HW 3

Ruijie Song

Mar.4.2021

Ex1.

ECE 595 HW 3

1. a) $tr[A] = \sum_{i=1}^{d} [A]_{i,i}$. $A \in \mathbb{R}^{d \times d}$

$tr(xx^TA) = \cancel{tr(x(A^Tx)^T) = x^TA^Tx = x^TAx}$

$\cancel{= tr(x(Ax^T)) = tr}$

$= tr(x(A^Tx)^T) = x^TA^Tx = x^TAx$.

b) $\Sigma \in \mathbb{R}^{d \times d}$

$$\sum_{n=1}^{n} (x_n - \mu)^T \Sigma^{-1} (x_n - \mu) = tr\left[\Sigma^{-1} \sum_{n=1}^{N} (x_n - \mu)(x_n - \mu)^T\right].$$

$$\therefore P(D|\Sigma) = \frac{1}{(2\pi)^{Nd/2}} |\Sigma^{-1}|^{N/2} \exp\left\{-\frac{1}{2} tr\left[\Sigma^{-1} \sum_{n=1}^{N} (x_n - \mu)(x_n - \mu)^T\right]\right\}$$

c) $\tilde{\Sigma} = \frac{1}{N} \sum_{n=1}^{N} (x_n - \mu)(x_n - \mu)^T$, $A = \Sigma^{-1}\tilde{\Sigma}$, $\lambda_1, \dots \lambda_d$ eigenvalue of A.

$tr[A] = \sum_{i=1}^{d} \lambda_i$

$$P(D|\Sigma) = \frac{1}{(2\pi)^{Nd/2}} |\Sigma^{-1}|^{N/2} \exp\left\{-\frac{N}{2} tr[A]\right\}$$

$$= \frac{1}{(2\pi)^{Nd/2}} |\Sigma^{-1}|^{N/2} \exp\left(-\frac{N}{2} \sum_{i=1}^{d} \lambda_i\right)$$

$A = \Sigma^{-1}\tilde{\Sigma}$ $\quad \Sigma^{-1} = \tilde{\Sigma}^{-1} A$. $\quad (\Sigma^{-1})^{N/2} = (\tilde{\Sigma}^{-1} A)^{N/2}$

$$P(D|\Sigma) = \frac{1}{(2\pi)^{Nd/2} |\tilde{\Sigma}|^{N/2}} \left(\prod_{i=1}^{d} \lambda_i\right)^{N/2} \exp\left\{-\frac{N}{2} \sum_{i=1}^{d} \lambda_i\right\}$$

d) $\log P(D|\Sigma) = -\frac{Nd}{2}\log(2\pi) - \frac{N}{2}\log(|\tilde{\Sigma}|) + \frac{N}{2}\sum_{i=1}^{d}\log(\lambda_i)$

$+ \left(-\frac{N}{2}\sum_{i=1}^{d}\lambda_i\right)$

$\underset{\lambda}{\arg\max}\ \log(P(D|\Sigma)) = \underset{\lambda}{\arg\min}\,(-P(D|\Sigma))$

$= \underset{\lambda}{\arg\min}\left(-\frac{N}{2}\sum_{i=1}^{d}\log(\lambda_i) + \frac{N}{2}\sum_{i=1}^{d}\lambda_i\right)$

$\frac{\partial P(D|\Sigma)}{\partial \lambda_i} = -\frac{N}{2}\sum_{i=1}^{d}\frac{1}{\lambda_i} + \frac{N}{2}d = 0$

$\lambda_i = 1$

since $A = I$, $\tilde{\Sigma} = A \cdot \Sigma = \hat{\Sigma}_{ML}$

e) $\cancel{\frac{d}{d\Sigma}\log(P(D|\Sigma)) = \frac{d}{d\Sigma}\left(\frac{-Nd}{2}\log 2\pi - \frac{N}{2}\log(|\tilde{\Sigma}|) + \frac{N}{2}\sum_{i=1}^{d}\log(\lambda_i)\right.}$

$\cancel{\left. - \frac{N}{2}\sum_{i=1}^{d}\lambda_i\right)}$

f) $\frac{d}{d\Sigma}\log(P(D|\Sigma)) = \frac{d}{d\Sigma}\left(\frac{N}{2}\log|\Sigma| + \frac{N}{2}\cancel{\log(2\pi)}d\right.$

$\left. + \sum_{n=1}^{N}\left\{\frac{1}{2}(x_n - \mu)^T\Sigma^{-1}(x_n-\mu)\right\}\right)$

$= \frac{N}{2}\cancel{\frac{1}{\Sigma}} + \sum_{n=1}^{N}(-1)\frac{1}{2}(x_n-\mu)^T(x_n-\mu)\cancel{\Sigma^{-2}}\ \Sigma^{-1} = 0$

$\cancel{\frac{N}{2}}\ \cancel{\frac{1}{2}}\sum_{n=1}^{N}(x_n-\mu)^T(x_n-\mu)\cancel{\Sigma} = \cancel{0}\ \frac{N}{\cancel{2}}\Sigma$

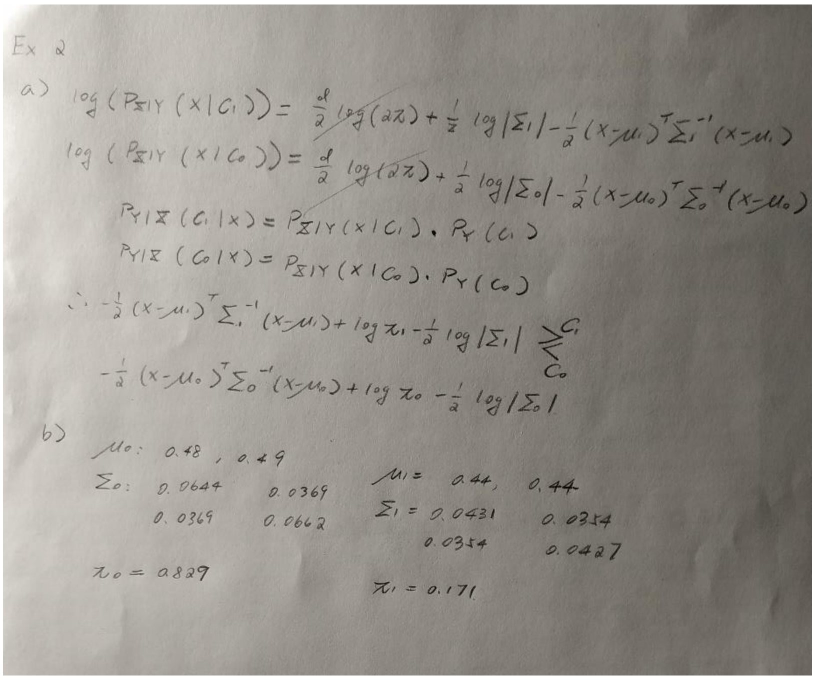
$\therefore \Sigma = \frac{1}{N}\sum_{n=1}^{N}(x_n-\mu)(x_n-\mu)^T$

g) $\hat{\mu} = \frac{1}{N}\sum_{n=1}^{N}x_n$,

$\hat{\Sigma}_{\cancel{unbiased}\ \text{unbias}} = \frac{1}{N-1}\sum_{n=1}^{N}(x_n-\hat{\mu})(x_n-\hat{\mu})^T$

Ex2.

a & b



Ex 2

a) $\log(P_{X|Y}(x|C_1)) = \cancel{\frac{d}{2}\log(2\pi)} + \frac{1}{2}\log|\Sigma_1| - \frac{1}{2}(x-\mu_1)^T\Sigma_1^{-1}(x-\mu_1)$

$\log(P_{X|Y}(x|C_0)) = \cancel{\frac{d}{2}\log(2\pi)} + \frac{1}{2}\log|\Sigma_0| - \frac{1}{2}(x-\mu_0)^T\Sigma_0^{-1}(x-\mu_0)$

$P_{Y|X}(C_1|x) = P_{X|Y}(x|C_1) \cdot P_Y(C_1)$

$P_{Y|X}(C_0|x) = P_{X|Y}(x|C_0) \cdot P_Y(C_0)$

$\therefore -\frac{1}{2}(x-\mu_1)^T\Sigma_1^{-1}(x-\mu_1) + \log\pi_1 - \frac{1}{2}\log|\Sigma_1| \underset{C_0}{\overset{C_1}{\gtrless}}$

$-\frac{1}{2}(x-\mu_0)^T\Sigma_0^{-1}(x-\mu_0) + \log\pi_0 - \frac{1}{2}\log|\Sigma_0|$

b) μ_0: 0.48 , 0.49

Σ_0: 0.0644 0.0369
0.0369 0.0662

$\pi_0 = 0.829$

$\mu_1 =$ 0.44, 0.44

$\Sigma_1 =$ 0.0431 0.0354
0.0354 0.0427

$\pi_1 = 0.171$

c.

d.

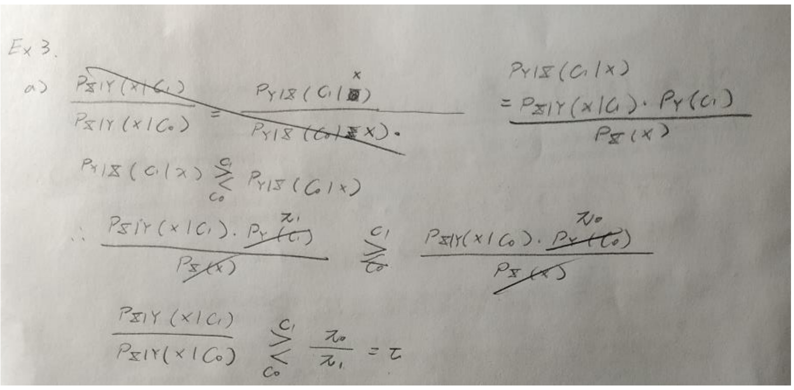
MAE = 0.08762355033904315

e.

No, it does not perform well. Maybe blurred background of this plot makes the prediction inaccurate.

Ex3.

a.



Ex 3.

a) $\cancel{\dfrac{P_{X|Y}(x|C_1)}{P_{X|Y}(x|C_0)} = \dfrac{P_{Y|X}(C_1|x)}{P_{Y|X}(C_0|x)}}$

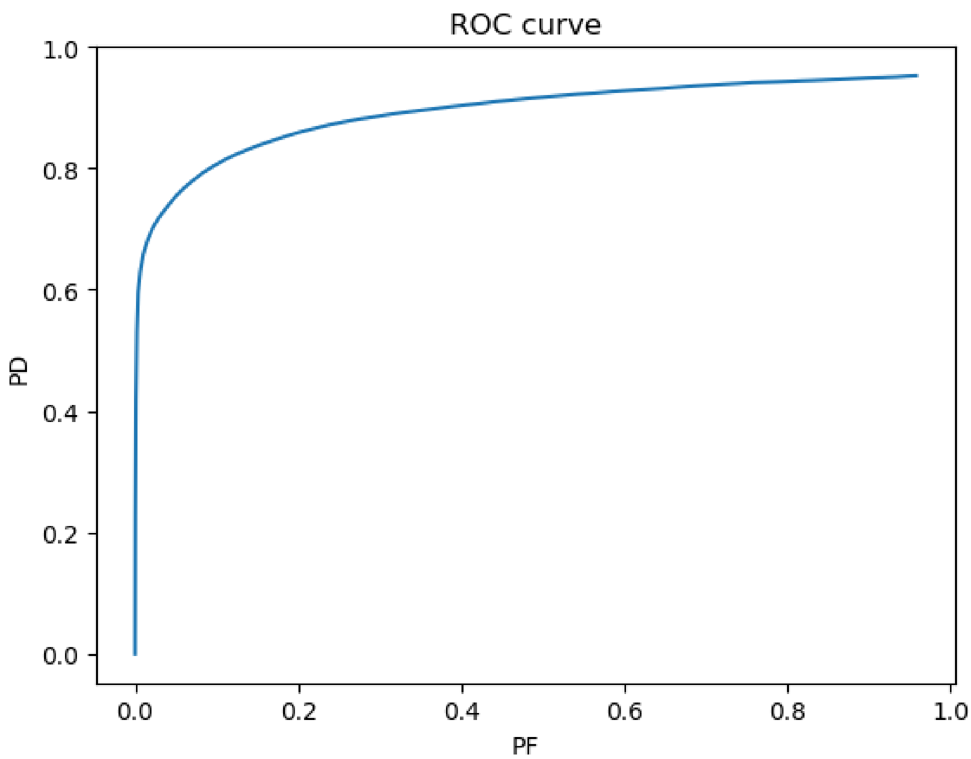
$$P_{Y|X}(C_1|x) = \frac{P_{X|Y}(x|C_1)\cdot P_Y(C_1)}{P_X(x)}$$

$$P_{Y|X}(C_1|x) \underset{C_0}{\overset{C_1}{\gtrless}} P_{Y|X}(C_0|x)$$

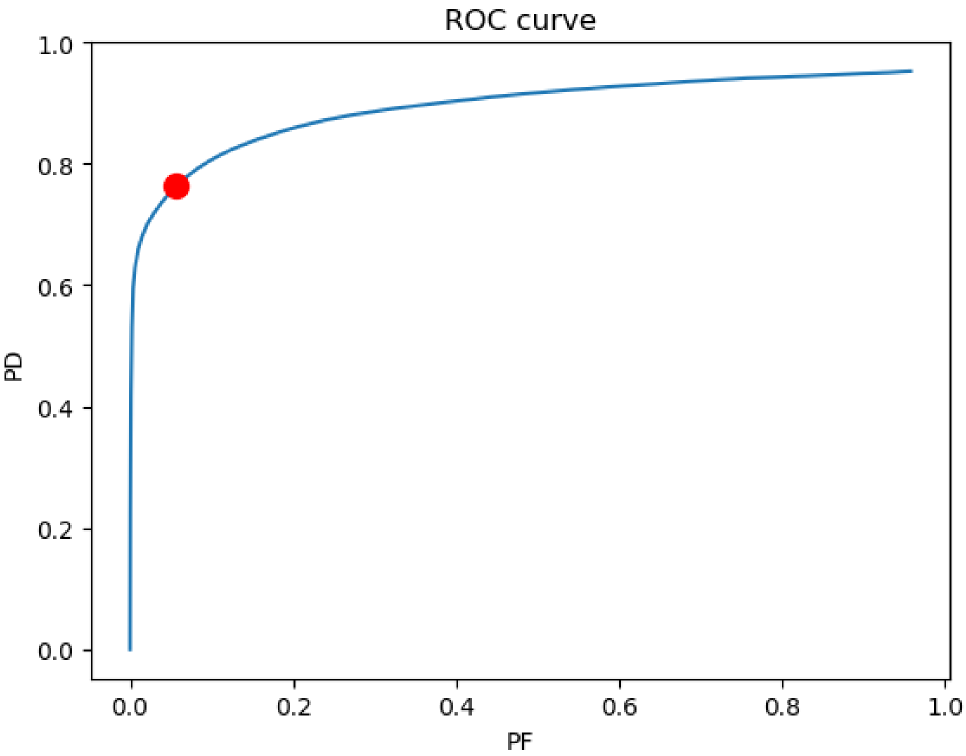
$$\therefore \frac{P_{X|Y}(x|C_1)\cdot \cancel{P_Y(C_1)}^{\pi_1}}{\cancel{P_X(x)}} \underset{C_0}{\overset{C_1}{\gtrless}} \frac{P_{X|Y}(x|C_0)\cdot \cancel{P_Y(C_0)}^{\pi_0}}{\cancel{P_X(x)}}$$

$$\frac{P_{X|Y}(x|C_1)}{P_{X|Y}(x|C_0)} \underset{C_0}{\overset{C_1}{\gtrless}} \frac{\pi_0}{\pi_1} = \tau$$

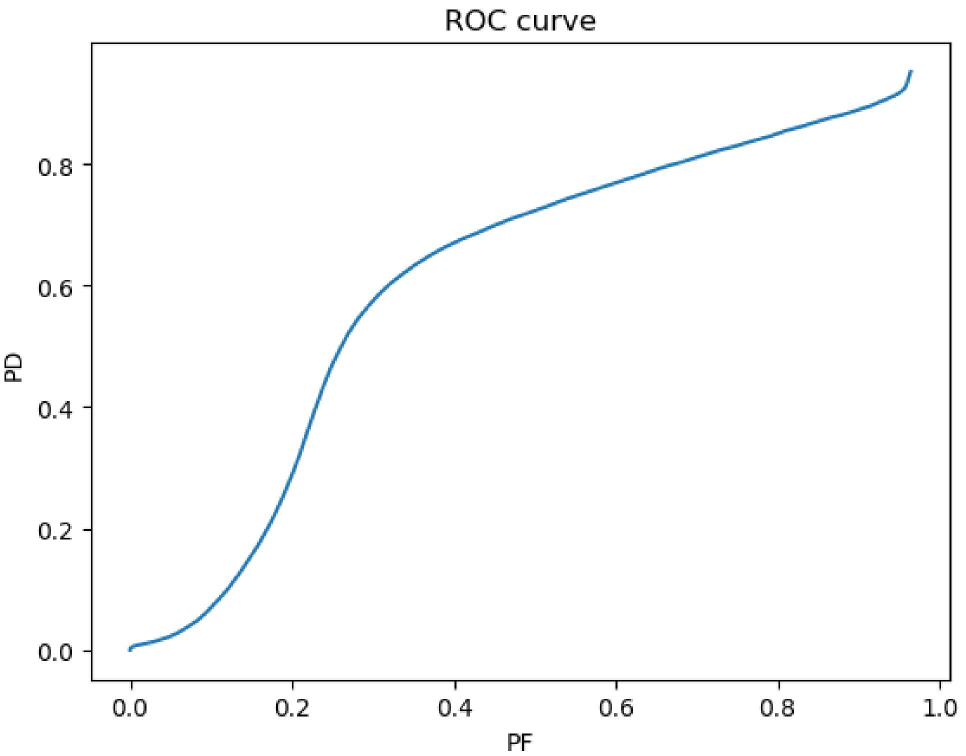
b.



c.


ROC curve
1.0
0.8
0.6
0.4
0.2
0.0
PD
0.0
0.2
0.4
0.6
0.8
1.0
PF

d.


ROC curve
0.8
0.6
0.4
0.2
0.0
PD
0.0
0.2
0.4
0.6
0.8
1.0
PF