Τεχνικές Βελτιστοποίησης

Αναφορά για την 3η εργαστηριακή άσκηση

Θεόδωρος Λιούπης
ΑΕΜ 9733

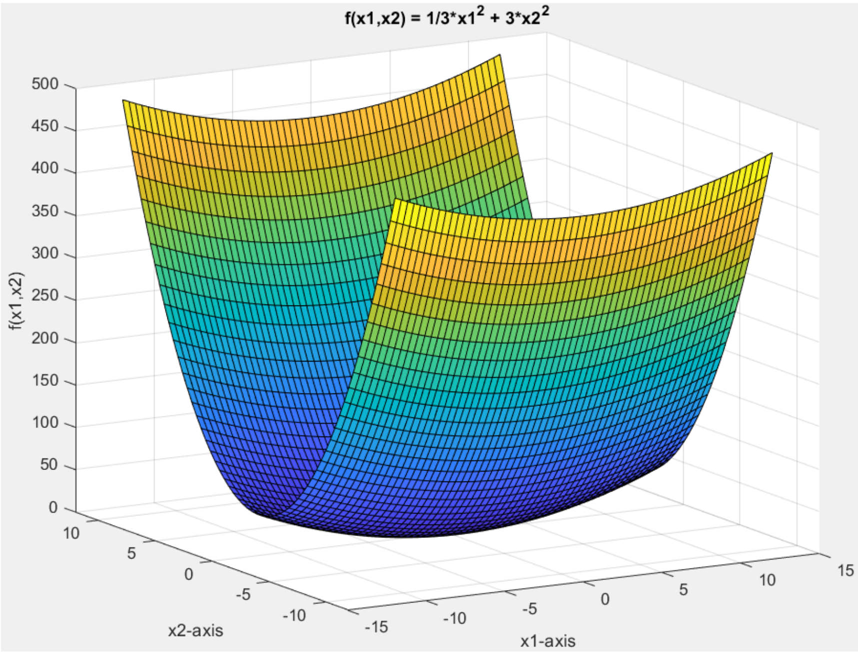
Χειμερινό εξάμηνο 2022-2023

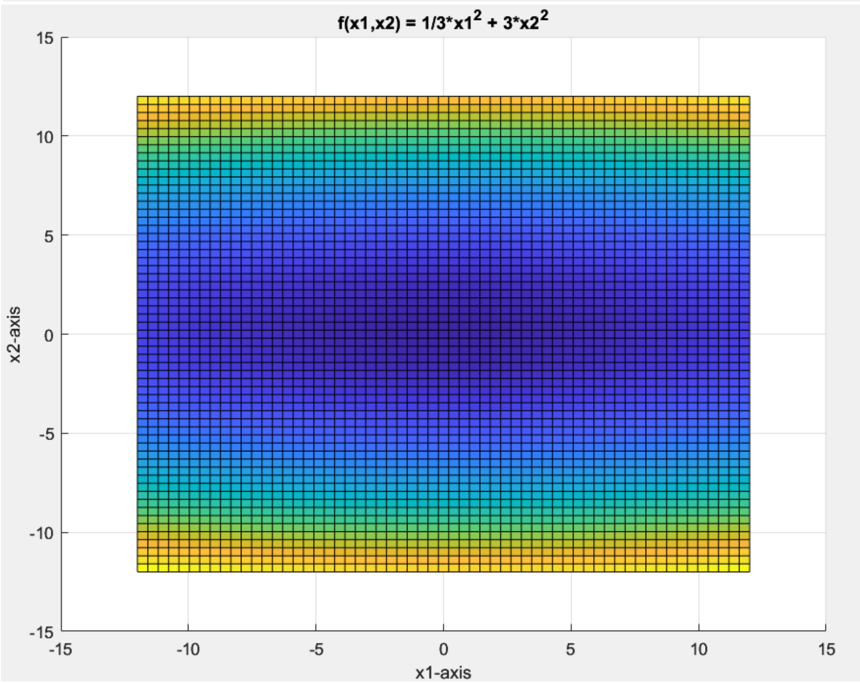
Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών ΑΠΘ

Ζητούμενο της $3^{ης}$ εργαστηριακής άσκησης ήταν: 1) η δοκιμή της Μεθόδου Μέγιστης Καθόδου χωρίς περιορισμούς σε συνάρτηση f:R^2 → R με τέσσερα διαφορετικά βήματα $\gamma_κ$ και τυχαίο σημείο εκκίνησης του αλγορίθμου, 2) η δοκιμή της Μεθόδου Μέγιστης Καθόδου με προβολή σε τρεις διαφορετικές περιπτώσεις (διαφορετικά $\gamma_κ$, $s_κ$, σημεία εκκίνησης).
Η συνάρτηση στην οποία έγιναν οι δοκιμές είναι η εξής:

$$f: \mathbb{R}^2 \rightarrow \mathbb{R},\ f(x) = \frac{1}{3}x_1^2 + 3x_2^2,\ x = [x_1\ x_2]^T.$$

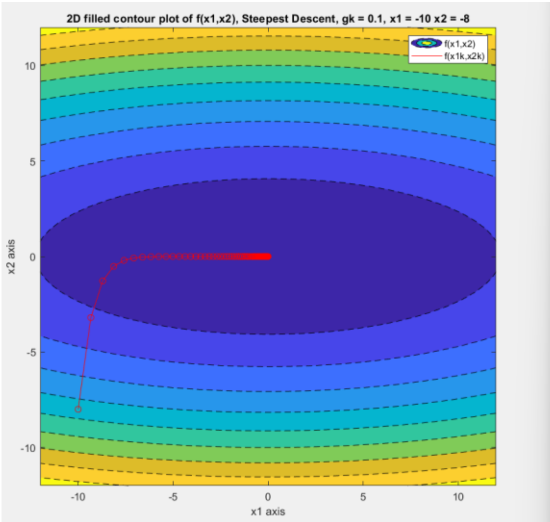
Για την απεικόνιση της γραφικής παράστασης της συνάρτησης γράφτηκε ξεχωριστό script στο MATLAB. Παρακάτω βλέπουμε την τρισδιάστατη απεικόνιση της συνάρτησης αλλά και το διάγραμμα του επιπέδου x1-x2.

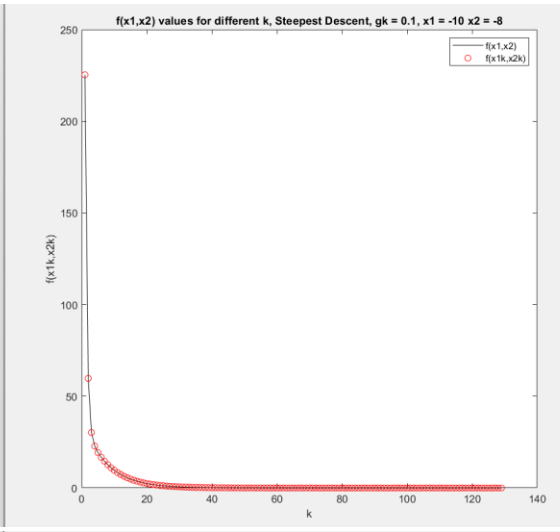


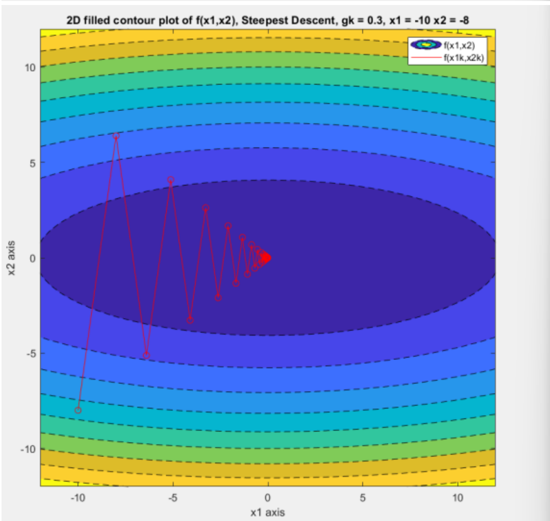


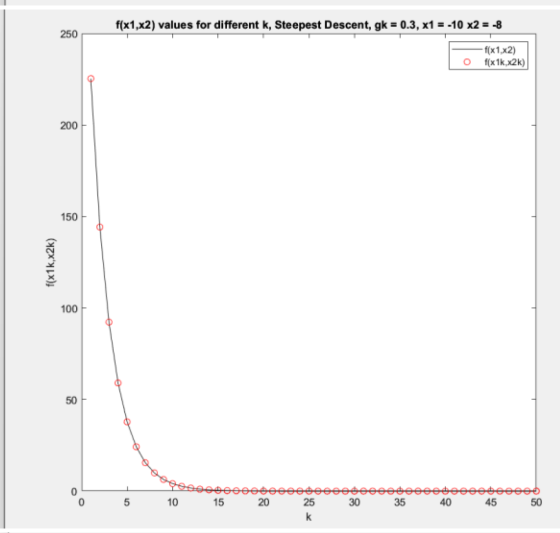
Μέγιστη Κάθοδος χωρίς περιορισμούς

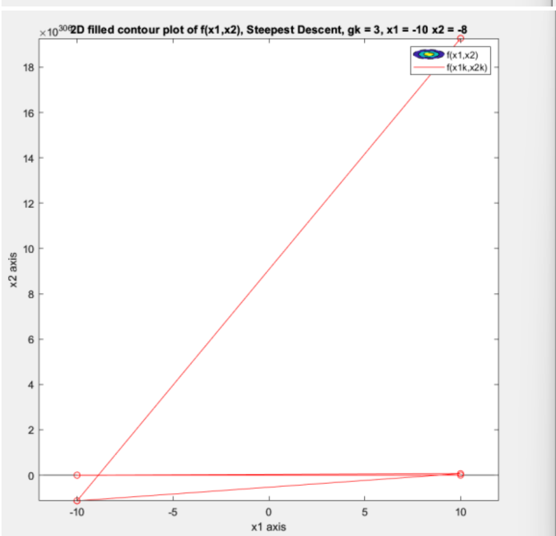
Με ακρίβεια ε = 0.001, βήμα i) γ_κ = 0.1, ii) γ_κ = 0.3, iii) γ_κ = 3, iv) γ_κ = 5 και σημείο εκκίνησης το (-10,-8).

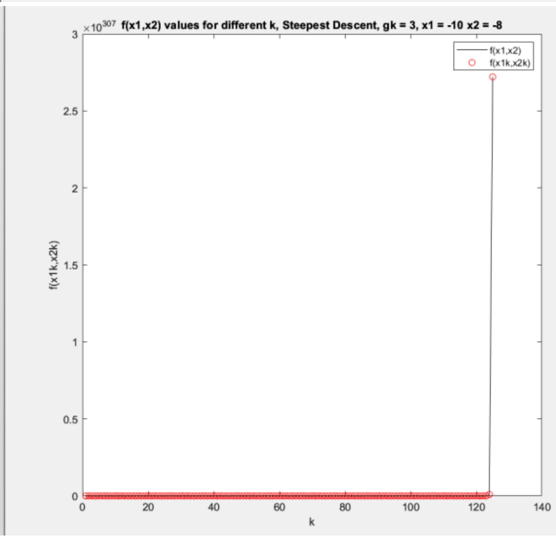


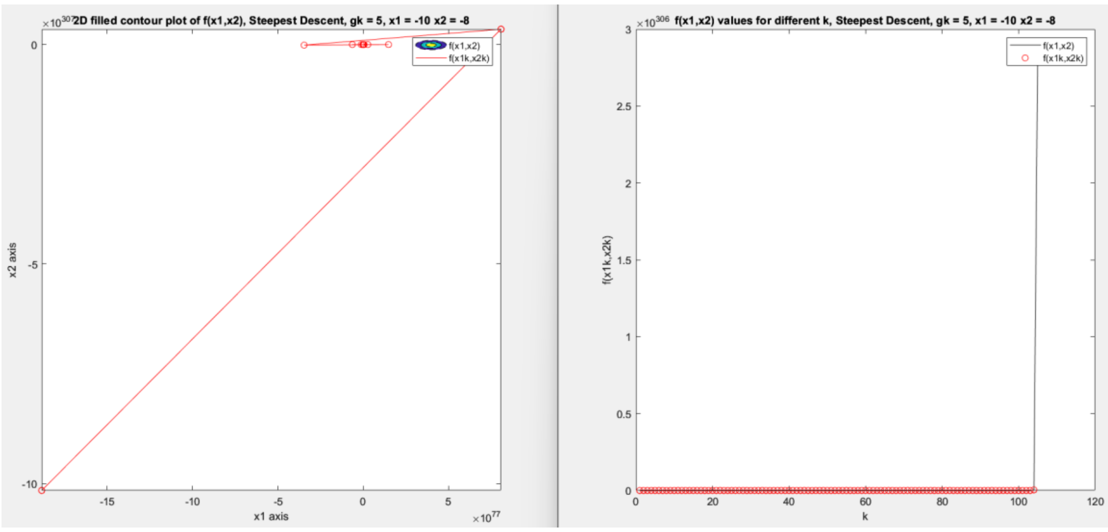












Σχολιασμός των αποτελεσμάτων:

Παρατηρούμε ότι ο αλγόριθμος για τις δύο τελευταίες περιπτώσεις βήματος γ_κ = 3 και γ_κ = 5 δεν συγκλίνει. Για να καταλάβουμε γιατί, πραγματοποιούμε την εξής μαθηματική ανάλυση:

$$f(x_1, x_2) = \frac{1}{3} * {x_1}^2 + 3 * {x_2}^2$$

$$\nabla f(x_1, x_2) = \begin{bmatrix} \frac{2}{3} * x_1 \\ 6 * x_2 \end{bmatrix}$$

$$X_{k+1} = X_k - \gamma_\kappa * \nabla f(x_{1k}, x_{2k}) = X_k - \begin{bmatrix} \frac{2}{3} * \gamma_\kappa * x_{1k} \\ 6 * \gamma_\kappa * x_{2k} \end{bmatrix} = \begin{bmatrix} 1 - \frac{2}{3} * \gamma_\kappa \\ 1 - 6 * \gamma_\kappa \end{bmatrix} * X_k$$

Για να συγκλίνει ο αλγόριθμος πρέπει:

$$\left|1 - \frac{2}{3} * \gamma_\kappa\right| < 1 \Rightarrow 0 < \gamma_\kappa < 3$$

$$|1 - 6 * \gamma_\kappa| < 1 \Rightarrow 0 < \gamma_\kappa < \frac{1}{3}$$

Τελικά $0 < \gamma_\kappa < \frac{1}{3} = 0.333$.

Για αυτό λοιπόν στις περιπτώσεις i) γ_κ = 0.1 και ii) γ_κ = 0.3 όπου γ_κ < 0.333 ο αλγόριθμος φτάνει στο ελάχιστο ενώ στις iii) γ_κ = 3 και iv) γ_κ = 5 όπου γ_κ > 0.333 η μέθοδος δεν συγκλίνει.

Μέγιστη Κάθοδος με προβολή

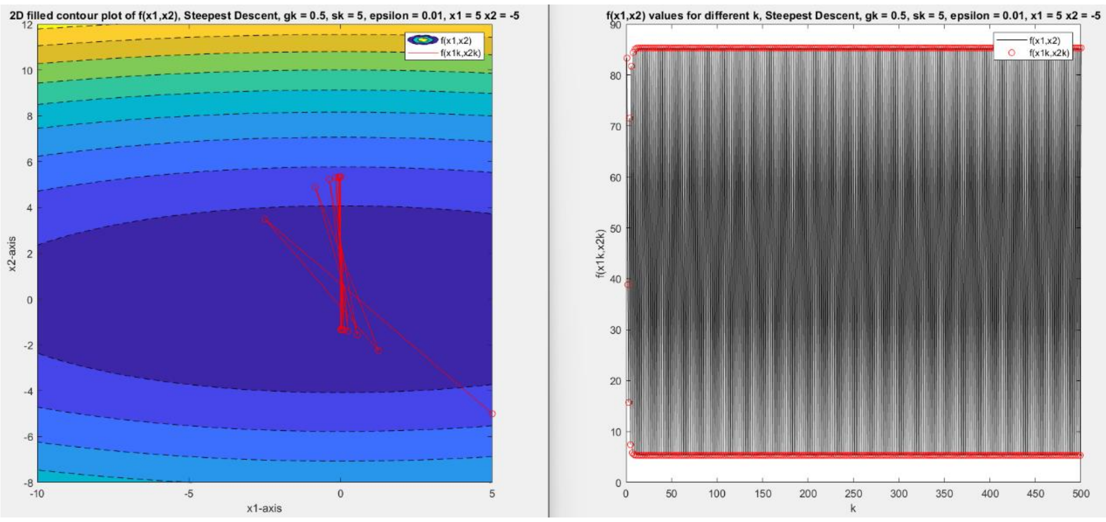
Εξετάζουμε τρεις περιπτώσεις:

1. $s_κ$ = 5, $\gamma_κ$,= 0.5, σημείο εκκίνησης το (5, −5) και ακρίβεια ε = 0.01
2. $s_κ$ = 15, $\gamma_κ$,= 0.1, σημείο εκκίνησης το (-5, 10) και ακρίβεια ε = 0.01
3. $s_κ$ = 0.1, $\gamma_κ$,= 0.2, σημείο εκκίνησης το (8, −10) και ακρίβεια ε = 0.01

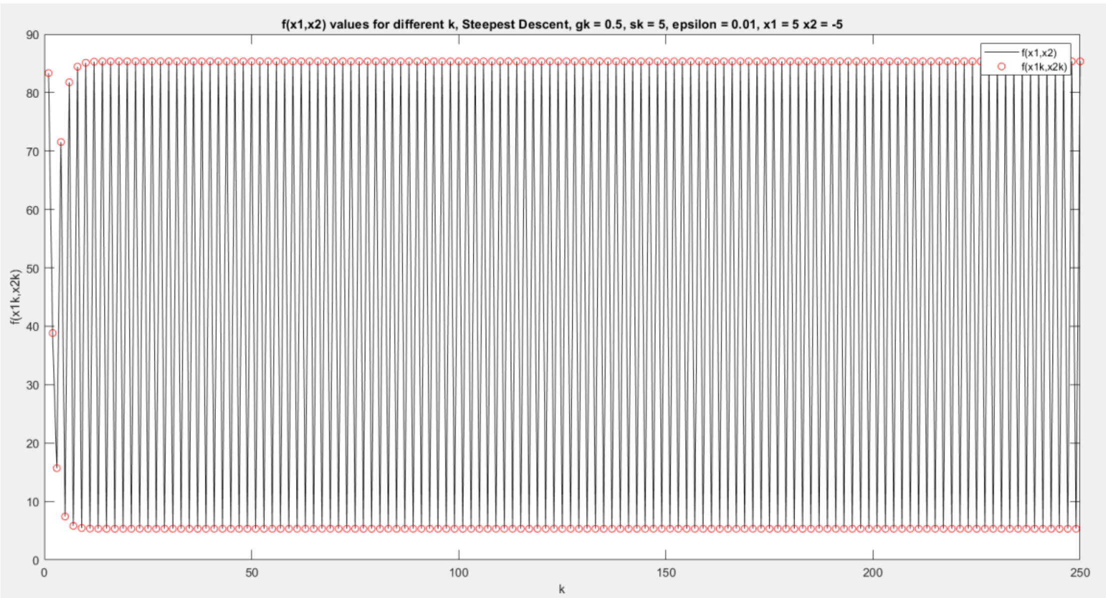
Θεωρούμε επίσης τον περιορισμό:

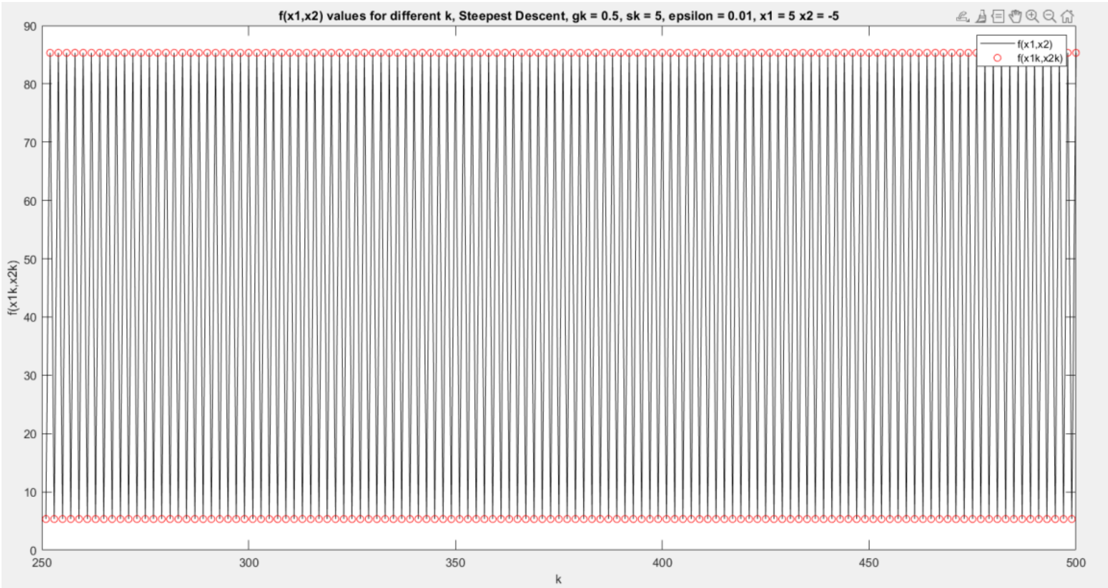
$$-10 \le X_1 \le 5 \text{ και } -8 \le X_2 \le 12$$

1η περίπτωση

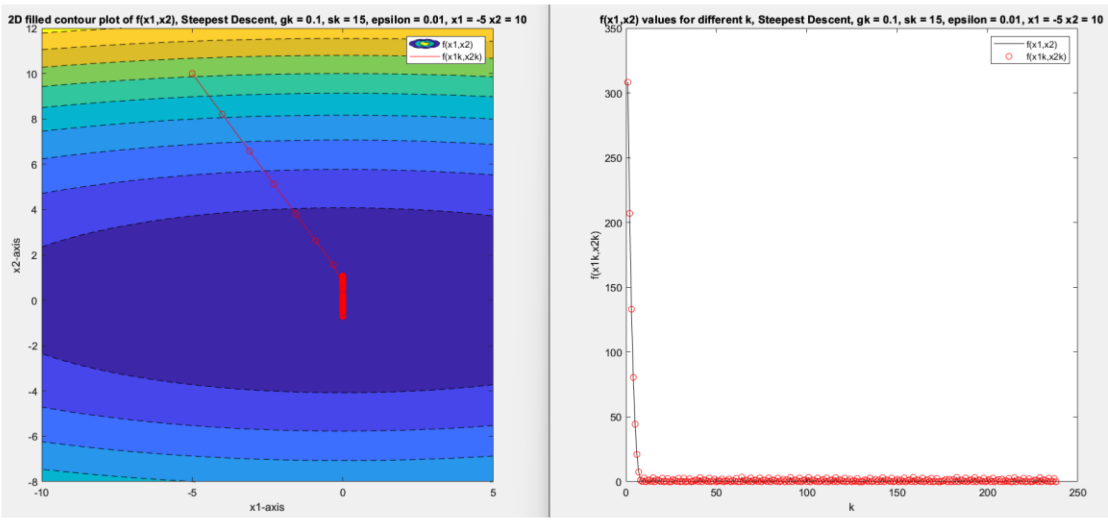


Το δεύτερο διάγραμμα της πάνω εικόνας «σπάει» σε δύο άλλα, για $0 \le κ \le 250$ και $250 \le κ \le 500$ ώστε να δούμε καλύτερα ότι η μέθοδος σε αυτήν την περίπτωση ταλαντώνεται μεταξύ δύο σημείων και δεν συγκλίνει στο ελάχιστο.

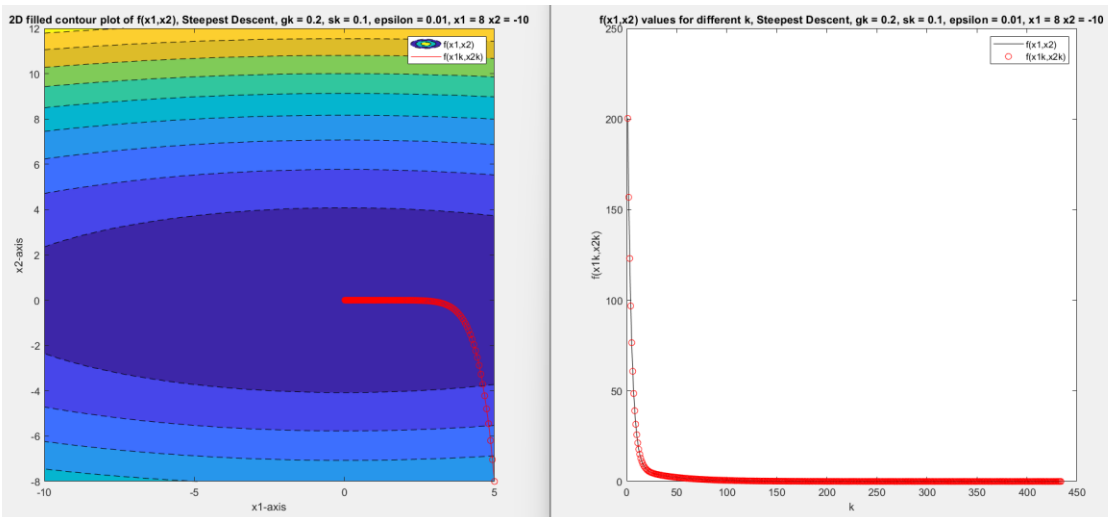




2η περίπτωση



3η περίπτωση

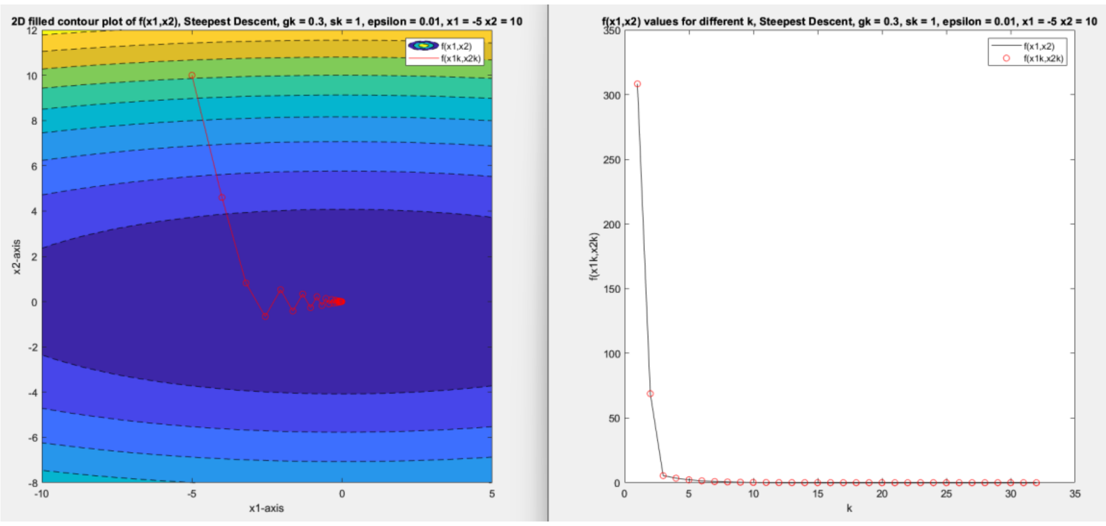


Σχολιασμός των αποτελεσμάτων:

Στην πρώτη περίπτωση η μέθοδος λειτουργεί εντός των περιορισμών και άρα δεν χρησιμοποιείται κάποια προβολή και ο αλγόριθμος τρέχει σαν την Μέθοδο Μέγιστης Καθόδου χωρίς περιορισμούς. Έτσι ισχύει και εδώ ο περιορισμός ότι $0 < \gamma_\kappa < 0.333$, με την μόνη διαφορά ότι έχουμε και το s_κ το οποίο πολλαπλασιάζεται με το γ_κ .Άρα τελικά, $0 < s_k * \gamma_\kappa < 0.333$. Επειδή χρησιμοποιούμε s_κ = 5 και $\gamma_\kappa = 0{,}5$, η μέθοδος δεν συγκλίνει ποτέ σε αυτήν την περίπτωση, αφού $s_k * \gamma_\kappa = 2.5 > 0.333$, και ταλαντώνεται μεταξύ δύο σημείων, όπως φαίνεται στο αντίστοιχο διάγραμμα παραπάνω.

Στην δεύτερη περίπτωση, ο αλγόριθμος συγκλίνει, παρόλο που χρησιμοποιούμε s_κ = 15 και $\gamma_\kappa = 0.1$ που μας δίνει $s_k * \gamma_\kappa = 1.5 > 0.333$. Αυτό ίσως συμβαίνει, διότι ικανοποιείται μία εκ των δύο ανισότητες ($0 < s_k * \gamma_\kappa < 3$). Παρατηρούμε βέβαια ότι η μέθοδος στην περίπτωση αυτή συγκλίνει πολύ αργά. Εάν χρησιμοποιηθεί συνδυασμός s_k, γ_κ ώστε $s_k * \gamma_\kappa = 0.3$ (π.χ. s_κ = 1 και $\gamma_\kappa = 0.3$) τότε η μέθοδος συγκλίνει γρήγορα.

Δοκιμή στο MATLAB:



Στην τρίτη περίπτωση, το σημείο εκκίνησης δεν είναι εφικτό και έτσι παίρνουμε την προβολή του. Πιο συγκεκριμένα η προβολή του σημείου (8,-10) είναι το (5,-8). Από εκεί και πέρα, αφού το ελάχιστο βρίσκεται εντός της περιορισμένης περιοχής, ο αλγόριθμος συγκλίνει στο ελάχιστο γιατί χρησιμοποιούμε s_κ = 0.1 και $\gamma_\kappa = 0.2$ και ισχύει ότι $s_k * \gamma_\kappa = 0.02 < 0.333$.