

Τεχνικές Βελτιστοποίησης: Εργασία 3

Ιωάννης Μιχάλαινας ΑΕΜ:10902

Δεκέμβριος 2025

Αποθετήριο Κώδικα:
github.com/ioannisam/optimization-techniques

Περίληψη

Στην εργασία αυτή υλοποιείται και εφαρμόζεται η βελτιστοποίηση αντικειμενικής συνάρτησης παρουσία περιορισμών. Ονομαστικά μελετάται η μέθοδος Μέγιστης Καθόδου με Προβολή. Εξετάζεται ο αριθμός υπολογισμών και η σύγχλιση, ως προς την ταχύτητα και την ακρίβεια, εφαρμόζοντάς την από διαφορετικά αρχικά σημεία και με διαφορετικές τιμές του βήματος. Τα αποτελέσματα δείχνουν τη σημασία της σωστής επιλογής του βήματος για τον αλγόριθμο. Ο σκοπός της εργασίας αυτής είναι η εξοικείωση με την παραπάνω μέθοδο και η μελέτη της, όσον αφορά την αποδοτικότητα.

Περιεχόμενα

1	Εισαγωγή	2
2	Μέθοδος Μέγιστης Καθόδου	4
2.1	Χωρίς Προβολή	4
2.2	Με Προβολή	4
3	Θέμα 1	5
3.1	Γραφικές Παραστάσεις	5
3.2	Μαθηματική Ανάλυση	6
3.3	Συμπεράσματα	6
4	Θέμα 2	7
4.1	Γραφικές Παραστάσεις	7
4.2	Συμπεράσματα	8
5	Θέμα 3	9
5.1	Γραφικές Παραστάσεις	9
5.2	Συμπεράσματα	10
6	Θέμα 4	11
6.1	Γραφικές Παραστάσεις	11
6.2	Συμπεράσματα	12
A'	Εργαλεία	12
B'	Βιβλιογραφία	12

1 Εισαγωγή

Στην εργασία αυτή μελετάμε τη μέθοδο Μέγιστης Καθόδου με Προβολή. Ο σκοπός είναι η προσέγγιση του x^* , της λύσης του προβλήματος ελαχιστοποίησης, παρουσία περιορισμών.

Έστω z ένα διάνυσμα στον R^n . Θα ονομάζουμε προβολή του z στο X οποιοδήποτε διάνυσμα $x \in X$ το οποίο απέχει τη μικρότερη Ευκλείδεια απόσταση από το z . Δηλαδή, αναζητούμε το $x \in X$ που λύνει το πρόβλημα:

$$\min_{x \in X} |z - x|^2.$$

Η επιλογή του βήματος g_k είναι νευραλγικής σημασίας για την ορθότητα του αλγορίθμου, αφού αν επιλεγεί πολύ μικρό θα οδηγήσει σε ασήμαντη μείωση της f και κατά συνέπεια αργή σύγκλιση. Αντίθετα, αν επιλεγεί πολύ μεγάλο μπορεί να οδηγήσει τον αλγόριθμο σε αστάθεια.

Δοσμένων των αρχικών σημείων $[5, -5]$, $[-5, 10]$ και $[8, -10]$ θα ελαχιστοποιηθεί η:

$$f(x_1, x_2) = \frac{1}{3}x_1^2 + 3x_2^2 = \frac{1}{2}x^\top Qx \quad \text{όπου} \quad Q = \begin{bmatrix} \frac{2}{3} & 0 \\ 0 & 6 \end{bmatrix}.$$

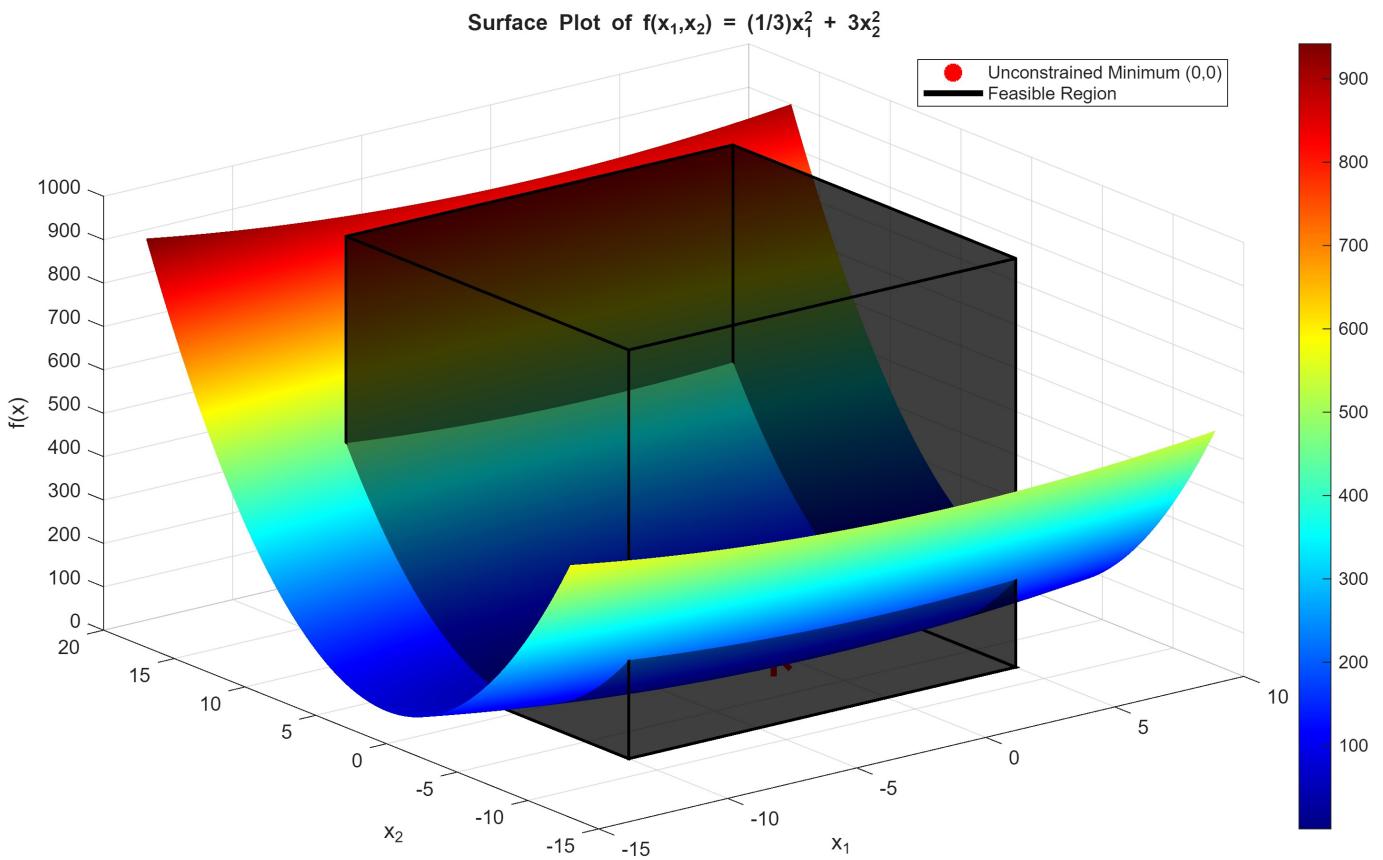
Η συνάρτηση f είναι τετραγωνική, συνεχής και δύο φορές παραγωγίσιμη. Παρακάτω θα παρουσιαστεί η γραφική της παράσταση τόσο στον τρισδιάστατο χώρο όσο και στον δισδιάστατο χώρο, υπό τη μορφή ισοβαρών καμπυλών.

Αυτή παρουσιάζει προφανές ελάχιστο στο $[0, 0]$ με τιμή $f_{min} = 0$. Εξατίας του τετραγωνικού της χαρακτήρα, η επιφάνεια σχηματίζει συμμετρική παραβολή που ανοίγει προς τα πάνω. Οι ισοβαρείς καμπύλες έχουν ελλειπτικό σχήμα, με μεγαλύτερη επιμήκυνση κατά τον άξονα x_1 λόγω του μικρότερου συντελεστή μπροστά στο x_1^2 , ενώ η γρήγορη αύξηση του όρου $3x_2^2$ δημιουργεί πιο «στενή» καμπυλότητα προς την κατεύθυνση του x_2 . Η μορφή αυτή αντικατοπτρίζει την ομαλή και μονοτονική αύξηση της συνάρτησης καθώς απομακρύνομαστε από το μοναδικό ελάχιστο.

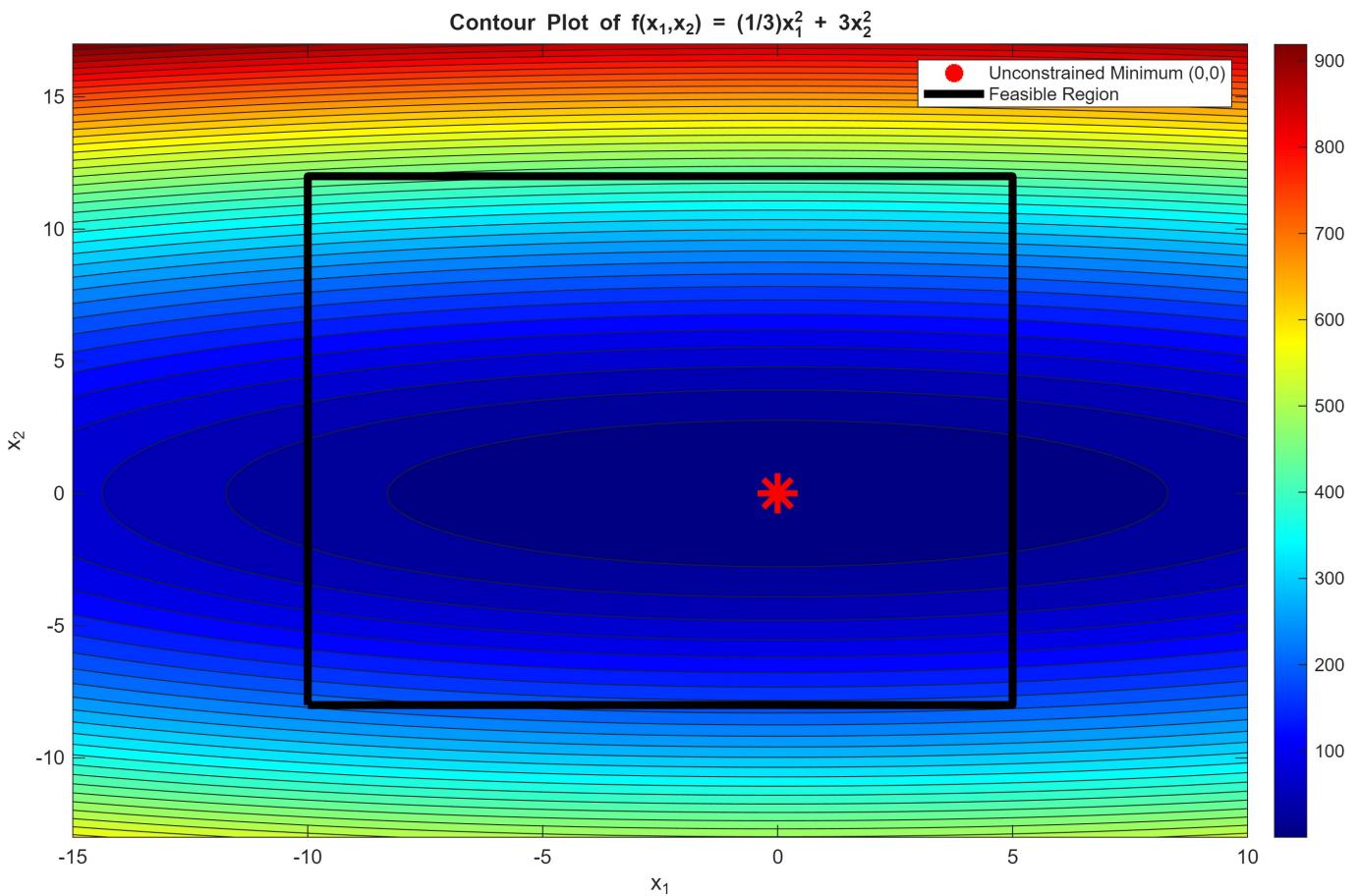
Μαζί με τη συνάρτηση έχει σχεδιαστεί, ως ημιδιαφανές μαύρο κουτί, και ο χώρος των εφικτών σημείων X , που ορίζεται από τους εξής περιορισμούς:

$$-10 < x_1 < 5$$

$$-8 < x_2 < 12$$



Σχήμα 1: Τρισδιάστατη απεικόνιση της συνάρτησης f



Σχήμα 2: Ισοεπίπεδη απεικόνιση της συνάρτησης f

2 Μέθοδος Μέγιστης Καθόδου

2.1 Χωρίς Προβολή

Ο αλγόριθμος και η μέθοδος αυτή παρουσιάστηκε εκτενώς στην προηγούμενη εργασία, οπότε στην ενότητα αυτή παραθέτουμε μόνο τον ψευδοκώδικα.

Algorithm 1 Μέθοδος Μέγιστης Καθόδου

Require: συνάρτηση f , αρχικό σημείο x_1 , σταθερά $\epsilon > 0$

Ensure: προσέγγιση ελαχίστου x^*

```
1: ορίστε  $x_k \leftarrow x$ ,  $k \leftarrow 1$ 
2: while  $|\nabla f(x_k)| > \epsilon$  do
3:    $d_k \leftarrow -\nabla f(x_k)$ 
4:    $\gamma_k \leftarrow \arg \min_{\gamma > 0} f(x_k + \gamma d_k)$ 
5:    $x_k \leftarrow x_k + \gamma_k d_k$ 
6:    $k \leftarrow k + 1$ 
7: end while
8: return  $x^* \leftarrow x_k$ 
```

2.2 Με Προβολή

Η παρακάτω αποτελεί τροποποίηση της μεθόδου Μέγιστης Καθόδου, ώστε αυτή να μπορεί να διαχειριστεί προβλήματα με περιορισμούς. Εξασφαλίζει την αναζήτηση του ελαχίστου χρησιμοποιώντας την έννοια της προβολής για την αποφυγή της εξόδου από το X , με την προϋπόθεση πως το σημείο εκκίνησης είναι εφικτό, δηλαδή $x_1 \in X$.

Algorithm 2 Μέθοδος Μέγιστης Καθόδου με Προβολή

Require: συνάρτηση f , αρχικό σημείο x_1 , σταθερά $s_k > 0$, σ σταθερά $\gamma_k > 0$, σταθερά $\epsilon > 0$

Ensure: προσέγγιση ελαχίστου x^*

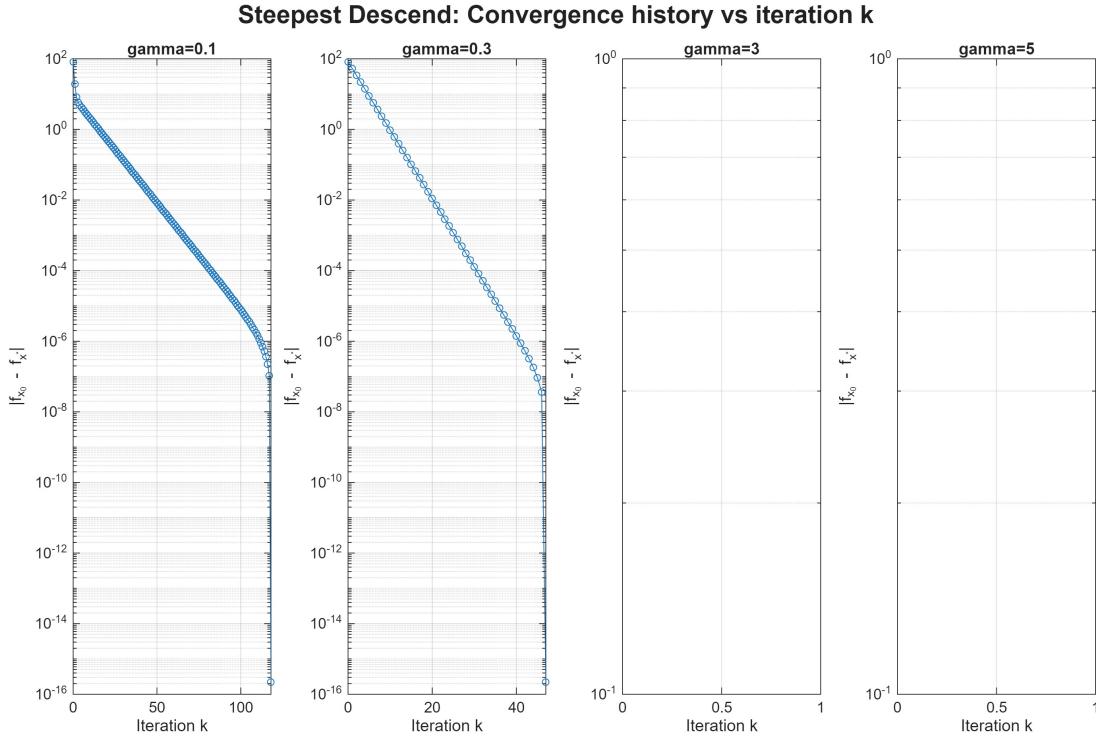
```
1: ορίστε  $x_k \leftarrow x_1$ ,  $k \leftarrow 1$ 
2: while true do
3:    $\bar{x}_k \leftarrow P(x_k - s_k \nabla f(x_k))$  ▷ προβολή του βήματος κατεύθυνσης
4:    $d_k \leftarrow \bar{x}_k - x_k$ 
5:   if  $\|d_k\| < \epsilon$  then
6:     break
7:   end if
8:    $x_{k+1} \leftarrow x_k + \gamma_k d_k$ 
9:    $k \leftarrow k + 1$ 
10: end while
11: return  $x^* \leftarrow x_k$ 
```

Όπου με P συμβολίζεται η πράξη της προβολής πάνω στο εφικτό σύνολο X . Η προβολή εξασφαλίζει πως το σημείο x_{k+1} παραμένει εφικτό.

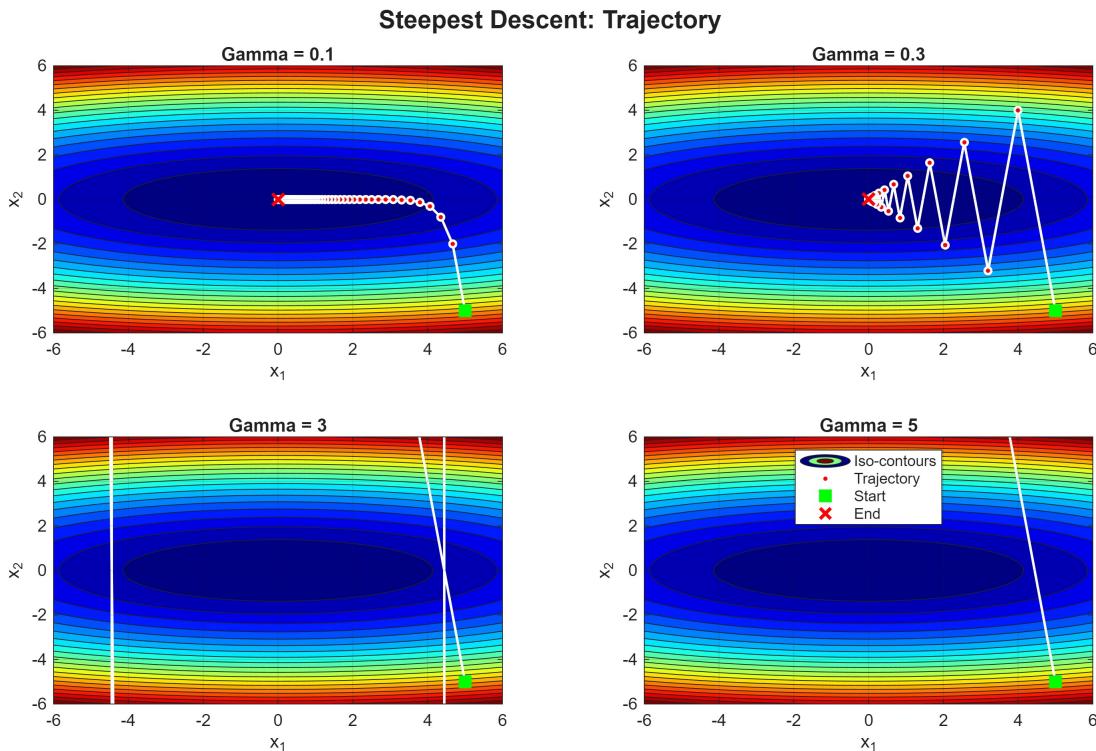
3 Θέμα 1

Στο Θέμα 1, εφαρμόζουμε τη μέθοδο Μέγιστης Καθόδου **χωρίς** Προβολή με βήματα $\gamma_k = [0.1, 0.3, 3, 5]$, από το αρχικό σημείο $x_1 = [5, -5]$ και ακρίβεια $\epsilon = 0.001$.

3.1 Γραφικές Παραστάσεις



Σχήμα 3: Σύγκλιση της f ανά τις επαναλήψεις για διάφορα γ_k



Σχήμα 4: Σύγκλιση της f για διάφορα γ_k

γ	Initial Point	Minimum Point x^*	Minimum Value $f(x^*)$	k	Function Evaluations
0.1	[5 -5]	[0.00146 -5.52e-47]	7.07e-07	118	238
0.3	[5 -5]	[0.000139 0.000139]	6.48e-08	47	96
3	[5 -5]	[-5 NaN]	NaN	251	504
5	[5 -5]	[5.13e+78 NaN]	NaN	212	426

Πίνακας 1: Συγκεντρωτικός πίνακας αποτελεσμάτων (Μέγιστη Κάθοδος)

Στα Σχήματα 3 και 4 παρουσιάζεται η σύγκλιση της f . Αυτή φαίνεται να συγκλίνει για $\gamma_k = 0.1$ και $\gamma_k = 0.3$, ενώ αποτυγχάνει για $\gamma_k = 3$ και $\gamma_k = 5$. Αυτό επιβεβαιώνεται και από τον Πίνακα 1 που δείχνει πως για μεγάλα γ_k δεν υπάρχει σύγκλιση.

3.2 Μαθηματική Ανάλυση

Ακολουθεί μαθηματική ανάλυση, ώστε να ερμηνεύσουμε τα παραπάνω αποτελέσματα. Ξεκινάμε από την αναδρομική σχέση της μεθόδου Μέγιστης Καθόδου:

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

Για μια τετραγωνική συνάρτηση $f(x) = \frac{1}{2}x^\top Qx$ το gradient είναι $\nabla f(x) = Qx$, οπότε η σχέση γίνεται:

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

Στη δεύτερη επανάληψη του αλγορίθμου έχουμε:

$$x_{k+2} = x_{k+1} - \gamma_k Qx_{k+1} = (I - \gamma_k Q)x_{k+1} = (I - \gamma_k Q)^2 x_k$$

Στη n-οστή επανάληψη του αλγορίθμου έχουμε:

$$x_{k+n} = (I - \gamma_k Q)^n x_k$$

Συνεπώς, για να συγκλίνει η μέθοδος θα πρέπει να ισχύει:

$$\|I - \gamma_k Q\| < 1$$

Στην περίπτωσή μας έχουμε

$$Q = \begin{bmatrix} \frac{2}{3} & 0 \\ 0 & 6 \end{bmatrix}$$

Άρα:

- $|1 - \frac{2}{3}\gamma_k| < 1 \implies -1 < 1 - \frac{2}{3}\gamma_k < 1 \implies 0 < \gamma_k < 3$
- $|1 - 6\gamma_k| < 1 \implies -1 < 1 - 6\gamma_k < 1 \implies 0 < \gamma_k < \frac{1}{3}$

Συγκεντρωτικά, για να συγκλίνει η μέθοδος θα πρέπει να ισχύει:

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

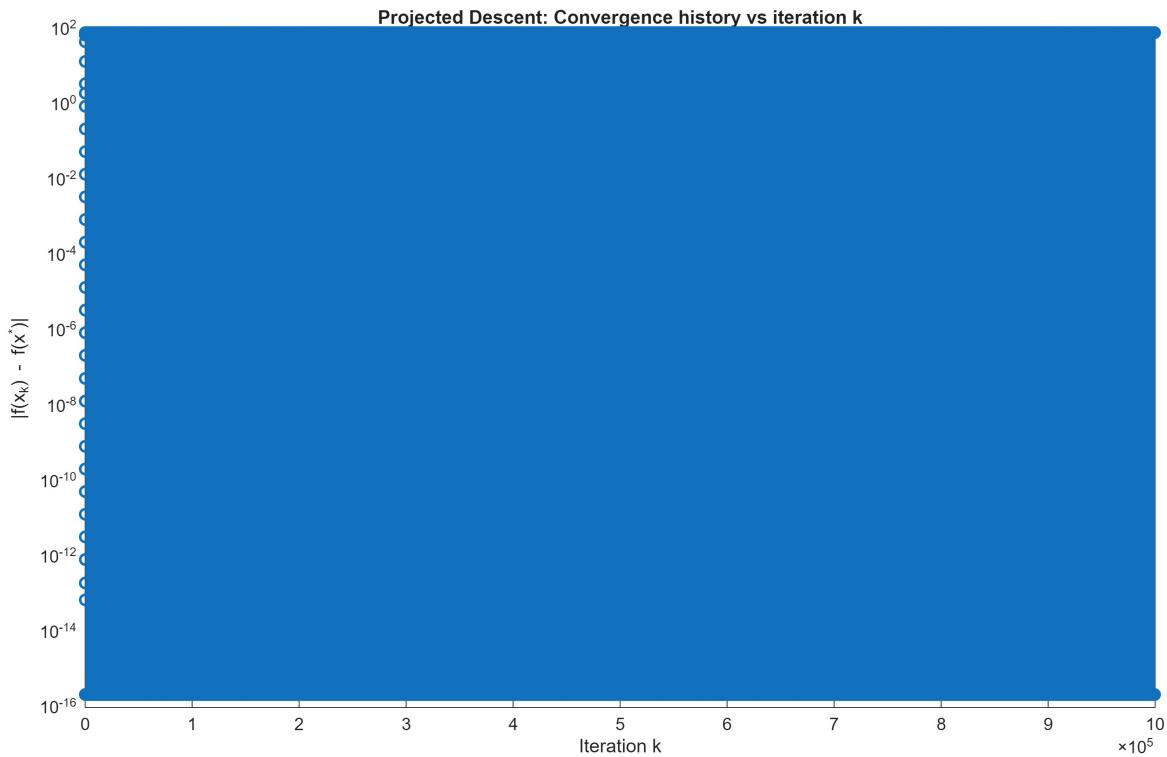
3.3 Συμπεράσματα

Η παραπάνω μαθηματική ανάλυση αποδεικνύει με μαθηματική αυστηρότητα πως για $\gamma_k > \frac{1}{3}$ η μέθοδος στη συγκεκριμένη συνάρτηση δεν συγκλίνει. Η παρατήρηση αυτή συνάδει με τα πειραματικά μας ευρήματα. Μάλιστα, παρατηρούμε πως για $\gamma_k = 0.3$, δηλαδή χοντά στο όριο $\frac{1}{3} = 0.333\dots$, υπάρχει ταλάντωση στην τροχιά της σύγκλισης, με ζιχ-ζαχ μοτίβο.

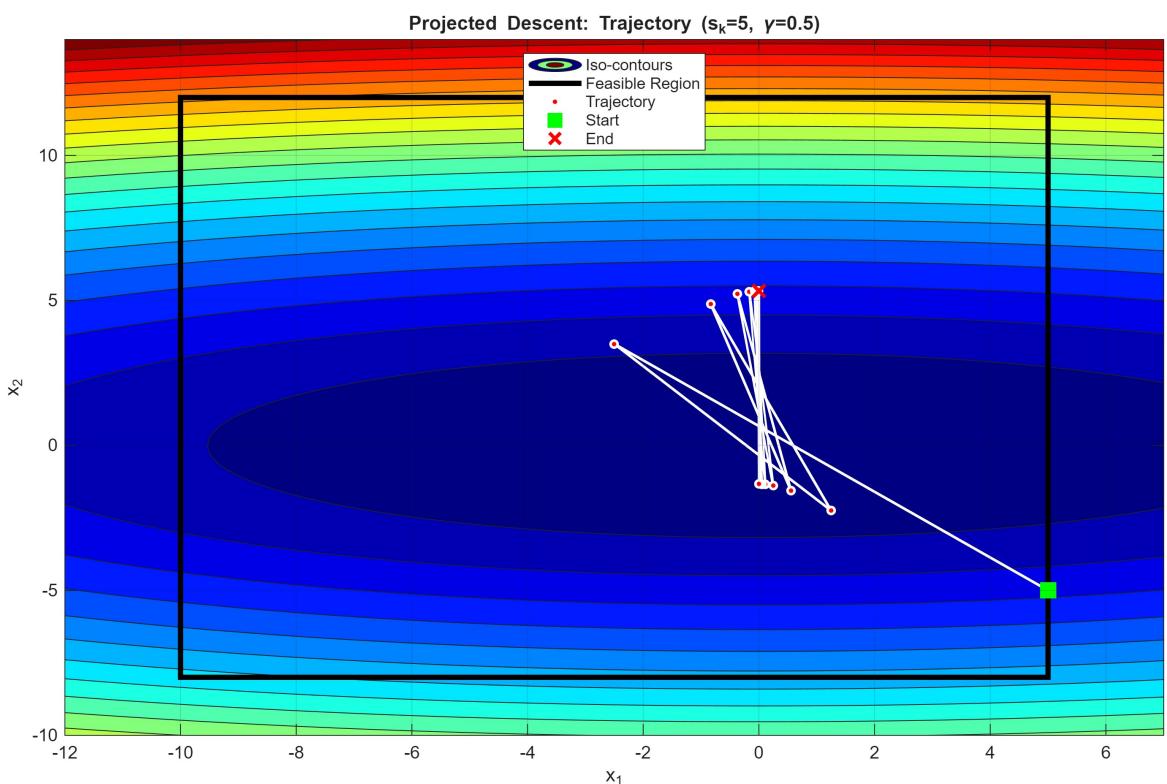
4 Θέμα 2

Στο Θέμα 2, εφαρμόζουμε τη μέθοδο Μέγιστης Καυθόδου με Προβολή με βήματα $s_k = 5$, $\gamma_k = 0.5$, από το αρχικό σημείο $x_1 = [5, -5]$ και ακρίβεια $\epsilon = 0.001$.

4.1 Γραφικές Παραστάσεις



Σχήμα 5: Σύγκλιση της f ανά τις επαναλήψεις



Σχήμα 6: Σύγκλιση της f

Στα Σχήματα 5 και 6 παρουσιάζεται η σύγκλιση της f . Αυτή φαίνεται να μην καταφέρνει να συγκλίνει για $s_k = 5$ και $\gamma_k = 0.5$, αντ' αυτού τριγυρνάει γύρω από το σημείο του ολικού ελαχίστου. Ο αλγόριθμος τερματίζει τεχνητά, αφού υπερβεί το όριο μέγιστων επαναλήψεων που έχει τεθεί. Το Σχήμα 5 φαίνεται «γεμισμένο» λόγω των πολλών επαναλήψεων και της περιοδικής κίνησης γύρω από το σημείο του ελαχίστου.

4.2 Συμπεράσματα

Παρατηρούμε πως παρόλο που η χρήση της προβολής εγγυάται την εφικτότητα των σημείων, δεν μπορεί να εγγυηθεί τη σύγκλιση. Η συγκεκριμένη επιλογή των βημάτων s_k και γ_k οδηγεί τον αλγόριθμο σε αστάθεια.

Στη μέθοδο με προβολή το νέο σημείο υπολογίζεται ως:

$$x_{k+1} = x_k + \gamma_k(\bar{x}_k - x_k), \quad \bar{x}_k = P(x_k - s_k \nabla f(x_k))$$

Όταν το σημείο $x_k - s_k \nabla f(x_k)$ βρίσκεται εντός του X η προβολή δεν αλλάζει τίποτα, οπότε ο κανόνας ενημέρωσης ταυτίζεται με τη μέθοδο μέγιστης καθόδου με βήμα:

$$\gamma'_k = s_k \gamma_k$$

Η σύγκλιση επομένως υπακούει στην ίδια συνθήκη του Θέματος 1, δηλαδή:

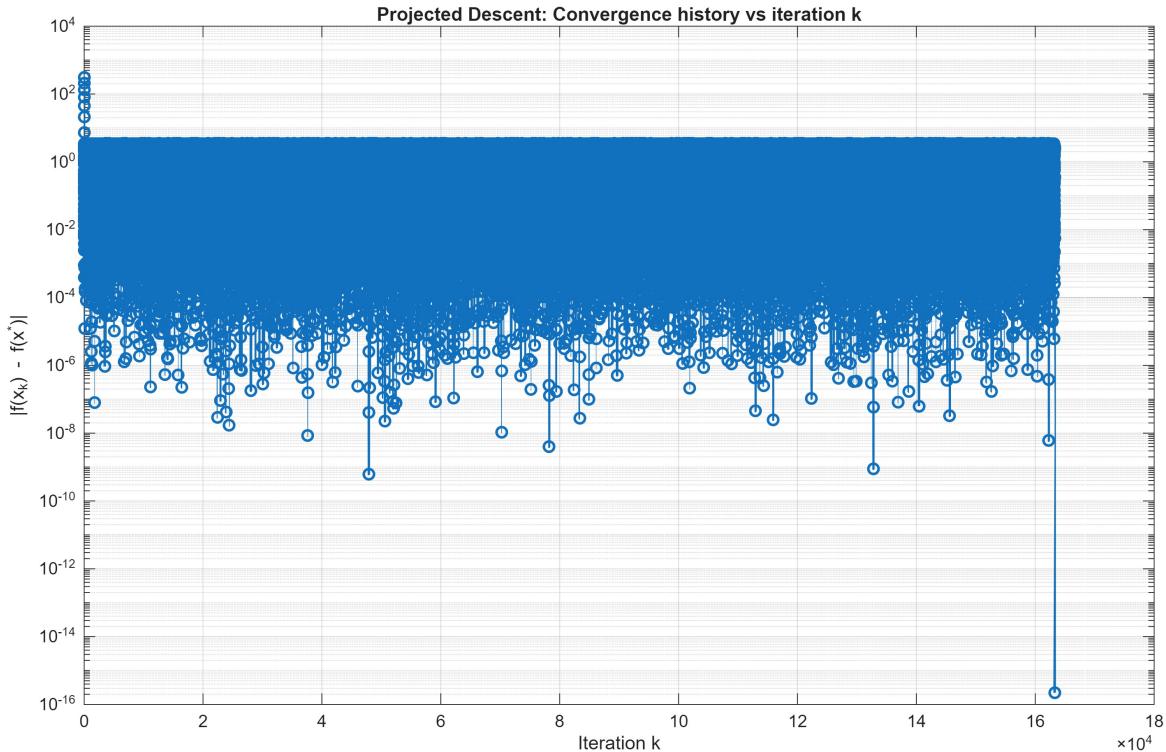
$$0 < \gamma'_k < \frac{1}{3}$$

Στην περίπτωσή μας $\gamma'_k = s_k \gamma_k = 2.5 > \frac{1}{3}$, οπότε η μη σύγκλιση είναι αναμενόμενη ανεξάρτητα από το ότι τα σημεία παραμένουν εφικτά.

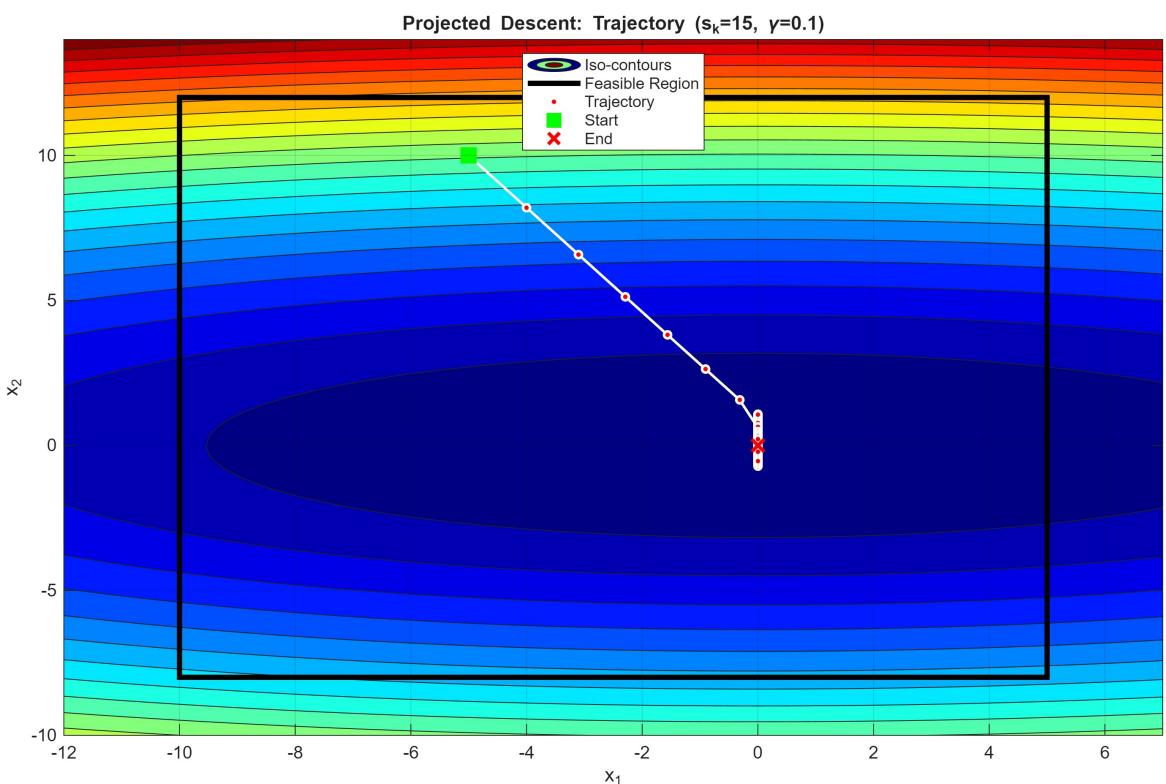
5 Θέμα 3

Στο Θέμα 3, εφαρμόζουμε τη μέθοδο Μέγιστης Καθόδου με Προβολή με βήματα $s_k = 15$, $\gamma_k = 0.1$, από το αρχικό σημείο $x_1 = [-5, 10]$ και ακρίβεια $\epsilon = 0.001$.

5.1 Γραφικές Παραστάσεις



Σχήμα 7: Σύγκλιση της f ανά τις επαναλήψεις



Σχήμα 8: Σύγκλιση της f

Στα Σχήματα 7 και 8 παρουσιάζεται η σύγκλιση της f . Αυτή φαίνεται να συγκλίνει για $s_k = 15$ και $\gamma_k = 0.1$. Ο αλγόριθμος επιτυγχάνει στο να προσδιορίσει το σημείο ελαχίστου, έπειτα όμως από πολλές επαναλήψεις, που επισημαίνουν την ανάγκη για περαιτέρω ρύθμιση των βημάτων s_k και γ_k .

5.2 Συμπεράσματα

Ο αλγόριθμος είναι επιτυχής, αλλά αναποτελεσματικός. Προσδιορίζει το ελάχιστο σε πολλές επαναλήψεις, πράγμα που τον καθιστά αργό.

Μία απλή προσέγγιση ώστε να λυθεί το θέμα που προκύπτει όταν να επιλέγαμε μικρότερο s_k (*π.χ. $s_k = 3$*) έτσι ώστε αν απαλειφθεί η ταλάντωση και η μέθοδος να συγκλίνει πιο ομαλά.

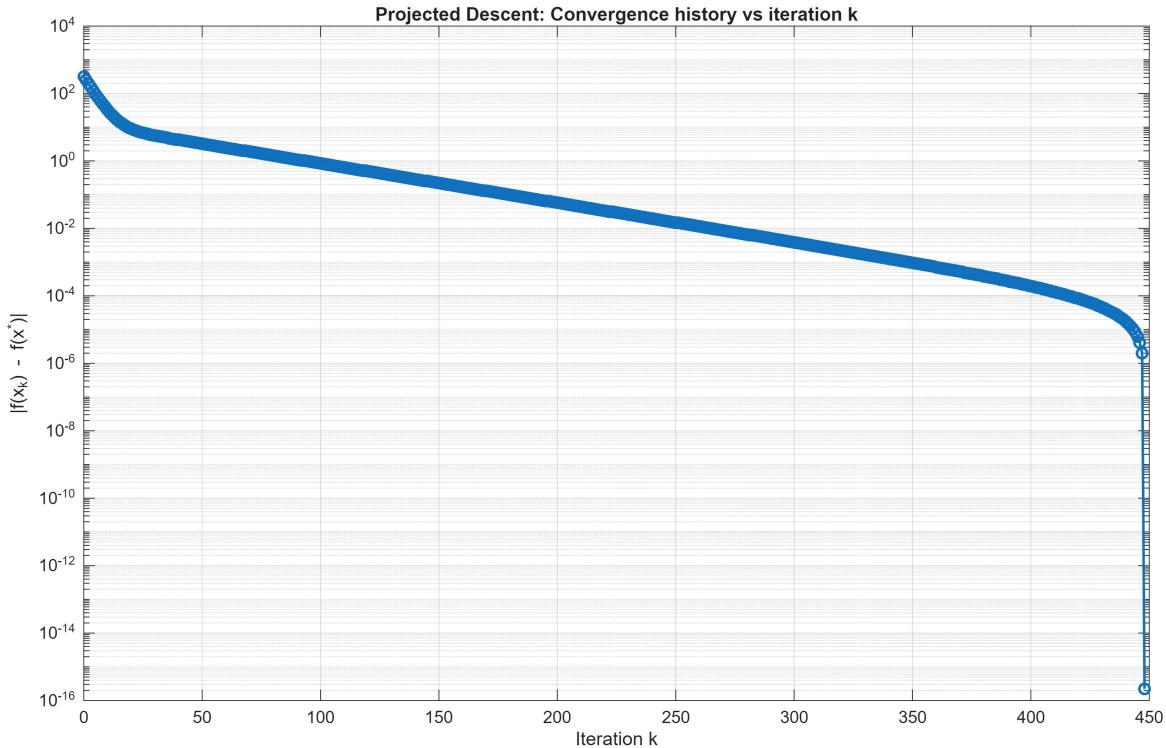
Ένας άλλος πρακτικός τρόπος ώστε η μέθοδος να συγκλίνει με αξιοπιστία είναι η δυναμική προσαρμογή του βήματος. Συγκεκριμένα, χάθε φορά που παρατηρείται αύξηση της τιμής της $f(x_k)$ ή ταλάντωση στην πορεία του x_k , μπορούμε να μειώσουμε το s_k κατά έναν σταθερό παράγοντα (*π.χ. $s_{k+1} = 0.5 s_k$*). Η διαδικασία αυτή εξασφαλίζει ότι τελικά το βήμα $\gamma'_k = s_k \gamma_k$ εισέρχεται εντός του διαστήματος σύγκλισης $0 < s_k \gamma_k < \frac{1}{3}$, επιταχύνοντας ταυτόχρονα τη σταθεροποίηση της τροχιάς.

Τέλος, παρατηρούμε πως όπως και στα Θέματα 1 και 2 η σύγκλιση υπακούει στην ίδια σχέση.

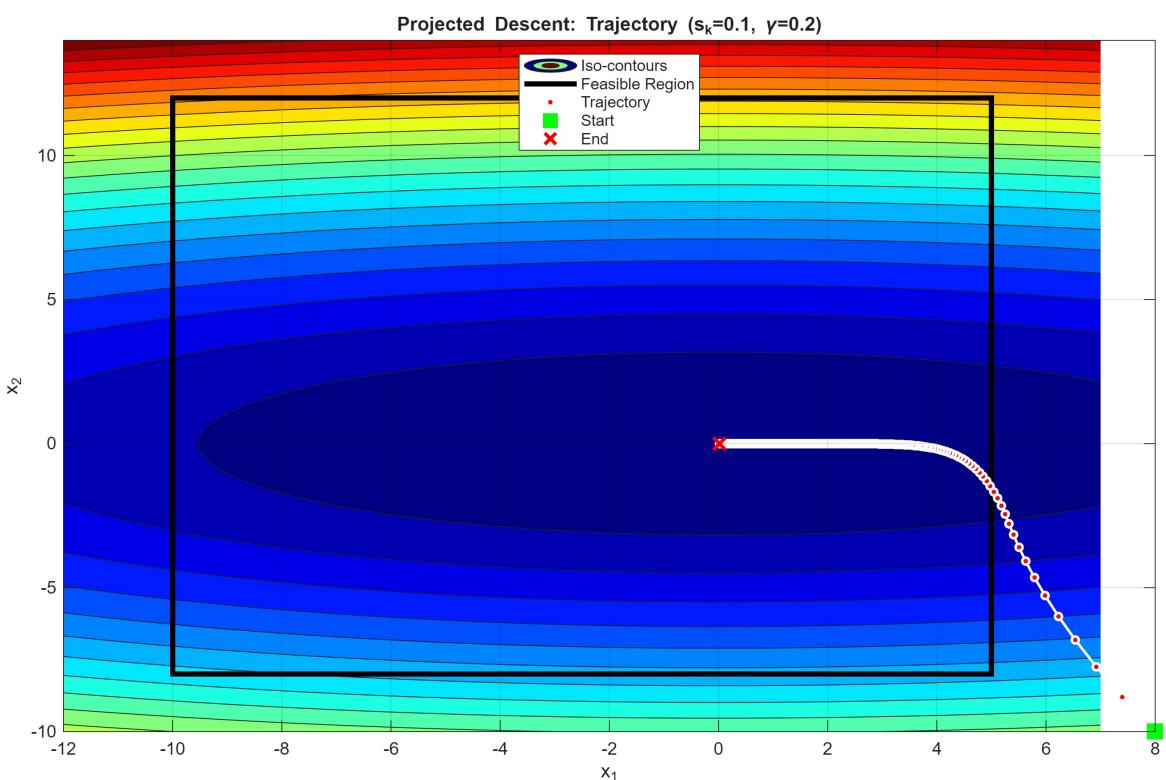
6 Θέμα 4

Στο Θέμα 4, εφαρμόζουμε τη μέθοδο Μέγιστης Καθόδου με Προβολή με βήματα $s_k = 0.1$, $\gamma_k = 0.2$, από το αρχικό σημείο $x_1 = [8, -10]$ και ακρίβεια $\epsilon = 0.001$.

6.1 Γραφικές Παραστάσεις



Σχήμα 9: Σύγκλιση της f ανά τις επαναλήψεις



Σχήμα 10: Σύγκλιση της f

Στα Σχήματα 9 και 10 παρουσιάζεται η σύγκλιση της f . Αυτή φαίνεται να συγκλίνει για $s_k = 0.1$ και $\gamma_k = 0.2$. Ο αλγόριθμος επιτυγχάνει στο να προσδιορίσει το σημείο ελαχίστου, παρά το γεγονός πως το αρχικό σημείο $x_1 = [8, -10]$ δεν ανήκει στο εφικτό σύνολο. Αυτό εξηγείται λόγω της χρήσης της προβολής, που οδηγεί τον αλγόριθμο στην εφικτή περιοχή.

6.2 Συμπεράσματα

Παρατηρούμε πως ακόμα και πριν την εκτέλεση μπορούμε να αξιολογήσουμε τη δυνατότητα σύγκλισης. Πράγματι, $s_k \gamma_k = 0.1 \cdot 0.2 = 0.02 < \frac{1}{3}$, άρα το βήμα βρίσκεται εντός του θεωρητικού ορίου σύγκλισης των προηγούμενων Θεμάτων. Επομένως αναμένεται ότι η μέθοδος θα συγκλίνει, εφόσον ο τελεστής προβολής διατηρεί τα σημεία εντός του X .

Επιπλέον το αρχικό σημείο $(8, -10)$ βρίσκεται εκτός του X , όμως η πρώτη πράξη προβολής το μεταφέρει στο κοντινότερο εφικτό σημείο. Από εκεί και πέρα η μέθοδος συμπεριφέρεται όπως η μέθοδος μέγιστης καθόδου με μικρό σταθερό βήμα, οπότε η σύγκλιση είναι αναμενόμενη από πριν.

Ένας τρόπος να βελτιώσουμε περαιτέρω την απόδοση του παραπάνω αλγορίθμου θα ήταν να βρούμε την προβολή του αρχικού σημείου $x_1 = [8, -10]$ στην εφικτή περιοχή X , που είναι το σημείο $[5, -8]$ (η γωνία του κουτιού), κι έπειτα να τρέξουμε τον αλγόριθμο από εκεί. Έτσι θα μειωνόταν οι συνολικές επαναλήψεις του αλγορίθμου.

A' Εργαλεία

Τα εργαλεία που χρησιμοποιήθηκαν κατά την εκπόνηση της εργασίας είναι τα εξής:

- MATLAB
- LaTeX
- Git

B' Βιβλιογραφία

Αναφορές

[1] Γ. Ροβισάκης, *Τεχνικές Βελτιστοποίησης*, Εκδόσεις Τζιόλα, 2007.