Homework 2

Nguyễn Hoàng Long - 11202352 - DSEB 62

24/08/2022

1 Problem 1

1.1 a.

Normalization of Univariate Gaussian distribution is given by:

$$\int_{-\infty}^{\infty} p(x|\mu, \sigma^2) dx = 1$$

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

We will first prove the base case when the mean equals to zero ($\mu = 0$), which means that:

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

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

Let:

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

Then we will take the square of both sides:

$$I^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\left(\frac{-1}{2\sigma^2} x^2\right) \exp\left(\frac{-1}{2\sigma^2} y^2\right) dxdy$$

$$\iff I^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) dxdy$$

To conduct the integration, we will make the transformation from Cartesian coordinates (x, y) to **polar coordinates** (r, θ) by assuming:

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

where r and θ are arbitrary number and angle. By the trigonometric identity we also have:

$$\cos^2\theta + \sin^2\theta = 1$$
$$x^2 + y^2 = r^2$$

While transforming integrals between two coordinate systems, the Jacobian the change of variables is given by:

$$
\begin{aligned}
dxdy &= |J|drd\theta \\
&= \begin{vmatrix} \frac{\partial(x)}{\partial(r)} & \frac{\partial(x)}{\partial(\theta)} \\ \frac{\partial(y)}{\partial(r)} & \frac{\partial(y)}{\partial(\theta)} \end{vmatrix} \\
&= \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} \\
&= r\cos^2\theta + r\sin^2\theta \\
&= r \\
&\implies dxdy = rdrd\theta
\end{aligned}
$$

Substituting the above results to the expression of I then:

$$
\begin{aligned}
I^2 &= \int_0^{2\pi}\int_0^{\infty} \exp\left(-\frac{r^2}{2\sigma^2}\right) rdrd\theta \\
&= 2\pi \int_0^{\infty} \exp\left(-\frac{r^2}{2\sigma^2}\right) rdr \\
&= 2\pi\pi \int_0^{\infty} \exp\exp\left(-\frac{r^2}{2\sigma^2}\right) \frac{d(r^2)}{2} \\
&= \pi\left[\exp\left(-\frac{r^2}{2\sigma^2}\right)(-2\sigma^2)\right]_0^{\infty} \\
&= 2\pi\sigma^2
\end{aligned}
$$

We have $I = \sqrt{2\pi\sigma^2}$, to prove the case when mean is non zero, suppose $t = x - \mu$ so that:

$$
\begin{aligned}
\int_{-\infty}^{\infty} p(x|\mu,\sigma^2)dx &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{t^2}{2\sigma^2}\right) dt \\
&= \frac{I}{\sqrt{2\pi\sigma^2}} \\
&= \frac{\sqrt{2\pi\sigma^2}}{\sqrt{2\pi\sigma^2}} = 1
\end{aligned}
$$

1.2 b.

From the definition of the Gaussian distribution, X has the probability density function:

$$
f_x(x) = \frac{1}{\sqrt{2\pi\sigma^2}}.exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)
$$

We have Expectation of Gaussian distribution:

$$
E(X) = \int\limits_{-\infty}^{\infty} x.f_x(x)dx
$$

So:

$E(X) = \int\limits_{-\infty}^{\infty} x.x.f_x(x)dx$

We substitute: $t = \frac{x-\mu}{\sqrt{2}\sigma}$ => $dt = \frac{dx}{\sqrt{2}\sigma}$ => $dx = \sqrt{2}$

$$
\begin{aligned}
=> E(X) &= \frac{\sqrt{2}\sigma}{\sigma\sqrt{2\pi}}. \int\limits_{-\infty}^{\infty} (\sqrt{2}\sigma t + \mu).exp(-t^2)dt \\
&= \frac{1}{\sqrt{\pi}}. \left(\sqrt{2}\sigma. \int\limits_{-\infty}^{\infty} t.exp(-t^2)dt + \mu \int\limits_{-\infty}^{\infty} exp(-t^2)dt \right)
\end{aligned}
$$

We have: $\int\limits_{-\infty}^{\infty} t.exp(-t^2)dt$ (1)

From (1): Set $u = e^{-x^2}$ => $du = -2xe^{-x^2}dx$ => $-\frac{1}{2}du = x.e^{-x^2}dx$

We have: $\int\limits_{-\infty}^{\infty} exp(-t^2)dt$ (2)

From (2): We have $\int\limits_{-\infty}^{\infty} exp(-t^2)dt = \sqrt{\pi}$

$$
\begin{aligned}
=> E(X) &= \frac{1}{\pi} \left(\sqrt{2}\sigma. \int\limits_{-\infty}^{\infty} \frac{-1}{2}du + \mu\sqrt{\pi} \right) \\
&= \frac{1}{\sqrt{\pi}}. \left(\sqrt{2}\sigma. \left[\frac{-1}{2}.exp(-t^2) \right]_{-\infty}^{\infty} + \mu\sqrt{pi} \right) \\
&= \frac{1}{\sqrt{\pi}}.\mu\sqrt{\pi} \\
&= \mu
\end{aligned}
$$

1.3 c.

We have the probability density function of Gaussian distribution is:

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

From the Varience as Expectation of Square minus Square of expactation:

$$var(X) = \int\limits_{-\infty}^{\infty} x^2 f_X(x)\mathrm{d}x - (E(X))^2$$

So:

$$
\begin{aligned}
var(X) &= \frac{1}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} x^2 \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\mathrm{d}x - \mu^2 \\
Set: & \\
t = \frac{x-\mu}{\sqrt{2}\sigma}: & \\
&= \frac{\sqrt{2}\sigma}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} (\sqrt{2}\sigma t + \mu)^2 \exp(-t^2)\mathrm{d}t - \mu^2 \\
&= \frac{1}{\sqrt{\pi}}(2\sigma^2 \int\limits_{-\infty}^{\infty} t^2\exp(-t^2)\mathrm{d}t + 2\sqrt{2}\sigma\mu\int\limits_{-\infty}^{\infty} t\exp(-t^2)\mathrm{d}t + \mu^2\int\limits_{-\infty}^{\infty}\exp(-t^2)\mathrm{d}t) - \mu^2 \\
&= \frac{1}{\sqrt{\pi}}(2\sigma^2 \int\limits_{-\infty}^{\infty} t^2\exp(-t^2)\mathrm{d}t + 2\sqrt{2}\sigma\mu[-\frac{1}{2}\exp(-t^2)] + \mu^2\sqrt{\pi}) - \mu^2 \\
&= \frac{1}{\sqrt{\pi}}(2\sigma^2 \int\limits_{-\infty}^{\infty} t^2\exp(-t^2)\mathrm{d}t + 2\sqrt{2}\sigma\mu\cdot 0) + \mu^2 - \mu^2 \\
&= \frac{2\sigma^2}{\sqrt{\pi}}\int\limits_{-\infty}^{\infty} t^2\exp(-t^2)\mathrm{d}t \\
&= \frac{2\sigma^2}{\sqrt{\pi}}([-\frac{t}{2}\exp -t^2] + \frac{1}{2}\int\limits_{-\infty}^{\infty}\exp(-t^2)\mathrm{d}t) \\
&= \frac{2\sigma^2}{\sqrt{\pi}}\cdot\frac{1}{2}\int\limits_{-\infty}^{\infty}\exp(-t^2)\mathrm{d}t \\
&= \frac{2\sigma^2\sqrt{\pi}}{2\sqrt{\pi}} \\
&= \sigma^2
\end{aligned}
$$

1.4 d.

Variance of Gaussian distribution is sigma^2($variance$)

We have

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

with $y_i = u_i^T(x-\mu)$ We also have $|\Sigma|^{\frac{1}{2}} = \prod_{i=1}^{D}\lambda_i^{\frac{1}{2}}$. For a D-dimensional vector x, the multivariate Gaussian distribution takes the form

$$
p(x \mid \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{D}{2}}|\Sigma|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)
$$

We replace $y_i = u_i^T(x-\mu)$ into the equation, we have

$$\begin{aligned} p(y) &= \frac{1}{(2\pi)^{\frac{D}{2}}\left(\prod_{i=1}^{D}\lambda_i\right)^{\frac{1}{2}}}\exp\left(-\frac{1}{2}\sum_{i=1}^{D}\frac{y_i^2}{\lambda_i}\right) \\ &= \frac{1}{(2\pi)^{\frac{D}{2}}\left(\prod_{i=1}^{D}\lambda_i\right)^{\frac{1}{2}}}\prod_{i=1}^{D}\exp\left(-\frac{1}{2}\frac{y_i^2}{\lambda_i}\right) \\ &= \prod_{j=1}^{D}\frac{1}{(2\pi\lambda_i)^{\frac{1}{2}}}\exp\left(-\frac{y_j^2}{2\lambda_j}\right) \\ \implies \int_{-\infty}^{\infty} p(y)dy &= \prod_{j=1}^{D}\int_{-\infty}^{\infty}\frac{1}{(2\pi\lambda_i)^{\frac{1}{2}}}\exp\left(-\frac{y_j^2}{2\lambda_j}\right)dy_j \\ &= 1 \end{aligned}$$

2 Problem 2

2.1 a.

Suppose x is a D-dimensional vector with Gaussian distribution N(x|,) and that we partition x into two disjoint subsets xa and xb
x = $\begin{pmatrix} x_a \\ x_b \end{pmatrix}$
We also define corresponding partitions of the mean vector μ given by
$\mu = \begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix}$
and of the covariance matrix Σ given by
$\Sigma = \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix}$ ⇒A = $\Sigma^{-1} = \begin{pmatrix} A_{aa} & A_{ab} \\ A_{ba} & A_{bb} \end{pmatrix}$
Σ is symmetric so $\Sigma_{aa} and \Sigma_{bb}$ are symmetric while $\Sigma_{ab} = \Sigma_{bb}^T$
We are looking for conditional distribution $p(x_a|x_b)$
We have

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

Compare with Gaussian disitrubtion
$\Delta^2 = \frac{-1}{2}x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu$

$\Rightarrow \Sigma_{a|b} = A_{aa}^{-1}$
$\Rightarrow \mu_{a|b} = \Sigma_{a|b}(A_{aa}\mu_a - A_{ab}(x_b - \mu_b)) = \mu_a - A_{aa}^{-1}A_{ab}(x_b - \mu_b)$

By using Schur complement,
$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} M & -MBD^{-1} \\ -D^{-1}CMD^{-1} & D^{-1}CMBD^{-1} \end{pmatrix}$

$\Rightarrow \mathrm{A}_{aa} = (\Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba})^{-1}$
$\Rightarrow \mathrm{A}_{ba} = -(\Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba})^{-1}\Sigma_{ab}\Sigma_{bb}^{-1}$

As the result:

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

$\Rightarrow \mathrm{p}(\mathrm{x}_a|x_b) = N(x_{a|b}|\mu_{a|b}, \Sigma_{a|b})$

2.2 b.

The marginal distribution given by:
$p(x_a) = \int p(x_a, x_b)dx_b$
We need to integrate out x_b by looking the quadratic form related to x_b
$-\frac{1}{2}x_b^T A_{bb}x_b + x_b^T m = -\frac{1}{2}(x_b - A_{bb}^{-1}m)^T A_{bb}(x_b - A_{bb}^{-1}m) + \frac{1}{2}m^T A_{bb}^{-1}m$
with $m = A_{bb}\mu_b - A_{ba}(x_a - \mu_a)$
We can integrate over unnormalized Gaussian:
$\int exp[-\frac{1}{2}(x_b - A_{bb}^{-1}m)^T A_{bb}(x_b - A_{bb}^{-1}m)]dx_b$
The remaining term
$-\frac{1}{2}x_a^T(A_{aa} - A_{ab}A_{bb}^{-1}A_{ba})x_a + x_a^T(A_{aa} - A_{ab}A_{bb}^{-1}A_{ba})^{-1}\mu_a + \text{const}$
Similarly, we have

$$\begin{aligned} E[x_a] &= \mu_a \\ cov[x_a] &= \Sigma_{aa} \end{aligned}$$

$\Rightarrow \mathrm{p}(\mathrm{x}_a) = N(x_a|\mu_a, \Sigma_{aa})$