Τεχνικές Βελτιστοποίησης, 7ο εξάμηνο

Ινώ Χαριτίδη, 10043

<u>2η Εργαστηριακή Άσκηση</u>

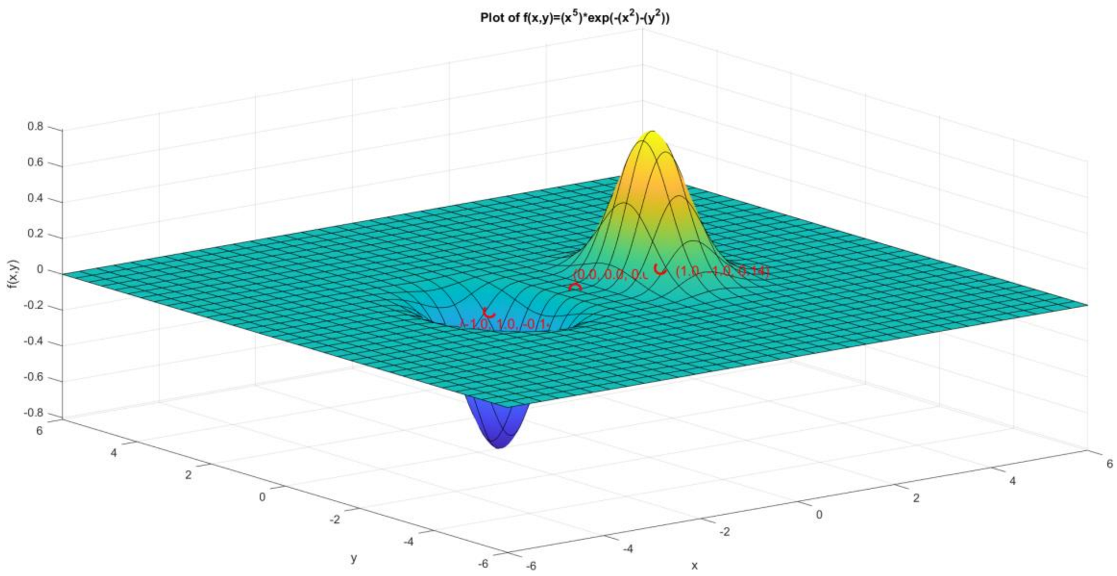
Αρχικά βρίσκω το σημείο ελαχίστου της f χρησιμοποιώντας το εργαλείο Wolfram Alpha: min=(-sqrt(5/2), 0), περίπου (-1.5811, 0).

Η ελάχιστη τιμή της f είναι περίπου -0.811174.

Η λύση αυτή θα χρησιμοποιηθεί για την επαλήθευση των αποτελεσμάτων μου.

Οι αλγόριθμοι λειτουργούν με ακρίβεια 0.01

<u>Θέμα 1</u>: Σχεδιάστε την f(x,y)=(x^5)*exp(-(x^2)-(y^2)) , για να πάρετε μια γενική εικόνα της μορφής της.



Γράφημα της f. Σημειωμένα επάνω είναι τα σημεία εκκίνησης (x,y)=(-1,1), (0,0) και (1,-1).

<u>Θέμα 2</u>: Ελαχιστοποιείστε την f με την μέθοδο μέγιστης καθόδου, χρησιμοποιώντας ως αρχικά σημεία (x0 , y0) τα

i) (0,0)

ii) (−1,1)

iii) (1,-1)

Το βήμα γ_k θα επιλεγεί:

α) σταθερό (της επιλογής σας)

β) τέτοιο ώστε να ελαχιστοποιεί την f (x_k+$γ_k$ d_k)

γ) βάσει του κανόνα Armijo.

Σχολιάστε τις διαφορές στα αποτελέσματα, σε περίπτωση που προκαλούνται, λόγω της επιλογής του σημείου έναρξης (x_0, y_0) του αλγορίθμου, καθώς επίσης και λόγω της επιλογής του βήματος $γ_κ$. Οδηγούμαστε πάντα σε σωστό αποτέλεσμα; Αν όχι, τι πιστεύετε ότι φταίει;

<u>Απάντηση:</u>

i) Ανεξάρτητα από την επιλογή βήματος, εάν ο αλγόριθμος εκκινήσει από το (0,0), δε θα πραγματοποιήσει καμία επανάληψη και θα επιστρέψει ως αποτέλεσμα το αρχικό σημείο.

Η συμπεριφορά αυτή οφείλεται στο γεγονός ότι, στο (0,0) , η παράγωγος της f είναι μηδενική (και άρα το ίδιο και η νόρμα της). Επομένως η συνθήκη τερματισμού norm(grad(:,k)) < accuracy ικανοποιείται άμεσα για οσοδήποτε μικρή ακρίβεια.

Από το γράφημα της f είναι εμφανές ότι το σημείο (0,0) βρίσκεται σε μια επίπεδη περιοχή, στην οποία ο αλγόριθμος εγκλωβίζεται, αφού δεν έχει κάποια κατεύθυνση να ακολουθήσει.

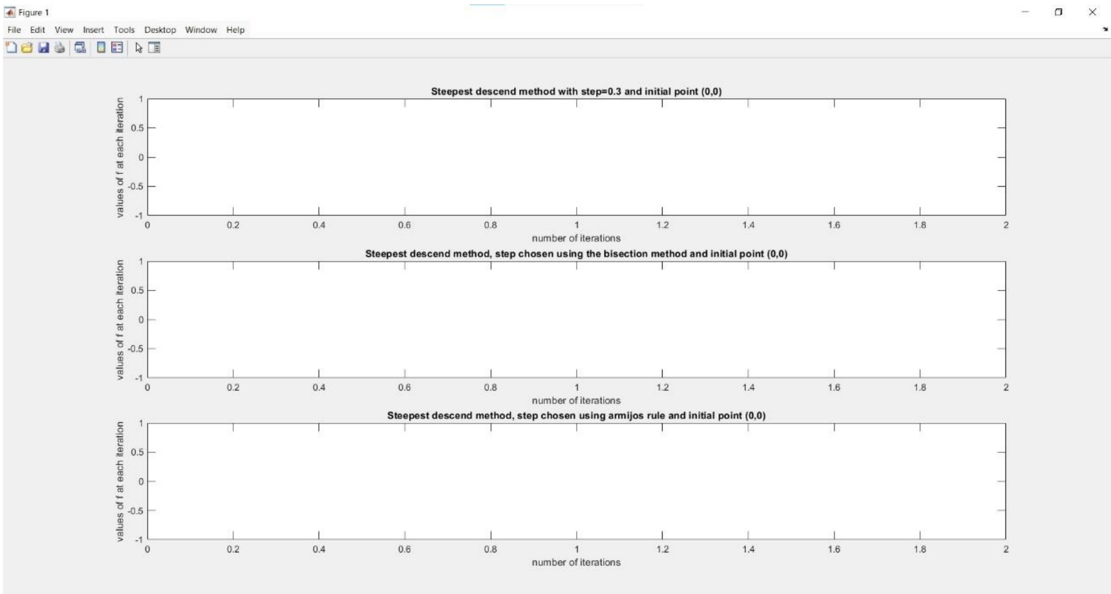


Figure 1: Γραφική παράσταση της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων που απαιτούνται μέχρι να τερματίσει ο αλγόριθμος μέγιστης καθόδου, για αρχικό σημείο (0,0).

ii) Για αρχικό σημείο (-1,1), ο αλγόριθμος δίνει ικανοποιητικά αποτελέσματα. Συγκεκριμένα:

α) Για σταθερό βήμα 0.5, min=(-1.5784, 0.0027). Αριθμός επαναλήψεων: k=7.

β) Για βήμα τέτοιο ώστε να ελαχιστοποιεί την f (x_k+ $γ_k$ d_k), min=(-1.5801, 0.0043).

Αριθμός επαναλήψεων: k=6.

Ο αριθμός επαναλήψεων είναι συγκρίσιμος με αυτόν που προκύπτει για σταθερό βήμα, και η ακρίβεια ελαφρώς καλύτερη.

(η ελαχιστοποίηση της f (f (x_k+ γ_k d_k),) έγινε με χρήση της μεθόδου της διχοτόμου)

γ)Για βήμα βάσει του κανόνα Armijo (με a=0.1 , b=0.1 , s=1), min=(-1.5811, 0.0053). Αριθμός επαναλήψεων: k=38.

Παρατηρώ ότι, αν το βήμα επιλεγεί βάσει του κανόνα Armijo, η σύγκλιση είναι πιο ομαλή, και ο αριθμός επαναλήψεων σημαντικά μεγαλύτερος. Δεν παρατηρώ ουσιαστική διαφορά στην ακρίβεια του αποτελέσματος σε σχέση με βήμα τέτοιο ώστε να ελαχιστοποιεί την f (x_k+ γ_k d_k).

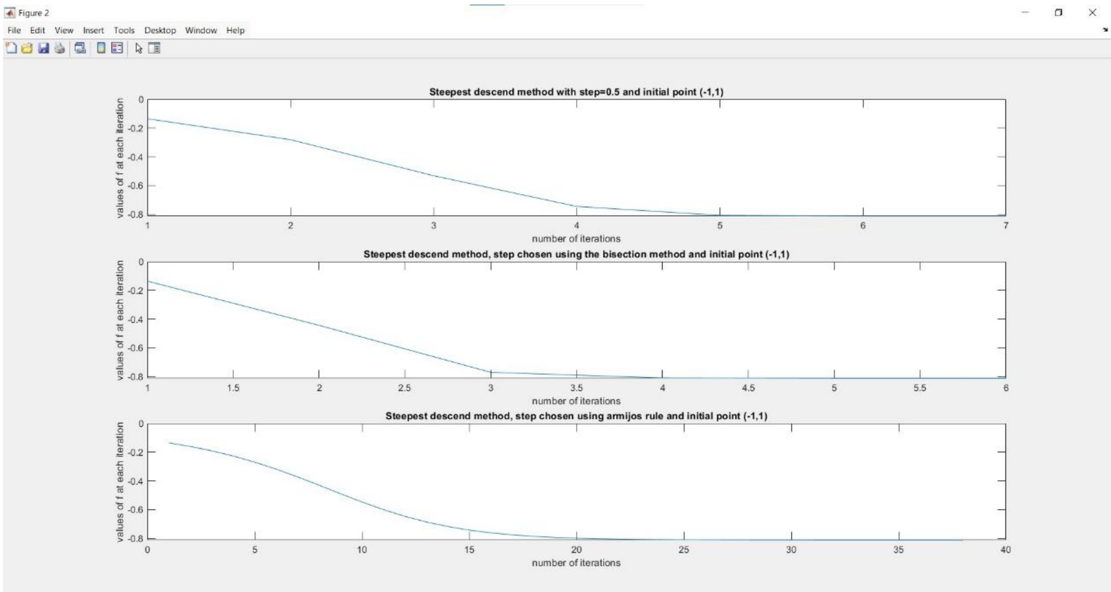


Figure 2: Γραφική παράσταση της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων που απαιτούνται μέχρι να τερματίσει ο αλγόριθμος μέγιστης καθόδου, για αρχικό σημείο (-1,1).

Όπως φαίνεται και στο γράφημα της f, το σημείο (-1,1) βρίσκεται μέσα στο «χωνί», ο πυθμένας του οποίου είναι το ελάχιστο της f. Το διάνυσμα κατεύθυνσης οδηγεί τον αλγόριθμο στη σωστή γενική κατεύθυνση, και δεν υπάρχουν τοπικά ακρότατα στα οποία θα μπορούσε να εγκλωβιστεί.

iii) Για αρχικό σημείο (1,-1), ο αλγόριθμος δεν προσδιορίζει επιτυχώς το σημείο ελαχίστου της f, ανεξάρτητα από τη μέθοδο επιλογής βήματος.

α) Για σταθερό βήμα 0.5, min=(0.3322 , -1.2943). Αριθμός επαναλήψεων: k=23.

β) Για βήμα τέτοιο ώστε να ελαχιστοποιεί την f (x_k+ γ_k d_k), min=(0.3431, -1.3403).

Αριθμός επαναλήψεων: k=11.

γ)Για βήμα βάσει του κανόνα Armijo (με a=0.1 , b=0.1 , s=1) , min=(0.3271 ,-1.2685). Αριθμός επαναλήψεων: k=117.

Βλέποντας και το σχήμα, φαίνεται ότι ο αλγόριθμος τερμάτισε σε λανθασμένο σημείο επειδή εγκλωβίστηκε στην επίπεδη περιοχή ανάμεσα στο «ύψωμα» και στο «χωνί» (όπου η κλίση της f είναι σχεδόν μηδέν).

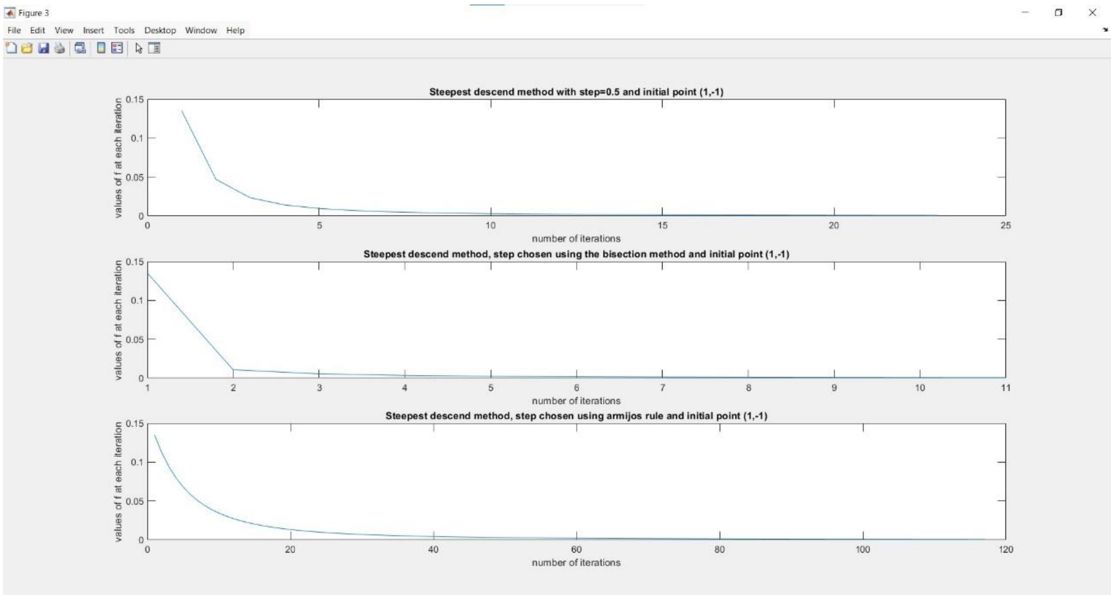


Figure 3: Γραφική παράσταση της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων που απαιτούνται μέχρι να τερματίσει ο αλγόριθμος μέγιστης καθόδου, για αρχικό σημείο (1,-1).

<u>Θέμα 3:</u> Επαναλάβετε τα ερωτήματα του Θέματος 2 χρησιμοποιώντας την μέθοδο Newton.

<u>Απάντηση:</u>

i)Για σημείο εκκίνησης (0,0) , όμοια με τη μέθοδο μέγιστης καθόδου και για τον ίδιο ακριβώς λόγο, ο αλγόριθμος επιστρέφει το σημείο (0,0), ασχέτως του τρόπου επιλογής βήματος.

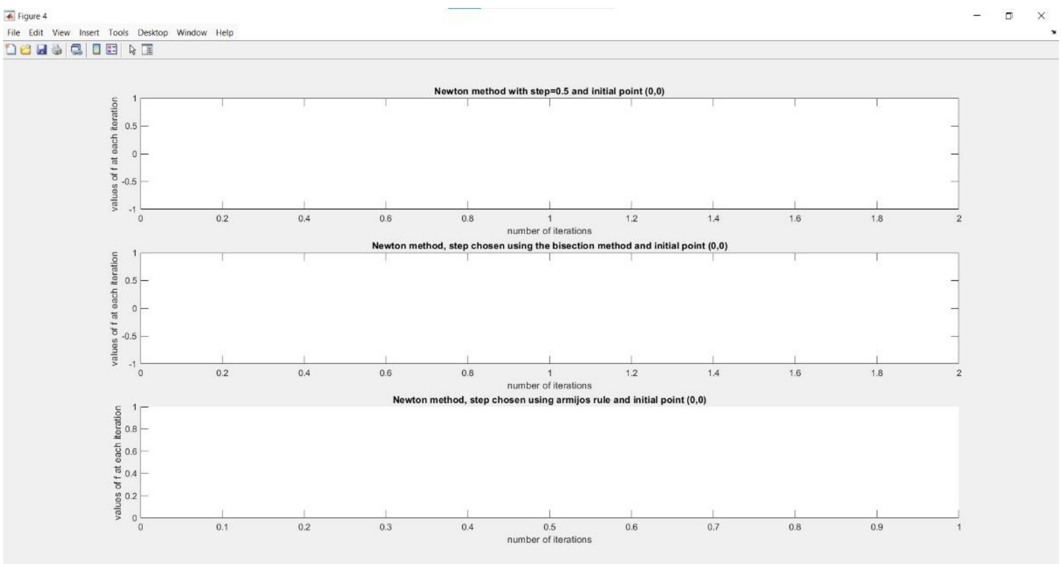


Figure 4: Γραφική παράσταση της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων που απαιτούνται μέχρι να τερματίσει ο αλγόριθμος Newton, για αρχικό σημείο (0,0).

ii) Για σημείο εκκίνησης (-1,1) , η μέθοδος Newton αποτυγχαίνει στον προσδιορισμό του σημείου ελαχίστου.

double(eig(hessian_f(-1,1))) = (0.5413, -1.0827), δηλαδή οι δύο ιδιοτιμές του εσσιανού της f στο αρχικό σημείο έχουν αντίθετο πρόσημο. Επομένως ο πίνακας δεν είναι θετικά ορισμένος, όπως προϋποθέτει η μέθοδος Newton.

α) Για σταθερό βήμα 0.5, min=(-0.6408, 2.0975). Αριθμός επαναλήψεων: k=9.

Συνθήκη Τερματισμού:

double(norm(gradient_f(-0.6408, 2.0975))) = 0.0068 < accuracy = 0.01

Ιδιοτιμές του εσσιανού πίνακα της f στο τελικό σημείο (είναι ετερόσημες):

double(eig(hessian_f(-0.6408, 2.0975))) = (0.0054, -0.0441)

β) Για βήμα τέτοιο ώστε να ελαχιστοποιεί την f (x_k+ γ_k d_k), min=(-0.6134, 2.1332).

Αριθμός επαναλήψεων: k=5.

Συνθήκη Τερματισμού:

double(norm(gradient_f(-0.6134, 2.1332))) = 0.0051 < accuracy = 0.01

Ιδιοτιμές του εσσιανού πίνακα της f στο τελικό σημείο:

double(eig(hessian_f(-0.6134, 2.1332))) = (0.0039, -0.0347) (ετερόσημες)

γ)Για βήμα βάσει του κανόνα Armijo (με a=0.1 , b=0.1 , s=1), min=(-0.6656, 2.0470). Αριθμός επαναλήψεων: k=40.

Συνθήκη Τερματισμού:

double(norm(gradient_f(-0.6656, 2.0470))) = 0.0094 < accuracy = 0.01

Ιδιοτιμές του εσσιανού πίνακα της f στο τελικό σημείο:

double(eig(hessian_f(-0.6656, 2.0470))) = (0.0076, -0.0580)

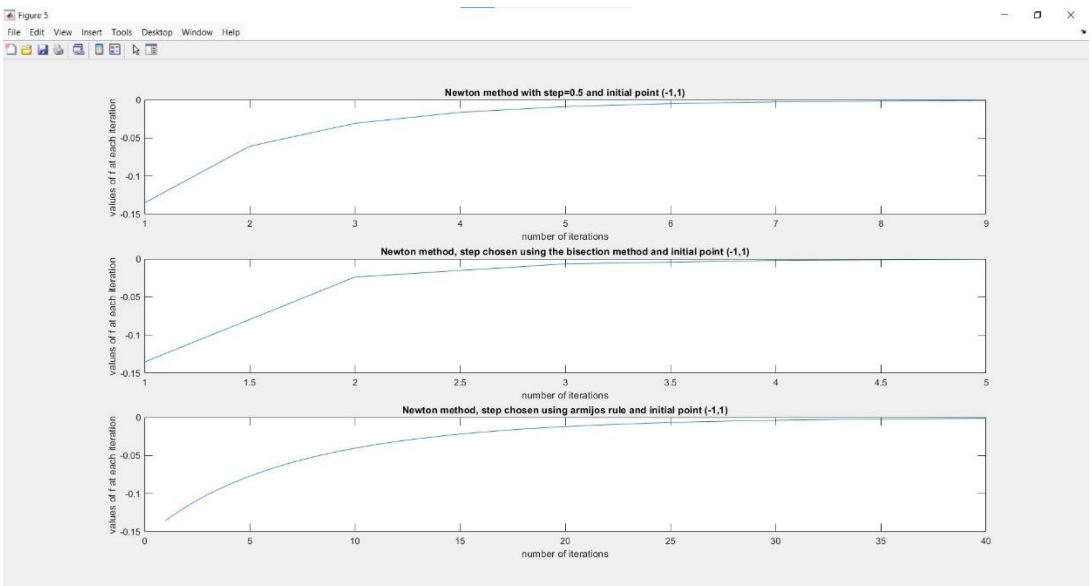


Figure 5: Γραφική παράσταση της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων που απαιτούνται μέχρι να τερματίσει ο αλγόριθμος Newton, για αρχικό σημείο (-1,1).

Παρατηρώ ότι, στο σημείο εκκίνησης και στο σημείο τερματισμού (και μάλλον και σε καποια από τα ενδιάμεσα σημεία) οι ιδιοτιμές του εσσιανού της f είναι ετερόσημες. Επομένως ο εσσιανός της f δεν είναι πάντα θετικά ορισμένος, όπως προϋποθέτει η μέθοδος Newton, και ο αλγόριθμος συμπεριφέρεται αντιπαραγωγικά. Στο figure 5 φαίνεται ότι σε κάθε επανάληψη, προκύπτει τιμή της f μεγαλύτερη από την προηγούμενη, δηλαδή παραβιάζεται το 1^o κριτήριο καλής λειτουργίας.

iii) Για αρχικό σημείο (1,-1), ο αλγόριθμος δεν προσδιορίζει επιτυχώς το σημείο ελαχίστου της f, ανεξάρτητα από τη μέθοδο επιλογής βήματος.

Σε αντίθεση με την περίπτωση εκκίνησης από το (-1,1), το 1^o κριτήριο καλής λειτουργίας ικανοποιείται.

double(eig(hessian_f(1,-1))) = (-0.5413 , 1.0827)   (μία θετική και μία αρνητική ιδιοτιμή.)

α) Για σταθερό βήμα 0.5, min=(0.6408, -2.0975). Αριθμός επαναλήψεων: k=9.

Συνθήκη Τερματισμού:

double(norm(gradient_f(0.6408, -2.0975))) =0.0068 < accuracy = 0.01

Ιδιοτιμές του εσσιανού πίνακα της f στο τελικό σημείο (ετερόσημες):

double(eig(hessian_f(0.6408, -2.0975))) = (-0.0054, 0.0441)

β) Για βήμα τέτοιο ώστε να ελαχιστοποιεί την f (x_k+ $γ_k$ d_k), min=(0.6134, -2.1332).

Αριθμός επαναλήψεων: k=5.

Συνθήκη Τερματισμού:

double(norm(gradient_f(0.6134, -2.1332))) =0.0051 < accuracy = 0.01

Ιδιοτιμές του εσσιανού πίνακα της f στο τελικό σημείο (ετερόσημες):

double(eig(hessian_f(0.6134, -2.1332))) = (-0.0039, 0.0347)

γ)Για βήμα βάσει του κανόνα Armijo, min=(0.6656, -2.0470). Αριθμός επαναλήψεων: k=40.

Συνθήκη Τερματισμού:

double(norm(gradient_f(0.6656, -2.0470))) =0.0094 < accuracy = 0.01

Ιδιοτιμές του εσσιανού πίνακα της f στο τελικό σημείο (ετερόσημες):

double(eig(hessian_f(0.6656, -2.0470))) = (-0.0076, 0.0580)

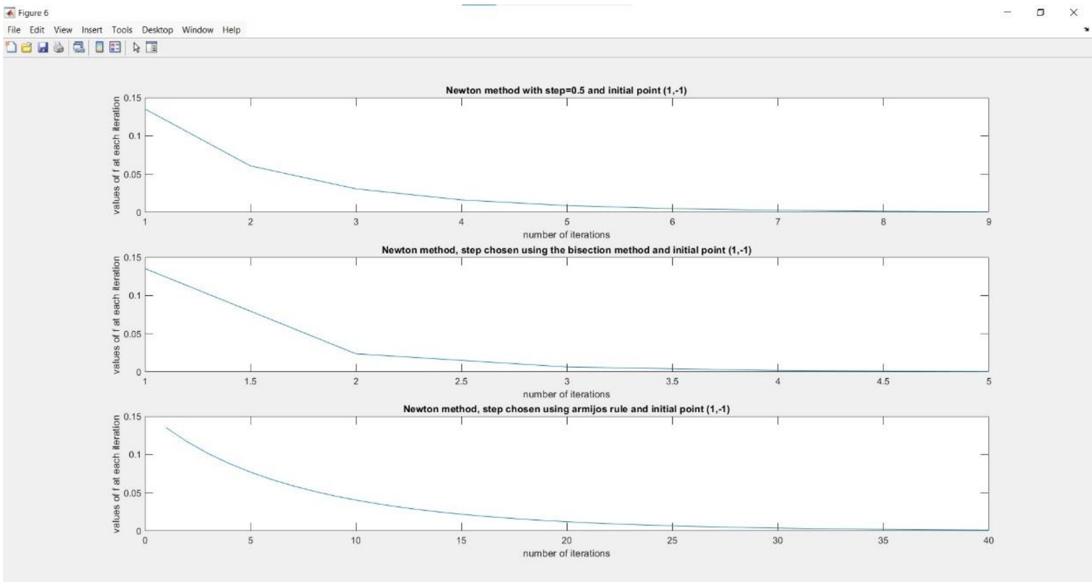


Figure 6: Γραφική παράσταση της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων που απαιτούνται μέχρι να τερματίσει ο αλγόριθμος Newton, για αρχικό σημείο (1,-1).

<u>Θέμα 4:</u> Επαναλάβετε τα ερωτήματα του Θέματος 2 χρησιμοποιώντας την μέθοδο Levenberg-Marquardt.

i) Για σημείο εκκίνησης [0,0], ο αλγόριθμος Levenberg-Marquardt παγιδεύεται, όπως ο αλγόριθμος μέγιστης καθόδου και ο αλγόριθμος Newton.

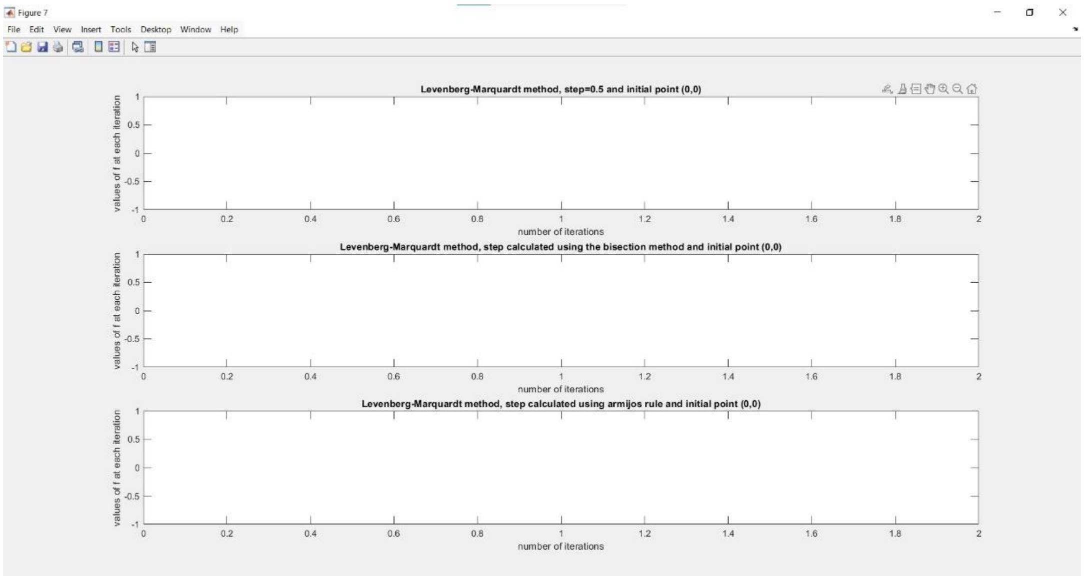


Figure 7: Γραφική παράσταση της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων που απαιτούνται μέχρι να τερματίσει ο αλγόριθμος Levenberg-Marquardt, για αρχικό σημείο (0,0).

ii) Για σημείο εκκίνησης [-1,1], ο αλγόριθμος επιστρέφει σωστό αποτέλεσμα.

α) Για σταθερό βήμα 0.5, min=(-1.5811, 0.0059). Αριθμός επαναλήψεων: k=33.

β) Για βήμα τέτοιο ώστε να ελαχιστοποιεί την $f(x_k + \gamma_k d_k)$,

min=(-1.5814, 0.0057).

Αριθμός επαναλήψεων: k=26.

γ)Για βήμα βάσει του κανόνα Armijo, min=(-1.5810, 0.0061). Αριθμός

επαναλήψεων: k=173. Η σύγκλιση είναι πιο ομαλή, και ο αριθμός επαναλήψεων σημαντικά μεγαλύτερος.

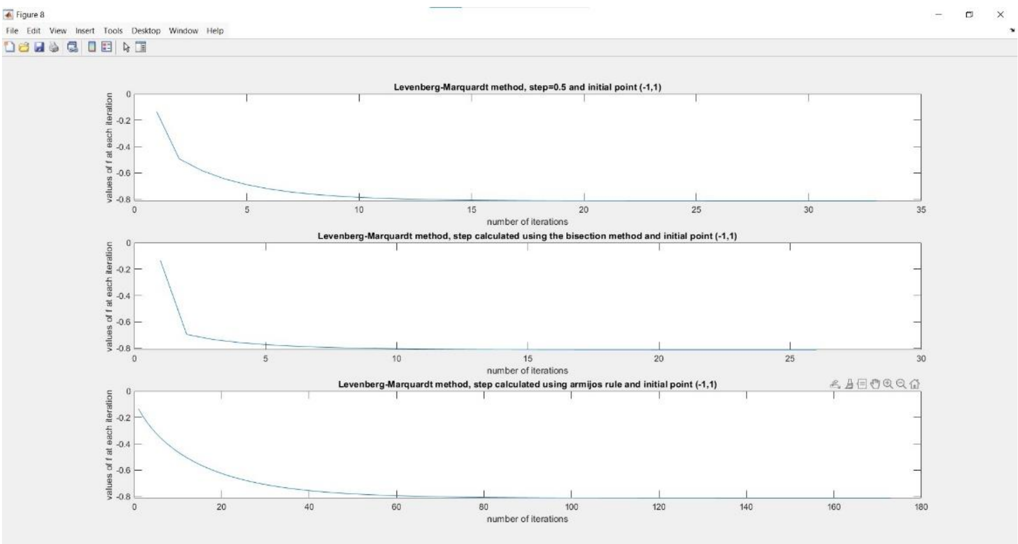


Figure 8: Γραφική παράσταση της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων που απαιτούνται μέχρι να τερματίσει ο αλγόριθμος Levenberg-Marquardt, για αρχικό σημείο (-1,1).

iii) Για αρχικό σημείο (1,-1), ο αλγόριθμος δεν προσδιορίζει επιτυχώς το σημείο ελαχίστου της f, ανεξάρτητα από τη μέθοδο επιλογής βήματος, διότι είναι επιρρεπής στα ίδια λάθη με τον αλγόριθμο μέγιστης καθόδου (εγκλωβισμός σε τοπικό ακρότατο)

α) Για σταθερό βήμα 0.5, min=(0.3142 , -1.2090). Αριθμός επαναλήψεων: k=38.

β) Για βήμα τέτοιο ώστε να ελαχιστοποιεί την f (x_k+ γ_k d_k), min=(0.3149, -1.2136).

Αριθμός επαναλήψεων: k=19.

γ)Για βήμα βάσει του κανόνα Armijo (με a=0.1 , b=0.1 , s=1) ,

min=(0.3136, -1.2056). Αριθμός επαναλήψεων: k=192.

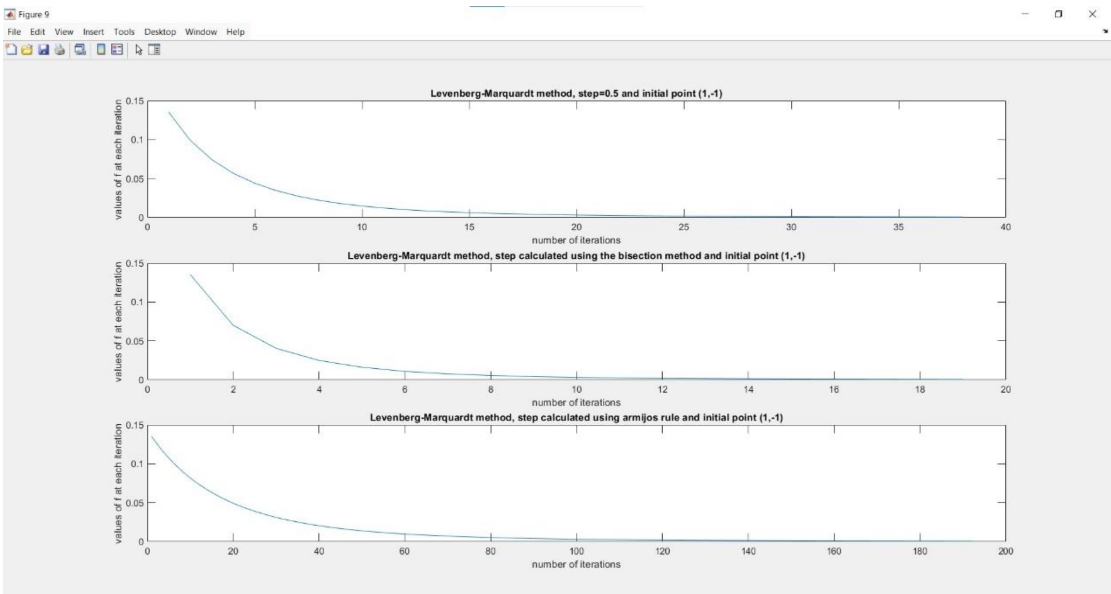


Figure 9: Γραφική παράσταση της σύγκλισης της αντικειμενικής συνάρτησης ως προς τον αριθμό των επαναλήψεων που απαιτούνται μέχρι να τερματίσει ο αλγόριθμος Levenberg-Marquardt, για αρχικό σημείο (1,-1).

Συνοπτικά, μόνο ο αλγόριθμος μέγιστης καθόδου και ο αλγόριθμος Levenberg-Marquardt προσδιορίζουν επιτυχώς το σημείο ελαχίστου, και μόνο αν εκκινήσουν από το σημείο (-1,1) (εκ των τριών δοθέντων αρχικών σημείων).

Από τους τρεις τρόπους επιλογής βήματος, ο κανόνας Armijo οδηγεί σε μεγάλο αριθμό επαναλήψεων, αλλά και σε πιο ομαλή σύγκλιση, επομένως έχει καλύτερες δυνατότητες σχετικά με την ακρίβεια.

Η επιλογή βήματος τέτοιου ώστε να ελαχιστοποιεί την f (x_k+ γ_k d_k) οδηγεί σε λιγότερο ομαλή σύγκλιση, αλλά με μικρότερο αριθμό απαναλήψεων από όλους τους τρόπους. Δε μπορώ να παρατηρήσω ουσιαστική διαφορά ανάμεσα στην ακρίβεια αυτού του τρόπου επιλογής βήματος και της επιλογής με βάσει τον κανόνα Armijo, αλλά πιθανόν αυτό να οφείλεται στις επιλογές που έκανα για τις παραμέτρους του Armijo και της συνάρτησης ελαχιστοποίησης

της f (x_k+ γ_k d_k) ως προς γ_k.

Η επιλογή σταθερού βήματος είναι η υπολογιστικά απλούστερη, και στην περίπτωση αυτή οδηγεί σε αριθμό επαναλήψεων λίγο μεγαλύτερο από (αλλά συγκρίσιμο με) την επιλογή βήματος τέτοιου ώστε να ελαχιστοποιεί την f $(x_k + \gamma_k d_k)$. Η σύγκλιση της αντικειμενικής συνάρτησης όμως είναι λιγότερο ομαλή από ότι με τους άλλους δύο τρόπους, και τα αποτελέσματα λιγότερο ακριβή.

Αξίζει να σημειωθεί ότι, στις περιπτώσεις που οι αλγόριθμοι ανταπεξέρχονται επιτυχώς, οι διαφορές στην ακρίβεια και στην ταχύτητα είναι αρκετά μικρές. Φυσικά, αν οι απαιτήσεις και η κλίμακα των προβλημάτων ελαχιστοποίησης ήταν μεγαλύτερες, τα πράγματα θα ήταν διαφορετικά.