Notes of Multivariate Statistical Analysis

Xie Zejian

2021-03-02

Contents

Chapter 1

Random Vectors

1.1 Covariance Matrix

Covariance matrix of random vector $\mathbf{x}$ with mean μ is defined as

$$\Sigma := \mathrm{Cov}\left[\mathbf{x}, \mathbf{x}\right] = \mathbb{E}(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})' = \mathbb{E}\,\mathbf{X}\mathbf{X}' - (\mathbb{E}\,\mathbf{X})(\mathbb{E}\,\mathbf{X})'$$

1. $\Sigma \geq 0$ and equality fail to hold when $|\Sigma| \neq 0$ 2. $\mathrm{Cov}\left[\mathbf{A}\mathbf{x}, \mathbf{B}\mathbf{y}\right] = \mathbf{A}\,\mathrm{Cov}\left[\mathbf{x}, \mathbf{y}\right]\mathbf{B}'$

The correction matrix of $\mathbf{x}$ is defined as

$$\boldsymbol{\rho} = \sqrt{\mathrm{diag}(\boldsymbol{\Sigma})^{-1}\,\boldsymbol{\Sigma}\,\mathrm{diag}(\boldsymbol{\Sigma})^{-1}}$$

1.2 Multivariate summary statistics

Suppose $(\mathbf{x}_i)_1^N$ be random sample from a multivariate distribution with mean μ, variance Σ and correction ρ. Then the sample mean and sample covariance is define as:

$$\hat{\boldsymbol{\mu}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i, \hat{\boldsymbol{\Sigma}} = \frac{1}{N-1}\sum_{i=1}^{N}(\mathbf{x_i} - \overline{\mathbf{x}})(\mathbf{x} - \overline{\mathbf{x}})'$$

Alternatively, arranging the observation vectors as the columns of a matrix, so that,

$$\mathbf{X}' = \begin{bmatrix}\mathbf{x}_1 & \mathbf{x}_2 & \ldots & \mathbf{x_N}\end{bmatrix}$$

then

$$\hat{\boldsymbol{\mu}} = \frac{\mathbf{X}'\mathbf{e}}{\mathbf{e}'\mathbf{e}}, \hat{\boldsymbol{\Sigma}} = \frac{\mathbf{X}(\mathbf{I} - \mathbf{P_e})\mathbf{X}'}{N-1}$$

and the sample correction matrix is $\hat{\boldsymbol{\rho}} = \sqrt{\mathrm{diag}(\hat{\boldsymbol{\Sigma}})^{-1}\,\hat{\boldsymbol{\Sigma}}\,\mathrm{diag}(\hat{\boldsymbol{\Sigma}})^{-1}}$

Proposition 1.1. *Those estimator are unbiased:*

1. $\mathbb{E}\,\hat{\boldsymbol{\mu}} = \boldsymbol{\mu}$, *Var* $\hat{\boldsymbol{\mu}} = \frac{1}{N}\boldsymbol{\Sigma}$
2. $\mathbb{E}\,\hat{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}$

Proof. **Mean**. Direct algebra yields:

$$\mathbb{E}\,\hat{\boldsymbol{\mu}} = \mathbb{E}\,\frac{\mathbf{X}'\mathbf{e}}{\mathbf{e}'\mathbf{e}} = \frac{1}{\mathbf{e}'\mathbf{e}}\,\mathbb{E}\,\mathbf{X}'\mathbf{e} = \frac{1}{\mathbf{e}'\mathbf{e}}\boldsymbol{\mu}\mathbf{e}'\mathbf{e} = \boldsymbol{\mu}$$

and

$$\operatorname{Var}\hat{\boldsymbol{\mu}} = \frac{1}{N^2}\operatorname{Var}\sum_{i=1}^{N}\mathbf{x_i} = \frac{1}{N^2}N\operatorname{Var}\mathbf{x_i} = \frac{1}{N}\boldsymbol{\Sigma}$$

Sample covariance. Note

$$\mathbb{E}\,\hat{\boldsymbol{\Sigma}} = \frac{1}{N-1}\,\mathbb{E}\,\mathbf{X}(\mathbf{I}-\mathbf{P_e})\mathbf{X}'$$

□

1.3 Exercises