ΑΡΙΣΤΟΤΕΛΕΙΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΕΣΣΑΛΟΝΙΚΗΣ

2η εργασία

«Ελαχιστοποίηση με χρήση παραγώγων»

Τεχνικές Βελτιστοποίησης

στο MATLAB

Δεκέμβριος 2022

2022-2023 | ΑΕΜ: 10265

ΣΤΑΥΡΟΣ ΚΙΗΓΜΑΣ

Φοιτητής ΤΗΜΜΥ ΑΠΘ

ΕΙΣΑΓΩΓΗ

Στην εργασία αυτή το ζητούμενο ήταν η μελέτη της ελαχιστοποίησης μιας δοσμένης συνάρτησης πολλών μεταβλητών $f: \mathbb{R}^n \to \mathbb{R}$ χωρίς περιορισμούς. Οι τρεις υπό μελέτη αλγόριθμοι βασίζονται στην ιδέα της επαναληπτικής καθόδου, σύμφωνα με την οποία ξεκινώντας από κάποιο αρχικό σημείο $x_0 \in \mathbb{R}^n$ παράγουμε διαδοχικά τα διανύσματα $x_1, x_2, \dots$ έτσι ώστε $f(x_{k+1}) < f(x_k), k = 1,2,3, \dots$

Οι αλγόριθμοι αναζήτησης που θα μελετήσουμε είναι:

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

Και οι τρεις αλγόριθμοι κάνουν χρήση παραγώγων στην υλοποίησή τους.

Η αντικειμενική συνάρτηση που ζητείται να χρησιμοποιήσουμε είναι η:

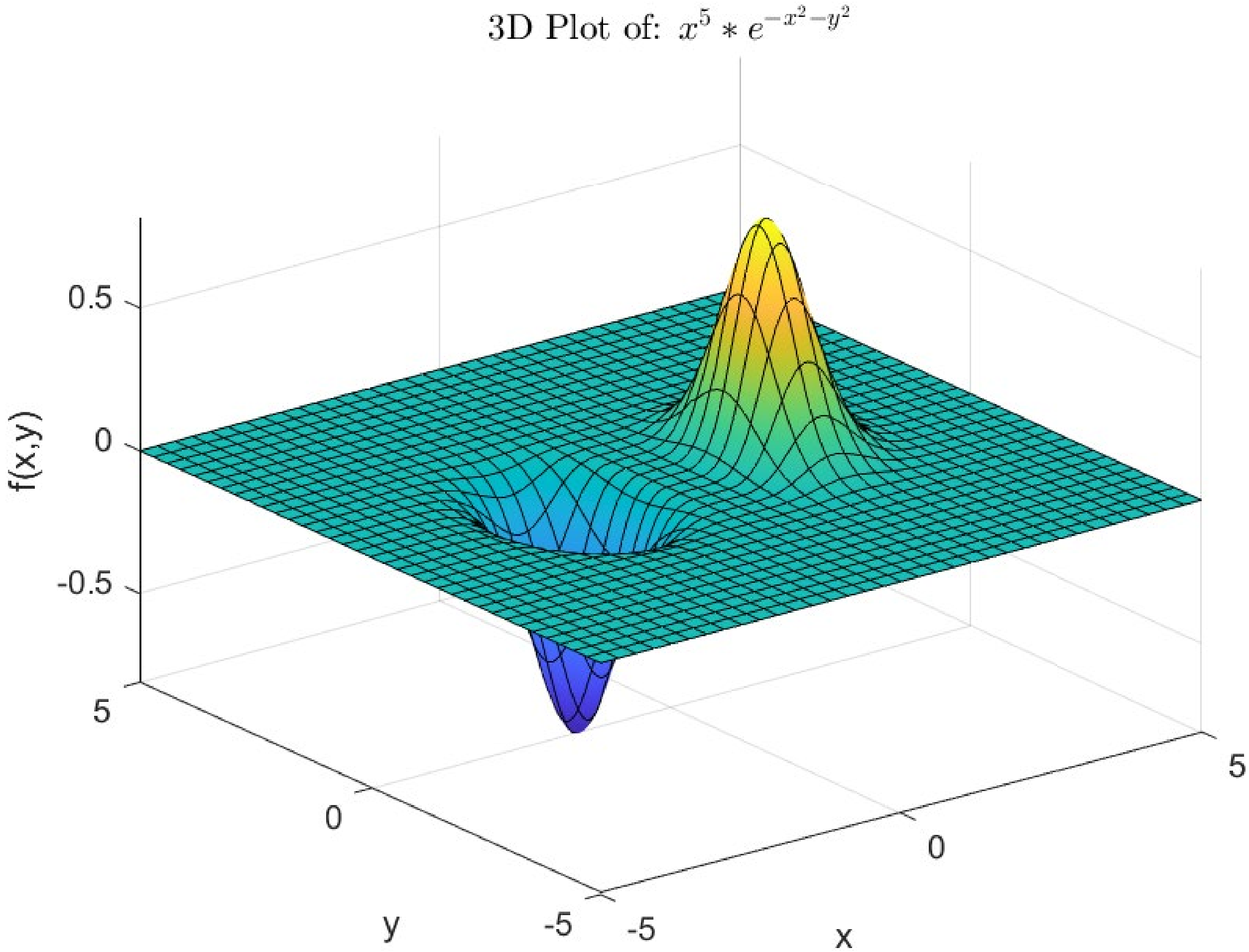
- $f(x, y) = x^5 e^{-x^2-y^2}$

Τα αρχικά σημεία που θα ελέγξουμε ορίζονται ως τα $(0,0)$, $(-1,1)$ $και$ $(1,-1)$ σύμφωνα με την εκφώνηση της εργασίας. Το βήμα γ_κ ζητείται να επιλεχθεί α) σταθερό, β) τέτοιο ώστε να ελαχιστοποιεί την $f(x_k + \gamma_k d_k)$, και γ) βάσει του κανόνα Armijo και στις τρεις αναλύσεις που θα ακολουθήσουν.

Η υλοποίηση όλων των μεθόδων είναι βασισμένη στο κεφάλαιο 5.1 του βιβλίου *Τεχνικές Βελτιστοποίησης.*

ΘΕΜΑ 1

Στο πρώτο θέμα σχεδιάζουμε την f, έτσι ώστε να έχουμε μια οπτική επαφή με το πρόβλημά μας και την μορφή της αντικειμενικής συνάρτησης.



Διάγραμμα 1 - Τρισδιάστατη απεικόνιση της αντικειμενικής συνάρτησης

Χρησιμοποιώντας την συνάρτηση `fsurf` (εναλλακτικά την `fmesh`) παρατηρούμε πως υπάρχει μια κοιλότητα και ένα όρος για τιμές του x $\{-1.581, 1.581\}$ αντίστοιχα, ενώ το y είναι στο 0 και στις δύο περιπτώσεις. Για απόλυτες τιμές του x λίγο μεγαλύτερες των προαναφερθεισών, επομένως, οι αλγόριθμοι μας θα επιφέρουν λάθος αποτελέσματα εξαιτίας της μηδενικής κλίσης και άρα μηδενικής παραγώγου εκεί.

ΘΕΜΑ 2

Στο δεύτερο θέμα πραγματοποιήθηκε η υλοποίηση της μεθόδου της Μέγιστης Καθόδου. Ζητήθηκε να ελαχιστοποιήσουμε την αντικειμενική συνάρτηση f χρησιμοποιώντας ως αρχικά σημεία (x_0, y_0) τα i) $(0{,}0)$, ii) $(-1{,}1)$, iii) $(1, -1)$. Αναλύθηκαν περιπτώσεις διαφορετικών διαδικασιών επιλογής του βήματος γ_k. Πιο συγκριμένα:

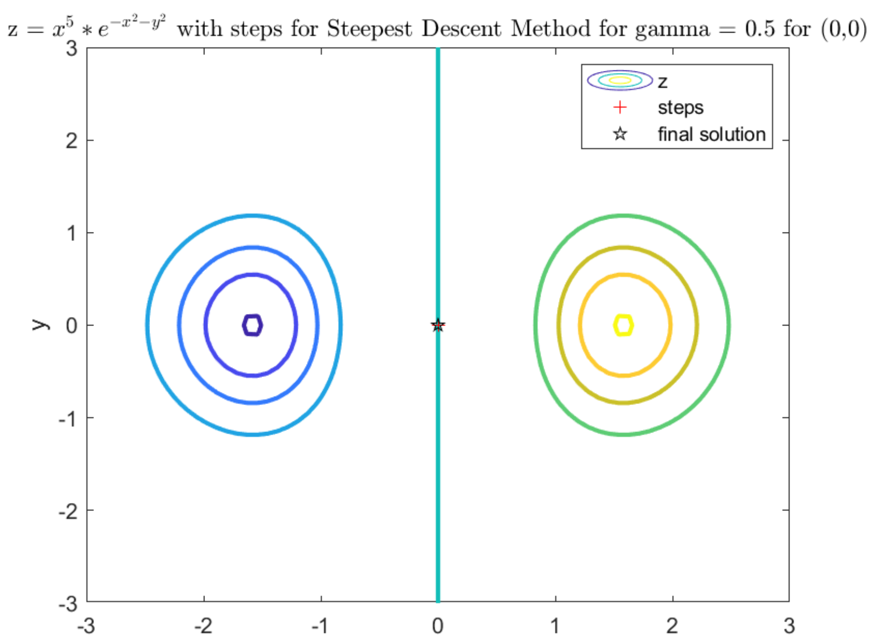
- ***Περίπτωση α***: Το βήμα επιλέγεται σταθερό (της επιλογής μας).
- ***Περίπτωση β***: Το βήμα επιλέγεται τέτοιο ώστε να ελαχιστοποιεί την $f(x_k + \gamma_k * d_k)$.
- ***Περίπτωση γ***: Το βήμα επιλέγεται βάσει του κανόνα Armijo.

Παρατηρώντας περαιτέρω το *Διάγραμμα 1* μπορούμε να υποθέσουμε πως γύρω από το σημείο $(0{,}0)$, αλλά και στις άλλες περιοχές που το σχήμα είναι επίπεδο είναι εξαιρετικά πιθανό οι αλγόριθμοι να μην αποδίδουν καλά. Τουναντίον, στα άλλα δυο σημεία έναρξης του αλγορίθμου, περιμένουμε ελαφρώς καλύτερα αποτελέσματα, αφού είναι κοντά στο όρος και στην κοιλότητα.

Σχετικά με το γ_k, επιλέχθηκε το 0.5 μετά από κάποιες δοκιμές μεταξύ αυτού και άλλων τιμών.

- ***Περίπτωση α – Σταθερό γ_k***

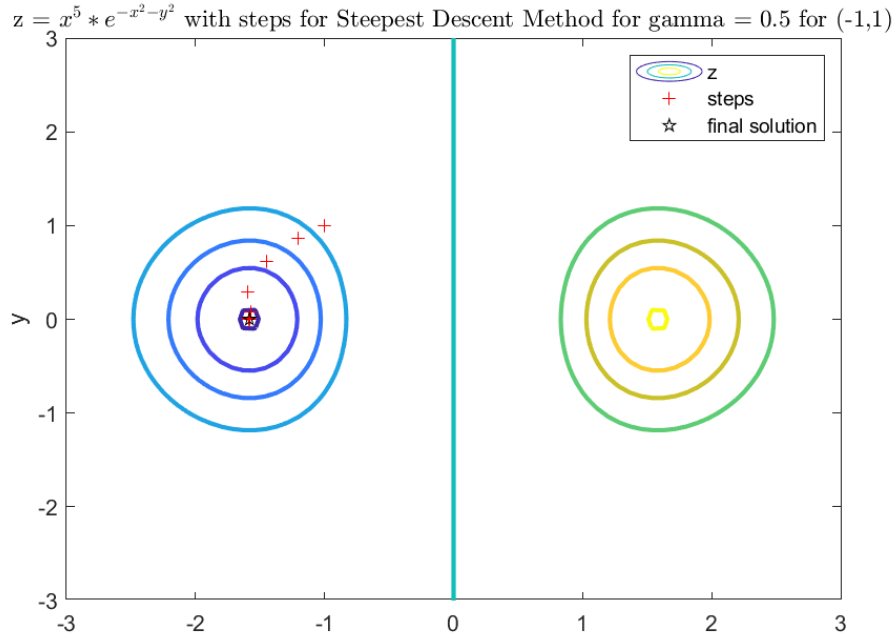
i) Αρχικό σημείο: (0,0)



Διάγραμμα 2

Πράγματι, ξεκινώντας από το σημείο (0,0) σε κάθε επανάληψη θα παραμένουμε σε αυτό, εξαιτίας της μηδενικής κλίσης όπως είπαμε.

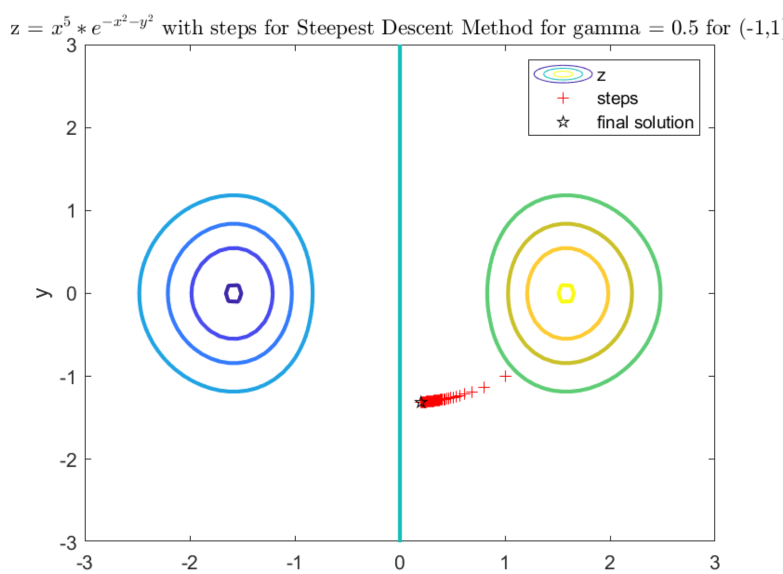
ii) Αρχικό σημείο: (-1,1)



Διάγραμμα 3

Από αυτό το σημείο ο αλγόριθμος λειτουργεί με μεγάλη αποτελεσματικότητα. Μετά από μόλις 5 επαναλήψεις έχει καταλήξει στο σωστό σημείο και έπειτα κινείται γύρω από αυτό.

iii) Αρχικό σημείο: (1,-1)

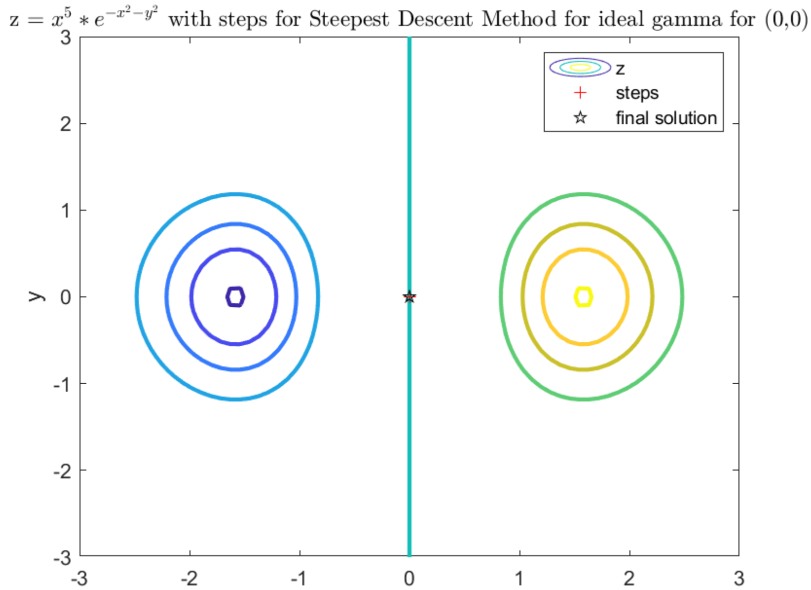


Διάγραμμα 4

Σε αυτήν την περίπτωση ο αλγόριθμος δεν κατάφερε να «ανεβεί» στο όρος, άλλα συνέκλινε προς ένα σημείο του οριζόντιου επιπέδου. Μετά από πολλές επαναλήψεις δεν αλλάζει το α-ποτέλεσμα.

- **Περίπτωση β – Βήμα γ_k που ελαχιστοποιεί την $f(x_k + \gamma_k * d_k)$**

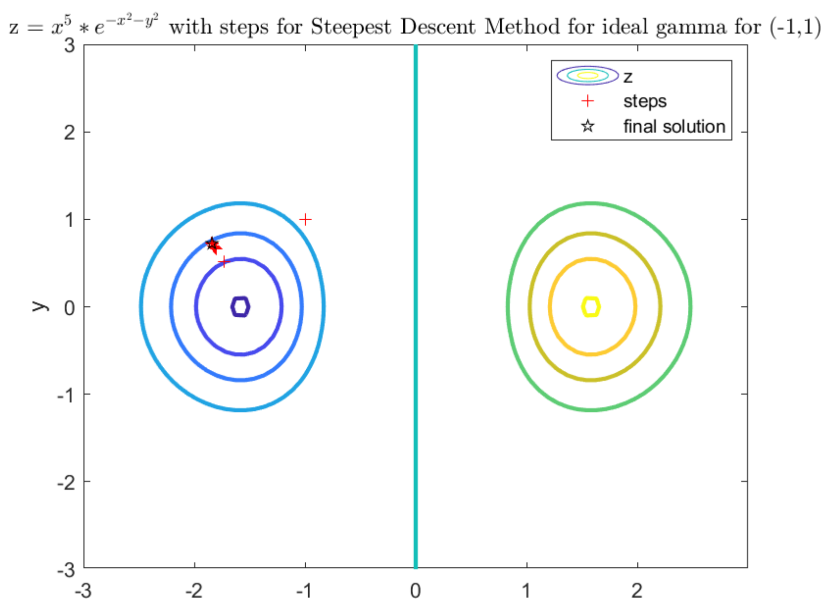
i) Αρχικό σημείο: (0,0)



Διάγραμμα 5

Όπως και στην προηγούμενη περίπτωση, για αρχικό σημείο το $(0,0)$ ο αλγόριθμος παραμένει σε αυτό.

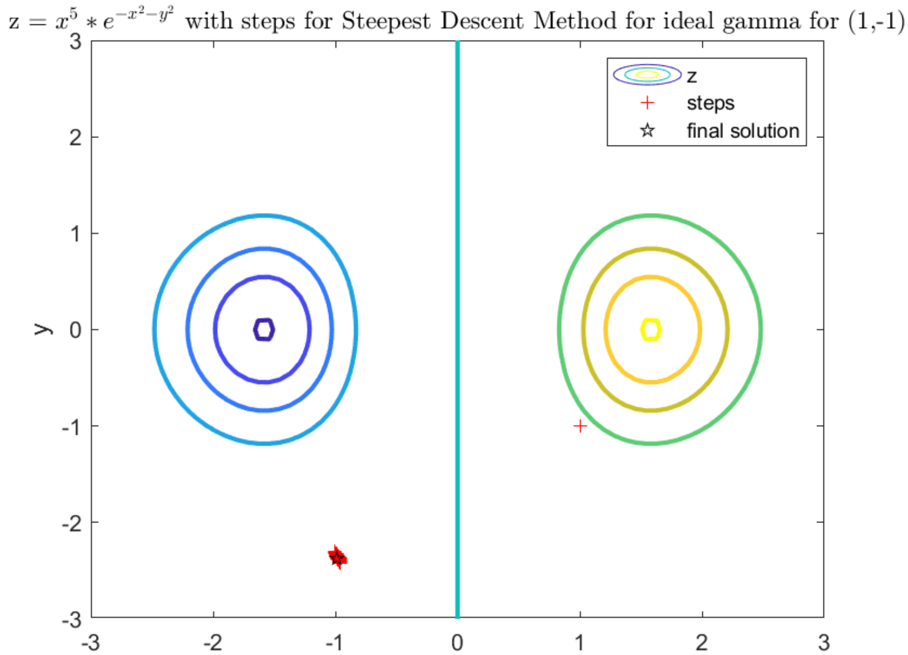
ii) Αρχικό σημείο: (-1,1)



Διάγραμμα 6

Ήδη από την δεύτερη επανάληψη ο αλγόριθμος συγκλίνει σε ένα σημείο εντός της κοιλότητας, αλλά όχι στο ελάχιστο. Μετά από 500 επαναλήψεις δεν άλλαξε κάτι.

iii) Αρχικό σημείο: (1,-1)

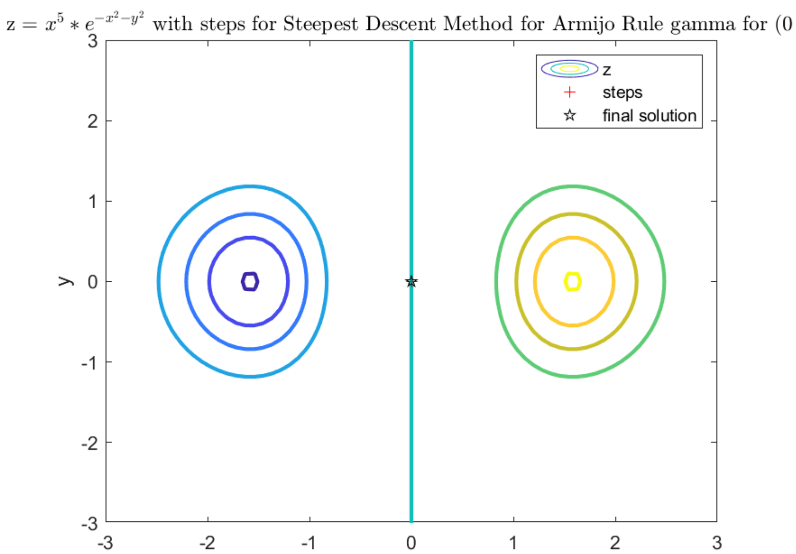


Διάγραμμα 7

Από την πρώτη επανάληψη ο αλγόριθμος καταλήγει σε ένα πολύ μακρινό σημείο.

- **Περίπτωση γ – Βήμα γ_k σύμφωνα με τον κανόνα Armijo**

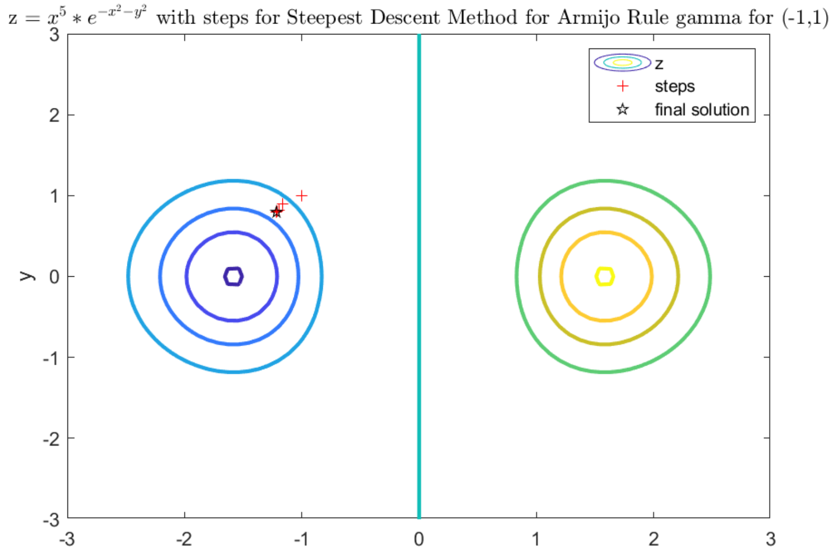
i) Αρχικό σημείο: (0,0)



Διάγραμμα 8

Αναμενόμενα, παραμένουμε στο (0,0).

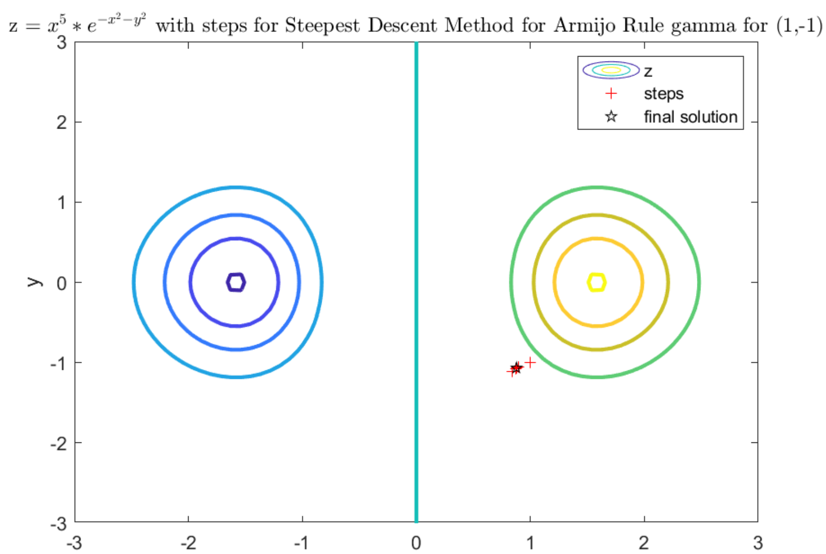
ii) *Αρχικό σημείο: (-1,1)*



Διάγραμμα 9

Εδώ, τα αποτελέσματα είναι παρόμοια με την *Περίπτωση β*. Ο αλγόριθμος συγκλίνει εντός κοιλότητας, αλλά όχι στο σημείο ελαχίστου από τις πρώτες επαναλήψεις.

iii) *Αρχικό σημείο: (1,-1)*



Διάγραμμα 10

Για άλλη μια φορά αρχίζοντας από αυτό το σημείο καταλήγουμε εκτός όρους και μάλιστα κοντά στο σημείο έναρξής μας.

ΘΕΜΑ 3

Στο τρίτο θέμα πραγματοποιήθηκε η υλοποίηση της μεθόδου Newton. Ζητήθηκε ξανά να ελαχιστοποιήσουμε την αντικειμενική συνάρτηση f χρησιμοποιώντας ως αρχικά σημεία (x_0, y_0) τα i) $(0,0)$, ii) $(-1,1)$, iii) $(1,-1)$. Ομοίως, αναλύθηκαν περιπτώσεις διαφορετικών διαδικασιών επιλογής του βήματος γ_k. Πιο συγκριμένα:

- ***Περίπτωση α***: Το βήμα επιλέγεται σταθερό (της επιλογής μας).
- ***Περίπτωση β***: Το βήμα επιλέγεται τέτοιο ώστε να ελαχιστοποιεί την $f(x_k + \gamma_k * d_k)$.
- ***Περίπτωση γ***: Το βήμα επιλέγεται βάσει του κανόνα Armijo.

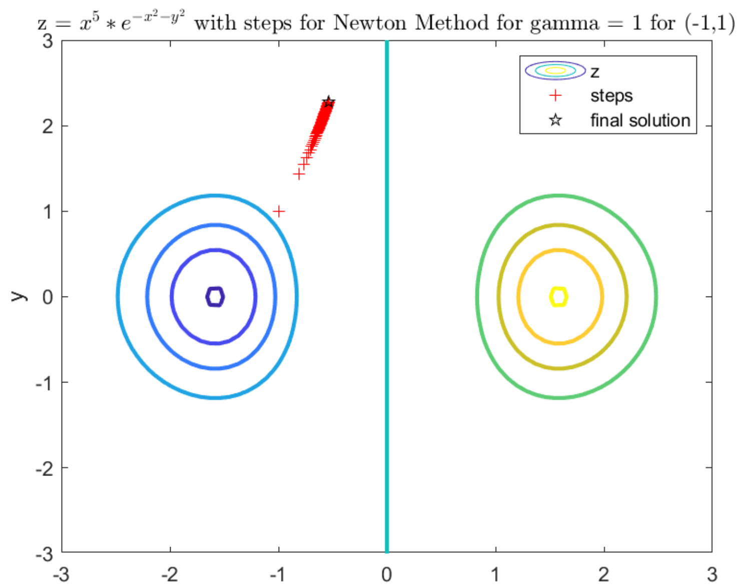
Σε αυτήν την μέθοδο κάνουμε χρήση του Εσσιανού πίνακα, δηλαδή του πίνακα των δεύτερων παραγώγων της συνάρτησης μας. Αυτός πρέπει να είναι θετικά ορισμένος ή ακόμα και θετικά ημιορισμένος.

- ***Περίπτωση α – Σταθερό γ_k***

i) Αρχικό σημείο: (0,0)

Για αυτό το σημείο ο πίνακας δεν είναι θετικά ορισμένος ή ημιορισμένος.

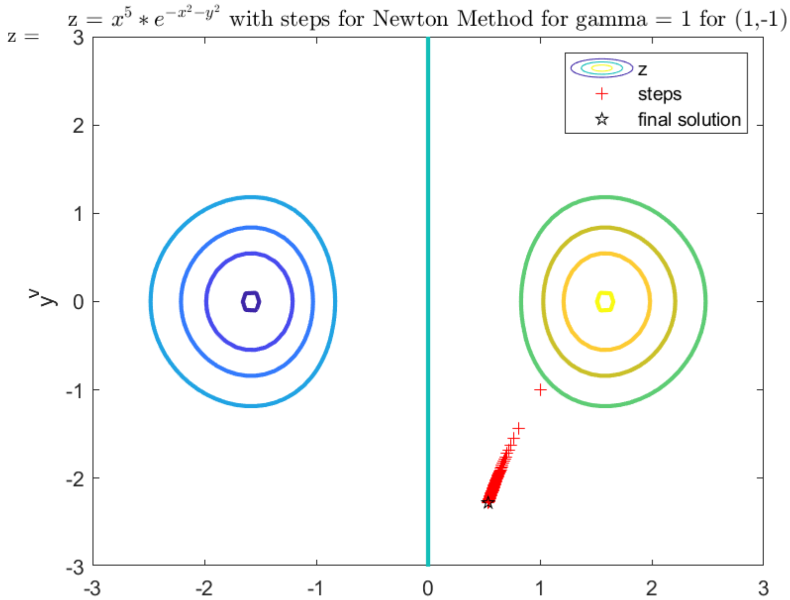
ii) Αρχικό σημείο: (-1,1)



Διάγραμμα 11

Από αυτό το σημείο ο αλγόριθμος δεν λειτουργεί με μεγάλη αποτελεσματικότητα. Συγκλίνει σε ένα σημείο μακριά από το επιθυμητό.

iii) Αρχικό σημείο: (1,-1)



Διάγραμμα 12

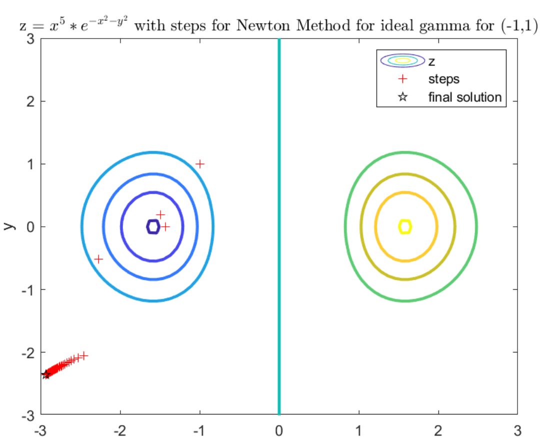
Σε αυτήν την περίπτωση ο αλγόριθμος δεν κατάφερε να «ανεβεί» στο όρος, άλλα συνέκλινε προς ένα σημείο του οριζόντιου επιπέδου. Μετά από πολλές επαναλήψεις δεν αλλάζει το α-ποτέλεσμα.

- **Περίπτωση β – Βήμα γ_k που ελαχιστοποιεί την $f(x_k + \gamma_k * d_k)$**

i) *Αρχικό σημείο: (0,0)*

Για αυτό το σημείο ο πίνακας δεν είναι θετικά ορισμένος ή ημιορισμένος.

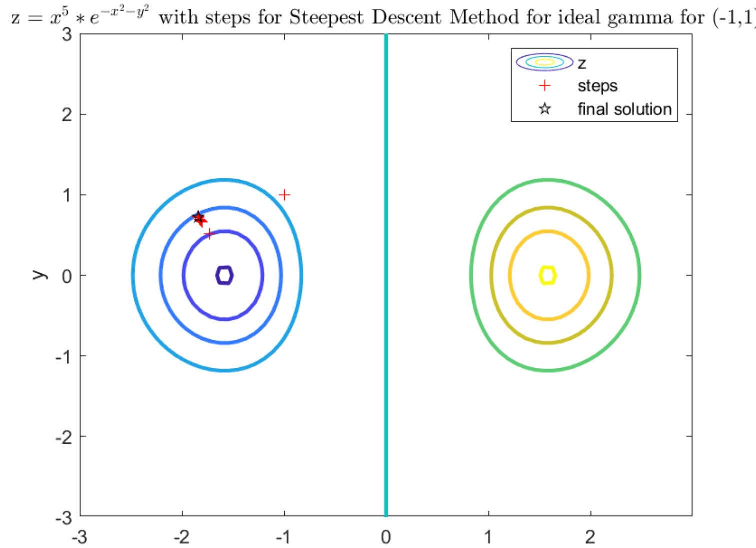
ii) *Αρχικό σημείο: (-1,1)*



Διάγραμμα 13

Εδώ, τα αποτελέσματα είναι πρωτοφανή. Ενώ αρ-χικά ο αλγόριθμος ήταν κοντά τελικά κατέληξε πολύ μακριά από το ελάχιστο.

iii) *Αρχικό σημείο: (1,-1)*



Διάγραμμα 14

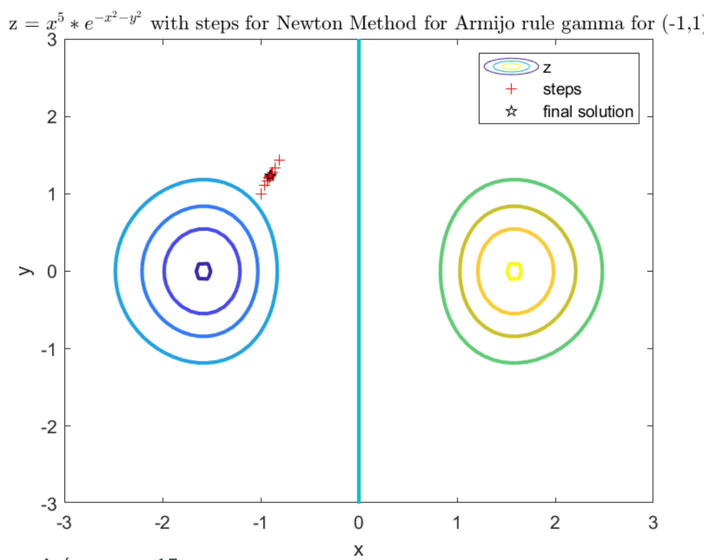
Από την πρώτη επανάληψη ο αλγόριθμος καταλήγει εντός κοιλότητας, όχι όμως εκεί που πρέ-πει.

- **Περίπτωση γ – Βήμα γ_k σύμφωνα με τον κανόνα Armijo**

i) *Αρχικό σημείο: (0,0)*

Για αυτό το σημείο ο πίνακας δεν είναι θετικά ορισμένος ή ημιορισμένος.

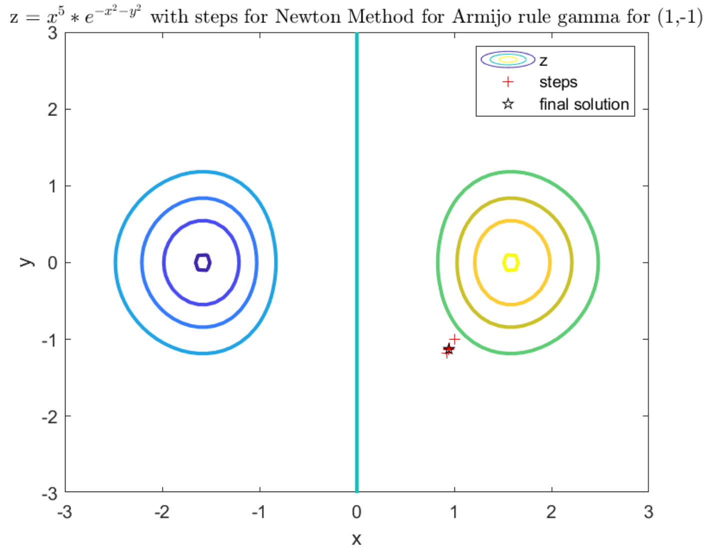
ii) *Αρχικό σημείο: (-1,1)*



Διάγραμμα 15

Εδώ, ο αλγόριθμος δεν μπαίνει ποτέ στην κοιλότητα.

iii) Αρχικό σημείο: (1,-1)



Διάγραμμα 16

Για άλλη μια φορά αρχίζοντας από αυτό το σημείο καταλήγουμε εκτός όρους και μάλιστα κοντά στο σημείο έναρξής μας.

.

ΘΕΜΑ 4

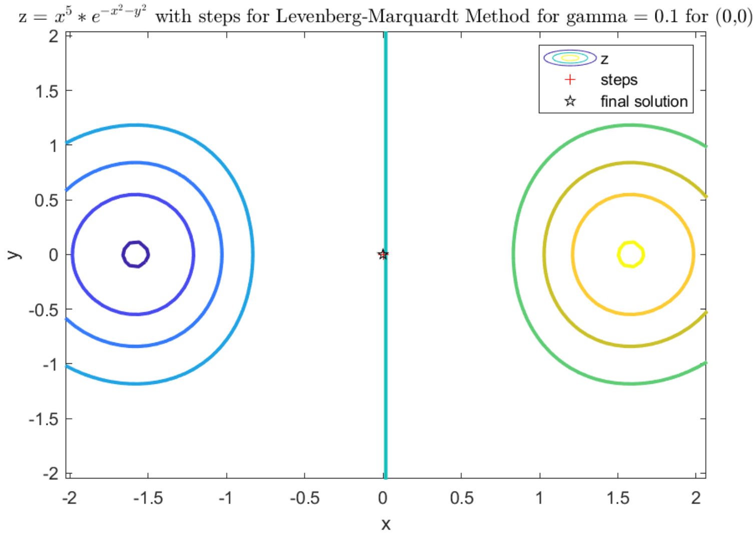
Στο τέταρτο θέμα πραγματοποιήθηκε η υλοποίηση της μεθόδου *Levenberg-Marquardt*. Πρόκειται για μια τροποποιημένη εκδοχή της μεθόδου *Newton* συνδυάζοντας στοιχεία της *Μέγιστης Καθόδου*. Ζητήθηκε ξανά να ελαχιστοποιήσουμε την αντικειμενική συνάρτηση f χρησιμοποιώντας ως αρχικά σημεία (x_0, y_0) τα i) $(0,0)$, ii) $(-1,1)$, iii) $(1,-1)$. Όπως και πριν, αναλύθηκαν περιπτώσεις διαφορετικών διαδικασιών επιλογής του βήματος γ_k. Πιο συγκριμένα:

- ***Περίπτωση α***: Το βήμα επιλέγεται σταθερό (της επιλογής μας).
- ***Περίπτωση β***: Το βήμα επιλέγεται τέτοιο ώστε να ελαχιστοποιεί την $f(x_k + \gamma_k * d_k)$.
- ***Περίπτωση γ***: Το βήμα επιλέγεται βάσει του κανόνα Armijo.

Σε αυτήν την μέθοδο κάνουμε επίσης χρήση του Εσσιανού πίνακα, δηλαδή του πίνακα των δεύτερων παραγώγων της συνάρτησης μας. Εδώ στην περίπτωση που δεν είναι θετικά ορισμένος ή ημιορισμένος δεν υπάρχει πρόβλημα.

- ***Περίπτωση α – Σταθερό γ_k***

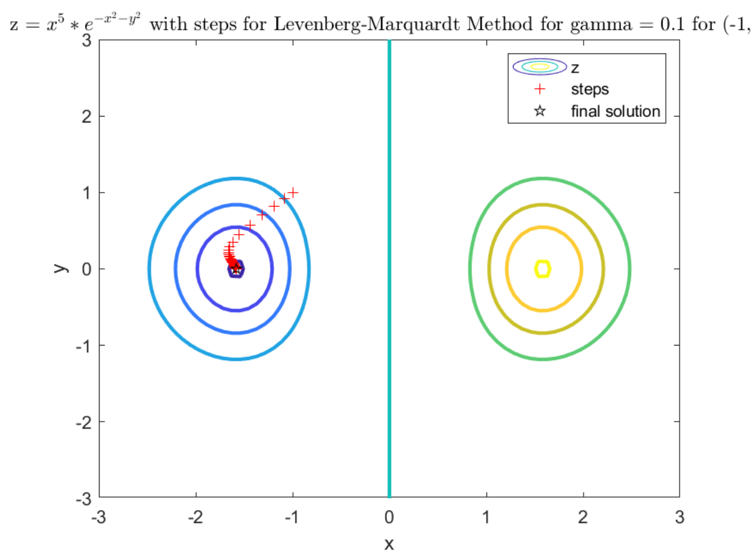
i) Αρχικό σημείο: (0,0)



Διάγραμμα 17

Για αυτό το σημείο, όπως συνήθως μένουμε σταθεροί.

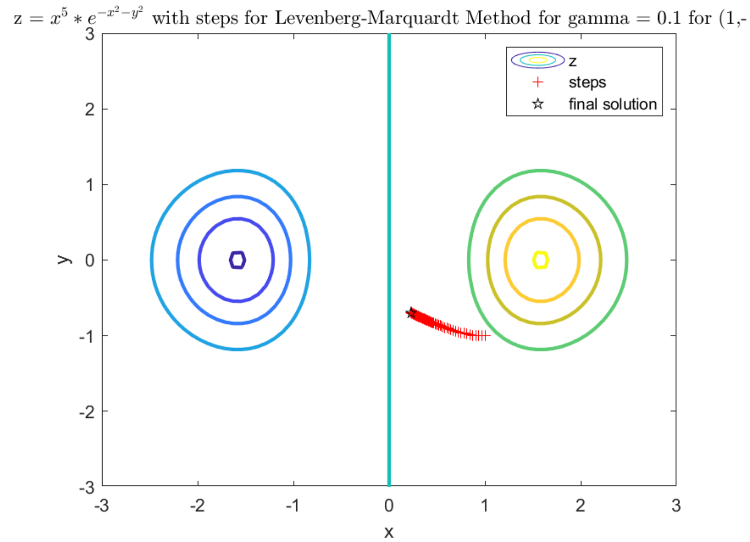
ii) Αρχικό σημείο: (-1,1)



Διάγραμμα 18

Εδώ βλέπουμε πως η μέθοδος λειτουργεί άριστα συγκλίνοντας στο σημείο ελαχίστου. Μετά την 8^{η} επανάληψη άρχισε να καταλήγει.

iii) Αρχικό σημείο: (1,-1)

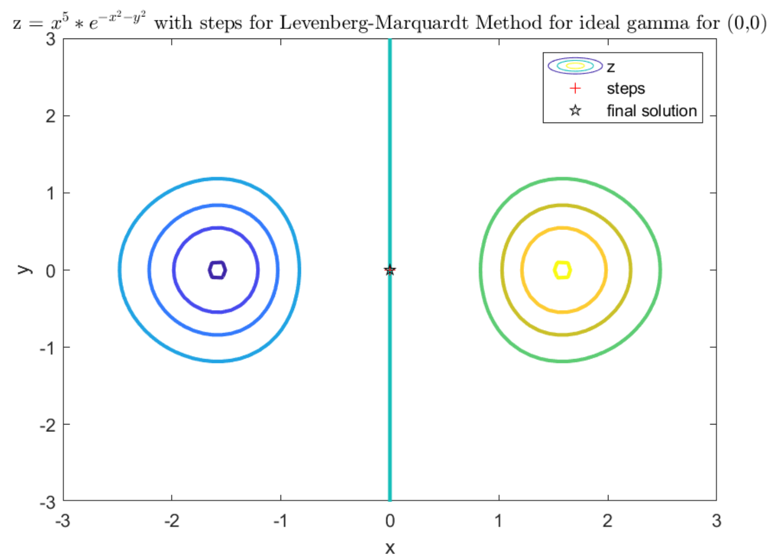


Διάγραμμα 19

Σε αυτήν την περίπτωση ο αλγόριθμος δεν κατάφερε να «ανεβεί» στο όρος, άλλα συνέκλινε προς ένα σημείο του οριζόντιου επιπέδου. Μετά από πολλές επαναλήψεις δεν αλλάζει το α-ποτέλεσμα. Αυτό συνέβη στις περισσότερες εφαρμογές των μεθόδων που άρχισαν από αυτό το σημείο.

- **Περίπτωση β – Βήμα γ_k που ελαχιστοποιεί την $f(x_k + \gamma_k * d_k)$**

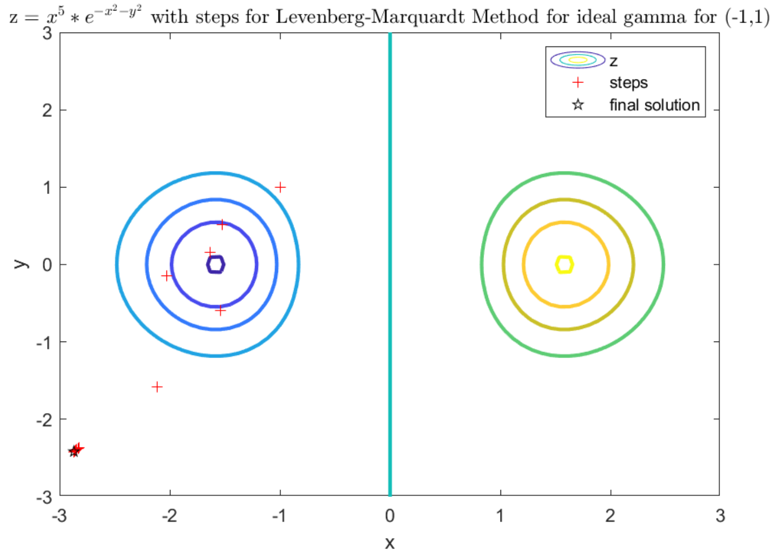
i) Αρχικό σημείο: (0,0)



Διάγραμμα 20

Για αυτό το σημείο ο αλγόριθμος δεν προχωράει πέρα από το σημείο εκκίνησής του.

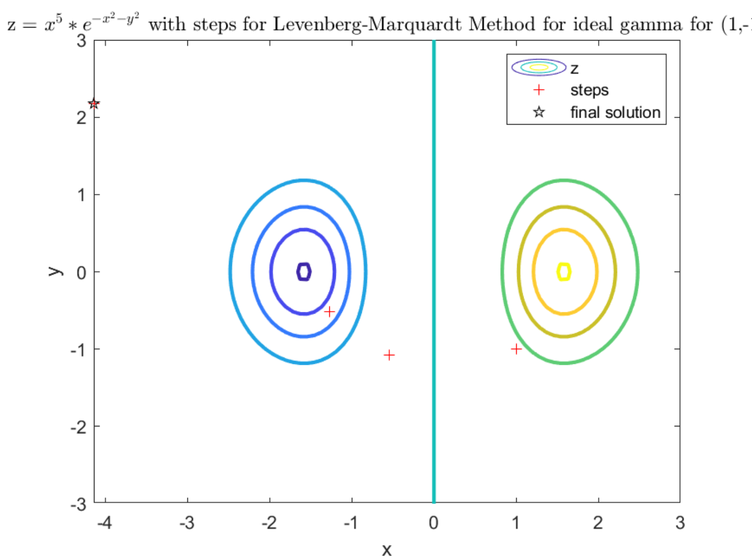
ii) Αρχικό σημείο: (-1,1)



Διάγραμμα 21

Εδώ, τα αποτελέσματα είναι ακριβώς ίδια με αυτά της *Μεθόδου Newton*. Εδώ φαίνεται ξεκάθαρα η σχέση των δύο αυτών μεθόδων, καθώς διαχειρίστηκαν την κλίση σε κάθε σημείο με τον ίδιο τρόπο.

iii) Αρχικό σημείο: (1,-1)

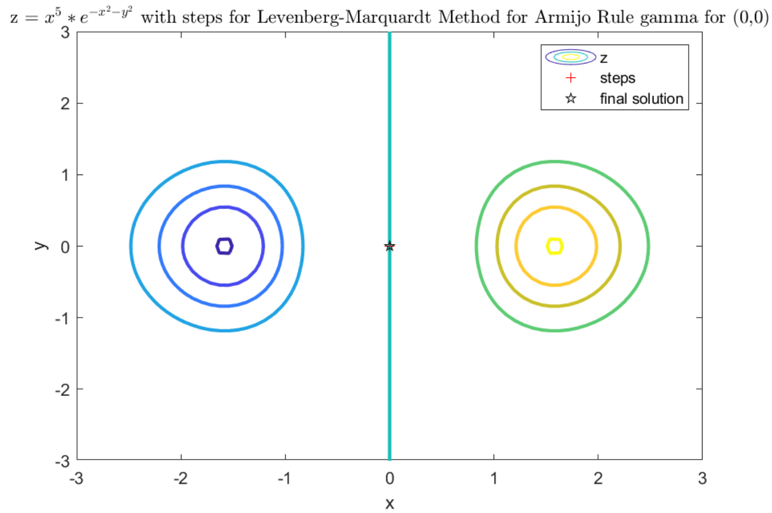


Διάγραμμα 22

Τώρα, ο αλγόριθμος είναι τελείως εκτός σωστής λειτουργίας. Βλέπουμε πως μετά την $4^η$ επανάληψη έχει φτάσει σε ένα σημείο πολύ μακρινό του ελαχίστου.

- **Περίπτωση γ – Βήμα γ_k σύμφωνα με τον κανόνα Armijo**

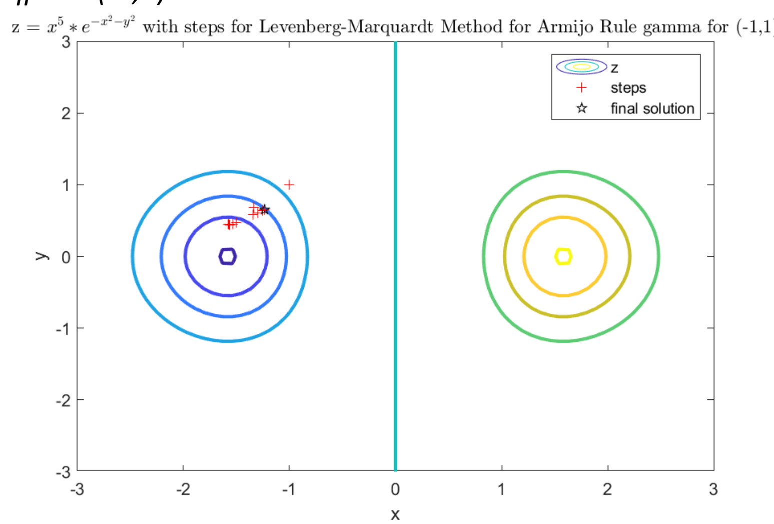
i) Αρχικό σημείο: (0,0)



Διάγραμμα 23

Για αυτό το σημείο ο αλγόριθμος δεν προχωράει πέρα από το σημείο εκκίνησής του.

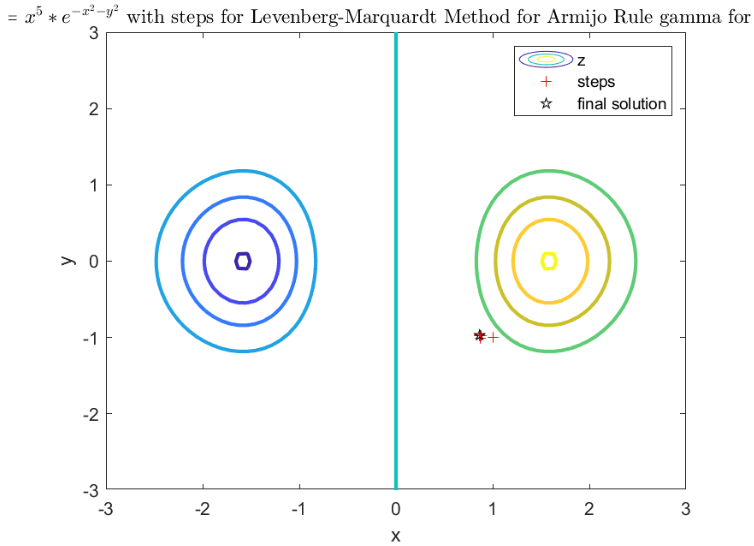
ii) Αρχικό σημείο: (-1,1)



Διάγραμμα 24

Εδώ, ο αλγόριθμος προσπάθησε να καταλήξει στο σωστό σημείο, αλλά τελικά έμεινε αρκετά κοντά, εντός της κοιλότητας.

iii) *Αρχικό σημείο: (1,-1)*



Διάγραμμα 25

Για άλλη μια φορά αρχίζοντας από αυτό το σημείο καταλήγουμε εκτός όρους και μάλιστα κοντά στο σημείο έναρξής μας. Παρόμοια κατάληξη με τις άλλες μεθόδους. Απλά σε αυτήν την περίπτωση μείναμε πολύ πιο κοντά.

ΣΥΓΚΡΙΤΙΚΟΣ ΣΧΟΛΙΑΣΜΟΣ

Μετά των παραπάνω αναλύσεων συμπεραίνουμε πως το δεύτερο σημείο εκκίνησης των αλγορίθμων επιφέρει πιο ορθολογικά αποτελέσματα, οπότε και η συγκριτική ανάλυση θα γίνει ξεκινώντας από αυτό.

- **Μέθοδος Μέγιστης Καθόδου**

Αρχικά μελετήσαμε τον αριθμό των επαναλήψεων συναρτήσει του γ_k για την περίπτωση που είναι σταθερό, έτσι ώστε να επιλέξουμε το κατάλληλο. Επιλέξαμε το 0.5.

γ	0.1	0.25	0.5	0.75	1.00
Reps	41	15	5	9	36

Τώρα θα συγκρίνουμε τις επιδόσεις των 3 τρόπων επιλογής του βήματος γ_k.

	Constant step	**Ideal step**	**Armijo step**
γ	0.5	-	-
Reps	5	6	4

Φαίνεται πως και οι τρεις για το σημείο $(-1,1)$ είναι αρκετά καλές.

- ## Μέθοδος Newton

Πάλι ξεκινάμε επιλέγοντας το κατάλληλο σταθερό γ_k. Αυτό είναι το

γ	0.1	0.25	0.5	0.75	1.00
Reps	85	82	82	82	85

Η επιλογή του βήματος δεν επηρεάζει τον αριθμό των επαναλήψεων. Είναι πολλές ανεξάρτητα από αυτό.

	Constant step	Ideal step	Armijo step
γ	1	-	-
Reps	85	77	10

Με μεγάλη διαφορά ακολουθώντας τον κανόνα του Armijo έχουμε λιγότερες επαναλήψεις.

- ## Μέθοδος Levenberg-Marquardt

Αυτή τη φορά μας βολεύει το μικρό βήμα.

γ	0.1	0.25	0.5	0.75	1.00
Reps	13	20	20	97	99

	Constant step	Ideal step	Armijo step
γ	0.1	-	-
Reps	13	5	11

Οι επιδόσεις για αυτή τη μέθοδο είναι καλές, ωστόσο για το βέλτιστο βήμα είναι η καλύτερη.

ΑΡΧΕΙΑ ΕΡΓΑΣΙΑΣ

Στο συμπιεσμένο φάκελο που ανέβηκε, εκτός από την αναφορά αυτή υπάρχουν τέσσερις φάκελοι συνδεδεμένοι με τα τέσσερα θέματα που ζητήθηκαν.

Στον φάκελο `1. Function Plot` υπάρχει η συνάρτηση functionPlot η οποία αναπαριστά γραφικά την αντικειμενική συνάρτηση f.

Στον φάκελο `2. Steepest Descent Method` περιλαμβάνεται η υλοποίηση της ομώνυμης μεθόδου. Τα αρχεία `steepestDescentConstGamma`, `steepestDescentMinGamma`, `steepestDescentArmijo` διαφέρουν στην επιλογή του βήματος γ_k, σύμφωνα με τις αντίστοιχες διαδικασίες του βιβλίου. Γενικά, σε όλες τις μεθόδους χρειάζεται να υπολογίζουμε το *gradient* της συνάρτησης και έπειτα το διάνυσμα σε κάθε σημείο που βρισκόμαστε και αυτό γίνεται ορίζοντας την `gradientVector`. Στους υπόλοιπους φακέλους γίνεται κλήση αυτής. Αναφορικά με την δεύτερη διαδικασία επιλογής του βήματος κάνουμε χρήση του `goldenSection` και του `nextGammaFinder` ώστε να βρούμε την τιμή του βήματος που πρέπει. Στο `Armijo` γίνεται η εφαρμογή του κανόνα του Armijo. Τέλος, στο `steepestDescentPlots` δημιουργούνται τα γραφήματά μας.

Στον φάκελο `3. Newton Method` αρχικά ορίζεται ο υπολογισμός του Εσσιανού πίνακα στο αρχείο `hessianMatrix`. Στα `newtonConstGamma`, `newtonMinGamma`, `newtonArmijo` γίνεται η υλοποίηση της *Μεθόδου Newton* διαφοροποιώντας την επιλογή του βήματος. Στο `newtonPlots` σχηματίζονται τα διαγράμματα που μελετήσαμε.

Ο φάκελος `4. Levenberg-Marquardt Method` περιλαμβάνει αντίστοιχα τα αρχεία για αυτήν την μέθοδο. Όπως και πριν, καθορίζουμε την επιλογή του βήματος στα `levenbergConstGamma`, `levenbergMinGamma`, `levenbergArmijo`, ενώ στο `levenbergPlots` γίνονται τα γραφήματα.