STAT 3690 Lecture 05

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 $p \times r$, $\mathbf{A}_{12}$ is of $p \times s$, $\mathbf{A}_{21}$ is of $q \times r$ and $\mathbf{A}_{22}$ is of $q \times s$. Make a new $(p+q) \times (r+s)$-matrix by organizing $\mathbf{A}_{ij}$'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 $\mathbf{A}_{ij}$ and $\mathbf{B}_{ij}$ 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 $\mathbf{A}_{ij}\mathbf{B}_{jk}$ makes sense for each i, j, k, then

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

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

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

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

```
options(digits = 4)
(Sigma = matrix(c(1, .5, .5,
              .5, 3, .5,
              .5, .5, 7),
            nrow = 3, ncol = 3))
```

```r
# Verify the inverse of partition matrix
## Method 1: by the above formula
(Sigma11 = Sigma[1:2, 1:2])
(Sigma12 = as.matrix(Sigma[1:2, 3]))
(Sigma21 = t(Sigma12))
(Sigma22 = as.matrix(Sigma[3, 3]))
(Sigma11.2 = Sigma11 - Sigma12 %*% solve(Sigma22) %*% Sigma21)
(Sigma22.1 = Sigma22 - Sigma21 %*% solve(Sigma11) %*% Sigma12)

(SigmaInv = rbind(
  cbind(solve(Sigma11.2), -solve(Sigma11.2) %*% Sigma12 %*% solve(Sigma22)),
  cbind(-solve(Sigma22) %*% Sigma21 %*% solve(Sigma11.2), solve(Sigma22.1))
))

## Method 2: solve()
solve(Sigma)
```

- 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., $\mathrm{rk}(\mathbf{A}) = q$
    - $\mathbf{X} \sim MVN_q(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, i.e.,

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

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

- Exercise:
  1. $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}^\top > 0 \Leftrightarrow \mathrm{rk}(\mathbf{A}) = q$ (Hint: SVD of $\mathbf{A}$);
  2. $\boldsymbol{\Sigma} > 0 \Rightarrow$ there exists a $q \times q$ 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 $\boldsymbol{\Sigma}$).

1. $\underset{q\times p}{A} = \underset{q\times q}{B}\,\underset{q\times p}{\Lambda}\,\underset{p\times p}{C^T}$, where $\Lambda = \left[\begin{array}{ccc|c} \lambda_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & \lambda_q & \end{array}\right]$ (SVD of A)

$\Rightarrow AA^T = B\Lambda C^T C \Lambda^T B^T$

$= B\Lambda\Lambda^T B^T$,

where $\Lambda\Lambda^T = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_q \end{bmatrix}$

$\therefore AA^T > 0 \Leftrightarrow \lambda_1, \cdots, \lambda_q > 0 \Leftrightarrow rk(A) = q$

2. $\Sigma = B\Lambda B^T$, where $\Lambda = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_q \end{bmatrix}$ (eigen-/spectral decomposition of Σ)

$\Rightarrow \Lambda^{-1} = \begin{bmatrix} \lambda_1^{-1} & & 0 \\ & \ddots & \\ 0 & & \lambda_q^{-1} \end{bmatrix}$ $(\because \Sigma > 0)$

$\Rightarrow \Sigma^{-1} = B\Lambda^{-1}B^T$ $(\because (B\Lambda^{-1}B^T)(B\Lambda B^T) = I)$

Let $\Lambda^{1/2} = \begin{bmatrix} \lambda_1^{-1} & & 0 \\ & \ddots & \\ 0 & & \lambda_q^{-1} \end{bmatrix}$, $\Lambda^{-\frac{1}{2}} = \begin{bmatrix} \lambda_1^{-\frac{1}{2}} & & 0 \\ & \ddots & \\ 0 & & \lambda_q^{-\frac{1}{2}} \end{bmatrix}$

$\Sigma^{1/2} = B\Lambda^{1/2}B^T$, $\Sigma^{-\frac{1}{2}} = B\Lambda^{-\frac{1}{2}}B^T$

then $\Sigma^{1/2}\Sigma^{1/2} = \Sigma$ and $\Sigma^{-\frac{1}{2}}\Sigma^{-\frac{1}{2}} = \Sigma^{-1}$