ΝΙΟΒΗ ΠΕΡΑΤΙΚΟΥ
ΑΕΜ: 10494

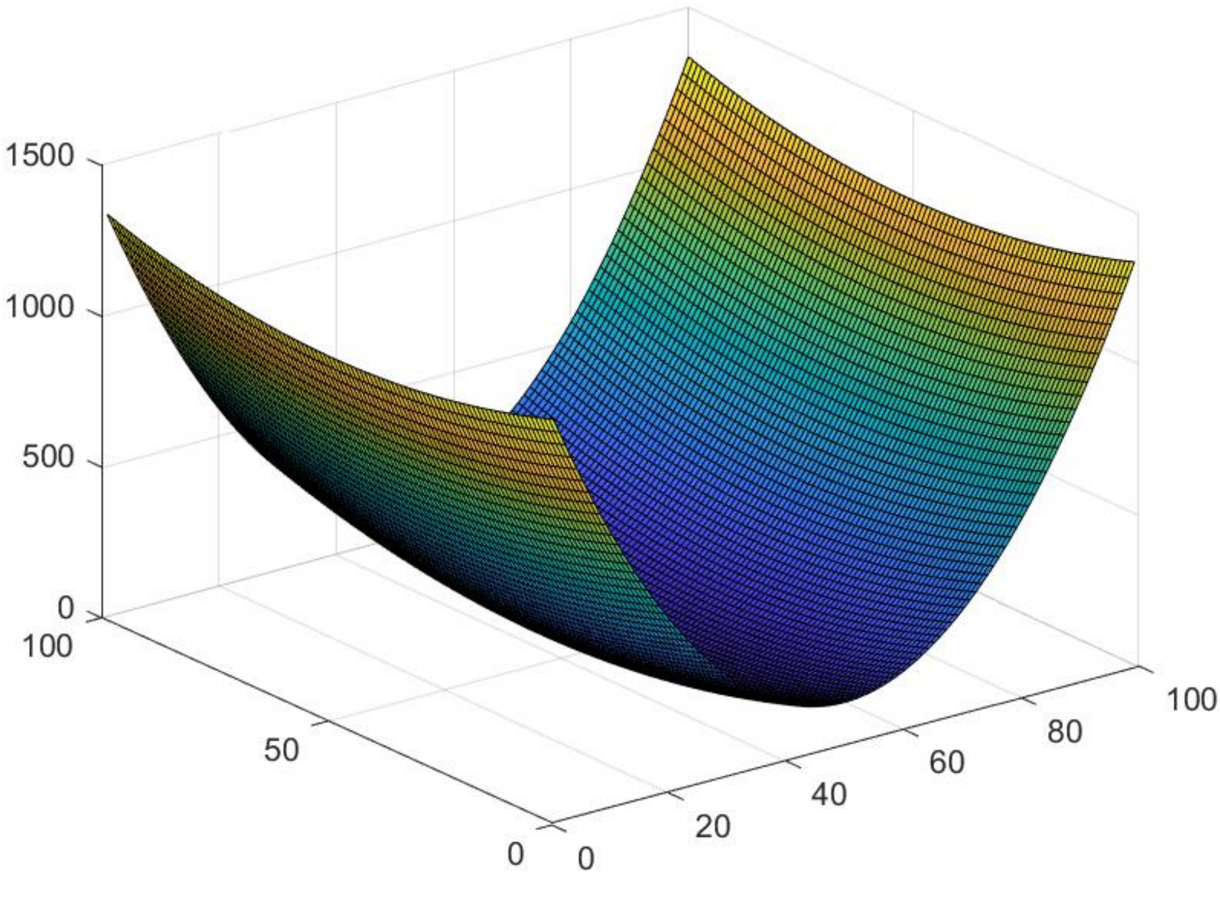
Τεχνικές Βελτιστοποίησης

Άσκηση 3 : Εύρεση ελαχίστου σε συνάρτηση δύο μεταβλητών με ύπαρξη περιορισμών

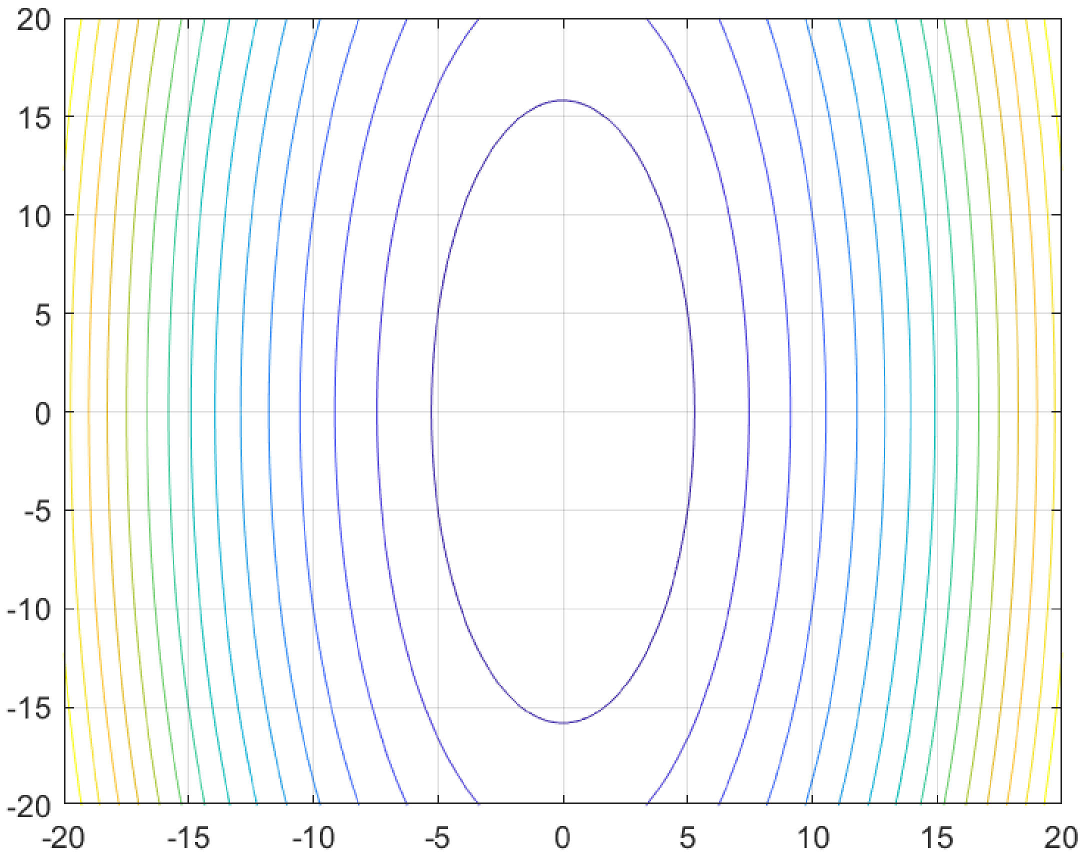
Αντικείμενο αυτής της άσκησης, είναι η εφαρμογή της Μεθόδου της Μέγιστης Καθόδου σε δύο περιπτώσεις και η σύγκριση των αποτελεσμάτων. Αρχικά, υλοποιείται για συνάρτηση *f* για την οποία δεν ισχύουν κάποιοι περιορισμοί και στην συνέχεια χρησιμοποιείται για την ελαχιστοποίηση της ίδιας συνάρτησης υπό την παρουσία, όμως, περιορισμών.

Η *f* έχει τετραγωνική μορφή και δίνεται από τον τύπο

$$f(x_1, x_2) = \frac{x_1^2}{3} + 3*x_2^2 \quad με\ [x_1\ x_2]^T$$



Όπως είχαμε δει, σε προηγούμενη εργασία η μέθοδος της μέγιστης καθόδου εκμεταλλεύεται την καθετότητα του διανύσματος κλίσης, $\nabla f(x_1, x_2)$, στην ισοβαρή καμπύλη της f και για να πετύχει την μεγαλύτερη δυνατή μείωση, εκτελεί κάθετα βήματα, με κατεύθυνση αντίθετη του $\nabla f(x_1, x_2)$. Πιο συγκεκριμένα, $x_{k+1} = x_k - \gamma_k \nabla f(x_k)$, με $x_k \in R^2$.



ΝΙΟΒΗ ΠΕΡΑΤΙΚΟΥ
AEM: 10494

<u>Θέμα 1:</u>

Εφαρμογή μεθόδου

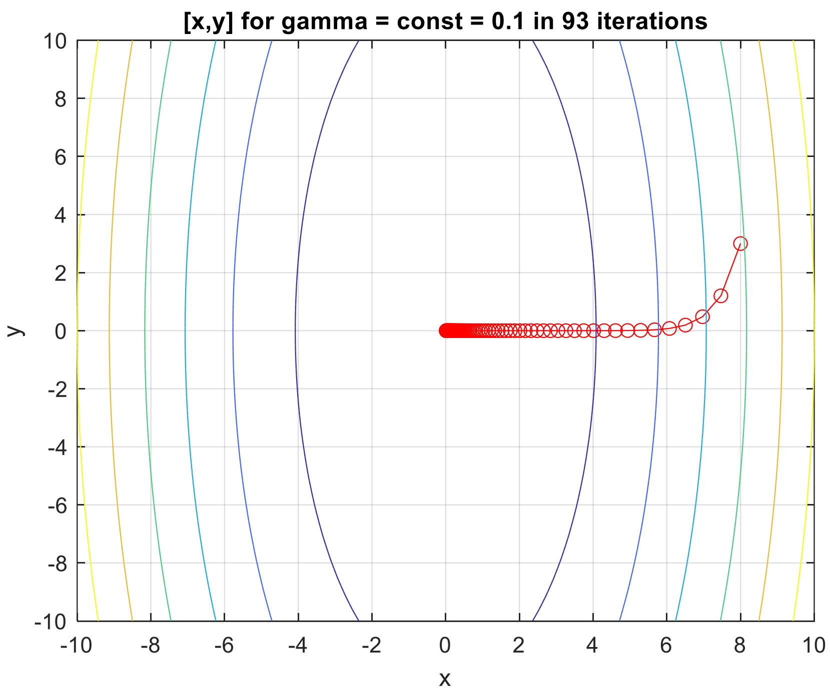
Για την μέγιστη κάθοδο, ισχύει $x_{k+1} = x_k - \gamma_k \nabla f(x_k)$. (1)

Γνωρίζοντας τον τύπο της f, υπολογίζουμε το $\nabla f(x) = x$, με x ∈ R^2.

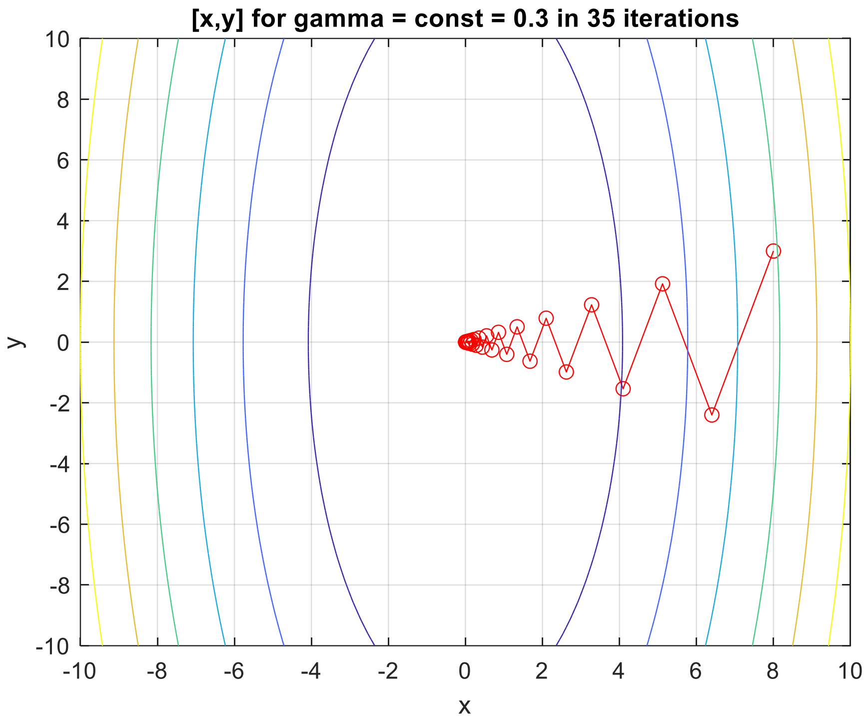
Τότε, η (1) → $x_{k+1} = x_k - \gamma_k x_k = (1 - \gamma_k) * x_k$ (2)

Θεωρώντας ακρίβεια ε = 0.001,επιλέγοντας αρχικό σημείο το (8,3) και ακολουθώντας τα βήματα της εκφώνησης έχουμε:

I. Για γ_k = 0.1, θα έχουμε σύγκλιση αλλά θα είναι αργή.



II. Για $\gamma_k = 0.3$, θα έχουμε σύγκλιση αλλά θα είναι αργή.

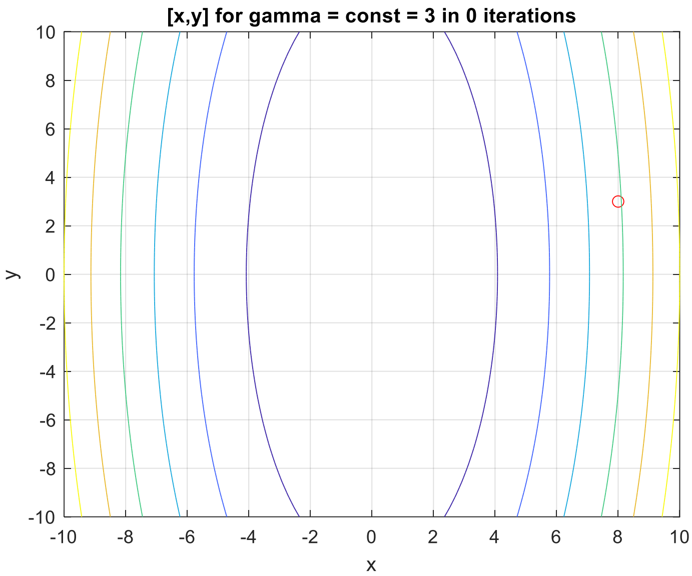


III. Για γ_k=3, θα είχα $x_{k+1} = -2x_k \Rightarrow |x_{k+1}| = 2 * |x_k|$

Δηλαδή, για αρχικό σημείο x_0 θα ίσχυε:

$$|x_1| = 2 * |x_0|$$
$$|x_2| = 2 * |x_1| = 2^2 * |x_0|$$
$$|x_3| = 2 * |x_2| = 2^3 * |x_0|$$
$$\ldots$$
$$|x_k| = 2 * |x_{k-1}| = 2^k * |x_0|$$
$$\text{Για } k \to \infty, \text{ θα έχω } \lim_{k\to\infty} 2^k * |x_0| = \infty \quad (3)$$

Το μέτρο του x_k θα αυξάνει συνεχώς και επομένως δεν θα υπάρξει σύγκλιση του αλγορίθμου.

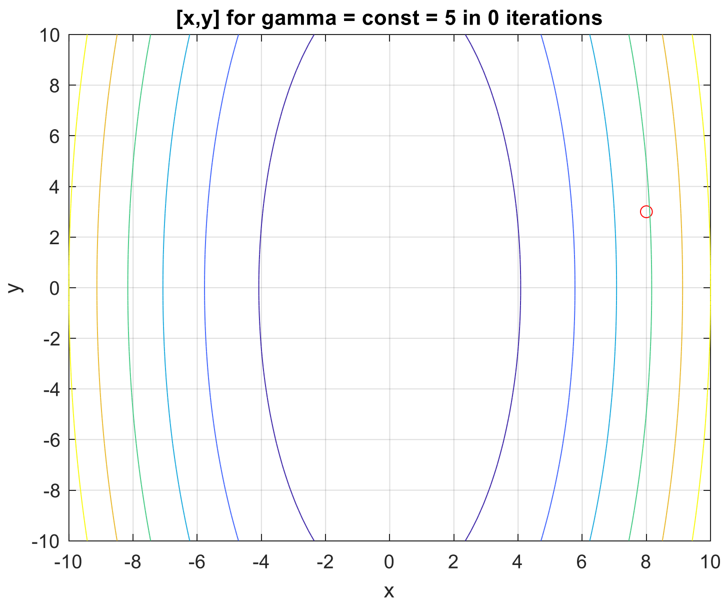


IV. Για γ

Δηλαδή, για αρχικό σημείο x_0 θα ίσχυε:

$$|x_1| = 4 * |x_0|$$
$$|x_2| = 4| = 4^2 * |x_0|$$
$$|x_3| = 4 * |x_2| = 4^3 * |x_0|$$
$$\ldots$$
$$|x_k| = 4 * |x_{k-1}| = 4^k * |x_0|$$
$$\text{Για } k \to \infty, \text{ θα έχω } \lim_{k\to\infty} 4^k * |x_0| = \infty \quad (3)$$

Το μέτρο του x_k θα αυξάνει συνεχώς και επομένως δεν θα υπάρξει σύγκλιση του αλγορίθμου



ΝΙΟΒΗ ΠΕΡΑΤΙΚΟΥ
ΑΕΜ: 10494

<u>Θέματα 2/3/4:</u>

Στην συνέχεια ισχύουν οι παρακάτω περιορισμοί για τις μεταβλητές της συνάρτησης.

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

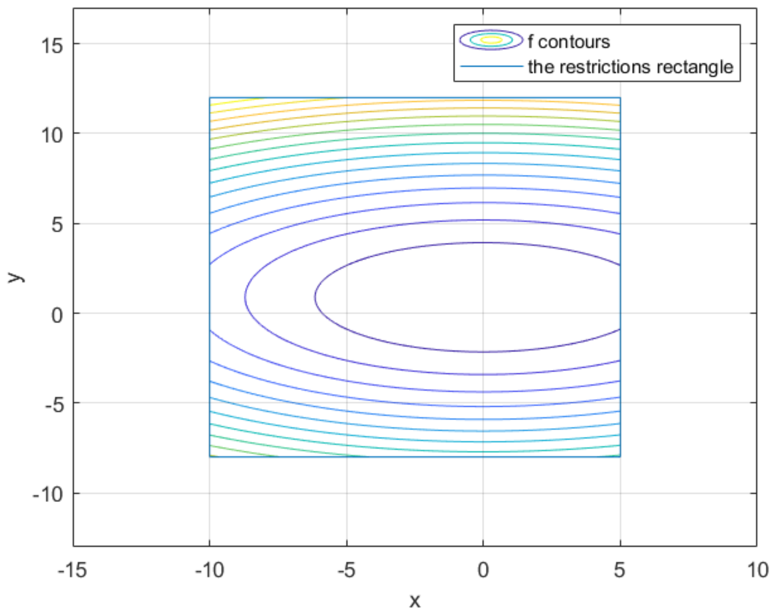
Δηλαδή, δημιουργείται το σύνολο X = { x ∈ R^2 : $-10 \leq x_1 \leq 5$ και $-8 \leq x_2 \leq 12$ }.

Εκμεταλλευόμαστε το γεγονός, ότι οι περιορισμοί εκφράζονται ως φράγματα και ορίζουμε:

$$[PrX\{x_1\}] = \begin{cases} -10, & x \leq -10 \\ x, & -10 < x < 5 \\ 5, & x \geq 5 \end{cases} \quad \text{και} \quad [PrX\{x_2\}] = \begin{cases} -8, & x \leq -8 \\ x2, & -8 < x < 12 \\ 12, & x \geq 12 \end{cases}$$

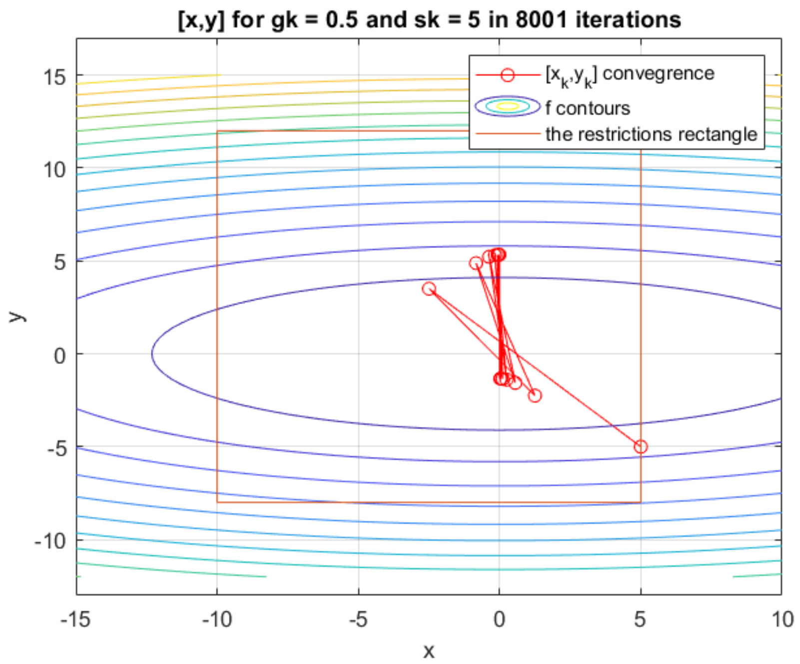
τις προβολές των x_1,x_2 στο X.

Τέτοιου είδους περιορισμοί, δημιουργούν παραλληλόγραμμα που φράζουν στο εσωτερικό τους, τις τιμές της f, όπως φαίνεται στο παρακάτω σχήμα που το ορθογώνιο των περιορισμών περιβάλλει τις ισοβαρείς της f.



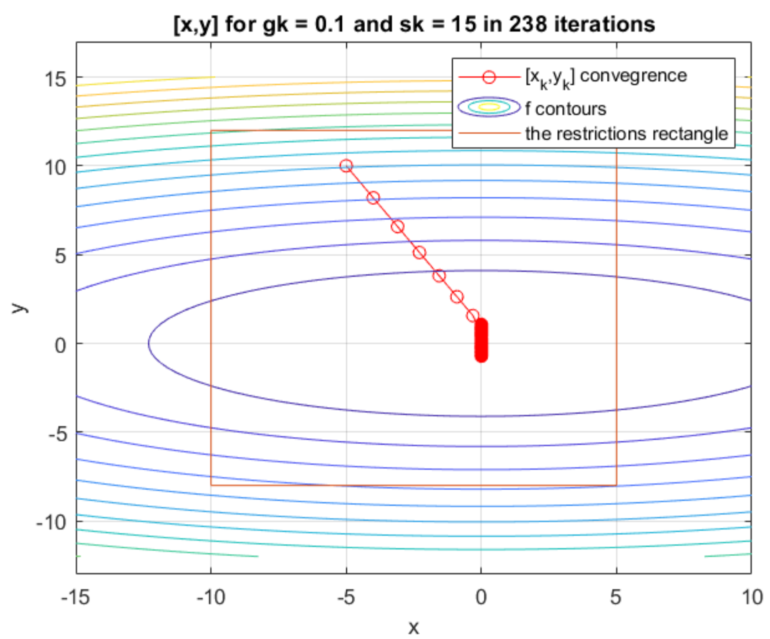
2. $\varepsilon = 0.01$, $\gamma_k = 0.5$ και $s_k = 5$. Αρχικό σημείο (5,-5).

Εκτελώντας το script steepest_descent_projection_plot_x_y παρατηρούμε, ότι η μέθοδος δεν συγκλίνει.



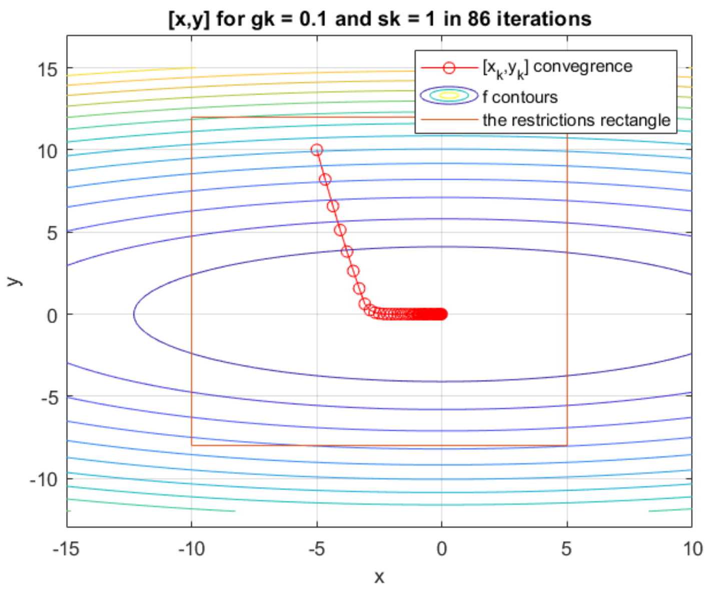
3. $\varepsilon = 0.01$, $\gamma_k = 0.1$ και $s_k = 15$. Αρχικό σημείο (-5,10).

Σκεπτόμενοι με την ίδια λογική, παρατηρούμε πως το βήμα s_k έχει πολύ μεγάλη τιμή και αναμένουμε να μην υπάρξει σύγκλιση. Τρέχοντας, τον αλγόριθμο της μέγιστης καθόδου με προβολή δεν συγκλίνουμε στο ελάχιστο.



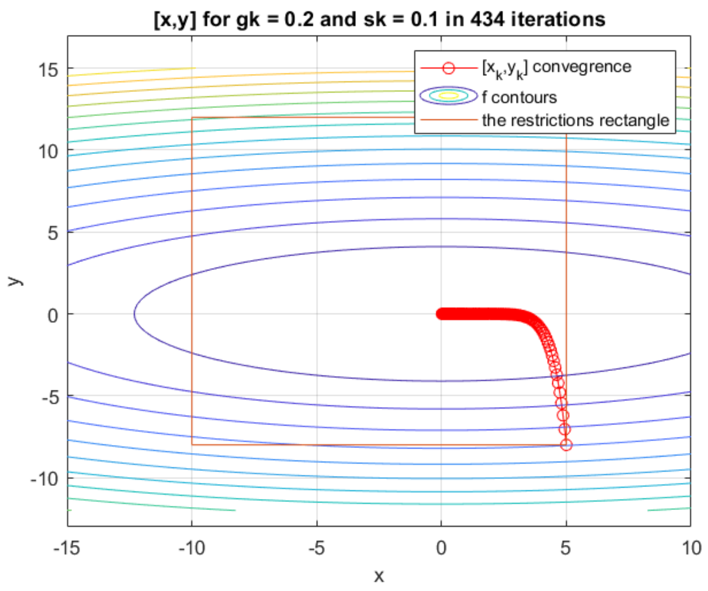
Μια πρακτική λύση για την σύγκλιση του αλγορίθμου είναι η αλλαγή του βήματος $s_k = 0.8$, έτσι ώστε η μέθοδος της μέγιστης καθόδου με προβολή να εκφυλιστεί στην μέθοδο της μέγιστης καθόδου, χωρίς περιορισμούς.

Με την χρήση του ίδιου script και $s_k = 1$



4. ε = 0.01, $\gamma_k = 0.2$ και $s_k = 0.1$. Αρχικό σημείο (8,-10).

Αρχικά, το σημείο που δίνεται <u>δεν</u> είναι εφικτό, άρα προβάλλεται στο σύνολο X, έτσι ώστε να πληροί τις προϋποθέσεις για την εκκίνηση του αλγορίθμου. Στην συνέχεια, το βήμα s_k έχει πολύ μικρή τιμή, που μπορεί να οδηγήσει σε αργή σύγκλιση.



Συμπεράσματα:

Ισχύει ότι, $\overline{xk}$ = Pr{ x_k -$s_k\nabla f(x_k)$} = Pr{ x_k - $s_k x_k$ } και x_{k+1} = x_k + $\gamma_k * (\overline{xk} - x_k)$

Το $\overline{xk}$, μπορεί να διακριθεί σε τρεις περιπτώσεις:

i. $x_k - s_k x_k \leq a_i$, τότε Pr{ $x_k - s_k x_k$ } = a_i

Για το επόμενο σημείο, θα έχω $x_{k+1} = x_k + \gamma_k * (a_i - x_k) = (1 - \gamma_k)\, x_k + \gamma_k\, a_i$

$$\Rightarrow |x_{k+1}|=|(1-\gamma_k)x_k+\gamma_k a_i|\leq|(1-\gamma_k)x_k|+|\gamma_k a_i|$$

Για $0 < \gamma_k \leq 1$ και (αγνοώντας τον όρο $|\gamma_k a_i|$), καταλήγουμε $|x_{k+1}|=|(1-\gamma_k)||x_k|$. Δηλαδή, συνεχώς το επόμενο σημείο θα μειώνεται και θα οδηγείται στο ελάχιστο.

ii. $a_i < x_k - s_k x_k < b_i$, τότε Pr{$x_k - s_k x_k$ } = $x_k - s_k x_{ki}$

Για το επόμενο σημείο,θα έχω $x_{k+1} = x_k + \gamma_k * (x_k - s_k x_k - x_k) = (1 - \gamma_k s_k)\, x_k$

$$\Rightarrow |x_{k+1}|=|(1-\gamma_k s_k)x_k| = |x_{k+1}| = |1-\gamma_k s_k| * x_k$$

Για $0 < 1-\gamma_k s_k < 2$, το $|x_k+1|<|x_k|$. Όμοια με πριν, υπάρχει σύγκλιση προς το ελάχιστο.

iii. $x_k - s_k x_k \geq b_i$, τότε Pr{ $x_k - s_k x_k$ } = b_i

Για το επόμενο σημείο, θα έχω $x_{k+1} = x_k + \gamma_k * (b_i - x_k) = (1 - \gamma_k)x_k + \gamma_k\, b_i$

$$\Rightarrow |x_{k+1}|=|(1-\gamma_k)x_k + \gamma_k b_i| \leq |(1 - \gamma_k)x_k| + |\gamma_k b_i|$$

Για $0 < \gamma_k \leq 1$ και (αγνοώντας τον όρο $|\gamma_k b_i|$), καταλήγουμε $|x_{k+1}| \leq |(1 - \gamma_k)x_k|$.

Το x_{k+1} θα μειώνεται με αποτέλεσμα στην j-οστή επανάληψη κι έπειτα να ισχύει $a_i < x_k - s_k x_k < b_i$, ,$\forall\, k>j$ και να ανήκει το $\overline{xk}$ στην (ii).

Σημείωση: Πρέπει το j < STEP_MAX, το $\overline{xk}$ και το x_{k+1} ορίζουν διανυσματικές σχέσεις και ο δείκτης i, των a,b, παίρνει τιμές i = 1,2 αναφερόμενος στο x_1 και στο x_2, αντίστοιχα.

Η παραπάνω ανάλυση, μας δείχνει ότι η σύγκλιση της μέγιστης καθόδου με προβολή, βασίζεται σε βήμα $\gamma_k s_k$. Εκμεταλλευόμενοι, το όριο της σχέσης (4) και χρησιμοποιώντας το βήμα $\gamma_k s_k$, βρίσκουμε

$$\lim_{k\to\infty} |1\text{-}\gamma s|^k * |x_0| = \begin{cases} 0, & \text{για } 0 < \gamma s < 2 \\ \infty, & \text{για } \gamma s \geq 2 \end{cases}$$