Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
Τμήμα Ηλεκτρολόγων Μηχανικών & Μηχανικών Υπολογιστών

Τεχνικές Βελτιστοποίησης
Εργασία 3

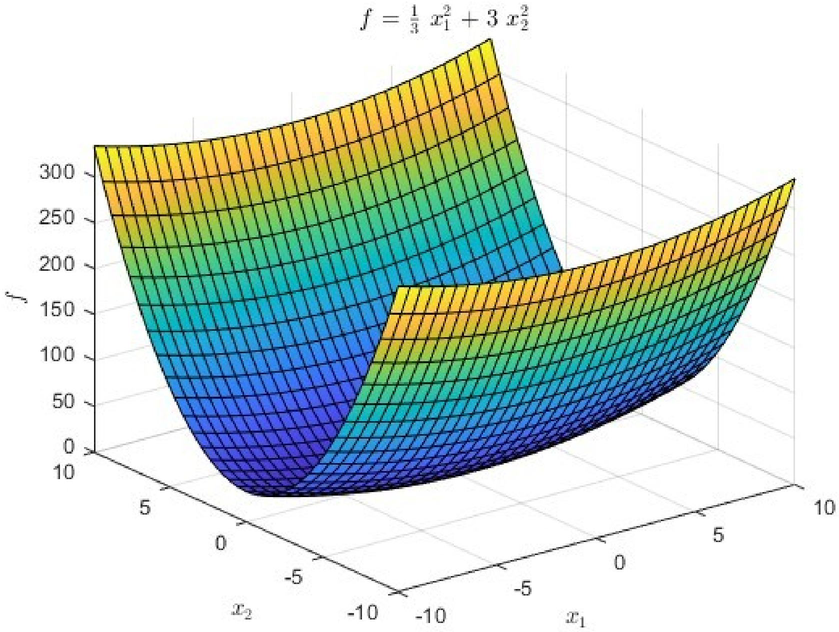
Χριστοφορίδης Χρήστος – ΑΕΜ 10395
christoscs@ece.auth.gr

29/11/2024

Θέμα 1

Ζητείται να χρησιμοποιήσουμε την μέθοδο της μεγίστης καθόδου (από την προηγούμενη εργασία) με ακρίβεια ε = 0.001 και διάφορα σταθερά γ_k με σημείο εκκίνησης της επιλογής μας στην συνάρτηση $f(x)=\frac{1}{3}x_1^2+3x_2^2,\ x=[x_1\ \ x_2]^T$

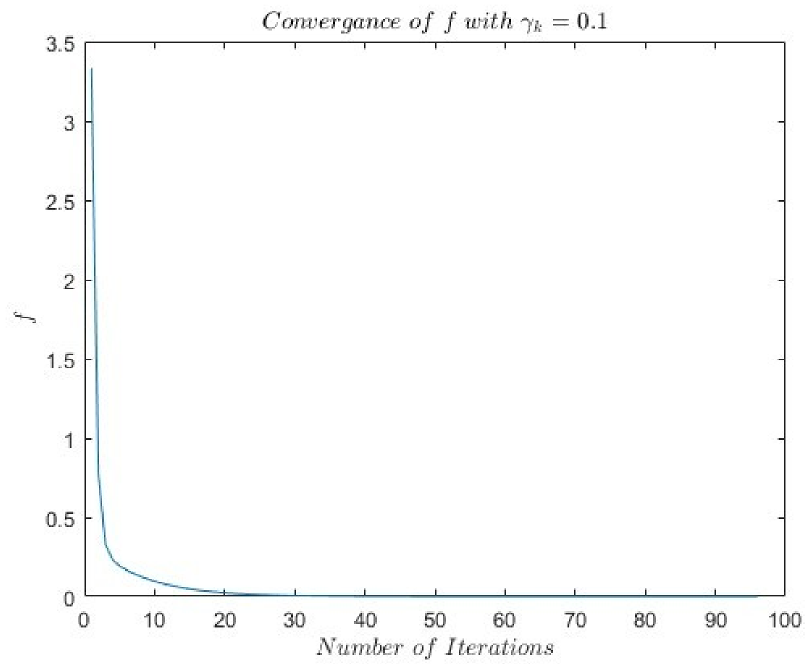
Παρακάτω φαίνεται η γραφική παράσταση της συνάρτησης f



Επιλέχτηκε ως αρχικό σημείο το (1,1) για κάθε περίπτωση.

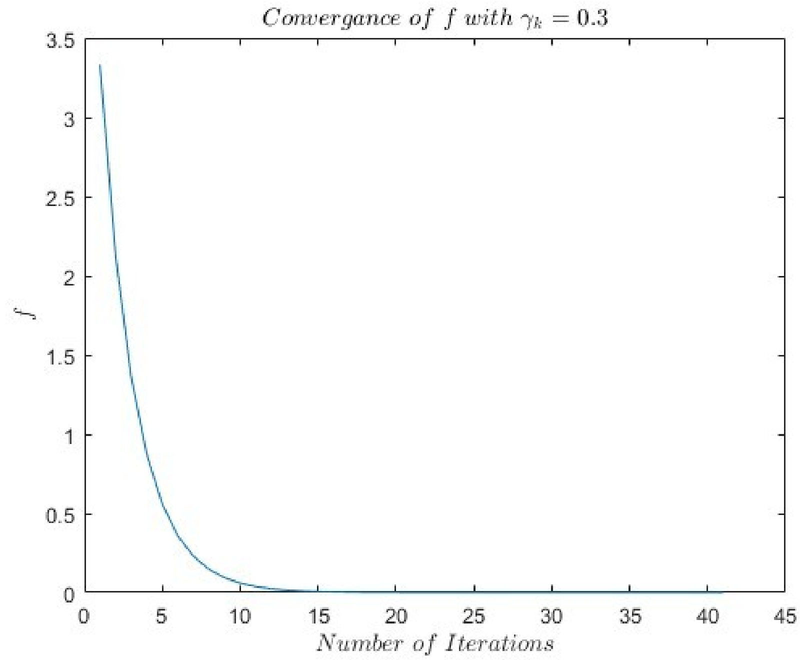
Τρέχω το script steepest_descent.m και παίρνω τα παρακάτω αποτελέσματα:

i) Για βήμα γ_k = 0.1



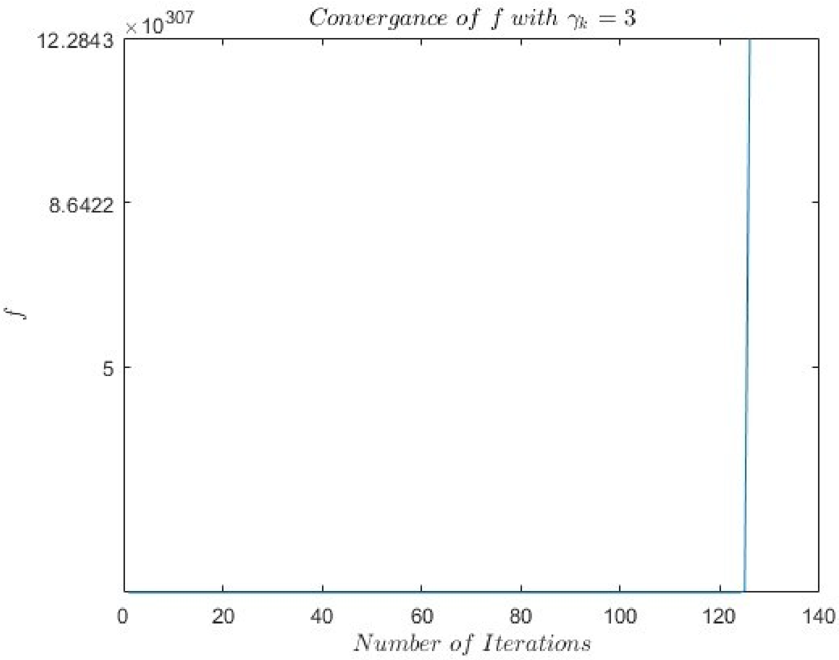
Παρατηρώ ότι η συνάρτηση f συγκλίνει σε μία τιμή κοντά στο 0 έπειτα από εύλογο αριθμό επαναλήψεων.

ii) Για βήμα $\gamma_k = 0.3$



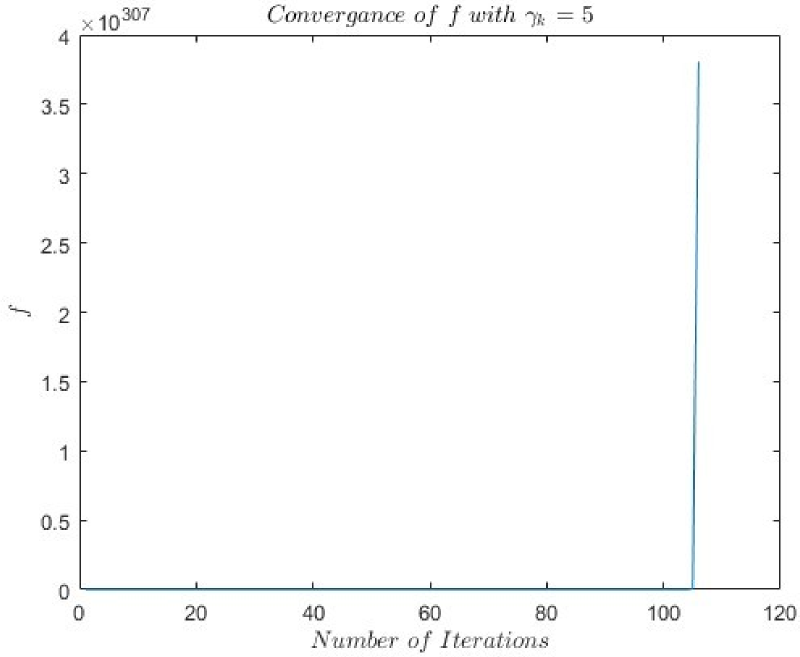
Παρατηρώ ότι η συνάρτηση f και πάλι συγκλίνει σε μία τιμή κοντά στο 0 έπειτα από εύλογο αριθμό επαναλήψεων (μικρότερο από ότι στην πρώτη περίπτωση).

iii) Για βήμα $\gamma_k = 3$



Παρατηρώ ότι η συνάρτηση f δεν συγκλίνει αντιθέτως αποκλίνει απότομα και απειρίζεται (η μέθοδος δεν λειτουργεί).

iv) Για βήμα $\gamma_k = 5$



Παρατηρώ και πάλι ότι η συνάρτηση f δεν συγκλίνει αντιθέτως αποκλίνει απότομα και απειρίζεται (η μέθοδος δεν λειτουργεί).

Θα εξηγήσουμε τα παραπάνω αποτελέσματα κάνοντας την αντίστοιχη μαθηματική ανάλυση.

Έχω:

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

Άρα η μέθοδος είναι:

$$\begin{bmatrix} x_{1,k+1} \\ x_{2,k+1} \end{bmatrix} = \begin{bmatrix} x_{1,k} \\ x_{2,k} \end{bmatrix} - \gamma_k \cdot \begin{bmatrix} \frac{2}{3} x_{1,k} \\ 6 x_{2,k} \end{bmatrix}$$

Δηλαδή:

$$x_{1,k+1} = x_{1,k} - \gamma_k \cdot \frac{2}{3} x_{1,k} \quad \text{και}$$

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

Ή αλλιώς:

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

$$x_{2,k+1} = (1 - 6\gamma_k) x_{2,k}$$

Για να έχουμε σύγκλιση της μεθόδου πρέπει να ισχύει:

$$\left|\frac{x_{k+1}}{x_k}\right|<1$$

Άρα από x_1 έχω:

$$\left|1-\frac{2}{3}\gamma_k\right|<1$$

$$\Rightarrow\ -1<1-\frac{2}{3}\gamma_k<1$$

$$\Rightarrow\ 0<\gamma_k<3$$

και από x_2:

$$|1-6\gamma_k|<1$$

$$\Rightarrow\ -1<1-6\gamma_k<1$$

$$\Rightarrow\ 0<\gamma_k<\frac{1}{3}$$

Επομένως για να συγκλίνει η μέθοδος πρέπει να ισχύει $0<\gamma_k<\frac{1}{3}\approx 0.33$

Για αυτό τον λόγο η μέθοδος συγκλίνει μόνο στις πρώτες δύο περιπτώσεις.

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

Στα παρακάτω θέματα θεωρούμε τους εξής περιορισμούς:

$$-10\leq x_1\leq 5 \quad \text{και} \quad -8\leq x_2\leq 12$$

Εφαρμόζουμε την παρακάτω σχέση

$$x_{k+1}=x_k+\gamma_k(\bar{x}_k-x_k)$$

όπου

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

Αν το x_k είναι εφικτό, δηλαδή τηρεί τους περιορισμούς, τότε $\bar{x}_k=x_k-s_k\nabla f(x_k)$ και επομένως προκύπτει:

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

Άρα έχω νέο $\gamma_k{}'=\gamma_k s_k$ για το οποίο από την ανάλυση στο θέμα 1 θα πρέπει να ισχύει ότι $\gamma_k{}'\in(0,\frac{1}{3})$

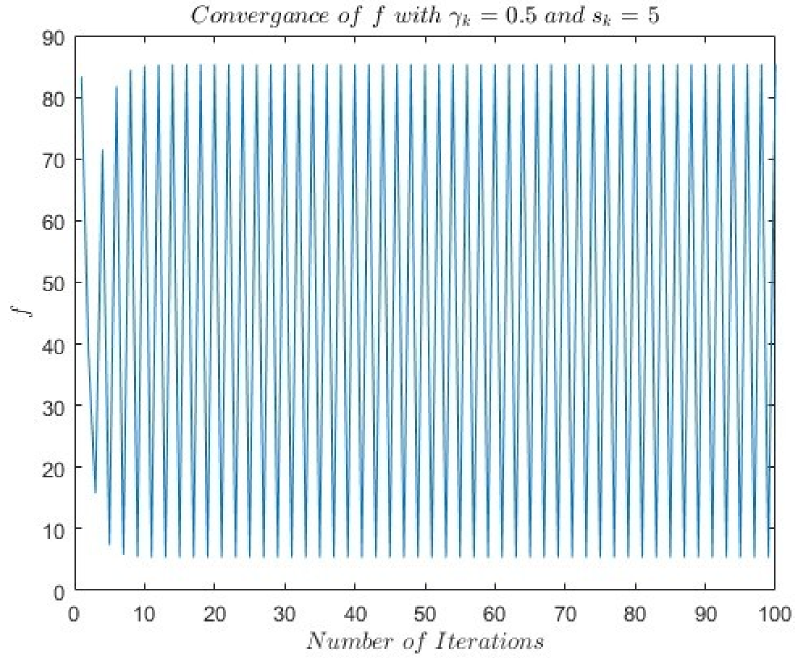
Από τους περιορισμούς για τις προβολές έχω:

$$\Pr\{x_1\}=\begin{cases}-10, & \text{αν } x_1\leq -10\\ 5, & \text{αν } x_1\geq 5\\ x_1 & \text{αλλόυ}\end{cases} \qquad \Pr\{x_2\}=\begin{cases}-8, & \text{αν } x_2\leq -8\\ 12, & \text{αν } x_2\geq 12\\ x_2 & \text{αλλόυ}\end{cases}$$

<u>Θέμα 2</u>

Θα εφαρμόσουμε την μέθοδο της μεγίστης καθόδου με προβολή, με s_k = 5, γ_k = 0.5, σημείο εκκίνησης (5,-5) και ακρίβεια ε = 0.01

Τρέχω το script steepest_descent_proj2.m και παίρνω το παρακάτω αποτέλεσμα για την σύγκλιση της συνάρτησης f

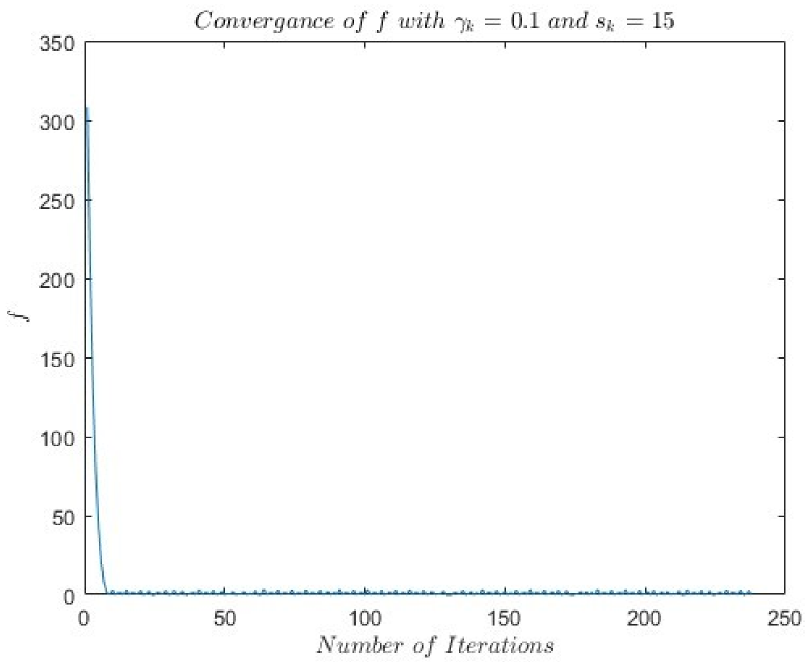


Παρατηρώ ότι η συνάρτηση δεν συγκλίνει δηλαδή η μέθοδος δεν λειτουργεί κάτι που αναμέναμε καθώς ισχύει ότι $\gamma_k' = \gamma_k s_k = 0.5 \cdot 5 = 2.5 > \frac{1}{3}$. Σε αντίθεση με όμως με το θέμα 1 παρατηρούμε ότι η f δεν απειρίζεται αλλά ταλαντώνεται γύρω από ένα σημείο ισορροπίας (χρειάστηκε να θέσουμε όριο στις 100 επαναλήψεις για να τερματίσει ο αλγόριθμος, εφόσον λόγω της μαθηματικής ανάλυσης και του γ_k γνωρίζουμε εκ των προτέρων ότι η f δεν θα συγκλίνει). Αυτό συμβαίνει επειδή η μέθοδος της προβολής δεν “αφήνει” τα x_1 και x_2 να “βγουν” εκτός περιορισμών.

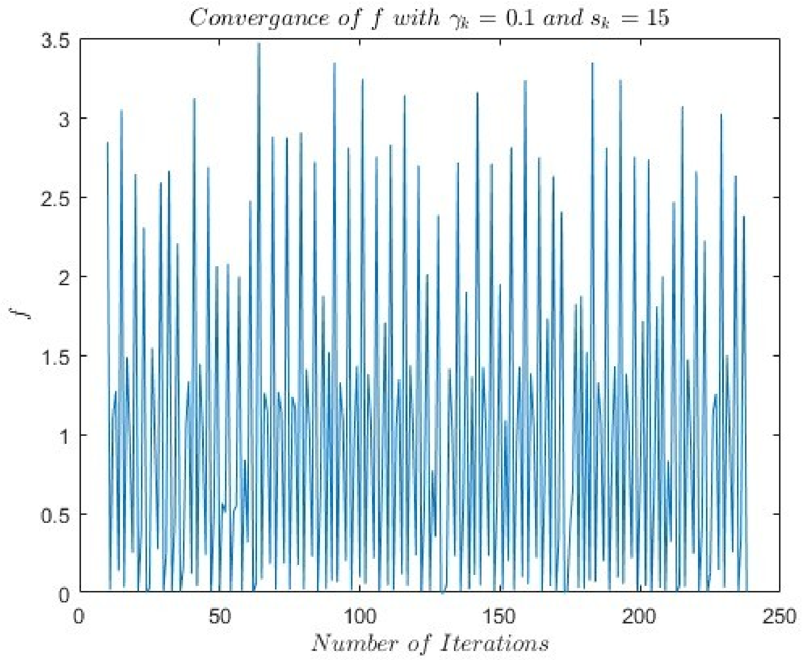
Θέμα 3

Θα εφαρμόσουμε την μέθοδο της μεγίστης καθόδου με προβολή, με $s_k = 15$, $\gamma_k = 0.1$, σημείο εκκίνησης (-5,10) και ακρίβεια ε = 0.01

Τρέχω το script steepest_descent_proj3.m και παίρνω το παρακάτω αποτέλεσμα για την σύγκλιση της συνάρτησης f



Παρατηρώ ότι η συνάρτηση δεν συγκλίνει δηλαδή η μέθοδος δεν λειτουργεί κάτι που αναμέναμε καθώς ισχύει ότι $\gamma_k{}' = \gamma_k s_k = 0.1 \cdot 15 = 1.5 > \frac{1}{3}$. Σε αντίθεση με όμως με το θέμα 1 παρατηρούμε ότι η f δεν απειρίζεται αλλά ταλαντώνεται, για τον ίδιο λόγο που εξηγήσαμε στο θέμα 2. Να σημειωθεί ότι ο αλγόριθμος τερμάτισε μετά από 238 επαναλήψεις κάτι που θεωρώ πως “έτυχε” καθώς σε αυτήν την επανάληψη ο αλγόριθμος κατάφερε να βρεθεί εντός αποδεκτού ορίου για το δεδομένο ε, ενώ σε διαφορετικές συνθήκες θα ταλαντωνόταν και αυτός επ’ αόριστον. Παραθέτουμε και το διάγραμμα της σύγκλισης από την 10η επανάληψη και μετά.



Παρατηρούμε λοιπόν πως σε αντίθεση με το θέμα 2 η μέθοδος ταλαντώνεται σε πολύ μικρότερη περιοχή από ότι στην προηγούμενη περίπτωση.

Ένας απλός και πρακτικός τρόπος για να συγκλίνει η μέθοδος θα ήταν να επιλεχτούν κατάλληλα τα γ_k και s_k έτσι ώστε να ισχύει η συνθήκη σύγκλισης, δηλαδή

$$\gamma_k \cdot s_k \in \left(0, \frac{1}{3}\right)$$

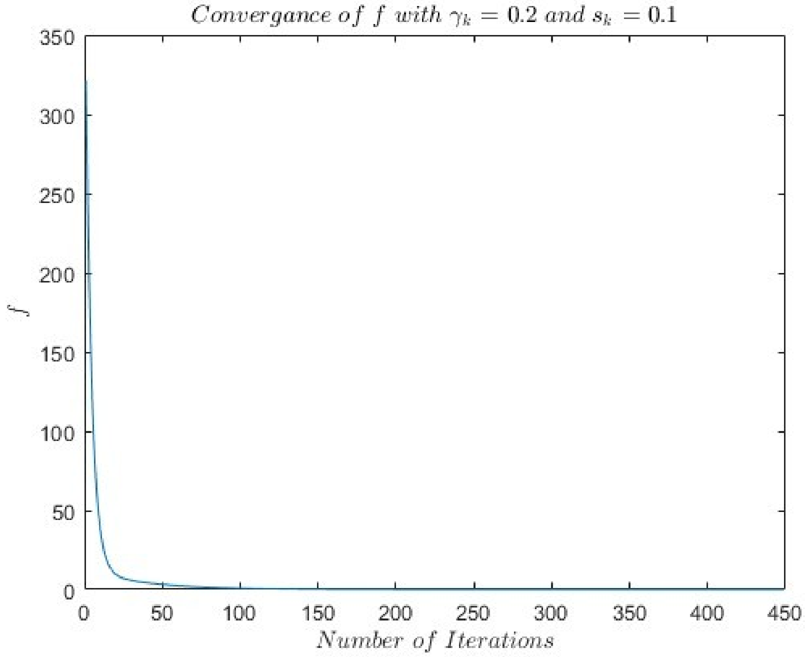
<u>Θέμα 4</u>

Θα εφαρμόσουμε την μέθοδο της μεγίστης καθόδου με προβολή, με $s_k = 0.1$, $\gamma_k = 0.2$, σημείο εκκίνησης (8,-10) και ακρίβεια ε = 0.01

Αρχικά παρατηρώ ότι το αρχικό μου σημείο είναι εκτός περιορισμών, κάτι που δεν περιμένω όμως να προκαλέσει πρόβλημα καθώς λόγω του διανύσματος προβολής θα οδηγηθεί εντός περιορισμών στην πρώτη επανάληψη.

Παρατηρώ επίσης ότι $\gamma_k{}' = \gamma_k \cdot s_k = 0.2 \cdot 0.1 = 0.002 < \frac{1}{3}$ άρα αναμένω η μέθοδος να συγκλίνει στο ελάχιστο σημείο.

Τρέχω το script steepest_descent_proj4.m και παίρνω το παρακάτω αποτέλεσμα για την σύγκλιση της συνάρτησης f



Σε αυτήν την περίπτωση, λοιπόν, η μέθοδος συγκλίνει στο ελάχιστο αντιθέτως με τα θέματα 2 και 3. Παρατηρώ όμως ότι χρειάζεται μεγάλος αριθμός επαναλήψεων (συγκεκριμένα χρειάστηκαν 449). Αυτό συμβαίνει επειδή το $\gamma_k{}' = \gamma_k s_k = 0.02$ είναι πολύ μικρό και “αργεί” η τιμή της συνάρτησης να φτάσει εντός του αποδεκτού διαστήματος.