

253 Oct 13, 2015

## Review gaussian

1-var Meaning of  $\sigma$ . Why divide by sigma in exp.  
multi-var. Covariance matrix  $\Sigma$

$$\frac{1}{N} (\bar{X} - \mu)^T (\bar{X} - \mu) \quad \text{What's } \mu?$$

$$\exp\left(-\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu)\right)$$

$$\frac{1}{2} \left( \sqrt{\Sigma^{-1}} (x - \mu) \right)^T \left( \sqrt{\Sigma^{-1}} (x - \mu) \right)$$

$$X = \sqrt{\Sigma} (z - \mu)$$

$\uparrow$   
 $N(0, 1^2)$

How far away  
from  $\mu$  in " $z$ " units  
"Mahalanobis distance"

LDA Which group is nearest in Mahalanobis distance

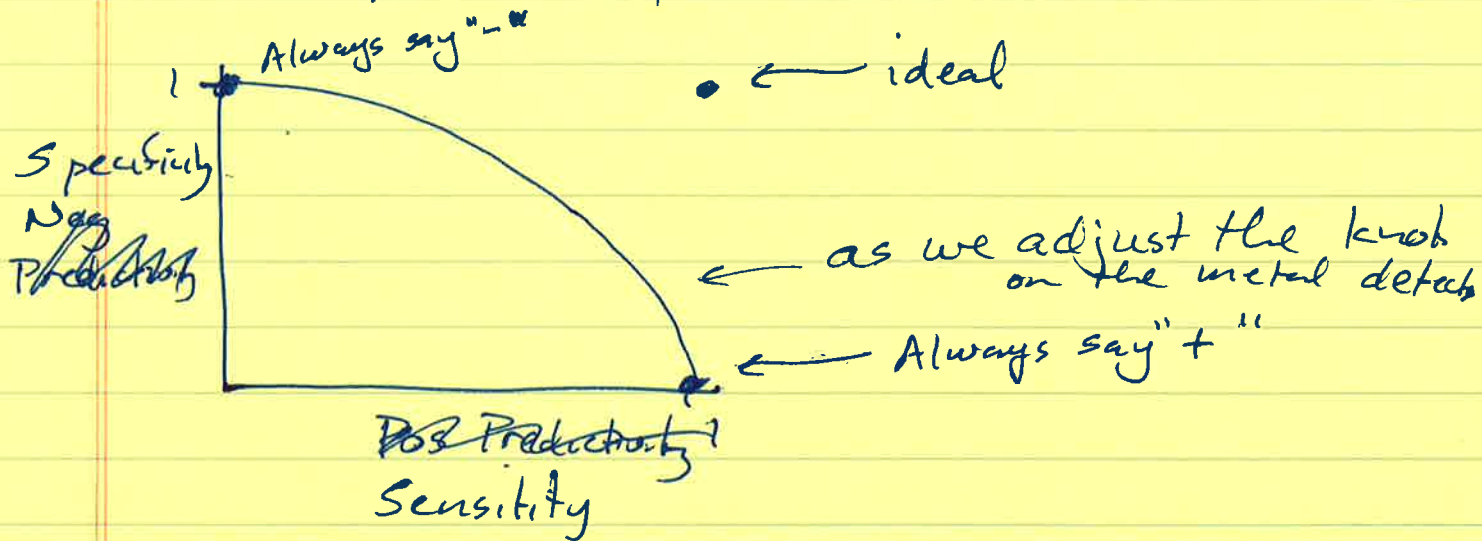
Fig. 4.6 interpret as ~~\*~~  $\rightarrow$   
All classes treated as same  $\Sigma$ .

QDA Same as LDA but allow classes to have different  $\Sigma_k$

Fig. 4.9 Interpret

Role of prior:  ~~$\frac{1}{N}$  Score~~  $\frac{1}{2} (\text{Mahalanobis})^2 + \log \pi_k$

# President Garfield and ROC



Loss function: Make an explicit evaluation of the cost of FP and FN. Set parameters to minimize this.

Importance of prevalence. It's not just the classifier  
A test of ~~little use~~ <sup>little</sup> use in one group can be very useful in another.

HIV testing  
0.1% in general pop.  
5% in special sub-pops.

Mad Cow Disease.

Prevalence 0.000001%  
vs. 0.0% in downers.

Actual

	+ Test -	
+	.08	.02
-	1	1,000,000