IAML PP 2012-2013

1.a. Principal component Analysis

$x_1 = \begin{bmatrix}1\\1\end{bmatrix}$ $x_2 = \begin{bmatrix}2\\1\end{bmatrix}$ $x_3 = \begin{bmatrix}3\\1\end{bmatrix}$ $x_4 = \begin{bmatrix}1\\3\end{bmatrix}$ $x_5 = \begin{bmatrix}2\\3\end{bmatrix}$ $x_6 = \begin{bmatrix}3\\3\end{bmatrix}$

i. covariance matrix

$$\mu = \begin{bmatrix}\frac{1+2+3+1+2+3}{6}\\[4pt] \frac{1+1+1+3+3+3}{6}\end{bmatrix} = \begin{bmatrix}2\\2\end{bmatrix}$$

$x_1 = \begin{bmatrix}-1\\-1\end{bmatrix}$ $x_2 = \begin{bmatrix}0\\-1\end{bmatrix}$ $x_3 = \begin{bmatrix}1\\-1\end{bmatrix}$ $x_4 = \begin{bmatrix}-1\\1\end{bmatrix}$ $x_5 = \begin{bmatrix}0\\1\end{bmatrix}$ $x_6 = \begin{bmatrix}1\\1\end{bmatrix}$

...

$$\sigma_A = \frac{1+1+1+1}{6} = 2/3 \qquad \sigma = \begin{bmatrix}2/3\\1\end{bmatrix} \qquad \frac{\cancel{(-1)(-1)+0(-1)+(-1)}+(-1)+0}{\cancel{(4)}} = 0/6 = 0$$

$$\sigma_B = \frac{1+1+1+1+1+1}{6} = 1$$

$$\text{cov matrix} = \begin{bmatrix}2/3 & 0\\ 0 & 1\end{bmatrix}$$

$\sigma_{A,B}$ $\sigma_{A,B}$

ii. all eigenvectors

$$\det\begin{bmatrix}2/3-\lambda & 0\\ 0 & 1-\lambda\end{bmatrix} = (2/3-\lambda)(1-\lambda)-(0)(0) = 2/3 - 2/3\lambda - \lambda + \lambda^2$$
$$= \lambda^2 - 5/3\lambda + 2/3$$

$\lambda_1 = 1$ $\lambda_2 = 2/3$

$$\begin{bmatrix}2/3 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}e_1\\e_2\end{bmatrix} = 1\begin{bmatrix}e_1\\e_2\end{bmatrix} \Rightarrow \begin{bmatrix}2/3\,e_1\\e_2\end{bmatrix} = \begin{matrix}e_1\\e_2\end{matrix} \quad \begin{bmatrix}0\\1\end{bmatrix}$$

$$\begin{bmatrix}2/3 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}e_1\\e_2\end{bmatrix} = 2/3\begin{bmatrix}e_1\\e_2\end{bmatrix} \rightarrow \begin{bmatrix}2/3\,e_1\\e_2\end{bmatrix} = \begin{matrix}2/3\,e_1\\2/3\,e_2\end{matrix} = \begin{bmatrix}1\\0\end{bmatrix}$$

1 2/3

iii. % of eigenvalues

$$\lambda_1 = \frac{1}{1+2/3} \times 100 = 60\%$$

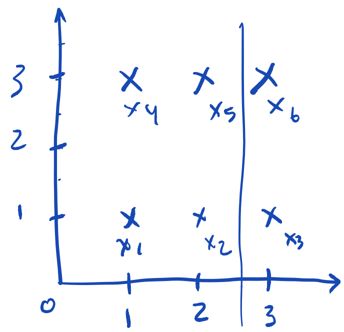
$$\lambda_2 = \frac{2/3+1}{1+2/3} \times 100 = 100\%$$

iv. PCA $\quad 2 \times 1 \quad 1 \times 2$ $\quad e^T x$

$$[0\ \ 1]\begin{bmatrix}1000\\1\end{bmatrix} = [1]$$

$$[0\ \ 1]\begin{bmatrix}1000\\3\end{bmatrix} = [3] \quad \rightarrow [1\ \ 3]$$

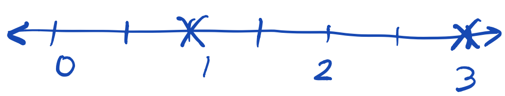
v.



$$[0\ 1]\begin{bmatrix}1\\1\end{bmatrix} = 1 \qquad [0\ 1]\begin{bmatrix}1\\3\end{bmatrix} = 3$$

$$[0\ 1]\begin{bmatrix}2\\1\end{bmatrix} = 1 \qquad [0\ 1] = \begin{bmatrix}2\\3\end{bmatrix} = 3$$

$$[0\ 1]\begin{bmatrix}3\\1\end{bmatrix} \ 1 \qquad [0\ 1] = \begin{bmatrix}3\\3\end{bmatrix} = 3$$



non-linearly separable, highest acc: $2/3$

vi. —

vii.

$$\begin{bmatrix}0\\0\end{bmatrix} \begin{bmatrix}0\\0\end{bmatrix} \begin{bmatrix}1\\0\end{bmatrix} \quad \begin{bmatrix}0\\1\end{bmatrix} \begin{bmatrix}0\\1\end{bmatrix} \begin{bmatrix}1\\1\end{bmatrix} = \begin{matrix}2/6 = 1/3\\ 3/6 = 1/2\end{matrix}$$

$$\begin{bmatrix}-1/3\\-1/2\end{bmatrix} \begin{bmatrix}-1/3\\-1/2\end{bmatrix} \begin{bmatrix}2/3\\-1/2\end{bmatrix} \begin{bmatrix}-1/3\\1/2\end{bmatrix} \begin{bmatrix}-1/3\\1/2\end{bmatrix} \begin{bmatrix}2/3\\1/2\end{bmatrix}$$

$$\Sigma = \begin{bmatrix}1/3 & 0\\ 0 & 1/2\end{bmatrix}$$

$$\det\begin{bmatrix}1/3-\lambda & 0\\ 0 & 1/2-\lambda\end{bmatrix} = (1/3-\lambda)(1/2-\lambda)$$

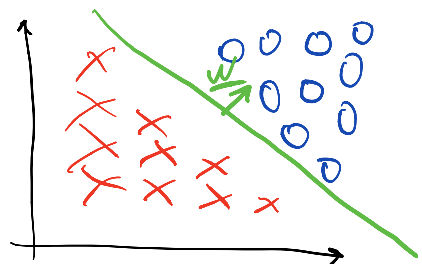
$$= 1/6 - 1/3\lambda - 1/2\lambda + \lambda^2$$

$$= \lambda^2 - 5/6\lambda + 1/6 \qquad \begin{matrix}\lambda_1 = 1/2\\ \lambda_2 = 1/3\end{matrix}$$

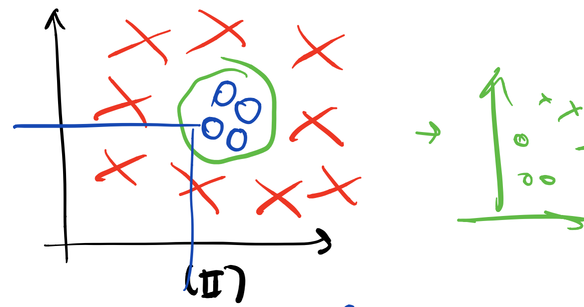
Percentages = $3/5$, $2/5$

60% $\quad$ 40%

2. a. Linear classification



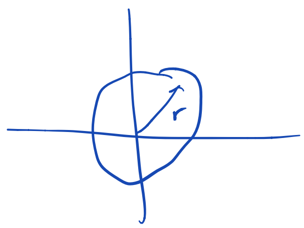
(I) $W = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$



(II)

$\begin{bmatrix} 2 \\ 2 \end{bmatrix}$

I. Non-linearly separable



$x_1^2 + x_2^2 = r^2 \rightarrow x^2 + y^2 = 1$

$x \rightarrow \begin{bmatrix} x_1^2 \\ x_2^2 \end{bmatrix}$ 4 4

$[4 \;\; 4]\begin{bmatrix} 2 \\ 2 \end{bmatrix} =$ 16 16

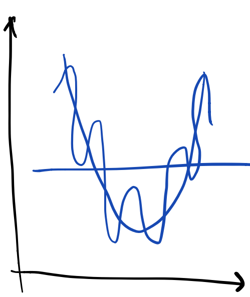
new vector $\rightarrow \begin{bmatrix} x_1^2 + x_2^2 \\ 0 \end{bmatrix}$

For all x_1, x_2, if $\underline{x_1^2 + x_2^2 \geq 1}$ ... else

map $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \rightarrow [x_1^2 + x_2^2]$ ↑

b. Overfitting

$y = w_0 x^2 + w_1 x + \varepsilon$



i. model that will
underfit: x
overfit: x^3

ii. Objective function

$$\frac{1}{n}\sum_{i=1}^{n}\Big(\underset{\text{true}}{y} - \underbrace{[w_0 x_i^2 + w_1 x_i]}_{\text{value we predicted}}\Big)^2 + \lambda\|w\|$$

c. Evaluation

i. Classification accuracy

if a lot of No ≈ 1

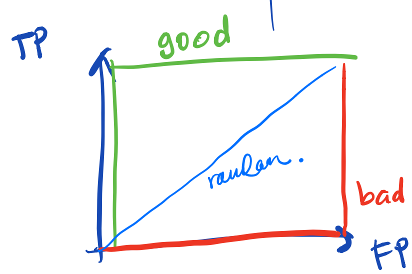
a lot of yes

$\frac{TN}{TN + FN} \rightarrow$ a lot ≈ 0

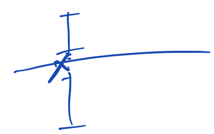
$$ACC = \frac{TP + TN}{TP + TN + FP + FN} = \frac{\text{correct}}{\text{total}} \qquad \frac{99}{100} \qquad FP + FN = \frac{1}{100} \times \text{no.}$$

not enough → class might be imbalanced

know TP and FP



* shows performance of system across all possible thresholds



ii. C 70% in 10, A 70% in 100, B 90% in 100

210 21 9

$$0.3 \pm \frac{0.3(0.7)}{10} \qquad 0.3 \pm \frac{0.3(0.7)}{100} \qquad 0.1 \pm \frac{0.1(0.9)}{100}$$

B < A < C

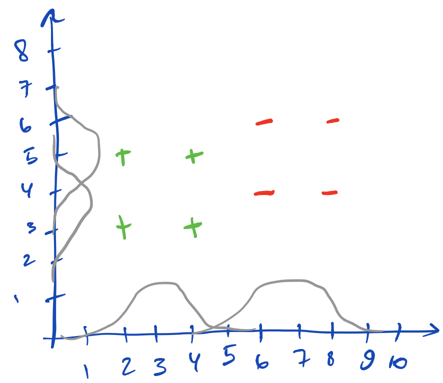
3. Naive Bayes

Gaussian distribution: $P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$

Diagnose cancer, tests $\rightarrow$ real number.

Training examples

Class	Test 1 (t_1)	Test 2 (t_2)
+	4.0	5.0
-	6.0	6.0
+	2.0	5.0
-	8.0	4.0
+	4.0	3.0
-	8.0	6.0
+	2.0	3.0
-	6.0	4.0



a. Estimate NB classifier

$\mu_{+_1}, \mu_{+_2}, \sigma_{+_1}, \sigma_{+_2}$

$\mu_{-_1}, \mu_{-_2}, \sigma_{-_1}, \sigma_{-_2}$

$$\mu_+ = \frac{4+5+2+5+4+3+2+3}{8} = 3.5$$

$$\sigma_+ = \frac{(0.5)^2 + (1.5)^2 + (1.5)^2 + (1.5)^2 + (0.5)^2 + (0.5)^2 + (1.5)^2 + (0.5)^2}{8}$$

$= 1.25$