ECBM 6040: Homework # 1

Name: Siddharth Aman Varshney Uni: sav2125

Problem C

(i) $p_X(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$

$$y = \frac{-1}{\lambda} \ln(x)$$

Now, $p_Y(y) = p_X(g^{-1}(y)) \left|\frac{\partial x}{\partial y}\right|$ — (1)

$$x = g^{-1}(y) = e^{-\lambda y}$$

$$\left|\frac{\partial x}{\partial y}\right| = \left|\frac{\partial (g^{-1}(y))}{\partial y}\right| = \left|\frac{\partial (e^{-\lambda y})}{\partial y}\right| = |-\lambda e^{-\lambda y}|$$

Substituting $\frac{\partial x}{\partial y}$ in (1), we get

$$p_Y(y) = p_X(e^{-\lambda y})(+\lambda e^{-\lambda y})$$

Now $x \in [0, 1]$

When $x = 0$

$e^{-\lambda y} = 0$

$\Rightarrow y = \infty$

When $x = 1$, $e^{-\lambda y} = 1$

$\Rightarrow$ $y = 0$

$\therefore$ $0 \leq x \leq 1$, $0 \leq y \leq \infty$

Hence,

$$p_y(y) = \begin{cases} +\lambda e^{-\lambda y} & 0 \leq y < \infty \\ 0 & \text{otherwise} \end{cases}$$

(ii) $$p(x = x, y = y) = \begin{cases} 3(xy^2 + yx^2) & \forall x, y \in [0,1] \\ 0 & \text{otherwise} \end{cases}$$

$$p(x = x) = \int_y p(x = x, y = y)\, dy$$

$$\Rightarrow p(x = x) = \int_0^1 3(xy^2 + yx^2)\, dy$$

$$= xy^3 + \frac{3x^2 y^2}{2} \Big|_0^1$$

$$= x + \frac{3x^2}{2}$$

$$p(Y=y) = \int_0^1 3(xy^2 + yx^2)\,dx$$

$$= \frac{3x^2y^2}{2} + yx^3 \Big|_0^1$$

$$= y + \frac{3y^2}{2}$$

$$E(x) = \int_0^1 x\,p(x)\,dx$$

$$= \int_0^1 x\left(x + \frac{3x^2}{2}\right) dx$$

$$= \int_0^1 x^2 + \frac{3x^3}{2}\,dx = \frac{x^3}{3} + \frac{3}{8}x^4 \Big|_0^1 = \frac{17}{24}$$

$$E(y) = \int_0^1 y\,p(y)\,dy = \frac{17}{24}$$

[since $p(x)$, $p(y)$ are symmetrical]

$$E(xy) = \int_0^1 \int_0^1 xy \cdot 3(xy^2 + yx^2)\, dy\, dx$$

$$= \int_0^1 \int_0^1 3x^2y^3 + 3x^3y^2 \, dy\, dx$$

$$= \int_0^1 \frac{3x^2y^4}{4} + x^3y^3 \Big|_0^1 \, dx$$

$$= \int_0^1 \frac{3x^2}{4} + x^3 \, dx$$

$$= \frac{x^4}{4} + \frac{x^3}{4} \Big|_0^1 = \frac{1}{2}$$

Since $E(xy) \neq E(x)E(y)$

x and y are not independent

Problem 2

(i) $X = \{x^{(1)}, \ldots, x^{(m)}\}$ $x^{(i)} \in \mathbb{R}^n$

is drawn from gaussian $N(\mu, \Sigma)$

Maximum likelihood estimate for μ

$$N(\mu, \Sigma) \sim \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right)$$

$$\frac{\partial}{\partial \mu}(N(\mu, \Sigma)) = 0$$

$\Rightarrow$ Maximizing log likelihood for μ, we get

$$\frac{\partial}{\partial \mu}\left(\sum_{i=1}^{m} \log \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x^{(i)}-\mu)^T \Sigma^{-1}(x^{(i)}-\mu)\right)\right)$$

$$= \frac{\partial}{\partial \mu}\left(\sum_{i=1}^{m} -\frac{n}{2}\log 2\pi - \frac{1}{2}\log \Sigma - \frac{1}{2}\left\{(x^{(i)}-\mu)^T \Sigma^{-1}(x^{(i)}-\mu)\right\}\right)$$

$$= \frac{\partial}{\partial \mu}\sum_{i=1}^{m}\left(-\frac{1}{2}(x^{(i)}-\mu)^T \Sigma^{-1}(x^{(i)}-\mu)\right) = 0$$

Now $\dfrac{\partial x^T x}{\partial x} = 2x^T$

$$\therefore \sum_{i=1}^{m} (x^{(i)} - \mu)^T \Sigma^{-1} = 0$$

$$\Rightarrow \mu = \frac{1}{m} \sum_{i=1}^{m} x^{(i)}$$

Maximising log likelihood for Σ

$$\frac{\partial}{\partial \Sigma} \sum_{i=1}^{m} \left(-\frac{n}{2} \log 2\pi - \frac{1}{2} \log |\Sigma| - \frac{1}{2} \left\{ (x^{(i)} - \mu)^T \Sigma^{-1} (x^{i} - \mu) \right\} \right)$$

$$= 0$$

Writing log likelihood in trace form, we get

$$l = \sum_{i=1}^{m} \left| -\frac{n}{2} \log 2\pi - \frac{1}{2} \log |\Sigma| - \frac{1}{2} \sum_{i=1}^{m} tr \left\{ (x^{(i)} - \mu)^T \right. \right.$$

$$\left. \Sigma^{-1} (x^{(i)} - \mu) \right\}$$

$$= -\frac{mn}{2} \log 2\pi + \frac{m}{2} \log |\Sigma^{-1}| - \frac{1}{2} \sum_{i=1}^{m} tr \left[(x^{(i)} - \mu) \right.$$

$$\left. (x^{(i)} - \mu)^T \Sigma^{-1} \right]$$

$$= -\frac{mn}{2}\log 2\pi + \frac{m}{2}\log|A| - \frac{1}{2}\sum_{i=1}^{m} tr\left[(x^{(i)}-\mu)(x^{(i)}-\mu)^T A\right]$$

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

Now $\frac{\partial \log A}{\partial A} = (A^{-1})^T$ and

$$\frac{\partial\, tr(BA)}{\partial A} = B^T$$

$$\therefore \frac{\partial l}{\partial A} = \frac{m}{2}(A^{-1})^T - \frac{1}{2}\sum_{i=1}^{m}\left[(x^{(i)}-\mu)(x^{(i)}-\mu)^T\right]$$

$$\frac{\partial l}{\partial A} = 0$$

$$\Rightarrow \frac{m}{2}\Sigma - \frac{1}{2}\sum_{i=1}^{m}(x^{(i)}-\mu)(x^{(i)}-\mu)^T = 0$$

$$\Rightarrow \Sigma = \frac{1}{m}\sum_{i=1}^{m}(x^{(i)}-\mu)(x^{(i)}-\mu)^T$$

(ii) μ Estimator:

An estimator is unbiased if ~~estimation~~ ~~of~~ ~~Exp~~ expectation of estimator equals true value.

$$E[(\bar{\mu})] = E\left[\frac{1}{m}\sum_{i=1}^{m} x^{(i)}\right]$$

$$= \sum_{i=1}^{m}\frac{E[x]}{m} = \frac{1}{m} \times m \times E[x]$$

$$= E[x] = \mu$$

Thus μ is unbiased

Σ estimator:

Let $s^2 = \frac{1}{m}\sum_{i=1}^{m}(x^{(i)}-\mu)^2$ be variance estimator

$$E[(s^2)] = E\left[\frac{1}{m}\sum_{i=1}^{m}(x^{(i)}-\mu)(x^{(i)}-\mu)^T\right]$$

$$= \frac{1}{m} E\left[\sum_{i=1}^{m} (x^{(i)})^2 - 2\sum_{i=1}^{m} x^{(i)}\mu + \sum_{i=1}^{m} \mu^2 \right]$$

$$= \frac{1}{m} \cdot E\left[\sum_{i=1}^{m} (x^{(i)})^2 - m\mu^2 \right]$$

$$= \frac{1}{m} E\left[\sum_{i=1}^{m} (x^{(i)})^2 \right] - E[\mu^2]$$

$$= E[x^{(i)2}] - E[\mu^2]$$

By definition of variance,

$$\sigma_x^2 = E[x^2] - E[x]^2$$

$$\therefore E[x^2] - E[\mu^2] = \sigma_x^2 + E[x^{(i)2}] - \sigma_x^2 - E[(x^{(i)})^2]$$

$$= \sigma_x^2 - \sigma_\mu^2$$

$$\sigma_\mu^2 = \mathrm{Var}[\mu] = \mathrm{Var}\left[\frac{1}{m} \cdot \sum_{i=1}^{m} x^{(i)}\right] = \frac{1}{m^2} \mathrm{Var}\left[\sum_{i=1}^{m} x^{(i)}\right]$$

Now, $\mathrm{Var}\left[\sum_{i=1}^{m} x^{(i)}\right] = \sum_{i=1}^{m} \mathrm{Var}[x] = N \times \mathrm{Var}[x]$

Thus $\sigma_{\mu}^2 = \frac{1}{m} \cdot Var[x] = \frac{1}{m} \cdot \sigma_x^2$

$\therefore \quad E[s^2] = \frac{m-1}{m} \sigma_x^2$

Since $E[s^2] \neq \sigma_x^2$,

Σ_1 is ~~un~~biased.