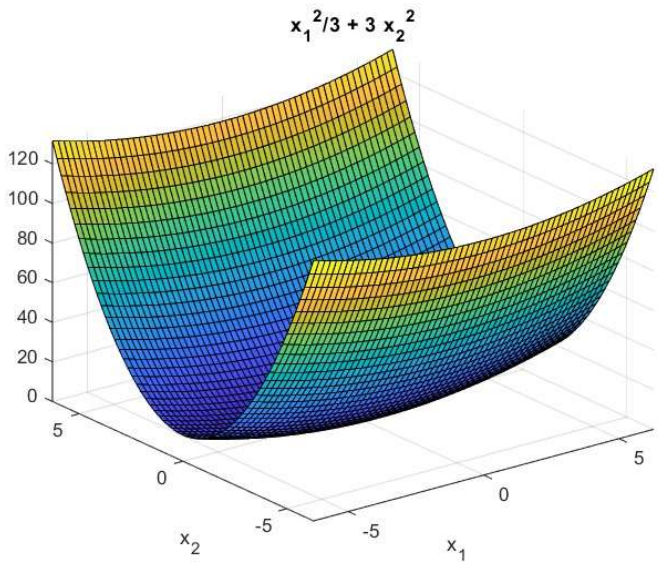
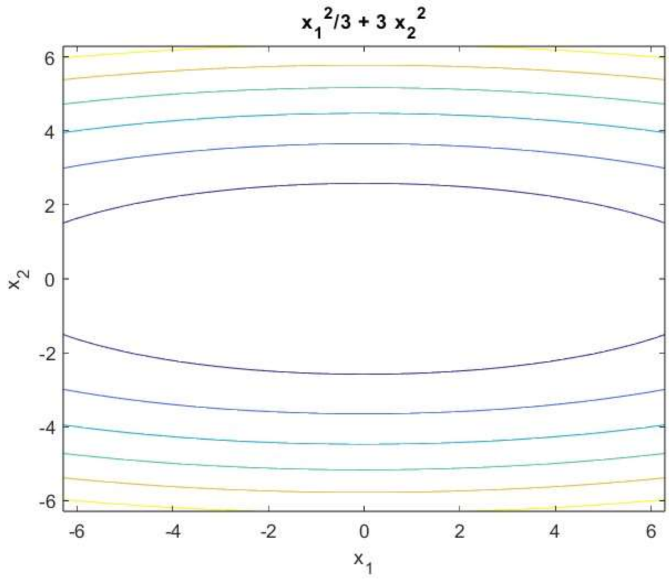
3η Εργαστηριακή Άσκηση:

Μέθοδος Μέγιστης Καθόδου με Προβολή

Σκοπός μας είναι να ελαχιστοποιήσουμε την συνάρτηση:

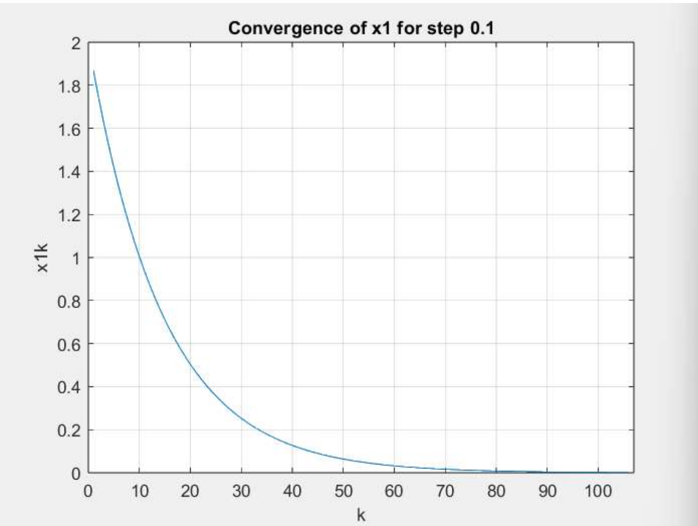
$$f(x) \;=\; \frac{1}{3}\,x_1^2 + 3\,x_2^2$$

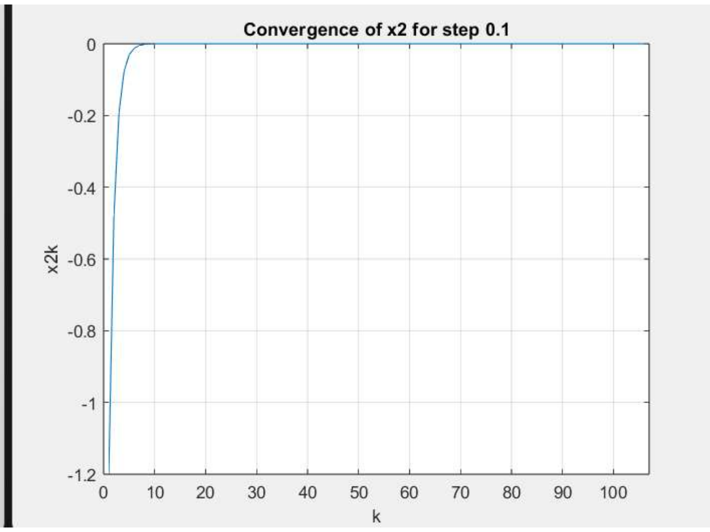




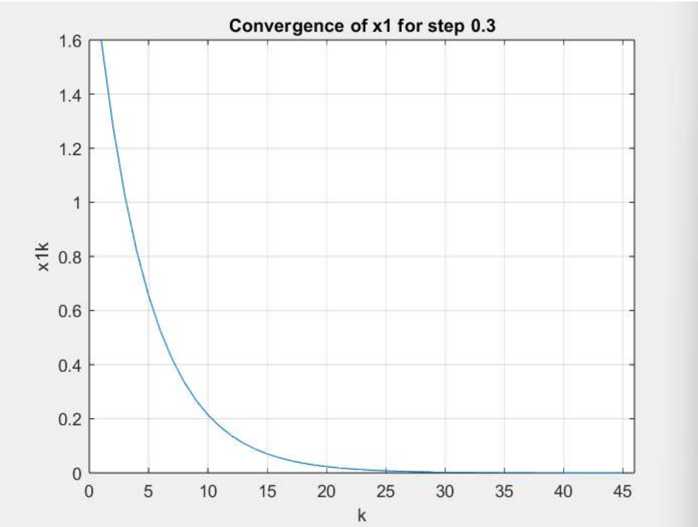
Θέμα 1

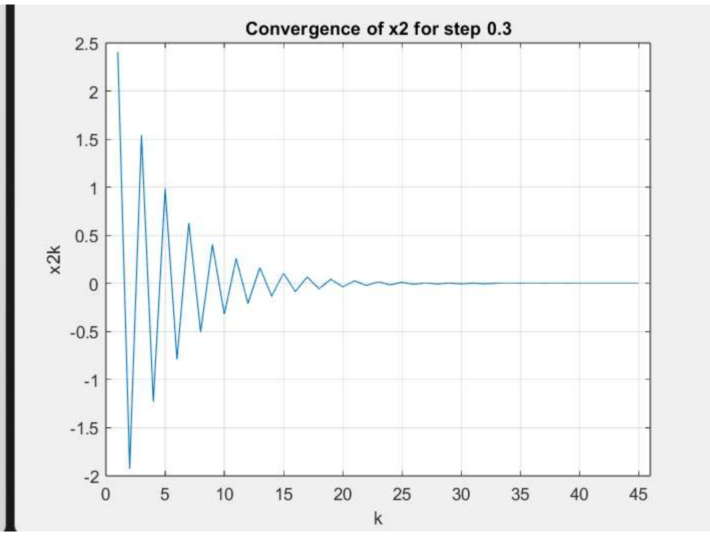
Προηγουμένως για αυτόν τον σκοπό χρησιμοποιήσαμε τον *αλγόριθμο της μέγιστης καθόδου* όπου σε κάθε επανάληψη βρίσκαμε ένα σημείο όλο και πιο κοντά στο ελάχιστο μέσω της διαδρομής που επέφερε μεγαλύτερη μείωση του gradient. Κάνοντας πάλι το ίδιο για βήμα 0,1, 0.3, 3, 5 και με τυχαίο αρχικό σημείο το [2, -3] παίρνουμε τα εξής γραφήματα.





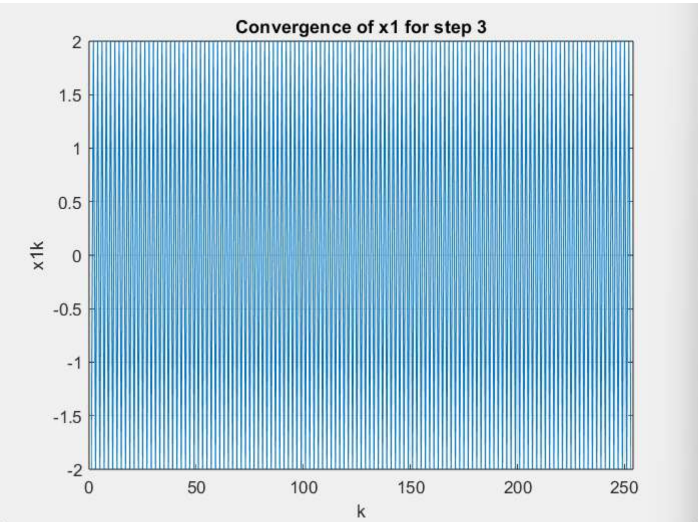
Με αποτέλεσμα 5.9256e-07

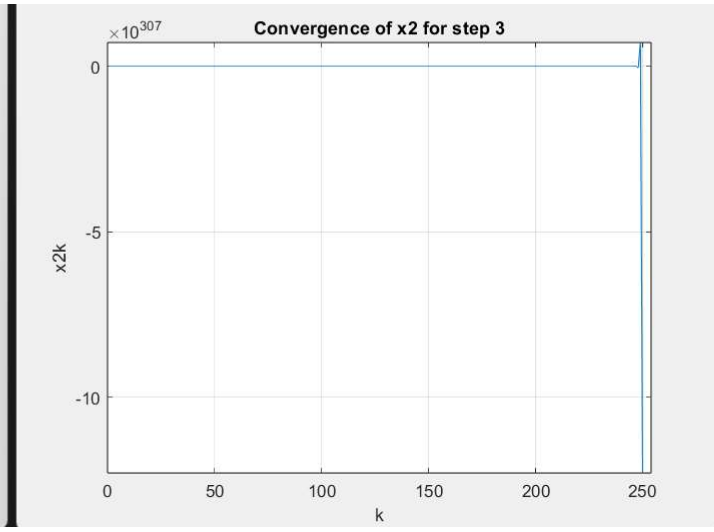


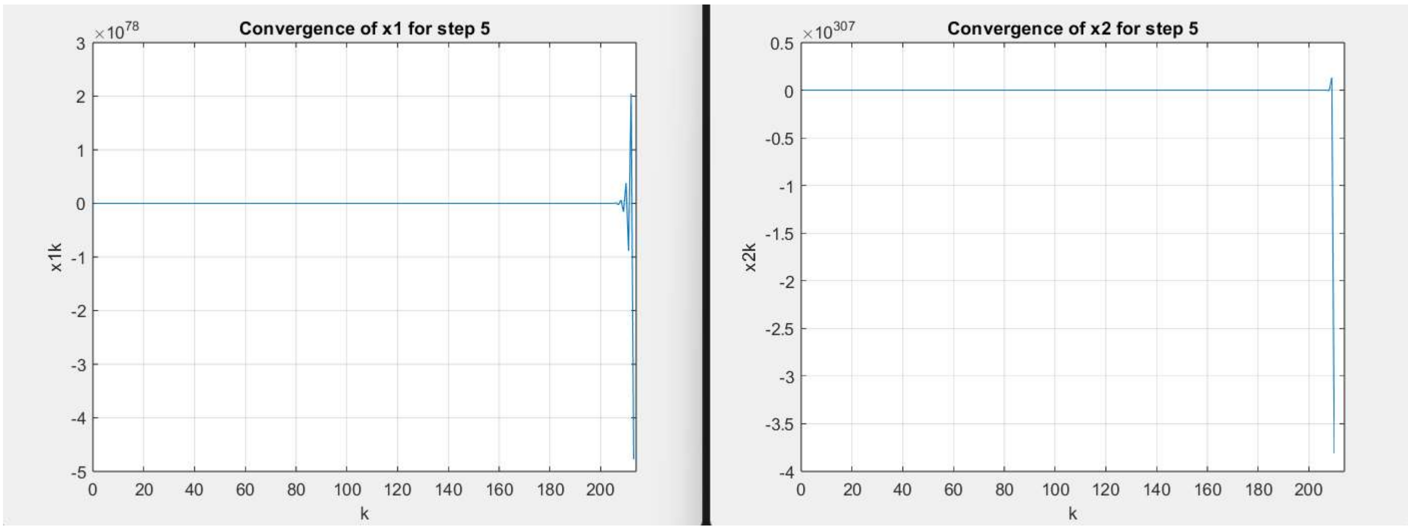


Με αποτέλεσμα 1.0e-03

Παρατηρούμε ήδη ότι όσο μεγαλώνει το βήμα τόσο πιο μακριά πηγαίνουμε από τα σημεία [0, 0] με τιμή 0 που είναι το πραγματικό μας ελάχιστο. Στον αλγόριθμο έχουμε ως χαρακτηριστικό να τερματίζει τη λούπα αν η τιμή της συνάρτησης στο νέο x είναι μεγαλύτερη της προηγούμενης, αφού αυτό έρχεται σε άμεση αντίθεση με τη μέθοδο της μέγιστης καθόδου. Ωστόσο για να πάρουμε μια ιδέα από το πόσο χαώδη γίνονται τα αποτελέσματα την απενεργοποιούμε και βλέπουμε τα εξής.







Όπως βλέπουμε, τα αποτελέσματα διαφοροποιήθηκαν από τα προηγούμενα ραγδαία, και αυτό γιατί με τόσο μεγάλο βήμα ο αλγόριθμος μας είναι πολύ ασταθής και είτε προκαλεί τα σημεία x_1 και x_2 σε ταλάντωση είτε τα ωθεί μακριά από τη λύση.

Για να αποδείξουμε τα ευρήματα μας παίρνουμε το gradient της αρχικής συνάρτησης και τον τύπο της μέγιστης καθόδου.

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

Βρίσκουμε ότι $\nabla f(x_k) = [\frac{2}{3}x_{1k}, 6x_{2k}]$

Άρα $x_{k+1} = [x_{1k} - \gamma_k \frac{2}{3}x_{1k}, x_{2k} - \gamma_k 6x_{2k}]$

ή καλύτερα $x_{(1)k+1} = (1 - \frac{2}{3}\gamma_k) x_{1k}$ και $x_{(2)k+1} = (1 - 6\gamma_k) x_{2k}$.

Είναι εύκολο να αποδείξουμε επαγωγικά ότι $x_{1k} = (1 - \frac{2}{3}\gamma_k)^k x_{(1)0}$ (1) και αντίστοιχα $x_{2k} = (1 - 6\gamma_k)^k x_{(2)0}$ (2).

Για τη σύγκλιση χρειαζόμαστε $\frac{|x_{k+1}|}{|x_k|} < 1$. Από επίλυση αυτού χρησιμοποιώντας τις σχέσεις (1) και (2) καταλήγουμε ότι για τη σύγκλιση του x_1 πρέπει $0 < \gamma_k < 3$ και για του x_2 πρέπει $0 < \gamma_k < \frac{1}{3}$.

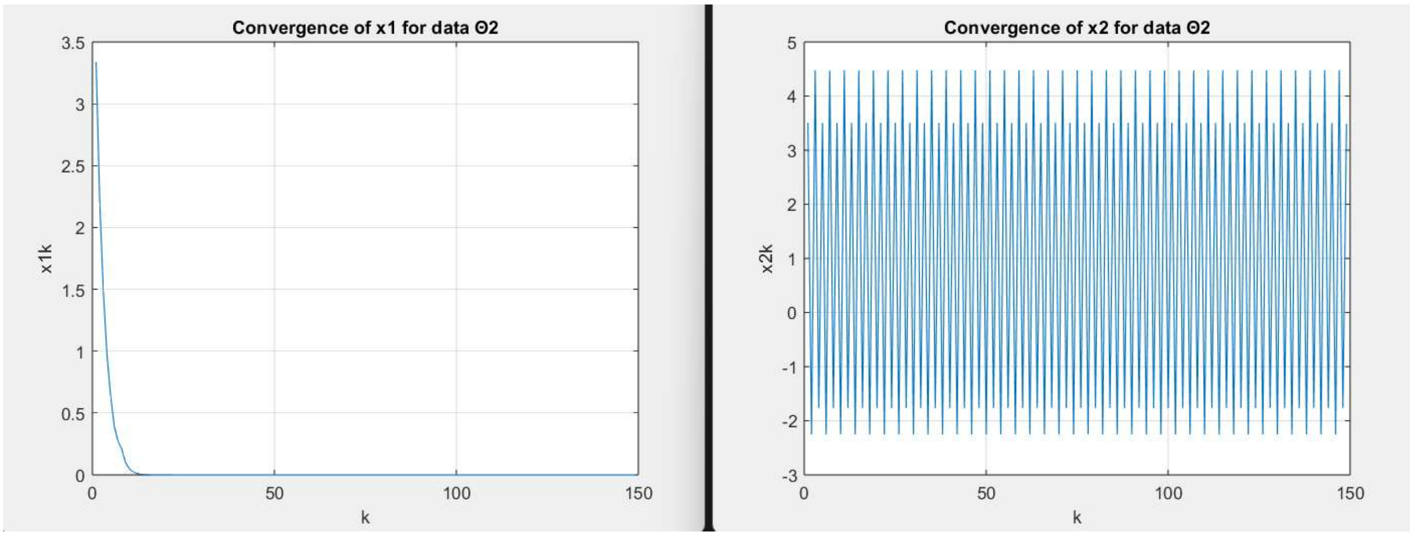
Από αυτό καταλαβαίνουμε γιατί για $\gamma_k = 0.1$ ο αλγόριθμος συγκλίνει τέλεια ενώ όταν $\gamma_k = 0.3$ -το άνω όριο του επιθυμητού βήματος για το x_2- και αντίστοιχα όταν $\gamma_k = 3$ -το άνω όριο του επιθυμητού βήματος για το x_1- οι αντίστοιχες μεταβλητές ξεκινούν να ταλαντώνονται γύρω από τις επιθυμητές τιμές. Για βήμα

5, ο αλγόριθμος πλέον ξεκάθαρα δεν μπορεί να συγκλίνει αφού είναι πολύ εκτός των ορίων του γ_k.

Τώρα θα δοκιμάσουμε να χρησιμοποιήσουμε τον *αλγόριθμο της μέγιστης καθόδου με προβολή*. Έχουμε πλέον περιορισμούς στο πρόβλημα ελαχιστοποίησης και πιο συγκεκριμένα: $-10 \leq x_1 \leq 5$ και $-8 \leq x_2 \leq 12$. Επομένως θέλουμε να μπορούμε να προβάλλουμε οποιοδήποτε σημείο προκύψει, ακόμα και ένα μη εφικτό, σε αυτό το σύνολο τιμών ώστε να μπορέσει ο αλγόριθμος να συγκλίνει στο ελάχιστο. Για αυτόν τον σκοπό βρίσκουμε την προβολή της κατεύθυνσης αναζήτησης και κινούμαστε κάθε φορά πάνω σε αυτή με κάποιο βήμα.

Θέμα 2

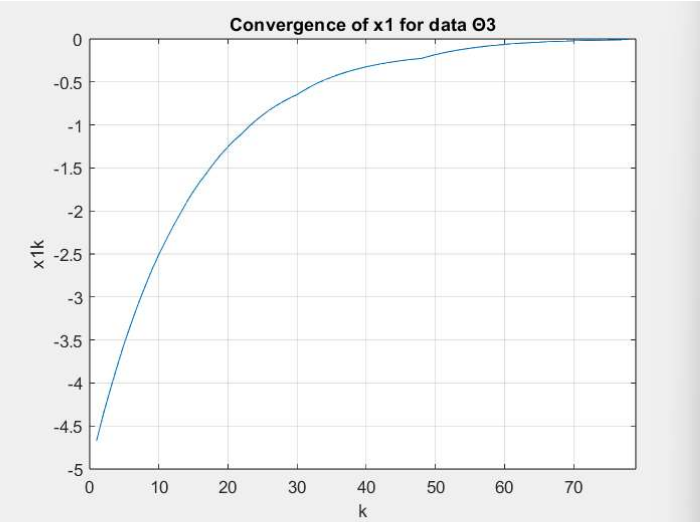
Τα δεδομένα μας είναι: $s_k = 5$, $\gamma_k = 0.5$, $\varepsilon = 0.01$ και το αρχικό σημείο επιλέγεται το (5, -5). Ήδη παρατηρούμε πως το $x_1 = 5$ βρίσκεται στα όρια της περιοχής, καθώς και πως το βήμα της κατεύθυνσης αναζήτησης s_k έχει μια αρκετά μεγάλη τιμή, τα οποία ίσως προκαλέσουν πρόβλημα στον αλγόριθμο. Πράγματι, τρέχοντας τον αλγόριθμο:

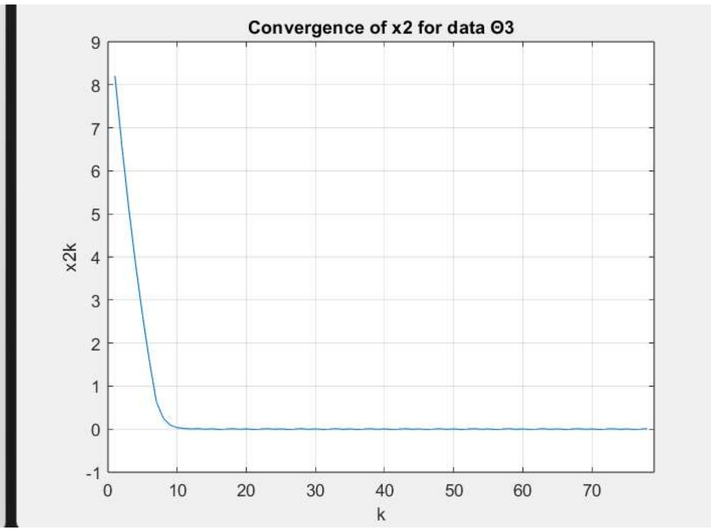


Η μέθοδος δεν καταφέρνει να συγκλίνει και ενώ το x_1 φτάνει σχετικά ομαλά στο σημείο 0 το x_2 ταλαντώνεται συνεχώς. Επομένως, σε αυτή την περίπτωση η χρήση της προβολής δεν οδηγεί στην επιθυμητή λύση, που συνάδει με όσα αποδείξαμε στο Θέμα 1 για το βήμα.

Θέμα 3

Τα δεδομένα μας είναι: $s_k = 15$, $\gamma_k = 0.1$, $\varepsilon = 0.01$ και το αρχικό σημείο επιλέγεται το (-5, 10). Εδώ το αρχικό σημείο είναι ήδη εφικτό, ωστόσο το βήμα s_k είναι ακόμα μεγαλύτερο. Τρέχοντας τον αλγόριθμο:

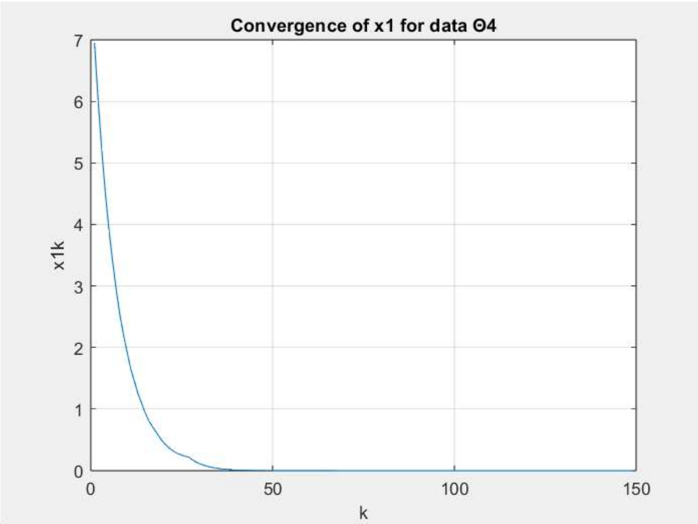


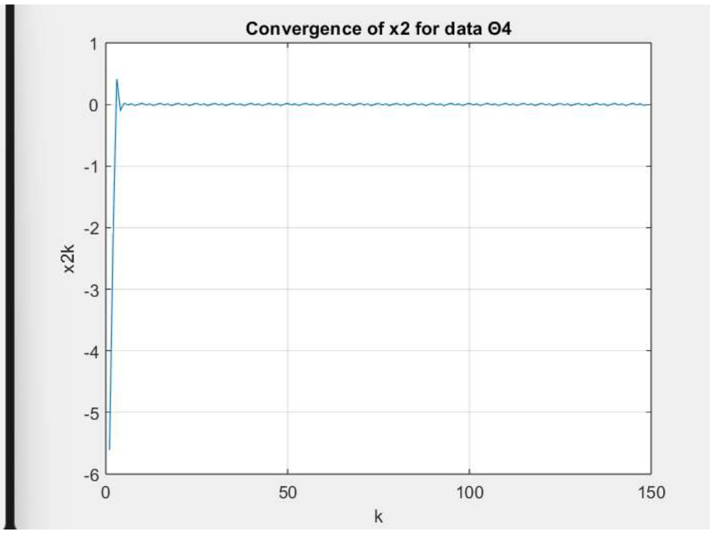


Τελικά τα σημεία x_1 και x_2 συγκλίνουν σε όχι πολλές επαναλήψεις στο (0, 0) και το ελάχιστο υπολογίζεται τελικά 0, όπως και θέλαμε. Επομένως καταλαβαίνουμε πως η αρχική μας υπόθεση για το s_k ήταν λάθος και τελικά το μέγεθος του δεν επηρεάζει αρνητικά την ευστάθεια του αλγορίθμου. Αξίζει ωστόσο να σημειωθεί ότι και το βήμα γ_k ήταν μικρότερο αυτή τη φορά.

Θέμα 4

Τα δεδομένα μας είναι: $s_k = 0.1$, $\gamma_k = 0.2$, $\varepsilon = 0.01$ και το αρχικό σημείο επιλέγεται το (8, -10). Τα βήματα είναι για πρώτη φορά και τα δύο αρκετά μικρά, ωστόσο είναι προφανές ότι το νέο αρχικό σημείο είναι εκτός του εφικτού συνόλου. Γενικά, ο αλγόριθμος υποθέτει ότι το αρχικό σημείο είναι εφικτό, επομένως αυτό το γεγονός προμηνύει πώς ο αλγόριθμος πιθανότατα θα δυσκολευτεί παραπάνω να συγκλίνει. Πράγματι, τρέχοντας τον αλγόριθμο:





Παρατηρούμε πως ο αλγόριθμος δεν συγκλίνει ακριβώς αλλά φτάνει σε ένα σημείο το x_2 να κάνει ταλάντωση σε μια πολύ μικρή περιοχή γύρω από το 0. Έτσι, παρότι χωρίς κάποια έξτρα συνθήκη τερματισμού (που έχουμε φροντίσει να υπάρχει) δεν θα καταφέρει να βγει από τη λούπα, το τελικό αποτέλεσμα προσεγγίζει πολύ κοντά το επιθυμητό.