Probability, kinda (2)

More like OLS (4), armed with some probability tools

Yue Jiang

February 14, 2023

The following material was used by Yue Jiang during a live lecture.

Without the accompanying oral comments, the text is incomplete as a record of the presentation.

An important disclaimer

This is not a probability class. There are semester-long (and multiple semester-long) courses on probability, and so what we cover in just two lectures scarcely touches on even the basics.

However, familiarity with some concepts from probability, such as probability distributions and certain aspects of them (e.g., expectation, variance, etc.) are needed to more fully grasp linear models. As such, we will be presenting a very abridged treatment of some of the fundamentals needed to proceed.

Review: The linear model in matrix form

$$\underbrace{\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}}_{\mathbf{y}} = \underbrace{\begin{bmatrix} 1 & x_{11} & \cdots & x_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{bmatrix}}_{\mathbf{X}} \underbrace{\begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix}}_{\boldsymbol{\epsilon}}$$

Vector-valued random variables

Suppose you have a $k-$vector $\mathbf{y}$, with some expectation $E(\mathbf{y})$ (element-wise expectations). Then its covariance matrix $Cov(\mathbf{y}) = \mathbf{\Sigma}$ is given by

$$\mathbf{\Sigma} = E\left((\mathbf{y} - E(\mathbf{y}))(\mathbf{y} - E(\mathbf{y}))^T\right)$$

- What is the dimension of $\mathbf{\Sigma}$?
- What are the diagonal entries? What is the $(i, j)th$ entry?
- Show that $\mathbf{\Sigma} = E(\mathbf{y}\mathbf{y}^T) - E(\mathbf{y})E(\mathbf{y})^T$.

Vector-valued random variables

Suppose you have a $k-$vector $\mathbf{x}$, with some expectation $E(\mathbf{x})$ (element-wise in the vector). Then its covariance matrix $\mathbf{\Sigma}$ is:

$$\mathbf{\Sigma} = E\left((\mathbf{x} - E(\mathbf{x}))(\mathbf{x} - E(\mathbf{x}))^T\right)$$

- For a fixed (i.e., non-random) matrix $\mathbf{A}$, what is $E(\mathbf{Ax})$?
- For a fixed matrix $\mathbf{A}$, what is $Cov(\mathbf{Ax})$?

Some assumptions

For now, assume that $\mathbf{X}$ is some fixed, non-random matrix of observations (we could let $\mathbf{X}$ "be random" and condition on it instead...more on this in office hours if you're interested, but long story short it doesn't matter in the end and I don't want to write out everything conditional on $\mathbf{X}$). As well,

- The model is truly $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$
- $E(\epsilon) = \mathbf{0}$
- $Cov(\epsilon) = \sigma^2\mathbf{I}$

Review: The OLS Estimator

The ordinary least squares estimator:

$$\widehat{\boldsymbol{\beta}} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y}$$

- What is the expectation of $\widehat{\boldsymbol{\beta}}$?
- What is the variance of $\widehat{\boldsymbol{\beta}}$?
- What assumptions did you use?

Review: Residuals

The **residuals** are defined as the differences between the actual value of the outcomes $\mathbf{y}$ and the values predicted by the model $\widehat{\mathbf{y}} = \mathbf{X}\widehat{\beta}$:

$$\widehat{\epsilon} = \mathbf{y} - \widehat{\mathbf{y}}$$

- Express the residuals in terms of observed quantities $\mathbf{y}$ and $\mathbf{X}$
- What is the expectation of the residuals? What is the covariance of them?

Homework 5: Due Feb. 21

1. Suppose $\mathbf{x}$ is a $k-$vector and $\mathbf{A}$ is a fixed $k \times k$ matrix (non-random). What is the expectation of the quadratic form $\mathbf{x}^T\mathbf{A}\mathbf{x}$? You may assume $E(\mathbf{x})$ and $Cov(\mathbf{x})$ exist.
2. Suppose the assumptions on Slide 6 hold. What is the variance of the i^{th} residual (hint: it is a function of σ^2)? Note: If we had a "good" estimate of σ^2, which we'll call $\widehat{\sigma}^2$. The i^{th} **standardized residual** might then be defined by $(y_i - \widehat{y}_i)$ divided by the square root of the answer to this exercise. Use this intuition to develop a notion of an *outlier* (no hard numbers required, just an idea).
3. Anything not finished from class today (note: an earlier version of these notes had a different problem that might have been a bit too difficult)