Machine Learning Assignment 6

10185101210 陈俊潼

4.10

在线性SVM中，我们假设分类边界的形式为$\mathbf{w}^T\mathbf{x}+b$，其中包含一个偏置项$b\in\mathbb{R}$. 但是，学习一个不含偏置项的SVM分类器，即使用$\mathbf{w}^T\mathbf{x}$作为分类边界也是可行的。

(a) 如果不含偏置项，原始空间中的优化问题是什么样的?
(b) 如果不含偏置项，给出其对偶形式，如果关于拉格朗日乘子λ的解是λ^*,则最优的决策边界$\mathbf{w}^*$是?
(c) 当希望存在偏置项时，如果希望上述方法仍然是有用的。给定一个训练数据集$(\mathbf{x}_i, y_i)(1\le i\le n)$,我们可将任意$\mathbf{x}\in\mathbb{R}$转换为$\mathbb{R}^{d+1}$空间中的$\hat{\mathbf{x}}$，这只需要通过对$\mathbf{x}$增加一个额外的维度即可，所增加的维度总有一个常数值1。假设λ^*是关于对偶形式的最优解，并假设分类边界为$\mathbf{w}^T\mathbf{x}+b$，则最优的b^*值是多少

a. 如果不含偏置项，优化问题为：

$$\min_{w,b}\gamma=\frac{2}{||\omega||}$$
$$s.t.\, y_i(w^t x_i)\ge 1$$

b. 拉格朗日函数为：

$$L(\omega,b,\lambda)=\frac{1}{2}||\omega||^2+\sum_{i=1}^{m}\lambda_i(1-y_i(\omega^T x_i+b))$$

令偏导数为 0，得到对偶问题：

$$\max_{\alpha}\sum_{i=1}^{m}\alpha_i-\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\lambda_i\lambda_j y_i y_j x_i^T x_j$$
$$s.t.\sum_{i=1}^{m}\lambda_i y_i=0$$
$$\lambda_i\ge 0, i=1,2,3,\dots m$$

最优决策边界中：

$$\omega^*=\sum_{i=1}^{m}\lambda_i y_i x_i^T x$$

4.11

Consider the general 2-nd polynomial kernel $K2(\mathbf{x};\mathbf{x}') = (\xi + \gamma\mathbf{x}^T\mathbf{x}')^2$. Which of the following transform can be used to derive this kernel?
A. $\phi(\mathbf{x}) = (1, \sqrt{2\gamma}x_1, \ldots, \sqrt{2\gamma}x_d, \gamma x_1^2, \ldots, \gamma x_d^2)$
B. $\phi(\mathbf{x}) = (\xi, \sqrt{2\gamma}x_1, \ldots, \sqrt{2\gamma}x_d, x_1^2, \ldots, x_d^2)$
C. $\phi(\mathbf{x}) = (\xi, \sqrt{2\gamma\xi}x_1, \ldots, \sqrt{2\gamma\xi}x_d, x_1^2, \ldots, x_d^2)$
D. $\phi(\mathbf{x}) = (\xi, \sqrt{2\gamma\xi}x_1, \ldots, \sqrt{2\gamma\xi}x_d, \gamma x_1^2, \ldots, \gamma x_d^2)$

D

4.12

At the optimal solution of

$$\begin{array}{ll} \min & \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{i=1}^{N}\xi_i \\ s.t. & y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1 - \xi_i, \xi_i \geq 0 \quad (i = 1, 2, \ldots, N) \end{array}$$

assume that $y1(wTz1 + b) = -10$. What is the corresponding ξ_1?
A. 1
B. 11
C. 21
D. 31

B