

A dark blue vertical bar is positioned on the left side of the page. A blue arrow-shaped box points to the right from this bar, containing the date. Below the arrow, several thin, curved lines in dark blue and light grey originate from the vertical bar and sweep upwards and to the right.

16/5/2018

## Άσκηση 3 : Εύρεση Ελαχίστου σε συνάρτηση δύο μεταβλητών

ΓΙΩΡΓΟΣ ΜΠΑΛΑΟΥΡΑΣ (8861)

# Τεχνικές Βελτιστοποίησης

Ονοματεπώνυμο: Μπαλαούρας Γεώργιος

AEM: 8861

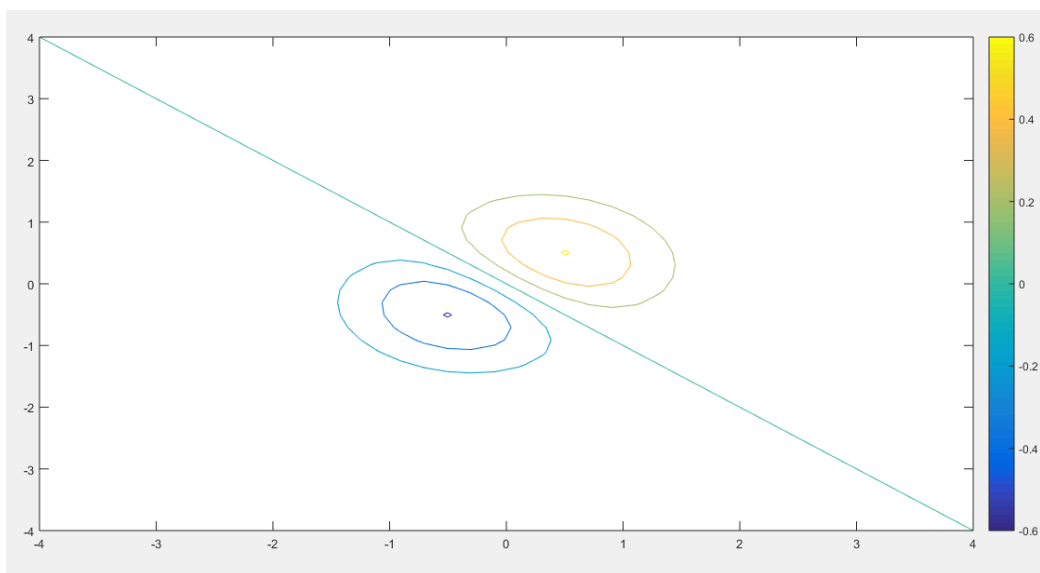
E-mail: [mpalaourg@ece.auth.gr](mailto:mpalaourg@ece.auth.gr)

Εξάμηνο: 6<sup>ο</sup> Ηλεκτρονικής

## Άσκηση 3 : Εύρεση ελαχίστου σε συνάρτηση δύο μεταβλητών

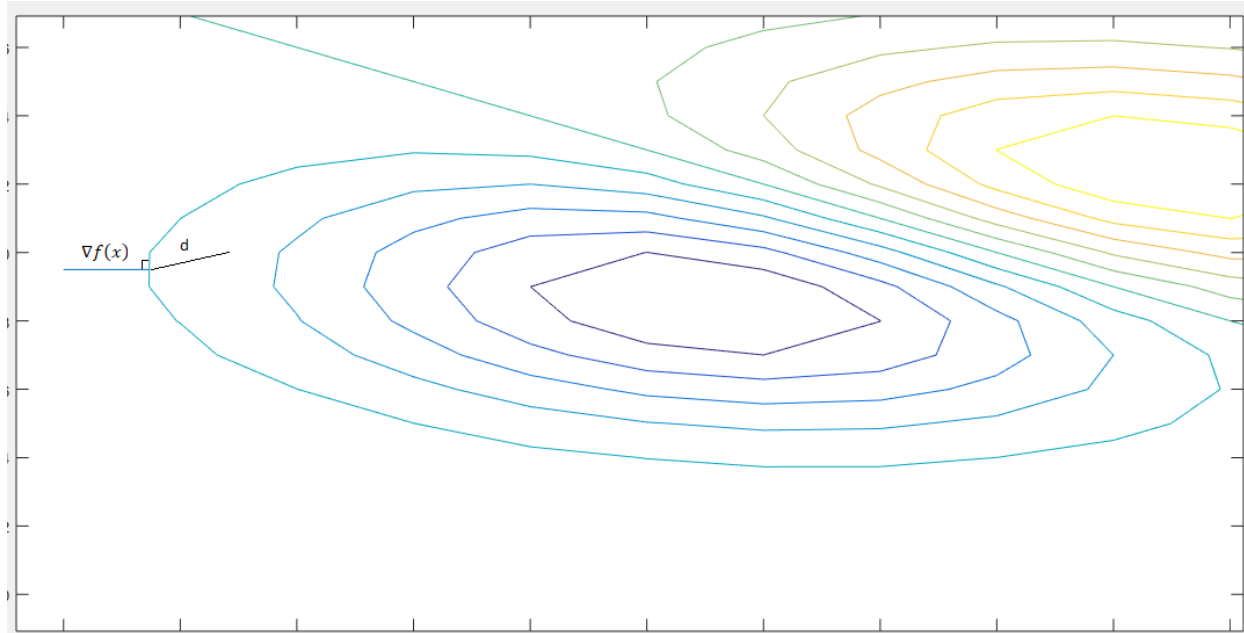
Αντικείμενο αυτής της άσκησης, ήταν η εφαρμογή τριών μεθόδων, της μεθόδου των Levenberg - Marquardt, της μεθόδου των συζυγών κλίσεων και της μεθόδου σχεδόν Newton, σε συνάρτηση  $f$ , για την οποία δεν υπήρχε αναλυτικός τύπος, αλλά υπήρχε πρόσβαση στις αριθμητικές τιμές των  $f(x)$ ,  $\nabla f(x)$  και  $\nabla^2 f(x)$ , με  $x \in \mathbb{R}^2$ .

Οι τρεις αυτοί μέθοδοι, ανήκουν στις μεθόδους κλίσης. Βασικό χαρακτηριστικό τους, όπως είδαμε κι στην προηγούμενη άσκηση, είναι η αξιοποίηση της καθετότητας του διανύσματος κλίσης,  $\nabla f(x)$ , στην ισοβαρή καμπύλη της  $f$ .



Σχήμα 1: Ισοβαρείς καμπύλες της  $f$

Γνωρίζοντας, ότι οι μέθοδοι στηρίζονται στη ιδέα της επαναληπτικής διαδικασίας , δηλαδή ισχύει  $f(x_{k+1}) < f(x_k)$ ,  $k = 0, 1, 2, \dots$  , ορίζουμε το  $\bar{x} = x + \gamma d$ , με  $\gamma \geq 0$  και διάνυσμα κατεύθυνσης  $d \in \mathbb{R}^n$  (εδώ  $d \in \mathbb{R}^2$ ). Για να μειωθεί στην επόμενη επανάληψη το  $\bar{x}$ , πρέπει το διάνυσμα κατεύθυνσης να σχηματίζει αμβλεία γωνία με το διάνυσμα κλίσης, δηλαδή  $\nabla f^T(x)d < 0$ , έτσι ώστε να «βλέπει» προς μικρότερη ισοβαρή καμπύλη.

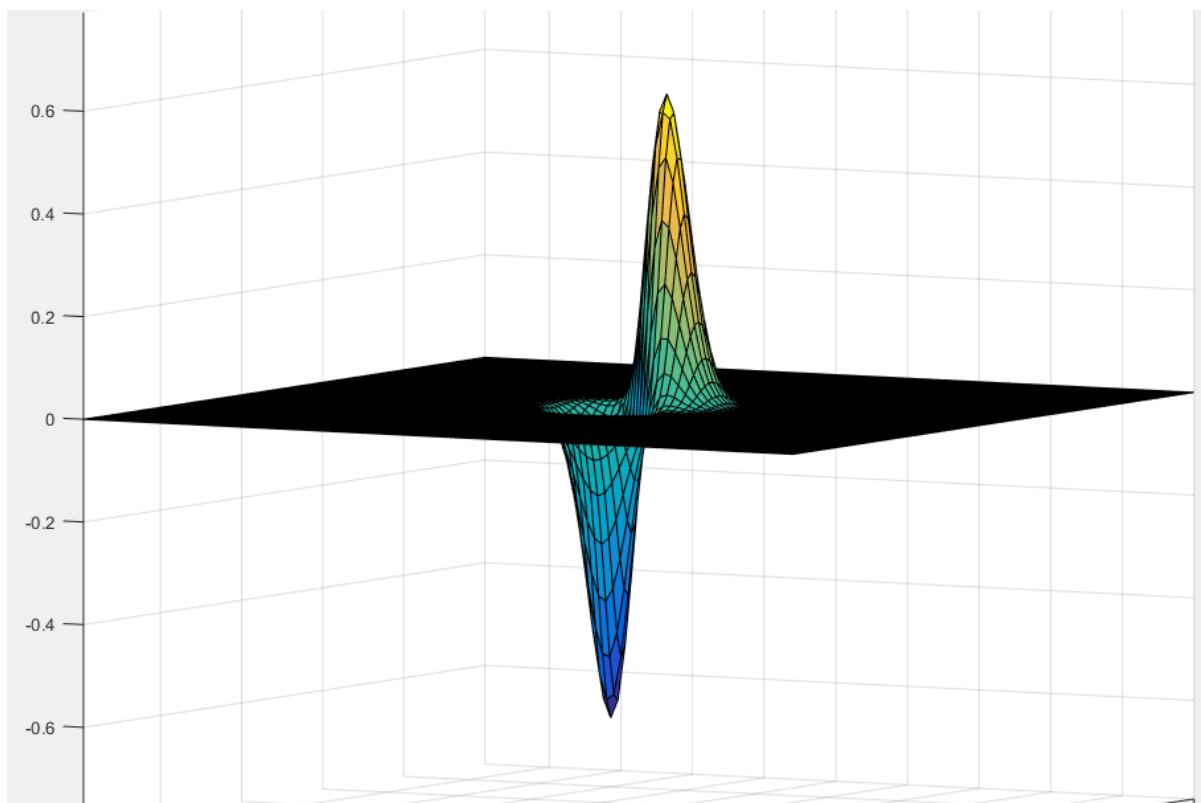


Σχήμα 2: Μεγεθυμένη η φωτογραφία του Σχήματος 1

Στο σχήμα 2, φαίνεται η αναγκαιότητα τα δύο διανύσματα να σχηματίζουν αμβλεία γωνία. Επίσης, γίνεται σαφές πως η επιλογή μικρού  $\gamma$  θα έχει ως αποτέλεσμα αργή σύγκλιση στο ζητούμενο σημείο, ενώ η επιλογή μεγάλου  $\gamma$ , μπορεί να οδηγήσει σε ταλαντώσεις ή σε ισοβαρείς καμπύλες με μεγαλύτερη τιμή.

Ένα ακόμα, χαρακτηριστικό (αρκετών) αλγορίθμων κλίσης είναι η σχέση που δίνει το επόμενο διάνυσμα  $x_{k+1}$ , δηλαδή  $x_{k+1} = x_k - \gamma_k \Delta_k \nabla f(x_k)$ , όπου  $\Delta_k = \Delta_k^T > 0$ . (1)

Επιπλέον, για να έχω μια γενική ιδέα της αντικειμενικής συνάρτησης χρησιμοποιώ το `plot_f.m`, όπου εκτός της γραφικής παράστασης της  $f$ , υπολογίζεται (στο περίπου) το  $\max\{f\} \cong 0.6065$  και το  $\min\{f\} \cong -0.6065$



Σχήμα 3: Γραφική παράσταση της  $f$

### Εφαρμογή των μεθόδων:

Στο πρώτο σημείο που μας δόθηκε,  $(-0.5, -0.5)$ :

Υπολογίζοντας το  $f(-0.5, -0.5) = -0.6065$ ,

Μέσω του  $\text{grad}f(-0.5, -0.5) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ ,

και  $\text{hessian}f(-0.5, -0.5) = \begin{bmatrix} 1.8196 & 0.6065 \\ 0.6065 & 1.8196 \end{bmatrix}$ , με ιδιοτιμές  $\lambda = \begin{bmatrix} 1.2131 \\ 2.4261 \end{bmatrix}$ , άρα θετικά ορισμένος. Δηλαδή, το σημείο είναι το (ολικό) ελάχιστο της  $f$ , άρα καμία μέθοδος δεν θα εκτελεστεί μ' αυτό το αρχικό σημείο.

Ομοίως, στο δεύτερο σημείο που μας δόθηκε,  $(0.5, 0.5)$ :

$$f(0.5, 0.5) = 0.6065,$$

$$\text{Μέσω του } \text{grad}f(0.5, 0.5) = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

και  $\text{hessian}f(0.5, 0.5) = \begin{bmatrix} -1.8196 & -0.6065 \\ -0.6065 & -1.8196 \end{bmatrix}$ , με ιδιοτιμές  $\lambda = \begin{bmatrix} -2.4261 \\ -1.2131 \end{bmatrix}$ , άρα αρνητικά ορισμένος. Δηλαδή, το σημείο είναι το (ολικό) μέγιστο της  $f$ , άρα και για αυτό το αρχικό σημείο, οι μέθοδοι δεν θα εκτελεστούν.

**Μέθοδος Levenberg - Marquardt:** Η μέθοδος αυτή, λύνει το «πρόβλημα» της μεθόδου Newton. Όπως είδαμε στην προηγούμενη εργασία, η μέθοδος Newton για να λειτουργήσει και να ισχύει η ιδιότητα της επαναληπτικής διαδικασίας, απαιτούσε  $\nabla^2 f(x) > 0$ , σε διαφορετική περίπτωση είτε υπήρχε ταλάντωση είτε παραγόταν ακολουθία σημείων  $\{x_k\}$  που οδηγούσε σε λύση μακριά από το ζητούμενο σημείο. Έτσι, η μέθοδος των Levenberg – Marquardt στην (1), θέτει όπου  $\Delta_k$ ,  $[\nabla^2 f(x_k) + \mu_k I]^{-1}$ , δηλαδή επιλέγουμε κατάλληλα το  $\mu_k$  για να είναι  $\Delta_k = [\nabla^2 f(x_k) + \mu_k I]^{-1} > 0$ . Η επιλογή του  $\mu_k$  δεν είναι τυχαία, αλλά επιλέγεται έτσι ώστε να ισχύει:  $\mu_k > |\lambda_{\max}\{\nabla^2 f(x_k)\}|$ , όπου  $\lambda_{\max}$ : η μέγιστη ιδιοτιμή του  $\nabla^2 f(x_k)$ .

- Αρχικό σημείο  $(-2, -1)$  και  $\varepsilon = 0.001$ .

Υπολογίζοντας τον εσσιανό στο συγκεκριμένο σημείο:

$\text{hessian}f(-2, -1) = \begin{bmatrix} -0.2291 & -0.1213 \\ -0.1213 & -0.0135 \end{bmatrix}$ , παρατηρούμε ότι οι ελάχιστες ορίζουσες είναι αρνητικές,  $D_1 = -0.2291 < 0$  και  $D_2 = \det\left(\begin{bmatrix} -0.2291 & -0.1213 \\ -0.1213 & -0.0135 \end{bmatrix}\right) = -0.0116 < 0$ . Άρα, ο εσσιανός αρνητικά ορισμένος, ωστόσο βάσει της παραπάνω ανάλυσης και εκτελώντας το `levenberg_marquardt_plot_x_y`, έχουμε:

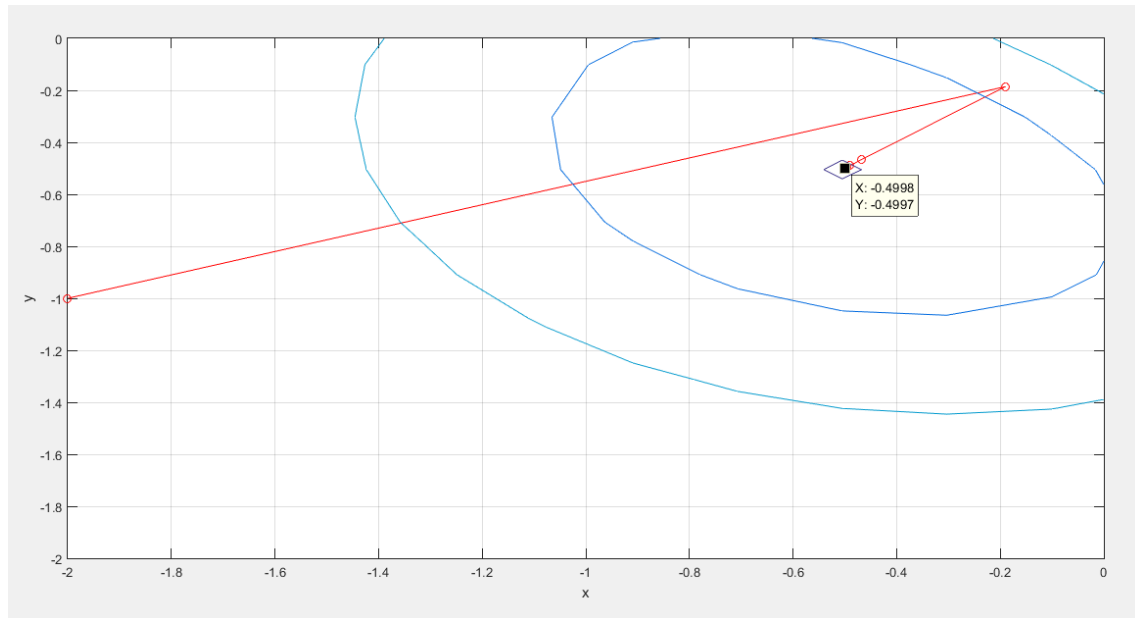
**Σημείωση:** Για τις παρακάτω περιπτώσεις, ελέγχεται μέσω της Boolean συνάρτησης, `criteria_testing.m` αν ικανοποιούνται τα κριτήρια σύγκλισης:

$$\exists \beta, \text{ με } 0 < \beta < 1 : d_k^T * \nabla f(x_{k+1}) > \beta * d_k^T * \nabla f(x_k)$$

$$\exists \alpha, \text{ με } 0 < \alpha < \beta < 1 : f(x_{k+1}) \leq f(x_k) + \alpha * \gamma_k * d_k^T * \nabla f(x_k)$$

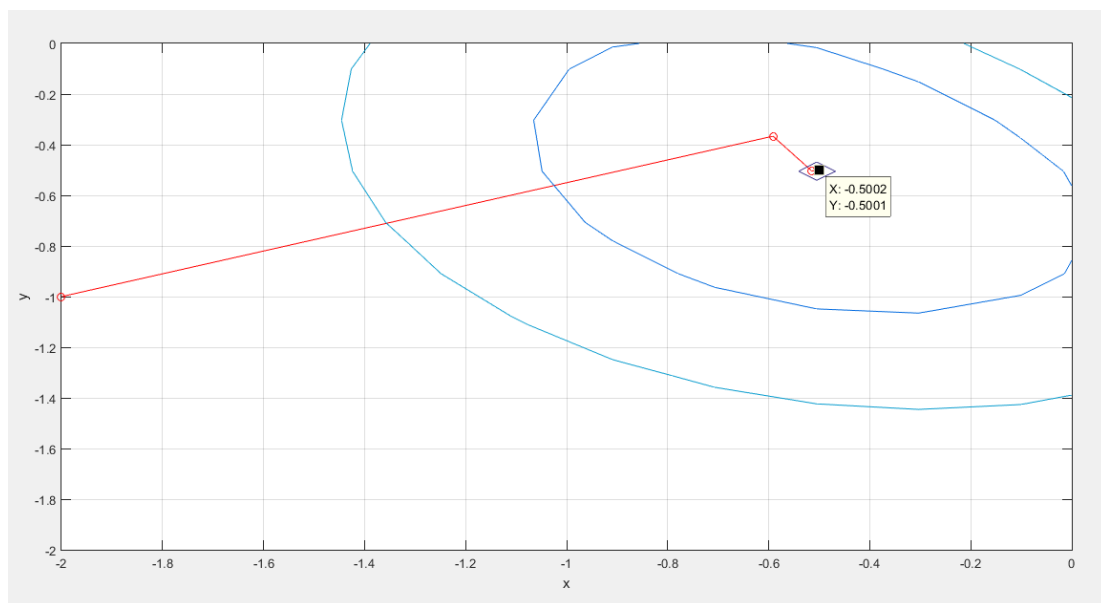
Αυτά κριτήρια, διασφαλίζουν το  $\gamma_k$  να μην επιλέγεται οσοδήποτε μικρό, αλλά και ότι δεν θα πάρει πολύ μεγάλες τιμές, αντίστοιχα. Δηλαδή, αποκλείονται περιπτώσεις ταλάντωσης και απόκλισης.

I. Για  $\gamma_k = \text{σταθερό} = 1.4$



Σχήμα 4: Σύγκλιση του ζεύγους  $(x, y)$  για σταθερό  $\gamma$  [Μέθοδος Levenberg – Marquardt]

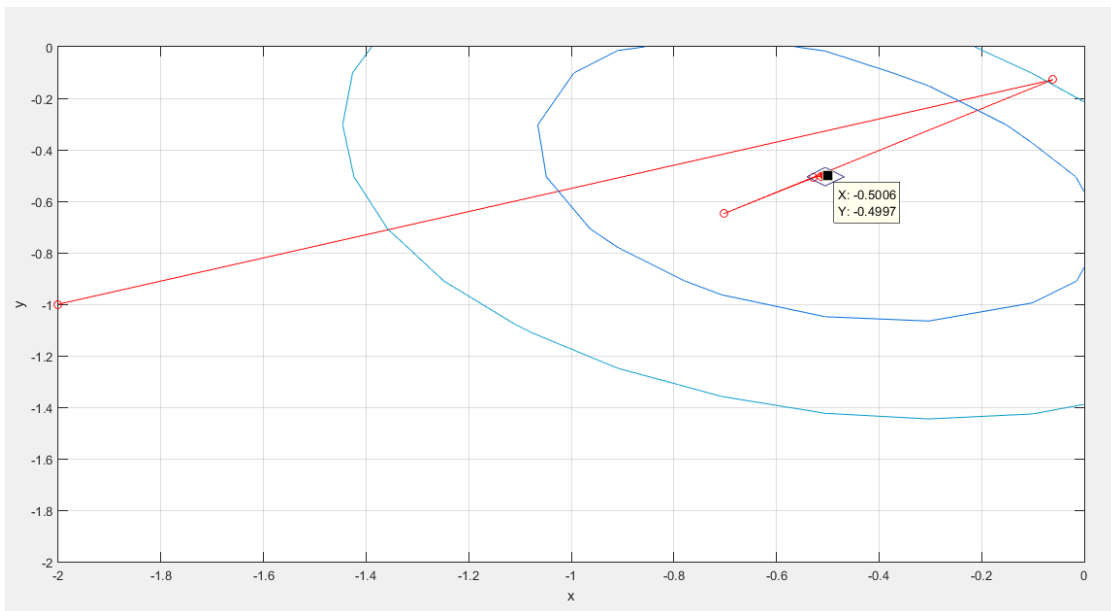
II. Χρησιμοποιώντας την `golden_section.m`, και επιλέγοντας εκείνο το  $\gamma_k$  που σε κάθε επανάληψη ελαχιστοποιεί -εσωτερικά- την  $h(\gamma) = f(x_k + \gamma_k d_k)$ .



Σχήμα 5: Σύγκλιση του ζεύγους  $(x, y)$  για ελαχιστοποιημένο  $\gamma$  [Μέθοδος Levenberg - Marquardt]

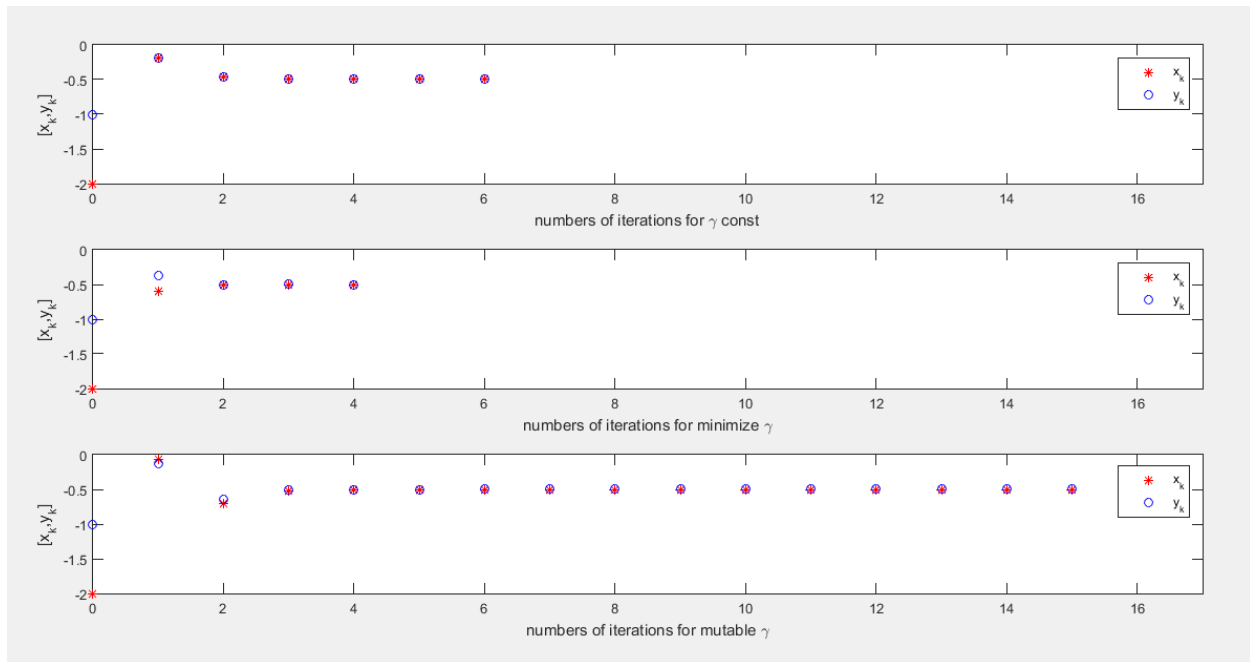
- III. Για το μεταβλητό  $\gamma$ , η ιδέα ήταν να ξεκινήσω από μια μεγάλη τιμή για το  $\gamma$ ,  $\gamma = 1.5$  και όσο προχωράω σε «μικρότερες» ισοβαρείς, να μειώνω το  $\gamma$ , κατά 10%. Η ιδέα, βασίστηκε στο γεγονός, πως αν είμαι μακριά από το σημείο θέλω να κάνω μεγαλύτερα βήματα και όσο πλησιάζω στο ζητούμενο σημείο, να μικραίνει το βήμα μου. Για να αποφύγω τα πολύ μικρά βήματα, ελέγχω αν το  $\gamma$ , πέσει κάτω από μια τιμή και αν γίνει κάτι τέτοιο, το τριπλασιάζω.

**Σημείωση:** Ο έλεγχος στο  $\gamma$ , ενεργοποιείται -κυρίως- για μικρά  $\varepsilon$ .



Σχήμα 6: Σύγκλιση του ζεύγους  $(x, y)$  για μεταβλητό  $\gamma$  [Μέθοδος Levenberg - Marquardt]

Για να γίνει πιο σαφής η σύγκριση μεταξύ των διαφορετικών «μορφών» της μεθόδου των Levenberg - Marquardt, χρησιμοποιώ το `levenberg_marquardt_plot_x_y_compare.m`, όπου το ζεύγος  $(x, y)$ , για κάθε μια από τις παραπάνω μορφές, σχεδιάζεται συναρτήσει των αριθμών επαναλήψεων.



Σχήμα 7: Σύγκλιση των  $(x_k, y_k)$  [Μέθοδος Levenberg - Marquardt]

**Μέθοδος των συζυγών κλίσεων:** Η μέθοδος αυτή, επιλύει ένα άλλο μειονέκτημα της μεθόδου Newton, τον υπολογισμό του εσσιανού πίνακα. Η προηγούμενη άσκηση, μας απέδειξε ότι εφ' όσον ο εσσιανός της αντικειμενικής συνάρτησης είναι θετικά ορισμένος, η μέθοδος Newton παρουσιάζει πολύ καλή ταχύτητα σύγκλισης, σε σύγκριση με την μέθοδο της μέγιστης καθόδου. Ωστόσο, απαιτούσε τον υπολογισμό  $n^2$  παραγώγων δεύτερης τάξης, που μπορεί ενίοτε είτε να μην μπορούν να υπολογιστούν είτε να είναι δύσκολος ο υπολογισμός τους. Η μέθοδος των συζυγών κλίσεων, στην πρώτη επανάληψη λειτουργεί σαν την *μέθοδο της μέγιστης καθόδου* και από την επόμενη κιόλας επανάληψη παράγει τις συζυγείς κατευθύνσεις.

**Σημείωση:** Συζυγείς κατευθύνσεις, ονομάζονται τα διανύσματα  $d_i, d_j$  τα οποία για δοσμένο πίνακα  $Q > 0$  ικανοποιούν την σχέση  $d_i^T Q d_j = 0, \forall i, j, \text{ με } i \neq j$ .

Για να δημιουργηθούν  $k$  συζυγείς κατευθύνσεις, χρειαζόμαστε  $k$  γραμμικά ανεξάρτητα διανύσματα. Αυτά τα  $k$  γραμμικά ανεξάρτητα διανύσματα, μπορούν να θεωρηθούν τα διανύσματα κλίσης  $\nabla f(x_k)$  κάθε επανάληψης και έτσι βρίσκουμε  $x_{k+1} = x_k + \gamma_k * d_k$ , με  $d_k = -\nabla f(x_k) + \beta_k * d_{k-1}$ . Για τον υπολογισμό του  $\beta_k$ , μπορούμε να αποφύγουμε την περίπλοκη διαδικασία των *Gram - Schmidt* και λαμβάνοντας υπόψιν ότι η  $f$  δεν είναι τετραγωνική, γράφουμε: 
$$\beta_k = \frac{\nabla f^T(x_k) * [\nabla f(x_k) - \nabla f(x_{k-1})]}{\nabla f^T(x_{k-1}) * \nabla f(x_{k-1})}.$$

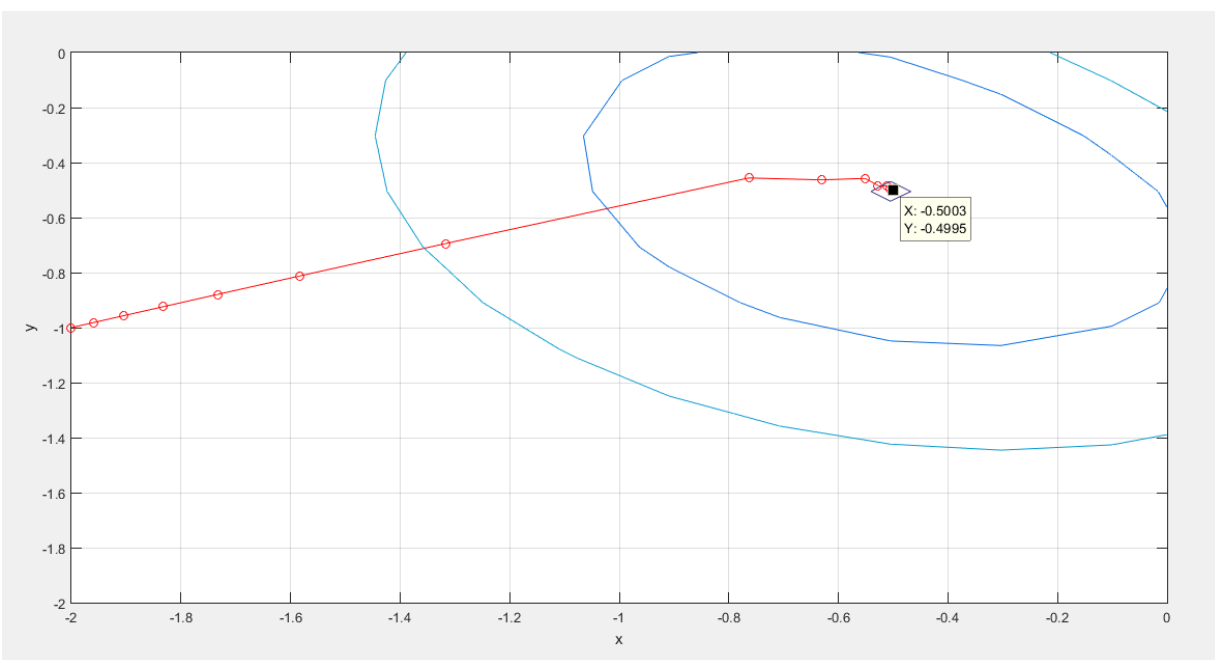


Η μέθοδος αυτή, εγγυάται για τετραγωνικές συναρτήσεις με το διάνυσμα  $x \in \mathbb{R}^n$ , σύγκλιση σε το πολύ  $n$  βήματα, δηλαδή επιταχύνει την ταχύτητα σύγκλισης της *Μέγιστης Καθόδου*, χωρίς τον υπολογισμό και την χρήση του εσσιανού.

- Αρχικό σημείο  $(-2, -1)$  και  $\varepsilon = 0.001$ .

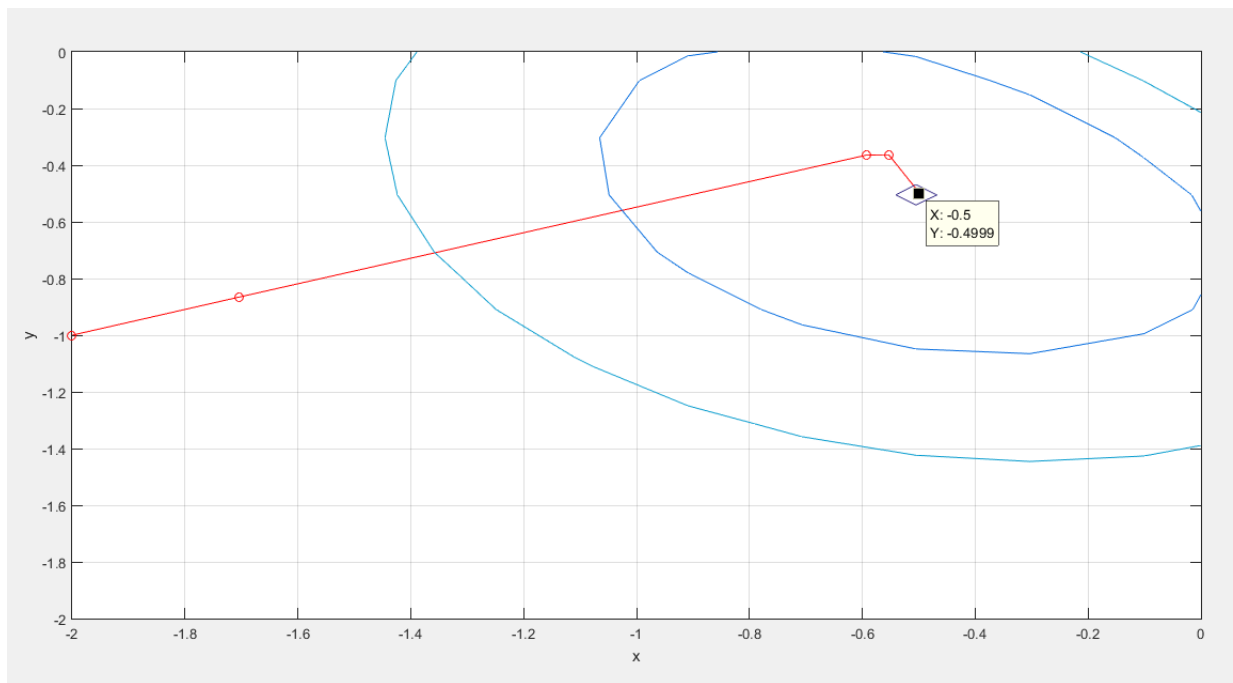
Χρησιμοποιώντας το `conjugate_plot_x_y.m` και αποφεύγοντας τον υπολογισμό του εσσιανού, με μόνο γνώση του `gradf`, παρατηρούμε:

- Για  $\gamma_k = \text{σταθερό} = 0.5459$



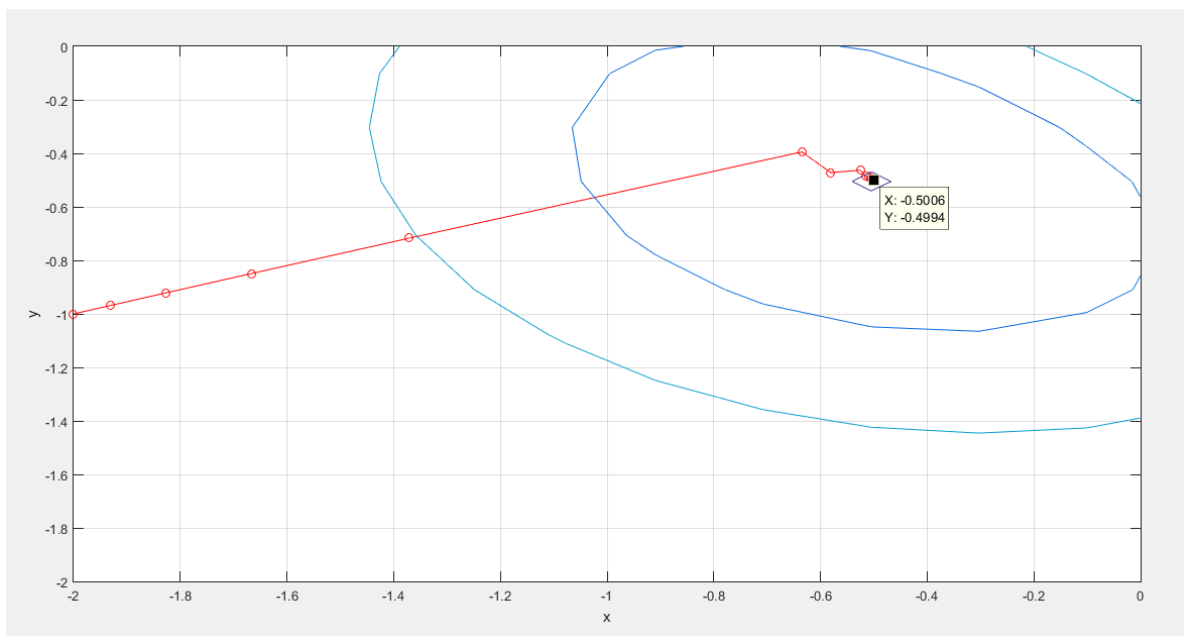
Σχήμα 8: Σύγκλιση του ζεύγους  $(x, y)$  για σταθερό  $\gamma$  [Μέθοδος Συζυγών Κλίσεων]

- Χρησιμοποιώντας την `golden_section.m`, και επιλέγοντας εκείνο το  $\gamma_k$  που σε κάθε επανάληψη ελαχιστοποιεί -εσωτερικά- την  $h(\gamma) = f(x_k + \gamma_k * d_k)$ .



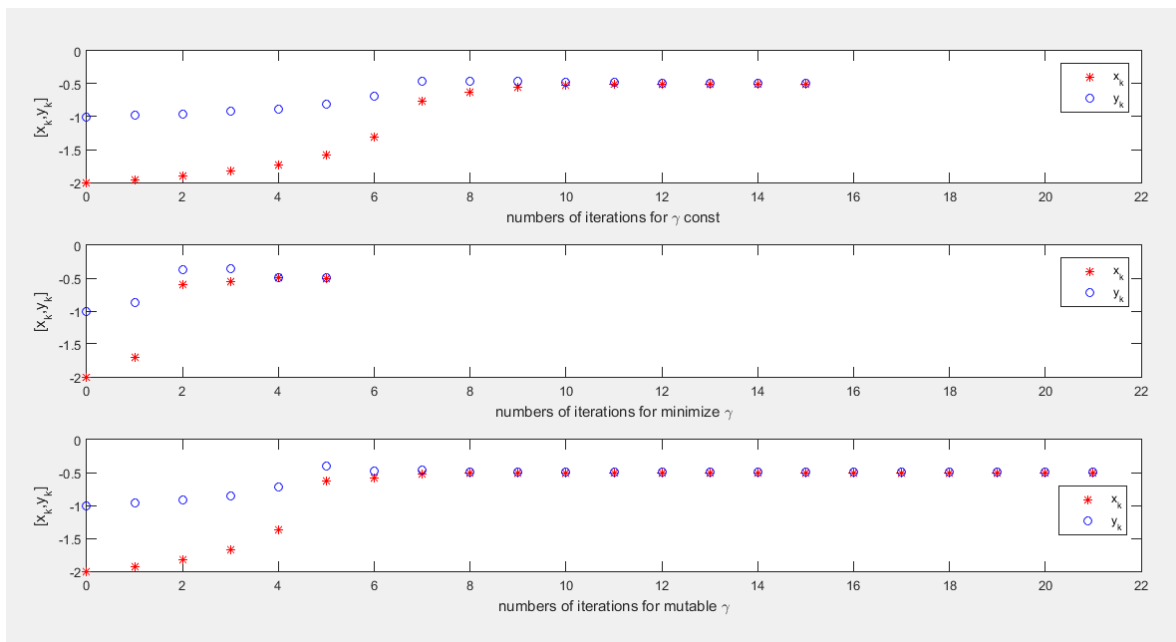
Σχήμα 9: Σύγκλιση του ζεύγους  $(x, y)$  για ελαχιστοποιημένο  $\gamma$  [Μέθοδος Συζυγών Κλίσεων]

- Για μεταβλητό  $\gamma_k$ , ακολουθώντας τον ίδιο κανόνα με πριν, μόνο που αρχικά  $\gamma = 0.9293$



Σχήμα 10: Σύγκλιση του ζεύγους  $(x, y)$  για μεταβλητό  $\gamma$  [Μέθοδος Συζυγών Κλίσεων]

Για να γίνει πιο εμφανής η σύγκριση μεταξύ των διαφορετικών «μορφών» της μεθόδου των συζυγών κλίσεων, χρησιμοποιώ το `conjugate_plot_x_y_compare.m`, όπου το ζεύγος  $(x, y)$ , για κάθε μια από τις παραπάνω μορφές, σχεδιάζεται συναρτήσει των αριθμών επαναλήψεων.



Σχήμα 11: Σύγκλιση των  $(x_k, y_k)$  [Μέθοδος Συζυγών Κλίσεων]

**Μέθοδος σχεδόν Newton:** Τέλος, η σχεδόν *Newton* έχει το ίδιο σκεπτικό με την μέθοδο των συζυγών κλίσεων, δηλαδή να αποφευχθούν οι πράξεις για τον υπολογισμό του εσσιανού, αλλά ακολουθεί διαφορετική τεχνική. Συγκεκριμένα, **δεν** υπολογίζουν τον εσσιανό, αλλά τον προσεγγίζουν, μέσω της σχέσης  $q_k \cong \nabla^2 f(x_{k+1}) * p_k$ , με  $q_k = \nabla f(x_{k+1}) - \nabla f(x_k)$  και  $p_k = x_{k+1} - x_k$ .

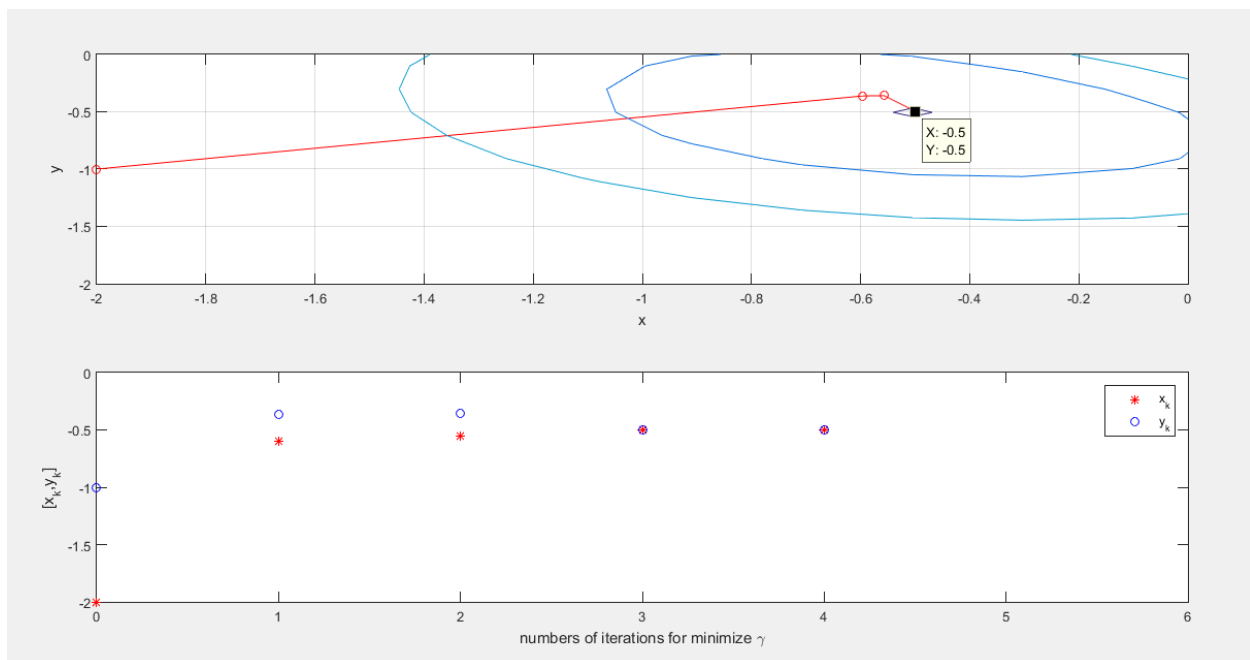
Το  $x_{k+1} = x_k + \gamma_k d_k$ , εφαρμόζεται κι σ' αυτήν την μέθοδο, το διάνυσμα κατεύθυνσης δίνεται από τον τύπο  $d_k = -\Delta_k \nabla f(x_k)$ , ο  $\Delta_k = \Delta_k^T$  είναι ένας θετικά ορισμένος πίνακας αλλά μεταβάλλεται από επανάληψη σε επανάληψη για να τείνει η κατεύθυνση  $d_k$  να προσεγγίσει την κατεύθυνση *Newton*. Στην πρώτη επανάληψη ο  $\Delta_0$ , συνήθως επιλέγεται ο μοναδιαίος πολλαπλασιασμένος με έναν θετικό ακέραιο αριθμό. (εδώ  $\Delta_0 = 5 \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ )

Για να υπολογίσουμε τον  $\Delta_{k+1} = \Delta_k + \frac{p_k p_k^T}{p_k^T q_k} - \frac{\Delta_k q_k q_k^T \Delta_k}{q_k^T \Delta_k q_k} + \xi_k \tau_k u_k u_k^T$ , με  $\tau_k = q_k^T \Delta_k q_k$   
 $u_k = \frac{p_k}{p_k^T q_k} - \frac{\Delta_k q_k}{\tau_k}$  και  $0 \leq \xi_k \leq 1$ . Θέτοντας,  $\xi_k = 0$ , για κάθε επανάληψη, θα έχουμε την  
 μέθοδο των *David- Fletcher- Powell*, όπου ο  $\Delta_{k+1} = \Delta_k + \frac{p_k p_k^T}{p_k^T q_k} - \frac{\Delta_k q_k q_k^T \Delta_k}{q_k^T \Delta_k q_k}$ .

Η όλη ιδέα της μεθόδου βασίζεται στον κατάλληλο υπολογισμό του  $\gamma_k$ , έτσι ώστε να ισχύει  $\frac{\partial f(x_{k+1})}{\partial \gamma} = 0$  και κατ' επέκταση να ισχύει  $\nabla f^T(x_k) d_k < \nabla f^T(x_{k+1}) d_k$ . Έτσι, η μέθοδος σχεδόν Newton για  $\gamma$  που προκύπτει από την ελαχιστοποίηση της  $h(\gamma) = f(x_k + \gamma_k * d_k)$ , εξασφαλίζει ότι αν ο  $\Delta_k$  είναι θετικά ορισμένος τότε και ο  $\Delta_{k+1}$  είναι θετικά ορισμένος.

- Αρχικό σημείο  $(-2, -1)$  και  $\varepsilon = 0.001$ .

Χρησιμοποιώντας το `quasi_newton_plot_x_y.m` και προσεγγίζοντας τον εσσιανό, με μόνο γνώση του `gradf`, παρατηρούμε:



Σχήμα 12: Σύγκλιση του ζεύγους  $(x, y)$  για ελαχιστοποιημένο  $\gamma$  [Μέθοδος Σχεδόν Newton]

Για σταθερό  $\gamma$  και για  $\gamma$  που υπακούει στον ευριστικό κανόνα, η σύγκλιση δεν είναι εφικτή. Αυτό οφείλεται στο γεγονός ότι δεν ισχύει η σχέση  $\nabla f^T(x_k) d_k < \nabla f^T(x_{k+1}) d_k$  σε κάθε επανάληψη,

δηλαδή δεν θα παράγω κάθε φορά θετικά ορισμένο  $\Delta_{k+1}$  και άρα η ιδέα της μεθόδου (να προσεγγιστεί η κατεύθυνση Newton) δεν γίνεται να εφαρμοστεί.

- Αρχικό σημείο  $(-3,3)$

Υπολογίζουμε  $f(-3,3) = 0$  και  $\text{grad}f(-3,3) = [0.1523, 0.1523] * 10^{-7}$ , όποτε το μέτρο του,  $|\text{grad}f(-2,-1)| = 2.1538 * 10^{-8}$ .

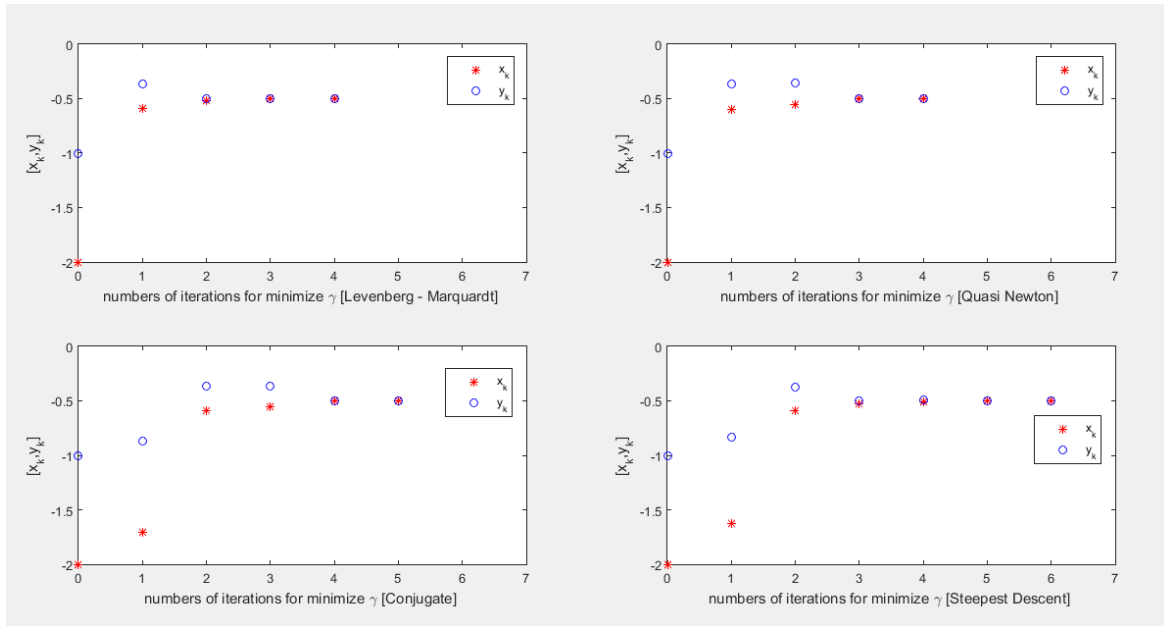
Για να καταφέρουμε τους αλγόριθμους των παραπάνω μεθόδων, έστω να ξεκινήσουν τις επαναλήψεις πρέπει να θέσουμε ένα αρκούντως μικρό  $\varepsilon$ . Ακόμα κι έτσι, οι αλγόριθμοι φτάνουν στο άνω όριο βημάτων (5000 επαναλήψεις) και παρατηρούμε ότι για αυτό το σημείο οι μέθοδοι δεν μπορούν να οδηγηθούν στο ελάχιστο, αφού δεν μπορούν να «ξεφύγουν» από την μικρή κλίση της περιοχής.

### Συγκρίσεις – Συμπεράσματα:

Παρατηρώντας τα παραπάνω διαγράμματα αλλά και της προηγούμενης εργασίας, βλέπουμε ότι η *σχεδόν Newton* και η μέθοδος των *Levenberg – Marquardt*, για ελαχιστοποιημένο  $\gamma$ , συγκλίνουν στον ίδιο αριθμό επαναλήψεων και χρειάζονται τον ίδιο αριθμό πράξεων  $O(n^2)$  για τον υπολογισμό του  $d_k$  και κατ' επέκταση του  $x_{k+1}$ .

Σε περίπτωση που ο υπολογισμός του εσσιανού είναι ανέφικτος ή περιλαμβάνει δύσκολους υπολογισμούς, προτιμάτε η *σχεδόν Newton*.

Ακολουθεί η *μέθοδος των συζυγών κλίσεων*, η οποία απαιτεί για τον υπολογισμό των συζυγών κατευθύνσεων και του  $x_{k+1}$ ,  $O(n)$  πράξεις. Η μέθοδος για τετραγωνικές συναρτήσεις εξασφαλίζει σύγκλιση σε το πολύ  $n$  βήματα, ωστόσο σε περίπτωση μη τετραγωνικών συναρτήσεων, όπως είδαμε, κάτι τέτοιο δεν ισχύει και για καλύτερα αποτελέσματα απαιτείται η συχνή επανεκκίνηση του αλγορίθμου. Η *μέθοδος των συζυγών κλίσεων*, προτιμάται συνήθως για μεγάλα  $n$ , όπου οι πράξεις που εκτελεί είναι σημαντικά λιγότερες σε σύγκριση με τις άλλες μεθόδους.



Σχήμα 13: Σύγκλιση του ζεύγους  $(x, y)$  για ελαχιστοποιημένο  $\gamma$  [Σύγκριση Μεθόδων]

Από το σχήμα 13, γίνεται αντιληπτό ότι ακόμα και για μη τετραγωνικές συναρτήσεις η μέθοδος των συζυγών κλίσεων είναι γρηγορότερη από την μέθοδο της Μέγιστης Καθόδου, επιβεβαιώνοντας το θεωρητικό υπόβαθρο και την ιδέα της μεθόδου των συζυγών κλίσεων, που ήθελε να επιταχύνει την σύγκλιση της Μέγιστης Καθόδου.

Σε ότι αφορά, τις υλοποιήσεις για σταθερό και μεταβλητό  $\gamma$  τα αποτελέσματα δεν είναι συγκρίσιμα, καθώς σε κάθε μέθοδο επιλέχθηκε το κατάλληλο  $\gamma$  έτσι ώστε να επιτυγχάνεται η σύγκλιση, μετά από διερεύνηση. Επιπλέον, στην μέθοδο σχεδόν *Newton*, δεν βρέθηκαν κατάλληλα  $\gamma$ , για τις δύο αυτές υλοποιήσεις, που θα εξασφάλιζαν σύγκλιση. Τέλος, για διαφορετικό αρχικό σημείο πρέπει να επαναληφθεί η διερεύνηση των τιμών του  $\gamma$ , το οποίο είναι και λογικό καθώς η ορθή λειτουργία των μεθόδων έχει επιβεβαιωθεί για  $\gamma$  που προκύπτει από την εσωτερική ελαχιστοποίηση της  $h(\gamma) = f(x_k + \gamma_k * d_k)$ .

Από την υλοποίηση των διαφόρων μεθόδων διαπιστώθηκε, ότι υπάρχουν κοινά σημεία και όμοια συμπεριφορά. Πιο συγκεκριμένα, η μέθοδος των *Levenberg – Marquardt* για μικρό  $\mu_k$  συμπεριφέρεται σαν την *Newton* και για μεγάλες τιμές του  $\mu_k$ , λειτουργεί όπως η μέθοδος της Μέγιστης Καθόδου. Τέλος, η μέθοδος των Συζυγών κλίσεων στην πρώτη επανάληψη «εκμεταλλεύεται» την βασική ιδέα της λειτουργίας της Μέγιστης Καθόδου.