



ΑΡΙΣΤΟΤΕΛΕΙΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΕΣΣΑΛΟΝΙΚΗΣ

Πολυτεχνική Σχολή

Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών
Υπολογιστών

2η Εργασία

Τζομίδης Νικόλαος-Φώτιος
ΑΕΜ: (9461)
tzomidis@ece.auth.gr

Τεχνικές Βελτιστοποίησης (7ο εξάμηνο)

7 Δεκεμβρίου 2021

Περιεχόμενα

Περιεχόμενα	i
Λίστα Σχημάτων	ii
1 2η Εργαστηριακή Άσκηση	1
1.1 Θέμα 1	1
1.2 Θέμα 2 (Μεγίστης Καθόδου)	2
1.3 Θέμα 3 (Newton)	6
1.4 Θέμα 4 (Levenberg-Marquardt)	8
1.5 Συμπεράσματα	11

Λίστα Σχημάτων

1.1	Σχεδίαση της f	2
1.2	Μέθοδος Μεγίστης Καθόδου για αρχικό σημείο $(-1,-1)$	3
1.3	Μέθοδος Μεγίστης Καθόδου για αρχικό σημείο $(1,1)$	5
1.4	Μέθοδος Newton για αρχικό σημείο $(-1,-1)$	7
1.5	Μέθοδος Newton για αρχικό σημείο $(1,1)$ με κανόνα Armijo	8
1.6	Μέθοδος Levenberg-Marquardt για αρχικό σημείο $(0,0)$	9
1.7	Μέθοδος Levenberg-Marquardt για αρχικό σημείο $(-1,-1)$	10
1.8	Μέθοδος Levenberg-Marquardt για αρχικό σημείο $(1,1)$ με κανόνα Armijo	11

Κεφάλαιο 1

2η Εργαστηριακή Άσκηση

Στα πλαίσια της 2η εργαστηριακής άσκησης στις Τεχνικές Βελτιστοποίησης θα ασχοληθούμε με την ελαχιστοποίηση μιας δοσμένης συνάρτησης $f : \mathbb{R}^n \rightarrow \mathbb{R}$ χωρίς περιορισμούς. Χρησιμοποιήθηκαν οι εξής 3 αλγόριθμοι:

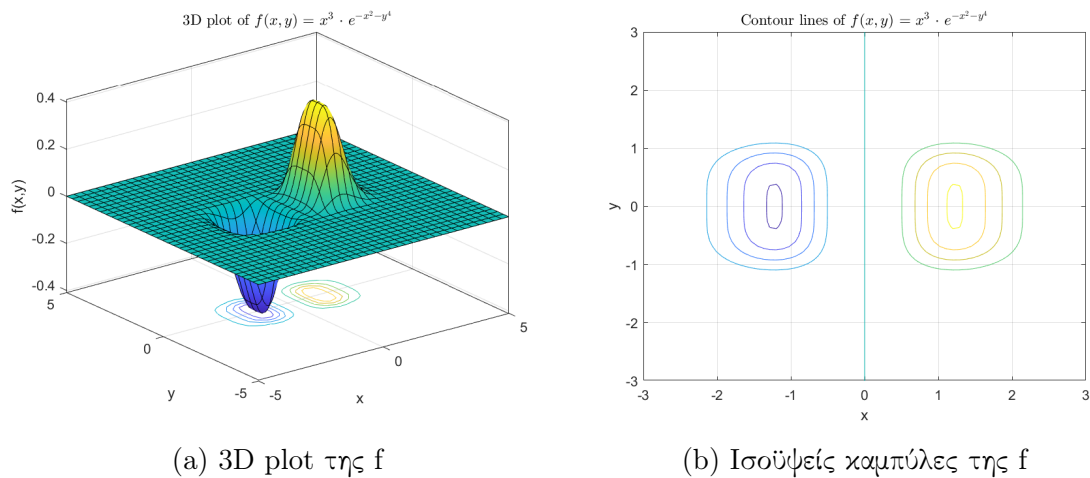
- Μέθοδος Μεγίστης Καθόδου(Steepest Descent)
- Μέθοδος Newton
- Μέθοδος Levenberg-Marquardt

Οι τρεις αυτές μέθοδοι ανήκουν στις μεθόδους κλίσης και βασική ιδιότητα τους είναι η χρήση του διανύσματος κλίσης, $\nabla f(x)$, στην ισοβαρή καμπύλη της f . Η αντικειμενική συνάρτηση που θα μελετήσουμε είναι η:

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

1.1 Θέμα 1

Με το αρχείο `part1_plots.m` σχεδιάζουμε την f για να πάρουμε μια καλύτερη οπτική μορφή της. Το αρχείο παράγει τόσο την μορφή της f στον τρισδιάστατο χώρο όσο και τις ισοϋψείς καμπύλες της.

(a) 3D plot της f (b) Ισοϋψείς καμπύλες της f Figure 1.1: Σχεδίαση της f

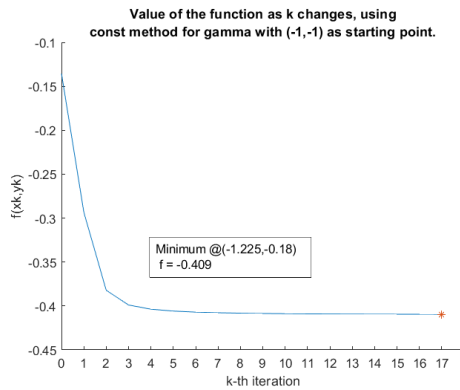
Από το διάγραμμα της f στον τρισδιάστατο χώρο παρατηρούμε πως η ελάχιστη τιμή εμφανίζεται στα αρνητικά x και κοντά στην γειτονιά του 0 στον άξονα των y . Αντίστοιχα παρουσιάζει μέγιστο σε θετικά x στην γειτονιά του 0 των y . Τέλος παρατηρούμε και ένα τοπικό ελάχιστο στο σημείο $(0,0)$.

1.2 Θέμα 2 (Μεγίστης Καθόδου)

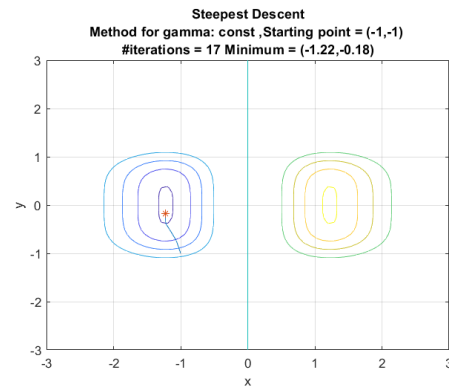
Στο Θέμα 2 ελαχιστοποιούμε την f με την χρήση της μεθόδου Μεγίστης Καθόδου. Υλοποιούμε 3 διαφορετικά .m αρχεία ανάλογα με τον τρόπο επιλογής του βήματος γ_k όπου (α) σταθερό(της επιλογής μας), (β) τέτοιο ώστε να ελαχιστοποιεί την $f(x_k + \gamma_k \cdot d_k)$, (γ) βάσει του κανόνα Armijo. Για κάθε μέθοδο χρησιμοποιούμε ως αρχικά σημεία (x_0, y_0) τα (i) $(0,0)$, (ii) $(-1,-1)$, (iii) $(1,1)$. Τα διαγράμματα που προκύπτουν από το part2_plots.m μας βοηθούν να δούμε τις διαφορές στα αποτελέσματα τόσο λόγω διαφορετικού σημείου έναρξης (x_0, y_0) όσο και λόγω διαφορετικού αλγορίθμου επιλογής του βήματος γ_k αλλά και το πως συγκλίνει η συνάρτηση στο ελάχιστο με κάθε επανάληψη.

Αρχικά για το $(0,0)$ η μέθοδος εγκλωβίζεται σε αυτό το σημείο μιας και αυτό είναι τοπικό ελάχιστο της συνάρτησης και έτσι δεν εκτελείται καμία επανάληψη (τα διαγράμματα λοιπόν παραλείπονται καθώς δεν παρουσιάζουν κανένα ενδιαφέρον). Εάν θέσουμε ως αρχικό σημείο το $(-1,-1)$ παρατηρούμε ότι ο αλγόριθμος φτάνει πολύ κοντά στο ελάχιστο σημείο για δικιά μας

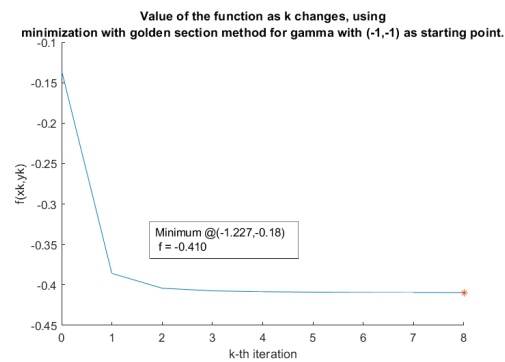
επιλογή $\epsilon = 0.01$ και για τις 3 διαφορετικές μεθόδους επιλογής του γ_k .



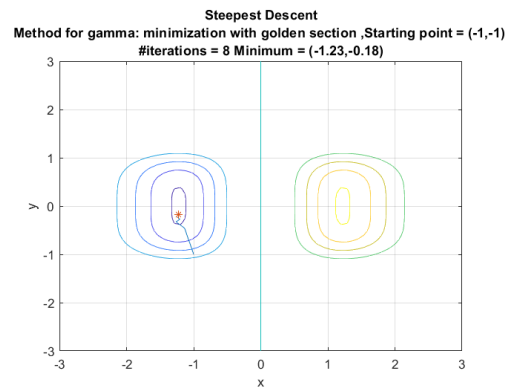
(α) Τιμές της f για μεταβαλλόμενο x (α)



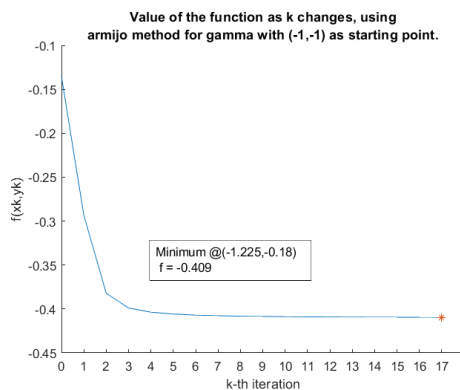
(β) Σύγκλιση του ζεύγους (x,y) (α)



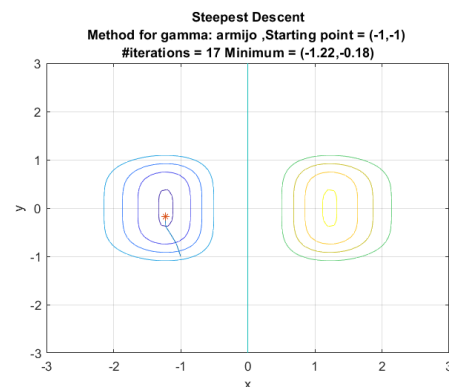
(γ) Τιμές της f για μεταβαλλόμενο x (β)



(δ) Σύγκλιση του ζεύγους (x,y) (β)



(ε) Τιμές της f για μεταβαλλόμενο x (γ)



(ς) Σύγκλιση του ζεύγους (x,y) (γ)

Figure 1.2: Μέθοδος Μεγίστης Καθόδου για αρχικό σημείο $(-1,-1)$

Από τα παραπάνω διαγράμματα φαίνεται πως η μέθοδος (β) έχει αρκετά πιο

γρήγορη σύγκλιση συγκριτικά με την μέθοδο όπου το γ_k είναι σταθερό ή την μέθοδο Armijo. Αυτό γίνεται εύκολα αντιληπτό τόσο από την απότομη κλίση του διαγράμματος όσο και από τον αριθμό των επαναλήψεων για την μέθοδο (8 έναντι 17 επαναλήψεων στις άλλες 2 μεθόδους). Να σημειωθεί εδώ ότι για την εύρεση του ελάχιστου γ_k στην (β) μέθοδο χρησιμοποιήθηκε ο αλγόριθμος του χρυσού τομέα από την 1η εργασία (gold_sect.m) για την ελαχιστοποίηση της $f(x_k + \gamma_k \cdot d_k)$.

Τέλος αν πάρουμε σαν αρχικό σημείο το (1,1) ο αλγόριθμος θα εγκλωβιστεί και πάλι στο τοπικό ελάχιστο της αρχής των αξόνων και αδυνατεί να συνεχίσει προς το πραγματικό ελάχιστο της συνάρτησης. Παρόλα αυτά η μέθοδος (β) αποδεικνύεται πιο αποτελεσματική παρότι δεν παίρνουμε 'πραγματικό αποτέλεσμα'.

* Σε όλα τα διαγράμματα του 2ου θέματος χρησιμοποιήθηκε $\epsilon = 0.01$ ενώ για τις περιπτώσεις (α) όπου ζητούταν σταθερό γ_k τέθηκε αυθαίρετα $\gamma_k = 0.5$. Στην μέθοδο Armijo επιλέχτηκαν αρχικές συνθήκες $\alpha = 0.002$ και $\beta = 0.25$. Ενώ η χρήση του αλγορίθμου του χρυσού τομέα έχει τροποποιηθεί ελάχιστα για να μας δίνει μια ελάχιστη τιμή και όχι διάστημα στο οποίο θα βρίσκεται μέσα η ελαχιστη τιμή μας.

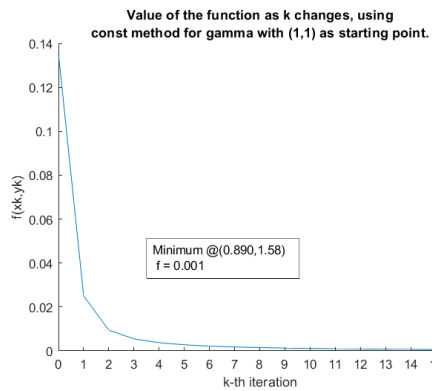
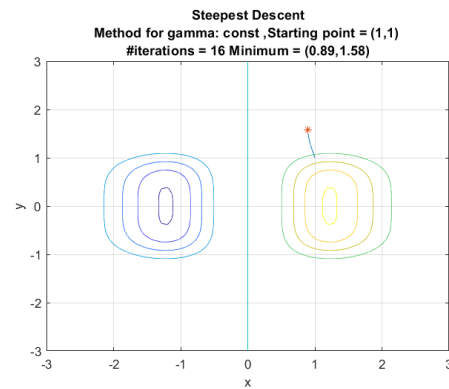
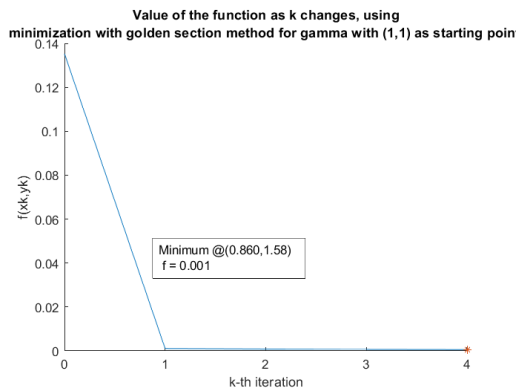
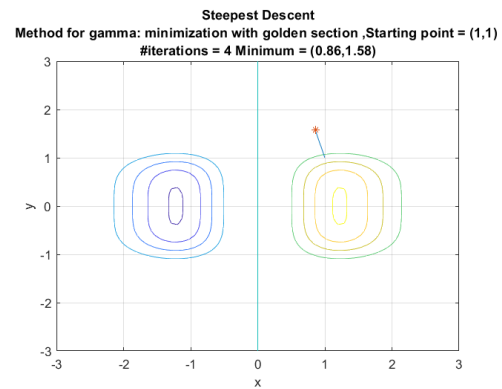
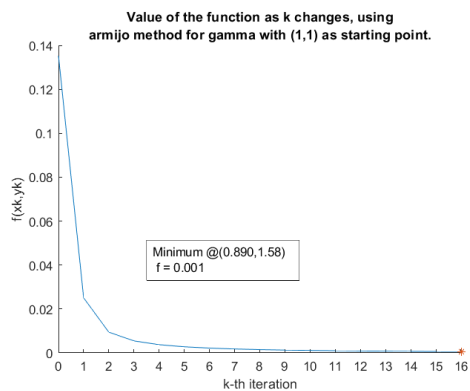
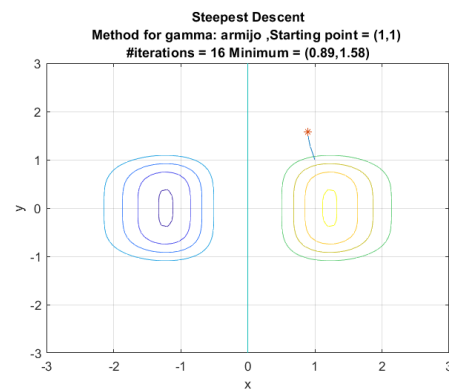
(α) Τιμές της f για μεταβαλλόμενο x (α)(b) Σύγκλιση του ζεύγους (x,y) (α)(c) Τιμές της f για μεταβαλλόμενο x (β)(d) Σύγκλιση του ζεύγους (x,y) (β)(e) Τιμές της f για μεταβαλλόμενο x (γ)(f) Σύγκλιση του ζεύγους (x,y) (γ)

Figure 1.3: Μέθοδος Μεγίστης Καθόδου για αρχικό σημείο (1,1)

1.3 Θέμα 3 (Newton)

Αντίθετα με την μέθοδο μεγίστης καθόδου η μέθοδος Newton αναζητεί σημείο ελαχίστου στην κατεύθυνση $d_k = -[\nabla^2 f(x_k)]^{-1} \cdot \nabla f(x_k)$ χρησιμοποιεί δηλαδή και τον εσσιανό πίνακα ο οποίος θα πρέπει όμως να είναι θετικά ορισμένος. Παρακάτω παράγουμε με την ίδια λογική τα διαγράμματα για να δούμε την συμπεριφορά της μεθόδου στα διαφορετικά αρχικά σημεία και τις διαφορετικές επιλογές του γ_k .

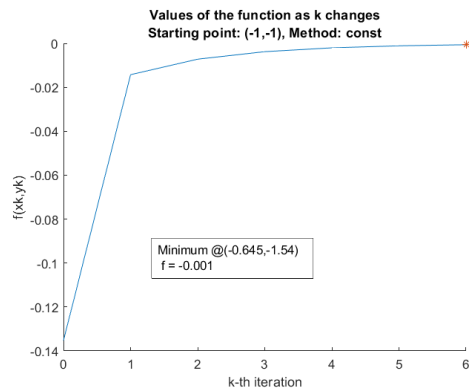
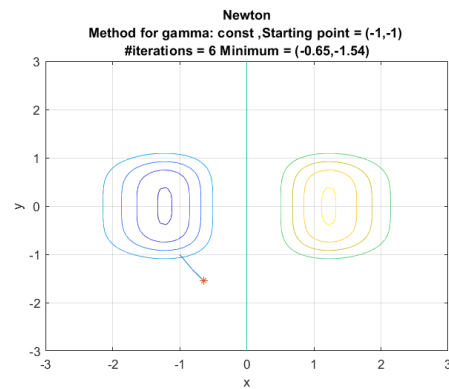
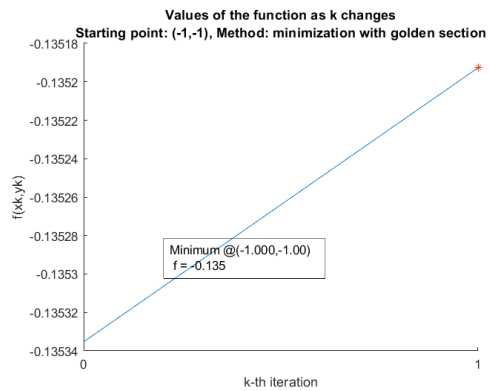
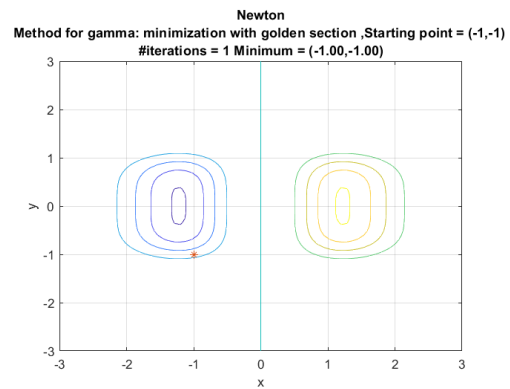
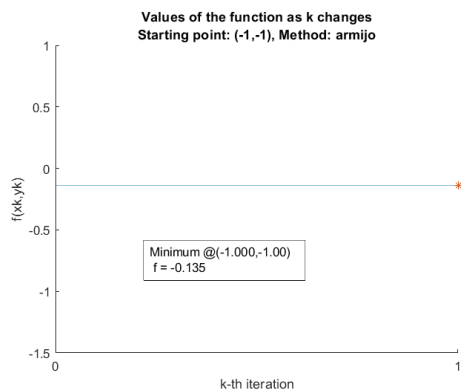
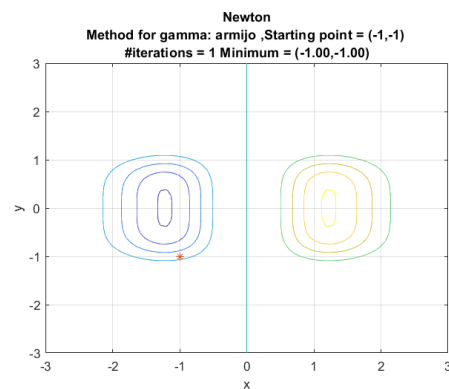
Στο σημείο (0,0) η συνάρτηση εγκλωβίζεται ξανά σε τοπικό ελάχιστο (μιας και η κλίση της συνάρτησης στο σημείο είναι 0).

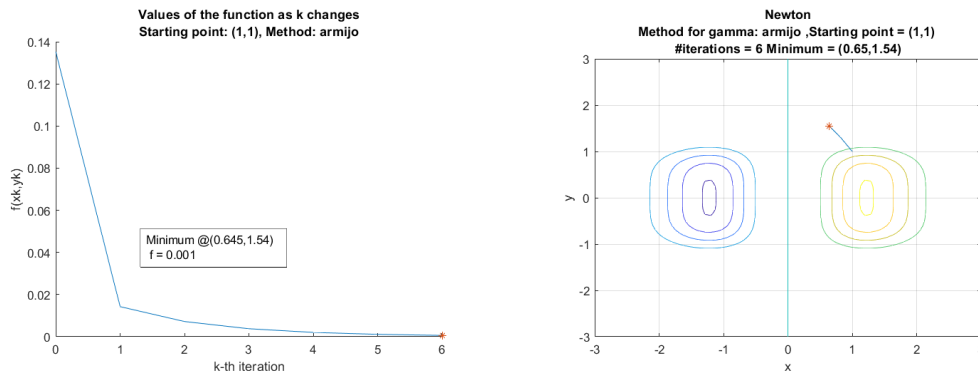
Πριν περάσουμε στα διαγράμματα για αρχικό σημείο (-1,-1) μπορούμε να παρατηρήσουμε ότι ο εσσιανος πίνακας μας δεν είναι θετικά ορισμένος και ο αλγόριθμος δεν συγκλίνει ποτέ σε σημείο ελαχίστου. Για αυτό τον λόγο έχουμε θέσει τον αλγόριθμο στις μεθόδους (β) και (γ) να σταματάει όταν η τιμή της f για x_k γίνει μεγαλύτερη ή ίση της τιμής της συνάρτησης στο x_{k-1} . Παρατηρούμε επίσης πως λόγω του μη θετικά ορισμένου εσσιανού πίνακα για σταθερό γ_k η τιμή της συνάρτησης αυξάνεται αντί να μειώνεται.

Για το σημείο (1,1) αντιμετωπίζουμε το ίδιο πρόβλημα εγκλωβισμού της μεθόδου σε τοπικό ελάχιστο. Για το τελευταίο αρχικό σημείο θα παρουσιάσουμε μόνο την μέθοδο Armijo μιας και δεν παρατηρούνται διαφορές στις υπόλοιπες.

Είναι εύλογο να πούμε ότι η μέθοδος Newton δεν μας δίνει καθόλου καλά συμπεράσματα για ελάχιστο της συνάρτησης εξαιτίας της μορφής αυτής και του μη θετικά ορισμένου εσσιανού πίνακα της.

* Παρόμοια με παραπάνω σε όλα τα διαγράμματα του 3ου θέματος χρησιμοποιήθηκε $\epsilon = 0.01$ ενώ για τις περιπτώσεις (α) όπου ζητούταν σταθερό γ_k τέθηκε αυθαίρετα $\gamma_k = 0.5$. Στην μέθοδο Armijo επιλέχτηκαν αρχικές συνθήκες $\alpha = 0.002$ και $\beta = 0.25$. Όλα τα παρακάτω διαγράμματα δημιουργήθηκαν μέσω του `part3_plots.m`.

(α) Τιμές της f για μεταβαλλόμενο x (α)(β) Σύγκλιση του ζεύγους (x,y) (α)(γ) Τιμές της f για μεταβαλλόμενο x (β)(δ) Σύγκλιση του ζεύγους (x,y) (β)(ε) Τιμές της f για μεταβαλλόμενο x (γ)(ς) Σύγκλιση του ζεύγους (x,y) (γ)Figure 1.4: Μέθοδος Newton για αρχικό σημείο $(-1,-1)$



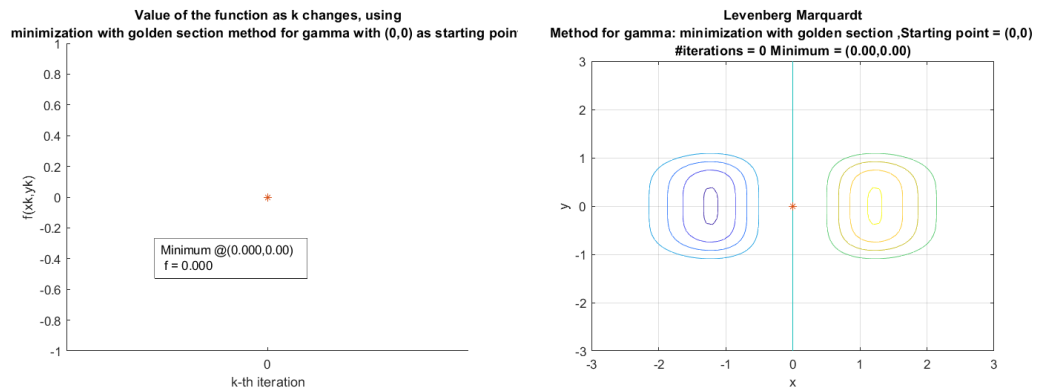
(a) Τιμές της f για μεταβαλλόμενο k (γ) (b) Σύγκλιση του ζεύγους (x, y) (γ)

Figure 1.5: Μέθοδος Newton για αρχικό σημείο $(1, 1)$ με κανόνα Armijo

1.4 Θέμα 4 (Levenberg-Marquardt)

Στο Θέμα 4 γίνεται η χρήση του Αλγορίθμου Levenberg-Marquardt για την ελαχιστοποίηση της f . Πρόκειται ουσιαστικά για έναν τροποποιημένο αλγόριθμο Newton που μας επιτρέπει να παρακάμψουμε την αδυναμία του Newton σε περίπτωση που ο $\nabla^2 f(x_k)$ δεν είναι θετικά ορισμένος. Αυτό επιτυγχάνεται με επιλογή $d_k = [\nabla^2 f(x_k) + \mu_k \cdot I]^{-1}$ όπου $\mu_k > 0$ και με κατάλληλη επιλογή του ο $\nabla^2 f(x_k) + \mu_k \cdot I$ θα είναι θετικά ορισμένος.

Για αρχικό σημείο εκκίνησης του αλγορίθμου το $(0, 0)$ παρουσιάζονται τα ίδια προβλήματα με προηγουμένως όπου υπάρχει δηλαδή εγκλωβισμός σε τοπικό ελάχιστο. Στο διάγραμμα παρακάτω βλέπουμε ότι δεν υπάρχει καμία επανάληψη και ο αλγόριθμος μένει στάσιμος.



(a) Τιμές της f για μεταβαλλόμενο x (y) (b) Σύγκλιση του ζεύγους (x,y) (y)

Figure 1.6: Μέθοδος Levenberg-Marquardt για αρχικό σημείο $(0,0)$

Για αρχικό σημείο $(-1,-1)$ αυτή την φορά ο αλγόριθμός λειτουργεί καλά λόγω του ότι ο αλγόριθμος παρακάμπτει τον μη θετικά ορισμένο εσσιανό πίνακα της συνάρτησης μας.

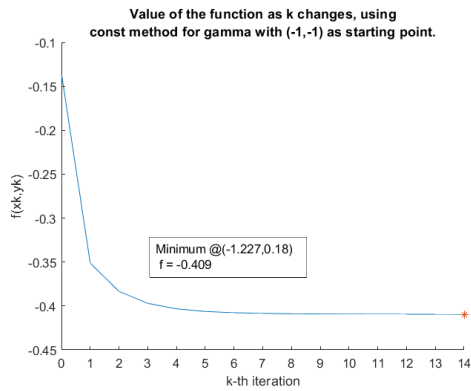
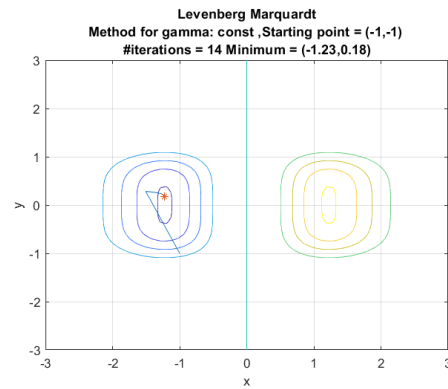
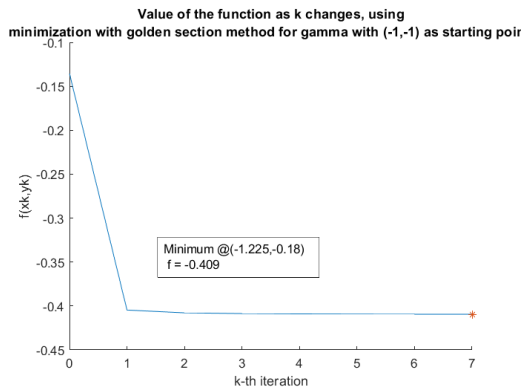
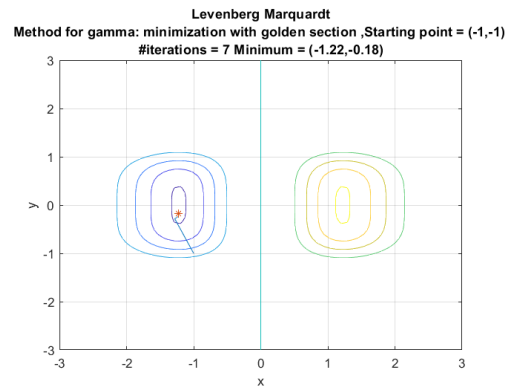
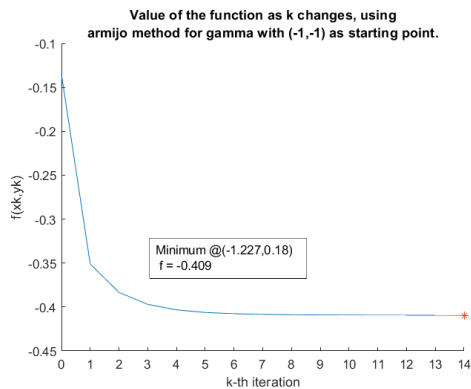
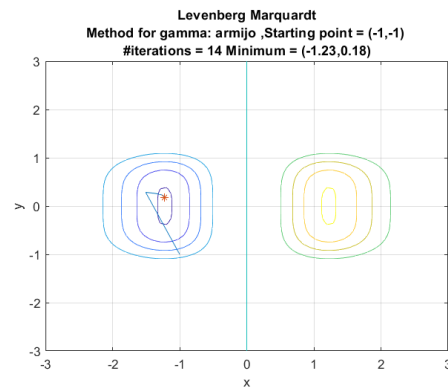
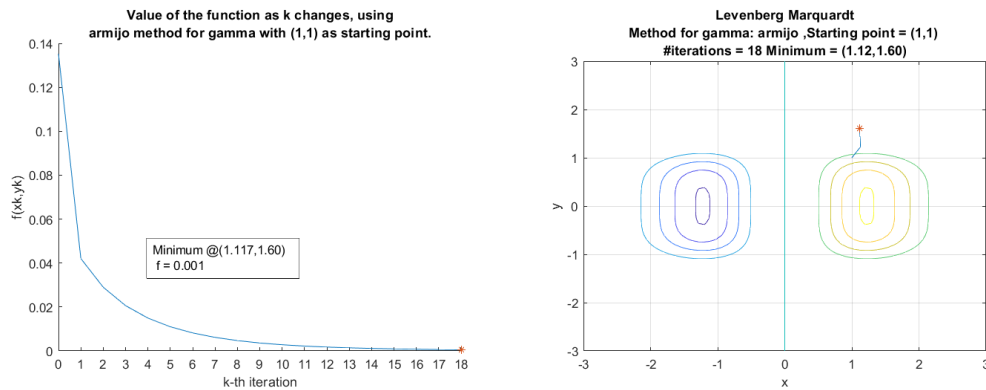
(α) Τιμές της f για μεταβαλλόμενο x (α)(β) Σύγκλιση του ζεύγους (x, y) (α)(γ) Τιμές της f για μεταβαλλόμενο x (β)(δ) Σύγκλιση του ζεύγους (x, y) (β)(ε) Τιμές της f για μεταβαλλόμενο x (γ)(ς) Σύγκλιση του ζεύγους (x, y) (γ)

Figure 1.7: Μέθοδος Levenberg-Marquardt για αρχικό σημείο (-1,-1)

Για αρχικό σημείο $(1,1)$ ο αλγόριθμος τρέχει μέχρι να φτάσει κοντά στην τιμή του 0 όπου και εγκλωβίζεται, παρακάτω παρουσιάζουμε αυτή την συμπεριφορά με την μέθοδο Armijo.



(a) Τιμές της f για μεταβαλλόμενο x (y) (b) Σύγκλιση του ζεύγους (x,y) (y)

Figure 1.8: Μέθοδος Levenberg-Marquardt για αρχικό σημείο $(1,1)$ με κανόνα Armijo

1.5 Συμπεράσματα

Καταλήγουμε ότι για το μόνο αρχικό σημείο που εξάγονται ασφαλή συμπεράσματα είναι το $(-1,-1)$. Για αρχικό σημείο $(0,0)$ οι μέθοδοι μένουν στάσιμοι χωρίς να έχουμε καμία επανάληψη. Ενώ για αρχικό σημείο $(1,1)$ οι αλγόριθμοι εγκλωβίζονται στην πλευρά που στη τρισδιάστατη αναπαράσταση της f βλέπουμε θετικές τιμές z .

Επίσης η μέθοδος Newton αποκλείεται εντελώς καθώς η $f(x,y)$ δεν ικανοποιεί την συνθήκη της μεθόδου όπου ο εσσιανός της πίνακας θα πρέπει να είναι θετικά ορισμένος. Έτσι δεν δίνει "σωστές λύσεις" για κανένα αρχικό σημείο.

Στις άλλες δύο μεθόδους όπου παίρνουμε επιθυμητές τιμές για το αρχικό σημείο $(-1,-1)$ φαίνεται η περίπτωση επιλογής του γ_k μέσω ελαχιστοποίησης της $f(x_k + \gamma_k \cdot d_k)$ να είναι η πιο αποδοτική με τον μικρότερο αριθμό επαναλήψεων. Να σημειωθεί βέβαια ότι εργαστήκαμε με σταθερό $\gamma_k = 0,5$, για μεγαλύτερες τιμές του κοντά στο 1 η μέθοδος συγκλίνει πιο γρήγορα χωρίς όμως να υπερτερεί της (β) μεθόδου.

Τέλος ο αλγόριθμος Levenberg-Marquardt μιας και είναι μια βελτιωμένη μορφή του Newton που χρησιμοποιεί και την μέγιστη κάθοδο φαίνεται ότι συγκλίνει πιο γρήγορα και με πιο ικανοποιητικό αριθμό βημάτων.