<u>**Matrices to Simplify Regression Notation**</u>

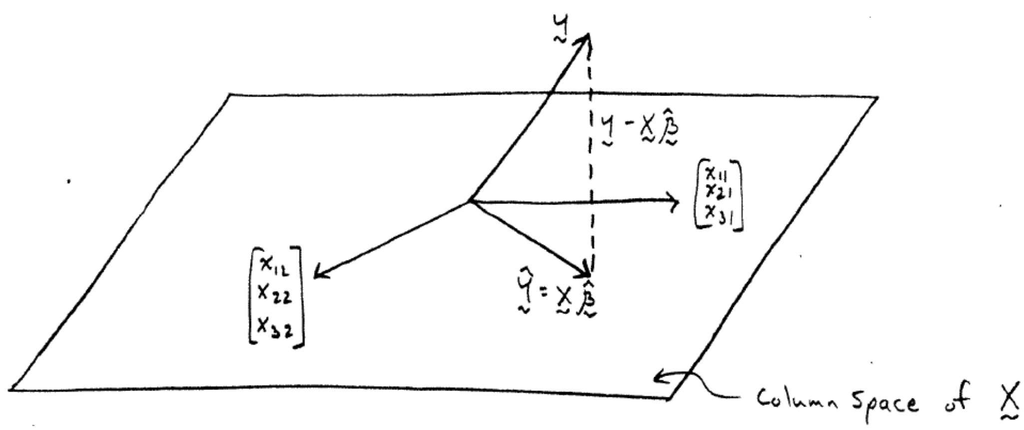
<u>SLR</u>:

$$
\begin{aligned}
Y_1 &= \beta_0 + \beta_1 x_1 + \varepsilon_1 \\
Y_2 &= \beta_0 + \beta_1 x_2 + \varepsilon_2 \\
&\vdots \qquad\qquad\quad \vdots \\
Y_n &= \beta_0 + \beta_1 x_n + \varepsilon_n
\end{aligned}
\qquad
\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} 1 & X_1 \\ 1 & X_2 \\ 1 & \vdots \\ 1 & X_n \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}
$$

$$\underset{\sim}{Y} = \underset{\sim}{X}\ \underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

- <u>Least Squares Estimates:</u> $\hat{\underset{\sim}{\beta}} = (\hat{\beta}_0, \hat{\beta}_1)'$ chosen to minimize SSE
  - $\underset{\sim}{Y} = \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{\varepsilon} \quad \rightarrow \quad \underset{\sim}{\varepsilon} = \underset{\sim}{Y} - \hat{\underset{\sim}{Y}} \quad = \quad \underset{\sim}{Y} - \underset{\sim}{X}\underset{\sim}{\beta} \quad \rightarrow \quad \underset{\sim}{\varepsilon}'\underset{\sim}{\varepsilon} = \varepsilon_1{}^2 + \varepsilon_2{}^2 + \cdots \varepsilon_n{}^2 = SSE$
  - $\hat{\underset{\sim}{\beta}} = (\hat{\beta}_0, \hat{\beta}_1)' = \left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}\underset{\sim}{X}'\underset{\sim}{Y}$
  - Geometry of LSEs:

$$\text{Let } \underset{\sim}{X}_{3\times 2} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{bmatrix} \quad \text{and} \quad \underset{\sim}{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$



Properties of Residuals

1. $\Sigma e_i = 0$
2. $\frac{\Sigma e_i^2}{n-(p+1)} = MSE = s^2 = \hat{\sigma}^2$
3. Same scale as the y_i observations
4. normalized residuals are scale independent

Normalized Residuals

	Text	R	AKA.
$r_i = \frac{e_i - 0}{s\sqrt{1-h_{ii}}}$	"standardized"	rstandard()	internally studentized
$r_{(-i)} = \frac{e_i - 0}{s_{(-i)}\sqrt{1-h_{ii}}}$	"jackknife"	rstudent()	externally studentized

- $s_{(-i)}$ is the MSE computed if we leave out the i^{th} observation

Testing for outliers (in the y-direction)

- $r_{(-i)}$ is better modeled by a T-distribution than r_i
- H_0: $r_{(-i)}$ is not from an outlier

Cook's Distance

- Measures the average standardized distance between $\hat{\underset{\sim}{\beta}}$ and $\hat{\underset{\sim}{\beta}}_{(-i)}$
- Measures influence in both the X and Y direction

$$D_i = \left(\frac{r_i^2}{p+1}\right)\left(\frac{h_{ii}}{1-h_{ii}}\right)$$

- If the model is correct then $D_i \sim F_{p,\, n-(p+1)}$

Transformations and Tukey's "Rule of the Bulge"

- Observe which way the curve bulges as suggested by a scatterplot of the data
- Transform ***y*** or ***x*** (or both) according to the signs of the corresponding quadrant:

  up → powers >1

  down → powers <1 (including logarithm)