Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
Πολυτεχνική Σχολή
Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Τεχνικές Βελτιστοποίησης

Αξιμιώτης Δημήτρης 10622

Νοέμβριος 2024

1 Μέθοδος μέγιστης καθόδου

Για την αντικειμενική συνάρτηση $f(x,y) = \frac{1}{3}x^2 + 3y^2$, $\nabla f(x,y) = [\frac{2}{3}x, 6y]$ έχουμε κρίσιμα σημεία για $\nabla f(x,y)$=0 οπότε x=y=0.Παίρνοντας τον εσσιανό πίνακα $\nabla^2 f(x,y)$=

$$\begin{pmatrix} 0.666 & 0 \\ 0 & 6 \end{pmatrix}$$

είναι θετικά ορισμένος οπότε το (0,0) είναι ολικό ελάχιστο της f.Για να αποδείξουμε την σύγκλιση της f στο (0,0) με την μέθοδο της μέγιστης καθόδου θα πρέπει για κάθε επανάληψη να ισχύει $|\frac{x_{k+1}}{x_k}| < 1$ και $|\frac{y_{k+1}}{y_k}| < 1$ το διάνυσμα κλίσης είναι το $d_k = -\nabla f(x,y)$ και για κάθε επανάληψη ισχύει ότι:

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

$$y_{k+1} = y_k - \gamma_k 6 y_k$$

οπότε ,

$$y_{k+1} = y_k(1 - \gamma_k 6)$$

$$x_{k+1} = x_k(1 - \gamma_k \frac{2}{3})$$

$|\frac{x_{k+1}}{x_k}| < 1 => 0 < \gamma_k < 3$

$|\frac{y_{k+1}}{y_k}| < 1 => 0 < \gamma_k < \frac{1}{3}$.

Άρα για να έχουμε σύγκλιση με σταθερό βήμα θα πρέπει γ< $\frac{1}{3}$. Έχοντας τις αναδρομικές σχέσεις των x_{k+1} και y_{k+1} μπορούμε να βρούμε τον ρυθμό σύγκλισης με $x_k = x_0(1 - \frac{2}{3}\gamma_k)^k$,$y_k = y_0(1 - 6\gamma_k)^k$.

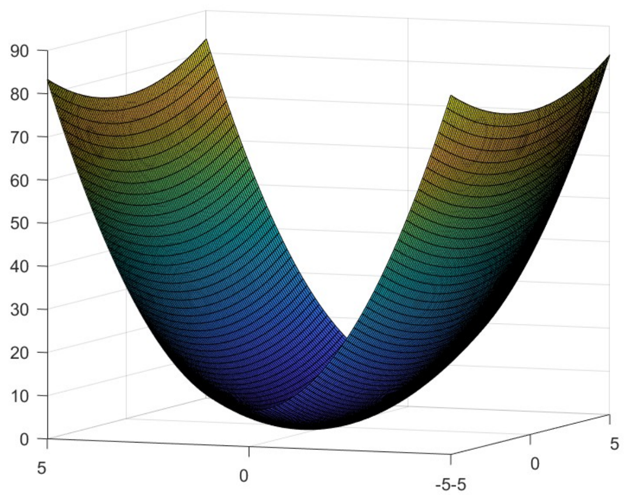

Figure 1: Γραφική παράσταση αντικειμενικής συνάρτησης f

1.1 $\gamma_k = 0.1$

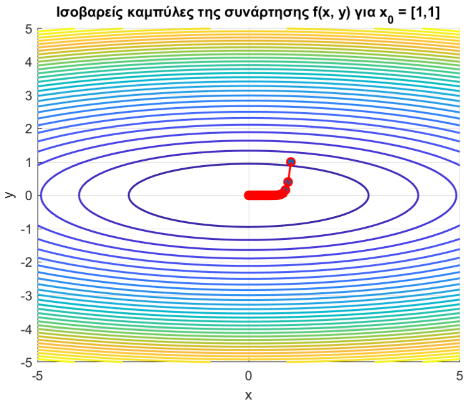

Figure 2: Σύγκλιση της f σε ισοβαρής καμπύλες

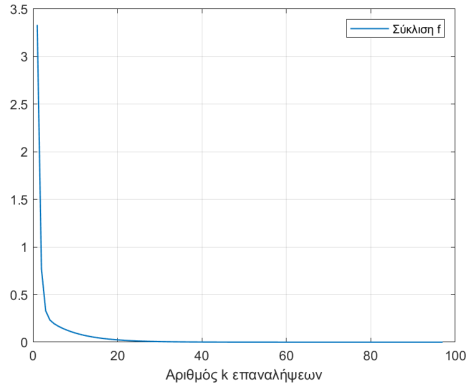

Figure 3: Σύγκλιση της f ανά k επαναλήψεις

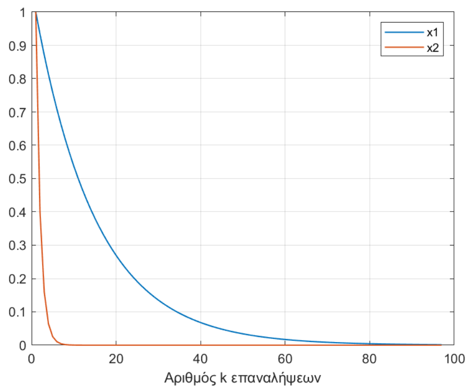

Figure 4: Μεταβολή x_1, x_2

Ο συνολικός αριθμός επαναλήψεων ειναι k=97.Το γ=0.1 είναι μικρότερο του 1/3 οπότε με αυτό το βήμα ο αλγόριθμος θα συγκλίνει το ολικό ελάχιστο.Το x_1 συγκλίνει πιο αργά καθώς ο ρυθμός σύγκλισης του είναι πιο αργός από αυτόν του x_2 αφού για γ=0.1 και $x_0 = [1, 1]$ $x_k = (0.9333)^k$ και $y_k = (0.4)^k$ συνεπώς το y συγκλίνει πιο γρήγορα στο 0 (αφού έχει μικρότερη βάση).

1.2 $\gamma_k = 0.3$

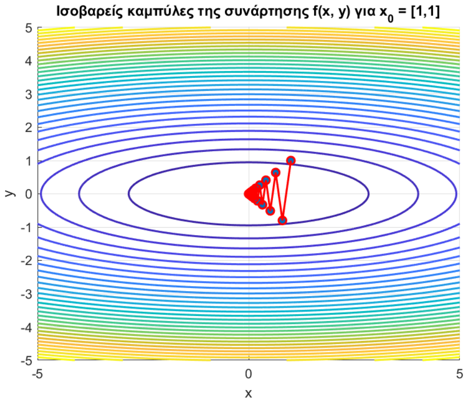

Figure 5: Σύγκλιση της f σε ισοβαρής καμπύλες

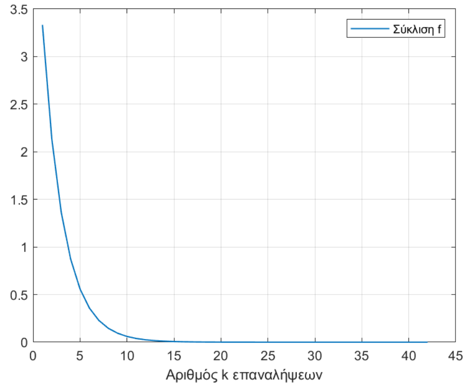

Figure 6: Σύγκλιση της f ανά k επαναλήψεις

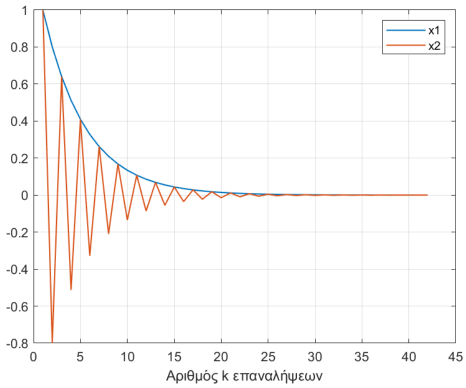

Figure 7: Μεταβολή x_1, x_2

Ο συνολικός αριθμός επαναλήψεων ειναι k=42.Το γ=0.3 είναι μικρότερο του 1/3 οπότε με αυτό το βήμα ο αλγόριθμος θα συγκλίνει το ολικό ελάχιστο.Το x_1 συγκλίνει με τον ίδιο ρυθμό που συγκλίνει το x_2 κατά απόλυτη τιμή αφού για γ=0.1 και $x_0 = [1, 1]$ $x_k = (0.8)^k$ και $y_k = (-0.8)^k$.Η ταλάντωση του y όπως φαίνεται και στο διάγραμμα οφείλεται στην αρνητική τιμή του ρυθμού σύγκλισης,οπότε παρόλο που και 2 μεταβλητές φτάνουν στο 0 σχεδόν ταυτόχρονα το y παρουσιάζει αυτην την εναλλαγή προσήμου ανά επανάληψη.

1.3 $\gamma_k = 3$

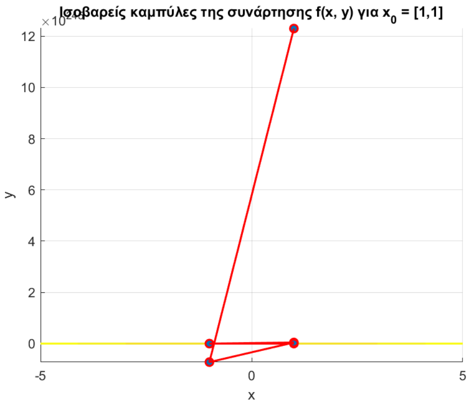

Figure 8: Σύγκλιση της f σε ισοβαρής καμπύλες

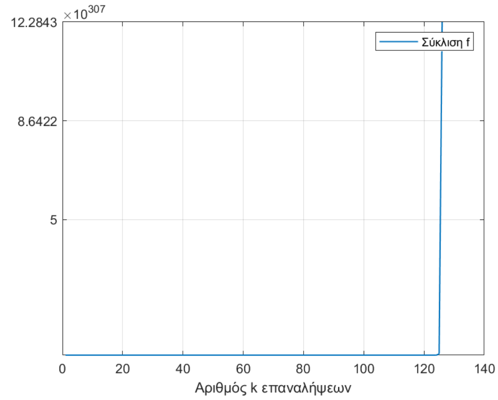

Figure 9: Σύγκλιση της f ανά k επαναλήψεις

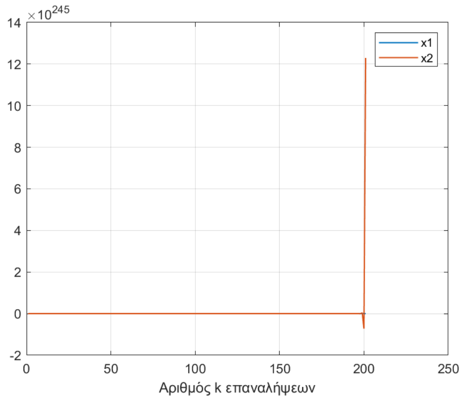

Figure 10: Μεταβολή x_1, x_2

Για $\gamma_k = 3$ όπως έχουμε αποδείξει παραπάνω ο αλγόριθμος δεν συγκλίνει.Η μεταβλητή y αποκλίνει σε πολύ μεγάλες τιμές ενώ για την χ στην ουσία έχουμε εναλλαγή προσήμου αφού για γ=3 $x_k = x_0 * (-1)^k$.

1.4 $\gamma_k = 5$

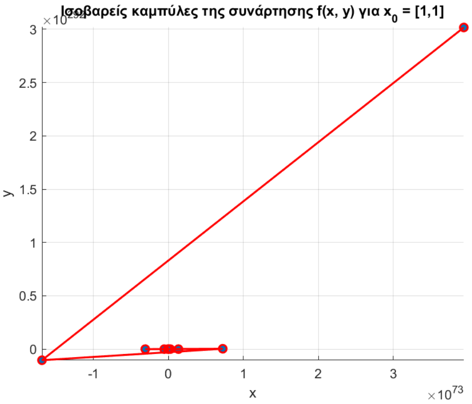

Figure 11: Σύγκλιση της f σε ισοβαρής καμπύλες

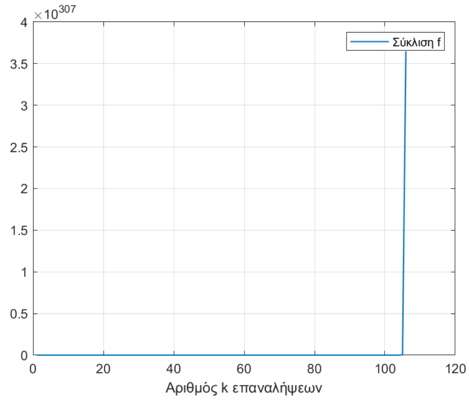

Figure 12: Σύγκλιση της f ανά k επαναλήψεις

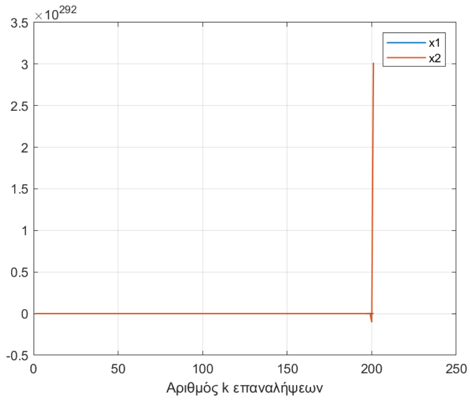

Figure 13: Μεταβολή x$_1, x_2$

Για $\gamma_k = 5$ όπως έχουμε αποδείξει παραπάνω ο αλγόριθμος δεν συγκλίνει.Η μεταβλητή y αποκλίνει πιό γρήγορα από την χ αφού στην ουσία $y_k = (-29)^k$,$x_k = (-2.333)^k$, το y στην ουσία έχει μεγαλύτερο ρυθμό απόκλισης.

2 Μέθοδος μέγιστης καθόδου με προβολή

Προκειμένου να μπορέσουμε να κάνουμε τον αλγόριθμο να συγκλίνει στο ολικό ελάχιστο θα χρησιμοποιήσουμε την μέθοδο προβολής.Το σύνολο

$$-10 < x_1 < 5$$

$$-8 < x_2 < 12$$

είναι κυρτό οπότε μπορούμε να εφαρμόσουμε το θεώρημα.Η λειτουργία του αλγορίθμου περιγράφεται ως εξής:

-Ξεκινάμε με ένα εφικτό σημείο x_0 που ανήκει στο κυρτό σύνολο.

-Ακολουθούμε αλγόριθμο ελαχιστοποίησης (στην προκειμένη περίπτωση μέγιστη κάθοδο)

-Αν το νέο σημείο που προκύπτει είναι εφικτό τότε συνεχίζουμε με την μέθοδο μέγιστης κλίσης.

-Εαν δεν είναι εφικτό βρίσκουμε την προβολή του σημείου στο κυρτό σύνολο και επαναλαμβάνουμε την διαδικασία.

Με τη μέθοδο προβολής στην ουσία μπορούμε να επιτύχουμε την σύγκλιση του αλγορίθμου χρησιμοποιώντας μεγαλύτερο βήμα γ ,ωστόσο αυτό δεν σημαίνει απαραίτητα ότι ο αλγόριθμος θα συγκλίνει ωστόσο θα παραμένουμε εντός του κυρτού συνόλου.Εαν για κάποια αλλη συνάρτηση που προσπαθούμε να κάνουμε ελαχιστοποίηση με μέθοδο προβολής και βλέπουμε ότι ο αλγόριθμος απόκλινει από το σύνολο τότε σημαίνει το ελάχιστο δεν είναι εντός του κυρτού συνόλου.Η τροποποίηση που θα ακολουθήσουμε είναι η εξής :

$$x_{k+1} = x_k + \gamma_k(\bar{x_k} - x_k)$$
$$y_{k+1} = y_k + \gamma_k(\bar{y_k} - y_k)$$

όπου

$$\bar{x_k} = Pr_X(x_k - s_k\nabla f(x_k, y_k)[1])$$
$$\bar{y_k} = Pr_X(y_k - s_k\nabla f(x_k, y_k)[2])$$

η

$$\bar{x_k} = Pr_X(x_k - \frac{2}{3}s_kx_k)$$
$$\bar{y_k} = Pr_X(y_k - 6s_ky_k)$$

$$\bar{x_k} = \begin{cases} -10 & \text{αν } x_k(1-\frac{2}{3}s_k) < -10, \\ x_k(1-\frac{2}{3}s_k) & \text{αν } -10 \leq x_k(1-\frac{2}{3}s_k) \leq 5, \\ 5 & \text{αν } x_k(1-\frac{2}{3}s_k) > 5. \end{cases}$$

$$\bar{y_k} = \begin{cases} -8 & \text{αν } y_k(1-6s_k) < -8, \\ y_k(1-6s_k) & \text{αν } -8 \leq y_k(1-6s_k) \leq 12, \\ 12 & \text{αν } y_k(1-6s_k) > 12. \end{cases}$$

Επομένως όταν θα έχουμε προβολή εντός του κυρτού συνόλου για το x_k και y_k θα έχουμε για την k+1 επανάληψη

$$x_{k+1} = x_k(1 - \frac{2}{3}s_k\gamma_k)$$
$$y_{k+1} = y_k(1 - 6s_k\gamma_k)$$

Συνεπώς όταν η προβολή βρεθεί εντός του κυρτού συνόλου θα έχουμε σύγκλιση στο ολικό ελάχιστο αν

$|\frac{x_{k+1}}{x_k}| < 1 => 0 < s_k\gamma_k < 3$

$|\frac{y_{k+1}}{y_k}| < 1 => 0 < s_k\gamma_k < \frac{1}{3}$.

2.1 x_0=[5,-5]$\gamma_k = 0.5$,$s_k = 5$

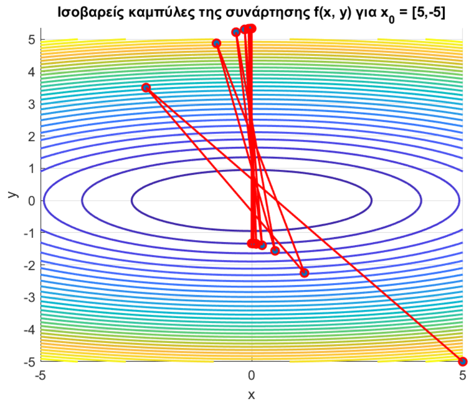

Figure 14: Σύγκλιση της f σε ισοβαρής καμπύλες

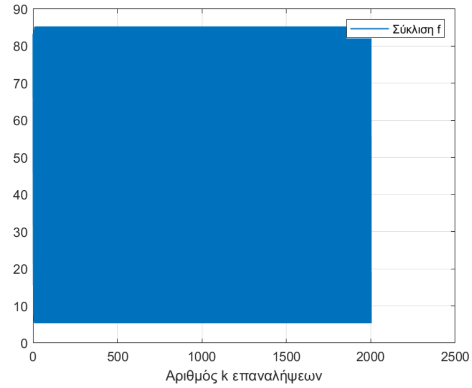

Figure 15: Σύγκλιση της f ανά k επαναλήψεις

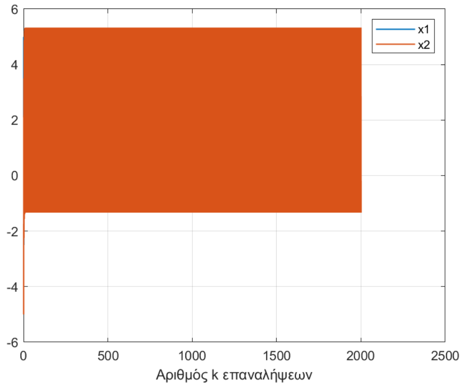

Figure 16: Μεταβολή x_1, x_2

Ο αλγόριθμος έχει ξεπεράσει το κατώφλι των 2000 επαναλήψεων και δεν συγκλίνει, οπότε τον τερματίζουμε για να μην έχουμε ατέρμονο βρόχο.Ο λόγος που συμβαίνει αυτό είναι επειδή η τιμή του y_k ταλαντώνεται μεταξύ των τιμών -1.333 και 5.333 ενώ το x_k συγκλίνει όπως περιμέναμε στο 0 αφού $s_k\gamma_k = 2.5$ άρα όταν η προβολή του $\bar{x_k}$ είναι εντός του κυρτού συνόλου θα ισχύει

$$x_{k+1} = x_k(-\frac{1}{3})$$

συνεπώς θα συγκλίνουμε ως προς x_k στο 0.Σε σχέση με το θέμα 1 αυτό που έχουμε καταφέρει είναι να κάνουμε τον αλγόριθμο να μην αποκλίνει και ταυτόχρονα να μένει εντός του κυρτού συνόλου X. Με μια πιο προσσεκτική ματιά στα παρακάτω διαγράμματα επιβεβαιώνονται τα παραπάνω.

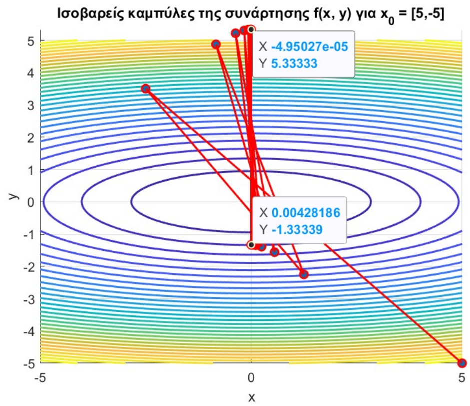

Figure 17: Σύγκλιση της f σε ισοβαρής καμπύλες

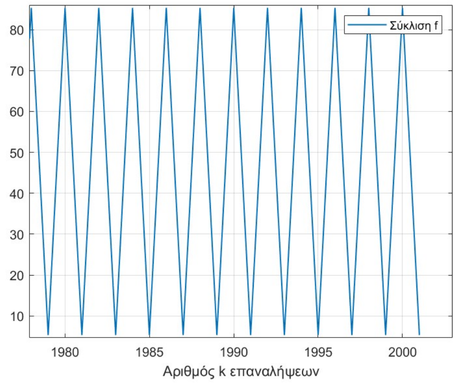

Figure 18: Σύγκλιση της f ανά k επαναλήψεις

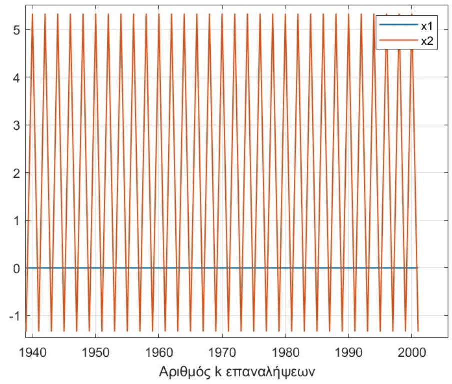

Figure 19: Μεταβολή x_1, x_2

Για $y_k = 5.3333$, η προβολή ισούται με $\bar{y_k} = -8$ αφού $5.333 * (1 - 6 * 5) < -8$ οπότε $y_{k+1} = 5.333 + 0.5 * (-8 - 5.3333) = -1.333$

για $y_k = -1.333$ $\bar{y_k} = 12$ αφού $-1.333 * (1 - 6 * 5) > 12$ άρα $y_{k+1} = -1.333 + 0.5 * (12 + 1.333) = 5.333$, συνεπώς το y θα εναλλάσεται μεταξύ των τιμών [-1.333,5.333] οπότε και ο αλγόριθμος δεν θα τερματίσει.

2.2 x_0=[-5,10]$\gamma_k = 0.1$,$s_k = 15$

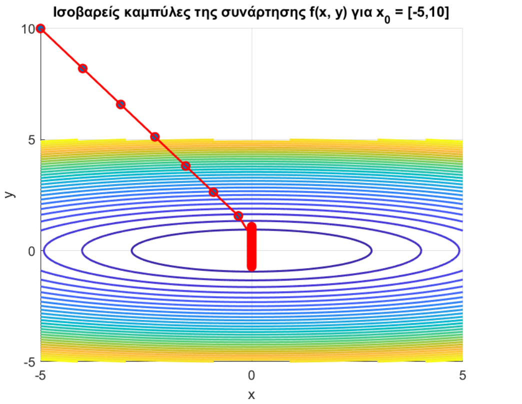

Figure 20: Σύγκλιση της f σε ισοβαρής καμπύλες

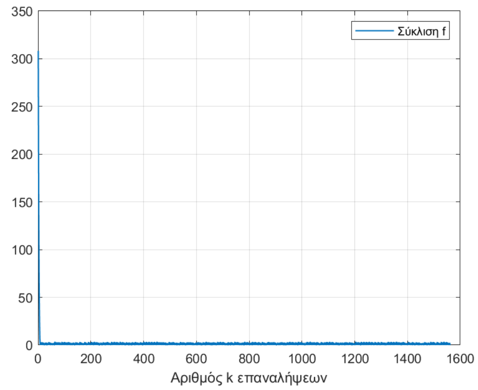

Figure 21: Σύγκλιση της f ανά k επαναλήψεις

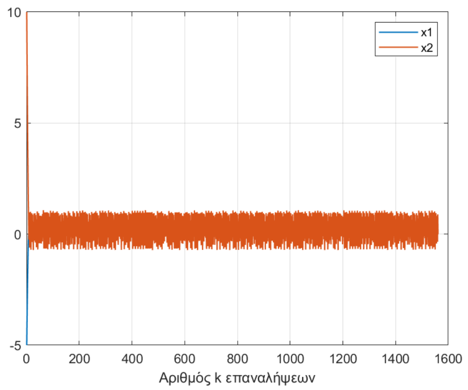

Figure 22: Μεταβολή x$_1, x_2$

Παρατηρούμε πως ο αλγόριθμος σε σχέση με το θέμα 1 και θέμα 2 συγκλίνει στο ολικό ελάχιστο ωστόσο με μεγάλο αριθμό επαναλήψεων,συνολικά 1542.Το x συγκλίνει σε πολύ μικρό αριθμό επαναλήψεων όπως φαίνεται και στο διάγραμμα με τις ισοβαρής καμπύλες , συγκεκριμένα 9 επαναλήψεις,αυτό οφείλεται στο ότι έχουμε επιλέξει τον βέλτιστο συνδυσμό s_k και γ_k καθώς όταν η προβολή $\bar{x_k}$ βρεθεί εντός του κυρτού συνόλου θα έχουμε

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

με $s_k\gamma_k = 1.5$,$x_{k+1} = 0$ οπότε θα βρεθούμε σε μία επανάληψη στο 0 για το x.Αυτό δεν ισχύει για το y που στην ουσία χρειάζεται μεγάλο αριθμό επαναλήψεων για να συγκλίνει στο 0.

α)Πρώτος τρόπος

Προκείμενου να μπορέσουμε να συγκλίνουμε πιό γρήγορα στο (0,0) μπορούμε να ακολουθήσουμε αντίστοιχη διαδικασία για το y δηλαδη:

$$y_{k+1} = y_k(1 - 6s_k\gamma_k)$$

, άρα κρατώντας το γ=0.1 σταθερό και επιλέγωντας $s_k = 10/6$, όταν η προβολή του y βρεθεί εντός του κυρτού συνόλου το y θα μηδενιστεί στην αμέσως επόμενη επανάληψη.Επίσης εφόσομ $\gamma_k s_k < 0.3333$ ξέρουμε ότι και το x_k θα συγκλίνει στο 0 σε κάποιο αριθμό επαναλήψεων,στην ουσία δλδ έχουμε επιταχύνει την σύγκλιση κάνοντας το y να συγκλίνει πιο γρήγορα.

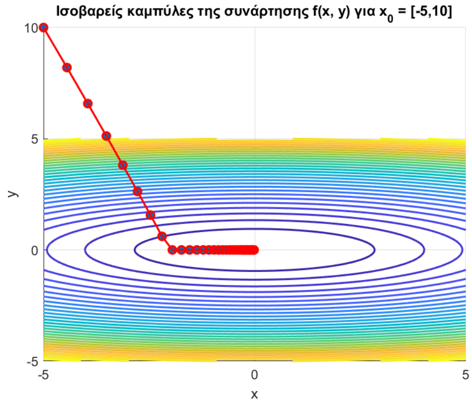

Figure 23: Σύγκλιση της f σε ισοβαρής καμπύλες

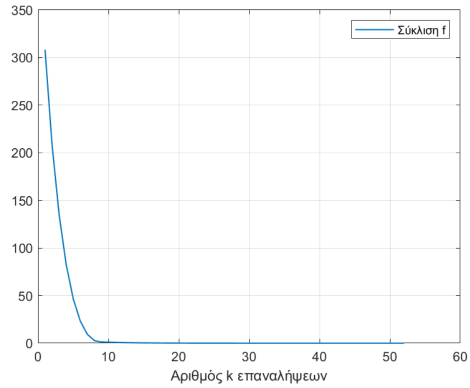

Figure 24: Σύγκλιση της f ανά k επαναλήψεις

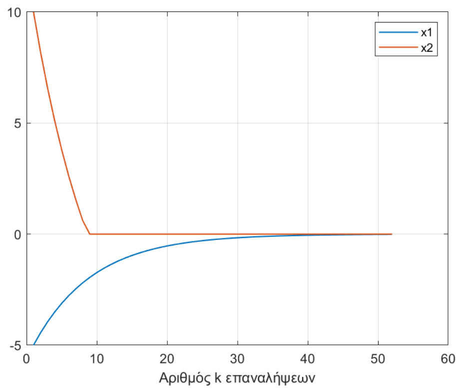

Figure 25: Μεταβολή x_1, x_2

Ο συνολικός αριθμός επαναλήψεων που προκύπτει είναι k=52 μικρότερος από πριν.Επίσης εάν επιλέγαμε γ=0.1 με μέθοδο μέγιστης καθόδου χωρίς προβολή με το ίδιο σημείο εκκίνησης ο αλγόριθμος θα τερμάτιζε σε 120 επαναλήψεις.Όπως παρατηρούμε με αυτόν τον τρόπο το y συγκλίνει γρήγορα στο 0 καθώς όταν το $\bar{y_k}$ ανήκει στο κυρτό σύνολο θα χρειαστεί μια επανάληψη ώστε το y να μηδενιστεί.

β)Δεύτερος τρόπος

Επίσης μπορούμε να βάλουμε τα x,y να συγκλίνουν με τον ίδιο ρυθμό σύγκλισης δηλαδή

$$|1 - \frac{2}{3} s_k \gamma_k| = |1 - 6 s_k \gamma_k|$$

οπότε προκύπτει λύση $s_k \gamma_k = 0.3$ και επιλέγοντας $s_k = 3$,$\gamma_k = 0.1$ ο αλγόριθμος συγκλίνει στο ολικό ελάχιστο με 33 επαναλήψεις!!!

Οπότε ανά k επανάληψη έχουμε

$x_{k+1} = 0.8 x_k$,και για

$y_{k+1} = -0.8 y_k$,και για ,συνεπως τα x,y συγκλίνουν με το ίδιο ρυθμό κατά απόλυτη τιμή στο (0,0) οπότε όσες επαναλήψεις χρειαστεί το y για να μηδενιστεί αντίστοιχο αριθμό θα χρειαστεί και το x (ελαφρώς διαφορετικό όμως επειδή δεν έχουμε $x_0 = y_0$).

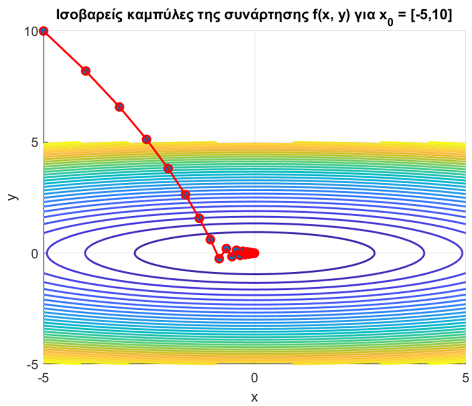

Figure 26: Σύγκλιση της f σε ισοβαρής καμπύλες

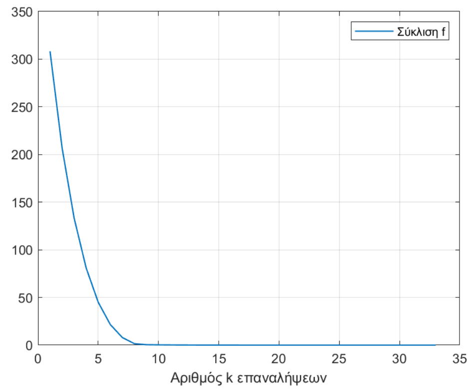

Figure 27: Σύγκλιση της f ανά k επαναλήψεις

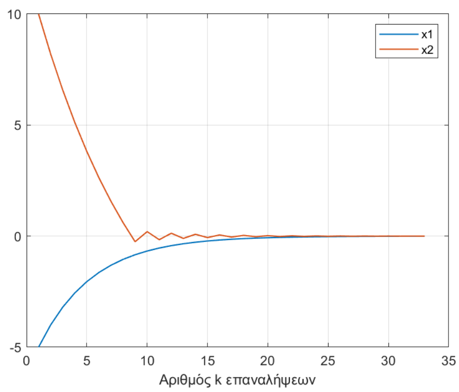

Figure 28: Μεταβολή x$_1$, x_2

γ)Τρίτος τρόπος

Ένας άλλος τρόπος για να συγκλίνουμε στο ελάχιστο πιο γρήγορα θα ήταν αυτή τη φορά αντί να αλλάξουμε το βήμα γ ή το s_k,να αλλάξουμε το σημείο εκκίνησης .Έστω $x_0 = [-5, 0]$ τότε για κάθε επανάληψη θα είχαμε $y_k = 0$ ενώς με $s_k = 15$,και $\gamma_k = 0.1$,(αυτά που έχουμε στην εκφώνηση) το x_k θα είχε παρόμοια συμπεριφορά με αυτην που αναλύσαμε προηγουμένως και θα μηδενίζοταν σε 9 επαναλήψεις.Οπότε σε 9 επαναλήψεις θα είμαστε στο ολικό ελάχιστο.

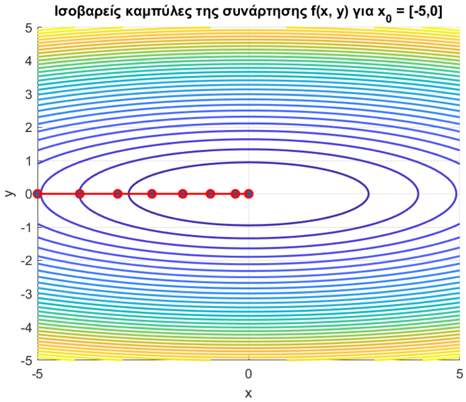

Figure 29: Σύγκλιση της f σε ισοβαρής καμπύλες

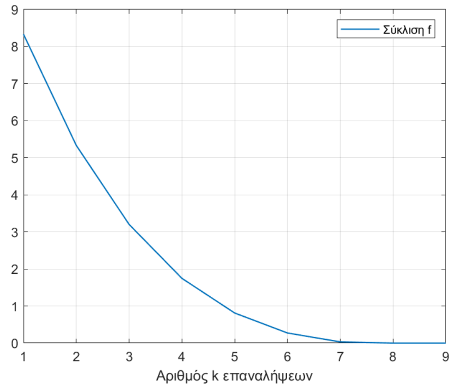

Figure 30: Σύγκλιση της f ανά k επαναλήψεις

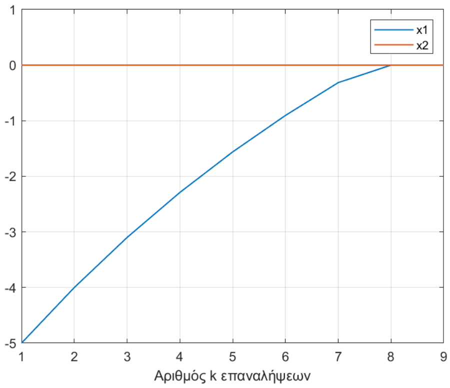

Figure 31: Μεταβολή x_1, x_2

2.3 x_0=[8,-10]$\gamma_k = 0.2$,$s_k = 0.1$

Για τα συγκεκριμένα s και γ περιμένουμε ότι αλγόριθμος θα σύγκλινει στο ολικό ελάχιστο μετά από καποιον αριθμό επαναλήψεων.Όταν τα $\bar{x_k}$ και $\bar{y_k}$ έρθουν εντός του κυρτού συνόλου θα ισχύει ότι :

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

$$y_{k+1} = y_k(1 - 6s_k\gamma_k)$$

,οπότε θα έχουμε σίγουρα σύγκλιση στο (0,0) αφού $\gamma_k s_k = 0.02 < 0.333$. Κάνοντας αντικατάσταση,

$$x_{k+1} = x_k(0.9866)$$

$$y_{k+1} = y_k(0.88)$$

Συνεπώς θα έχουμε εκθετική σύγκλιση στο 0.Όπως φαίνεται από τις εξίσωσεις το y καθώς έχει μικρότερη βάση συγκλίνει στο 0 πιο γρήγορα από το x,πράγμα που επιβεβαιώνεται στην γραφική παράσταση.Επίσης αξίζει να τονίσουμε ότι το x_k μειώνεται με πολύ αργό ρυθμό οπότε περιμένουμε σχετικά μεγάλο αριθμό επαναλήψεων.

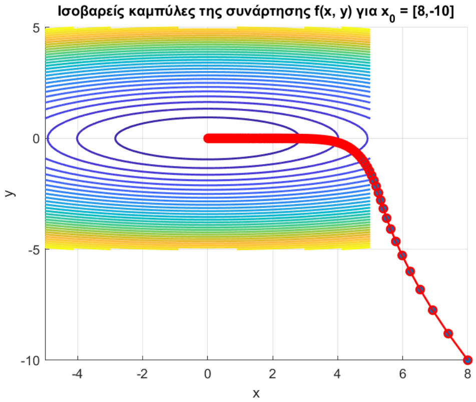

Figure 32: Σύγκλιση της f σε ισοβαρής καμπύλες

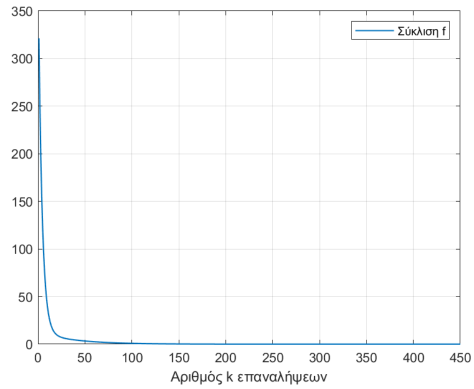

Figure 33: Σύγκλιση της f ανά k επαναλήψεις

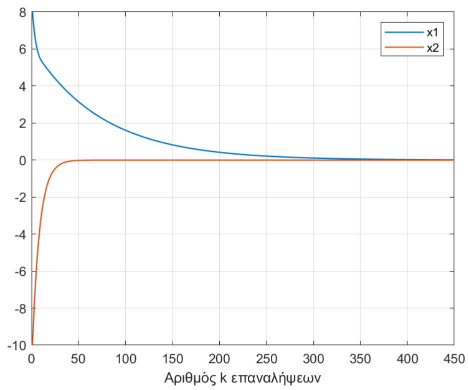

Figure 34: Μεταβολή x_1, x_2

Ο συνολικός αριθμός επαναλήψεων είναι 450.Επειδή το βήμα γ_k και το s_k είναι πολύ μικρά αριθμητικά περιμέναμε ούτως η άλλως αργή σύγκλιση.