SVM outline

Date: 1-06-2018

(14)

Primal Problem:
$$\min_{\omega} \frac{1}{2} \|\omega\|^2$$
$$\text{s.t } y^{(i)}(\omega^T x^{(i)} + b) \geq 1 \quad \forall i \in \{1, 2, \ldots m\}$$

Dual Problem:
$$\max_{\alpha} \Theta_D(\alpha) = \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_i \alpha_j y^{(i)} y^{(j)} \langle x^{(i)}, x^{(j)} \rangle$$
$$\text{s.t } \alpha_i \geq 0 \text{ and } \sum_{i=1}^{m} \alpha_i y^{(i)} = 0$$

Replace $\langle x^{(i)}, x^{(j)} \rangle$ with Kernels

Kernels:

linear: $K(x, z) = x^T z + b$

Polynomial: $K(x, z) = (x^T z + b)^d$

gaussian or rbf: $K(x, z) = \exp\left(\frac{-\|x - z\|^2}{2\sigma^2}\right)$

All $K(x, z) \in \mathbb{R}$

tells how close given two feature vectors are

SMO: use SMO algo to solve dual Problem
find α's

$$\max_{\alpha} \Theta_D(\alpha_1, \alpha_2, \ldots \alpha_m)$$

~~Fix two~~ vary two α's and Fix rest and update α's.

$$\sum_{i=1}^{m} \alpha_i y^{(i)} = 0$$

~~Ex~~ Vary two alphas at a time.

$$\alpha_i y^{(i)} + \alpha_j y^{(j)} = -\sum_{\substack{k=1 \\ k \neq i,j}}^{m} \alpha_k y^k = \text{const} \rightarrow (1)$$

Let $S = y^{(i)} y^{(j)}$

Multiply eqn (1) with $y^{(i)}$

$$\alpha_i y^{(i)} y^{(i)} + \alpha_j y^{(i)} y^{(j)} = \gamma$$

$$\alpha_i = \gamma - \alpha_j S$$

Dual objective function written as:

$$\theta_D(\alpha) = \alpha_i + \alpha_j + \text{const} - \frac{1}{2}\Big(y_i y_i x_i^T x_i \alpha_i \alpha_i + y_j y_j x_j^T x_j \alpha_j \alpha_j + 2 y_i y_j x_i^T x_j \alpha_i \alpha_j + 2 \Big(\sum_{\substack{k=3 \\ k \neq i,j}}^{m} \alpha_k y_k x_k^T\Big)\big(y_i x_i \alpha_i + y_j x_j \alpha_j\big) + \text{const} \Big)$$

Solving :

$$\theta_D(\alpha) = \frac{1}{2}\eta\alpha_j^2 + \left(y_j\left(E_i^{old} - E_j^{old}\right) - \eta\alpha_j^{old}\right)\alpha_j + \text{const}$$

where

$$\eta = 2K_{ij} - K_{ii} - K_{jj}$$

$$K_{ii} = x_i^T x_i, \quad K_{jj} = x_j^T x_j, \quad K_{ij} = x_i^T x_j$$

$$E_i^{old} = \left(x_i^T w^{old} - b^{old}\right) - y^{(i)}$$

$$E_j^{old} = \left(x_j^T w^{old} - b^{old}\right) - y^{(j)}$$

SMO α update rule :

$$\alpha_j := \alpha_j + \frac{y^{(j)}\left(E_j^{old} - E_i^{old}\right)}{\eta}$$

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

$$b = -\frac{\max_{i, y^{(i)}=-1} w^T x^{(i)} + \min_{i, y^{(i)}=1} w^T x^{(i)}}{2}$$

Prediction : $w^T x + b = \sum_{i=1}^{m} \alpha_i y^{(i)} K(x^{(i)}, x) + b$

Scanned by CamScanner