24/5/2018

Άσκηση 4 : Εύρεση ελαχίστου σε συνάρτηση δύο μεταβλητών με ύπαρξη περιορισμών

ΓΙΩΡΓΟΣ ΜΠΑΛΑΟΥΡΑΣ (8861)

Τεχνικές Βελτιστοποίησης

Ονοματεπώνυμο: *Μπαλαούρας Γεώργιος*

AEM: *8861*

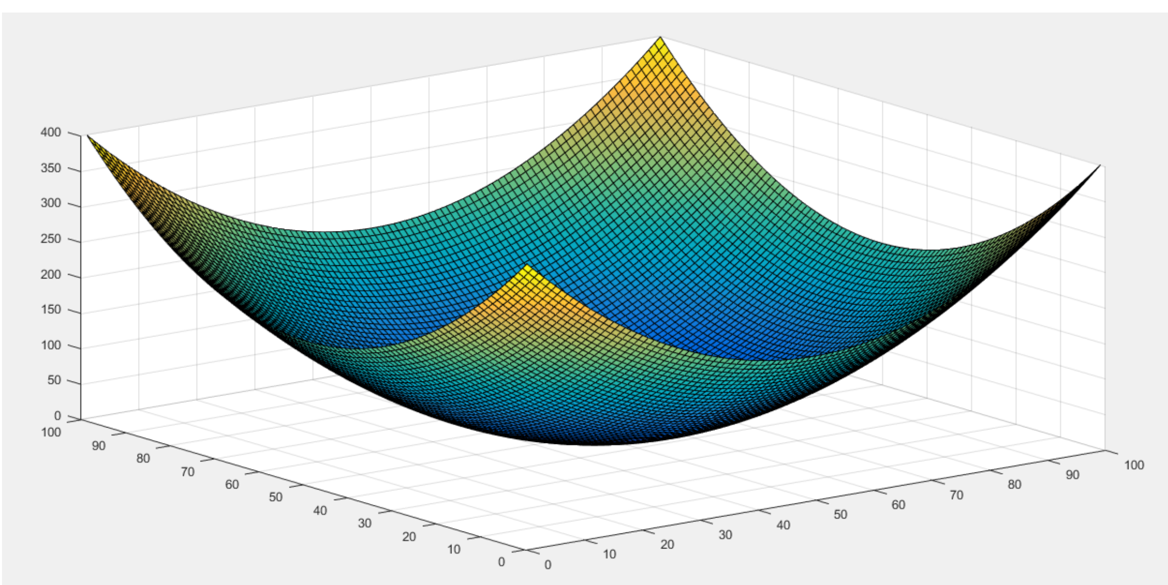
E-mail: *mpalaourg@ece.auth.gr*

Εξάμηνο: *6ο Ηλεκτρονικής*

Άσκηση 4 : *Εύρεση ελαχίστου σε συνάρτηση δύο μεταβλητών με ύπαρξη περιορισμών*

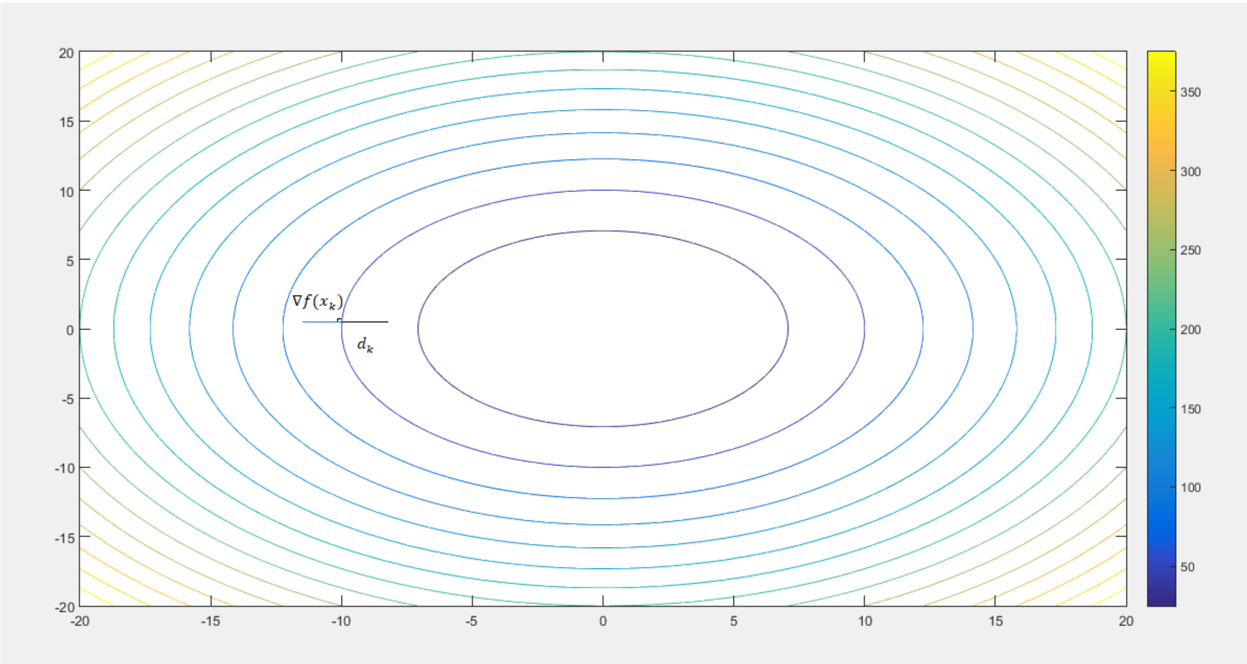
Αντικείμενο αυτής της άσκησης, ήταν η εφαρμογή της *μεθόδου της μέγιστης καθόδου* σε δύο περιπτώσεις και η σύγκριση των αποτελεσμάτων. Αρχικά, υλοποιείται για συνάρτηση f για την οποία δεν ισχύουν κάποιοι περιορισμοί και στην συνέχεια χρησιμοποιείται για την ελαχιστοποίηση της ίδια συνάρτησης υπό την παρουσία, όμως, περιορισμών.

Η f έχει τετραγωνική μορφή και δίνεται από τον τύπο $f(x_1, x_2) = \frac{x_1^2}{2} + \frac{x_2^2}{2}$ με $x_1, x_2 \in \mathbb{R}$.



Σχήμα 1: Η γραφική παράσταση της f

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 <u>αντίθετη</u> του $\nabla f(x_1, x_2)$. Πιο συγκεκριμένα, $x_{k+1} = x_k - \gamma_k \nabla f(x_k)$, με $x_k \in \mathbb{R}^2$.



Σχήμα 2: Ισοβαρείς καμπύλες της f

Εφαρμογή μεθόδου

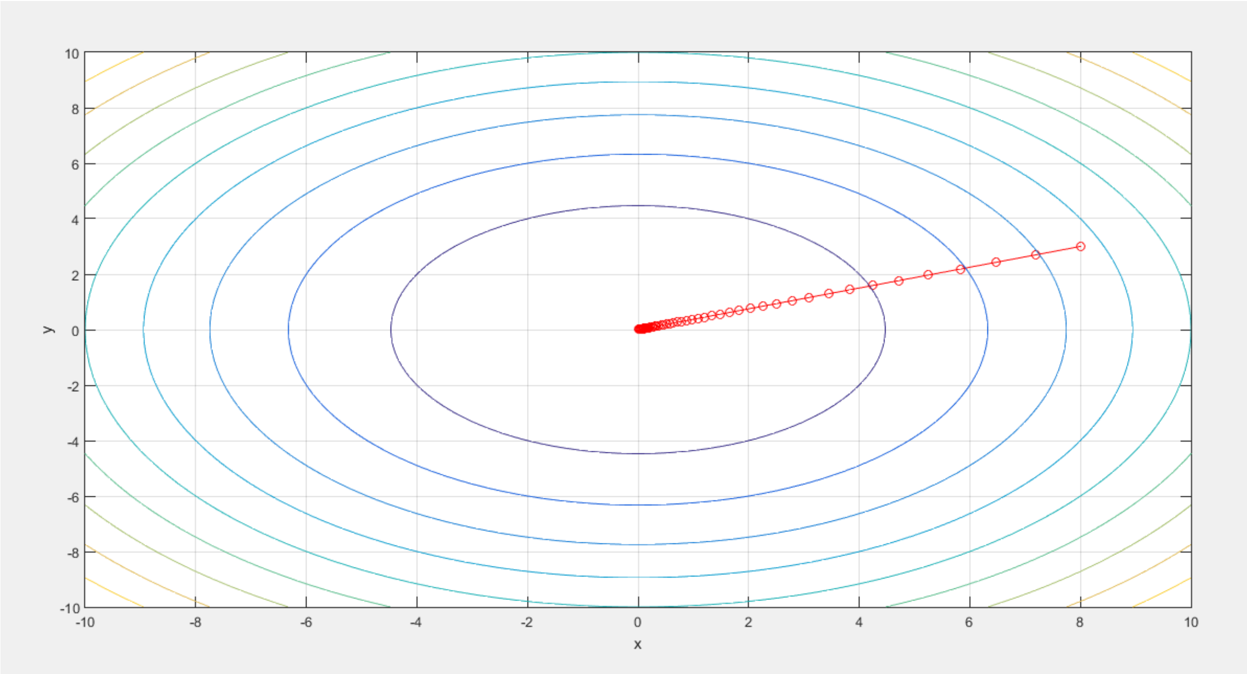
Για την *μέγιστη κάθοδο*, ισχύει $x_{k+1} = x_k - \gamma_k \nabla f(x_k)$. (1)

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

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

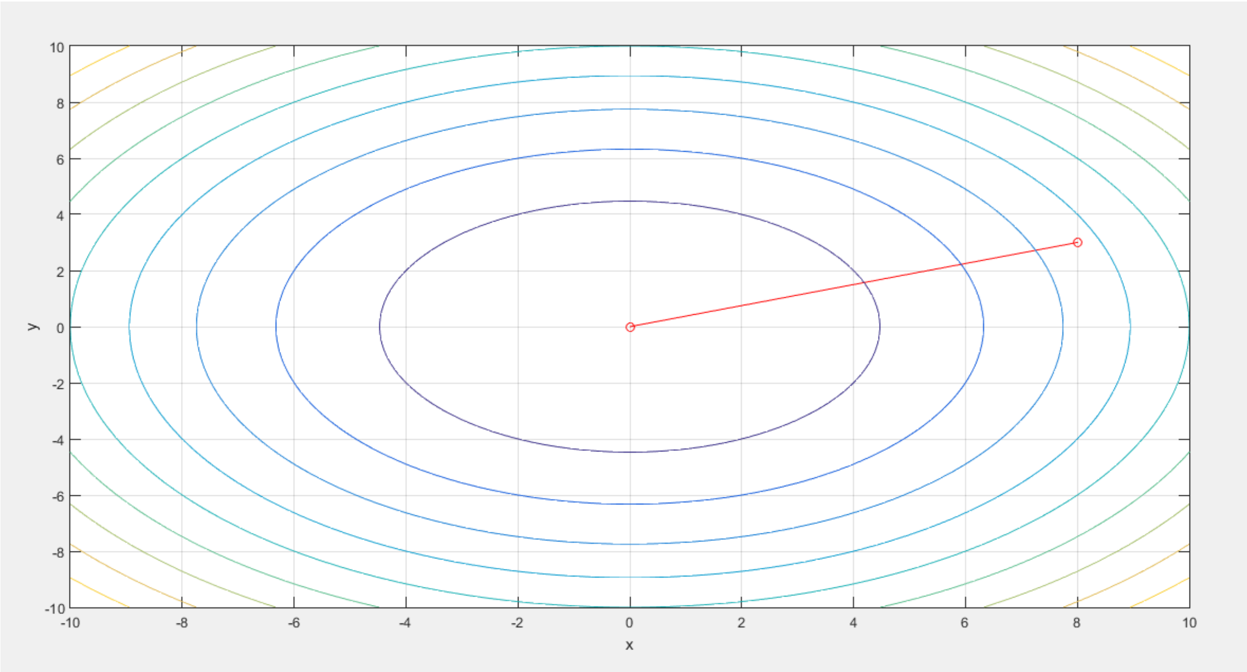
Θεωρώντας ακρίβεια $\varepsilon = 0.01$ και ακολουθώντας τα βήματα της εκφώνησης έχουμε:

I. Για $\gamma = 0.1$, η $(2) \Rightarrow\ x_{k+1} = 0.9 * x_k$. Δηλαδή, θα έχουμε σύγκλιση *αλλά* θα είναι αργή.



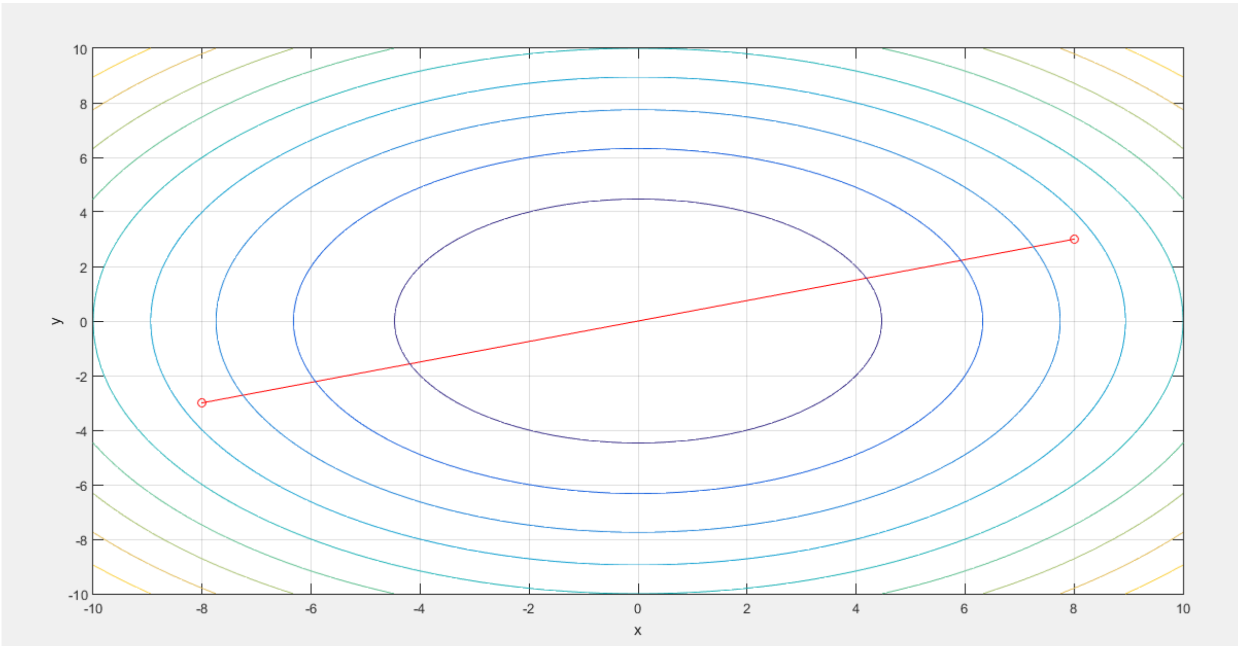
Σχήμα 3: Σύγκλιση του ζεύγους (x, y) για σταθερό γ = 0.1

II. Για $\gamma = 1$, η (2) μηδενίζει και περιμένω σύγκλιση σε ένα και μόνο βήμα.



Σχήμα 4: Σύγκλιση του ζεύγους (x, y) για σταθερό γ = 1

III. Για $\gamma = 2$, η $(2) \Rightarrow\ x_{k+1} = -\ x_k$. Δηλαδή θα έχω ταλάντωση μεταξύ δύο συμμετρικών σημείων και ο αλγόριθμος θα τερματίσει λόγω του ανώτατου ορίου βημάτων.



Σχήμα 5: Σύγκλιση του ζεύγους (x, y) για σταθερό γ = 2

IV. Ενώ, για $\gamma = 10$, θα είχα $x_{k+1} = -9x_k \Rightarrow |x_{k+1}| = 9 * |x_k|$

Δηλαδή, για αρχικό σημείο x_0 θα ίσχυε:

$$|x_1| = 9 * |x_0|$$
$$|x_2| = 9 * |x_1| = 9^2 * |x_0|$$
$$|x_3| = 9 * |x_2| = 9^3 * |x_0|$$
$$\ldots$$
$$|x_k| = 9 * |x_{k-1}| = 9^k * |x_0|$$
$$\text{Για } k \to \infty, \text{θα έχω } \lim_{k\to\infty} 9^k * |x_0| = \infty \quad (3)$$

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

Πιο συγκεκριμένα, το όριο της σχέσης (3), για διάφορες τιμές του γ:

$$\lim_{k\to\infty} |1-\gamma|^k * |x_0| = \begin{cases} 0, & \text{για } 0 < \gamma < 2 \\ \infty, & \text{για } \gamma \geq 2 \end{cases} \qquad (4)$$

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

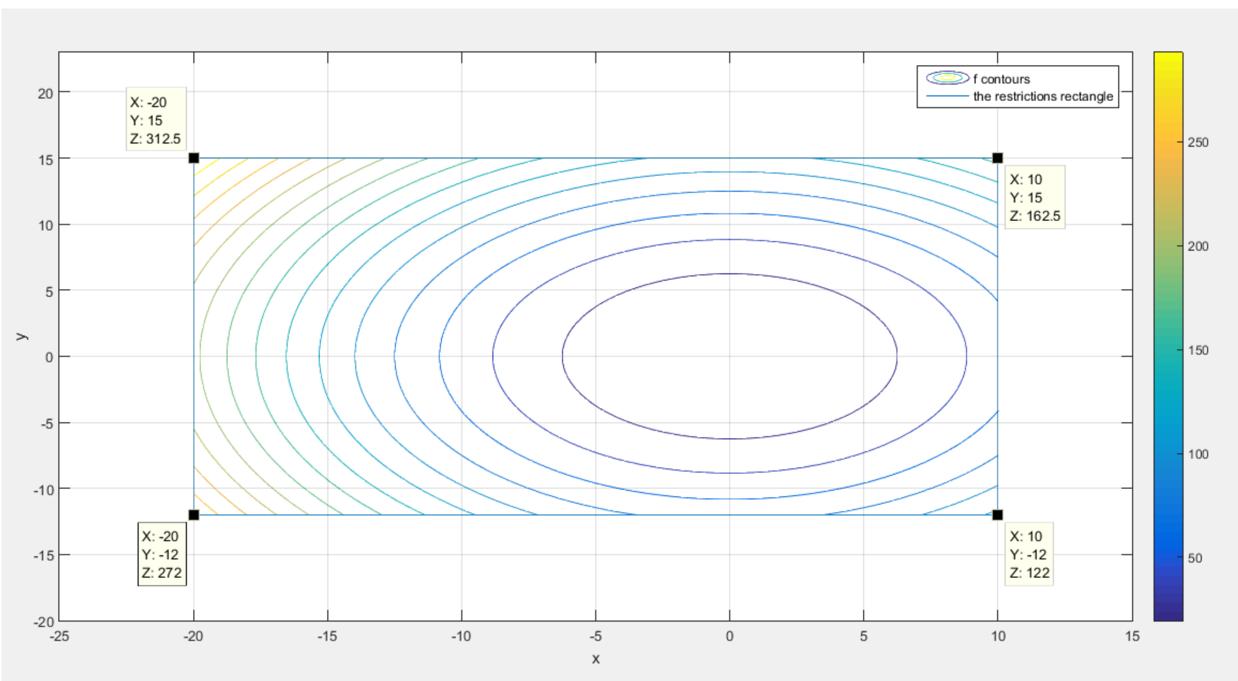
$$-20 \leq x_1 \leq 10 \text{ και } -12 \leq x_2 \leq 15.$$

Δηλαδή, δημιουργείται το σύνολο $X = \{x \in \mathbb{R}^2 : -20 \leq x_1 \leq 10 \text{ και } -12 \leq x_2 \leq 15\}$. Εκμεταλλευόμαστε το γεγονός, ότι οι περιορισμοί εκφράζονται ως φράγματα και ορίζουμε:

$$[Pr_X\{x\}]_1 = \begin{cases} -20, & x_1 \leq -20 \\ x_1, & -20 < x_1 < 10 \\ 10, & x_1 \geq 10 \end{cases} \quad \underline{\textbf{και}} \quad [Pr_X\{x\}]_2 = \begin{cases} -12, & x_2 \leq -12 \\ x_2, & -12 < x_2 < 15 \\ 15, & x_2 \geq 15 \end{cases}$$

τις προβολές των x_1, x_2 στο X.

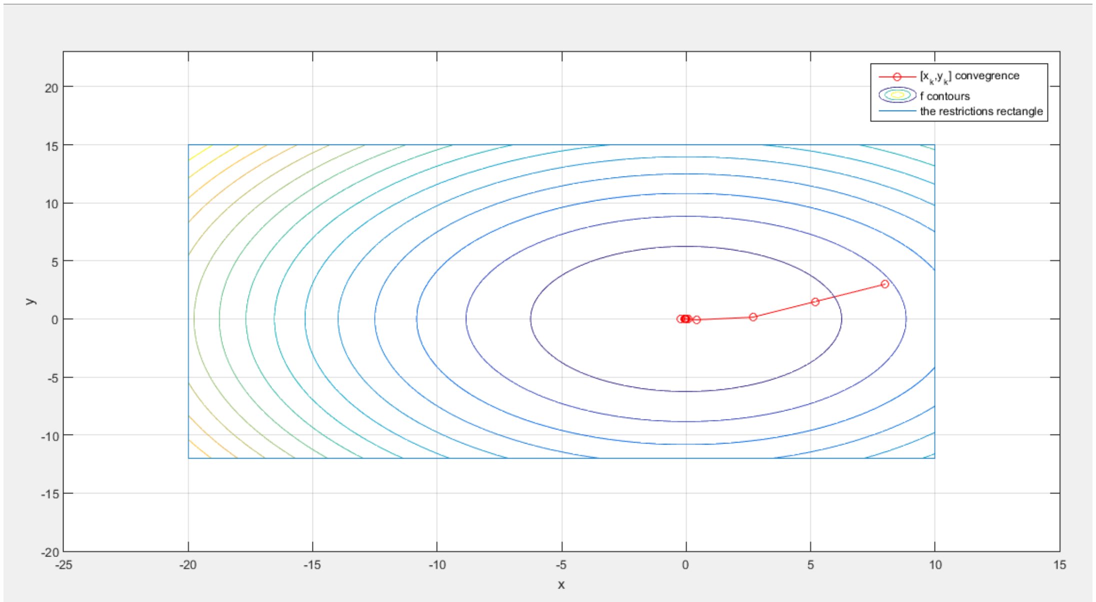
Τέτοιου είδους περιορισμοί, δημιουργούν παραλληλόγραμμά που φράζουν στο εσωτερικό τους, τις τιμές της f, όπως φαίνεται στο παρακάτω σχήμα που το ορθογώνιο των περιορισμών περιβάλλει τις ισοβαρείς της f.



Σχήμα 6: Ο χώρος των περιορισμών που περιβάλλει τις τιμές της f

A. $\varepsilon = 0.01$, $\gamma_k = 0.1$ και $s_k = 15$. Αρχικό σημείο $(8{,}3)$.

Πριν την εκτέλεση του αλγορίθμου, παρατηρούμε ότι το βήμα s_k είναι μεγαλύτερο σε σχέση με την (IV) περίπτωση εκτέλεσης της *μέγιστης καθόδου*. Ωστόσο, εκτελώντας το `script steepest_descent_with_projection_plot_x_y` παρατηρούμε, ότι η μέθοδος συγκλίνει.



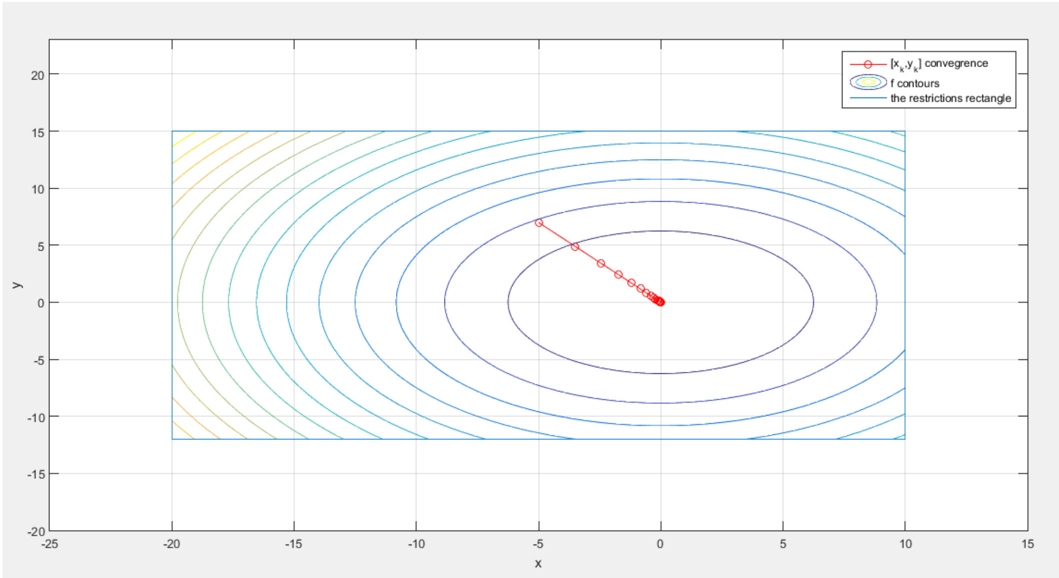
Σχήμα 7: Σύγκλιση του σημείου (8,3) για $\gamma_k = 0.1$ *και* $s_k = 15$

B. $\varepsilon = 0.02$, $\gamma_k = 0.3$ και $s_k = 20$. Αρχικό σημείο $(-5{,}7)$.

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.

Παρατηρώντας ότι το αρχικό σημείο που μας δίνεται είναι κι εφικτό, μια πρακτική λύση για την σύγκλιση του αλγορίθμου είναι η αλλαγή του βήματος $s_k = 1$, έτσι ώστε η *μέθοδος της μέγιστης καθόδου με προβολή* να εκφυλιστεί στην *μέθοδο της μέγιστης καθόδου, χωρίς περιορισμούς*.

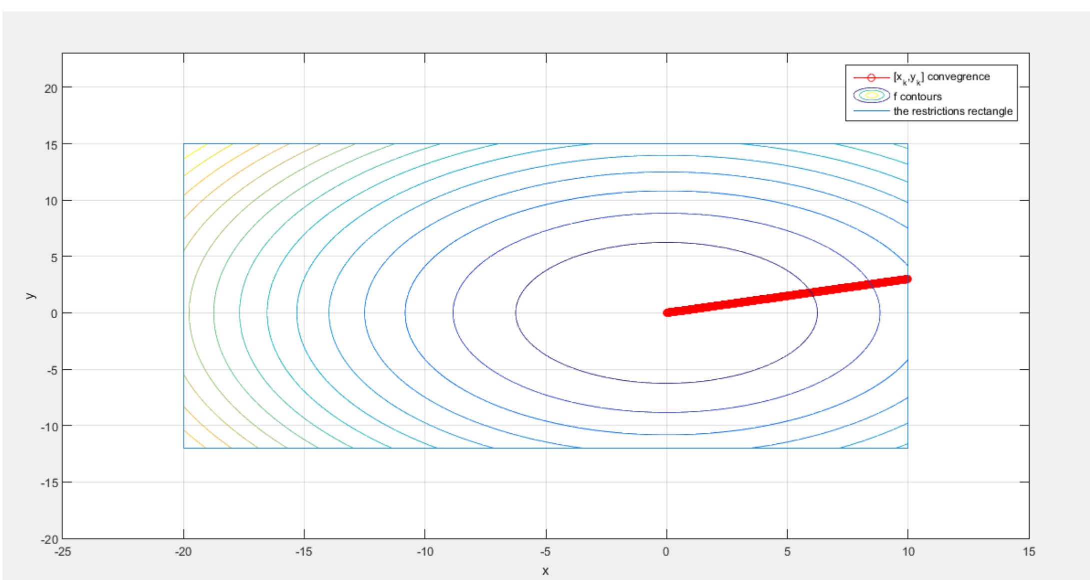
Με την χρήση του ίδιου `script` και $s_k = 1$



Σχήμα 8: Σύγκλιση του σημείου (-5,7) για $\gamma_k = 0.3$ *και* $s_k = 1$

C. $\varepsilon = 0.01$, $\gamma_k = 0.01$ και $s_k = 0.1$. Αρχικό σημείο $(11,3)$.

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.



Σχήμα 9: Σύγκλιση του σημείου (11,3) για $\gamma_k = 0.01$ *και* $s_k = 0.1$

Συμπεράσματα

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

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

i. $x_k - s_k x_k \leq \alpha_i$, τότε $\Pr_X\{x_k - s_k x_k\} = \alpha_i$

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

Δηλαδή, συνεχώς το επόμενο σημείο θα μειώνεται και θα οδηγείται στο ελάχιστο.

ii. $a_i < x_k - s_k x_k < b_i$, τότε $\Pr_X\{x_k - s_k x_k\} = x_k - s_k x_{k_i}$

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 $\Pr_X\{x_k - s_k x_k\} = b_i$

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 $\overline{x_k}$ στην (ii).

<u>Σημείωση</u>: Πρέπει το $j < STEP\_MAX$, το $\overline{x_k}$ 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 $\lim_{k\to\infty} |1 - \gamma s|^k * |x_0| = \begin{cases} 0, & \gamma\iota\alpha\ 0 < \gamma s < 2 \\ \infty, & \gamma\iota\alpha\ \gamma s \geq 2 \end{cases}$

Παρατηρώντας τα παραπάνω ερωτήματα, στα (A), (C) το κριτήριο για την σύγκλιση ικανοποιείται, αφού είναι 1.5 και 0.001 αντίστοιχα. Ωστόσο, στο (B) το βήμα είναι ίσο με 5 και η μη-σύγκλιση του αλγορίθμου δικαιολογείται. Τέλος, στο (C) το βήμα είναι πολύ μικρό και οι 6949 επαναλήψεις που χρειάζονται το αποδεικνύουν.