

SVM - Tóm tắt

## 1. Bài toán Gốc (Primal)

**Input:**  $X \in \mathbb{R}^{N \times (M-1)}$ ,  $t \in \{-1, +1\}^N$ . **Giả thiết:** Khả tách tuyến tính.

**Khoảng cách:**  $d_x = \frac{|w^T x + b|}{\|w\|}$ . **Hàm quyết định:**  $y(x) = w^T x + b$

Bài toán tối ưu:

$$w^*, b^* = \arg \min_{w, b} \frac{1}{2} \|w\|^2 \quad \text{s.t.: } t_n(w^T x_n + b) \geq 1, \forall n$$

**Độ rộng lề:**  $\delta = \frac{2}{\|w\|}$

**CVXOPT format:**  $\min_x \frac{1}{2} x^T K x + p^T x \text{ s.t.: } Gx \leq h$

$K = \text{diag}(1, \dots, 1, 0)_{M \times M}$ ,  $p = \mathbf{0}_{M \times 1}$ ,  $h = -\mathbf{1}_{N \times 1}$

$G_{N \times M}$ : dòng thứ  $n$  là  $[-t_n x_{n,1}, \dots, -t_n x_{n,M-1}, -t_n]$

## 2. Bài toán Đôi ngẫu (Dual)

**Lagrangian:**  $L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{n=1}^N \alpha_n \{t_n(w^T x_n + b) - 1\}$

**KKT:** (1)  $\nabla_{w,b} L = 0$ , (2)  $t_n(w^T x_n + b) \geq 1$ , (3)  $\alpha_n \geq 0$ , (4)

$$\alpha_n \{1 - t_n(w^T x_n + b)\} = 0$$

Từ KKT-1:  $w = \sum_{n=1}^N \alpha_n t_n x_n$ ,  $\sum_{n=1}^N \alpha_n t_n = 0$

Hàm đôi ngẫu:  $g(\alpha) = -\frac{1}{2} \sum_{r,c} \alpha_r \alpha_c t_r t_c x_r^T x_c$

Bài toán đôi ngẫu:

$$\alpha^* = \arg \min_{\alpha} \frac{1}{2} \alpha^T K \alpha - \mathbf{1}^T \alpha$$

$$\text{s.t.: } \alpha_n \geq 0, \sum_n \alpha_n t_n = 0$$

$K = K_{\text{Gram}} \odot T$  với  $K_{\text{Gram}} = XX^T$ ,  $T = tt^T$

**Support vectors:**  $\alpha_n > 0 \Rightarrow w^T x_n + b = 1$  (diễn là SV)

**Dự báo:**  $w = \sum_{s \in S} \alpha_s t_s x_s$ ,  $b = \frac{1}{|S|} \sum_{m \in S} (t_m - \sum_{n \in S} \alpha_n t_n x_n^T x_m)$

$y(x) = \sum_{n \in S} \alpha_n t_n x_n^T x + b$ , label = sign( $y$ )

## 3. Kernel Method

**Đồng lực:** Phi tuyến hóa decision boundary bằng ánh xạ  $\Phi(x)$

**Kernel trick:**  $k(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle$  (không cần tính  $\Phi$  tương minh)

**Điều kiện Mercer:** (1)  $k(x_i, x_j) = k(x_j, x_i)$ , (2)

$$\sum_i \sum_j c_i c_j k(x_i, x_j) \geq 0$$

**Kernel thông dụng:**

- Linear:  $k(x, x') = x^T x'$
- Polynomial:  $k(x, x') = (\gamma x^T x' + r)^d$
- RBF:  $k(x, x') = \exp(-\gamma \|x - x'\|^2)$
- Sigmoid:  $k(x, x') = \tanh(\gamma x^T x' + r)$

**Training với kernel:**  $K_{\text{Gram}} = k(X_{\text{train}}, X_{\text{train}})$

**Prediction:**  $K_{BS} = k(X_{\text{test}}, X_S)$ ,  $y = K_{BS}[a_S] + b$  với  $a_S = \alpha_S \odot t_S$

	Primal	Dual
Số biến	$M$	$N$
Số ràng buộc	$N$	$N + 1$
Dùng khi	$M \ll N$	$M > N$
Kernel	Khó	Dễ