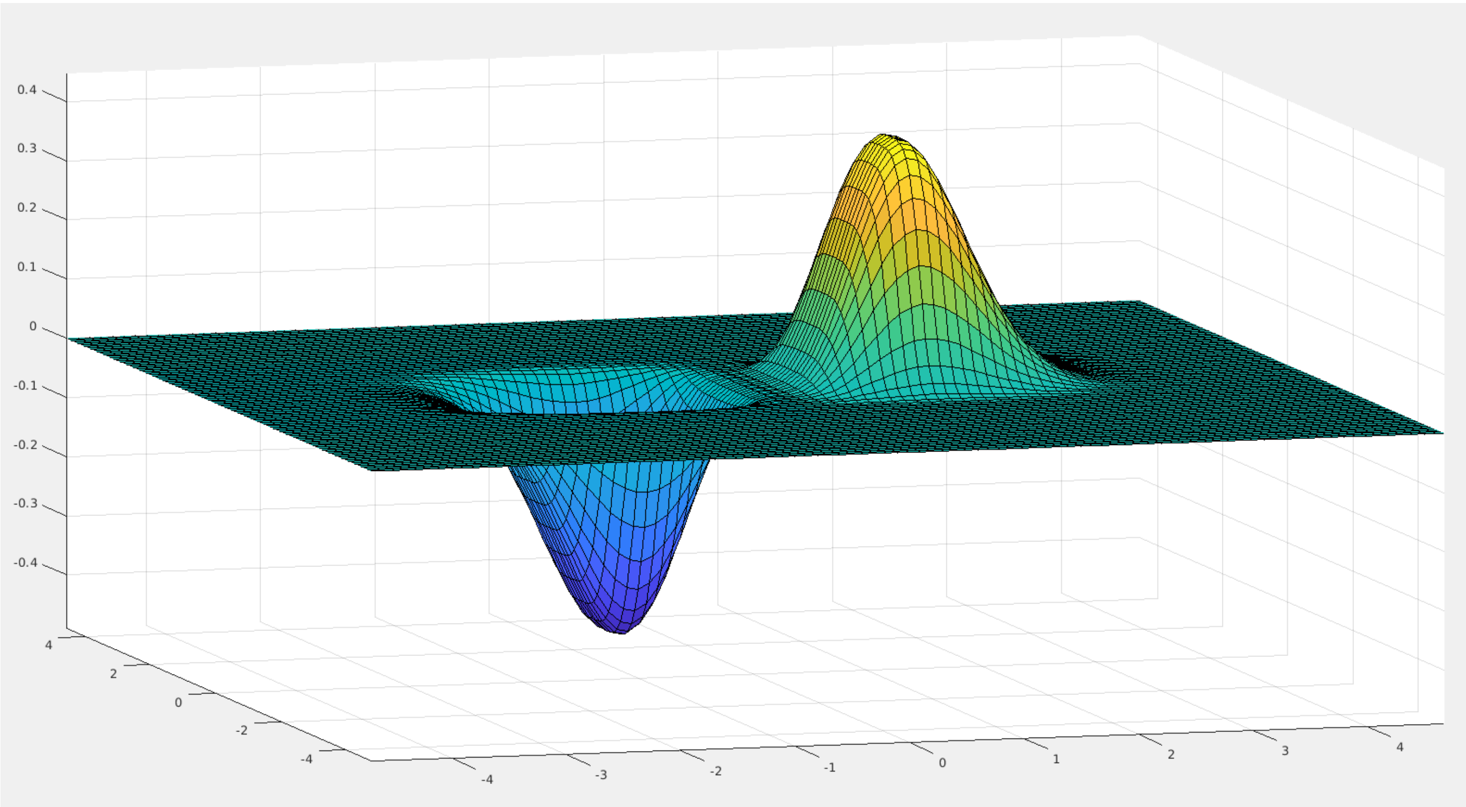
Τεχνικές Βελτιστοποίησης

2η Εργασία

Ιορδάνης Κωνσταντινίδης

ΑΕΜ: 9492

iordaniak@ece.auth.gr



Εισαγωγή

Στα πλαίσια της 2ης εργασίας ασχολούμαστε με το πρόβλημα ελαχιστοποίησης μιας συνάρτησης πολλών μεταβλητών χρησιμοποιώντας αλγόριθμους που βασίζονται στην ιδέα της επαναληπτικής καθόδου. Η συνάρτηση μας είναι 2 μεταβλητών δίνεται από τον τύπο:

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

Οι αλγόριθμοι αναζήτησης που θα χρησιμοποιήσουμε είναι οι :

- Μέθοδος Μέγιστης Καθόδου
- Μέθοδος Newton
- Μέθοδος Levenberg - Marquardt

Για το Θέμα 1 της εργασίας το ζητούμενο είναι να σχεδιαστεί η γραφική παράσταση της f.

Για το Θέμα 2 μας ζητείται να ελαχιστοποιηθεί η f με την χρήση της μεθόδου Μέγιστης Καθόδου χρησιμοποιώντας ως αρχικά σημεία τα
$x_0 = (0, 0)$, $x_0 = (1, 1)$, $x_0 = (-1, -1)$
και βήμα γ_k :

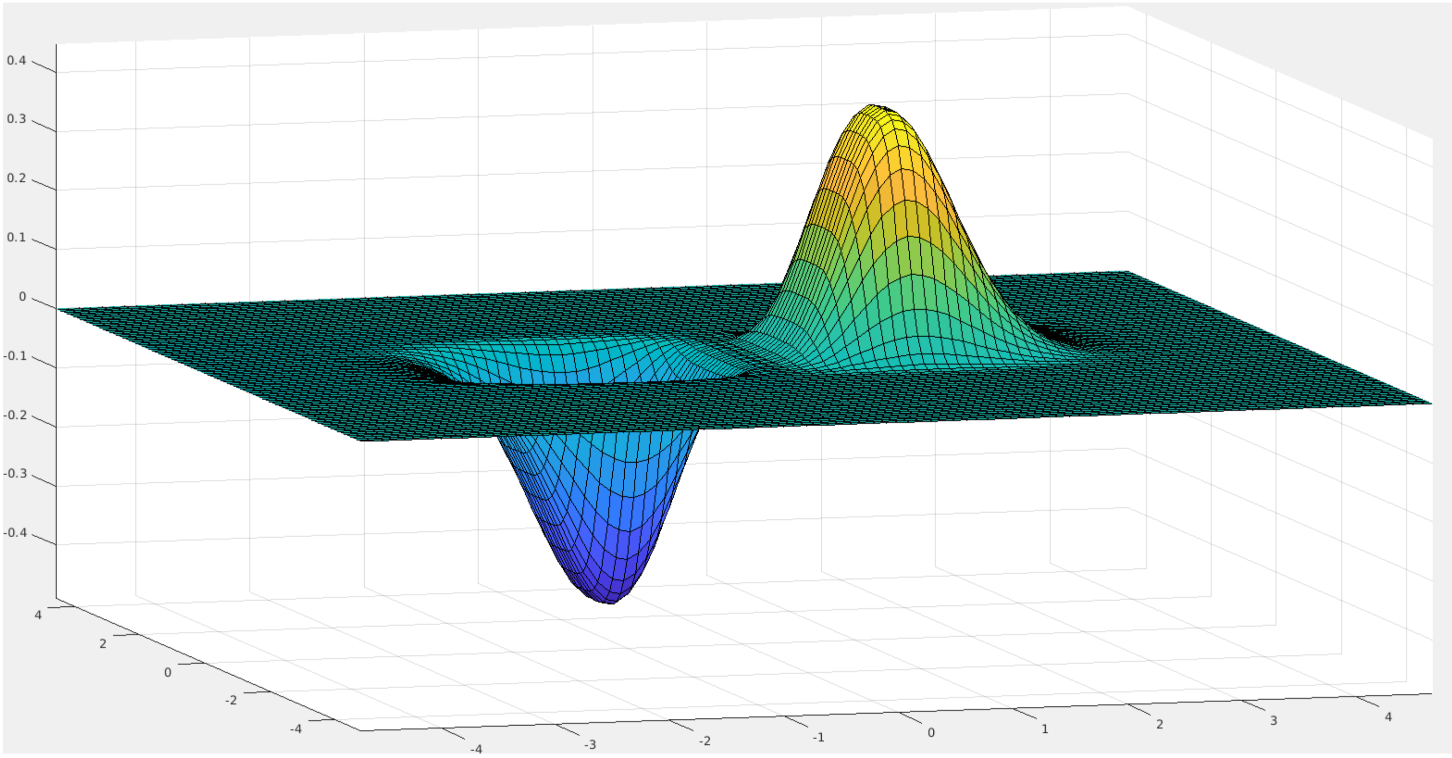
A. σταθερό
B. τέτοιο ώστε να ελαχιστοποιεί την $f\ (x_k + \gamma_k d_k)$
C. βάσει του κανόνα Armijo

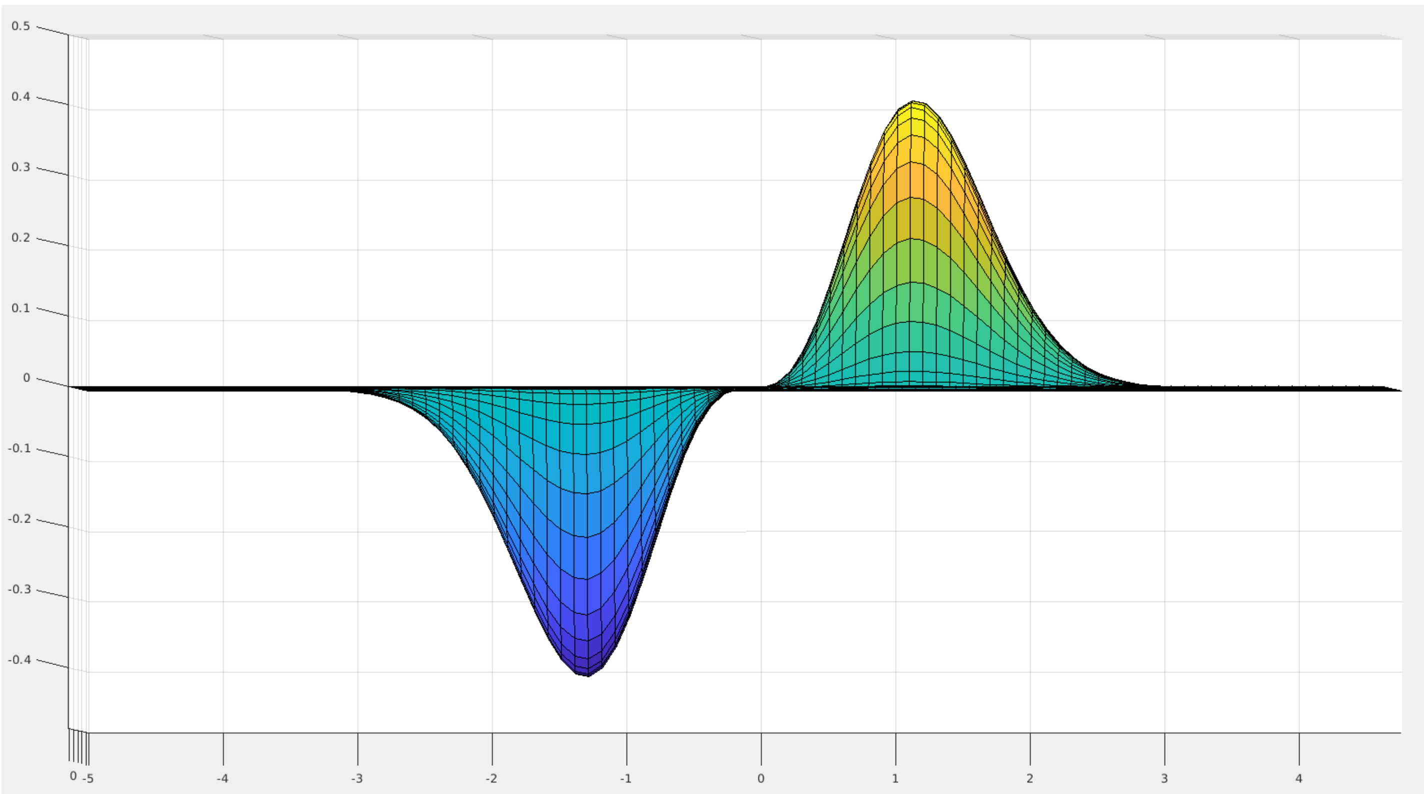
Για το Θέμα 3 τα ζητούμενα είναι ίδια με αυτά του θέματος 2 με την διαφορά ότι θα χρησιμοποιήσουμε την μέθοδο Newton

Για το Θέμα 4 τα ζητούμενα είναι ίδια με αυτά του θέματος 2 με την διαφορά ότι θα χρησιμοποιήσουμε την μέθοδο Levenberg - Marquardt

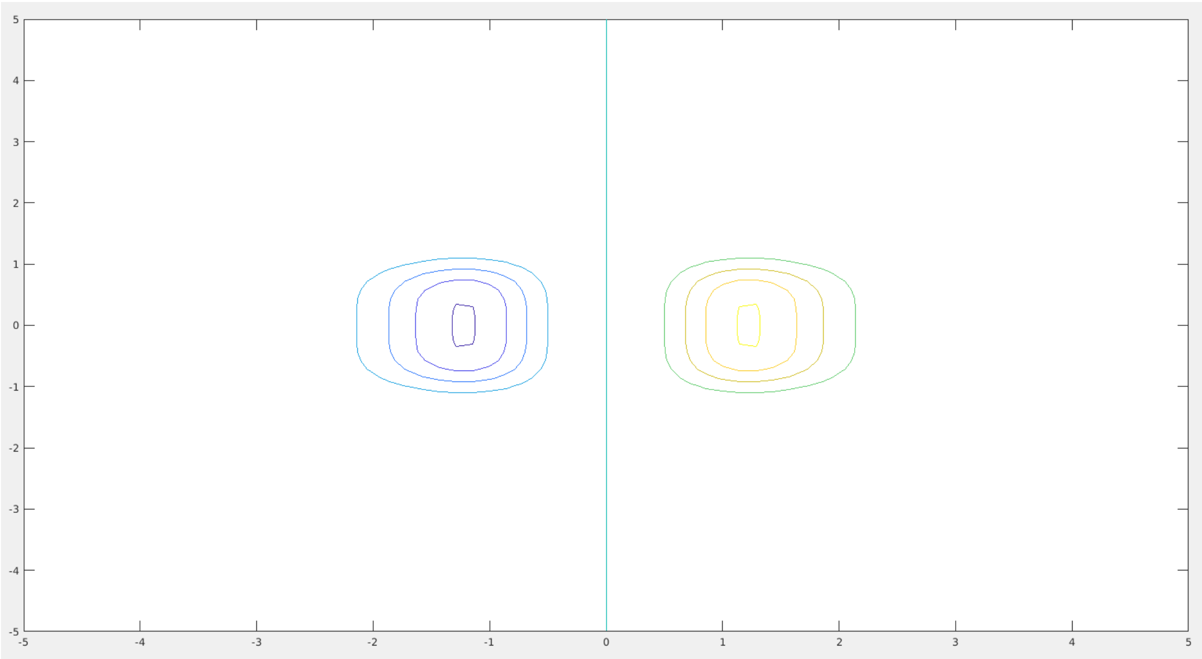
Θέμα 1 - Γραφική Παράσταση της Συνάρτησης

Το αρχείο matlab βρίσκεται στον φάκελο thema1 με τίτλο thema1.m

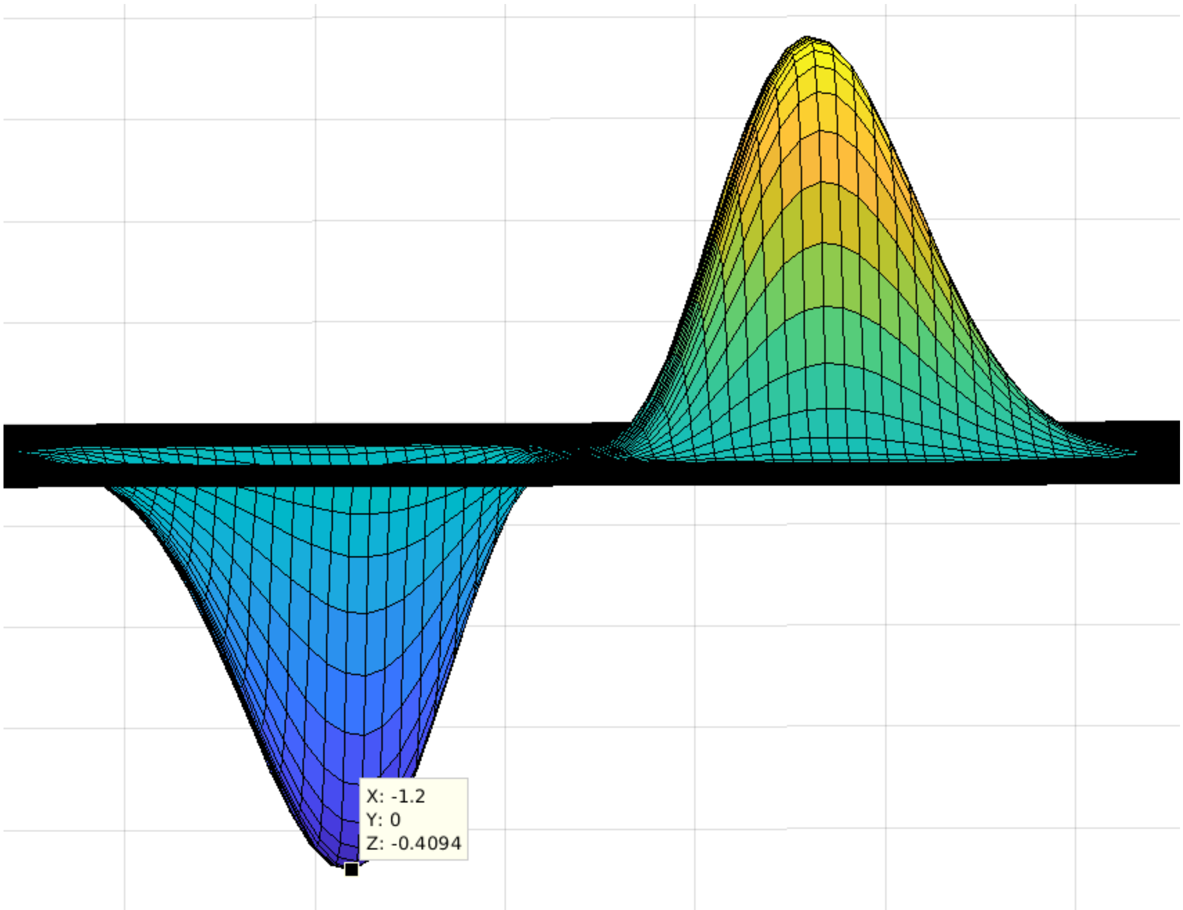




Η συνάρτηση αποτελείται από δύο κώνους έξω από τους οποίους είναι πρακτικά 0 καθώς υπερισχύει ο όρος: $e^{-x^2-y^4}$. Παρακάτω φαίνονται οι οριακοί κύκλοι της συνάρτησης

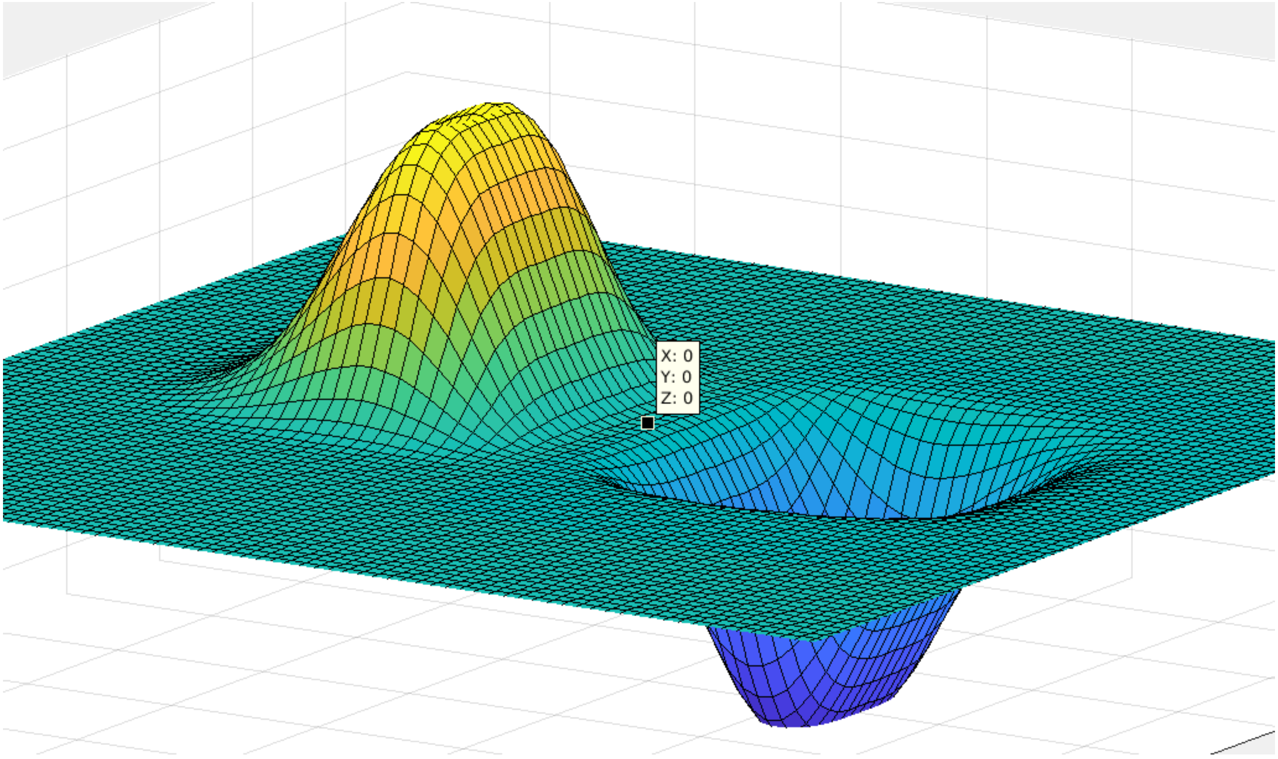


Παρακάτω φαίνεται το ολικό ελάχιστο της συνάρτησης στο σημείο $x^* \approx (-\,1.2, 0)$ με $f\ (x^*)\ \approx\ \ -\ 0.4094$

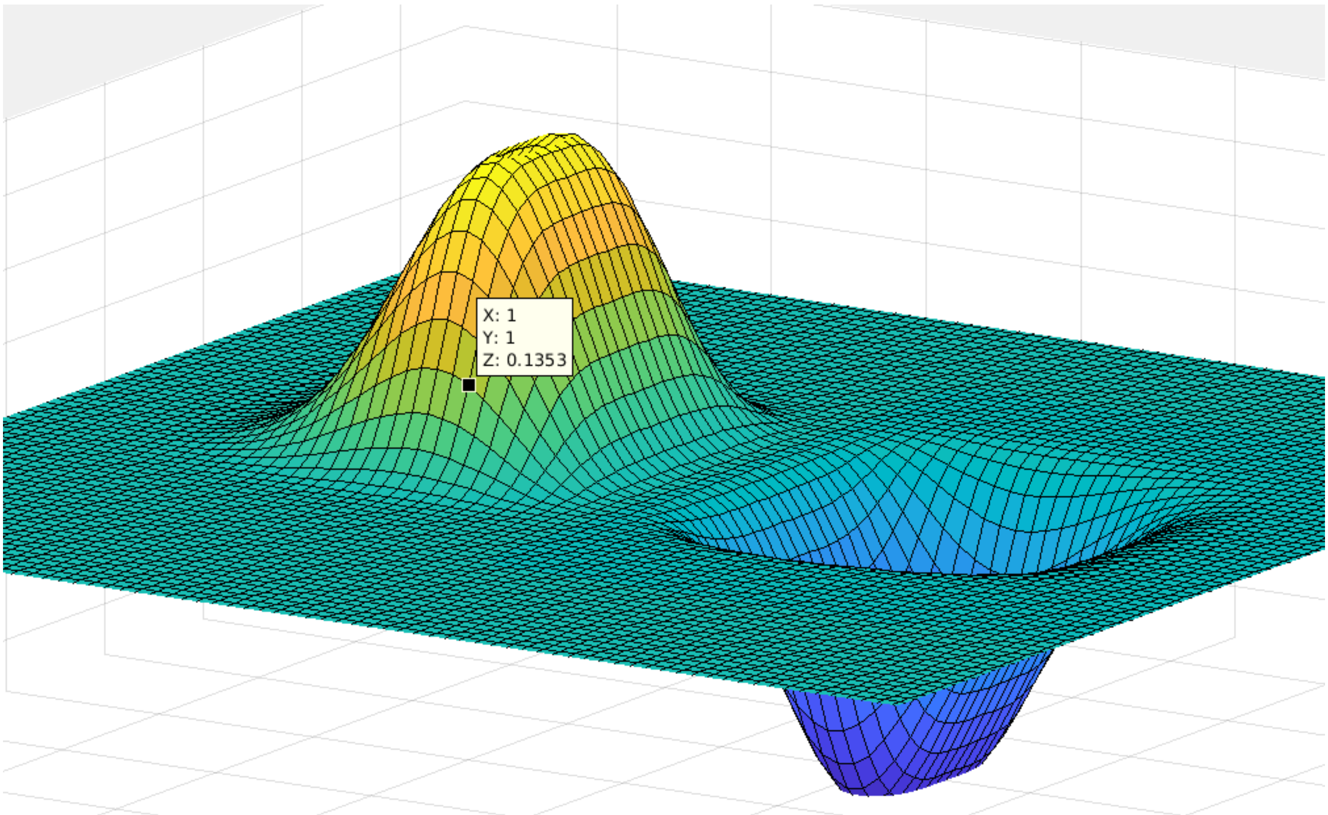


Παρακάτω φαίνονται τα αρχικά σημεία x_0 τα οποία μαζί με το βήμα γ_k θα παίξουν καθοριστικό ρόλο στα αποτελέσματα των αλγορίθμων μας

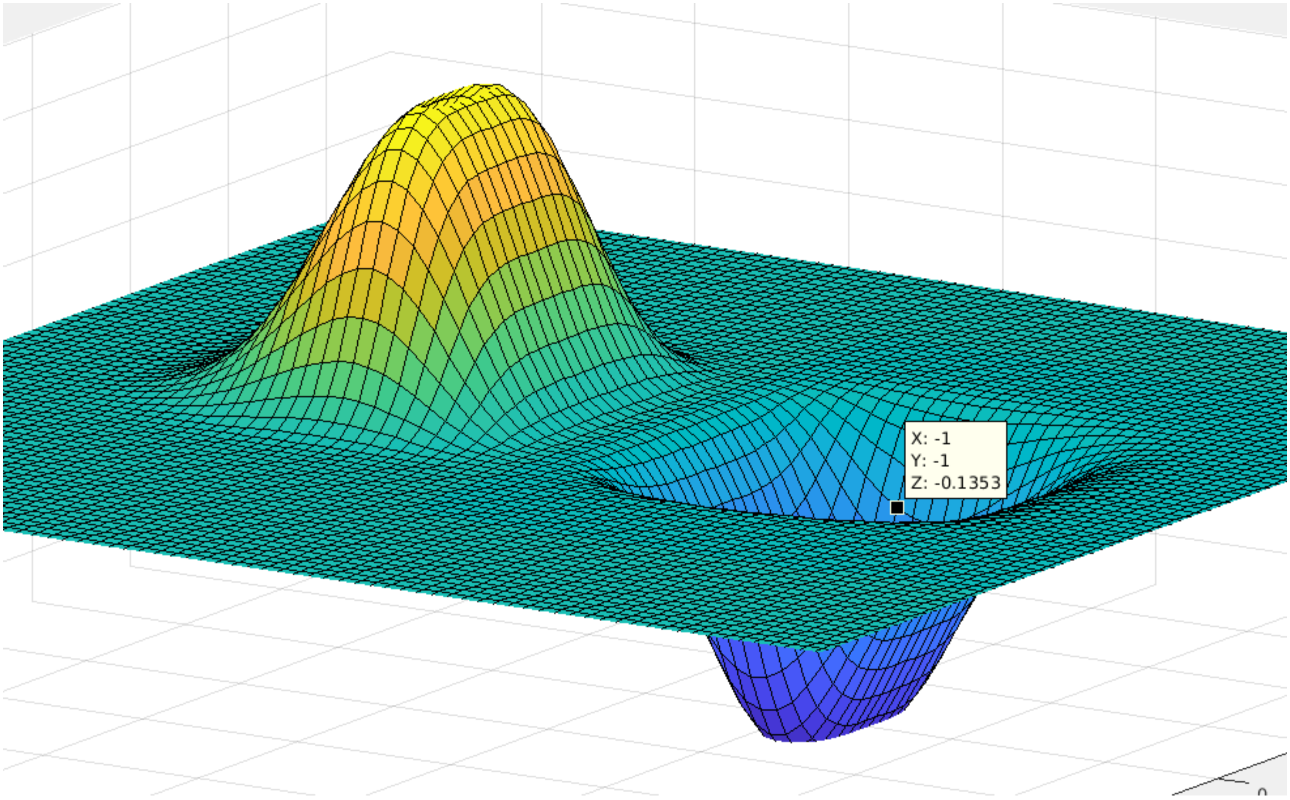
$x_0 = (0, 0)$



$x_0 = (1, 1)$



$x_0 = (-1, -1)$



Θέμα 2 - Μέθοδος Μέγιστης Καθόδου

Α) Το αρχείο matlab βρίσκεται στον φάκελο thema2 με τίτλο thema2_A.m
Στο Α ζητούμενο καλούμαστε να επιλέξουμε ένα σταθερό βήμα γ_k. Από την θεωρία γνωρίζουμε ότι το γ_k δεν μπορεί να επιλεγεί οσοδήποτε μικρό σύμφωνα με το 3ο κριτήριο. Επίσης δεν μπορεί να επιλεγεί ούτε και πολύ μεγάλο αφού αυτό θα οδηγούσε σε μεγάλο βήμα από το x_k στο x_{k+1}, χωρίς αυτό πάντα να αποτελεί και την ενδεδειγμένη κίνηση προς την ελαχιστοποίηση της $f\ (x)$. Είδαμε ότι μπορεί να οδηγηθούμε σε ταλαντώσεις ή ακόμα και σε αποκλίσεις όταν το βήμα μας είναι μεγάλο.

Για τα 3 διαφορετικά αρχικά σημεία x_0 καταλήγουμε στα εξής σημεία x_k :

- Για $\gamma_k = \gamma\ = 0.5$

x0 = (0,0) ----> xk = (0.000000, 0.000000)

x0 = (-1,-1) ----> xk = (-1.224745, -0.084684)

x0 = (1,1) ----> xk = (0.877512, 1.720780)

- Για $\gamma_k = \gamma\ = 1$

x0 = (0,0) ----> xk = (0.000000, 0.000000)

x0 = (-1,-1) ----> xk = (-1.224745, -0.084014)

x0 = (1,1) ----> xk = (0.846096, 1.718477)

- Για $\gamma_k = \gamma \ = 3$

x0 = (0,0) ----> xk = (0.000000, 0.000000)

x0 = (-1,-1) ----> xk = (-1.224580, -0.015504)

x0 = (1,1) ----> xk = (0.593994, 2.624023)

- Για $\gamma_k = \gamma \ = 4$

x0 = (0,0) ----> xk = (0.000000, 0.000000)

x0 = (-1,-1) ----> xk = (0.015525, -0.060331)

x0 = (1,1) ----> xk = (0.458659, 3.165365)

- Για $\gamma_k = \gamma \ = 5$

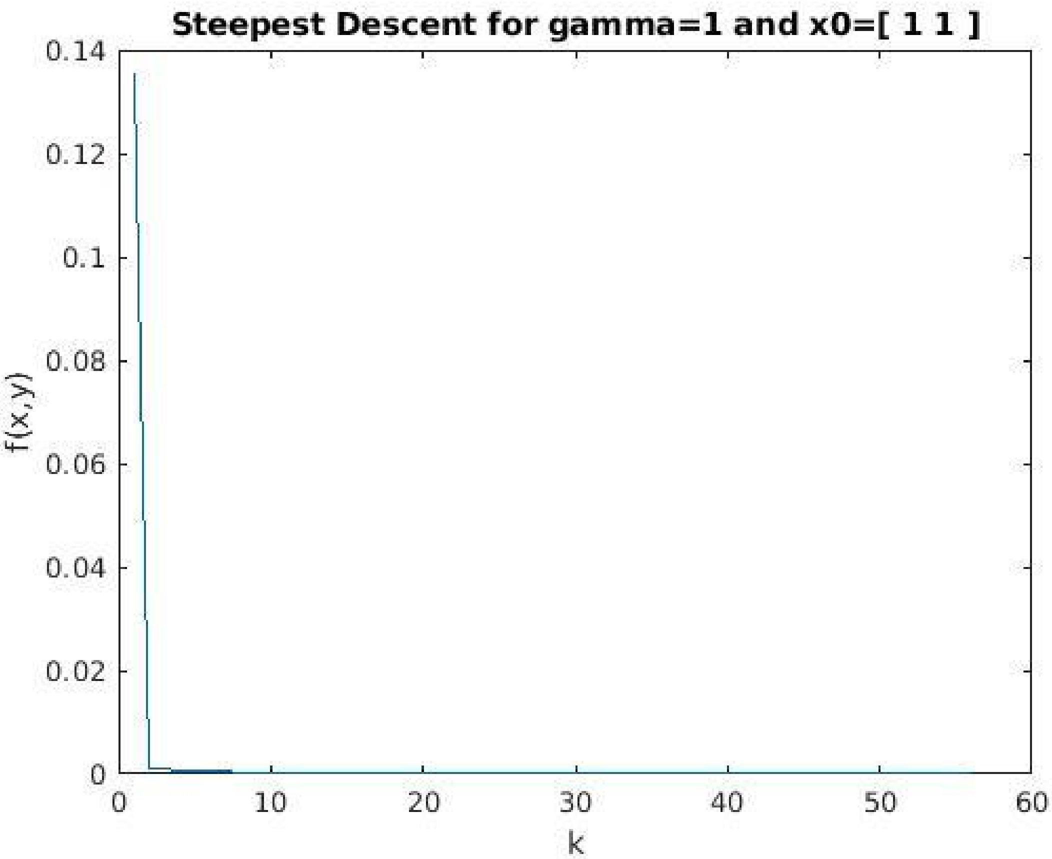
x0 = (0,0) ----> xk = (0.000000, 0.000000)

x0 = (-1,-1) ----> xk = (-3.753575, -0.098742)

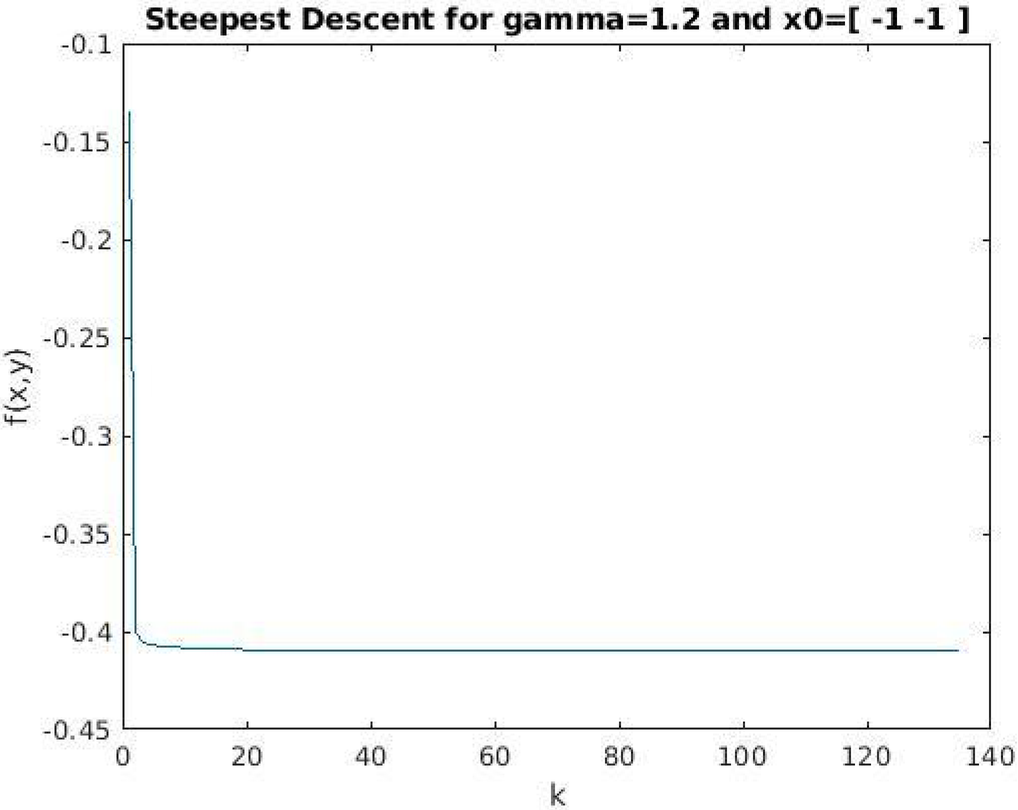
x0 = (1,1) ----> xk = (0.323324, 3.706706)

Βλέποντας το αρχικό σημείο $x_0 = (0, 0)$ στην γραφική παράσταση και παρατηρώντας την δοσμένη συνάρτηση f η οποία περιλαμβάνει τον όρο x^3 καταλαβαίνουμε ότι το σημείο αυτό είναι ιδιαίτερο. Συγκεκριμένα, τόσο η συνάρτηση f όσο και οι πρώτοι παράγωγοι της μηδενίζονται στο σημείο αυτο. Καταλαβαίνουμε ότι πρόκειται για ένα κρίσιμο σημείο και όχι για ένα ακρότατο, καθώς η f παίρνει μεγαλύτερες και μικρότερες τιμές κοντά στο (0,0). Συμπεραίνουμε ότι ο αλγόριθμος μας εγκλωβίζεται σε αυτό το κρίσιμο σημείο ανεξαρτήτως του βήματος γ αφού η παράγωγος είναι 0 < ε. Το ίδιο συμβαίνει και στις υπόλοιπες μεθόδους οπότε δεν θα αναφερθούμε ξανά στο σημείο αυτό.

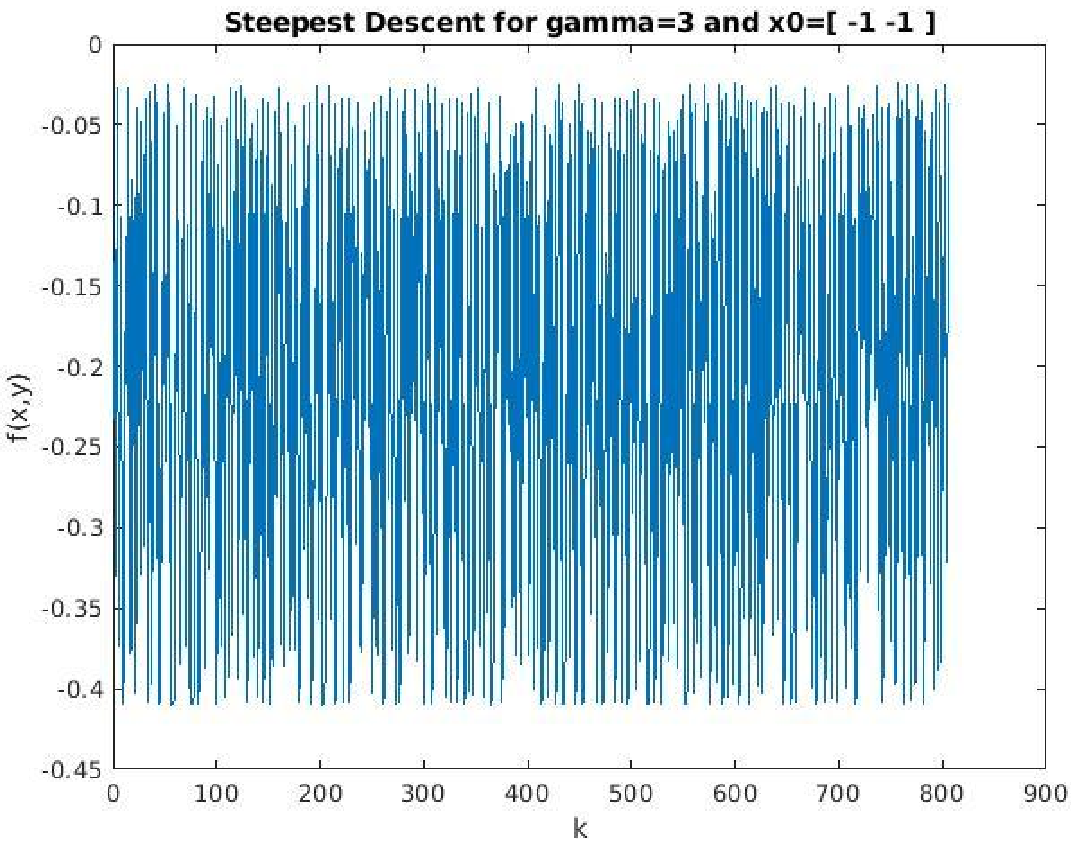
Για αρχικό σημείο το $x_0 = (1, 1)$, κοιτώντας την γραφική παράσταση, παρατηρούμε ότι βρίσκεται πάνω στον θετικό λόφο και τα αποτελέσματα μας δείχνουν ότι καταλήγει, σαν x_k , σε 2 διαφορετικές περιοχές. Σίγουρα καμία από τις περιοχές αυτές δεν περιέχει το ολικό ελάχιστο όμως έχουν κάτι κοινό μεταξύ τους. Και οι δύο περιοχές βρίσκονται έξω από τον λόφο, στο επίπεδο όπου η f μηδενίζεται. Έτσι ξεκινώντας από το $x_0 = (1, 1)$ το οποίο βρίσκεται στον θετικό ως προς z λόφο, κατεβαίνουμε με διαφορετικό βήμα γ κάθε φορά στο επίπεδο όπου η κλίση της f είναι μηδενική (ε > 0) και εκεί εγκλωβιζόμαστε σε ένα κρίσιμο σημείο. Παρακάτω βλέπουμε την $f\ (x_k)$ ως προς k



Για αρχικό σημείο το $x_0 = (-1, -1)$, το οποίο βρίσκεται 'μέσα' στον αρνητικό λόφο ο αλγόριθμός μας βγάζει σωστά αποτελέσματα για ένα ορισμένο εύρος του γ_k . Για μεγάλο όμως βήμα παρατηρούμε ταλαντώσεις και αποκλίσεις.



Ταλάντωση για γ = 3



B) Το αρχείο matlab βρίσκεται στον φάκελο thema2 με τίτλο thema2_B.m
Στο B ζητούμενο καλούμαστε να υπολογίσουμε το βήμα γ_k τέτοιο ώστε να ελαχιστοποιεί την $f\ (x_k + \gamma_k d_k)$. Η ελαχιστοποίηση αυτή γίνεται με την βοήθεια της συνάρτησης minimize.m η οποία βρίσκεται στον ίδιο φάκελο και αυτό που κάνει είναι να ελαχιστοποιεί την δοσμένη συνάρτηση ως προς γ_k με την χρήση της μεθόδου της Χρυσής Τομής η οποία αναπτύχθηκε στην προηγούμενη εργασία.

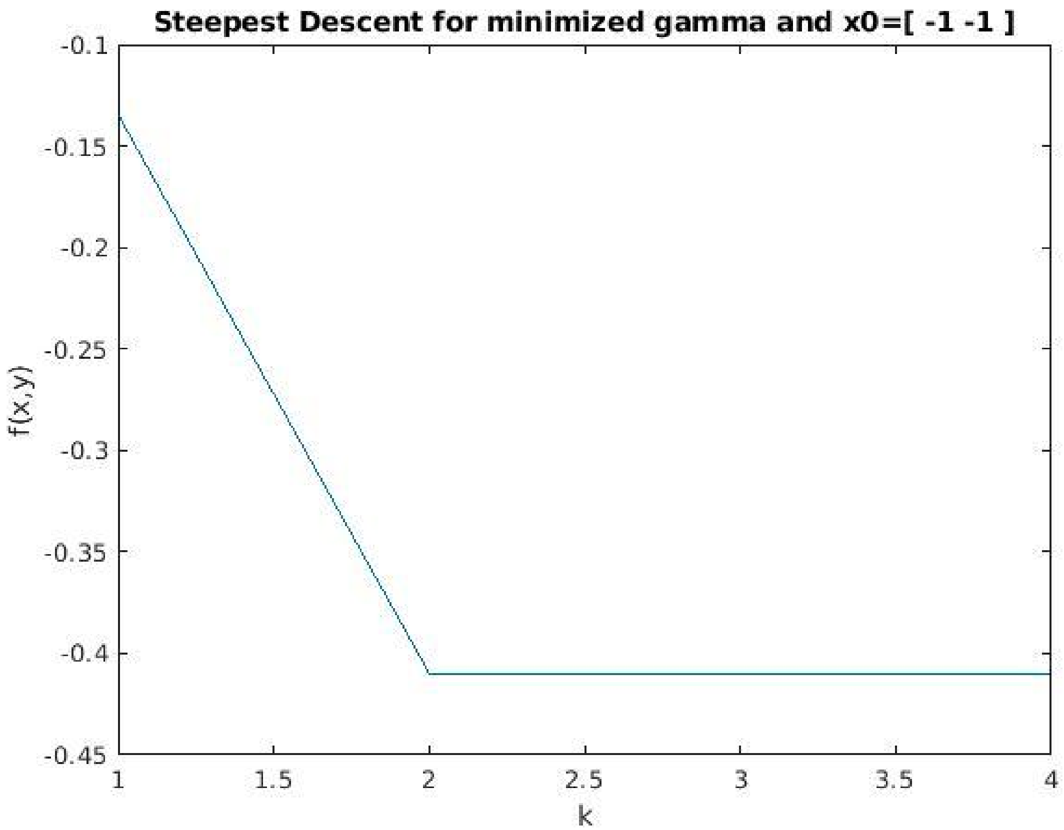
Για τα 3 διαφορετικά αρχικά σημεία x_0 καταλήγουμε στα εξής σημεία x_k :

x0 = (0,0) ----> xk = (0.000000, 0.000000)

x0 = (-1,-1) ----> xk = (-1.224625, -0.079031)

x0 = (1,1) ----> xk = (-0.001236, 5.004944)

Παρατηρούμε ότι ο αριθμός τον επαναλήψεων έχει μειωθεί σημαντικά.
Επίσης $f\ (1, 1)\ \approx\ 0$ με μόλις μία επανάληψη.



C) Το αρχείο matlab βρίσκεται στον φάκελο thema2 με τίτλο thema2_C.m
Στο C ζητούμενο καλούμαστε να υπολογίσουμε το βήμα γ_k βάσει του κανόνα Armijo. Ο υπολογισμός αυτός γίνεται με την βοήθεια της συνάρτησης armijo.m η οποία βρίσκεται στον ίδιο φάκελο. Έχουμε ελευθερία στην επιλογή των τιμών για τις μεταβλητές s, a, β και έπειτα από δοκιμές καταλήξαμε στις:
s = 1, a = 0.01 , β = 0.3 και s = 2, a = 0.01 , β = 0.3
με την δεύτερη όπου s = 2 να έχουμε σημαντικά λιγότερες επαναλήψεις

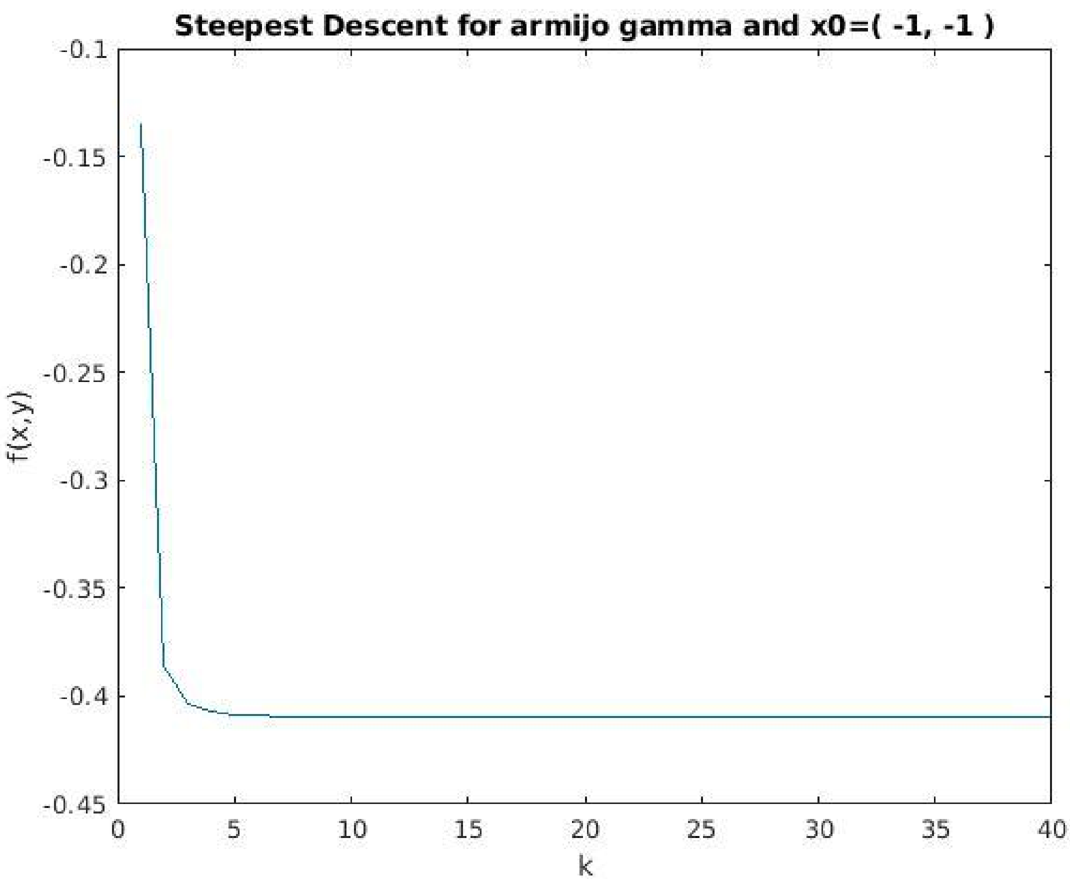
Για τα 3 διαφορετικά αρχικά σημεία x_0 καταλήγουμε στα εξής σημεία x_k :

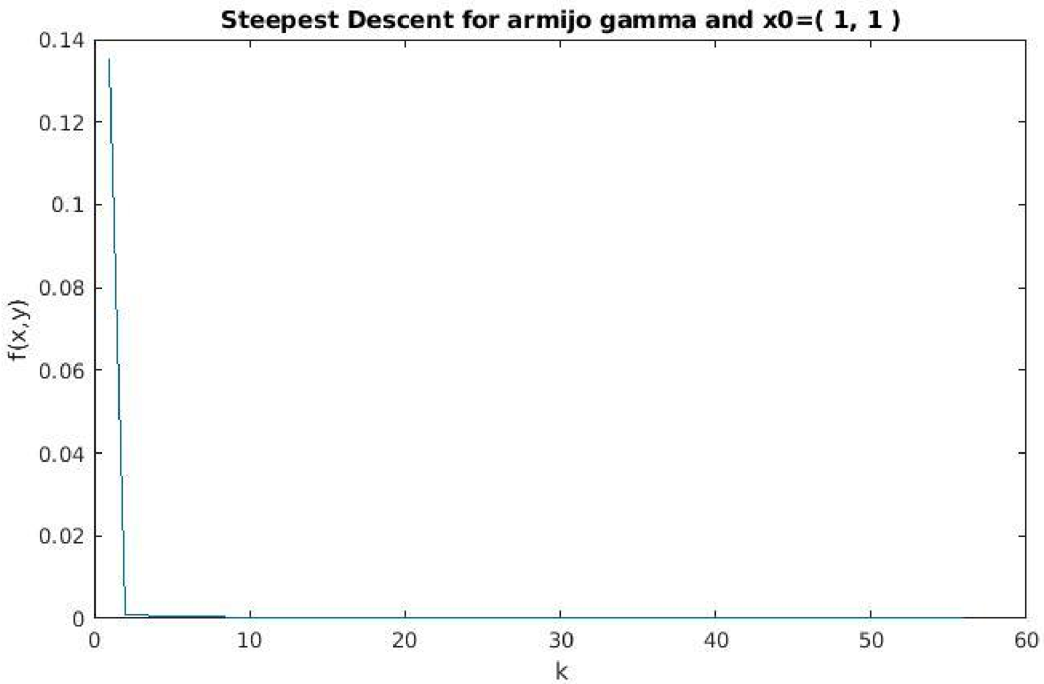
- Για $s = 1$

x0 = (0,0) ----> xk = (0.000000, 0.000000)

x0 = (-1,-1) ----> xk = (-1.224745, -0.084014)

x0 = (1,1) ----> xk = (0.846096, 1.718477)



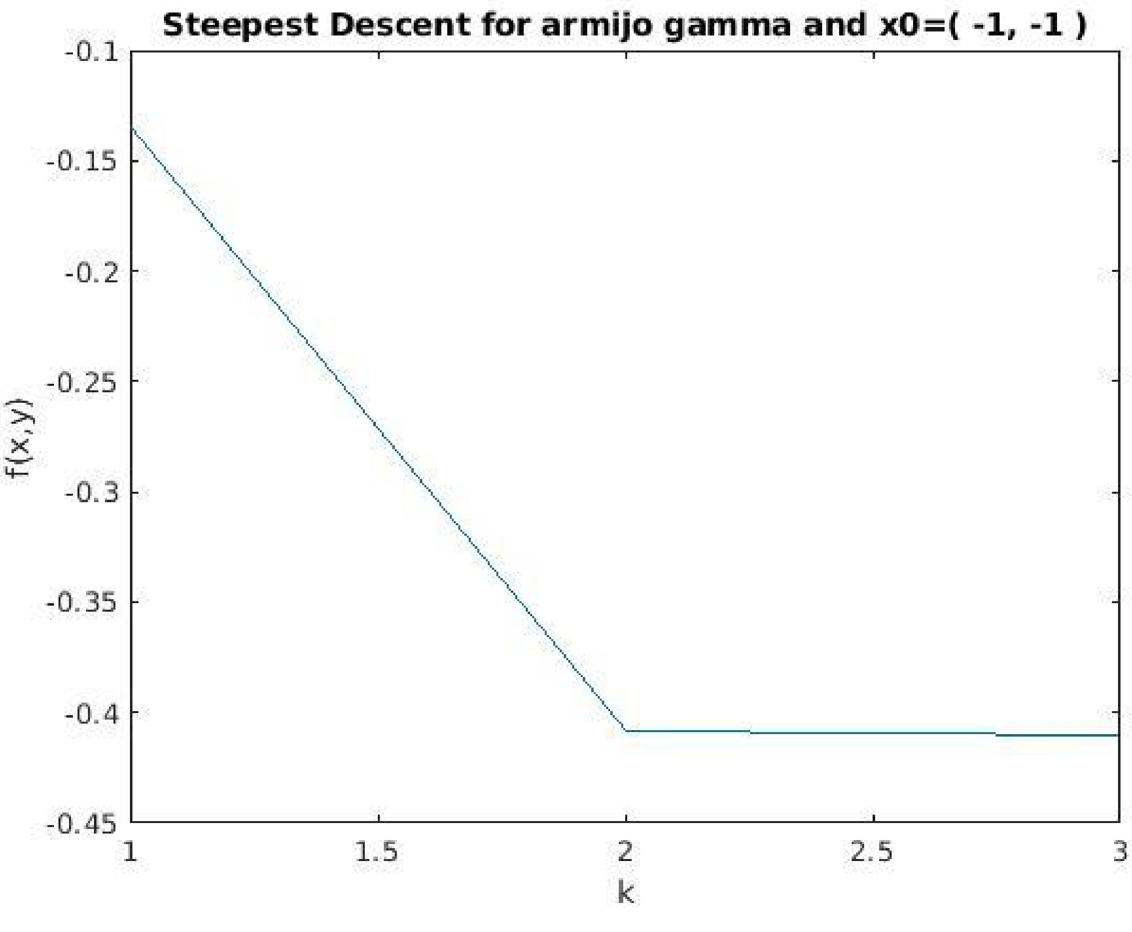


- Για $s = 2$

x0 = (0,0) ----> xk = (0.000000, 0.000000)

x0 = (-1,-1) ----> xk = (-1.224774, 0.081584)

x0 = (1,1) ----> xk = (0.729329, 2.082682)



Θέμα 3 - Μέθοδος Newton

Το αρχείο matlab βρίσκεται στον φάκελο thema3 με τίτλο thema3.m
Η μέθοδος Newton έχει μία βασική προϋπόθεση για να είναι καλώς ορισμένη: Ο εσσιανός πίνακας της συνάρτησης να είναι θετικά ορισμένος. Αυτό βέβαια δεν σημαίνει ότι η μέθοδος αυτή δεν δουλεύει, αλλά ότι δεν μπορεί να προχωρήσει με τα συγκεκριμένα σημεία εκκίνησης.

Θέμα 4 - Μέθοδος Levenberg - Marquardt

Η συγκεκριμένη μέθοδος είναι ο τροποποιημένος αλγόριθμος του Newton. Στην περίπτωση μας, λύνει το πρόβλημα καθώς καταφέρνει να βρει έναν ισοδύναμο πίνακα του εσσιανού ο οποίος να είναι θετικά ορισμένος.

A) Το αρχείο matlab βρίσκεται στον φάκελο thema4 με τίτλο thema4_A.m
Στο Α ζητούμενο καλούμαστε να επιλέξουμε ένα σταθερό βήμα γ_k.

Για τα 3 διαφορετικά αρχικά σημεία x_0 καταλήγουμε στα εξής σημεία x_k :

- Για $\gamma_k = \gamma \ = 0.2$

x0 = (0,0) ----> xk = (0.000000, 0.000000)

x0 = (-1,-1) ----> xk = (-1.224631, -0.083565)

x0 = (1,1) ----> xk = (1.018439, 1.730091)

- Για $\gamma_k = \gamma \ = 0.6$

x0 = (0,0) ----> xk = (0.000000, 0.000000)

x0 = (-1,-1) ----> xk = (-1.224735, 0.071732)

x0 = (1,1) ----> xk = (1.108503, 1.733271)

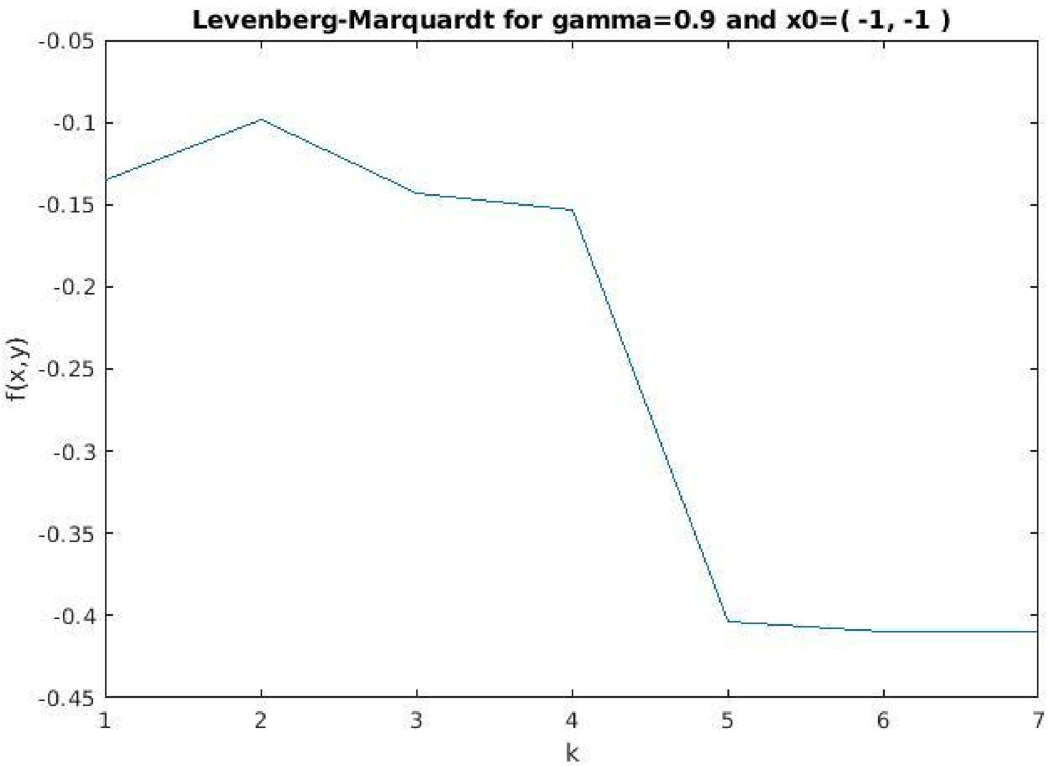
- Για $\gamma_k = \gamma \ = 1$

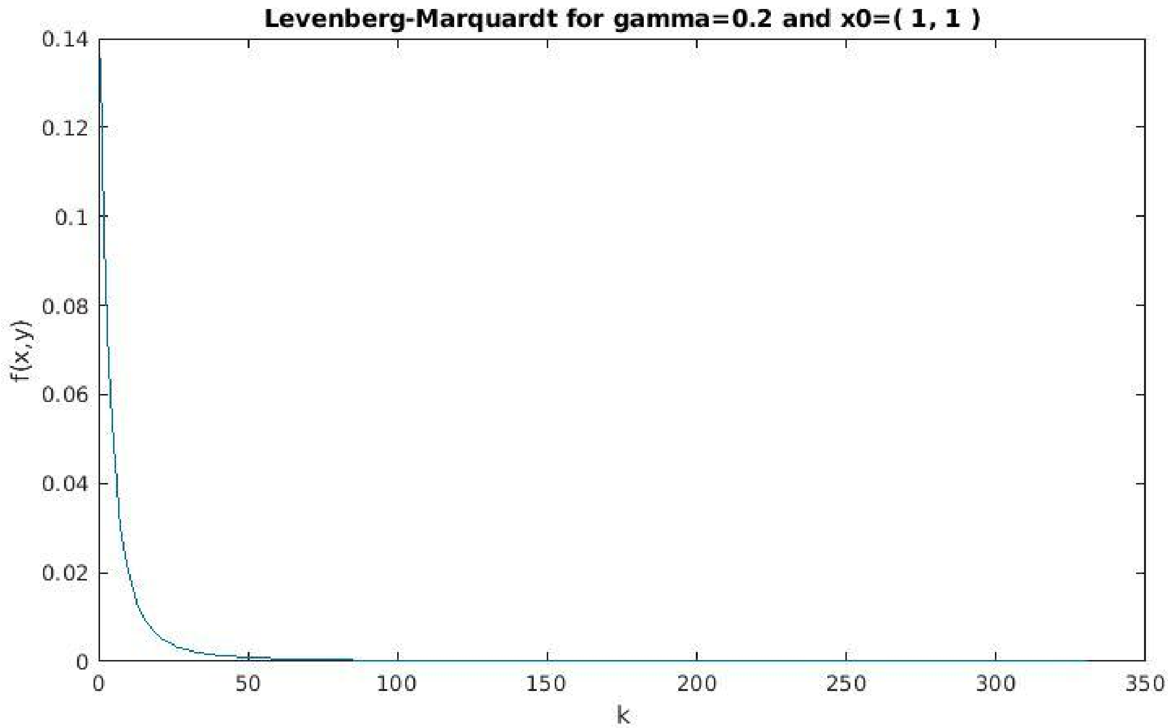
x0 = (0,0) ----> xk = (0.000000, 0.000000)

x0 = (-1,-1) ----> xk = (-0.610870, -1.784068)

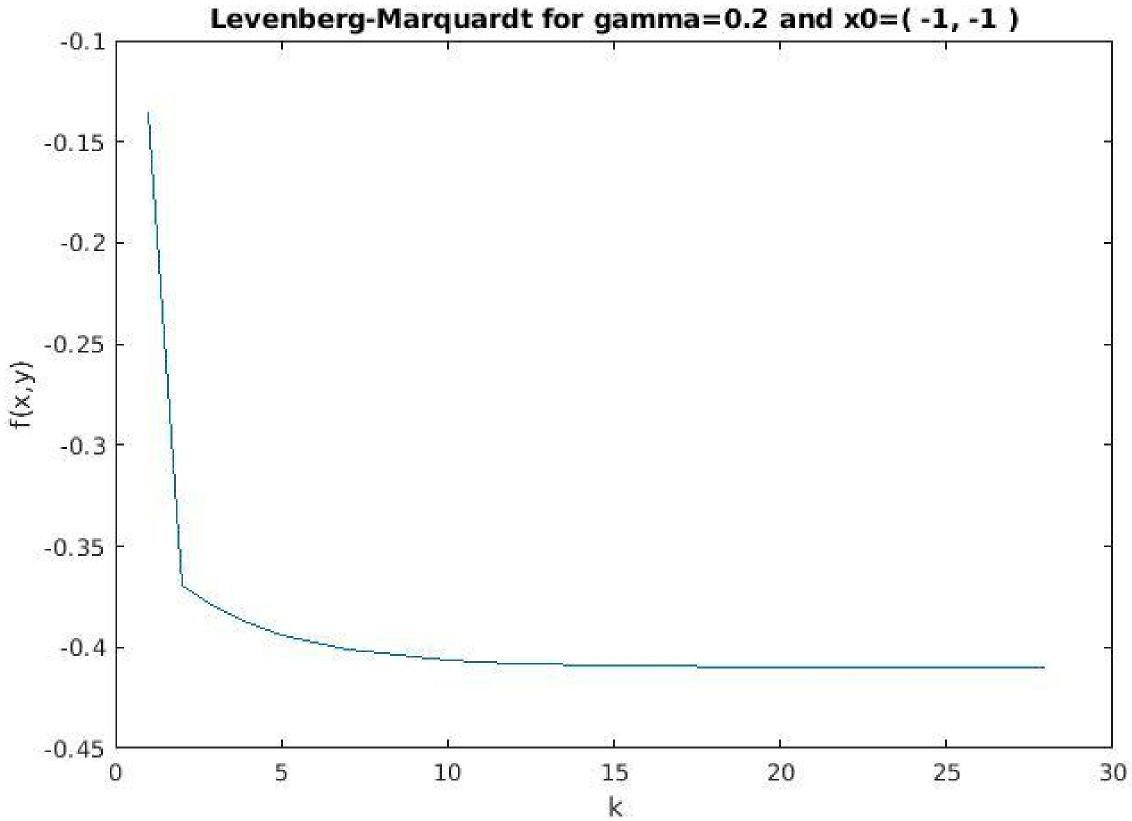
x0 = (1,1) ----> xk = (1.202347, 1.735140)

Παρατηρούμε ότι το γ από μία τιμή και πάνω μας επιστρέφει λάθος αποτελέσματα.
Στην συγκεκριμένη ενώ καταλήγει στο σωστό αποτέλεσμα παρατηρούμε ότι δεν πληροί το κριτήριο 1: $f\ (x_{k+1}) < f\ (x_k)$ και επομένως απορρίπτεται.





Levenberg-Marquardt for gamma=0.2 and x0=(1, 1)
f(x,y)
k
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0
50
100
150
200
250
300
350



Levenberg-Marquardt for gamma=0.2 and x0=(-1, -1)
f(x,y)
k
-0.1
-0.15
-0.2
-0.25
-0.3
-0.35
-0.4
-0.45
0
5
10
15
20
25
30

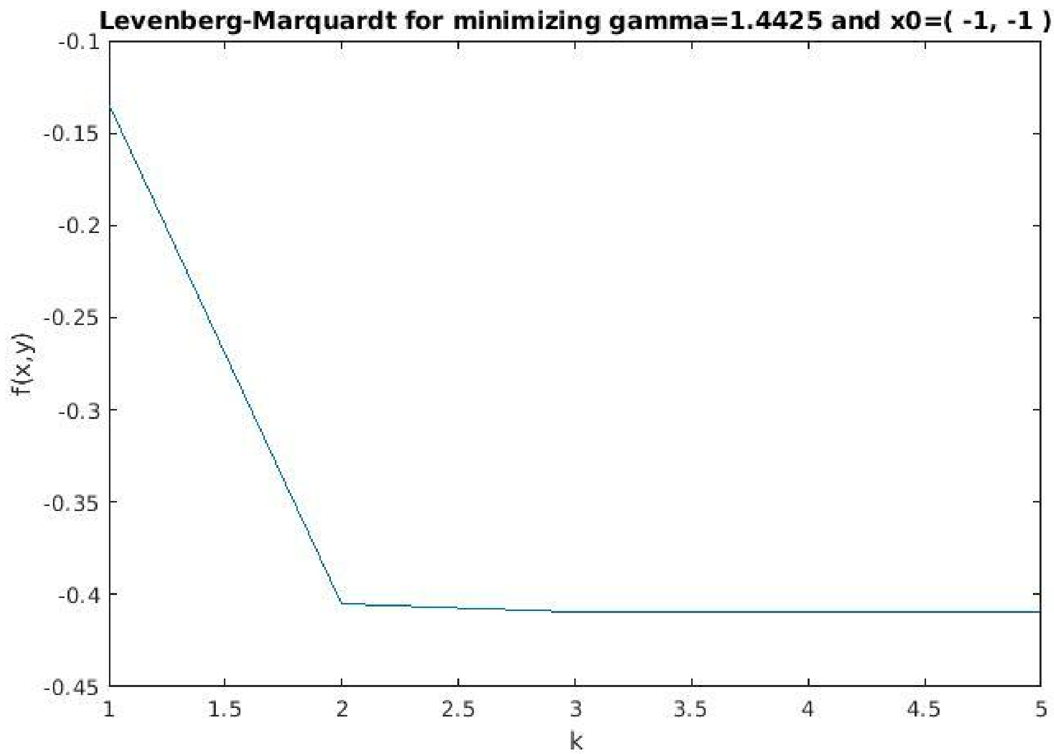
B) Το αρχείο matlab βρίσκεται στον φάκελο thema4 με τίτλο thema4_B.m
Στο B ζητούμενο καλούμαστε να υπολογίσουμε το βήμα γ_k τέτοιο ώστε να ελαχιστοποιεί την $f\ (x_k + \gamma_k d_k)$. Η ελαχιστοποίηση αυτή γίνεται με την βοήθεια της συνάρτησης minimize.m η οποία βρίσκεται στον ίδιο φάκελο και αυτό που κάνει είναι να ελαχιστοποιεί την δοσμένη συνάρτηση ως προς γ_k με την χρήση της μεθόδου της Χρυσής Τομής η οποία αναπτύχθηκε στην προηγούμενη εργασία.

Για τα 3 διαφορετικά αρχικά σημεία x_0 καταλήγουμε στα εξής σημεία x_k :

x0 = (0,0) ----> xk = (0.000000, 0.000000)

x0 = (-1,-1) ----> xk = (-1.225119, -0.026450)

x0 = (1,1) ----> xk = (3.030601, 5.208439)



Παρατηρούμε ότι ο αριθμός τον επαναλήψεων έχει μειωθεί σημαντικά.

C) Το αρχείο matlab βρίσκεται στον φάκελο thema4 με τίτλο thema4_C.m
Στο C ζητούμενο καλούμαστε να υπολογίσουμε το βήμα γ_k βάσει του κανόνα Armijo. Ο υπολογισμός αυτός γίνεται με την βοήθεια της συνάρτησης armijo.m η οποία βρίσκεται στον ίδιο φάκελο. Έχουμε ελευθερία στην επιλογή των τιμών για τις μεταβλητές s, a, β και έπειτα από δοκιμές καταλήξαμε στις:
s = 0.5, a = 0.01 , β = 0.3   και   s = 1, a = 0.01 , β = 0.3
με την δεύτερη όπου s = 1 να έχουμε λιγότερες επαναλήψεις

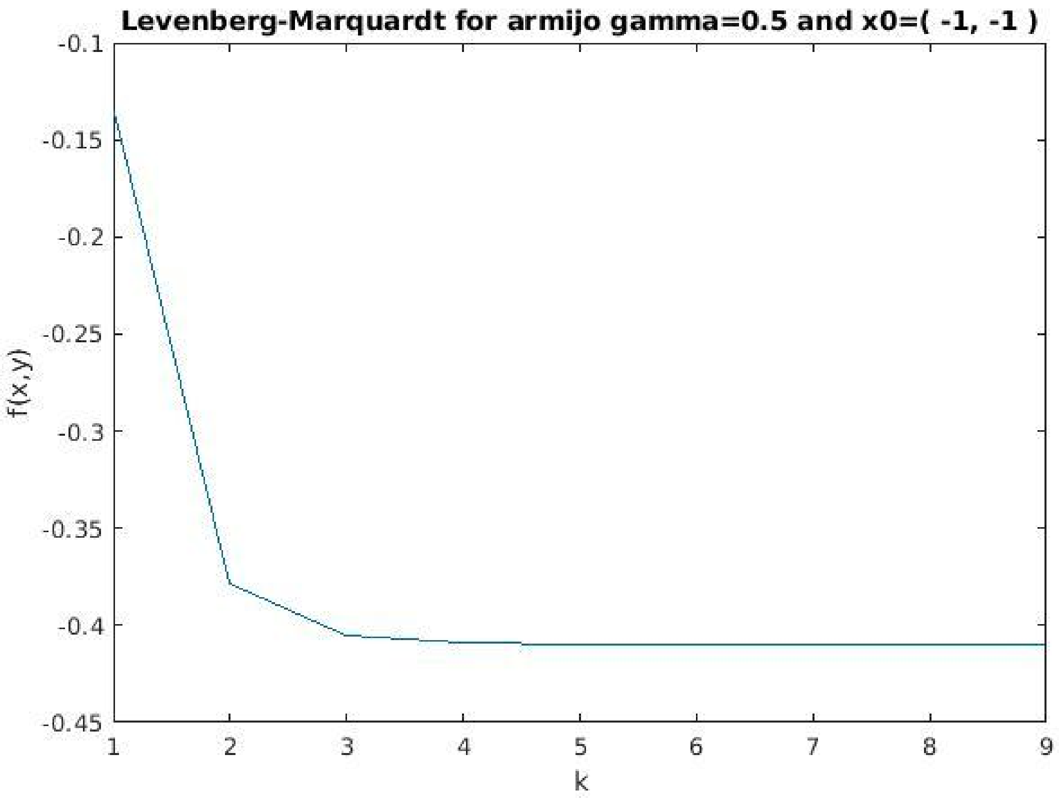
Για τα 3 διαφορετικά αρχικά σημεία x_0 καταλήγουμε στα εξής σημεία x_k :

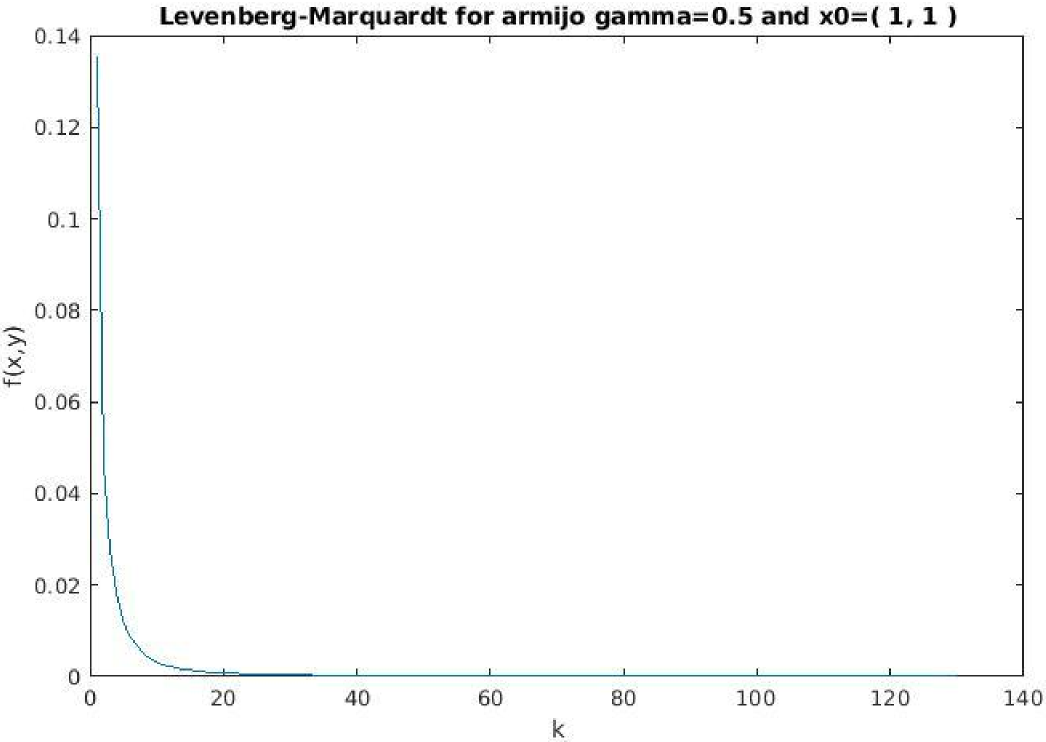
- Για $s = 0.5$

x0 = (0,0) ----> xk = (0.000000, 0.000000)

x0 = (-1,-1) ----> xk = (-1.225269, 0.021420)

x0 = (1,1) ----> xk = (1.084343, 1.732865)



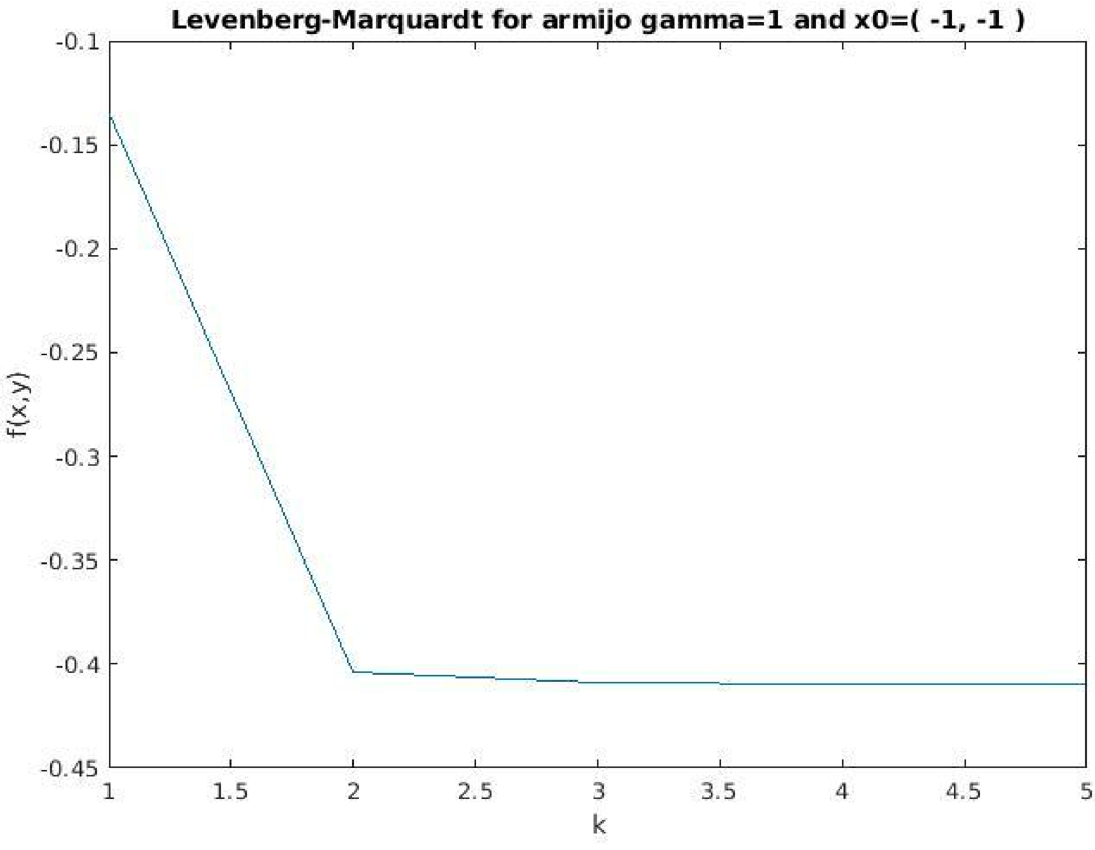


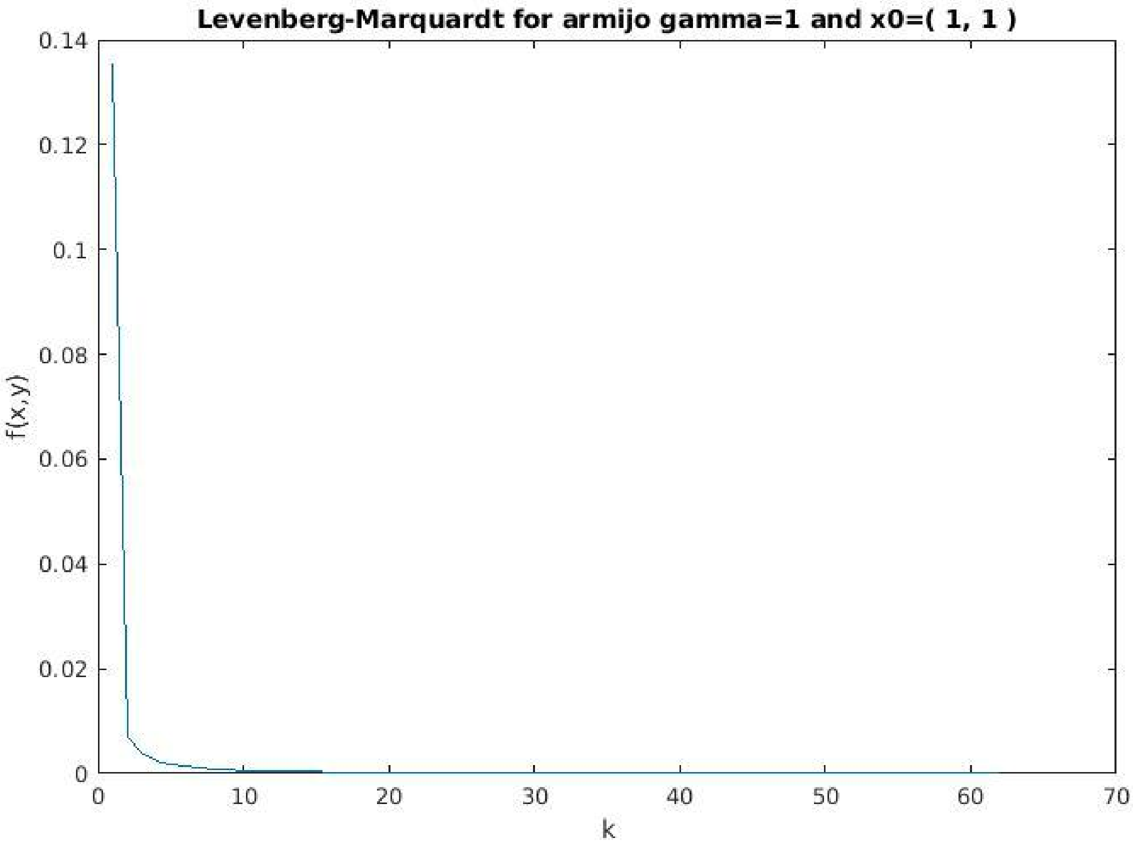
- Για $s = 1$

x0 = (0,0) ----> xk = (0.000000, 0.000000)

x0 = (-1,-1) ----> xk = (-1.224745, -0.066658)

x0 = (1,1) ----> xk = (1.202347, 1.735140)





Levenberg-Marquardt for armijo gamma=1 and x0=(1, 1)
f(x,y)
k
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
10
20
30
40
50
60
70