Problem 1
Proof multivariate gaussian distribution normalization

Solution Multivariate Gaussian Distribution:

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

With:
μ is a D-dimensional mean vector
Σ is a D $\times$ D covariance matrix
$|\Sigma|$ denotes the determinant of Σ
We have:
+) Σ symmetric so eigenvalues $(\lambda_1, \lambda_2, ..., \lambda_d)$ real and eigenvector orthornomal $(u_1, u_2, .., u_d)$ ($||u_i|| = 1$ and $u_i u_j = 0$)
+) Eigendecomposition decomposes a matrix Σ

$$\Sigma = PDP^{-1} = PDP^T$$

*, P is a matrix of eigenvectors
*. D is a diagonal matrix of eigenvalues
Where

$$P = \begin{bmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_n \\ | & | & & | \end{bmatrix}$$

and

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \vdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$

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

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

So that:

$$\Delta^2 = (x-\mu)^T\Sigma^{-1}(x-\mu) = \sum_{i=1}^{D}\frac{1}{\lambda_i}(x-\mu)^T u_i u_i^T (x-\mu)$$

$$= \sum_{i=1}^{D}\frac{y_i^2}{\lambda_i} \quad with \quad y_i = u_i^T(x-\mu)$$

After that we have:

$$det(\Sigma) = det(P)det(D)det(P^{-1}) = det(D)$$

$$\rightarrow |\Sigma|^{\frac{1}{2}} = \prod_{j=1}^{D}\lambda_i^{\frac{1}{2}}$$

Finally we have:

$$p(y) = \prod_{j=1}^{D}(\frac{1}{2\pi\lambda_j})^{\frac{1}{2}}e^{-\frac{y_i^2}{2\lambda_i}}$$

$$\rightarrow \int_{-\infty}^{+\infty} p(y)dy = \prod_{j=1}^{D} \int_{-\infty}^{+\infty} (\frac{1}{2\pi\lambda_j})^{\frac{1}{2}} e^{-\frac{y_i^2}{2\lambda_i}} dy_j = 1$$

$$\rightarrow Multivariate \quad gaussian \quad distribution \quad normalization$$

Problem 2
Calculate conditional normal distribution.

Solution
Set:

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

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

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

We have Σ symmetric so Σ_{aa} and Σ_{bb} are symmetric while $\Sigma_{ab} = \Sigma_{ba}^T$ We have:

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

From (*) and (**):

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

By using Schur complement

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

with M = $(A - BD^{-1}C)^{-1}$

$$\rightarrow A_{aa} = (\Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba})^{-1}$$
$$A_{ab} = -(\Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba})^{-1}\Sigma_{ab}\Sigma_{bb}^{-1}$$

As a result:

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

Problem 3
Calculate marginal normal distribution

Solution The margianl 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

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

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

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

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

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