



ΑΡΙΣΤΟΤΕΛΕΙΟ  
ΠΑΝΕΠΙΣΤΗΜΙΟ  
ΘΕΣΣΑΛΟΝΙΚΗΣ

ΤΜΗΜΑ ΗΜΜΥ. ΤΟΜΕΑΣ ΗΛΕΚΤΡΟΝΙΚΗΣ  
ΤΕΧΝΙΚΕΣ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ

---

## 3η Εργαστηριακή Άσκηση Μέθοδος Μέγιστης Καθόδου με Προβολή

---

Συντάκτης:  
Χρήστος Χουτουρίδης  
ΑΕΜ:8997  
[cchoutou@ece.auth.gr](mailto:cchoutou@ece.auth.gr)

Διδάσκοντες:  
Γ. Ροβιθάκης  
[rovithak@auth.gr](mailto:rovithak@auth.gr)  
Θ. Αφορόζη  
[taforozi@ece.auth.gr](mailto:taforozi@ece.auth.gr)

## 1. ΕΙΣΑΓΩΓΗ

Η παρούσα εργασία αφορά το πρόβλημα της ελαχιστοποίησης μιας δοσμένης συνάρτησης πολλών μεταβλητών  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  με περιορισμούς χρησιμοποιώντας τη μέθοδο μέγιστης καθόδου με προβολή. Η μέθοδος αυτή θα εκτελεστεί σε αντιπαραβολή με την αντίστοιχη μέθοδο χωρίς περιορισμούς από την προηγούμενη εργασία. Για το λόγο αυτό χρησιμοποιούμε τον κώδικα της προηγούμενης εργασίας με κάποιες τροποποιήσεις, όπως θα δούμε και παρακάτω.

### 1.1. Παραδοτέα

Τα παραδοτέα της εργασίας αποτελούνται από:

- Την παρούσα αναφορά.
- Τον κατάλογο `scripts/`, που περιέχει τον κώδικα της MATLAB.
- Το [σύνδεσμο](#) με το αποθετήριο που περιέχει όλο το project με τον κώδικα της MATLAB, της αναφοράς και τα παραδοτέα.

### 1.2. Προγραμματιστική προσέγγιση

Για τον προγραμματισμό και εκτέλεση των μεθόδων της παρούσας εργασίας έγινε χρήση της MATLAB. Στον κατάλογο `scripts`, περιέχονται όλες οι μέθοδοι και οι τεχνικές υπολογισμού βημάτων με τη μορφή συναρτήσεων καθώς και `scripts` που τις καλούν. Για κάθε ένα θέμα της εργασίας, υπάρχει το αντίστοιχο script που περιέχει τους υπολογισμούς, τις κλήσεις των μεθόδων και τη δημιουργία των διαγραμμάτων. Για το πρώτο θέμα το αρχείο `Script_1_SteepDesc.m` για το δεύτερο το `Script_2_SteepDesc_Proj.m` και ούτω καθεξής. Η μέθοδος μέγιστης καθόδου (αρχείο: `method_SteepDesc.m`) είναι η ίδια με αυτή της προηγούμενης εργασίας με τη μόνη διαφορά ότι τροποποιήθηκε ώστε η αντικειμενική συνάρτηση να δέχεται διάνυσμα ως όρισμα και όχι δύο διαφορετικές μεταβλητές  $x, y$ . Αυτό ακολουθήθηκε και για την έκδοση με προβολή και φυσικά είχε αντίκτυπο και στις υπόλοιπες συναρτήσεις, όπως η κλίση ή ο Εσσιανός. Στην παρούσα εργασία η υλοποίηση του κώδικα ακολουθεί την προσέγγιση των προηγούμενων εργασιών. Πιο συγκεκριμένα.

### 1.3. Μέθοδοι επιλογής βήματος

Εφόσον οι υπάρχουσες μέθοδοι επιλογής βήματος είναι ανεξάρτητες από την μέθοδο υπολογισμού του ελάχιστου και εφόσον χρησιμοποιούμε τον ίδιο κώδικα και στην παρούσα εργασία, αυτός ο τρόπος σχεδίασης παρέμεινε. Ουσιαστικά για κάθε ένα τρόπο υπολογισμού του  $\gamma_k$ , υπάρχει αντίστοιχη συνάρτηση, με κοινό interface. Αυτό έχει τη μορφή: `gamma_ <method> (f, grad_f, dk, xk)`, όπου το `f` είναι η αντικειμενική συνάρτηση, `grad_f` η συνάρτηση κλίσης της, `dk` η τιμή της συνάρτησης κλίσης στο `xk` και `xk` το σημείο ενδιαφέροντος. Έτσι οι μέθοδοι αντιγράφηκαν και εδώ για ολότητα, ακόμα και αν για την παρούσα εργασία χρησιμοποιείται μόνο το σταθερό βήμα  $\gamma_k$ .

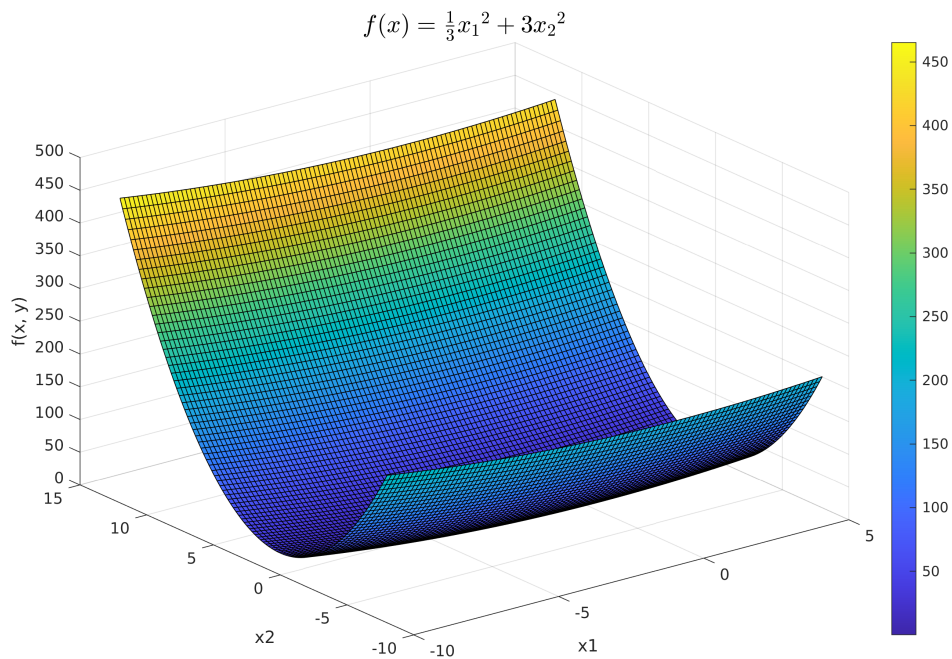
## 2. ΑΠΕΙΚΟΝΙΣΗ ΤΗΣ ΣΥΝΑΡΤΗΣΗΣ

Η συνάρτηση με την οποία ασχολούμαστε στην παρούσα εργασία είναι η:

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, f(x) = \frac{1}{3}x_1^2 + 3x_2^2, \quad x = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T \quad (1)$$

Με σύνολο περιορισμών  $X$ :

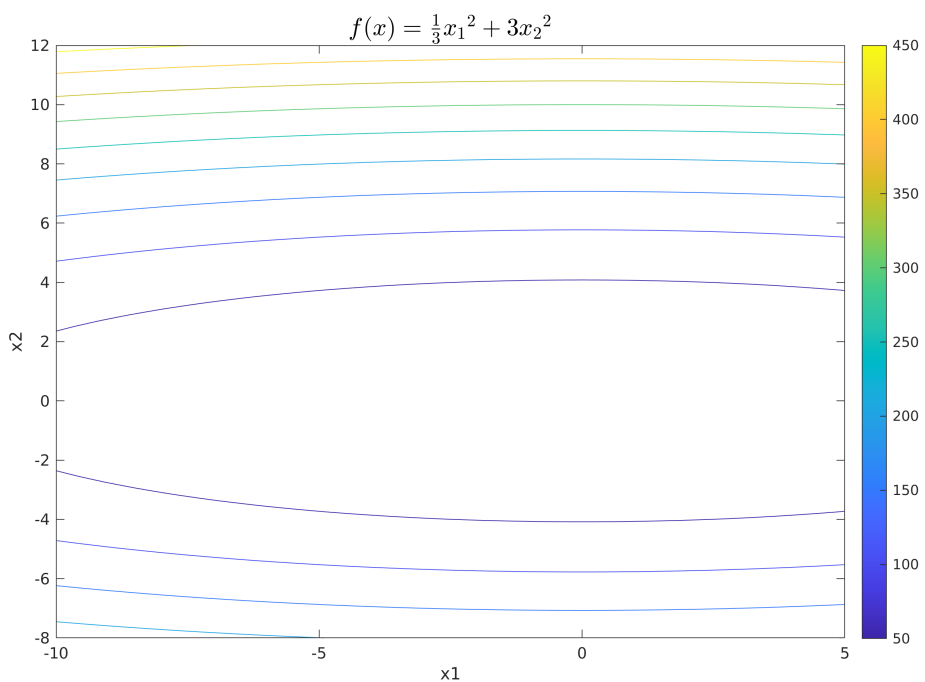
$$\forall x = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T \in X \subset \mathbb{R}^2 : \quad -10 \leq x_1 \leq 5, \quad -8 \leq x_2 \leq 12$$



Σχήμα 1: Γραφική παράσταση της  $f$

Στο σχήμα 1 φαίνεται η τρισδιάστατη απεικόνιση της συνάρτησης.

Από το σχήμα μπορούμε πολύ εύκολα να διακρίνουμε ότι η συνάρτηση είναι κυρτή στο σύνολο των περιορισμών της εκφώνησης  $-10 \leq x_1 \leq 5$  και  $-8 \leq x_2 \leq 12$ . Για να πάρουμε μια καλύτερη αίσθηση για τις κλίσεις της  $f$ , παρακάτω παραθέτουμε ένα γράφημα με τις ισοβαρείς καμπύλες της  $f$ .

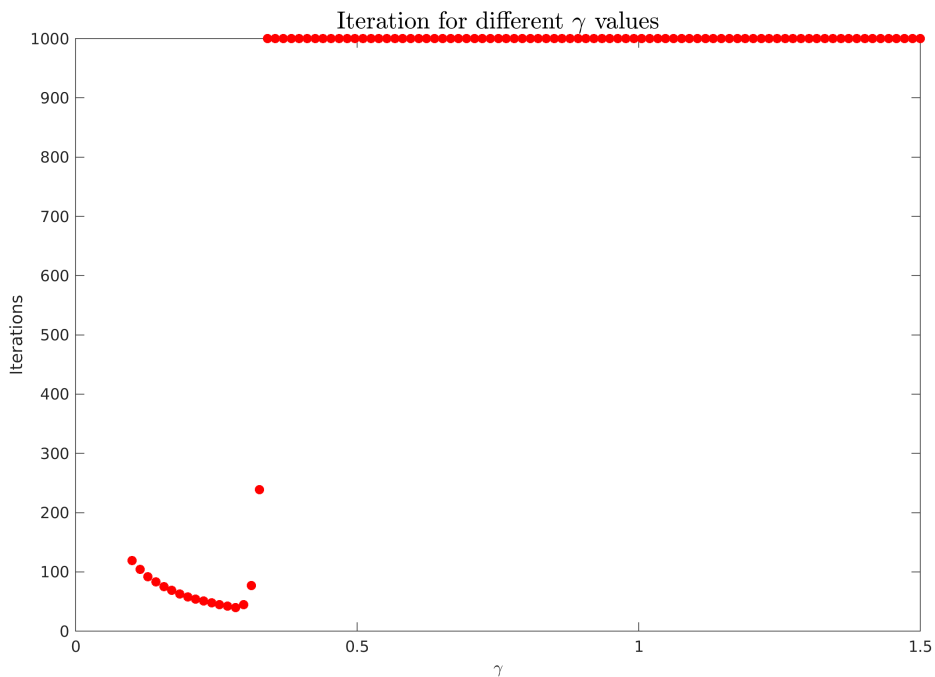


Σχήμα 2: Ισοβαρείς της  $f$

Από το παραπάνω σχήμα 2 φαίνονται και γραφικά οι μικρές κλίσης που παρουσιάζει η συνάρτηση κοντά στο ελάχιστο σημείο (0, 0). Τα διαγράμματα για τη μέθοδο δημιουργούνται εκτελώντας το αρχείο `Script_0_Plots.m`

### 3. ΘΕΜΑ 1 - ΜΕΘΟΔΟΣ ΜΕΓΙΣΤΗΣ ΚΑΘΟΔΟΥ ΧΩΡΙΣ ΠΕΡΙΟΡΙΣΜΟΥΣ

Εφαρμόζοντας την μέθοδο μέγιστης καθόδου από την προηγούμενη εργασία, με ακρίβεια  $\epsilon = 0.001$ , για τα βήματα  $\gamma_k$  της εκφώνησης, παρατηρούμε ότι η μέθοδος συγκλίνει στο ελάχιστο για μικρά  $\gamma_k$  ενώ **αποκλίνει για μεγάλα  $\gamma_k > 0.33$** . Από τις δοκιμές φαίνεται ότι το σημείο εκκίνησης δεν παίζει ρόλο και για αυτό επιλέξαμε να παραθέσουμε τα ευρήματά μας από το σημείο (5, -5), για αντιπαραβολή με το επόμενο βήμα της εκφώνησης.



**Σχήμα 3:** Αριθμός επαναλήψεων για διαφορετικές τιμές  $\gamma_k$  [Μέγιστη Κάθοδος].

Επίσης παρατηρούμε ότι για μικρό  $\gamma_k = 0.1$  η σύγκλιση είναι ομαλή, ενώ για μεγάλο  $\gamma_k = 0.3$  παρουσιάζει ταλάντωση κατά την σύγκλιση. Παρακάτω στο σχήμα 4 παραθέτουμε την πορεία σύγκλισης και απόκλισης για τις διαφορετικές τιμές του  $\gamma_k$ .

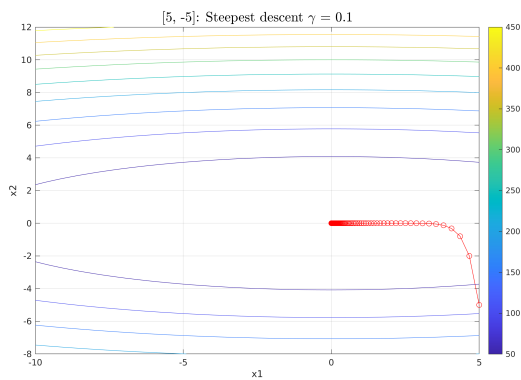
#### 3.1. Μαθηματική ανάλυση

Τα παραπάνω αποτελέσματα επιβεβαιώνονται και θεωρητικά. Πιο συγκεκριμένα για τη σύγκλιση της μεθόδου μέγιστης καθόδου όπου το κάθε σημείο υπολογίζεται από την σχέση:

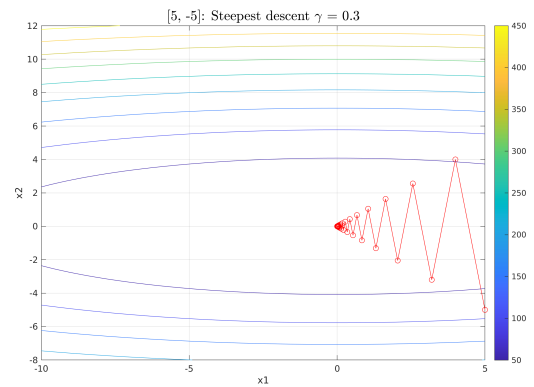
$$\mathbf{x}_{k+1} = \mathbf{x}_k - \gamma_k \nabla f(\mathbf{x}_k)$$

Πρέπει να ισχύουν:

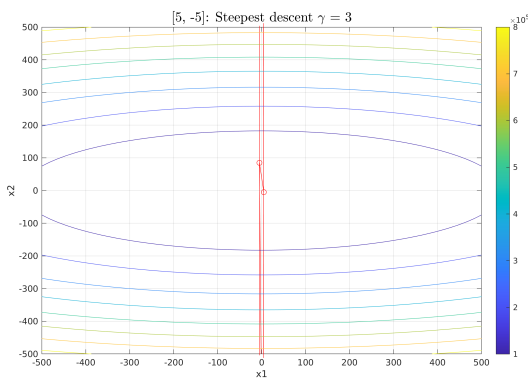
1. Η  $f$  να είναι κυρτή.
2. Η  $f$  να είναι συνεχής και διαφορίσιμη και η κλίση της υπολογίσιμη.



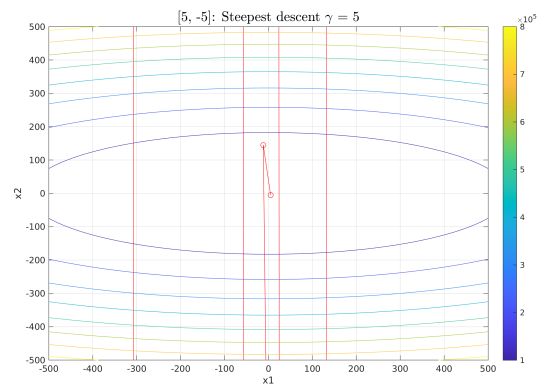
$\gamma_k = 0.1$



$\gamma_k = 0.3$



$\gamma_k = 3$



$\gamma_k = 5$

**Σχήμα 4:** Σύγκριση της μεθόδου Steepest descent για διαφορετικά  $\gamma_k$

3. Για το βήμα υπολογισμού να ισχύει η σχέση:

$$0 < \gamma_k < \frac{2}{L} \tag{2}$$

Όπου  $L$  το άνω φράγμα της Lipschitz για την κλίση  $\nabla f(x)$  (αν είναι γνωστή), η οποία είναι η μέγιστη ιδιοτιμή του Εσσιανού και δίνεται από τη σχέση:

$$L = \max_x \{\lambda_{\max}(H(x))\} \tag{3}$$

Έτσι για τη δική μας περίπτωση, όπου η  $f$  είναι κυρτή και διαφορίσιμη, η κλίση της  $f(x)$  είναι:

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

Ο Εσσιανός πίνακας  $H(x)$  της  $f(x)$  είναι:

$$H(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & 0 \\ 0 & 6 \end{bmatrix}$$

Και εφόσον είναι διαγώνιος οι ιδιοτιμές του είναι οι τιμές της διαγωνίου. Δηλαδή:

$$\det(H - \lambda I) = 0 \Leftrightarrow \det\left(\begin{bmatrix} \frac{2}{3} - \lambda & 0 \\ 0 & 6 - \lambda \end{bmatrix}\right) = 0 \Leftrightarrow \left(\frac{2}{3} - \lambda\right)(6 - \lambda) = 0 \Leftrightarrow \lambda_{\min} = \frac{2}{3}, \quad \lambda_{\max} = 6.$$

Έτσι από τις εξισώσεις (2) και (3) προκύπτει τελικά ότι  $\frac{2}{L} = \frac{2}{\lambda_{max}} = \frac{1}{3}$  και επομένως για σύγκλιση της μεθόδου πρέπει να ισχύει:

$$0 < \gamma_k < \frac{1}{3} \tag{4}$$

Βλέπουμε ότι από την ανάλυσή μας, η τιμή που βρήκαμε εμπειρικά για την απόκλιση, από την εκτέλεση του αλγορίθμου για το  $\gamma > 0.33$  επιβεβαιώνεται.

Επίσης το γεγονός ότι ο Εσσιανός είναι διαγώνιος, μας δίνει επίσης πληροφορία ότι:

- Η ιδιοτιμή  $\lambda_1 = \frac{2}{3}$  εφαρμοσμένη στη σχέση (2) μας δίνει:

$$0 < \gamma_{k,x_1} < 3 \tag{5}$$

που είναι η τιμή για την οποία η μέθοδος συγκλίνει στη διάσταση  $x_1$

- Η ιδιοτιμή  $\lambda_2 = 6$  εφαρμοσμένη στη σχέση (2) μας δίνει:

$$0 < \gamma_{k,x_2} < \frac{1}{3} \tag{6}$$

που είναι η τιμή για την οποία η μέθοδος συγκλίνει στη διάσταση  $x_2$

Εναλλακτικά

Αν θέλαμε να βρούμε το κριτήριο σύγκλισης για το βήμα  $\gamma_k$  ξεχωριστά για την κάθε διάσταση θα μπορούσαμε να θεωρήσουμε ισοδύναμα με την παραπάνω ανάλυση, πως για να συγκλίνει η μέθοδος θα πρέπει να ισχύει:

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

Από την παραπάνω εξίσωση προκύπτει πάλι ένα σύστημα εξισώσεων για την κάθε διάσταση:

$$\begin{cases} \left| \frac{x_{1,k} - \frac{2}{3}\gamma_k x_{1,k}}{x_{1,k}} \right| < 1 \\ \left| \frac{x_{2,k} - 6\gamma_k x_{2,k}}{x_{2,k}} \right| < 1 \end{cases} \Leftrightarrow \begin{cases} \left| 1 - \frac{2}{3}\gamma_k \right| < 1 \\ |1 - 6\gamma_k| < 1 \end{cases} \Leftrightarrow \begin{cases} 0 < \gamma_k < 3 \\ 0 < \gamma_k < \frac{1}{3} \end{cases}$$

Που επιβεβαιώνει την προηγούμενη ανάλυσή μας για τη συνολική σύγκλιση, αλλά και για την κάθε διάσταση ξεχωριστά.

4. ΜΕΘΟΔΟΣ ΜΕΓΙΣΤΗΣ ΚΑΘΟΔΟΥ ΜΕ ΠΡΟΒΟΛΗ

Πριν περάσουμε στις υπόλοιπες απαιτήσεις της εργασίας θα θέλαμε να παραθέσουμε κάποιες πληροφορίες για την υλοποίηση της μεθόδου μέγιστης καθόδου με προβολή (αρχείο: `method_SteepDesc_Proj.m`). Η συνάρτηση αυτή δέχεται ως είσοδο την αντικειμενική συνάρτηση και την συνάρτηση κλίσης καθώς και το σημείο εκκίνησης  $x_k$  και το βήμα  $s_k$ . Με τη βοήθεια της συνάρτησης `ProjectionPoint()` παίρνει πρώτα την προβολή του  $x_k$  στο διάστημα των περιορισμών, αν αυτό χρειάζεται, και έπειτα εφαρμόζει τον αλγόριθμο. Αυτό σημαίνει ότι μπορεί να χρησιμοποιηθεί και για σημεία εκκίνησης εκτός του συνόλου των περιορισμών. Ο αλγόριθμος είναι παρόμοιος με αυτόν της προηγούμενης εργασίας με τη διαφορά ότι η διεύθυνση  $d_k$  επιλέγεται από τη σχέση:

$$d_k = Pr_X\{x_k - s_k \nabla f(x_k)\} - x_k = \bar{x}_k - x_k$$

και τελικά έχουμε:

$$x_{k+1} = x_k + \gamma_k d_k$$

Δηλαδή εφαρμόζουμε πρώτα τη μέθοδο μέγιστης καθόδου με βήμα  $s_k$  στην κατεύθυνση  $-\nabla f$  και έπειτα προβάλλουμε το σημείο στο σύνολο  $X$  και χρησιμοποιούμε αυτό ως διεύθυνση με βήμα  $\gamma_k$ .

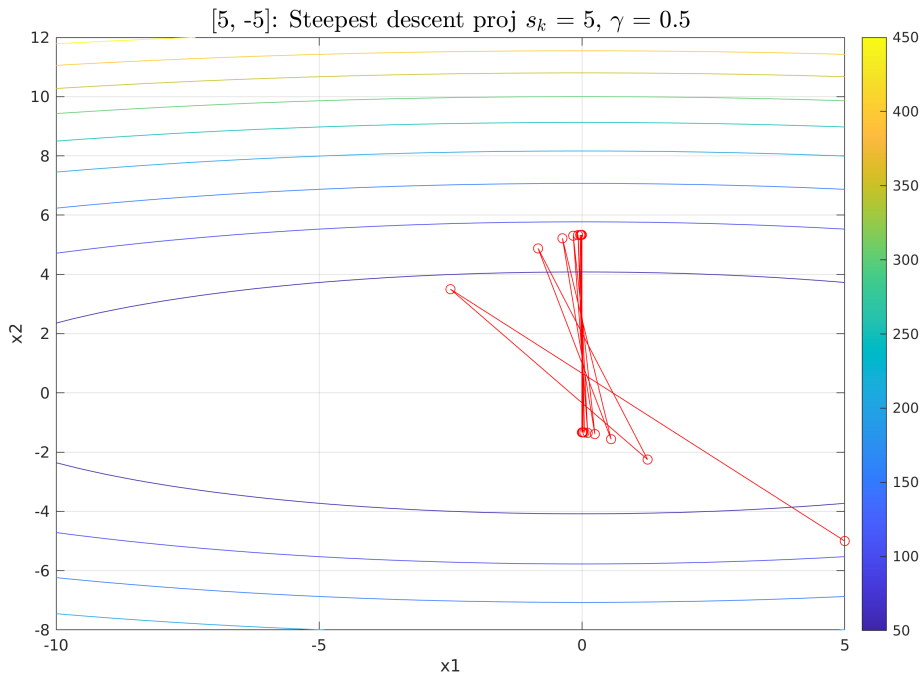
Αξίζει να παρατηρήσουμε πως αν το σημείο  $x_k$  είναι εντός του συνόλου περιορισμών τότε  $\bar{x}_k = x_k$ , δηλαδή η προβολή του σημείου στο σύνολο, είναι το ίδιο το σημείο. Έτσι:

$$\begin{aligned} x_{k+1} &= x_k + \gamma_k (Pr_X\{x_k - s_k \nabla f(x_k)\} - x_k) \\ &= x_k + \gamma_k ((x_k - s_k \nabla f(x_k)) - x_k) \\ &= x_k - \gamma_k s_k \nabla f(x_k) \\ &= x_k - \gamma'_k \nabla f(x_k) \end{aligned}$$

Βλέπουμε λοιπόν ότι για σημεία εσωτερικά του  $X$ , έχουμε ουσιαστικά τη μέθοδο της μέγιστης καθόδου με βήμα  $\gamma'_k = \gamma_k s_k$ . Εφόσον όμως για μη εφικτά σημεία, πριν εφαρμόσουμε τη μέθοδο, τα προβάλλουμε στο σύνολο των περιορισμών  $X$ , βλέπουμε πως η μέθοδος τελικά προσεγγίζει τη μέγιστη κάθοδο με βήμα  $\gamma'_k$ , που είναι και το βήμα για το οποίο ελέγχουμε αν ισχύουν τα κριτήρια σύγκλισης.

5. ΘΕΜΑ 2 - ΣΗΜΕΙΟ (5, -5),  $s_k = 5, \gamma_k = 0.5$

Εφαρμόζοντας τη μέθοδο για ακρίβεια  $\epsilon = 0.01, s_k = 5$  και  $\gamma_k = 0.5$  έχουμε  $\gamma'_k = 2.5 > \frac{1}{3}$ , άρα το κριτήριο σύγκλισης δεν πληρείται.

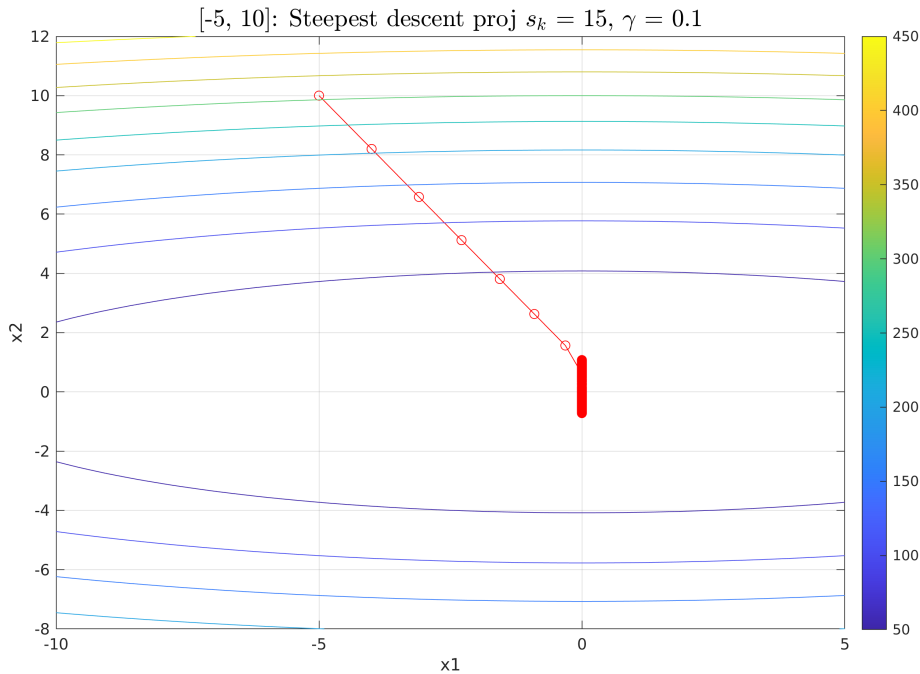


Σχήμα 5: Μέθοδος μέγιστης καθόδου με προβολή για  $s_k = 5, \gamma_k = 0.5$ .

Παρατηρούμε πως ενώ η μέθοδος ταλαντώνει και δεν συγκλίνει στο ελάχιστο όπως και η αντίστοιχη εκτέλεση της μέγιστης καθόδου χωρίς περιορισμούς με το ίδιο  $\gamma_k$ . Παρόλα αυτά όμως, η ταλάντωση λαμβάνει χώρα εντός του συνόλου των περιορισμών της εκφώνησης.

6. ΘΕΜΑ 3 - ΣΗΜΕΙΟ (-5, 10),  $s_k = 15, \gamma_k = 0.1$

Εφαρμόζοντας τη μέθοδο για ακρίβεια  $\epsilon = 0.01, s_k = 15$  και  $\gamma_k = 0.1$  έχουμε  $\gamma'_k = 1.5 > \frac{1}{3}$ , άρα το κριτήριο σύγκλισης δεν πληρείται και πάλι.



Σχήμα 6: Μέθοδος μέγιστης καθόδου με προβολή για  $s_k = 15, \gamma_k = 0.1$ .

Και εδώ παρατηρούμε πως ενώ το  $\gamma_k$  έχει επιλεγεί θεωρητικά στο εύρος που οδηγεί σε σύγκλιση, το αντίστοιχο βήμα  $s_k$  είναι πολύ μεγάλο, με αποτέλεσμα το γινόμενο τους  $\gamma'_k = 15 * 0.1 = 1.5$  συνολικά να μην πληροί το κριτήριο  $\gamma'_k < \frac{1}{3}$  και η μέθοδος να ταλαντώνει και πάλι. Αυτή τη φορά μόνο στον άξονα  $x_2$ . Αυτό βέβαια εξηγείται από την παραπάνω ανάλυσή καθώς βλέπουμε πως ενώ το γινόμενο 1.5 δεν πληροί τις προϋποθέσεις για σύγκλιση, πληροί όμως τις προϋποθέσεις της εξίσωσης (5) με αποτέλεσμα να ταλαντώνει μόνο στον άξονα  $x_2$ .

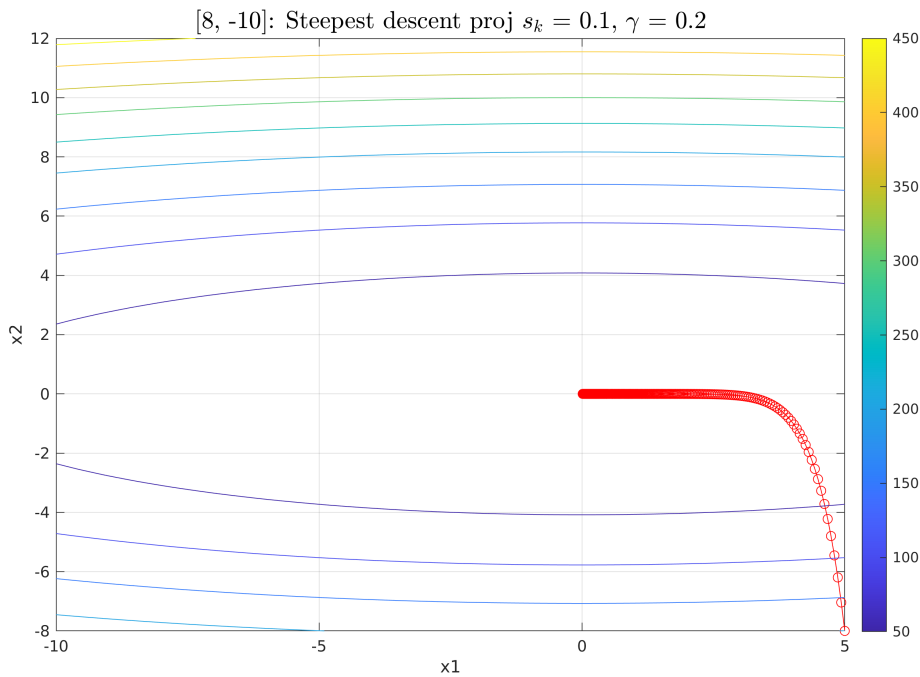
Αυτό φυσικά είναι αληθές και για την προηγούμενη περίπτωση όπου το  $\gamma'_k = 5 * 0.5 = 2.5$  και όπου πάλι η σύγκλιση ήταν μερική, μόνο για την διάσταση  $x_1$ .

7. ΘΕΜΑ 4 - ΣΗΜΕΙΟ (8, -10),  $s_k = 0.1, \gamma_k = 0.2$

Αρχικά παρατηρούμε πως το σημείο **δεν είναι εφικτό**, καθώς είναι εκτός του συνόλου των περιορισμών της εκφώνησης. Αυτό βέβαια δεν μας αποτρέπει από την εφαρμογή της μεθόδου, καθώς αρχικά μπορούμε να προβάλουμε το σημείο στο σύνολο και να εφαρμόσουμε τη μέθοδο έπειτα. Η προβολή του (8, -10) είναι  $Pr_X\{(8, -10)\} = (5, -8)$ , που είναι και το σημείο εκκίνησης του αλγορίθμου όπως φαίνεται και στο σχήμα 7 παρακάτω.

Ακόμα, αυτή τη φορά οι τιμές των βημάτων  $s_k, \gamma_k$ , έχουν επιλεγεί μέσα στο εύρος για το οποίο έχουμε σύγκλιση, καθώς:  $\gamma'_k = 0.1 * 0.2 = 0.02 < \frac{1}{3}$ , επομένως αναμένουμε η μέθοδος να συγκλίνει στο ελάχιστο. Εφαρμόζοντας τη μέθοδο για ακρίβεια  $\epsilon = 0.01$  έχουμε:





**Σχήμα 7:** Μέθοδος μέγιστης καθόδου με προβολή για  $s_k = 0.1, \gamma_k = 0.2$ .

Όπου βλέπουμε πως η μέθοδος συγκλίνει επιβεβαιώνοντας την παραπάνω ανάλυση.

### 8. ΣΥΜΠΕΡΑΣΜΑΤΑ

Η μέθοδος μέγιστης καθόδου, με και χωρίς προβολή, παρουσιάζει τις αναμενόμενες συμπεριφορές σύγκλισης ανάλογα με την επιλογή του βήματος  $\gamma_k$ . Για την περίπτωση χωρίς προβολή, επιβεβαιώθηκε θεωρητικά και εμπειρικά ότι η μέθοδος συγκλίνει μόνο όταν το  $\gamma_k$  βρίσκεται εντός του εύρους που ορίζεται από την ανάλυση της Lipschitz σταθεράς, δηλαδή  $0 < \gamma_k < \frac{1}{3}$ . Επιπλέον, παρατηρήθηκε πως για μεγαλύτερες τιμές  $\gamma_k$  η μέθοδος αποκλίνει, όπως προβλέπεται από την μαθηματική ανάλυση.

Αντίστοιχα, για τη μέθοδο με προβολή, αποδεικνύεται πως το κριτήριο σύγκλισης εξαρτάται από τον συνδυασμό του  $s_k$  (βήμα αντίθετο στην κατεύθυνση της κλίσης) και του  $\gamma_k$  (βήμα κατά την προβολή). Η ανάλυση επιβεβαίωσε ότι η μέθοδος προσεγγίζει τη μέθοδο χωρίς προβολή για εφικτά σημεία, ενώ για μη εφικτά σημεία εξασφαλίζει ότι οι διαδοχικές προσεγγίσεις παραμένουν εντός των περιορισμών.

Τέλος, μέσα από τα παραδείγματα, αναδείχθηκε η σημασία της σωστής επιλογής των παραμέτρων  $\gamma_k$  και  $s_k$  για τη συνολική σταθερότητα και σύγκλιση της μεθόδου, καθώς και ο τρόπος με τον οποίο η ανάλυση των ιδιοτιμών του Εσσιανού επιτρέπει την καλύτερη κατανόηση της συμπεριφοράς της μεθόδου.