**Τεχνικές Βελτιστοποίησης**

**Εργασία 2**

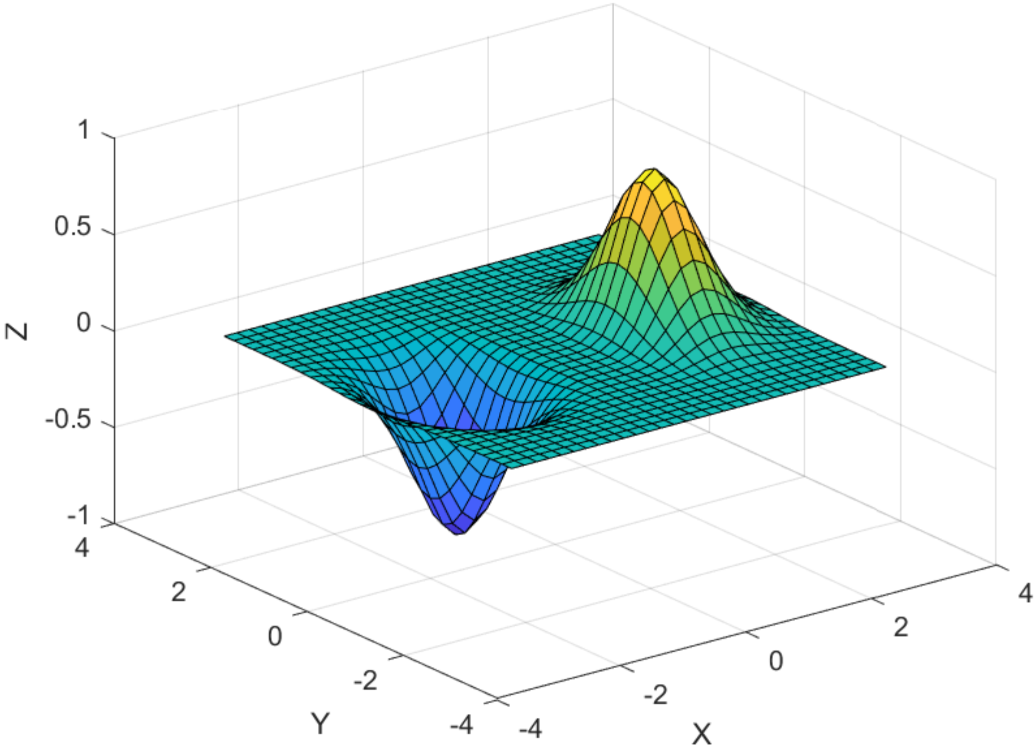
**Ηλιάνα Κόγια (10090)**

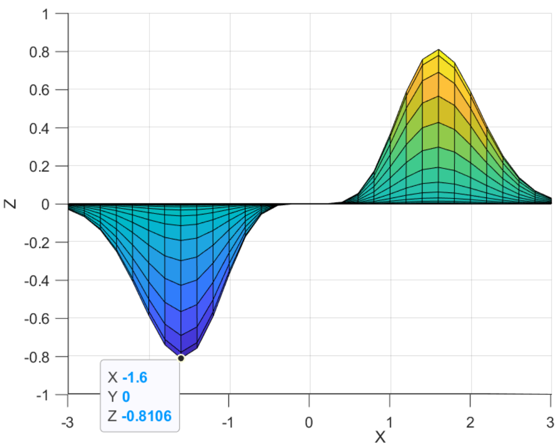
*Στη συγκεκριμένη εργασία* καλούμαστε να ελαχιστοποιήσουμε μια συνάρτηση δύο μεταβλητών χωρίς περιορισμούς. Οι αλγόριθμοι που χρησιμοποιούνται βασίζονται στην ιδέα της επαναληπτικής μεθόδου.

## Θέμα 1

Παρακάτω παρατηρούμε τη μορφή της δοθείσας συνάρτησης:

$f(x, y) = 5^x \exp(-x^2-y^2) = z$



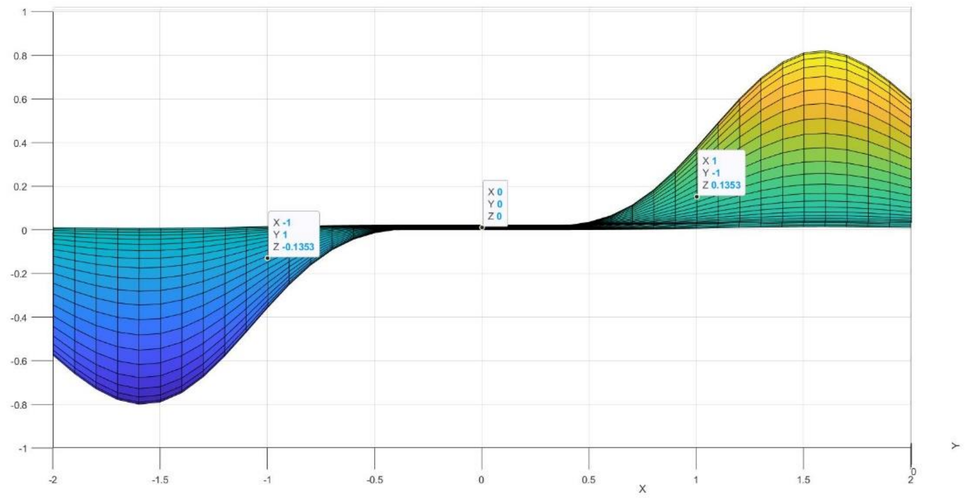

Στην αριστερή εικόνα φαίνεται το σημείο όπου η αντικειμενική συνάρτηση ελαχιστοποιείται.

f(-1.6, 0) = -0.8106

**Θέμα 2**

Ο πρώτος αλγόριθμος για την ελαχιστοποίηση της f, είναι η Μέθοδος Μέγιστης Καθόδου (Steepest Descent).

*Αρχικά σημεία:*




| $x_o$ | 0 | -1 | +1 |
|---|---|---|---|
| $y_o$ | 0 | +1 | -1 |
| f | 0 | -0.1353 | +0.1353 |

*Επιλέγουμε ακρίβεια ε = 0,001

Διερευνούμε για καθένα αρχικό σημείο τις παρακάτω περιπτώσεις του βήματος γ:

a) Για σταθερό βήμα $\gamma_k$

**$(x_o, y_o) = (0, 0)$**

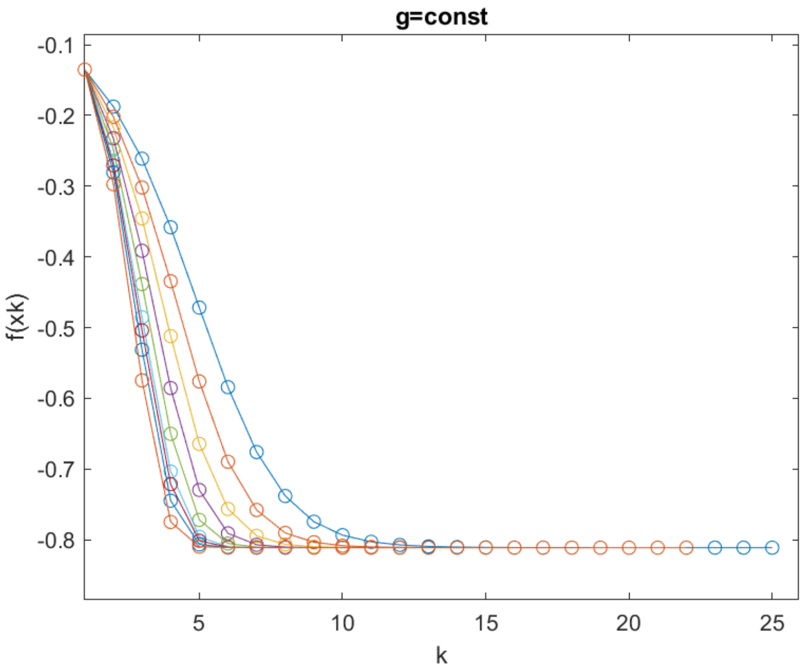
Ο αλγόριθμος εγκλωβίζεται και δεν παράγει το σωστό αποτέλεσμα, διότι το gradient της f(x,y) ισούται με 0 οπότε τερματίζει στην πρώτη επανάληψη.

*Αυτό θα συμβαίνει σε οποιαδήποτε από τις 3 περιπτώσεις για το βήμα γ. Οπότε μελετάμε τα άλλα δύο αρχικά σημεία στη συνέχεια.*

**$(x_o, y_o) = (-1, +1)$**

Ο αριθμός των επαναλήψεων k που απαιτούνται, ώστε να υπολογίσει ο αλγόριθμος το ελάχιστο της συνάρτησης, για τη συγκεκριμένη ακρίβεια ε, διαφέρει ανάλογα με τη (σταθερή) τιμή του βήματος γ. Μετά από δοκιμές παρατηρείται πως πιο αποδοτική ως προς τον αριθμό των επαναλήψεων είναι η τιμή **γ = 0,47** με **k=9**. Για το αρχικό σημείο (-1,1) η μέθοδος δε συγκλίνει στο ελάχιστο για μεγάλες τιμές του γ, ενώ για πολύ μικρά γ η μέθοδος απαιτεί πιο πολλές επαναλήψεις.

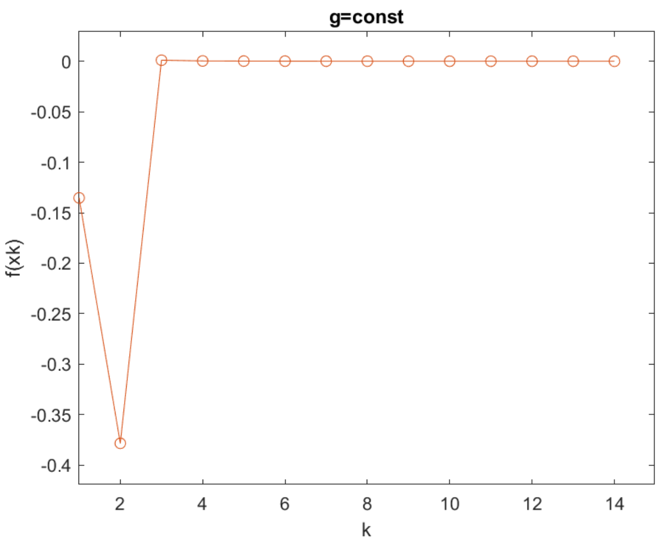
*Σύγκλιση συνάρτησης ως προς k, για διάφορες τιμές του γ*




| **γ** | 0.2 | 0.25 | 0.3 | 0.35 | 0.4 | 0.45 | **0.47** | 0.5 | 0.55 |
|---|---|---|---|---|---|---|---|---|---|
| **k** | 25 | 19 | 16 | 13 | 11 | 10 | **9** | 12 | 22 |

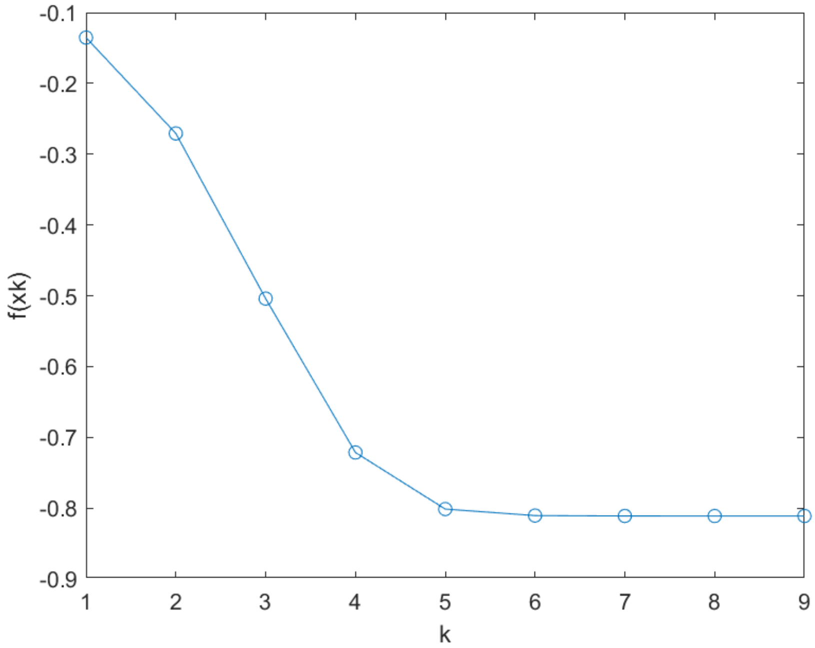
** Για μεγάλες τιμές του γ (π.χ. γ=3) η συνάρτηση με αρχικό σημείο (-1,1) *<u>δεν συγκλίνει στο -0.81, αλλά στο 0:</u>*

γ = 3



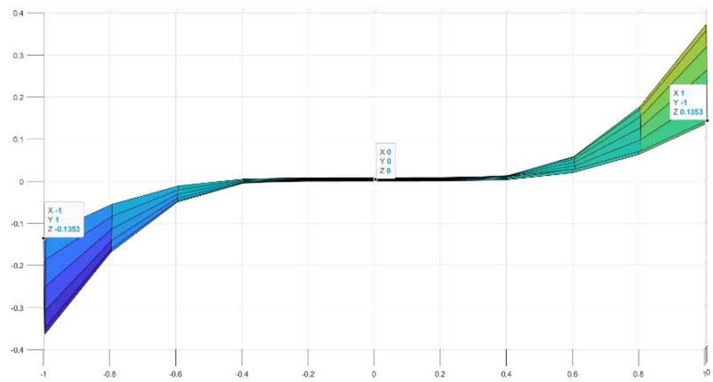
*Σύγκλιση συνάρτησης ως προς k*

**γ = 0,47**




**$(x_o, y_o)$ = (+1, -1)**

Για σταθερό βήμα $\gamma_k$ , αν έχουμε σαν αφετηρία το σημείο (1,-1) ο αλγόριθμος Μέγιστης Καθόδου δεν βρίσκει το ελάχιστο της συνάρτησης, αλλά εγκλωβίζεται στην επιφάνεια που ισχύει z = 0. Αυτό συμβαίνει, καθώς το σημείο (1,-1) βρίσκεται δεξιά της επιφάνειας που ισχύει f(x,y) = 0, οπότε δεν υπάρχει σταθερό βήμα για κάθε επανάληψη ώστε να μην εγκλωβιστεί σε αυτό το τοπικό ελάχιστο. Το σχήμα επεξηγεί και γραφικά τα παραπάνω:




Ακόμη και για μεγάλες τιμές του γ μπορεί να προσπερνά το f=0, αλλά ξανά δεν βρίσκει το ελάχιστο, καθώς έχοντας μεγάλο βήμα προσπερνά και το ελάχιστο και φτάνει ξανά σε επιφάνεια με f = 0 πιο αριστερά του ελαχίστου.

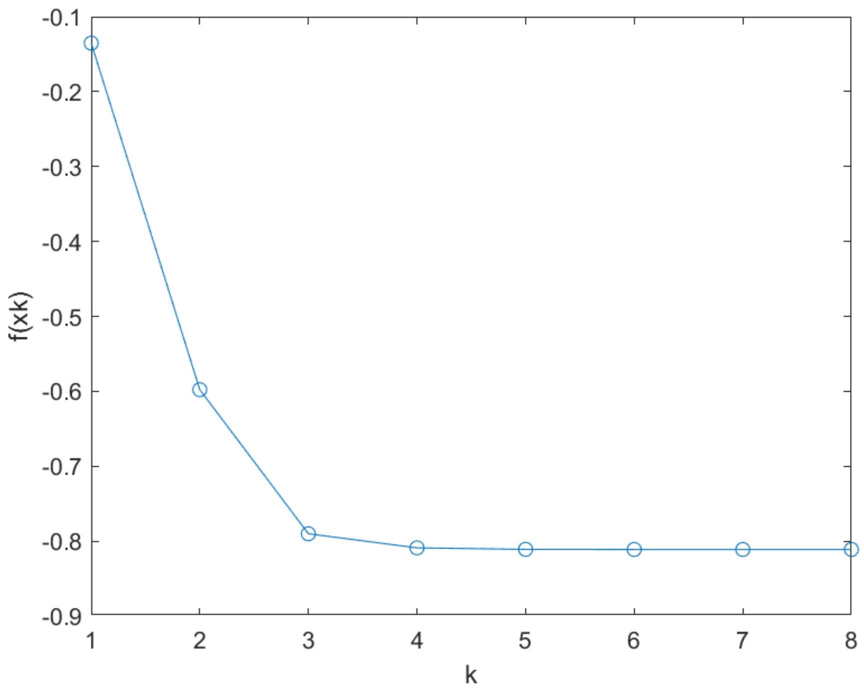
b) Για βήμα $\gamma_k$ ώστε να ελαχιστοποιείται η συνάρτηση $\phi(\gamma) = f(x_k + \gamma_k d_k)$

Κάνοντας χρήση κάποιον αλγόριθμο ελαχιστοποίησης, όπως Fibonacci ή Διχοτόμο με χρήση παραγώγων, για διάστημα αναζήτησης (0, w) βρίσκουμε σε κάθε επανάληψη την τιμή του γ που ελαχιστοποιεί τη συνάρτηση $f(x_k + \gamma_k d_k)$.

**$(x_o, y_o)$ = (-1, +1)**

Μετά από δοκιμές καταλήγουμε σε διάστημα αναζήτησης (0, 2). Τότε οι απαιτούμενες επαναλήψεις για την ελαχιστοποίηση της f είναι **k=8**

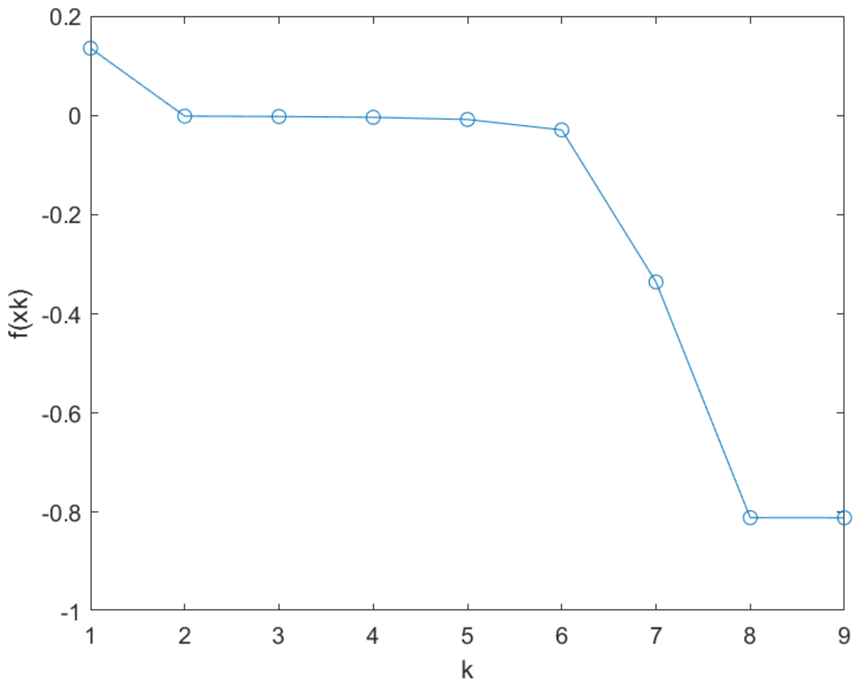
*Σύγκλιση συνάρτησης ως προς k με <u>διάστημα αναζήτησης στη Fibonacci για την εύρεση του $\gamma_{min}$</u> :* **(0, 2)**




**$(x_o, y_o)$ = (+1, -1)**

Μετά από δοκιμές καταλήγουμε σε διάστημα αναζήτησης (0, 8). Για το αρχικό σημείο (1,-1) δεν λειτουργεί για διάστημα αναζήτησης (0,2), διότι το τοπικό ελάχιστο $\gamma_{min}$ που επιστρέφει η μέθοδος Fibonacci είναι αρκετά μικρό και έτσι εγκλωβίζεται, ενώ για διάστημα αναζήτησης (0,8) παράγει μεγαλύτερο $\gamma_{min}$ και έτσι συγκλίνει τελικά η Μέθοδος Μέγιστης Καθόδου στο -0,81. Τότε οι απαιτούμενες επαναλήψεις για την ελαχιστοποίηση της f είναι **k=9**.

*Σύγκλιση συνάρτησης ως προς k με <u>διάστημα αναζήτησης για την εύρεση του $\gamma_{min}$</u> :* **(0, 8)**




Παρατηρούμε, ότι ο αλγόριθμος, για αυτή την περίπτωση του βήματος γ, υπολογίζει τελικά το σωστό ελάχιστο της αντικειμενικής συνάρτησης και για αρχικό σημείο (1,-1), χωρίς να εγκλωβίζεται σε κάποιο τοπικό ελάχιστο.

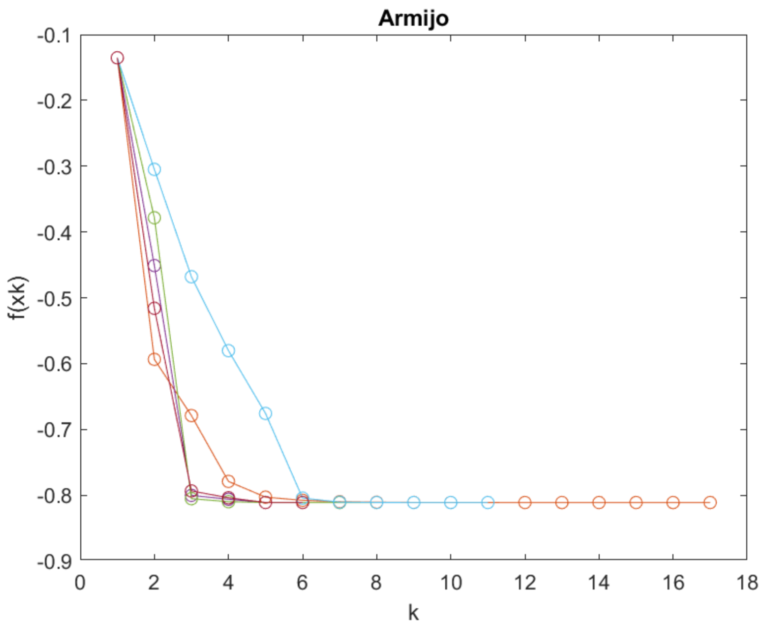
c) Για βήμα $\gamma_k$ βάσει του κανόνα Armijo, επιλέγουμε **a = 10^-2 και b=1/2** (μετά από δοκιμές και σφάλματα για διάφορους συνδυασμούς a, b, s).

Ανάλογα με το αρχικό βήμα s, για τα δεδομένα a και b ο αριθμός των επαναλήψεων μεταβάλλεται κάθε φορά.

**$(x_o, y_o)$ = (-1, +1)**

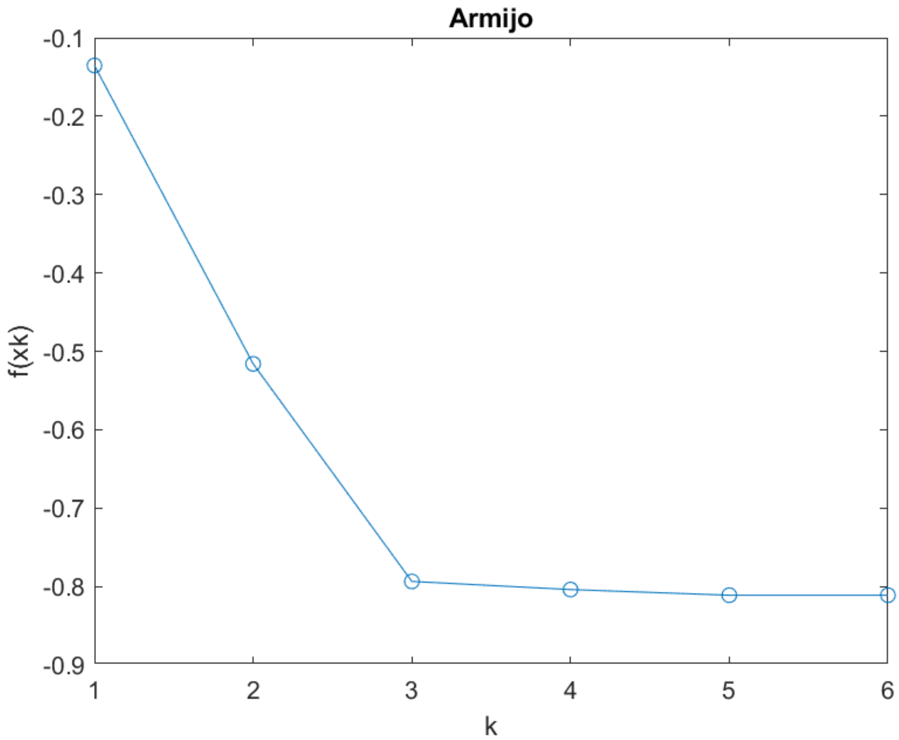
| **s** | 2 | 4 | **5** | 5.5 | 6 | 6.5 |
|---|---|---|---|---|---|---|
| **k** | 17 | 17 | **6** | 8 | 9 | 11 |

*Διάγραμμα Σύγκλισης ως προς k*




Ο βέλτιστος αριθμός επαναλήψεων που βρέθηκε, ώστε ο αλγόριθμος της Μέγιστης Καθόδου να βρίσκει το ελάχιστο -0,8112 με τον κανόνα Armijo για αρχικό σημείο (-1,1) , είναι **k=6**

*Διάγραμμα Σύγκλισης ως προς k με **s = 5, a = 10^-2 και b=1/2***

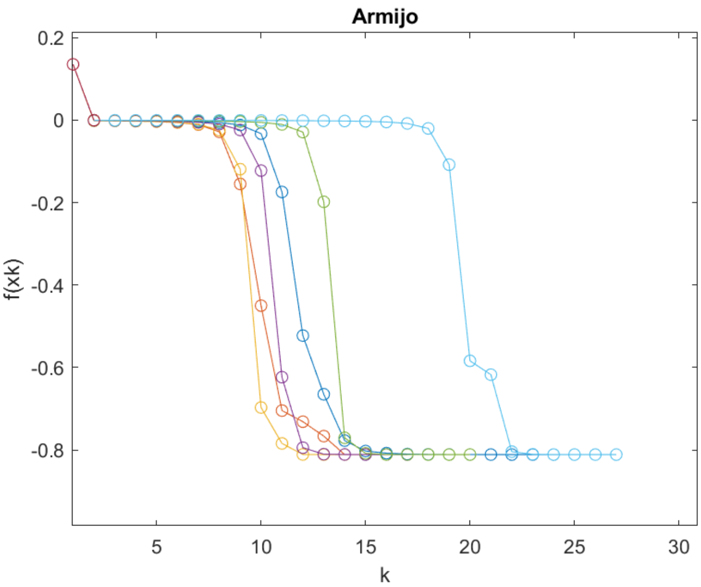


**$(x_o, y_o) = (+1, -1)$**

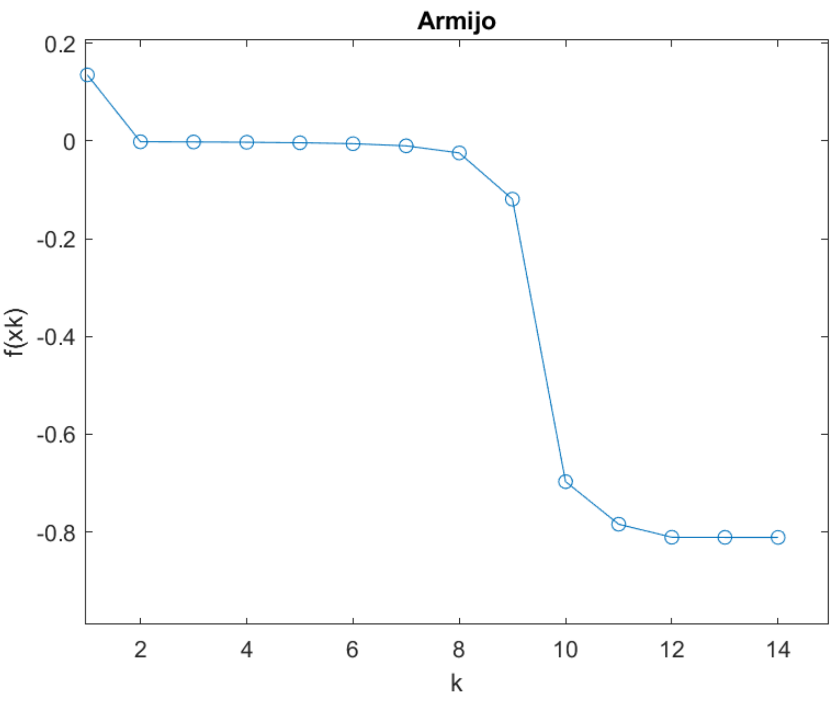
Επιλέγουμε ξανά **a = 10^-2 και b=1/2**. Για διάφορες τιμές του s, ο αριθμός των επαναλήψεων μεταβάλλεται. (Για ορισμένες τιμές των s μπορεί να μην εμφανίζει το σωστό ελάχιστο, αλλά το f=0, γι' αυτό με τη μέθοδο trial and error βρίσκουμε τους καλύτερους συνδυασμούς των a,b,s που μας οδηγούν στο ελάχιστο -0,81)

| | | | | | | |
|---|---|---|---|---|---|---|
| **s** | 4 | 5 | **5.1** | 5.5 | 6 | 6.5 |
| **k** | 27 | 15 | **14** | 17 | 20 | 27 |

*Διάγραμμα Σύγκλισης ως προς k, για διάφορα s*




*Διάγραμμα Σύγκλισης ως προς k με **s = 5.1, a = 10^-2 και b=1/2***



**Παρατηρήσεις**

Από τα παραπάνω διαπιστώνουμε ότι η επιλογή του αρχικού σημείου επηρεάζει τον αριθμό των επαναλήψεων σε κάθε περίπτωση επιλογής βήματος γ. Λιγότερη επιρροή παρουσιάζει για τη (b) περίπτωση που βρίσκουμε το $γ_{min}$, όπου ο βέλτιστος αριθμός των επαναλήψεων k που βρέθηκε διαφέρει κατά 1 για τα αρχικά σημεία (-1,1) και (1,-1).

Συγκεκριμένα, για σταθερό γ η επιλογή του αρχικού σημείου είναι μείζονος σημασίας, διότι είναι πιθανό να εγκλωβιστεί σε κάποιο τοπικό ελάχιστο. Αν το βήμα επιλεγεί πολύ μεγάλο μπορεί να παρατηρηθεί και απόκλιση της διαδικασίας αναζήτησης, ενώ αν επιλεγεί πολύ μικρό ο ρυθμός σύγκλισης μικραίνει αυξάνοντας σημαντικά των αριθμό των επαναλήψεων. Μπορούμε και για γ σταθερό να πετύχουμε αριθμό επαναλήψεων k παρόμοιο με τις άλλες μεθόδους επιλογής βήματος, ωστόσο χρειάζεται αρκετές δοκιμές, δεν είναι πάντα εφικτό, και για διαφορετικά αρχικά σημεία διαφέρει το σταθερό βήμα που δίνει "καλό" αριθμό επαναλήψεων. Για τη συγκεκριμένη συνάρτηση δεν βρέθηκε κάποιο σταθερό βήμα που να μας δίνει το σωστό ελάχιστο για αρχικό σημείο (+1,-1).

Η περίπτωση βήματος με την ελαχιστοποίηση της φ(γ) είναι αρκετά αποδοτική, καθώς με αυτή την μέθοδο βήματος και για τα δύο αρχικά σημεία ο αλγόριθμος Steepest Descent παράγει το σωστό ελάχιστο. Επίσης, απαιτούνται λιγότερες δοκιμές, καθώς χρειάζεται να βρούμε μόνο το διάστημα αναζήτησης (0,w), w>0. Για τα διαστήματα αναζήτησης στη μέθοδο της Fibonacci (0,5) (0,6) (0,7) (0,8) καταλήγουμε στο σωστό ελάχιστο -0,81 και για τα δύο αρχικά σημεία.

Η επιλογή βήματος με τον κανόνα Armijo μας δίνει επίσης μικρό αριθμό επαναλήψεων, ωστόσο είναι πιο πολύπλοκο σε σχέση με την δεύτερη περίπτωση μιας και απαιτείται η εύρεση κατάλληλου συνδυασμού των a, b, s, ώστε να παράγει η μέθοδος Steepest Descent το σωστό ελάχιστο με μικρό σχετικά αριθμό επαναλήψεων.

Να σημειώσουμε τέλος ότι για αρχικό σημείο (0,0) καμία μέθοδο επιλογής βήματος δεν λειτουργεί, καθώς το (0,0) αποτελεί κρίσιμο σημείο της συνάρτησης και ο αλγόριθμος τερματίζει στην πρώτη επανάληψη.

**Θέμα 3**

Ο δεύτερος αλγόριθμος για την ελαχιστοποίηση της f, είναι η Μέθοδος Newton.

Προϋπόθεση για τη σωστή λειτουργία της μεθόδου είναι ο εσσιανός πίνακας της f να είναι θετικά ορισμένος. Η υπόθεση αυτή προσδίδει στη μέθοδο Newton την ιδιότητα της επαναληπτικής καθόδου.

Στη συγκεκριμένη συνάρτηση δεν προκύπτει θετικά ορισμένος ο εσσιανός πίνακας ακόμη και για τα αρχικά σημεία. Οπότε η μέθοδος σε αυτή την περίπτωση δίνει είτε λύσεις μακριά από το σημείο ολικού ελαχίστου, είτε αποκλίνει, για οποιαδήποτε επιλογή βήματος γ.

Έτσι, μπορεί να χρησιμοποιηθεί ο παρακάτω τροποποιημένος αλγόριθμος Newton για να αντιμετωπιστεί αυτό το πρόβλημα.

**<u>Θέμα 4</u>**

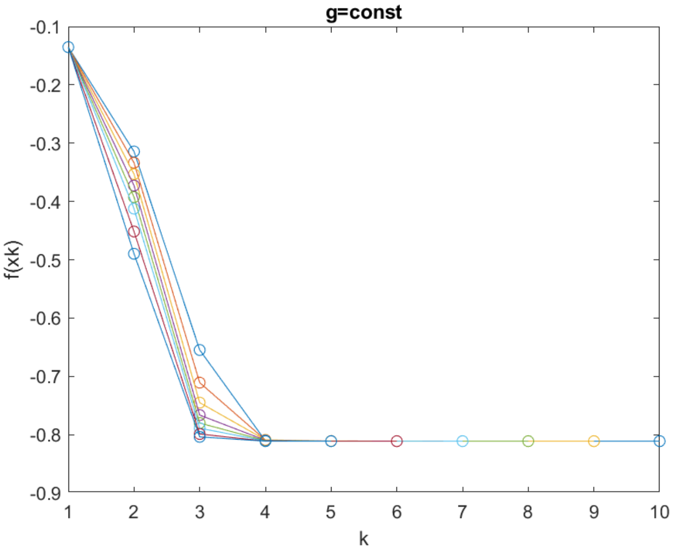
Ο δεύτερος αλγόριθμος για την ελαχιστοποίηση της f, είναι η Μέθοδος Levenberg-Marquardt. Σε αυτή τη μέθοδο, υπολογίζουμε το $\mu_k$ ώστε ο $(\nabla^2 f(x_k) + \mu_k I)$ να είναι θετικά ορισμένος

Διερευνώνται για καθένα αρχικό σημείο οι παρακάτω περιπτώσεις του βήματος γ:

a) Για σταθερό βήμα $\gamma_k$

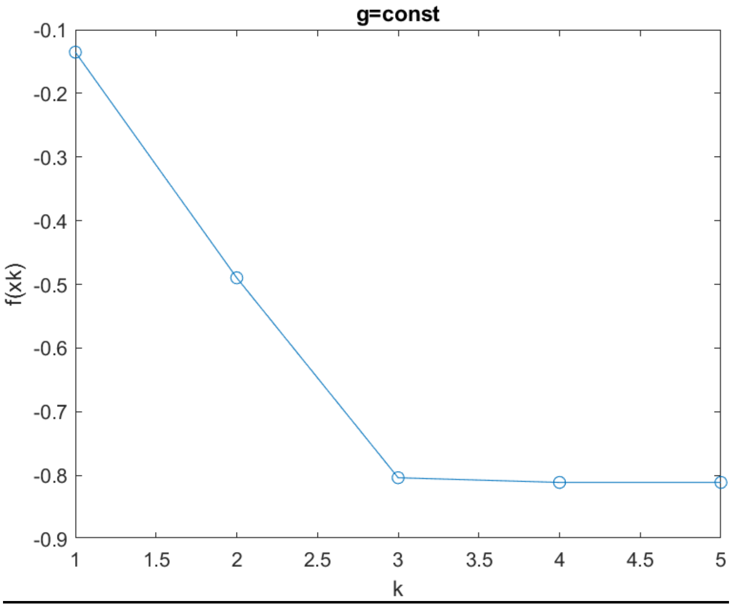
**$(x_o, y_o) = (-1, +1)$**

Δοκιμάζοντας διάφορες τιμές του γ επιλέγοντας τελικά εκείνη που μας δίνει το σωστό ελάχιστο σε μικρό σχετικά αριθμό επαναλήψεων.




| **γ** | 0.55 | 0.6 | 0.65 | 0.7 | 0.75 | 0.8 | 0.9 | **1** |
|---|---|---|---|---|---|---|---|---|
| **k** | 10 | 9 | 9 | 8 | 8 | 7 | 6 | **5** |

*Σύγκλιση συνάρτησης ως προς k με* **γ = 1**




**$(x_o, y_o) = (1, -1)$**

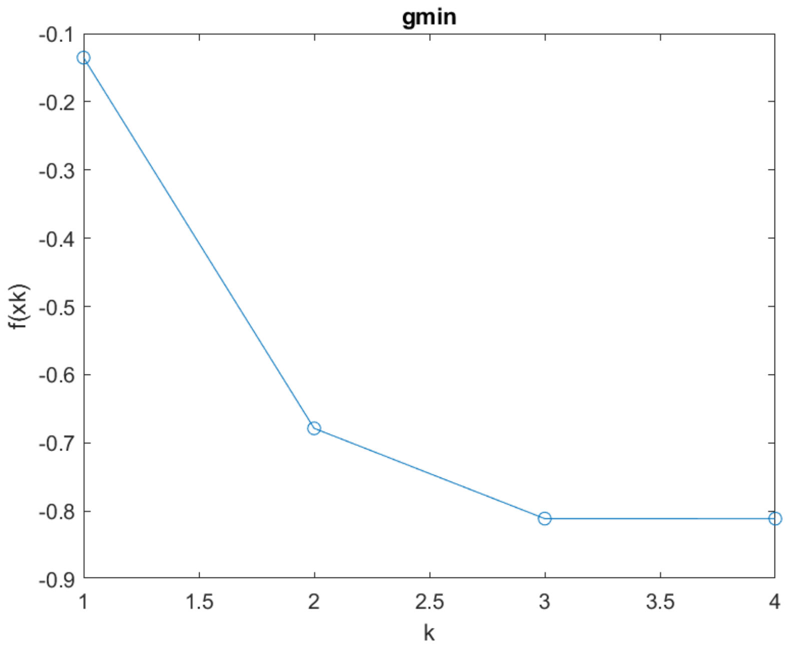
Για το συγκεκριμένο αρχικό σημείο δεν βρέθηκε σταθερό βήμα γ, ώστε ο αλγόριθμος Levenberg-Marquardt να παράγει το σωστό ελάχιστο για την αντικειμενική συνάρτηση, διότι εγκλωβίζεται στην επιφάνεια z=0. Ακόμα και να ξεπεράσει αυτή την επιφάνεια με μεγάλο βήμα, δεν καταφέρνει να βρει το ελάχιστο -0,81, διότι το βήμα είναι πολύ μεγάλο και το προσπερνά πηγαίνοντας ξανά σε επιφάνεια z=0.

b) Για βήμα $\gamma_k$ ώστε να ελαχιστοποιείται η συνάρτηση $\phi(\gamma) = f(x_k + \gamma_k d_k)$

**$(x_o, y_o) = (-1, +1)$**

Μετά από δοκιμές βρίσκουμε τα διαστήματα αναζήτησης για τη μέθοδο Fibonacci για την ελαχιστοποίηση της φ(γ).

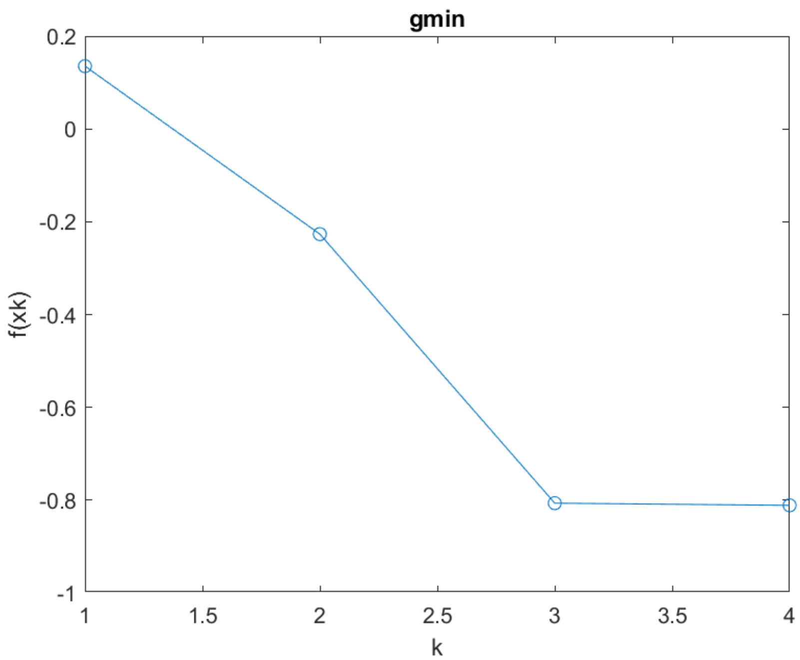
*Σύγκλιση συνάρτησης ως προς k με <u>διάστημα αναζήτησης για την εύρεση του $\gamma_{min}$</u> :* **(0, 8)**




Ο αριθμός των επαναλήψεων είναι **k = 4.**

**$(x_o, y_o) = (+1, -1)$**

*Σύγκλιση συνάρτησης ως προς k με <u>διάστημα αναζήτησης για την εύρεση του $\gamma_{min}$</u> :* **(0, 8)**




Ο αριθμός των επαναλήψεων είναι και για αυτό το αρχικό σημείο **k = 4.**

c) Για βήμα $\gamma_k$ βάσει του κανόνα Armijo, επιλέγουμε **a = 10^-2 και b=1/2** (μετά από δοκιμές και σφάλματα για διάφορους συνδυασμούς a, b, s).

Ανάλογα με το αρχικό βήμα s, για δεδομένα a και b ο αριθμός των επαναλήψεων μεταβάλλεται κάθε φορά.

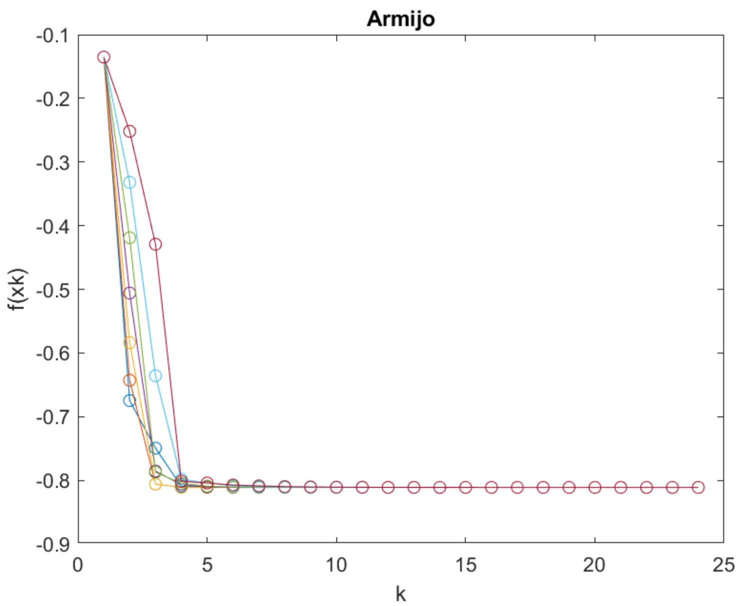
**$(x_o, y_o)$ = (-1, +1)**

| **s** | **4** | 4.5 | 5 | 5.5 | 6 | 6.5 | 7 |
|---|---|---|---|---|---|---|---|
| **k** | **5** | 6 | 7 | 10 | 12 | 17 | 24 |

Τα s που επιλέχθηκαν (με τη μέθοδο trial and error) είναι εκείνα που οδηγούν στο σωστό ελάχιστο μέσω του αλγορίθμου Levenberg-Marquardt. Για κάποια s ο αλγόριθμος μπορεί να αποκλίνει ή να εγκλωβίζεται σε σημείο που ισχύει f = 0.

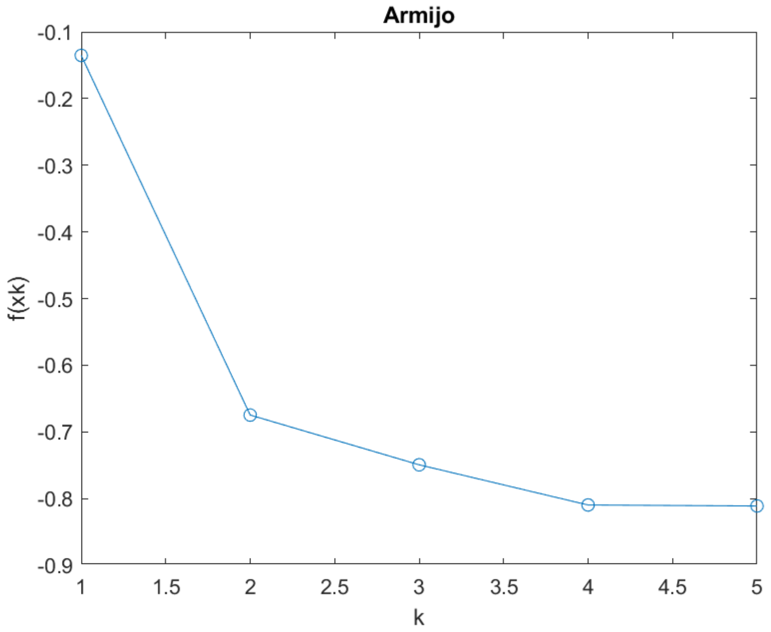
<u>Σημείωση: Ανάλογα και με τον τρόπο που βρίσκουμε το $\mu_k$ προγραμματιστικά μπορεί να διαφέρουν τα γ και s για τα οποία ο αλγόριθμος Levenberg παράγει σωστό ελάχιστο.</u>

*Διάγραμμα Σύγκλισης ως προς k,* για διάφορα s




Ο βέλτιστος αριθμός επαναλήψεων που βρέθηκε, ώστε ο αλγόριθμος Levenberg-Marquardt να βρίσκει το ελάχιστο -0,8112 με τον κανόνα Armijo, είναι **k=5**

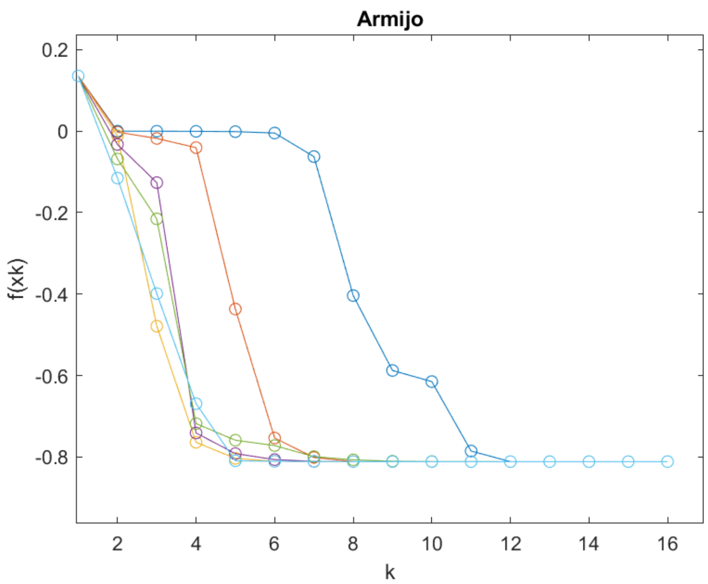
*Διάγραμμα Σύγκλισης ως προς k με* ***s = 4, a = 10^-2 και b=1/2***




**$(x_o, y_o)$ = (+1, -1)**

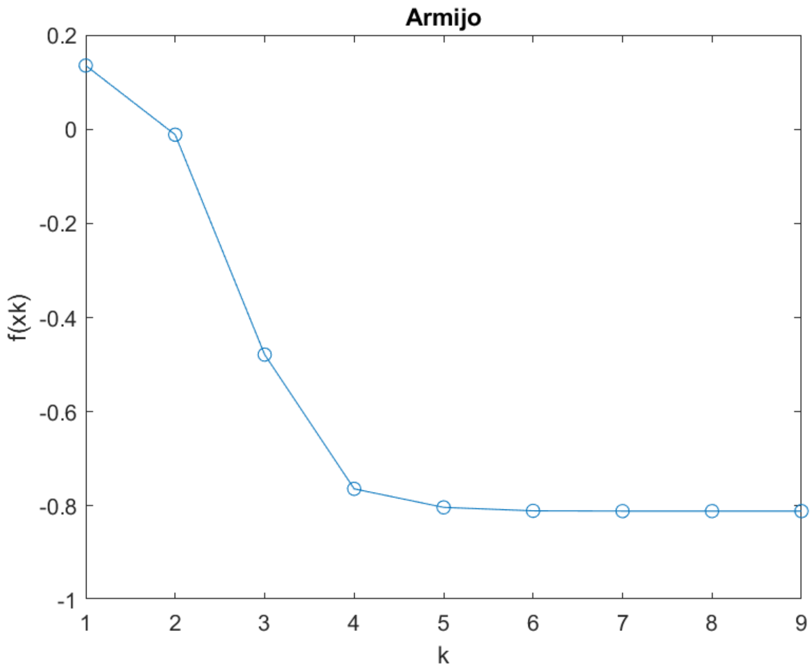
| | | | | | | |
|---|---|---|---|---|---|---|
| **s** | 4 | 4.5 | **5** | 5.5 | 6 | 6.5 |
| **k** | 13 | 10 | **9** | 12 | 16 | 16 |

*Διάγραμμα Σύγκλισης ως προς k,* για διάφορα s



Ο βέλτιστος αριθμός επαναλήψεων που βρέθηκε, ώστε ο αλγόριθμος της Μέγιστης Καθόδου να βρίσκει το ελάχιστο -0,8112 με τον κανόνα Armijo για το αρχικό σημείο (1,-1), είναι **k=9**

*Διάγραμμα Σύγκλισης ως προς k με* ***s = 4, a = 10^-2 και b=1/2***




## <u>Σύγκριση Μεθόδων (Steepest Descent & Levenberg-Marquardt)</u>

Για γ = σταθερό και οι δύο μέθοδοι δεν καταλήγουν σε σωστό ελάχιστο για το αρχικό σημείο (+1,-1), ενώ για το σημείο (-1,1) η μέθοδος Levenberg απαιτεί λιγότερες επαναλήψεις (k=5 < k=9) για επιλεγμένο γ. Ωστόσο, αυτό μπορεί να αλλάζει ανάλογα με τα σταθερά γ που επιλέγουμε κάθε φορά για κάθε μέθοδο και με το αρχικό σημείο.

Για γ που ελαχιστοποιεί τη συνάρτηση φ(γ) και οι δύο μέθοδοι λειτουργούν σωστά για τα αρχικά σημεία (-1,1) και (1,-1), με τη μέθοδο Levenberg να απαιτεί λιγότερο αριθμό επαναλήψεων (σχεδόν το μισό) για την εύρεση του ελαχίστου -0,81 (k=4 και για τα 2 αρχικά σημεία, ενώ η Steepest Descent k = 8 και k=9 αντίστοιχα, για τα εκάστοτε διαστήματα αναζήτησης που επιλέχθηκαν για τη Fibonacci).

Για γ μέσω του κανόνα Armijo με κατάλληλη επιλογή των s, a, b για καθέναν από τους δύο αλγορίθμους παρατηρούμε ότι η Levenberg είναι ελάχιστα καλύτερη και κυρίως για το σημείο (+1,-1). (Steepest k = 6 και k=14, Levenberg k = 5 και k=9 για τα σημεία (-1,1) και (1,-1) αντίστοιχα).