Exercise 01

1.

(a)

Let $g_A(\boldsymbol{x}) = g_B(\boldsymbol{x})$, we have:

$$\log(p(\boldsymbol{x}|y = A)) + \log \pi_A = \log(p(\boldsymbol{x}|y = B)) + \log \pi_B.$$

Thus,

$$(\boldsymbol{x} - \mu_A)^{\mathrm{T}} \Sigma_A^{-1} (\boldsymbol{x} - \mu_A) + \log \pi_A = (\boldsymbol{x} - \mu_B)^{\mathrm{T}} \Sigma_B^{-1} (\boldsymbol{x} - \mu_B) + \log \pi_B.$$

Expanding both sides gives,

$$\begin{aligned} x^{\mathrm{T}} \left(\Sigma_A^{-1} - \Sigma_B^{-1}\right) x &- 2x^{\mathrm{T}} \Sigma_A^{-1} \mu_A + 2x^{\mathrm{T}} \Sigma_B^{-1} \mu_B \\ &+ \left(\mu_A^{\mathrm{T}} \Sigma_A^{-1} \mu_A - \mu_B^{\mathrm{T}} \Sigma_B^{-1} \mu_B\right) \\ &+ \log \mid \Sigma_0 \mid - \log \mid \Sigma_1 \mid + \log \pi_A - \log \pi_B = 0. \end{aligned}$$

Comparing the equation above with what is given in (a) gives:

$$\begin{aligned} \Lambda &= \left(\Sigma_A^{-1} - \Sigma_B^{-1}\right), \\ \omega^{\mathrm{T}} &= -2\left(\mu_A^{\mathrm{T}} \Sigma_A^{-1} - \mu_B^{\mathrm{T}} \Sigma_B^{-1}\right), \\ b &= \mu_A^{\mathrm{T}} \Sigma_A^{-1} \mu_A - \mu_B^{\mathrm{T}} \Sigma_B^{-1} \mu_B + \log \mid \Sigma_0 \mid - \log \mid \Sigma_1 \mid + \log \frac{\pi_A}{\pi_B}. \end{aligned}$$

(b)

If $\Sigma_A = \Sigma_B$, two terms in Λ cancel out, which results in $\omega^{\mathrm{T}} x + b = 0$. Also, if we denote $\Sigma = \Sigma_A = \Sigma_B$, the weight ω and bias b becomes

$$\begin{aligned} \omega^{\mathrm{T}} &= -2\left(\mu_A^{\mathrm{T}} - \mu_B^{\mathrm{T}}\right) \Sigma^{-1}, \\ b &= \left(\mu_A^{\mathrm{T}} \Sigma_A^{-1} \mu_A - \mu_B^{\mathrm{T}} \Sigma_B^{-1} \mu_B\right) + \log \frac{\pi_A}{\pi_B}, \end{aligned}$$

respectively.

2.

(a) Decision Boundaries for QDA and LDA by Python

```
Q_qda:
 [[ 0.42016807 -0.42016807]
 [-0.42016807  0.42016807]]
w_qda:
 [[3.07692308 3.07692308]]
b_qda:
 [[0.8873032]]
w_lda:
 [[1.53846154 1.53846154]]
b_lda:
 [[0.]]
```

(b) Plot

See the results in jupyter-notebook.

(c) Result Analysis

Bigger diagonal elements in Σ_B indicates class B is more widespread than A, whilst the negative off-diagonal element furthermore implies a more diverged tendency of class B, which makes the QDA area of class B much bigger.