Linear Models

Prof. Lauren Perry

Isn't this class called *Linear* Models?

You may have noticed that our broad definition of regression didn't say anything about linearity.

Let's define the current problem a little more closely.

Suppose for convenience we have three predictors, X_1, X_2, X_3.

We often use the very general form

$$Y = f(X_1, X_2, X_3) + \epsilon$$

where f is some unknown function and ϵ is the error in this representation.

This is still a super broad definition!

Instead, we often assume some more restricted form of f, such as linearity:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \epsilon$$

where each β_i is an unknown parameter.

Here, the problem is reduced to a finite parameter estimation, instead of the infinite possibilities for f.

Linearity

The linear model is linear *under the parameters*. That means the predictors don't need to be!

$$Y = \beta_0 + \beta_1 X_1^2 + \epsilon$$

and

$$Y = \beta_0 + \beta_2 \log X_2 + \beta_3 X_3 + \epsilon$$

and

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 X_3 + \epsilon$$

are all linear models!

This makes linear models much more flexible than they may seem at first pass.

Matrix Representation

We can also represent linear models as matrices. This looks like

$$y = X\beta + \epsilon$$

where $y = (y_1, \ldots, y_n)^T$, $\epsilon = (\epsilon_1, \ldots, \epsilon_n)^T$, $\beta = (\beta_0, \ldots, \beta_p)^T$ and

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \ldots & x_{1p} \\ 1 & x_{21} & x_{22} & \ldots & x_{2p} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \ldots & x_{np} \end{bmatrix}$$