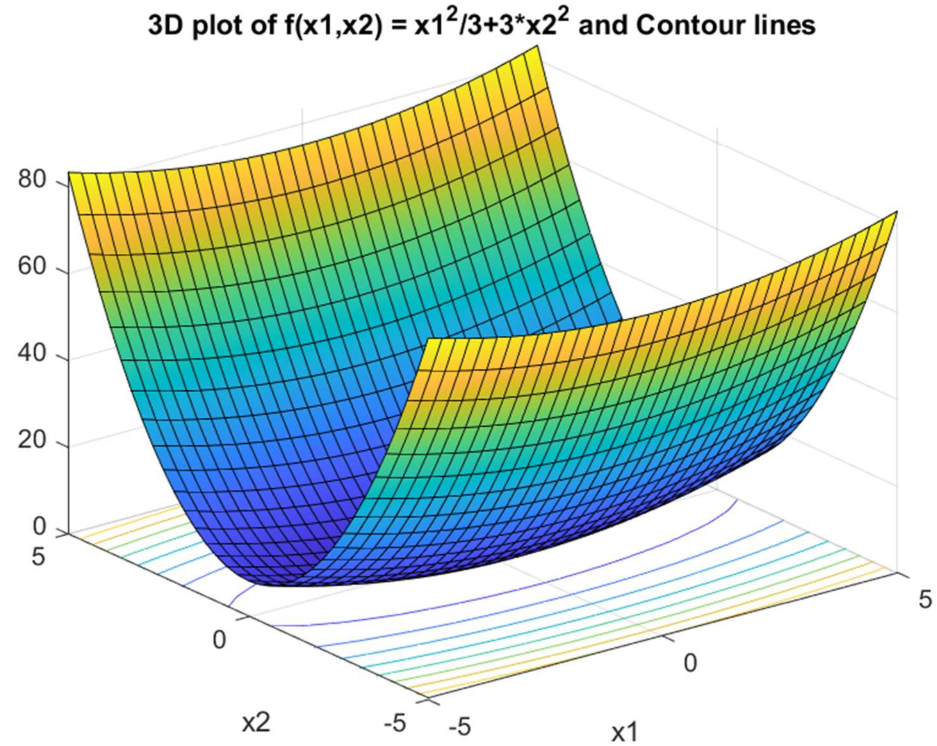
# 3η Εργαστηριακή Άσκηση

Εργασία της Πλευρίδη Βασιλική Βαρβάρα (ΑΕΜ:10454)

Η άσκηση αυτή έχει ως στόχο των μελέτη  και σύγκριση της **μεθόδου Μέγιστης Καθόδου με και χωρίς προβολή** για διαφορετικά $\gamma_k$ και αρχικά σημεία εκκίνησης. Συγκεκριμένα, στο θέμα 1, χρησιμοποιώντας τον κώδικα για την Μέθοδο Μέγιστης Καθόδου που υλοποιήθηκε στην προηγούμενη εργαστηριακή άσκηση και κάνοντας μικρές προσαρμοστικές αλλαγές, μελετάμε την σύγκλιση ή  μη, για διάφορα $\gamma_k$ για τυχαίο σημείο εκκίνησης. Στα θέματα 2, 3, 4, χρησιμοποιώντας διαφορετικές παραμέτρους για τα $\gamma_k$ και $s_k$, και αρχικό σημείο εκκίνησης, παρατηρούμε το πως επηρεάζει αυτή την φορά την Μέθοδο Μέγιστης Καθόδου με προβολή.

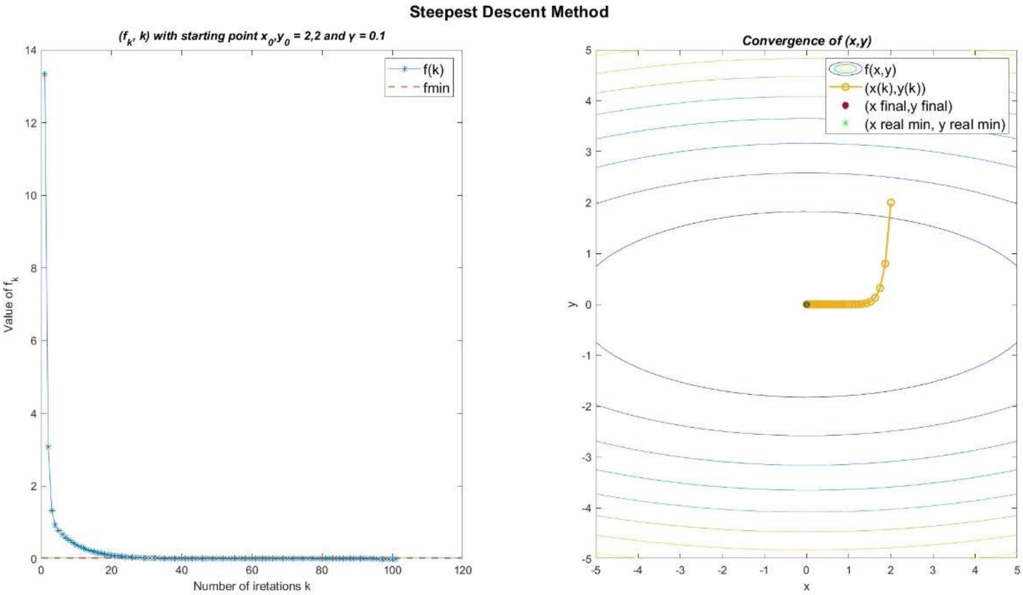
Η αντικειμενική συνάρτηση που χρησιμοποιείται για τον παραπάνω σκοπό και δηλαδή αναζητούμε το ελάχιστό της είναι η $\boldsymbol{f(x_1,x_2) = x_1^2/3 + 3x_2^2}$. Παρακάτω παρουσιάζεται το τρισδιάστατο γράφημα της $f$ για να την μας βοηθήσει στην οπτικοποίηση της μορφής της:

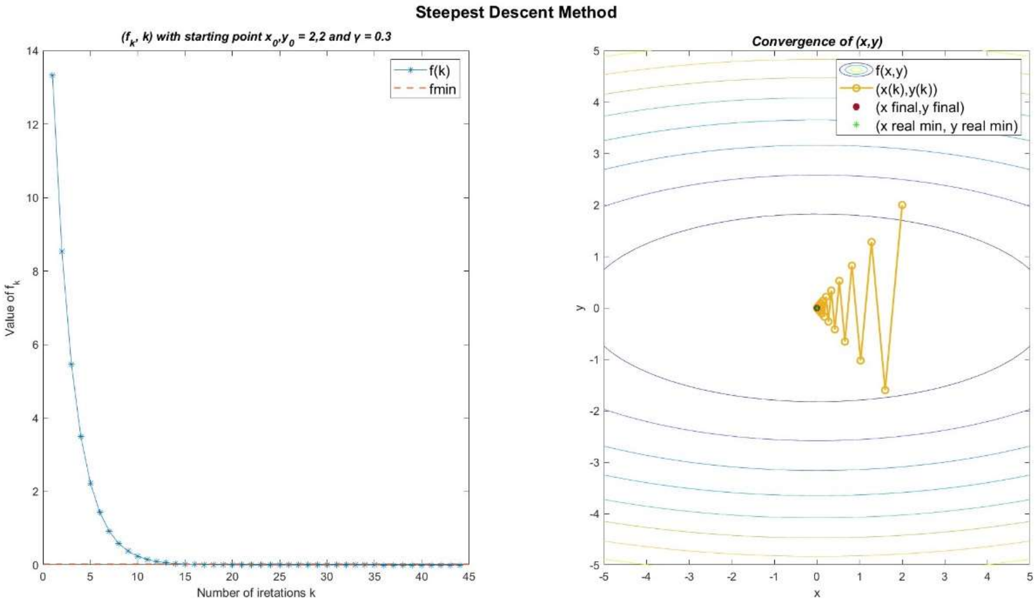


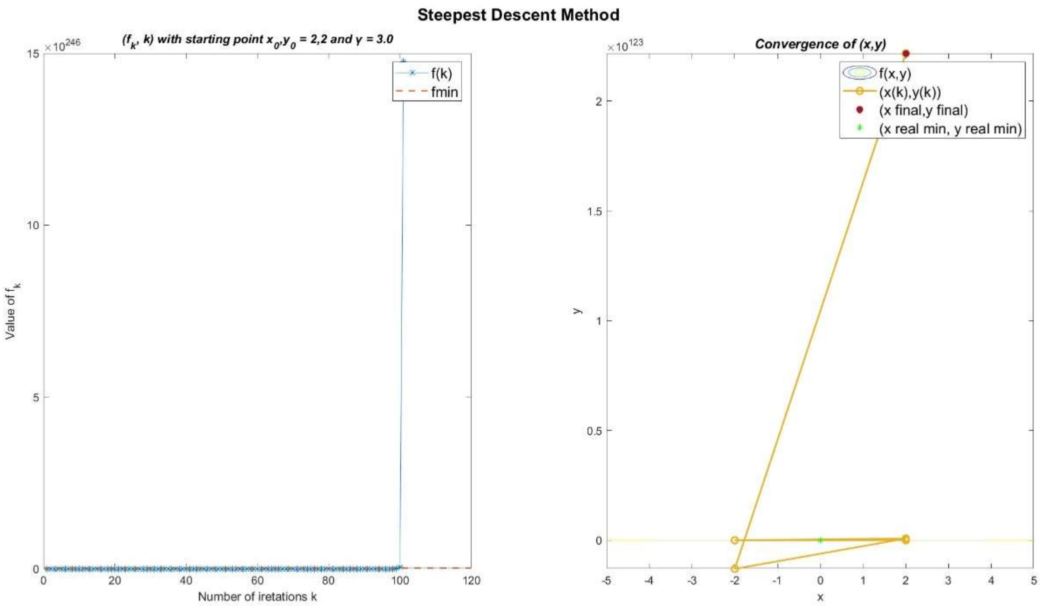

Όπως είναι εμφανές η $f$ παρουσιάζει ελάχιστο στο $(0,0)$και άρα σε αυτό το σημείο αναμένουμε να τερματίσει ο αλγόριθμος μας κάθε φορά.

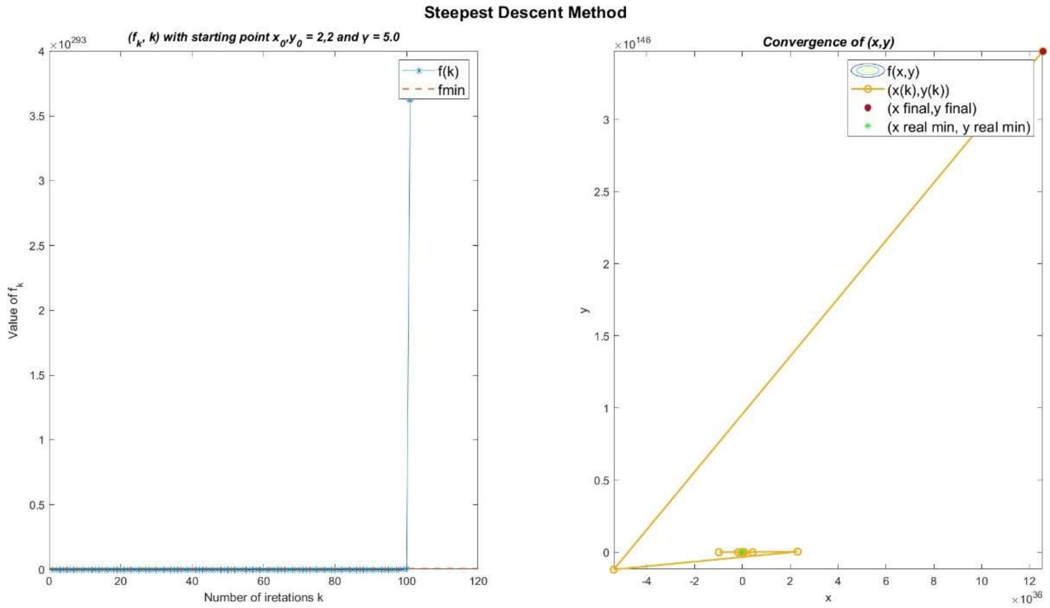
## Θέμα 1

Ως σημείο εκκίνησης λοιπόν για το θέμα αυτό, επιλέχθηκε τυχαία το $(x_1, x_2) = (2,2)$ ενώ από την εκφώνηση θεωρείται σταθερά τερματισμού $\varepsilon = 0.001$ και το $\gamma_k$ να παίρνει τις τιμές 0.1 , 0.3, 3 και 5 αντίστοιχα. Τα ζητούμενα γραφικά αποτελέσματα παρουσιάζονται παρακάτω:


Steepest Descent Method
(f_k, k) with starting point x_0,y_0 = 2,2 and γ = 0.1
f(k)
fmin
Value of f_k
Number of iretations k
Convergence of (x,y)
f(x,y)
(x(k),y(k))
(x final,y final)
(x real min, y real min)
x
y


Steepest Descent Method
(f_k, k) with starting point x_0,y_0 = 2,2 and γ = 0.3
f(k)
fmin
Value of f_k
Number of iretations k
Convergence of (x,y)
f(x,y)
(x(k),y(k))
(x final,y final)
(x real min, y real min)
x
y


Steepest Descent Method
(f_k, k) with starting point x_0,y_0 = 2,2 and γ = 3.0
×10^246
f(k)
fmin
Value of f_k
Number of iretations k
Convergence of (x,y)
×10^123
f(x,y)
(x(k),y(k))
(x final,y final)
(x real min, y real min)
x
y




Όπως φαίνεται στα παραπάνω γραφήματα, για τις περιπτώσεις που $\boldsymbol{\gamma_k = [0.1, 0.3]}$, **η μέθοδος συγκλίνει κανονικά στο ελάχιστο ενώ για** $\boldsymbol{\gamma_k = [3, 5]}$ **όχι**. Τα δύο τελευταία γραφήματα ουσιαστικά δεν οδηγούν σε κανένα άλλο συμπέρασμα παρά μονό ότι η επιλογή βήματος δεν είναι σωστή και ότι η τιμή της $f$ αντί να μειώνεται, αυξάνεται εκθετικά. Ενδιαφέρον παρουσιάζει η ταλάντωση γύρω από το 0 στην δεύτερη περίπτωση όπου $\gamma_k = 0.3$. Όλα αυτά, έρχονται ως λογικό επακόλουθο της επεξήγησης των αποτελεσμάτων με μαθηματική ακρίβεια για το όρια του $\gamma_k$ που οδηγούν στην σωστή λειτουργία του αλγορίθμου που παρουσιάζεται παρακάτω ( και σε κώδικα στο MATLAB):

**<u>Μαθηματική Απόδειξη</u>**

Το gradient της $f$ δίνεται από τον τύπο:

$$\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1/3 \\ 6x_2 \end{bmatrix} \ (1)$$

Τα σημεία που χρησιμοποιεί ο αλγόριθμος είναι τα $x_{k+1} = x_k - \gamma_k \nabla f x_k$ και άρα με χρήση της $(1)$ θα είναι της μορφής:

$$\begin{bmatrix} x_{1k+1} \\ x_{2k+1} \end{bmatrix} = \begin{bmatrix} x_{1k} - \gamma \ 2x_1/3 \\ x_{2k} - \gamma \ 6x_2 \end{bmatrix} \ (2)$$

Το ελάχιστο της συνάρτησης βρίσκεται στο σημείο $(x_1, x_2) = (0,0)$ και άρα θέλουμε για $k \to \infty, x_1 = 0, x_2 = 0$.

Για να ισχύει αυτό, αρκεί να συγκλίνουν οι σχέσεις $(2)$ και άρα:

$$\left|1 - \gamma \frac{2}{3}\right| < 1 \text{ και } |1 - 6\gamma| < 1$$

Έπειτα από πράξεις οδηγούμαστε στα τελικά όρια $\boldsymbol{0 < \gamma_k < \frac{1}{3}}$ $(3)$

Η παραπάνω μαθηματική απόδειξη έρχεται σε απόλυτη συμφωνία με τα γραφικά αποτελέσματα καθώς για τιμές μικρότερες του 1/3 ο αλγόριθμος συγκλίνει κανονικά, ενώ για αρκετά μεγαλύτερες τιμές όπως είναι τα 3 και 5 αποκλίνει. Για την οριακή τιμή $\gamma_k = 0.3$, η οποία είναι πολύ κοντά στο άνω όριο, καταλαβαίνουμε πλέον γιατί παρατηρείται αυτή η αποσβεννύμενη ταλάντωση.

# <u>Θέμα 2</u>

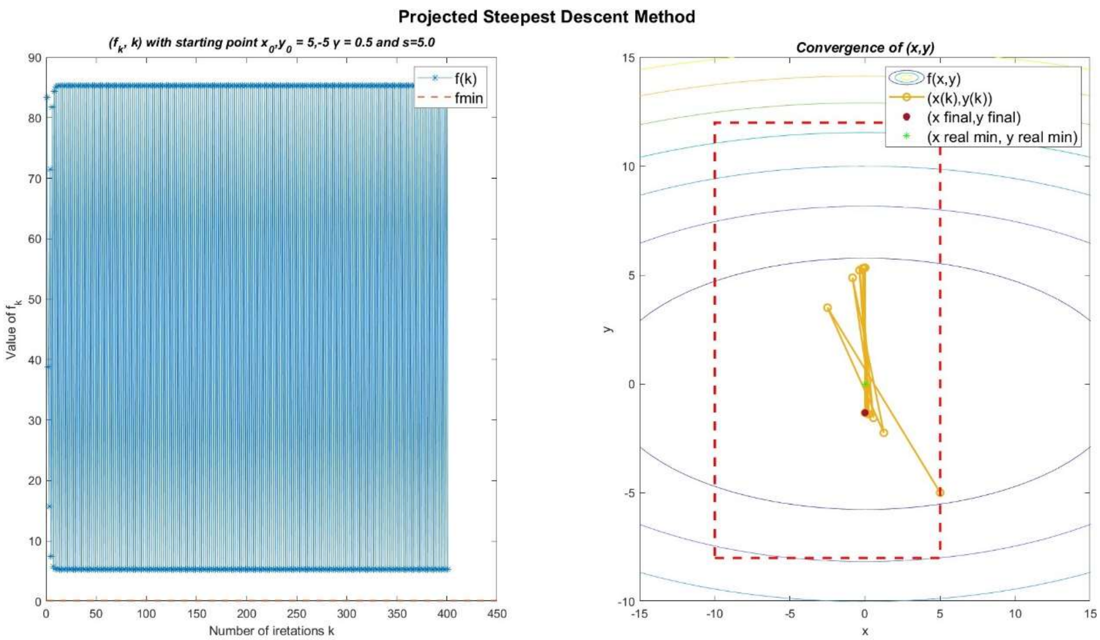
Σε αυτό το θέμα, όπως προαναφέρθηκε, υλοποιείται και χρησιμοποιείται η **Μέθοδος Μέγιστης Καθόδου με προβολή**. Σε αυτή την περίπτωση τα σημεία του αλγορίθμου είναι της μορφής:

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

όπου $\widehat{x_k} = \Pr_X\{x_k - s_k \nabla f(x_k))\}$ με $X = x \in \mathbb{R}^2 : -10 \leq \mathrm{x}_1 \leq 5 \kappa\alpha\iota - 8 \leq \mathrm{x}_2 \leq 12$

Με τα εξής αρχικοποιημένα στοιχεία:

- $\varepsilon = 0.01$
- $\gamma_k = 0.5$
- $s_k = 5$
- $(x_{1\,0}, x_{2\,0}) = (5, -5)$




Η διακεκομμένη κόκκινη γραμμή στο δεξιά γράφημα αποτελεί τα φράγματα για τις τιμές των $x_1, x_2$.

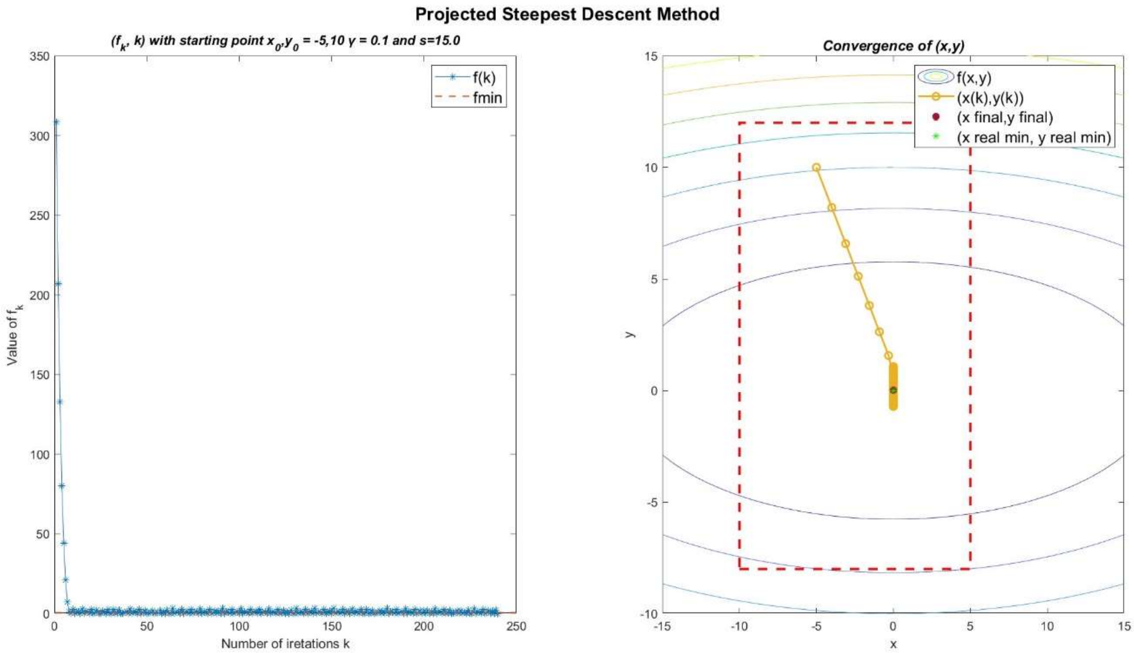
Όπως φαίνεται από τα παραπάνω γραφήματα, **ο αλγόριθμος δεν συγκλίνει ποτέ** και μάλιστα θα συνέχιζε να **ταλαντώνεται επ άπειρον** αν δεν θέταμε έναν μέγιστο αριθμό επαναλήψεων k για τον τερματισμό της μεθόδου. Το αποτέλεσμα αυτό είναι αναμενόμενο καθώς το νέο μας βήμα $\boldsymbol{\gamma_k' = \gamma_k s_k = 0.5 * 5 = 2.5}$ **είναι αρκετά μεγαλύτερο του άνω ορίου της σχέσης** $\boldsymbol{(3)}$**.** Συγκριτικά με το Θέμα 1 για τις περιπτώσεις που δεν έχουμε σύγκλιση, η διαφορά είναι ότι τώρα οι τιμές της $f$ δεν αυξάνονται εκθετικά αλλά λόγω των φραγμάτων από το επίπεδο X, εγκλωβίζονται και ταλαντώνονται στο εύρος των τιμών αυτών.

# Θέμα 3

Στο θέμα αυτό το μόνο που αλλάζει συγκριτικά με το θέμα 2 είναι οι αρχικές συνθήκες οι οποίες αυτή την φορά είναι οι εξής:

- $\varepsilon = 0.01$
- $\gamma_k = 0.1$
- $s_k = 15$
- $(x_{1\,0}, x_{2\,0}) = (-5{,}10)$

Το γραφικό αποτέλεσμα παρουσιάζεται παρακάτω:




Η περίπτωση αυτή παρουσιάζει ιδιαίτερα μεγάλο ενδιαφέρον καθώς φαινομενικά, από το δεξί διάγραμμα, μπορεί εύκολα κάποιος να πει ότι η μέθοδος συγκλίνει κανονικά στο ελάχιστο. Με μία δεύτερη πιο προσεχτική ματιά όμως στο αριστερό διάγραμμα, βλέπουμε ότι πάλι η τιμή της $f$ ταλαντώνεται απλά με μικρότερη peak τιμή ( Παρακάτω παρουσιάζεται και μια μεγέθυνση αυτού για καλύτερη κατανόηση). Για το βήμα πάλι ισχύει:

$$\boldsymbol{\gamma_k' = \gamma_k s_k = 0.1 * 15 = 1.5 > \frac{1}{3}}$$

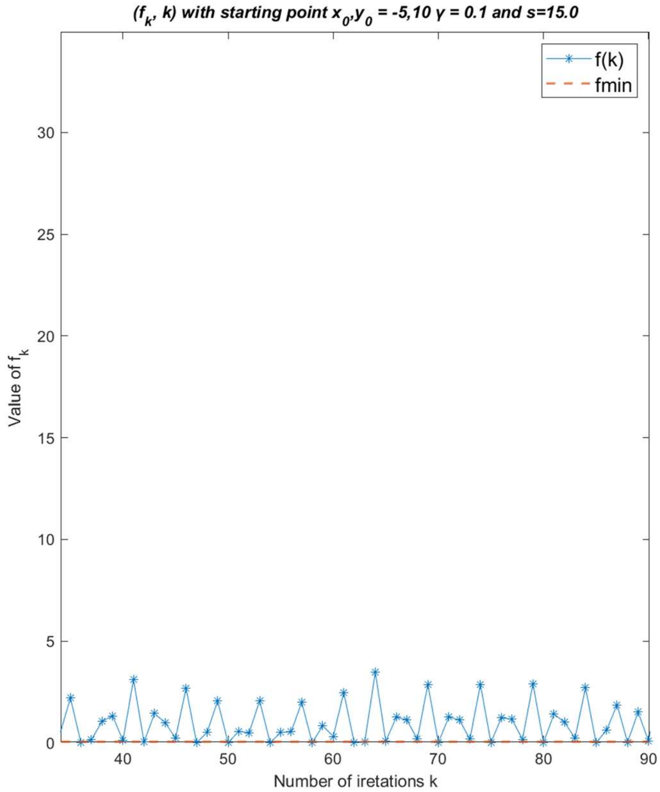
Στην περίπτωση αυτή **το τελικό βήμα $\gamma_k'$ δεν απέχει τόσο πολύ από το άνω όριο και για αυτό ίσως παρατηρείται μικρότερη ταλάντωση** και γενικότερα καλύτερα αποτελέσματα.

Ένας εύκολος και άμεσος τρόπος έτσι ώστε να οδηγηθούμε σε σύγκλιση στο ελάχιστο είναι να μεταβάλουμε την τιμή του $s_k$ έτσι ώστε να το καινούριο $\gamma_k$ να πληρεί την προυπόθεση της σχέσης (3). Παραδείγματος χάριν, αλλάζοντας το $s_k$ σε 2 θα είχαμε:

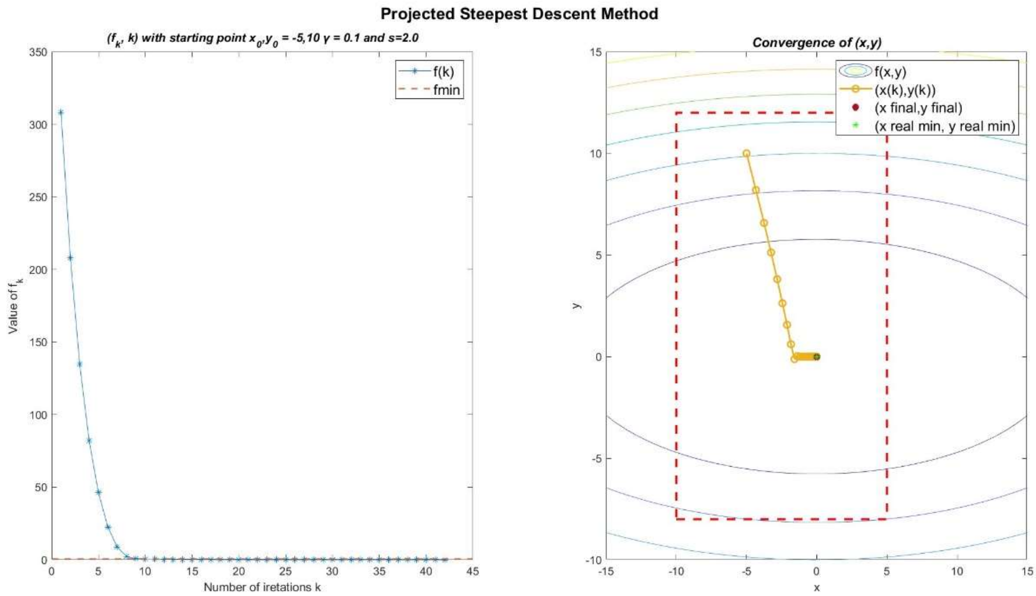
$$\boldsymbol{\gamma_k' = \gamma_k s_k = 0.1 * 2 = 0.2 < \frac{1}{3}}$$

Παρακάτω φαίνεται και το αντίστοιχο γράφημα για να επαληθευτεί ο συλλογισμός.

Μεγέθυνση της περίπτωσης $\gamma_k = 0.1$ και $s_k = 15$:




Για $\gamma_k = 0.1\ και\ s_k = 2$:

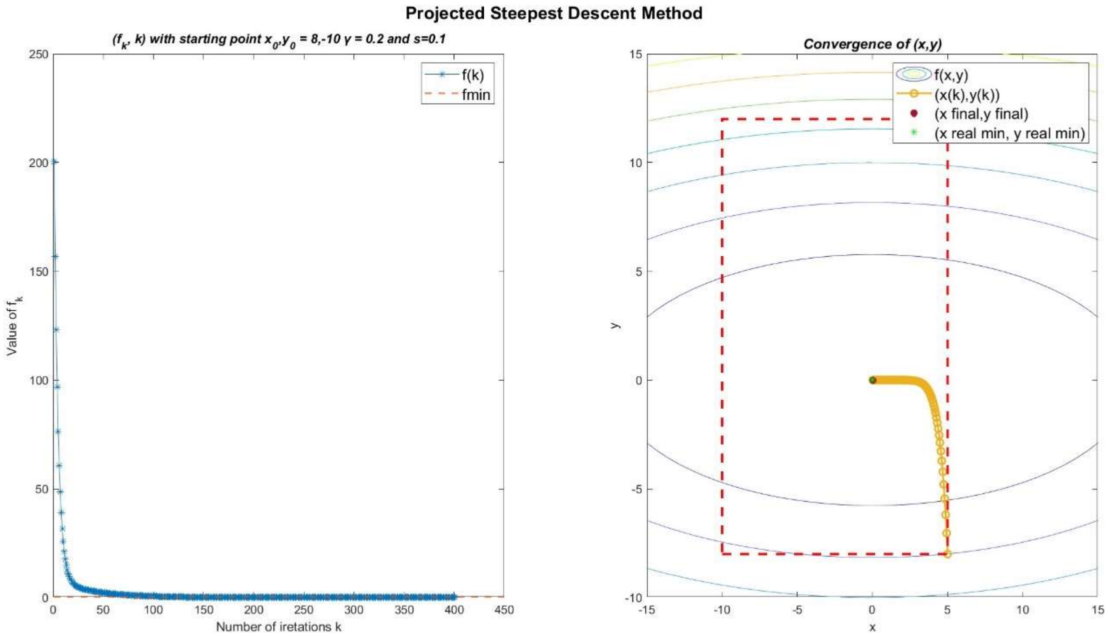


## Θέμα 4

Στο θέμα αυτό το μόνο που αλλάζει συγκριτικά με τα θέματα 2 και 3 είναι οι αρχικές συνθήκες οι οποίες αυτή την φορά είναι οι εξής:

- $\varepsilon = 0.01$
- $\gamma_k = 0.2$
- $s_k = 0.1$
- $(x_{1\,0}, x_{2\,0}) = (8, -10)$

Το γραφικό αποτέλεσμα παρουσιάζεται παρακάτω:




Η χρήση της μεθόδου Μέγιστης Καθόδου με προβολή με τις παραπάνω αρχικοποιημένες συνθήκες, όπως φαίνεται και στο γράφημα, **συγκλίνει στο ελάχιστο** γεγονός που επαληθεύεται και με μαθηματική ακρίβεια από την σχέση: $\boldsymbol{\gamma'_k = \gamma_k s_k = 0.2 * 0.1 = 0.02 < \frac{1}{3}}$. Η **χρήση τόσο μικρού βήματος** όμως, έχει οδηγήσει σε **πολύ αργή εκτέλεση του κώδικα**, ο οποιός δεν θα τερματιζόταν εύκολα αν δεν είχε προστεθεί μέγιστος αριθμός επαναλήψεων στην υλοποίηση στο MATLAB.

Να σημειωθεί ότι το σημείο εκκίνησης έχει συντεταγμένη που δεν ανήκει στο σύνολο X (8 > 5). Η μέθοδος λειτουργεί έτσι ώστε κάθε φορά που έχουμε σημείο που δεν είναι εφικτό, να παίρνουμε την προβολή αυτού στο επίπεδο X και να συνεχίζουμε με τη μέθοδο Μεγίστης Καθόδου. Αυτό γίνεται και άμα το σημείο εκκίνησης δεν είναι εφικτό. **Ο αλγόριθμος της προβολής, με την πρώτη επανάληψη, μας παραπέμπει μέσω της προβολής σε ένα σημείο που είναι εφικτό.**

## Σχολιασμός και συμπεράσματα

Γενικότερα, για την σωστή χρήση της μεθόδου Μέγιστης Καθόδου με ή χωρις προβολή, **σημαντικό είναι να μελετατούνται τα όρια του βήματος $\gamma_k$ για τα οποία ο αλγόριθμος οδηγεί στην σύγκλιση**. Η επιλογή του βήματος λοιπόν κρίνεται ιδιαίτερα σημαντική και προτείνεται **να μην κίνεται στα όρια της ανίσωσης αλλά να παίρνει ενδιάμεσες τιμές**. Ό́σον αφορά τη μέθοδο Μέγιστης Καθόδου χωρίς προβολή, καθώς αποτελεί αλγόριθμο που εφαρμόζεται απουσία περιορισμών, **έχει ως αποτέλεσμα να εκτοξεύει τις τιμές της f προς το άπειρο όταν το $\gamma_k$ είναι τέτοιο ώστε να μη συγκλίνει η ακολουθία των $\{x_k\}$.**

## Βιβλιογραφία

Για την υλοποίηση των αλγορίθμων όπως και για κάποια σχόλια για τον αριθμό των αντικειμενικών συναρτήσεων χρησιμοποιήθηκε το βιβλίο «Τεχνικές Βελτιστοποίησης» του κ.Ροβιθάκη, εκδόσεις ΤΖΙΟΛΑ.

.