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

$$
\begin{aligned}
r &= \parallel (\boldsymbol{x}-\boldsymbol{x_0}) \parallel_2 cos\theta \\
&= \parallel (\boldsymbol{x}-\boldsymbol{x_0}) \parallel_2 \frac{|\boldsymbol{w^T}(\boldsymbol{x}-\boldsymbol{x_0})|}{\parallel \boldsymbol{w} \parallel_2 \parallel \boldsymbol{x}-\boldsymbol{x_0} \parallel_2} \\
&= \frac{|\boldsymbol{w^T x} + b - (\boldsymbol{w^T x_0} + b)|}{\parallel \boldsymbol{w} \parallel_2} \\
&= \frac{|\boldsymbol{w^T x} + b|}{\parallel \boldsymbol{w} \parallel_2}
\end{aligned}
$$

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$\boldsymbol{w}_1$与SVM的超平面$\boldsymbol{w}_2$垂直时，且满足各自的约束条件时，两者等价，即：

$$
\begin{aligned}
LDA &: \boldsymbol{w_1} = S_w^{-1}(\boldsymbol{\mu_0} - \boldsymbol{\mu_1}) \\
SVM &: \boldsymbol{w_2} = \sum_{i=1}^{m} \alpha_i y_i \boldsymbol{x_i} \\
s.t. &\quad \boldsymbol{w_1^T w_2} = 0
\end{aligned}
$$

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：

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

6.9　原始的对率回归模型为：

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

现在引入核函数，则核对率回归的模型如下所示：

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

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

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

相应地，极大似然函数变为：

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

由表示定理，有：

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

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。