1 Prove that

1.1 Gaussian distribution is normalized.

Let:

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

Squaring the above expression:

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

We utilize:

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

Based on the fact that:

$$\int_0^{+\infty} \exp\left(-\frac{r^2}{2\sigma^2}\right) rdr = -\,\sigma^2 \exp\left(-\frac{r^2}{2\sigma^2}\right)\Big|_0^{+\infty} = -\sigma^2(0-1) = \sigma^2$$

Therefore:

$$I^2 = \int_0^{2\pi} \sigma^2 d\theta = 2\pi\sigma^2, \quad => I = \sqrt{2\pi}\sigma$$

We can prove normalized by:

$$\begin{aligned}
\int_{-\infty}^{+\infty} N\left(x \mid \mu, \sigma^2\right) dx &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\} dx \\
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}y^2\right\} dy \quad (y = x-\mu) \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} \exp\left\{-\frac{1}{2\sigma^2}y^2\right\} dy \\
&= 1
\end{aligned}$$

1.2 Expectation of Gaussian distribution is mu (mean)

We have the Expectation formula of a continuous random variable:

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

And the Gaussian distribution formula:

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

Therefore:

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

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

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

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

From Variance as Expectation of Square minus Square of Expectation:

$$\mathrm{var}(X) = \int_{-\infty}^{\infty}(x^2-\mu^2)p_X(x|\mu,\sigma^2)\mathrm{d}x$$

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

1.4 Multivariate Gaussian distribution is normalized

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

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

2 Calculate

2.1 The conditional of Gaussian distribution

Suppose x is a D-dimensional vector with Gaussian distribution $\mathcal{N}(x \mid \mu, \Sigma)$ and that we partition x into two disjoint subsets x_a and x_b

$$x = \begin{pmatrix} x_a \\ x_b \end{pmatrix}$$

We also define corresponding partitions of the mean vector μ given by

$$\mu = \begin{pmatrix} \mu_{\mathrm{a}} \\ \mu_{\mathrm{b}} \end{pmatrix}$$

and of the covariance matrix Σ given by

$$\Sigma = \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix} \Rightarrow A = \Sigma^{-1} = \begin{pmatrix} A_{aa} & A_{ab} \\ A_{ba} & A_{bb} \end{pmatrix}$$

Σ is symmetric so Σ_{aa} and Σ_{bb} are symmetric while $\Sigma_{ab} = \Sigma_{ba}^T$ We are looking for conditional distribution $p\left(x_a \mid x_b\right)$ We have

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

It is quadratic form of x_a hence conditional distribution $\mathrm{p}\left(x_a \mid x_b\right)$ will be Gaussian, because this distribution is characterized by its mean and its variance. Compare with Gaussian disitrubtion

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

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

By using Schur complement,
$\Rightarrow A_{aa} = \left(\Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba}\right)^{-1}$ $A_{ab} = -\left(\Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba}\right)^{-1}\Sigma_{ab}\Sigma_{bb}^{-1}$
As a result

$$
\begin{aligned}
&\mu_{a|b} = \mu_a + \Sigma_{ab}\Sigma_{bb}^{-1}\left(x_b-\mu_b\right)\\
&\Sigma_{a|b} = \Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba}\\
&\Rightarrow p\left(x_a \mid x_b\right) = N\left(x_{a|b} \mid \mu_{a|b}, \Sigma_{a|b}\right)
\end{aligned}
$$

2.2 The marginal of Gaussian distribution

The margianl distribution given by

$$p\left(x_a\right)=\int p\left(x_a, x_b\right) d x_b$$

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

$$-\frac{1}{2} x_b^T A_{b b} x_b+x_b^T m=-\frac{1}{2}\left(x_b-A_{b b}^{-1} m\right)^T A_{b b}\left(x_b-A_{b b}^{-1} m\right)+\frac{1}{2} m^T A_{b b}^{-1} m$$

with $m=A_{b b} \mu_b-A_{b a}\left(x_a-\mu_a\right)$ We can integrate over unnormalized Gaussian

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

The remaining term

$$-\frac{1}{2} x_a^T\left(A_{a a}-A_{a b} A_{b b}^{-1} A_{b a}\right) x_a+x_a^T\left(A_{a a}-A_{a b} A_{b b}^{-1} A_{b a}\right)^{-1} \mu_a+\text { const }$$

Similarly, we have

$$\begin{aligned}
& \mathbb{E}\left[x_a\right]=\mu_a \\
& \operatorname{cov}\left[x_a\right]=\Sigma_{a a} \\
& \Rightarrow p\left(x_a\right)=N\left(x_a \mid \mu_a, \Sigma_{a a}\right)
\end{aligned}$$