Βελτιστοποίηση

Ιωάννης Γ. Τσούλος

Τμήμα Πληροφορικής και τηλεπικοινωνιών
Πανεπιστήμιο Ιωαννίνων

2022

Περίληψη

Ελαχιστοποίηση σε μία διάσταση

1. Μπορεί να χρησιμοποιηθεί όχι μόνο για εύρεση ριζών αλλά και για ελαχιστοποίηση συναρτήσεων.
2. Στην μονοδιάστατη περίπτωση γίνεται ενημέρωση του τρέχοντος σημείου ως ακολούθως:

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}$$

3. Βασική απαίτηση $f''(x) > 0$, δηλαδή για κυρτές συναρτήσεις.

Παράδειγμα σε μία διάσταση

1. Έστω η συνάρτηση

$$f(x) = \frac{x^2}{2} - 0.8x + 0.2\cos(x)$$

2. Η δεύτερη παράγωγος είναι

$$f''(x) = 1 - 0.2\cos(x)$$

και είναι μονίμως θετική

3. Με αρχικό διάστημα [0,10] γίνονται τα ακόλουθα βήματα

ITER	f(x)
1	0.05766426794
2	-0.1913444931
3	-0.1925042934
4	-0.1925043213

Μέθοδος Newton σε πολλές διαστάσεις

1. **Θέσε** $k = 0$
2. **Αρχικοποίηση** x_0
3. **Όσο** δεν ικανοποιούνται τα κριτήρια τερματισμού
    1. **Επίλυση** του συστήματος $H_f\left(x_k\right)\Delta x_k = \nabla f\left(x_k\right)$, χρήση Εσσιανού πίνακα
    2. **Θέσε** $x_{k+1} = x_k + \Delta x_k$
    3. k=k+1
4. **Τέλος Επανάληψης**

Παράδειγμα εφαρμογής σε δύο διαστάσεις

1. Έστω η συνάρτηση

$$f(x) = 2x_1^2 + 3x_1x_2 - 5x_1 + 4x_2^2 + 2x_2$$

2. Ο Εσσιανός πίνακας δίνεται από την εξίσωση

$$H_f = \begin{bmatrix} 4 & 3 \\ 3 & 8 \end{bmatrix}$$

3. Αν θέσουμε $x_0 = (1, 0)$τότε το ελάχιστο $(2, -1)$ θα βρεθεί μόλις σε ένα βήμα

Προβλήματα μεθόδου Newton

- Θα πρέπει να γίνει αναζήτηση κοντά στο όρια ενός ελαχίστου.
- Θα πρέπει ο Εσσιανός πίνακας να είναι μονίμως θετικά ορισμένος.
- Ο Εσσιανός απαιτεί n^2 υπολογισμούς και σε πολλές περιπτώσεις δεν παρέχεται.
- Λύση είναι να χρησιμοποιηθεί μια προσέγγιση του Εσσιανού.

Η μέθοδος BFGS

1. Η μέθοδος ξεκινά θεωρώντας πως μια προσέγγιση του Εσσιανό είναι ο πίνακας I.
2. Ο πίνακας I ειναι θετικά ορισμένος.
3. Θέτουμε:

$$\Delta g_k = g_{k+1} - g_k \, , \; \Delta x_k = x_{k+1} - x_k$$

4. Σε κάθε επανάληψη διορθώνει αυτήν την εκτίμηση

$$B_{(k+1)} = \left[I - \frac{\Delta g_k \Delta x_k^T}{\Delta x_k^T \Delta g_k} \right] B_k \left[I - \frac{\Delta g_k \Delta x_k^T}{\Delta x_k^T \Delta g_k} \right] + \frac{\Delta g_k \Delta g_k^T}{\Delta x_k^T \Delta g_k}$$

Τα βήματα της μεθόδου BFGS

1. Έστω x_0 ένα αρχικό σημείο της συνάρτησης.
2. Θέσε $B_0 = I$
3. Μέχρι Να ισχύσουν τα κριτήρια τερματισμού επανέλαβε
   1. Επίλυση του συστήματος $B_k s_k = -g_k$
   2. Υπολογισμός του λ^* με γραμμική αναζήτηση ως προς την γραμμή $x_k + \lambda s_k$
   3. Θέσε $x_{k+1} = x_k + \lambda^* s_k$
   4. Ενημέρωση του B_{k+1}
   5. k=k+1
4. Τέλος Επανάληψης

Ιακωβιανός πίνακας

- Έστω η διανυσματική συνάρτηση $f = [f_1, f_2, ..., f_m]$, πχ σε προσαρμογή δεδομένων.
- Ο Ιακωβιανός πίνακας ορίζεται ως

$$J_f(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

Η μέθοδος

1. Στην μέθοδο Newton ο Εσσιανός πρέπει να είναι θετικά ορισμένος.
2. Χρησιμοποιείται ο τροποποιημένος Εσσιανός:
$$H_f + \mu I$$
ο οποίος έχει θετικές ιδιοτιμές αυστηρά για κατάλληλη επιλογή του μ.
3. Η μέθοδος Levenger Marquardt κάνει την ακόλουθη ενημέρωση στα σημεία:
$$x_{k+1} = x_k - \left(H_f\left(x_k\right) + \mu_k I\right)\right)^{-1} \nabla f\left(x_k\right)$$
4. Η μέθοδος αυτή εξασφαλίζει ότι θα υπάρχει μείωση στις τιμές της συνάρτησης
5. Αν $\mu_k \rightarrow 0$, τότε έχει την συμπεριφορά της μεθόδου του Newton.
6. Το μεγάλο πρόβλημα είναι όμως ότι απαιτεί την χρήση

Εφαρμογή σε προβλήματα ελαχίστων τετραγώνων

- Το πρόβλημα ορίζεται ως

$$\min_{x} \sum_{i=1}^{m} (r_i(x))^2$$

- Η παράγωγος δίνεται από:

$$\nabla f\,(x_i) = 2\sum r_{j=1}^{m} r_j(x)\frac{\partial r_j(x)}{\partial x_i}$$

- Με χρήση του Ιακωβιανού πίνακα:

$$\nabla f(x) = 2J_r\,(x)^T\, r(x)$$

  Ο Εσσιανός υπο προυποθέσεις μπορεί να γραφεί ως:
  $H_f(x) \simeq 2J_r(x)^T J_r(x)$

- Επομένως

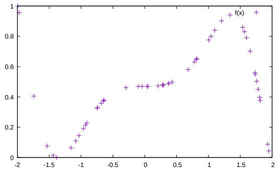
$$x_{k+1} = x_k - \left(J_r\,(x_k)^T\, J_r\,(x_k) + \mu_k I\right)^{-1} J_r(x_k)^T\, r\,(x_k)$$

Προσαρμογή δεδομένων

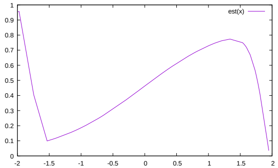
1. Στην γενική μορφή υπάρχουν μια σειρα από δεδομένα (x_i, y_i), $i = 1, .., M$ από μια άγνωστη συνάρτηση.
2. Το ζητούμενο είναι να βρεθεί ένα μοντέλο που να διέρχεται από αυτά τα σημεία αλλά και που να μπορεί να εκτιμήσει την συμπεριφορά της άγνωστης συνάρτησης και σε άλλα σημεία.
3. Αυτού του είδους τα προβλήματα ονομάζονται προβλήματα προσαρμογής δεδομένω ή data fitting.

Παράδειγμα Data Fitting

- Σημεία της συνάρτησης



- Σημεία που μπορεί να μάθει ένα μοντέλο



Προσαρμογή δεδομένων με πολυώνυμο

- Δημιουργείται ένα πολυώνυμο βαθμού $k < M$
- Το πολυώνυμο δίνεται από τον τύπο:

$$p_k(x) = a_k x^k + a_{k-1} x^{k-1} + \ldots + a_0$$

- Γίνεται ελαχιστοποίηση της συνάρτησης σφάλματος:

$$E\left(p_k(x)\right) = \sum_{i=1}^{M} \left(y_i - p_k\left(x_i\right)\right)^2$$

- Το ολικό ελάχιστο αυτής της συνάρτησης είναι γνωστό και είναι μηδέν.

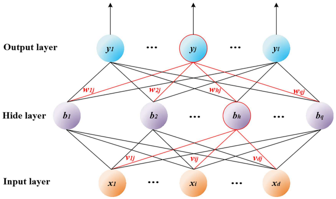
Συνάρτηση σφάλματος σε πολυώνυμο

- Ελαχιστοποίηση της $E\left(p_k(x)\right) = \sum_{i=1}^{M}\left(y_i - p_k\left(x_i\right)\right)^2$ με άγνωστους τα στοιχεία $a_i, \; i = 1, .., k$
- Μερικές παράγωγοι:

$$\frac{\partial E}{\partial a_j} = 2\sum_{i=1}^{M}\left(y_i - p_k\left(x_i\right)\right)\left(-x_i^j\right)$$

Τεχνητό νευρωνικό δίκτυο

- Είναι παραμετρικό μοντέλο όπως το πολυώνυμο αλλά με μεγαλύτερη ευελιξία
- Σχηματικά:



Τεχνητό νευρωνικό δίκτυο (εξισώσεις)

- Μια μορφή εξίσωσης των τεχνητών νευρωνικών δικτυων μπορεί να είναι:

$$N\left(\overrightarrow{x},\overrightarrow{w}\right)=\sum_{i=1}^{H} w_{(d+2)i-(d+1)}\sigma\left(\sum_{j=1}^{d} x_j w_{(d+2)i-(d+1)+j}+w_{(d+2)i}\right)$$

- Η ποσότητα $w_{(d+2)i}$ ονομάζεται πόλωση.
- Η συνάρτηση $\sigma(x)$ είναι η σιγμοειδής συνάρτηση με τύπο:

$$\sigma(x)=\frac{1}{1+e^{-x}}$$

Σύνοψη

- Η μέθοδος Newton
- Η μέθοδος Levenger Marquadrt
- Εφαρμογές

Βιβλιογραφία I

- Βελτιστοποίηση τεχνικών συστημάτων, Άγγελος Πρωτοπαπάς, 2015, Κάλλιπος, ISBN: 978-960-603-493-0
- Καθολική βελτιστοποίηση: μέθοδοι λογισμικό και εφαρμογές, Ιωάννης Γ. Τσούλος, διδακτορική διατριβή, εθνικό αρχείο διδακτορικών διατριβών.
- Αριθμητικές μέθοδοι βελτιστοποίησης, Ισαάκ Λαγαρής, ιστοσελίδα σημειώσεων διαθέσιμη από http://www.cs.uoi.gr/~lagaris/OPT_UNDER/
- Τεχνικές βελτιστοποίησης, Γεώργιος Α. Ροβιθάκης, 2007, εκδόσεις Τζιόλα, ISBN-13: 978-960-418-141-4 .

Βιβλιογραφία II

- Μαθηματική θεωρία βελτιστοποίησης , DingZhu Du, Panos M. Pardalos, Weili Wu, 2005, Εκδόσεις Νέων Τεχνολογιών, ISBN 960-8105-79-X, ISBN-13 978-960-8105-79-9
- Boender, C.G.E., Rinnooy Kan, A.H.G. Bayesian stopping rules for multistart global optimization methods. Mathematical Programming 37, 59–80 (1987)
- Rinnooy Kan, A. H. G., & Timmer, G. T. (1985). The Multi Level Single Linkage Method For Unconstrained And Constrained Global Optimization (No. 2099-2018-3232).
- Ioannis G. Tsoulos, Isaac E. Lagaris, MinFinder: Locating all the local minima of a function, Computer Physics Communications 174, pp. 166-179, 2006.