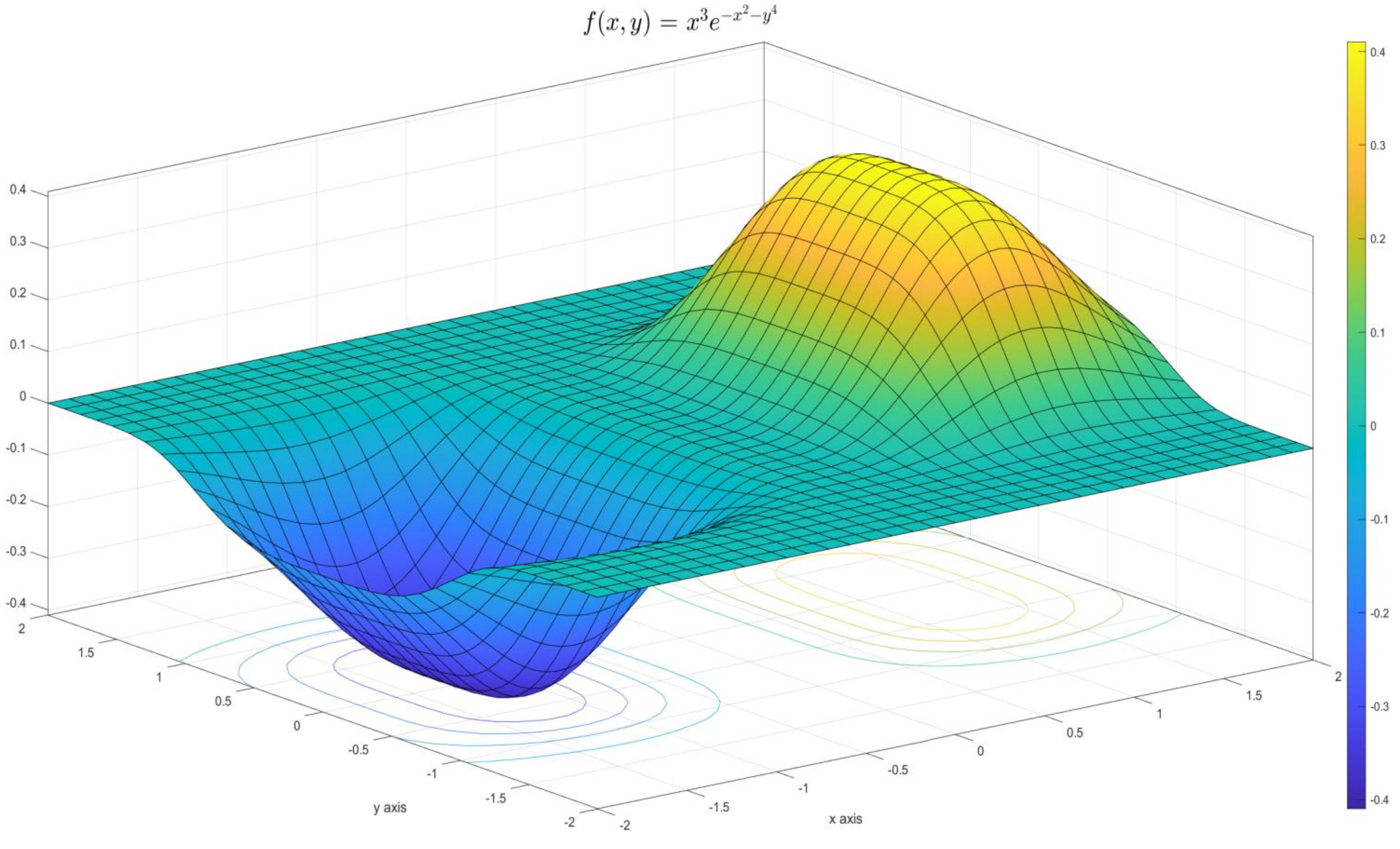
2η Εργαστηριακή Άσκηση: Ελαχιστοποίηση συναρτήσεων πολλών μεταβλητών χωρίς περιορισμούς με χρήση παραγώγων.

Θέμα 1ο :

Σχεδιασμός Συνάρτησης:

Στο συγκεκριμένο θέμα μας ζητείται να σχεδιάσουμε την αντικειμενική συνάρτηση $\boldsymbol{f(x,y) = x^3 e^{-x^2-y^4}}$ έτσι ώστε να πάρουμε μια γενική εικόνα για την μορφή της και το που βρίσκονται τα τοπικά ελάχιστα, τα τοπικά μέγιστα καθώς επίσης και τα σαγματικά σημεία της συνάρτησης. Με αυτό τον τρόπο καθίσταται πιο εύκολη η αξιολόγηση των μεθόδων με βάση το αρχικό σημείο και την ταχύτητα σύγκλισης τους καθώς είναι δυνατό να επαληθευτεί άμα όντως συγκλίνει στο ελάχιστο η μέθοδος. Παρακάτω φαίνεται η συνάρτηση στον τρισδιάστατο χώρο καθώς επίσης και οι προβολές των ισοβαρών καμπυλών πάνω στο επίπεδο **xy.**

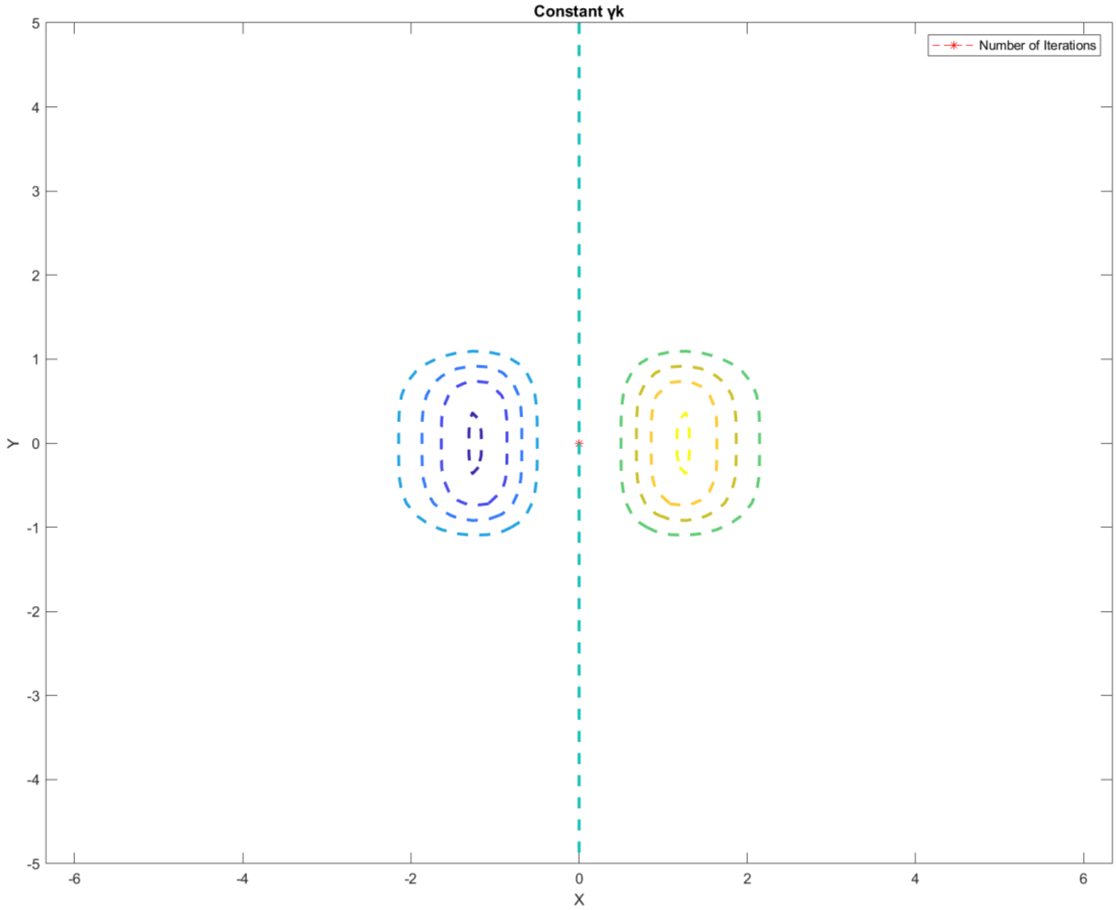


Θέμα 2ο :

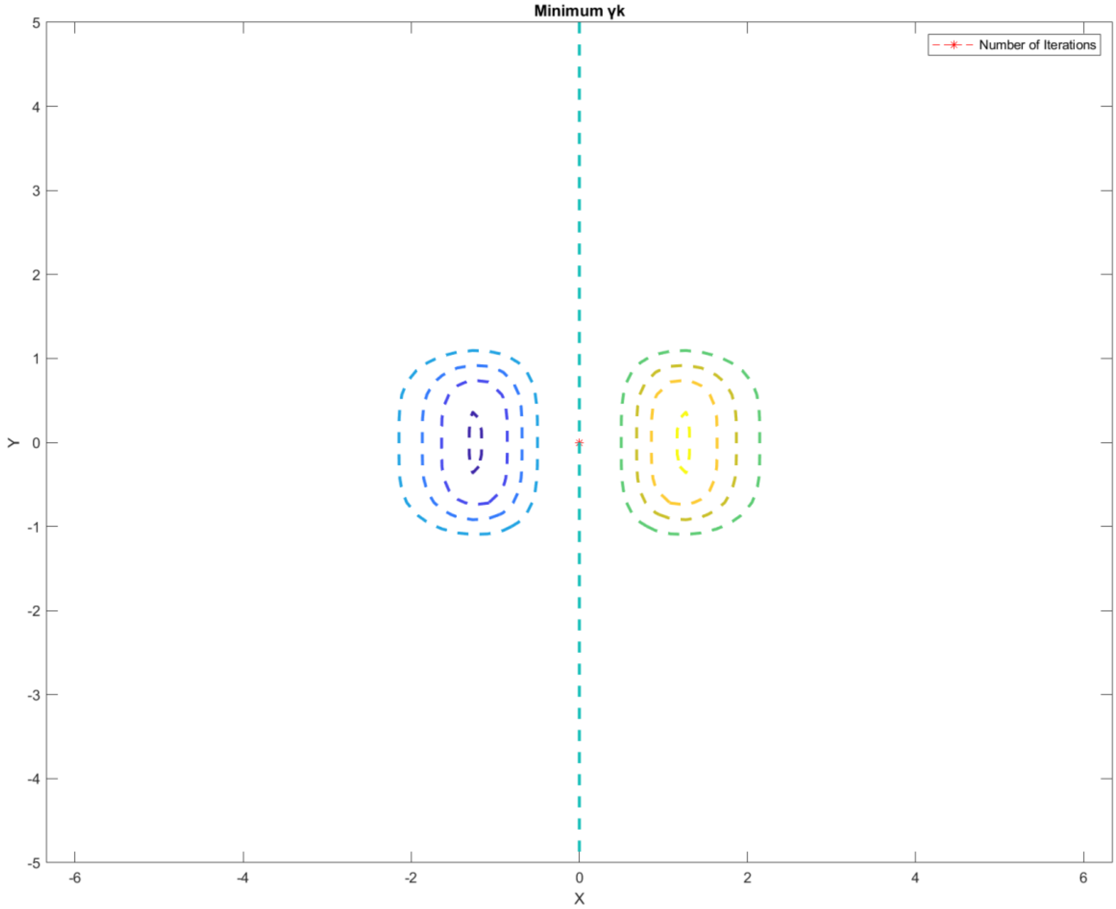
Μέθοδος Μέγιστης Καθόδου:

Η μέθοδος μέγιστης καθόδου είναι ένας από τους τρόπους για τον προσδιορισμό του ελαχίστου μιας συνάρτησης η οποία ικανοποιεί την ιδιότητα της μέγιστης καθόδου δηλαδή $\forall \boldsymbol{k} \in \mathbb{N}$ $\boldsymbol{\iota\sigma\chi}$ύ$\boldsymbol{\varepsilon\iota}$ $\boldsymbol{f}(\boldsymbol{x_k}) > \boldsymbol{f}(\boldsymbol{x_{k+1}})$ και έχει ως διάνυσμα κατεύθυνσης αναζήτησης (δηλαδή την κατεύθυνση στην οποία αναζητούμε το ελάχιστο) το $\boldsymbol{d_k} = -\nabla \boldsymbol{f}(\boldsymbol{x_k})$ έτσι ώστε με αυτό τον τρόπο η γωνία που σχηματίζουν το διάνυσμα κατεύθυνσης αναζήτησης με το $\nabla \boldsymbol{f}(\boldsymbol{x_k})$ να είναι $\mathbf{90}°$

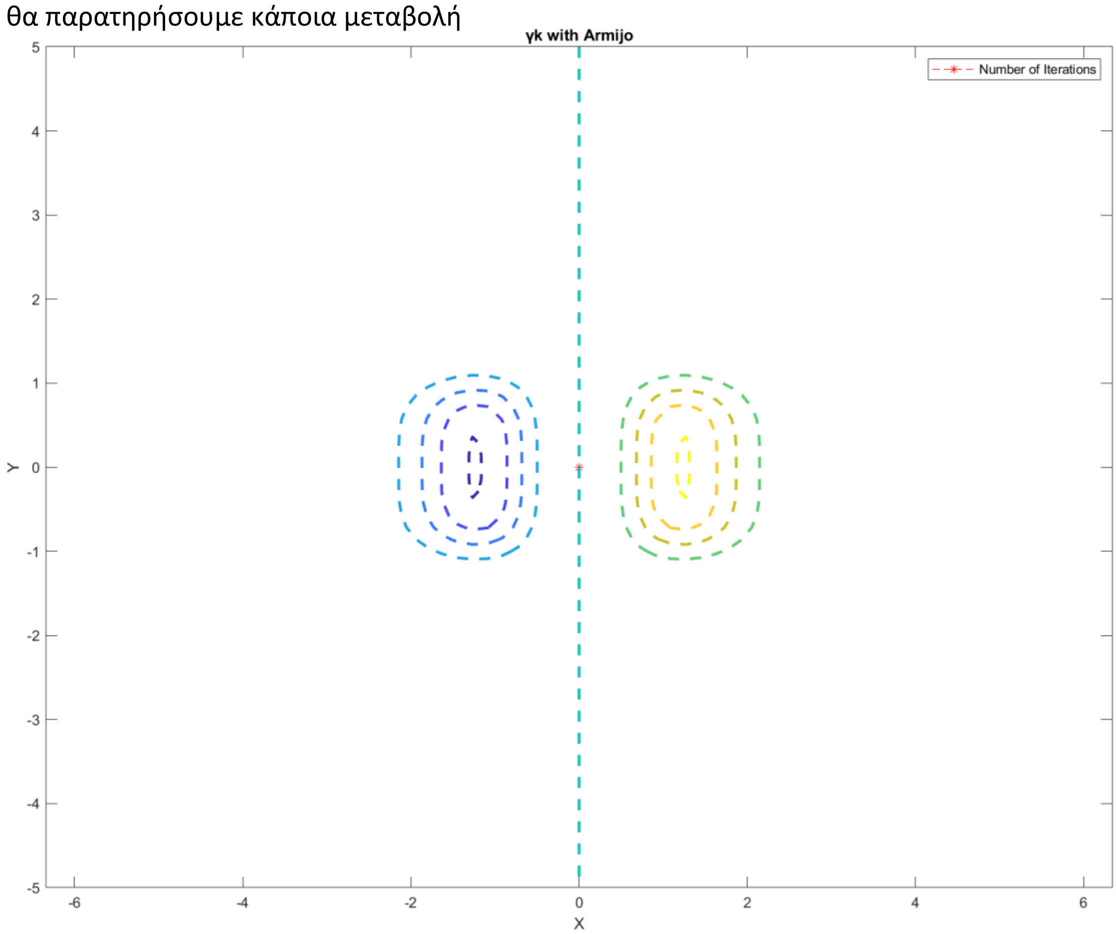
i. **α)** Για το σημείο **(0,0)** δεν έχει σημασία η επιλογή του $\boldsymbol{\gamma_k}$ καθότι σε εκείνο το σημείο όπως φαίνεται και από την γραφική παράσταση της αντικειμενικής συνάρτησης ισχύει ότι $\nabla \boldsymbol{f}(\boldsymbol{x_k}) = \mathbf{0}$ είναι δηλαδή ένα σαγματικό, οπότε η μέθοδος «παγιδεύεται» και δεν μπορεί να προχωρήσει περαιτέρω με αποτέλεσμα να μην μας δίνει το ελάχιστο το οποίο εμείς αναζητούμε.



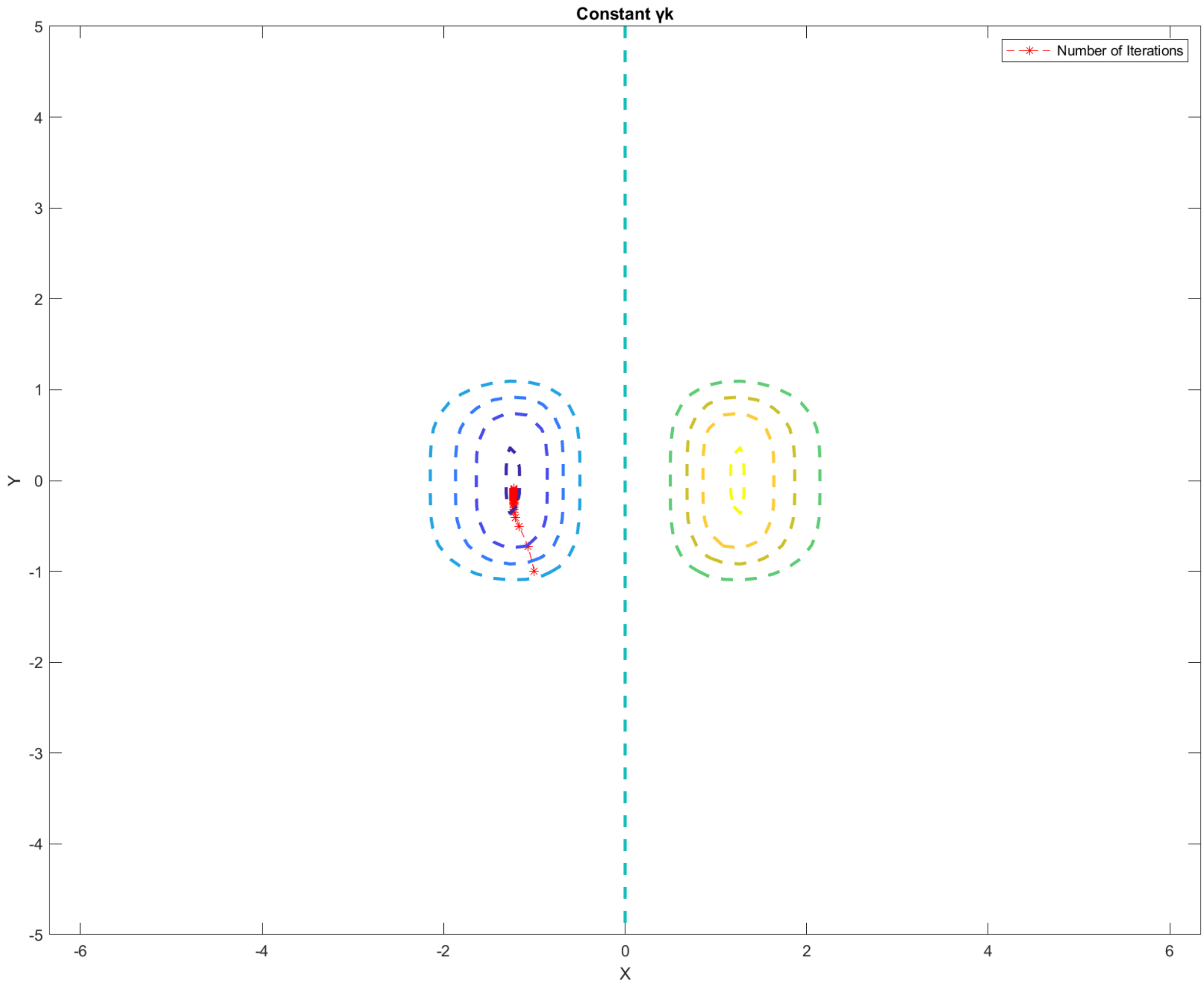
β) Ομοίως, για το ίδιο πάλι σημείο, η επιλογή του $\boldsymbol{\gamma_k}$ τέτοια ώστε να ελαχιστοποιεί την $\boldsymbol{f}(\boldsymbol{x_k} + \boldsymbol{\gamma_k}\boldsymbol{d_k})$ δεν έχει κάποιο νόημα διότι ισχύει $\nabla \boldsymbol{f}(\boldsymbol{x_k}) = \mathbf{0}$ οπότε όπως και πριν η μέθοδος θα «παγιδεύεται» στο σαγματικό σημείο και δεν θα μπορεί να προχωρήσει περαιτέρω.



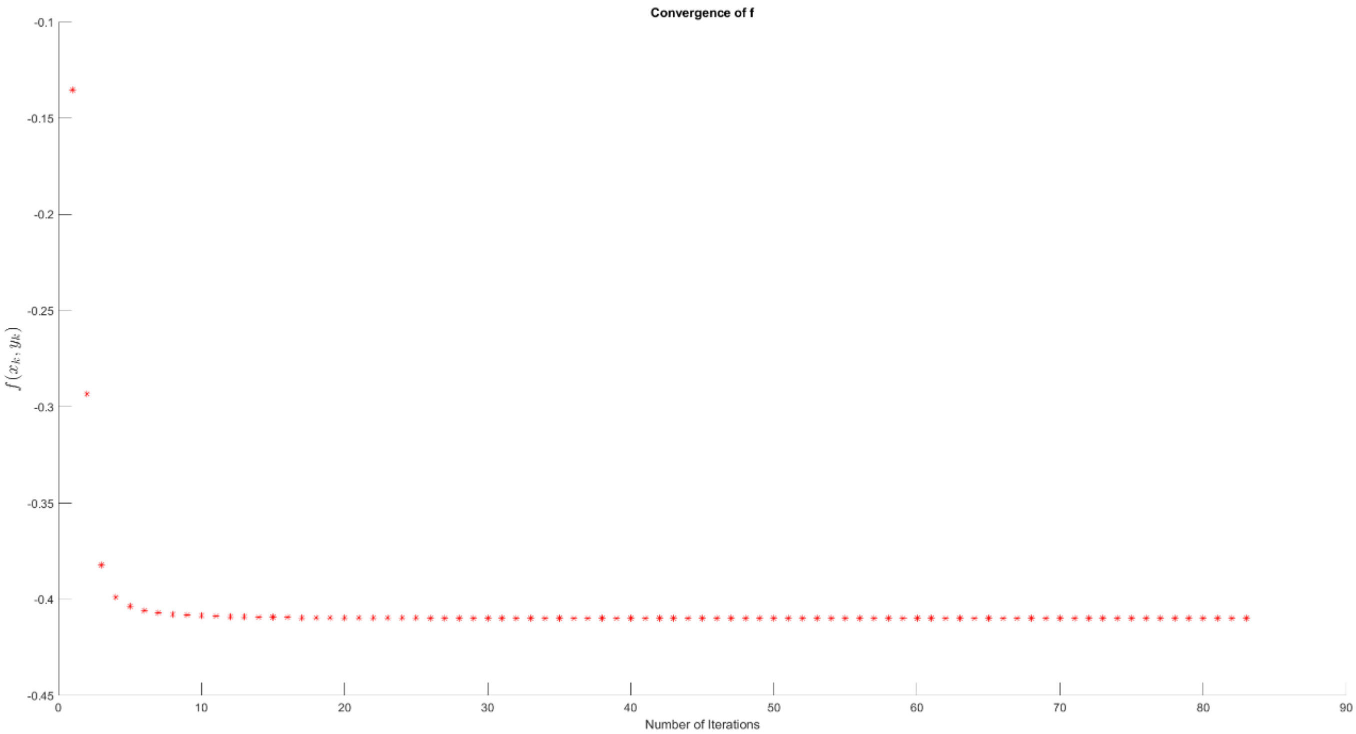
γ) Τέλος, όπως και με τους δύο προηγούμενους τρόπους έτσι και με τον **κανόνα Armijo** δεν θα παρατηρήσουμε κάποια μεταβολή



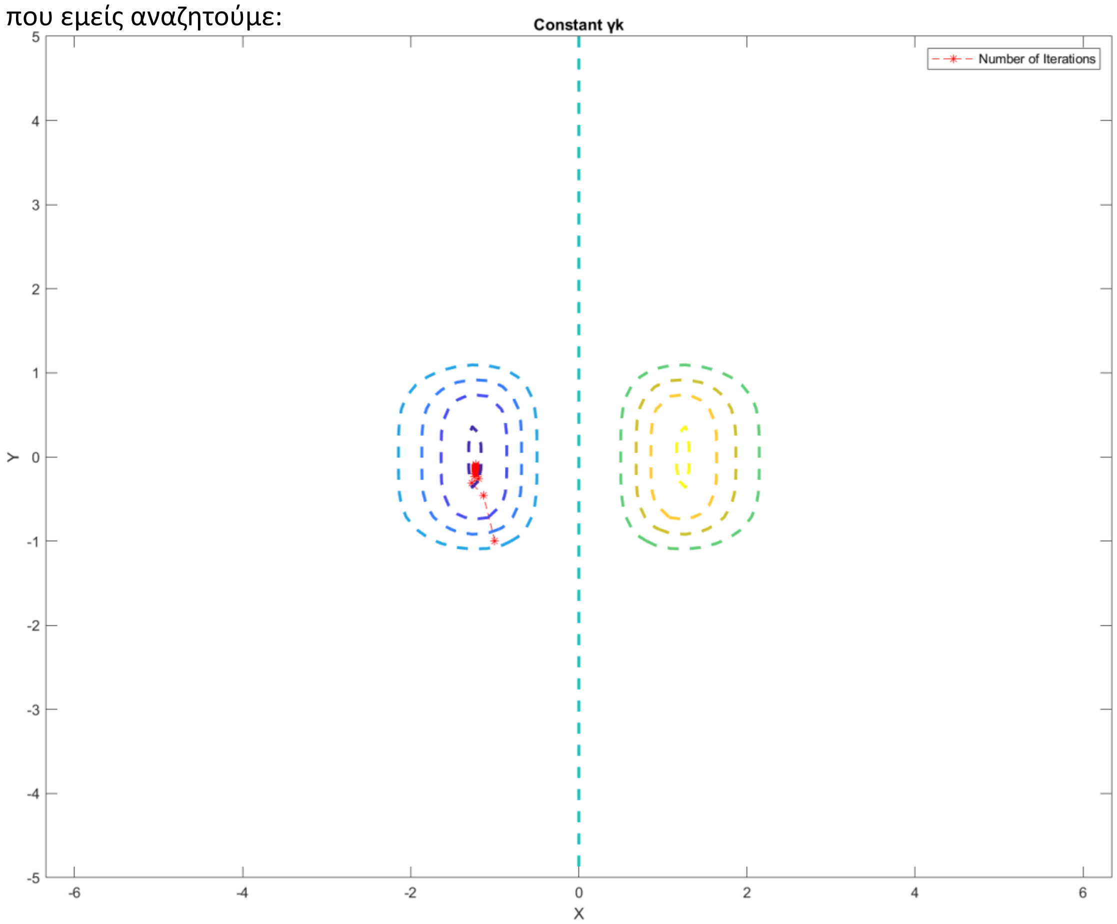
ii. **α)** Για το σημείο **(-1,-1)** παρατηρούμε ότι είναι το **μοναδικό σημείο** που μπορεί να μας οδηγήσει στο **ελάχιστο της συνάρτησης** και για το οποίο η μέθοδος δουλεύει με ικανοποιητικό τρόπο. Η επιλογή του $\boldsymbol{\gamma_k}$ έγινε αρχικά με τυχαίο τρόπο, δηλαδή δοκίμασα πρώτα τιμές οι οποίες να μην είναι ούτε πολύ μεγάλες αλλά ούτε και πολύ μικρές και οι οποίες βρίσκονται στο διάστημα $\boldsymbol{0 < \gamma_k < 1}$ δηλαδή τέτοιες ώστε να ικανοποιούνται τα κριτήρια **(3)** και **(4)** του βιβλίου. Αρχικά δοκίμασα για $\boldsymbol{\gamma_k = 0.5}$ οπότε η γραφική παράσταση της σύγκλισης της αντικειμενικής συνάρτησης που προέκυψε ως προς τον αριθμό των επαναλήψεων, μέχρι να τερματίσει ο αλγόριθμος, είναι η εξής όπου με μεγέθυνση του σχήματος φαίνεται το zik-zak που κάνει ο αλγόριθμος μέχρι να καταλήξει στο ελάχιστο της συνάρτησης:

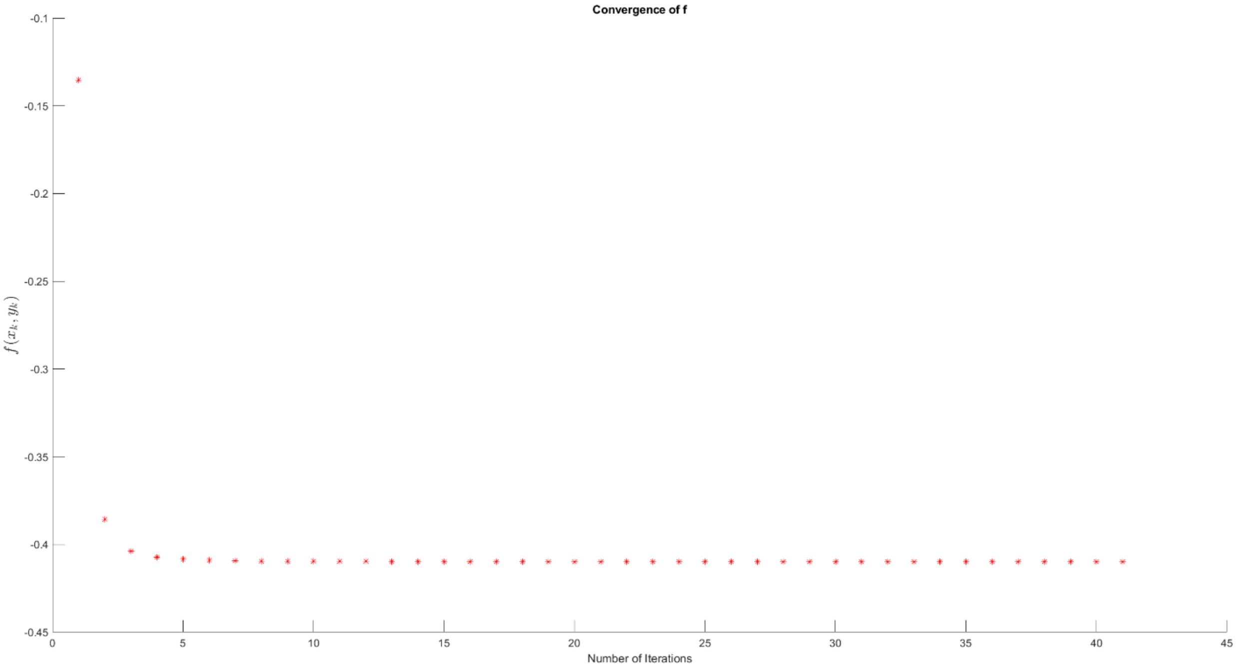


ΣΗΜΕΙΩΣΗ: Στην πραγματικότητα θα έπρεπε να φαίνεται και η καθετότητα των ευθυγράμμων τμημάτων καθώς ο αλγόριθμος συγκλίνει προς το ελάχιστο, αλλά αυτό φαίνεται να μην συμβαίνει ίσως λόγω των προσεγγίσεων που κάνει το Matlab(αυτή η παρατήρηση ισχύει και για τα υπόλοιπα για την μέθοδο της Μέγιστης Καθόδου)

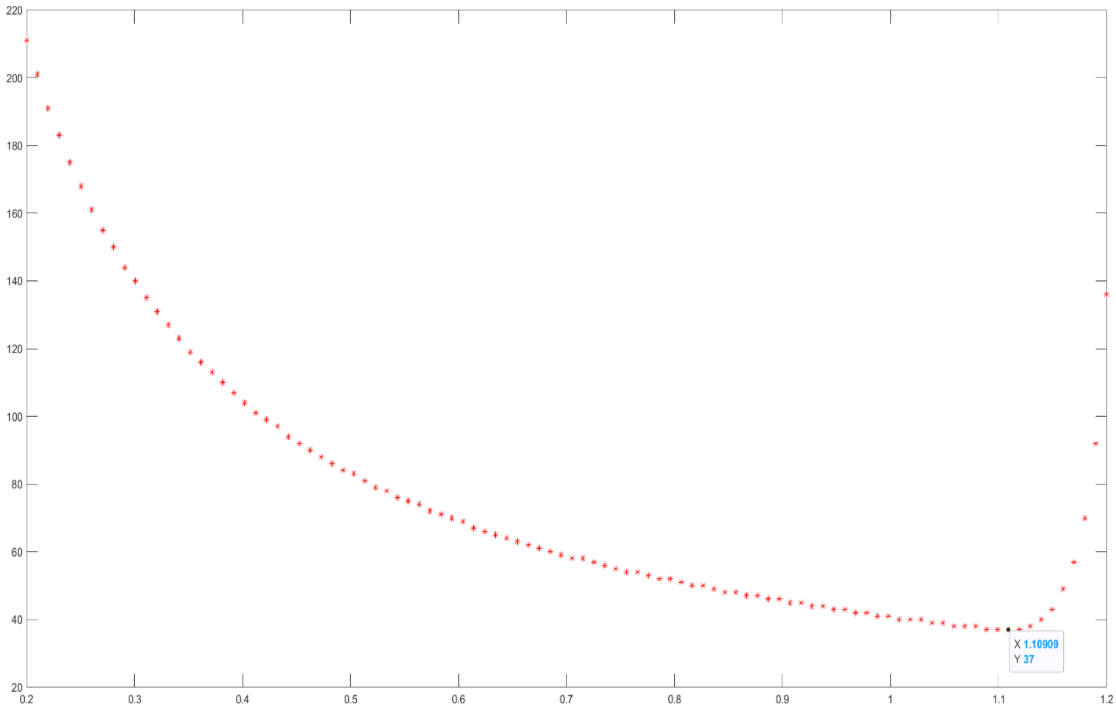


Στην συνέχεια δοκίμασα για λίγο μεγαλύτερο $\boldsymbol{\gamma_k}$ για να παρατηρήσω την συμπεριφορά του αλγορίθμου για να ελέγξω άμα συγκλίνει γρηγορότερα ή πιο αργά για μεγαλύτερα $\boldsymbol{\gamma_k}$. Συνεπώς, δοκίμασα για $\boldsymbol{\gamma_k = 1}$ και το αποτέλεσμα φαίνεται στην παρακάτω γραφική παράσταση, όπου φαίνεται ότι ο αλγόριθμος συγκλίνει σαφώς γρηγορότερα στο ελάχιστο που εμείς αναζητούμε:

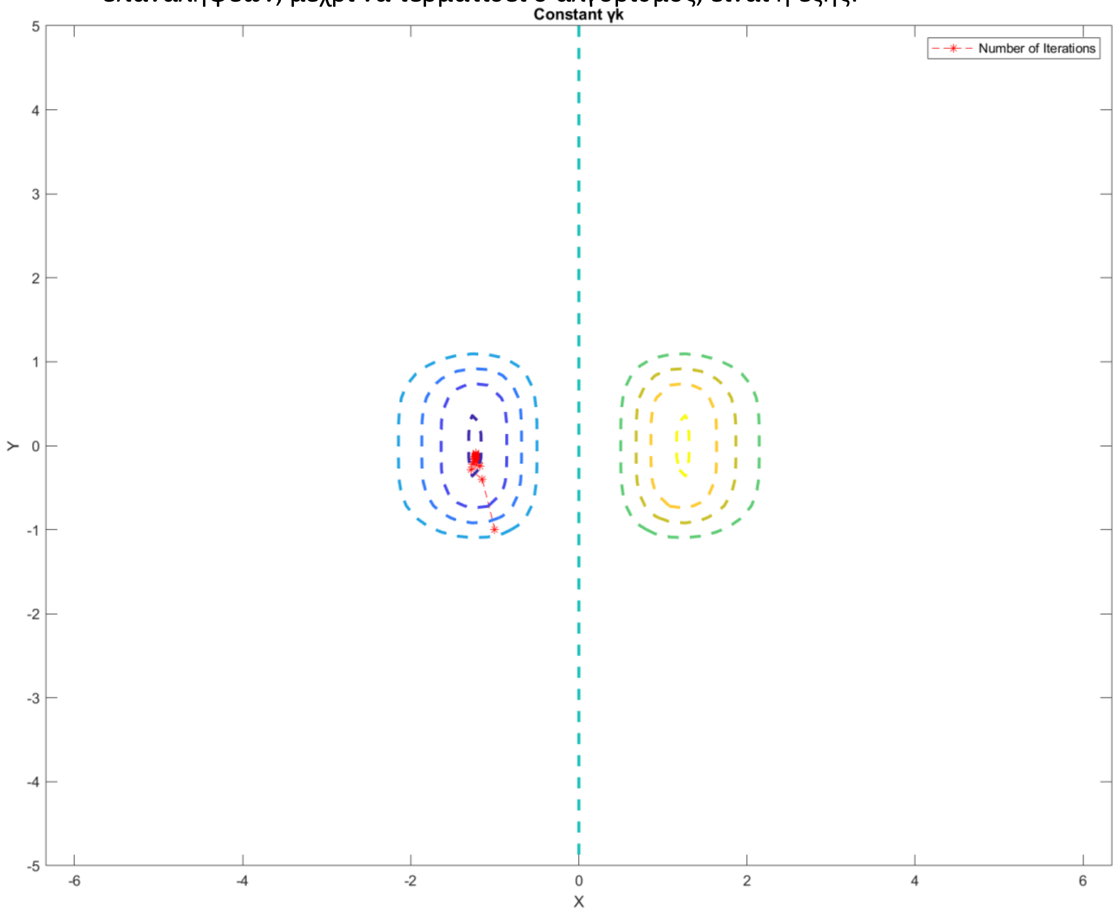


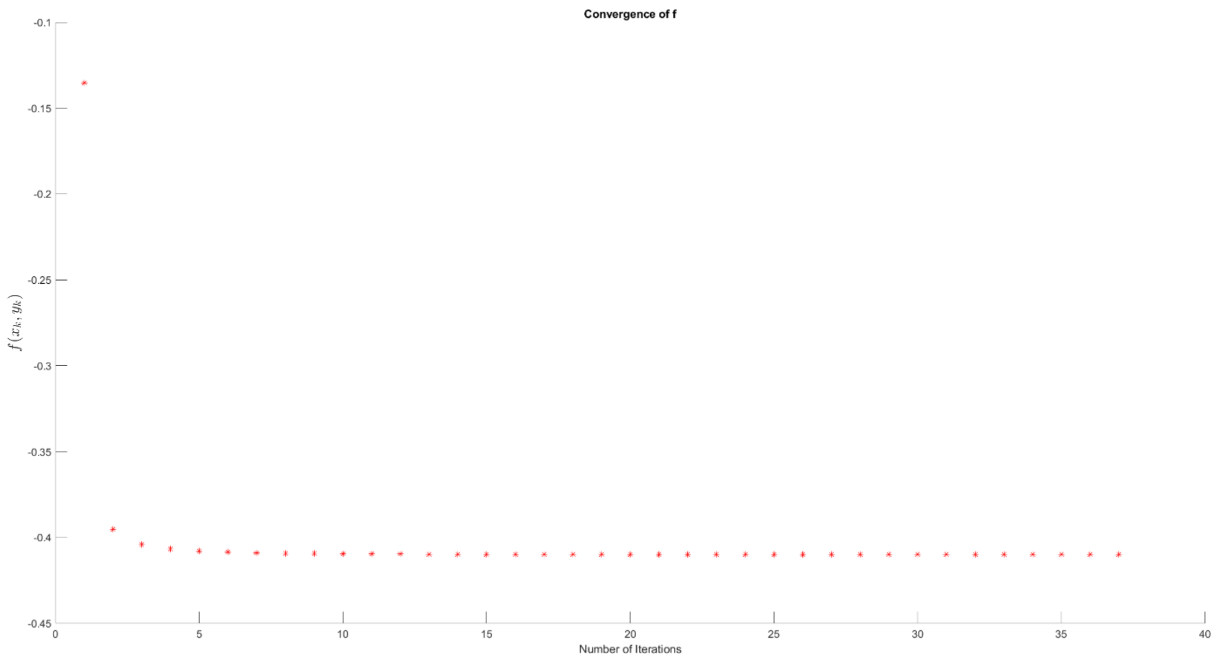


Τέλος, αποφάσισα να βρω για ποιο $\boldsymbol{\gamma_k}$ ο αλγόριθμος μέγιστης καθόδου θα κάνει τις λιγότερες επαναλήψεις για να συγκλίνει από το οποίο φαίνεται ότι αυτό συμβαίνει για $\boldsymbol{\gamma_k = 1.10909}$ με αριθμό επαναλήψεων ίσο με **37.**



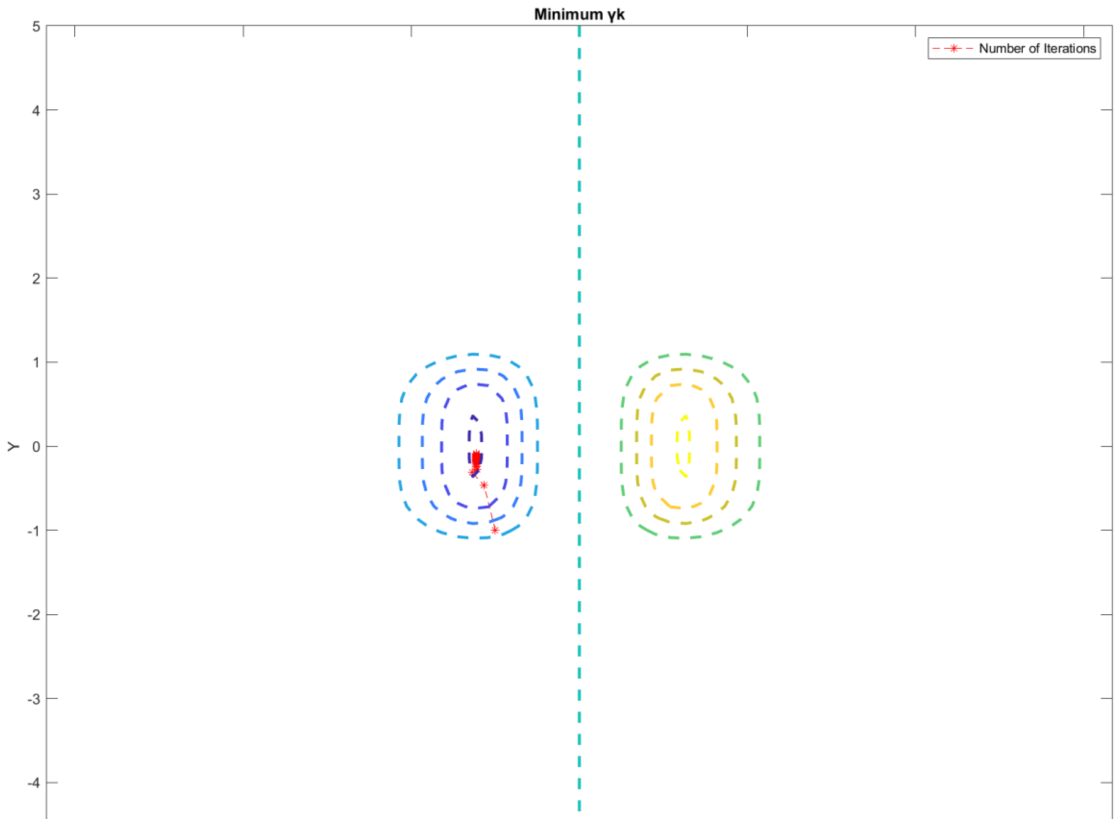
Οπότε για αυτό το $\boldsymbol{\gamma_k = 1.10909}$ προκύπτει ότι η γραφική παράσταση της σύγκλισης της αντικειμενικής συνάρτησης που προέκυψε ως προς τον αριθμό των επαναλήψεων, μέχρι να τερματίσει ο αλγόριθμος, είναι η εξής:

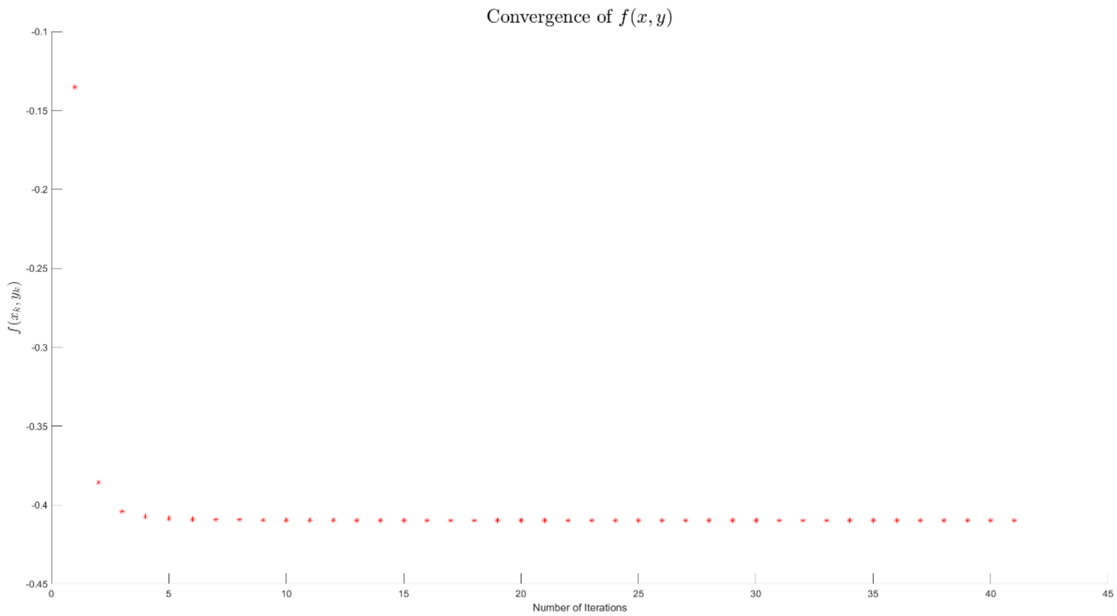




β) Για το ίδιο πάλι σημείο, δηλαδή για το **(-1,-1)** η επιλογή του $\boldsymbol{\gamma_k}$ τέτοια ώστε να ελαχιστοποιεί την $\boldsymbol{f}(\boldsymbol{x_k} + \boldsymbol{\gamma_k} \boldsymbol{d_k})$ γίνεται με την μέθοδο της χρυσής τομής που υλοποιήσαμε στην προηγούμενη εργασία, τροποποιημένη έτσι ώστε αντί να επιστρέφει διάστημα στο οποίο μπορεί να βρίσκεται το ελάχιστο, να επιστρέφει σημείο το οποίο προκύπτει από την μέση τιμή των άκρων του διαστήματος στο οποίο η μέθοδος εντοπίζει το ελάχιστο. Μια παρατήρηση που μπορεί να γίνει για αυτό τον τρόπο σύγκλισης είναι ότι μπορεί να χρειαστεί μεγαλύτερο χρόνο εκτέλεσης καθότι σε κάθε επανάληψη υπολογίζεται το $\boldsymbol{\gamma_k}$ που θα ελαχιστοποιεί την $\boldsymbol{f}(\boldsymbol{x_k} + \boldsymbol{\gamma_k} \boldsymbol{d_k})$**.** Ωστόσο, με αυτό τον τρόπο εσωτερικής βελτιστοποίησης ο αλγόριθμος θα συγκλίνει πιο γρήγορα στο ελάχιστο που αναζητούμε εφόσον κάθε φορά βρίσκει το βέλτιστο $\boldsymbol{\gamma_k}$ για το οποίο θα οδηγούμαστε στον ελάχιστο. Παρακάτω φαίνεται η γραφική της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων μέχρι να τερματίσει ο αλγόριθμος όπου με μεγέθυνση του σχήματος φαίνεται το zik-zak που κάνει ο αλγόριθμος μέχρι να καταλήξει στο ελάχιστο της συνάρτησης:

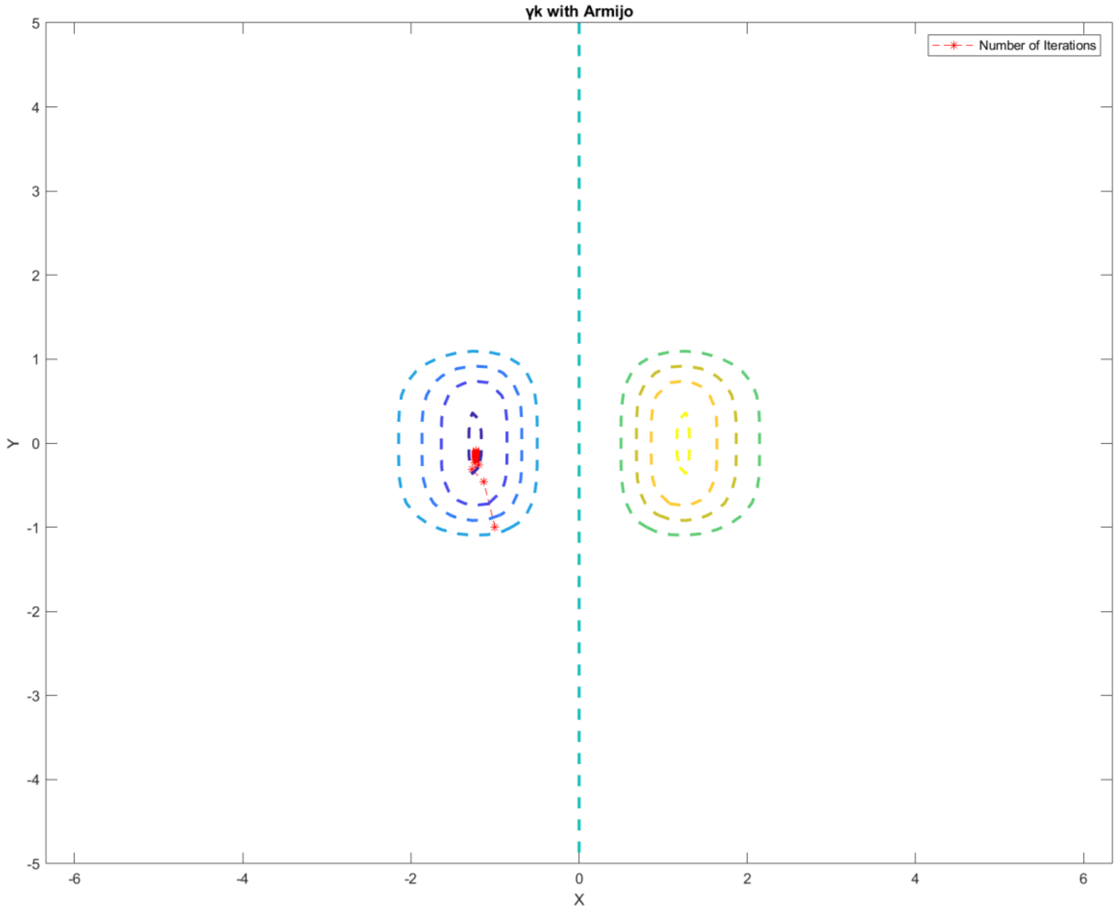
(Ο λόγος που η μέθοδος του χρυσού τομέα λειτουργεί και βρίσκουμε το γ_k που ελαχιστοποιεί την $f(x_k + \gamma_k d_k)$,είναι γιατί από το σημείο (-1,-1) η συνάρτηση είναι αυστηρά σχεδόν κυρτή και επομένως η μέθοδος μπορεί να εφαρμοστεί και θα μας δώσει σωστά αποτελέσματα, ενώ αντίθετα για το σημείο (1,1) η μέθοδος δεν μπορεί να λειτουργήσει διότι η συνάρτηση είναι κοίλη από εκείνο το σημείο και μετέπειτα. Αυτό ισχύει και για την Newton αλλά και για την Levenberg-Marquardt)

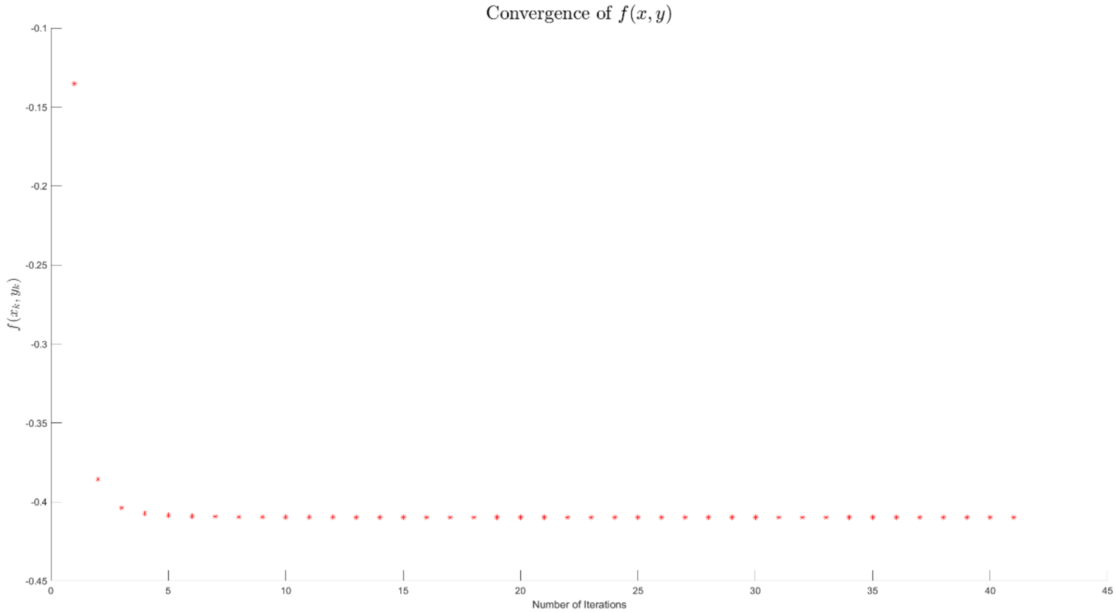




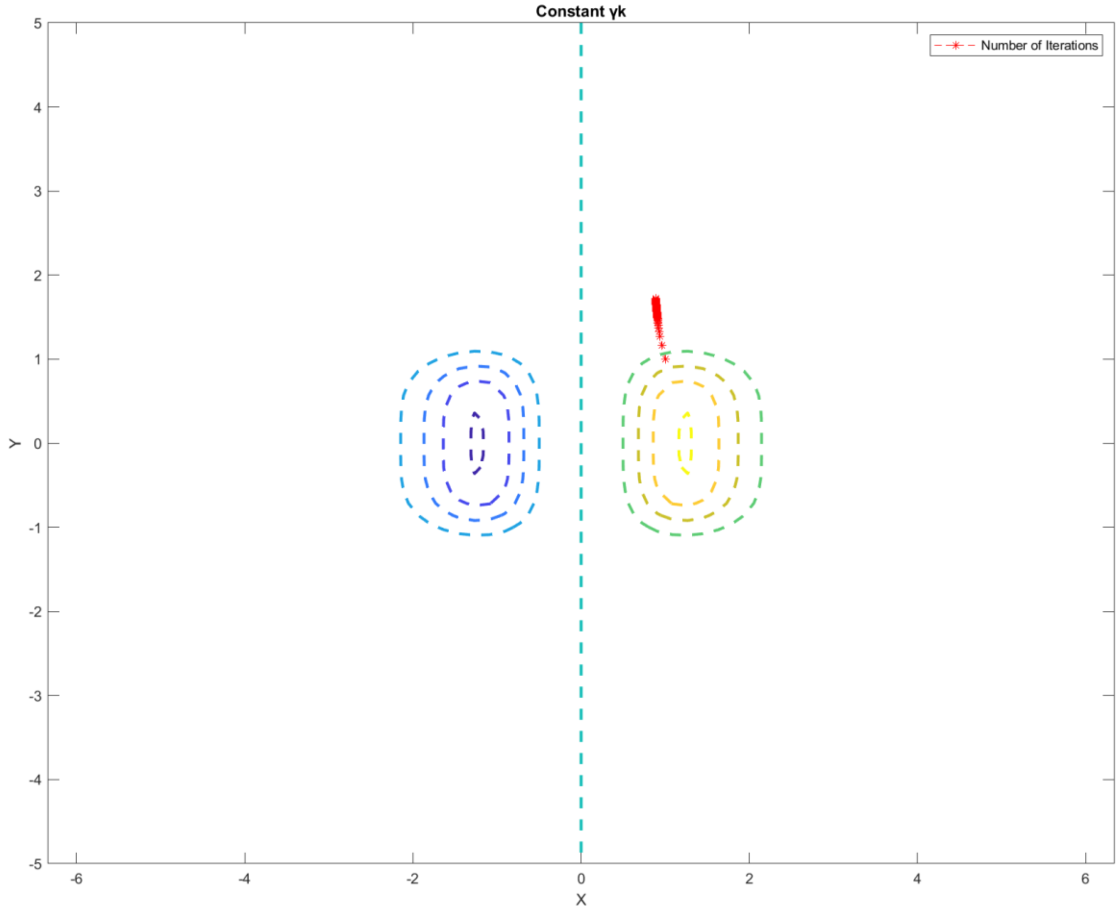
Όπως φαίνεται με αυτό τον τρόπο οδηγούμαστε στο σωστό αποτέλεσμα διότι ο αλγόριθμος συγκλίνει στο ελάχιστο της συνάρτησης με ικανοποιητικό αριθμό επαναλήψεων και με πολύ καλή ακρίβεια, καλύτερη από την περίπτωση του σταθερού $\boldsymbol{\gamma_k}$

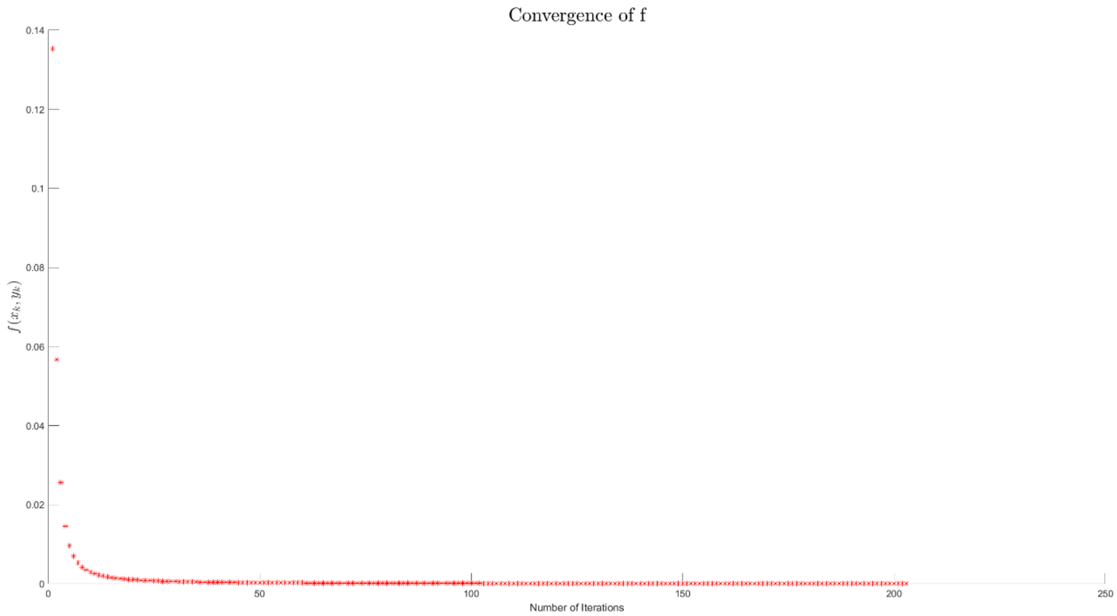
γ) Τέλος, βάσει του **κανόνα Armijo**, με τιμές για τα $\boldsymbol{s, a \text{ και } b}$ με βάση τα διαστήματα που προτείνονται στο βιβλίο επέλεξα $\boldsymbol{s = 1, a = 10^{-3} \text{ και } b = 0.3}$ με τα οποία παρατηρώ ότι παρόλο που το m παραμένει 0 καθ' όλη την διάρκεια της μεθόδου (δηλαδή με αυτές τις επιλογές των $\boldsymbol{s, a \text{ και } b}$ ικανοποιείται συνεχώς το **κριτήριο (4)** το οποίο λέει ότι τα $\boldsymbol{\gamma_k}$ δεν μπορούν να γίνουν πολύ μεγάλα) ο αλγόριθμος συγκλίνει στο ελάχιστο της αντικειμενικής συνάρτησης με τον ίδιο βέβαια ρυθμό με την περίπτωση του σταθερού $\boldsymbol{\gamma_k}$ το οποίο είναι λογικό αφού δεν μεταβάλλεται καθόλου κατά την διάρκεια της μεθόδου. Παρακάτω φαίνεται η γραφική παράσταση σύγκλισης ως προς τον αριθμό των επαναλήψεων όπου με μεγέθυνση του σχήματος φαίνεται το zik-zak που κάνει ο αλγόριθμος μέχρι να καταλήξει στο ελάχιστο της συνάρτησης





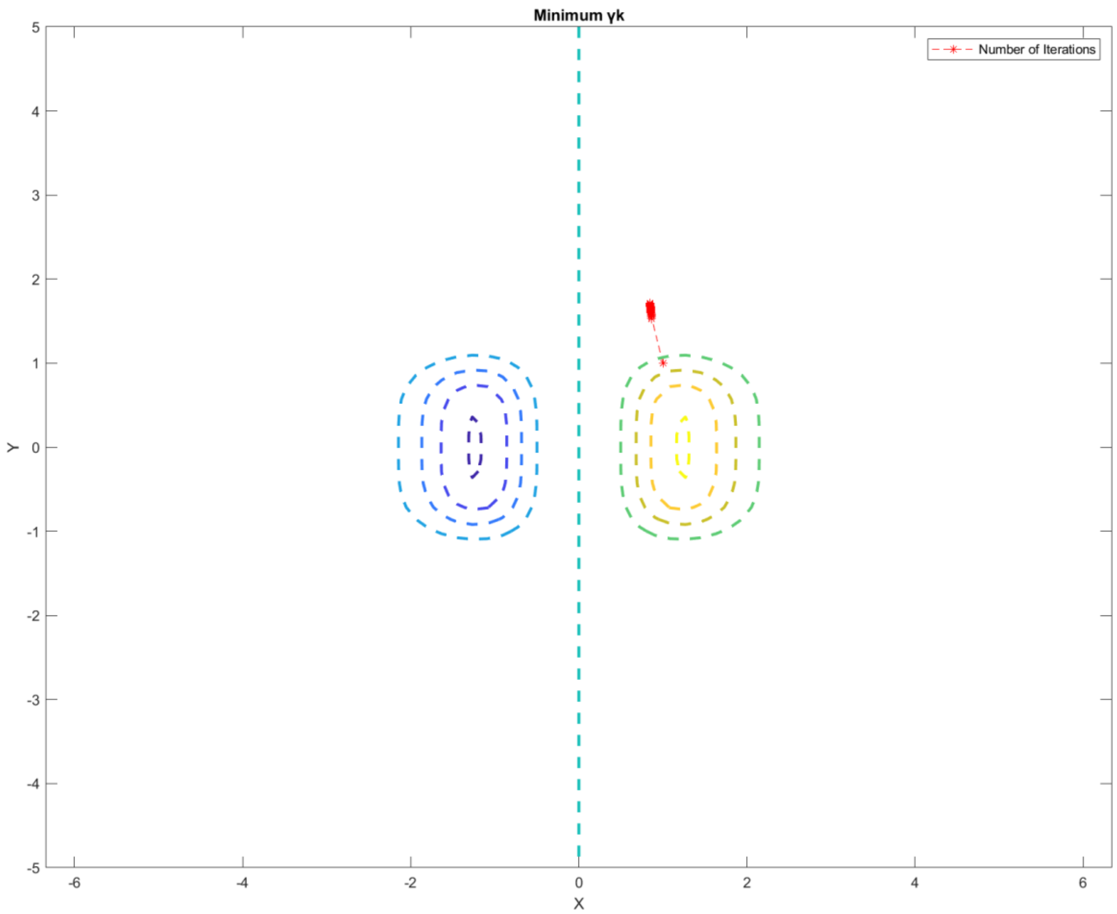
iii. **α)** Για το σημείο **(1,1)** παρατηρούμε ότι ο αλγόριθμος της μέγιστης καθόδου δεν θα συγκλίνει ποτέ στο ελάχιστο της αντικειμενικής συνάρτησης και αυτό διότι γιατί το σημείο από το οποίο επιλέγουμε να ξεκινήσουμε είναι πολύ κοντά σε ένα σαγματικό σημείο της συνάρτησης (κοντά στο 0) στο οποίο πληρείται η συνθήκη τερματισμού του αλγορίθμου δηλαδή $|\boldsymbol{\nabla f(x_k)}| \leq \boldsymbol{\varepsilon}$**.** Επομένως, ο αλγόριθμος θα «ταλαντώνεται» συνεχώς γύρω από αυτό το σημείο χωρίς ποτέ να καταλήξει στο ελάχιστο της συνάρτησης μέχρι να ισχύσει η συνθήκη τερματισμού ή μέχρι να ξεπεράσει τον αριθμό των επιτρεπόμενων επαναλήψεων που έχουμε δώσει στην συνάρτηση. Παρακάτω φαίνεται η γραφική παράσταση σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων από την οποία είναι φανερός ο εγκλωβισμός του αλγορίθμου σε μια ισοβαρή καμπύλη πολύ κοντά στο μηδέν(το $\boldsymbol{\gamma_k}$ επιλέχθηκε σταθερό και ίσο με **0.3** για την συγκεκριμένη γραφική παράσταση)

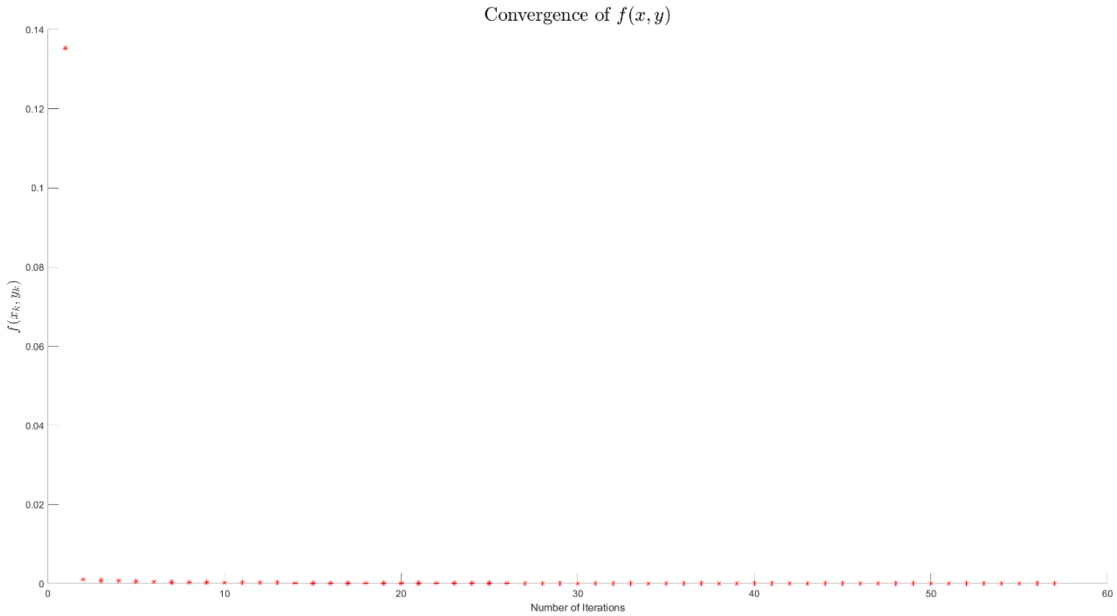




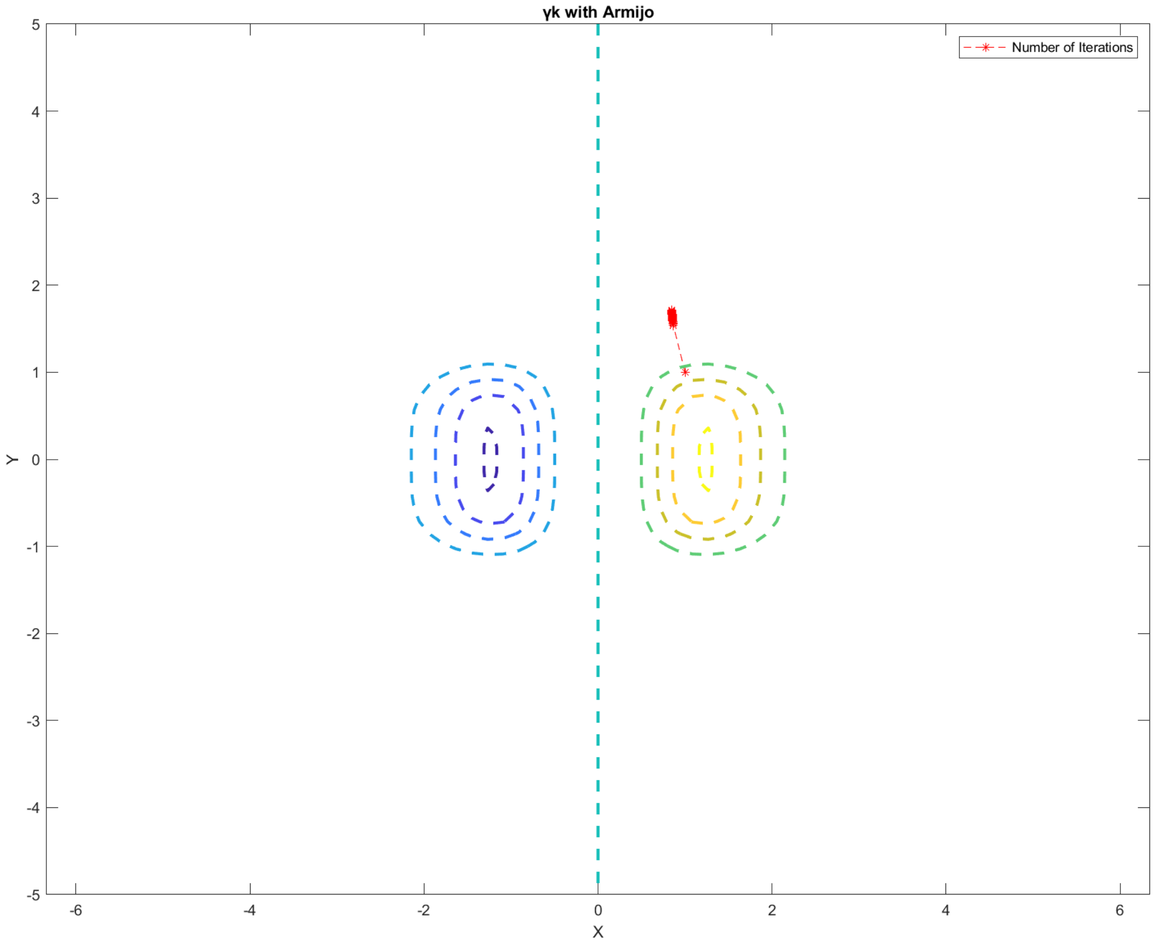
β) Για το ίδιο το σημείο, δηλαδή για το **(1,1),** με επιλογή του $\boldsymbol{\gamma_k}$ τέτοια ώστε να ελαχιστοποιείται η $\boldsymbol{f}(\boldsymbol{x_k} + \boldsymbol{\gamma_k d_k})$, παρατηρούμε, όπως και ήταν αναμενόμενο, ότι ο αλγόριθμος εξακολουθεί να εγκλωβίζεται σε μία ισοβαρή καμπύλη πολύ κοντά στο μηδέν και δεν παίζει ρόλο που το $\boldsymbol{\gamma_k}$ είναι σε κάθε επανάληψη το βέλτιστο. Από αυτά μπορούμε να συμπεράνουμε πόσο **«κακή»** είναι η επιλογή αυτού του σημείου ως αρχικό σημείο αναζήτησης καθότι δεν θα μας οδηγήσει ποτέ στο ελάχιστο το οποίο εμείς ψάχνουμε να βρούμε. Στην συνέχεια, ακολουθεί η σχετική γραφική παράσταση στην οποία φαίνεται, όπως και προηγουμένως , η «ταλάντωση» που κάνει ο αλγόριθμος μέχρι να ισχύσει η συνθήκη τερματισμού του $|\boldsymbol{\nabla f}(\boldsymbol{x_k})| \leq \boldsymbol{\varepsilon}$ η οποία όμως απαιτεί λιγότερο αριθμό επαναλήψεων καθότι έχουμε το βέλτιστο $\boldsymbol{\gamma_k}$ σε κάθε επανάληψη.

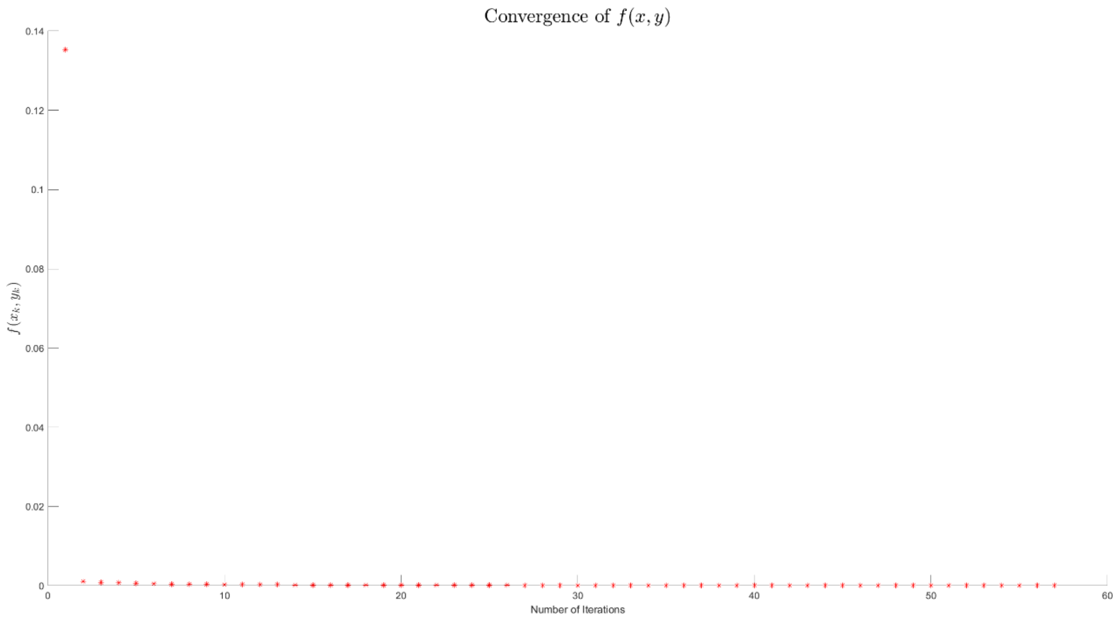
Σημείωση : Ο υπολογισμός του γ_k έγινε πάλι με την τροποποιημένη μέθοδο της χρυσής τομής όπως δηλαδή και για το σημείο (-1,-1).





γ) Τέλος, ο προσδιορισμός του $\boldsymbol{\gamma_k}$ με αρχικό σημείο αναζήτησης το **(1,1)** και με τον **κανόνα Armijo** μας δίνει παρόμοια αποτελέσματα με τα αποτελέσματα που μας έδωσαν και η επιλογή σταθερού $\boldsymbol{\gamma_k}$ αλλά και επιλογή του$\boldsymbol{\gamma_k}$ βάσει της ελαχιστοποίησης της συνάρτησης $\boldsymbol{f(x_k + \gamma_k d_k)}$. Και σε αυτή την περίπτωση επιλέχθηκαν τα ίδια $\boldsymbol{s, a\ και\ b}$ με την περίπτωση του σημείου **(1, 1)** δηλαδή $\boldsymbol{s = 1, a = 10^{-3}\ και\ b = 0.3}$ **.** Παρακάτω φαίνεται η γραφική παράσταση της σύγκλισης της συνάρτησης ως προς τον αριθμό των επαναλήψεων όπου φαίνεται ξεκάθαρα ότι ο αλγόριθμος εγκλωβίζεται για ακόμα μία φορά σε μια ισοβαρή καμπύλη πολύ κοντά στο μηδέν και δεν μπορεί να βρει το ελάχιστο της αντικειμενικής συνάρτησης, δείχνοντας μας για ακόμα μια φορά το πόσο «κακή» είναι η επιλογή αυτού του σημείου ως αρχικό σημείο αναζήτησης.



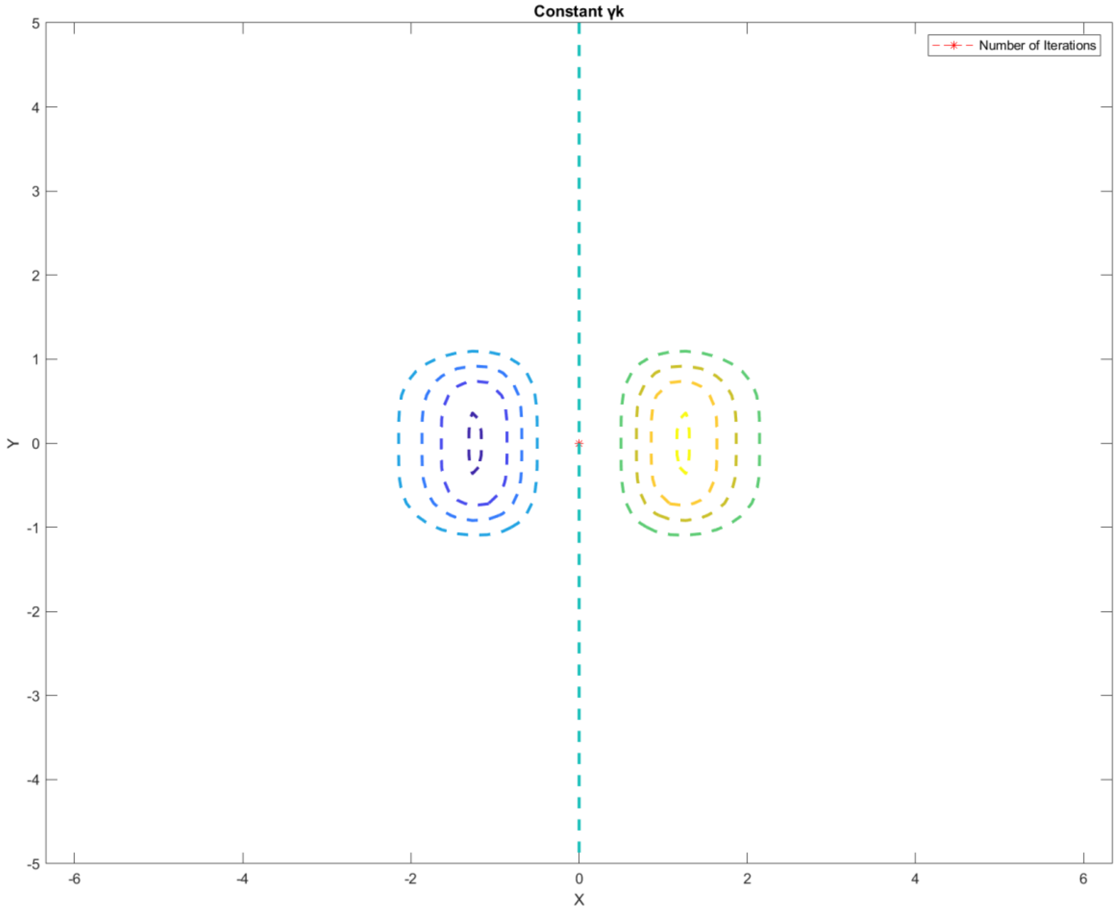


Θέμα 3ο :

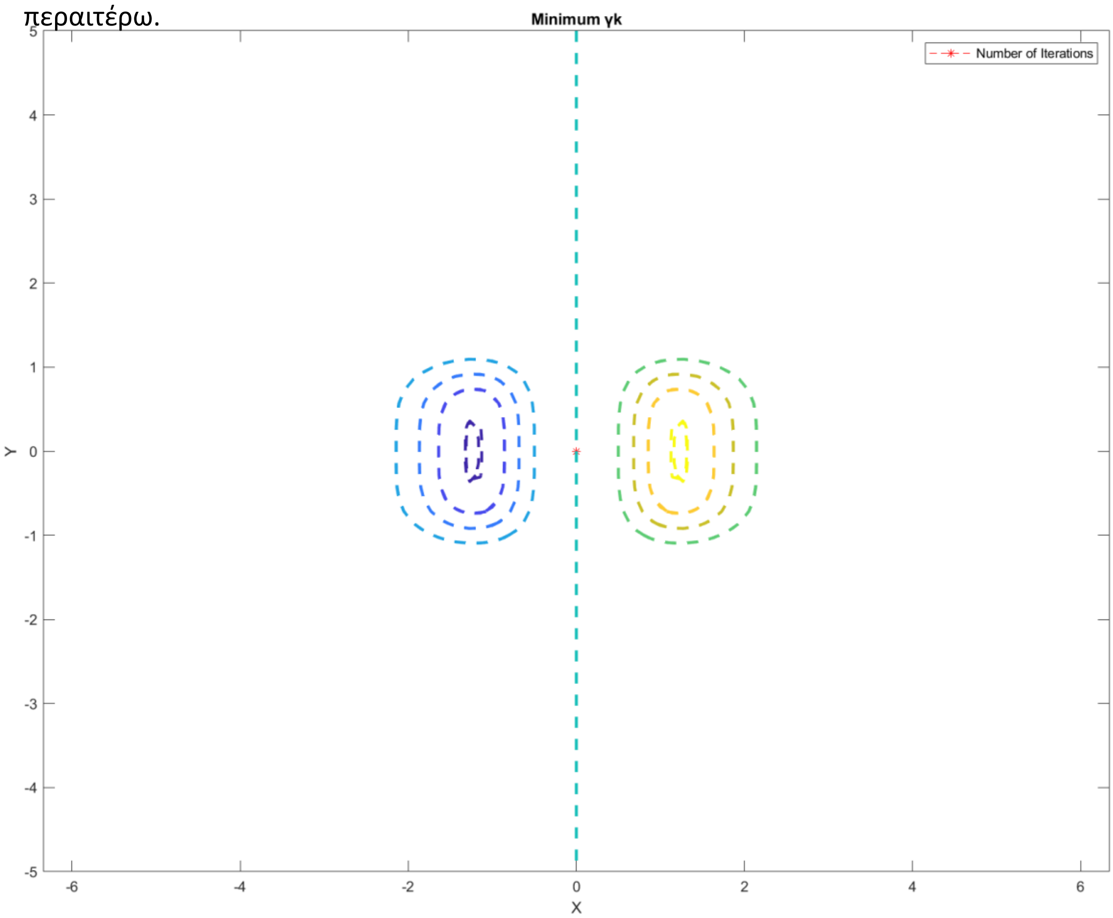
Μέθοδος Newton:

Η μέθοδος **Newton** είναι ένας από τους τρόπους για τον προσδιορισμό του ελαχίστου μιας συνάρτησης η οποία ικανοποιεί την ιδιότητα της μέγιστης καθόδου δηλαδή $\forall \boldsymbol{k} \in \mathbb{N}$ $\boldsymbol{ισχύει}$ $\boldsymbol{f}(\boldsymbol{x_k}) > \boldsymbol{f}(\boldsymbol{x_{k+1}})$ και έχει ως διάνυσμα κατεύθυνσης αναζήτησης (δηλαδή την κατεύθυνση στην οποία αναζητούμε το ελάχιστο) το $\boldsymbol{d_k} = -[\nabla^2 \boldsymbol{f}(\boldsymbol{x_k})]^{-1} \nabla \boldsymbol{f}(\boldsymbol{x_k})$ όπου ο πίνακας $[\nabla^2 \boldsymbol{f}(\boldsymbol{x_k})]^{-1}$ είναι ο αντίστροφος Εσσιανός πίνακας της αντικειμενικής συνάρτησης $\boldsymbol{f}(\boldsymbol{x_k})$. Η μέθοδος αυτή για να δουλέψει απαιτεί να είναι θετικά ορισμένος ο Εσσιανός πίνακας και να είναι αντιστρέψιμος(γεγονός που εξασφαλίζεται διότι αφού ο Εσσιανός πίνακας είναι συμμετρικός άμα είναι και θετικά ορισμένος θα έχει σίγουρα αντίστροφο). Ωστόσο, παρατηρούμε ότι για όλα τα σημεία αναζήτησης που μας δίνονται για την εύρεση του ελαχίστου ο Εσσιανός δεν είναι θετικά ορισμένος σε αυτά τα σημεία οπότε κανονικά η μέθοδος δεν μπορεί να εφαρμοστεί διότι δεν είναι σωστή. Παρόλα αυτά, δοκιμάζω να την εφαρμόσω για να δω τα αποτελέσματα που θα προκύψουν, ωστόσο η σύγκριση της με τις άλλες μεθόδους καθώς επίσης και ο ρυθμός σύγκλισης της στο ελάχιστο δεν έχουν πολύ νόημα αφού **η μέθοδος δεν λειτουργεί**.

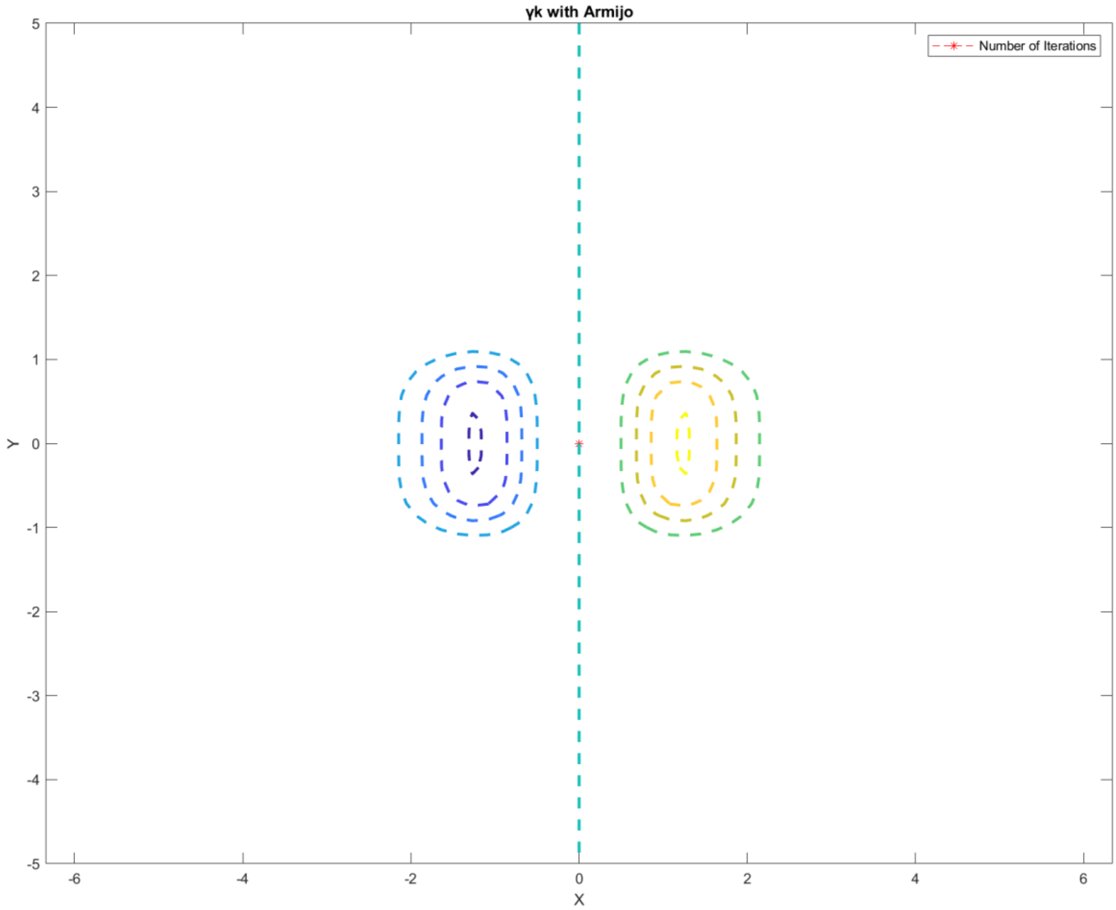
i. **α)** Για το σημείο **(0,0)** δεν έχει σημασία η επιλογή του $\boldsymbol{\gamma_k}$ καθότι σε εκείνο το σημείο όπως φαίνεται και από την γραφική παράσταση της αντικειμενικής συνάρτησης ισχύει ότι $\nabla \boldsymbol{f}(\boldsymbol{x_k}) = \boldsymbol{0}$ είναι δηλαδή ένα σαγματικό σημείο, οπότε η μέθοδος «παγιδεύεται» και δεν μπορεί να προχωρήσει περαιτέρω με αποτέλεσμα να μην μας δίνει το ελάχιστο το οποίο εμείς αναζητούμε. Επίσης ο πίνακας $\nabla^2 \boldsymbol{f}(\boldsymbol{x_k}) = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}$ είναι ο μηδενικός οπότε προφανώς δεν θα υπάρχει ο αντίστροφος του. Παρακάτω φαίνεται το γράφημα της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων. **(** Η γραφική παράσταση έγινε για σταθερό $\boldsymbol{\gamma_k = 0.7}$**)**



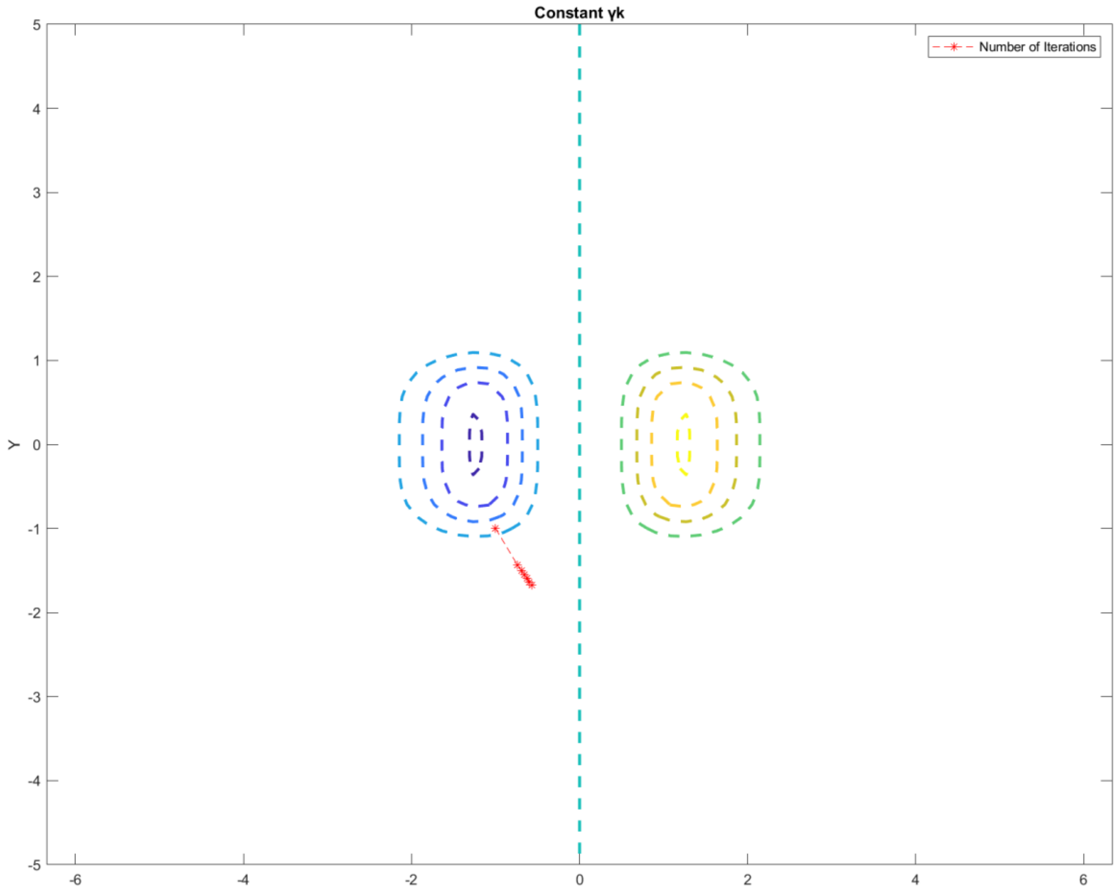
β) Ομοίως, για το ίδιο πάλι σημείο, η επιλογή του $\boldsymbol{\gamma_k}$ τέτοια ώστε να ελαχιστοποιεί την $\boldsymbol{f(x_k + \gamma_k d_k)}$ δεν έχει κάποιο νόημα διότι ισχύει $\boldsymbol{\nabla f(x_k) = 0}$ οπότε όπως και πριν η μέθοδος θα «παγιδεύεται» στο σαγματικό σημείο και δεν θα μπορεί να προχωρήσει περαιτέρω.

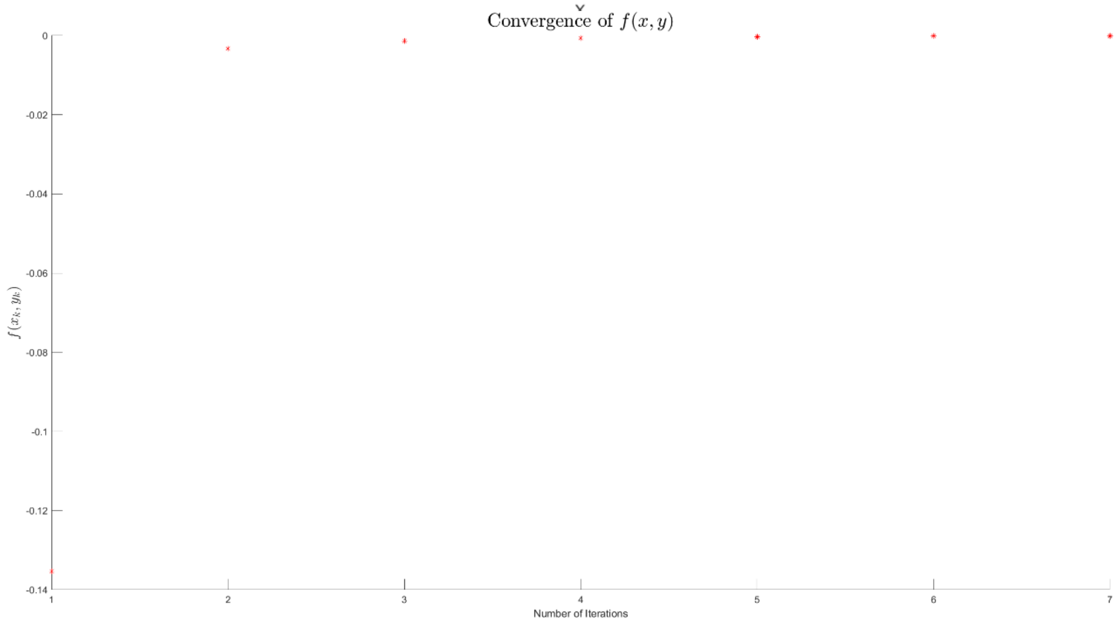


γ) Τέλος, όπως και με τους δύο προηγούμενους τρόπους έτσι και με τον **κανόνα Armijo** δεν θα παρατηρήσουμε κάποια μεταβολή. Με $s = 1, a = 10^{-3}$ $και$ $b = 0.3$



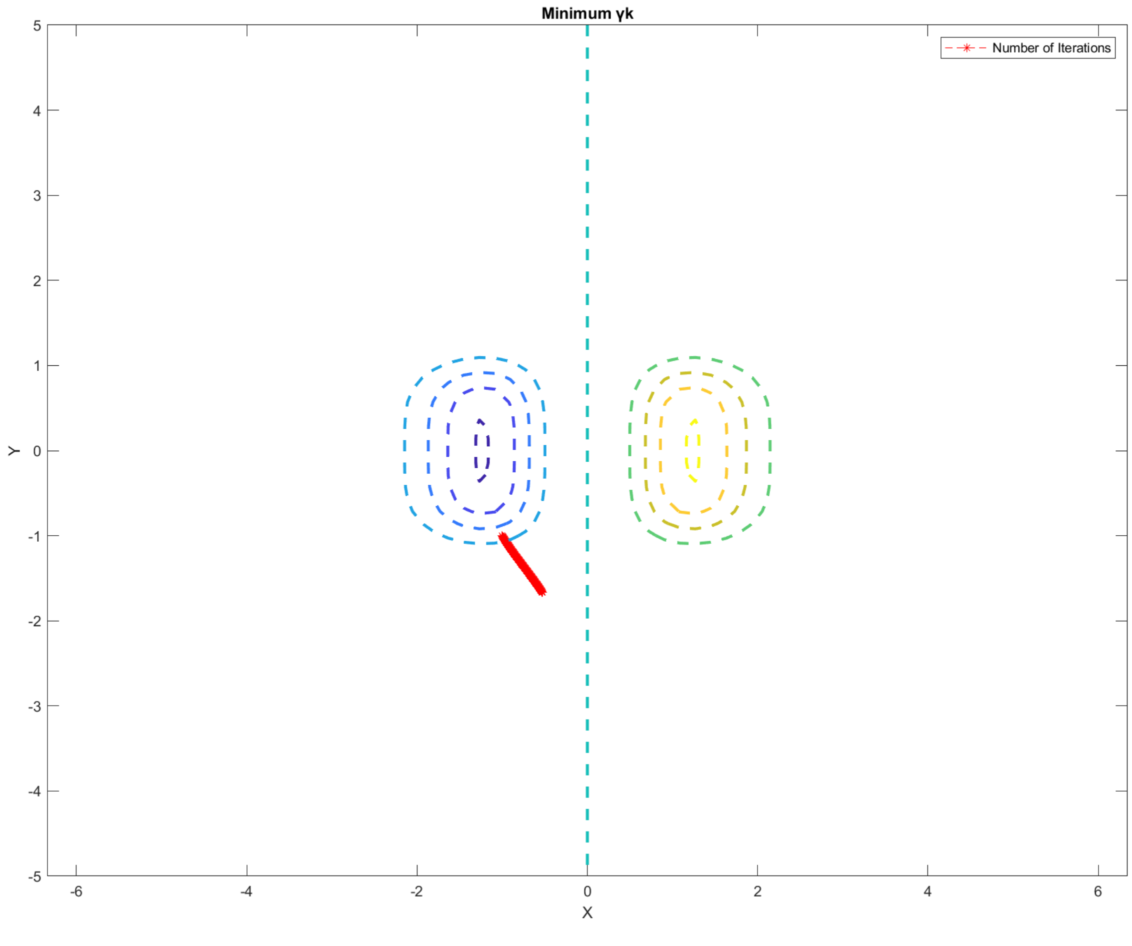
ii. **α)** Για το σημείο **(-1,-1)** παρατηρούμε όπως αναφέρθηκε και στην αρχή ότι ο Εσσιανός πίνακας $\boldsymbol{\nabla^2 f(x_k)}$ δεν θα είναι θετικά ορισμένος οπότε ούτε και για αυτό το αρχικό σημείο αναζήτησης έχει νόημα να εφαρμοστεί η μέθοδος **Newton.** Ωστόσο, αμελώ προς το παρόν την απαίτηση ο Εσσιανός να είναι πάντοτε θετικά ορισμένος προκειμένου να δω ποια θα είναι η συμπεριφορά του αλγορίθμου. Παρακάτω φαίνεται η γραφική παράσταση της σύγκλισης της αντικειμενικής συνάρτησης που προέκυψε ως προς τον αριθμό των επαναλήψεων, μέχρι να τερματίσει ο αλγόριθμος, η οποία είναι η εξής **(**Η γραφική παράσταση έγινε για σταθερό $\boldsymbol{\gamma_k = 0.7}$**)**:

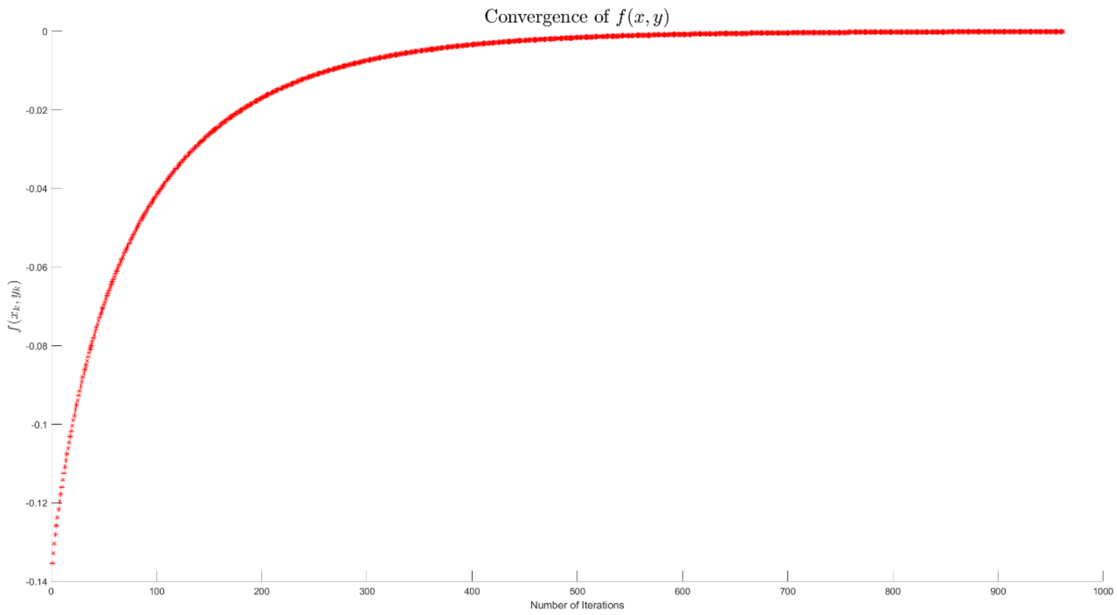




Όπως φαίνεται ο αλγόριθμος δεν μας δίνει σωστά αποτελέσματα κάτι που ήταν αναμενόμενο αφού η μέθοδος κανονικά δεν μπορεί να εφαρμοστεί διότι ο πίνακας $\boldsymbol{\nabla^2 f(x_k)}$ δεν είναι θετικά ορισμένος για το αρχικό σημείο αναζήτησης.

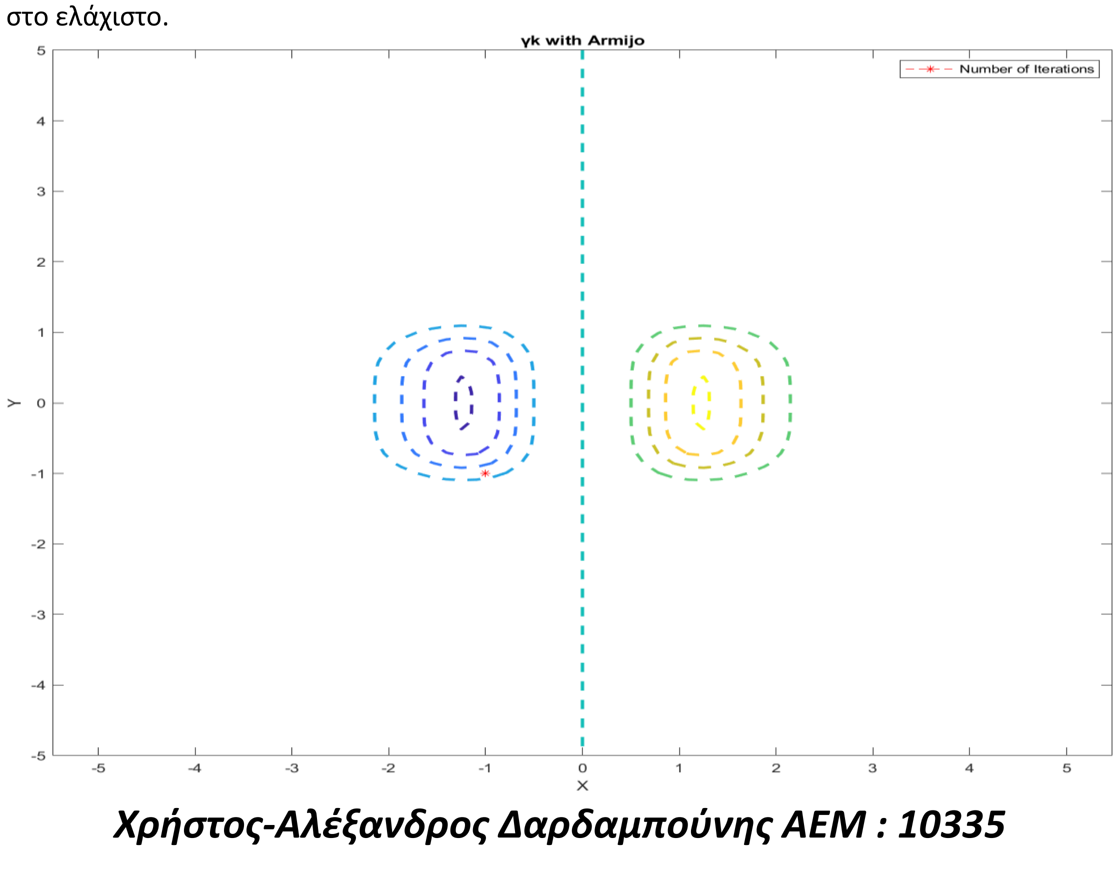
β) Για το ίδιο πάλι σημείο, δηλαδή για το **(-1,-1)** η επιλογή του $\boldsymbol{\gamma_k}$ τέτοια ώστε να ελαχιστοποιεί την $\boldsymbol{f(x_k + \gamma_k d_k)}$ γίνεται με την μέθοδο της χρυσής τομής που υλοποιήσαμε στην προηγούμενη εργασία, τροποποιημένη έτσι ώστε αντί να επιστρέφει διάστημα στο οποίο μπορεί να βρίσκεται το ελάχιστο, να επιστρέφει σημείο το οποίο προκύπτει από την μέση τιμή των άκρων του διαστήματος στο οποίο η μέθοδος εντοπίζει το ελάχιστο. Ωστόσο, όπως αναφέραμε και προηγουμένως ο Εσσιανός πίνακας δεν θα είναι θετικά ορισμένος οπότε από αυτό που λαμβάνουμε δεν έχει κάποιο νόημα διότι δεν μπορεί να εφαρμοστεί η μέθοδος. Παρακάτω φαίνεται η γραφική της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων μέχρι να τερματίσει ο αλγόριθμος:

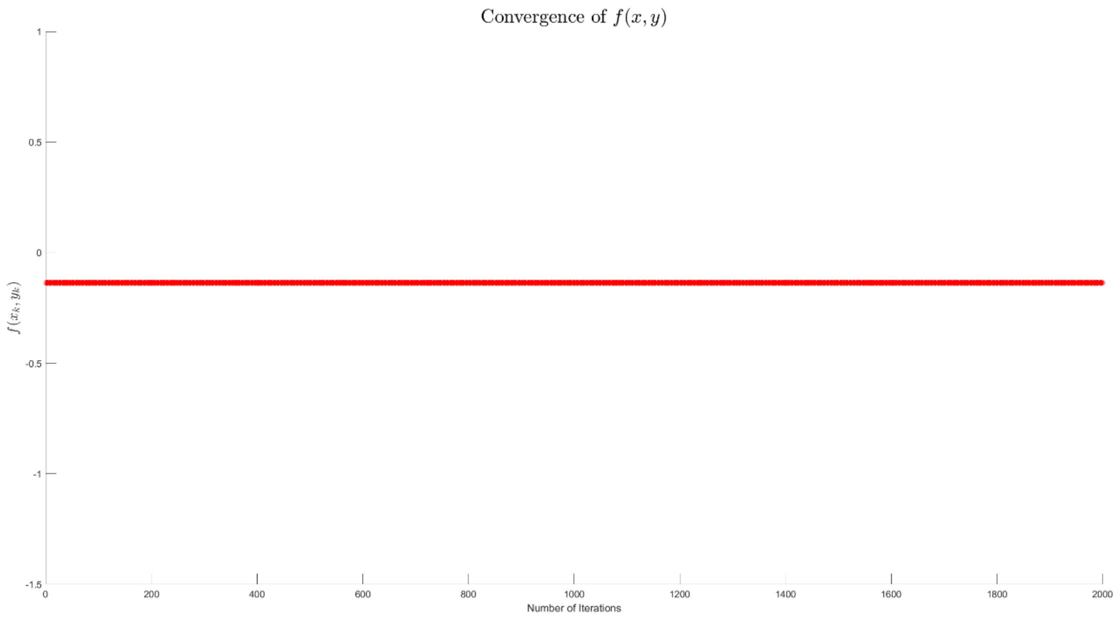




Από το οποίο φαίνεται ότι η μέθοδος κάνει πολλές επαναλήψεις ψάχνοντας το βέλτιστο $\boldsymbol{\gamma_k}$ τέτοιο ώστε να ελαχιστοποιεί την $\boldsymbol{f(x_k + \gamma_k d_k)}$ χωρίς όμως να καταφέρνει να συγκλίνει στο σημείο ισορροπίας.

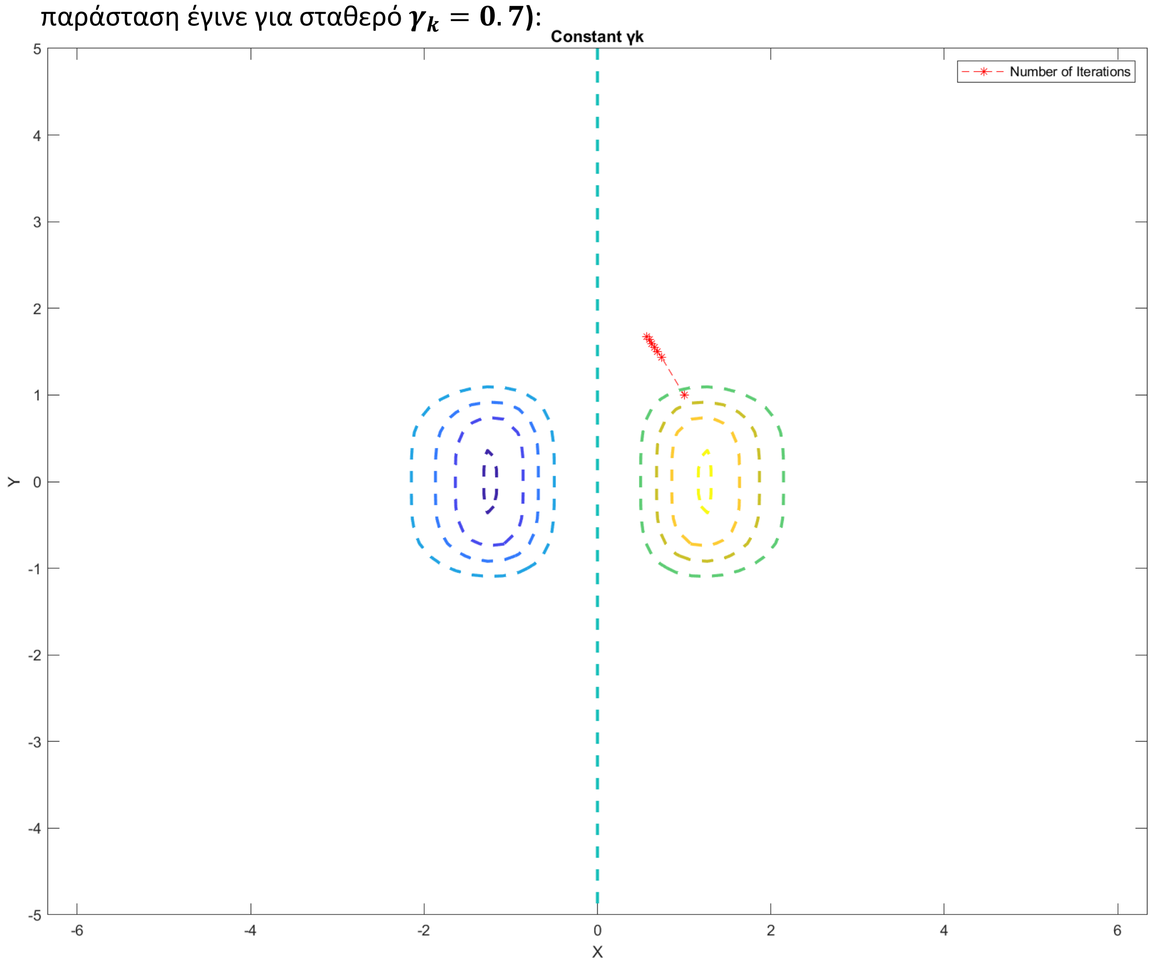
γ) Τέλος, ο προσδιορισμός του $\boldsymbol{\gamma_k}$ με αρχικό σημείο αναζήτησης το **(-1,-1)** και με τον **κανόνα Armijo** δεν μας δίνει παρόμοια αποτελέσματα με τα αποτελέσματα που μας έδωσαν και η επιλογή σταθερού $\boldsymbol{\gamma_k}$ αλλά και επιλογή του $\boldsymbol{\gamma_k}$ βάσει της ελαχιστοποίησης της συνάρτησης $\boldsymbol{f(x_k + \gamma_k d_k)}$ καθότι δεν μπορεί να εκτελεστεί καθόλου διότι η μέθοδος παγιδεύεται σε έναν ατέρμονα βρόχο . Και σε αυτή την περίπτωση επιλέχθηκαν τα ίδια $\boldsymbol{s, a \; και \; b}$ με την περίπτωση του σημείου **(0,0)** δηλαδή $\boldsymbol{s = 1, a = 10^{-3} \; και \; b = 0.3}$ **.** Παρακάτω φαίνεται η γραφική παράσταση της σύγκλισης της συνάρτησης ως προς τον αριθμό των επαναλήψεων όπου φαίνεται ξεκάθαρα ότι ο αλγόριθμος δεν θα μας οδηγήσει στο ελάχιστο.

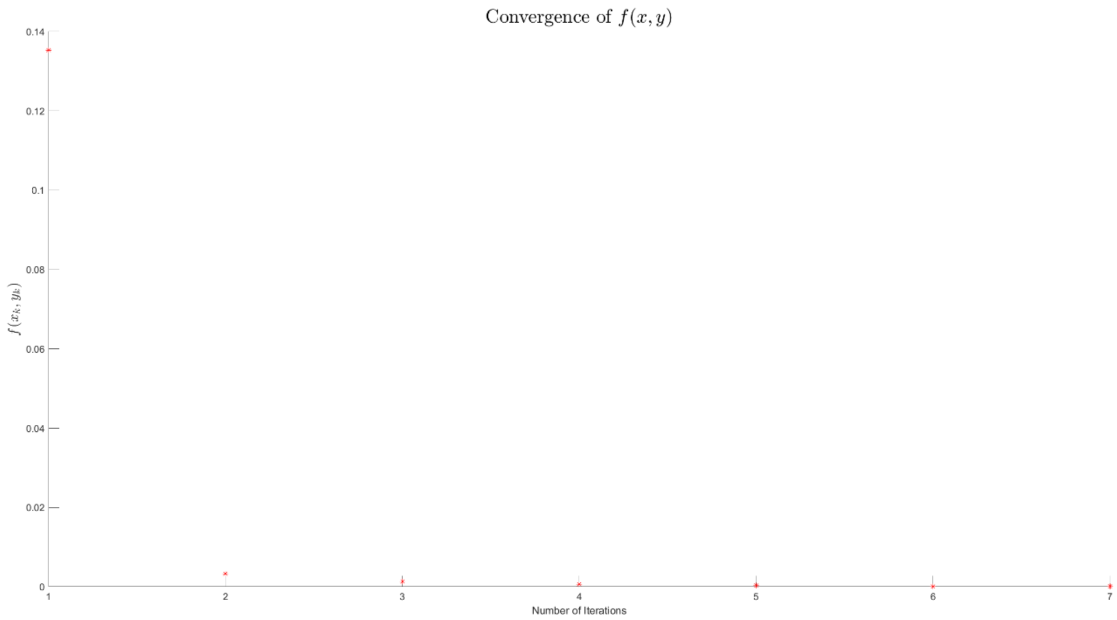




Όπως ήταν και εδώ αναμενόμενο η μέθοδος δεν μπορεί να εφαρμοστεί και άρα δεν μας δίνει σωστά αποτελέσματα.

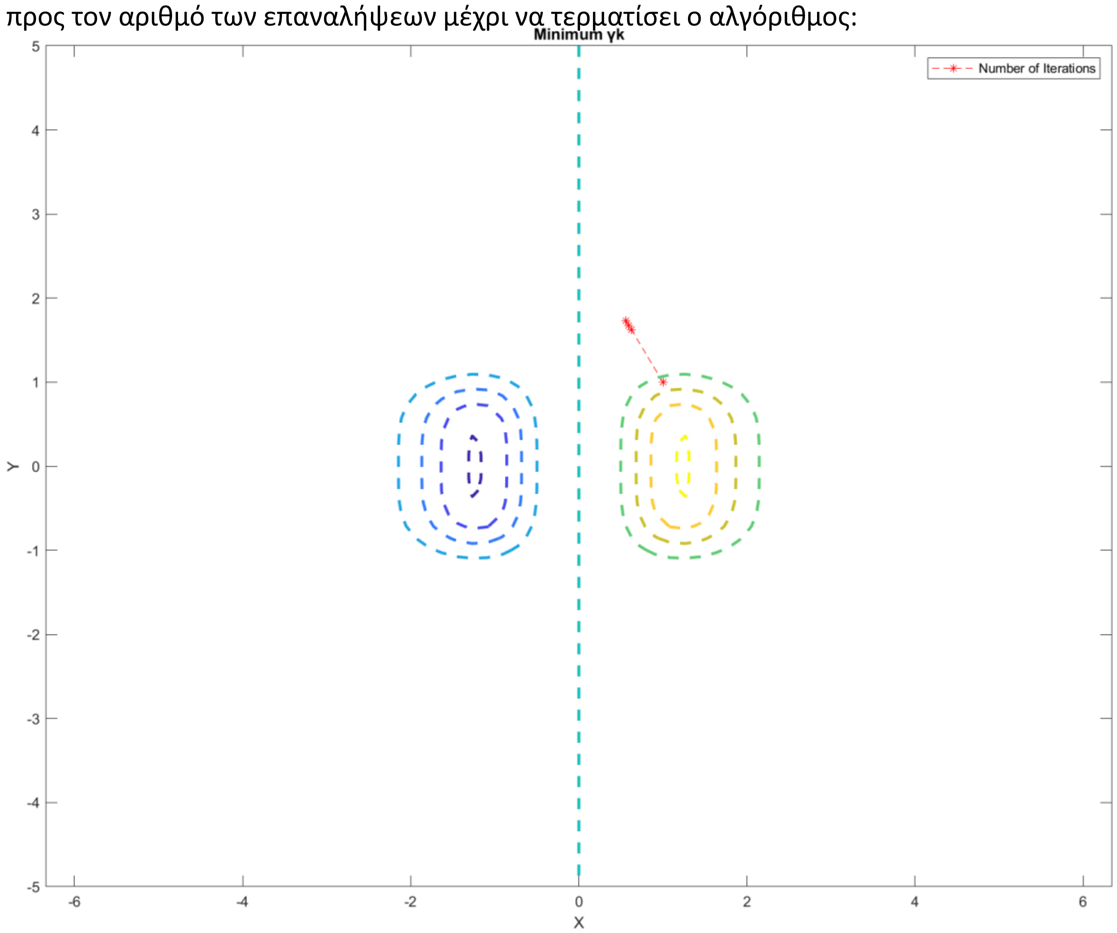
α) Για το σημείο **(1, 1)** παρατηρούμε όπως αναφέρθηκε και στην αρχή ότι ο Εσσιανός πίνακας $\nabla^2 f(x_k)$ δεν θα είναι θετικά ορισμένος οπότε ούτε και για αυτό το αρχικό σημείο αναζήτησης έχει νόημα να εφαρμοστεί η μέθοδος **Newton.** Ωστόσο, αμελώ προς το παρόν την απαίτηση ο Εσσιανός να είναι πάντοτε θετικά ορισμένος προκειμένου να δω ποια θα είναι η συμπεριφορά του αλγορίθμου. Παρακάτω φαίνεται η γραφική παράσταση της σύγκλισης της αντικειμενικής συνάρτησης που προέκυψε ως προς τον αριθμό των επαναλήψεων, μέχρι να τερματίσει ο αλγόριθμος, η οποία είναι η εξής **(**Η γραφική παράσταση έγινε για σταθερό $\boldsymbol{\gamma_k = 0.7}$**)**:

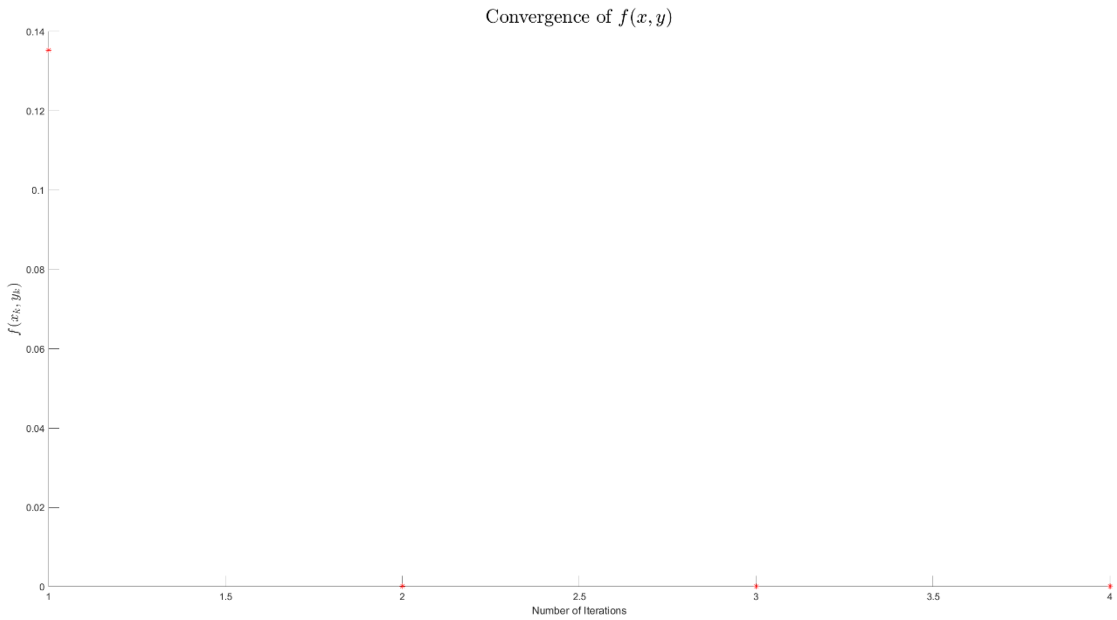




Όπως φαίνεται ο αλγόριθμος δεν μας δίνει σωστά αποτελέσματα κάτι που ήταν αναμενόμενο αφού η μέθοδος κανονικά δεν μπορεί να εφαρμοστεί διότι ο πίνακας $\nabla^2 f(x_k)$ δεν είναι θετικά ορισμένος για το αρχικό σημείο αναζήτησης.

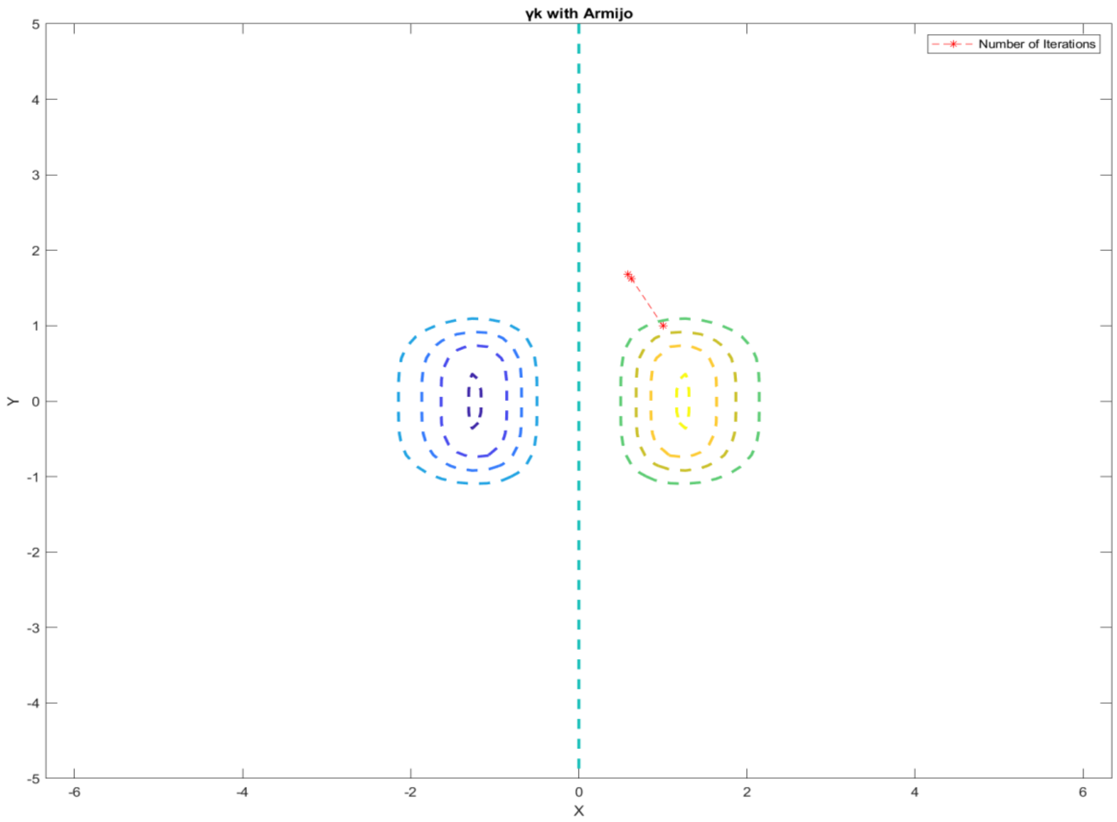
β) Για το ίδιο πάλι σημείο, δηλαδή για το **(1,1)** η επιλογή του γ_k τέτοια ώστε να ελαχιστοποιεί την $f(x_k + \gamma_k d_k)$ γίνεται με την μέθοδο της χρυσής τομής που υλοποιήσαμε στην προηγούμενη εργασία, τροποποιημένη έτσι ώστε αντί να επιστρέφει διάστημα στο οποίο μπορεί να βρίσκεται το ελάχιστο, να επιστρέφει σημείο το οποίο προκύπτει από την μέση τιμή των άκρων του διαστήματος στο οποίο η μέθοδος εντοπίζει το ελάχιστο. Ωστόσο, όπως αναφέραμε και προηγουμένως ο Εσσιανός πίνακας δεν θα είναι θετικά ορισμένος οπότε από αυτό που λαμβάνουμε δεν έχει κάποιο νόημα διότι δεν μπορεί να εφαρμοστεί η μέθοδος. Παρακάτω φαίνεται η γραφική της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων μέχρι να τερματίσει ο αλγόριθμος:

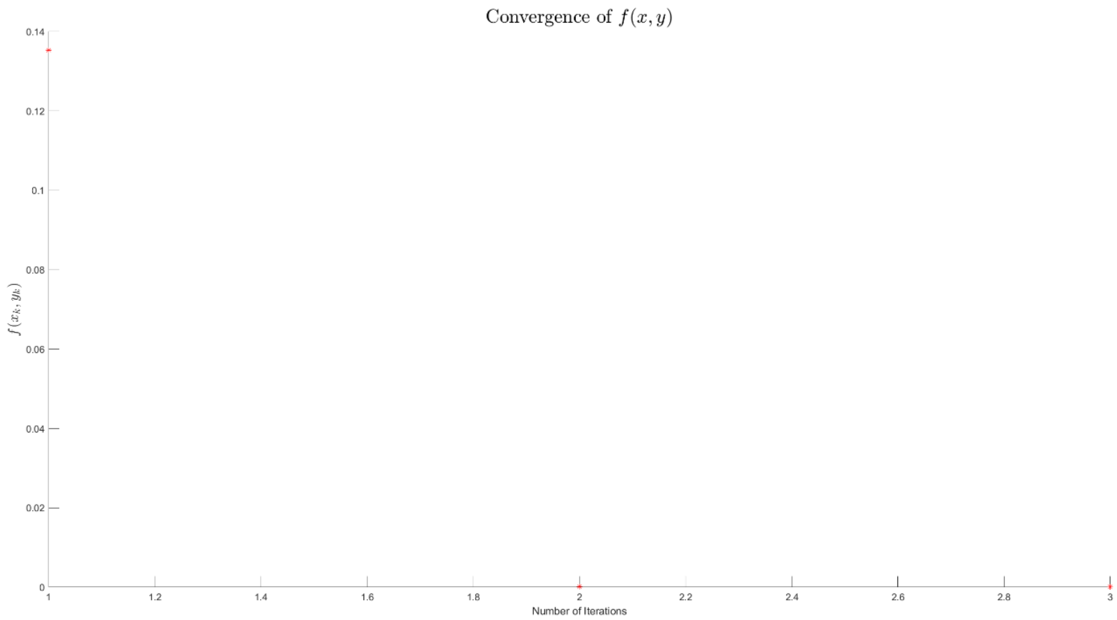




Από το οποίο φαίνεται ότι η μέθοδος κάνει λιγότερες επαναλήψεις (γεγονός που σημαίνει ότι πληρείται πιο «γρήγορα» η συνθήκη τερματισμού $|\nabla f(x_k)| \leq \varepsilon$)ψάχνοντας το βέλτιστο γ_k τέτοιο ώστε να ελαχιστοποιεί την $f(x_k + \gamma_k d_k)$ χωρίς όμως να καταφέρνει να συγκλίνει στο σημείο ισορροπίας.

γ) Τέλος, ο προσδιορισμός του γ_k με αρχικό σημείο αναζήτησης το **(1,1)** και με τον **κανόνα Armijo** μας δίνει παρόμοια αποτελέσματα με τα αποτελέσματα που μας έδωσαν και η επιλογή σταθερού γ_k αλλά και επιλογή τουγ_k βάσει της ελαχιστοποίησης της συνάρτησης $f(x_k + \gamma_k d_k)$. Και σε αυτή την περίπτωση επιλέχθηκαν τα ίδια s, a *και* b με την περίπτωση του σημείου **(-1,-1)** δηλαδή $s = 1, a = 10^{-3}$ *και* $b = 0.3$ **.** Παρακάτω φαίνεται η γραφική παράσταση της σύγκλισης της συνάρτησης ως προς τον αριθμό των επαναλήψεων όπου φαίνεται ξεκάθαρα ότι ο αλγόριθμος δεν θα μας οδηγήσει στο ελάχιστο ενώ επίσης παρατηρείται και ο μικρός αριθμός επαναλήψεων σε σχέση με τους άλλους δύο τρόπους υπολογισμού του γ_k**.**





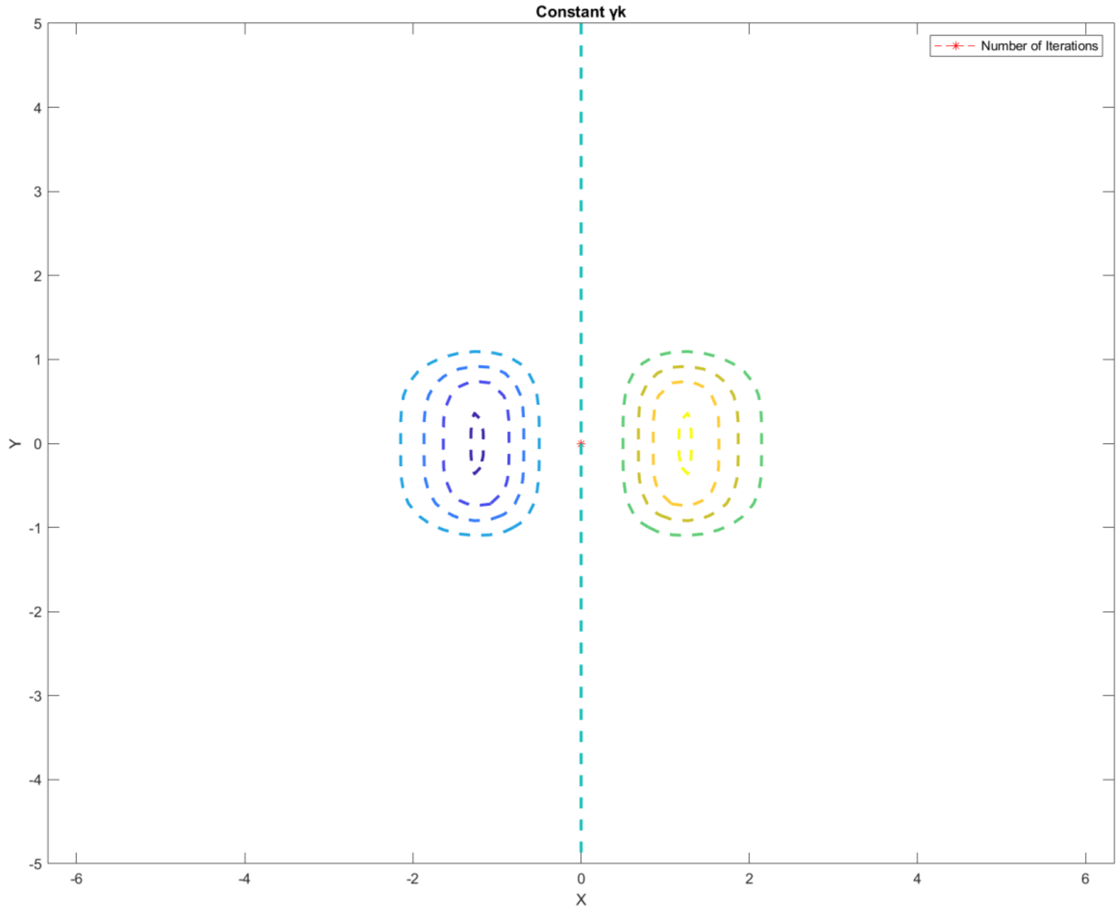
Όπως ήταν και εδώ αναμενόμενο η μέθοδος δεν μπορεί να εφαρμοστεί και άρα δεν μας δίνει σωστά αποτελέσματα, ωστόσο η συνθήκη τερματισμού $|\boldsymbol{\nabla f(x_k)}| \leq \boldsymbol{\varepsilon}$ φαίνεται ότι ικανοποιείται με μικρότερο αριθμό επαναλήψεων με τον **κανόνα Armijo** σε σχέση με τους άλλους δύο τρόπους οι οποίοι απαιτούσαν πολλές περισσότερες επαναλήψεις, με την εσωτερική βελτιστοποίηση να απαιτεί τις περισσότερες.

Θέμα 4ο :

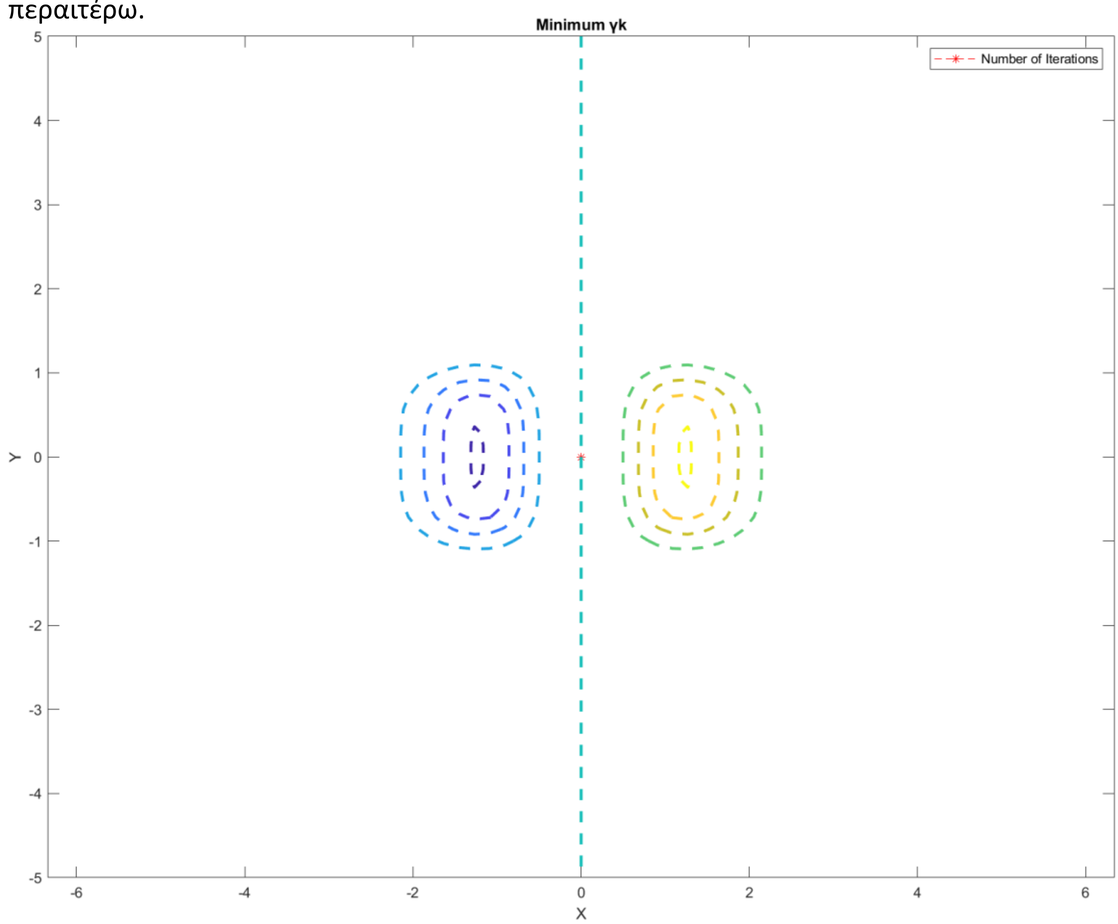
Μέθοδος Levenberg-Marquardt:

Η μέθοδος **Levenberg-Marquardt** είναι ένας από τους τρόπους για τον προσδιορισμό του ελαχίστου μιας συνάρτησης η οποία ικανοποιεί την ιδιότητα της μέγιστης καθόδου δηλαδή $\forall \boldsymbol{k} \in \mathbb{N}$ $\boldsymbol{\iota\sigma\chi}$ύ$\boldsymbol{\varepsilon\iota}$ $\boldsymbol{f}(\boldsymbol{x_k}) > \boldsymbol{f}(\boldsymbol{x_{k+1}})$ και έχει ως διάνυσμα κατεύθυνσης αναζήτησης (δηλαδή την κατεύθυνση στην οποία αναζητούμε το ελάχιστο) το $\boldsymbol{d_k} = -\{\boldsymbol{\nabla^2 f}(\boldsymbol{x_k}) + \boldsymbol{\mu_k I}\}^{-1}\boldsymbol{\nabla f}(\boldsymbol{x_k})$ όπου ο πίνακας $[\boldsymbol{\nabla^2 f}(\boldsymbol{x_k})]^{-1}$ είναι ο αντίστροφος Εσσιανός πίνακας της αντικειμενικής συνάρτησης $\boldsymbol{f}(\boldsymbol{x_k})$, $\boldsymbol{\mu_k}$ είναι ένας θετικός αριθμός μεγαλύτερος κατά μια ποσότητα(όσο μεγάλη θέλουμε εμείς) από την μεγαλύτερη κατά απόλυτη τιμή ιδιοτιμή του Εσσιανού πίνακα και $\boldsymbol{I}$ είναι ο μοναδιαίος πίνακας. Η μέθοδος αυτή εξασφαλίζει ότι σε κάθε επανάληψη ο πίνακας θα είναι θετικά ορισμένος και άρα αντιστρέψιμος(γεγονός που εξασφαλίζεται διότι αφού ο Εσσιανός και ο πολλαπλασιασμένος μοναδιαίος πίνακας είναι συμμετρικοί άμα και το άθροισμα τους είναι θετικά ορισμένο θα έχει σίγουρα αντίστροφο, πράγμα που εξασφαλίζεται σε κάθε επανάληψη διότι το πετυχαίνουμε με την επιλογή του $\boldsymbol{\mu_k}$). Συνεπώς, αυτή η μέθοδος είναι μια βελτιωμένη παραλλαγή της μεθόδου **Newton** οπότε θα έχει τα «καλά» χαρακτηριστικά της τα οποία είναι η μεγαλύτερη ταχύτητα σε σχέση με την μέθοδο **Μέγιστης Καθόδου** διότι έχουμε την πληροφορία της δεύτερης παραγώγου της συνάρτησης. Ωστόσο, όσο μεγαλύτερο επιλέγουμε τον θετικό αριθμό κατά τον οποίο θα είναι μεγαλύτερο το $\boldsymbol{\mu_k}$ από την μέγιστη κατά απόλυτο ιδιοτιμή του Εσσιανού τόσο περισσότερο «μειώνεται» η ταχύτητα σύγκλισης του αλγορίθμου και πλησιάζει τον ρυθμό σύγκλισης της **Μέγιστης Καθόδου** επειδή θα κυριαρχεί ο όρος του μοναδιαίου πίνακα . Συνεπώς, ως προς τον ρυθμό σύγκλισης του αλγορίθμου αναμένουμε να συγκλίνει με ρυθμό ο οποίος βρίσκεται μεταξύ των δύο προηγουμένων μεθόδων.

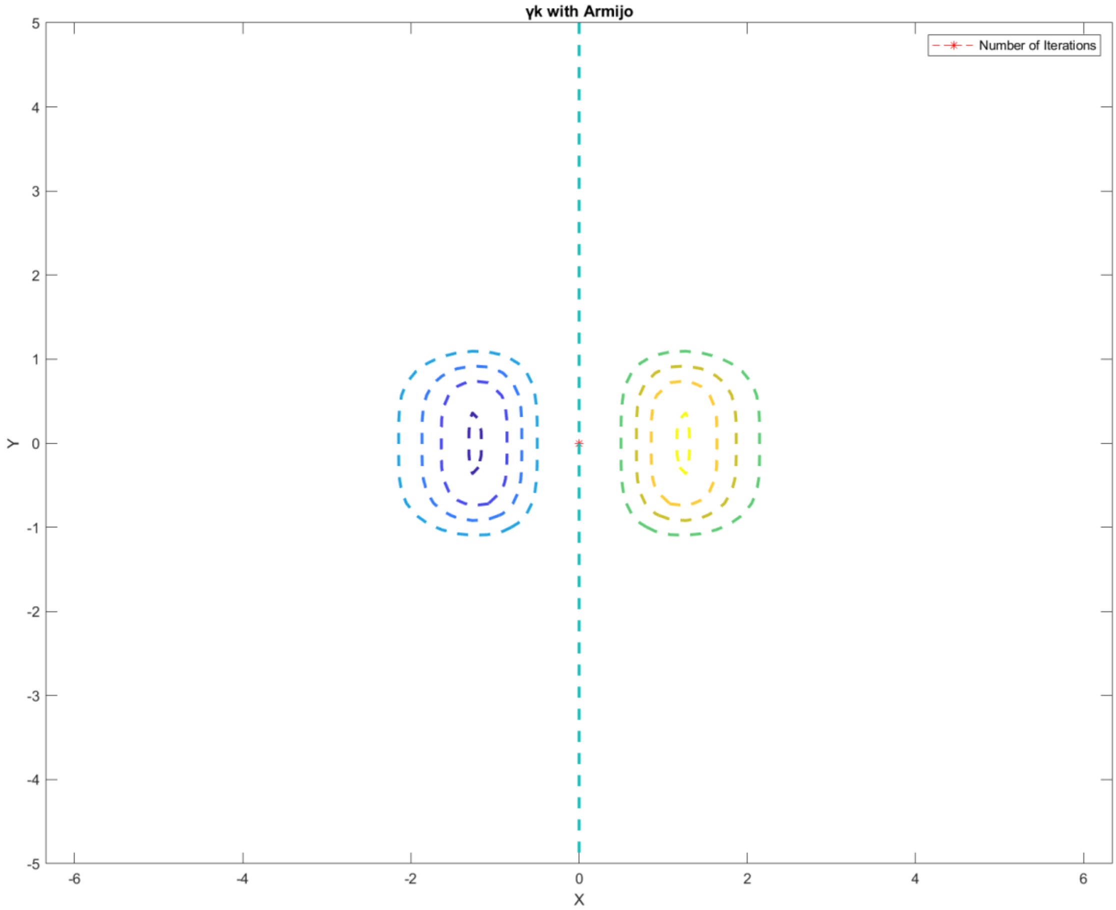
i. **α)** Για το σημείο **(0,0)** δεν έχει σημασία η επιλογή του $\boldsymbol{\gamma_k}$ καθότι σε εκείνο το σημείο όπως φαίνεται και από την γραφική παράσταση της αντικειμενικής συνάρτησης ισχύει ότι $\boldsymbol{\nabla f}(\boldsymbol{x_k}) = \boldsymbol{0}$ είναι δηλαδή ένα σαγματικό σημείο, οπότε η μέθοδος «παγιδεύεται» και δεν μπορεί να προχωρήσει περαιτέρω με αποτέλεσμα να μην μας δίνει το ελάχιστο το οποίο εμείς αναζητούμε. . **(** Η γραφική παράσταση έγινε για σταθερό $\boldsymbol{\gamma_k = 0.7}$**)**



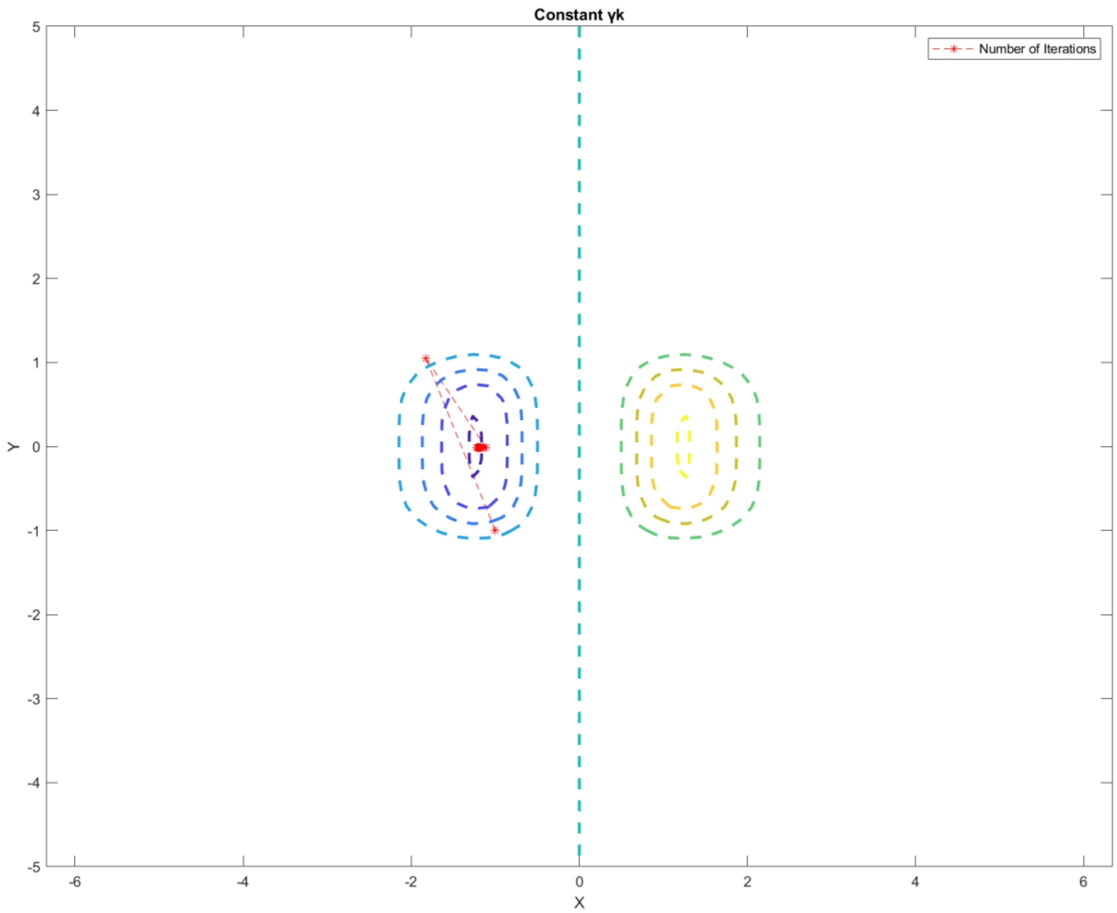
β) Ομοίως, για το ίδιο πάλι σημείο, η επιλογή του $\boldsymbol{\gamma_k}$ τέτοια ώστε να ελαχιστοποιεί την $\boldsymbol{f}(\boldsymbol{x_k} + \boldsymbol{\gamma_k d_k})$ δεν έχει κάποιο νόημα διότι ισχύει $\boldsymbol{\nabla f}(\boldsymbol{x_k}) = \mathbf{0}$ οπότε όπως και πριν η μέθοδος θα «παγιδεύεται» στο σαγματικό σημείο και δεν θα μπορεί να προχωρήσει περαιτέρω.



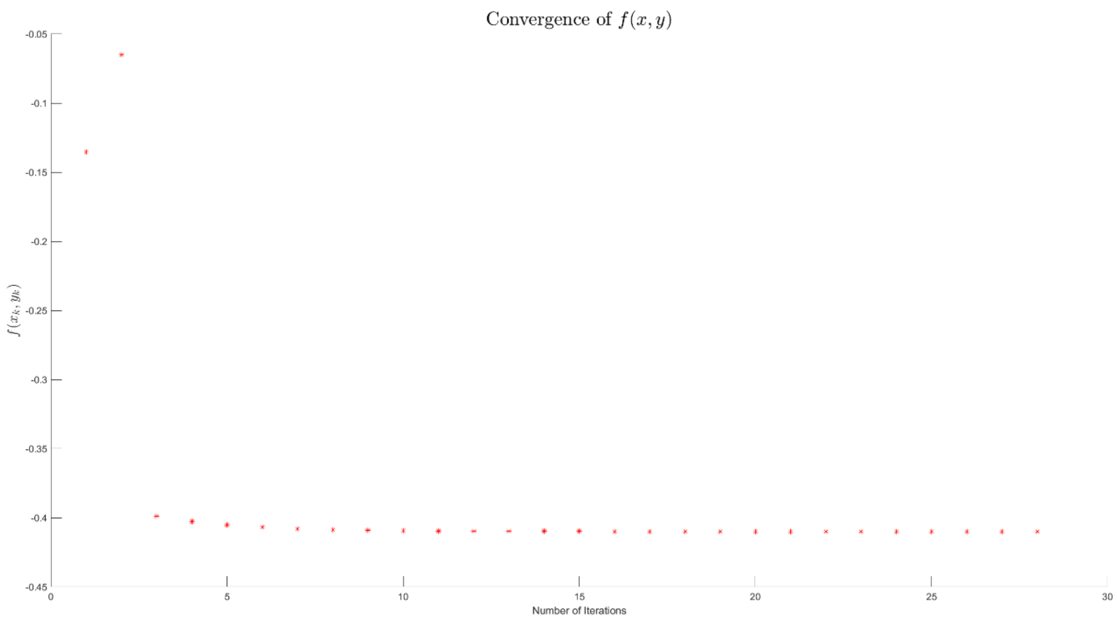
γ) Τέλος, όπως και με τους δύο προηγούμενους τρόπους έτσι και με τον **κανόνα Armijo** δεν θα παρατηρήσουμε κάποια μεταβολή. Με $\boldsymbol{s = 1, a = 10^{-3}\ \kappa\alpha\iota\ b = 0.3}$.



ii. **α)** Για το σημείο **(-1,-1)** παρατηρούμε ότι η μέθοδος θα συγκλίνει στην ελάχιστη τιμή της συνάρτησης, όπως δηλαδή και η μέθοδος **Μέγιστης Καθόδου,** και μάλιστα αρκετά πιο γρήγορα από αυτήν κάτι που ήταν αναμενόμενο για τους λόγους που αναφέραμε στην αρχή του **θέματος 4.**Παρακάτω φαίνεται η γραφική παράσταση της σύγκλισης της αντικειμενικής συνάρτησης που προέκυψε ως προς τον αριθμό των επαναλήψεων, μέχρι να τερματίσει ο αλγόριθμος, η οποία είναι η εξής (Η γραφική παράσταση έγινε για σταθερό $\boldsymbol{\gamma_k = 0.4}$):

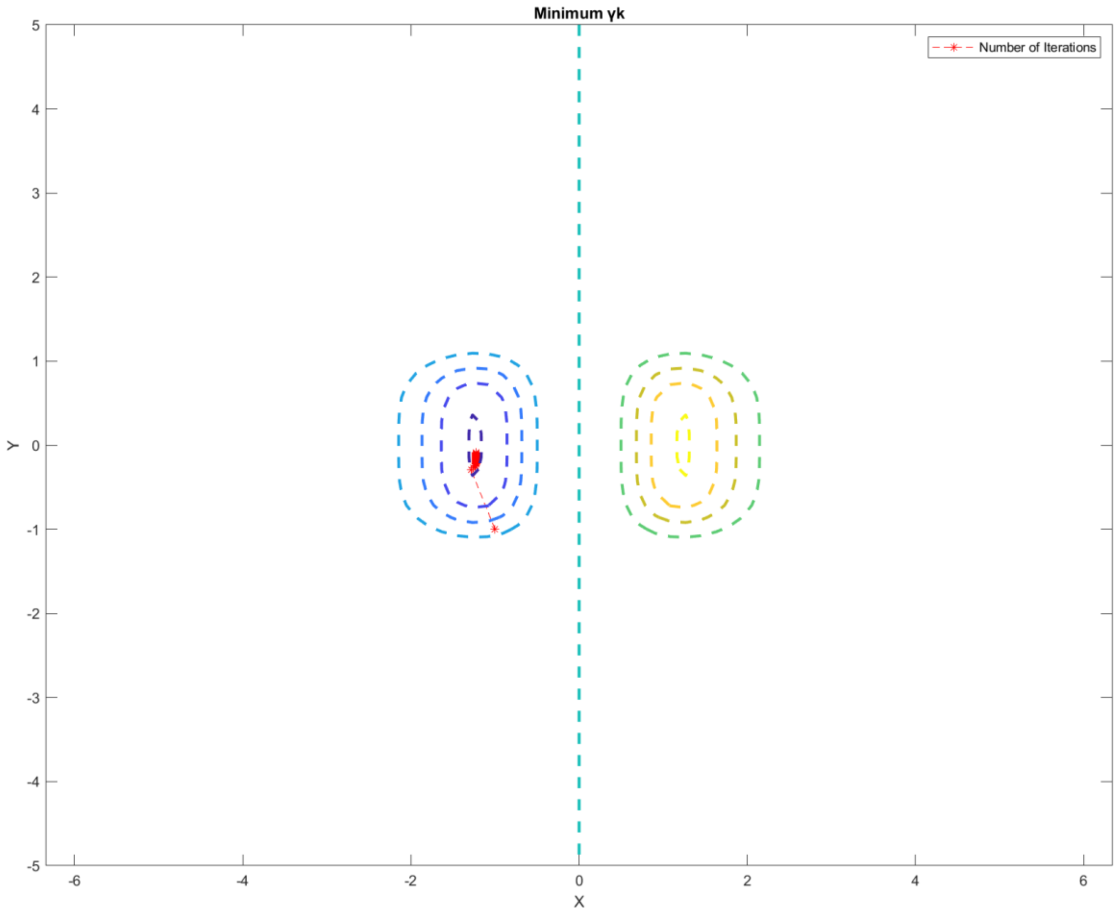


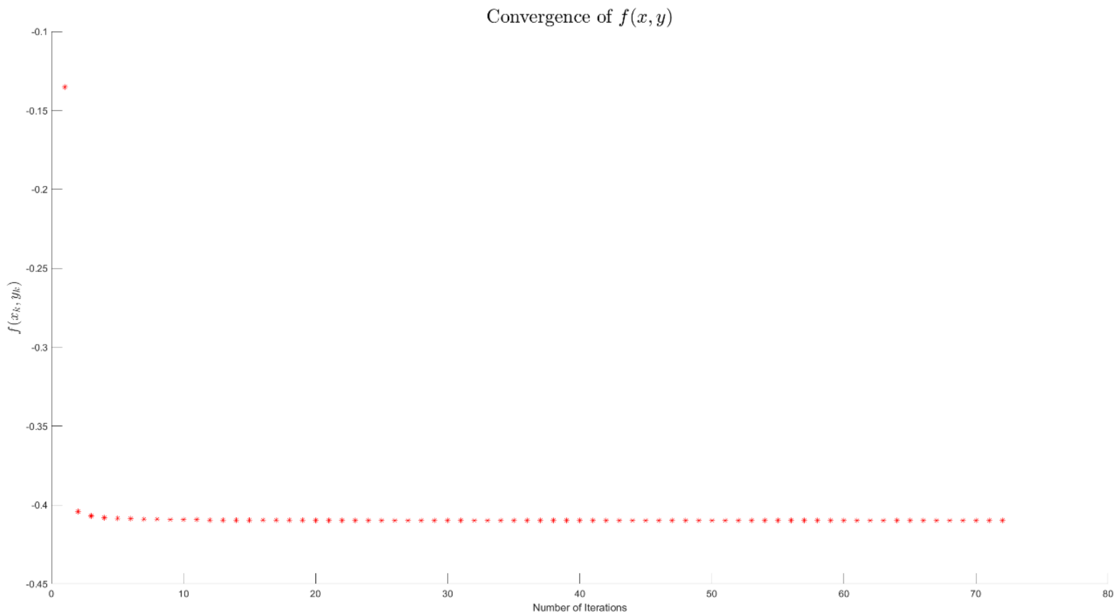
Constant γk
Number of Iterations
X
Y



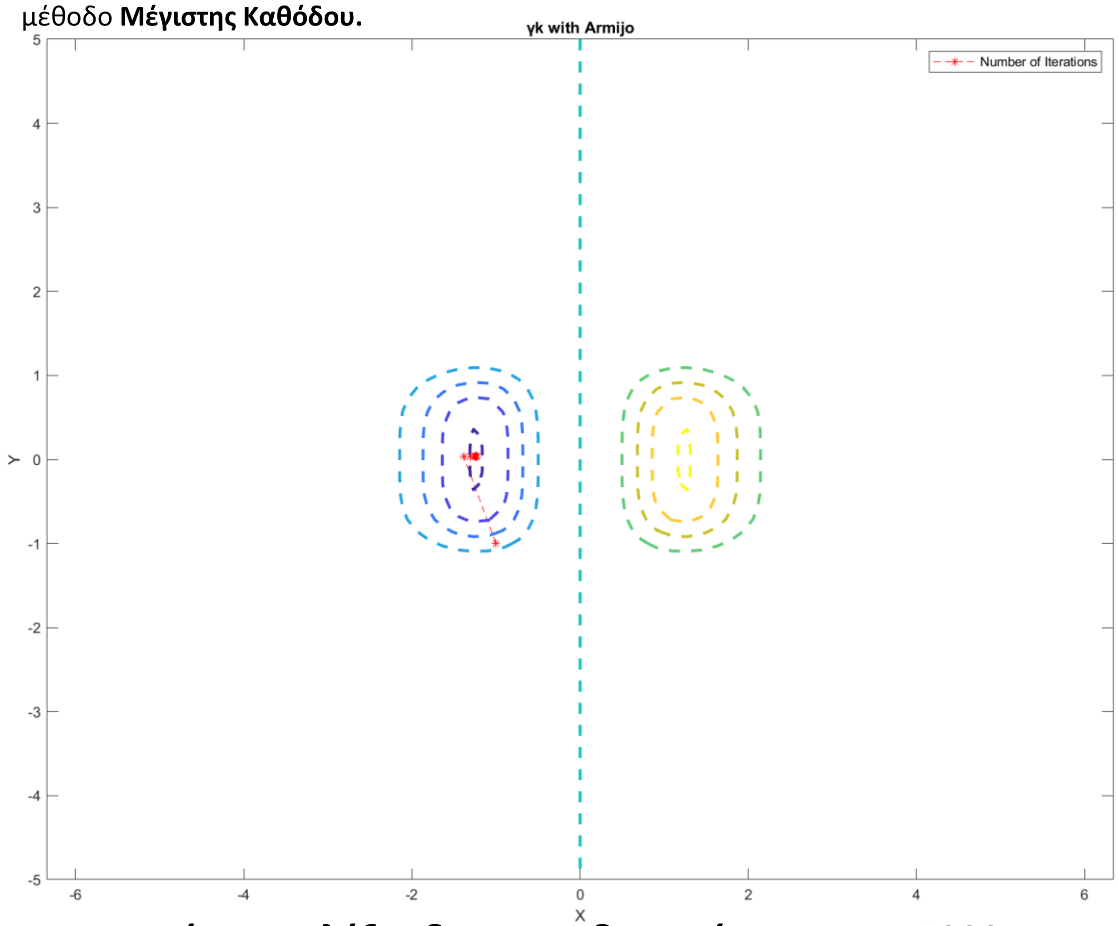
Convergence of $f(x,y)$
$f(x_k, y_k)$
Number of Iterations

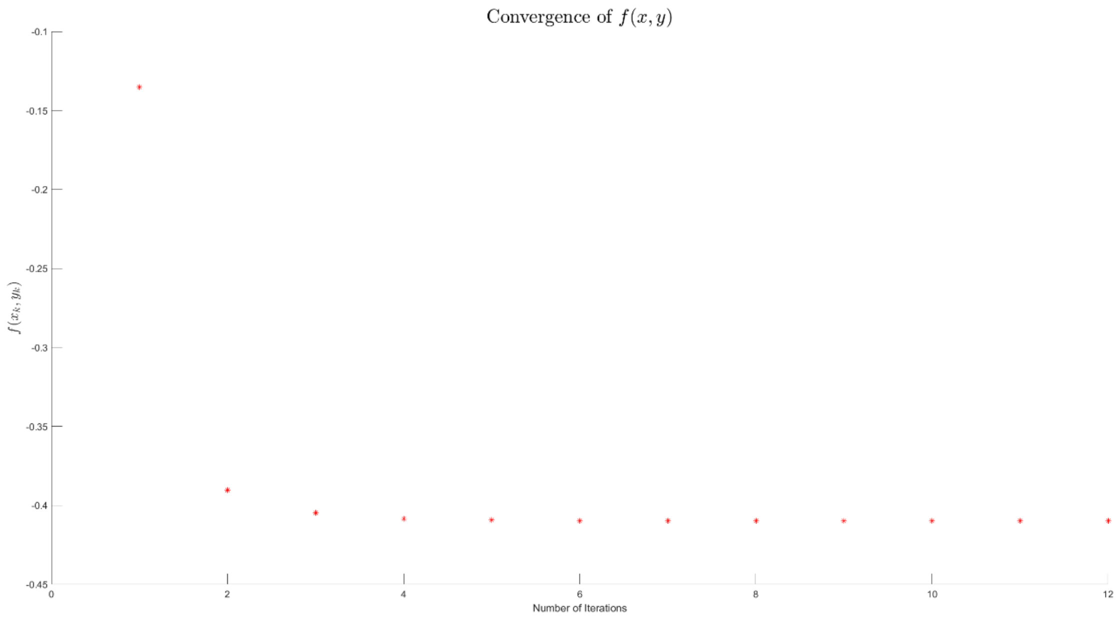
β) Για το ίδιο πάλι σημείο, δηλαδή για το **(-1,-1)** η επιλογή του $\boldsymbol{\gamma_k}$ τέτοια ώστε να ελαχιστοποιεί την $\boldsymbol{f}(\boldsymbol{x_k} + \boldsymbol{\gamma_k d_k})$ γίνεται με την μέθοδο της χρυσής τομής που υλοποιήσαμε στην προηγούμενη εργασία, τροποποιημένη έτσι ώστε αντί να επιστρέφει διάστημα στο οποίο μπορεί να βρίσκεται το ελάχιστο, να επιστρέφει σημείο το οποίο προκύπτει από την μέση τιμή των άκρων του διαστήματος στο οποίο η μέθοδος εντοπίζει το ελάχιστο. Ωστόσο, παρατηρώ ότι με αυτό τον τρόπο εσωτερικής βελτιστοποίησης η μέθοδος απαιτεί περισσότερες επαναλήψεις από ότι η μέθοδος **Μέγιστης Καθόδου**,κάτι που δεν ήταν αναμενόμενο,αλλά πιθανόν να οφείλεται στον θετικό αριθμό που προσθέτω στην μέγιστη κατά απόλυτο ιδιοτιμή του Εσσιανού. Παρακάτω φαίνεται η γραφική της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων μέχρι να τερματίσει ο αλγόριθμος:



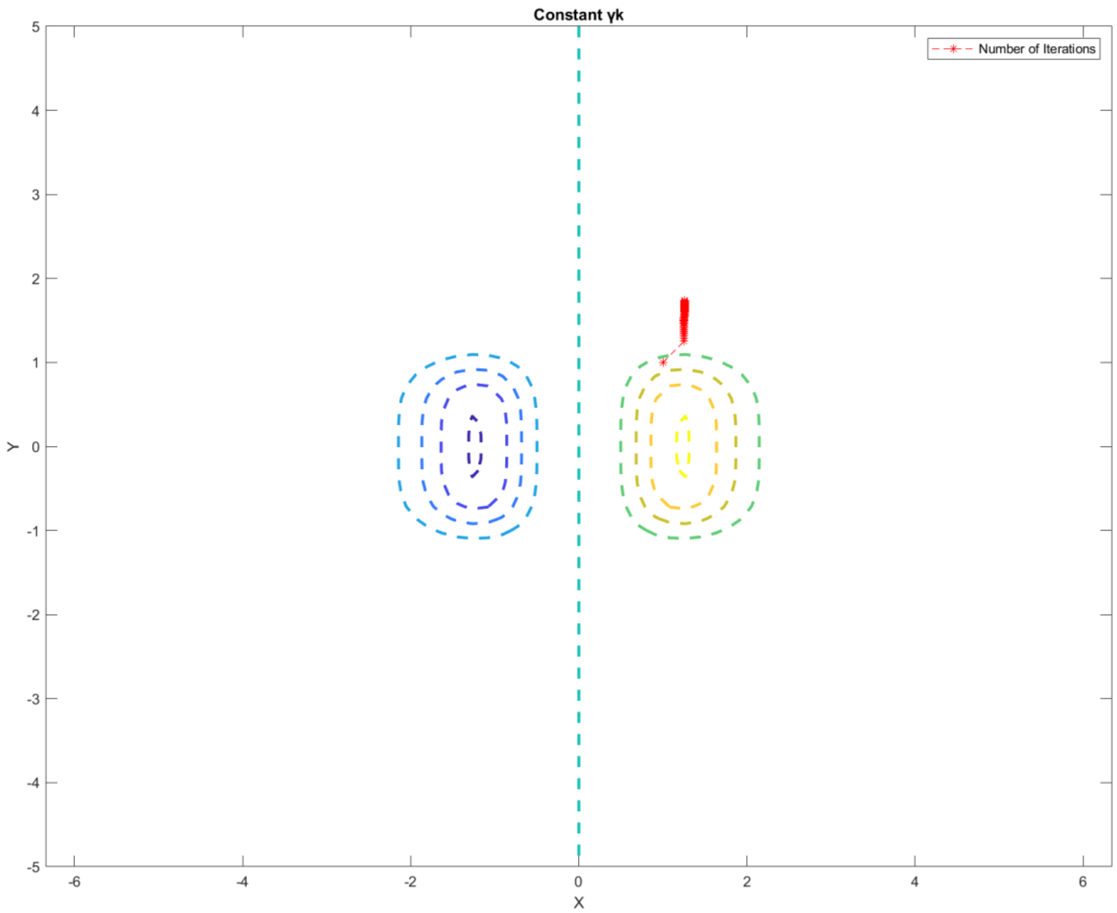


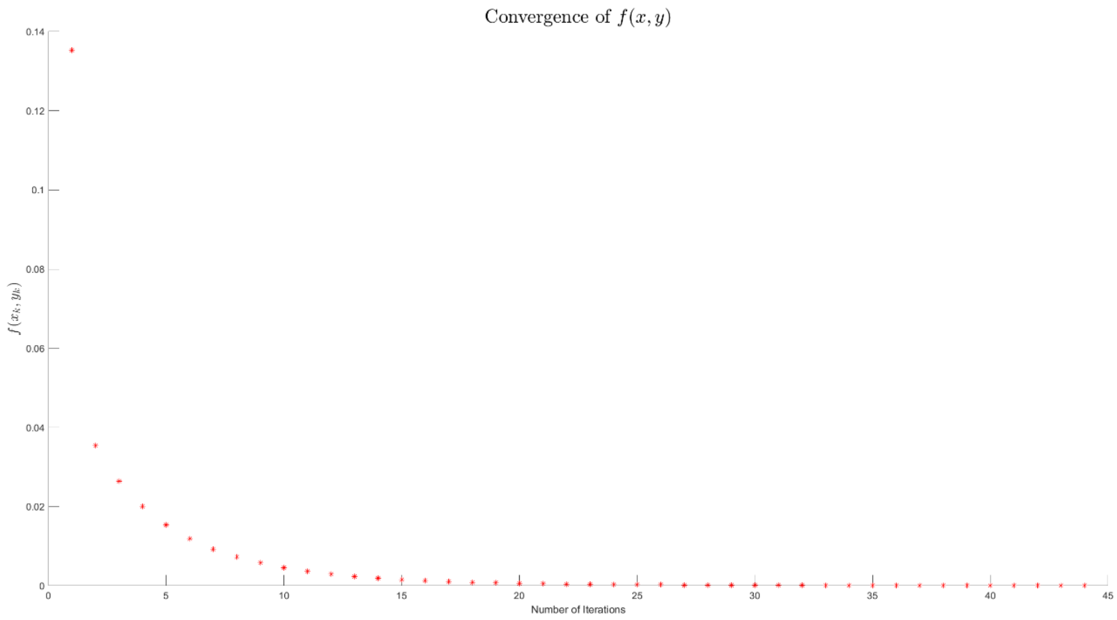
γ) Τέλος, ο προσδιορισμός του $\boldsymbol{\gamma_k}$ με αρχικό σημείο αναζήτησης το **(-1,-1)** και με τον **κανόνα Armijo** μας δίνει παρόμοια αποτελέσματα με τα αποτελέσματα που μας έδωσαν και η επιλογή σταθερού $\boldsymbol{\gamma_k}$ αλλά και επιλογή του $\boldsymbol{\gamma_k}$ βάσει της ελαχιστοποίησης της συνάρτησης $\boldsymbol{f(x_k + \gamma_k d_k)}$. Και σε αυτή την περίπτωση επιλέχθηκαν τα ίδια $\boldsymbol{s, a \;\kappa\alpha\iota\; b}$ με την περίπτωση του σημείου **(0,0)** δηλαδή $\boldsymbol{s = 1, a = 10^{-3} \;\kappa\alpha\iota\; b = 0.3}$ **.** Παρακάτω φαίνεται η γραφική παράσταση της σύγκλισης της συνάρτησης ως προς τον αριθμό των επαναλήψεων όπου φαίνεται και πάλι ότι η μέθοδος **Levenberg-Marquardt** είναι πάλι πιο γρήγορη από την μέθοδο **Μέγιστης Καθόδου.**



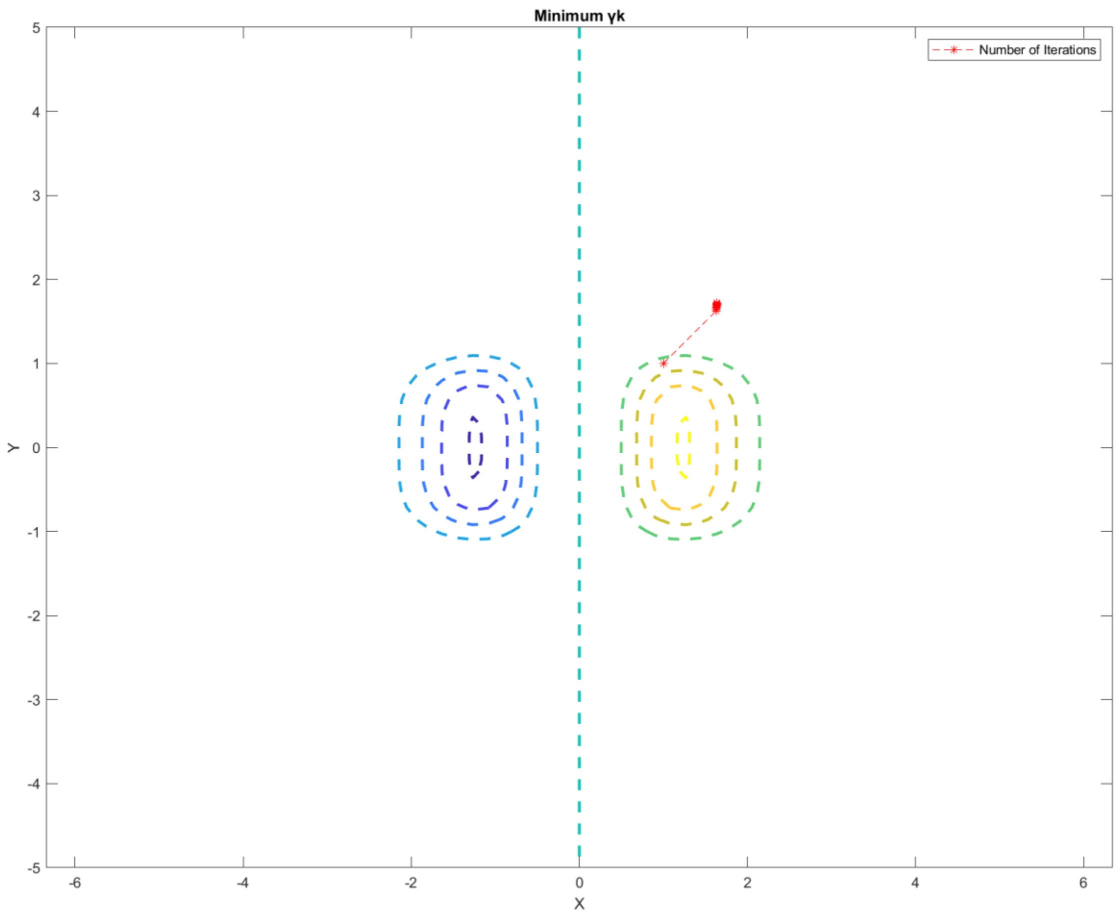


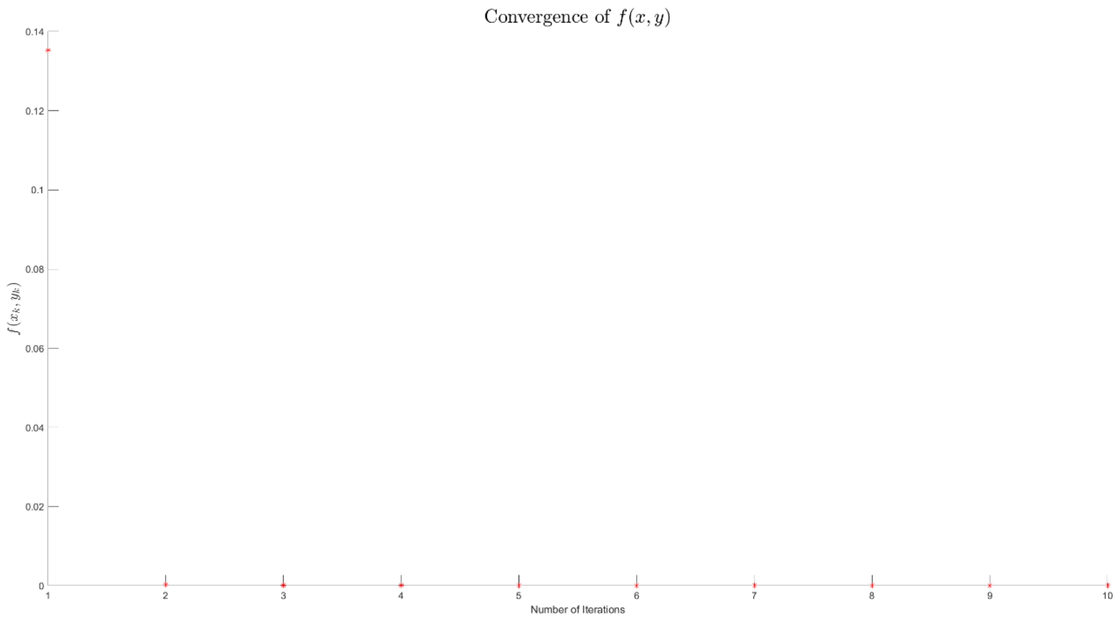
iii. **α)** Για το σημείο **(1, 1)** παρατηρούμε όπως αναφέρθηκε και προηγουμένως ότι η μέθοδος δεν θα συγκλίνει στο ελάχιστο γιατί λόγω της μορφής της συνάρτησης σε εκείνο το σημείο αναζήτησης η μέθοδος εγκλωβίζεται σε ένα σαγματικό σημείο πολύ κοντά στο μηδέν. Παρακάτω φαίνεται η γραφική παράσταση της σύγκλισης της αντικειμενικής συνάρτησης που προέκυψε ως προς τον αριθμό των επαναλήψεων, μέχρι να τερματίσει ο αλγόριθμος, η οποία είναι η εξής (Η γραφική παράσταση έγινε για σταθερό $\boldsymbol{\gamma_k = 0.4}$**)** όπου και πάλι φαίνεται ότι έχει μεγαλύτερο ρυθμό σύγκλισης από την **Μέγιστη Κάθοδο**:



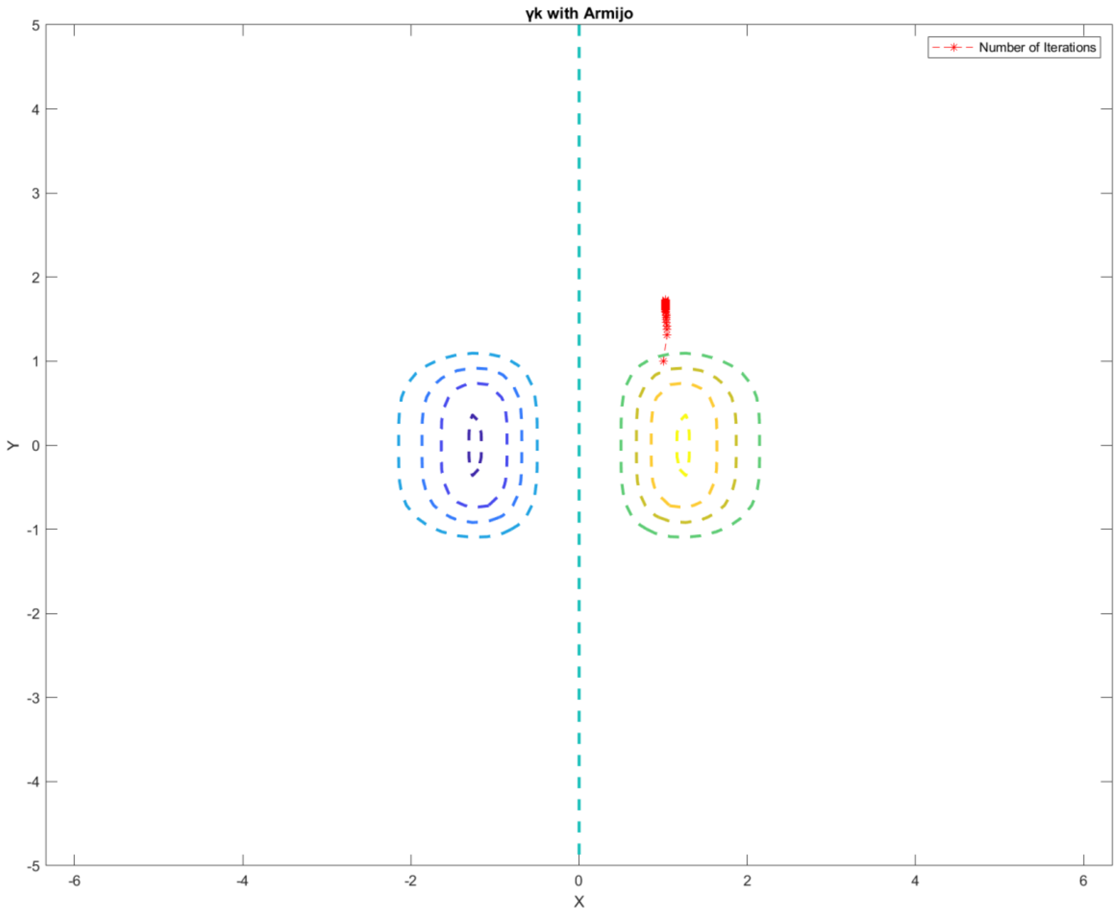


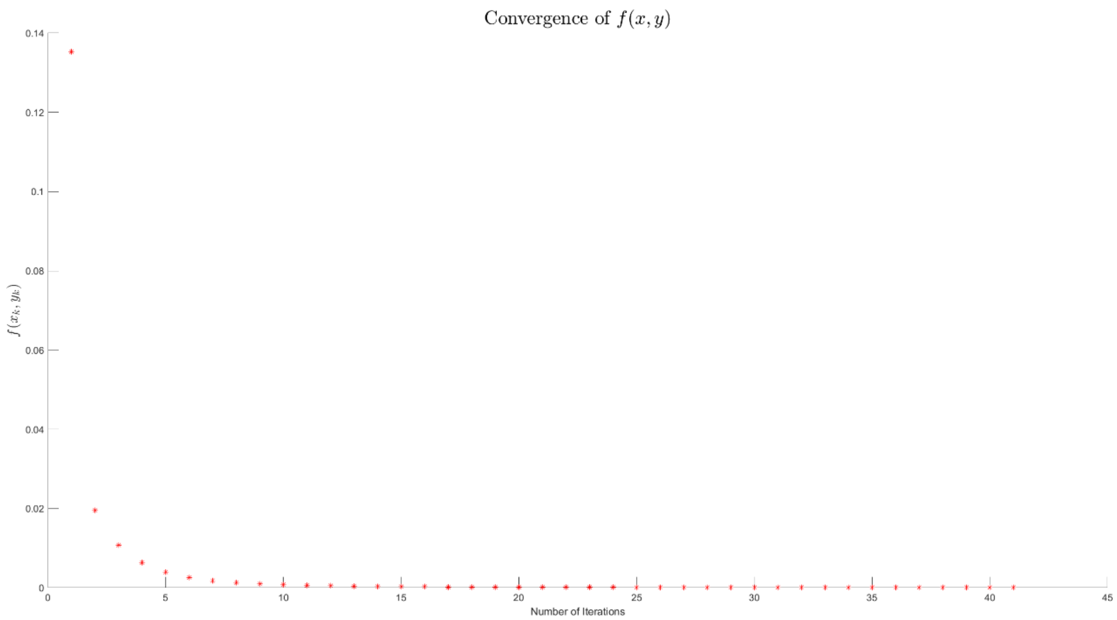
β) Για το ίδιο πάλι σημείο, δηλαδή για το **(1,1)** η επιλογή του $\boldsymbol{\gamma_k}$ τέτοια ώστε να ελαχιστοποιεί την $\boldsymbol{f}(\boldsymbol{x_k} + \boldsymbol{\gamma_k}\boldsymbol{d_k})$ γίνεται με την μέθοδο της χρυσής τομής που υλοποιήσαμε στην προηγούμενη εργασία, τροποποιημένη έτσι ώστε αντί να επιστρέφει διάστημα στο οποίο μπορεί να βρίσκεται το ελάχιστο, να επιστρέφει σημείο το οποίο προκύπτει από την μέση τιμή των άκρων του διαστήματος στο οποίο η μέθοδος εντοπίζει το ελάχιστο. Παρακάτω φαίνεται η γραφική της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων μέχρι να τερματίσει ο αλγόριθμος:





γ) Τέλος, ο προσδιορισμός του $\boldsymbol{\gamma_k}$ με αρχικό σημείο αναζήτησης το **(1,1)** και με τον **κανόνα Armijo** μας δίνει παρόμοια αποτελέσματα με τα αποτελέσματα που μας έδωσαν και η επιλογή σταθερού $\boldsymbol{\gamma_k}$ αλλά και επιλογή του$\boldsymbol{\gamma_k}$ βάσει της ελαχιστοποίησης της συνάρτησης $\boldsymbol{f(x_k + \gamma_k d_k)}$. Και σε αυτή την περίπτωση επιλέχθηκαν τα ίδια $\boldsymbol{s, a\ και\ b}$ με την περίπτωση του σημείου **(-1,-1)** δηλαδή $\boldsymbol{s = 1, a = 10^{-3}\ και\ b = 0.3}$ **.** Παρακάτω φαίνεται η γραφική παράσταση της σύγκλισης της συνάρτησης ως προς τον αριθμό των επαναλήψεων όπου φαίνεται ξεκάθαρα ότι ο αλγόριθμος δεν θα μας οδηγήσει στο ελάχιστο ενώ επίσης παρατηρείται και ο μικρός αριθμός επαναλήψεων σε σχέση με την μέθοδο **Μέγιστης Καθόδου.**




Convergence of $f(x,y)$
$f(x_k, y_k)$
Number of Iterations
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0
5
10
15
20
25
30
35
40
45