

Multivariate Normal Distribution II

Nipun Batra and teaching staff

IIT Gandhinagar

August 30, 2025

Detour: Inverse of partioned symmetric matrix ¹

Consider an $n \times n$ symmetric matrix \mathbf{A} and divide it into four blocks

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{12}^T & \mathbf{A}_{22} \end{bmatrix}$$

For example, let $n = 3$, we have

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 6 \\ 3 & 6 & 8 \end{bmatrix}$$

We could for example have

$$\mathbf{A}_{11} = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} \text{ and } \mathbf{A}_{12} = \begin{bmatrix} 3 \\ 6 \end{bmatrix} \text{ and } \mathbf{A}_{22} = [8]$$

¹Courtesy: [http:](http://fourier.eng.hmc.edu/e161/lectures/gaussianprocess/node6.html)

[//fourier.eng.hmc.edu/e161/lectures/gaussianprocess/node6.html](http://fourier.eng.hmc.edu/e161/lectures/gaussianprocess/node6.html)

Detour: Inverse of partitioned symmetric matrix

Question: Write $\mathbf{B} = \mathbf{A}^{-1}$ in terms of the four blocks

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{12}^{\top} & \mathbf{B}_{22} \end{bmatrix} = \mathbf{A}^{-1}$$

\mathbf{A}_{11} and $\mathbf{B}_{11} \in \mathbb{R}^{p \times p}$

\mathbf{A}_{22} and $\mathbf{B}_{22} \in \mathbb{R}^{q \times q}$

$\mathbf{A}_{12} = \mathbf{A}_{21}^{\top}$ and $\mathbf{B}_{12} = \mathbf{B}_{21}^{\top} \in \mathbb{R}^{p \times q}$

and, $p + q = n$

Detour: Inverse of partitioned symmetric matrix

$$\begin{aligned} \mathbf{I}_n &= \mathbf{A}\mathbf{A}^{-1} = \mathbf{A}\mathbf{B} \\ &= \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{12}^\top & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{12}^\top & \mathbf{B}_{22} \end{bmatrix} = \\ &\begin{bmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{12}^\top & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} \\ \mathbf{A}_{12}^\top\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{12}^\top & \mathbf{A}_{12}^\top\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_q \end{bmatrix} \end{aligned}$$

Thus, we have

$$\begin{aligned} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{12}^\top &= \mathbf{I}_p \\ \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} &= \mathbf{0}^{p \times q} \\ \mathbf{A}_{12}^\top\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{12}^\top &= \mathbf{0}^{q \times p} \\ \mathbf{A}_{12}^\top\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} &= \mathbf{I}_q \end{aligned}$$

Detour: Inverse of partitioned symmetric matrix

Moving the expressions around we get the following results.

$$\begin{aligned}\mathbf{B}_{11} &= (\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{12}^{\top})^{-1} \\ &= \mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1}\mathbf{A}_{12}(\mathbf{A}_{22} - \mathbf{A}_{12}^{\top}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}\mathbf{A}_{12}^{\top}\mathbf{A}_{11}^{-1}\end{aligned}$$

$$\begin{aligned}\mathbf{B}_{22} &= (\mathbf{A}_{22} - \mathbf{A}_{12}^{\top}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1} \\ &= \mathbf{A}_{22}^{-1} + \mathbf{A}_{22}^{-1}\mathbf{A}_{12}^{\top}(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{12}^{\top})^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\end{aligned}$$

$$\mathbf{B}_{12}^{\top} = -\mathbf{A}_{22}^{-1}\mathbf{A}_{12}^{\top}(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{12}^{\top})^{-1}$$

$$\mathbf{B}_{12}^{\top} = -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}^{\top}(\mathbf{A}_{22} - \mathbf{A}_{12}^{\top}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}$$

Determinant of Partitioned Symmetric Matrix

Theorem: Determinant of a partitioned symmetric matrix can be written as follows

$$\begin{aligned}\det(\mathbf{A}) &= \det \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} \\ &= \det(\mathbf{A}_{11}) \det(\mathbf{A}_{22} - \mathbf{A}_{12}^{\top} \mathbf{A}_{11}^{-1} \mathbf{A}_{12}) \\ &= \det(\mathbf{A}_{22}) \det(\mathbf{A}_{11} - \mathbf{A}_{12} \mathbf{A}_{22}^{-1} \mathbf{A}_{12}^{\top})\end{aligned}$$

Determinant of Partitioned Symmetric Matrix

Proof: Note that

$$\begin{aligned}\mathbf{A} &= \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{A}_{12}^\top & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_{22} - \mathbf{A}_{12}^\top \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{I} & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{A}_{11} - \mathbf{A}_{12} \mathbf{A}_{22}^{-1} \mathbf{A}_{12}^\top & \mathbf{0} \\ \mathbf{A}_{22}^{-1} \mathbf{A}_{21} & \mathbf{I} \end{bmatrix}\end{aligned}$$

The theorem is proved as we also know that

$$\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B})$$

and

$$\det \begin{pmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \det \begin{pmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{0} & \mathbf{D} \end{pmatrix} = \det(\mathbf{B}) \det(\mathbf{D})$$

Marginalisation and Conditional of multivariate normal²

Assume an n -dimensional random vector

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 \end{bmatrix}$$

has a normal distribution $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with

$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix} \text{ and } \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}$$

where \mathbf{x}_1 and \mathbf{x}_2 are two subvectors of respective dimensions p and q with $p + q = n$. Note that $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}^\top$, and $\boldsymbol{\Sigma}_{21} = \boldsymbol{\Sigma}_{12}^\top$.

²Courtesy: <http://fourier.eng.hmc.edu/e161/lectures/gaussianprocess/node7.html>.

Marginalisation and Conditional of multivariate normal

Theorem:

part a: The marginal distributions of \mathbf{x}_1 and \mathbf{x}_2 are also normal with mean vector $\boldsymbol{\mu}_i$ and covariance matrix $\boldsymbol{\Sigma}_{ii}$ ($i = 1, 2$), respectively.

part b: The conditional distribution of \mathbf{x}_i given \mathbf{x}_j is also normal with mean vector

$$\boldsymbol{\mu}_{i|j} = \boldsymbol{\mu}_i + \boldsymbol{\Sigma}_{ij}\boldsymbol{\Sigma}_{jj}^{-1}(\mathbf{x}_j - \boldsymbol{\mu}_j)$$

Marginalisation and Conditional of multivariate normal

Proof:

The joint density of \mathbf{x} is:

$$f(\mathbf{x}) = f(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{(2\pi)^{n/2} \det(\Sigma)^{1/2}} \exp \left[-\frac{1}{2} Q(\mathbf{x}_1, \mathbf{x}_2) \right]$$

where Q is defined as

$$\begin{aligned} Q(\mathbf{x}_1, \mathbf{x}_2) &= (\mathbf{x} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) \\ &= [(\mathbf{x}_1 - \boldsymbol{\mu}_1)^\top, (\mathbf{x}_2 - \boldsymbol{\mu}_2)^\top] \begin{bmatrix} \Sigma^{11} & \Sigma^{12} \\ \Sigma^{21} & \Sigma^{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 - \boldsymbol{\mu}_1 \\ \mathbf{x}_2 - \boldsymbol{\mu}_2 \end{bmatrix} \\ &= (\mathbf{x}_1 - \boldsymbol{\mu}_1)^\top \Sigma^{11} (\mathbf{x}_1 - \boldsymbol{\mu}_1) + 2(\mathbf{x}_1 - \boldsymbol{\mu}_1)^\top \Sigma^{12} (\mathbf{x}_2 - \boldsymbol{\mu}_2) + (\mathbf{x}_2 - \boldsymbol{\mu}_2)^\top \cdots \\ &\quad \cdots \Sigma^{22} (\mathbf{x}_2 - \boldsymbol{\mu}_2) \end{aligned}$$

Marginalisation and Conditional of multivariate normal

Here we have assumed

$$\Sigma^{-1} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \Sigma^{11} & \Sigma^{12} \\ \Sigma^{21} & \Sigma^{22} \end{bmatrix}$$

According to inverse of a partitioned symmetric matrix we have,

$$\begin{aligned} \Sigma^{11} &= \left(\Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{12}^{\top} \right)^{-1} \\ &= \Sigma_{11}^{-1} + \Sigma_{11}^{-1} \Sigma_{12} \left(\Sigma_{22} - \Sigma_{12}^{\top} \Sigma_{11}^{-1} \Sigma_{12} \right)^{-1} \Sigma_{12}^{\top} \Sigma_{11}^{-1} \\ \Sigma^{22} &= \left(\Sigma_{22} - \Sigma_{12}^{\top} \Sigma_{11}^{-1} \Sigma_{12} \right)^{-1} \\ &= \Sigma_{22}^{-1} + \Sigma_{22}^{-1} \Sigma_{12}^{\top} \left(\Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{12}^{\top} \right)^{-1} \Sigma_{12} \Sigma_{22}^{-1} \\ \Sigma^{12} &= -\Sigma_{11}^{-1} \Sigma_{12} \left(\Sigma_{22} - \Sigma_{12}^{\top} \Sigma_{11}^{-1} \Sigma_{12} \right)^{-1} = \left(\Sigma^{21} \right)^{\top} \end{aligned}$$

Marginalisation and Conditional of multivariate normal

Substituting the second expression for Σ^{11} , first expression for Σ^{22} , and Σ^{12} into $Q(\mathbf{x}_1, \mathbf{x}_2)$ to get:

$$\begin{aligned} Q(\mathbf{x}_1, \mathbf{x}_2) &= \\ &(\mathbf{x}_1 - \boldsymbol{\mu}_1)^\top \left[\Sigma_{11}^{-1} + \Sigma_{11}^{-1} \Sigma_{12} (\Sigma_{22} - \mathbf{A}_{12}^\top \Sigma_{11}^{-1} \Sigma_{12})^{-1} \Sigma_{12}^\top \Sigma_{11}^{-1} \right] (\mathbf{x}_1 - \boldsymbol{\mu}_1) \\ &- 2(\mathbf{x}_1 - \boldsymbol{\mu}_1)^\top \left[\Sigma_{11}^{-1} \Sigma_{12} (\Sigma_{22} - \Sigma_{12}^\top \Sigma_{11}^{-1} \Sigma_{12})^{-1} \right] (\mathbf{x}_2 - \boldsymbol{\mu}_2) \\ &+ (\mathbf{x}_2 - \boldsymbol{\mu}_2)^\top \left[(\Sigma_{22} - \Sigma_{12}^\top \Sigma_{11}^{-1} \Sigma_{12})^{-1} \right] (\mathbf{x}_2 - \boldsymbol{\mu}_2) \\ &= (\mathbf{x}_1 - \boldsymbol{\mu}_1)^\top \Sigma_{11}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1) \\ &+ (\mathbf{x}_1 - \boldsymbol{\mu}_1)^\top \Sigma_{11}^{-1} \Sigma_{12} (\Sigma_{22} - \mathbf{A}_{12}^\top \Sigma_{11}^{-1} \Sigma_{12})^{-1} \Sigma_{12}^\top \Sigma_{11}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1) \\ &- 2(\mathbf{x}_1 - \boldsymbol{\mu}_1)^\top \left[\Sigma_{11}^{-1} \Sigma_{12} (\Sigma_{22} - \Sigma_{12}^\top \Sigma_{11}^{-1} \Sigma_{12})^{-1} \right] (\mathbf{x}_2 - \boldsymbol{\mu}_2) \\ &+ (\mathbf{x}_2 - \boldsymbol{\mu}_2)^\top \left[(\Sigma_{22} - \Sigma_{12}^\top \Sigma_{11}^{-1} \Sigma_{12})^{-1} \right] (\mathbf{x}_2 - \boldsymbol{\mu}_2) \end{aligned}$$

Marginalisation and Conditional of multivariate normal

$$\begin{aligned} &= (\mathbf{x}_1 - \boldsymbol{\mu}_1)^\top \boldsymbol{\Sigma}_{11}^{-1} \\ &\quad + \left[(\mathbf{x}_2 - \boldsymbol{\mu}_2) - \boldsymbol{\Sigma}_{12}^\top \boldsymbol{\Sigma}_{11}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1) \right]^\top \cdot \left(\boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{12}^\top \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{12} \right)^{-1} \\ &\quad \cdot \left[(\mathbf{x}_2 - \boldsymbol{\mu}_2) - \boldsymbol{\Sigma}_{12}^\top \boldsymbol{\Sigma}_{11}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1) \right] \end{aligned}$$

The last equal sign is due to the following equations for any vectors \mathbf{u} and \mathbf{v} and a symmetric matrix $\mathbf{A} = \mathbf{A}^\top$:

$$\begin{aligned} &\mathbf{u}^\top \mathbf{A} \mathbf{u} - 2\mathbf{u}^\top \mathbf{A} \mathbf{v} + \mathbf{v}^\top \mathbf{A} \mathbf{v} = \mathbf{u}^\top \mathbf{A} \mathbf{u} - \mathbf{u}^\top \mathbf{A} \mathbf{v} - \mathbf{u}^\top \mathbf{A} \mathbf{v} + \mathbf{v}^\top \mathbf{A} \mathbf{v} \\ &= \mathbf{u}^\top \mathbf{A} (\mathbf{u} - \mathbf{v}) - (\mathbf{u} - \mathbf{v})^\top \mathbf{A} \mathbf{v} = \mathbf{u}^\top \mathbf{A} (\mathbf{u} - \mathbf{v}) - \mathbf{v}^\top \mathbf{A} (\mathbf{u} - \mathbf{v}) \\ &= (\mathbf{u} - \mathbf{v})^\top \mathbf{A} (\mathbf{u} - \mathbf{v}) = (\mathbf{v} - \mathbf{u})^\top \mathbf{A} (\mathbf{v} - \mathbf{u}) \end{aligned}$$

Marginalisation and Conditional of multivariate normal

We define

$$\mathbf{b} \triangleq \boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{12}^\top \boldsymbol{\Sigma}_{11}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1)$$

$$\mathbf{A} \triangleq \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{12}^\top \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{12}$$

and

$$\left\{ \begin{array}{l} Q_1(\mathbf{x}_1) \triangleq (\mathbf{x}_1 - \boldsymbol{\mu}_1)^\top \boldsymbol{\Sigma}_1^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1) \\ \mathbf{N} = \left[(\mathbf{x}_2 - \boldsymbol{\mu}_2) - \boldsymbol{\Sigma}_{12}^\top \boldsymbol{\Sigma}_{11}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1) \right] \\ Q_2(\mathbf{x}_1, \mathbf{x}_2) \triangleq \mathbf{N}^\top \left(\boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{12}^\top \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{12} \right)^{-1} \mathbf{N} \\ = (\mathbf{x}_2 - \mathbf{b})^\top \mathbf{A}^{-1} (\mathbf{x}_2 - \mathbf{b}) \end{array} \right.$$

and get

$$Q(\mathbf{x}_1, \mathbf{x}_2) = Q_1(\mathbf{x}_1) + Q_2(\mathbf{x}_1, \mathbf{x}_2)$$

Marginalisation and Conditional of multivariate normal

Now the joint distribution can be written as:

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{(2\pi)^{n/2} \det(\Sigma)^{1/2}} \exp \left[-\frac{1}{2} Q(\mathbf{x}_1, \mathbf{x}_2) \right] \\ &= \frac{1}{(2\pi)^{n/2} \det(\Sigma_{11})^{1/2} \det(\Sigma_{22} - \Sigma_{12}^\top \Sigma_{11}^{-1} \Sigma_{12})^{1/2}} \exp \left[-\frac{1}{2} Q(\mathbf{x}_1, \mathbf{x}_2) \right] \\ &= \frac{1}{(2\pi)^{p/2} \det(\Sigma_{11})^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{x}_1 - \boldsymbol{\mu}_1)^\top \Sigma_{11}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1) \right] \\ &\quad \times \frac{1}{(2\pi)^{q/2} \det(\mathbf{A})^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{x}_2 - \mathbf{b})^\top \mathbf{A}^{-1} (\mathbf{x}_2 - \mathbf{b}) \right] \\ &= \mathcal{N}(\boldsymbol{\mu}_1, \Sigma_{11})_{\mathbf{x}_1} \mathcal{N}(\mathbf{b}, \mathbf{A})_{\mathbf{x}_2} \end{aligned}$$

The third equal sign is due to Determinant of a partitioned symmetric matrix:

$$\det(\Sigma) = \det(\Sigma_{11}) \det(\Sigma_{22} - \Sigma_{12}^\top \Sigma_{11}^{-1} \Sigma_{12})$$

Marginalisation and Conditional of multivariate normal

The marginal distribution of \mathbf{x}_1 is

$$\begin{aligned} f_1(\mathbf{x}_1) &= \int f(\mathbf{x}_1, \mathbf{x}_2) d\mathbf{x}_2 \\ &= \frac{1}{(2\pi)^{p/2} \det(\Sigma_{11})^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{x}_1 - \boldsymbol{\mu}_1)^\top \Sigma_{11}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1) \right] \end{aligned}$$

and the conditional distribution of \mathbf{x}_2 given \mathbf{x}_1 is

$$\begin{aligned} f_{2|1}(\mathbf{x}_2|\mathbf{x}_1) &= \frac{f(\mathbf{x}_1, \mathbf{x}_2)}{f(\mathbf{x}_1)} \\ &= \frac{1}{(2\pi)^{q/2} \det(\mathbf{A})^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{x}_2 - \mathbf{b})^\top \mathbf{A}^{-1} (\mathbf{x}_2 - \mathbf{b}) \right] \end{aligned}$$

Marginalisation and Conditional of multivariate normal

with

$$\mathbf{b} = \boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{12}^{\top} \boldsymbol{\Sigma}_{11}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1)$$

$$\mathbf{A} = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{12}^{\top} \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{12}$$