ΕΡΓΑΣΙΑ 2

Ελαχιστοποίηση συνάρτησης δύο μεταβλητών με χρήση παραγώγων

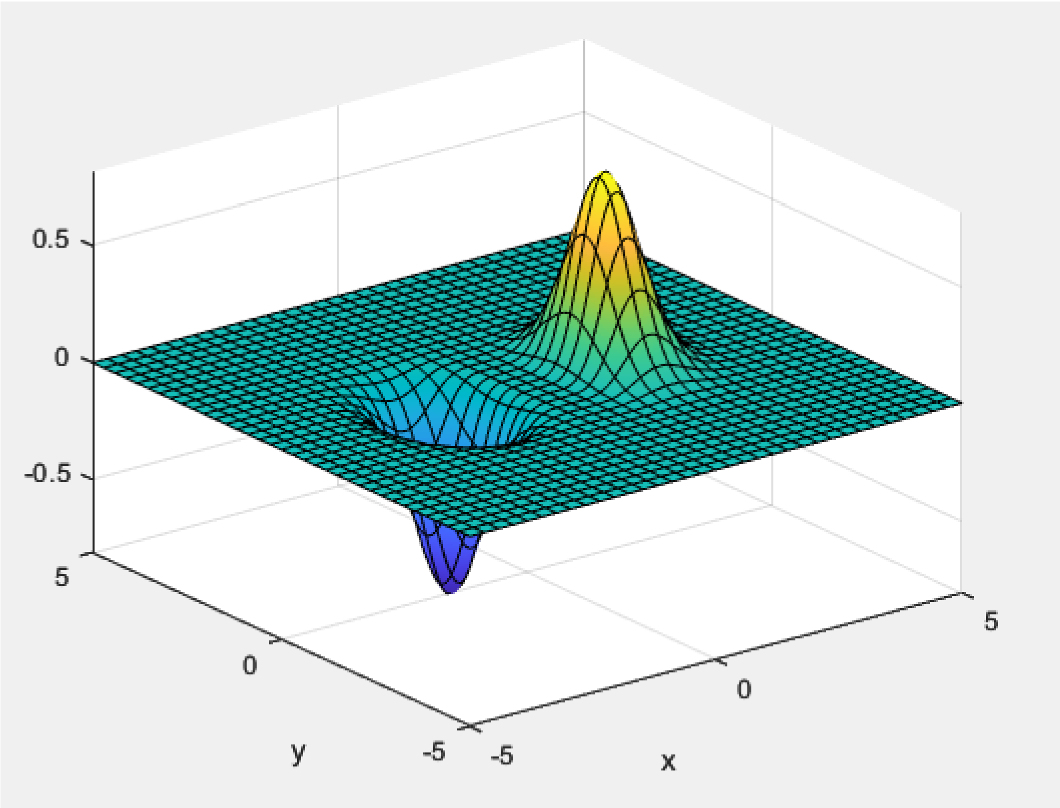
Ονοματεπώνυμο: Παπούλιας Μιχαήλ

ΑΕΜ: 10204

email: mpapouli@ece.auth.gr

Θέμα 1

Αρχικά, σχεδιάζουμε τη γραφική παράσταση της συνάρτησης f(x,y) = x^5 * e^(-x^2-y^2), ώστε να έχουμε μια γενική εικόνα της μορφής της:

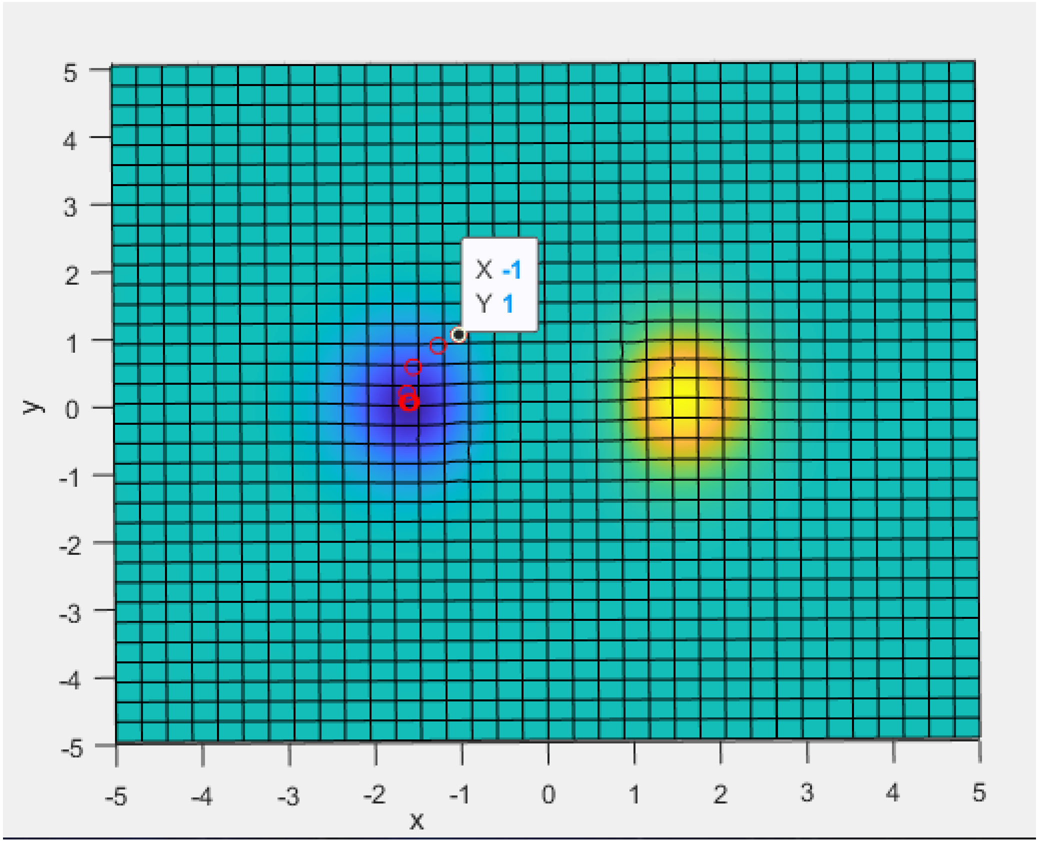


Θέμα 2

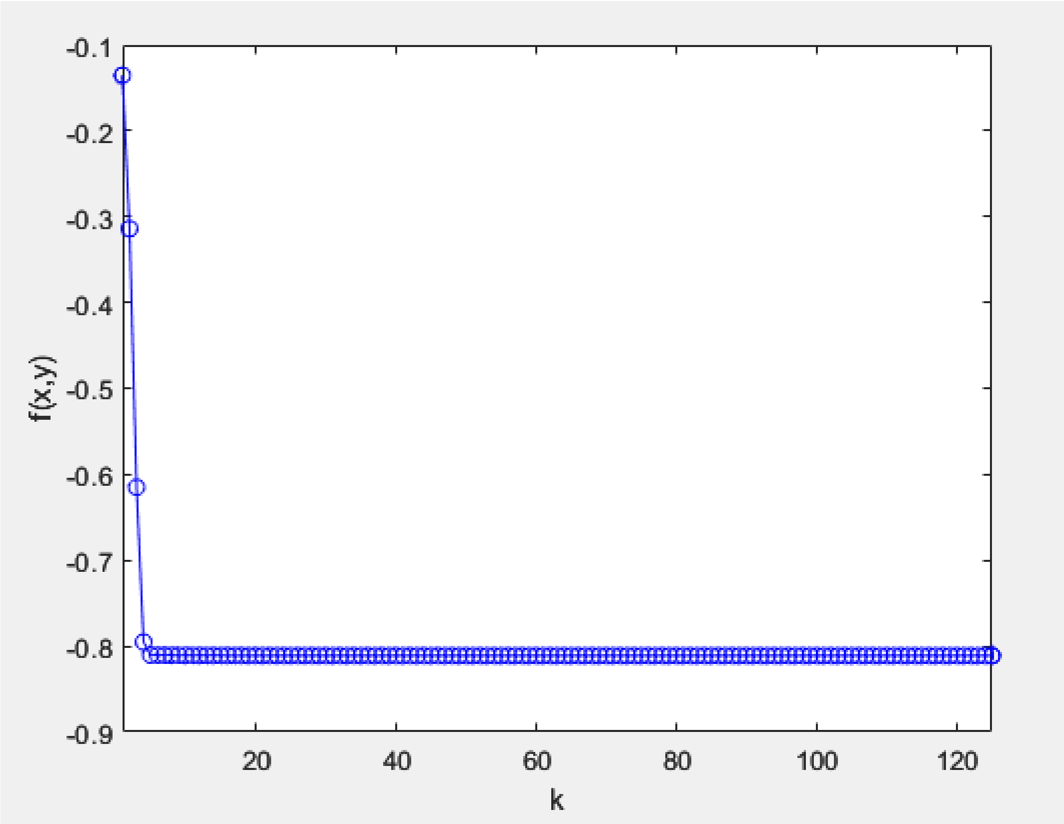
Κατασκευάζουμε τη μέθοδο της μέγιστης καθόδου, με τρεις διαφορετικούς τρόπους, όσον αφορά την επιλογή του βήματος gk: αρχικά, θεωρούμε gk σταθερό, της επιλογής μας, στη συνέχεια το επιλέγουμε τέτοιο ώστε να ελαχιστοποιεί την f(xk + gk*dk), και τέλος, επιλέγεται gk βάση του κανόνα Armijo. Τα σημεία εκκίνησης που ζητούνται είναι: [-1 1], [0 0] και [1 -1].

<u>Σημείο εκκίνησης [-1 1]</u>

Α) Θέτοντας σταθερό gk = 0.6, τρέχουμε τον αλγόριθμο της μέγιστης καθόδου και εμφανίζουμε τη σύγκλιση των διαδοχικών σημείων μέχρι την εύρεση του ελαχίστου, για σφάλμα e = 0.0001:

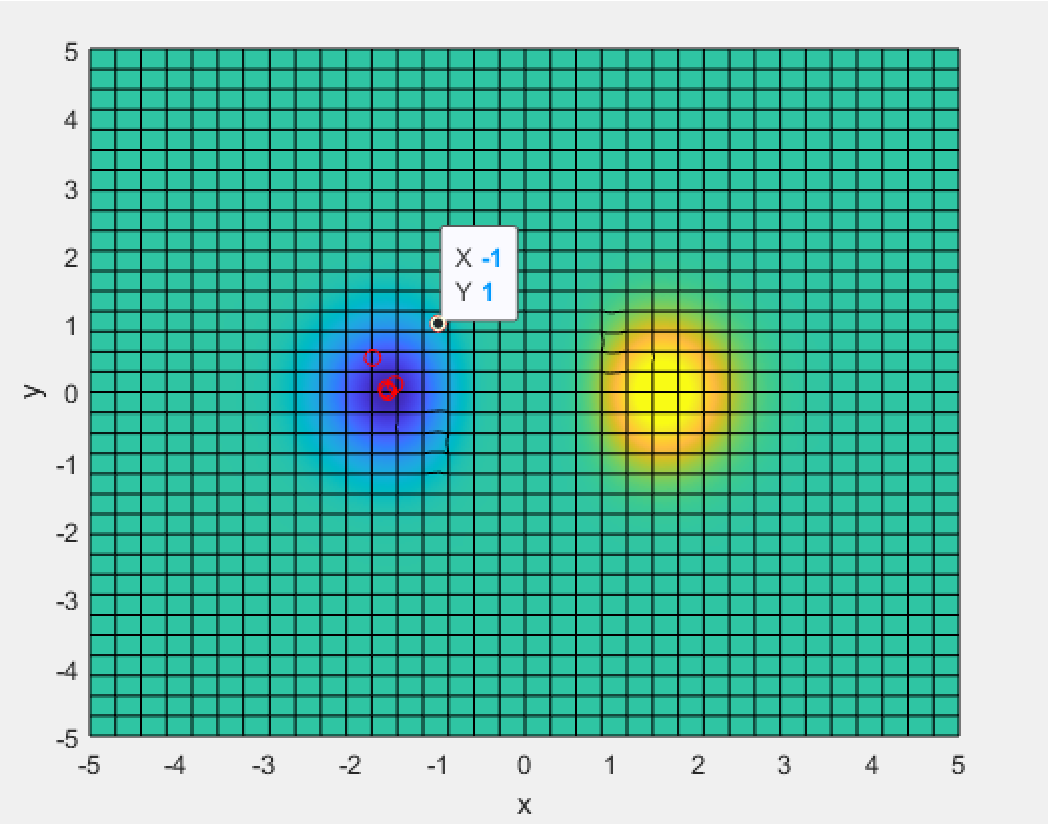


Σύμφωνα με τη μέθοδο και τις επιλεγμένες σταθερές, το ελάχιστο βρίσκεται στο σημείο [-1.58110844, 8.11078187e-192], και ο συνολικός αριθμός διαδοχικών σημείων είναι k = 125. Εμφανίζουμε, ακόμη, τη γραφική παράσταση της αριθμητικής τιμής της συνάρτησης f(x,y) στα σημεία αυτά, συναρτήσει του βήματος k:

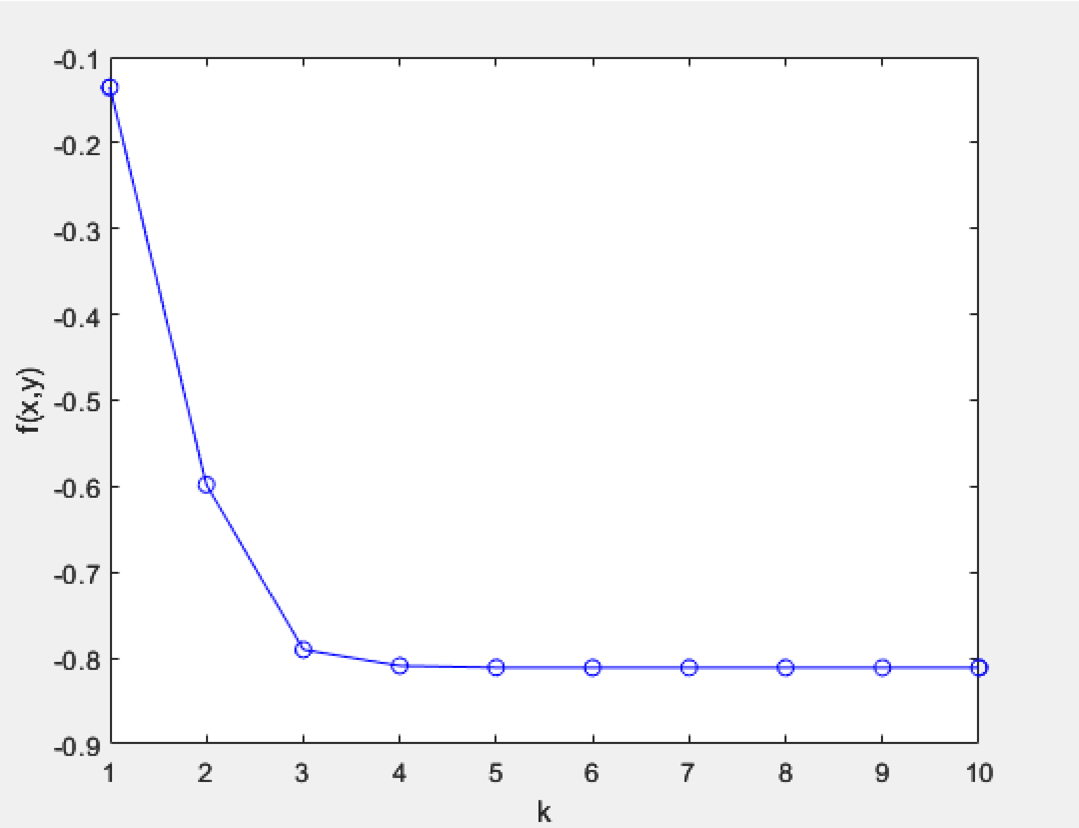


Βλέπουμε ότι τα σημεία πράγματι συγκλίνουν στο ελάχιστο, όσο αυξάνεται το k. Μάλιστα, παρατηρώντας τη γραφική παράσταση της f, φαίνεται ότι το ελάχιστο αυτό είναι το ολικό της. Αυτό συμβαίνει, καθώς ξεκινήσαμε από μια περιοχή σχετικά κοντά στο ολικό ελάχιστο, μέσα στη 'κοιλάδα' της f. Η σύγκλιση, βέβαια, είναι αρκετά αργή, εξαιτίας της φύσης του αλγορίθμου. Το βήμα gk δεν επιλέχθηκε πραγματικά τυχαία, αλλά τέτοιο ώστε να οδηγήσει σε σύγκλιση του αλγορίθμου στο ελάχιστο. Για μεγαλύτερες ή μικρότερες τιμές του gk μπορεί κανείς να διαπιστώσει ότι η μέθοδος δε θα συγκλίνει.

B) Σε κάθε επανάληψη του αλγορίθμου, υπολογίζουμε το gk, ώστε να ελαχιστοποιείται η f(xk + gk*dk). Τρέχουμε τον αλγόριθμο της μέγιστης καθόδου και εμφανίζουμε τη σύγκλιση των διαδοχικών σημείων μέχρι την εύρεση του ελαχίστου, για σφάλμα e = 0.0001:

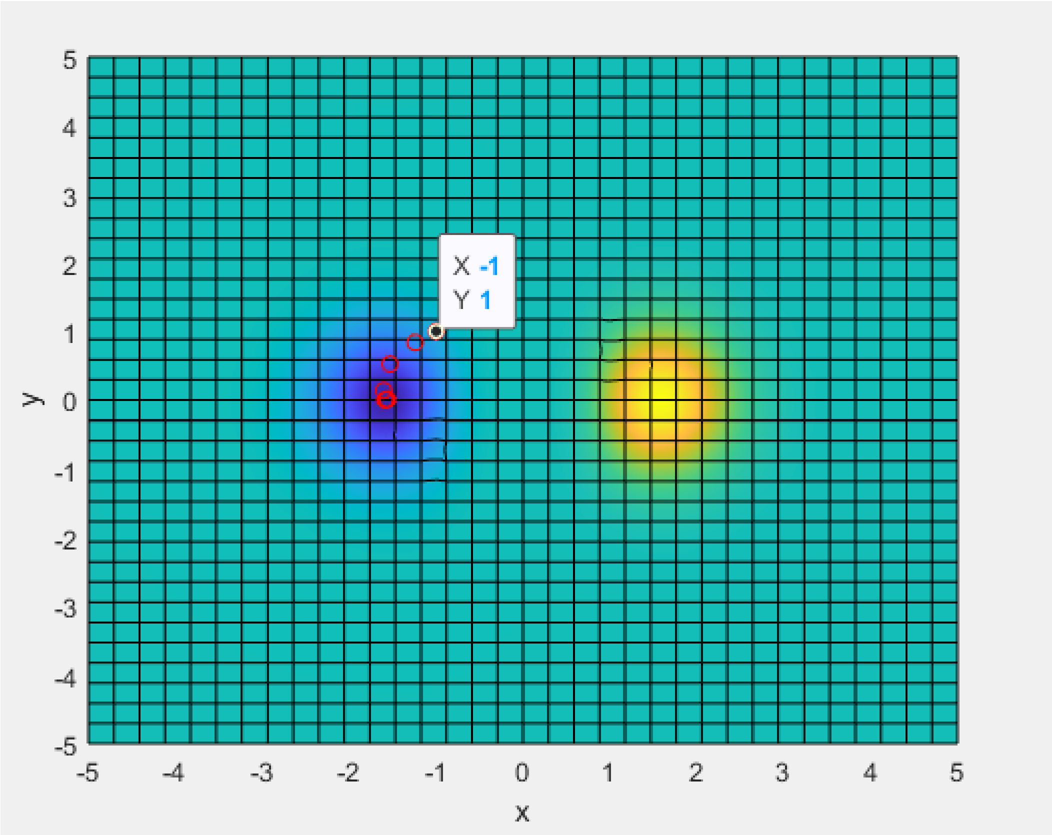


Σύμφωνα με τη μέθοδο και τις επιλεγμένες σταθερές, το ελάχιστο βρίσκεται στο σημείο [-1.58115282, 0.00004197], και ο συνολικός αριθμός διαδοχικών σημείων είναι k = 10. Βλέπουμε, ακόμη, ότι τα διαδοχικά σημεία σχηματίζουν κάθετα μεταξύ τους διανύσματα, το οποίο είναι λογικό, από την επιλογή του μεγέθους gk. Εμφανίζουμε, ακόμη, τη γραφική παράσταση της αριθμητικής τιμής της συνάρτησης f(x,y) στα σημεία αυτά, συναρτήσει του βήματος k:

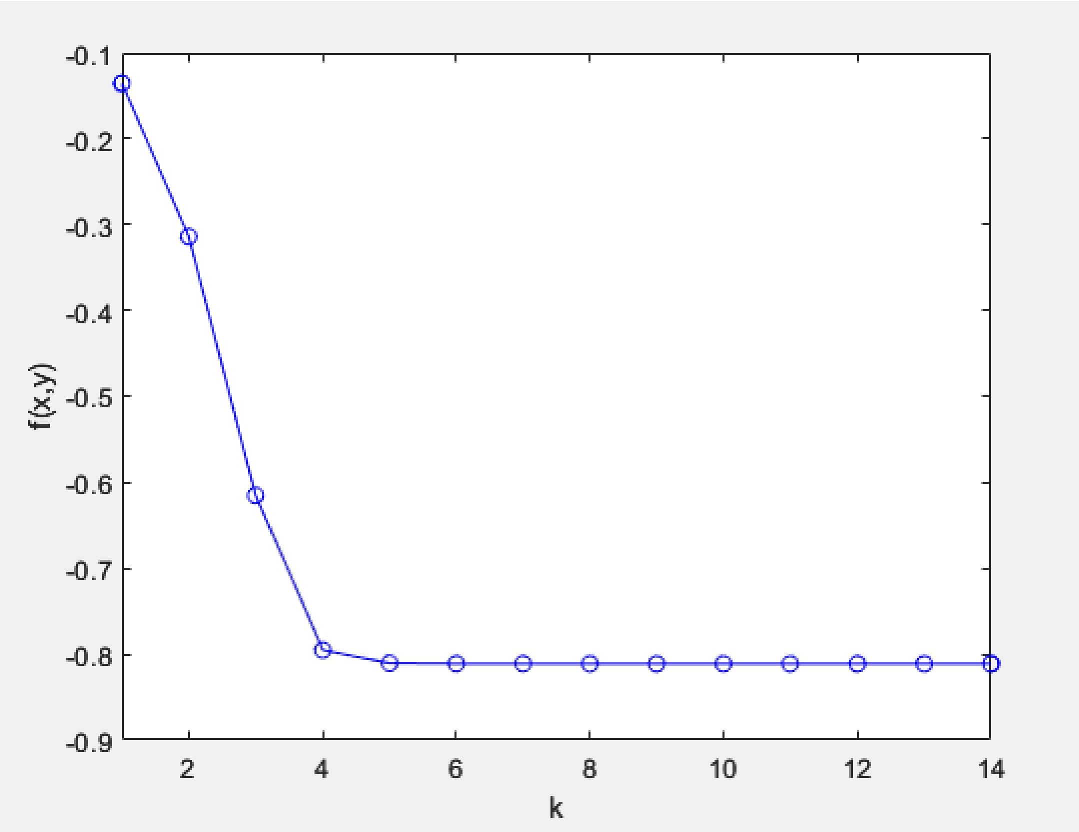


Βλέπουμε ότι τα σημεία πράγματι συγκλίνουν στο ελάχιστο, όσο αυξάνεται το k. Μάλιστα, παρατηρώντας τη γραφική παράσταση της f, φαίνεται ότι το ελάχιστο αυτό είναι το ολικό της. Η σύγκλιση εδώ είναι πολύ πιο γρήγορη, μιας και θέτουμε gk κάθε φορά τέτοιο ώστε να ελαχιστοποιεί την f(xk + gk*dk). Φαίνεται ότι η μέθοδος αυτή πάντοτε θα συγκλίνει στο ολικό ελάχιστο της f, καθώς χρησιμοποιεί τη βέλτιστη επιλογή του gk κάθε φορά.

Γ) Υπολογίζουμε, τώρα, το βήμα gk με βάση τον κανόνα Armijo. Με αρχικές συνθήκες a = 0.05, b = 0.3 και s = 0.6, τρέχουμε τον αλγόριθμο της μέγιστης καθόδου και εμφανίζουμε τη σύγκλιση των διαδοχικών σημείων μέχρι την εύρεση του ελαχίστου, για σφάλμα e = 0.0001:



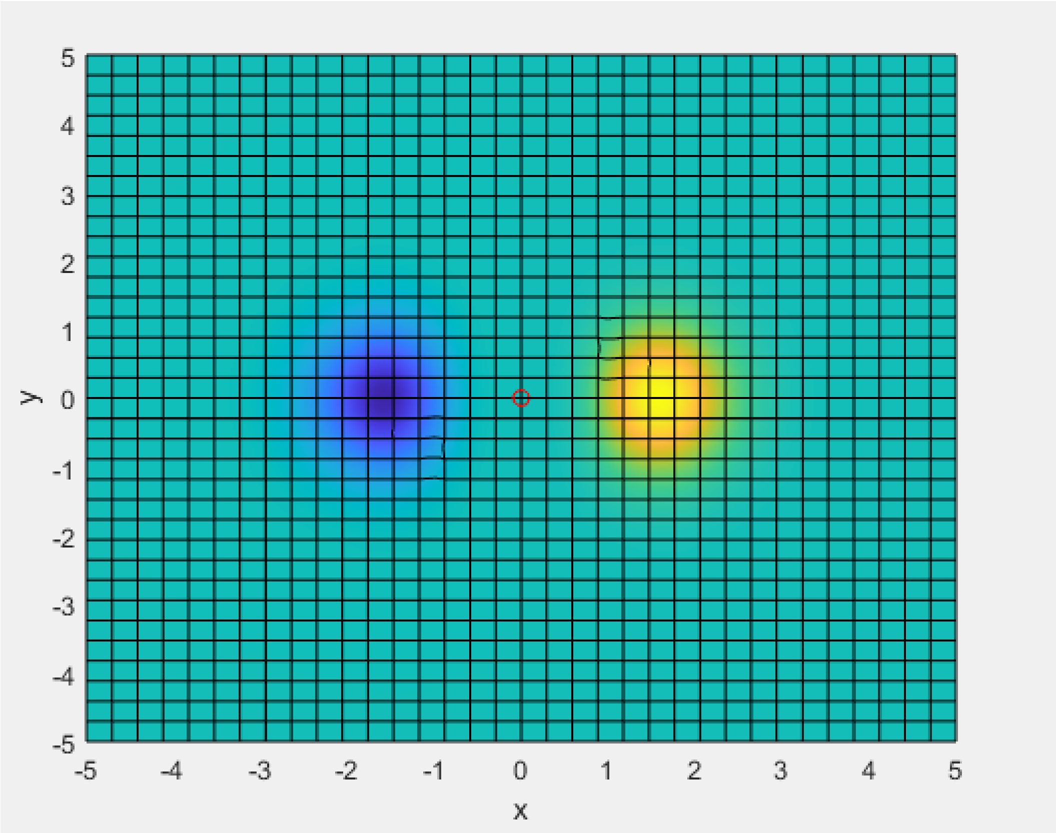
Σύμφωνα με τη μέθοδο και τις επιλεγμένες σταθερές, το ελάχιστο βρίσκεται στο σημείο [-1.58115727, 0.00001058], και ο συνολικός αριθμός διαδοχικών σημείων είναι k = 14. Εμφανίζουμε, ακόμη, τη γραφική παράσταση της αριθμητικής τιμής της συνάρτησης f(x,y) στα σημεία αυτά, συναρτήσει του βήματος k:



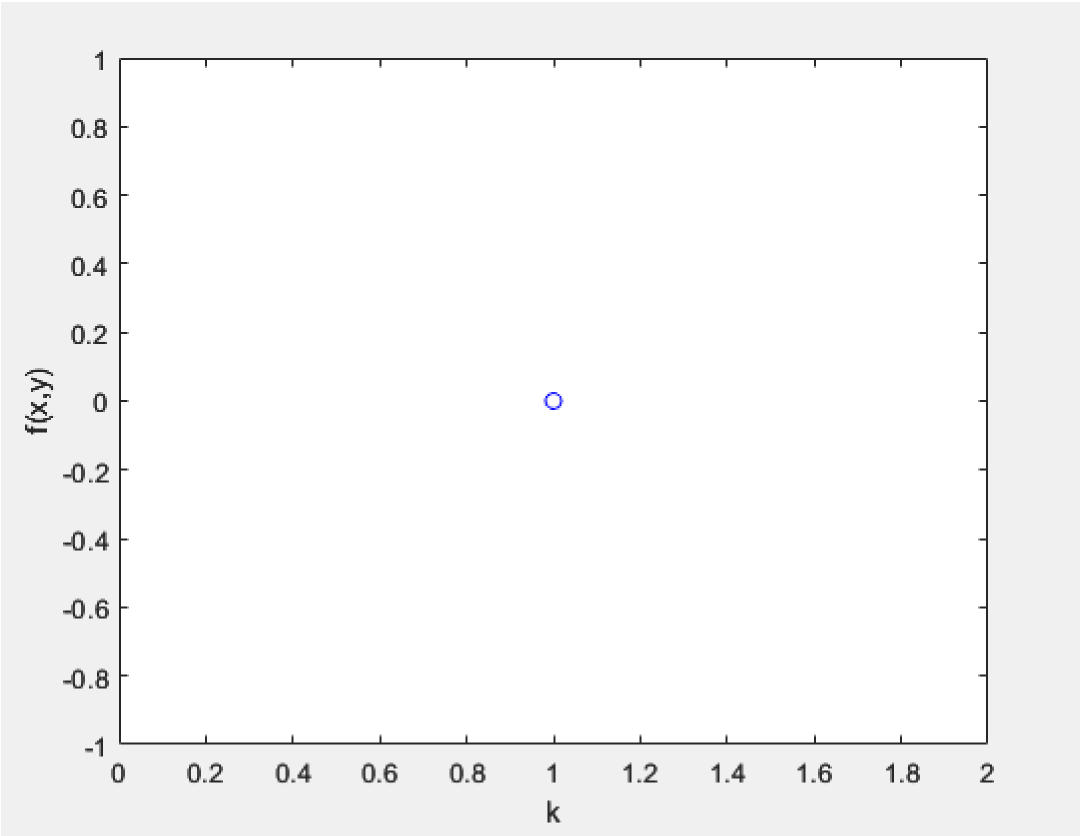
Βλέπουμε ότι τα σημεία πράγματι συγκλίνουν στο ελάχιστο, όσο αυξάνεται το k. Μάλιστα, παρατηρώντας τη γραφική παράσταση της f, φαίνεται ότι το ελάχιστο αυτό είναι το ολικό της. Η σύγκλιση εδώ είναι πολύ γρήγορη, καθώς το gk μεταβάλλεται κάθε φορά, με βάση τον κανόνα Armijo. Εδώ θα πρέπει να προσέξουμε η αρχική τιμή του s να μην είναι ούτε πολύ μεγάλη, ούτε πολύ μικρή, μιας και ενδέχεται να οδηγήσει την όλη διαδικασία σε αστάθεια, από το πρώτο βήμα.

<u>Σημείο εκκίνησης [0 0]</u>

Α, Β, Γ) Θέτοντας σταθερό gk = 0.6, τρέχουμε τον αλγόριθμο της μέγιστης καθόδου και εμφανίζουμε τη σύγκλιση των διαδοχικών σημείων μέχρι την εύρεση του ελαχίστου, για σφάλμα e = 0.0001:



Σύμφωνα με τη μέθοδο και τις επιλεγμένες σταθερές, το ελάχιστο βρίσκεται στο σημείο [0, 0], και ο συνολικός αριθμός διαδοχικών σημείων είναι k = 1. Αυτό δε θα πρέπει να μας εκπλήσσει, καθώς, εφαρμόζοντας ως αρχικό σημείο το [0, 0], ικανοποιείται αυτομάτως η συνθήκη για τον τερματισμό της συνάρτησης. Μάλιστα, το σημείο αυτό αποτελεί και τοπικό ελάχιστο της f. Εμφανίζουμε, ακόμη, τη γραφική παράσταση της αριθμητικής τιμής της συνάρτησης f(x,y) στο σημείο αυτό, συναρτήσει του βήματος k:

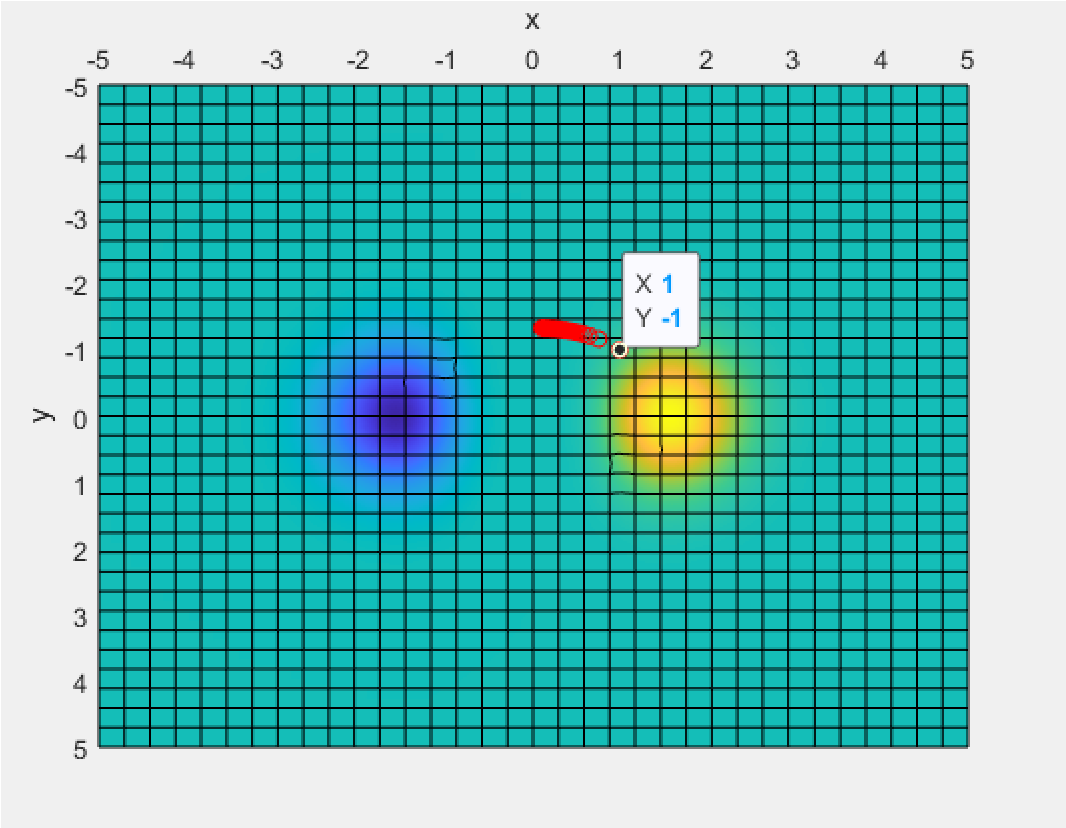


Προφανώς f(0,0) = 0. Το σημείο αυτό είναι τοπικό ελάχιστο της συνάρτησης στη περιοχή κοντά στην αρχή των αξόνων.

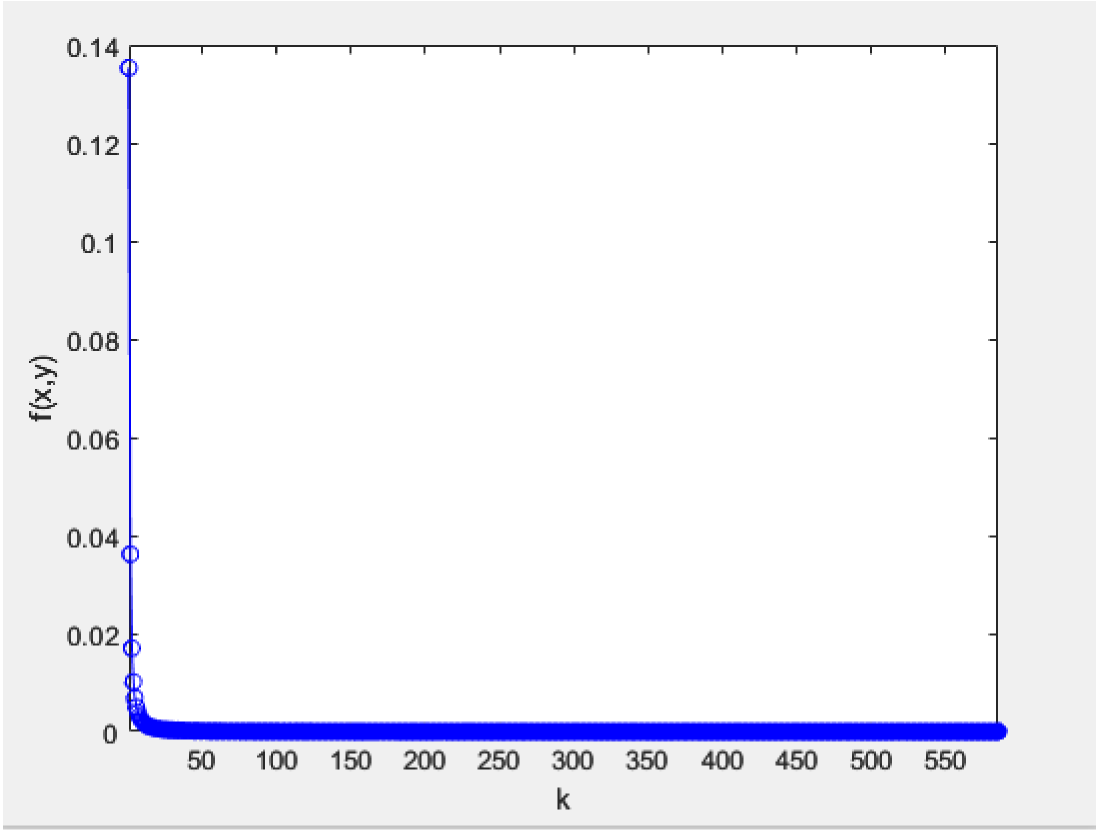
Ακριβώς τα ίδια αποτελέσματα - συμπεράσματα θα εμφανιστούν και στις άλλες δύο μεθόδους μέγιστης καθόδου, αυτής με gk ώστε να ελαχιστοποιεί την f(xk + gk*dk), και αυτής με gk με βάση τον κανόνα Armijo.

<u>Σημείο εκκίνησης [1 -1]</u>

A) Θέτοντας σταθερό gk = 0.6, τρέχουμε τον αλγόριθμο της μέγιστης καθόδου και εμφανίζουμε τη σύγκλιση των διαδοχικών σημείων μέχρι την εύρεση του ελαχίστου, για σφάλμα e = 0.0001:

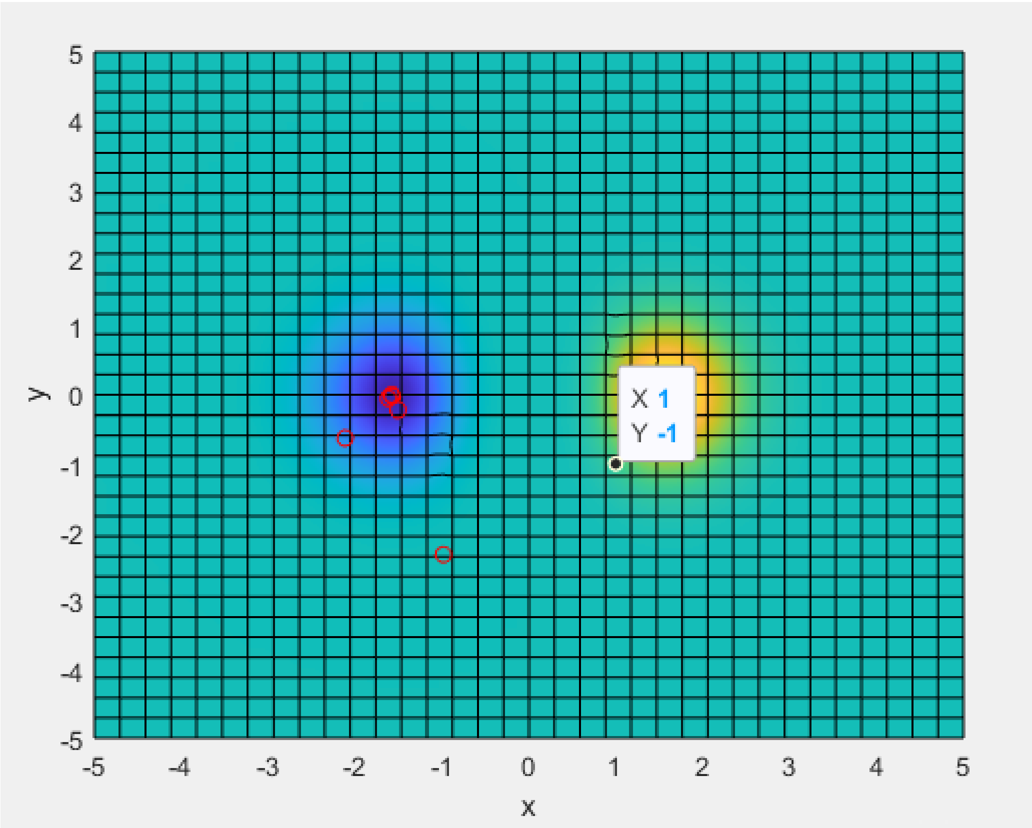


Σύμφωνα με τη μέθοδο και τις επιλεγμένες σταθερές, το ελάχιστο βρίσκεται στο σημείο [0.10436734, -1.32945256], και ο συνολικός αριθμός διαδοχικών σημείων είναι k = 585. Εμφανίζουμε, ακόμη, τη γραφική παράσταση της αριθμητικής τιμής της συνάρτησης f(x,y) στα σημεία αυτά, συναρτήσει του βήματος k:

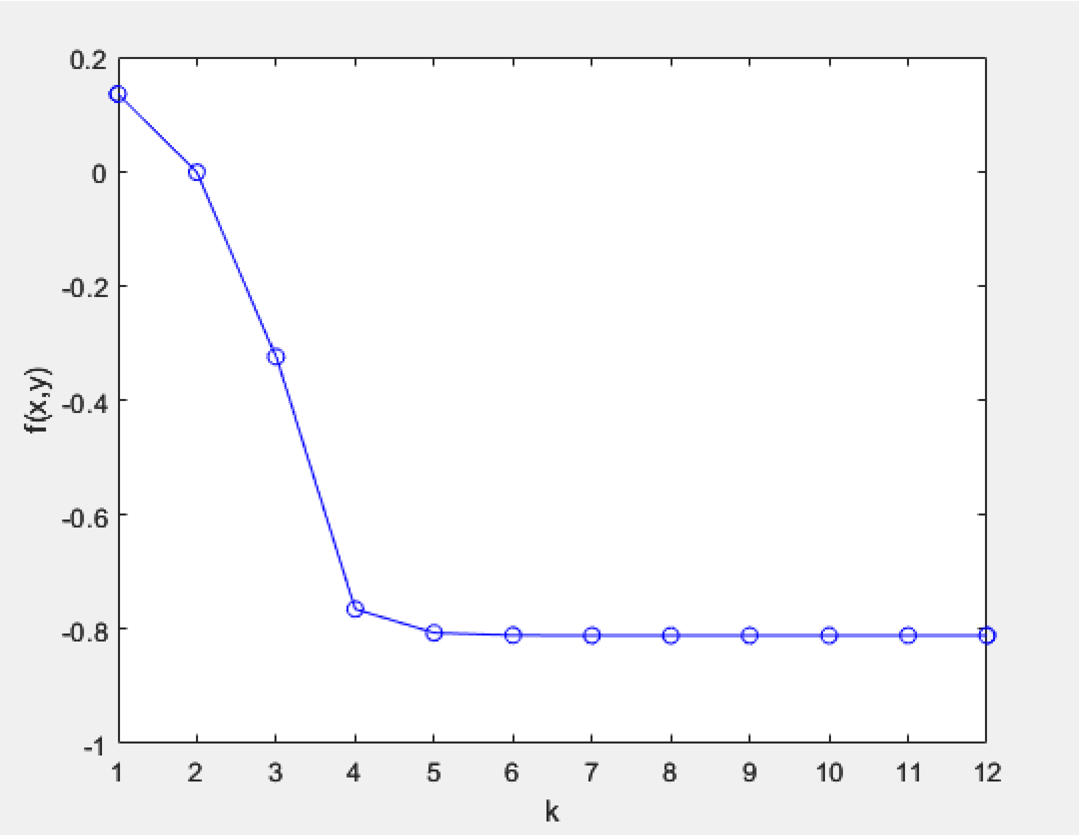


Βλέπουμε ότι τα σημεία συγκλίνουν στο μηδέν, όσο αυξάνεται το k. Δηλαδή ο αλγόριθμος δε συγκλίνει στο ολικό ελάχιστο της f, αλλά σε μία ισοσταθμική της, όπου η τιμή της είναι μηδέν. Αυτό συμβαίνει, καθώς ξεκινήσαμε από μια περιοχή μακριά από το ολικό ελάχιστο, έξω από την 'κοιλάδα' της f, και συγκεκριμένα μέσα στο 'όρος' της. Η σύγκλιση, βέβαια, είναι αρκετά αργή, εξαιτίας της φύσης του αλγορίθμου. Το βήμα gk δεν επιλέχθηκε πραγματικά τυχαία, αλλά τέτοιο ώστε να οδηγήσει σε σύγκλιση του αλγορίθμου στο ελάχιστο. Για μεγαλύτερες ή μικρότερες τιμές του gk μπορεί κανείς να διαπιστώσει ότι η μέθοδος δε θα συγκλίνει.

B) Σε κάθε επανάληψη του αλγορίθμου, υπολογίζουμε το gk, ώστε να ελαχιστοποιείται η f(xk + gk*dk). Τρέχουμε τον αλγόριθμο της μέγιστης καθόδου και εμφανίζουμε τη σύγκλιση των διαδοχικών σημείων μέχρι την εύρεση του ελαχίστου, για σφάλμα e = 0.0001:

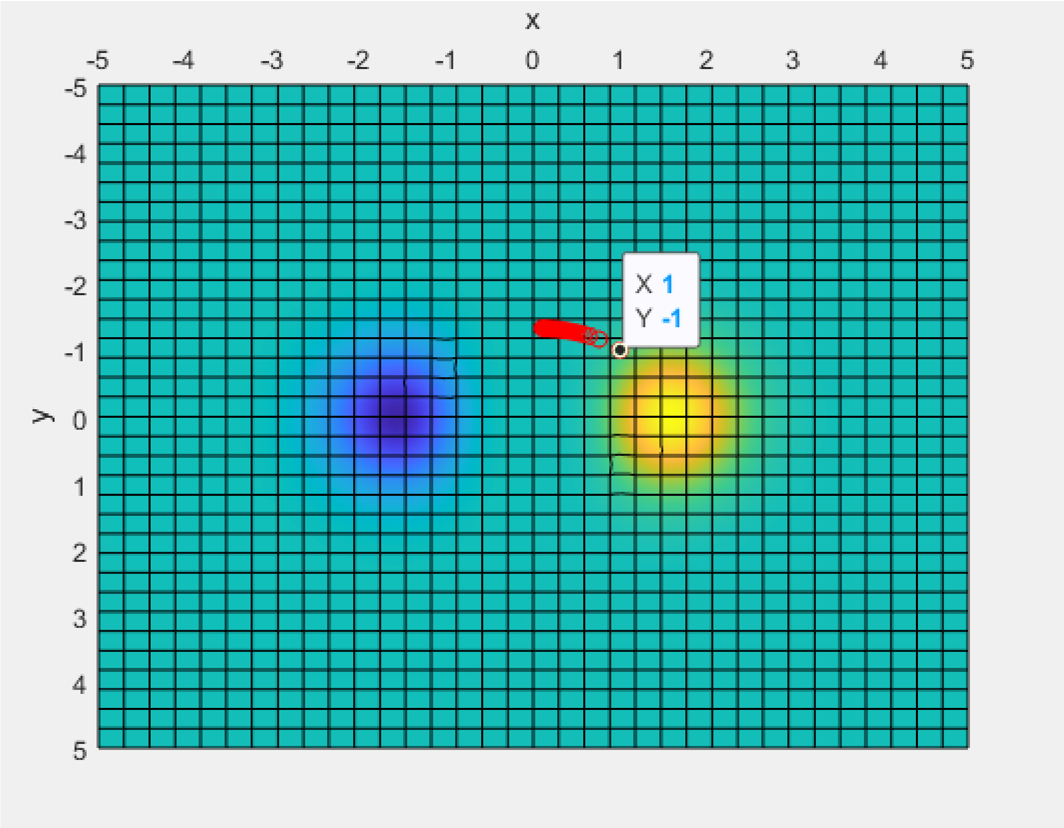


Σύμφωνα με τη μέθοδο και τις επιλεγμένες σταθερές, το ελάχιστο βρίσκεται στο σημείο [-1.58113251, -0.00001894], και ο συνολικός αριθμός διαδοχικών σημείων είναι k = 12. Βλέπουμε, ακόμη, ότι τα διαδοχικά σημεία σχηματίζουν κάθετα μεταξύ τους διανύσματα, το οποίο είναι λογικό, από την επιλογή του μεγέθους gk. Εμφανίζουμε, ακόμη, τη γραφική παράσταση της αριθμητικής τιμής της συνάρτησης f(x,y) στα σημεία αυτά, συναρτήσει του βήματος k:

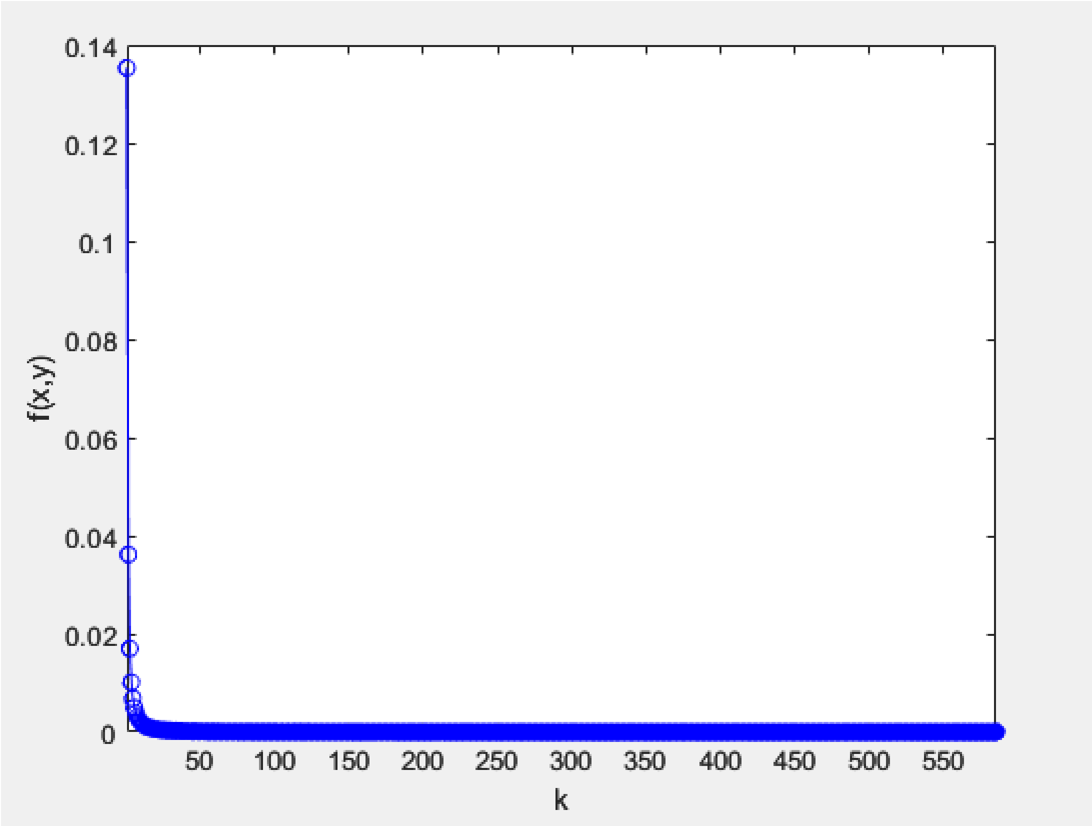


Βλέπουμε ότι τα σημεία πράγματι συγκλίνουν στο ελάχιστο, όσο αυξάνεται το k. Μάλιστα, παρατηρώντας τη γραφική παράσταση της f, φαίνεται ότι το ελάχιστο αυτό είναι το ολικό της. Η σύγκλιση εδώ είναι πολύ πιο γρήγορη, μιας και θέτουμε gk κάθε φορά τέτοιο ώστε να ελαχιστοποιεί την f(xk + gk*dk). Παρά το γεγονός ότι ξεκινήσαμε από μία περιοχή εκτός της ‘κοιλάδας’ της f, συγκεκριμένα μέσα από το ‘όρος’ της, η μέθοδος κατάφερε να περάσει από το τοπικό ελάχιστο της αρχής των αξόνων και να βρεθεί στο ολικό ελάχιστο. Δηλαδή η μέθοδος αυτή πάντοτε θα συγκλίνει στο ολικό ελάχιστο της f, καθώς χρησιμοποιεί τη βέλτιστη επιλογή του gk κάθε φορά.

Γ) Υπολογίζουμε, τώρα, το βήμα gk με βάση τον κανόνα Armijo. Με αρχικές συνθήκες a = 0.05, b = 0.3 και s = 0.6, τρέχουμε τον αλγόριθμο της μέγιστης καθόδου και εμφανίζουμε τη σύγκλιση των διαδοχικών σημείων μέχρι την εύρεση του ελαχίστου, για σφάλμα e = 0.0001:



Σύμφωνα με τη μέθοδο και τις επιλεγμένες σταθερές, το ελάχιστο βρίσκεται στο σημείο [0.10436734, -1.32945256], και ο συνολικός αριθμός διαδοχικών σημείων είναι k = 585. Εμφανίζουμε, ακόμη, τη γραφική παράσταση της αριθμητικής τιμής της συνάρτησης f(x,y) στα σημεία αυτά, συναρτήσει του βήματος k:



Βλέπουμε ότι τα σημεία συγκλίνουν στο μηδέν, όσο αυξάνεται το k. Δηλαδή ο αλγόριθμος δε συγκλίνει στο ολικό ελάχιστο της f, αλλά σε μία ισοσταθμική της, όπου η τιμή της είναι μηδέν. Αυτό συμβαίνει, καθώς ξεκινήσαμε από μια περιοχή μακριά από το ολικό ελάχιστο, έξω από την 'κοιλάδα' της f, και συγκεκριμένα μέσα στο 'όρος' της, και γι' αυτό η σύγκλιση είναι αρκετά αργή. Εδώ θα πρέπει να προσέξουμε η αρχική τιμή του s να μην είναι ούτε πολύ μεγάλη, ούτε πολύ μικρή, μιας και ενδέχεται να οδηγήσει την όλη διαδικασία σε αστάθεια, από το πρώτο βήμα.

Θέμα 3

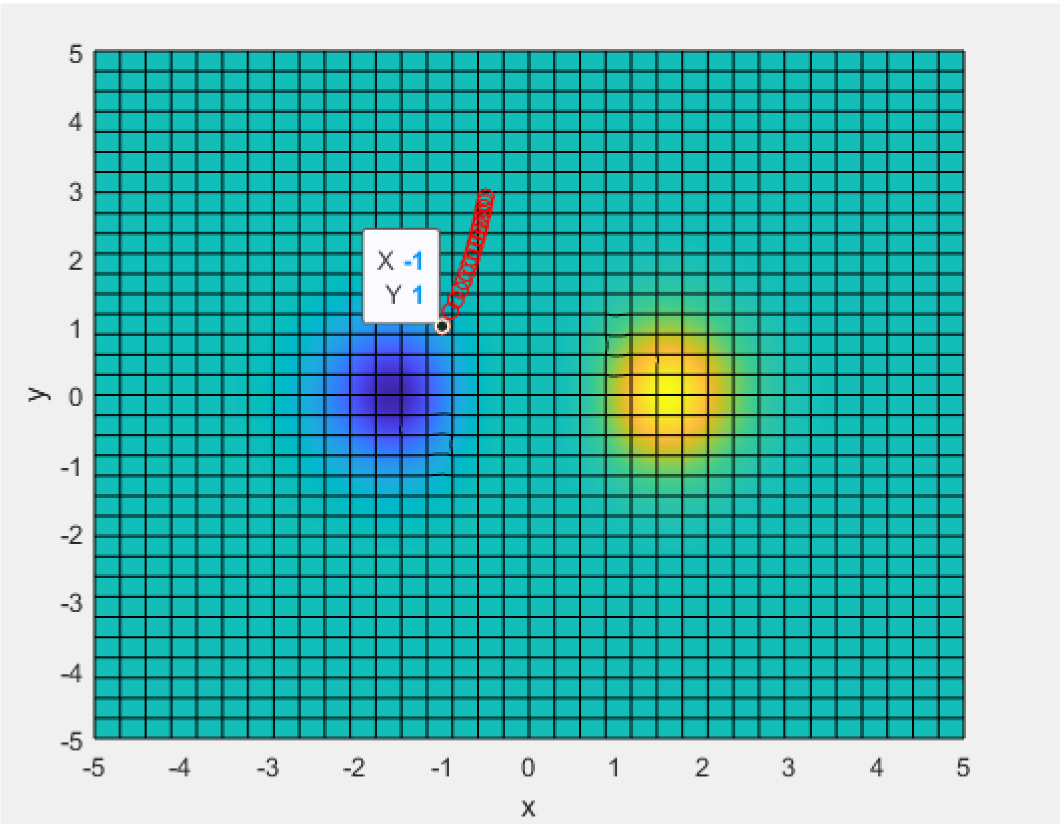
Κατασκευάζουμε τη μέθοδο Newton, με τρεις διαφορετικούς τρόπους, όσον αφορά την επιλογή του βήματος gk: αρχικά, θεωρούμε gk σταθερό, της επιλογής μας, στη συνέχεια το επιλέγουμε τέτοιο ώστε να ελαχιστοποιεί την f(xk + gk*dk), και τέλος, επιλέγεται gk βάση του κανόνα Armijo. Τα σημεία εκκίνησης που ζητούνται είναι: [-1 1], [0 0] και [1 -1].

Για να συγκλίνει σωστά η μέθοδος, πρέπει ο εσσιανός της f, υπολογισμένος στο σημείο Xk, κάθε φορά, να είναι θετικά ορισμένος. Η ιδιότητα αυτή προσδίδει προσδίδει στη μέθοδο Newton την ιδιότητα της επαναληπτικής καθόδου.

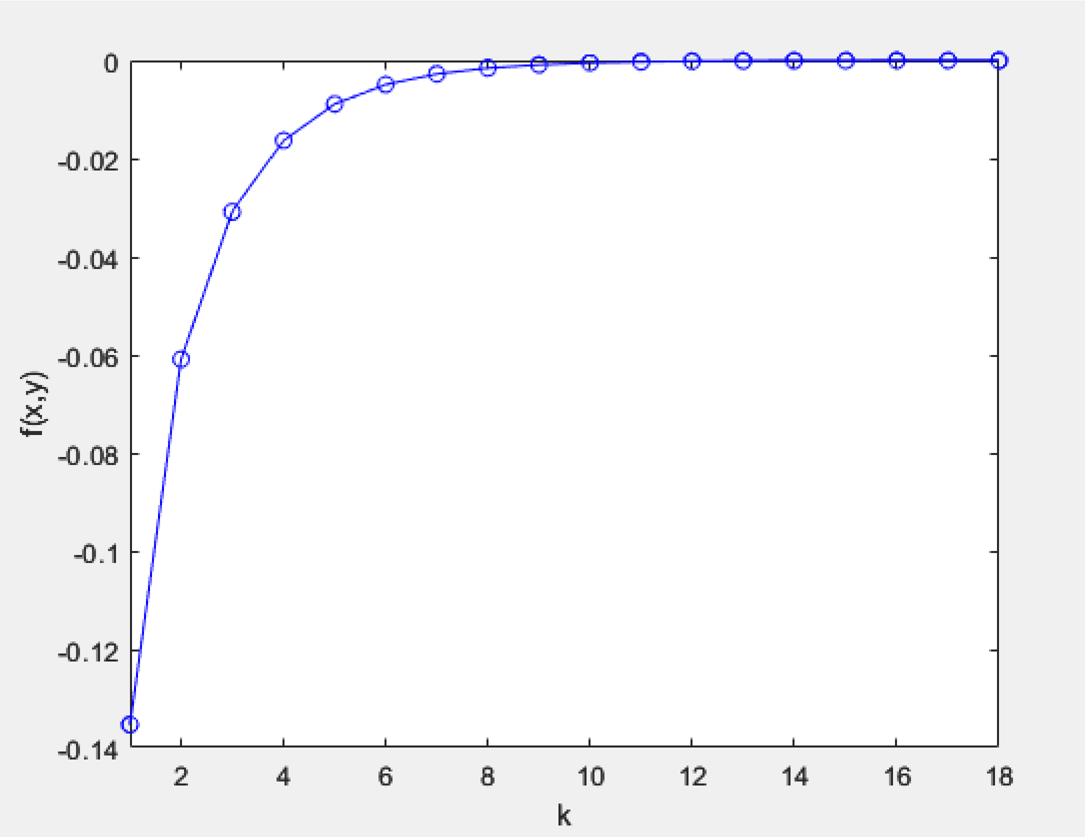
Όμως, μπορούμε εύκολα να υπολογίσουμε ότι ο εσσιανός της f στα σημεία [-1 1] και [1 -1] δεν είναι θετικά ορισμένος. Στο [0 0], ο εσσιανός είναι θετικά ορισμένος. Επομένως, τόσο στο [-1 1], όσο και στο [1, -1], η μέθοδος θα αποκλίνει.

<u>Σημείο εκκίνησης [-1 1]</u>

Α) Θέτοντας σταθερό gk = 0.6, τρέχουμε τον αλγόριθμο Newton και εμφανίζουμε τη σύγκλιση των διαδοχικών σημείων μέχρι την εύρεση του ελαχίστου, για σφάλμα e = 0.0001:

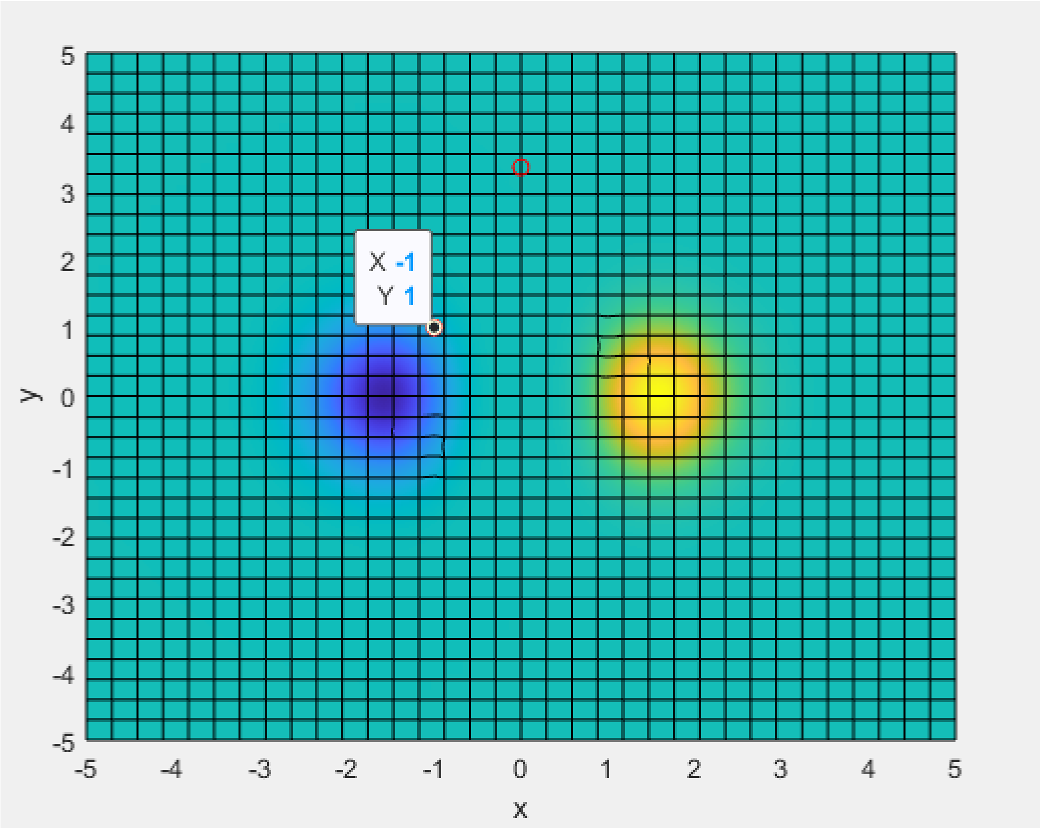


Η μέθοδος εμφανίζει (εσφαλμένο αποτέλεσμα) ως ελάχιστο το σημείο [-0.50059391, 2.86864785], και ο συνολικός αριθμός διαδοχικών σημείων είναι k = 18. Φαίνεται ότι τα διαδοχικά σημεία αποκλίνουν από το ολικό ελάχιστο. Εμφανίζουμε, ακόμη, τη γραφική παράσταση της αριθμητικής τιμής της συνάρτησης f(x,y) στα σημεία αυτά, συναρτήσει του βήματος k:

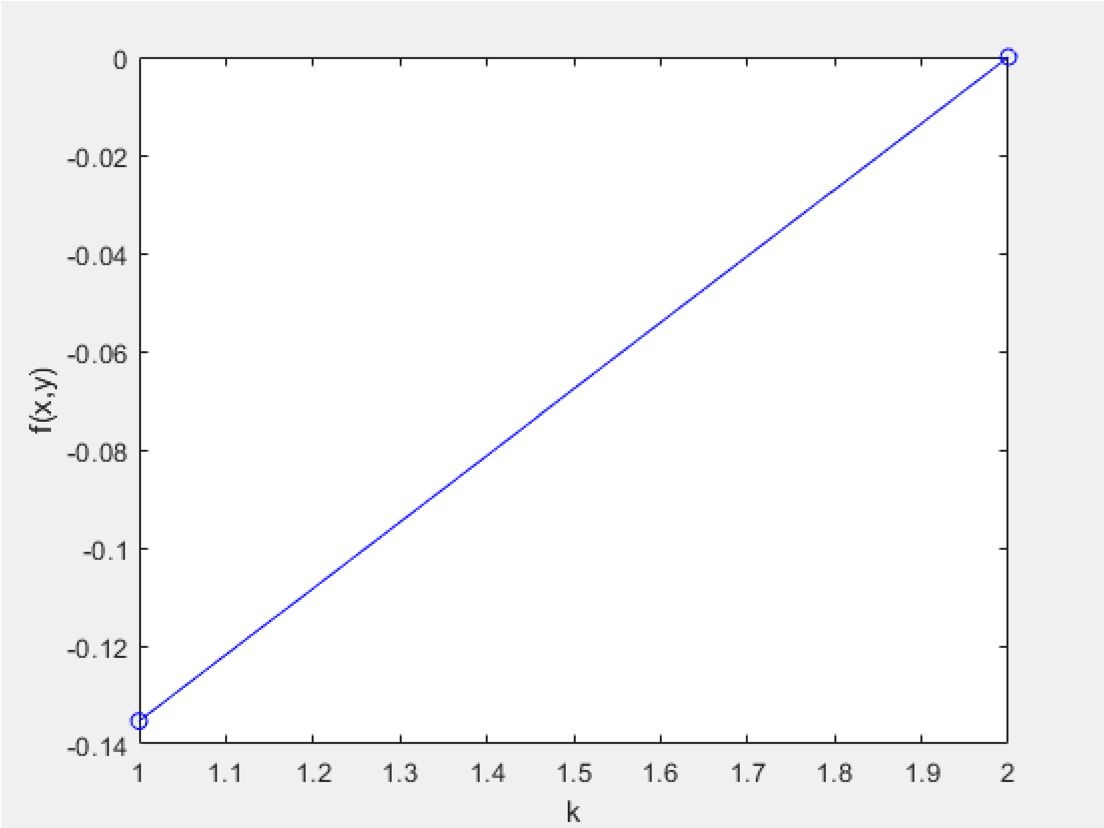


Βλέπουμε ότι τα σημεία πράγματι αποκλίνουν από το ελάχιστο, όσο αυξάνεται το k , καθώς η αριθμητική τιμή της f σε κάθε σημείο είναι μεγαλύτερη από την προηγούμενή της. Φαίνεται, δηλαδή, ότι η ιδιότητα της επαναληπτικής καθόδου δεν ισχύει.

B) Σε κάθε επανάληψη του αλγορίθμου, υπολογίζουμε το gk, ώστε να ελαχιστοποιείται η f(xk + gk*dk). Τρέχουμε τον αλγόριθμο Newton και εμφανίζουμε τη σύγκλιση των διαδοχικών σημείων μέχρι την εύρεση του ελαχίστου, για σφάλμα e = 0.0001:

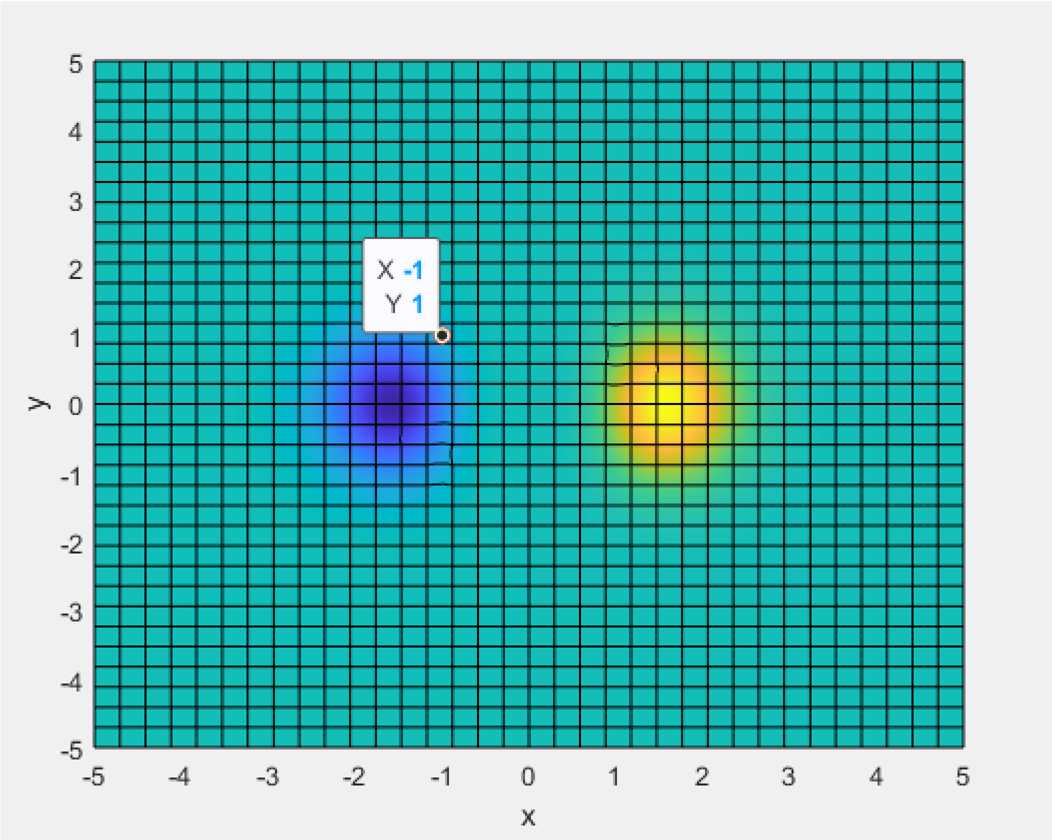


Η μέθοδος εμφανίζει (εσφαλμένο αποτέλεσμα) ως ελάχιστο το σημείο [-3.67341984e-40, 3.33333333], και ο συνολικός αριθμός διαδοχικών σημείων είναι k = 2. Φαίνεται ότι τα διαδοχικά σημεία αποκλίνουν από το ολικό ελάχιστο. Εμφανίζουμε, ακόμη, τη γραφική παράσταση της αριθμητικής τιμής της συνάρτησης f(x,y) στα σημεία αυτά, συναρτήσει του βήματος k:

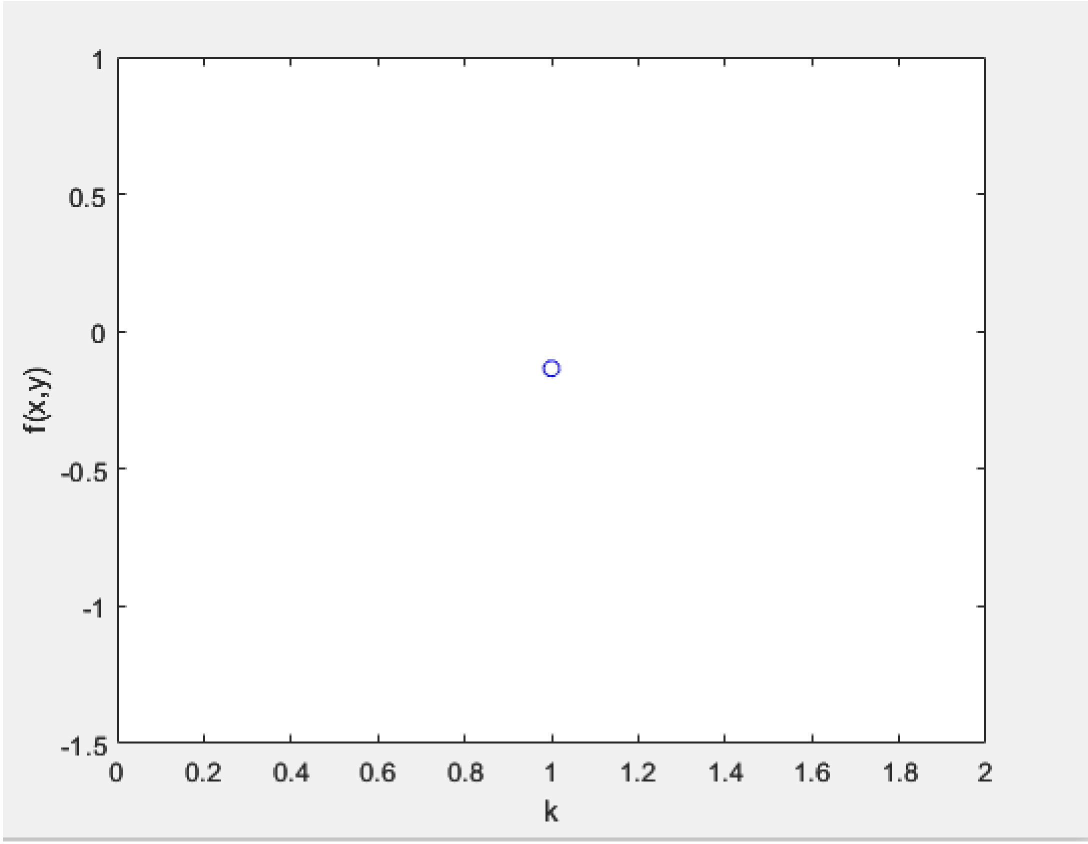


Βλέπουμε ότι τα σημεία πράγματι αποκλίνουν από το ελάχιστο, όσο αυξάνεται το k , καθώς η αριθμητική τιμή της f σε κάθε σημείο είναι μεγαλύτερη από την προηγούμενή της. Φαίνεται, δηλαδή, ότι η ιδιότητα της επαναληπτικής καθόδου δεν ισχύει.

Γ) Υπολογίζουμε, τώρα, το βήμα gk με βάση τον κανόνα Armijo. Με αρχικές συνθήκες a = 0.05, b = 0.3 και s = 0.6, τρέχουμε τον αλγόριθμο Newton και εμφανίζουμε τη σύγκλιση των διαδοχικών σημείων μέχρι την εύρεση του ελαχίστου, για σφάλμα e = 0.0001:



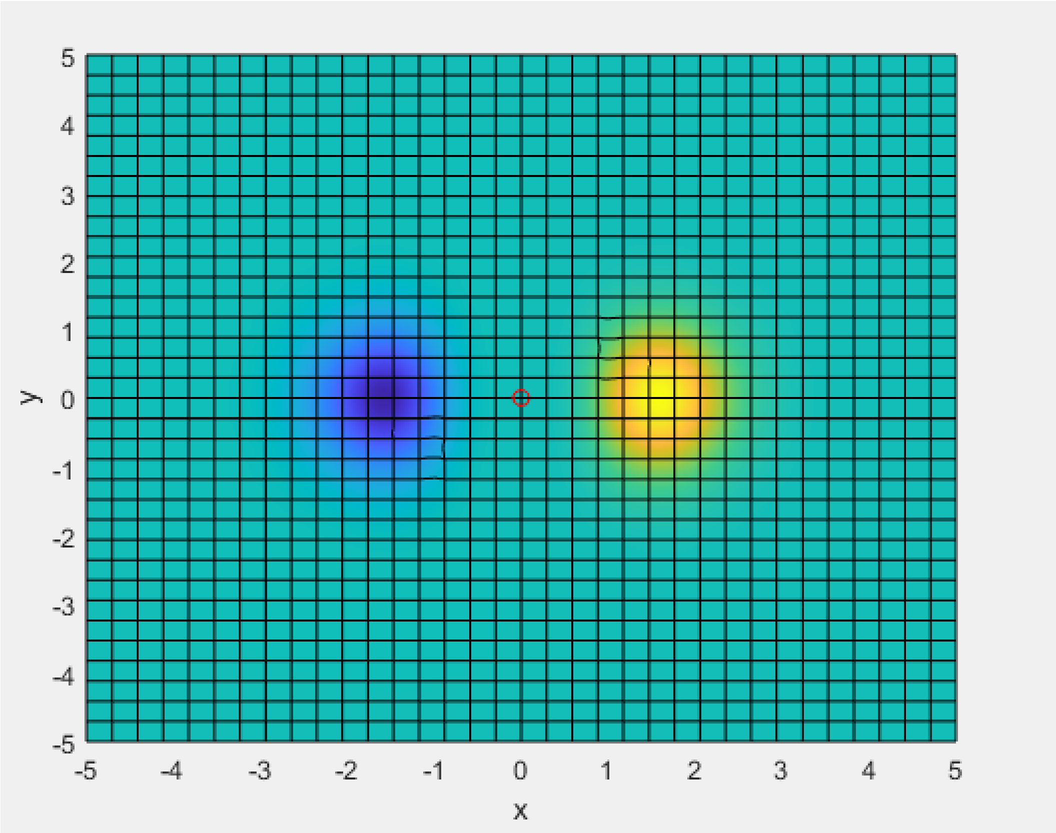
Η μέθοδος, στη πραγματικότητα, δεν τελειώνει ποτέ, καθώς το επόμενο διαδοχικό σημείο είναι τόσο κοντά στο προηγούμενο, που το πρόγραμμα Matlab τα ταυτίζει. Έτσι, εμφανίζει (εσφαλμένο αποτέλεσμα) ως ελάχιστο το σημείο [-1, 1], δηλαδή το σημείο εκκίνησης. Φαίνεται ότι τα άπειρα διαδοχικά σημεία δε θα φτάσουν ποτέ στο ολικό ελάχιστο. Εμφανίζουμε, ακόμη, τη γραφική παράσταση της αριθμητικής τιμής της συνάρτησης f(x,y) στα σημεία αυτά, συναρτήσει του βήματος k:



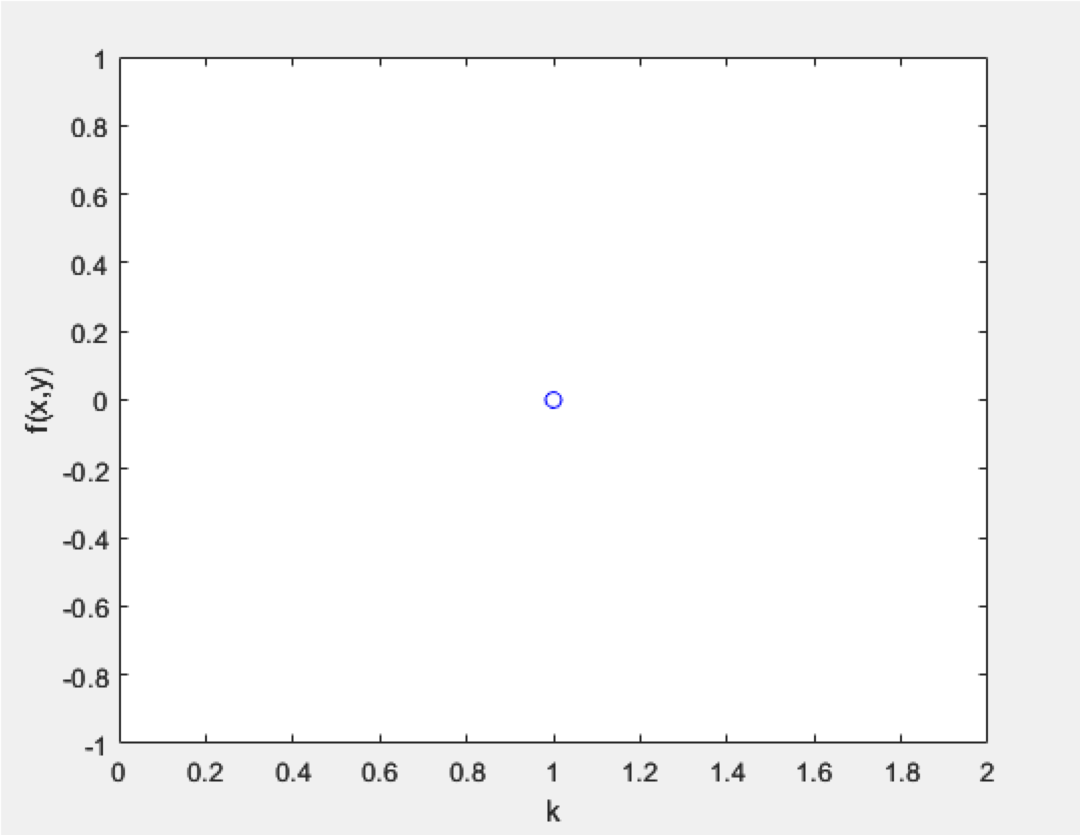
Βλέπουμε ότι τα σημεία πράγματι δε θα φτάσουν ποτέ στο ελάχιστο, όσο κι αν αυξάνεται το k , καθώς η αριθμητική τιμή της f παραμένει διαρκώς ίδια. Η ιδιότητα της επαναληπτικής καθόδου δεν ισχύει.

<u>Σημείο εκκίνησης [0 0]</u>

Α, Β, Γ) Θέτοντας σταθερό gk = 0.6, τρέχουμε τον αλγόριθμο Newton και εμφανίζουμε τη σύγκλιση των διαδοχικών σημείων μέχρι την εύρεση του ελαχίστου, για σφάλμα e = 0.0001:



Σύμφωνα με τη μέθοδο και τις επιλεγμένες σταθερές, το ελάχιστο βρίσκεται στο σημείο [0, 0], και ο συνολικός αριθμός διαδοχικών σημείων είναι k = 1. Αυτό δε θα πρέπει να μας εκπλήσσει, καθώς, εφαρμόζοντας ως αρχικό σημείο το [0, 0], ικανοποιείται αυτομάτως η συνθήκη για τον τερματισμό της συνάρτησης. Μάλιστα, το σημείο αυτό αποτελεί και τοπικό ελάχιστο της f. Εμφανίζουμε, ακόμη, τη γραφική παράσταση της αριθμητικής τιμής της συνάρτησης f(x,y) στο σημείο αυτό, συναρτήσει του βήματος k:

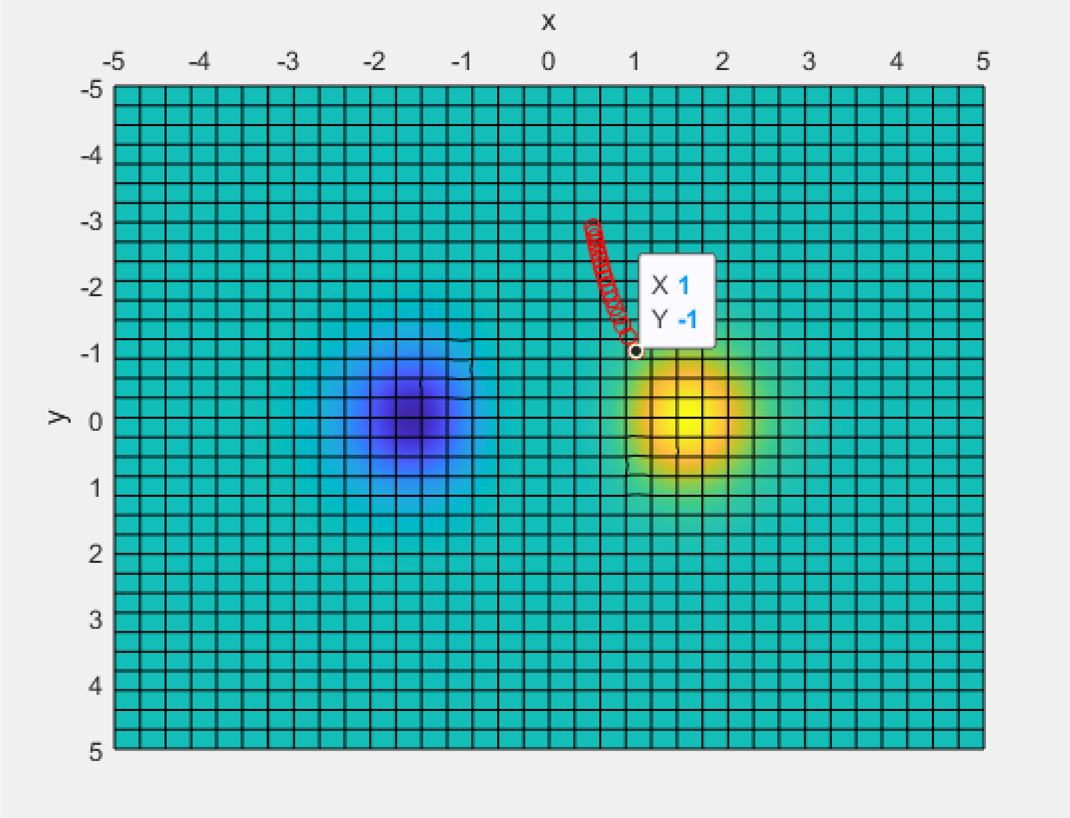


Προφανώς f(0,0) = 0. Το σημείο αυτό είναι τοπικό ελάχιστο της συνάρτησης στη περιοχή κοντά στην αρχή των αξόνων.

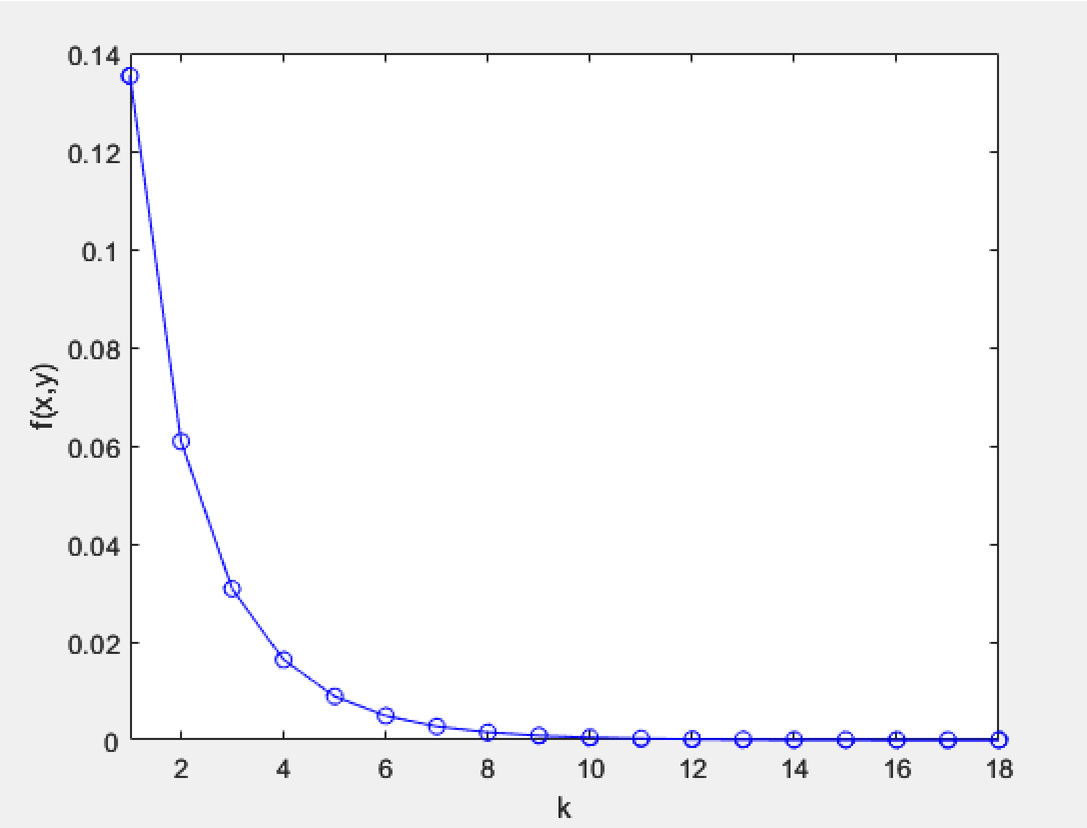
Ακριβώς τα ίδια αποτελέσματα - συμπεράσματα θα εμφανιστούν και στις άλλες δύο μεθόδους Newton, αυτής με gk ώστε να ελαχιστοποιεί την f(xk + gk*dk), και αυτής με gk με βάση τον κανόνα Armijo.

<u>Σημείο εκκίνησης [1 -1]</u>

Α) Θέτοντας σταθερό gk = 0.6, τρέχουμε τον αλγόριθμο Newton και εμφανίζουμε τη σύγκλιση των διαδοχικών σημείων μέχρι την εύρεση του ελαχίστου, για σφάλμα e = 0.0001:

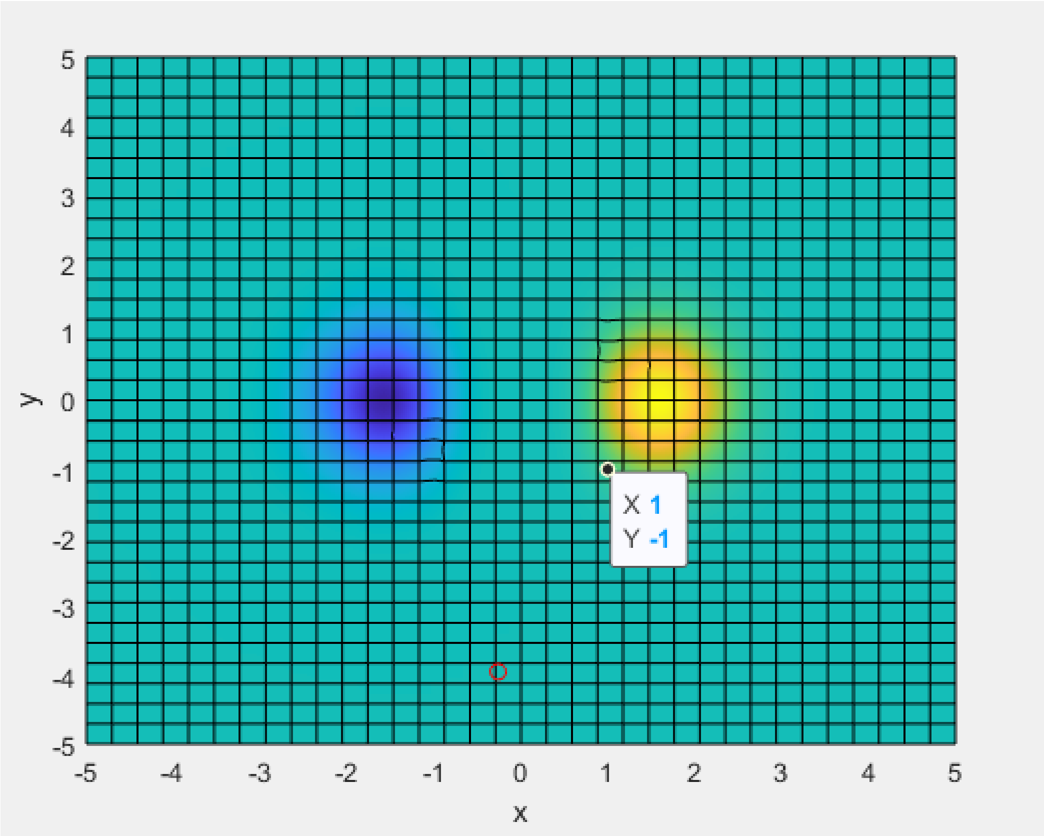


Η μέθοδος εμφανίζει (εσφαλμένο αποτέλεσμα) ως ελάχιστο το σημείο [0.50059391, -2.86864785], και ο συνολικός αριθμός διαδοχικών σημείων είναι k = 18. Φαίνεται ότι τα διαδοχικά σημεία αποκλίνουν από το ολικό ελάχιστο. Εμφανίζουμε, ακόμη, τη γραφική παράσταση της αριθμητικής τιμής της συνάρτησης f(x,y) στα σημεία αυτά, συναρτήσει του βήματος k:

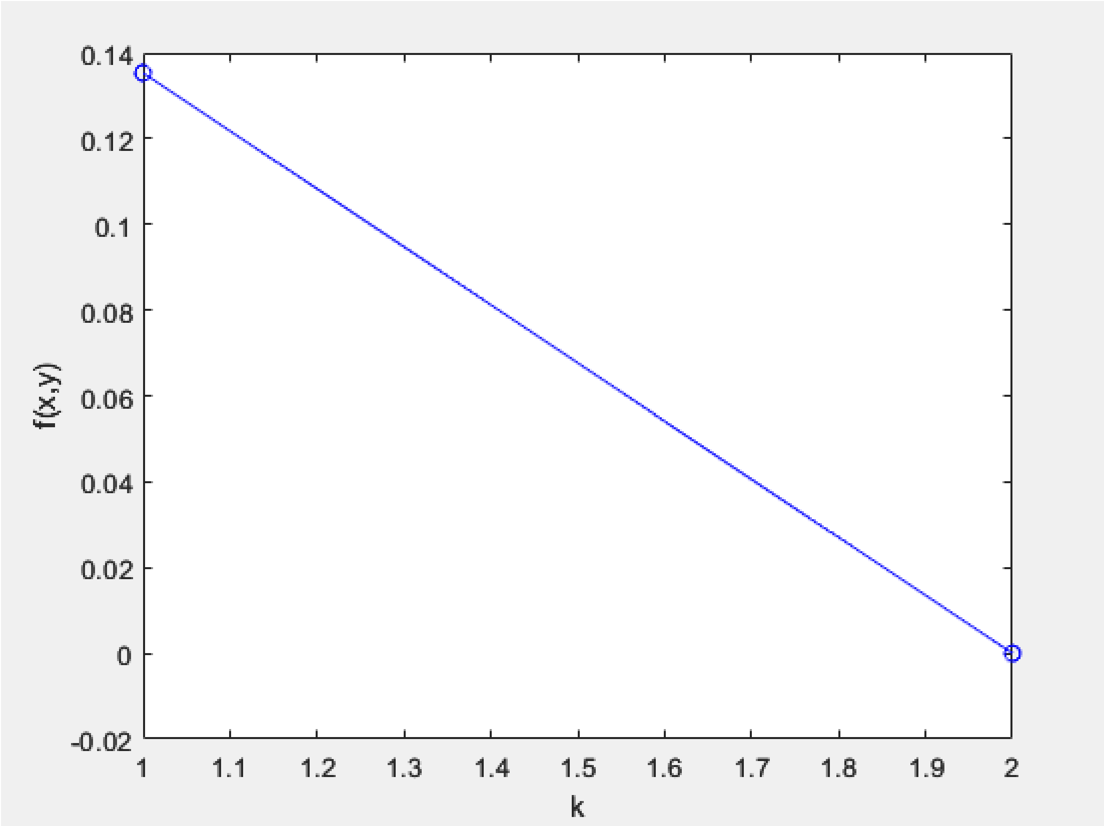


Αν και μπορεί να φαίνεται ότι τα σημεία συγκλίνουν στο τοπικό ελάχιστο 0, όσο αυξάνεται το k, εντούτοις γνωρίζουμε ότι δεν τηρείται η ιδιότητα της επαναληπτικής καθόδου στο [1 -1], και άρα δεν έχουμε πραγματικά σωστό αποτέλεσμα. Το αποτέλεσμα του γραφήματος μπορεί να εξηγηθεί, αν σκεφτούμε ότι ξεκινάμε από μία περιοχή μέσα στο ‘όρος’ της f, οπότε, αποκλίνοντας από το ολικό ελάχιστο, το οποίο βρίσκεται στη ‘κοιλάδα’ της συνάρτησης, αναγκαστικά θα βρεθούμε στο ισοσταθμικό επίπεδο του μηδενός. Έτσι, η αριθμητική τιμή της f διαρκώς μειώνεται, αλλά δεν υπάρχει πραγματική σύγκλιση.

B) Σε κάθε επανάληψη του αλγορίθμου, υπολογίζουμε το gk, ώστε να ελαχιστοποιείται η f(xk + gk*dk). Τρέχουμε τον αλγόριθμο Newton και εμφανίζουμε τη σύγκλιση των διαδοχικών σημείων μέχρι την εύρεση του ελαχίστου, για σφάλμα e = 0.0001:

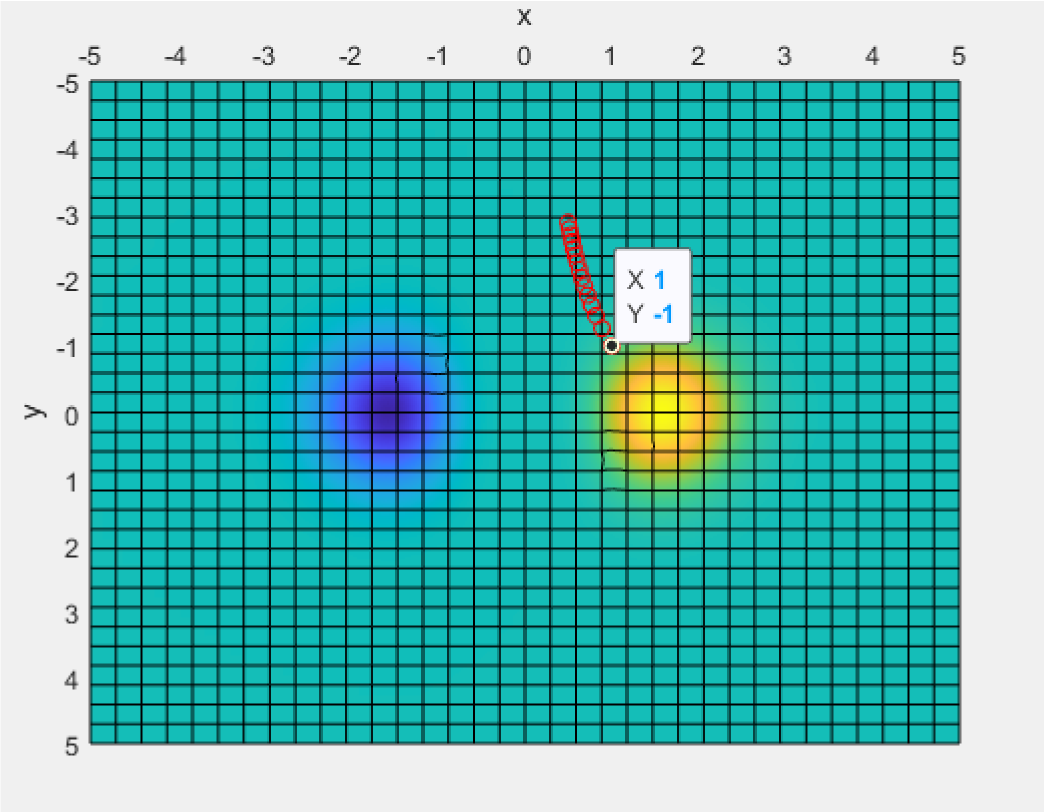


Η μέθοδος εμφανίζει (εσφαλμένο αποτέλεσμα) ως ελάχιστο το σημείο[-0.26377767, -3.94881458], και ο συνολικός αριθμός διαδοχικών σημείων είναι k = 2. Φαίνεται ότι τα διαδοχικά σημεία αποκλίνουν από το ολικό ελάχιστο. Εμφανίζουμε, ακόμη, τη γραφική παράσταση της αριθμητικής τιμής της συνάρτησης f(x,y) στα σημεία αυτά, συναρτήσει του βήματος k:

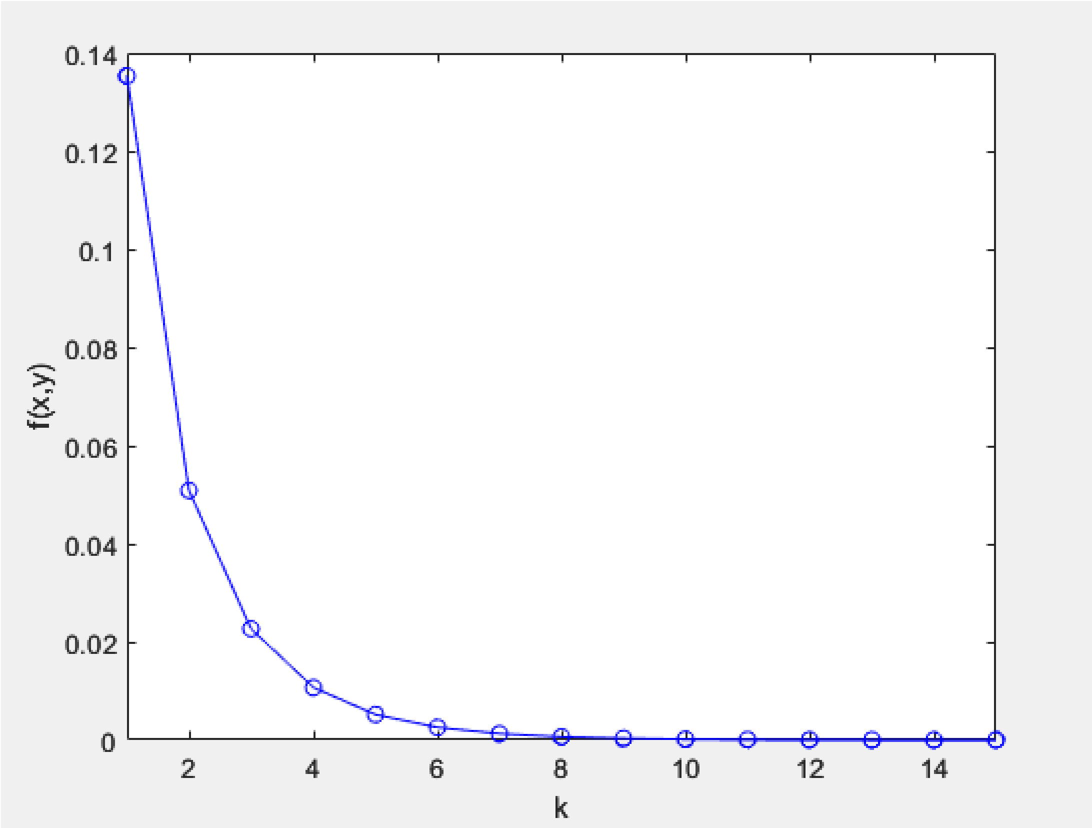


Αν και μπορεί να φαίνεται ότι τα σημεία συγκλίνουν στο τοπικό ελάχιστο 0, όσο αυξάνεται το k, εντούτοις γνωρίζουμε ότι δεν τηρείται η ιδιότητα της επαναληπτικής καθόδου στο [1 -1], και άρα δεν έχουμε πραγματικά σωστό αποτέλεσμα. Το αποτέλεσμα του γραφήματος μπορεί να εξηγηθεί, αν σκεφτούμε ότι ξεκινάμε από μία περιοχή μέσα στο ‘όρος’ της f, οπότε, αποκλίνοντας από το ολικό ελάχιστο, το οποίο βρίσκεται στη ‘κοιλάδα’ της συνάρτησης, αναγκαστικά θα βρεθούμε στο ισοσταθμικό επίπεδο του μηδενός. Έτσι, η αριθμητική τιμή της f διαρκώς μειώνεται, αλλά δεν υπάρχει πραγματική σύγκλιση.

Γ) Υπολογίζουμε, τώρα, το βήμα gk με βάση τον κανόνα Armijo. Με αρχικές συνθήκες a = 0.05, b = 0.3 και s = 0.6, τρέχουμε τον αλγόριθμο Newton και εμφανίζουμε τη σύγκλιση των διαδοχικών σημείων μέχρι την εύρεση του ελαχίστου, για σφάλμα e = 0.0001:



Η μέθοδος εμφανίζει (εσφαλμένο αποτέλεσμα) ως ελάχιστο το σημείο[0.49775402, -2.86118164], και ο συνολικός αριθμός διαδοχικών σημείων είναι k = 15. Φαίνεται ότι τα διαδοχικά σημεία αποκλίνουν από το ολικό ελάχιστο. Εμφανίζουμε, ακόμη, τη γραφική παράσταση της αριθμητικής τιμής της συνάρτησης f(x,y) στα σημεία αυτά, συναρτήσει του βήματος k:



Αν και μπορεί να φαίνεται ότι τα σημεία συγκλίνουν στο τοπικό ελάχιστο 0, όσο αυξάνεται το k, εντούτοις γνωρίζουμε ότι δεν τηρείται η ιδιότητα της επαναληπτικής καθόδου στο [1 -1], και άρα δεν έχουμε πραγματικά σωστό αποτέλεσμα. Το αποτέλεσμα του γραφήματος μπορεί να εξηγηθεί, αν σκεφτούμε ότι ξεκινάμε από μία περιοχή μέσα στο 'όρος' της f, οπότε, αποκλίνοντας από το ολικό ελάχιστο, το οποίο βρίσκεται στη 'κοιλάδα' της συνάρτησης, αναγκαστικά θα βρεθούμε στο ισοσταθμικό επίπεδο του μηδενός. Έτσι, η αριθμητική τιμή της f διαρκώς μειώνεται, αλλά δεν υπάρχει πραγματική σύγκλιση.

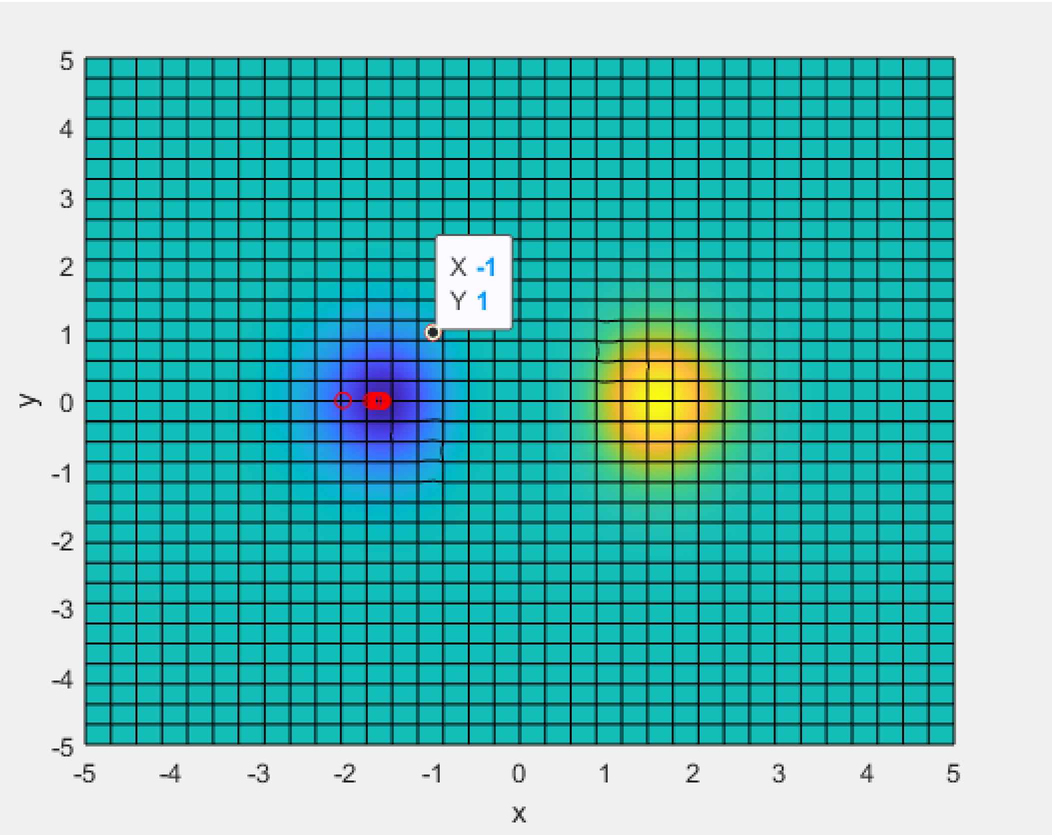
Θέμα 4

Κατασκευάζουμε τη μέθοδο Levenberg - Marquardt, με τρεις διαφορετικούς τρόπους, όσον αφορά την επιλογή του βήματος gk: αρχικά, θεωρούμε gk σταθερό, της επιλογής μας, στη συνέχεια το επιλέγουμε τέτοιο ώστε να ελαχιστοποιεί την f(xk + gk*dk), και τέλος, επιλέγεται gk βάση του κανόνα Armijo. Τα σημεία εκκίνησης που ζητούνται είναι: [-1 1], [0 0] και [1 -1].

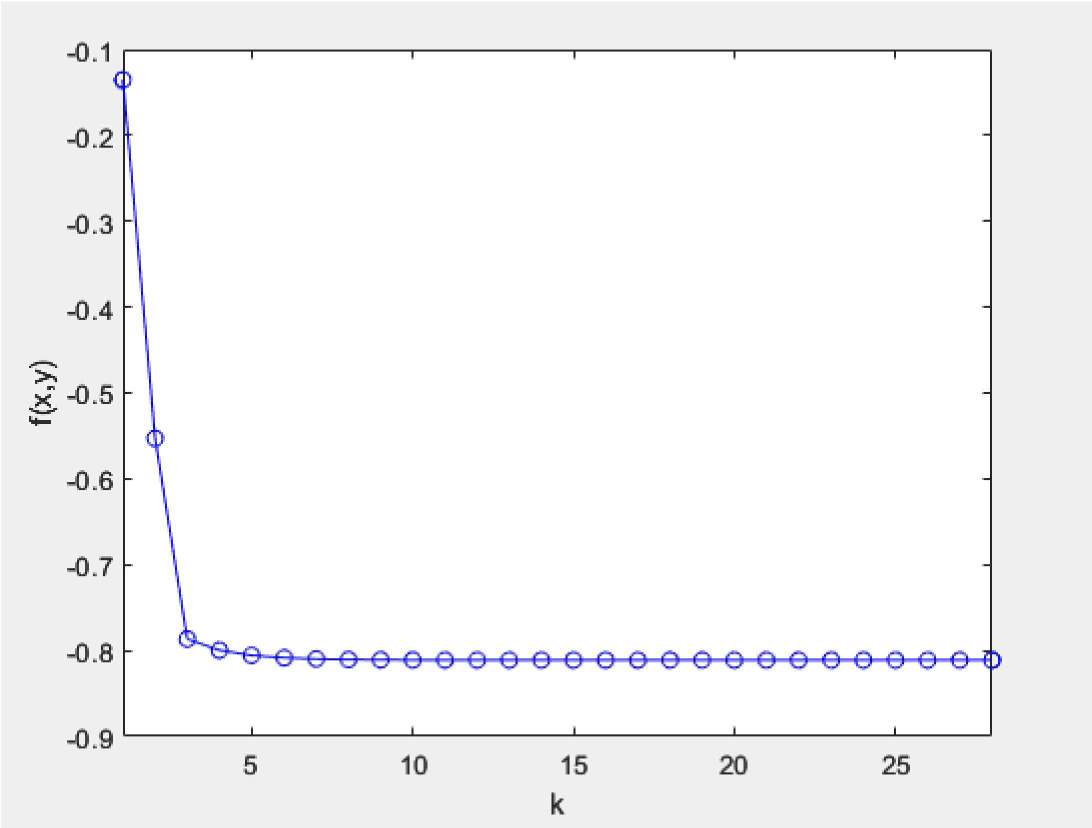
Η μέθοδος αυτή αποτελεί ένα τροποποιημένο αλγόριθμο Newton, όταν ο εσσιανός της f δεν είναι θετικά ορισμένος. Η βασική ιδέα είναι να υπολογίζουμε μια τιμή mk κάθε φορά, έτσι ώστε: [εσσιανός_στο_σημείο_Xk + mk * μοναδιαίος_πίνακας] να είναι θετικά ορισμένος. Για το σκοπό αυτό, επιλέγουμε mk = max(απολυτη_τιμή(μεγαλύτερης_ιδιοτιμής(εσσιανός_στο_σημείο_Xk))) + 0.2. Έτσι, η μέθοδος αυτή θα συγκλίνει σε ελάχιστο (είτε τοπικό, είτε ολικό), ανεξαρτήτως του αρχικού σημείου εκκίνησης, σε αντίθεση με τη μέθοδο Newton.

Σημείο εκκίνησης [-1 1]

A) Θέτοντας σταθερό gk = 0.6, τρέχουμε τον αλγόριθμο Levenberg - Marquardt και εμφανίζουμε τη σύγκλιση των διαδοχικών σημείων μέχρι την εύρεση του ελαχίστου, για σφάλμα e = 0.0001:



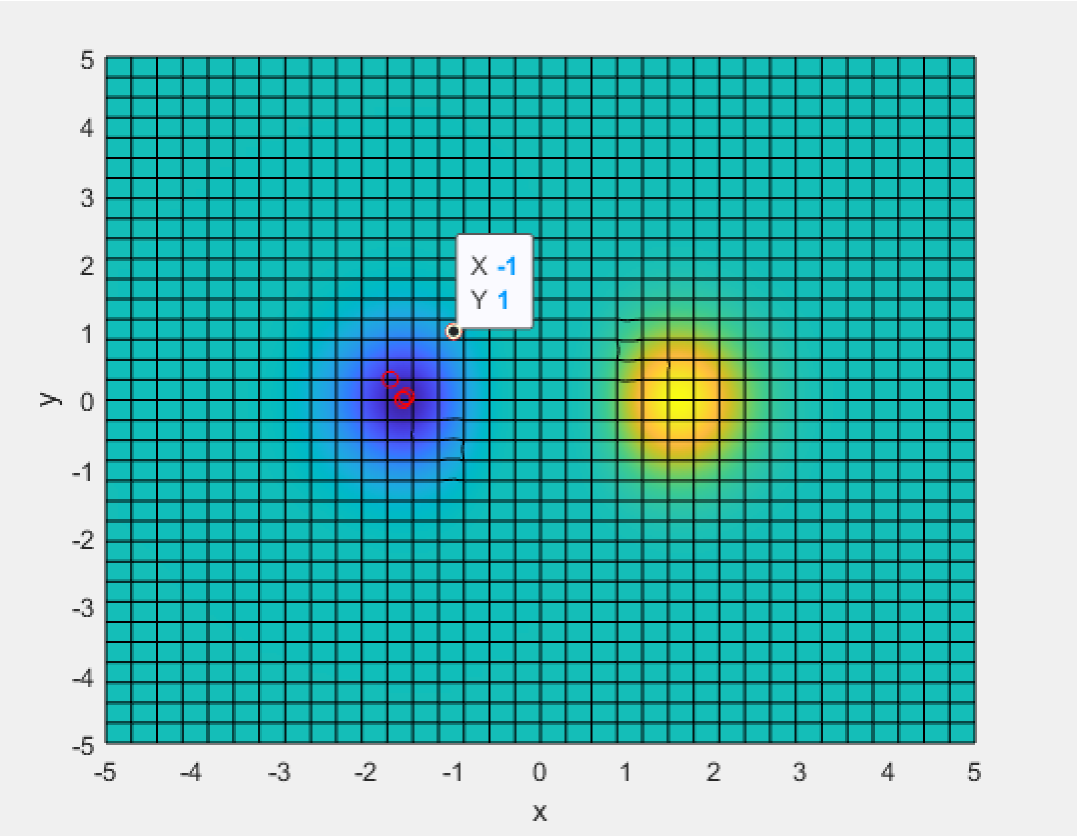
Σύμφωνα με τη μέθοδο και τις επιλεγμένες σταθερές, το ελάχιστο βρίσκεται στο σημείο[-1.58115979, 0.00001559], και ο συνολικός αριθμός διαδοχικών σημείων είναι k = 28. Εμφανίζουμε, ακόμη, τη γραφική παράσταση της αριθμητικής τιμής της συνάρτησης f(x,y) στα σημεία αυτά, συναρτήσει του βήματος k:



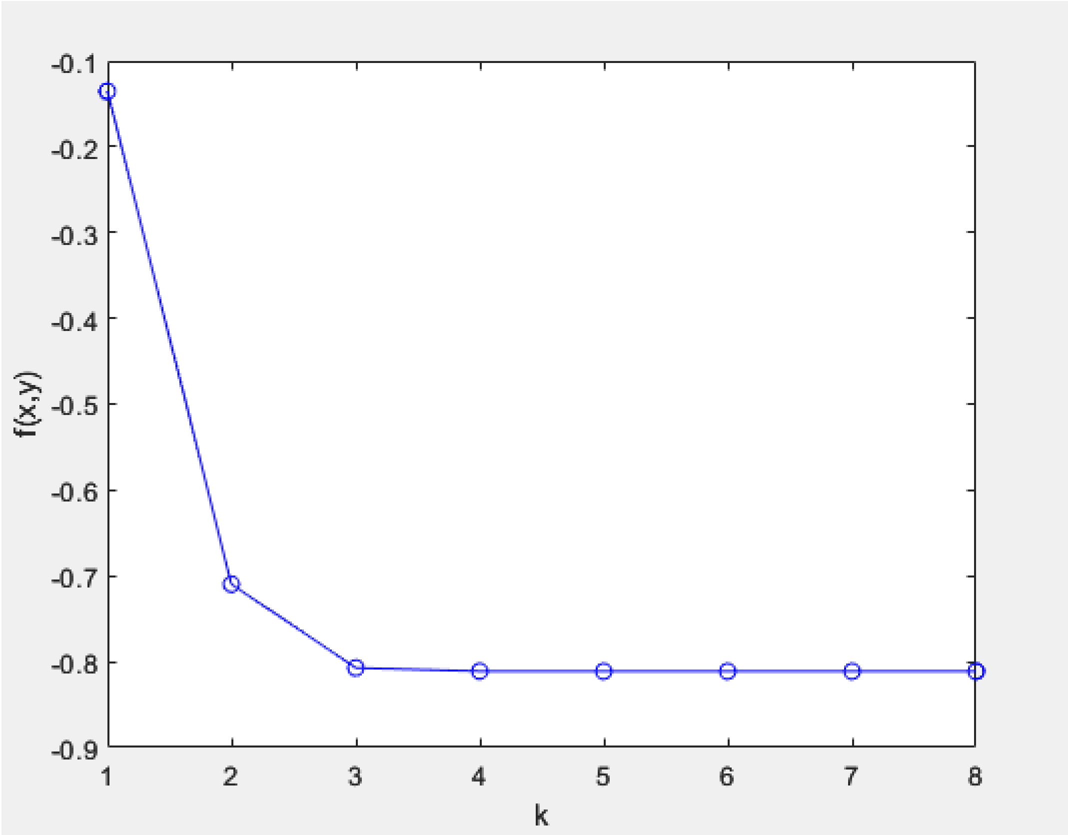
Βλέπουμε ότι τα σημεία πράγματι συγκλίνουν στο ελάχιστο, όσο αυξάνεται το k. Μάλιστα, παρατηρώντας τη γραφική παράσταση της f, φαίνεται ότι το ελάχιστο αυτό είναι το ολικό της. Αυτό συμβαίνει, καθώς ξεκινήσαμε από μια περιοχή σχετικά κοντά στο ολικό ελάχιστο, μέσα στη ‘κοιλάδα’ της f. Η σύγκλιση εδώ είναι σχετικά γρήγορη. Το βήμα gk δεν επιλέχθηκε πραγματικά τυχαία, αλλά τέτοιο ώστε να οδηγήσει σε σύγκλιση του αλγορίθμου στο ελάχιστο. Για μεγαλύτερες ή μικρότερες τιμές του gk μπορεί κανείς να διαπιστώσει ότι η μέθοδος δε θα συγκλίνει.

B) Σε κάθε επανάληψη του αλγορίθμου, υπολογίζουμε το gk, ώστε να ελαχιστοποιείται η f(xk + gk*dk). Τρέχουμε τον αλγόριθμο

Levenberg – Marquardt και εμφανίζουμε τη σύγκλιση των διαδοχικών σημείων μέχρι την εύρεση του ελαχίστου, για σφάλμα e = 0.0001:

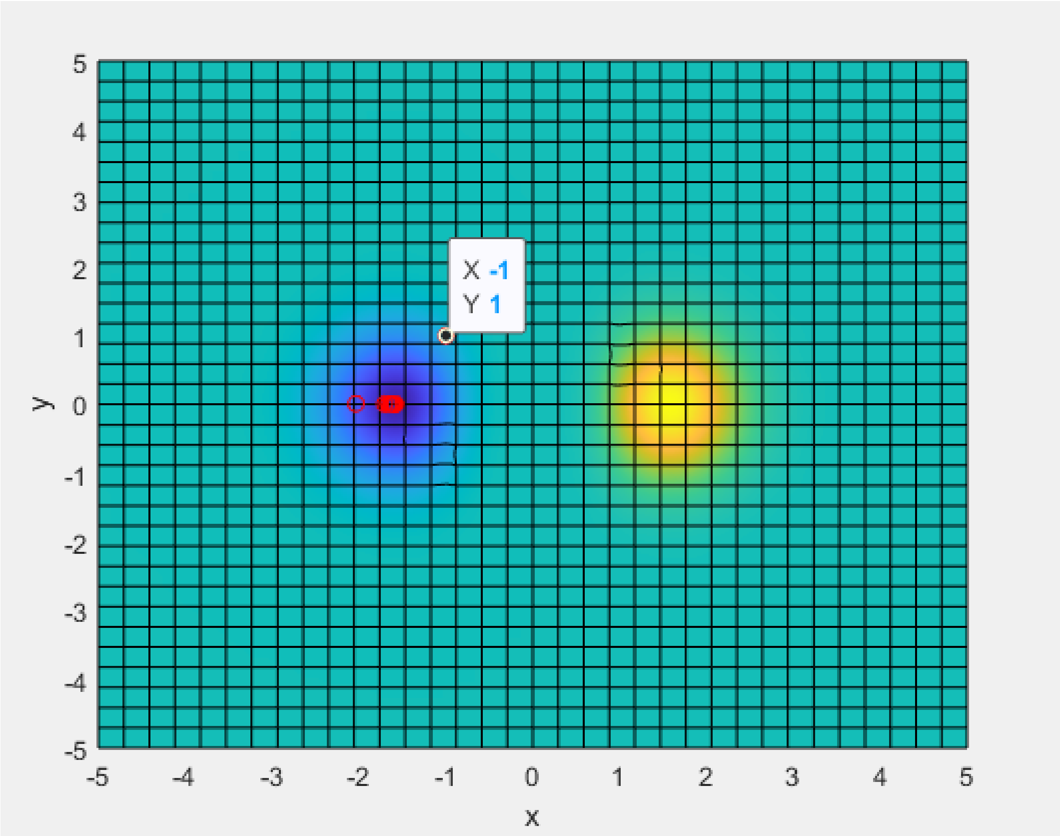


Σύμφωνα με τη μέθοδο και τις επιλεγμένες σταθερές, το ελάχιστο βρίσκεται στο σημείο[-1.58114856, 0.00001974], και ο συνολικός αριθμός διαδοχικών σημείων είναι k = 8. Βλέπουμε, ακόμη, ότι τα διαδοχικά σημεία σχηματίζουν κάθετα μεταξύ τους διανύσματα, το οποίο είναι λογικό, από την επιλογή του μεγέθους gk. Εμφανίζουμε, ακόμη, τη γραφική παράσταση της αριθμητικής τιμής της συνάρτησης f(x,y) στα σημεία αυτά, συναρτήσει του βήματος k:

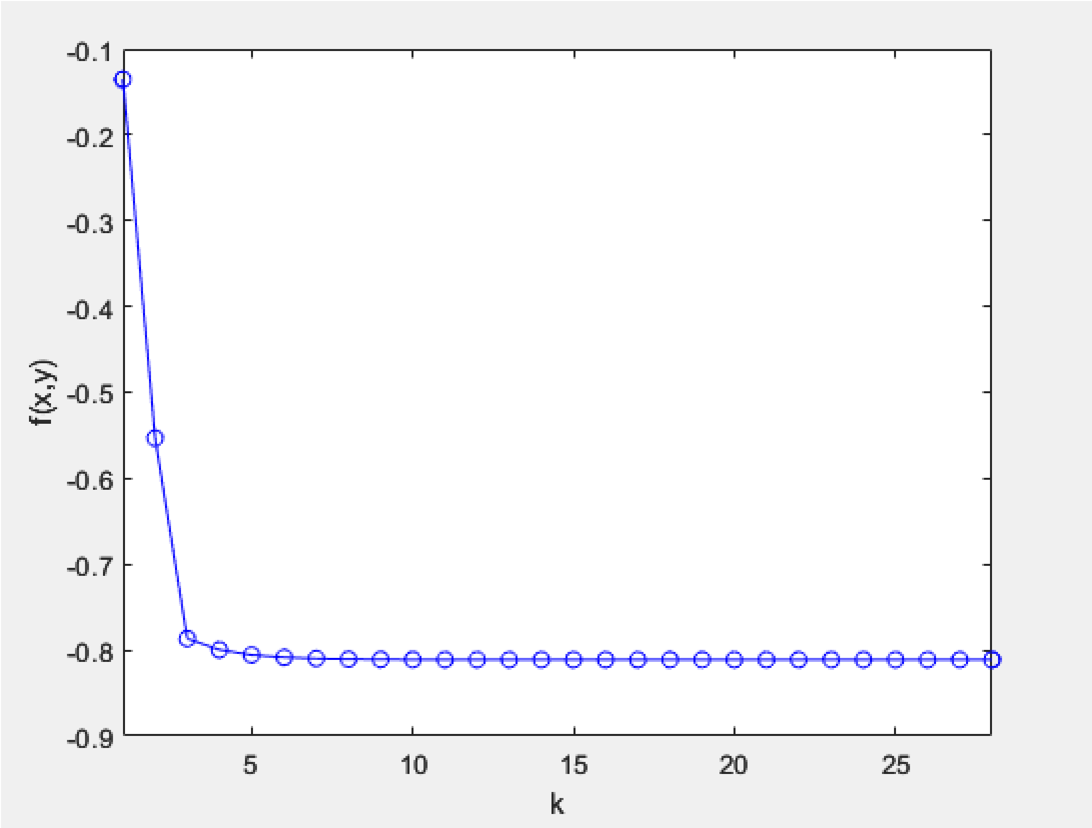


Βλέπουμε ότι τα σημεία πράγματι συγκλίνουν στο ελάχιστο, όσο αυξάνεται το k. Μάλιστα, παρατηρώντας τη γραφική παράσταση της f, φαίνεται ότι το ελάχιστο αυτό είναι το ολικό της. Η σύγκλιση εδώ είναι ακόμα πιο γρήγορη, μιας και θέτουμε gk κάθε φορά τέτοιο ώστε να ελαχιστοποιεί την f(xk + gk*dk). Φαίνεται ότι η μέθοδος αυτή πάντοτε θα συγκλίνει στο ολικό ελάχιστο της f, καθώς χρησιμοποιεί τη βέλτιστη επιλογή του gk κάθε φορά.

Γ) Υπολογίζουμε, τώρα, το βήμα gk με βάση τον κανόνα Armijo. Με αρχικές συνθήκες a = 0.05, b = 0.3 και s = 0.6, τρέχουμε τον αλγόριθμο Levenberg – Marquardt και εμφανίζουμε τη σύγκλιση των διαδοχικών σημείων μέχρι την εύρεση του ελαχίστου, για σφάλμα e = 0.0001:



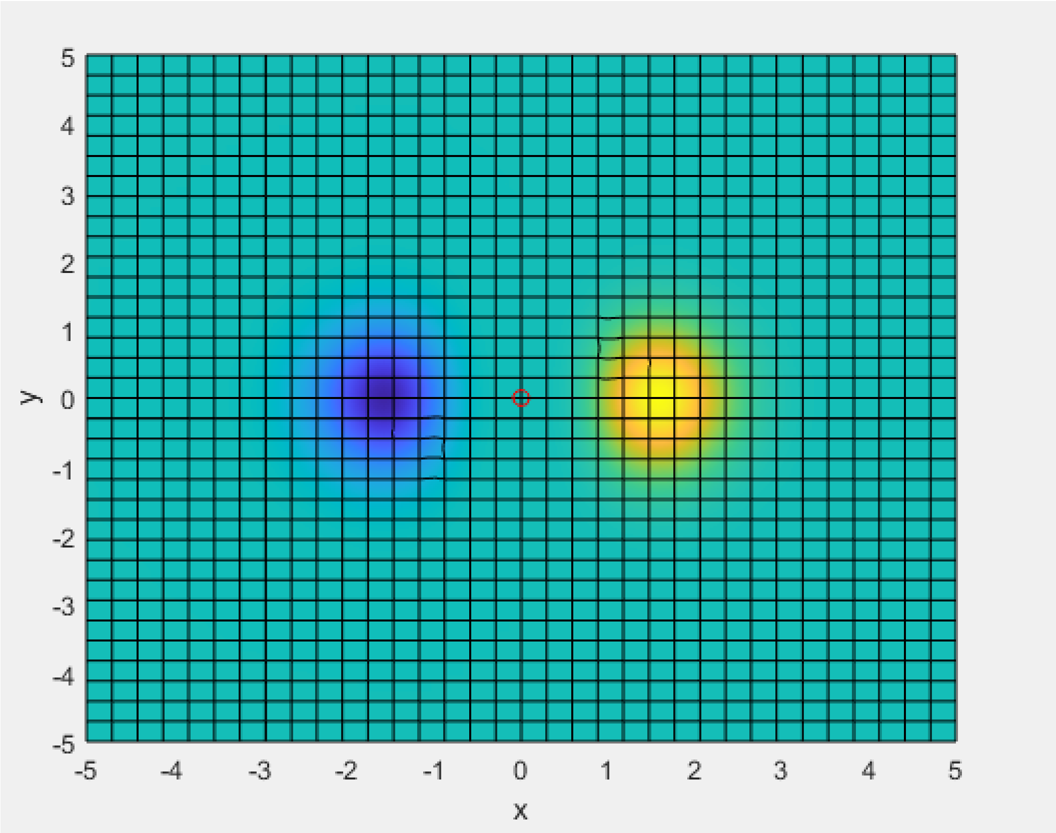
Σύμφωνα με τη μέθοδο και τις επιλεγμένες σταθερές, το ελάχιστο βρίσκεται στο σημείο [-1.58115979, 0.00001559], και ο συνολικός αριθμός διαδοχικών σημείων είναι k = 28. Εμφανίζουμε, ακόμη, τη γραφική παράσταση της αριθμητικής τιμής της συνάρτησης f(x,y) στα σημεία αυτά, συναρτήσει του βήματος k:



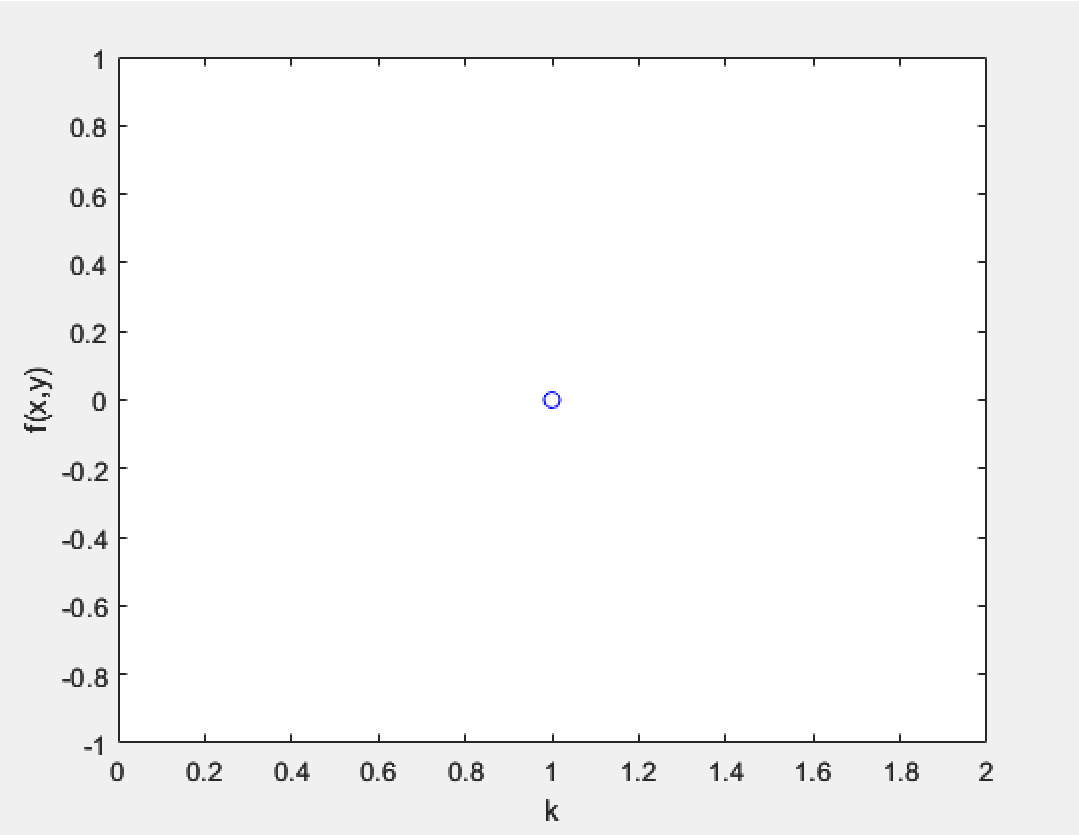
Βλέπουμε ότι τα σημεία πράγματι συγκλίνουν στο ελάχιστο, όσο αυξάνεται το k. Μάλιστα, παρατηρώντας τη γραφική παράσταση της f, φαίνεται ότι το ελάχιστο αυτό είναι το ολικό της. Η σύγκλιση εδώ είναι πολύ γρήγορη, καθώς το gk μεταβάλλεται κάθε φορά, με βάση τον κανόνα Armijo. Εδώ θα πρέπει να προσέξουμε η αρχική τιμή του s να μην είναι ούτε πολύ μεγάλη, ούτε πολύ μικρή, μιας και ενδέχεται να οδηγήσει την όλη διαδικασία σε αστάθεια, από το πρώτο βήμα.

<u>Σημείο εκκίνησης [0 0]</u>

Α, Β, Γ) Θέτοντας σταθερό gk = 0.6, τρέχουμε τον αλγόριθμο Levenberg – Marquardt και εμφανίζουμε τη σύγκλιση των διαδοχικών σημείων μέχρι την εύρεση του ελαχίστου, για σφάλμα e = 0.0001:



Σύμφωνα με τη μέθοδο και τις επιλεγμένες σταθερές, το ελάχιστο βρίσκεται στο σημείο [0, 0], και ο συνολικός αριθμός διαδοχικών σημείων είναι k = 1. Αυτό δε θα πρέπει να μας εκπλήσσει, καθώς, εφαρμόζοντας ως αρχικό σημείο το [0, 0], ικανοποιείται αυτομάτως η συνθήκη για τον τερματισμό της συνάρτησης. Μάλιστα, το σημείο αυτό αποτελεί και τοπικό ελάχιστο της f. Εμφανίζουμε, ακόμη, τη γραφική παράσταση της αριθμητικής τιμής της συνάρτησης f(x,y) στο σημείο αυτό, συναρτήσει του βήματος k:

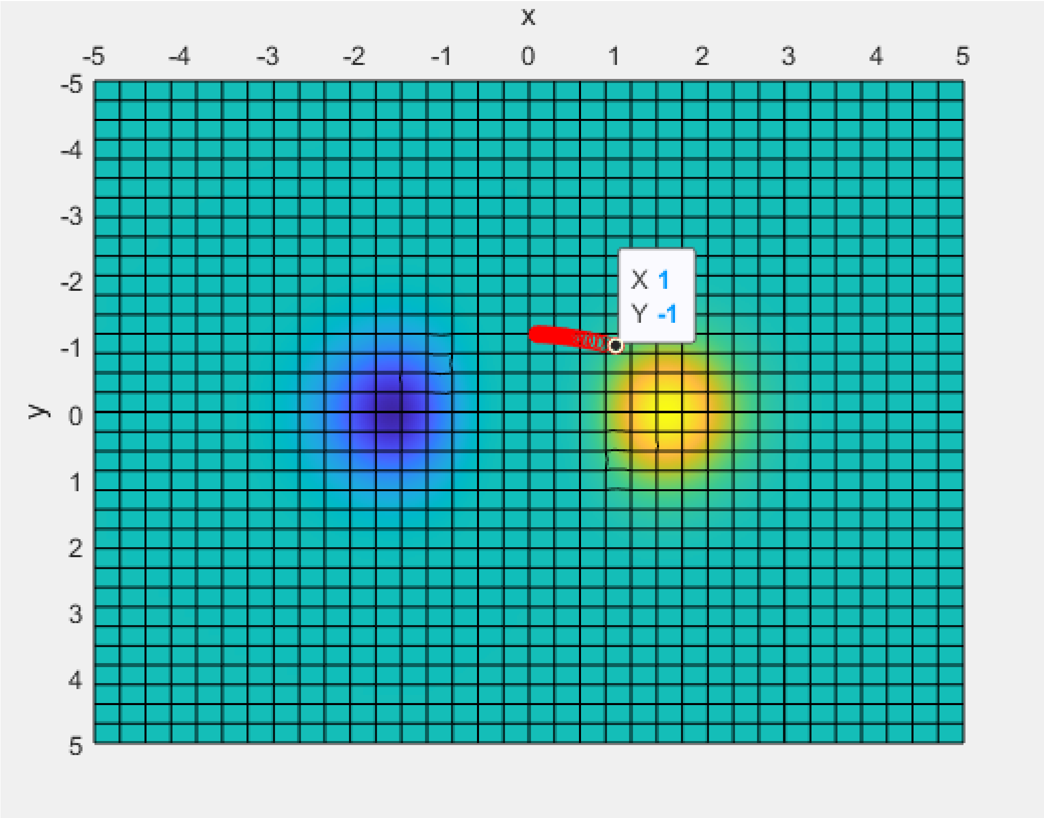


Προφανώς f(0,0) = 0. Το σημείο αυτό είναι τοπικό ελάχιστο της συνάρτησης στη περιοχή κοντά στην αρχή των αξόνων.

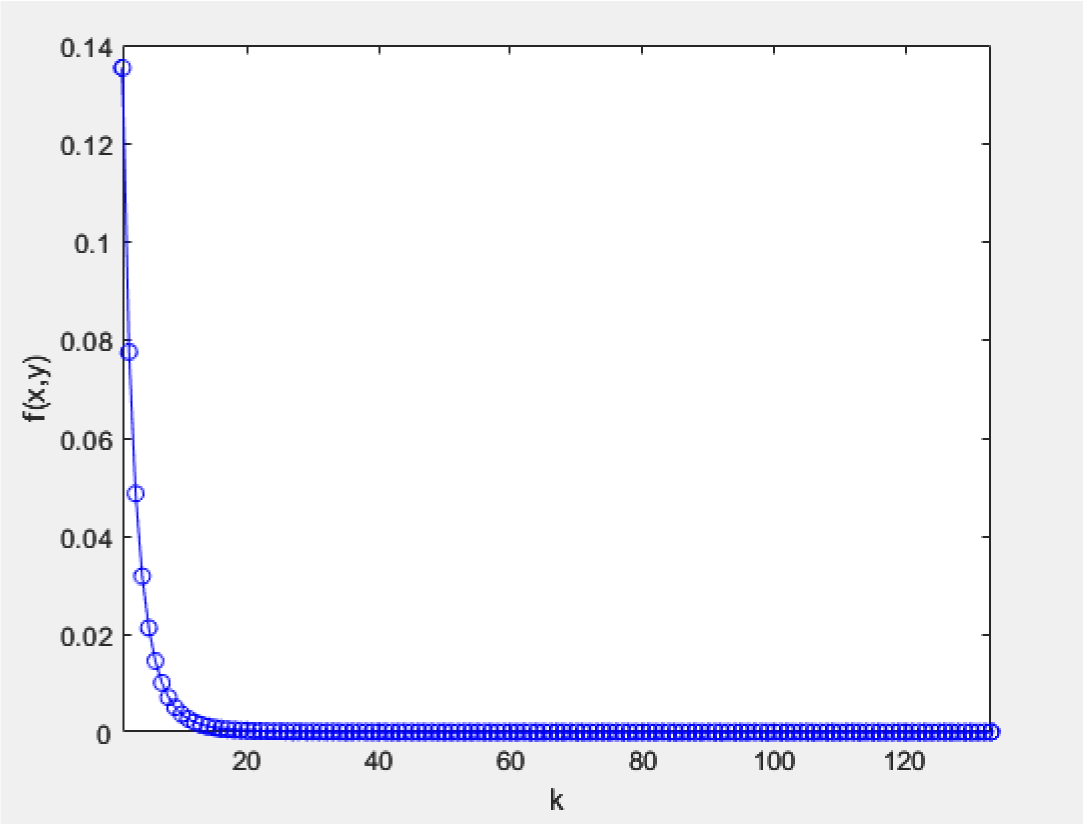
Ακριβώς τα ίδια αποτελέσματα - συμπεράσματα θα εμφανιστούν και στις άλλες δύο μεθόδους Levenberg – Marquardt, αυτής με gk ώστε να ελαχιστοποιεί την f(xk + gk*dk), και αυτής με gk με βάση τον κανόνα Armijo.

Σημείο εκκίνησης [1 -1]

A) Θέτοντας σταθερό gk = 0.6, τρέχουμε τον αλγόριθμο Levenberg – Marquardt και εμφανίζουμε τη σύγκλιση των διαδοχικών σημείων μέχρι την εύρεση του ελαχίστου, για σφάλμα e = 0.0001:

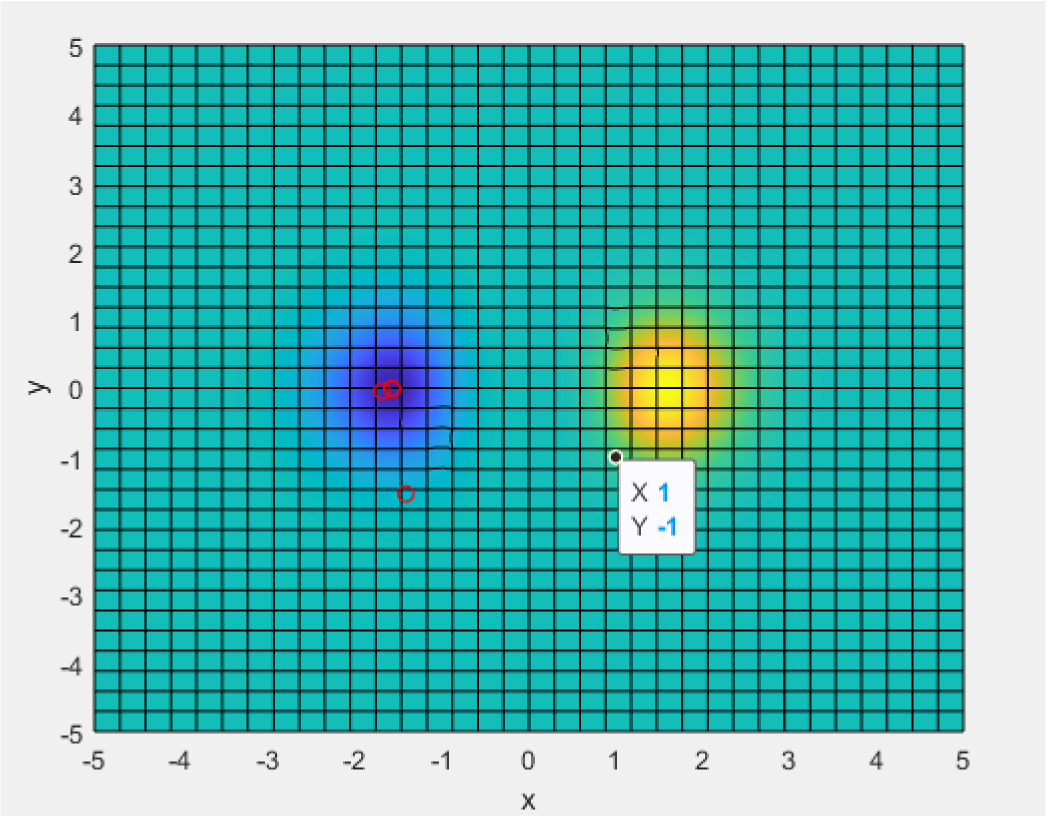


Σύμφωνα με τη μέθοδο και τις επιλεγμένες σταθερές, το ελάχιστο βρίσκεται στο σημείο [0.09448356, -1.17187149], και ο συνολικός αριθμός διαδοχικών σημείων είναι k = 133. Εμφανίζουμε, ακόμη, τη γραφική παράσταση της αριθμητικής τιμής της συνάρτησης f(x,y) στα σημεία αυτά, συναρτήσει του βήματος k:

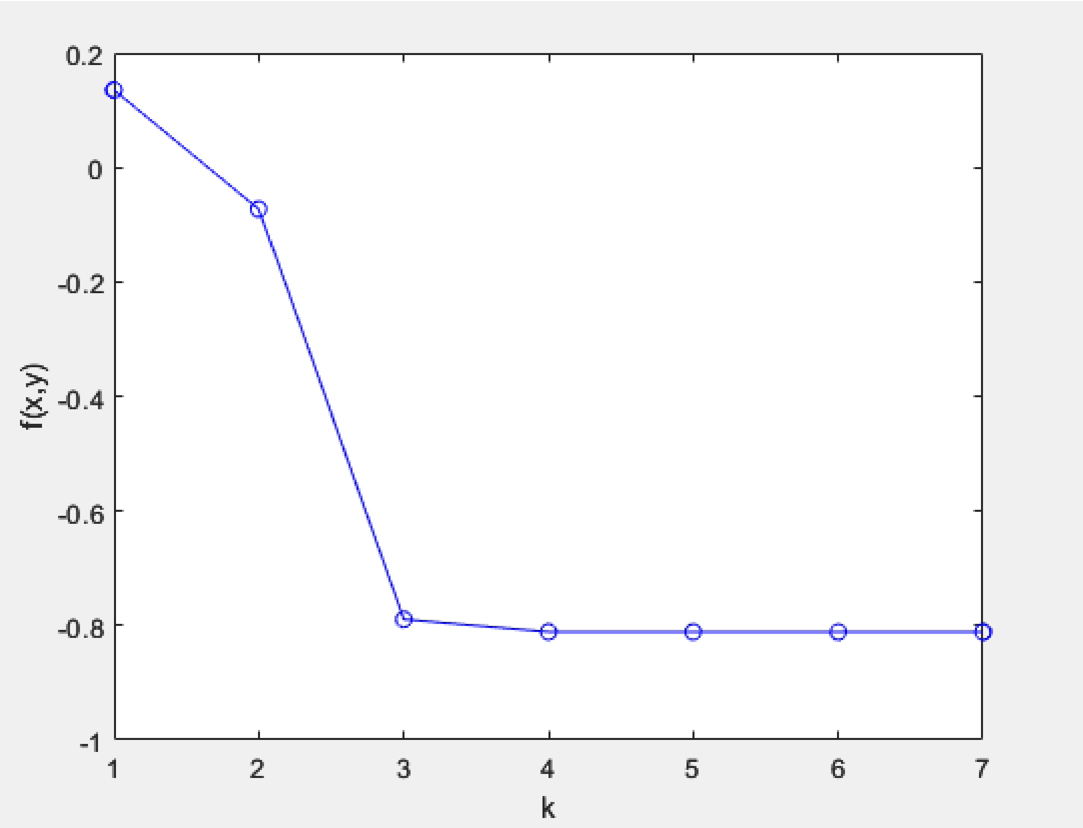


Βλέπουμε ότι τα σημεία συγκλίνουν στο μηδέν, όσο αυξάνεται το k. Δηλαδή ο αλγόριθμος δε συγκλίνει στο ολικό ελάχιστο της f, αλλά σε μία ισοσταθμική της, όπου η τιμή της είναι μηδέν. Αυτό συμβαίνει, καθώς ξεκινήσαμε από μια περιοχή μακριά από το ολικό ελάχιστο, έξω από την ‘κοιλάδα’ της f, και συγκεκριμένα μέσα στο ‘όρος’ της. Η σύγκλιση, βέβαια, είναι αρκετά αργή, εξαιτίας της φύσης του αλγορίθμου. Το βήμα gk δεν επιλέχθηκε πραγματικά τυχαία, αλλά τέτοιο ώστε να οδηγήσει σε σύγκλιση του αλγορίθμου στο ελάχιστο. Για μεγαλύτερες ή μικρότερες τιμές του gk μπορεί κανείς να διαπιστώσει ότι η μέθοδος δε θα συγκλίνει.

B) Σε κάθε επανάληψη του αλγορίθμου, υπολογίζουμε το gk, ώστε να ελαχιστοποιείται η f(xk + gk*dk). Τρέχουμε τον αλγόριθμο Levenberg – Marquardt και εμφανίζουμε τη σύγκλιση των διαδοχικών σημείων μέχρι την εύρεση του ελαχίστου, για σφάλμα e = 0.0001:

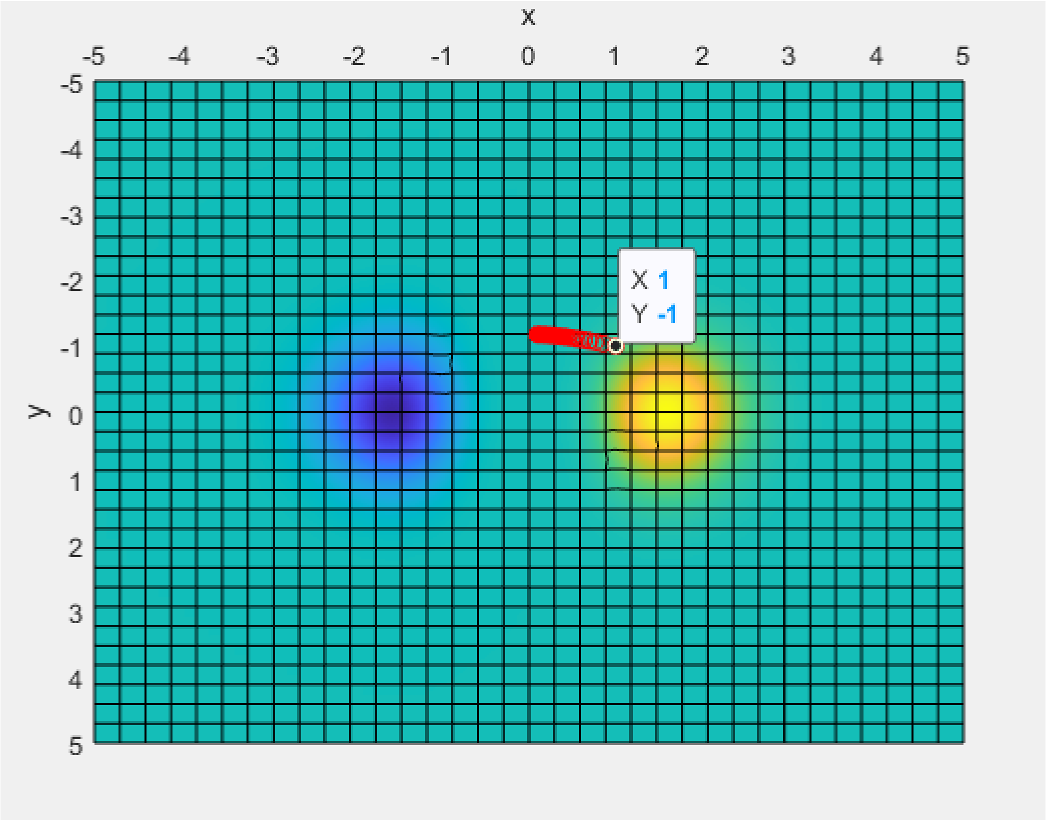


Σύμφωνα με τη μέθοδο και τις επιλεγμένες σταθερές, το ελάχιστο βρίσκεται στο σημείο [-1.58114407, -0.00000190], και ο συνολικός αριθμός διαδοχικών σημείων είναι k = 7. Βλέπουμε, ακόμη, ότι τα διαδοχικά σημεία σχηματίζουν κάθετα μεταξύ τους διανύσματα, το οποίο είναι λογικό, από την επιλογή του μεγέθους gk. Εμφανίζουμε, ακόμη, τη γραφική παράσταση της αριθμητικής τιμής της συνάρτησης f(x,y) στα σημεία αυτά, συναρτήσει του βήματος k:

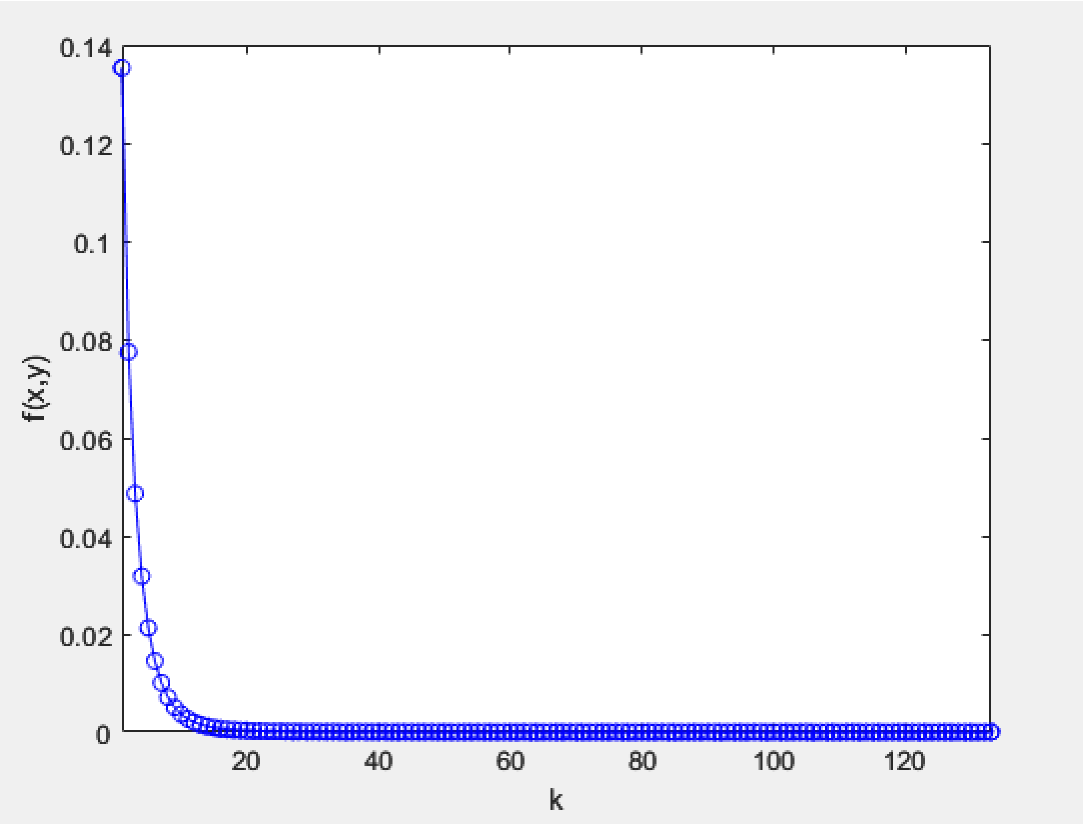


Βλέπουμε ότι τα σημεία πράγματι συγκλίνουν στο ελάχιστο, όσο αυξάνεται το k. Μάλιστα, παρατηρώντας τη γραφική παράσταση της f, φαίνεται ότι το ελάχιστο αυτό είναι το ολικό της. Η σύγκλιση εδώ είναι πολύ πιο γρήγορη, μιας και θέτουμε gk κάθε φορά τέτοιο ώστε να ελαχιστοποιεί την f(xk + gk*dk). Παρά το γεγονός ότι ξεκινήσαμε από μία περιοχή εκτός της 'κοιλάδας' της f, συγκεκριμένα μέσα από το 'όρος' της, η μέθοδος κατάφερε να περάσει από το τοπικό ελάχιστο της αρχής των αξόνων και να βρεθεί στο ολικό ελάχιστο. Δηλαδή η μέθοδος αυτή πάντοτε θα συγκλίνει στο ολικό ελάχιστο της f, καθώς χρησιμοποιεί τη βέλτιστη επιλογή του gk κάθε φορά.

Γ) Υπολογίζουμε, τώρα, το βήμα gk με βάση τον κανόνα Armijo. Με αρχικές συνθήκες a = 0.05, b = 0.3 και s = 0.6, τρέχουμε τον αλγόριθμο Levenberg – Marquardt και εμφανίζουμε τη σύγκλιση των διαδοχικών σημείων μέχρι την εύρεση του ελαχίστου, για σφάλμα e = 0.0001:



Σύμφωνα με τη μέθοδο και τις επιλεγμένες σταθερές, το ελάχιστο βρίσκεται στο σημείο [0.09448356, -1.17187149], και ο συνολικός αριθμός διαδοχικών σημείων είναι k = 133. Εμφανίζουμε, ακόμη, τη γραφική παράσταση της αριθμητικής τιμής της συνάρτησης f(x,y) στα σημεία αυτά, συναρτήσει του βήματος k:



Βλέπουμε ότι τα σημεία συγκλίνουν στο μηδέν, όσο αυξάνεται το k. Δηλαδή ο αλγόριθμος δε συγκλίνει στο ολικό ελάχιστο της f, αλλά σε μία ισοσταθμική της, όπου η τιμή της είναι μηδέν. Αυτό συμβαίνει, καθώς ξεκινήσαμε από μια περιοχή μακριά από το ολικό ελάχιστο, έξω από την 'κοιλάδα' της f, και συγκεκριμένα μέσα στο 'όρος' της, και γι' αυτό η σύγκλιση είναι αρκετά αργή. Εδώ θα πρέπει να προσέξουμε η αρχική τιμή του s να μην είναι ούτε πολύ μεγάλη, ούτε πολύ μικρή, μιας και ενδέχεται να οδηγήσει την όλη διαδικασία σε αστάθεια, από το πρώτο βήμα.

Από τα παραπάνω, συμπεραίνουμε ότι η καλύτερη επιλογή βήματος gk, που μπορούμε να κάνουμε, είναι σαφώς αυτή που ελαχιστοποιεί την f(xk + gk*dk). Στη περίπτωση αυτή, είτε στη μέθοδο Μέγιστης Καθόδου, είτε στη μέθοδο Levenberg –

Marquardt, ο αλγόριθμος θα βρίσκει πάντοτε το ολικό ελάχιστο της f, και αυτό θα συμβαίνει τόσο αν ξεκινάμε μέσα από την ‘κοιλάδα’ της, όσο και μέσα από το ‘όρος’ της. Εξαίρεση αποτελεί το σημείο εκκίνησης [0 0], στο οποίο όλες οι μέθοδοι τερματίζουν από το πρώτο βήμα.

Αντίθετα, για επιλογή αυθαίρετου gk, ή επιλογή gk σύμφωνα με τον κανόνα Armijo, οι μέθοδοι Μέγιστης Καθόδου και Levenberg – Marquardt δεν βρίσκουν πάντοτε το ολικό ελάχιστο. Για να συμβεί αυτό, θα πρέπει να ξεκινήσουμε μέσα από την ‘κοιλάδα’ της f. Αντιθέτως (αν ξεκινήσουμε από το ‘όρος’ της f, ή από το [0 0]), οι δύο μέθοδοι θα συγκλίνουν στο μηδέν, λόγω εγκλωβισμού τους στο τοπικό αυτό ακρότατο.

Τέλος, gk που ελαχιστοποιεί την f(xk + gk*dk) οδηγεί σε μικρότερο αριθμό k διαδοχικών σημείων και άρα σε πιο γρήγορη εύρεση του επιθυμητού αποτελέσματος.