HW #1 Solns., EE556, Fa

#1) $g(\underline{x}) = \ln\left(\dfrac{f(\underline{x}/w_1)}{f(\underline{x}/w_2)}\right) = 0 \Rightarrow$

$$-\frac{1}{2\sigma^2}\|\underline{x}-\underline{\mu}_1\|^2 + \frac{1}{2\sigma^2}\|\underline{x}-\underline{\mu}_2\|^2 = 0$$

or

$$\|\underline{x}-\underline{\mu}_1\|^2 = \|\underline{x}-\underline{\mu}_2\|^2$$

or

$$-2\underline{x}^T\underline{\mu}_1 + \|\underline{\mu}_1\|^2 + 2\underline{x}^T\underline{\mu}_2 - \|\underline{\mu}_2\|^2 = 0$$

or

$$\underline{x}^T(\underline{\mu}_2-\underline{\mu}_1) + \underbrace{\frac{\|\underline{\mu}_1\|^2-\|\underline{\mu}_2\|^2}{2}}_{-\frac{(\underline{\mu}_2+\underline{\mu}_1)^T(\underline{\mu}_2-\underline{\mu}_1)}{2}} = 0$$

or

$$\left(\underline{x} - \left(\frac{\underline{\mu}_1+\underline{\mu}_2}{2}\right)\right)^T(\underline{\mu}_2-\underline{\mu}_1) = 0$$

#2) $y = \left(\underline{x} - \left(\frac{\underline{\mu}_1 + \underline{\mu}_2}{2}\right)\right)^T(\underline{\mu}_2 - \underline{\mu}_1)$

$y \geq 0 \Rightarrow$ "decide w_2"

else "decide w_1"

$$y = x^T(\underline{\mu}_2 - \underline{\mu}_1) - \left(\frac{\|\underline{\mu}_2\|^2 - \|\underline{\mu}_1\|^2}{2}\right)$$

Q: How is y distributed ?

A: consider y given $\underline{x}$ is from class w_2 :

$$f_{y/w_2}(y/w_2) \sim N(\mu_{y/2}, \sigma^2_{y/2})$$

$$\mu_{y/2} = E[y/w_2] = E[\underline{x}^T(\underline{\mu}_2 - \underline{\mu}_1)/w_2] - \frac{(\|\underline{\mu}_2\|^2 - \|\underline{\mu}_1\|^2)}{2}$$

$$= \underline{\mu}_2^T(\underline{\mu}_2 - \underline{\mu}_1) - \frac{(\underline{\mu}_2 + \underline{\mu}_1)^T(\underline{\mu}_2 - \underline{\mu}_1)}{2}$$

$$= \frac{1}{2}(\underline{\mu}_2 - \underline{\mu}_1)^T(\underline{\mu}_2 - \underline{\mu}_1) = \frac{1}{2}\|\underline{\mu}_2 - \underline{\mu}_1\|^2$$

Also, $E[(y - \mu_{y/2})^2/w_2] =$

$$E\Bigg[\Bigg(\underline{x}^T(\underline{\mu}_2 - \underline{\mu}_1) - (\underline{\mu}_2 - \underline{\mu}_1)^T\frac{(\underline{\mu}_1 + \underline{\mu}_2)}{2} - \underbrace{\frac{(\underline{\mu}_2 - \underline{\mu}_1)^T(\underline{\mu}_2 - \underline{\mu}_1)}{2}}_{\mu_{y/2}}\Bigg)^2/w_2\Bigg]$$

$$= E\left[\left(\underline{x}^T(\underline{\mu}_2 - \underline{\mu}_1) - \frac{(\underline{\mu}_2 - \underline{\mu}_1)^T}{2}(2\underline{\mu}_2)\right)^2/w_2\right]$$

$$= E\left[\left((\underline{\mu}_2 - \underline{\mu}_1)^T(\underline{x} - \underline{\mu}_2)\right)^2/w_2\right]$$

$$= (\underline{\mu}_2 - \underline{\mu}_1)^T E\left[(\underline{x} - \underline{\mu}_2)(\underline{x} - \underline{\mu}_2)^T/w_2\right](\underline{\mu}_2 - \underline{\mu}_1)$$

$$= \|\underline{\mu}_2 - \underline{\mu}_1\|^2 \cdot \sigma^2$$

So:

$$\text{Prob}[\text{error}] = \frac{1}{2} \cdot \frac{1}{\sqrt{2\pi}\,\sigma_{Y/2}} \int_{-\infty}^{0} e^{-\frac{(Y-\mu_{Y/2})^2}{2\sigma_{Y/2}^2}}\, dy$$

$$+ \frac{1}{2} \cdot \frac{1}{\sqrt{2\pi}\,\sigma_{Y/1}} \int_{0}^{\infty} e^{-\frac{(Y-\mu_{Y/1})^2}{2\sigma_{Y/1}^2}}\, dy$$

by symmetry

$$= \frac{1}{\sqrt{2\pi}\,\sigma_{Y/2}} \int_{-\infty}^{0} e^{-(Y-\mu_{Y/2})^2/2\sigma_{Y/2}^2}\, dy$$

Let $r = \dfrac{Y - \mu_{Y/2}}{\sigma_{Y/2}} \Rightarrow \sigma_{Y/2}\, dr = dy$

Can then rewrite:

$$\text{Prob}[\text{error}) = \frac{1}{\sqrt{2\pi}\,\sigma_{Y/2}} \int_{-\infty}^{-\mu_{Y/2}/\sigma_{Y/2}} e^{-\frac{r^2}{2}}\, \sigma_{Y/2}\, dr$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\|\underline{\mu}_2 - \underline{\mu}_1\|/2\sigma} e^{-r^2/2}\, dr$$

$$= \underbrace{\frac{1}{\sqrt{2\pi}} \int_{\|\underline{\mu}_2 - \underline{\mu}_1\|/2\sigma}^{\infty} e^{-r^2/2}\, dr}_{\text{can calculate via standard tables for normal c.d.f.}}$$

#3) $g_i(\underline{x}) = \ln\left(f(\underline{x}/w_i)\,P[w_i]\right)$

$$= -\frac{1}{2}(\underline{x}-\underline{\mu}_i)^T\Sigma^{-1}(\underline{x}-\underline{\mu}_i) + \left\{\begin{array}{l}\text{nonessential}\\ \text{terms (in}\\ \text{this case)}\end{array}\right\}$$

$g_1(\underline{x}) - g_2(\underline{x}) =$

$$\underbrace{-\tfrac{1}{2}\underline{x}^T\Sigma^{-1}\underline{x} + \tfrac{1}{2}\underline{x}^T\Sigma^{-1}\underline{x}}_{\text{cancel}} + \tfrac{1}{2}\underline{\mu}_1^T\Sigma^{-1}\underline{x} + \tfrac{1}{2}\underline{x}^T\Sigma^{-1}\underline{\mu}_1 - \tfrac{1}{2}\underline{\mu}_1^T\Sigma^{-1}\underline{\mu}_1$$
$$-\tfrac{1}{2}\underline{\mu}_2^T\Sigma^{-1}\underline{x} - \tfrac{1}{2}\underline{x}^T\Sigma^{-1}\underline{\mu}_2 + \tfrac{1}{2}\underline{\mu}_2^T\Sigma^{-1}\underline{\mu}_2$$

same

$\underline{\mu}_1^T\Sigma^{-1}\underline{x} = \underline{x}^T\Sigma^{-1}\underline{\mu}_1$ -- why?

A: $(\underline{\mu}_1^T\Sigma^{-1}\underline{x})^T = (\Sigma^{-1}\underline{x})^T(\underline{\mu}_1^T)^T$

$$= \underline{x}^T\underbrace{(\Sigma^{-1})^T}_{\text{symmetric}}\underline{\mu}_1 = x^T\Sigma^{-1}\mu_1$$

$\Rightarrow$

$$g_1(\underline{x}) - g_2(\underline{x}) = \underline{x}^T\Sigma^{-1}\underline{\mu}_1 - \tfrac{1}{2}\underline{\mu}_1^T\Sigma^{-1}\underline{\mu}_1 - \underline{x}^T\Sigma^{-1}\underline{\mu}_2 + \tfrac{1}{2}\underline{\mu}_2^T\Sigma^{-1}\underline{\mu}_2$$

$$= \underbrace{\underline{x}^T(\Sigma^{-1}\underline{\mu}_1 - \Sigma^{-1}\underline{\mu}_2)}_{x^T\Sigma^{-1}(\underline{\mu}_1-\underline{\mu}_2)} \underbrace{- \tfrac{1}{2}\underline{\mu}_1^T\Sigma^{-1}\underline{\mu}_1 + \tfrac{1}{2}\underline{\mu}_2^T\Sigma^{-1}\underline{\mu}_2}_{=\ \frac{1}{2}(\underline{\mu}_2-\underline{\mu}_1)^T\Sigma^{-1}(\underline{\mu}_1+\underline{\mu}_2)}$$

$$= (\underline{\mu}_1-\underline{\mu}_2)^T\Sigma^{-1}\left(\underline{x} - \left(\frac{\underline{\mu}_1+\underline{\mu}_2}{2}\right)\right) = 0$$

#4)

$$P[x/w_1] \sim N\left(\begin{pmatrix}2\\0\end{pmatrix}, \begin{pmatrix}1&0\\0&1\end{pmatrix}\right)$$

$$P[x/w_3] \sim N\left(\begin{pmatrix}0\\2\end{pmatrix}, \begin{pmatrix}1&0\\0&1\end{pmatrix}\right)$$

$$P[x/w_2] \sim N\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1&0\\0&2\end{pmatrix}\right)$$

$$P[w_1] = P(w_2) = P(w_3) = 1/3$$

Let's find boundaries between all pairs of classes:

Between classes 1 + 3, we know (from class) that the decision boundary is the line:

$$(\mu_1 - \mu_2)^T(\underline{x} - \underline{x}_0) = 0.$$

We can choose $\underline{x}_0 = \begin{pmatrix}0\\0\end{pmatrix}$ in this case $\Rightarrow$

$$B_{13} = \left\{ \underline{x} = \begin{pmatrix}x_1\\x_2\end{pmatrix} \text{ s.t. } (1\ -1)\begin{pmatrix}x_1\\x_2\end{pmatrix} = 0 \right\}$$

The boundary between classes 1 + 2 is given by:

$$-\frac{1}{2}(\underline{x} - \underline{\mu}_1)^T \Sigma^{-1}(\underline{x} - \mu_1) - \frac{1}{2}\log|\Sigma_1| =$$

$$-\frac{1}{2}(\underline{x} - \mu_2)^T \Sigma^{-1}(\underline{x} - \mu_2) - \frac{1}{2}\log|\Sigma_2|$$

or

$$(x_1 - 2)^2 + x_2^2 + \underbrace{\tfrac{1}{2}\log 1}_{0} = x_1^2 + \frac{1}{2}x_2^2 + \log 2$$

or

$$\frac{1}{2}x_2^2 - 4x_1 + 4 - \log 2 = 0$$

$$B_{12} = \left\{ \underline{x} = \begin{pmatrix}x_1\\x_2\end{pmatrix} \text{ s.t. } \frac{1}{8}x_2^2 = \left(x_1 - 1 + \frac{\log 2}{4}\right)\right.$$

The boundary between classes 2 and 3 is given by:

$$x_1^2 + (x_2 - 2)^2 = x_1^2 + \frac{1}{2} x_2^2 + \log 2$$

or

$$\frac{1}{8} x_2^2 = x_2 - 1 + \frac{\log 2}{4}$$

$$B_{23} = \{ \underline{x} : x_2 = 0.96 \text{ or } x_2 = 7.04 \}$$

Rough sketch :

$A = (.94, .94)$

$B = (7.06, 7.06)$