ΤΕΧΝΙΚΕΣ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ

Νικόλαος Γιακουμόγλου 9043

11 Δεκεμβρίου 2020

Part I
Steepest Descent

Steepest Descent

Στην μέθοδο αυτή έχουμε

$$x(k+1) = x(k) - \gamma(k) \cdot \Delta(k) \cdot \nabla f\left(x(k)\right) = x(k) + \gamma(k) \cdot d(k)$$

$$\Delta(k) = I > 0$$

$$d(k) = -\Delta(k) \cdot \nabla f\left(x(k)\right) = -\nabla f\left(x(k)\right)$$

Άρα

$$x(k+1) = x(k) - \gamma(k) \cdot \nabla f\left(x(k)\right)$$

όπου $\gamma(k) = \gamma = ct$. Κριτήριο τερματισμού αποτελεί αν το μέτρο της κλίσης $\nabla f(x, y) < \varepsilon$ όπου ε η ακρίβεια ή αν υπερβούμε έναν ανώτερο αριθμό επαναλήψεων. Η μέθοδος Steepest Descent είναι η ίδια με προηγουμένως μόνο που αφαιρέθηκε η παράμετρος για τον τρόπο επιλογής του γ αφού θα τον θεωρήσουμε σταθερό.

1 ΘΕΜΑ 1

1.1 Εκτύπωση συνάρτησης $f(x_1, x_2)$

Εκτελούμε τον παρακάτω κώδικα για να δούμε την σνάρτηση $f(x_1, x_2)$

```
syms x y
f = 0.5*x^2 + 0.5*y^2;
figure
title('Function f(x,y)');
fsurf(f)
xlabel('x');
ylabel('y');
zlabel('f(x,y)');
saveas(gcf,[pwd '\function.png'])
```

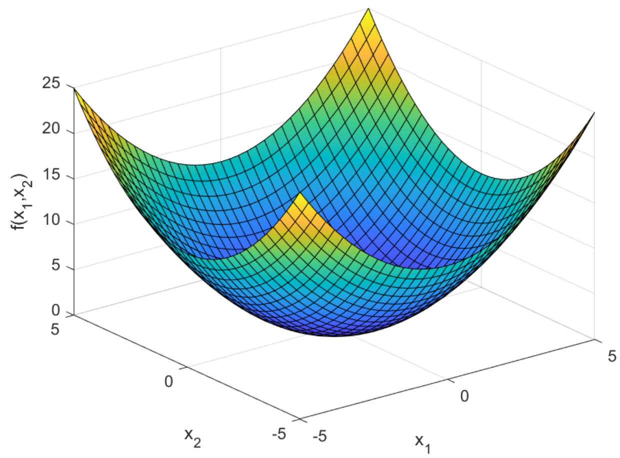


Figure 1: $f(x_1, x_2) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2$

1.2 Πειραματικά αποτελέσματα

Επιλέγουμε ως αρχικό σημείο οποιοδήποτε εκτός του $(0, 0)$ έστω $(x_0, y_0) = (1, 1)$. Παρατίθενται τα αποτελέσματα.

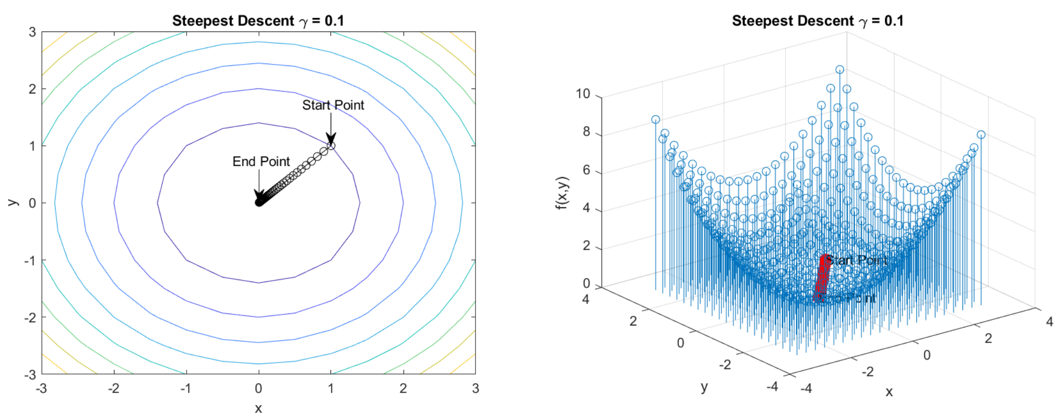


Figure 2: Αποτελέσματα για $\gamma = 0.1$

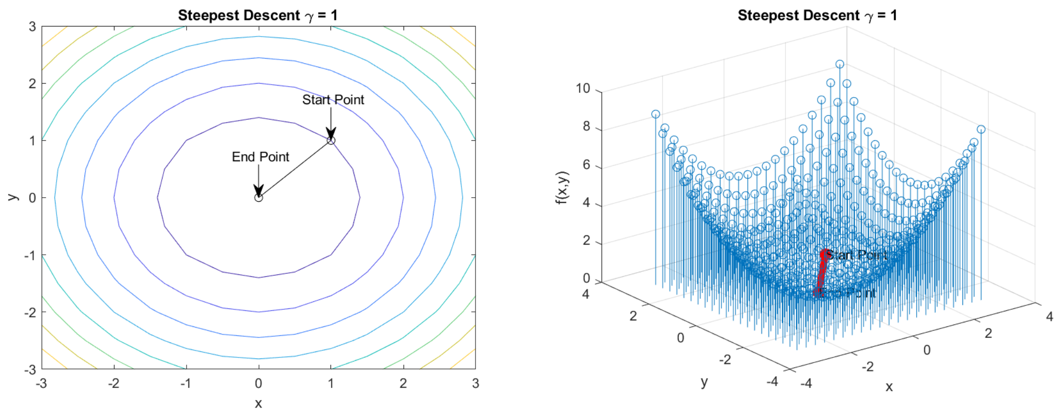


Figure 3: Αποτελέσματα για $\gamma = 1$

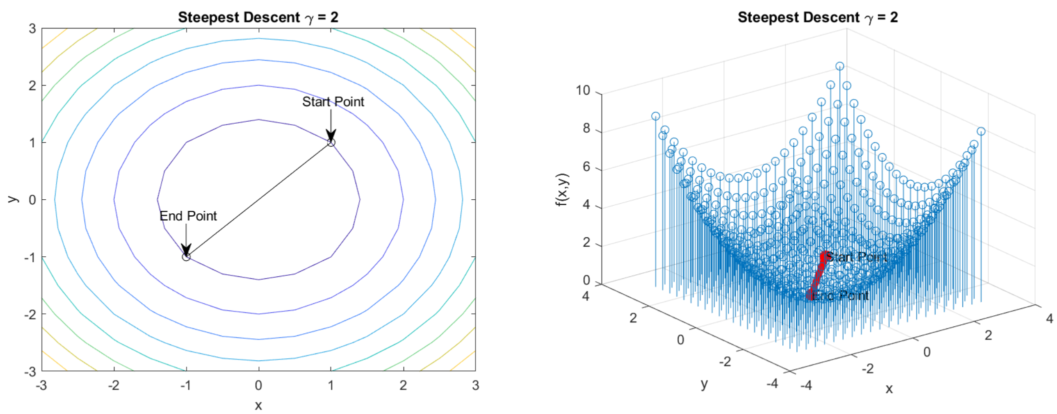


Figure 4: Αποτελέσματα για $\gamma = 2$

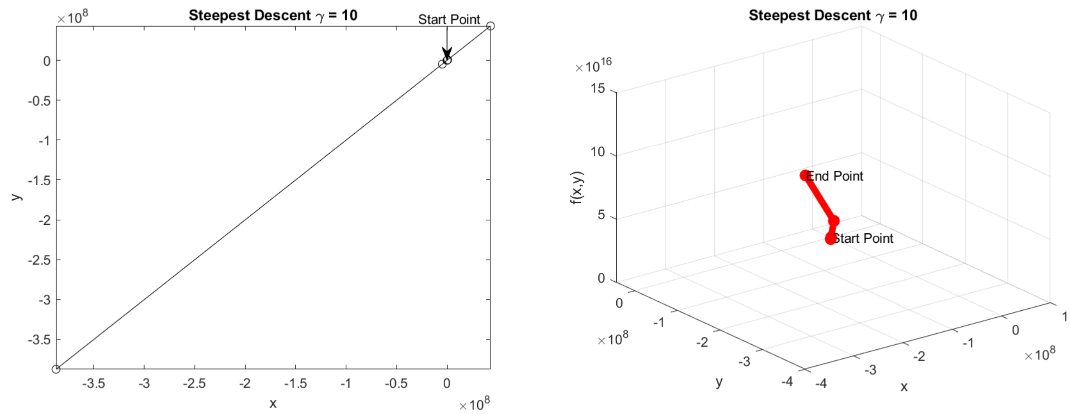


Figure 5: Αποτελέσματα για $\gamma = 10$

Βλέπουμε ότι για $\gamma = 0.1$ και 1 ο αλγόριθμος συγκλίνει σε λίγα μόλις βήματα. Για $\gamma \geq 2$ ο αλγόριθμος αποκλίνει. Συγκεκριμένα για $\gamma = 2$ τον σταματάμε γιατί υπερβαίνει τις μέγιστες επαναλήψεις που θέσαμε ως 10.000. Αυτό συμβαίνει γιατί υπάρχει ταλάντωση μεταξύ 2 σημείων. Για $\gamma = 10$ ο αλγόριθμος φεύγει πολύ γρήγορα στο άπειρο. Στα σχήματα μάλιστα δεν βλέπουμε την συνάρτηση γιατί ο αλγόριθμος εκτυπώνει ένα διάστημα αυτής, αλλά ούτως ή άλλως οι αριθμητικές τιμές φθάνουν το 10^{16}! Ο αλγόριθμος σταματάει γιατί ικανοποιείται η συνθήκη με την ακρίβεια.

γ	0.1	1	2	10
Iterations	48	2	MAX_ITER	325

Table 1: Αποτελέσματα

1.3 Θεωρητική Ανάλυση

$$f(x_1, x_2) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2$$

$$\nabla f(x_1, x_2) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} > 0$$

Έχουμε

$$x(k+1) = x(k) - \gamma(k) \cdot \nabla f\left(x(k)\right) \Rightarrow$$

$$\Rightarrow x(k+1) = x(k) - \gamma(k) \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \Rightarrow$$

$$\Rightarrow \begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} - \gamma \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} \Rightarrow$$

$$x(k+1) = (1-\gamma) \cdot x(k)$$

Ο αλγόριθμος συγκλίνει αν

$$\frac{|x(k+1)|}{|x(k)|} < 1 \Rightarrow$$
$$\Rightarrow |1-\gamma| < 1 \Rightarrow$$
$$\gamma \in (0,2)$$

Πράγματι επιβεβαιώνουμε τα πειραματικά αποτελέσματα σύγκλισης για σταθερό γ.

Part II
Steepest Descent Projection

Steepest Descent Projection

Στην μέθοδο αυτή έχουμε περιορισμούς οπότε εισάγουμε την έννοια της προβολής.
Αρχικά προβάλλουμε το $x(k)$ στο X όπου X το σύνολο των περιορισμών. Μετά έχουμε:

$$x(k+1) = x(k) + \gamma(k) \cdot (\overline{x} - x(k))$$

όπου

$$\overline{x} = Pr_X \{x(k) - s(k) \cdot \nabla f\left(x(k)\right)\}$$

Για την εργασία $\gamma(k) = \gamma = ct$ και $s(k) = s = ct$. Η προβολή υπολογίζεται ως εξής. Έστω $a \leq x \leq b$. Τότε

$$Pr_X\{x\} = \begin{cases} a, & x \leq a \\ x, & a < x < b \\ b, & x \geq b \end{cases}$$

Κριτήριο τερματισμού αποτελεί αν το μέτρο της κλίσης $\nabla f(x,y) < \varepsilon$ όπου ε η ακρίβεια ή αν υπερβούμε έναν ανώτερο αριθμό επαναλήψεων. Οι περιορισμοί για την εργασία είναι της μορφής

$$-20 \leq x_1 \leq 10$$

$$-12 \leq x_2 \leq 15$$

Σχέση με Steepest Descent

Αν αντικαταστήσουμε το $\overline{x}$ στο $x(k+1)$ και θεωρήσουμε ότι δεν παίρνουμε την προβολή, δηλαδή $Pr_X\{x\} = x$, έχουμε

$$x(k+1) = x(k) + \gamma(k)\left(\cancel{x(k)} - s(k) \cdot \nabla f\left(x(k)\right) - \cancel{x(k)}\right) = x(k) - \gamma(k) \cdot s(k) \cdot \nabla f\left(x(k)\right)$$

που ουσιαστικά προκύπτει από τον απλό με την αντικατάσταση

$$\gamma(k) \leftarrow \gamma(k) \cdot s(k)$$

2 ΘΕΜΑ 2

2.1 Πειραματικά αποτελέσματα

Θεωρούμε $(x_0, y_0) = (8, 3)$, $\varepsilon = 0.01$, $s(k) = 15$ και $\gamma(k) = 0.1$. Ο αλγόριθμος συγκλίνει σε 10 μόλις βήματα. Αν τρέχαμε τον απλό Steepest Descent θα είχαμε $\gamma'(k) = 1.5$. Η σύγκλιση ήταν επομένως αναμενόμενη

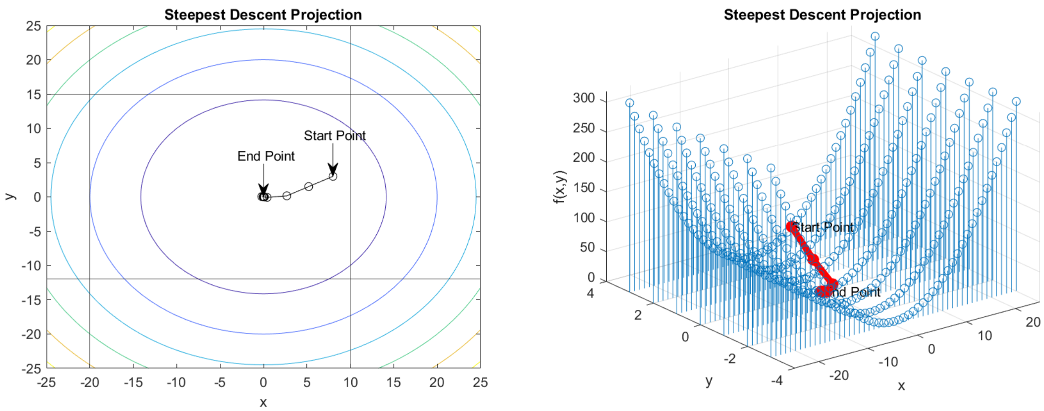


Figure 6: Αποτελέσματα

3 ΘΕΜΑ 3

Θεωρούμε $(x_0, y_0) = (-5, 7)$, $\varepsilon = 0.02$, $s(k) = 20$ και $\gamma(k) = 0.3$. Ο αλγόριθμος συγκλίνει σε 1000 βήματα. Αν τρέχαμε τον απλό Steepest Descent θα είχαμε $\gamma'(k) = 6$. Η προβολή αποτυγχάνει να βρεί το ελάχιστο. Για να βρούμε το ελάχιστο μια προφανής λύση θα ήταν να μειώσουμε το $s(k)$ στο 1 για παράδειγμα. Με αυτόν τον τρόπο ο αλγόριθμος συγκλίνει σε 19 επαναλήψεις.

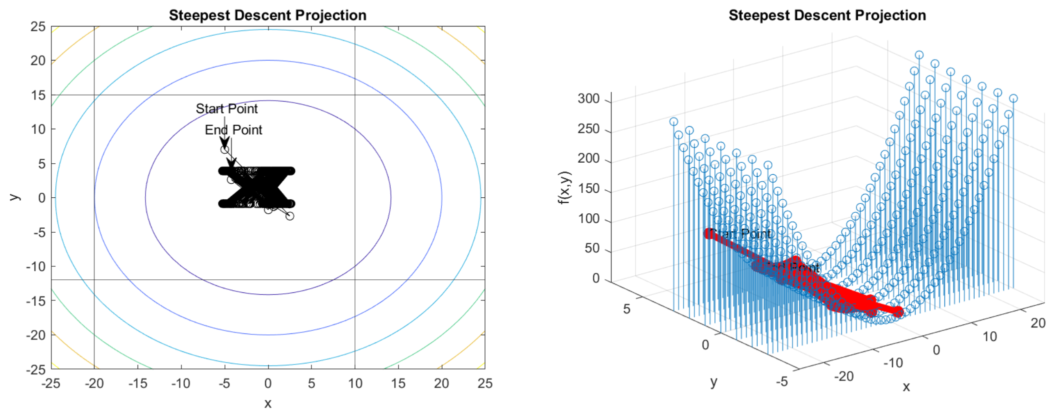


Figure 7: Αποτελέσματα

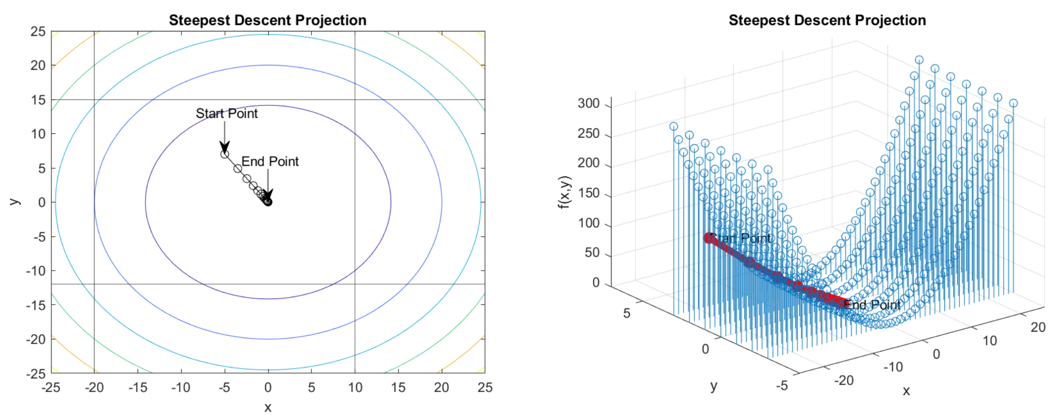
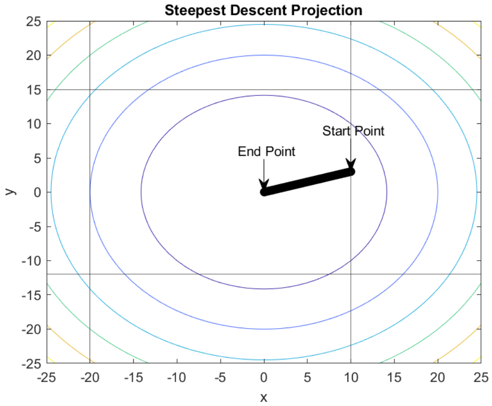


Figure 8: Βελτιωμένα Αποτελέσματα

4 ΘΕΜΑ 4

Θεωρούμε $(x_0, y_0) = (11, 3)$, $\varepsilon = 0.01$, $s(k) = 0.1$ και $\gamma(k) = 0.01$. Το αρχικό σημείο είναι εκτός των περιορισμών άρα προβάλλεται στο σύνολο X. Ο αλγόριθμος συγκλίνει σε ~ 7000 βήματα. Αν τρέχαμε τον απλό Steepest Descent θα είχαμε $\gamma'(k) = 0.001$. Η σύγκλιση ήταν επομένως αναμενόμενη αλλά λόγω του μικρού βήματος ο αλγόριθμος κάνει πολλές επαναλήψεις.



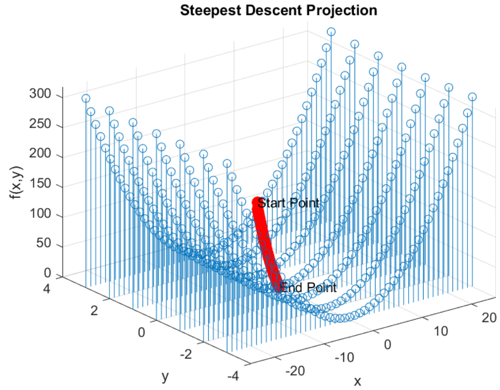


Figure 9: Αποτελέσματα