Τσαντίκης Γεώργιος
ΑΕΜ: 10722

3η Εργαστηριακή Άσκηση

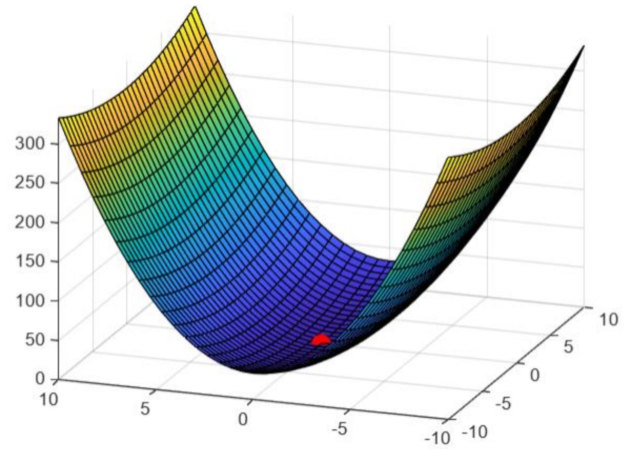
Μέθοδος Μέγιστης Καθόδου με Προβολή

ΠΕΡΙΓΡΑΦΗ ΠΡΟΒΛΗΜΑΤΟΣ

Στην παρούσα εργασία ζητείται να επιλυθεί το πρόβλημα ελαχιστοποίησης της συνάρτησης δύο μεταβλητών $f(\mathrm{x}) = \frac{1}{3}x_1^2 + 3x_2^2$, $f{:}\mathbb{R}^2 \to \mathbb{R}$ και $\mathrm{x} = [x_1, x_2]^T$, υπό τους περιορισμούς $-10 \leq x_1 \leq 5$ και $-8 \leq x_2 \leq 12$ (ορθογώνιο σύνολο), με χρήση της μεθόδου **Μέγιστης Καθόδου με Προβολή.** Προϋποθέσεις της μεθόδου αποτελούν το σύνολο X να είναι κυρτό και το x^*(η λύση) να ανήκει στο X. Όπως είναι φανερό από τους περιορισμούς που διατυπώθηκαν παραπάνω, ισχύουν και οι δύο.

Η μέθοδος της μέγιστης καθόδου με προβολή είναι ένας αλγόριθμο εφικτών σημείων, της μορφής $x_{k+1} = x_k + \gamma_k(\overline{x_k} - x_k)$ με $\gamma_k \in (0, 1]$, όπου $\overline{x_k} = Pr_X\{x_k - s_k \nabla f(x_k)\}$ με $s_k > 0$ και $Pr_X(\ )$ ο τελεστής προβολής. Επομένως το $\overline{x_k}$ προσδιορίζεται από το x_k αν κινηθούμε στην κατεύθυνση της αρνητικής κλίσης με βήμα s_k και στη συνέχεια βρούμε την προβολή του $x_k - s_k \nabla f(x_k)$ στο X, καταλήγοντας με τον τρόπο αυτό σε εφικτό σημείο. Έχοντας βρει την εφικτή κατεύθυνση αναζήτησης $\overline{x_k} - x_k$, κινούμαστε με βήμα $\gamma_k \in (0, 1]$ σε αυτήν και προσδιορίζουμε το νέο εφικτό σημείο x_{k+1}. Αν το $x_k - s_k \nabla f(x_k)$ είναι εφικτό, η μέθοδος μέγιστης καθόδου με προβολή μετατρέπεται στη μέθοδο μέγιστης καθόδου χωρίς περιορισμούς. Η μέθοδος τερματίζεται όταν $\|\nabla f(x_k)\| \leq \varepsilon$, με ε να αποτελεί την ακρίβεια που έχει επιλεχθεί για την αναζήτηση του στάσιμου σημείου (x^*) για το οποίο θεωρητικά ισχύει $\nabla f(x^*) = 0$.

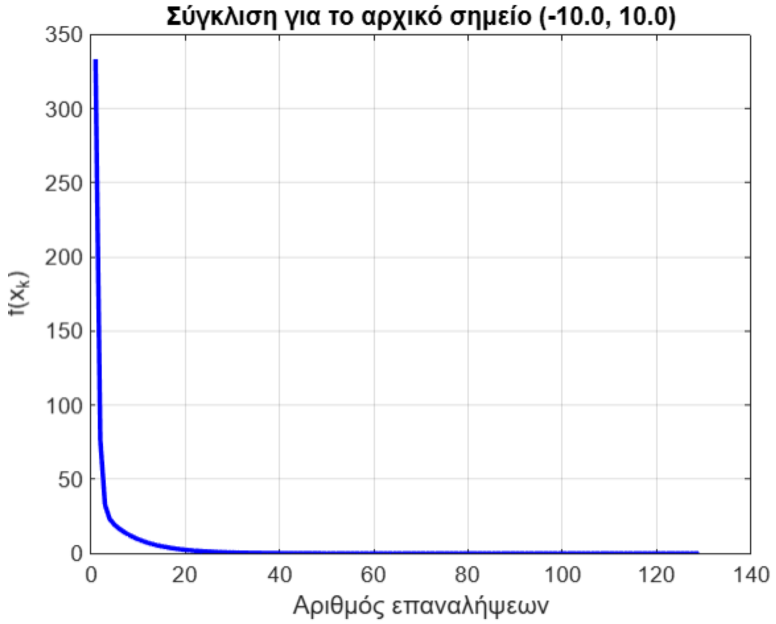
Θα δοκιμαστούν διάφορες τιμές των βημάτων γ_k, s_k και από διαφορετικά αρχικά σημεία, με $\gamma_k = \gamma$ σταθερό . Τα αποτελέσματα πρέπει να συγκριθούν τόσο μεταξύ τους όσο και με την μέθοδο της **Μέγιστης Καθόδου** από την προηγούμενη εργασία. Παρακάτω απεικονίζεται η συνάρτηση f η οποία παρουσιάζει **ολικό ελάχιστο** στο σημείο $(\boldsymbol{x_1^*}, \boldsymbol{x_2^*}) = (\mathbf{0}, \mathbf{0})$.



ΘΕΜΑ 1

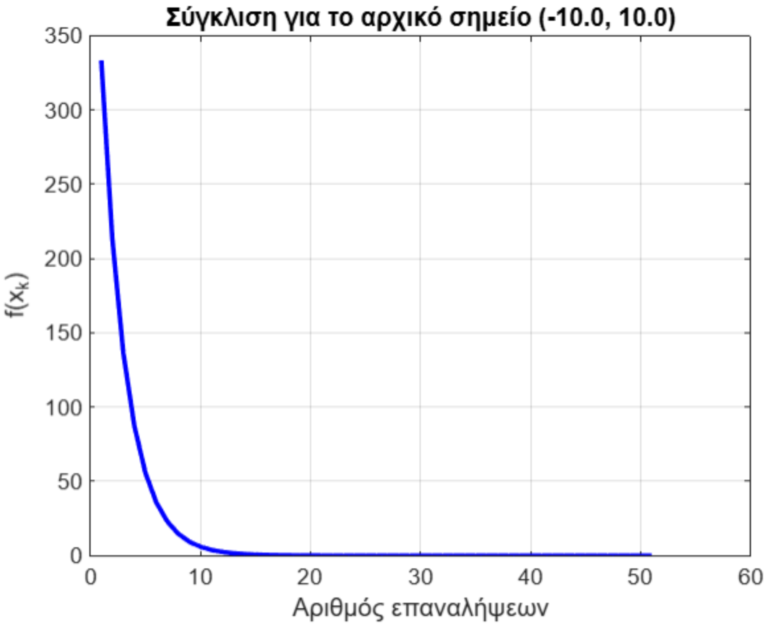
Στο παρόν θέμα χρησιμοποιήθηκε η μέθοδος της **Μέγιστης Καθόδου**, όπως ακριβώς παρουσιάστηκε στην προηγούμενη εργασία. Παρακάτω παρουσιάζονται οι γραφικές παραστάσεις της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων που απαιτούνται μέχρι να τερματίσει ο αλγόριθμος με τυχαίο αρχικό σημείο εκκίνησης $(-10, 10)$ για διάφορες τιμές του γ_k με ακρίβεια $\varepsilon = 0.001$:

i) $\gamma_k = 0$.



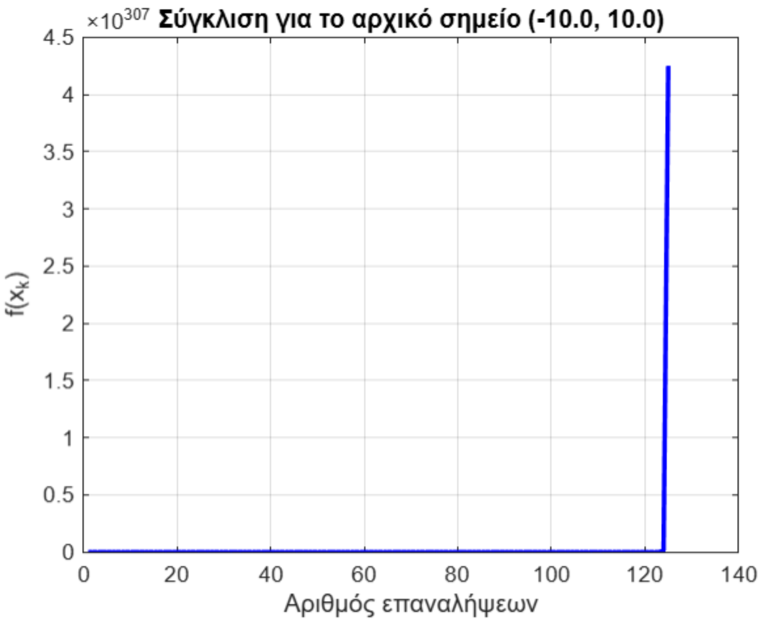
```
Αρχικό σημείο: (-10.00, 10.00)
Τελικό σημείο: x* = (-0.0015, 0.0000)
Τελική τιμή: f(x*) = 0.000001
Συνολικές επαναλήψεις: 128
```

ii) $\gamma_k = 0.3$



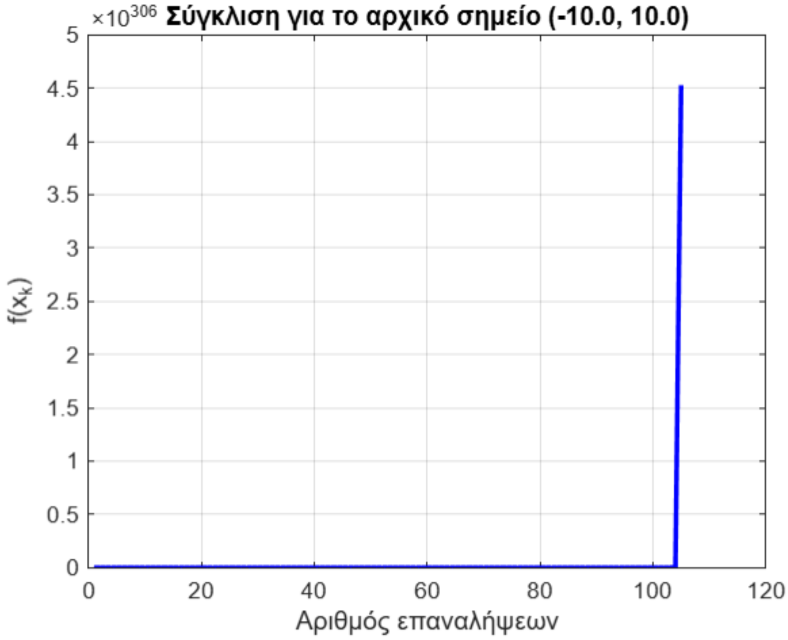

```
Αρχικό σημείο: (-10.00, 10.00)
Τελικό σημείο: x* = (-0.0001, 0.0001)
Τελική τιμή: f(x*) = 0.000000
Συνολικές επαναλήψεις: 50
```

iii) $\gamma_k = 3$



```
Αρχικό σημείο: (-10.00, 10.00)
Τελικό σημείο: x* = (10.0000, NaN)
Τελική τιμή: f(x*) = NaN
Συνολικές επαναλήψεις: 251
```

iv) $\gamma_k = 5$




```
Αρχικό σημείο: (-10.00, 10.00)
Τελικό σημείο: x* = (-10258372911528904832294143205596794320667428444705335605533858716663796985757696.0000, NaN)
Τελική τιμή: f(x*) = NaN
Συνολικές επαναλήψεις: 212
```

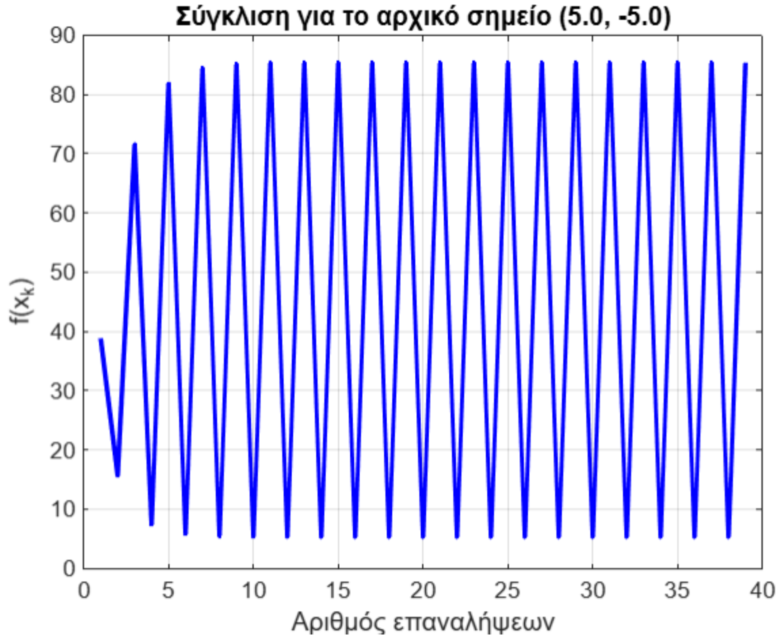
Παρατηρούμε ότι για τις περιπτώσεις i) $\gamma_k = 0.1$ και ii) $\gamma_k = 0.3$ ο αλγόριθμος συγκλίνει ενώ για τις περιπτώσεις iii) $\gamma_k = 3$ και iv) $\gamma_k = 5$ δεν συγκλίνει καθώς το $\gamma_k \in (0, 1]$ σύμφωνα και με την *Παρατήρηση **6.1.6** (σελ. 202 ΒΙΒΛΙΟΥ).* Τα παραπάνω αποτελέσματα αποδεικνύονται με μαθηματική αυστηρότητα ως εξής:

Θέλουμε να ελαχιστοποιήσουμε την $f(\mathrm{x}) = \frac{1}{3}x_1^2 + 3x_2^2$, $f{:}\mathbb{R}^2 \rightarrow \mathbb{R}$ και $\mathrm{x} = [x_1, x_2]^T$, χρησιμοποιώντας τη μέθοδο μέγιστης καθόδου με σταθερό βήμα $\gamma_k = \gamma$, $\forall\, k$.
Έχουμε: $\nabla f(x_1, x_2) = [\frac{2}{3}x_1,\ 6x_2\,]^T$. Ο αλγόριθμος της μέγιστης καθόδου είναι : $x_{k+1} = x_k - \gamma_k \nabla \mathrm{f}(x_k)$ και αντικαθιστώντας προκύπτει: $x_{1,k+1} = x_{1,k}(1 - \frac{2}{3}\gamma)$ και $x_{2,k+1} = x_{2,k}(1 - 6\gamma)$. Για να συγκλίνει η μέθοδος στο $(x_1^*, x_2^*) = (0, 0)$ που αποτελέι προφανώς το ελάχιστο της f θα πρέπει : $\frac{|x_{1,k+1}|}{|x_{1,k}|} < 1 \;=> \cdots \;=>\; 0 < \gamma < 3$ και $\frac{|x_{2,k+1}|}{|x_{2,k}|} < 1 \;=> \cdots \;=>\; 0 < \gamma < \frac{1}{3}$, άρα τελικά **για να συγκλίνει η μέθοδος** πρέπει:

$$\mathbf{0 < \gamma < \frac{1}{3}}$$

<u>ΘΕΜΑ 2</u>

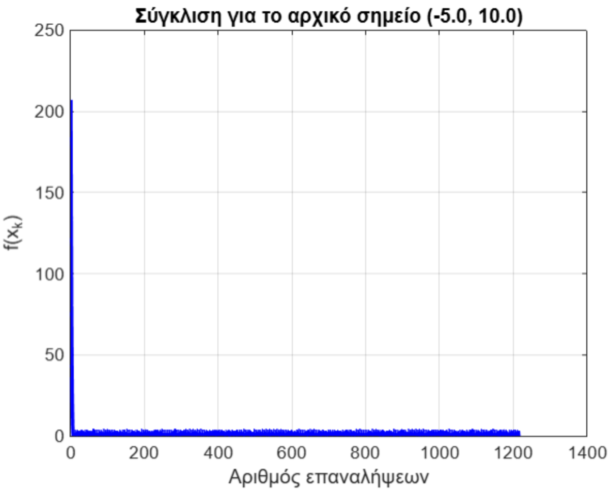
Στο παρόν θέμα χρησιμοποιήθηκε η μέθοδος της **Μέγιστης Καθόδου με Προβολή** όπως παρουσιάστηκε στην *ΠΕΡΙΓΡΑΦΗ ΠΡΟΒΛΗΜΑΤΟΣ,* με $s_k = 5$, $\gamma_k = 0.5$, σημείο εκκίνησης το $(5, -5)$ και ακρίβεια $\varepsilon = 0.01$. Παρακάτω παρουσιάζεται η γραφική παράσταση της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων έχοντας τεθεί ένα όριο στον αριθμό επαναλήψεων καθώς ο αλγόριθμος δεν τερματίζει, ώστε να φαίνεται η διακύμανση των τιμών της f:



Σε σχέση με το *ΘΕΜΑ 1* παρατηρούμε ότι ο αλγόριθμος αποκλίνει. Εφόσον βρισκόμαστε εξ' αρχής στο εφικτό σύνολο X, το νέο σημείο υπολογίζεται κάθε φορά με τη σχέση: $x_{k+1} = x_k - \gamma_k s_k \nabla f(x_k)$, καθώς όταν ένα σημείο βρίσκεται μέσα στο εφικτό σύνολο X η προβολή του σε αυτό είναι το ίδιο το σημείο και η σχέση προκύπτει μετά από απλές πράξεις στην $x_{k+1} = x_k + \gamma_k(\overline{x_k} - x_k)$. Επομένως το **βήμα** μας για στην κατεύθυνση καθόδου θα είναι $\gamma_k \cdot s_k$ και χρησιμοποιώντας τα δεδομένα του συγκεκριμένου θέματος έχουμε: $\gamma_k \cdot s_k = 0.5 \cdot 5 = 2.5 > \frac{1}{3}$, οπότε ο αλγόριθμος θα αποκλίνει σύμφωνα με την μαθηματική ανάλυση στο *ΘΕΜΑ 1.*

ΘΕΜΑ 3

Στο παρόν θέμα χρησιμοποιήθηκε η μέθοδος της **Μέγιστης Καθόδου με Προβολή** όπως παρουσιάστηκε στην *ΠΕΡΙΓΡΑΦΗ ΠΡΟΒΛΗΜΑΤΟΣ,* με $s_k = 15,\ \gamma_k = 0.1$, σημείο εκκίνησης το $(-5, 10)$ και ακρίβεια $\varepsilon = 0.01$. Παρακάτω παρουσιάζεται η γραφική παράσταση της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων (αριστερά) και η ίδια σε μεγέθυνση για να γίνουν εμφανείς όταν αρχίζουν οι ταλαντώσεις (δεξιά):



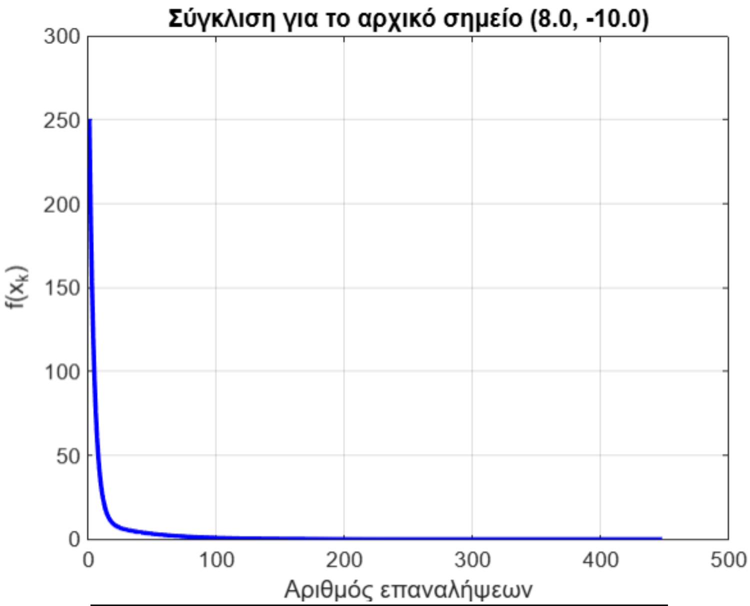
Παρομοίως με το *ΘΕΜΑ 2* για τον υπολογισμό του νέου σημείου, προκύπτει για το βήμα $\gamma_k \cdot s_k$ της κατεύθυνσης καθόδου ότι: $\gamma_k \cdot s_k = 0.1 \cdot 15 = 1.5 > \frac{\mathbf{1}}{\mathbf{3}}$. Σύμφωνα με την μαθηματική ανάλυση στο *ΘΕΜΑ 1* θα αναμέναμε ο αλγόριθμος να αποκλίνει λόγω του προηγούμενου αποτελέσματος (όπως και στο *ΘΕΜΑ 2*), ωστόσο πλησιάζει γρήγορα σε μία περιοχή κοντά στο ολικό ελάχιστο της f, μετά αρχίζει να ταλαντώνεται γύρω από αυτή και καταλήγει να συγκλίνει:

```
Αρχικό σημείο: (-5.00, 10.00)
Τελικό σημείο: x* = (0.0000, 0.0009)
Τελική τιμή: f(x*) = 0.000003
Συνολικές επαναλήψεις: 1216
```

Ένας απλός πρακτικός τρόπος για να αντιμετωπιστεί η ταλάντωση που παρατηρείται στη σύγκλιση της μεθόδου μέγιστης καθόδου με προβολή είναι με τη χρήση δυναμικής ρύθμισης του βήματος $\gamma_k \cdot s_k$, κάτι το οποίο μπορεί να πραγματοποιηθεί με τον **κανόνα Armijo** που αποτελεί μία μέθοδο διαδοχικής μείωσης του βήματος και υλοποιήθηκε στην προηγούμενη εργασία. Ο κανόνας Armijo προφανώς δεν εγγυάται από μόνος του ότι το βήμα θα ανήκει στο διάστημα $\left(0, \frac{1}{3}\right)$ αλλά μπορούμε είτε να ρυθμίσουμε τις παραμέτρους του κατάλληλα είτε να προσθέσουμε έναν έλεγχο πέρα από τον ήδη υπάρχοντα της μεθόδου ώστε να συνεχίζει να μειώνει το συνολικό βήμα όσο αυτό είναι μεγαλύτερο ή ίσο του 1/3. Τελικά, εξασφαλίζει την ταχύτερη και σταθερή σύγκλιση στον στόχο.

ΘΕΜΑ 4

Στο παρόν θέμα χρησιμοποιήθηκε η μέθοδος της **Μέγιστης Καθόδου με Προβολή** όπως παρουσιάστηκε στην *ΠΕΡΙΓΡΑΦΗ ΠΡΟΒΛΗΜΑΤΟΣ,* με $s_k = 0.1,\ \gamma_k = 0.2$, σημείο εκκίνησης το $(8, -10)$ και ακρίβεια $\varepsilon = 0.01$. Παρακάτω παρουσιάζεται η γραφική παράσταση της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων:




```
Αρχικό σημείο: (8.00, -10.00)
Τελικό σημείο: x* = (0.0149, -0.0000)
Τελική τιμή: f(x*) = 0.000074
Συνολικές επαναλήψεις: 449
```

Σε αντίθεση με τα προηγούμενα θέματα το σημείο εκκίνησης δεν ανήκει στο εφικτό σύνολο X, οπότε γίνεται σταδιακό βήμα γ_k από το τρέχον σημείο x_k προς την προβολή $\overline{x_k}$ στο X. Αν είχαμε $\gamma_k = 1$ τότε το x_k θα μετακινούνταν αμέσως στο $\overline{x_k}$, ωστόσο έχουμε $\gamma_k = 0.2$ που είναι μικρό, συνεπώς ο αλγόριθμος θα χρειαστεί αρκετές επαναλήψεις για να εισέλθει στο εφικτό σύνολο. Όταν το x_k βρεθεί εντός του εφικτού συνόλου υπολογίζουμε το νέο σημείο όπως και στα προηγούμενα θέματα με βήμα $\gamma_k \cdot s_k = 0.1 \cdot 0.2 = 0.02$, το οποίο είναι επίσης μικρό. Λόγω των παραπάνω αναμένουμε ο αλγόριθμος να συγκλίνει με πολύ αργό ρυθμό προς το ολικό ελάχιστο της f, κάτι το οποίο επιβεβαιώνεται από την παραπάνω γραφική παράσταση. Παράλληλα, εφόσον η f είναι κυρτή και το συνολικό βήμα $\gamma_k \cdot s_k = 0.02 < \frac{1}{3}$ είναι μικρό και ικανοποιεί τη συνθήκη σύγκλισης, η πληροφορία που έχουμε εκ των προτέρων είναι ότι ο αλγόριθμος θα συγκλίνει και μάλιστα αργά.