# Deep Learning Homework1

**UNI: yb2356** **Name: Yang Bai**

**PROBLEM C**

(i)

$$p_y(y) = p_{x(y)}\big(x(y)\big) * \left|\frac{dx}{dy}\right|$$

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

$$\left|\frac{dx}{dy}\right| = |\lambda e^{-\lambda y}|$$

$$p_y(y) = \begin{cases} 1 * \lambda e^{-\lambda y}, & 0 \le e^{-\lambda y} \le 1 \\ 0, & otherwise \end{cases}$$

$If\ \lambda > 0,$

$$p_y(y) = \begin{cases} \lambda e^{-\lambda y}, & y \ge 0 \\ 0, & otherwise \end{cases}$$

$If\ \lambda < 0,$

$$p_y(y) = \begin{cases} -\lambda e^{-\lambda y}, & y \le 0 \\ 0, & otherwise \end{cases}$$

(ii)

$$p(X = x) = \int p(X = x, Y = y)dy = 3\int_0^1 xy^2 + x^2y\ dy$$

$$= xy^3 + \frac{3x^2y^2}{2}\ |_0^1 = x + \frac{3x^2}{2}\ (0 \le x \le 1)$$

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

$$= yx^3 + \frac{3x^2y^2}{2}\ |_0^1 = y + \frac{3y^2}{2}(0 \le y \le 1)$$

$$\mathrm{E}[x] = \int p(X = x)\ x\ dx = 3\int_0^1 \frac{x^2}{3} + \frac{x^3}{2}\ dx$$

$$\mathrm{E}[x] = \frac{3}{9} + \frac{3}{8} = \frac{17}{24}$$

$$\mathrm{E}[y] = \int p(Y = y)\ y\ dx = 3\int_0^1 \frac{y^2}{3} + \frac{y^3}{2}\ dy$$

$$\mathrm{E}[y] = \frac{3}{9} + \frac{3}{8} = \frac{17}{24}$$

$$\mathrm{E}[xy] = \int p(X = x, Y = y)\ xy\ dxdy$$

$$\mathrm{E}[xy] = 3\int_0^1\int_0^1 x^2y^3 + x^3y^2\, dxdy = 3\int_0^1 \frac{y^3}{3} + \frac{y^2}{4}\, dy = \frac{1}{2}$$

X, y are not independent. Because $p(X=x) * p(Y=y) \neq p(X=x, Y=y)$.

## PROBLEM D

(i)

$$P(X) = \prod_{i=1}^{m} p(\vec{x_i}) = \prod_{i=1}^{m} \frac{1}{(2\pi)^{n/2}\sqrt{|\Sigma|}} \exp(-\frac{1}{2}(\vec{x_i} - \mu)^T \Sigma^{-1} (\vec{x_i} - \mu))$$

$$\log\ P(X) = \sum_{i=1}^{m} -\frac{1}{2}\log|\Sigma| - \frac{1}{2}(\vec{x_i} - \mu)^T \Sigma^{-1} (\vec{x_i} - \mu) + constant$$

Max over $\mu$:

$$\frac{\partial \log\ P(X)}{\partial \mu} = \sum_{i=1}^{m} -(\vec{x_i} - \mu)^T \Sigma^{-1} = 0$$

$$\mu = \frac{1}{m}\sum_{i=1}^{m} \vec{x_i}$$

Max over $\Sigma$:

$$\frac{\partial \log\ P(X)}{\partial \Sigma} = \frac{m}{2}\Sigma - \sum_{i=1}^{m} \frac{1}{2}(\vec{x_i} - \mu)^T (\vec{x_i} - \mu) = 0$$

$$\Sigma = \frac{1}{\mathrm{m}}\sum_{i=1}^{m} (\vec{x_i} - \vec{\mu})^T (\vec{x_i} - \vec{\mu})$$

(ii)

$\mu$ is an unbiased estimator. $\Sigma$ is a biased estimator.

Set $\bar{X} = \frac{1}{m}\sum_{i=1}^{m} \vec{x_i}$, $S = \frac{1}{\mathrm{m}}\sum_{i=1}^{m} (\vec{x_i} - \vec{\mu})^T (\vec{x_i} - \vec{\mu})$

$\bar{X}$ is a normal distribution: $\bar{X} \sim Normal(\vec{\mu}, \frac{1}{m}\Sigma)$

$$\mathrm{E}[\bar{X}] = \vec{\mu}$$

The expectation of estimator $\bar{X}$ is the same as its value, so $\vec{\mu}$ is an unbiased estimator.

According to the definition of Wishart distribution, $S \sim \mathrm{Wishart}(\mathrm{m} - 1, \frac{1}{\mathrm{m}}\Sigma)$.

$$\mathrm{E}[S] = \frac{\mathrm{m} - 1}{\mathrm{m}}\Sigma$$

$$\mathrm{E}[S] \neq S$$

The expectation of estimator S is different from its value, so $\Sigma$ is a biased estimator.