

Τεχνικές Βελτιστοποίησης: Εργασία 2

Ιωάννης Μιχάλαϊνας AEM:10902

Νοέμβριος 2025

Αποθετήριο Κώδικα:

github.com/ioannisam/optimization-techniques

Περίληψη

Στην εργασία αυτή υλοποιούνται και συγκρίνονται n -διάστατες μέθοδοι βελτιστοποίησης - συγκεκριμένα η μέθοδος Μέγιστης Καθόδου, η μέθοδος Newton και η μέθοδος Levenberg-Marquardt. Για καθεμία από αυτές εξετάζεται ο αριθμός υπολογισμών και η σύγκλιση, ως προς την ταχύτητα και την ακρίβεια, εφαρμόζοντάς τις από διαφορετικά αρχικά σημεία και με διαφορετικές μεθόδους υπολογισμού του βήματος. Τα αποτελέσματα δείχνουν τις ομοιότητες και τις διαφορές μεταξύ των μεθόδων και αναδεικνύουν τα πλεονεκτήματα και τους περιορισμούς της καθεμιάς. Ο σκοπός της εργασίας αυτής είναι η εξοικείωση με τις παραπάνω μεθόδους και η σύγκριση μεταξύ τους, όσον αφορά την αποδοτικότητα.

Περιεχόμενα

1	Εισαγωγή	2
2	Μέθοδος Μέγιστης Καθόδου	4
2.1	Αλγόριθμος	4
2.2	Γραφικές Παραστάσεις	4
2.3	Συμπεράσματα	5
3	Μέθοδος Newton	7
3.1	Αλγόριθμος	7
3.2	Γραφικές Παραστάσεις	7
3.3	Συμπεράσματα	8
4	Μέθοδος Levenberg-Marquardt	10
4.1	Αλγόριθμος	10
4.2	Γραφικές Παραστάσεις	10
4.3	Συμπεράσματα	11
5	Σύγκριση Μεθόδων	13
A'	Εργασία	14
B'	Βιβλιογραφία	14

1 Εισαγωγή

Στην εργασία αυτή μελετάμε διάφορες τεχνικές βελτιστοποίησης, ονομαστικά τη μέθοδο Μέγιστης Καθόδου, τη μέθοδο Newton και τη μέθοδο Levenberg-Marquardt. Ο σκοπός είναι η προσέγγιση του x^* , της λύσης του προβλήματος ελαχιστοποίησης.

Οι μέθοδοι που θα εξετάσουμε βασίζονται στην ιδέα της επαναληπτικής καθόδου, η οποία οδηγεί σε ολοένα και βελτιωμένες τιμές της f .

$$f(x_{k+1}) < f(x_k), \quad k = 0, 1, 2, \dots$$

Στόχος των μεθόδων αυτών είναι με κάθε επανάληψη συγκλίνουμε προς το x^* . Αυτό επιτυγχάνεται μέσω της επαναληπτικής σχέσης:

$$x_{k+1} = x_k + \gamma_k d_k$$

Όπου k ο αριθμός των επαναλήψεων, d_k το διάνυσμα της κατεύθυνσης του αλγορίθμου που ικανοποιεί τη συνθήκη $\nabla f^\top(x_k) d_k < 0$ και $\gamma_k > 0$, τέτοιο ώστε να ισχύει η συνθήκη $f(x_k + \gamma_k d_k) < f(x_k)$.

Οι αλγόριθμοι που θα μελετήσουμε είναι της γενικής μορφής

$$x_{k+1} = x_k - \gamma_k \Delta_k \nabla f(x_k)$$

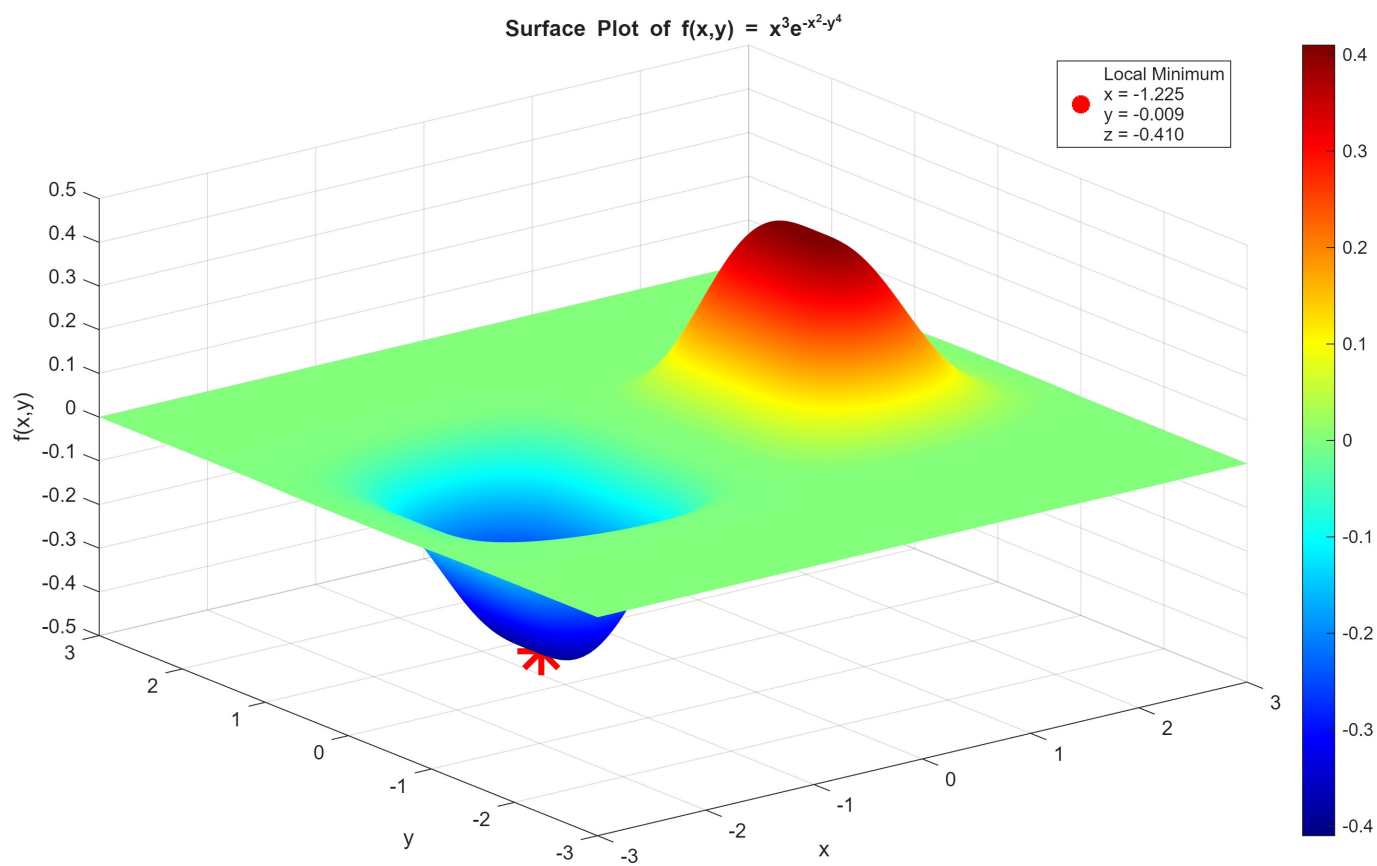
Η επιλογή του βήματος γ_k είναι νευραλγικής σημασίας για την ορθότητα του αλγορίθμου, αφού αν επιλεγεί πολύ μικρό θα οδηγήσει σε ασήμαντη μείωση της f και κατά συνέπεια αργή σύγκλιση. Αντίθετα, αν επιλεγεί πολύ μεγάλο μπορεί να οδηγήσει τον αλγόριθμο σε αστάθεια. Θα εξετάσουμε τρεις διαφορετικούς τρόπους προσδιορισμού του βήματος:

- **Σταθερό Βήμα:** Κρατάμε σταθερό το βήμα γ_k καθ' όλη τη διάρκεια της αναζήτησης. Αποτελεί την απλούστερη μέθοδο, αφού ελαχιστοποιεί την υπολογιστική πολυπλοκότητα του αλγορίθμου. Δεδομένης της ευαισθησίας του ζητήματος της επιλογής βήματος, αυτή η μέθοδος προϋποθέτει μεγάλη εμπειρία και ακρίβεια για την ορθή εφαρμογή της.
- **Βέλτιστο Βήμα:** Επιλέγουμε βήμα τέτοιο ώστε να ελαχιστοποιηθεί η $f(x_k + \gamma_k d_k)$ ως προς γ_k . Αυτό αποτελεί μονοδιάστατο πρόβλημα βελτιστοποίησης που σε αυτή την εργασία θα επιλύεται με τη βοήθεια της μεθόδου Χρυσής Τομής, υλοποιημένης στην προηγούμενη εργασία. Επιλέγουμε τη μέθοδο Χρυσής Τομής διότι είναι η ταχύτερη μέθοδος χωρίς χρήση παραγώγων που δεν απαιτεί τον εκ των προτέρων υπολογισμό των συνολικών επαναλήψεων.
- **Κανόνας Armijo:** Σύμφωνα με τον κανόνα Armijo, το βήμα επιλέγεται ως $\gamma_k = s\beta^{\mu_k}$, όπου $\mu_k \in \mathbb{Z}_{\geq 0}$ τέτοιο ώστε $f(x_k) - f(x_{k+1}) \geq -\alpha\beta^{\mu_k} s d_k^\top \nabla f(x_k)$. Εδώ ο συντελεστής s δηλώνει το αρχικό βήμα. Συνήθως ισχύει $\alpha \in [10^{-5}, 10^{-1}]$ και $\beta \in [\frac{1}{10}, \frac{1}{2}]$.

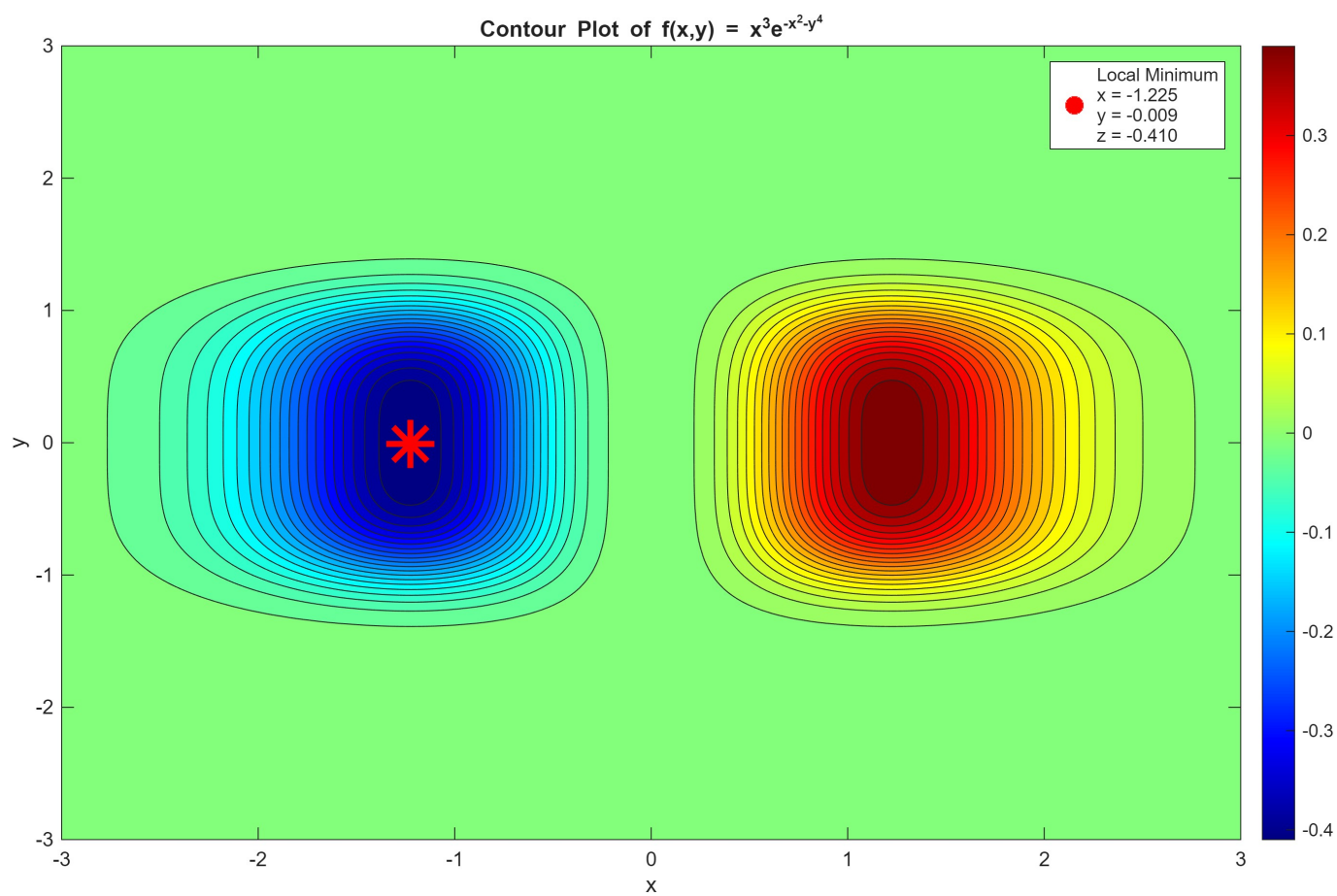
Δοσμένων των αρχικών σημείων $[0, 0]$, $[-1, -1]$ και $[1, 1]$ θα ελαχιστοποιηθεί η:

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

Η συνάρτηση f είναι δύο φορές παραγωγίσιμη, συνθήκη απαραίτητη για την εφαρμογή των αλγορίθμων που θα εξετάσουμε. Παρακάτω παρουσιάζεται η γραφική της παράσταση τόσο στον τρισδιάστατο χώρο όσο και στον δισδιάστατο χώρο, υπό τη μορφή ισοβαρών καμπυλών. Αυτή φαίνεται να παρουσιάζει ελάχιστο στο $[-1.225, 0.009]$ με τιμή $f_{\min} = -0.41$. Για θετικές τιμές του x , η συνάρτηση αποκτά θετικές τιμές και σχηματίζει έναν ομαλό λόφο, ενώ για αρνητικές τιμές του x εμφανίζεται μια βαθιά κοιλάδα η οποία οδηγεί στο τοπικό ελάχιστο που εντοπίστηκε.



Σχήμα 1: Τρισδιάστατη απεικόνιση της συνάρτησης f



Σχήμα 2: Ισοεπίπεδα απεικόνιση της συνάρτησης f

2 Μέθοδος Μέγιστης Καθόδου

2.1 Αλγόριθμος

Η μέθοδος της Μέγιστης Καθόδου προκύπτει επιλέγοντας $\Delta_k = I$, $k = 1, 2, \dots$ και η επαναληπτική σχέση, λοιπόν, γίνεται:

$$x_{k+1} = x_k - \gamma_k \nabla f(x_k)$$

Το διάνυσμα κατεύθυνσης $d_k = -\nabla f(x_k)$ ορίζεται από την αρνητική κλίση της f στο σημείο x_k .

Ένα βασικό μειονέκτημα της μεθόδου Μέγιστης Καθόδου είναι πως η πορεία που «χαράζει» προς το ελάχιστο είναι μια τεθλασμένη γραμμή, όπου το διάνυσμα που ενώνει το x_k με το x_{k+1} είναι πάντα κάθετο στο διάνυσμα που ενώνει το x_{k+1} με το x_{k+2} . Αυτή η συμπεριφορά καθιστά τη μέθοδο αυτή αργή ως προς τη σύγκλιση σε πολλές περιπτώσεις.

Algorithm 1 Μέθοδος Μέγιστης Καθόδου

Require: συνάρτηση f , αρχικό σημείο x_1 , σταθερά $\epsilon > 0$

Ensure: προσέγγιση ελαχίστου x^*

```
1: ορίστε  $x_k \leftarrow x$ ,  $k \leftarrow 1$ 
2: while  $|\nabla f(x_k)| > \epsilon$  do
3:    $d_k \leftarrow -\nabla f(x_k)$ 
4:    $\gamma_k \leftarrow \arg \min_{\gamma > 0} f(x_k + \gamma d_k)$ 
5:    $x_k \leftarrow x_k + \gamma_k d_k$ 
6:    $k \leftarrow k + 1$ 
7: end while
8: return  $x^* \leftarrow x_k$ 
```

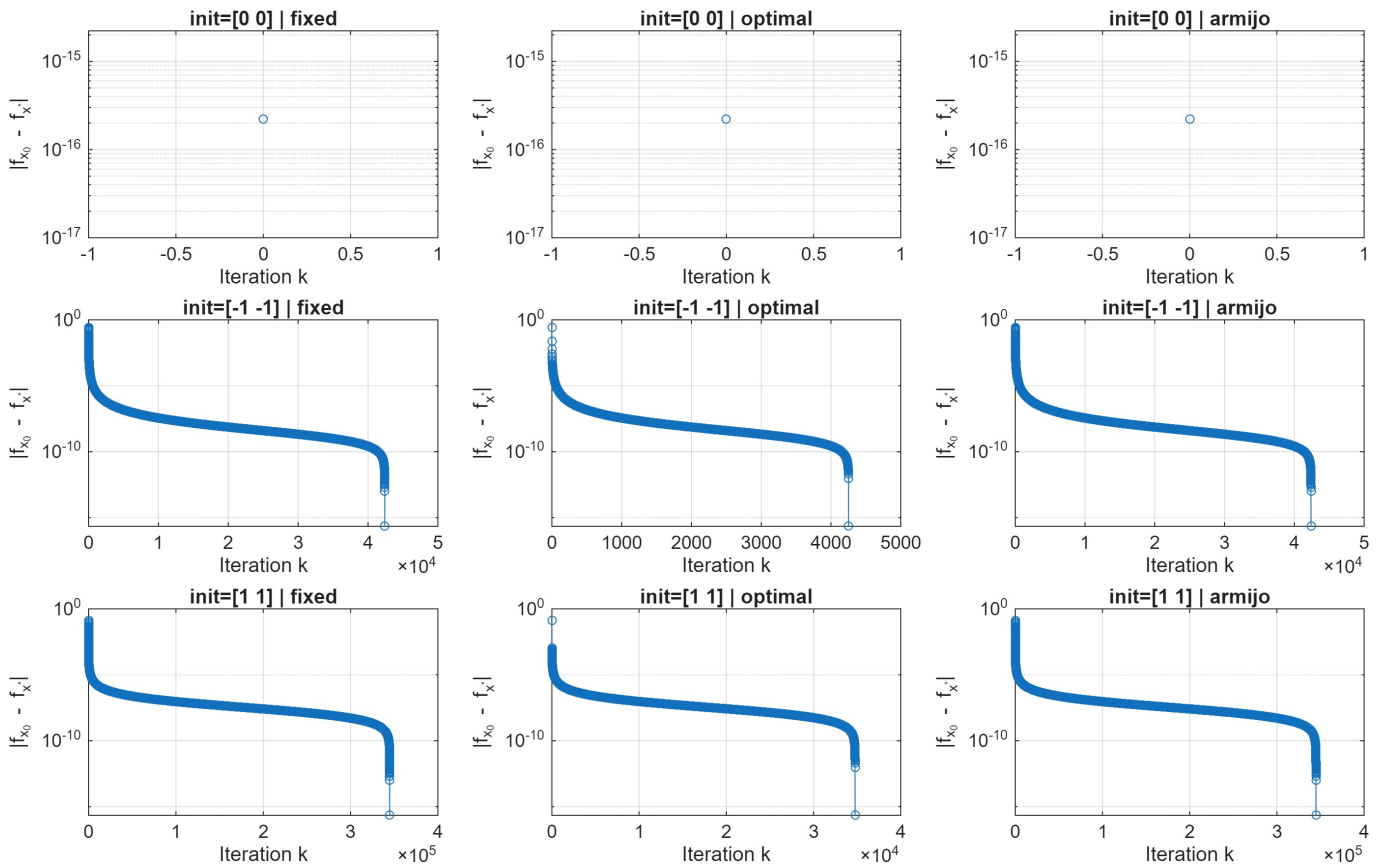
2.2 Γραφικές Παραστάσεις

Στον Πίνακα 1 και στο Σχήμα 3 παρουσιάζονται τα αποτελέσματα της μεθόδου Μέγιστης Καθόδου για τη συνάρτηση f . Καταγράφεται το ελάχιστο σημείο και οι επαναλήψεις για κάθε αρχικό σημείο και μέθοδο προσδιορισμού του βήματος.

Initial Point	Method	Minimum Point x^*	Minimum Value $f(x^*)$	k	Function Evaluations	Step Evaluations
[0 0]	fixed	[0 0]	0	0	2	0
[0 0]	optimal	[0 0]	0	0	2	0
[0 0]	armijo	[0 0]	0	0	2	0
[-1 -1]	fixed	[-1.22 -0.00848]	-0.41	42397	84796	0
[-1 -1]	optimal	[-1.22 -0.00848]	-0.41	4252	8506	51024
[-1 -1]	armijo	[-1.22 -0.00848]	-0.41	42397	84796	84794
[1 1]	fixed	[0.875 2]	3.1e-08	344812	689626	0
[1 1]	optimal	[0.825 2]	3.12e-08	34765	69532	417180
[1 1]	armijo	[0.875 2]	3.1e-08	344812	689626	689624

Πίνακας 1: Συγκεντρωτικός πίνακας αποτελεσμάτων (Μέγιστη Κάθοδος)

Steepest Descend: Convergence history vs iteration k



Σχήμα 3: Σύγκλιση της f ανά τις επαναλήψεις

2.3 Συμπεράσματα

Βάσει των παραπάνω αποτελεσμάτων παρατηρούμε ότι η συμπεριφορά του αλγορίθμου εξαρτάται έντονα τόσο από το αρχικό σημείο όσο και από τον τρόπο υπολογισμού του βήματος γ_k .

Επίδραση Αρχικού Σημείου

Αρχικό σημείο $(0,0)$: Και οι τρεις μέθοδοι βήματος σταματούν αμέσως, καθώς στο σημείο αυτό ισχύει $\nabla f(0,0) = 0$. Το σημείο όμως δεν είναι ελάχιστο, συνεπώς παρατηρείται εγκλωβισμός. Ο αλγόριθμος, βασιζόμενος αποκλειστικά στο μηδενικό gradient, θεωρεί λανθασμένα ότι έχει βρεθεί βέλτιστη λύση. Άρα, η μέθοδος δεν εγγυάται πάντα σύγκλιση.

Αρχικό σημείο $(-1,-1)$: Όλες οι παραλλαγές βήματος συγκλίνουν στο παγκόσμιο ελάχιστο

$$x^* \approx (-1.22, -0.00848), \quad f(x^*) \approx -0.41.$$

Ωστόσο, οι διαφορές στον αριθμό επαναλήψεων είναι μεγάλες. Το σταθερό βήμα και ο κανόνας Armijo απαιτούν περίπου 4.2×10^4 επαναλήψεις, ενώ το βέλτιστο βήμα οδηγεί σε περίπου ένα δέκατο αυτού του πλήθους. Η αργή σύγκλιση οφείλεται στη χαρακτηριστική τεθλασμένη πορεία της μεθόδου.

Αρχικό σημείο $(1,1)$: Ο αλγόριθμος δεν οδηγείται στο πραγματικό ελάχιστο. Και οι τρεις μέθοδοι συγκλίνουν σε σημείο της μορφής

$$x^* \approx (0.9, 2), \quad f(x^*) \approx 3 \times 10^{-8},$$

μια περιοχή όπου η f μηδενίζεται εκθετικά. Η μέθοδος κατευθύνεται μακριά από το πραγματικό ελάχιστο, με αποτέλεσμα να κολλάει σε επίπεδη περιοχή. Ο μεγάλος αριθμός επαναλήψεων (έως και 3.4×10^5) επιβεβαιώνει την δυσκολία πλοήγησης σε ανεπιθύμητες περιοχές του χώρου.

Επίδραση του βήματος

Σταθερό Βήμα: Απλό στην εφαρμογή, αλλά εξαιρετικά αργό. Σε στενές κοιλάδες προκαλεί μεγάλες ταλαντώσεις και χιλιάδες επαναλήψεις.

Βέλτιστο βήμα: Πετυχαίνει σημαντικά λιγότερες επαναλήψεις, αλλά απαιτεί πολύ μεγάλο αριθμό *step evaluations*, λόγω της μονοδιάστατης αναζήτησης (μέθοδος Χρυσής Τομής). Έφερε τα καλύτερα αποτελέσματα για αυτή τη μέθοδο, παρ' όλα αυτά, δεν εξασφαλίζει σύγκλιση στο σωστό ελάχιστο όταν το αρχικό σημείο είναι ακατάλληλο.

Κανόνας Armijo: Εξασφαλίζει μονοτονική μείωση της f , αλλά στη συγκεκριμένη συνάρτηση η συμπεριφορά του είναι σχεδόν ίδια με το σταθερό βήμα. Συχνά επιλέγει πολύ μικρά βήματα, οδηγώντας σε μεγάλο αριθμό επαναλήψεων.

Γενικά Συμπεράσματα

- Η μέθοδος Μέγιστης Καθόδου είναι ιδιαίτερα ευαίσθητη στα αρχικά σημεία. Δεν οδηγεί πάντα στο σωστό ελάχιστο.
- Το σταμάτημα στο $\nabla f(x_k) < \epsilon$ δεν αρκεί για να εγγυηθεί ελάχιστο (χαρακτηριστικό παράδειγμα το $(0, 0)$).
- Η μορφή της συνάρτησης (στενές κοιλάδες, επίπεδες περιοχές) επηρεάζει δραματικά τη σύγκλιση.
- Το βήμα παίζει καθοριστικό ρόλο: μικρό βήμα σημαίνει πολύ αργή πρόοδο, μεγάλο βήμα μπορεί να οδηγήσει σε αστάθεια.
- Η χρήση βέλτιστου βήματος μειώνει τις επαναλήψεις, αλλά αυξάνει το συνολικό υπολογιστικό κόστος.

Συνολικά, η Μέθοδος Μέγιστης Καθόδου είναι αξιόπιστη μόνο όταν το πεδίο της συνάρτησης είναι «καλοήθες» και όταν το αρχικό σημείο βρίσκεται κοντά στην επιθυμητή περιοχή. Διαφορετικά, είτε οδηγείται σε λανθασμένα αποτελέσματα είτε παρουσιάζει εξαιρετικά αργή σύγκλιση.

3 Μέθοδος Newton

3.1 Αλγόριθμος

Η μέθοδος Newton προκύπτει επιλέγοντας $\Delta_k = [\nabla^2 f(x_k)]^{-1}$, $k = 1, 2, \dots$ και η επαναληπτική σχέση, λοιπόν, γίνεται:

$$x_{k+1} = x_k - \gamma_k [\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$$

Το διάνυσμα κατεύθυνσης, συνεπώς είναι: $d_k = -[\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$.

Ένα βασικό μειονέκτημα της μεθόδου Newton είναι πως μεγάλα προβλήματα προκύπτουν όταν ο $\nabla^2 f(x)$ δεν είναι θετικά ορισμένος ή ακόμα και αντιστρέψιμος, πράγμα που συμβαίνει στην περίπτωση της συνάρτησης που εξετάζουμε σε αυτή την εργασία. Σε αυτές τις περιπτώσεις η μέθοδος Newton δεν ορίζεται ή δίνει λύσεις μακριά από το σημείο ολικού ελαχίστου.

Algorithm 2 Μέθοδος Newton

Require: συνάρτηση f , αρχικό σημείο x_1 , σταθερά $\epsilon > 0$

Ensure: προσέγγιση ελαχίστου x^*

- 1: ορίστε $x_k \leftarrow x$, $k \leftarrow 1$
 - 2: **while** $|\nabla f(x_k)| > \epsilon$ **do**
 - 3: $d_k \leftarrow -[\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$
 - 4: $\gamma_k \leftarrow \arg \min_{\gamma > 0} f(x_k + \gamma d_k)$
 - 5: $x_k \leftarrow x_k + \gamma_k d_k$
 - 6: $k \leftarrow k + 1$
 - 7: **end while**
 - 8: **return** $x^* \leftarrow x_k$
-

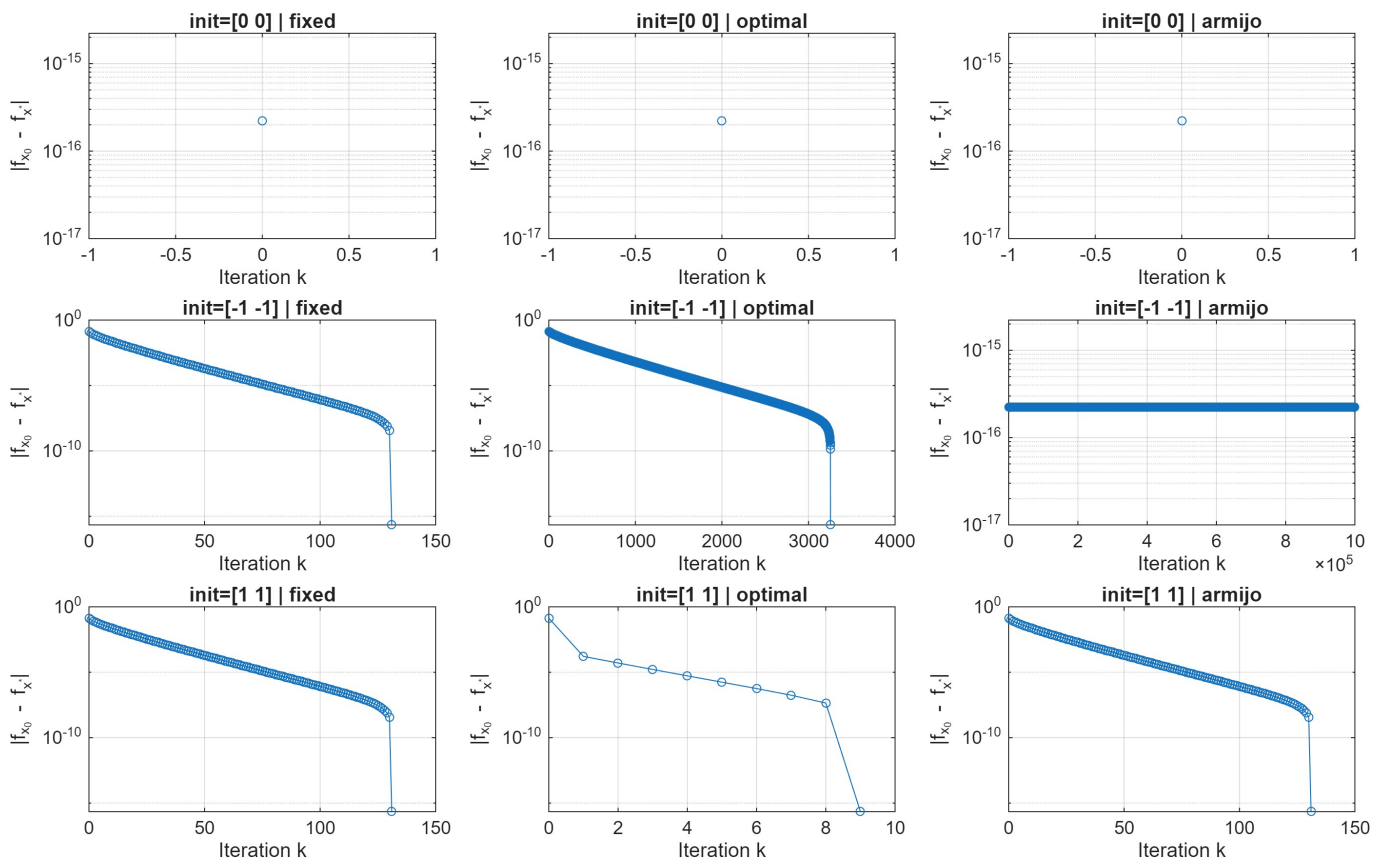
3.2 Γραφικές Παραστάσεις

Στον Πίνακα 2 και στο Σχήμα 4 παρουσιάζονται τα αποτελέσματα της μεθόδου Μέγιστης Καθόδου για τη συνάρτηση f . Καταγράφεται το ελάχιστο σημείο και οι επαναλήψεις για κάθε αρχικό σημείο και μέθοδο προσδιορισμού του βήματος.

Initial Point	Method	Minimum Point x^*	Minimum Value $f(x^*)$	k	Function Evaluations	Step Evaluations
[0 0]	fixed	[0 0]	0	0	2	0
[0 0]	optimal	[0 0]	0	0	2	0
[0 0]	armijo	[0 0]	0	0	2	0
[-1 -1]	fixed	[-0.376 -1.94]	-3.23e-08	131	264	0
[-1 -1]	optimal	[-0.374 -1.94]	-3.32e-08	3254	6510	39048
[-1 -1]	armijo	[-1 -1]	-0.135	999999	2000000	30999969
[1 1]	fixed	[0.376 1.94]	3.23e-08	131	264	0
[1 1]	optimal	[0.42 1.96]	2.31e-08	9	20	108
[1 1]	armijo	[0.376 1.94]	3.23e-08	131	264	262

Πίνακας 2: Συγκεντρωτικός πίνακας αποτελεσμάτων (Newton)

Newton: Convergence history vs iteration k



Σχήμα 4: Σύγκλιση της f ανά τις επαναλήψεις

3.3 Συμπεράσματα

Τα αποτελέσματα δείχνουν ότι η μέθοδος Newton παρουσιάζει σημαντικές δυσκολίες στη συγκεκριμένη συνάρτηση. Ο βασικός λόγος είναι ότι ο $\nabla^2 f$ δεν είναι θετικά ορισμένος στα περισσότερα σημεία, με αποτέλεσμα ο υπολογισμός της κατεύθυνσης

$$d_k = -[\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$$

να οδηγεί σε βήματα που δεν κατευθύνονται προς ελάχιστο, αλλά προς περιοχές όπου η f έχει επίπεδες ή ασταθείς καμπυλότητες.

Επίδραση Αρχικού Σημείου / Βήματος

Αρχικό σημείο (0,0): Όπως και στη μέθοδο Μέγιστης Καθόδου, και οι τρεις μέθοδοι βήματος τερματίζουν αμέσως, επειδή $\nabla f(0,0) = 0$. Το σημείο όμως δεν είναι ελάχιστο, συνεπώς παρατηρείται εγκλωβισμός.

Αρχικό σημείο (-1,-1): Η συμπεριφορά διαφέρει έντονα ανάλογα με το βήμα:

- **Σταθερό βήμα:** Η μέθοδος συγκλίνει γρήγορα, αλλά όχι στο πραγματικό ελάχιστο. Το σημείο στο οποίο καταλήγει,

$$x^* \approx (-0.376, -1.94), \quad f(x^*) \approx -3.23 \times 10^{-8},$$

είναι μια επίπεδη περιοχή όπου η f είναι σχεδόν μηδενική, μακριά από τη βαθιά κοιλότητα του πραγματικού ελάχιστου.

- **Βέλτιστο βήμα:** Η μονοδιάστατη αναζήτηση οδηγεί σε χιλιάδες επαναλήψεις (3254) και τεράστιο αριθμό step evaluations ($\approx 39k$), καταλήγοντας σε παρόμοιο σημείο με το σταθερό βήμα. Η μέθοδος δηλαδή «συμβιβάζεται» με μηδενικές τιμές της f αντί να κινείται προς το αρνητικό ελάχιστο.
- **Κανόνας Armijo:** Η μέθοδος αποτυγχάνει πλήρως. Ο αλγόριθμος φτάνει το μέγιστο πλήθος επαναλήψεων χωρίς ουσιαστική μεταβολή:

$$x_k = (-1, -1), \quad f(x_k) = -0.135.$$

Ο λόγος είναι ότι ο μη θετικά ορισμένος Εσσιανός δίνει κατευθύνσεις που η μέθοδος Armijo απορρίπτει συνεχώς, με αποτέλεσμα ο αλγόριθμος να μην κινείται καθόλου, έως ότου φτάσει στο όριο επαναλήψεων.

Αρχικό σημείο (1,1): Η συμπεριφορά είναι παρόμοια για κάθε βήμα:

- **Σταθερό βήμα:** Σύγκλιση σε σημείο της μορφής

$$x^* \approx (0.376, 1.94), \quad f(x^*) \approx 3.23 \times 10^{-8},$$

δηλαδή σε επίπεδη περιοχή μακριά από το παγκόσμιο ελάχιστο.

- **Βέλτιστο βήμα:** Μόνο εδώ η μέθοδος Newton συγκλίνει ταχύτατα (μόλις 9 επαναλήψεις). Ωστόσο, ακόμη και αυτή η σύγκλιση οδηγεί όχι στο πραγματικό ελάχιστο αλλά σε σημείο με σχεδόν μηδενική τιμή της f .
- **Armijo:** Σύγκλιση σε σημείο της μορφής

$$x^* \approx (0.376, 1.94), \quad f(x^*) \approx 3.23 \times 10^{-8},$$

δηλαδή σε επίπεδη περιοχή μακριά από το παγκόσμιο ελάχιστο.

Γενικά Συμπεράσματα

- Η μέθοδος Newton δεν είναι κατάλληλη για συναρτήσεις με μη θετικά ορισμένο Εσσιανό.
- Η μέθοδος Newton λειτουργεί σωστά μόνο όταν το αρχικό σημείο βρίσκεται σε περιοχή όπου το Εσσιανό είναι θετικά ορισμένο. Εδώ αυτό δεν συμβαίνει.
- Παρά την ταχύτητα σύγκλισης σε ιδανικές περιπτώσεις, εδώ καταλήγει σε λανθασμένα σημεία.
- Η χρήση Armijo με Newton μπορεί να οδηγήσει σε πλήρη ακινητοποίηση του αλγορίθμου.
- Το βέλτιστο βήμα δεν διορθώνει το πρόβλημα του Εσσιανού, απλώς αυξάνει το κόστος.

Συνολικά, η μέθοδος Newton αποτυγχάνει να προσεγγίσει το παγκόσμιο ελάχιστο της f και οδηγείται σε πλατώ. Αυτό την καθιστά σαφώς λιγότερο αξιόπιστη από τις υπόλοιπες μεθόδους για το συγκεκριμένο πρόβλημα.

4 Μέθοδος Levenberg-Marquardt

4.1 Αλγόριθμος

Η μέθοδος Levenberg-Marquardt είναι μια τροποποίηση του αλγορίθμου Newton που αποσκοπεί στο να ξεπεράσει το πρόβλημα που προκύπτει αν ο $\nabla^2 f(x_k)$ δεν είναι θετικά ορισμένος. Επιλέγουμε $\Delta_k = [\nabla^2 f(x_k) + \mu_k I]^{-1}$, $k = 1, 2, \dots$, με μ_k τέτοιο ώστε ο $[\nabla^2 f(x_k) + \mu_k I]$ να είναι θετικά ορισμένος. Η επαναληπτική σχέση γίνεται:

$$x_{k+1} = x_k - \gamma_k [\nabla^2 f(x_k) + \mu_k I]^{-1} \nabla f(x_k)$$

Το διάνυσμα κατεύθυνσης, συνεπώς είναι: $d_k = -[\nabla^2 f(x_k) + \mu_k I]^{-1} \nabla f(x_k)$.

Η μέθοδος αυτή αποτελεί τον συγκερασμό της μεθόδου Newton και Μέγιστης Καθόδου. Όταν το μ_k είναι αρκετά μεγάλο, ο παράγοντας $\mu_k I$ κυριαρχεί σε σχέση με τον $\nabla^2 f(x_k)$. Κατά συνέπεια η μέθοδος Levenberg-Marquardt θα λειτουργεί σχεδόν όπως η μέθοδος της Μέγιστης Καθόδου. Αν όμως το μ_k είναι αρκετά μικρό, τότε ο $\nabla^2 f(x_k)$ υπερισχύει και η μέθοδος συμπεριφέρεται σαν τη μέθοδο Newton.

Algorithm 3 Μέθοδος Levenberg-Marquardt

Require: συνάρτηση f , αρχικό σημείο x_1 , σταθερά $\epsilon > 0$

Ensure: προσέγγιση ελαχίστου x^*

- 1: ορίστε $x_k \leftarrow x$, $k \leftarrow 1$
 - 2: **while** $|\nabla f(x_k)| > \epsilon$ **do**
 - 3: ορίστε μ_k τέτοιο ώστε ο $[\nabla^2 f(x_k) + \mu_k I]$ θετικά ορισμένος
 - 4: $d_k \leftarrow -[\nabla^2 f(x_k) + \mu_k I]^{-1} \nabla f(x_k)$
 - 5: ορίστε γ_k τέτοιο ώστε να ικανοποιούνται τα κριτήρια 3, 4¹
 - 6: $x_k \leftarrow x_k + \gamma_k d_k$
 - 7: $k \leftarrow k + 1$
 - 8: **end while**
 - 9: **return** $x^* \leftarrow x_k$
-

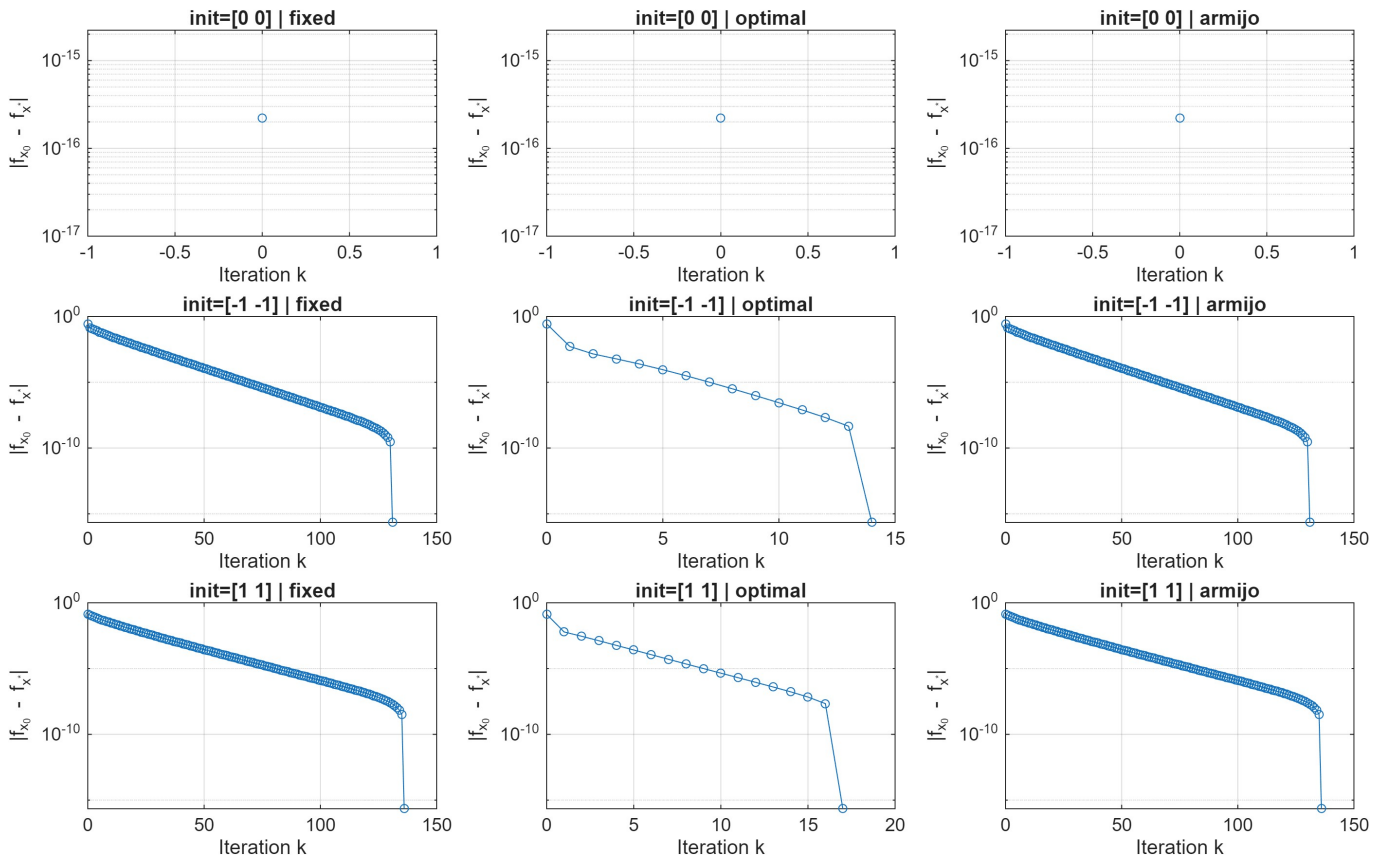
4.2 Γραφικές Παραστάσεις

Στον Πίνακα 3 και στο Σχήμα 5 παρουσιάζονται τα αποτελέσματα της μεθόδου Μέγιστης Καθόδου για τη συνάρτηση f . Καταγράφεται το ελάχιστο σημείο και οι επαναλήψεις για κάθε αρχικό σημείο και μέθοδο προσδιορισμού του βήματος.

Initial Point	Method	Minimum Point x^*	Minimum Value $f(x^*)$	k	Function Evaluations	Step Evaluations
[0 0]	fixed	[0 0]	0	0	2	0
[0 0]	optimal	[0 0]	0	0	2	0
[0 0]	armijo	[0 0]	0	0	2	0
[-1 -1]	fixed	[-1.22 -0.00832]	-0.41	131	264	0
[-1 -1]	optimal	[-1.22 -0.00803]	-0.41	14	30	168
[-1 -1]	armijo	[-1.22 -0.00832]	-0.41	131	264	262
[1 1]	fixed	[1.22 2.01]	2.92e-08	136	274	0
[1 1]	optimal	[1.22 2.03]	1.86e-08	17	36	204
[1 1]	armijo	[1.22 2.01]	2.92e-08	136	274	272

Πίνακας 3: Συγκεντρωτικός πίνακας αποτελεσμάτων (Levenberg-Marquardt)

Levenberg-Marquardt: Convergence history vs iteration k



Σχήμα 5: Σύγκλιση της f ανά τις επαναλήψεις

4.3 Συμπεράσματα

Η μέθοδος Levenberg-Marquardt συνδυάζει τα χαρακτηριστικά της μεθόδου Μέγιστης Καθόδου και Newton. Σε σημεία όπου ο Εσσιανός δεν είναι θετικά ορισμένος, ο όρος απόσβεσης $\mu_k I$ σταθεροποιεί την κατεύθυνση και επιτρέπει στον αλγόριθμο να συνεχίσει, αποφεύγοντας τις αποτυχίες της μεθόδου Newton. Τα αποτελέσματα επιβεβαιώνουν ότι η μέθοδος είναι πιο ανθεκτική και σταθερή, χωρίς όμως να εξαλείφει εντελώς τα προβλήματα που σχετίζονται με τη μορφή της συνάρτησης.

Επίδραση Αρχικού Σημείου

Αρχικό σημείο (0,0). Όπως και στις προηγούμενες μεθόδους, η διαδικασία τερματίζει αμέσως επειδή στο σημείο αυτό ισχύει $\nabla f(0,0) = 0$. Το σημείο όμως δεν είναι ελάχιστο.

Αρχικό σημείο (-1,-1). Όλες οι παραλλαγές βήματος συγκλίνουν στο ίδιο σημείο:

$$x^* \approx (-1.22, -0.008), \quad f(x^*) \approx -0.41,$$

που αποτελεί το παγκόσμιο ελάχιστο της συνάρτησης.

Αρχικό σημείο (1,1). Η μέθοδος συγκλίνει στο

$$x^* \approx (1.22, 2.0-2.03), \quad f(x^*) \approx 2 \times 10^{-8}.$$

Πρόκειται για επίπεδη περιοχή όπου η f σχεδόν μηδενίζεται, αλλά δεν αποτελεί το αρνητικό παγκόσμιο ελάχιστο της συνάρτησης. Η μέθοδος Levenberg-Marquardt, όπως και η μέθοδος Μέγιστης Καθόδου, ακολουθεί την τοπική γεωμετρία της συνάρτησης και οδηγείται σε αυτήν την επίπεδη περιοχή αντί να κινηθεί προς την κοιλότητα γύρω από $(-1.22, -0.008)$.

Επίδραση Βήματος

- **Σταθερό βήμα:** Πραγματοποιεί ≈ 130 επαναλήψεις. Η σύγκλιση είναι σταθερή, χωρίς ταλαντώσεις, αλλά σχετικά αργή.
- **Βέλτιστο βήμα:** Είναι ξεκάθαρα η πιο αποδοτική παραλλαγή. Χρειάζεται μόλις 14 επαναλήψεις και περιορισμένο πλήθος function evaluations. Η μέθοδος επωφελείται ιδιαίτερα από την εύρεση του ιδανικού μεγέθους βήματος σε κάθε επανάληψη.
- **Κανόνας Armijo:** Έχει ίδια συμπεριφορά με το σταθερό βήμα, με ≈ 130 επαναλήψεις και παρόμοια ακρίβεια. Ο Armijo τείνει να επιλέγει σχετικά μικρά βήματα, οπότε η σύγκλιση είναι σταθερή αλλά όχι γρήγορη.

Γενικά Συμπεράσματα

- Η μέθοδος Levenberg-Marquardt είναι πολύ πιο σταθερή από τη μέθοδο Newton, επειδή ο όρος απόσβεσης μ_k διορθώνει τα προβλήματα μη θετικά ορισμένου Εσσιανού.
- Παρά τη σταθερότητα, η μέθοδος παραμένει ευαίσθητη στο σημείο εκκίνησης. Από το $(1, 1)$ δεν προσεγγίζει το παγκόσμιο ελάχιστο, αλλά μια επίπεδη περιοχή με σχεδόν μηδενική τιμή της f .
- Το βέλτιστο βήμα είναι μακράν η πιο αποδοτική παραλλαγή: μειώνει τις επαναλήψεις κατά έναν παράγοντα 10 σε σχέση με τις άλλες επιλογές.
- Ο Armijo έχει συμπεριφορά σχεδόν ταυτόσημη με το σταθερό βήμα, επιλέγει συνήθως μικρά βήματα, κάτι που οδηγεί σε αργή, αλλά ασφαλή σύγκλιση.

Συνολικά, η μέθοδος Levenberg-Marquardt αποδεικνύεται η πιο σταθερή και αξιόπιστη από όλες τις μεθόδους που εφαρμόστηκαν, χωρίς όμως να εξασφαλίζει σύγκλιση στο παγκόσμιο ελάχιστο όταν το αρχικό σημείο βρίσκεται σε περιοχή όπου η συνάρτηση εμφανίζει επίπεδες γεωμετρίες.

5 Σύγκριση Μεθόδων

Σε αυτή την ενότητα θα χρησιμοποιήσουμε τα ευρήματα από τα όλα τα θέματα για να συγκρίνουμε τις μεθόδους μεταξύ τους ως προς την ορθότητα και την αποδοτικότητα.

- **Παρατήρηση 1:** Η επιλογή αρχικού σημείου επηρεάζει καθοριστικά και τις τρεις μεθόδους. Στην περιοχή του $(0, 0)$ όλες οι μέθοδοι εγκλωβίζονται σε σημείο όπου το gradient μηδενίζεται χωρίς να υπάρχει ελάχιστο.
- **Παρατήρηση 2:** Από το $(-1, -1)$ τόσο η Μέγιστη Κάθοδος όσο και η Levenberg--Marquardt συγκλίνουν στο πραγματικό ελάχιστο. Η μέθοδος Newton αποτυγχάνει λόγω μη θετικά ορισμένου Εσσιανού.
- **Παρατήρηση 3:** Από το $(1, 1)$ όλες οι μέθοδοι εκτός από τη Newton καταλήγουν σε επίπεδη περιοχή με σχεδόν μηδενική τιμή της f , αλλά όχι στο παγκόσμιο ελάχιστο. Η Newton συγκλίνει γρήγορα αλλά επίσης καταλήγει λανθασμένα.
- **Παρατήρηση 4:** Το βέλτιστο βήμα μειώνει δραστικά τις επαναλήψεις σε όλες τις μεθόδους, αλλά αυξάνει τον αριθμό αξιολογήσεων λόγω της μονοδιάστατης αναζήτησης.
- **Παρατήρηση 5:** Ο Armijo είναι σταθερός αλλά συντηρητικός: συνήθως επιλέγει μικρά βήματα και οδηγεί σε αργή αλλά σταθερή σύγκλιση. Στη Newton μπορεί να προκαλέσει πλήρη στασιμότητα.

Ακολουθεί πίνακας που συνοψίζει τα ευρήματα και τα συμπεράσματα της εργασίας αυτής:

Μέθοδος	Πλεονεκτήματα	Μειονεκτήματα	Χρήση
Μέγιστη Κάθοδος	Πάντα δίνει κατεύθυνση καθόδου, απλή υλοποίηση, δεν απαιτεί Εσσιανό.	Πολύ αργή σύγκλιση, ευαίσθησία στο βήμα, τεθλασμένη πορεία και ταλαντώσεις, εύκολος εγκλωβισμός σε επίπεδες περιοχές.	Όταν δεν υπάρχει Εσσιανός ή η συνάρτηση είναι απλή.
Newton	Πολύ γρήγορη σύγκλιση όταν $\nabla^2 f > 0$, τετραγωνική σύγκλιση κοντά στο ελάχιστο.	Αποτυγχάνει όταν $\nabla^2 f < 0$, μπορεί να απομακρυνθεί από το ελάχιστο, απαιτεί υπολογισμό και αντιστροφή Εσσιανού.	Όταν ο Εσσιανός είναι θετικά ορισμένος ή γενικά καλά συμπεριφερόμενο.
Levenberg-Marquardt	Πιο σταθερή από τη Newton, αντέχει μη θετικά ορισμένο Εσσιανό, έχει πολύ καλή αποδοτικότητα.	Μπορεί να συμπεριφερθεί σαν Steepest Descent και να είναι αργή, δεν εγγυάται σύγκλιση στο παγκόσμιο ελάχιστο, απαιτεί προσεκτική επιλογή του μ_k .	Για προβλήματα όπου η Newton είναι ασταθής.

Πίνακας 4: Συνοπτική σύγκριση των μεθόδων βελτιστοποίησης

A' Εργαλεία

Τα εργαλεία που χρησιμοποιήθηκαν κατά την εκπόνηση της εργασίας είναι τα εξής:

- MATLAB
- LaTeX
- Git

B' Βιβλιογραφία

Αναφορές

[1] Γ. Ροβιθάκης, *Τεχνικές Βελτιστοποίησης*, Εκδόσεις Τζιόλα, 2007.