Μέθοδος μέγιστης καθόδου με προβολή

Τεχνικές Βελτιστοποίησης

3η Εργαστηριακή Άσκηση

20 Δεκέμβρη 2021

Περιληπτικά:
Στην εργασία αυτή ζητούμενο είναι η ελαχιστοποίηση μιας δοσμένης συνάρτησης δύο μεταβλητών με τη μέθοδο της Μέγιστης Καθόδου με προβολή. Θα δοκιμαστούν διάφορες επιλογές παραμέτρων και τα αποτελέσματά τους θα συγκριθούν με αυτά που περιμενουμένουμε από τη θεωρητική ανάλυση. Θα συγκριθεί η μέθοδος της Μέγιστης Καθόδου με προβολή με αυτή χωρίς προβολή και θα αναδειχθούν ποια προβλήματα προκύπτουν από την κάθε μία.

Σχεδίαση της f

Αρχικά, είναι χρήσιμο να σχεδιάσουμε τη γραφική παράσταση της δοσμένης συνάρτησης $f(x_1, x_2)$ ή $f(x, y)$ προκειμένου να έχουμε μία εποπτεία του προβλήματος της ελαχιστοποίησής της, και να ξέρουμε τι περίπου να περιμένουμε. Έτσι, στα παρακάτω διαγράμματα παρατηρούμε πως η συνάρτηση έχει ολικό ελάχιστο στο $(0, 0)$, με τιμή ίση με 0. Η περιοχή κοντά στο $(0, 0)$, περιμένουμε να είναι και η κύρια περιοχή αναζήτησής μας, πέραν από τους περιορισμούς που μας θέτει η Μέθοδο Μέγιστης Καθόδου με Προβολή.

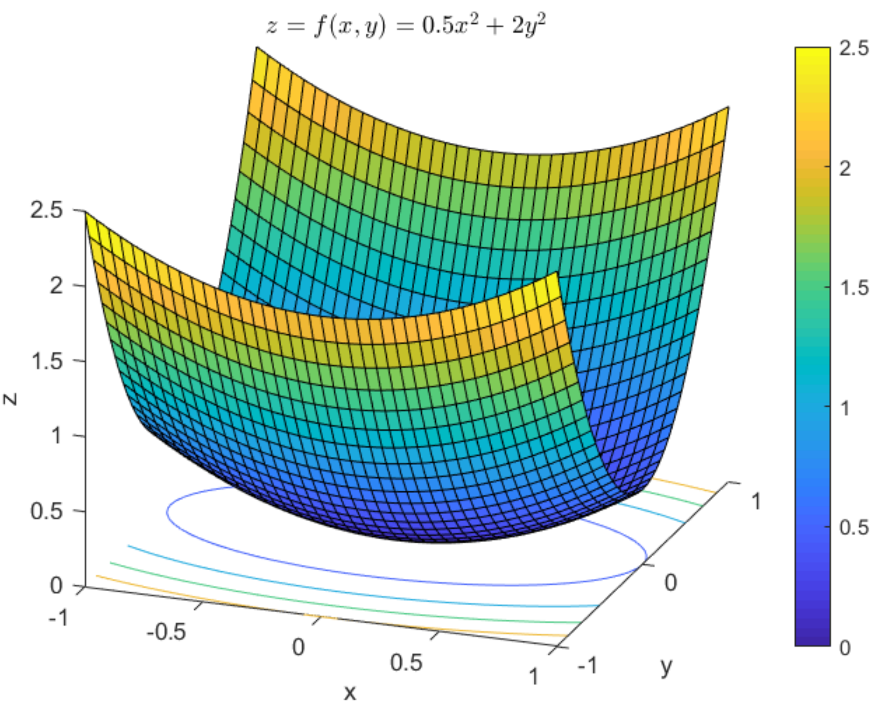
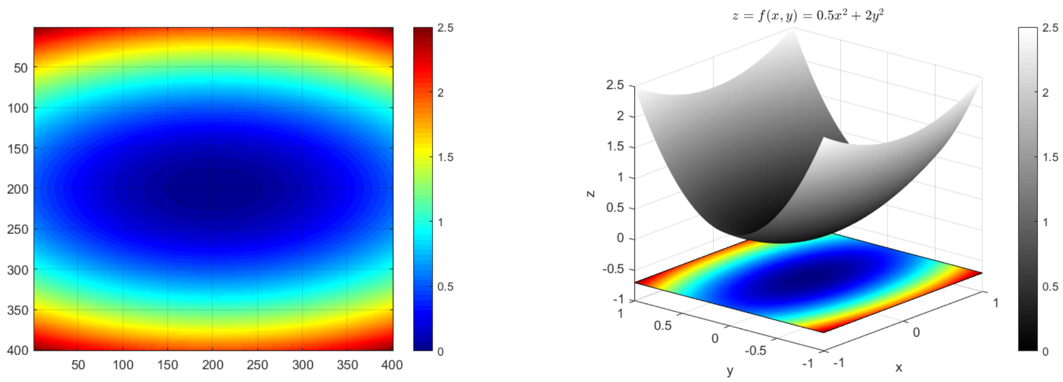


Figure 1: Γραφική παράσταση της $f(x_1, x_2)$



(a) (b)

Figure 2: Επιπλέον γραφικές παραστάσεις

Για κάθε μέθοδο, για να τρέξει, πρέπει να κάνετε uncomment το αντίστοιχο σημείο στο `mainScript.m`, τα αντίστοιχα σχόλια θα σας καθοδηγήσουν.

1. Μέθοδος Μέγιστης Καθόδου - Steepest Descent

Παρακάτω, θα χρησιμοποιήσουμε τη μέθοδο της μέγιστης καθόδου (από την προηγούμενη εργασία) με ακρίβεια $\epsilon = 0.01$ και διαφορετικά βήματα γ, η σημασία των οποίων θα αναλυθεί στην πορεία.

1.1. Επιλογές γ

Αρχικά, επιλέγουμε $\gamma = 0.05$, και παρατηρούμε τα εξής αποτελέσματα ξεκινώντας από το σημείο $(-1, 1)$:

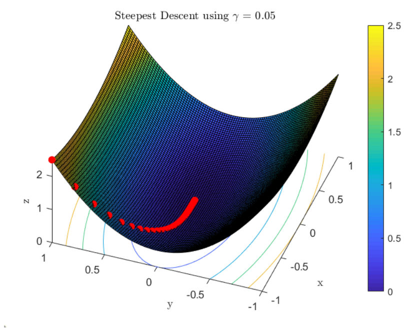


Figure 3: Διάγραμμα επιφάνειας για εκκίνηση από το (-1,1) και γ = 0.05

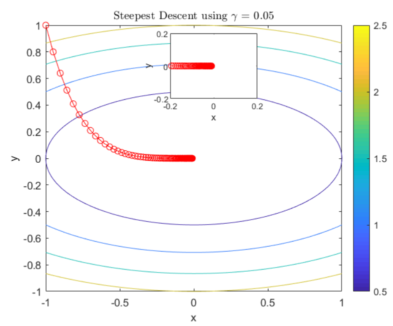


Figure 4: Ισοβαρείς καμπύλες για εκκίνηση από το (-1,1) και γ = 0.05

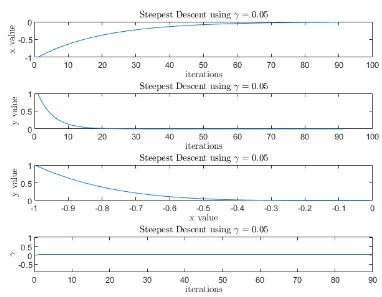


Figure 5: Αλλαγές στις παραμέτρους για εκκίνηση από το (-1,1) και γ = 0.05

Επιλέγοντας $\gamma = 0.5$, παρατηρούμε τα εξής αποτελέσματα ξεκινώντας από το σημείο $(-1, 1)$:
Αν θέσουμε μέγιστο αριθμό επαναλήψεων ίσο με 10, μπορούμε να δούμε πιο ξεκάθαρα τις μεταβολές

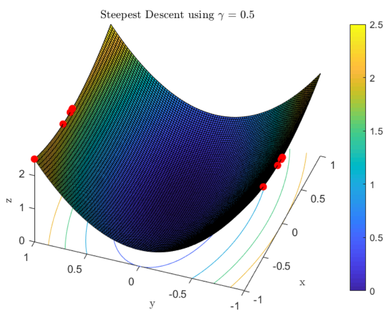


Figure 6: Διάγραμμα επιφάνειας για εκκίνηση από το (-1,1) και γ = 0.5

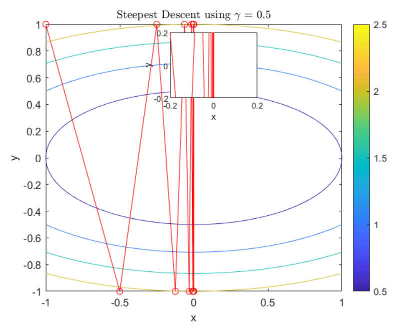


Figure 7: Ισοβαρείς καμπύλες για εκκίνηση από το (-1,1) και γ = 0.5

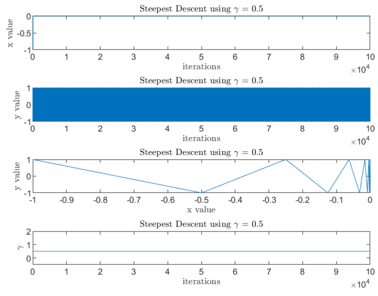


Figure 8: Αλλαγές στις παραμέτρους για εκκίνηση από το (-1,1) και γ = 0.5

στις δύο μεταβλητές της συνάρτησης.

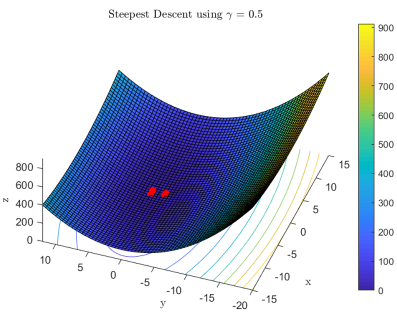


Figure 9: Διάγραμμα επιφάνειας για εκκίνηση από το (-1,1) και γ = 0.5

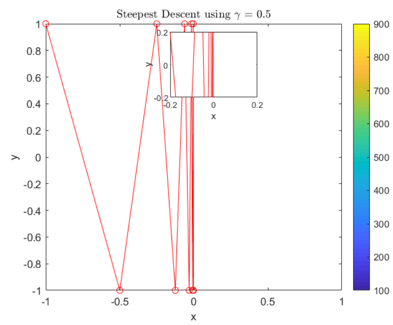


Figure 10: Ισοβαρείς καμπύλες για εκκίνηση από το (-1,1) και γ = 0.5

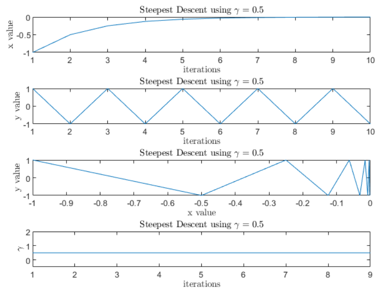


Figure 11: Αλλαγές στις παραμέτρους για εκκίνηση από το (-1,1) και γ = 0.5

Επιλέγοντας $\gamma = 2$, παρατηρούμε τα εξής αποτελέσματα ξεκινώντας από το σημείο $(-1, 1)$ και θέτοντας μέγιστο όριο τις 10 επαναλήψεις:

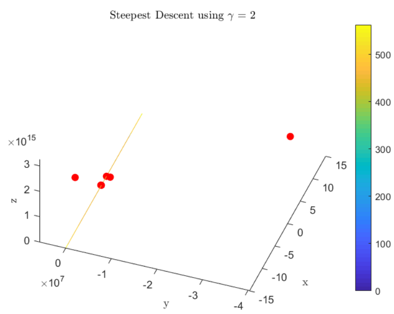


Figure 12: Διάγραμμα επιφάνειας για εκκίνηση από το (-1,1) και γ = 2

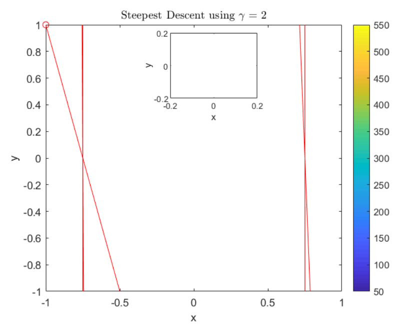


Figure 13: Ισοβαρείς καμπύλες για εκκίνηση από το (-1,1) και γ = 2

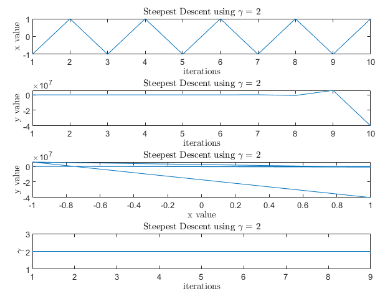


Figure 14: Αλλαγές στις παραμέτρους για εκκίνηση από το (-1,1) και γ = 2

Επιλέγοντας $\gamma = 10$, παρατηρούμε τα εξής αποτελέσματα ξεκινώντας από το σημείο $(-1, 1)$ και θέτοντας μέγιστο όριο τις 10 επαναλήψεις:

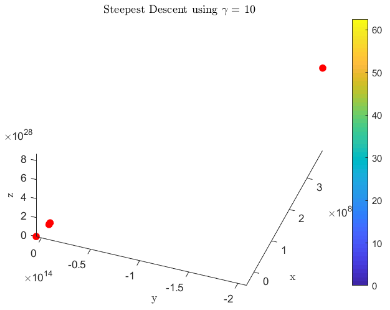


Figure 15: Διάγραμμα επιφάνειας για εκκίνηση από το (-1,1) και γ = 10

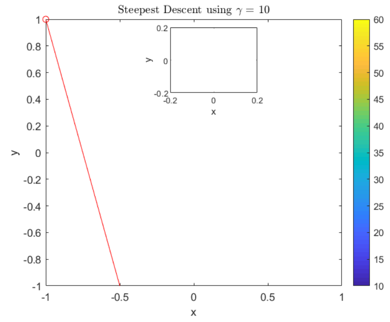


Figure 16: Ισοβαρείς καμπύλες για εκκίνηση από το (-1,1) και γ = 10

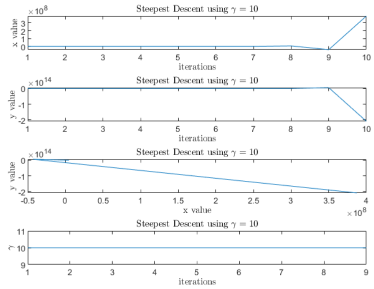


Figure 17: Αλλαγές στις παραμέτρους για εκκίνηση από το (-1,1) και γ = 10

1.2. Παρατηρήσεις και Συμπεράσματα

Με τη χρήση της μεθόδου της μέγιστης καθόδου με σταθερό βήμα γ_k (υλοποίηση της προηγούμενης εργασίας) αναζητούμε το ελάχιστο της συνάρτησης με αρχικό σημείο το (-1,1). Στα παραπάνω σχήματα βλέπουμε τα αποτελέσματα για $\gamma = 0.05$, $\gamma = 0.5$, $\gamma = 2$, $\gamma = 10$.

Παρατηρούμε ότι:

1. Για $\gamma = 0.05$ ο αλγόριθμος συγκλίνει επιτυχώς στο ελάχιστο σε 90 βήματα.
2. Για $\gamma = 0.5$ ο αλγόριθμος δε συγκλίνει ποτέ στο ελάχιστο και φτάνει το ανώτερο όριο που έχουμε ορίσει ίσο με 100000 επαναλήψεις. Ο αλγόριθμος, ενώ για το x_1 καταφέρνει και βρίσκει σε λίγα μόνο βήματα το 0, το x_2 ταλαντώνεται μεταξύ του -1 και του 1 (τα δύο αυτά σημεία ταλάντωσης εξαρτώνται από το σημείο εκκίνησης, και αν το αλλάξουμε, θα αλλάξουν και τα δύο αυτά σημεία).
3. Για $\gamma = 2$ ο αλγόριθμος επίσης δε συγκλίνει. Το x_1 ταλαντώνεται μεταξύ των 1 και -1, ενώ το x_2 αυξάνεται εκθετικά παίρνοντας εναλλάξ θετική και αρνητική τιμή. Για τα διαγράμματα έχουμε θέσει μέγιστο αριθμό επαναλήψεων ίσο με 10, που είναι αρκετές για να παρατηρήσουμε τα προβλήματα που προκύπτουν από τον αλγόριθμο.
4. Για $\gamma = 10$ ο αλγόριθμος απομακρύνεται ταχύτατα από το αρχικό του σημείο. Η ακολουθία x_k παίρνει τιμές $(-1, 1) \to (9, -39) \to (-81, 1521) \to (729, -59319) \to ... \to (\infty, -\infty)$ (όπου ∞ θεωρούμε τις μέγιστες τιμές που μπορούν να χρησιμοποιηθούν στο MATLAB).

Τα αποτελέσματα μπορούν να εξηγηθούν με βάση τον τύπο για την ακολουθία της μέγιστης καθόδου:

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

όπου αν αντικαταστήσουμε την τιμή του $\nabla f(x_k) = \begin{bmatrix} x_{1,k} \\ 4x_{2,k} \end{bmatrix}$ και θεωρήσουμε σταθερό $\gamma_k = \gamma$ έχουμε:

$$x_{k+1} = \begin{bmatrix} x_{1,k+1} \\ x_{2,k+1} \end{bmatrix} = \begin{bmatrix} x_{1,k} \\ x_{2,k} \end{bmatrix} - \gamma \begin{bmatrix} x_{1,k} \\ 4x_{2,k} \end{bmatrix} = \begin{bmatrix} (1-\gamma)x_{1,k} \\ (1-4\gamma)x_{2,k} \end{bmatrix} \tag{2}$$

Έτσι, αντικαθιστώντας για τις διάφορες τιμές του γ έχουμε:

1. $\gamma = 0.05 \Rightarrow x_{k+1} = \begin{bmatrix} 0.95x_{1,k} \\ 0.8x_{2,k} \end{bmatrix}$
   Οπότε, αν έχουμε ένα αρχικό σημείο x_1 η ακολουθία x_k έχει ως εξής:
   - $x_{1,2} = 0.95 \cdot x_{1,1}$
   - $x_{1,3} = 0.95 \cdot x_{1,2} = 0.95^2 \cdot x_{1,1}$
   - $x_{1,4} = 0.95 \cdot x_{1,3} = 0.95^3 \cdot x_{1,1}$
   - ...

- $x_{1,k} = 0.95^{k-1} \cdot x_{1,1}$

και αντίστοιχα,

- $x_{2,2} = 0.8 \cdot x_{2,1}$
- $x_{2,3} = 0.8 \cdot x_{2,2} = 0.8^2 \cdot x_{2,1}$
- $x_{2,4} = 0.8 \cdot x_{2,3} = 0.8^3 \cdot x_{2,1}$
- $\ldots$
- $x_{2,k} = 0.8^{k-1} \cdot x_{2,1}$

Παρατηρούμε ότι $|x_{k+1}| < |x_k|\ \forall k \in \mathbb{N}$ και ότι η ακολουθία συγκλίνει προς το (0,0) για $k \to +\infty$.

2. $\gamma = 0.5 \Rightarrow x_{k+1} = \begin{bmatrix} 0.95x_{1,k} \\ 0.8x_{2,k} \end{bmatrix}$

Οπότε, αν έχουμε ένα αρχικό σημείο x_1 η ακολουθία x_k έχει ως εξής:

- $x_{1,2} = 0.5 \cdot x_{1,1}$
- $x_{1,3} = 0.5 \cdot x_{1,2} = 0.5^2 \cdot x_{1,1}$
- $x_{1,4} = 0.5 \cdot x_{1,3} = 0.5^3 \cdot x_{1,1}$
- $\ldots$
- $x_{1,k} = 0.5^{k-1} \cdot x_{1,1}$

και αντίστοιχα,

- $x_{2,2} = -1 \cdot x_{2,1}$
- $x_{2,3} = -1 \cdot x_{2,2} = -1^2 \cdot x_{2,1}$
- $x_{2,4} = -1 \cdot x_{2,3} = -1^3 \cdot x_{2,1}$
- $\ldots$
- $x_{2,k} = -1^{k-1} \cdot x_{2,1}$

Παρατηρούμε ότι $|x_{1,k+1}| < |x_{1,k}|\ \forall k \in \mathbb{N}$ και ότι η ακολουθία συγκλίνει προς το 0 για $k \to +\infty$. Ωστόσο, $|x_{2,k+1}| = |x_{2,k}|\ \forall\ k \in \mathbb{N}$, και έτσι παρατηρούμε ταλάντωση μεταξύ δύο σημείων.

3. $\gamma = 2 \Rightarrow x_{k+1} = \begin{bmatrix} 0.95x_{1,k} \\ 0.8x_{2,k} \end{bmatrix}$

Οπότε, αν έχουμε ένα αρχικό σημείο x_1 η ακολουθία x_k έχει ως εξής:

- $x_{1,2} = -1 \cdot x_{1,1}$
- $x_{1,3} = -1 \cdot x_{1,2} = -1^2 \cdot x_{1,1}$
- $x_{1,4} = -1 \cdot x_{1,3} = -1^3 \cdot x_{1,1}$
- $\ldots$
- $x_{1,k} = -1^{k-1} \cdot x_{1,1}$

και αντίστοιχα,

- $x_{2,2} = -7 \cdot x_{2,1}$
- $x_{2,3} = -7 \cdot x_{2,2} = -7^2 \cdot x_{2,1}$
- $x_{2,4} = -7 \cdot x_{2,3} = -7^3 \cdot x_{2,1}$
- $\ldots$
- $x_{2,k} = -7^{k-1} \cdot x_{2,1}$

Παρατηρούμε ότι $|x_{1,k+1}| = |x_{1,k}|\ \forall k \in \mathbb{N}$, και έτσι παρατηρούμε ταλάντωση μεταξύ δύο σημείων. Έπειτα, $|x_{k+1}| > |x_k|\ \forall k \in \mathbb{N}$, άρα αποκλίνει από το 0 για $k \to +\infty$.

4. $\gamma = 10 \Rightarrow x_{k+1} = \begin{bmatrix} -9x_{1,k} \\ -39x_{2,k} \end{bmatrix}$

Οπότε, αν έχουμε ένα αρχικό σημείο x_1 η ακολουθία x_k έχει ως εξής:

- $x_{1,2} = -9 \cdot x_{1,1}$
- $x_{1,3} = -9 \cdot x_{1,2} = -9^2 \cdot x_{1,1}$
- $x_{1,4} = -9 \cdot x_{1,3} = -9^3 \cdot x_{1,1}$
- $\ldots$
- $x_{1,k} = -9^{k-1} \cdot x_{1,1}$

και αντίστοιχα,

- $x_{2,2} = -39 \cdot x_{2,1}$
- $x_{2,3} = -39 \cdot x_{2,2} = -39^2 \cdot x_{2,1}$
- $x_{2,4} = -39 \cdot x_{2,3} = -39^3 \cdot x_{2,1}$
- $\ldots$
- $x_{2,k} = -39^{k-1} \cdot x_{2,1}$

Παρατηρούμε ότι $|x_{k+1}| > |x_k|$ $\forall k \in \mathbb{N}$ και ότι η ακολουθία αποκλίνει από το (0,0) για $k \to +\infty$.
Αντικαθιστώντας το $x_1 = (-1, 1)$:

- $x_2 = (9, -39)$
- $x_3 = (-81, 1521)$
- $x_4 = (729, -59319)$
- ...

Έτσι, εξηγείται η συμπεριφορά του αλγορίθμου και για αυτή την τιμή του γ.

Γενικά, αφού ισχύει

$$x_{1,k} = (1-\gamma)^{k-1} \cdot x_{1,1}$$

και θέλουμε $x_{1,k} \to 0$ όταν $k \to \infty$ πρέπει:

$$\lim_{k\to\infty} x_{1,k} = 0 \Rightarrow$$
$$\lim_{k\to\infty} (1-\gamma)^{k-1} \cdot x_{1,1} = 0, \text{ και } x_{1,1} \neq (0,0)$$
$$\lim_{k\to\infty} (1-\gamma)^{k} = 0 \Rightarrow$$
$$|1-\gamma| < 1 \Rightarrow$$
$$0 < \gamma < 2$$

για να συγκλίνει ο αλγόριθμος για x_1
και αντίστοιχα,

$$x_{2,k} = (1-4\gamma)^{k-1} \cdot x_{2,1}$$

και θέλουμε $x_{1,k} \to 0$ όταν $k \to \infty$ πρέπει:

$$\lim_{k\to\infty} x_{2,k} = 0 \Rightarrow$$
$$\lim_{k\to\infty} (1-4\gamma)^{k-1} \cdot x_{2,1} = 0, \text{ και } x_{2,1} \neq (0,0)$$
$$\lim_{k\to\infty} (1-4\gamma)^{k} = 0 \Rightarrow$$
$$|1-4\gamma| < 1 \Rightarrow$$
$$0 < \gamma < 0.5$$

για να συγκλίνει ο αλγόριθμος για x_2.

Έτσι, για να συγκλίνει η f στο ελάχιστο **θα πρέπει**

$$0 < \gamma < 0.5$$

2. Μέθοδος μέγιστης καθόδου με προβολή

2.1. Θεωρητική ανάλυση

Θεωρούμε τους περιορισμούς

$$-15 \le x_1 \le 15$$

και

$$-20 \le x_2 \le 12$$

και θα χρησιμοποιήσουμε τον αλγόριθμο της μέγιστης καθόδου με προβολή για να βρούμε το ελάχιστο.

Η ακολουθία των σημείων x_k αυτής της μεθόδου δίνεται από την σχέση

$$x_{k+1} = x_k + \gamma_k(\bar{x}_k - x_k) \tag{3}$$

ή

$$x_{k+1} = (1-\gamma)x_k + \gamma\bar{x}_k \tag{4}$$

όπου

$$\bar{x}_k = \Pr\{x_k - s_k \nabla f(x_k)\} \tag{5}$$

Αντικαθιστώντας την κλίση στην (5) παίρνουμε:

$$\bar{x}_k = \Pr\left\{x_k - s\begin{bmatrix} x_{1,k} \\ 4x_{2,k} \end{bmatrix}\right\}$$

και τελικά

$$\bar{x}_k = \Pr\left\{\begin{bmatrix} (1-s)x_{1,k} \\ (1-4s)x_{2,k} \end{bmatrix}\right\} \tag{6}$$

Στο πρόβλημά μας οι περιορισμοί εκφράζονται από φράγματα στις μεταβλητές της $f(x)$. Το σύνολο των εφικτών σημείων,

$$X = \{(x_1, x_2) \in \mathbb{R}^n : -15 \le x_1 \le 15, -20 \le x_2 \le 12\}$$

είναι κυρτό και η προβολή για κάθε σημείο $x_k = (x_1, x_2)$ δίνεται από την σχέση στη σελίδα 202 του βιβλίου:

$$[\Pr_X\{x_1\}]_1 = \begin{cases} -15, & \text{αν } x_1 \le -15 \\ x_1, & \text{αν } -15 < x_1 < 15 \\ 15, & \text{αν } x_1 \ge 15 \end{cases} \tag{7}$$

$$[\Pr_X\{x_2\}]_2 = \begin{cases} -20, & \text{αν } x_2 \le -20 \\ x_2, & \text{αν } -20 < x_2 < 12 \\ 12, & \text{αν } x_2 \ge 12 \end{cases} \tag{8}$$

Αντικαθιστώντας στις σχέσεις (7) και (8) έχουμε:

$$[\Pr_X\{(1-s)x_1\}]_1 = \begin{cases} -15, & \text{αν } (1-s)x_1 \le -15 \\ (1-s)x_1, & \text{αν } -15 < (1-s)x_1 < 15 \\ 15, & \text{αν } (1-s)x_1 \ge 15 \end{cases} \tag{9}$$

$$[\Pr_X\{(1-4s)x_2\}]_2 = \begin{cases} -20, & \text{αν } (1-4s)x_2 \le -20 \\ (1-4s)x_2, & \text{αν } -20 < (1-4s)x_2 < 12 \\ 12, & \text{αν } (1-4s)x_2 \ge 12 \end{cases} \tag{10}$$

Αντικαθιστώντας την (9), (10) και (5) στην (4) παίρνουμε:

$$x_{1_{k+1}} = (1-\gamma)x_{1_k} + \begin{cases} -15\gamma, & \text{αν } (1-s)x_{1_k} \le -15 \\ (1-s)\gamma x_{1_k}, & \text{αν } -15 < (1-s)x_{1_k} < 15 \\ 15\gamma, & \text{αν } (1-s)x_{1_k} \ge 15 \end{cases} \tag{11}$$

και

$$x_{2_{k+1}} = (1-\gamma)x_{2_k} + \begin{cases} -20\gamma, & \text{αν } (1-s)x_{2_k} \le -20 \\ (1-4s)\gamma x_{2_k}, & \text{αν } -20 < (1-s)x_{2_k} < 12 \\ 12\gamma, & \text{αν } (1-s)x_{2_k} \ge 12 \end{cases} \tag{12}$$

Αν πάρουμε τα x_{1_k}, x_{2_k} ώστε να ικανοποιούν τους περιορισμούς ισχύει

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

είναι προφανές ότι για να έχουμε και πάλι σύγκλιση προς το ελάχιστο αφού βρεθούμε σε κάποιο εφικτό σημείο x_k θέλουμε να ισχύει

$$|1 - 4s_k\gamma| < 1$$

για να συγκλίνουν και οι δύο μεταβλητές στο μηδέν, δηλαδή:

$$0 < s_k\gamma < 0.5 \tag{13}$$

2.2. Θέμα 2

Χρησιμοποιούμε τα ορίσματα $s_k = 8$, $\gamma_k = 0.05$, $\epsilon = 0.01$ και αρχικό σημείο το $(10, -5)$.

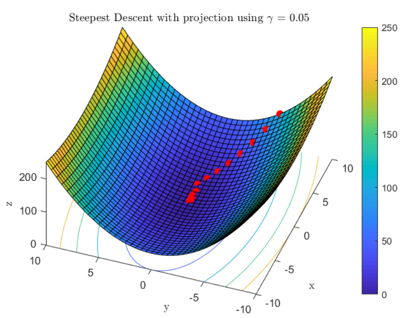


Figure 18: Διάγραμμα επιφάνειας για εκκίνηση από το (10,-5) και γ = 0.05

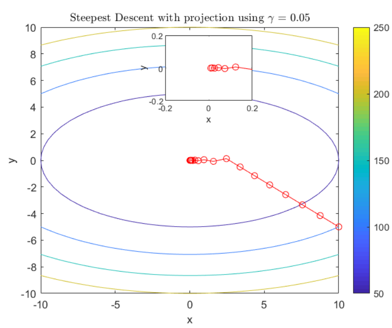


Figure 19: Ισοβαρείς καμπύλες για εκκίνηση από το (10,-5) και γ = 0.05

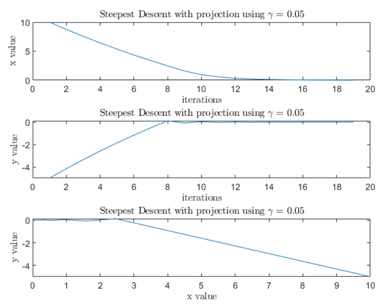


Figure 20: Αλλαγές στις παραμέτρους για εκκίνηση από το (10,-5) και γ = 0.05

Τα s και γ ικανοποιούν τη σχέση (13). Όταν εκτελούμε τον αλγόριθμο τερματίζεται για $k = 19$ (για το δοσμένο αρχικό σημείο x_1) βρίσκοντας επιτυχώς το ελάχιστο.

Σε σχέση με το (1,i) ο αλγόριθμος συγκλίνει γρηγορότερα καθώς σε κάθε βήμα το x_k μειώνεται με μεγαλύτερο ρυθμό, καθώς στο (1,i) για το δεδομένο αρχικό σημείο χρειάστηκε 136 επαναλήψεις. Δηλαδή, στο (1,i) είχαμε $x_{1_{k+1}} = 0.95x_{1_k}$ και $x_{2_{k+1}} = 0.8x_{2_k}$ ενώ εδώ έχουμε $x_{1_{k+1}} = 0.6x_{1_k}$ και $x_{2_{k+1}} = -0.6x_{2_k}$ (όταν η προβολή είναι μέσα στο X), το οποίο δικαιολογεί και πως η Μέθοδος της μέγιστης καθόδου με προβολή είναι γρηγορότερη.

Στο (1,iv) είδαμε το λόγο που ο αλγόριθμος χωρίς προβολή δε συγκλίνει προς το ελάχιστο. Στην περίπτωση με την προβολή συγκλίνει καθώς ικανοποιείται η συνθήκη (13).
Συνεπώς, τα αποτελέσματα που πήραμε από την προσομοίωση είναι αναμενόμενα.

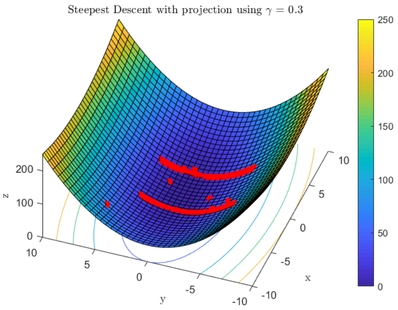


Figure 21: Διάγραμμα επιφάνειας για εκκίνηση από το (-7,5) και γ = 0.3

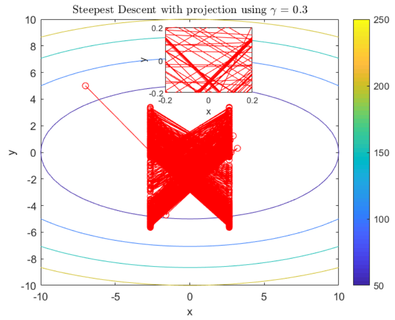


Figure 22: Ισοβαρείς καμπύλες για εκκίνηση από το (-7,5) και γ = 0.3

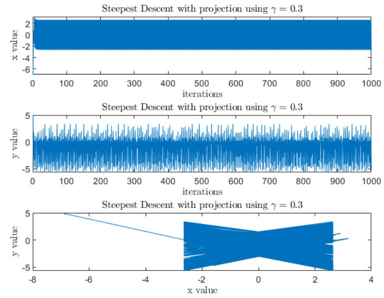


Figure 23: Αλλαγές στις παραμέτρους για εκκίνηση από το (-7,5) και γ = 0.3

2.3. Θέμα 3

Χρησιμοποιούμε τα ορίσματα $s_k = 10$, $\gamma_k = 0.3$, $\epsilon = 0.02$ και αρχικό σημείο το $(-7, 5)$.

Τα s και γ ΔΕΝ ικανοποιούν τη σχέση (13). Έτσι, αν εκτελέσουμε τον αλγόριθμο ΔΕ θα τερματίσει ποτέ. Στην προσομοίωση θέσαμε ανώτερο όριο επαναλήψεων ίσο με 1000.

Στο (1,iv) ο αλγόριθμος συγκλίνει σε 129 επαναλήψεις (με $\epsilon = 0.01$) και σε 116 (με $\epsilon = 0.02$). Ωστόσο, δεν τίθεται ζήτημα σύγκρισης, καθώς η μέθοδος με προβολή δεν τερματίζει ποτέ.

Στο (1,iv) είδαμε τον λόγο που ο αλγόριθμος χωρίς προβολή δε συγκλίνει προς το ελάχιστο. Αυτή η περίπτωση είναι κατά μία έννοια όμοια με την περίπτωση της μεθόδου με προβολή που εξετάζουμε, καθώς και στις δύο περιπτώσεις δεν ικανοποιείται η βασική συνθήκη για το γ και το sγ αντίστοιχα.
Στην περίπτωση (1,iv) ο αλγόριθμος δε συγκλίνει και απειρίζεται. Στην περίπτωσή μας ο αλγόριθμος δε συγκλίνει αλλά δεν απειρίζεται. Αυτό εξηγείται εύκολα καθώς το $\bar{x}_k$ μπορεί να πάρει περιορισμένες τιμές, λόγω της προβολής στο X.

Για να λύσουμε αυτό το πρόβλημα αρκεί να επιλέξουμε τιμές που ικανοποιούν τη (13). Πχ, για $s = 0.5$, $\gamma = 0.4$ ο αλγόριθμος συγκλίνει στο ελάχιστο για $k = 28$.

Τέλος, και πάλι, μπορούμε να πούμε ότι τα αποτελέσματα που πήραμε από την προσομοίωση είναι αναμενόμενα, με βάση τους υπολογισμούς που κάναμε παραπάνω.

2.4. Θέμα 4

Χρησιμοποιούμε τα ορίσματα $s_k = 0.5$, $\gamma_k = 0.1$, $\epsilon = 0.01$ και αρχικό σημείο το $(17, -5)$. Πριν ακόμη τρέξουμε τον αλγόριθμο, παρατηρούμε πως το σημείο που δίνεται δεν είναι εφικτό, άρα προβάλλεται στο σύνολο X έτσι ώστε να πληροί τις προϋποθέσεις για την εκκίνηση του αλγορίθμου.
Ακόμα, το βήμα $s\gamma = 0.05$ έχει πολύ μικρή τιμή, που μπορεί να οδηγήσει σε αργή σύγκλιση. Τα s και γ ικανοποιούν τη σχέση (13), οπότε περιμένουμε να υπάρξει σύγκλιση, αλλά να μην είναι γρήγορη.

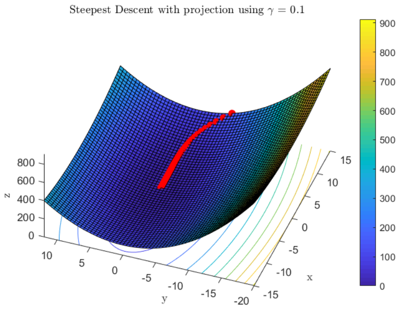


Figure 24: Διάγραμμα επιφάνειας για εκκίνηση από το (17,-5) και γ = 0.1

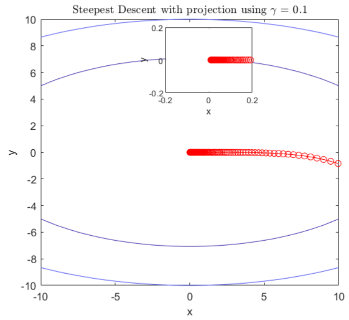


Figure 25: Ισοβαρείς καμπύλες για εκκίνηση από το (17,-5) και γ = 0.1

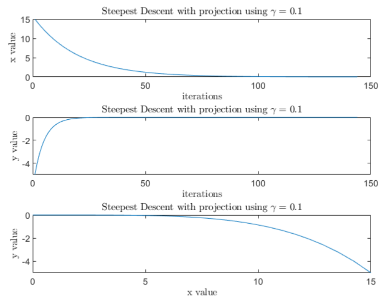


Figure 26: Αλλαγές στις παραμέτρους για εκκίνηση από το (17,-5) και γ = 0.1

Αν εκτελέσουμε τον αλγόριθμο επιβεβαιώνονται οι υποθέσεις μας αφού αυτός τερματίζει για $k = 144$. Αυτή η ταχύτητα σύγκλισης είναι σχετικά αργή σε σχέση με τις προηγούμενες (αν και θα βλέπαμε ακόμα αργότερη σύγκλιση αν είχαμε $s\gamma$ ακόμα μικρότερο) και για να τη βελτιώσουμε θα πρέπει να αυξήσουμε τις τιμές του γ και s οι οποίες όμως πρέπει να τηρούν τη σχέση (13).

2.5. Συμπεράσματα και σχολιασμός αποτελεσμάτων

Σύμφωνα με τα παραπάνω, συμπεραίνουμε πως ο αλγόριθμος της **Μέγιστης Καθόδου με προβολή** μπορεί να συγκλίνει στο ελάχιστο μιας συνάρτησης αρκετά **γρηγορότερα** (λιγότερα βήματα), αν γίνει **σωστή επιλογή των παραμέτρων** s_k και γ, ενώ σε αντίθετη περίπτωση (βλ. θέμα 4) μπορεί να γίνει εξαιρετικά αργός. Το γινόμενο $s_k\gamma$ πρέπει να ικανοποιεί τους **περιορισμούς** που θέτει η κάθε επιλεγμένη συνάρτηση προς ελαχιστοποίηση. Ακόμη, θα πρέπει να είναι γνωστό το σύνολο X μέσα στο οποίο θα γίνει η αναζήτηση του ελαχίστου. Αν οι παραπάνω προϋποθέσεις ισχύουν ο αλγόριθμος της μέγιστης καθόδου με προβολή, που είναι αλγόριθμος αναζήτησης με περιορισμούς, είναι **αποδοτικότερος** έναντι του αντίστοιχου χωρίς περιορισμούς, για την ίδια συνάρτηση.