ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ
ΤΗΜΜΥ ΤΟΜΕΑΣ ΗΛΕΚΤΡΟΝΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΕΧΝΙΚΕΣ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ

2Η ΕΡΓΑΣΙΑ

Ιωάννης Δεϊρμεντζόγλου   10015

ΑΕΜ 10015   Email deirmentz@ece.auth.gr

Εισαγωγή

Σκοπός της δεύτερης εργασίας είναι η ελαχιστοποίηση μιας συνάρτησης 2 μεταβλητών (δισδιάστατος χώρος), χωρίς περιορισμούς. Υλοποιήθηκαν οι παρακάτω τρεις μέθοδοι ελαχιστοποίησης :

- η μέθοδος της μέγιστης καθόδου (Steepest Descent),
- η μέθοδος Newton
- η μέθοδος Levenberg-Marquardt

Οι τρείς αυτές μέθοδοι ανήκουν στις μεθόδους κλίσης και βασική ιδιότητα τους είναι η χρήση του διανύσματος κλίσης, $\nabla f(x)$, στην ισοβαρή καμπύλη της f. Η αντικειμενική συνάρτηση που θα μελετήσουμε είναι :

$$\boldsymbol{f(x,y) = x^3 \cdot e^{(x^2 - y^4)}}$$

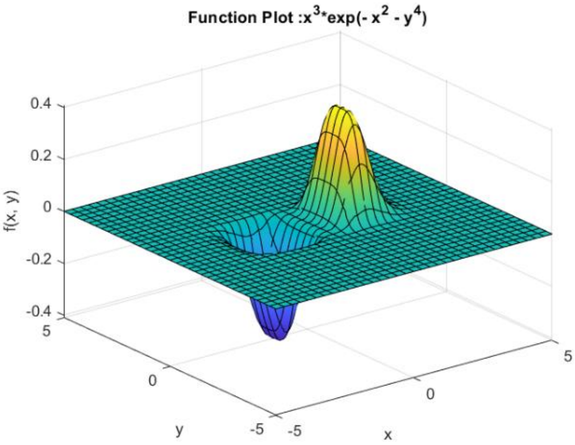
Για κάθε μία μελετήθηκαν 3 τρόποι επιλογής της παραμέτρου $\gamma\kappa$ αλλά και 3 διαφορετικά σημεία εκκίνησης. Τα σημεία για τα οποία θα πραγματοποιηθεί η μελέτη είναι τα **(0,0) , (−1,−1) , (1,1)** και η τιμή του $\gamma\kappa$ επιλέγεται :

- σταθερή
- έτσι ώστε να ελαχιστοποιεί την $f(\chi\kappa+\gamma\kappa*dk)$
- βάσει του κανόνα Armijo .Για τον Armijo επιλέχθηκαν οι τιμές α=0.001 $\kappa\alpha\iota$ β=0.2.

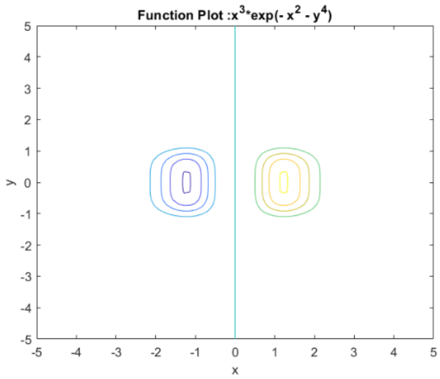
Το κατώφλι τερματισμού (epsilon = η τιμή για την οποία το gradient της συνάρτησης στο σημείο $\chi\kappa$ θεωρείται μηδέν) τέθηκε σε κάθε περίπτωση που εξετάστηκε στο 0.01. Για την ελαχιστοποίηση της συνάρτησης και την εύρεση του $\gamma\kappa$ χρησιμοποιήθηκε η μέθοδος της διχοτόμου με χρήση παραγώγου που είχε υλοποιηθεί στην πρώτη εργασία.

1. Task 1 – Γραφική παράσταση της συνάρτησης

Στο πρώτο θέμα ζητήθηκε να σχεδιαστή η συνάρτηση f για να δοθεί μια γενική εικόνα της μορφής στον τρισδιάστατο χώρο.
Παρακάτω παρουσιάζεται η μορφή της συνάρτησης στον τρισδιάστατο χώρο αλλά και οι ισοβαρείς καμπύλες της. Από τη γραφική παράσταση της συνάρτησης μπορεί να εξαχθεί το συμπέρασμα ότι η ελάχιστη τιμή βρίσκεται σε αρνητικά x και y. Ο κώδικας βρίσκεται στο αρχείο Task1_Plot.m .



Σχήμα 1.1 : Plot της αντικειμενικής συνάρτησης f



Σχήμα 1.2 : Ισοβαρείς καμπύλες της f

Από τη γραφική παράσταση της συνάρτησης μπορεί να εξαχθεί το συμπέρασμα ότι η ελάχιστη τιμή βρίσκεται σε αρνητικά x και y. Πράγματι , το ελάχιστο της f έχει τιμή -0.4099 και βρίσκεται στο σημείο (-1.2247,-5.539e-05)(περίπου – ο κώδικας βρίσκεται στο παραπάνω αρχείο) . Παρατηρείται πως δεν είναι κυρτή σε όλο το πεδίο ορισμού της καθώς έχει ένα όρος και μια κοιλότητα. Εκ των προτέρων αναμένεται πως τιμές με θετικά x μεγαλύτερα του 1 θα έχουν πιθανόν λανθασμένο αποτέλεσμα με τους αλγορίθμους που θα χρησιμοποιηθούν.

<u>Σημείωση</u>: Όταν το αρχικό σημείο είναι (0, 0) , παρατηρούμε σε κάθε περίπτωση πως ο αλγόριθμος τερματίζει αμέσως (k=1 iteration). Αυτό βγάζει νόημα αν σκεφτούμε την συνθήκη τερματισμού, $|\nabla f| < \epsilon\ ,\ \epsilon > 0$. Για το σημείο (0, 0) προφανώς έχουμε το μηδενικό διάνυσμα $(\nabla f\ (0,0)=\vec{0})$. Βλέπουμε λοιπόν ότι το σημείο αυτό είναι σημείο εγκλωβισμού.

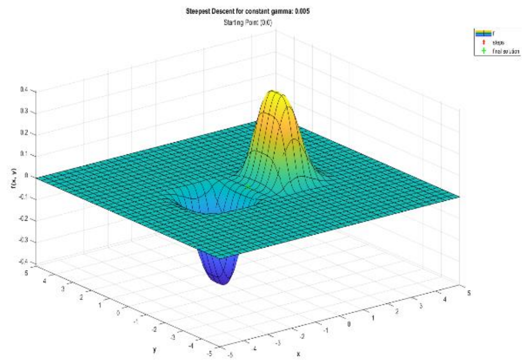
2. Task 2 – Μέθοδος Steepest Descent (Μέγιστης Καθόδου)

Στο Θέμα 2 η αντικειμενική συνάρτηση f ελαχιστοποιείται με την χρήση της μεθόδου Μεγίστης Καθόδου . Υλοποιήθηκαν 3 διαφορετικά .m αρχεία ανάλογα με τον τρόπο επιλογής του βήματος γ_k . Ο κώδικας για σταθερό βήμα βρίσκεται στο αρχείο Task2_ConstGamma.m , για βημα που ελαχιστοποιεί την $f(\chi\kappa+\gamma\kappa*dk)$
στο αρχειο Task2_MinGamma.m και για βημα υπολογισμένο βάσει του κανόνα του Armijo στο Task2_ArmijoGamma.m .

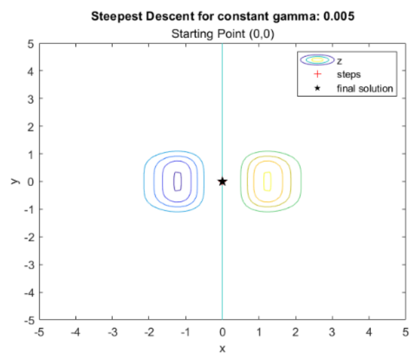
Σε όλα τα διαγράμματα του 2ου θέματος χρησιμοποιήθηκε $\epsilon = 0.01$ ενώ για τις περιπτώσεις (α) όπου ζητούταν σταθερό γ_κ τέθηκε αυθαίρετα $\gamma_\kappa = 0.05$. Στην μέθοδο Armijo επιλέχτηκαν αρχικές συνθήκες $\alpha = 0.001$ και $\beta = 0.2$.

- **Αρχικό σημείο [0, 0]**
  Το σημείο δεν θα μπει μέσα στον αλγόριθμο γιατί θα έχει κλίση 0, οπότε μένει στο σημείο [0, 0] (σημείο εγκλωβισμού). Το ίδιο συμβαίνει για οποιοδήποτε βήμα. Ως μηδενική επανάληψη θεωρείται αυτή για το σημείο εκκίνησης. Όταν, λοιπόν, ως αρχικό σημείο τεθεί το (0,0), η μέθοδος της μέγιστης καθόδου εγκλωβίζεται σε αυτό . Ενδεικτικά , παρουσιάζονται τα παρακάτω διαγράμματα (για κάθε μέθοδο τα διαγράμματα που προκύπτουν είναι όπως είναι λογικό είναι ακριβώς τα ίδια).

Α) Σταθερό γ_κ
Β) γ_κ που ελαχιστοποιεί τη συνάρτηση $f(\chi\kappa+\gamma\kappa*dk)$
Γ) γ_κ υπολογισμένο με τη μέθοδο armijo



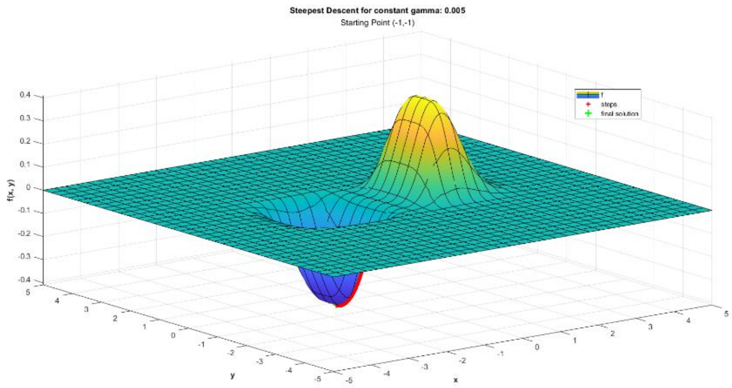
Σχήμα 2.1 :Steepest Descend, γ_k= 0.005,starting point = (0, 0)



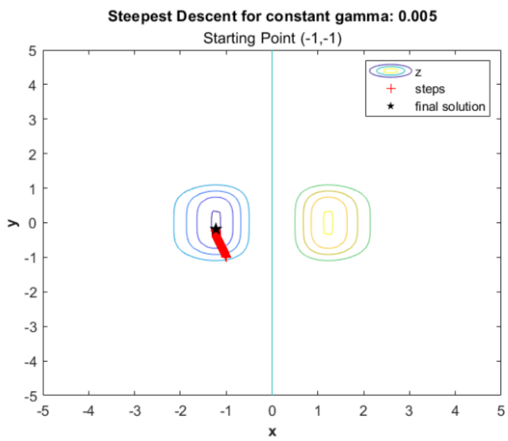
Σχήμα 2.2 :Steepest Descend, γ_k= 0.005, starting point = (0, 0)

- **Αρχικό σημείο [-1, -1]**
  Όταν για αρχικό σημείο τεθεί το (−1,−1), παρατηρείται ότι ο αλγόριθμος φτάνει πολύ κοντά στο ελάχιστο (ανάλογα με το epsilon που έχει επιλεχτεί) και με τις τρεις μεθόδους επιλογής του $\gamma\kappa$. Από τα διαγράμματα της σύγκλισης της συνάρτησης συναρτήσει του αριθμού της επανάληψης φαίνεται ότι η μέθοδος του σταθερού $\gamma\kappa$ και χρειάζεται πολλές επαναλήψεις και επομένως είναι αργή , ενώ οι άλλες 2 μέθοδοι είναι σαφώς πιο γρήγορες αλλά το ίδιο αποτελεσματικές. Η μέθοδος με τον κανόνα Armijo χρειάζεται πολύ λιγότερες επαναλήψεις και από την $2^{η}$ μέθοδο.

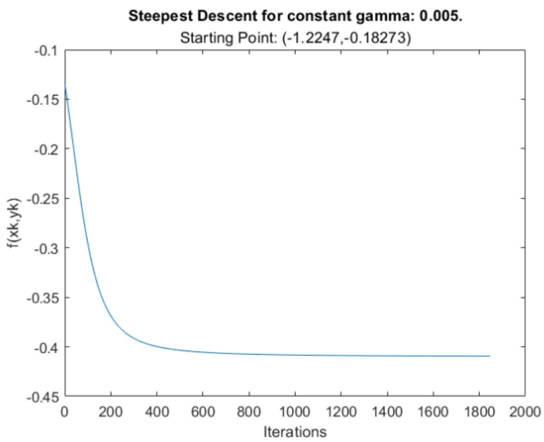
Α) Σταθερό γ_κ



Σχήμα 2.3 :Steepest Descend, γ_k= 0.005,starting point = (-1, -1)

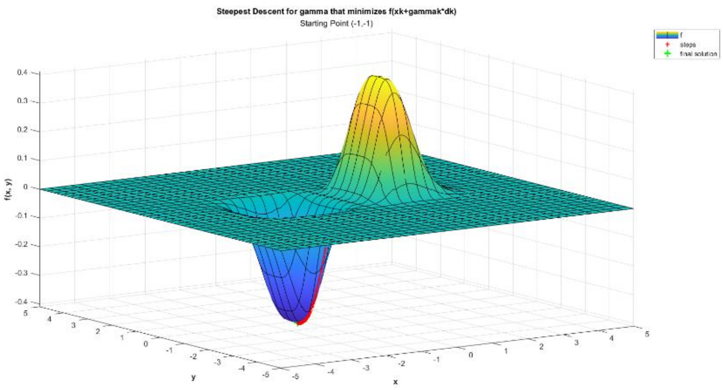


Σχήμα 2.4 :Steepest Descend, γ_k= 0.005,starting point = (-1, -1)

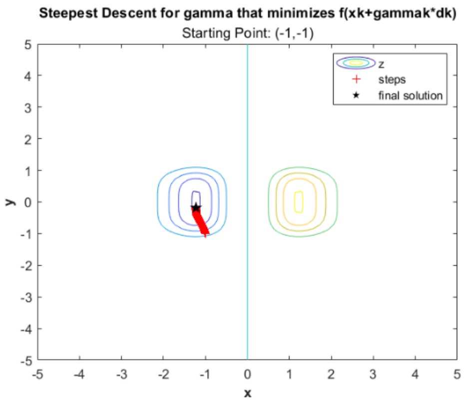


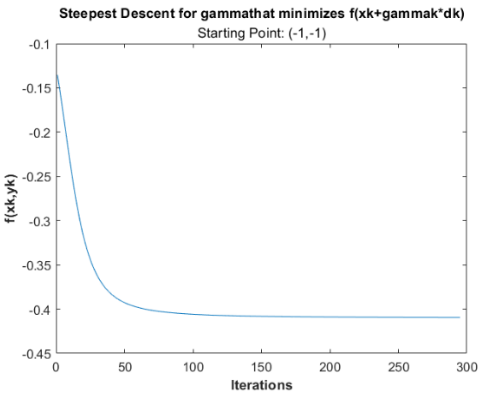
Σχήμα 2.5 :Steepest Descend, γ_k= 0.005,starting point = (-1 ,-1)

Β) $γ_κ$ που ελαχιστοποιεί τη συνάρτηση $f(χκ+γκ*dk)$



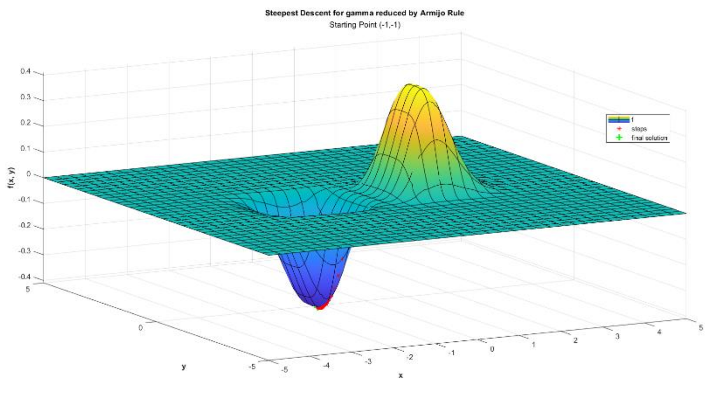
Σχήμα 2.6 :Steepest Descend, $γ_k$= minimizes minimizes f(xk + gkdk)



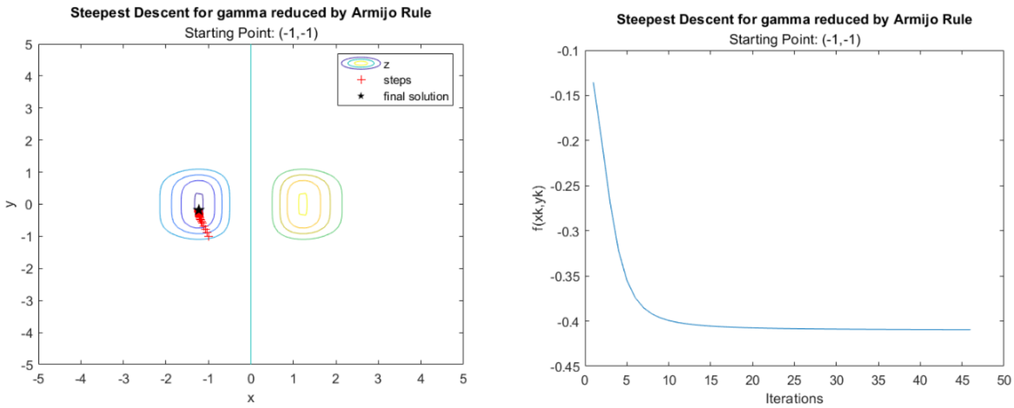


Σχήματα 2.6 , 2.7 :Steepest Descend, $γ_k$= minimizes minimizes f(xk + gkdk)

Γ) $γ_κ$ υπολογισμένο με τη μέθοδο armijo



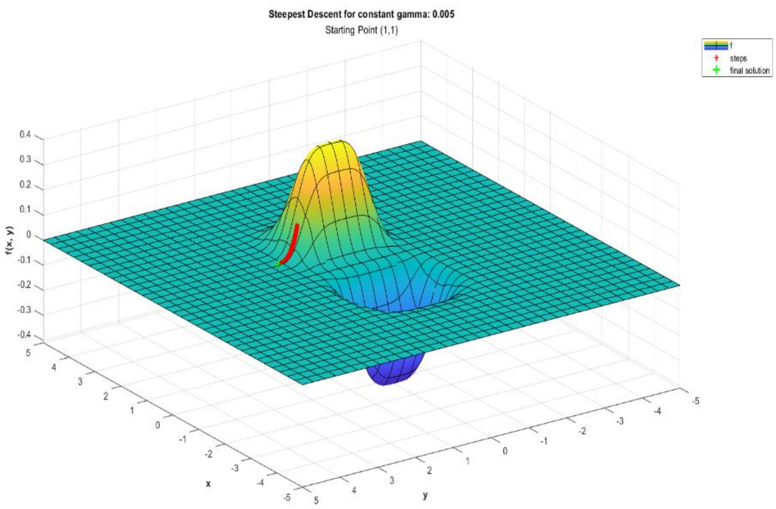
Σχήμα 2.8 : :Steepest Descend, γ_k reduced by Armijo's Rule



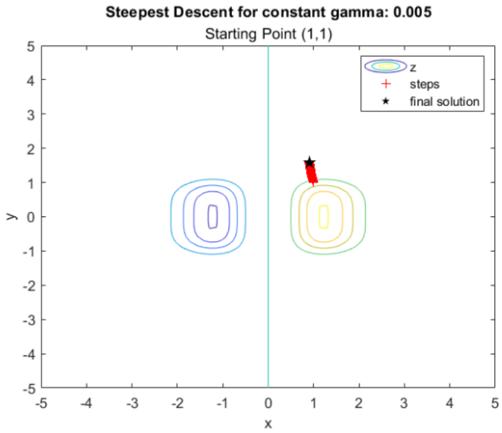
Σχήματα 2.9, 2.10 :Steepest Descend, γ_k reduced by Armijo's Rule

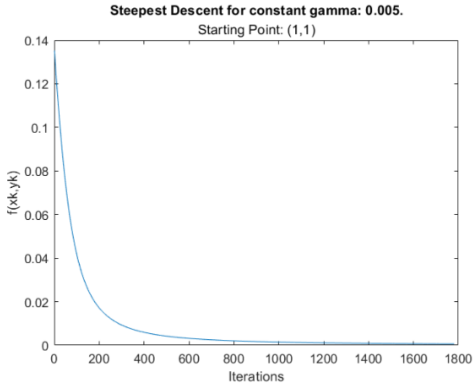
- **Αρχικό σημείο [1, 1]**
  Με αρχικό σημείο το (1,1), η μέθοδος εγκλωβίζεται πάλι όπως και με το (0,0), σε τοπικό ελάχιστο και αδυνατεί να βρει το πραγματικό. Εξάγεται, λοιπόν το συμπέρασμα πως η επιλογή του σημείου έναρξης του αλγορίθμου παίζει καθοριστικό ρόλο στο αν τελικά ο αλγόριθμος θα καταλήξει να βρει το ολικό ελάχιστο της συνάρτησης. Σε ότι αφορά την ταχύτητα της σύγκλισης ανάλογα με τη μέθοδο επιλογής του $\gamma\kappa$ ισχύει ότι αναφέρθηκε και προηγουμένως και για το σημείο $(-1,-1)$.

Α) Σταθερό γ_κ



Σχήμα 2.11 :Steepest Descend, γ_k= 0.005,starting point = (1, 1)

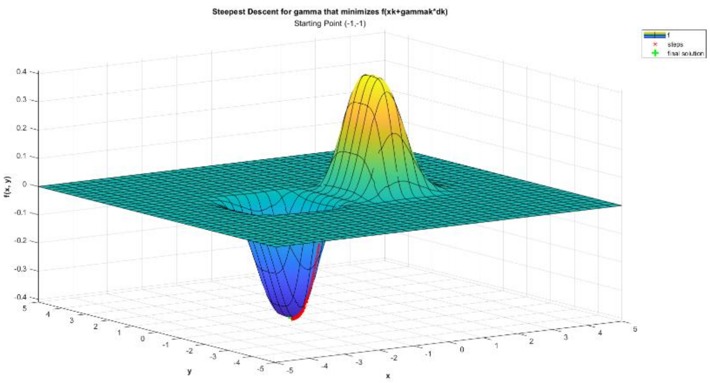




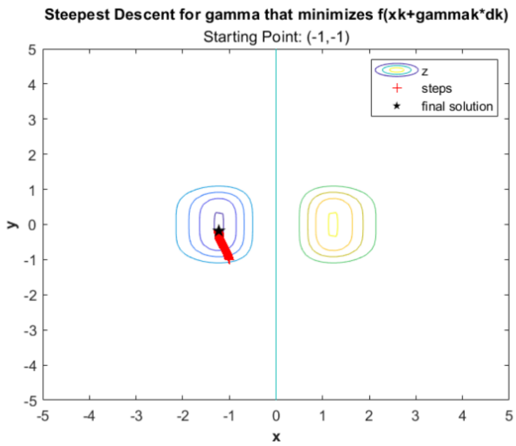
Σχήμα 2.12 :Steepest Descend, γ_k= 0.005,starting point = (1, 1)

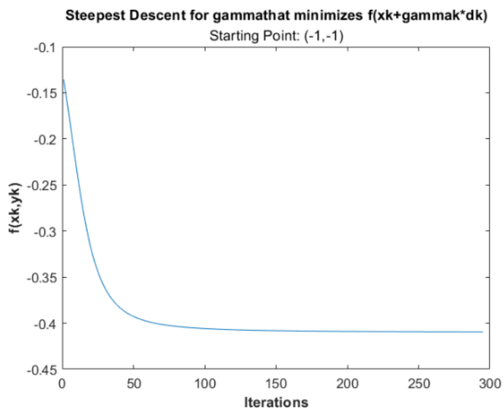
Σχήμα 2.13 :Steepest Descend, γ_k= 0.005

B) γ_κ που ελαχιστοποιεί τη συνάρτηση $f(\chi\kappa+\gamma\kappa*dk)$



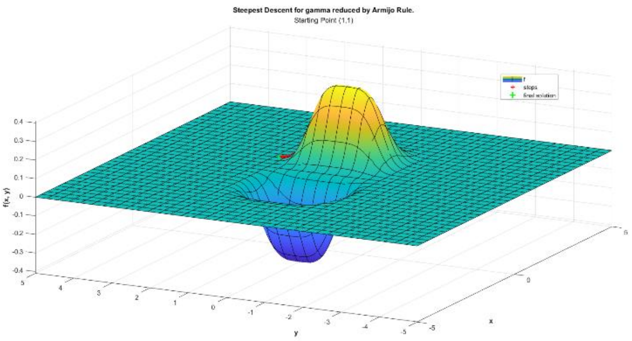
Σχήμα 2.6 :Steepest Descend, γ_k= minimizes minimizes f(xk + gkdk)



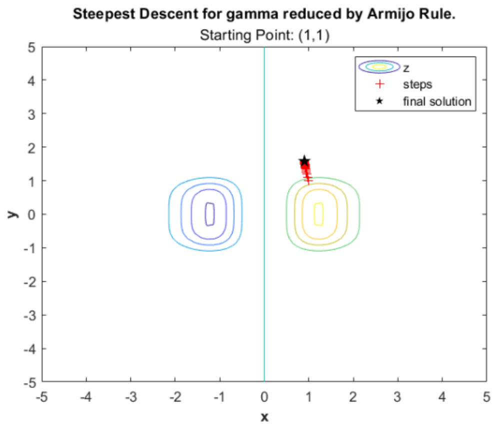


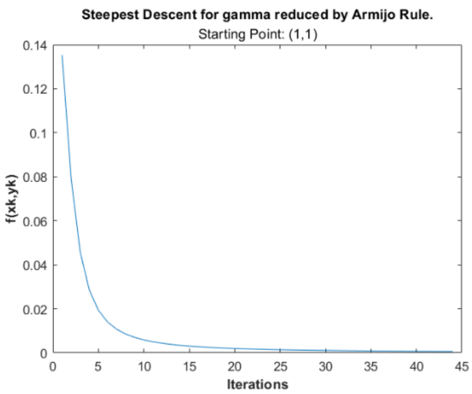
Σχήματα 2.15 , 2.16 :Steepest Descend, γ_k= minimizes minimizes f(xk + gkdk)

Γ) γ_k υπολογισμένο με τη μέθοδο armijo



Σχήμα 2.17 : :Steepest Descend, γ_k reduced by Armijo's Rule





Σχήματα 2.18, 2.19 :Steepest Descend, γ_k reduced by Armijo's Rule

3. Task 3 – Μέθοδος Newton

Αντίθετα με την μέθοδο μεγίστης καθόδου η μέθοδος Newton αναζητεί σημείο ελαχίστου στην κατεύθυνση **dk = −[∇^2f(xk)]−1 • ∇f(xk)** χρησιμοποίει δηλαδή και τον εσσιανό πίνακα ο οποίος θα πρέπει όμως να είναι θετικά ορισμένος(ή ημιορισμένος). Στην συγκεκριμένη συνάρτηση ο πίνακας δεν είναι ημιορισμένος

Για της μεθόδους ελαχιστοποίησης της $f(\chi\kappa+\gamma\kappa*dk)$ και του Armijo τέθηκε ως κριτήριο ο αλγόριθμος να σταματάει όταν η τιμή της συνάρτησης για $\chi\kappa$ που υπολογίστηκε είναι μεγαλύτερη ή ίση με την τιμή της συνάρτησης στο $\chi\kappa-1$. Σε διαφορετική περίπτωση, ο αλγόριθμος δεν συγκλίνει ποτέ σε σημείο ελαχίστου. Επιπλέον, στη μέθοδο α όπου το $\gamma\kappa$ παραμένει σταθερό αντί να η συνάρτηση να ελαχιστοποιείται, η τιμή της στο τελικό σημείο έχει μεγαλύτερη τιμή από την τιμή στο σημείο έναρξης.

Ωστόσο, ο αλγόριθμος υλοποιήθηκε και βρίσκεται στα αρχεία Task3_ConstGamma.m , Task3_MinGamma.m , Task3_ArmijoGamma.m . Στο σημείο (0,0) η συνάρτηση εγκλωβίζεται ξανά σε τοπικό ελάχιστο (μιας και η κλίση της συνάρτησης στο σημείο είναι 0). Επίσης , παρουσιάζονται και τα αντίστοιχα διαγράμματα για λογούς πληρότητας πάρα το γεγονός ότι μέθοδος δεν δουλεύει σωστά .

Σε όλα τα διαγράμματα του 3ου θέματος χρησιμοποιήθηκε ϵ = 0.01 ενώ για τις περιπτώσεις (α) όπου ζητούταν σταθερό γ_κ τέθηκε αυθαίρετα γ_κ = 0.05. Στην μέθοδο Armijo επιλέχτηκαν αρχικές συνθήκες α = 0.001 και β = 0.2.

Για τα 3 σημεία που έχουμε ο Εσσιανός πίνακας είναι:

- Για το [0, 0]:

```
[0, 0]
[0, 0]
```

- Για το [-1, -1]:

```
[4*exp(-2),  4*exp(-2)]
[4*exp(-2), -4*exp(-2)]
```

- Για το [1, 1]:

```
[-4*exp(-2), -4*exp(-2)]
[-4*exp(-2),  4*exp(-2)]
```

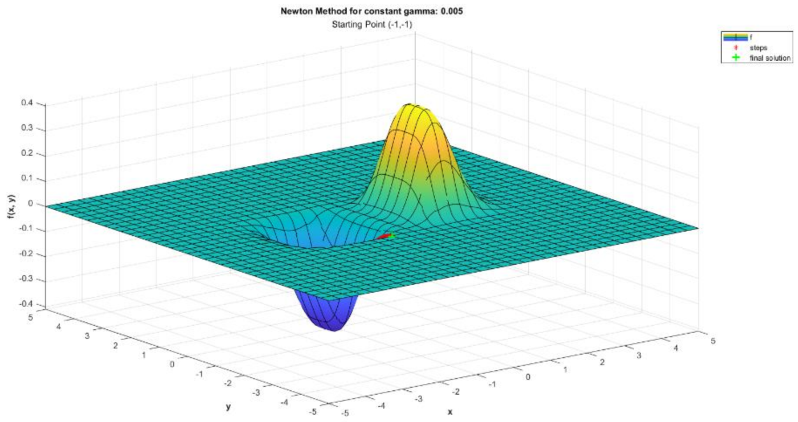
Για κανένα από τα τρία αρχικά σημεία δεν έχουμε θετικά ορισμένο πίνακα. Το αποτέλεσμα είναι η μέθοδος να μην τερματίζει ή να βγάζει λανθασμένα αποτελέσματα .

- **Αρχικό σημείο [0, 0]**

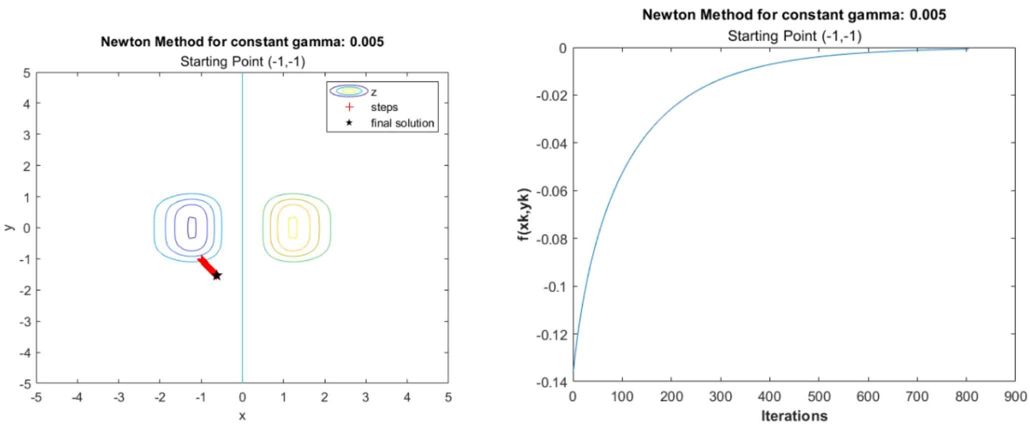
Το σημείο δεν θα μπει μέσα στον αλγόριθμο γιατί θα έχει κλίση 0, οπότε μένει στο σημείο [0, 0] (σημείο εγκλωβισμού). Το ίδιο συμβαίνει για οποιοδήποτε βήμα. Ως μηδενική επανάληψη θεωρείται αυτή για το σημείο εκκίνησης. Όταν, λοιπόν, ως αρχικό σημείο τεθεί το (0,0), η μέθοδος Newton εγκλωβίζεται σε αυτό . Τα διαγράμματα προφανώς είναι πανομοιότυπα με αυτά στην ενότητα 2 (2.1 2.2) .

- **Αρχικό σημείο [-1, -1]**

Α) Σταθερό $γ_κ$

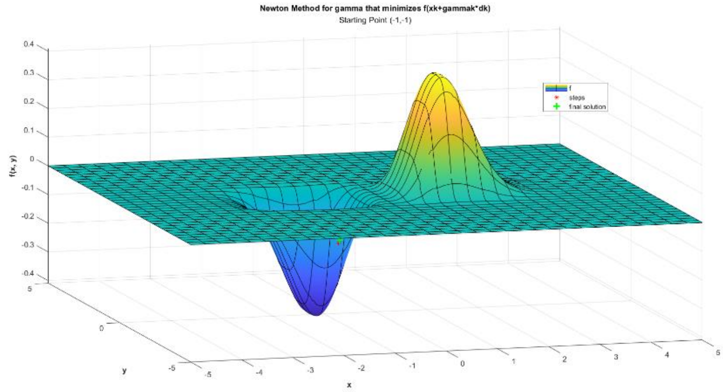


Σχήμα 3.1 :Newton method, $γ_k$= 0.005 constant

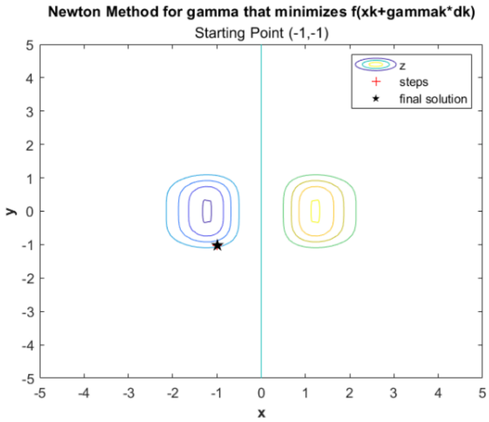


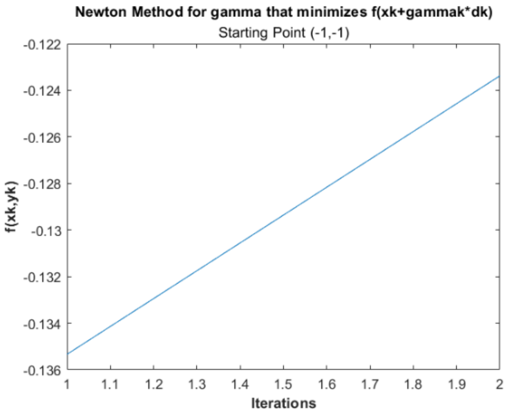
Σχήματα 3.2 , 3.3 :Newton method, $γ_k$= 0.005 constant

B) γ_κ που ελαχιστοποιεί τη συνάρτηση $f(\chi\kappa+\gamma\kappa*dk)$



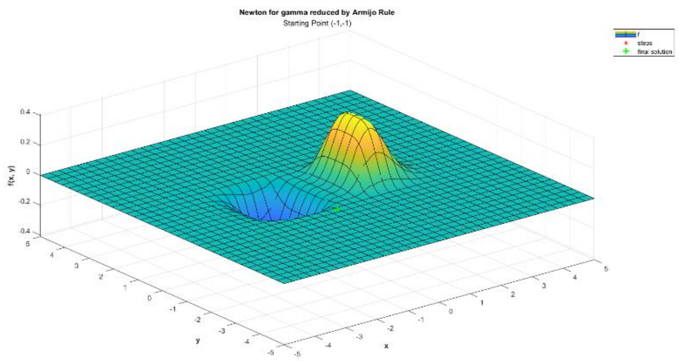
Σχήμα 3.4 : Newton method, γ_k= minimizes minimizes f(xk + gkdk)



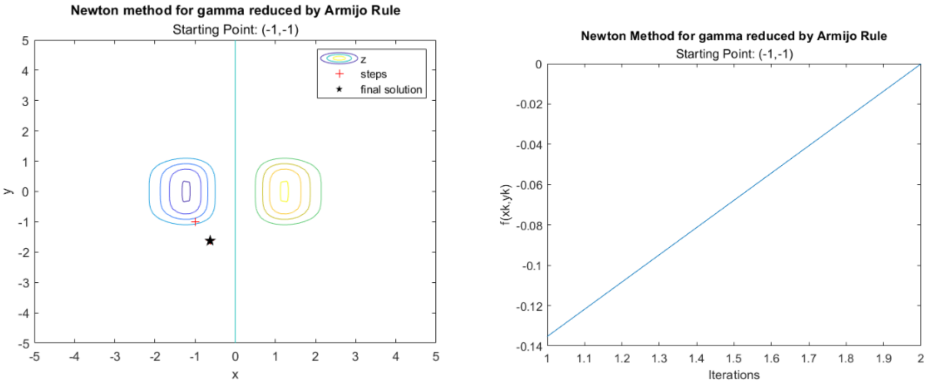


Σχήματα 3.5 , 3.6 :Newton method, γ_k= minimizes minimizes f(xk + gkdk)

Γ) γ_κ υπολογισμένο με τη μέθοδο armijo



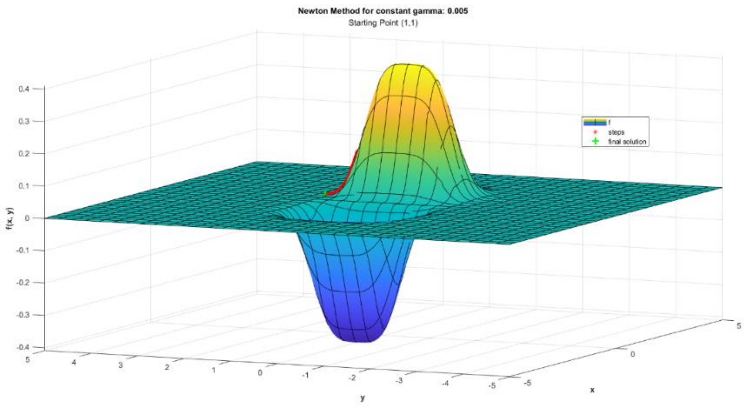
Σχήμα 3.7 : Newton method, γ_k reduced by Armijo's Rule



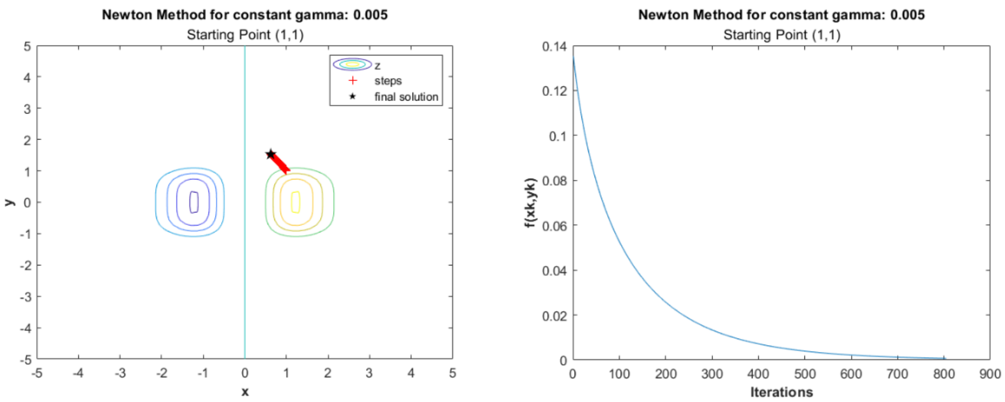
Σχήματα 3.8 , 3.9 :Newton method, γ_k reduced by Armijo's Rule

- **Αρχικό σημείο [1, 1]**
  Για το σημείο (1,1) ο αλγόριθμος έχει παρόμοια αποτελέσματα με αυτόν τη μέγιστης καθόδου, δηλαδή δεν βρίσκει το ολικό ελάχιστο αλλά εγκλωβίζεται σε τοπικό .

A) Σταθερό γ_k

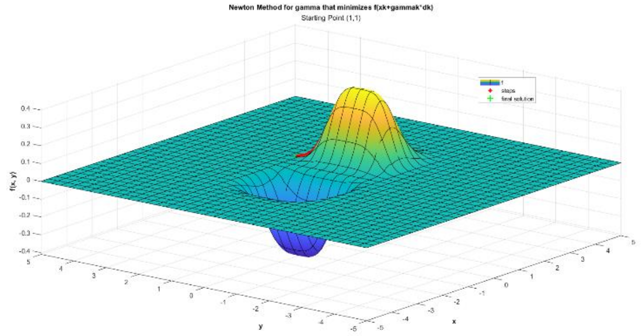


Σχήμα 3.10 : :Newton Method, γ_k constant

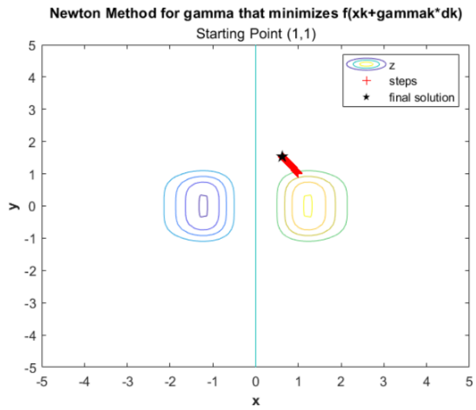


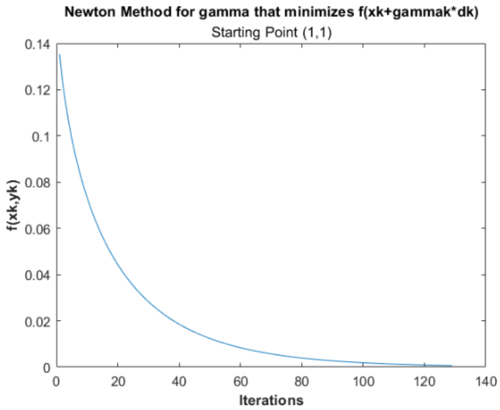
Σχήματα 3.11 , 3.12 : :Newton Method, γ_k constant

Β) γ_κ που ελαχιστοποιεί τη συνάρτηση $f(\chi\kappa+\gamma\kappa*dk)$



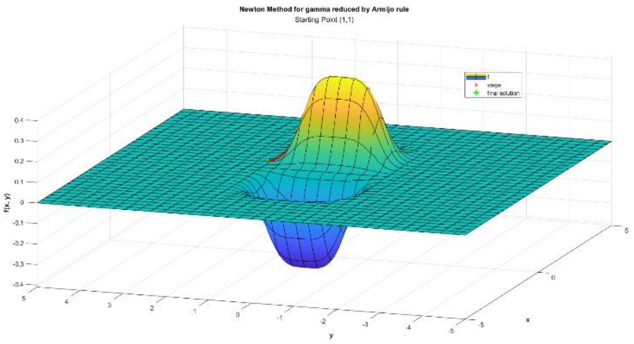
Σχήμα 3.13 : Newton method, γ_k= minimizes minimizes f(xk + gkdk)



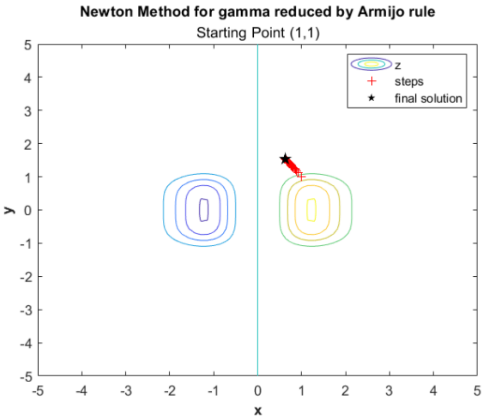


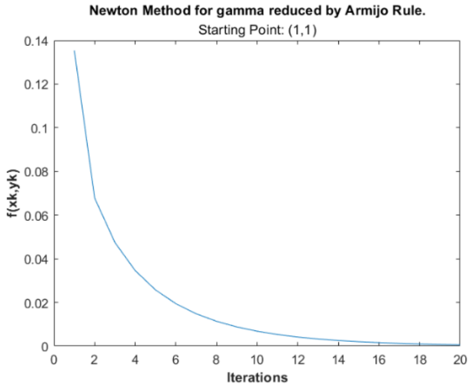
Σχήματα 3.14, 3.15 :Newton method, γ_k= minimizes minimizes f(xk + gkdk)

Γ) γ_κ υπολογισμένο με τη μέθοδο armijo



Σχήμα 3.16 : Newton method, γ_k reduced by Armijo's Rule





Σχήματα 3.17 , 3.18 :Newton method, γ_k reduced by Armijo's Rule

4. Task 4 – Μέθοδος Levenberg-Marquardt

Η συγκεκριμένη μέθοδος είναι ο τροποποιημένος αλγόριθμος του Newton. Η μέθοδος Levenberg Marquardt λύνει το πρόβλημα της μεθόδου του Newton όταν ο πίνακας δεν είναι θετικά ορισμένος. Υπολογίζεται το $\mu\kappa$ ώστε ο πίνακας $[\nabla^2 f(x)+\mu\kappa I]$ να είναι θετικά ορισμένος .

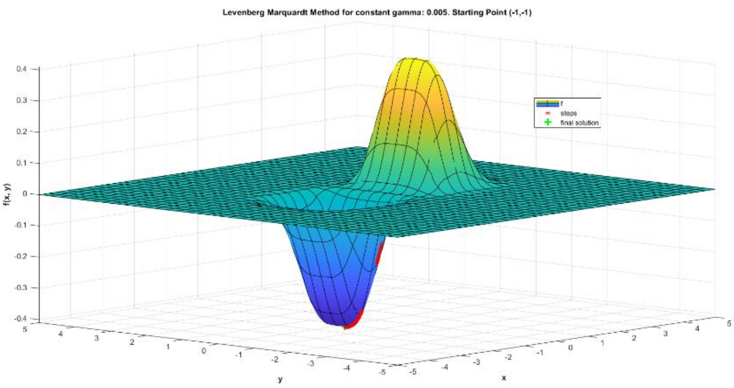
Ο αλγόριθμος που υλοποιήθηκε και βρίσκεται στα αρχεία Task4_ConstGamma.m , Task4_MinGamma.m , Task4_ArmijoGamma_v2.m , Armijo.m , func.m.

- **Αρχικό σημείο [0, 0]**

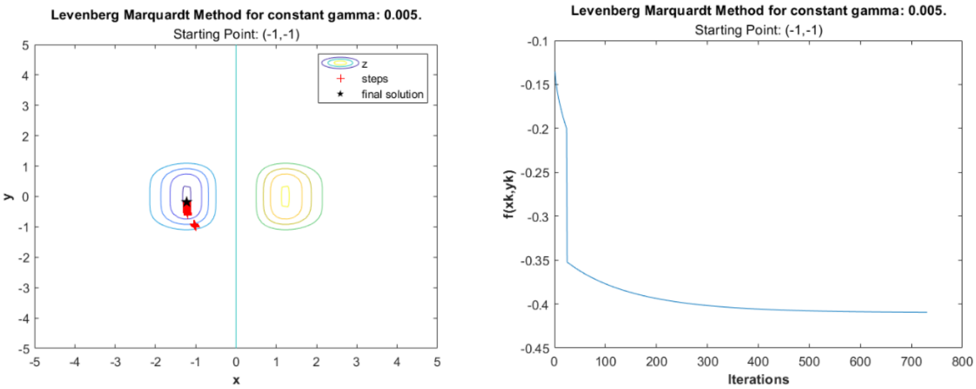
Το σημείο δεν θα μπει μέσα στον αλγόριθμο γιατί θα έχει κλίση 0, οπότε μένει στο σημείο [0, 0] (σημείο εγκλωβισμού). Το ίδιο συμβαίνει για οποιοδήποτε βήμα. Ως μηδενική επανάληψη θεωρείται αυτή για το σημείο εκκίνησης. Όταν, λοιπόν, ως αρχικό σημείο τεθεί το (0,0), η μέθοδος Levenberg Marquardt εγκλωβίζεται σε αυτό . Τα διαγράμματα προφανώς είναι πανομοιότυπα με αυτά στην ενότητα 2 (2.1 2.2) .

- **Αρχικό σημείο [-1, -1]**

Α) Σταθερό γ_κ

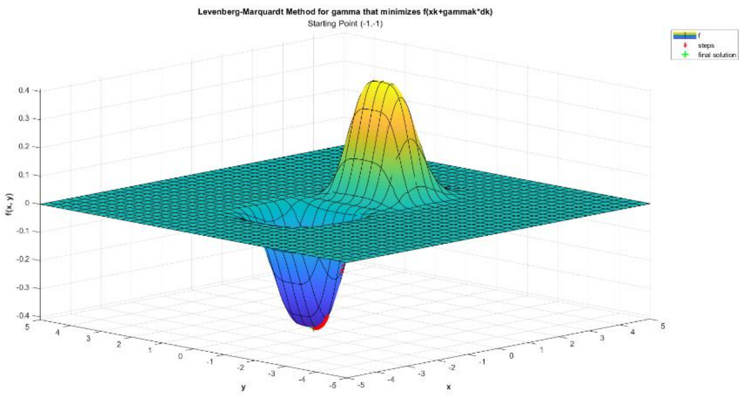


Σχήμα 4.1 :Levenberg Marquardt method, γ_k= 0.005 constant

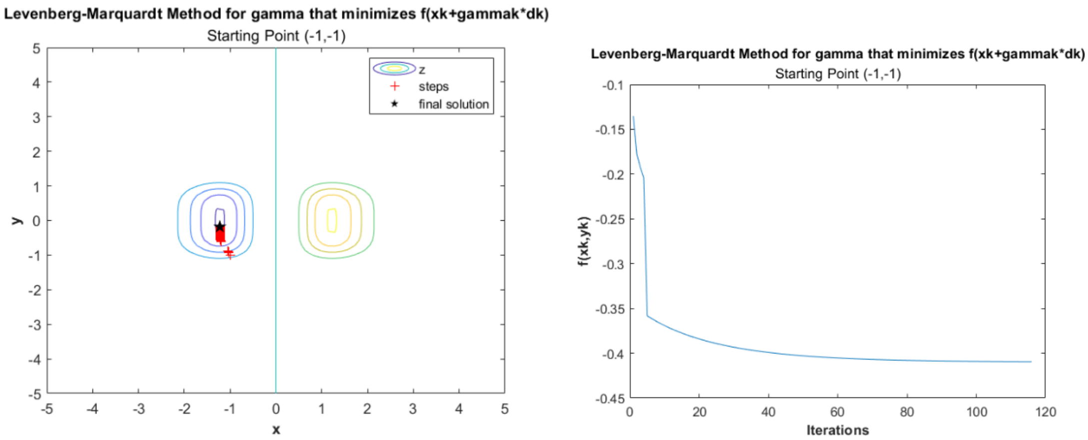


Σχήματα 4.2 , 4.3 : Levenberg Marquardt method, γ_k= 0.005 constant

B) γ_κ που ελαχιστοποιεί τη συνάρτηση $f(\chi\kappa + \gamma\kappa * dk)$

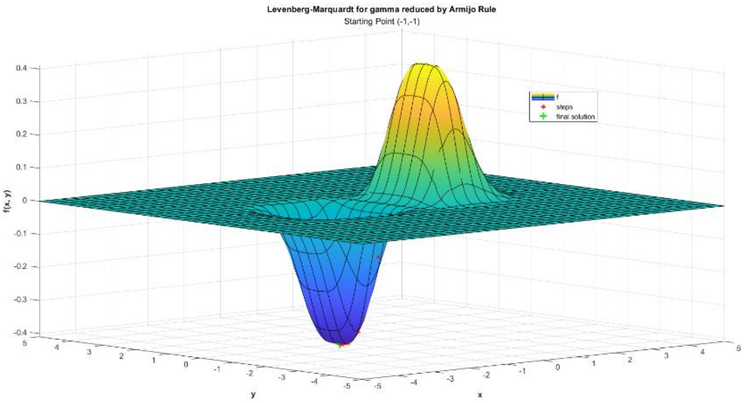


Σχήμα 4.4 : Levenberg Marquardt method, γ_k= minimizes minimizes f(xk + gkdk)

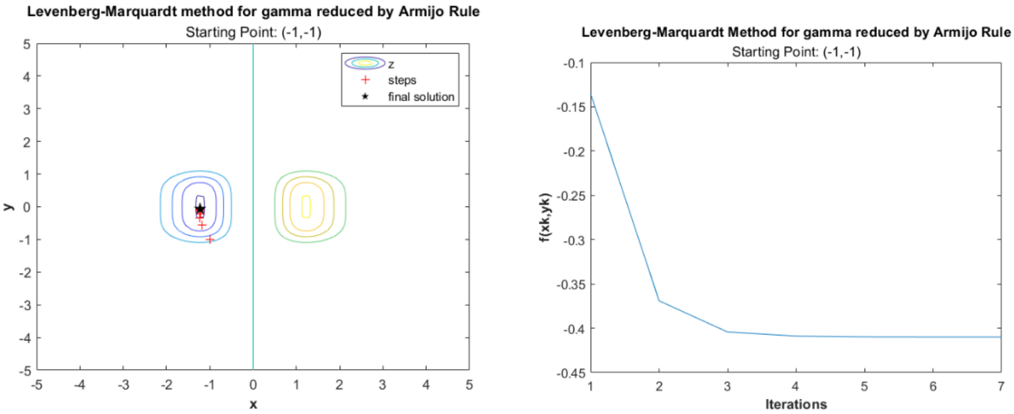


Σχήματα 4.5 , 4.6 : Levenberg Marquardt method, γ_k= minimizes minimizes f(xk + gkdk)

Γ) $γ_κ$ υπολογισμένο με τη μέθοδο armijo



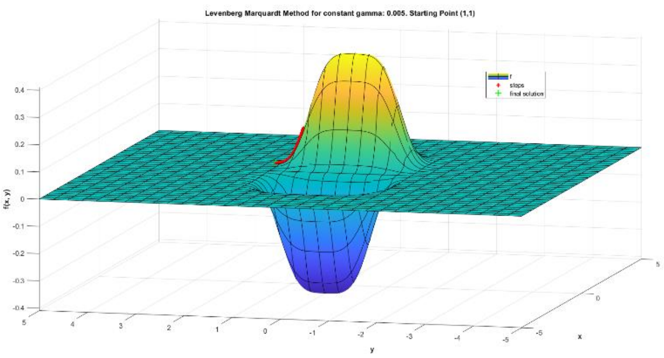
Σχήμα 4.7 : Levenberg Marquardt method, $γ_k$ reduced by Armijo's Rule



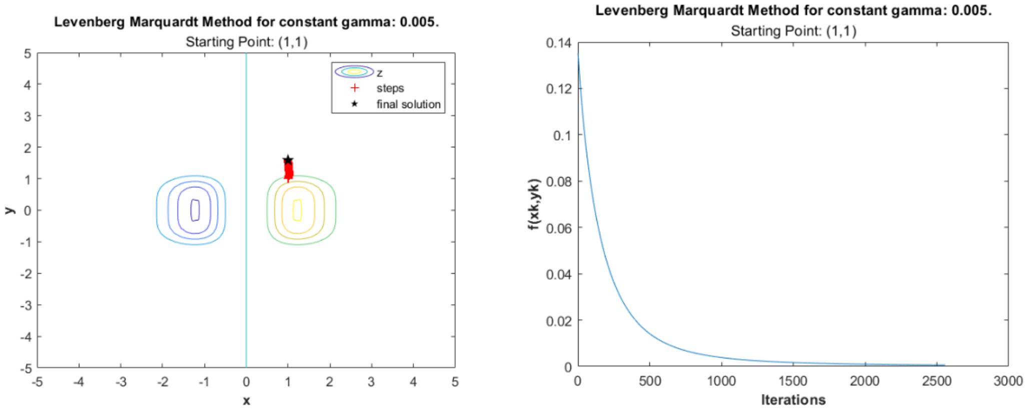
Σχήματα 4.8 , 4.9 : Levenberg Marquardt method, $γ_k$ reduced by Armijo's Rule

- **Αρχικό σημείο [1, 1]**
  Για το σημείο (1,1) και η μέθοδος Levenberg Marquardt βρίσκει τοπικό ελάχιστο της συνάρτησης και εγκλωβίζεται σε αυτό.

Α) Σταθερό $γ_κ$

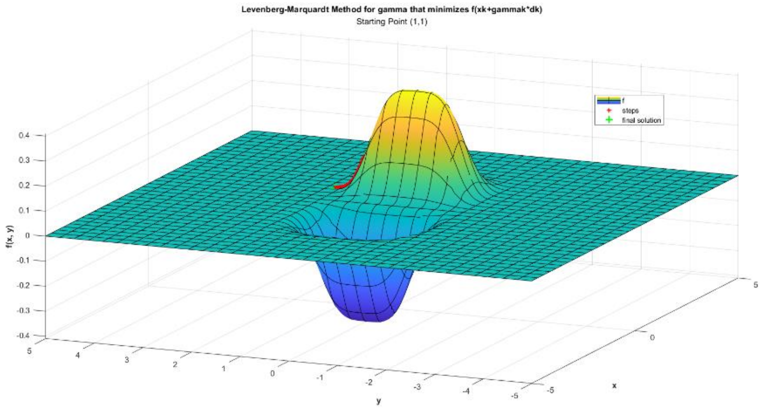


Σχήμα 4.10 : : Levenberg Marquardt Method, $γ_k$ constant

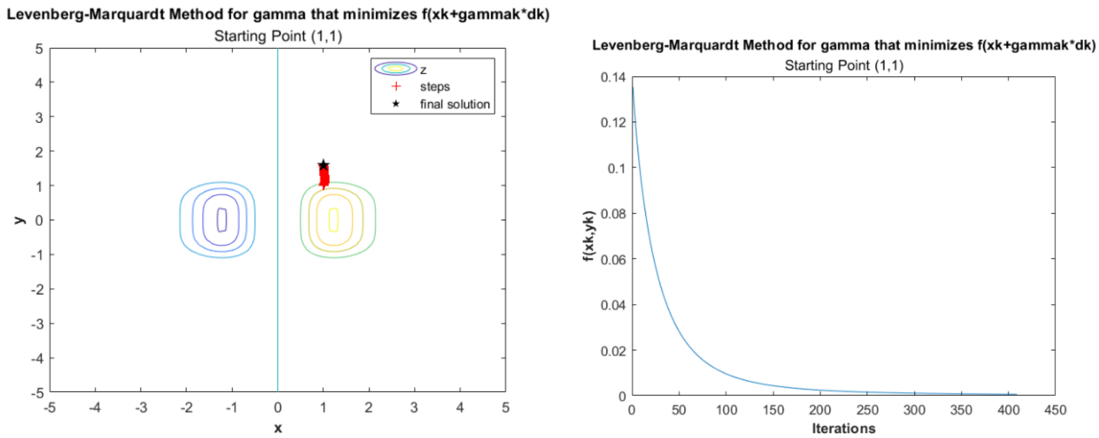


Σχήματα 4.11 , 4.12 : : Levenberg Marquardt Method, γ_k constant

B) γ_κ που ελαχιστοποιεί τη συνάρτηση $f(\chi\kappa+\gamma\kappa*dk)$

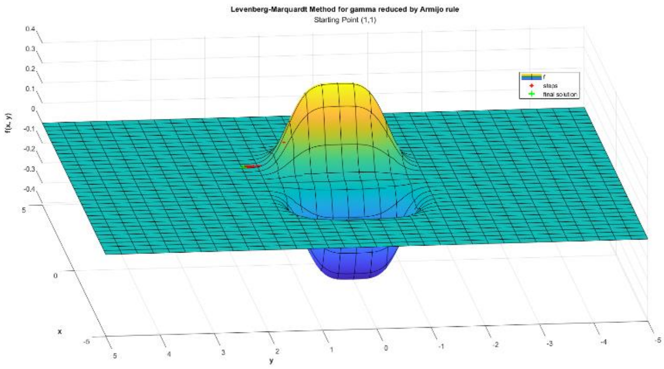


Σχήμα 4.13 : Levenberg Marquardt method, γ_k= minimizes minimizes f(xk + gkdk)

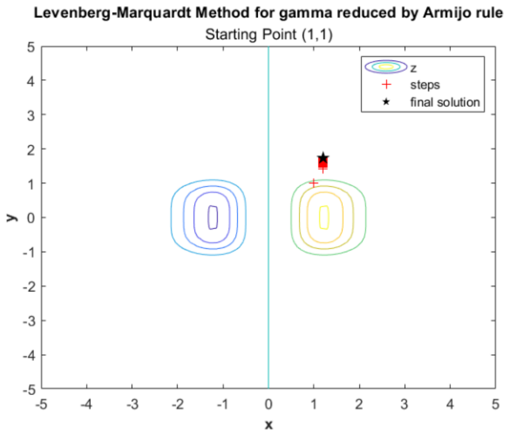


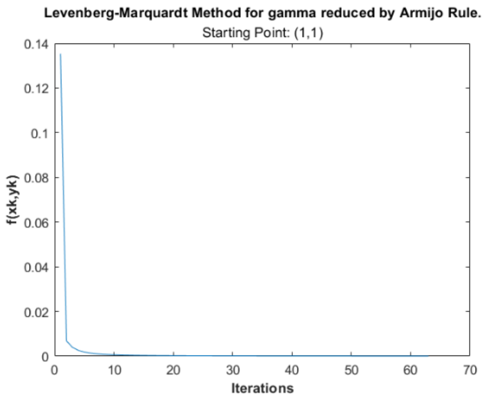
Σχήματα 4.14, 4.15 : Levenberg Marquardt method, γ_k= minimizes minimizes f(xk + gkdk)

Γ) γ_k υπολογισμένο με τη μέθοδο armijo



Σχήμα 4.16 : Levenberg Marquardt method, γ_k reduced by Armijo's Rule





Σχήματα 4.17 , 4.18 : Levenberg Marquardt method, γ_k reduced by Armijo's Rule

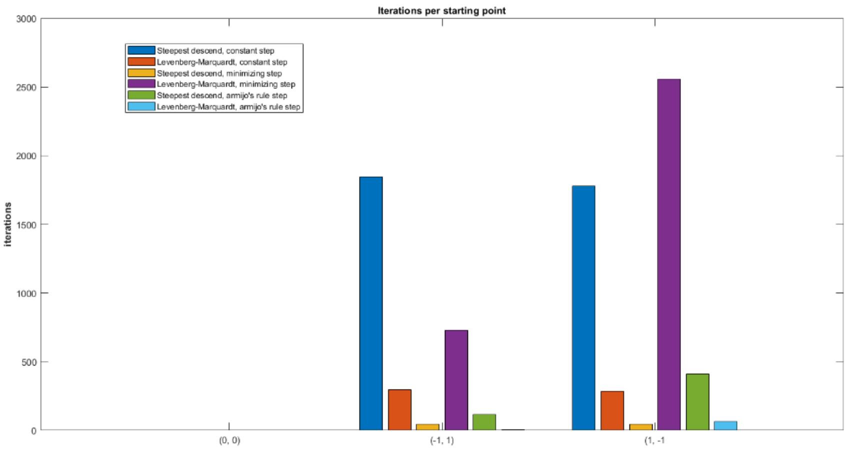
5 . Συμπεράσματα

Συγκρίνοντας τις μεθόδους επιλογής του $\gamma\kappa$ παρατηρείται ότι η μέθοδος με σταθερό γ_κ και γ_κ τέτοιο ώστε να ελαχιστοποιείται η $f(x_k+g_k d_k)$ χρειάζονταν περισσότερο αριθμό επαναλήψεων για την εύρεση του ελαχίστου ή τοπικού ελαχίστου σε σχέση με τη μέθοδο εύρεσης του γ_κ με βάση τον κανόνα του Armijo. Παρατηρείται ότι μέθοδοι υπολογισμού του κατάλληλου βήματος, όπως η ελαχιστοποίηση της συνάρτησης ή ο κανόνας του Armijo μπορούν να μειώσουν αισθητά τον αριθμό των επαναλήψεων που απαιτούνται για να συγκλίνει ο εκάστοτε αλγόριθμος .

Όσον αφορά τη σύγκριση των κύριων μεθόδων ελαχιστοποίησης, είναι προφανές ότι αυτός με το περισσότερο υπολογιστικό φόρτο είναι αυτός της μέγιστης καθόδου καθώς απαιτεί τα περισσότερα βήματα (συνήθως) για την εύρεση του ελαχίστου. Ο αλγόριθμος του Newton, αν και υπολογιστικά γρηγορότερος δεν λειτουργεί σωστά όταν ο εσσιανός της f δεν είναι θετικά ορισμένος. Από την άλλη πλευρά, ο αλγόριθμος Levenberg Marquardt συνδυάζει τους δύο αυτούς αλγορίθμους και καταλήγει στο ελάχιστο σε ικανοποιητικό αριθμό βημάτων, ξεκινώντας στην αρχή κάνοντας κάθετα βήματα όπως ο αλγόριθμος της μέγιστης καθόδου και έπειτα ακολουθεί τον αλγόριθμο του Newton, καθώς πλησιάζει στο ελάχιστο.

Παράλληλα, εξάγεται το συμπέρασμα ότι η επιλογή του σημείου έναρξης του αλγορίθμου παίζει καθοριστικό ρόλο στην εύρεση του ολικού ελαχίστου της συνάρτησης καθώς οι μέθοδοι εγκλωβίζονται εύκολα σε τοπικά ακρότατα.

Στο παρακάτω διάγραμμα απεικονίζονται τα iterations για τις μεθόδους ελαχιστοποίησης Steepest Descent και Levenberg Marquardt (για την Newton δεν έχει νόημα από την στιγμή που δεν λειτουργεί σωστά) για κάθε αρχικό σημείο και καθεμιά από τις μεθόδους του βήματος γ_κ. Το διάγραμμα επιβεβαιώνει τις παραπάνω παρατηρήσεις .



Σχήμα 5.1