ΤΕΧΝΙΚΕΣ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ

Νικόλαος Γιακουμόγλου 9043

28 Νοεμβρίου 2020

1 Εισαγωγή

Στόχος της παρούσας εργασίας είναι η ελαχιστοποίηση 1 συνάρτήσεις με 3 διαφορετικές μεθόδους:

- Μέθοδος Steepest Descent → function **SteepestDescent.m**
- Μέθοδος Newton → function **Newton.m**
- Μέθοδος Levenberg-Marqardt → functionpt **LevenbergMarqardt.m**

Σημειώνεται πως οι αλγόριμοι υλοποιήθηκαν βάσει αναζήτησης στο διαδίκτυο και της παράδοσης στο μάθημα και όχι αναλυτική βάσει το βιβλίο (όπως αναφέρεται στην εκφώνηση) καθώς το τελευταίο δεν έχει παραδοθεί ακόμα.

2 Σύντομη Επεξήγηση Αλγορίθμων

Κάθε αλγόριθμος είναι της μορφής

$$x(k+1) = x(k) - \gamma(k) \cdot \Delta(k) \cdot \nabla f\left(x(k)\right)$$

Η διαφορά τους έγγυται στην επιλογή του $\Delta(k)$ όπου $\Delta(k) = \Delta(k)^T > 0$. Επίσης ορίζουμε την κατέθυνση ως

$$d(k) = -\Delta(k) \cdot \nabla f\left(x(k)\right)$$

και τελικά

$$x(k+1) = x(k) + \gamma(k) \cdot d(k)$$

2.1 Μέθοδος Steepest Descent

Στην μέθοδο αυτή έχουμε

$$\Delta(k) = I \Rightarrow d(k) = -\nabla f\left(x(k)\right)$$

Άρα

$$x(k+1) = x(k) - \gamma(k) \cdot \nabla f\left(x(k)\right)$$

2.2 Μέθοδος Newton

Στην μέθοδο αυτή έχουμε

$$\Delta(k) = \left[\nabla^2 f\left(x(k)\right)\right]^{-1} \Rightarrow d(k) = -\left[\nabla^2 f\left(x(k)\right)\right]^{-1} \cdot \nabla f\left(x(k)\right)$$

Άρα

$$x(k+1) = x(k) - \gamma(k) \cdot \left[\nabla^2 f\left(x(k)\right)\right]^{-1} \cdot \nabla f\left(x(k)\right)$$

2.3 Μέθοδος Levenberg-Marqardt

Η μέθοδος αυτή είναι εξέλιξη της μεθόδου του Newton. Στην μέθοδο αυτή έχουμε

$$\Delta(k) = \left[\nabla^2 f\left(x(k)\right) + \mu(k) \cdot I\right]^{-1} \Rightarrow d(k) = -\left[\nabla^2 f\left(x(k)\right) + \mu(k) \cdot I\right]^{-1} \cdot \nabla f\left(x(k)\right)$$

όπου $\mu(k) > \bar{\mu}(k)$, $\bar{\mu}(k)$ η μεγαλύτερη κατά απόλυτη τιμή του Εσσιανού. Με αυτόν τον τρόπο εξασφαλίζουμε ότι $\Delta(k) > 0$.

Άρα

$$x(k+1) = x(k) - \gamma(k) \cdot \left[\nabla^2 f\left(x(k)\right) + \mu(k) \cdot I\right]^{-1} \cdot \nabla f\left(x(k)\right)$$

Σημειώνεται ότι

- Αν $\nabla^2 f\left(x(k)\right) \ll \mu(k) \cdot I$ τότε η μέθοδος μοιάζει με την μέθοδο Steepest Descent.
- Αν $\nabla^2 f\left(x(k)\right) \gg \mu(k) \cdot I$ τότε η μέθοδος μοιάζει με την μέθοδο Newton.

2.4 Αλγόριθμοι επιλογής γ

Ο τρόπος επιλογής του $\gamma(k)$ μπορεί να προκύψει σύμφωνα με 3 τρόπους:

1. $\gamma(k) = \gamma = \sigma\tau\alpha\theta\varepsilon\rho\acute{o}\ \forall k$ όπου για να συγκλίνει ο αλγόριθμος πρέπει να επιλέξουμε κατάλληλα το γ έτσι ώστε $\left|\frac{x(k+1)}{x(k)}\right| < 1$
2. $\gamma(k) = argmin\, h(\gamma) = f\left(x(k) - \gamma \cdot \nabla f\left(x(k)\right)\right)$
3. Σύμφωνα με τον κανόνα Armijo $\gamma(k) = s \cdot \beta^{m(k)}, 0 < \beta < 1$ και $m(k)$ ο μικρότερος μη αρνητικός ακέραιος που ικανοποιεί το κριτήριο 4. Για το α επιλέγεται το διάστημα $[10^{-5}, 10^{-1}]$ ενώ για το β το διάστημα $[10^{-1}, 1/2]$. Το s τίθεται αυθαίρετα ίσο με 0.1. Για τον κανόνα ισχύει:

$$f\left(x(k+1)\right) \leq f\left(x(k)\right) + a \cdot \gamma(k) \cdot d(k)^T \cdot \nabla f\left(x(k)\right), 0 < a < b < 1 \Rightarrow$$

$$\Rightarrow f\left(x(k) + sb^{m(k)}\right) \leq f\left(x(k)\right) + a \cdot \gamma(k) \cdot d(k)^T \cdot \nabla f\left(x(k)\right)$$

2.5 Κριτήριο τερματισμού

Κριτήριο τερματισμού αποτελεί αν το μέτρο της κλίσης $\nabla f(x,y) < \varepsilon$ όπου ε η ακρίβεια ή αν υπερβούμε έναν ανώτερο αριθμό επαναλήψεων.

2.6 Κώδικας

Η κλήση των συναρτώσεων είναι οι εξής:

```
function [x, y, F] = SteepestDescent(f, e, x0, y0, ALGORITHM, g, KMAX)
function [x, y, F] = Newton(f, e, x0, y0, ALGORITHM, g, KMAX)
function [x, y, F] = LevenbergMarquardt(f, e, x0, y0, ALGORITHM, g, KMAX)
```

όπου η είσοδος ALGORITHM παίρνει τιμες $1, 2$ ή 3 και εφαρμόζει τους αντίστοιχους τρόπους επιλογής του γ όπως αναφέρθηκε παραπάνω. Στην περίπτωση που ο αλγόριθμος επιλεγεί ο 2 ή ο 3, το γ δεν λαμβάνεται υπ'όψη στην είσοδο.

3 Βοηθητικές Συναρτήσεις σε MATLAB

Για την εποπτική επικύρωση των αποτελεσμάτων, χρησιμοποιήθηκε το **wolfram**. Τα αποτελέσματα είναι ότι η $f(x, y) = x^3 \cdot e^{-x^2-y^4}$ παρουσιάζει τοπικό ελάχιστο στο $(x, y) = (1.22, 0)$ και τοπικό μέγιστο στο $(x, y) = (-1.22, 0)$.
Για τον σκοπό των προσομοιώσεων φτιάχτηκε η συνάρτηση **PlotStem3.m** που εμφανίζει στο την τρισδιάστατη συνάρτηση στον χώρο με τα σημεία που βρήκε ο αλγόριθμος σε κάθε επανάληψη. Επίσης για να υπάρχει καλύτερη επισκόπηση των αποτελεσμάτων φτιάχτηκε και η συνάρτηση **PlotContour.m** που εμφανίζει το contour plot με τα σημεία που βρήκε ο αλγόριθμος σε κάθε επανάληψη δείχνοντας το τελικό και το αρχικό σημείο με βέλη. Για τα βέλη βρέθηκε η συνάρτηση **Arrow.m** απο το math-exchange και χρησιμοποιήθηκε αυτούσια. Τέλος για την εύρεση του γ από τον κανόνα του Armijo (ALGORITHM=3), δημιουργήθηκε ξεχωριστή συνάρτηση **Armijo.m** ενώ για την ελαχιστοποίηση (ALGORITHM=2) χρησιμοποιήθηκε η συνάρτηση από το προηγούμενο παραδοτέο αυτούσια μετονομασμένη σε **Golden.m**.

4 Συνάρτηση $f(x, y) = x^3 \cdot e^{-x^2-y^4}$ (ΘΕΜΑ 1)

4.1 Εκτύπωση συνάρτησης $f(x, y)$

Η προσομοιώσεις βρίσκονται στο αρχείο **thema1.m**. Τα figures αποθηκεύονται αυτόματα στον τωρινό φάκελο. Εκτυπώνουμε την δοσμένη συνάρτηση με την χρήση της *fsurf*. Σημειώνεται ότι stem plots και contour plots της f θα δούμε στην συνέχεια. Το αποτέλεσμα είναι το εξής

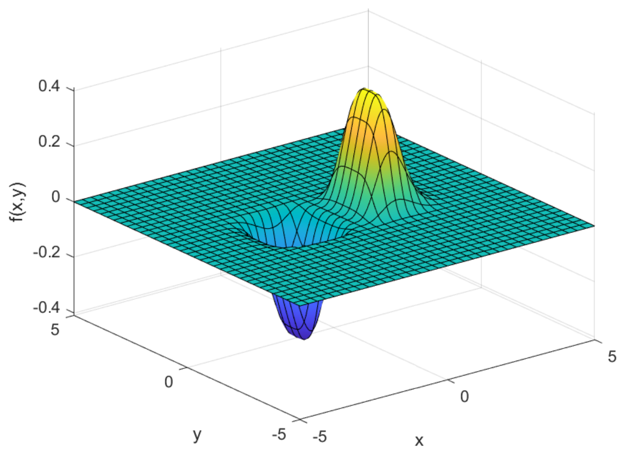

Figure 1: $f(x, y) = x^3 \cdot e^{-x^2-y^4}$

4.2 Υπολογισμός $\nabla f(x, y)$ και $\nabla^2 f(x, y)$

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

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

$$\nabla^2 f(x, y) = e^{-x^2-y^4} \cdot \begin{bmatrix} 6x + 4x^5 - 14x^3 & 8x^4y^3 - 12x^2y^3 \\ 8x^4y^3 - 12x^2y^3 & 16x^3y^6 - 12x^3y^2 \end{bmatrix}$$

5 Μέθοδος Steepest Descent (ΘΕΜΑ 2)

Η προσομοιώσεις βρίσκονται στο αρχείο **thema2.m**. Τα figures αποθηκεύονται αυτόματα στον φάκελο *PlotsSteepestDescent*.

5.1 Σημείο 1 $(x, y) = (0, 0)$

Επειδή ισχύει ότι $|\nabla f(0, 0)| = 0$ ο αλγόριθμος θα τερματίζει απευθείας ανεξαρτήτου μεθόδου επιλογής γ. Οπότε έχουμε κοινά αποτελέσματα τα εξής:

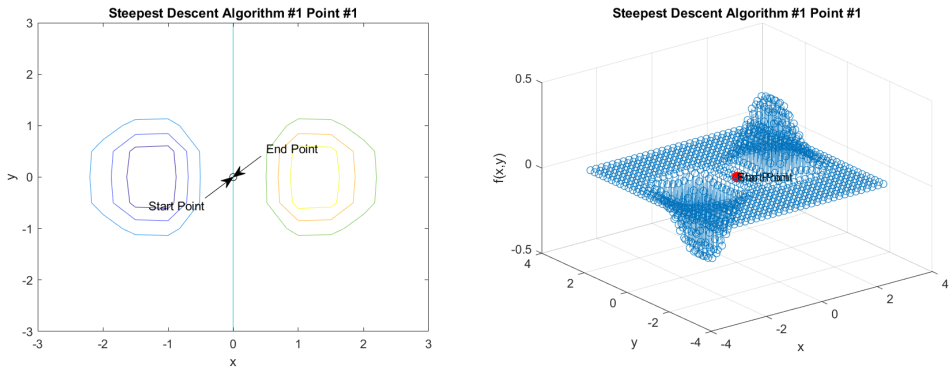

Figure 2: Σημείο 1 Αλγόριθμοι 1 έως 3

5.2 Σημείο 2 $(x, y) = (-1, -1)$

Καταρχάς, για να βρούμε σταθερό γ, τρέχουμε τον αλγόριθμο για διάφορα γ από 0.1 έως 2 περιορίζοντας την ακρίβεια για ταχύτερα αποτελέσματα. Επίσης θέτουμε τον μέγιστο αριθμό επαναλήψεων στις 1000 που σημαίνει ότι για 1000 επαναλήψεις ο αλγόριθμος μάλλον δημιουργεί ταλαντώσεις (εμφανίζονται για $\gamma \approx 1.3$). Επιλέγουμε $\gamma = 1.1$. Σημειώνεται πως αύτο μπορεί να γίνει θεωρητικά έτσι ώστε $\left|\frac{x(k+1)}{x(k)}\right| < 1$ αλλά εμείς περιοριζόμαστε σε πειραματικά αποτελέσματα. Για αυτόν τον λόγο παραθέτουμε το γράφημα των αριθμό των επαναλήψεων για διάφορα γ. Ο αλγόριθμος συγκλίνει. Αν επιλέξουμε σταθερό βήμα συγκλίνει σε 176 επαναλήψεις, αν ελαχιστοποιήσουμε την συνάρτηση $h(\gamma) = f\left(x(k) - \gamma \cdot \nabla f\left(x(k)\right)\right)$ με την βοήθεια της μεθόδου της Χρυσής τομής συγκλίνει σε 3061 επαναλήψεις και σε περισσότερο χρόνο και για τον κανόνα του Armijo συγκλίνει σε 2012 επαναλήψεις. Παραθέτουμε τα αποτελέσματα για κάθε αλγόριθμο (ALGORITHM=1,2,3):

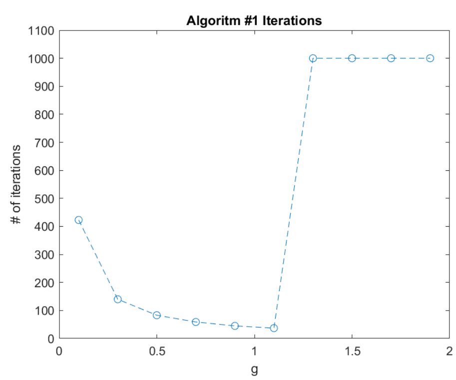

Figure 3: Σημείο 2 Αλγόριθμος 1 Επαναλήψεις

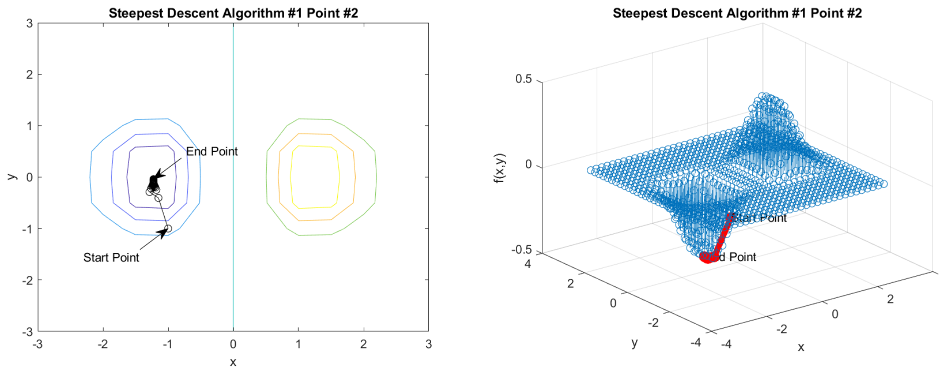

Figure 4: Σημείο 2 Αλγόριθμος 1

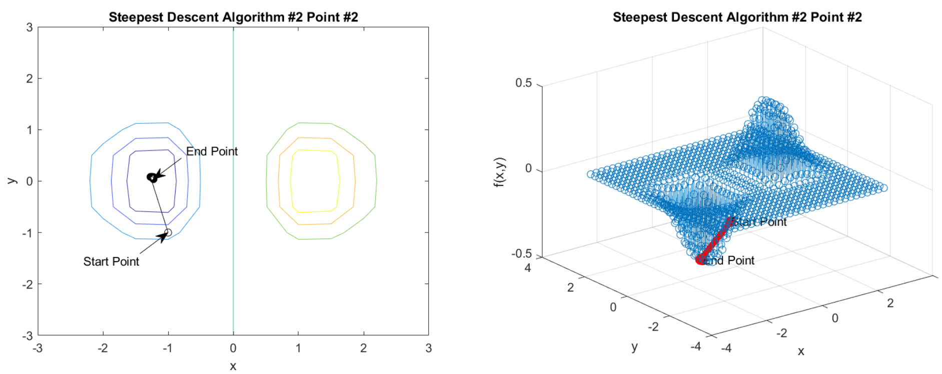

Figure 5: Σημείο 2 Αλγόριθμος 2

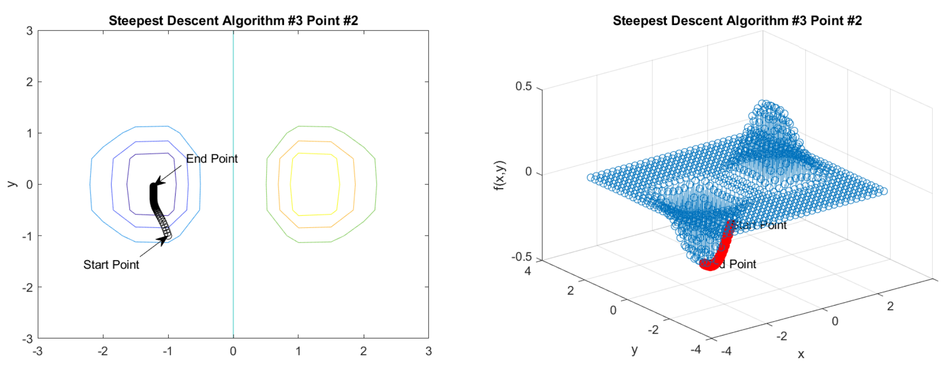

Figure 6: Σημείο 2 Αλγόριθμος 3

5.3 Σημείο 3 $(x, y) = (1, 1)$

Καταρχάς, για να βρούμε σταθερό γ, τρέχουμε τον αλγόριθμο για διάφορα γ από 0.1 έως 2 περιορίζοντας την ακρίβεια για ταχύτερα αποτελέσματα. Επίσης θέτουμε τον μέγιστο αριθμό επαναλήψεων στις 1000 που σημαίνει ότι για 1000 επαναλήψεις ο αλγόριθμος μάλλον δημιουργεί ταλαντώσεις. Επιλέγουμε $\gamma = 0.1$. Σημειώνεται πως αύτο μπορεί να γίνει θεωρητικά έτσι ώστε $\left|\frac{x(k+1)}{x(k)}\right| < 1$ αλλά εμείς περιοριζόμαστε σε πειραματικά αποτελέσματα. Για αυτόν τον λόγο παραθέτουμε το γράφημα των αριθμό των επαναλήψεων για διάφορα γ. Ωστόσο επειδή ο αλγόριθμος αποκλίνει, μεγαλύτερο γ οδηγεί σε πιο γρήγορη απόκλιση (η ταχύτερη σύγλιση για μεγαλύτερο γ είναι ψευδαίσθηση). Ο αλγόριθμος αποκλίνει και σταματάει σε 4774 επαναλήψεις για σταθερό γ, σε 2 επαναλήψεις αν ελαχιστοποιήσουμε την $h(\gamma) = f\left(x(k) - \gamma \cdot \nabla f\left(x(k)\right)\right)$ ή 4885 για τον κανόνα Armijo. Το ίδιο ισχύει και για μικρά βήματα, όταν το βήμα είναι σταθερό, καθώς παγιδεύεται σε περιοχή μικρών κλίσεων. Παραθέτουμε τα αποτελέσματα για κάθε αλγόριθμο (ALGORITHM=1,2,3):

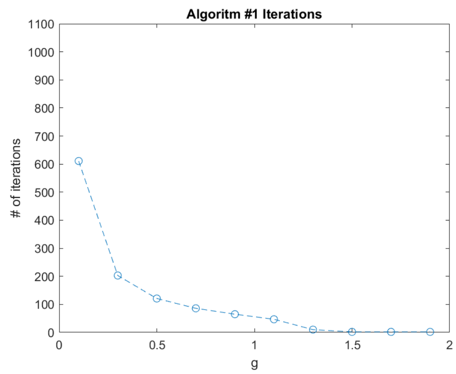

Figure 7: Σημείο 3 Αλγόριθμος 1 Επαναλήψεις

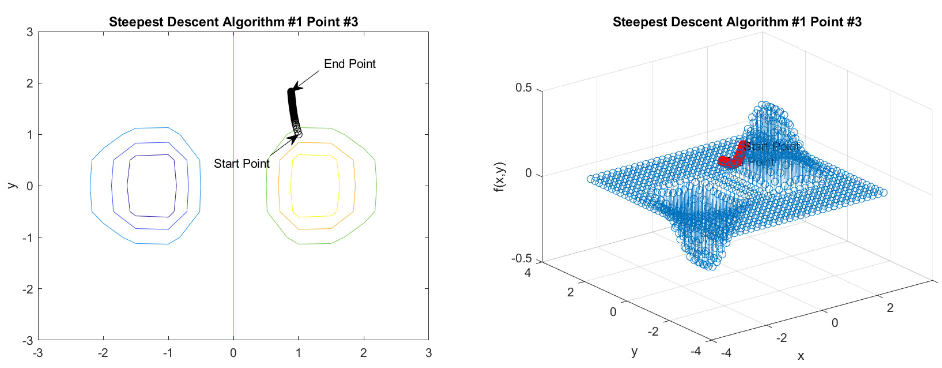

Figure 8: Σημείο 3 Αλγόριθμος 1

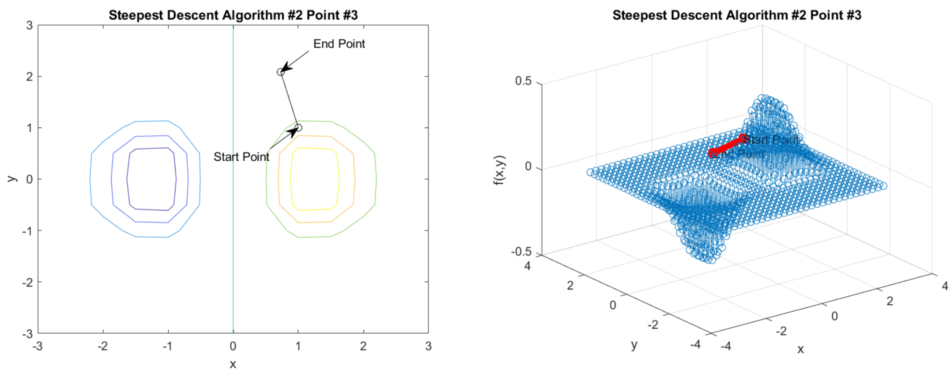

Figure 9: Σημείο 3 Αλγόριθμος 2

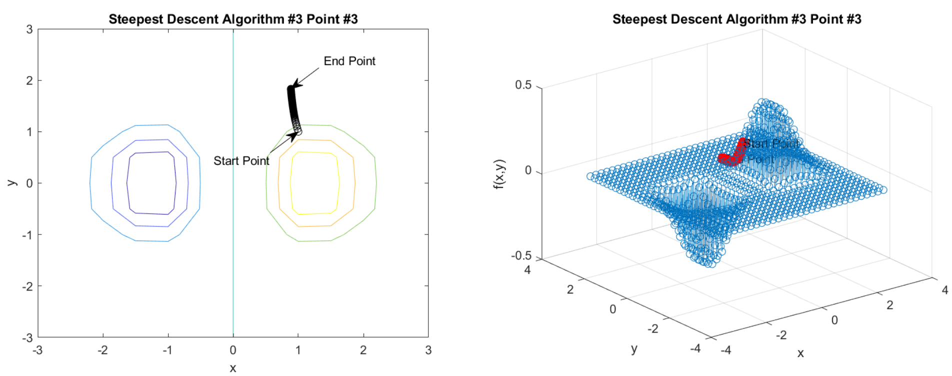

Figure 10: Σημείο 3 Αλγόριθμος 3

6 Μέθοδος Newton (ΘΕΜΑ 3)

Η προσομοιώσεις βρίσκονται στο αρχείο **thema3.m**. Τα figures αποθηκεύονται αυτόματα στον φάκελο *PlotsNewton*.

6.1 Σημείο 1 $(x, y) = (0, 0)$

Επειδή ισχύει ότι $|\nabla f(0,0)| = 0$ ο αλγόριθμος θα τερματίζει απευθείας ανεξαρτήτου μεθόδου επιλογής γ (ALGORITHM=1,2,3). Οπότε έχουμε κοινά αποτελέσματα τα εξής (παραθέτουμε μόνο για ALGORITHM=1):

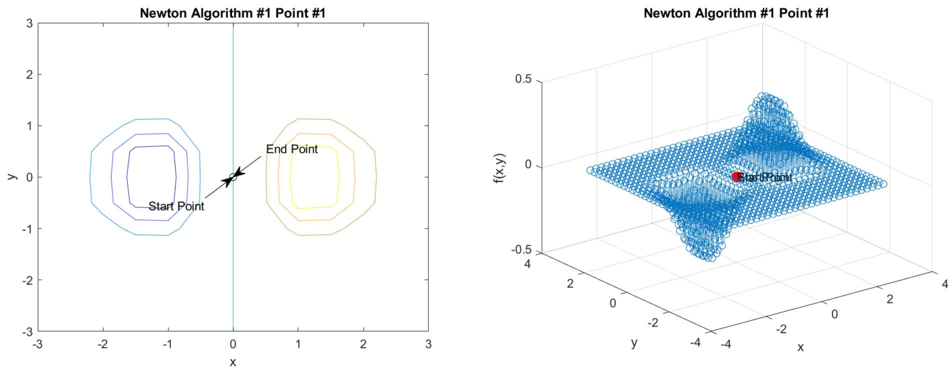

Figure 11: Σημείο 1 Αλγόριθμοι 1 έως 3

6.2 Σημείο 2 $(x, y) = (-1, -1)$

Ο αλγόριθμος δεν θα ξεκινήσει καθώς ο Εσσιανός δεν είναι θετικά ορισμένος και συνεπώς εμφανίζεται το αντίστοιχο μήνυμα στην οθόνη μας. Πράγματι

$$\nabla^2 f(-1,-1) = \begin{bmatrix} 0.5413 & 0.5413 \\ 0.5413 & -0.5413 \end{bmatrix}$$

αφού $\det\left[\nabla^2 f(-1,-1)\right] = -0.5861 < 0$

6.3 Σημείο 3 $(x, y) = (1, 1)$

Ο αλγόριθμος δεν θα ξεκινήσει καθώς ο Εσσιανός είναι αρνητικά ορισμένος και συνεπώς εμφανίζεται το αντίστοιχο μήνυμα στην οθόνη μας. Πράγματι

$$\nabla^2 f(1,1) = \begin{bmatrix} -0.5413 & -0.5413 \\ -0.5413 & 0.5413 \end{bmatrix} < 0$$

αφού $\det\left[\nabla^2 f(-1,-1)\right] = -0.5861 < 0$ και $-0.5413 < 0$

6.4 Σημείο 4 $(x, y) = (-0.9, 0.3)$

Για να αποδείξουμε την ορθή χρήση του αλγορίθμου βρίσκουμε ένα σημείο όπου ο εσσιανός είναι θετικά ορισμένος. Πράγματι

$$\nabla^2 f(-0.9, 0.3) = \begin{bmatrix} 1.0785 & -0.0533 \\ -0.0533 & 0.3437 \end{bmatrix} > 0$$

αφού $\det\left[\nabla^2 f(-0.9, 0.3\right] = 0.3678 > 0$ και $1.0785 > 0$.
Καταρχάς, για να βρούμε σταθερό γ, τρέχουμε τον αλγόριθμο για διάφορα γ από 0.1 έως 2 περιορίζοντας την ακρίβεια για ταχύτερα αποτελέσματα. Επίσης θέτουμε τον μέγιστο αριθμό επαναλήψεων στις 1000 που σημαίνει ότι για 1000 επαναλήψεις ο αλγόριθμος μάλλον δημιουργέι ταλαντώσεις (εμφανίζονται για $\gamma \approx 1.3$). Επιλέγουμε $\gamma = 0.1$ καθώς για μεγάλα γ ο εσσιανός δεν είναι θετικά ορισμένος.

Σημειώνεται πως αύτο μπορεί να γίνει θεωρητικά έτσι ώστε $\left|\frac{x(k+1)}{x(k)}\right| < 1$ αλλά εμείς περιοριζόμαστε σε πειραματικά αποτελέσματα. Για αυτόν τον λόγο παραθέτουμε το γράφημα των αριθμό των επαναλήψεων για διάφορα γ (στα "μεγάλα" γ ο αλγόριθμος κάνει abort οπότε $k = 2$). Ο αλγόριθμος συγκλίνει σε 7 επαναλήψεις για σταθερό γ, 9 αν ελαχιστοποιήσουμε την $h(\gamma) = f\left(x(k) - \gamma \cdot \nabla f\left(x(k)\right)\right)$ και 84 με τον κανόνα Armijo. Παραθέτουμε τα αποτελέσματα για κάθε αλγόριθμο (ALGORITHM=1,2,3):

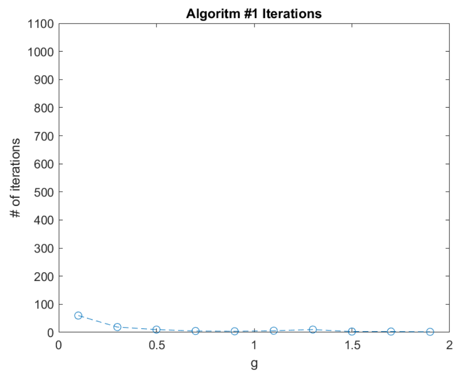

Figure 12: Σημείο 4 Αλγόριθμος 1 Επαναλήψεις

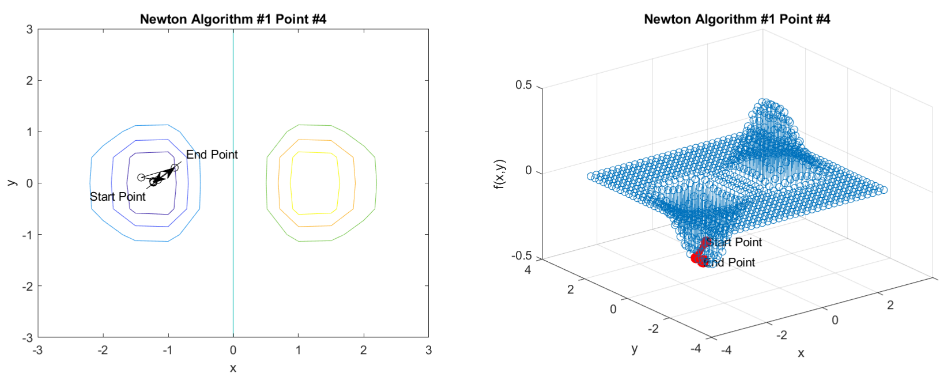

Figure 13: Σημείο 4 Αλγόριθμος 1

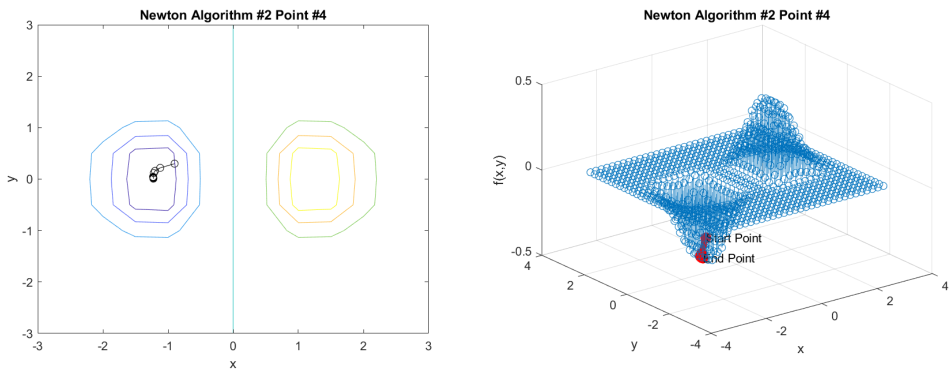

Figure 14: Σημείο 4 Αλγόριθμος 2

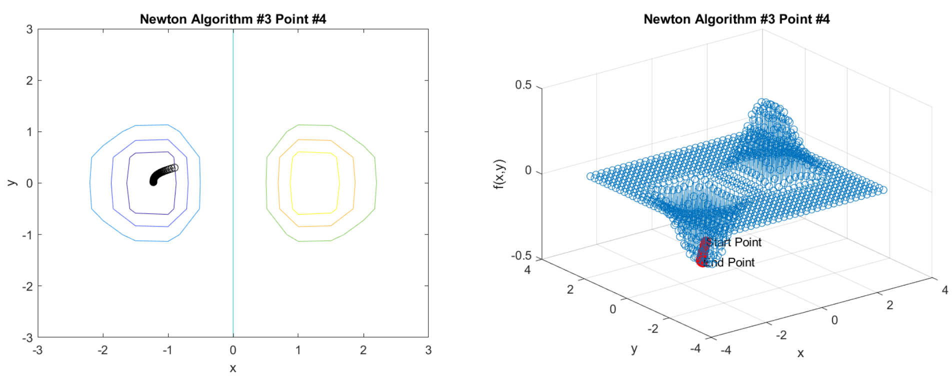

Figure 15: Σημείο 4 Αλγόριθμος 3

7 Μέθοδος Levenberg-Marquardt (ΘΕΜΑ 4)

Η προσομοιώσεις βρίσκονται στο αρχείο **thema4.m**. Τα figures αποθηκεύονται αυτόματα στον φάκελο *PlotsLevenbergMarquardt*.

7.1 Σημείο 1 $(x, y) = (0, 0)$

Επειδή ισχύει ότι $|\nabla f(0,0)| = 0$ ο αλγόριθμος θα τερματίζει απευθείας ανεξαρτήτου μεθόδου επιλογής γ. Οπότε έχουμε κοινά αποτελέσματα τα εξής:

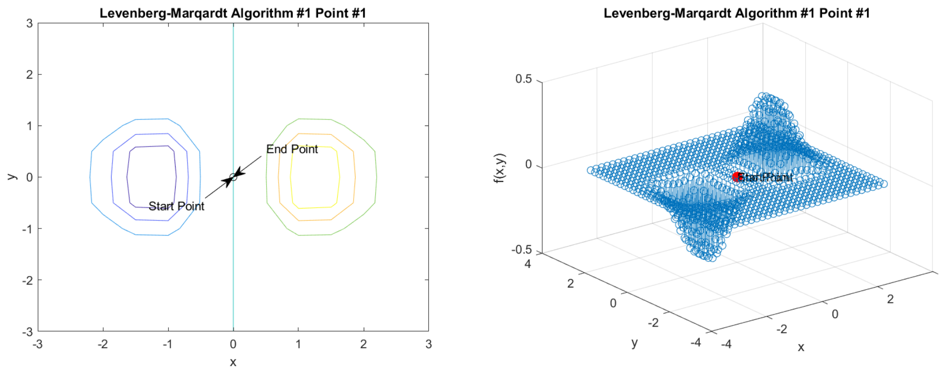

Figure 16: Σημείο 1 Αλγόριθμοι 1 έως 3

7.2 Σημείο 2 $(x, y) = (-1, -1)$

Εδώ λύνεται το πρόβλημα αν ο εσσινανός δεν είναι θετικά ορισμένος. Καταρχάς, για να βρούμε σταθερό γ, τρέχουμε τον αλγόριθμο για διάφορα γ από 0.1 έως 2 περιορίζοντας την ακρίβεια για ταχύτερα αποτελέσματα. Επίσης θέτουμε τον μέγιστο αριθμό επαναλήψεων στις 1000 που σημαίνει ότι για 1000 επαναλήψεις ο αλγόριθμος μάλλον δημιουργέι ταλαντώσεις. Εδώ όμως τα μεγάλα γ δηλώνουν ταχύτερη απόκλιση. Επιλέγουμε $\gamma = 0.1$. Σημειώνεται πως αύτο μπορεί να γίνει θεωρητικά έτσι ώστε $\left|\frac{x(k+1)}{x(k)}\right| < 1$ αλλά εμείς περιοριζόμαστε σε πειραματικά αποτελέσματα. Ο αλγόριθμος συγκλίνει σε 114 επαναλήψεις για σταθερό γ, 328 αν ελαχιστοποιήσουμε την $h(\gamma) = f\left(x(k) - \gamma \cdot \nabla f\left(x(k)\right)\right)$ και 793 με τον κανόνα Armijo. Παραθέτουμε τα αποτελέσματα για κάθε αλγόριθμο (ALGORITHM=1,2,3):

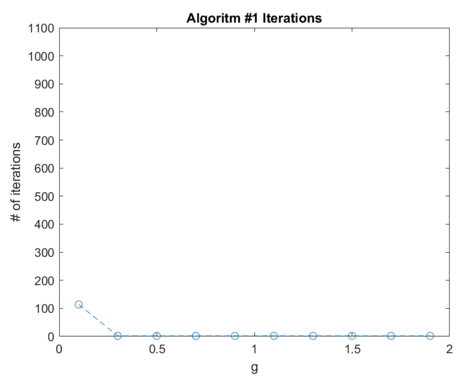

Figure 17: Σημείο 3 Αλγόριθμος 1 Επαναλήψεις

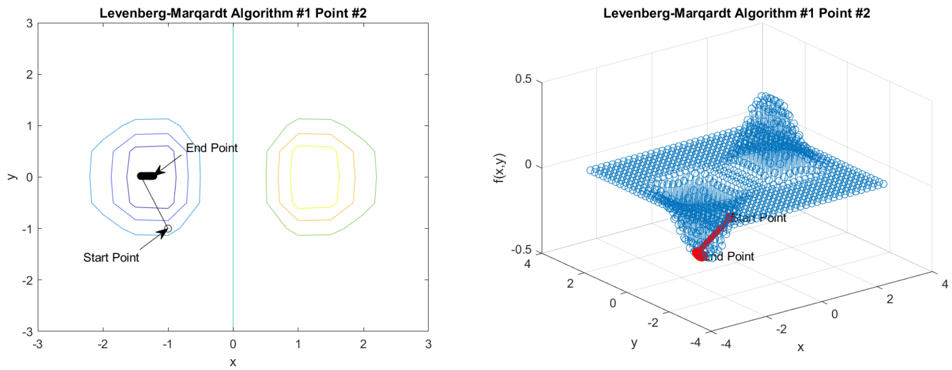

Figure 18: Σημείο 2 Αλγόριθμος 1

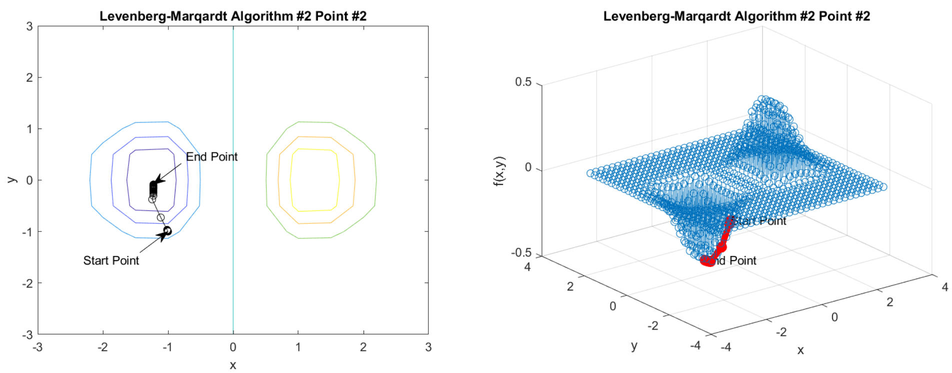

Figure 19: Σημείο 2 Αλγόριθμος 2

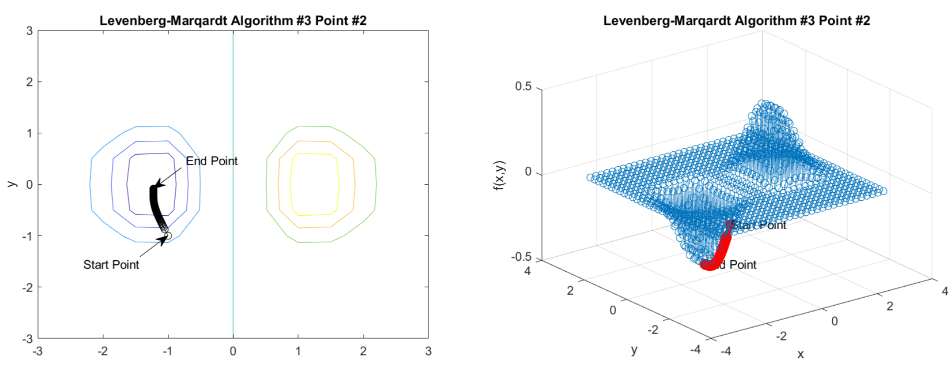

Figure 20: Σημείο 2 Αλγόριθμος 3

7.3 Σημείο 3 $(x, y) = (1, 1)$

Καταρχάς, για να βρούμε σταθερό γ, τρέχουμε τον αλγόριθμο για διάφορα γ από 0.1 έως 2 περιορίζοντας την ακρίβεια για ταχύτερα αποτελέσματα. Επίσης θέτουμε τον μέγιστο αριθμό επαναλήψεων στις 1000 που σημαίνει ότι για 1000 επαναλήψεις ο αλγόριθμος μάλλον δημιουργεί ταλαντώσεις. Επιλέγουμε $\gamma = 0.1$. Σημειώνεται πως αύτο μπορεί να γίνει θεωρητικά έτσι ώστε $\left|\frac{x(k+1)}{x(k)}\right| < 1$ αλλά εμείς περιοριζόμαστε σε πειραματικά αποτελέσματα. Για αυτόν τον λόγο παραθέτουμε το γράφημα των αριθμό των επαναλήψεων για διάφορα γ. Ωστόσο επειδή ο αλγόριθμος αποκλίνει, μεγαλύτερο γ οδηγεί σε πιο γρήγορη απόκλιση (η ταχύτερη σύγλιση για μεγαλύτερο γ είναι ψευδαίσθηση). Παρόλο που λύσαμε το πρόβλημα περί θετικά ορισμένου εσσιανού, ο αλγόριθμος αποκλίνει και σταματάει σε 155 επαναλήψεις για σταθερό γ, σε 2 επαναλήψεις αν ελαχιστοποιήσουμε την $h(\gamma) = f\left(x(k) - \gamma \cdot \nabla f\left(x(k)\right)\right)$ ή 159 για τον κανόνα Armijo. Παραθέτουμε τα αποτελέσματα για κάθε αλγόριθμο (ALGORITHM=1,2,3):

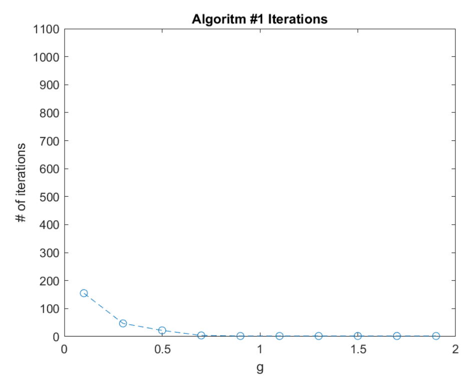

Figure 21: Σημείο 3 Αλγόριθμος 1 Επαναλήψεις

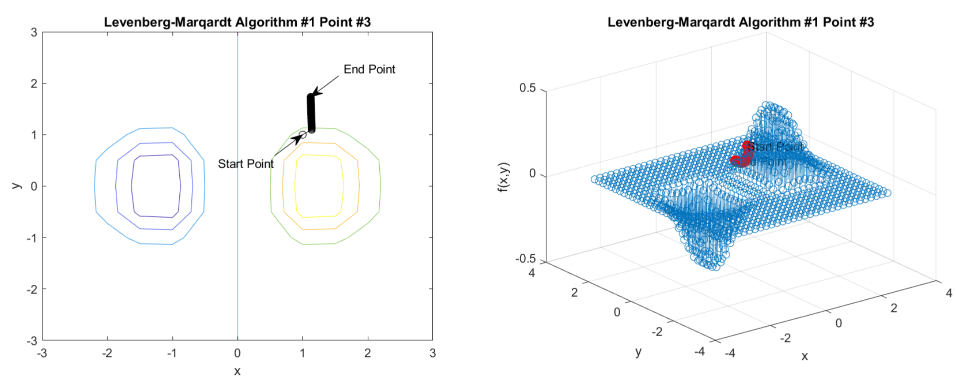

Figure 22: Σημείο 3 Αλγόριθμος 1

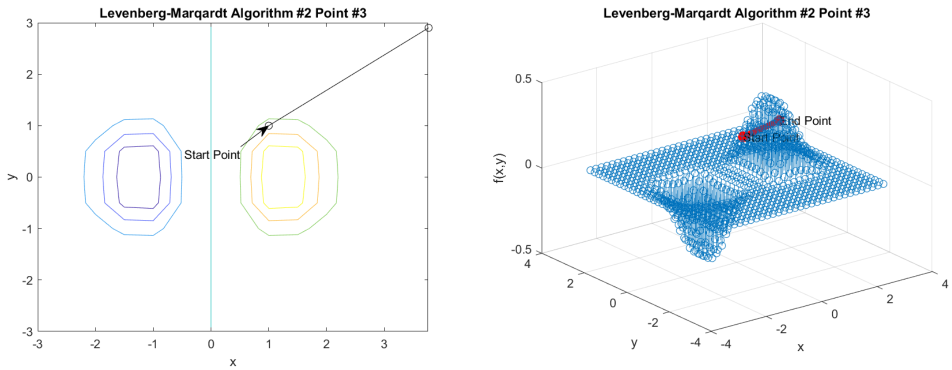

Figure 23: Σημείο 3 Αλγόριθμος 2

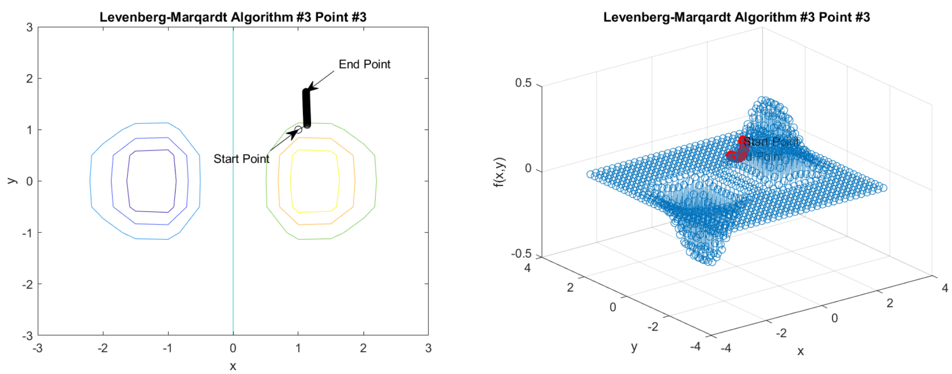

Figure 24: Σημείο 3 Αλγόριθμος 3

8 Συμπεράσματα

ALGORITHM	#1	#2	#3
Point #1 $(0,0)$	1	1	1
Point #2 $(-1,-1)$	176	3061	2012
Point #3 $(1,1)$	diverges	diverges	diverges

Table 1: Steepest Descent Iterations

ALGORITHM	#1	#2	#3
Point #1 $(0,0)$	1	1	1
Point #2 $(-1,-1)$	hessian<0		
Point #3 $(1,1)$	hessian<0		
Point #4 $(-0.9, 0.3)$	7	9	84

Table 2: Newton Iterations

ALGORITHM	#1	#2	#3
Point #1 $(0,0)$	1	1	1
Point #2 $(-1,-1)$	114	328	793
Point #3 $(1,1)$	diverges	diverges	diverges

Table 3: Levenberg-Marquardt Iterations

Η μέθοδος Steepest Descent συγκλίνει σε λιγότερες επαναλήψεις αν χρησιμοποιήσουμε τον αλγόριθμο με σταθερό γ σε σχέση με την ελαχιστοπίηση της $h(\gamma) = f\left(x(k) - \gamma \cdot \nabla f\left(x(k)\right)\right)$ ή τον κανόνα του Armijo που μάλιστα χρειάζονται τον περισσότερο χρόνο. Ανάμεσα στις 2 καλύτερες επιλογές, προτιμούμε την πιο απλή υλοποίηση με σταθερο γ. Προσοχή ότι για να είναι αυτός ο πιο αποδοτικός αλγόριθμος πρέπει να βρούμε από πριν γ που η μέθοδος συγκλίνει. Αξιζει βέβαια να σημειωθεί πως ο αλγόριθμος με την ελαχιστοποίηση της $h(\gamma) = f\left(x(k) - \gamma \cdot \nabla f\left(x(k)\right)\right)$ αργεί καθώς το γ που υπολογίζει είναι πολύ μικρό στις τελευταίες επαναλήψεις.
Η μέθοδος Newton παρουσιάζει παρόμοια αποτελέσματα για σταθερό γ και τον κανόνα του Armijo ενώ για την ελαχιστοποίηση της $h(\gamma) = f\left(x(k) - \gamma \cdot \nabla f\left(x(k)\right)\right)$ συγκλίνει σε μόλις 9 επαναλήψεις. Προτιμούμε τον τελευταίο αλγόριθμο για προφανείς λόγους.
Η μέθοδος Levenberg-Marqardt αποτελεί εξέίξη της μεθόδου του Newton καθως εξαλείφει το πρόβλημα περί θετικά ορισμένου εσσιανού. Για σταθερό γ ο αλγόριθμος τερματίζει στις λιγότερες επαναλήψεις, ακολουθεί ο αλγόριθμος με την ελαχιστοποίηση της $h(\gamma) = f\left(x(k) - \gamma \cdot \nabla f\left(x(k)\right)\right)$ και τέλος βρίσκεται ο κανόνας του Armijo.
Ανάμεσα στις 3 μεθόδους, η μέθοδος Levenberg-Marqardt φαίνεται να αντιμετωπίζει καλύτερα το πρόβλημα τις ελαχιστοποίησης αφενός γιατι συγκλίνει σε λιγότερες επαναλήψεις, αφετέρου γιατί προσεγγίζει και τις δύο άλλες μεθόδους βάσει την επιλογή του $m(k)$ όπως αναφέρθηκε στην ενότητα 2.3.
Ωστόσο και οι 3 μέθοδοι αποτυγχάνουν να βρουν το ελάχιστο στο σημείο 3 $(x, y) = (1, 1)$ καθώς παγιδεύονται σε περιοχή πολύ χαμηλού gradient - το διάνυσμα κλίσης “αποπροσανατολίζεται”.