ΤΕΧΝΙΚΕΣ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ

3η Εργασία

Παναγιώτης

Καρβουνάρης

ΑΕΜ 10193

Αντικειμενική συνάρτηση: f(x1, x2) = (1/3)*(x1^2) + 3*(x2^2)

Αρχικά, χρησιμοποιούμε την Μέθοδο Μέγιστης Καθόδου χωρίς περιορισμούς στις τιμές των x1 και x2 με ακρίβεια ε = 0.001 και με επιλογές βήματος γκ, i) γκ = 0.1 ii) γκ = 0.3 iii) γκ = 3 iv) γκ = 5. Παίρνουμε τρία τυχαία αρχικά σημεία της επιλογής μας (-5, 5), (-5, 10) και (8, -10).

i) Για γκ = 0.1 παρατηρούμε ότι και τα τρία σημεία φτάνουν με επιτυχία στο (0, 0) που αποτελεί και το ολικό ελάχιστο της συνάρτησης. Όμως, τα σημεία (-5, 5) και (-5, 10) έφτασαν στο σωστό αποτέλεσμα σε 118 επαναλήψεις ενώ το (8, -10) σε 125.

ii) Για γκ = 0.3 παρατηρούμε πάλι ότι και τα τρία σημεία φτάνουν με επιτυχία στο (0, 0) που αποτελεί και το ολικό ελάχιστο της συνάρτησης. Όμως, με την αλλαγή στην τιμή του γκ παρατηρείται μεγάλη βελτίωση στην ταχύτητα σύγκλισης του αλγορίθμου καθώς οι επαναλήψεις που χρειάστηκαν είναι 47, 50 και 50 αντίστοιχα για τα τρία σημεία.

iii) Για γκ = 3 παρατηρούμε ότι μετά από 251 επαναλήψεις και για τα τρία σημεία το MatLab μας έδωσε αποτελέσματα τύπου NaN που σημαίνει απροσδιοριστίες, άρα το συγκεκριμένο γκ είναι ακατάλληλο για την βελτιστοποίηση αυτής της συνάρτησης.

iv) Για γκ = 5 παρατηρούμε ότι όπως και στην προηγούμενη περίπτωση και για τα τρία σημεία το MatLab μας έδωσε αποτελέσματα τύπου NaN που σημαίνει απροσδιοριστίες αυτήν την φορά μετά από 212 επαναλήψεις.

Η μαθηματική απόδειξη έγινε στο χαρτί και παρακάτω υπάρχει σε μορφή χειρόγραφων σημειώσεων.

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

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

Η γ_κ πρέπει να ελαχιστοποιεί την $f(x_\kappa - \gamma_\kappa \nabla f(x_\kappa))$ $\gamma_\kappa \geq 0$

Για αρχικό σημείο $x_0 = (5,-5)$ γ_κ=σταθερό

$$f(x_0 - \gamma_\kappa \nabla f(x_0)) = f\left(\begin{bmatrix} 5 \\ -5 \end{bmatrix} - \gamma_\kappa \begin{bmatrix} \frac{10}{3} \\ -30 \end{bmatrix} \right) =$$

$$= f\left(\begin{bmatrix} 5 - \gamma_\kappa \frac{10}{3} \\ -5 + \gamma_\kappa 30 \end{bmatrix} \right) = f\left(5 - \gamma_\kappa \frac{10}{3}, -5 + 30\gamma_\kappa\right) =$$

$$= \frac{1}{3} \left(5 - \gamma_\kappa \frac{10}{3}\right)^2 + 3(-5 + 30\gamma_\kappa)^2 =$$

$$= \frac{1}{3} \left(25 - \frac{100}{3}\gamma_\kappa + \gamma_\kappa^2 \frac{100}{9}\right) + 3(25 - 300\gamma_\kappa + 900\gamma_\kappa^2) =$$

$$= \overset{8,3}{\frac{25}{3}} - \overset{11,1}{\frac{100}{9}} \gamma_\kappa + \overset{3,7}{\frac{100}{27}} \gamma_\kappa^2 + 75 - 900\gamma_\kappa + 2700\gamma_\kappa^2 =$$

$$= 2703,7\gamma_\kappa^2 - 911,1\gamma_\kappa + 83,3$$

$$D = 528.990,37$$

$$\gamma_\kappa = \frac{911,1 \pm 727,3}{1.807,4} = \begin{cases} \approx 0,9 \\ \approx 0,1 \end{cases}$$

Για $\gamma_\kappa = 0,5$ $f\left(5 - \frac{10}{6}, -5 + 15\right) = f\left(\frac{10}{3}, +10\right) = -146,325$

Άρα ξέρουμε για ποιες τιμές του γ_κ η f είναι θετική, αρνητική και μηδέν

$$f\left(5-\gamma\kappa\frac{10}{3},\ -5+30\gamma\kappa\right) = 903{,}7\gamma\kappa^2 - 911{,}1\gamma\kappa + 83{,}3$$

$$\frac{df}{d\gamma\kappa} = 1807{,}4\gamma\kappa - 911{,}1$$

$$\frac{df}{d\gamma\kappa} = 0 \Rightarrow \gamma\kappa = \frac{911{,}1}{1807{,}4} \approx 0{,}5$$

Για $\gamma\kappa = 0{,}5$ έχουμε τοπικό ακρότατο

$$\left.\begin{array}{l} \text{Για } \gamma\kappa < 0{,}5 \quad \frac{df}{d\gamma\kappa} < 0 \\ \text{Για } \gamma\kappa > 0{,}5 \quad \frac{df}{d\gamma\kappa} > 0 \end{array}\right\} \Rightarrow \text{Άρα έχουμε ολικό ελάχιστο}$$

Για οποιοδήποτε $\gamma\kappa > 0{,}5 \Rightarrow \frac{df}{d\gamma\kappa} > 0$, άρα f αύξουσα, άρα δεν θα πετύχει ο αλγόριθμος.

Για $\gamma\kappa < 0{,}5 \Rightarrow \frac{df}{d\gamma\kappa} < 0$, άρα f φθίνουσα, άρα θα πετύχει ο αλγόριθμος αλλά θα χρειαστεί περισσότερα βήματα από ότι με $\gamma\kappa = 0{,}5$, πάντα όμως $\gamma\kappa \geq 0$. Γι' αυτό αν $\gamma\kappa = 3$ ή $\gamma\kappa = 5$ ο αλγόριθμος δεν βγάζει σωστά αποτελέσματα. Αντίστοιχα, μπορούν να βρεθούν όρια και για τα υπόλοιπα αρχικά σημεία.

Για την συνέχεια θα θεωρήσουμε ότι έχουμε κάποιους περιορισμούς στις τιμές που μπορούν να πάρουν τα σημεία που χρησιμοποιούνται στον αλγόριθμο: -10<=x1<=5 και -8<=x2<=12.

Ως συνθήκη τερματισμού του αλγορίθμου χρησιμοποιήθηκε η παρακάτω σχέση που προέκυψε από σχέσεις του βιβλίου και φαίνεται με την μορφή γραπτών σημειώσεων.

Έχουμε ότι: $\bar{x}_k = Pr_X\{x_k - s_k \nabla f(x_k)\}$, $s_k > 0$

Ξέρουμε ότι η συνθήκη τερματισμού είναι

$$\nabla f^T(x^*)(x - x^*) \geq 0, \ \forall x \in X$$

όπου $x^* = Pr_X\{x^* - s\nabla f(x_k)\}$, $\forall s > 0$

Άρα το $\bar{x}_k$ γίνεται x^* στην τελευταία επανάληψη όπου x^* στάσιμο

Άρα $\nabla f^T(\bar{x}_k) \cdot (x_k - \bar{x}_k) \geq 0 \Rightarrow \nabla f^T(\bar{x}_k) \cdot (x_k - x_k + s_k \nabla f(x_k)) \geq 0 \Rightarrow$

$\Rightarrow \nabla f^T(\bar{x}_k) \cdot s_k \nabla f(x_k) \geq 0$, συνθήκη τερματισμού

Από την 6.1.12:

$$((x^* - s\nabla f(x^*)) - x^*)^T (x - x^*) \leq 0, \ \forall x \in X, s > 0$$

$$((\bar{x}_k - s_k \nabla f(\bar{x}_k)) - \bar{x}_k)^T \cdot (x_k - x_k + s_k \nabla f(x_k)) \leq 0 \Rightarrow$$

$$\Rightarrow (-s_k \nabla f(\bar{x}_k))^T \cdot s_k \cdot \nabla f(x_k) \leq 0 \Rightarrow -s_k^2 (\nabla f(\bar{x}_k))^T \cdot \nabla f(x_k) \leq 0$$

συνθήκη τερματισμού

Με παραμέτρους sk=5, γκ=0.5, ε=0.01 και σημείο εκκίνησης το (5,-5) παρατηρούμε ότι ενώ στην προηγούμενη περίπτωση είχαμε σύγκλιση του αλγορίθμου σε επιθυμητό σημείο με γκ=0.5, εδώ αυτό δεν συμβαίνει, καθώς η τιμή το x1 μειώνεται και τείνει στο μηδέν ενώ η τιμή του x2 κάνει μία εναλλαγή μεταξύ δύο τιμών αρκετά μακριά του μηδέν, με αποτέλεσμα να μην έχουμε σύγκλιση και των δύο τιμών άρα και του σημείου.

Στην συνέχεια, αλλάζουμε τις παραμέτρους και έχουμε sk=15, γκ=0.1, ε=0.01 και σημείο εκκίνησης το (-5,10). Παρατηρούμε ότι δεν έχουμε σύγκλιση του σημείου σε στο (0,0) (ο αλγόριθμος έτρεξε για περίπου 10000 επαναλήψεις οπότε βγήκε αυτό το συμπέρασμα) που είναι το επιθυμητό, όμως βλέποντας τις τιμές των σημείων καθώς προχωράει ο αλγόριθμος παρατηρούμε ότι η τιμή του x1 μηδενίζει ενώ η τιμή του x2 είναι κοντά στο μηδέν (<1) όμως χωρίς να πλησιάζει το μηδέν ώστε να ισχύσει η συνθήκη τερματισμού και να τερματίσει ο αλγόριθμος. Μετά από μερικές δοκιμές παρατηρούμε ότι το πρόβλημα είναι η τιμή το sk και καταλήγουμε στο συμπέρασμα ότι για μεγάλες τιμές του sk δεν έχουμε σύγκλιση, ενώ αν αλλάξουμε την τιμή του σε 10 για παράδειγμα ο αλγόριθμος συγκλίνει στο επιθυμητό σημείο

Τέλος, αλλάζουμε πάλι τις παραμέτρους και πλέον έχουμε sk=0.1, γκ=0.2, ε=0.01 και σημείο εκκίνησης το (8, -10). Εδώ, πριν τρέξουμε την προσομοίωση παρατηρούμε ότι το σημείο εκκίνησης είναι εκτός των περιορισμών που έχουμε θέσει και στις δύο συντεταγμένες, επομένως ο αλγόριθμος δεν είναι σίγουρο ότι θα συγκλίνει, καθώς σύμφωνα με το βιβλίο σελ.200 βασική προϋπόθεση για να λειτουργήσει σωστά είναι το σημείο εκκίνησης να βρίσκεται εντός των περιορισμών του χώρου που επιθυμούμε. Όταν τρέξουμε τον αλγόριθμο παρατηρούμε ότι βγάζει ως αποτέλεσμα το σημείο (1.527345, -1.9e-05) που δεν αποτελεί σωστή λύση. Αυτό συμβαίνει λόγω του αρχικού σημείου που αναφέρθηκε παραπάνω και της επιλογής το sk, το οποίο είναι πολύ μικρό και επέτρεψε τον αλγόριθμο να τερματίσει. Έπειτα, άλλαξα λίγο τον κώδικα ώστε να ελέγχει αν το σημείο εκκίνησης είναι έξω από την επιθυμητή περιοχή, να βρίσκει την προβολή του και από εκεί να ξεκινάει ο αλγόριθμος όμως δημιουργήθηκε το ίδιο πρόβλημα αυτήν την φορά λογικά δεν προέκυψε από το αρχικό σημείο, καθώς αυτό το πρόβλημα το λύσαμε, αλλά από την επιλογή του sk=0.1, το οποίο από ότι φαίνεται δεν αποτελεί καλή επιλογή σε αυτήν την περίπτωση.