Τεχνικές Βελτιστοποίησης

2η εργασία

Ονοματεπώνυμο: Παναγιώτης Λεκός
ΑΕΜ: 10347
Email: panaleko@ece.auth.gr

Στην παρούσα εργασία θα μελετήσουμε τρεις διαφορετικές μεθόδους ελαχιστοποίησης πολλών μεταβλητών χωρίς περιορισμούς με χρήση παραγώγων, σε τρεις διαφορετικές παραλλαγές, μιας δοσμένης συνάρτησης με τη χρήση του MATLAB. Θα μελετήσουμε τα πλεονεκτήματα και τα μειονεκτήματα αυτών, θα δούμε πώς το αρχικό σημείο αναζήτησης επηρεάζει το αποτέλεσμα, και θα μελετήσουμε το υπολογιστικό κόστος τους.

Οι τρεις αυτές μέθοδοι είναι οι:

- Μέθοδος Μέγιστης καθόδου (Steepest Descent)
- Μέθοδος Newton
- Μέθοδος Levenberg-Marquardt

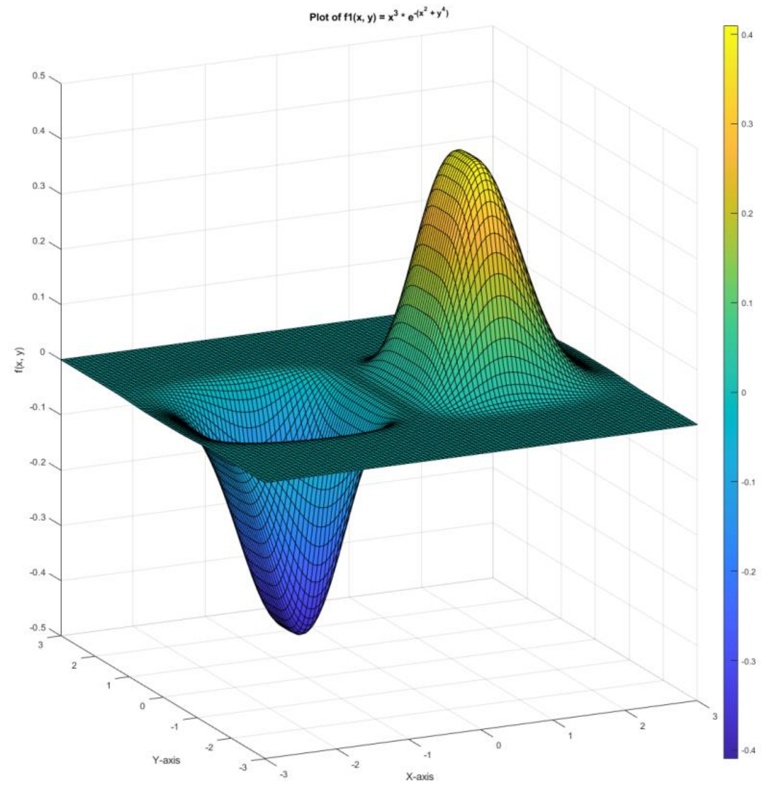
Οι διαφορετικές παραλλαγές στην υλοποίηση των αλγορίθμων αυτών αφορούν το βήμα γ_κ. Στην πρώτη παραλλαγή επιλέγουμε ένα σταθερό βήμα. Στην δεύτερη παραλλαγή σε κάθε επανάληψη του αλγορίθμου υπολογίζουμε το γ_κ το οποίο ελαχιστοποιεί την συνάρτηση $f(x_k + \gamma_\kappa * d_k)$, χρησιμοποιώντας έναν αλγόριθμο ελαχιστοποίησης που υλοποιήσαμε στην πρώτη εργασία. Στην Τρίτη υπολογίζουμε το γ_κ σε κάθε επανάληψη με βάση τον κανόνα του Armijo.

Η αντικειμενική συνάρτηση που θα μελετήσουμε είναι η:

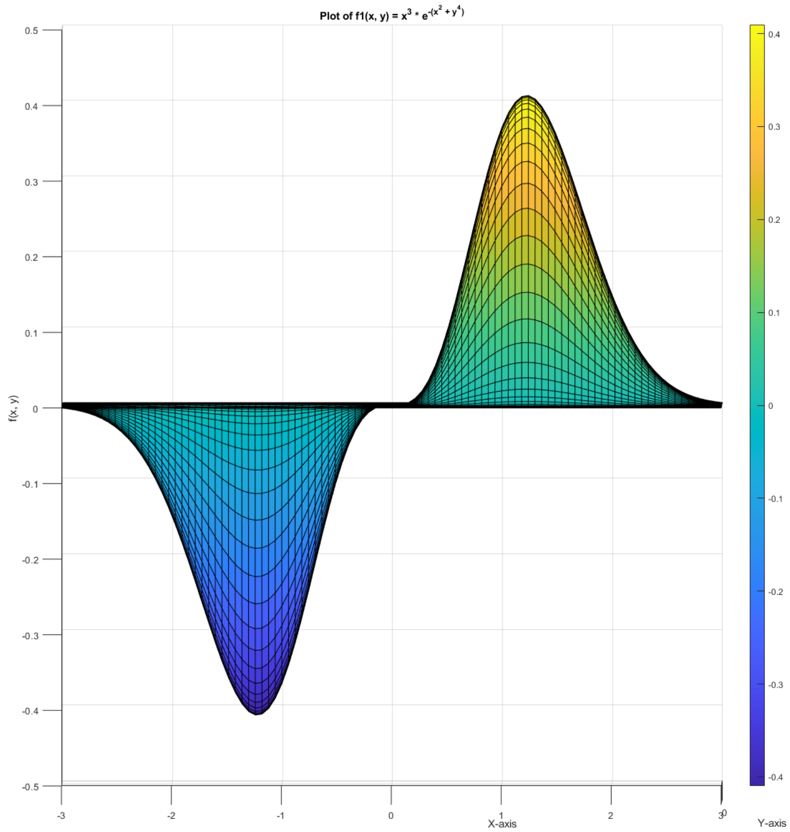
$$f(x,y) = x^3 * e^{-x^2-y^4}$$

Η συνάρτηση $f(x,y)$ υλοποιήθηκε στο αρχείο f.m, η $\nabla f(x,y)$ στο grad_f.m, η $\nabla^2 f(x,y)$ στο hessian_f.m. Στα αρχεία steepest_descent_x.m , newton_x.m , levenberg_marquardt_x.m υλοποιήσαμε τις μεθόδους, όπου x είναι a, b, c, για την πρώτη, δεύτερη, και τρίτη παραλλαγή αντίστοιχα. Η μέθοδος που επιλέχθηκε για την δεύτερη παραλλαγή είναι η μέθοδος του χρυσού τομέα (golden_section_search.m). Επίσης στα αρχεία Thema_x_y.m γίνεται η εκτέλεση των μεθόδων ώστε να εξάγουμε τα αποτελέσματα (π.χ. Thema_2_c.m εκτέλεση της steepest_descent με την παραλλαγή του κανόνα του Armijo).

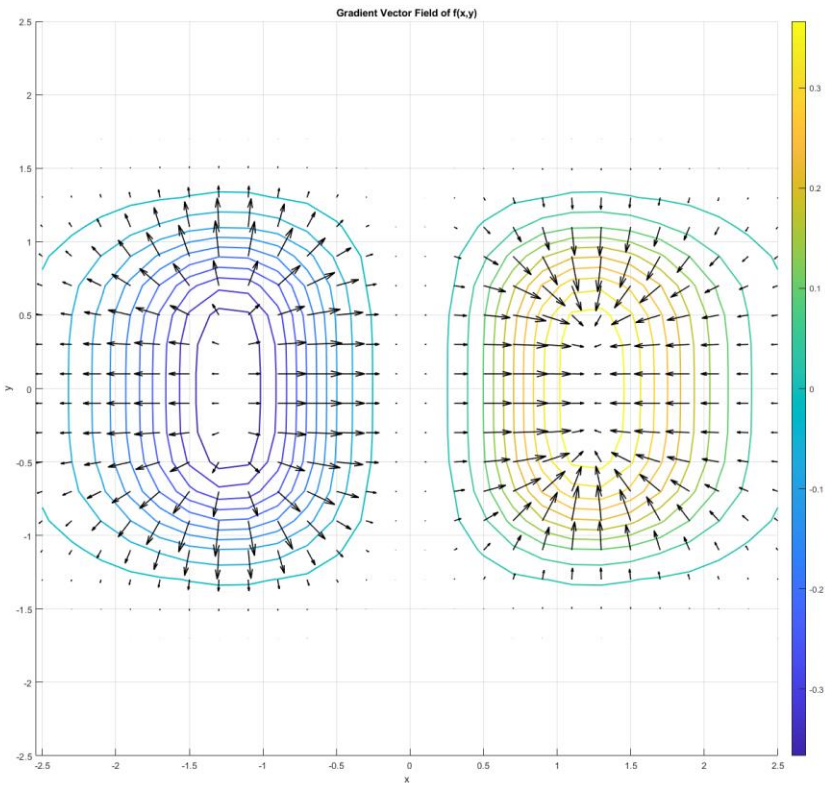
Για την καλύτερη οπτικοποίηση του προβλήματος, στο αρχείο plot_f σχεδιάζουμε την αντικειμενική συνάρτηση, και στο plot_grad_f.m σχεδιάζουμε την κατεύθυνση των κλίσεων της f. Παρακάτω φαίνονται τα αντίστοιχα διαγράμματα. Παρατηρούμε ότι το ελάχιστο έχει προσεγγιστικά τιμή λίγο μικρότερη από το -0.4.



Plot of f1(x, y) = x^3 * e^-(x^2 + y^4)
f(x, y)
Y-axis
X-axis



Plot of f1(x, y) = x^3 * e^-(x^2 + y^4)
f(x, y)
X-axis
Y-axis



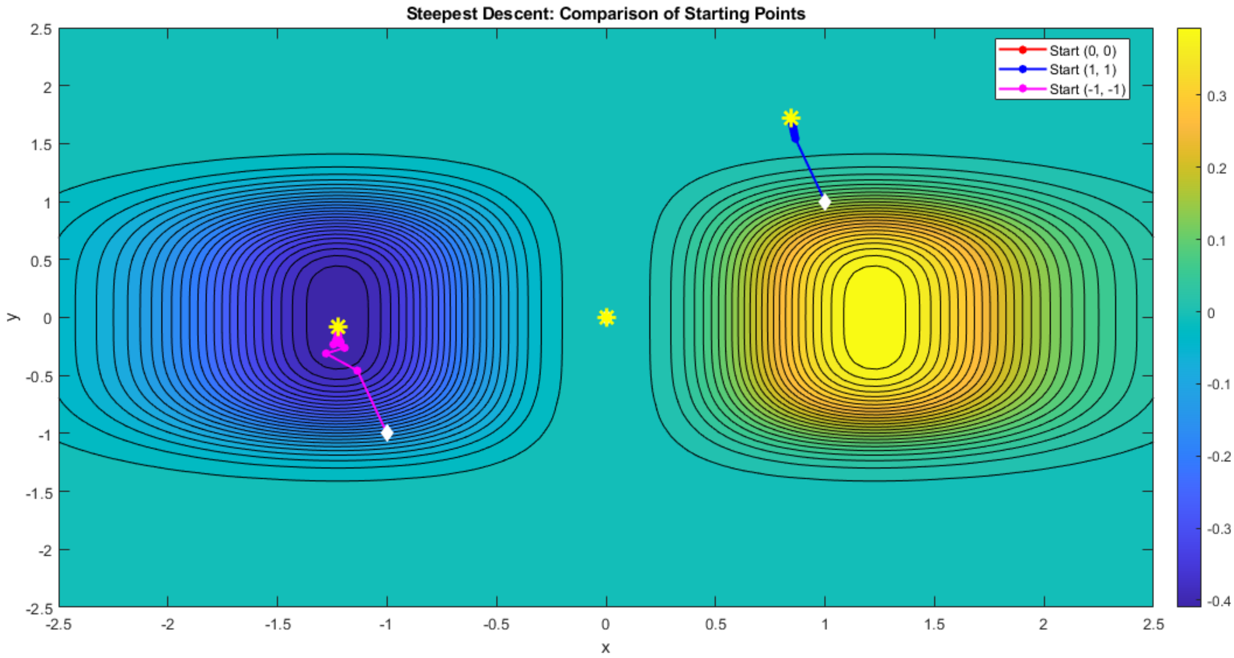
<u>Μέθοδος Μέγιστης Καθόδου (Steepest Descent)</u>

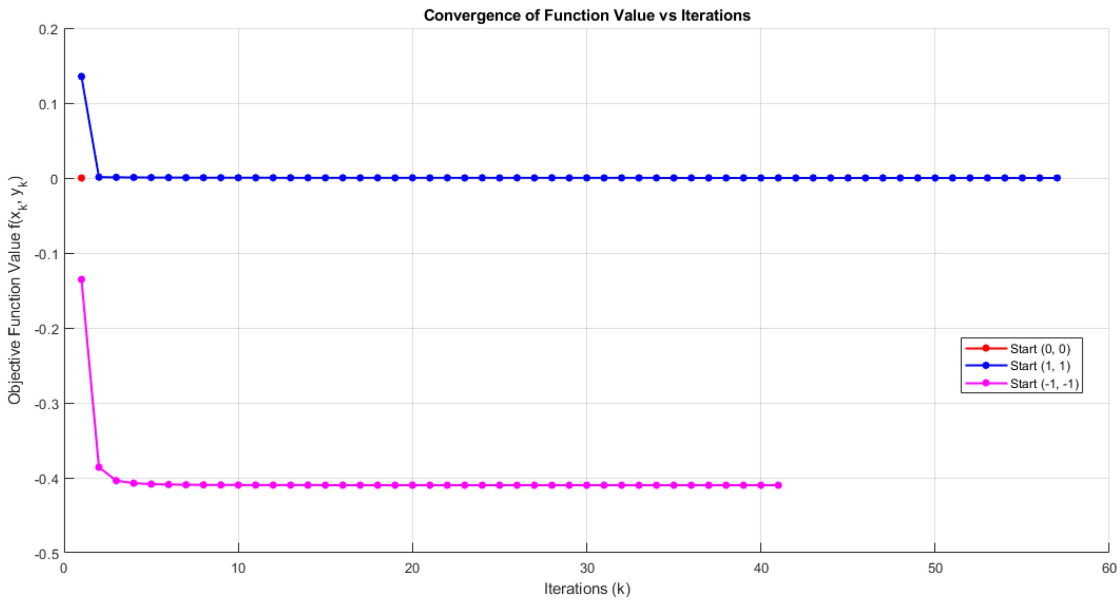
Σε αυτήν την μέθοδο ξεκινάμε από ένα τυχαίο σημείο της f. Σε κάθε επανάληψη του αλγορίθμου υπολογίζουμε την κλίση της συνάρτησης στο σημείο αυτό. Αν το μέτρο της κλήσης σε αυτό το σημείο είναι ικανοποιητικά μικρό ($|\nabla \mathrm{f}(x_k)| < \varepsilon$) τότε ο αλγόριθμος τερματίζει και το παρόν σημείο είναι το ελάχιστο. Διαφορετικά, υπολογίζουμε την κατεύθυνση αναζήτησης του $d_k = -\nabla \mathrm{f}(\mathrm{x}_k)$ του νέου σημείου, το οποίο δίνεται από τον τύπο $x_{k+1} = x_k + \gamma_\kappa * d_k$. Επαναλαμβάνουμε την διαδικασία, μέχρι να βρεθεί το ελάχιστο.

Παρακάτω φαίνονται τα αποτελέσματα της εφαρμογής του αλγορίθμου, με τις παραλλαγές του υπολογισμού του γ_κ και για 3 διαφορετικά σημεία εκκίνησης. Τα διαγράμματα δείχνουν την εξέλιξη των λύσεων του αλγορίθμου, από τα αρχικά έως τα τελικά ελάχιστα σημεία, και την σύγκλιση των λύσεων ως προς τον αριθμό επαναλήψεων του αλγορίθμου.

A) Σταθερό γ_κ

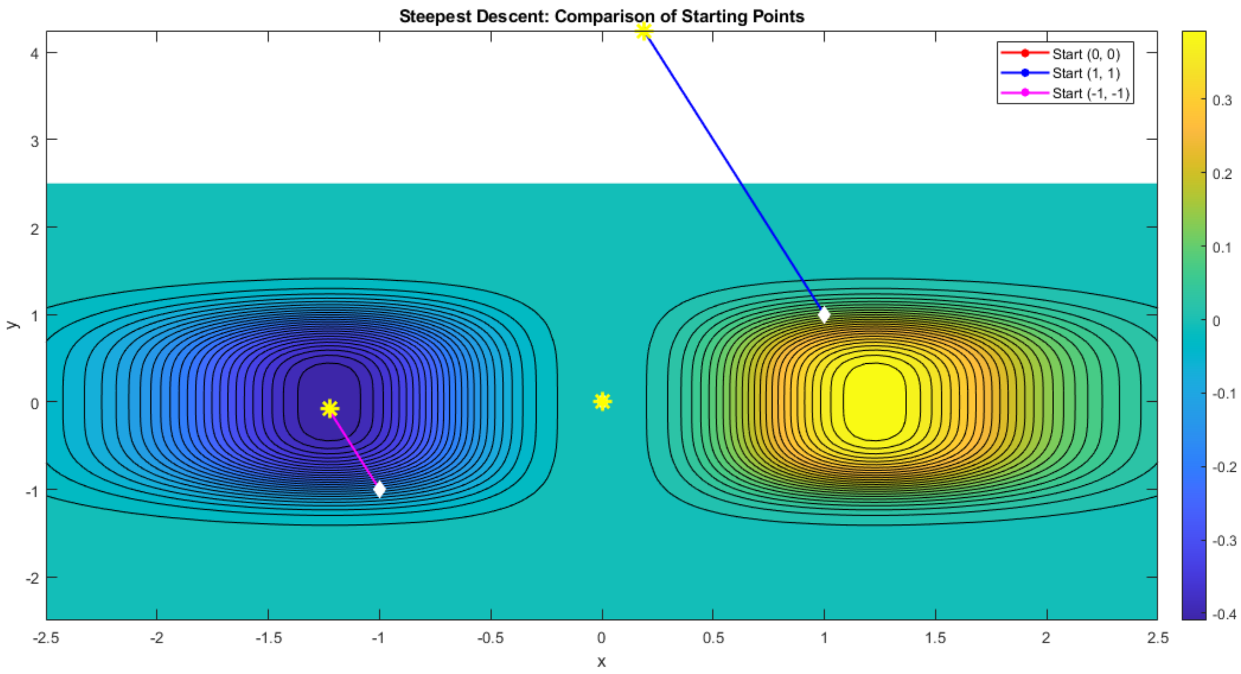
Σημείο εκκίνησης	Επαναλήψεις αλγορίθμου	Υπολογισμοί f	Υπολογισμοί ∇f	Ελάχιστο
(0,0)	1	1	1	0
(1,1)	57	1	57	$4.8 * 10^{-5}$
(-1, -1)	41	1	41	-0.409896

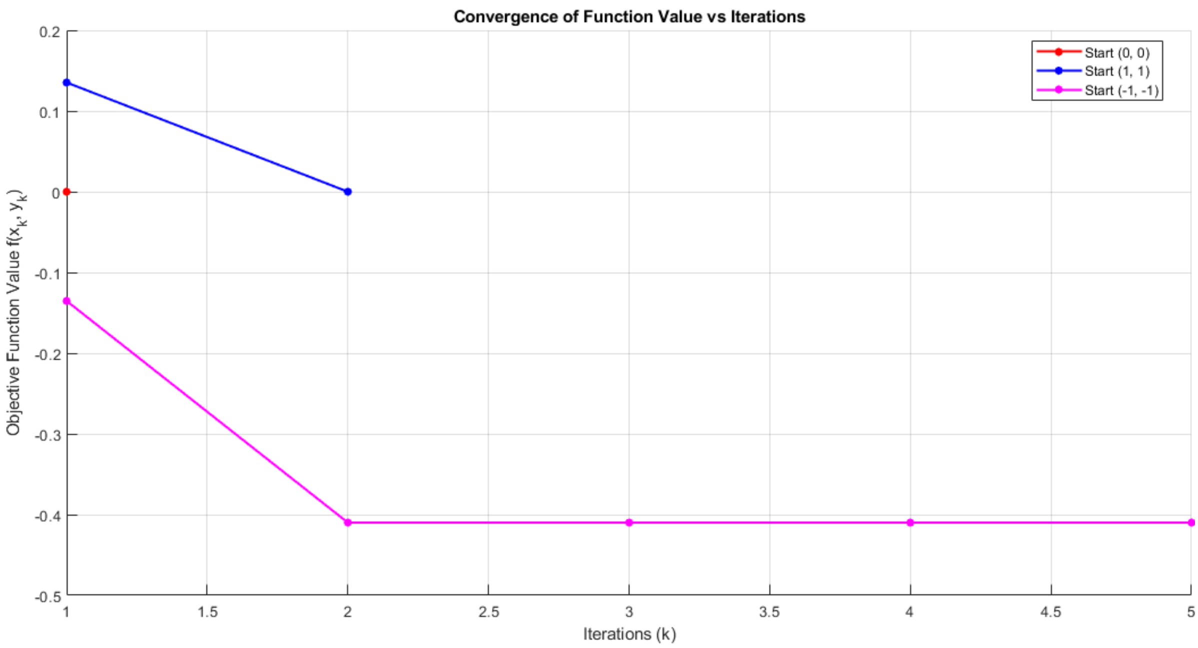




B) Το γ_κ υπολογίζεται σε κάθε επανάληψη ώστε να ελαχιστοποιείται η $f(x_k + \gamma_k * d_k)$, με τη χρήση της golden_section_search.

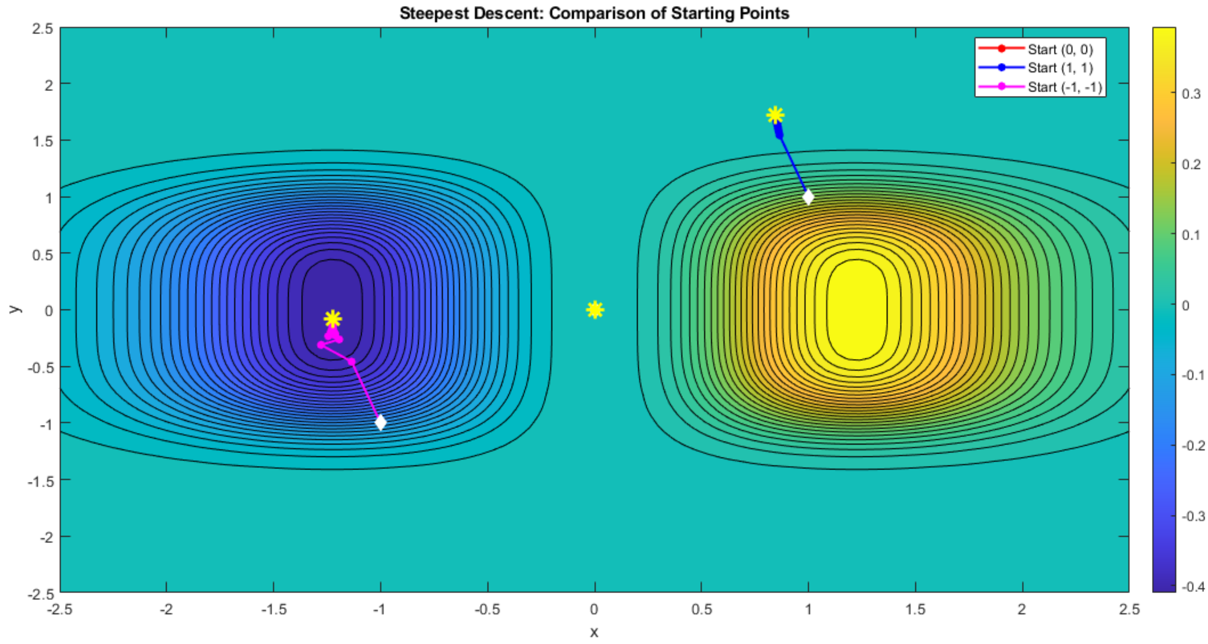
Σημείο εκκίνησης	Επαναλήψεις αλγορίθμου	Υπολογισμοί f	Υπολογισμοί ∇f	Ελάχιστο
(0,0)	1	1	1	0
(1,1)	2	22	2	$2.5 * 10^{-144}$
(-1, -1)	5	85	5	-0.409898

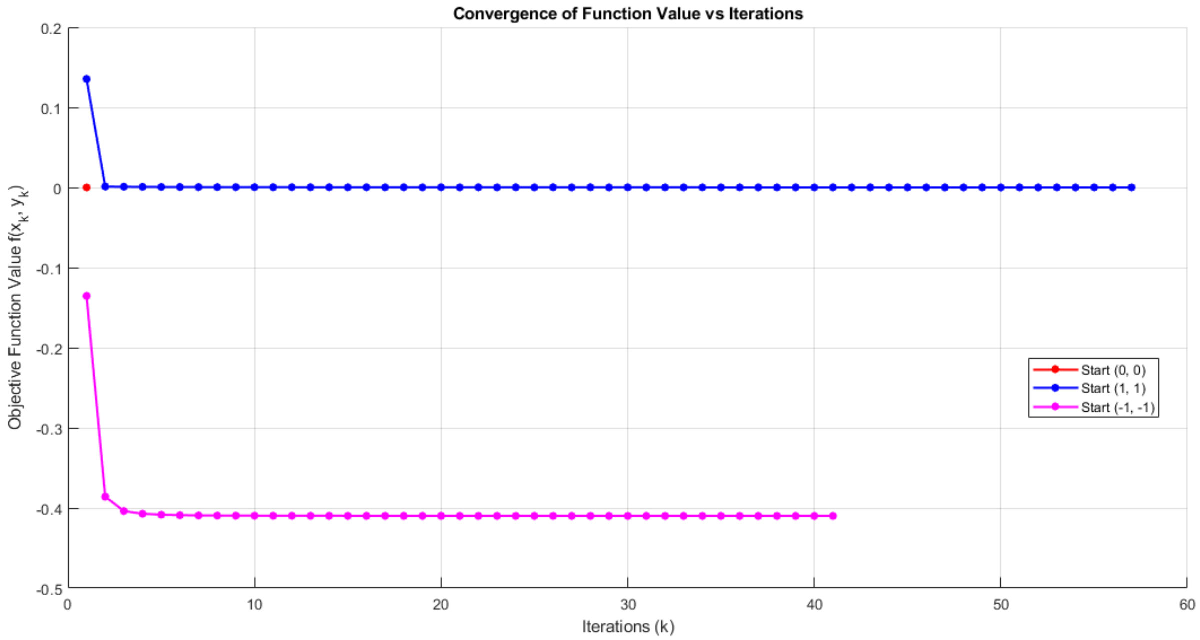




Γ) Το γ_κ υπολογίζεται σύμφωνα με τον κανόνα του Armijo.

Σημείο εκκίνησης	Επαναλήψεις αλγορίθμου	Υπολογισμοί f	Υπολογισμοί ∇f	Ελάχιστο
(0,0)	1	1	1	0
(1,1)	57	57	57	$4.8284 * 10^{-5}$
(-1, -1)	41	41	41	-0.409896





Παρατηρήσεις

Αρχικά, σε κάθε παραλλαγή του αλγορίθμου για το σημείο εκκίνησης (0,0) ο αλγόριθμος τερματίζει αμέσως αδυνατώντας να βρει το ελάχιστο. Αυτό είναι αναμενόμενο, αφού το (0,0) είναι στάσιμο σημείο με $|\nabla \mathrm{f}(x_k)| = 0$. Την ίδια συμπεριφορά αναμένουμε και στις επόμενες μεθόδους, αφού το κριτήριο τερματισμού τους είναι το ίδιο.

Ξεκινώντας από το σημείο (1,1) ο αλγόριθμος πάλι αδυνατεί να εντοπίσει το ελάχιστο. Λόγω της μορφής της συνάρτησης ο αλγόριθμος εγκλωβίζεται διότι η καμπύλη γίνεται πολύ επίπεδη στα σημεία αναζήτησης, με αποτέλεσμα η κλίση να γίνεται σχεδόν μηδενική. Έτσι ικανοποιείται η συνθήκη τερματισμού $|\nabla \mathrm{f}(x_k)| < \varepsilon$ και ο αλγόριθμος τερματίζει.

Ξεκινώντας από το (-1,-1) είναι η μόνος τρόπος που πέτυχαμε να βρούμε το ελάχιστο.

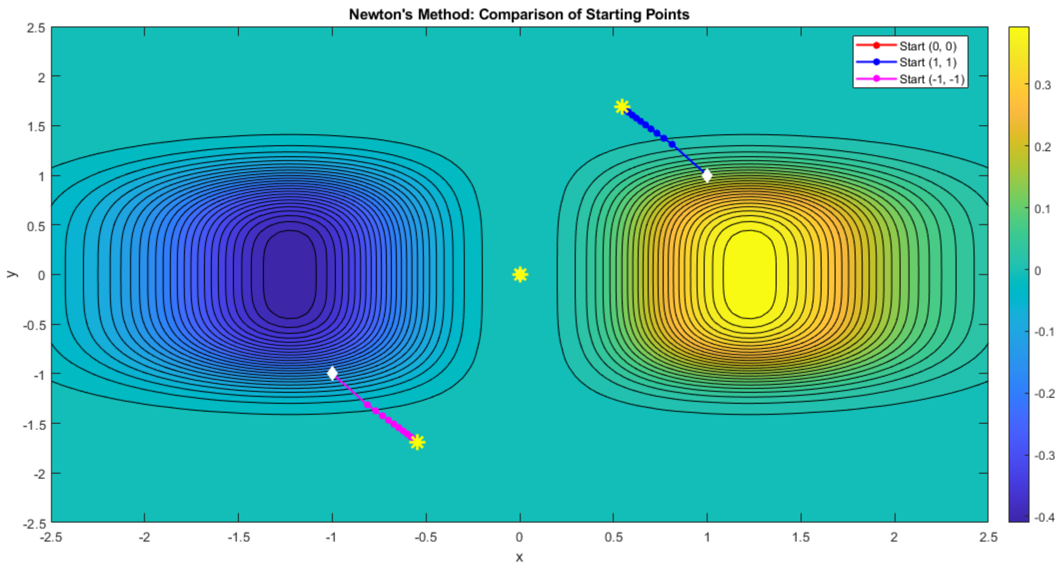
Χρησιμοποιώντας σταθερό γ_κ χρειάστηκαν 41 επαναλήψεις, και υπολογισμοί ∇f, αρκετά λιγότερες απ' ότι με golden section search, που χρειάστηκαν 5. Όμως στη δεύτερη περίπτωση χρειάστηκαν και έξτρα 84 υπολογισμοί της f. Στη συγκεκριμένη περίπτωση η παραλλαγή με τον κανόνα του Armijo μας δίνει ίδια αποτελέσματα με της παραλλαγή σταθερού γ_κ. Παρατήρησα πως σε κάθε επανάληψη το γ_κ υπολογιζόταν ίσο με 1. Η παραλλαγή του Armijo είναι υπολογιστικά ακριβότερη εφόσον χρειάστηκε έξτρα 40 υπολογισμούς της f. Από αυτό παρατηρούμε πόσο καλή ήταν η επιλογή του βήματος 1 στην πρώτη περίπτωση, διότι δοκιμάζοντας με διαφορετικά σταθερά βήματα, τα αποτελέσματα είναι πολύ χειρότερα. Η παραλλαγή του Armijo έχει το πλεονέκτημα ότι δεν χρειάζεται να γνωρίζουμε τι βήμα να ορίσουμε.

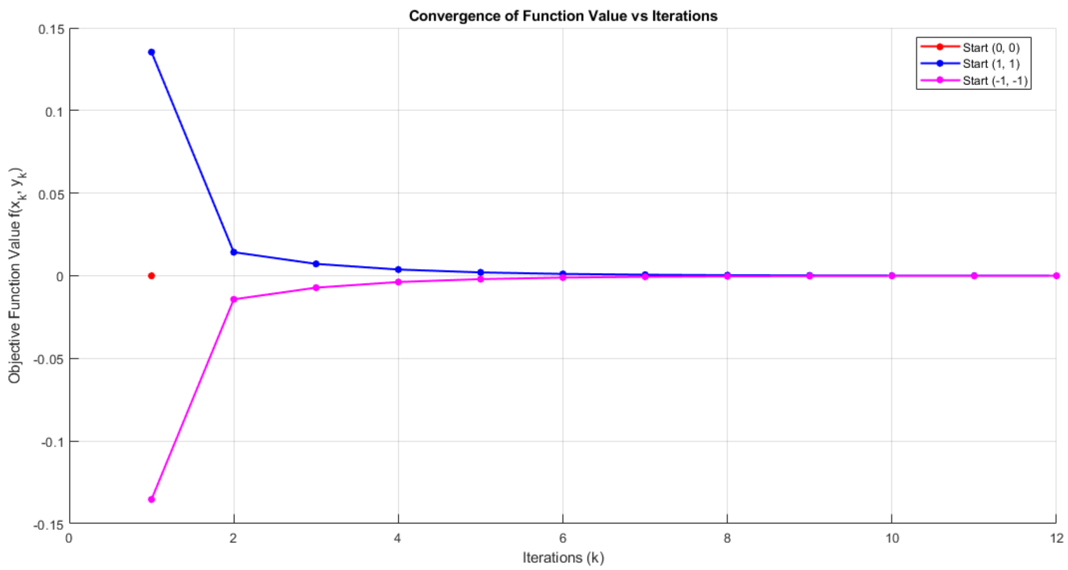
Μέθοδος Newton

Η μέθοδος αυτή είναι σχεδόν ίδια με την steepest descent, με την διαφορά ότι η κατεύθυνση αναζήτησης του νέου σημείου υπολογίζεται ως $d_k = -[\nabla^2 f(x_k)]^{-1} * \nabla f(x_k)$. Εδώ απαιτείτε επίσης ο υπολογισμός του εσσιανού πίνακα $\nabla^2 f(x_k)$, ο οποίος πρέπει να είναι θετικά ορισμένος. Ο τύπος υπολογισμού του κάθε νέου σημείου δίνεται πάλι από τον τύπο $x_{k+1} = x_k + \gamma_\kappa * d_k$.

Α) Σταθερό γ_κ

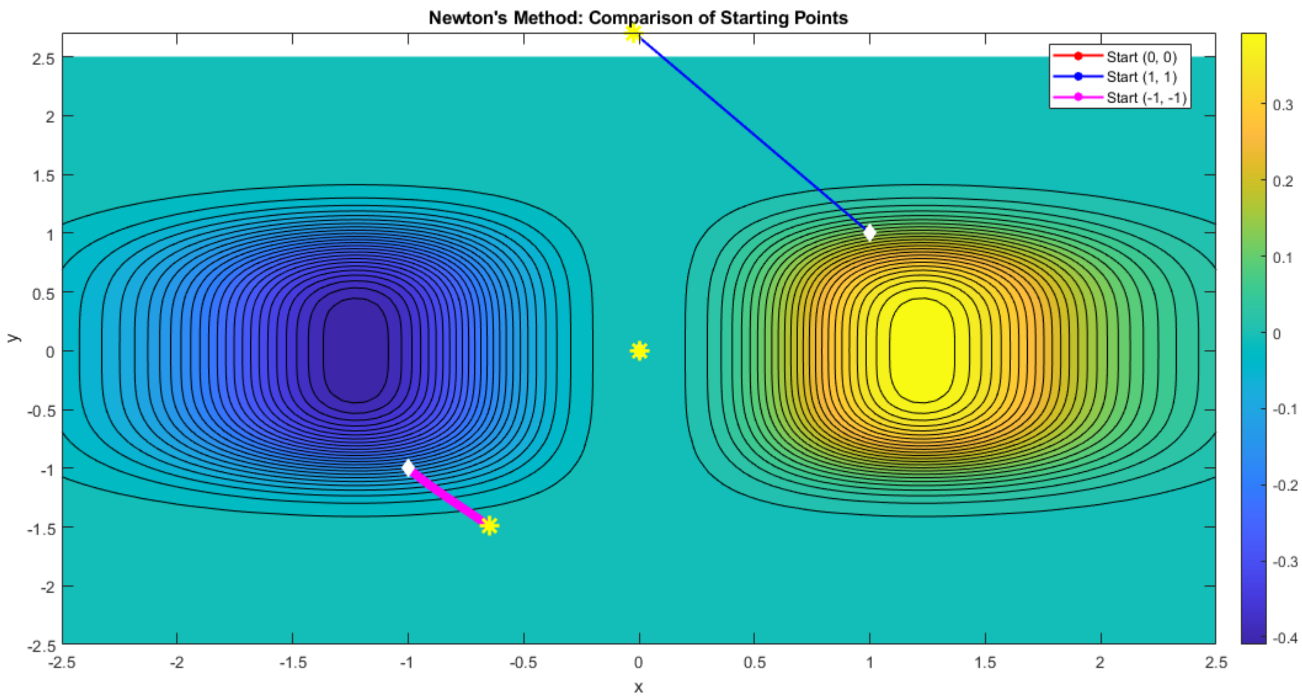
Σημείο εκκίνησης	Επαναλήψεις αλγορίθμου	Υπολογισμοί f	Υπολογισμοί ∇f	Υπολογισμοί ∇^2f	Ελάχιστο
(0,0)	1	1	1	0	0
(1,1)	12	1	12	11	$3.5 * 10^{-5}$
(-1, -1)	12	1	12	11	$-3.5 * 10^{-5}$

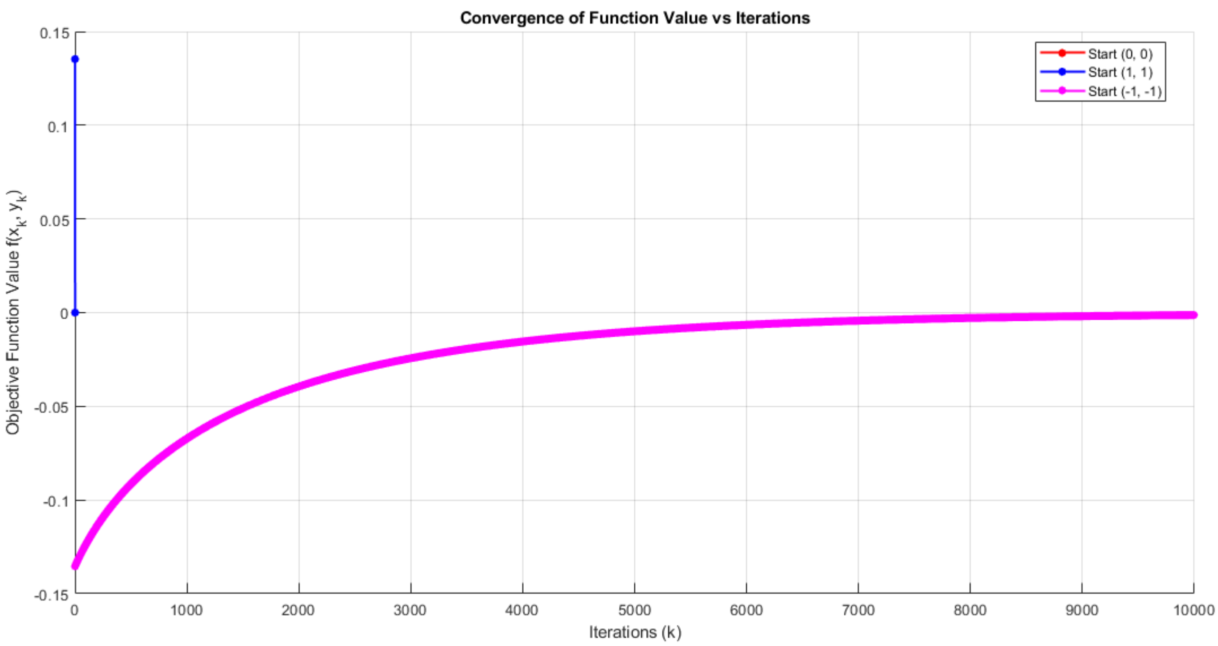




B) Το γ_κ υπολογίζεται σε κάθε επανάληψη ώστε να ελαχιστοποιείται η $f(x_k + \gamma_k * d_k)$, με τη χρήση της golden_section_search.

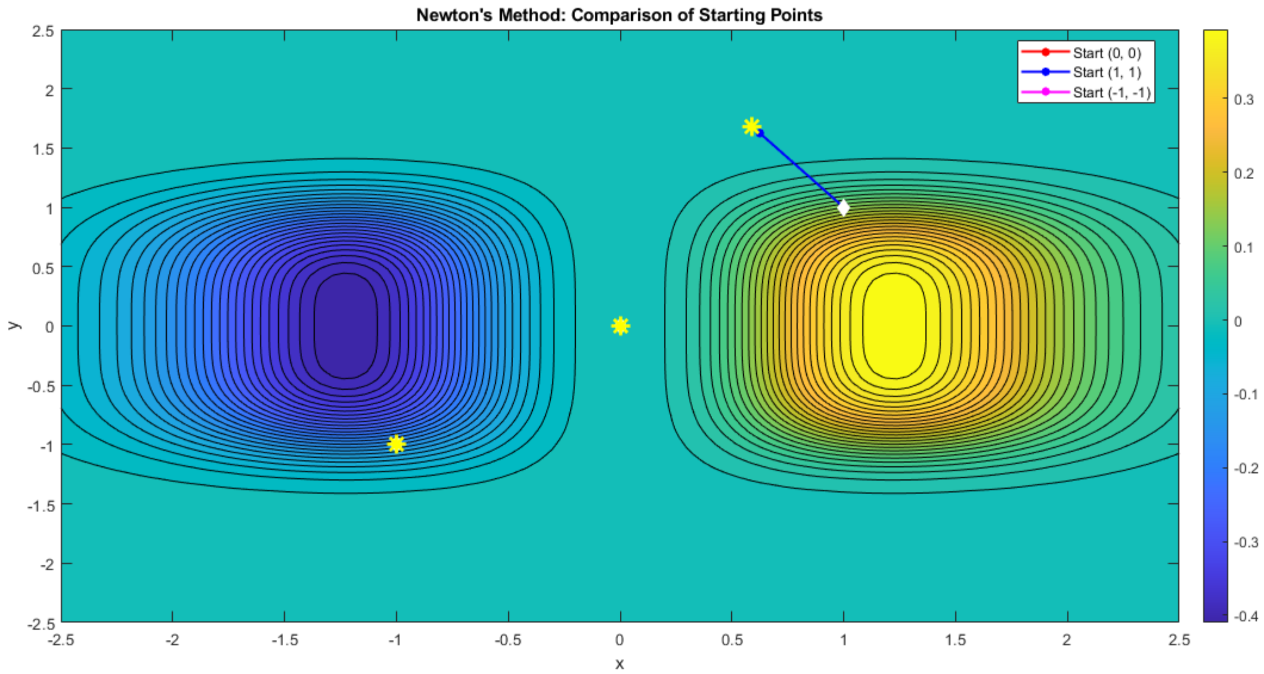
Σημείο εκκίνησης	Επαναλήψεις αλγορίθμου	Υπολογισμοί f	Υπολογισμοί ∇f	Υπολογισμοί $\nabla^2 f$	Ελάχιστο
(0,0)	1	1	1	0	0
(1,1)	2	21	2	1	$-6.8 * 10^{-29}$
(-1, -1)	10000	199981	10000	9999	$-12 * 10^{-4}$

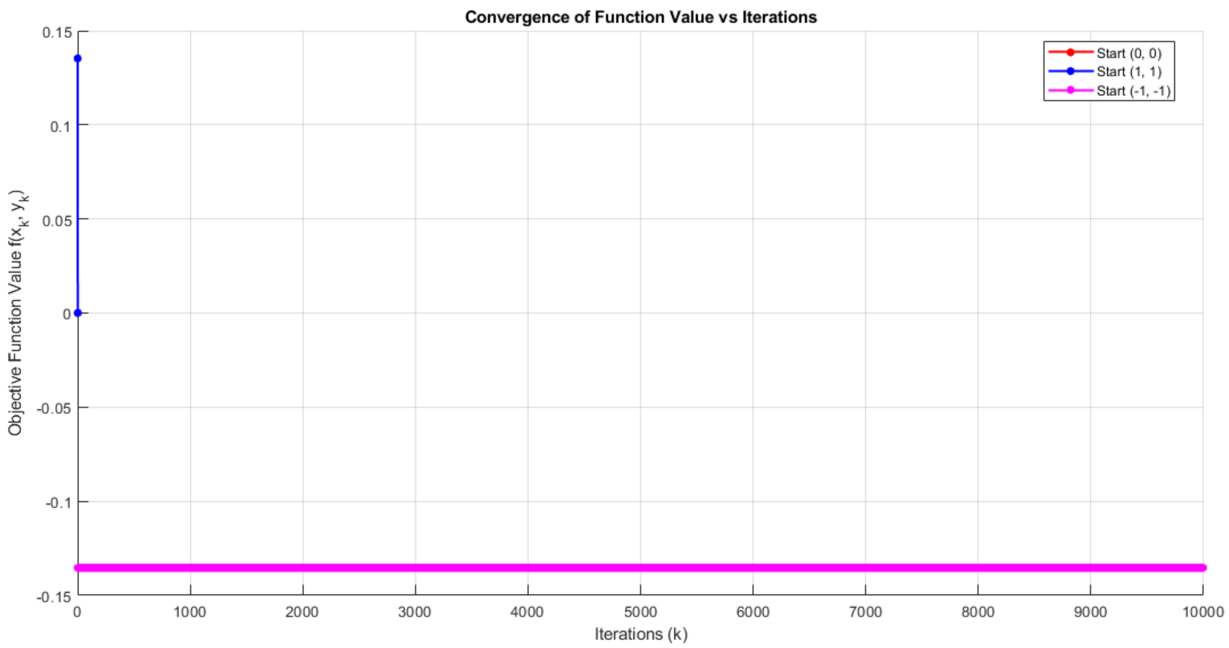




Γ) Το γ_κ υπολογίζεται σύμφωνα με τον κανόνα του Armijo.

Σημείο εκκίνησης	Επαναλήψεις αλγορίθμου	Υπολογισμοί f	Υπολογισμοί ∇f	Υπολογισμοί $\nabla^2 f$	Ελάχιστο
(0,0)	1	1	1	0	0
(1,1)	3	3	3	2	$4.8 * 10^{-5}$
(-1, -1)	10000	519949	10000	9999	-0.1353





Παρατηρήσεις

Όπως ήταν αναμενόμενο για το σημείο εκκίνησης (0,0) ο αλγόριθμος τερματίζει αμέσως όπως προηγουμένως.

Επίσης για το αρχικό σημείο (1,1) η συμπεριφορά του αλγορίθμου είναι πολύ παρόμοια με της steepest descent. Αυτό διακρίνεται εύκολα και από τα διαγράμματα στα οποία βλέπουμε ευδιάκριτα την "πορεία" του .

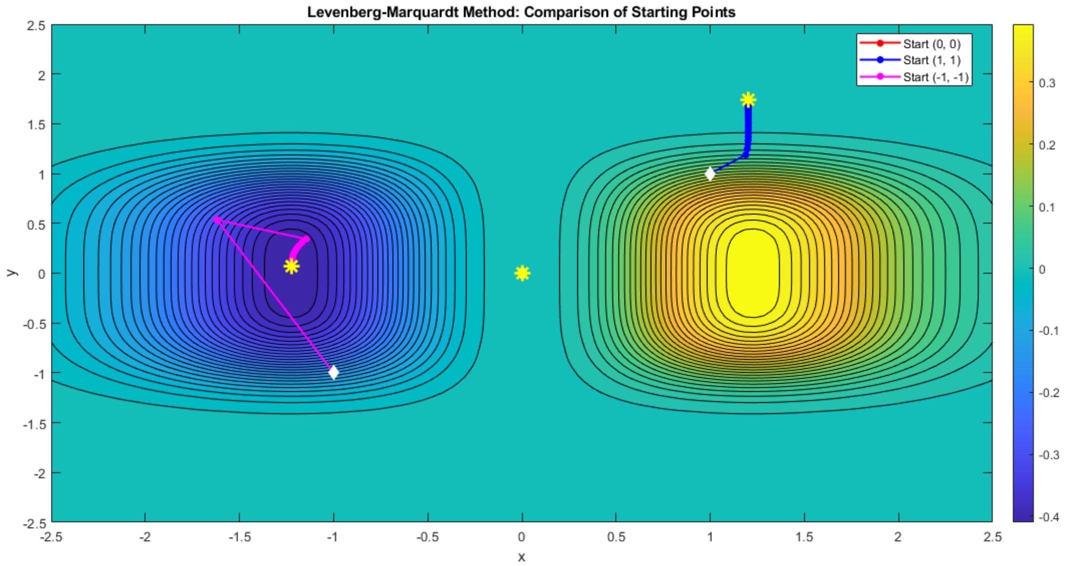
Ακόμα όμως και για το σημείο (-1,-1) τα αποτελέσματα είναι απογοητευτικά. Εκτελέστηκε ο μέγιστος αριθμός επαναλήψεων και δεν επιτεύχθηκε η εύρεση του ελαχίστου. Από τα διαγράμματα μάλιστα φαίνεται πως καθώς εκτελείται ο αλγόριθμος το σημείο απομακρύνεται από το ελάχιστο αντί να το πλησιάζει. Ακόμη και με την χρήση των golden section search και Armijo το πρόβλημα δεν διορθώθηκε. Αυτό οφείλεται στο ότι ο Εσσιανός πίνακας δεν είναι θετικά ορισμένος στα σημεία που υπολογίστηκε με αποτέλεσμα η κατεύθυνση d_k να μην οδηγεί σε κάθοδο της συνάρτησης.

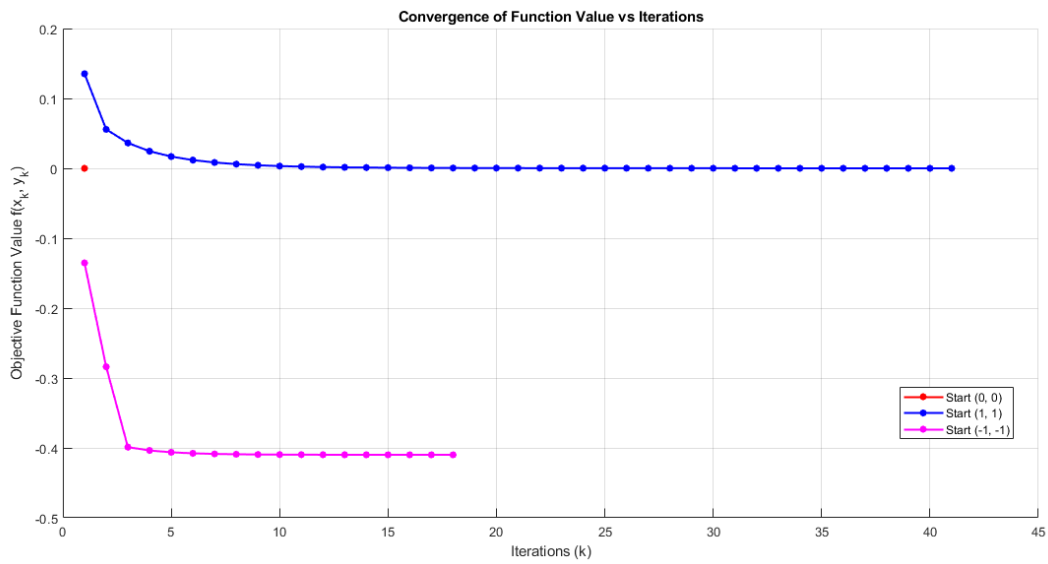
Μέθοδος Levenberg-Marquardt

Η παρούσα μέθοδος καταφέρνει να αντιμετωπίσει το πρόβλημα που αντιμετωπίσαμε προηγουμένως στη μέθοδο Newton (μη θετικά ορισμένοι εσσιανοί πίνακες). Η μέθοδος Levenberg-Marquardt τροποποιεί τον εσσιανό εξασφαλίζοντας να είναι πάντα θετικά ορισμένος. Με αυτόν τον τρόπο εγγυάται πως πάντα η κατεύθυνση αναζήτησης d_k θα έχει κατεύθυνση καθόδου.

Α) Σταθερό γ_k

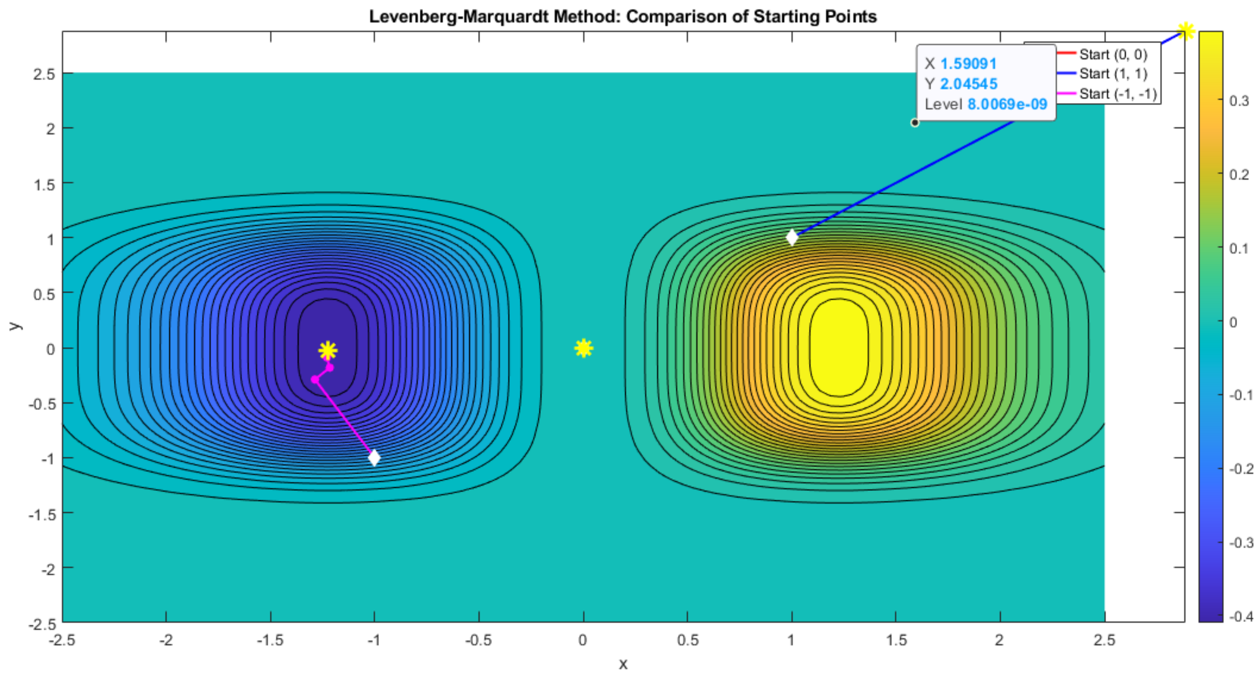
Σημείο εκκίνησης	Επαναλήψεις αλγορίθμου	Υπολογισμοί f	Υπολογισμοί ∇f	Υπολογισμοί $\nabla^2 f$	Ελάχιστο
(0,0)	1	1	1	0	0
(1,1)	41	1	41	40	$4.6 * 10^{-5}$
(-1, -1)	18	1	18	17	-0.409906

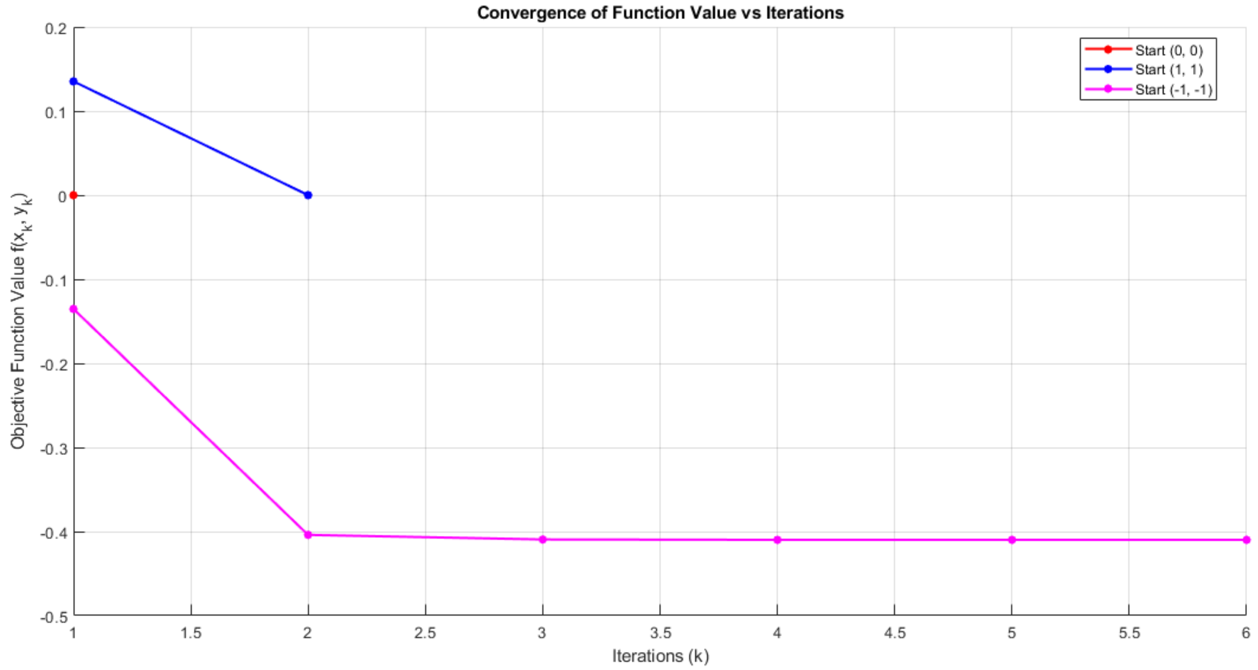




B) Το γ_k υπολογίζεται σε κάθε επανάληψη ώστε να ελαχιστοποιείται η $f(x_k + \gamma_k * d_k)$, με τη χρήση της golden_section_search.

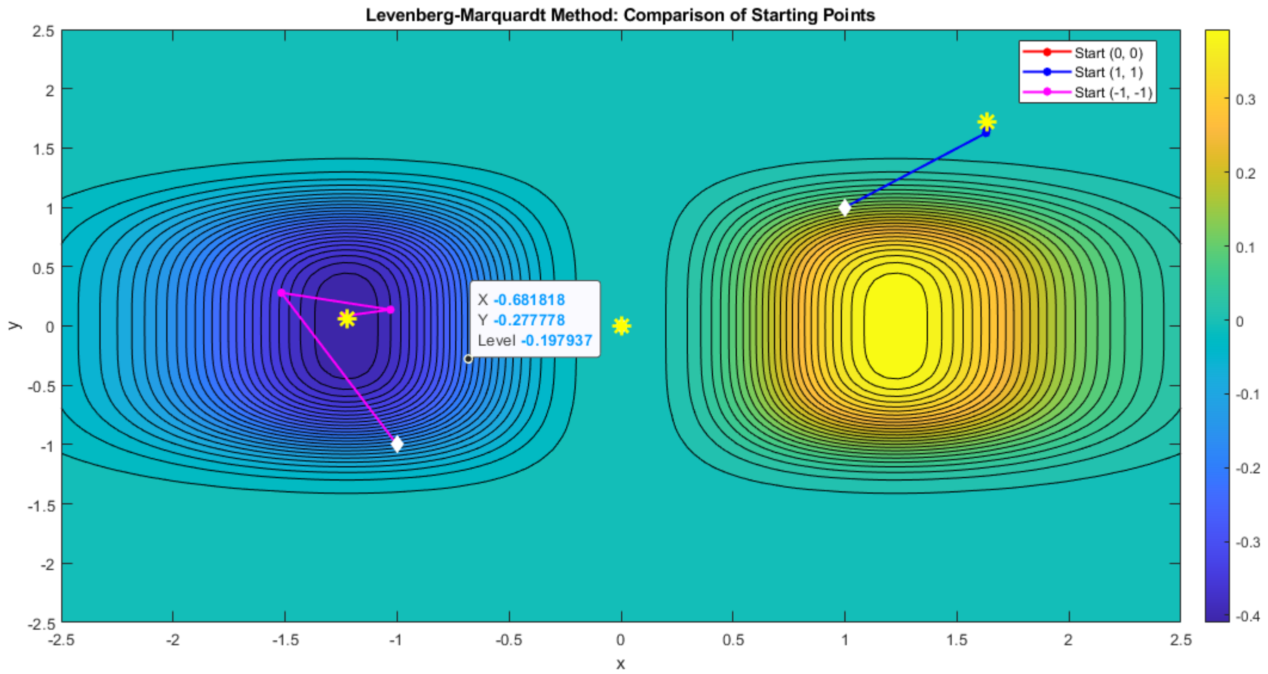
Σημείο εκκίνησης	Επαναλήψεις αλγορίθμου	Υπολογισμοί f	Υπολογισμοί ∇f	Υπολογισμοί $\nabla^2 f$	Ελάχιστο
(0,0)	1	1	1	0	0
(1,1)	2	20	2	1	$-7.3 * 10^{-33}$
(-1, -1)	6	96	6	5	-0.409915

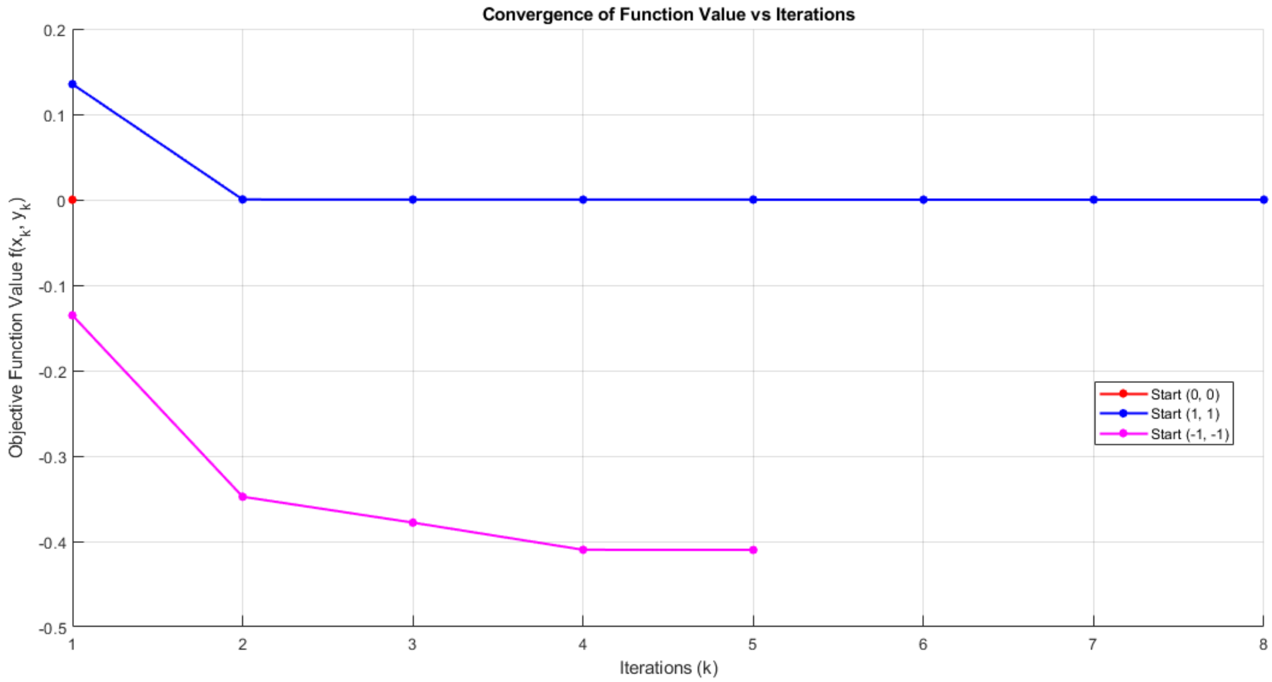




Γ) Το γ_κ υπολογίζεται σύμφωνα με τον κανόνα του Armijo.

Σημείο εκκίνησης	Επαναλήψεις αλγορίθμου	Υπολογισμοί f	Υπολογισμοί ∇f	Υπολογισμοί $\nabla^2 f$	Ελάχιστο
(0,0)	1	1	1	0	0
(1,1)	8	8	8	7	$4.4 * 10^{-5}$
(-1, -1)	5	7	5	4	-0.409912





Παρατηρήσεις

Για αρχικό σημείο (0,0) εγκλωβίζεται όπως είναι αναμενόμενο.

Για αρχικό σημείο (1,1) ο αλγόριθμος εγκλωβίζεται για τον ίδιο λόγο όπως και στις προηγούμενες περιπτώσεις.

Για αρχικό σημείο (-1,-1) ο αλγόριθμος κατάφερε να συγκλίνει στο ελάχιστο και στις τρείς παραλλαγές του. Η βέλτιστη παραλλαγή αποδείχτηκε αυτή που χρησιμοποιεί τον κανόνα Armijo εφόσον χρειάστηκε τους λιγότερους υπολογισμούς στις f, $\nabla^2 f$, και $\nabla^2 f$. Η παραλλαγή με την golden section search είχε επίσης ικανοποιητικά αποτελέσματα, αν και το κόστος υπολογισμών της f είναι αρκετά μεγαλύτερο. Με σταθερό γ_κ ήταν επίσης ικανοποιητικά τα αποτελέσματα αλλά έχει μεγαλύτερο κόστος για τον υπολογισμό των $\nabla^2 f$ και $\nabla^2 f$.

Συνοψίζοντας τα πειραματικά δεδομένα, η μέθοδος Levenberg-Marquardt αναδείχθηκε ως ο αποδοτικότερος και πιο στιβαρός αλγόριθμος. Κατάφερε να διορθώσει την αρνητική καμπυλότητα του Εσσιανού πίνακα στο σημείο (-1,-1) επιτυγχάνοντας σύγκλιση εκεί όπου η μέθοδος Newton απέτυχε, ενώ παράλληλα διατήρησε πολύ υψηλότερη ταχύτητα σύγκλισης από την Steepest Descent.