6.6) Applications to Linear Models

We're going to transform $Ax = b$ to $X\beta = y$, where

- X is the design matrix
- β is the parameter vector
- y is the observation vector

Least squares Line : $y = \beta_0 + \beta_1 x$

Given a set of points (x_1, y_1) through (x_n, y_n), we want to determine the two $\beta_0 + \beta_1 x$ that makes the line as close to the points as possible

y_j is the observed values of y, if we substitute x_j into $y = \beta_0 + \beta_1 x$

- difference between observed and predicted is the residual

If all points were on a line, the predicted would be equal to the observed :

- $\beta_0 + \beta_1 x_1 = y_1$
- $\beta_0 + \beta_1 x_2 = y_2$

$$\vdots$$

$$\beta_0 + \beta_1 x_n = y_n$$

In practicality, all your data points won't be on this line,

Thus this becomes a least squares problem:

$$X\beta = y$$

So in order to find the right β or coefficients,

we can use: $X^T X \beta = X^T y$ and find

$X^T y$ and $X^T X$ to solve for β.