



Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης

Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών  
Υπολογιστών

Τομέας Ηλεκτρονικής και Υπολογιστών

## Τεχνικές Βελτιστοποίησης-Εργασία 3

Λώλος Ιωάννης 10674

4 Δεκεμβρίου 2024

# Περιεχόμενα

<b>1</b>	<b>Εισαγωγή</b>	<b>2</b>
<b>2</b>	<b>Θεωρητική Μελέτη</b>	<b>3</b>
2.1	Η Συνάρτηση . . . . .	3
2.2	Μαθηματική μελέτη σύγκλισης της μεθόδου . . . . .	4
<b>3</b>	<b>Θέμα 1</b>	<b>5</b>
<b>4</b>	<b>Η μέθοδος Μέγιστης Καθόδου με Προβολή</b>	<b>8</b>
4.1	Βασικό Θεωρητικό Πλαίσιο . . . . .	8
4.2	Μεθοδολογία . . . . .	8
<b>5</b>	<b>Θέματα 2, 3, 4</b>	<b>9</b>

# Εισαγωγή

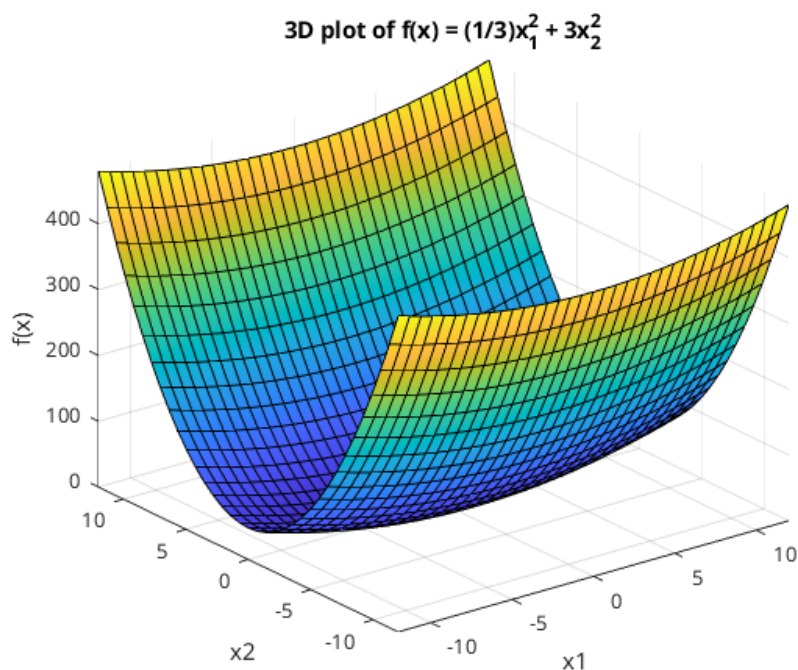
Ζητούμενο της παρούσας εργασίας ήταν η περαιτέρω μελέτη της μεθόδου Μέγιστης Καθόδου, όσον αφορά την διαφοροποίηση του βήματος  $\gamma$ , καθώς και η υλοποίηση της μεθόδου Μέγιστης Καθόδου με Προβολή στο περιβάλλον Matlab, για την μελέτη της ελαχιστοποίησης συναρτήσεων με περιορισμούς.

# Θεωρητική Μελέτη

## 2.1 Η Συνάρτηση

Η συνάρτηση προς μελέτη αυτή τη φορά ήταν η:

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, f(x) = \frac{1}{3} \cdot x_1^2 + 3 \cdot x_2^2 \quad (2.1)$$



Σχήμα 2.1: Τρισδιάστατη αναπαράσταση της  $f(x)$

Από την γραφική παράσταση μπορούμε να έχουμε μία ιδέα της μορφής της συνάρτησης και να παρατηρήσουμε ότι εμφανίζει ελάχιστο στο  $x^* = (0, 0)$  το  $f(x^*) = 0$ . Η συνάρτηση είναι επίσης προφανώς κυρτή.

## 2.2 Μαθηματική μελέτη σύγκλισης της μεθόδου

Για την εύρεση των τιμών του  $\gamma$ , για τις οποίες η μέθοδος Μέγιστης καθόδου συγκλίνει, αρχικά θα υπολογίσουμε το gradient της συνάρτησης.

$$\nabla f(x_1, x_2) = \left[ \frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \right]^T = \left[ \frac{2}{3} \cdot x_1 \quad 6 \cdot x_2 \right]^T \quad (2.2)$$

Επομένως:

$$\nabla f(x_{1k}, x_{2k}) = \left[ \frac{2}{3} \cdot x_{1k} \quad 6 \cdot x_{2k} \right]^T$$

Δηλαδή για τη μέθοδο ισχύει:

$$\begin{aligned} \begin{bmatrix} x_{1_{k+1}} & x_{2_{k+1}} \end{bmatrix}^T &= \begin{bmatrix} x_{1k} & x_{2k} \end{bmatrix}^T - \gamma_k \cdot \left[ \frac{2}{3} \cdot x_{1k} \quad 6 \cdot x_{2k} \right]^T \Leftrightarrow \\ \begin{cases} x_{1_{k+1}} = x_{1k} - \gamma_k \cdot \frac{2}{3} \cdot x_{1k} \\ x_{2_{k+1}} = x_{2k} - \gamma \cdot 6 \cdot x_{2k} \end{cases} &\Leftrightarrow \begin{cases} x_{1_{k+1}} = (1 - \frac{2}{3}\gamma_k)x_{1k} \\ x_{2_{k+1}} = (1 - 6\gamma)x_{2k} \end{cases} \end{aligned} \quad (2.3)$$

Εφόσον γνωρίζουμε ότι  $x^* = (0, 0)$  θέλουμε για  $k \rightarrow \infty$ :

$$\begin{cases} x_{1k} \rightarrow 0 \\ x_{2k} \rightarrow 0 \end{cases}$$

Άρα:

$$|1 - 6\gamma| < 1 \Leftrightarrow -1 < 1 - 6\gamma < 1 \Leftrightarrow 0 < \gamma < \frac{1}{3} \quad (2.4)$$

Βλέπουμε δηλαδή ότι η σύγκλιση της μεθόδου Μέγιστης Καθόδου για τη συγκεκριμένη συνάρτηση απαιτεί τιμές του  $\gamma$  μικρότερες του  $\frac{1}{3}$

# Θέμα 1

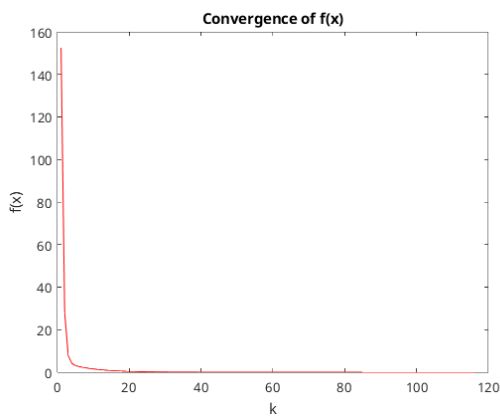
Στο θέμα 1 έπρεπε να ελαχιστοποιήσουμε την  $f$  χρησιμοποιώντας την μέθοδο Μέγιστης Καθόδου που υλοποιήσαμε στο δεύτερο παραδοτέο. Η ελαχιστοποίηση έπρεπε να γίνει με ακρίβεια  $e = 0.001$  για κάθε ένα από τα παρακάτω βήματα  $\gamma$ :

- $\gamma = 0.1$
- $\gamma = 0.3$
- $\gamma = 3$
- $\gamma = 5$

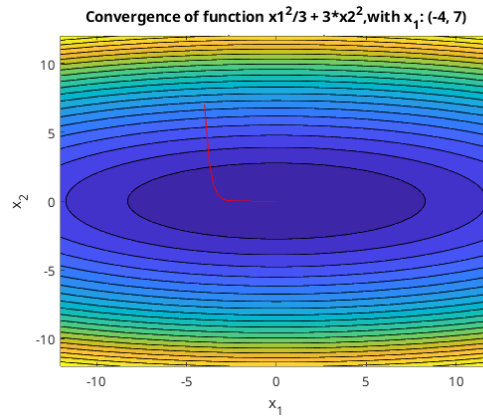
Το αρχικό σημείο που επιλέχθηκε ήταν το  $(-4, 7)$ .

Για  $\gamma = 0.1$ :

Η μέθοδος βρήκε ελάχιστο το 0 στο σημείο  $(-0.001, 0)$  έπειτα από 116 επαναλήψεις.



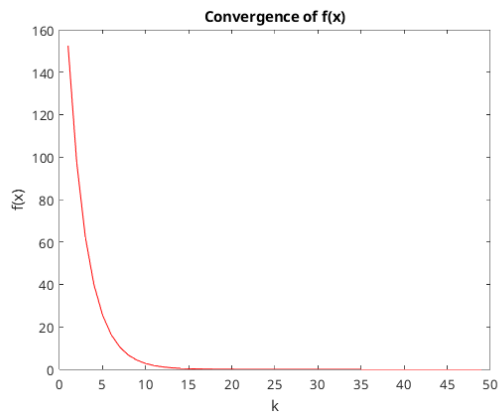
(α') Convergence in 1D space



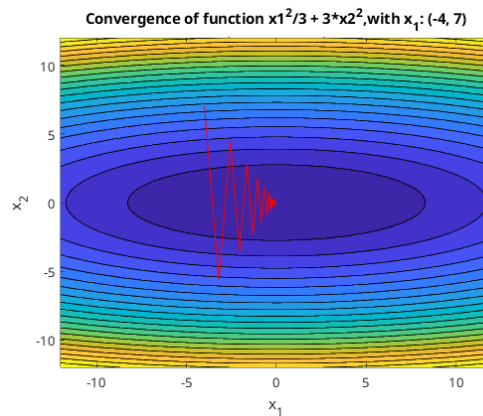
(β') Convergence in 2D space

Για  $\gamma = 0.3$ :

Η μέθοδος βρήκε ελάχιστο το 0 στο σημείο  $(0, 0)$  έπειτα από 49 επαναλήψεις.



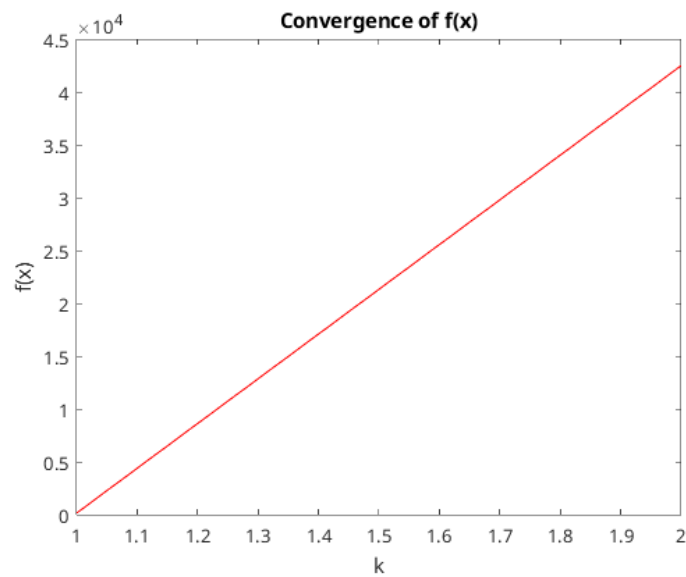
(α') Convergence in 1D space



(β') Convergence in 2D space

Για  $\gamma = 3$ :

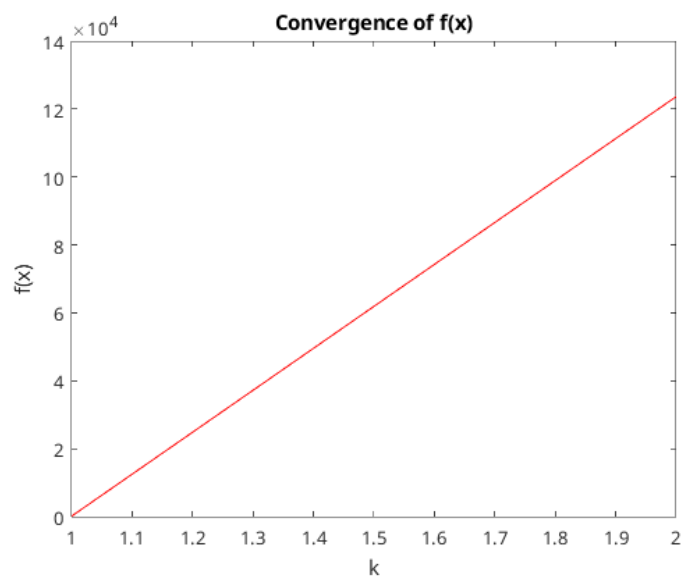
Όπως ήταν αναμενόμενο από την ανάλυση που έγινε στο section 2.2 η μέθοδος δεν συνέκλινε. Η τιμή του  $x$  απομακρύνθηκε από το  $(0, 0)$  και η μέθοδος τερμάτισε λόγω μίας συνθήκης που προστέθηκε στον κώδικα για να αποτρέψει το  $x$  από το να απομακρυνθεί αρκετά από το σημείο ελαχίστου που είχε παρατηρηθεί αρχικά



Σχήμα 3.1: Η μεταβολή του  $f(x_k)$  για  $\gamma = 3$

Για  $\gamma = 5$ :

Παρατηρήθηκαν παρόμοια αποτελέσματα όπως και στην παραπάνω περίπτωση.



Σχήμα 3.2: Η μεταβολή του  $f(x_k)$  για  $\gamma = 5$



# Η μέθοδος Μέγιστης Καθόδου με Προβολή

## 4.1 Βασικό Θεωρητικό Πλαίσιο

Η μέθοδος Μέγιστης Καθόδου με Προβολή (Maximum Descent with Projection) αποτελεί τεχνική βελτιστοποίησης που επιλύει προβλήματα περιορισμένης ελαχιστοποίησης. Αποτελεί τροποποίηση της απλής μεθόδου Μέγιστης Καθόδου, ώστε να επιτρέπει την εφαρμογή της σε προβλήματα με περιορισμούς. Η σύγκλιση της επηρεάζεται από παράγοντες όπως η επιλογή του βήματος  $\gamma$  και η κυρτότητα της συνάρτησης.

## 4.2 Μεθοδολογία

Η παρούσα μέθοδος "βρίσκει" τα νέα εφικτά σημεία με την εξής σχέση [1]:

$$x_{k+1} = x_k + \gamma_k \cdot (\bar{x} - x_k) \quad (4.1)$$

όπου

$$\bar{x} = Pr_x\{x_k - s_k \nabla f(x_k)\} \quad (4.2)$$

Το  $s_k$  επιλέγεται από εμάς, ενώ ο τελεστής  $Pr_x(\cdot)$  ουσιαστικά "προβάλλει" οποιοδήποτε σημείο εκτός των περιορισμών στο κοντινότερο του εντός των περιορισμών. Προφανώς τα σημεία που βρίσκονται ήδη μέσα στους περιορισμούς δεν επηρεάζονται από τον τελεστή.

Με τον παραπάνω τρόπο εγγυόμαστε πως η αναζήτηση ελαχίστου παραμένει πάντοτε εντός της περιοχής που επιβάλλουν οι περιορισμοί.

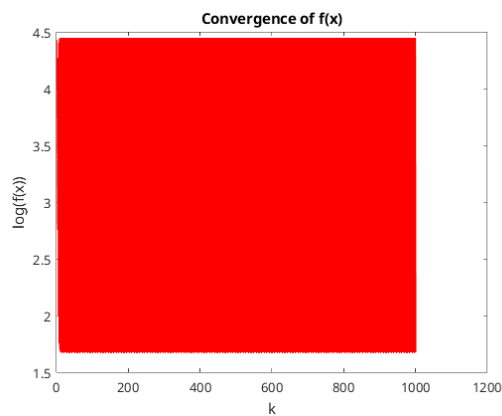
## Θέματα 2, 3, 4

Στα θέματα 2, 3, 4 κληθήκαμε να εφαρμόσουμε την μέθοδο Μέγιστης Καθόδου που υλοποιήσαμε για να βρούμε το ελάχιστο της γνωστής συνάρτησης  $f$ . Το παραπάνω έπρεπε να γίνει με ακρίβεια  $\epsilon = 0.01$  για κάθε μία από τις παρακάτω περιπτώσεις  $\gamma$ :

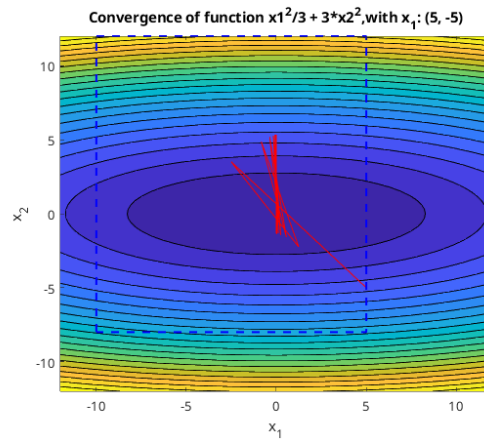
1.  $\gamma_k = 0.5, s_k = 5, x_0 = (5, -5)$
2.  $\gamma_k = 0.1, s_k = 15, x_0 = (-5, 10)$
3.  $\gamma_k = 0.2, s_k = 0.1, x_0 = (8, -10)$

Για την περίπτωση 1:

Η μέθοδος δεν συνέκλινε. Αντιθέτως το σημείο άρχισε να ταλαντώνεται γύρω από το σημείο ελαχίστου επ' αόριστον και η μέθοδος τερμάτισε μόνο όταν έφτασε τον μέγιστο αριθμό επαναλήψεων που είχε οριστεί ως δικλίδα ασφαλείας. Το αποτέλεσμα είναι λογικό, καθώς η ανάλυση του section 2.2 για το  $\gamma$  ισχύει και στην παρούσα περίπτωση, ωστόσο για το  $\gamma' = \gamma_k \cdot s_k$ , το οποίο στην περίπτωση μας ισούται με 2.5.



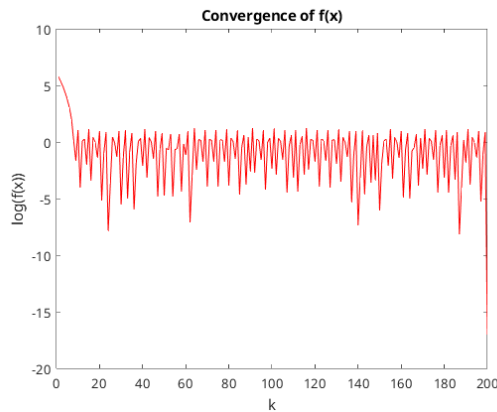
(α') Convergence in 1D space



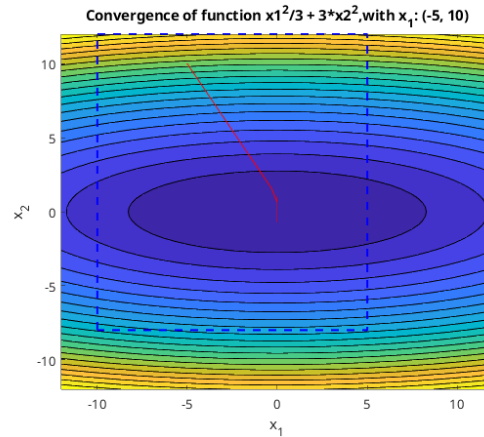
(β') Convergence in 2D space

Για την περίπτωση 2:

Η μέθοδος συνέκλινε και βρήκε ελάχιστο το 0 στο σημείο  $x^* = (0, 0)$ , έπειτα από 200 επαναλήψεις. Η μέθοδος ενδεχομένως συνέκλινε από τύχη (κάτι που παρατηρείται και στο μονοδιάστατο διάγραμμα σύγκλισης που δεν είναι ομαλό), καθώς θεωρητικά το  $\gamma_k \cdot s_k = 1.5$ , δηλαδή πάνω από 0.33. Ένας απλός τρόπος να σιγουρέψουμε την σύγκλιση θα ήταν η μείωση του  $s_k$  πχ.  $s_k = 3$ .



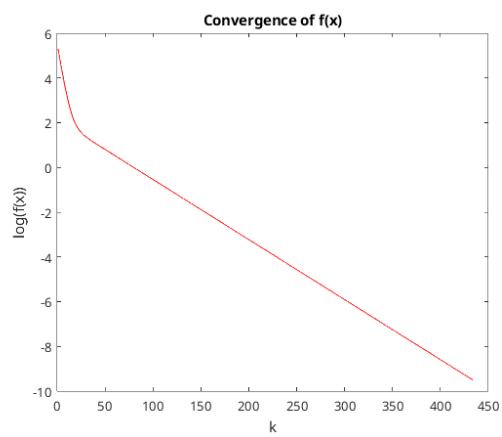
(α') Convergence in 1D space



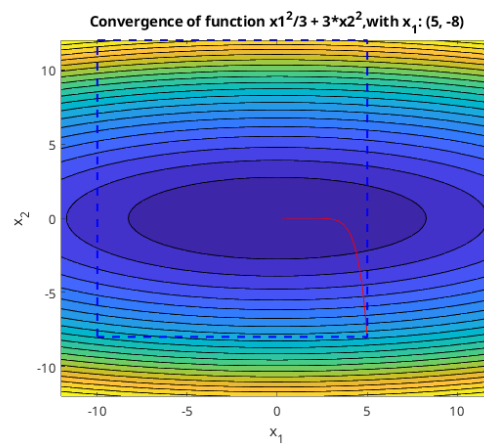
(β') Convergence in 2D space

Για την περίπτωση 3:

Παρατηρούμε ότι το αρχικό σημείο βρίσκεται εκτός περιορισμών. Θα εφαρμόσουμε τη μέθοδο επομένως με αρχικό σημείο την προβολή του, δηλαδή το  $x_0 = (5, -8)$ . (Παρόλο που λόγω περιέργειας η μέθοδος δοκιμάστηκε και στο αρχικό σημείο και συγκλίνει κανονικά). Η μέθοδος συνέκλινε και βρήκε ελάχιστο το 0 στο σημείο  $x^* = (0.015, 0)$ , έπειτα από 434 επαναλήψεις. Η σύγκλιση στο ελάχιστο γίνεται πολύ πιο αργά και ομαλά σε σχέση με την προηγούμενη περίπτωση, καθώς το  $s_k$  είναι 150 φορές μικρότερο, ενώ το  $\gamma$  μόλις δύο φορές μεγαλύτερο. Βλέπουμε λοιπόν πως η ταχύτητα σύγκλισης εξαρτάται και από την επιλογή της παραμέτρου  $s$ , πέραν του  $\gamma$ .



( $\gamma'$ ) Convergence in 1D space



( $\delta'$ ) Convergence in 2D space

# Bibliography

- [1] Γεώργιος Α. Ροβιθάκης. *Τεχνικές Βελτιστοποίησης*. ΕΚΔΟΣΕΙΣ Α. ΤΖΙΟΛΑ & ΥΙΟΙ Α.Ε, 2007.