

机器学习课程 第6次作业

黄昊 20204205

选择题目 6.1 6.4 6.6 6.7 6.9

6.1 显然超平面 $w^T x + b = 0$ 的法向量为 w ，考虑超平面上一点 x_0 及空间中任意一点 x ，有向量 $x - x_0$ ，考虑该向量与法向量的夹角 θ ，可求得点 x 到超平面的距离 r 为：

$$\begin{aligned} r &= \| (x - x_0) \|_2 \cos \theta \\ &= \| (x - x_0) \|_2 \frac{|w^T (x - x_0)|}{\|w\|_2 \|x - x_0\|_2} \\ &= \frac{|w^T x + b - (w^T x_0 + b)|}{\|w\|_2} \\ &= \frac{|w^T x + b|}{\|w\|_2} \end{aligned}$$

6.4 考虑同一个线性可分的二分类问题，当LDA的投影直线的 w_1 与SVM的超平面 w_2 垂直时，且满足各自的约束条件时，两者等价，即：

$$LDA: w_1 = S_w^{-1}(\mu_0 - \mu_1)$$

$$SVM: w_2 = \sum_{i=1}^m \alpha_i y_i x_i$$

$$s.t. \quad w_1^T w_2 = 0$$

6.6 直观上看，SVM的最终模型仅与少数的支持向量有关。此时若噪声成为了最终的支持向量，将会严重影响模型，甚至会导致线性不可分。因此，SVM对噪声敏感。

6.7 KKT条件为：

$$\begin{cases} \alpha_i (f(x_i) - y - \epsilon - \xi_i) = 0 \\ \hat{\alpha}_i (y_i - f(x_i) - \epsilon - \hat{\xi}_i) = 0 \\ \mu_i \xi_i = 0 \\ \hat{\mu}_i \hat{\xi}_i = 0 \end{cases}$$

6.9 原始的对率回归模型为：

$$\ln \frac{y}{1-y} = w^T x + b$$

现在引入核函数，则核率回归的模型如下所示：

$$\ln \frac{y}{1-y} = w^T \phi(x) + b$$

不妨令 $\beta^T = (w^T, b)$ ， $\hat{x} = (x^T, 1)$ ，则模型变为：

$$\ln \frac{y}{1-y} = \beta^T \phi(\hat{x})$$

相应地，极大似然函数变为：

$$l(\beta) = \sum_{i=1}^m (-y_i \beta^T \phi \hat{x} + \ln(1 + e^{\beta^T \phi \hat{x}}))$$

由表示定理，有：

$$\ln \frac{y}{1-y} = h(\hat{x}) = \sum_{i=1}^m \alpha_i \kappa(\hat{x}, \hat{x}_i)$$

因此，损失函数变为：

$$l(\alpha) = \sum_{i=1}^m (-y_i \sum_{j=1}^m \alpha_j \kappa(\hat{x}_i, \hat{x}_j)) + \sum_{i=1}^m \ln(1 + e^{\sum_{j=1}^m \alpha_j \kappa(\hat{x}_i, \hat{x}_j)})$$

对参数求导：

$$\frac{\partial l}{\partial \alpha_j} = - \sum_{i=1}^m y_i \kappa(\hat{x}_i, \hat{x}_j) + \sum_{i=1}^m \frac{\kappa(\hat{x}_i, \hat{x}_j)}{1 + e^{\sum_{j=1}^m \alpha_j \kappa(\hat{x}_i, \hat{x}_j)}} e^{\sum_{j=1}^m \alpha_j \kappa(\hat{x}_i, \hat{x}_j)}$$

按照该公式即可实现更新。