4 Random Vector and Matrices

- Expectation: Let $\mathbf{Y}$ and $\mathbf{X}$ be $p \times 1$ random vectors. The expected value of

$$\boldsymbol{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_p \end{pmatrix} \quad \text{is given by} \quad E(\boldsymbol{Y}) = \begin{pmatrix} E(Y_1) \\ E(Y_2) \\ \vdots \\ E(Y_p) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix} = \boldsymbol{\mu}$$

Note: $\underline{E(\boldsymbol{X} + \boldsymbol{Y}) = E(\boldsymbol{X}) + E(\boldsymbol{Y})}$

- Covariance Matrix:

Σ - symmetric, $|\Sigma| \geq 0$

$$\boldsymbol{\Sigma} = Cov(\boldsymbol{Y}) = E\{[\boldsymbol{Y} - E(\boldsymbol{Y})][\boldsymbol{Y} - E(\boldsymbol{Y})]'\} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{pmatrix}$$

$\sigma_{ij} = E(y_i - EY_i)(y_j - EY_j) = cov(y_i, y_j)$ if $i \neq j$

$\sigma_{ii} = Var(y_i) = \sigma_i^2$

a set of observations. $\underset{\sim}{y}_i = \begin{pmatrix} y_{i1} \\ \vdots \\ y_{ip} \end{pmatrix}$, $i = 1, \ldots, n$.

n - sample size

$\underset{\sim}{\bar{Y}} = \begin{pmatrix} \bar{y}_1 \\ \vdots \\ \bar{y}_p \end{pmatrix}$ $\quad \bar{y}_j = \frac{1}{n} \sum_{i=1}^{n} y_{ij}$

$\hat{\Sigma} = \underset{p\times p}{\hat{S}} = \frac{1}{n} \sum_{i=1}^{n} (\underset{\sim}{Y}_i - \underset{\sim}{\bar{Y}})(\underset{\sim}{Y}_i - \underset{\sim}{\bar{Y}})'$

sample covariance matrix

- Let $\underset{q\times p}{\boldsymbol{A}}$ be a constant matrix, then

$$Cov(\boldsymbol{A}\underset{p\times 1}{\boldsymbol{Y}}) = \boldsymbol{A}[Cov\boldsymbol{Y}]\boldsymbol{A}'$$

- Let $\boldsymbol{A}, \boldsymbol{B}$ be constant matrices, then

$$Cov(\boldsymbol{AX}, \boldsymbol{BY}) = \mathbf{A}Cov(\boldsymbol{X}, \boldsymbol{Y})\boldsymbol{B}'$$

$Y = (y_1 \cdots y_p) \sim (\underset{\sim}{\mu}, \underset{\sim}{\Sigma})$

- Generalized variance: overall measure of variability

$$\text{Generalized variance} = |\boldsymbol{\Sigma}|$$

* The larger the generalized variance, the more dispersed the data are.

* We can also use $\sqrt{|\underset{\sim}{\Sigma}|}$ or $\sqrt{|\underset{\sim}{S}|}$

* special case: $y_1 \cdots y_p$ are independent

$Var(y_j) = \sigma_j^2 = \sigma_{jj}$

$\underset{\sim}{\Sigma} = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_p^2 \end{pmatrix}$ $|\underset{\sim}{\Sigma}| = \sigma_1^2 \cdots \sigma_p^2$

— Mahalanobis distance (1936) (standardized distance)

$$(\underset{\sim}{y} - \underset{\sim}{\mu})' \underset{\sim}{\Sigma}^{-1} (\underset{\sim}{y} - \underset{\sim}{\mu})$$

$p\times 1$ $p\times p$ $p\times 1$

- Correlation matrices

$$\boldsymbol{\Omega} = \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & 1 & \cdots & \rho_{2p} \\ \vdots & \vdots & & \vdots \\ \rho_{p1} & \rho_{p2} & \cdots & 1 \end{pmatrix}$$

where

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{jj}}} = \frac{\sigma_{ij}}{\sigma_i \cdot \sigma_j}$$

for $i \neq j$.

- Partitioned random vectors

$$\boldsymbol{V} = \begin{pmatrix} \boldsymbol{Y} \\ \boldsymbol{X} \end{pmatrix}$$

$$\boldsymbol{\mu} = E(\boldsymbol{V}) = E\begin{pmatrix} \boldsymbol{Y} \\ \boldsymbol{X} \end{pmatrix} = \begin{pmatrix} E(\boldsymbol{Y}) \\ E(\boldsymbol{X}) \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_{\boldsymbol{Y}} \\ \boldsymbol{\mu}_{\boldsymbol{X}} \end{pmatrix}$$

$$\boldsymbol{\Sigma} = Cov(\boldsymbol{V}) = Cov\begin{pmatrix} \boldsymbol{Y} \\ \boldsymbol{X} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Sigma}_{\boldsymbol{YY}} & \boldsymbol{\Sigma}_{\boldsymbol{YX}} \\ \boldsymbol{\Sigma}_{\boldsymbol{XY}} & \boldsymbol{\Sigma}_{\boldsymbol{YY}} \end{pmatrix}$$

Σ_{XX}

$\underset{\sim}{\Sigma}^{-1}$ (use formula discussed in ch3)

- Let $\boldsymbol{Y}$ be a random vector with mean $\boldsymbol{\mu} = E(\boldsymbol{Y})$ and $\boldsymbol{\Sigma} = Cov(\boldsymbol{Y})$, then $E(\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}) = tr(\boldsymbol{A\Sigma}) + \boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}$ where $\boldsymbol{A}$ is a symmetric matrix.

Since $E(\underset{\sim}{Y}-\underset{\sim}{\mu})'\underset{\sim}{A}(\underset{\sim}{Y}-\underset{\sim}{\mu}) = E(\underset{\sim}{Y}'\underset{\sim}{A}\underset{\sim}{Y}) - \underset{\sim}{\mu}'\underset{\sim}{A}\underset{\sim}{\mu}$

then

$E(\underset{\sim}{Y}'\underset{\sim}{A}\underset{\sim}{Y}) = E(\underset{\sim}{Y}-\underset{\sim}{\mu})'\underset{\sim}{A}(\underset{\sim}{Y}-\underset{\sim}{\mu}) + \underset{\sim}{\mu}'\underset{\sim}{A}\underset{\sim}{\mu}$

$E(\underset{\sim}{Y}-\underset{\sim}{\mu})'\underset{\sim}{A}(\underset{\sim}{Y}-\underset{\sim}{\mu})$

$= tr[E\underbrace{(\underset{\sim}{Y}-\underset{\sim}{\mu})'}_{1\times p}\underbrace{\underset{\sim}{A}}_{p\times p}\underbrace{(\underset{\sim}{Y}-\underset{\sim}{\mu})}_{p\times 1}]$

$\boxed{tr(AB) = tr(BA)}$

$= E\ tr[(\underset{\sim}{Y}-\underset{\sim}{\mu})'\underset{\sim}{A}(\underset{\sim}{Y}-\underset{\sim}{\mu})]$

$= E\ tr[\underset{\sim}{A}(\underset{\sim}{Y}-\underset{\sim}{\mu})(\underset{\sim}{Y}-\underset{\sim}{\mu})']$

$= tr\{\underset{\sim}{A}\ E[(\underset{\sim}{Y}-\underset{\sim}{\mu})(\underset{\sim}{Y}-\underset{\sim}{\mu})']\}$

$= tr(\underset{\sim}{A}\underset{\sim}{\Sigma})$

#

- MGF: The moment generating function of a random vector $\boldsymbol{Y}$ is given by

$$M_{\mathbf{Y}}(\mathbf{t}) = E(e^{\mathbf{t}'\boldsymbol{Y}}) = h(t_1, \cdots, t_n)$$

$$= \int e^{\underset{\sim}{t}'\underset{\sim}{y}} p(\underset{\sim}{y}) d\underset{\sim}{y}$$

where $\mathbf{t} = \begin{pmatrix} \mathbf{t}_1 \\ \mathbf{t}_2 \\ \vdots \\ \mathbf{t}_n \end{pmatrix}$ if the expectation exists

for $-h < t_i < h$ where $h > 0$ and $i = 1, \cdots, n$

- Theorem
  Let $g_1(\boldsymbol{Y}_1), \cdots, g_m(\boldsymbol{Y}_m)$ be m functions of the random vectors $\boldsymbol{Y}_1, \cdots, \boldsymbol{Y}_m$, respectively. If $\boldsymbol{Y}_1, \cdots, \boldsymbol{Y}_m$ are mutually independent, then $g_1, \cdots, g_m$ are mutually independent.