Stat 571
12-3-13

From last time:

① $Y = c(\vec{\mu}_1 - \vec{\mu}_2)' \Sigma^{-1} \vec{X}$ was Fisher's Discriminant Function

The <u>sample</u> <u>linear</u> dis<u>criminant</u> <u>function</u> is

$$y = c(\bar{X}_1 - \bar{X}_2)' S^{-1} \vec{X}$$

where $S = \dfrac{(n_1-1)S_1 + (n_2-1)S_2}{n_1+n_2-2}$

②

Let $\vec{X}_0$ be a new observation

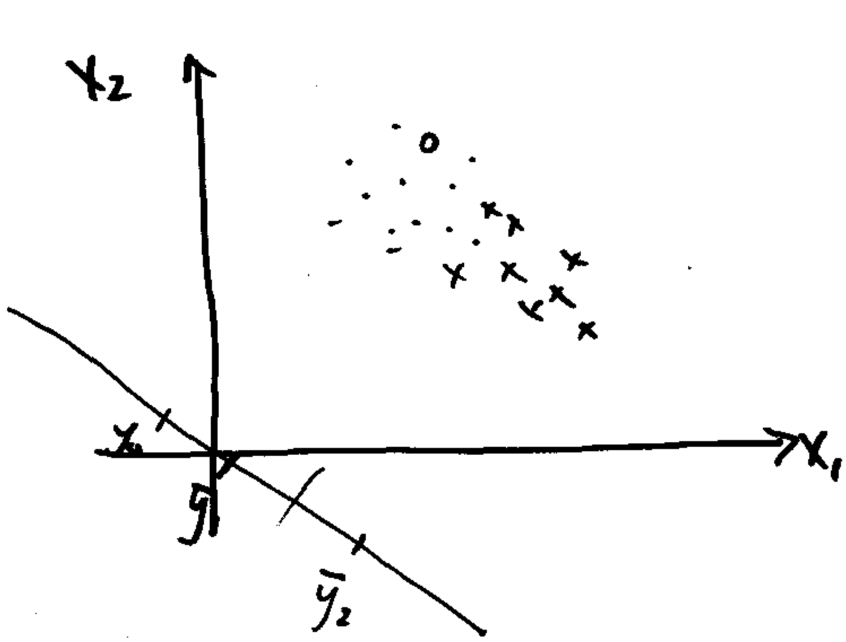
Compute $Y_0 = c(\vec{\mu}_1 - \vec{\mu}_2)' \Sigma^{-1} \vec{X}_0$ or

$$y_0 = c(\bar{X}_1 - \bar{X}_2)' S^{-1} \vec{X}_0$$

Let $\bar{y}_i$ be the average y value for all members of population i

Classify the new obs. according to the closest $\bar{y}_i$

3



4

let $P(2|1) = P(\text{classify as } \pi_2 \mid \text{actually } \pi_1)$

$P(1|2) = P(\text{classify as } \pi_1 \mid \text{actually } \pi_2)$

Let $c(2|1)$ and $c(1|2)$ be the associated costs

		Classify as π_1	Classify as π_2	
Actual	π_1	$p_1 P(1|1)$	$p_1 P(2|1)$	p_1
	π_2	$p_2 P(1|2)$	$p_2 P(2|2)$	p_2
				1

Costs

0	$c(2|1)$
$c(1|2)$	0

(5)

Use these tables to find the

ECM = expected cost of misclassification

$$ECM = c(2|1)\,p_1 P(2|1) + c(1|2)\,p_2 P(1|2)$$

You could the dividing point to minimize ECM.

If the costs are unknown, or about equal, then you could minimize the total probability of misclassification $= p_1 P(2|1) + p_2 P(1|2)$

(6)

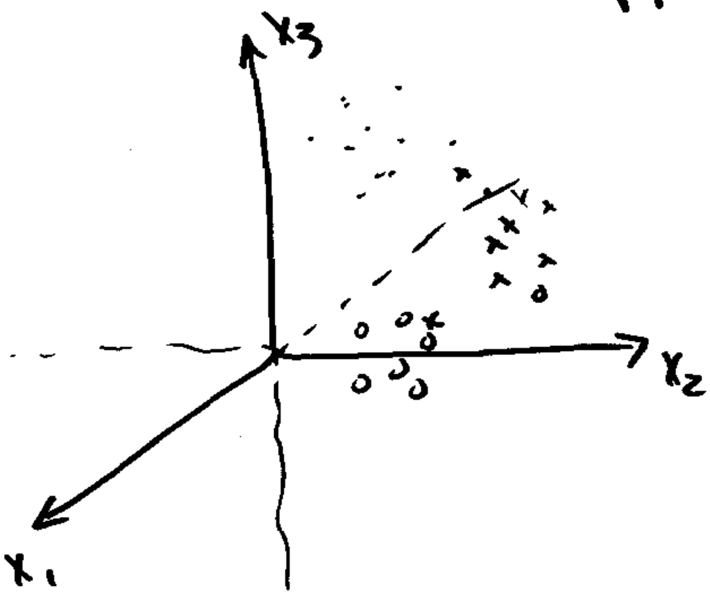
Part of the output will be the <u>confusion matrix</u>

	Predicted π_1	π_2	
True π_1	n_{1c}	n_{1m}	n_1
π_2	n_{2m}	n_{2c}	n_2
			n

c = correct
m = misclassified

<u>Defn</u>: The apparent error rate $= \dfrac{n_{1m} + n_{2m}}{n}$

What if there are more than 2 populations? (7)



For g groups, you can construct $g-1$ linear discriminant functions

(8)

Let $\bar{\mu}$ be the mean vector of all the populations combined.

$$\bar{\mu} = \frac{1}{g}\sum_{i=1}^{g}\vec{\mu}_i$$

Let $$B_\mu = \sum_{i=1}^{g}(\vec{\mu}_i - \bar{\mu})(\vec{\mu}_i - \bar{\mu})'$$

$$Y = \vec{a}'\vec{X}, \quad E(Y|\pi_i) = \vec{a}'\vec{\mu}_i$$

$$V(Y) = \vec{a}'\Sigma\vec{a}$$

9

Fisher suggested forming:

$$\frac{\sum_{i=1}^{g}\left(\vec{a}'\vec{\mu}_i - \vec{a}'\bar{\mu}\right)^2}{\vec{a}'\Sigma\vec{a}} = \frac{\sum_{i=1}^{g}\left(\vec{a}'(\vec{\mu}_i-\bar{\mu})\right)^2}{\vec{a}'\Sigma\vec{a}}$$

$$= \frac{\sum_{i=1}^{g}\vec{a}'(\vec{\mu}_i-\bar{\mu})(\vec{\mu}_i-\bar{\mu})'\vec{a}}{\vec{a}'\Sigma\vec{a}}$$

$$= \frac{\vec{a}'B_\mu\vec{a}}{\vec{a}'\Sigma\vec{a}}$$ Then maximize this.

10

The result: $\vec{a}_1 = \vec{e}_1$, the 1st eigenvector of $\Sigma^{-1}B$

$\vec{a}_2 = \vec{e}_2$, etc.