3η Εργαστηριακή Άσκηση: Μέθοδος Μέγιστης Καθόδου με Προβολή

Θέμα 1ο :

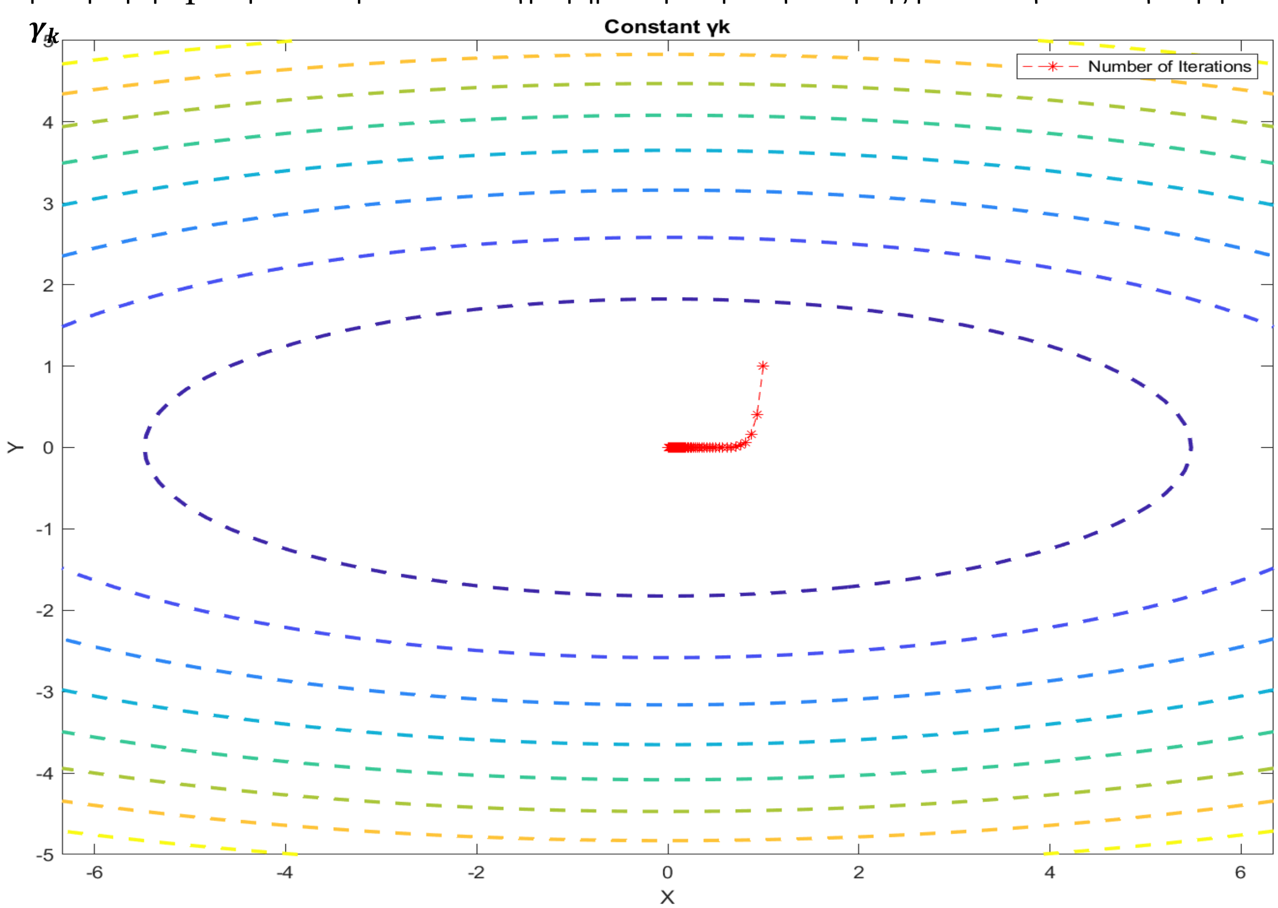
Στην παρούσα εργασία, η αντικειμενική συνάρτηση στην οποία θα εφαρμόσουμε την μέθοδο της μέγιστης καθόδου είναι η εξής : $\boldsymbol{f(\vec{x}) \ = \ \frac{1}{3}x_1^2 + 3x_2^2}$

Η συνάρτηση $f(x)$ είναι γνήσια κυρτή συνάρτηση διότι ο Εσσιανός πίνακας της είναι θετικά ορισμένος αφού ισχύει $\boldsymbol{\nabla^2 f(x_1, x_2) = \begin{bmatrix} 2/3 & 0 \\ 0 & 6 \end{bmatrix}}$ από το οποίο φαίνεται ότι όλες οι ιδιοτιμές του είναι θετικές και άρα θα είναι θετικά ορισμένος. Επιπλέον, αφού η συνάρτηση που μελετάμε είναι μια γνήσια κυρτή συνάρτηση το τοπικό ελάχιστο της (το οποίο προφανώς είναι το σημείο $(\mathbf{0}, \mathbf{0})$) θα είναι επίσης και το ολικό ελάχιστο της.

Αρχικά θέλουμε να δούμε αν για συγκεκριμένα $\boldsymbol{\gamma_k}$, τα οποία μας δίνονται στην εκφώνηση, η μέθοδος θα βρεί το ολικό ελάχιστο της αντικειμενικής συνάρτησης. Θεωρώντας ως αρχικό σημείο εκκίνησης το $(\mathbf{1}, \mathbf{1})$ παρατηρούμε τα εξής για :

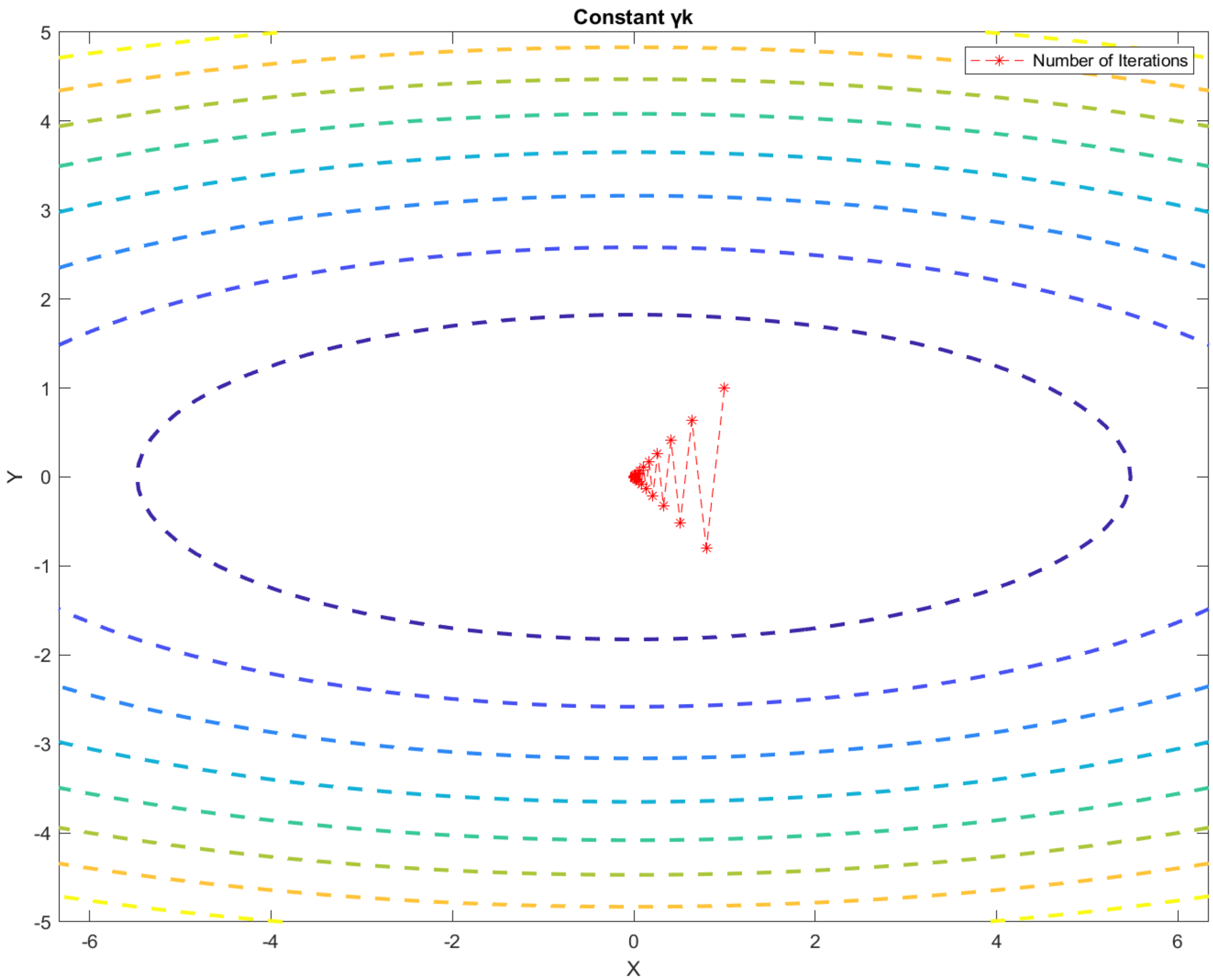
a) $\boldsymbol{\gamma_k = 0.1}$:

Παρατηρούμε ότι για αυτή την τιμή του $\boldsymbol{\gamma_k}$ η μέθοδος μας συγκλίνει στο ελάχιστο της συνάρτησης και συγκεκριμένα βλέπουμε ότι το $\boldsymbol{x_2}$ πηγαίνει πιο γρήγορα στο ελάχιστο της συνάρτησης σε σχέση με την μεταβλητή $\boldsymbol{x_1}$. Παρακάτω φαίνονται τα γραφήματα με την σύγκλιση της μεθόδου για το συγκεκριμένο $\boldsymbol{\gamma_k}$



b) $\gamma_k = 0.3$**:**

Παρατηρούμε ότι για αυτή την τιμή του γ_k η μέθοδος μας συγκλίνει στο ελάχιστο της συνάρτησης και συγκεκριμένα βλέπουμε ότι και το x_1αλλά και το x_2 συγκλίνουν με τον ίδιο ρυθμό στο ελάχιστο της συνάρτησης, ωστόσο το x_2 ταλαντώνεται περισσότερο από το x_1 κάτι που θα εξηγήσουμε παρακάτω όπου θα αποδείξουμε γιατί συμβαίνει αυτό με μαθηματική αυστηρότητα. Παρακάτω φαίνονται τα γραφήματα με την σύγκλιση της μεθόδου για το συγκεκριμένο γ_k.



c) $\gamma_k = 3$:

Για το συγκεκριμένο βήμα παρατηρούμε ότι η μέθοδος μέγιστης καθόδου δεν συγκλίνει στο ελάχιστο της συνάρτησης καθώς βλέπουμε ότι το x_1 ταλαντώνεται μεταξύ του **1** και του **-1(**σαν δηλαδή να ταλαντώνεται γύρω από το αρχικό σημείο εκκίνησης που επιλέξαμε**)** ενώ αντίθετα το x_2 αποκλίνει και τείνει στο άπειρο. Για τον λόγο αυτό δεν έχει νόημα να σχεδιάσουμε το σχήμα με τις ισοβαρείς καμπύλες και την σύγκλιση της μεθόδου διότι η μέθοδος , όπως γίνεται φανερό δεν συγκλίνει στο ελάχιστο της συνάρτησης.

d) $\gamma_k = 5$:

Με παρόμοιο τρόπο παρατηρούμε ότι ούτε για αυτή την τιμή του γ_k έχουμε σύγκλιση, ενώ επίσης τώρα παρατηρούμε ότι και το x_1 αποκλίνει και τείνει και αυτό πλέον στο άπειρο. Επομένως και για αυτή την τιμή του γ_k η μέθοδος μέγιστης καθόδου δεν θα συγκλίνει στο ελάχιστο της συνάρτησης και άρα δεν έχει νόημα η γραφική παράσταση των ισοβαρών καμπυλών που αναπαριστά την σύγκλιση της μεθόδου στο ελάχιστο.

Από τα παραπάνω συμπεραίνουμε ότι έχουμε σύγκλιση στο ελάχιστο της αντικειμενικής συνάρτησης μόνο για τα πρώτα δύο γ_k ενώ για τα υπόλοιπα όπως αναφέρθηκε και προηγουμένως η μέθοδος δεν θα συγκλίνει και θα αποκλίνει. Ο λόγος που συμβαίνει αυτό είναι γιατί για να έχουμε σύγκλιση θα πρέπει να ισχύουν τα εξής:

$$\left|\frac{x_{k+1}}{x_k}\right| < 1 \Rightarrow \frac{\sqrt{x_{1k}^2\left(1-\frac{2\gamma_k}{3}\right)^2 + x_{2k}^2(1-6\gamma_k)^2}}{\sqrt{x_{1k}^2 + x_{2k}^2}} < 1 \Rightarrow$$

$$x_{1k}^2\left(1-\frac{2\gamma_k}{3}\right)^2 + x_{2k}^2(1-6\gamma_k)^2 < x_{1k}^2 + x_{2k}^2 \Rightarrow x_{1k}^2\left[\left(1-\frac{2\gamma_k}{3}\right)^2 - 1\right] + x_{2k}^2\left[(1-6\gamma_k)^2 - 1\right] < 0$$

Άρα για να είναι μικρότερα του μηδενός και για να συγκλίνει ο αλγόριθμος θα πρέπει να οι ποσότητες βρίσκονται μέσα στις αγκύλες να είναι αρνητικές δηλαδή θα πρέπει :

$$\left[\left(1-\frac{2\gamma_k}{3}\right)^2 - 1\right] < 0 \Rightarrow \left(1-\frac{2\gamma_k}{3}\right)^2 < 1 \Rightarrow \left|1-\frac{2\gamma_k}{3}\right| < 1 \Rightarrow 0 < \gamma_k < 3$$

Το οποίο μας εξασφαλίζει την σύγκλιση του x_1**.**

Και ταυτόχρονα θα πρέπει να ισχύει:

$$\left[(1-6\gamma_k)^2 - 1\right] < 0 \Rightarrow (1-6\gamma_k)^2 < 1 \Rightarrow |1-6\gamma_k| < 1 \Rightarrow 0 < \gamma_k < \frac{1}{3}$$

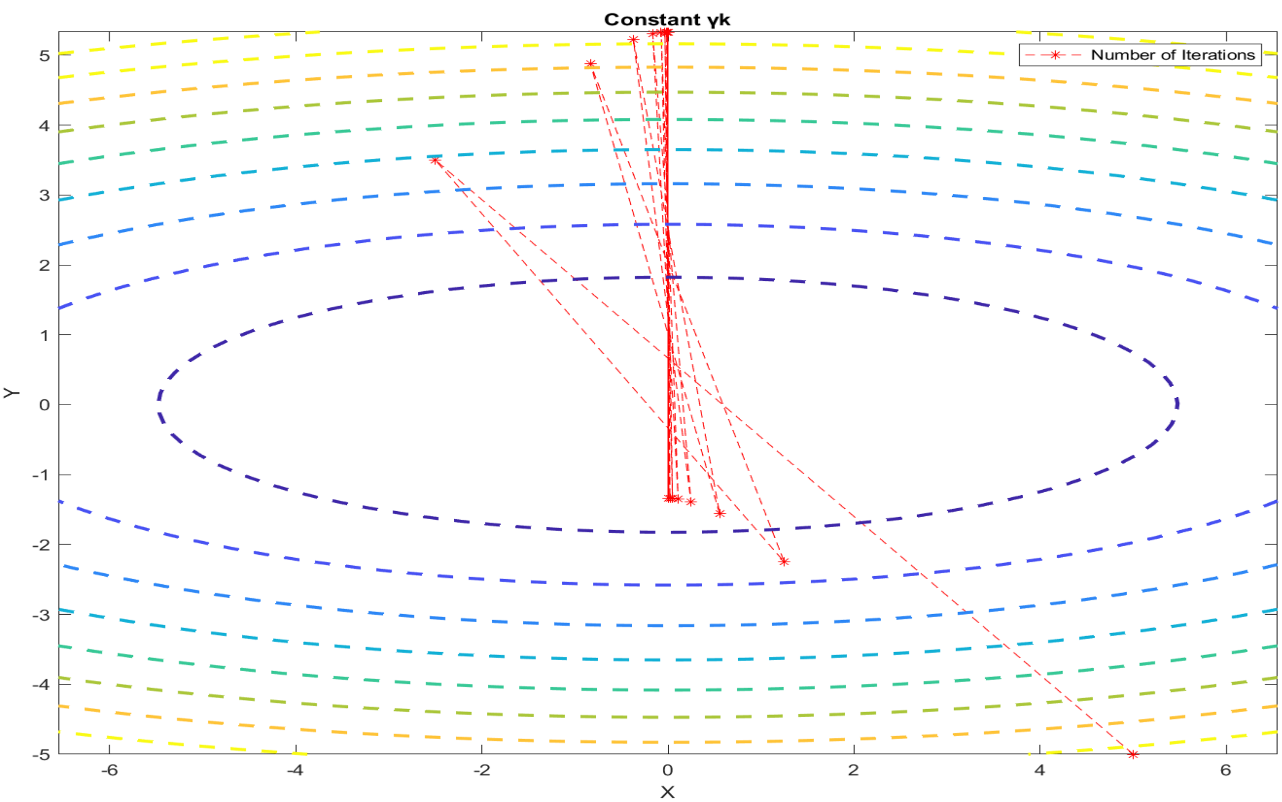
Το οποίο μας εξασφαλίζει την σύγκλιση του x_2

Οπότε από αυτές τις δύο σχέσεις συνεπάγεται ότι για να έχουμε σύγκλιση θα πρέπει να ισχύει
$0 < \gamma_k < \frac{1}{3}$

Επομένως είναι λογικό να μην έχουμε σύγκλιση για τα δύο τελευταία $\boldsymbol{\gamma_k}$ της εκφώνησης αφού θα θέλαμε για να συγκλίνει η μέθοδος της μέγιστης καθόδου το $\boldsymbol{\gamma_k}$ να βρίσκεται μεταξύ των προαναφερθέντων τιμών. Για αυτό κιόλας τον λόγο όταν το $\boldsymbol{\gamma_k}$ είναι μικρό το $\boldsymbol{x_2}$ συγκλίνει πιο γρήγορα στο ελάχιστο από το $\boldsymbol{x_1}$ ενώ όσο αυξάνει βλέπουμε ότι το $\boldsymbol{x_2}$ πηγαίνει με πιο αργό ρυθμό και επίσης βλέπουμε ότι ταλαντώνεται περισσότερο καθώς πλησιάζει στο ολικό ελάχιστο της αντικειμενικής συνάρτησης . Επίσης βλέπουμε ότι για $\boldsymbol{\gamma_k = 3}$ το $\boldsymbol{x_1}$ ταλαντώνεται κάτι που είναι αναμενόμενο γιατί η μέγιστη τιμή που μπορεί να πάρει το $\boldsymbol{\gamma_k}$ για την σύγκλιση του $\boldsymbol{x_1}$ είναι το **3**.

Θέμα 2ο :

Θεωρώντας τώρα τους περιορισμούς $\boldsymbol{-10 \leq x_1 \leq 5}$ και $\boldsymbol{-8 \leq x_2 \leq 12}$ θα επιδιώξουμε να βρούμε το ελάχιστο με την μέθοδο μέγιστης καθόδου με την βοήθεια της προβολής. Το σύνολο στο οποίο βρισκόμαστε είναι ένα κυρτό σύνολο και άρα μπορεί να εφαρμοστεί η μέθοδος της προβολής. Επιπλέον στο συγκεκριμένο θέμα το σημείο εκκίνησης, δηλαδή το $\boldsymbol{(5, -5)}$ είναι ένα εφικτό σημείο καθότι βρίσκεται μέσα στο κυρτό σύνολο. Ωστόσο, παρατηρούμε ότι για τα συγκεκριμένα $\boldsymbol{s_k}$ και $\boldsymbol{\gamma_k}$ που μας δίνονται η μέθοδος δεν θα συγκλίνει στο ελάχιστο αλλά θα ταλαντώνεται γύρω από αυτό. Αυτή η συμπεριφορά της μεθόδου φαίνεται παρακάτω στο σχήμα:



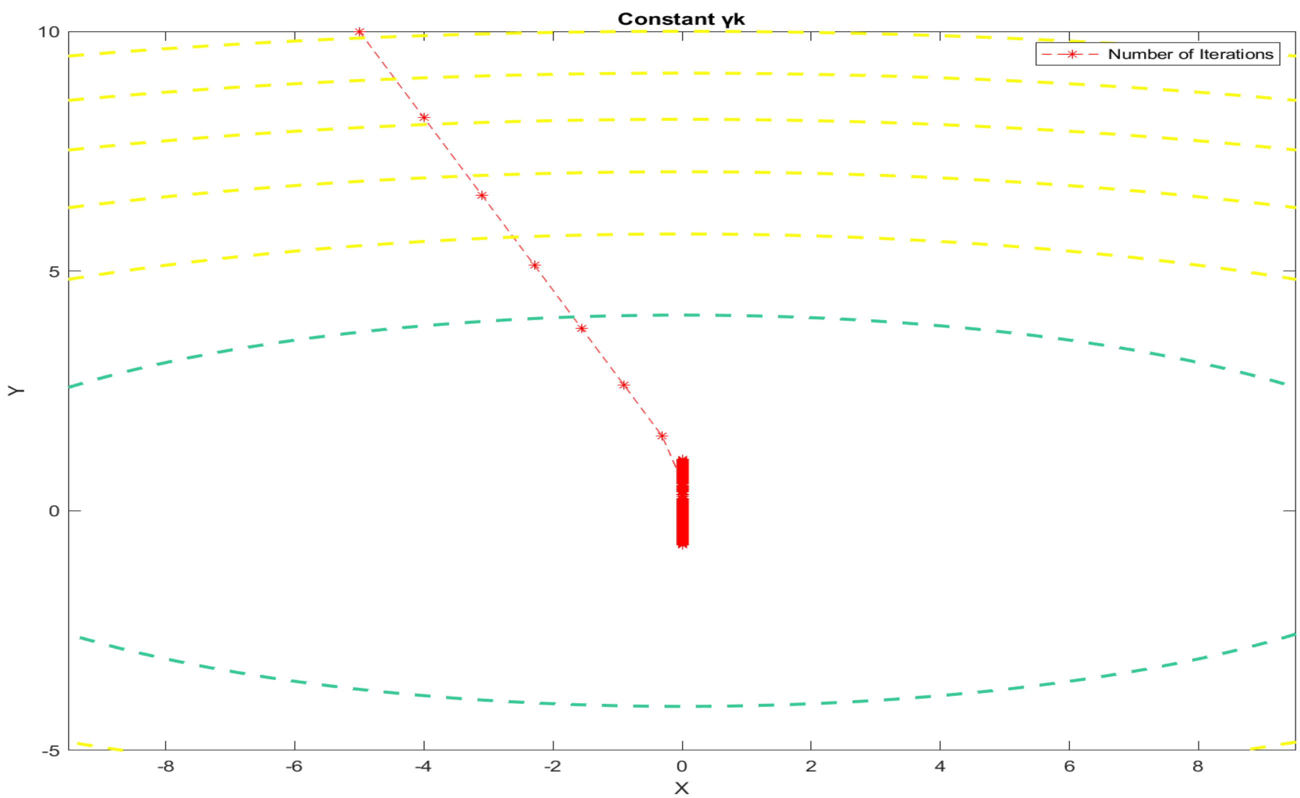
Ο λόγος που συμβαίνει αυτό είναι επειδή :

$\boldsymbol{x_{k+1} = x_k + \gamma_k(\overline{x_k} - x_k)}$ όπου $\boldsymbol{\overline{x_k} = \Pr_X\{x_k - s_k \nabla f(x_k)\}}$

Επομένως από τον συνδυασμό των δύο αυτών εξισώσεων προκύπτει $x_{k+1} = x_k - \widetilde{\gamma_k}\, \nabla f(x_k)$ όπου το $\widetilde{\gamma_k} = \gamma_k s_k$ θεωρώντας προφανώς πως κάθε φορά για την προβολή παίρνουμε το ίδιο το σημείο που προκύπτει από την μέθοδο μέγιστης καθόδου και ότι δηλαδή βρισκόμαστε συνεχώς μέσα στο κυρτό σύνολο που ορίζουν οι περιορισμοί του προβλήματος μας . Συνεπώς στο συγκεκριμένο θέμα το $\widetilde{\gamma_k}$ θα είναι ίσο με $\widetilde{\gamma_k} = 2.5$ και άρα η μέθοδος μέγιστης καθόδου με προβολή δεν θα συγκλίνει αφού όπως αναφέραμε και προηγουμένως για να συγκλίνει η μέθοδος θα πρέπει να ισχύει $0 < \widetilde{\gamma_k} < \frac{1}{3}$. Άρα θα συγκλίνει το x_1 και δεν θα συγκλίνει το x_2 γεγονός που εξηγεί τις μεγάλες ταλαντώσεις που έχει καθώς πλησιάζει στο ελάχιστο και επομένως δεν θα καταλήγουμε στο ελάχιστο της συνάρτησης.

Θέμα 3ο :

Σε αυτή την περίπτωση εφαρμόζουμε και πάλι την μέθοδο προβολής μόνο που τώρα το βήμα της προβολής είναι $s_k = 15$ και το $\gamma_k = 0.1$. Επομένως σε αυτή την συγκεκριμένη περίπτωση το $\widetilde{\gamma_k}$ θα είναι ίσο με **1.5** και άρα αναμένουμε και εδώ το x_1 να τείνει στο μηδέν και το x_2 να μην φτάνει ποτέ στο ελάχιστο. Η υπόθεση μας αυτή επιβεβαιώνεται και από την γραφική παράσταση όπου βλέπουμε ότι το x_1 θα συγκλίνει στο ελάχιστο σε μερικές μόνο επαναλήψεις, ενώ αντίθετα βλέπουμε το x_2 να ταλαντώνεται γύρω από το ελάχιστο και μετά από πολλές επαναλήψεις καταλήγουμε τελικά με τελείως τυχαίο τρόπο σε μια τιμή πολύ κοντά στο ολικό ελάχιστο της αντικειμενικής συνάρτησης για την οποία πληρείται η συνθήκη $\|\nabla f(x_k)\| < \varepsilon$. Παρακάτω φαίνεται η γραφική παράσταση της σύγκλισης στο ελάχιστο της μεθόδου με προβολή:

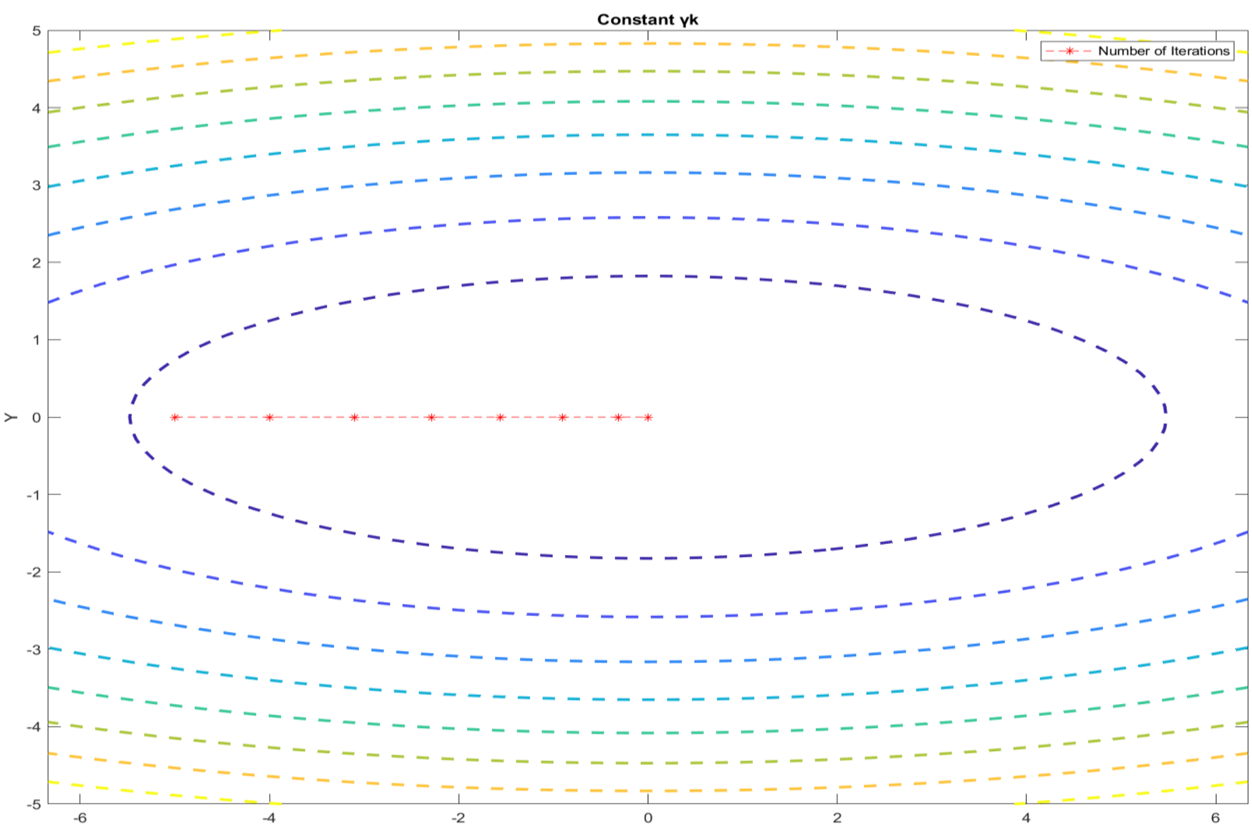


- Οι διαφορές σε σχέση με το **Θέμα 1** είναι ότι στην συγκεκριμένη περίπτωση ο αλγόριθμος, εκμεταλλευόμενος την βοήθεια της μεθόδου της προβολής ,πλησιάζει πολύ κοντά στο πραγματικό ελάχιστο της αντικειμενικής συνάρτησης, παρόλο που το $\widetilde{\gamma_k}$ είναι μεγαλύτερο από το απαιτούμενο

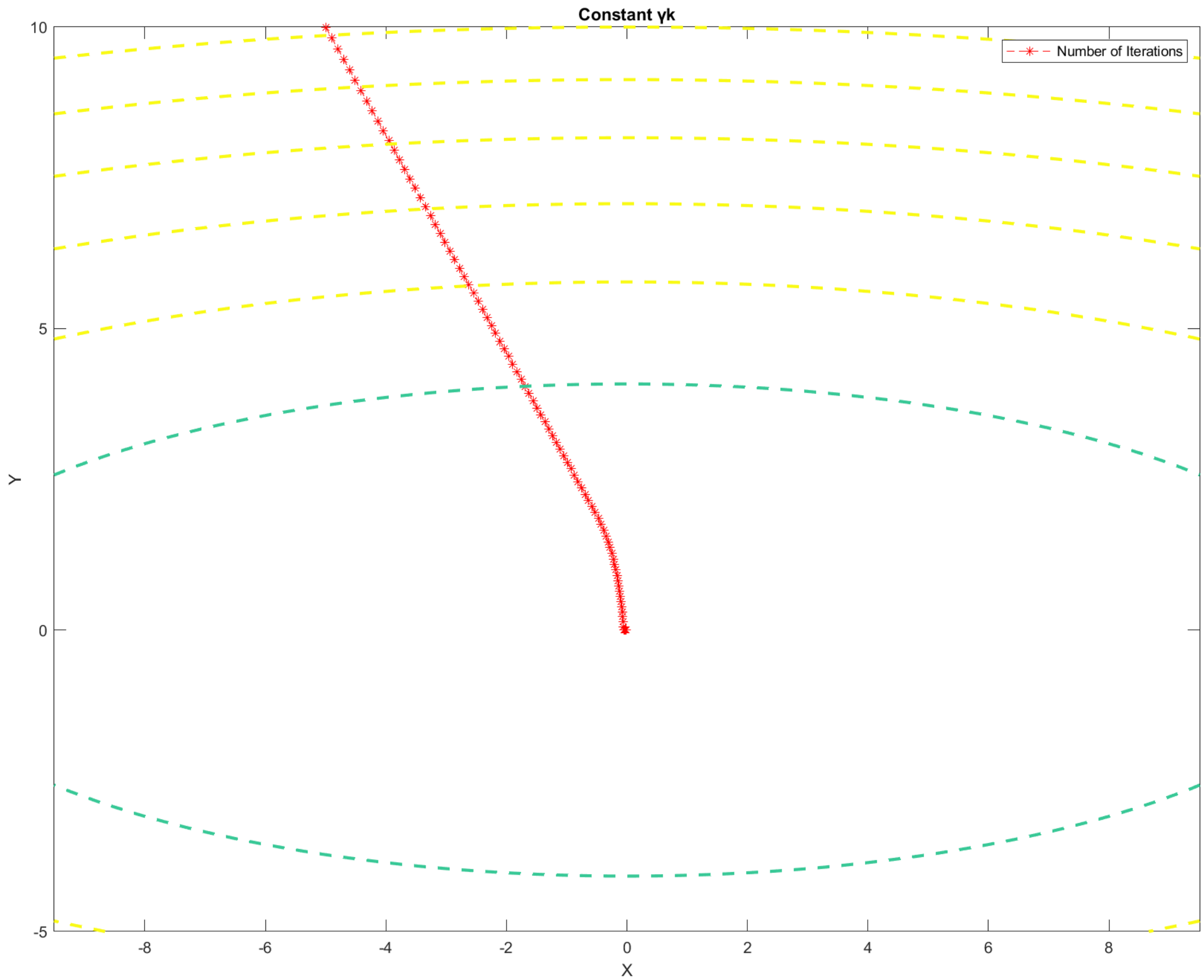
$\frac{1}{3}$ που θέλουμε έτσι ώστε να έχουμε την σύγκλιση στο ολικό ελάχιστο της κυρτής συνάρτησης, ενώ επίσης εδώ πέρα αναζητούμε το ελάχιστο της συνάρτησης σε ένα κυρτό σύνολο το οποίο προσδιορίζεται από τους περιορισμούς του προβλήματος και εμπεριέχει το ολικό ελάχιστο της αντικειμενικής συνάρτησης. Αυτό το γεγονός σε συνδυασμό με την Μέθοδο Μέγιστης Καθόδου με προβολή ίσως να μας επιτρέπει, παρόλο που το $\widetilde{\gamma_k}$ δεν είναι κατάλληλο για σύγκλιση, να συγκλίνουμε στο ελάχιστο της συνάρτησης.

- Επιπρόσθετα, οι διαφορές με το **Θέμα 2** είναι ότι επιτυγχάνουμε και πάλι την σύγκλιση σε ένα σημείο πολύ κοντά στο ελάχιστο. Αυτή η διαφορά ίσως να οφείλεται στο γεγονός ότι το $\widetilde{\gamma_k}$ είναι πολύ κοντά στο άνω όριο που θέλουμε για την σύγκλιση του x_1 (αφού από το **Θέμα 1** είχαμε βρει ότι για να συγκλίνει το x_1 θα έπρεπε να ισχύει $0 < \gamma_k < 3$) ενώ στο συγκεκριμένο θέμα είναι μικρότερο κάτι που ίσως εξηγεί γιατί πηγαίνουμε σε σημείο κοντά στο πραγματικό ολικό ελάχιστο της αντικειμενικής συνάρτησης.

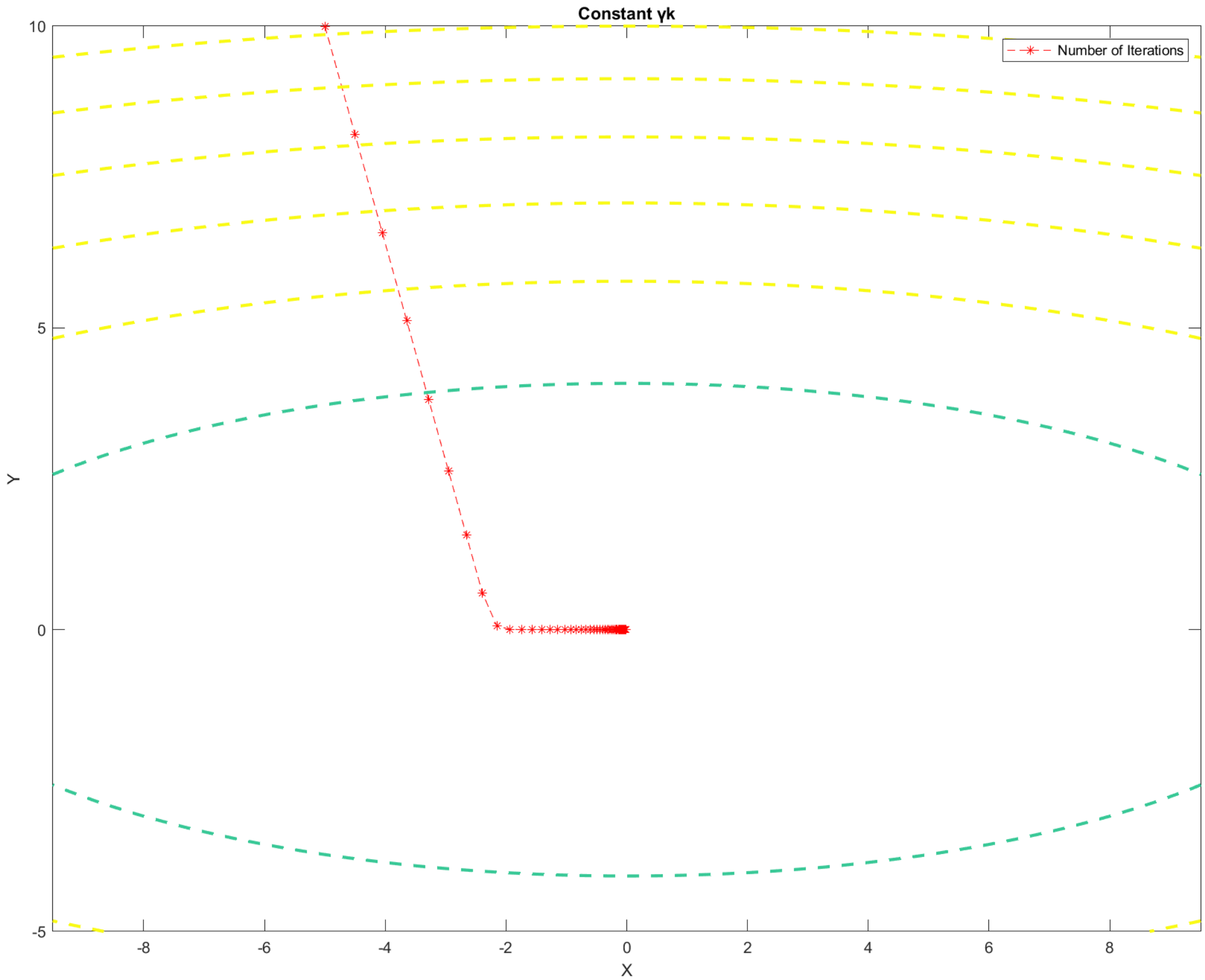
Ένας απλός και πρακτικός τρόπος έτσι ώστε η μέθοδος μέγιστης καθόδου με προβολή να συγκλίνει στο ελάχιστο είναι να ξεκινήσουμε πάνω στον άξονα των x_1 διότι με αυτό τον τρόπο εξασφαλίζουμε ότι το x_2 θα βρίσκεται σίγουρα στο ελάχιστο και επειδή το καθένα συγκλίνει ανεξάρτητα από το άλλο και με διαφορετικό ρυθμό(κάτι που αποδείξαμε στο **Θέμα1**) και άρα λόγω του $\widetilde{\gamma_k} = 1.5$ θα συγκλίνουμε με βεβαιότητα στο ολικό ελάχιστο και μάλιστα με μικρό αριθμό επαναλήψεων διότι η τιμή που παίρνει είναι μέσα στα όρια που επιθυμούμε για την σύγκλιση του x_1. Ένας ακόμη τρόπος με τον οποίο μπορούμε να εξασφαλίσουμε την σύγκλιση είναι παίρνοντας το $\gamma_k = 0.01$ και άρα με αυτό τον τρόπο θα έχουμε $\widetilde{\gamma_k} = 0.15$ συνθήκη που μας εξασφαλίζει την σύγκλιση. Παρακάτω φαίνεται το σχήμα για όταν το σημείο εκκίνησης βρίσκεται πάνω στον άξονα των x_1 για παράδειγμα το **(-5,0)**



Στην συνέχεια φαίνεται η γραφική παράσταση της σύγκλισης για $\widetilde{\gamma_k} = \mathbf{0.15}$**(**δηλαδή ο δεύτερος πρακτικός τρόπος**)** με σημείο εκκίνησης του αλγορίθμου αυτό που δίνεται στην εκφώνηση. Όπου η σύγκλιση επιτυγχάνεται για μόλις 94 επαναλήψεις

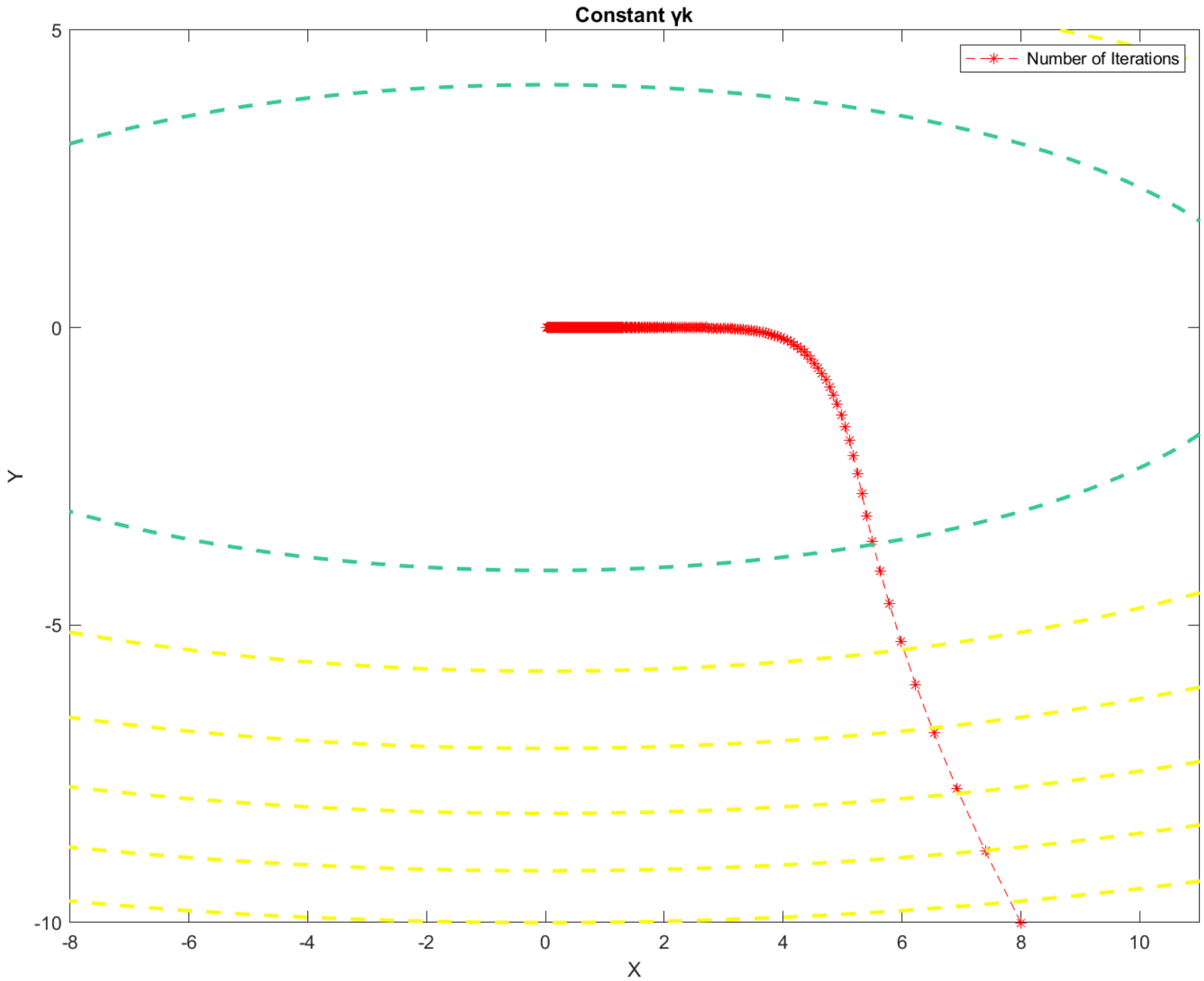


Τέλος παρατηρούμε ότι άμα θέσουμε αντί για το $\boldsymbol{\gamma_k} = \mathbf{0.01}$ αλλά θέσουμε το βήμα της προβολής $\boldsymbol{s_k} = \mathbf{1.5}$**,** δηλαδή έχουμε και πάλι $\widetilde{\gamma_k} = \mathbf{0.15}$**,** θα έχουμε λίγο μικρότερο αριθμό επαναλήψεων για σύγκλιση στο ολικό ελάχιστο της συνάρτησης. Παρακάτω φαίνεται το σχετικό σχήμα:



Θέμα 4ο :

Στο συγκεκριμένο ερώτημα παρατηρούμε ότι το σημείο εκκίνησης βρίσκεται εκτός του κυρτού συνόλου οπότε θα ξεκινήσουμε εξαρχής με την μέθοδο της προβολής έτσι ώστε να βρεθούμε μέσα στο κυρτό σύνολο και να μπορέσουμε να εφαρμόσουμε την μέθοδο μέγιστης καθόδου για να συγκλίνουμε στο ελάχιστο της κυρτής συνάρτησης. Ωστόσο, παρατηρούμε ότι το βήμα $\widetilde{\gamma_k}$ στην συγκεκριμένη περίπτωση είναι πολύ μικρό και άρα αναμένουμε να συγκλίνουμε με πολύ αργό ρυθμό στο ελάχιστο. Παρόλα αυτά η τιμή που παίρνει το $\widetilde{\gamma_k}$ μας εγγυάται σίγουρα την σύγκλιση καθώς βρίσκεται μέσα στα όρια που έχουμε βρεί στο **Θέμα 1** απλά λόγω της πολύ μικρής τιμής του ($\widetilde{\gamma_k} = \mathbf{0.02}$) θα χρειαστεί σίγουρα μεγαλύτερο αριθμό επαναλήψεων.



Αυτό που παρατηρούμε είναι ότι ο αλγόριθμος θα συγκλίνει στο ολικό ελάχιστο της συνάρτησης μετά από έναν μεγάλο αριθμό επαναλήψεων το οποίο προφανώς οφείλεται στην πολύ μικρή τιμή που έχει και το βήμα $\boldsymbol{\gamma_k}$ αλλά και το $\widetilde{\boldsymbol{\gamma_k}}$ το οποίο μας εγγυάται την σύγκλιση στο ελάχιστο της αντικειμενικής κυρτής συνάρτησης . Άμα θέλουμε να συγκλίνουμε γρηγορότερα στο ελάχιστο της συνάρτησης θα έπρεπε να έχουμε μεγαλύτερο βήμα προβολής έτσι ώστε να οδηγηθούμε γρηγορότερα μέσα στο κυρτό σύνολο και άρα να απαιτείται μικρότερος αριθμός επαναλήψεων για την σύγκλιση του αλγορίθμου στο ολικό ελάχιστο. Ωστόσο θα πρέπει να εξακολουθεί να ισχύει ότι το $\widetilde{\boldsymbol{\gamma_k}}$ θα βρίσκεται μέσα στα όρια που βρήκαμε στο **Θέμα 1**, τα οποία θα μας εξασφαλίζουν σε κάθε περίπτωση την σύγκλιση της συνάρτησης.