Exercise Sheet 2
Ex 1

1 (a) The decision boundary occurs where $g_A(\underline{x}) = g_B(\underline{x})$.
We will find the locus of such $\underline{x}$.

Note: $\underline{X} \sim \mathcal{N}(\mu, \Sigma) \Rightarrow p(\underline{X}=\underline{x}) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \exp(-\frac{1}{2}(\underline{x}-\underline{\mu})^T \Sigma^{-1} (\underline{x}-\underline{\mu}))$

$$\Rightarrow \log p(\underline{X}=\underline{x}) = \frac{p}{2}\log 2\pi - \frac{1}{2}\log|\Sigma| - \frac{1}{2}(\underline{x}-\underline{\mu})^T \Sigma^{-1} (\underline{x}-\underline{\mu}).$$

$\therefore$ $g_A(\underline{x}) = g_B(\underline{x}) \iff \log p(\underline{x} \mid y=A) + \log \pi_A = \log p(\underline{x} \mid y=B) + \log \pi_B$

$$\iff \log p(\underline{x} \mid y=A) - \log p(\underline{x} \mid y=B) + \log \frac{\pi_A}{\pi_B} = 0$$

$$\iff \frac{1}{2}\log|\Sigma_B| - \frac{1}{2}\log|\Sigma_A| + \frac{1}{2}(\underline{x}-\underline{\mu}_B)^T \Sigma_B^{-1} (\underline{x}-\underline{\mu}_B)$$

$$- \frac{1}{2}(\underline{x}-\underline{\mu}_A)^T \Sigma_A^{-1} (\underline{x}-\underline{\mu}_A) + \log \frac{\pi_A}{\pi_B} = 0 \quad (*)$$

Simplify one part of $(*)$:

$$(\underline{x}-\underline{\mu}_B)^T \Sigma_B^{-1} (\underline{x}-\underline{\mu}_B) - (\underline{x}-\underline{\mu}_A)^T \Sigma_A^{-1} (\underline{x}-\underline{\mu}_A)$$

$$= \underline{x}^T(\Sigma_B^{-1} - \Sigma_A^{-1})\underline{x} - \underline{\mu}_B^T \Sigma_B^{-1} \underline{x} + \underline{\mu}_B^T \Sigma_B^{-1} \underline{\mu}_B - \underline{x}^T \Sigma_B^{-1} \underline{\mu}_B$$

$$+ \underline{\mu}_A^T \Sigma_A^{-1} \underline{x} - \underline{\mu}_A^T \Sigma_A^{-1} \underline{\mu}_A + \underline{x}^T \Sigma_A^{-1} \underline{\mu}_A \quad (**)$$

Note that $\underline{x}^T \Sigma_A^{-1} \underline{\mu}_A$ is a scalar, ~~and~~ Σ_A is symmetric $\Rightarrow \Sigma_A^{-1}$ is symmetric $[\Sigma_A^{-1} = \Sigma_A^{-1\,T}]$ and $c^T = c$ for any scalar c. Thus

And equiv. for B

$$\underline{x}^T \Sigma_A^{-1} \underline{\mu}_A = (\underline{x}^T \Sigma_A^{-1} \underline{\mu}_A)^T = \underline{\mu}_A^T \Sigma_A^{-1} \underline{x}.$$

$\therefore (**) = x^T(\Sigma_B^{-1} - \Sigma_A^{-1})x - 2\mu_B^T\Sigma_B^{-1}x$

$+ 2\mu_A^T\Sigma_A^{-1}x + \mu_B^T\Sigma_B^{-1}\mu_B - \mu_A^T\Sigma_A^{-1}\mu_A$

Substituting into $(*)$:

$\frac{1}{2}x^T(\Sigma_B^{-1} - \Sigma_A^{-1})x$

$+ (\mu_A^T\Sigma_A^{-1} - \mu_B^T\Sigma_B^{-1})x$

$+\frac{1}{2}(\mu_B^T\Sigma_B^{-1}\mu_B) - \frac{1}{2}\mu_A^T\Sigma_A^{-1}\mu_A + \frac{1}{2}\log|\Sigma_B| - \frac{1}{2}\log|\Sigma_A|$

$+ \log\frac{\pi_A}{\pi_B} = 0$

So $\boxed{\Lambda = \frac{1}{2}(\Sigma_B^{-1} - \Sigma_A^{-1})}$

$w^T = \mu_A^T\Sigma_A^{-1} - \mu_B^T\Sigma_B^{-1} \Rightarrow \boxed{w = \Sigma_A^{-1}\mu_A - \Sigma_B^{-1}\mu_B}$

$\boxed{b = \frac{1}{2}[\mu_B^T\Sigma_B^{-1}\mu_B - \mu_A^T\Sigma_A^{-1}\mu_A + \log|\Sigma_B| - \log|\Sigma_A|] + \log\frac{\pi_A}{\pi_B}}$

Brief explanation:

Here, linear is used to mean affine, ie linear up to a translating constant (in this case b). A linear function f satisfies $f(\lambda x) = \lambda f(x)$. If $f(x) = w^T x$, then $f(\lambda x) = w^T(\lambda x) = \lambda w^T x = \lambda f(x)$, ie linearity fulfilled. However for a component $c(x) = x^T \Lambda x$ (ie a quadratic term)
$c(\lambda x) = (\lambda x)^T \Lambda (\lambda x) = \lambda^2 x^T \Lambda x = \lambda^2 c(x)$, ie quadratic scaling, not linear.

b) In this case, we have

$$\Lambda = \frac{1}{2}(\Sigma^{-1} - \Sigma^{-1}) = 0$$

$$\underline{w} = \Sigma^{-1}\mu_A - \Sigma^{-1}\mu_B = \Sigma^{-1}(\mu_A - \mu_B)$$

$$b = \frac{1}{2}\left[\mu_B^T \Sigma^{-1}\mu_B - \mu_A^T \Sigma^{-1}\mu_A + 2\log|\Sigma| + \log\frac{\pi_A}{\pi_B}\right].$$

In particular, $\Lambda = 0$, so we recover

$$\underline{w}^T\underline{x} + b = 0,$$

as desired.

2. ~~a)~~ See notebook