Τεχνικές Βελτιστοποίησης

Εργασία 3

Ηλιάνα Κόγια (10090)

Στη συγκεκριμένη εργασία καλούμαστε να ελαχιστοποιήσουμε μια συνάρτηση δύο μεταβλητών χωρίς περιορισμούς, με τη Μέθοδο Μέγιστης Καθόδου και έπειτα με περιορισμούς, με τη μέθοδο Μέγιστης Καθόδου με Προβολή.

Η αντικειμενική συνάρτηση είναι:

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

Θέμα 1

Η δοθείσα συνάρτηση είναι τετραγωνική:

$$f(x) = \frac{1}{2} x^T Q\, x, \quad \text{με πίνακα } Q = \begin{bmatrix} 2/3 & 0 \\ 0 & 6 \end{bmatrix}$$

$$\nabla f(x) = Q\, x$$

$$\nabla^2 f(x) = Q$$

Ο πίνακας Q είναι θετικά ορισμένος, διότι οι ελάσσονες ορίζουσες είναι θετικές, 2/3>0 και $\det\begin{pmatrix} 2/3 & 0 \\ 0 & 6 \end{pmatrix}$ = 2/18>0 (ή οι ιδιοτιμές τιμές του διαγώνιου πίνακα Q είναι θετικές).

Άρα η συνάρτηση f είναι γνήσια κυρτή.

Επίσης, για την εύρεση του ελαχίστου έχουμε:

$$\nabla f(x) = Q\, x = 0 \Rightarrow \quad \frac{2}{3} x_1 = 0 \text{ και } 6x_2 = 0 \Rightarrow \mathbf{x_1^* = 0} \text{ και } \mathbf{x_2^* = 0}$$

Το (0,0) αποτελεί κρίσιμο σημείο και επειδή η f είναι γνήσια κυρτή αποτελεί ολικό και μοναδικό ελάχιστο της αντικειμενικής συνάρτησης.

$$\mathbf{x^* = 0 \text{ και } f(x^*) = 0}$$

Χρησιμοποιούμε, αρχικά, τη Μέθοδο Μεγίστης Καθόδου.

Η μέθοδος της μέγιστης καθόδου έχει τη μορφή:

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

είναι:

$$|x_{k+1}|^2 = x_{k+1}^T x_{k+1} = x_k^T (I - \gamma_k Q)^2 x_k \le \lambda_{max} ((I - \gamma_k Q)^2) |x_k|^2$$

Όπου λ_{max} είναι η μέγιστη ιδιοτιμή του πίνακα $(I - \gamma_k Q)^2$

Αν λ_i είναι οι ιδιοτιμές του πίνακα Q, τότε οι ιδιοτιμές του πίνακα $(I - \gamma_k Q)^2$ έχουν τη μορφή $(I - \gamma_k \lambda_i)^2$

Επομένως,

$$\lambda_{max}((I - \gamma_k Q)^2) = \max\{(1 - \gamma_k m)^2,\ (1 - \gamma_k M)^2\}$$

m : η μικρότερη ιδιοτιμή του Q => m = 2/3

M : η μεγαλύτερη ιδιοτιμή του Q => M = 6

Τελικά, για κάθε $x_k \neq 0$

$$\text{Ισχύει} \quad \frac{|x_{k+1}|}{|x_k|} \leq \max\{|1-\gamma_k m|,\ |1-\gamma_k M|\} = \max\{|1-\gamma_k \tfrac{2}{3}|,\ |1-\gamma_k 6|\}$$

Οι τιμές των βημάτων που εγγυώνται ότι η μέθοδος συγκλίνει είναι εκείνες που αντιστοιχούν σε $\frac{|x_{k+1}|}{|x_k|} \leq 1$.

Άρα, για $\gamma_k M - 1 \leq 1$

ή

$$\gamma_k \leq \frac{2}{M} = \frac{2}{6}$$

⇨ $\mathbf{0 \leq \gamma_k \leq \frac{1}{3}}$, αυτές οι τιμές του γ εγγυώνται την ευστάθεια του αλγορίθμου.

Επιπλέον, το άνω φράγμα στη ταχύτητα σύγκλισης ελαχιστοποιείται όταν:

$$1 - \gamma^* m = \gamma^* M - 1$$

$$\gamma^* = \frac{2}{m+M} = \frac{2}{\frac{2}{3}+6} = 0.3$$

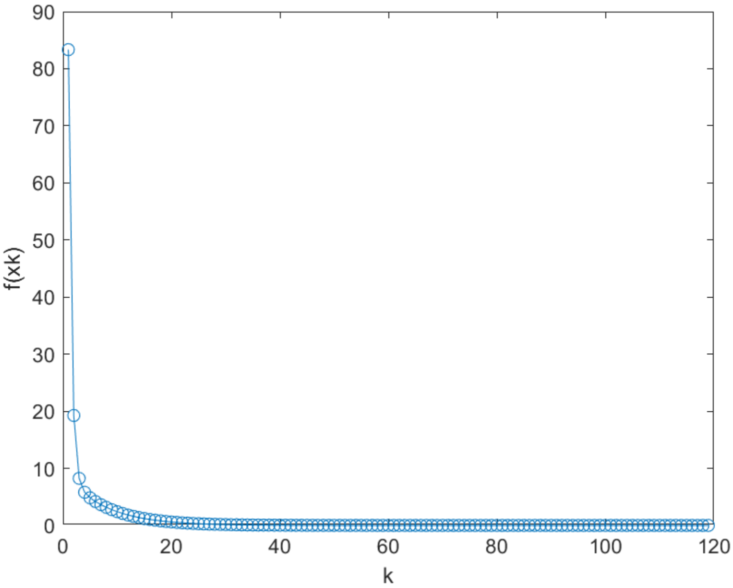
Τότε ισχύει $\frac{|x_{k+1}|}{|x_k|} \leq \frac{M-m}{M+m} = 0.8$ το οποίο είναι το καλύτερο άνω φράγμα που μπορεί να επιτευχθεί όταν χρησιμοποιούμε σταθερό βήμα γ.

Χρησιμοποιώντας την Μέθοδο Μεγίστης Καθόδου με ακρίβεια ε = 0.001 και αρχικό σημείο (5,-5) για διαφορετικό σταθερό βήμα, παρατηρούμε ότι για τις τιμές του γ για τις οποίες ισχύει $\gamma < \frac{1}{3}$ ο αλγόριθμος συγκλίνει στο ελάχιστο (0,0), διαφορετικά αποκλίνει.

Συγκεκριμένα, για τις δοθείσες τιμές του βήματος γ συμπεραίνουμε ότι ο αλγόριθμος συγκλίνει στο ελάχιστο για γ=0.1 και γ=0.5, ενώ δεν συγκλίνει για γ=3 και γ=5, όπως ήταν και θεωρητικά αναμενόμενο. Επιπλέον για γ=3 και γ=5, παρατηρούμε ότι οι τιμές των x_1 και x_2 αποκλίνουν, δηλαδή αυξάνονται συνεχώς σε κάθε επανάληψη. Ο αλγόριθμος τότε δεν είναι ευσταθής.

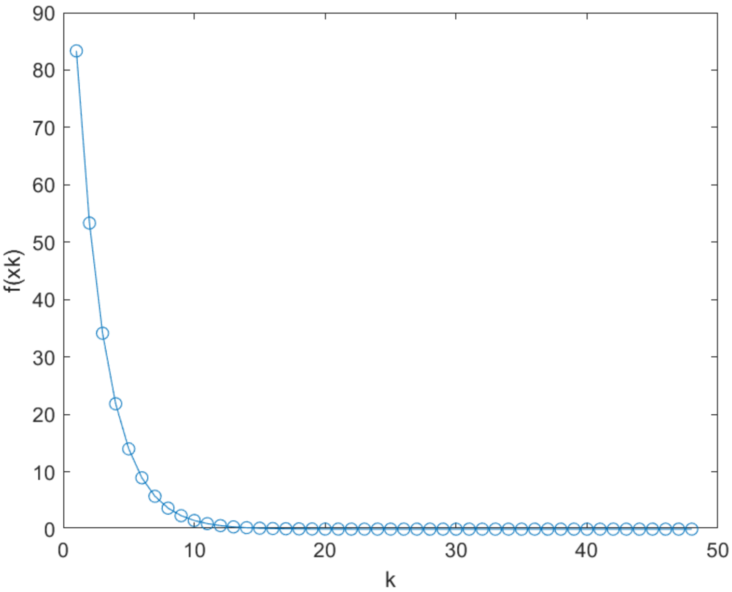
Διαγράμματα σύγκλισης της συνάρτησης ως προς k:

i) γ=0.1 x_o = (5,-5)



Προκύπτει: k= 119

ii) γ=0.3 x_o = (5,-5)



Προκύπτει: k= 48

Παρατηρούμε ότι για βήμα γ=0.3 προκύπτει πολύ μικρότερος αριθμός επαναλήψεων k, διότι η τιμή αυτή είναι ίση με το γ^*, που προέκυψε θεωρητικά με την ελαχιστοποίηση του άνω φράγματος της ταχύτητας σύγκλισης (πρακτικά για <u>γ=0.29</u> προκύπτει ο μικρότερος αριθμός επαναλήψεων <u>k = 39</u>).

<u>Θέμα 2</u>

Θεωρώντας, επιπλέον, τους περιορισμούς

$-10 \leq x_1 \leq 5$

και

$-8 \leq x_2 \leq 12$

Έστω το σύνολο $X = \{x \in R^2 : -10 \leq x_1 \leq 5, \ -8 \leq x_2 \leq 12\}$ το οποίο αποτελεί ένα ορθογώνιο. Το X είναι κυρτό σύνολο, διότι για κάθε x_1, $x_2 \in X$ το ευθύγραμμο τμήμα που ενώνει τα x_1, x_2 ανήκει επίσης στο X.

Χρησιμοποιούμε, αρχικά, τη Μέθοδο Μεγίστης Καθόδου με Προβολή με ακρίβεια

$\varepsilon = 0.01$, $\gamma_k = 0.5$, $s_k = 5$ και σημείο εκκίνησης (5,-5).

Σημείωση: Το $x^* = (0,0)$ βρίσκεται εντός του συνόλου X των περιορισμών και επειδή το σύνολο X είναι κυρτό, επιτυγχάνεται η λειτουργικότητα του αλγορίθμου.

<u>Θεωρητική Ανάλυση</u>

Τα εφικτά σημεία του αλγορίθμου είναι της μορφής:

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

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

$$\bar{x}_k = Pr_X \{\begin{bmatrix} x_{1k}\ (1-2s/3) \\ x_{2k}\ (1-6s) \end{bmatrix}\}$$

Επομένως:

$$\bar{x}_{1k} = \begin{cases} -10\,, & x_{1k}\ \left(1-\frac{2}{3}s_k\right) \leq -10 \\ x_{1k}\ \left(1-\frac{2}{3}s_k\right), & -10 < x_{1k}\ \left(1-\frac{2}{3}s_k\right) < 5 \\ 5\,, & x_{1k}\ \left(1-\frac{2}{3}s_k\right) \geq 5 \end{cases}$$

$$\bar{x}_{2k} = \begin{cases} -8\,, & x_{2k}(1-6s_k) \leq -8 \\ x_{2k}\ (1-6s_k), & -8 < x_{2k}\ (1-6s_k) < 12 \\ 12\,, & x_{2k}\ (1-6s_k) \geq 12 \end{cases}$$

$$x_{1,k+1} = \begin{cases} -10\gamma_k + (1-\gamma_k)x_{1k}, & x_{1k}\ \left(1-\frac{2}{3}s_k\right) \leq -10 \\ x_{1k}\ \left(1-\frac{2}{3}s_k\gamma_k\right), & -10 < x_{1k}\ \left(1-\frac{2}{3}s_k\right) < 5 \\ 5\gamma_k + (1-\gamma_k)x_{1k}\,, & x_{1k}\ \left(1-\frac{2}{3}s_k\right) \geq 5 \end{cases}$$

$$x_{2,k+1} = \begin{cases} -8\gamma_k + (1-\gamma_k)x_{2k}, & x_{2k}(1-6s_k) \leq -8 \\ x_{2k}\,(1-6s_k\gamma_k), & -8 < x_{2k}\,(1-6s_k) < 12 \\ 12\gamma_k + (1-\gamma_k)x_{2k}\,, & x_{2k}\,(1-6s_k) \geq 12 \end{cases}$$

Όταν το $\bar{x}_k$ είναι εντός του περιορισμού, για να συγκλίνει ο αλγόριθμος πρέπει:

$$\frac{|x_{1,k+1}|}{|x_{1,k}|} = \left|1 - \frac{2}{3}s_k\gamma_k\right| \leq 1$$

$$\Rightarrow \mathbf{s_k\gamma_k \leq 3}$$

<u>και</u>

$$\frac{|x_{2,k+1}|}{|x_{2,k}|} = |1 - 6s\gamma_k| \leq 1$$

$$\Rightarrow \mathbf{s_k\gamma_k \leq \frac{1}{3}}$$

Συνολικά, για να συγκλίνουν τα (x,y) θέλουμε $\mathbf{s_k\gamma_k \leq \frac{1}{3}}$

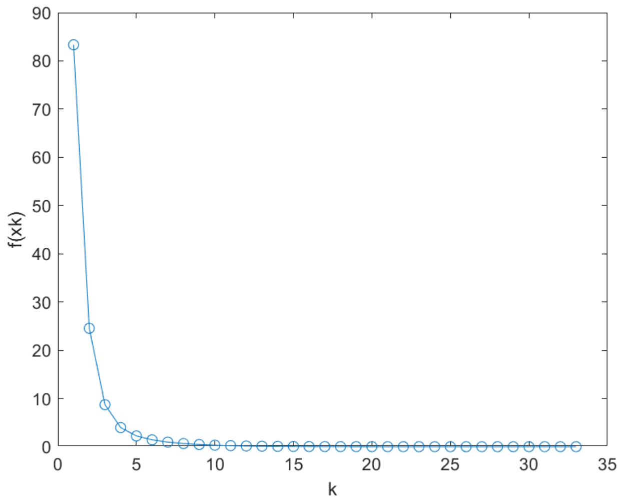
Παρατηρούμε ότι για τα δοθέντα s = 5 και γ = 0.5 ισχύει: **γs = 0.5 · 5 = 2.5**. Εκτελούμε τον αλγόριθμο για σημείο εκκίνησης (5,-5) και τα δοθέντα γ και s και παρατηρούμε ότι το x_1 συγκλίνει στο 0 σε 20 επαναλήψεις, ενώ δεν ισχύει το ίδιο για το x_2 το οποίο ταλαντώνεται μεταξύ των 5.3 και -1.3. Άρα, ο αλγόριθμος <u>δεν τερματίζει</u> και δεν βρίσκει το ελάχιστο της συνάρτησης. Αυτό είναι αναμενόμενο και θεωρητικά, καθώς ισχύει γs > $\frac{1}{3}$ άρα το x_2 δεν συγκλίνει στο x_2^*, ενώ γs < 3 οπότε το x_1 συγκλίνει στο x_1^*.

Η διαφορά σε σχέση με το Θέμα 1, είναι ότι τα (x_1 ,x_2) στην Steepest Descent αποκλίνουν, δηλαδή αυξάνονται συνεχώς για τιμές γ = 3 και γ = 5, ενώ στη μέθοδο με προβολή το x_1 συγκλίνει και το x_2 ταλαντώνεται μεταξύ δύο σημείων.

Αν θέσουμε **s=0.5 και γ=0.5** τότε τα (x_1 ,x_2) συγκλίνουν στο σημείο ελαχίστου, διότι ισχύει γs < $\frac{1}{3}$ <3

Διάγραμμα σύγκλισης της συνάρτησης ως προς k, τροποποιώντας τα s,γ :

γ_k = 0.5 **s_k = 0.5** **x_o = (5,-5)**



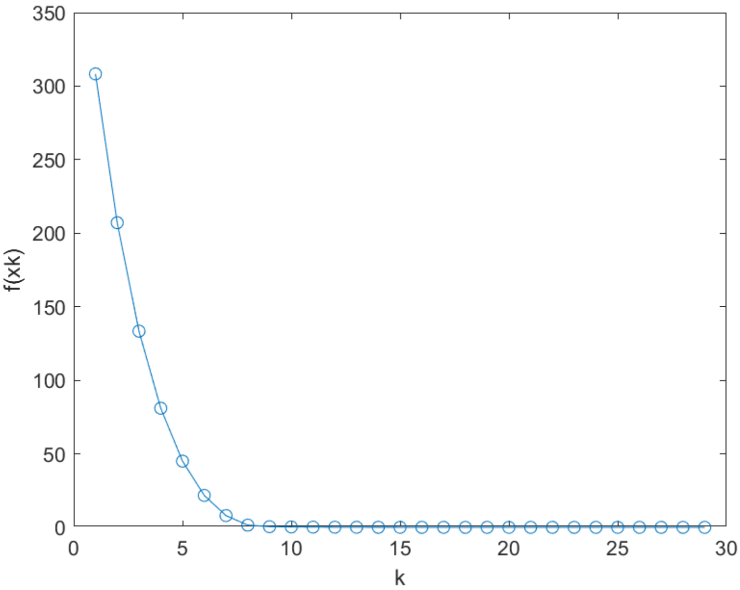
Θέμα 3

Χρησιμοποιούμε τη Μέθοδο Μεγίστης Καθόδου με Προβολή με ακρίβεια $\varepsilon = 0.01$, $\gamma_k = 0.1$, $s_k = 15$ και σημείο εκκίνησης (-5,10). Εκτελούμε τον αλγόριθμο και παρατηρούμε ότι το x_1 συγκλίνει στο $x_1^* = 0$ στην 8^η επανάληψη, ενώ το x_2 δεν συγκλίνει, αλλά ταλαντώνει και συγκεκριμένα όχι μεταξύ των ίδιων σημείων, όπως παρατηρήθηκε στο θέμα 2. Οπότε, *ο αλγόριθμος δεν συγκλίνει τελικά στο ελάχιστο της f.* Αυτό ήταν θεωρητικά αναμενόμενο, διότι ισχύει $\gamma s = 0.1 \cdot 15 = 1.5 > 1/3$. (Μπορεί τυχαία όπως ταλαντώνει το x_2 κοντά στο 0 να βρει το 0, όμως για πολύ μεγάλες τιμές του k).

Αν θέσουμε $\gamma_k = 0.1$ και $s_k = 2.9$, τότε το γινόμενο είναι $\gamma s = 0.29 < 1/3$. Συνεπώς, ο αλγόριθμος θα συγκλίνει στο ελάχιστο.

Διάγραμμα σύγκλισης της συνάρτησης ως προς k, τροποποιώντας τα s,γ :

$\gamma_k = 0.1$ $\quad$ $s_k = 2.9$ $\quad$ $x_o = (-5,10)$



Προκύπτουν k = 29 επαναλήψεις

Εναλλακτικά, για να βρούμε το βέλτιστο βήμα, εντός των περιορισμών ισχύει:

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

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

$$x_{2,1} = x_{2o} \left(1 - 6s\gamma_k\right)$$

για το βέλτιστο βήμα:

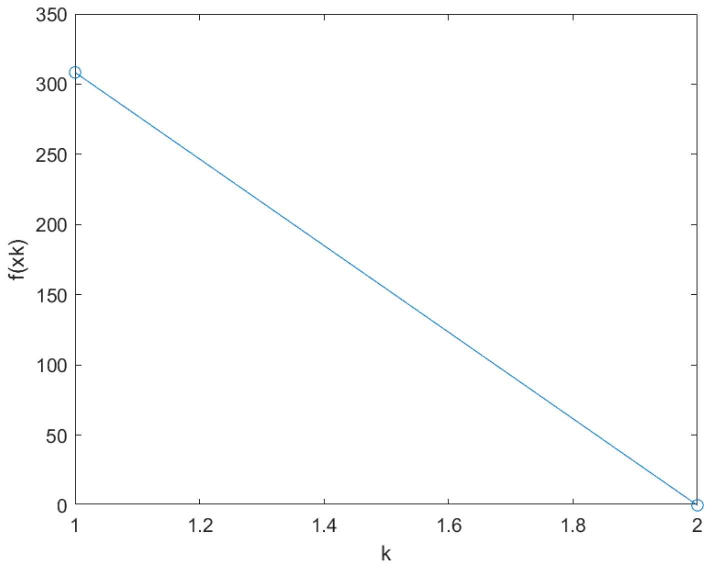
$$\frac{df(x_{k+1})}{d\gamma'} = 0$$

⇨ $-\nabla^T f(x_{k+1}) \, \nabla f(x_{k+1}) = 0$

⇨ **$\gamma_k\, s_k$ = 1.5 και $\gamma_k\, s_k$ = 1/6**

Για να συγκλίνει σε ένα βήμα στο (0,0) παρατηρούμε ότι πρέπει να ισχύουν ταυτόχρονα οι δύο παραπάνω σχέσεις. Για να επιτευχθεί αυτό θέτουμε $\gamma_k = 1$ και διαφορετικά s ($s_{1k} = 1.5$, $s_{2k} = 1/6$) για τα x_1 και x_2, ώστε να πληρούνται οι δύο σχέσεις. Προκύπτει έτσι k=2, δηλαδή απαιτείται μία επανάληψη για την εύρεση του ελαχίστου f(0,0)=0.

Διάγραμμα σύγκλισης της συνάρτησης ως προς k:



Θέμα 4

Χρησιμοποιώντας γ = 0.2 και s = 0.1 ισχύει γs = 0.02<1/3. Συνεπώς, ο αλγόριθμος θα συγκλίνει στο ελάχιστο, αλλά υστερεί αρκετά ως προς την ταχύτητα σύγκλισης.

Εντός των περιορισμών ισχύει:

$$x_{1,k+1} = x_{1k}\left(1-\frac{2s}{3}\gamma_k\right) = 0.9867x_{1k}$$

$$\Rightarrow\ x_{1,k+1} = (0.9867)^k\, x_{1o}$$

$$x_{2,k+1} = x_{2k}\left(1-6s\gamma_k\right) = 0.88x_{2k}$$

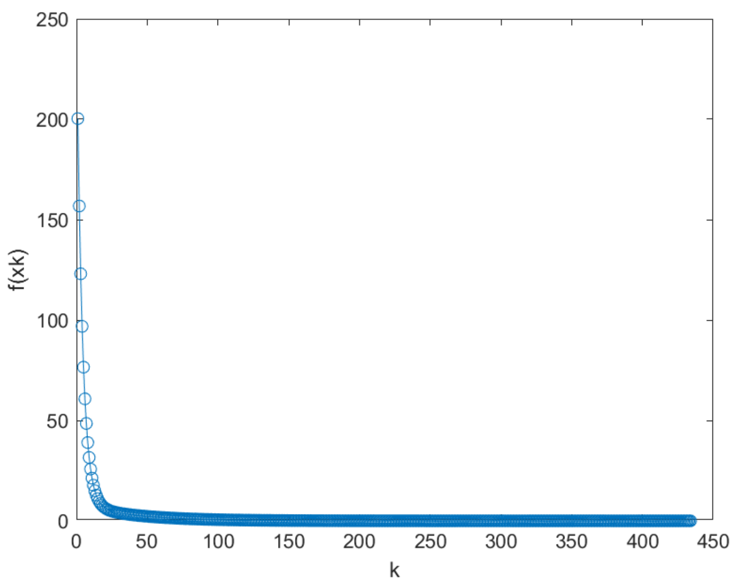
$$\Rightarrow\ x_{2,k+1} = (0.88)^k\, x_{2o}$$

Από τις παραπάνω σχέσεις εξάγουμε μια εκ' των προτέρων πληροφορία σχετικά με τη σύγκλιση του αλγορίθμου. Παρατηρούμε ότι για k = 450 το x_1 συγκλίνει περίπου στο 0.002 (0.9867^{450} = 0.002). Οπότε δεδομένης της ακρίβειας ε = 0.01, περιμένουμε περίπου 450 επαναλήψεις. Το σημείο εκκίνησης επηρεάζει την ταχύτητα σύγκλισης, αλλά όχι σε μεγάλο βαθμό όπως το γινόμενο γs. Σημειώνεται, ακόμη, ότι το δοθέν αρχικό σημείο (8,-10) είναι εκτός του κυρτού συνόλου X. Ωστόσο, με την εύρεση της προβολής του αρχικού σημείου στο X, έχουμε εφικτό αρχικό σημείο.

Με την εκτέλεση του αλγορίθμου επαληθεύονται τα παραπάνω.

Διάγραμμα σύγκλισης της συνάρτησης ως προς k:

γ_k = 0.2 s_k = 0.1 x_o = (8,-10)



Προκύπτουν k = 434 επαναλήψεις