Fisher线性判别 LDA的一种，二分类 n维→1维

次优分类器

$g(x) = w^T x + w_0$

关键是：估计出 w 和 w_0

D包含m个n维样本

第一类 D_1 N_1

第二类 D_2 N_2

$y_i = w^T x_i$

$\|w\|=1$ 时 y_i 是 x_i 在直线上的投影

n维X空间中：

$$\mu_i = \frac{1}{N_i} \sum_{x_j \in D_i} x_j$$

类内离散度矩阵

$$S_i = \sum_{x_j \in D_i} (x_j - \mu_i)(x_j - \mu_i)^T$$

满秩

||

$S_w = S_1 + S_2$

S_w 是对称半正定矩阵，$m>n$ 通常是非奇异的

类间离散度矩阵

$$S_b = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T$$

S_b 是对称半正定矩阵

一维Y空间

$$\tilde{\mu}_i = \frac{1}{N_i} \sum_{y_j \in D_i} y_j$$

$$\tilde{S}_i^2 = \sum_{y_j \in D_i} (y_j - \tilde{\mu}_i)^2$$

$$\tilde{S}_w = \tilde{S}_1^2 + \tilde{S}_2^2$$

$$\tilde{S}_b = (\tilde{\mu}_1 - \tilde{\mu}_2)^2$$

目标函数

$$J(w) = \frac{(\tilde{\mu}_1 - \tilde{\mu}_2)^2}{\tilde{S}_1^2 + \tilde{S}_2^2} \quad max$$

$$w^* = \underset{w}{argmax}\, J(w)$$

$$\tilde{\mu}_i = w^T \mu_i$$

$$(\tilde{\mu}_1 - \tilde{\mu}_2)^2 = w^T (\tilde{\mu}_1 - \tilde{\mu}_2)(\tilde{\mu}_1 - \tilde{\mu}_2)^T w$$

$$= w^T S_b w$$

$$\tilde{S}_1^2 + \tilde{S}_2^2 = \sum_{y_j \in D_1}^{m} (y_j - \tilde{\mu}_1)^2 + \sum_{y_j \in D_2}^{m} (y_j - \tilde{\mu}_2)^2$$

$$= \sum_{i=1}^{2} \sum_{x_j \in D_i}^{m} (w^T x_j - w^T \mu_i)^2$$

$$= \sum_{i=1}^{2} \sum^{m} w^T (x_j - \mu_i)(x_j - \mu_i)^T w$$

$$= \sum_{i=1}^{2} w^T S_i w = w^T S_w w$$