STAT 3690 Lecture 04

zhiyanggeezhou.github.io

Zhiyang Zhou (zhiyang.zhou@umanitoba.ca)

Feb 2nd, 2022

Block/partitioned matrix

- A partition of matrix: Suppose $\mathbf{A}_{11}$ is of $p \times r$, A_{12} is of $p \times s$, A_{21} is of $q \times r$ and A_{22} is of $q \times s$. Make a new $(p+q) \times (r+s)$-matrix by organizing A_{ij}'s in a 2 by 2 way:

$$\mathbf{A} = \left[\begin{array}{c|c} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22} \end{array}\right]$$

  e.g.,

$$\mathbf{A} = \left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 3 \\ \hline 4 & 5 & 6 \end{array}\right]$$

  if

$$\mathbf{A}_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{A}_{12} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad \mathbf{A}_{21} = \begin{bmatrix} 4 & 5 \end{bmatrix}, \quad \text{and} \quad \mathbf{A}_{22} = \begin{bmatrix} 6 \end{bmatrix}.$$

- Operations with block matrices
  - Working with partitioned matrices just like ordinary matrices
  - Matrix addition: if dimensions of A_{ij} and B_{ij} are quite the same, then

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

  - Matrix multiplication: if $A_{ij}B_{jk}$ makes sense for each i, j, k, then

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

  - Inverse: if $\mathbf{A}$, $\mathbf{A}_{11}$ and $\mathbf{A}_{22}$ are all invertible, then

$$\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{A}_{11.2}^{-1} & -\mathbf{A}_{11.2}^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{A}_{11.2}^{-1} & \mathbf{A}_{22.1}^{-1} \end{bmatrix}$$

    * $\mathbf{A}_{11.2} = \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}$
    * $\mathbf{A}_{22.1} = \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}$

- Conditional mean vectors and covariance matrices: If $\mathbf{X} \sim (\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix}, \quad \boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} > 0,$$

  where $\mathrm{E}(\mathbf{X}_i) = \boldsymbol{\mu}_i$ and $\mathrm{cov}(\mathbf{X}_i, \mathbf{X}_j) = \boldsymbol{\Sigma}_{ij}$, then
  - $\mathrm{E}(\mathbf{X}_i \mid \mathbf{X}_j = \boldsymbol{x}_j) = \boldsymbol{\mu}_i + \boldsymbol{\Sigma}_{ij}\boldsymbol{\Sigma}_{jj}^{-1}(\boldsymbol{x}_j - \boldsymbol{\mu}_j)$ for $i \neq j$ and $\boldsymbol{\Sigma}_{jj} > 0$
  - $\mathrm{cov}(\mathbf{X}_i \mid \mathbf{X}_j = \boldsymbol{x}_j) = \boldsymbol{\Sigma}_{ii} - \boldsymbol{\Sigma}_{ij}\boldsymbol{\Sigma}_{jj}^{-1}\boldsymbol{\Sigma}_{ji}$ for $i \neq j$ and $\boldsymbol{\Sigma}_{jj} > 0$

Multivariate normal (MVN) distribution

- Standard normal random vector
  - $\mathbf{Z} = [Z_1, \ldots, Z_p]^\top \sim MVN_p(\mathbf{0}, \mathbf{I}) \Leftrightarrow Z_1, \ldots, Z_p \overset{\text{iid}}{\sim} N(0,1) \Leftrightarrow$

$$\phi_{\mathbf{Z}}(\boldsymbol{z}) = (2\pi)^{-p/2} \exp(-\boldsymbol{z}^\top \boldsymbol{z}/2), \quad \boldsymbol{z} = [z_1, \ldots, z_p]^\top \in \mathbb{R}^p$$

- (General) normal random vector
  - Def: The distribution of $\mathbf{X}$ is MVN iff there exists $q \in \mathbb{Z}^+$, $\boldsymbol{\mu} \in \mathbb{R}^q$, $\mathbf{A} \in \mathbb{R}^{q\times p}$ and $\mathbf{Z} \sim MVN_p(\mathbf{0}, \mathbf{I})$ such that $\mathbf{X} = \mathbf{A}\mathbf{Z} + \boldsymbol{\mu}$
    - Limit the discussion to non-degenerate cases, i.e., $\text{rk}(\mathbf{A}) = q$
    - $\mathbf{X} \sim MVN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, i.e.,

$$f_{\mathbf{X}}(\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^p \det(\boldsymbol{\Sigma})}} \exp\{-(\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu})/2\}, \quad \boldsymbol{x} \in \mathbb{R}^p$$

      · $\boldsymbol{\Sigma} = \text{var}(\mathbf{X}) = \mathbf{A}\mathbf{A}^\top > 0$

- Exercise:
  1. $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}^\top > 0 \Leftrightarrow \text{rk}(\mathbf{A}) = q$ (Hint: SVD of $\mathbf{A}$);
  2. there exists a $p \times p$ positive definite matrix, say $\boldsymbol{\Sigma}^{1/2}$, such that $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}^{1/2}\boldsymbol{\Sigma}^{1/2}$ and $\boldsymbol{\Sigma}^{-1} = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}^{-1/2}$ (Hint: spectral decomposition of $\boldsymbol{\Sigma}$).

- Useful properties of MVN
  - $\mathbf{X} \sim MVN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Leftrightarrow \mathbf{Z} = \boldsymbol{\Sigma}^{-1/2}(\mathbf{X} - \boldsymbol{\mu}) \sim MVN_p(\mathbf{0}, \mathbf{I})$. So, we have a stochastic representation of arbitrary $\mathbf{X} \sim MVN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$: $\mathbf{X} = \boldsymbol{\Sigma}^{1/2}\mathbf{Z} + \boldsymbol{\mu}$, where $\mathbf{Z} \sim MVN_p(\mathbf{0}, \mathbf{I})$.
  - $\mathbf{X} \sim MVN$ iff, for all $a \in \mathbb{R}^p$, $a^\top \mathbf{X}$ has a (univariate) normal distribution.
  - If $\mathbf{X} \sim MVN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then $\mathbf{A}\mathbf{X} + \boldsymbol{b} \sim MVN_q(\mathbf{A}\boldsymbol{\mu} + \boldsymbol{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^\top)$ for $\mathbf{A} \in \mathbb{R}^{q\times p}$ and $\text{rk}(\mathbf{A}) = q$.
- Exercise: Generate six iid samples following bivariate normal $MVN_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with

$$\boldsymbol{\mu} = [3, 6]^\top, \quad \boldsymbol{\Sigma} = \begin{bmatrix} 10 & 2 \\ 2 & 5 \end{bmatrix}.$$

- Exercise:
  1. Prove that $(\mathbf{X} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) \sim \chi^2(p)$ if $\mathbf{X} \sim MVN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.
  2. Suppose $X_1 \sim N(0,1)$ and $\mathbf{X} = [X_1, X_2]^\top$. Does $\mathbf{X}$ follow an MVN in the following two cases?
     a. $X_2 = -X_1$;
     b. $X_2 = (2Y - 1)X_1$, where $Y \sim Ber(p)$ is independent of $\mathbf{X}$.

Joint, marginal and conditional MVN

- If $\mathbf{X} \sim MVN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and

$$\mathbf{X} = \left[\begin{array}{c} \mathbf{X}_1 \\ \hline \mathbf{X}_2 \end{array}\right], \quad \boldsymbol{\mu} = \left[\begin{array}{c} \boldsymbol{\mu}_1 \\ \hline \boldsymbol{\mu}_2 \end{array}\right] \quad \text{and} \quad \boldsymbol{\Sigma} = \left[\begin{array}{c|c} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \hline \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{array}\right]$$

  with $\boldsymbol{\Sigma}_{11} > 0$ and $\boldsymbol{\Sigma}_{22} > 0$, then
  - $\mathbf{X}_i \sim MVN_{p_i}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_{ii})$, i.e., marginals of MVN are MVN.

- $\mathbf{X}_i \mid \mathbf{X}_j = \boldsymbol{x}_j \sim MVN_{p_i}(\boldsymbol{\mu}_{i|j}, \boldsymbol{\Sigma}_{i|j})$, i.e., conditionals of MVN ar MVN.
    * $\boldsymbol{\mu}_{i|j} = \boldsymbol{\mu}_i + \boldsymbol{\Sigma}_{ij}\boldsymbol{\Sigma}_{jj}^{-1}(\boldsymbol{x}_j - \boldsymbol{\mu}_j)$
    * $\boldsymbol{\Sigma}_{i|j} = \boldsymbol{\Sigma}_{ii} - \boldsymbol{\Sigma}_{ij}\boldsymbol{\Sigma}_{jj}^{-1}\boldsymbol{\Sigma}_{ji}$
- $\mathbf{X}_i \perp \mathbf{X}_j \Leftrightarrow \boldsymbol{\Sigma}_{ij} = 0$