Linear Regression - Review

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

- What do each of these symbols mean?

  x_i = predictor variable value at i

  y_i = response variable value at i

  β_0 = y-intercept of best-fit line

  β_1 = slope of best fit line

  ε_i = residual of observed x_i to $\hat{y}_i$ based on line

- What are we estimating?

  β_0, β_1

- How do we do the estimation?

  OLS or ML

Linear regression assumptions:

1. Normality
2. Homogeneity of variance
3. Independence

Others:

- linearity
- fixed X-values

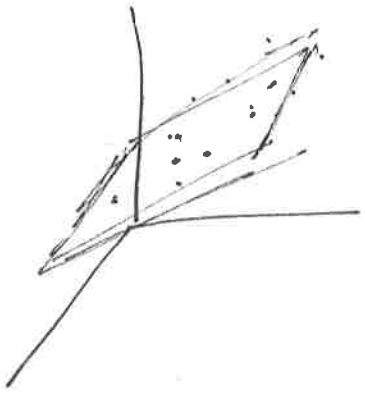
- Aside: standardized residuals = residual / standard deviation of residuals
- Aside: F-statistic returned by `lm` of regression → comparisson to model w/ only mean.
- Aside: adjusted R^2 = R^2 that is adjusted for the number of predictors

How do we know if our model fits our assumptions?

- Regression diagnotics.

Multiple Linear Regression

Rather than predicting the relationship between a response and the predictors with a single line, we are now looking at a plane



The slope of a line in one-dimension of the plane is the partial regression coefficient.