TỐI ƯU KHÔNG CÓ RÀNG BUỘC

THUẬT TOÁN

Dương Thị Kim Huyền – Tạ Thúy Anh

Khoa công nghệ thông tin
Đại học PHENIKAA

Hà Nội - 2023

1. Phương pháp gradient
2. Phương pháp Newton
3. Phương pháp hướng gradient liên hợp

Phương pháp gradient

Các thuật toán lặp (iterative algorithms)

Cho $f : \mathbb{R}^n \to \mathbb{R}$ là hàm số khả vi liên tục.

Xét bài toán tối ưu không có ràng buộc

$$\min\{f(x) : x \in \mathbb{R}^n\}.$$

- Đi tìm điểm dừng

Giải phương trình $\nabla f(x) = 0$.

- Nếu f là hàm lồi thì điểm dừng là nghiệm của bài toán.

Các thuật toán lặp (iterative algorithms)

Cho $f : \mathbb{R}^n \to \mathbb{R}$ là hàm số khả vi liên tục.
Xét bài toán tối ưu không có ràng buộc

$$\min\{f(x) : x \in \mathbb{R}^n\}.$$

- Đi tìm điểm dừng

Giải phương trình $\nabla f(x) = 0$.

- Nếu f là hàm lồi thì điểm dừng là nghiệm của bài toán.

Các thuật toán lặp (iterative algorithms)

Cho $f : \mathbb{R}^n \to \mathbb{R}$ là hàm số khả vi liên tục.
Xét bài toán tối ưu không có ràng buộc

$$\min\{f(x) : x \in \mathbb{R}^n\}.$$

- Đi tìm điểm dừng

Giải phương trình $\nabla f(x) = 0$.

- Nếu f là hàm lồi thì điểm dừng là nghiệm của bài toán.

Các thuật toán lặp (iterative algorithms)

Cho $f : \mathbb{R}^n \to \mathbb{R}$ là hàm số khả vi liên tục.
Xét bài toán tối ưu không có ràng buộc

$$\min\{f(x) : x \in \mathbb{R}^n\}.$$

- Đi tìm điểm dừng

Giải phương trình $\nabla f(x) = 0$.

- Nếu f là hàm lồi thì điểm dừng là nghiệm của bài toán.

Các thuật toán lặp (iterative algorithms)

Cho $f : \mathbb{R}^n \to \mathbb{R}$ là hàm số khả vi liên tục.
Xét bài toán tối ưu không có ràng buộc

$$\min\{f(x) : x \in \mathbb{R}^n\}.$$

- Đi tìm điểm dừng

Giải phương trình $\nabla f(x) = 0$.

- Nếu f là hàm lồi thì điểm dừng là nghiệm của bài toán.

Thuật toán lặp đi tìm điểm dừng

$$x_{k+1} = x_k + t_k d_k$$

- d_k: hướng (direction)
- t_k: bước lặp (stepsize)

Hướng giảm I

Definition

Cho $f : \mathbb{R}^n \to \mathbb{R}$ là hàm số khả vi liên tục. Một véctơ $d \in \mathbb{R}^n$ được gọi là một hướng giảm (descent direction) của f tại x nếu đạo hàm theo hướng $f'(x; d) < 0$, tức là

$$f'(x; d) = \nabla f(x)^T d < 0.$$

Tính chất của hướng giảm

Lemma

Cho f là là hàm số khả vi liên tục trên $\mathbb{R}^n$ và $x \in \mathbb{R}^n$. Giả sử rằng d là một hướng giảm của f tại x. Khi đó, tồn tại $c > 0$ sao cho

$$f(x + td) < f(x)$$

với mọi $t \in (0, \epsilon]$.

Phương pháp hướng giảm

Descent Directions Method

Initialization: Pick $x_0 \in \mathbb{R}^n$ arbitrarily.

General step: For any $k = 0, 1, 2, \ldots$

(a) Pick a descent direction d_k.

(b) Find a stepsize t_k satisfying $f(x_k + t_k d_k) < f(x_k)$.

(c) Set $x_{k+1} = x_k + t_k d_k$.

(d) If stopping criterion is satisfied, then STOP and x_{k+1} is the output

Phương pháp hướng giảm

Descent Directions Method

Initialization: Pick $x_0 \in \mathbb{R}^n$ arbitrarily.

General step: For any $k = 0, 1, 2, \ldots$

(a) Pick a descent direction d_k.

(b) Find a stepsize t_k satisfying $f(x_k + t_k d_k) < f(x_k)$.

(c) Set $x_{k+1} = x_k + t_k d_k$.

(d) If stopping criterion is satisfied, then STOP and x_{k+1} is the output

Phương pháp hướng giảm

Descent Directions Method

Initialization: Pick $x_0 \in \mathbb{R}^n$ arbitrarily.

General step: For any $k = 0, 1, 2, \ldots$

(a) Pick a descent direction d_k.

(b) Find a stepsize t_k satisfying $f(x_k + t_k d_k) < f(x_k)$.

(c) Set $x_{k+1} = x_k + t_k d_k$.

(d) If stopping criterion is satisfied, then STOP and x_{k+1} is the output

Phương pháp hướng giảm

Descent Directions Method

Initialization: Pick $x_0 \in \mathbb{R}^n$ arbitrarily.

General step: For any $k = 0, 1, 2, \ldots$

(a) Pick a descent direction d_k.

(b) Find a stepsize t_k satisfying $f(x_k + t_k d_k) < f(x_k)$.

(c) Set $x_{k+1} = x_k + t_k d_k$.

(d) If stopping criterion is satisfied, then STOP and x_{k+1} is the output

Phương pháp hướng giảm

Descent Directions Method

Initialization: Pick $x_0 \in \mathbb{R}^n$ arbitrarily.

General step: For any $k = 0, 1, 2, \ldots$

(a) Pick a descent direction d_k.

(b) Find a stepsize t_k satisfying $f(x_k + t_k d_k) < f(x_k)$.

(c) Set $x_{k+1} = x_k + t_k d_k$.

(d) If stopping criterion is satisfied, then STOP and x_{k+1} is the output

Phương pháp hướng giảm

Descent Directions Method

Initialization: Pick $x_0 \in \mathbb{R}^n$ arbitrarily.

General step: For any $k = 0, 1, 2, \ldots$

(a) Pick a descent direction d_k.

(b) Find a stepsize t_k satisfying $f(x_k + t_k d_k) < f(x_k)$.

(c) Set $x_{k+1} = x_k + t_k d_k$.

(d) If stopping criterion is satisfied, then STOP and x_{k+1} is the output

Phương pháp hướng giảm

Descent Directions Method

Initialization: Pick $x_0 \in \mathbb{R}^n$ arbitrarily.

General step: For any $k = 0, 1, 2, \ldots$

(a) Pick a descent direction d_k.

(b) Find a stepsize t_k satisfying $f(x_k + t_k d_k) < f(x_k)$.

(c) Set $x_{k+1} = x_k + t_k d_k$.

(d) If stopping criterion is satisfied, then STOP and x_{k+1} is the output

??????????????

- Điểm bắt đầu (xuất phát) là điểm nào?
- Bước lặp (stepsize) chọn như thế nào?
- Tiêu chuẩn dừng thuật toán?

?????????????

- Điểm bắt đầu (xuất phát) là điểm nào?
- Bước lặp (stepsize) chọn như thế nào?
- Tiêu chuẩn dừng thuật toán?

?????????????

- Điểm bắt đầu (xuất phát) là điểm nào?
- Bước lặp (stepsize) chọn như thế nào?
- Tiêu chuẩn dừng thuật toán?

- Điểm bắt đầu (xuất phát) là điểm bất kỳ $x_0 \in \mathbb{R}^n$

Chọn bước lặp (stepsize)

- Bước lặp hằng $t_k = t$ với mọi k
- Bước lặp theo *"exact line search"* t_k là cực tiểu dọc theo tia $x_k + td_k$:

$$t_k = \operatorname{argmin}_{t \geqslant 0} f(x_k + td_k)$$

- Backtracking

Chọn bước lặp (stepsize)

- Bước lặp hằng $t_k = t$ với mọi k
- Bước lặp theo *"exact line search"* t_k là cực tiểu dọc theo tia $x_k + td_k$:

$$t_k = \operatorname{argmin}_{t \geqslant 0} f(x_k + td_k)$$

- Backtracking

Chọn bước lặp (stepsize)

- Bước lặp hằng $t_k = t$ với mọi k
- Bước lặp theo *"exact line search"* t_k là cực tiểu dọc theo tia $x_k + td_k$:

$$t_k = \operatorname{argmin}_{t \geqslant 0} f(x_k + td_k)$$

- Backtracking

- Tiêu chuẩn dừng thuật toán:

$$\|\nabla f(x_{k+1})\| < \epsilon$$

ϵ là số dương đủ bé, thường chọn $\epsilon = 10^{-6}$ hoặc $\epsilon = 10^{-5}$

Chú ý: Chuẩn của một véctơ $x = (a_1, a_2, \ldots, a_n)$ trong $\mathbb{R}^n$ được định nghĩa là

$$\|x\| = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}.$$

- Tiêu chuẩn dừng thuật toán:

$$\|\nabla f(x_{k+1})\| < \epsilon$$

ϵ là số dương đủ bé, thường chọn $\epsilon = 10^{-6}$ hoặc $\epsilon = 10^{-5}$

Chú ý: Chuẩn của một véctơ $x = (a_1, a_2, \ldots, a_n)$ trong $\mathbb{R}^n$ được định nghĩa là

$$\|x\| = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}.$$

- Tiêu chuẩn dừng thuật toán:

$$\|\nabla f(x_{k+1})\| < \epsilon$$

ϵ là số dương đủ bé, thường chọn $\epsilon = 10^{-6}$ hoặc $\epsilon = 10^{-5}$

Chú ý: Chuẩn của một véctơ $x = (a_1, a_2, \ldots, a_n)$ trong $\mathbb{R}^n$ được định nghĩa là

$$\|x\| = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}.$$

Hướng giảm: ngược hướng gradient

$d_k = -\nabla f(x_k)$ là một hướng giảm tại x_k nếu $\nabla f(x_k) \neq 0$.

Thật vậy

$$f'(x_k; -\nabla f(x_k)) = -\nabla f(x_k)^T \nabla f(x_k) = -\|\nabla f(x_k)\|^2 < 0.$$

Hướng giảm: ngược hướng gradient

$d_k = -\nabla f(x_k)$ là một hướng giảm tại x_k nếu $\nabla f(x_k) \neq 0$.

Thật vậy

$$f'(x_k; -\nabla f(x_k)) = -\nabla f(x_k)^T \nabla f(x_k) = -\|\nabla f(x_k)\|^2 < 0.$$

Thuật toán gradient (gradient method) I

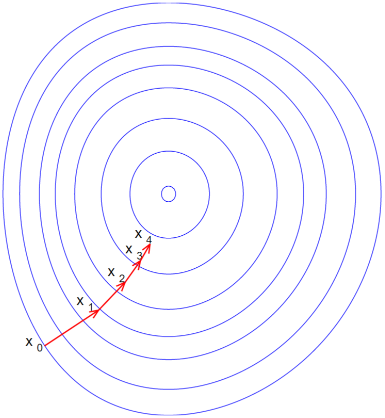
Input: $\epsilon > 0$ - tolerance parameter.

Initialization: Pick $x_o \in \mathbb{R}^n$ arbitrarily.

For any $k = 0, 1, 2, \dots,$ execute the following steps:

- Pick a stepsize t_k by a *line search procedure* on the function
$$g(t) = f(x_k - t_k \nabla f(x_k))$$
- Set $x_{k+l} = x_k - t_k \nabla f(x_k)$.
- If $\|\nabla f(x_{k+1})\| < \epsilon$, then STOP and x_{k+1} is the output

Thuật toán gradient (gradient method) II



Hình: Gradient Method

Thuật toán gradient (gradient method) III

Ví dụ: Xây dựng dãy lặp theo phương pháp gradient cho bài toán

$$\{\min f(x,y) = x^2 + 2y^2 : (x,y) \in \mathbb{R}^2\}$$

Start point: $(x_0, y_0) = (2, 1)$
Constant stepsize: $t = 0.1$

Phương pháp Newton

Phương pháp Newton I

- Start: $x_0 \in \mathbb{R}^n$
- $x_{k+1} = x_k - (\|\nabla^2 f(x_k)\|)^{-1} \nabla f(x_k)$

Phương pháp Newton II

Chú ý: Thuật toán Newton cũng là một thuật toán gradient với stepsize

$$t_k = \frac{1}{\|\nabla^2 f(x_k)\|}$$

Phương pháp Newton III

Ví dụ: Xây dựng dãy lặp theo phương pháp Newton cho bài toán

$$\{\min f(x, y) = x^2 + 2y^2 : (x, y) \in \mathbb{R}^2\}$$

Start point: $(x_0, y_0) = (2, 1)$

Phương pháp hướng gradient liên hợp