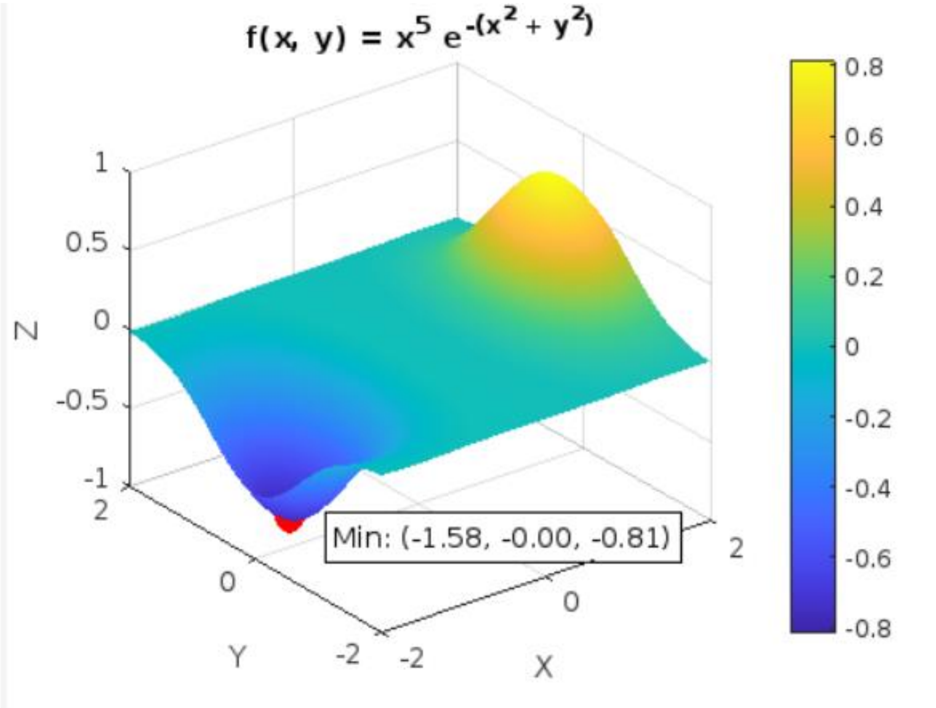
1η Εργαστηριακή Άσκηση

Χρυσολόγου Γεώργιος (ΑΕΜ: 10782)

Περιγραφή προβλήματος: Για μία δοσμένη συνάρτηση f δύο μεταβλητών ($f: R^2 \to R$),η οποία είναι δύο φορές παραγωγίσιμη, καλούμαστε να εφαρμόσουμε τρεις διαφορετικές μεθόδους κλίσης με σκοπό την ελαχιστοποίηση της. Οι μέθοδοι αυτοί βασίζονται στην ιδέα της επαναληπτικής καθόδου και είναι η Μέθοδος της Μέγιστης Καθόδου, η Μέθοδος Newton και η Μέθοδος Levenberg-Marquardt.
$f(x,y) = x^5 * e^{-x^2-y^2}$

Θέμα 1 (Σχεδίαση της f)



Παραπάνω απεικονίζεται η γραφική παράσταση της f για $-2 \leq x \leq 2$ και $-2 \leq y \leq 2$. Παρατηρούμε ότι η συνάρτηση ελαχιστοποιείται για $x = -1.58$, $y = 0$ όπου παίρνει τη τιμή $f(-1.58, 0) = -0.81$.

Θέμα 2 (Μέθοδος Μέγιστης Καθόδου)

Εξήγηση τρόπου λειτουργίας: Κατά την εφαρμογή της μεθόδου αυτής, επιλέγεται σε κάθε επανάληψη k+1 ένα νέο σημείο x_{k+1} με βάση το αντίστοιχο σημείο της προηγούμενης επανάληψης k, σύμφωνα με την σχέση $x_{k+1} = x_k + \gamma_k * d_k$, όπου $d_k = -\nabla f(x_k)$ το διάνυσμα κατεύθυνσης. Αναλυτικότερα, επιλέγεται, αρχικά, μία τιμή για την σταθερά τερματισμού ε καθώς και το σημείο έναρξης x_1. Στην αρχή κάθε επανάληψης k ελέγχεται εάν ικανοποιείται η ανισότητα $|\nabla f(x_k)| < \varepsilon$, γεγονός που οδηγεί στον τερματισμό της μεθόδου. Κατόπιν, σε περίπτωση δηλαδή που η ανισότητα δεν ικανοποιείται, ορίζεται το διάνυσμα κατεύθυνσης ως $d_k = -\nabla f(x_k)$, επιλέγεται η τιμή του γ_k (με τον έναν εκ των τριών τρόπων που ζητούνται) και υπολογίζεται το σημείο που θα εξεταστεί στην επόμενη επανάληψη $x_{k+1} = x_k + \gamma_k * d_k$. Η διαδικασία αυτή, με εξαίρεση την αρχικοποίηση που αναφέρθηκε πρώτη, ακολουθείται σε κάθε επανάληψη ώσπου να ικανοποιηθεί η παραπάνω ανισότητα. Με την Μέθοδο Μέγιστης Καθόδου, συμπερασματικά, το σημείο x κάθε νέας επανάληψης ορίζεται με τέτοιο τρόπο ώστε να πλησιάζει ολοένα και πιο κοντά στο σημείο του ελαχίστου της αντικειμενικής συνάρτησης, ώσπου η τιμή $|\nabla f(x_k)|$ να είναι όσο επιθυμούμε μικρή.

Επιλέχθηκε η τιμή $\varepsilon = 0.01$ για την σταθερά τερματισμού.

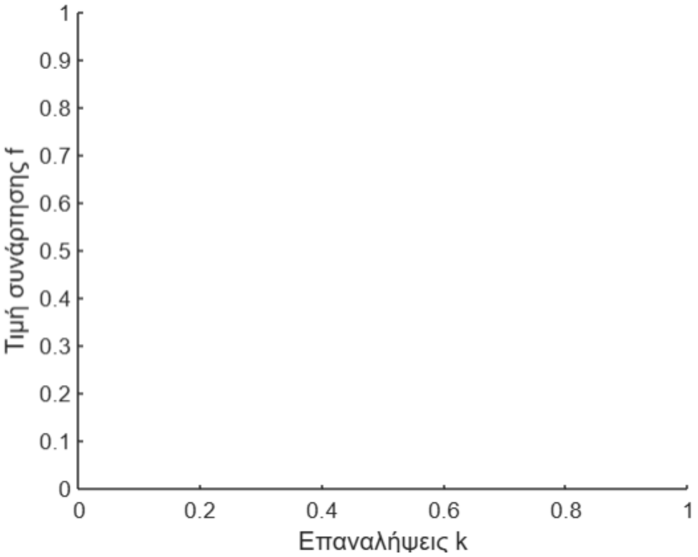
Όσον αφορά την επιλογή γ_k τέτοιου ώστε να ελαχιστοποιείται η $f(x_k + \gamma_k * d_k)$, το κατάλληλο γ_k υπολογίστηκε με χρήση της μεθόδου διχοτόμησης, εφαρμοσμένης στην παραπάνω συνάρτηση με τα x_kκαι d_k γνωστά σε κάθε επανάληψη. Για την μέθοδο διχοτόμησης, επιλέχθηκαν $l = 0.01$, $\varepsilon = 0.001$ και αρχικό διάστημα αναζήτησης το διάστημα [0, 1].

Όσον αφορά την επιλογή γ_k βάσει του κανόνα Armijo, επιλέχθηκαν $\alpha = 0.01$, $\beta = 0.5$ και $s = 0.5$.

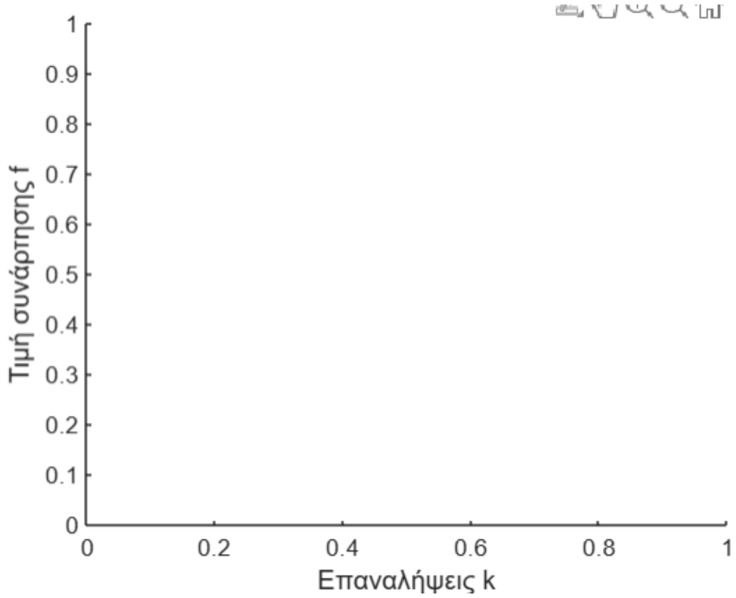
Ακολουθούν οι γραφικές παραστάσεις της σύγκλισης της αντικειμενικής συνάρτησης προς τον αριθμό των επαναλήψεων για τα διαφορετικά αρχικά σημεία και τους διαφορετικούς τρόπους επιλογής του γ_k.

1) Γραφικές παραστάσεις για αρχικό σημείο (0, 0):

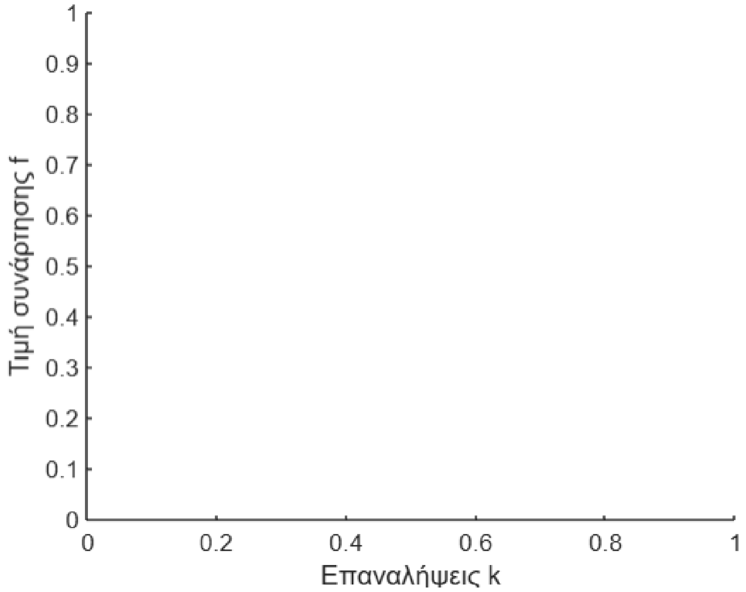
i) Επιλογή σταθερού $\gamma_k = 0.01$



ii) Επιλογή γ_k τέτοιου ώστε να ελαχιστοποιείται η $f(x_k + \gamma_k * d_k)$



iii) Επιλογή γ_k βάσει του κανόνα Armijo

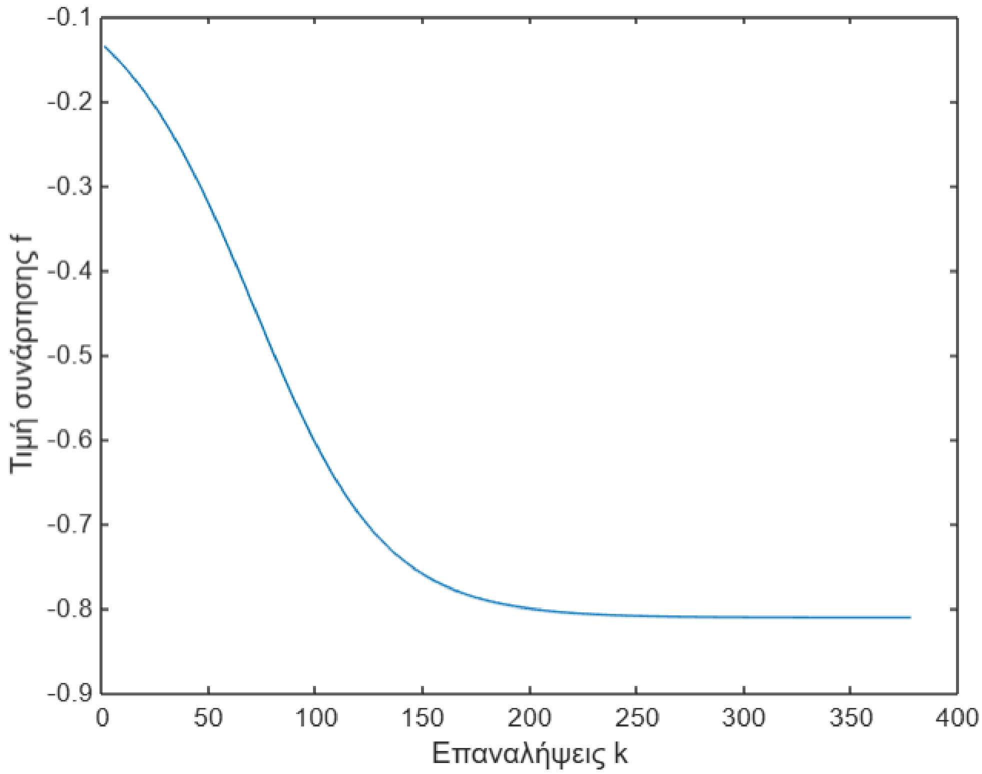


Παρατηρούμε ότι η επιλογή του σημείου (0, 0) ως σημείο έναρξης έχει ως αποτέλεσμα την πραγματοποίηση μόνο μίας επανάληψης της μεθόδου και την αδυναμία σύγκλισης της μεθόδου. Το γεγονός αυτό είναι λογικό καθώς κατά την 1η επανάληψη, όπου $x_1 = (0,0)$, ισχύει $|\nabla f(x_1)| = 0$. Ως συνέπεια, ικανοποιείται η ανισότητα $|\nabla f(x_1)| < \varepsilon$ και ο αλγόριθμος τερματίζεται. Γίνεται αντιληπτό, λοιπόν, ότι το σημείο έναρξης αυτό συνιστά κακή επιλογή όσον αφορά την επίλυση του προβλήματος ελαχιστοποίησης.

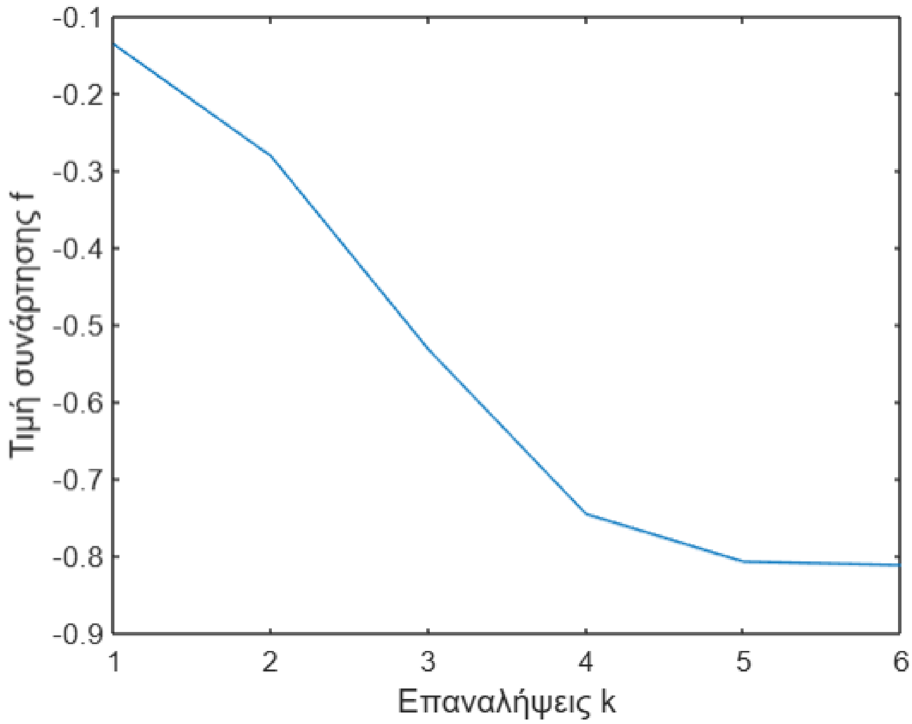
2) Γραφικές παραστάσεις για αρχικό σημείο (-1, 1):

i) Επιλογή σταθερού γ_k

α) $\gamma_k = 0.01$

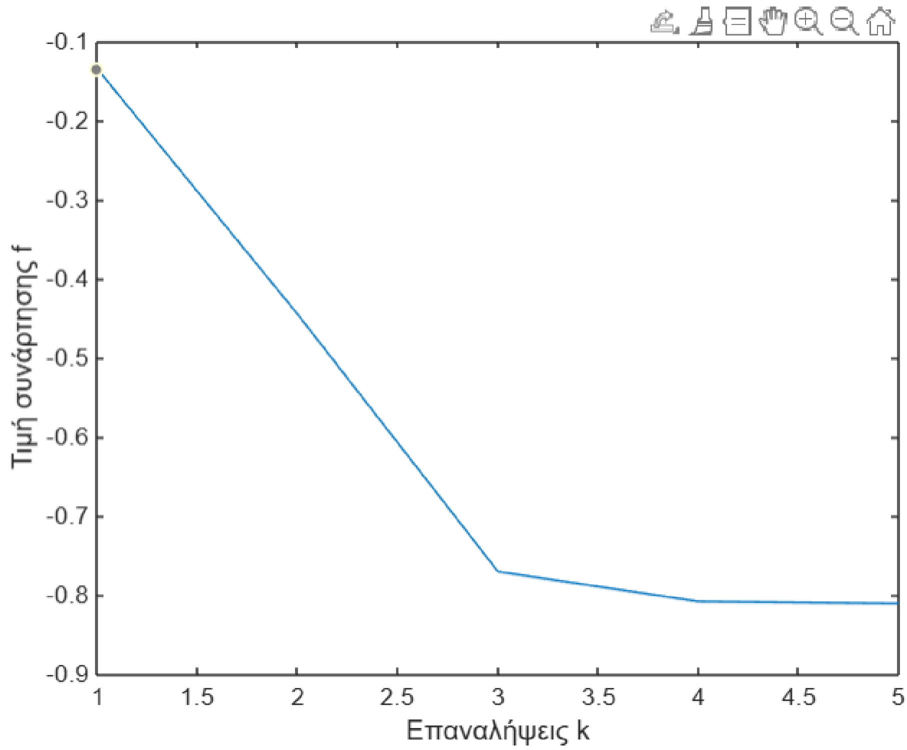


β) $\gamma_k = 0.5$

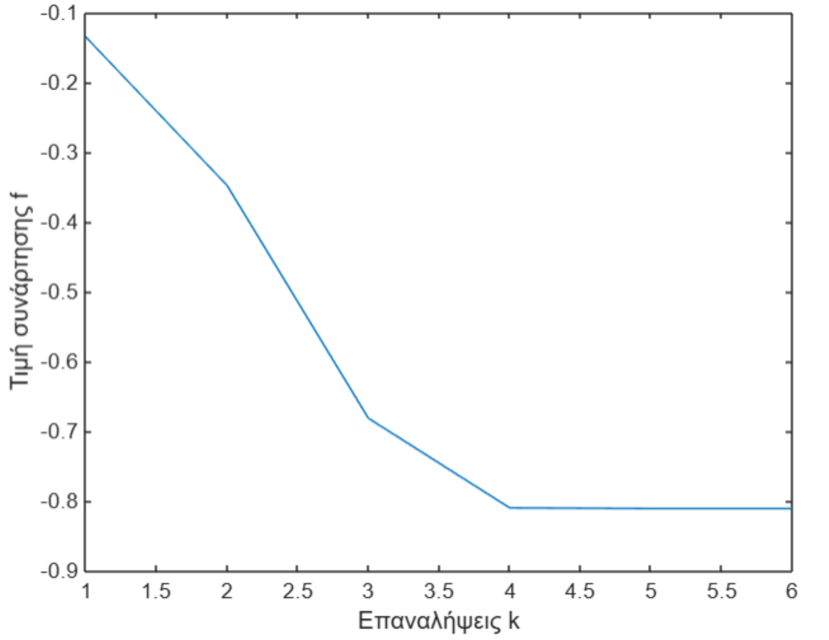


Παρατηρούμε ότι με κατάλληλη επιλογή της σταθερής τιμής του γ_k ο συνολικός αριθμός επαναλήψεων μειώνεται ραγδαία.(Για δοκιμές με τιμές $\gamma_k > 0.5$ η μέθοδος συνέκλινε με πιο αργό ρυθμό. Συμπεραίνουμε λοιπόν ότι η επιλογή $\gamma_k = 0.5$ είναι η βέλτιστη (ή προσεγγίζει την βέλτιστη).

ii) Επιλογή γ_k τέτοιου ώστε να ελαχιστοποιείται η $f(x_k + \gamma_k * d_k)$



iii) Επιλογή γ_k βάσει του κανόνα Armijo

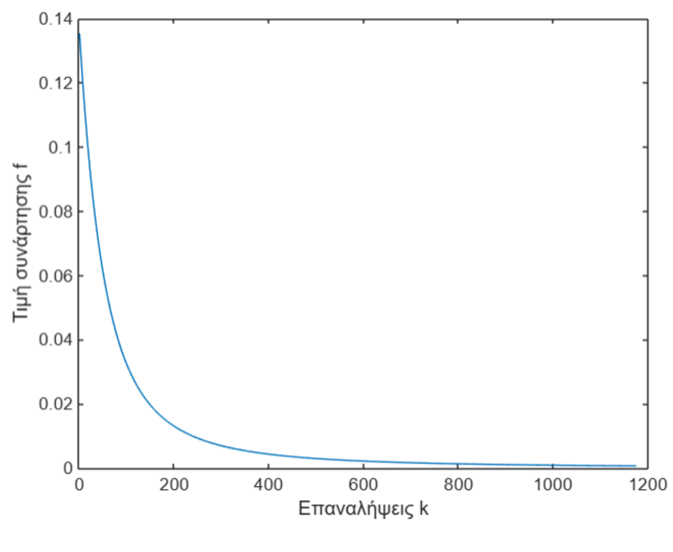


Παρατηρούμε ότι η χρήση και των τριών τρόπων επιλογής του βήματος γ_k οδηγεί σε σύγκλιση της μεθόδου, καθώς η τιμή της συνάρτησης τείνει προς το ελάχιστο της -0.81. Συγκεκριμένα, η επιλογή σταθερού βήματος $\gamma_k = 0.01$ προκαλεί την πιο αργή σύγκλιση εκ των τριών, απαιτώντας 378 επαναλήψεις για τον τερματισμό της μεθόδου. Αξίζει να σημειωθεί, όμως, ότι η πραγματοποίηση πολλών δοκιμών με διαφορετικές τιμές σταθερού βήματος μπορεί να οδηγήσει στην εύρεση της βέλτιστης τιμής του, καθιστώντας την επιλογή σταθερό βήματος εξίσου αποτελεσματικό τρόπο σε σχέση με τους υπόλοιπους δύο τρόπους επιλογής βήματος (όπως και συνέβη για $\gamma_k = 0.5$). Μεταξύ των τελευταίων, ταχύτερη σύγκλιση προκαλείται μέσω της επιλογής του γ_k ώστε να ελαχιστοποιείται η $f(x_k + \gamma_k * d_k)$ καθώς η μέθοδος τερματίζεται σε 5 επαναλήψεις, έναντι των 6 επαναλήψεων με επιλογή του βήματος βάσει του κανόνα Armijo. Όσον αφορά τον τελευταίο, πραγματοποιήθηκαν πολλές δοκιμές με διάφορες τιμές των παραμέτρων α, β, s και οι συγκεκριμένες επιλογές οδήγησαν στον μικρότερο αριθμό επαναλήψεων.

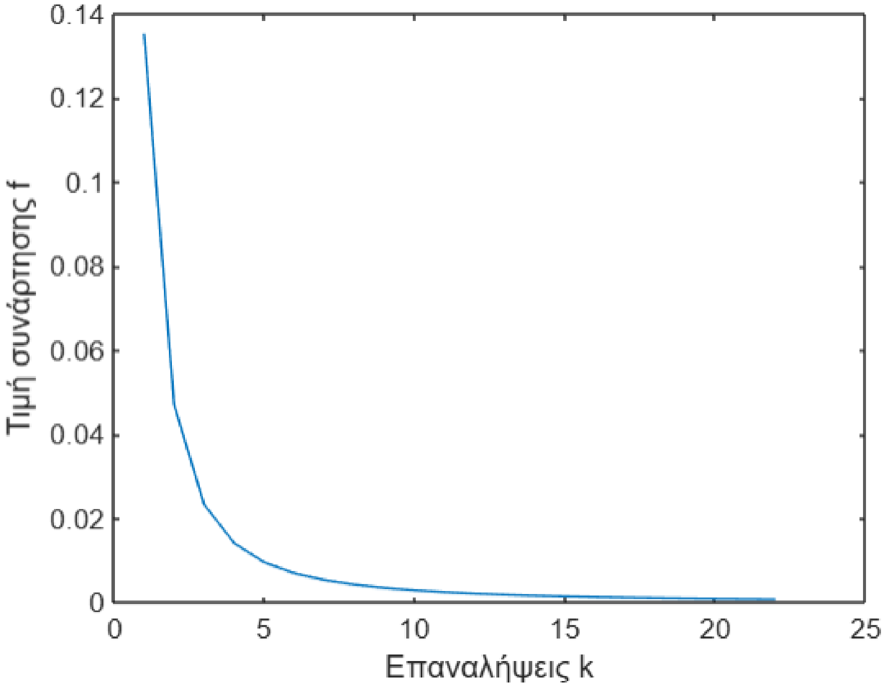
3) Γραφικές παραστάσεις για αρχικό σημείο (1, -1):

i) Επιλογή σταθερού γ_k

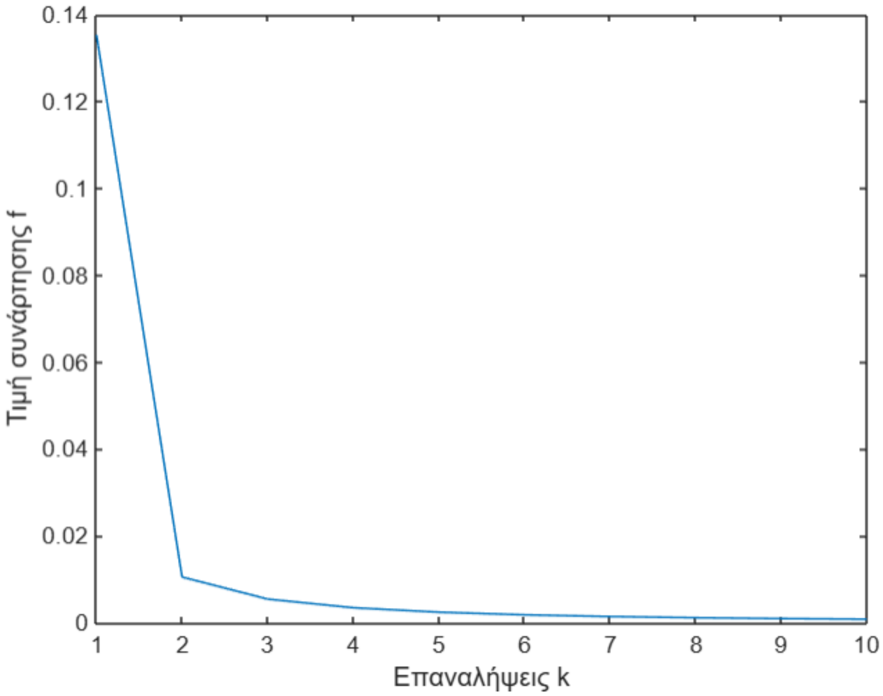
α) $\gamma_k = 0.01$



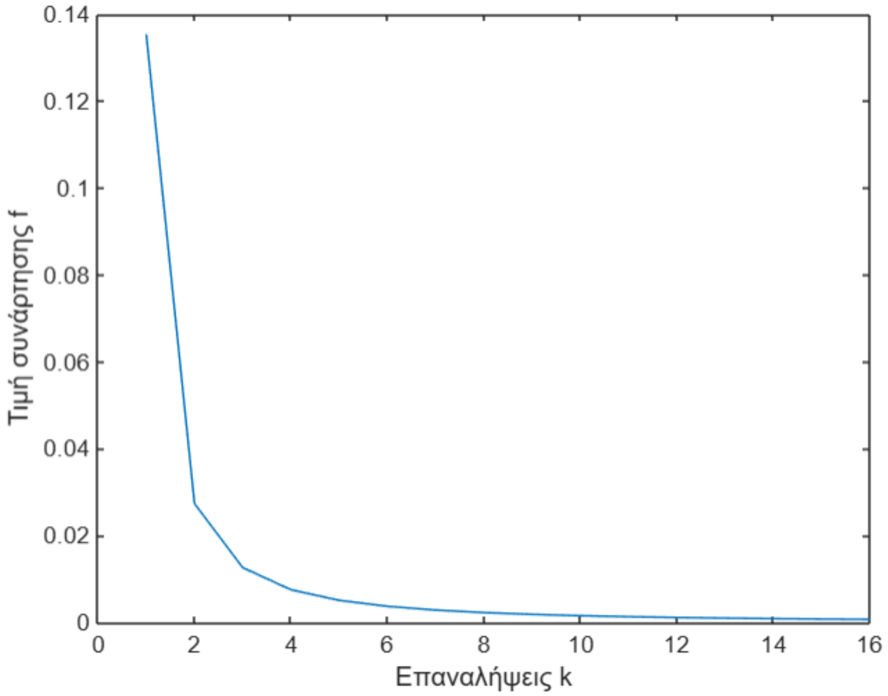
β) $\gamma_k = 0.5$



ii) Επιλογή γ_k τέτοιου ώστε να ελαχιστοποιείται η $f(x_k + \gamma_k * d_k)$



iii) Επιλογή γ_k βάσει του κανόνα Armijo



Παρατηρούμε ότι η χρήση και των τριών τρόπων επιλογής του βήματος γ_k οδηγεί σε σύγκλιση της μεθόδου. Όπως και με το προηγούμενο σημείο έναρξης, η επιλογή του γ_k ώστε να ελαχιστοποιείται η $f(x_k + \gamma_k * d_k)$ οδηγεί σε γρηγορότερη σύγκλιση (λιγότερες επαναλήψεις), ενώ ακολουθεί η επιλογή του γ_k βάσει του κανόνα Armijo. Η επιλογή σταθερού βήματος οδηγεί σε διαφορετική ταχύτητα σύγκλισης με βάση τον ορισμό της σταθερής τιμής, όπως φαίνεται και από τα δύο παραδείγματα για $\gamma_k = 0.01$ και $\gamma_k = 0.5$. Ωστόσο, και στις τρεις περιπτώσεις η τιμή της αντικειμενικής συνάρτησης οδηγείται περίπου προς το σημείο (0.34, -1.34),γεγονός που υποδεικνύει ότι σε εκείνη την περιοχή βρίσκεται κάποιο κρίσιμο σημείο της. Η μέθοδος ‘’εγκλωβίζεται’’ κοντά σε αυτό το κρίσιμο σημείο και δεν μπορεί να απομακρυνθεί, καθώς η τιμή $|\nabla f(x_k)|$ γίνεται μικρή, με αποτέλεσμα να ικανοποιείται η ανισότητα $|\nabla f(x_k)| < \varepsilon$ και ο αλγόριθμος , τελικά, να τερματίζεται. Ως συμπέρασμα, η επιλογή του σημείου (1, -1) ως σημείο έναρξης αποτελεί κακή επιλογή για την επίλυση του προβλήματος ελαχιστοποίησης.

Θέμα 3 (Μέθοδος Newton)

Εξήγηση τρόπου λειτουργίας: Η Μέθοδος Newton ακολουθεί την ίδια λογική με την Μέθοδο Μέγιστης Καθόδου. Όπως και προηγουμένως, επιλέγεται σε κάθε επανάληψη k ένα νέο σημεία x_{k+1} με βάση το αντίστοιχο σημείο της επανάληψης k, σύμφωνα με την σχέση $x_{k+1} = x_k + \gamma_k * d_k$. Η διαφορά της Μεθόδου Newton, ωστόσο, είναι ότι το διάνυσμα κατεύθυνσης ορίζεται ως $d_k = -[\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$. Με ανάλογο τρόπο με την 1η μέθοδο, ορίζεται μία σταθερά τερματισμού ε και στην αρχή κάθε επανάληψης k ελέγχεται εάν ικανοποιείται η ανισότητα $|\nabla f(x_k)| < \varepsilon$, γεγονός που οδηγεί στον τερματισμό της μεθόδου. Σημαντική προϋπόθεση για την ορθή λειτουργίας της Μεθόδου Newton αποτελεί ο εσσιανός πίνακας $\nabla^2 f(x_k)$ να είναι θετικά ορισμένος σε κάθε επανάληψη k. Σε αντίθετη περίπτωση, η μέθοδος δίνει λύσεις μακριά από το σημείο του ολικού ελαχίστου.

Ο υπολογισμός του εσσιανού πίνακα οδήγησε στο συμπέρασμα ότι δεν είναι θετικά ορισμένος για κανένα από τα τρία σημεία έναρξης (0, 0), (-1, 1) και (1, -1). Επομένως, η εφαρμογή της μεθόδου αναμένεται να μην οδηγήσει στο επιθυμητό αποτέλεσμα.

Για την Μέθοδο Newton επιλέχθηκαν οι ίδιες τιμές με την προηγούμενη μέθοδο για τις διάφορες σταθερές (σταθερά τερματισμού ε, σταθερές παραμέτρους της Μεθόδου Διχοτόμησης και σταθερές παραμέτρους του κανόνα Armijo).

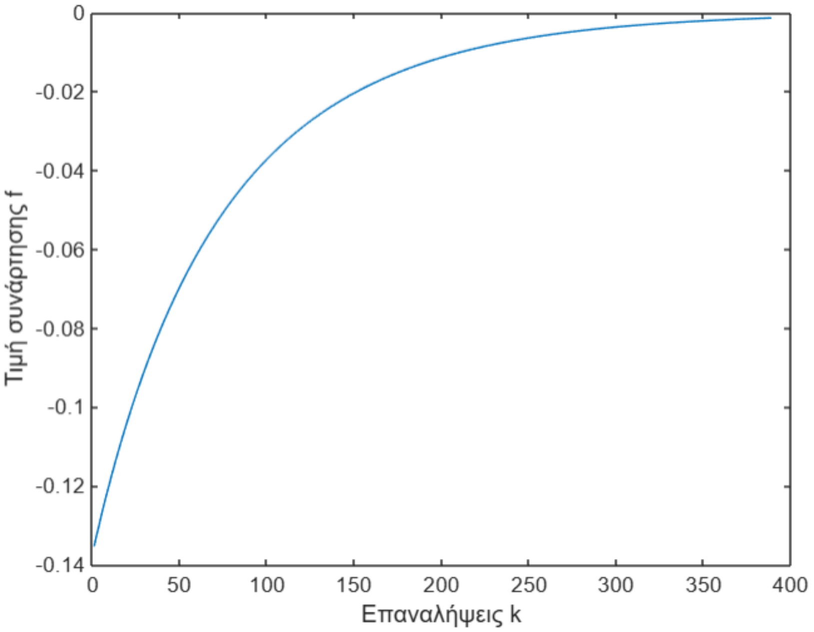
Ακολουθούν οι γραφικές παραστάσεις της σύγκλισης της αντικειμενικής συνάρτησης προς τον αριθμό των επαναλήψεων για τα διαφορετικά αρχικά σημεία και τους διαφορετικούς τρόπους επιλογής του γ_k.

1) Γραφικές παραστάσεις για αρχικό σημείο (0, 0):

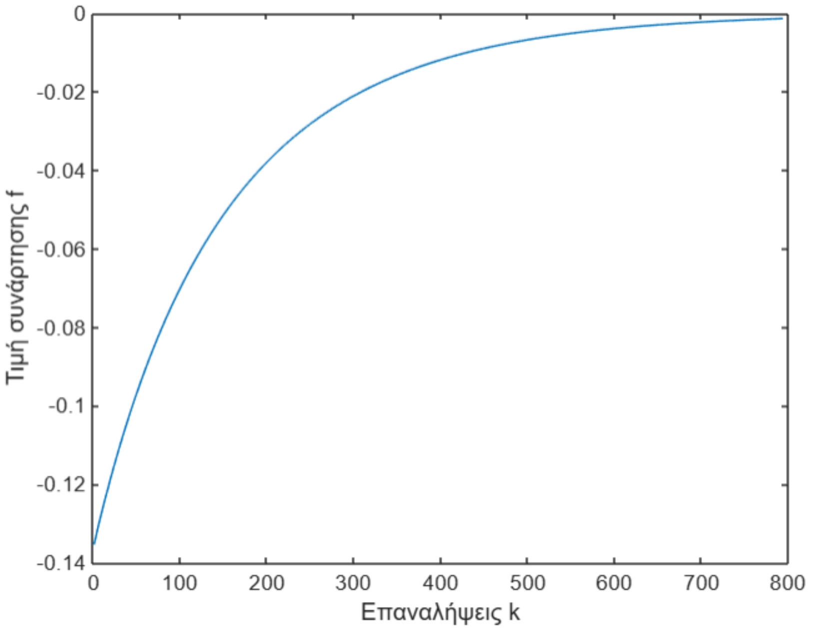
Ανάλογα με την Μέθοδο Μέγιστης Καθόδου, η επιλογή του σημείου (0, 0) ως σημείο έναρξης οδηγεί σε τερματισμό του αλγορίθμου στη 1η κιόλας επανάληψη.

2) Γραφικές παραστάσεις για αρχικό σημείο (-1, 1):

i) Επιλογή σταθερού $\gamma_k = 0.01$



ii) Επιλογή γ_k τέτοιου ώστε να ελαχιστοποιείται η $f(x_k + \gamma_k * d_k)$



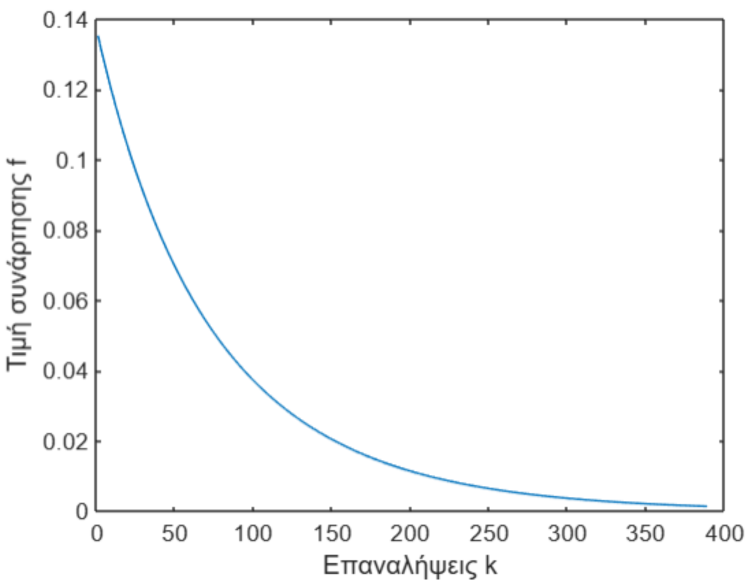
iii) Επιλογή γ_k βάσει του κανόνα Armijo

Παρά την δοκιμή πολλών διαφορετικών τιμών για τις παραμέτρους του κανόνα, η μέθοδος δείχνει να μην συγκλίνει καθώς η εκτέλεση του προγράμματος στο Matlab δεν τερματίζει μέσα σε λογικά χρονικά πλαίσια.

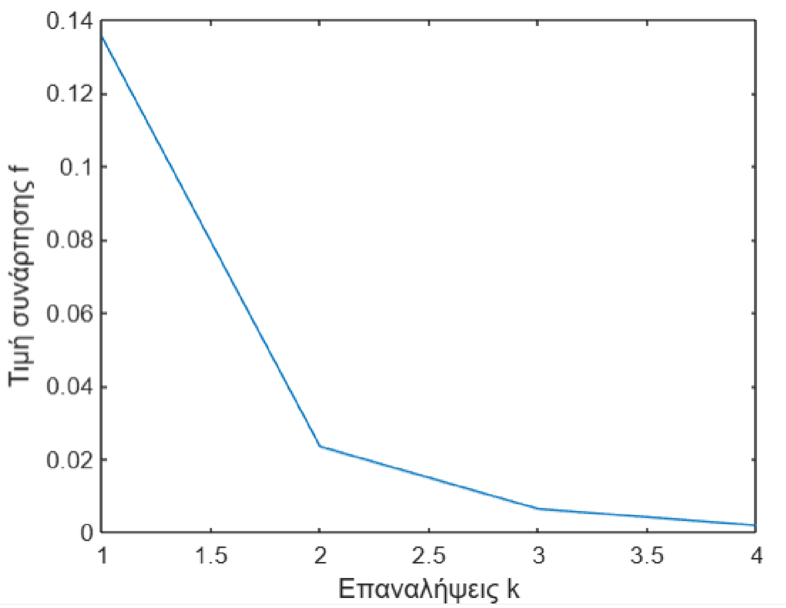
Παρατηρούμε ότι με τους δύο τρόπους επιλογής του γ_k που οδηγούν σε σύγκλιση της μεθόδου, η τιμή της αντικειμενικής συνάρτησης τείνει προς το μηδέν ενώ το x_k τείνει προς το σημείο (x, y)=(-0.67, 2.03). Συμπερασματικά, η μέθοδος, λόγω του μη θετικά ορισμένου εσσιανού πίνακα, δίνει λύση μακριά από το σημείο ολικού ελαχίστου.

3) Γραφικές παραστάσεις για αρχικό σημείο (1, -1):

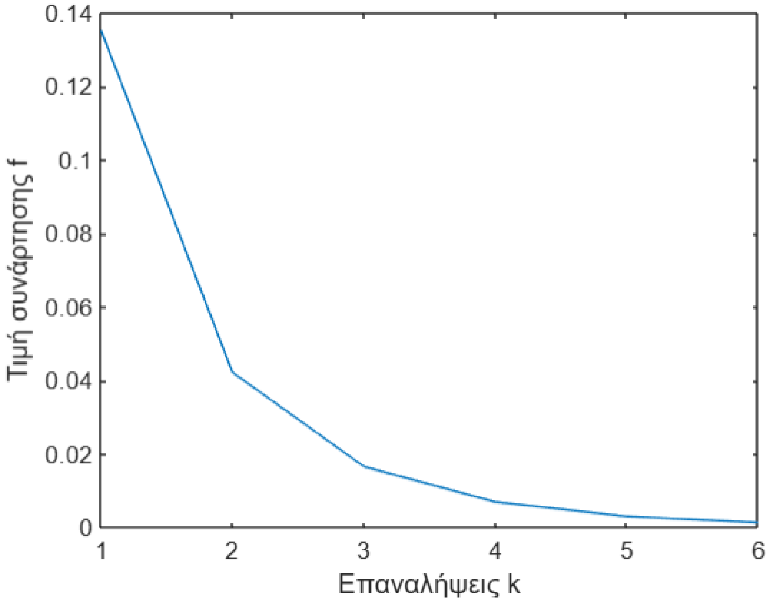
i) Επιλογή σταθερού $\gamma_k = 0.01$



ii) Επιλογή γ_k τέτοιου ώστε να ελαχιστοποιείται η $f(x_k + \gamma_k * d_k)$



iii) Επιλογή γ_k βάσει του κανόνα Armijo



Παρατηρούμε ότι η μέθοδος συγκλίνει γρηγορότερα στην 2η περίπτωση ενώ ακολουθεί η 3η . Η μέθοδος δίνει και για αυτό το σημείο έναρξης λύση μακριά από το σημείο του ελαχίστου και συγκεκριμένα καταλήγει περίπου στο σημείο (x, y)=(0.62, -2.15), ενώ η τιμή της αντικειμενικής συνάρτησης τείνει να μηδενιστεί.

Αξιοσημείωτο είναι το γεγονός ότι το σημείο αυτό βρίσκεται πολύ κοντά με το αντιδιαμετρικό σημείο του (-0.67, 2.03), στο οποίο συνέκλινε η μέθοδος για σημείο έναρξης (-1, 1), ως προς την αρχή των αξόνων πάνω στο επίπεδο xy.

Θέμα 4 (Μέθοδος Levenberg-Marquardt)

Εξήγηση τρόπου λειτουργίας: Η μέθοδος αυτή αποτελεί παραλλαγή της Μεθόδου Newton, ακολουθώντας την βασική ιδέα της με κάποιες διαφοροποιήσεις. Συγκεκριμένα, το διάνυσμα κατεύθυνσης υπολογίζεται μέσω της επίλυσης του συστήματος $[\nabla^2 f(x_k) + \mu_k * I] * d_k = -\nabla f(x_k)$, όπου η τιμή του μ_k επιλέγεται ώστε ο πίνακας $[\nabla^2 f(x_k) + \mu_k * I]$ να είναι θετικά ορισμένος. Όπως ακριβώς και στις προηγούμενες μεθόδους, σε κάθε επανάληψη k ορίζεται το σημείο που θα εξεταστεί στην επόμενη επανάληψη, μέσω της σχέσης $x_{k+1} = x_k + \gamma_k * d_k$. Επιπλέον, ορίζεται σταθερά τερματισμού ε και στην αρχή κάθε επανάληψης γίνεται έλεγχος της ανισότητας $|\nabla f(x_k)| < \varepsilon$, σε περίπτωση ικανοποίησης της οποίας ο αλγόριθμος τερματίζεται. Συμπερασματικά, η Μέθοδος Levenberg-Marquardt ακολουθεί την λογική της Μεθόδου Newton, αντικαθιστώντας τον εσσιανό πίνακα $\nabla^2 f(x_k)$ με τον πίνακα $[\nabla^2 f(x_k) + \mu_k * I]$ ο οποίος έχει σχεδιαστεί με τέτοιο τρόπο ώστε να είναι θετικά ορισμένος.

Στην ανάλυση αυτή, το μ_k επιλέχθηκε κατάλληλα με βάση το χωρίο του βιβλίου στην σελίδα 139, στο οποίο σημειώνεται ότι το μ_k εξασφαλίζει θετικά ορισμένο πίνακα $[\nabla^2 f(x_k) + \mu_k * I]$ αν είναι μεγαλύτερο από την απόλυτη τιμή της μεγαλύτερης ιδιοτιμής του εσσιανού $\nabla^2 f(x_k)$. Επομένως, ορίστηκε $\mu_k = \max(|ιδιοτιμές\ εσσιανού|) + 0.06$.

Για την Μέθοδο Levenberg-Marquardt επιλέχθηκε η ίδια τιμή με τις δύο παραπάνω μεθόδους όσον αφορά την τιμή της σταθεράς τερματισμού ε. Για το αρχικό διάστημα αναζήτησης $[\alpha, \beta]$ στη Μεθόδου διχοτόμησης και την τιμή του αρχικού βήματος που δηλώνεται με τον συντελεστή s στον κανόνα Armijo, πραγματοποιήθηκαν δοκιμές με διάφορες τιμές για το αρχικό διάστημα αναζήτησης καθώς παρατηρήθηκαν μεγάλες διαφοροποιήσεις των αποτελεσμάτων των μεθόδων με βάση την επιλογή των τιμών αυτών των παραμέτρων. Οι υπόλοιπες παράμετροι της Μεθόδου Διχοτόμησης (ε, l) και του κανόνα Armijo (α, β) έλαβαν ίδιες τιμές με τις αντίστοιχες στις δύο προηγούμενες μεθόδους.

Ακολουθούν οι γραφικές παραστάσεις της σύγκλισης της αντικειμενικής συνάρτησης προς τον αριθμό των επαναλήψεων για τα διαφορετικά αρχικά σημεία και τους διαφορετικούς τρόπους επιλογής του γ_k.

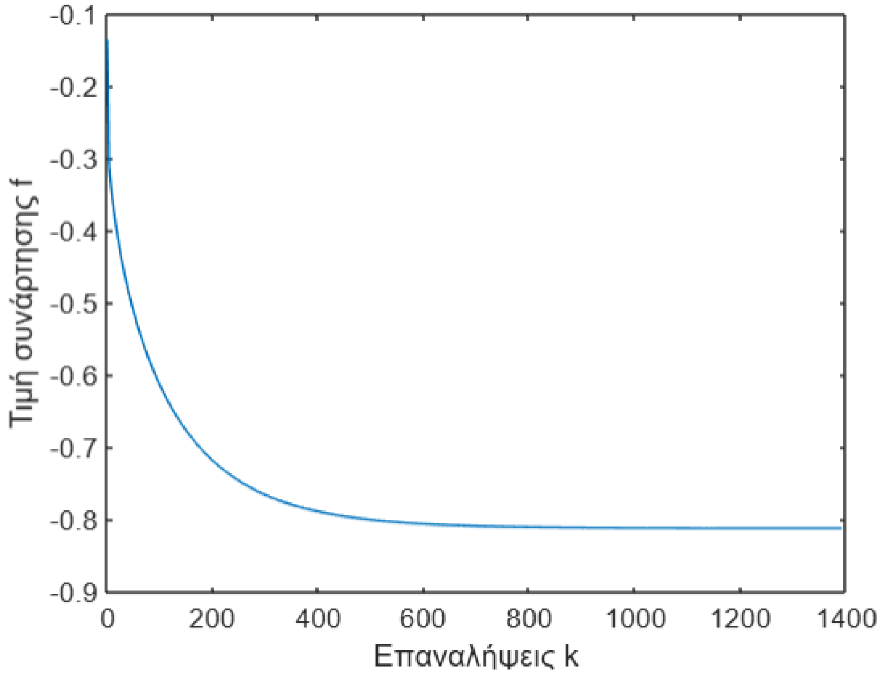
1) Γραφικές παραστάσεις για αρχικό σημείο (0, 0):

Ανάλογα με τις δύο προηγούμενες μεθόδους, η επιλογή του σημείου (0, 0) ως σημείο έναρξης οδηγεί σε τερματισμό του αλγορίθμου στη 1η κιόλας επανάληψη.

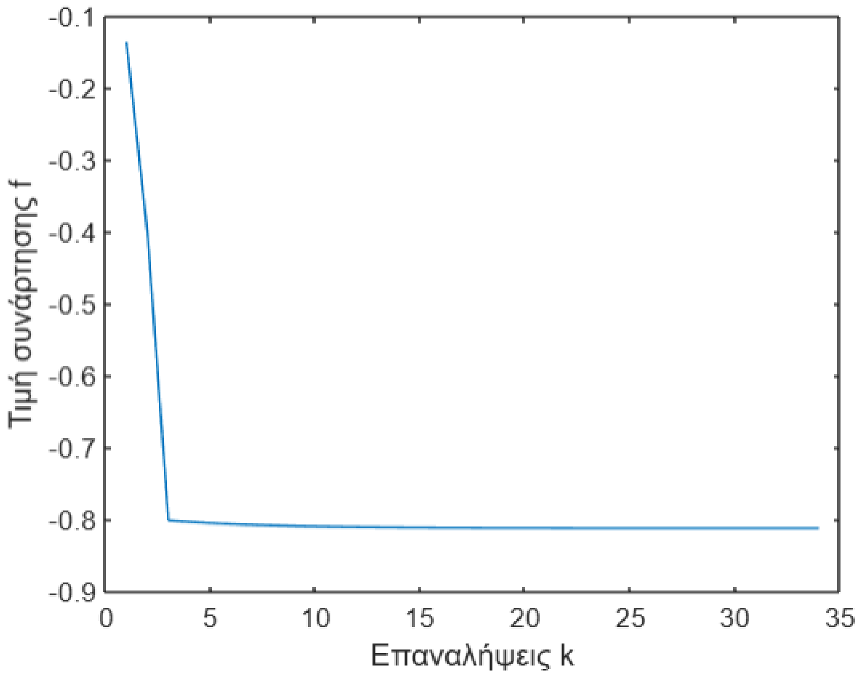
2) Γραφικές παραστάσεις για αρχικό σημείο (-1, 1):

i) Επιλογή σταθερού γ_k

α) $\gamma_k = 0.01$



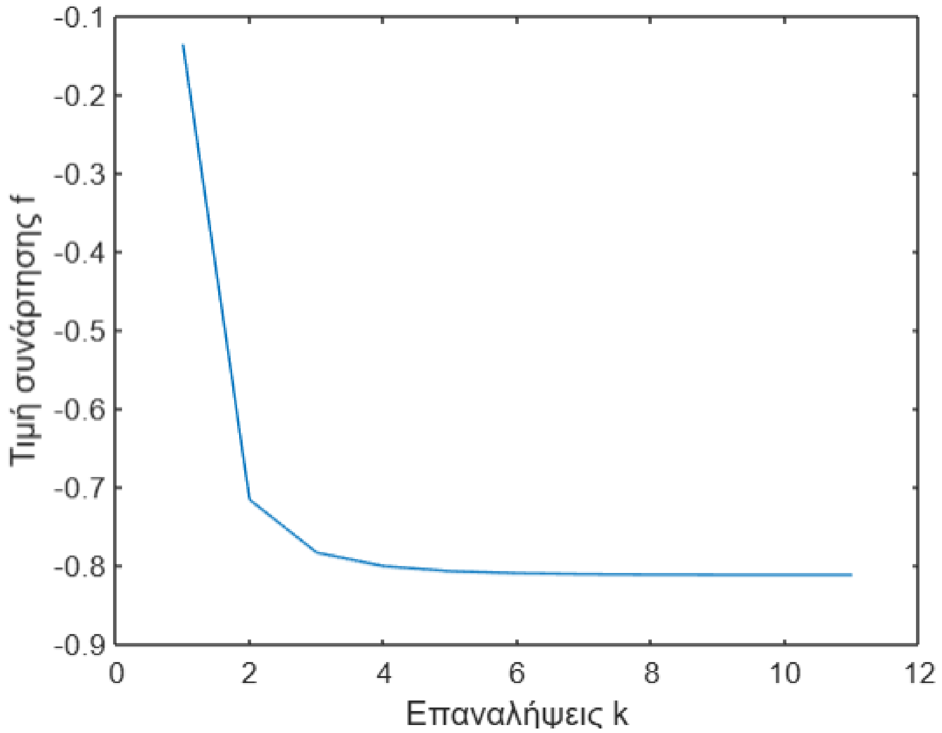
β) $\gamma_k = 0.21$



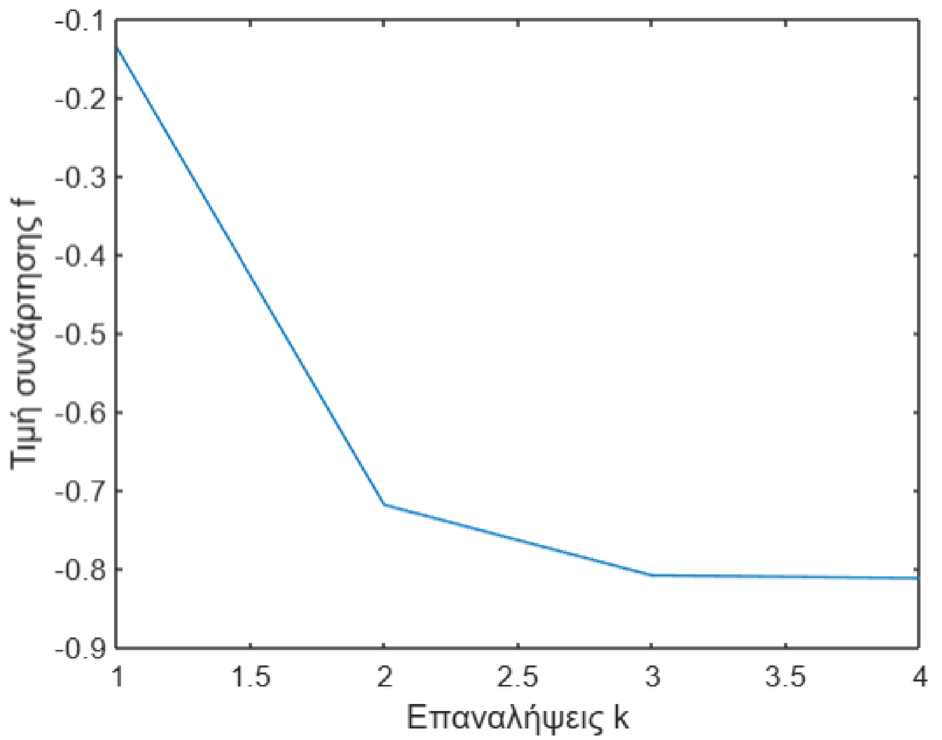
Μέσω δοκιμών, παρατηρήθηκε ότι η σταθερή τιμή του βήματος γ_k(με ακρίβεια δύο δεκαδικών ψηφίων) που οδηγεί στον μικρότερο αριθμό επαναλήψεων χωρίς την εμφάνιση ταλαντώσεων είναι η τιμή $\gamma_k = 0.21$. Επομένως, μπορούμε να συμπεράνουμε ότι η βέλτιστη τιμή του βήματος βρίσκεται κοντά στην τιμή 0.21.

ii) Επιλογή γ_k τέτοιου ώστε να ελαχιστοποιείται η $f(x_k + \gamma_k * d_k)$

α) Παράμετροι Μεθόδου Διχοτόμησης: $\alpha = 0, \beta = 1$



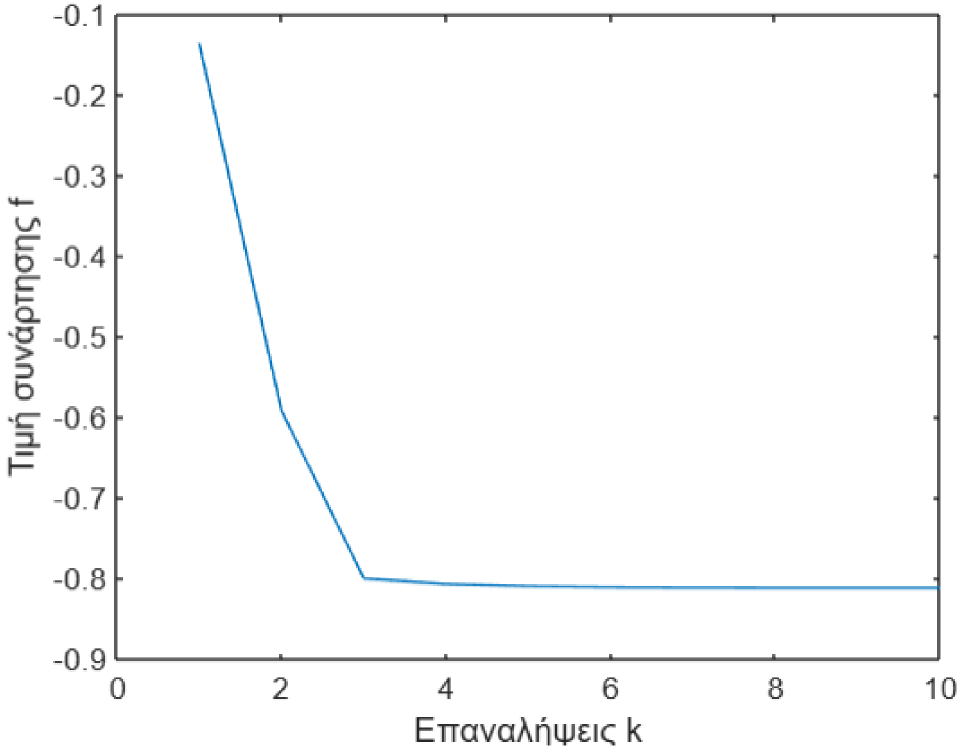
β) Παράμετροι Μεθόδου Διχοτόμησης: $\alpha = 0, \beta = 5$



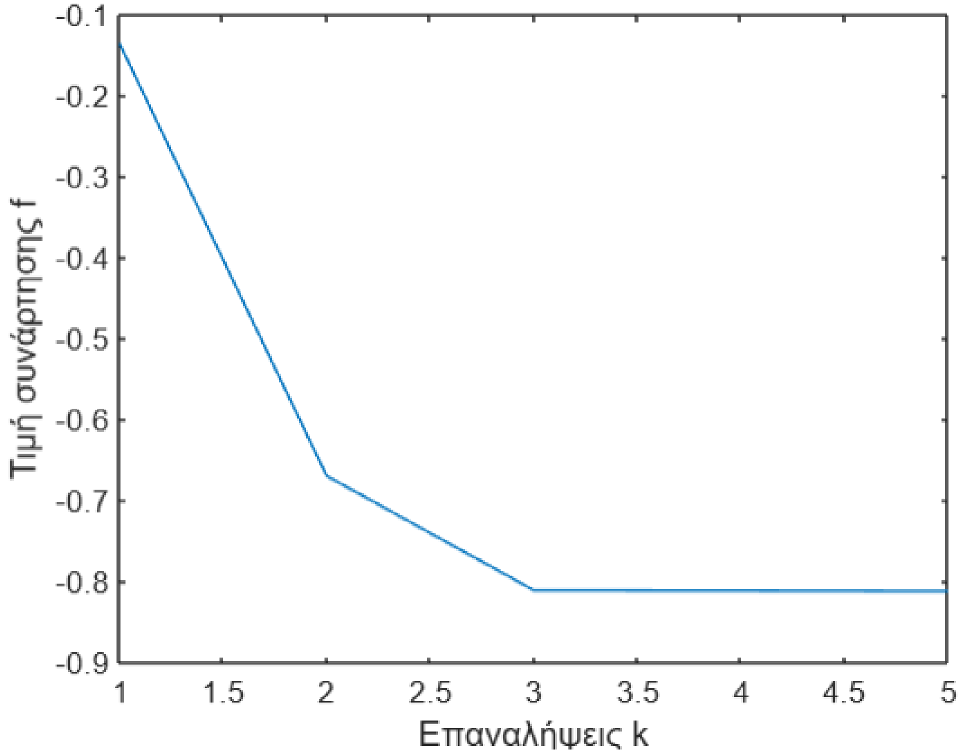
Παρατηρούμε ότι η αύξηση του εύρους του αρχικού διαστήματος αναζήτησης οδηγεί σε μείωση του συνολικού αριθμού επαναλήψεων της μεθόδου. Δεν παρατηρείται καλύτερη απόδοση με περαιτέρω αύξηση της τιμής του β. (Το α παραμένει σταθερό.)

iii) Επιλογή γ_k βάσει του κανόνα Armijo

α) $s = 0.5$



β) $s = 5$



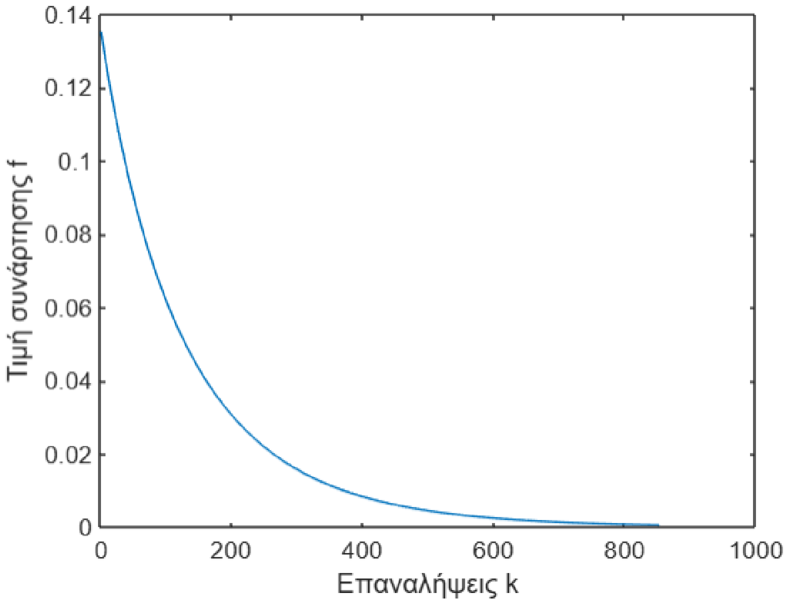
Παρατηρούμε ότι η αύξηση της τιμής του συντελεστή s οδηγεί σε μείωση του συνολικού αριθμού επαναλήψεων. Αυτό συμβαίνει μέχρι την τιμή $s = 5$.

Παρατηρούμε ότι η χρήση και των τριών τρόπων επιλογής του βήματος γ_k οδηγεί σε σύγκλιση της μεθόδου, καθώς η τιμή της συνάρτησης τείνει προς το ελάχιστο της -0.81. Ταχύτερη σύγκλιση σημειώνεται με επιλογή γ_k τέτοιου ώστε να ελαχιστοποιείται η $f(x_k + \gamma_k * d_k)$ ενώ ακολουθεί η περίπτωση επιλογής του γ_k βάσει του κανόνα Armijo. Κατά την χρήση της Μεθόδου Levenberg-Marquardt, η επιλογή της τιμής γ_k=0.21 για το σταθερό βήμα οδήγησε σε καλή επίδοση, με τον αλγόριθμο να τερματίζεται σε μόλις 34 επαναλήψεις.

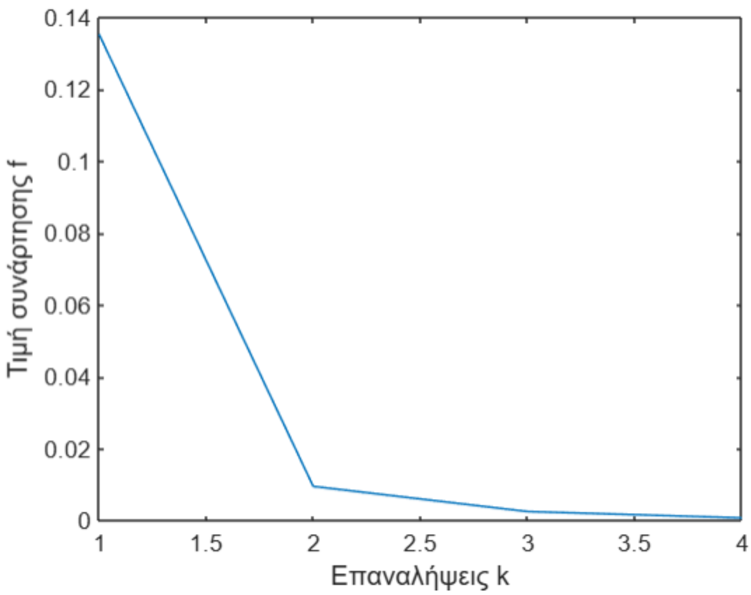
3) Γραφικές παραστάσεις για αρχικό σημείο (1, -1):

i) Επιλογή σταθερού γ_k

α)$\gamma_k = 0.01$



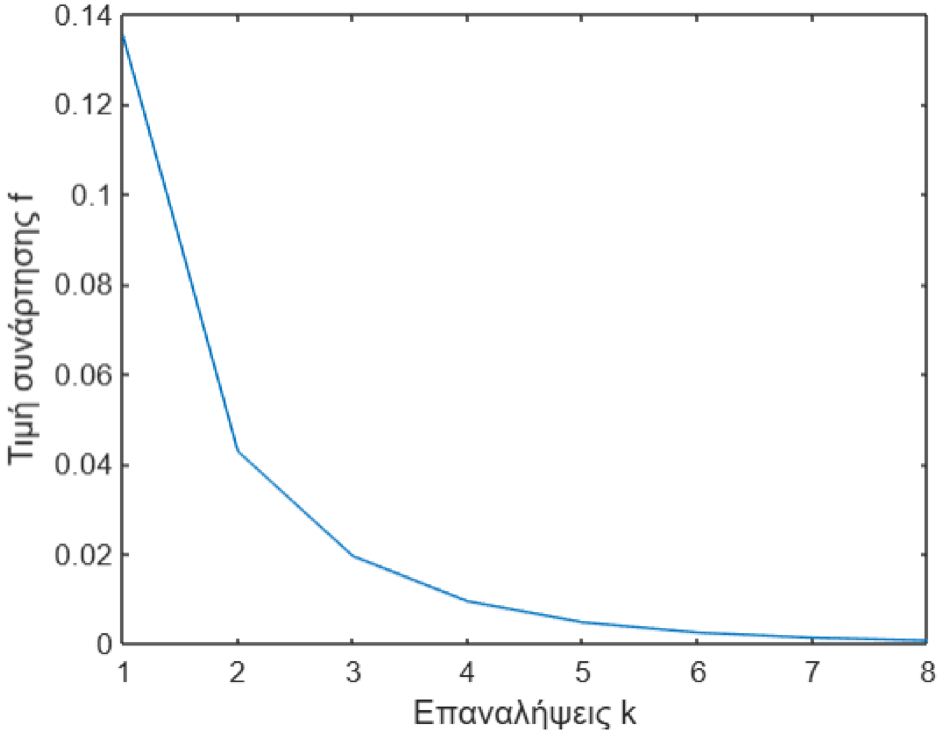
β)$\gamma_k = 1.8$



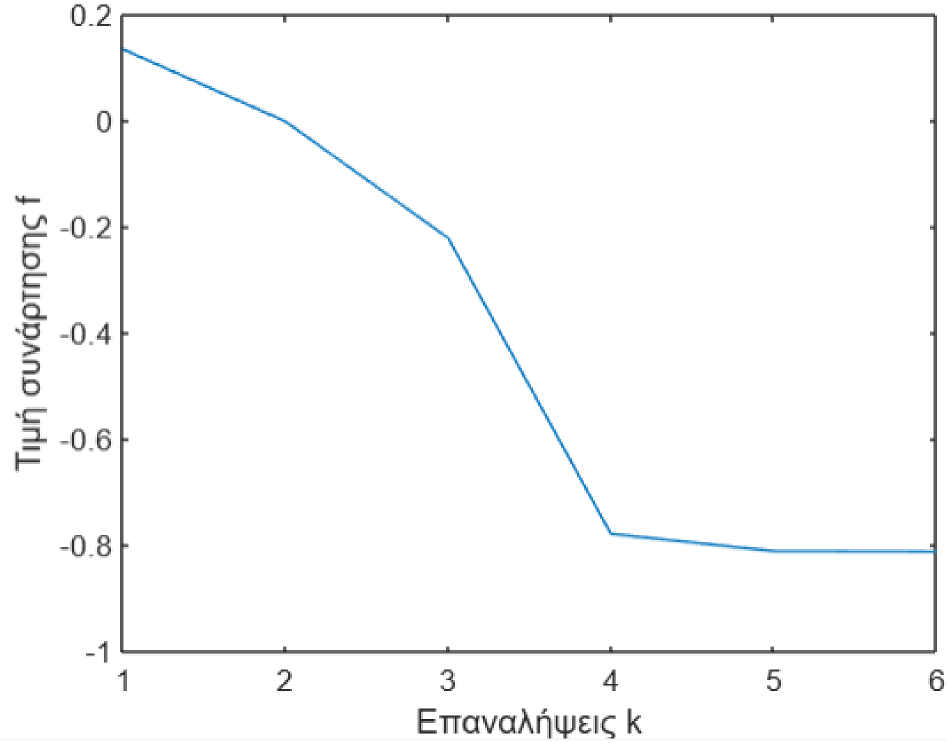
Η αύξηση της τιμής του σταθερού βήματος οδηγεί σε μείωση του συνολικού αριθμού των επαναλήψεων, έως την τιμή $\gamma_k = 1.8$. Η μέθοδος, ωστόσο, δεν συγκλίνει προς το ολικό ελάχιστο της αντικειμενικής συνάρτησης, όπως ακριβώς και η Μέθοδος Μέγιστης Καθόδου για το σημείο έναρξης (1, -1).

ii) Επιλογή γ_k τέτοιου ώστε να ελαχιστοποιείται η $f(x_k + \gamma_k * d_k)$

Με ορισμό του αρχικού διαστήματος αναζήτησης του γ_k στην μέθοδο διχοτόμησης ως το διάστημα [0,1]:



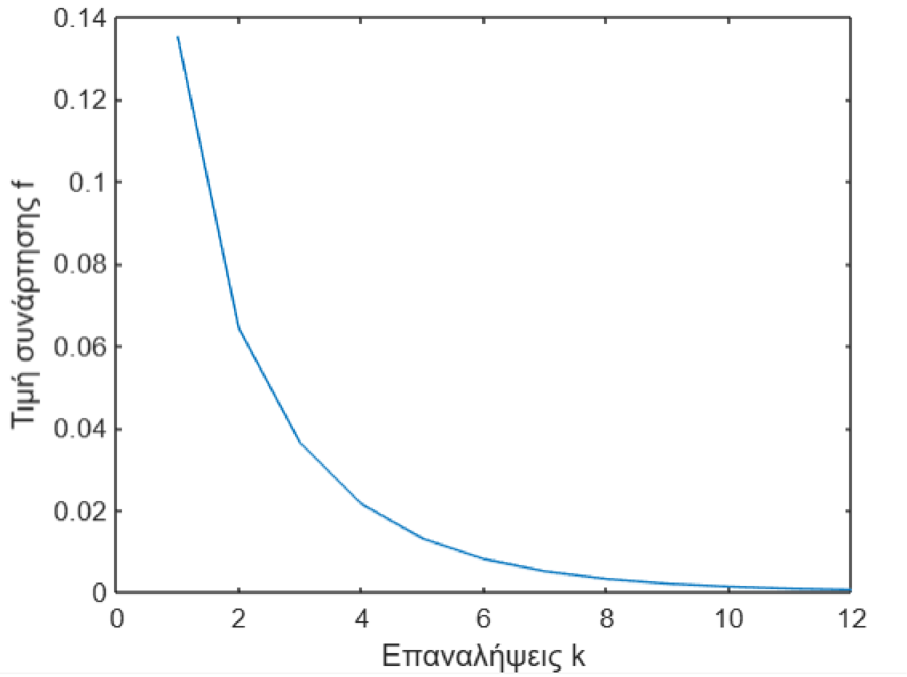
Με ορισμό του αρχικού διαστήματος αναζήτησης του γ_k στην μέθοδο διχοτόμησης ως το διάστημα [0,5]:



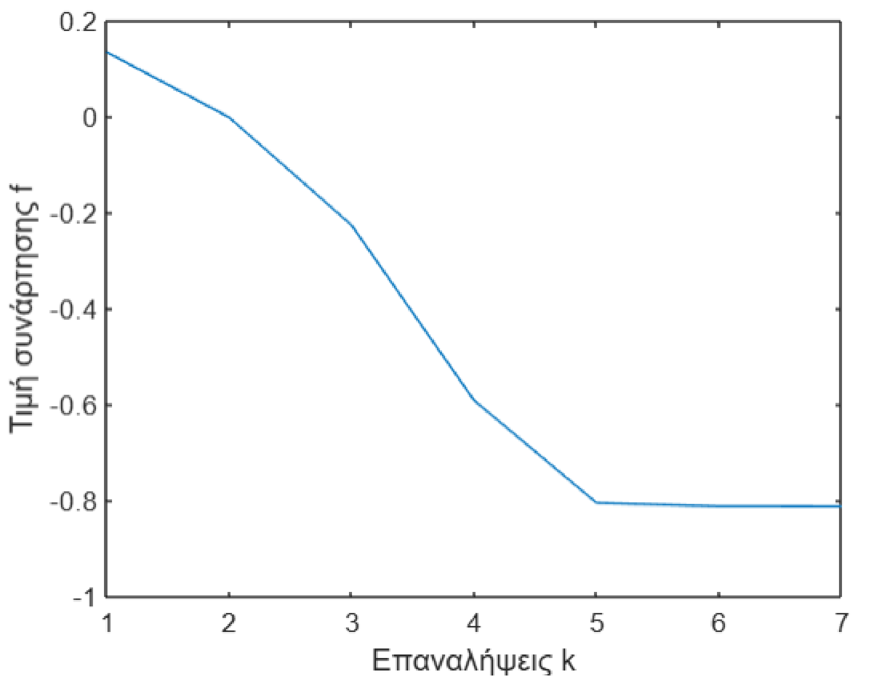
Παρατηρούμε ότι η επιλογή του αρχικού διαστήματος αναζήτησης στην δεύτερη περίπτωση οδήγησε στην προσέγγιση του ολικού ελαχίστου της αντικειμενικής συνάρτησης, σε αντίθεση με την πρώτη περίπτωση. Αυτό οφείλεται στο γεγονός ότι για το αρχικό διάστημα αναζήτησης [0, 1] η τιμή του γ_k περιορίζεται σε μικρές τιμές, με αποτέλεσμα η μέθοδος να εγκλωβιστεί σε κρίσιμο σημείο που βρίσκεται πιο κοντά στο σημείο έναρξης σε σχέση με το σημείο του ολικού ελαχίστου. Αντιθέτως, ο ορισμός του αρχικού διαστήματος ως [0, 5] επιτρέπει στο βήμα γ_k να πάρει αρκετά μεγάλη τιμή ώστε η μέθοδος να κατευθυνθεί μακριά από το ανεπιθύμητο κρίσιμο σημείο και προς το ολικό ελάχιστο.

iii) Επιλογή γ_k βάσει του κανόνα Armijo

Με ορισμό του συντελεστή s=1:



Με ορισμός του συντελεστή s=5:



Με την χρήση του κανόνα Armijo για την επιλογή του βήματος γ_k, παρατηρείται το ίδιο φαινόμενο με παραπάνω. Η τιμή του συντελεστή s=5 επιτρέπει στην μέθοδο να αποφύγει το ανεπιθύμητο κρίσιμο σημείο και, τελικά, να δώσει λύση κοντά στο σημείο του ολικού ελαχίστου της αντικειμενικής συνάρτησης.

Σύγκριση των μεθόδων με βάση την αποδοτικότητα τους: Όσον αφορά την σύγκριση των μεθόδων με βάση την αποδοτικότητα τους στην επίλυση του συγκεκριμένου προβλήματος ελαχιστοποίησης, η

Μέθοδος Newton αναδεικνύεται ως η χειρότερη μέθοδος καθώς η μορφή της αντικειμενικής συνάρτησης την οδηγεί σε λάθος αποτελέσματα. Μεταξύ των μεθόδων Μέγιστης Καθόδου και Levenberg-Marquardt, θα συγκρίνουμε τις επιδόσεις τους για καθένα από τα τρία διαφορετικά σημεία έναρξης.

Σχετικά με το σημείο έναρξης (0, 0), καμία από τις δύο μεθόδους δεν συγκλίνει στη λύση του προβλήματος.

Για το σημείο έναρξης (-1, 1), η Μέθοδος Μέγιστης Καθόδου οδηγεί σε μικρότερο αριθμό επαναλήψεων τόσο για την επιλογή σταθερού $\gamma_k = 0.01$ όσο και για την επιλογή της βέλτιστης τιμής του γ_k καθεμίας από τις δύο μεθόδους. Με την επιλογή γ_k τέτοιου ώστε να ελαχιστοποιείται η $f(x_k + \gamma_k * d_k)$, στην περίπτωση ίδιου αρχικού διαστήματος αναζήτησης [0,1] της Μεθόδου Διχοτόμησης η Μέθοδος Μέγιστης Καθόδου υπερισχύει ξανά. Η επιλογή άλλων αρχικών διαστημάτων αναζήτησης δεν οδηγεί την Μέθοδο Μέγιστης Καθόδου σε καλύτερη απόδοση, σε αντίθεση με την Μέθοδο Levenberg-Marquardt η οποία με το ορισμό του αρχικού διαστήματος [0, 5] συγκλίνει σε πολύ λίγες επαναλήψεις. Το ίδιο φαινόμενο παρατηρείται και με την επιλογή του γ_k βάσει του κανόνα Armijo. Για $s = 0.5$ και για τις δύο μεθόδους, η μέθοδος Μέγιστης Καθόδου αποδεικνύεται γρηγορότερη. Αντίθετα, για προσέγγιση βέλτιστης επιλογή s για καθεμία από τις δύο μεθόδους ($s = 0.5$ όπως πριν για Μέγιστης Καθόδου και $s = 5$ για Levenberg-Marquardt), η δεύτερη μέθοδος συγκλίνει σε λιγότερες επαναλήψεις.

Τέλος, η επιλογή του σημείου (1, -1) για σημείο έναρξης έχει ως συνέπεια η Μέθοδος Μέγιστης Καθόδου να μην επιτυγχάνει σύγκλιση προς το ελάχιστο για καμία επιλογή των διαφόρων παραμέτρων. Ωστόσο, η Μέθοδος Levenberg-Marquardt, με κατάλληλες τιμές των παραμέτρων, συγκλίνει στις περιπτώσεις επιλογής του βήματος με την Μέθοδο Διχοτόμησης ώστε να ελαχιστοποιείται η $f(x_k + \gamma_k * d_k)$ καθώς και βάσει του κανόνα Armijo. Καθίσταται, επομένως, καλύτερη όσον αφορά την επίλυση του προβλήματος με αυτό το σημείο έναρξης.

Συνοψίζοντας, η Μέθοδος Newton κρίνεται ακατάλληλη για το συγκεκριμένο πρόβλημα ενώ καθεμία από τις υπόλοιπες δύο υπερτερεί σε διαφορετικές περιπτώσεις. Η Μέθοδος Μέγιστης Καθόδου διαθέτει το πλεονέκτημα της απλούστερης υλοποίησης καθώς δεν απαιτεί τον υπολογισμού του εσσιανού πίνακα της αντικειμενικής συνάρτησης, απαραίτητο βήμα της Μεθόδου Levenberg. Η δεύτερη, με την σειρά της, μπορεί να επιφέρει, όπως αποδείχθηκε, την επιθυμητή λύση σε περισσότερες περιπτώσεις.