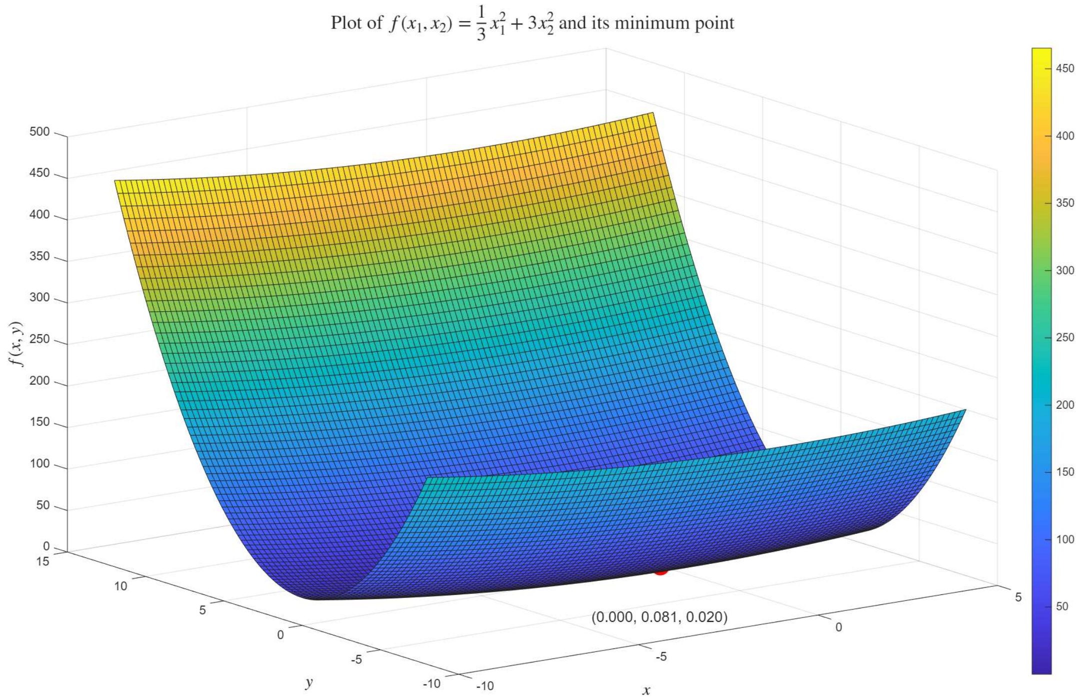
Τεχνικές Βελτιστοποίησης 3η Εργαστηριακή Άσκηση

Σπύρος Κούγιας, ΑΕΜ: 10124

I. ΠΕΡΙΓΡΑΦΗ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ

Σκοπός της παρούσας εργασίας είναι η μελέτη και η επίλυση προβλημάτων ελαχιστοποίησης για τη μη γραμμική συνάρτηση δύο μεταβλητών:

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



Η συγκεκριμένη συνάρτηση είναι τετραγωνική, διαχωρίσιμη και αυστηρά κυρτή. Το μοναδικό ολικό ελάχιστο της συνάρτησης βρίσκεται στο σημείο (0,0) με τιμή f(0,0)=0, (υπολογισμένο μαθηματικά). Μορφής ελλειπτικού παραβολειδούς, όπου η καμπυλότητα είναι πιο έντονη στον άξονα x2 σε σχέση με τον άξονα x1.

Η εργασία αποτελείτε από δύο μέρη:

1. Στο πρώτο μέρος, εξετάζουμε την ελαχιστοποίηση χρησιμοποιώντας την κλασική Μέθοδο της Μέγιστης Καθόδου με σταθερό βήμα, σε ολόκληρο το πεδίο ορισμού της αντικειμενικής συνάρτησης.
2. Στο δεύτερο μέρος χρησιμοποιούμε τη Μέθοδο της Μέγιστης Καθόδου με Προβολή εντός ενός συγκεκριμένου υποσυνόλου του πεδίου ορισμού, το οποίο, για την παρούσα ανάλυση, είναι το σύνολο X:

$$-10 \le x_1 \le 5 \ \text{ και } \ -8 \le x_2 \le 12$$

Οι αλγόριθμοι (όπως υλοποιήθηκαν στο Matlab) επιστρέφουν όλα τα σημεία xk που επισκέφτηκαν κατά τη διάρκεια της εκτέλεσής τους.

II. ΘΕΜΑ 1: ΜΕΘΟΔΟΣ ΜΕΓΙΣΤΗΣ ΚΑΘΟΔΟΥ

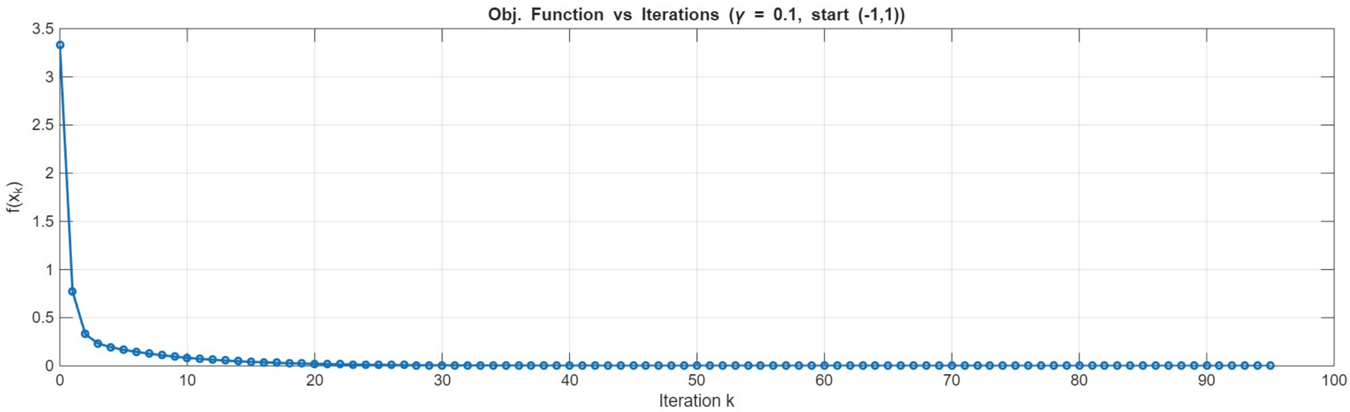
Η μέθοδος Μέγιστης Καθόδου έχει ήδη εξηγηθεί στην προηγούμενη εργασία. Γενικώς όμως, η μέθοδος στηρίζεται στην ιδέα της επαναληπτικής καθόδου με αναδρομική σχέση:

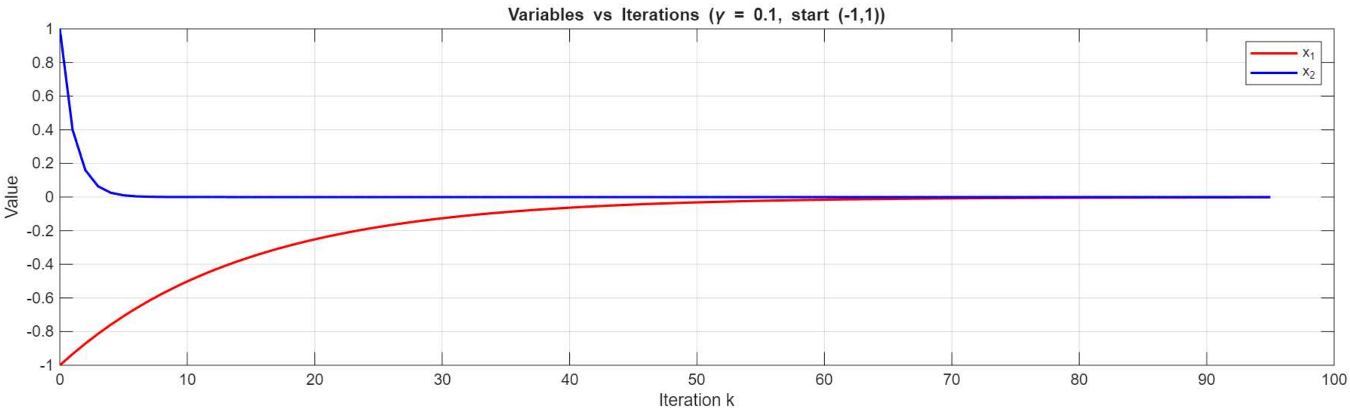
$$x_{k+1} = x_k - \gamma \nabla f(x_k)$$

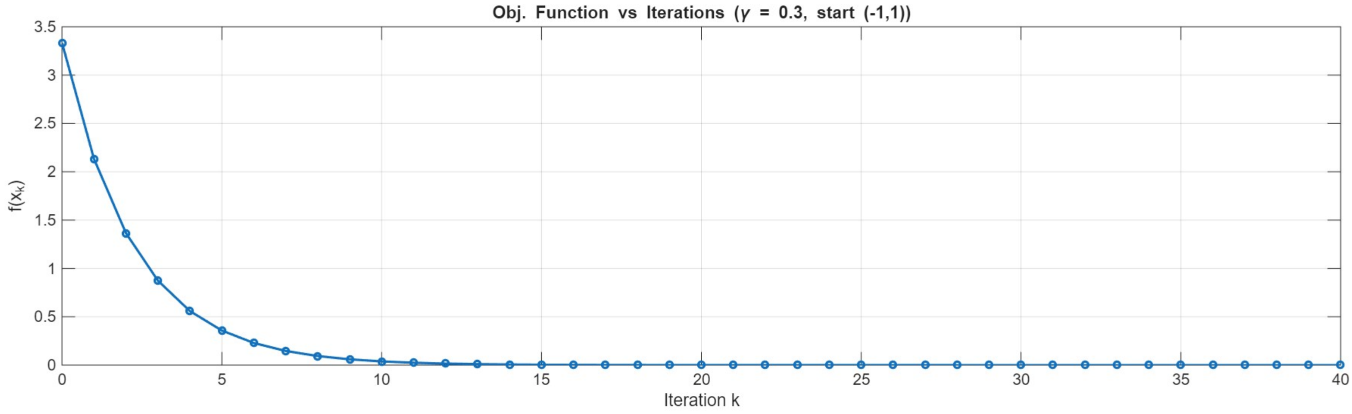
Όπου η κλίση της συνάρτησης είναι:

$$\nabla f(x_k) = \begin{bmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} \dfrac{2}{3} x_1 \\ 6x_2 \end{bmatrix}$$

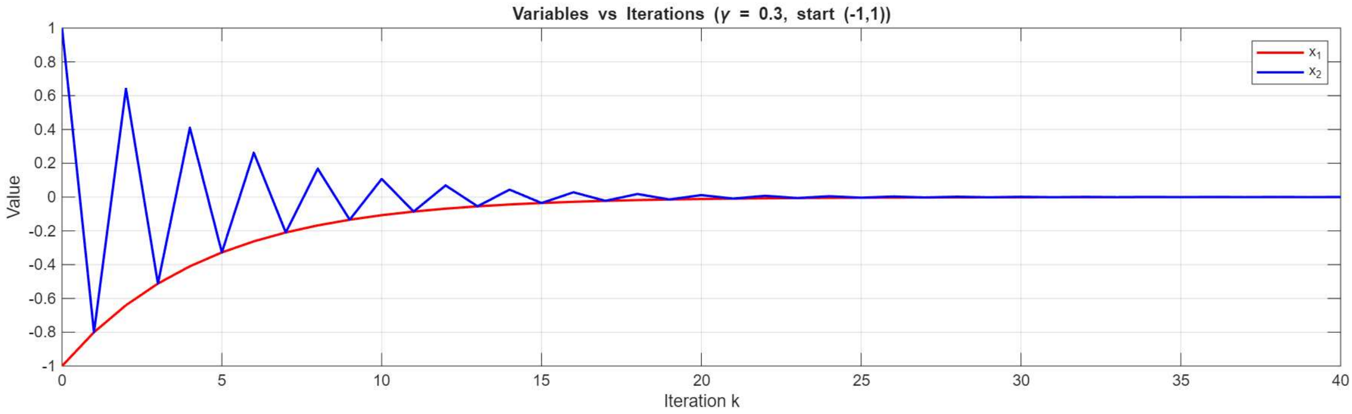
Εφαρμόσαμε τον αλγόριθμο για τέσσερις διαφορετικές τιμές του βήματος γ = 0.1, 0.3, 3 και 5, με ακρίβεια ε = 0.001.



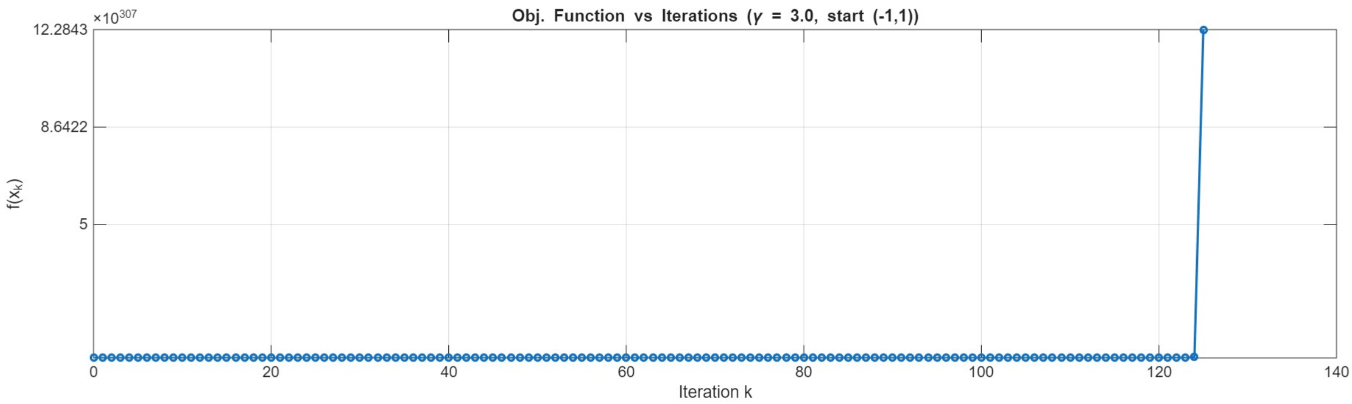




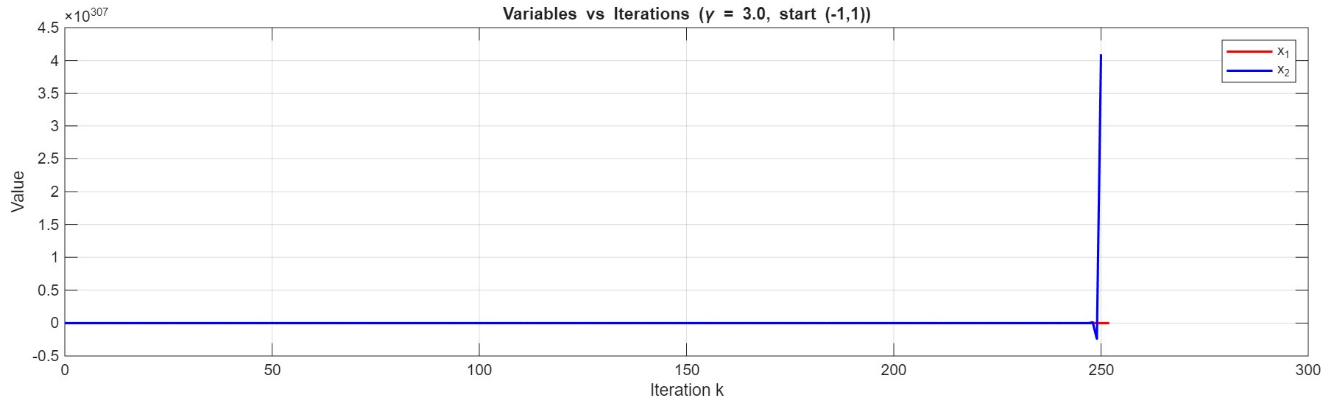
Obj. Function vs Iterations (γ = 0.3, start (-1,1))
f(x_k)
Iteration k



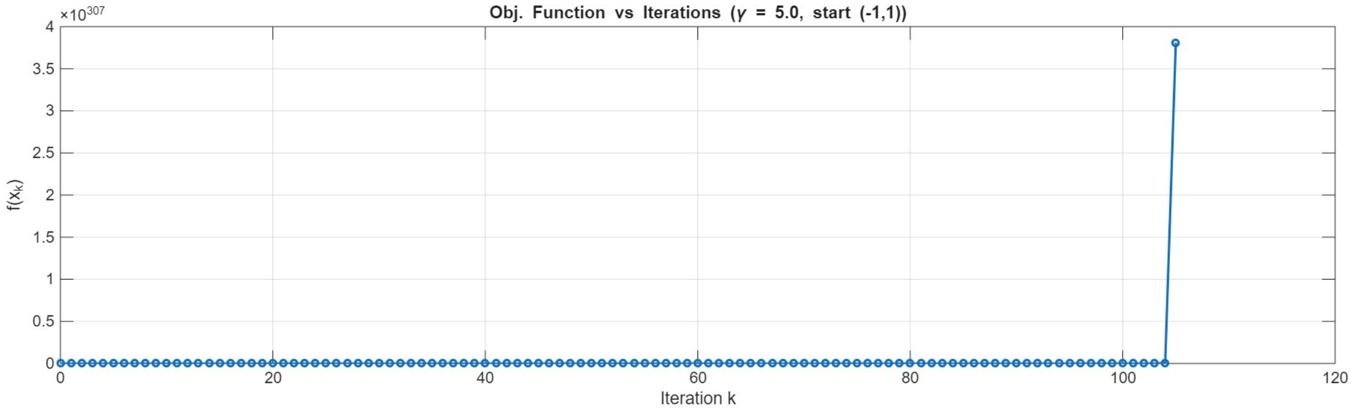
Variables vs Iterations (γ = 0.3, start (-1,1))
Value
Iteration k
x_1
x_2

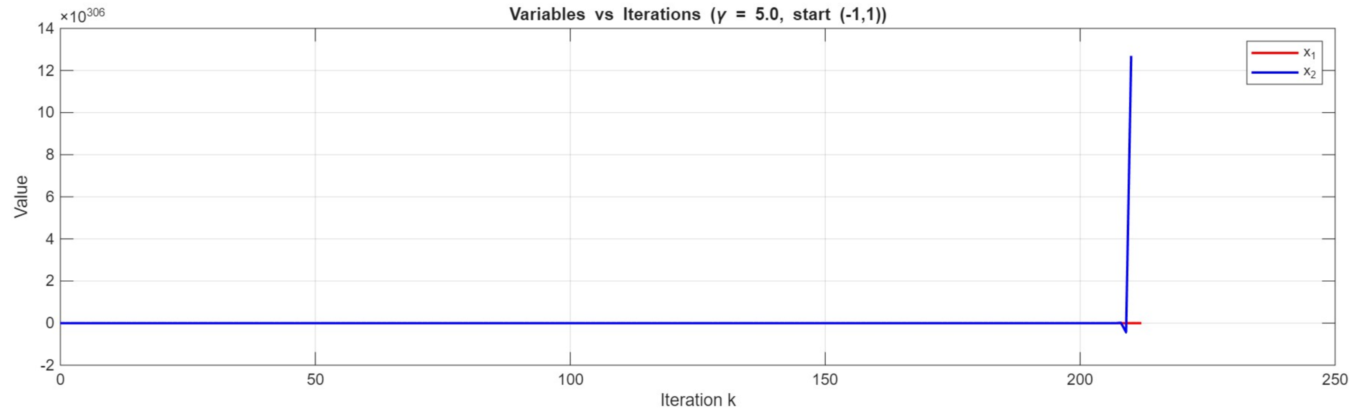


Obj. Function vs Iterations (γ = 3.0, start (-1,1))
×10^307
f(x_k)
Iteration k



Variables vs Iterations (γ = 3.0, start (-1,1))
×10^307
Value
Iteration k
x_1
x_2

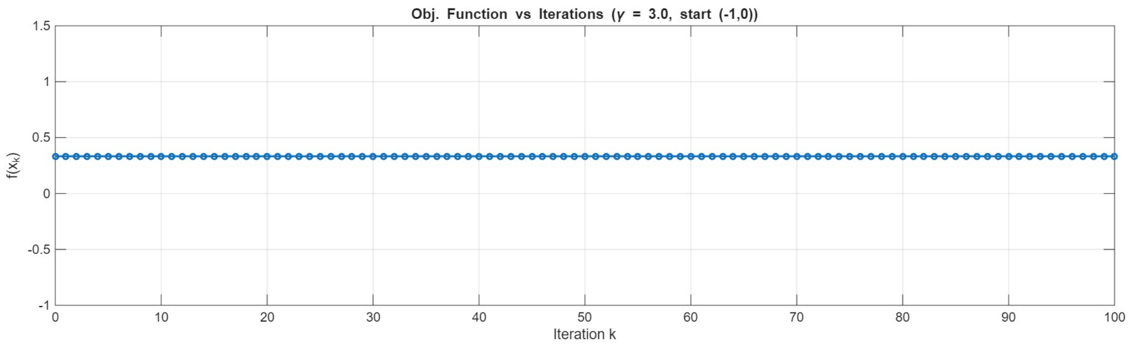


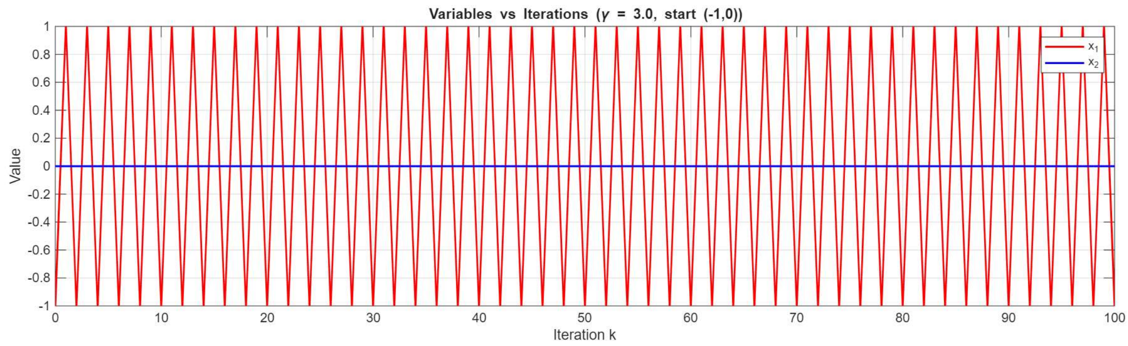


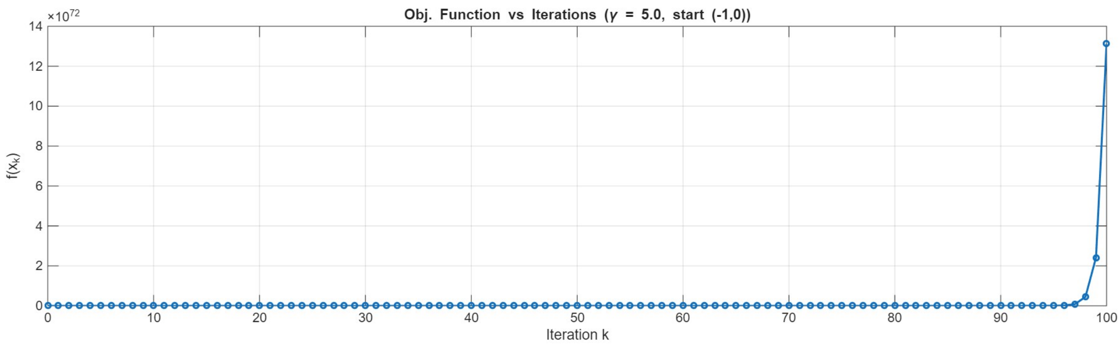
Παρατηρούμε τα εξής για αρχικό σημείο (-1,1):

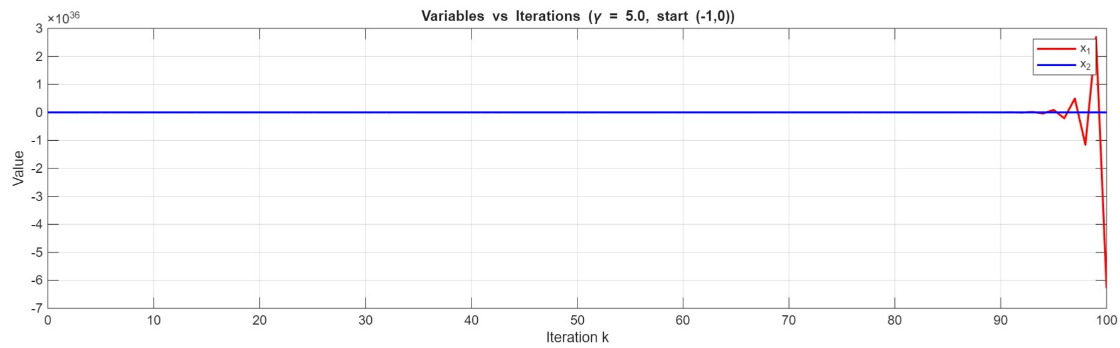
- Για γ = 0.1: Η μέθοδος συγκλίνει ομαλά και μονότονα στο ελάχιστο (0,0). Οι μεταβλητές μειώνονται σταθερά προς το μηδέν.
- Για γ = 0.3: Η μέθοδος τελικά συγκλίνει, ωστόσο παρατηρείται έντονη ταλάντωση (x2). Αυτό πιθανότατα συμβαίνει διότι το βήμα είναι πολύ κοντά στο θεωρητικό όριο ευστάθειας (1/3).
- Για γ = 3 και γ = 5: Η μέθοδος αποκλίνει βίαια.

Αξίζει να σημειωθεί ότι για αρχικά σημεία με μορφή (a, 0) και βήμα γ = 3 παρατηρείται ατέρμονη ταλάντωση και όχι απόκλιση. Κάτι παρόμοιο δεν συμβαίνει για γ = 5 όπως μαρτυρούν και τα παρακάτω διαγράμματα:









Όμοιες δοκιμές έγιναν με αρχικά σημεία x = (a, 0) με a = {1, 2, -2, -7} με παρόμοια αποτελέσματα.

Μαθηματική Απόδειξη Σύγκλισης

Αντικαθιστώντας την κλίση στον τύπο ανανέωσης, έχουμε για κάθε συνιστώσα:

$$x_{1,k+1} = x_{1,k} - \frac{2}{3}\gamma x_{1,k} = \left(1 - \frac{2}{3}\gamma\right) x_{1,k}$$

$$x_{2,k+1} = (1-6)x_{2,k}$$

Η συνθήκη σύγκλισης είναι:

$$f(x_{k+1}) < f(x_k)$$

Και έτσι καταλήγουμε στην σχέση:

$$(\gamma - 3)x_{1,k}^2 + (729\gamma - 243)x_{2,k}^2 < 0$$

Για να είναι και οι δύο όροι αρνητικοί, πρέπει:

$$[\gamma < 3]\ \&\left[\gamma < \frac{1}{3}\right] \Rightarrow \gamma < \frac{1}{3}$$

Για γ < 1/3 λοιπόν, έχουμε σύγκλισή στο ελάχιστο ανεξάρτητα από την επιλογή του αρχικού σημείου. Αντιθέτως για γ > 3 και οι δύο όροι είναι θετικοί και έτσι έχουμε απόκλιση για κάθε αρχικό σημείο.

Τέλος για γ = 1/3 και γ = 3 θα έχουμε ατέρμονη ταλάντωση για αρχικά σημεία (0, b) και (a, 0) αντίστοιχα. Αυτό συμβαίνει διότι όταν το γ πάρει μία από αυτές τις δύο τιμές, ένας από τους όρους μηδενίζεται.

Για παράδειγμα, αν γ = 1/3 η νέα συνθήκη σύγκλισης είναι:

$$-\frac{8}{3}x_{1,k}^2 < 0$$

Η οποία ικανοποιείται για κάθε x1 διάφορο του μηδενός (συγκλίνει). Ωστόσο, αν το σημείο εκκίνησης είναι της μορφής (0, b) έχουμε:

$$x_{1,k+1} = \frac{7}{9}x_{1,k}\ \ \mu\varepsilon\ \ x_{1,0} = 0$$

Και

$$x_{2,k+1} = -x_{2,k} = (-1)^k \cdot x_{2,0} = (-1)^k \cdot b$$

Άρα η μέθοδος θα ταλαντώνει μεταξύ των (0, b) και (0, -b) επ' άπειρον. Με αντίστοιχο τρόπο προκύπτει το ίδιο αποτέλεσμα και για γ = 3 για αρχικά σημεία (a, 0), με διαφορά την απόκλιση της μεθόδου για κάθε άλλη επιλογή αρχικού σημείου, καθώς η νέα συνθήκη σύγκλισης θα είναι αδύνατο να ικανοποιηθεί για οποιοδήποτε x2.

Τέλος, για οποιαδήποτε τιμή του γ ανάμεσα στο 3 και στο 1/3, η ανισότητα σύγκλισης προσδιορίζει μία περιοχή στον δισδιάστατο χώρο (x1, x2), στο εσωτερικό της οποίας ανήκουν όλα τα σημεία που οδηγούν σε σύγκλιση.

III. ΕΙΣΑΓΩΓΗ ΓΙΑ ΤΗΝ ΜΕΘΟΔΟ ΜΕΓΙΣΤΗΣ ΚΑΘΟΔΟΥ ΜΕ ΠΡΟΒΟΛΗ

Η μέθοδος της Μέγιστης Καθόδου με Προβολή είναι μια επέκταση της προηγούμενης για προβλήματα βελτιστοποίησης με περιορισμούς. Ενώ η κλασική μέθοδος αναζητά το ελάχιστο σε ολόκληρο τον χώρο χωρίς να λαμβάνει υπόψη περιορισμούς, η μέθοδος με προβολή αναζητά το ελάχιστο εντός ενός συγκεκριμένου υποσυνόλου.

Στην περίπτωσή μας, το σύνολο X ορίζεται από τους περιορισμούς:

$$X = \{x \in \mathbb{R}^2 : -10 \leq x_1 \leq 5, -8 \leq x_2 \leq 12\}$$

Και ακολουθεί την επαναληπτική σχέση:

$$x_{k+1} = x_k + \gamma_k(\overline{x_k} - x_k)$$

Όπου γk το μέγεθος του βήματος και $\bar{x}_k \epsilon X$ το εφικτό σημείο που υπολογίζεται μέσω της σχέσης:

$$\bar{x}_k = P_{rX}\big(x_k - s_k \nabla f(x_k)\big)$$

Με sk > 0 και Prx(z) είναι η προβολή του σημείου z στο υποσύνολο X. Έτσι εξασφαλίζουμε ότι κάθε σημείο θα παραμένει εντός του X για οποιοδήποτε k. Η μέθοδος τερματίζει όταν φτάσουμε σε σημείο x*, για το οποίο ισχύει:

$$\nabla f^T(x^*)(x - x^*) \geq 0, \;\; \forall x \in X$$

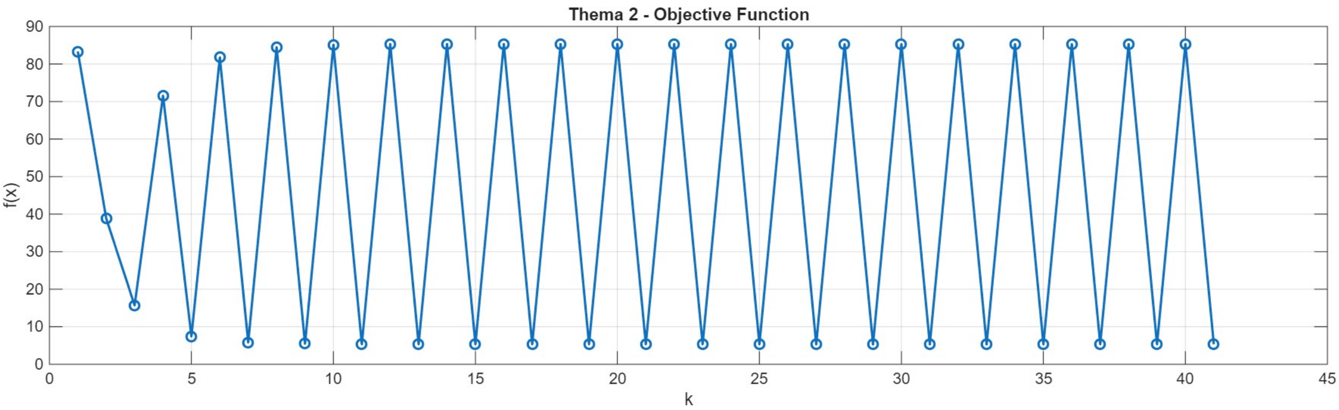
Για περιορισμούς τύπου κουτιού συγκεκριμένα ([a, b]), η προβολή μπορεί να υπολογιστεί ως:

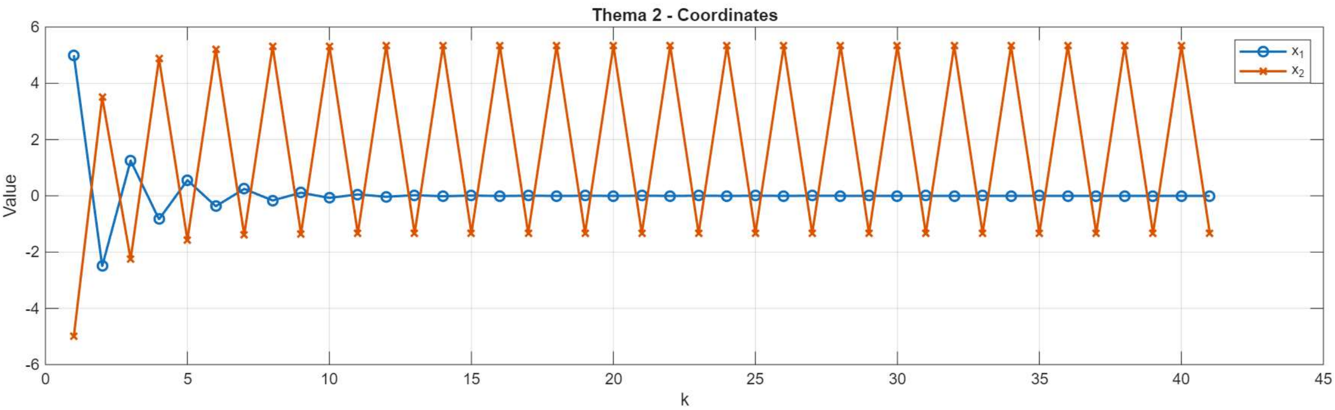
$$[P_{rX}(z)]_i = min(max(z_i, a_i), b_i)$$

IV. ΘΕΜΑ 2

Εφαρμόσαμε τη Μέθοδο με Προβολή με τις εξής παραμέτρους:

- $s_k = 5, \gamma_k = 0.5$
- Αρχικό Σημείο: x = (5, -5)





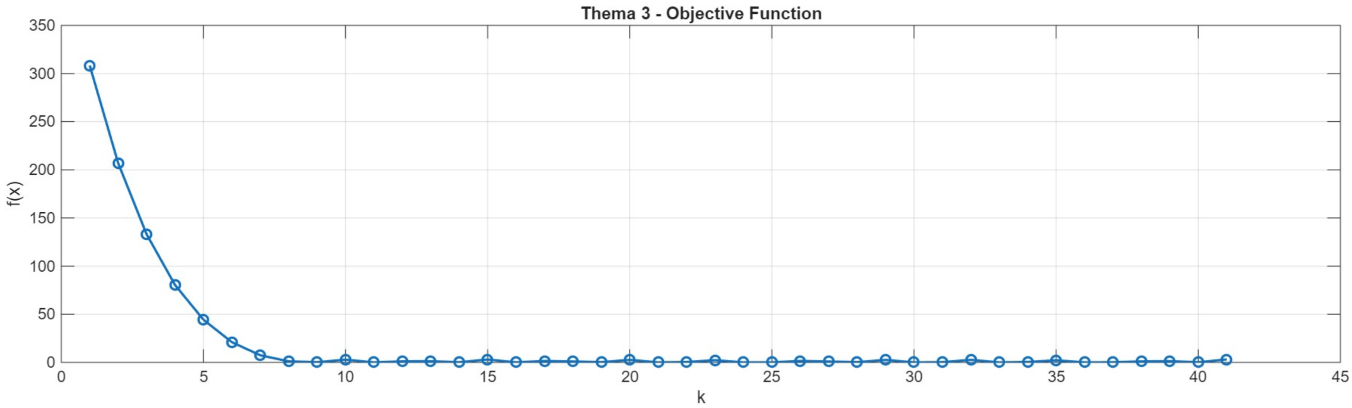
Παρατηρούμε ότι, ενώ η συντεταγμένη x1 συγκλίνει στο 0, η x2 ταλαντώνεται μεταξύ των τιμών −1.33 και 5.33. Στις πρώτες επαναλήψεις η x2 προσεγγίζει αυτές τις δύο τιμές και έπειτα εναλλάσσεται μεταξύ τους ατέρμονα. Αντίστοιχα και το ύψος του "υποψήφιου" ελαχίστου (δηλαδή η τιμή της f(x1,x2)) ταλαντώνεται μεταξύ 5.33 και 85.33. Να σημειωθεί ότι πανομοιότυπη συμπεριφορά βλέπουμε και για μεγαλύτερο αριθμό επαναλήψεων, καθώς έγιναν πειράματα για τιμές max_iters = {100, 500, 1000, 2000} χωρίς να παρουσιαστεί οποιαδήποτε διαφορά.

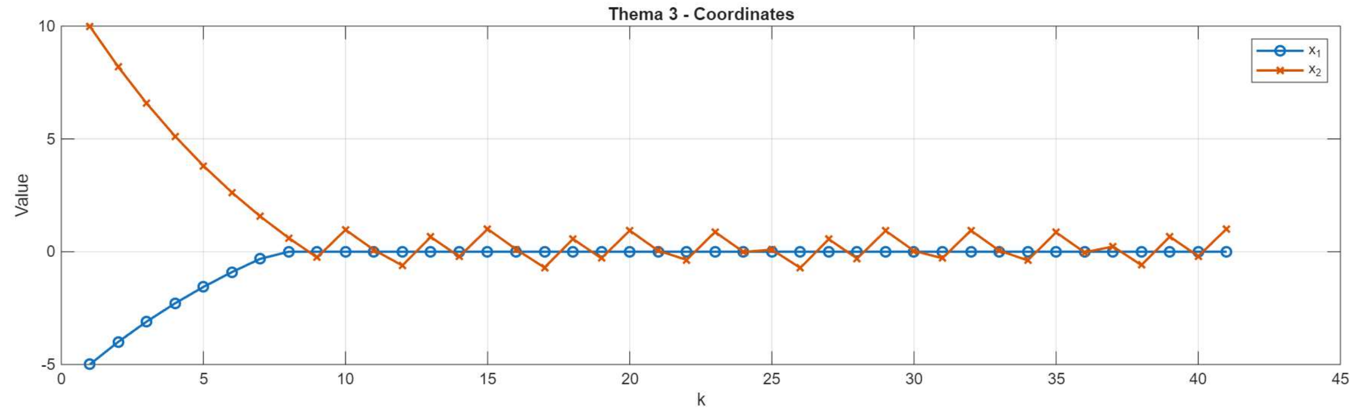
Από τις παραπάνω παρατηρήσεις καταλαβαίνουμε ότι η συμπεριφορά της τροποποιημένης μεθόδου είναι αρκετά διαφορετική της αρχικής. Με βήμα 0.5 και αρχικό σημείο (-5, 5), στην περίπτωση χωρίς περιορισμούς θα περιμέναμε να αποκλείσουμε. Αντίθετα όμως, επειδή η νέα μέθοδος μας επιβάλει να μείνουμε εντός του χώρου που ορίζουν οι περιορισμοί, η τιμή της εγκλωβίζεται και ταλαντεύεται εντός του εφικτού συνόλου.

V. ΘΕΜΑ 3

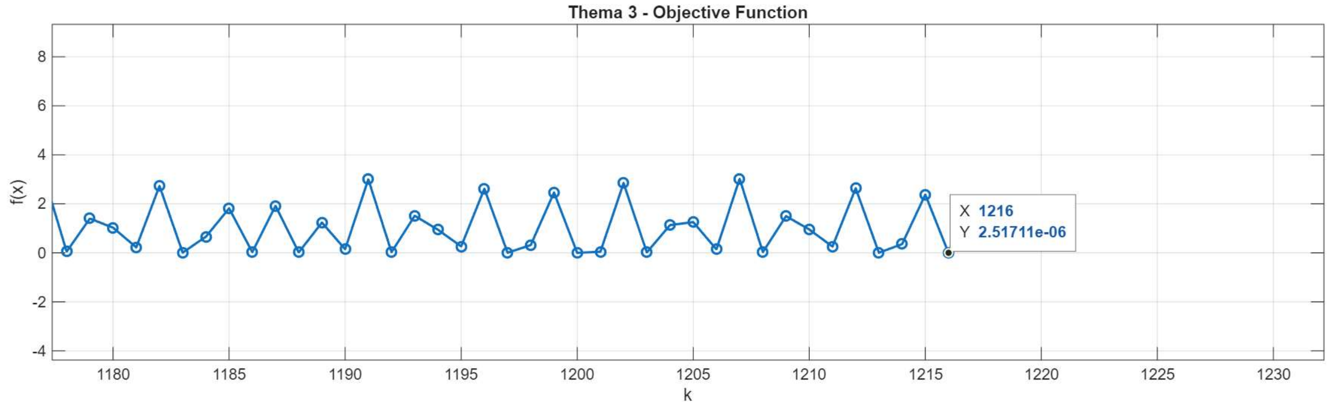
Εφαρμόσαμε τη Μέθοδο με Προβολή με τις εξής παραμέτρους:

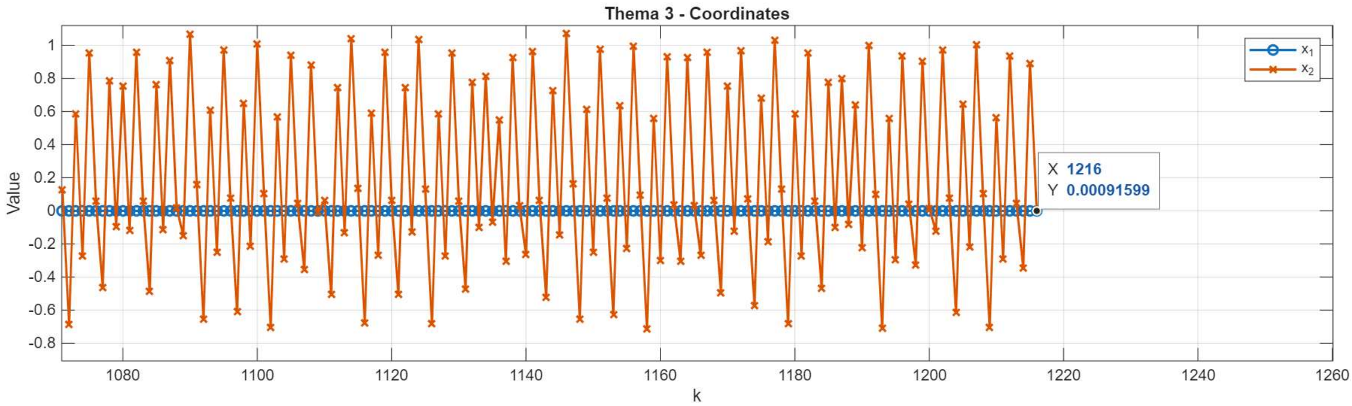
- $s_k = 15, \gamma_k = 0.1$
- Αρχικό Σημείο: x = (-5, 10)



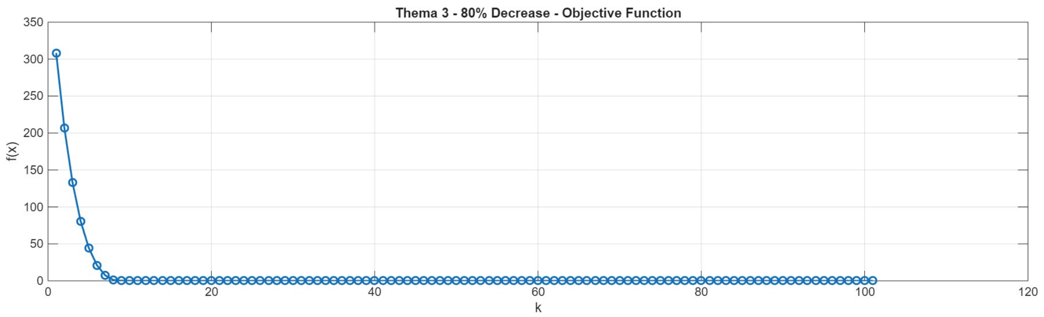


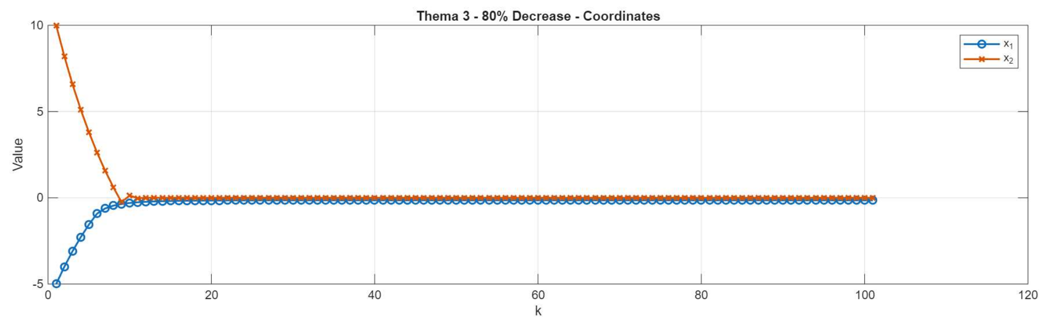
Όπως στο Θέμα 2, η συντεταγμένη x1 συγκλίνει στο 0. Με διαφορά, την x2 που δεν εναλλάσσεται μεταξύ δύο σημείων, αλλά μεταξύ περισσότερων. Η συνάρτησης f συγκλίνει επιθυμητά στο 0 αλλά χωρίς να επιτυγχάνει την απαιτούμενη ακρίβεια, καθώς κάνει φαινομενικά χαοτικές ελλειπτικές τροχιές γύρω από το πραγματικό ελάχιστο. Το φαινόμενο αυτό φαίνεται ωστόσο να τερματίζει σε 1216 επαναλήψεις όπως βλέπουμε και παρακάτω.

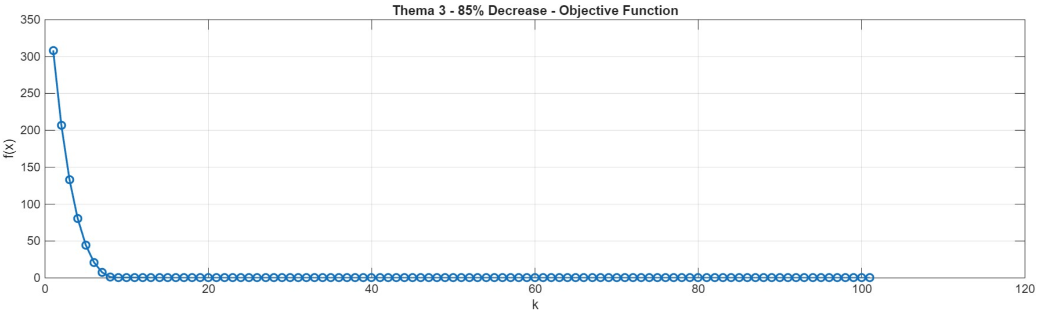


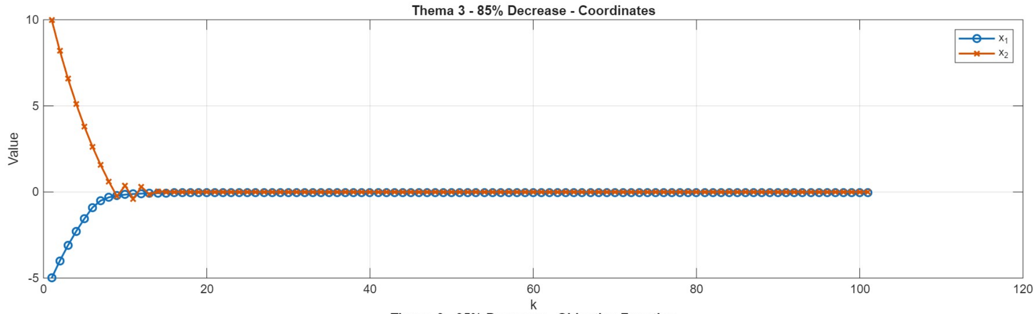


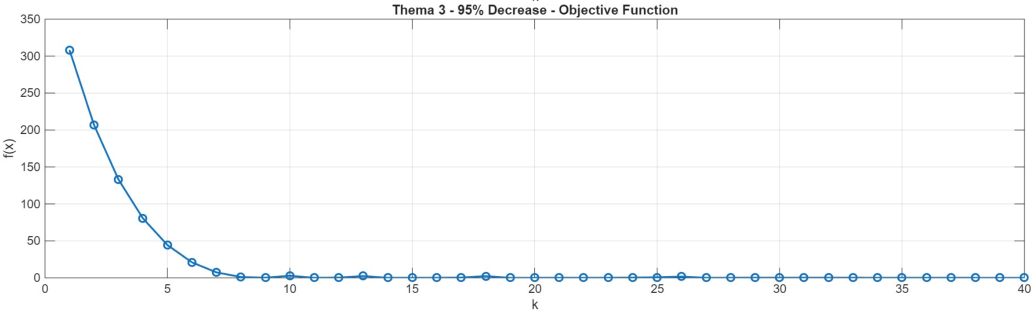
Μία καλή στρατηγική για σύγκλιση, θα ήταν η τροποποίηση των τιμών των βημάτων sk και γk. Συγκεκριμένα, εφαρμόσαμε σταδιακή μείωση του sk σε κάθε επανάληψη. Δοκιμάστηκαν οι εξής τιμές του συντελεστή μείωσης {0.80, 0.85, 0.90, 0.95}:

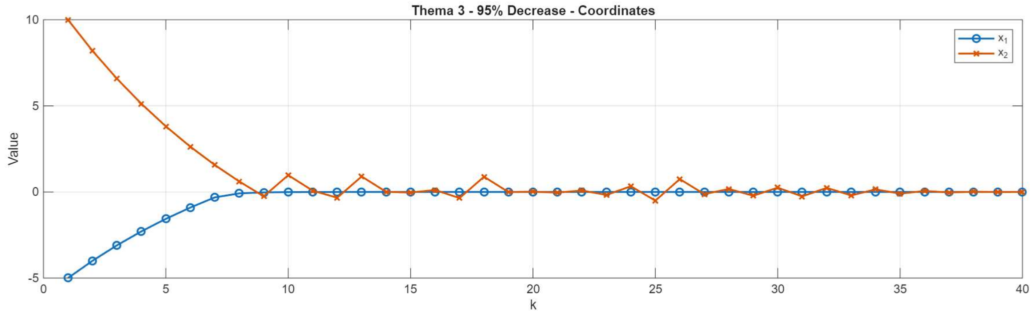




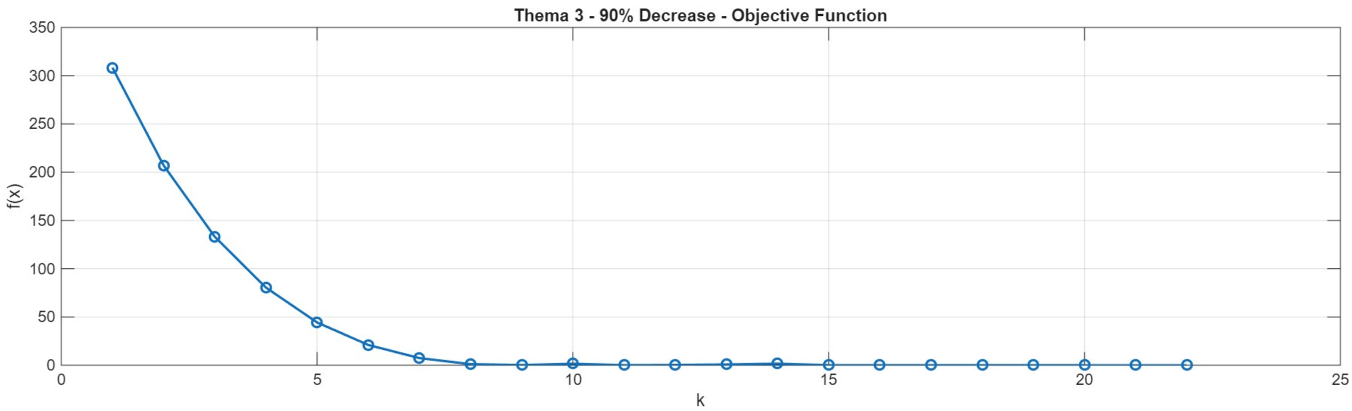


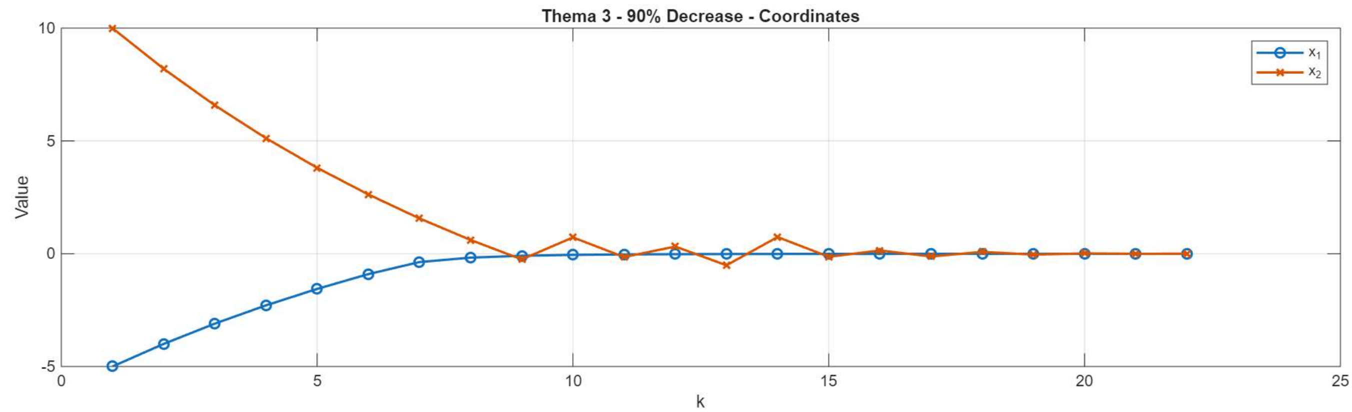






Παρατηρούμε ότι για τις τιμές συντελεστή μείωσης < 0.9 η μέθοδος συνεχίζει να μην συγκλίνει. Για μεγαλύτερες τιμές όμως η μέθοδος τείνει στο ελάχιστο και ικανοποιεί την ακρίβεια που ζητάμε. Επιλέγουμε λοιπόν **0.90 ή 90%** μείωση διότι μας προσφέρει την ταχύτερη σύγκλιση σε σχέση με το 0.95:

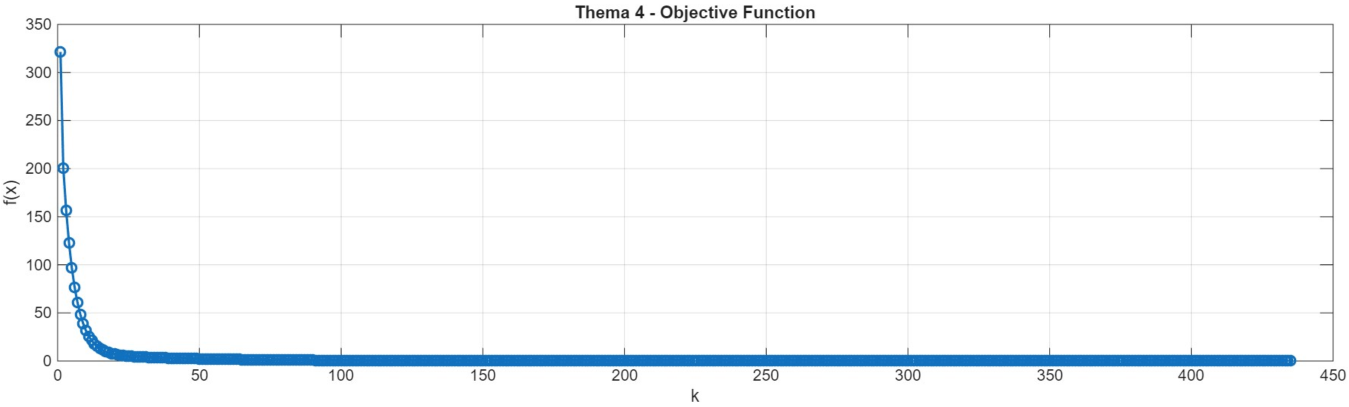


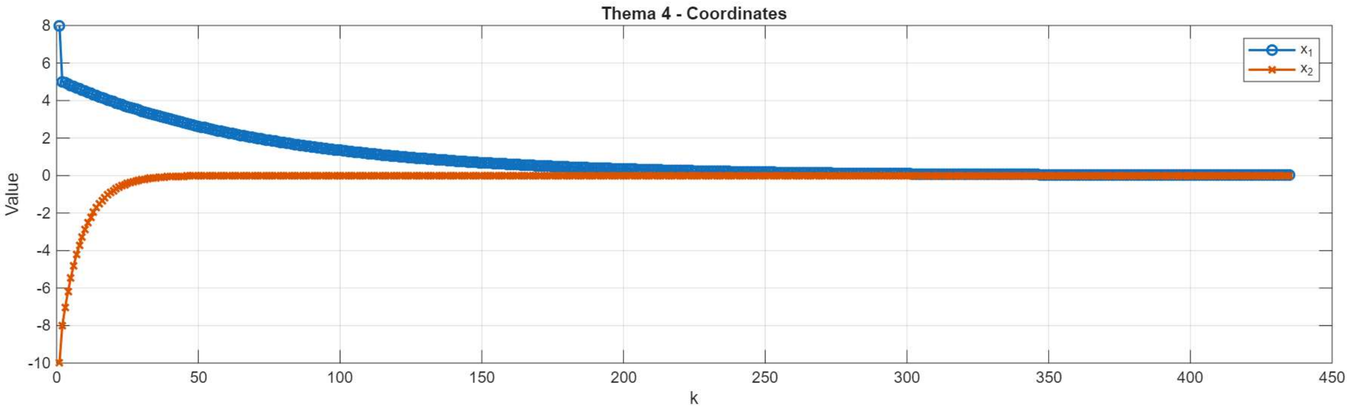


VI. ΘΕΜΑ 4

Εφαρμόσαμε τη Μέθοδο με Προβολή με τις εξής παραμέτρους:

- $s_k = 0.1, \gamma_k = 0.2$
- Αρχικό Σημείο: x = (8, -10)





Στο Θέμα 4 έχουμε μη εφικτό σημείο εκκίνησης, καθώς η συντεταγμένη x1 είναι μεγαλύτερο του 5. Παρόλο που το x1 δεν βρίσκεται εντός αρχικά των περιορισμών μας, παρατηρούμε ότι ο αλγόριθμος συγκλίνει επιτυχώς στο ελάχιστο και τερματίζει μέσα σε 435 επαναλήψεις. Από την πρώτη κιόλας επανάληψη, η μέθοδος επιβάλει στο x1 να μπει εντός του X, παίρνοντας την οριακή τιμή x1 = 5, επιδεικνύοντας έτσι ανοχή σε σφάλματα αρχικοποίησης, καθώς διορθώνει αυτόματα μη εφικτά αρχικά σημεία. Μετά την αρχική διόρθωση, η ακολουθία των σημείων φαίνεται να συγκλίνει ομαλά και ασυμπτωτικά στο ολικό ελάχιστο (0,0).