

Derivation of the conditional probabilities of multivariate Gaussian distribution \mathbf{x} : $p(\mathbf{x}) = p(\mathbf{x}_1, \mathbf{x}_2)$. Suppose $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2)$ is jointly Gaussian:

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}, \quad \Lambda = \Sigma^{-1} = \begin{pmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{pmatrix}$$

The marginal probabilities are:

$$\begin{aligned} p(\mathbf{x}_1) &= \mathcal{N}(\mathbf{x}_1 | \mu_1, \Sigma_{11}) \\ p(\mathbf{x}_2) &= \mathcal{N}(\mathbf{x}_2 | \mu_2, \Sigma_{22}) \end{aligned}$$

the conditional probabilities are:

$$\begin{aligned} p(\mathbf{x}_1 | \mathbf{x}_2) &= \mathcal{N}(\mathbf{x}_1 | \mu_{1|2}, \Sigma_{1|2}) \\ \mu_{1|2} &= \mu_1 + \Sigma_{12} \Sigma_{22}^{-1} (\mathbf{x}_2 - \mu_2) \\ &= \mu_1 - \Lambda_{11}^{-1} \Lambda_{12} (\mathbf{x}_2 - \mu_2) \\ &= \Sigma_{1|2} (\Lambda_{11} \mu_1 - \Lambda_{12} (\mathbf{x}_2 - \mu_2)) \\ \Sigma_{1|2} &= \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \\ &= \Lambda_{11}^{-1} \end{aligned} \tag{1}$$

Proof:

Part 1. derive $\mu_{1|2}$ and $\Sigma_{1|2}$

use the following equality to transform Σ :

$$M^{-1} = \begin{pmatrix} E & F \\ G & H \end{pmatrix}^{-1} = \begin{pmatrix} I & 0 \\ -H^{-1}G & I \end{pmatrix} \begin{pmatrix} (M/H)^{-1} & 0 \\ 0 & H^{-1} \end{pmatrix} \begin{pmatrix} I & -FH^{-1} \\ 0 & I \end{pmatrix}$$

where

$$M/H \equiv E - FH^{-1}G$$

By a simple substitution:

$$M \rightarrow \Sigma, \quad E \rightarrow \Sigma_{11}, \quad F \rightarrow \Sigma_{12}, \quad G \rightarrow \Sigma_{21}, \quad H \rightarrow \Sigma_{22}$$

$$\Sigma^{-1} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}^{-1} = \begin{pmatrix} I & 0 \\ -\Sigma_{22}^{-1}\Sigma_{21} & I \end{pmatrix} \begin{pmatrix} (\Sigma/\Sigma_{22})^{-1} & 0 \\ 0 & \Sigma_{22}^{-1} \end{pmatrix} \begin{pmatrix} I & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I \end{pmatrix}$$

Therefore, $p(\mathbf{x}) = p(\mathbf{x}_1, \mathbf{x}_2)$:

$$\begin{aligned} p(\mathbf{x}_1, \mathbf{x}_2) &\propto \exp \left\{ -\frac{1}{2} \begin{pmatrix} \mathbf{x}_1 - \mu_1 \\ \mathbf{x}_2 - \mu_2 \end{pmatrix}^T \begin{pmatrix} I & 0 \\ -\Sigma_{22}^{-1}\Sigma_{21} & I \end{pmatrix} \begin{pmatrix} (\Sigma/\Sigma_{22})^{-1} & 0 \\ 0 & \Sigma_{22}^{-1} \end{pmatrix} \begin{pmatrix} I & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 - \mu_1 \\ \mathbf{x}_2 - \mu_2 \end{pmatrix} \right\} \\ &= \exp \left\{ -\frac{1}{2} ((\mathbf{x}_1 - \mu_1)^T, (\mathbf{x}_2 - \mu_2)^T) \begin{pmatrix} I & 0 \\ -\Sigma_{22}^{-1}\Sigma_{21} & I \end{pmatrix} \begin{pmatrix} (\Sigma/\Sigma_{22})^{-1} & 0 \\ 0 & \Sigma_{22}^{-1} \end{pmatrix} \begin{pmatrix} (\mathbf{x}_1 - \mu_1) - \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \mu_2) \\ (\mathbf{x}_2 - \mu_2) \end{pmatrix} \right\} \\ &= \exp \left\{ -\frac{1}{2} ((\mathbf{x}_1 - \mu_1)^T - (\mathbf{x}_2 - \mu_2)^T \Sigma_{22}^{-1}\Sigma_{21}, (\mathbf{x}_2 - \mu_2)^T) \begin{pmatrix} (\Sigma/\Sigma_{22})^{-1}(\mathbf{x}_1 - \mu_1 - \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \mu_2)) \\ \Sigma_{22}^{-1}(\mathbf{x}_2 - \mu_2) \end{pmatrix} \right\} \\ &= \exp \left\{ -\frac{1}{2} (\underbrace{(\mathbf{x}_1 - \mu_1)^T - (\mathbf{x}_2 - \mu_2)^T \Sigma_{22}^{-1}\Sigma_{21}}_{(\mathbf{x}_1 - \mu_1)^T - [(\Sigma_{22}^{-1}\Sigma_{21})^T(\mathbf{x}_2 - \mu_2)]^T}, (\mathbf{x}_2 - \mu_2)^T) \begin{pmatrix} (\Sigma/\Sigma_{22})^{-1}(\mathbf{x}_1 - \mu_1 - \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \mu_2)) \\ \Sigma_{22}^{-1}(\mathbf{x}_2 - \mu_2) \end{pmatrix} \right\} \\ &= \exp \left\{ -\frac{1}{2} ((\mathbf{x}_1 - \mu_1 - \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \mu_2))^T, (\mathbf{x}_2 - \mu_2)^T) \begin{pmatrix} (\Sigma/\Sigma_{22})^{-1}(\mathbf{x}_1 - \mu_1 - \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \mu_2)) \\ \Sigma_{22}^{-1}(\mathbf{x}_2 - \mu_2) \end{pmatrix} \right\} \\ &= \exp \left[-\frac{1}{2} (\mathbf{x}_1 - \mu_1 - \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \mu_2))^T (\Sigma/\Sigma_{22})^{-1} (\mathbf{x}_1 - \mu_1 - \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \mu_2)) \right] \exp \left[-\frac{1}{2} (\mathbf{x}_2 - \mu_2)^T \Sigma_{22}^{-1} (\mathbf{x}_2 - \mu_2) \right] \\ &\propto \mathcal{N}(\mathbf{x}_1 | \mu_{1|2}, \Sigma_{1|2}) \times \mathcal{N}(\mathbf{x}_2 | \mu_2, \Sigma_{22}) \end{aligned}$$

where

$$\begin{aligned} \mu_{1|2} &= \mu_1 + \Sigma_{12} \Sigma_{22}^{-1} (\mathbf{x}_2 - \mu_2) \\ \Sigma_{1|2} &= \Sigma/\Sigma_{22} = \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \end{aligned}$$

Part 2. derive other forms of $\mu_{1|2}$ and $\Sigma_{1|2}$.

use the identity:

$$M^{-1} = \begin{pmatrix} E & F \\ G & H \end{pmatrix}^{-1} = \begin{pmatrix} (M/H)^{-1} & -(M/H)^{-1}FH^{-1} \\ -H^{-1}G(M/H)^{-1} & H^{-1} + H^{-1}G(M/H)^{-1}FH^{-1} \end{pmatrix}, \quad M/H \equiv E - FH^{-1}G$$

By the following substitution:

$$M \rightarrow \Sigma, \quad E \rightarrow \Sigma_{11}, \quad F \rightarrow \Sigma_{12}, \quad G \rightarrow \Sigma_{21}, \quad H \rightarrow \Sigma_{22}$$

$$\begin{aligned} \Sigma^{-1} &= \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}^{-1} = \begin{pmatrix} (\Sigma/\Sigma_{22})^{-1} & -(\Sigma/\Sigma_{22})^{-1}\Sigma_{12}\Sigma_{22}^{-1} \\ -\Sigma_{22}^{-1}\Sigma_{21}(\Sigma/\Sigma_{22})^{-1} & \Sigma_{22}^{-1} + \Sigma_{22}^{-1}\Sigma_{21}(\Sigma/\Sigma_{22})^{-1}\Sigma_{12}\Sigma_{22}^{-1} \end{pmatrix} \\ = \Lambda &= \begin{pmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{pmatrix} \end{aligned}$$

Therefore

$$\underline{\Lambda_{11}^{-1}\Lambda_{12}} = (\Sigma/\Sigma_{22})^{-1}[-(\Sigma/\Sigma_{22})^{-1}\Sigma_{12}\Sigma_{22}^{-1}] = \underline{\Sigma_{12}\Sigma_{22}^{-1}}$$

$$\Lambda_{11} = (\Sigma/\Sigma_{22})^{-1} = (\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1} = \Sigma_{1|2}^{-1}$$

$$\Rightarrow \underline{\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}} = \underline{\Lambda_{11}^{-1}} = \Sigma_{1|2}$$

$$\begin{aligned} \mu_1 - \Lambda_{11}^{-1}\Lambda_{12}(\mathbf{x}_2 - \mu_2) &= \Lambda_{11}^{-1}\Lambda_{11}\mu_1 - \Lambda_{11}^{-1}\Lambda_{12}(\mathbf{x}_2 - \mu_2) \\ &= \Lambda_{11}^{-1}[\Lambda_{11}\mu_1 - \Lambda_{12}(\mathbf{x}_2 - \mu_2)] \\ &= \underline{\Sigma_{1|2}[\Lambda_{11}\mu_1 - \Lambda_{12}(\mathbf{x}_2 - \mu_2)]} \end{aligned}$$