- Fisher 准则函数

Fisher 准则函数定义为：

$$J_{\mathrm{F}}(\boldsymbol{w})=\frac{(\tilde{m}_1-\tilde{m}_2)^2}{\tilde{S}_1+\tilde{S}_2}$$

其中，$(\tilde{m}_1-\tilde{m}_2)$ 是两类均值之差，$\tilde{S}_i$ 是样本类内离散度。显然，应该使 $J_{\mathrm{F}}(\boldsymbol{w})$的分子尽可能大而分母尽可能小，即应寻找使 $J_{\mathrm{F}}(\boldsymbol{w})$尽可能大的 $\boldsymbol{w}$ 作为投影方向。但上式中并不显含 $\boldsymbol{w}$，因此须设法将 $J_{\mathrm{F}}(\boldsymbol{w})$变成 $\boldsymbol{w}$ 的显函数。

由各类样本的均值可推出：

$$\tilde{m}_i=\frac{1}{N_i}\sum_{y\in\Gamma_i'}y=\frac{1}{N_i}\sum_{x\in\Gamma_i}\boldsymbol{w}^T\boldsymbol{x}=\boldsymbol{w}^T\frac{1}{N_i}\left(\sum_{x\in\Gamma_i}\boldsymbol{x}\right)=\boldsymbol{w}^T\boldsymbol{m}_i$$

这样，Fisher 准则函数 $J_{\mathrm{F}}(\boldsymbol{w})$的分子可写成：

$$\begin{aligned}\left(\tilde{m}_1-\tilde{m}_2\right)^2&=\left(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{m}_1-\boldsymbol{w}^{\mathrm{T}}\boldsymbol{m}_2\right)^2\\&=\left(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{m}_1-\boldsymbol{w}^T\boldsymbol{m}_2\right)\left(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{m}_1-\boldsymbol{w}^{\mathrm{T}}\boldsymbol{m}_2\right)^{\mathrm{T}}\\&=\left(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{m}_1-\boldsymbol{w}^T\boldsymbol{m}_2\right)\left(\boldsymbol{m}_1^{\mathrm{T}}\boldsymbol{w}-\boldsymbol{m}_2^{\mathrm{T}}\boldsymbol{w}\right)\\&=\boldsymbol{w}^{\mathrm{T}}\left(\boldsymbol{m}_1-\boldsymbol{m}_2\right)\left(\boldsymbol{m}_1-\boldsymbol{m}_2\right)^{\mathrm{T}}\boldsymbol{w}=\boldsymbol{w}^{\mathrm{T}}\boldsymbol{S}_b\boldsymbol{w}\end{aligned}$$

现在再来考察 $J_{\mathrm{F}}(\boldsymbol{w})$的分母与 $\boldsymbol{w}$ 的关系：

$$\begin{aligned}\tilde{S}_i&=\sum_{y\in\Gamma_i'}\left(y-\tilde{m}_i\right)^2=\sum_{x\in\Gamma_i}\left(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}-\boldsymbol{w}^{\mathrm{T}}\boldsymbol{m}_i\right)^2\\&=\boldsymbol{w}^{\mathrm{T}}\left[\sum_{\boldsymbol{x}\in\Gamma_i}\left(\boldsymbol{x}-\boldsymbol{m}_i\right)\left(\boldsymbol{x}-\boldsymbol{m}_i\right)^{\mathrm{T}}\right]\boldsymbol{w}\\&=\boldsymbol{w}^{\mathrm{T}}\left[\sum_{\boldsymbol{x}\in\Gamma_i}\left(\boldsymbol{x}-\boldsymbol{m}_i\right)\left(\boldsymbol{x}-\boldsymbol{m}_i\right)^{\mathrm{T}}\right]\boldsymbol{w}\\&=\boldsymbol{w}^{\mathrm{T}}\boldsymbol{S}_i\boldsymbol{w}\end{aligned}$$

因此，

$$\tilde{S}_1+\tilde{S}_2=\boldsymbol{w}^{\mathrm{T}}(\boldsymbol{S}_1+\boldsymbol{S}_2)\boldsymbol{w}=\boldsymbol{w}^{\mathrm{T}}\boldsymbol{S}_w\boldsymbol{w}$$

将上述各式代入 $J_{\mathrm{F}}(\boldsymbol{w})$，可得：

$$J_{\mathrm{F}}(\boldsymbol{w})=\frac{\boldsymbol{w}^{\mathrm{T}}\boldsymbol{S}_{b}\boldsymbol{w}}{\boldsymbol{w}^{\mathrm{T}}\boldsymbol{S}_{w}\boldsymbol{w}}$$

其中 $\boldsymbol{S}_{\mathrm{b}}$ 为样本类间离散度矩阵，$\boldsymbol{S}_{\mathrm{w}}$ 为总样本类内离散度矩阵。