Τεχνικές Βελτιστοποίησης
Εργασία 3η

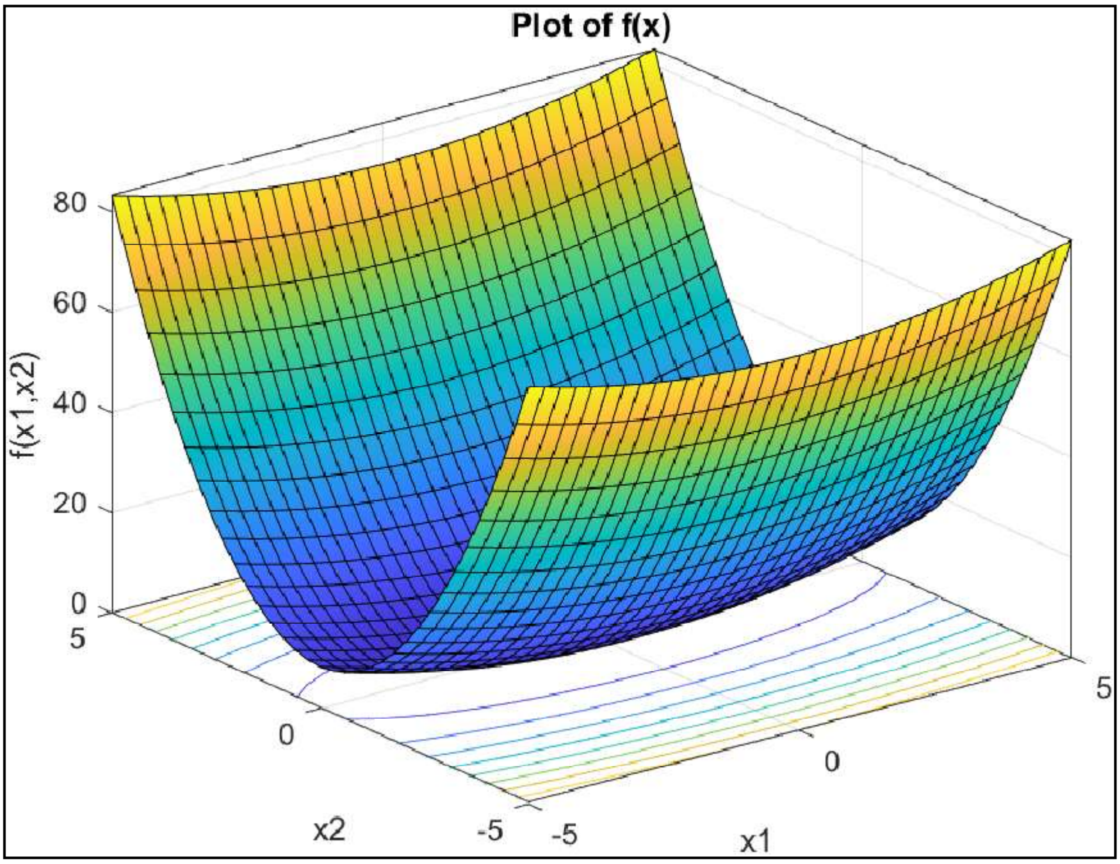
Παπουτσή Νικολέτα AEM : 10858

Ζητούμενο : Σύγκριση του αλγορίθμου της **μεθόδου μέγιστης καθόδου** με ή χωρίς προβολή για διαφορετικές τιμές του βήματος γ_k ή των αρχικών σημείων εκκίνησης για δοσμένη συνάρτηση f : $\mathbb{R}^2 \rightarrow \mathbb{R}$

Η αντικειμενική συνάρτηση της οποίας ψάχνουμε το ελάχιστο είναι :

$$f(x) = \frac{1}{3}x_1^2 + 3x_2^2 \text{ , όπου } x = [x_1\ x_2]^T$$

Παρακάτω παρουσιάζεται η μορφή της συνάρτησης f στον τρισδιάστατο χώρο, καθώς και οι ισοβαρείς καμπύλες της προκειμένου να εντοπιστεί οπτικά το σημείο ελαχίστου.



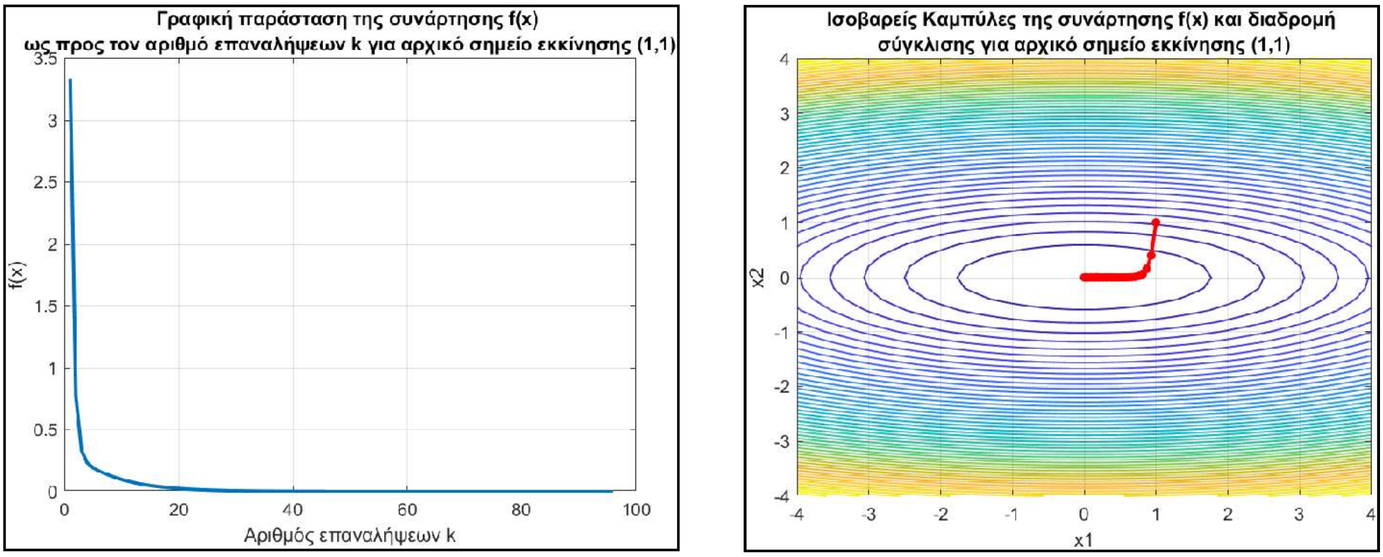
Σχήμα 1. Γραφική παράσταση της f

Η f παρουσιάζει ελάχιστο στο (0,0).

Θέμα 1ο : Μέθοδος μέγιστης καθόδου

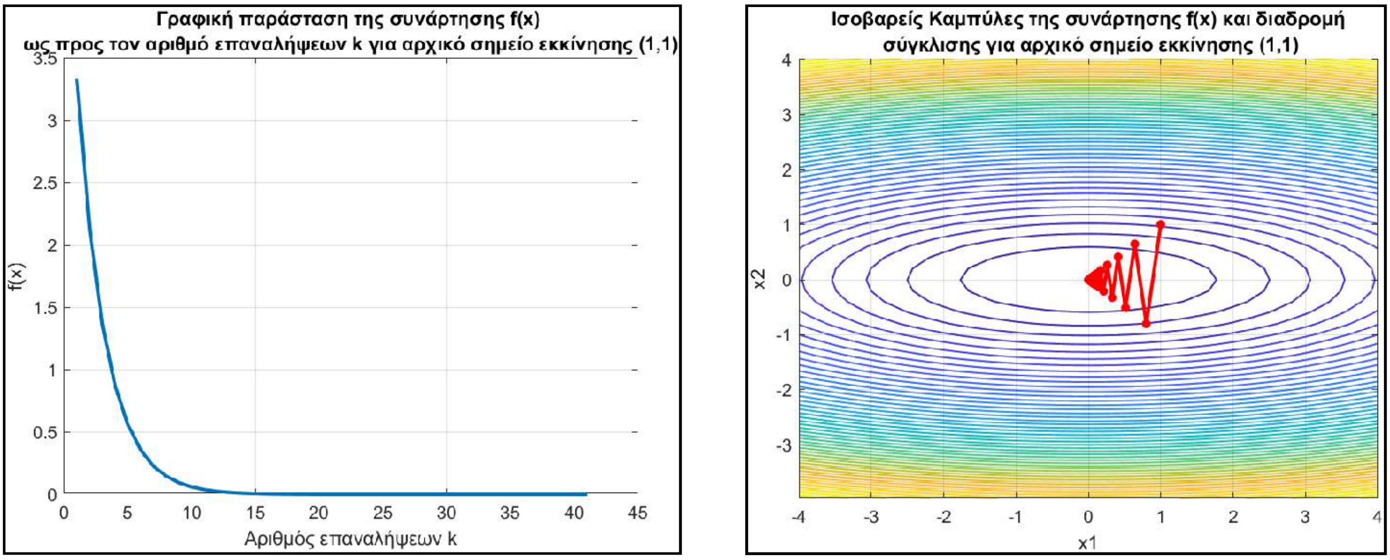
Στο πλαίσιο της παρούσας εργασίας , αξιοποιείται η υλοποίηση της μεθόδου μέγιστης καθόδου από την 2η εργασία με ελάχιστες τροποποιήσεις . Συγκεκριμένα , η μέθοδος εφαρμόστηκε για τη συνάρτηση f με αρχικό σημείο εκκίνησης το (1,1) για διαφορετικές σταθερές τιμές του γ_k και με ακρίβεια ε = 0.001 .

i. $\gamma_k = 0.1$



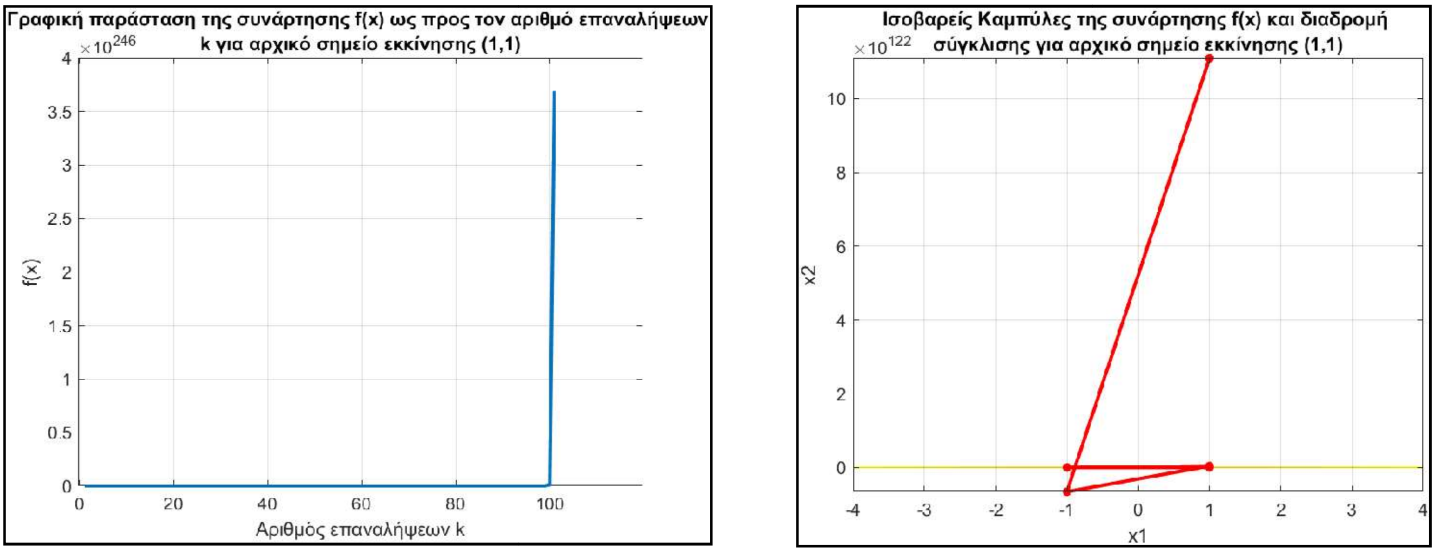
Σχήμα 2.Τιμή της *f(x)* ως προς τις επαναλήψεις k για σταθερό βήμα γ_k = 0.1 και πορεία σύγκλισης στις ισοβαρείς καμπύλες για το σημείο (1,1)

ii. $\gamma_k = 0.3$



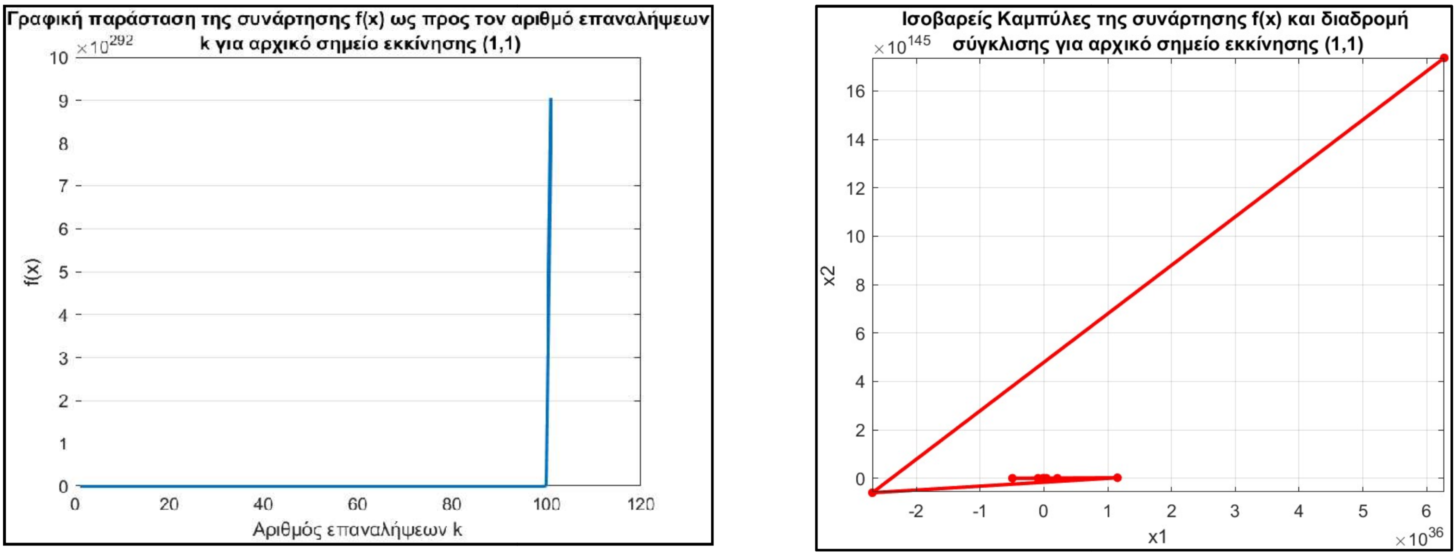
Σχήμα 3.Τιμή της *f(x)* ως προς τις επαναλήψεις k για σταθερό βήμα γ_k = 0.3 και πορεία σύγκλισης στις ισοβαρείς καμπύλες για το σημείο (1,1)

iii. $\gamma_k = 3$



Σχήμα 4.Τιμή της *f(x)* ως προς τις επαναλήψεις k για σταθερό βήμα $\gamma_k = 3$ και πορεία σύγκλισης στις ισοβαρείς καμπύλες για το σημείο (1,1)

iv. $\gamma_k = 5$



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

γ_k	Αριθμός επαναλήψεων k	Τελική τιμή της f	Τελικό σημείο (x_1,x_2)
0.1	**96**	0	(0.0014 , 0.0)
0.3	**41**	0	(0.0001 , 0.0001)
3	**101**	$\sim 10^{246}$	(0.00,1.108)
5	**101**	$\sim 10^{292}$	-

Όπως, φαίνεται από τα παραπάνω γραφήματα :

- Για μικρές τιμές του βήματος $\boldsymbol{\gamma_k}$ η μέθοδος συγκλίνει στο ελάχιστο με σχετικά μικρό αριθμό επαναλήψεων. Για $\boldsymbol{\gamma_k}$ = 0.1 συγκλίνει αργά αλλά σταθερά. Για $\boldsymbol{\gamma_k}$ = 0.3 επιτυγχάνεται η βέλτιστη σύγκλιση.
- Αντιθέτως, για μεγαλύτερες τιμές του $\boldsymbol{\gamma_k}$, όπως 3 και 5, παρουσιάζεται απόκλιση από το ελάχιστο και η τιμή της συνάρτησης καταλήγει να αυξάνεται εκθετικά.

Παρακάτω παρουσιάζεται η απόδειξη των αποτελεσμάτων με **μαθηματική αυστηρότητα**

Ο gradient της συνάρτησης f είναι :

$$\nabla f(x_1, x_2) = \begin{bmatrix} \dfrac{\partial f(x_1, x_2)}{\partial x_1} \\ \dfrac{\partial f(x_1, x_2)}{\partial x_2} \end{bmatrix} = \begin{bmatrix} \frac{2}{3} x_1 \\ 6x_2 \end{bmatrix}$$

άρα,
$$\nabla f\left(x_{1_k}, x_{2_k}\right) = \begin{bmatrix} \frac{2}{3} x_{1_k} \\ 6x_{2_k} \end{bmatrix}$$

Η μέθοδος της μέγιστης καθόδου παράγει σημεία που ικανοποιούν :

$$x_{k+1} = x_k - \gamma_k \nabla f(x_k) \qquad (1.1)$$

Η (1.1) λοιπόν γίνεται

$$\begin{bmatrix} x_{1_{k+1}} \\ x_{2_{k+1}} \end{bmatrix} = \begin{bmatrix} x_{1_k} \\ x_{2_k} \end{bmatrix} - \gamma_k \begin{bmatrix} \frac{2}{3} x_{1k} \\ 6x_{2k} \end{bmatrix} \qquad (1.2)$$

Το ελάχιστο της συνάρτησης $f(x_1,x_2)$ βρίσκεται στο σημείο (0,0), επομένως ο αλγόριθμος πρέπει να συγκλίνει σε αυτό το σημείο. Για να συμβεί αυτό η (1.2) πρέπει να συγκλίνει στο (0,0) καθώς το $k \to \infty$. Άρα πρέπει να ισχύει :

$$x_{1_{k+1}} = (1 - \tfrac{2}{3}\gamma_k)x_{1k} \qquad\qquad \lim_{k\to\infty} x_{1_k} = 0$$

θέλουμε :

$$x_{2_{k+1}} = (1 - 6\gamma_k)x_{2k} \qquad\qquad \lim_{k\to\infty} x_{2_k} = 0$$

Άρα , πρέπει να ισχύουν οι εξής ανισότητες :

$$\left|1 - \tfrac{2}{3}\gamma_k\right| < 1 \qquad \text{και} \qquad |1 - 6\gamma_k| < 1$$

Λύνοντας , καταλήγουμε στο συμπέρασμα ότι το βήμα γ_k πρέπει να ικανοποιεί την ακόλουθη συνθήκη για να εξασφαλιστεί η σύγκλιση :

$$\boxed{\mathbf{0 < \boldsymbol{\gamma_k} < \frac{1}{3}}}$$

Επομένως, η παραπάνω μαθηματική απόδειξη συμφωνεί πλήρως με τα γραφικά αποτελέσματα , καθώς επιβεβαιώνεται ότι για τιμές του βήματος γ_k μικρότερες από το όριο $\mathbf{\frac{1}{3}}$, ο αλγόριθμος συγκλίνει και καταλήγει στο ελάχιστο. Αντίθετα, για μεγαλύτερες τιμές του βήματος γ_k παρατηρείται έντονη απόκλιση , καθώς η τιμή της συνάρτησης αυξάνεται εκθετικά, αντί να μειώνεται.

Επιπλέον, η αποσβεννύμενη ταλάντωση που παρατηρείται για τη τιμή $\gamma_k = 0.3$ εξηγείται από το γεγονός ότι το βήμα γ_k είναι αρκετά μεγάλο για να προκαλέσει ταλάντωση , αλλά όχι τόσο μεγάλο ώστε να οδηγήσει σε αστάθεια του αλγορίθμου.

<u>Μέθοδος Μέγιστης Καθόδου με Προβολή</u>

Υλοποιείται ο αλγόριθμος της μεθόδου μέγιστης καθόδου με προβολή για την συνάρτηση *f(x)* με περιορισμούς :

$-10 \leq x_1 \leq 5$ και $-8 \leq x_2 \leq 12$

Η μέθοδος είναι αλγόριθμος εφικτών σημείων της μορφής :

$$x_{k+1} = x_k + \gamma_k(\bar{x}_k - x_k), \quad \gamma_k \in (0, 1]$$

όπου

$$\bar{x}_k = Pr_X\{x_k - s_k \nabla f(x_k)\}, \quad s_k > 0$$

Έχοντας βρει την εφικτή κατεύθυνση αναζήτησης $\bar{x}_k - x_k$, κινούμαστε με βήμα γ_k σ' αυτή και προσδιορίζουμε το νέο εφικτό σημείο x_{k+1}.

- Αν το νέο σημείο είναι εφικτό, η μέθοδος της μέγιστης καθόδου με προβολή μετατρέπεται στη μέθοδο της μέγιστης καθόδου χωρίς περιορισμούς με βήμα $\gamma'_k = \gamma_k s_k$.

- Αν το νέο σημείο δεν είναι εφικτό, βρίσκουμε την προβολή του στο κυρτό σύνολο και επαναλαμβάνουμε τη διαδικασία μέχρι να καταλήξουμε σε στάσιμο σημείο.

Για το γ'_k ισχύουν οι περιορισμοί του Θέματος 1. Δηλαδή:

$$\mathbf{0 < \gamma'_k < \frac{1}{3}} \qquad (1.3)$$

- $\bar{x}_{1k} = Pr_X\{x_{1k} - s_k \nabla f(x_k)\} = Pr_X\left\{x_{1k} - \frac{2}{3} s_k x_{1k}\right\}$
- $\bar{x}_{2k} = Pr_X\{x_{2k} - s_k \nabla f(x_k)\} = Pr_X\{x_{2k} - 6\, s_k x_{2k}\}$

Οι προβολές των x_1 , x_2 *προκύπτουν από τον τύπο στη σελίδα 202 του βιβλίου*:

$$[Pr_X\{x]_i = \begin{cases} \alpha_i, & \alpha\nu\ x_i \leq \alpha_i \\ \beta_i, & \alpha\nu\ x_i \geq \beta_i \\ x_i, & \alpha\nu\ \alpha_i < x_i < \beta_i \end{cases}$$

Επομένως, από όλα τα παραπάνω προκύπτει, αν το νέο σημείο δεν είναι εφικτό :

- $x_{1k+1} = x_{1k} - \frac{2}{3} s_k x_{1k} \gamma_k$
- $x_{2k+1} = x_{2k} - 6 s_k x_{2k} \gamma_k$

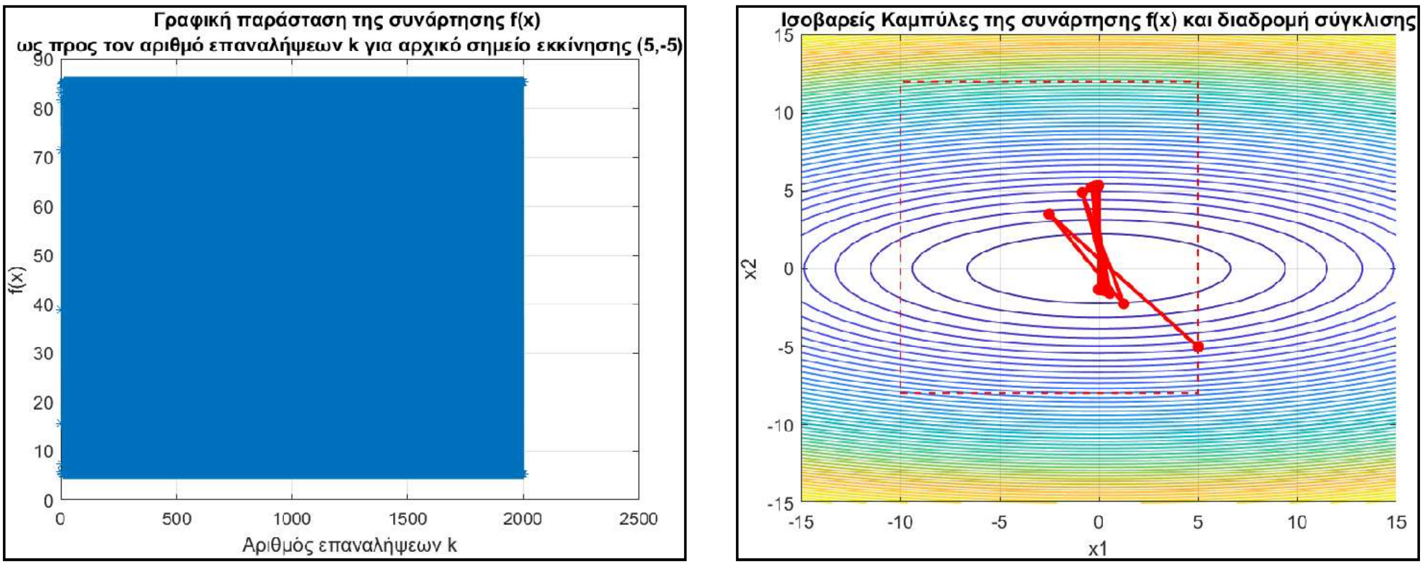
<u>**Θέμα 2ο**</u>

- Αρχικό σημείο : (5 , -5)
- Βήμα s_k =5
- Βήμα γ_k = 0.5
- Ακρίβεια ε = 0.01

Από τα παραπάνω προκύπτει :

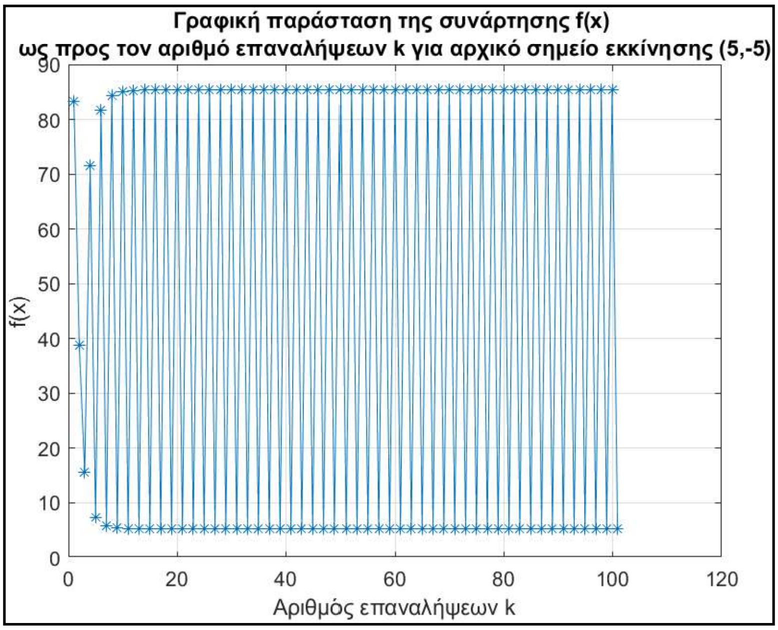
$$\gamma'_k = \gamma_k s_k = 0.5 \cdot 5 = 2.5$$

Δηλαδή, αναμένουμε ο αλγόριθμος να μην συγκλίνει καθώς το βήμα γ'_k είναι πολύ μεγαλύτερο του άνω ορίου της (1.3).



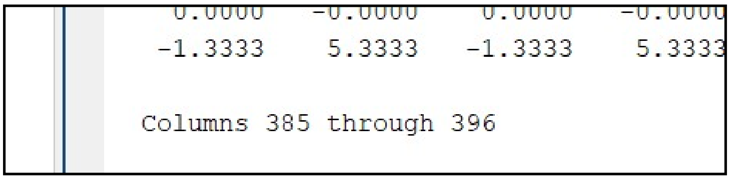
Σχήμα 6.Τιμή της *f(x)* ως προς τις επαναλήψεις k για βήματα γ_k = 0.5 , s_k = 5 και πορεία σύγκλισης στις ισοβαρείς καμπύλες για αρχικό σημείο (5,-5)

Το κόκκινο διακεκομμένο παραλληλόγραμμο εκφράζει τους περιορισμούς για τις τιμές των x_1 , x_2.



Σχήμα 7.Τιμή της *f(x)* ως προς τις επαναλήψεις k =100 για βήματα $\gamma_k = 0.5$, $s_k = 5$ για αρχικό σημείο (5,-5) , ώστε να διακρίνεται η ταλάντωση

Από τα γραφήματα παρατηρούμε ότι ο αλγόριθμος δεν συγκλίνει όπως αναμέναμε. Αντίθετα, ταλαντώνεται επ' αόριστον και τελικά τερματίζει λόγω του κριτηρίου τερματισμού που θέσαμε για τον αριθμό επαναλήψεων k. Η ταλάντωση προκύπτει από τη χρήση της προβολής, η οποία εγκλωβίζει τα σημεία x_k εντός των προκαθορισμένων ορίων. Συγκεκριμένα, παρατηρούμε ότι το x_{2k} εναλλάσσεται μεταξύ των τιμών -1.3333 και 5.3333 ,ενώ το x_{1k} συγκλίνει σταδιακά στο 0.



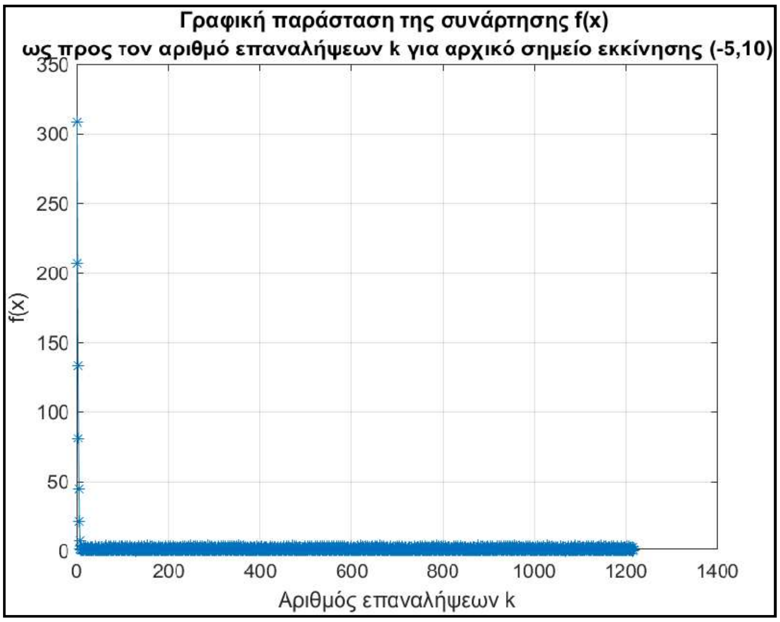
Συγκριτικά με το Θέμα 1 όπου παρατηρήθηκε απόκλιση και εκθετική αύξηση των τιμών της συνάρτησης , η διαφορά σε αυτή τη περίπτωση είναι ότι η μέθοδος της προβολής εμποδίζει την εκθετική αύξηση των τιμών της συνάρτησης f . Αντίθετα, λόγω των περιορισμών του κυρτού συνόλου , οι τιμές εγκλωβίζονται και ταλαντώνονται εντός του συνόλου, αποτρέποντας την απόκλιση του αλγορίθμου.

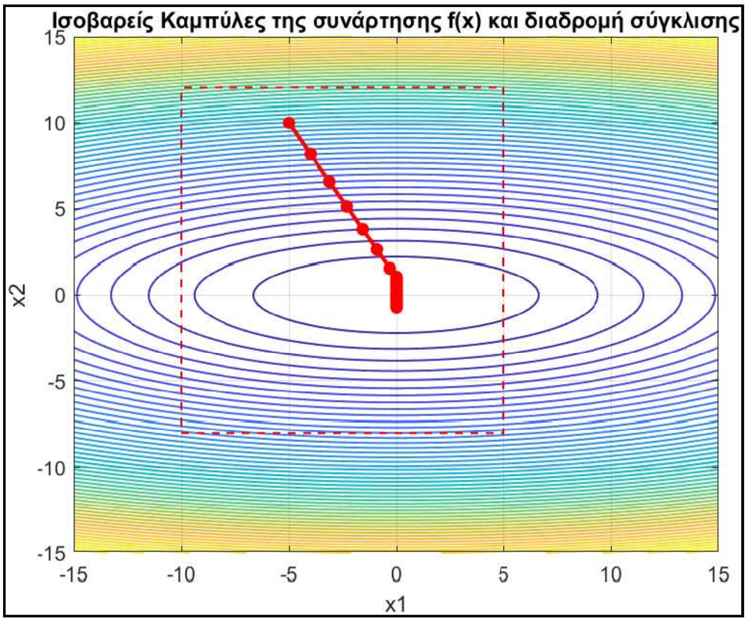
<u>Θέμα 3ο</u>

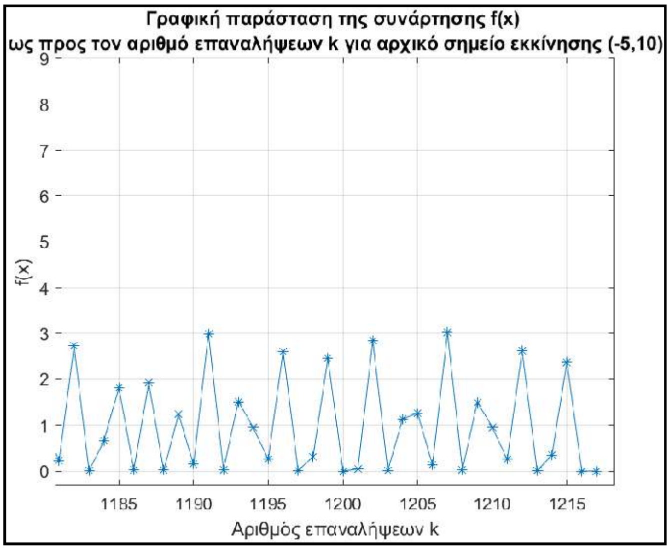
- Αρχικό σημείο : (-5 , 10)
- Βήμα s_k =15
- Βήμα γ_k = 0.1
- Ακρίβεια ε = 0.01

Από τα παραπάνω προκύπτει :

$$\gamma'_k = \gamma_k s_k = 0.1 \cdot 15 = 1.5$$







Σχήμα 8.Τιμή της *f(x)* ως προς τις επαναλήψεις k για βήματα γ_k = 0.1 , s_k = 15 , πορεία σύγκλισης στις ισοβαρείς καμπύλες για αρχικό σημείο (-5,10) και προβολή με zoom στη περιοχή ταλάντωσης

Στη περίπτωση αυτή το τελικό βήμα γ'_k είναι αρκετά κοντά στο άνω όριο, γεγονός που εξηγεί τη μικρότερη ταλάντωση και τη σύγκλιση που παρατηρείται. Συγκριτικά με τα Θέματα 1 και 2, παρατηρούμε ότι ο αλγόριθμος τελικά συγκλίνει στο ολικό ελάχιστο, αν και απαιτούνται k = 1217 επαναλήψεις.

Από την ανάλυση και τις ισοβαρείς καμπύλες γίνεται φανερό ότι το x_{1k} συγκλίνει πολύ γρήγορα στο 0, ενώ το x_{2k} απαιτεί σημαντικά περισσότερες επαναλήψεις. Η γρήγορη σύγκλιση του x_{1k} μπορεί να αποδοθεί στην κατάλληλη επιλογή των s_k και γ_k. Όταν η προβολή $\bar{x}_k$ βρίσκεται εντός του κυρτού συνόλου από την παρακάτω σχέση προκύπτει:

$$x_{1k+1} = x_{1k} - \frac{2}{3} s_k x_{1k} \gamma_k$$

ότι για $s_k\gamma_k = 1.5$, το x_{1k} γίνεται μηδέν στην επόμενη επανάληψη.

Αντίθετα, το x_{2k} εξαρτάται από τον παράγοντα $6s_k\gamma_k$, ο οποίος είναι πολύ μεγαλύτερος , οδηγώντας σε πιο αργή σύγκλιση.

Πρόταση ώστε η μέθοδος να συγκλίνει στο ελάχιστο :

Αν προσαρμόσουμε τις παραμέτρους s_k και γ_k ώστε να εξισορροπηθεί η σύγκλιση και στις δύο συντεταγμένες , θα μπορούσε να μειωθεί σημαντικά ο συνολικός αριθμός επαναλήψεων που απαιτούνται για σύγκλιση.

Συγκεκριμένα,

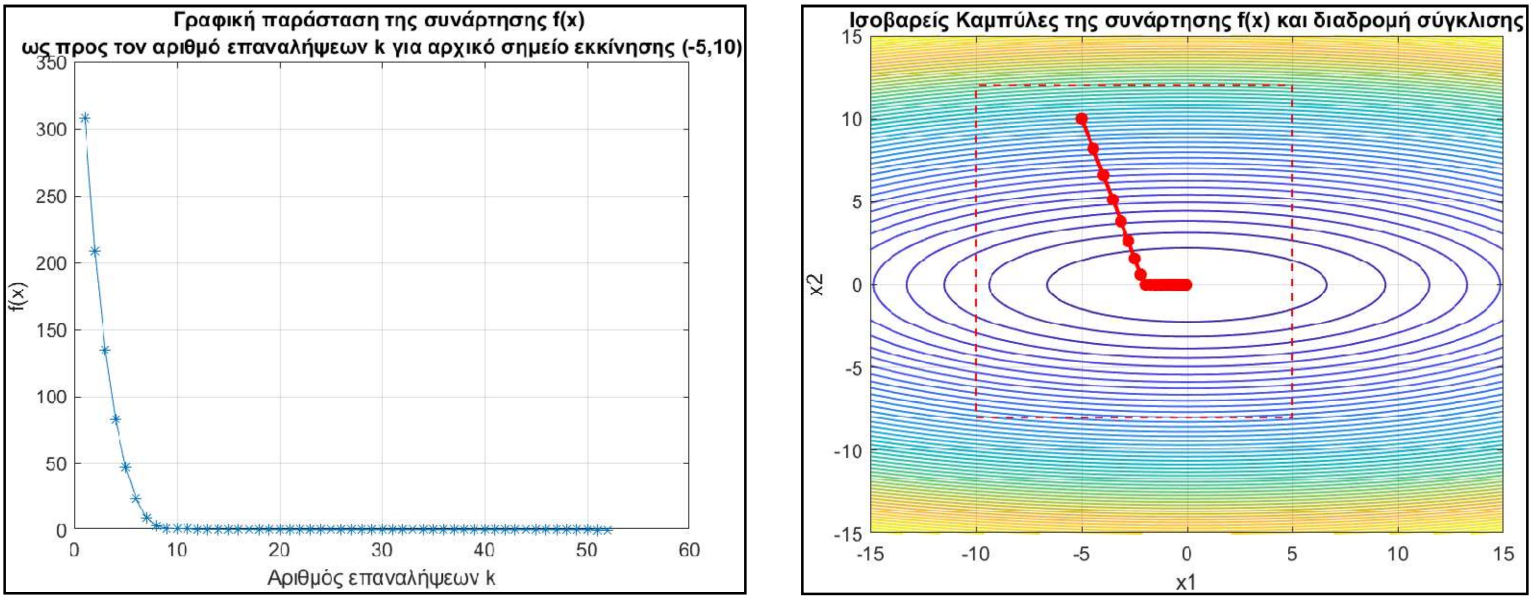
$$x_{2k+1} = x_{2k} - 6 s_k x_{2k} \gamma_k$$

Μπορούμε να οδηγηθούμε σε σύγκλιση στο ελάχιστο αν μεταβάλλουμε τη τιμή του s_k έτσι ώστε το καινούργιο βήμα να πληρεί τα όρια της (1.3).

$$\gamma'_k = \gamma_k s_k = 0.1 \cdot s_k < \tfrac{1}{3}$$

Επιπλέον, για βήμα $\boldsymbol{\gamma_k = 0.1}$ **και** $\boldsymbol{s_k = \frac{10}{6}}$ η παρακάτω σχέση συγκλίνει με μικρό αριθμό επαναλήψεων όπως έγινε με το x_{1k} .

$$x_{2k+1} = x_{2k}(1 - 6 s_k \gamma_k) = x_{2k}(1 - 6 \cdot \frac{1}{10} s_k)$$



Σχήμα 9.Τιμή της *f(x)* ως προς τις επαναλήψεις k για βήματα γ_k = 0.1 , s_k = 1.66 και πορεία σύγκλισης στις ισοβαρείς καμπύλες για αρχικό σημείο (-5,10)

<u>Παρατηρήσεις</u>:

- Με τη βελτιστοποίηση της παραμέτρου s_k συγκλίνει στο ελάχιστο με σημαντικά μικρότερο αριθμό επαναλήψεων, k = 52. Έτσι, το βήμα προσαρμόζεται ώστε να αποφεύγεται η ταλάντωση, οδηγώντας σε ταχύτερη σύγκλιση. Να σημειωθεί ότι για s_k=2 παρατηρείται βέλτιστος αριθμός επαναλήψεων, k =47.

- Η σύγκλιση μπορεί να επιτευχθεί πιο αποδοτικά και με την αλλαγή του σημείου εκκίνησης. Επιλέγοντας ένα σημείο που βρίσκεται πιο κοντά στο ολικό ελάχιστο, μειώνεται ο αριθμός επαναλήψεων. Για παράδειγμα, εάν το x_{2k} ξεκινά από το μηδέν, η σύγκλιση για αυτό είναι άμεση και το x_{1k} συγκλίνει στο ελάχιστο σε μικρό αριθμό επαναλήψεων.

Θέμα 4ο

- Αρχικό σημείο : (8 , -10)
- Βήμα s_k =0.1
- Βήμα $γ_k$ = 0.2
- Ακρίβεια ε = 0.01

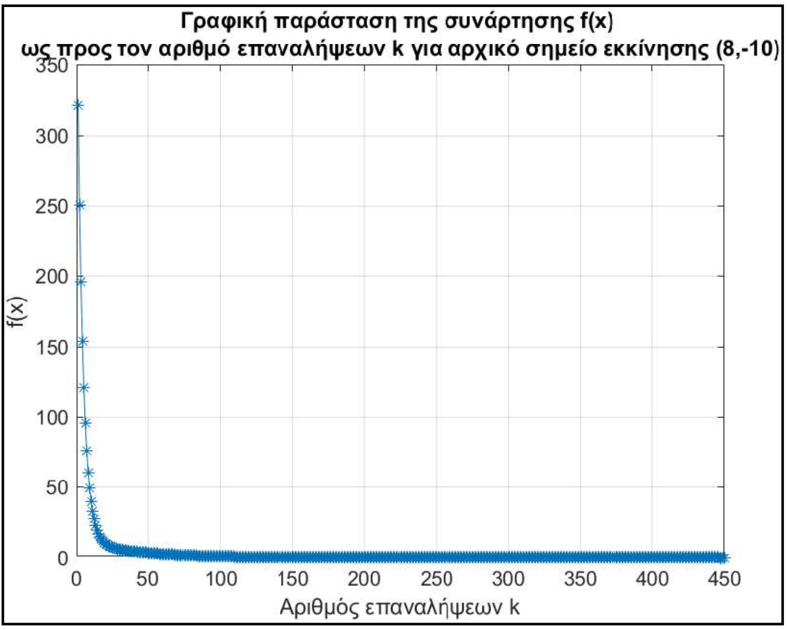
Από τα παραπάνω προκύπτει :

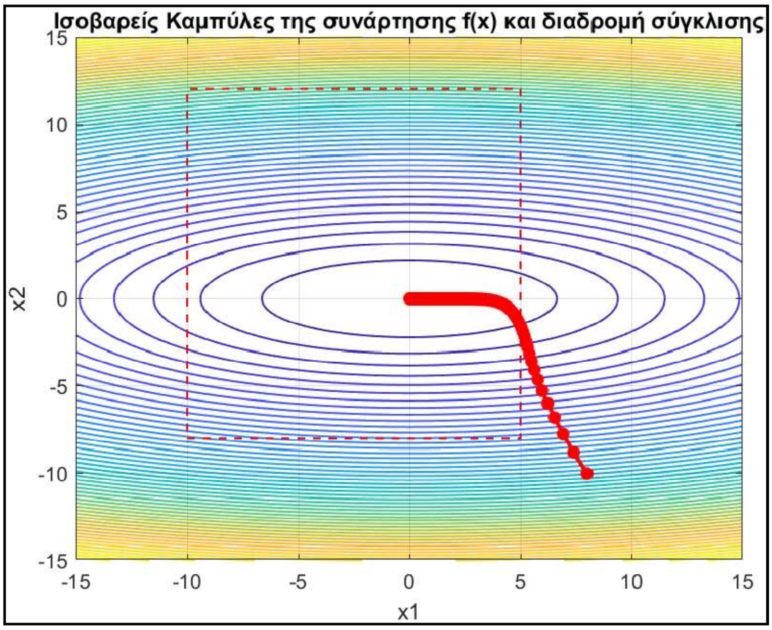
$$\gamma'_k = \gamma_k s_k = 0.1 \cdot 0.2 = 0.02$$

Σε αυτή τη περίπτωση , υπάρχει εκ των προτέρων πληροφορία για τη σύγκλιση του αλγορίθμου. Οι περιορισμοί της σύγκλισης $\gamma'_k < \frac{1}{3}$ πληρούνται , γεγονός που υποδηλώνει εκθετική σύγκλιση στο ολικό ελάχιστο. Συγκεκριμένα :

$$x_{1k+1} = (0.9866)x_{1k}$$
$$x_{2k+1} = (0.88)x_{2k}$$

Το x_{2k+1} συγκλίνει γρηγορότερα από το x_{1k+1}.





Σχήμα 10.Τιμή της *f(x)* ως προς τις επαναλήψεις k για βήματα $γ_k$ = 0.2 , s_k = 0.1 και πορεία σύγκλισης στις ισοβαρείς καμπύλες για αρχικό σημείο (8,-10)

Παρόλο που η σύγκλιση είναι αναμενόμενη, ο συνολικός αριθμός των επαναλήψεων, k = 450, είναι σημαντικά μεγάλος , λόγω της μικρής μείωσης ανά βήμα του x_{1k+1} . Το x_{1k+1} μειώνεται αργά, ενώ το x_{2k+1} μηδενίζεται σε λίγες επαναλήψεις. Αυτό είναι αποτέλεσμα της μεγαλύτερης επίδρασης του όρου $6\gamma_k s_k$ στον υπολογισμό του x_{2k+1} .

<u>Βιβλιογραφία</u>

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