7/12/2023

Τεχνικές Βελτιστοποίησης

ΕΡΓΑΣΙΑ 3

Μάριος Τζαμτζής ΑΕΜ: 10038

ΕΙΣΑΓΩΓΗ

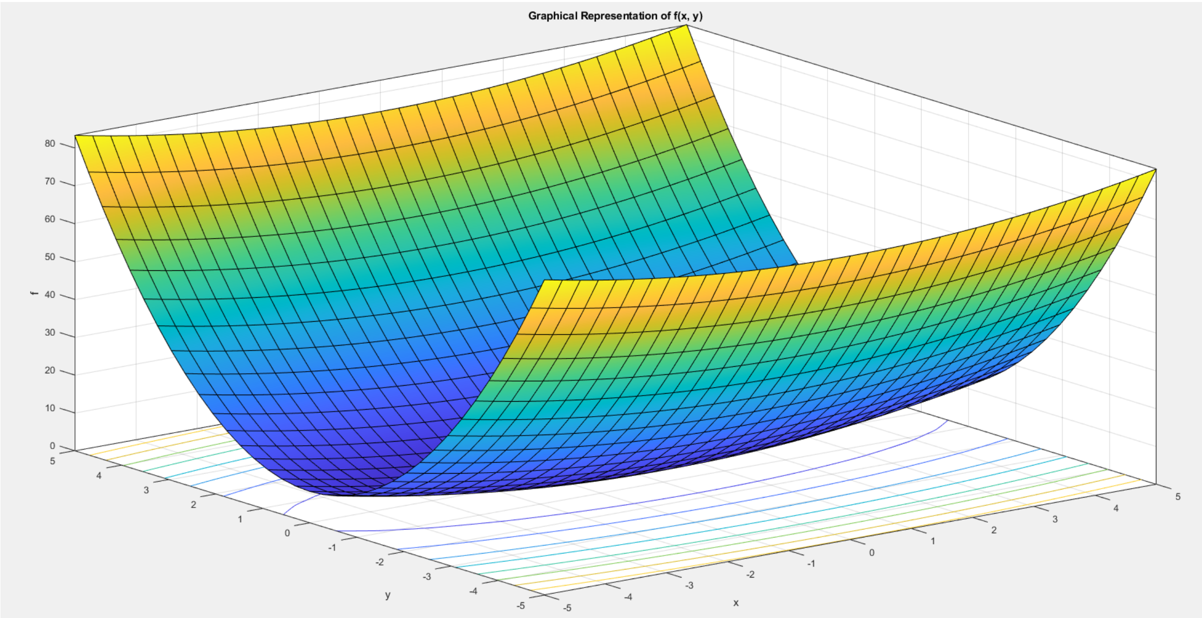
Στην παρούσα εργασία ζητήθηκε η ελαχιστοποίηση της συνάρτησης:

$$f(x) = \frac{1}{3}x_1^2 + 3x_2^2$$

με τη χρήση μεθόδων αναζήτησης ελαχίστου. Οι μέθοδοι που χρησιμοποιήθηκαν ήταν:

- Η μέθοδος της Μέγιστης Καθόδου
- Η μέθοδος της Μέγιστης Καθόδου με προβολή

Παρακάτω γίνεται μια σύντομη ανάλυση κάθε θέματος της εργαστηριακής άσκησης. ‘Ολες οι μέθοδοι υλοποιήθηκαν στο περιβάλλον του MATLAB.



ΘΕΜΑ 1

Στο 1ο θέμα της εργασίας καλούμαστε να χρησιμοποιήσουμε στην συνάρτηση την μέθοδο της Μέγιστης Καθόδου που υλοποιήσαμε στην $2^{η}$ εργασία. Προσπαθούμε λοιπόν να ελαχιστοποιήσουμε την f με ακρίβεια $\epsilon = 0.001$ και διαφορετικά βήματα:

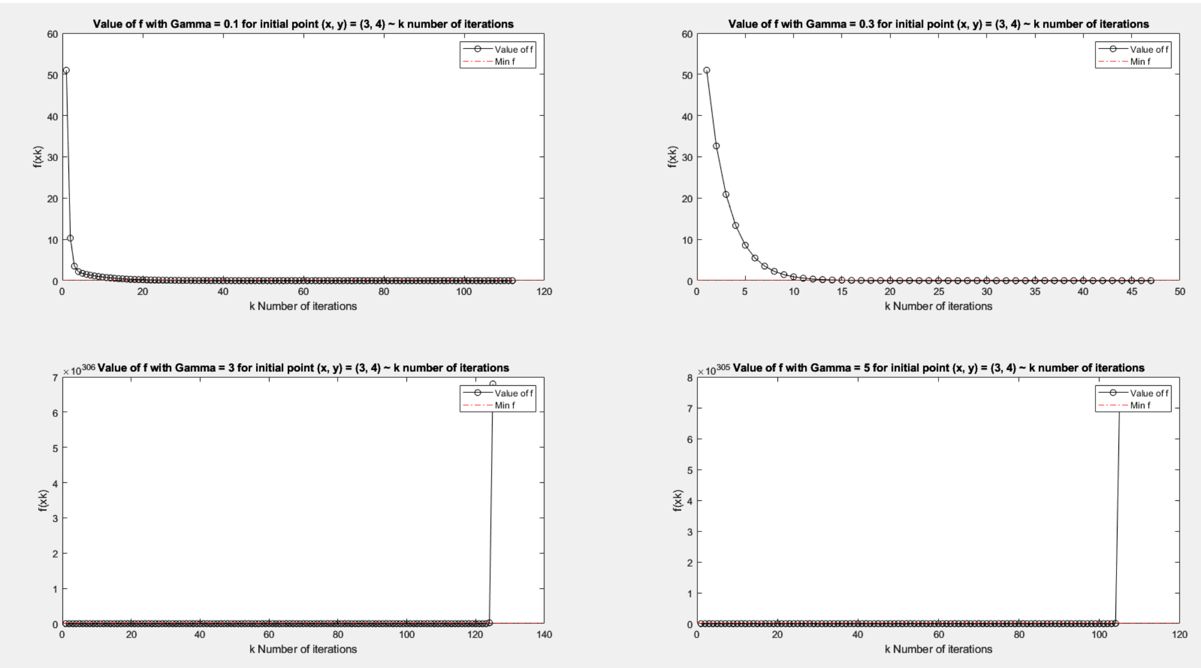
(i) $\gamma_k = 0.1$

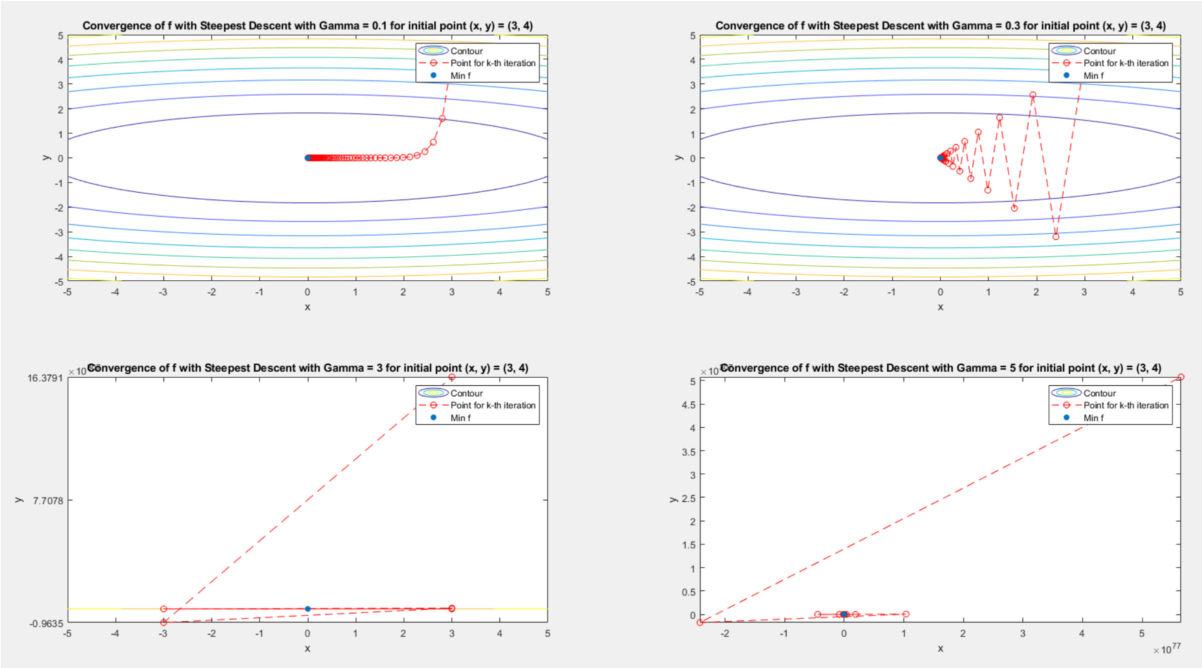
(ii) $\gamma_k = 0.3$

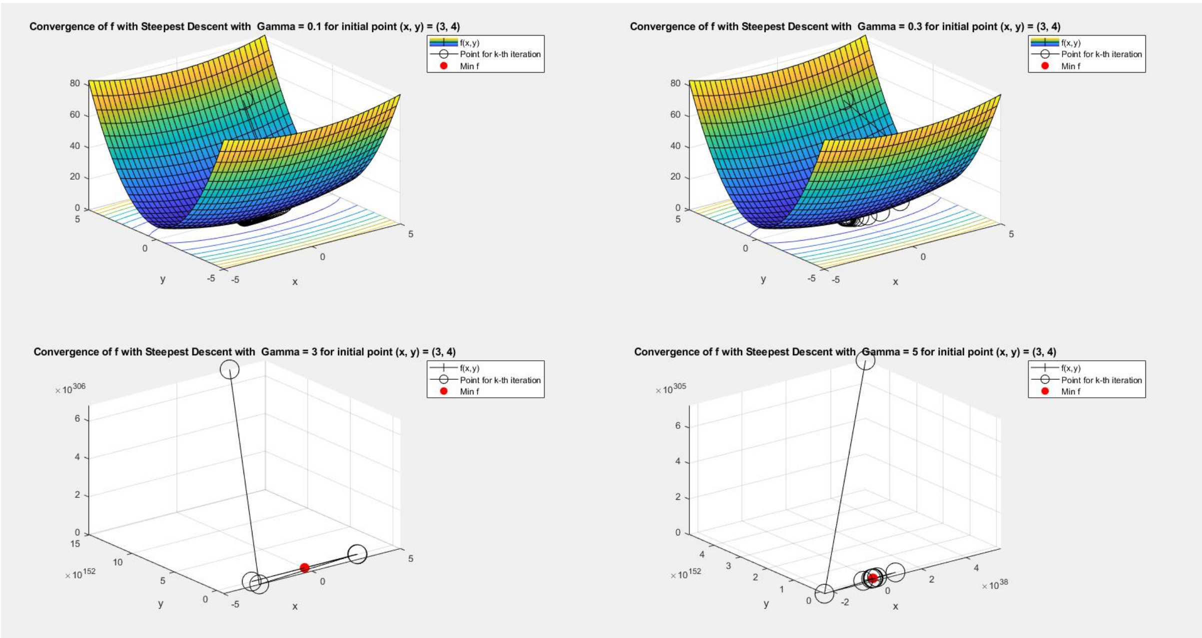
(iii) $\gamma_k = 3$

(iv) $\gamma_k = 5$

Σαν σημείο εκκίνησης για την μέθοδο επιλέξαμε αυθαίρετα το (3,4) και παίρνουμε τα παρακάτω διαγράμματα για τις τιμές του $\gamma_k = 0.1$ και $\gamma_k = 0.3$. Για $\gamma_k = 3$ και $\gamma_k = 5$ η μέθοδος αποκλίνει. Στην περίπτωση (i) η μέθοδος βρίσκει το ελάχιστο με σχετικά καλή ακρίβεια. Στην περίπτωση (ii) επίσης η μέθοδος ξεκινάει να ελαχιστοποιεί την συνάρτηση μας και πλησιάζει στο πραγματικό ελάχιστο σημείο της και με λιγότερες επαναλήψεις από την (i). Αντίθετα στις περιπτώσεις (iii) και (iv) για τα πιο μεγάλα γ_k η μέθοδος αποκλίνει κατά πολύ από το ελάχιστο ενώ οι τιμές της συνάρτησης παρουσιάζουν εκθετική αύξηση.







Θεωρητική ανάλυση:

Αρχικά υπολογίζω το gradient της συνάρτησης:

$$\nabla f\left(x_{1_\kappa}, x_{2_\kappa}\right) = \begin{bmatrix} \dfrac{\partial f}{\partial x_{1_\kappa}}\left(x_{1_\kappa}, x_{2_\kappa}\right) \\ \dfrac{\partial f}{\partial x_{2_\kappa}}\left(x_{1_\kappa}, x_{2_\kappa}\right) \end{bmatrix} = \begin{bmatrix} \dfrac{2}{3} \cdot x_{1_\kappa} \\ 6 \cdot x_{2_\kappa} \end{bmatrix}$$

Οπότε για την μέθοδο θα είναι: $\begin{pmatrix} x_{1\,k+1} \\ x_{2\,k+1} \end{pmatrix} = \begin{pmatrix} x_{1\,k} \\ x_{2\,k} \end{pmatrix} - \gamma_k \begin{pmatrix} \frac{2}{3}\, x_{1\,k} \\ 6\, x_{2\,k} \end{pmatrix}$

Από την τρισδιάστατη αναπαράσταση της f παραπάνω μπορούμε να δούμε πως η συνάρτηση εμφανίζει ελάχιστο στο σημείο $(x_1 ,x_2) = (0, 0)$. Οπότε αυτό που θέλουμε να επιτύχουμε είναι για $k \to \infty$: $x_{1k} = 0$ και $x_{2k} = 0$ Για να γίνει αυτό πρέπει να συγκλίνουν οι 2 σχέσεις, οπότε:

$$\begin{cases} |1 - \frac{2}{3}\gamma_\kappa| < 1 \\ |1 - 6\gamma_\kappa| < 1 \end{cases} \Leftrightarrow 0 < \gamma_\kappa < \frac{1}{3}$$

Βλέπουμε λοιπόν πως η τιμή του γ_k πρέπει να κυμαίνεται μεταξύ του 0 και του 0.33 για να συγκλίνει η μέθοδος μας, κάτι που επαληθεύει και τα αποτελέσματα από τα διαγράμματα μας παραπάνω. Για τις περιπτώσεις (iii) και (iv) η πολύ μεγάλη απόκλιση εξηγείται από το ότι τα συγκεκριμένα γ_k απέχουν πολύ από το διάστημα που καταλήξαμε παραπάνω.

ΜΕΘΟΔΟΣ ΜΕΓΙΣΤΗΣ ΚΑΘΟΔΟΥ ΜΕ ΠΡΟΒΟΛΗ

Η μέθοδος Μέγιστης Καθόδου παρότι είναι ένας χρήσιμος αλγόριθμος στον τομέα της ελαχιστοποίησης αντιμετωπίζει προβλήματα όταν το πρόβλημα που προσπαθούμε να επιλύσουμε θέτει περιορισμούς, παραδείγματος χάριν όταν το διάνυσμα μας x_k πρέπει να βρίσκεται συνεχώς εντός ενός κυρτού συνόλου. Αυτό το πρόβλημα έρχεται να λύσει η μέθοδος μέγιστης καθόδου με προβολή η οποία ξεκινά με ένα εφικτό σημείο και συνεχίζει με τον αλγόριθμο της μέγιστης καθόδου έως ότου βρει μη εφικτό σημείο x_k οπότε και βρίσκει την προβολή αυτού στο κυρτό σύνολο και επαναλαμβάνει την ίδια διαδικασία. Τα εφικτά σημεία δηλαδή του αλγορίθμου είναι της μορφής:

$x_{k+1} = x_k + \gamma_k(\bar{x} - x_k)$ όπου $\bar{x} = P_{r_x}\{x_k - S_k \nabla f(x_k)\}$

Για εφικτό σημείο x_k όμως θα είναι $\bar{x} = x_k - S_k \nabla f(x_k)$ ακολουθούμε δηλαδή την μέθοδο της μέγιστης καθόδου με βήμα $\gamma'_k = \gamma_k \cdot s_k$ και για το γ'_k θα ισχύουν οι περιορισμοί του θέματος 1 ($0 < \gamma'_k < 0.33$). Για την συνέχεια της εργασίας ισχύουν οι παρακάτω περιορισμοί για τα x1, x2 :

$-10 \le x_1 \le 5$ και $-8 \le x_2 \le 12$

Βασιζόμενοι στο βιβλίο ορίζουμε τις προβολές των x_1, x_2 στο X ως εξής:

$$[P_{r_x}\{x\}]_1 = \begin{cases} -10, & x_1 \le -10 \\ x_1, & -10 \le x_1 \le 5 \\ 5, & x_1 \ge 5 \end{cases}$$

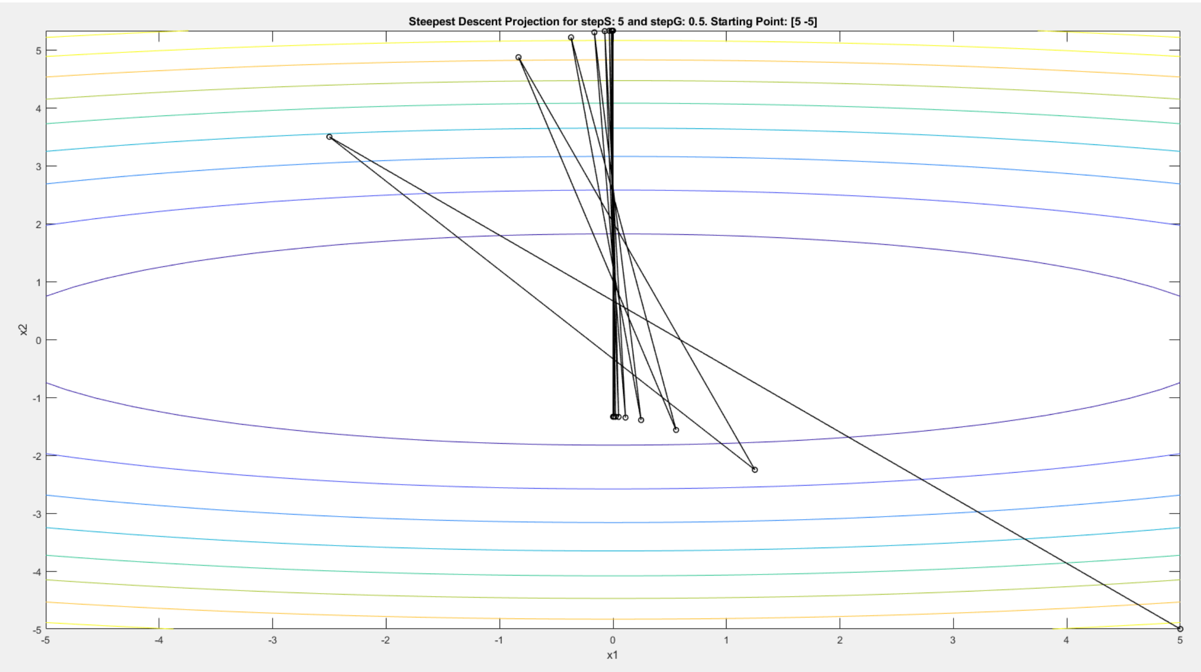
$$[P_{r_x}\{x\}]_2 = \begin{cases} -8, & x_2 \le -8 \\ x_2, & -8 \le x_2 \le 12 \\ 12, & x_2 \ge 12 \end{cases}$$

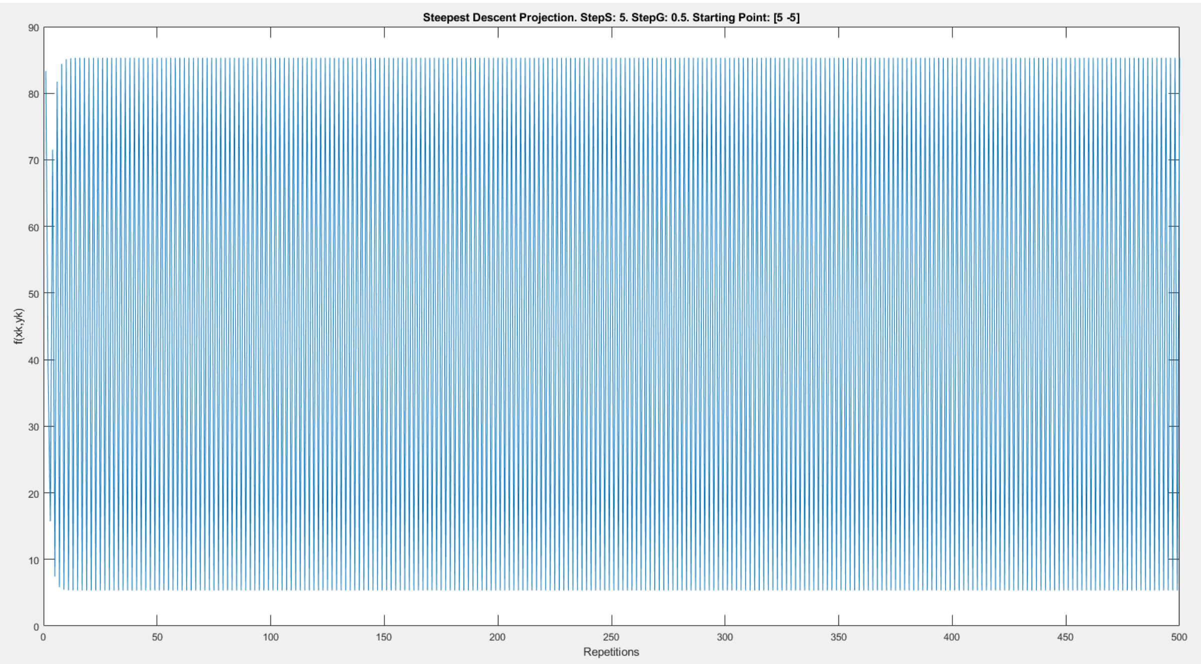
Τέτοιοι περιορισμοί δημιουργούν παραλληλόγραμμα που φράζουν εσωτερικά τους τις τιμές της συνάρτησης.

ΘΕΜΑ 2

Στο Θέμα 2 χρησιμοποιούμε την μέθοδο μέγιστης καθόδου με προβολή με s_k= 5, γ_k = 0.5, ε = 0.01 και αρχικό σημείο εκκίνησης του αλγορίθμου το (5,-5). Λαμβάνοντας υπόψιν και τα παραπάνω θα είναι $\gamma'_k = \gamma_k \cdot s_k = 0.5 \cdot 5 = 2.5$ δηλαδή $\gamma'_k > 0.33$ και θα αναμένουμε να μην συγκλίνει η μέθοδος μας με αυτές τις αρχικές συνθήκες.

Ο αλγόριθμος αδυνατεί να τερματίσει για αυτό και τέθηκε όριο στις 500 επαναλήψεις.



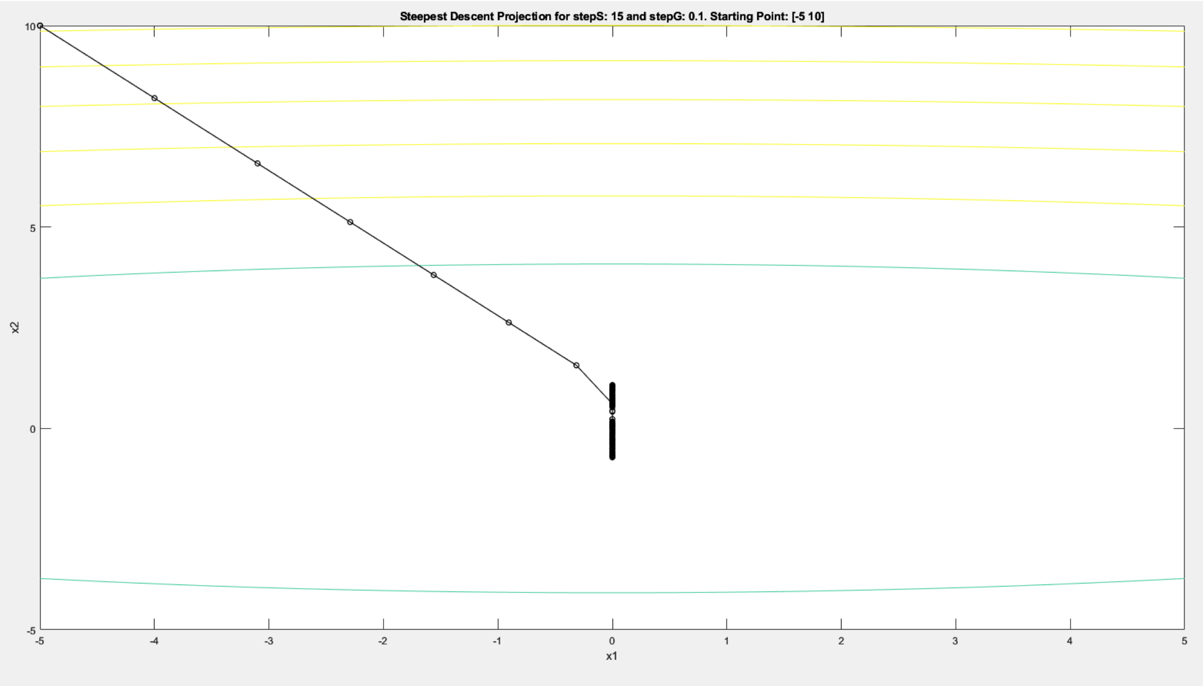


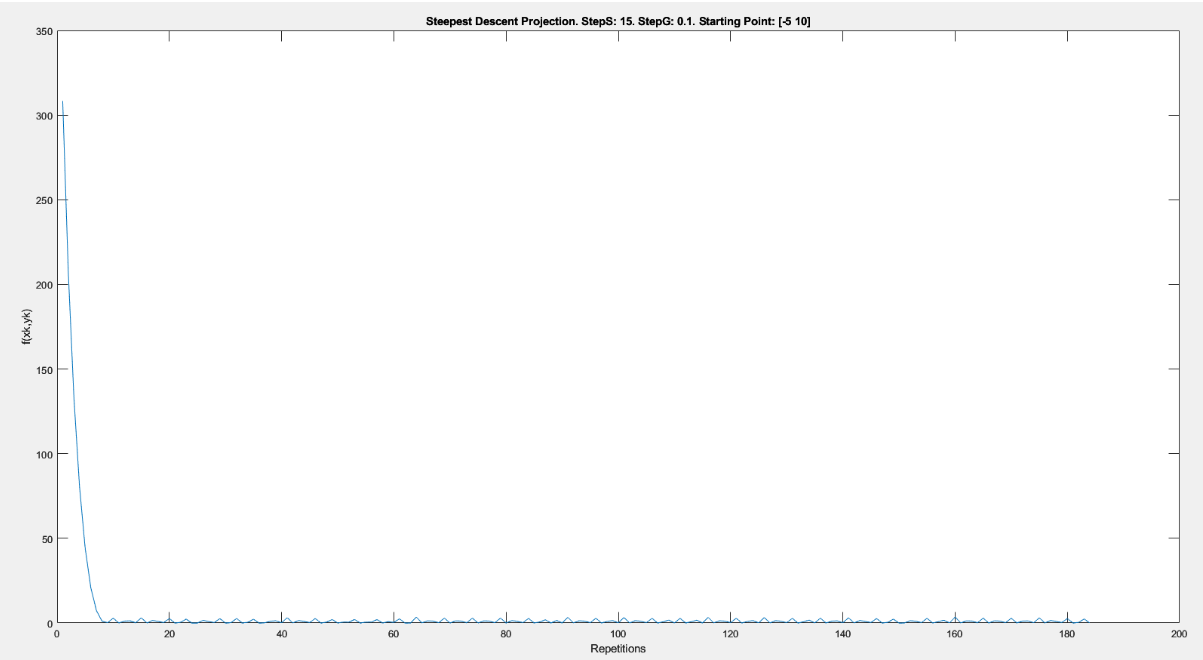
Σε αυτή όμως την περίπτωση παρότι $\gamma'_k > 0.33$ και θα περιμέναμε οι τιμές της συνάρτησης να αυξάνονται εκθετικά, βλέπουμε ότι η μέθοδος ταλαντεύεται μιας και η χρήση της προβολής της μεθόδου και το μεγάλο s_k εγκλωβίζει το x_k εντός των περιορισμών. Επομένως σε σύγκριση με το Θέμα 1(i) εδώ δεν καταλήγουμε σε ελάχιστο με μια ομαλή πορεία και σε πεπερασμένο αριθμό βημάτων αλλά υπάρχει συνεχής αλλαγή κατεύθυνσης και μη σύγκλιση στο σημείο ελαχίστου. Συγκριτικά με το Θέμα 1(iv) τώρα και οι 2 περιπτώσεις έχουν μεγαλύτερο γ_k από το επιθυμητό για την σύγκλιση και ενώ στην μια περίπτωση οι τιμές αυξάνονται εκθετικά απομακρυνόμενες κατά πολύ από το (0,0) στην περίπτωση της μέγιστης καθόδου με προβολή οι τιμές διατηρούνται εντός του παραλληλογράμμου που περιγράψαμε παραπάνω και μπορεί σε κάποιο πολύ μεγάλο αριθμό επαναλήψεων να βρεθούν έστω και "τυχαία" στο επιθυμητό σημείο τερματίζοντας έτσι τον αλγόριθμο.

΄Οπως είδαμε το μεγάλο s_k σε συνδυασμό με το $\gamma k = 0.5$ αποτρέπουν την μέθοδο από το να συγκλίνει. Οπότε ένας πρακτικός τρόπος να το αλλάξουμε αυτό θα ήταν εάν θέταμε π.χ $s_k = 0.6$ και τότε θα επιτυγχάναμε την σύγκλιση όπως μπορούμε να δούμε και παρακάτω αφού θα είναι $\gamma'_k = 0.5 \cdot 0.6 = 0.3$

ΘΕΜΑ 3

Στο Θέμα 3 χρησιμοποιούμε την μέθοδο μέγιστης καθόδου με προβολή με s_k= 15, γ_k = 0.1, ε = 0.01 και αρχικό σημείο εκκίνησης του αλγορίθμου το (-5,10). Λαμβάνοντας υπόψιν και τα παραπάνω θα είναι $\gamma'_k = \gamma_k \cdot s_k = 0.5 \cdot 5 = 2.5$ δηλαδή $\gamma'_k > 0.33$ και θα αναμένουμε να μην συγκλίνει η μέθοδος μας με αυτές τις αρχικές συνθήκες.

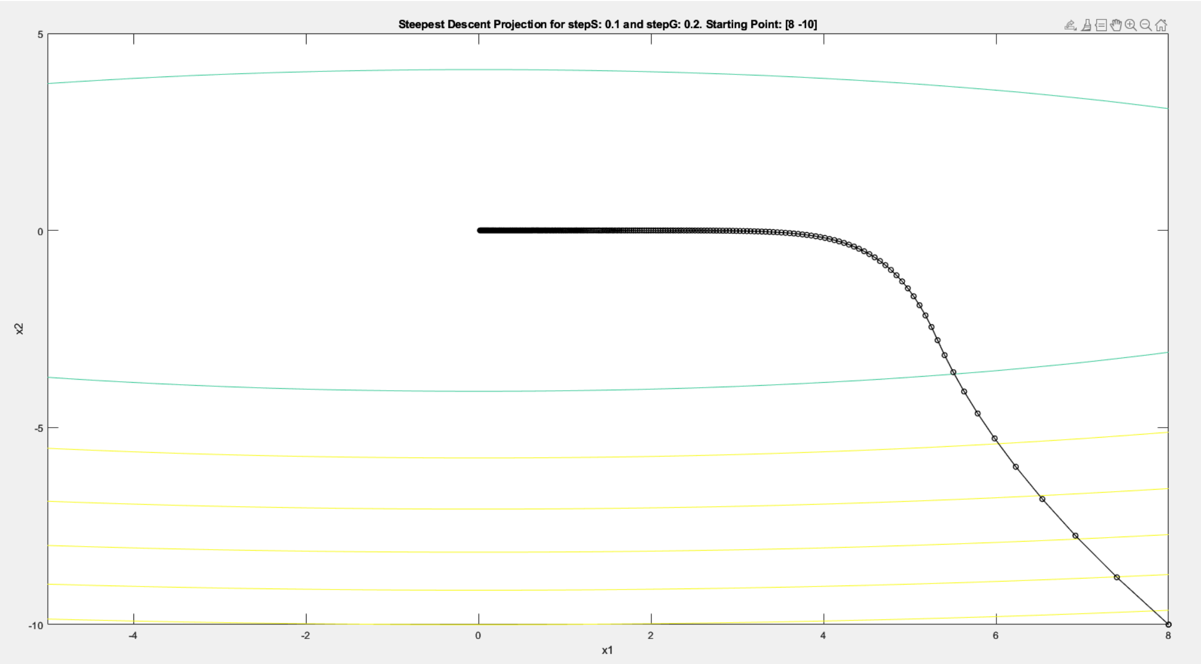


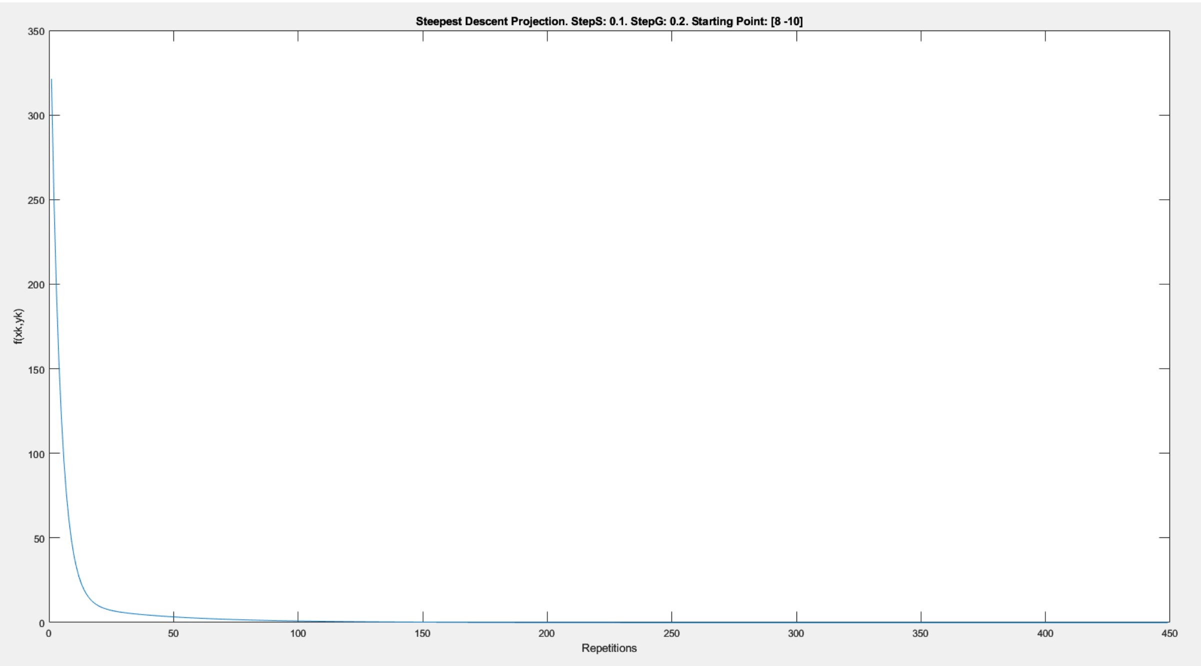


Σε αυτήν την περίπτωση, λόγω του μεγάλου βήματος κάνει λίγες επαναλήψεις αλλά δεν μπορεί να φτάσει εύκολα κοντά στο ελάχιστο, οπότε είναι σαν να ταλαντώνεται το x_2 κοντά στο 0. Με μείωση του s_k μπορεί να λυθεί αυτό, αλλά θα αυξηθούν οι επαναλήψεις.

ΘΕΜΑ 4

Στο Θέμα 3 χρησιμοποιούμε την μέθοδο μέγιστης καθόδου με προβολή με s_k= 0.1, γ_k = 0.2, ε = 0.01 και αρχικό σημείο εκκίνησης του αλγορίθμου το (8,-10). Λαμβάνοντας υπόψιν και τα παραπάνω θα είναι $\gamma'_k = \gamma_k \cdot s_k = 0.2 \cdot 0.1 = 0.02$ δηλαδή $\gamma'_k < 0.33$ και θα αναμένουμε να συγκλίνει η μέθοδος μας με αυτές τις αρχικές συνθήκες.





΄Οπως αναμέναμε η μέθοδος συγκλίνει. Αυτό οφείλεται στην ύπαρξη του s_k το οποίο έχει μικρή τιμή επιβραδύνοντας έτσι την σύγκλιση του αλγορίθμου.

ΣΥΜΠΕΡΑΣΜΑ

Για μεγάλα βήματα, η μέθοδος μέγιστης καθόδου με προβολή βοηθάει στο να μην ξεφύγει στο άπειρο η συνάρτηση, αλλά δεν συγκλίνει. Για μικρά βήματα δουλεύει επιθυμητά ο αλγόριθμος. Επίσης όσο μεγαλύτερο s_k έχουμε τόσο πιο γρήγορα συγκλίνει ο αλγόριθμος.