Αριθμητικές Μέθοδοι Βελτιστοποίησης

Σύγκλιση (υπεργραμμική ταχύτητα). Τροποποιημένη μέθοδος Νεύτωνα - Εισαγωγή στη γραμμική αναζήτηση

Σύγκλιση (υπεργραμμική)

Θεώρημα: Έστω $f : S \subset \mathbb{R}^n \to \mathbb{R}$, S κυρτό και $\nabla^2 f$ είναι Lipshitz συνεχής στο S. Έστω $\{x_k\} \subset S$, ακολουθία στο S, η οποία συγκλίνει στο $x^\star \in S$, $x_k \neq x^\star$, η οποία δίνεται $x_{k+1} = x_k + p_k$. Επίσης $\nabla^2 f(x^\star)$ είναι θετικά ορισμένος. Τότε

$$x_k \to x^\star, \text{ (υπεργραμμικά) και } \nabla f(x^\star) = 0$$

αν και μόνο αν

$$\lim_{k\to\infty} \frac{\|p_k - (p_N)_k\|}{\|p_k\|} = 0$$

όπου $(p_N)_k$ είναι η κατεύθυνση Νεύτωνα στο σημείο x_k, δηλαδή

$$\nabla^2 f(x_k)(p_N)_k = -\nabla f(x_k)$$

Σύγκλιση (υπεργραμμική) - Απόδειξη

Απόδειξη: Ας υποθέσουμε ότι

$$\lim_{k\to\infty} \frac{\|p_k - (p_N)_k\|}{\|p_k\|} = 0$$

Από τον ορισμό του $(p_N)_k$ έχουμε

$$\nabla^2 f(x_k)(p_N)_k = -\nabla f(x_k) \text{ και } x_{k+1} = x_k + p_k$$

Σύγκλιση (υπεργραμμική) - Απόδειξη

$$\begin{aligned}\nabla f(x_{k+1}) &= [\,\nabla f(x_{k+1}) - \nabla f(x_k)] - \nabla^2 f(x_k)(x_{k+1} - x_k)] + \nabla f(x_k) + \nabla^2 f(x_k)(x_{k+1} - x_k) \\ &= [\,\nabla f(x_{k+1}) - \nabla f(x_k)] - \nabla^2 f(x_k)(x_{k+1} - x_k)] + \nabla^2 f(x_k)(p_k - (p_N)_k)\end{aligned}$$

Επομένως

$$\frac{\|\nabla f(x_{k+1})\|}{\|p_k\|} \leq \frac{\|\nabla f(x_{k+1}) - \nabla f(x_k) - \nabla^2 f(x_k)(x_{k+1} - x_k)\|}{\|p_k\|} + \|\nabla^2 f(x_k)\| \frac{\|p_k - (p_N)_k\|}{\|p_k\|}$$

Σύγκλιση (υπεργραμμική) - Απόδειξη

Χρησιμοποιώντας γνωστό λήμμα του Απειροστικού Λογισμού επειδή η ∇f είναι Lipschitz συνεχής θα έχουμε

$$\|\nabla f(x_{k+1}) - \nabla f(x_k) - \nabla^2 f(x_k)(x_{k+1} - x_k)\| \leq \gamma \|p_k\|^2$$

Άρα

$$\lim_{k\to\infty} \frac{\|\nabla f(x_{k+1})\|}{\|p_k\|} \leq \lim_{k\to\infty} \gamma \|p_k\| + \lim_{k\to\infty} \|\nabla^2 f(x_k)\| \lim_{k\to\infty} \frac{\|p_k - (p_N)_k\|}{\|p_k\|}$$

Συνεπώς λόγω συνέχειας της $\nabla^2 f$ και $p_k \to 0$ έχουμε

$$\|\nabla f(x_k)\| \to \|\nabla f(x^\star)\| = 0$$

Σύγκλιση (υπεργραμμική) - Απόδειξη

Στη συνέχεια δείχνουμε την υπεργραμμική σύγκλιση.

Γνωρίζουμε από λήμμα του απειροστικού λογισμού ότι θα υπάρχει σταθερά $a > 0$ τέτοια ώστε

$$\|\nabla f(x_{k+1})\| = \|\nabla f(x_{k+1}) - \nabla f(x^\star)\| \geq a\|x_{k+1} - x^\star\|, \quad k > k_0$$

Οπότε

$$\frac{\|\nabla f(x_{k+1})\|}{\|p_k\|} \geq \frac{a\|x_{k+1} - x^\star\|}{\|p_k\|} \geq \frac{a\|x_{k+1} - x^\star\|}{\|x_{k+1} - x^\star\| + \|x_k - x^\star\|}$$

Σύγκλιση (υπεργραμμική) - Απόδειξη

$$\frac{\|\nabla f(x_{k+1})\|}{\|p_k\|} \geq \frac{a\|x_{k+1} - x^\star\|}{\|x_{k+1} - x^\star\| + \|x_k - x^\star\|}$$

$$= \frac{a\frac{\|x_{k+1} - x^\star\|}{\|x_k - x^\star\|}}{1 + \frac{\|x_{k+1} - x^\star\|}{\|x_k - x^\star\|}}$$

Οπότε εύκολα έπεται ότι $\frac{\|x_{k+1} - x^\star\|}{\|x_k - x^\star\|} \to 0$ δηλαδή έχουμε υπεργραμμική σύγκλιση.

Σύγκλιση (υπεργραμμική) - Απόδειξη

Στη συνέχεια θα δείξουμε την αντίθετη κατεύθυνση.

Έστω ότι $\{x_k\}$ συγκλίνει υπεργραμμικά και ότι $\nabla f(x^\star) = 0$

Από λήμμα του Απειροστικού Λογισμού έχουμε ότι

$$\|\nabla f(x_{k+1})\| = \|\nabla f(x_{k+1}) - \nabla f(x^\star)\| \leq \beta \|x_{k+1} - x^\star\|, \quad k > k_0$$

Επομένως,

$$\frac{\|x_{k+1} - x^\star\|}{\|x_k - x^\star\|} \geq \frac{1}{\beta} \frac{\|\nabla f(x_{k+1})\|}{\|x_k - x^\star\|} = \frac{1}{\beta} \frac{\|\nabla f(x_{k+1})\|}{\|p_k\|} \frac{\|p_k\|}{\|x_k - x^\star\|}$$

Σύγκλιση (υπεργραμμική) - Απόδειξη

Λόγω της υπεργραμμικότητας θα έχουμε

$$\frac{\|\nabla f(x_{k+1})\|}{\|p_k\|}\frac{\|p_k\|}{\|x_k - x^\star\|} \to 0$$

Αν δείξουμε ότι $\dfrac{\|p_k\|}{\|x_k - x^\star\|} \to 1$, τότε $\dfrac{\|\nabla f(x_{k+1})\|}{\|p_k\|} \to 0$

Σύγκλιση (υπεργραμμική) - Απόδειξη

Πράγματι έχουμε

$$\left| \frac{\|x_{k+1} - x_k\|}{\|x_k - x^\star\|} - 1 \right| = \left| \frac{\|x_{k+1} - x_k\| - \|x_k - x^\star\|}{\|x_k - x^\star\|} \right|$$

$$\leq \frac{\|x_{k+1} - x^\star\|}{\|x_k - x^\star\|} \to 0$$

Σύγκλιση (υπεργραμμική) - Απόδειξη

Επομένως επειδή $\dfrac{\|\nabla f(x_{k+1})\|}{\|p_k\|} \to 0$

Στη συνέχεια χρησιμοποιώντας παρόμοιο επιχείρημα όπως προηγουμένως θα έχουμε

$$\frac{\|\nabla^2 f(x_k)(p_k - (p_N)_k)\|}{\|p_k\|} \leq \frac{\|\nabla f(x_{k+1}) - \nabla f(x_k) - \nabla^2 f(x_k)(x_{k+1} - x_k)\|}{\|p_k\|}$$

$$+\frac{\|\nabla f(x_{k+1})\|}{\|p_k\|}$$

Σύγκλιση (υπεργραμμική) - Απόδειξη

Οπότε εύκολα προκύπτει ότι

$$\frac{\|\nabla^2 f(x_k)(p_k - (p_N)_k)\|}{\|p_k\|} \to 0$$

Τέλος λόγω του γεγονότος ότι $\nabla^2 f(x^\star)$ είναι θετικά ορισμένος, λόγω συνέχειας θα είναι και $\nabla^2 f(x_k)$ θετικά ορισμένος για κατάλληλα μεγάλο k, οπότε θα είναι και αντιστρέψιμος και τότε $\dfrac{\|p_k - (p_N)_k\|}{\|p_k\|} \to 0$

Είναι η κατεύθυνση Νεύτωνα, κατεύθυνση καθόδου;

Έστω $p \in \mathbb{R}^n$ μια κατεύθυνση τότε αν $p^T \nabla f(x_k) < 0$, τα σημεία $x_k + ap$, με $a > 0$ ορίζουν μια κατεύθυνση καθόδου της f γιατί η συνάρτηση

$$F(a) = f(x_k + ap),$$

έχει αρνητική παράγωγο $F'(0) = p^T \nabla f(x_k) < 0$.

Οπότε ελαττώνεται κοντά στο x_k.

Είναι η κατεύθυνση Νεύτωνα, κατεύθυνση καθόδου;

Στη μέθοδο του Νεύτωνα η κατεύθυνση p_k από την οποία προκύπτει η επόμενη προσέγγιση $x_{k+1} = x_k + p_k$ είναι η λύση του γραμμικού συστήματος την $\nabla^2 f(x_k) p_k = -\nabla f(x_k)$.

Αν ο $\nabla^2 f(x_k)$ είναι θετικά ορισμένος τότε

$$p_k^T \nabla^2 f(x_k) p_k = -p_k^T \nabla f(x_k) > 0$$

Οπότε η p_k είναι κατεύθυνση καθόδου αν ο $\nabla^2 f(x_k)$ είναι θετικά ορισμένος.

Τροποποίηση της μεθόδου Νεύτωνα

Αν $\nabla^2 f(x_k)$ δεν είναι θετικά ορισμένος θα μπορούσαμε να τροποποιήσουμε την Εσσιανή της f, ώστε προκύψει ένας θετικά ορισμένος πίνακας, π.χ. θεωρούμε έναν πίνακα E, ώστε ο $\nabla^2 f(x_k) + E$ να είναι θετικά ορισμένος.

Η λύση p του $(\nabla^2 f(x_k) + E)p = -\nabla f(x_k)$, θα είναι κατεύθυνση καθόδου

Τροποποίηση της μεθόδου Νεύτωνα

Τροποποίηση της μεθόδου Νεύτωνα, ώστε να παίρνουμε κατεύθυνση καθόδου

Έστω x_0 αρχική προσέγγιση δημιουργούμε την ακολουθία $\{x_k\}$σύμφωνα με τον αλγόριθμο:

$$x_{k+1} = x_k + p_k$$

$$(\nabla^2 f(x_k) + E_k)p_k = -\nabla f(x_k)$$

όπου E_k τέτοιος ώστε $\nabla^2 f(x_k) + E_k$ συμμετρικός και θετικά ορισμένος

Γραμμική αναζήτηση

Βρίσκοντας μια κατεύθυνση καθόδου p_k και χρησιμοποιώντας ένα σταθερό βήμα για τον υπολογισμό της επόμενης προσέγγισης

$$x_{k+1} = x_k + p_k$$

δεν αρκεί για να μειώνεται η τιμή της f στην επόμενη προσέγγιση,

$$f(x_k + p_k) < f(x_k)$$

Παράδειγμα

Θεωρούμε τη συνάρτηση

$$f(x_1, x_2) = 5x_1^2 + 7x_2^2 - 3x_1x_2$$

Έστω $x_k = (2,3)^T$, $p_k = (5, -7)^T$. Τότε $f(x_k) = 65$ και επειδή

$$\nabla f(x_k) = (10x_1 - 3x_2, 14x_2 - 3x_1),$$

$\nabla f(x_k) = (11,36)$. Οπότε $p_k^T \nabla f(x_k) < 0$, δηλαδή p_k κατεύθυνση καθόδου.

Αν θέσουμε $x_{k+1} = x_k + p_k$, τότε $f(x_{k+1}) = 121 > f(x_k) = 65$.

Όμως αν $x_{k+1} = x_k + \dfrac{1}{2}p_k$, τότε $f(x_{k+1}) = \dfrac{9}{4} < f(x_k) = 65$

Γενικός Αλγόριθμος

1. Θεωρούμε μια αρχική προσέγγιση $x_0 \in \mathbb{R}^n$ και ένα $\epsilon > 0$

2. Για $k = 0,1,2,\ldots$

   (i) Αν το $\|\nabla f(x_k)\| < \epsilon$ έχουμε βρει "σχεδόν" το ελάχιστο και σταματάμε

   (ii) Βρείτε μια κατεύθυνση καθόδου p_k (Μέθοδος Νεύτωνα ή κάποια άλλη)

   (iii) Βρείτε ένα βήμα a_k, ώστε $x_{k+1} = x_k + a_k p_k$ να είναι καλύτερη προσέγγιση από το x_k