

Τεχνικές Βελτιστοποίησης — 2ο Παραδοτέο

Ελαχιστοποίηση συναρτήσεων πολλών μεταβλητών χωρίς περιορισμούς
με χρήση παραγώγων: Steepest Descent, Newton, Levenberg-Marquardt

Χρήστος Αλεξόπουλος-10618

Contents

1	Εισαγωγή	2
2	Μέθοδος Steepest Descent	3
3	Μέθοδος Newton	8
4	Μέθοδος Levenberg–Marquardt	13
5	Παρατηρήσεις και Συμπεράσματα	18
6	Αρχιτεκτονική και Οργάνωση Κώδικα	21

1 Εισαγωγή

Στόχος της εργασίας είναι η πειραματική μελέτη τριών μεθόδων πολυδιάστασης αναζήτησης ελαχίστου: Steepest Descent, Newton, Levenberg–Marquardt. Φτιάχνουμε διαγράμματα για τη σύγκλιση των αλγορίθμων αναζήτησης και τον αριθμό επαναλήψεων. Η υλοποίηση των μεθόδων βασίζεται στις παραγράφους (αλγόριθμοι) του βιβλίου 5.2.1, 5.2.2, 5.2.3 και η επιλογή βήματος από τις σελίδες 140–141.

Η αντικειμενική συνάρτηση που θα μελετήσουμε είναι:

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

Παρακάτω παρουσιάζονται δύο γραφικές παραστάσεις της συνάρτησης, ώστε να αποκτήσουμε μία γενική εικόνα της μορφής της.

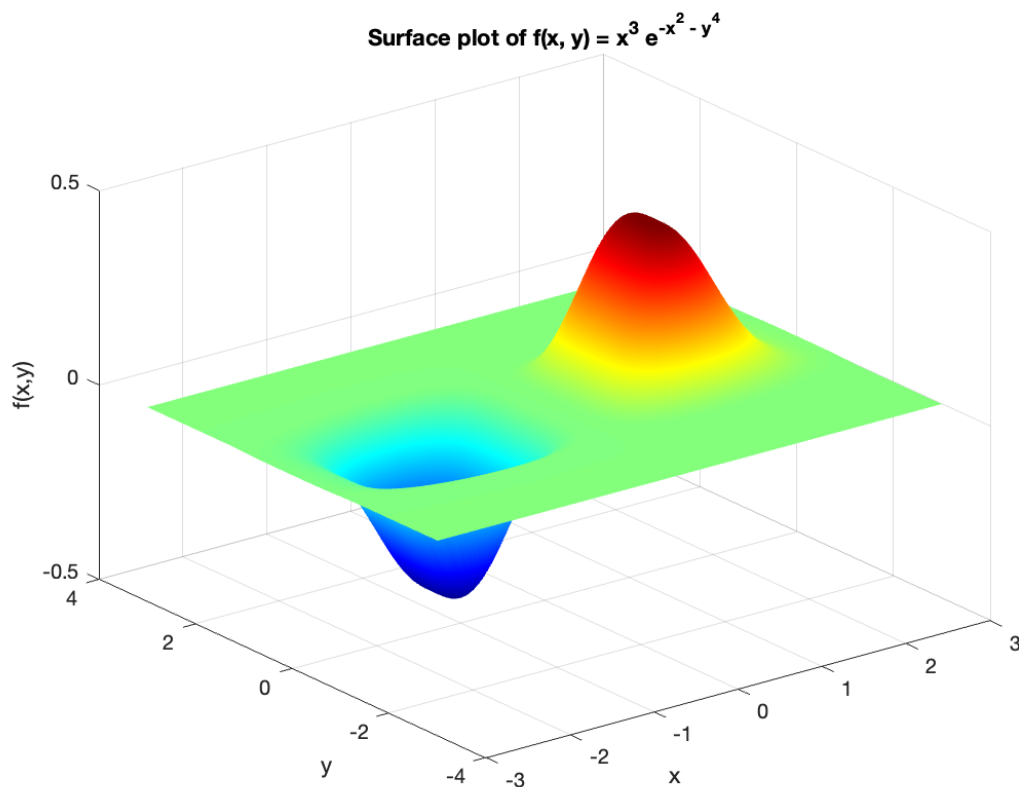


Figure 1: Τρισδιάστατη σχεδίαση της συνάρτησης $f(x, y)$.

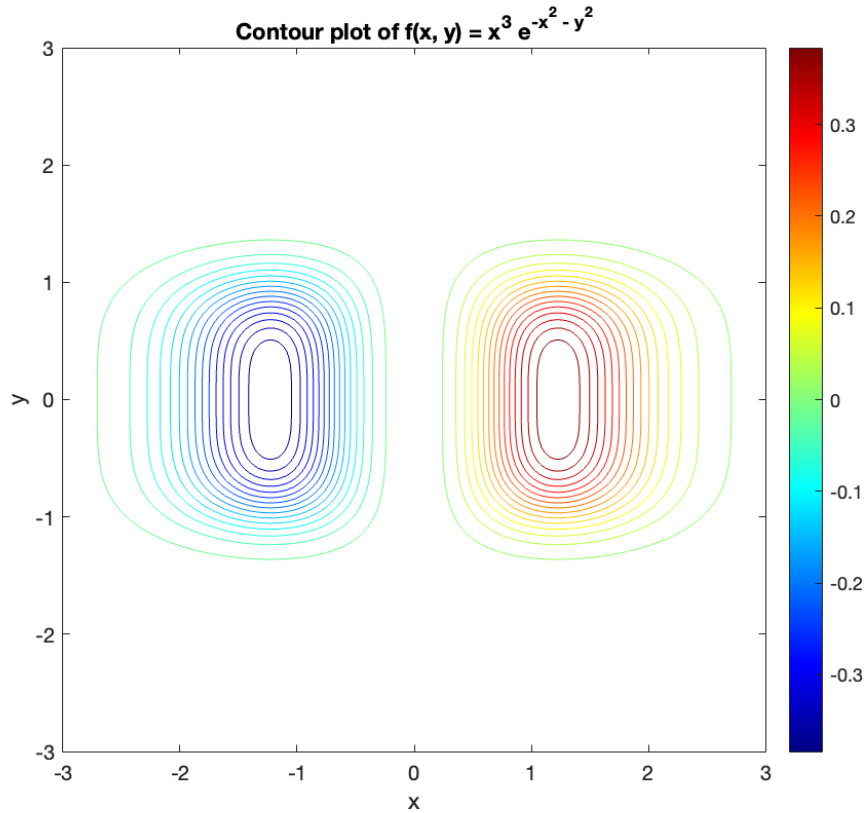


Figure 2: Ισοσταθμικές καμπύλες (contour plot) της συνάρτησης $f(x, y)$.

2 Μέθοδος Steepest Descent

1) Σύνοψη

Ακολουθούμε πιστά τον αλγόριθμο 5.2.1 στη σελίδα 121. Υπολογίζουμε το

$$\nabla f(x, y) = \begin{pmatrix} \frac{\partial f}{\partial x}(x, y) \\ \frac{\partial f}{\partial y}(x, y) \end{pmatrix}$$

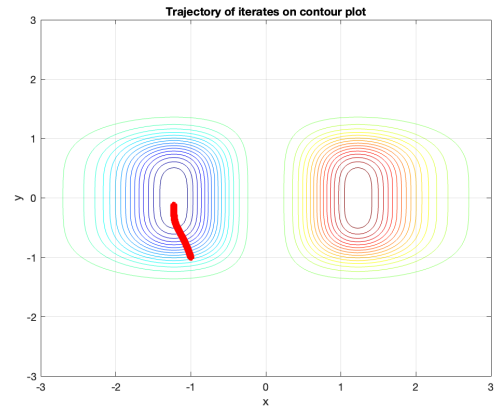
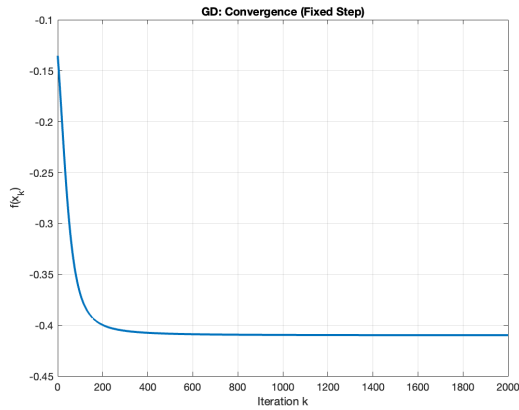
να είναι ίσο με

$$\nabla f(x, y) = \begin{pmatrix} (3x^2 - 2x^4) e^{-(x^2+y^4)} \\ -4x^3 y^3 e^{-(x^2+y^4)} \end{pmatrix}$$

Αποτελέσματα Μεθόδου Steepest Descent

Για τα σημεία εκκίνησης $(-1, -1)$ $(0, 0)$ $(1, 1)$ έχουμε τα εξής αποτελέσματα για κάθε επιλογή βήματος.

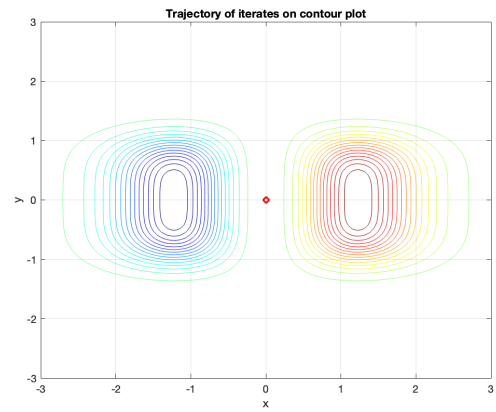
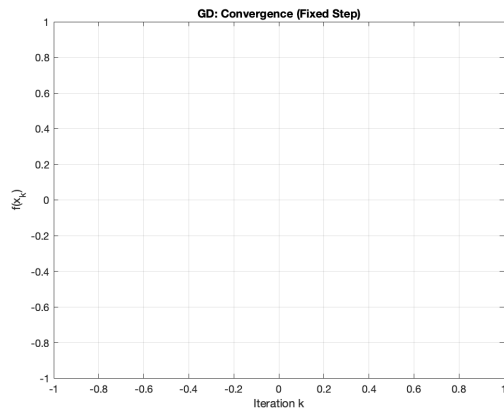
Μέθοδος Steepest Descent με σταθερό βήμα



(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.

(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

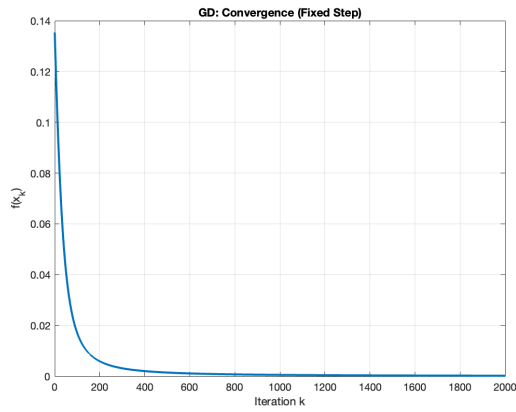
Figure 3: Αποτελέσματα για σταθερό βήμα, αρχικό σημείο $(-1, -1)$.



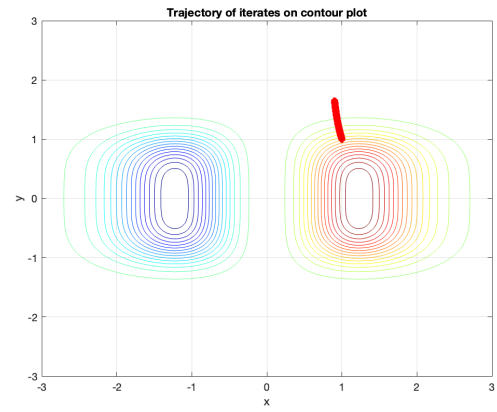
(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.

(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

Figure 4: Αποτελέσματα για σταθερό βήμα, αρχικό σημείο $(0, 0)$.



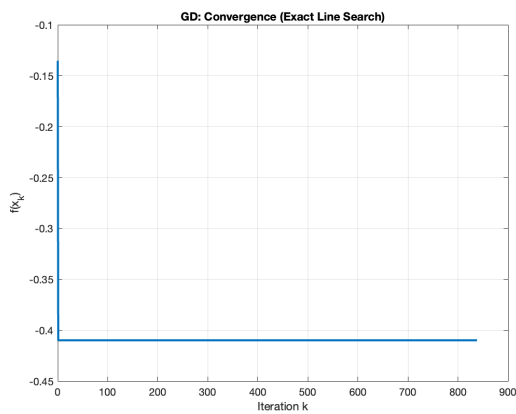
(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.



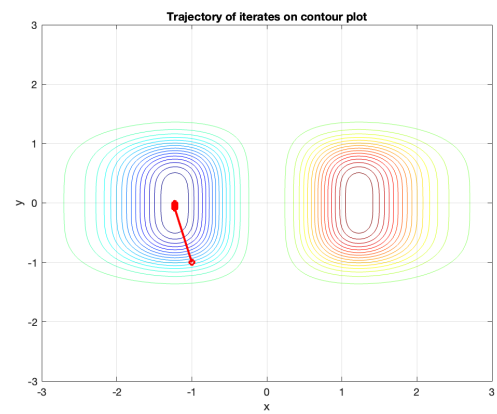
(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

Figure 5: Αποτελέσματα για σταθερό βήμα, αρχικό σημείο $(1, 1)$.

Μέθοδος Steepest Descent με exact line search

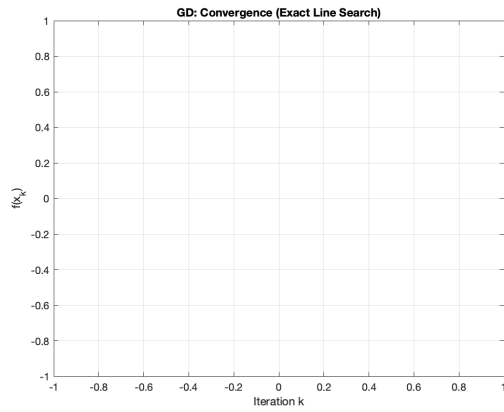


(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.

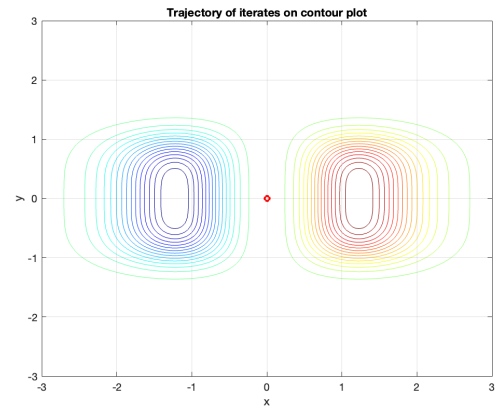


(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

Figure 6: Αποτελέσματα για exact line search, αρχικό σημείο $(-1, -1)$.

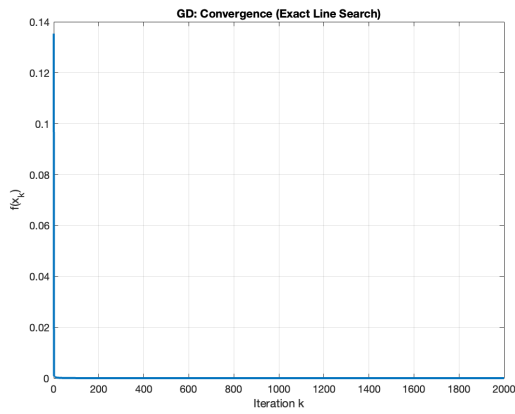


(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.

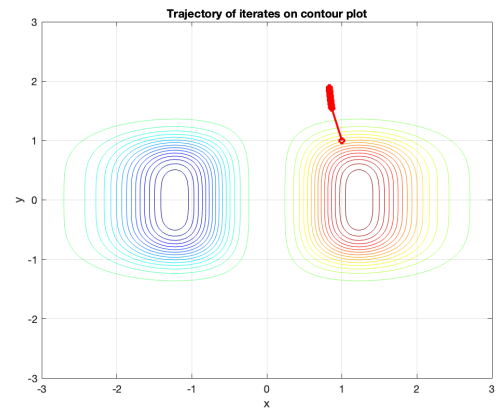


(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

Figure 7: Αποτελέσματα για exact line search, αρχικό σημείο $(0,0)$.



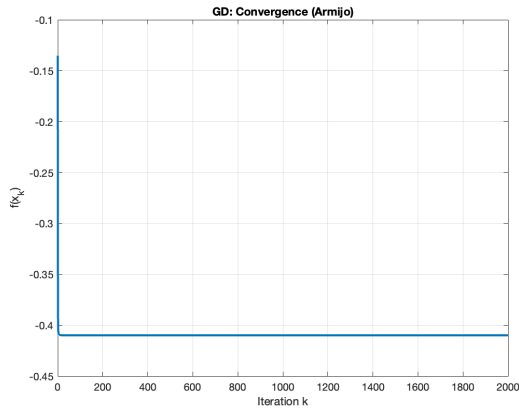
(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.



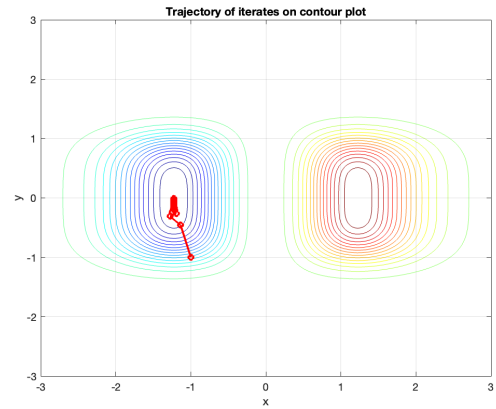
(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

Figure 8: Αποτελέσματα για exact line search, αρχικό σημείο $(1,1)$.

Μέθοδος Steepest Descent με Armijo rule

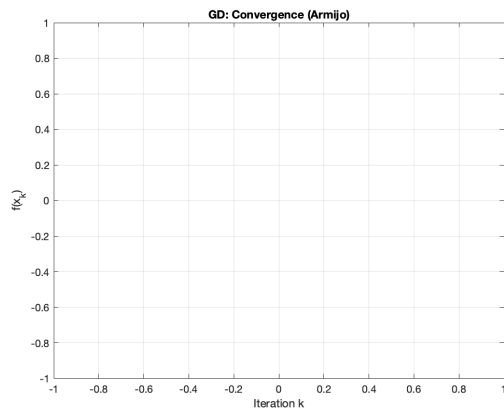


(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.

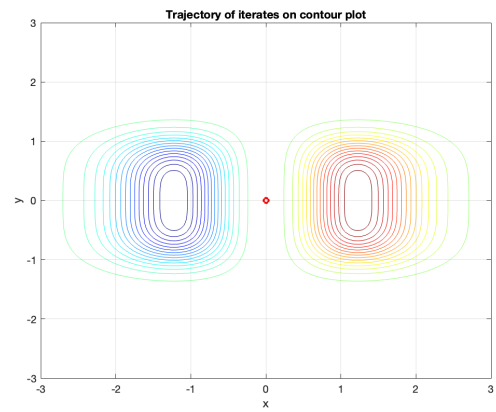


(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

Figure 9: Αποτελέσματα με Armijo rule, αρχικό σημείο $(-1, -1)$.

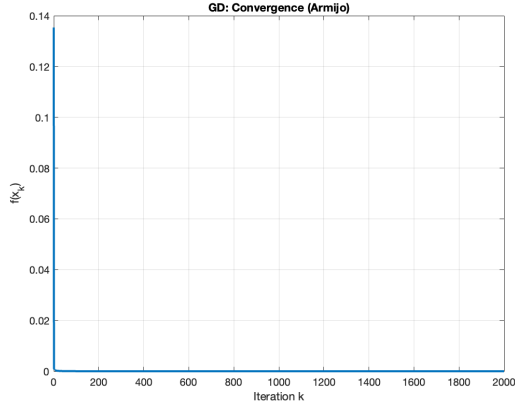


(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.

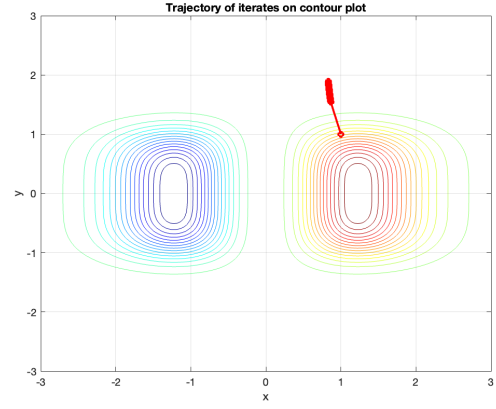


(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

Figure 10: Αποτελέσματα με Armijo rule, αρχικό σημείο $(0, 0)$.



(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.



(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

Figure 11: Αποτελέσματα με Armijo rule, αρχικό σημείο (1, 1).

3 Μέθοδος Newton

3) Σύνοψη

Ακολουθούμε πιστά τον αλγόριθμο 5.2.2 στη σελίδα 126. Υπολογίζουμε τον Εσσιανό πίνακα

$$\nabla^2 f(x, y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(x, y) & \frac{\partial^2 f}{\partial x \partial y}(x, y) \\ \frac{\partial^2 f}{\partial y \partial x}(x, y) & \frac{\partial^2 f}{\partial y^2}(x, y) \end{pmatrix}.$$

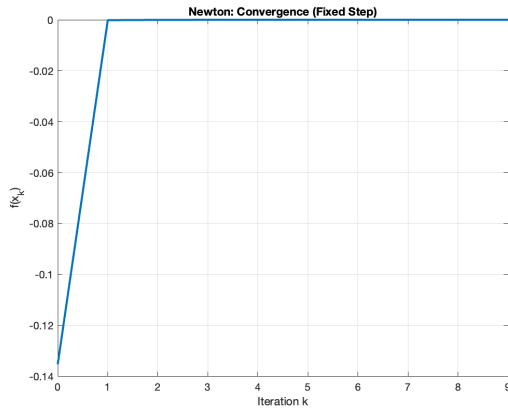
Με βάση τις υπολογισμένες παραγώγους, ο Hessian δίνεται από:

$$\nabla^2 f(x, y) = \begin{pmatrix} (6x - 14x^3 + 4x^5) e^{-(x^2+y^4)} & 4x^2 y^3 (2x^2 - 3) e^{-(x^2+y^4)} \\ 4x^2 y^3 (2x^2 - 3) e^{-(x^2+y^4)} & (-12x^3 y^2 + 16x^3 y^6) e^{-(x^2+y^4)} \end{pmatrix}.$$

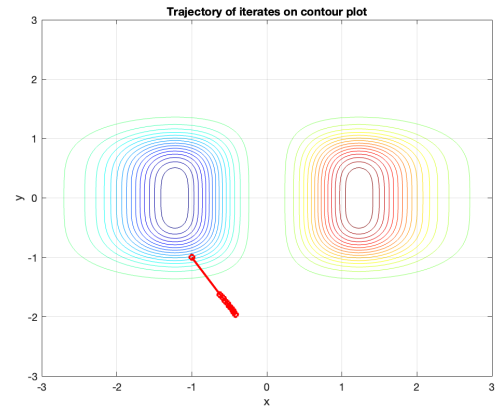
Αποτελέσματα Μεθόδου Newton

Για τα σημεία εκκίνησης $(-1, -1)$ $(0, 0)$ $(1, 1)$ έχουμε τα εξής αποτελέσματα για κάθε επιλογή βήματος.

Μέθοδος Newton με σταθερό βήμα

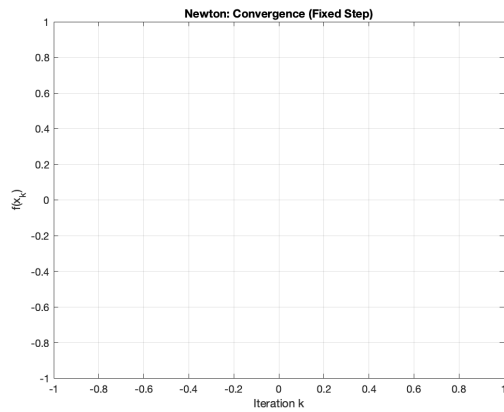


(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.

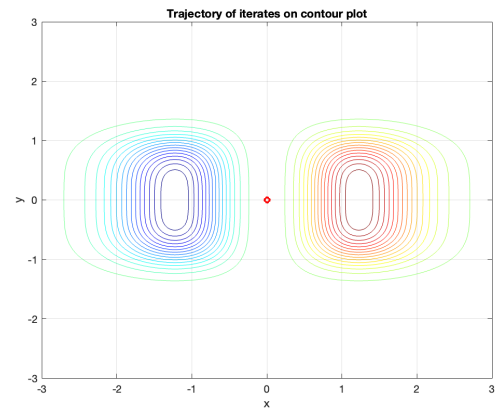


(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

Figure 12: Αποτελέσματα για σταθερό βήμα, αρχικό σημείο $(-1, -1)$.

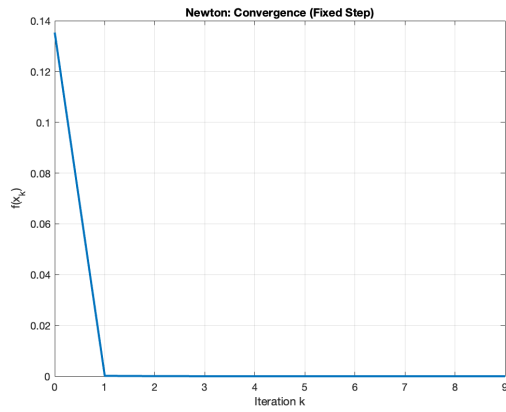


(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.

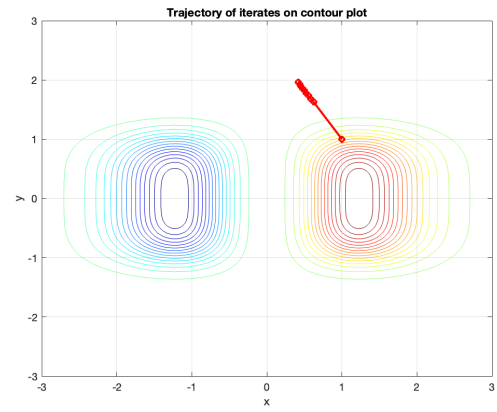


(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

Figure 13: Αποτελέσματα για σταθερό βήμα, αρχικό σημείο $(0, 0)$.



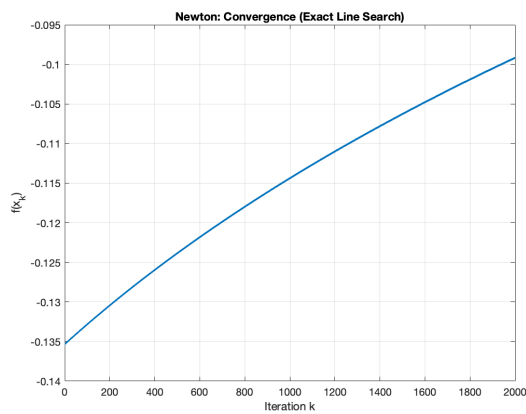
(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.



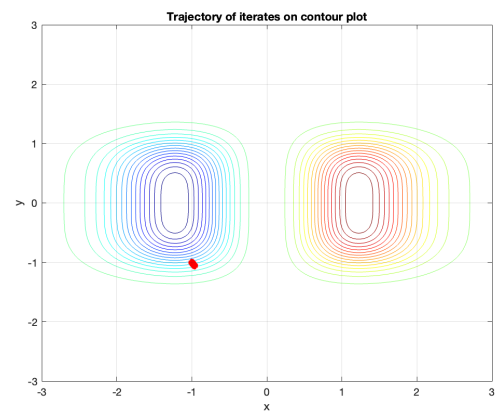
(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

Figure 14: Αποτελέσματα για σταθερό βήμα, αρχικό σημείο $(1, 1)$.

Μέθοδος Newton με exact line search

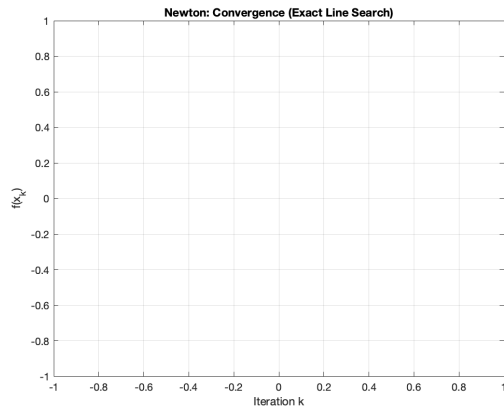


(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.

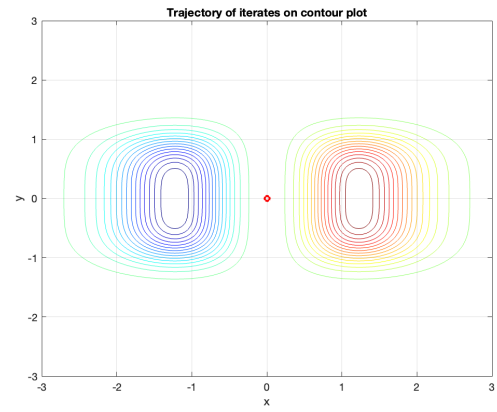


(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

Figure 15: Αποτελέσματα για exact line search, αρχικό σημείο $(-1, -1)$.

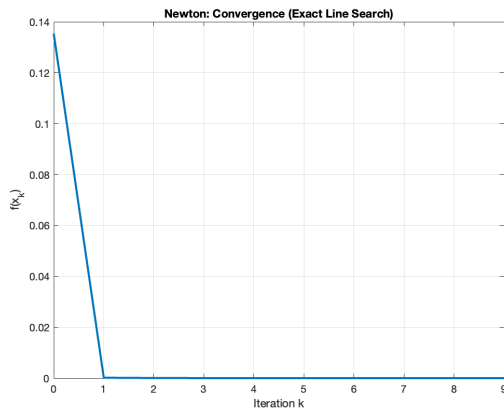


(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.

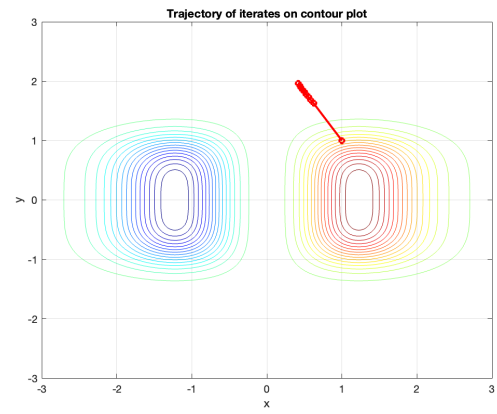


(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

Figure 16: Αποτελέσματα για exact line search, αρχικό σημείο (0,0).



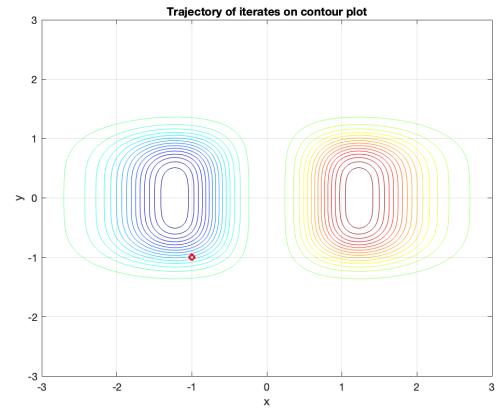
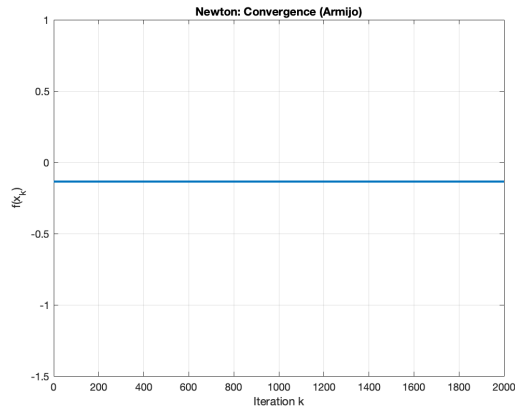
(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.



(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

Figure 17: Αποτελέσματα για exact line search, αρχικό σημείο (1,1).

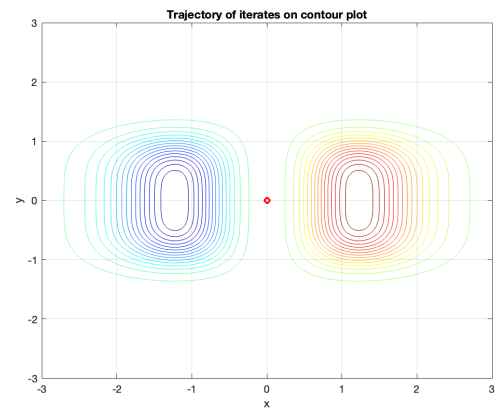
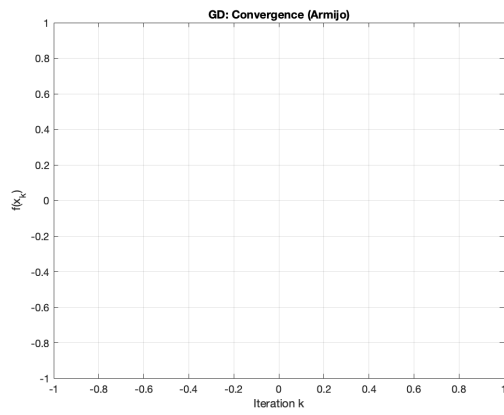
Μέθοδος Newton με Armijo rule



(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.

(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

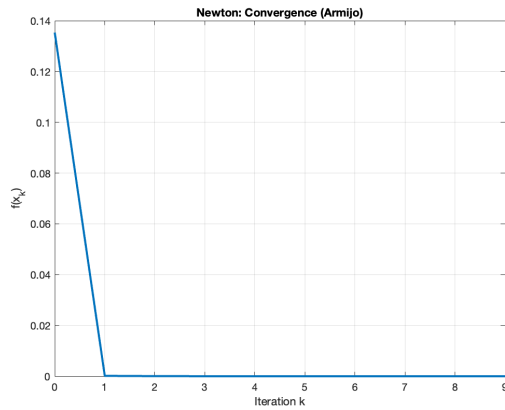
Figure 18: Αποτελέσματα με Armijo rule, αρχικό σημείο $(-1, -1)$.



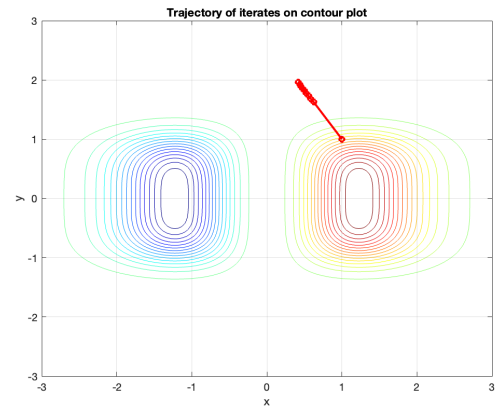
(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.

(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

Figure 19: Αποτελέσματα με Armijo rule, αρχικό σημείο $(0, 0)$.



(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.



(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

Figure 20: Αποτελέσματα με Armijo rule, αρχικό σημείο $(1, 1)$.

4 Μέθοδος Levenberg–Marquardt

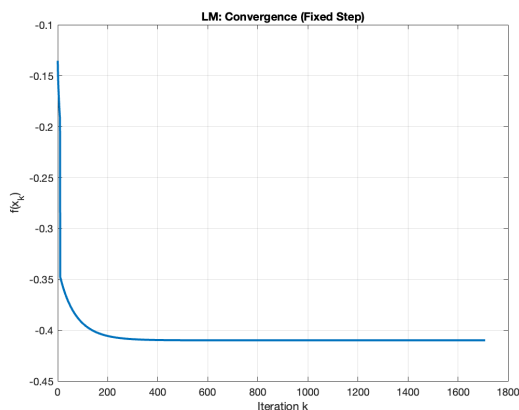
1) Σύνοψη

Αν ο Hessian $\nabla^2 f(x)$ δεν είναι θετικά ορισμένος για όλα τα $x \in \mathbb{R}^n$, τότε η μέθοδος Newton δεν μπορεί να εγγυηθεί την σύγκλιση. Για τον λόγο αυτό χρησιμοποιούμε τον αλγόριθμο Levenberg–Marquardt, ο οποίος εισάγει ένα διορθωτικό όρο ώστε να διασφαλίζει ένα τροποποιημένο, πάντα θετικά ορισμένο, Hessian. Πάλι ακολουθούμε πιστά τον αλγόριθμο 5.2.3 στη σελίδα 139.

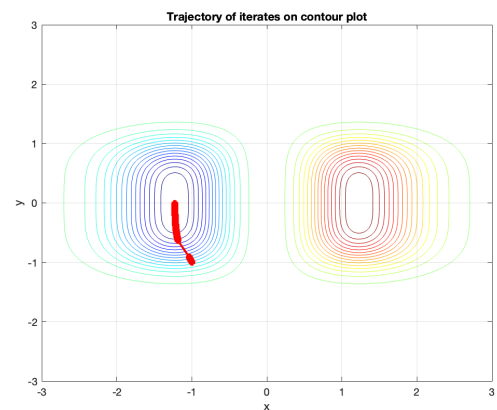
Αποτελέσματα Μεθόδου Levenberg–Marquardt

Για τα σημεία εκκίνησης $(-1, -1)$ $(0, 0)$ $(1, 1)$ έχουμε τα εξής αποτελέσματα για κάθε επιλογή βήματος.

Μέθοδος Levenberg–Marquardt με σταθερό βήμα

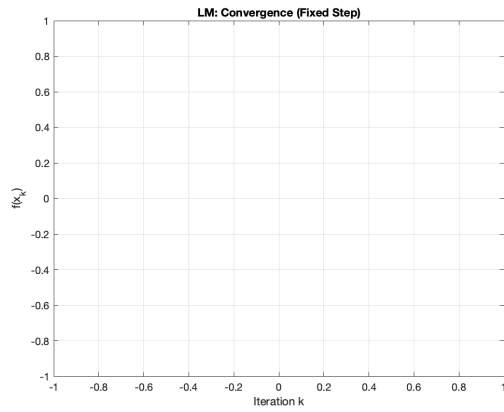


(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.

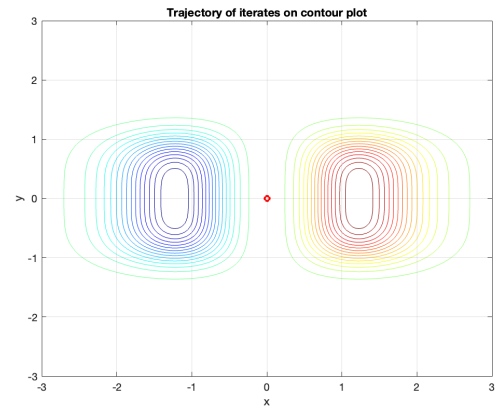


(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

Figure 21: Αποτελέσματα για σταθερό βήμα, αρχικό σημείο $(-1, -1)$.

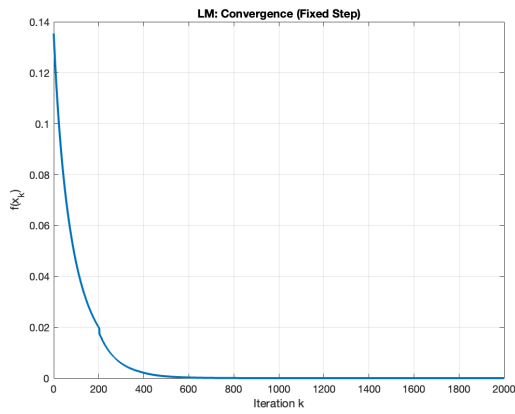


(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.

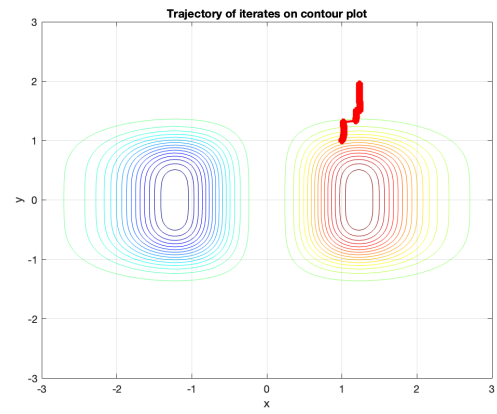


(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

Figure 22: Αποτελέσματα για σταθερό βήμα, αρχικό σημείο $(0, 0)$.



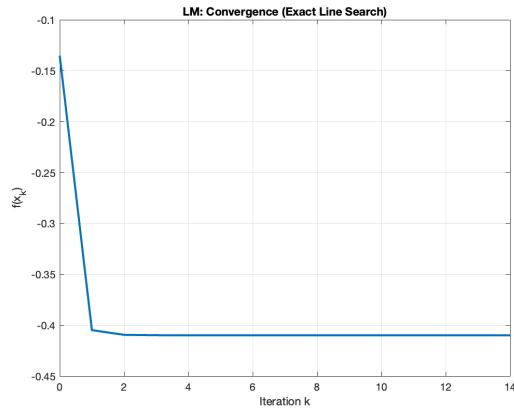
(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.



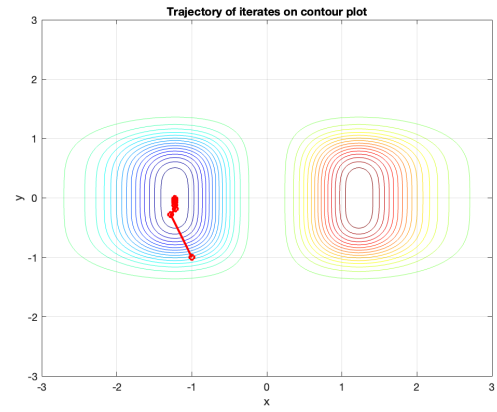
(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

Figure 23: Αποτελέσματα για σταθερό βήμα, αρχικό σημείο $(1, 1)$.

Μέθοδος Levenberg–Marquardt με exact line search

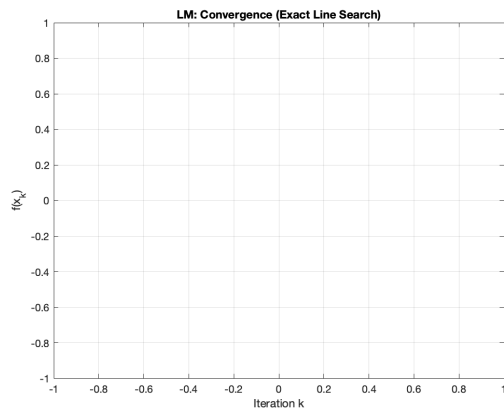


(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.

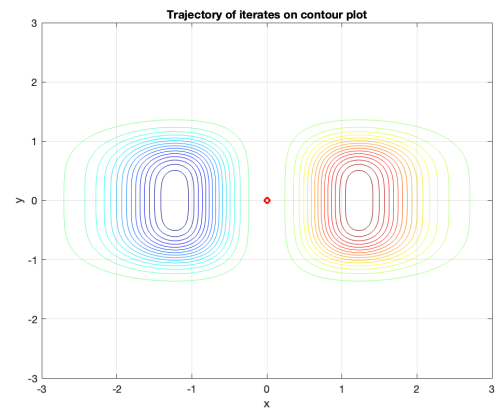


(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

Figure 24: Αποτελέσματα για exact line search, αρχικό σημείο $(-1, -1)$.

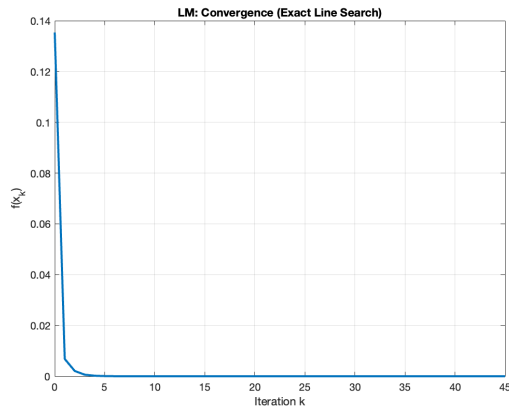


(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.

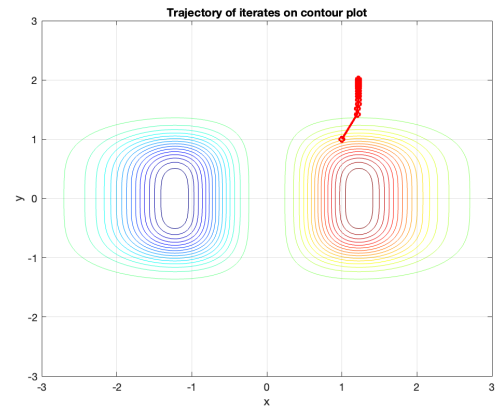


(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

Figure 25: Αποτελέσματα για exact line search, αρχικό σημείο $(0, 0)$.



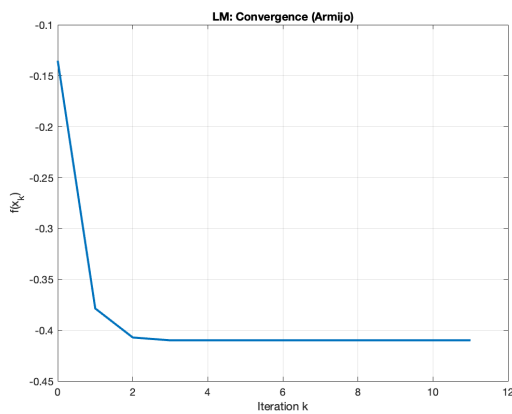
(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.



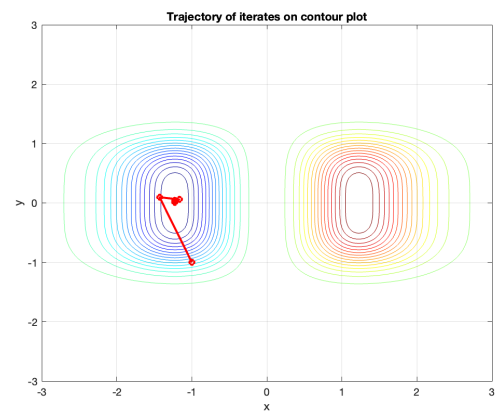
(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

Figure 26: Αποτελέσματα για exact line search, αρχικό σημείο $(1, 1)$.

Μέθοδος Levenberg–Marquardt με Armijo rule

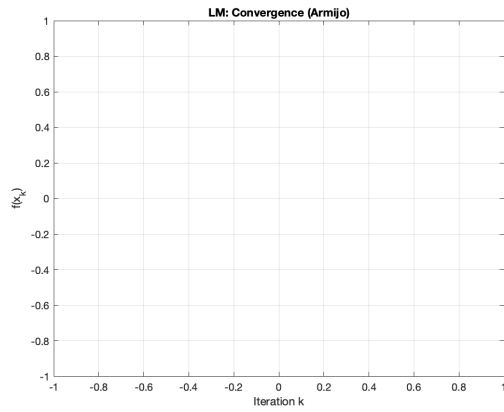


(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.

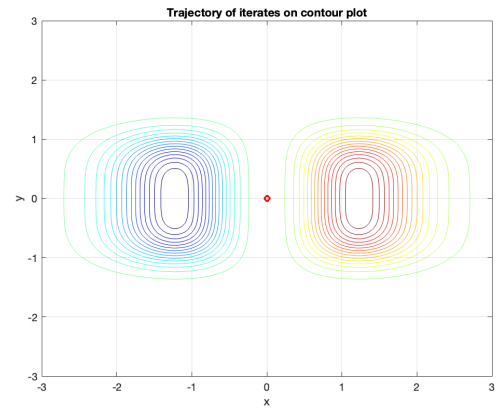


(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

Figure 27: Αποτελέσματα με Armijo rule, αρχικό σημείο $(-1, -1)$.

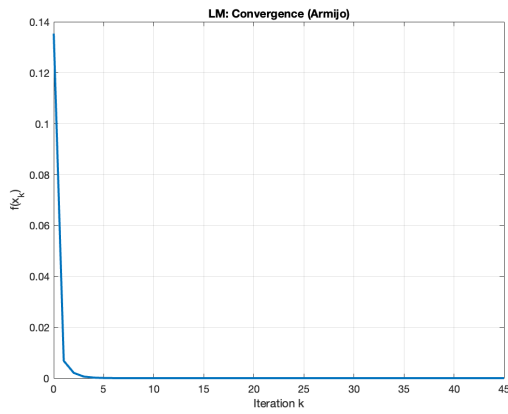


(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.

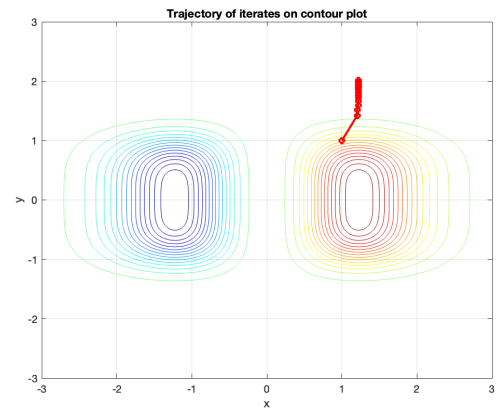


(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

Figure 28: Αποτελέσματα με Armijo rule, αρχικό σημείο $(0, 0)$.



(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.



(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

Figure 29: Αποτελέσματα με Armijo rule, αρχικό σημείο $(1, 1)$.

5 Παρατηρήσεις και Συμπεράσματα

Στην ενότητα αυτή συνοψίζουμε τα συμπεράσματα που προκύπτουν από την πειραματική μελέτη των μεθόδων Steepest Descent, Newton και Levenberg–Marquardt για τις τρεις επιλογές βήματος (stationary step, exact line search, Armijo rule).

Σταθερό βήμα (Fixed Step Rule)

Για σταθερό βήμα επιλέξαμε τις τιμές:

$$\alpha_{\text{GD}} = 0.01, \quad \alpha_{\text{Newton}} = 1, \quad \alpha_{\text{LM}} = 0.01.$$

Exact Line Search

Για την εύρεση του βέλτιστου βήματος

$$\gamma_k = \arg \min_{\gamma \geq 0} f(x_k + \gamma d_k),$$

χρησιμοποιήθηκε η συνάρτηση `fminbnd`, υλοποιώντας

$$\phi(\gamma) = f(x_k + \gamma d_k), \quad \gamma = \text{fminbnd}(\phi, 0, 1).$$

Κανόνας Armijo (Backtracking Line Search)

Ο κανόνας Armijo προσαρμόζει δυναμικά το βήμα, εξασφαλίζοντας τη συνθήκη: Στη μέθοδο αυτή το βήμα γ_k διατηρείται της μορφής

$$\gamma_k = s\beta^{m_k}, \quad 0 < \beta < 1$$

και m_k να είναι ο μικρότερος μη-αρνητικός ακέραιος που ικανοποιεί το Κριτήριο 4. Σύμφωνα με τη μέθοδο, το βήμα επιλέγεται ώστε:

$$f(x_k) - f(x_k + \gamma_k d_k) \geq -\alpha \beta^{m_k} d_k^T \nabla f(x_k).$$

Με άλλα λόγια, δοκιμάζουμε διαδοχικές τιμές του m_k με προεπιλεγμένα α, β και s . Στην εργασία αυτή χρησιμοποιήσαμε τις τιμές

$$\alpha = 10^{-4}, \quad \beta = 0.5, \quad s = 1.$$

Συμπεριφορά στα σημεία εκκίνησης.

1. **Σημείο εκκίνησης (0,0)** Το σημείο (0,0) είναι ιδιαίτερο, διότι

$$\nabla f(0,0) = 0, \quad \nabla^2 f(0,0) = 0_{2 \times 2},$$

άρα ούτε η Steepest Descent, ούτε η Newton, ούτε η Levenberg–Marquardt παράγουν μη μηδενική κατεύθυνση. Συνεπώς, ανεξάρτητα από την επιλογή βήματος (fixed, exact, Armijo),

$$x_{k+1} = x_k,$$

και ο αλγόριθμος μένει στάσιμος στο $(0, 0)$. Αυτό επιβεβαιώνεται από όλα τα contour plots της αναφοράς για $(0, 0)$, όπου το σημείο δεν μετακινείται από την αρχική θέση.

2. Σημείο εκκίνησης $(-1, -1)$ Οι μέθοδοι GD και LM συγκλίνουν στο ελάχιστο με όλες τις επιλογές βήματος:

$$(x^*, y^*) \approx (-1.22474, 0), \quad f(x^*, y^*) \approx -0.409916.$$

Η μέθοδος Newton αποκλίνει.

3. Σημείο εκκίνησης $(1, 1)$ Σύμφωνα με όλα τα αποτελέσματα στις αντίστοιχες σελίδες της αναφοράς, οι μέθοδοι δεν οδηγούνται στο ελάχιστο από αυτό το σημείο. Τα plots δείχνουν ότι τα βήματα φεύγουν προς περιοχές όπου η f αυξάνεται σε μέτρο και οι τιμές του $f(x_k)$ αποκλίνουν.

Σχόλιο για την μέθοδο Newton: Σύμφωνα με το Συμπέρασμα 5.2.1 του βιβλίου (σελ 128), για μη τετραγωνικές, όπως η

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

η μέθοδος Newton δεν είναι σχεδιασμένη ώστε να συγκλίνει σε πεπερασμένο αριθμό βημάτων και συχνά μπορεί να μη συγκλίνει καθόλου. Αυτό εξηγεί το ένα μέρος της συμπεριφορά που βλέπουμε πειραματικά. Επίσης, ο Εσσιανός πίνακας δεν είναι θετικά ορισμένος ούτε στο $(-1, -1)$ και $(1, 1)$ όπου με βάση το Συμπέρασμα 5.2.3 σελ 129 δίνει λύσεις μακριά από το σημείου ολικού ελαχίστου.

Επίδραση του αρχικού σημείου στη σύγκλιση: Από τα αποτελέσματα που παίρνουμε, παρατηρούμε ότι η επιλογή του αρχικού σημείου παίζει καθοριστικό ρόλο στην σύγκλιση.

Το σημείο $(-1, -1)$ βρίσκεται στην αριστερή μεριά της συνάρτησης δηλαδή στην μεριά της κοιλάδας που βρίσκεται το τοπικό ελάχιστο.

$$(x^*, y^*) \approx (-1.22474, 0), \quad f(x^*) \approx -0.409916.$$

Άρα ξεκινάμε ήδη από την σωστή μεριά και σιγα σιγά πέφτουμε μέσα στην κοιλάδα και τελικά στο ολικό ελάχιστο.

Αντίθετα, το σημείο $(1, 1)$ βρίσκεται στην αντίθετη περιοχή, δηλαδή στη περιοχή του μεγίστου. Τα αντίστοιχα plots δείχνουν ότι οι γραμμές επίπεδου γύρω από $(1, 1)$ «ωθούν» τις επαναλήψεις προς τα θετικά x , δηλαδή προς μια κατεύθυνση όπου η $f(x, y)$ αυξάνεται. Αυτό σημαίνει ότι οι κατευθύνσεις κατάβασης δεν οδηγούν προς την περιοχή του ελαχίστου. Το αποτέλεσμα είναι απόκλιση για όλες τις μεθόδους, ανεξάρτητα από το είδος της επιλογής βήματος.

Συνολικά, η μορφή της συνάρτησης $f(x, y)$ παρουσιάζει δύο διαφορετικές «πλευρές»: μία για αρνητικές τιμές του x , η οποία οδηγεί φυσικά προς το ελάχιστο, και μία για θετικές τιμές, όπου οι επαναλήψεις τείνουν να απομακρύνονται. Έτσι εξηγείται γιατί το $(-1, -1)$ συγκλίνει με όλες τις μεθόδους, ενώ το $(1, 1)$ οδηγεί σε απόκλιση.

Σύγκριση των τριών μεθόδων

Μελετώντας τα διαγράμματα σύγκλισης και τις τροχιές των επαναλήψεων, προκύπτουν τα ακόλουθα συμπεράσματα για τη συμπεριφορά των τριών μεθόδων.

Steepest Descent (GD). Η μέθοδος GD παρουσιάζει τη μεγαλύτερη σταθερότητα αλλά την μικρότερη ταχύτητα σύγκλισης. Από το σημείο $(-1, -1)$ συγκλίνει στο τοπικό ελάχιστο

$$(x^*, y^*) \approx (-1.22474, 0),$$

ενώ από το $(1, 1)$ αποκλίνει, αφού το gradient εκεί «ωθεί» τη μέθοδο εκτός της λεκάνης έλξης του ελαχίστου. Η χρήση exact line search ή Armijo βελτιώνει την σταθερότητα αλλά όχι την ταχύτητα της μεθόδου.

Newton. Η μέθοδος Newton θεωρητικά έχει την μεγαλύτερη ταχύτητα σύγκλισης, αλλά στην παρούσα συνάρτηση η συμπεριφορά της είναι ασταθής. Όπως είδαμε στα αντίστοιχα διαγράμματα της αναφοράς, από το σημείο εκκίνησης $(-1, -1)$ η μέθοδος αποκλίνει και δεν κατευθύνεται προς το τοπικό ελάχιστο, ενώ από το $(1, 1)$ αποκλίνει ακόμα περισσότερο. Αυτό οφείλεται στο ότι ο Hessian της συνάρτησης δεν είναι θετικά ορισμένος σε μεγάλο μέρος του χώρου, οπότε η Newton δεν δείχνει πάντα στην κατεύθυνση μείωσης της f , δηλαδή του ελαχίστου. Επιπλέον, σύμφωνα με το Συμπέρασμα 5.2.1 του βιβλίου, για μη τετραγωνικές συναρτήσεις όπως η $f(x, y) = x^3 e^{-(x^2+y^4)}$ η μέθοδος Newton δεν εγγυάται σύγκλιση και μπορεί να οδηγηθεί σε τιμές μακριά από τη λεκάνη έλξης του ελαχίστου. Έτσι, παρά τη θεωρητική της ταχύτητα, η μέθοδος σε αυτή τη συνάρτηση και με αυτές τις αρχικές τιμές δεν είναι αξιόπιστη.

Levenberg–Marquardt (LM). Η μέθοδος LM αποτελεί ενδιάμεση λύση μεταξύ Newton και GD. Παρουσιάζει πολύ καλή σταθερότητα, ακόμη και όταν ο Hessian δεν είναι θετικά ορισμένος, και συγκλίνει ομαλά από το $(-1, -1)$. Από το $(1, 1)$ δεν οδηγείται στο ελάχιστο, αλλά αποκλίνει πιο αργά και πιο ελεγχόμενα σε σχέση με τις άλλες μεθόδους. Από το $(0, 0)$ παραμένει στάσιμη για τον ίδιο λόγο όπως και οι άλλες μέθοδοι.

Εγκλωβισμός σε τοπικά ακρότατα. Η συνάρτηση $f(x, y) = x^3 e^{-(x^2+y^4)}$ έχει με μία «λεκάνη έλξης» που βρίσκεται στην περιοχή $x < 0$. Έτσι, οι μέθοδοι βελτιστοποίησης οδηγούνται στο μοναδικό τοπικό ελάχιστο μόνο όταν το αρχικό σημείο βρίσκεται εντός αυτής της περιοχής. Αντίθετα, για αρχικές τιμές με $x > 0$, οι μέθοδοι αποκλίνουν, διότι οι κατευθύνσεις κατάβασης δεν οδηγούν προς το ελάχιστο αλλά προς περιοχές όπου η f αυξάνεται σε μέτρο. Συγκεκριμένα, βλέπουμε ότι στη μέθοδο Steepest Descent στο $(1, 1)$ με exact line search και armijo έχουμε μάλλον εγκλωβισμό σε τοπικό ακρότατο αφού αμέσως πηγαίνει η συνάρτηση εκεί και σταματάμε το ψάξιμο. Τα ίδια μπορούμε μάλλον να πούμε και για τη μέθοδο Newton από το $(-1, -1)$ και $(1,$

1) με σταθερό βήμα και για το (1, 1) με exact line search και armijo. Αντίστοιχα συμπεράσματα βγάζουμε και για την μέθοδο Levenberg–Marquardt.

Ταχύτητα σύγκλισης των μεθόδων Από τα διαγράμματα σύγκλισης και τις τροχιές των επαναλήψεων παρατηρούμε ότι η μέθοδος Steepest Descent παρουσιάζει τη μικρότερη ταχύτητα σύγκλισης. Η μέθοδος Newton εμφανίζει θεωρητικά πολύ ταχύτερη σύγκλιση, εδώ όμως αποκλίνει για τα περισσότερα αρχικά σημεία λόγω του ότι ο Hessian δεν είναι θετικά ορισμένος. Η μέθοδος Levenberg–Marquardt πετυχαίνει το ενδιάμεσο, δηλαδή είναι σημαντικά πιο σταθερή από τη Newton και ταυτόχρονα συγκλίνει ταχύτερα από τη Steepest Descent.

6 Αρχιτεκτονική και Οργάνωση Κώδικα

Η δομή του κώδικα φαίνεται παρακάτω:

PROJECT_2/

functions/

f_xy.m

grad_f_xy.m

hessian_f_xy.m

methods/

Steepest_descent.m

Newton.m

Levenberg_Marquarddt.m

steps/

step_fixed.m

step_exact_line_search.m

step_armijo.m

plots/

plot_f.m

plot_contours.m

plot_convergence.m

plot_trajectory.m

main.m

Περιγραφή Φακέλων

functions/ Περιέχει τις βασικές μαθηματικές συναρτήσεις: την αντικειμενική $f(x, y)$, το gradient και τον Hessian πίνακα.

methods/ Περιλαμβάνει τις υλοποιήσεις των μεθόδων βελτιστοποίησης: Steepest Descent, Newton και Levenberg–Marquardt. Κάθε μέθοδος έχει διακριτό αρχείο για λόγους καθαρότητας και modularity.

steps/ Περιέχει τους διαφορετικούς κανόνες επιλογής του μήκους βήματος (step size) δηλαδή fixed step, exact line search και Armijo backtracking. Οι μέθοδοι (στον φάκελο **methods/**) δέχονται έναν *step rule* ως συνάρτηση, με αυτό τον τρόπο δοκιμάζουμε τα διαφορετικά είδους γ_k .

plots/ Συγκεντρώνει όλες τις βοηθητικές ρουτίνες για παραγωγή γραφημάτων: επίπεδο καμπυλών, 3D απεικόνιση της συνάρτησης, διαγράμματα σύγκλισης και απεικόνιση της τροχιάς των επαναλήψεων.

main.m Το κεντρικό αρχείο εκτέλεσης. Ρυθμίζει τις παραμέτρους, επιλέγει αρχικές τιμές, καλεί τις μεθόδους βελτιστοποίησης και παράγει τα αντίστοιχα γραφήματα και αποτελέσματα.