EEL6825: Pattern Recognition

Homework 1

Problem 25 in Chapter 2, on Page 72 of the following textbook: Richard O. Duda, Peter E. Hart, David G. Stork, "Pattern Classification", 2nd Edition, Wiley-Interscience, October 2000.

Problem 25 is given as below:

Fill in the steps in the derivation from Eq. 59 to Eqs. 60 - 65, which are given as follows:

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^t \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}_i) + \ln \boldsymbol{P}(\omega_i). \tag{59}$$

$$g_i(\mathbf{x}) = \mathbf{w}_i^t \mathbf{x} + w_{i0}, \tag{60}$$

where

$$\mathbf{w}_i = \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i \tag{61}$$

$$w_{i0} = -\frac{1}{2}\boldsymbol{\mu}_i^t \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_i + \ln \boldsymbol{P}(\omega_i). \tag{62}$$

$$\mathbf{w}^t(\mathbf{x} - \mathbf{x}_0) = 0, \tag{63}$$

where

$$\mathbf{w} = \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j) \tag{64}$$

$$\mathbf{x}_0 = \frac{1}{2}(\boldsymbol{\mu}_i + \boldsymbol{\mu}_j) - \frac{\ln[\boldsymbol{P}(\omega_i)/\boldsymbol{P}(\omega_j)]}{(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)^t \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j). \tag{65}$$

Ans. =

The discriminant function is given as –

$$g_i(x) = \ln p(x|\omega_i) + \ln P(\omega_i)$$

when $p(x|\omega_i) \sim N(\mu_i, \Sigma_i)$

and $p(x) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(x-\mu)^t \Sigma^{-1}(x-\mu)\right]$

The general form of discriminant function =

$$g_i(x) = -\frac{1}{2}(x-\mu_i)^t \Sigma_i^{-1}(x-\mu_i) - \frac{d}{2}\ln 2\pi$$
$$- \frac{1}{2}\ln|\Sigma_i| + \ln P(\omega_i)$$

when covariances are same, i.e. $\Sigma_i = \Sigma$

Terms marked as yellow could be removed

and

$$\boxed{g_i(x) = -\frac{1}{2}(x-\mu_i)^t \Sigma_i^{-1}(x-\mu_i) + \ln P(\omega_i)}$$

$$g_i(x) = \frac{-1}{2}\underbrace{(x-\mu_i)^t \Sigma_i^{-1}(x-\mu_i)}_{\text{Squared mahalanobis distance}} + \ln P(\omega_i)$$

Expanding squared mahalanobis distance $\Rightarrow$

$$(x-\mu_i)^t \Sigma^{-1} (x-\mu_i)$$

$$= x^t \Sigma^{-1} x - x^t \Sigma^{-1} \mu_i - \mu_i^t \Sigma_i^{-1} x + \mu_i^t \Sigma^{-1} \mu_i$$

Covariance matrix are symmetric hence its inverse is also symmetric.

$$\Sigma^t = \Sigma, \quad (\Sigma^{-1})^t = \Sigma^{-1}$$

Hence the expansion could be written as.

$$\Rightarrow \underline{x^t \Sigma^{-1} x} - 2(\Sigma^{-1}\mu_i)^t x + \mu_i^t \Sigma^{-1} \mu_i$$

The term $x^t \Sigma^{-1} x$ is an additive constant independent of class hence can be removed

$$\Rightarrow -2(\Sigma^{-1}\mu_i)^t x + \mu_i^t \Sigma^{-1} \mu_i$$

Thus Squared malhalanobin distance is expanded to ⇒

$$(x-\mu_i)^t \Sigma^{-1}(x-\mu_i) =$$

$$-2(\Sigma^{-1}\mu_i)^t x + \mu_i^t \Sigma^{-1}\mu_i$$

Putting this value in $g_i(x)$ we get =

$$g_i(x) = \frac{-1}{2}(x-\mu_i)^t \Sigma_i^{-1}(x-\mu_i) + \ln P(\omega_i)$$

$$g_i(x) = \frac{-1}{2}\left[-2(\Sigma^{-1}\mu_i)^t x + \mu_i^t \Sigma^{-1}\mu_i\right] + \ln P(\omega_i)$$

$$g_i(x) = (\Sigma^{-1}\mu_i)^t x - \frac{1}{2}\mu_i^t \Sigma^{-1}\mu_i + \ln P(\omega_i)$$

linear discriminant function $g_i(x)$ is given as =

$$g_i(x) = \omega_i^t x + \omega_{i0}$$

and we obtained

$$g_i(x) = (\Sigma^{-1}\mu_i)^t x - \frac{1}{2}\mu_i^t \Sigma^{-1}\mu_i + \ln P(\omega_i)$$

After comparing both equations we get =

$$w_i = \Sigma^{-1} \mu_i$$

$$w_{i0} = -\frac{1}{2} \mu_i^t \Sigma^{-1} \mu_i + \ln P(\omega_i)$$

Part-2 =

we know that Decision boundary:

$$g_i(x) = g_j(x)$$

$$\Rightarrow (\Sigma^{-1}\mu_i)^t x - \frac{1}{2} \mu_i^t \Sigma^{-1} \mu_i + \ln P(\omega_i)$$

$$= (\Sigma^{-1}\mu_j)^t x - \frac{1}{2} \mu_j^t \Sigma^{-1} \mu + \ln P(\omega_j)$$

$$\Rightarrow \Sigma^{-1}\mu_i^t x - \frac{1}{2} \mu_i^t \Sigma^{-1} \mu_i + \ln P(\omega_i)$$

$$= \Sigma^{-1}\mu_j^t x - \frac{1}{2} \mu_j^t \Sigma^{-1} \mu + \ln P(\omega_j)$$

$$\Rightarrow (\mu_i^t - \mu_j^t)\,\Sigma^{-1}x - \frac{1}{2}\mu_i^t\Sigma^{-1}\mu_i + \frac{1}{2}\mu_j^t\Sigma^{-1}\mu_j + \ln\left(\frac{\omega_i}{\omega_j}\right) = 0$$

$$\Rightarrow (\mu_i - \mu_j)\,\Sigma^{-1}\left[x - \frac{1}{2}(\mu_i - \mu_j) + \frac{\ln\left[\frac{P(\omega_i)}{P(\omega_j)}\right](\mu_i - \mu_j)}{(\mu_i - \mu_j)^t\Sigma^{-1}(\mu_i - \mu_j)}\right] - \frac{1}{2}\mu_j^T\Sigma^{-1}\mu_i + \frac{1}{2}\mu_i^t\Sigma^{-1}\mu_j = 0$$

we know that =

$$\boxed{-\frac{1}{2}\mu_j^T\Sigma^{-1}\mu_i + \frac{1}{2}\mu_i^t\Sigma^{-1}\mu_j = 0}$$

This is the form of a linear discriminant

$$\omega^t(x - x_0) = 0$$

where weight (ω) and bias (x_0) are =

$$\omega = \Sigma^{-1}(\mu_i - \mu_j)$$

$$x_0 = \frac{1}{2}(\mu_i + \mu_j) - \frac{\ln\left[P(\omega_i)/P(\omega_j)\right](\mu_i - \mu_j)}{(\mu_i - \mu_j)^t\Sigma^{-1}(\mu_i - \mu_j)}$$