Homework week 2

Nguyễn Minh Đức
Student code: 11204838

August 2022

1 Question 1:

a) Gaussian distribution is normalized

We have :

$$\mathcal{N}(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\bigg(\frac{-(x-\mu)^2}{2\sigma^2}\bigg)$$

To proof Gaussian distribution is normalized, we have to proof

$$\int_{-\infty}^{\infty} \mathcal{N}(x|\mu,\sigma^2)dx = 1 \Leftrightarrow \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} exp\bigg(\frac{-(x-\mu)^2}{2\sigma^2}\bigg)dx = 1$$

Assume $\mu = 0$, $\sigma^2 = 1$

Let denote I = $\int_{-\infty}^{\infty} exp(\frac{-x^2}{2})dx$

We :

$$I^2 = \bigg[\int_{-\infty}^{\infty} exp\Big(\frac{-x^2}{2}\Big)dx\bigg]\bigg[\int_{-\infty}^{\infty} exp\Big(\frac{-y^2}{2}\Big)dy\bigg] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} exp\Big(-\frac{x^2+y^2}{2}\Big)dxdy$$

Converting to polar coordinates by substitute:

$$x = r\cos(\theta)$$
$$y = r\sin(\theta)$$

$$\begin{aligned}\rightarrow I^2 &= \int_0^{2\pi}\int_0^{\infty} exp\Big(\frac{-r^2}{2}\Big)rdrd\theta \\ &= \int_0^{2\pi} d\theta \\ &= 2\pi \\ \rightarrow I &= \sqrt{2\pi}\end{aligned}$$

$\rightarrow \int_{-\infty}^{\infty} \mathcal{N}(x|\mu, \sigma^2)dx = 1 \rightarrow$ Gaussian distribution is normalized
b) Expectation of Gaussian distribution is μ

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx$$

So:

$$E(X) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} xexp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right) dx$$

Substitute $t = \frac{x-\mu}{\sqrt{2}\sigma}$

$$\begin{aligned} E(X) &= \frac{\sqrt{2}\sigma}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (\sqrt{2}\sigma t + \mu) exp(-t^2) dt \\ &= \frac{1}{\sqrt{\pi}} \left(\sqrt{2}\sigma \int_{-\infty}^{\infty} texp(-t^2) + \mu \int_{-\infty}^{\infty} exp(-t^2) \right) \\ &= \frac{1}{\sqrt{\pi}} \left(\sqrt{2}\sigma \bigg] \frac{-1}{2} exp(-t^2) \bigg]_{-\infty}^{\infty} + \mu\sqrt{\pi} \right) \\ &= \frac{\mu\sqrt{\pi}}{\sqrt{\pi}} \\ &= \mu \end{aligned}$$

c) Variance of Gaussian distribution is σ^2

$$Var(X) = \int_{-\infty}^{\infty} x^2 f_X(x) - (E(X))^2$$

So:

$$Var(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x^2 exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right) dx - \mu^2$$

Subtitute t = $\frac{x-\mu}{\sqrt{2}\sigma}$

$$\begin{aligned}
Var(x) &= \frac{\sqrt{2}\sigma}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}(\sqrt{2}\sigma t+\mu)^2 exp(-t^2)dx-\mu^2 \\
&= \frac{1}{\sqrt{\pi}}\left(2\sigma^2\int_{-\infty}^{\infty}t^2 exp(-t^2)dt+2\sqrt{2}\sigma\mu\int_{-\infty}^{\infty}texp(-t^2)dt+\mu^2\int_{-\infty}^{\infty}exp(-t^2)\right)-\mu^2 \\
&= \frac{1}{\sqrt{\pi}}\left(2\sigma^2\int_{-\infty}^{\infty}t^2 exp(-t^2)dt+2\sqrt{2}\sigma\mu[\frac{-1}{2}exp(-t^2)]+\mu^2\sqrt{\pi}\right)-\mu^2 \\
&= \frac{2\sigma^2}{\sqrt{\pi}}\int_{-\infty}^{\infty}t^2 exp(-t^2)dt \\
&= \frac{2\sigma^2}{\sqrt{\pi}}\left(\left[\frac{-t}{2}exp(-t^2)\right]_{-\infty}^{\infty}+\frac{1}{2}\int_{-\infty}^{\infty}exp(-t^2)dt\right) \\
&= \frac{2\sigma^2}{\sqrt{\pi}}\frac{1}{2}\int_{-\infty}^{\infty}exp(-t^2)dt \\
&= \frac{2\sigma^2\sqrt{\pi}}{2\sqrt{\pi}} \\
&= \sigma^2
\end{aligned}$$

d) Multivariate Gaussian distribution is normalized

For D-dimensional vector, the multivariate Gaussian distribution takes the form

$$\mathcal{N}(x|\mu,\Sigma)=\frac{1}{(2\pi)^{D/2}}\frac{1}{|\Sigma|^{1/2}}exp(\frac{-1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu))$$

where μ is a D-dimensional vector, Σ is D x D matrix and $|\Sigma|$ denote the determinant of Σ.

The functional dependence of the Gaussian on X is throuugh the quadratic form:

$$\Delta^2=(x-\mu)^T\Sigma^{-1}(x-\mu)$$

Consider eigenvalue and eigenvector of Σ:

$$\Sigma u_i=\lambda u_i$$

Because Σ is a real, symmetric matrix:

$$\Sigma=\sum_{i=1}^{D}\lambda_i u_i u_i^T$$

$$\rightarrow \Sigma^{-1}=\sum_{i=1}^{D}\frac{1}{\lambda_i}u_i u_i^T \quad \text{and} \quad |\Sigma|^{\frac{1}{2}}=\prod_{j=1}^{D}\lambda_i^{1/2}$$

$$\begin{aligned}
\rightarrow \Delta^2 &= (x-\mu)^T \sum_{i=1}^{D} \frac{1}{\lambda_i} u_i u_i^T (x-\mu) \\
&= \sum_{i=1}^{D} \frac{y_i^2}{\lambda_i} \quad \text{where} \quad y_i = u_i^T (x-\mu)
\end{aligned}$$

In new Y coordinate system, the Gaussian distribution takes the form:

$$P(y) = \prod_{j=1}^{D} \frac{1}{(2\pi\lambda_j)^{1/2}} exp(-\frac{y_j^2}{2\lambda_j})$$

$$\rightarrow \int_{-\infty}^{\infty} p(y)dy = \prod_{j=1}^{D} \frac{1}{(2\pi\lambda_j)^{1/2}} exp(-\frac{y_j^2}{2\lambda_j}) dy_j = 1$$

2 Question 2:

a) The conditional of Gaussian distribution
Suppose X is a D-dimensional vector with Gaussian distribution $\mathcal{N}(x|\mu, \Sigma)$ and we partition X into two disjoint subsets X_a and X_b. Without the loss of generality, we can take X_a to form the first M components of X, with X_b comprising the remaining D - M components, so that

$$X = \begin{pmatrix} x_a \\ x_b \end{pmatrix} \quad \text{and} \quad \mu = \begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix}$$

For convenient, we denote: $\Lambda = \Sigma^{-1} \rightarrow \Lambda = \begin{pmatrix} \Lambda_{aa} & \Lambda_{ab} \\ \Lambda_{ba} & \Lambda_{bb} \end{pmatrix}$
We have:

$$\begin{aligned}
\frac{-1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu) &= -\frac{1}{2}(x_a-\mu_a)^T \Lambda_{aa} (x_a-\mu_a) - \frac{1}{2}(x_a-\mu_a)^T \Lambda_{ab} (x_b-\mu_b) \\
&\quad -\frac{1}{2}(x_b-\mu_b)^T \Lambda_{ba} (x_a-\mu_a) - \frac{1}{2}(x_b-\mu_b)^T \Lambda_{bb} (x_b-\mu_b) \\
&= -\frac{1}{2} x_a^T \Lambda_{aa}^{-1} x_a + x_a^T (\Lambda_{aa}\mu_a - \Lambda_{ab}(x_b-\mu_b)) + const
\end{aligned}$$

We can see that as a function of x_a, this is again a quadratic form, and hence the corresponding conditional disatribution $P(X_a|X_b)$ will be Gaussian,because this distribution is characterized by its mean and its variance. Compare with functional dependence of Gaussian.

$$\Delta^2 = -\frac{1}{2} x^T \Sigma^{-1} x + x^T \Sigma^{-1} \mu + const$$

$$
\begin{aligned}
\rightarrow \Sigma_{a|b} &= \Lambda_{aa}^{-1} \\
\mu_{a|b} &= \Sigma_{a|b}(\Lambda_{aa}\mu_a - \Lambda_{ab}(x_b - \mu_b)) \\
&= \mu_a - \Lambda_{aa}^{-1}\Lambda_{ab}(x_b - \mu_b)
\end{aligned}
$$

By using Schur complement:

$$
\begin{aligned}
\rightarrow \Lambda_{aa} &= (\Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba})^{-1} \\
\Lambda_{ab} &= -(\Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba})^{-1}\Sigma_{ab}\Sigma_{bb}^{-1}
\end{aligned}
$$

As a result:

$$
\begin{aligned}
\mu_{a|b} &= \mu_a + \Sigma_{ab}\Sigma_{bb}^{-1}(x_b - \mu_b) \\
\Sigma_{a|b} &= \Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba} \\
P(X_a|X_b) &= \mathcal{N}(x_{a|b}|\mu_{a|b}, \Sigma_{a|b})
\end{aligned}
$$

b) The marginal of Gaussian distribution

$$P(x_a) = \int p(X_a, X_b)dx_b$$

We need to integrate out xb by looking the quadratic form related to X_b

$$-\frac{1}{2}X_b^T\Lambda_{bb}x_b + x_b^Tm = -\frac{1}{2}(X_b - \Lambda_{bb}^{-1}m)^T\Lambda_{bb}(X_b - \Lambda_{bb}^{-1}m) + \frac{1}{2}m^T\Lambda_{bb}^{-1}m$$

$$\text{where} \quad m = \Lambda_{bb}\mu_b - \Lambda_{ba}(x_a - \mu_a)$$

The integration over x_b will take the form:

$$\int exp(-\frac{1}{2}(X_b - \Lambda_{bb}^{-1}m)^T\Lambda_{bb}(X_b - \Lambda_{bb}^{-1}m))dx_b$$

and the remaining term depend on X_a is:

$$
\begin{aligned}
&\frac{1}{2}(\Lambda_{bb}\mu_b - \Lambda_{ba}(X_a - \mu_a))^T\Lambda_{bb}^{-1}(\Lambda_{bb}\mu_b - \Lambda_{ba}(X_a - \mu_a)) - \frac{1}{2}X_a^T\Lambda_{aa}X_a \\
&\qquad + X_a^T(\Lambda_{aa}\mu_a + \Lambda_{ab}\mu_b) + const \\
&= -\frac{1}{2}X_a^T(\Lambda_{aa} - \Lambda_{ab}\Lambda_{bb}^{-1}\Lambda_{ba})X_a + X_a^T(\Lambda_{aa} - \Lambda_{ab}\Lambda_{bb}^{-1}\Lambda_{ba})^{-1}\mu_a + const
\end{aligned}
$$

So the covariance of the marginal distribution of $P(x_a)$ is given by:

$$
\begin{aligned}
E(X_a) &= \mu_a \\
\Sigma_a &= (\Lambda_{aa} - \Lambda_{ab}\Lambda_{bb}^{-1}\Lambda_{ba})^{-1} = \Sigma_{aa} \\
P(X_a) &= \mathcal{N}(X_a|\mu_a, \Sigma_{aa})
\end{aligned}
$$