3η Εργαστηριακή Άσκηση

Χρυσολόγου Γεώργιος (ΑΕΜ: 10782)

Περιγραφή προβλήματος: Για μία δοσμένη συνάρτηση f δύο μεταβλητών ($f: R^2 \to R$),η οποία είναι παραγωγίσιμη, καλούμαστε να εφαρμόσουμε δύο μεθόδους κλίσης, μία χωρίς περιορισμούς και μία με την θεώρηση περιορισμών, με σκοπό την ελαχιστοποίηση της. Η πρώτη πρόκειται για την Μέθοδο Μέγιστης Καθόδου και η δεύτερη είναι η Μέθοδος Μέγιστης Καθόδου με προβολή. Η συνάρτηση που ζητείται να ελαχιστοποιηθεί είναι η εξής: $f(x) = \frac{1}{3}x_1^2 + 3x_2^2$, $x = [x1\ x2]^T$

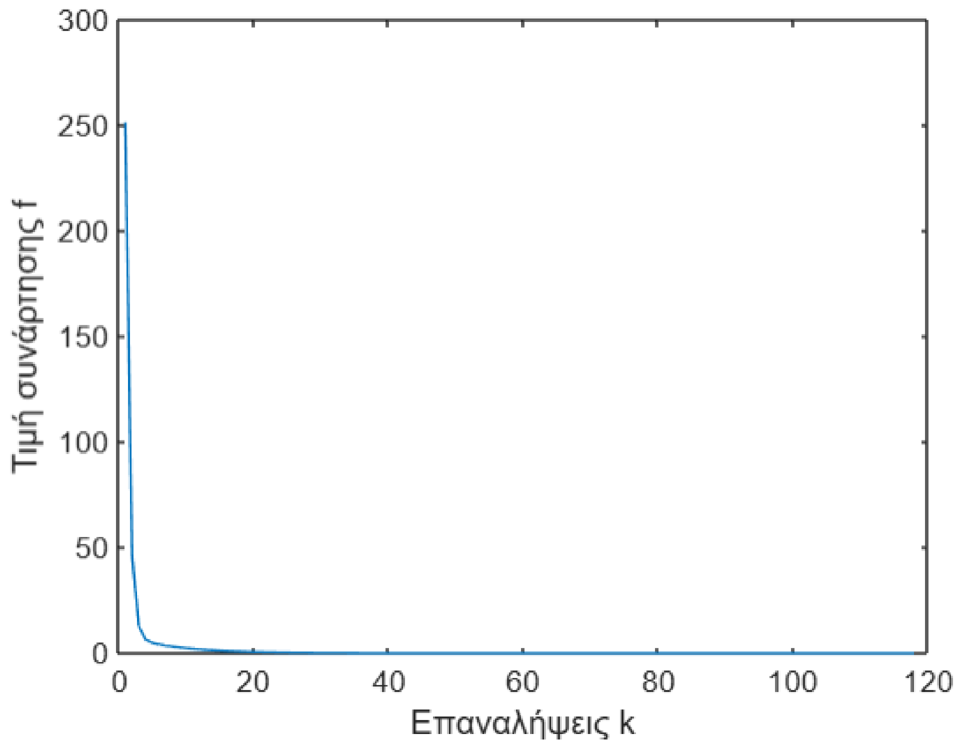
Θέμα 1 (Μέθοδος Μέγιστης Καθόδου)

Η εξήγηση του τρόπου λειτουργίας της μεθόδου αυτής παραλείπεται καθώς αποτέλεσε αντικείμενο ανάλυσης της προηγούμενης εργαστηριακής άσκησης.

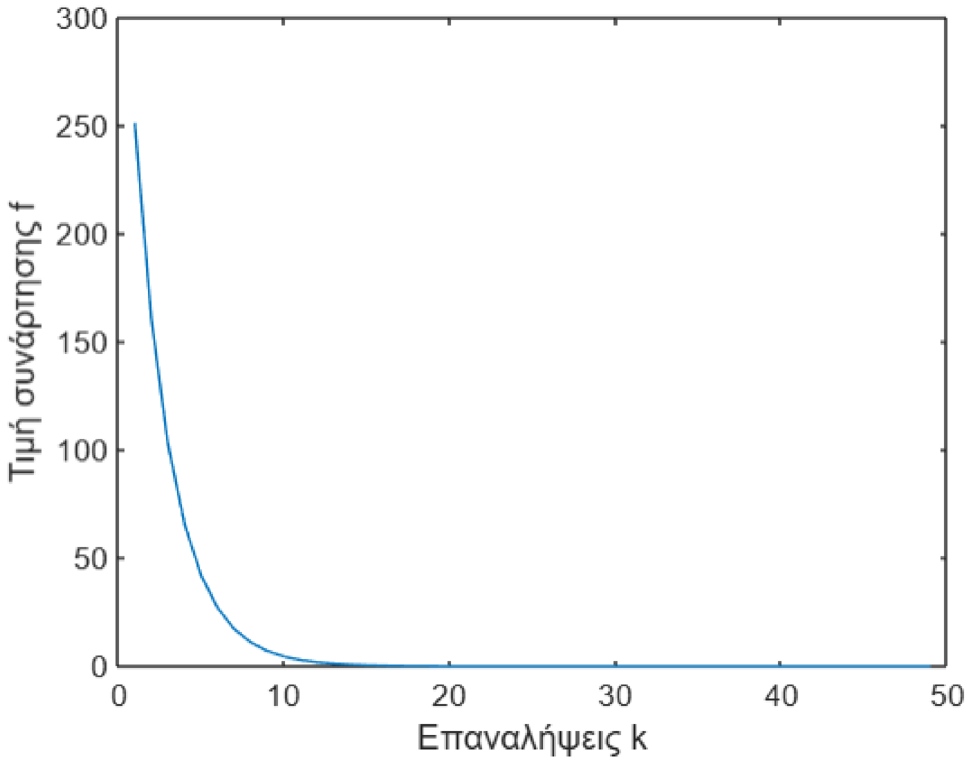
Σύμφωνα με την εκφώνηση της άσκησης, ορίστηκε η σταθερά τερματισμού $\varepsilon = 0.001$ ενώ επιλέχθηκε σταθερό βήμα $\gamma_k = \gamma$, με $\gamma = 0.1,\ 0.3,\ 3,\ 5$. Ως σημείο έναρξης χρησιμοποιήθηκε το σημείο (5, -9) (διαφορετικό του (0, 0)).

Ακολουθούν οι γραφικές παραστάσεις της σύγκλισης της αντικειμενικής συνάρτησης προς τον αριθμό των επαναλήψεων για τις διαφορετικές τιμές του σταθερού βήματος γ_k.

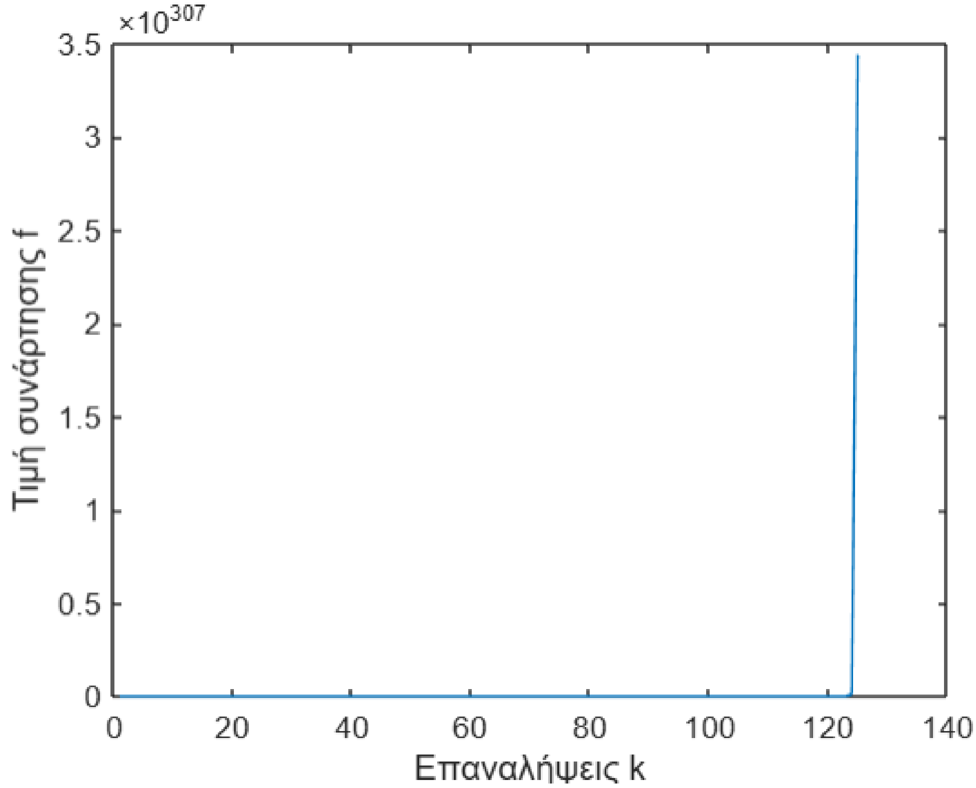
i) $\boldsymbol{\gamma_k = 0.1}$



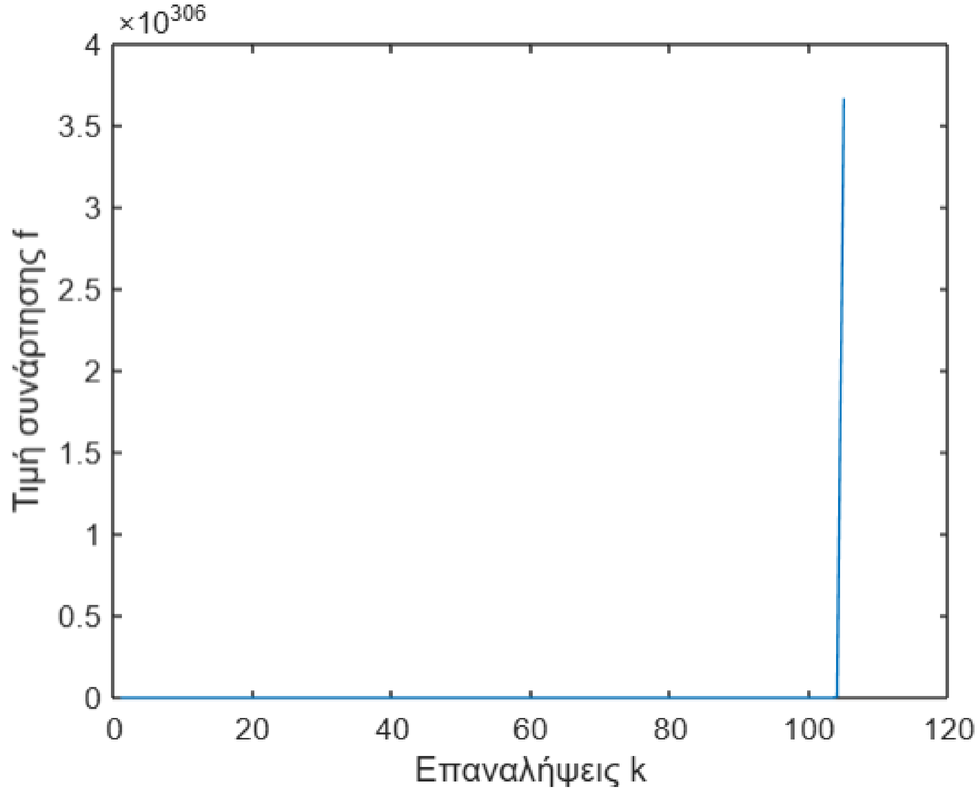
ii) $\boldsymbol{\gamma_k = 0.3}$



iii) $\boldsymbol{\gamma_k = 3}$



iv) $\gamma_k = 5$



Παρατηρούμε ότι, μεταξύ των δύο τιμών του σταθερού βήματος γ_k που οδηγούν σε σύγκλιση της μεθόδου, καλύτερη είναι η $\gamma_k = 0.3$ καθώς απαιτεί αρκετά λιγότερες επαναλήψεις για τον τερματισμό του αλγορίθμου. Αντίθετα με τις δύο παραπάνω επιλογές τιμών του βήματος, οι τιμές $\gamma_k = 3$ και $\gamma_k = 5$ έχουν ως συνέπεια την αποτυχία της μεθόδου να συγκλίνει προς το ελάχιστο της αντικειμενική συνάρτησης. Στις δύο αυτές περιπτώσεις, το πρόγραμμα στο MATLAB τερματίζεται όταν η τιμή της αντικειμενικής συνάρτησης προσεγγίζει το άπειρο.

Ακολουθεί μαθηματική ανάλυση κατά την οποία υπολογίζονται τα όρια της τιμής του γ_k ώστε η μέθοδος να συγκλίνει:

$f(x) = \frac{1}{3}x_1^2 + 3x_2^2$, $x = [x1, x2]^T$
$\nabla f(x) = [\frac{2}{3}x_1,\ 6x_2]$

Για την Μέθοδο Μέγιστης Καθόδου ισχύει:

$$x_{k+1} = x_k - \gamma_k * \nabla f(x_k) \Rightarrow [x_{1\,k+1},\ x_{2\,k+1}] = [x_{1\,k},\ x_{2\,k}] - \gamma_k * [\frac{2}{3}x_{1\,k},\ 6x_{2\,k}]$$

Άρα:

$x_{1\,k+1} = x_{1\,k} - \gamma_k * \frac{2}{3}x_{1\,k} \Rightarrow x_{1\,k+1} = x_{1\,k}\,(1 - \frac{2}{3}\gamma_k)$ **(1)**

$x_{2\,k+1} = x_{2\,k} - \gamma_k * 6x_{2\,k} \Rightarrow x_{2\,k+1} = x_{2\,k}\,(1 - 6\gamma_k)$ **(2)**

Το ελάχιστο της αντικειμενικής συνάρτησης βρίσκεται στο (0, 0). Επομένως, το σημείο που εξετάζεται στην επανάληψη k+1 πρέπει να βρίσκεται πιο κοντά στο (0, 0) σε σχέση με το σημείο που εξετάστηκε στην επανάληψη k, ώστε η μέθοδος να συγκλίνει.

Άρα:
$\frac{|x_{k+1}|}{|x_k|} < 1$

Η ανισότητα αυτή θα ικανοποιείται με απόλυτη βεβαιότητα εάν ισχύουν ταυτόχρονα οι παρακάτω σχέσεις:

$$\frac{|x_{1\,k+1}|}{|x_{1\,k}|} < 1 \Rightarrow^{(1)} \frac{|\,x_{1\,k}\,(1-\ \frac{2}{3}\gamma_k)\ |}{|x_{1\,k}|} < 1 \Rightarrow \left|1-\ \ \frac{2}{3}\gamma_k\right| < 1 \Rightarrow 0 < \gamma_k < 3$$

$$\frac{|x_{2\,k+1}|}{|x_{2\,k}|} < 1 \Rightarrow^{(2)} \frac{|\,x_{2\,k}\,(1-\ 6\gamma_k)\ \ |}{|x_{2\,k}|} < 1 \Rightarrow |1-\ \ 6\gamma_k| < 1 \Rightarrow 0 < \gamma_k < \frac{1}{3}$$

Οι δύο σχέσεις ισχύουν ταυτόχρονα για $0 < \gamma_k < \frac{1}{3}$.

Τα αποτελέσματα των δοκιμών για τις τέσσερεις διαφορετικές τιμές του σταθερού βήματος γ_k συμφωνούν με το συμπέρασμα της παραπάνω μαθηματικής ανάλυσης. Οι τιμές $\gamma_k = 0.1$ και $\gamma_k = 0.3$, οι οποίες βρίσκονται εντός των ορίων $0 < \gamma_k < \frac{1}{3}$, οδηγούν σε σύγκλιση της μεθόδου, σε αντίθεση με τις τιμές $\gamma_k = 3$ και $\gamma_k = 5$, οι οποίες βρίσκονται εκτός των ορίων $0 < \gamma_k < \frac{1}{3}$.

Για τα τρία παρακάτω θέματα χρησιμοποιείται η Μέθοδος Μέγιστης Καθόδου με προβολή, ενώ έχουν θεωρηθεί οι εξής περιορισμοί:
$\boldsymbol{-10 \le x_1 \le 5}$ **και** $\boldsymbol{-8 \le x_2 \le 12}$

Το σύνολο που ορίζεται από τους περιορισμούς ονομάζεται $\boldsymbol{X}$.

Εξήγηση τρόπου λειτουργίας μεθόδου: Σύμφωνα με την μέθοδο αυτή, επιλέγεται, αρχικά, ένα σημείο ως σημείο έναρξης και ορίζεται η τιμή της σταθεράς τερματισμού ε. Σε κάθε επανάληψη k του αλγορίθμου, υπολογίζεται ένα νέο σημείο x_{k+1} βάσει του αντίστοιχου σημείου της επανάληψης k, x_k, σύμφωνα με την σχέση $x_{k+1} = x_k + \gamma_k * (\overline{x_k} - x_k)$, όπου $\overline{x_k} = \Pr_X(x_k - s_k * \nabla f(x_k))$. Αναλυτικότερα, το $(x_k - s_k * \nabla f(x_k))$ πρόκειται για το σημείο που προκύπτει από την κίνηση προς την κατεύθυνση της αρνητικής κλίσης με βήμα s_k ξεκινώντας από το σημείο x_k. Το $\overline{x_k}$ υπολογίζεται ως η προβολή του $(x_k - s_k * \nabla f(x_k))$ στο X. Κατόπιν, το σημείο x_{k+1} υπολογίζεται με κίνηση από το σημείο x_k προς την εφικτή κατεύθυνση $\overline{x_k} - x_k$ με βήμα γ_k. Στην αρχή κάθε επανάληψης ελέγχεται αν ικανοποιείται η ανισότητα $|\nabla f(x_k)| < \varepsilon$, γεγονός που οδηγεί στον τερματισμό της μεθόδου καθώς έχει καταλήξει κοντά σε στάσιμο σημείο.

Θέμα 2

Στο θέμα αυτό ζητείται η εφαρμογή της Μεθόδου Μέγιστης Καθόδου με προβολή, με $s_k = 5$, $\gamma_k = 0.5$, σημείο εκκίνησης το (5, -5), το οποίο αποτελεί εφικτό σημείο, και σταθερά τερματισμού $\varepsilon = 0.01$.

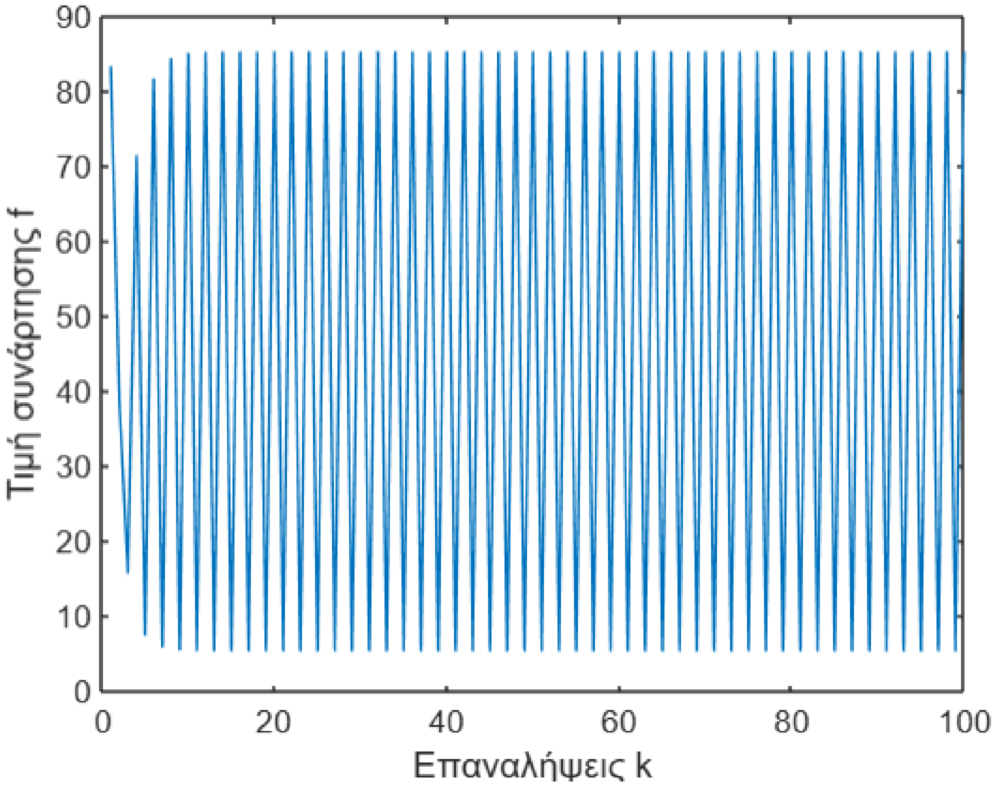
Η εκτέλεση του προγράμματος στο MATLAB δεν τερματίζει καθώς εμφανίζεται ατέρμονας βρόγχος. Αυτό οφείλεται στο γεγονός ότι η μέθοδος δεν επιτυγχάνει να συγκλίνει προς κάποιο στάσιμο σημείο της αντικειμενικής συνάρτησης, σε αντίθεση με το Θέμα 1. Συγκεκριμένα, η τιμή της αντικειμενικής συνάρτησης σε κάθε επανάληψη ταλαντώνεται μεταξύ των τιμών 5.33 και 85.33. Ο συνδυασμός των δύο επιλογών για τις τιμές των βημάτων γ_k και s_k ευθύνεται για αυτήν την αποτυχία σύγκλισης του αλγορίθμου. Το συμπέρασμα αυτό επιβεβαιώνεται από την σημείωση του βιβλίου στην οποία αναφέρεται ότι όταν το σημείο $(x_k - s_k * \nabla f(x_k))$ είναι εφικτό, η Μέθοδος Μέγιστης Καθόδου με προβολή ανάγεται στην Μέθοδο Μέγιστης Καθόδου χωρίς περιορισμούς όπου το x_{k+1} υπολογίζεται, μετά από πράξεις, μέσω της σχέσης $x_{k+1} = x_k - \gamma_k * s_k * \nabla f(x_k)$. Δεδομένου ότι στην Μέθοδο Μέγιστης Καθόδου ισχύει η σχέση $x_{k+1} = x_k - \gamma_k * \nabla f(x_k)$ και σύμφωνα με την μαθηματική ανάλυση που προηγήθηκε στο Θέμα 1, γίνεται αντιληπτό ότι για την σύγκλιση της Μεθόδου Μέγιστης Καθόδου με προβολή απαιτείται να ισχύει $0 < \gamma_k * s_k < \frac{1}{3}$. Είναι σημαντικό να ληφθεί υπόψιν ότι το ολικό ελάχιστο (0, 0) αποτελεί εφικτό σημείο οπότε η μέθοδος θα αναζητήσει το ελάχιστο εντός του συνόλου X και θα αναχθεί στην Μέθοδο Μέγιστης Καθόδου χωρίς περιορισμούς.

Ο ισχυρισμός αυτός συμφωνεί με τα αποτελέσματα δύο ειδών δοκιμών εφαρμογής της μεθόδου με τις ίδιες τιμές για τις παραμέτρους, με μόνη διαφορά τον ορισμού του s_k και του γ_k, αντίστοιχα. Η υλοποίηση της μεθόδου με $s_k = 0.5$ οδηγεί σε σύγκλιση της μεθόδου προς το ολικό ελάχιστο (0, 0). Το ίδιο αποτέλεσμα σημειώνεται και από την υλοποίηση με $\gamma_k = 0.02$.

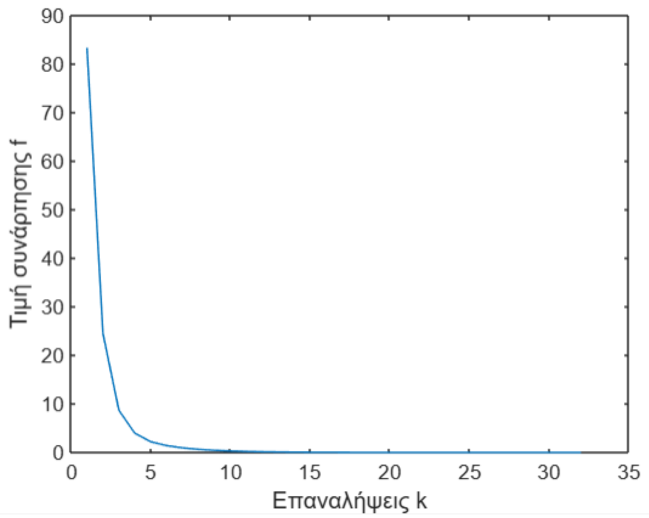
Ακόμη, σύγκλιση παρατηρείται και με τον ταυτόχρονο ορισμό $s_k = 0.5$ και $\gamma_k = 0.02$.

Ως συμπέρασμα, ο συνδυασμός των τιμών $s_k = 5$, $\gamma_k = 0.5$ οδηγεί την μέθοδο στην αποτυχία να δώσει μία επιθυμητή λύση.

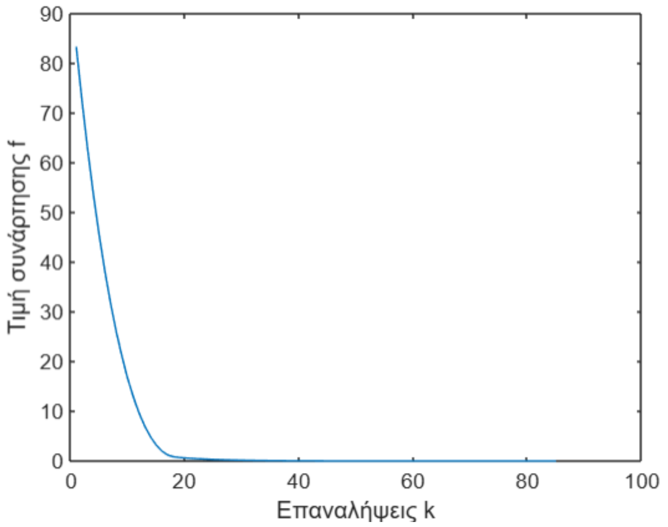
$s_k = 5, \gamma_k = 0.5$ (τερματισμός στις 100 επαναλήψεις):



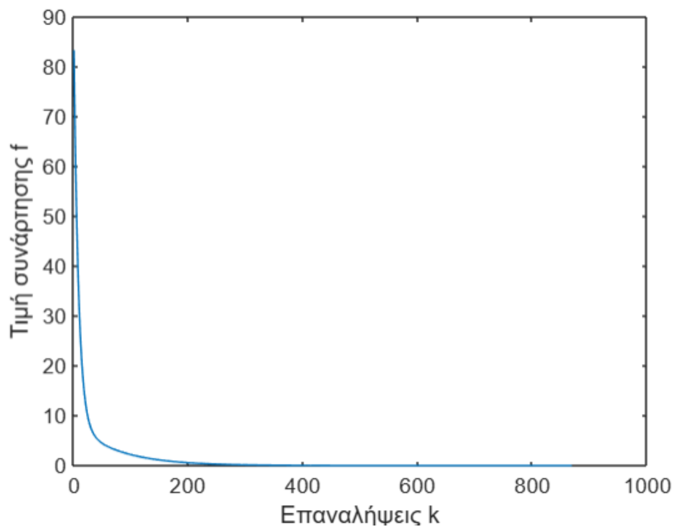
$s_k = 0.5, \gamma_k = 0.5$



$s_k = 5, \gamma_k = 0.02$



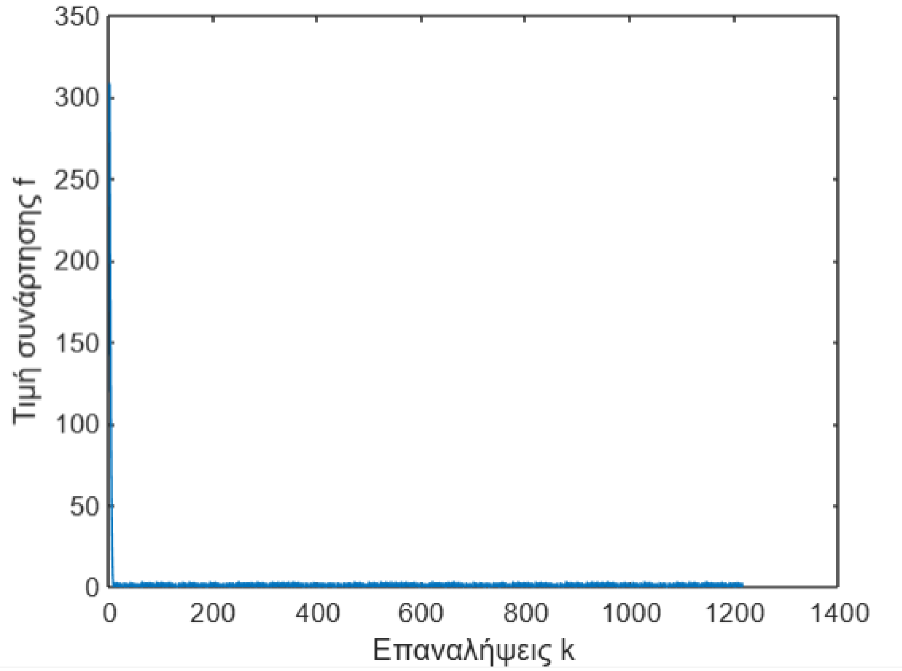
$s_k = 0.5, \gamma_k = 0.02$



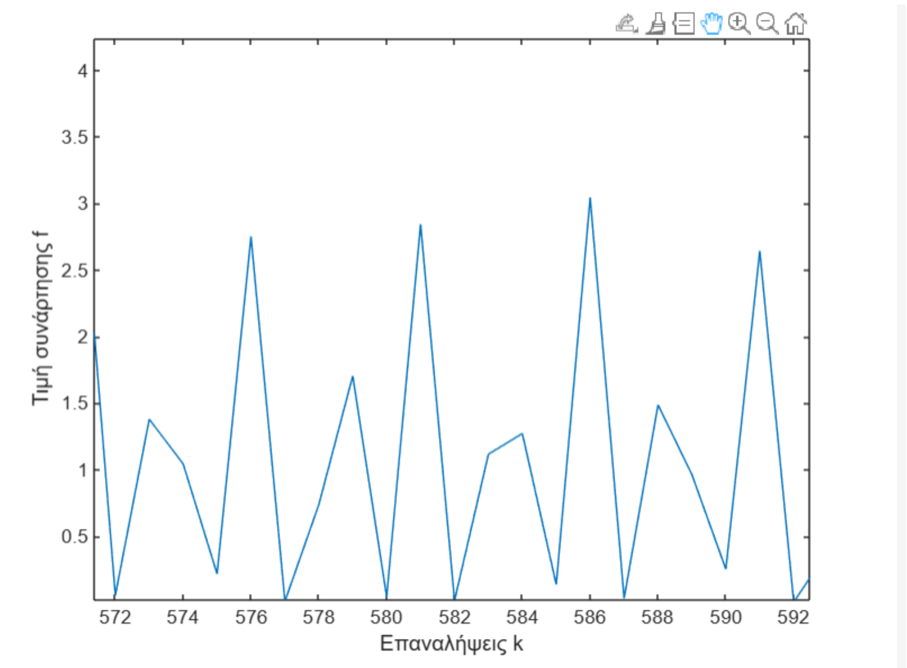
Θέμα 3

Στο θέμα αυτό ζητείται η εφαρμογή της Μεθόδου Μέγιστης Καθόδου με προβολή, με $s_k = 15$, $\gamma_k = 0.1$, σημείο εκκίνησης το (-5, 10), το οποίο αποτελεί εφικτό σημείο, και σταθερά τερματισμού $\varepsilon = 0.01$.

Η γραφική παράσταση της σύγκλισης της αντικειμενικής συνάρτησης προς τον αριθμό των επαναλήψεων είναι η εξής:



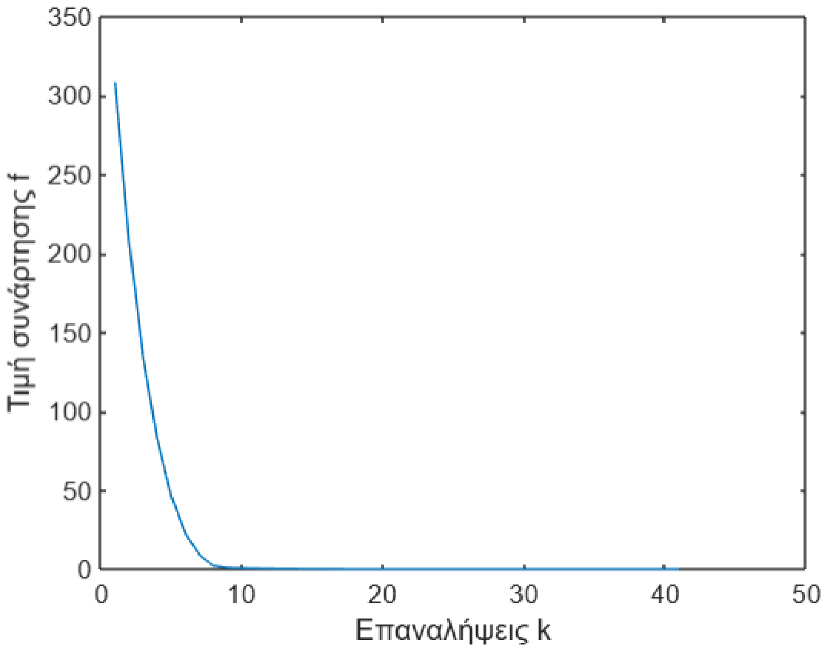
Η εικόνα που ακολουθεί αποτελεί μεγέθυνση της παραπάνω, στο τμήμα της γραφικής που φαίνεται σχεδόν παράλληλο με τον οριζόντιο άξονα των επαναλήψεων k:



Η μέθοδος, για αυτές τις τιμές των παραμέτρων της, τελικά επιτυγχάνει σύγκλιση στο ολικό ελάχιστο (0, 0). Συγκεκριμένα, δίνεται η λύση $(0\,, 9.15 * 10^{-4})$ η οποία προσεγγίζει την πραγματική λύση του προβλήματος ελαχιστοποίησης. Ωστόσο, παρατηρούμε, όπως είναι εμφανές από την 2η εικόνα, ότι σημειώνεται ταλάντωση μεταξύ των τιμών της αντικειμενικής συνάρτησης $f(x_k)$ για κάθε επανάληψη k.

Αυτό αποτελεί συνέπεια κακής λειτουργίας της μεθόδου καθώς δεν εξασφαλίζεται ότι η τιμή $|f(x_{k+1})|$ είναι μικρότερη από την τιμή $|f(x_k)|$. Επιπλέον, ο συνολικός αριθμός επαναλήψεων προκύπτει πολύ μεγάλος. Παρά το γεγονός, λοιπόν, ότι, σε αντίθεση με το Θέμα 2, η μέθοδος στο θέμα αυτό επιτυγχάνει σύγκλιση, η λειτουργία της δεν είναι η ενδεδειγμένη, όπως παρατηρείται στο Θέμα 1.

Έναν πρακτικό τρόπο ώστε η μέθοδος να συγκλίνει στο ελάχιστο, σε λιγότερες επαναλήψεις και χωρίς την εμφάνιση ταλαντώσεων, αποτελεί η επιλογή μικρότερου βήματος s_k, βάσει όσων αναφέρθηκαν στα παραπάνω θέματα. Συγκεκριμένα, η επιλογή της τιμής $s_k = 2$ ($\gamma_k * s_k = 0.1 * 2 = 0.2 < \frac{1}{3}$) οδηγεί σε ομαλή σύγκλιση της μεθόδου, δίνοντας ως τελική λύση το σημείο (-0.0142, 0) σε 41 επαναλήψεις. Η γραφική παράσταση της σύγκλισης της αντικειμενικής συνάρτησης προς τον αριθμό των επαναλήψεων για την ίδια δοκιμή με την διαφορά $s_k = 2$ είναι η εξής:

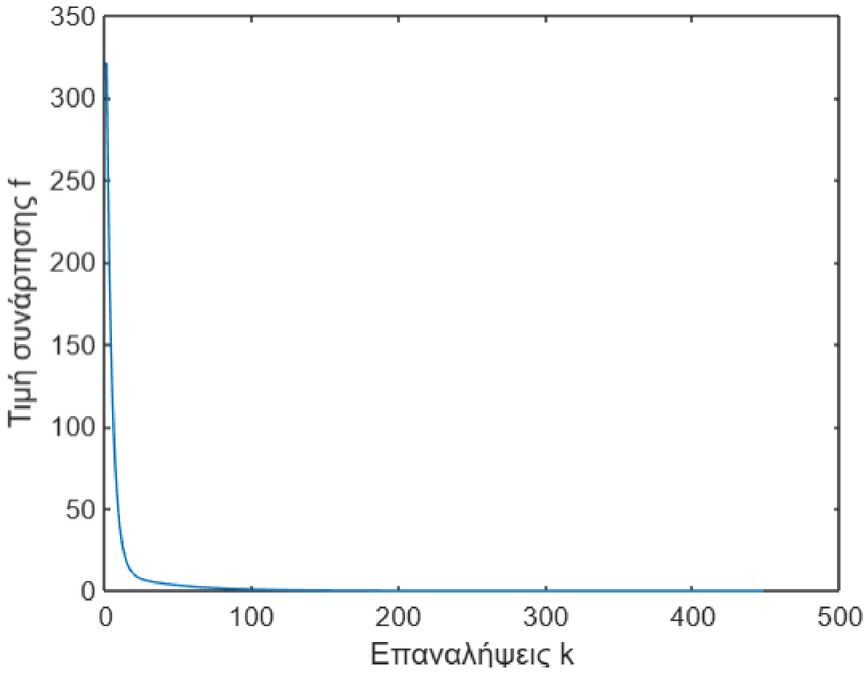


Θέμα 4

Στο θέμα αυτό ζητείται η εφαρμογή της Μεθόδου Μέγιστης Καθόδου με προβολή, με $s_k = 0.1$, $\gamma_k = 0.2$, σημείο εκκίνησης το (8, -10) και σταθερά τερματισμού $\varepsilon = 0.01$.

Παρατηρούμε ότι το σημείο εκκίνησης δεν αποτελεί εφικτό σημείο καθώς δεν ανήκει στο σύνολο X. Το γεγονός αυτό δεν αποτελεί πρόβλημα για την Μέθοδο Μέγιστης Καθόδου με προβολή, καθώς υπολογίζεται η προβολή του μη εφικτού σημείου στο X και η τελική λύση, αν η μέθοδος επιτύχει σύγκλιση, θα ανήκει, εγγυημένα, στο X. Δεδομένου ότι η επιλογή των βημάτων $s_k = 0.1$ και $\gamma_k = 0.2$ έχει ως αποτέλεσμα το γινόμενο τους να ικανοποιεί την σχέση $0 < \gamma_k * s_k < \frac{1}{3}$, είναι βέβαιο ότι η μέθοδος θα συγκλίνει προς το ελάχιστο (0, 0),αφού η αντικειμενική συνάρτηση δεν διαθέτει άλλο στάσιμο σημείο. Ωστόσο, επειδή οι τιμές θεωρούνται σχετικά μικρές και το σημείο εκκίνησης δεν βρίσκεται εντός του συνόλου X, αναμένεται η σύγκλιση να είναι αργή.

Η γραφική παράσταση της σύγκλισης της αντικειμενικής συνάρτησης προς τον αριθμό των επαναλήψεων είναι η εξής:



Παρατηρούμε ότι η μέθοδος συγκλίνει επιτυχώς προς το ολικό ελάχιστο (0, 0) παρά την επιλογή σημείου εκκίνησης που δεν ανήκει στο σύνολο X. Ακόμη, επιβεβαιώνεται η εκτίμηση ότι ο συνδυασμός μικρής τιμής των βημάτων s_k, γ_k και η εκκίνηση από μη εφικτό σημείο θα οδηγήσει σε αργή σύγκλιση, καθώς απαιτούνται 448 επαναλήψεις έως τον τερματισμό του αλγορίθμου.

Συμπεραίνουμε ότι η επιλογή μικρών τιμών για τα βήματα s_k και γ_k που πραγματοποιούνται στο θέμα αυτό επιτρέπουν στην μέθοδο να συγκλίνει ομαλά στο ελάχιστο, σε αντίθεση με τα προηγούμενα δύο θέματα στα οποία οι τιμές των βημάτων δεν ορίστηκαν κατάλληλα ώστε να επιτρέψουν την σωστή λειτουργία της μεθόδου.

Συνοψίζοντας, τόσο η Μέθοδος Μέγιστης Καθόδου, απουσία περιορισμών, όσο και η Μέθοδος Μέγιστης Καθόδου με προβολή, υπό την θεώρηση περιορισμών, απαιτούν κατάλληλη επιλογή των παραμέτρων τους ώστε να εξασφαλιστεί η ορθή λειτουργία τους. Για την πρώτη, το βήμα γ_k πρέπει να οριστεί μεταξύ συγκεκριμένων ορίων ώστε η μέθοδος να συγκλίνει. Αντίστοιχα, για την δεύτερη είναι απαραίτητος ο κατάλληλος συνδυασμός τιμών για τα βήματα s_k και γ_k ώστε η μέθοδος να επιτύχει ομαλή σύγκλιση, ανεξάρτητα από την επιλογή εφικτού ή μη εφικτού σημείου έναρξης του αλγορίθμου.