



Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
Πολυτεχνική Σχολή
Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Τομέας Ηλεκτρονικής

Τεχνικές Βελτιστοποίησης - 2η Εργασία

Θεόδωρος Παπαφωτίου
ΑΕΜ: 9708
papafotit@ece.auth.gr

7 Δεκεμβρίου 2021

Περιεχόμενα

1 Περιγραφή 1ης Εργασίας	2
1.1 Μέθοδοι	2
1.2 Υλοποίηση	2
1.3 Μελέτη	2
2 Μελέτη Μεθόδων	5
2.1 Μέθοδος Μέγιστης Καθόδου (Steepest Descent)	6
2.1.1 Σύντομη περιγραφή της μεθόδου	6
2.1.2 Γραφήματα	7
2.2 Μέθοδος Newton	14
2.2.1 Σύντομη περιγραφή της μεθόδου	14
2.2.2 Γραφήματα	14
2.3 Μέθοδος Levenberg-Marquardt	18
2.3.1 Σύντομη περιγραφή της μεθόδου	18
2.3.2 Γραφήματα	19
2.4 Γενικά συμπεράσματα	26
2.4.1 Αρχικό σημείο	26
2.4.2 Αριθμός επαναλήψεων	27
3 Σύγκριση Μεθόδων	29
3.1 Γενικές παρατηρήσεις	29
3.2 Επιλογή του γ	32

Κεφάλαιο 1

Περιγραφή 1ης Εργασίας

Στόχος της παρούσας εργασίας είναι η υλοποίηση και μελέτη μεθόδων **κλίσης** για την **εύρεση ελαχίστου** μια δωσμένης συνάρτησης πολλών μεταβλητών **χωρίς περιορισμούς**.

1.1 Μέθοδοι

Οι μέθοδοι που καλούμαστε να υλοποιήσουμε και να μελετήσουμε είναι οι παρακάτω:

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

Σε όλες τις παραπάνω μεθόδους η μελέτη γίνεται σε συνάρτηση με 2 μεταβλητές $f(x, y)$ χωρίς περιορισμούς. Η συνάρτηση f χρειάζεται αν είναι τουλάχιστον 1 φορά παραγωγίσιμη για την πρώτη μέθοδο, ενώ χρειάζεται να είναι τουλάχιστον 2 φορές παραγωγίσιμη για τις υπόλοιπες 2 μεθόδους. Με τη χρησιμοποίηση ενός ακολουθιακού αλγόριθμου καταλήγουμε σε μια κλίση $\nabla f(x_k)$ της συνάρτησης στην k -στη επανάληψη η οποία τείνει στο μηδέν (0) με προδιαγεγραμμένη **ακρίβεια** ϵ , δηλαδή $\nabla f(x_k) \leq \epsilon$.

1.2 Υλοποίηση

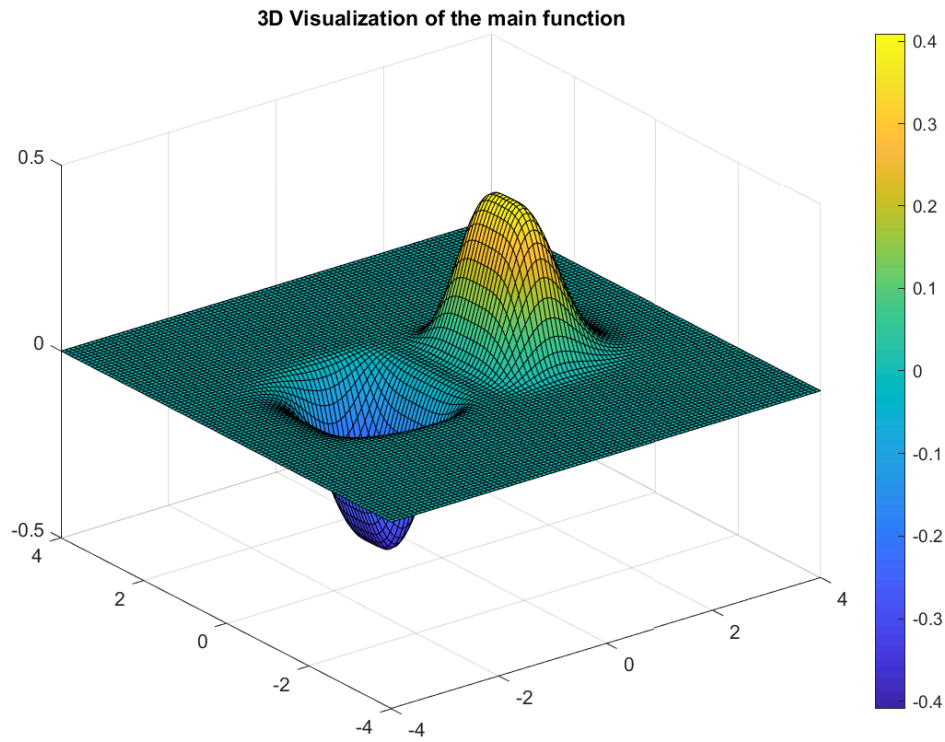
Η υλοποίηση των παραπάνω μεθόδων πραγματοποιήθηκε με τη χρήση του λογισμικού **MATLAB** (v.2019b). Για κάθε μέθοδο έχει υλοποιηθεί η αντίστοιχη συνάρτηση (function), καθώς και ένα αρχείο (Project2_i, $i = 2, 3, 4$) που καλεί τη συνάρτηση αυτή. Όλα τα αρχεία επισυνάπτονται με την παρούσα τεχνική αναφορά στο αποθετήριο της πλατφόρμας eLearning.

1.3 Μελέτη

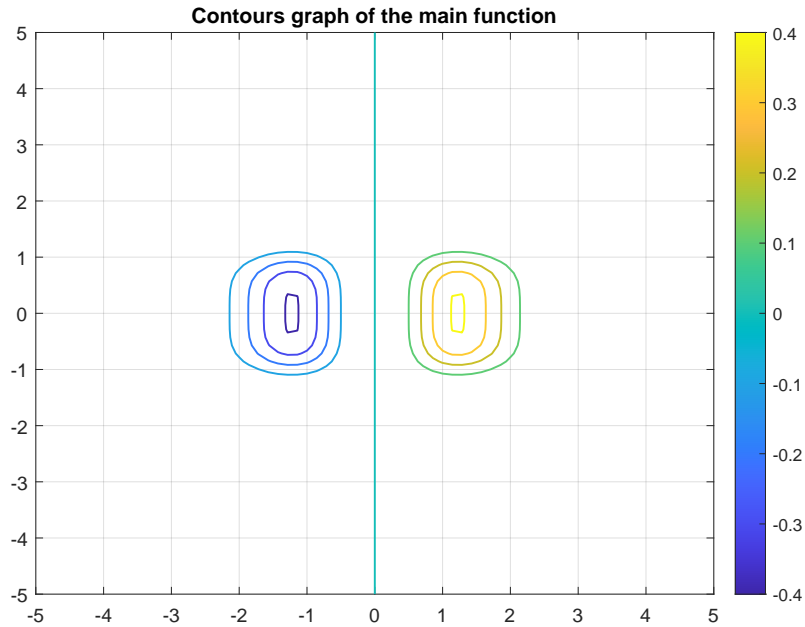
Η συνάρτηση προς ελαχιστοποίηση είναι η παρακάτω:

$$f(x, y) = x^3 \cdot e^{-x^2 - y^4} \quad (1.1)$$

Παρακάτω μπορούμε να παρατηρήσουμε τη μορφή της συνάρτησης τόσο σε 3D μορφή, 1.1, όσο και μέσω κάποιων ισοβαρών καμπυλών της (2D μορφή), 1.2.



Σχήμα 1.1: 3D Visualisation of the function



Σχήμα 1.2: Contours graph of the function

Η συνάρτηση παρουσιάζει ένα ελάχιστο και ένα μέγιστο, όπως μπορούμε να παρατηρήσουμε στα παραπάνω διαγράμματα. Συγκεκριμένα :

$$f_{\min} \simeq -0.409 \quad f_{\max} \simeq 0.409$$

Στην παρούσα εργασία καλούμαστε να αναζητήσουμε το ελάχιστό της παραπάνω συνάρτησης, χρησιμοποιώντας της παραπάνω μεθόδους. Η συγκεκριμένη συνάρτηση είναι **τουλάχιστον 2 φορές παραγωγίσιμη**, οπότε μπορούμε να εφαρμόσουμε και τις 3 μεθόδους.

Έπειτα από την αναλυτική μελέτη που θα πραγματοποιηθεί στα επόμενα κεφάλαια για κάθε μέθοδο με τη χρήση διαγραμμάτων και αποτελεσμάτων, η αναφορά θα ολοκληρωθεί με έναν συγκριτικό σχολιασμό πάνω στην αποδοτικότητα των υπό μελέτη μεθόδων για τη δοσμένη συνάρτηση.

Κεφάλαιο 2

Μελέτη Μεθόδων

Οι μέθοδοι που μελετούμε στην παρούσα εργασία βασίζονται στην ιδέα της επαναληπτικής καθόδου. Ξεκινάμε από κάποιο σημείο $x_0 \in \mathbb{R}^n$ και παράγουμε διαδοχικά τα διανύσματα x_1, x_2, x_3, \dots έτσι ώστε:

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

Ο αλγόριθμος μας οδηγεί σε βελτιωμένες τιμές της f προς ελαχιστοποίησή της. Η διαδικασία επιλογής των $x \in \mathbb{R}^n$ συνοψίζεται από την παρακάτω αναδρομική σχέση:

$$x_{k+1} = x_k + \gamma_k d_k, k = 0, 1, 2, \dots$$

όπου k ο αριθμός της επανάληψης, $\gamma_k > 0$ το βήμα και d_k το διάνυσμα κατεύθυνσης, ο υπολογισμός του οποίου μεταβάλλεται μεταξύ των μεθόδων, ικανοποιώντας πάντα τη συνθήκη:

$$\nabla f^T(x_k) d_k < 0, k = 0, 1, 2, \dots$$

Οι περισσότεροι αλγόριθμοι κλίσης είναι της μορφής (μεσα σε αυτούς και οι υπό μελέτη μέθοδοι της συγκεκριμένης εργασίας):

$$x_{k+1} = x_k - \gamma_k \Delta_k \nabla f(x_k) \quad (2.1)$$

όπου ο πίνακας Δ_k είναι συμμετρικός και **θετικά ορισμένος**.

Αξίζει να σημειωθεί εδώ ότι κατά τη μελέτη των μεθόδων μεταβάλουμε τη επιλογή του βήματος, με τις παρακάτω μεθόδους:

- $\gamma_k = \text{const}$
- $\gamma_k = \min(f(x_k + \gamma_k d_k))$, ως προς $\gamma_k > 0$
- Κανόνας Armijo

Ο κανόνας Armijo αποτελεί μια μέθοδο διαδοχικής μείωσης του γ_k , όπου το βήμα επιλέγεται ως:

$$\gamma_k = s\beta^{m_k}$$

όπου m_k είναι ο μικρότερος μη-αρνητικός ακέραιος που ικανοποιεί την παρακάτω σχέση:

$$f(x_k) - f(x_{k+1}) \geq -a\beta^{m_k} s d_k^T \nabla f(x_k)$$

Ο συντελεστής s δηλώνει το αρχικό βήμα, ενώ τα a, β τα προκαθορίζουμε. Για την παρούσα εργασία χρησιμοποιήθηκαν:

$$s = \gamma_0 = \text{const} \quad a = 10^{-3} \quad \beta = 0.1$$

Να σημειώσουμε εδώ ότι στη δεύτερη μέθοδο υπολογισμού του γ αξιοποιείται η μέθοδος **Golden Section** που υλοποιήθηκε στην προηγούμενη εργασία.

Τέλος, σε όλους τους αλγορίθμους χρησιμοποιήθηκε **μέγιστο όριο 1000 βημάτων** για τον υπολογισμό του ελαχίστου.

2.1 Μέθοδος Μέγιστης Καθόδου (Steepest Descent)

2.1.1 Σύντομη περιγραφή της μεθόδου

Η μέθοδος της μέγιστης καθόδου (steepest descent) προκύπτει επιλέγοντας:

$$\Delta_k = I, k = 1, 2, 3, \dots$$

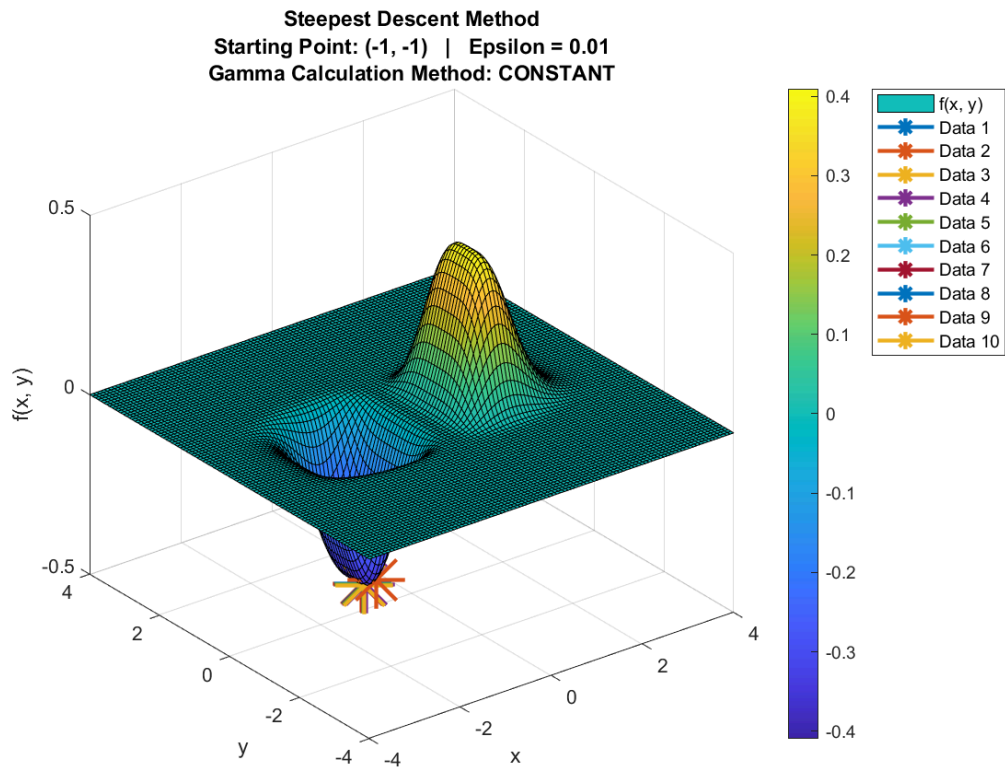
όπου I μοναδιαίος $n \times n$ πίνακας, οπότε από την εξίσωση 2.1 προκύπτει ότι:

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

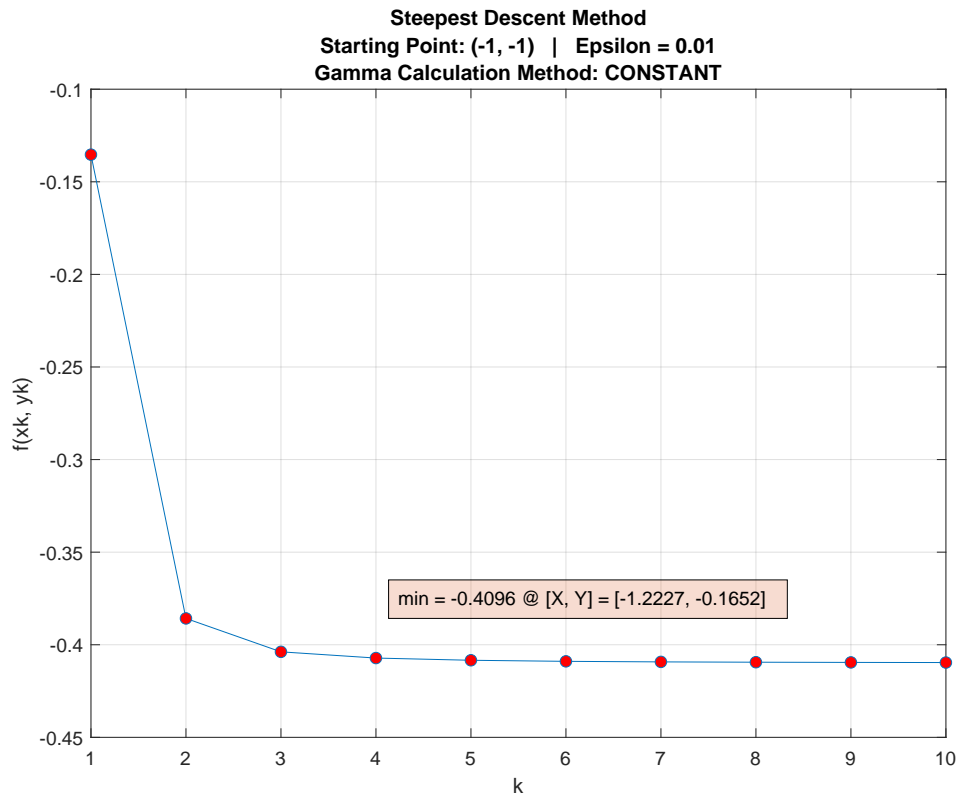
Η γενική ιδέα της μεθόδου είναι ότι η μεγαλύτερη αύξηση, $\|dx\|$, της μεταβλητής x θα υπάρξει όταν το διάνυσμα dx είναι συγγραμικό με το διάνυσμα κλίσης, ∇f , της f .

2.1.2 Γραφήματα

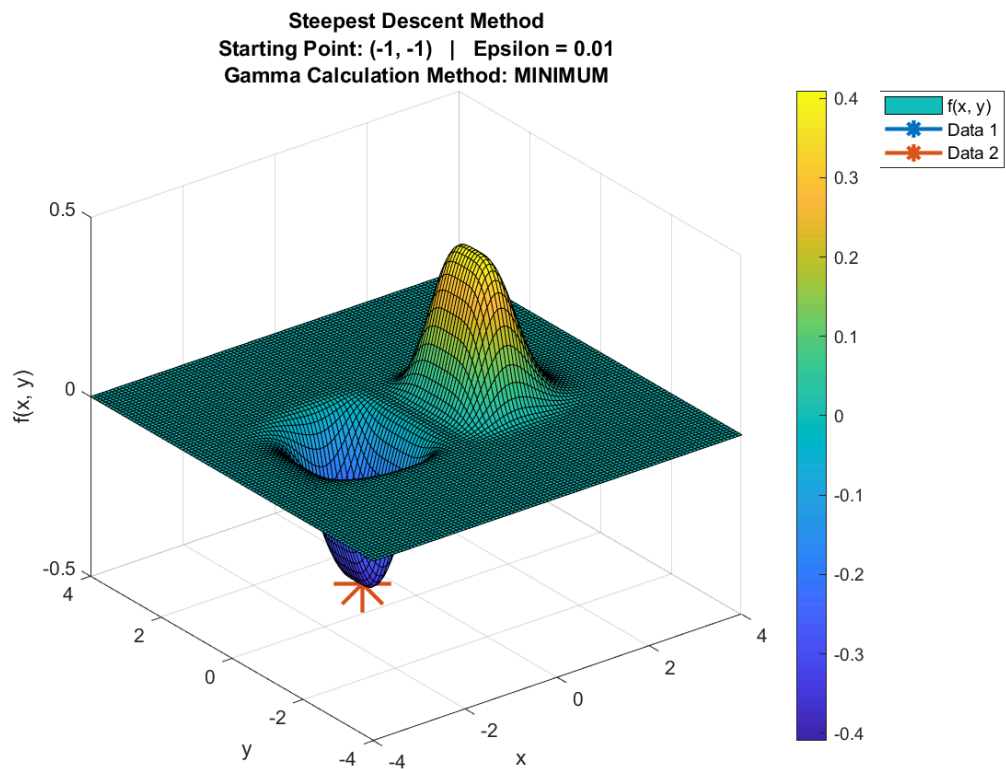
Σημείο εκκίνησης $[x_0, y_0] = [-1, -1]$



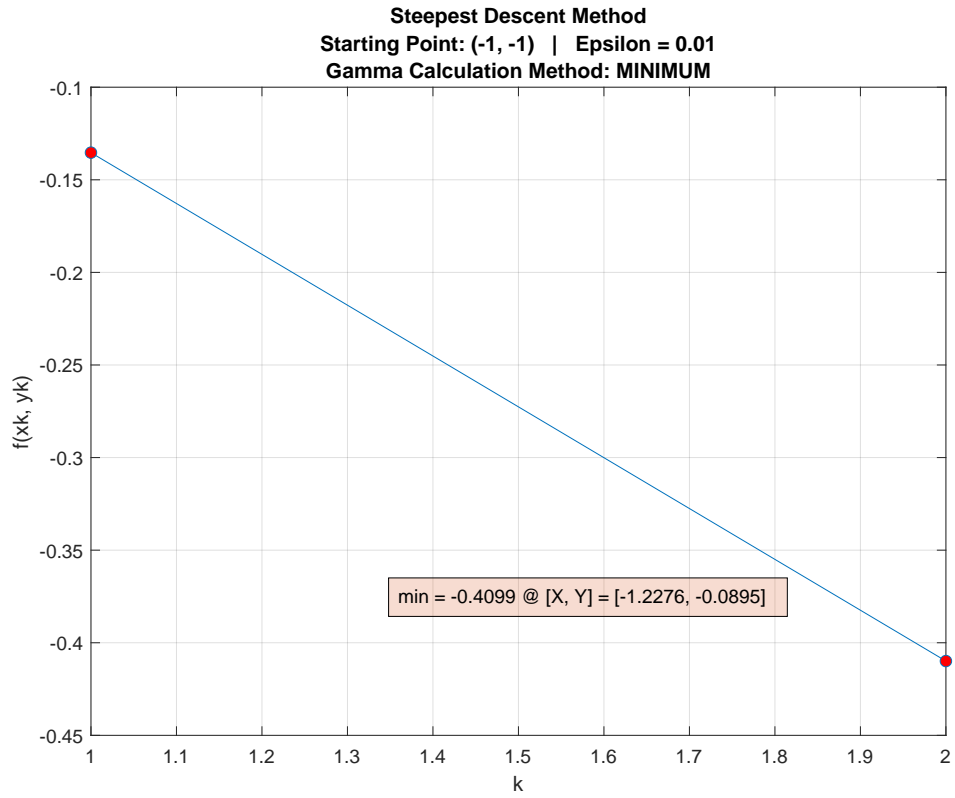
Σχήμα 2.1: 3D visualisation – γ_k method: CONSTANT – $[-1, -1]$



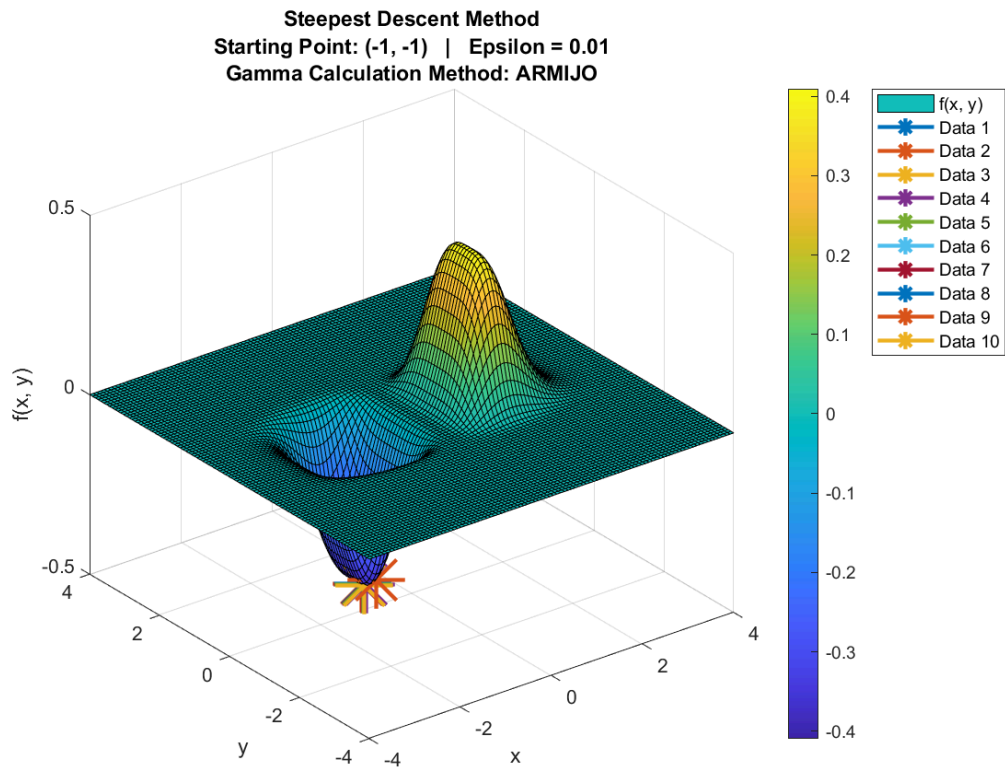
Σχήμα 2.2: 2D visualisation – γ_k method: CONSTANT – [-1, -1]



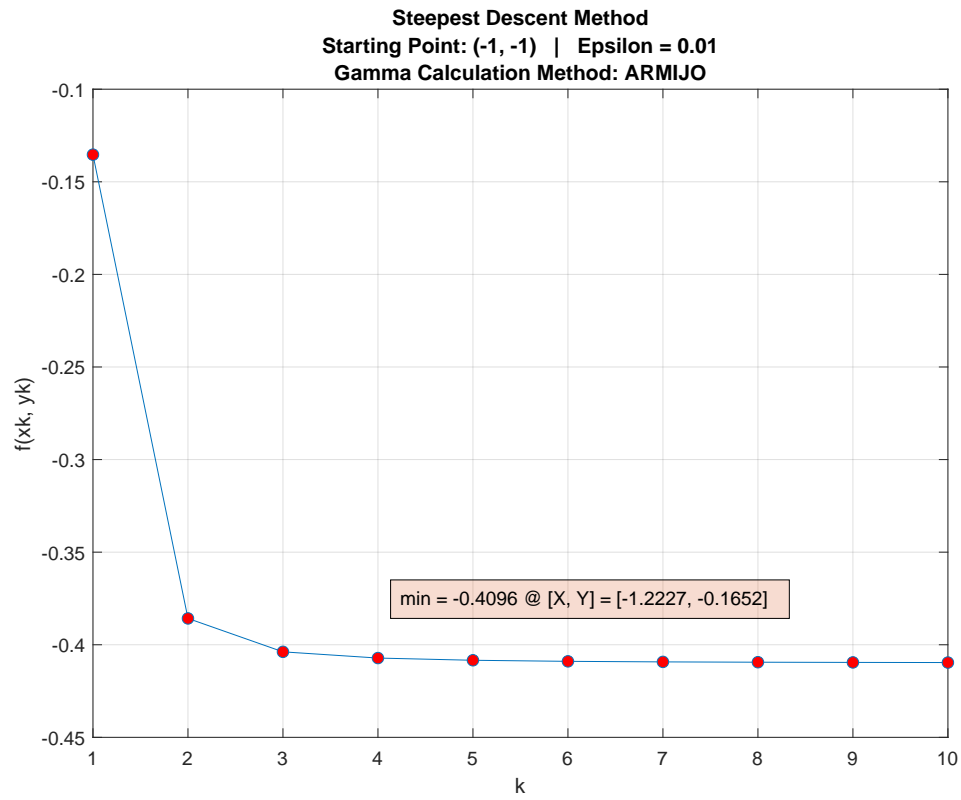
Σχήμα 2.3: 3D visualisation – γ_k method: MINIMUM – [-1, -1]



Σχήμα 2.4: 2D visualisation – γ_k method: MINIMUM – [-1, -1]

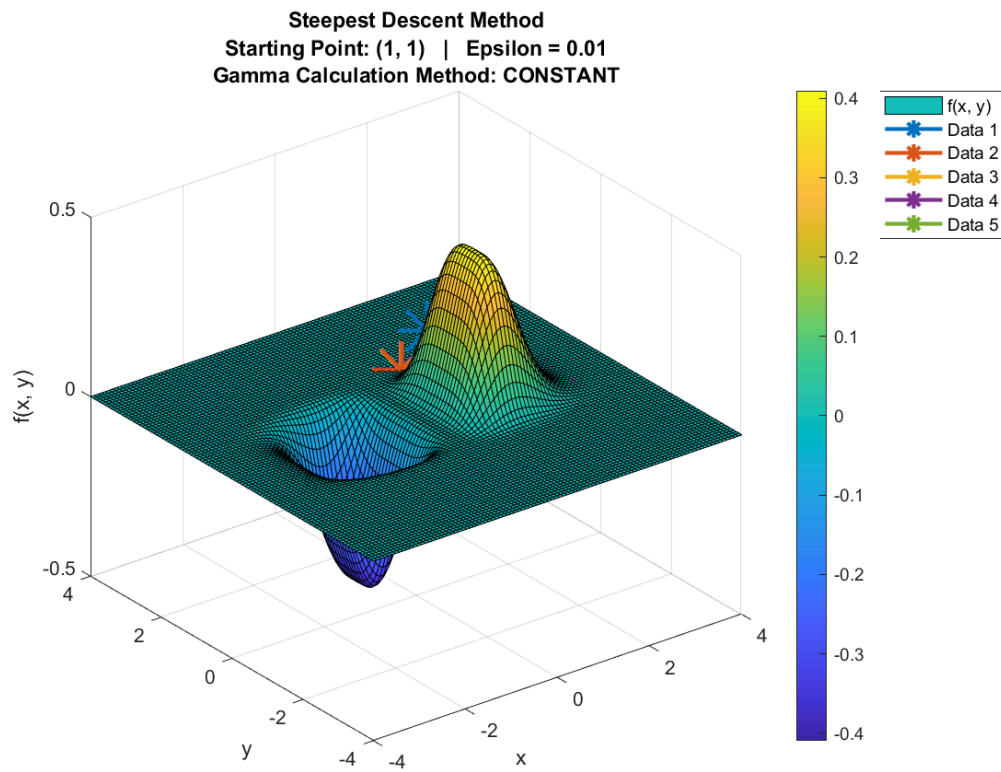


Σχήμα 2.5: 3D visualisation – γ_k method: ARMIJO – [-1, -1]

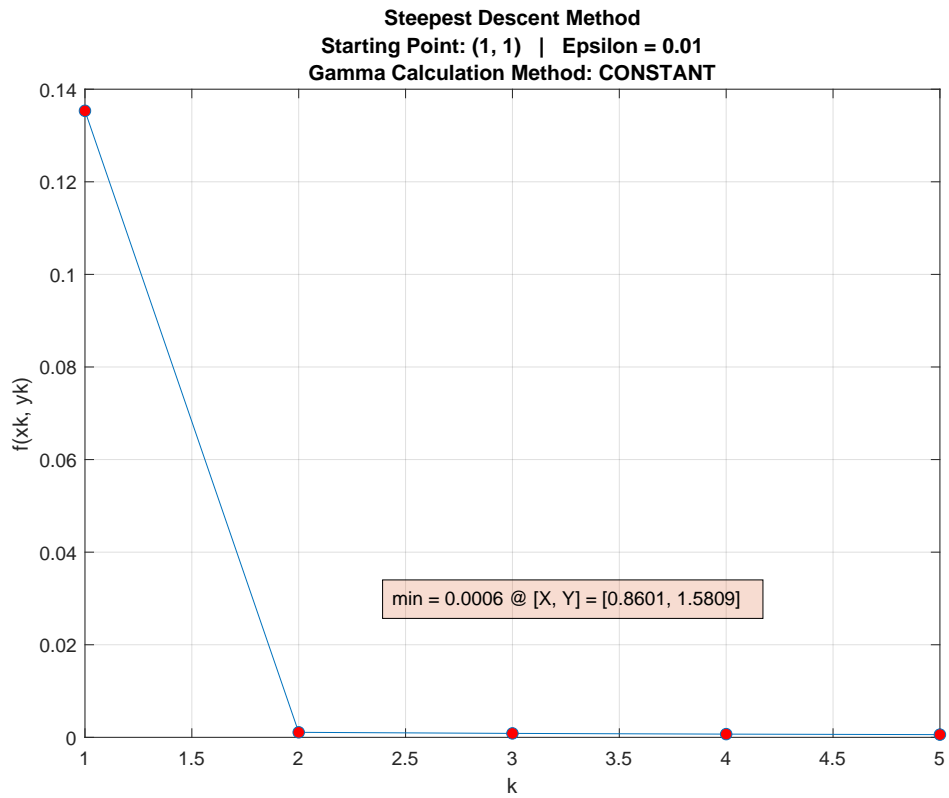


Σχήμα 2.6: 2D visualisation – γ_k method: ARMIJO – [-1, -1]

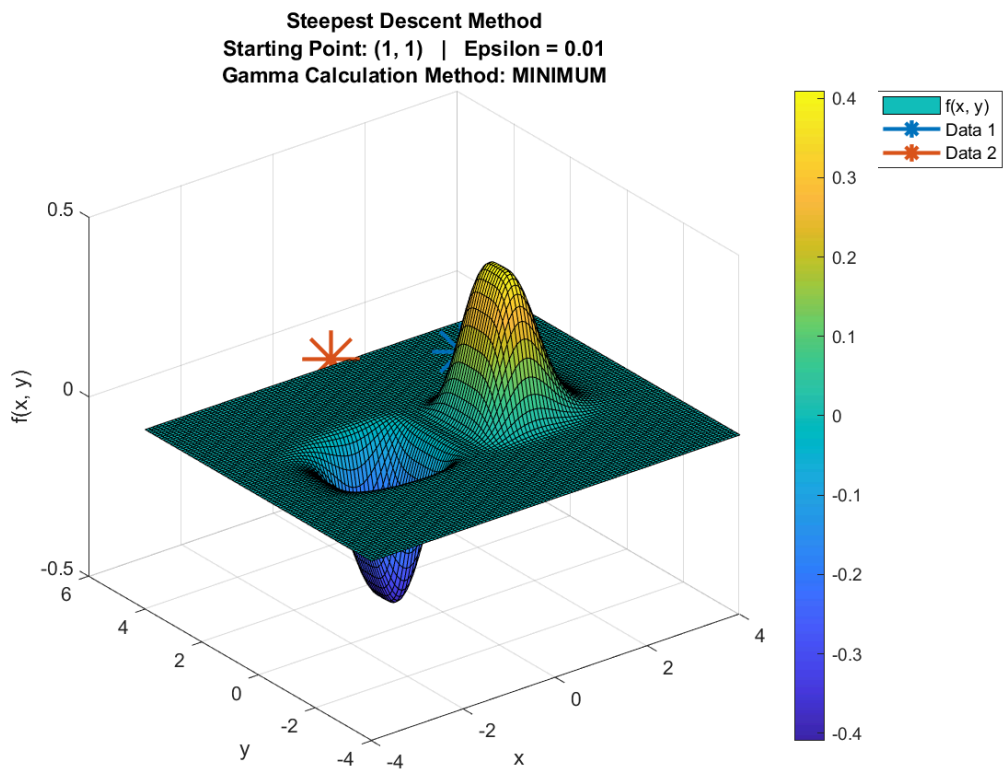
Σημείο εκκίνησης $[x_0, y_0] = [1, 1]$



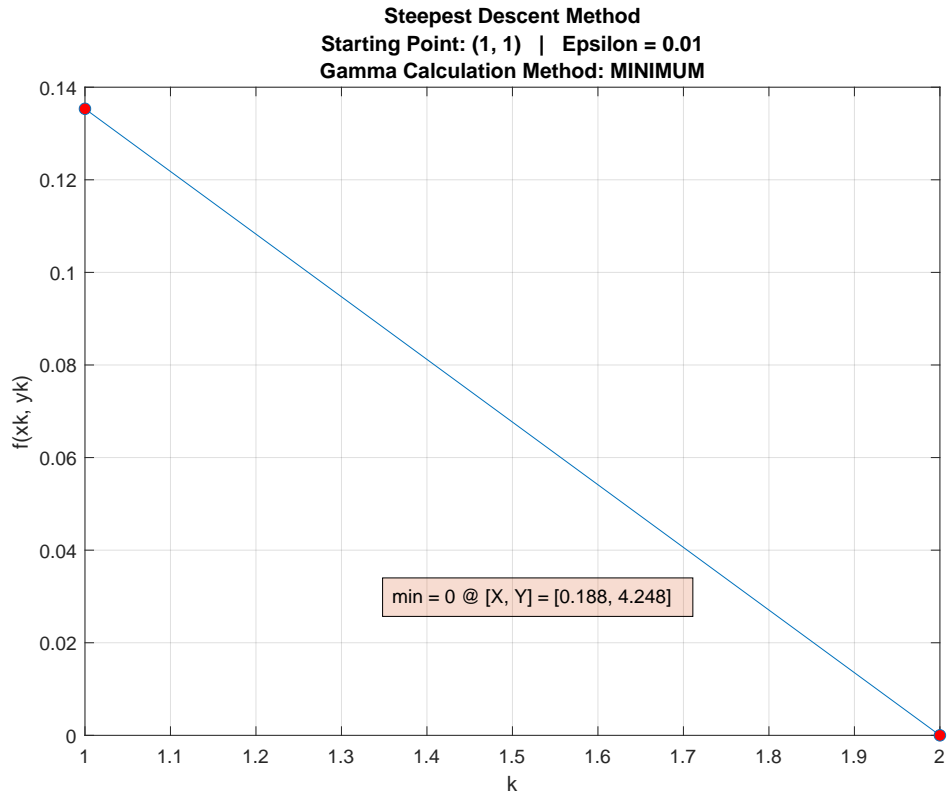
Σχήμα 2.7: 3D visualisation – γ_k method: CONSTANT – [1, 1]



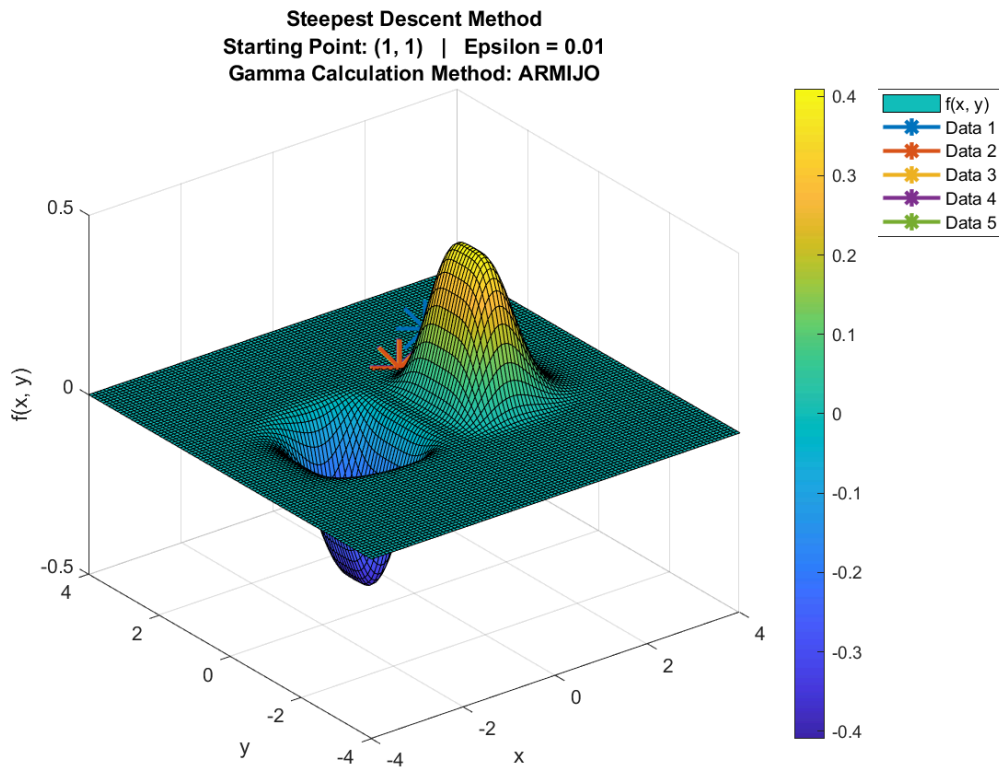
Σχήμα 2.8: 2D visualisation – γ_k method: CONSTANT – [1, 1]



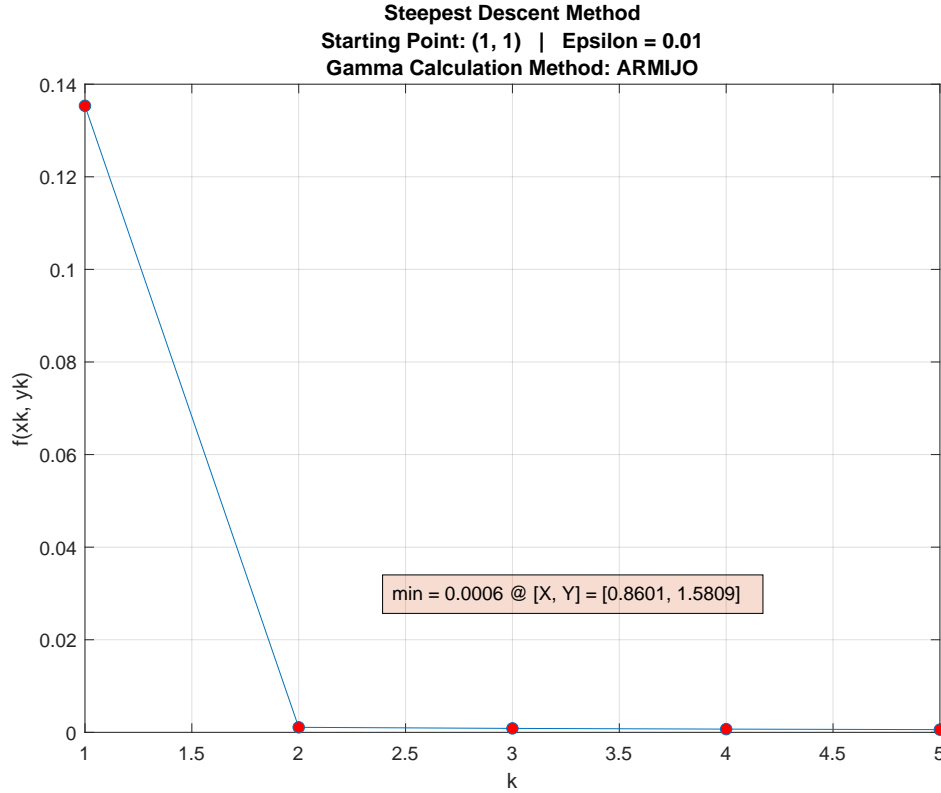
Σχήμα 2.9: 3D visualisation – γ_k method: MINIMUM – [1, 1]



Σχήμα 2.10: 2D visualisation – γ_k method: MINIMUM – [1, 1]



Σχήμα 2.11: 3D visualisation – γ_k method: ARMIJO – [1, 1]



Σχήμα 2.12: 2D visualisation – γ_k method: ARMIJO – [1, 1]

2.2 Μέθοδος Newton

2.2.1 Σύντομη περιγραφή της μεθόδου

Η μέθοδος Newton προκύπτει επιλέγοντας:

$$\Delta_k = [\nabla^2 f(x_k)]^{-1}, k = 0, 1, 2, \dots$$

με την προϋπόθεση ότι ο εσσιανός πίνακας $\nabla^2 f(x_k)$ είναι θετικά ορισμένος και αντιστρέψιμος. Συνεπώς, από την εξίσωση 2.1 προκύπτει ότι:

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

Σ' αυτήν την μέθοδο πρακτικά προσεγγίζουμε την αντικειμενική συνάρτηση με την τετραγωνική της μορφή, με σκοπό την σύγκλιση στο σημείο ελαχίστου σε ένα και μόνο βήμα.

2.2.2 Γραφήματα

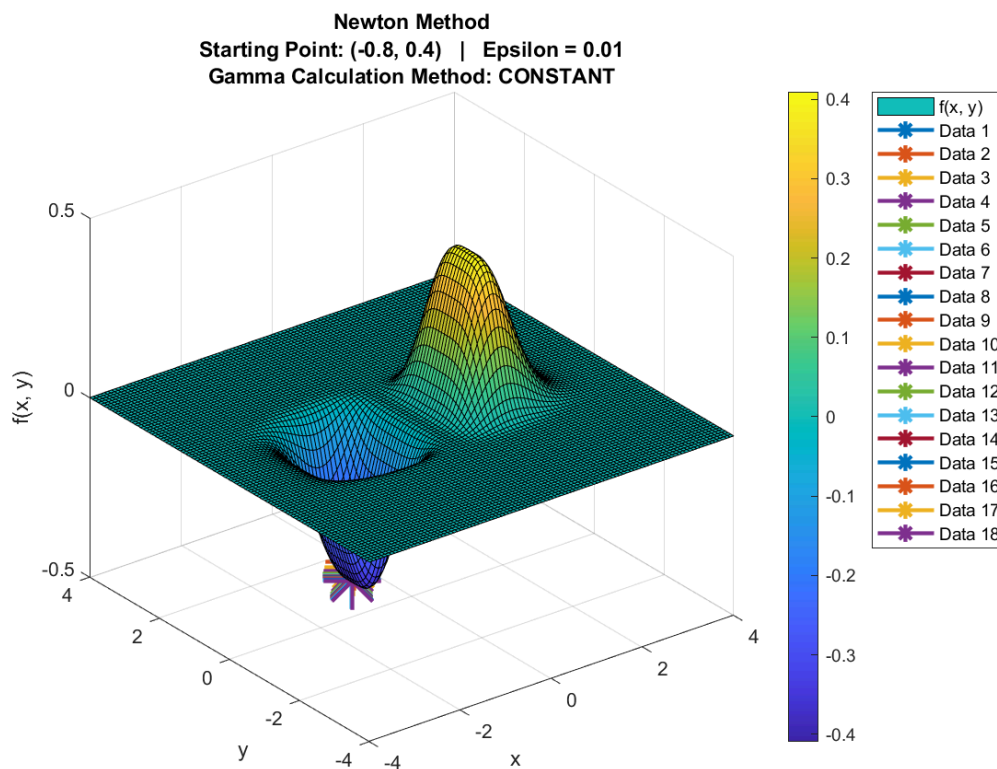
Όπως αναφέρθηκε παραπάνω, για να μπορεί να εφαρμοστεί η μέθοδος Newton χρειάζεται ο εσσιανός πίνακας στο σημείο x_k να είναι θετικά ορισμένος. Για όλες τις αρχικές συνθήκες που δόθηκαν, ο εσσιανός πίνακας στο σημείο x_0 είναι **μη θετικά ορισμένος**, καθώς:

$$\det(\nabla^2 f(x_0)) < 0 \quad \text{and} \quad \nabla^2 f(x_0)(1, 1) < 0$$

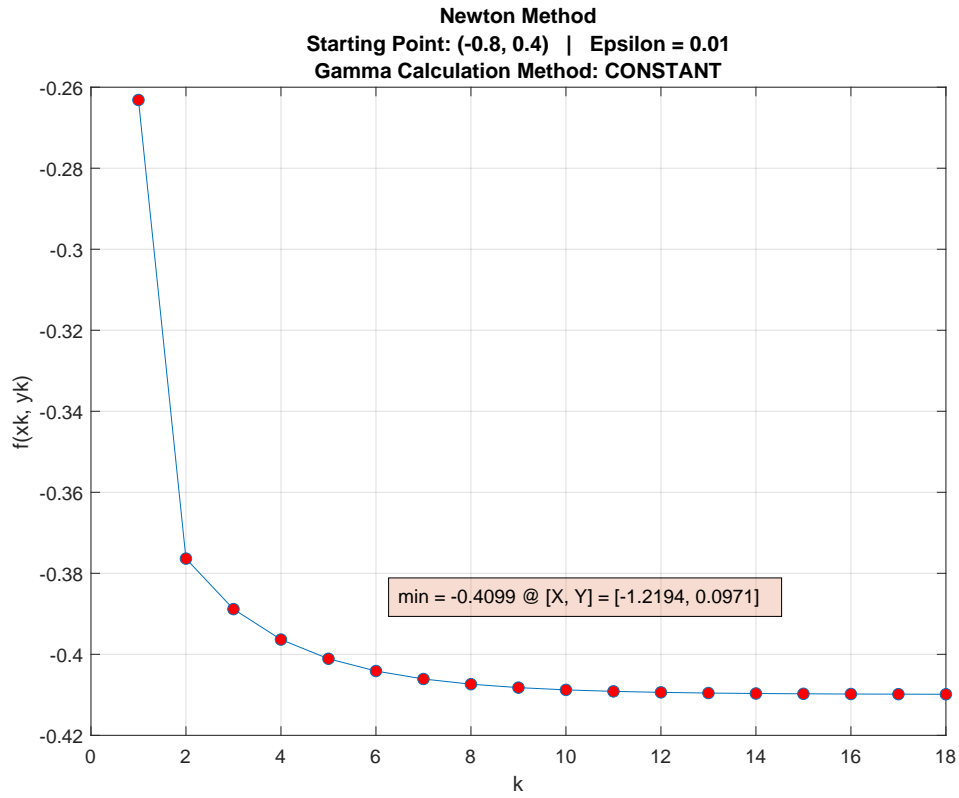
Συνεπώς δεν μπορεί να εφαρμοστεί η μέθοδος για τα συγκεκριμένα αρχικά σημεία.

Για να μπορέσουμε, ωστόσο, να δούμε σε εφαρμογή τη μέθοδο, έπειτα από κατάλληλη αναζήτηση, εντοπίστηκαν κάποιες αρχικές συνθήκες για τις οποίες ο αρχικός εσσιανός πίνακας και οι υπόλοιποι που ακολουθούν είναι **θετικά ορισμένοι**, άρα η μέθοδος Newton εφαρμόζεται και μάλιστα συγκλίνει στο ελάχιστο. Παρακάτω βλέπουμε τα αντίστοιχα γραφήματα:

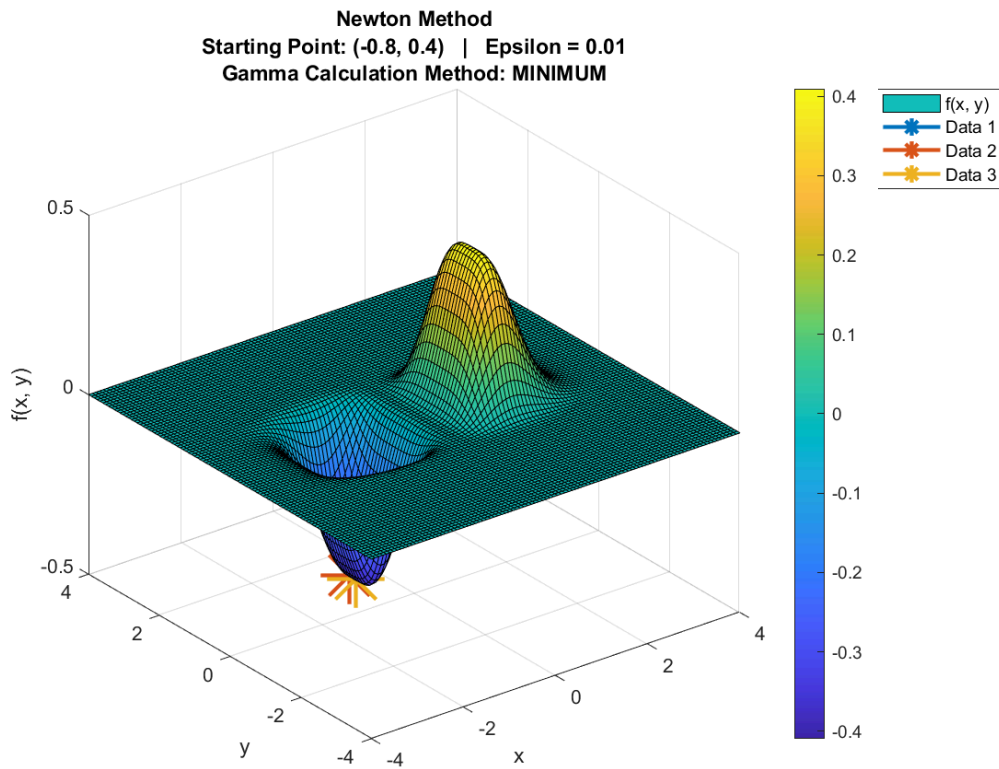
Σημείο εκκίνησης $[x_0, y_0] = [-0.8, 0.4]$



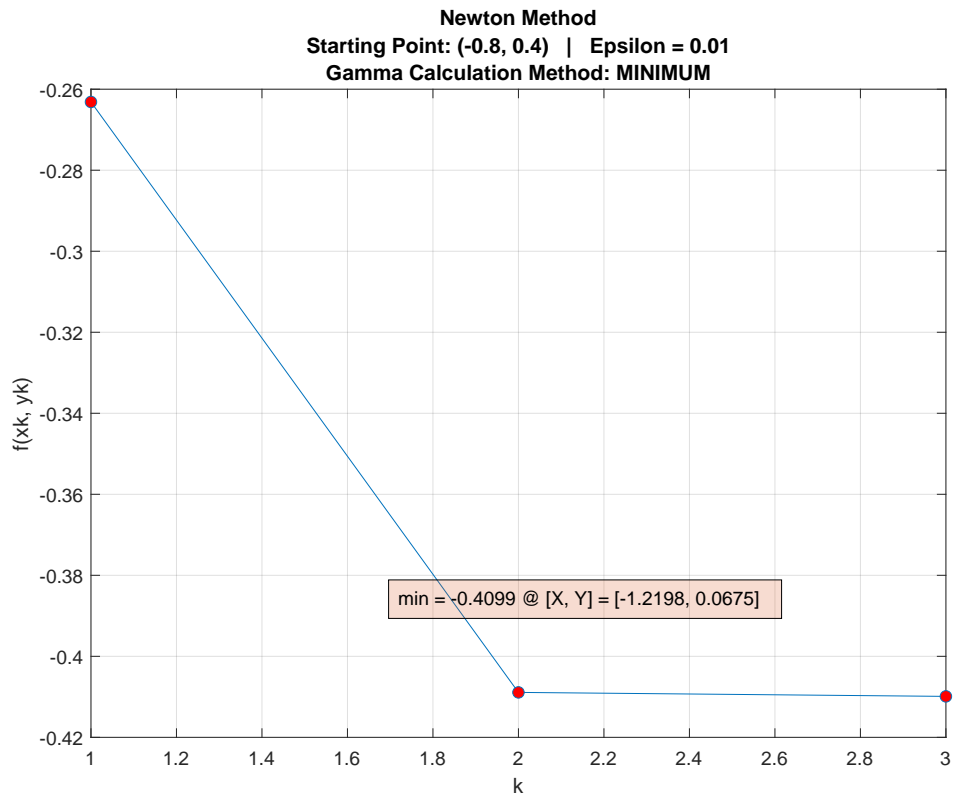
Σχήμα 2.13: 3D visualisation – γ_k method: CONSTANT – $[-0.8, 0.4]$



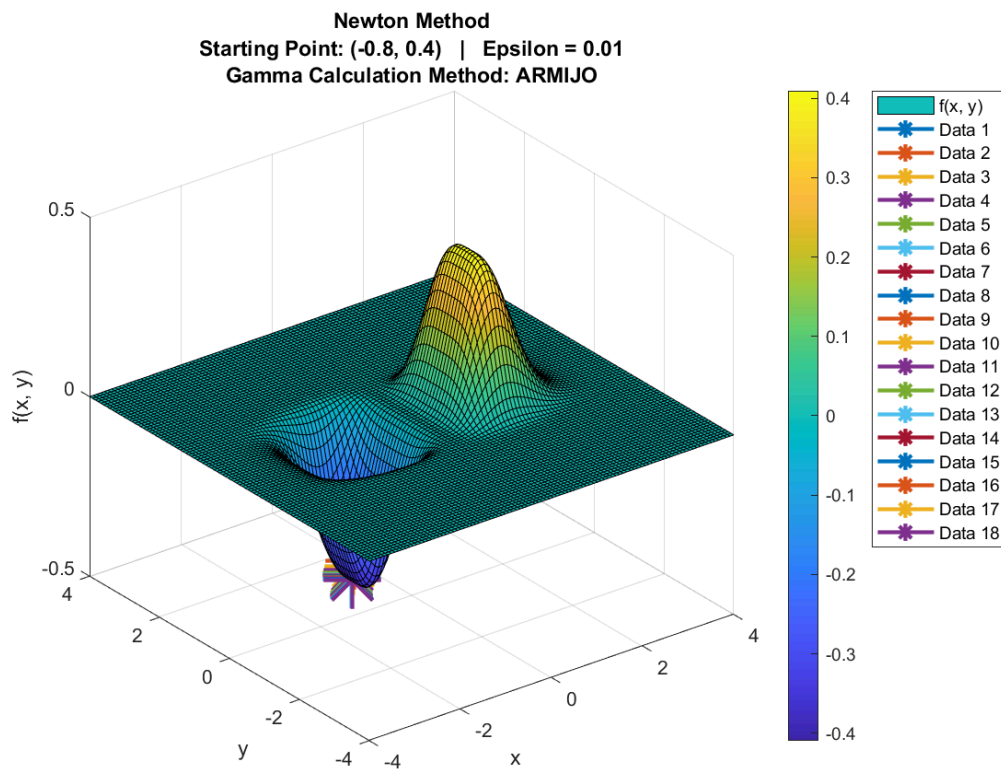
Σχήμα 2.14: 2D visualisation – γ_k method: CONSTANT – [-0.8, 0.4]



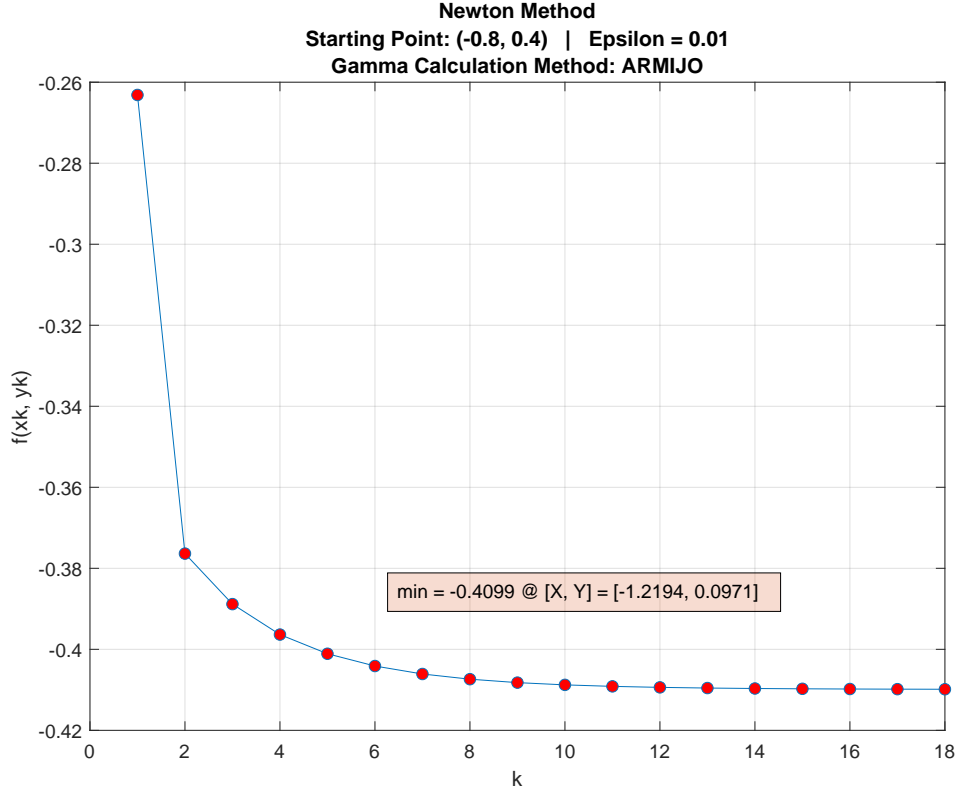
Σχήμα 2.15: 3D visualisation – γ_k method: MINIMUM – [-0.8, 0.4]



Σχήμα 2.16: 2D visualisation – γ_k method: MINIMUM – [-0.8, 0.4]



Σχήμα 2.17: 3D visualisation – γ_k method: ARMIJO – [-0.8, 0.4]



Σχήμα 2.18: 2D visualisation – γ_k method: ARMIJO – [-0.8, 0.4]

2.3 Μέθοδος Levenberg-Marquardt

2.3.1 Σύντομη περιγραφή της μεθόδου

Η μέθοδος Levenberg-Marquardt τροποποιεί τη μέθοδο Newton, ώστε να μπορεί να είναι λειτουργική ακόμη και αν ο $\nabla^2 f(x_k)$ δεν είναι θετικά ορισμένος.

Πρακτικά επιλέγουμε:

$$\Delta_k = [\nabla^2 f(x_k) + \mu_k I]^{-1}, k = 0, 1, 2, \dots$$

όπου I ο μοναδιαίος πίνακας και $\mu_k > 0$. Για κατάλληλη επιλογή του μ_k τέτοια ώστε:

$$\mu_k > |\hat{\lambda}_{\max}\{\nabla^2 f(x_k)\}|$$

όπου $\hat{\lambda}_{\max}$ η μεγαλύτερη ιδιοτιμή του εσσανού πίνακα $\nabla^2 f(x_k)$.

Για την εφαρμογή της μεθόδου, φροντίζουμε να ισχύουν τα παρακάτω κριτήρια:

$$d_k^T \nabla f(x_{k+1}) > \beta d_k^T \nabla f(x_k) \quad (2.2)$$

για κάποιο $\beta \in (0, 1)$ και:

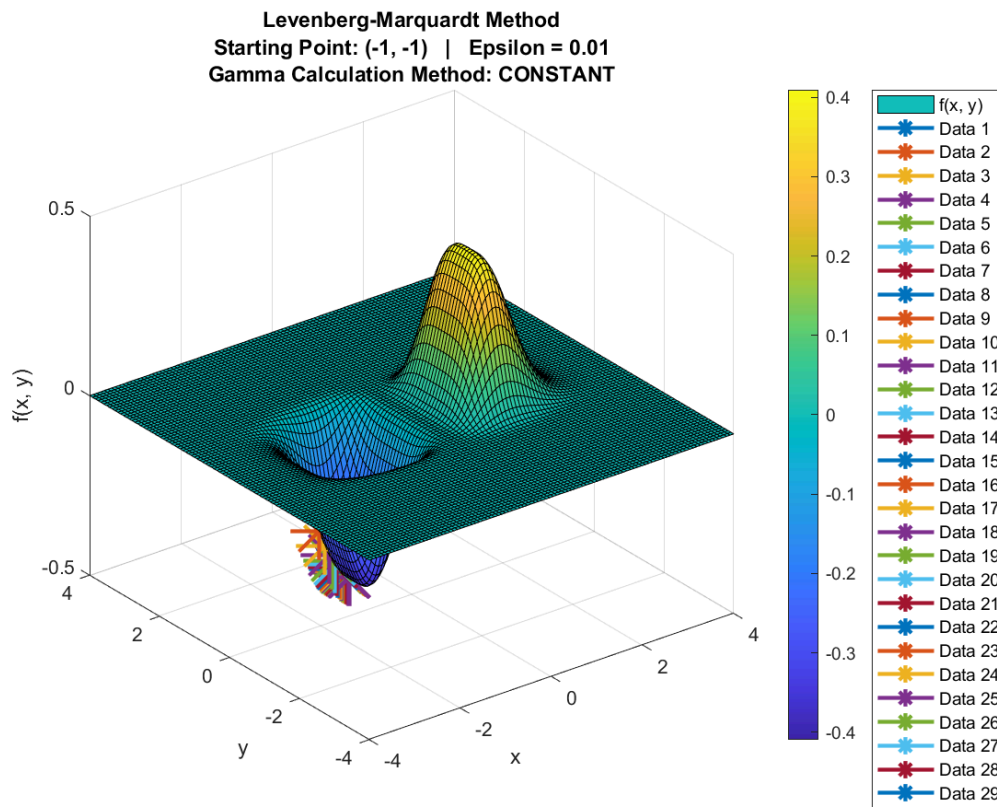
$$f(x_{k+1}) \leq f(x_k) + a \gamma_k d_k^T \nabla f(x_k) \quad (2.3)$$

για κάποιο $a \in (0, \beta)$

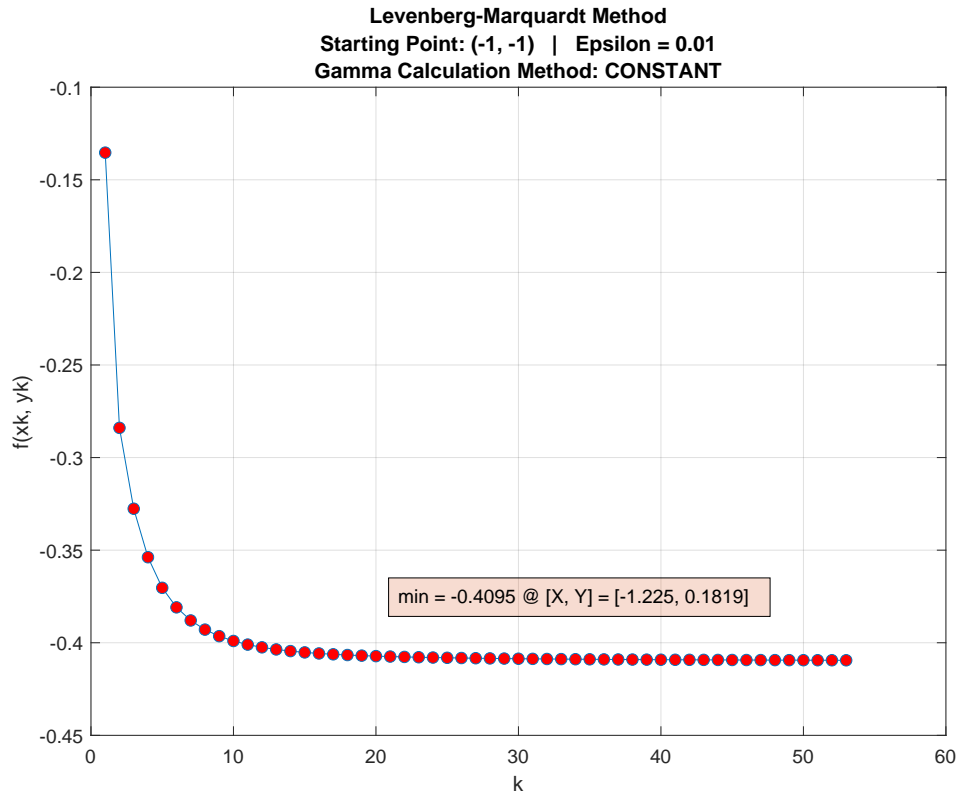
Αυτά τα κριτήρια διασφαλίζουν το γ_k να μην επιλέγεται οσοδήποτε μικρό, αλλά και ότι δεν θα πάρει πολύ μεγάλες τιμές, αντίστοιχα. Δηλαδή, **αποκλείονται περιπτώσεις ταλάντωσης και απόκλισης**.

2.3.2 Γραφήματα

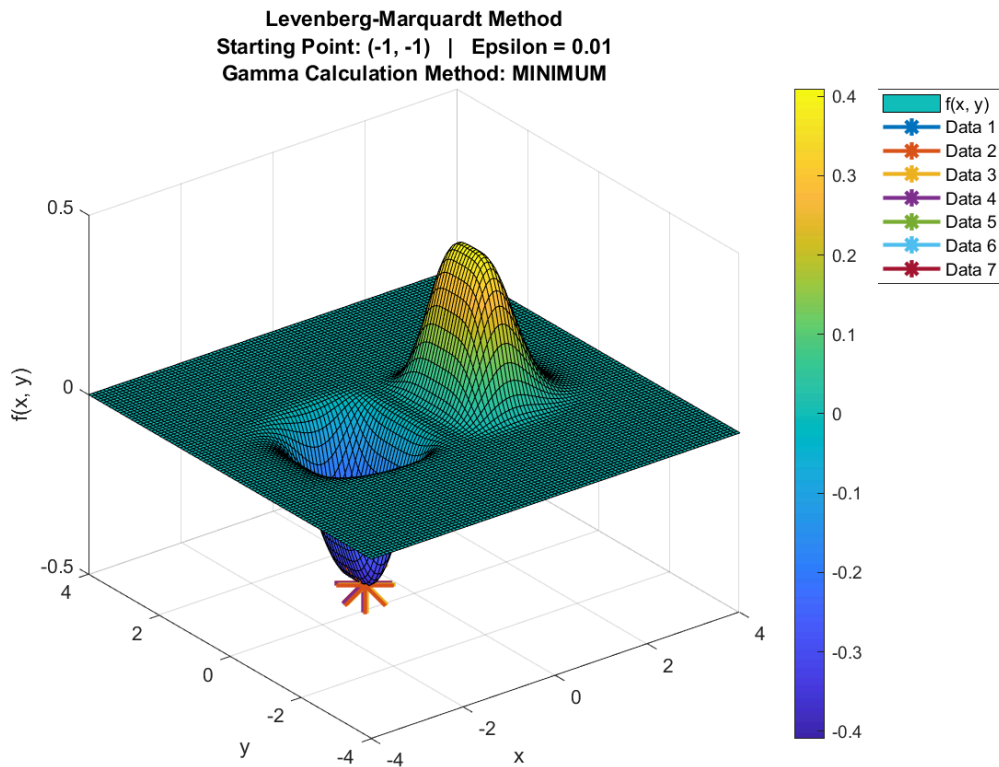
Σημείο εκκίνησης $[x_0, y_0] = [-1, -1]$



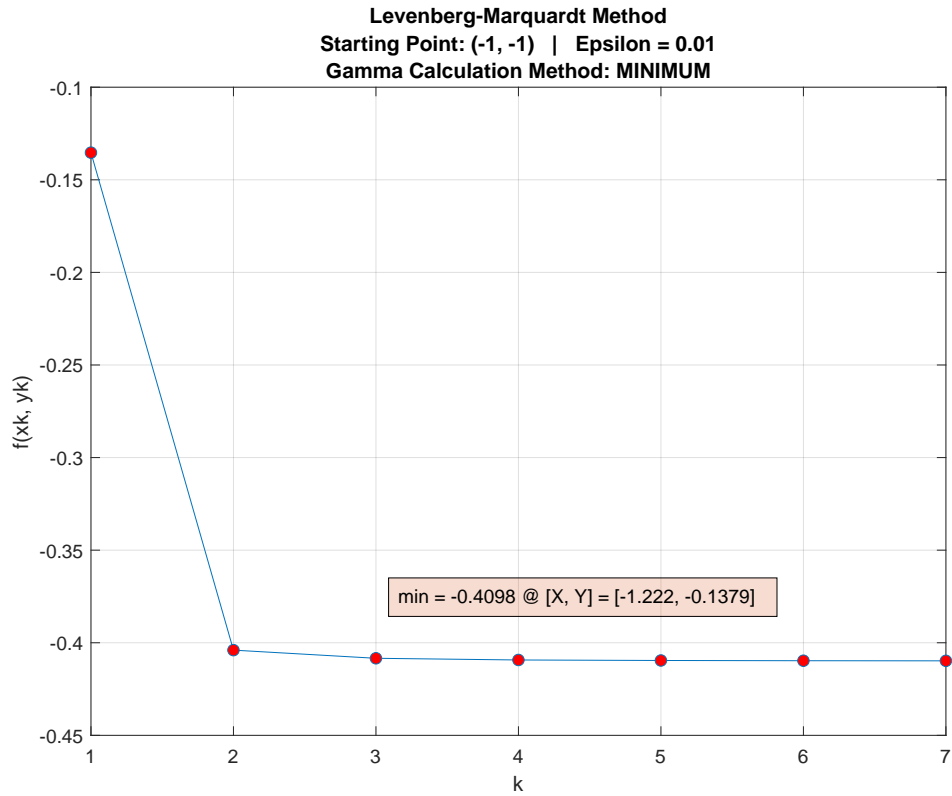
Σχήμα 2.19: 3D visualisation – γ_k method: CONSTANT – $[-1, -1]$



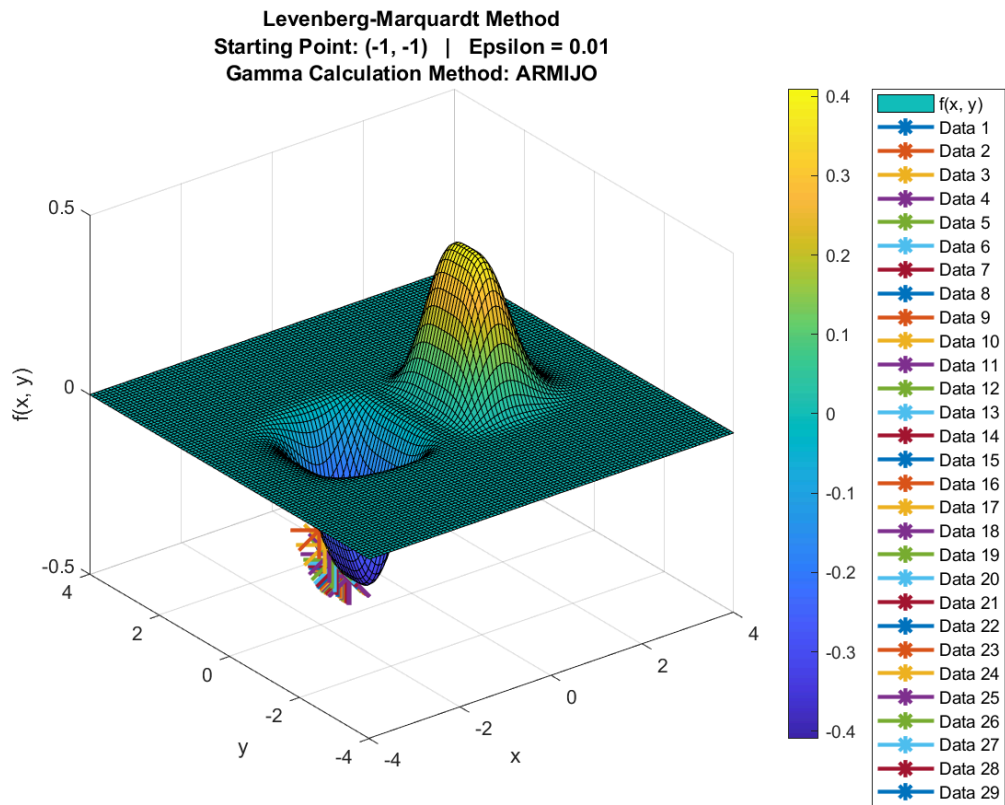
Σχήμα 2.20: 2D visualisation – γ_k method: CONSTANT – [-1, -1]



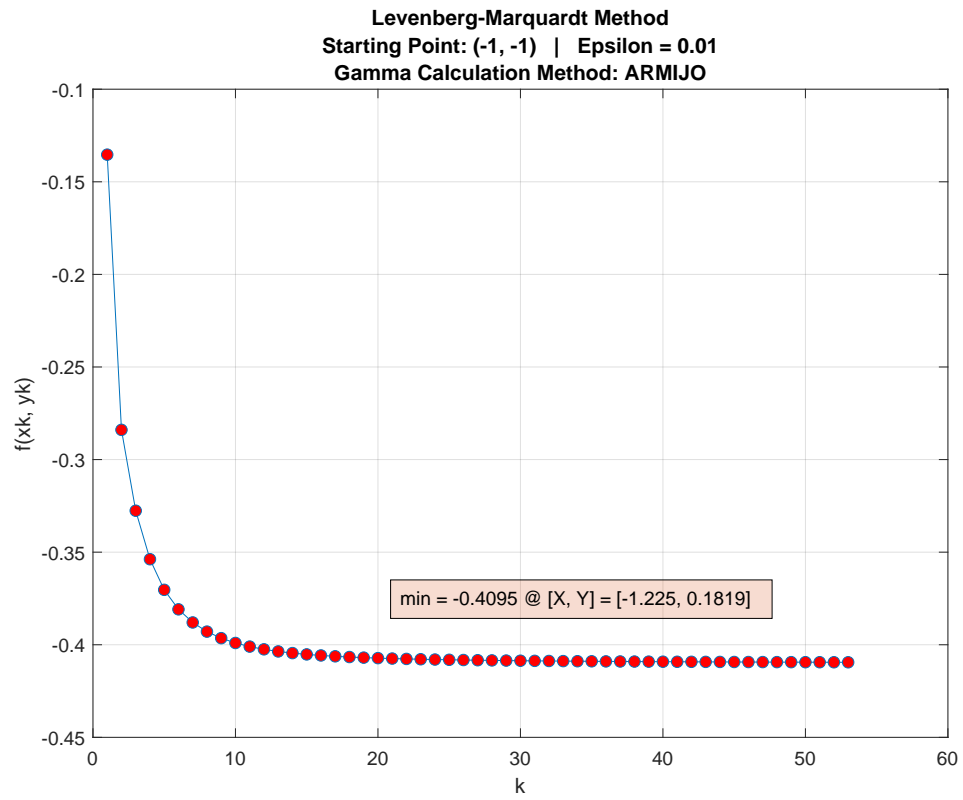
Σχήμα 2.21: 3D visualisation – γ_k method: MINIMUM – [-1, -1]



Σχήμα 2.22: 2D visualisation – γ_k method: MINIMUM – [-1, -1]

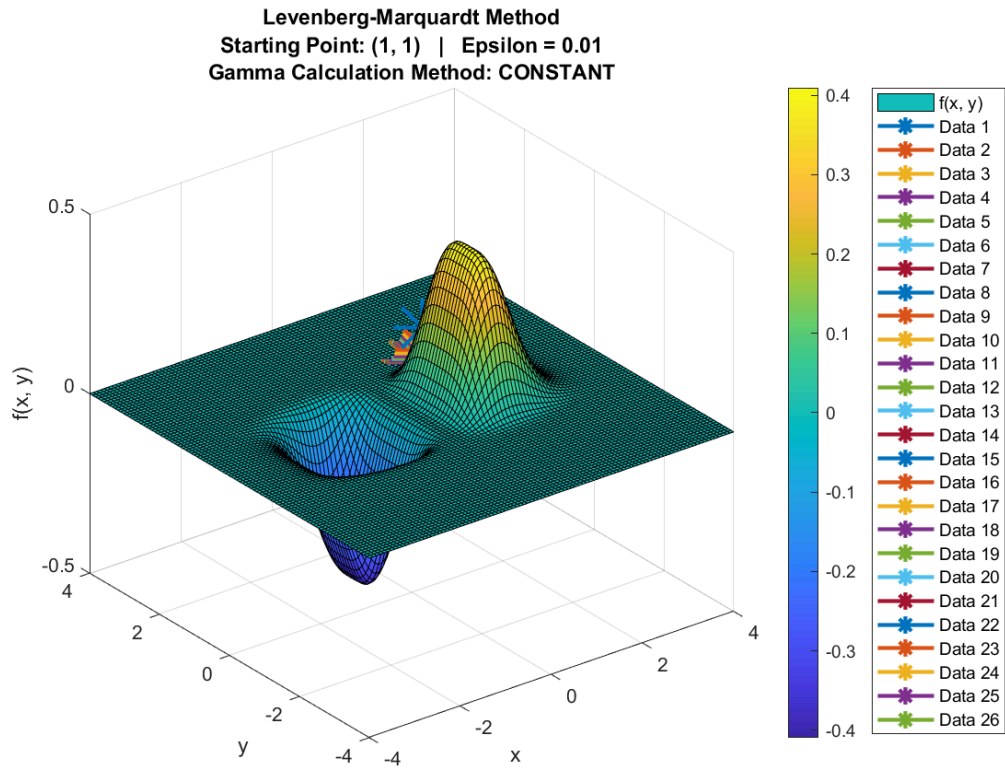


Σχήμα 2.23: 3D visualisation – γ_k method: ARMIJO – [-1, -1]

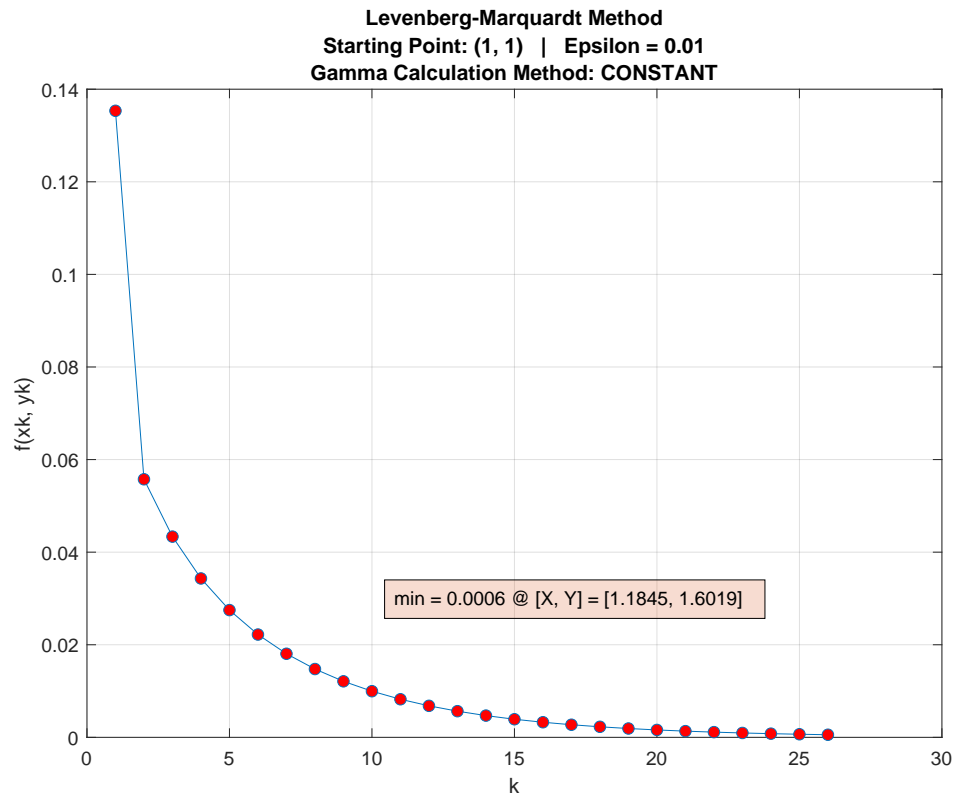


Σχήμα 2.24: 2D visualisation – γ_k method: ARMIJO – [-1, -1]

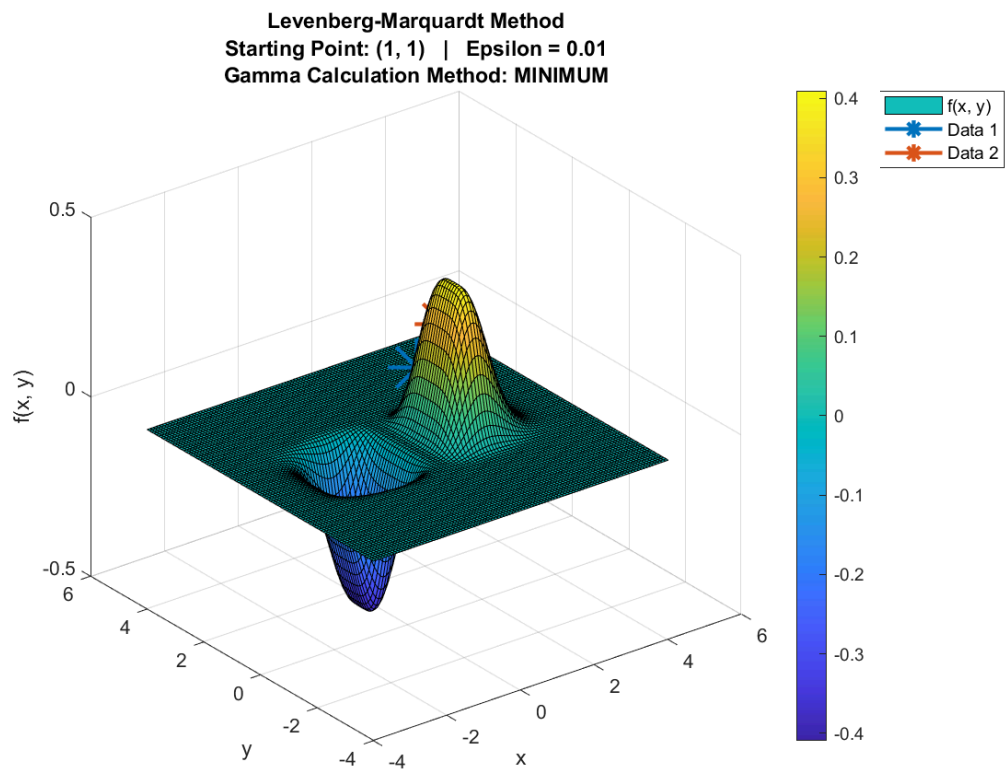
Σημείο εκκίνησης $[x_0, y_0] = [1, 1]$



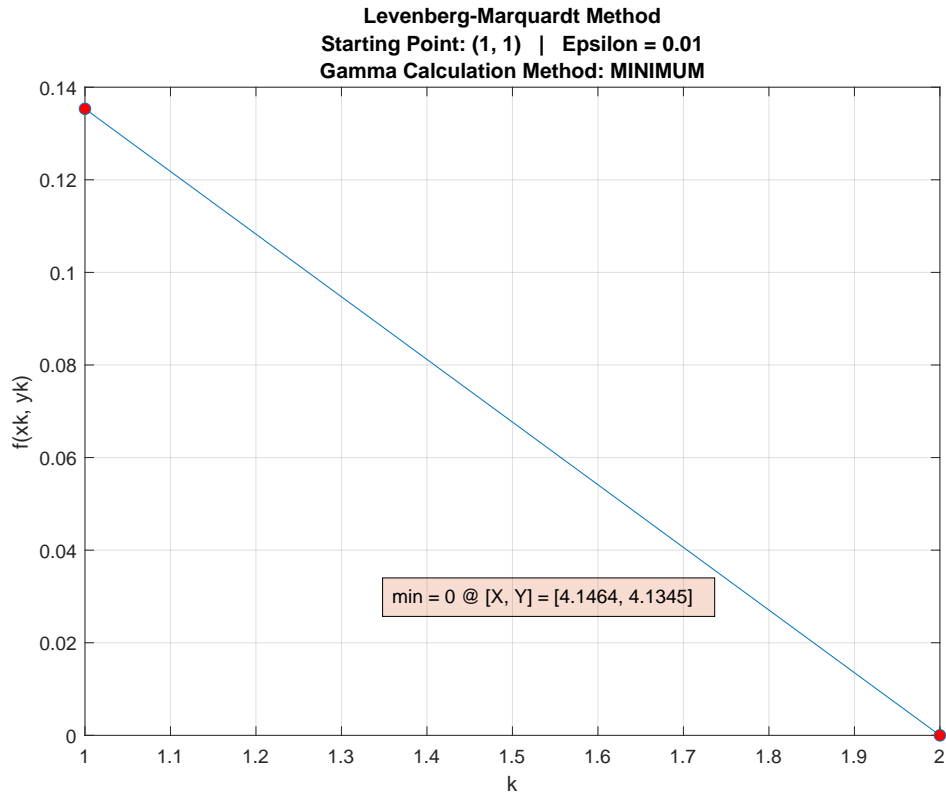
Σχήμα 2.25: 3D visualisation – γ_k method: CONSTANT – [1, 1]



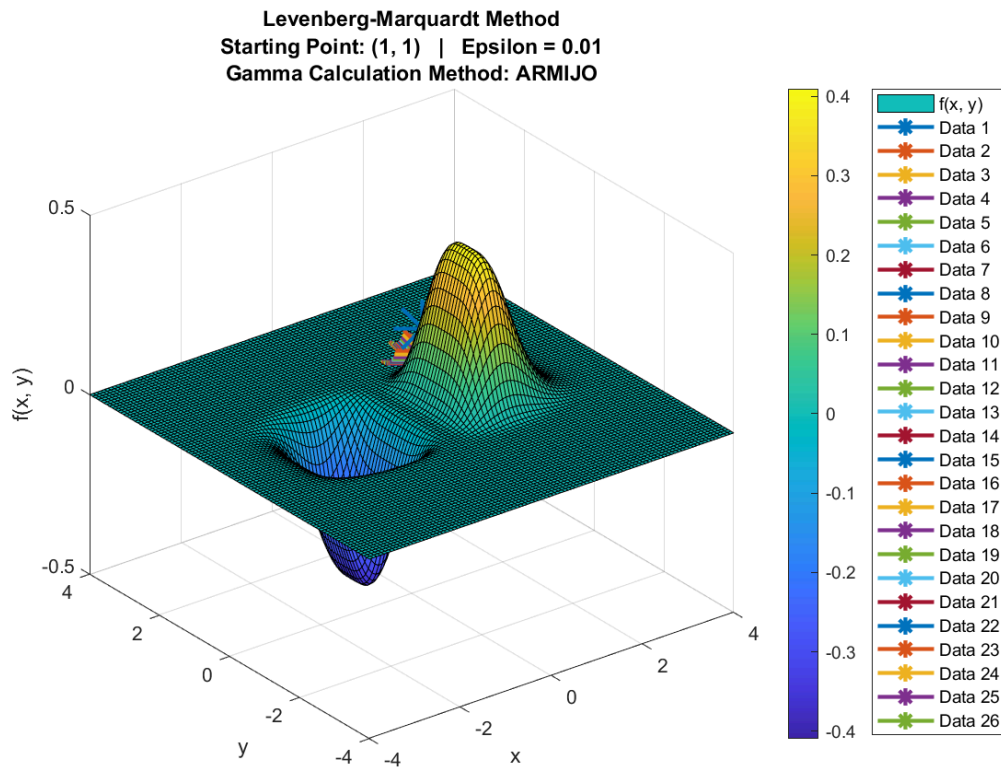
Σχήμα 2.26: 2D visualisation – γ_k method: CONSTANT – [1, 1]



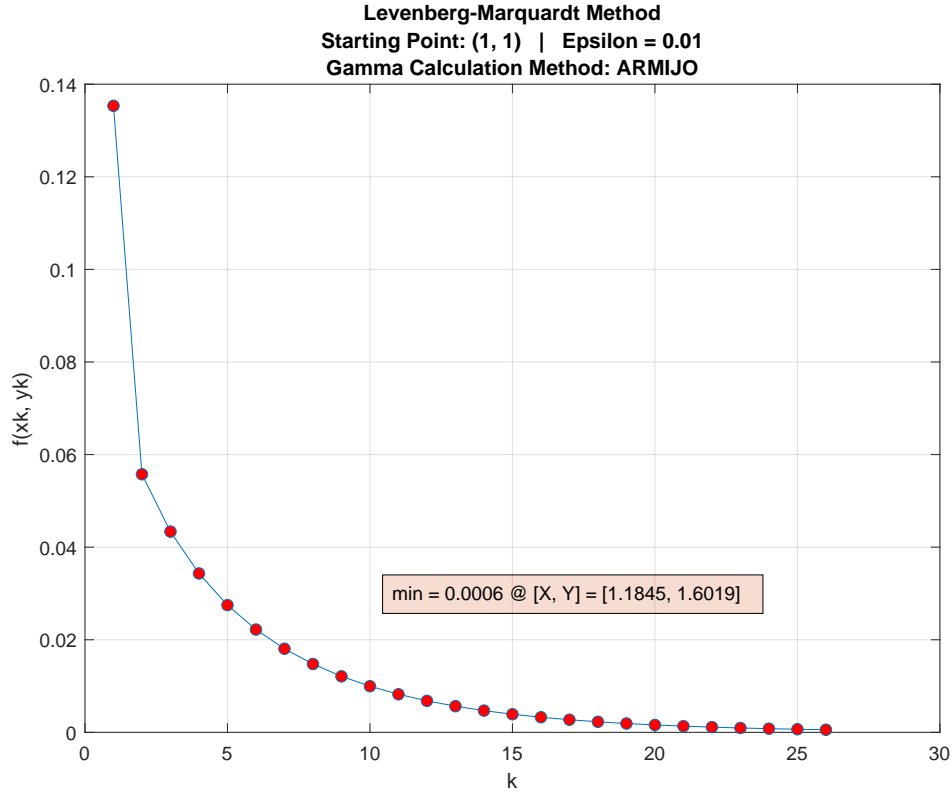
Σχήμα 2.27: 3D visualisation – γ_k method: MINIMUM – [1, 1]



Σχήμα 2.28: 2D visualisation – γ_k method: MINIMUM – [1, 1]



Σχήμα 2.29: 3D visualisation – γ_k method: ARMIJO – [1, 1]



Σχήμα 2.30: 2D visualisation – γ_k method: ARMIJO – [1, 1]

2.4 Γενικά συμπεράσματα

2.4.1 Αρχικό σημείο

Για το πρώτο αρχικό σημείο που δίνεται, το $x_{0_1} = (0, 0)$ υπολογίζουμε ότι :

$$f(x_{0_1}) = 0 \quad \nabla f(x_{0_1}) = [0, 0]^T$$

Συνεπώς το $\text{norm}(\nabla f(x_k))$ ήδη από την πρώτη κιόλας επανάληψη είναι $< \epsilon$ και ο αλγόριθμος τερματίζει, χωρίς πρακτικά να έχει βρει το πραγματικό ελάχιστο της συνάρτησης.

Αντίστοιχα με το $x_{0_3} = (1, 1)$ μπορούμε να παρατηρήσουμε από τα παραπάνω διαγράμματα σε όλες τις μεθόδους (για παράδειγμα στα διαγράμματα 2.12, 2.11 και 2.26, 2.25) ότι η συνάρτηση ξεκινάει αρχικά από τα θετικά της $f(x, y)$ και συγκεκριμένα από το σημείο

$$f(x_{0_3}) = 0.1353 > 0$$

και κινείται προς το ελάχιστο. Ωστόσο, εγκλωβίζεται στο τοπικό ελάχιστο επίπεδο 0 και επειδή προκύπτει

$$\nabla f(x_k) = [0, 0]^T$$

το $\text{norm}(\nabla f(x_k))$ γίνεται $< \epsilon$ και ο αλγόριθμος τερματίζει, χωρίς πρακτικά να έχει βρει το πραγματικό ελάχιστο της συνάρτησης.

Τέλος, όταν το αρχικό σημείο είναι το $x_{0_2} = (-1, -1)$, ο αλγόριθμός μας ξεκινάει από το σημείο

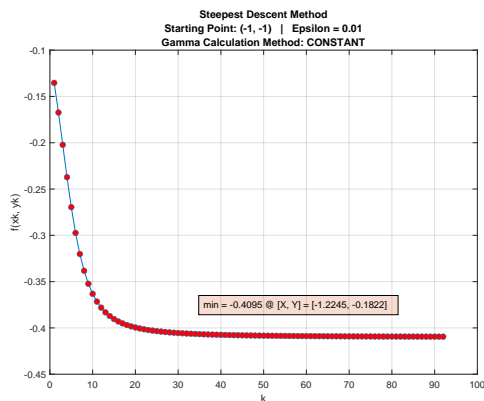
$$f(x_{0_2}) = -0.1353 < 0$$

και κινείται προς το ελάχιστο, το οποίο είναι και το ολικό της $f(x, y)$, όπως μπορούμε να παρατηρήσουμε στα παραπάνω διαγράμματα, όπως για παράδειγμα στα διαγράμματα 2.6, 2.5 και 2.20, 2.19.

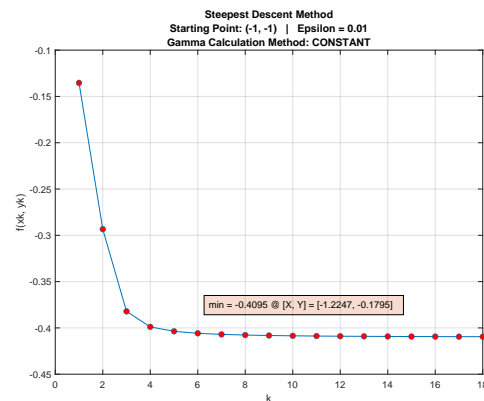
2.4.2 Αριθμός επαναλήψεων

Έχοντας θέσει στον κανόνα Armijo τις σταθερές που αναφέρθηκαν στην αρχή αυτού του κεφαλαίου και θεωρώντας $\gamma_0 = 0.3$, είναι προφανές από τα διαγράμματα ότι οι ελάχιστες επαναλήψεις για την εύρεση του (τοπικού) ελαχίστου παρουσιάζονται όταν το γ υπολογίζεται με βάση τη δεύτερη μέθοδο (ελαχιστοποίησης). Ενδεικτικά μπορεί κανείς να το παρατηρήσει αυτό από τα διαγράμματα 2.2, 2.4, 2.6 της μεθόδου Μεγίστης Καθόδου και τα διαγράμματα 2.20, 2.22, 2.24 της μεθόδου Levenberg-Marquardt.

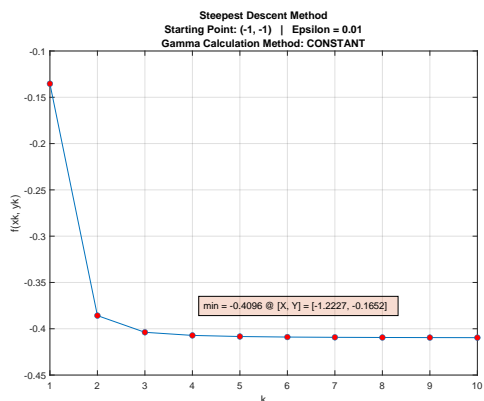
Με μεταβολή του $\gamma_0 = ct$ είναι προφανές ότι αλλάζει και ο αριθμός επαναλήψεων, ωστόσο η αύξηση του γ_0 δεν μας δίνει απαραίτητα κάποια πληροφορία για το πώς επηρεάζεται ο αριθμός επαναλήψεων. Παρακάτω μπορούμε να δούμε τη μεταβολή στον αριθμό των επαναλήψεων στη μέθοδο Μεγίστης Καθόδου χρησιμοποιώντας 4 διαφορετικά, αλλά σταθερά γ , με αρχικό σημείο το $(-1, -1)$:



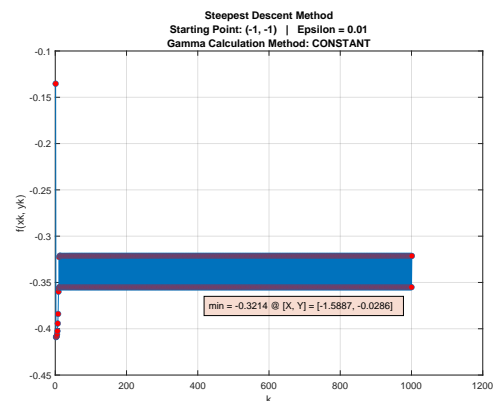
Σχήμα 2.31: $\gamma_k = 0.1$



Σχήμα 2.32: $\gamma_k = 0.5$



Σχήμα 2.33: $\gamma_k = 1$



Σχήμα 2.34: $\gamma_k = 1.5$

Είναι εύκολα αντιληπτό ότι για $\gamma_k = 0.1$ οι επαναλήψεις που εκτελούνται είναι σημαντικά περισσότερες από τη χρήση $\gamma_k = 0.5$ *or* $\gamma_k = 1$, ωστόσο για $\gamma_k = 1.5$ παρατηρούμε ότι η μέθοδος ταλαντώνεται μεταξύ 2 τιμών και δεν μπορεί να καταλήξει σε κάποιο ελάχιστο (ο αλγόριθμος τερματίζει στις 1000 επαναλήψεις).

Κεφάλαιο 3

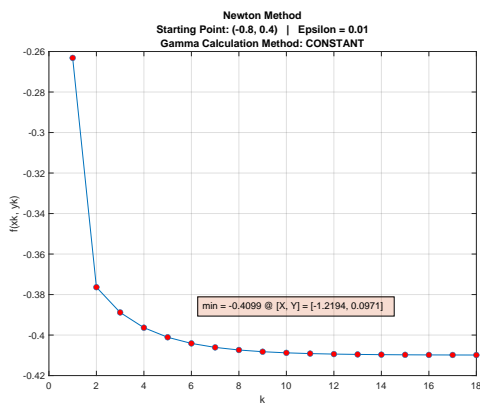
Σύγκριση Μεθόδων

3.1 Γενικές παρατηρήσεις

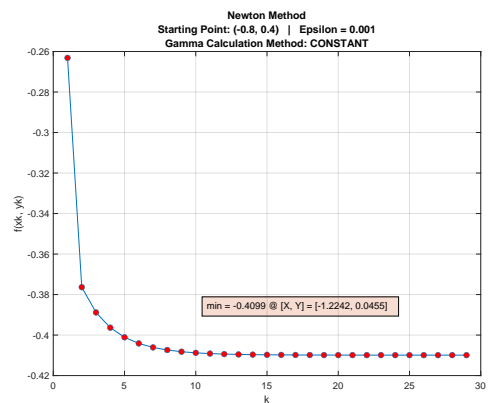
Όσον αφορά στα γραφήματα που παρατέθηκαν στο προηγούμενο κεφάλαιο, αναφέρουμε ότι έχει χρησιμοποιηθεί όριο $epsilon = 0.01$ για την ολοκλήρωση του αλγορίθμου και την προσεγγιστική εύρεση του ελαχίστου όταν:

$$norm(\nabla f(x_k)) < epsilon$$

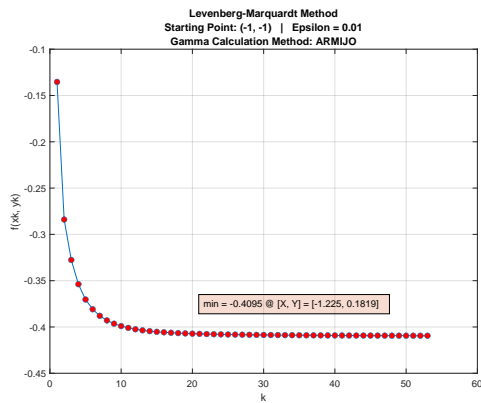
Είναι ευνόητο ότι η περαιτέρω μείωση αυτού του ορίου θα οδηγήσει τους αλγορίθμους σε εκτέλεση περισσότερων βημάτων προς την ολοκλήρωσή τους, ωστόσο η τελική προσέγγιση του ελαχίστου θα είναι πολύ πιο ακριβής. Ενδεικτικά παρατίθενται παρακάτω γραφήματα που υποδηλώνουν τη διαφορά στις μεθόδους για διαφορετικά $epsilon$.



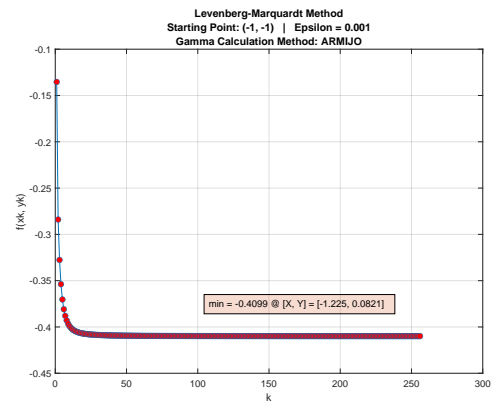
Σχήμα 3.1: Newton – $\epsilon = 0.01$



Σχήμα 3.2: Newton – $\epsilon = 0.001$



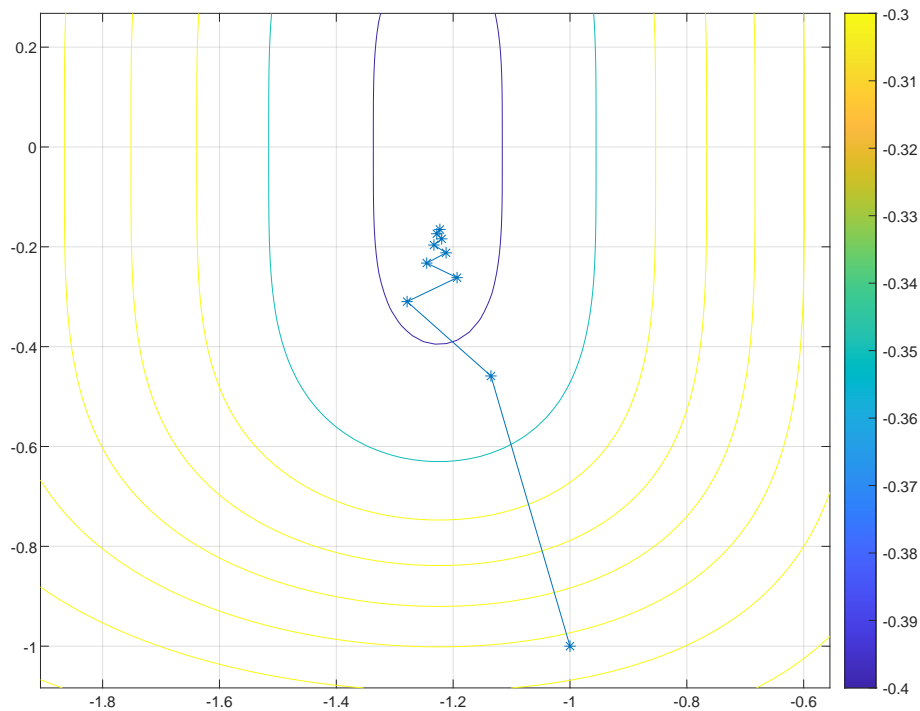
Σχήμα 3.3: L.M. - $\epsilon = 0.01$



Σχήμα 3.4: L.M. - $\epsilon = 0.001$

Εύκολα παρατηρούμε ότι οι μείωση του ϵ οδηγεί σε σημαντική **αύξηση των επαναλήψεων** μέχρι να υπολογιστεί το ελάχιστο.

Παρατηρώντας τα διαγράμματα που προκύπτουν από τις 3 μεθόδους που μελετήσαμε, διαπιστώνουμε ότι η μέθοδος της μέγιστης καθόδου, αν και αυτή με τους απλούστερους υπολογισμούς, απαιτεί σημαντικά **περισσότερα βήματα** από τις υπόλοιπες 2, καθυστερώντας της σύγκλιση στο (τοπικό) ελάχιστο. Αυτό συμβαίνει, καθώς η πορεία προς το ελάχιστο είναι μια **τεθλασμένη γραμμή**, με τα διαδοχικά ευθύγραμμα τμήματα να σχηματίζουν γωνία 90 deg, όπως μπορούμε να διακρίνουμε στο παρακάτω διάγραμμα:



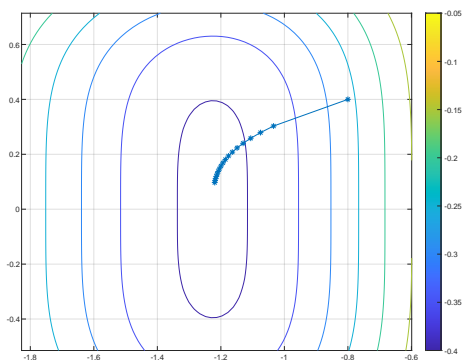
Σχήμα 3.5: Zik-Zak during Steepest Descent method

Αντίστοιχα η μέθοδος Newton, αν και **συγκλίνει ταχύτερα** στο (τοπικό) ελάχιστο, χρειάζεται θετικά ορισμένο εσσιανό πίνακα, συνθήκη που δεν μπορούμε να εξασφαλίσουμε σε κάθε επανάληψη και, όπως αναφέρθηκε στο προηγούμενο κεφάλαιο, κυρίως δεν μπορούμε να εξασφαλίσουμε για όλες τα σημεία εκκίνησης του αλγορίθμου.

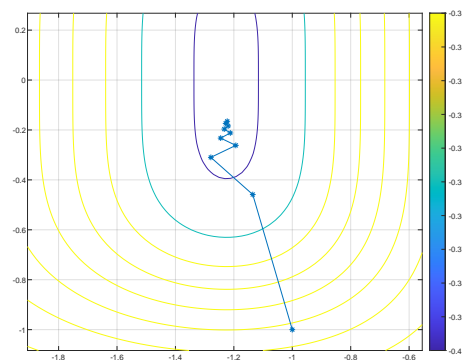
Η μέθοδος Levenberg-Marquardt με τη σειρά της **διορθώνει** το παραπάνω πρόβλημα που εμφανίζεται στη μέθοδο Newton, ωστόσο τα επιλεγόμενα γ_k περιορίζονται αρκετά από τα κριτήρια που χρειάζεται να ικανοποιούνται σε κάθε περίπτωση.

Επίσης, είναι σημαντικό να τονίσουμε ότι και οι 3 μέθοδοι έχουν τοπικό χαρακτήρα, καθώς **εγκλωβίζονται στην γειτονιά του τοπικού ελαχίστου**, επομένως δεν μπορούμε να είμαστε σίγουροι για το ότι το ελάχιστο που υπολογίστηκε είναι ολικό. Το τοπικό ελάχιστο στο οποίο εγκλωβίζονται οι μέθοδοι εξαρτάται άμεσα από το αρχικό σημείο που ξεκινάει η διαδικασία.

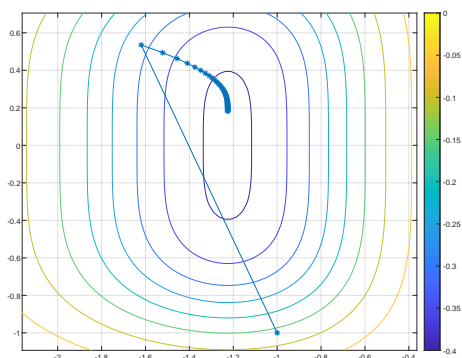
Αξίζει να τονίσουμε βέβαια ότι από την υλοποίηση των διαφόρων μεθόδων διαπιστώθηκε ότι υπάρχουν κοινά σημεία και όμοια συμπεριφορά. Πιο συγκεκριμένα, η μέθοδος των Levenberg – Marquardt για μικρό μ_k συμπεριφέρεται σαν την Newton και για μεγάλες τιμές του μ_k , λειτουργεί όπως η μέθοδος της Μέγιστης Καθόδου. Μπορούμε να παρατηρήσουμε τις ομοιότητες μεταξύ των μεθόδων στα παρακάτω διαγράμματα :



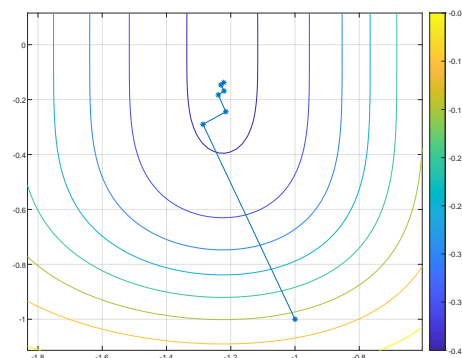
Σχήμα 3.6: Newton converge



Σχήμα 3.7: Steepest Descent converge



Σχήμα 3.8: L.M. converge (sim. Newton)



Σχήμα 3.9: L.M. converge (sim. S.D.)

3.2 Επιλογή του γ

Όλες οι παραπάνω μέθοδοι για σταθερό γ οδηγήθηκαν στο ελάχιστο, ωστόσο αυτό συνέβη έπειτα από κατάλληλη επιλογή από μέρους μας. **Ανάλογα με την επιλογή του γ** οι αλγόριθμοι μπορεί να καταλήγουν μετά από > 500 βήματα στο ελάχιστο ή μπορεί και να ταλαντώνονται μεταξύ 2 σημείων και να τερματίζουν λόγω του ανώτατου ορίου βημάτων (fig. 2.34).

Με τον κανόνα Armijo όσο πιο μακριά βρισκόμαστε από το σημείο ελαχίστου τόσο μεγαλύτερα βήματα θέλουμε να κάνουμε. Ωστόσο, η αρχική επιλογή του γ επηρεάζει σημαντικά τα βήματα μέχρι τη σύγκλιση στο ελάχιστο, καθώς μπορεί να οδηγηθούμε σε μεγαλύτερη ισοβαρή καμπύλη, άρα να συγκλίνουμε πιο αργά. Φυσικά, τα αποτελέσματα μεταβάλλονται για αλλαγές των παραμέτρων α, β, s , οι οποίες παρέμειναν **σταθερές** κατά την εκπόνηση αυτής της εργασίας.

Τέλος, με την ελαχιστοποίηση της συνάρτησης φαίνεται πετυχαίνουμε τα καλύτερα αποτελέσματα σε όλες τις μεθόδους, δεδομένων των συνθηκών που λάβαμε υπόψιν σε κάθε μέθοδο.