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., $\boldsymbol{\mu}_\mathbf{X} = \mathrm{E}(\mathbf{X}) = [\mathrm{E}(X_1), \ldots, \mathrm{E}(X_p)]^\top$.
  - $\mathrm{E}(\mathbf{A}\mathbf{X} + \boldsymbol{a}) = \mathbf{A}\mathrm{E}(\mathbf{X}) + \boldsymbol{a}$ for arbitrary non-random legit $\mathbf{A}$ and $\boldsymbol{a}$
- Covariance matrix: the (i, j)-entry is the covariance between the i-th entry of $\mathbf{X}$ and j-th entry of $\mathbf{Y}$
  - $\boldsymbol{\Sigma}_{\mathbf{XY}} = [\mathrm{cov}(X_i, Y_j)]_{p\times q} = \mathrm{E}[\{\mathbf{X} - \mathrm{E}(\mathbf{X})\}\{\mathbf{Y} - \mathrm{E}(\mathbf{Y})\}^\top] = \mathrm{E}(\mathbf{X}\mathbf{Y}^\top) - \boldsymbol{\mu}_\mathbf{X}\boldsymbol{\mu}_\mathbf{Y}^\top$
  - $\boldsymbol{\Sigma}_{\mathbf{AX}+\boldsymbol{a},\mathbf{BY}+\boldsymbol{b}} = \mathbf{A}\Sigma_{\mathbf{XY}}\mathbf{B}^\top$ for arbitrary non-random legit $\mathbf{A}$, $\boldsymbol{a}$, $\mathbf{B}$ and $\boldsymbol{b}$
  - $\boldsymbol{\Sigma}_{\mathbf{XX}} \geq 0$, i.e., $\boldsymbol{\Sigma}_{\mathbf{XX}}$ is positive semi-definite

---

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

Sample covariance matrix

- $(\mathbf{X}_i, \mathbf{Y}_i) \overset{\text{iid}}{\sim} (\mathbf{X}, \mathbf{Y})$, $i = 1, \ldots, n$
- Sample means: $\bar{\mathbf{X}} = n^{-1}\sum_{i=1}^n X_i$ and $\bar{\mathbf{Y}} = n^{-1}\sum_{i=1}^n Y_i$
- 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: $\mathrm{E}(\mathbf{S}_{\mathbf{XY}}) = \boldsymbol{\Sigma}_{\mathbf{XY}}$
  - $\mathbf{S}_{\mathbf{AX}+\boldsymbol{a},\mathbf{BY}+\boldsymbol{b}} = \mathbf{A}\mathbf{S}_{\mathbf{XY}}\mathbf{B}^\top$ for arbitrary non-random legit $\mathbf{A}$, $\boldsymbol{a}$, $\mathbf{B}$ and $\boldsymbol{b}$
  - $\mathbf{S}_{\mathbf{XX}} \geq 0$
  - Implementation in *R*: cov() (or var() if $\mathbf{X} = \mathbf{Y}$)

- Exercise: Verify the properties of sample covariance matrix
  1. $\mathrm{E}(\mathbf{S}_{\mathbf{XY}}) = \boldsymbol{\Sigma}_{\mathbf{XY}}$. (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. $\mathbf{S}_{\mathbf{AX}+\boldsymbol{a},\mathbf{BY}+\boldsymbol{b}} = \mathbf{A}\mathbf{S}_{\mathbf{XY}}\mathbf{B}^\top$ for arbitrary non-random legit $\mathbf{A}$, $\boldsymbol{a}$, $\mathbf{B}$ and $\boldsymbol{b}$.
  3. $\mathbf{S}_{\mathbf{XX}} \geq 0$.

Method of moments (MOM) estimator for mean vectors and covariance matrices

- MOM imposes no specific distribution on $\mathbf{X}$ or $\mathbf{Y}$

- Steps
  1. Equate raw moments to their sample counterparts:
$$\begin{cases} \mathrm{E}(\mathbf{X}) = \bar{\mathbf{X}} \\ \mathrm{E}(\mathbf{Y}) = \bar{\mathbf{Y}} \\ \mathrm{E}(\mathbf{X}\mathbf{Y}^\top) = n^{-1}\sum_i \mathbf{X}_i\mathbf{Y}_i^\top \end{cases} \Leftrightarrow \begin{cases} \boldsymbol{\mu}_{\mathbf{X}} = \bar{\mathbf{X}} \\ \boldsymbol{\mu}_{\mathbf{Y}} = \bar{\mathbf{Y}} \\ \boldsymbol{\Sigma}_{\mathbf{XY}} + \boldsymbol{\mu}_{\mathbf{X}}\boldsymbol{\mu}_{\mathbf{Y}}^\top = n^{-1}\sum_i \mathbf{X}_i\mathbf{Y}_i^\top \end{cases}$$
  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*

Identities of block/partitioned matrices

- A partition of covariance matrix
$$\boldsymbol{\Sigma} = \left[\begin{array}{c|c} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \hline \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{array}\right]$$
with $\boldsymbol{\Sigma}_{11}$ and $\boldsymbol{\Sigma}_{22}$ both square matrices. Then the inverse of $\boldsymbol{\Sigma} > 0$ is
$$\boldsymbol{\Sigma}^{-1} = \left[\begin{array}{c|c} \boldsymbol{\Sigma}_{11.2}^{-1} & -\boldsymbol{\Sigma}_{11.2}^{-1}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \\ \hline -\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11.2}^{-1} & \boldsymbol{\Sigma}_{22.1}^{-1} \end{array}\right] > 0$$
  - $\boldsymbol{\Sigma}_{11.2} = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$
  - $\boldsymbol{\Sigma}_{22.1} = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}$
- Conditional mean vectors and covariance matrices: If $\mathbf{X} \sim (\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and
$$\mathbf{X} = \left[\begin{array}{c} \mathbf{X}_1 \\ \hline \mathbf{X}_2 \end{array}\right], \quad \boldsymbol{\mu} = \left[\begin{array}{c} \boldsymbol{\mu}_1 \\ \hline \boldsymbol{\mu}_2 \end{array}\right] \quad \text{and} \quad \boldsymbol{\Sigma} = \left[\begin{array}{c|c} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \hline \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{array}\right] > 0,$$
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 $i \neq j$
  - $\mathrm{cov}(\mathbf{X}_i \mid \mathbf{X}_j = \boldsymbol{x}_j) = \boldsymbol{\Sigma}_{ii} - \boldsymbol{\Sigma}_{ij}\boldsymbol{\Sigma}_{jj}^{-1}\boldsymbol{\Sigma}_{ji}$ for $i \neq j$