۱

تمرین چهارم مبانی یادگیری ماشین

کسری حسنی ۹۹۳۶۲۲۰۳۰

①

a) $\mathcal{L}(w,b,\varepsilon,\alpha) = \frac{1}{2}w^Tw + \frac{C}{2}\sum_{i=1}^{m} \varepsilon_i^2 - \sum_{i=1}^{m}\alpha_i\left[y^{(i)}(w^Tx^{(i)}+b)-1+\varepsilon_i\right]$

b) $\nabla_w \mathcal{L} = 0 = w - \sum_{i=1}^{m}\alpha_i y^{(i)}x^{(i)}$

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

$\frac{\partial \mathcal{L}}{\partial b} = 0 = -\sum_{i=1}^{m}\alpha_i y^{(i)}$

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

$\nabla_{\varepsilon}\mathcal{L} = C\varepsilon - \alpha_i = 0 \longrightarrow C\varepsilon_i - \alpha_i = 0 \longrightarrow C\varepsilon_i = \alpha_i$

c) $w(\alpha) = \min_{w,b,\varepsilon}\mathcal{L}(w,b,\varepsilon,\alpha) = \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}(\alpha_i y^{(i)}x^{(i)})^T(\alpha_j y^{(j)}x^{(j)}) + \frac{1}{2}\sum_{i=1}^{m}$

$\frac{\alpha_i}{\varepsilon_i}\varepsilon_i^2 - \sum_{i=1}^{m}\alpha_i\left[y^{(i)}((\sum_{j=1}^{m}\alpha_j y^{(j)}x^{(j)})^T x^{(i)} + b) - 1 + \varepsilon_i\right] =$

$= \sum_{i=1}^{m}\alpha_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_i\alpha_j y^{(i)}y^{(j)}(x^{(i)})^T x^{(j)} - \frac{1}{2}\sum_{i=1}^{m}\alpha_i\varepsilon_i =$

$= \sum_{i=1}^{m}\alpha_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_i\alpha_j y^{(i)}y^{(j)}(x^{(i)})^T x^{(j)} - \frac{1}{2}\sum_{i=1}^{m}\frac{\alpha_i^2}{C}$

۲

$$Pr(\sim P_{۳}) = \sum_{P_{۱},P_{۲},P_{۴}} Pr(P_{۱},P_{۲},\sim P_{۳},P_{۴}) = \sum_{P_{۱},P_{۲},P_{۴}} Pr(P_{۱})Pr(P_{۲}|P_{۱})Pr(\sim P_{۳}|P_{۲})Pr(P_{۴}|P_{۲})$$

$$= Pr(P_{۱})Pr(P_{۲}|P_{۱})Pr(\sim P_{۳}|P_{۲})Pr(P_{۴}|P_{۲}) + Pr(P_{۱})Pr(P_{۲}|P_{۱})Pr(\sim P_{۳}|P_{۲})Pr(\sim P_{۴}|P_{۲})$$
$$+ Pr(P_{۱})Pr(\sim P_{۲}|P_{۱})Pr(\sim P_{۳}|\sim P_{۲})Pr(P_{۴}|\sim P_{۲}) + Pr(P_{۱})Pr(\sim P_{۲}|P_{۱})Pr(\sim P_{۳}|\sim P_{۲})Pr(\sim P_{۴}|\sim P_{۲})$$
$$+ Pr(\sim P_{۱})Pr(P_{۲}|\sim P_{۱})Pr(\sim P_{۳}|P_{۲})Pr(P_{۴}|P_{۲}) + Pr(\sim P_{۱})Pr(P_{۲}|\sim P_{۱})Pr(\sim P_{۳}|P_{۲})Pr(\sim P_{۴}|P_{۲})$$
$$+ Pr(\sim P_{۱})Pr(\sim P_{۲}|\sim P_{۱})Pr(\sim P_{۳}|\sim P_{۲})Pr(P_{۴}|\sim P_{۲}) + Pr(\sim P_{۱})Pr(\sim P_{۲}|\sim P_{۱})Pr(\sim P_{۳}|\sim P_{۱})$$
$$Pr(\sim P_{۴}|\sim P_{۲}) = ۰٫۲۰۴۸ + ۰٫۰۵۱۲ + ۰٫۰۲۸ + ۰٫۰۲۸ + ۰٫۱۹۲ + ۰٫۰۴۸ + ۰٫۱۰۵ + ۰٫۱۰۵ = ۰٫۷۶۲$$

$$Pr(P_{۲}|\sim P_{۳}) = \frac{Pr(P_{۲},\sim P_{۳})}{Pr(\sim P_{۳})}$$

$$Pr(\sim P_{۳}) = ۰٫۷۶۲$$

$$Pr(P_{۲},\sim P_{۳}) = \sum_{P_{۱},P_{۴}} Pr(P_{۱},P_{۲},\sim P_{۳},P_{۴}) = \sum_{P_{۱},P_{۴}} Pr(P_{۱})Pr(P_{۲}|P_{۱})Pr(\sim P_{۳}|P_{۲})Pr(P_{۴}|P_{۲})$$
$$= ۰٫۲۰۴۸ + ۰٫۰۵۱۲ + ۰٫۱۹۲ + ۰٫۰۴۸ = ۰٫۴۹۶$$

$$Pr(P_{۲}|\sim P_{۳}) = \frac{۰٫۴۹۶}{۰٫۷۶۲}$$

۳

a) $$S = \begin{bmatrix} Cov(X_{۱},X_{۱}) & Cov(X_{۱},X_{۲}) \\ Cov(X_{۲},X_{۱}) & Cov(X_{۲},X_{۲}) \end{bmatrix} = \begin{bmatrix} ۱۴ & -۱۱ \\ -۱۱ & ۲۳ \end{bmatrix}$$

b) $$\det(S-\lambda I) = ۰ = \begin{vmatrix} ۱۴-\lambda & -۱۱ \\ -۱۱ & ۲۳-\lambda \end{vmatrix} = (۱۴-\lambda)(۲۳-\lambda) - (-۱۱)\times(-۱۱) =$$

$$= \lambda^{۲} - ۳۷\lambda + ۲۰۱ \rightarrow \lambda_{۱} = ۳۰٫۳۸, \quad \lambda_{۲} = ۶٫۶۱$$

$$U = \begin{bmatrix} u_{۱} \\ u_{۲} \end{bmatrix} \qquad (S-\lambda_{۱}I)X = \begin{bmatrix} ۰ \\ ۰ \end{bmatrix} = \begin{bmatrix} ۱۴-\lambda_{۱} & -۱۱ \\ -۱۱ & ۲۳ \end{bmatrix} \begin{bmatrix} u_{۱} \\ u_{۲} \end{bmatrix} = \begin{bmatrix} (۱۴-\lambda_{۱})u_{۱} - ۱۱u_{۲} \\ -۱۱u_{۱} + (۲۳-\lambda_{۱})u_{۲} \end{bmatrix}$$

$$(۱۴-\lambda_۱)u_۱ - ۱۱u_۲ = ۰$$
$$-۱۱u_۱ + (۲۳-\lambda_۱)u_۲ = ۰$$

$$U_۱ = \begin{bmatrix} ۱۱ \\ ۱۴-\lambda_۱ \end{bmatrix} \qquad \|u_۱\| = \sqrt{۱۱^۲ + (۱۴-\lambda_۱)^۲} = \sqrt{۱۱^۲ + (۱۴-۳۰,۳۸)^۲} = ۱۹,۷۳$$

$$e_۱ = \begin{bmatrix} ۱۱/\|u_۱\| \\ (۱۴-\lambda_۱)/\|u_۱\| \end{bmatrix} = \begin{bmatrix} ۰,۵۵ \\ -۰,۸۳ \end{bmatrix}$$

$$e_۲ = \begin{bmatrix} ۰,۸۳ \\ ۰,۵۵ \end{bmatrix}$$