Τεχνικές Βελτιστοποίησης — 3ο Παραδοτέο

Μέθοδος Μέγιστης Καθόδου με Προβολή

Χρήστος Αλεξόπουλος-10618

Contents

1 Εισαγωγή

Στόχος της εργασίας είναι η πειραματική μελέτη της Μεθόδου Μέγιστης Καθόδου με Προβολή. Φτιάχνουμε διαγράμματα για τη σύγκλιση των αλγορίθμων αναζήτησης και τον αριθμό επαναλήψεων. Η υλοποίηση των μεθόδων βασίζεται στις παραγράφους (αλγόριθμοι) του βιβλίου

Η αντικειμενική συνάρτηση που θα μελετήσουμε είναι:

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

Παρακάτω παρουσιάζονται δύο γραφικές παραστάσεις της συνάρτησης, ώστε να αποκτήσουμε μία γενική εικόνα της μορφής της.

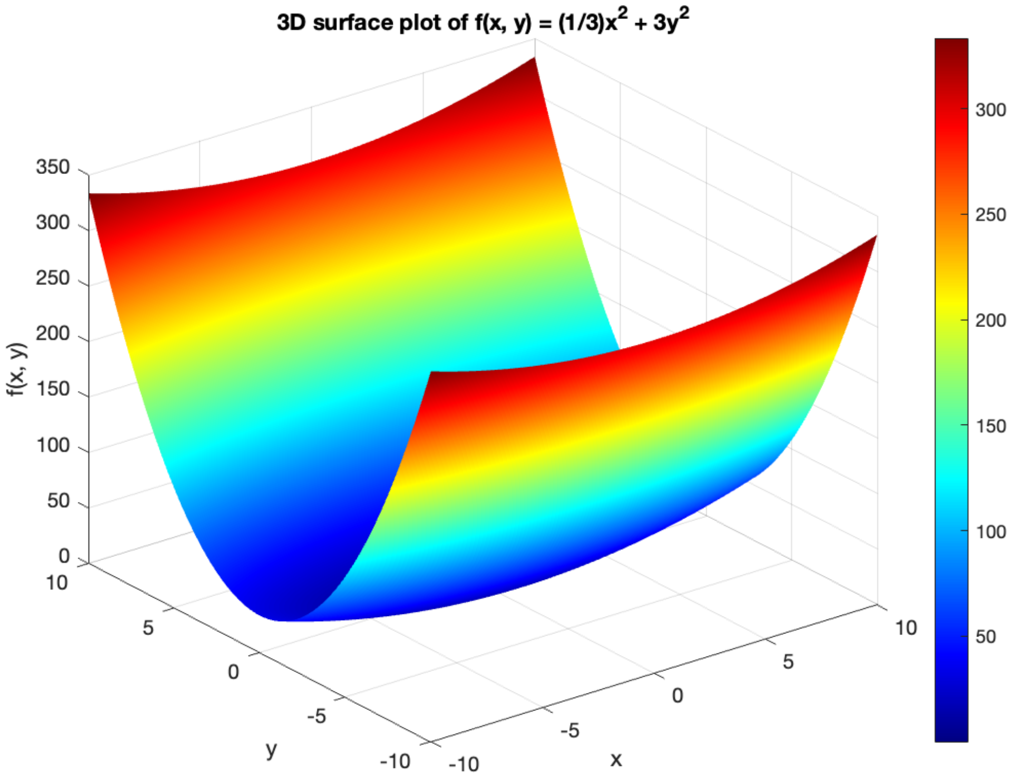


Figure 1: Τρισδιάστατη σχεδίαση της συνάρτησης $f(x,y)$.

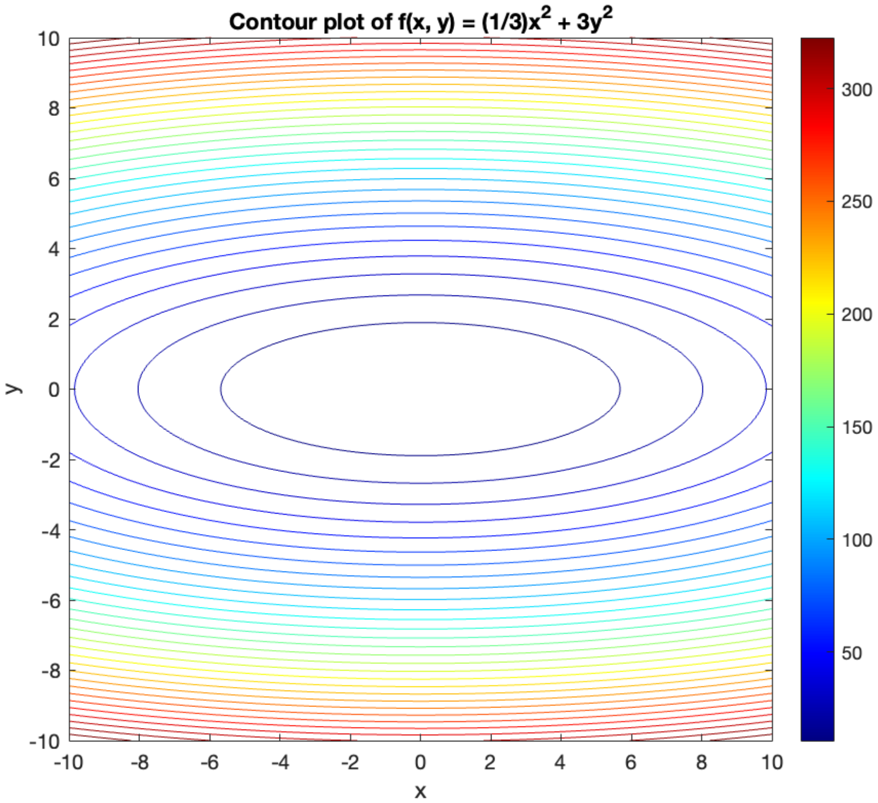


Figure 2: Ισοσταθμικές καμπύλες (contour plot) της συνάρτησης $f(x, y)$.

2 Μέθοδος Μέγιστης Καθόδου

2.1 Μαθηματική Ανάλυση

Όπως και στο 2ο παραδοτέο ακολουθούμε πιστά τον αλγόριθμο 5.2.1 του βιβλίου σελ. 121. Επίσης, το ελάχιστο της συνάρτησης είναι το 0. Η ανάλυση για τον περιορισμό του γ που συγκλίνει είναι:

Θεωρούμε τη συνάρτηση:

$$f(x, y) = \frac{1}{3}x^2 + 3y^2,$$

με gradient

$$\nabla f(x, y) = \left[\frac{2}{3}x,\ 6y\right].$$

Η μέθοδος Μέγιστης Καθόδου δίνει τον επαναληπτικό τύπο:

$$(x_{k+1}, y_{k+1}) = (x_k, y_k) - \gamma \nabla f(x_k, y_k).$$

Άρα:

$$x_{k+1} = \left(1 - \frac{2}{3}\gamma\right) x_k, \qquad y_{k+1} = (1 - 6\gamma)\, y_k.$$

Με επανάληψη για N βήματα:

$$x_{k+N} = \left(1 - \frac{2}{3}\gamma\right)^N x_k, \qquad y_{k+N} = (1 - 6\gamma)^N y_k.$$

Για να συγκλίνει η ακολουθία στο ελάχιστο $(0, 0)$, απαιτείται:

$$\lim_{N\to\infty}(1-\alpha\gamma)^N = 0 \quad \Longleftrightarrow \quad |1-\alpha\gamma| < 1,$$

όπου

$$\alpha_x = \frac{2}{3}, \qquad \alpha_y = 6.$$

Συνθήκη σύγκλισης για τη μεταβλητή x

$$|1 - \tfrac{2}{3}\gamma| < 1$$

Λύνοντας:

$$0 < \gamma < 3.$$

Συνθήκη σύγκλισης για τη μεταβλητή y

$$|1 - 6\gamma| < 1$$

Λύνοντας:

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

Τελικό συμπέρασμα

Για να συγκλίνουν *και οι δύο* μεταβλητές στο ελάχιστο σημείο, πρέπει να ισχύουν και οι δύο συνθήκες. Άρα η συνολική περιοχή σύγκλισης της μεθόδου είναι:

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

Η θεωρητική αυτή περιοχή σύγκλισης επιβεβαιώνεται πλήρως από τα πειραματικά αποτελέσματα:

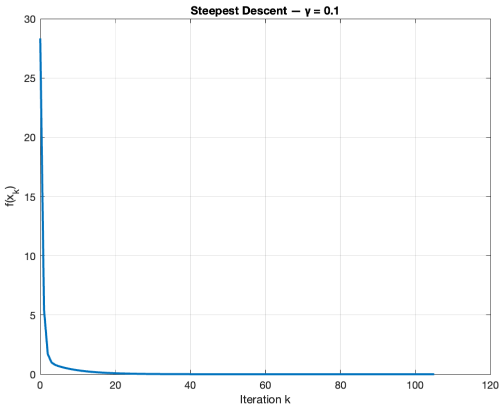
- Για $\gamma = 0.1$ και $\gamma = 0.3$: η μέθοδος συγκλίνει.
- Για $\gamma = 3$ και $\gamma = 5$: η μέθοδος αποκλίνει.

Η συμπεριφορά αυτή συμφωνεί απόλυτα με τη θεωρία για τις τετραγωνικές συναρτήσεις και τους όρους σύγκλισης της μεθόδου Steepest Descent.

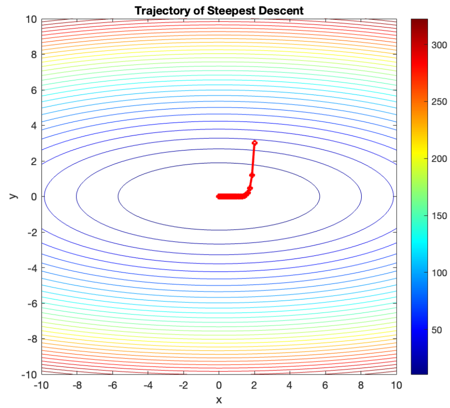
2.2 Αποτελέσματα Προσομοίωσης

Για το σημείο εκκίνησης $(2, 3)$ και για ακρίβεια $\epsilon = 0.001$ έχουμε τα εξής αποτελέσματα για κάθε επιλογή βήματος.

2.2.1 Μέθοδος Steepest Descent με βήμα $\gamma_k = 0.1$

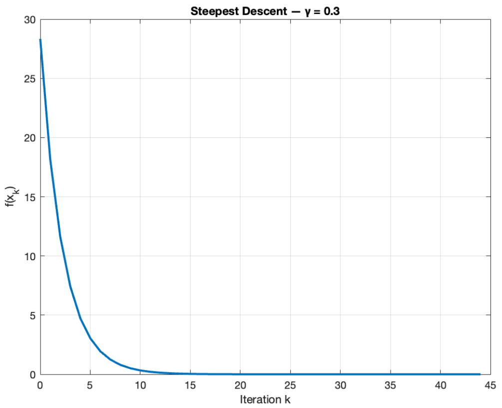


(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.

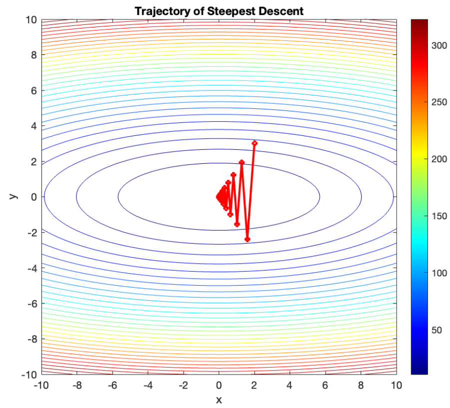


(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

2.2.2 Μέθοδος Steepest Descent με βήμα $\gamma_k = 0.3$

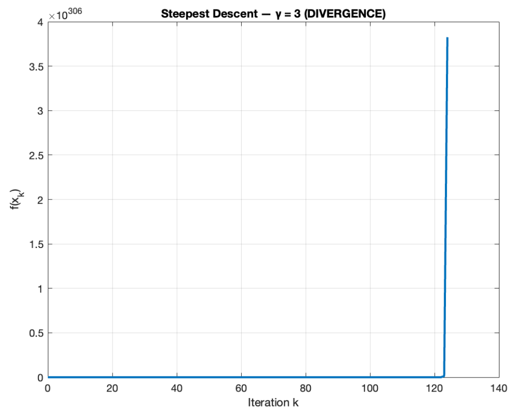


(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.



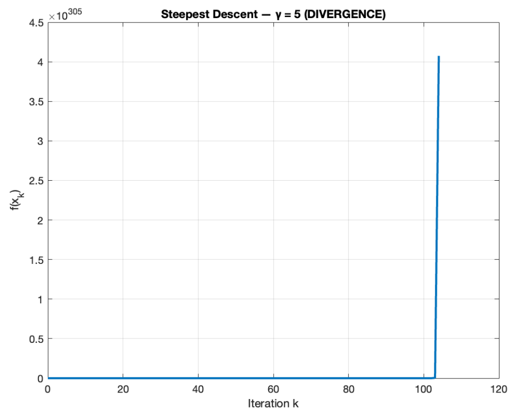
(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

2.2.3 Μέθοδος Steepest Descent με βήμα $\gamma_k = 3$



(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.

2.2.4 Μέθοδος Steepest Descent με βήμα $\gamma_k = 5$



(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.

2.3 Παρατηρήσεις

Όπως είδαμε και πιο πάνω η μέθοδος για γ=3 και γ=5 αποκλίνει(αναμενόμενο απο την μαθηματική ανάλυση), μάλιστα δεν μπορούμε να δούμε την τροχιά επαναλήψεων πάνω στο contour plot αφού οι τιμές γίνονται τεράστιες(της τάξης 10^{70}). Από άποψη ταχύτητας η μέθοδος είναι πιο γρήγορη με $\gamma = 0.3$, αναμενόμενο αφού οι τιμές αλλάζουν περισσότερο σε κάθε επανάληψη μέχρι να φτάσουμε στο ελάχιστο και για αυτό βλέπουμε και μεγαλύτερες ταλαντώσεις(λιγότερο ομαλή πορεία σε σχέση με το $\gamma = 0.1$).

3 Μέθοδος Μέγιστης Καθόδου με Προβολή

Η λύση που αναζητούμε δεν μπορεί να κινείται ελεύθερα στο $\mathbb{R}^2$, αλλά πρέπει να ικανοποιεί περιορισμούς της μορφής:

$$X = \{(x, y) \in \mathbb{R}^2 : a_x \leq x \leq b_x,\ a_y \leq y \leq b_y\}.$$

Οι περιορισμοί εδώ είναι:

$$-10 \leq x \leq 5, \qquad -8 \leq y \leq 12.$$

Ορισμός της προβολής σε κουτί (box projection)

Η προβολή ενός σημείου (u, v) στο σύνολο X δίνεται συντεταγμένα ως:

$$\mathrm{Proj}_X(u, v) = \left(\mathrm{Proj}_X(u),\ \mathrm{Proj}_X(v)\right),$$

όπου

$$\mathrm{Proj}_X(z) = \begin{cases} a, & z < a, \\ b, & z > b, \\ z, & \text{αλλιώς}. \end{cases}$$

Ισοδύναμα, μπορεί να γραφτεί με τη χρήση των τελεστών min και max:

$$\mathrm{Proj}_X(z) = \min\bigl(\max(z, a),\ b\bigr),$$

που αποτελεί τη μορφή που χρησιμοποιήθηκε στην υλοποίηση.

Για τις δύο μεταβλητές:

$$x_{\text{proj}} = \min\bigl(\max(x, -10),\ 5\bigr), \qquad y_{\text{proj}} = \min\bigl(\max(y, -8),\ 12\bigr).$$

Αλγόριθμος Μέγιστης Καθόδου με Προβολή

Η μέθοδος ξεκινά από ένα σημείο $x_k \in X$ και υπολογίζει το επόμενο βήμα ως:

$$y_k = x_k - s_k \nabla f(x_k),$$

όπου $s_k > 0$ είναι η μετατόπιση.

Το σημείο y_k μπορεί να βρίσκεται εκτός των περιορισμών, άρα εφαρμόζεται προβολή:

$$x'_k = \mathrm{Proj}_X(y_k).$$

Στη συνέχεια, το νέο σημείο υπολογίζεται με μεταβολή της μορφής:

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

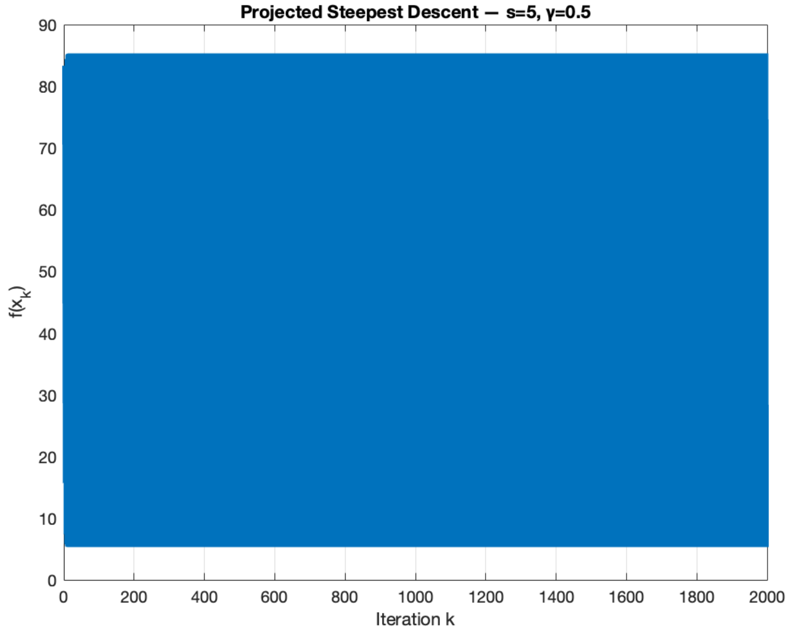
όπου $0 < \gamma_k \leq 1$(ελέγχει πόσο γρήγορα κινούμαστε προς το προβεβλημένο σημείο).

Συνοψίζοντας η διαδικασία που ακολουθούμε είναι η εξής:

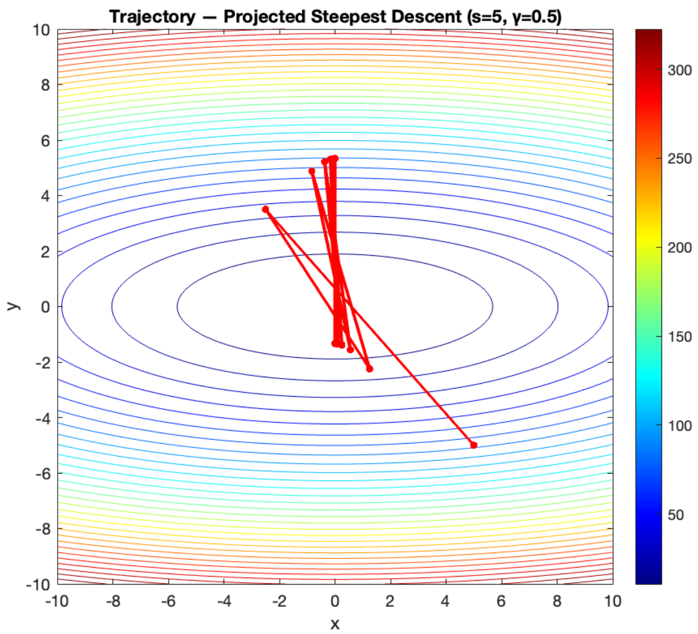
$$\boxed{\begin{aligned} y_k &= x_k - s_k \nabla f(x_k), \\ x'_k &= \mathrm{Proj}_X(y_k), \\ x_{k+1} &= (1 - \gamma_k) x_k + \gamma_k x'_k. \end{aligned}}$$

4 Μέθοδος Μέγιστης Καθόδου με Προβολή: $s_k = 5$, $\gamma_k = 0.5$, σημείο εκκίνησης $(5, -5)$, ακρίβεια $\epsilon = 0.01$

4.1 Αποτελέσματα Προσομοίωσης



(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.



(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

4.2 Παρατηρήσεις

Στο Θέμα 2 εφαρμόζεται η Μέθοδος Μέγιστης Καθόδου με Προβολή, με βήμα $s_k = 5$ και $\gamma_k = 0.5$. Το αρχικό σημείο βρίσκεται στην άκρη του ορθογωνίου που ορίζουν οι περιορισμοί, γεγονός που σε συνδυασμό με το μεγάλο s_k έχει ως αποτέλεσμα το ενδιάμεσο σημείο

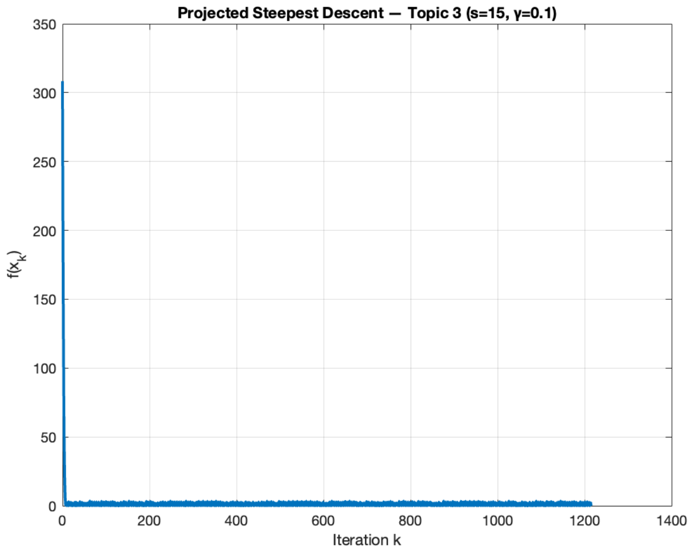
$$y_k = x_k - s_k \nabla f(x_k)$$

να βρίσκεται σχεδόν πάντα εκτός του επιτρεπτού συνόλου. Έτσι η προβολή ενεργοποιείται σε κάθε επανάληψη και η μέθοδος δεν ακολουθεί την πραγματική κατεύθυνση καθόδου.

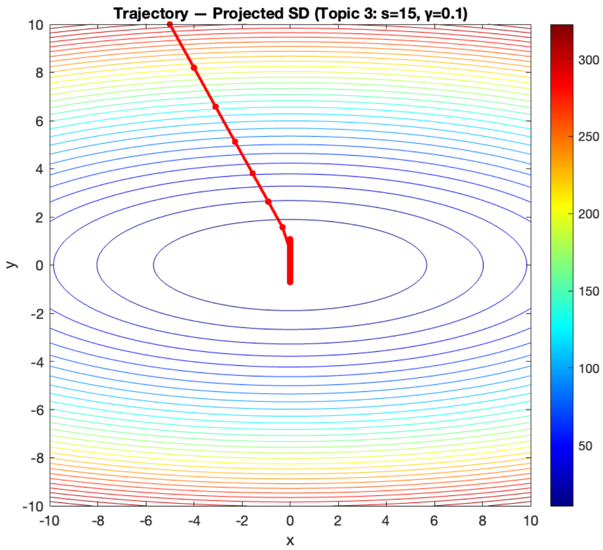
Σε σχέση με το Θέμα 1, όπου οι τιμές για $\gamma < 1/3$ οδηγούν σε σύγκλιση, εδώ η επιλογή $\gamma_k = 0.5$ (μη αποδεκτή θεωρητικά) σε συνδυασμό με το μεγάλο s_k προκαλεί πλήρη αστάθεια: η τιμή της $f(x_k)$ δεν μειώνεται και η τροχιά των επαναλήψεων παραμένει κολλημένη στα όρια των περιορισμών.

5 Μέθοδος Μέγιστης Καθόδου με Προβολή: $s_k = 15$, $\gamma_k = 0.1$, σημείο εκκίνησης $(-5, 10)$, ακρίβεια $\epsilon = 0.01$

5.1 Αποτελέσματα Προσομοιώσεων



(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.



(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

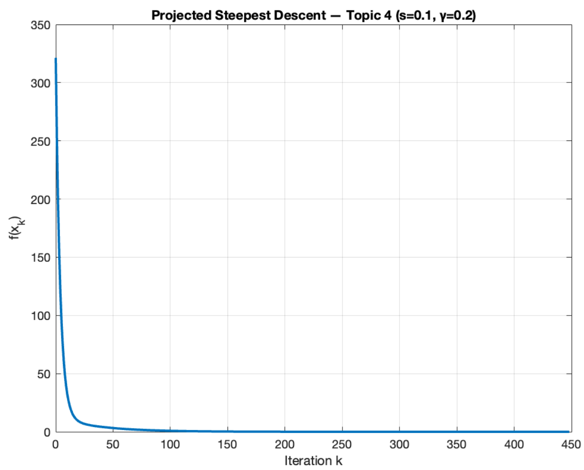
5.2 Παρατηρήσεις

Σε σχέση με τα Θέματα 1 και 2 παρατηρούμε πολύ καλύτερα αποτελέσματα· το σύστημα φαίνεται να συγκλίνει στο ελάχιστο (μάλλον αναμενόμενο αφού $\gamma < 0.1$), αν και δεν σταθεροποιείται πλήρως αλλά εμφανίζει μικρές ταλαντώσεις γύρω από αυτό. Αυτό οφείλεται στο μεγάλο s_k, το οποίο προκαλεί υπερπήδηση του ελαχίστου, ενώ η προβολή εμποδίζει την απόκλιση και διατηρεί τις τιμές κοντά στο ελάχιστο.

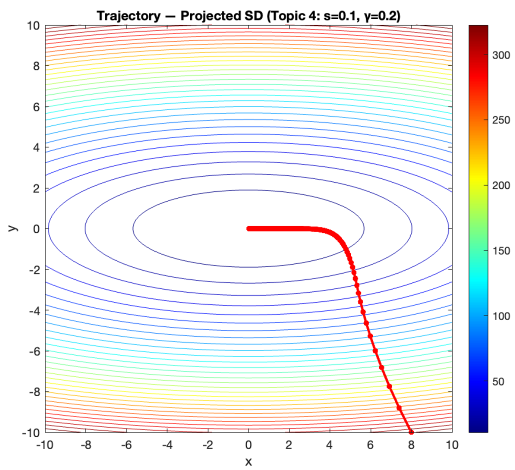
Ένας απλός τρόπος ώστε η μέθοδος να συγκλίνει και να σταθεροποιείται είναι να επιλεγεί μικρότερο s_k (όπως στο Θέμα 4). Αν όμως κρατήσουμε το ίδιο s_k, μπορούμε να χρησιμοποιήσουμε έναν προσαρμοστικό κανόνα επιλογής βήματος, όπως **Exact Line Search** ή **Armijo Backtracking**, οι οποίοι μειώνουν αυτόματα το αποτελεσματικό βήμα $\gamma_k s_k$ και αποτρέπουν τις ταλαντώσεις.

6 Μέθοδος Μέγιστης Καθόδου με Προβολή: $s_k = 0.1$, $\gamma_k = 0.2$, σημείο εκκίνησης $(8, -10)$, ακρίβεια $\epsilon = 0.01$

6.1 Αποτελέσματα Προσομοιώσεων



(a) Σύγκλιση της τιμής $f(x_k)$ ως προς τις επαναλήψεις.



(b) Τροχιά των επαναλήψεων πάνω στο contour plot.

6.2 Παρατηρήσεις

Σε αυτή την περίπτωση το πραγματικό βήμα της μεθόδου είναι

$$\gamma_k' = \gamma_k s_k = 0.02,$$

το οποίο είναι πολύ μικρότερο από το όριο σύγκλισης

$$\gamma_k' < \frac{1}{3}.$$

Άρα εκ των προτέρων γνωρίζουμε ότι η μέθοδος θα συγκλίνει.

Επιπλέον, επειδή το βήμα είναι μικρό, το

$$x_k - s_k \nabla f(x_k)$$

παραμένει σχεδόν πάντα μέσα στα όρια. Άρα η προβολή δεν ενεργοποιείται. Η μέθοδος λειτουργεί σχεδόν σαν απλή μέθοδος μέγιστης καθόδου χωρίς περιορισμούς.

- Η απενεργοποίηση της προβολής λόγω μικρού s_k σημαίνει ότι το βήμα δεν «χτυπά» τα όρια και δεν προκαλεί ταλαντώσεις (όπως συνέβη στο Θέμα 3).
- Αφού ισχύει $\gamma_k' = 0.02 < \frac{1}{3}$, η σύγκλιση είναι εγγυημένη και μάλιστα σταθερή και γρήγορη.
- Σε αντίθεση με τα προηγούμενα θέματα, εδώ η τροχιά είναι ομαλή και οδηγεί απευθείας προς το $(0, 0)$.

7 Αρχιτεκτονική και Οργάνωση Κώδικα

Η δομή του κώδικα του Project 3 έχει οργανωθεί σε σαφώς διαχωρισμένους φακέλους, ώστε οι μαθηματικές συναρτήσεις, οι μέθοδοι, οι κανόνες βήματος και τα γραφήματα να είναι πλήρως modular και επεκτάσιμα.

```
PROJECT_3/

 functions/
    f_xy.m
    grad_f_xy.m
    proj_box.m

 methods/
    Steepest_descent.m
    Steepest_descent_projected.m

 steps/
    step_fixed.m
    step_projected.m
```

```
plots/
    plot_f.m
    plot_contours.m
    plot_convergence.m
    plot_trajectory.m

main.m
```

Περιγραφή Φακέλων

functions/ Περιέχει τις βασικές μαθηματικές συναρτήσεις: την αντικειμενική $f(x, y)$, το διάνυσμα βαθμίδας και τη συνάρτηση προβολής στο κουτί περιορισμών $[-10, 5] \times [-8, 12]$.

methods/ Περιλαμβάνει τις υλοποιήσεις της μεθόδου Μέγιστης Καθόδου και της Μέγιστης Καθόδου με Προβολή. Κάθε μέθοδος υλοποιείται σε ανεξάρτητο αρχείο για καθαρότητα και εύκολο debugging.

steps/ Περιέχει τους διαφορετικούς κανόνες επιλογής μήκους βήματος: σταθερό βήμα και βήμα προβολής (όπου πρώτα γίνεται κίνηση κατά $-s_k \nabla f(x_k)$ και στη συνέχεια προβολή στο σύνολο περιορισμών). Οι μέθοδοι δέχονται αυτά τα step rules ως function handles για μέγιστη ευελιξία.

plots/ Συγκεντρώνει όλες τις συναρτήσεις για παραγωγή γραφημάτων: επίπεδων καμπύλων, 3D απεικόνισης, γραφημάτων σύγκλισης και τροχιάς επαναλήψεων.

main.m Το κεντρικό πρόγραμμα εκτέλεσης. Ορίζει παραμέτρους, καλεί τις μεθόδους, παράγει τα γραφήματα για κάθε θέμα και οργανώνει το συνολικό πείραμα της εργασίας.