PHƯƠNG PHÁP GRADIENT

TRẦN HÀ SƠN

Ngày 14 tháng 4 năm 2024

Overview

1. Bài toán quy hoạch không ràng buộc

2. Khái niệm và cách tìm hướng giảm

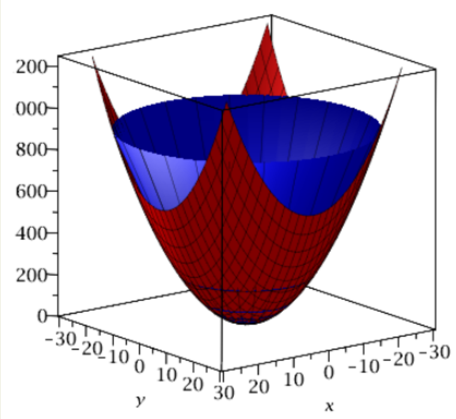
3. Xác định độ dài bước theo quy tắc Armijo

4. Thuật toán gradient với thủ tục quay lui

Bài toán quy hoạch không ràng buộc

Bài toán

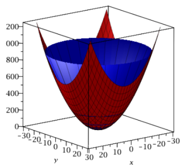
Cho hàm số $f : \mathbb{R}^n \longrightarrow \mathbb{R}$*, khả vi trên* $\mathbb{R}^n$*. Tìm giá trị nhỏ nhất của* $f(x)$ *với* $x \in \mathbb{R}^n$*.*



Bài toán quy hoạch không ràng buộc

Bài toán

Cho hàm số $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, khả vi trên $\mathbb{R}^n$. Tìm giá trị nhỏ nhất của $f(x)$ với $x \in \mathbb{R}^n$.



Ý tưởng: Xuất phát từ một điểm $x^0 \in \mathbb{R}^n$ bất kì, ta xây dựng một dãy điểm $x^1, x^2, \ldots, x^k, \ldots$ sao cho

$$f(x^0) > f(x^1) > f(x^2) > \cdots$$

và dãy số $\{x^k\}$ hội tụ đến điểm dừng $x^* \in \mathbb{R}^n$ của hàm f, thỏa $\nabla f(x^*) = 0$.

Mô hình chung

- Bước khởi đầu: *Chọn một điểm* $x^0 \in \mathbb{R}^n$ *tùy ý. Gán* $k := 0$.
- *Bước lặp* k $(k = 0, 1, 2, \ldots)$
  1. (k_1) *If* x^k *thỏa điều kiện dừng:*

     Dừng thuật toán.

     Else:

     Xác định$x^{k+1} := x^k + t_k d^k$ *sao cho* $f(x^{k+1}) < f(x^k)$.
  2. (k_2) *Gán* $k := k + 1$ *và quay lại Bước lặp* k.

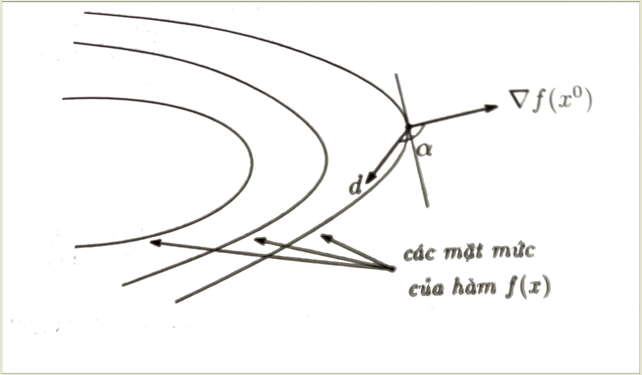
Lưu ý:

- Điều kiện dừng tại Bước (k_1) thường là $\bigtriangledown f(x^k) \approx 0$ hoặc $\|x^k - x^{k-1}\|$ đủ nhỏ.
- $d^k \in \mathbb{R}^n$ là hướng giảm của f tại x^k và $t_k > 0$ được gọi là độ dài bước.

Tìm hướng giảm

Định nghĩa

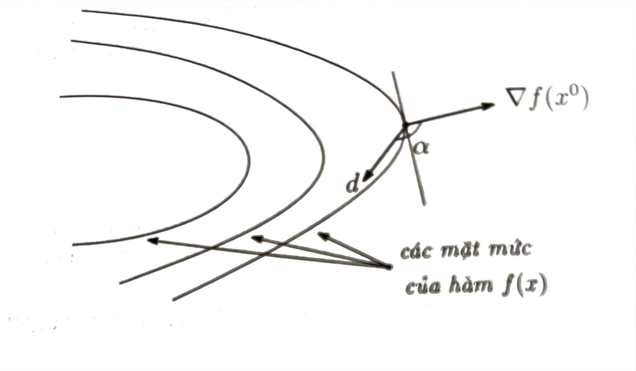
Cho $x^0 \in \mathbb{R}^n$. Ta gọi vector $d \in \mathbb{R}^n$, khác vector không là hướng giảm của hàm f tại x^0 nếu tồn tại $\epsilon > 0$ sao cho với mọi t thỏa $0 < t < \epsilon$ thì $f(x^0 + td) < f(x^0)$.



Tìm hướng giảm

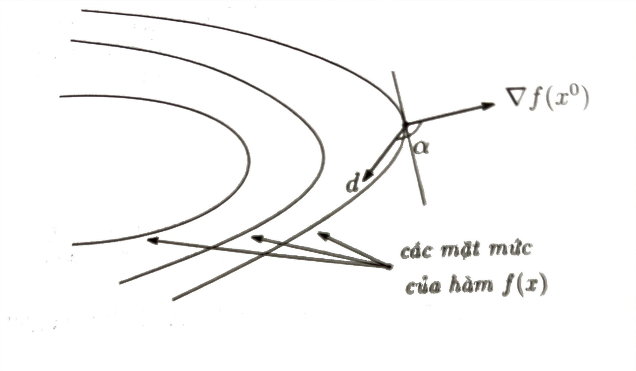
Định lý

Cho hàm f khả vi trên $\mathbb{R}^n$, điểm $x^0 \in \mathbb{R}^n$ và hướng $d \in \mathbb{R}^n/\{0\}$. Nếu $< \bigtriangledown f(x^0), d >< 0$ thì d là hướng giảm của f tại x^0.



Hệ quả

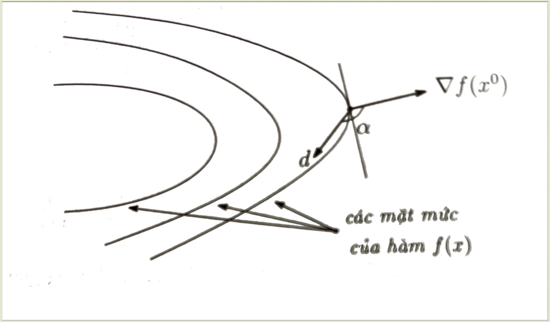
Cho hàm f khả vi trên $\mathbb{R}^n$, điểm $x^0 \in \mathbb{R}^n$. Nếu $\bigtriangledown f(x^0) \neq 0$ thì $d = -\bigtriangledown f(x^0)$ là một hướng giảm của f tại x^0.



Tìm hướng giảm

Hệ quả

Cho hàm f khả vi trên $\mathbb{R}^n$, điểm $x^0 \in \mathbb{R}^n$ và $\bigtriangledown f(x^0) \neq 0$. Trong các hướng giảm d của hàm f tại x^0 có $\|d\| = 1$ thì hàm f giảm nhanh nhất theo hướng $d = -\dfrac{\bigtriangledown f(x^0)}{\| \bigtriangledown f(x^0)\|}$.



Tìm hướng giảm

Ví dụ

Cho điểm $x^0 = (2, 1)$, vector $d = (3, -1)$ và hàm số

$$f(x_1, x_2) = x_2 e^{-(x_1 + x_2)}.$$

a) Vẽ tập hợp các điểm $\{x = x_0 + td | t \geq 0\}$.

b) Kiểm tra xem vector d có phải là hướng giảm của hàm f tại x^0 không? Vì sao?

Xác định độ dài bước theo quy tắc Armijo

Thuật toán (Thủ tục quay lui)

Input *Điểm* $x^k \in \mathbb{R}^n$ *và hướng giảm* d^k *của hàm* f *tại* x^k.

Ouput *Điểm* x^{k+1} *trên tia* $\{x^k + td^k | t > 0\}$ *thỏa* $f(x^{k+1}) < f(x_k)$.

1. *Chọn tùy ý* $m_1 \in (0, 1)$ *và* $\alpha \in (0, 1)$ *và đặt* $t_k := 1$.
2. *Tính* $x^{k+1} = x^k + t_k d^k$ *và* $f(x^{k+1})$.
3. *If* $f(x^{k+1}) \leq f(x^k) + m_1 t_k < \nabla f(x^k), d^k >$
   Then Dừng thuật toán
   Else $t_k = \alpha t_k$ *và quay về Bước 2.*

Thuật toán (Thủ tục quay lui)

Bước khởi đầu *Chọn $m_1 \in (0,1)$ và $\alpha \in (0,1)$, chọn số thực $\epsilon > 0$ đủ nhỏ, chọn một điểm xuất phát tùy ý $x^0 \in \mathbb{R}^n$ sao cho $\bigtriangledown f(x^0) \neq \overrightarrow{0}$. Đặt $k := 0$.*

Bước lặp k với $k = 0, 1, 2 \ldots$

- k_1 *Đặt $t_k := 1$.*
- k_2 *Tính $x^{k+1} = x^k - t_k \bigtriangledown f(x^k)$ và $f(x^{k+1})$.*
- k_3 *If*

$$f(x^{k+1}) - f(x^k) \leq +m_1 t_k < \bigtriangledown f(x^k), -\bigtriangledown f(x^k) >= -m_1 t_k \|\bigtriangledown f(x^k)\|^2$$

Then Chuyển qua Bước k_4
Else $t_k = \alpha t_k$ và quay về Bước k_2.

Thuật toán gradient với thủ tục quay lui

Thuật toán (Thủ tục quay lui)

Bước lặp k với $k = 0, 1, 2 \ldots$

1. *Đặt* $t_k := 1$.
2. *Tính* $x^{k+1} = x^k - t_k \bigtriangledown f(x^k)$ *và* $f(x^{k+1})$.
3. *If*

$$f(x^{k+1}) - f(x^k) \leq +m_1 t_k < \bigtriangledown f(x^k), -\bigtriangledown f(x^k) >= -m_1 t_k \|\bigtriangledown f(x^k)\|^2$$

   Then Chuyển qua Bước k_4
   Else $t_k = \alpha t_k$ *và quay về Bước* k_2.
4. *Tính* $\bigtriangledown f(x^{k+1})$.
5. *If* $\| \bigtriangledown f(x^{k+1})\| < \epsilon$ *Then Dừng thuật toán.*
   Else $k = k + 1$, *quay về Bước lặp* k.

Ví dụ

Tìm giá trị nhỏ nhất của các hàm số sau trên $\mathbb{R}^2$.

a) $f(x_1, x_2) = x_1^3 + x_2^2 - 3x_1 - 2x_2 + 12.$

b) $f(x_1, x_2) = x_1^2 + x_1x_2 + x_2^2 + 3(x_1 + x_2 - 2).$