Τεχνικές Βελτιστοποίησης - Εργασία 2

Ονοματεπώνυμο: Τσαμήτρος Νικόλαος

ΑΕΜ: 10781

Η εργασία αυτή αφορά το πρόβλημα ελαχιστοποίησης μια δοσμένης συνάρτησης πολλών μεταβλητών $f:\mathbb{R}^n \to \mathbb{R}$ χωρίς περιορισμούς. Οι αλγόριθμοι που χρησιμοποιούνται βασίζονται στην ιδέα της επαναληπτικής καθόδου, βάσει τις οποίας ξεκινάμε από κάποιο σημείο $x_0 \in \mathbb{R}^n$ και παράγουμε διαδοχικά τα διανύσματα $x_1, x_2, \ldots$ έτσι ώστε $f(x_{k+1}) < f(x_k)$, $k = 1,2,..$

Οι αλγόριθμοι μου μελετούνται είναι οι:

- Μέθοδος Μέγιστης Καθόδου
- Μέθοδος Newton
- Μέθοδος Levenberg-Marquardt

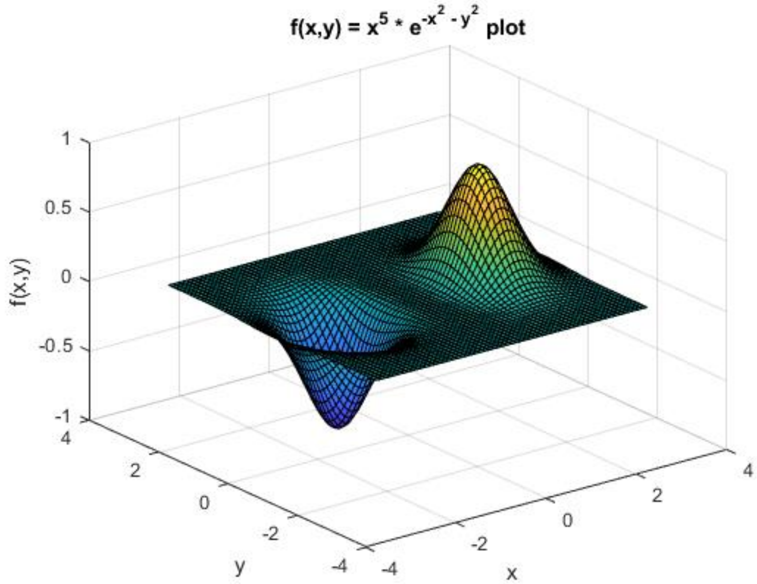
Η αντικειμενική συνάρτηση που μελετάται είναι η: $f(x) = x^5 e^{(-x^2-y^2)}$

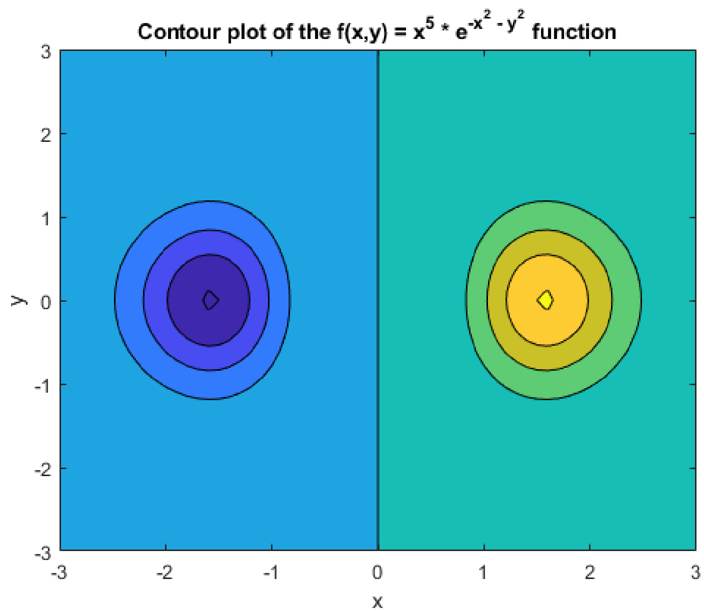
Σημείωση: Για το σταθερό βήμα έχω επιλέξει $\gamma_\kappa = 0.1$, η ελαχιστοποίση της $f(x_k + \gamma_\kappa d_k)$ γίνεται με την μέθοδο του χρυσού τομέα από την 1η εργασία, και για τον κανόνα του Armijo επέλεξα $a = 1e - 4$, $b = 0.2$, $s = 100$. Για την συνθήκη τερματισμού έχουμε $\varepsilon = 0.01$.

Θέμα 1

Στο πρώτο θέμα ζητείται ο σχεδιασμός της αντικειμενικής συνάρτησης ώστε να πάρουμε μια γενική εικόνα της μορφής της. Παρακάτω παρουσιάζονται δύο διαγράμματα της f , ένα στον $\mathbb{R}^3$ και ένα contour plot αντίστοιχα. Παρατηρούμε πως η f έχει:

- Ολικό ελάχιστο με τιμή περίπου -0.81
- Ολικό μέγιστο με τιμή περίπου +0.81
- Σημείο σάγματος στο 0.





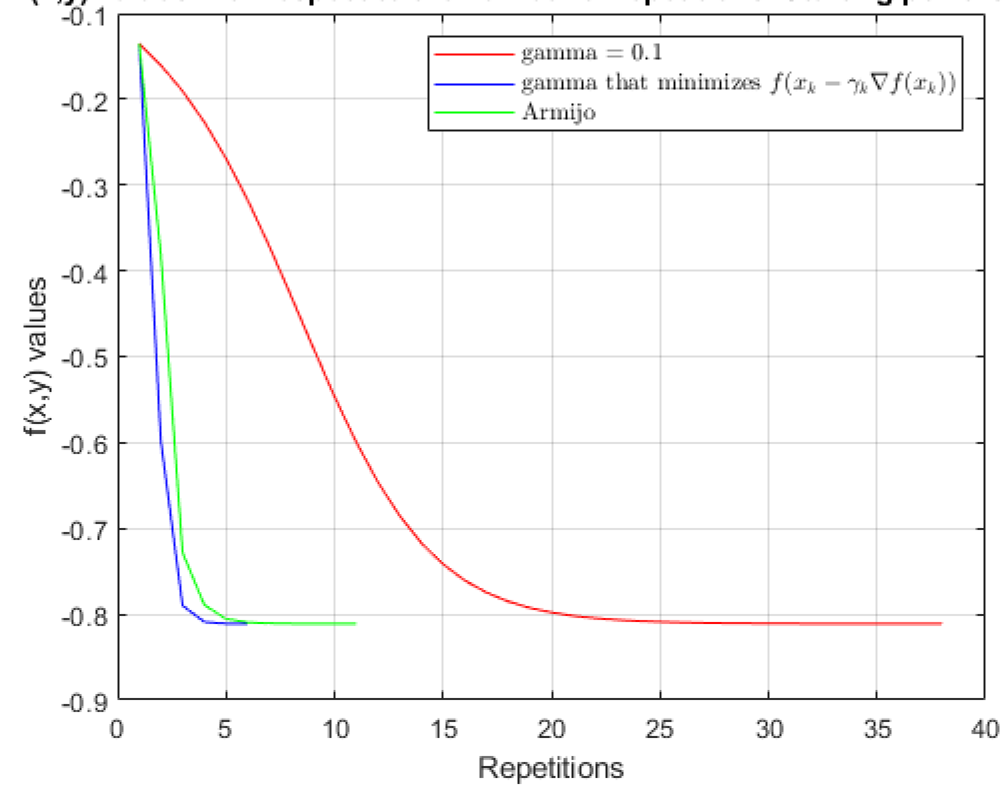
Θέμα 2

Στο δεύτερο θέμα ζητείται η υλοποίηση της μεθόδου της μέγιστης καθόδου χρησιμοποιώντας ως αρχικά σημεία (x_0, y_0) τα i) (0,0), ii) (−1,1), και iii) (1, −1). Επίσης το βήμα γ_κ επιλέγεται ως εξής: α) σταθερό (της επιλογής μας), β) τέτοιο ώστε να ελαχιστοποιεί την $f(x_k + \gamma_\kappa d_k)$ και γ) βάσει του κανόνα Armijo.

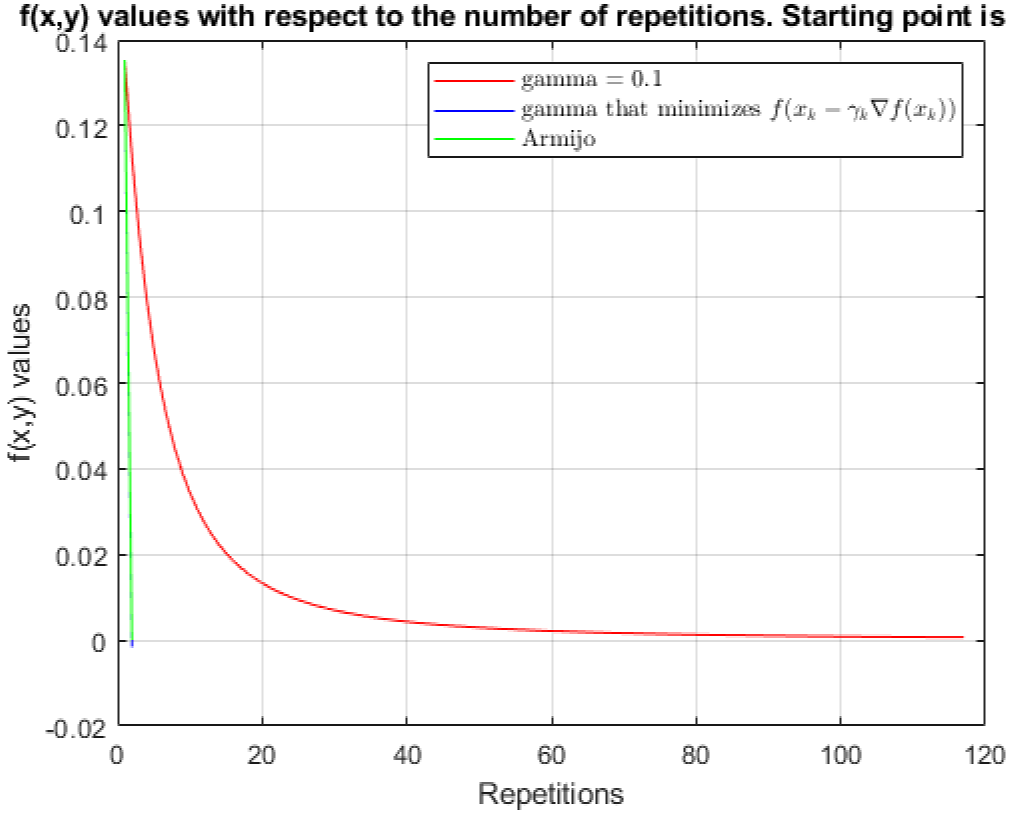
1. Όταν το αρχικό μας σημείο είναι το (0,0), ο αλγόριμος δεν θα ξεκινήσει καθόλου, αφού αυτό είναι σημείο σάγματος της συνάρτησης άρα το $\nabla f(0,0) = 0$, οπότε πληρείται η συνθήκη τερματισμού του αλγορίθμου. Άρα δεν υπάρχουν κάποια διαγράμματα σε αυτό το υποερώτημα.
2. Σημείο (-1,1). Ο αλγόριθμος βρίσκει επιτυχώς το ολικό ελάχιστο της συνάρτησης, βέβαια υπάρχει διαφοροποίηση στον αριθμό τον επαναλήψεων που εξαρτάται από την επιλογή του βήματος γ_κ. Για σταθερό $\gamma_\kappa = 0.1$, ο αλγόριθμος συγκλίνει σε 38 επαναλήψεις, με την ελαχιστοποίηση σε 6 επαναλήψεις, ενώ με τον κανόνα του Armijo σε 11 επαναλήψεις. Προφανώς ο τρόπος επιλογής του βήματος με την ελαχιστοποίηση συγκλίνει πιο γρήγορα από τους υπόλοιπους αφού επιλέγεται το βέλτιστο βήμα κάθε φορά(η ελαχιστοποίηση έγινε με την μέθοδο του Χρυσού Τομέα από την πρώτη εργασία). Ακολουθεί ο κανόνας του Armijo, ο οποίος επιλέγει ένα βήμα σιγουρεύοντας την μείωση της συνάρτησης. Τέλος έχουμε το σταθερό $\gamma_\kappa = 0.1$, που χρειάζεται τις περισσότερες επαναλήψεις, το οποίο είναι αναμενόμενο, εφόσον έχω επιλέξει μικρό βήμα, ώστε να διασφαλιστεί η σταθερή μείωση της συνάρτησης και να αποφευχθούν ταλαντώσεις εάν το βήμα πάρει κάποια μεγάλη τιμή.
3. Σημείο (1,-1). Ο αλγόριθμος και με τους 3 τρόπους καταλήγει στο 0 το οποίο είναι σημείο σάγματος και πληρεί την συνθήκη τερματισμού. Το σταθερό βήμα $\gamma_\kappa = 0.1$ χρειάζεται προφανώς τις περισσότερες επαναλήψεις (117), ενώ η ελαχιστοποίηση και ο κανόνας του Armijo είναι ταυτόσημα με 2 επαναλήψεις το καθένα.

Παρακάτω παρουσιάζονται τα διαγράμματα της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων μέχρι τον τερματισμό του αλγορίθμου.

f(x,y) values with respect to the number of repetitions. Starting point is (-1,1)



f(x,y) values with respect to the number of repetitions. Starting point is (1,-1)

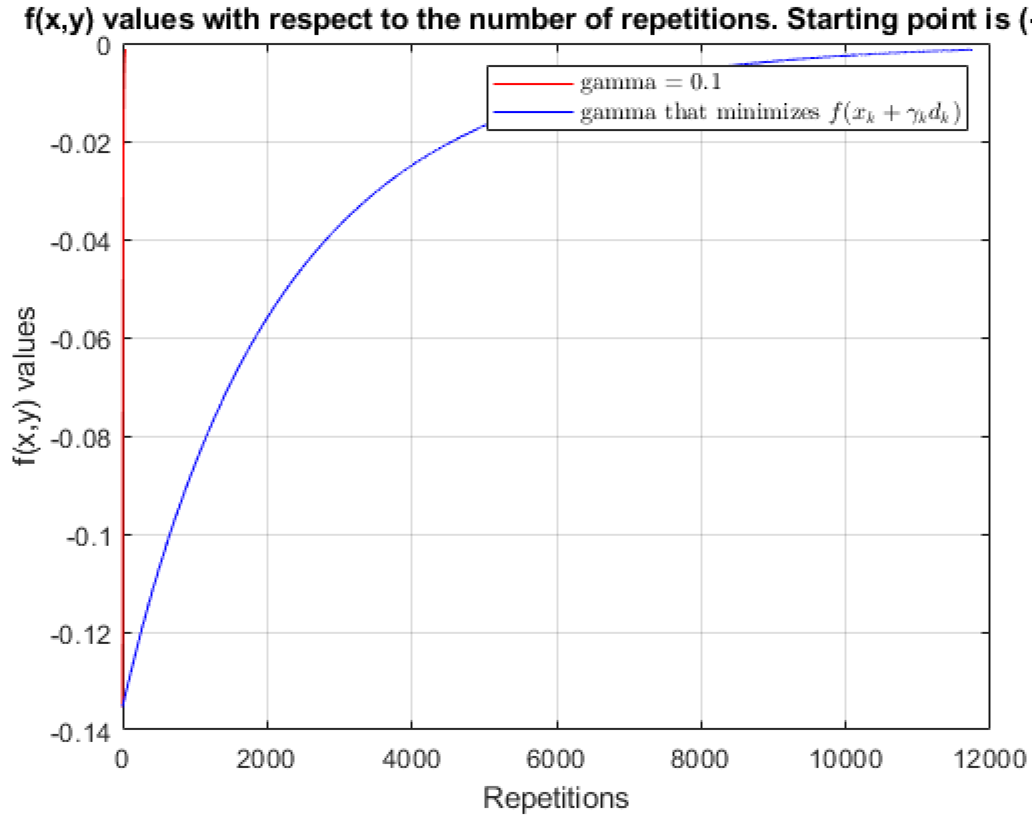


Θέμα 3

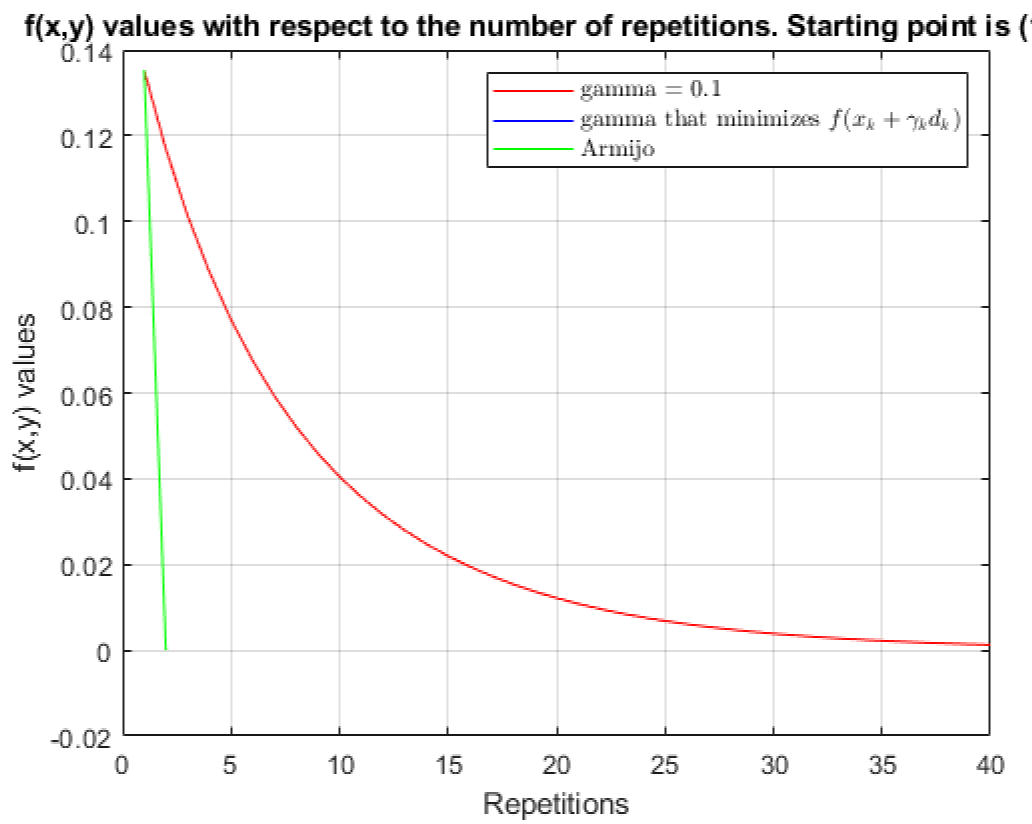
Τα ερωτήματα του θέματος 3 είναι τα ίδια με το 2, απλά με την μέθοδο Newton. Προϋπόθεση για την εφαρμογή αυτής είναι ο εσσιανός πίνακας της συνάρτησης να είναι θετικά ορισμένος, κάτι το οποίο δεν ισχύει σε αυτήν την περίπτωση. Έτσι ο αλγόριθμος έχει λάθος συμπεριφορά η οποία εξηγείται παρακάτω.

1. Σημείο (0,0). Ο αλγόριθμος δεν αρχίζει διότι είναι σημείο σάγματος και πληρεί την συνθήκη τερματισμού.
2. Σημείο (-1,1). Ο αλγόριθμος κινείται προς την κατεύθυνση της ανόδου $-\nabla f(x_k)$ επειδή ο εσσιανός δεν είναι θετικά ορισμένος. Έτσι καταλήγει στο 0.
   - Με το σταθερό βήμα, ο αλγόριθμος χρειάζεται 40 επαναλήψεις για να φτάσει στο σημείο σάγματος 0.
   - Με την ελαχιστοποίηση, ο αλγόριθμος χρειάζεται **11752 επαναλήψεις!** Κάτι τέτοιο συμβαίνει διότι η διαδικασία της βέλτιστης εύρεσης γ_κ γίνεται πολύ συντηρητική και το βήμα επιλέγεται πολύ μικρό (μάλιστα $\gamma_\kappa = 3.3083e-04$ σε κάθε επανάληψη), ώστε να αποφευχθεί η απότομη αύξηση της συνάρτησης (ή η μείωση της σε μη κατάλληλη κατεύθυνση).
   - Με τον κανόνα του Armijo η μέθοδος δεν συγκλίνει, διότι δεν μπορεί να βρεθεί βήμα το οποίο μειώνει την τιμή της συνάρτησης προς την κατεύθυνση που θέλουμε.
3. Σημείο (1,-1). Εδώ θα περιμέναμε ο αλγόριθμος να βρει το ολικό μέγιστο της συνάρτησης, αφού ο αλγόριθμος κινείται προς την κατεύθυνση της ανόδου $-\nabla f(x_k)$. Κάτι τέτοιο όμως δεν συμβαίνει και ο αλγόριθμος καταλήγει πάλι στο σημείο σάγματος 0. Αυτό οφείλεται στις πολύ μικρές αρνητικές ιδιοτιμές του εσσιανού πίνακα, καθώς και στην μικρή απόσταση του σημείου εκκίνησης με το σημείο σάγματος 0.
   - Με το σταθερό βήμα, ο αλγόριθμος χρειάζεται 40 επαναλήψεις.
   - Η ελαχιστοποίηση και ο κανόνας Armijo είναι ταυτόσημα και καταλήγουν στο 0 σε μόλις δύο επαναλήψεις.

Παρακάτω παρουσιάζονται τα διαγράμματα της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων μέχρι τον τερματισμό του αλγορίθμου. Προφανώς στο (-1,1) ο κανόνας armijo δεν παρουσιάζεται αφού δεν συγκλίνει.



f(x,y) values with respect to the number of repetitions. Starting point is (-1,1)
gamma = 0.1
gamma that minimizes $f(x_k + \gamma_k d_k)$
f(x,y) values
Repetitions
0
-0.02
-0.04
-0.06
-0.08
-0.1
-0.12
-0.14
0
2000
4000
6000
8000
10000
12000



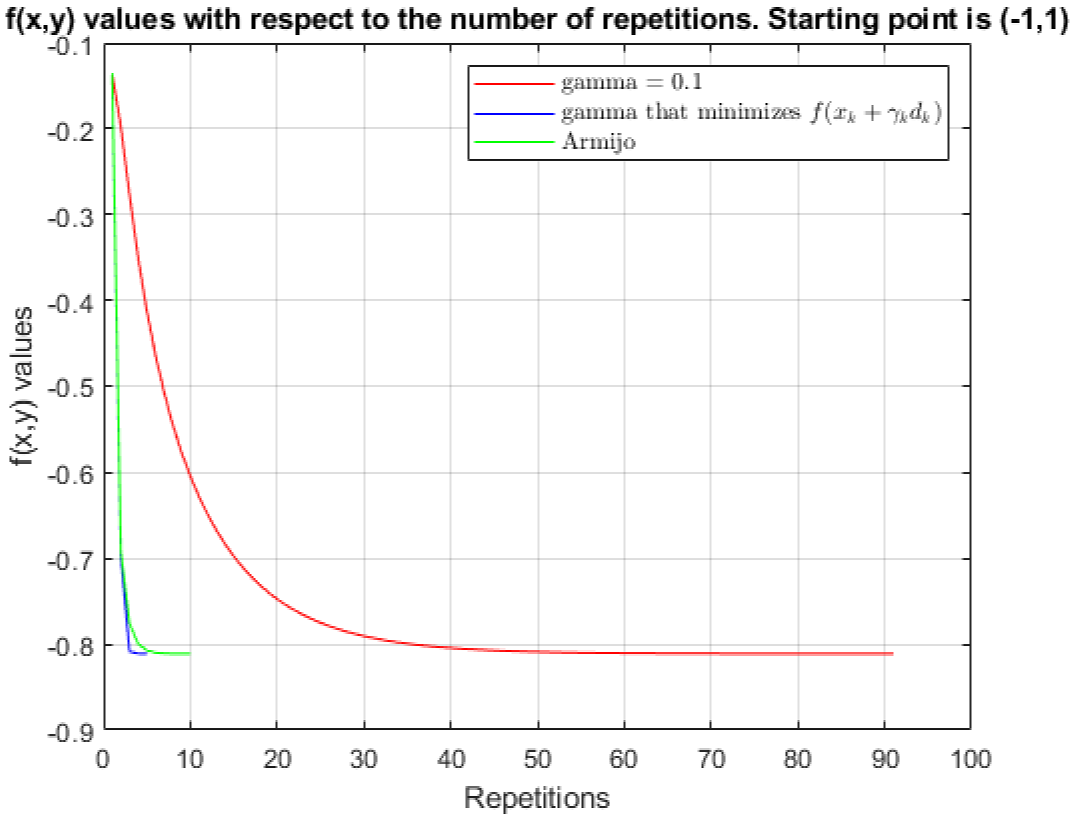
f(x,y) values with respect to the number of repetitions. Starting point is (1,-1)
gamma = 0.1
gamma that minimizes $f(x_k + \gamma_k d_k)$
Armijo
f(x,y) values
Repetitions
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
-0.02
0
5
10
15
20
25
30
35
40

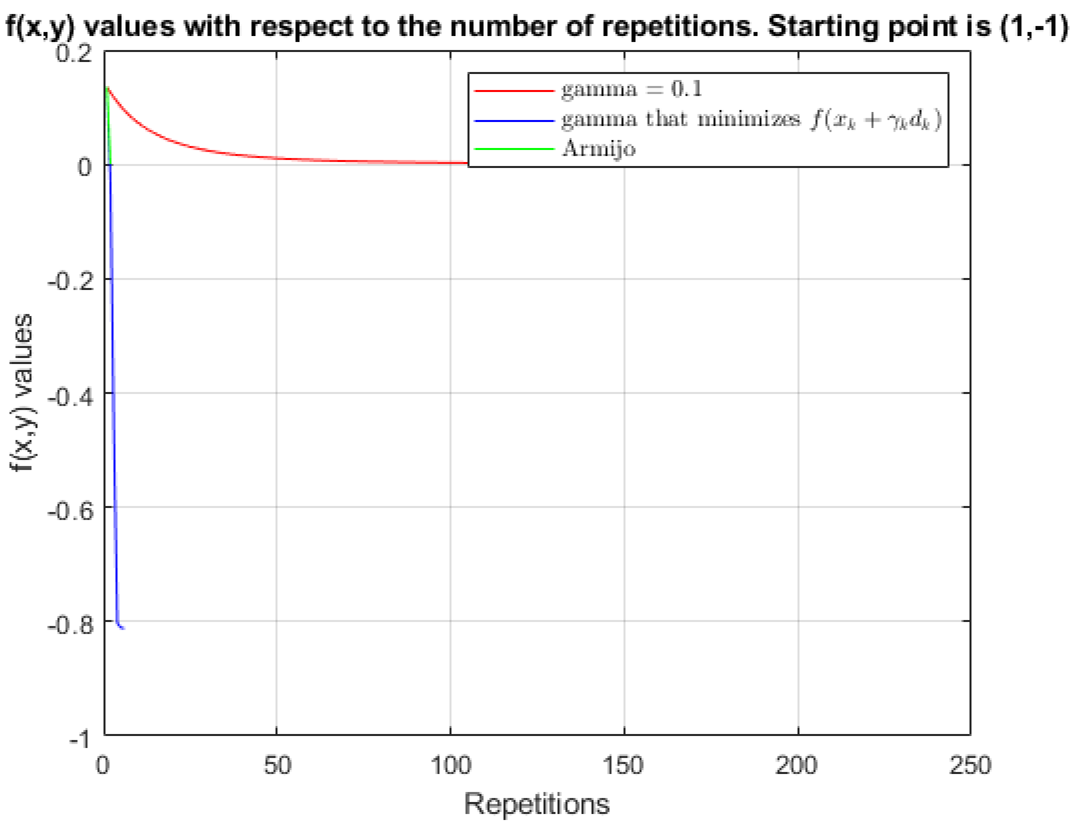
Θέμα 4

Τα ερωτήματα του θέματος 4 είναι τα ίδια με του 2ου και του 3ου, απλά με την μέθοδο Levenberg-Marquardt. Αυτή είναι ίδια με την Newton, με την μόνη διαφορά ότι $\Delta_\kappa = [\nabla^2 f(x_k) + \mu_k I]^{-1}$. Ο παράγοντας $\mu_k I$ προστίθεται ώστε ο εσσιανός να είναι θετικά ορισμένος. Φυσικά πρέπει το μ_k να είναι μεγαλύτερο της μεγαλύτερης ιδιοτιμής κατ’απόλυτο του εσσιανού. Στην συγκεκριμένη άσκηση θέτω $\mu_k = 1.5$ σε κάθε επανάληψη.

1. Σημείο (0,0). Ο αλγόριθμος δεν αρχίζει διότι είναι σημείο σάγματος και πληρεί την συνθήκη τερματισμού.
2. Σημείο (-1,1). Ο αλγόριθμος καταλήγει επιτυχώς στο ολικό ελάχιστο της συνάρτησης. Με σταθερό βήμα χρειάζεται, όπως είναι αναμενόμενο, τις περισσότερες επαναλήψεις, οι οποίες είναι 91. Η ελαχιστοποίηση χρειάζεται τις λιγότερες, 5 επαναλήψεις. Ο κανόνας Armijo χρειάζεται 10 επαναλήψεις.
3. Σημείο (1,-1). Με το σταθερό βήμα ο αλγόριθμος καταλήγει στο σημείο σάγματος 0 σε 207 επαναλήψεις. Με τον κανόνα του Armijo καταλήγει πάλι στο 0 σε 2 επαναλήψεις. Όμως με την ελαχιστοποίηση, ο αλγόριθμος καταλήγει επιτυχώς στο ολικό ελάχιστο της συνάρτησης σε 6 επαναλήψεις.

Παρακάτω παρουσιάζονται τα διαγράμματα της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων μέχρι τον τερματισμό του αλγορίθμου.


f(x,y) values with respect to the number of repetitions. Starting point is (-1,1)
gamma = 0.1
gamma that minimizes $f(x_k + \gamma_k d_k)$
Armijo
f(x,y) values
Repetitions


f(x,y) values with respect to the number of repetitions. Starting point is (1,-1)
gamma = 0.1
gamma that minimizes $f(x_k + \gamma_k d_k)$
Armijo
f(x,y) values
Repetitions

Παρακάτω παρουσιάζονται πίνακες στους οποίους φαίνεται ο αριθμός των επαναλήψεων σε κάθε μέθοδο, ανάλογα με το βήμα.

Σημείο εκκίνησης (-1,1)	Μέγιστη κάθοδος	Newton (μη θετικά ορισμένος εσσιανός)	Levenberg Marquardt
Σταθερό $\gamma_k = 0.1$	38 επαναλήψεις → Ολικό ελάχιστο	40 επαναλήψεις → Σημείο σάγματος	91 επαναλήψεις→ Ολικό ελάχιστο
Ελαχιστοποίηση $f(x_k + \gamma_k d_k)$	6 επαναλήψεις → Ολικό ελάχιστο	11752 επαναλήψεις→ Σημείο σάγματος	5 επαναλήψεις→ Ολικό ελάχιστο
Κανόνας Armijo	11 επαναλήψεις → Ολικό ελάχιστο	Δεν συγκλίνει	10 επαναλήψεις→ Ολικό ελάχιστο

Σημείο εκκίνησης (1,-1)	Μέγιστη κάθοδος	Newton (μη θετικά ορισμένος εσσιανός)	Levenberg Marquardt
Σταθερό $\gamma_k = 0.1$	117 επαναλήψεις→ Σημείο σάγματος	40 επαναλήψεις → Σημείο σάγματος	207 επαναλήψεις→ Σημείο σάγματος
Ελαχιστοποίηση $f(x_k + \gamma_k d_k)$	2 επαναλήψεις → Σημείο σάγματος	2 επαναλήψεις→ Σημείο σάγματος	6 επαναλήψεις→ Ολικό ελάχιστο
Κανόνας Armijo	2 επαναλήψεις → Σημείο σάγματος	2 επαναλήψεις → Σημείο σάγματος	2 επαναλήψεις→ Σημείο σάγματος

Από τους παραπάνω μπορούμε να συμπεράνουμε πως ο πιο αποδοτικός αλγόριθμος είναι ο Levenberg-Marquardt, καθώς καταλήγει στο ολικό ελάχιστο περισσότερες φορές από τους άλλους και σε λιγότερες επαναλήψεις (πέρα από όταν το βήμα είναι σταθερό). Ακολουθεί η μέθοδος της μέγιστης καθόδου, η οποία είναι αρκετά αξιόπιστη και τερματίζει σε λίγες επαναλήψεις. Έπειτα ακολουθεί η μέθοδος Newton η οποία καταλήγει συνέχεια σε σημείο σάγματος. Αυτό είναι λογικό στην συγκεκριμένη περίπτωση λόγω του μη θετικά ορισμένου εσσιανού.

Σημείωση: Η Levenberg-Marquardt λειτουργεί όπως η μέθοδος της μέγιστης καθόδου όταν το μ_k είναι αρκετά μεγάλο, διότι ο παράγοντας $\mu_k I$ κυριαρχεί σε σχέση με τον $\nabla^2 f(x_k)$. Όταν όμως το μ_k είναι πιο μικρό, τότε ο $\nabla^2 f(x_k)$ υπερισχύει και η μέθοδος συμπεριφέρεται σαν τη μέθοδο Newton.