22 March 2021

1

Bias

a) $$\left\{ x^1 = \overset{1}{\begin{pmatrix} 1 \\ 1 \end{pmatrix}},\ x^2 = \overset{1}{\begin{pmatrix} 2 \\ 2 \end{pmatrix}},\ x^3 = \overset{1}{\begin{pmatrix} 0 \\ -1 \end{pmatrix}},\ x^4 = \overset{1}{\begin{pmatrix} -1 \\ 0 \end{pmatrix}} \right\}$$

$$\{ t^1 = 1,\ t^2 = 1,\ t^3 = -1,\ t^4 = -1 \}$$

$$\eta = 1 \ ,\ w = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \ ;\ w = w + \eta\, \underbrace{(t + \sigma)}_{\delta} x$$

1st Epoch

- x^1: $\sigma^1 = \text{sign}(w\, x^1) = \text{sign}(1 \cdot 1 + 1 \cdot 1 + 1 \cdot 1) = +1$

  As $t^1 = \sigma^1$, no mistake was made
- x^2: $\sigma^2 = \text{sign}(w\, x^2) = \text{sign}(1 \cdot 1 + 2 \cdot 1 + 2 \cdot 1) = +1$

  As $t^2 = \sigma^2$, no mistake was made
- x^3: $\sigma^3 = \text{sign}(w \cdot x^3) = \text{sign}(1 \cdot 1 + 0 \cdot 1 + (-1) \cdot 1) = \text{sign}(0) = +1$

  $w^{new} = w + \eta \cdot (t - \sigma) x \Rightarrow w = (1\ 1\ 1)^T + 1 \cdot (1\ 0\ -1)^T\ (-1 - 1)$

  $\Leftrightarrow w = (1\ 1\ 1)^T + (-2\ 0\ 2)^T = (-1\ 1\ 3)$
- x^4: $\sigma^4 = \text{sign}(w \cdot x^4) = \text{sign}(1 \cdot -1 + (-1) \cdot 1 + 0 \cdot 3) = \text{sign}(-2) = -1$

  As $t^4 = \sigma^4$, no mistake was made

2nd Epoch

- x^1: $\sigma^1 = \text{sign}(w\, x^1) = \text{sign}(1 \cdot (-1) + 1 \cdot 1 + 1 \cdot 3) = +1$

  As $t^1 = \sigma^1$, no mistake was made
- x^2: $\sigma^2 = \text{sign}(w\, x^2) = \text{sign}(1 \cdot (-1) + 2 \cdot 1 + 2 \cdot 3) = +1$

  As $t^2 = \sigma^2$, no mistake was made
- x^3: $\sigma = \text{sign}(w\, x^3) = \text{sign}(1 \cdot (-1) + 0 \cdot 1 + (-1) \cdot 3) = -1$

  As $t^3 = \sigma^3$, no mistake was made
- x^4: $\sigma = \text{sign}(w\, x^4) = \text{sign}(1\ (-1) + (-1) \cdot 1 + 0 \cdot 3) = -1$

  As $t^4 = \sigma^4$, no mistake was made

We have converged. $w = \begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix}$

HW 1 - Page 2

22 March 2021

1

b) $\left\{ x^1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, x^2 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, x^3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, x^4 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\}$ → Bias

$\{ t^1 = 1, t^2 = 1, t^3 = -1, t^4 = -1 \}$

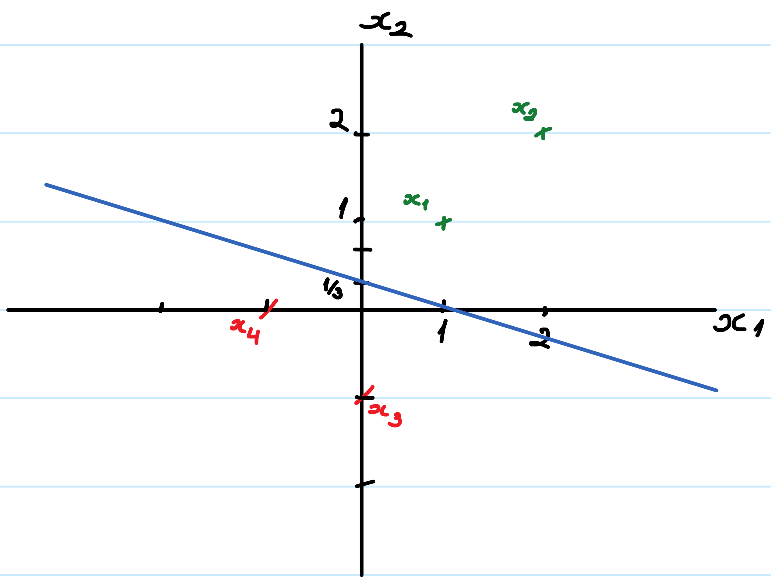
$\eta = 1 \quad , \quad w = \begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix}$

Seperation Hyperplane

$$w \cdot x = 0 \Leftrightarrow w_0 \cdot x_0 + w_1 \cdot x_1 + w_2 x_2 = 0$$

$$\Leftrightarrow (-1)\,1 + 1 \cdot x_1 + 3\, x_2 = 0 \Leftrightarrow x_1 + 3x_2 = 1$$

$$\Leftrightarrow x_2 = \frac{1}{3} - \frac{1}{3} x_1$$



22 March 2021

2

a)

F_1	F_2	F_3	F_4	Output
c	a	b	x	m
a	a	c	a	t
a	b	b	a	t
c	b	c	x	m
a	b	b	c	f

1^{st} - Compute initial entropy

$$E_{start} = E\left(\frac{\#\{O=f\}}{\sum_o \#\{O=o\}}, \frac{\#\{O=m\}}{\sum_o \#\{O=o\}}, \frac{\#\{O=m\}}{\sum_o \#\{O=o\}}, \frac{\#\{O=t\}}{\sum_o \#\{O=o\}}\right)$$

$$= E\left(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{2}{5}\right) = -\left(\frac{1}{5}\log_2\frac{1}{5} + \frac{1}{5}\log_2\frac{1}{5} + \frac{1}{5}\log_2\frac{1}{5} + \frac{2}{5}\cdot\log_2\frac{2}{5}\right)$$

$$= -\left(\frac{3}{5}\log_2\frac{1}{5} + 2/5\log_2\frac{2}{5}\right) = -\left(-1.39 + (-0.53)\right) = 1.92 \text{ Bits}$$

2^{nd} - Test each attribute

- F_1:

$F_1 = a$
$\#\{O=f\} = 1$
$\#\{O=m\} = 0$
$\#\{O=m\} = 0$
$\#\{O=t\} = 2$

$F_1 = c$
$\#\{O=f\} = 0$
$\#\{O=m\} = 1$
$\#\{O=m\} = 1$
$\#\{O=t\} = 0$

$$E\left(\frac{1}{3}, \frac{2}{3}\right) = -\left(1/3\log_2 1/3 + 2/3\log_2 2/3\right)$$
$$= 0.92 \text{ Bits}$$

$$E(1/2, 1/2) = -\left(1/2\log_2 1/2 + 1/2\log_2 1/2\right)$$
$$= 1 \text{ Bit}$$

$$E_{F_1} = 3/5 \cdot E(1/3, 2/3) + 2/5\ E(1/2, 1/2) = 0.952 \text{ Bits}$$

$$G(F_1) = E_{start} - E_{F_1} = 1.92 - 0.95 \approx 0.97 \text{ Bits}$$

2

a cont.) 2^{nd} - Test each attribute (cont)

- F_2:

$F_2 = a$	$F_2 = b$
$\#\{O=j\} = 0$	$\#\{O=j\} = 0$
$\#\{O=m\} = 0$	$\#\{O=m\} = 1$
$\#\{O=m\} = 1$	$\#\{O=m\} = 1$
$\#\{O=t\} = 1$	$\#\{O=t\} = 1$

$E(1/2, 1/2) = -(1/2 \log_2 1/2 + 1/2 \log_2 1/2)$
$= 1$ Bit

$E(1/3, 1/3, 1/3) = -(1/3 \log_2 1/3 + 1/3 \log_2 1/3 + 1/3 \log_2 1/3)$
$= 1.58$ Bits

$$E_{F_2} = 3/5 \cdot E(1/3, 1/3, 1/3) + 2/5 \; E(1/2, 1/2) = 1.348 \text{ Bits}$$
$$\underline{\underline{G}}(F_2) = E_{Start} - E_{F_2} = 1.92 - 1.35 \approx 0.57 \text{ Bits}$$

- F_3:

$F_3 = b$	$F_3 = c$
$\#\{O=j\} = 1$	$\#\{O=j\} = 0$
$\#\{O=m\} = 0$	$\#\{O=m\} = 1$
$\#\{O=m\} = 1$	$\#\{O=m\} = 0$
$\#\{O=t\} = 1$	$\#\{O=t\} = 1$

$E(1/3, 1/3, 1/3) = -(1/3 \log_2 1/3 + 1/3 \log_2 1/3 + 1/3 \log_2 1/3)$
$= 1.58$ Bits

$E(1/2, 1/2) = -(1/2 \log_2 1/2 + 1/2 \log_2 1/2)$
$= 1$ Bit

$$E_{F_3} = 3/5 \cdot E(1/3, 1/3, 1/3) + 2/5 \; E(1/2, 1/2) = 1.348 \text{ Bits}$$
$$\underline{\underline{G}}(F_3) = E_{Start} - E_{F_3} = 1.92 - 1.35 \approx 0.57 \text{ Bits}$$

- F_4:

$F_4 = x$	$F_4 = d$	$F_4 = c$
$\#\{O=j\} = 0$	$\#\{O=j\} = 0$	$\#\{O=j\} = 1$
$\#\{O=m\} = 1$	$\#\{O=m\} = 0$	$\#\{O=m\} = 0$
$\#\{O=m\} = 1$	$\#\{O=m\} = 0$	$\#\{O=m\} = 0$
$\#\{O=t\} = 0$	$\#\{O=t\} = 2$	$\#\{O=t\} = 0$

$E(1/2, 1/2) = -(1/2 \log_2 1/2 + 1/2 \log_2 1/2)$
$= 1$ Bit

$E(2/2) = -(1 \log_2 1)$
$= 0$ Bits

$E(1/1) = -(1 \log_2 1)$
$= 0$ Bits

$$E_{F_4} = 2/5 \cdot E(1/2, 1/2) + \cancelto{0}{2/5 \; E(2/2)} + \cancelto{0}{1/5 \; E(1/1)} = 2/5 \text{ Bits}$$
$$\underline{\underline{G}}(F_4) = E_{Start} - E_{F_4} = 1.92 - 2/5 \approx 1.52 \text{ Bits}$$

2

a cont.) 3^{rd} - Pick root

$G(F_1) = 0{,}97$

$G(F_2) = G(F_3) = 0{,}57$

$G(F_4) = 1{,}52$

Since F_4 has the highest info gain, F_4 is picked as root ✓

b) Picking F_4 as our root, we have the following partitions

$F_4 = x$

F_1	F_2	F_3	O
c	a	b	m
c	b	c	m

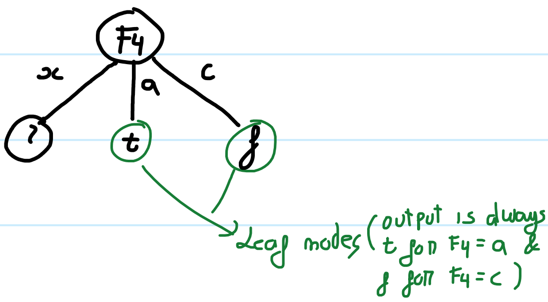
$F_4 = a$

F_1	F_2	F_3	O
a	a	c	t
a	b	b	t

$F_4 = c$

F_1	F_2	F_3	O
a	b	b	f

- Right now we have the tree.
- So we need to pick which feature to test on node ?



- Following the same logic as in exercise 2 a), we have

1^{st} - Initial Entropy

- $E_{start} = E(1/2, 1/2) = 1$ Bit

2^{nd} - Test each feature

- F_1

$F_1 = c$
$\#\{O = m\} = 1$
$\#\{O = m\} = 1$

$E_{F_1} = 1 \cdot E(1/2, 1/2) = 1$ Bit

$G_{F_1} = E_{start} - E_{F_1} = 0$

HW 1 - Page 6

22 March 2021

2

b cont.) 2^{nd} - Test each feature (cont)

- F_2

$F_2 = a$	$F_2 = b$
$\#\{0=m\} = 1$	$\#\{0=m\} = 0$
$\#\{0=m\} = 0$	$\#\{0=m\} = 1$
$E(1/1) = 0$ Bit	$E(1/1) = 0$ Bit

$$E_{F_2} = 1/2 \; 0 + 1/2 \; 0 = 0$$

$$G_{F_2} = E_{start} - E_{F_2} = 1 \text{ Bit} \quad (\text{MAX!})$$

- F_3

$F_3 = a$	$F_3 = b$
$\#\{0=m\} = 0$	$\#\{0=m\} = 1$
$\#\{0=m\} = 1$	$\#\{0=m\} = 0$
$E(1/1) = 0$ Bit	$E(1/1) = 0$ Bit

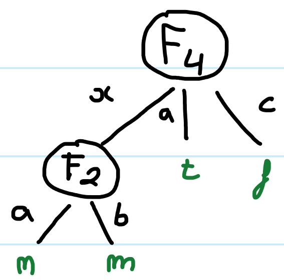
$$E_{F_3} = 1/2 \; 0 + 1/2 \; 0 = 0$$

$$G_{F_3} = E_{start} - E_{F_3} = 1 \text{ Bit} \quad (\text{MAX!})$$

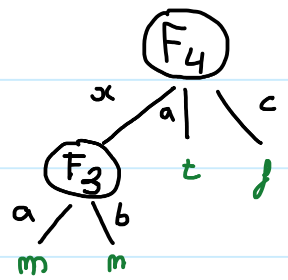
3^{rd} - Pick Feature to test

Both F_2 & F_3 have maximum info. gain (whilst F_1 has minimum)

So, we can pick between either of these 2 features.



or

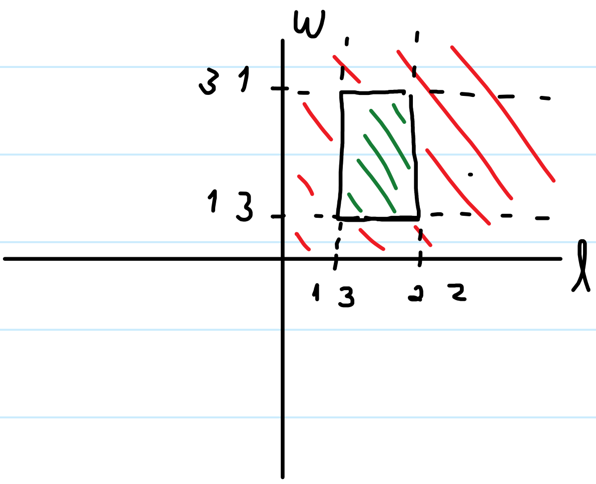


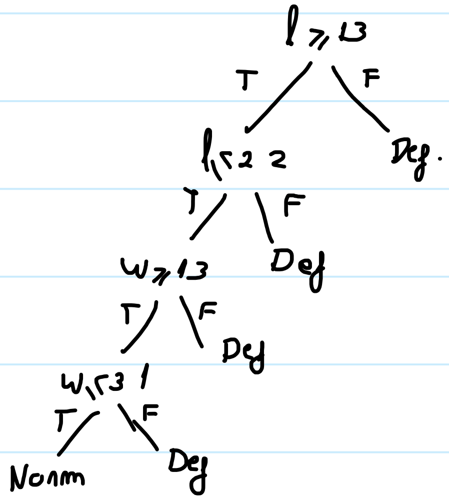
HW 1 - Page 7

26 March 2021

3

Only the Decision Tree. This problem demands a classifier able to represent joint conditions across multiple dimensions. By this we mean that the condition to classify a piece as defective or not can be written as a conjunction of AND statements (l < 2.2 AND l > 1.3 AND w < 3.1 AND w > 1.3), which would be too complex for a single perceptron (which can only classify linearly separable points). Decision trees, however, have no problem dealing with conjunctions such as this one.





4

Teacher Note: Assume for standard deviation and covariance that we have population, use N instead of N-1 (divide by N and not N-1)

a)

x_1	x_2	Class
0	10	A
0	20	A
10	10	A
5	20	A
30	30	B
40	40	B
50	30	B
50	50	B

$$x = \begin{pmatrix} 5 \\ 10 \end{pmatrix}$$

$$\mu = \frac{\sum_{x \in X} x}{N} \qquad \sigma = \sqrt{\frac{\sum_{x \in X} (x-\mu)^2}{N}}$$

- ① Prior

$$P(A) = \frac{\#A}{\#Examples} = \frac{4}{8} = 1/2 \qquad P(B) = \frac{\#B}{\#Examples} = \frac{4}{8} = 1/2$$

- ② Estimate likelihoods

• $p(X_1 | Class = A)$

$\mu \quad \frac{0+0+10+5}{4} = \underline{3.75}$

$\sigma \quad \sqrt{\left(\frac{(0-3.75)^2+(0-3.75)^2+(10-3.75)^2+(5-3.75)^2}{4}\right)} = \underline{4.15}$

• $P(X_1 | Class = B)$

$\mu \quad \frac{30+40+50+50}{4} = \underline{42.5}$

$\sigma \quad \sqrt{\left(\frac{(30-42.5)^2+(40-42.5)^2+(50-42.5)^2+(50-42.5)^2}{4}\right)} = \underline{8.29}$

26 March 2021

4

Teacher Note: Assume for standard deviation and covariance that we have population, use N instead of N-1 (divide by N and not N-1)

a cont.)

x_1	x_2	Class
0	10	A
0	20	A
10	10	A
5	20	A
30	30	B
40	40	B
50	30	B
50	50	B

$$x = \begin{pmatrix} 5 \\ 10 \end{pmatrix}$$

$$\mu = \frac{\sum x}{N} \qquad \sigma = \sqrt{\frac{\sum (x-\mu)^2}{N}}$$

$$p(x \mid \mu, \sigma^2) = \mathcal{N}(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2 \cdot \pi}\, \sigma} \exp\left(-\frac{1}{2 \cdot \sigma^2} (x-\mu)^2\right)$$

- ② Estimate likelihoods

- $p(X_2 \mid Class = A)$

μ $\frac{10+20+10+20}{4} = \underline{15.0}$

σ $\sqrt{\left(\frac{(10-15)^2 + (20-15)^2 + (10-15)^2 + (20-15)^2}{4}\right)} = \underline{5}$

- $p(X_2 \mid Class = B)$

μ $\frac{30+40+30+50}{4} = \underline{37.5}$

σ $\sqrt{\left(\frac{(30-37.5)^2 + (40-37.5)^2 + (30-37.5)^2 + (50-37.5)^2}{4}\right)} = \underline{8.29}$

- $$P(Class = A \mid X_1 = 5, X_2 = 10) = \frac{P(Class = A)\, P(X_1 = 5, X_2 = 10 \mid Class = A)}{P(X_1 = 5, X_2 = 10)}$$

$$= \frac{1/2 \cdot p(X_1 = 5 \mid Class = A)\ p(X_2 = 10 \mid Class = A)}{P(X_1 = 5, X_2 = 10)} = \frac{1/2\ \mathcal{N}(5 \mid \mu = 3.75, \sigma = 4.15) \cdot \mathcal{N}(10 \mid \mu = 15, \sigma = 5)}{P(X_1 = 5, X_2 = 10)}$$

$$= \frac{1/2 \frac{1}{\sqrt{2\pi} \cdot 4.15} \cdot \exp\left(-\frac{1}{2 \cdot 4.15^2} \cdot (5-3.75)^2\right) \cdot \frac{1}{\sqrt{2\pi} \cdot 5} \cdot \exp\left(-\frac{1}{2 \cdot 5^2} \cdot (10-15)^2\right)}{P(x_1 = 5, x_2 = 10)} \Rightarrow \frac{1/2\ 0.092 \cdot 0.048}{P(x_1 = 5, x_2 = 10)}$$

$$= \frac{0.002}{P(X_1 = 5, X_2 = 10)}$$

- $$P(Class = B \mid X_1 = 5, X_2 = 10) = \frac{P(Class = B)\, P(X_1 = 5, X_2 = 10 \mid Class = B)}{P(X_1 = 5, X_2 = 10)}$$

$$= \frac{1/2 \cdot p(X_1 = 5 \mid Class = B)\ p(X_2 = 10 \mid Class = B)}{P(X_1 = 5, X_2 = 10)} = \frac{1/2\ \mathcal{N}(5 \mid \mu = 42.5, \sigma = 8.29) \cdot \mathcal{N}(10 \mid \mu = 37.5, \sigma = 8.29)}{P(X_1 = 5, X_2 = 10)}$$

$$= \frac{1/2 \frac{1}{\sqrt{2\pi} \cdot 8.29} \cdot \exp\left(-\frac{1}{2 \cdot 8.29^2} \cdot (5-42.5)^2\right) \cdot \frac{1}{\sqrt{2\pi} \cdot 8.29} \cdot \exp\left(-\frac{1}{2 \cdot 8.29^2} \cdot (10-37.5)^2\right)}{P(x_1 = 5, x_2 = 10)} \Rightarrow \frac{1/2\ 1.73 \cdot 10^{-6}\ 0.00019}{P(x_1 = 5, x_2 = 10)}$$

$$= \frac{1.64 \times 10^{-10}}{P(X_1 = 5, X_2 = 10)}$$

As $0.002 > 1.64 \times 10^{-6}$, this query vector will be classified as A

$$= \frac{1.64 \times 10^{-10}}{P(x_1=5, x_2=10)}$$

As $0.0027 > 1.64 \times 10^{-6}$, this query vector will be Classified as A

HW 1 - Page 9

26 March 2021

4

Teacher Note: Assume for standard deviation and covariance that we have population, use N instead of N-1 (divide by N and not N-1)

b)

x_1	x_2	Class
0	10	A
0	20	A
10	10	A
5	20	A
30	30	B
40	40	B
50	30	B
50	50	B

$x = \begin{pmatrix} 5 \\ 10 \end{pmatrix}$

$\mu = \frac{\sum x}{N}$ $\quad \Sigma_{i,j} = \frac{1}{N} \sum_{l=1}^{N} (x_i - \mu_i)(x_j - \mu_j)$

$p(x \mid \mu, \sigma^2) = \mathcal{N}(x \mid \mu, \Sigma) = \frac{1}{(2\pi)^{D/2}} \cdot \frac{1}{|\Sigma|^{1/2}} \exp\left(-\frac{1}{2} \cdot (x-\mu)^T \Sigma^{-1} \cdot (x-\mu)\right)$

$D = 2$

① Prior

$P(A) = \frac{\# A}{\# \text{Examples}} = \frac{4}{8} = \frac{1}{2}$ $\qquad P(B) = \frac{\# B}{\# \text{Examples}} = \frac{4}{8} = \frac{1}{2}$

② Estimate likelihoods

- $p(X_1, X_2 \mid Class = A)$

$$\mu \cdot \begin{bmatrix} \frac{0+0+10+5}{4} \\ \frac{10+20+10+20}{4} \end{bmatrix} = \begin{bmatrix} 3{,}75 \\ 15 \end{bmatrix}$$

$$\Sigma \cdot \begin{bmatrix} \frac{(0-3.75)^2 + (0-3.75)^2 + (10-3.75)^2 + (5-3.75)^2}{4} & \frac{(0-3.75)\cdot(10-15) + (0-3.75)\cdot(20-15) + (10-3.75)\cdot(10-15) + (5-3.75)(20-15)}{4} \\ \frac{(0-3.75)\cdot(10-15) + (0-3.75)\cdot(20-15) + (10-3.75)\cdot(10-15) + (5-3.75)(20-15)}{4} & \frac{(10-15)^2 + (20-15)^2 + (10-15)^2 + (20-15)^2}{4} \end{bmatrix} = \begin{bmatrix} 17.19 & -6.25 \\ -6.25 & 25 \end{bmatrix}$$

- $p(X_1, X_2 \mid Class = B)$

$$\mu \cdot \begin{bmatrix} \frac{30+40+50+50}{4} \\ \frac{30+40+30+50}{4} \end{bmatrix} = \begin{bmatrix} 42.5 \\ 37{,}5 \end{bmatrix}$$

$$\Sigma \cdot \begin{bmatrix} \frac{(30-42.5)^2 + (40-42.5)^2 + (50-42.5)^2 + (50-42.5)^2}{4} & \frac{(30-42.5)(30-37.5) + (40-42.5)\cdot(40-37.5) + (50-42.5)\cdot(30-37.5) + (50-42.5)(50-37.5)}{4} \\ \frac{(30-42.5)(30-37.5) + (40-42.5)\cdot(40-37.5) + (50-42.5)\cdot(30-37.5) + (50-42.5)(50-37.5)}{4} & \frac{(30-37{,}5)^2 + (40-37{,}5)^2 + (30-37{,}5)^2 + (50-37{,}5)^2}{4} \end{bmatrix} = \begin{bmatrix} 68.75 & 31.25 \\ 31.25 & 68.75 \end{bmatrix}$$

$$\left[\frac{(30-42.5)(30-37.5)+(40-42.5)\cdot(40-37.5)+(50-42.5)\cdot(30-37.5)+(50-42.5)(50-37.5)}{4} \quad \frac{(30-37{,}5)^2+(40-37{,}5)^2+(30-37{,}5)^2+(50-37{,}5)^2}{4} \right]$$

HW 1 - Page 10

26 March 2021

4

b cont.)

- ② Estimate likelihoods

• $p(X_2 | Class = A)$

μ $\frac{10+20+10+20}{4} = \underline{15.0}$

σ $\sqrt{\left(\frac{(10-15)^2 + (20-15)^2 + (10-15)^2 + (20-15)^2}{4}\right)} = \underline{5}$

• $P(X_2 | Class = B)$

μ $\frac{30+40+30+50}{4} = \underline{37.5}$

σ $\sqrt{\left(\frac{(30-37.5)^2 + (40-37.5)^2 + (30-37.5)^2 + (50-37.5)^2}{4}\right)} = \underline{8.29}$

• $P(Class = A \mid X_1 = 5, X_2 = 10) = \frac{P(Class = A)\, P(X_1 = 5, X_2 = 10 \mid Class = A)}{P(X_1 = 5, X_2 = 10)}$

$$= \frac{\frac{1}{2} \mathcal{N}\left(\begin{bmatrix} 5 \\ 10 \end{bmatrix} \middle| \mu = \begin{bmatrix} 3.75 \\ 15 \end{bmatrix}, \Sigma = \begin{bmatrix} 17.19 & -6.25 \\ -6.25 & 25 \end{bmatrix}\right)}{P(x_1 = 5, x_2 = 10)} =$$

$$= \frac{\frac{1}{2}\left(\frac{1}{2\pi^{D/2}} \cdot \frac{1}{|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}\left(\begin{bmatrix} 5 \\ 10 \end{bmatrix} - \begin{bmatrix} 3.75 \\ 15 \end{bmatrix}\right)^T \Sigma^{-1} \cdot \left(\begin{bmatrix} 5 \\ 10 \end{bmatrix} - \begin{bmatrix} 3.75 \\ 15 \end{bmatrix}\right)\right)\right)}{P(x_1 = 5, x_2 = 10)} = \frac{\frac{1}{2}\, 0.0049}{P(x_1 = 5, x_2 = 10)}$$

$= 2/2$

$17.19 \cdot 25 - (-6.25 \cdot -6.25)$

• $P(Class = B \mid X_1 = 5, X_2 = 10) = \frac{P(Class = B)\, P(X_1 = 5, X_2 = 10 \mid Class = B)}{P(X_1 = 5, X_2 = 10)}$

$$= \frac{\frac{1}{2} \mathcal{N}\left(\begin{bmatrix} 5 \\ 10 \end{bmatrix} \middle| \mu = \begin{bmatrix} 42.5 \\ 37.5 \end{bmatrix}, \Sigma = \begin{bmatrix} 68.75 & 31.25 \\ 31.25 & 68.75 \end{bmatrix}\right)}{P(x_1 = 5, x_2 = 10)} =$$

$$= \frac{\frac{1}{2}\left(\frac{1}{2\pi^{D/2}} \cdot \frac{1}{|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}\left(\begin{bmatrix} 5 \\ 10 \end{bmatrix} - \begin{bmatrix} 42.5 \\ 37.5 \end{bmatrix}\right)^T \Sigma^{-1} \cdot \left(\begin{bmatrix} 5 \\ 10 \end{bmatrix} - \begin{bmatrix} 42.5 \\ 37.5 \end{bmatrix}\right)\right)\right)}{P(x_1 = 5, x_2 = 10)} = \frac{\frac{1}{2}\, 3.45 \cdot 10^{-8}}{P(x_1 = 5, x_2 = 10)}$$

$2/2$

$68.75 \cdot 68.75 - (31.25 \cdot 31.25)$

$\boxed{\text{As } 0.0049 > 3.45 \times 10^{-8}, \text{ this query vector will be classified as } A}$