Homework 3 - Gaussian

Thiều Ngọc Mai

August 2023

Problem 1

a. Prove that the Univariate Gaussian PDF is normalized.

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

When Univariate Gaussian PDF is normalize, we need to show that:

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

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

Assume $\mu = 0$, have

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

$$\Rightarrow I^2 = \int\limits_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx \int\limits_{-\infty}^{\infty} \exp\left(-\frac{y^2}{2\sigma^2}\right) dy$$

$$= \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2\sigma^2}\right) \exp\left(-\frac{y^2}{2\sigma^2}\right) dxdy$$

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

Let $x = r\cos\theta$, $y = r\sin\theta$, then $x^2 + y^2 = r^2(\cos^2\theta + \sin^2\theta) = r^2$:

$$
\begin{aligned}
I^2 &= \int_0^{2\pi}\int_0^{\infty} \exp\left(-\frac{1}{2\sigma^2}r^2\right) dr d\theta \\
&= \int_0^{2\pi} -\sigma^2 \left[\exp\left(-\frac{1}{2\sigma^2}r^2\right)\right]_0^{\infty} d\theta \\
&= \int_0^{2\pi} \sigma^2 d\theta \\
&= \left[\sigma^2\theta\right]_0^{2\pi} \\
&= 2\pi\sigma^2 \\
\Rightarrow I &= \sqrt{2\pi\sigma^2}
\end{aligned}
$$

From there, normalize of Univariate Gaussian PDF is proved.

b. A random variable X follows Gaussian distribution (notation: $X \sim \mathcal{N}(\mu, \sigma^2)$). Prove that the expected value of X is μ and the standard deviation of X is σ.

- Have:

$$
\begin{aligned}
E(X) &= \int_X x f_X(x) dx \\
&= \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right) dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right) dx
\end{aligned}
$$

Let $z = x - \mu \Rightarrow x = z + \mu$, we have:

$$E(X) = \frac{1}{\sqrt{2\pi\sigma^2}} \int\limits_{-\infty}^{\infty} (z+\mu) \exp\left(-\frac{1}{2\sigma^2} z^2\right) d(z+\mu)$$
$$= \frac{1}{\sqrt{2\pi\sigma^2}} \left(\int\limits_{-\infty}^{\infty} z \exp\left(-\frac{1}{2\sigma^2} z^2\right) dz + \mu \int\limits_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2} z^2\right) dz \right)$$

The general antiderivatives are:

$$\int x \exp(-ax^2) dx = -\frac{1}{2a} \exp(-ax^2)$$
$$\int \exp(-ax^2) dx = -\frac{1}{2}\sqrt{\frac{\pi}{a}} \text{erf}(\sqrt{a}x)$$

Using this, we have:

$$E(X) = \frac{1}{\sqrt{2\pi\sigma^2}} \left(\left[-\sigma^2 \exp\left(-\frac{1}{2\sigma^2} z^2\right) \right]_{-\infty}^{\infty} + \mu \left[\sqrt{\frac{\pi}{2}} \sigma \text{erf}\left(\frac{1}{\sqrt{2}\sigma} z\right) \right]_{-\infty}^{\infty} \right)$$
$$= \frac{1}{\sqrt{2\pi\sigma^2}} \left(0 + 2\mu \sqrt{\frac{\pi}{2}\sigma} \right)$$
$$= \mu$$

- Have:

$$Var(X) = \int\limits_{\mathbb{R}} (x - E(X))^2 f_X dx$$
$$= \int\limits_{-\infty}^{\infty} (x-\mu)^2 \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx$$

Let $z = x - \mu$, we have:

$$Var(X) = \int\limits_{-\infty}^{\infty} z^2 \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{z^2}{2\sigma^2}\right) dz$$

Let $z = \sqrt{2}\sigma x$:

$$
\begin{aligned}
Var(X) &= \int\limits_{-\infty}^{\infty} (\sqrt{2}\sigma x)^2 \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\sqrt{2}\sigma x)^2}{2\sigma^2}\right) d(\sqrt{2}\sigma x) \\
&= \frac{1}{\sigma\sqrt{2\pi}} \cdot 2\sigma^2 \cdot \sqrt{2}\sigma \int\limits_{-\infty}^{\infty} x^2 \exp(-x^2) dx \\
&= \frac{2\sigma^2}{\sqrt{\pi}} \int\limits_{-\infty}^{\infty} x^2 \exp(-x^2) dx \\
&= \frac{4\sigma^2}{\sqrt{\pi}} \int\limits_{0}^{\infty} x^2 \exp(-x^2) dx
\end{aligned}
$$

Define $t = x^2 \Rightarrow x = \sqrt{t} \Rightarrow dx = (2\sqrt{t})^{-1} dt$. Subtituting:

$$
\begin{aligned}
Var(X) &= \frac{4\sigma^2}{\sqrt{\pi}} \int\limits_{0}^{\infty} t \exp(-t)(2\sqrt{t})^{-1} dt \\
&= \frac{4\sigma^2}{\sqrt{\pi}} \frac{1}{2} \int\limits_{0}^{\infty} t^{\frac{1}{2}} \exp(-t) dt
\end{aligned}
$$

Definition of the Γ function:

$$
\Gamma(X) = \int\limits_{0}^{\infty} x^{(k-1)} \exp(-x) dx
$$

Thus:

$$
Var(X) = \frac{2\sigma^2}{\sqrt{\pi}} \cdot \Gamma\left(\frac{3}{2}\right) = \frac{2\sigma^2}{\sqrt{\pi}} \cdot \frac{\sqrt{\pi}}{2} = \sigma^2
$$

Therefore, $\text{std}(X) = \sigma$

Problem 2

a. Prove that the Multivariate Gaussian PDF is normalized.
We need to prove:

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{(2\pi)^n|\Sigma|}} \cdot \exp\left[-\frac{1}{2}(X-\mu)^T\Sigma^{-1}(X-\mu)\right] dX = 1$$

$$\leftrightarrow I = \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}(X-\mu)^T\Sigma^{-1}(X-\mu)\right] dX = \sqrt{(2\pi)^n|\Sigma|}$$

Let $y = X - \mu$:

$$I = \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}y^T\Sigma^{-1}y\right] dy$$

Since Σ is symmetric and positive definite, then: $\Sigma = ESE^T$.

Where $E = [E_1, E_2, \ldots, E_n]$ is an orthogonal matrix of eigenvectors and $S = diag[\lambda_1, \lambda_2, \ldots, \lambda_n]$ is diagonal matrix of eigenvalues.

Then, $\Sigma^{-1} = ES^{-1}E^T$
Have:

$$I = \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}y^TUS^{-1}U^Ty\right] dX$$

$$= \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}(U^Ty)^TS^{-1}(U^Ty)\right] dX$$

Let $z = U^T y$:

$$I = \int\limits_{-\infty}^{\infty} \exp\left[-\frac{1}{2} z^T S^{-1} z\right] dz$$
$$= \int\limits_{-\infty}^{\infty} \exp\left[-\frac{z_1^2}{2\lambda_1} - \frac{z_2^2}{2\lambda_2} - \cdots - \frac{z_n^2}{2\lambda_n}\right] dz$$
$$= \int\limits_{-\infty}^{\infty} \prod_{i=1}^{n} \exp\left[-\frac{z_i^2}{2\lambda_i}\right] dz$$
$$= \sqrt{2\pi\lambda_1} \cdot \sqrt{2\pi\lambda_2} \cdot \cdots \sqrt{2\pi\lambda_n}$$
$$= \sqrt{(2\pi)^n \lambda_1 \lambda_2 \ldots \lambda_n}$$
$$= \sqrt{(2\pi)^n |S|}$$

Have: $|\Sigma| = |ESE^T| = |E||S||E^T| = |S|$, then $I = \sqrt{(2\pi)^n |\Sigma|}$ (proved)

b. Find the formula of marginal distribution in Multivariate Gaussian distribution.

$$p(x_a) = \int p(x_a, x_b) dx_b$$

We need to integrate out x_b by looking the quadratic form related to x_b

$$\Delta^2 = -\frac{1}{2}(x - \mu)^T A (x - \mu)$$
$$= -\frac{1}{2} x_b^T A_{bb} x_b + x_b^T m + const \quad (with\, m = A_{bb}\mu_b - A_{ba}(x_a - \mu_a))$$
$$= -\frac{1}{2}(x_b - A_{bb}^{-1} m)^T A_{bb} (x_b - A_{bb}^{-1} m) + \frac{1}{2} m^T A_{bb}^{-1} m$$

We can integrate over unnormalized Gaussian

$$\int exp\left\{ -\frac{1}{2}(x_b - A_{bb}^{-1} m)^T A_{bb} (x_b - A_{bb}^{-1} m)\right\} dx_b$$

The remaining term

$$-\frac{1}{2} x_a^T (A_{aa} - A_{ab} A_{bb}^{-1} A_{ba}) x_a + x_a^T (A_{aa} - A_{ab} A_{bb}^{-1} A_{ba})^{-1} \mu_a + const$$

Similarly we have

$$\mathbb{E}[x_a] = \mu_a$$
$$cov[x_a] = \Sigma_{aa}$$
$$\Rightarrow p(x_a) = \mathcal{N}(x_a|\mu_a, \Sigma_{aa})$$

c. Find the formula of conditionnal distribution in Multivariate Gaussian distribution.
Assume that

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{1}$$

where x_1 is $n_1 \times 1$ vector and x_2 is $n_2 \times 1$ vector and x is an $n_1 + n_2 = n \times 1$ vector.
The joint distribution of x_1 and x_2 is:

$$x_1, x_2 \sim \mathcal{N}(\mu, \Sigma) \tag{2}$$

The marginal distribution of x_2 is:

$$x_2 \sim \mathcal{N}(\mu_2, \Sigma_{22}) \tag{3}$$

Apply (1), (2), (3) to the law of conditional probability, it holds that:

$$\begin{aligned}
p(x_1|x_2) &= \frac{\mathcal{N}(x;\mu,\Sigma)}{\mathcal{N}(x_2;\mu_2,\Sigma_22)} \\
&= \frac{1/\sqrt{(2\pi)^n|\Sigma|} \cdot \exp\left[-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right]}{1/\sqrt{(2\pi)^{\frac{n}{2}}_2|\Sigma_{22}|} \cdot \exp\left[-\frac{1}{2}(x_2-\mu_2)^T\Sigma_{22}^{-1}(x_2-\mu_2)\right]} \\
&= \frac{1}{\sqrt{(2\pi)^{n-n_2}}} \cdot \sqrt{\frac{|\Sigma_{22}|}{|\Sigma|}} \cdot \exp\left[-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu) + \frac{1}{2}(x_2-\mu_2)^T\Sigma_{22}^{-1}(x_2-\mu_2)\right] \\
&= \frac{1}{\sqrt{(2\pi)^{n-n_2}}} \cdot \sqrt{\frac{|\Sigma_{22}|}{|\Sigma|}} \cdot \exp\left[-\frac{1}{2}\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}\right)^T \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}\right)\right. \\
&+ \frac{1}{2}(x_2-\mu_2)^T\Sigma_{22}^{-1}(x_2-\mu_2)
\end{aligned}$$

Apply the formula for inverse of a block matrix:

$$\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} \left(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)^{-1} & -\left(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)^{-1}\Sigma_{12}\Sigma_{22}^{-1} \\ -\Sigma_{22}^{-1}\Sigma_{21}\left(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)^{-1} & \Sigma_{22}^{-1} + \Sigma_{22}^{-1}\Sigma_{21}\left(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)^{-1}\Sigma_{12}\Sigma_{22}^{-1} \end{bmatrix}$$

Then,

$$\begin{aligned}
p(x_1|x_2) =& \frac{1}{\sqrt{(2\pi)^{n-n_2}}} \cdot \sqrt{\frac{|\Sigma_{22}|}{|\Sigma|}} \cdot \\
& \exp\Bigg[-\frac{1}{2}\Big((x_1-\mu_1)^{\mathrm{T}}(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1}(x_1-\mu_1) \; - \\
& \qquad 2(x_1-\mu_1)^{\mathrm{T}}(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1}\Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu_2) + \\
& \qquad (x_2-\mu_2)^{\mathrm{T}}\left[\Sigma_{22}^{-1}+\Sigma_{22}^{-1}\Sigma_{21}(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1}\Sigma_{12}\Sigma_{22}^{-1}\right](x_2-\mu_2)\Big) \\
& \quad + \frac{1}{2}\Big((x_2-\mu_2)^{\mathrm{T}}\Sigma_{22}^{-1}(x_2-\mu_2)\Big)\Bigg] \\
=& \frac{1}{\sqrt{(2\pi)^{n-n_2}}} \cdot \sqrt{\frac{|\Sigma_{22}|}{|\Sigma|}} \cdot \\
& \exp\Bigg[-\frac{1}{2}\Big((x_1-\mu_1)^{\mathrm{T}}(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1}(x_1-\mu_1) \; - \\
& \qquad 2(x_1-\mu_1)^{\mathrm{T}}(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1}\Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu_2) + \\
& \qquad (x_2-\mu_2)^{\mathrm{T}}\Sigma_{22}^{-1}\Sigma_{21}(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1}\Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu_2)\Big)\Bigg] \; . \\
=& \frac{1}{\sqrt{(2\pi)^{n-n_2}}} \cdot \sqrt{\frac{|\Sigma_{22}|}{|\Sigma|}} \cdot \exp\Bigg[-\frac{1}{2} \cdot \\
& \left[(x_1-\mu_1)-\Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu_2)\right]^{\mathrm{T}}(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1}\big[(x_1-\mu_1)-\Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu_2) \\
=& \frac{1}{\sqrt{(2\pi)^{n-n_2}}} \cdot \sqrt{\frac{|\Sigma_{22}|}{|\Sigma|}} \cdot \exp\Bigg[-\frac{1}{2} \cdot \\
& \left[x_1-\left(\mu_1+\Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu_2)\right)\right]^{\mathrm{T}}(\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1}\big[x_1-\big(\mu_1+\Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu_2)
\end{aligned}$$

Since Σ is covariance matrix, $\Sigma_{21} = \Sigma_{12}^T$
Determinant of block matrix is:

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |D| \cdot |A - BD^{-1}C|$$

Thus,

$$\begin{vmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{vmatrix} = |\Sigma_{22}| \cdot |\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}|$$

With $n - n_2 = n_1$:

$$\begin{aligned} p(x_1|x_2) =& \frac{1}{\sqrt{(2\pi)^{n_1}|\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}|}} \cdot \exp\left[-\frac{1}{2}\cdot\right. \\ & \left[x_1 - \left(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2)\right)\right]^{\mathrm{T}} (\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1} \\ & \left[x_1 - \left(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2)\right)\right] \end{aligned}$$

Which is the PDF of multivariate Normal distribution. Therefore,

$$p(x_1|x_2) = \mathcal{N}(x_1; \mu_{1|2}, \Sigma_{1|2})$$