Lecture 16 (GDA Continuted)

If $\Sigma_0 = \Sigma_1 = \Sigma$, we get a linear separator and Σ is given as:

$$\Sigma = \frac{1}{m}\sum_{i=0}^{m}(x_i - \mu_{y_i})(x_i - \mu_{y_i})^T$$

1 How to Use Model

We know Θ, how do we obtain the prediction?

$$P(y|x;\Theta) = \frac{P(x|y;\Theta)P(y|\Theta)}{P(x;\Theta)}$$

Now, $P(x) = P(x|y=0;\Theta)\phi + P(x|y=1;\Theta)(1-\phi)$. The above expression will simplify to:

$$P(y|x;\Theta) = \frac{1}{1 + \dfrac{P(x|y=1;\Theta)(1-\phi)}{P(x|y=0;\Theta)\phi}}$$

$$P(y|x;\Theta) = \frac{1}{1+A}$$

Now, the decision boundary is given by $\log A = 0$. On simplifying this equation, we get:

$$\log A = \log\left(\frac{1-\phi}{\phi}\sqrt{\frac{|\Sigma_1|}{|\Sigma_0|}}\right) + \frac{1}{2}\left(x^T(\Sigma_1^{-1} - \Sigma_0^{-1})x - 2(\mu_1^T\Sigma_1^{-1} - \mu_0^T\Sigma_0^{-1})x + \mu_1^T\Sigma_1^{-1}\mu_1 - \mu_0^T\Sigma_0^{-1}\mu_0\right)$$

Therefore, the separator in general will be quadratic. If $\Sigma_0 = \Sigma_1 = \Sigma$, then $\log A$ simplifies to:

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

This form is similar to $\log A = \theta^T x$ where the x_0 term is separate

2 Comparison with Logistic Regression (for special case)

1. GDA makes stronger assumptions than logistic regression
2. GDA is less likely to overfit since it is more constrained
3. GDA is better only if assumptions are correct