MATH 523: RESIDUALS

Residual checking is a key component of assessment of model fit for the linear model. The vector of fitted values is $\widehat{Y} = \mathbf{X}\widehat{\beta}$, and the $(n \times 1)$ vector of residuals is

$$R = Y - \widehat{Y} = Y - \mathbf{X}\widehat{\beta} = (\mathbf{I}_n - \mathbf{X}(\mathbf{X}^\mathsf{T}\mathbf{X})^{-1}\mathbf{X}^\mathsf{T})Y = (\mathbf{I}_n - \mathbf{H})Y$$

with observed values

$$r = y - \widehat{y} = (\mathbf{I}_n - \mathbf{H})y.$$

Under the assumption of a correct model specification, we have that

$$\begin{aligned} \widehat{y}^\mathsf{T} r = (\mathbf{X}\widehat{\beta})^\mathsf{T}(y - \mathbf{X}\widehat{\beta}) &= \widehat{\beta}^\mathsf{T}\mathbf{X}^\mathsf{T}y - \widehat{\beta}^\mathsf{T}(\mathbf{X}^\mathsf{T}\mathbf{X})\widehat{\beta} \\ &= y^\mathsf{T}\mathbf{X}(\mathbf{X}^\mathsf{T}\mathbf{X})^{-1}\mathbf{X}^\mathsf{T}y - y^\mathsf{T}\mathbf{X}(\mathbf{X}^\mathsf{T}\mathbf{X})^{-1}(\mathbf{X}^\mathsf{T}\mathbf{X})(\mathbf{X}^\mathsf{T}\mathbf{X})^{-1}\mathbf{X}^\mathsf{T}y \\ &= 0 \end{aligned}$$

by the previous orthogonality results. By similar arguments, we have that

$$\mathbf{X}^\mathsf{T} r = \mathbf{X}^\mathsf{T}(y - \mathbf{X}\widehat{\beta}) = \mathbf{0}_p$$

that is, the observed residuals are

- orthogonal to fitted values $\widehat{y}$
- orthogonal to the columns of $\mathbf{X}$.

so therefore plots of residuals versus fitted values or the columns of $\mathbf{X}$ should show no pattern, and be centered at zero.

Note, however, that

$$y^\mathsf{T} r = y^\mathsf{T}(y - \mathbf{X}\widehat{\beta}) = y^\mathsf{T}y - y^\mathsf{T}\mathbf{X}(\mathbf{X}^\mathsf{T}\mathbf{X})^{-1}\mathbf{X}^\mathsf{T}y = y^\mathsf{T}(\mathbf{I}_n - \mathbf{H})y \neq 0$$

in general, so the residuals are **not** orthogonal to the observed responses.

Note also that by similar calculations for the random variables

$$\mathbb{E}[\widehat{Y}^\mathsf{T} R] = 0 \qquad\qquad \mathbb{E}[\mathbf{X}^\mathsf{T} R] = \mathbf{0}_p$$

and that

$$\text{Var}[R] = \text{Var}[(\mathbf{I}_n - \mathbf{H})Y] = (\mathbf{I}_n - \mathbf{H})\Sigma(\mathbf{I}_n - \mathbf{H})^\mathsf{T} = \mathbf{V}_R$$

say. This suggests that the residuals can be *standardized* by dividing them by their (estimated) standard errors; the standard error for R_i is the square root of the (i, i) element of the matrix $\mathbf{V}_R$

$$r_i^{(S)} = \frac{y_i - (\mathbf{X}\widehat{\beta})_i}{\sqrt{V_{Rii}}}.$$

If $\Sigma = \sigma^2\mathbf{I}_n$, then

$$\mathbf{V}_R = \sigma^2(\mathbf{I}_n - \mathbf{H})(\mathbf{I}_n - \mathbf{H})^\mathsf{T} = \sigma^2(\mathbf{I}_n - \mathbf{H} - \mathbf{H} + \mathbf{H}\mathbf{H}^\mathsf{T}) = \sigma^2(\mathbf{I}_n - \mathbf{H})$$

as $\mathbf{H}\mathbf{H}^\mathsf{T} = \mathbf{H}$ so that the i^{th} standardized residual is

$$r_i^{(S)} = \frac{y_i - (\mathbf{X}\widehat{\beta})_i}{\widehat{\sigma}\sqrt{1 - H_{ii}}}.$$