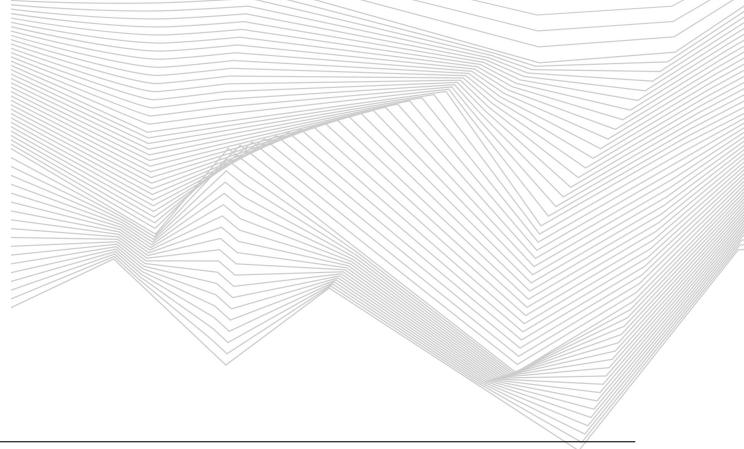


Ελαχιστοποίηση με χρήση παραγώγων

ΤΕΧΝΙΚΕΣ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ 2Η ΕΡΓΑΣΤΗΡΙΑΚΗ ΑΣΚΗΣΗ



7 Δεκέμβρη 2021

Περιληπτικά:

Στην εργασία αυτή ζητούμενο είναι η ελαχιστοποίηση μιας δοσμένης συνάρτησης πολλών μεταβλητών χωρίς περιορισμούς. Οι αλγόριθμοι που θα χρησιμοποιήσουμε βασίζονται στην ιδέα της επαναληπτικής καθόδου και είναι οι εξής :

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

Σχεδίαση της f

Αρχικά, είναι χρήσιμο να σχεδιάσουμε τη γραφική παράσταση της δοσμένης συνάρτησης $f(x,y)$ προκειμένου να έχουμε μία εποπτεία του προβλήματος της ελαχιστοποίησής της, και έτσι να ξέρουμε τι περίπου να περιμένουμε. Έτσι, στα παρακάτω διαγράμματα παρατηρούμε πως η συνάρτηση είναι σχετικά επίπεδη με εξαίρεση δύο περιοχές. Είναι προφανές ότι στη μία εξ' αυτών υπάρχει το μέγιστο και στην άλλη το ελάχιστο της, με τιμή περίπου στο -0.4. Αυτή η δεύτερη περιοχή, περιμένουμε να είναι και η κύρια περιοχή αναζήτησής μας.

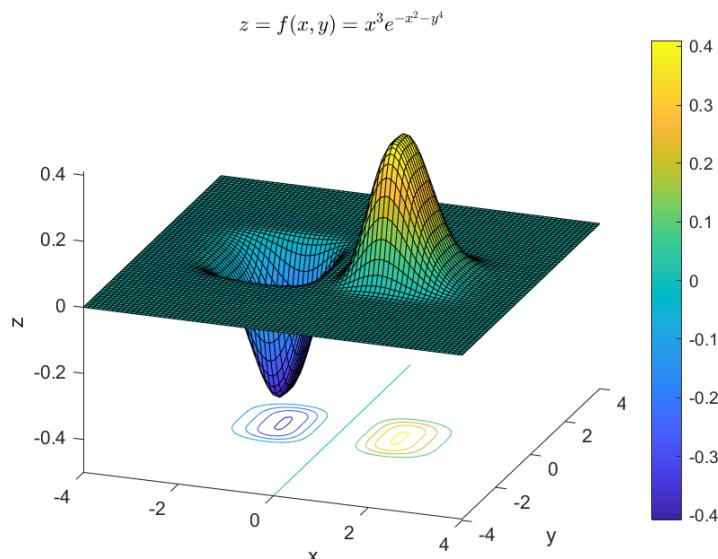


Figure 1: Γραφική παράσταση της $f(x,y)$

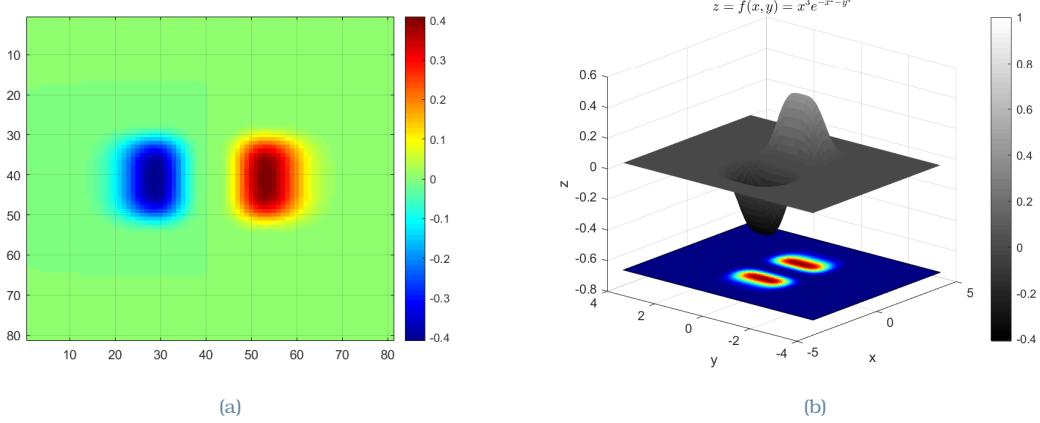


Figure 2: Επιπλέον γραφικές παραστάσεις

Επιλογή αρχικών παραμέτρων και διαδικασία

Πριν δούμε τη συμπεριφορά των διαφόρων μεθόδων, πρέπει να καταγράψουμε τις παραμέτρους που χρησιμοποιήθηκαν στις δοκιμές. Οι διάφορες αντίστοιχες παράμετροι κρατήθηκαν ίδιες σε όλες τις παρακάτω δοκιμές. Η ακρίβεια ε θα είναι στο εξής ίση με 10^{-5} . Όταν το γ θα είναι σταθερό, θα είναι ίσο με 1. Η μέθοδος της Χρυσής τομής (που χρησιμοποιήθηκε για την ελαχιστοποίηση της $f(y)$), θα ψάχνει το ελάχιστο στο διάστημα $(0, 100]$ (ώστε το γ να είναι θετικό) με ακρίβεια $\epsilon = 10^{-3}$, δίνοντας ως αποτέλεσμα το αριστερό άκρο του τελικού διαστήματος. Ο κανόνας του Armijo θα καλείται με $\alpha = 0.1$, $\beta = 0.5$ και $s = 50$. Σεκινώντας από το σημείο $(-1, -1)$, θα προσπαθήσουμε να ελαχιστοποιήσουμε τη συνάρτηση αρχικά με σταθερό γ, έπειτα θα είναι τέτοιο ώστε να ελαχιστοποιεί την $f(x_k + \gamma d_k)$ και, τέλος, θα ικανοποιεί τον κανόνα Armijo. Θα συνεχίσουμε όμοια στα σημεία $(0, 0)$ και $(1, 1)$.

Για κάθε μέθοδο, για να τρέξει, πρέπει να κάνετε uncomment το αντίστοιχο σημείο στο mainScript.m, τα αντίστοιχα σχόλια θα σας καθοδηγήσουν.

1. Μέθοδος Μέγιστης Καθόδου - Steepest Descent

Στην ενότητα αυτή θα παρουσιάσουμε τα διαγράμματα που αφορούν στη μέθοδο μέγιστης καθόδου.

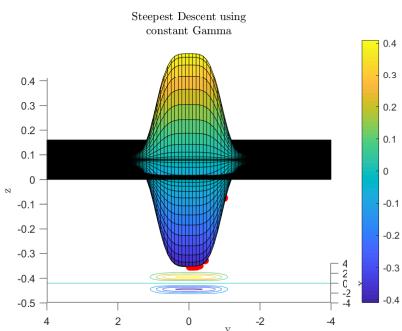


Figure 3: Διάγραμμα επιφάνειας για εκκίνηση από το $(-1, -1)$ και σταθερό γ

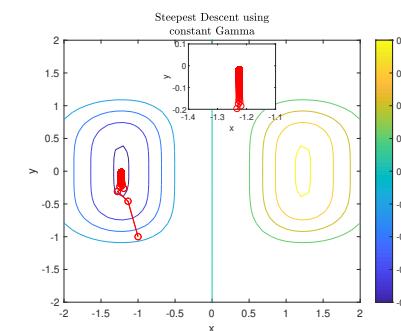


Figure 4: Ισοθαρείς καμπύλες για εκκίνηση από το $(-1, -1)$ και σταθερό γ

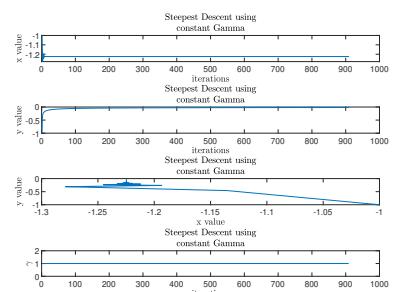


Figure 5: Αλλαγές στις παραμέτρους για εκκίνηση από το $(-1, -1)$ και σταθερό γ

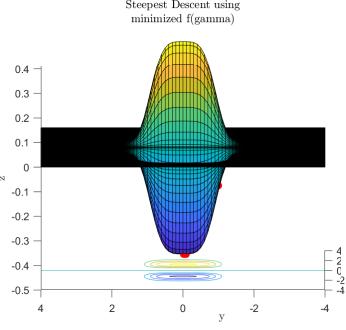


Figure 6: Διάγραμμα επιφάνειας για εκκίνηση από το $(-1, -1)$ και ελαχιστοποίηση της $f(\gamma)$

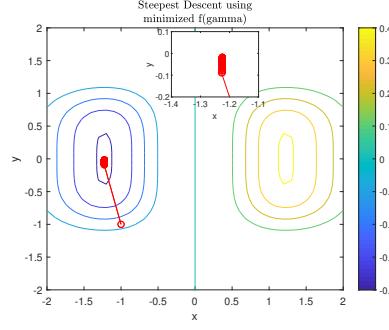


Figure 7: Ισοβαρείς καμπύλες για εκκίνηση από το $(-1, -1)$ και ελαχιστοποίηση της $f(\gamma)$

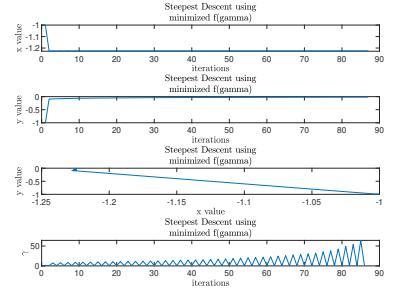


Figure 8: Αλλαγές στις παραμέτρους για εκκίνηση από το $(-1, -1)$ και ελαχιστοποίηση της $f(\gamma)$

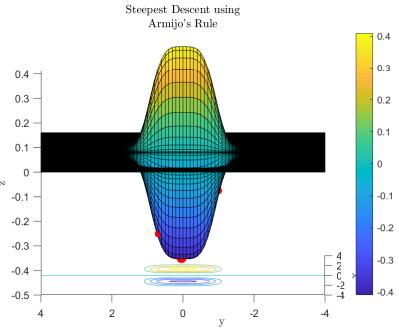


Figure 9: Διάγραμμα επιφάνειας για εκκίνηση από το $(-1, -1)$ και κανόνα Armijo

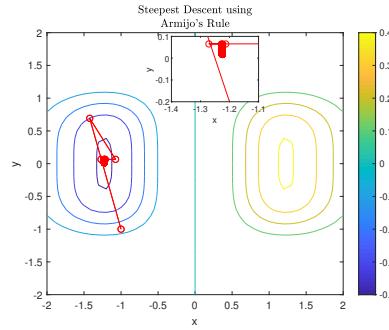


Figure 10: Ισοβαρείς καμπύλες για εκκίνηση από το $(-1, -1)$ και κανόνα Armijo

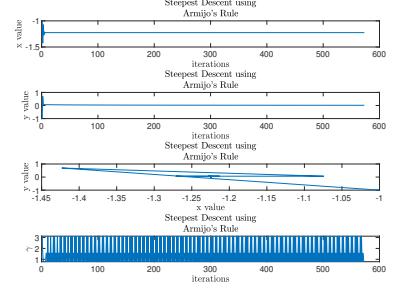
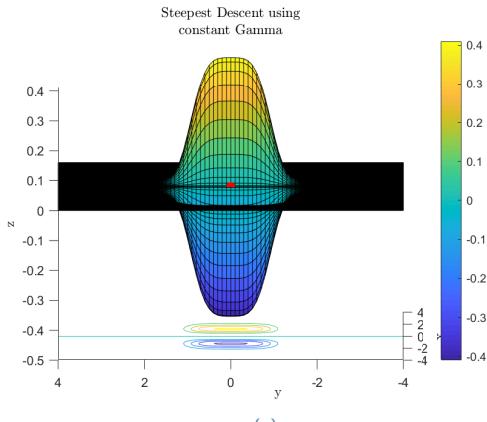
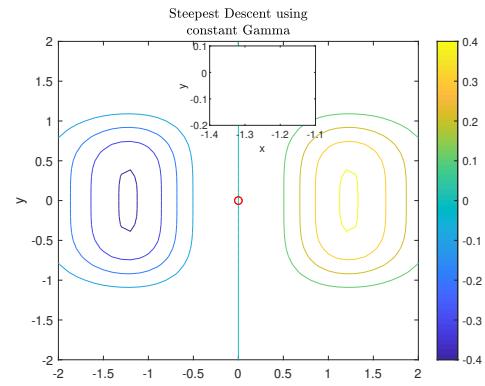


Figure 11: Αλλαγές στις παραμέτρους για εκκίνηση από το $(-1, -1)$ και κανόνα Armijo

Ξεκινώντας από το σημείο $(0, 0)$ η κλίση είναι μηδενική, ο αλγόριθμος σταματάει πριν καν ξεκινήσει.



(a)



(b)

Figure 12: Εκκίνηση από το $(0,0)$ και σταθερό γ

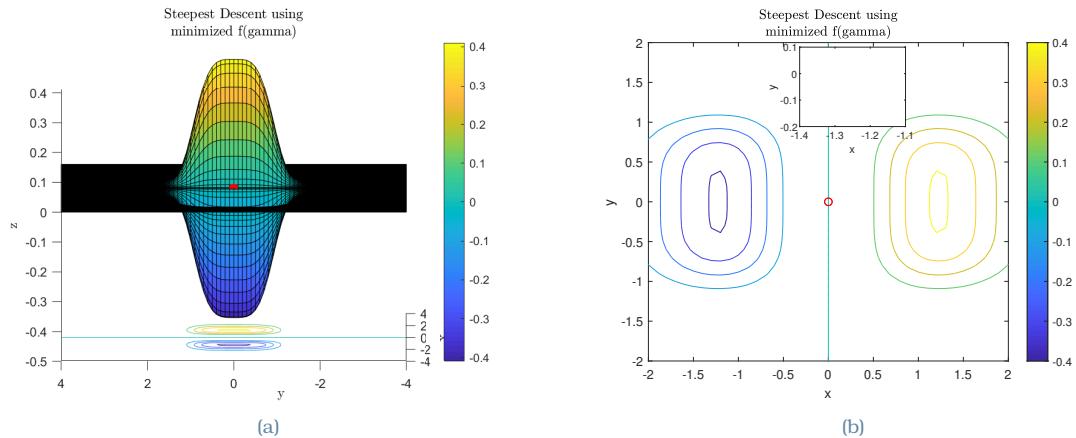


Figure 13: Εκκίνηση από το $(0,0)$ και ελαχιστοποίηση της $f(y)$

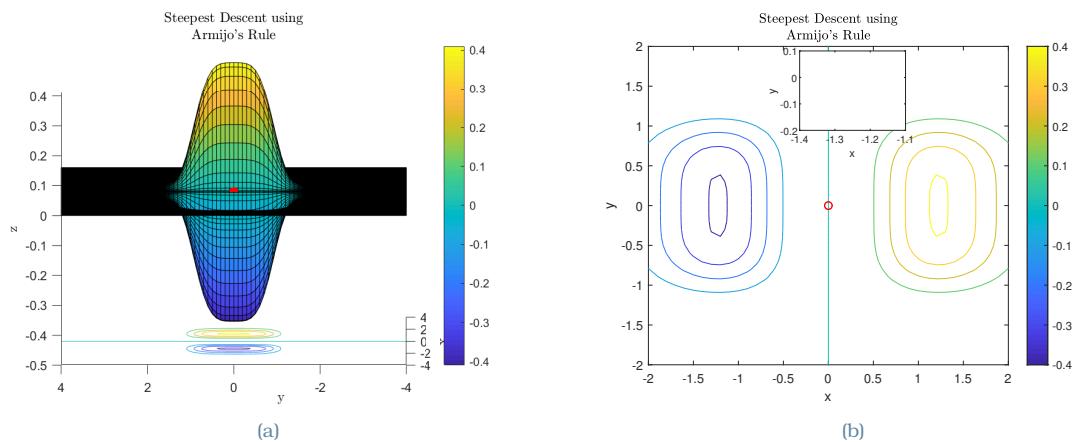


Figure 14: Εκκίνηση από το $(0,0)$ και κανόνα Armijo

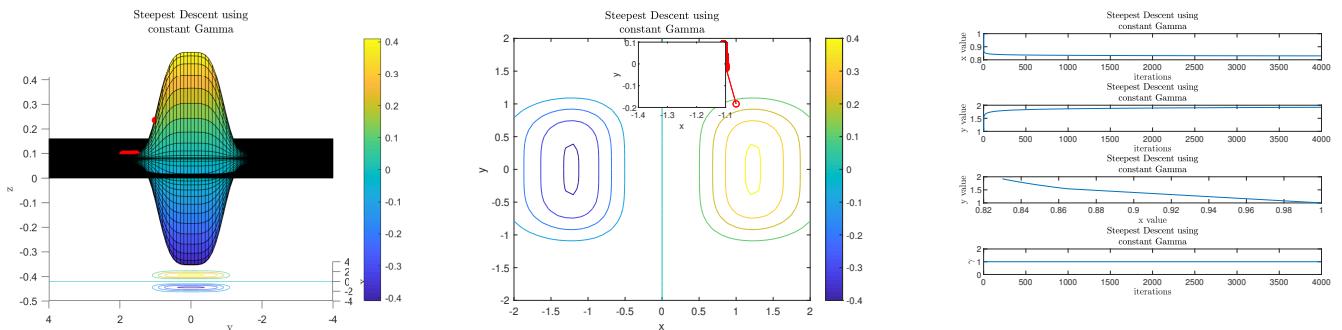


Figure 15: Διάγραμμα επιφάνειας για εκκίνηση από το $(1, 1)$ και σταθερό γ

Figure 16: Ισοβαρείς καμπύλες για εκκίνηση από το $(1, 1)$ και σταθερό γ

Figure 17: Αλλαγές στις παραμέτρους για εκκίνηση από το $(1, 1)$ και σταθερό γ

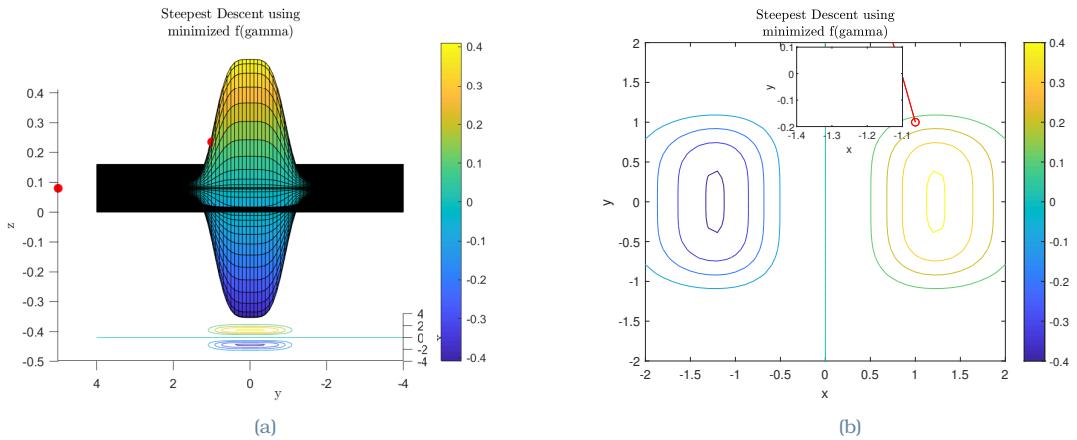


Figure 18: Εκκίνηση από το $(1, 1)$ και ελαχιστοποίηση της $f(y)$

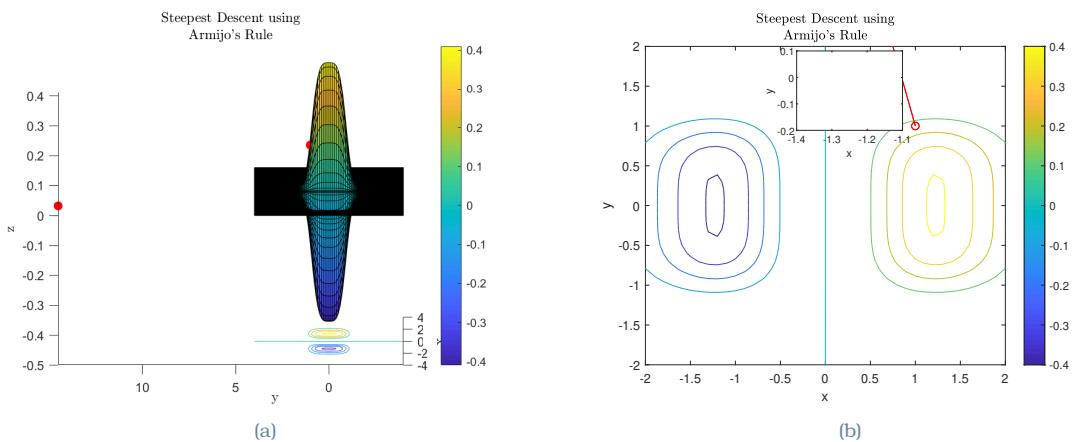


Figure 19: Εκκίνηση από το $(1, 1)$ και κανόνα Armijo

2. Μέθοδος Levenberg-Marquardt

Στην ενότητα αυτή θα παρουσιάσουμε τα διαγράμματα που αφορούν στη μέθοδο Levenberg-Marquardt.

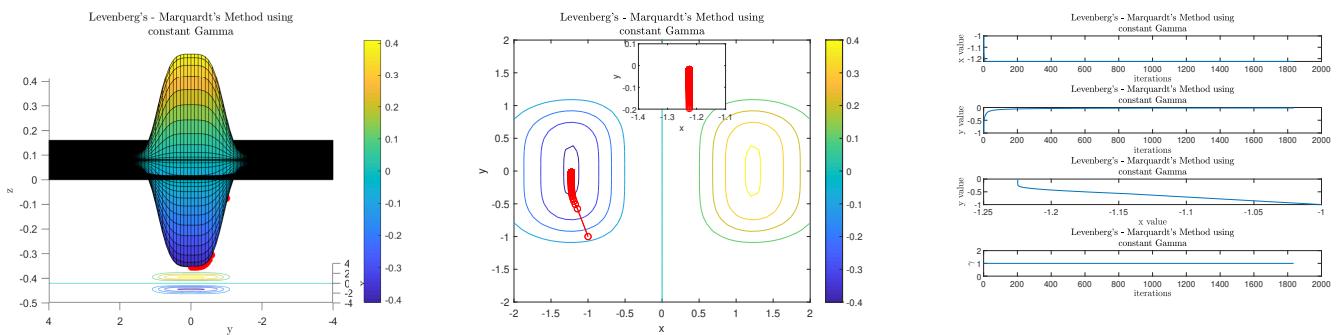


Figure 20: Διάγραμμα επιφάνειας για εκκίνηση από το $(-1, -1)$ και σταθερό γ

Figure 21: Ισοβαρείς καμπύλες για εκκίνηση από το $(-1, -1)$ και σταθερό γ

Figure 22: Άλλαγές στις παραμέτρους για εκκίνηση από το $(-1, -1)$ και σταθερό γ

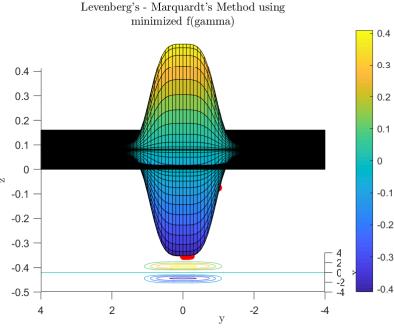


Figure 23: Διάγραμμα επιφάνειας για εκκίνηση από το $(-1, -1)$ και ελαχιστοποίηση της $f(\gamma)$

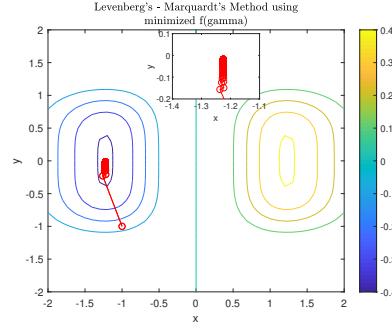


Figure 24: Ισοβαρείς καμπύλες για εκκίνηση από το $(-1, -1)$ και ελαχιστοποίηση της $f(\gamma)$

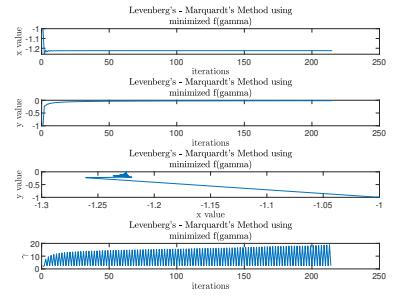


Figure 25: Αλλαγές στις παραμέτρους για εκκίνηση από το $(-1, -1)$ και ελαχιστοποίηση της $f(\gamma)$

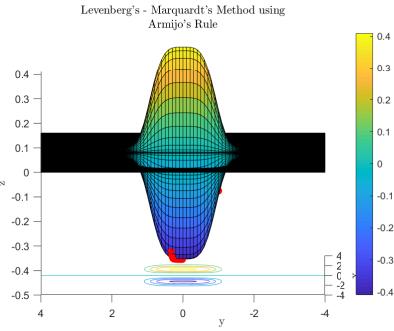


Figure 26: Διάγραμμα επιφάνειας για εκκίνηση από το $(-1, -1)$ και κανόνα Armijo

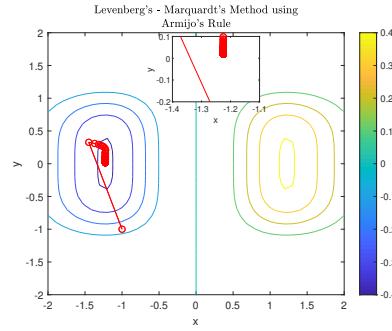


Figure 27: Ισοβαρείς καμπύλες για εκκίνηση από το $(-1, -1)$ και κανόνα Armijo

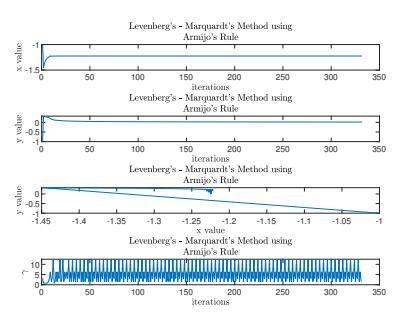


Figure 28: Αλλαγές στις παραμέτρους για εκκίνηση από το $(-1, -1)$ και κανόνα Armijo

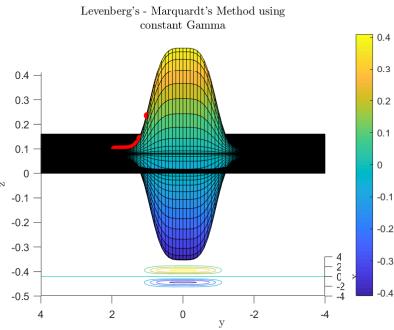


Figure 29: Διάγραμμα επιφάνειας για εκκίνηση από το $(1, 1)$ και σταθερό γ

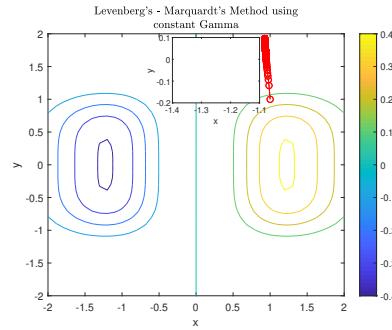


Figure 30: Ισοβαρείς καμπύλες για εκκίνηση από το $(1, 1)$ και σταθερό γ

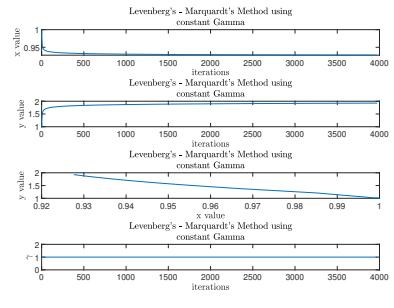


Figure 31: Αλλαγές στις παραμέτρους για εκκίνηση από το $(1, 1)$ και σταθερό γ

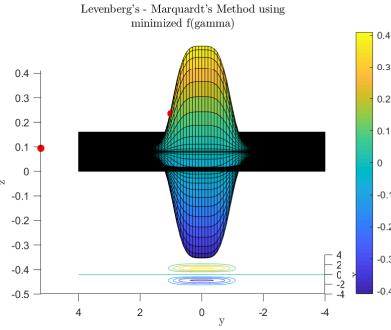


Figure 32: Διάγραμμα επιφάνειας για εκκίνηση από το $(1,1)$ και ελαχιστοποίηση της $f(\gamma)$

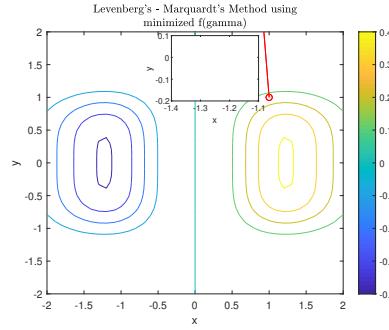


Figure 33: Ισοβαρείς καμπύλες για εκκίνηση από το $(1,1)$ και ελαχιστοποίηση της $f(\gamma)$

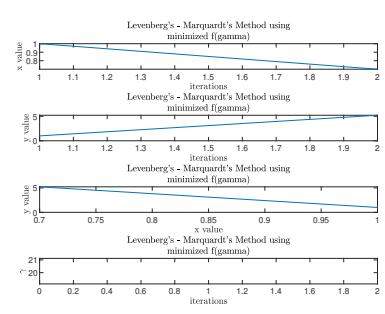


Figure 34: Αλλαγές στις παραμέτρους για εκκίνηση από το $(1,1)$ και ελαχιστοποίηση της $f(\gamma)$

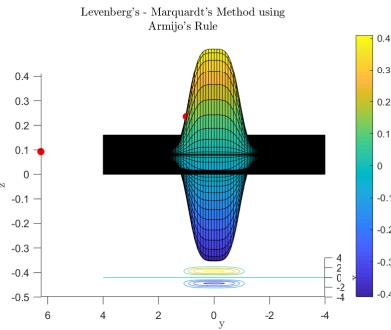


Figure 35: Διάγραμμα επιφάνειας για εκκίνηση από το $(1,1)$ και κανόνα Armijo

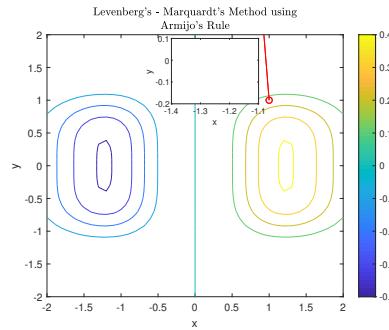


Figure 36: Ισοβαρείς καμπύλες για εκκίνηση από το $(1,1)$ και κανόνα Armijo

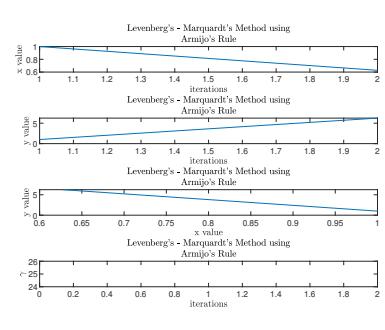


Figure 37: Αλλαγές στις παραμέτρους για εκκίνηση από το $(1,1)$ και κανόνα Armijo

3. Μέθοδος Νεύτωνα - Newton's Method

Στην ενότητα αυτή θα παρουσιάσουμε τα διαγράμματα που αφορούν στη μέθοδο Newton.

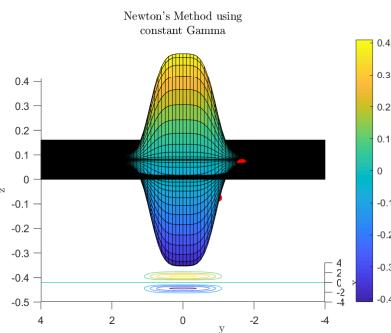


Figure 38: Διάγραμμα επιφάνειας για εκκίνηση από το $(-1,-1)$ και σταθερό γ

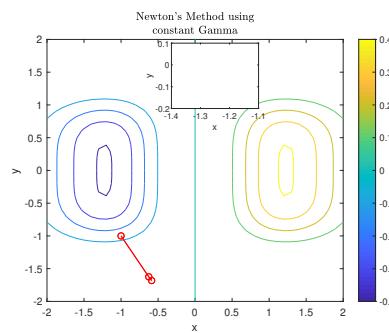


Figure 39: Ισοβαρείς καμπύλες για εκκίνηση από το $(-1,-1)$ και σταθερό γ

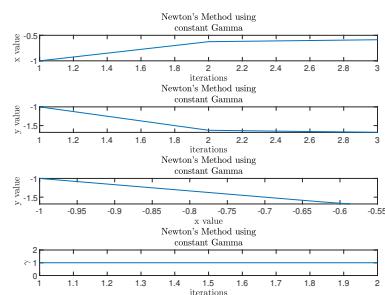


Figure 40: Αλλαγές στις παραμέτρους για εκκίνηση από το $(-1,-1)$ και σταθερό γ

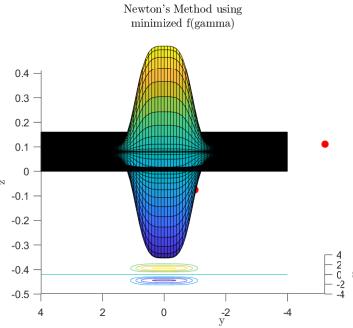


Figure 41: Διάγραμμα επιφάνειας για εκκίνηση από το $(-1, -1)$ και ελαχιστοποίηση της $f(y)$

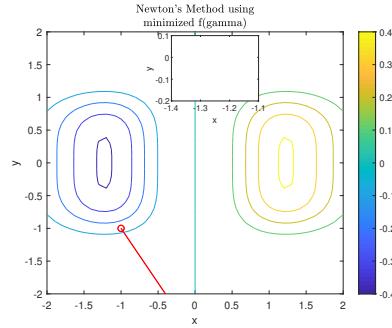


Figure 42: Ισοβαρείς καμπύλες για εκκίνηση από το $(-1, -1)$ και ελαχιστοποίηση της $f(y)$

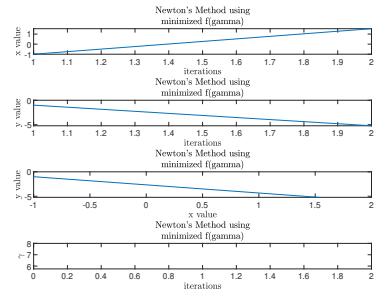


Figure 43: Αλλαγές στις παραμέτρους για εκκίνηση από το $(-1, -1)$ και ελαχιστοποίηση της $f(y)$

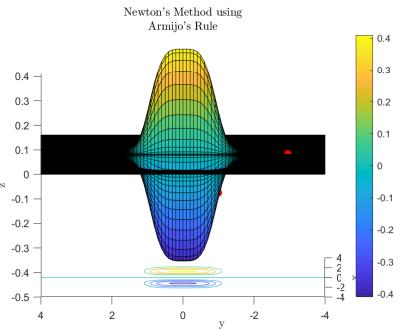


Figure 44: Διάγραμμα επιφάνειας για εκκίνηση από το $(-1, -1)$ και κανόνα Armijo

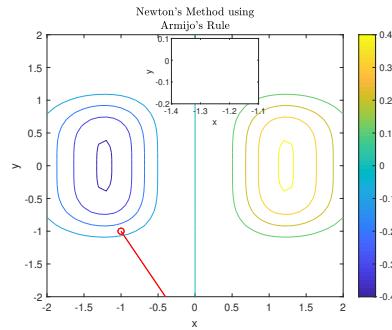


Figure 45: Ισοβαρείς καμπύλες για εκκίνηση από το $(-1, -1)$ και κανόνα Armijo

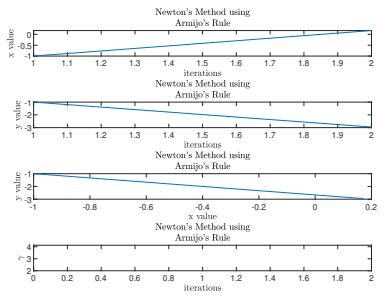


Figure 46: Αλλαγές στις παραμέτρους για εκκίνηση από το $(-1, -1)$ και κανόνα Armijo

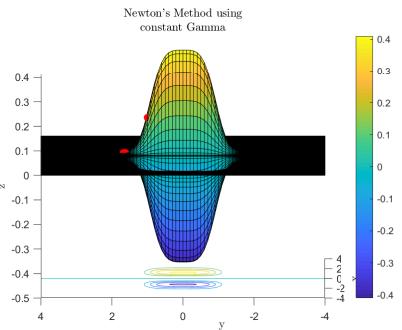


Figure 47: Διάγραμμα επιφάνειας για εκκίνηση από το $(1, 1)$ και σταθερό γ

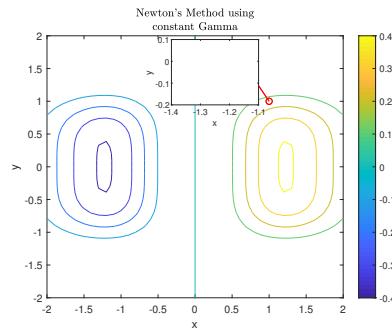


Figure 48: Ισοβαρείς καμπύλες για εκκίνηση από το $(1, 1)$ και σταθερό γ

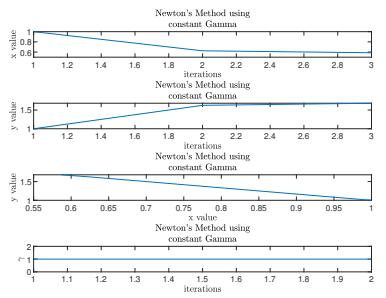


Figure 49: Αλλαγές στις παραμέτρους για εκκίνηση από το $(1, 1)$ και σταθερό γ

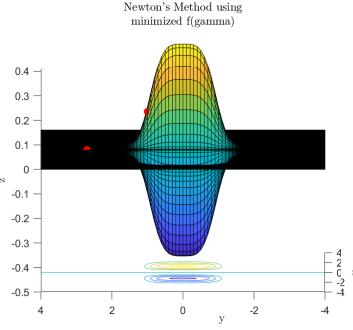


Figure 50: Διάγραμμα επιφάνειας για εκκίνηση από το $(1, 1)$ και ελαχιστοποίηση της $f(y)$

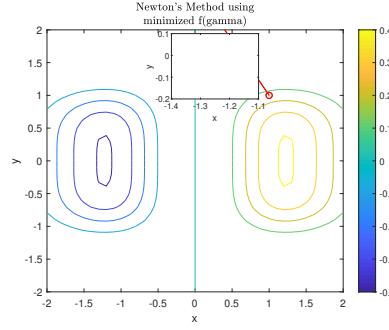


Figure 51: Ισοβαρείς καμπύλες για εκκίνηση από το $(1, 1)$ και ελαχιστοποίηση της $f(y)$

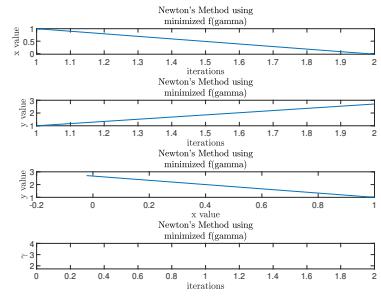


Figure 52: Αλλαγές στις παραμέτρους για εκκίνηση από το $(1, 1)$ και ελαχιστοποίηση της $f(y)$

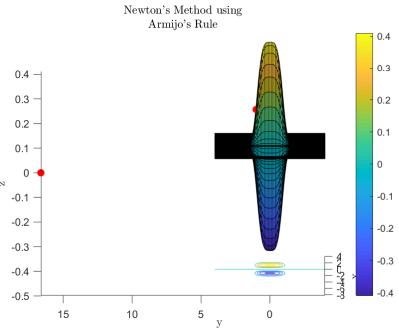


Figure 53: Διάγραμμα επιφάνειας για εκκίνηση από το $(1, 1)$ και κανόνα Armijo

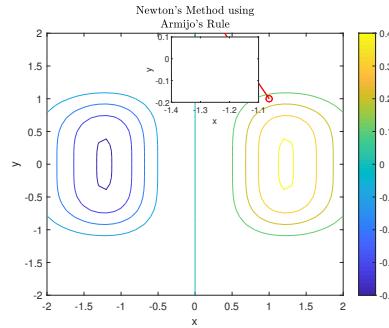


Figure 54: Ισοβαρείς καμπύλες για εκκίνηση από το $(1, 1)$ και κανόνα Armijo

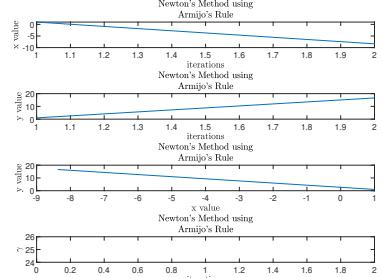


Figure 55: Αλλαγές στις παραμέτρους για εκκίνηση από το $(1, 1)$ και κανόνα Armijo

4. Παρατηρήσεις και Συμπεράσματα

Παρατηρούμε εύκολα πως όταν το σημείο εκκίνησης είναι $(0, 0)$, όπου η κλίση της συνάρτησης είναι μηδενική, η συνθήκη τερματισμού του αλγορίθμου ικανοποιείται εξ' αρχής. Άρα, η μέθοδος ολοκληρώνεται πριν καν ξεκινήσει.

Από της άλλη, όταν σημείο εκκίνησης είναι το $(1, 1)$, η μέθοδος εγκλωβίζεται και δε θα περάσει στα αρνητικά x καθώς η κλίση θα μηδενιστεί πριν φτάσει στην περιοχή του ελαχίστου. Ωστόσο, αν επιλεγεί αρκούντως μεγάλο γ ώστε να περάσει την περιοχή με τη μηδενική κλίση, θα μπορέσει τελικά να βρει το ελάχιστο.

Αναφορικά με τη μέθοδο Newton, παρατηρούμε ότι δεν ικανοποιείται η βασική προϋπόθεση να είναι ο Εσσιανός πίνακας θετικά ορισμένος. Και έτοι, δεν περιμένουμε να οδηγηθούμε σε εύρεση του ελαχίστου. Ενδιαφέρουσα παρατήρηση μετά από προσομοίωση αποτελεί το γεγονός ότι αν καταφέρουμε να είναι ο Εσσιανός πίνακας θετικά ορισμένος, μπορούμε να βρούμε τελικά το ελάχιστο. Αυτό μπορεί να γίνει αν αλλάζουμε το πρόστιμο του γ όταν ο Εσσιανός είναι αρνητικά ορισμένος, και έτοι να κάνουμε τον Εσσιανό θετικά ορισμένο. Στην πράξη αυτό μπορεί να γίνει αν στην κλήση της μεθόδου της Χρυσής Τομής για την ελαχιστοποίηση της f ως προς γ , δεν περιορίσουμε το γ να είναι θετικό.

Επιπλέον, μπορούμε να οδηγηθούμε στο ελάχιστο αν ξεκινήσουμε από ένα σημείο όπου ο Εσσιανός είναι θετικά ορισμένος. Μερικά τέτοια σημεία φαίνονται παρακάτω, όπου με κόκκινο κυκλάκι φαίνεται το σημείο $(-1, -1)$ που αποτελεί και το σημείο εκκίνησης μας.

ακρίβεια ϵ	Steepest descent	L-M
10^{-3}	40	172
10^{-5}	909	3653
10^{-7}	19440	8000

Table 1: σταθερό γ

ακρίβεια ϵ	Steepest descent	L-M
10^{-3}	4	12
10^{-5}	86	194
10^{-7}	154	2000

Table 2: επιλογή του γ που ελαχιστοποιεί την συνάρτηση $f(y)$

Από τους 3 πίνακες πιο αποτελεσματικός τρόπος επιλογής του γ είναι μέσω της ελαχιστοποίησης της f ως προς γ , ακολουθούμενη από

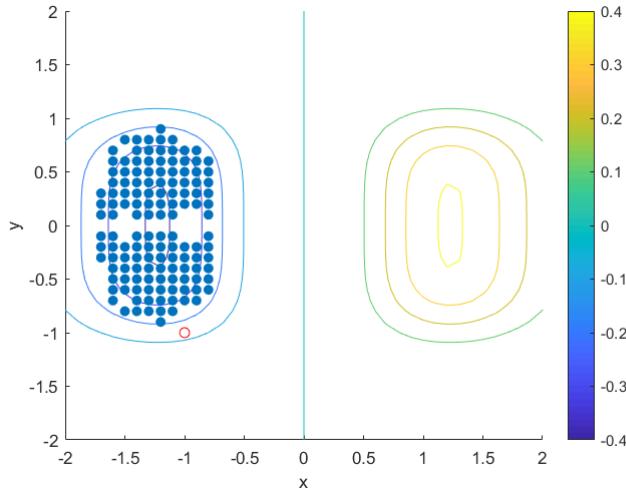


Figure 56: Σημεία εκκίνησης για θετικά ορισμένο Εσσιανό με σταθερό γ

ακρίβεια ϵ	Steepest descent	L-M
10^{-3}	7	26
10^{-5}	573	443
10^{-7}	12530	8051

Table 3: επιλογή γ με τη μέθοδο Armijo

των κανόνα του Armijo. Ποια θα είναι η πιο αποτελεσματική μέθοδος ως προς τις επαναλήψεις (για σημείο εκκίνησης το $(-1, -1)$) θα εξαρτηθεί από τις αρχικές παραμέτρους που επιλέξαμε. Από τα δεδομένα του πίνακα φαίνεται πως υπερέχει η Steepest descent όσο έχουμε μικρά ϵ (για $\epsilon = 10^{-3}$ και $\epsilon = 10^{-5}$) με εξαίρεση των κανόνα Armijo που η διαφορά τους είναι συγκρίσιμη, αντιθέτως η L-M υπερισχύει για $\epsilon = 10^{-7}$, με εξαίρεση την επιλογή του γ που ελαχιστοποιεί την f .

Γ' αυτό θα πρέπει σε κάθε περίπτωση να επιλέγονται οι παράμετροι και να ρυθμίζονται συγκεκριμένα για ώστε να έχουμε το βέλτιστο αποτέλεσμα. Παραδείγματος χάρη, αυξάνοντας την τιμή της σταθεράς η_k στη μέθοδο L-M μπορεί πιθανόν να μιας αυξήσει την ταχύτητα σύγκλισης, ωστόσο μεγάλη αύξηση της σταθεράς αυτής μπορεί να οδηγήσει σε ταλαντώσεις και άρα μείωση της ταχύτητα σύγκλισης.