Διάλεξη 9

Αλγόριθμοι αναζήτησης σε γραμμή, Τρίτη 20/3/18

Περιεχόμενα

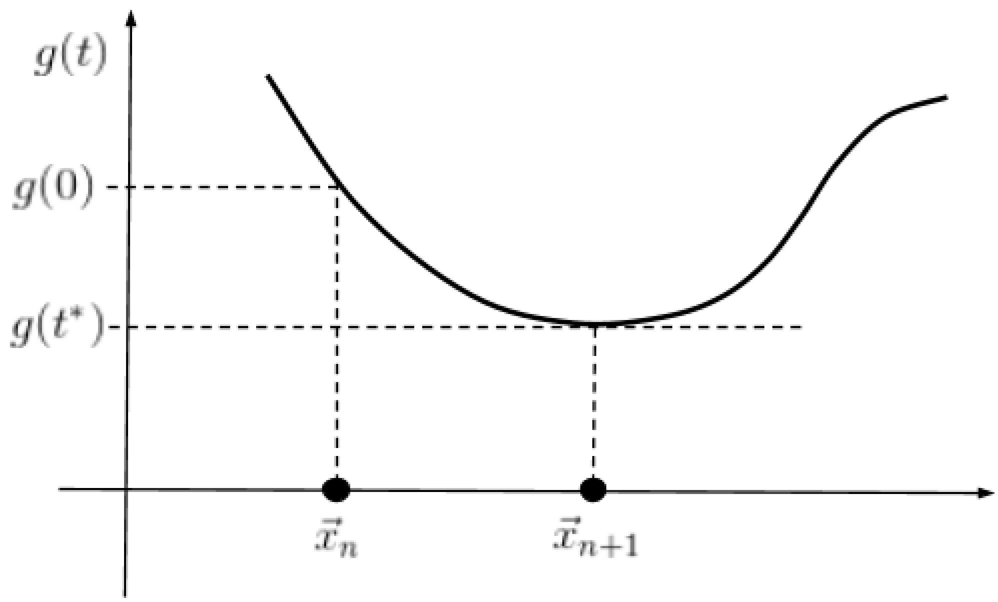
1. Αλγόριθμοι αναζήτησης σε γραμμή

Οι αλγόριθμοι αναζήτησης σε γραμμή παίρνουν ως είσοδο μια διεύθυνση $\vec{\delta}$ και παράγουν ως έξοδο ένα πολλαπλάσιο t της διεύθυνσης. Με αυτόν τον τρόπο η νέα εκτίμηση υπολογίζεται ως $\vec{x}_{n+1} = \vec{x}_n + t\vec{\delta}$.

Η τιμή t επιλέγεται με τέτοιον τρόπο ώστε η εκτίμηση $\vec{x}_{n+1}$ να είναι "αρκετά καλύτερη" από τη $\vec{x}_n$. Ανάλογα με τι θεωρούμε "αρκετά καλύτερη" θα περιγράψουμε δύο τρόπους αναζήτησης: την ακριβή αναζήτηση και την οπισθοδρόμηση.

1.1 Ακριβής αναζήτηση

Εδώ κατά μήκος της διεύθυνσης $\vec{\delta}$, αναζητείται το σημείο $\vec{x}_{n+1}$ που ελαχιστοποιεί την τιμή της αντικειμενικής συνάρτησης. Πιο συγκεκριμένα, έστω $g(t) = f(\vec{x}_n + t\vec{\delta})$ ο περιορισμός της f πάνω στη διεύθυνση $\vec{\delta}$. Η ακριβής αναζήτηση επιλύει το πρόβλημα $\min_t g(t)$ και επιλέγει τη βέλτιστη λύση t^*.



ΜΕΘΟΔΟΣ ΤΗΣ ΑΚΡΙΒΟΥΣ ΑΝΑΖΗΤΗΣΗΣ
Είσοδος: $\vec{x}_n, \vec{\delta}$
$\vec{x}_{n+1} = \vec{x}_n + t^* \vec{\delta}$ **, όπου** t^* **η βέλτιστη λύση του προβλήματος** $\min_t g(t)$ **για** $g(t) = f(\vec{x}_n + t\vec{\delta})$

Παράδειγμα: πιο απότομη κατάβαση με ακριβή αναζήτηση

Θεωρείστε τη συνάρτηση $f(x, y) = x^2 + 2y^2 - 2xy$*. Εκτελέστε 2 βήματα του αλγορίθμου της πιο απότομης κατάβασης χρησιμοποιώντας ακριβή αναζήτηση σε γραμμή, αρχίζοντας από το σημείο* $\vec{x}_0 = (1, 1)$*.*

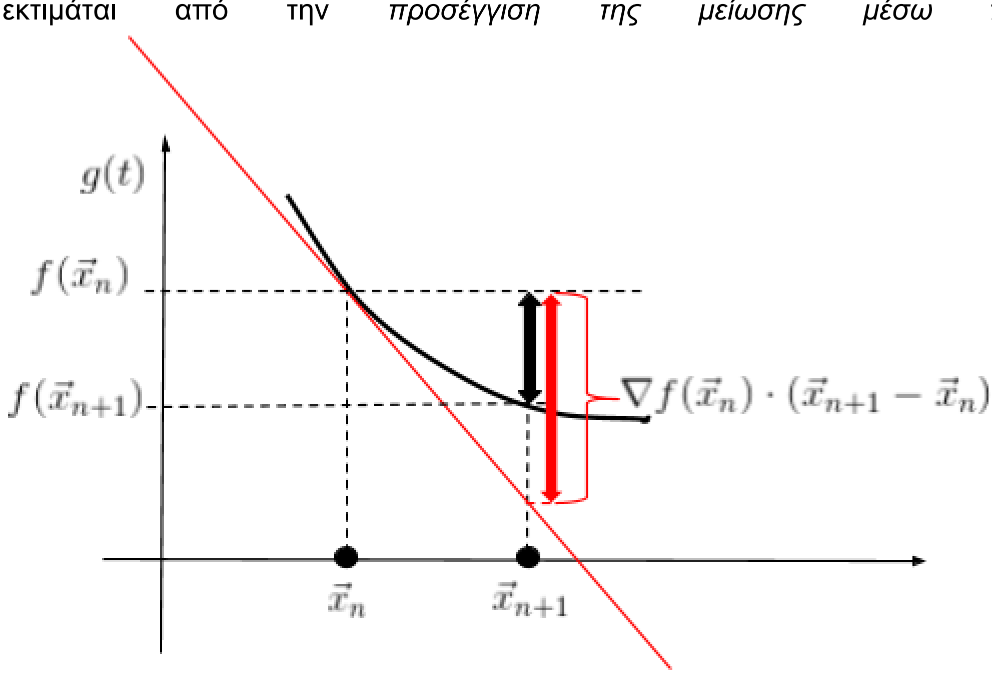
Βήμα 1:

Η διεύθυνση είναι $\vec{\delta} = -\nabla f(1,1) = \left(2x - 2y|_{x=y=1}, 4y - 2x|_{x=y=1}\right) = (0, -2)$. Ο περιορισμός της f στη γραμμή $\vec{x}_0 + t\vec{\delta} = (1, 1-2t)$ είναι η συνάρτηση $g(t) = f(1, 1-2t) = 8t^2 - 4t + 1$ η οποία είναι κυρτή και η παράγωγος της μηδενίζεται στο $t^* = 1/4$. Συνεπώς η ελάχιστη τιμή της f στην παραπάνω γραμμή λαμβάνεται στο σημείο $\vec{x}_1 = (1, 1/2)$.

Βήμα 2:
Τώρα $\delta = -\nabla f(1, 1/2) = (-1, 0)$ και η ελάχιστη τιμή της f στη γραμμή $(1-t, 1/2)$ επιτυγχάνεται στο σημείο $\vec{x}_2 = (1/2, 1/2)$ (γιατί;)

1.2 Οπισθοδρόμηση (backtracking)

Ενώ η ακριβής αναζήτηση βρίσκει το σημείο $\vec{x}_{n+1}$ κατά μήκος της $\vec{\delta}$ που επιτυγχάνει τη μεγαλύτερη μείωση στην τιμή της αντικειμενικής συνάρτησης, η μέθοδος της οπισθοδρόμησης επιστρέφει ένα σημείο όπου απλά η μείωση είναι αρκετή. Μια μείωση χαρακτηρίζεται αρκετή εάν ικανοποιείται η συνθήκη Armijo-Goldstein για δεδομένο $c \in (0, 1)$. Η συνθήκη αυτή εξετάζει εάν η μείωση $f(\vec{x}_{n+1}) - f(\vec{x}_n) = g(t) - g(0)$ υπερβαίνει το $100c$% της μείωσης όπως εκτιμάται από την *προσέγγιση της μείωσης μέσω της παραγώγου*:



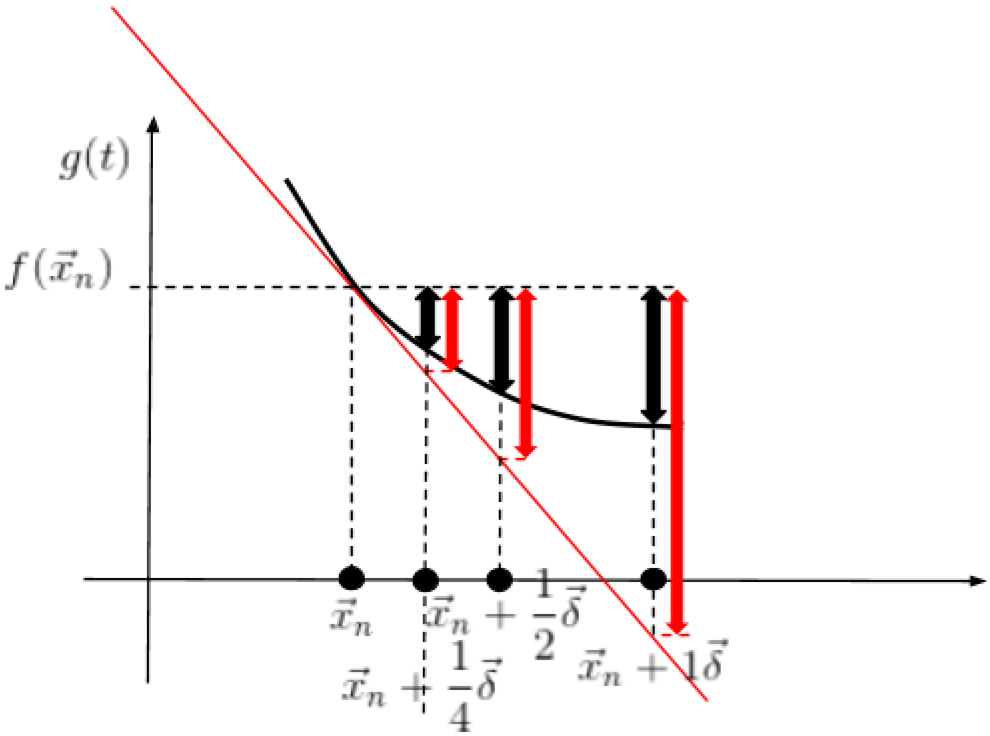
$tg'(0) = t\nabla f(\vec{x}_n) \cdot \vec{\delta} = \nabla f(\vec{x}_n) \cdot (\vec{x}_{n+1} - \vec{x}_n)$.

Δηλαδή εάν $\dfrac{|f(\vec{x}_{n+1}) - f(\vec{x}_n)|}{|\nabla f(\vec{x}_n) \cdot (\vec{x}_{n+1} - \vec{x}_n)|} \geq c$ ή ισοδύναμα:

ΣΥΝΘΗΚΗ ARMIJO-GOLDSTEIN ΓΙΑ $c \in (0, 1)$:

$$f(\vec{x}_{n+1}) \leq f(\vec{x}_n) + c\nabla f(\vec{x}_n) \cdot (\vec{x}_{n+1} - \vec{x}_n)$$

Η μέθοδος της οπισθοδρόμησης ελέγχει διαδοχικά τα σημεία $\vec{x}_{n+1} = \vec{x}_n + t\vec{\delta}$ για $t = 1, \frac{1}{2^1}, \frac{1}{2^2}, \ldots$ έως ότου βρεθεί κάποιο που ικανοποιεί τη συνθήκη Armijo-Goldstein.



ΜΕΘΟΔΟΣ ΤΗΣ ΟΠΙΣΘΟΔΡΟΜΗΣΗΣ (BACKTRACKING)

Είσοδος: $\vec{x}_n, \vec{\delta}, c$

1. **Θέσε** $t = 1$**.**
2. **Υπολογισμός** $\vec{x}_{n+1} = \vec{x}_n + t\vec{\delta}$ **.**
3. **Εάν** $f(\vec{x}_{n+1}) \leq f(\vec{x}_n) + c\nabla f(\vec{x}_n) \cdot (\vec{x}_{n+1} - \vec{x}_n)$ **τότε ΤΕΛΟΣ.**
4. **Θέσε** $t = t/2$ **και μετάβαση στο βήμα 2.**

Παρατηρήστε ότι η μέθοδος πάντα τερματίζει αφού εάν δε συνέβαινε αυτό, λόγω της συνεχούς οπισθοδρόμησης θα είχαμε $t \to 0$ και συνεπώς $(g(t) - g(0))/t \to g'(0)$. Σε αυτή την περίπτωση το ποσοστό μείωσης θα έτεινε στο 100%. Εφόσον $100 > 100c$, η μέθοδος θα τερματίσει μετά από έναν πεπερασμένο αριθμό βημάτων.