Stat 6021:Leave One Out Formula and PRESS Residual

Read this after Section 2 from Guided Notes

1 Notation

Let $\hat{\boldsymbol{\beta}}_{(\boldsymbol{i})}$, $\boldsymbol{X}_{(\boldsymbol{i})}$, and $\boldsymbol{Y}_{(\boldsymbol{i})}$ denote the least-squares estimator for the regression model, the design matrix, and the vector of responses, with observation i removed from the data set. Therefore, the least-squares estimator with observation i removed can be written as

$$\begin{aligned} \hat{\boldsymbol{\beta}}_{(\boldsymbol{i})} &= (\boldsymbol{X}'_{(\boldsymbol{i})}\boldsymbol{X}_{(\boldsymbol{i})})^{-1}\boldsymbol{X}'_{(\boldsymbol{i})}\boldsymbol{Y}_{(\boldsymbol{i})} \\ &= (\boldsymbol{X}'\boldsymbol{X} - \boldsymbol{x}_{\boldsymbol{i}}\boldsymbol{x}'_{\boldsymbol{i}})^{-1}(\boldsymbol{X}'\boldsymbol{Y} - \boldsymbol{x}_{\boldsymbol{i}}y_i) \end{aligned} \tag{1}$$

2 Leave-One-Out Formula

We will use the following result on matrices without proof:

$$(\boldsymbol{A} + \boldsymbol{BCB'})^{-1} = \boldsymbol{A^{-1}} - \boldsymbol{A^{-1}B}(\boldsymbol{C^{-1}} + \boldsymbol{B'A^{-1}B})^{-1}\boldsymbol{B'A^{-1}} \tag{2}$$

Apply (2) to $(\boldsymbol{X}'\boldsymbol{X} - \boldsymbol{x}_{\boldsymbol{i}}\boldsymbol{x}'_{\boldsymbol{i}})^{-1}$, the first part of (1). Let $\boldsymbol{A} = \boldsymbol{X}'\boldsymbol{X}$, $\boldsymbol{B} = \boldsymbol{x}_{\boldsymbol{i}}$, and $\boldsymbol{C} = -1$, so

$$\begin{aligned} (\boldsymbol{X}'\boldsymbol{X} - \boldsymbol{x}_{\boldsymbol{i}}\boldsymbol{x}'_{\boldsymbol{i}})^{-1} &= (\boldsymbol{X}'\boldsymbol{X})^{-1} - (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x}_{\boldsymbol{i}}\left[-1 + \boldsymbol{x}'_{\boldsymbol{i}}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x}_{\boldsymbol{i}}\right]^{-1}\boldsymbol{x}'_{\boldsymbol{i}}(\boldsymbol{X}'\boldsymbol{X})^{-1} \\ &= (\boldsymbol{X}'\boldsymbol{X})^{-1} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x}_{\boldsymbol{i}}(1 - h_{ii})^{-1}\boldsymbol{x}'_{\boldsymbol{i}}(\boldsymbol{X}'\boldsymbol{X})^{-1} \end{aligned} \tag{3}$$

since the leverage $h_{ii} = \boldsymbol{x}'_{\boldsymbol{i}}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x}_{\boldsymbol{i}}$. Next, apply (3) to (1) to obtain the Leave-One-Out formula,

$$
\begin{aligned}
\hat{\boldsymbol{\beta}}_{(\boldsymbol{i})} &= \left[(\boldsymbol{X'X})^{-1} + (\boldsymbol{X'X})^{-1}\boldsymbol{x_i}(1-h_{ii})^{-1}\boldsymbol{x_i'}(\boldsymbol{X'X})^{-1}\right](\boldsymbol{X'Y} - \boldsymbol{x_i}y_i) \\
&= (\boldsymbol{X'X})^{-1}\boldsymbol{X'Y} \\
&\quad -(\boldsymbol{X'X})^{-1}\boldsymbol{x_i}y_i \\
&\quad +(\boldsymbol{X'X})^{-1}\boldsymbol{x_i}(1-h_{ii})^{-1}\boldsymbol{x_i'}(\boldsymbol{X'X})^{-1}\boldsymbol{X'Y} \\
&\quad -(\boldsymbol{X'X})^{-1}\boldsymbol{x_i}(1-h_{ii})^{-1}\boldsymbol{x_i'}(\boldsymbol{X'X})^{-1}\boldsymbol{x_i}y_i \\
&= \hat{\boldsymbol{\beta}} \\
&\quad -(1-h_{ii})^{-1}(1-h_{ii})(\boldsymbol{X'X})^{-1}\boldsymbol{x_i}y_i \\
&\quad +(1-h_{ii})^{-1}(\boldsymbol{X'X})^{-1}\boldsymbol{x_i}\boldsymbol{x_i'}\hat{\boldsymbol{\beta}} \\
&\quad -(1-h_{ii})^{-1}(\boldsymbol{X'X})^{-1}\boldsymbol{x_i}h_{ii}y_i \\
&= \hat{\boldsymbol{\beta}} - (1-h_{ii})^{-1}(\boldsymbol{X'X})^{-1}\boldsymbol{x_i}y_i\left[(1-h_{ii}+h_{ii}\right] \\
&\quad +(1-h_{ii})^{-1}(\boldsymbol{X'X})^{-1}\boldsymbol{x_i}\hat{y}_i \\
&= \hat{\boldsymbol{\beta}} - (1-h_{ii})^{-1}(\boldsymbol{X'X})^{-1}\boldsymbol{x_i}(y_i - \hat{y}_i) \\
&= \hat{\boldsymbol{\beta}} - (1-h_{ii})^{-1}(\boldsymbol{X'X})^{-1}\boldsymbol{x_i}e_i. \qquad (4)
\end{aligned}
$$

Note that $h_{ii}y_i = \hat{y}_i$.

3 PRESS Residual

Next, we show the two equivalent formulas for the PRESS residual. The PRESS residual is

$$
\begin{aligned}
e_{(i)} &= y_i - \hat{y}_{(i)} \\
&= y_i - \boldsymbol{x_i'}\hat{\boldsymbol{\beta}}_{(\boldsymbol{i})} \\
&= y_i - \boldsymbol{x_i'}\left[\hat{\boldsymbol{\beta}} - (1-h_{ii})^{-1}(\boldsymbol{X'X})^{-1}\boldsymbol{x_i}e_i\right] \text{ using (4)} \\
&= y_i - \hat{y}_i + \frac{1}{1-h_{ii}}\boldsymbol{x_i'}(\boldsymbol{X'X})^{-1}\boldsymbol{x_i}e_i \\
&= e_i + \frac{h_{ii}}{1-h_{ii}}e_i \\
&= \frac{e_i}{1-h_{ii}}. \qquad (5)
\end{aligned}
$$