Dennis Cornwell
CSE802 HW 2
1/23/13

8.

a. $1 = \int_{-\infty}^{\infty} 1/\pi b * 1/(1+(((x-a_i)/b)^2)$

$= 1/\pi b \int_{-\infty}^{\infty} 1/(1+(((x-a_i)/b)^2)$

$Let\ ((x-a_i)/b)^2 = y^2$

$= 1/\pi b \int_{-\infty}^{\infty} 1/(1+y^2)$

$= 1/\pi b * tan^{-1}(y)|{-\infty\, to\, \infty}$

$= 1/\pi b *(\pi/2 + \pi/2)$

$1 = 1/\pi * \pi$

b. $1/\pi b * 1/(1+(((x-a_1)/b)^2) = 1/\pi b * 1/(1+(((x-a_2)/b)^2)$

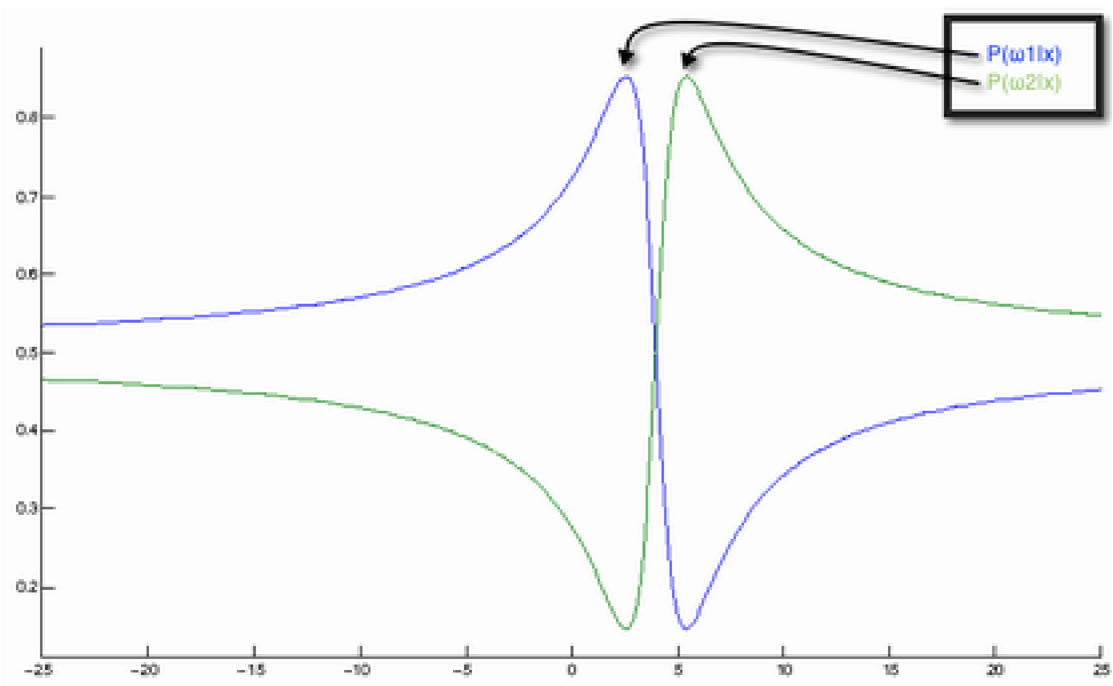
$1/(1+(((x-a_1)/b)^2) = 1/(1+(((x-a_2)/b)^2)$

$Given\ x = a_1 + a_2/2, Let\ k = x - a_1/b \Rightarrow x - a_2/b = -k$

$1/(1+(k)^2) = 1/(1+(-k)^2)$

$1/(1+k^2) = 1/(1+k^2)$

c.



d. $P(\omega_1|x)$ and $P(\omega_2|x)$ converge on 0.5 as x approaches $\pm\infty$ meaning that the further feature x gets from μ the less likely we will be to distinguish $\omega_1$from ω_2.

10.

Dennis Cornwell
CSE802 HW 2
1/23/13

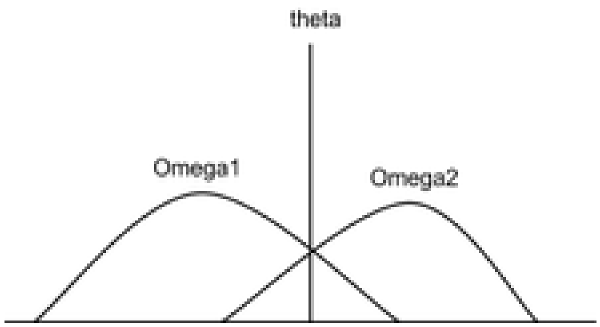
a. There are two types of error in this problem: $\omega_1$chosen with $x<\theta$ and ω_2 chosen with $x>\theta$. Let region $E_1=(-\infty,\theta)\ and\ E_2=(\theta,-\infty)$
$$P(error)=P(\omega_1|x)\ in\ E_1+P(\omega_2|x)\ in\ E_2$$
$$=p(x|\omega_1)P(\omega_1)\ in\ E_1+p(x|\omega_2)P(\omega_2)\ in\ E_2$$
$$=\int_{-\infty}^{\theta} P(\omega_1)p(x|\omega_1)dx + \int_{\theta}^{\infty} P(\omega_2)p(x|\omega_2)dx$$
$$=P(\omega_1)\int_{-\infty}^{\theta} p(x|\omega_1)dx + P(\omega_2)\int_{\theta}^{\infty} p(x|\omega_2)dx$$

b. $$P(error)=P(\omega_1)\int_{-\infty}^{\theta} p(x|\omega_1)dx + P(\omega_2)\int_{\theta}^{\infty} p(x|\omega_2)dx$$

$$1=P(\omega_1)p(\theta|\omega_1)d\theta+1-P(\omega_2)p(\theta|\omega_2)$$
$$0=P(\omega_1)p(\theta|\omega_1)d\theta-P(\omega_2)p(\theta|\omega_2)$$
$$P(\omega_2)p(\theta|\omega_2)=P(\omega_1)p(\theta|\omega_1)d\theta$$

c. No, the equation in part B only states that θ must lie where $p(\omega_1|x)$and $p(\omega_2|x)$meet. Knowing the curves for $p(\omega_1|x)$and $p(\omega_2|x)$would fully define θ.

d. Since the rule states only to select $\omega_1$if $x>\theta$ any curve similar to:

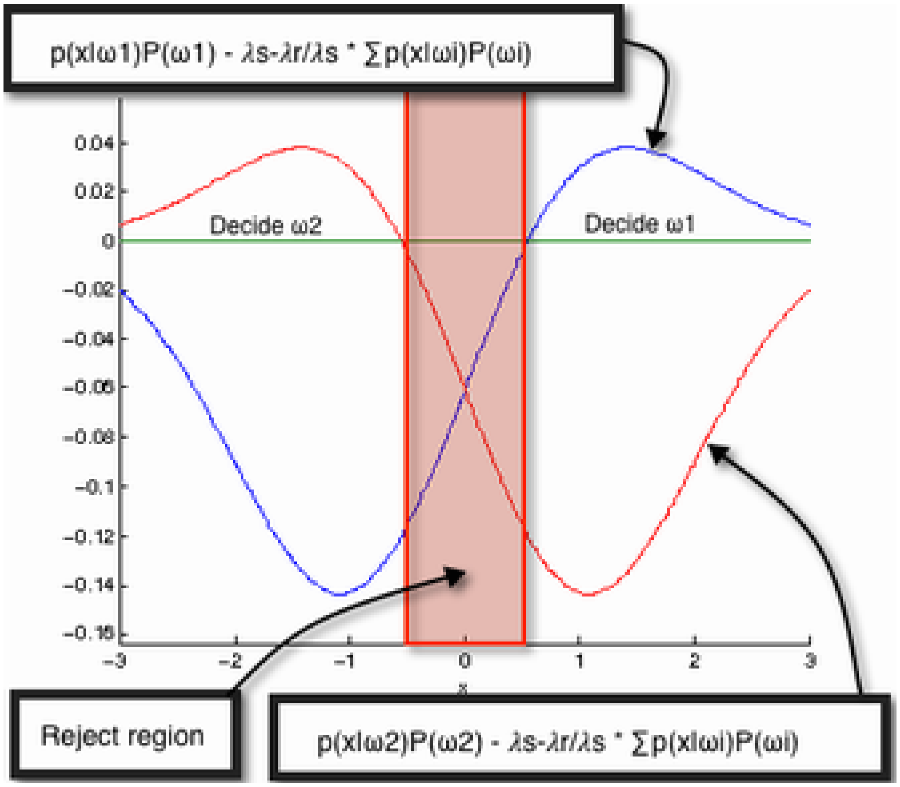


will maximize error rather than minimize it.

14.

a.
$$g(x)=p(x|\omega_i)P(\omega_i)-\frac{\lambda_s-\lambda_r}{\lambda_s}\sum_{j=1}^{c} p(x|\omega_j)P(\omega_j)$$
$$g(x)=P(\omega_i|x)-\frac{\lambda_s-\lambda_r}{\lambda_s}\sum_{j=1}^{c} P(\omega_j|x)$$
$$g(x)=P(\omega_i|x)-(1-\frac{\lambda_r}{\lambda_s})\sum_{j=1}^{c} P(\omega_j|x)$$
$$P(\omega_i|x)-(1-\frac{\lambda_r}{\lambda_s})\sum_{j=1}^{c} P(\omega_j|x)>0\Rightarrow select\ \omega_i, reject\ otherwise$$

b.



c. As the risk of selecting the reject option increases relative to the risk of making a substitution error, the number of values of x for which we would choose to reject decreases. Eventually (where $\lambda_r / \lambda_s > 0.5$), it becomes sub-optimal to reject any samples.

24.

a. $$\frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} * e^{\left(-\frac{1}{2}(x-\mu)^t\Sigma^{-1}(x-\mu)\right)}$$

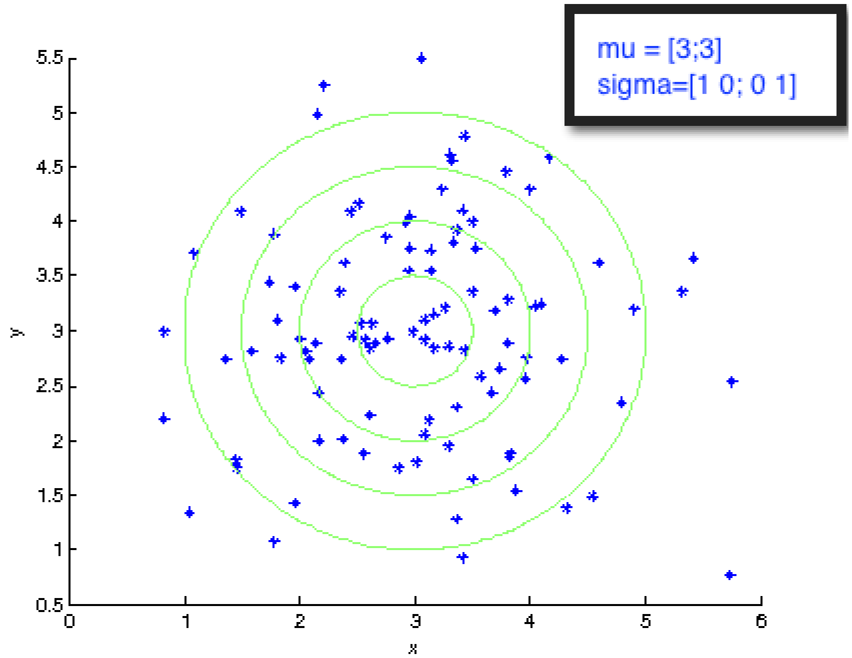
$$\frac{1}{(2\pi)^{d/2}\sqrt{\sigma 2}} * e^{\left(-\frac{1}{2} * \frac{1}{\sigma 2}(x-\mu)^t(x-\mu)\right)}$$

$$\frac{1}{(2\pi)^{d/2}\sigma} * e^{\left(-\frac{1}{2} * \frac{1}{\sigma 2}\sum_{j}^{d}(x_j-\mu_j)^2\right)}$$

$$\frac{1}{(2\pi)^{d/2}\sigma} * e^{\left(-\frac{1}{2} * \sum_{j}^{d}\left(\frac{x_j-\mu_j}{\sigma}\right)^2\right)}$$

$$\frac{1}{\prod_i^d \sqrt{2\pi\sigma}} * e^{(-\frac{1}{2} * \sum_j^d (\frac{x_j-\mu_j}{\sigma})^2)}$$

b. The contours of the constant density are basically a series of circles about (μ_1, μ_2) because the covariance matrix is $\sigma^2 * I$.



c. $M = \sqrt{(x-\mu)^t \Sigma^{-1} (x-\mu)}$

$= \sqrt{\sum_i^d (\frac{x_i-\mu_i}{\sigma_i})}$ ⇐ Becomes a euclidean distance

2.

a.

$(x_1-\mu_1)^t\Sigma^{-1}(x_1-\mu_1) = [x-1\, x-2]*[1\ -1; -1\, 2]\ [x-1; x-2] = x^2-4x+4$

$(x_2-\mu_2)^t\Sigma^{-1}(x_2-\mu_2) = [x-1\ x+2]*\ [2\ -1; -1\, 1]*[x-1; x+2] = \ = x^2-2x+1$

Max Likelihood

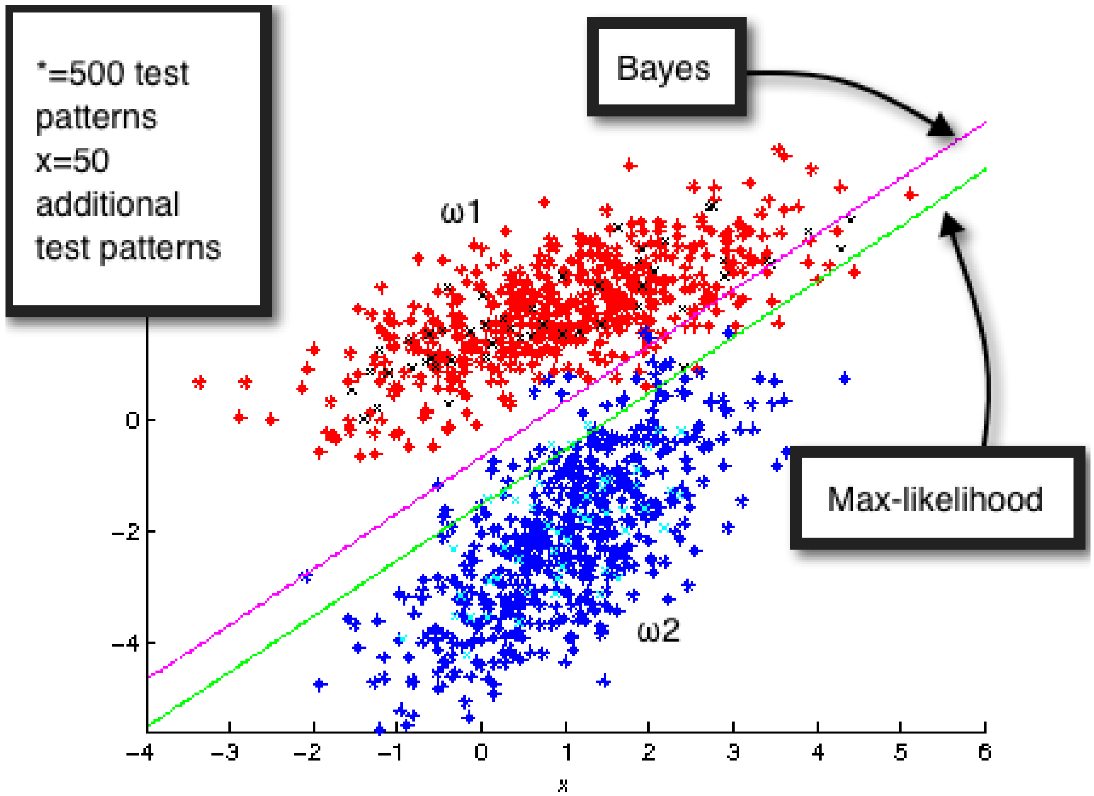
$g(x) = -\frac{1}{2}(x^2-4x+4) - ln(1) + \frac{1}{2}(x^2-2x+1) + ln(1)) < x \Rightarrow \omega_1, \omega_2\, otherwise$

$= -\frac{1}{2}(x^2-4x+4) + \frac{1}{2}(x^2-2x+1) < x \Rightarrow \omega_1, \omega_2\, otherwise$

Bayes

$$= -\frac{1}{2}(x^2-4x+4)+ln(P(\omega_1))+\frac{1}{2}(x^2-2x+1)-ln(P(\omega_2))$$
$$< x \Rightarrow \omega_1, \omega_2\, otherwise$$

b.



c.

i. Likelihood $E_{empirical} = E(\omega_1) + E(\omega_2) = 1/50 + 2/50 = 0.03$

ii. Bayes $E_{empirical} = E(\omega_1) + E(\omega_2) = 2/50 + 0/50 = 0.02$

d. $Av\Sigma = \frac{\Sigma_1 + \Sigma_2}{2} = [1.51;11.5], |\Sigma_1| = |\Sigma_2| = 1$

$$k(1/2) = 1/8 * (\mu_1 - \mu_2)^t Av\Sigma^{-1} * (\mu_2 - \mu_1) + \frac{1}{2} * ln(\frac{|Av\Sigma|}{\sqrt{|\Sigma_1| * |\Sigma_2|}})$$

$$k(1/2) = 1/8 * (\mu_1 - \mu_2)^t Av\Sigma^{-1} * (\mu_2 - \mu_1) + \frac{1}{2} * ln(|Av\Sigma|)$$

$$k(1/2) = 1/8 * (\mu_1 - \mu_2)^t Av\Sigma^{-1} * (\mu_2 - \mu_1) + 0.2231$$

$$k(1/2) = 1/8 * 19.2 + 0.2231$$

$$k(1/2) = 2.6231$$

$$P(error) \le \sqrt{P(\omega_1)P(\omega_2)} * e^{-k(1/2)}$$

$$= \sqrt{0.3 * 0.7} * e^{-2.6231}$$

$= 0.0333$ ⇐ Consistent with the 2-3% empirical error obtained in part C

3. $p(1) = 0.2*0.4 + 0.3*0.3 + 0.6*0.3 = 0.35$
$p(2) = 0.1*0.4 + 0.5*0.3 + 0.2*0.3 = 0.25$
$p(3) = 0.7*0.4 + 0.2*0.3 + 0.2*0.3 = 0.40$

$x \backslash \omega$	ω_1	ω_2	ω_3
1	$\frac{0.2*0.4}{0.35}$	$\frac{0.3*0.3}{0.35}$	$\frac{0.6*0.3}{0.35}$
2	$\frac{0.1*0.4}{0.25}$	$\frac{0.5*0.3}{0.25}$	$\frac{0.2*0.3}{0.25}$
3	$\frac{0.7*0.4}{0.4}$	$\frac{0.2*0.3}{0.4}$	$\frac{0.2*0.3}{0.4}$

Bayes:

$x \backslash \omega$	ω_1	ω_2	ω_3
1	0.229	0.257	**0.514**
2	0.16	**0.6**	0.24
3	**0.7**	0.15	0.15

Risk:

$x \backslash \alpha$	α_1	α_2	α_3	α_4
1	3.087	**0.743**	1.200	1.743
2	2.400	**0.400**	1.680	2.360
3	1.050	**0.850**	2.400	2.000

The Bayes decision rule leads to 3 different classes being picked at different values of x whereas the loss function is taken into account, it becomes most optimal to always chose class ω_2. This is likely because the loss when decision α_2 is wrong is considerably lower than the losses for rejection and selection errors in other decisions.