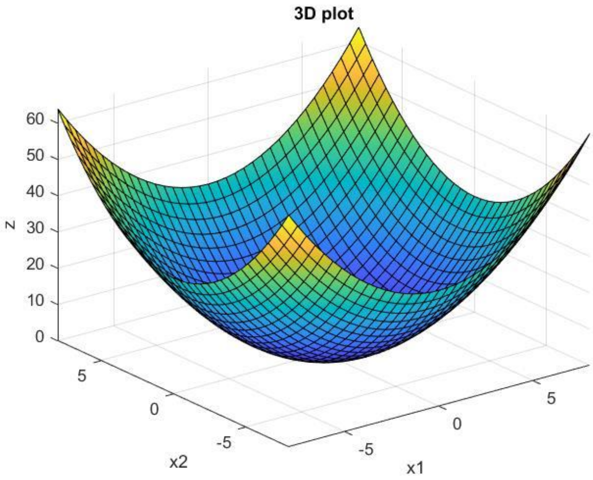
ΕΡΓΑΣΙΑ 3η
Μέθοδος Μέγιστης Καθόδου με Προβολή

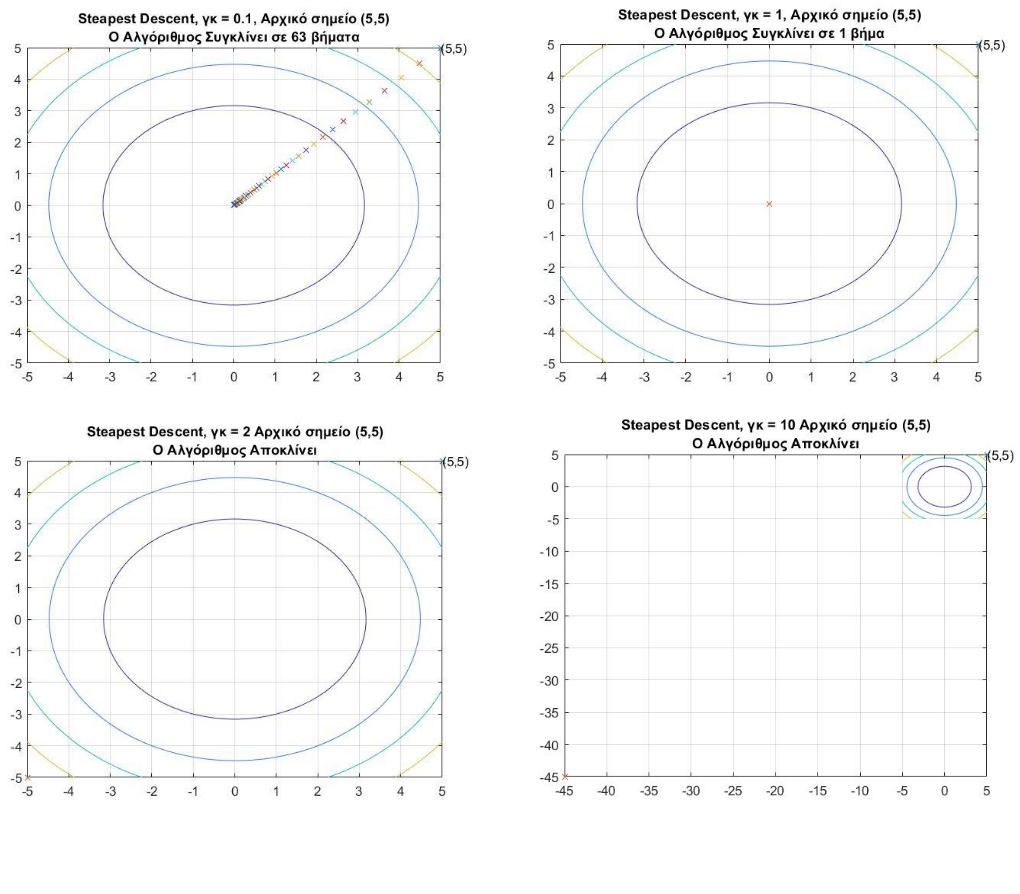
Φόρογλου Γεώργιος-Βησσαρίων, ΑΕΜ: 9557

$f(x) = x_1^2/2 + x_2^2/2$



ΘΈΜΑ 1Ο

Μέθοδος Μέγιστης Καθόδου



Αρχικά ελαχιστοποιούμε την f με χρήση της μεθόδου της Μέγιστης Καθόδου με αρχικό σημείο το (5,5), το $γ_κ$ να παίρνει τιμές 0.1, 1, 2 και 10 (βλ. αντίστοιχα διαγράμματα) και ακρίβεια ε = 0.01. Παρατηρούμε ότι για $γ_κ = 0.1$ ο αλγόριθμος συγκλίνει στο ελάχιστο μετά από 63 βήματα. Αντίθετα, για $γ_κ = 2$ και $γ_κ = 10$ ο αλγόριθμος αποκλίνει και τείνει στο άπειρο. Ο αλγόριθμός μου προγραμματίστηκε έτσι ώστε να σταματάει στην αμέσως επόμενη επανάληψη από την στιγμή πού θα περάσει το ελάχιστο της f καθώς σε αντίθετη περίπτωση δεν θα τερμάτιζε ποτέ. Βλέπουμε ότι για $γ_κ = 10$ ο αλγόριθμος αποκλίνει με πολύ πιο γρήγορο ρυθμό από ότι με $γ_κ = 2$. Τέλος, σημαντική είναι παρατήρηση ότι με $γ_κ = 1$, ο αλγόριθμος συγκλίνει στο ελάχιστο σε μόλις ένα βήμα, κάτι που μας επιδεικνύει ότι αυτή είναι και η βέλτιστη επιλογή για το $γ_κ$.

Αυτές οι παρατηρήσεις μπορούν να αποδειχθούν και μαθηματικά. Όπως βλέπουμε παρακάτω ο αλγόριθμος συγκλίνει για κάθε $0 < γ_κ < 2$ και η βέλτιστη επιλογή του $γ_κ$ είναι το 1.

$f: \mathbb{R}^2 \to \mathbb{R}, \quad f(x) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2$

Φλόγκα Γεώργιος-Βικτώριος
9557

Μέγιστη Κάθοδος: $x_{k+1} = x_k - γ_k \nabla f(x_k)$

Υπολογισμός $\nabla f(x)$: $\nabla f(x) = [x_1 \;\; x_2]^T$

Επομένως ορίζουμε $\nabla f_1(x) = x_1$, $\nabla f_2(x) = x_2$ τ.ω. $\nabla f(x) = [\nabla f_1(x) \;\; \nabla f_2(x)]^T$

<u>Έτσι</u> $x_{k+1} = x_k - γ_k \nabla f(x_k)$, όπου $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

<u>Επομένως</u>: $\begin{bmatrix} x_{1,k+1} \\ x_{2,k+1} \end{bmatrix} = \begin{bmatrix} x_{1,k} \\ x_{2,k} \end{bmatrix} - γ_k \begin{bmatrix} \nabla f_1(x_{1,k}) \\ \nabla f_2(x_{2,k}) \end{bmatrix}$

$$\Rightarrow \begin{cases} x_{1,k+1} = x_{1,k} - γ_k x_{1,k} \\ x_{2,k+1} = x_{2,k} - γ_k x_{2,k} \end{cases} \Rightarrow \begin{cases} x_{1,k+1} = (1-γ_k)x_{1,k} \quad (1) \\ x_{2,k+1} = (1-γ_k)x_{2,k} \quad (2) \end{cases}$$

Για να συγκλίνει ο αλγόριθμος θα πρέπει $\left|\frac{x_{k+1}}{x_k}\right| < 1$. Επομένως εδώ:

$$\begin{cases} \left|\frac{x_{1,k+1}}{x_{1,k}}\right| < 1 \\ \left|\frac{x_{2,k+1}}{x_{2,k}}\right| < 1 \end{cases} \Rightarrow \begin{cases} |1-γ_k| < 1 \\ |1-γ_k| < 1 \end{cases}$$

<u>Τελικά</u>: $0 < γ_k < 2$, δηλ. $\boxed{γ_k < 2}$. Ο αλγόριθμος συγκλίνει για $γ_k < 2$ και για $γ_k \geq 2$ αποκλίνει από το ελάχιστο.

Φίρογλου Γεώργιος Βησσαρίων
9557

Εύρεση βέλτιστης επιλογής γ_k :

$$\frac{df(x_{1,k+1}, x_{2,k+1})}{d\gamma} = 0 \quad (3)$$

Υπολογίζω: $f(x_{1,k+1}, x_{2,k+1}) = \frac{1}{2} x_{1,k+1}^2 + \frac{1}{2} x_{2,k+1}^2 =$

$$\overset{(1)}{\underset{(2)}{=}} \frac{1}{2}(1-\gamma_k)^2 x_{1,k}^2 + \frac{1}{2}(1-\gamma_k)^2 x_{2,k}^2 =$$

$$= \frac{1}{2}(1-\gamma_k)^2 (x_{1,k}^2 + x_{2,k}^2)$$

και: $\frac{df(x_{1,k+1}, x_{2,k+1})}{d\gamma} = -(1-\gamma_k)(x_{1,k}^2 + x_{2,k}^2)$

Από (3), $\frac{df(x_{1,k+1}, x_{2,k+1})}{d\gamma} = 0 \Rightarrow -(1-\gamma_k)(x_{1,k}^2 + x_{2,k}^2) = 0$

$\Rightarrow \boxed{\gamma_k = 1}$, Απορρίπτεται την τετριμμένη λύση $\begin{vmatrix} x_{1,k}=0 \\ x_{2,k}=0 \end{vmatrix}$

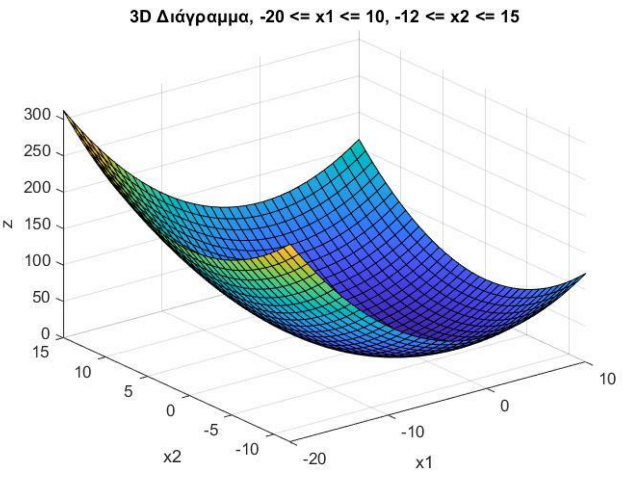
Επαλήθευση: $\begin{cases} x_{1,k+1} = (1-\gamma_k)x_{1,k} \\ x_{2,k+1} = (1-\gamma_k)x_{2,k} \end{cases} \xRightarrow{\gamma_k=1} \begin{vmatrix} x_{1,k+1}=0 \\ x_{2,k+1}=0 \end{vmatrix}$

Δηλαδή; $x_{k+1} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Δηλαδή ο αλγόριθμος για $\gamma_k=1$ συγκλίνει σε ένα βήμα

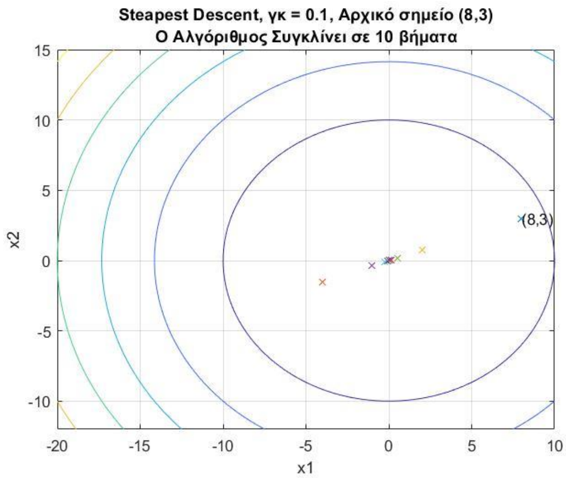
ΘΈΜΑ 2^{O}

Μέθοδος Μέγιστης Καθόδου με χρήση Προβολής

Αρχικά, μπορούμε να δούμε την γραφική παράσταση της f εάν βάλουμε τους περιορισμούς που μας δόθηκαν ($-20 <= x_1 <= 10$ και $-12 <= x_2 <= 15$).



Σε αυτό το θέμα θα ελαχιστοποιηθεί η f με χρήση της μεθόδου της Μέγιστης Καθόδου με χρήση Προβολής. Έχουμε ακρίβεια ε = 0.01, σημείο εκκίνησης του αλγορίθμου το (8,3), $s_κ$ = 15 και $γ_κ$ = 0.1. Ξέρουμε ότι το βήμα αναζήτησης αυτού του αλγορίθμου ορίζεται από το γινόμενο $s_κ$ $γ_κ$ = 1.5. Βλέπουμε ότι το βήμα αναζήτησης είναι μικρότερο του 2 (βλ. θέμα 1), επομένως περιμένουμε ο αλγόριθμος να συγκλίνει στο ελάχιστο.



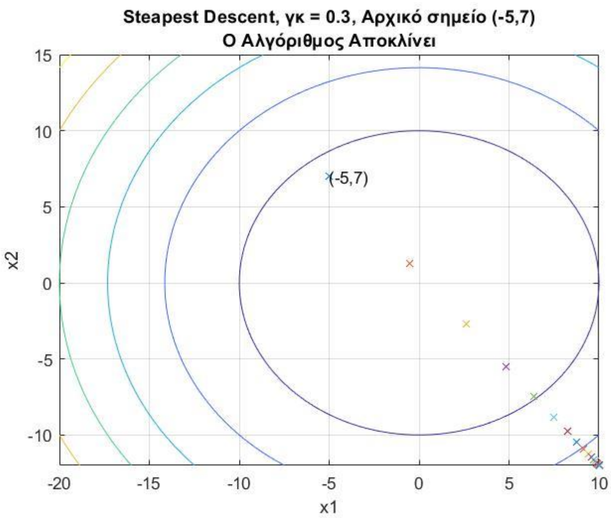
Πράγματι ο αλγόριθμος συγκλίνει σε 10 βήματα στο ελάχιστο. Παρατηρούμε ότι σε σχέση με το θέμα 1(iv) ο αλγόριθμος ο αλγόριθμος συγκλίνει στο ελάχιστο, όπως περιμέναμε, καθώς το βήμα αναζήτησης του είναι μικρότερο του 2. Επιπλέον, βλέπουμε ότι και στο θέμα 1(i), που $γ_κ$ = 0.1 < 2, υπήρχε σύγκλιση στο ελάχιστο, και εδώ που $γ_κ s_κ$ = 1.5 < 2 υπάρχει σύγκλιση. Παρόλα αυτά επειδή το 1.5 είναι πιο κοντά στο ιδανικό βήμα αναζήτησης $γ_κ$ = 1 βλέπουμε ότι αυτός ο αλγόριθμος συγκλίνει πολύ πιο γρήγορα σε σχέση με αυτόν του θέματος 1(i) (10 επαναλήψεις στο θέμα 2, 63 στο θέμα 1(i)).

Όλες οι παραπάνω παρατηρήσεις είναι αναμενόμενες βλέποντας το βήμα αναζήτησης του αλγορίθμου και συγκρίνοντας το με αυτό του θέματος 1.

ΘΈΜΑ 3^{O}

Μέθοδος Μέγιστης Καθόδου με χρήση Προβολής

Σε αυτό το θέμα θα ελαχιστοποιηθεί η f με χρήση της μεθόδου της Μέγιστης Καθόδου με χρήση Προβολής. Έχουμε ακρίβεια ε = 0.02, σημείο εκκίνησης του αλγορίθμου το (-5,7), $s_κ$ = 20 και $γ_κ$ = 0.3. Βλέπουμε ότι $s_κ$ $γ_κ$ = 6, άρα το βήμα αναζήτησης είναι μεγαλύτερο του 2 (βλ. θέμα 1), επομένως περιμένουμε ο αλγόριθμος να αποκλίνει.



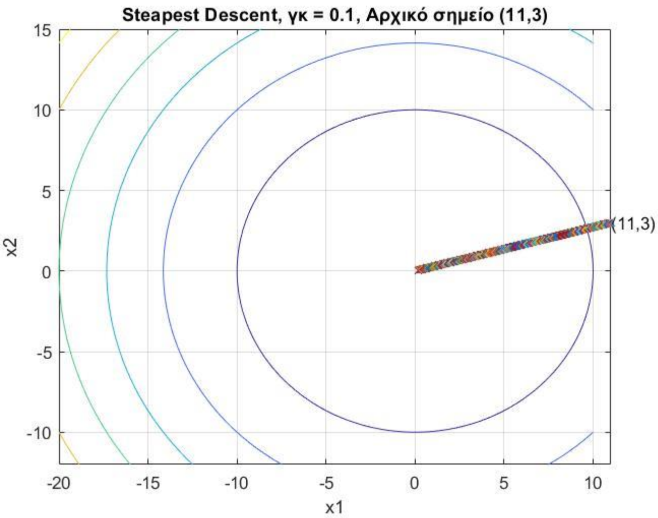
Πράγματι ο αλγόριθμος αποκλίνει. Βέβαια, λόγω των περιορισμών δεν τείνει στο άπειρο αλλά εγκλωβίζεται στο σημείο (10,-12) που είναι και τα όρια για τα x_1 και x_2, αντίστοιχα. Οι βασικές διαφορές με τον αλγόριθμο του πρώτου θέματος είναι 2. Πρώτον στο θέμα 1(i) υπήρχε σύγκλιση λόγω του $γ_κ$ = 0.1 < 2, ενώ εδώ δεν υπάρχει αφού το βήμα αναζήτησης είναι μεγαλύτερο του 2. Δεύτερον στο θέμα 1(iv) ο αλγόριθμος έτεινε στο άπειρο, κάτι που όπως είδαμε εδώ δεν εμφανίζεται λόγω της ύπαρξης των περιορισμών για τα x_1 και x_2.

Ένας πρακτικός τρόπος έτσι ώστε να συγκλίνει η μέθοδος στο ελάχιστο θα ήταν να μειώσουμε το βήμα αναζήτησης έως ότου να γίνει μικρότερο του 2. Όπως είδαμε προηγουμένως, όσο πιο κοντά στο βέλτιστο βήμα 1 καταφέρουμε να το φτάσουμε τόσο γρηγορότερα θα συγκλίνει και ο αλγόριθμος. Για να μειώσουμε το βήμα αναζήτησης αρκεί να μειώσουμε έναν από τους όρους του γινομένου $γ_κ s_κ$.

ΘΈΜΑ 4^{O}

Μέθοδος Μέγιστης Καθόδου με χρήση Προβολής

Σε αυτό το θέμα θα ελαχιστοποιηθεί η f με χρήση της μεθόδου της Μέγιστης Καθόδου με χρήση Προβολής. Έχουμε ακρίβεια ε = 0.01, σημείο εκκίνησης του αλγορίθμου το (11,3), $s_κ$ = 0.1 και $γ_κ$ = 0.01. Βλέπουμε ότι $s_κ$ $γ_κ$ = 0.001, άρα το βήμα αναζήτησης είναι μικρότερο του 2 αλλά πολύ κοντά στο 0 (βλ. θέμα 1), επομένως περιμένουμε ο αλγόριθμος να τείνει στο ελάχιστο αλλά να μην τερματίζει.



Πράγματι βλέπουμε ότι ο αλγόριθμος τείνει στο σημείο ισορροπίας αλλά ποτέ δεν φτάνει σε αυτό. Όπως αναφέρθηκε και προηγουμένως, αυτό είναι λογικό καθώς το βήμα αναζήτησης είναι οριακά εντός των ορίων επιλογής των $γ_κ$ για να συγκλίνει ο αλγόριθμος (βλ. Θέμα 1).

ΣΥΜΠΈΡΑΣΜΑ

Όπως είδαμε στις περιπτώσεις που εξετάσαμε, η σύγκλιση ή απόκλιση των δύο αλγορίθμων εξαρτάται εξ ολοκλήρου από την επιλογή του βήματος αναζήτησης και όχι από το αρχικό σημείο ούτε από την ακρίβεια που επιλέγουμε. Αποδείξαμε και μαθηματικά (βλ. Θέμα 1) ότι για βήμα αναζήτησης μεγαλύτερο του 2 οι αλγόριθμοι αποκλίνουν είτε στο άπειρο είτε στο όριο του χωρίου των περιορισμών. Επιπλέον, είδαμε ότι όσο πιο κοντά ήταν το βήμα αναζήτησης στο βέλτιστο 1 (βλ. Θέμα 2 και Θέμα 1), τόσο πιο γρήγορα σύγκλινε ο αλγόριθμος. Τέλος, στο Θέμα 4, είδαμε ότι για βήμα αναζήτησης οριακά μεγαλύτερο του 0, ο αλγόριθμος τείνει στο σημείο ισορροπίας αλλά δεν συγκλίνει σε αυτό καθώς δεν τερματίζει ποτέ.