ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ

Τεχνικές Βελτιστοποίησης Αναφορά

Εργασία 3

Διακολουκάς Δημήτριος
ΑΕΜ 10642

Email: ddiakolou@ece.auth.gr

Περιεχόμενα

Κεφάλαιο 1

Εισαγωγή και σχεδιασμός συνάρτησης

Η τρίτη εργασία εστιάζει στη χρήση της **Μεθόδου Μέγιστης Καθόδου με Προβολή** για τη βελτιστοποίηση μιας συνάρτησης δύο μεταβλητών. Η συγκεκριμένη συνάρτηση που θα μελετηθεί είναι η $f(x) = \frac{1}{3}x_1^2 + 3x_2^2$, όπου $x = [x_1, x_2]^T$. Η εργασία περιλαμβάνει διάφορες παραλλαγές της μεθόδου, με στόχο τη σύγκριση της απόδοσης της σε διαφορετικά σενάρια βελτιστοποίησης, λαμβάνοντας υπόψη και περιορισμούς.

Ο βασικός στόχος της εργασίας είναι η διερεύνηση:

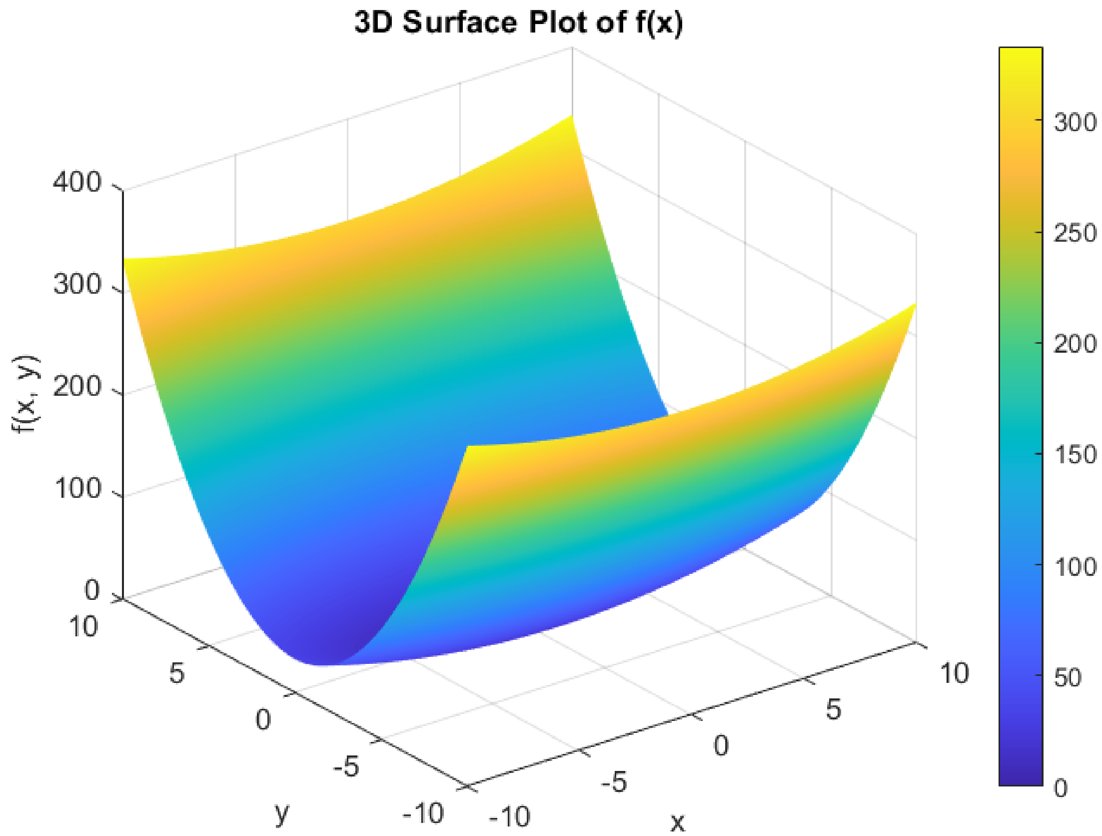
- Της σύγκλισης της **Μεθόδου Μέγιστης Καθόδου** υπό διαφορετικές τιμές για την παράμετρο γ_k και με συγκεκριμένο βαθμό ακρίβειας ε.
- Της συμπεριφοράς της μεθόδου όταν εισάγονται περιορισμοί στις μεταβλητές x_1 και x_2, δηλαδή $-10 \leq x_1 \leq 5$ και $-8 \leq x_2 \leq 12$.
- Της βελτίωσης της σύγκλισης μέσω αλλαγών στις παραμέτρους γ_k και s_k, καθώς και της χρήσης προβολών.

Η εργασία οργανώνεται σε τέσσερα θέματα:

1. Στο πρώτο θέμα, εφαρμόζεται η **Μέθοδος Μέγιστης Καθόδου** χωρίς περιορισμούς, εξετάζοντας διαφορετικές τιμές της παραμέτρου γ_k και την επίδρασή τους στη σύγκλιση.
2. Στο δεύτερο θέμα, εξετάζεται η μέθοδος υπό περιορισμούς στις μεταβλητές x_1 και x_2, χρησιμοποιώντας τον συντελεστή προβολής $s_k = 5$ και διαφορετικό αρχικό σημείο.
3. Στο τρίτο θέμα, εισάγεται η βελτίωση της μεθόδου με $s_k = 15$, και προτείνεται ένας πρακτικός τρόπος για την επιτάχυνση της σύγκλισης.
4. Στο τέταρτο θέμα, μελετάται η συμπεριφορά της μεθόδου όταν χρησιμοποιούνται εξαιρετικά μικρές τιμές για το s_k, με στόχο την ανάλυση της ακρίβειας του αλγορίθμου.

Η ανάλυση κάθε θέματος συνοδεύεται από διαγράμματα, αποτελέσματα και κώδικα $MATLAB$ ο οποίος επισυνάπτεται ξεχωριστά.

Παρακάτω επισυνάπτεται και μία γραφική αναπαράσταση της συνάρτησης στο Σχήμα 1.1 που μελετάται σε αυτή την εργασία και εμφανίζεται κατά την εκτέλεση του αρχείου **MyFunction.m**.



Σχήμα 1.1: Γενική εικόνα της μορφής της f.

Κεφάλαιο 2

Steepest Descent για πολλαπλά βήματα γ_k

Ανάλυση του Αλγορίθμου (Θέμα 1)

Ο παρακάτω κώδικας υλοποιεί τη **Μέθοδο Μέγιστης Καθόδου** για τη βελτιστοποίηση μιας συνάρτησης δύο μεταβλητών. Αναλύουμε τα μαθηματικά βήματα και τη λειτουργία του αλγορίθμου.

Περιγραφή της Μεθόδου

Η **Μέθοδος Μέγιστης Καθόδου** βασίζεται στη χρήση της κλίσης ($\nabla f(x)$) μιας συνάρτησης $f(x)$ για τον προσδιορισμό της διεύθυνσης στην οποία η συνάρτηση $f(x)$ μειώνεται περισσότερο. Ο αλγόριθμος εκτελεί επαναληπτικά βήματα σύμφωνα με τον τύπο:

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

όπου:

- x_k: το τρέχον σημείο,
- $\gamma_k > 0$: ο ρυθμός μάθησης ή το μήκος βήματος,
- $\nabla f(x_k)$: η κλίση της συνάρτησης στο σημείο x_k,
- $d_k = -\nabla f(x_k)$: η διεύθυνση μέγιστης καθόδου.

Επεξήγηση Υλοποίησης του Αλγορίθμου

1. **Αρχικοποίηση:** $x_0 = \text{start\_point}, k = 0.$
2. **Υπολογισμός της Κλίσης:**
$$\nabla f(x_k) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x_k) \\ \frac{\partial f}{\partial x_2}(x_k) \end{bmatrix}.$$
3. **Έλεγχος Σύγκλισης:** Εάν το μέτρο της κλίσης $\|\nabla f(x_k)\| = \sqrt{\left(\frac{\partial f}{\partial x_1}(x_k)\right)^2 + \left(\frac{\partial f}{\partial x_2}(x_k)\right)^2}$ είναι μικρότερο ή ίσο από ε, τότε ο αλγόριθμος σταματά.
4. **Υπολογισμός Διεύθυνσης Καθόδου:**
$$d_k = -\nabla f(x_k).$$

5. **Υπολογισμός Επόμενου Σημείου:** Το επόμενο σημείο x_{k+1} υπολογίζεται ως εξής:

$$x_{1,k+1} = x_{1,k} - \gamma_k \frac{\partial f}{\partial x_1}(x_k),$$

$$x_{2,k+1} = x_{2,k} - \gamma_k \frac{\partial f}{\partial x_2}(x_k).$$

6. **Επανάληψη:** Ενημερώνεται το πλήθος των επαναλήψεων $k \leftarrow k+1$ και επαναλαμβάνονται τα βήματα μέχρι τη σύγκλιση.

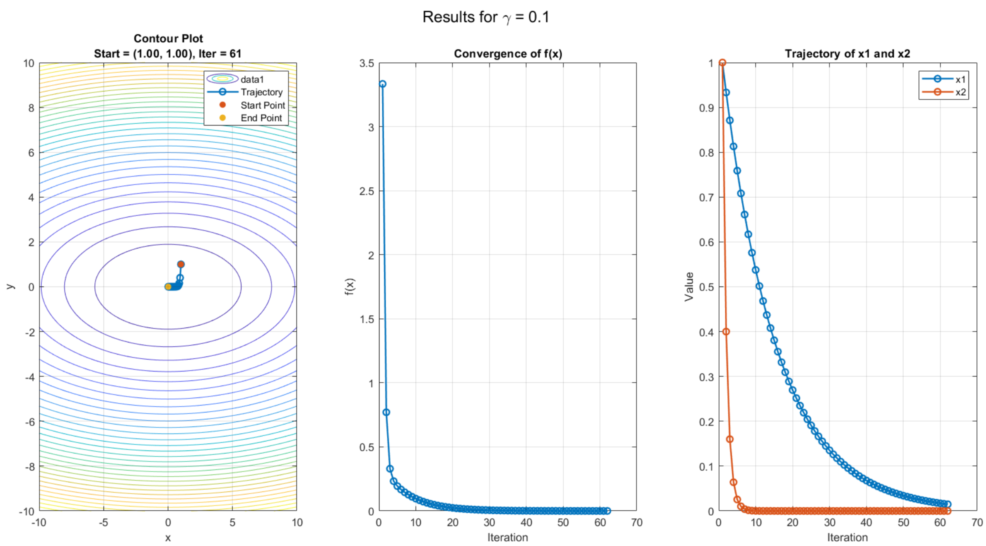
Περιορισμοί της Μεθόδου

Η **Μέθοδος Μέγιστης Καθόδου** είναι εύκολη στην υλοποίηση και αποτελεσματική για κυρτές συναρτήσεις, αλλά παρουσιάζει ορισμένα μειονεκτήματα:

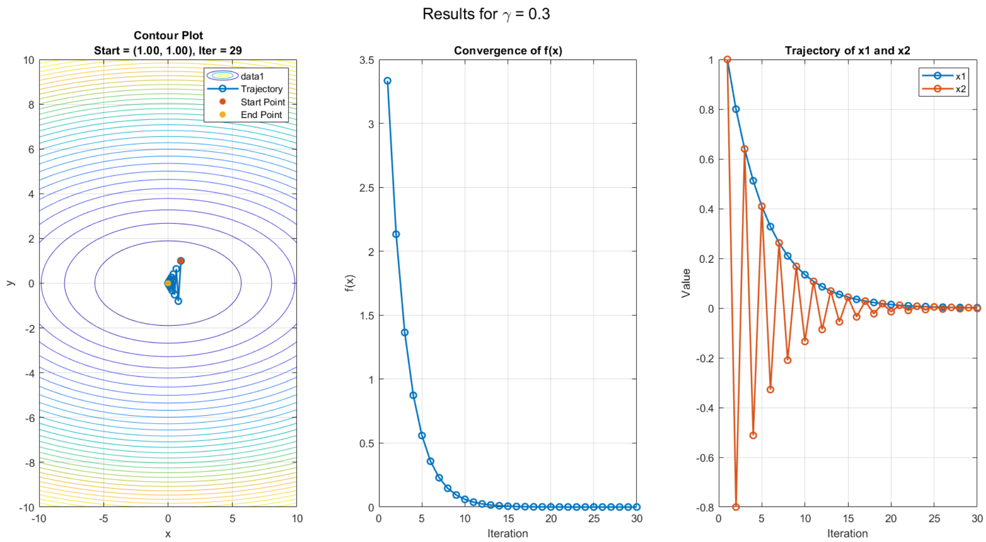
- Η ταχύτητα σύγκλισης εξαρτάται έντονα από την επιλογή του γ_k. Μικρές τιμές οδηγούν σε αργή σύγκλιση, ενώ μεγάλες τιμές μπορεί να προκαλέσουν αστάθεια.
- Μπορεί να παγιδευτεί σε τοπικά ελάχιστα όταν η συνάρτηση δεν είναι κυρτή.

Επίδειξη και Μαθηματική απόδειξη αποτελεσμάτων

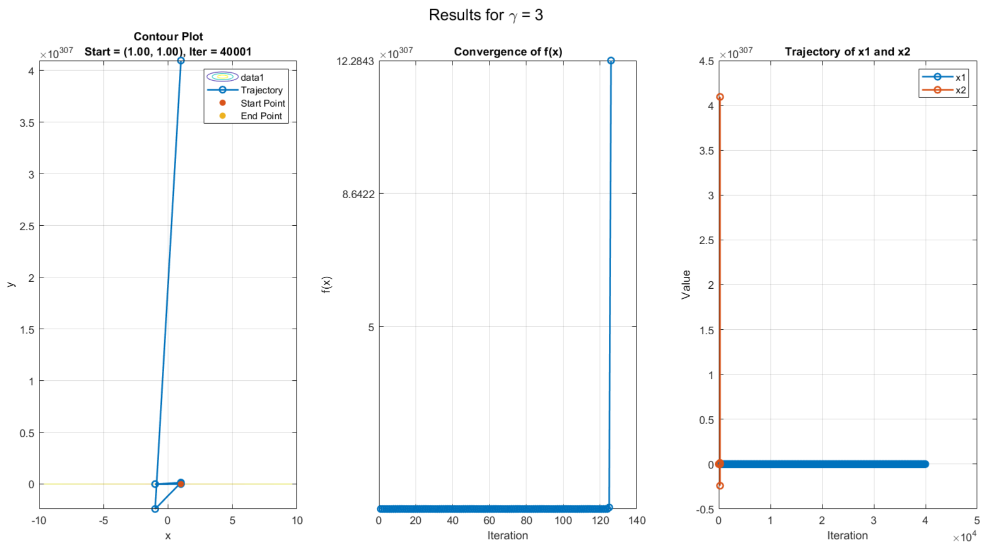
Έχοντας λοιπόν την γενικότερη ιδέα για την λειτουργικότητα της μεθόδου Μεγίστης Καθόδου παρακάτω επισυνάπτωνται και τα αποτελέσματα από την εφαρμογή της στο *MATLAB* για διάφορες τιμές γ_k. Ο αντίστοιχος κώδικας βρίσκεται στο αρχείο **Part1.m**. Σε κάθε περίπτωση θα παρατηρήσετε ότι επιλέχθηκε αρχικό σημείο το $(1, 1)$. Στο Σχήμα 2.1 αναπαρίστανται τα Plots για γ_k ίσο με 0.1 στο Σχήμε 2.2 για γ_k ίσο με 0.3 στο Σχήμα 2.3 για γ_k ίσο με 3 και στο σχήμα 2.4 για γ_k ίσο με 5.



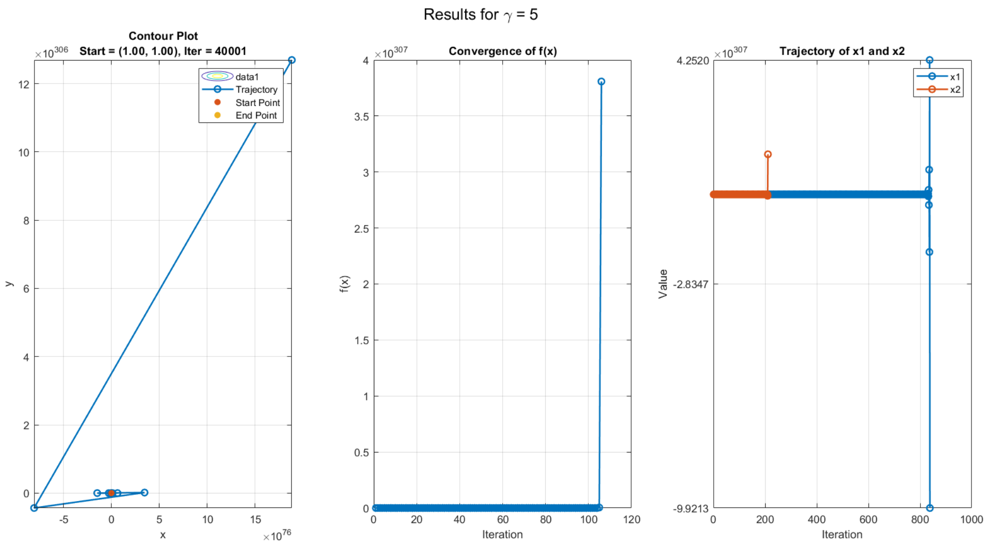
Σχήμα 2.1: Αναπαράσταση για γ_k ίσο με 0.1.



Σχήμα 2.2: Αναπαράσταση για γ_k ίσο με 0.3.



Σχήμα 2.3: Αναπαράσταση για γ_k ίσο με 3.



Σχήμα 2.4: Αναπαράσταση για γ_k ίσο με 5.

Παρατηρώ ότι στο Σχήμα 2.1 ο συνολικός αριθμός επαναλήψεων για $\gamma_k = 0.1$ ανέρχεται σε $k = 61$. Έπίσης παρατηρώ ότι για $\gamma = 0.1$, το μήκος βήματος είναι αρκετά μικρό, και ο αλγόριθμος συγκλίνει προς το συνολικό ελάχιστο σταθερά και αργά με την x_1 να συγκλίνει με βραδύτερο ρυθμό από την x_2. Επομένως, η μεταβλητή x_2 φτάνει στο ελάχιστο (0, 0) πιο γρήγορα από τη μεταβλητή x_1. Στο Σχήμα 2.2 φαίνεται πως ο συνολικός αριθμός επαναλήψεων για $\gamma_k = 0.3$ ανέρχεται σε $k = 29$. Επιπλέον παρατηρώ ταλαντώσεις του x_2 σε αντίθεση με το x_1 που δεν παρουσιάζει. Αυτό συμβαίνει διότι για $\gamma_k = 0.3$ προκαλείται μεγάλο βήμα, με αποτέλεσμα ο x_2 να μην συγκλίνει απευθείας, αλλά να υπερβαίνει και να επιστρέφει προς το ελάχιστο, προκαλώντας ταλάντωση. Τέλος, στις περιπτώσεις όπου το $\gamma_k = 3$ ή $\gamma = 5$ δεν παρατηρείται σύγκλιση του αλγορίθμου όπως άλλωστε φαίνεται ξεκάθαρα και στα σχήματα Σχήμα 2.3 και 2.4 αντίστοιχα.

Ακολουθεί απόδειξη των αποτελεσμάτων που είδαμε παραπάνω και προέκυψαν από τον $MATLAB$ κώδικα για τις διάφορες τιμές γ_k που ακολούθησαν με μαθηματική ακρίβεια. Έτσι μπορούν να εξηγηθούν και τα αποτελέσματα που πήραμε για τιμές γ_k ίσο με 0.1, 0.3, 3 και 5 αντίστοιχα. Έτσι, λοιπόν έχω μία αξιόλογη λύση για τον λόγο που για κάθε $\gamma_k > \frac{1}{3}$ δεν έχω σύγκλιση. Παρακάτω, λοιπόν επισυνάπτω την απόδειξη όπως άλλωστε φαίνεται στο Σχήμα 2.5.

Αντικειμενική Συνάρτηση $f(x_1,x_2)=\frac{1}{3}x_1^2+3x_2^2$, $\boxed{\begin{array}{l} x_k=[x_{1k},\ x_{2k}]^T \\ x_{k+1}=[x_{1k+1},\ x_{2k+1}]^T \end{array}}$

Steepest Descent: $x_{k+1}=x_k-\gamma_k\nabla f(x_k)$

$$\nabla f(x_1,x_2)=\begin{bmatrix}\partial f/\partial x_1\\ \partial f/\partial x_2\end{bmatrix}=\begin{bmatrix}2x_1/3\\ 6x_2\end{bmatrix}\Rightarrow \nabla f(x_k)=\begin{bmatrix}0{,}6667\cdot x_{1k}\\ 6x_{2k}\end{bmatrix}=\begin{bmatrix}0{,}6667 & 0\\ 0 & 6\end{bmatrix}\underbrace{\begin{bmatrix}x_{1k}\\ x_{2k}\end{bmatrix}}_{x_k}$$

Hessian $\nabla^2 f(x_1,x_2)=\begin{bmatrix}0{,}6667 & 0\\ 0 & 6\end{bmatrix}$ Θετικά ορισμένος $\longrightarrow (0,0)$ ολικό ελάχιστο της συνάρτησης f

(*αφού $\nabla f(x_1,x_2)=\left[\frac{2}{3}x_1\quad 6x_2\right]^T$ τότε $\nabla f(x_1,x_2)=0$ έχω κρίσιμα σημεία και $x_1,x_2=0$)

Για να συγκλίνει η μέθοδος Μεγίστης Καθόδου θέλω: $\left|\frac{x_{k+1}}{x_k}\right|<1$ και άρα

$$\text{θέλω}\begin{cases}\left|\frac{x_{1k+1}}{x_{1k}}\right|<1 \quad (1) \quad \text{με } d_k=-\nabla f(x_{1k},x_{2k})\\[2ex] \left|\frac{x_{2k+1}}{x_{2k}}\right|<1\end{cases}$$

Ακολουθώντας τα βήματα της μεθόδου από θεωρία έχω:

$$x_{k+1}=x_k-\gamma_k\nabla f x_k \Rightarrow \begin{cases}x_{1k+1}=x_{1k}-\gamma_k\cdot 0{,}6667\,x_{1k}\\[1ex] x_{2k+1}=x_{2k}-\gamma_k\cdot 6\,x_{2k}\end{cases}\Rightarrow\begin{cases}\frac{x_{1k+1}}{x_{1k}}=1-\gamma_k\cdot 0{,}6667\\[1ex] \frac{x_{2k+1}}{x_{2k}}=1-\gamma_k\cdot 6\end{cases}\Rightarrow$$

$$\overset{(1)}{\Rightarrow}\begin{cases}|1-\gamma_k\cdot 0{,}6667|<1\\[1ex] |1-\gamma_k\cdot 6|<1\end{cases}\Rightarrow\begin{cases}0<\gamma_k<3\\[1ex] 0<\gamma_k<\frac{1}{3}\end{cases}$$

οπότε για σύγκλιση με σταθερό βήμα θέλω $\boxed{\gamma_k<\frac{1}{3}}$

Σχήμα 2.5: Μαθηματική Απόδειξη.

Κεφάλαιο 3

Μέθοδος Steepest Descent with Projection

Ανάλυση του Αλγορίθμου

Περιγραφή της Μεθόδου

Στα υποερωτήματα 2, 3 και 4 θα περιγραφεί η χρήση της Μεθόδου Μέγιστης Καθόδου με Προβολή για την ελαχιστοποίηση μιας συνάρτησης υπό συγκεκριμένους περιορισμούς. Εξετάζονται διάφορες τιμές των παραμέτρων γ_k και s_k, καθώς και διαφορετικά αρχικά σημεία, για να αναλυθεί η σύγκλιση του αλγορίθμου, οι ιδιαιτερότητές του και να προταθούν πρακτικές βελτιώσεις. Όπως και στην απλή μέθοδο Μεγίστης Καθόδου έτσι και στην μέθοδο Μεγίστης Καθόδου με Προβολή θα ακολουθήσουμε την παρόμοια υλοποίηση στον αλγόριθμό μας με μικρές διαφοροποιήσεις.

Επεξήγηση Υλοποίησης του Αλγορίθμου

1. **Αρχικοποίηση:** Ορίζεται το αρχικό σημείο $x_0 \in \mathbb{R}^2$ και ο αριθμός των επαναλήψεων $k = 0$. Η αρχική τιμή της συνάρτησης $f(x_0)$ υπολογίζεται ως:

$$f(x_0) = \frac{1}{3}x_{1,0}^2 + 3x_{2,0}^2.$$

2. **Υπολογισμός της Κλίσης:** Για κάθε επανάληψη k, υπολογίζεται η κλίση της συνάρτησης $f(x)$ στο σημείο x_k:

$$\nabla f(x_k) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x_k) \\ \frac{\partial f}{\partial x_2}(x_k) \end{bmatrix} = \begin{bmatrix} \frac{2}{3}x_{1,k} \\ 6x_{2,k} \end{bmatrix}.$$

3. **Κριτήριο Τερματισμού:** Εξετάζεται εάν το μέτρο της κλίσης $\|\nabla f(x_k)\|$ ικανοποιεί την προϋπόθεση σύγκλισης:

$$\|\nabla f(x_k)\| = \sqrt{\left(\frac{2}{3}x_{1,k}\right)^2 + (6x_{2,k})^2} \leq \varepsilon.$$

   Εάν η παραπάνω συνθήκη ικανοποιηθεί ή $k > 40000$, η διαδικασία τερματίζεται.

4. **Υπολογισμός Διεύθυνσης Καθόδου:** Η κατεύθυνση καθόδου d_k είναι η αρνητική κλίση:

$$d_k = -\nabla f(x_k).$$

5. **Ενδιάμεσο Σημείο:** Υπολογίζεται το προσωρινό σημείο $x_{\tau\epsilon\mu\pi}$ χρησιμοποιώντας το βήμα s_k:
$$x_{\text{temp},1} = x_{1,k} + s_k d_{1,k}, \quad x_{\text{temp},2} = x_{2,k} + s_k d_{2,k}.$$

6. **Προβολή στα Όρια:** Λαμβάνονται υπόψη οι περιορισμοί του χώρου:
$$-10 \leq x_1 \leq 5$$
$$-8 \leq x_2 \leq 12$$
για τις συνιστώσες του x_{temp}:
$$x_{\text{next},1} = \min(\max(x_{\text{temp},1}, x_{\min}), x_{\max}),$$
$$x_{\text{next},2} = \min(\max(x_{\text{temp},2}, y_{\min}), y_{\max}).$$

7. **Ενημέρωση Σημείου:** Το επόμενο σημείο x_{k+1} υπολογίζεται με βάση τη χαλάρωση:
$$x_{1,k+1} = x_{1,k} + \gamma_k(x_{\text{next},1} - x_{1,k}),$$
$$x_{2,k+1} = x_{2,k} + \gamma_k(x_{\text{next},2} - x_{2,k}).$$

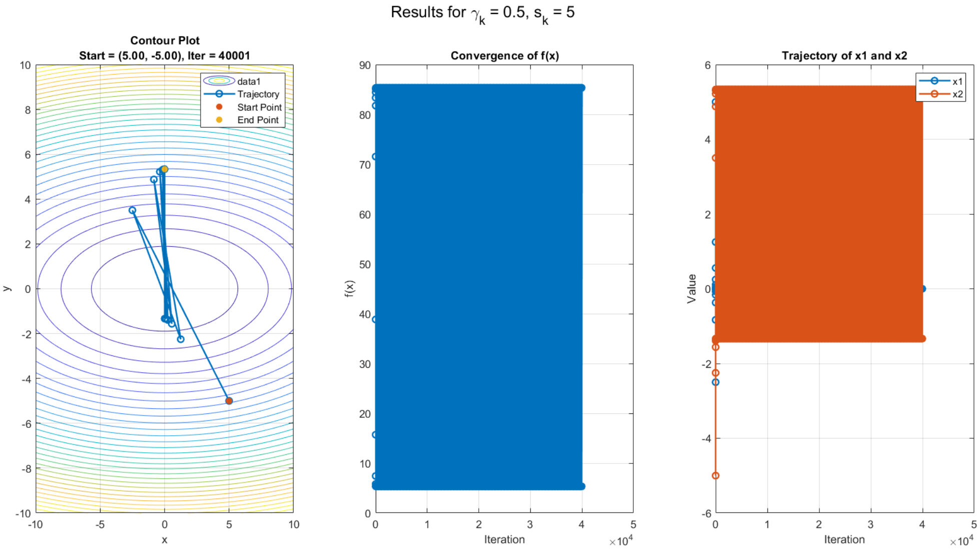
8. **Επανάληψη:** Καταγράφονται οι τιμές των x_k, y_k και της συνάρτησης $f(x_k)$:
$$f(x_k) = \frac{1}{3}x_{1,k}^2 + 3x_{2,k}^2.$$
Το πλήθος των επαναλήψεων ενημερώνεται $k \leftarrow k+1$, και η διαδικασία επαναλαμβάνεται μέχρι να ικανοποιηθεί το κριτήριο σύγκλισης.

Οι περιορισμοί και αδυναμίες και αυτής της μεθόδου είναι αντίστοιχοι με αυτούς που συζητήθηκαν στην απλή μέθοδο Μεγίστης Καθόδου παραπάνω. Ο αντίστοιχος αλγόριθμος βρίσκεται στο αρχείο **Part2.m**.

Επίδειξη αποτελεσμάτων και συμπεράσματα συγκρίσεων

Θέμα 2

Στο Θέμα αυτό έχουμε $s_k = 5$, $\gamma_k = 0.5$ και σημείο εκκίνησης το $(5, -5)$ παρατηρούμε στα διαγράμματα παρακάτω ότι δεν υπάρχει καμία συγκλιση στις 40000 επαναλήψεις. Σε αυτό το ερώτημα ακολουθώντας την ίδια ανάλυση που κάναμε για το πρώτο θέμα εύκολα απορρέει το συμπέρασμα ότι θέλουμε να ισχύει $s_k \cdot \gamma_k < \frac{1}{3}$. Ωστόσο με βάση το Σχήμα 3.1 φαίνεται πως $s_k \cdot \gamma_k = 2.5 > \frac{1}{3}$. Άρα ο αλγόριθμος αποκλίνει. Όπως είδαμε σε ένα απο τα παραδείγματα και στο Θέμα 1 έτσι και εδώ έχουμε ταλαντώσεις που δεν καταλήγουν σε σύγκλιση. Το πρόβλημα προκαλείται από το μεγάλο βήμα που οδηγεί την προβολή x_{2k} εκτός των ορίων του επιτρεπτού συνόλου και σε αυτή την περίπτωση. Ωστόσο το x_{1k} φαίνεται να συγκλίνει στο 0 λογικό καθώς $s_k \cdot \gamma_k = 2.5 < 3$ που ισχύει ως περιορισμός για το x_{1k} όπως αποδείξαμε και στο Θέμα 1 κατ΄ αντιστοιχία.

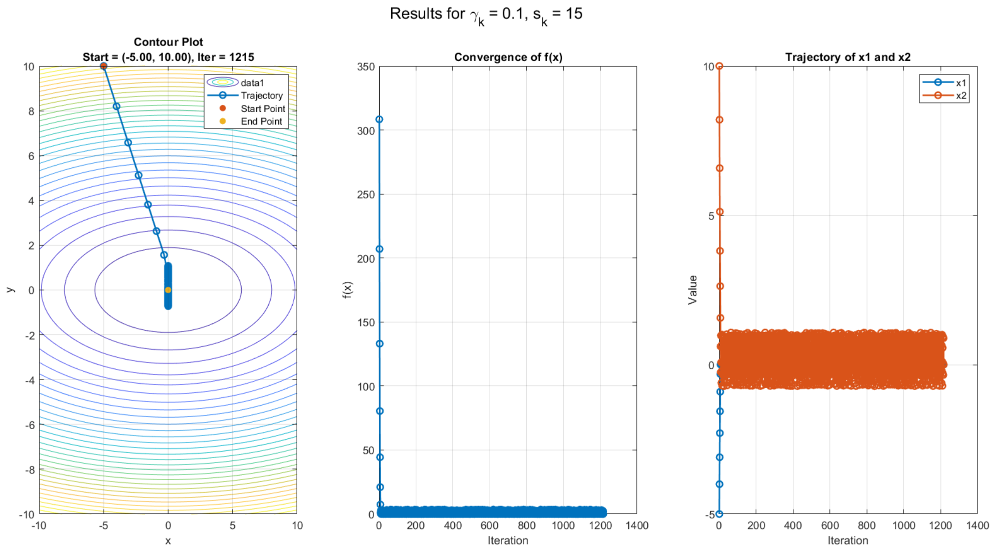


Σχήμα 3.1: Αναπαράσταση για γ_k ίσο με 0.5 και s_k ίσο με 5.

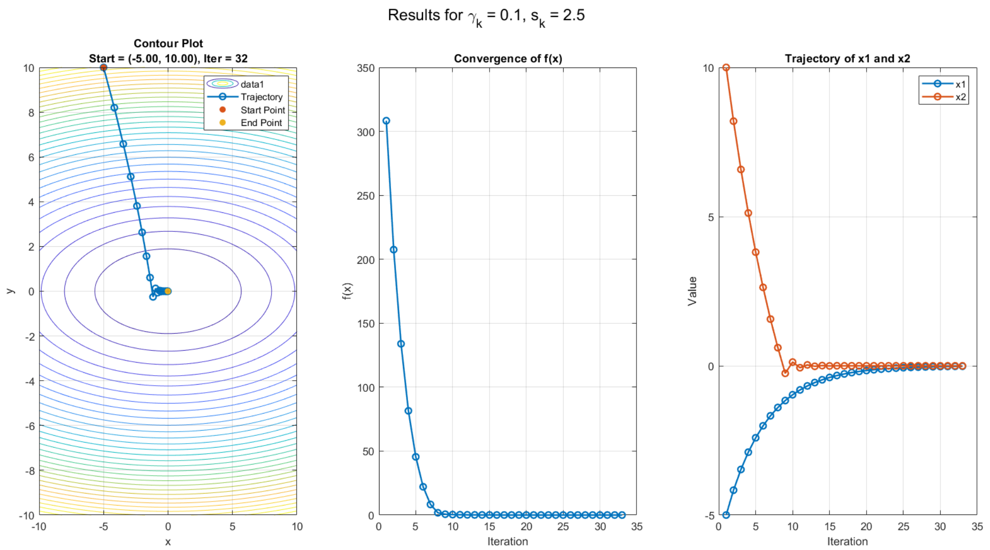
Θέμα 3

Στο Θέμα αυτό έχουμε $s_k = 15$, $\gamma_k = 0.1$ και σημείο εκκίνησης το $(-5, 10)$ αυτό παρατηρούμε στα διαγράμματα παρακάτω ότι η αντικειμενική συνάρτηση συγκλίνει στις 1215 επαναλήψεις όπως φαίνεται και στο Σχήμα 3.2. Σε αυτό το ερώτημα ακολουθώντας την ίδια ανάλυση που κάναμε για το προηγούμενο παρατηρούμε ότι $s_k \cdot \gamma_k = 1.5 > \frac{1}{3}$. Για αυτό το x_{2k} ταλαντώνεται και χρειάζεται πολλές επαναλήψεις για σύγκλιση στο 0. Επίσης μπορούμε να καταλάβουμε την σύγκλιση που έχει το x_{1k} στο 0 καθώς $s_k \cdot \gamma_k = 1.5 < 3$. Συνεπώς, έχω πολύ αργή σύγκλιση.

Ένας άλλος τρόπος να πετύχουμε σύγκλιση αλλά με λιγότερες επαναλήψεις έκρινα ότι είναι μειώνωντας το $s_k = 15$ σε $s_k = 2.5$ διατηρώντας το $\gamma_k = 0.1$. Σε αυτή την περίπτωση έχουμε $s_k \cdot \gamma_k = 0.25 < \frac{1}{3}$ και παρατηρούμε σύγκλιση στις 32 επαναλήψεις όπως φαίνεται και στο Σχήμα 3.3 που συγκριτικά φαίνεται ως μία αρκετά μεγάλη βελτίωση.



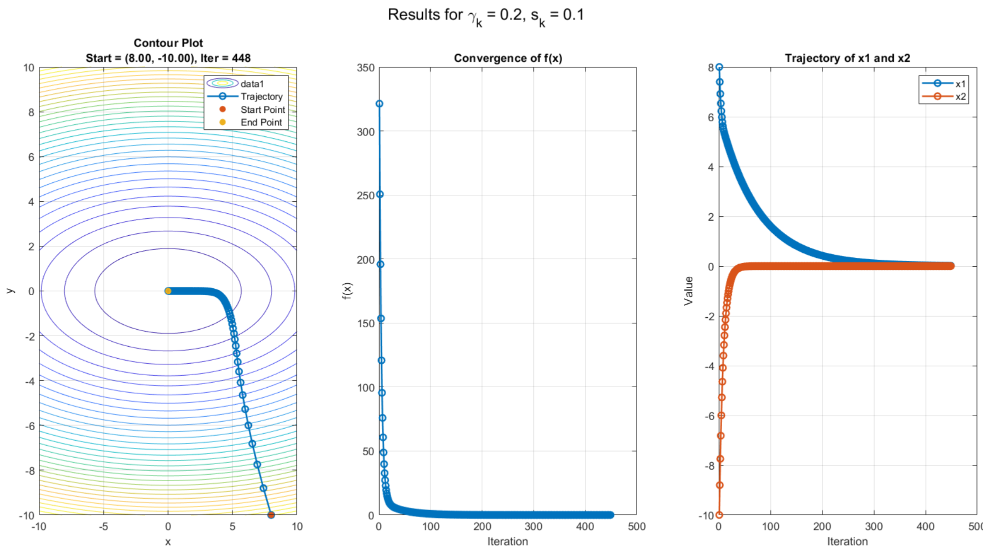
Σχήμα 3.2: Αναπαράσταση για γ_k ίσο με 0.1 και s_k ίσο με 15.



Σχήμα 3.3: Αναπαράσταση για γ_k ίσο με 0.1 και s_k ίσο με 2.5.

Θέμα 4

Στο Θέμα αυτό έχουμε $s_k = 0.1$, $\gamma_k = 0.2$ και σημείο εκκίνησης το $(8, -10)$ παρατηρούμε στα διαγράμματα παρακάτω ότι υπάρχει συγκλιση στις 448 επαναλήψεις. Σε αυτό το ερώτημα ακολουθώντας την ίδια ανάλυση που κάναμε και στα προηγούμενα εύκολα απορρέει το συμπέρασμα ότι θέλουμε να ισχύει $s_k \cdot \gamma_k < \frac{1}{3}$ πράγμα που ισχύει αφού $s_k \cdot \gamma_k = 0.02$ και τα αποτελέσματα αναπαρίστανται και στο Σχήμα 3.4. Ωστόσο οι μικρές τιμές σε s_k αλλά και γ_k μας δείχνουν ότι συγκλίνουμε αργά.



Σχήμα 3.4: Αναπαράσταση για γ_k ίσο με 0.2 και s_k ίσο με 0.1.

Βιβλιογραφία

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