Conditional Gaussian

- $\boldsymbol{X} = \left[\begin{smallmatrix} \boldsymbol{X}_a \\ \boldsymbol{X}_b \end{smallmatrix}\right]$ jointly Gaussian. We want to find $p(\boldsymbol{x}_a \mid \boldsymbol{x}_b)$. The joint pdf is

$$\mathcal{N}\left(\boldsymbol{x} \,|\, \boldsymbol{\mu}, \boldsymbol{\Sigma}\right) = \frac{1}{(2\pi)^{n/2}(\det \boldsymbol{\Sigma})^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^{\intercal}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\right). \tag{1}$$

- We fix $\boldsymbol{X}_b = \boldsymbol{x}_b$ as a constant in $\mathcal{N}\left(\boldsymbol{x} \,|\, \boldsymbol{\mu}, \boldsymbol{\Sigma}\right)$ and find its *functional form*, which tells us what $p(\boldsymbol{x}_a \mid \boldsymbol{x}_b)$ is.
- From (1), we know that $p(\boldsymbol{x}_a \mid \boldsymbol{x}_b)$ is a Gaussian pdf, so we only need to determine its mean $\boldsymbol{\mu}_{a|b}$ and covariance $\boldsymbol{\Sigma}_{a|b}$.

$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_a \\ \boldsymbol{\mu}_b \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{aa} & \boldsymbol{\Sigma}_{ab} \\ \boldsymbol{\Sigma}_{ba} & \boldsymbol{\Sigma}_{bb} \end{bmatrix}$$

$$\boldsymbol{\mu}_a = \mathbb{E}[\boldsymbol{X}_a]$$

$$\boldsymbol{\Sigma}_{aa} = \mathbb{E}[(\boldsymbol{X}_a - \boldsymbol{\mu}_a)(\boldsymbol{X}_a - \boldsymbol{\mu}_a)^{\intercal}], \; \boldsymbol{\Sigma}_{ab} = \mathbb{E}[(\boldsymbol{X}_a - \boldsymbol{\mu}_a)(\boldsymbol{X}_b - \boldsymbol{\mu}_b)^{\intercal}] = \boldsymbol{\Sigma}_{ba}^{\intercal}$$

Conditional Gaussian

- Let $\boldsymbol{\Lambda} = \boldsymbol{\Sigma}^{-1} = \begin{bmatrix} \boldsymbol{\Lambda}_{aa} & \boldsymbol{\Lambda}_{ab} \\ \boldsymbol{\Lambda}_{ba} & \boldsymbol{\Lambda}_{bb} \end{bmatrix}$ and consider the exponent:

$$
\begin{aligned}
-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathsf{T}}\boldsymbol{\Lambda}(\boldsymbol{x}-\boldsymbol{\mu}) = -\frac{1}{2}\big[&(\boldsymbol{x}_a-\boldsymbol{\mu}_a)^{\mathsf{T}}\boldsymbol{\Lambda}_{aa}(\boldsymbol{x}_a-\boldsymbol{\mu}_a) \\
&+(\boldsymbol{x}_a-\boldsymbol{\mu}_a)^{\mathsf{T}}\boldsymbol{\Lambda}_{ab}(\boldsymbol{x}_b-\boldsymbol{\mu}_b) \\
&+(\boldsymbol{x}_b-\boldsymbol{\mu}_b)^{\mathsf{T}}\boldsymbol{\Lambda}_{ba}(\boldsymbol{x}_a-\boldsymbol{\mu}_a)\big] + \text{const.}
\end{aligned}
$$

  Compare with

$$
-\frac{1}{2}(\boldsymbol{x}_a-\boldsymbol{\mu}_{a|b})^{\mathsf{T}}\boldsymbol{\Sigma}_{a|b}^{-1}(\boldsymbol{x}_a-\boldsymbol{\mu}_{a|b}) = -\frac{1}{2}\boldsymbol{x}_a^{\mathsf{T}}\boldsymbol{\Sigma}_{a|b}^{-1}\boldsymbol{x}_a + \boldsymbol{x}_a^{\mathsf{T}}\boldsymbol{\Sigma}_{a|b}^{-1}\boldsymbol{\mu}_{a|b} + \text{const.}
$$

- Quadratic term in $\boldsymbol{x}_a$: $-\frac{1}{2}\boldsymbol{x}_a^{\mathsf{T}}\boldsymbol{\Lambda}_{aa}\boldsymbol{x}_a \implies \boldsymbol{\Sigma}_{a|b} = \boldsymbol{\Lambda}_{aa}^{-1}$.
- Linear term in $\boldsymbol{x}_a$: $\boldsymbol{x}_a^{\mathsf{T}}(\boldsymbol{\Lambda}_{aa}\boldsymbol{\mu}_a - \boldsymbol{\Lambda}_{ab}(\boldsymbol{x}_b-\boldsymbol{\mu}_b)) \implies$

$$
\begin{aligned}
\boldsymbol{\mu}_{a|b} &= \boldsymbol{\Sigma}_{a|b}(\boldsymbol{\Lambda}_{aa}\boldsymbol{\mu}_a - \boldsymbol{\Lambda}_{ab}(\boldsymbol{x}_b-\boldsymbol{\mu}_b)) \\
&= \boldsymbol{\mu}_a - \boldsymbol{\Lambda}_{aa}^{-1}\boldsymbol{\Lambda}_{ab}(\boldsymbol{x}_b-\boldsymbol{\mu}_b).
\end{aligned}
$$

Conditional Gaussian

- Inverse of partitioned matrix:

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{M} & -\mathbf{M}\mathbf{B}\mathbf{D}^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}\mathbf{M} & \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{C}\mathbf{M}\mathbf{B}\mathbf{D}^{-1} \end{bmatrix},$$

  where $\mathbf{M} = (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}$ is called the Schur complement w.r.t. $\mathbf{D}$.
- Applying the above, we obtain

$$\begin{aligned} \mathbf{\Lambda}_{aa} &= (\mathbf{\Sigma}_{aa} - \mathbf{\Sigma}_{ab}\mathbf{\Sigma}_{bb}^{-1}\mathbf{\Sigma}_{ba})^{-1}, \\ \mathbf{\Lambda}_{ab} &= -(\mathbf{\Sigma}_{aa} - \mathbf{\Sigma}_{ab}\mathbf{\Sigma}_{bb}^{-1}\mathbf{\Sigma}_{ba})^{-1}\mathbf{\Sigma}_{ab}\mathbf{\Sigma}_{bb}^{-1}, \end{aligned}$$

  and finally

$$\boldsymbol{\mu}_{a|b} = \boldsymbol{\mu}_a + \mathbf{\Sigma}_{ab}\mathbf{\Sigma}_{bb}^{-1}(\boldsymbol{x}_b - \boldsymbol{\mu}_b), \tag{2}$$

$$\mathbf{\Sigma}_{a|b} = \mathbf{\Sigma}_{aa} - \mathbf{\Sigma}_{ab}\mathbf{\Sigma}_{bb}^{-1}\mathbf{\Sigma}_{ba}. \tag{3}$$