

# Τεχνικές Βελτιστοποίησης

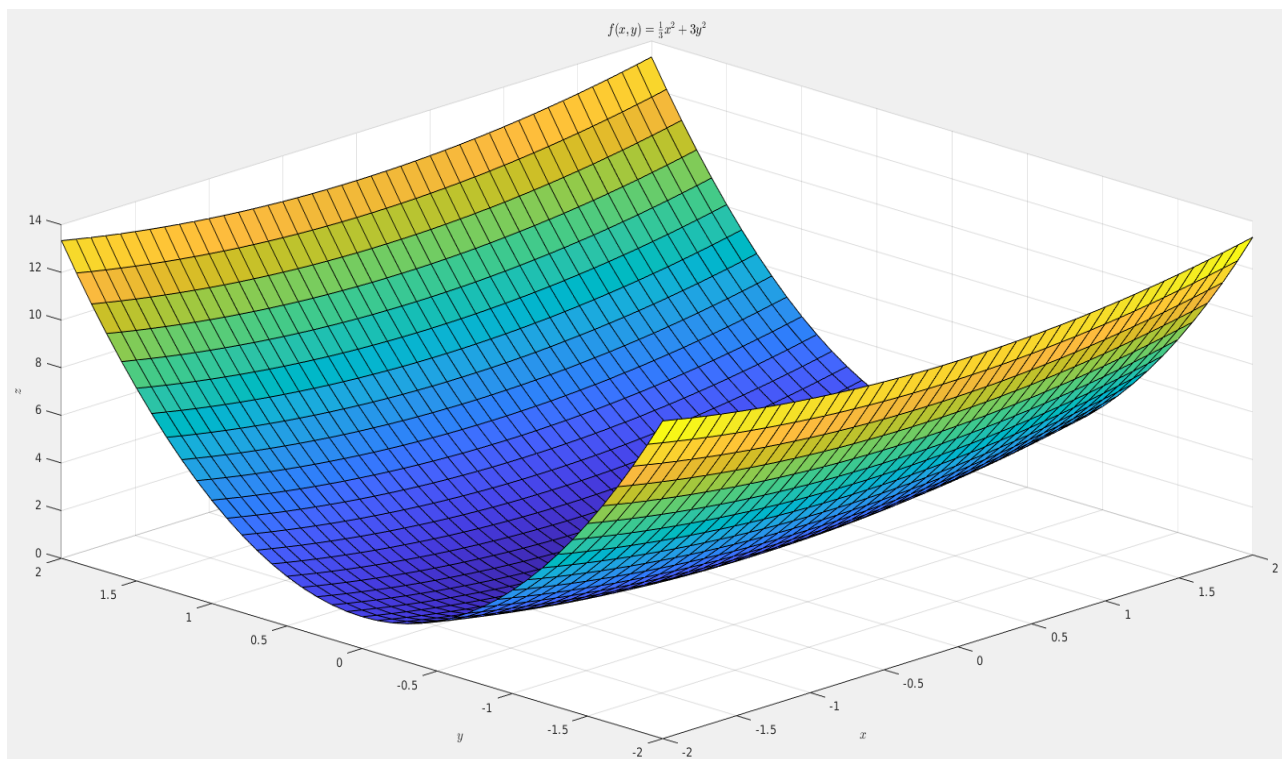
## 3ο παραδοτέο

Τουλκερίδης Νικόλαος

AEM: 10718

### Εισαγωγή

Η εργασία αυτή ασχολείται με την μελέτη της αντικειμενικής συνάρτησης δύο μεταβλητών:  
 $f(x, y) = (1/3)x^2 + 3y^2$ .



Συγκεκριμένα, εξετάζεται η ελαχιστοποίηση της συνάρτησης, αρχικά με χρήση της μεθόδου της μέγιστης καθόδου και απουσία περιορισμών και έπειτα με χρήση της μεθόδου μέγιστης καθόδου με προβολή στο σύνολο  $X : -10 \leq x \leq 5$  και  $-8 \leq y \leq 12$ .

## Θέμα 1 - Μέθοδος μέγιστης καθόδου

Ξεκινώντας την υλοποίηση της μεθόδου ορίζουμε την ακρίβεια - σταθερά τερματισμού  $e$  και το αρχικό σημείο  $(x, y)$ . Θέτουμε

$d = -\nabla f(x, y)$ , για όσο ισχύει η σχέση  $|\nabla f(x, y)| > e$ . Ορίζουμε την αναδρομική σχέση  $x_{k+1} = x_k + \gamma_k * d_k$ , όπου  $\gamma_k$  είναι σταθερά που επιλέγεται ανάμεσα από τις 4 τιμές:  $(0.1, 0.3, 3, 5)$ .

Η διαδικασία εύρεσης του επόμενου σημείου επαναλαμβάνεται ώσπου να ξεπεράσουμε τον μέγιστο αριθμό επαναλήψεων (1300, 2000) που ορίζεται από εμάς ώστε να αποφύγουμε τους ατέρμονες βρόγχους, εάν η μέθοδος δεν συγκλίνει.

Επιλέχθηκε ακρίβεια  $e$  ίση με 0.001 και το αρχικό σημείο  $(4, 4)$ . Για τις διαφορετικές τιμές του  $\gamma$  (0.1, 0.3, 3, 5) λαμβάνουμε διαφορετικές προσομοιώσεις. Παρατηρήθηκε ότι για μικρές τιμές του  $\gamma$ , δηλαδή για 0.1 και 0.3, ο αλγόριθμος φαίνεται να συγκλίνει στο ελάχιστο μετά από λίγες επαναλήψεις ενώ για τις μεγάλες τιμές του  $\gamma$  οι επαναλήψεις αυξάνονται σημαντικά χωρίς να συγκλίνει η μέθοδος.

Για να εξηγήσουμε αυτή τη συμπεριφορά υπολογίζουμε αρχικά το gradient της συνάρτησης

$$\nabla f(x, y) = \begin{bmatrix} (2/3)x \\ 6y \end{bmatrix}, \text{ και αντικαθιστώντας το στην αναδρομική σχέση έχουμε:}$$

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ y_k \end{bmatrix} - \gamma_k \begin{bmatrix} (2/3)x_k \\ 6y_k \end{bmatrix} \implies \begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix} = \begin{bmatrix} x_k(1 - (2/3)\gamma_k) \\ y_k(1 - 6\gamma_k) \end{bmatrix}$$

Οι τιμές του βήματος που εγγυώνται την ευστάθεια είναι αυτές που ικανοποιούν την συνθήκη:

$$|x_{k+1}/x_k| \leq 1 \implies |x_k(1 - (2/3)\gamma_k)/x_k| \leq 1 \implies |1 - (2/3)\gamma_k| \leq 1 \implies 0 \leq \gamma_k < 3$$

$$|y_{k+1}/y_k| \leq 1 \implies |y_k(1 - 6\gamma_k)/y_k| \leq 1 \implies |1 - 6\gamma_k| \leq 1 \implies 0 \leq \gamma_k < \frac{1}{3}$$

**Οι δύο ανισώσεις συναλθεύουν όταν  $0 \leq \gamma_k < \frac{1}{3}(1)$ .**

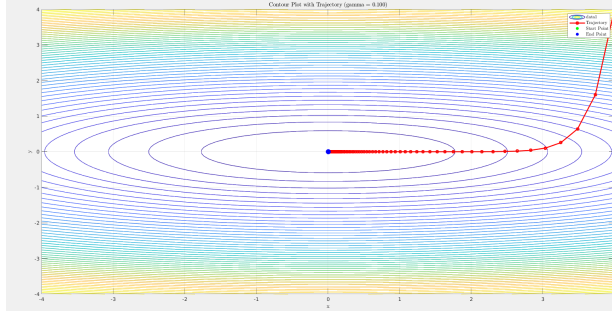
Επιπρόσθετα, μπορούμε να προσδιορίσουμε τους ρυθμούς σύγκλισης μέσω των αναδρομικών σχέσεων των  $x_{k+1}, y_{k+1}$ :  $\mathbf{x}_k = \mathbf{x}_0(1 - \frac{2}{3}\gamma_k)^k$  και  $\mathbf{y}_k = \mathbf{y}_0(1 - 6\gamma_k)^k$ .

Η ανισότητα (1) εξηγεί τα αποτελέσματα των προσομοιώσεων αφού φαίνεται πώς για επιλογή  $\gamma_k < \frac{1}{3}$  η μέθοδος συγκλίνει στο ελάχιστο ενώ για μεγαλύτερες τιμές η σύγκλιση δεν είναι εγγυημένη.

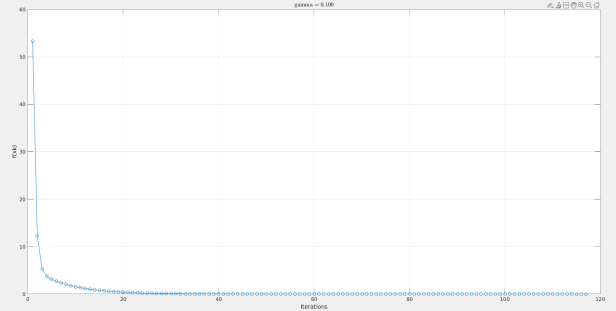
Παρακάτω παρατίθενται τα διαγράμματα της σύγκλισης της συνάρτησης για την εκάστοτε επιλογή του  $\gamma_k$ . Παρατηρούμε ότι για  $\gamma_k$  γύρω από την οριακή τιμή  $\frac{1}{3}$ , για - 0.001 η τροχιά των  $[x_k, y_k]$  δείχνει να ταλαντώνει αλλά τελικά συγκλίνει αργά στο ελάχιστο (με 1676 επαναλήψεις), ενώ για +0.001 η τροχιά ταλαντώνει χωρίς σύγκλιση και ο αλγόριθμος τερματίζει στις 2000 προκαθορισμένες επαναλήψεις.

$$\gamma = 0.1$$

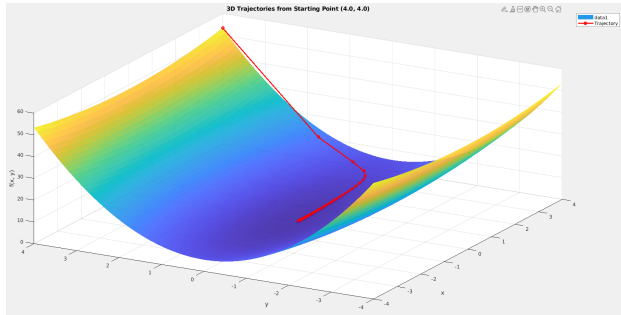
Συγκλίνει ο αλγόριθμος σε 115 επαναλήψεις. Ισχύει ότι  $0.1 < \frac{1}{3}$ , το οποίο επαληθεύει την σύγκλιση. Παρατηρούμε από το Σχήμα 4 ότι το  $x$  συγκλίνει πιο αργά από το  $y$ . Αυτό επαληθεύεται και από τους τύπους σύγκλισης αφού για  $\gamma = 0.1$  και  $(x_0, y_0) = (4, 4)$ , παίρνουμε:  $x_k = 4(0.933)^k$  και  $y_k = 4(0.4)^k$  (μικρότερη βάση, πιο απότομη σύγκλιση).



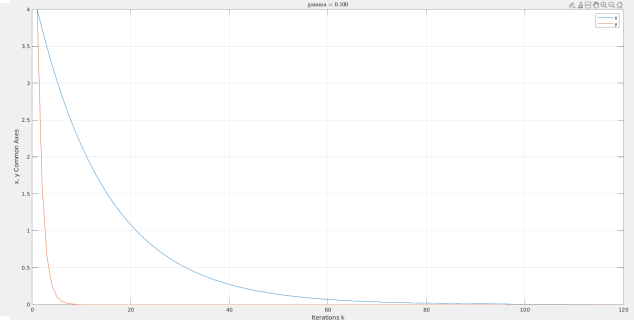
Σχήμα 1: Απεικόνιση σύγκλισης σε ισοβαρές καμπύλες



Σχήμα 2: Σύγκλιση αλγορίθμου



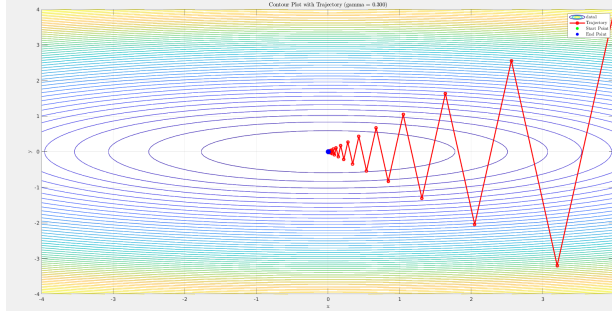
Σχήμα 3: Τρισδιάστατη Αναπαράσταση Σύγκλισης



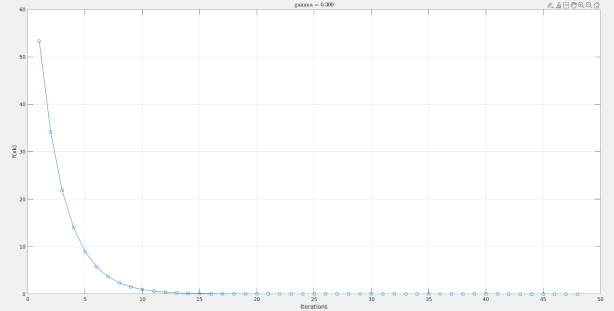
Σχήμα 4: Μεταβολή  $x, y$  για  $x_0=[4,4]$

$$\gamma = 0.3$$

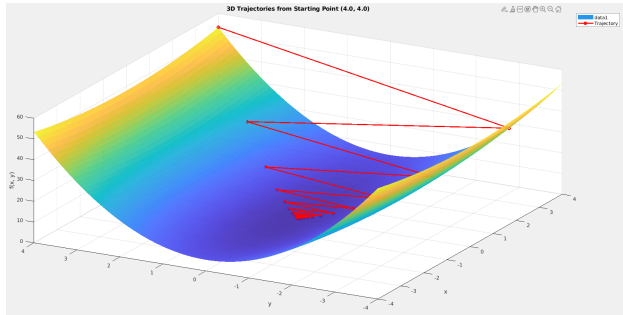
Συγκλίνει ο αλγόριθμος σε 46 επαναλήψεις. Ισχύει ότι  $0.3 < \frac{1}{3}$ , το οποίο επαληθεύει την σύγκλιση. Παρατηρούμε από το Σχήμα 4 ότι το  $x$  συγκλίνει με τον ίδιο ρυθμό που συγκλίνει το  $y$ , αν και το  $y$  ταλαντώνεται ανά επανάληψη. Αυτό επαληθεύεται και από τους τύπους σύγκλισης αφού για  $\gamma = 0.3$  και  $(x_0, y_0) = (4, 4)$ , παίρνουμε:  $x_k = 4(0.8)^k$  και  $y_k = 4(-0.8)^k$  (η εναλλαγή προσήμου στο  $y_k$  του προκαλεί την ταλάντωση).



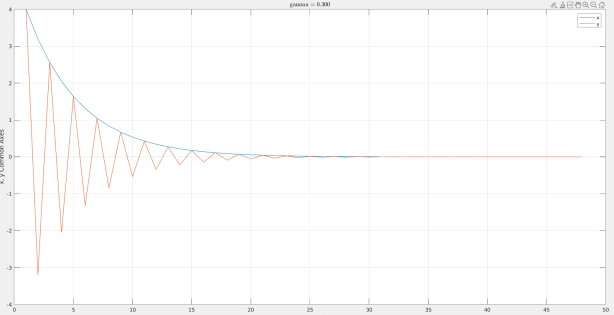
Σχήμα 5: Απεικόνιση σύγκλισης σε ισοβαρές καμπύλες



Σχήμα 6: Σύγκλιση αλγορίθμου



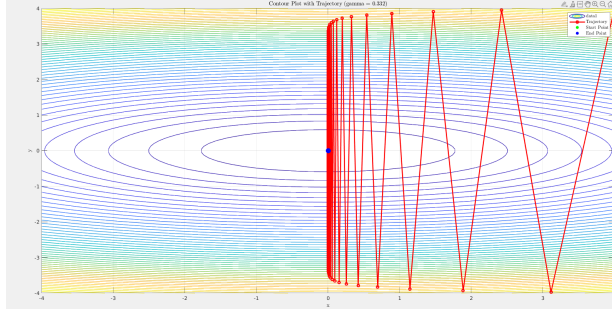
Σχήμα 7: Τρισδιάστατη Αναπαράσταση Σύγκλισης



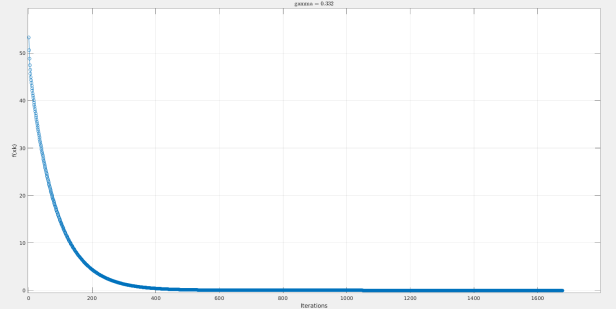
Σχήμα 8: Μεταβολή  $x, y$  για  $x_0=[4,4]$

$$\gamma = 0.333 - 0.001$$

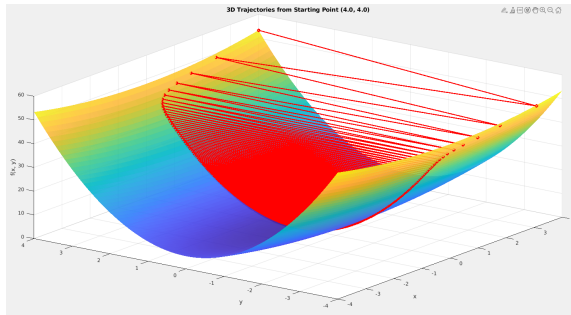
Συγκλίνει ο αλγόριθμος σε 1676 επαναλήψεις. Ισχύει ότι  $\frac{1}{3} - 0.001 < \frac{1}{3}$ , το οποίο επαληθεύει την σύγκλιση. Παρατηρούμε από το Σχήμα 4 ότι το  $x$  συγκλίνει πλέον πιο γρήγορα από το  $y$ , και το  $y$  ταλαντώνεται ανά επανάληψη για πολλές επαναλήψεις. Αυτό επαληθεύεται και από τους τύπους σύγκλισης αφού για  $\gamma = 0.333 - 0.001$  και  $(x_0, y_0) = (4, 4)$ , παίρνουμε:  $x_k = 4(0.7784)^k$  και  $y_k = 4(-0.994)^k$ . Η εναλλαγή προσήμου στο  $y_k$  του προκαλεί την ταλάντωση. Επιπλέον, επειδή ο ρυθμός σύγκλισης του  $y_k$  τείνει στο 1, παρατηρείται δυσκολία σύγκλισης.



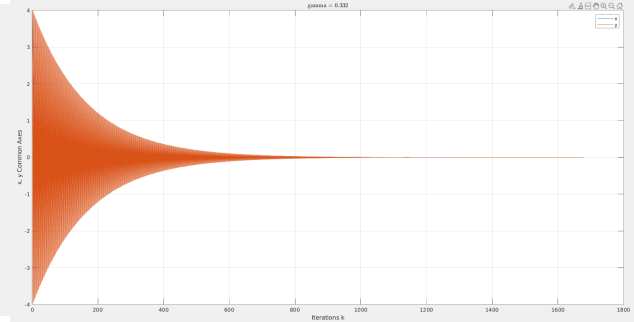
Σχήμα 9: Απεικόνιση σύγκλισης σε ισοβαρές καμπύλες



Σχήμα 10: Σύγκλιση αλγορίθμου



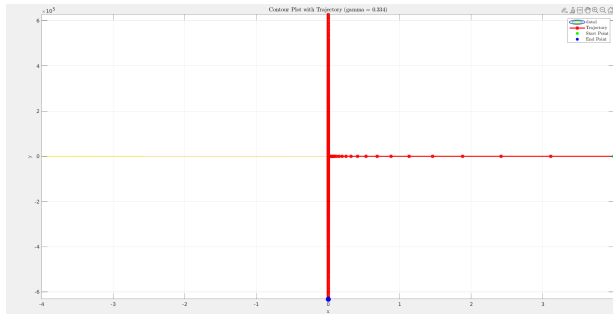
Σχήμα 11: Τρισδιάστατη Αναπαράσταση Σύγκλισης



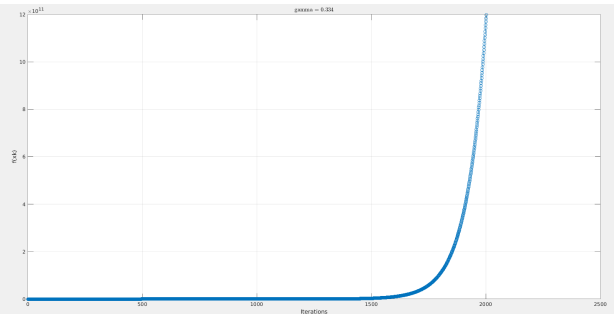
Σχήμα 12: Μεταβολή  $x, y$  για  $x_0=[4,4]$

$$\gamma = 0.333+0.001$$

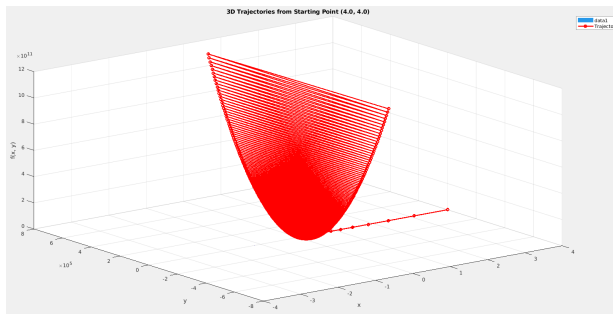
ΔΕΝ συγκλίνει ο αλγόριθμος για 2000 επανλήψεις (maximum numbers of iterations). Ισχύει ότι  $\frac{1}{3}+0.001 > \frac{1}{3}$ , το οποίο επαληθεύει την απόκλιση. Παρατηρούμε από το Σχήμα 4 ότι το  $x$  συγκλίνει ενώ το  $y$  αποκλίνει ταλαντώνοντας ανά επανάληψη. Αυτό επαληθεύεται και από τους τύπους σύγκλισης αφού για  $\gamma = 0.333+0.001$  και  $(x_0, y_0) = (4, 4)$ , παίρνουμε:  $x_k = 4(0.7771)^k$  και  $y_k = 4(-1.006)^k$ . Ο ρυθμός σύγκλισης είναι μεγαλύτερος της μονάδας, οπότε ανά επανάληψη το  $y_k$  θα αυξάνεται αντί να μειώνεται. Η απόκλιση θα έχει αργό ρυθμό.



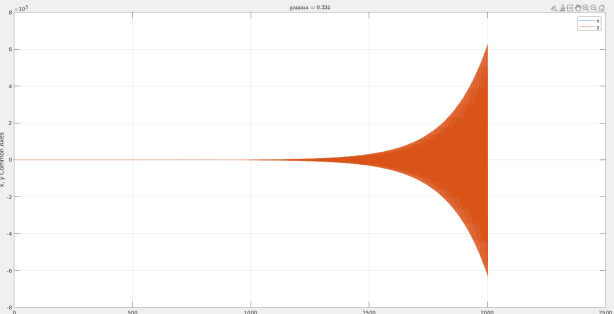
Σχήμα 13: Απεικόνιση σύγκλισης σε ισοβαρές καμπύλες



Σχήμα 14: Σύγκλιση αλγορίθμου



Σχήμα 15: Τρισδιάστατη Αναπαράσταση Σύγκλισης

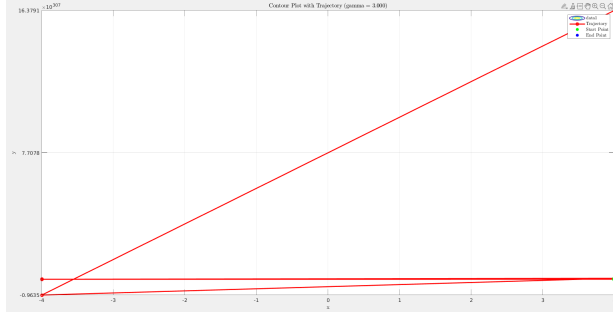


Σχήμα 16: Μεταβολή  $x, y$  για  $x_0=[4,4]$

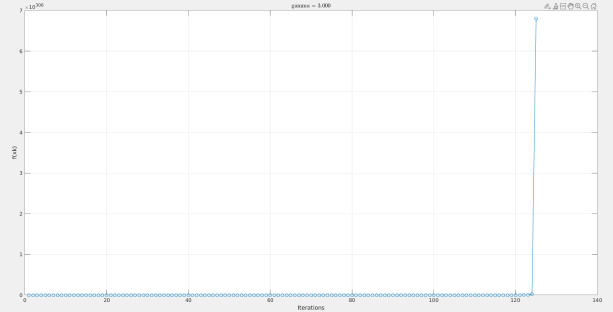
$$\gamma = 3$$

Αποκλίνει ακραία ο αλγόριθμος για 252 επαναλήψεις (calculation memory outflows). Ισχύει ότι  $3 > \frac{1}{3}$ , το οποίο επαληθεύει την απόκλιση. Όσο αυξάνουμε το  $\gamma_k$ , τόσο επιταγχύνουμε την απόκλιση.

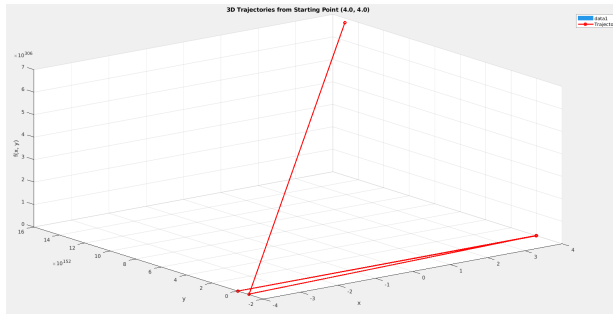
Από τους ρυθμούς σύγκλισης, βλέπουμε ότι το  $x$  παραμένει στάσιμο και ταλαντώνεται μεταξύ των 2 τιμών  $-x_0, +x_0$ , ενώ το  $y$  αποκλίνει απότομα, ταλαντώνοντας ανά επανάληψη. Για  $\gamma = 3$  και  $(x_0, y_0) = (4, 4)$ , παίρνουμε:  $x_k = 4(-1)^k$  και  $y_k = 4(-17)^k$ .



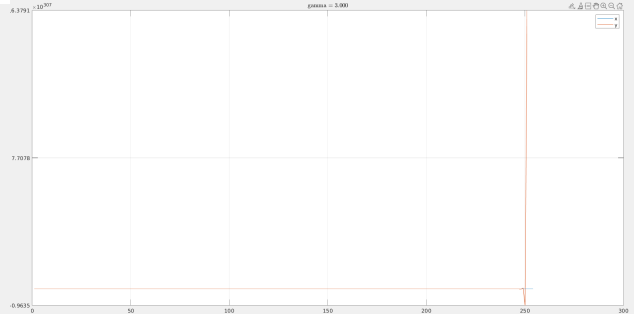
Σχήμα 17: Απεικόνιση σύγκλισης σε ισοβαρές καμπύλες



Σχήμα 18: Σύγκλιση αλγορίθμου



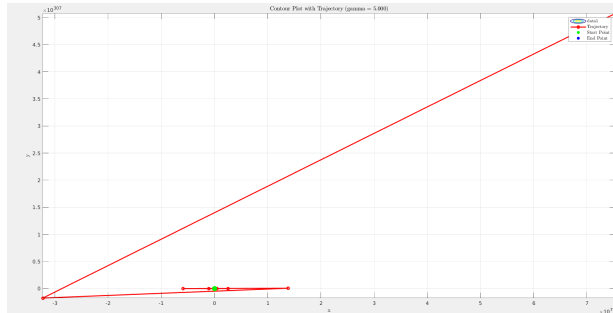
Σχήμα 19: Τρισδιάστατη Αναπαράσταση Σύγκλισης



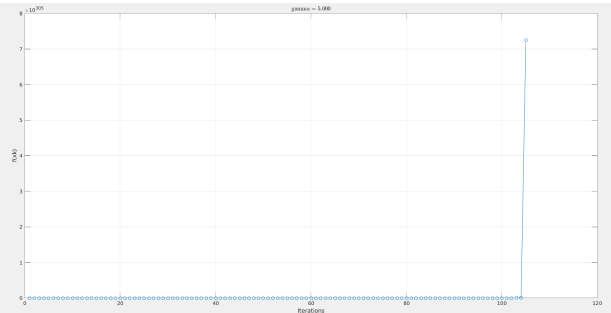
Σχήμα 20: Μεταβολή  $x, y$  για  $x_0=[4,4]$

$$\gamma = 5$$

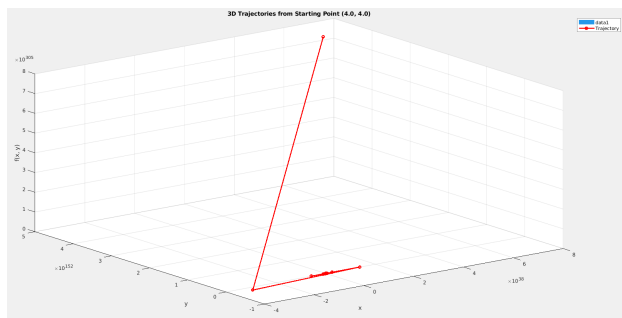
Αποκλίνει ακραία ο αλγόριθμος για 212 επαναλήψεις. Ισχύει ότι  $5 > \frac{1}{3}$ , το οποίο επαληθεύει την απόκλιση. Αυξήσαμε κι'άλλο το  $\gamma_k$ , και επιτεύχθηκε πιο γρήγορη απόκλιση. Από τους ρυθμούς σύγκλισης, βλέπουμε ότι και το  $x$  και το  $y_k$  αποκλίνουν ταλαντώνοντας. Για  $\gamma = 5$  και  $(x_0, y_0) = (4, 4)$ , παίρνουμε:  $x_k = 4(-2.333)^k$  και  $y_k = 4(-29)^k$ .



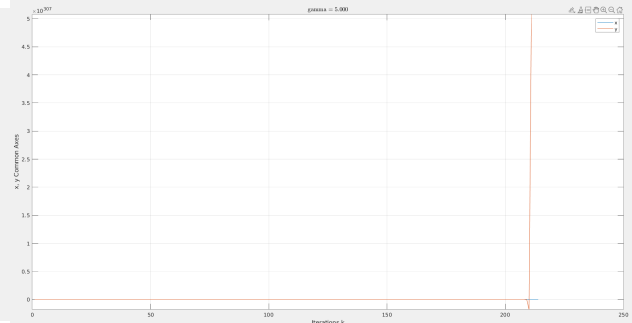
Σχήμα 21: Απεικόνιση σύγκλισης σε ισοβαρές καμπύλες



Σχήμα 22: Σύγκλιση αλγορίθμου



Σχήμα 23: Τρισδιάστατη Αναπαράσταση Σύγκλισης



Σχήμα 24: Μεταβολή  $x, y$  για  $x_0=[4,4]$



## Μέθοδος Μέγιστης Καθόδου με Προβολή

Με στόχο τη σύγκλιση του αλγορίθμου Μεγ. Καθόδου στο ολικό ελάχιστο, χρησιμοποιούμε τη μέθοδο Μέγιστης Καθόδου με Προβολή. Το σύνολο

$$-10 < x < 5$$

$$-8 < y < 12$$

είναι κυρτό οπότε μπορούμε να εφαρμόσουμε το Θεώρημα. Ο αλγόριθμος λειτουργεί ως εξής:

Ξεκινώντας την υλοποίηση της μεθόδου αυτής ορίζουμε πάλι την ακρίβεια - σταθερά τερματισμού  $e$  και το αρχικό σημείο  $(x, y)$ . Για όσο ισχύει η σχέση  $|\nabla f(x, y)| > e$ , υπολογίζεται το νέο σημείο σύμφωνα με την:

$$x_{k+1} = x_k + \gamma_k(x'_k - x_k), \text{ όπου } x'_k, \text{ η προβολή του } x_k \text{ στο κυρτό σύνολο } X \text{ και δίνεται από τη σχέση:}$$

$$x'_k = \text{Pr}(x_k - s_k \nabla f(x_k)), s_k \geq 0$$

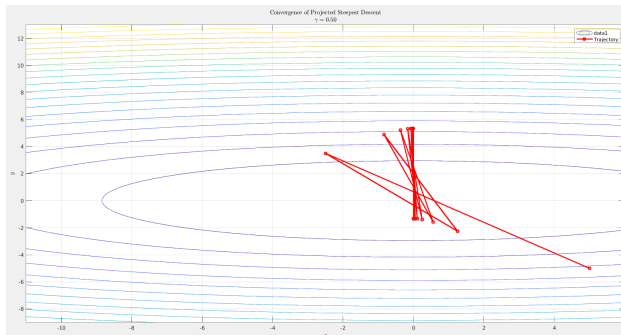
Στην περίπτωση που έχουμε προβολή εντός του κυρτού  $X$  και για το  $x_k$  και για το  $y_k$ , θα έχουμε για την  $k+1$  επανάληψη:

$$x_{k+1} = x_k(1 - \frac{2}{3}s_k\gamma_k)$$

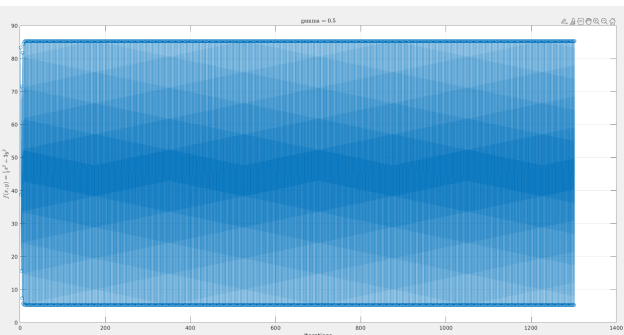
$$y_{k+1} = y_k(1 - 6s_k\gamma_k).$$

## Θέμα 2

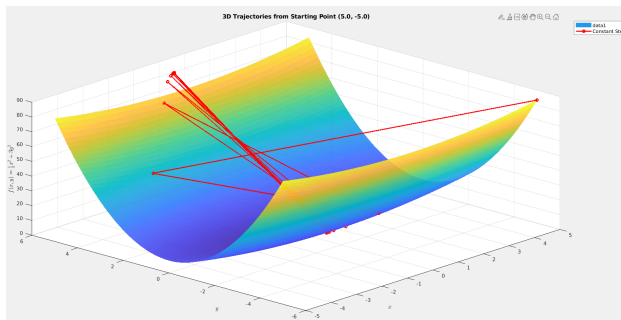
Θέτουμε το  $s_k = 5$  και το  $\gamma_k = 0.5$ . Παίρνουμε ως αρχικό σημείο το  $(5, -5)$  και όπως φαίνεται από τις προσομοιώσεις, η  $f(x, y)$  ταλαντώνεται χωρίς να συγκλίνει. Αφού φτάσει τον προκαθορισμένο αριθμό των 1300 επαναλήψεων, ο αλγόριθμος τερματίζει και αποτυγχάνει στον εντοπισμό του ελαχίστου.



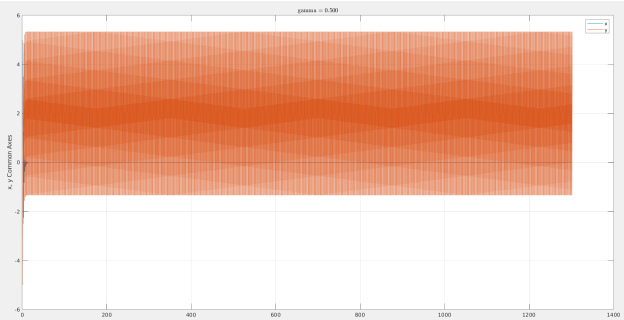
Σχήμα 25: Απεικόνιση σύγκλισης σε ισοβαρές καμπύλες



Σχήμα 26: Σύγκλιση αλγορίθμου

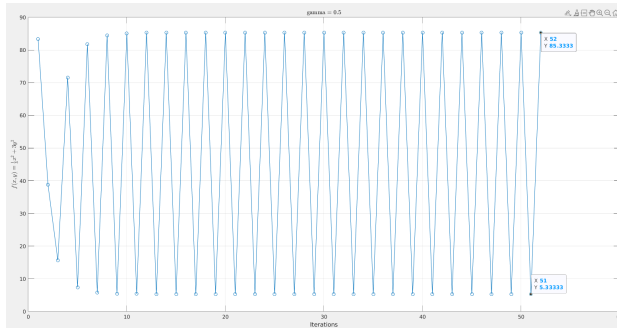


Σχήμα 27: Τρισδιάστατη Αναπαράσταση Σύγκλισης

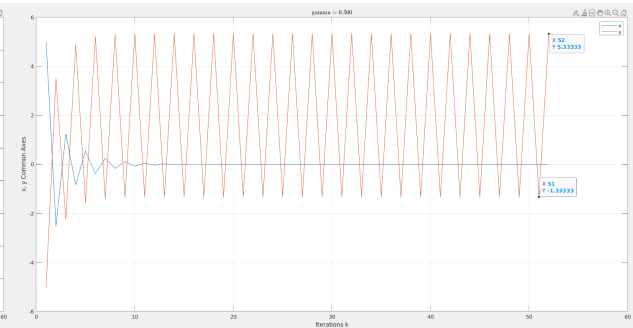


Σχήμα 28: Μεταβολή  $x, y$  για  $x_0=[5,-5]$

Πιο αναλυτικά, στο σχήμα 29, ήδη από τις πρώτες 10 επαναλήψεις η ταλάντωση της  $f$  καθιερώνεται ανάμεσα στις  $f(x, y) = 5.333$  και  $f(x, y) = 85.333$ . Παρατηρώντας και το σχήμα 30, καταλαβαίνουμε ότι το  $x_k$  συγκλίνει σε 10 επαναλήψεις ενώ το  $y_k$  ξεκινάει και ταλαντώνεται μεταξύ των τιμών  $-1.333$  και  $5.333$ .



Σχήμα 29: Σύγκλιση  $f(x, y)$



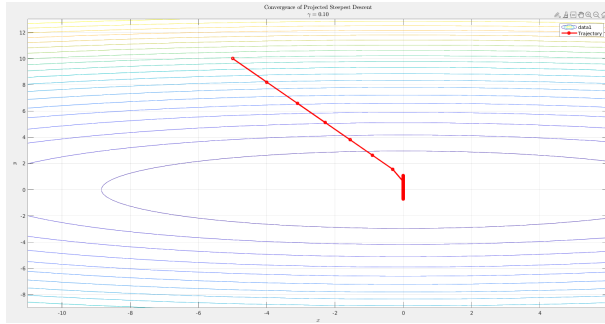
Σχήμα 30: Μεταβολή  $x, y$  για  $x_0=[5,-5]$

Είναι ξεκάθαρο ότι, σε αντίθεση με το θέμα 1, τα  $x_k, y_k$  **δεν αποκλίνουν ποτέ**. Δεν απομακρύνονται ποτέ από το κυρτό σύνολο  $X$ , κάτι που είναι εφικτό εξαιτίας της προβολής των  $x_k, y_k$  στο σύνολο αυτό, σε περίπτωση που κάποιο από τα δύο πέφτει εκτός αυτού στην εκάστοτε επανάληψη. Δηλαδή, το κάθε νέο σημείο που προκύπτει θα ανήκει στο κυρτό σύνολο και έτσι θα είναι εφικτό.

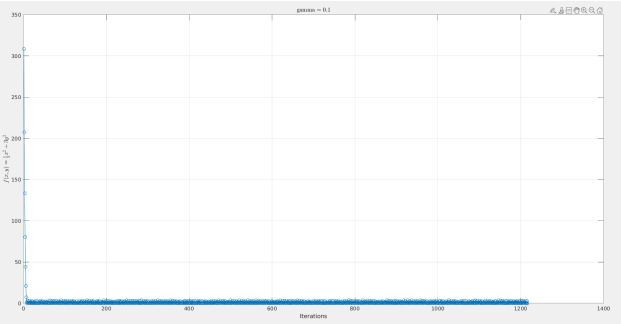
Η χρήση της προβολής, ωστόσο, δεν υπόσχεται τη σύγκλιση του αλγορίθμου, όπως βλέπουμε.

### Θέμα 3

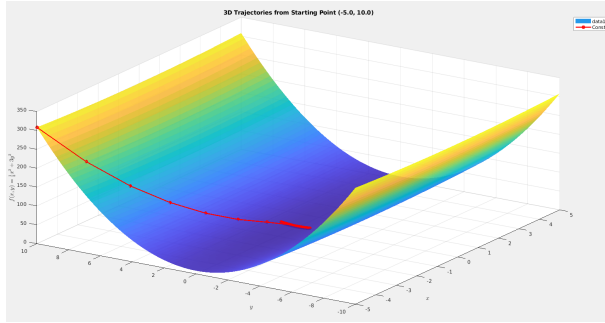
Θέτουμε το  $s_k = 15$  και το  $\gamma_k = 0.1$ . Παίρνουμε ως αρχικό σημείο το  $(-5, 10)$  και όπως φαίνεται από τις προσομοιώσεις, η  $f(x, y)$  ταλαντώνεται αλλά τελικά συγκλίνει στο ολικό ελάχιστο, μετά από 1215 επαναλήψεις. Το  $x$  συγκλίνει σε μόλις 8 επαναλήψεις όπως φαίνεται και στο σχήμα 35. Αυτό δικαιολογείται από το γεγονός ότι ο συνδυασμός  $s_k, \gamma_k$  είναι ιδανικός καθώς όταν το  $\bar{x}_k$  βρεθεί εντός του κυρτού συνόλου  $X$ , θα έχουμε:  $x_{k+1} = x_k(1 - \frac{2}{3}s_k\gamma_k)$ , όπου  $s_k\gamma_k = 1.5$ , άρα  $x_{k+1} = 0$ . Ακαριαία σύγκλιση στο 0 δηλαδή. Από την άλλη, το  $y$  είναι αυτό που καθυστερεί τη σύγκλιση του αλγορίθμου μας, και καταλήγουμε στις 1215 επαναλήψεις.



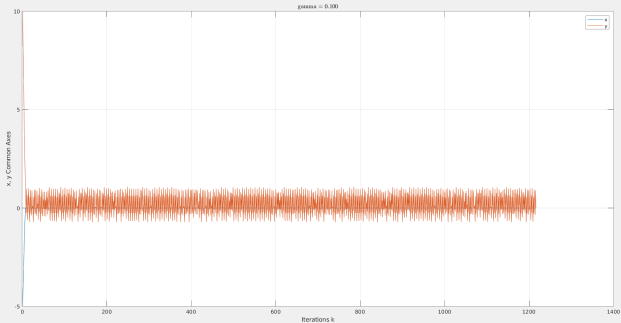
Σχήμα 31: Απεικόνιση σύγκλισης σε ισοβαρές καμπύλες



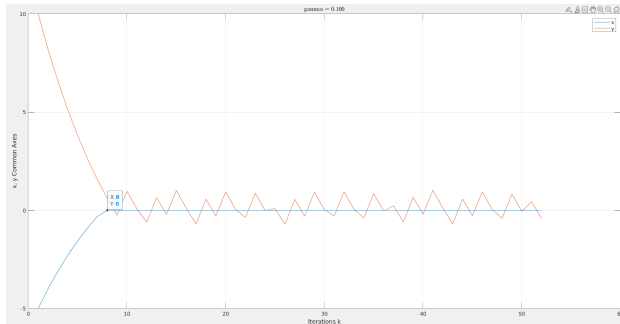
Σχήμα 32: Σύγκλιση αλγορίθμου



Σχήμα 33: Τρισδιάστατη Αναπαράσταση Σύγκλισης

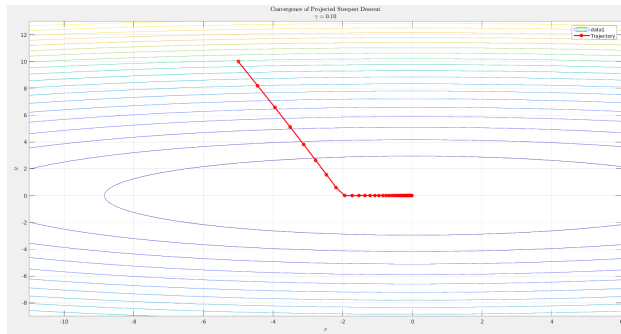


Σχήμα 34: Μεταβολή  $x, y$  για  $x_0 = [-5, 10]$

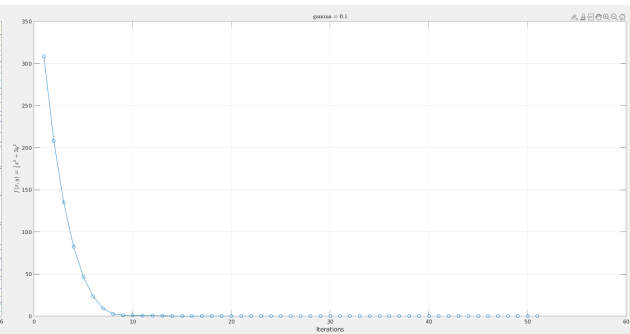


Σχήμα 35: 50 πρώτες επαναλήψεις: Μεταβολή  $x, y$  για  $x_0 = [-5, 10]$

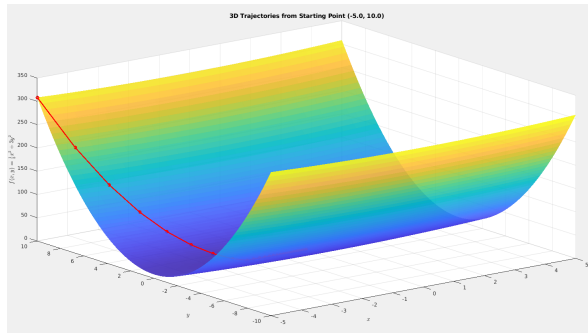
Για να αποφύγουμε τον μεγάλο αυτό αριθμό επαναλήψεων, μπορούμε να εφαρμόσουμε παρόμοια τεχνική για το  $y$ . Γνωρίζουμε ότι όταν το  $\bar{y}_k$  βρεθεί εντός συνόλου  $X$ , ισχύει:  $y_{k+1} = y_k(1 - 6s_k\gamma_k)$ . Κρατώντας  $\gamma_k = 0.1 = \text{const.}$  και θέτοντας  $s_k = 10/6$ , προκύπτουν τα παρακάτω αποτελέσματα:



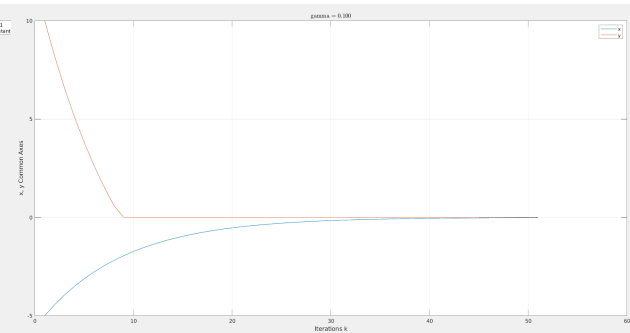
Σχήμα 36: Απεικόνιση σύγκλισης σε ισοβαρές καμπύλες



Σχήμα 37: Σύγκλιση αλγορίθμου



Σχήμα 38: Τρισδιάστατη Αναπαράσταση Σύγκλισης



Σχήμα 39: Μεταβολή  $x, y$  για  $x_0 = [-5, 10]$

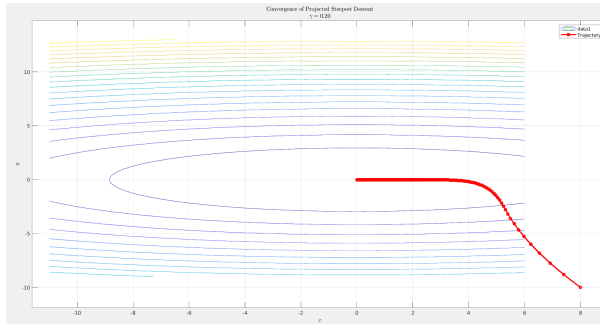
Ο συνολικός αριθμός επαναλήψεων που προκύπτει είναι  $k=50$ , μικρότερος από πριν. Το  $y$  είχε ήδη συγκλίνει από την 9η επανάληψη. Όπως παρατηρούμε με αυτόν τον τρόπο το  $y$  συγκλίνει γρήγορα στο 0 καθώς όταν το  $\bar{y}_k$  βρεθεί στο κυρτό σύνολο, θα χρειαστεί μια επανάληψη ώστε το  $y_{k+1}$  να μηδενιστεί.

## Θέμα 4

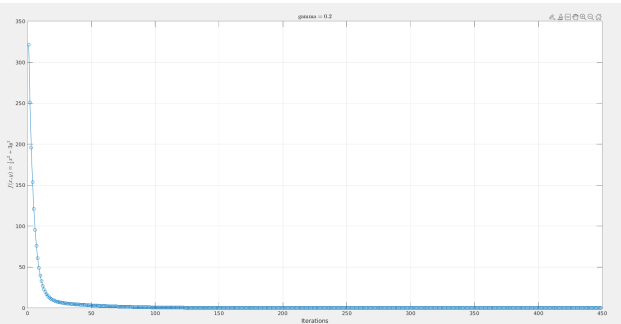
Θέτουμε τώρα το  $s_k = 0.1$  και το  $\gamma_k = 0.2$ . Παίρνουμε ως αρχικό σημείο το  $(8, -10)$ . Εκ των προτέρων μπορούμε να προβλέψουμε την επιτυχία σύγκλισης αλλά και τον σχετικά μεγάλο αριθμό επαναλήψεων, λόγω των μικρών τιμών  $s_k, \gamma_k$ .

Πιο συγκεκριμένα, όταν τα  $\bar{x}_k$  και  $\bar{y}_k$  βρεθούν εντός του κυρτού συνόλου  $X$ , θα ισχύουν οι γνωστοί τύποι:  $x_{k+1} = x_k(1 - \frac{2}{3}s_k\gamma_k)$  και  $y_{k+1} = y_k(1 - 6s_k\gamma_k)$ . Με αντικατάσταση των  $s_k, \gamma_k$ , παίρνουμε:  $x_{k+1} = x_k(0.9866)$  και  $y_{k+1} = y_k(0.88)$ . Συνεπώς, από την επανάληψη αυτή και έπειτα, θα έχουμε δεδομένη σύγκλιση στο ολικό ελάχιστο  $(0,0)$ . Ακόμα, στο σχήμα 43 φαίνεται η ταχύτερη σύγκλιση του  $y$ , που επιβεβαιώνει από το γεγονός ότι έχει μικρότερη βάση στον ρυθμό σύγκλισής του ( $0.88 < 0.9866$ ).

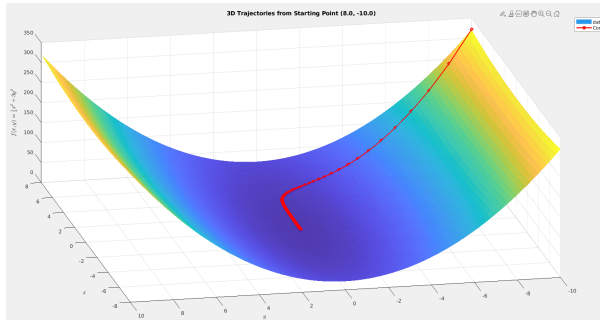
Όπως φαίνεται από τις προσομοιώσεις, η  $f(x, y)$  συγκλίνει στο ολικό ελάχιστο, μετά από 448 επαναλήψεις.



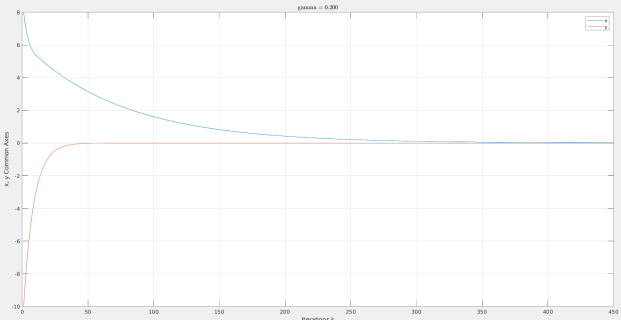
Σχήμα 40: Απεικόνιση σύγκλισης σε ισοβαρές καμπύλες



Σχήμα 41: Σύγκλιση αλγορίθμου



Σχήμα 42: Τρισδιάστατη Αναπαράσταση Σύγκλισης



Σχήμα 43: Μεταβολή  $x, y$  για  $x_0=[8,-10]$