ΤΕΧΝΙΚΕΣ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ – ΕΡΓΑΣΙΑ 3

ΕΡΓΑΣΙΑ 3 - ΑΝΑΦΟΡΑ

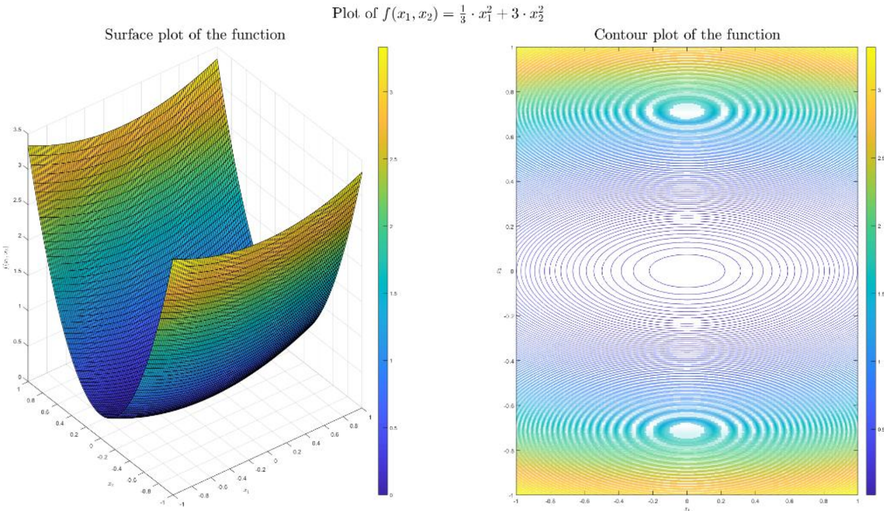
ΧΕΙΜΕΡΙΝΟ ΕΞΑΜΗΝΟ 2023-24

ΑΝΑΣΤΑΣΙΟΣ ΣΤΕΡΓΙΟΥ| 10079
ΤΗΜΜΥ ΑΠΘ | anasster@ece.auth.gr

ΕΙΣΑΓΩΓΗ

Το παρόν έγγραφο αποτελεί αναφορά για την Τρίτη εργασία του μαθήματος *Τεχνικές Βελτιστοποίησης* με θέμα «Μέθοδος Μέγιστης Καθόδου με Προβολή». Η συγκεκριμένη εργασία αφορά αρχικά την εφαρμογή του απλού αλγορίθμου Μέγιστης Καθόδου (Steepest Descent) σε μία αντικειμενική συνάρτηση του χώρου $\mathbb{R}^2$, και στη συνέχεια μελέτη του αλγορίθμου Μέγιστης καθόδου με προβολή (Steepest Descent with Projection), το οποίο αποτελεί και το κυρίως αντικείμενο της εργασίας.

Η αντικειμενική συνάρτηση που μελετάται είναι η $f(x) = \frac{1}{3} \cdot x_1^2 + 3 \cdot x_2^2, x = [x_1 \quad x_2]^T$, η οποία είναι μία τετραγωνική συνάρτηση της μορφής $f(x) = x^T \cdot Q \cdot x$ με $Q = \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & 1 \end{bmatrix} = Q^T > 0$, άρα εξ' ορισμού κυρτή. Η γραφική παράσταση της επιφάνειας της f, καθώς και ένα γράφημα με τις ισομετρικές καμπύλες της f παρουσιάζεται παρακάτω.



Εικόνα 1 - Γραφική παράσταση της f

Όπως φαίνεται και από το γράφημα λοιπόν, η f είναι πράγματι κυρτή, και μάλιστα λύνοντας αναλυτικά την εξίσωση $\nabla f(x) = 0$ καταλήγουμε εύκολα στη λύση $x^* = [0\ 0]^T$ με $f(x^*) = 0$, άρα η f έχει ελάχιστο το 0.

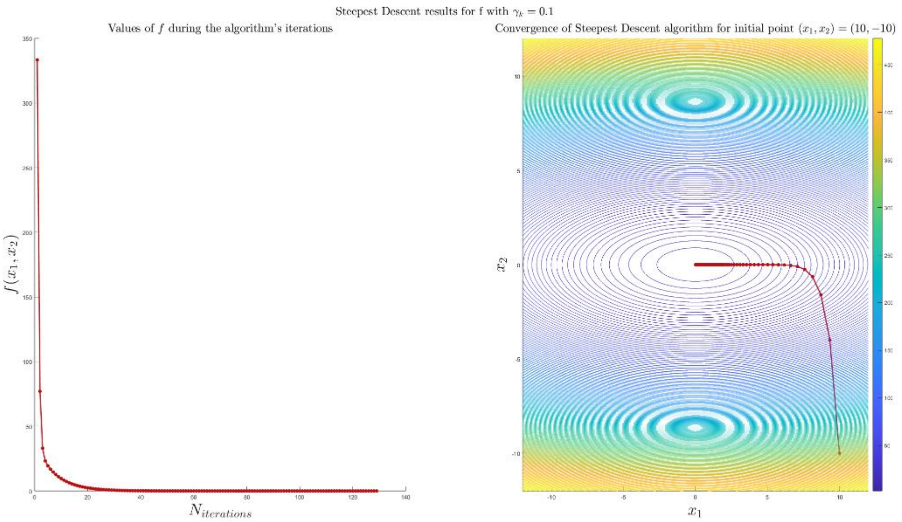
ΑΛΓΟΡΙΘΜΟΣ ΜΕΓΙΣΤΗΣ ΚΑΘΟΔΟΥ

Ως πρώτο ζήτημα στην εργασία αυτή είναι η δοκιμή του αλγορίθμου Μέγιστης Κλίσης για οποιοδήποτε αρχικό σημείο εκτός του $[0\ 0]^T$, για $\varepsilon = 0.001$ και για βήμα $\gamma_k \in \{0.1, 0.3, 3, 5\}$.

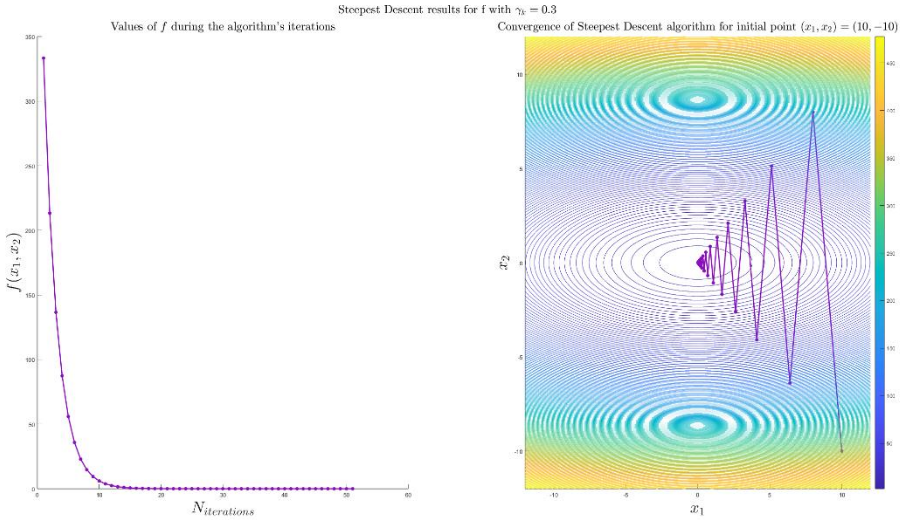
Αρχικά, θα γίνει μαθηματική θεμελίωση του προβλήματος. Για οποιαδήποτε σημεία x_k, x_{k+1}, θα πρέπει να ισχύει $f(x_{k+1}) < f(x_k)$, όπου k ο αριθμός της επανάληψης του αλγορίθμου. Στο συγκεκριμένο αλγόριθμο το επόμενο σημείο επιλέγεται ως $x_{k+1} = x_k - \gamma_k \cdot \nabla f(x_k)$. Είναι $\nabla f(x_k) = 2Qx_k$, άρα $x_{k+1} = (I - 2\gamma_k \cdot Q)x_k$, όπου I ο μοναδιαίος πίνακας. Θεωρώντας $B = I - 2\gamma_k \cdot Q = \begin{bmatrix} 1 - \frac{2}{3}\gamma_k & 0 \\ 0 & 1 - 6\gamma_k \end{bmatrix} = B^T$ και αντικαθιστώντας στην εξίσωση $f(x_{k+1}) < f(x_k)$, έχουμε $x^T BQBx < x^T Qx$. Φέρνοντας όλους τους όρους από τη μία πλευρά της ανίσωσης προκύπτει $x^T(Q - BQB)x > 0$, κάτι που μας οδηγεί στο συμπέρασμα ότι ο πίνακας $A = Q - BQB$ πρέπει να είναι *θετικά ορισμένος*. Υπολογίζεται εύκολα ότι $A = \begin{bmatrix} \frac{1}{3} \cdot \left[1 - \left(1 - \frac{2}{3}\gamma_k\right)^2\right] & 0 \\ 0 & 3 \cdot [1 - (1 - 6\gamma_k)^2] \end{bmatrix}$, και άρα ως διαγώνιος πίνακας, τα στοιχεία του ταυτίζονται με τις ιδιοτιμές του. Επομένως θα πρέπει:

$$\begin{cases} 1 - \left(1 - \frac{2}{3}\gamma_k\right)^2 > 0 \\ 1 - (1 - 6\gamma_k)^2 > 0 \end{cases} \Leftrightarrow \begin{cases} \left|1 - \frac{2}{3}\gamma_k\right| < 1 \\ |1 - 6\gamma_k| < 1 \end{cases} \Leftrightarrow \begin{cases} 0 < \gamma_k < 3 \\ 0 < \gamma_k < \frac{1}{3} \end{cases} \Leftrightarrow 0 < \gamma_k < \frac{1}{3}$$

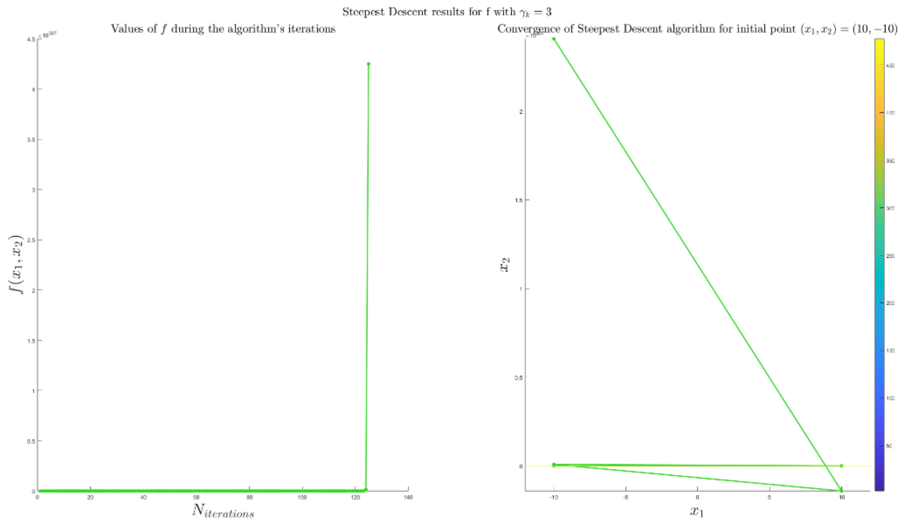
Πράγματι, τα σημεία γ_k που δοκιμάζονται για τον αλγόριθμο φαίνεται να πληρούν αυτόν τον περιορισμό.



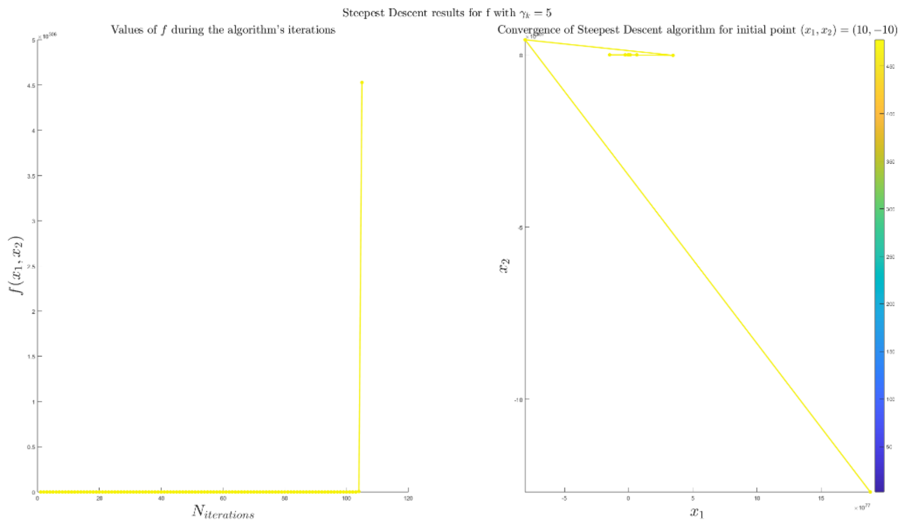
Εικόνα 2 - Αλγόριθμος Μέγιστης Καθόδου για $\gamma_k = 0.1$



Εικόνα 3 - Αλγόριθμος Μέγιστης Καθόδου για $\gamma_k = 0.3$



Εικόνα 4 - Αλγόριθμος Μέγιστης Καθόδου για $\gamma_k = 3$



Εικόνα 5 - Αλγόριθμος Μέγιστης Καθόδου για $\gamma_k = 5$

Όπως φαίνεται, για $\gamma_k = 0.1$ και $\gamma_k = 0.3$, ο αλγόριθμος συγκλίνει και τα γ_k βρίσκονται εντός των περιορισμών που υπολογίστηκαν. Αντιθέτως, για $\gamma_k = 3$ και $\gamma_k = 5$, δεν επιτυγχάνεται σύγκλιση, άρα και πάλι επαληθεύεται η μαθηματική θεμελίωση του προβλήματος.

ΑΛΓΟΡΙΘΜΟΣ ΜΕΓΙΣΤΗΣ ΚΑΘΟΔΟΥ ΜΕ ΠΡΟΒΟΛΗ

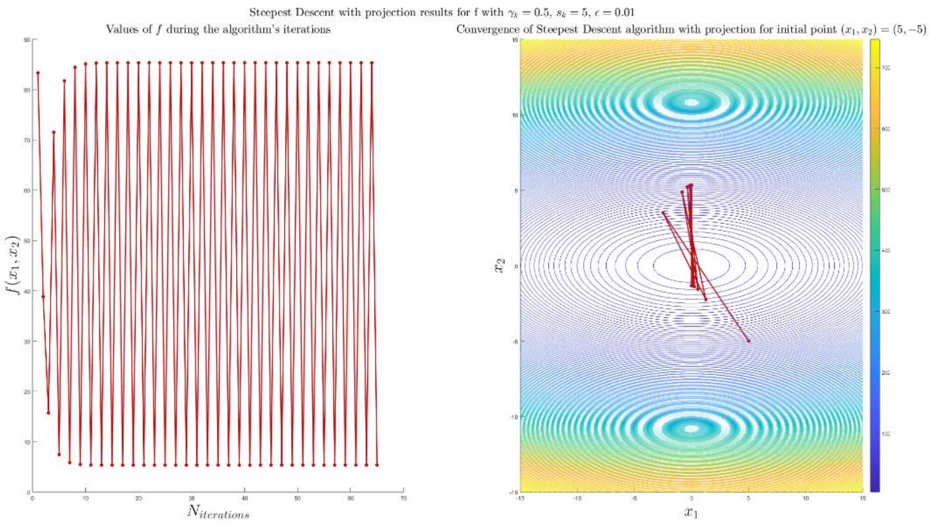
Εισάγονται τώρα οι εξής περιορισμοί για το πρόβλημα ελαχιστοποίησης:

$$-10 \leq x_1 \leq 5$$

$$-8 \leq x_2 \leq 12$$

Οι περιορισμοί αυτοί αποτελούν ένα ορθογώνιο παραλληλόγραμμο στο επίπεδο με κορυφές τα σημεία που προκύπτουν από τους συνδυασμούς των οριακών τιμών των x_1 και x_2.

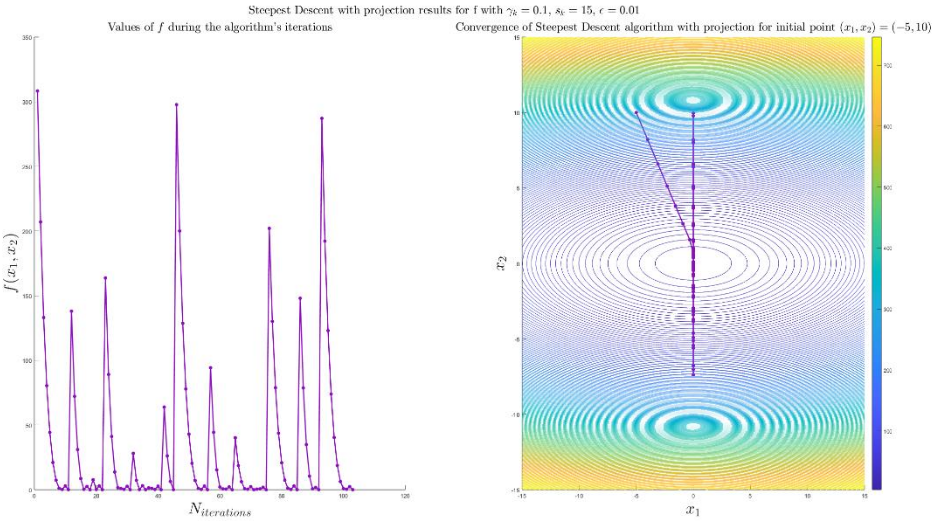
Αρχικά, ζητείται η δοκιμή του αλγορίθμου για αρχικό σημείο το $[5 - 5]^T$, και παραμέτρους $\gamma_k = 0.5$, $s_k = 5$ και $\varepsilon = 0.01$. Το αποτέλεσμα που προκύπτει είναι το εξής:



Εικόνα 6 – Αλγόριθμος Μέγιστης Καθόδου με προβολή για $\gamma_k = 0.5$, $s_k = 5$ και $\varepsilon = 0.01$

Όπως φαίνεται υπάρχει ένα φαινόμενο ταλάντωσης, όπου τα σημεία επιλογής διαρκώς ταλαντώνονται κατά τη διεύθυνση του x_2, και ποτέ δεν συγκλίνουν στο σημείο ελαχίστου που είναι το $[0\ 0]^T$. Στον αλγόριθμο προστέθηκε μια γραμμή ελέγχου ταλάντωσης, ώστε να διακόπτεται ο αλγόριθμος σε περίπτωση που πολλά διαδοχικά σημεία x_k έχουν ίσες μεταξύ τους ευκλείδειες αποστάσεις. Συγκριτικά με τον αλγόριθμο Μέγιστης καθόδου χωρίς περιορισμούς, φαίνεται ότι το s_k είναι αυτό που δημιουργεί το πρόβλημα, καθώς κάθε σημείο το στέλνει εκτός της εφικτής περιοχής X που ορίζεται από το παραλληλόγραμμο.

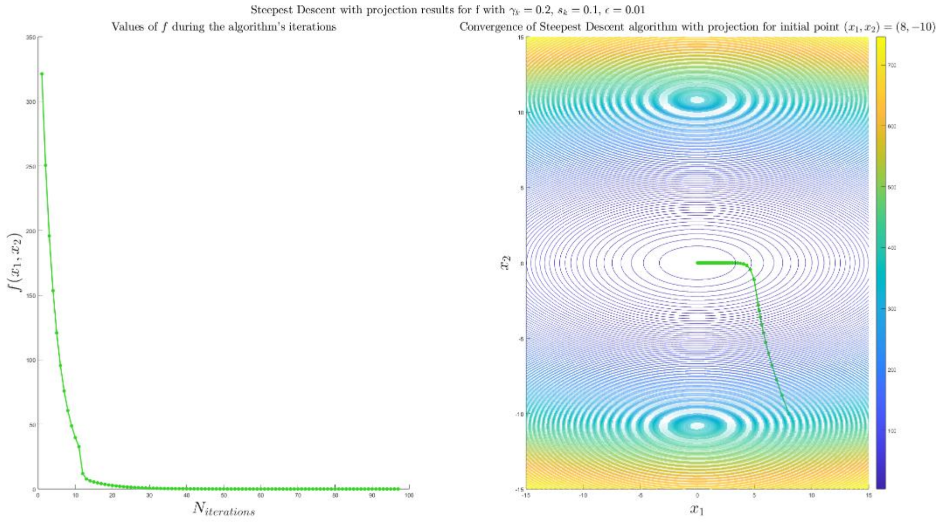
Στη συνέχεια, ζητήθηκε η δοκιμή του αλγορίθμου για αρχικό σημείο το $[-5\ 10]^T$, και παραμέτρους $\gamma_k = 0.1$, $s_k = 15$ και $\varepsilon = 0.01$. Το αποτέλεσμα που προκύπτει είναι το εξής:



Εικόνα 7 - Αλγόριθμος Μέγιστης Καθόδου με προβολή για $\gamma_k = 0.1$, $s_k = 15$ και $\varepsilon = 0.01$

Όπως φαίνεται και εδώ πέρα υπάρχει φαινόμενο ταλάντωσης, το οποίο είναι ξεκάθαρο ότι οφείλεται στην υπερβολικά μεγάλη τιμή του s_k. Ένας πρακτικός τρόπος να υπάρχει σύγκλιση στο ελάχιστο θα ήταν η επιλογή μικρότερου s_k, για παράδειγμα υποδεκαπλάσιο ($s_k = 1.5$). Στην περίπτωση αυτή, αφού το s_k θα βρίσκεται μεταξύ των περιορισμών που τέθηκαν κατά τον αλγόριθμο Μέγιστης Καθόδου για το γ_k, και το αρχικό σημείο x είναι εφικτό, θεωρητικά θα υπάρχει σύγκλιση για τον αλγόριθμο.

Τέλος, ζητείται η δοκιμή του αλγορίθμου για αρχικό σημείο το $[8 - 10]^T$, και αρχικές συνθήκες $\gamma_k = 0.2$, $s_k = 0.1$ και $\varepsilon = 0.01$. Μπορούν να εξαχθούν κάποια εκ των προτέρων συμπεράσματα για την εκτέλεση του αλγορίθμου αυτού. Αρχικά, το αρχικό σημείο είναι σίγουρα μη εφικτό, άρα θα γίνει η προβολή του στο εφικτό παραλληλόγραμμο. Επιπλέον, το μικρό γ_k εξασφαλίζει ότι θα μείνει το σημείο στο χώρο αυτό. Τώρα, το $s_k = 0.1$, από τη στιγμή που ο αλγόριθμος θα μπει μέσα στον εφικτό χώρο, σημαίνει ότι θα εκτελείται ο αλγόριθμος Μέγιστης Καθόδου χωρίς περιορισμούς, και με βήμα s_k εντός των περιορισμών που υπολογίστηκαν παραπάνω, άρα θα υπάρχει βέβαιη σύγκλιση. Πράγματι, αποδεικνύεται από τα αποτελέσματα:



Εικόνα 8 - Αλγόριθμος Μέγιστης Καθόδου με προβολή για $\gamma_k = 0.2$, $s_k = 0.1$ *και* $\varepsilon = 0.01$

Υπάρχει λοιπόν σύγκλιση του αλγορίθμου, άρα και πάλι οι μαθηματικοί υπολογισμοί επιβεβαιώνονται.

ΤΕΛΙΚΑ ΣΥΜΠΕΡΑΣΜΑΤΑ

Ο αλγόριθμος Μέγιστης Καθόδου είναι ένας αλγόριθμος ο οποίος εξασφαλίζει καλή και γρήγορη σχετικά σύγκλιση, εφόσον υπάρχει ένα σχετικά μικρό βήμα γ_k, ώστε να ικανοποιείται η θεμελιώδης αρχή των αλγορίθμων κλίσης:

$$f(x_{k+1}) < f(x_k)$$

Η επιλογή μεγάλων τιμών βήματος σε ένα πρόβλημα χωρίς περιορισμούς μπορεί να οδηγήσει σε απόκλιση του αλγορίθμου.

Όταν εισάγονται περιορισμοί, είναι πάλι σημαντική η επιλογή του s_k, καθώς ένα πολύ μεγάλο βήμα θα στέλνει τον αλγόριθμο εκτός του εφικτού χώρου των περιορισμών, και θα προκαλείται ταλάντωση μεταξύ των σημείων προβολής και έτσι πάλι δεν θα υπάρχει σύγκλιση ποτέ.

ΤΕΛΟΣ ΑΝΑΦΟΡΑΣ