1. Linh tinh

1.1. Chuẩn Euclid của vector

$$\|x\| = \sqrt{\langle \text{x, x} \rangle} = \sqrt{\sum_{i=1}^{n} |x|^2}$$

1.2. Bất đẳng thức Cauchy - Bunjakowski - Schwarz

$$|\langle x, y \rangle| \leq \|x\|\|y\|$$

1.3. Góc giữa 2 vector

$$\cos\alpha = \frac{\langle x, y \rangle}{\|x\|\|y\|}$$

1.4. Tích vô hướng

$$\langle x, y \rangle = \begin{cases} x_1 y_1 + ... + x_n y_n \\ \|x\|\|y\| \cos\left(\widehat{\text{x, y}}\right) \end{cases}$$

Một số tính chất:

$$\begin{cases} \langle x, y \rangle = \langle y, x \rangle \\ \langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle \\ \langle \lambda x, z \rangle = \lambda \langle x, z \rangle \\ \langle x, x \rangle \geq 0 \text{ và } \langle x, x \rangle = 0 \text{ khi và chỉ khi } x = 0 \end{cases}$$

1.5. Gradient

$$\nabla f(x^0) = \left(\frac{\partial f(x^0)}{\partial x_1}, \frac{\partial f(x^0)}{\partial x_2}, ..., \frac{\partial f(x^0)}{\partial x_n} \right)^T$$

1.6. Ma trận Hesse

$$\nabla^2 f(x^0) = \begin{pmatrix} \frac{\partial^2 f(x^0)}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 f(x^0)}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x^0)}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(x^0)}{\partial x_n \partial x_n} \end{pmatrix}$$

1.7. Đoạn thẳng nối 2 điểm

Cho x^1, $x^2 \in \mathbb{R}$. Đoạn thẳng nối 2 điểm là tập hợp các điểm có dạng

$$x = \lambda x^1 + (1 - \lambda) x^2, \quad 0 \leq \lambda \leq 1$$

1.8. Hàm tuyến tính

$f(x)$ là tuyến tính nếu

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$$

Một hàm tuyến tính xác định trên $\mathbb{R}^n$ luôn có dạng $f(x) = \langle c, x \rangle$, với $c \in \mathbb{R}^n$

1.9. Hàm affine

Hàm affine có dạng

$$f(x) = \langle c, x \rangle + \alpha, \quad c \in \mathbb{R}^n, \alpha \in \mathbb{R}$$

Nếu $f(x)$ là affine thì $\forall x, y \in \mathbb{R}^n$, $\forall \lambda, \mu \in \mathbb{R}$ mà $\lambda + \mu = 1$, ta có

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$$

2. Giữa kì

2.1. Nêu định nghĩa tập affine, lồi, hàm lồi và các tính chất cơ bản

2.1.1. Tập affine

Tập $M \in \mathbb{R}^n$ được gọi là tập affine nếu chứa trọn cả **đường thẳng đi qua 2 điểm bất kì** của M, tức:

$$\forall x, y \in \mathbb{R}^n, \lambda \in \mathbb{R} \Rightarrow \lambda x + (1 - \lambda) y \in M$$

Tổ hợp affine:

$$x = \sum_{i=1}^{k} \lambda_i x_i, \text{với } \lambda_1, ..., \lambda_k \in \mathbb{R}, \text{ và } \sum_{i=1}^{k} \lambda_i = 1$$

2.1.2. Tập lồi

Tập $M \in \mathbb{R}^n$ được gọi là tập lồi nếu nó chứa trọn **đoạn thẳng nối 2 điểm bất kì** thuộc nó, tức $\forall x, y$ và $0 \leq \lambda \leq 1$, ta có

$$\lambda x + (1 - \lambda) y \in M$$

Tổ hợp lồi:

$$x = \sum_{i=1}^{k} \lambda_i x_i, \text{với } \lambda_1, ..., \lambda_k \geq 0, \text{ và } \sum_{i=1}^{k} \lambda_i = 1$$

Nếu $\lambda_i > 0$ thì x là tổ hợp lồi chặt

- Một tập M là lồi khi và chỉ khi nó chứa tất cả các tổ hợp lồi của những phần tử thuộc nó
- Nếu M là tập lồi thì αM cũng là tập lồi
- Nếu M_1, M_2 lồi thì $M_1 + M_2$ cũng là lồi

Định lý tách:

- Định lý tách I: Nếu 2 tập lồi không rỗng và rời nhau thì có 1 siêu phẳng tách chúng
- Định lý tách II: Nếu 2 tập lồi không rỗng và rời nhau và 1 trong 2 tập ấy là compact thì có 1 siêu phẳng tách chúng

2.1.3. Hàm lồi

f xác định trên tập lồi $X \in \mathbb{R}^n$ được gọi là lồi nếu

$$f(\lambda x + (1 - \lambda) y) \leq \lambda f(x) + (1 - \lambda) f(y), \quad \forall x, y \in X, \ \lambda \in [0, 1]$$

Hàm f được gọi là lõm nếu $-f$ là lồi. Ta nói f là lồi chặt nếu

$$f(\lambda x + (1 - \lambda) y) < \lambda f(x) + (1 - \lambda) f(y), \quad \forall x, y \in X, \ \lambda \in [0, 1]$$

- Miền xác định hữu hiệu của hàm lồi f là

$$\mathrm{dom} f = \{x \in X \mid f(x) < \infty\}$$

- Trên đồ thị của hàm lồi f là tập

$$\mathrm{epi} f = \{(x, \psi) \in X \times \mathbb{R} \mid f(x) \leq \psi\}$$

- f lồi $\Leftrightarrow$ $\mathrm{epi} f$ là tập lồi
- Nếu f lồi thì tập mức dưới $L_\alpha f = \{x \in X \mid f(x) \leq \alpha\}$ là tập lồi, $\forall \alpha \in \mathbb{R}$
- Nếu f_1 lồi trên X_1, f_2 lồi trên X_2 và $\lambda, \mu > 0$ thì các hàm $\lambda f_1 + \mu f_2$, $\max\{f_1, f_2\}$ lồi trên $X_1 \cap X_2$
- Nếu f là hàm lồi xác định trên tập lồi mở $X \subseteq \mathbb{R}^n$ thì f liên tục trên X
- $$f'(x^0, d) \leq f(x^0 + d) - f(x^0)$$
- Cho f là hàm khả vi trên tập lồi mở $X \subseteq \mathbb{R}^n$. Khi đó, f là hàm lồi trên X khi và chỉ khi

$$f(y) - f(x) \geq \langle \nabla f(x), y - x \rangle, \ \forall x, y \in X$$

- Nếu f là hàm khả vi 2 lần trên tập lồi mở $X \in \mathbb{R}^n$, khi đó:
  - Hàm f lồi trên X khi và chỉ khi ma trận Hesse $\nabla^2 f(x)$ là **nửa xác định dương** trên X, tức là $\forall x \in X$

$$y^T \nabla^2 f(x) y \geq 0, \quad \forall y \in \mathbb{R}^n$$

  lồi chặt thì ma trận Hesse **xác định dương** và

$$y^T \nabla^2 f(x) y > 0, \quad \forall y \in \mathbb{R}^n$$

  lõm thì ngược lại $(\leq, <)$
- Giả sử f là hàm lồi khả vi trên $\mathbb{R}^n$. Khi đó $x^* \in \mathbb{R}^n$ là nghiệm cực tiểu toàn cục của bài toán khi và chỉ khi

$$\nabla f(x^*) = 0$$

2.2. Với dữ liệu đã cho, phát biểu mô hình bài toán tối ưu

Bài toán tối ưu tổng quát được phát biểu như sau:

$$\min f(x), \text{ v.đ.k } x \in D$$

Trong đó $D \subseteq \mathbb{R}^n$ được gọi là tập nghiệm chấp nhận được hay tập ràng buộc và $f : D \to \mathbb{R}$ là hàm mục tiêu

2.3. Phát biểu điều kiện cần và điều kiện đủ của sự tồn tại điểm cực tiểu của bài toán khả vi không ràng buộc

2.3.1. Điều kiện cần (Điều kiện bậc nhất)

Cho hàm f xác định, khả vi trên $\mathbb{R}^n$. Nếu $x^* \in \mathbb{R}^n$ là nghiệm cực tiểu địa phương của bài toán (P^{krb}) thì $\nabla f(x^*) = 0$.

Điểm $x^* \in \mathbb{R}^n$ thoả $\nabla f(x^*) = 0$ được gọi là **điểm dừng** của hàm f.

2.3.2. Điều kiện đủ (Điều kiện bậc hai)

Giả sử hàm f khả vi liên tục 2 lần trên $\mathbb{R}^n$. Khi đó:
- Nếu $x^* \in \mathbb{R}^n$ là **điểm cực tiểu địa phương** của f trên $\mathbb{R}^n$ thì

$$\begin{cases} \nabla f(x^*) = 0 \text{ và} \\ \nabla^2 f(x^*) \text{ nửa xác định dương} \end{cases}$$

- Ngược lại, nếu

$$\begin{cases} \nabla f(x^*) = 0 \text{ và} \\ \nabla^2 f(x^*) \text{ xác định dương} \end{cases}$$

thì x^* là **điểm cực tiểu địa phương chặt** của f trên $\mathbb{R}^n$

2.4. Trình bày thuật toán gradient với thủ tục tìm chính xác theo tia và thuật toán gradient với thủ tục quay lui

2.4.1. Thuật toán gradient với thủ tục tìm chính xác theo tia

Trong thuật toán này, tại mỗi bước lặp k, điểm lặp tiếp theo được xác định bởi

$$x^{k+1} = x^k - t_k \nabla f(x^k)$$

trong đó t_k là nghiệm cực tiểu của hàm một biến $\varphi_k(t) = f(x^k - t\nabla f(x^k))$ với $t > 0$

- Bước khởi đầu: chọn $\varepsilon > 0$ đủ nhỏ, xuất phát từ 1 điểm $x^0 \in \mathbb{R}^n$ tuỳ ý có $\nabla f(x^0) \neq 0$, đặt $k := 0$
- Bước lặp $k(k = 0, 1, 2, ...)$:
  - (k_1): Tính $x^{k+1} = x^k - t_k \nabla f(x^k)$, với $t_k = \text{argmin}\ \{\varphi_k(x),\ t > 0\}$
  - (k_2): Tính $\nabla f(x^{k+1})$
  - (k_3): **If** $\|\nabla f(x^{k+1})\| < \varepsilon$
    Then Dừng thuật toán (lấy điểm dừng $x^* \approx x^{k+1}$)
    Else $k := k + 1$ và quay lại Bước lặp k

Nếu hàm mục tiêu của bài toán là hàm toàn phương lồi chặt

$$f(x) = \frac{1}{2}x^T Ax - b^T x + c$$

thì ta có công thức tính độ dài bước chính xác t_k tại mỗi Bước lặp k là

$$t_k = \frac{(Ax^k - b)^T \nabla f(x^k)}{(\nabla f(x^k))^T A \nabla f(x^k)} > 0$$

2.4.2. Thuật toán gradient với thủ tục quay lui

Trong thuật toán này, tại mỗi bước lặp k, chọn hướng giảm $d^k = -\nabla f(x^k)$ và độ dài bước t_k được xác định theo thủ tục quay lui

- Bước khởi đầu: chọn tuỳ ý $m_1 \in (0, 1)$ và $\alpha \in (0, 1)$, chọn $\varepsilon > 0$ đủ nhỏ, xuất phát từ 1 điểm $x^0 \in \mathbb{R}^n$ tuỳ ý có $\nabla f(x^0) \neq 0$, đặt $k := 0$
- Bước lặp $k(k = 0, 1, 2, ...)$:
  - (k_1): Đặt $t_k = 1$
  - (k_2): Tính $x^{k+1} = x^k - t_k \nabla f(x^k)$ và $f(x^{k+1})$
  - (k_3): **If** $f(x^{k+1}) - f(x^k) \leq m_1 t_k \langle \nabla f(x^k), -\nabla f(x^k) \rangle = -m_1 t_k \|\nabla f(x^k)\|^2$
    Then Chuyển Bước k_4
    Else $t_k := \alpha t_k$ và quay về Bước k_2
  - (k_4): Tính $\nabla f(x^{k+1})$
  - (k_5): **If** $\|\nabla f(x^{k+1})\| < \varepsilon$
    Then Dừng thuật toán (lấy điểm dừng $x^* \approx x^{k+1}$)
    Else $k := k + 1$ và quay lại Bước lặp k

3. Cuối kì

3.1. Nêu định nghĩa tập affine, lồi, hàm lồi và các tính chất cơ bản (Section 2.1)

3.2. Thuật toán gradient với thủ tục tìm chính xác theo tia và thuật toán gradient với thủ tục quay lui (Section 2.4)

3.3. Phương pháp Newton cổ điển giải hệ phương trình phi tuyến

3.3.1. Trường hợp $n = 1$

Xét phương trình 1 biến số

$$f(x) = 0$$

Giả sử nghiệm của phương trình này là $x^* \in \mathbb{R}$. Xuất phát từ điểm x^0 đủ gần x^* và sinh ra 1 dãy nghiệm xấp sỉ $x^0, x^1, x^2, ...$ hội tụ đến x^*

Đặt

$$x^{k+1} = x^k - \frac{f(x^k)}{f'(x^k)}$$

Gán $k := k + 1$ và lặp lại quá trình tính toán đối với điểm x^k mới

- **Giả sử hàm f khả vi liên tục cấp 2**
- **x^* là nghiệm của phương trình $f(x) = 0$, tức $f(x^*) = 0$**
- **$f'(x^k) \neq 0$**
- **Điểm xuất phát ban đầu x^0 phải gần phải đủ gần nghiệm x^* của hệ**. Nếu không, thuật toán có thể không hội tụ

3.3.2. Trường hợp $n > 1$

Ma trận Jacobi của F tại x:

$$DF(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

Dòng thứ i của ma trận Jacobi chính là $[\nabla f_i(x)]^T$

Xét hệ phương trình n ẩn, n phương trình

$$F(x) = 0$$

trong đó $F(x) = (f_1(x), f_2(x), ..., f_n(x))^T$ là hàm vector

Giả sử nghiệm của phương trình này là $x^* \in \mathbb{R}$. Xuất phát từ điểm x^0 đủ gần x^* và sinh ra 1 dãy nghiệm xấp sỉ $x^0, x^1, x^2, ...$ hội tụ đến x^*

Bước lặp:

$$x^{k+1} = x^k - [DF(x^k)]^{-1} F(x^k)$$

Đặt $x^{k+1} = x^k$ và lặp lại quá trình tính toán đối với điểm x^k mới

- **Điểm xuất phát ban đầu x^0 phải gần phải đủ gần nghiệm x^* của hệ**
- **Ma trận Jacobi $DF(x^k)$ không suy biến tại mọi bước lặp k**. Nếu không, thuật toán sẽ không thực hiện được

3.4. Thuật toán Newton thuần tuý giải bài toán tối ưu không ràng buộc

- Bước khởi đầu: chọn $\varepsilon > 0$ đủ nhỏ, xuất phát từ 1 điểm $x^0 \in \mathbb{R}^n$ tuỳ ý đủ gần điểm dừng x^* và $\nabla f(x^0) \neq 0$, đặt $k := 0$
- Bước lặp $k(k = 0, 1, 2, ...)$:
  - (k_1): Tính hướng Newton p^k của f tại x^k bằng việc giải hệ phương trình tuyến tính
  $$[\nabla^2 f(x^k)]p^k = -\nabla f(x^k)$$
  - (k_2): Xác định $x^{k+1} := x^k + p^k$ và $\nabla f(x^{k+1})$
  - (k_3): **If** $\|\nabla f(x^{k+1})\| < \varepsilon$
    Then Dừng thuật toán (lấy điểm dừng $x^* \approx x^{k+1}$)
    Else $k := k + 1$ và quay lại Bước lặp k

3.5. Thuật toán chia đôi

- Bước khởi đầu: lấy $\varepsilon > 0$ đủ nhỏ, đặt $a^1 := a,\ b^1 := b,\ k := 1$
- Bước lặp $k(k = 1, 2, ...)$:
  - (k_1): Đặt $c := \frac{a^k+b^k}{2}$, $x^k = c - \frac{\varepsilon}{2}$, $y^k = c + \frac{\varepsilon}{2}$ (có $a^k < x^k < y^k < b^k$)
  - (k_2): Tính $z_1 = f(x^k)$, $z_2 = f(y^k)$
  - (k_3): **If** $z_1 \leq z_2$
    Then Chuyển Bước k_4
    Else Chuyển Bước k_5
  - (k_4): (Có $x^* \in [a^k, y^k]$)
    If $y^k - a^k \leq \varepsilon$
    Then Dừng thuật toán (lấy $x^* := x^k$ và $f_* := z_1$, ở đây f_* là giá trị tối ưu)
    Else Đặt $a^{k+1} = a^k$, $b^{k+1} = y^k$, $k := k + 1$, chuyển về Bước lặp k
  - (k_5): (Có $x^* \in [x^k, b^k]$)
    If $b^k - x^k \leq \varepsilon$
    Then Dừng thuật toán (lấy $x^* := y^k$ và $f_* := z_2$)
    Else Đặt $a^{k+1} = x^k$, $b^{k+1} = b^k$, $k := k + 1$, chuyển về Bước lặp k

3.6. Thuật toán lát cắt vàng

- Bước khởi đầu: lấy $\varepsilon > 0$ đủ nhỏ, đặt $a^1 := a,\ b^1 := b,\ k := 1$ và $\alpha = \frac{\sqrt{5}-1}{2}$
- Bước lặp $k(k = 1, 2, ...)$:
  - (k_1): Chia $[a^k, b^k]$ bởi các điểm chia $x^k := a^k + (1-\alpha)(b^k - a^k)$, $y^k := a^k + \alpha(b^k - a^k)$
  - (k_2): Tính $z_1 = f(x^k)$, $z_2 = f(y^k)$
  - (k_3): **If** $z_1 \leq z_2$
    Then Chuyển Bước k_4
    Else Chuyển Bước k_5
  - (k_4): (Có $x^* \in [a^k, y^k]$)
    If $y^k - a^k \leq \varepsilon$
    Then Dừng thuật toán (lấy $x^* := x^k$ và $f_* := z_1$)
    Else Đặt $a^{k+1} = a^k$, $b^{k+1} = y^k$, $k := k + 1$, chuyển về Bước lặp k
  - (k_5): (Có $x^* \in [x^k, b^k]$)
    If $b^k - x^k \leq \varepsilon$
    Then Dừng thuật toán (lấy $x^* := y^k$ và $f_* := z_2$)
    Else Đặt $a^{k+1} = x^k$, $b^{k+1} = b^k$, $k := k + 1$, chuyển về Bước lặp k