STAT 3690 Lecture 06

zhiyanggeezhou.github.io

Zhiyang Zhou (zhiyang.zhou@umanitoba.ca)

Feb 4nd, 2022

Multivariate normal (MVN) distribution (J&W Sec 4.2)

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

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

- (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.,

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

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

Derive the pdf of MVN.

$X = AZ + \mu$ and $\mathrm{rk}(A) = q \Rightarrow \mathrm{supp}(X) = \mathbb{R}^q$

If $q = p$: $Z = A^{-1}(X - \mu)$ and $J = \frac{\partial X}{\partial Z} = A^{-1}$ (matrix calculus)

$\because \phi_Z(z) = (2\pi)^{-q/2} \exp(-\frac{1}{2} z^T z)$, $z \in \mathbb{R}^q$

$\therefore f_X(x) = (2\pi)^{-q/2} |\det(A^{-1})| \exp\{-\frac{1}{2}(x-\mu)^T (A^T)^{-1} A^{-1} (x-\mu)\}$

$= (2\pi)^{-q/2} |\det(A)|^{-1} \exp\{-\frac{1}{2}(x-\mu)^T (AA^T)^{-1} (x-\mu)\}$

$= (2\pi)^{-q/2} |\det(AA^T)|^{-1/2} \exp\{-\frac{1}{2}(x-\mu)^T (AA^T)^{-1} (x-\mu)\}$ ($\because \det(A) = \det(A^T) = \sqrt{\det(AA^T)}$)

for all $x \in \mathbb{R}^q$

If $q < p$: $A = \underset{q\times q}{U} \underset{q\times p}{\Lambda} \underset{p\times p}{V^T}$ (SVD of A, $\Lambda = [\underset{q\times q}{\Lambda_1} \vdots \underset{q\times(p-q)}{0}]$)

Due to the results for "$q=p$", the pdf of $Y = V^T Z$ is

$f_Y(y) = (2\pi)^{-p/2} \exp(-\frac{1}{2} y^T y)$ for $y \in \mathbb{R}^p$ ($\because VV^T = V^TV = I$)

i.e., $Y = [Y_1, \dots, Y_p]^T \sim MVN_p(0, I) \Leftrightarrow Y_1, \dots, Y_p \overset{\text{iid}}{\sim} N(0,1)$

$\Rightarrow X = AZ + \mu = U\Lambda Y + \mu = U\Lambda_1 [Y_1, \dots, Y_q]^T + \mu$

$\Rightarrow f_X(x) = (2\pi)^{-q/2} |\det(AA^T)|^{-1/2} \exp\{-\frac{1}{2}(x-\mu)^T (AA^T)^{-1} (x-\mu)\}$ ($\because AA^T = U\Lambda_1\Lambda_1^T U^T = U\Lambda_1^2 U^T$)

for all $x \in \mathbb{R}^q$

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

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

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

$= B\Lambda\Lambda^\top B^\top$,

where $\Lambda\Lambda^\top = \begin{bmatrix} \lambda_1^2 & & 0 \\ & \ddots & \\ 0 & & \lambda_q^2 \end{bmatrix}$

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

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

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

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

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

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

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

```r
options(digits = 4)
(Sigma = matrix(c(10, 2 ,2, 5), ncol = 2, nrow = 2))
(spectral = eigen(Sigma))
(SigmaRoot = spectral$vectors %*% diag(spectral$values^.5) %*% t(spectral$vectors))
(SigmaRootInv = spectral$vectors %*% diag(spectral$values^-.5) %*% t(spectral$vectors))
# Check properties of root of Sigma
(SigmaRoot %*% SigmaRoot - Sigma)
(solve(SigmaRoot) - SigmaRootInv)
(SigmaRootInv %*% SigmaRootInv - solve(Sigma))
# SVD <=> spectral decomposition if Sigma is positive (semi-)definite
svd(Sigma)
eigen(Sigma)
```

- Useful properties of MVN
    - $\mathbf{X} \sim MVN_p(\boldsymbol{\mu}, \mathbf{\Sigma}) \Leftrightarrow \mathbf{Z} = \mathbf{\Sigma}^{-1/2}(\mathbf{X} - \boldsymbol{\mu}) \sim MVN_p(\mathbf{0}, \mathbf{I})$. So, we have a stochastic representation of arbitrary $\mathbf{X} \sim MVN_p(\boldsymbol{\mu}, \mathbf{\Sigma})$: $\mathbf{X} = \mathbf{\Sigma}^{1/2}\mathbf{Z} + \boldsymbol{\mu}$, 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 $\mathbf{X} \sim MVN_p(\boldsymbol{\mu}, \mathbf{\Sigma})$, then $\mathbf{A}\mathbf{X} + \boldsymbol{b} \sim MVN_q(\mathbf{A}\boldsymbol{\mu} + \boldsymbol{b}, \mathbf{A}\mathbf{\Sigma}\mathbf{A}^\top)$ for $\mathbf{A} \in \mathbb{R}^{q \times p}$ and $\mathrm{rk}(\mathbf{A}) = q$.

- Exercise: Generate six iid samples following bivariate normal $MVN_2(\boldsymbol{\mu}, \mathbf{\Sigma})$ with

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

```r
options(digits = 4)
set.seed(1)
Mu = matrix(c(3, 6), ncol = 1, nrow = 2)
Sigma = matrix(c(10, 2 ,2, 5), ncol = 2, nrow = 2)
n = 1000
# Method 1: via rnorm()
spectral = eigen(Sigma)
SigmaRoot = spectral$vectors %*% diag(spectral$values^.5) %*% t(spectral$vectors)
A1 = matrix(0, nrow = n, ncol = length(Mu))
for (i in 1:n) {
  A1[i, ] = t(SigmaRoot %*% matrix(rnorm(2), nrow = 2, ncol = 1) + Mu)
}
# Method 2: via MASS::mvrnorm()
A2 = MASS::mvrnorm(n, Mu, Sigma)
```