Tối ưu hóa (6)

TS. Đỗ Đức Đông
dongdoduc@gmail.com

Các cách tiếp cận

Tối ưu tổ hợp	Tối ưu liên tục
• Các phương pháp truyền thống	• Quy hoạch tuyến tính
Chứng minh hội tụ hoặc tỷ lệ tối ưu	• Quy hoạch phi tuyến
• Các phương pháp dựa trên thực nghiệm	+ Tìm kiếm địa phương
+ Heuristic kiến trúc (constructive heuristic)	- Phương pháp gradient ….
+ Tìm kiếm địa phương (local search)	+ Tối ưu toàn cục
+ Metaheurisics:	- Quy hoạch lồi, hiệu lồi
- GA (Genetic Algorithms)	- Tìm kiếm ngẫu nhiên (Monter-Carlo)
- ACO (Ant Colony Optimization)	- GA (Genetic Algorithms)
- Memetic algorithm	…

Bài toán tối ưu phi tuyến tổng quát dạng chính tắc (1)

Cho các hàm số f, $g_j : R^n \rightarrow R$, j = 1, 2, ..., m.

Max (Min) f(x), với các ràng buộc

$$\begin{cases} (i)\quad g_j(x) \le 0, & j = 1, 2, \ldots, k, \\ (ii)\quad g_j(x) = 0, & j = k+1, k+2, \ldots, m. \end{cases}$$

Hàm mục tiêu f(x) hoặc ít nhất một trong các hàm ràng buộc $g_j(x)$, j=1,2,…,m là phi tuyến

Ký hiệu D thuộc R^n là miền ràng buộc (miền phương án khả thi) xác định bởi ràng buộc (i) và (ii), viết gọn lại: ***f(x)→Max(Min), với x thuộc D.*** Nếu D=R^n thì bài là BTQHPT không ràng buộc.

Bài toán tối ưu phi tuyến tổng quát dạng chính tắc (2)

Xét bài toán *Max f(x)*

- Điểm $x^* \in R^n$ được gọi là điểm tối ưu toàn cục (hay phương án tối ưu toàn cục) nếu $x^* \in D$ và $f(x^*) \geq f(x), \forall x \in D$.
- Điểm $x^* \in R^n$ được gọi là điểm tối ưu địa phương(hay phương án tối ưu địa phương) nếu $x^* \in D$ và $f(x^*) \geq f(x), \forall x \in N_\varepsilon \cap D$ với N_ε là một lân cận đủ nhỏ của điểm x .
- Có nhiều phương pháp giải các lớp BTQHPT, nhưng chưa có phương pháp nào tỏ ra hữu hiệu cho mọi BTQHPT, phân ra thành hai lớp: phương pháp tất định và phương pháp Heuristic
- Phương pháp tất định sử dụng các tính chất giải tích của hàm mục tiêu và các hàm ràng buộc.
- Phương pháp Heuristic có thể áp dụng để giải các bài toán tối ưu toàn cục dạng bất kỳ, không đòi hỏi các tính chất đặc biệt của hàm mục tiêu hay các hàm ràng buộc → phương án “gần” tối ưu.

Cực trị không điều kiện

Cực trị hàm 2 biến z=f(x,y)

Xét hàm f(x,y) xác định trong miền D và điểm $M_0(x_0, y_0)$ thuộc D

- Ta nói $M_0(x_0, y_0)$ là cực tiểu (hoặc cực đại) nếu tồn tại lân cận $\mathrm{B}(\mathrm{M}_0, \varepsilon)$ của M_0 sao cho: $f(x_0, y_0) \leq f(x, y)\ \forall (x, y) \in \mathrm{B}(\mathrm{M}_0, \varepsilon)$

  (Cực đại $f(x_0, y_0) \geq f(x, y)\ \forall (x, y) \in \mathrm{B}(\mathrm{M}_0, \varepsilon)$)

- Hàm số z=f(x,y) đạt cực tiểu (cực đại) tại $M_0(x_0, y_0)$ nếu: $\Delta f(x_0, y_0) = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) \geq 0\ (\leq 0)\quad \forall \Delta x, \Delta y$
- Nếu $\Delta f(x_0, y_0)$ đổi dấu khi $\Delta x, \Delta y$ thay đổi thì hàm số không đạt cực trị tại $M_0(x_0, y_0)$.

Cực trị không điều kiện

Cực trị hàm 2 biến z=f(x,y)

Ví dụ, xét hàm số $z=x^3 + y^3$ có đạt cực trị tại M(0,0) hay không?

Xét $\mathrm{N}(0+\Delta x, 0+\Delta y)$ là một điểm trong lân cận M(0,0), ta có:

$$\Delta f(0,0) = f(\Delta x, \Delta y) - f(0,0) = (\Delta x)^3 + (\Delta y)^3$$

Với $\Delta x > 0, \Delta y > 0$ thì $\Delta f(0,0) > 0$

Với $\Delta x < 0, \Delta y < 0$ thì $\Delta f(0,0) < 0$

Vậy $\Delta f(0,0)$ thay đổi dấu nên hàm f không đạt cực trị tại điểm M(0,0).

Cực trị không điều kiện

Cực trị hàm 2 biến z=f(x,y)

Điều kiện cần

Nếu hàm f(x,y) đạt cực trị địa phương tại $M_0(x_0, y_0)$ và nếu f có đạo hàm riêng tại M_0 thì:

$$\frac{\partial f}{\partial x}(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0) = 0$$

Điều kiện đủ

Giả sử z=f(x,y) có đạo hàm riêng cấp hai liên tục trong lân cận điểm dừng $M_0(x_0, y_0)$, đặt

$A = \frac{\partial^2 f}{\partial x^2}(x_0, y_0); B = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0);\ \ C = \frac{\partial^2 f}{\partial y^2}(x_0, y_0);$

- Nếu $AC - B^2 > 0$ và A>0 (hay C>0) thì f đạt cực tiểu tại M_0
- Nếu $AC - B^2 > 0$ và A<0 (hay C<0) thì f đạt cực đại tại M_0
- Nếu $AC - B^2 < 0$ thì f không đạt cực trị tại M_0

Cực trị không điều kiện

- Xét bài toán cực trị hàm nhiều biến: Max(min) $\{f(x)/x=(x_1, x_2,\ldots, x_n) \in R^n\}$ với giả thiết f có các đạo hàm riêng cấp một và cấp hai liên tục.
- Ký hiệu $H_k(x) = |\frac{\partial^2 f(x)}{\partial x_i \partial x_j}|$ với $i \le k; j \le k$ là định thức của ma trận các đạo hàm riêng cấp hai của f tại x theo k biến đầu tiên.
- Nếu x là cực trị của bài toán thì $\frac{\partial f(x)}{\partial x_i} = 0\ \forall i \le n$ (các điểm thỏa mãn điều kiện này gọi là điểm dừng). Điều kiện đủ để một điểm dừng là điểm cực trị địa phương như sau: Giả sử x^* là điểm dừng của f và $H_k(x^*) \ne 0$ với mọi $k \le n$ thì:
    1) x^* là điểm cực tiểu của f nếu $H_k(x^*) > 0$ với mọi $k \le n$
    2) x^* là điểm cực đại của f nếu $sign(H_k(x^*)) = (-1)^k$ với mọi $k \le n$
    3) Nếu không thỏa mãn 1) và 2) thì x^* không phải là điểm cực trị (điểm yên ngựa)

Cực trị không điều kiện

Ví dụ 1

Tìm cực trị của hàm số f(x,y)=$x^2y + xy^3 - xy$.

Ta có : $\dfrac{\partial f}{\partial x} = 2xy + y^3 - y$

$$\dfrac{\partial f}{\partial y} = x^2 + 3xy^2 - x \,.$$

Vậy (x,y) là điểm dừng của f nếu nó là nghiệm của hệ :

$$\begin{cases} 2xy + y^3 - y = 0 \\ x^2 + 3xy^2 - x = 0 \end{cases}$$

Hệ trên có các nghiệm sau : (0,0) ; (1,0) ; (0,1) ; (0,-1) ; $\left(\dfrac{2}{5}, \dfrac{\sqrt{5}}{5}\right)$ và $\left(\dfrac{2}{5}, -\dfrac{\sqrt{5}}{5}\right)$.

Cực trị không điều kiện
Ví dụ 2

Tìm cực trị của hàm số

$$z = x^3 + y^3 - 3xy$$

Cực trị có điều kiện (1)

Xét bài toán tìm cực trị hàm nhiều biến:

$Max(\min)\ \{f(x)x = (x_1, x_2, \dots, x_n) \in R^n$ với điều kiện $g_i(x) = 0$.

Trong đó f, g_i là các hàm có đạo hàm riêng cấp một và cấp hai liên tục, ta thiết lập hàm Lagrange của bài toán:

$L(x_1, x_2, \dots, x_n, \lambda_1, \lambda_2, \dots, \lambda_n) = f(x) + \sum_{i=1}^{n} \lambda_i g_i(x)$

Các thừa số $\lambda_1, \lambda_2, \dots, \lambda_n$ được gọi là nhân tử Lagrange.

Điều kiện cần với điểm cực trị địa phương:

Nếu x^* là điểm cực trị địa phương của hàm f thì tồn tại $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$ để (x^*, λ) là điểm dừng của hàm Lagrange.

Chú ý rằng: $\frac{\partial L(x,\lambda)}{\partial \lambda_i} = g_i(x)$ nên để (x^*, λ) là điểm dừng của hàm Lagrange là tồn tại λ sao cho x^* thỏa mãn $g_i(x) = 0$ và $\frac{\partial L(x,\lambda)}{\partial x_i} = 0$.

Cực trị có điều kiện (2)

Khi f là hàm lồi hay lõm và g_i là các hàm tuyến tính thi ta có các cực trị toàn cục. Giả sử ***(x,λ)*** là điểm dừng của hàm Lagrange và g_i là các hàm tuyến tính với mọi i= 1,..,m thì các kết luận sau đúng:

- ***x*** là điểm cực tiểu toàn cục nếu *f* là hàm lồi.
- ***x*** là điểm cực đại toàn cục nếu *f* là hàm lõm.

Phương pháp Gradient (đường dốc nhất)

Phương pháp Gradient là phương pháp phổ thông, đơn giản và dễ ứng dụng giải bài toán phi tuyến tính không ràng buộc.

Xét bài toán cực trị không điều kiện: $\min\{f(x) | x \in R^n\}$

d là vector hướng giảm của hàm f tại x nếu $\exists\delta > 0$ sao cho $f(x + \lambda d) < f(x), \forall\lambda \in (0, \delta)$

Bước 1: Chon trước $\alpha > 0$ và ε $(0 < \varepsilon < 1)$, lấy xấp xỉ ban đầu $x_0 \in R^n$ tùy ý

Bước 2: Xây dựng dãy $x_k \in R^n; k = 1,2, \dots,$

$$x_{k+1} = x_k - \alpha_k f'(x_k) \text{ với } \alpha_k > 0$$

Trong đó α_k được chọn như sau: Nếu $f\big(x_k - \alpha f'(x)\big) < f(x_k) - \varepsilon \|f'(x_k)\|^2$ thì $a_k = a$, ngược lại giảm α (chẳng hạn nhân α với $\lambda \in (0,1)$), lặp cho tới khi thỏa mãn.

Thuật toán kết thúc khi $\|f'(x_k)\|^2$ đủ bé.

Phương pháp Gradient (đường dốc nhất)

$$\min \mathrm{f(x)} = \mathrm{x}_1^2 + 25\mathrm{x}_2^2$$

Chọn $x_0 = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, \alpha = 0.01$

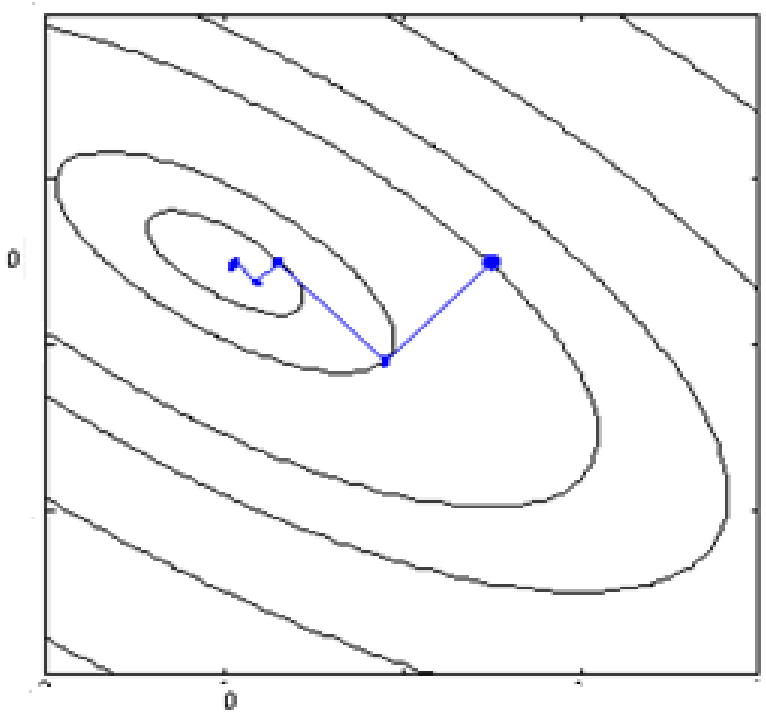
$f'(x) = \begin{bmatrix} 2x_1 \\ 50x_2 \end{bmatrix};\ f'(x_0) = \begin{bmatrix} 1 \\ 25 \end{bmatrix}$

$x_1 = x_0 - \alpha f'(x_0) = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} - 0.01 \begin{bmatrix} 1 \\ 25 \end{bmatrix} = \begin{bmatrix} 0.49 \\ 0.25 \end{bmatrix};$

…

Phương pháp Gradient (đường dốc nhất)

Nhận xét: Phương pháp phụ thuộc vào tốc độ học α. Nếu chọn α nhỏ $f'(x)$ gần như trực giao với đường mức $f(x)$, thuật toán hội tụ chậm nhưng khi đến gần điểm cực tiểu thì không dao động.



Phương pháp hàm phạt

Xét bài toán cực trị có điều kiện

$\min\{f(x)|x \in X\}\,; X \in R^n$

X được xác định bởi các điều kiện:

1) $g_j(x) \leq 0\ (j = 1,2, \dots, s)$

2) $g_j(x) = 0\ (j = s + 1, \dots, m)$

Tư tưởng của phương pháp hàm phạt là thay thế việc giải bài toán có điều kiện bởi bài toán cực trị không điều kiện phụ thuộc tham số

$M(x, \beta) = f(x) + \frac{1}{\beta}\emptyset(x); \beta > 0$

Hàm $\emptyset(x)$ được chọn sao cho lời giải của bài toán $\min\{M(x, \beta)/x \in G\} = m(\beta)$

Phương pháp monte-carlo (1)

Bài toán quy hoạch mà hàm mục tiêu phức tạp, không cho được dưới dạng hiển trên một miền giới nội D nào đó thì các phương pháp đã nêu không dùng được. Khi đó phương pháp monte-carlo là một phương pháp có hiệu quả.

$\max\{f(x)|x \in D\}\,; D \in R^n$

Trong đó f là hàm liên tục, D là miền giới nội trong R^n: $D \subset \prod_{i=1}^{n}[a_i, b_i]$

Phương pháp monte-carlo (2)

Tạo một tập đủ lớn n véc tơ ngẫu nhiên có phân bố đều trên D và chọn véc tơ có hàm mục tiêu lớn nhất để làm lời giải gần đúng. Với số bước lặp N cho trước, thuật toán thực hiện như sau.

- *Bước 1*. Khởi tạo $j = 0$; $f = m$ đủ nhỏ.
- *Bước 2*. Với mỗi $i = 1,..,n$ tạo số ngẫu nhiên $r_i \in [0,1]$, tính $y_i = a_i + ri_{(b_i - a_i)}$ và xác định $y = (y_1, \dots, y_n)$.
- *Bước 3*. Kiểm tra nếu y thuộc D thì tăng j:=j+1 và sang bước 4, nếu không thuộc D thì trở lại bước 2.
- *Bước 4.* Tính $f(y)$, nếu $f(y) > f$ thì gán $x = y$ và $f = f(y)$
- *Bước 5*. Kiểm tra điều kiện kết thúc $j = N$, nếu đúng thi in kết quả x và f tương ứng, chưa đúng thì trở lại bước 2.

Phương pháp monte-carlo (3)

- Nếu ta chạy lai thuật toán thì giá tri khởi tạo của f có thể lấy kết quả của lần chạy trước .
- Điều kiện kết thúc có thể thay việc đếm số lần lặp bởi điều kiện khác.
- Kết quả của mỗi lần chạy không giống nhau.
- **Sự hội tụ**. Người ta chứng minh được khi N dần ra vô hạn thì f hội tụ theo xác suất tới giá trị tối ưu nhưng không co đánh giá sai số cụ thể.

Tối ưu đa mục tiêu (1)

Bài toán tối ưu đa mục tiêu tổng quát có thể xét dưới dạng sau

Cực đại hoá các hàm $f_i \rightarrow \max(i = 1,2, \dots , k)$ với $x \in X \subset R^n$

Nói chung không có lời giải đồng thời đạt cực đại của cả k hàm f_i. Lời giải của nó được tìm theo nghĩa tối ưu pareto như sau: Điểm $x^* \in X$ gọi là tối ưu pareto của bài toán đa mục tiêu trên tập X nếu không tồn tại điểm $y \in X$ sao cho có ít nhất $i \leq k$ mà

$f_i(y) > f_i(x^*)$

và $f_i(y) \geq f_i(x^*)$ với $\forall j \neq i$

Tối ưu đa mục tiêu (2)

Cách 1: Đưa các mục tiêu thứ yếu vào điều kiện buộc

Theo cách này, ta chọn hàm mục tiêu f_j mà ta cho là quan trọng nhất và xét bài toán $f_j(x) \rightarrow max$

với các điều kiện $\begin{cases} f_i(x) \geq c, \forall i \neq j \\ x \in X \end{cases}$

Cách 2: Chọn trọng số ưu tiên

Chọn $p_1, p_2, \dots, p_k$ tương ứng với k hàm mục tiêu sao cho $p_i > 0$ và $p_1 + p_2 + \cdots + p_k = 1$ (độ lớn của ρ_i phụ thuộc vào mức độ quan trọng của hàm mục tiêu f_i)

Đi giải bài toán $\max\left\{ \sum_{i=1}^{k} \rho_i f_i(x) \middle/ x \in X \right\}$