Notes 6

Maryclare Griffin

3/7/2023

These notes are based on Chapters 1, 5, and 6 of KNNL.

Recall the linear regression model, which we can now write out using linear algebra. Let $\boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$.

> ***Note:*** Given an $r \times 1$ vector $\boldsymbol{A}$ with elements A_i that are random variables, the expected value of $\boldsymbol{A}$, denoted by $\boldsymbol{E}\left\{A\right\}$, is an $r \times 1$ vector:
>
> $$\boldsymbol{E}\left\{\boldsymbol{A}\right\} = \boldsymbol{E}\left\{\begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_r \end{pmatrix}\right\} = \begin{pmatrix} E\left\{A_1\right\} \\ E\left\{A_2\right\} \\ \vdots \\ E\left\{A_r\right\} \end{pmatrix}$$
>
> ***Note:*** Given an $r \times 1$ vector $\boldsymbol{A}$ with elements A_i that are random variables, the variance of $\boldsymbol{A}$, denoted by $\boldsymbol{\sigma}^2\left\{A\right\}$, is an $r \times r$ matrix:
>
> $$\boldsymbol{\sigma}^2\left\{\boldsymbol{A}\right\} = \boldsymbol{\sigma}^2\left\{\begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_r \end{pmatrix}\right\} = \begin{pmatrix} \sigma^2\left\{A_1\right\} & \sigma\left\{A_1, A_2\right\} & \dots & \sigma\left\{A_1, A_r\right\} \\ \sigma\left\{A_1, A_2\right\} & \sigma^2\left\{A_2\right\} & \dots & \sigma\left\{A_2, A_r\right\} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma\left\{A_1, A_r\right\} & \sigma\left\{A_2, A_r\right\} & \dots & \sigma^2\left\{A_r\right\} \end{pmatrix}$$

The linear regression model for a dependent variable or response Y and independent variables, predictors, or covariates $X_1, \dots X_{p-1}$ is defined as:

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

where:

- The elements of $\boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{pmatrix}$ are parameters
- The elements of the $n \times p$ matrix $\boldsymbol{X} = \begin{pmatrix} 1 & X_{11} & X_{12} & \dots & X_{1,p-1} \\ 1 & X_{21} & X_{22} & \dots & X_{2,p-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \dots & X_{n,p-1} \end{pmatrix}$ are known constants
- $\boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$ is a random error term with mean $\boldsymbol{E}\left\{\boldsymbol{\epsilon}\right\} = \boldsymbol{0}$ and variance $\boldsymbol{\sigma}^2\left\{\boldsymbol{\epsilon}\right\} = \sigma^2\boldsymbol{I}$.

The least squares minimizing estimator $\boldsymbol{b}$ is given by:

$$\boldsymbol{b} = \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{Y}$$

Note: Given an $c \times s$ matrix $\boldsymbol{B}$ with elements B_{ij} that are random variables, and $r \times c$ and $s \times f$ matrices $\boldsymbol{A}$ and $\boldsymbol{D}$ with elements A_{ij} and D_{ij} that are fixed constants, the expected value of $\boldsymbol{ABD}$, denoted by $\boldsymbol{E}\left\{\boldsymbol{ABD}\right\}$, is an $r \times f$ matrix:

$$\boldsymbol{E}\left\{\boldsymbol{ABD}\right\} = \boldsymbol{AE}\left\{\boldsymbol{B}\right\}\boldsymbol{D}$$

Using the above fact about expectations of products of matrices, we can show that the least squares estimator is unbiased for $\boldsymbol{\beta}$ when the linear regression model holds. We have:

$$\begin{aligned}
E\left\{\boldsymbol{b}\right\} &= \mathbb{E}\left\{\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{Y}\right\} \\
&= \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\mathbb{E}\left\{\boldsymbol{Y}\right\} \\
&= \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\mathbb{E}\left\{\boldsymbol{X\beta}+\boldsymbol{\epsilon}\right\} \\
&= \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\left(\boldsymbol{X\beta}+\mathbb{E}\left\{\boldsymbol{\epsilon}\right\}\right) \\
&= \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{X\beta} \\
&= \boldsymbol{\beta}
\end{aligned}$$

It follows that $\hat{\boldsymbol{Y}} = \boldsymbol{Xb}$, which we will refer to as the **fitted values**, **estimated regression function**, or the **estimated mean response** is an unbiased estimator for $E\left\{\boldsymbol{Y}\right\} = \boldsymbol{X\beta}$.

Note: Given an $c \times 1$ vector $\boldsymbol{B}$ with elements B_i that are random variables, and $r \times c$ matrix $\boldsymbol{A}$ with elements A_{ij} that are fixed constants, the variance of $\boldsymbol{AB}$, denoted by $\boldsymbol{\sigma}^2\left\{\boldsymbol{AB}\right\}$, is an $r \times r$ matrix:

$$\boldsymbol{\sigma}^2\left\{\boldsymbol{AB}\right\} = \boldsymbol{A\sigma}^2\left\{\boldsymbol{B}\right\}\boldsymbol{A}'$$

Using the above fact about the variance of a product of a fixed matrix and a random vector, we can also derive the variance of the least squares estimator,

$$\begin{aligned}
\sigma^2\left\{\boldsymbol{b}\right\} &= \sigma^2\left\{\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{Y}\right\} \\
&= \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\sigma^2\left\{\boldsymbol{Y}\right\}\boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1} \\
&= \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\sigma^2\left\{\boldsymbol{\epsilon}\right\}\boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1} \\
&= \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\left(\sigma^2\boldsymbol{I}\right)\boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1} \\
&= \sigma^2\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1} \\
&= \sigma^2\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}
\end{aligned}$$

Without knowing σ^2, the variance of the least squares estimator $\boldsymbol{b}$ is unknown. Nonetheless, we can state that the least squares estimator $\boldsymbol{b}$ is has the lowest variance of all estimators of $\boldsymbol{\beta}$ that are linear in $\boldsymbol{Y}$.