SVM: What if X is not separable w.r.t. y? What about noise?

14

We can introduce slack variables, ξ_i :

$$\min_{\mathbf{w},b} \|\mathbf{w}\|_2^2 + c \sum_{i=1}^{M} \xi_i$$

$$\text{s.t.} \quad \forall i \; y_i\left(\mathbf{w}^T \mathbf{x}_i + b\right) \geq 1 - \xi_i$$

$$\text{where} \quad \xi_i = \begin{cases} 1 - y_i(\mathbf{w}^T \mathbf{x}_i + b) & if \quad y_i(\mathbf{w}^T \mathbf{x}_i + b) < 1 \\ 0 & if \quad y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1 \end{cases}$$

which can be reduced to:

$$\min_{\mathbf{w},b} \|\mathbf{w}\|_2^2 + c \sum_{i=1}^{M} \max\left(1 - y_i(\mathbf{w}^T\mathbf{x}_i + b), 0\right)$$

$$\text{s.t.} \quad y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1 - \max\left(1 - y_i(\mathbf{w}^T\mathbf{x}_i + b), 0\right)$$

SVM: What if X is not separable w.r.t. y? What about noise?

We can introduce slack variables, ξ_i :

$$\min_{\mathbf{w},b} \|\mathbf{w}\|_2^2 + c\sum_{i=1}^{M} \xi_i$$

s.t. $\forall i \; y_i\left(\mathbf{w}^T\mathbf{x}_i + b\right) \geq 1 - \xi_i$

where $\xi_i = \begin{cases} 1 - y_i(\mathbf{w}^T\mathbf{x}_i + b) & if \quad y_i(\mathbf{w}^T\mathbf{x}_i + b) < 1 \\ 0 & if \quad y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1 \end{cases}$

which can be reduced to:

$$\min_{\mathbf{w},b} \|\mathbf{w}\|_2^2 + c\sum_{i=1}^{M} \max\left(1 - y_i(\mathbf{w}^T\mathbf{x}_i + b), 0\right)$$

s.t. $y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1 - \max\left(1 - y_i(\mathbf{w}^T\mathbf{x}_i + b), 0\right)$

Georgetown University

Lab: Text Normalization