

MLCheatsheet

1. Bài toán Gốc: Toán học (Theory)

1. Thiết lập:

- $x, w \in \mathbb{R}^{M-1}$; Siêu phẳng: $w^T x + b = 0$.
- Max Margin $\Leftrightarrow \text{Min } \|w\| \Leftrightarrow \text{Min } \frac{1}{2} \|w\|^2$.

3. Công thức Primal:

$$\begin{aligned} & \min_{w,b} \frac{1}{2} \|w\|^2 \\ \text{s.t. } & t_n(w^T x_n + b) \geq 1, \forall n \end{aligned}$$

2. Bài toán Gốc: Implementation (CVXOPT)

Solver: $\min \frac{1}{2} u^T P u + q^T u$ s.t. $Gu \leq h$. Biến: $u = [\mathbf{w}; \mathbf{b}]_{(M \times 1)}$.

Mapping:

$$\begin{aligned} P &= \text{diag}(\mathbf{1}, \dots, \mathbf{1}, \mathbf{0})_{(M \times M)} \\ q &= \mathbf{0}_{(M \times 1)} \\ G &= -[\text{diag}(\mathbf{t}) \cdot \mathbf{X}, \quad \mathbf{t}]_{(N \times M)} \\ h &= -\mathbf{1}_{(N \times 1)} \end{aligned}$$

3. Lagrangian & Điều kiện KKT

1. Hàm Lagrangian:

$$\begin{aligned} \mathcal{L} &= \frac{1}{2} \|w\|^2 + \sum_{n=1}^N \underbrace{\alpha_n}_{\substack{\geq 0 \\ \text{dùng: } \leq 0 \\ \text{sai: } > 0}} \{1 - t_n(w^T x_n + b)\} \\ &= \frac{1}{2} \|w\|^2 + E_{\text{data}} \text{ (cost trên dữ liệu)} \end{aligned}$$

Tính chất: Với $\alpha_n \geq 0$, (w^*, b^*) tối ưu:

$$\mathcal{L}(w^*, b^*, \alpha) \leq \frac{1}{2} \|w^*\|^2$$

2. Tính đạo hàm:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial w} &= w - \sum_{n=1}^N \alpha_n t_n x_n \\ \frac{\partial \mathcal{L}}{\partial b} &= - \sum_{n=1}^N \alpha_n t_n \end{aligned}$$

3. Điều kiện dừng (KKT-1):

- $\nabla_w \mathcal{L} = 0 \Rightarrow w = \sum_{n=1}^N \alpha_n t_n x_n$ (1)

- $\nabla_b \mathcal{L} = 0 \Rightarrow \sum_{n=1}^N \alpha_n t_n = 0$ (2)

4. Các điều kiện KKT khác:

- (KKT-2) Ràng buộc gốc: $1 - t_n(w^T x_n + b) \leq 0$
- (KKT-3) Ràng buộc dual: $\alpha_n \geq 0, \forall n$
- (KKT-4) Điều kiện bù: $\alpha_n \{1 - t_n(w^T x_n + b)\} = 0$

Ý nghĩa: Thỏa KKT $\Rightarrow (w, b, \alpha) = (w^*, b^*, \alpha^*)$

5. Tiêu chuẩn Slater (cho SVM):

Nếu phần trong của tập khả thi không rỗng thì:

$$\exists (w, b) \text{ s.t. } 1 - t_i(w^T x_i + b) < 0, \forall i$$

$$\Rightarrow \text{strong duality } (p^* = d^*) \Leftrightarrow \text{duality gap} = 0$$

4. Bài toán Dual: Biến đổi (Math)

Thay (1), (2) vào \mathcal{L} :

A. **Thay w vào $\frac{1}{2} w^T w$:**

$$\frac{1}{2} \left(\sum \alpha_i t_i x_i \right)^T \left(\sum \alpha_j t_j x_j \right) = \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j t_i t_j (x_i^T x_j) (*)$$

B. **Thay w vào $-\sum \alpha_n t_n w^T x_n$:**

$$-\sum_n \alpha_n t_n \underbrace{\left(\sum_m \alpha_m t_m x_m \right)^T}_{w^T} x_n = -2 \times (*)$$

C. Kết quả ($A + B = -A$):

$$\begin{aligned} \max_{\alpha} g(\alpha) &= \sum_n \alpha_n - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j t_i t_j (x_i^T x_j) \\ \text{s.t. } & \alpha_n \geq 0; \quad \sum \alpha_n t_n = 0 (g(\alpha) \text{ luôn là hàm lồi}) \end{aligned}$$

5. Bài toán Dual: Implementation (CVXOPT)

Solver: $\min \frac{1}{2} \alpha^T P \alpha + q^T \alpha$ s.t. $G\alpha \leq h, A\alpha = b$. Biến: $\alpha \in \mathbb{R}^N$.

Mapping:

$$\begin{aligned} P &= \mathbf{K}_{\text{Gram}}_{(N \times N)} \quad (K_{ij} = t_i t_j x_i^T x_j) \\ q &= -\mathbf{1}_{(N \times 1)} \quad (\text{Max } \Sigma \rightarrow \text{Min } -\Sigma) \\ G &= -I_{(N \times N)}; \quad h = \mathbf{0}_{(N \times 1)} \\ A &= t^T; \quad b = 0 \end{aligned}$$

6. Soft Margin & Kernel

1. **Soft Margin (ý tưởng):** Cho phép một số điểm vi phạm margin bằng biến slack ξ_n , đổi lại bị phạt trong hàm mục tiêu.

2. Bài toán gốc (Primal – soft margin):

$$\begin{aligned} \min_{w,b,\xi} & \frac{1}{2} \|w\|^2 + C \sum_{n=1}^N \xi_n \\ \text{s.t. } & t_n(w^T x_n + b) \geq 1 - \xi_n, \quad \xi_n \geq 0 \end{aligned}$$

Ý nghĩa: C : C lớn \Rightarrow phạt mạnh (ít sai, lề hẹp); C nhỏ \Rightarrow cho sai nhiều (lề rộng).

3. Dạng đối ngẫu (Dual):

$$\begin{aligned} \mathcal{L} &= \frac{1}{2} \|w\|^2 + C \sum_n \xi_n - \sum_n \alpha_n \{t_n(w^T x_n + b) - 1 + \xi_n\} - \sum_n \mu_n \xi_n \\ &\quad \max_{\alpha} \sum_{n=1}^N \alpha_n - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j t_i t_j x_i^T x_j \\ &\quad \text{s.t. } \boxed{0 \leq \alpha_n \leq C}, \quad \sum_{n=1}^N \alpha_n t_n = 0 \end{aligned}$$

Phân loại theo KKT:

- $\alpha_n = 0$: điểm ngoài margin (không support)
- $0 < \alpha_n < C$: support vector trên margin
- $\alpha_n = C$: support vector vi phạm / phân loại sai

4. Kernel Trick: Thay tích vô hướng:

$$x_i^T x_j \longrightarrow k(x_i, x_j)$$

$$f(x) = \sum_{n \in SV} \alpha_n t_n k(x_n, x) + b, \quad y = \text{sign}(f(x))$$