SVM 02

Agenda:

1. Soft Svm
2. Dual formulation.
3. Kernel Trick
4. Polynomial Kernel
5. RBF Kernel

Primal SVM

$$\underset{w,b}{\text{argmin}} \; \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m} \xi_i$$

$$\text{s.t} \quad \forall i \; y_i(\langle w, x_i\rangle + b) \geq 1 - \xi_i$$

$$\forall i \; \xi_i \geq 0$$

$$\forall i \; \text{Loss}_{\text{hinge}}(w, b, x_i, y_i)$$

$$\text{Loss}_{\text{hinge}}(w, b, x, y) = \max(0, 1 - y(\langle w, x\rangle + b))$$

Primal + Dual Equivalence

$$\underset{w,b}{\text{argmin}} \left[\frac{1}{2}\|w\|^2 + \underbrace{\sum_{i=1}^{m} \max(0, 1 - y_i(\langle w, x_i\rangle + b))}_{\text{Hinge loss}} \right]$$

$$\text{s.t} \quad \forall i \; \xi_i \geq 1 - y_i(\langle w, x_i\rangle + b)$$

$$\forall i \; \xi_i \geq 0$$

$$\forall i \; \text{Loss}_{\text{hinge}}$$

Derivation.

① $\forall i \; \xi_i = \xi_i$

② $\forall i \; \xi_i \geq \text{Loss}_{\text{hinge}}$

③ $\forall i \; \xi_i \geq \text{Loss}_{\text{hinge}}$

$$\boxed{\forall i \; \xi_i = \max(0, 1 - y_i(\langle w, x_i\rangle + b))}$$

Support vectors

$$\forall i \; \alpha_i^* = 0 \quad \text{for all} \quad \text{Non-Support Vectors}$$

$$\forall i \; \alpha_i > 0 \quad \text{for all} \quad \text{Support vectors}$$

$$w_{\text{SVM}} = \text{argmax}(\alpha_1, \ldots, \alpha_m)$$

$$\text{argmax}(1, 1, 0, 0)$$

$$\text{argmax}(0, 0, 1, 1)$$

$$\alpha_1 = 0$$

$$\alpha_2 = 0$$

$$\alpha_3 = 1$$

$$\alpha_4 = 1$$

$$\text{Support vectors}$$

$$\text{Non-Support vectors}$$

Dual Form (Intuition)

$$\underset{\alpha}{\min} \; L(\alpha) = \sum_{i=1}^{m} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{m} \alpha_i \alpha_j y_i y_j \langle x_i, x_j\rangle$$

$$\text{s.t.} \quad \forall i \; 0 \leq \alpha_i \leq C$$

$$\sum_{i=1}^{m} \alpha_i^* y_i = 0$$

$$\forall i \; \alpha_i^* \geq 0$$

$$\forall i \; \alpha_i^* \leq C$$

$$\forall i \; \alpha_i^* \geq 0$$

$$\forall i \; \alpha_i^* \leq C$$

$$\forall i \; \text{Loss}