1. N-D Fisher's Linear Discriminant (FLD)

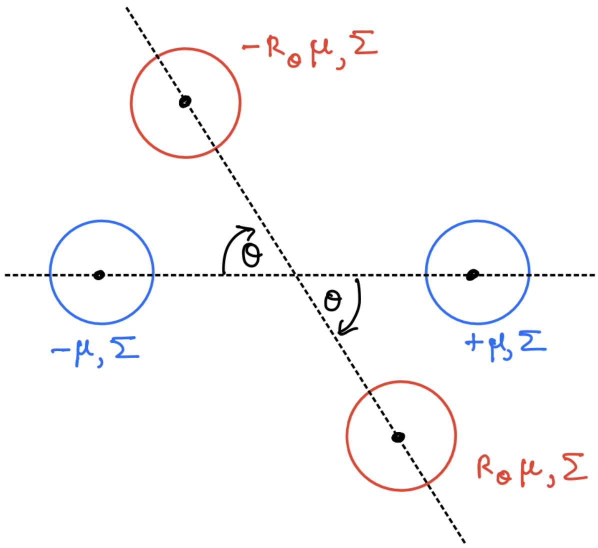

Figure 1: A rough illutration of the in-distribution task (blue) and the out-of-distribution task (red)

Consider an in-distribution task and an out-of-distribution task specified by the distributions F_{in} and F_{out}, respectively. F_{in} is characterized by the class conditional densities,

$$f_{0,in} = \mathcal{N}(-\mu, \Sigma) \tag{1}$$

$$f_{1,in} = \mathcal{N}(\mu, \Sigma) \tag{2}$$

and F_{out} is characterized by the class conditional densities,

$$f_{0,out} = \mathcal{N}(-R_\theta \mu, \Sigma) \tag{3}$$

$$f_{1,out} = \mathcal{N}(R_\theta \mu, \Sigma) \tag{4}$$

where,

$$R_\theta = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \tag{5}$$

Suppose that we have n samples $S_{in} = \{X_i, Y_i\}_{i=1}^{n}$ drawn from F_{in} and m samples $S_{out} = \{X_j, Y_j\}_{j=1}^{m}$ drawn from F_{out}. The samples are class-balanced. We are interested in generalizing on the in-distribution task using both S_{in} and S_{out}.

Let M_0 and M_1 be the estimated means of classes 0 and 1 respectively. Note that each class comprises of samples from both in- and out-of-distribution tasks. Consider M_1, which is given by,

$$M_1 = \frac{\sum_{i=1}^{n/2} X_i + \sum_{j=1}^{m/2} X_j}{n/2 + m/2} \tag{6}$$

The mean and variance of M_1 are given by,

$$\mathbb{E}[M_1] = \frac{(nI + mR_\theta)\mu}{n+m} \tag{7}$$

$$\text{Var}[M_1] = \frac{2}{n+m}\Sigma \tag{8}$$

By the central limit theorem,

$$M_1 \sim \mathcal{N}\Big(\frac{(nI + mR_\theta)\mu}{n+m}, \frac{2}{n+m}\Sigma\Big) \tag{9}$$

Similary, for M_0,

$$M_0 \sim \mathcal{N}\Big(-\frac{(nI + mR_\theta)\mu}{n+m}, \frac{2}{n+m}\Sigma\Big) \tag{10}$$

The decision rule of the FLD is given by,

$$g(x) = \begin{cases} 1, & \omega^\top x > c \\ 0, & \text{otherwise} \end{cases} \tag{11}$$

where, $\omega = (\Sigma_0 + \Sigma_1)^{-1}(M_1 - M_0)$ and $c = \omega^\top(M_1 + M_0)/2$. ($\Sigma_1$ and Σ_1 are the variances of class 0 and 1 respectively). By letting, $h = (M_1 + M_0)/2$, the decision rule can be written as,

$$g(x) = \begin{cases} 1, & \omega^\top x > \omega^\top h \\ 0, & \text{otherwise} \end{cases} \tag{12}$$

Now, consider a test input X from the in-distribution task, i.e. $X \sim F_{in}$. The generalization risk $L(w, h)$ is then given by,

$$
\begin{aligned}
L(w,h) &= P[Y \neq g(X) | X = x] && (13)\\
&= \frac{1}{2}\big(P_{X\sim f_{1,in}}[\omega^\top X < \omega^\top h] + P_{X\sim f_{0,in}}[\omega^\top X > \omega^\top h]\big) && (14)\\
&= \frac{1}{2}\big(1 + P_{X\sim f_{1,in}}[\omega^\top X < \omega^\top h] - P_{X\sim f_{0,in}}[\omega^\top X < \omega^\top h]\big) && (15)
\end{aligned}
$$