*Τσαντίκης Γεώργιος*
*ΑΕΜ: 10722*

# 2η Εργαστηριακή Άσκηση

## Ελαχιστοποίηση συναρτήσεων πολλών μεταβλητών χωρίς περιορισμούς με χρήση παραγώγων

## ΠΕΡΙΓΡΑΦΗ ΠΡΟΒΛΗΜΑΤΟΣ

Στην παρούσα εργασία ζητείται να επιλυθεί το πρόβλημα ελαχιστοποίησης μιας δοσμένης συνάρτησης πολλών μεταβλητών $f$:$\mathbb{R}^n \to \mathbb{R}$ , χωρίς περιορισμούς, με χρήση αλγορίθμων που βασίζονται στην ιδέα της **επαναληπτικής καθόδου**, βάσει της οποίας ξεκινάμε από κάποιο σημείο $x_0 \epsilon \mathbb{R}^n$ και παράγουμε διαδοχικά τα διανύσματα $x_1, x_2$, ... έτσι ώστε $f(x_{k+1})$ < $f(x_k)$, $k$ = 1,2,… . Με λίγα λόγια σε κάθε επανάληψη θέλω να παίρνω κάτι καλύτερο (βελτίωση). Τα διανύσματα παράγονται σύμφωνα με τη σχέση: $x_{k+1} = x_k + \gamma_k d_k$, όπου $\boldsymbol{d} \epsilon \mathbb{R}^n$ είναι το διάνυσμα κατεύθυνσης και το $\boldsymbol{\gamma}$ (βήμα) δηλώνει πόσο πολύ θα κινηθώ σε αυτήν την κατεύθυνση αναζήτησης της λύσης. Επομένως, πρέπει: $\nabla f^T(x_k)d_k < 0$, ( $\nabla f(x)$ → η κλίση της $\boldsymbol{f}$ που δείχνει προς την αύξηση των τιμών της συνάρτησης), με σκοπό το $\boldsymbol{d_k}$ να με οδηγεί σε κατεύθυνση μείωσης, δηλαδή προς ισοβαρείς καμπύλες με μικρότερη τιμή για την συνάρτηση.

Ωστόσο, χρειάζεται ιδιαίτερη προσοχή η **επιλογή** του βήματος $\boldsymbol{\gamma_k}$ σε κάθε επανάληψη των αναδρομικών αλγορίθμων για την ελαχιστοποίηση της συνάρτησης $\boldsymbol{f}$, η οποία στην εργασία θα εξεταστεί για τρεις περιπτώσεις:

a) **Σταθερό** (μία ενδεικτική τιμή της επιλογής μας). Αν το $\boldsymbol{\gamma_k}$ επιλεγεί πολύ μικρό θα οδηγήσει σε ασήμαντη μείωση της $\boldsymbol{f}$ και κατά συνέπεια σε αργή σύγκλιση και αύξηση του πλήθους των επαναλήψεων, ενώ αν επιλεγεί πολύ μεγάλο μπορεί να οδηγήσει τον αλγόριθμο σε αστάθεια, μέχρι και απόκλιση της διαδικασίας αναζήτησης (Τυπικά, το $\gamma_k \in [\mathbf{10^{-3}}, \mathbf{1}]$). Γενικά ως μέθοδος είναι η απλούστερη δυνατή που ελαχιστοποιεί την υπολογιστική πολυπλοκότητα του αλγορίθμου ελαχιστοποίησης καθώς το βήμα παραμένει σταθερό σε όλη την διάρκεια της αναζήτησης.

b) Τέτοιο ώστε να **ελαχιστοποιεί** την $\boldsymbol{f(x_k + \gamma_k d_k)}$ ως προς $\gamma_k > 0$, βάσει της **μεθόδου** της **διχοτόμου** (*Αλγόριθμος 5.1.1* από το *ΒΙΒΛΙΟ*) και

c) Βάσει του κανόνα **Armijo,** τεχνική η οποία προσπαθεί να ικανοποιήσει κάποια **κριτήρια καλής λειτουργίας** (βλ. *σελ.* 133 *ΒΙΒΛΙΟΥ*). Αποτελεί μία μέθοδο διαδοχικής μείωσης του $\gamma_k$ , το οποίο επιλέγεται ως : $\gamma_k = s\beta^{m_k}$, όπου $s$ δηλώνει το αρχικό βήμα, με $m_k$ να είναι ο μικρότερος μη αρνητικός ακέραιος που ικανοποιεί το Κριτήριο 4, δηλαδή: $\mathrm{f}(x_k) - \mathrm{f}(x_{k+1}) \geq -\alpha\beta^{m_k} s d_k^T \nabla f(x_k)$ , όπου συνήθως $a \in [10^{-5}, 10^{-1}]$ $και\ \beta \in [\frac{1}{10}, \frac{1}{2}]$. Επομένως, δοκιμάζουμε διαδοχικές τιμές του $m_k$ με προεπιλεγμένα $\alpha, \beta, s$ μέχρι την ικανοποίηση της παραπάνω σχέσης. Όπως ειπώθηκε αποσκοπεί στην ικανοποίηση του **Κριτηρίου 4** (βλ. σελ. 135 *ΒΙΒΛΙΟΥ*), το οποίο είναι το πιο κρίσιμο από όλα και μας βοηθάει να αποφύγουμε την επιλογή επικίνδυνα μεγάλων τιμών που συνεπάγονται ταλαντώσεις ή αποκλίσεις.

Σύμφωνα με το ***Θεώρημα 5.2.6*** (βλ. *σελ.* 136 *ΒΙΒΛΙΟΥ*) η επιλογή των $\gamma_k$, $d_k$ ώστε να ικανοποιούνται τα **Κριτήρια 1-4** είναι πάντα εφικτή.

Οι τρεις επαναληπτικοί αλγόριθμοι που θα μελετηθούν με **κριτήριο τερματισμού** $\| \nabla \mathrm{f}(x_k) \| < \varepsilon$ , όπου $\varepsilon > 0$ η σταθερά τερματισμού την οποία έχω ορίσει **για όλους** ίση με **0.01**, είναι**:**

1) Μέθοδος **Μέγιστης Καθόδου (Steepest Descent),**

2) Μέθοδος **Newton** και

3) Μέθοδος **Levenberg-Marquardt**

Στο πλαίσιο της εργασίας, με στόχο να συγκριθεί η απόδοσή τους, πέρα από τις διαφορετικές περιπτώσεις επιλογής $\boldsymbol{\gamma_k}$, η διαδικασία θα εκτελεστεί και από τρία διαφορετικά αρχικά σημεία:

i. το **κρίσιμο σημείο** $(0, 0)$ που είναι σημείο σάγματος,

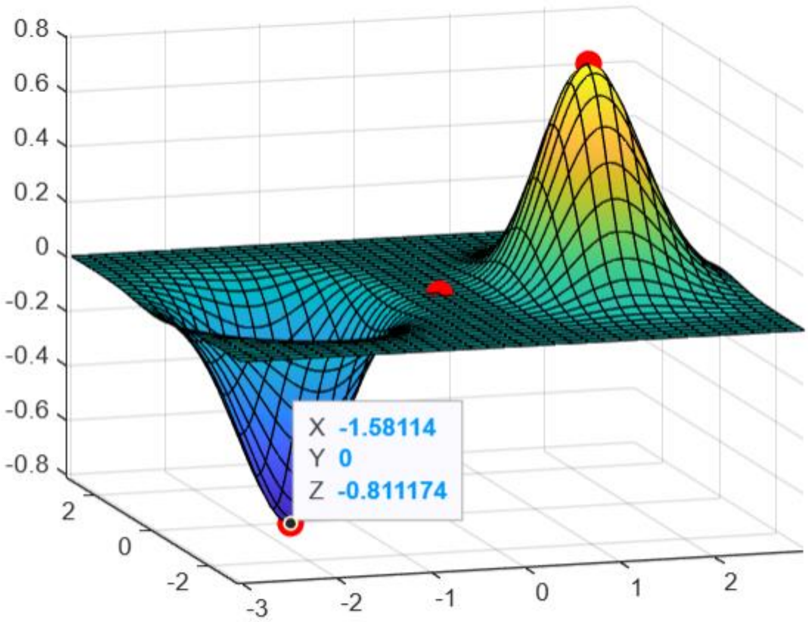
ii. το σημείο $(-1, 1)$ , που βρίσκεται εκτός της επιφάνειας της συνάρτησης, σχετικά κοντά στο ολικό ελάχιστο συγκριτικά με τα υπόλοιπα κρίσιμα σημεία, (παρουσιάζονται στο ***ΘΕΜΑ 1***) και επιτρέπει τη μελέτη της πορείας των αλγορίθμων από μια γενική θέση, όπως και

iii. το σημείο $(1, -1)$ με τη διαφορά ότι βρίσκεται πιο μακριά από το προηγούμενο σημείο

# ΘΕΜΑ 1

Παρακάτω παρουσιάζεται η τρισδιάστατη αναπαράσταση της αντικειμενικής προς μελέτη συνάρτησης $\boldsymbol{f}(\mathbf{x}, \mathbf{y}) = \boldsymbol{x}^{5} \cdot \boldsymbol{e}^{-x^2-y^2}$ και τα κρίσιμα σημεία της που επισημαίνονται με κόκκινες κουκκίδες. Συμπεραίνουμε τα παρακάτω:

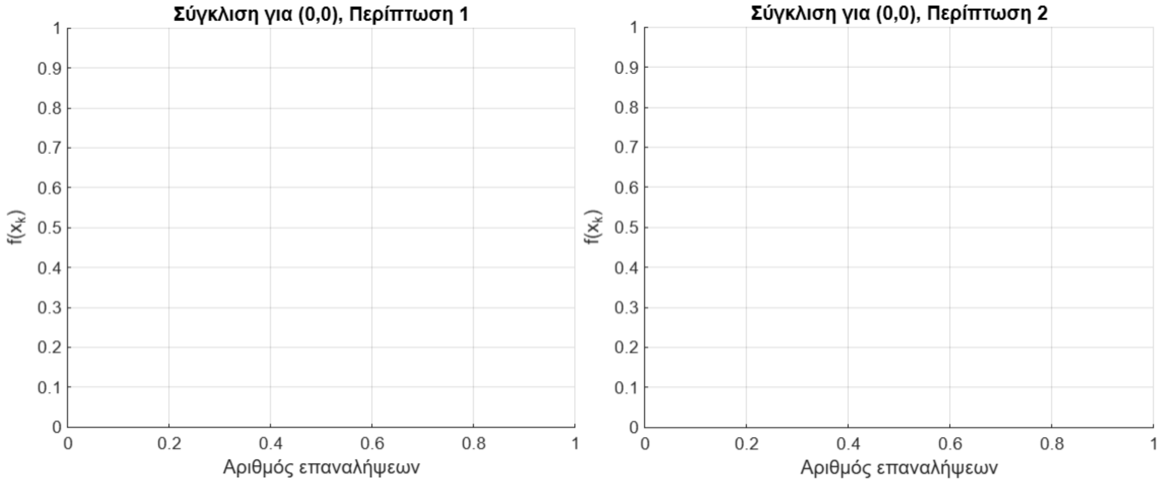
- Στο σημείο $(\mathrm{x} = -1.58114,\ \mathrm{y} = 0)$ η συνάρτηση παρουσιάζει **ελάχιστο** (ολικό) με $\mathrm{z} = \mathrm{f}(\mathrm{x}, \mathrm{y}) = -0.811174$
- Στο σημείο $(\mathrm{x} = 1.58114,\ \mathrm{y} = 0)$ η συνάρτηση παρουσιάζει **μέγιστο** (ολικό) με $\mathrm{z} = \mathrm{f}(\mathrm{x}, \mathrm{y}) = 0.811174$
- Στο σημείο $(\mathrm{x} = 0,\ \mathrm{y} = 0)$ η επιφάνεια εμφανίζει την τυπική γεωμετρία ενός **σαγματικού σημείου**, καθώς σε κάποιες κατευθύνσεις αυξάνεται και σε άλλες μειώνεται:
  - αν $x > 0, τότε\ f(x, 0) = x^5 \cdot e^{-x^2} > 0$
  - αν $x < 0, τότε\ f(x, 0) = x^5 \cdot e^{-x^2} < 0$



## ΘΕΜΑ 2

Στο παρόν θέμα υλοποιήθηκε στο MATLAB η **μέθοδος** της **Μέγιστης Καθόδου**, όπως παρουσιάζεται στον ***Αλγόριθμο 5.2.1*** του *ΒΙΒΛΙΟΥ* στη σελίδα 121. Η ιδέα της μεθόδου βασίζεται στην μετακίνηση προς την κατεύθυνση της μεγαλύτερης μείωσης των τιμών της συνάρτησης, με **διάνυσμα κατεύθυνσης** το $d_k = -\nabla f(x_k)$ . Ο αλγόριθμος κλίσης που χρησιμοποιείται έχει τη μορφή: $x_{k+1} = x_k - \gamma_k \Delta_k \nabla f(x_k)$, όπου $\mathbf{\Delta_k}$ είναι συμμετρικός και θετικά ορισμένος πίνακας. Πιο συγκεκριμένα, η μέθοδος της μέγιστης καθόδου προκύπτει επιλέγοντας $\mathbf{\Delta_k} = \boldsymbol{I}$**,** όπου $\boldsymbol{I}$ είναι ο μοναδιαίος $2 \times 2$ πίνακας , οπότε ο αλγόριθμος κλίσης παίρνει τη μορφή: $x_{k+1} = x_k - \gamma_k \nabla f(x_k)$. Σύμφωνα με τον ***Αλγόριθμο 5.2.1****,* ορίζουμε την σταθερά τερματισμού $\varepsilon > 0$ επιλέγουμε το αρχικό σημείο $x_1$ και θέτουμε $k = 1$ (1$^{η}$ επανάληψη). Έπειτα πηγαίνουμε στο κύριο βήμα όπου αν ικανοποιείται το κριτήριο τερματισμού $\| \nabla \mathrm{f}(x_k) \| < \varepsilon$, ο αλγόριθμος τερματίζει. Διαφορετικά υπολογίζουμε το $\boldsymbol{d_k}$ όπως προαναφέρθηκε, ορίζουμε την τιμή του $\boldsymbol{\gamma_k}$ σύμφωνα με τις τρεις περιπτώσεις που παρουσιάστηκαν στην ***ΠΕΡΙΓΡΑΦΗ ΠΡΟΒΛΗΜΑΤΟΣ***, θέτουμε $x_{k+1} = x_k - \gamma_k \nabla f(x_k)$, $k = k + 1$ και επαναλαμβάνουμε το κύριο βήμα. Παρακάτω παρουσιάζονται οι γραφικές παραστάσεις της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων που απαιτούνται μέχρι να τερματίσει ο αλγόριθμος για τις τρεις περιπτώσεις επιλογής του $\gamma_k$:

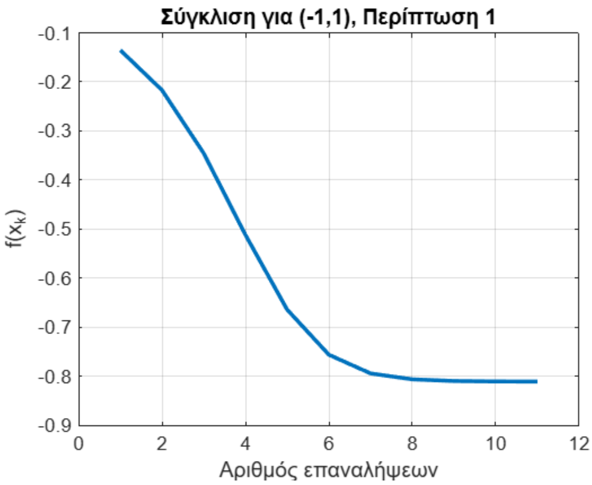
i. αρχικό σημείο το $(0, 0)$

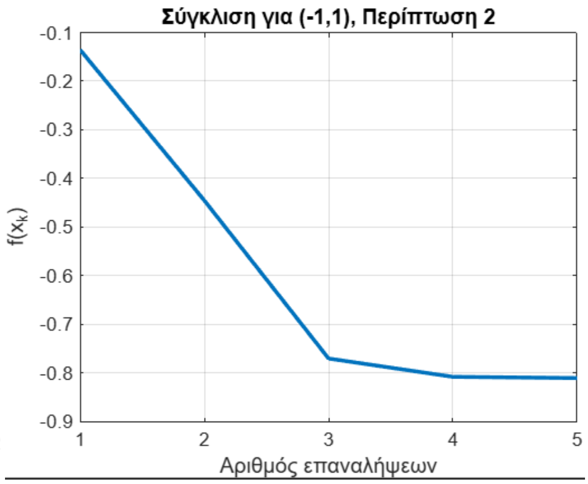


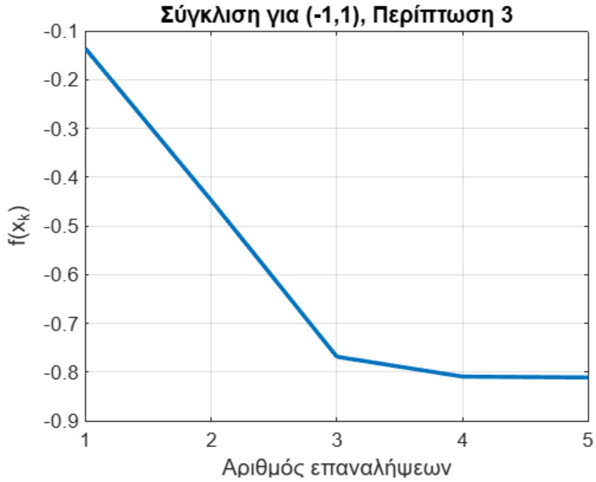

(*κανονικά σύμφωνα με τον αλγόριθμο η επανάληψη είναι μία παραπάνω από αυτή που εμφανίζεται στις γραφικές παραστάσεις καθώς πρώτα αυξάνεται η επανάληψη και ύστερα ελέγχεται το κριτήριο τερματισμού, ωστόσο αποφάσισα να συμπεριλάβω τον πραγματικό αριθμό επαναλήψεων μέχρι το επιθυμητό αποτέλεσμα*)

Η κλίση (gradient) της $f$ στο $(0,0)$ είναι: $\nabla f(0,0) = 0$, οπότε από το πρώτο βήμα της μεθόδου $d_{\mathrm{k}} = 0$, συνεπώς δεν υπάρχει κατεύθυνση καθόδου. Το $(0,0)$ όπως προαναφέρθηκε στο ***ΘΕΜΑ 1***, είναι σαγματικό σημείο και η μέθοδος δεν "ξεκολλαέι" από αυτό διότι θεωρεί ότι έχει φτάσει ήδη στο ελάχιστο λόγω της μηδενικής κλίσης αφού $x_{k+1} = x_k - \gamma_k \cdot 0 = x_k$ $(1)$. Τελικά έχουμε απόκλιση από την επιθυμητή τιμή ελαχίστου που είναι: $f(-1.58114, 0) = -0.811174$, λόγω εγκλωβισμού του αλγορίθμου στο κρίσιμο σημείο $(0,0)$ που είναι σαγματικό. Αυτό είναι το αποτέλεσμα, ανεξάρτητα από την επιλογή του $\gamma_k$, όπως είναι φανερό από την $(1)$. (*Περίπτωση 1→σταθερό $\gamma_k$ , Περίπτωση 2 → ελαχιστοποίηση της $f(x_{k+1})$ ως προς $\gamma_k$ με τη μέθοδο της διχοτόμου, Περίπτωση 3 → κανόνας Armijo*)

ii. αρχικό σημείο το $(-1,1)$






(*κανονικά σύμφωνα με τον αλγόριθμο η επανάληψη είναι μία παραπάνω από αυτή που εμφανίζεται στις γραφικές παραστάσεις καθώς πρώτα αυξάνεται η επανάληψη και ύστερα ελέγχεται το κριτήριο τερματισμού, ωστόσο αποφάσισα να συμπεριλάβω τον πραγματικό αριθμό επαναλήψεων μέχρι το επιθυμητό αποτέλεσμα*)

Η κλίση (gradient) της $f$ στο $(-1, 1)$ είναι: $\nabla f(-1, 1) = \begin{bmatrix} 3e^{-2} \\ 2e^{-2} \end{bmatrix}$ (μειώνεται εκθετικά), οπότε η αρχική κατεύθυνση καθόδου είναι $d_{\mathrm{k}} = \begin{bmatrix} -3e^{-2} \\ -2e^{-2} \end{bmatrix}$. Για την ενημέρωση του $x_{\mathrm{k}}$ απαιτείται ο υπολογισμός του βήματος $\gamma_{\mathrm{k}}$ η οποία γίνεται σύμφωνα με την ανάλυση που προηγήθηκε για κάθε περίπτωση στην ***ΠΕΡΙΓΡΑΦΗ ΠΡΟΒΛΗΜΑΤΟΣ***:

a) Η **σταθερή** ενδεικτική τιμή που επέλεξα για το $\boldsymbol{\gamma_{\mathrm{k}}}$ , ώστε να πραγματοποιηθεί η απεικόνιση στη γραφική παράσταση, είναι **0.3**. Όπως αναμενόταν παρατηρήθηκε ότι για πολύ μικρές τιμές του $\gamma_{\mathrm{k}}$ (πχ 0.001) η σύγκλιση ήταν πολύ αργή με τον αριθμό των επαναλήψεων να είναι πολύ μεγάλος, ενώ για τιμές περίπου 0.6 και μεγαλύτερες η διαδικασία αναζήτησης είναι ασταθής και φτάνει σε σημείο να αποκλίνει. Συγκριτικά με τις άλλες δύο περιπτώσεις θεωρήτικα είναι η λιγότερο αποδοτική εκτός αν γίνει ορθή επιλογή βήματος που προϋποθέτει μεγάλη εμπειρία, οπότε είναι συγκρίσιμη με τις άλλες δύο περιπτώσεις (για $\gamma_{\mathrm{k}} = 0.5$ προκύπτουν 6 επαναλήψεις για σύγκλιση).
b) Μέσω της **ελαχιστοποίησης** την $f(x_k + \gamma_k d_k)$ ως προς $\gamma_k > 0$, με τη **μέθοδο** της **διχοτόμου**, το $\gamma_{\mathrm{k}}$ είναι προσαρμοσμένο ώστε να επιτυγχάνει τη μεγαλύτερη δυνατή μείωση της $f(x, y)$ σε κάθε βήμα, για αυτό και αναμένεται γρήγορη σύγκλιση (αξιόπιστα), αν και υπολογιστικά είναι πιο δαπανηρό από τις άλλες δύο περιπτώσεις καθώς απαιτεί πολλαπλούς υπολογισμούς της $f(x, y)$ σε κάθε επανάληψη. Επιλέχθηκαν η σταθερά $\varepsilon = 0.001$, το τελικό εύρος του διαστήματος $l = 0.01$ και αρχικό διάστημα αναζήτησης του $\gamma_k$ το [0, 1] καθώς δεν παρατηρήθηκε διαφορά με την χρήση άλλων διαστημάτων (πχ [-1,1], [0, 5],… κ.α.). Γενικά είναι πιο αποδοτικό σε σύγκριση με το σταθερό $\gamma_k$.

c) Για τον κανόνα **Armijo** δοκιμάστηκαν διάφορες τιμές για τα $\alpha, \beta, s$ και επέλεξα να κρατήσω ενδεικτικά τις $\alpha = 0.01, \beta = 0.4 \; και \; s = 1$ καθώς έτσι συγκλίνει στις βέλτιστες επαναλήψεις που μάλιστα είναι ίδιες με την (b). Γενικά, εξασφαλίζει μείωση της $f(x, y)$, αποφεύγοντας ταλαντώσεις ή υπερβολικά μεγάλα βήματα με αυτόματη προσαρμογή του $\gamma_k$, συγκλίνει αξιόπιστα και είναι λιγότερο υπολογιστικά δαπανηρή μέθοδος από την (b).

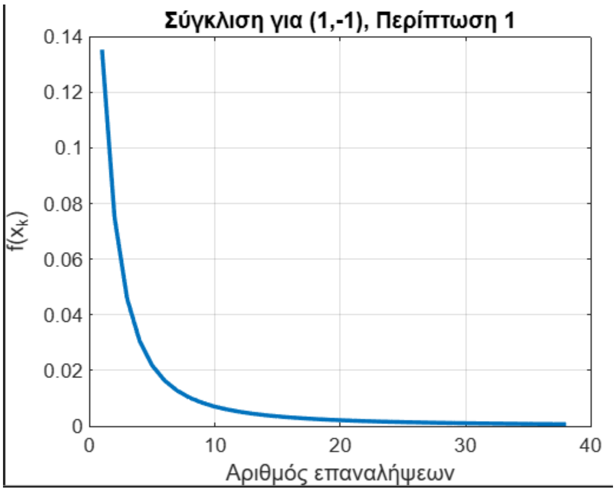
Παρακάτω εμφανίζονται συγκεντρωτικά τα αποτελέσματα για τις ενδεικτικές τιμές:

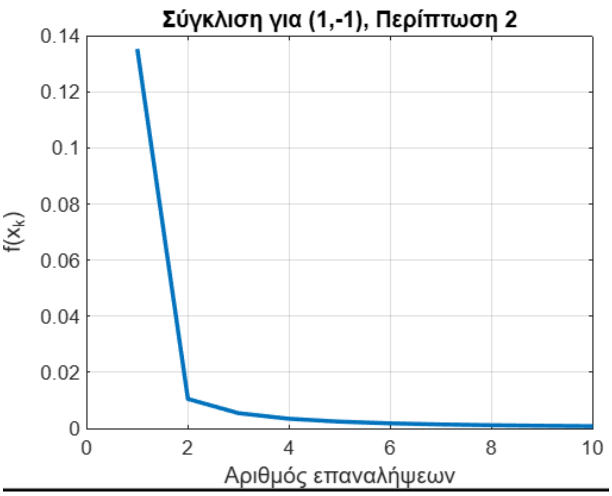
```
Αρχικό σημείο: (-1.00, 1.00)
  Περίπτωση 1: x* = (-1.5811, 0.0053), f(x*) = -0.811151, Επαναλήψεις = 11
  Περίπτωση 2: x* = (-1.5802, 0.0039), f(x*) = -0.811160, Επαναλήψεις = 5
  Περίπτωση 3: x* = (-1.5787, 0.0037), f(x*) = -0.811153, Επαναλήψεις = 5
```

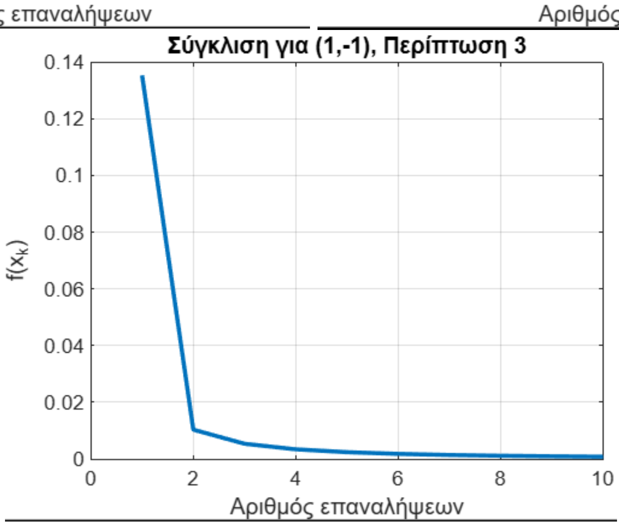
Τελικά έχουμε σύγκλιση στην επιθυμητή τιμή ελαχίστου που είναι:

$$f(-1.58114, 0) = -0.811174$$

iii. αρχικό σημείο το $(1, -1)$



(*κανονικά σύμφωνα με τον αλγόριθμο η επανάληψη είναι μία παραπάνω από αυτή που εμφανίζεται στις γραφικές παραστάσεις καθώς πρώτα αυξάνεται η επανάληψη και ύστερα ελέγχεται το κριτήριο τερματισμού, ωστόσο αποφάσισα να συμπεριλάβω τον πραγματικό αριθμό επαναλήψεων μέχρι το επιθυμητό αποτέλεσμα*)

Η κλίση (gradient) της $f$ στο $(-1, 1)$ είναι: $\nabla f(-1,1) = \begin{bmatrix} 3e^{-2} \\ 2e^{-2} \end{bmatrix}$ (μειώνεται εκθετικά καθώς πλησιάζει στο ελάχιστο), οπότε η αρχική κατεύθυνση καθόδου είναι $d_{\mathrm{k}} = \begin{bmatrix} -3e^{-2} \\ -2e^{-2} \end{bmatrix}$. Για την ενημέρωση του $x_{\mathrm{k}}$ απαιτείται ο υπολογισμός του βήματος $\gamma_{\mathrm{k}}$ η οποία γίνεται σύμφωνα με την ανάλυση που προηγήθηκε για κάθε περίπτωση στην ***ΠΕΡΙΓΡΑΦΗ ΠΡΟΒΛΗΜΑΤΟΣ.*** Η διαφορά με το προηγούμενο σημείο είναι ότι η μέθοδος δεν συγκλίνει στο ελάχιστο σημείο αλλά φτάνει σε μία περιοχή (συγκλίνει σε σημείο) όπου η συνάρτηση φαίνεται επίπεδη ($\nabla f(\mathrm{x}, y) \sim 0$) σε σχέση με την κατεύθυνση καθόδου οπότε και σταματάει η εκτέλεση του αλγορίθμου ενώ έχει φτάσει κοντά στο σαγματικό σημείο (βλ. ***ΘΕΜΑ 1***):

a) Η **σταθερή** ενδεικτική τιμή που επέλεξα για το $\boldsymbol{\gamma_{\mathrm{k}}}$ , ώστε να πραγματοποιηθεί η απεικόνιση στη γραφική παράσταση, είναι **0.3**. Όπως αναμενόταν παρατηρήθηκε ότι για πολύ μικρές τιμές του $\gamma_{\mathrm{k}}$ (πχ 0.001) η σύγκλιση ήταν πολύ αργή με τον αριθμό των επαναλήψεων να είναι πολύ μεγάλος, ενώ για τιμές περίπου 0.6 και μεγαλύτερες η διαδικασία αναζήτησης είναι ασταθής και φτάνει σε σημείο να αποκλίνει. Συγκριτικά με τις άλλες δύο περιπτώσεις θεωρήτικα είναι η λιγότερο αποδοτική όπως και αποδεικνύεται στην πράξη καθώς δεν βρέθηκε τιμή του $\gamma_{\mathrm{k}}$ ώστε να είναι παρόμοια αποδοτικά (για $\gamma_{\mathrm{k}} = 0.5$ που στο σημείο $(1, -1)$ ήταν συγκρίσιμη, εδώ προκύπτουν 22 επαναλήψεις για σύγκλιση που είναι αρκετές παραπάνω).

b) Μέσω της **ελαχιστοποίησης** την $f(x_k + \gamma_k d_k)$ ως προς $\gamma_k > 0$, με τη **μέθοδο** της **διχοτόμου**, το $\gamma_{\mathrm{k}}$ είναι προσαρμοσμένο ώστε να επιτυγχάνει τη μεγαλύτερη δυνατή μείωση της $f(x, y)$ σε κάθε βήμα, για αυτό και αναμένεται γρήγορη σύγκλιση (αξιόπιστα), αν και υπολογιστικά είναι πιο δαπανηρό από τις άλλες δύο περιπτώσεις καθώς απαιτεί πολλαπλούς υπολογισμούς της $f(x, y)$ σε κάθε επανάληψη. Επιλέχθηκαν η σταθερά $\varepsilon = 0.001$, το τελικό εύρος του διαστήματος $l = 0.01$ και αρχικό διάστημα αναζήτησης του $\gamma_k$ το [0, 1]. Με την χρήση άλλων διαστημάτων (πχ [-1,1], [0, 2],… κ.α.) η μέθοδος παρατηρήθηκε να συγκλίνει σε διαφορετικό σημείο αλλά στην ίδια επίπεδη περιοχή με προηγουμένως. Γενικά είναι πιο αποδοτικό σε σύγκριση με το σταθερό $\gamma_k$.

c) Για τον κανόνα **Armijo** δοκιμάστηκαν διάφορες τιμές για τα $\alpha, \beta, s$ και επέλεξα να κρατήσω ενδεικτικά τις $\alpha = 0.01, \beta = 0.4 \; και \; s = 1$. Με αλλαγή των τιμών στο αρχικό βήμα $\mathrm{s}$ η μέθοδος παρατηρήθηκε να συγκλίνει σε διαφορετικό σημείο αλλά στην ίδια επίπεδη περιοχή με προηγουμένως Γενικά, εξασφαλίζει μείωση της $f(x, y)$, αποφεύγοντας ταλαντώσεις ή υπερβολικά μεγάλα βήματα με αυτόματη προσαρμογή του $\gamma_k$, συγκλίνει αξιόπιστα και είναι λιγότερο υπολογιστικά δαπανηρή μέθοδος από την (b).

Ένα παράδειγμα για αρχικό διάστημα [0,2] για την (b) και για $s = 2$ για την (c) είναι το εξής:

```
Αρχικό σημείο: (1.00, -1.00)
  Περίπτωση 1: x* = (0.3285, -1.2806), f(x*) = 0.000666, Επαναλήψεις = 38
  Περίπτωση 2: x* = (0.1900, -1.5400), f(x*) = 0.000022, Επαναλήψεις = 1
  Περίπτωση 3: x* = (0.1880, -1.5413), f(x*) = 0.000021, Επαναλήψεις = 1
```

Παρακάτω εμφανίζονται συγκεντρωτικά τα αποτελέσματα για τις ενδεικτικές τιμές που χρησιμοποίησα:

```
Αρχικό σημείο: (1.00, -1.00)
  Περίπτωση 1: x* = (0.3285, -1.2806), f(x*) = 0.000666, Επαναλήψεις = 38
  Περίπτωση 2: x* = (0.3421, -1.3416), f(x*) = 0.000689, Επαναλήψεις = 10
  Περίπτωση 3: x* = (0.3416, -1.3423), f(x*) = 0.000683, Επαναλήψεις = 10
```

Όπως είναι φανερό δεν έχουμε σύγκλιση στην επιθυμητή τιμή ελαχίστου που είναι η $f(-1.58114, 0) = -0.811174$, αλλά έχουμε σύγκλιση περίπου στο σημείο με τιμή $f(0.34, -1.3) = 0$ .

# <u>ΘΕΜΑ 3</u>

Στο παρόν θέμα υλοποιήθηκε στο MATLAB η **μέθοδος Newton**, όπως παρουσιάζεται στον ***Αλγόριθμο 5.2.2*** του *ΒΙΒΛΙΟΥ* στη σελίδα 126. Η ιδέα της μεθόδου βασίζεται στη χρήση πληροφοριών της παραγώγου δεύτερης τάξης (**Εσσιανός** πίνακας), για να προσαρμόζει τη διαδικασία καθόδου. Ο αλγόριθμος κλίσης που χρησιμοποιείται έχει τη μορφή: $x_{k+1} = x_k - \gamma_k \Delta_k \nabla f(x_k)$, όπου $\mathbf{\Delta_k}$ είναι συμμετρικός και θετικά ορισμένος πίνακας. Πιο συγκεκριμένα, η μέθοδος της μέγιστης καθόδου προκύπτει επιλέγοντας $\mathbf{\Delta_k} = \boldsymbol{\nabla^2 f(x_k)}$**,** με την **προϋπόθεση** ότι $\boldsymbol{\nabla^2 f(x_k)}$ είναι **θετικά ορισμένος**, ώστε να υπάρχει ο αντίστροφος του και να έχουμε εγγυημένα κατεύθυνση καθόδου. Τελικά ο αλγόριθμος κλίσης παίρνει τη μορφή: $x_{k+1} = x_k - \gamma_k [\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$, με την αναζήτηση ελαχίστου να γίνεται στην **κατεύθυνση:** $d_k = -[\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$. Σύμφωνα με τον ***Αλγόριθμο 5.2.2****,* ορίζουμε τη σταθερά τερματισμού $\varepsilon$, επιλέγουμε το σημείο εκκίνησης $x_1$ και θέτουμε $k = 1$ ($1^{η}$ επανάληψη). Έπειτα πηγαίνουμε στο *Βήμα 1* όπου υπολογίζουμε την $\nabla f(x_k)$ και αν ικανοποιείται το κριτήριο τερματισμού $\| \nabla \mathrm{f}(x_k) \| < \varepsilon$, ο αλγόριθμος τερματίζει. Διαφορετικά πηγαίνουμε στο *Βήμα 2* και υπολογίζουμε τον εσσιανό πίνακα $\nabla^2 f(x_k) = \left[\frac{\partial^2 f}{\partial x_i \partial y_i}\right]$, μετά στο *Βήμα 3* υπολογίζουμε το $\boldsymbol{d_k}$ όπως προαναφέρθηκε, και στο *Βήμα 4* θέτουμε $x_{k+1} = x_k + \gamma_k d_\mathrm{k}$, όπου το $\gamma_k \geq 0$ υπολογίζεται σύμφωνα με τις τρεις περιπτώσεις που παρουσιάστηκαν στην ***ΠΕΡΙΓΡΑΦΗ ΠΡΟΒΛΗΜΑΤΟΣ***. Τέλος, στο *Βήμα 5* θέτουμε $k = k + 1$ και πηγαίνουμε στο *Βήμα 1.* Για τη συγκεκριμένη μέθοδο δεν θα παρουσιαστούν οι γραφικές παραστάσεις αναλυτικά καθώς για **κανένα σημείο Εσσιανός** πίνακας **δεν** είναι **θετικά ορισμένος** (η συνάρτηση $\boldsymbol{hessian\_f(x,y)}$ που φαίνεται παρακάτω είναι για τον υπολογισμό των τιμών του Εσσιανού $2 \times 2$ στο σημείο $(x, y)$). Για την εξαγωγή του συμπεράσματος της θετικής οριστικότητας του Εσσιανού ελέγχουμε τις ιδιοτιμές του, οπότε μόνο αν **όλες** οι **ιδιοτιμές** του είναι **θετικές**, ο **Εσσιανός** είναι **θετικά ορισμένος**:

i. αρχικό σημείο το $(0, 0)$

```
>> hessian_f(0,0)

ans =

     0     0
     0     0
```

Ο παραπάνω πίνακας είναι μηδενικός όποτε και δύο οι ιδιοτιμές του είναι μηδέν και συνεπώς ο πίνακας δεν είναι θετικά ορισμένος.

ii. αρχικό σημείο το $(-1, 1)$

Αριστερά απεικονίζεται ο Εσσιανός πίνακας και δεξιά οι ιδιοτιμές του με χρήση της συνάρτησης $eig()$ του MATLAB:

```
>> hessian_f(-1, 1)

ans =

   -0.2707   -0.8120
   -0.8120   -0.2707
```

```
>> eig(hessian_f(1, -1))

ans =

   -0.5413
    1.0827
```

Παρατηρείται ότι υπάρχει μία αρνητική ιδιοτιμή και συνεπώς ο πίνακας δεν είναι θετικά ορισμένος.

iii. αρχικό σημείο το $(1, -1)$

Αριστερά απεικονίζεται ο Εσσιανός πίνακας και δεξιά οι ιδιοτιμές του με χρήση της συνάρτησης $eig()$ του MATLAB:

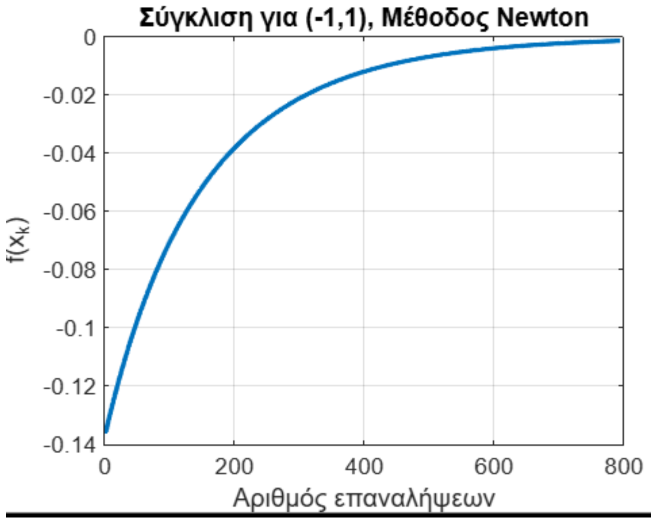
```
>> hessian_f(1, -1)

ans =

    0.2707    0.8120
    0.8120    0.2707
```

```
>> eig(hessian_f(1, -1))

ans =

   -0.5413
    1.0827
```

Παρατηρείται ότι υπάρχει μία αρνητική ιδιοτιμή και συνεπώς ο πίνακας δεν είναι θετικά ορισμένος.

Παρακάτω παρουσιάζεται συνοπτικά η **ελαττωματική εφαρμογή** της μεθόδου εάν δεν ληφθεί υπόψη ο έλεγχος ότι ο Εσσιανός πρέπει να είναι θετικά ορισμένος με ένα παράδειγμα γραφικής παράστασης για το *σημείο* $(-1, 1)$ και την *περίπτωση 2* (της **ελαχιστοποίησης** την $f(x_k + \gamma_k d_k)$ ως προς $\gamma_k > 0$, με τη **μέθοδο** της **διχοτόμου**):

```
Αρχικό σημείο: (-1.00, 1.00)
  Περίπτωση 1: x* = (-0.5962, -1.9745), f(x*) = -0.001070, Επαναλήψεις = 21
  Περίπτωση 2: x* = (-0.6689, -2.0361), f(x*) = -0.001356, Επαναλήψεις = 795
  Περίπτωση 3: x* = (-1.0000, -1.0000), f(x*) = -0.135335, Επαναλήψεις = 1199
Αρχικό σημείο: (1.00, -1.00)
  Περίπτωση 1: x* = (0.5962, 1.9745), f(x*) = 0.001070, Επαναλήψεις = 21
  Περίπτωση 2: x* = (0.3214, 1.4997), f(x*) = 0.000327, Επαναλήψεις = 5
  Περίπτωση 3: x* = (0.3190, 1.4996), f(x*) = 0.000315, Επαναλήψεις = 5
```




Για την *περίπτωση 3* στο *σημείο* $(-1,1)$, ο αλγόριθμος δεν τερμάτιζε για αυτό και έθεσα ένα όριο μέγιστων επαναλήψεων (1200-1). Όπως είναι φανερό και σύμφωνα με το ***Συμπέρασμα 5.2.3*** (σελ. 129 του *ΒΙΒΛΙΟΥ*) η μέθοδος Newton δεν ορίζεται για το πρόβλημα μας.

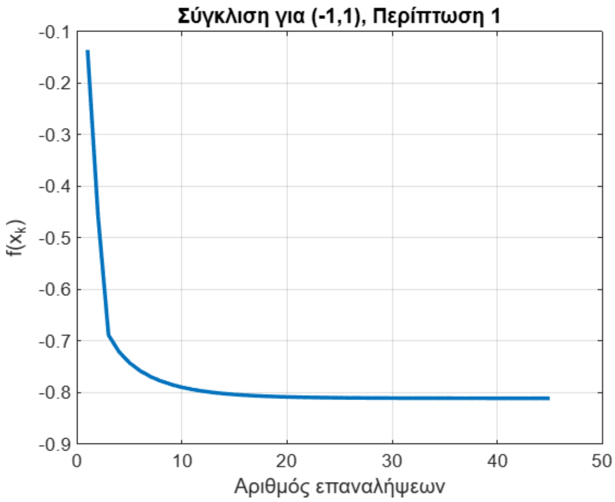
# <u>ΘΕΜΑ 4</u>

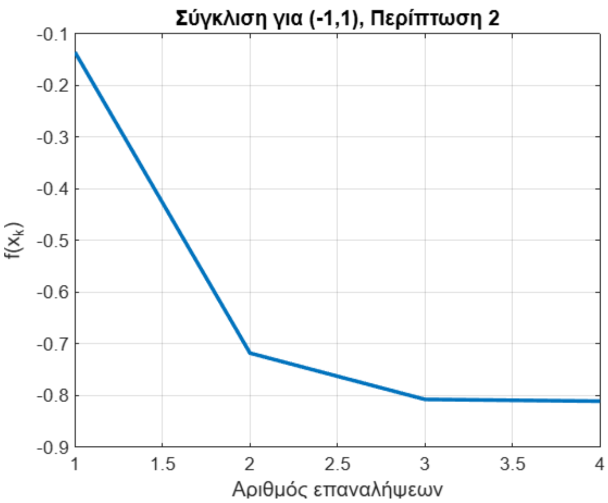
Στο παρόν θέμα υλοποιήθηκε στο MATLAB η **μέθοδος Levenberg-Marquardt**, όπως παρουσιάζεται στον ***Αλγόριθμο 5.2.3*** του *ΒΙΒΛΙΟΥ* στη σελίδα 139, και πρόκειται για τροποποιημένο αλγόριθμο Newton που χρησιμοποιείται **αν** ο **Εσσιανός** $\nabla^2 f(x_k)$ **δεν** είναι **θετικά ορισμένος**. Επομένως, επιλύει το πρόβλημα που αντιμετωπίστηκε στο ***Θέμα 3*** . Έτσι, όπως και στην Newton, η ιδέα της μεθόδου βασίζεται στη χρήση πληροφοριών της παραγώγου δεύτερης τάξης (**Εσσιανός** πίνακας), για να προσαρμόζει τη διαδικασία καθόδου. Πλέον, δοσμένου του $x_k$ στο *Βήμα 1,* πηγαίνουμε στο *Βήμα 2* και υπολογίζουμε το $\boldsymbol{\mu_k}$ έτσι ώστε ο $\nabla^2 f(x_k) + \mu_k I$ να είναι θετικά ορισμένος, όπου $I$ είναι ο μοναδιαίος $2 \times 2$ πίνακας και $\mu_k$ σταθερά που πρέπει απλώς να είναι μεγαλύτερη από την απόλυτη τιμή της μεγαλύτερης ιδιοτιμής του $\nabla^2 f(x_k)$ για να εξασφαλιστεί ότι ο $\nabla^2 f(x_k) + \mu_k I$ θα είναι θετικά ορισμένος ($\boldsymbol{delta} \rightarrow$ η θετική σταθερά που όρισα για να προσθέσω στην μεγαλύτερη κατ' απόλυτη τιμή ιδιοτιμή). Στο *Βήμα 3* υπολογίζουμε την **κατεύθυνση καθόδου** (αναζήτησης ελαχίστου) $d_k$ λύνοντας το σύστημα: $[\nabla^2 f(x_k) + \mu_k I] d_k = -\nabla f(x_k)$. Έπειτα, στο *Βήμα 4* το $\gamma_k$ υπολογίζεται σύμφωνα με τις τρεις περιπτώσεις που παρουσιάστηκαν στην ***ΠΕΡΙΓΡΑΦΗ ΠΡΟΒΛΗΜΑΤΟΣ*** και κατ' επέκταση ώστε να ικανοποιούνται τα Κριτήρια 3-4. Το Κριτήριο 3 που δεν αναφέρθηκε προηγουμένως είναι το δεύτερο σε κρισιμότητα κριτήριο γιατί αν επιλέξουμε αρκετά μικρό το βήμα δεν διακινδυνεύουμε την ευστάθεια του αλγορίθμου ή την σύγκλιση του, αλλά διακυβεύεται μόνο η ταχύτητα σύγκλισης. Τέλος, στο *Βήμα 5* θέτουμε $x_{k+1} = x_k + \gamma_k d_k$ και $k = k + 1$ και πηγαίνουμε στο *Βήμα 1.* Παρακάτω παρουσιάζονται οι γραφικές παραστάσεις της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων που απαιτούνται μέχρι να τερματίσει ο αλγόριθμος για τις τρεις περιπτώσεις επιλογής του $\gamma_k$:

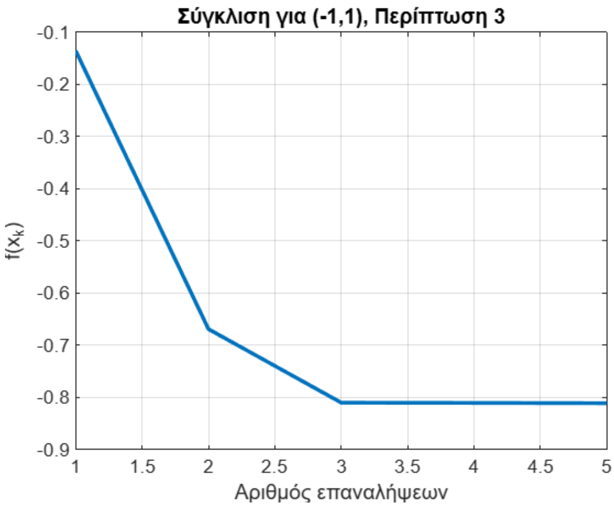
i. αρχικό σημείο το $(0, 0)$

Δεν θα παρουσιαστούν γραφικές παραστάσεις για το συγκεκριμένο σημείο διότι είναι ίδιες όπως και για το σημείο $(0, 0)$ για τη μέθοδο Μέγιστης Καθόδου στο ***Θέμα 2*** και αντίστοιχα η κλίση (gradient) της $f$ στο $(0, 0)$ είναι: $\nabla f(0, 0) = 0$. Το $(0, 0)$ όπως προαναφέρθηκε στο ***ΘΕΜΑ 1***, είναι σαγματικό σημείο και η μέθοδος δεν "ξεκολλαέι" από αυτό διότι θεωρεί ότι έχει φτάσει ήδη στο ελάχιστο λόγω της μηδενικής κλίσης. Τελικά έχουμε απόκλιση από την επιθυμητή τιμή ελαχίστου που είναι: $f(-1.58114, 0) = -0.811174$, λόγω εγκλωβισμού του αλγορίθμου στο κρίσιμο σημείο $(0, 0)$ που είναι σαγματικό. Αυτό είναι το αποτέλεσμα, ανεξάρτητα από την επιλογή του $\gamma_k$.

ii. αρχικό σημείο το $(-1, 1)$




Για κάθε επανάληψη υπολογίζεται η κλίση της $f$ και ο εσσιανός της συνάρτησης στο τρέχον σημείο $x_{\mathrm{k}}$. Ανάλογα με τη μέγιστη κατά απόλυτη τιμή του εσσιανού ($\lambda_{max}$) υπολογίζεται το $\mu_k$ ως: $\mu_k = \lambda_{max} + delta$ ώστε ο διορθωμένος εσσιανός να είναι θετικά ορισμένος και έπειτα υπολογίζεται η κατεύθυνση $d_k$ από το σύστημα που αναφέρθηκε παραπάνω. Για την ενημέρωση του $x_{\mathrm{k}}$ απαιτείται ο υπολογισμός του βήματος $\gamma_{\mathrm{k}}$ η οποία γίνεται σύμφωνα με την ανάλυση που προηγήθηκε για κάθε περίπτωση στην ***ΠΕΡΙΓΡΑΦΗ ΠΡΟΒΛΗΜΑΤΟΣ***:

a) Η **σταθερή** ενδεικτική τιμή που επέλεξα για το $\boldsymbol{\gamma_{\mathrm{k}}}$ , ώστε να πραγματοποιηθεί η απεικόνιση στη γραφική παράσταση, είναι **0.2** και η τιμή για το ***delta*** είναι 0.06. Όπως αναμενόταν παρατηρήθηκε ότι για πολύ μικρές τιμές του $\gamma_{\mathrm{k}}$ (πχ 0.001) η σύγκλιση ήταν πολύ αργή με τον αριθμό των επαναλήψεων να είναι πολύ μεγάλος, ενώ για τιμές περίπου 0.22 και μεγαλύτερες η διαδικασία αναζήτησης είναι ασταθής και φτάνει σε σημείο να αποκλίνει . Συγκριτικά με τις άλλες δύο περιπτώσεις θεωρήτικα είναι η λιγότερο αποδοτική όπως και αποδεικνύεται στην πράξη καθώς δεν βρέθηκε τιμή του $\gamma_{\mathrm{k}}$ ώστε να είναι παρόμοια αποδοτικά. Ο καλύτερος συνδυασμός που μπόρεσα να βρω ήταν με $\boldsymbol{\gamma_{\mathrm{k}} = 0.34}$ και $delta = 0.1$ όπου οι επαναλήψεις ήταν 18.

b) Μέσω της **ελαχιστοποίησης** την $f(x_k + \gamma_k d_k)$ ως προς $\gamma_k > 0$, με τη **μέθοδο** της **διχοτόμου**, το $\gamma_{\mathrm{k}}$ είναι προσαρμοσμένο ώστε να επιτυγχάνει τη μεγαλύτερη δυνατή μείωση της $f(x, y)$ σε κάθε βήμα, για αυτό και αναμένεται γρήγορη σύγκλιση (αξιόπιστα), αν και υπολογιστικά είναι πιο δαπανηρό από τις άλλες δύο περιπτώσεις καθώς απαιτεί πολλαπλούς υπολογισμούς της $f(x, y)$ σε κάθε επανάληψη. Επιλέχθηκαν η σταθερά $\varepsilon = 0.001$, το τελικό εύρος του διαστήματος $l = 0.01$, $delta = 0.06$ και αρχικό διάστημα αναζήτησης του $\gamma_k$ το [0, 5] καθώς η μέθοδος συγκλίνει στις λιγότερες επαναλήψεις (τέσσερις). Με την χρήση άλλων διαστημάτων πχ [0, 1] ή $delta$ η μέθοδος παρατηρήθηκε να συγκλίνει πιο αργά ή και να αποκλίνει. Γενικά είναι πολύ πιο αποδοτικό σε σύγκριση με το σταθερό $\gamma_k$.

c) Για τον κανόνα **Armijo** δοκιμάστηκαν διάφορες τιμές για τα $\alpha, \beta, s, delta$ και επέλεξα να κρατήσω ενδεικτικά τις $\alpha = 0.01, \beta = 0.5 \;\kappa\alpha\iota\; s = 5$ αν και συγκλίνει στις βέλτιστες επαναλήψεις για $s = 8$ και $delta = 0.1$ που μάλιστα είναι μία λιγότερη από την (b). Για άλλες τιμές $\beta$ και $s$ πχ $\beta = 0.3 \;\kappa\alpha\iota\; s = 1$ η μέθοδος παρατηρήθηκε να συγκλίνει πιο αργά. Γενικά, εξασφαλίζει μείωση της $f(x, y)$, αποφεύγοντας ταλαντώσεις ή υπερβολικά μεγάλα βήματα με αυτόματη προσαρμογή του $\gamma_k$, συγκλίνει αξιόπιστα και είναι λιγότερο υπολογιστικά δαπανηρή μέθοδος από την (b).

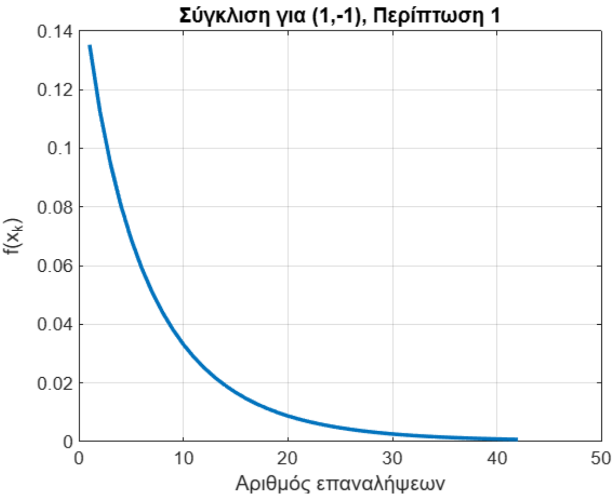
Παρακάτω εμφανίζονται συγκεντρωτικά τα αποτελέσματα για τις ενδεικτικές τιμές:

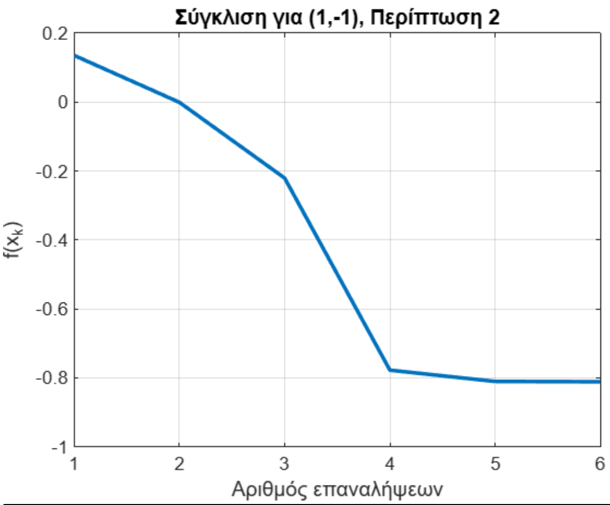
```
Αρχικό σημείο: (-1.00, 1.00)
  Περίπτωση 1: x* = (-1.5836, -0.0034), f(x*) = -0.811155, Επαναλήψεις = 45
  Περίπτωση 2: x* = (-1.5799, 0.0021), f(x*) = -0.811168, Επαναλήψεις = 4
  Περίπτωση 3: x* = (-1.5818, -0.0007), f(x*) = -0.811173, Επαναλήψεις = 5
```

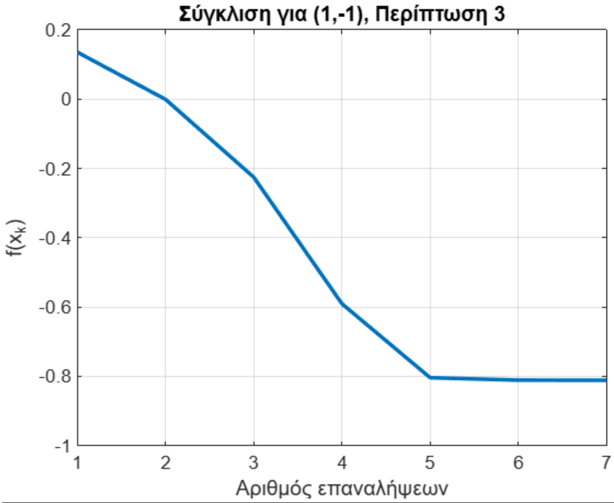
Τελικά έχουμε σύγκλιση στην επιθυμητή τιμή ελαχίστου που είναι:

$$f(-1.58114, 0) = -0.811174$$

iii. αρχικό σημείο το $(1, -1)$




Για κάθε επανάληψη υπολογίζεται η κλίση της $f$ και ο εσσιανός της συνάρτησης στο τρέχον σημείο $x_{\mathrm{k}}$. Ανάλογα με τη μέγιστη κατά απόλυτη τιμή του εσσιανού ($\lambda_{max}$) υπολογίζεται το $\mu_k$ ως: $\mu_k = \lambda_{max} + delta$ ώστε ο διορθωμένος εσσιανός να είναι θετικά ορισμένος και έπειτα υπολογίζεται η κατεύθυνση $d_k$ από το σύστημα που αναφέρθηκε παραπάνω. Για την ενημέρωση του $x_{\mathrm{k}}$ απαιτείται ο υπολογισμός του βήματος $\gamma_{\mathrm{k}}$ η οποία γίνεται σύμφωνα με την ανάλυση που προηγήθηκε για κάθε περίπτωση στην ***ΠΕΡΙΓΡΑΦΗ ΠΡΟΒΛΗΜΑΤΟΣ***:

a) Η **σταθερή** ενδεικτική τιμή που επέλεξα για το $\boldsymbol{\gamma_{\mathrm{k}}}$ , ώστε να πραγματοποιηθεί η απεικόνιση στη γραφική παράσταση, είναι **0.2** και η τιμή για το ***delta*** είναι 0.06. Όπως αναμενόταν παρατηρήθηκε ότι για πολύ μικρές τιμές του $\gamma_{\mathrm{k}}$ (πχ 0.001) η σύγκλιση ήταν πολύ αργή με τον αριθμό των επαναλήψεων να είναι πολύ μεγάλος, ενώ για τιμές περίπου 0.22 και μεγαλύτερες η διαδικασία αναζήτησης είναι ασταθής και φτάνει σε σημείο να αποκλίνει. Η διαφορά με το προηγούμενο σημείο είναι ότι η μέθοδος δεν συγκλίνει στο ελάχιστο σημείο αλλά φτάνει σε μία περιοχή (συγκλίνει σε σημείο) όπου η συνάρτηση φαίνεται επίπεδη ($\nabla f(\mathrm{x}, y) \sim 0$) σε σχέση με την κατεύθυνση καθόδου οπότε και σταματάει η εκτέλεση του αλγορίθμου ενώ έχει φτάσει κοντά στο σαγματικό σημείο (βλ. ***ΘΕΜΑ 1***). Συγκριτικά με τις άλλες δύο περιπτώσεις δεν βρέθηκε τιμή του $\gamma_{\mathrm{k}}$ ώστε να συγκλίνει στο ολικό ελάχιστο.

b) Μέσω της **ελαχιστοποίησης** την $f(x_k + \gamma_k d_k)$ ως προς $\gamma_k > 0$, με τη **μέθοδο** της **διχοτόμου**, το $\gamma_{\mathrm{k}}$ είναι προσαρμοσμένο ώστε να επιτυγχάνει τη μεγαλύτερη δυνατή μείωση της $f(x, y)$ σε κάθε βήμα, για αυτό και αναμένεται γρήγορη σύγκλιση (αξιόπιστα), αν και υπολογιστικά είναι πιο δαπανηρό από τις άλλες δύο περιπτώσεις καθώς απαιτεί πολλαπλούς υπολογισμούς της $f(x, y)$ σε κάθε επανάληψη. Επιλέχθηκαν η σταθερά $\varepsilon = 0.001$, το τελικό εύρος του διαστήματος $l = 0.01$, $delta = 0.06$ και αρχικό διάστημα αναζήτησης του $\gamma_k$ το [0, 5] καθώς η μέθοδος συγκλίνει στις λιγότερες επαναλήψεις (τέσσερις). Με την χρήση άλλων διαστημάτων πχ [0, 1] ή $delta$ η μέθοδος παρατηρήθηκε να συγκλίνει πιο αργά ή και να αποκλίνει. Σε σύγκριση με το σταθερό $\gamma_k$, η μέθοδος αυτή συγκλίνει στο ολικό ελάχιστο και είναι αρκετά αποδοτική.

c) Για τον κανόνα **Armijo** δοκιμάστηκαν διάφορες τιμές για τα $\alpha, \beta, s, delta$ και επέλεξα να κρατήσω ενδεικτικά τις $\alpha = 0.01, \beta = 0.5$ , $s = 5$ και $delta = 0.06$ καθώς έτσι συγκλίνει στις βέλτιστες επαναλήψεις (έξι). Για άλλες τιμές $\beta$, $s$ και $delta$ πχ $\beta = 0.3$ , $s = 1$ η μέθοδος παρατηρήθηκε να συγκλίνει πιο αργά ή και να αποκλίνει. Γενικά, εξασφαλίζει μείωση της $f(x, y)$, αποφεύγοντας ταλαντώσεις ή υπερβολικά μεγάλα βήματα με αυτόματη προσαρμογή του $\gamma_k$, συγκλίνει αξιόπιστα και είναι λιγότερο υπολογιστικά δαπανηρή μέθοδος από την (b).

Παρακάτω εμφανίζονται συγκεντρωτικά τα αποτελέσματα για τις ενδεικτικές τιμές:

```
Αρχικό σημείο: (1.00, -1.00)
  Περίπτωση 1: x* = (0.2979, -1.1357), f(x*) = 0.000591, Επαναλήψεις = 42
  Περίπτωση 2: x* = (-1.5808, -0.0009), f(x*) = -0.811173, Επαναλήψεις = 6
  Περίπτωση 3: x* = (-1.5819, 0.0003), f(x*) = -0.811173, Επαναλήψεις = 7
```

Τελικά για τις περιπτώσεις 2 και 3 (ελαχιστοποίηση με μέθοδο διχοτόμου και Armijo) έχουμε σύγκλιση στην επιθυμητή τιμή ελαχίστου που είναι:

$$f(-1.58114, 0) = -0.811174,$$

Ενώ για την περίπτωση 1 (σταθερό βήμα) έχουμε σύγκλιση στην τιμή:

$$f(0.2979, -1.1357) = 0$$

# <u>ΣΥΓΚΡΙΣΗ ΑΠΟΔΟΤΙΚΟΤΗΤΑΣ</u>

Συνοψίζοντας, μη λαμβάνοντας υπόψη τη μέθοδο Newton καθώς δεν ορίζεται για εσσιανό πίνακα που δεν είναι θετικά ορισμένος, από τη σύγκριση των αποτελεσμάτων που προκύπτουν από τις μεθόδους Μέγιστης Καθόδου και Levenberg-Marquardt, παρατηρούμε τα εξής:

- Όσον αφορά τους παράγοντες **σύγκλιση και αξιοπιστία** ο αλγόριθμος **Levenberg-Marquardt** καταφέρνει να συγκλίνει σταθερά στο ολικό ελάχιστο της συνάρτησης στις περισσότερες περιπτώσεις. Για παράδειγμα, τόσο για το αρχικό σημείο $(-1, 1)$ όσο και για το$(1, -1)$ με εξαίρεση την πρώτη περίπτωση, συγκλίνει με διαφορετικές επιλογές $\gamma_k$. Αντίθετα ο αλγόριθμος **Μέγιστης Καθόδου** δεν καταφέρνει να συγκλίνει για καμία επιλογή $\gamma_k$ στο σημείο $(1, -1)$. Αυτό δείχνει ότι η μέθοδος είναι πιο ευαίσθητη στις επιλογές των παραμέτρων και λιγότερο αξιόπιστη.

- Όσον αφορά την **ταχύτητα σύγκλισης** ο αλγόριθμος **Levenberg-Marquardt** εμφανίζεται ελαφρώς πιο αποδοτικός. Για παράδειγμα, στο σημείο $(-1, 1)$ , η μέθοδος **Levenberg-Marquardt** με τον κανόνα Armijo (Περίπτωση 3) συγκλίνει σε μόλις 3 επαναλήψεις, ενώ η **Μέγιστης Καθόδου** χρειάζεται 5 επαναλήψεις. Γενικά η μέθοδος **Μέγιστης Καθόδου** λόγω της απλής κατεύθυνσης που υπολογίζει, χρειάζεται περισσότερες επαναλήψεις για να πλησιάσει την ελάχιστη τιμή, εάν συγκλίνει.

- Όσον αφορά την **ακρίβεια τελικού αποτελέσματος** και οι δύο μέθοδοι όταν συγκλίνουν στο ολικό ελάχιστο $f(-1.58114, 0) = -0.811174$ καταλήγουν σε παρόμοια τιμή.