ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

CHƯƠNG 7:
CÁC PHƯƠNG PHÁP CỰC TIỂU HÓA KHÔNG RÀNG BUỘC

Vũ Văn Thiệu, Đinh Viết Sang, Nguyễn Khánh Phương

TÍNH TOÁN KHOA HỌC

Giới thiệu

1. Nhắc lại một số khái niệm từ giải tích
2. Bài toán qui hoạch phi tuyến không ràng buộc
3. Các phương pháp cực tiểu một biến
   - Hàm đơn cực trị
   - Phương pháp Fibonacci
   - Phương pháp lát cắt vàng
4. Các phương pháp số cực tiểu không ràng buộc
   - Các phương pháp gradient
   - Phương pháp Newton

Nhắc lại một số khái niệm từ giải tích

Không gian Euclid n-chiều

Ký hiệu $\mathbb{R}^n$ - tập các vec tơ thực n-chiều

$$R^n = \{x = (x_1, x_2, \cdots, x_n)^T : x_i \in \mathbb{R}, i = 1, 2, \cdots, n\}$$

trong đó $\mathbb{R}$ là tập số thực. Trên đó ta xác định các phép toán

- Phép cộng hai vec tơ $u = (u_1, u_2, \cdots, u_n)^T$ và $v = (v_1, v_2, \cdots, v_n)^T$

$$u + v = (u_1 + v_1, u_2 + v_2, \cdots, u_n + v_n)$$

- Phép nhân vec tơ với một số thực α

$$\alpha u = (\alpha u_1, \alpha u_2, \cdots, \alpha u_n)^T$$

$\mathbb{R}^n$ cùng các phép toán vừa định nghĩa lập thành một không gian tuyến tính. Các phần tử của $\mathbb{R}^n$ đôi khi là các điểm.

Nhắc lại một số khái niệm từ giải tích

Không gian Euclid n-chiều (tiếp)

- Nếu ta đưa vào thêm khái niệm tích vô hướng của hai vec tơ $u, v \in \mathbb{R}^n$:

$$< u, v >= \sum_{i=1}^{n} u_i \times v_i$$

  thì $\mathbb{R}^n$ cùng với tích vô hướng sẽ trở thành không gian Euclid n-chiều.

- Độ dài chuẩn của vec tơ $u \in \mathbb{R}^n$ là số

$$||u|| =< u, u >^{1/2}= \left(\sum_{i=1}^{n} u_i^2 \right)^{1/2}$$

Nhắc lại một số khái niệm từ giải tích

Không gian Euclid n-chiều (tiếp)

- Khoảng cách giữa hai điểm $u, v \in \mathbb{R}^n$

$$\rho(u, v) = ||u - v|| = \left(\sum_{i=1}^{n} (u_i - v_i)^2 \right)^{1/2}$$

- Đối với $u, v, w \in \mathbb{R}^n$ ta có bất đẳng thức tam giác

$$||u - v|| \leq ||u - w|| + ||w - v||$$

Nhắc lại một số khái niệm từ giải tích

Không gian Euclid n-chiều (tiếp)

- Giả sử $\{u^k, k = 1, 2, \cdots\}$ dãy điểm trong $\mathbb{R}^n$, nghĩa là $u^k \in \mathbb{R}^n$, $k = 1, 2, \cdots$, điểm v được gọi là điểm tới hạn của dãy $\{u^k\}$ nếu tìm được dãy con $\{u^{k(i)}\}$ hội tụ đến v.
- Dãy $\{u^k\}$ được gọi là bị chặn nếu tìm được hằng số $M \geq 0$ sao cho $||u^k|| \leq M$, với mọi $k = 1, 2, \cdots$
- Tập $O(x, \epsilon) = \{u \in \mathbb{R}^n : ||u - x|| < \epsilon\}$ là khối cầu tâm tại x và bán kính $\epsilon > 0$ được gọi là **lân cận** ϵ của x.
- Điểm $v \in \mathbb{R}^n$ được gọi là **điểm tới hạn** của tập $U \subset \mathbb{R}^n$, nếu mọi lân cận ϵ của nó luôn chứa điểm của U khác với v.

Nhắc lại một số khái niệm từ giải tích

Không gian Euclid n-chiều (tiếp)

- Điểm $x \in X$ được gọi là **điểm trong** của tập X nếu tồn tại một ϵ lân cận của nó nằm trọn trong X. Tập các điểm trong của X được ký hiệu là $int(X)$.
- Điểm $x \in X$ được gọi là **điểm biên** của tập X nếu trong mọi ϵ lân cận của nó có điểm những điểm thuộc X và không thuộc X. Tập các điểm trong của X được ký hiệu là $\partial(X)$.
- Tập X được gọi là **tập mở** nếu mỗi điểm $x \in X$ đều là điểm trong của X.
- Tập X trong không gian $\mathbb{R}^n$ được gọi là **bị chặn** hay **giới nội**, nếu tìm được hằng số $L > 0$ sao cho $||u|| \leq L$ với mọi $u \in X$.

Nhắc lại một số khái niệm từ giải tích

Không gian Euclid n-chiều (tiếp)

- Tập X trong không gian $\mathbb{R}^n$ được goi là **tập đóng** nếu nó chứa tất cả các điểm tới hạn.
- Giả sử $\{x^k\}$ là dãy điểm trong tập đóng X và $\lim_{k \to +\infty} x = \overline{x}$, khi đó $\overline{x} \in X$
- Tập X được gọi là **compact** nếu nó đóng và giới nội.
- Giả sử $\{x^k\}$ là dãy điểm trong tập compact X. Khi đó từ $\{x^k\}$ ta luôn có thể trích ra dãy con hội tụ $\{x^{k(i)}\}$ sao cho $\lim_{k(i) \to +\infty} x^{k(i)} = \overline{x}$, khi đó $\overline{x} \in X$

Không gian Euclid n-chiều

Nhắc lại một số khái niệm từ giải tích

Vi phân hàm nhiều biến

Định nghĩa 1 : Giả sử hàm f xác định tại lân cận $O(x, \epsilon)$ của điểm x. Ta nói hàm f là khả vi tại x nếu tìm được vec tơ $f'(x) \in \mathbb{R}^n$ sao cho số gia của hàm số tại x : $\Delta f(x) = f(x + \Delta x) - f(x), ||\Delta x|| \leq \epsilon$ có thể viết dưới dạng

$$\Delta f(x) = < f'(x), \Delta x > + o(x, \Delta x)$$

trong đó $o(x, \Delta x)$ là vô cùng bé bậc cao hơn $||\Delta x||$, nghĩa là $\lim_{||\Delta x|| \to 0} \frac{o(x, \Delta x)}{||\Delta x||} = 0$.
Hàm $f'(x)$ được gọi là gradient của hàm f tại x và thường được ký hiệu là $\nabla f(x)$.

Nhắc lại một số khái niệm từ giải tích

Vi phân hàm nhiều biến (tiếp)

Định nghĩa 2 : Giả sử hàm f xác định tại lân cận $O(x, \epsilon)$ của điểm x. Ta nói hàm f là hai lần khả vi tại x nếu cùng với vec tơ $f'(x)$, tồn tại ma trận đối xứng $f"(x) \in \mathbb{R}^{n \times n}$ sao cho số gia của hàm số tại x có thể viết dưới dạng

$$\Delta f(x) = f(x+\Delta x) - f(x) = < f'(x), \Delta x > + \frac{< f"(x)\Delta x, \Delta x >}{2} + o(x, \Delta x)$$

trong đó $\lim_{||\Delta x|| \to 0} \frac{o(x,\Delta x)}{||\Delta x||^2} = 0$.

Ma trận $f"(x)$ được gọi là ma trận đạo hàm cấp hai hay Hessian của hàm f tại x và đôi khi còn được ký hiệu là $\nabla^2 f(x)$

Nhắc lại một số khái niệm từ giải tích

Vi phân hàm nhiều biến (tiếp)

Định nghĩa 3 : Giả sử hàm f xác định trên tập mở X. Ta nói hàm f là khả vi liên tục trên tập X nếu f là khả vi tại mọi điểm x của X và

$$||f'(x+\Delta x) - f'(x)|| \to 0 \text{ khi } ||\Delta x|| \to 0, \quad \forall x, x+\Delta x \in X$$

Tập các hàm thỏa mãn tính chất này được ký hiệu là $C^1(X)$.
Định nghĩa 4 : Giả sử hàm f xác định trên tập mở X. Ta nói hàm f là hai lần khả vi liên tục trên tập X nếu f là hai lần khả vi tại mọi điểm x của X và

$$||f''(x+\Delta x) - f''(x)|| \to 0 \text{ khi } ||\Delta x|| \to 0, \quad \forall x, x+\Delta x \in X$$

Tập các hàm thỏa mãn tính chất này được ký hiệu là $C^2(X)$.

Nhắc lại một số khái niệm từ giải tích

Vi phân hàm nhiều biến (tiếp)

Công thức Taylor : Giả sử $f(x)$ là hai lần khả vi liên tục tại một ϵ lân cận nào đó của x^o, khi đó ta có

$$\begin{aligned} f(x) =& f(x^o) + < f'(x^o), x - x^o > \\ &+ \frac{1}{2} < f"(x^o)(x - x^o), x - x^o > + \alpha(x, x^o)||x - x^o||^2 \end{aligned}$$

trong đó $\lim_{x \to x^o} \alpha(x, x^o) = 0$, sai số có thể được viết dưới dạng $o(||x - x^o||^2)$

Nhắc lại một số khái niệm từ giải tích

Vi phân hàm nhiều biến (tiếp)

Công thức số gia hữu hạn : Giả sử hàm f là khả vi liên tục trên tập mở S và x là một vec tơ nào đó trong S. Khi đó mọi vec tơ y thỏa mãn $x + y \in S$, luôn tìm được số $\alpha \in [0, 1]$ sao cho

$$f(x+y) - f(x) =< f'(x+\alpha y), y >= \int_0^1 < f'(x+ty), y > dt$$

Nếu f hai lần khả vi liên tục thì ta có công thức:

$$f(x+y) - f(x) =< f'(x), y > + \frac{1}{2} < f''(x+\alpha y)y, y > .$$

Vi phân hàm nhiều biến

Nhắc lại một số khái niệm từ giải tích

Bài toán cực trị hàm nhiều biến

Xét bài toán tối ưu

$$f(x) \to \min, x \in X$$

trong đó $X \subset \mathbb{R}^n$, còn f là hàm xác định trên X.

Định nghĩa 5 : Điểm $x^* \in X$ được gọi là điểm **cực tiểu toàn cục** của f trên X nếu $f(x^*) \leq f(x), \quad \forall x \in X$.

- Giá trị $f(x^*)$ là giá trị cực tiểu của f trên X và ta sẽ ký hiệu $\min\{f(x) : x \in X\}$
- Điểm $x^* \in X$ được gọi là điểm **cực tiểu địa phương** của f trên X nếu tìm được lân cận $O(x, \epsilon), \epsilon > 0$ sao cho $f(x^*) \leq f(x)$, với $x \in O(x, \epsilon) \cap X$.

Nhắc lại một số khái niệm từ giải tích

Bài toán cực trị hàm nhiều biến (tiếp)

Định nghĩa 6 : Giả sử hàm f bị chặn trên X. Số f^* được gọi là cận dưới của f trên X nếu

1. $f^* \leq f(x)$ với mọi $x \in X$
2. Với mọi số $\epsilon > 0$ luôn tìm đc $u^\epsilon \in X$ sao cho $f(u^\epsilon) < f^* + \epsilon$

Khi đó ta ký hiệu : $\inf_{x \in X} f(x) = f^*$

Chú ý

Rõ ràng nếu hàm f đạt cực tiểu toàn cục trên X thì

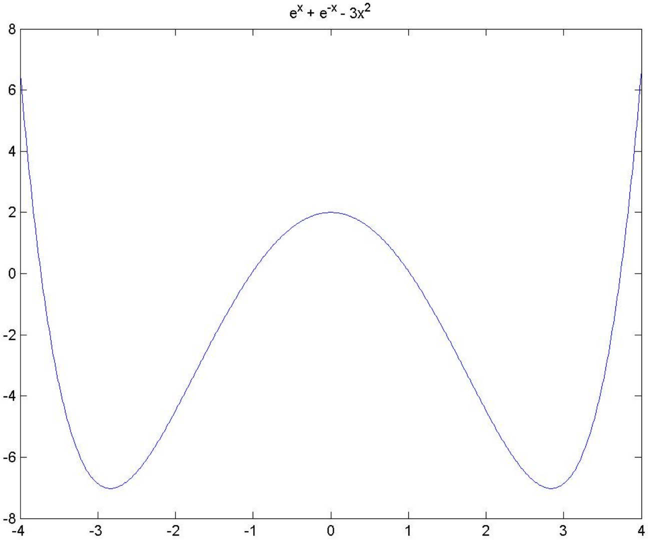
$$\inf_{x \in X} f(x) = \min_{x \in X} f(x)$$

Nhắc lại một số khái niệm từ giải tích

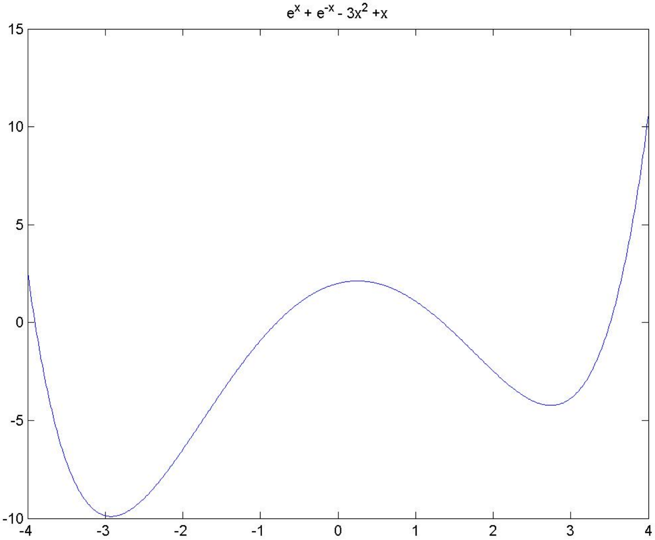
Một số ví dụ

- $f(x) = (x-1)^2$ có cực tiểu toàn cục tại $x^* = 1$ với $f(x^*) = 0$.
- $f(x) = e^x + e^{-x} - 3x^2$. Giá trị tối ưu của hàm $f(x) = -7.02$. Bài toán có cực tiểu toàn cục tại hai điểm $x = \pm 2.84$, không có cực tiểu địa phương.
- $f(x) = e^{-x}$ cận dưới bằng không nhưng không đạt được. Không có cực tiểu địa phương cũng như cực tiểu toàn cục.
- $f(x) = -x + e^{-x}$ Hàm mục tiêu không bị chặn dưới, không có cực tiểu địa phương cũng như cực tiểu toàn cục.
- $f(x) = e^x + e^{-x} - 3x^2 + x$ Bài toán có hai cực tiểu địa phương $x_1 = -2.9226$ và $x_2 = 2.7418$, trong đó x_1 là cực tiểu toàn cục. Giá trị tối ưu của hàm là -9.9040

Nhắc lại một số khái niệm từ giải tích



Nhắc lại một số khái niệm từ giải tích



Bài toán cực trị hàm nhiều biến

Bài toán qui hoạch phi tuyến không ràng buộc

Định nghĩa

Xét bài toán qui hoạch phi tuyến không ràng buộc (unconstrained nonlinear programming)

$$\min\{f(x) : x \in \mathbb{R}^n\} \tag{1}$$

trong đó $f(x)$ khả vi liên tục.

Các định lý

Định lý 1 (Điều kiện cần tối ưu) : Điều kiện cần để x^0 là cực tiểu địa phương là

$$\nabla f(x^0) = 0 \tag{2}$$

Điều kiện (2) được gọi là điều kiện dừng, điểm x^0 thỏa mãn (2) đc gọi là điểm dừng. Như vậy việc giải bài toán (1) có thể qui về giải hệ phương trình (2).

Bài toán qui hoạch phi tuyến không ràng buộc

Các ví dụ

- $f(x) = x^2 - 3x - 1$ phương trình $f'(x) = 2x - 3 = 0$ có nghiệm duy nhất $x^0 = 3/2$ là điểm cực tiểu địa phương, đồng thời là điểm cực tiểu toàn cục.
- $f(x_1, x_2) = x_1^2 + x_2^2 - 2x_1x_2 + x_1$ phương trình
  $\nabla f(x) = \begin{pmatrix} 2x_1 - 2x_2 + 1 \\ -2x_1 + 2x_2 \end{pmatrix} = 0$ có nghiệm duy nhất
  $x^0 = (-1/4, 1/4)$. Tuy nhiên, nghiệm x^0 không là phương án tối ưu của bài toán $\min\{f(x) : x \in \mathbb{R}^2\}$ vì ta có
  $f(-1/4, 1/4) = -1/8 > -1 = f(0, 1)$.

Bài toán qui hoạch phi tuyến không ràng buộc

Các định lý (tiếp)

Định lý 2 (Điều kiện đủ tối ưu) : Giả sử f là hai lần khả vi liên tục. Điểm dừng x^0 là cực tiểu địa phương nếu ma trận $f''(x^0)$ là ma trận xác định dương.

Để biết ma trận có tính xác định dương hay không có thể sử dụng tiêu chuẩn Silvestra sau đây.

Tiêu chuẩn Silvestra : Ma trận $A = (a_{ij})_{n\times n}$ là xác định không âm (bán xác định dương) khi và chỉ khi tất cả các định thức con của nó là không âm

$$\triangle_{i_1,i_2,\cdots,i_k} = det\begin{vmatrix} a_{i_1,i_1} & a_{i_1,i_2} & \cdots & a_{i_1,i_k} \\ a_{i_2,i_1} & a_{i_2,i_2} & \cdots & a_{i_2,i_k} \\ \cdots & & & \\ a_{i_k,i_1} & a_{i_k,i_2} & \cdots & a_{i_k,i_k} \end{vmatrix} \geq 0$$

trong đó $\forall 1 \leq i_1 < i_2 < \cdots < i_k \leq n, \forall k = 1, 2, \cdots, n$

Bài toán qui hoạch phi tuyến không ràng buộc

Các ví dụ

- Xét $f(x_1, x_2) = e^{x_1^2+x_2^2}$ giải hệ phương trình

$$\nabla f(x) = \begin{pmatrix} 2x_1 e^{x_1^2+x_2^2} \\ 2x_2 e^{x_1^2+x_2^2} \end{pmatrix} = 0$$

có nghiệm duy nhất $x^0 = (0,0)$ do $f"(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ có định thức $det|f"(x^0)| > 0$ suy ra x^0 là điểm cực tiểu địa phương đồng thời là phương án tối ưu của bài toán.

Bài toán qui hoạch phi tuyến không ràng buộc

Các ví dụ (tiếp)

- Xét $f(x_1, x_2) = -x_1^2 + x_2^2 - 2x_1x_2 - x_1$ giải hệ phương trình

$$\nabla f(x) = \begin{pmatrix} -2x_1 - 2x_2 - 1 \\ -2x_1 + 2x_2 \end{pmatrix} = 0$$

có nghiệm duy nhất $x^0 = (-1/4, -1/4)$ do $f"(x^0) = \begin{pmatrix} -2 & -2 \\ -2 & 2 \end{pmatrix}$ có định thức $det|f"(x^0)| < 0$ vậy x^0 không là cực tiểu của hàm $f(x)$.

Bài toán qui hoạch phi tuyến không ràng buộc

Các phương pháp cực tiểu một biến

Hàm đơn cực trị

Hàm đơn cực trị (unimodal function) là hàm chỉ có một điểm cực đại hay cực tiểu trên đoạn xác định.

Định nghĩa : Hàm $f(x)$ được gọi là đơn cực tiểu nếu

- $x_1 < x_2 < x^*$ kéo theo $f(x_2) < f(x_1)$
- $x_2 > x_1 > x^*$ kéo theo $f(x_1) < f(x_2)$

trong đó x^* là điểm cực tiểu.

Chú ý

Các phương pháp tìm kiếm (sẽ được giới thiệu) đều sử dụng giả thuyết là hàm đơn cực trị.

Các phương pháp cực tiểu một biến

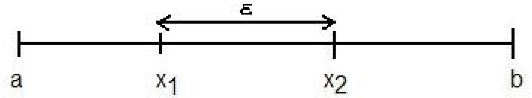
Ví dụ

Giả sử đoạn chứa cực tiểu [0,1], ta tính giá trị hàm tại hai điểm $x_1 < x_2 : f(x_1) = f_1, f(x_2) = f_2$ có 3 khả năng xảy ra :

1. $f_1 < f_2$: Điểm cực tiểu x không thể nằm bên phải x_2 vì thế đoạn chứa cực tiểu mới là $[0, x_2]$
2. $f_1 > f_2$: Điểm cực tiểu x không thể nằm bên trái x_1 vì thế đoạn chứa cực tiểu mới là $[x_1, 1]$
3. $f_1 = f_2$: Hai đoạn $[0, x_1)$ và $(x_2, 1]$ có thể loại bỏ và đoạn chứa cực tiểu mới là $[x_1, x_2]$

Các phương pháp cực tiểu một biến

Sơ đồ tìm kiếm



Giả sử ta có đoạn chứa cực tiểu xuất phát là $[a, b]$

1. Tính $x_1 = a + (b - a)/2 - e/2$ và $x_2 = a + (b - a)/2 + e/2$ với e là sai số.
2. Tính $f_1 = f(x_1)$ và $f_2 = f(x_2)$
3.
   - Nếu $f_1 < f_2$ thì đặt $b = x_2$ (loại bỏ đoạn $x > x_2$);
   - Nếu $f_1 > f_2$ thì đặt $a = x_1$ (loại bỏ đoạn $x < x_1$);
   - Nếu $f_1 = f_2$ thì đặt $a = x_1, b = x_2$ (loại bỏ đoạn $x < x_1$ và $x > x_2$);
4. Nếu $|b - a| < 2e$ thì kết thúc; trái lại quay về bước 1

Các phương pháp cực tiểu một biến

Ví dụ :

Viết đoạn kịch bản tìm cực tiểu của $f(x) = x(x - 1.5)$ trên đoạn $(a, b) = (0, 1)$ với chênh lệch $e = 0.01$ so với nghiệm đúng $x^* = 0.75$

```
f = inline('x.*(x-1.5)','x');
eps = 0.01;
a = 0; b = 1; k = 0;
while abs(b-a)>= 2*eps
        x1=a + (b-a)/2 - eps/2; x2=a + (b-a)/2 + eps/2;
        f1=f(x1); f2=f(x2); k=k+1;
        if f1<f2 b=x2;
                elseif f1>f2 a=x1;
                else a=x1;b=x2;
        end
end
...
```

Các phương pháp cực tiểu một biến

Ví dụ (tiếp) :

```
...
fprintf('So buoc lap k= %d ',k);
fprintf('Do dai doan : b-a = %d',b-a);
fprintf('x= %d',x1);
```

.

Kết quả

» So buoc lap k=7

Do dai doan : b-a = 1.773438e-002

x=7.502344e-001

Các phương pháp cực tiểu một biến

Phương pháp Fibonacci

Định nghĩa : Dãy Fibonacci được định nghĩa đệ quy như sau

$$F_0 = 1; F_1 = 1;$$
$$F_k = F_{k-1} + F_{k-2}, k \geq 2;$$

Ta phải xác định số bước lặp N trước khi thực hiện tính cực tiểu. Việc chọn hai điểm $x_1^{(k)}$ và $x_2^{(k)}$ ở bước lặp k đc xác định theo công thức :

$$x_1^{(k)} = \frac{F_{N-1-k}}{F_{N+1-k}}(b_k - a_k) + a_k, k = 0, 1, \cdots, N-1$$
$$x_2^{(k)} = \frac{F_{N-k}}{F_{N+1-k}}(b_k - a_k) + a_k, k = 0, 1, \cdots, N-1$$

Các phương pháp cực tiểu một biến

Phương pháp lát cắt vàng

Để cải tiến phương pháp Fibonacci, khi không cần cho trước số bước lặp N, ta áp luôn tỉ lệ cố định khi phân chia khoảng $b_k - a_k$

$$\lim_{N\to\infty} \frac{F_{N-1}}{F_{N+1}} = 0.382; \quad \lim_{N\to\infty} \frac{F_N}{F_{N+1}} = 0.618$$

Do đó phương pháp lát cắt vàng đề xuất chọn hai điểm thử $x_1^{(k)}$ và $x_2^{(k)}$ theo công thức cập nhật tại bước lặp như sau

$$x_1^{(k)} = 0.382(b_k - a_k) + a_k,$$
$$x_2^{(k)} = 0.618(b_k - a_k) + a_k, \quad k = 0, 1, 2, \cdots$$

Các phương pháp cực tiểu một biến

Ví dụ

Cài đặt thuật toán lát cắt vàng để tìm cực tiểu hàm trong đoạn [0,2]

$$f(x) = x^2 - 2x + 1$$

```
f = inline('x^2-2*x+1');
a = 0; b = 2; eps = 0.00001;
x1 = a + (b-a)*0.382;
x2 = a + (b-a)*0.618;
f1 = f(x1);
f2 = f(x2);
while abs(b-a)>2*eps
        if f1 > f2
                a = x1; x1 = x2; f1 = f2;
                x2 = a + (b-a)*0.618;
                f2 = f(x2);...
```

Ví dụ (tiếp)

```
        ...
    else
        b = x2; x2 = x1; f2 = f1;
        x1 = a + (b-a)*0.382;
        f1 = f(x1);
    end
end
fprintf('Khoang co hep nghiem : [%f,%f]',a,b);
»
Khoang co hep nghiem : [0.999990,1.000004]
```

Các phương pháp cực tiểu một biến

Các phương pháp số cực tiểu không ràng buộc

Mở đầu

Xét bài toán qui hoạch phi tuyến không ràng buộc

$$f(x) \to \min, x \in \mathbb{R}^n \tag{3}$$

trong đó $f(x)$ là khả vi liên tục. Để giải (3), nếu tồn tại có thể tìm được trong số các nghiệm của phương trình

$$\nabla f(x) = 0 \tag{4}$$

Tuy vậy, việc giải hệ phương trình (4) trong trường hợp tổng quát cũng không kém phần phức tạp. Dẫn đến ta phải dùng các phương pháp hiệu quả để giải (3).

Mở đầu (tiếp)

Hướng thường dùng để giải quyết (3) là dùng các phương pháp lặp từ giá trị khởi tạo x^0 rồi dịch chuyển dẫn 'về hướng' giá trị tối ưu x^*, theo mỗi bước lặp cập nhật là :

$$x^{k+1} = x^k + \alpha_k p^k, \quad k = 1, 2, \cdots \tag{5}$$

trong đó

- p^k là vec tơ định hướng dịch chuyển từ điểm x^k.
- α_k là độ dài của bước dịch chuyển theo hướng p^k.

Rõ ràng thủ tục (5) là xác định khi ta xác định được hướng dịch chuyển p^k và cách tính độ dài bước dịch chuyển α_k.

Mở đầu (tiếp)

Phụ thuộc vào các cách xây dựng p^k và α^k khác nhau mà ta có các thủ tục lặp với các đặc tính khác nhau. Ta đặc biệt quan tâm đến hai đặc tính sau :

- Sự thay đổi giá trị của hàm mục tiêu f của dãy $\{x^k\}$
- Sự hội tụ của dãy $\{x^k\}$ đến lời giải x^*.

Cũng cần chú ý là việc xác định p^k và α_k khác nhau cũng đòi hỏi khối lượng tính toán khác nhau.

Các phương pháp số cực tiểu không ràng buộc

Các phương pháp gradient

Ta chọn hướng p^k sao cho

$$< \nabla f(x^k), p^k >< 0 \tag{6}$$

Bởi vì khi chọn α^k đủ nhỏ, ta có

$$f(x^{k+1}) = f(x^k + \alpha_k) = f(x^k) + \alpha_k < \nabla f(x^k), p^k > + o(\alpha_k) < f(x^k)$$

tức là dịch chuyển theo hướng p^k với độ dài đủ nhỏ ta sẽ đến được điểm x^{k+1} với giá trị hàm mục tiêu nhỏ hơn. Vậy hướng p^k thỏa mãn (6) được gọi là hướng giảm (hướng tụt) của hàm mục tiêu $f(x)$.

Các phương pháp gradient (tiếp)

Một trong các vec tơ thỏa mãn bất đẳng thức (6) có thể chọn là vec tơ đối gradient của hàm f tại x^k :

$$p_k = -\alpha_k \nabla f(x^k), \alpha_k > 0, k = 0, 1, 2, \cdots$$

Khi đó ta có thủ tục lặp

$$x^{k+1} = x^k - \alpha_k \nabla f(x^k), \alpha_k > 0, k = 0, 1, 2, \cdots \tag{7}$$

Thử tục lặp tuân theo công thức (7) được gọi là **các phương pháp gradient**.

Các phương pháp số cực tiểu không ràng buộc

Các phương pháp gradient (tiếp)

Do hướng dịch chuyển là cố định, nên các phương pháp gradient khác nhau do cách chọn α_k. Ta liệt kê ra một số cách chọn cơ bản sau

- Thủ tục 1 : Giải bài toán cực tiểu hàm một biến

  $$\min\{\varphi_k(\lambda) : \lambda \geq 0\}, \text{ với } \varphi_k(\lambda) - f(x^k - \lambda \nabla f(x^k))$$

  Phương án tối ưu của bài toán được lấy làm giá trị của α_k.
- Thủ tục 2 : Để chọn α_k ta tiến hành theo thủ tục sau
  1. Đặt $\alpha = \overline{\alpha} > 0$
  2. Đặt $u = x^k - \alpha \nabla f(x^k)$, tính $f(u)$
  3. Kiểm tra

     $$f(u) - f(x^k) \leq -\epsilon\alpha ||\nabla f(x^k)||^2, \text{ với } 0 < \epsilon < 1 \tag{8}$$

  4. Nếu bđt (8) thỏa mãn thì đặt $\alpha_k = \alpha$, ngược lại đặt $\alpha = \alpha/2$ và quay lại bước 2.

Các phương pháp số cực tiểu không ràng buộc

Các định lý với phương pháp gradient

- Định lý 1 : Giả sử $f(x)$ bị chặn dưới và gradient của nó $f'(x)$ thỏa mãn điều kiện Lipchitz :

$$||f'(x) - f'(y)|| \leq L||x - y||$$

với mọi $x, y \in \mathbb{R}^n$, việc chọn α_k được tiến hành theo thủ tục 2. Khi đó theo công thức lặp (7) sẽ sinh ra dãy $\{x^k\}$ thỏa mãn điều kiện

$$||f'(x)|| \to 0, \text{ khi } k \to \infty$$

với mọi điểm xuất phát x^0

Các phương pháp số cực tiểu không ràng buộc

Các định lý với phương pháp gradient (tiếp)

Định lý 2 : Giả sử hàm f là hai lần khả vi liên tục và ma trận Hessian của nó thỏa mãn điều kiện

$$m||y||^2 \leq < H(x)y, y > \leq M||y||^2, M \geq m > 0$$

với mọi $x, y \in \mathbb{R}^n$, dãy $\{x^k\}$ được xây dựng theo thủ tục lặp (7) với α_k được xác định theo thủ tục 2. Khi đó với mọi điểm xuất phát x^0 ta có

$$x^k \to x^*, f(x^k) \to f(x^*), \text{ khi } k \to \infty$$

trong đó x^* là điểm cực tiểu của $f(x)$, đồng thời ta có các đánh giá sau:

$$f(x^k) - f(x^*) \leq q^k[f(x^0) - f(x^*)],$$
$$||x^k - x^*|| \leq Cq^{1/2}$$

với $0 < C < +\inf, 0 < q < 1$ là các hằng số.

Các phương pháp số cực tiểu không ràng buộc

Ví dụ :

Xét bài toán cực tiểu hóa không ràng buộc

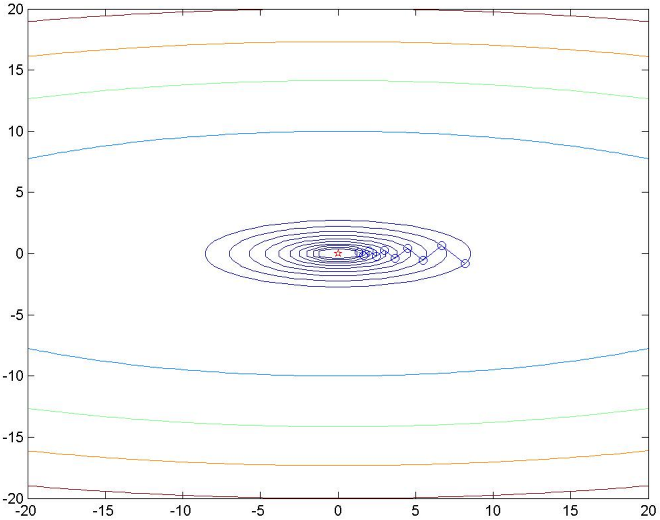
$$f(x_1, x_2) = \frac{1}{2}(x_1^2 + \gamma x_2^2) \quad (\gamma > 0)$$

Thực hiện thuật toán gradient bắt đầu từ phương án $x^0 = (\gamma, 1)$, độ dài bước xác định theo thủ tục 1, ta thu được dãy các phương án xấp xỉ $x^k = (x_1^k, x_2^k)$ trong đó

$$x_1^k = \gamma \left(\frac{\gamma - 1}{\gamma + 1}\right)^k, x_2^k = \left(-\frac{\gamma - 1}{\gamma + 1}\right)^k$$

Dãy xấp xỉ hội tụ chậm đến phương án tối ưu $x^* = (0, 0)$ khi $\gamma >> 1$ và $\gamma << 1$

Các phương pháp số cực tiểu không ràng buộc



Các phương pháp số cực tiểu không ràng buộc

Phương pháp Newton

Trong trường hợp hàm f là hai lần khả vi liên tục và việc tính toán $f'(x)$ và $f"(x)$ là không khó khăn ta có thể sử dụng đến số hạng bậc hai của khai triển Taylor.

$$f_k(x) \approx f(x^k) + < f'(x^k), x - x^k > + \frac{1}{2} < H(x^k)(x - x^k), x - x^k > \quad (9)$$

là xấp xỉ bậc hai của hàm f tại lân cận điểm x^k.

Phương pháp Newton (tiếp)

Dễ thấy là khi x^k rất gần với x^* thì $f(x^k)$ rất gần với 0 và vì thế số hạng bậc hai của (9) xấp xỉ sẽ cho ta thông tin chính xác hơn về sự biến thiên của hàm f trong lân cận của x^k.
Ta xác định vec tơ xâp xỉ phụ u^k từ điều kiện

$$f_k(u^k) = \min f_k(x) \tag{10}$$

và xây dựng phương án xấp xỉ tiếp theo

$$x^{k+1} = x^k + \alpha_k(u^k - x^k) \tag{11}$$

Vậy phụ thuộc các phương án chọn α_k mà ta có các phương pháp khác nhau. Nếu $\alpha_k = 1$, với mọi k, ta có phương pháp Newton.

Phương pháp Newton (tiếp)

Từ (11) ta có $x^{k+1} = u^k$ khi chọn $\alpha_k = 1$ thế vào điều kiện (10) trở thành

$$f_k(x^{k+1}) = \min f_k(x), k = 1, 2, \cdots \qquad (12)$$

Từ (12) ta suy ra x^{k+1} là điểm dừng của hàm $f_k(x)$, tức là

$$\nabla f_k(x) = \nabla f(x^k) + H(x^k)(x - x^k) = 0$$

Vậy nếu $H(x^k)$ không suy biến thì ta có công thức Newton sau đây

$$x^{k+1} = x^k - [H(x^k)]^{-1} \nabla f(x^k), k = 1, 2, \cdots \qquad (13)$$

Các phương pháp số cực tiểu không ràng buộc

Phương pháp Newton (tiếp)

Định lý hội tụ : Giả sử $H(x)$ là khả nghịch, $[H(x)]^{-1}$ bị chặn, $H(x)$ thỏa mãn điều kiện Liptchitz

$$||H(x) - H(y)|| \leq L||x - y||, \forall x, y \in \mathbb{R}^n$$

Khi đó dãy $\{x^k\}$ xây dựng theo (13) sẽ thỏa mãn

$$||x^{k+1} - x^*|| \leq c||x^k - x^*||^2$$

với x^* là nghiệm của phương trình $f'(x) = 0$. Vậy tốc độ hội tụ của phương pháp Newton là bình phương khoảng cách sau mỗi bước lặp.

Các phương pháp số cực tiểu không ràng buộc

Ví dụ

Minh họa giải thuật Newton cho hàm mục tiêu không ràng buộc sau

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

trong giải thuật có các hàm $f(x), gf(x)$ và $hf(x)$.

f = inline('100*(x(2)-x(1)^2)^2 + (1-x(1))^1');
gf = inline('[100*2*(x(2)-x(1)^2)*(-2*x(1)) + 2*x(1)-2;200*(x(2)-x(1)^2)]');
hf = inline('[1200*x(1)^2 - 400*x(2)+ 2, -400*x(1);-400*x(1),200]');

Các phương pháp số cực tiểu không ràng buộc

Ví dụ (tiếp)

```
fval = f(x);gval = gf(x);H = hf(x);ng = norm(gval);nf = 1;tol =
0.05;iter = 0; alpha = 1;
while ng >= tol
   iter = iter + 1;nf = 0;alpha = 1;p = -inv(H)*gval;pass = 0;
   while pass == 0
     ftest = f(x+alpha*p);
     nf = nf+1;
     if ftest <= fval + 0.01*alpha*gval'*p
        pass = 1;
        x = x+alpha*p;
        fval = ftest;
        gval = gf(x);
        H = hf(x);
        ...
```

Các phương pháp số cực tiểu không ràng buộc

Ví dụ (tiếp)

```
        ...
        ng = norm(gval);
    else
        alpha = alpha/2;
    end
  end
fprintf('%3i %3.2e %3.2e %3.2e %3.2e
%i',iter,fval,ng,norm(x-xstart),alpha,nf);
end
```

Các phương pháp số cực tiểu không ràng buộc

iter	f	||g||	||x-x*||	alpha	nf
1	2.18e+000	4.64e+000	2.21e+000	1.00e+000	1
2	2.08e+000	1.10e+001	2.06e+000	6.25e-002	5
3	2.00e+000	1.51e+001	1.93e+000	2.50e-001	3
4	1.91e+000	1.63e+001	1.83e+000	5.00e-001	2
5	1.79e+000	1.68e+001	1.72e+000	1.00e+000	1
6	1.46e+000	7.79e+000	1.65e+000	1.00e+000	1
7	1.36e+000	8.04e+000	1.59e+000	5.00e-001	2
8	1.23e+000	8.23e+000	1.50e+000	1.00e+000	1
9	9.87e-001	3.70e+000	1.40e+000	1.00e+000	1
10	9.82e-001	1.06e+001	1.23e+000	1.00e+000	1
11	6.72e-001	1.17e+000	1.12e+000	1.00e+000	1
.....					
16	1.42e-001	4.59e+000	2.82e-001	1.00e+000	1
17	9.04e-002	4.41e-001	1.96e-001	1.00e+000	1
18	2.24e-002	2.13e+000	4.88e-002	1.00e+000	1
19	9.92e-003	2.86e-002	2.22e-002	1.00e+000	1

Các phương pháp số cực tiểu không ràng buộc

Các phương pháp tìm kiếm

Trong thực tế không phải lúc nào hàm f cũng trơn và khả vi hai lần. Ngoài ra, việc tính toán cấc đạo hàm này đòi hỏi tính toán khối lượng lớn nên không thích hợp. Vậy ta cần các phương pháp chỉ cần đòi hỏi tính giá trị hàm.

$$f(x) \to \min, x \in \mathbb{R}^n \tag{14}$$

Ký hiệu

$$e^i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

là vec tơ chỉ gồm các số không và số 1 duy nhất ở hàng thứ i. Đương nhiên $e^i \in \mathbb{R}^n$, đây là vec tơ trục tọa độ đơn vị của không gian n-chiều.

Các phương pháp tìm kiếm

Giải thuật

Giả sử ta có x^0 là một xấp xỉ xuất phát. Chọn $\alpha_0 > 0$ là một thông số của thuật toán. Trong các bước lặp $k = 0, 1, 2, \cdots$ tại các xấp xỉ x^k và $\alpha_k > 0$. Đặt

$$p^k = e^{i_k}, \quad i_k = k - \left[\frac{k}{n}\right] n + 1 \tag{15}$$

trong đó $[k/n]$ là số nguyên lớn nhất không vượt quá k/n.
Tính $f(x^k + \alpha_k p^k)$ nếu

$$f(x^k + \alpha_k p^k) < f(x^k) \tag{16}$$

thỏa mãn thì gán

$$x^{k+1} = x^k + \alpha_k p^k, \quad \alpha_{k+1} = \alpha_k$$

rồi chuyển sang $k + 1$.

Các phương pháp tìm kiếm

Giải thuật (tiếp)

Nếu bđt (16) không đc thực hiện thì tính $f(x^k - \alpha_k p^k)$ và kiểm bất đẳng thức

$$f(x^k - \alpha_k p^k) < f(x^k) \tag{17}$$

thỏa mãn thì gán

$$x^{k+1} = x^k - \alpha_k p^k, \quad \alpha_{k+1} = \alpha_k$$

rồi chuyển sang $k+1$.

Một bước lặp thành công nếu một trong hai bđt (16) và (17) thỏa mãn. Ngược lại, ta xử lý như sau

$$x^{k+1} = x^k,$$

Các phương pháp tìm kiếm

Giải thuật (tiếp)

$$\alpha_{k+1} = \begin{cases} \lambda\alpha_k, & \text{nếu } i_k = n, x^k = x^{k-n+1} \\ \alpha_k, & \text{nếu } i_k \neq n, \text{ hay } x^k \neq x^{k-n+1} \end{cases}$$

trong đó $0 < \lambda < 1$ là thông số của thuật toán.

- Công thức (15) đảm bảo việc thực hiện chuyển đổi hướng dịch chuyển một cách có chu kỳ

$$p^0 = e^1, p^1 = e^2, \cdots, p^{n-1} = e^n, p^n = e^1, \cdots, p^{2n-1} = e^n, p^{2n} = e^1,$$

- Khi độ dài $\alpha_{k+1} = \lambda\alpha_k$ cần thay đổi theo công thức trên chỉ khi sau n chu kỳ liên tiếp mà ta không thành công trong việc giảm giá trị hàm mục tiêu khi thỏa mãn một trong hai bđt (16) và (17).

Các phương pháp số cực tiểu không ràng buộc

Các phương pháp tìm kiếm (tiếp)

Định lý hội tụ : Giả sử f là hàm lồi khả vi liên tục trên $\mathbb{R}^n$, còn x^0 được chọn sao cho tập

$$M(x^0) = \{x : f(x) \leq f(x^0)\}$$

là giới nội. Khi đó dãy $\{x^k\}$ xây dựng theo phương pháp mô tả ở trên là dãy cực tiểu hóa của bài toán (14), nghĩa là

$$f(x^k) \to f^*, k \to \infty$$

trong đó f^* là giá trị tối ưu của bài toán (14).

Các phương pháp số cực tiểu không ràng buộc

Bài tập về nhà :
tìm hiểu các hàm tối ưu hóa trong Matlab

- FMINBND
- FMINUNC
- FMINSEARCH
- OPTIMSET
- INLINE

Xét bài toán cực tiểu hóa không điều kiện:

$$f(x) = e^{x_1}(4x_1^2 + 2x_2^2 + 4x_1x_2 + 2x_2 + 1)$$

Điểm xuất phát $x^0 = (-1, 1)$

Xét bài toán quy hoạch phi tuyến không ràng buộc:

$$f(x_1, x_2) = (3x_1 - 9)^2 + (4x_2 - 10)^2 \rightarrow min, x = (x_1, x_2) \in \mathbb{R}^2.$$

Thực hiện giải thuật gradient giải bài toán đặt ra bắt đầu từ phương án xuất phát $x^0 = (1, 2)$. Để xác định độ dài bước hãy sử dụng thủ tục 1 (trong đó đòi hỏi giải bài toán cực tiểu hóa một chiều)

SOICT

VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG
SCHOOL OF INFORMATION AND COMMUNICATION TECHNOLOGY

Thank you for your attentions!

solct.hust.edu.vn/ fb.com/groups/solct