Some LS Outputs

Prediction at a new point $\mathbf{x}^*$

$$\hat{y}^* = \hat{\beta}_0 + x_{i1}^*\hat{\beta}_1 + \cdots + x_{ip}^*\hat{\beta}_p.$$

Fitted value at $\mathbf{x}_i$:

$$\hat{y}_i = \hat{\beta}_0 + x_{i1}\hat{\beta}_1 + \cdots + x_{ip}\hat{\beta}_p.$$

Residual at $\mathbf{x}_i$: $r_i = y_i - \hat{y}_i$.

RSS $= \sum_{i=1}^n r_i^2$.

The error variance is estimated by

$$\hat{\sigma}^2 = \frac{\text{RSS}}{n-p-1} = \frac{\sum_{i=1}^n r_i^2}{n-p-1}$$

The degree of freedom (df) of the residuals is $n-(p+1)$. In general

$$\begin{aligned} df(\text{residuals}) &= (\text{sample-size}) \\ &\quad -(\text{number-of-linear-coefs}) \end{aligned}$$

The Residual Vector

$\mathbf{X}^t\mathbf{r} = \mathbf{0}_{(p+1)\times 1}$ implies that the residual vector $\mathbf{r}$ is subject to $(p+1)$ equality constraints, therefore it loses $(p+1)$ degrees of freedom.