非线性判别方法习题

姓名： 甘云冲

学号： 2101213081

1.

(1)

$$
\begin{aligned}
\delta_k^L(t) &= \frac{\partial E(t)}{\partial u_k^L(t)} \\
&= -y_k(t)\frac{\partial}{\partial u_k^L(t)}\ln o_k(t) \\
&= -y_k(t)\frac{\partial}{\partial u_k^L(t)}\left[\ln\frac{1}{1+e^{-\alpha u_k^L(t)}}\right] \\
&= y_k(t)\frac{\partial}{\partial u_k^L(t)}\ln\left(1+e^{-\alpha u_k^L(t)}\right) \\
&= y_k(t)\frac{1}{1+e^{-\alpha u_k^L(t)}}(-\alpha)e^{-\alpha u_k^L(t)} \\
&= -\alpha y_k(t)\frac{e^{-\alpha u_k^L(t)}}{1+e^{-\alpha u_k^L(t)}} \\
&= -\alpha y_k(t)(1-o_k(t))
\end{aligned}
$$

(2)

$$
\begin{aligned}
\delta_k^L(t) &= \frac{\partial E(t)}{\partial u_k^L(t)} \\
&= -\sum_{i=1}^{K} y_i(t) \frac{\partial}{\partial u_k^L(t)} \ln o_i(t) \\
&= -\sum_{i=1}^{K} \frac{y_i(t)}{o_i(t)} \frac{\partial}{\partial u_k^L(t)} o_i(t) \\
&= -\sum_{i \neq k} \frac{y_i(t)}{o_i(t)} \left(-\frac{e^{u_i^L(t)} e^{u_k^L(t)}}{\left(\sum_{j=1}^{K} e^{u_j^L(t)}\right)^2} \right) - \frac{y_k(t)}{o_k(t)} \frac{e^{u_k^L(t)} \left(\sum_{j=1}^{K} e^{u_j^L(t)}\right) - e^{u_k^L(t)} e^{u_k^L(t)}}{\left(\sum_{j=1}^{K} e^{u_j^L(t)}\right)^2} \\
&= \sum_{i=1}^{K} \frac{y_i(t)}{o_i(t)} \frac{\left(\sum_{j=1}^{K} e^{u_j^L(t)}\right) e^{u_k^L(t)}}{\left(\sum_{j=1}^{K} e^{u_j^L(t)}\right)^2} - \frac{y_k(t)}{o_k(t)} \frac{e^{u_k^L(t)}}{\sum_{j=1}^{K} e^{u_j^L(t)}} \\
&= \sum_{i=1}^{K} \frac{y_i(t)}{o_i(t)} o_k(t) - \frac{y_k(t)}{o_k(t)} o_k(t) \\
&= e^{u_k^L(t)} \sum_{i=1}^{K} y_i(t) e^{-u_i^L(t)} - y_k(t)
\end{aligned}
$$

2.

拉格朗日函数为：

$$
L = \frac{1}{2} ||\mathbf{w}||^2 + C \sum_{i=1}^{n} \xi_i^2 - \sum_{i=1}^{n} \alpha_i [y_i(\mathbf{w}^T \mathbf{x}_i + b) - 1 + \xi_i]
$$

于是：

$$
\begin{aligned}
\frac{\partial L}{\partial \mathbf{w}} &= \mathbf{w} - \sum_{i=1}^{n} \alpha_i y_i \mathbf{x}_i \\
\frac{\partial L}{\partial b} &= \sum_{i=1}^{n} \alpha_i y_i \\
\frac{\partial L}{\partial \xi_i} &= 2C\xi_i - \alpha_i
\end{aligned}
$$

令上式都为0，则有：

$$
\begin{aligned}
L &= \frac{1}{2}||\mathbf{w}||^2 + C\sum_{i=1}^{n}\xi_i^2 - \sum_{i=1}^{n}\alpha_i[y_i(\mathbf{w}^T\mathbf{x}_i + b) - 1 + \xi_i] \\
&= \frac{1}{2}||\mathbf{w}||^2 + C\sum_{i=1}^{n}\xi_i^2 - \mathbf{w}^T\sum_{i=1}^{n}\alpha_i y_i \mathbf{x}_i - b\sum_{i=1}^{n}\alpha_i y_i + \sum_{i=1}^{n}\alpha_i(1 - \xi_i) \\
&= -\frac{1}{2}||\mathbf{w}||^2 + C\sum_{i=1}^{n}\xi_i^2 + \sum_{i=1}^{n}2C\xi_i(1 - \xi_i) \\
&= -\frac{1}{2}||\mathbf{w}||^2 + \sum_{i=1}^{n}2C\xi_i - C\xi_i^2 \\
&= -\frac{1}{2}\sum_{i,j=1}^{n}y_i y_j \alpha_i \alpha_j \mathbf{x}_i^T\mathbf{x}_j - C\sum_{i=1}^{n}(\xi_i - 1)^2
\end{aligned}
$$

所以对偶问题为：

$$
\begin{aligned}
\min_{\alpha} \left[\frac{1}{2}\sum_{i,j=1}^{n}y_i y_j \alpha_i \alpha_j \mathbf{x}_i^T\mathbf{x}_j + C\sum_{i=1}^{n}(\xi_i - 1)^2\right] \\
s.t. \quad \xi_i > 0 \qquad i = 1, \ldots, n \\
\sum_{i=0}^{n}\alpha_i y_i = 0
\end{aligned}
$$

对应的核函数形式为：

$$
\begin{aligned}
\min_{\alpha} \left[\frac{1}{2}\sum_{i,j=1}^{n}y_i y_j \alpha_i \alpha_j K(\mathbf{x}_i, \mathbf{x}_j) + C\sum_{i=1}^{n}(\xi_i - 1)^2\right] \\
s.t. \quad \xi_i > 0 \qquad i = 1, \ldots, n \\
\sum_{i=0}^{n}\alpha_i y_i = 0
\end{aligned}
$$