STAT 3690 Lecture 04

zhiyanggeezhou.github.io

Zhiyang Zhou (zhiyang.zhou@umanitoba.ca)

Jan 31st, 2022

Covariance matrix of random vectors $\mathbf{X}$ and $\mathbf{Y}$

- Random p-vector $\mathbf{X} = [X_1, \ldots, X_p]^\top$ and q-vector $\mathbf{Y} = [Y_1, \ldots, Y_q]^\top$
- Expectations of random vectors/matrices are taken entry-wisely, e.g., $\boldsymbol{\mu}_\mathbf{X} = \mathrm{E}(\mathbf{X}) = [\mathrm{E}(X_1), \ldots, \mathrm{E}(X_p)]^\top$.
  - $\mathrm{E}(\mathbf{AX} + \boldsymbol{a}) = \mathbf{A}\mathrm{E}(\mathbf{X}) + \boldsymbol{a}$ as long as both $\mathbf{AX} + \boldsymbol{a}$ and $\mathbf{BY} + \boldsymbol{b}$ exist.
- Covariance matrix: the (i, j)-entry is the covariance between the i-th entry of $\mathbf{X}$ and j-th entry of $\mathbf{Y}$
  - $\boldsymbol{\Sigma}_{\mathbf{XY}} = [\mathrm{cov}(X_i, Y_j)]_{p\times q} = \mathrm{E}[\{\mathbf{X} - \mathrm{E}(\mathbf{X})\}\{\mathbf{Y} - \mathrm{E}(\mathbf{Y})\}^\top] = \mathrm{E}(\mathbf{XY}^\top) - \boldsymbol{\mu}_\mathbf{X}\boldsymbol{\mu}_\mathbf{Y}^\top$
  - $\boldsymbol{\Sigma}_{\mathbf{AX}+\boldsymbol{a},\mathbf{BY}+\boldsymbol{b}} = \mathbf{A}\Sigma_{\mathbf{XY}}\mathbf{B}^\top$ as long as both $\mathbf{AX} + \boldsymbol{a}$ and $\mathbf{BY} + \boldsymbol{b}$ exist.
  - $\boldsymbol{\Sigma}_{\mathbf{XX}} \geq 0$, i.e., $\boldsymbol{\Sigma}_{\mathbf{XX}}$ is positive semi-definite

---

- Exercise: Verify the properties of covariance matrix
  1. $\boldsymbol{\Sigma}_{\mathbf{AX}+\boldsymbol{a},\mathbf{BY}+\boldsymbol{b}} = \mathbf{A}\Sigma_{\mathbf{XY}}\mathbf{B}^\top$ as long as both $\mathbf{AX} + \boldsymbol{a}$ and $\mathbf{BY} + \boldsymbol{b}$ exist.
  2. $\boldsymbol{\Sigma}_{\mathbf{XX}} \geq 0$.

1. $\Sigma_{AX+a,BY+b} = E[\{AX+\cancel{a} - E(AX+\cancel{a})\}\{BY+\cancel{b} - E(BY+\cancel{b})\}^T]$

$= E[A\{X-E(X)\}\{Y-E(Y)\}^T B^T]$

$= AE[\{X-E(X)\}\{Y-E(Y)\}^T] B^T$

$= A\Sigma_{XY} B^T$

2. For all $a \in \mathbb{R}^p$,

$a^T\Sigma_{XX}a = E[a^T\{X-E(X)\}\{X-E(X)\}^T a]$

$= E[a^T\{X-E(X)\}]^2$ $(\because a^T\{X-E(X)\}$ is a scalar$)$

≥ 0

$\Leftrightarrow$ Eigenvalues of Σ_{XX} are all nonnegative

Sample covariance matrix

- $(\mathbf{X}_i, \mathbf{Y}_i) \overset{\text{iid}}{\sim} (\mathbf{X}, \mathbf{Y})$, $i = 1, \ldots, n$
- Sample means: $\bar{\mathbf{X}} = n^{-1} \sum_{i=1}^{n} \mathbf{X}_i$ and $\bar{\mathbf{Y}} = n^{-1} \sum_{i=1}^{n} \mathbf{Y}_i$
- Sample covariance matrix:

$$\mathbf{S}_{\mathbf{XY}} = \frac{1}{n-1} \sum_{i=1}^{n} \{(\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{Y}_i - \bar{\mathbf{Y}})^\top\}$$

  - Unbiasedness: $\mathrm{E}(\mathbf{S}_{\mathbf{XY}}) = \mathbf{\Sigma}_{\mathbf{XY}}$
  - $\mathbf{S}_{\mathbf{AX}+\boldsymbol{a},\mathbf{BY}+\boldsymbol{b}} = \mathbf{A}\mathbf{S}_{\mathbf{XY}}\mathbf{B}^\top$ as long as both $\mathbf{AX} + \boldsymbol{a}$ and $\mathbf{BY} + \boldsymbol{b}$ exist.
  - $\mathbf{S}_{\mathbf{XX}} \geq 0$
  - Implementation in *R*: cov() (or var() if $\mathbf{X} = \mathbf{Y}$)

- Exercise: Verify the properties of sample covariance matrix
  1. $\mathrm{E}(\mathbf{S}_{\mathbf{XY}}) = \mathbf{\Sigma}_{\mathbf{XY}}$. (Hint: $(n-1)\mathbf{S}_{\mathbf{XY}} = \sum_{i=1}^{n} \mathbf{X}_i \mathbf{Y}_i^\top - n\bar{\mathbf{X}}\bar{\mathbf{Y}}^\top = \sum_{i=1}^{n} \mathbf{X}_i \mathbf{Y}_i^\top - n^{-1} \sum_{i,j} \mathbf{X}_i \mathbf{Y}_j^\top$)
  2. $\mathbf{S}_{\mathbf{AX}+\boldsymbol{a},\mathbf{BY}+\boldsymbol{b}} = \mathbf{A}\mathbf{S}_{\mathbf{XY}}\mathbf{B}^\top$ as long as both $\mathbf{AX} + \boldsymbol{a}$ and $\mathbf{BY} + \boldsymbol{b}$ exist.
  3. $\mathbf{S}_{\mathbf{XX}} \geq 0$.

1. $(n-1) E(S_{XY}) = \sum_i E(X_i Y_i^T) - n^{-1} \sum_{i,j} E(X_i Y_j^T)$

$= \sum_i E(X_i Y_i^T) - n^{-1} \sum_{i=j} E(X_i Y_j^T) - n^{-1} \sum_{i \neq j} E(X_i Y_j^T)$

$= \sum_i E(X_i Y_i^T) - n^{-1} \sum_i E(X_i Y_i^T) - n^{-1} \sum_{i \neq j} E(X_i) E(Y_j^T)$ $(\because X_i \perp Y_j \text{ for } i \neq j)$

$= (n-1) E(X Y^T) - n^{-1}(n^2 - n) E(X) E(Y^T)$ $(\because (X_i, Y_i) \overset{iid}{\sim} (X, Y))$

$= (n-1) \Sigma_{XY}$

Method of moments (MM) estimators for mean vectors and covariance matrices

- MM imposes no specific distribution on $\mathbf{X}$ or $\mathbf{Y}$
- Steps
  1. Equate raw moments to their sample counterparts:

$$\begin{cases} \mathrm{E}(\mathbf{X}) = \bar{\mathbf{X}} \\ \mathrm{E}(\mathbf{Y}) = \bar{\mathbf{Y}} \\ \mathrm{E}(\mathbf{X}\mathbf{Y}^\top) = n^{-1} \sum_i \mathbf{X}_i \mathbf{Y}_i^\top \end{cases} \Leftrightarrow \begin{cases} \boldsymbol{\mu}_{\mathbf{X}} = \bar{\mathbf{X}} \\ \boldsymbol{\mu}_{\mathbf{Y}} = \bar{\mathbf{Y}} \\ \mathbf{\Sigma}_{\mathbf{XY}} + \boldsymbol{\mu}_{\mathbf{X}} \boldsymbol{\mu}_{\mathbf{Y}}^\top = n^{-1} \sum_i \mathbf{X}_i \mathbf{Y}_i^\top \end{cases}$$

2. Solve the above equations w.r.t. $\boldsymbol{\mu}_{\mathbf{X}}$, $\boldsymbol{\mu}_{\mathbf{Y}}$ and $\boldsymbol{\Sigma}_{\mathbf{XY}}$ and obtain estimators

$$\begin{cases} \hat{\boldsymbol{\mu}}_{\mathbf{X}} = \bar{\mathbf{X}} \\ \hat{\boldsymbol{\mu}}_{\mathbf{Y}} = \bar{\mathbf{Y}} \\ \widehat{\boldsymbol{\Sigma}}_{\mathbf{XY}} = n^{-1}\sum_i \mathbf{X}_i\mathbf{Y}_i^\top - \bar{\mathbf{X}}\bar{\mathbf{Y}}^\top = n^{-1}(n-1)\mathbf{S}_{\mathbf{XY}} \end{cases}$$

Computing means and covariance matrices by *R*

```
options(digits = 4)
install.packages(c('rgl', 'MASS'))
set.seed(1)

# parameters
n = 1000
Mu = 1:3
Sigma = matrix(c(1, .5, .5,
              .5, 3, .5,
              .5, .5, 7),
            nrow = 3, ncol = 3)

# check the eligibility of Sigma and review the spectral decomposition
isSymmetric.matrix(Sigma)
(eigen.Sig = eigen(Sigma))
(Lambda = diag(eigen.Sig$values))
(U = eigen.Sig$vectors)
(U %*% t(U))
(U %*% Lambda %*% t(U))

# generation of samples
samples = MASS::mvrnorm(n, Mu, Sigma)

# reference for various scatterplots https://www.statmethods.net/graphs/scatterplot.html
# scatterplots for paired RVs
pairs(samples)
# (spinning) 3D scatterplot
rgl::plot3d(samples[,1], samples[,2], samples[,3], col = "red", size = 6)

# sample mean vector for [V1,V2,V3]^T
(muHat = apply(samples, 2, mean))
(muHat = colMeans(samples))
# sample covariance matrix for [V1,V2,V3]^T
(S = var(samples))
(S = cov(samples))

# sample covariance matrix for V1 & [V2,V3]^T
cov(samples[,1], samples[,2:3])
# sample covariance matrix for V2 & [V3,V1]^T
cov(samples[,2], samples[,c(3,1)])
```