

The maximum is obtained when  $l = c\Sigma^{-1}(\mu_1 - \mu_2)$ , where  $c$  is an arbitrary constant. The constant indicates that the solution is not unique but depends on a scaling factor. A commonly used scaling factor is to choose  $c$  so that the variance of each  $Y$  variable is unity. The maximum possible value is  $D^2 = (\mu_1 - \mu_2)^T \Sigma^{-1}(\mu_1 - \mu_2)$ , which is the Mahalanobis distance between the class means that we discussed earlier.