Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης

Πολυτεχνική σχολή

Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

ΤΕΧΝΙΚΕΣ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ

ΑΝΑΦΟΡΑ ΤΡΙΤΗΣ ΕΡΓΑΣΙΑΣ:

Ημερομηνία 5/12/2024

Λιαροπούλου Κλεοπάτρα (ΑΕΜ : 10066, Email : liaropou@ece.auth.gr)

ΠΕΡΙΕΧΟΜΕΝΑ : <u>Σελίδα</u>

1. Εισαγωγή:

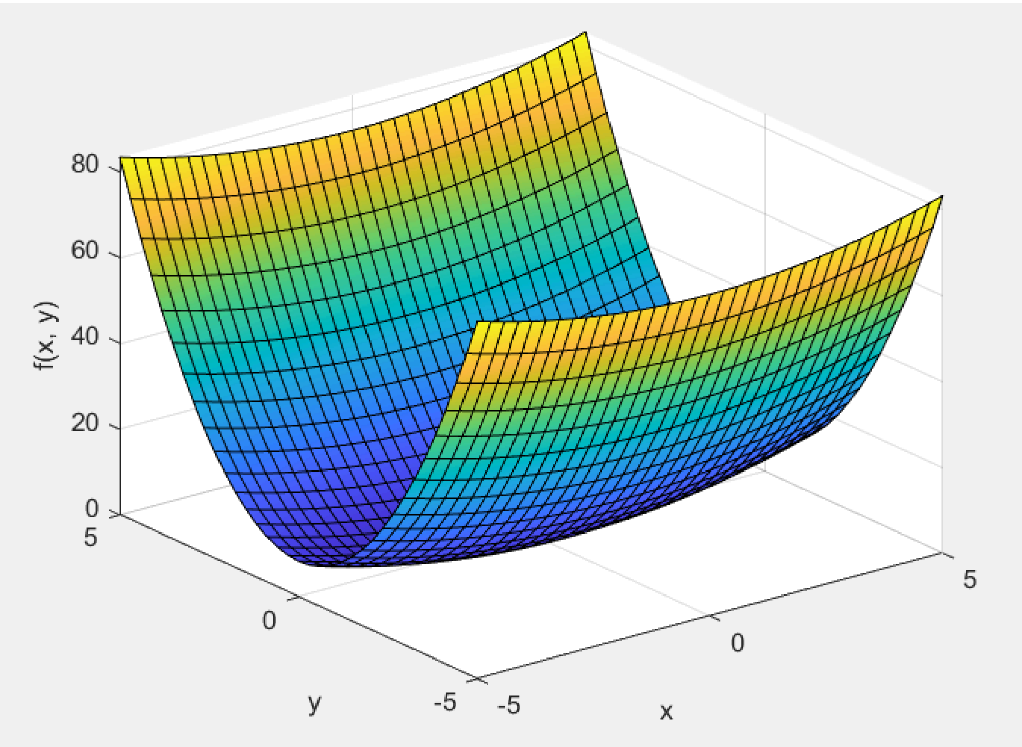
Στην προκειμένη εργασία θα παρουσιαστεί η σύγκριση των αποτελεσμάτων των εξής αλγορίθμων ελαχιστοποίησης:

- Μέθοδος μεγιστης καθόδου
- Μέθοδος μεγιστης καθόδου με προβολή

Η γραφική παράσταση που θα μελετηθεί είναι η:

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

Με τους εξής περιορισμούς: $-10 \leq x \leq 5$, $-8 \leq y \leq 12$ της οποίας η γραφική παράσταση απεικονίζεται στο *σχημα 1*:



Σχήμα 1: Γραφική παράσταση της f(x,y)

Για την $f(x, y)$ ισχύουν τα εξής:

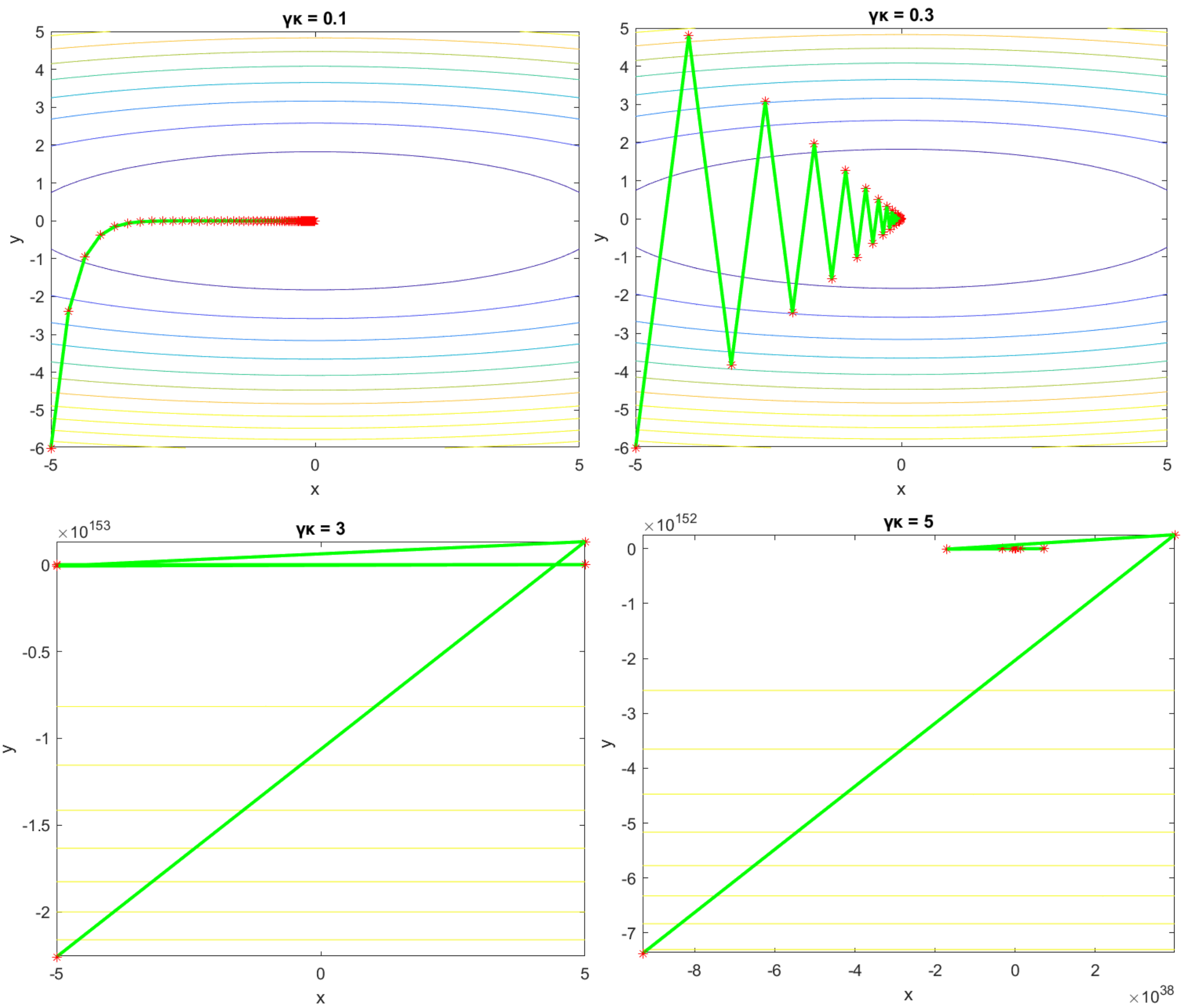
- Ο εσσιανός $H = \nabla^2 f(x, y) = \begin{bmatrix} \frac{2}{3} & 0 \\ 0 & 6 \end{bmatrix}$ έχει det=4≠0 καθώς και θετικές ιδιοτιμές (τα στοιχεία της διαγωνίου). Συνεπως, ο εσσιανός είναι θετικά ορισμένος γεγονός που καθιστά την $f(x, y)$ γνησίως κυρτή με μοναδικό ολικό ελάχιστο.
- Το ελάχιστο μπορεί να υπολογιστεί από το $\nabla f(x, y) = 0 \Rightarrow \begin{bmatrix} \frac{2}{3}x \\ 6y \end{bmatrix} = 0.$

Άρα το (0,0) είναι το μοναδικό ολικό ελάχιστο της συνάρτησης το οποίο επίσης ανήκει στους περιορισμούς των συντεταγμένων που δόθηκαν.

2. Μέθοδος μέγιστης καθόδου

Εκτελούμε την μέθοδο μέγιστης καθόδου για τις εξής τιμές του $\gamma_k = 0.1, 0.3, 3, 5$:

Σημείο εκκίνησης (-5,-6):



Συμπέρασμα: Γενικότερα παρατηρούμε ότι έχουμε επιτυχή αποτελέσματα σύγκλισης μόνο για τις δύο πρώτες περιπτώσεις (γ_k = 0.1,γ_k =0.3). Για την μέθοδο της μέγιστης καθόδου για την προκειμένη συναρτηση $f(x) = \frac{1}{3}x^2 + 3y^2$ ισχύουν οι εξής σχέσεις:

$$x_{k+1} = x_k - \frac{2}{3}\gamma_k x_k \quad (1) \qquad\qquad y_{k+1} = y_k - 6\gamma_k y_k \quad (2)$$

Γενικότερα σκοπός είναι να ικανοποιείται η εξής σχέση:

- $f(x_{k+1}, y_{k+1}) < f(x_k, y_k) \quad (3)$

Έπειτα από αντικατάσταση της (1),(2) στην (3) προκύπτει το εξής:

- $\frac{1}{3}\, x_{k+1}^{\ \ 2} + 3\, y_{k+1}^{\ \ 2} < \frac{1}{3}\, x_k^{\ 2} + 3y_k^{\ 2} \Rightarrow$

  $\frac{1}{3}\, x_k^{\ 2}(1 - \frac{2}{3}\gamma_k)^2 + 3y_k^{\ 2}((1 - 6\gamma_k^{\ 2}) < \frac{1}{3}\, x_k^{\ 2} + 3y_k^{\ 2} \Rightarrow$

  $\frac{1}{3}\, x_k^{\ 2}((1 - \frac{2}{3}\gamma_k)^2 - 1) + 3y_k^{\ 2}((1 - 6\gamma_k^{\ 2}) - 1) < 0$

  Άρα πρέπει:

  - $((1 - \frac{2}{3}\gamma_k)^2 - 1) < 0 \Rightarrow 0 < \gamma_k < 3$
  - $((1 - 6\gamma_k)^2 - 1) < 0 \Rightarrow 0 < \gamma_k < \frac{1}{3}$

Από την παραπάνω ανάλυση προκύπτει το συμπέρασμα ότι για να έχουμε σίγουρα επιτυχή σύγκλιση πρέπει για το βήμα να ισχύει: $0 < \gamma_k < \frac{1}{3}$

Βλέπουμε, λοιπόν, ότι η τιμή του βήματος γk πρέπει να βρίσκεται μεταξύ 0 και 0.333 για να εξασφαλιστεί η σύγκλιση της μεθόδου μας. Αυτό επιβεβαιώνεται και από τα αποτελέσματα που προκύπτουν από τα γραφήματα που αναφέραμε παραπάνω.

Συγκεκριμένα, για την τιμή γk=0.3, η ταλάντωση που παρατηρήσαμε μπορεί να εξηγηθεί από το γεγονός ότι η τιμή αυτή βρίσκεται στο όριο του εύρους τιμών για το οποίο η μέθοδος συγκλίνει. Από την άλλη, για τις τιμές γk στις περιπτώσεις (iii) και (iv), η μεγάλη απόκλιση που παρατηρείται οφείλεται στο ότι οι τιμές αυτές απέχουν πολύ από το διάστημα τιμών που έχουμε καθορίσει για τη σύγκλιση της μεθόδου.

Παρατηρηση: για την (iv) περιπτωση φαίνεται να προσεγγίζεται το ελάχιστο το οποίο είναι αντίθετο με το αναμενόμενο από την θεωρητική ανάλυση. Αν η μέθοδος χρησιμοποιεί ένα πολύ μεγάλο βήμα (όπως γk=5), μπορεί να κάνει πολύ μεγάλες κινήσεις στην κατεύθυνση του αρνητικού βαθμωτού, διασχιζοντας το αρκετές φορές, με αποτέλεσμα να προκαλεί την **αναγνώριση μιας προσεγγιστικής λύσης** που τελικά οδηγεί σε συγκέντρωση κοντά στο τοπικό ελάχιστο, αλλά χωρίς να ακολουθεί την θεωρητική κατεύθυνση και να παρουσιάζει μια φαινομενική σύγκλιση.

3.Μέθοδος μέγιστης καθόδου με προβολή

Η **μέθοδος Μέγιστης Καθόδου με προβολή** χρησιμεύει όταν το πρόβλημα περιλαμβάνει **περιορισμούς**. Για παράδειγμα, αν το διάνυσμα x_k πρέπει να παραμείνει εντός ενός κυρτού συνόλου, η μέθοδος της μέγιστης καθόδου δεν είναι επαρκής για να διασφαλίσει ότι οι λύσεις παραμένουν εντός αυτού του συνόλου. Όμως, η μέθοδος **Μέγιστης Καθόδου με Προβολή** όταν φτάνει σε ένα σημείο που δεν είναι εφικτό (δηλαδή εκτός του κυρτού συνόλου), τότε υπολογίζεται η προβολή του σημείου αυτού στο κυρτό σύνολο, και η διαδικασία συνεχίζεται με αυτό το νέο εφικτό σημείο.
Για να συγκλίνει η μέθοδος στο ελάχιστο βοηθά η κατάλληλη επιλογή των παραμέτρων s_k και γ_k οι οποίες προκύπτουν με αντιστοιχη συλλογιστική πορεία με την ανάλυση της παραγράφου 2 συνυπολογίζοντας τα επιπρόσθετα χαρακτηριστικά της μεθόδου μέγιστης καθόδου με χρηση προβολής.
Ξεκινώντας λοιπόν από την σχέση:

- $f\,(x_{k+1}\,,\,y_{k+1}) < f\,(x_k, y_k)$

καταλήγουμε στους εξής περιορισμούς:

- $((1 - \frac{2}{3}\,s_k\gamma_k)^2 - 1) < 0 \Rightarrow 0 < s_k\,\gamma_k < 3$
- $((1 - 6\,s_k\gamma_k)^2 - 1) < 0 \Rightarrow 0 < s_k\,\gamma_k < \frac{1}{3}$

Συνεπώς, προκειμένου να έχουμε σύγκλιση πρέπει για το γινόμενο να ισχύει το εξής: $s_k\,\gamma_k < \frac{1}{3}$

Για τα εφικτά σημεία x_k, η διαδικασία ακολουθεί την κλασική μέθοδο της μέγιστης καθόδου, και το βήμα γ'_k προσαρμόζεται ως $\gamma'_k = \gamma_k\, s_k$, όπου s_k είναι ένας παράγοντας που διασφαλίζει ότι οι περιορισμοί του προβλήματος παραμένουν ικανοποιημένοι. Εδώ, οι περιορισμοί για το βήμα γ'_k είναι οι ίδιοι με αυτούς της πρώτης περίπτωσης, δηλαδή $0 < \gamma'_k < \frac{1}{3}$

3.1: Σημείο εκκίνησης (5,-5)

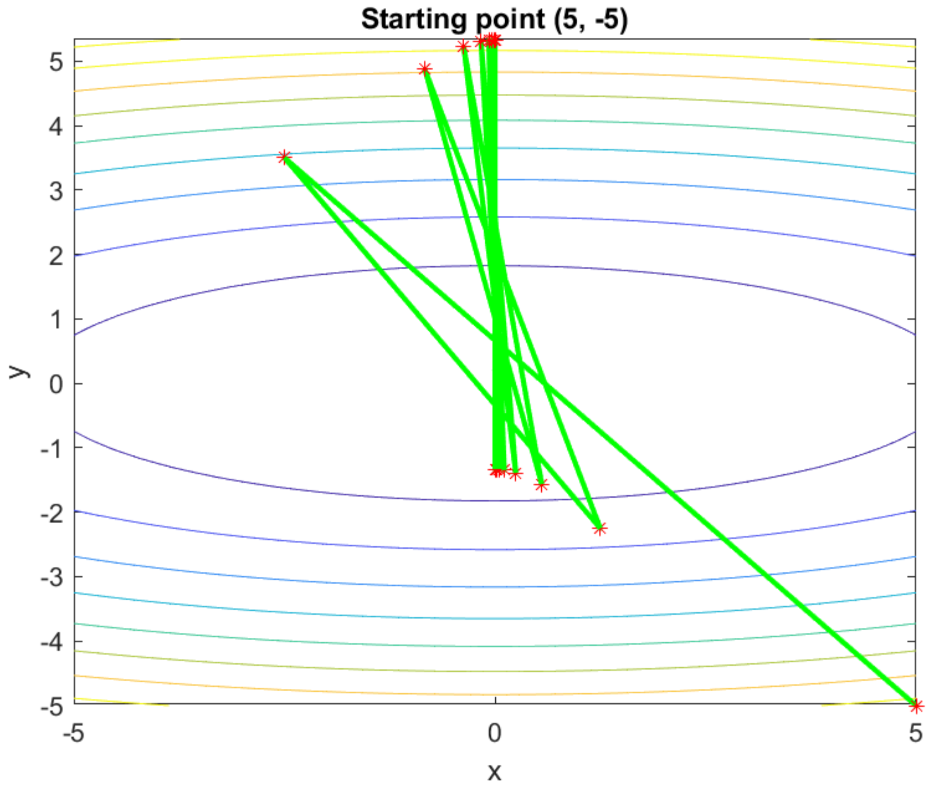
Εφαρμόζουμε την μέθοδο της μέγιστης καθόδου με προβολή για σημείο εκκίνησης (5,-5) και χαρακτηριστικά:

- $s_k = 5$
- $\gamma_k = 0.5$
- $\varepsilon = 0.01$

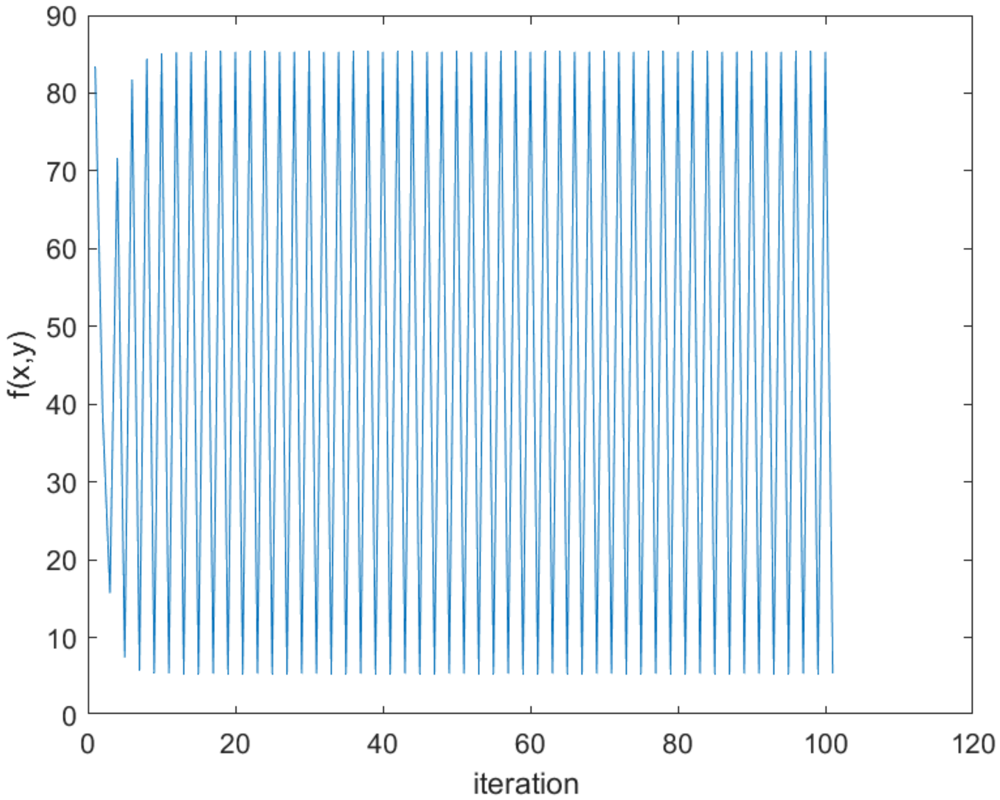
Παρατηρούμε ότι **η μέθοδος αποκλίνει** και το ελάχιστο δεν προσεγγίζεται. Συγκεκριμένα, παρατηρείται **ταλάντωση** μεταξύ δύο συγκεκριμένων σημείων όπως φαίνεται και στο σχημα 3. Λύση σε αυτό το πρόβλημα πιθανώς να αποτελούσε η επιλογή βήματος με τον κανόνα Armijo ώστε το βήμα να μην είναι σταθερό σε κάθε επανάληψη και να μπορέσουμε να διαφύγουμε από την ταλάντωση που εγκλωβιζόμαστε.

Στο σχήμα 2 φαίνεται αυτό το φαινόμενο καθώς όλα τα σημεία είναι συγκεντρωμένα (εντονα κόκκινα) και δεν εχουν διασκορπιστει στο χώρο σε προσπάθεια εντοπισμού του ελαχίστου. Το γεγονός οτι αποκλίνει είναι αναμενόμενο καθώς λόγω των αρχικών συνθηκών ισχύει ότι:

γ'_k = 5*0.5 = 2.5 > 0.333 το οποίο δεν πληροί τους περιορισμούς.



Σχήμα 2: Μέθοδος μέγιστης καθόδου με προβολή



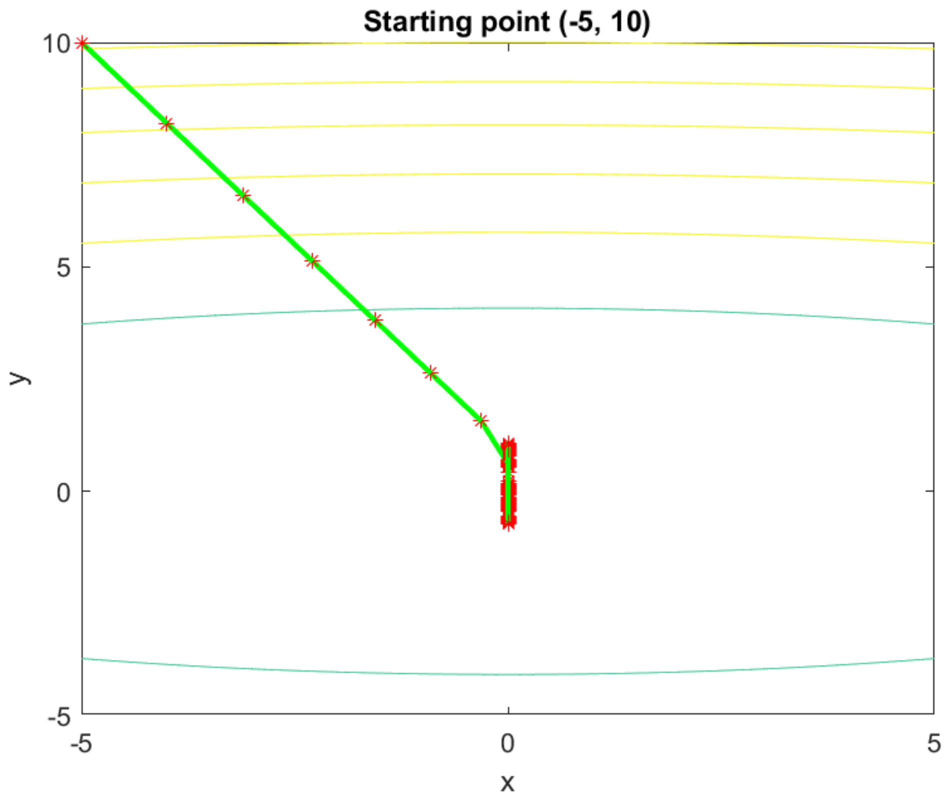
Σχήμα 3: Μέθοδος μέγιστης καθόδου με προβολή - τιμή συνάρτησης κατα τις επαναλήψεις

3.2: Σημείο εκκίνησης (-5,10)

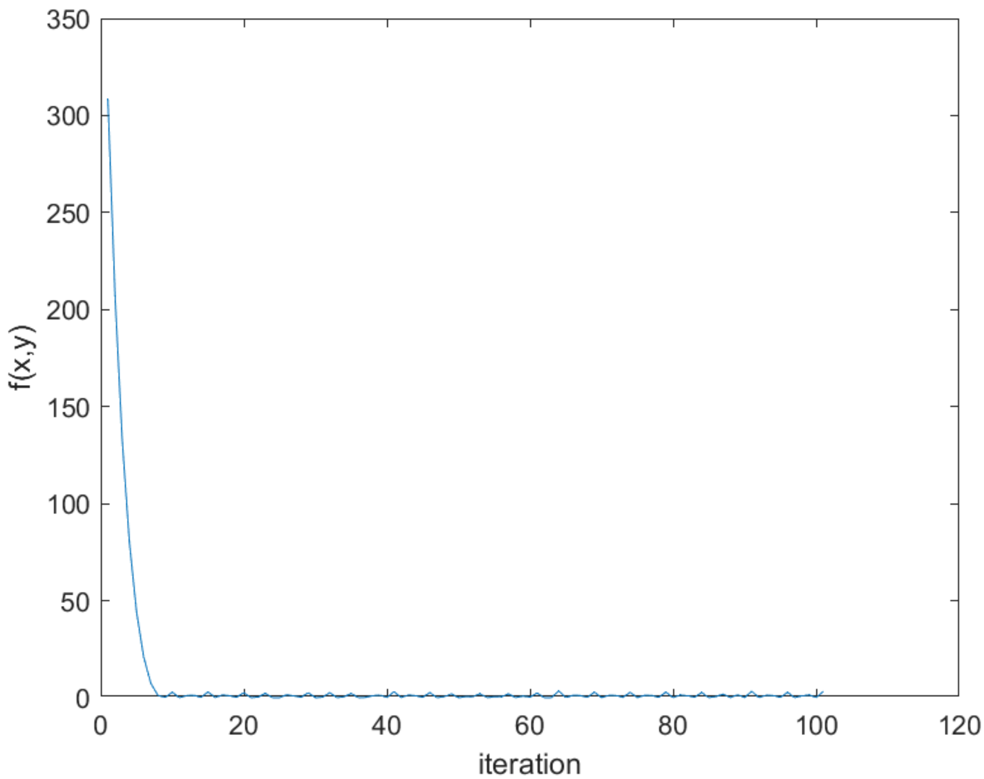
Εφαρμόζουμε την μέθοδο της μέγιστης καθόδου για σημείο εκκίνησης (-5,10) και χαρακτηριστικά:

- $s_k = 15$
- $\gamma_k = 0.1$
- $\varepsilon = 0.01$

Παρατηρούμε σε αυτή τη περίπτωση οτι **η μέθοδος συγκλίνει λίγο περισσότερο από πρίν**. Ωστόσο, είναι διακριτή ξανά μία **ταλάντωση** γύρω από το ελάχιστο η οποία όμως έχει μικρότερο εύρος διαστήματος από πρίν και είναι διακριτή στο σχήμα 5. Επίσης,η ταλάντωση γύρω από το (0,0) είναι εμφανής και στο σχήμα 4 όπου συσσωρεύονται τα σημεία.
Για αυτές τις αρχικές συνθήκες, αν γίνει αντίστοιχη ανάλυση των περιορισμών προκύπτει ότι: γ'_k= 15*0.1= 1.5 > 0.333 το οποίο και πάλι δεν είναι εντός του επιθυμητού εύρους. Ωστόσο, επειδή είναι πίο κοντά στο όριο του $\gamma'_k < \frac{1}{3}$ παρατηρείται μικρότερη ταλάντωση γύρω από το ελάχιστο.
Μια πιθανή λύση για σύγκλιση, εκτός από την κατάλληλη επιλογή των s_k, γ_k , ίσως είναι και πάλι η επιλογή βήματος βάσει του κανόνα Armijo ώστε να μεταβάλλεται το μέγεθος του βήματος κάθε φορά και πιθανώς να περιορίζεται σταδιακά η ταλάντωση.



Σχήμα 4: Μέθοδος μέγιστης καθόδου με προβολή



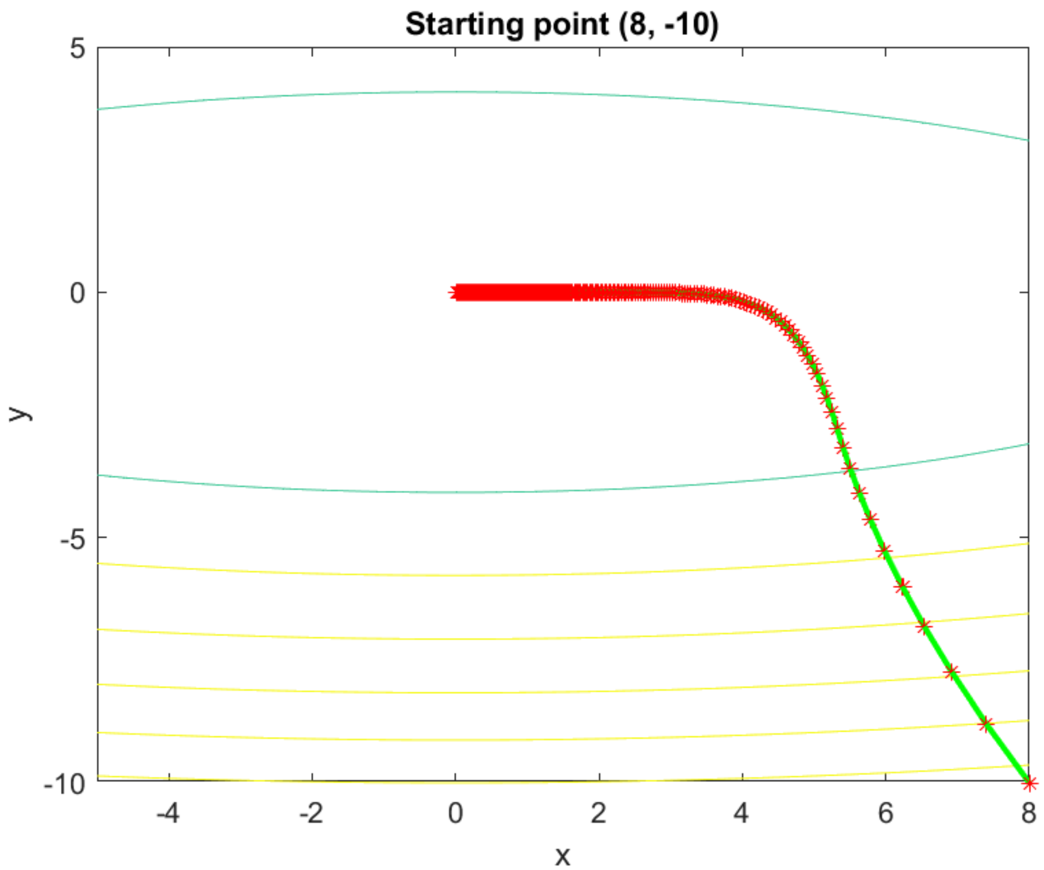
Σχήμα 5: Μέθοδος μέγιστης καθόδου με προβολή - τιμή συνάρτησης κατα τις επαναλήψεις

3.3: Σημείο εκκίνησης (8,-10)

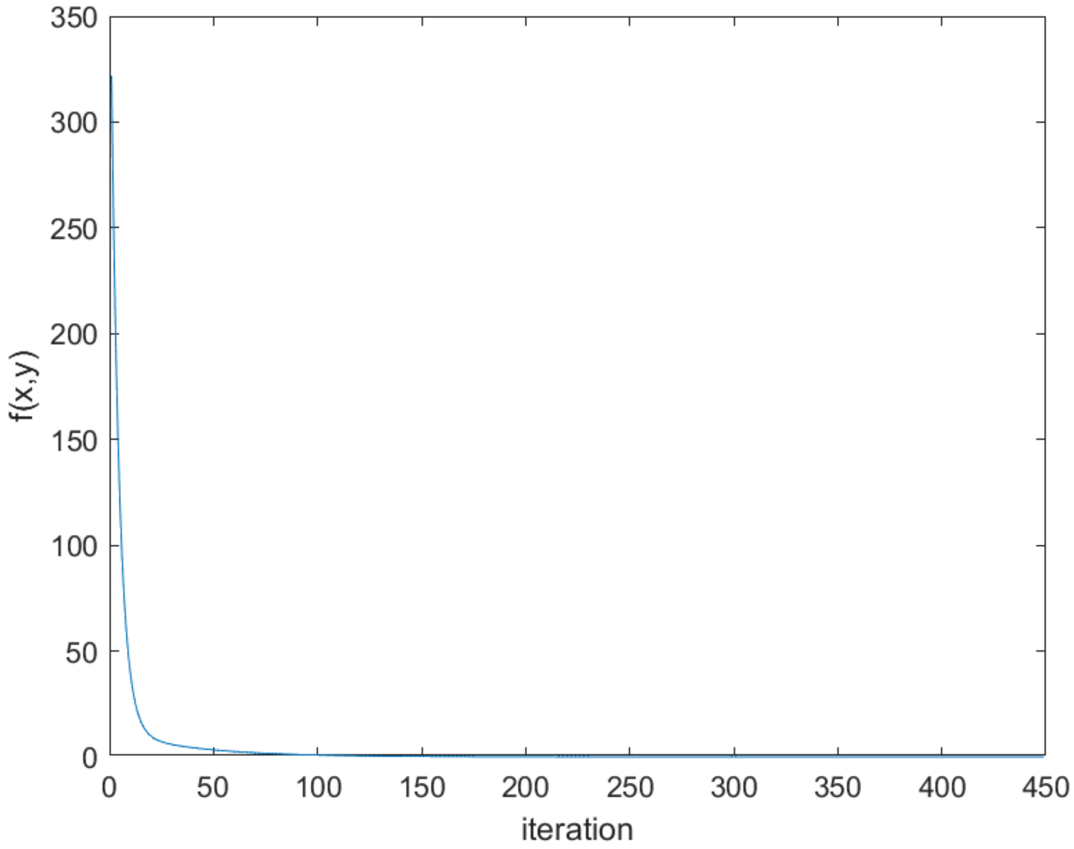
Εφαρμόζουμε την μέθοδο της μέγιστης καθόδου για σημείο εκκίνησης (8,-10) και χαρακτηριστικά:

- $s_k = 0.1$
- $\gamma_k = 0.2$
- $\varepsilon = 0.01$

Σε αυτη της περίπτωση έχουμε ως σημείο εκκίνησης ένα μη εφικτό σημείο σε αντίθεση με τις προηγούμενες, στις οποίες τα σημεία εκκίνησης του αλγορίθμου ανήκουν στο πεδίο ορισμού. Το γεγονός ότι χρησιμοποιούμε την μέθοδο με προβολή μας επιτρέπει ακριβως να **επιστρέψουμε στο πεδίο ορισμού**. Παρατηρούμε ότι **η μέθοδος συγκλίνει** και μάλιστα με μεγαλύτερη ακρίβεια από τις υπόλοιπες περιπτώσεις. Αυτό είναι αναμενόμενο καθώς σύμφωνα με την προηγούμενη ανάλυση ισχύει $s_k \gamma_k = 0.1 * 0.2 = 0.02 < \frac{1}{3}$



Σχήμα 6: Μέθοδος μέγιστης καθόδου με προβολή



Σχήμα 7: Μέθοδος μέγιστης καθόδου με προβολή - τιμή συνάρτησης κατα τις επαναλήψεις

4.Βιβλιογραφικές πηγές:

[1] Γεώργιος Α. Ροβιθάκης. Τεχνικές Βελτιστοποίησης. Εκδόσεις ΤΖΙΟΛΑ.