Fisher 线性判别习题解答

1，已知两类样本的数据如下：

$\omega_1:\{(5,37),(7,30),(10,35),(11.5,40),(14,38),(12,31)\}$

$\omega_2:\{(35,21.5),(39,21.7),(34,16),(37,17)\}$

试用 Fisher 判别函数法，求出最佳投影方向 W，及分类阈值 y_0

解：由题意知：

$$\mu_1 = \frac{1}{6}\sum_{n=1}^{6} X_n^{(1)} = (9.92\ 35.17)^T$$

$$\mu_{-1} = \frac{1}{4}\sum_{n=1}^{4} X_n^{(0)} = (36.25\ 19.05)^T$$

则可计算出类内离差阵：

$$\Sigma_1 = \sum_{n=1}^{6}(X_n^{(1)} - \mu_1)\cdot(X_n^{(1)} - \mu_1)^T = \begin{pmatrix} 56.21 & 16.58 \\ 16.58 & 78.83 \end{pmatrix}$$

$$\Sigma_{-1} = \sum_{n=1}^{4}(X_n^{(0)} - \mu_{-1})\cdot(X_n^{(0)} - \mu_{-1})^T = \begin{pmatrix} 14.75 & 9.55 \\ 9.55 & 26.53 \end{pmatrix}$$

$$S_w = \Sigma_1 + \Sigma_{-1} = \begin{pmatrix} 70.96 & 26.13 \\ 26.13 & 105.36 \end{pmatrix}$$

$${S_w}^{-1} = \begin{pmatrix} 0.0155 & -0.0038 \\ -0.0038 & 0.0104 \end{pmatrix}$$

从而可计算出最佳投影方向：

$$W^* = {S_w}^{-1}(\mu_1 - \mu_{-1}) = (-0.4704, 0.2696)^T$$

$$y_0 = W^{*T}\frac{(\mu_1 + \mu_{-1})}{2} = -3.55$$

2，在 Fisher 判别中，用向量梯度的计算法则证明：$\boldsymbol{S}_B\boldsymbol{w} = \lambda \boldsymbol{S}_w\boldsymbol{w}$

证明：$L(\boldsymbol{w},\lambda) = \boldsymbol{w}^T\boldsymbol{S}_B\boldsymbol{w} + \lambda(K - \boldsymbol{w}^T\boldsymbol{S}_w\boldsymbol{w}) = \boldsymbol{w}^T(\boldsymbol{S}_B - \lambda\boldsymbol{S}_w)\boldsymbol{w} + \lambda K$

$$\nabla L_w(\boldsymbol{w},\lambda) = \frac{\partial L(\boldsymbol{w},\lambda)}{\partial \boldsymbol{w}} = \boldsymbol{0}^T$$

$$\frac{\partial L(\boldsymbol{w},\lambda)}{\partial \boldsymbol{w}} = \frac{\partial(\boldsymbol{w}^T(\boldsymbol{S}_B - \lambda\boldsymbol{S}_w)\boldsymbol{w})}{\partial \boldsymbol{w}} + 0$$

根据$\frac{\partial \boldsymbol{u}^T\boldsymbol{v}}{\partial \mathbf{x}} = \boldsymbol{u}^T\frac{\partial \boldsymbol{v}}{\partial \mathbf{x}} + \boldsymbol{v}^T\frac{\partial \boldsymbol{u}}{\partial \mathbf{x}}$，$\frac{\partial \boldsymbol{A}\mathbf{x}}{\partial \mathbf{x}} = \boldsymbol{A}$，以及$\frac{\partial \mathbf{x}}{\partial \mathbf{x}} = \boldsymbol{I}$，上式：

$$\frac{\partial(\boldsymbol{w}^T(\boldsymbol{S}_B - \lambda\boldsymbol{S}_w)\boldsymbol{w})}{\partial \boldsymbol{w}} = \boldsymbol{w}^T(\boldsymbol{S}_B - \lambda\boldsymbol{S}_w)\boldsymbol{I} + \left((\boldsymbol{S}_B - \lambda\boldsymbol{S}_w)\boldsymbol{w}\right)^T\boldsymbol{I}$$

$$= \boldsymbol{w}^T(\boldsymbol{S}_B - \lambda\boldsymbol{S}_w)\boldsymbol{I} + \boldsymbol{w}^T(\boldsymbol{S}_B - \lambda\boldsymbol{S}_w)^T\boldsymbol{I}$$

因为：$\boldsymbol{S}_B$和$\boldsymbol{S}_w$均为对称矩阵，

所以：$(\boldsymbol{S}_B - \lambda\boldsymbol{S}_w)^T = (\boldsymbol{S}_B - \lambda\boldsymbol{S}_w)$，

又：$(\boldsymbol{S}_B - \lambda\boldsymbol{S}_w)\boldsymbol{I} = (\boldsymbol{S}_B - \lambda\boldsymbol{S}_w)$

所以：
$$\frac{\partial L(\boldsymbol{w},\lambda)}{\partial \boldsymbol{w}} = 2\boldsymbol{w}^T(\boldsymbol{S}_B - \lambda\boldsymbol{S}_w) = \boldsymbol{0}^T$$

$$2(\boldsymbol{S}_B - \lambda\boldsymbol{S}_w)\boldsymbol{w} = \boldsymbol{0}$$

$$\boldsymbol{S}_B\boldsymbol{w} = \lambda\boldsymbol{S}_w\boldsymbol{w}$$

3，在 Fisher 判别中，将目标函数$J(\boldsymbol{w})=\frac{\boldsymbol{w}^T\boldsymbol{S}_B\boldsymbol{w}}{\boldsymbol{w}^T\boldsymbol{S}_w\boldsymbol{w}}$作为分式求梯度，推导出$\boldsymbol{w}^*=\boldsymbol{S}_w^{-1}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)$

解：

将$J(\boldsymbol{w})=\frac{\boldsymbol{w}^T\boldsymbol{S}_B\boldsymbol{w}}{\boldsymbol{w}^T\boldsymbol{S}_w\boldsymbol{w}}$**作为分式求梯度：**

$$\frac{\partial J(\boldsymbol{w})}{\partial \boldsymbol{w}}=\frac{(\boldsymbol{w}^T\boldsymbol{S}_w\boldsymbol{w})2\boldsymbol{w}^T\boldsymbol{S}_B-(\boldsymbol{w}^T\boldsymbol{S}_B\boldsymbol{w})2\boldsymbol{w}^T\boldsymbol{S}_w}{(\boldsymbol{w}^T\boldsymbol{S}_w\boldsymbol{w})^2}=\boldsymbol{0}^T$$

$$\frac{(\boldsymbol{w}^T\boldsymbol{S}_w\boldsymbol{w})2\boldsymbol{w}^T(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)^T-(\boldsymbol{w}^T(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)^T\boldsymbol{w})2\boldsymbol{w}^T\boldsymbol{S}_w}{(\boldsymbol{w}^T\boldsymbol{S}_w\boldsymbol{w})^2}$$

$$=\frac{2\boldsymbol{w}^T(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)}{\boldsymbol{w}^T\boldsymbol{S}_w\boldsymbol{w}}\left((\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)^T-\frac{(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)^T\boldsymbol{w}\boldsymbol{w}^T\boldsymbol{S}_w}{\boldsymbol{w}^T\boldsymbol{S}_w\boldsymbol{w}}\right)$$

$$=\frac{2\boldsymbol{w}^T(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)}{\boldsymbol{w}^T\boldsymbol{S}_w\boldsymbol{w}}\left((\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)^T-\frac{\boldsymbol{w}^T(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)}{\boldsymbol{w}^T\boldsymbol{S}_w\boldsymbol{w}}\boldsymbol{w}^T\boldsymbol{S}_w\right)=\boldsymbol{0}^T$$

由于$\frac{\boldsymbol{w}^T(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)}{\boldsymbol{w}^T\boldsymbol{S}_w\boldsymbol{w}}$为标量，令其为$a$，

$$2a((\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)^T-a\boldsymbol{w}^T\boldsymbol{S}_w)=\boldsymbol{0}^T$$

$$(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)^T-a\boldsymbol{w}^T\boldsymbol{S}_w=\boldsymbol{0}^T$$

$$a\boldsymbol{w}^T\boldsymbol{S}_w=(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)^T$$

$$a\boldsymbol{S}_w^T\boldsymbol{w}=\boldsymbol{\mu}_1-\boldsymbol{\mu}_2$$

$\boldsymbol{S}_w$为对称矩阵，则$\boldsymbol{S}_w^T=\boldsymbol{S}_w$

如果$\boldsymbol{S}_w^{-1}$存在，则：$\boldsymbol{w}=\frac{1}{a}\boldsymbol{S}_w^{-1}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)$

即：$\boldsymbol{w}^*=\boldsymbol{S}_w^{-1}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_2)$