

Kernel SVM

Our goal is

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m [1 - y_i(w \cdot x_i + b)]_+$$

By Lagrange multiply method,

$$w = \sum_{i=1}^m \alpha_i y_i x_i$$

Take w back to our goal

$$\min \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \alpha_i \alpha_j y_i y_j x_i \cdot x_j + C \sum_{i=1}^m [1 - y_i (\sum_{j=1}^m \alpha_j y_j x_j \cdot x_i + b)]_+$$

In the following, use K to denote kernel function of two vector

$$K(x_i, x_j) = \phi(x_i, x_j)$$

For single example x_i :

$$1. \text{ if } \left[1 - y_i \left(\sum_{j=1}^m \alpha_j y_j x_j \cdot x_i + b \right) \right] > 0:$$

$$\frac{\partial L}{\partial \alpha_i} = \frac{1}{2} \alpha_i K(x_i, x_i) + \frac{1}{2} y_i \sum_{j=1}^m [\alpha_j y_j K(x_i, x_j)] + C(-y_i) y_i K(x_i, x_j)$$

$$\frac{\partial L}{\partial b} = C(-y_i)$$

2. else:

$$\frac{\partial L}{\partial \alpha_i} = \frac{1}{2} \alpha_i K(x_i, x_i) + \frac{1}{2} y_i \sum_{j=1}^m [\alpha_j y_j K(x_i, x_j)]$$

$$\frac{\partial L}{\partial b} = 0$$