<u>ΕΡΓΑΣΙΑ 3</u>

Μέθοδος Μέγιστης Καθόδου με Προβολή

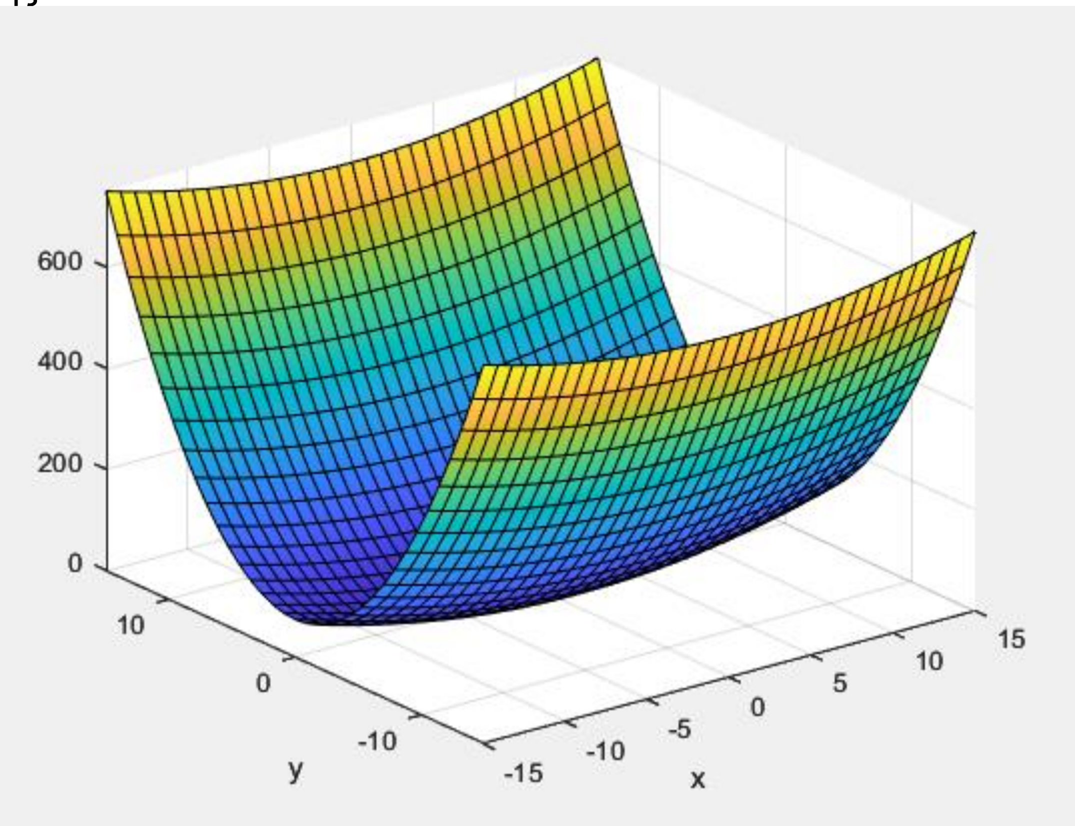
Ονοματεπώνυμο: Παπούλιας Μιχαήλ

ΑΕΜ: 10204

email: mpapouli@ece.auth.gr

Θέμα 1

Αρχικά, σχεδιάζουμε τη γραφική παράσταση της συνάρτησης $f(x,y) = \frac{1}{3}x^2 + 3y^2$, ώστε να έχουμε μια γενική εικόνα της μορφής της:



Από το γράφημα της f, φαίνεται ότι το ολικό ελάχιστό της (και το μοναδικό αυτής) είναι η αρχή των αξόνων (0, 0).

Εφαρμόζουμε τον αλγόριθμο της Μέγιστης Καθόδου με ακρίβεια e=0.001, σημείο εκκίνησης το [2, 2] και i) $\gamma_k = 0.1$, ii) $\gamma_k = 0.3$, iii) $\gamma_k = 3$ και iv) γ_k=5, και λαμβάνουμε τα εξής αποτελέσματα:

i) $\gamma_k = 0.1$:

Σημείο ελαχίστου το [0.00142853, 3.29100911e-42] k=106 επαναλήψεις

ii) $\gamma_k = 0.3$:

Σημείο ελαχίστου το [0.00013611, -0.00013611]] k=44 επαναλήψεις

iii) $\gamma_k = 3$:

Ο αλγόριθμος αποκλίνει για το συγκεκριμένο γ_k

iv) γ_k=5:

Ο αλγόριθμος αποκλίνει για το συγκεκριμένο γ_k

Τα παραπάνω αποτελέσματα μπορούν εύκολα να αποδειχθούν μαθηματικά. Θεωρώντας ότι $x_k = [x, y]$ και λαμβάνοντας υπόψιν τη μέθοδο της Μέγιστης Καθόδου, θα βρούμε εκείνα τα γ_k τα οποία οδηγούν σε σύγκλιση του αλγορίθμου. Έχουμε, λοιπόν, διαδοχικά:

$$x_{k+1} = x_k - \gamma_k \cdot \nabla_f(x_k) \quad \text{και} \quad \nabla_f(x_k) = \begin{bmatrix} \frac{2}{3}x \\ 6y \end{bmatrix} =>$$

$$=> \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} - \gamma_k \cdot \begin{bmatrix} \frac{2}{3}x \\ 6y \end{bmatrix} =>$$

$$x_{k+1} = x_k \cdot (1 - \tfrac{2}{3}\gamma_k)$$
$$y_{k+1} = y_k \cdot (1 - 6\gamma_k)$$

Γνωρίζουμε ότι για να συγκλίνει η μέθοδος στο ελάχιστο, θα πρέπει: $\frac{|x_{k+1}|}{|x_k|} < 1$ και $\frac{|y_{k+1}|}{|y_k|} < 1$, επομένως:

$$\left|1 - \tfrac{2}{3}\gamma_k\right| < 1 \;=>\; -1 < 1 - \tfrac{2}{3}\gamma_k < 1 \;=>\; \gamma_k > 0 \;\; \kappa\alpha\iota \;\; \gamma_k < 3$$

και

$$|1 - 6\gamma_k| < 1 \;=>\; -1 < 1 - 6\gamma_k < 1 \;=>\; \gamma_k > 0 \;\; \kappa\alpha\iota \;\; \gamma_k < \tfrac{1}{3}$$

Από τα παραπάνω, φαίνεται ότι το γ_k θα πρέπει να είναι μεγαλύτερο του μηδενός (προφανώς, αφού αποτελεί προϋπόθεση) και μικρότερο του $1/3$, ώστε η μέθοδος να συγκλίνει στο ελάχιστο. Γι' αυτό και στις δύο πρώτες περιπτώσεις η μέθοδος πράγματι συγκλίνει στο ελάχιστο, ενώ στις άλλες δύο αποκλίνει.

Θέμα 2

Κατασκευάζουμε τη μέθοδο της Μέγιστης Καθόδου με Προβολή. Ο αλγόριθμος αυτός προβάλλει το αρχικό σημείο εκκίνησης Xo στο σύνολο *Π* που ορίζεται ως:

$$\Pi = \begin{cases} -10 \leq x \leq 5 \\ -8 \leq y \leq 12 \end{cases}$$

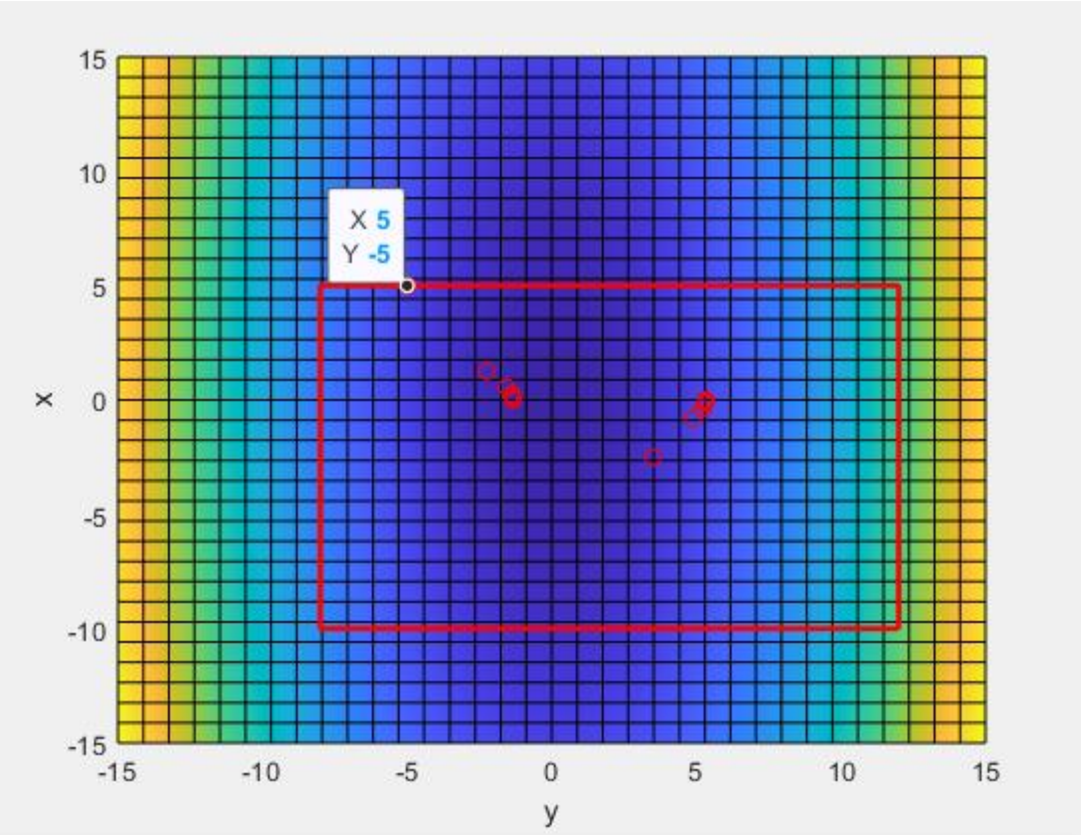
καθώς και κάθε διάνυσμα:

$$x_k + s_k \cdot d_k$$

στο ίδιο σύνολο περιορισμών, *Π*.

Θέτουμε $s_k = 5$, γ_k = 0.5 και αρχικό σημείο εκκίνησης Xo = [5, -5] και τρέχουμε τον αλγόριθμο. Το αποτέλεσμα δείχνει ότι αυτές οι αρχικές τιμές οδηγούν σε απόκλιση του αλγορίθμου από το ολικό ελάχιστο. Πράγματι το ελάχιστο, στα k=600 βήματα, εμφανίζεται ότι θα είναι το διάνυσμα [-9.34175249e-106, 5.33333333]. Η πρώτη συνιστώσα του διανύσματος συγκλίνει πράγματι στο 0, όμως δε συμβαίνει το ίδιο και για τη δεύτερη συνιστώσα. Αυτό εξηγείται παρακάτω με μαθηματικό τρόπο.

Εμφανίζουμε το γράφημα της σύγκλισης των διαδοχικών σημείων:



Φαίνεται ότι δεν υπάρχει σύγκλιση στο ελάχιστο [0, 0], αλλά μία διαρκής ταλάντωση γύρω από την αρχή των αξόνων, όσο το k αυξάνεται. Φυσικά, τα διαδοχικά σημεία θα είναι άπειρα, αλλά λόγω ανάγκης τερματισμού του αλγορίθμου για εξαγωγή συμπεράσματος, περιορίσαμε την τιμή αυτή του k στις 600 επαναλήψεις. Συνεπώς, τα σημεία στο σχήμα στη πραγματικότητα είναι 600 στο σύνολο, όσο κι αν αυτό δε φαίνεται εξαρχής, αφού μετά από ένα μικρό αριθμό επαναλήψεων τα σημεία 'πέφτουν' το ένα πάνω στο άλλο. Το κόκκινο ορθογώνιο αποτελεί το σύνολο *Π* των περιορισμών, μέσα στο οποίο πρέπει να βρίσκονται όλα τα σημεία x_k.

Το γεγονός ότι ο αλγόριθμος αποκλίνει για τις συγκεκριμένες αρχικές τιμές δε θα πρέπει να μας εκπλήσσει. Θα βρούμε εκείνα τα s_k, τα οποία οδηγούν σε σύγκλιση του αλγορίθμου. Έχουμε, λοιπόν, για $\gamma_k = 0.5$:

$$x_{k+1} = x_k + \gamma_k \cdot (\bar{x}_k - x_k)$$

με $\bar{x}_k = P_{r_\Pi}(x_k - s_k \cdot \nabla_f(x_k)) \;=>\; \bar{x}_k = P_{r_\Pi}(\begin{bmatrix} x \\ y \end{bmatrix} - s \cdot \begin{bmatrix} \frac{2}{3}x \\ 6y \end{bmatrix}) \;=>$

$$=> \bar{x}_k = P_{r_\Pi}(\begin{bmatrix} x(1-\frac{2}{3}s) \\ y(1-6s) \end{bmatrix}) \;=>\; \bar{x}_k = \begin{bmatrix} x(1-\frac{2}{3}s) \\ y(1-6s) \end{bmatrix}$$

Επομένως, έχουμε:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0.5x \\ 0.5y \end{bmatrix} + \begin{bmatrix} 0.5x(1-\frac{2}{3}s) \\ 0.5y(1-6s) \end{bmatrix} \;\; =>$$

$$x_{k+1} = x_k \cdot (1 - \tfrac{1}{3}s)$$
$$y_{k+1} = y_k \cdot (1 - 3s)$$

Γνωρίζουμε ότι για να συγκλίνει η μέθοδος στο ελάχιστο, θα πρέπει: $\frac{|x_{k+1}|}{|x_k|} < 1$ και $\frac{|y_{k+1}|}{|y_k|} < 1$, επομένως:

$$\left|1-\frac{1}{3}s\right| < 1 \;\; => \;\; -1 < 1-\frac{1}{3}s < 1 \;\; => \;\; s > 0 \;\; \kappa\alpha\iota \;\; s < 6$$

και

$$|1-3s| < 1 \;\; => \;\; -1 < 1-3s < 1 \;\; => \;\; s > 0 \;\; \kappa\alpha\iota \;\; s < \frac{2}{3}$$

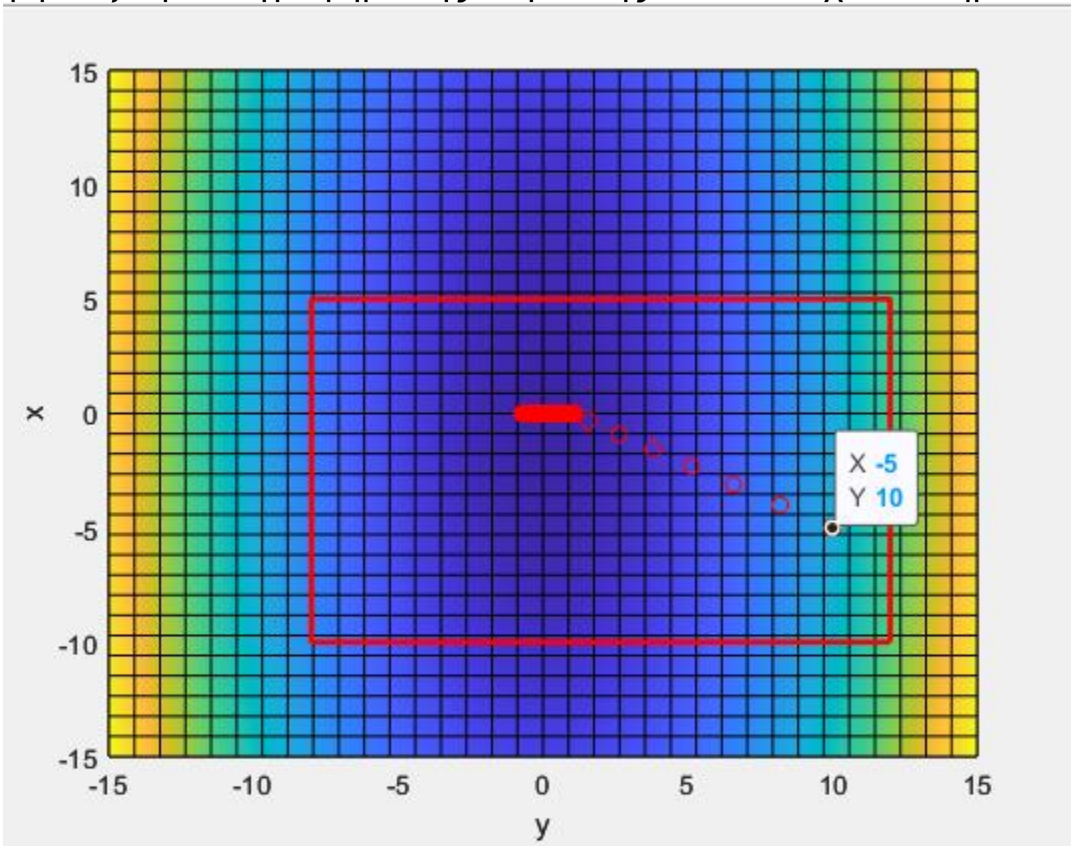
Από τα παραπάνω, φαίνεται ότι το s_k θα πρέπει να είναι μεγαλύτερο του μηδενός (προφανώς, αφού αποτελεί προϋπόθεση) και μικρότερο του $2/3$, ώστε η μέθοδος να συγκλίνει στο ελάχιστο. Γι' αυτό και εδώ έχουμε απόκλιση του αλγορίθμου. Μόνο η 1η συνιστώσα θα συγκλίνει στο 0, αφού η συνθήκη σύγκλισης για αυτή είναι $s < 6$. Εδώ έχουμε s=5. Αντιθέτως, η 2η συνιστώσα δε θα συγκλίνει στο 0, αφού η συνθήκη σύγκλισης για αυτή είναι $s < \frac{2}{3}$.

Παρατηρούμε ότι, σε σχέση με το Θέμα 1, όπου το όριο του γ_k αποτελεί τη συνθήκη σύγκλισης, η αντίστοιχη συνθήκη εγώ είναι το επιτρεπτό όριο του s_k. Η ίδια ανάλυση για τη σύγκλιση ή μη των συνιστωσών των διαδοχικών σημείων κάθε φορά ισχύει και στο Θέμα 1, στις τιμές του γ_k.

Θέμα 3

Με την ίδια ακριβώς λογική, όπως στο Θέμα 2, για τη μέθοδο της Μέγιστης Καθόδου με Προβολή, θέτουμε $s_k = 15$, γ_k = 0.1 και αρχικό σημείο εκκίνησης Xo = [-5, 10] και τρέχουμε τον αλγόριθμο. Το αποτέλεσμα δείχνει ότι αυτές οι αρχικές τιμές οδηγούν σε σύγκλιση του αλγορίθμου στο ολικό ελάχιστο [0, 0]. Πράγματι το ελάχιστο, στα k=599 βήματα, εμφανίζεται ότι θα είναι το [0, 0.00023207].

Εμφανίζουμε το γράφημα της σύγκλισης των διαδοχικών σημείων:



Φαίνεται ότι όντως υπάρχει σύγκλιση στο ελάχιστο [0, 0], αλλά με αργό ρυθμό. Τα διαδοχικά σημεία θα είναι k = 599. Το κόκκινο ορθογώνιο αποτελεί το σύνολο *Π* των περιορισμών, μέσα στο οποίο πρέπει να βρίσκονται όλα τα σημεία x_k.

Ακολουθώντας ξανά την ίδια μέθοδο του Θέματος 2 για τον υπολογισμό των επιτρεπτών s_k που οδηγούν τον αλγόριθμο σε σύγκλιση, όταν $\gamma_k = 0.1$, έχουμε:

$$x_{k+1} = x_k + \gamma_k \cdot (\bar{x}_k - x_k)$$

με $\bar{x}_k = P_{r_\Pi}(x_k - s_k \cdot \nabla_f(x_k))$ => $\bar{x}_k = P_{r_\Pi}(\begin{bmatrix} x \\ y \end{bmatrix} - s \cdot \begin{bmatrix} \frac{2}{3}x \\ 6y \end{bmatrix})$ =>

$$=> \bar{x}_k = P_{r_\Pi}(\begin{bmatrix} x(1-\frac{2}{3}s) \\ y(1-6s) \end{bmatrix}) \;\; => \;\; \bar{x}_k = \begin{bmatrix} x(1-\frac{2}{3}s) \\ y(1-6s) \end{bmatrix}$$

Επομένως, έχουμε:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0.9x \\ 0.9y \end{bmatrix} + \begin{bmatrix} 0.1x(1-\frac{2}{3}s) \\ 0.1y(1-6s) \end{bmatrix} \quad =>$$

$$x_{k+1} = x_k \cdot (1 - \frac{2}{30}s)$$
$$y_{k+1} = y_k \cdot (1 - 0.6s)$$

Γνωρίζουμε ότι για να συγκλίνει η μέθοδος στο ελάχιστο, θα πρέπει: $\frac{|x_{k+1}|}{|x_k|} < 1$ και $\frac{|y_{k+1}|}{|y_k|} < 1$, επομένως:

$$\left|1-\frac{1}{15}s\right| < 1 \;\; => \;\; -1 < 1-\frac{1}{15}s < 1 \;\; => \;\; s > 0 \;\; \kappa\alpha\iota \;\; s < 30$$

και

$$|1-0.6s| < 1 \;\; => \;\; -1 < 1-0.6s < 1 \;\; => \;\; s > 0 \;\; \kappa\alpha\iota \;\; s < \frac{10}{3}$$

Από τα παραπάνω, φαίνεται ότι το s_k θα πρέπει να είναι μεγαλύτερο του μηδενός (προφανώς, αφού αποτελεί προϋπόθεση) και μικρότερο του $10/3$, ώστε η μέθοδος να συγκλίνει στο ελάχιστο. Ωστόσο, εμείς έχουμε $s_k = 15$, και παρόλα αυτά η μέθοδος συγκλίνει(!). Αυτό μπορεί να εξηγηθεί με κάποιο τρόπο. Αν παρατηρήσουμε το γράφημα παραπάνω, βλέπουμε ότι η μέθοδος συγκλίνει στο 0 αρχικά μόνο στη 1η συνιστώσα των διαδοχικών σημείων. Αυτό είναι λογικό, αφού το κριτήριο σύγκλισης της 1ης συνιστώσας είναι $s < 30$. Στη συνέχεια, παρατηρείται μία ταλάντωση στον y άξονα και γύρω από το 0, η οποία τελικά θα οδηγήσει και τη 2η συνιστώσα στο 0 (αν και δεν ικανοποιείται το κριτήριο σύγκλισης εδώ). Η οδήγηση αυτή στο 0, της 2ης συνιστώσας, είναι αποτέλεσμα του πολύ μικρού γ_k που θέσαμε εξαρχής. Το μικρό βήμα γ_k οδηγεί την ταλάντωση της μιας συνιστώσας σε ολοένα και πιο μικρές τιμές, μέχρι που προσεγγίζει το 0, όταν η άλλη συνιστώσα εξαρχής συγκλίνει.

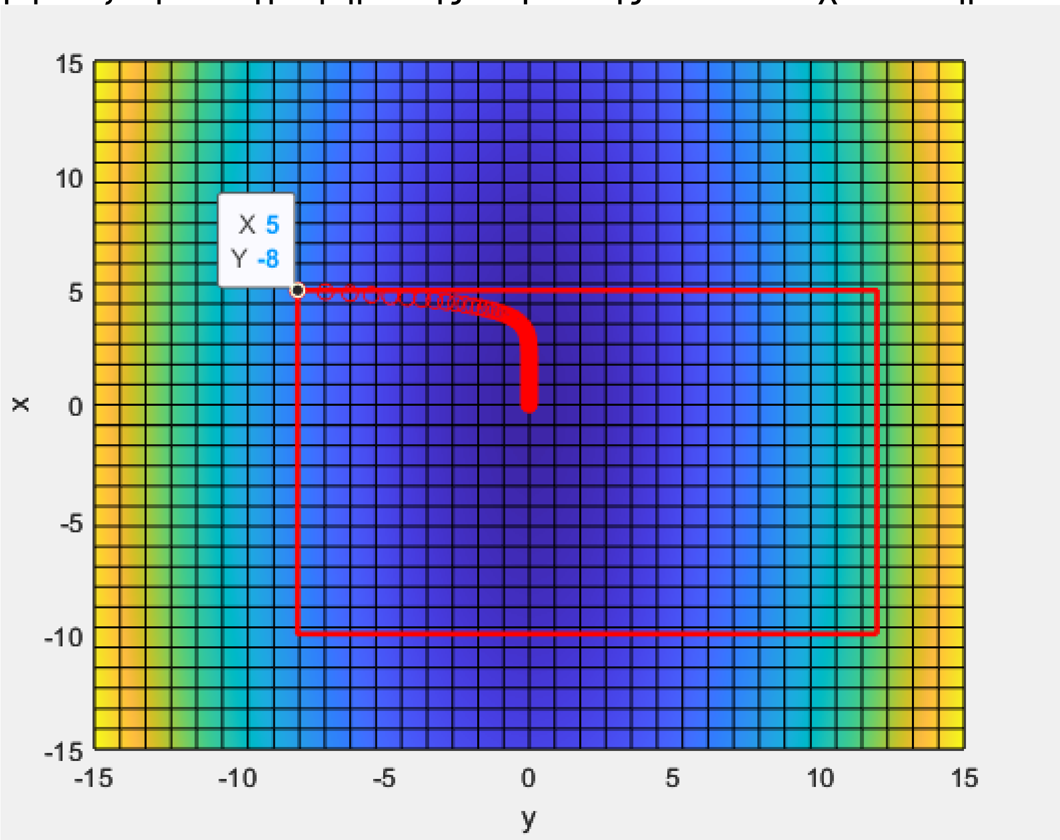
Για να συγκλίνει λοιπόν, η μέθοδος, θα πρέπει:

1) Να ισχύει το κριτήριο σύγκλισης του s_k και για τις δύο συνιστώσες
2) Να ισχύει το κριτήριο σύγκλισης του s_k για τη μία μόνο συνιστώσα, αλλά υποχρεωτικά να έχουμε ικανοποιητικά μικρό βήμα γ_k.

Θέμα 4

Παρόμοια, για τη μέθοδο της Μέγιστης Καθόδου με Προβολή, θέτουμε $s_k = 0.1$, γ_k = 0.2 και αρχικό σημείο εκκίνησης Xo = [8, -10] και τρέχουμε τον αλγόριθμο. Το αποτέλεσμα δείχνει ότι αυτές οι αρχικές τιμές οδηγούν σε σύγκλιση του αλγορίθμου στο ολικό ελάχιστο [0, 0]. Πράγματι το ελάχιστο, στα k=434 βήματα, εμφανίζεται ότι θα είναι το [0.01495469, 0]

Εμφανίζουμε το γράφημα της σύγκλισης των διαδοχικών σημείων:



Φαίνεται ότι όντως υπάρχει σύγκλιση στο ελάχιστο [0, 0], αλλά με αργό ρυθμό. Τα διαδοχικά σημεία θα είναι k = 434. Το κόκκινο ορθογώνιο αποτελεί το σύνολο *Π* των περιορισμών, μέσα στο οποίο πρέπει να βρίσκονται όλα τα σημεία x_k.

Ακολουθώντας ξανά την ίδια μέθοδο του Θέματος 2 για τον υπολογισμό των επιτρεπτών s_k που οδηγούν τον αλγόριθμο σε σύγκλιση, όταν $\gamma_k = 0.2$, έχουμε:

$$x_{k+1} = x_k + \gamma_k \cdot (\bar{x}_k - x_k)$$

με $\bar{x}_k = P_{r_\Pi}(x_k - s_k \cdot \nabla_f(x_k))$ => $\bar{x}_k = P_{r_\Pi}(\begin{bmatrix} x \\ y \end{bmatrix} - s \cdot \begin{bmatrix} \frac{2}{3}x \\ 6y \end{bmatrix})$ =>

$$=> \bar{x}_k = P_{r_\Pi}\left(\begin{bmatrix} x(1-\frac{2}{3}s) \\ y(1-6s) \end{bmatrix}\right) \quad => \quad \bar{x}_k = \begin{bmatrix} x(1-\frac{2}{3}s) \\ y(1-6s) \end{bmatrix}$$

Επομένως, έχουμε:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0.8x \\ 0.8y \end{bmatrix} + \begin{bmatrix} 0.2x(1-\frac{2}{3}s) \\ 0.2y(1-6s) \end{bmatrix} \quad =>$$

$$x_{k+1} = x_k \cdot (1 - \frac{2}{15}s)$$
$$y_{k+1} = y_k \cdot (1 - 1.2s)$$

Γνωρίζουμε ότι για να συγκλίνει η μέθοδος στο ελάχιστο, θα πρέπει: $\frac{|x_{k+1}|}{|x_k|} < 1$ και $\frac{|y_{k+1}|}{|y_k|} < 1$, επομένως:

$$\left|1-\frac{2}{15}s\right| < 1 \;=>\; -1 < 1-\frac{2}{15}s < 1 \;=>\; s > 0 \;\;\kappa\alpha\iota\;\; s < 15$$

και

$$|1-1.2s| < 1 \;=>\; -1 < 1-1.2s < 1 \;=>\; s > 0 \;\;\kappa\alpha\iota\;\; s < \frac{5}{3}$$

Από τα παραπάνω, φαίνεται ότι το s_k θα πρέπει να είναι μεγαλύτερο του μηδενός (προφανώς, αφού αποτελεί προϋπόθεση) και μικρότερο του $5/3$, ώστε η μέθοδος να συγκλίνει στο ελάχιστο. Εδώ, έχουμε $s_k = 0.1$, επομένως ισχύει το κριτήριο σύγκλισης και για τις δύο συνιστώσες. Η αργή σύγκλιση είναι αποτέλεσμα του μικρού βήματος γ_k.

Παρατηρούμε ότι το αρχικό σημείο εκκίνησης βρίσκεται εκτός του περιορισμού *Π*, επομένως ο αλγόριθμος αρχικά θα το προβάλει πάνω στο *Π* και έπειτα θα συνεχίσει. Για το λόγο αυτό στο παραπάνω σχήμα φαίνεται το αρχικό σημείο ως το [5, -8]. Από εκεί και πέρα, γνωρίζουμε ότι η μέθοδος θα συγκλίνει στο ελάχιστο με αργό ρυθμό (εξαιτίας του μικρού γ_k). Η σύγκλιση είναι σχετικά ομοιόμορφη και για τις δύο συνιστώσες του διανύσματος του διαδοχικού, κάθε φορά, σημείου.