Τεχνικές Βελτιστοποίησης
Εργασία 2η

Παπουτσή Νικολέτα **ΑΕΜ : 10858**

Ζητούμενο: Ελαχιστοποίηση μιας δοσμένης συνάρτησης $f : \mathbb{R}^n \to \mathbb{R}$ χωρίς περιορισμούς με τη χρήση των παρακάτω μεθόδων:

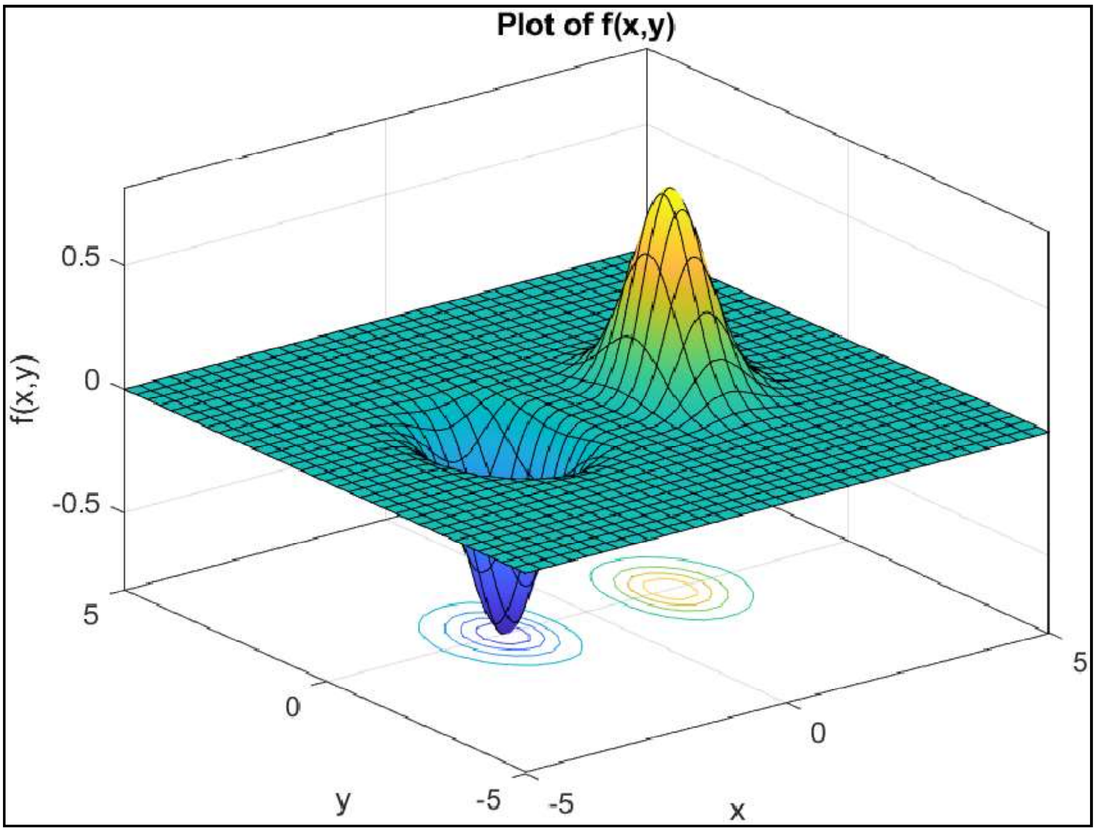
1. **Μέθοδος Μέγιστης Καθόδου (steepest descent)**
2. **Μέθοδος Newton**
3. **Μέθοδος Levenberg – Marquardt**

Η αντικειμενική συνάρτηση της οποίας ψάχνουμε το ελάχιστο είναι :

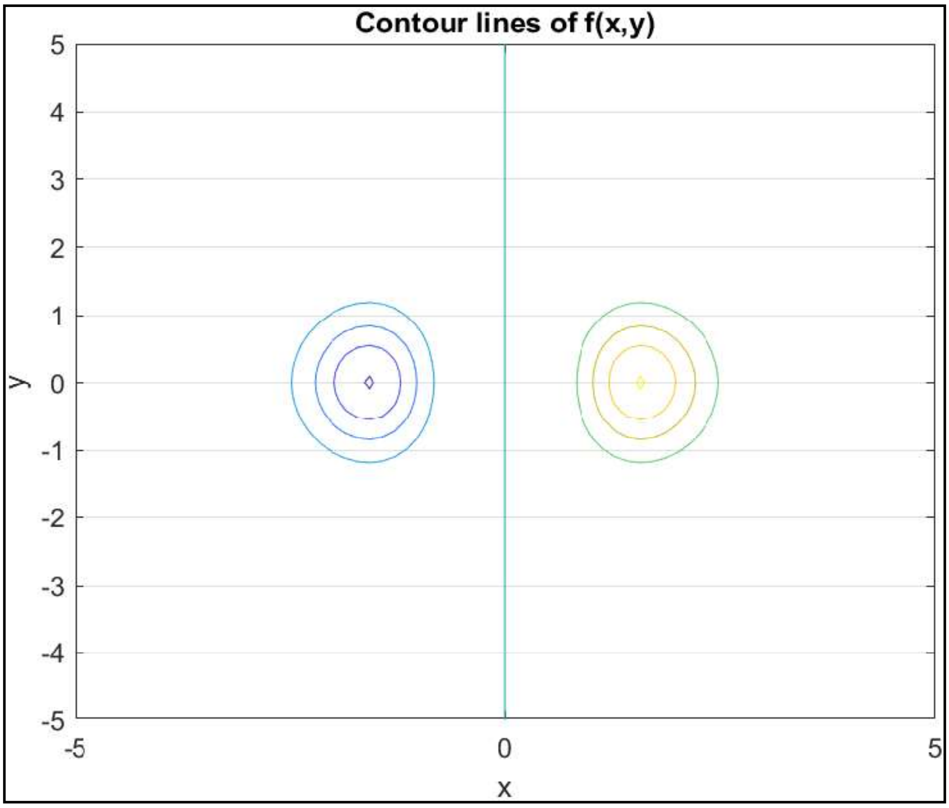
$$f(x, y) = x^5 e^{-x^2-y^2}$$

Θέμα 1ο

Παρακάτω παρουσιάζεται η μορφή της f στον τρισδιάστατο χώρο και οι ισοβαρείς καμπύλες της.



Σχήμα 1. Γραφική παράσταση της f



Σχήμα 2. Ισοβαρείς καμπύλες της f

Όπως, παρατηρείται παραπάνω η f ελαχιστοποιείται , όταν το x παίρνει αρνητικές τιμές και το y βρίσκεται κοντά στο μηδέν. Επιπλέον, παρουσιάζει τοπικό μέγιστο για x θετικά και y επίσης κοντά στο 0, ενώ στην αρχή των αξόνων (0,0) παρουσιάζει τοπικό ελάχιστο.

<u>Θέμα 2ο : Μέθοδος Μέγιστης Καθόδου</u>

Η μέθοδος μέγιστης καθόδου, όπως και οι άλλες μέθοδοι κλίσης , βασίζονται στην ιδέα της επαναληπτικής καθόδου . Η κατεύθυνση αναζήτησης για κάθε επανάληψη υπολογίζεται ως $d_k=-\Delta_k\nabla f(x_k)$, όπου Δk=I (ο μοναδιαίος πίνακας). Ως σταθερά τερματισμού για όλες τις μεθόδους επιλέχθηκε **ε = 0.0001** .

Παρακάτω μελετάται η συμπεριφορά της μεθόδου μέγιστης καθόδου για διαφορετικά αρχικά σημεία $(\mathbf{x_0}, \mathbf{y_0})$ και διαφορετικές επιλογές του βήματος $\boldsymbol{\gamma_k}$. Συγκεκριμένα, εξετάζουμε τη συμπεριφορά της μεθόδου για :

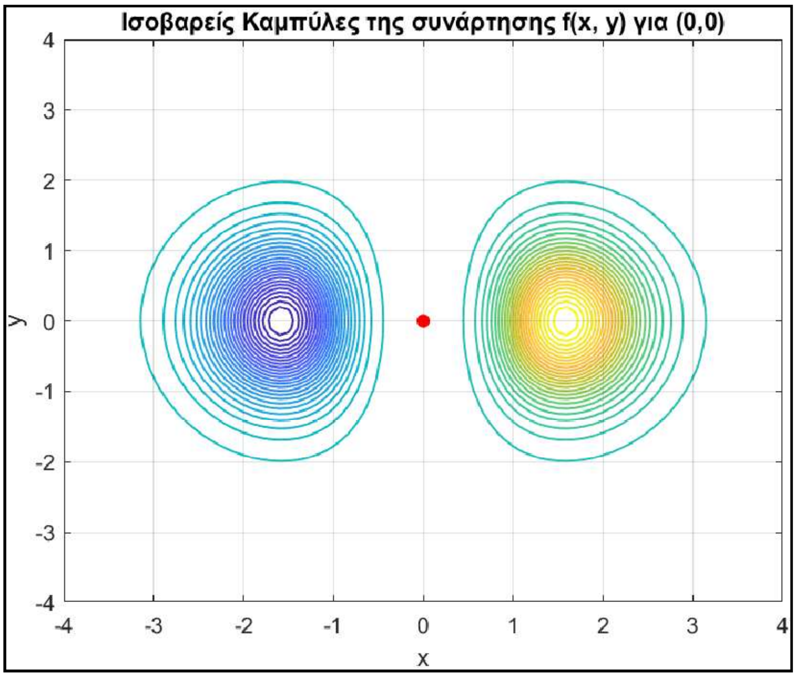
- $(\mathbf{x_0}, \mathbf{y_0}) = (0 , 0)$
- $(\mathbf{x_0}, \mathbf{y_0}) = (-1 , 1)$
- $(\mathbf{x_0}, \mathbf{y_0}) = (1 , -1)$

και για τρεις διαφορετικές επιλογές του $\boldsymbol{\gamma_k}$:

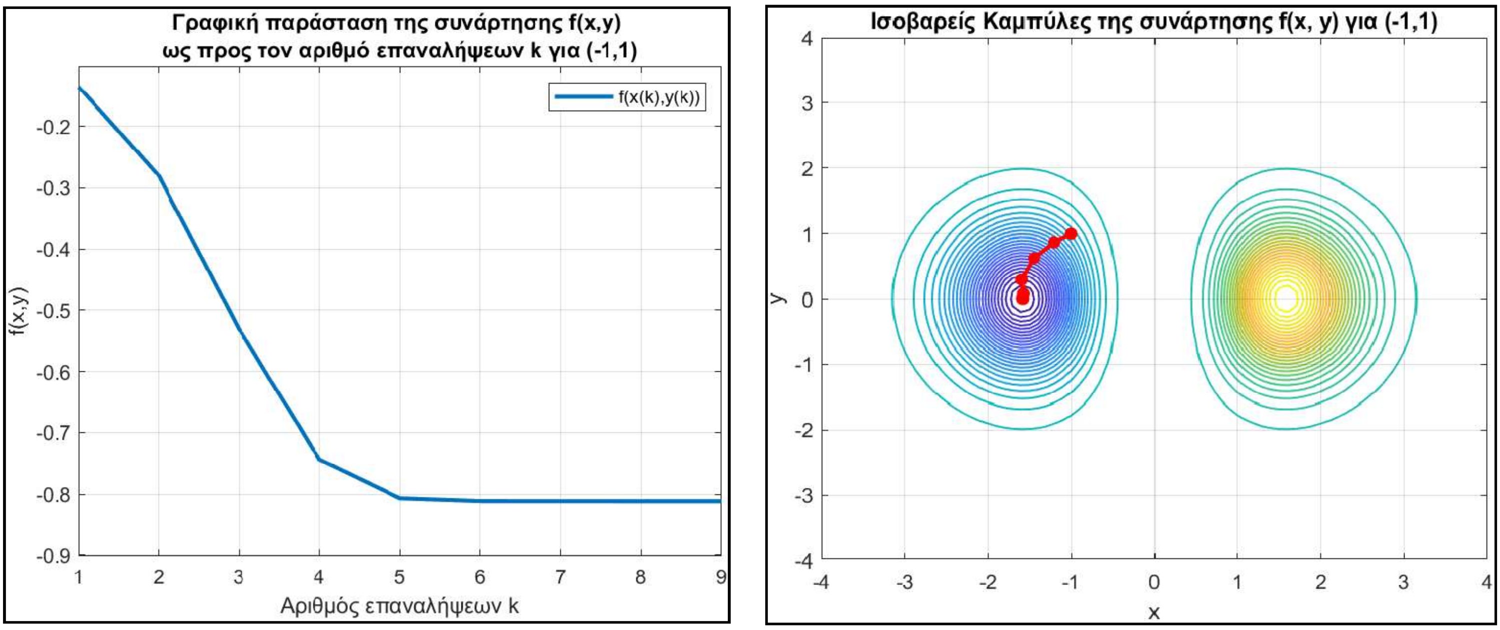
- Σταθερό $\boldsymbol{\gamma_k}$ = 0.5
- Τέτοιο ώστε να ελαχιστοποιεί την $\mathbf{f\ (x_k + \gamma_k\, d_k\ )}$
- Βάσει του κανόνα **Armijo**

1) <u>Σταθερό $\boldsymbol{\gamma_k}$ = 0.5</u>

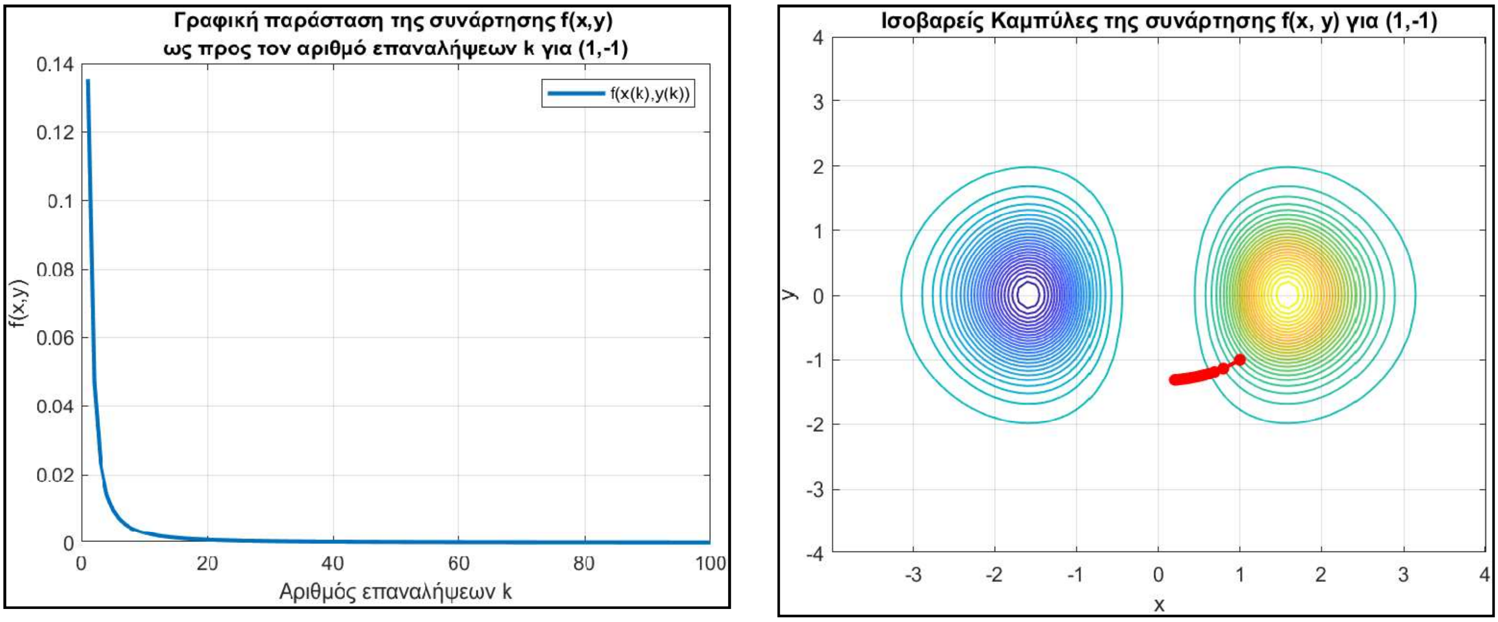
- Για το σημείο (0, 0) ο αλγόριθμος τερματίζει από την πρώτη επανάληψη. Αυτό συμβαίνει επειδή η κλίση της συνάρτησης ∇f(0,0) είναι μηδενική, γεγονός που υποδεικνύει ότι η f(x,y) παρουσιάζει τοπικό ελάχιστο στο σημείο αυτό. Συνεπώς, ο αλγόριθμος εγκλωβίζεται στο σημείο (0,0) για κάθε βήμα.



Σχήμα 3. Πορεία σύγκλισης για σταθερό βήμα γ_k = 0.5 στο σημείο (0,0) στις ισοβαρείς καμπύλες



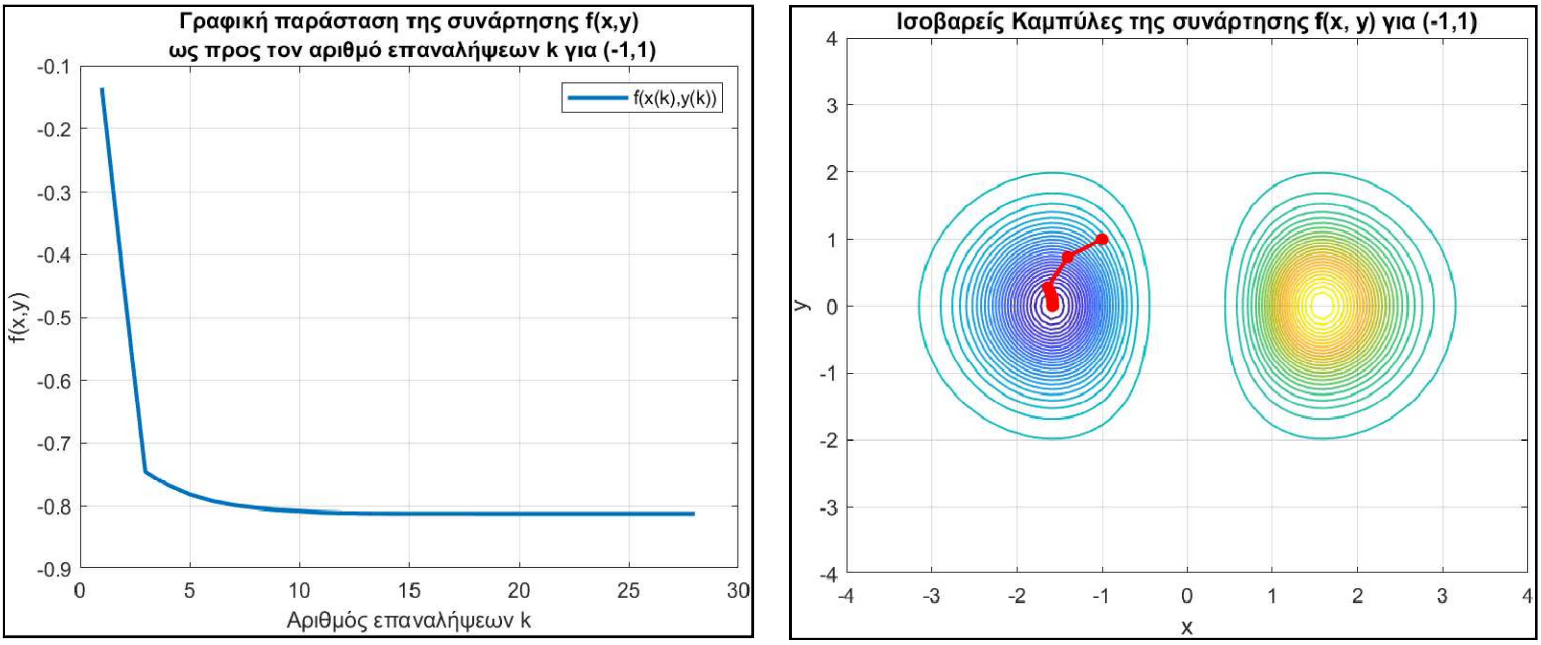
Σχήμα 4.Τιμή της f(x,y) ως προς τις επαναλήψεις k για σταθερό βήμα $\gamma_k = 0.5$ και πορεία σύγκλισης στις ισοβαρείς καμπύλες για το σημείο (-1,1)



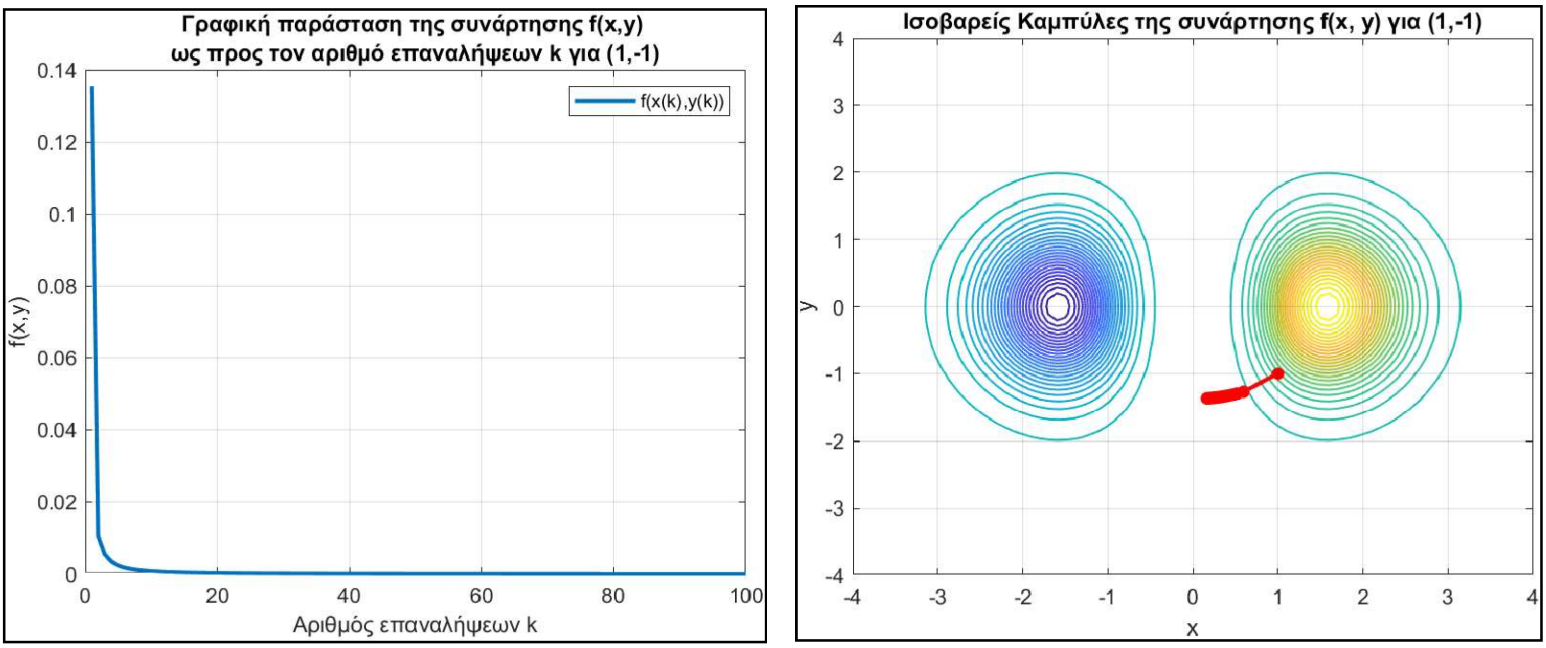
Σχήμα 5.Τιμή της f(x,y) ως προς τις επαναλήψεις k για σταθερό βήμα $\gamma_k = 0.5$ και πορεία σύγκλισης στις ισοβαρείς καμπύλες για το σημείο (1,-1)

Αρχικό Σημείο	**Επαναλήψεις k**	**Ακρότατο**
(0,0)	1	(0.0 , 0.0)
(-1,1)	10	(-1.581 , 0.00)
(1,-1)	101	(0.199,-1.313)

- Για την εύρεση του βέλτιστου $\boldsymbol{\gamma_k}$,ώστε να ελαχιστοποιεί την $f\,(x_k + \gamma_k\, d_k)$ χρησιμοποιήθηκε η μέθοδος του Χρυσού Τομέα από την προηγούμενη εργασία.

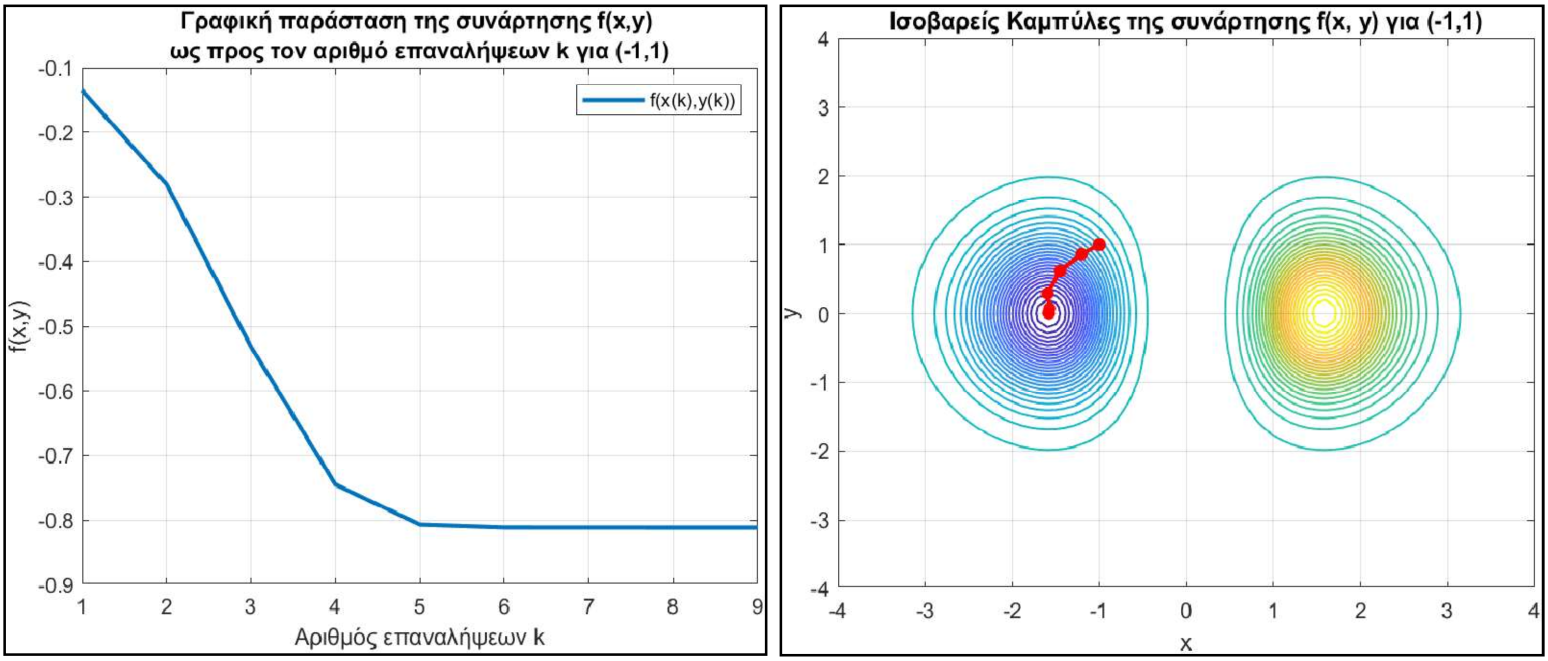


Σχήμα 6.Τιμή της f(x,y) ως προς τις επαναλήψεις k για βήμα γ_k που ελαχιστοποιεί την $f(x_k+\gamma_k d_k)$ και πορεία σύγκλισης στις ισοβαρείς καμπύλες για το σημείο (-1,1)

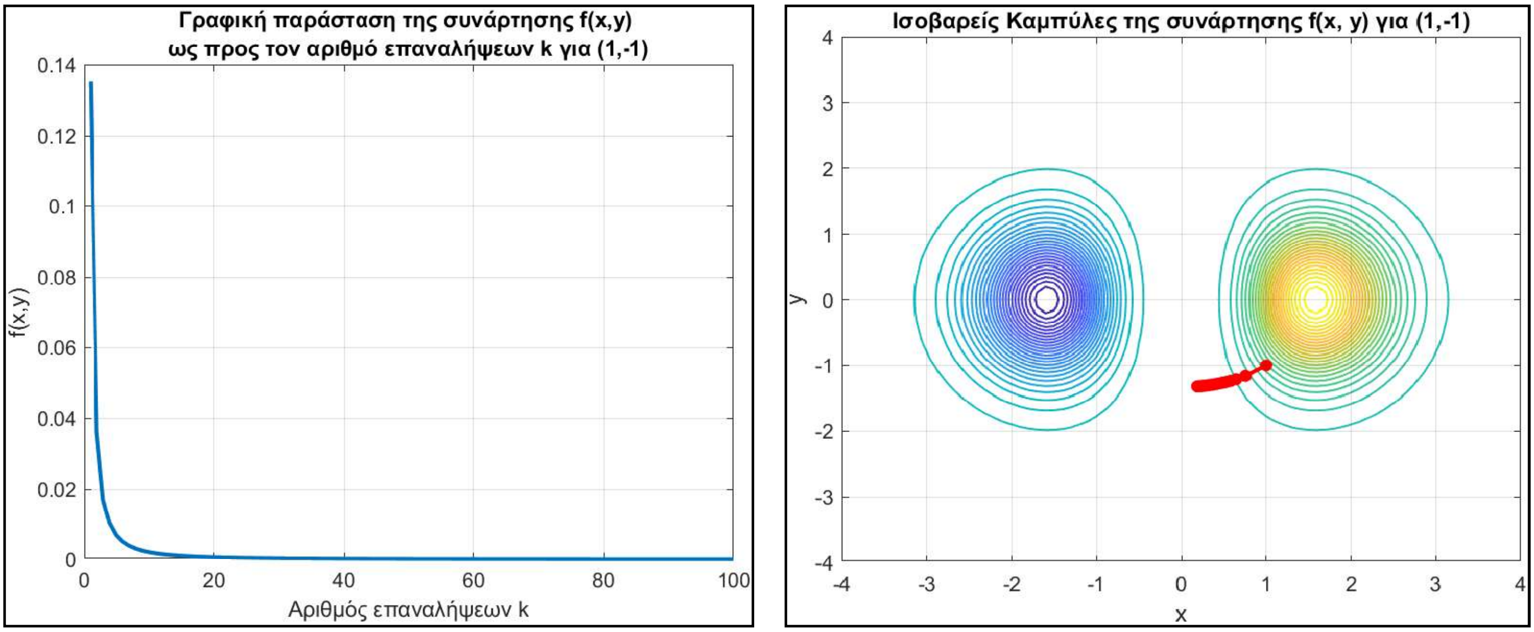


Σχήμα 7.Τιμή της f(x,y) ως προς τις επαναλήψεις k για βήμα γ_k που ελαχιστοποιεί την $f(x_k+\gamma_k d_k)$ και πορεία σύγκλισης στις ισοβαρείς καμπύλες για το σημείο (1,-1)

o Επιλογή βήματος $\boldsymbol{\gamma_k}$ μέσω του Κανόνα Armijo



Σχήμα 8.Τιμή της f(x,y) ως προς τις επαναλήψεις k για επιλογή βήματος γ_k μέσω του κανόνα Armijo και πορεία σύγκλισης στις ισοβαρείς καμπύλες για το σημείο (-1,1)



Σχήμα 9.Τιμή της f(x,y) ως προς τις επαναλήψεις k για επιλογή βήματος γ_k μέσω του κανόνα Armijo και πορεία σύγκλισης στις ισοβαρείς καμπύλες για το σημείο (1,-1)

- **Παρατηρήσεις**

-γ_k που ελαχιστοποιεί την $f(x_k+\gamma_k d_k)$

Αρχικό Σημείο	Επαναλήψεις k	Ακρότατο
(0,0)	1	(0.0 , 0.0)
(-1,1)	29	(-1.581 , 0.00)
(1,-1)	101	(0.1643,-1.367)

-γ_k μέσω του κανόνα Armijo

Στις προσομοιώσεις έχουν επιλεχθεί α = 0.0001, β = 0.1 και αρχικό βήμα s = 0.6 . Να σημειωθεί ότι ανάλογα την επιλογή των παραμέτρων α , β , s ο αλγόριθμος μπορεί να συγκλίνει στο ολικό ελάχιστο και για τα δύο σημεία αλλά αυξάνονται οι επαναλήψεις.

Αρχικό Σημείο	Επαναλήψεις k	Ακρότατο
(0,0)	1	(0.0 , 0.0)
(-1,1)	11	(-1.581 , 0.00)
(1,-1)	102	(0.1997,-1.3134)

Σύγκριση με βάση τα αρχικά σημεία

- Για το (0,0) όπως αναφέρθηκε παραπάνω είναι τοπικό ακρότατο οπότε ο αλγόριθμος τερματίζει από τη πρώτη επανάληψη .
- Για το (-1,1) ο αλγόριθμος συγκλίνει στο ολικό ελάχιστο (-1.581,0) και με τις τρεις επιλογές του γ_κ.
- Για το (1,-1) ο αλγόριθμος εγκλωβίζεται και στις τρεις περιπτώσεις σε τοπικό ακρότατο.

Σύγκριση με βάση την επιλογή του βήματος

Αρχικό Σημείο	γκ = 0.5	γ_k που ελαχιστοποιεί την $f(x_k+\gamma_k d_k)$	γ_k μέσω του κανόνα Armijo
(0,0)	1	1	1
(-1,1)	10	29	11
(1,-1)	101	101	102

<u>Θέμα 3ο : Μέθοδος Newton</u>

Η μέθοδος Newton αναζητά το σημείο ελαχίστου στην κατεύθυνση :

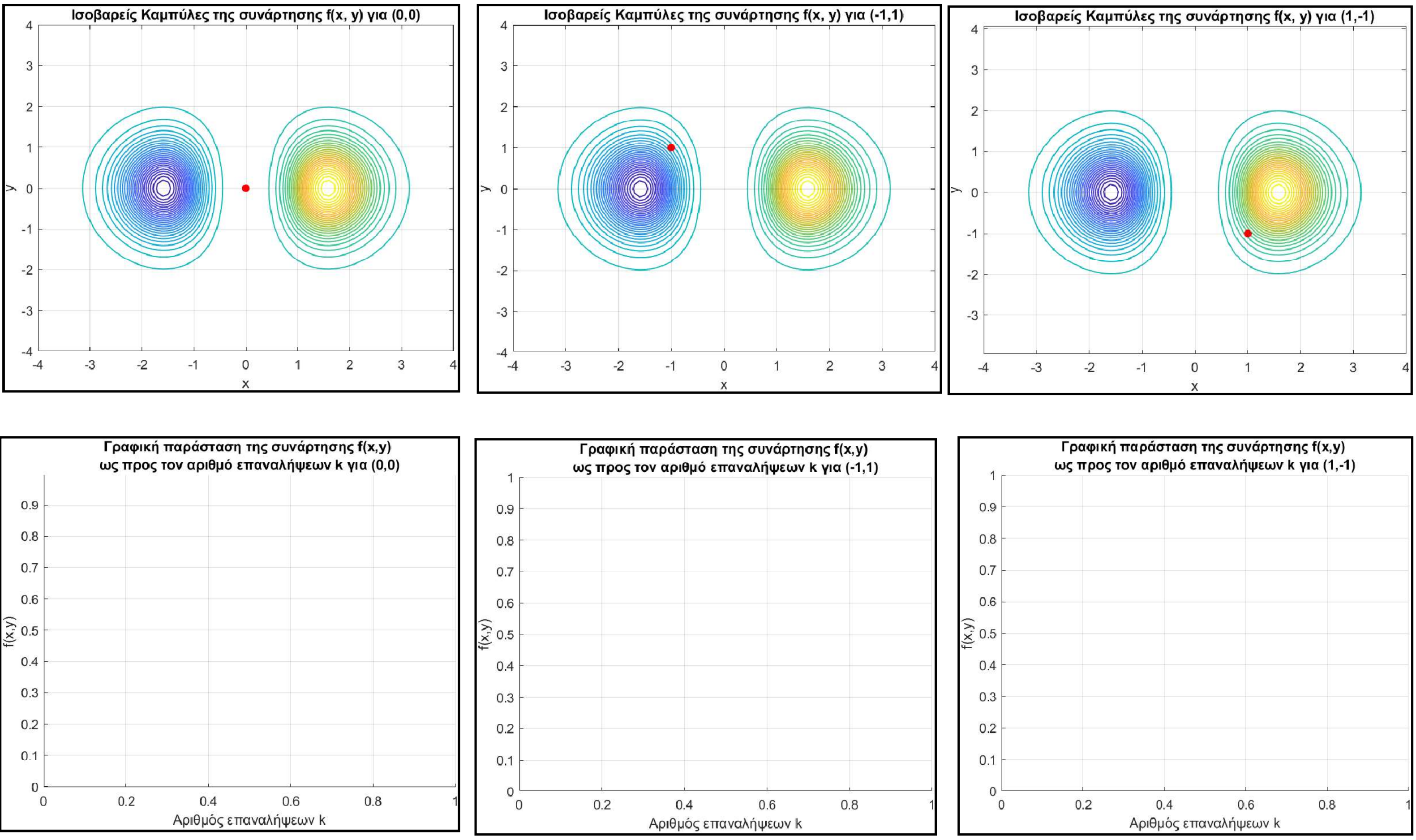
$$d_k = - [\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$$

με την προϋπόθεση ότι ο Εσσιανός πίνακας $\nabla^2 f(x_k)$ είναι θετικά ορισμένος.

Συνεπώς, εάν ο εσσιανός δεν είναι θετικά ορισμένος, ο αλγόριθμος της μεθόδου Newton θα τερματίσει στη πρώτη επανάληψη, επειδή δεν μπορεί να υπολογίσει την κατεύθυνση του βήματος.

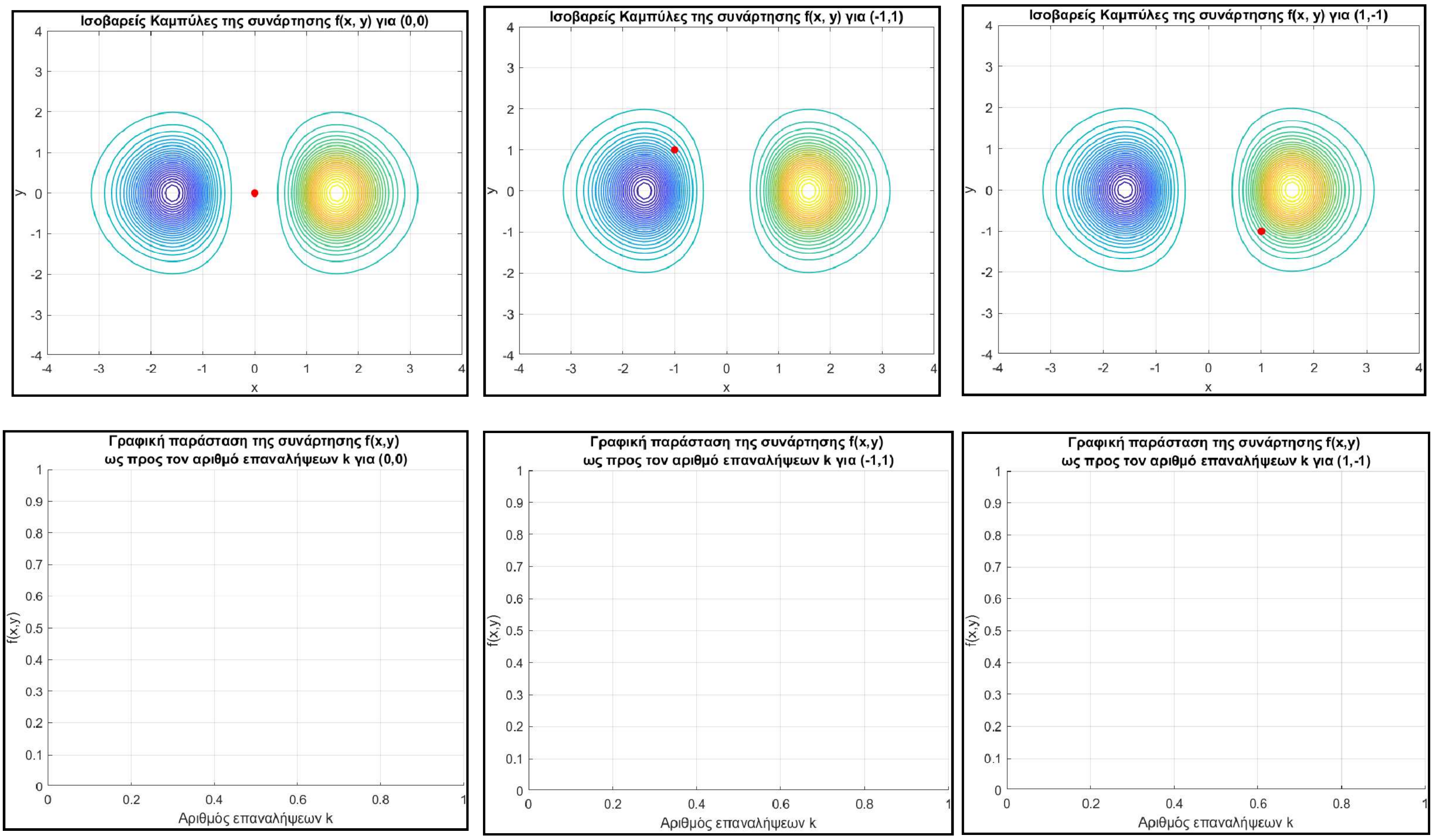
Όπως παρατηρείται παρακάτω, για τα δοσμένα αρχικά σημεία ο αλγόριθμος τερματίζει από την πρώτη επανάληψη καθώς ο εσσιανός δεν είναι θετικά ορισμένος από την αρχή.

- $\boldsymbol{\gamma_\kappa} = 0.5$



Σχήμα 10.Τιμή της f(x,y) ως προς τις επαναλήψεις k για σταθερό βήμα γ_k = 0.5 και πορεία σύγκλισης στις ισοβαρείς καμπύλες για τα σημεία {(0,0) , (-1,1) , (1,-1)}

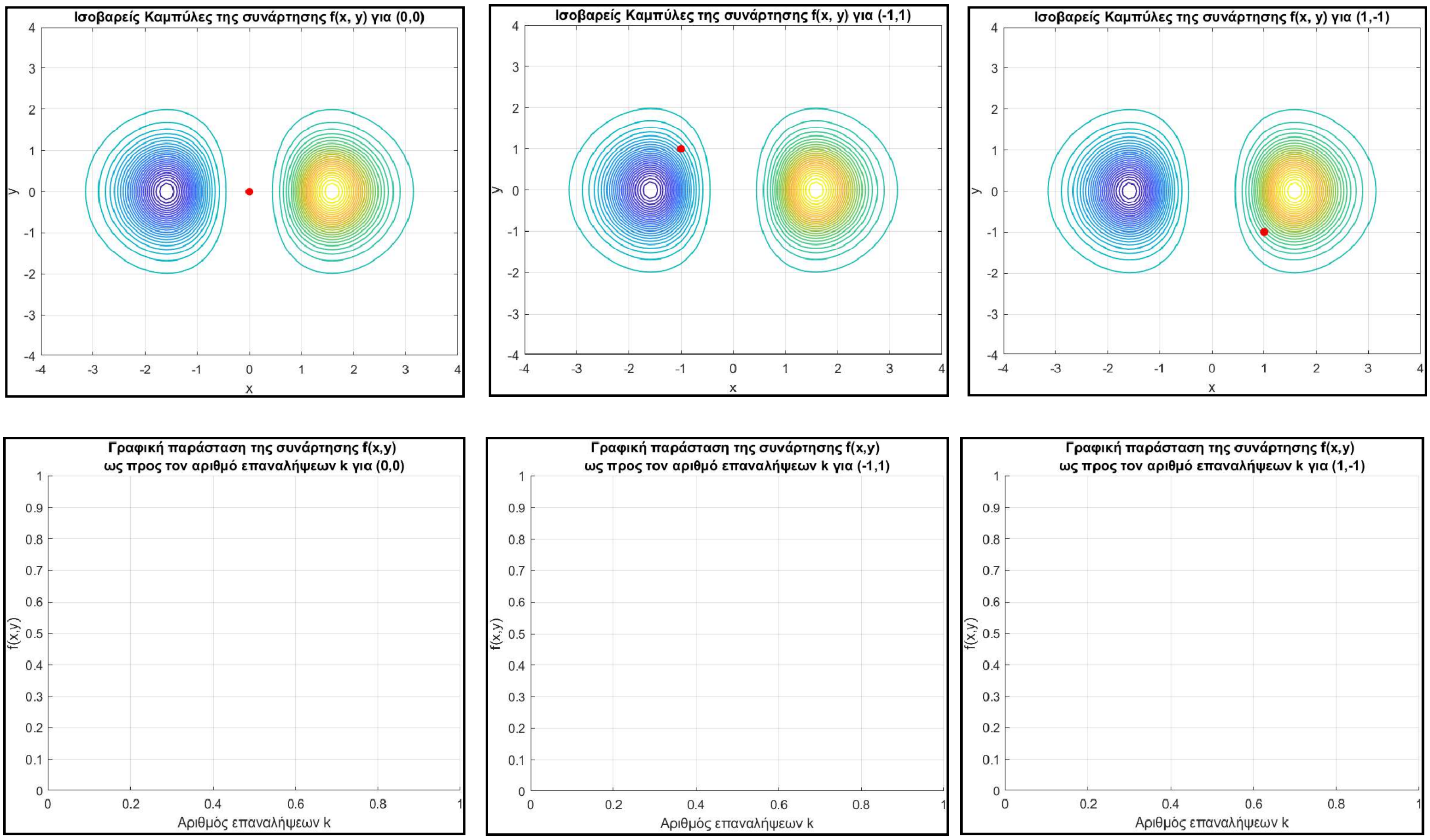
- Για την εύρεση του βέλτιστου $\boldsymbol{\gamma_k}$,ώστε να ελαχιστοποιεί την $f\,(x_k + \gamma_k\, d_k)$ χρησιμοποιήθηκε η μέθοδος του Χρυσού Τομέα από την προηγούμενη εργασία.



Σχήμα 11.Τιμή της f(x,y) ως προς τις επαναλήψεις k για βήμα γ_k που ελαχιστοποιεί την $f(x_k+\gamma_k d_k)$ και πορεία σύγκλισης στις ισοβαρείς καμπύλες για τα σημεία { (0,0), (-1,1) , (1,-1) }

Αρχικό Σημείο	Επαναλήψεις k	Εσσιανός
(0,0)	1	0 0
(-1,1)	1	-0.2707 -0.8120 -0.8120 -0.2707
(1,-1)	1	0.2707 0.8120 0.8120 0.2707

- Επιλογή βήματος $\boldsymbol{\gamma_k}$ μέσω του Κανόνα Armijo



Σχήμα 12.Τιμή της f(x,y) ως προς τις επαναλήψεις k για επιλογή βήματος γ_k μέσω του κανόνα Armijo και πορεία σύγκλισης στις ισοβαρείς καμπύλες για τα σημεία {(0,0) , (-1,1) , (1,-1)}

- **Παρατηρήσεις**

Για κάθε αρχικό σημείο $\{(0,\ 0),\ (-1,\ 1),\ (1,\ -1)\}$ ο αλγόριθμος τερματίζεται στη πρώτη επανάληψη , αφού ο εσσιανός που προκύπτει δεν είναι θετικά ορισμένος.

Σύγκριση με βάση τα αρχικά σημεία

Η επιλογή του αρχικού σημείου είναι σημαντικός παράγοντας στη λειτουργία της μεθόδου Newton. Όπως, παρατηρήθηκε για τα αρχικά σημεία που μας δόθηκαν ο εσσιανός πίνακας δεν είναι θετικά ορισμένος , με αποτέλεσμα ο αλγόριθμος να τερματίζει. Έτσι, η σωστή επιλογή αρχικού σημείου μπορεί να οδηγήσει πιθανόν σε ταχύτερη σύγκλιση και , ενδεχομένως , να επιτρέψει στον αλγόριθμο να συγκλίνει σε τοπικό ή ολικό ελάχιστο.

<u>Θέμα 4ο : Μέθοδος Levenberg – Marquardt</u>

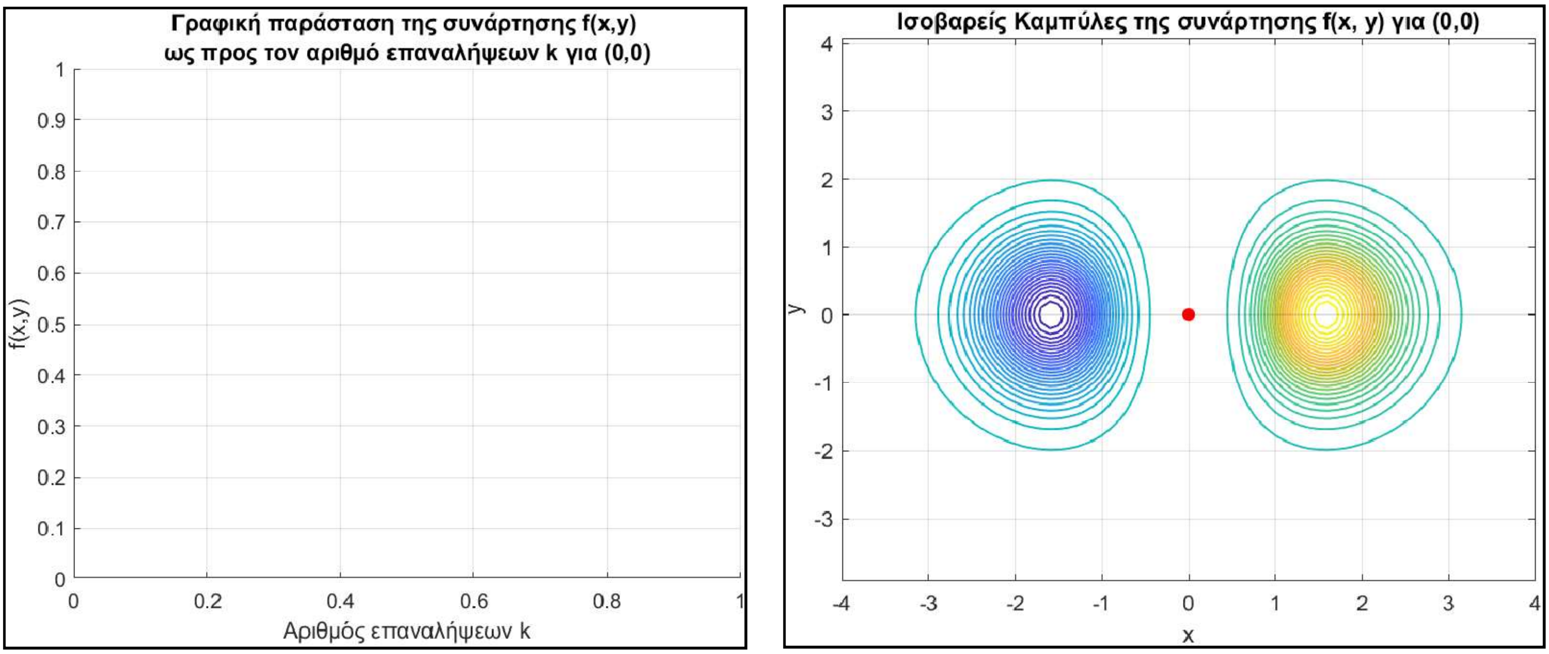
Η μέθοδος Levenberg – Marquardt είναι ο τροποποιημένος αλγόριθμος της μεθόδου Newton.

Η κύρια διαφορά είναι ότι δεν χρησιμοποιείται ο εσσιανός πίνακας απευθείας, αλλά εισάγεται μια θετική παράμετρος μ_κ τέτοια ώστε , ο $\nabla^2 f(x_k) + \mu_\kappa I$ να είναι θετικά ορισμένος.

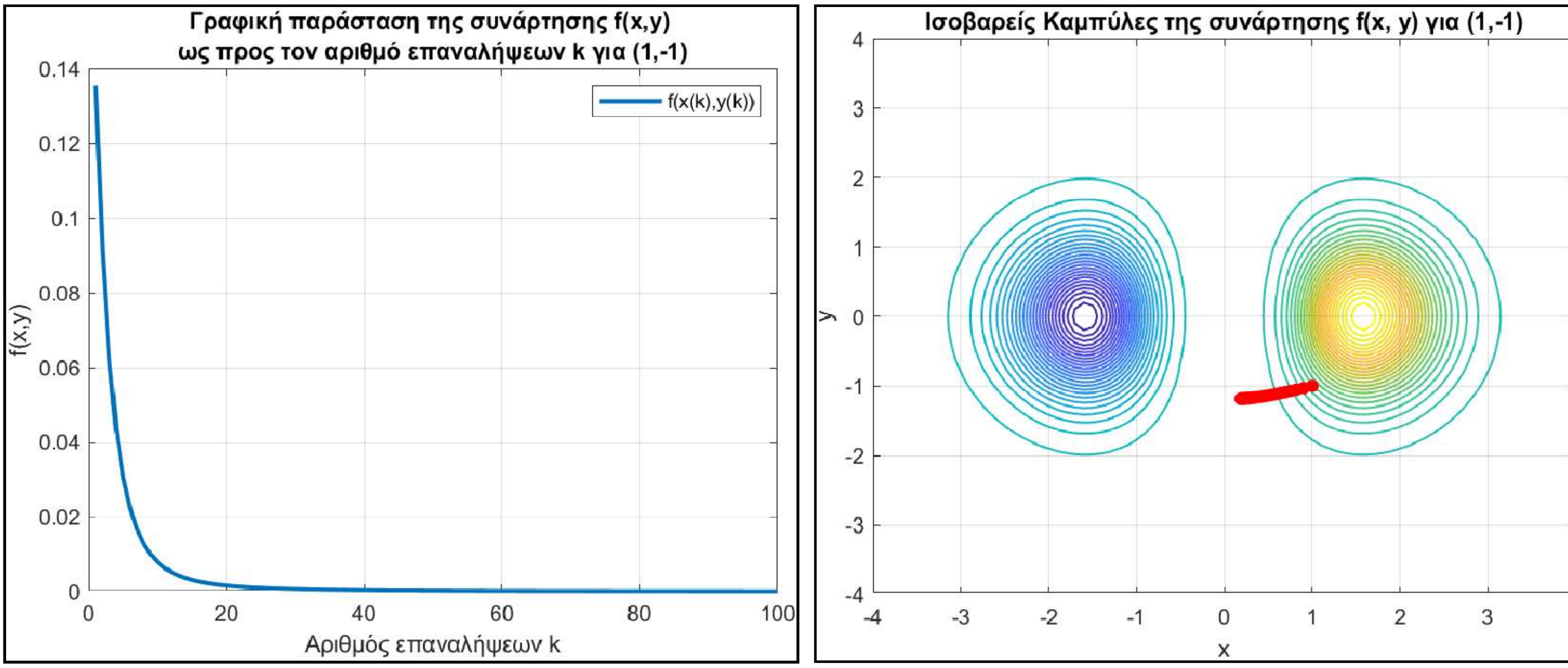
Αναζητά το σημείο ελαχίστου στην κατεύθυνση :

$$d_k = -[\nabla^2 f(x_k) + \mu_\kappa I]^{-1} \nabla f(x_k)$$

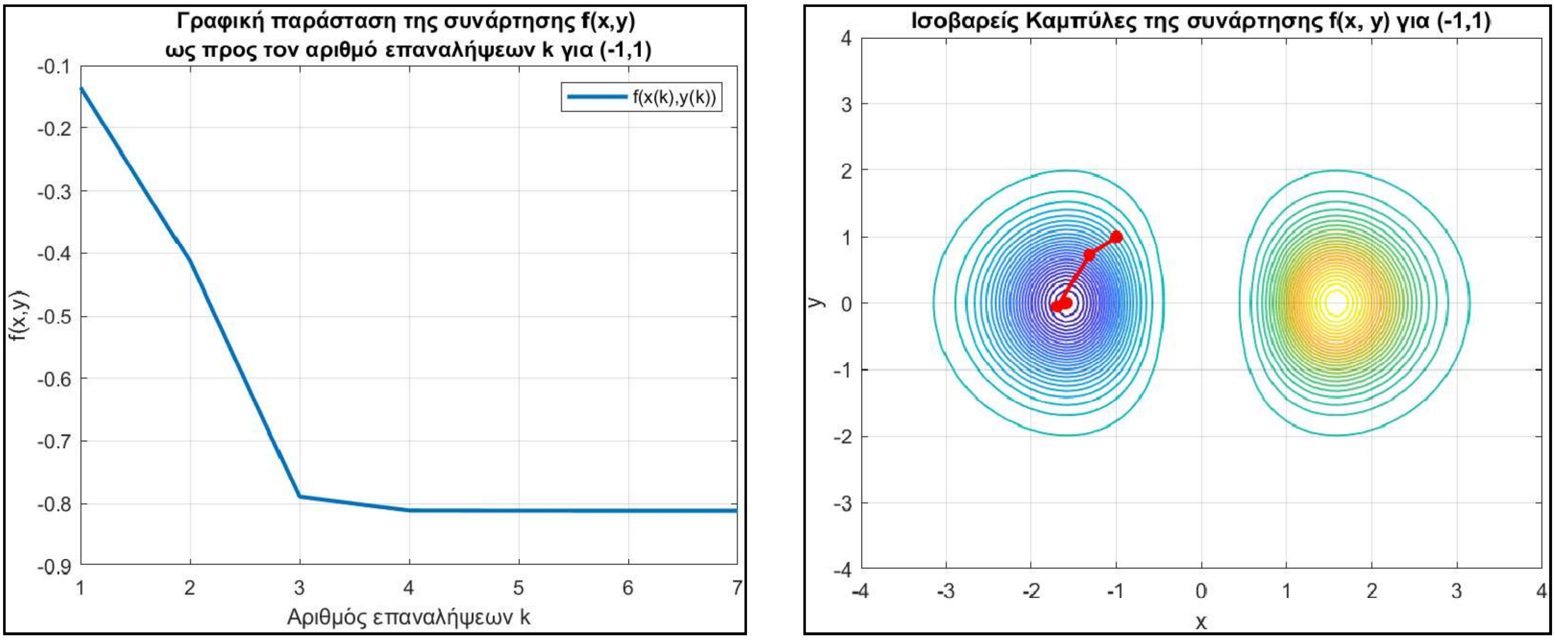
- $\boldsymbol{\gamma_\kappa} = 0.8$



Σχήμα 13.Τιμή της f(x,y) ως προς τις επαναλήψεις k για σταθερό βήμα $\gamma_k = 0.5$ και πορεία σύγκλισης στις ισοβαρείς καμπύλες για το σημείο (0,0)



Σχήμα 14.Τιμή της f(x,y) ως προς τις επαναλήψεις k για σταθερό βήμα $\gamma_k = 0.5$ και πορεία σύγκλισης στις ισοβαρείς καμπύλες για το σημείο (-1,1)

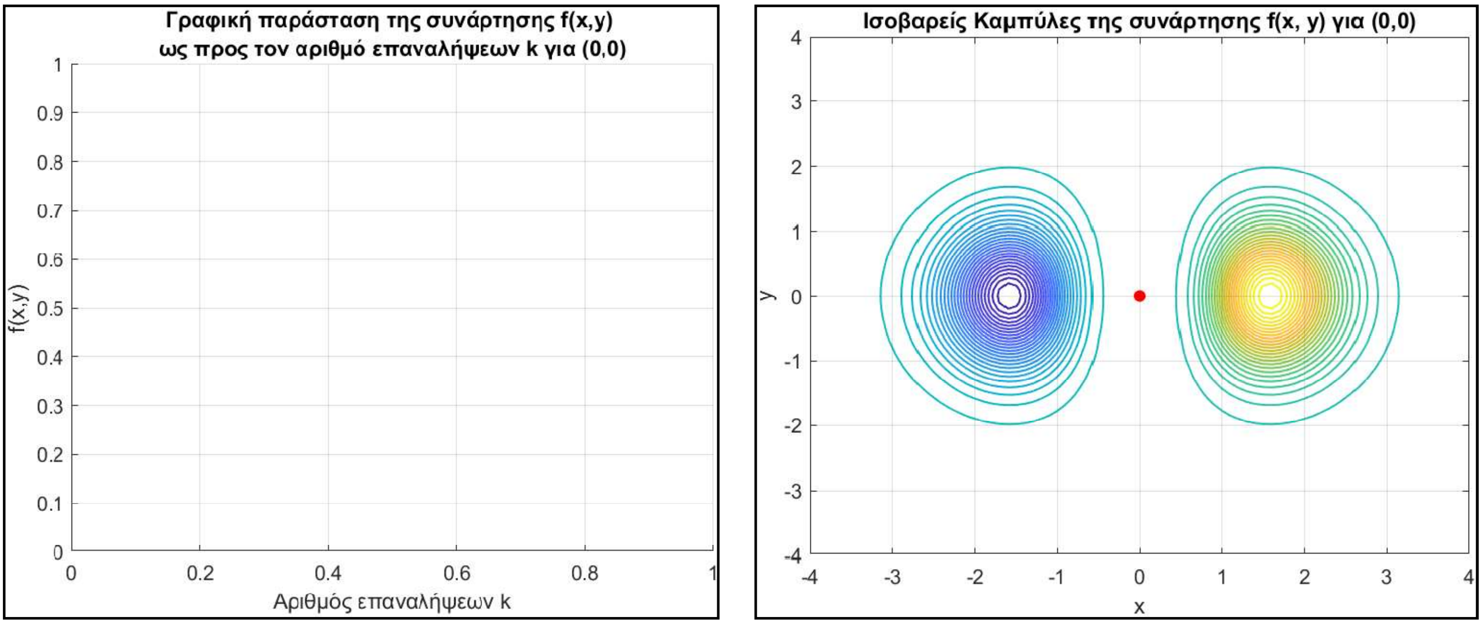


Σχήμα 15.Τιμή της f(x,y) ως προς τις επαναλήψεις k για σταθερό βήμα $\gamma_k = 0.5$ και πορεία σύγκλισης στις ισοβαρείς καμπύλες για το σημείο (1,-1)

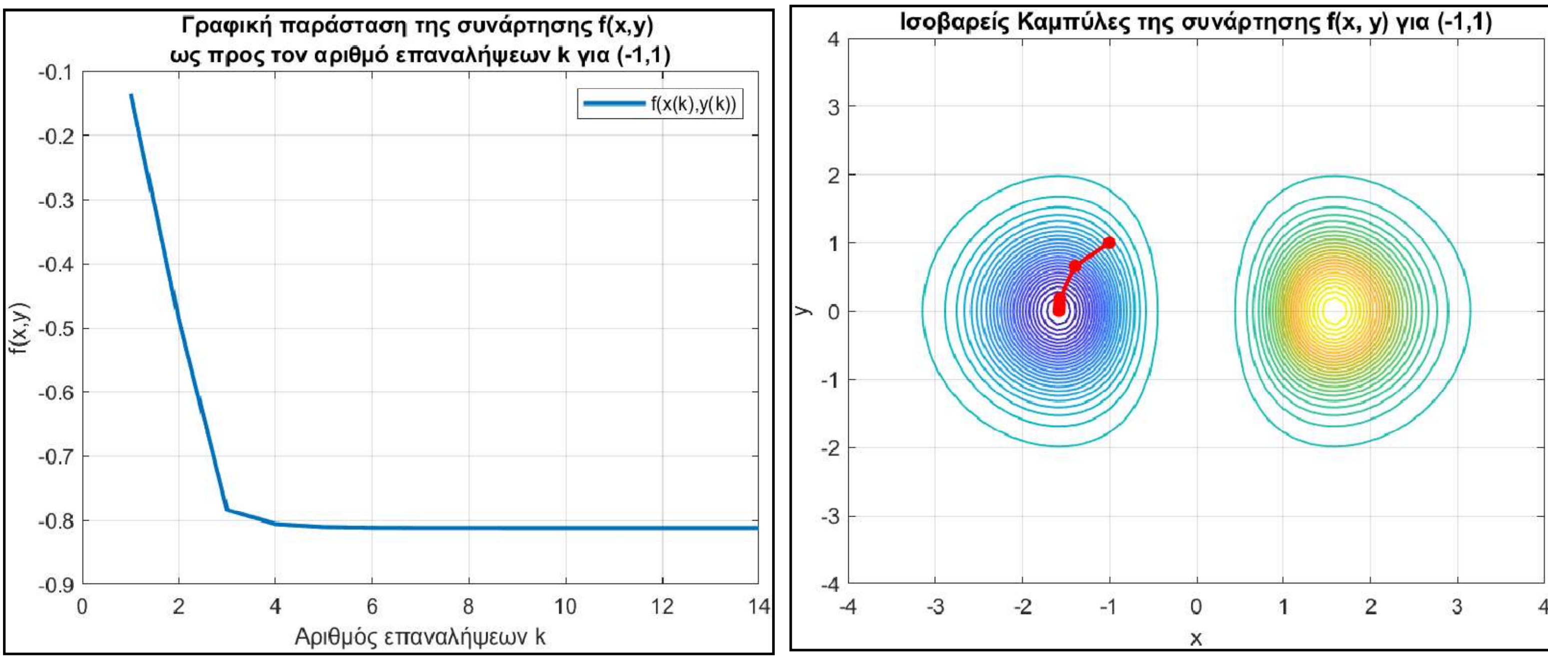
Συμπεραίνεται ότι για το (-1,1) υπάρχει σύγκλιση στο ολικό ελάχιστο για μικρό αριθμό επαναλήψεων k. Αντίθετα, για το $(1, -1)$ ο αλγόριθμος συγκλίνει σε τοπικό ακρότατο ύστερα από πολλές επαναλήψεις.

Αρχικό Σημείο	Επαναλήψεις k	Ακρότατο
(0,0)	2	(0.0 , 0.0)
(-1,1)	9	(-1.5812 , 0.00)
(1,-1)	102	(0.1478,-1.1144)

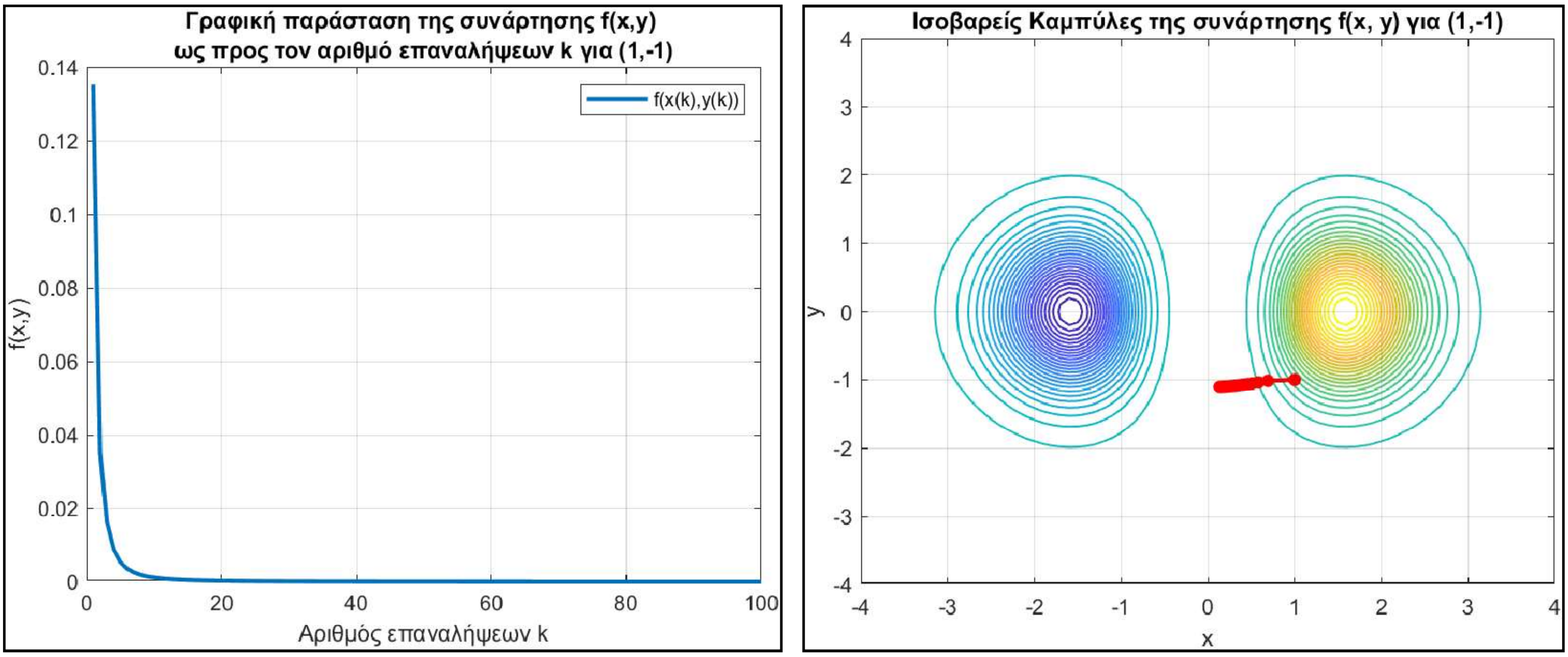
- Για την εύρεση του βέλτιστου $\boldsymbol{\gamma_k}$, ώστε να ελαχιστοποιεί την f (x_k + γ_k d_k) χρησιμοποιήθηκε η μέθοδος του Χρυσού Τομέα από την προηγούμενη εργασία.



Σχήμα 16.Τιμή της f(x,y) ως προς τις επαναλήψεις k για βήμα γ_k που ελαχιστοποιεί την f(x_k+$\gamma_k d_k$) και πορεία σύγκλισης στις ισοβαρείς καμπύλες για το σημείο (0,0)

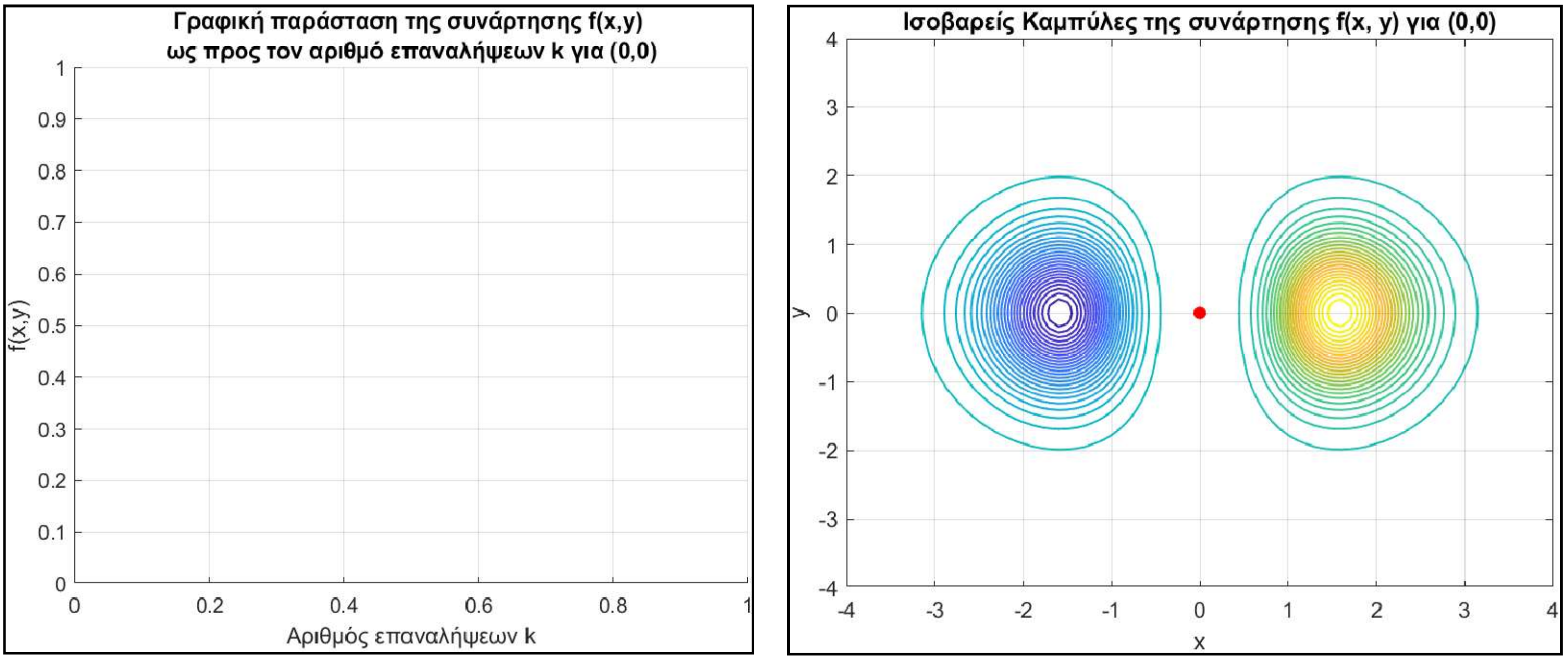


Σχήμα 17.Τιμή της f(x,y) ως προς τις επαναλήψεις k για βήμα γ_k που ελαχιστοποιεί την f(x_k+$\gamma_k d_k$) και πορεία σύγκλισης στις ισοβαρείς καμπύλες για το σημείο (-1,1)

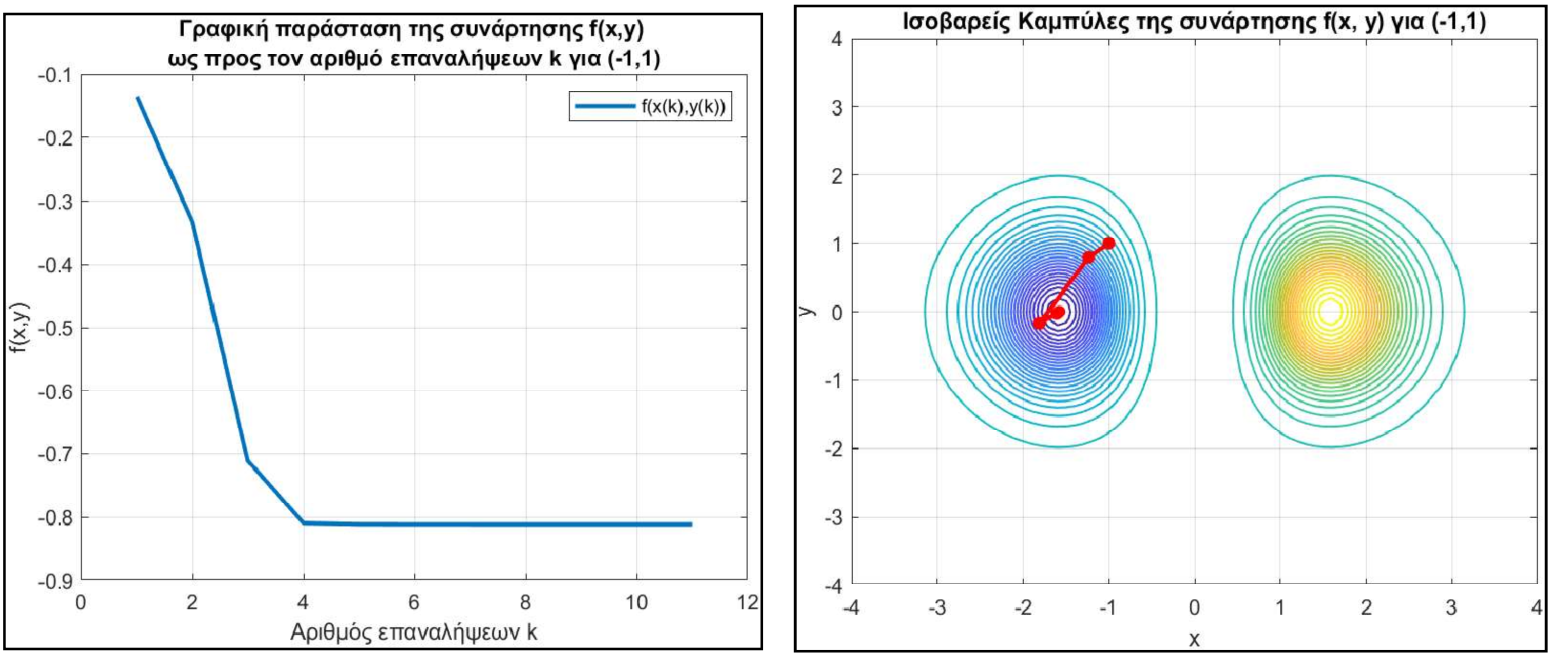


Σχήμα 18.Τιμή της f(x,y) ως προς τις επαναλήψεις k για βήμα γ_k που ελαχιστοποιεί την $f(x_k+\gamma_k d_k)$ και πορεία σύγκλισης στις ισοβαρείς καμπύλες για το σημείο (1,-1)

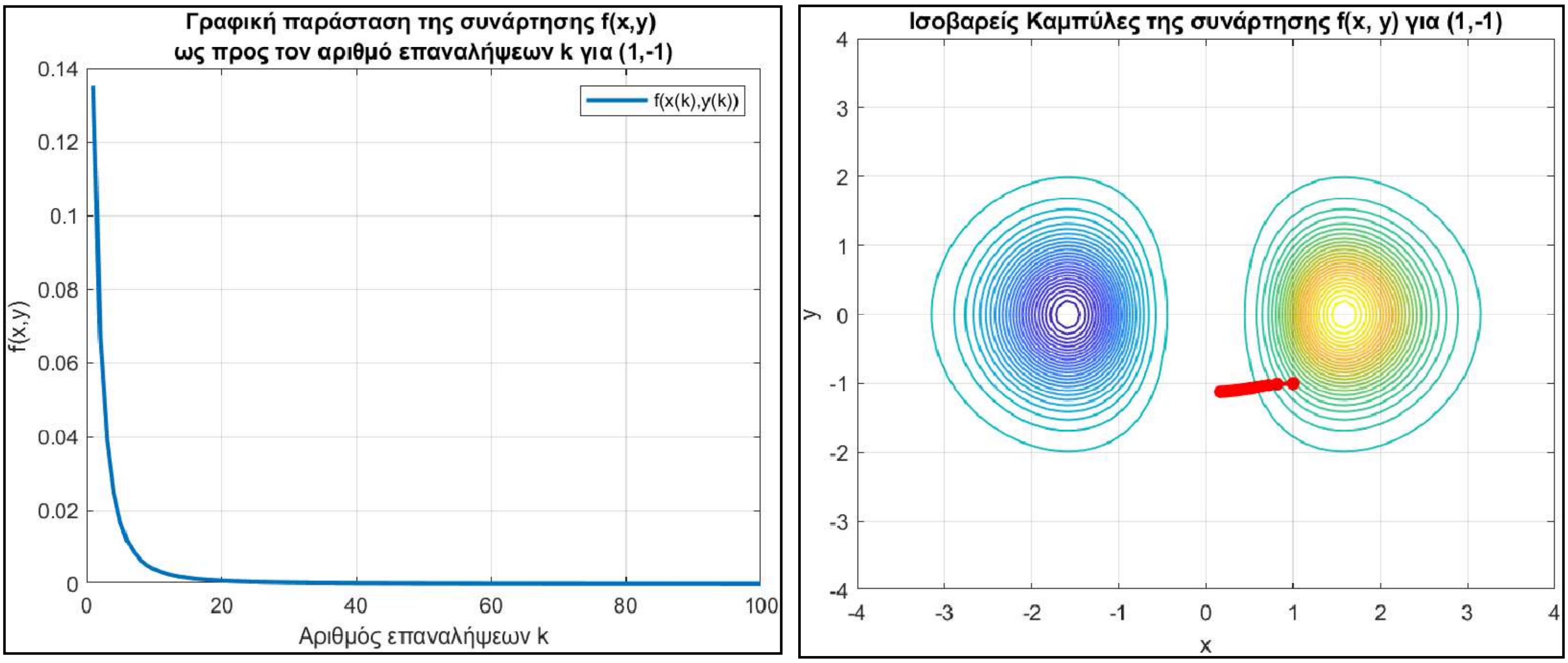
- Επιλογή βήματος $\boldsymbol{\gamma_k}$ μέσω του Κανόνα Armijo



Σχήμα 19.Τιμή της f(x,y) ως προς τις επαναλήψεις k για επιλογή βήματος γ_k μέσω του κανόνα Armijo και πορεία σύγκλισης στις ισοβαρείς καμπύλες για το σημείο (0,0)



Σχήμα 20.Τιμή της f(x,y) ως προς τις επαναλήψεις k για επιλογή βήματος γ_k μέσω του κανόνα Armijo και πορεία σύγκλισης στις ισοβαρείς καμπύλες για το σημείο (-1,1)



Σχήμα 21.Τιμή της f(x,y) ως προς τις επαναλήψεις k για επιλογή βήματος γ_k μέσω του κανόνα Armijo και πορεία σύγκλισης στις ισοβαρείς καμπύλες για το σημείο (1,-1)

○ Παρατηρήσεις

-γ_k που ελαχιστοποιεί την $f(x_k+\gamma_k d_k)$

Αρχικό Σημείο	Επαναλήψεις k	Ακρότατο
(0,0)	2	(0.0 , 0.0)
(-1,1)	16	(-1.5811 , 0.00)
(1,-1)	102	(0.165,-1.153)

-γ_k μέσω του κανόνα Armijo

Στις προσομοιώσεις έχουν επιλεχθεί α = 0.0001, β = 0.1 και αρχικό βήμα s = 0.6 .

Αρχικό Σημείο	Επαναλήψεις k	Ακρότατο
(0,0)	2	(0.0 , 0.0)
(-1,1)	13	(-1.5812 , 0.00)
(1,-1)	102	(0.1658,-1.1228)

Σύγκριση με βάση τα αρχικά σημεία

- Για το (0,0) που είναι τοπικό ελάχιστο , ο αλγόριθμος τερματίζει από τη πρώτη επανάληψη .
- Για το (-1,1) ο αλγόριθμος προσεγγίζει σωστά το ολικό ελάχιστο της αντικειμενικής συνάρτησης (-1.581,0) και με τις τρεις επιλογές του γ_κ.
- Για το (1,-1) ο αλγόριθμος εγκλωβίζεται και στις τρεις περιπτώσεις σε τοπικό ελάχιστο.

Σύγκριση με βάση την επιλογή του βήματος

Αρχικό Σημείο	γκ = 0.8	γ_k που ελαχιστοποιεί την $f(x_k+\gamma_k d_k)$	γ_k μέσω του κανόνα Armijo
(0,0)	2	2	2
(-1,1)	9	16	13
(1,-1)	102	102	102

Είναι σημαντικό να αναφερθεί ότι έχει τεθεί ένα πρόσθετο κριτήριο τερματισμού το οποίο τερματίζει τον αλγόριθμο στις περιπτώσεις όπου η τιμή του k γίνει 100. Άρα, στην περίπτωση του αρχικού σημείου (1,-1) ο αλγόριθμος ολοκληρώνεται λόγω του παραπάνω.

Συμπεράσματα και συγκρίσεις μεθόδων

I. Η μέθοδος μέγιστης καθόδου απαιτεί μεγαλύτερο αριθμό επαναλήψεων για να συγκλίνει, συγκριτικά με τους άλλους αλγορίθμους. Αυτό συμβαίνει επειδή η μέθοδος επιλέγει κάθετη κατεύθυνση προς το gradient σε κάθε βήμα, οπότε συγκλίνει με πιο αργό ρυθμό. Όμως, με τη κατάλληλη επιλογή του βήματος γ_k και του αρχικού σημείου ο αλγόριθμος μας οδηγεί σε ελάχιστο.

II. Η μέθοδος Newton έχει τη δυνατότητα να μας οδηγήσει σε ελάχιστο με μικρό αριθμό επαναλήψεων, αλλά όχι πάντα. Η κύρια δυσκολία της μεθόδου είναι ότι απαιτείται ο εσσιανός πίνακας να είναι θετικά ορισμένος αλλιώς ο αλγόριθμος τερματίζει στη πρώτη επανάληψη. Έτσι, η απόδοση της μεθόδου Newton εξαρτάται σημαντικά από το αρχικό σημείο που θα επιλέξουμε. Η μέθοδος Newton αποδείχθηκε μη αποτελεσματική για τη συγκεκριμένη ανάλυση βάση των αρχικών σημείων που επιλέχθηκαν, καθώς δεν κατάφερε να οδηγήσει σε σωστά αποτελέσματα. Ο λόγος αποτυχίας της οφείλεται στον εσσιανό πίνακα ο οποίος δεν ήταν θετικά ορισμένος. Αυτό οδήγησε σε αδυναμία υπολογισμού της κατεύθυνσης βήματος.

III. Η μέθοδος Levenberg-Marquardt συγκλίνει γρηγορότερα σε σχέση με τη μέθοδο της μέγιστης καθόδου ενώ ταυτόχρονα μας εξασφαλίζει ένα θετικά ορισμένο πίνακα $\nabla^2 f(x_k) + \mu_k I$ μέσω της επιλογής του μ_k. Όμως, η επιλογή του μ_k είναι σημαντική αφού για πολύ μικρό μ_k η μέθοδος συμπεριφέρεται όπως η μέθοδος Newton και αντίστοιχα για μεγαλό μ_k λειτουργεί όπως η μέθοδος μέγιστης καθόδου. Επομένως, ο αλγόριθμος είναι ο πιο αποδοτικός συγκριτικά με τους υπόλοιπους αλγορίθμους αρκεί να επιλεγεί το κατάλληλο μ_k.

Αναφορικά με τα αρχικά σημεία συμπεραίνεται :

I. Το σημείο (0,0) είναι τοπικό ακρότατο οπότε οδηγεί σε άμεσο τερματισμό κάθε αλγόριθμου.

II. Το σημείο (-1,1), όντας κοντά στο ολικό ελάχιστο της συνάρτησης, απαιτεί λιγότερες επαναλήψεις και είναι το μόνο που μας οδηγεί στην ελαχιστοποίηση της συνάρτησης, καθώς τα άλλα δύο σημεία εγκλωβίζονται σε τοπικό ελάχιστο.

III. Το σημείο (1,-1) απαιτεί πολλές επαναλήψεις για να φτάσει σε ελάχιστο.

Αναφορικά με την επιλογή του βήματος $γ_k$:

I. Αναλόγως του αρχικού σημείου απαιτείται αντίστοιχα μικρό βήμα για το (-1,1) ή μεγάλο βήμα για το (1,-1) καθώς τα αντίθετα μπορούν να οδηγήσουν σε απόκλιση ή εγκλωβισμό σε τοπικό ελάχιστο αντιστοίχως , όπως φάνηκε και από την προσομοίωση.

II. Η καλύτερη μέθοδος επιλογής βήματος είναι ο κανόνας Armijo. Στη περίπτωση μας θα αναμενόταν η μέθοδος επιλογής βήματος μέσω ελαχιστοποίησης της $f(x_k+γ_k d_k)$ να είναι η πιο αποδοτική. Ωστόσο, στη συγκεκριμένη ανάλυση επιλέχθηκε η μέθοδος του χρυσού τομέα που είχε υλοποιηθεί στη πρώτη εργασία. Είναι πιθανό, αν είχε επιλεγεί η μέθοδος της διχοτόμου με τη χρήση παραγώγων να είχε επιτευχθεί ταχύτερη σύγκλιση.

Βιβλιογραφία

Η υλοποίηση των αλγορίθμων, καθώς και οι παρατηρήσεις βασίστηκαν στο βιβλίο *«Τεχνικές Βελτιστοποίησης»* του κ. Ροβιθάκη, Εκδόσεις ΤΖΙΟΛΑ.