Problem 1
Multivariate Gaussian Distribution:
Proof that Multivariate Gaussian Distribution is normalize:

Solution
First, we have the PDF of the Gaussian Distribution is:

$$p(x \mid \mu, \sigma^2) = \frac{1}{(2\pi)^{\dfrac{D}{2}} \mid \Sigma \mid^{\dfrac{1}{2}}} \times e^{\dfrac{-1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)}$$

The Gaussian Distribution is normalize

$$\Longleftrightarrow \int_{-\infty}^{+\infty} p(x \mid \mu, \sigma^2) = 1$$

Where μ is a D-dimensional mean vector
Σ is a D x D covariance matrix
$\mid \Sigma \mid$ denotes the determinant of Σ
Set

$$\begin{aligned}\Delta^2 &= \frac{-1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\\ &= \frac{-1}{2} x^T \Sigma^{-1} x + x^T \Sigma^{-1} \mu + constant\end{aligned}$$

Consider eigenvalues and eigenvvectors of Σ we have:

$$\Sigma u_i = \lambda_i u_i, i = 1, ..., D$$

Because Σ is a real, symmetric matrix
$\rightarrow$ its eigenvalues will be real and its eigenvectors form an orthnormal set.
Proof:
1. its eigenvalues will be real
Example:

$$\begin{pmatrix} \sigma_1^2 & \sigma_{1,2} \\ \sigma_{2,1} & \sigma_2^2 \end{pmatrix}$$

$\Longrightarrow$ The equation to find the eigenvalues is :

$$(\sigma_1^2 - \lambda) \times (\sigma_2^2 - \lambda) - (\sigma_{1,2})^2 = 0$$

$$\Longleftrightarrow (\sigma_1^2 - \lambda) \times (\sigma_2^2 - \lambda) = (\sigma_{1,2})^2$$

$\Longrightarrow$ With $\lambda = \lambda_1$:

$$\begin{pmatrix} \sigma_1^2 - \lambda_1 & cov(\sigma_{1,2}) \\ cov(\sigma_{1,2}) & \sigma_2^2 - \lambda_1 \end{pmatrix}$$

$$(\sigma_1^2 - \lambda_1)x_1 + (\sigma_{1,2})x_2 = 0 \tag{1}$$

$$(\sigma_{1,2})x_1 + (\sigma_2^2 - \lambda_1)x_2 = 0 \tag{2}$$

From (1) we have:

$$\begin{aligned} x_1 &= \frac{-y \times cov(\sigma_{1,2})}{\sigma_1^2 - \lambda_1} \\ x_2 &= x_2 \end{aligned}$$

So the eigenvector in this case is:

$$\begin{pmatrix} \dfrac{-cov(\sigma_{1,2})}{\sigma_1^2 - \lambda_1} \\ 1 \end{pmatrix}$$

With $\lambda = \lambda_2$:

$$\begin{pmatrix} \sigma_1^2 - \lambda_2 & cov(\sigma_{1,2}) \\ cov(\sigma_{1,2}) & \sigma_2^2 - \lambda_2 \end{pmatrix}$$

$$(\sigma_1^2 - \lambda_2)x_1 + (\sigma_{1,2})x_2 = 0 \tag{3}$$

$$(\sigma_{1,2})x_1 + (\sigma_2^2 - \lambda_2)x_2 = 0 \tag{4}$$

From (3) we have:

$$x_1 = \frac{-y \times cov(\sigma_{1,2})}{\sigma_1^2 - \lambda_2}$$

$$\longrightarrow x_2 = x_2$$

So the eigenvector in this case is:

$$\begin{pmatrix} \dfrac{-cov(\sigma_{1,2})}{\sigma_1^2 - \lambda_2} \\ 1 \end{pmatrix}$$

And:

$$\begin{pmatrix} \dfrac{-cov(\sigma_{1,2})}{\sigma_1^2 - \lambda_2} \\ 1 \end{pmatrix}^T \times \begin{pmatrix} \dfrac{-cov(\sigma_{1,2})}{\sigma_1^2 - \lambda_2} \\ 1 \end{pmatrix} = 1$$

So its eigenvectors form an orthnormal set.

$$\Sigma = \Sigma_{i=1}^{D} \lambda_i u_i (u_i)^T \longrightarrow \Sigma^{-1} = \Sigma_{i=1}^{D} \frac{1}{\lambda i} u_i u_i^T$$

So that:

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

$$= \sum_{i=1}^{D} \frac{1}{\lambda_i}(x-\mu)^T u_i u_i^T (x-\mu)$$

Let:

$$y_i = u_i^T (x-\mu)$$

$$\longrightarrow \Delta^2 = \sum_{i=1}^{D} \frac{y_i^2}{\lambda_i}$$

$$\mid \Sigma \mid^{1/2} = \prod_{j=1}^{D} \lambda_j^{1/2}$$

Now,we have:

$$p(x \mid \mu, \sigma^2) = p(y) = \prod_{j=1}^{D} \frac{1}{(2\pi\lambda_j)^{1/2}} e^{\dfrac{-(y_j)^2}{2\lambda_j}}$$

$$\Longleftrightarrow \int_{-\infty}^{+\infty} p(y)dy = \prod_{j=1}^{D} \int_{-\infty}^{+\infty} \frac{1}{(2\pi\lambda_j)^{1/2}} e^{\dfrac{-(y_j)^2}{2\lambda_j}}$$

But in the last homework we have proofed that:

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{\dfrac{-y^2}{2}}\,dy = 1$$

So:

$$\int_{-\infty}^{+\infty} \frac{1}{(2\pi\lambda_j)^{1/2}} e^{\dfrac{-(y_j)^2}{2\lambda_j}} = 1$$

$$\Longleftrightarrow \prod_{j=1}^{D}\int_{-\infty}^{+\infty} \frac{1}{(2\pi\lambda_j)^{1/2}} e^{\dfrac{-(y_j)^2}{2\lambda_j}} = 1$$

Problem 2
Calculate marginal normal distribution

Solution Let:

$$\begin{aligned}\Delta^2 &= -\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\\ &= -\frac{1}{2}(x^T-\mu^T)\Sigma^{-1}(x-\mu)\\ &= -\frac{1}{2}x^T\Sigma^{-1}x + \frac{1}{2}\left(x^T\Sigma^{-1}\mu + \mu^T\Sigma^{-1}x\right) - \frac{1}{2}\mu^T\Sigma^{-1}\mu \qquad (5)\end{aligned}$$

Where:

(∗) x is a D × 1 matrix → x^T is a 1 × D matrix
(∗) μ is a D × 1 matrix
(∗) Σ^{-1} is a D × D covariance matrix which positive definite and symmetric

So, the dimension of

$$\begin{aligned}x^T\Sigma^{-1}\mu &= 1\times D\otimes D\times D\otimes D\times 1\\ &= 1\times 1\end{aligned}$$

→ $x^T\Sigma^{-1}\mu$ equal a numeric value

$$\begin{aligned}\Rightarrow x^T\Sigma^{-1}\mu &= (x^T\Sigma^{-1}\mu)^T\\ &= \mu^T\Sigma^{-1}x\end{aligned}$$

Therefore, equation (2) can be rewritten as

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

Suppose x is a D-dimensional vector with Gaussian distribution $\mathcal{N}(x|\mu,\Sigma)$ 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

$$\mu = \begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix}$$

and of the covariance matrix Σ given by

$$\Sigma = \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix}$$

Let

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

We are looking for conditional distribution $p(x_a|x_b)$. We have:

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

$$= -\frac{1}{2}(x-\mu)^T A(x-\mu)$$

$$= -\frac{1}{2}\begin{bmatrix} x_a-\mu_a \\ x_b-\mu_b \end{bmatrix}^T \begin{bmatrix} A_{aa} & A_{ab} \\ A_{ba} & A_{bb} \end{bmatrix} \begin{bmatrix} x_a-\mu_a \\ x_b-\mu_b \end{bmatrix}$$

$$= -\frac{1}{2}(x_a-\mu_a)^T A_{aa}(x_a-\mu_a) - \frac{1}{2}(x_a-\mu_a)^T A_{ab}(x_b-\mu_b)$$

$$-\frac{1}{2}(x_b-\mu_b)^T A_{ba}(x_a-\mu_a) - \frac{1}{2}(x_b-\mu_b)^T A_{bb}(x_b-\mu_b)$$

$$= -\frac{1}{2}x_a^T A_{aa}x_a + x_a^T[A_{aa}\mu_a - A_{ab}(x_b-\mu_b)] + const$$

Compare with Gaussian distribution

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

$$\rightarrow \begin{cases} -\frac{1}{2}x^T\Sigma^{-1}x = -\frac{1}{2}x_a^T A_{aa}x_a \\ x^T\Sigma^{-1}\mu = x_a^T[A_{aa}\mu_a - A_{ab}(x_b-\mu_b)] \end{cases}$$

$$\Leftrightarrow \begin{cases} \Sigma^{-1} = A_{aa} \\ \Sigma^{-1}\mu = A_{aa}\mu_a - A_{ab}(x_b-\mu_b) \end{cases}$$

$$\Leftrightarrow \begin{cases} \Sigma^{-1} = A_{aa} \\ \mu = \mu_a - A_{aa}^{-1}A_{ab}(x_b-\mu_b) \end{cases}$$

By using Schur complement,

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} M & -MBD^{-1} \\ -D^{-1}CMD^{-1} & D^{-1}CMBD^{-1} \end{pmatrix}, with\ M = (A - BD^{-1}C)^{-1}$$

$$\Rightarrow \begin{cases} A_{aa} = (\Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba})^{-1} \\ A_{ab} = -(\Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba})^{-1}\Sigma_{ab}\Sigma_{bb}^{-1} \end{cases}$$

As a result,

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

$$\Rightarrow p(x_a|x_b) = \mathcal{N}(x_{a|b}|\mu_{a|b}, \Sigma_{a|b})$$

Solution

Problem 3
Calculate conditional normal distribution

Solution 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

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

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

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

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

$$\mathbb{E}[x_a] = \mu_a$$

$$cov[x_a] = \Sigma_{aa}$$

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