3η Εργαστηριακή Άσκηση

Θέμα:

Μέθοδος Μέγιστης Καθόδου με Προβολή

Μάθημα: τεχνικές βελτιστοποίησης

ΑΕΜ: 10216

Όνομα: Καλαϊτζίδης Ελευθέριος

Email: eleftherk@ece.auth.gr

ΘΕΜΑ 1

$$f(x_1, x_2) = \frac{1}{3}{x_1}^2 + 3{x_2}^2$$

$$\nabla f(x_1, x_2) = \begin{bmatrix} {}^2/_3\, x_1 \\ 6x_2 \end{bmatrix}$$

$$x_{k+1} = x_k - \gamma \nabla f(x_1, x_2)$$

$$\begin{cases} x_{1,k+1} = x_{1,k} - \gamma \dfrac{2}{3} x_{1,k} \\ x_{2,k+1} = x_{2,k} - \gamma 6 x_{2,k} \end{cases}$$

Για να συγκλείνει πρέπει :

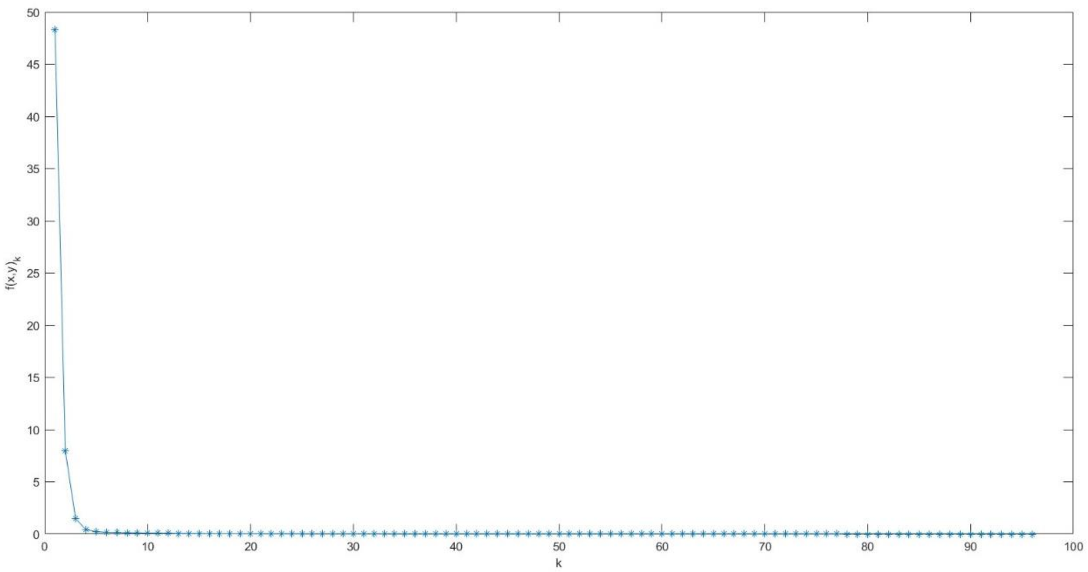
$$\begin{cases} \dfrac{|\chi_{1,\kappa+1}|}{|\chi_{1,\kappa}|} < 1 \\ \dfrac{|\chi_{2,\kappa+1}|}{|\chi_{2,\kappa}|} < 1 \end{cases} => \begin{cases} \dfrac{|(1-\gamma\, {}^2/_3)\chi_{1,\kappa}|}{|\chi_{1,\kappa}|} < 1 \\ \dfrac{|(1-\gamma 6)\chi_{2,\kappa}|}{|\chi_{2,\kappa}|} < 1 \end{cases} => \begin{cases} |1-\gamma\, {}^2/_3| < 1 \\ |1-\gamma 6| < 1 \end{cases} => \begin{cases} 0 < \gamma < 3 \\ 0 < \gamma < 0{,}333 \end{cases} =>$$

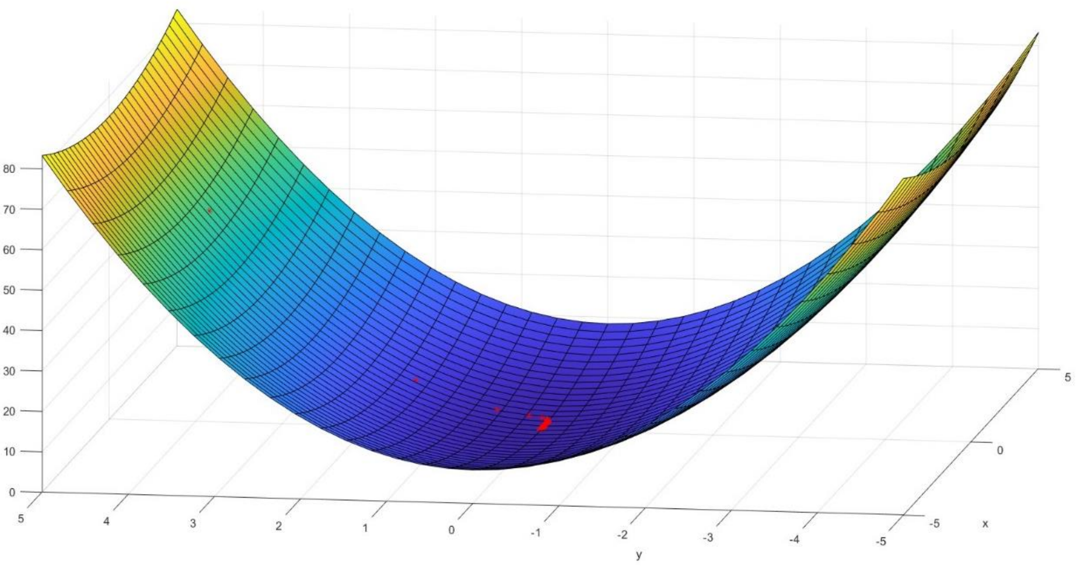
$$\gamma \epsilon (0, 0.333)$$

Παρατηρώ ότι για να συγκλείνει η μέθοδος θα πρέπει το γ_κ να παίρνει τιμές μεταξύ 0 και 0.333 αλλιώς αποκλίνει.

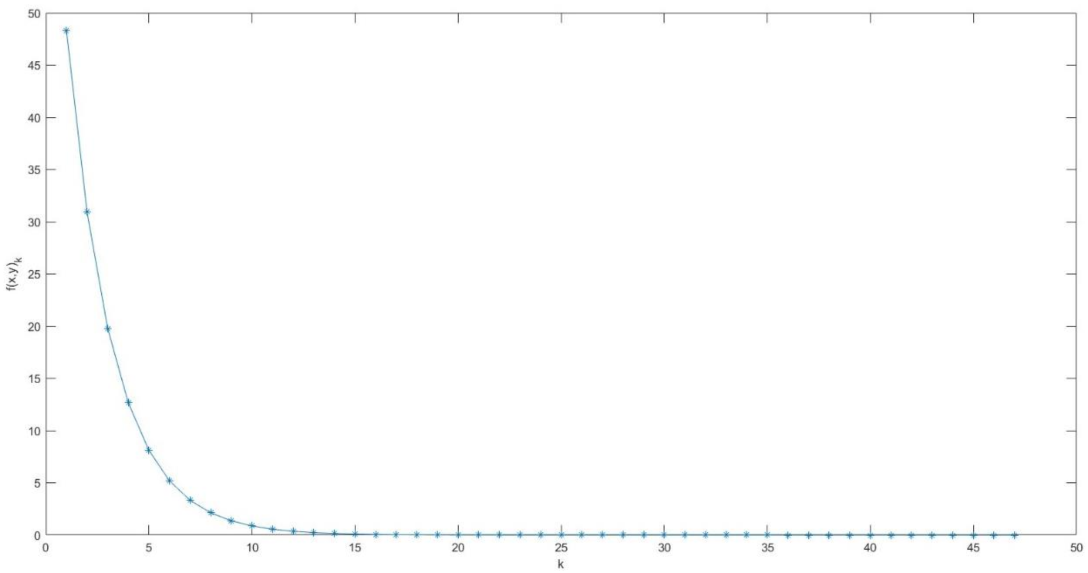
Αυτό φαίνεται και από τις δοκιμές μας που μόνο για τιμές του γ_κ ίσες με 0.1 και 0.3 βρίσκει το ελάχιστο ενώ για γ_κ ίσο με 3 και 5 αποκλίνει.

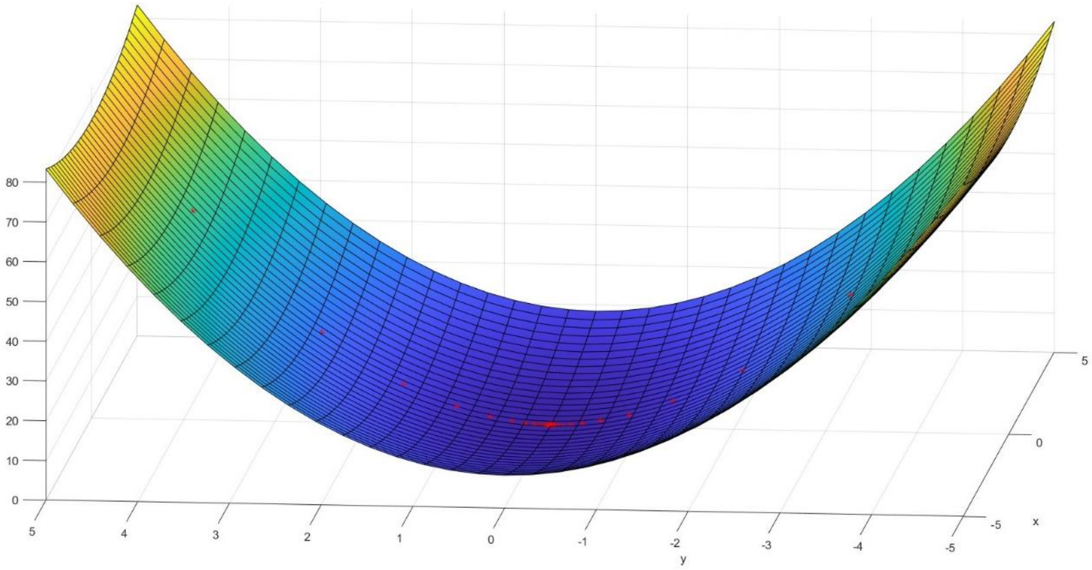
Για γκ = 0.1:





Για γ_κ=0.3:





ΘΕΜΑ 2

$$f(x_1, x_2) = \frac{1}{3}{x_1}^2 + 3{x_2}^2$$

$$\nabla f(x_1, x_2) = \begin{bmatrix} {}^2/_3\, x_1 \\ 6x_2 \end{bmatrix}$$

$$x_{k+1} = x_k + \gamma_\kappa(\bar{\chi}_\kappa - \chi_\kappa)$$

$$\bar{\chi}_\kappa = \Pr\left[\chi_\kappa - s_k * \nabla f(\chi_\kappa)\right]$$

$$\bar{\chi}_\kappa = \Pr\begin{bmatrix} \chi_1 - s_k * {}^2/_3 * \chi_1 \\ \chi_2 - s_k * 6 * \chi_2 \end{bmatrix}$$ *έστω ότι το σημείο είναι η ίδια η προβολή*

(για να πάρουμε και τους δυο τους περιορισμούς)

$$\bar{\chi}_\kappa = \begin{bmatrix} \chi_1 - s_k * {}^2/_3 * \chi_1 \\ \chi_2 - s_k * 6 * \chi_2 \end{bmatrix}$$

$$x_{k+1} = x_k + \gamma_\kappa(\bar{\chi}_\kappa - \chi_\kappa) \Rightarrow$$

$$\begin{bmatrix} \chi_{1,\kappa+1} \\ \chi_{2,\kappa+1} \end{bmatrix} = \begin{bmatrix} \chi_{1,\kappa} \\ \chi_{2,\kappa} \end{bmatrix} + \gamma_\kappa(\chi_\kappa \begin{bmatrix} 1 - s_k * {}^2/_3 \\ 1 - s_k * 6 \end{bmatrix} - \chi_\kappa)$$

$$\begin{bmatrix} \chi_{1,\kappa+1} \\ \chi_{2,\kappa+1} \end{bmatrix} = \begin{bmatrix} \chi_{1,\kappa} \\ \chi_{2,\kappa} \end{bmatrix} + 0.5(\chi_\kappa \begin{bmatrix} 1 - s_k * {}^2/_3 \\ 1 - s_k * 6 \end{bmatrix} - \begin{bmatrix} \chi_{1,\kappa} \\ \chi_{2,\kappa} \end{bmatrix})$$

$$\begin{bmatrix} \chi_{1,\kappa+1} \\ \chi_{2,\kappa+1} \end{bmatrix} = \begin{bmatrix} \chi_{1,\kappa} \\ \chi_{2,\kappa} \end{bmatrix} + 0.5\chi_\kappa \begin{bmatrix} 1 - s_k * {}^2/_3 \\ 1 - s_k * 6 \end{bmatrix} - 0.5\begin{bmatrix} \chi_{1,\kappa} \\ \chi_{2,\kappa} \end{bmatrix}$$

$$\begin{bmatrix} \chi_{1,\kappa+1} \\ \chi_{2,\kappa+1} \end{bmatrix} = 0.5\begin{bmatrix} \chi_{1,\kappa} \\ \chi_{2,\kappa} \end{bmatrix} + 0.5\begin{bmatrix} \chi_{1,\kappa} \\ \chi_{2,\kappa} \end{bmatrix}\begin{bmatrix} 1 - s_k * {}^2/_3 \\ 1 - s_k * 6 \end{bmatrix}$$

$$\begin{cases} x_{1,k+1} = 0.5x_{1,k} + 0.5\chi_{1,\kappa}(1 - s_k * {}^2/_3) \\ x_{2,k+1} = 0.5x_{2,k} + 0.5\chi_{2,\kappa}(1 - s_k * 6) \end{cases} \Rightarrow$$

$$\begin{cases} x_{1,k+1} = \chi_{1,\kappa}(1 - s_k * {}^1/_3) \\ x_{2,k+1} = \chi_{2,\kappa}(1 - s_k * 3) \end{cases}$$

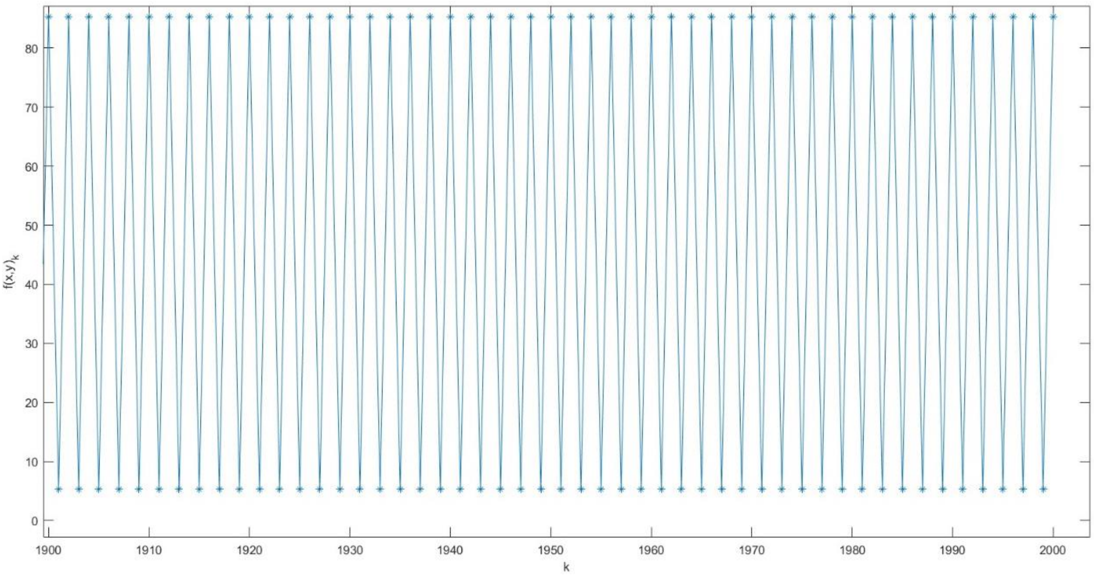
$$\begin{cases} \dfrac{|\chi_{1,\kappa+1}|}{|\chi_{1,\kappa}|} < 1 \\ \dfrac{|\chi_{2,\kappa+1}|}{|\chi_{2,\kappa}|} < 1 \end{cases} \Rightarrow \begin{cases} \dfrac{|\chi_{1,\kappa}(1 - s_k * {}^1/_3)|}{|\chi_{1,\kappa}|} < 1 \\ \dfrac{|(1 - s_k * 3)|}{|\chi_{2,\kappa}|} < 1 \end{cases} \Rightarrow$$

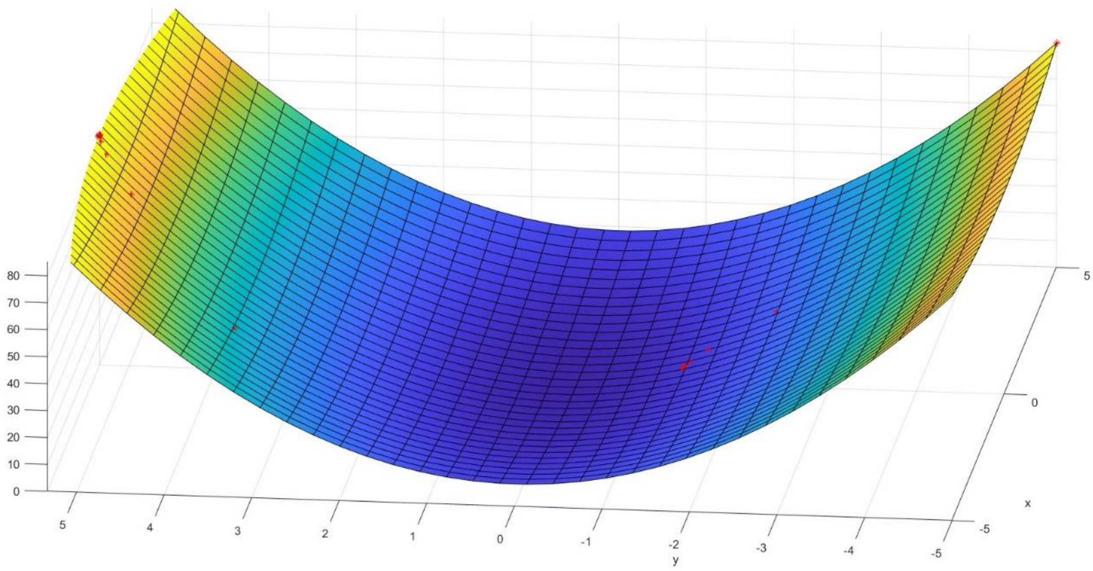
$$\begin{cases} \left|1 - s_k * {}^{1}/_{3}\right| < 1 \\ |1 - s_k * 3| < 1 \end{cases} => \begin{cases} 0 < s_k < 6 \\ 0 < s_k < 0{,}666 \end{cases} =>$$

$s_k \in (0{,}0.666)$

Όμως στο θέμα 2 το $s_k = 5$(εκτός περιορισμού)άρα δικαιολογεί ότι η μέθοδος αποκλίνει για τα δοσμένα νούμερα.

Όπως φαίνεται και παρακάτω σε αρκετά από τα τελευταία βήματα παρατηρούμε ταλάντωση και δεν καταλήγουμε στο ελάχιστο.





Σε σχέση με το θέμα 1 παρατηρώ ότι όπως στε εκείνο η μέθοδος απέκλεινε για μεγάλο γ_κ εδώ η μέθοδος αποκλείνει για μεγάλο s_k. Στο θέμα 1 είχαμε περιορισμό για το γ_κ ενώ τωρα για το s_k.

ΘΕΜΑ 3

$f(x_1, x_2) = \frac{1}{3}{x_1}^2 + 3{x_2}^2$

$\nabla f(x_1, x_2) = \begin{bmatrix} {}^2/_3\, x_1 \\ 6x_2 \end{bmatrix}$

$x_{k+1} = x_k + \gamma_\kappa(\bar{\chi}_\kappa - \chi_\kappa)$

$\bar{\chi}_\kappa = \Pr\,[\chi_\kappa - s_k * \nabla f(\chi_\kappa)\,]$

$\bar{\chi}_\kappa = \Pr \begin{bmatrix} \chi_1 - s_k * {}^2/_3 * \chi_1 \\ \chi_2 - s_k * 6 * \chi_2 \end{bmatrix}$ *έστω ότι το σημείο είναι η ίδια η προβολή*

(για να πάρουμε και τους δυο τους περιορισμούς)

$\bar{\chi}_\kappa = \begin{bmatrix} \chi_1 - s_k * {}^2/_3 * \chi_1 \\ \chi_2 - s_k * 6 * \chi_2 \end{bmatrix}$

$x_{k+1} = x_k + \gamma_\kappa(\bar{\chi}_\kappa - \chi_\kappa)$=>

$\begin{bmatrix} \chi_{1,\kappa+1} \\ \chi_{2,\kappa+1} \end{bmatrix} = \begin{bmatrix} \chi_{1,\kappa} \\ \chi_{2,\kappa} \end{bmatrix} + \gamma_\kappa(\chi_\kappa \begin{bmatrix} 1 - s_k * {}^2/_3 \\ 1 - s_k * 6 \end{bmatrix} - \chi_\kappa)$

$\begin{bmatrix} \chi_{1,\kappa+1} \\ \chi_{2,\kappa+1} \end{bmatrix} = \begin{bmatrix} \chi_{1,\kappa} \\ \chi_{2,\kappa} \end{bmatrix} + 0.1(\chi_\kappa \begin{bmatrix} 1 - s_k * {}^2/_3 \\ 1 - s_k * 6 \end{bmatrix} - \begin{bmatrix} \chi_{1,\kappa} \\ \chi_{2,\kappa} \end{bmatrix})$

$\begin{bmatrix} \chi_{1,\kappa+1} \\ \chi_{2,\kappa+1} \end{bmatrix} = \begin{bmatrix} \chi_{1,\kappa} \\ \chi_{2,\kappa} \end{bmatrix} + 0.1\chi_\kappa \begin{bmatrix} 1 - s_k * {}^2/_3 \\ 1 - s_k * 6 \end{bmatrix} - 0.1 \begin{bmatrix} \chi_{1,\kappa} \\ \chi_{2,\kappa} \end{bmatrix}$

$\begin{bmatrix} \chi_{1,\kappa+1} \\ \chi_{2,\kappa+1} \end{bmatrix} = 0.9 \begin{bmatrix} \chi_{1,\kappa} \\ \chi_{2,\kappa} \end{bmatrix} + 0.1 \begin{bmatrix} \chi_{1,\kappa} \\ \chi_{2,\kappa} \end{bmatrix} \begin{bmatrix} 1 - s_k * {}^2/_3 \\ 1 - s_k * 6 \end{bmatrix}$

$\begin{cases} x_{1,k+1} = 0.9x_{1,k} + 0.1\chi_{1,\kappa}(1 - s_k * {}^2/_3) \\ x_{2,k+1} = 0.9x_{2,k} + 0.1\chi_{2,\kappa}(1 - s_k * 6) \end{cases}$=>

$\begin{cases} x_{1,k+1} = \chi_{1,\kappa}(1 - s_k * {}^2/_{30}) \\ x_{2,k+1} = \chi_{2,\kappa}(1 - s_k * 0.6) \end{cases}$

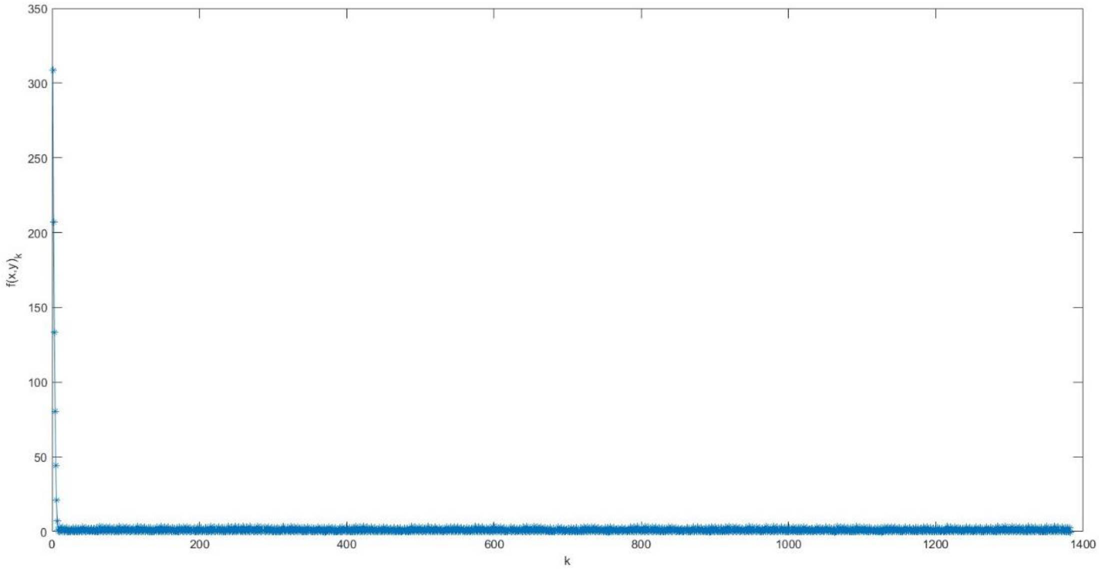
$\begin{cases} \dfrac{|\chi_{1,\kappa+1}|}{|\chi_{1,\kappa}|} < 1 \\ \dfrac{|\chi_{2,\kappa+1}|}{|\chi_{2,\kappa}|} < 1 \end{cases} => \begin{cases} \dfrac{|\chi_{1,\kappa}(1 - s_k * {}^2/_{30})|}{|\chi_{1,\kappa}|} < 1 \\ \dfrac{|(1 - s_k * 0.6)|}{|\chi_{2,\kappa}|} < 1 \end{cases} =>$

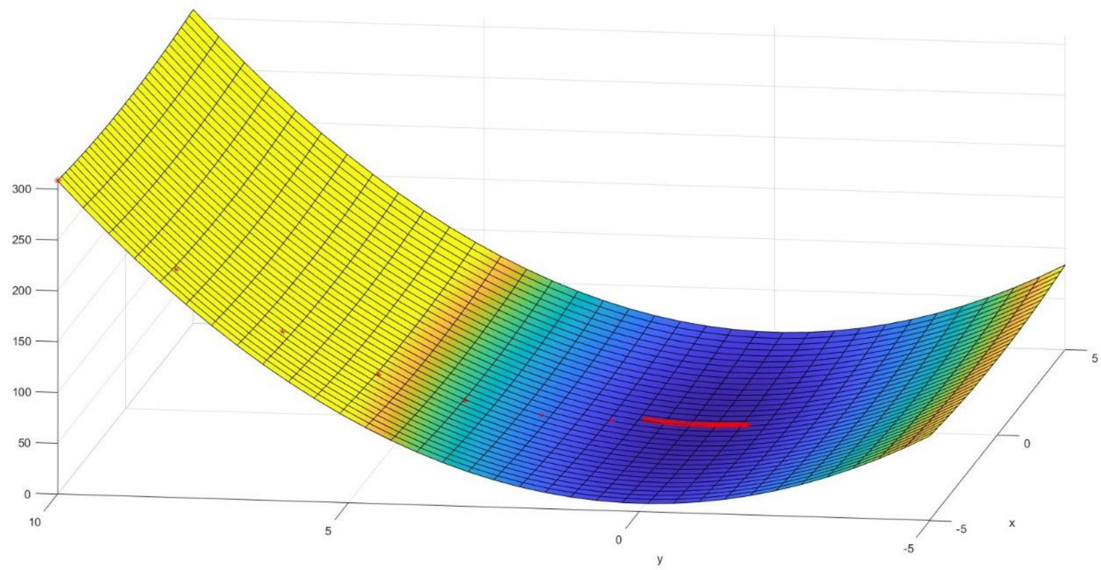
$\begin{cases} |1 - s_k * {}^2/_{30}| < 1 \\ |1 - s_k * 0.6| < 1 \end{cases} => \begin{cases} 0 < s_k < 30 \\ 0 < s_k < 3.333 \end{cases} =>$

$s_k \in (0{,}3.333)$

Βλέπουμε ότι το s_k στο θέμα αυτό είναι ίσο με 15 το οποίο δεν πληροί και τις δύο προϋποθέσεις (και θα περιμέναμε η μέθοδος να αποκλίνει) ωστόσο η μέθοδος συγκλείνει μετά όμως από πολλά βήματα. Η μία συνιστώσα καταλήγει στο ελάχιστο ενώ η άλλη ταλαντώνεται για αρκετά βήματα μέχρις ότου καταλήξει και αυτή στο ελάχιστο.

Παρακάτω δίνονται και γραφικές που δείχνουν αυτά τα οποία αναφέρονται.





Αντίστοιχα και με το προηγούμενο θέμα η διαφορά με το θέμα 1 εδώ είναι ότι ο περιορισμός αφορά το s_k και όχι το γ_k.

Για να συγκλείνει πιο γρήγορα στο ελάχιστο μπορούμε να μειώσουμε το s_k ώστε να πληροί και τους δύο περιορισμούς και όχι μόνο τον έναν όπως γίνεται παραπάνω.

ΘΕΜΑ 4

$$f(x_1, x_2) = \frac{1}{3}{x_1}^2 + 3{x_2}^2$$

$$\nabla f(x_1, x_2) = \begin{bmatrix} {}^2/_3\, x_1 \\ 6x_2 \end{bmatrix}$$

$$x_{k+1} = x_k + \gamma_\kappa(\bar{\chi}_\kappa - \chi_\kappa)$$

$$\bar{\chi}_\kappa = \Pr\left[\chi_\kappa - s_k * \nabla f(\chi_\kappa)\,\right]$$

$$\bar{\chi}_\kappa = \Pr\begin{bmatrix} \chi_1 - s_k * {}^2/_3 * \chi_1 \\ \chi_2 - s_k * 6 * \chi_2 \end{bmatrix}$$ *έστω ότι το σημείο είναι η ίδια η προβολή*

(για να πάρουμε και τους δυο τους περιορισμούς)

$$\bar{\chi}_\kappa = \begin{bmatrix} \chi_1 - s_k * {}^2/_3 * \chi_1 \\ \chi_2 - s_k * 6 * \chi_2 \end{bmatrix}$$

$$x_{k+1} = x_k + \gamma_\kappa(\bar{\chi}_\kappa - \chi_\kappa) \Rightarrow$$

$$\begin{bmatrix} \chi_{1,\kappa+1} \\ \chi_{2,\kappa+1} \end{bmatrix} = \begin{bmatrix} \chi_{1,\kappa} \\ \chi_{2,\kappa} \end{bmatrix} + \gamma_\kappa(\chi_\kappa \begin{bmatrix} 1 - s_k * {}^2/_3 \\ 1 - s_k * 6 \end{bmatrix} - \chi_\kappa)$$

$$\begin{bmatrix} \chi_{1,\kappa+1} \\ \chi_{2,\kappa+1} \end{bmatrix} = \begin{bmatrix} \chi_{1,\kappa} \\ \chi_{2,\kappa} \end{bmatrix} + 0.2(\chi_\kappa \begin{bmatrix} 1 - s_k * {}^2/_3 \\ 1 - s_k * 6 \end{bmatrix} - \begin{bmatrix} \chi_{1,\kappa} \\ \chi_{2,\kappa} \end{bmatrix})$$

$$\begin{bmatrix} \chi_{1,\kappa+1} \\ \chi_{2,\kappa+1} \end{bmatrix} = \begin{bmatrix} \chi_{1,\kappa} \\ \chi_{2,\kappa} \end{bmatrix} + 0.2\chi_\kappa \begin{bmatrix} 1 - s_k * {}^2/_3 \\ 1 - s_k * 6 \end{bmatrix} - 0.2\begin{bmatrix} \chi_{1,\kappa} \\ \chi_{2,\kappa} \end{bmatrix}$$

$$\begin{bmatrix} \chi_{1,\kappa+1} \\ \chi_{2,\kappa+1} \end{bmatrix} = 0.8\begin{bmatrix} \chi_{1,\kappa} \\ \chi_{2,\kappa} \end{bmatrix} + 0.2\begin{bmatrix} \chi_{1,\kappa} \\ \chi_{2,\kappa} \end{bmatrix}\begin{bmatrix} 1 - s_k * {}^2/_3 \\ 1 - s_k * 6 \end{bmatrix}$$

$$\begin{cases} x_{1,k+1} = 0.8x_{1,k} + 0.2\chi_{1,\kappa}(1 - s_k * {}^2/_3) \\ x_{2,k+1} = 0.8x_{2,k} + 0.2\chi_{2,\kappa}(1 - s_k * 6) \end{cases} \Rightarrow$$

$$\begin{cases} x_{1,k+1} = \chi_{1,\kappa}(1 - s_k * {}^4/_{30}) \\ x_{2,k+1} = \chi_{2,\kappa}(1 - s_k * 1.2) \end{cases}$$

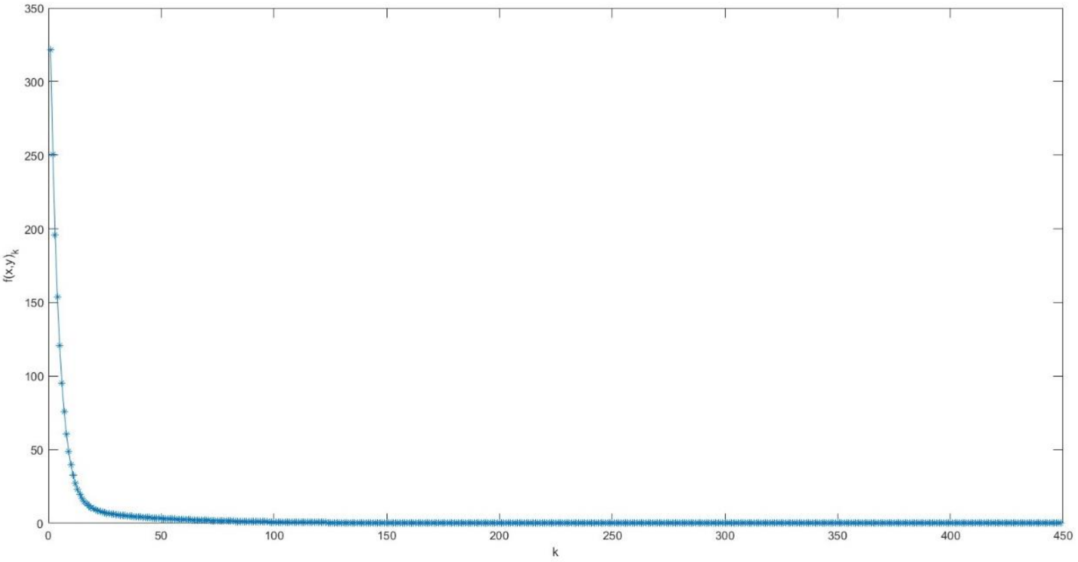
$$\begin{cases} \dfrac{|\chi_{1,\kappa+1}|}{|\chi_{1,\kappa}|} < 1 \\ \dfrac{|\chi_{2,\kappa+1}|}{|\chi_{2,\kappa}|} < 1 \end{cases} \Rightarrow \begin{cases} \dfrac{|\chi_{1,\kappa}(1 - s_k * {}^4/_{30})|}{|\chi_{1,\kappa}|} < 1 \\ \dfrac{|(1 - s_k * 1.2)|}{|\chi_{2,\kappa}|} < 1 \end{cases} \Rightarrow$$

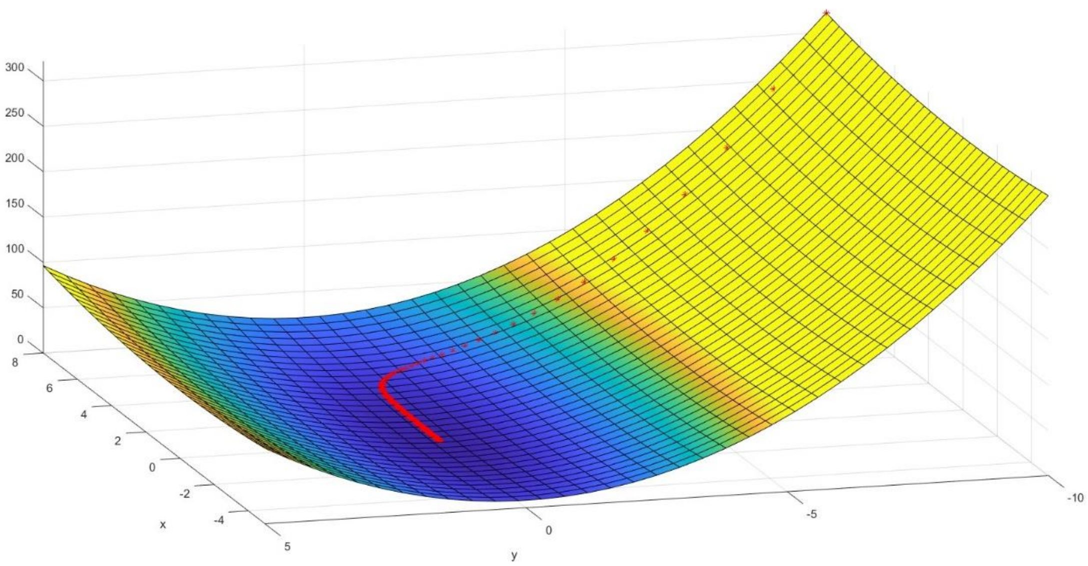
$$\begin{cases} |1 - s_k * {}^4/_{30}| < 1 \\ |1 - s_k * 1.2| < 1 \end{cases} \Rightarrow \begin{cases} 0 < s_k < 15 \\ 0 < s_k < 1.666 \end{cases} \Rightarrow$$

$$s_k \in (0, 1.666)$$

Στην περίπτωση αυτή περιμένουμε πως η μέθοδος θα συγκλίνει στο ελάχιστο καθώς το s_k είναι ίσο με 0.2 το οποίο ικανοποιεί και τους δύο περιορισμούς. Επίσης παρατηρούμε ότι η μέθοδος συγκλείνει αν και ξεκινάμε από μη επιτρεπτό σημείο (εκτός του πεδίου στο οποίο αναζητούμε το ελάχιστο). Αντίστοιχα με τις προηγούμενες δυο περιπτώσεις έχουμε περιορισμό για το s_k ο οποίος δίνεται από την παραπάνω μαθηματική ανάλυση.

Παρακάτω παρουσιάζονται διαγράμματα που αποδεικνύουν την σύγκληση του αλγορίθμου για τις δοσμένες συνθήκες.





Σε αύτη την περίπτωση βλέπουμε ότι φτάνουμε πολύ γρηγορότερα στο ελάχιστο καθώς πληρούνται οι προϋποθέσεις.