Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
Πολυτεχνική Σχολή
Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Μέθοδος Μέγιστης Καθόδου με προβολή

Αμαραντίδου Ευθυμία
7ο Εξάμηνο
Α.Ε.Μ.: 9762

Τεχνικές Βελτιστοποίησης
3η Εργαστηριακή Άσκηση

20 Δεκεμβρίου 2021

Περιεχόμενα

1 Εισαγωγή **2**

1.1 Περιγραφή του προβλήματος 2

1.1.1 Μέδοθος Μέγιστης Καθόδου 2

1.1.2 Μέθοδος Μέγιστης Καθόδου με προβολή 6

2 Συμπεράσματα & Παρατηρήσεις **10**

Κεφάλαιο 1

Εισαγωγή

Στην παρακάτω αναφορά θα μελετηθεί η Μέθοδος Μέγιστης Καθόδου που αναπτύχθηκε στην Εργαστηριακή Άσκηση 2 και θα αποδειχθεί σε θεωρητικό επίπεδο, ενώ στη συνέχεια θα μελετηθεί η εύρεση ελαχίστου μιας συνάρτησης παρουσία περιορισμών. Η ανάλυση που ακολουθεί και οι μέθοδοι που αναφέρονται, έχουν υλοποιηθεί και εφαρμοστεί στο περιβάλλον του `Matlab`.

1.1 Περιγραφή του προβλήματος

Μεταβάλλοντας κάποιες βασικές **παραμέτρους** των μεθόδων **Μέγιστης Καθόδου** και **Μέγιστης Καθόδου με προβολή**, θα παρατηρήσουμε τη συμπεριφορά μιας συνάρτησης απουσία **και** παρουσία περιορισμών αντίστοιχα. Η αντικειμενική συνάρτηση που θα μελετηθεί είναι η εξής:

$$f(x) = \frac{1}{2}{x_1}^2 + 2x_2^2, \quad f : \mathbb{R}^2 \to \mathbb{R} \tag{1.1}$$

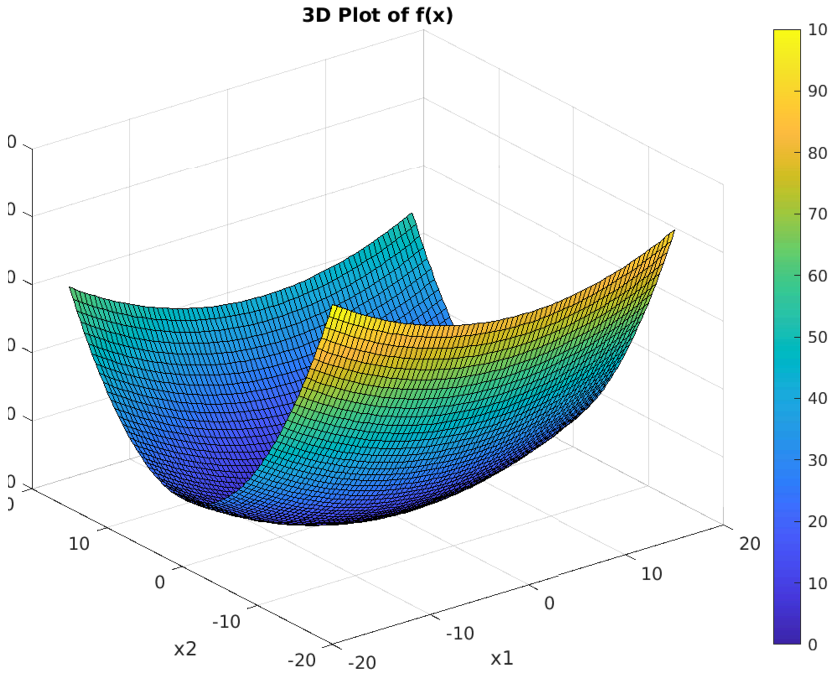
Μια καλύτερη εικόνα για τη μορφή της 1.1 μπορούμε να πάρουμε και από το σχήμα 1.1. Συμπεραίνουμε λοιπόν και οπτικά πως το ελάχιστο αυτής βρίσκεται στο (0,0) και είναι **ολικό.**

Στην ανάλυση που ακολουθεί θεωρούμε δεδομένο πως η αντικειμενική συνάρτηση είναι τ**ουλάχιστον μία φορά παραγωγίσιμη**.

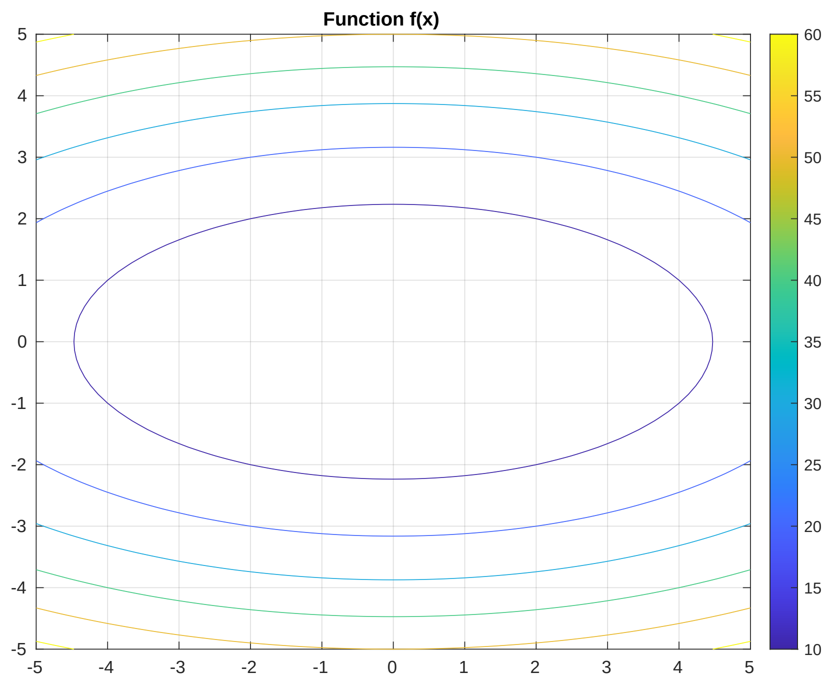
1.1.1 Μέδοθος Μέγιστης Καθόδου

Για την αντικειμενική συνάρτηση 1.1, χρησιμοποιήθηκε η Μέθοδος Μέγιστης Καθόδου που αναπτύχθηκε στην Εργαστηριακή Άσκηση 2, με στόχο την εύρεση ελαχίστου. Πιο συγκεκριμένα, χρησιμοποιήθηκαν **4 διαφορετικές τιμές** του $\gamma_k = [0.05, 0.5, 2, 10]$, ακρίβεια $\epsilon = 0.01$ και σημείο εκκίνησης (2, 4).

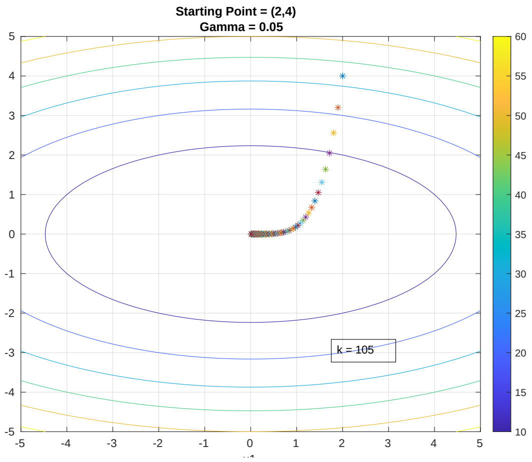
Οι εκτιμήσεις ελαχίστου που προέκυψαν φαίνονται στα γραφήματα 1.3, 1.4, 1.5 παρακάτω.



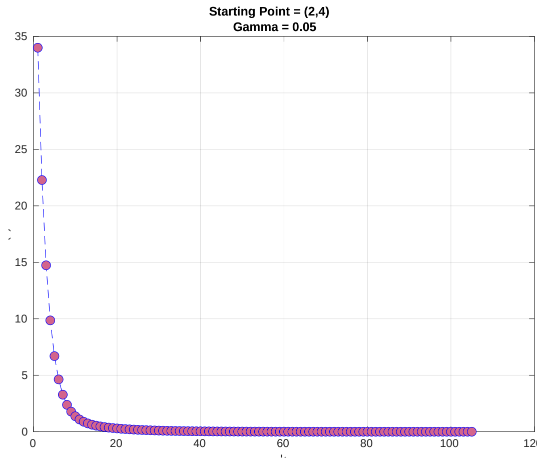
Σχήμα 1.1: Τρισδιάσταση απεικόνιση της f



Σχήμα 1.2: Ισοβαρείς καμπύλες της f

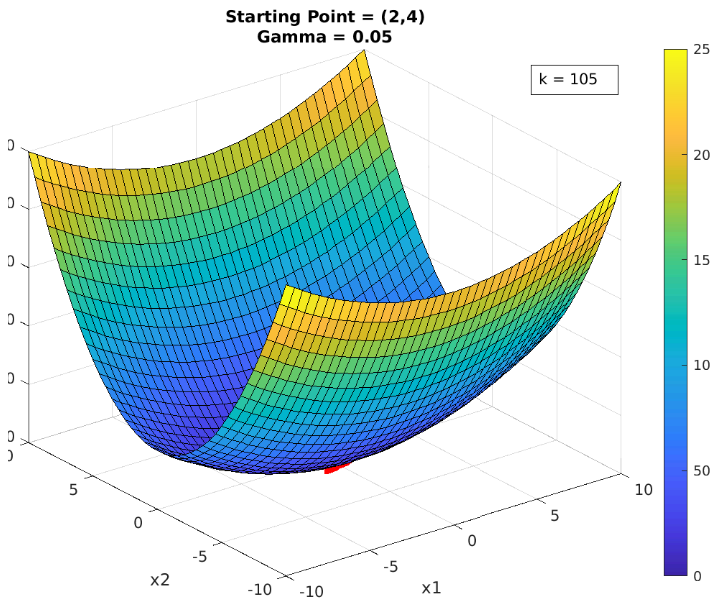


(α΄)

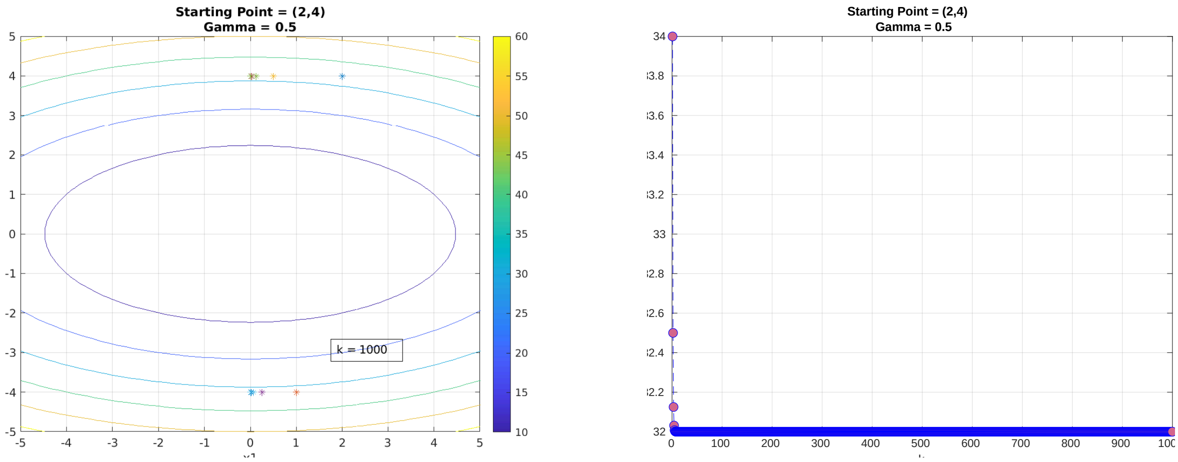


(β΄)

Σχήμα 1.3: Μέθοδος Μέγιστης Καθόδου για $\gamma_k = 0.05$



Σχήμα 1.4: Τρισδιάσταση απεικόνιση με τα διαδοχικά σημεία για $\gamma_k = 0.05$



(α′) (β′)

Σχήμα 1.5: Μέθοδος Μέγιστης Καθόδου για $\gamma_k = 0.5$

Παρατηρούμε πως για τιμές του γ_k **μεγαλύτερες από 0.5**, η μέθοδος **αδυνατεί** να προσεγγίσει το ελάχιστο. Συγκεκριμένα, για $\gamma_k = 0.5$ ο αλγόριθμος οδηγείται σε **ταλάντωση** σε τιμή κοντά στο 32 και τερματίζει όταν ξεπεράσει το όριο του μεγίστου αριθμού επαναλήψεων που έχουμε θέσει, ενώ για $\gamma_k = 2$ και $\gamma_k = 10$ η μέθοδος **δεν** επιτυγχάνει σύγκλιση και προσεγγίζει έναν πολύ μεγάλο αριθμό, γι΄ αυτό και **απορρίπτεται** μέσω του αλγορίθμου. Επιπλέον, κάτι που αξίζει να αναφέρουμε είναι ότι για $\gamma_k = 0.05$ χρειάζονται **πάνω από 100 επαναλήψεις** ώσπου να προσεγγίσουμε το ελάχιστο με την επιθυμητή ακρίβεια, το οποίο είναι λογικό δεδομένου της αρκετά μικρής τιμής βήματος.

Το παραπάνω αποτέλεσμα και τη μη σύγκλιση του αλγορίθμου για ορισμένες τιμές του βήματος γ_k μπορούν να επιβεβαιωθούν και **θεωρητικά**.

Σύμφωνα με τη θεωρία γνωρίζουμε ότι:

$$x_{k+1} = x_k - \gamma_k \nabla f(x_k) \tag{1.2}$$

,όπου γ_k το αντίστοιχο βήμα και:

$$\nabla f(\mathbf{x}) = [x_1, 4x_2] \tag{1.3}$$

Από την 1.2 θα έχουμε:

$$\begin{aligned}
\mathbf{x}_1 &= (x_{10}, x_{20}) - \gamma_k(x_{10}, 4x_{20})) = ((1-\gamma_k)x_{10}, (1-4\gamma_k x_{20}) && (1.4)\\
\mathbf{x}_2 &= (x_{11}, x_{21}) - \gamma_k(x_{11}, 4x_{21})) = ((1-\gamma_k)^2 x_{10}, (1-4\gamma_k)^2 x_{20}) && (1.5)\\
&\vdots && (1.6)\\
\mathbf{x}_N &= ((1-\gamma_k)^N x_{10}, (1-4\gamma_k)^N x_{20}) && (1.7)
\end{aligned}$$

Επομένως:

$$x_1 : \lim_{N\to\infty} |1-\gamma_k|^N |x_{1k}| = \begin{cases} 0 & \text{εάν } 0 < \gamma_k < 2 \\ \infty & \text{εάν } \gamma_k > 2 \end{cases} \tag{1.8}$$

$$x_2 : \lim_{N\to\infty} |1-4\gamma_k|^N |x_{1k}| = \begin{cases} 0 & \text{εάν } 0 < \gamma_k < 0.5 \\ \infty & \text{εάν } \gamma_k > 0.5 \end{cases} \tag{1.9}$$

Θα πρέπει όμως να βρούμε ένα γ_k τέτοιο ώστε να συγκλίνει τόσο η 1.8 όσο και η 1.9 εάν θέλουμε ο αλγόριθμος να καταλήγει επιτυχώς στο ελάχιστο της συνάρτησης. Άρα συμπεραίνουμε πως **το εύρος των επιτρεπτών τιμών βήματος** θα είναι:

$$0 < \gamma_k < 0.5 \tag{1.10}$$

το οποίο **επαληθεύεται** και από τα αποτελέσματα της ανάλυσης που πραγματοποιήσαμε νωρίτερα, όπου παρατηρήσαμε πως για τιμές βήματος μεγαλύτερες του 0.5 δεν οδηγούμασταν στο ελάχιστο.[1]

1.1.2 Μέθοδος Μέγιστης Καθόδου με προβολή

Η Μέθοδος Μέγιστης Καθόδου **με προβολή** εφαρμόζεται στις περιπτώσεις όπου λόγω της ύπαρξης περιορισμών χρειάζεται να διασφαλίσουμε πως **δε θα φύγουμε έξω από τα όρια του δυνατού συνόλου** X. Πρακτικά, η μέθοδος αυτή, κάθε φορά που ο αλγόριθμος μέγιστης καθόδου οδηγεί σε ένα **μη εφικτό σημείο**, υπολογίζει την **προβολή** του σημείου αυτού στο X [2] και συνεχίζει.

Τα διαδοχικά σημεία:

$$x_{k+1} = x_k + \gamma_k(\bar{x}_k - x_k), \quad \gamma_k \in (0, 1] \tag{1.11}$$

όπου

$$\bar{x}_k = Pr_X\{x_k - s_k \nabla f(x_k)\}, \quad s_k > 0 \tag{1.12}$$

Για την αντικειμενική συνάρτηση 1.1 που μελετάμε, θεωρούμε τους περιορισμούς:

$$-15 \le x_1 \le 15 \quad \text{και} \quad -20 \le x_2 \le 12 \tag{1.13}$$

Στη συνέχεια εφαρμόζουμε τη Μέθοδο Μέγιστης Καθόδου με προβολή για 3 διαφορετικές περιπτώσεις, οι παράμετροι των οποίων αναφέρονται αναλυτικά στον Πίνακα 1.1. Ο αλγόριθμος τερματίζει εάν καταλήξει σε στάσιμο σημείο ή εάν ξεπεράσει το μέγιστο αριθμό των 1000 επαναλήψεων που έχει οριστεί.

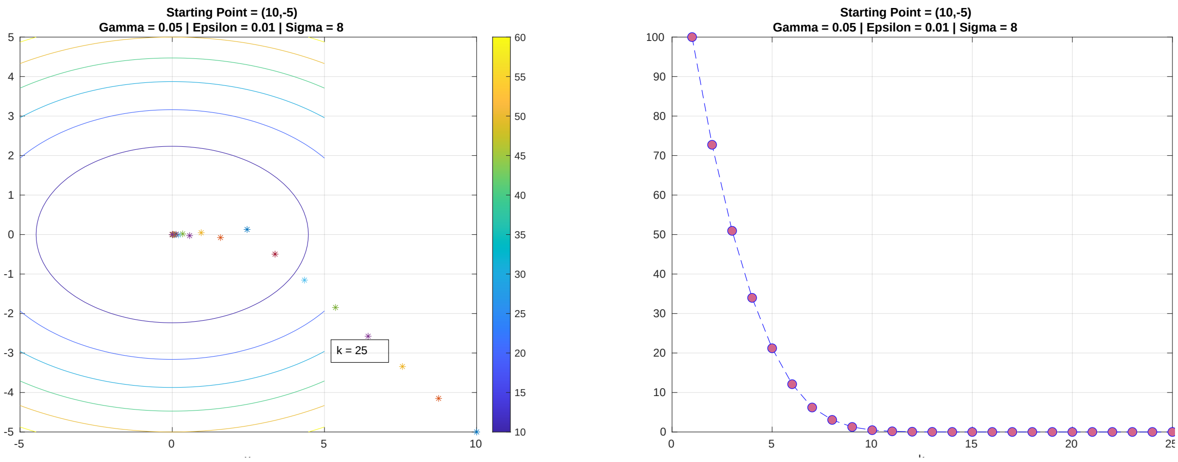
[1] Για να επιβεβαιωθεί η ισχύς του συμπεράσματος αυτού, ο αλγόριθμος δοκιμάστηκε και για τιμές λίγο μικρότερες του 0.5 όπου πράγματι οδηγούσε στο ελάχιστο.

[2] Η προβολή αυτή προκύπτει μέσω των φραγμάτων της αντικειμενικής συνάρτησης.

	I	**II**	**III**
$\mathbf{s_k}$	8	10	0.5
γ_k	0.05	0.3	0.1
$\mathbf{x}_0$	(10, -5)	(-7, 5)	(17, -5)
ϵ	0.01	0.02	0.01

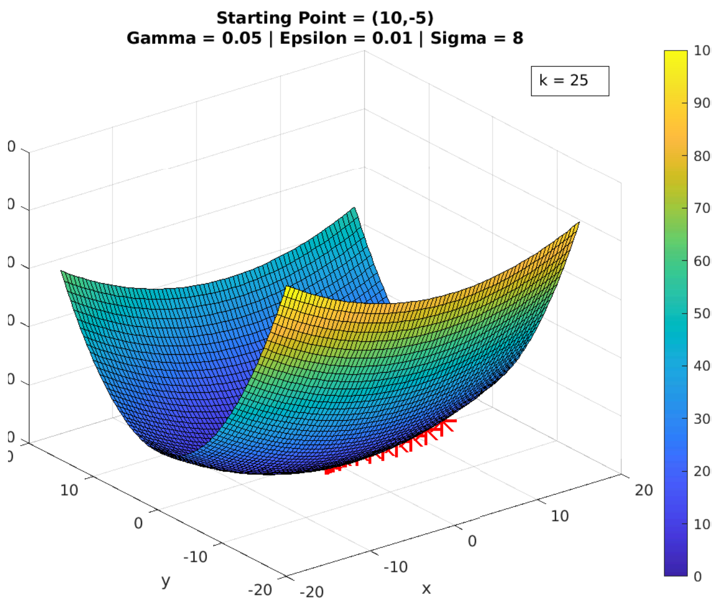
Πίνακας 1.1: Παράμετροι για κάθε εφαρμογή της μεθόδου

Ακολουθούν διαγράμματα από τα αποτελέσματα κάθε περίπτωσης:

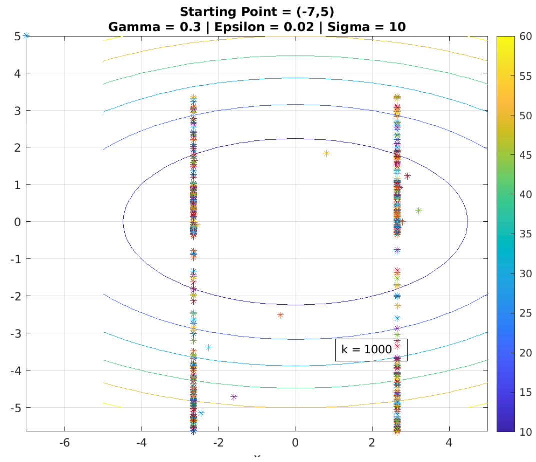


(α′) (β′)

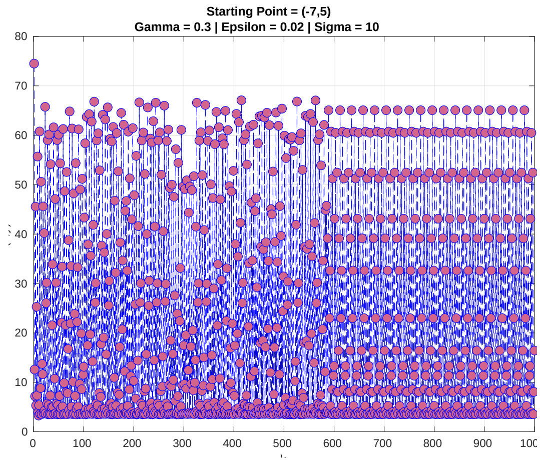
Σχήμα 1.6: Μέθοδος Μέγιστης Καθόδου για την Περίπτωση (I)



Σχήμα 1.7: Τρισδιάσταση απεικόνιση με τα διαδοχικά σημεία για την Περίπτωση (I)

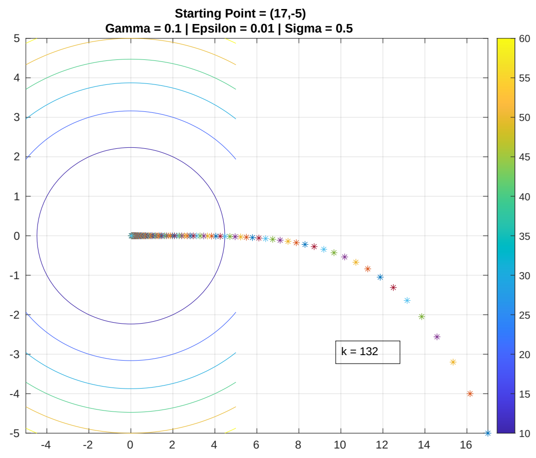


(α΄)

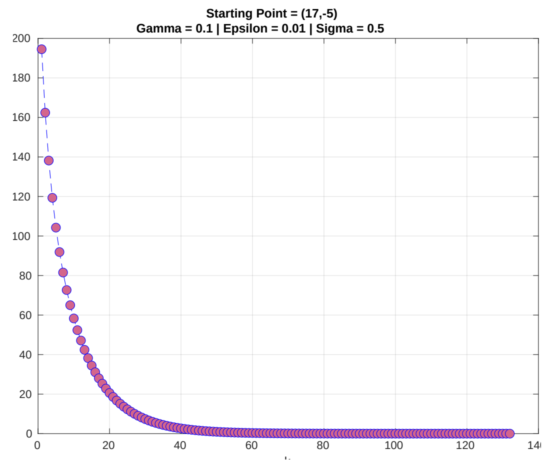


(β΄)

Σχήμα 1.8: Μέθοδος Μέγιστης Καθόδου για την Περίπτωση (II)

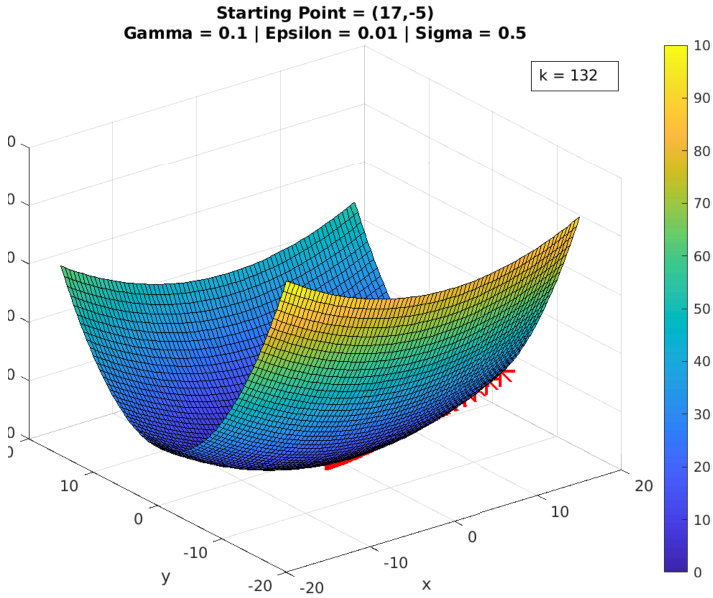


(α΄)



(β΄)

Σχήμα 1.9: Μέθοδος Μέγιστης Καθόδου για την Περίπτωση (III)



Σχήμα 1.10: Τρισδιάσταση απεικόνιση με τα διαδοχικά σημεία για την Περίπτωση (III)

Περίπτωση I Στην περίπτωση αυτή το σημείο εκκίνησης είναι **εφικτό** και ο αλγόριθμος **συγκλίνει στο ελάχιστο** και μάλιστα με σχετικά **μεγάλη** ταχύτητα (25 επαναλήψεις), όπως φαίνεται και στα διαγράμματα 1.6 και 1.7.

Περίπτωση II Στην δεύτερη εκτέλεση του αλγορίθμου, με μεταβολή των παραμέτρων, ο αλγόριθμος - ξεκινώντας από ένα εφικτό σημείο- **αδυνατεί να τερματίσει**. Μάλιστα, όπως φαίνεται έντονα στο διάγραμμα 1.6α΄, καταλήγει σε **ταλάντωση** με αποτέλεσμα να τερματίζει καθώς ξεπερνάει το ανώτατο όριο των 1000 επαναλήψεων που έχει οριστεί.

Περίπτωση III Στην τελευταία περίπτωση, ο αλγόριθμος παρόλο που έχει ένα **μη εφικτό** σημείο για σημείο εκκίνησης, οδηγείται στο ελάχιστο μετά από έναν σχετικά μεγαλύτερο αριθμό επαναλήψεων συγκριτικά με την πρώτη περίπτωση. Η σύγκλιση είναι εμφανής στο Σχήμα 1.9.

Κεφάλαιο 2

Συμπεράσματα & Παρατηρήσεις

Θέμα 1 Στην πρώτη ανάλυση της αναφοράς αυτής, παρατηρήσαμε την **επιρροή της τιμής του βήματος** γ_k στη σύγκλιση του αλγορίθμου Μέγιστης Καθόδου. Αποδείχθηκε και θεωρητικά, πως ανάλογα με την αντικειμενική συνάρτηση που μελετάται, υπάρχει ένα **άνω όριο**[1] στην τιμή του βήματος, εάν θέλουμε να επιτύχουμε εύρεση του ελαχίστου. Επιπλέον, αξίζει να αναφερθεί πως όταν το βήμα πάρει την οριακή τιμή, είναι εμφανές πως οδηγείται σε ταλάντωση γύρω από ένα σημείο.

Θέμα 2 Στο θέμα αυτό εξετάζουμε πλέον τη σύγκλιση της **Μεθόδου Μέγιστης Καθόδου με προβολή**. Έχοντας υπόψιν τον περιορισμό 1.10 που προέκυψε από την Θεωρητική ανάλυση του προηγούμενου θέματος, και τις τιμές των παραμέτρων που χρησιμοποιήθηκαν στην περίπτωση αυτή ($s_k = 8 \quad \& \quad \gamma_k = 0.05$) αναμένουμε η μέθοδος να **συγκλίνει**, εφόσον: $s_k \gamma_k = 0.4 < 0.5$. Η εκτίμηση αυτή συμφωνεί, λοιπόν, και με τα αποτελέσματα που προέκυψαν μετά την εφαρμογή του αλγορίθμου. Αξίζει επιπλέον να σημειωθεί πως το σημείο εκκίνησης του αλγορίθμου (10, -5) αποτελεί ένα εφικτό σημείο, **εντός** του συνόλου που ορίζεται από τους περιορισμούς 1.13.

Παρατηρούμε, επίσης, ότι συγκριτικά με το Θέμα 1(i), ο αλγόριθμος συγκλίνει με μεγαλύτερη ταχύτητα, καθώς απαιτεί σχεδόν 4 φορές μικρότερο αριθμό επαναλήψεων. Το γεγονός αυτό δικαιολογείται συγκρίνοντας τις τιμές των βημάτων σε κάθε περίπτωση.

Θέμα 3 Στην ανάλυση αυτή, ακολουθώντας την ίδια προσέγγιση με προηγουμένως και υπολογίζοντας το γινόμενο $s_k \gamma_k = 3 > 0.5$, αναμένουμε **μη σύγκλιση** του αλγορίθμου, το οποίο επιβεβαιώνει και η ανάλυση. Επομένως, για να οδηγηθούμε σε σύγκλιση, διατηρώντας σταθερή την τιμή του γ_k, θα δοκιμάζαμε να περιορίσουμε την τιμή του s_k στο εύρος [0, 1.66] για να ικανοποιείται η 1.10.[2] Μια σημαντική παρατήρηση που αξίζει να κάνουμε είναι πως ενώ στο Θέμα 1, οι δοκιμές με πολύ μεγάλες τιμές βήματος, οδηγούσαν σε **ακραίες τιμές της αντικειμενικής συνάρτησης**, τις οποίες **δεν** μπορούσαμε να αποτυπώσουμε, σε αυτή την περίπτωση, η Μέθοδος Μέγιστης Καθόδου με προβολή, **δεν επιτρέπει την ακραία αύξηση** αυτή και περιορίζεται σε

[1] Θεωρούμε δεδομένο πως το βήμα είναι πάντα θετικό.

[2] Αυτό πράγματι επιβεβαιώνεται μετά από εκτέλεση του αλγορίθμου για διάφορες τιμές του s_k οι οποίες ανήκουν στο επιθυμητό εύρος.

τιμές **εντός του ορισμένου συνόλου** X. Το σημείο εκκίνησης, είναι και πάλι **εντός** του συνόλου αυτού.

Θέμα 4 Κύρια διαφορά της δοκιμής αυτής αποτελεί πως το σημείο εκκίνησης **δεν** βρίσκεται εντός του συνόλου που ορίζουν οι περιορισμοί 1.10. Ωστόσο, αναμένουμε πως η μέθοδος προβολής που αναπτύξαμε, θα καταλήξει σε ένα **εφικτό σημείο** από την πρώτη κι όλας εκτέλεση. Κάνοντας και έναν έλεγχο του γινομένου $s_k \gamma_k = 0.05 < 0.5$, αναμένουμε ο αλγόριθμος ελαχιστοποίησης να **συγκλίνει στο ελάχιστο επιτυχώς**. Μετά την εκτέλεση της υλοποίησης του αλγορίθμου, η εκτίμηση αυτή επιβεβαιώνεται.

Ο κώδικας που αναπτύχθηκε για τις ανάγκες τις εργασίας αυτής είναι διαθέσιμος και στο αποθετήριο του GitHub optimization-techniques.