ECE 50024

Homework 3

Parth Sagar Hasabnis phasabni@purdue.edu

Exercise 1:

a) $x \in \mathbb{R}^{d\times 1}$ $\quad \therefore x^T \in \mathbb{R}^{1\times d}$

$A \in \mathbb{R}^{d\times d}$

$$\underset{1\times d}{x^T} \; \underset{d\times d}{A} \; \underset{d\times 1}{x} \in \mathbb{R}^{1\times 1}$$

$\therefore x^T A x$ is a scalar

$tr[x^T A x] = x^T A x$, Trace of a scalar is itself

$tr[x^T A x] = tr[A x x^T]$, $tr[AB] = tr[BA]$

$\therefore x^T A x = tr[A x x^T]$

b) The likelihood func^n is given by

$$p(\mathcal{D}|\Sigma) = \prod_{i=1}^{N} \left(\frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp\left\{ \frac{-1}{2} (x_i - \mu)^T \Sigma^{-1} (x_i - \mu) \right\} \right.$$

$$= \frac{1}{(2\pi)^{Nd/2} |\Sigma|^{-N/2}} \exp\left\{ \sum_{i=1}^{n} \frac{-1}{2} (x_i - \mu)^T \Sigma^{-1} (x_i - \mu) \right\}$$

$$= \frac{1}{(2\pi)^{Nd/2}} |\Sigma^{-1}|^{N/2} \exp\left\{ \frac{-1}{2} \sum_{i=1}^{n} tr\left[\Sigma^{-1} (x_i - \mu)(x_i - \mu)^T \right] \right\} \qquad x^T A x = tr[A x x^T]$$

$$= \frac{1}{(2\pi)^{Nd/2}} |\Sigma^{-1}|^{N/2} \exp\left\{ \frac{-1}{2} tr\left[\Sigma^{-1} \sum_{i=1}^{n} (x_i - \mu)(x_i - \mu)^T \right] \right\} \qquad tr(A+B) = tr(A) + tr(B)$$

c) $p(D|\Sigma) = \frac{1}{(2\pi)^{Nd/2}} \times \frac{1}{|\Sigma|^{N/2}} \exp\left\{ \frac{-1}{2} tr\left[\Sigma^{-1} \sum_{i=1}^{n} (x_i - \mu)(x_i-\mu)^T \right]\right\}$

$= \frac{1}{(2\pi)^{Nd/2}} \times \frac{1}{|\Sigma|^{N/2}} \exp\left\{ \frac{-1}{2} tr\left[\Sigma^{-1} N \tilde{\Sigma} \right]\right\}$ ... $\tilde{\Sigma} = \frac{1}{N} \Sigma_{n=1}^{N} (x_n - \mu)(x_n - \mu)^T$

$= \frac{1}{(2\pi)^{Nd/2}} \times \frac{1}{|\Sigma|^{N/2}} \exp\left\{ -\frac{N}{2} tr[A]\right\}$ $\quad tr(NA) = N\,tr(A), \quad A = \Sigma^{-1}\tilde{\Sigma}$

$A = \Sigma^{-1}\tilde{\Sigma}$

$\Sigma A = \tilde{\Sigma}$

$\Sigma = \tilde{\Sigma} A^{-1}$

$|\Sigma| = |\tilde{\Sigma}|\,|A^{-1}|$

$|A^{-1}| = |A|^{-1}$

$|A| = \prod_{i=1}^{d} \lambda_i$ $\quad \lambda_1, \dots, \lambda_d$ are the eigenvalues of A

$\therefore p(D|\Sigma) = \frac{1}{(2\pi)^{Nd/2}} \frac{1}{|\tilde{\Sigma}|^{N/2}} \left(\prod_{i=1}^{d} \lambda_i\right)^{N/2} \exp\left\{ -\frac{N}{2} \sum_{i=1}^{d} \lambda_i \right\}$ $\quad tr[A] = \sum_{i=1}^{d} \lambda_i$

d) $\lambda_i = \underset{\lambda_i}{\operatorname{argmax}}\; p(D|\Sigma)$

$= \underset{\lambda_i}{\operatorname{argmin}}\; -\log p(D|\Sigma)$

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

$\frac{d(-\log p(D|\Sigma))}{d\lambda_i} = -\frac{N}{2} \times \frac{1}{\lambda_i} + \frac{N}{2} = 0$

$\frac{N}{2} = \frac{N}{2\lambda_i}$

$\therefore \lambda_i = 1$

$\therefore \lambda = [\lambda_1, \lambda_2, \dots, \lambda_d]^T_{1\times d} = [1 \;\; 1 \;\; 1 \dots \; 1]^T_{1\times d}$

e) Consider the eigen decomposition of A

$$A = Q \Lambda Q^{-1}$$

$$\Lambda = \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots \; \lambda_i \end{bmatrix} = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} = I$$

$$\therefore \; A = Q Q^{-1}$$

But $A = \Sigma^{-1} \tilde{\Sigma}$

$$\Sigma^{-1} \tilde{\Sigma} = Q Q^{-1}$$

This implies that $\Sigma^{-1} = \tilde{\Sigma}^{-1}$

$$\therefore \; \Sigma = \tilde{\Sigma}$$

This is the value of Σ that maximizes $p(D|\Sigma)$

$$\therefore \; \tilde{\Sigma} = \hat{\Sigma}_{ML} = \frac{1}{N} \Sigma_{n=1}^{N} (x_n - \mu)(x_n - \mu)^T$$

f) An alternative way of finding $\hat{\Sigma}_{ML}$ would be :

$$\hat{\Sigma}_{ML} = \underset{\Sigma}{\operatorname{argmax}} \; p(D|\Sigma)$$

$$= \underset{\Sigma}{\operatorname{argmin}} \; -\log(p(D|\Sigma))$$

We can take the gradient of $-\log(p(D|\Sigma))$ and equate it to 0. Under first order optimality conditions and if $-\log(p(D|\Sigma))$ is convex, we can get $\hat{\Sigma}_{ML}$

g) The unbiased estimate of Σ is given by:

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

Exercise 2:

a)

$$p_{Y|X}(C_1|x) = \frac{p_{X|Y}(x|C_1)\ p(C_1)}{p_X(x)} \qquad p_{Y|X}(C_0|x) = \frac{p_{X|Y}(x|C_0)\ p(C_0)}{p_X(x)}$$

$$p_X(x) = p_{X|Y}(x|C_1) + p_{X|Y}(x|C_0)$$

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

$$\frac{p_{X|Y}(x|C_1)\ p(C_1)}{\cancel{p_X(x)}} \underset{C_0}{\overset{C_1}{\gtrless}} \frac{p_{X|Y}(x|C_0)\ p(C_0)}{\cancel{p_X(x)}}$$

$$\pi_1 \frac{1}{\cancel{(2\pi)^{d/2}}|\Sigma_1|^{1/2}} \exp\left\{\frac{-1}{2}(x-\mu_1)^T\Sigma_1^{-1}(x-\mu_1)\right\} \underset{C_0}{\overset{C_1}{\gtrless}} \pi_0 \frac{1}{\cancel{(2\pi)^{d/2}}|\Sigma_0|^{1/2}} \exp\left\{\frac{-1}{2}(x-\mu_0)^T\Sigma_0^{-1}(x-\mu_0)\right\}$$

Taking log on both sides

$$\frac{-1}{2}(x-\mu_1)^T\Sigma_1^{-1}(x-\mu_1) + \log\pi_1 - \frac{1}{2}\log|\Sigma_1| \underset{C_0}{\overset{C_1}{\gtrless}} \frac{-1}{2}(x-\mu_0)^T\Sigma_0^{-1}(x-\mu_0) + \log\pi_0 - \frac{1}{2}\log|\Sigma_0|$$

b)

- The first few values in μ_0 are: [[0.48249575] [0.4864399]]
- The first few values in μ_1 are: [[0.44080734] [0.43871359]]
- The first few values in Σ_0 are:

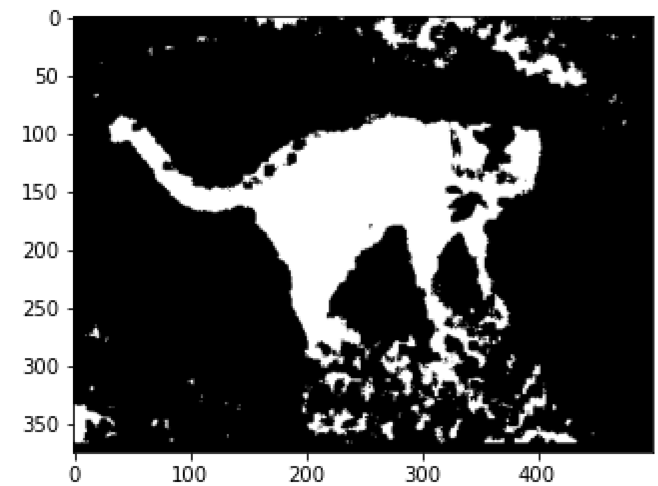
  [[0.064484 0.0369168]

  [0.0369168 0.06623457]]
- The first few values in Σ_1 are:

  [[0.04307832 0.03535405]

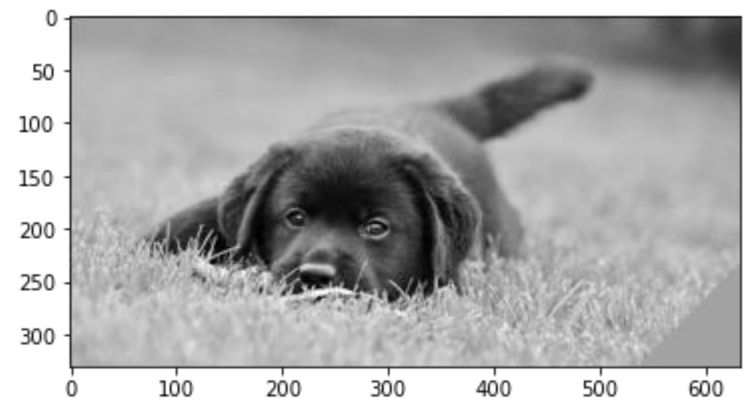
  [0.03535405 0.0424875]]
- π_0 = 0.83, π_1 = 0.17

c) Predicted Binary Mask

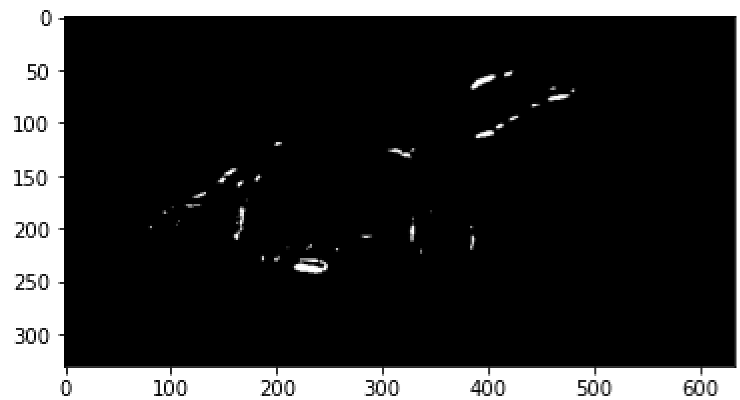


d) Mean Absolute Error = 0.087642

e) Test image and mask



Test image



Test image mask

The testing image does not perform well with the given Bayesian decision rule. This could be because the Class 1 pixels in the test image do not follow a similar distribution to the training pixels, and hence as misclassified.

Exercise 3:

a)

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

$$\frac{P_{X|Y}(x|C_1)}{P_{X|Y}(x|C_0)} = \frac{(2\pi)^{d/2}}{(2\pi)^{d/2}} \frac{|\Sigma_0|^{1/2}}{|\Sigma_1|^{1/2}} \frac{\exp\left\{-\frac{1}{2}(x-\mu_1)^T \Sigma_1^{-1}(x-\mu_1)\right\}}{\exp\left\{-\frac{1}{2}(x-\mu_0)^T \Sigma_0^{-1}(x-\mu_0)\right\}}$$

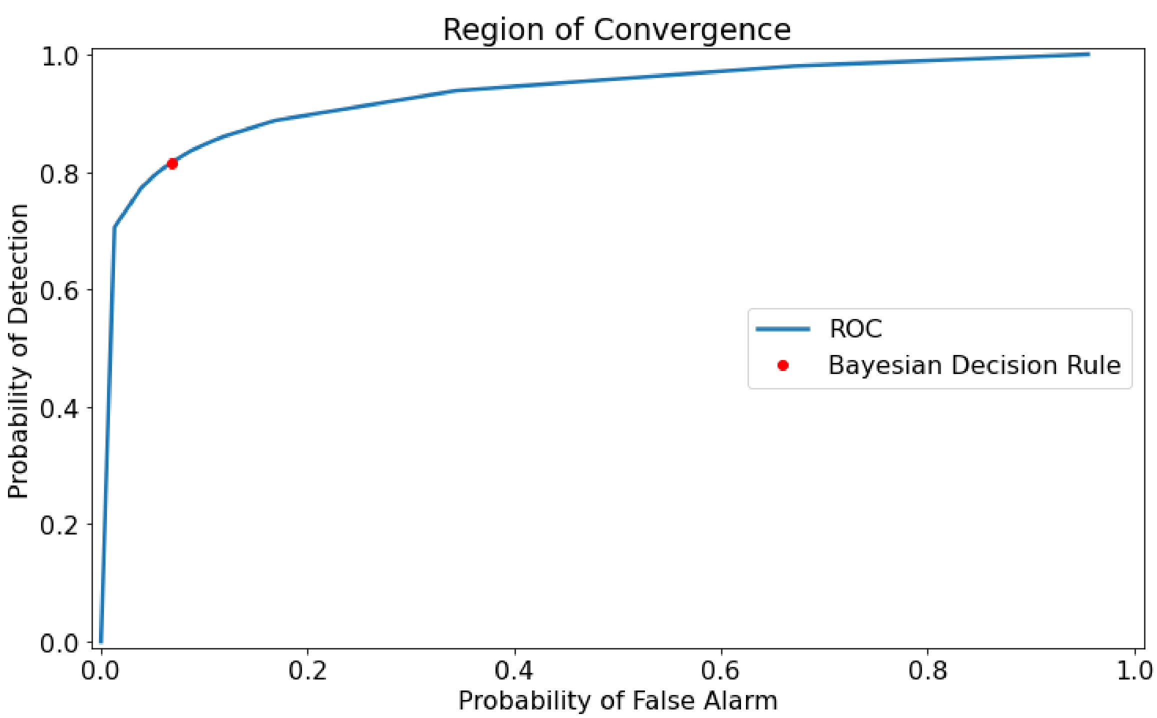
$$= \left(\frac{|\Sigma_0|}{|\Sigma_1|}\right)^{1/2} \exp\left\{\frac{1}{2}\left[(x-\mu_0)^T \Sigma_0^{-1}(x-\mu_0) - (x-\mu_1)^T \Sigma_1^{-1}(x-\mu_1)\right]\right\} = \tau$$

But, $\frac{(2\pi)^{d/2}}{(2\pi)^{d/2}} \frac{|\Sigma_0|^{1/2}}{|\Sigma_1|^{1/2}} \frac{\exp\left\{-\frac{1}{2}(x-\mu_1)^T \Sigma_1^{-1}(x-\mu_1)\right\}}{\exp\left\{-\frac{1}{2}(x-\mu_0)^T \Sigma_0^{-1}(x-\mu_0)\right\}} - \frac{\pi_0}{\pi_1}$

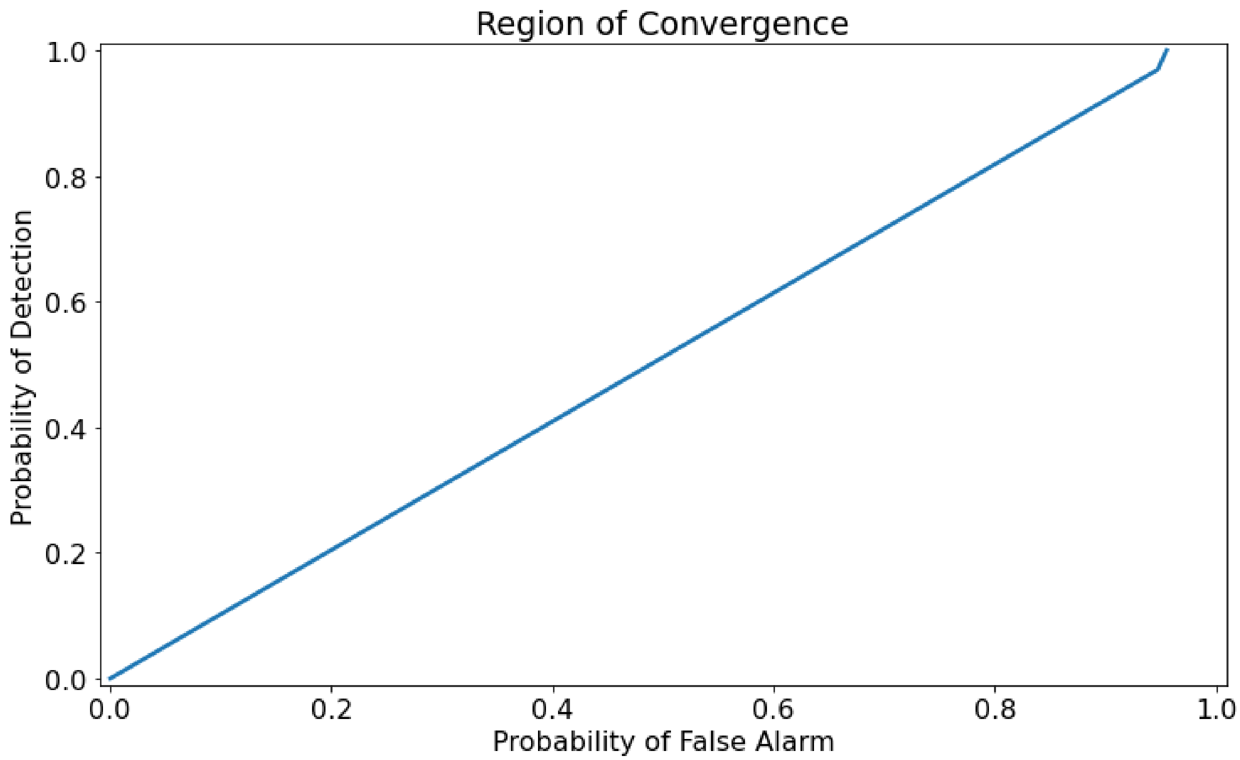
(From Exercise 2)

$$\therefore \boxed{\tau = \frac{\pi_0}{\pi_1}}$$

c)



d)


Region of Convergence
Probability of Detection
Probability of False Alarm
0.0
0.2
0.4
0.6
0.8
1.0

APPENDIX