Mary Stirling Brown

Assume the class covariance matrices are equal. We can prove that the GDA decision boundary is linear.

If x is on the decision boundary, then:

$$p(y=1|x) = p(y=0|x)$$

$$\frac{p(x|y=1)\,p(y=1)}{p(x)} = \frac{p(x|y=0)\,p(y=0)}{p(x)}$$

$$p(x|y=1)\,p(y=1) = p(x|y=0)\,p(y=0), \quad \text{where} \quad p(x) \neq 0$$

We can expand these probabilities to get:

$$\frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} \exp\left(\frac{-1}{2}(x-\mu_1)^T\Sigma^{-1}(x-\mu_1)\right)\varphi = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} \exp\left(\frac{-1}{2}(x-\mu_0)^T\Sigma^{-1}(x-\mu_0)\right)(1-\varphi)$$

$$\exp\left(\frac{-1}{2}(x-\mu_1)^T\Sigma^{-1}(x-\mu_1)\right)\varphi = \exp\left(\frac{-1}{2}(x-\mu_0)^T\Sigma^{-1}(x-\mu_0)\right)(1-\varphi)$$

By taking the log of both sides, we get:

$$\log\left(\exp\left(\frac{-1}{2}(x-\mu_1)^T\Sigma^{-1}(x-\mu_1)\right)\varphi\right) = \log\left(\exp\left(\frac{-1}{2}(x-\mu_0)^T\Sigma^{-1}(x-\mu_0)\right)(1-\varphi)\right)$$

$$\log\left(\exp\left(\frac{-1}{2}\right)(x-\mu_1)^T\Sigma^{-1}(x-\mu_1)\right) + \log(\varphi) = \log\left(\exp\left(\frac{-1}{2}(x-\mu_0)^T\Sigma^{-1}(x-\mu_0)\right)\right) + \log(1-\varphi)$$

$$-\frac{1}{2}(x-\mu_1)^T\Sigma^{-1}(x-\mu_1) + \log(\varphi) = \frac{-1}{2}(x-\mu_0)^T\Sigma^{-1}(x-\mu_0) + \log(1-\varphi)$$

Let Σ^{-1} be $\begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$, and $[(x_1-\mu_{11}) \quad (x_2-\mu_{12}) \dots (x_n-\mu_{1n})]$

Then, $(x-\mu_1)^T\Sigma^{-1}(x-\mu_1) = (x_1-\mu_{11})\sigma_{11}(x_1-\mu_{11}) + (x_2-\mu_{12})\sigma_{12}(x_2-\mu_{12}) + \dots$

$$= \sigma_{11}x_1^2 - 2\mu_{11}x_1 + \mu_{11}^2\sigma_{11} + \sigma_{12}x_2^2 - 2\mu_{12}\sigma_{12}x_2 + \mu_{12}^2\sigma_{12} + \dots$$

and

$$(x-\mu_0)^T\Sigma^{-1}(x-\mu_0) = (x_1-\mu_{01})\sigma_{11}(x_1-\mu_{01}) + (x_2-\mu_{02})\sigma_{12}(x_2-\mu_{02}) + \dots$$

$$= \sigma_{11}x_1^2 - 2\mu_{01}\sigma x_1 + \mu_{01}^2\sigma_{11} + \sigma_{12}x_2^2 - 2\mu_{02}\sigma_{12}x_2 + \mu_{02}^2 + \sigma_{12} + \dots$$

Because the two Σ^{-1} are equal, we can set these equal:

$$(\sigma_{11} x_1^2 - 2\mu_{11}\sigma_{11}x_1 + \mu_{11}^2\sigma_{11} + \sigma_{11}x_2^2 - 2\mu_{12}\sigma_{12}x_2 + \mu_{12}^2\sigma_{12} + \dots) + \log(\varphi)$$
$$= (\sigma_{11}x_1^2 - 2\mu_{01}\sigma_{11}x_1 + \mu_{01}^2\sigma_{11} + \sigma_{11}x^2 - 2\mu_{02}\sigma_{12}x_2 + \mu_{02}^2\sigma_{12} + \dots) + \log(1-\varphi)$$

By cancelling out quadratic terms that appear or both sides:

$$(-2\mu_{11}\sigma_{11}x_1 + \mu_{11}^2\sigma_{11} - 2\mu_{12}\sigma_{12}x_2 + \mu_{12}^2\sigma_{12} + \dots) + \log(\varphi)$$
$$= (-2\mu_{01}\sigma_{11}x_1 + \mu_{01}^2\sigma_{11} - 2\mu_{02}\sigma_{12}x_2 + \mu_{02}^2\sigma_{12} + \dots) + \log(1-\varphi)$$

We can simplify these by creating new vectors:

$$\begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \vdots \end{bmatrix} \begin{bmatrix} x_1 & x_2 & x_3 & \dots & 1 \end{bmatrix} + \log(\varphi) = \begin{bmatrix} a_{10} \\ a_{20} \\ a_{30} \\ \vdots \end{bmatrix} \begin{bmatrix} x_1 & x_2 & x_3 & \dots & 1 \end{bmatrix} + \log(1-\varphi)$$

By getting all the terms on one side, we get:

$$0 = \begin{bmatrix} a_{10} \\ a_{20} \\ a_{30} \\ \vdots \end{bmatrix} \begin{bmatrix} x_1 & x_2 & x_3 & \dots & 1 \end{bmatrix} + \log(1-\varphi) - \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \vdots \end{bmatrix} \begin{bmatrix} x_1 & x_2 & x_3 & \dots & 1 \end{bmatrix} - \log(\varphi)$$

$$0 = \begin{bmatrix} a_{10} - a_{11} \\ a_{20} - a_{21} \\ a_{30} - a_{31} \\ \vdots \end{bmatrix} \begin{bmatrix} x_1 & x_2 & x_3 & \dots & 1 \end{bmatrix} + \log(1-\varphi) - \log(\varphi)$$

Thus, the boundary is a line since the quadratic terms are gone and we get a matrix filled with a values where $b = \log(1-\varphi) - \log(\varphi)$

$$0 = [a]\bar{x} + b$$