<u>ΤΕΧΝΙΚΕΣ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ</u>

2η Εργαστηριακή Άσκηση

Ονοματεπώνυμο: Φωτιάδης Αλέξανδρος

ΑΕΜ:10392

E-mail: afotiadis@ece.auth.gr

Σε αυτό το report παρουσιάζονται οι λύσεις για την 2η εργαστηριακή άσκηση του μαθήματος, τα ζητούμενα διαγράμματα και ο σχολιασμός και τα συμπεράσματα που εξάγονται από τα αποτελέσματα.

Αντικείμενο της εργασίας είναι η ελαχιστοποίηση μιας δοσμένης κυρτής συνάρτησης $f : \mathbb{R}^n \to \mathbb{R}$ χωρίς περιορισμούς. Χρησιμοποιήθηκαν οι εξής 3 αλγόριθμοι:

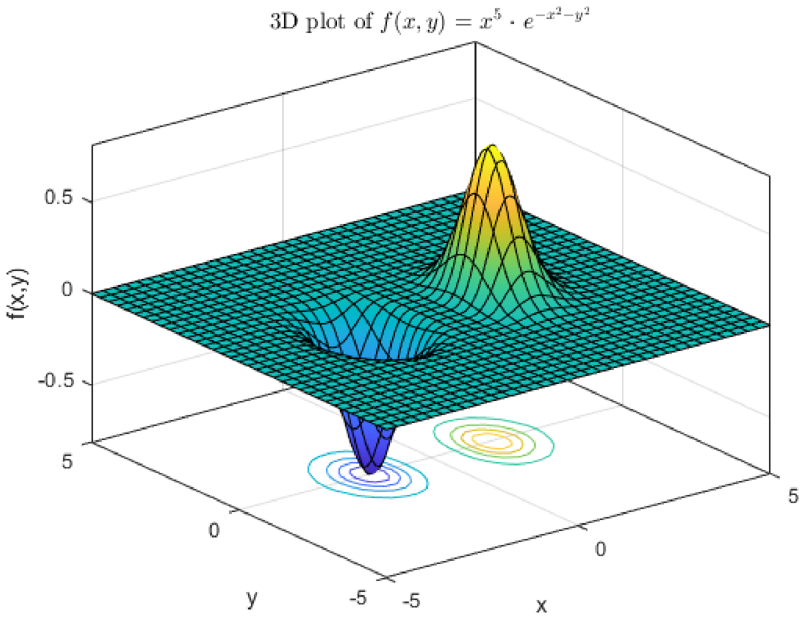
- Μέθοδος Μέγιστης Καθόδου (Steepest Descent)
- Μέθοδος Newton
- Μέθοδος Levenberg-Marquardt

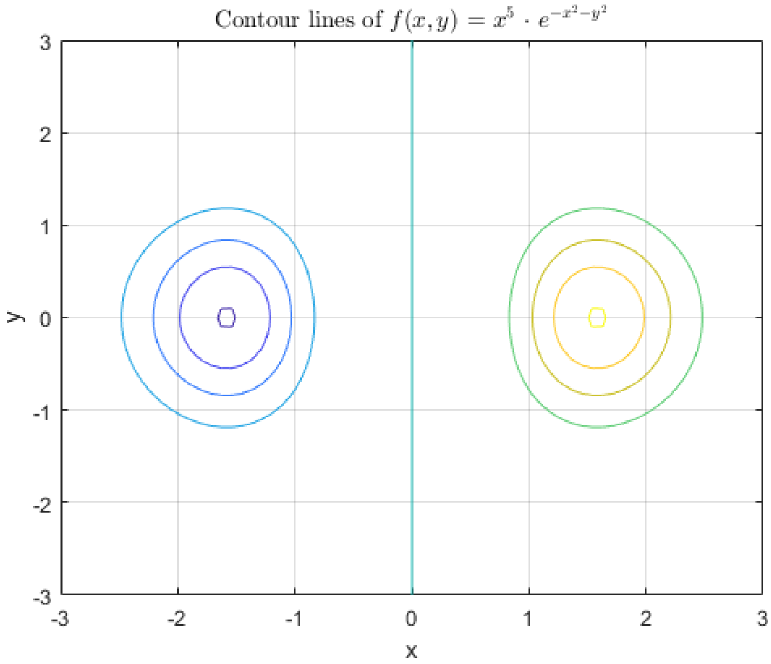
Οι 3 αυτές μέθοδοι ανήκουν στις μεθόδους κλίσης και βασική ιδιότητά τους είναι η χρήση του διανύσματος κλίσης $\nabla f(x)$, στην ισοβαρή καμπύλη της f. Η αντικειμενική συνάρτηση που θα μελετήσουμε παρακάτω είναι η

$$f(x,y) = x^5 e^{-x^2-y^2}$$

<u>ΘΕΜΑ 1ο:</u>

Στο πρώτο θέμα μας ζητήθηκε να σχεδιάσουμε την f ώστε μέσω της οπτικοποίησής της να έχουμε μια ιδέα για το τι να περιμένουμε από τις μεθόδους. Με το αρχείο plotting_the_func.m παράγουμε τόσο την μορφή της f όσο και τις ισουψείς καμπύλες της.



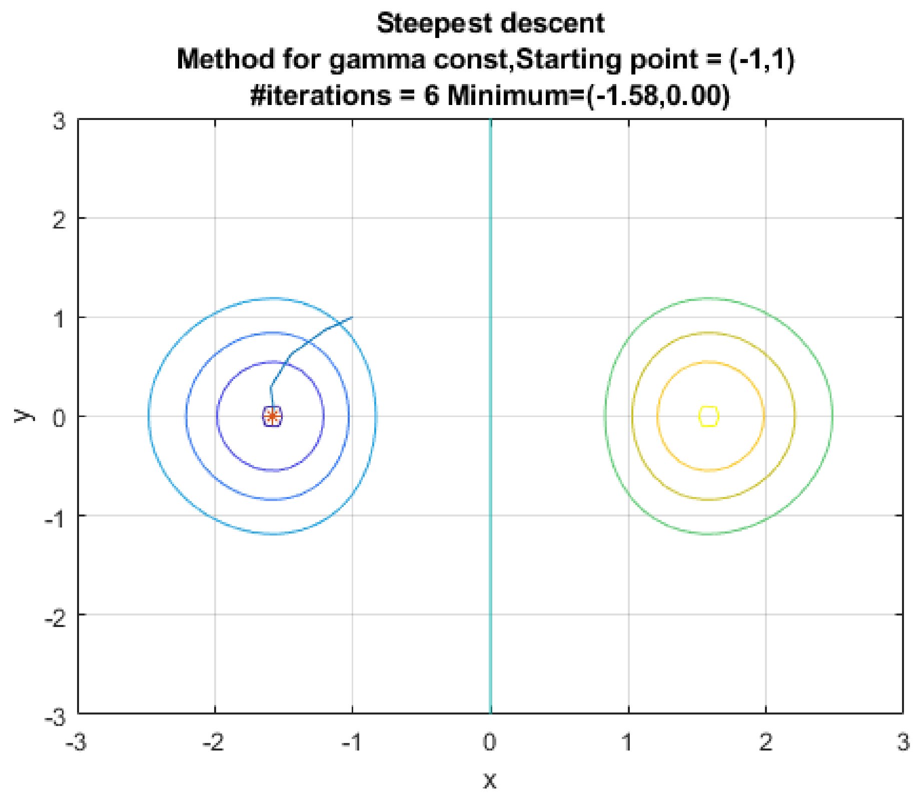


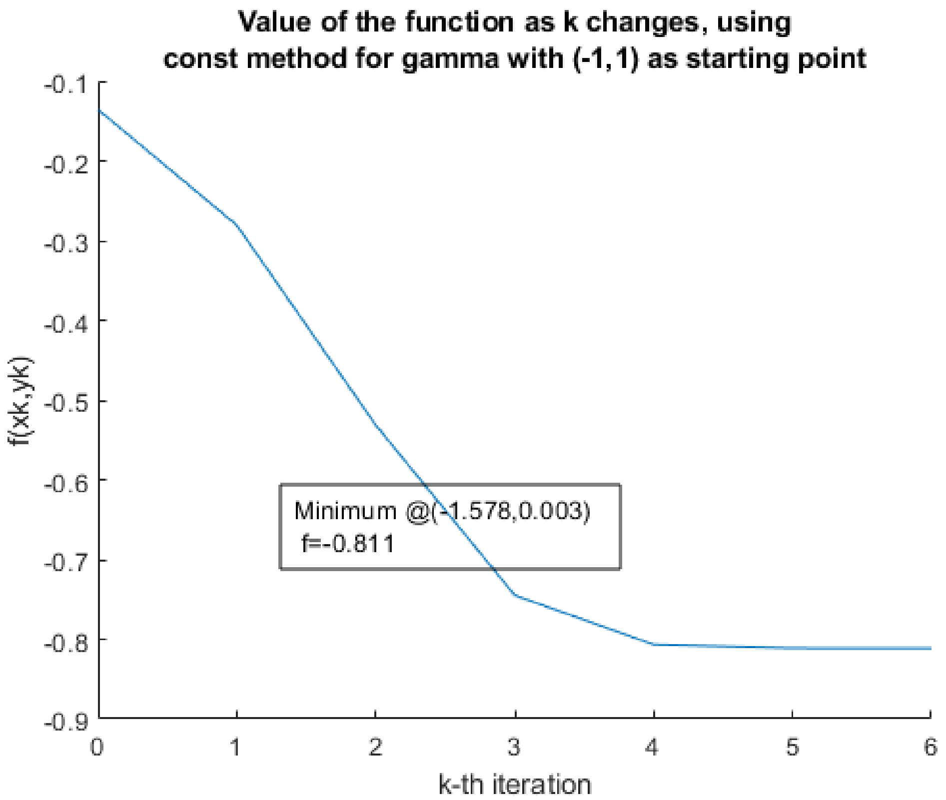
Από το διάγραμμα της f στον τρισδιάστατο χώρο παρατηρούμε πως η ελάχιστη τιμή εμφανίζεται στα αρνητικά x και κοντά στην γειτονιά του 0 στον άξονα των y. Αντίστοιχα παρουσιάζει μέγιστο σε θετικά x στην γειτονιά του 0 των y. Τέλος παρατηρούμε και ένα τοπικό ελάχιστο στο σημείο (0,0) όπου προφανώς η τιμή της αντικειμενικής συνάρτησης μηδενίζεται.

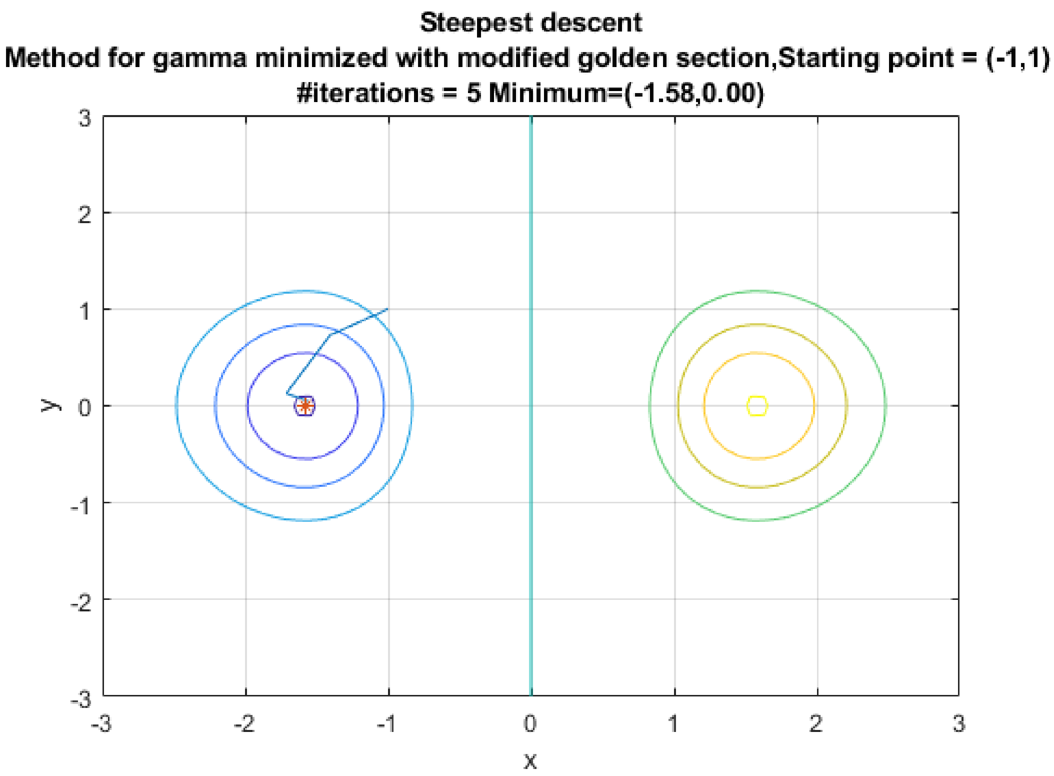
ΘΕΜΑ 2^{o}:

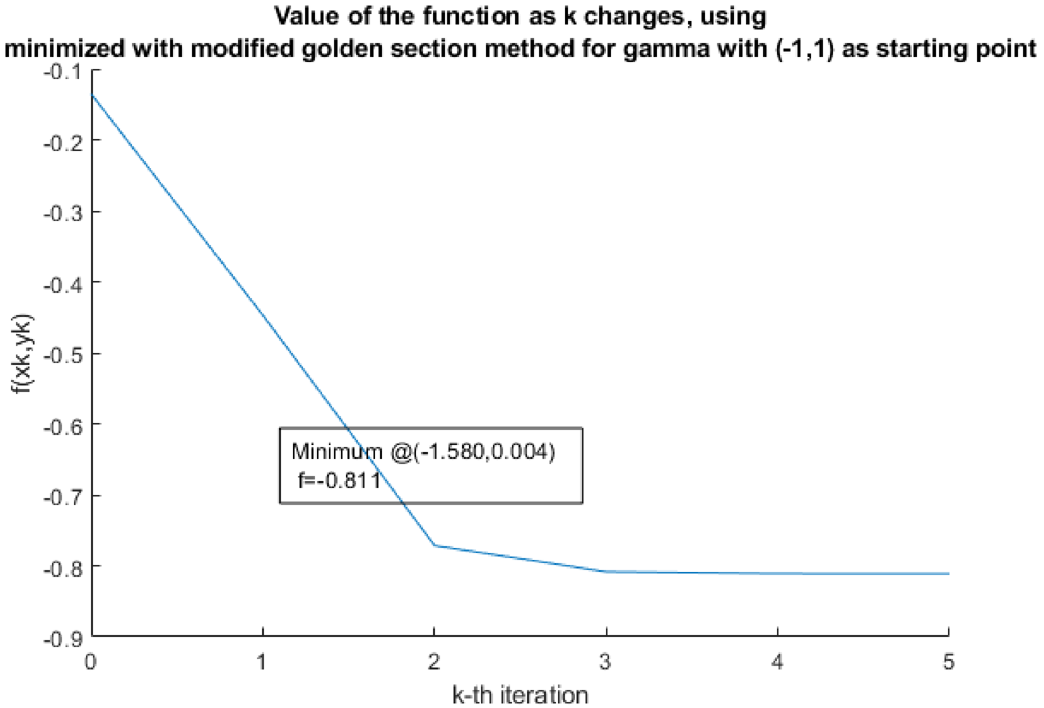
Στο 2^{o} θέμα μας ζητείται να ελαχιστοποιήσουμε την f με την χρήση της μεθόδου Μέγιστης Καθόδου. Επιπλέον μας ζητείται να επιλέξουμε το βήμα γ_k με τρεις διαφορετικούς τρόπους: (i) σταθερό (της επιλογής μας), (ii) τέτοιο ώστε να ελαχιστοποιεί την συνάρτηση $f(x_k + \gamma_k d_k)$, (iii) βάσει του κανόνα Armijo. Τέλος, ζητείται να υλοποιήσουμε την μέθοδο (και κάθε επιλεγμένο γ_k) 3 φορές έχοντας κάθε φορά διαφορετικό σημείο εκκίνησης (x_0,y_0) τα οποία είναι τα εξής: (i) (0,0), (ii) (-1,1), (iii) (1,-1). Τα διαγράμματα που προκύπτουν από το Αρχείο steepest_descent_method.m μας βοηθούν να εντοπίσουμε τις διαφορές στα αποτελέσματα. Διαφορές που μπορεί να προκύπτουν είτε λόγω του διαφορετικού σημείου εκκίνησης είτε λόγω του διαφορετικού βήματος επανάληψης. Ακόμη, οπτικοποιεί τη σύγκλιση της συνάρτησης στο ελάχιστο με κάθε επανάληψη.

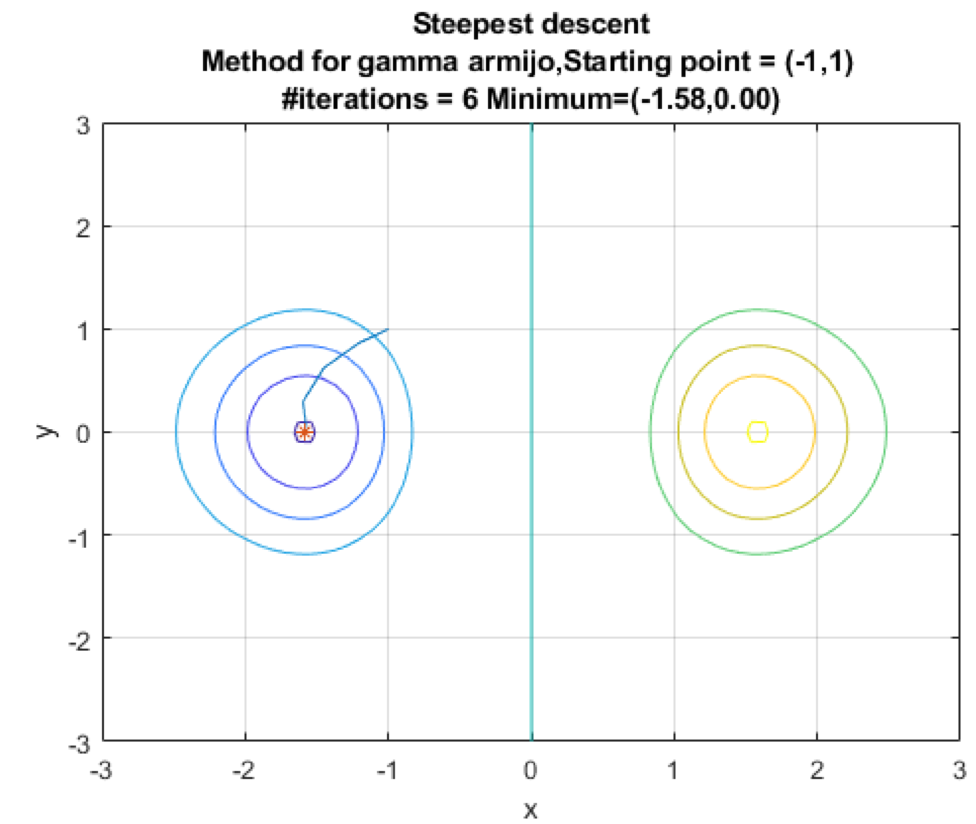
Αρχικά, εφόσον το (0,0) αποτελεί τοπικό ακρότατο και συγκεκριμένα ελάχιστο της αντικειμενικής συνάρτησης, η μέθοδος εγκλωβίζεται σε αυτό και δεν μπορεί να εκτελέσει τις επαναλήψεις που απαιτούνται και συνεπώς τα διαγράμματα που παράγονται δεν έχουν κανένα ενδιαφέρον και δεν παρατίθενται στην παρούσα αναφορά (παρόλα αυτά το αρχείο Matlab τα παράγει κανονικά). Εάν θέσουμε σημείο εκκίνησης το (-1,1) παρατηρούμε ότι ο αλγόριθμος φτάνει πολύ κοντά στο ελάχιστο σημείο για δική μας επιλογή ε=0.01 και για τις 3 διαφορετικές μεθόδους επιλογής του βήματος επανάληψης γ_k.

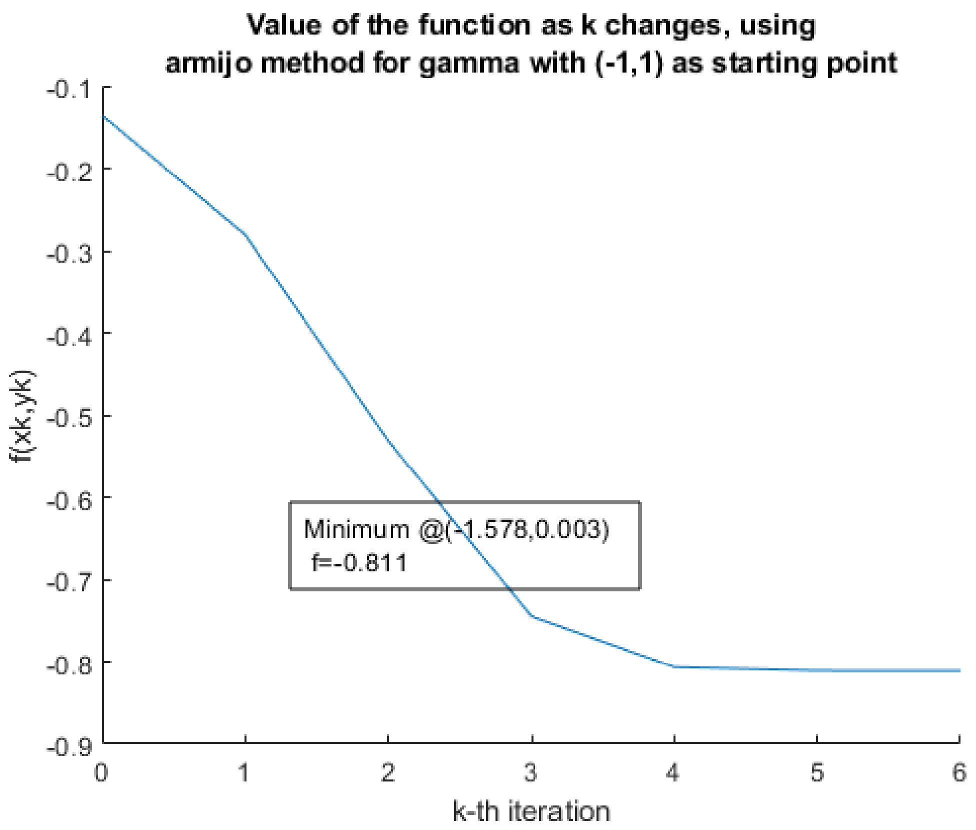

Steepest descent
Method for gamma const,Starting point = (-1,1)
#iterations = 6 Minimum=(-1.58,0.00)
3
2
1
0
-1
-2
-3
y
-3
-2
-1
0
1
2
3
x


Value of the function as k changes, using
const method for gamma with (-1,1) as starting point
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.7
-0.8
-0.9
f(xk,yk)
Minimum @(-1.578,0.003)
f=-0.811
0
1
2
3
4
5
6
k-th iteration


Steepest descent
Method for gamma minimized with modified golden section,Starting point = (-1,1)
#iterations = 5 Minimum=(-1.58,0.00)
3
2
1
0
-1
-2
-3
y
-3
-2
-1
0
1
2
3
x

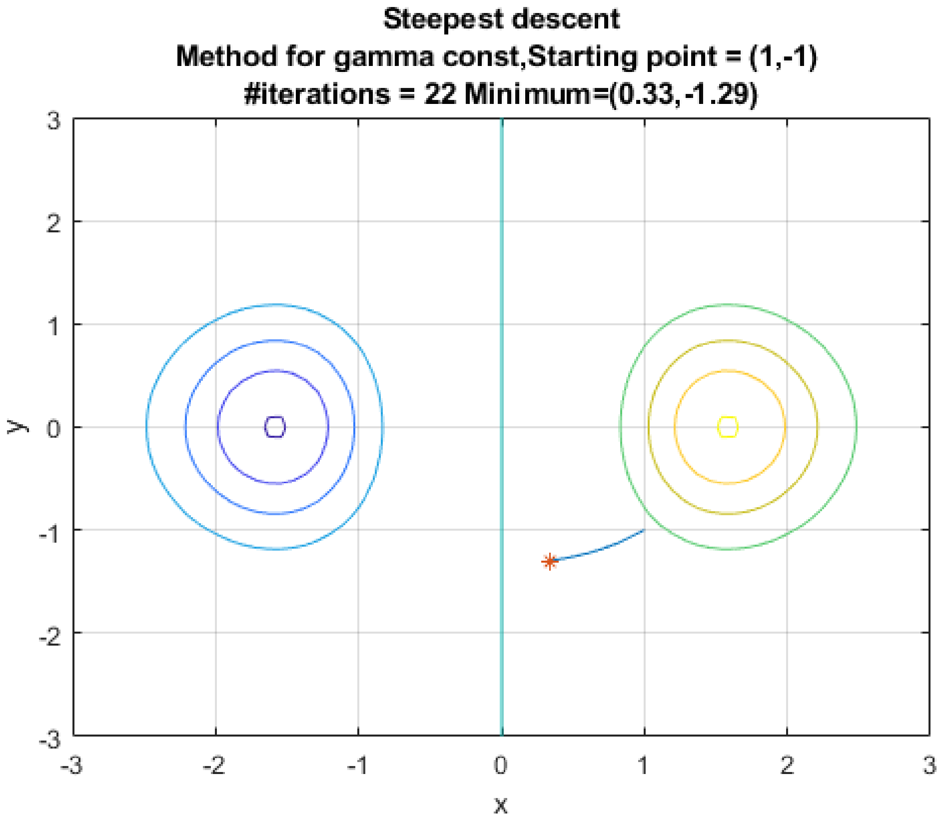

Value of the function as k changes, using
minimized with modified golden section method for gamma with (-1,1) as starting point
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.7
-0.8
-0.9
f(xk,yk)
Minimum @(-1.580,0.004)
f=-0.811
0
1
2
3
4
5
k-th iteration

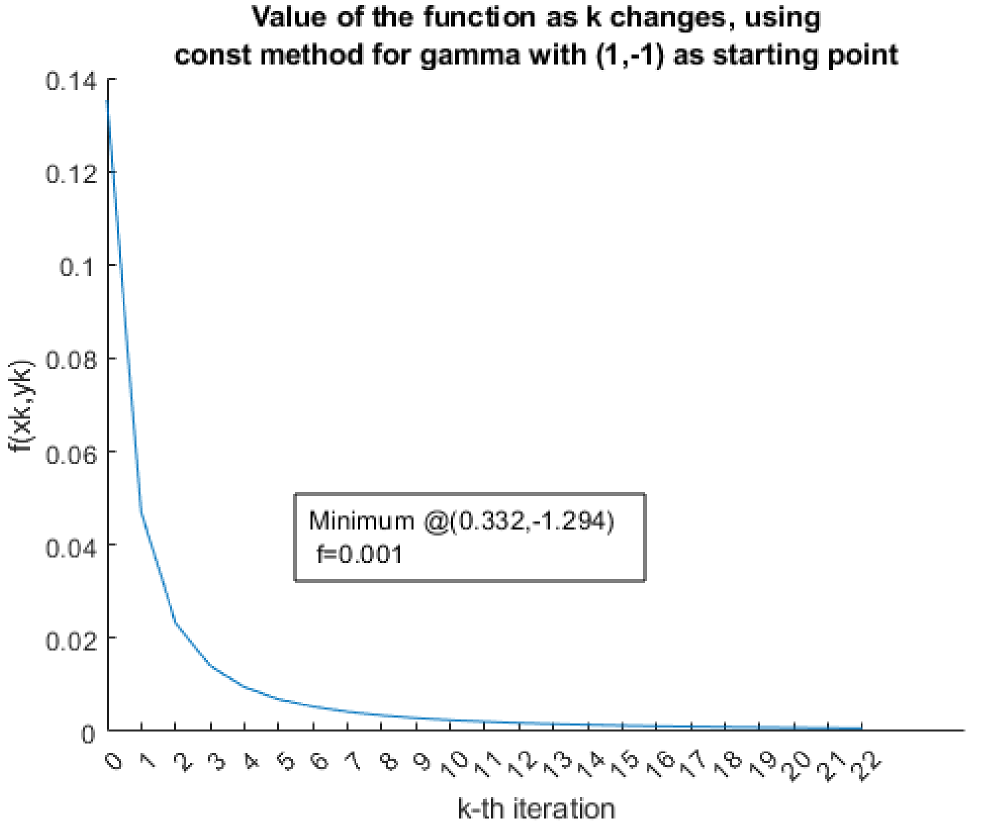




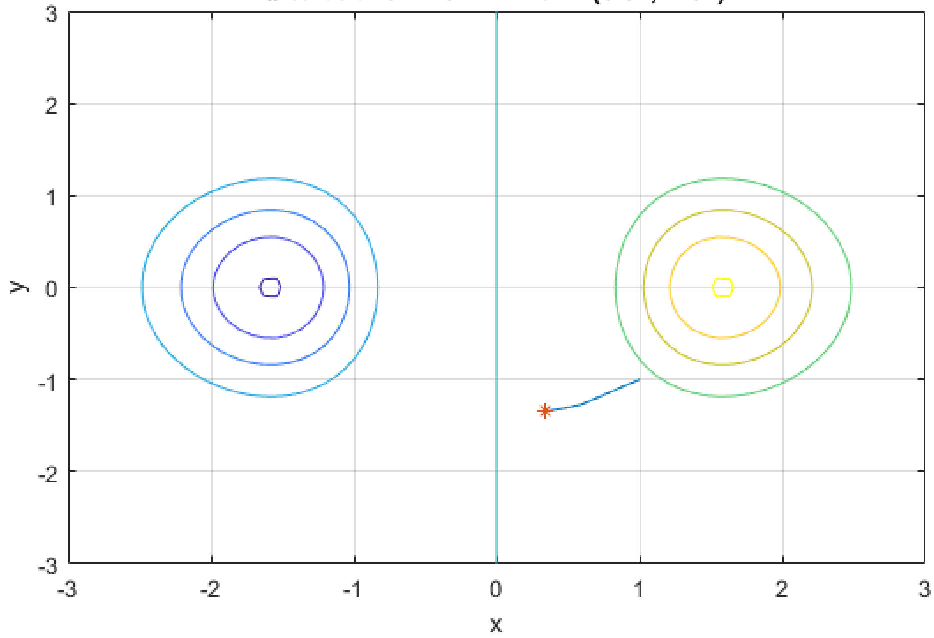
Από τα παραπάνω διαγράμματα γίνεται αντιληπτό πως έστω και για μια επανάληψη λιγότερο, η μέθοδος (ii) όπου επιλέγεται το βέλτιστο βήμα συγκλίνει γρηγορότερα από τις άλλες 2. Παρόλα αυτά και οι 3 μπορούν να χαρακτηριστούν αρκετά αποτελεσματικές. Κάπου εδώ πρέπει να διευκρινιστεί πως για τη μέθοδο (i) το βήμα αυθαίρετα ορίστηκε ως $γ_k$=0.5 ενώ για την μέθοδο (ii) και την εύρεση του βέλτιστου βήματος χρησιμοποιήθηκε ο αλγόριθμος του χρυσού τομέα από την προηγούμενη εργασία με μια μικρή τροποποίηση καθώς στο τέλος υπολογίζεται το βέλτιστο γ ως ο μέσος όρος των άκρων του τελικού διαστήματος (οποιαδήποτε από τις 4 μεθόδους της προηγούμενης εργασίας με την κατάλληλη τροποποίηση θα μπορούσε να χρησιμοποιηθεί. Εδώ επιλέχτηκε ο χρυσός τομέας κάπως αυθαίρετα).Ακόμα, στη μέθοδο Armijo επιλέχτηκαν οι αρχικές συνθήκες α=0.002 και β=0.25 (αυτά ισχύουν και για τα θέματα 3,4).

Τέλος αν πάρουμε ως αρχικό σημείο το (1,-1) ο αλγόριθμος δεν αποδίδει τόσο καλά καθώς όπως φαίνεται και παρακάτω στα διαγράμματα το σημείο εκκίνησης βρίσκεται στις ισουψείς καμπύλες που περιέχουν το μέγιστο της αντικειμενικής συνάρτησης, μια ατυχής επιλογή για να επιτύχουμε σωστή σύγκλιση. Αν συνυπολογίσουμε ότι μεταξύ των ισουψών καμπυλών του μεγίστου και του ελαχίστου παρεμβάλλεται και το τοπικό ελάχιστο (0,0) καταλαβαίνουμε ότι ο αλγόριθμος είναι πιο πιθανό να προσπαθήσει να πλησιάσει το (0,0) καθώς κατά τη λειτουργία του αντιλαμβάνεται πως αυτό είναι το επιθυμητό ελάχιστο που θέλει να φτάσει.

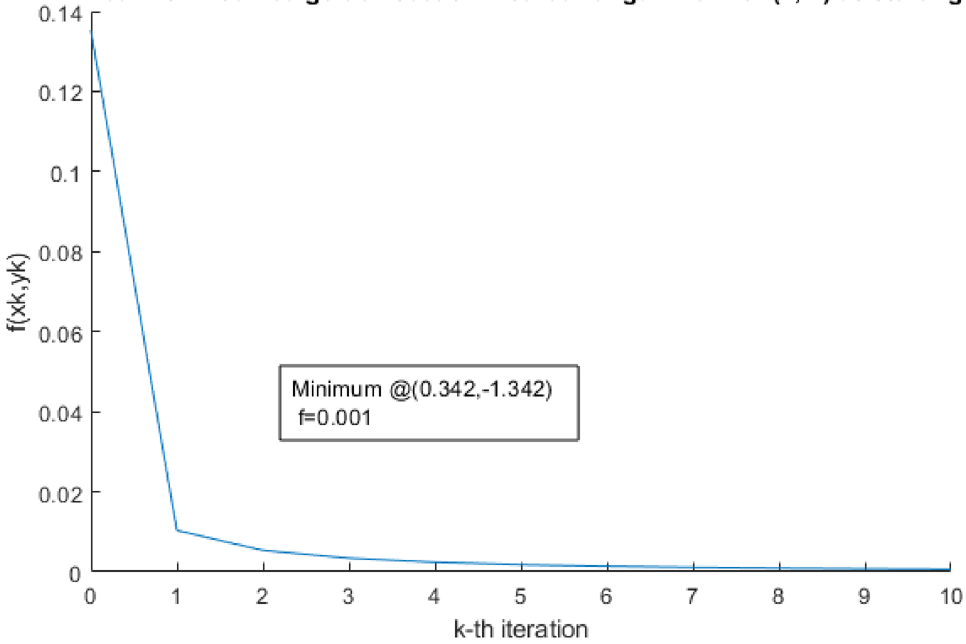

Steepest descent
Method for gamma const,Starting point = (1,-1)
#iterations = 22 Minimum=(0.33,-1.29)
y
x

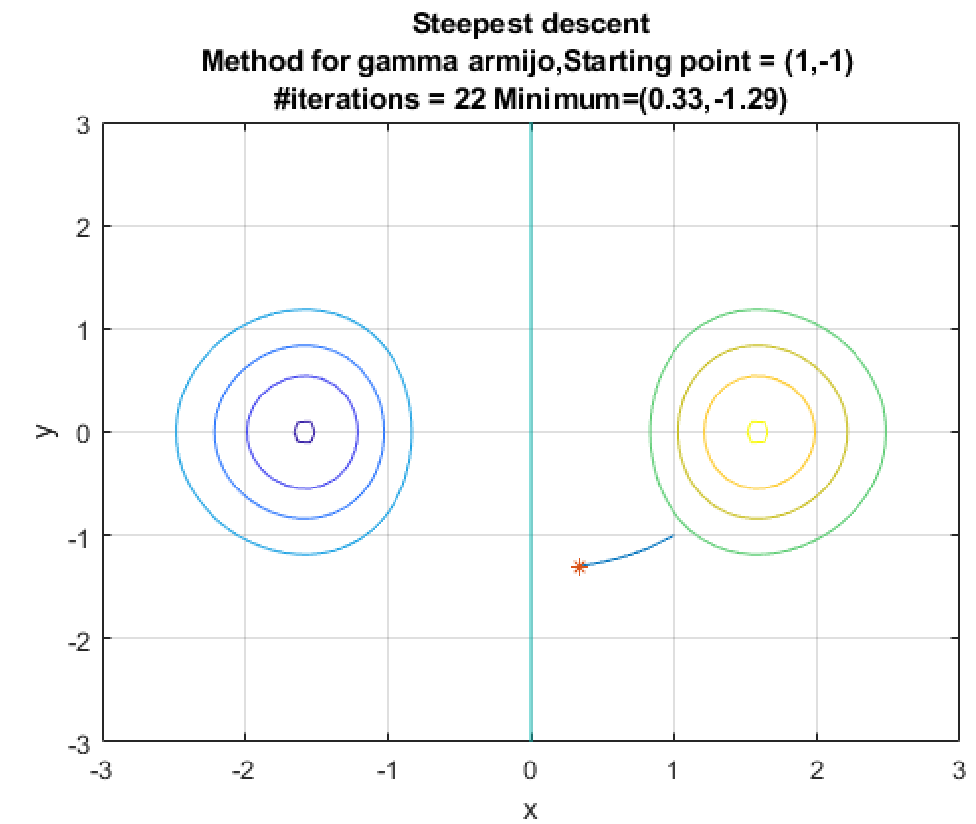

Value of the function as k changes, using
const method for gamma with (1,-1) as starting point
f(xk,yk)
Minimum @(0.332,-1.294)
f=0.001
k-th iteration

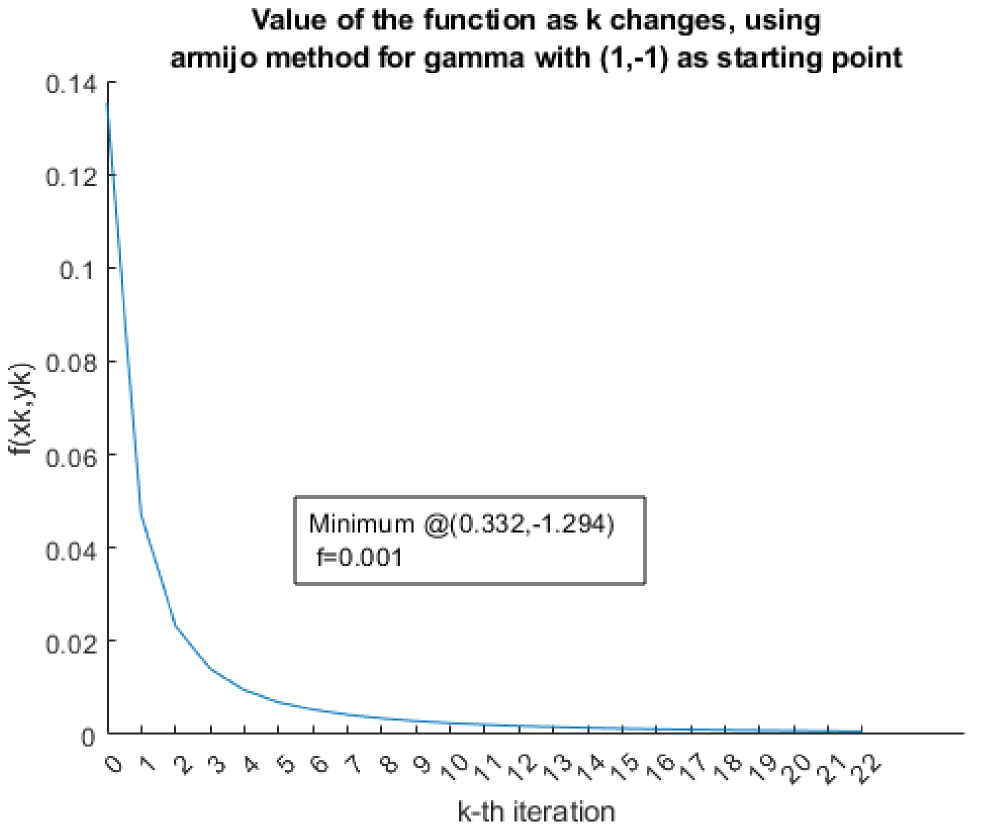
Steepest descent
Method for gamma minimized with modified golden section,Starting point = (1,-1)
#iterations = 10 Minimum=(0.34,-1.34)



Value of the function as k changes, using
minimized with modified golden section method for gamma with (1,-1) as starting point







Όπως αναμενόταν λοιπόν η μέθοδος της μέγιστης καθόδου προσπαθεί να πλησιάσει τιμές για τις οποίες η αντικειμενική συνάρτηση τείνει να μηδενιστεί. Πάντως και σε αυτό το παράδειγμα φαίνεται πως η μέθοδος (ii) υπερτερεί των άλλων όσον αφορά την αποτελεσματικότητα καθώς εδώ για να πλησιάσει την ίδια τιμή της αντικειμενικής συνάρτησης χρειάζεται λιγότερες των μισών επαναλήψεων σε σχέση με τις μεθόδους (i) και (iii).

<u>ΘΕΜΑ 3ο:</u>

Στο 2ο θέμα μας ζητείται να ελαχιστοποιήσουμε την f με την χρήση της μεθόδου Newton. Σε αντίθεση με την μέγιστη κάθοδο, αυτή η μέθοδος αναζητεί ελάχιστο στην κατεύθυνση $d_k = -[\nabla^2 f(x_k)]^{-1} \cdot \nabla f(x_k)$. Επίσης, πρέπει να τηρείται η προϋπόθεση ότι ο Εσσιανός πίνακας της f είναι θετικά ορισμένος.

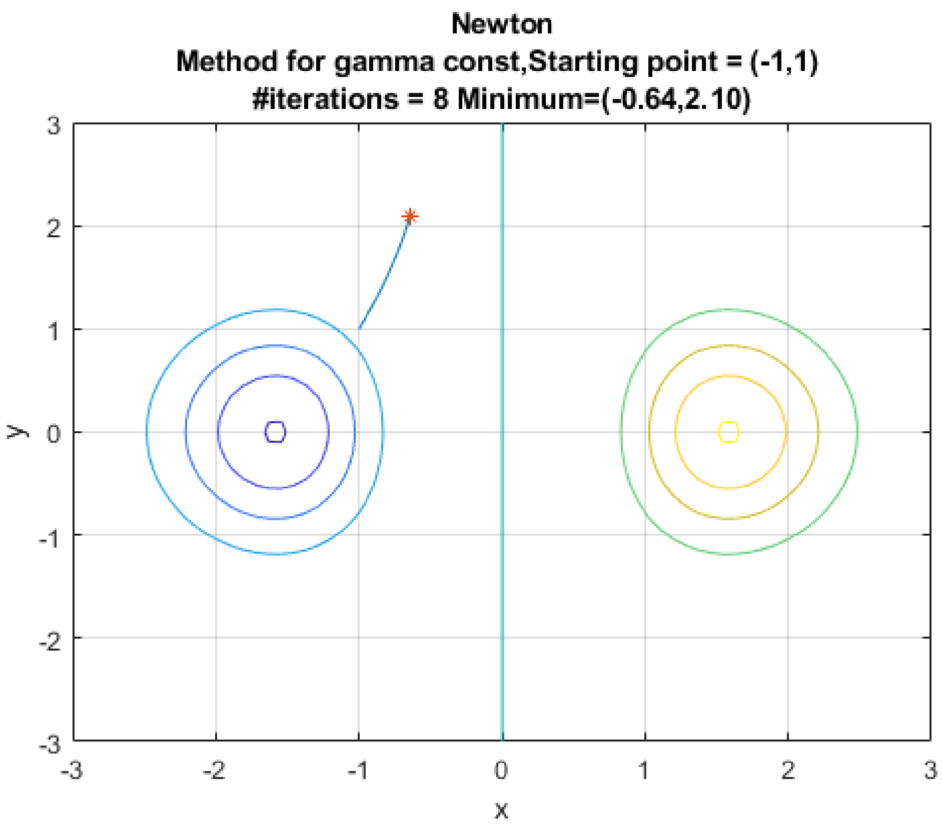
Τα διαγράμματα που προκύπτουν από το Αρχείο newton_method.m μας βοηθούν να εντοπίσουμε τις διαφορές στα αποτελέσματα.

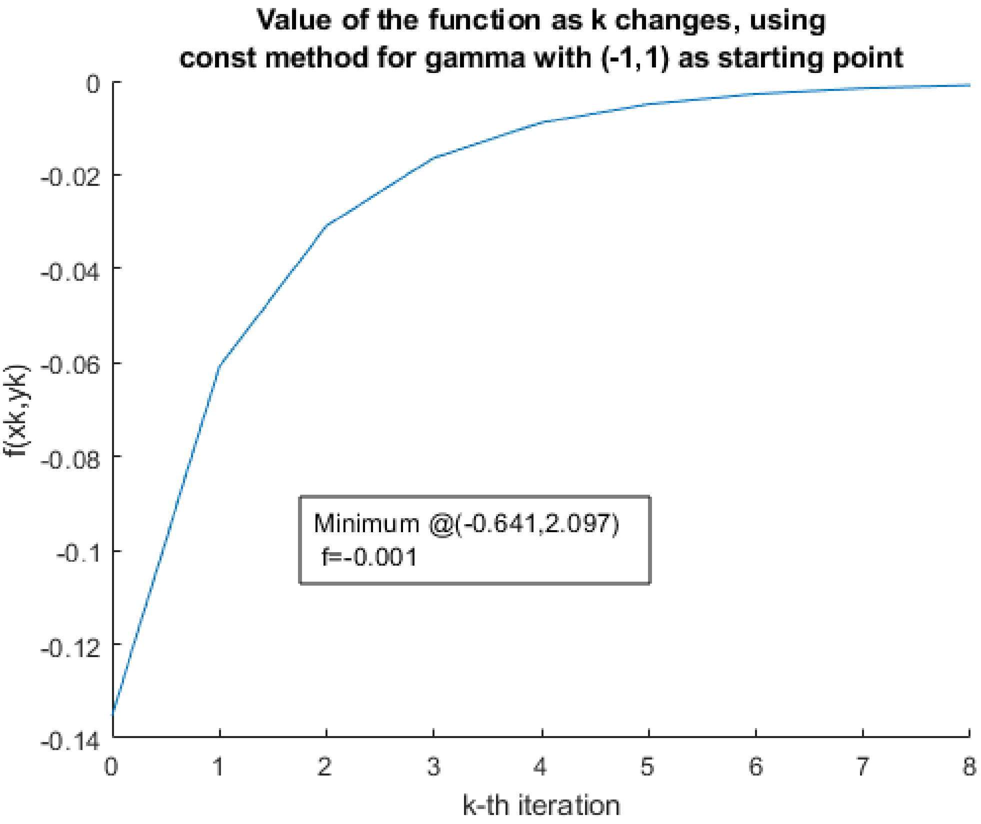
Αρχικά, εφόσον το (0,0) αποτελεί τοπικό ακρότατο και συγκεκριμένα ελάχιστο της αντικειμενικής συνάρτησης, η μέθοδος εγκλωβίζεται σε αυτό και δεν μπορεί να εκτελέσει τις επαναλήψεις που απαιτούνται και συνεπώς τα διαγράμματα που παράγονται δεν έχουν κανένα ενδιαφέρον και δεν παρατίθενται στην παρούσα αναφορά (παρόλα αυτά το αρχείο Matlab τα παράγει κανονικά).

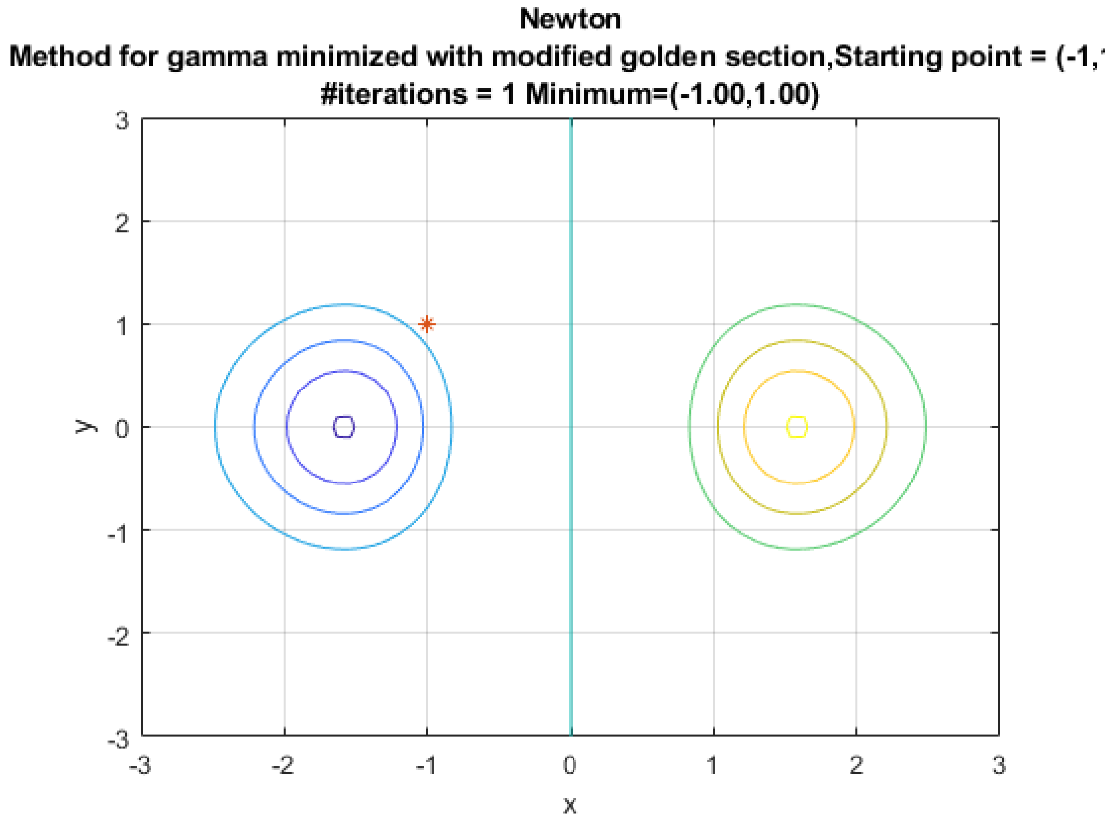
Πριν συνεχίσουμε την ανάλυση αξίζει να αναφερθεί ότι ο Εσσιανός πίνακας της συνάρτησης δεν είναι Θ.Ο. και αυτό εύκολα διαπιστώνεται τρέχοντας τις κατάλληλες εντολές στο Matlab και για αυτό περιμένουμε να μην είναι αποδοτικός ο αλγόριθμος και έχουμε θέσει στις μεθόδος (ii) και (iii) ένα break που θα σταματάει την διαδικασία αναζήτησης όταν $f(x_k) > f(x_{k-1})$.

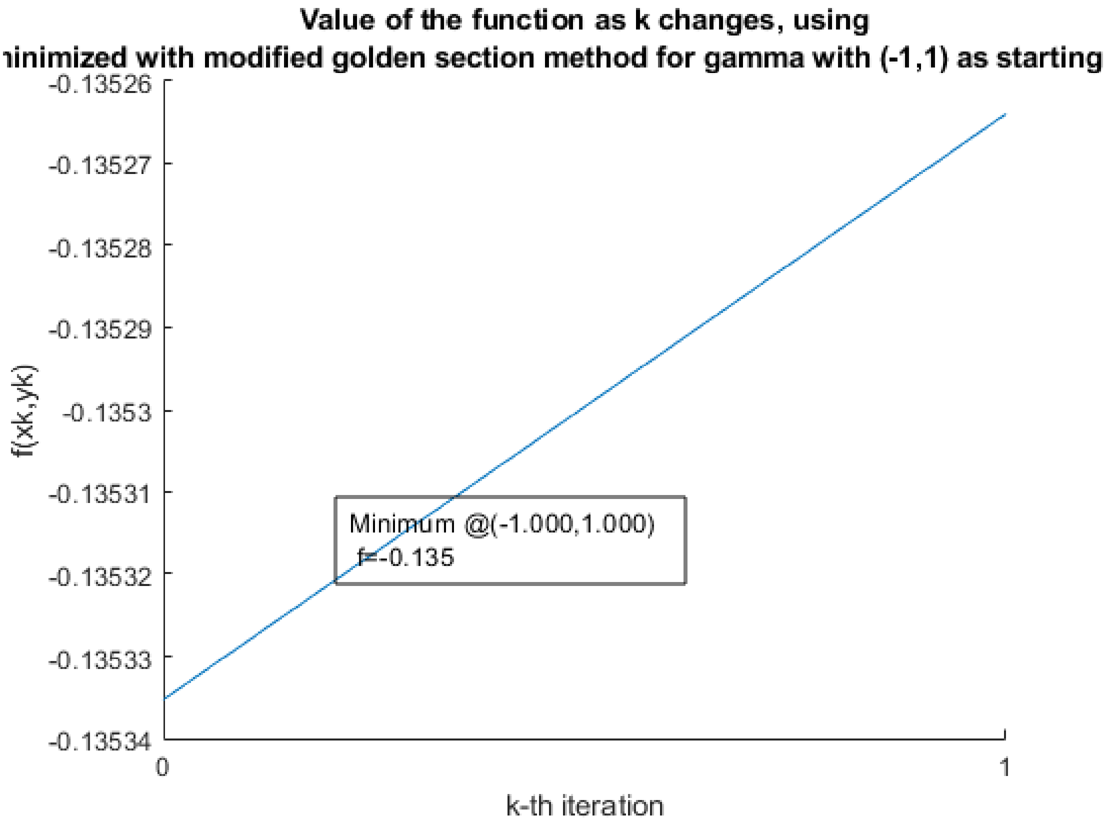
```
>> h=hessian(f(x,y))

h =

[20*x^3*exp(- x^2 - y^2) - 22*x^5*exp(- x^2 - y^2) + 4*x^7*exp(- x^2 - y^2), 4*x^6*y*exp(- x^2 - y^2) - 10*x^4*y*exp(- x^2 - y^2)]
[                       4*x^6*y*exp(- x^2 - y^2) - 10*x^4*y*exp(- x^2 - y^2),  4*x^5*y^2*exp(- x^2 - y^2) - 2*x^5*exp(- x^2 - y^2)]

>> d=det(h)

d =

-4*exp(- 2*x^2 - 2*y^2)*(2*x^12 + 2*x^10*y^2 - 11*x^10 + 5*x^8*y^2 + 10*x^8)

>>
```

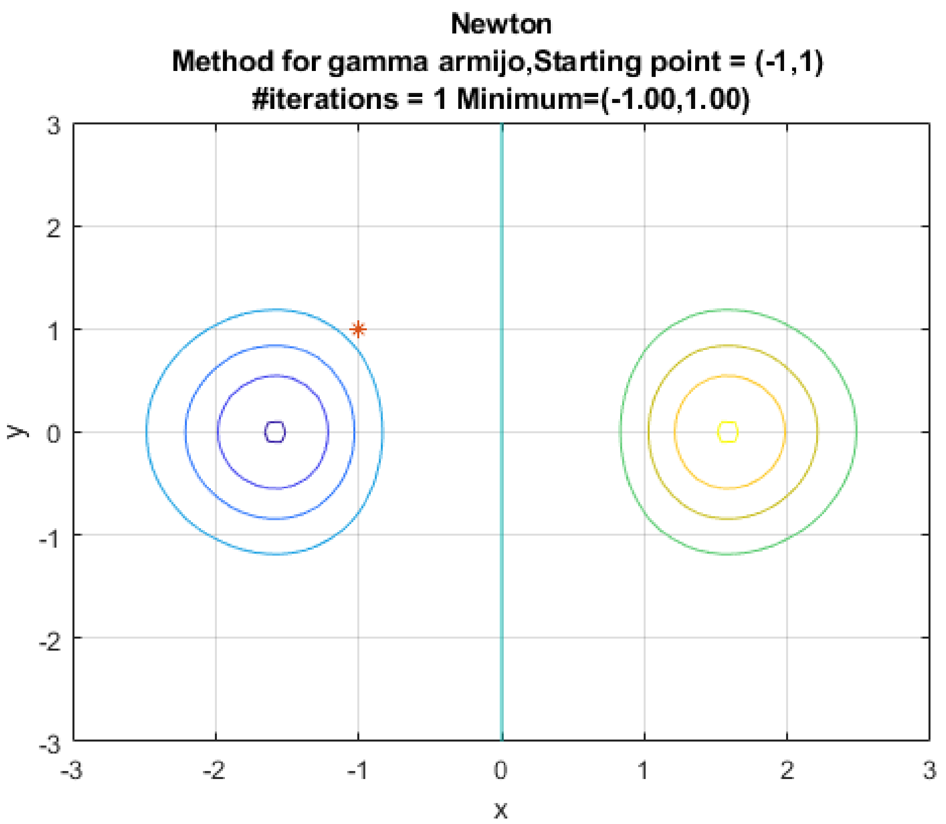
Εφόσον ο Εσσιανός δεν είναι Θ.Ο. παρακάτω θα δειχθούν όλα τα διαγράμματα μόνο για το σημείο (-1,1) ενώ θα δειχθούν και για τη μέθοδο Armijo τα διαγράμματα για το σημείο (1,-1). Επίσης αξίζει να σημειωθεί ότι για επιλογή βέλτιστου γ η μέθοδος τερματίζει σε ένα βήμα καθώς όπως γνωρίζουμε από την θεωρία, τηρουμένων όλων των προϋποθέσεων, με βέλτιστο γ η εύρεση ελαχίστου επιτυγχάνεται σε 1 βήμα με τη μέθοδο Newton.

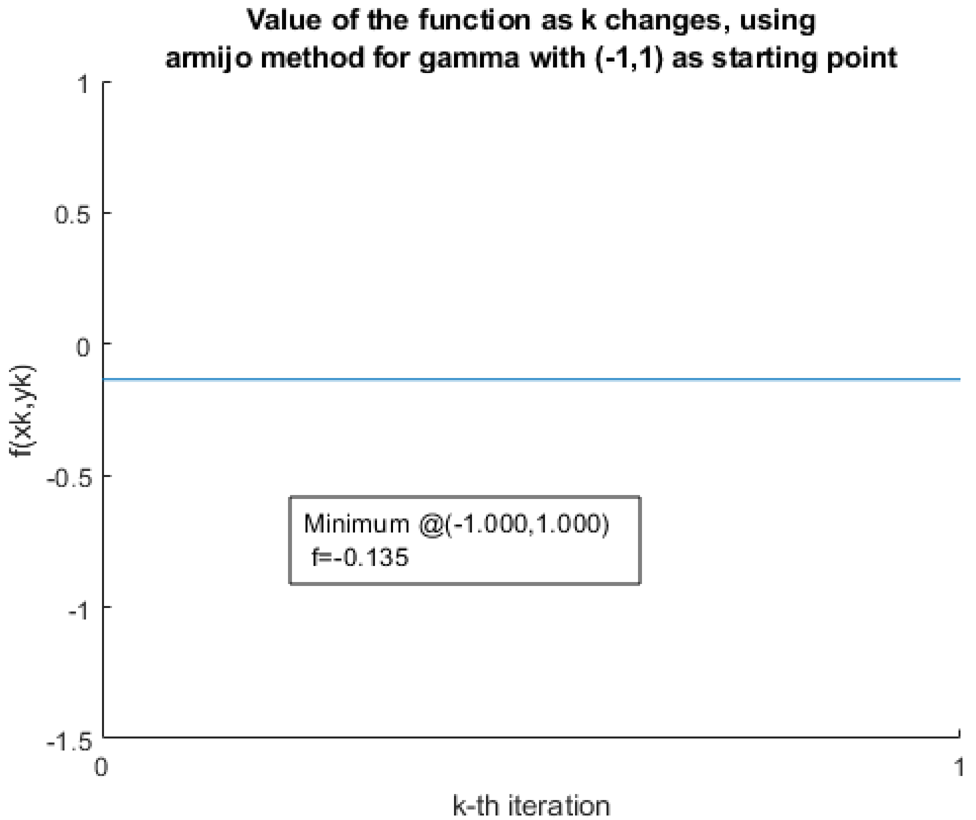

Newton
Method for gamma const,Starting point = (-1,1)
#iterations = 8 Minimum=(-0.64,2.10)
y
x

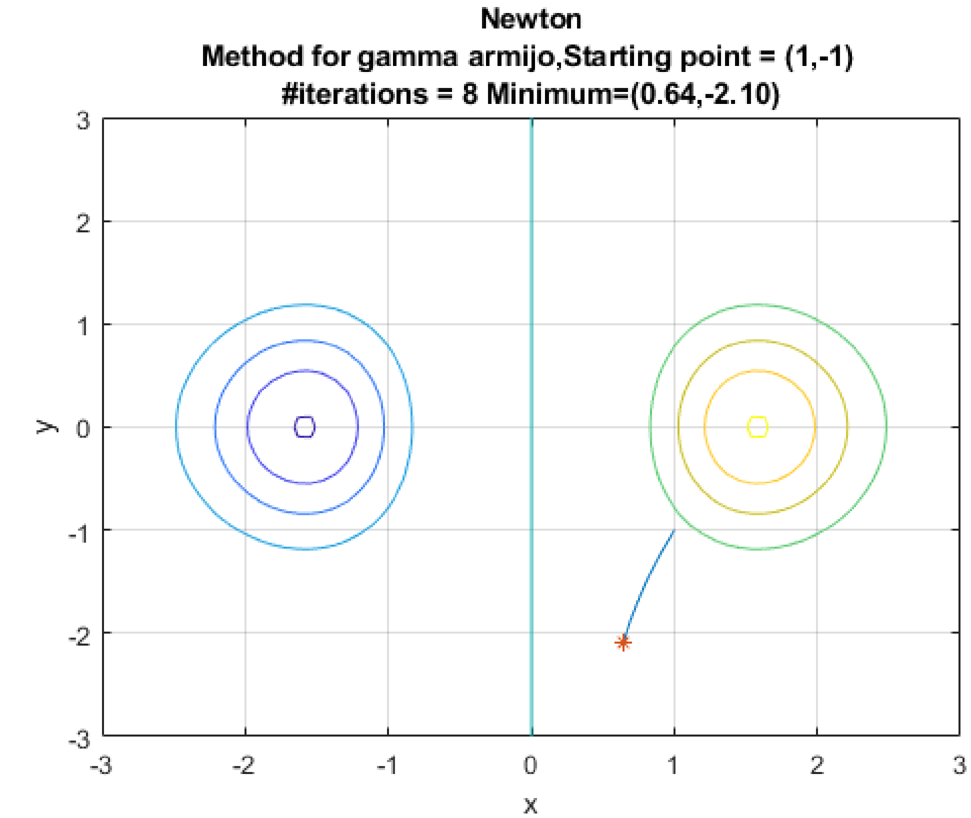

Value of the function as k changes, using
const method for gamma with (-1,1) as starting point
f(xk,yk)
k-th iteration
Minimum @(-0.641,2.097)
f=-0.001

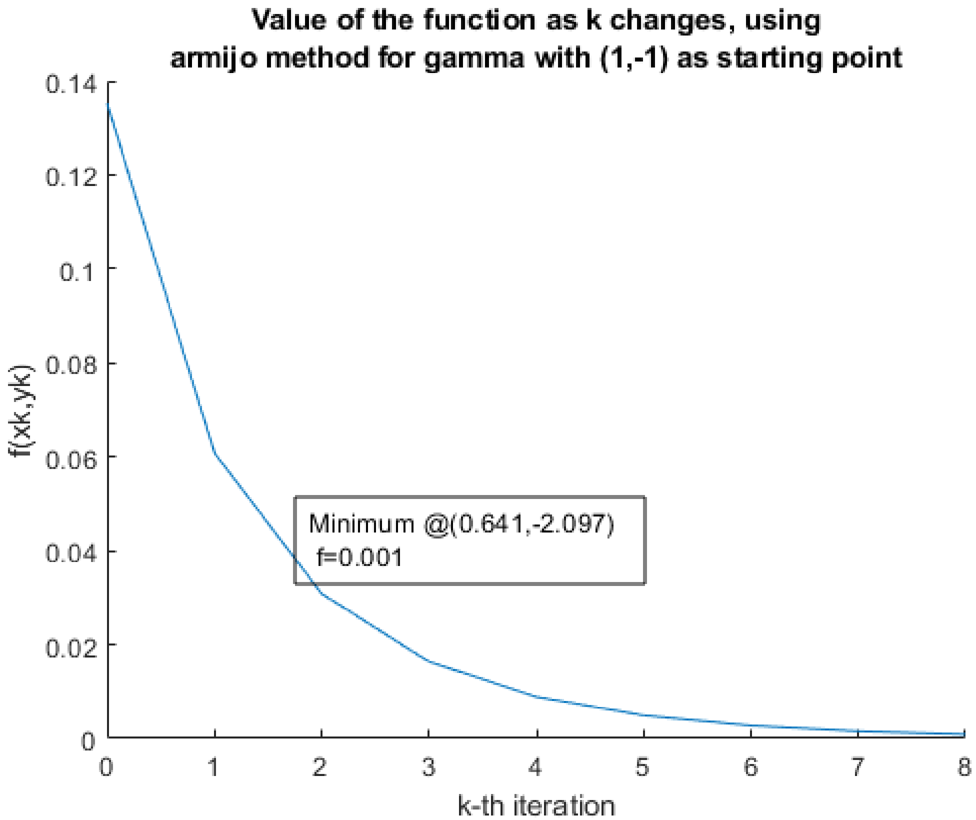

Newton
Method for gamma minimized with modified golden section,Starting point = (-1,1
#iterations = 1 Minimum=(-1.00,1.00)
3
2
1
0
-1
-2
-3
y
-3
-2
-1
0
1
2
3
x


Value of the function as k changes, using
ninimized with modified golden section method for gamma with (-1,1) as starting
-0.13526
-0.13527
-0.13528
-0.13529
-0.1353
-0.13531
-0.13532
-0.13533
-0.13534
f(xk,yk)
Minimum @(-1.000,1.000)
f=-0.135
0
1
k-th iteration


Newton
Method for gamma armijo,Starting point = (-1,1)
#iterations = 1 Minimum=(-1.00,1.00)
3
2
1
0
-1
-2
-3
y
-3
-2
-1
0
1
2
3
x


Value of the function as k changes, using
armijo method for gamma with (-1,1) as starting point
1
0.5
0
-0.5
-1
-1.5
f(xk,yk)
Minimum @(-1.000,1.000)
f=-0.135
0
1
k-th iteration

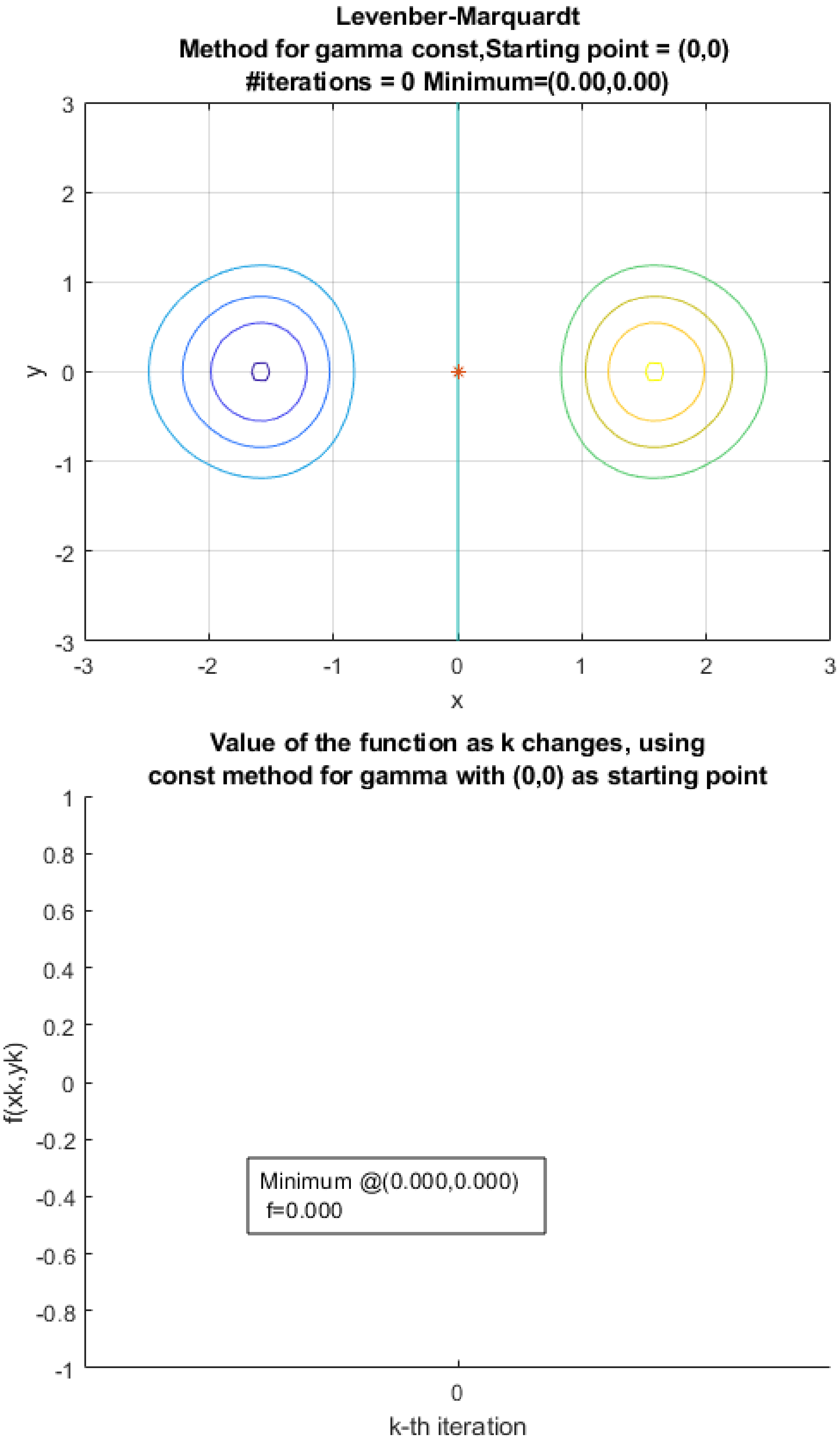

Newton
Method for gamma armijo,Starting point = (1,-1)
#iterations = 8 Minimum=(0.64,-2.10)
y
x
3
2
1
0
-1
-2
-3
-3
-2
-1
0
1
2
3


Value of the function as k changes, using
armijo method for gamma with (1,-1) as starting point
f(xk,yk)
k-th iteration
Minimum @(0.641,-2.097)
f=0.001
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0
1
2
3
4
5
6
7
8

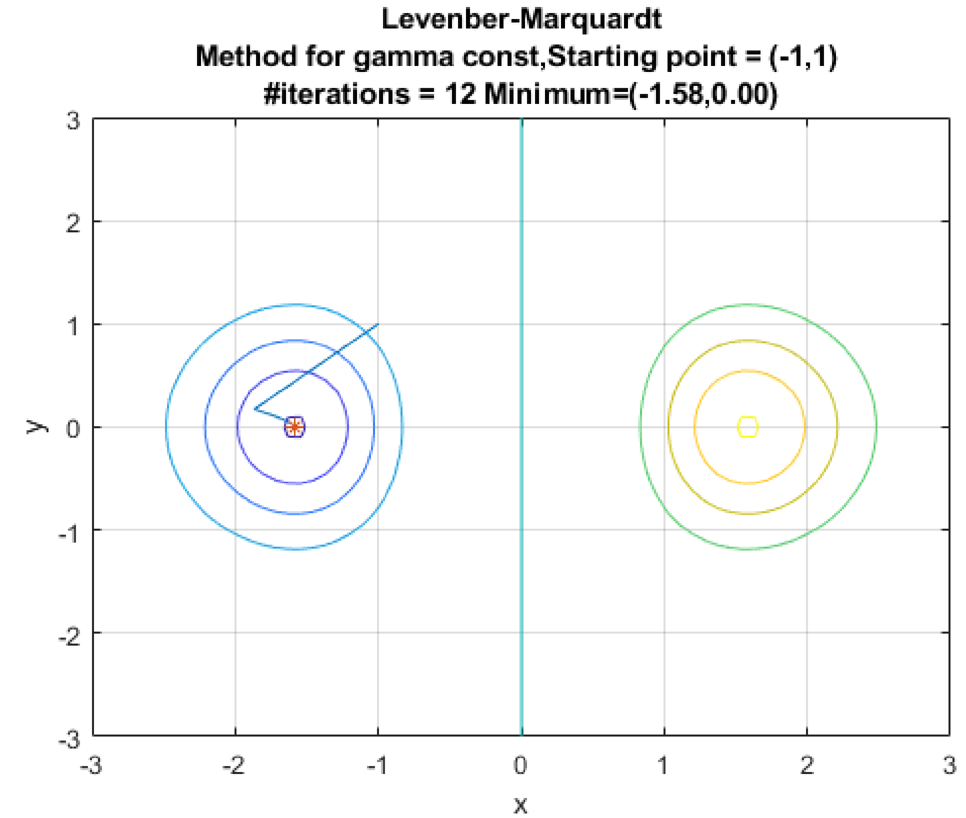
ΘΕΜΑ 4ο:

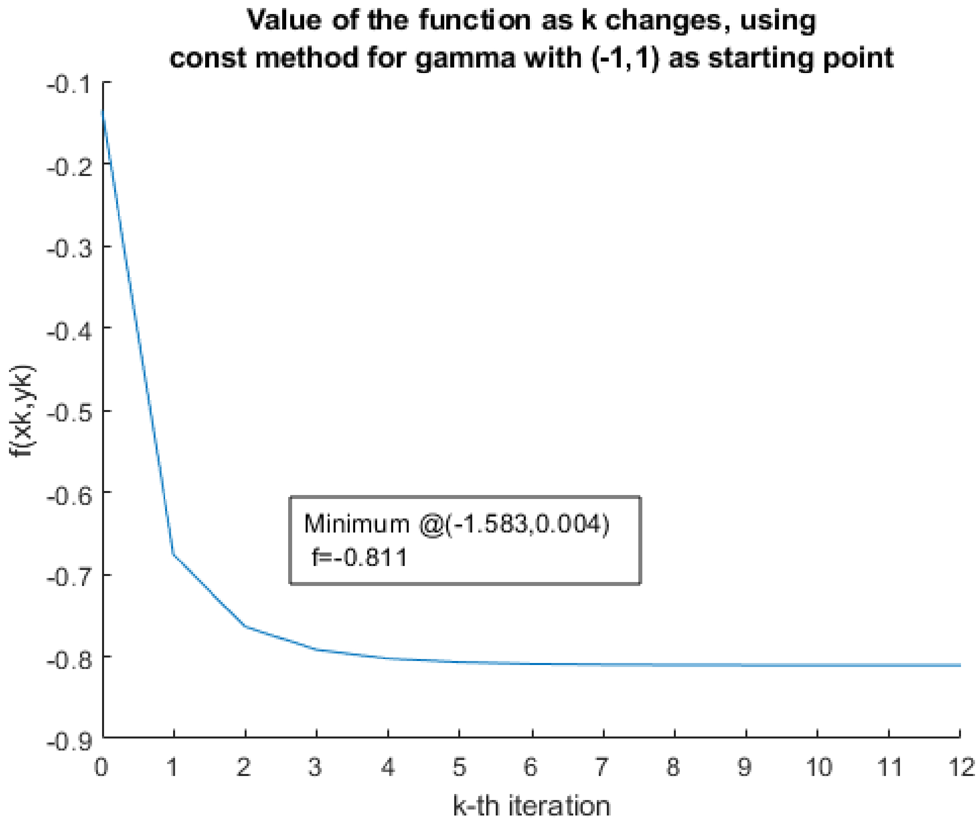
Σε αυτό το θέμα, υλοποιούμε τη μέθοδο Levenberg-Marquardt για να ελαχιστοποιήσουμε την f. Πρόκειται για μια τροποποίηση της μεθόδου Newton η οποία παρακάμπτει την προϋπόθεση του Θ.Ο. Εσσιανού πίνακα καθώς σε περίπτωση που αυτός δεν είναι Θ.Ο. επιλέγεται $d_k = -[\nabla^2 f(x_k) + \mu_k \cdot I]^{-1}$ όπου $\mu_k>0$ και με κατάλληλη επιλογή του ο πίνακας $\nabla^2 f(x_k) + \mu_k \cdot I$ θα είναι σίγουρα Θ.Ο.

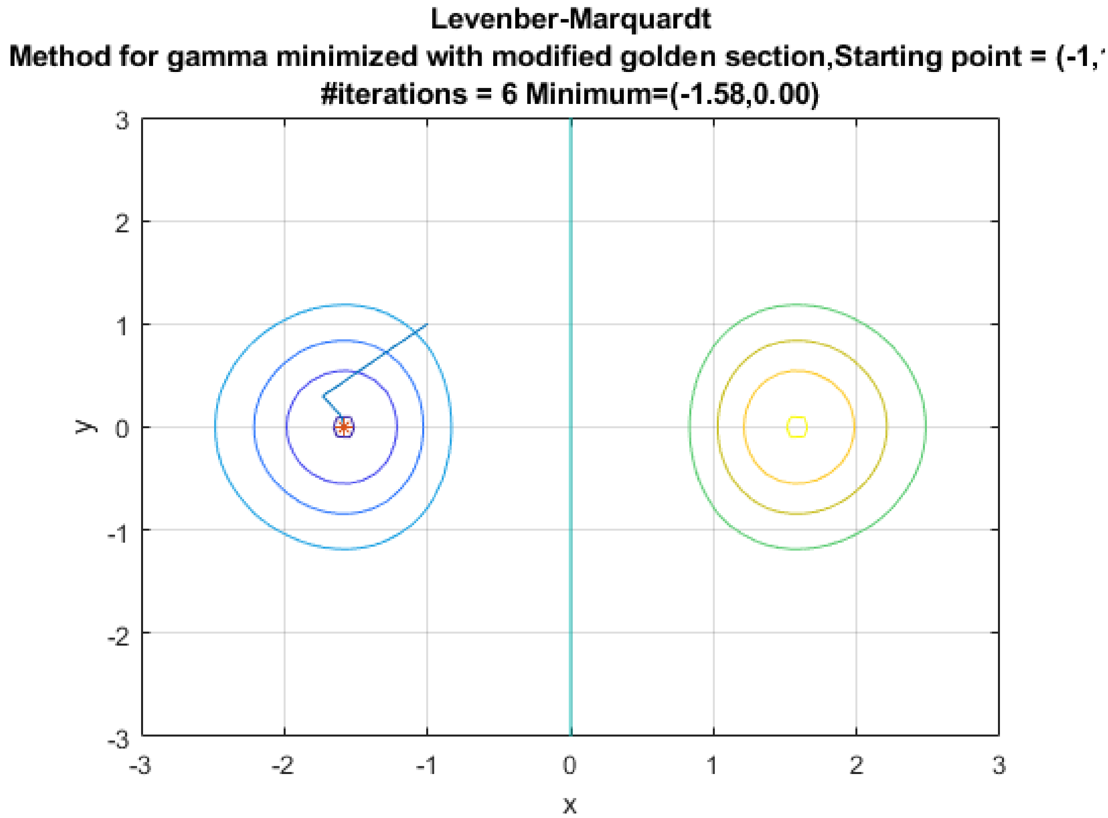
Για αρχικό σημείο εκκίνησης το (0,0) εμφανίζονται ξανά τα ίδια προβλήματα εγκλωβισμού και ο αλγόριθμος μένει στάσιμος και δεν εκτελεί καμία επανάληψη.

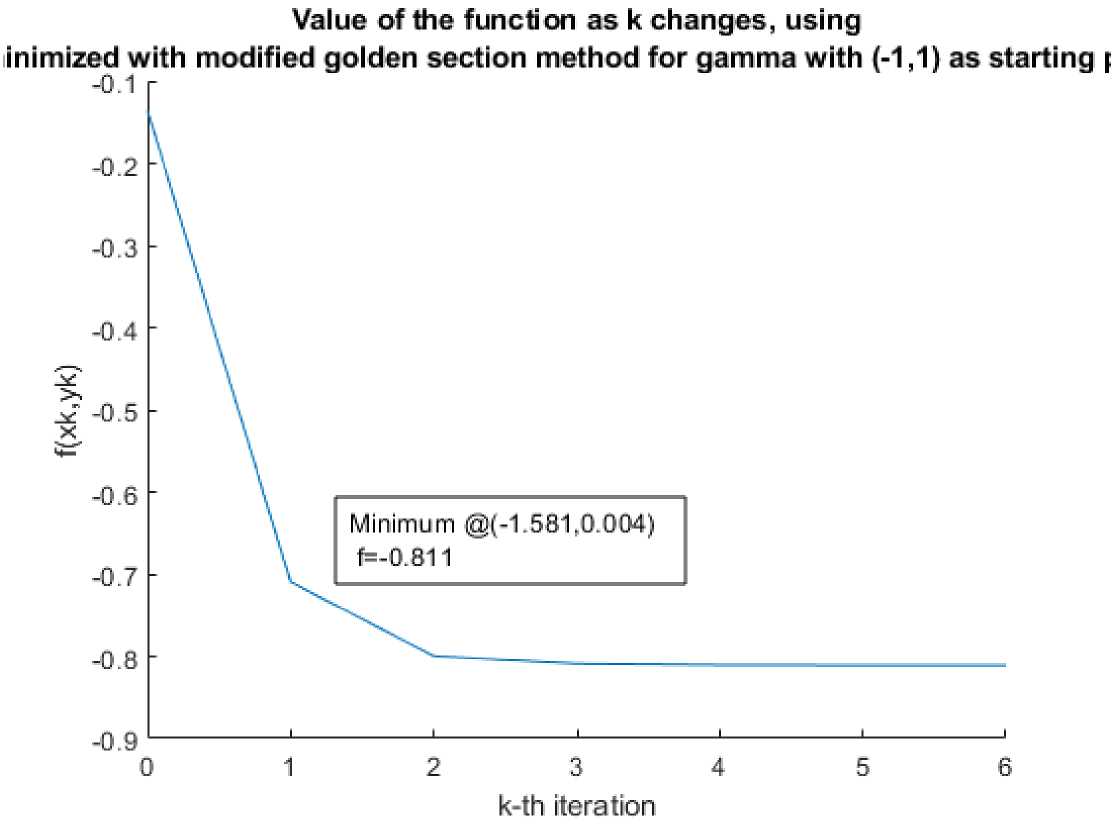


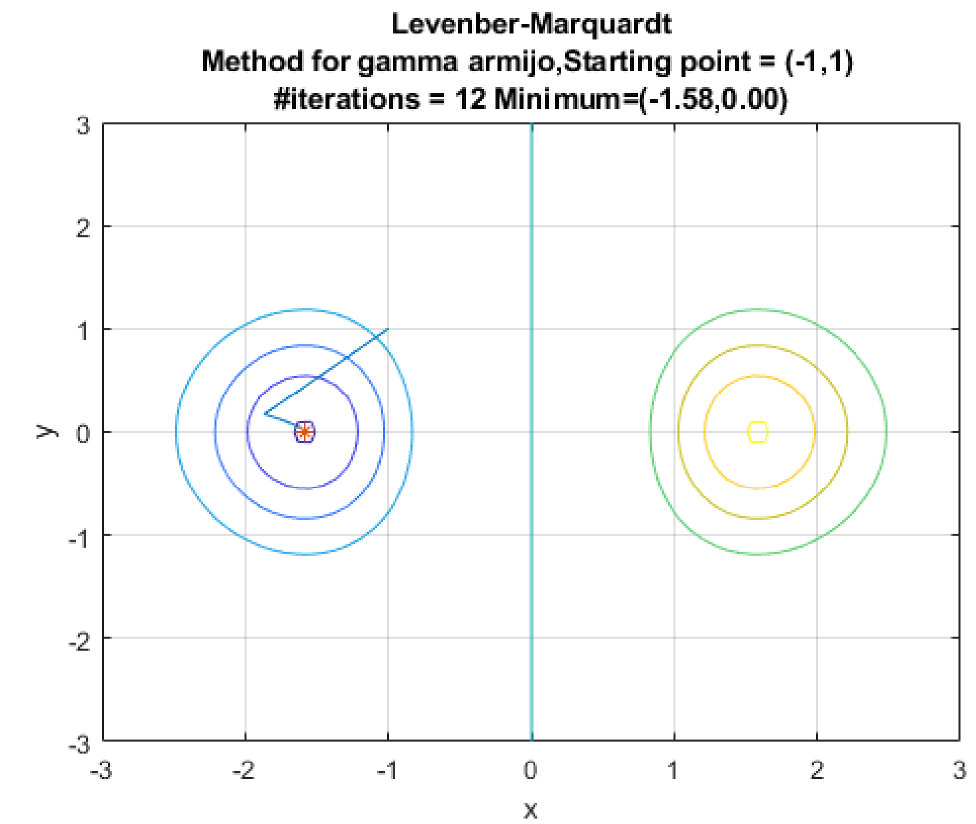
Για σημείο εκκίνησης το (-1,1) ο αλγόριθμος λειτουργεί καλά και προσεγγίζει το ζητούμενο ελάχιστο.

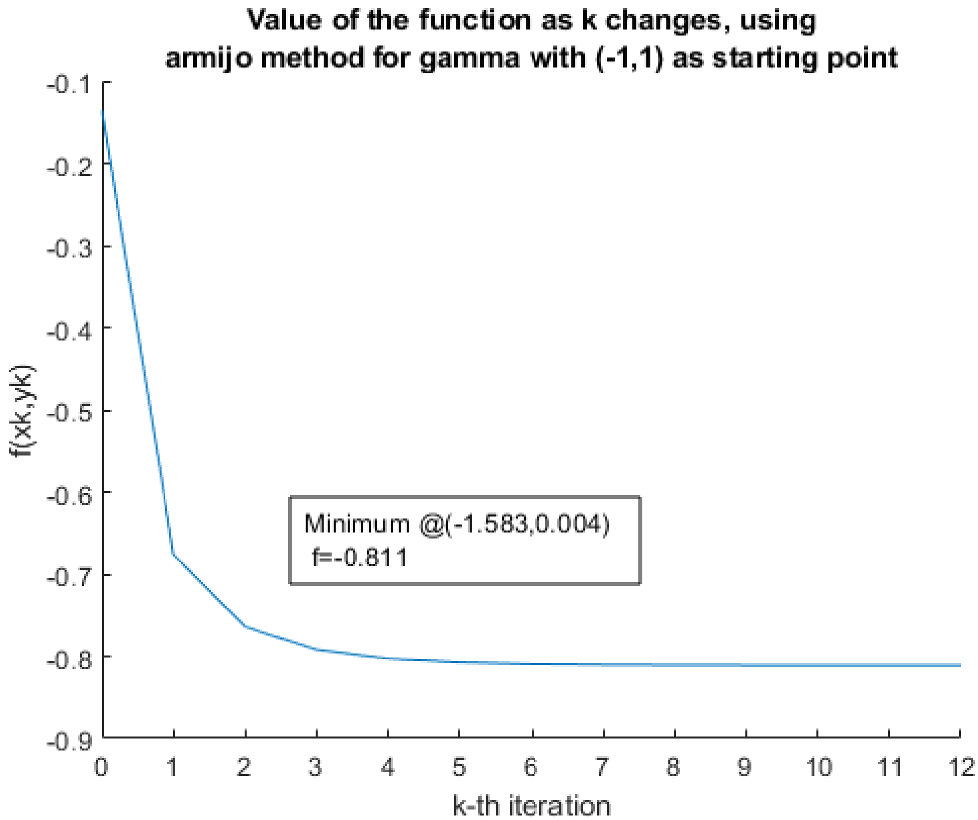

Levenber-Marquardt
Method for gamma const,Starting point = (-1,1)
#iterations = 12 Minimum=(-1.58,0.00)
3
2
1
0
-1
-2
-3
-3
-2
-1
0
1
2
3
x
y


Value of the function as k changes, using
const method for gamma with (-1,1) as starting point
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.7
-0.8
-0.9
0
1
2
3
4
5
6
7
8
9
10
11
12
k-th iteration
f(xk,yk)
Minimum @(-1.583,0.004)
f=-0.811


Levenber-Marquardt
Method for gamma minimized with modified golden section,Starting point = (-1,
#iterations = 6 Minimum=(-1.58,0.00)
y
x
3
2
1
0
-1
-2
-3
-3
-2
-1
0
1
2
3

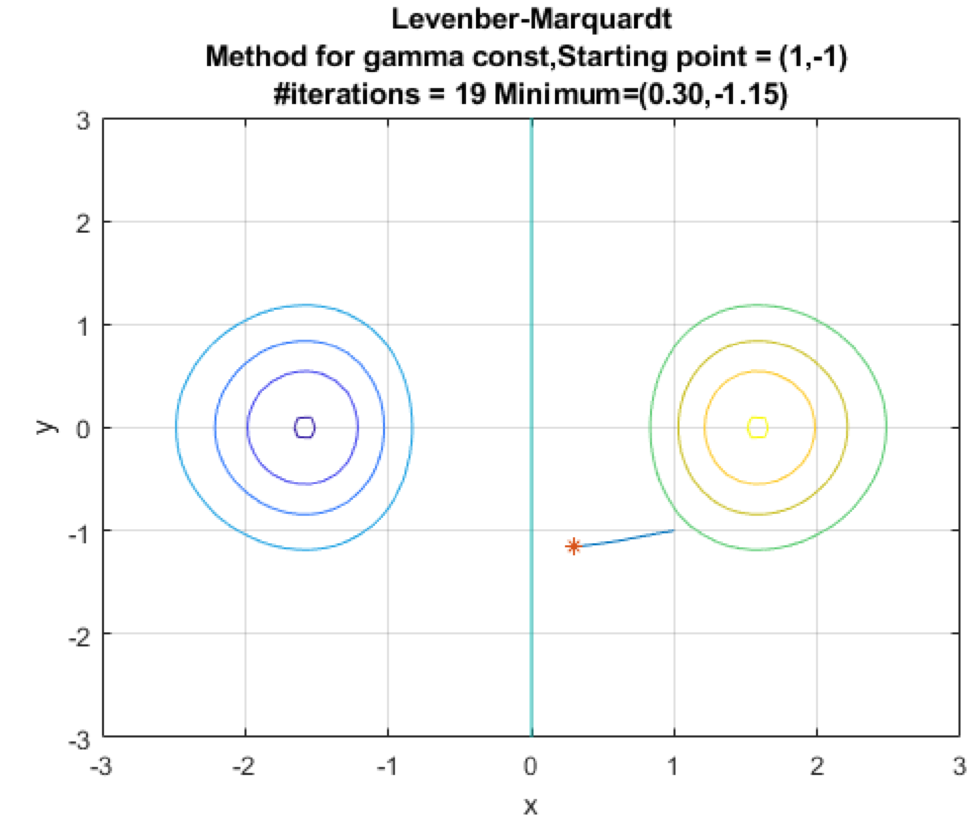

Value of the function as k changes, using
inimized with modified golden section method for gamma with (-1,1) as starting
f(xk,yk)
k-th iteration
Minimum @(-1.581,0.004)
f=-0.811
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.7
-0.8
-0.9
0
1
2
3
4
5
6

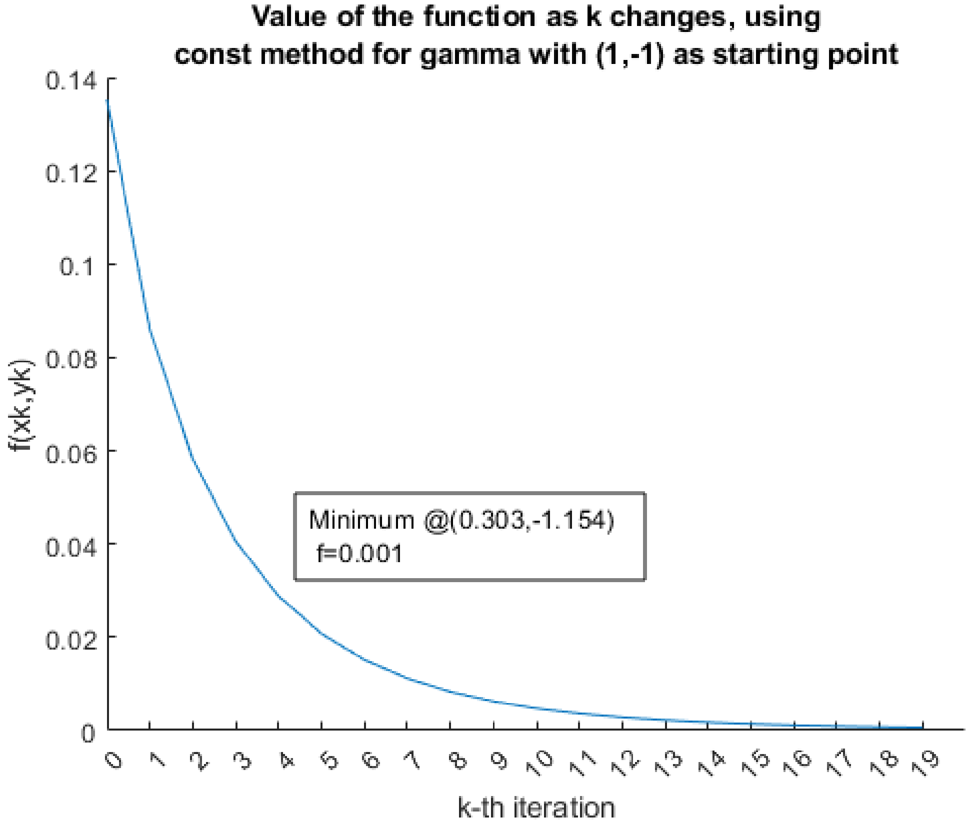


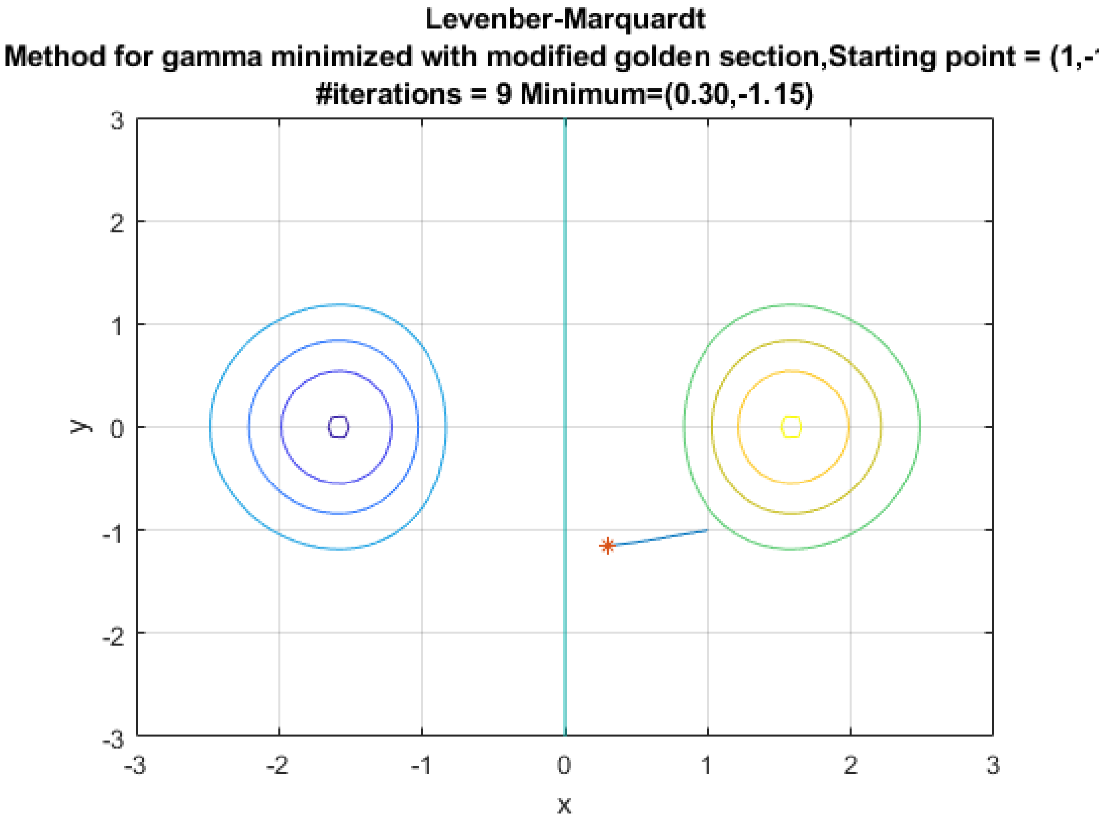


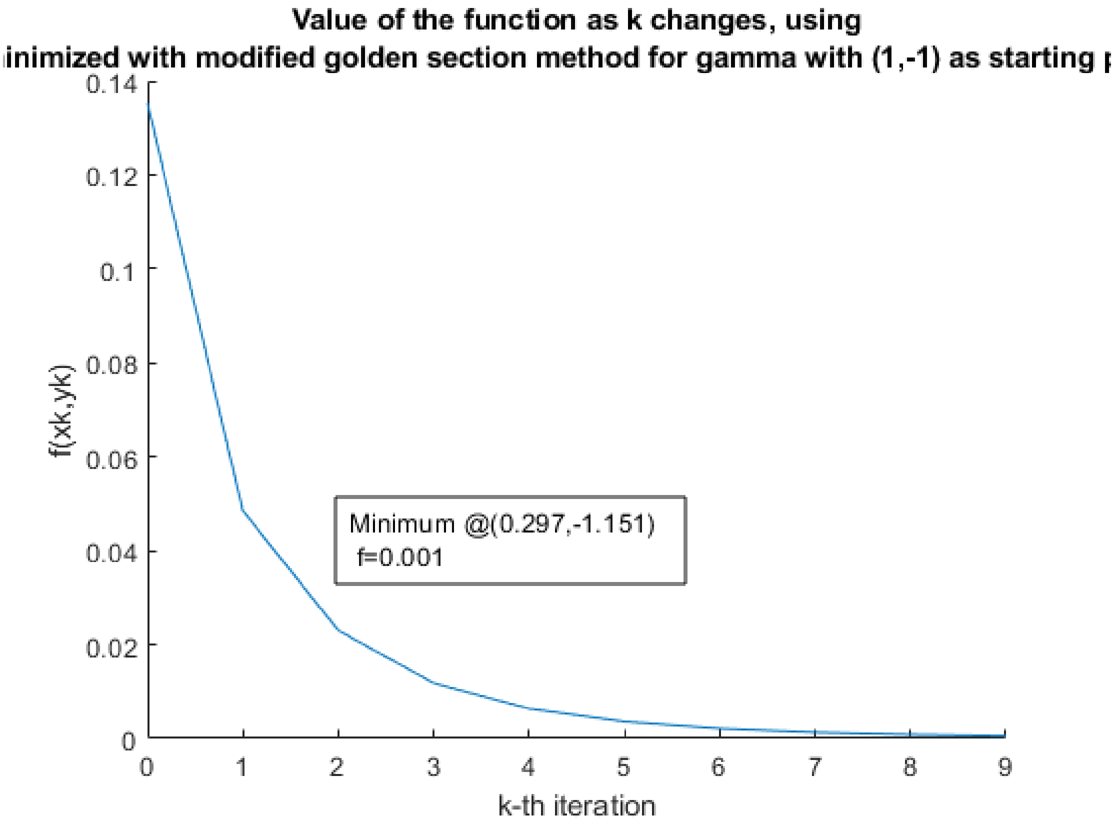
Όπως και στις προηγούμενες θεματικές της αναφοράς, παρατηρούμε πως η υλοποίηση με γ επιλεγμένο μέσω της μεθόδου (ii) είναι αυτή που παρουσιάζει την ταχύτερη σύγκλιση στο ζητούμενο σημείο.

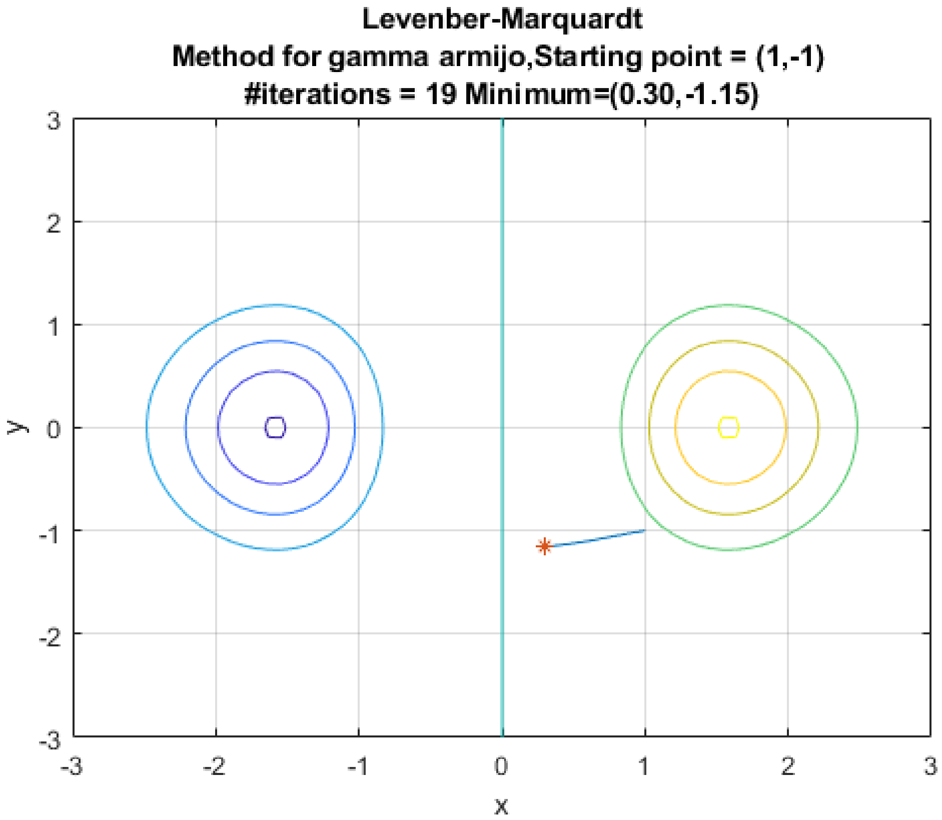
Τέλος, για αρχικό σημείο (1,-1), αντιμετωπίζουμε το ίδιο πρόβλημα που παρουσιάστηκε στο θέμα 2 με τον αλγόριθμο να εγκλωβίζεται σε μια περιοχή μεταξύ τοπικού ελαχίστου και ισουψών καμπυλών μεγίστου και συνεπώς να υπολειτουργεί, με την κυριαρχία της μεθόδου (ii) όσον αφορά την επιλογή βήματος να κυριαρχεί και εδώ παρά το «πλασματικό» των αποτελεσμάτων.

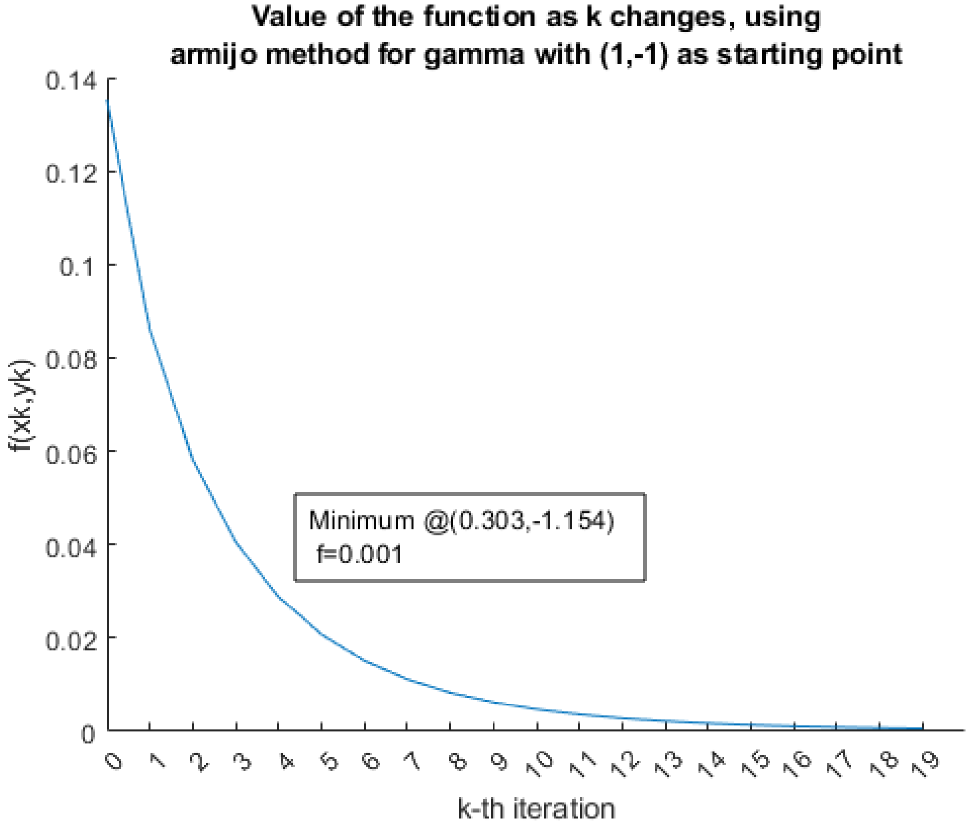

Levenber-Marquardt
Method for gamma const,Starting point = (1,-1)
#iterations = 19 Minimum=(0.30,-1.15)
y
x


Value of the function as k changes, using
const method for gamma with (1,-1) as starting point
f(xk,yk)
Minimum @(0.303,-1.154)
f=0.001
k-th iteration


Levenber-Marquardt
Method for gamma minimized with modified golden section,Starting point = (1,-
#iterations = 9 Minimum=(0.30,-1.15)
y
x
3
2
1
0
-1
-2
-3
-3
-2
-1
0
1
2
3


Value of the function as k changes, using
inimized with modified golden section method for gamma with (1,-1) as starting
f(xk,yk)
k-th iteration
Minimum @(0.297,-1.151)
f=0.001
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0
1
2
3
4
5
6
7
8
9


Levenber-Marquardt
Method for gamma armijo,Starting point = (1,-1)
#iterations = 19 Minimum=(0.30,-1.15)
3
2
1
0
-1
-2
-3
y
-3
-2
-1
0
1
2
3
x


Value of the function as k changes, using
armijo method for gamma with (1,-1) as starting point
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
f(xk,yk)
Minimum @(0.303,-1.154)
f=0.001
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
k-th iteration

<u>ΣΥΜΠΕΡΑΣΜΑΤΑ:</u>

Από τις παραπάνω αναλύσεις, εξάγεται το συμπέρασμα ότι μόνο για αρχικό σημείο εκκίνησης (-1,1) οι αλγόριθμοι που παρουσιάστηκαν μπορούν να δώσουν ασφαλή και ορθά αποτελέσματα. Για σημείο εκκίνησης το (0,0) οι αλγόριθμοι μένουν στάσιμοι και δεν εκτελούν καμία επανάληψη. Ενώ, για αρχικό σημείο (1,-1) οδηγούμαστε πάλι σε εγκλωβισμό στην πλευρά που η τρισδιάστατη απεικόνιση της δοθείσας συνάρτησης παίρνει θετικές τιμές.

Επίσης, διαπιστώνουμε πως η μέθοδος Newton καθίσταται απαγορευτική ως προς την χρήση της για το συγκεκριμένο πρόβλημα μιας και ο Εσσιανός πίνακας της $f(x,y) = x^5 e^{-x^2-y^2}$ δεν είναι Θ.Ο.

Στην υλοποίηση των άλλων 2 μεθόδων και για σημείο εκκίνησης (-1,1) γίνεται αντιληπτό πως η αποδοτικότερη μέθοδος επιλογής του βήματος επανάληψης είναι η μέθοδος υπ' αριθμόν (ii) καθώς μας οδηγεί σε μικρότερο αριθμό επαναλήψεων.

Τέλος, για την παρούσα αναφορά και την δοθείσα συνάρτηση συμπεραίνουμε ότι ο πιο αποδοτικός αλγόριθμος είναι αυτός της μέγιστης καθόδου καθώς οδηγεί σε σωστό αποτέλεσμα με μόλις 6 επαναλήψεις (περίπτωση επιλογής βέλτιστου γ) έναντι 9 της μεθόδου Levenberg-Marquardt.