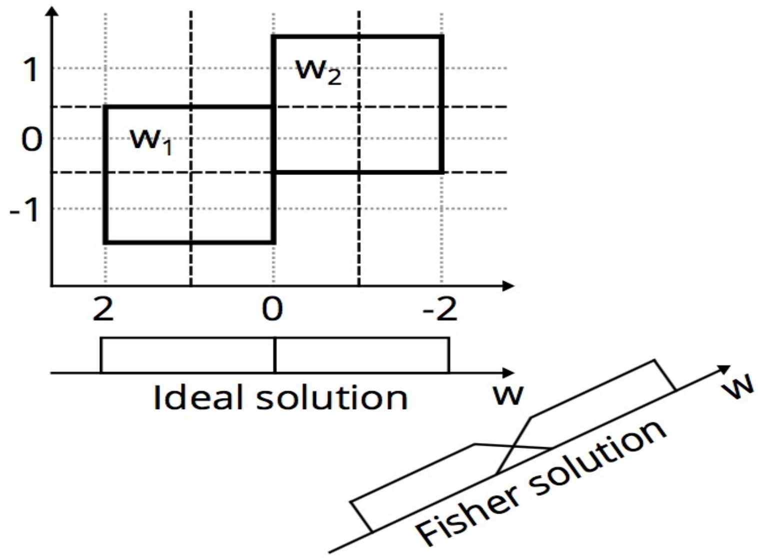
1 Problem 1

2 Problem 2

a)



Consider two distributions $p(x|w_1) \sim U(a,b)^2 = U(-1,1)^2$ and $p(x|w_2) \sim U(-1,1)^2$, then we define the expectation as:

$$\mu_1 = \begin{pmatrix} -1 \\ -0.5 \end{pmatrix} \qquad \text{and} \qquad \mu_2 = \begin{pmatrix} 1 \\ 0.5 \end{pmatrix}$$

We define X as a random variable for first dimension, Y as a random variable for second dimension, and we assume, that two variables are iid. So, we can say, $Cov(X,Y) = Cov(Y,X) = 0$, therefore, we know the variance of X and Y, $Var(X) = Var(Y) = \frac{1}{12}(a-b)^2 = \frac{1}{3}$, The covariance matrix is on the below.

$$\Sigma = \begin{bmatrix} Var(X) & Cov(X,Y) \\ Cov(Y,X) & Var(Y) \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{3} \end{bmatrix} = \frac{1}{3}I$$

We can compute w^*:

$$w^* = \Sigma_w^{-1}(\mu_2 - \mu_1)\Sigma = 3I\left(\begin{pmatrix} 1 \\ 0.5 \end{pmatrix} - \begin{pmatrix} -1 \\ -0.5 \end{pmatrix}\right) = 3\begin{pmatrix} 2 \\ 1 \end{pmatrix} \neq \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

where $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ would be the optimal Bayes solution (from the pic). The linear prediction in Bayes can easily obtained by $p(w_1|x) > p(w_2|x)$, which is different with Fisher discriminant analysis.

b)
When the two classes are generated by two d-dimensional Gaussian distributions,

$$
\begin{aligned}
p(x|w_1) &= \frac{1}{\sqrt{2\pi det(\Sigma_1)}} exp\left(-\frac{1}{2}(x-\mu_1)^\intercal \Sigma_1^{-1}(x-\mu_1)\right)\\
p(x|w_2) &= \frac{1}{\sqrt{2\pi det(\Sigma_2)}} exp\left(-\frac{1}{2}(x-\mu_2)^\intercal \Sigma_2^{-1}(x-\mu_2)\right)
\end{aligned}
$$

According to question a, if we decide w_1, therefore,

$$
\begin{aligned}
p(w_1|x) &> p(w_2|x)\\
\frac{p(x|w_1)p(w_1)}{p(x)} &> \frac{p(x|w_2)p(w_2)}{p(x)}\\
p(x|w_1)p(w_1) &> p(x|w_2)p(w_2)
\end{aligned}
$$

where we take the logarithm to simplify the computation:

$$
\begin{aligned}
\ln p(x|w_1) + \ln p(w_1) &> \ln p(x|w_2) + \ln p(w_2)\\
\ln p(x|w_1) - \ln p(x|w_2) &> \ln p(w_2) - \ln p(w_1)\\
-\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln det(\Sigma_1) - \frac{1}{2}(x-\mu_1)^\intercal \Sigma_1^{-1}(x-\mu_1)&\\
+\frac{1}{2}\ln 2\pi + \frac{1}{2}\ln det(\Sigma_2) + \frac{1}{2}(x-\mu_2)^\intercal \Sigma_2^{-1}(x-\mu_2) &> \ln p(w_2) - \ln p(w_1) \qquad \text{(PS: plug into above pdf)}
\end{aligned}
$$

We can identify the first part of the inequality as the mapping $\phi : \mathbb{R}^d \to \mathbb{R}$

$$
\phi(x) = (x-\mu_2)^\intercal \Sigma_2^{-1}(x-\mu_2) - (x-\mu_1)^\intercal \Sigma_1^{-1}(x-\mu_1)
$$

which is the optimal solutions in Bayes sense.