EE 503

HOMEWORK-3

1.

(a) Baye's optimal classifier's decision rule according to Bayesian decision theory, posterior probability can be calculated as

$$P(w_j|x) = \frac{P(x|w_j)P(w_j)}{P(x)} \rightarrow \text{Prior.}$$

let's use this for given problem $\rightarrow$ (or) $P(w_i|x) = \frac{P(x|w_i)P(w_i)}{P(x)}$

$$P(x) = \sum_{j=1}^{c} P(x|w_j)P(w_j)$$

$\therefore$ we need to chose the one which maximises the posterior probability

here, we can see,

~~$P(w_i|x) \propto P(x|w_i)P(w_i) =$~~

~~Given: $X_i|w_i \sim \text{Gamma}(p_i, \lambda_j)$~~

~~$\therefore P(x|w_j) = \prod_{i=1}^{k} \lambda_i$~~

$$P(w_i|x) \propto P(x|w_i)P(w_i)$$

Given: $x_j|w_i \sim \text{Gamma}(p_i, \lambda_j)$

$$\therefore P(x|w_i) = \prod_{j=1}^{k} \frac{\lambda_j^{p_i} x_j^{p_i-1} e^{-\lambda_j x_j}}{\Gamma(p_i)}.$$

$\therefore$ decision rule is

$$w = \arg\max_i \left[\log P(w_i) + \sum_{j=1}^{k}\left(p_i \log\lambda_j + (p_i-1)\log x_j - \lambda_j x_j - \log\left[\Gamma(p_i)\right]\right)\right]$$

(b) when are decision boundaries linear functions of $x_1, x_2, \ldots x_k$?

$$\log P(w_i|x) \propto \log P(w_i) + \sum_{j=1}^{k} (p_i \log \lambda_j + (p_i - 1)\log x_j - \lambda_j x_j - \log[\Gamma(p_i)])$$

here we see that when $(p_i - 1) = 0$ i.e $p_i = 1$, the equation becomes linear in x_j. & eq^n becomes $w^* = \arg\max\left[\log P(w_i) + \sum_{j=1}^{k}(-\lambda_j x_j)\right]$

Anyway, Suppose there are 2 classes, ~~there~~ w_i & w_k

$\therefore$ decision boundary becomes,

$$\underbrace{\log(P(w_i))}_{\text{const}} + \sum_{j=1}^{k}(p_i \log \lambda_j + (p_i - 1)\log x_j - \overbrace{\lambda_j x_j}^{\text{of class } w_i} - \log[\Gamma(p_i)]) = \overbrace{\log(P(w_j))}^{\text{const.}} - \log[\Gamma(p_k)] + \sum_{j=1}^{k}(p_k \log \lambda_j + (p_k - 1)\log x_j - \underbrace{\lambda_j x_j}_{\text{class } w_k}$$

$$\Rightarrow \sum_{j=1}^{k}(p_i - p_k)\log \lambda_j + (p_i - p_k)\log x_j - (\underbrace{\lambda_j p_i}_{\lambda_j x_j \text{ of } w_i(p_i) \text{ class}} - \underbrace{\lambda_j p_k}_{\lambda_j p_k \text{ of } w_i(p_k) \text{ class}})x_j) + \sum_{j=1}^{k}[-\log(\Gamma(p_i))$$

$$- \log(\Gamma(p_k)))] + \log(P(w_i)) - \log(P(w_k))) = 0$$

for the decision boundary to be linear function of $x_1, x_2 \ldots x_k$, the terms $\log(x_j) = 0$, this can happen, when $p_i = p_k$. leaving out $\boxed{\lambda_j x_j}$ terms which is linear in $x_1\ x_2 \ldots x_k$.

(c) classify:

$x = (0.1, 0.2, 0.3, 4)$

Given:

$P_1 = 4, \quad p_2 = 2$

$c = 2$

$k = 4$

$\lambda_1 = \lambda_3 = 1$

$\lambda_2 = \lambda_4 = 2$

equal prob for each class.

from part (i) we have,

$$\log P(w_1|x) \propto \log P(w_1) + \sum_{j=1}^{4} (P_1 \log \lambda_j + (P_1 - 1)\log x_j - \lambda_j x_j - \log\Gamma(P_1))$$

$$\because \log P(w_1) = \log(1/2) = -\log(2)$$

$$\therefore \log P(w_1|x) \propto -\log 2 + \sum_{j=1}^{4} (4\log\lambda_j + 3\log x_j - \lambda_j x_j - \log\Gamma(4))$$

$$\propto \underbrace{-\log 2}_{-0.3} + \underbrace{(4\log 1 + 3\log 0.1 - (1)0.1)}_{-3.1} + \underbrace{(4\log 2 + 3\log(0.2) - 0.4}_{-2.455}$$

$$+ \underbrace{(4\log 1 + 3\log(0.3) - 0.3(1))}_{-1.86} + \underbrace{(4\log 2 + 3\log(4) - 8)}_{-1.0685}$$

$$+ \underbrace{\log\Gamma(4))}_{-1.7918}$$

$$= \log P(w_1|x) \overset{\propto}{=} \boxed{-10.9777} \quad —①$$

for w_2,

$$\log(P(w_2|x)) \overset{\propto}{=} \log P(w_2) + \sum_{j=1}^{4} (P_2 \log\lambda_j + (P_2 - 1)\log x_j - \lambda_j x_j - \log\Gamma(P_2))$$

$$= \underbrace{\log(1/2)}_{-0.3} + \underbrace{(2\log 1 + 1\log(0.1) - (1)(0.1))}_{-1.1} + \underbrace{(2\log 2 + \log(0.2) - 0.4}_{-0.623}$$

$$+ \underbrace{2\log 1 + \log(0.3) - 0.3(1)}_{-0.822} + \underbrace{(2\log 2 + \log(4) - 8)}_{-5.222} + \underbrace{\log\Gamma(2)}_{0}$$

$$\Rightarrow \log P(w_2|x) \overset{\propto}{=} \boxed{-8.4665} \quad —②$$

$\because \log P(w_2|x) > \log P(w_1|x)$, we classify X as belonging to class w_2.

(d) $P_1 = 3.2$

$P_2 = 8$

$c = 2$

$k = 1$

$\lambda_1 = 1$

equal prior prob

To find: decision boundary $x = x^*$

for class w_1:

$$\log(w_1|x) \propto \log P(w_1) + \cancel{\log}\, P_1 \log\lambda_1 + (P_1 - 1)\log x_1 - \lambda_1 x_1 - \log\Gamma(P_1)$$

$$\propto \log(1/2) + (3.2)\log 1 + 2.2\log x - x - \log\Gamma(3.2)$$

$$\propto -\log 2 + 2.2\log x - x - \log\Gamma(3.2)$$

Illy,

for w_2;

$$\log(w_2|x) = \log(1/2) + 7\log x - x - \log\Gamma(8)$$

decision boundary,

$x = x^*$

$$\log P(w_1|x) = \log P(w_2|x)$$

$$-\cancel{\log 2} + 2.2\log x - \cancel{x} - \log\Gamma(3.2) = -\cancel{\log 2} + 7\log x - \cancel{x} - \log\Gamma(8)$$

$$\Rightarrow -4.8\log x = -\log\Gamma(8) + \log\Gamma(3.2)$$

$$\Rightarrow x = e^{\frac{0.8859 - 8.5251}{-4.8}} = e^{1.5914}$$

$$\boxed{\Rightarrow x = 4.909}$$

Type-1 & Type-2 Errors

$\frac{1}{1}=1$

$P(\text{Type-1 Error}) = P(x > x^* | w_1)$

i.e $P(x > 4.909)$ where $x \sim \gamma(3.2, x)$

$= 1 - P(x < 4.909)$

cdf of γ function $= \frac{\gamma(k, x)}{\Gamma(k)} = \frac{\gamma(3.2, 99)}{\Gamma(3.2)} = \frac{2.30}{2.42} = 0.95$

$\therefore$ Type-1 error $= 1 - 0.95 = \boxed{0.05}$

Type-2 error $= P(x < x^* | w_2)$

i.e $P(x < 4.909)$

$= \frac{\gamma(8, 4.909)}{\Gamma(8)} = \frac{0.008}{5040}$

$\boxed{\beta = 0.002}$

(e) $P_1 = P_2 = 4$; $c = 2$, $k = 2$, $\lambda_1 = 8$, $\lambda_2 = 0.3$ $P(w_1) = 1/4$ $P(w_2) = 3/4$

decision boundary; using result from part (i)

$-\log 4 + \sum_{j=1}^{2} 4 \log \lambda_j + 3 \log x_j - \lambda_j x_j - \log \Gamma(4)) = \log 3/4 + \sum_{j=1}^{2} 4 \log \lambda_j + 3 \log x_j - \lambda_j x_j - \log \Gamma(4)$

$- 8x_1 + 4 \log 0.3 + 3 \log x_2 - 0.3 x_2 = \log 3/4$

$= -\log 4 + 4 \log 8 + 3 \log x_1 - 8x_1 + 4 \log 0.3 + 3 \log x_2 - \lambda_2 x_2$
$+ 4 \log \lambda_1 + 3 \log x_1 + 4 \log 0.3 + 3 \log x_2 - \lambda_2 x_2$

$\Rightarrow -\log 4 = \log 3/4$

$\boxed{\Rightarrow -\log 4 = \log \frac{3}{4}}$ cannot happen.

-0.6

there are no x terms.

$\boxed{\Rightarrow f(x_1, x_2) = 0}$

2.

$x_i|w_j \sim Lap(m_{ij}, \lambda_i)$

$P(x_i|w_j) = \frac{\lambda_i}{2} e^{-\lambda_i|x_i - m_{ij}|}$, $\lambda_i > 0$

$i \in \{1, 2, --- K\}$ feature i,
$j \in \{1, 2, --- C\}$ class j

For minimum error rate classifier, we choose Baye's rule

$$P(w_j|x) = \frac{P(x|w_j)P(w_j)}{P(x)} \propto P(x|w_j)P(w_j)$$

Since independent,

$$P(x|w_j) = \prod_{i=1}^{K} P_{x_i|w_j}(x_i|w_j) = \prod_{i=1}^{K} \frac{\lambda_i}{2} e^{-\lambda_i|x_i - m_{ij}|} \quad \text{— (1)}$$

to maximise $P(x|w_j)$, we need to minimise exponents.

i.e $\sum_{i=1}^{K} -\lambda_i|x_i - m_{ij}|$ $\boxed{= \min \sum_{i=1}^{K} \lambda_i|x_i - m_{ij}|}$ — (2)

the above expression is similar/ same as that of Manhattan Distance. {Sum of absolute differences weighted by λ_i

When Does it become, minimum Manhattan Distance classifier?

From (1) when $\lambda_i = \lambda$

(2) becomes,

$$w^* = \arg\min \lambda \sum_{i=1}^{K} |x_i - m_{ij}|$$

Since λ = const, it can be ignored and thus it becomes the minimum Manhattan distance classifier. when all λ_i are equal

(3)

a.

$P(\omega_1) = 1/10$

$P(\omega_2) = 1/5$

$P(\omega_3) = 1/2$

$P(\omega_4) = 1/5$

$$R(\alpha_i|x) = \sum_{j=1}^{4} (\alpha_i|\omega_j)\, P(\omega_j|x)$$

$$\Rightarrow R(\alpha_1|x) = \lambda(\alpha_1|\omega_1)P(\omega_1|x) + \lambda(\alpha_1|\omega_2)P(\omega_2|x) + \lambda(\alpha_1|\omega_3)P(\omega_3) + \lambda(\alpha_1|\omega_4)P(\omega_4)$$

& so on for $\alpha_2, \alpha_3, \alpha_4$

to calculate, $P(\omega_j|x)$,

$$P(\omega_j|x) = \frac{P(x|\omega_j)\,P(\omega_j)}{P(x)}$$

$$P(x) = \sum_{j=1}^{4} P(x|\omega_j)P(\omega_j)$$

$$\Rightarrow P(x) = P(x|\omega_1)P(\omega_1) + P(x|\omega_2)P(\omega_2) + P(x|\omega_3)P(\omega_3) + P(x|\omega_4)P(\omega_4)$$

$$\therefore P(1) = (1/3)(1/10) + (1/2)(1/5) + (1/6)(1/3) + (2/5)(1/5)$$

$$\Rightarrow P(1) = 0.2966$$

$$P(2) = (1/3)(1/10) + (1/4)(1/5) + (1/3)(1/2) + (2/5)(1/5)$$

$$= 1/30 + 1/20 + 1/6 + 2/25$$

$$\Rightarrow P(2) = 0.33$$

for $x=3$,

$$P(3) = (1/3)(1/10) + (1/4)(1/5) + (1/2)(1/2) + (1/5)(1/5)$$

$$\Rightarrow P(3) = 0.3733.$$

?S?

NAN

①

$$\therefore P(w_j|x) = \frac{P(x|w_j)\,P(w_j)}{P(x)}$$

---

for $w_1, x=1$

$$P(w_1|1) = \frac{P(1|w_1)\,P(w_1)}{P(1)} = \frac{(1/3)(1/10)}{0.2966} = 0.1123$$

for $w_2, x=1$

$$P(w_2|1) = \frac{P(1|w_2)\,P(w_2)}{P(1)} = \frac{(1/2)(1/5)}{0.2966} = 0.3373$$

for $w_3, x=1$

$$P(w_3|1) = \frac{P(1|w_3)\,P(w_3)}{P(1)} = \frac{(1/6)(1/2)}{0.2966} = 0.2810$$

$$P(w_4|1) = \frac{P(1|w_4)\,P(w_4)}{P(1)} = \frac{(2/5)(1/5)}{0.2966} = 0.2696$$

---

for $x=2$,

$$P(w_1|2) = \frac{P(2|w_1)\,P(w_1)}{P(2)} = \frac{(1/3)(1/10)}{0.33} = 0.1010$$

$$P(w_2|2) = \frac{P(2|w_2)\,P(w_2)}{P(2)} = \frac{(1/4)(1/5)}{0.33} = 0.1515$$

$$P(w_3|2) = \frac{P(2|w_3)\,P(w_3)}{P(2)} = \frac{(1/3)(1/2)}{0.33} = 0.5051$$

$$P(w_4|2) = \frac{P(2|w_4)\,P(w_4)}{P(2)} = \frac{(2/5)(1/5)}{0.33} = 0.2424$$

---

for $x=3$,

$$P(w_1|3) = \frac{P(3|w_1)\,P(w_1)}{P(3)} = \frac{(1/3)(1/10)}{0.3733} = 0.0893$$

$$P(w_2|3) = \frac{P(3|w_2)\,P(w_2)}{P(3)} = \frac{(1/4)(1/5)}{0.3733} = 0.1339$$

$$P(w_3|3) = \frac{P(3|w_3)\,P(w_3)}{P(3)} = \frac{(1/2)(1/2)}{0.3733} = 0.6696$$

$$P(w_4|3) = \frac{P(3|w_4)\,P(w_4)}{P(3)} = \frac{(1/5)(1/5)}{0.3733} = 0.1072$$

Conditional Risks,

$R(\alpha_i|x_1)$

when $\alpha_i = \alpha_1$

$\boxed{x = x_1} = 1$

$$R(\alpha_1|1) = \sum_{j=1}^{4} \lambda(\alpha_1|\omega_j)\, P(\omega_j|1)$$

$$= \lambda(\alpha_1|\omega_1)P(\omega_1|1) + \lambda(\alpha_2|\omega_2)P(\omega_2|1) + \lambda(\alpha_1|\omega_3)P(\omega_3|1) + \lambda(\alpha_1|\omega_4)P(\omega_4|1)$$

$$= 0 + 2(0.3373) + 3(0.2810) + 4(0.2696)$$

$$\boxed{\Rightarrow R(\alpha_1|1) = 2.5961}$$

$R(\alpha_2|1)$

$$= \lambda(\alpha_2|\omega_1)P(\omega_1|1) + \lambda(\alpha_2|\omega_2)P(\omega_2|1) + \lambda(\alpha_2|\omega_3)P(\omega_3|1) + \lambda(\alpha_2|\omega_4)P(\omega_4|1)$$

$$= 1(0.1123) + 0 + (1)(0.2810) + 8(0.2696)$$

$$= \underline{2.5501}$$

$$R(\alpha_3|1) = \lambda(\alpha_3|\omega_1)P(\omega_1|1) + \lambda(\alpha_3|\omega_2)P(\omega_2|1) + \lambda(\alpha_3|\omega_3)P(\omega_3|1) + \lambda(\alpha_3|\omega_4)P(\omega_4|1)$$

$$= 3(0.1123) + 2(0.3373) + 0 + 2(0.2696)$$

$$\Rightarrow \boxed{R(\alpha_3|1) = 1.5507}$$

$\therefore$ for $x = 1$,

$\cancel{R(\alpha_1|1) = 2.596}$

$\cancel{R(\alpha_2|1) = 2.}$

$R(\alpha_4|1)$:

$= \lambda(\alpha_4|w_1)P(w_1|1) + \lambda(\lambda_4|w_2)P(w_2|1) + \lambda(\lambda_4|w_3)P(w_3|1)$
$+ \lambda(\lambda_4|w_4)P(w_4|1)$

$= 5(0.1123) + 3(0.3373) + 1(0.2810) + 0$

$= 1.8544.$

for $x = 2$,

$R(\alpha_1|2) = \lambda(\alpha_1|w_1)P(w_1|2) + \lambda(\alpha_2|w_2)P(w_2|2) + \lambda(\alpha_3|w_3)P(w_3|2)$
$+ \lambda(\alpha_1|w_4)P(w_4|2)$

$= 0 + 2(0.1515) + 3(0.505) + 4(0.2424)$

$= 2.7879$

$R(\alpha_2|2) = \lambda(\alpha_2|w_1)P(w_1|2) + \lambda(\alpha_2|w_2)P(w_2|2) + \lambda(\alpha_2|w_3)P(w_3|2)$
$+ \lambda(\alpha_2|w_4)P(w_4|2)$

$= 1(0.1010) + 0 + 1(0.505) + 8(0.2424)$

$= 2.5453$

$R(\alpha_3|2) = \lambda(\alpha_3|w_1)P(w_1|2) + \lambda(\alpha_3|w_2)P(w_2|2) + \lambda(\lambda_3|w_3)P(w_3|2)$
$+ \lambda(\alpha_3|w_4)P(w_4|2)$

$= 3(0.1010) + 2(0.1515) + 0 + 2(0.2424)$

$\boxed{= 1.0908}$

$R(\alpha_4|2) = \lambda(\alpha_4|w_1)P(w_1|2) + \lambda(\alpha_3|w_2)P(w_2|2) + \lambda(\lambda_4|w_3)P(w_3|2)$
$+ \lambda(\lambda_4|w_4)P(w_4|2)$

$= 5(0.1010) + 3(0.1515) + 1(0.5051) + 0$

$= 1.4646$

for $x=3$

$$R(\alpha_1|3) = \lambda(\alpha_1|w_1)P(w_1|3) + \lambda(\alpha_1|w_2)P(w_2|3) + \lambda(\alpha_1|w_3)P(w_3|3)$$
$$\lambda(\alpha_1|w_4)P(w_1|4)$$

$$= 0(0.0893) + 2(0.1339) + 3(0.6696) + 4(0.1072)$$

$$= 2.7054$$

$$R(\alpha_2|3) = 1(0.0893) + 0(0.1339) + 1(0.6696) + 8(0.1072)$$

$$= 1.6165$$

$$R(\alpha_3|3) = 3(0.0893) + 2(0.1339) + 0(0.6696) + 2(0.1072)$$

$$\boxed{= 0.7501}$$

$$R(\alpha_4|3) = 5(0.0893) + 3(0.1339) + 1(0.6696) + 0(0.1072)$$
$$= 1.5178$$

b) Overall Risk

$$R = \sum_{i=1}^{3} R(\alpha(x_i)|x_i)\, P(x_i)$$

$\alpha(x_i)$ = decision rule minimising conditional risk for x_i

(marked in color "red" in part (a))

$$R = R(\alpha_3|1)P(1) + R(\alpha_3|2)P(2) + R(\alpha_3|3)P(3)$$
$$= 1.5507(0.2966) + 1.0908(0.33) + (0.7501)(0.3733)$$
$$= 0.4599 + 0.3600 + 0.2799$$
$$\boxed{R = 1.0998}$$