<u>ΤΕΧΝΙΚΕΣ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ</u>

3η Εργαστηριακή Άσκηση

Ονοματεπώνυμο: Φωτιάδης Αλέξανδρος

ΑΕΜ:10392

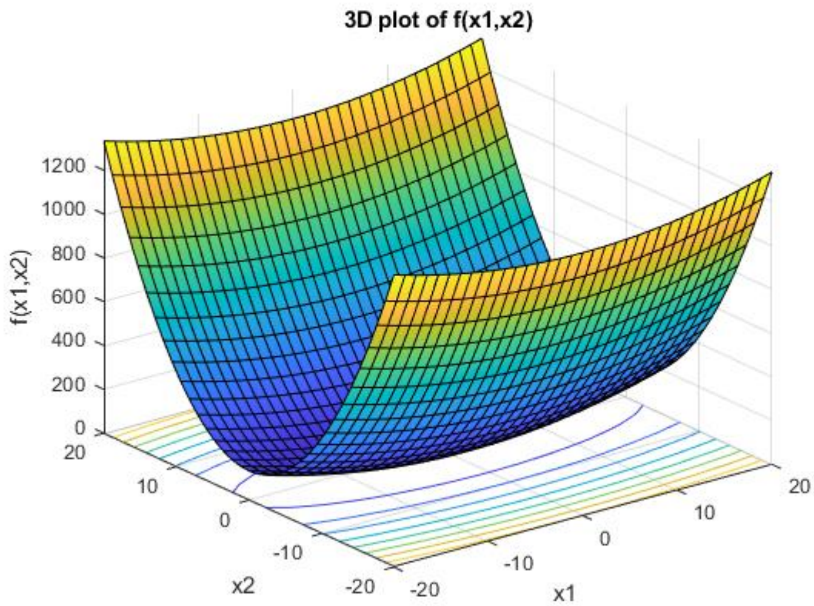
E-mail: afotiadis@ece.auth.gr

Σε αυτό το report παρουσιάζονται οι λύσεις για την 3η εργαστηριακή άσκηση του μαθήματος, τα ζητούμενα διαγράμματα και ο σχολιασμός και τα συμπεράσματα που εξάγονται από τα αποτελέσματα.

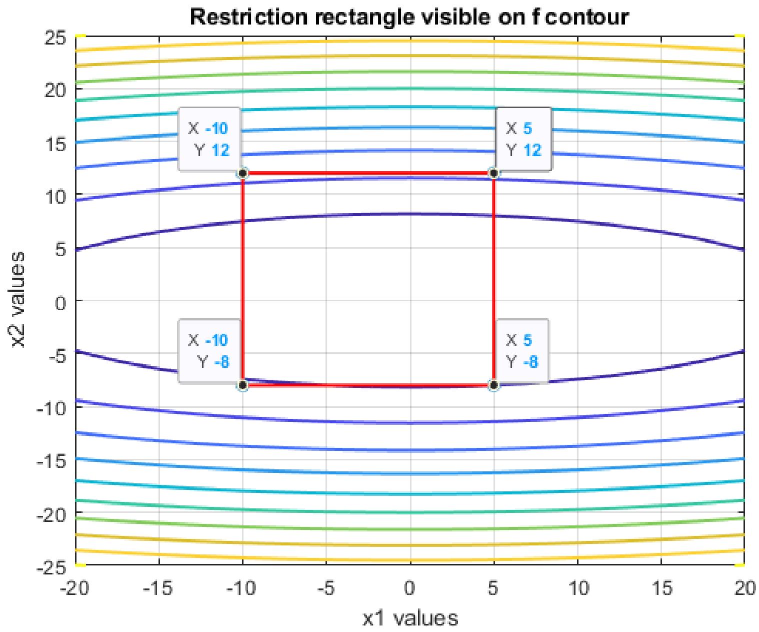
Αντικείμενο της εργασίας είναι η ελαχιστοποίηση μιας δοσμένης κυρτής συνάρτησης $f : \mathbb{R}^2 \to \mathbb{R}$ με περιορισμούς. Πιο συγκεκριμένα θα μελετήσουμε τη Μέθοδο Μέγιστης Καθόδου με προβολή. Η συνάρτηση που θα προσπαθήσουμε να ελαχιστοποιήσουμε με αυτή τη μέθοδο είναι η:

$$f(\boldsymbol{x}) = \frac{1}{3}x_1^2 + 3x_2^2$$
$$subject\ to: -10 \le x_1 \le 5\,, -8 \le x_2 \le 12$$

Με το αρχείο plotting_f.m αναπαριστούμε τρισδιάστατα την f για να οπτικοποιήσουμε τη μορφή της και να έχουμε μια πρώτη ιδέα του ελαχίστου που ψάχνουμε.



Παρατηρούμε ότι η συνάρτησή μας έχει ένα ελλειπτικό παραβολοειδές σχήμα και εμφανίζει ελάχιστο στο (0,0). Με το ίδιο πρόγραμμα σχεδιάζουμε και τις ισοϋψείς καμπύλες της f και θέτουμε εντός ορθογωνίου την περιοχή των περιορισμών για να γνωρίζουμε σε ποια περιοχή αναζητούμε το ελάχιστο.

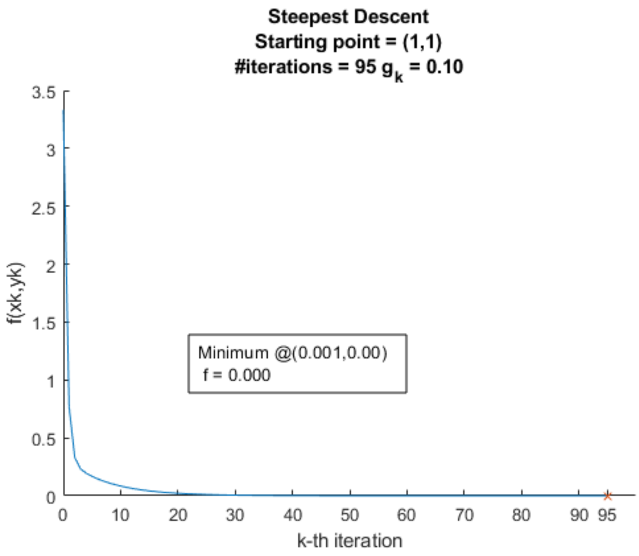


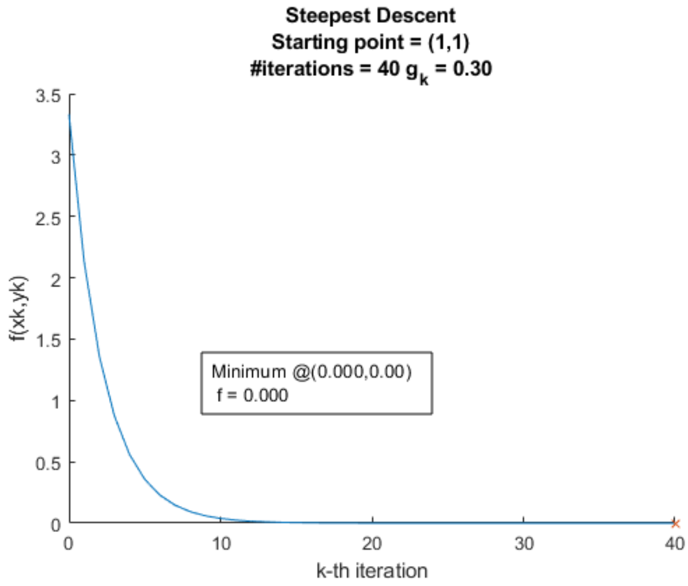
ΘΕΜΑ 1^{ο}:

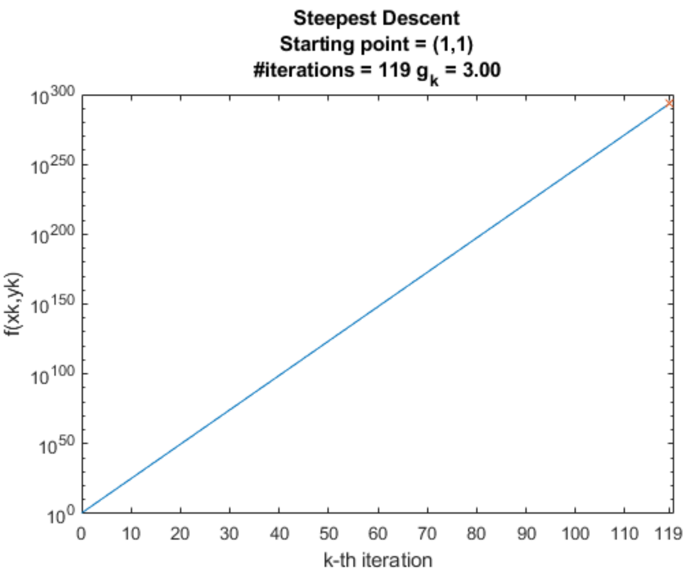
Στο 1^{ο} θέμα μας ζητείται να ελαχιστοποιήσουμε την συνάρτηση χρησιμοποιώντας την απλή Μέθοδο Μέγιστης Καθόδου (βλ. εργασία 2) με ακρίβεια ε=0.001 και 4 διαφορετικά γ_κ :

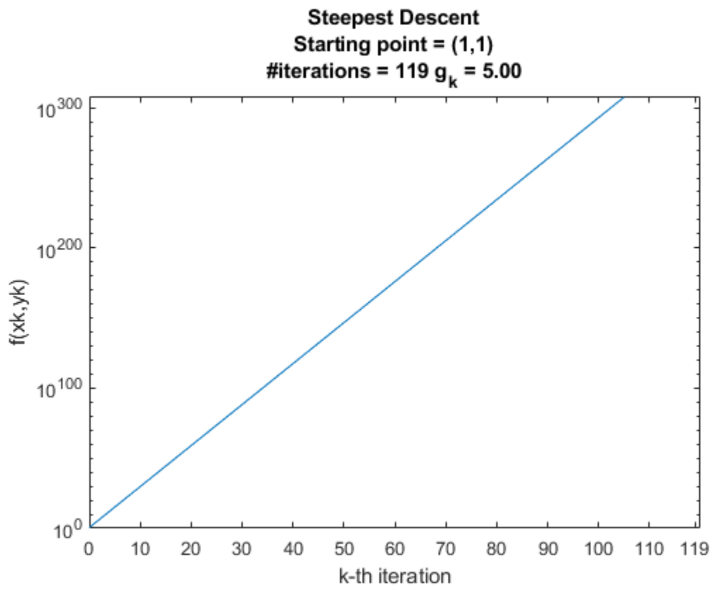
i. $\gamma_\kappa = 0.1$
ii. $\gamma_\kappa = 0.3$
iii. $\gamma_\kappa = 3$
iv. $\gamma_\kappa = 5$

Σαν σημείο εκκίνησης για την μέθοδο επιλέξαμε αυθαίρετα το (1,1). Για να μπορέσει να τερματίσει η μέθοδος τέθηκαν ως όριο οι 120 επαναλήψεις αλλιώς το πρόγραμμά μας θα οδηγούνταν σε ατέρμων βρόγχο. Στην 1^{η} και 2^{η} περίπτωση ο αλγόριθμος λειτουργεί αρκετά καλά και εντοπίζει με ακρίβεια το ζητούμενο ελάχιστο. Απλώς στην περίπτωση (i) επειδή το βήμα γ_κ είναι μικρό παρατηρείται πιο αργή σύγκλιση. Αντιθέτως στην 3^{η} και 4^{η} περίπτωση, επειδή το βήμα γ_κ γίνεται υπερβολικά μεγάλο, ο αλγόριθμος αποκλίνει και οδηγείται σε σημεία όπου η τιμή της αντικειμενικής συνάρτησης αυξάνεται με ρυθμό εκθετικό. Για αυτόν τον λόγο χρειάστηκε να βάλουμε λογαριθμική κλίμακα στον άξονα y=f(x1,x2) για τις περιπτώσεις (iii) και (iv) ώστε να μπορέσει να δημιουργηθεί το γράφημα και να είναι κάπως αναγνώσιμο.

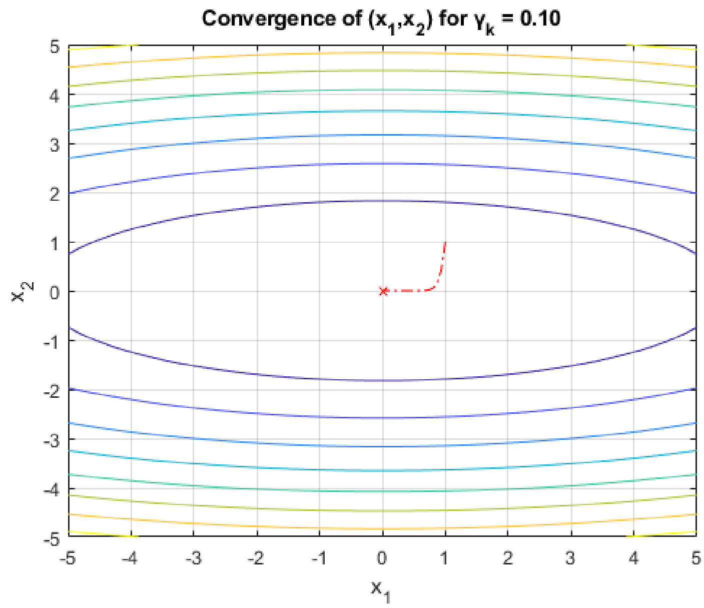

Steepest Descent
Starting point = (1,1)
#iterations = 95 g_k = 0.10
f(xk,yk)
Minimum @(0.001,0.00)
f = 0.000
k-th iteration

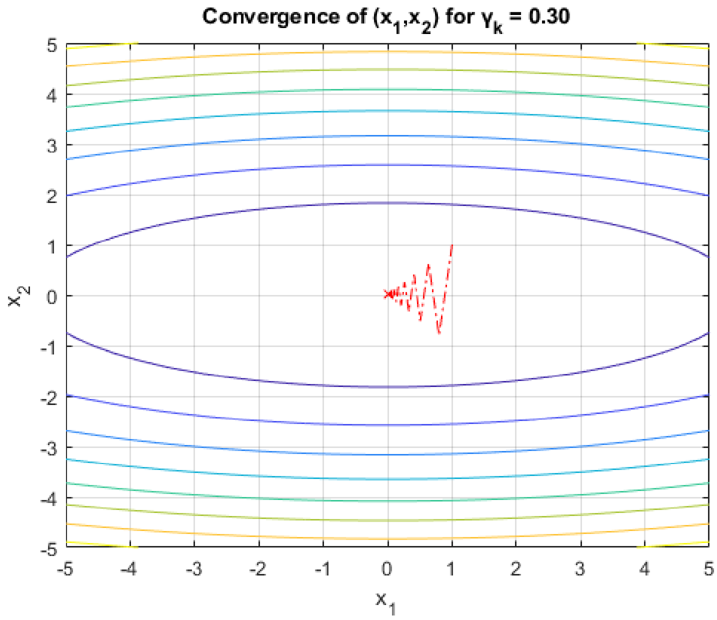

Steepest Descent
Starting point = (1,1)
#iterations = 40 g_k = 0.30
f(xk,yk)
Minimum @(0.000,0.00)
f = 0.000
k-th iteration


Steepest Descent
Starting point = (1,1)
#iterations = 119 g_k = 3.00
f(xk,yk)
k-th iteration



Στη συνέχεια θα παρουσιαστούν τα διαγράμματα σύγκλιση για τις περιπτώσεις (i) και (ii). Όπως γίνεται αντιληπτό για γ_κ=0.1 παρατηρείται ομαλή σύγκλιση προς το (0,0) ενώ για γ_κ=0.3 παρατηρείται μια μικρή ταλάντωση που αποσβένει μέχρι και την τελική σύγκλιση στο σημείο ελαχιστοποίησης. Η ταλάντωση αυτή παρατηρείται στο διάστημα $-1 \leq x_2 \leq 1$





Η μαθηματική εξήγηση πίσω από αυτές τις παρατηρήσεις είναι η εξής:

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

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

Οπότε για τη μέθοδο θα είναι

$$\begin{bmatrix} x_{1,k+1} \\ x_{2,k+1} \end{bmatrix} = \begin{bmatrix} x_{1,k} \\ x_{2,k} \end{bmatrix} - \gamma_\kappa \begin{bmatrix} \frac{2}{3} x_{1k} \\ 6x_{2k} \end{bmatrix} \Leftrightarrow$$

$$x_{1,k+1} = x_{1,k} - \gamma_\kappa \frac{2}{3} x_{1k}$$

$$x_{2,k+1} = x_{2,k} - \gamma_\kappa 6x_{2k}$$

Άρα τελικά $x_{1,k+1} = \left(1 - \gamma_\kappa \frac{2}{3}\right) x_{1k} \quad x_{2,k+1} = (1 - 6\gamma_\kappa) x_{2k}$

Για να επιτύχουμε σύγκλιση στο σημείο ελαχιστοποίησης (0,0) πρέπει να συγκλίνουν οι 2 παραπάνω σχέσεις δηλαδή:

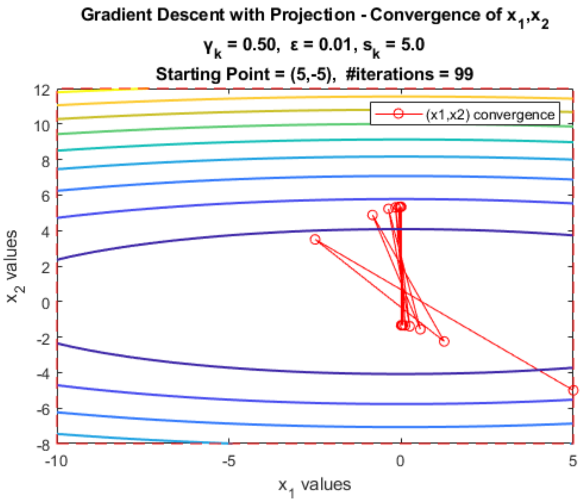
$$\left|1 - \frac{2}{3}\gamma_\kappa\right| < 1$$

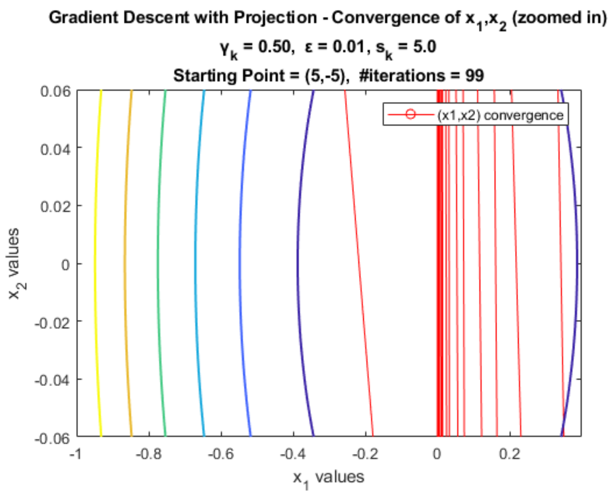
$$|1 - 6\gamma_\kappa| < 1$$

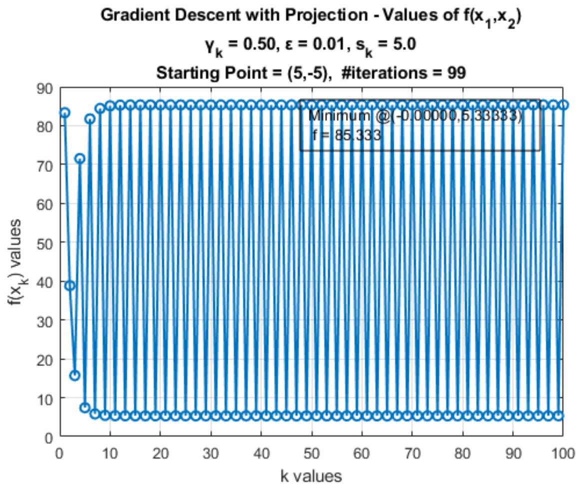
Λύνοντας τις 2 ανισότητες και βρίσκοντας που συναληθεύουν προκύπτει $0<\gamma_\kappa<1/3$. Η τιμή αυτή επαληθεύει και τα παραπάνω αποτελέσματά μας και δικαιολογεί και την ταλάντωση που προκύπτει για γ_κ=0.3 καθώς η τιμή αυτή είναι πολύ κοντά στο άνω φράγμα που έχουμε θέσει για το βήμα μας.

<u>ΘΕΜΑ 2°:</u>

Στο θέμα 2 χρησιμοποιούμε τη Μέγιστη Κάθοδο με Προβολή για s_κ=5 και γ_κ=0.5, ενώ ως αρχικό σημείο εκκίνησης έχουμε το (5,-5) και η επιζητούμενη παράμετρος ακρίβειας ε=0.01. Όπως γνωρίζουμε για τη συγκεκριμένη μέθοδο το τελικό συνολικό βήμα που πραγματοποιείται κατά την επαναληπτική διαδικασία είναι: $\gamma_\kappa' = \gamma_\kappa s_k$. Και αυτό το βήμα πρέπει να συμμορφώνεται με τους περιορισμούς που αναπτύχθηκαν στο θέμα 1 δηλαδή θα πρέπει το καινούργιο βήμα να είναι θετικό και μικρότερο του 1/3. Σε αυτή την περίπτωση βλέπουμε ότι $\gamma_\kappa' = 0.5 * 5 = 2.5 > 1/3$ άρα περιμένουμε ότι δεν θα υπάρχει σύγκλιση και ότι θα έχουμε μια περίπτωση αντίστοιχη με αυτές των (1,iii) και (1,iv). Παρακάτω παρουσιάζονται τα διαγράμματα που προκύπτουν για το θέμα 2 από το αρχείο tasks2_3_4.m



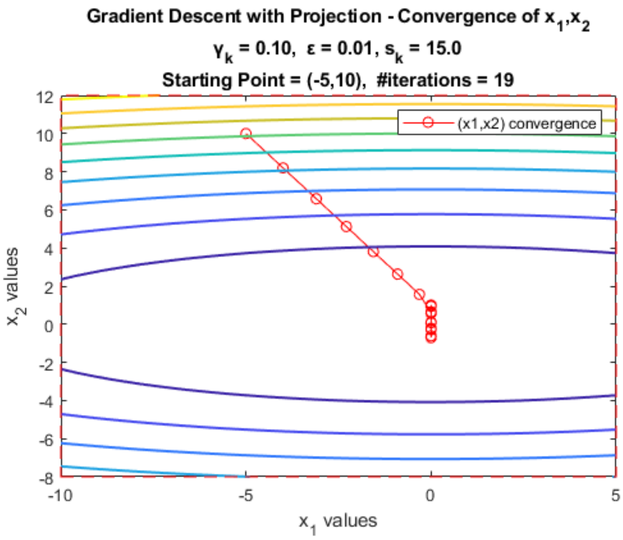


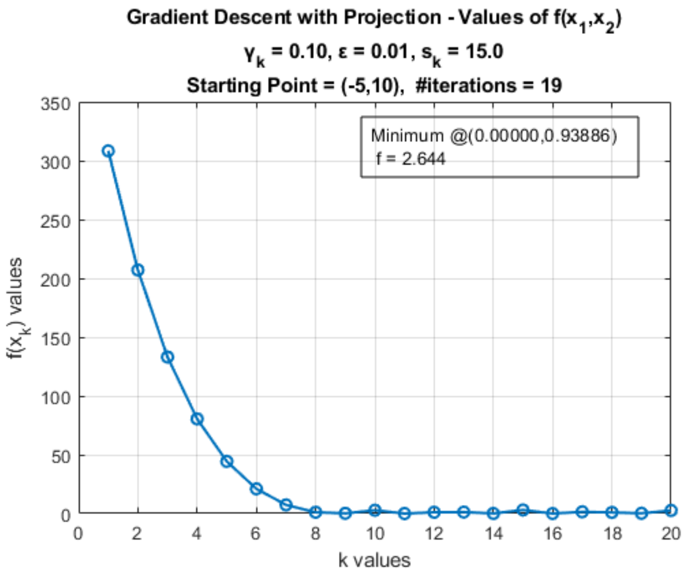


Όπως φαίνεται, έχουμε βάλει άνω όριο τις 100 επαναλήψεις και στα διαγράμματα είναι εμφανής η αποτυχία σύγκλισης του αλγορίθμου λόγω της μεγάλης τιμής του επαναληπτικού βήματος. Παρόλα αυτά η διαφοροποίηση με το Θέμα 1 είναι πως εδώ όσες επαναλήψεις κι αν επιχειρήσουμε να κάνουμε, οι τιμές της αντικειμενικής συνάρτησης πάντα θα υπολογίζονται μέσα στο ορθογώνιο των περιορισμών και δεν θα αυξάνονται εκθετικά. Αυτό είναι αποτέλεσμα της προβολής που κάνουμε σε αυτή τη μέθοδο καθώς , όπως γνωρίζουμε από την θεωρία η προβολή διατηρεί τις τιμές των μεταβλητών εντός των περιορισμών.

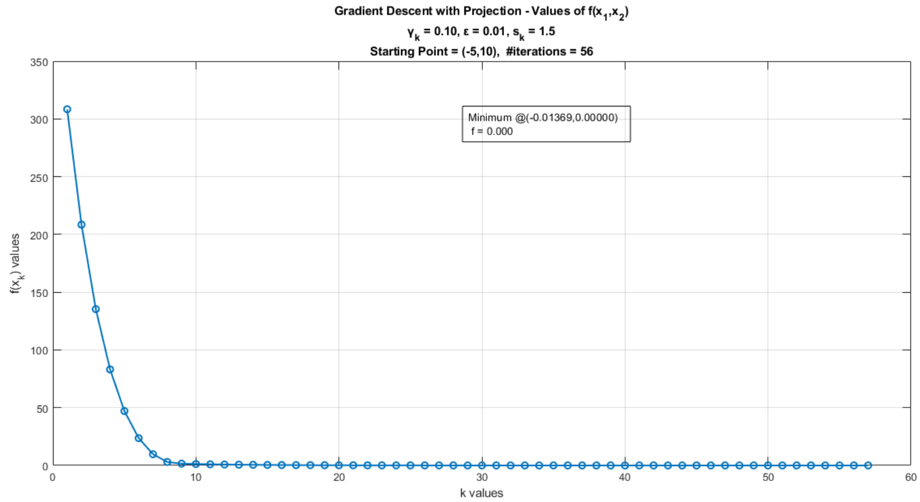
ΘΕΜΑ 3^{o}:

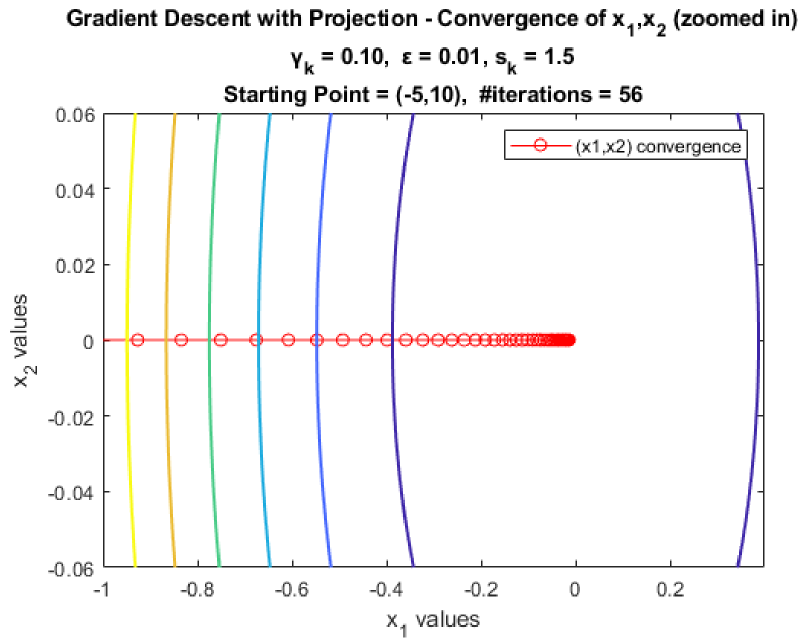
Στο θέμα 3 χρησιμοποιούμε τη Μέγιστη Κάθοδο με Προβολή για s_k=15 και $γ_κ$=0.1, ενώ ως αρχικό σημείο εκκίνησης έχουμε το (-5,10) και η επιζητούμενη παράμετρος ακρίβειας ε=0.01. Όπως γνωρίζουμε για τη συγκεκριμένη μέθοδο το τελικό συνολικό βήμα που πραγματοποιείται κατά την επαναληπτική διαδικασία είναι: $\gamma_\kappa' = \gamma_\kappa s_k$. Και αυτό το βήμα πρέπει να συμμορφώνεται με τους περιορισμούς που αναπτύχθηκαν στο θέμα 1 δηλαδή θα πρέπει το καινούργιο βήμα να είναι θετικό και μικρότερο του 1/3. Σε αυτή την περίπτωση βλέπουμε ότι $\gamma_\kappa' = 0.1 * 15 = 1.5 > 1/3$ άρα περιμένουμε ότι δεν θα υπάρχει σύγκλιση και ότι θα έχουμε μια περίπτωση που μοιάζει με αυτές των (1,iii) και (1,iv). Παρακάτω παρουσιάζονται τα διαγράμματα που προκύπτουν για το θέμα 3 από το αρχείο tasks2_3_4.m





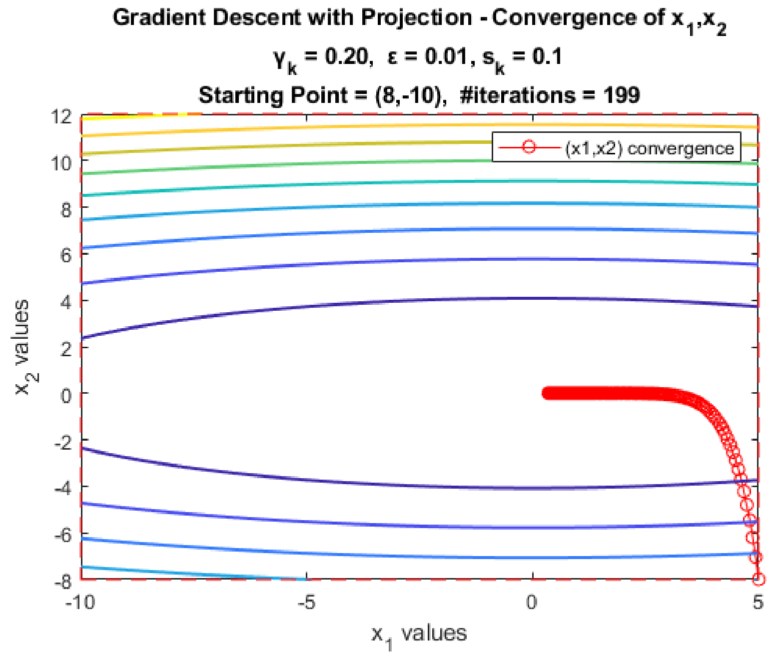
Παρατηρούμε πως η μέθοδος μοιάζει να συγκλίνει προς το ελάχιστο και αυτό είναι κάτι το οποίο δεν περιμέναμε δεδομένων των συνθηκών. Ωστόσο όσο κι αν πλησιάσει (αυξάνοντας δηλαδή το k από 20 σε 100) παρατηρούμε ότι ο αλγόριθμος θα τρέχει ατέρμονα και δεν θα πιάσει ποτέ την επιθυμητή ακρίβεια. Σε σύγκριση με τα θέματα 1 και 2 και τις αντίστοιχες περιπτώσεις αυτών, παρατηρούμε πως και εδώ οι τιμές της αντικειμενικής συνάρτησης υπολογίζονται εντός των περιορισμών λόγω της μεθόδου της προβολής και πως (ίσως επειδή εδώ γ_κ'=1.5 άρα πιο μικρό από ότι στο θέμα 2) υπάρχει μια κάπως ανεκτή σύγκλιση η οποία ωστόσο δεν πρέπει να μας ξεγελάει. Ο αλγόριθμος παραμένει και εδώ αναποτελεσματικός. Ένας πρακτικός τρόπος να οδηγήσουμε την μέθοδο στο να συγκλίνει αποτελεσματικά είναι το να αλλάξουμε το s_k=1.5 και έτσι να έχουμε γ_κ'=0.15 βήμα εντός των περιορισμών. Αν κάνουμε την σχετική αλλαγή στο matlab script θα παρατηρήσουμε ότι ο αλγόριθμος συγκλίνει σε 56 επαναλήψεις (δεν χρειάζεται να πιάσει το άνω όριο επαναλήψεων που έχουμε θέσει).

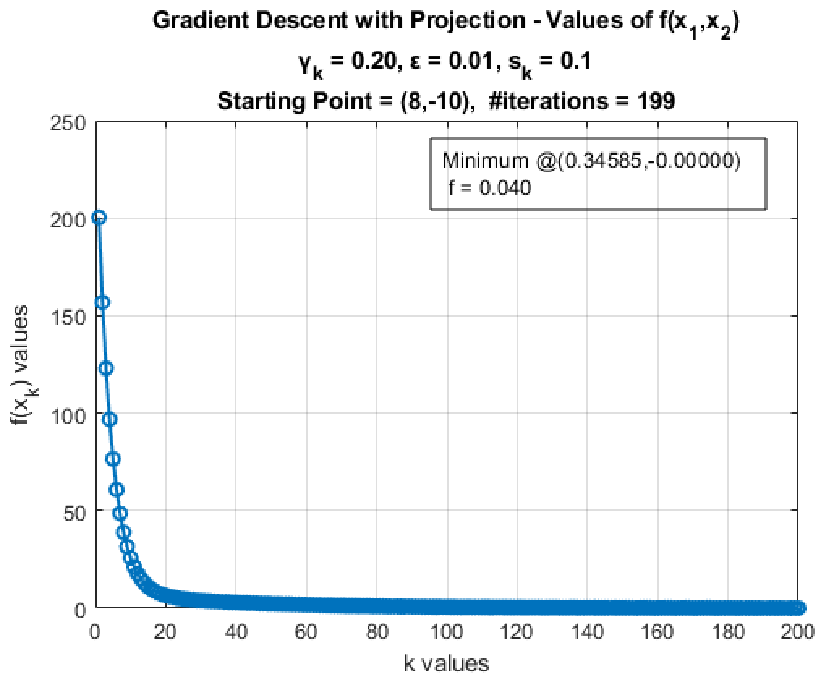




ΘΕΜΑ 4°:

Στο θέμα 4 χρησιμοποιούμε τη Μέγιστη Κάθοδο με Προβολή για $s_k=0.1$ και $\gamma_k=0.2$, ενώ ως αρχικό σημείο εκκίνησης έχουμε το (8,-10) και η επιζητούμενη παράμετρος ακρίβειας ε=0.01. Όπως γνωρίζουμε για τη συγκεκριμένη μέθοδο το τελικό συνολικό βήμα που πραγματοποιείται κατά την επαναληπτική διαδικασία είναι: $\gamma_k' = \gamma_k s_k$. Και αυτό το βήμα πρέπει να συμμορφώνεται με τους περιορισμούς που αναπτύχθηκαν στο θέμα 1 δηλαδή θα πρέπει το καινούργιο βήμα να είναι θετικό και μικρότερο του 1/3. Σε αυτή την περίπτωση βλέπουμε ότι $\gamma_k' = 0.2 * 0.1 = 0.02 < 1/3$. Συνεπώς αναμένουμε ο αλγόριθμος να συγκλίνει όμως με πολύ αργό ρυθμό καθώς η τιμή του βήματος είναι πάρα πολύ μικρή. Πριν δοκιμάσουμε να τρέξουμε τον αλγόριθμο για τις παραπάνω αρχικές συνθήκες παρατηρούμε πως η τιμή $x_1 = 8$ και $x_2 = -10$ δεν ανήκει στο κυρτό σύνολο των περιορισμών που ορίσαμε παραπάνω. Η πρώτη κλήση του αλγορίθμου θα χρησιμοποιήσει την προβολή στο Χ οπότε το αρχικό σημείο θα γίνει το (5,-8) οδηγώντας δηλαδή τον αλγόριθμο μέσα στο διάστημα των περιορισμών έστω και αν αυτό είναι στο όριο του. Από εκεί και έπειτα μπορούμε να δούμε την μέθοδο σαν της απλής μέγιστης καθόδου με γ'ₖ = 0.02 όπως είδαμε παραπάνω και βασιζόμενοι στο θέμα 1(i) να πούμε ότι κ ο αλγόριθμος του θέματος 4 θα συγκλίνει αργά λόγω μικρού βήματος.





Τα παραπάνω διαγράμματα επιβεβαιώνουν όσα περιμέναμε να δούμε σε αυτό το θέμα καθώς η αργή σύγκλιση είναι ορατή (σταματήσαμε χειροκίνητα τον αλγόριθμο στις 200 επαναλήψεις και παρά την σύγκλιση που εμφανίζεται ακόμα δεν είχε εντοπίσει με την επιθυμητή ακρίβεια το ελάχιστο) και φαίνεται πως το πρώτο βήμα εκτελείται από το σημείο (5,-8) καθώς το αρχικό σημείο (8,-10) προβλήθηκε εντός του ορθογωνίου των περιορισμών.