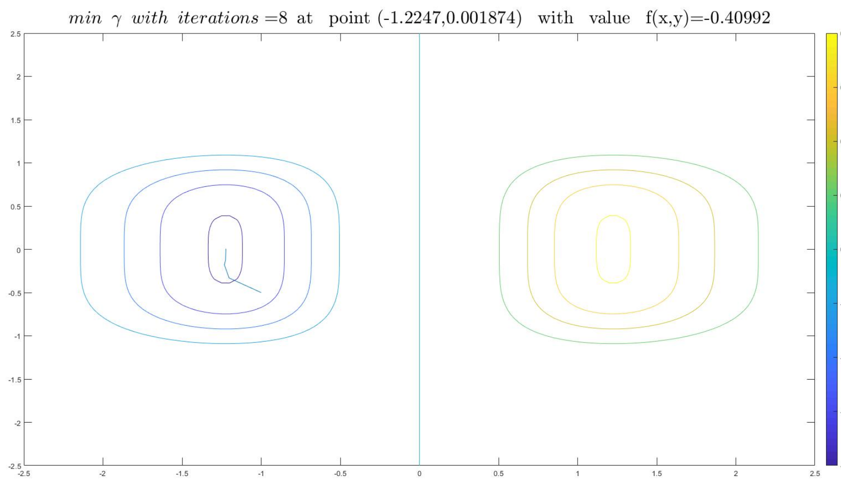
Αρχικό σημείο (-1,-1) για την συνάρτηση f

Για αρχικό σημείο (-1,-1) ο αλγόριθμος τερματίζει αφού έχουμε αρνητικά ορισμένο εσσιανό πίνακα.

Αρχικό σημείο (1,1) για την συνάρτηση f

Για αρχικό σημείο (1,1) ο αλγόριθμος τερματίζει αφού έχουμε αρνητικά ορισμένο εσσιανό πίνακα.

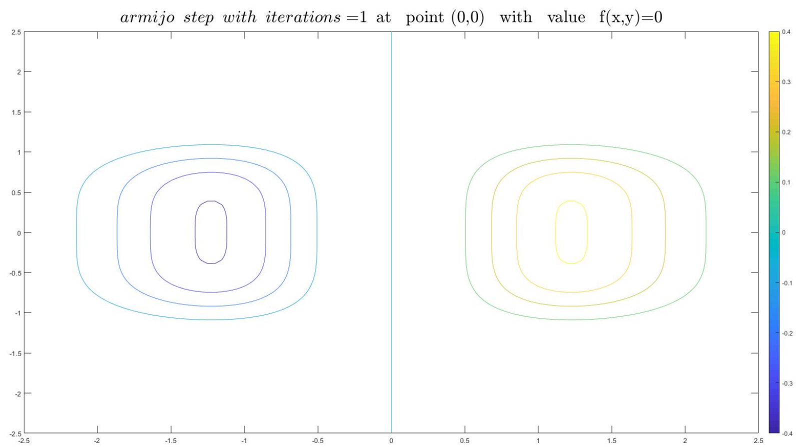
Αρχικό σημείο (-1,-0.5) για την συνάρτηση f



Για αρχικό σημείο (-1,-0.5) μετά απο 8 επαναλήψεις καταλήγουμε στο σημείο (-1.22,-0.03) για δοσμένη ακρίβεια e=10^{-4} με τιμή f(x,y)=-0.409 που είναι το **ολικό ελάχιστο**

Κανόνας Armijo

Αρχικό σημείο (0,0) για την συνάρτηση f



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.

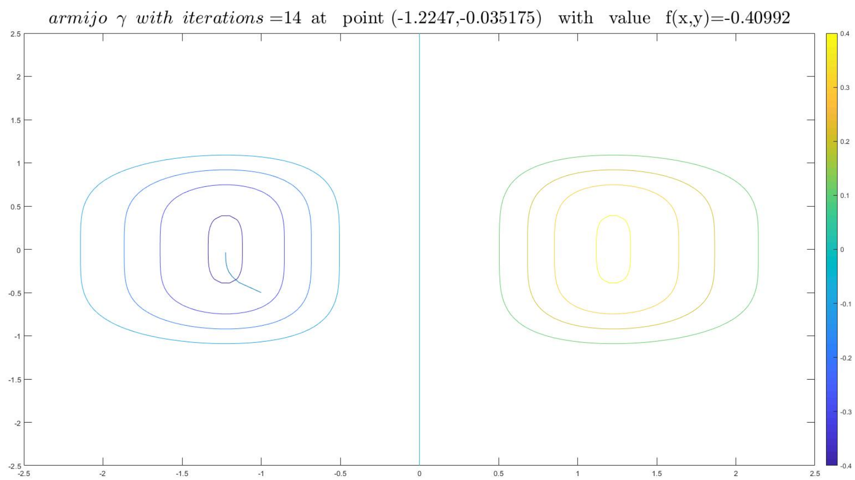
Αρχικό σημείο (-1,-1) για την συνάρτηση f

Για αρχικό σημείο (-1,-1) ο αλγόριθμος τερματίζει αφού έχουμε αρνητικά ορισμένο εσσιανό πίνακα.

Αρχικό σημείο (1,1) για την συνάρτηση f

Για αρχικό σημείο (1,1) ο αλγόριθμος τερματίζει αφού έχουμε αρνητικά ορισμένο εσσιανό πίνακα.

Αρχικό σημείο (-1,-0.5) για την συνάρτηση f



Για αρχικό σημείο (-1,-0.5) μετά απο 14 επαναλήψεις καταλήγουμε στο σημείο (-1.22,-0.03) για δοσμένη ακρίβεια e=10^{-4} με τιμή f(x,y)=-0.409 που είναι το **ολικό ελάχιστο**

Θέμα 4

Ζητούμενα

Σε αυτό το θέμα μας ζητείται να υλοποιήσουμε και να εφαρμόσουμε τη μέθοδο Levenberg-Marquardt για να ελαχιστοποιήσουμε την συναρτηση f 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 $f(x_k + g_k \cdot d_k)$
- βάσει του κανόνα Armijo

Περιγραφή αλγορίθμου

Βήμα 1 Δοσμένου του x_k.

Βήμα 2 Υπολογίστε το μ_k έτσι ώστε ο $\nabla^2 f(x_k) + \mu_k I$ να είναι θετικά ορισμένος.

Βήμα 3 Λύστε το σύστημα $[\nabla^2 f(x_k) + \mu_k I]d_k = -\nabla f(x_k)$.

Βήμα 4 Υπολογίστε το γ_k ώστε να ικανοποιούνται τα Κριτήρια 3, 4 του Θεωρήματος 5.2.6.

Βήμα 5 Θέστε $x_{k+1} = x_k + \gamma_k d_k$.

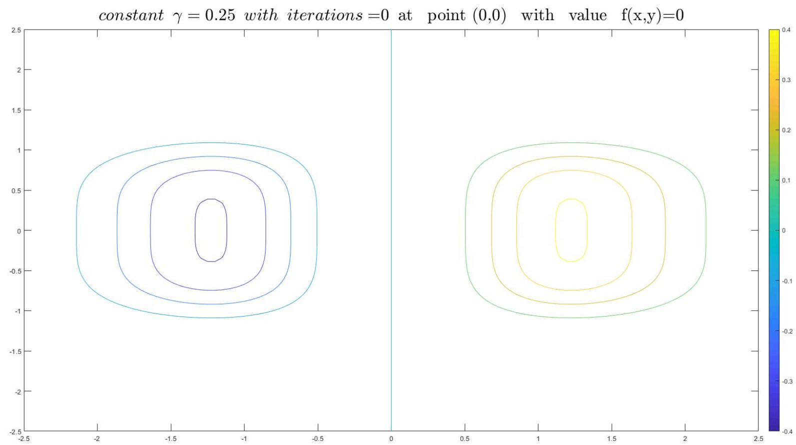
Βήμα 6 Πηγαίνετε στο Βήμα 1.

Σταθερό γάμμα

Θέτουμε ένα σταθερό $\boxed{\gamma_f = 0.25}$ της επιλογής μας για την συνάρτηση f. Όταν η τιμή του γ είναι πολύ μικρή τότε τα βήματα των επαναλήψεων για την εύρεση ελαχίστου αυξάνουν σημαντικά. Απο την άλλη η επιλογή ενος μεγάλου γ για το σύστημα προκαλεί αστάθεια καθώς με μεγάλο βήμα ο αλγόριθμος αδυνατεί να βρει τον ελάχιστο. Επίσης για την ακρίβεια e, δηλαδή πόσο κοντά θα είμαστε στο ελάχιστο επιλέξαμε αρκούντος μικρή τιμή $\boxed{e = 10^{-4}}$

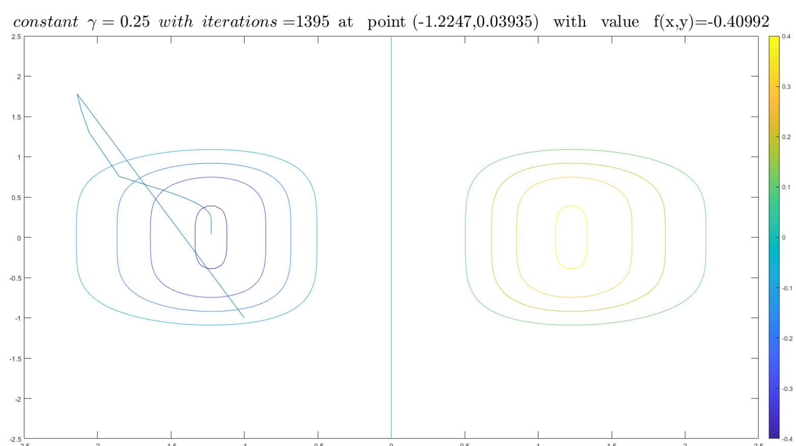
Σταθερό γαμμα

Αρχικό σημείο (0,0) για την συνάρτηση f



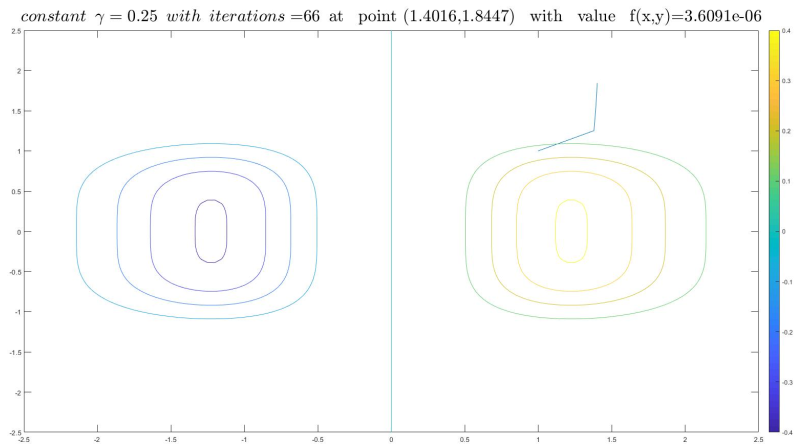
Για αρχικό σημείο (0,0) η παράγωγος της f είναι μηδέν οπότε ο αλγόριθμος τερματίζεται πρόωρα και συνεπώς εγκλωβιζόμαστε στο σημείο f(x,y)=0.

Αρχικό σημείο (-1,-1) για την συνάρτηση f



Για αρχικό σημείο (-1,-1) μετά απο 1395 επαναλήψεις καταλήγουμε στο σημείο (-1.22,-0.03) για δοσμένη ακρίβεια e=10^{-4} με τιμή f(x,y)=-0.409 που είναι το **ολικό ελάχιστο**

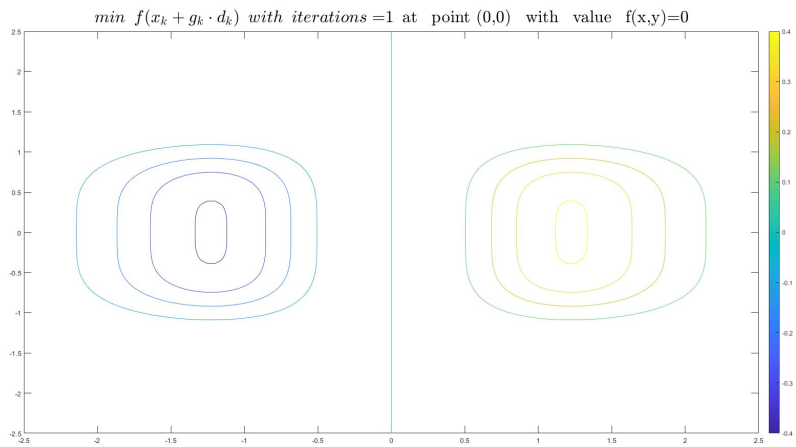
Αρχικό σημείο (1,1) για την συνάρτηση f



Για αρχικό σημείο (1,1) μετά απο 66 επαναλήψεις καταλήγουμε σε τοπικό ελάχιστο (0.85,1.83) για δοσμένη ακρίβεια e=10^{-4} με τιμή f(x,y)=$4{\cdot}10^{-6}$. Αποτέλεσμα λογικό αφού με τόσο μικρό βήμα είναι αδύνατον να βρεθούμε στην απέναντι περιοχή του ολικού ελάχιστου χωρίς να εγκλωβιστούμε.

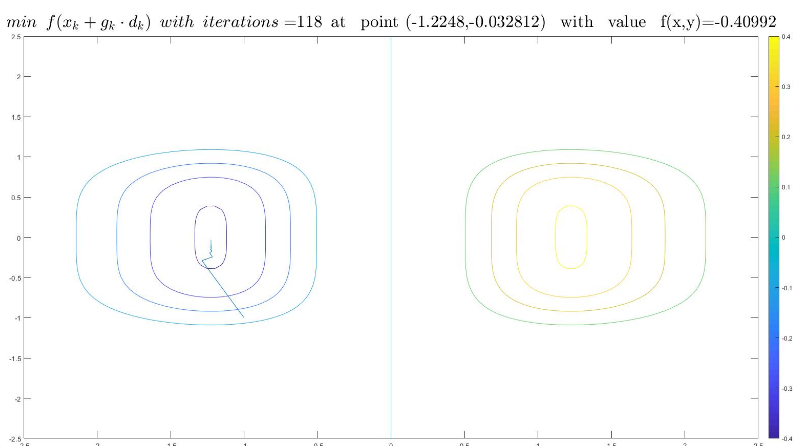
Ελαχιστοποίηση $f(x_k + g_k \cdot d_k)$

Αρχικό σημείο (0,0) για την συνάρτηση f



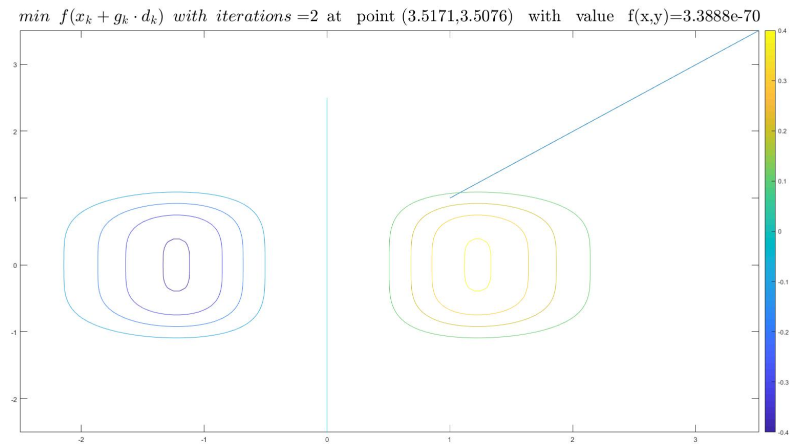
Για αρχικό σημείο (0,0) η παράγωγος της f είναι μηδέν οπότε ο αλγόριθμος τερματίζεται πρόωρα και συνεπώς εγκλωβιζόμαστε στο σημείο f(x,y)=0.

Αρχικό σημείο (-1,-1) για την συνάρτηση f



Για αρχικό σημείο (-1,-1) μετά απο 33 επαναλήψεις καταλήγουμε στο σημείο (-1.22,-0.03) για δοσμένη ακρίβεια e=10^{-4} με τιμή f(x,y)=-0.409 που είναι το **ολικό ελάχιστο**

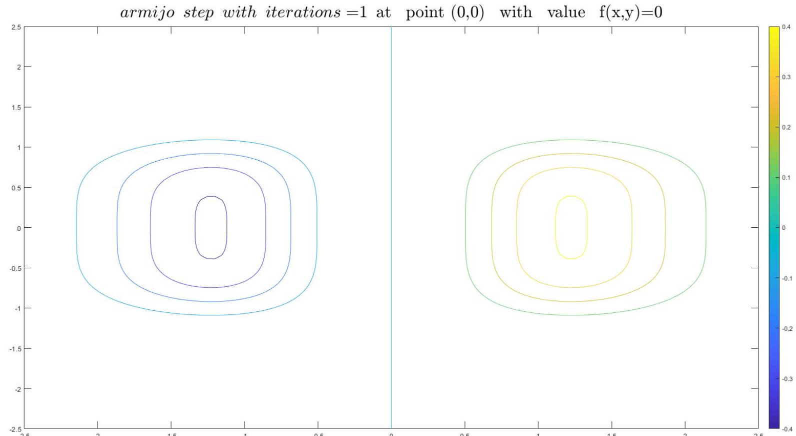
Αρχικό σημείο (1,1) για την συνάρτηση f



Για αρχικό σημείο (1,1) βλέπουμε και από την παραπάνω γραφική ο αλγόριθμος δεν συγκλίνει στο ελάχιστο. Μπορεί το γάμμα να μην είναι σταθερό και να ακολουθεί έναν κανόνα αύξησης(όταν αυτό κρίνεται απαραίτητο) ωστόσο ένας απλός κανόνας δεν φτάνει ώστε να καταφέρουμε την σύγκλιση του αλγορίθμου. Ο εκάστοτε κανόνας,συνήθως για να είναι αποδοτικός, πρέπει να είναι καλά σχεδιασμένος και να βασίζεται σε μαθηματικές σχέσεις και περιορισμούς.

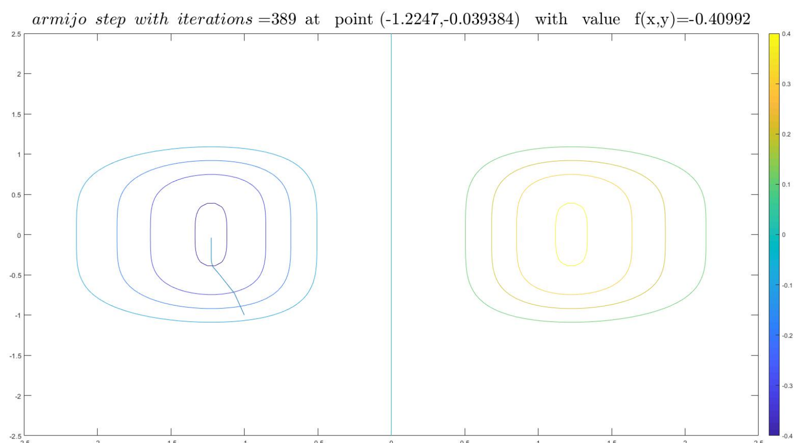
Κανόνας Armijo

Αρχικό σημείο (0,0) για την συνάρτηση f



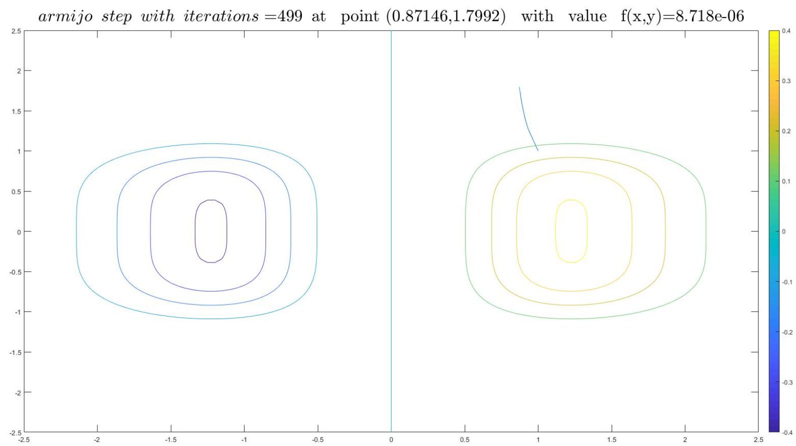
Για αρχικό σημείο (0,0) η παράγωγος της f είναι μηδέν οπότε ο αλγόριθμος τερματίζεται πρόωρα και συνεπώς εγκλωβιζόμαστε στο σημείο f(x,y)=0.

Αρχικό σημείο (-1,-1) για την συνάρτηση f



Για αρχικό σημείο (-1,-1) μετά απο 389 επαναλήψεις καταλήγουμε στο σημείο (-1.22,-0.03) για δοσμένη ακρίβεια e=10^{-4} με τιμή f(x,y)=-0.409 που είναι το **ολικό ελάχιστο**

Αρχικό σημείο (1,1) για την συνάρτηση f



Για αρχικό σημείο (1,1) μετά απο 499 επαναλήψεις καταλήγουμε σε τοπικό ελάχιστο (0.87,1.79) για δοσμένη ακρίβεια e=10^{-4} με τιμή f(x,y)=$8.7 \cdot 10^{-6}$.