Multiple Linear Regression II: An Overview

- Extra sums of squares
  - Uses of extra sums of squares in tests for regression coefficients
- Coefficients of partial determination

Extra Sums of Squares

- An extra sum of squares measures the marginal reduction (increase) in the SSE (SSR) when one or several predictor variables are added to the regression model, given other predictor variables are already in the model.
- Extra sums of squares are useful for constructing tests about subsets of regression coefficients.

Example 1

- Response variable Y and 2 predictor variables X_1, X_2.
- The reduced model is $Y = \beta_0 + \beta_1 X_1 + \epsilon$, and compute $SSE(X_1)$.
- The full model is $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$, and compute $SSE(X_1, X_2)$.
- Adding another predictor variable will *never* increase SSE (and hence never reduce R^2).
  - In fact

$$SSE(X_1) = SSE(X_1, X_2) + SS?$$

- Define SS? as the extra sum of squares and denote it by $SSR(X_2|X_1)$. That is,

$$SSR(X_2|X_1) = SSE(X_1) - SSE(X_1, X_2).$$

  - $SSR(X_2|X_1)$ measures the decrease in the SSE when X_2 is added to the regression model, given X_1 is already in the model.

- Equivalently

  - That is, equivalently, $SSR(X_2|X_1)$ measures the increase in the SSR when X_2 is added to the regression model, given X_1 is already in the model.

Example 2

- Response variable Y and 3 predictor variables X_1, X_2, X_3.
- The full model is

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

  and compute $SSE(X_1, X_2, X_3)$.
- The reduced model is

$$Y = \beta_0 + \beta_2 X_2 + \epsilon$$

  and compute $SSE(X_2)$.

Example 2

- The extra sum of squares $SSR(X_1, X_3|X_2)$ is

$$SSR(X_1, X_3|X_2) = SSE(X_2) - SSE(X_1, X_2, X_3)$$

  - $SSR(X_1, X_3|X_2)$ measures the decrease in the SSE when X_1 and X_3 are added to the regression model, given X_2 is already in the model.

- Equivalently,

$$SSR(X_1, X_3|X_2) = SSR(X_1, X_2, X_3) - SSR(X_2)$$

  - Thus $SSR(X_1, X_3|X_2)$ measures the increase in the SSR when X_1 and X_3 are added to the regression model, given X_2 is already in the model.

Decomposition of SSR into Extra Sums of Squares

- Begin with $SSTO = SSR(X_1) + SSE(X_1)$.
  - Since $SSE(X_1) = SSR(X_2|X_1) + SSE(X_1, X_2)$, we have

- Or begin with $SSTO = SSR(X_1, X_2) + SSE(X_1, X_2)$.
  - Since $SSR(X_1, X_2) = SSR(X_1) + SSR(X_2|X_1)$, we have

ANOVA Table

For $X_1, \ldots, X_{p-1}$ in general, we may summarize the decomposition of SSR into extra sums of squares in an ANOVA table:

Source	SS	df
Regression	$SSR(X_1, X_2, \ldots, X_{p-1})$	$p-1$
X_1	$SSR(X_1)$	1
X_2	$SSR(X_2 \mid X_1)$	1
$\ldots$	$\ldots$	
X_{p-1}	$SSR(X_{p-1} \mid X_1, \ldots, X_{p-2})$	1
Error	$SSE(X_1, X_2, \ldots, X_{p-1})$	$n-p$
Total	$SSTO$	$n-1$

- Order matters: small to big – Forward.

Decomposition into Extra Sums of Squares: Remarks

- Choose order of predictor variables to enable desired inferences. For example,

$$\begin{aligned} SSTO &= SSR(X_1) + SSR(X_2|X_1) + SSE(X_1, X_2) \\ SSTO &= SSR(X_2) + SSR(X_1|X_2) + SSE(X_1, X_2). \end{aligned}$$

  - Generally, SS values will depend on order of predictor variables.
  - The number of possible orderings becomes large as the number of predictor variables increases. (How many exactly in terms of p?)
  - The extra sums of squares in the ANOVA table above are called Type I SS in SAS (or, sequential SS).
  - Decomposition into Type I SS is useful when there is a pre-determined order for selecting predictor variables (e.g. polynomial regression).

General Linear Test Approach

- Consider the full model

$$Y = \beta_0 + \beta_1 X_1 + \epsilon,$$

  and obtain $SSE(F)$.
- Consider the reduced model when $\beta_1 = 0$

$$Y = \beta_0 + \epsilon$$

  and obtain $SSE(R)$.
- Recall that, under $H_0 : \beta_1 = 0$,

Partial F Test: Example 1

- Return to Example 1.
- The test statistic for $H_0 : \beta_2 = 0$ is
- Under H_0, $F^* \sim$
- The decision rule is to
    - This is equivalent to a t-test for β_2 in the full model, as $(t^*)^2 = F^*$.

Partial F Test: Example 2

- Return to Example 2.
- The test statistic for $H_0 : \beta_1 = \beta_3 = 0$ is
- Under H_0, $F^* \sim$
- The decision rule is to

Partial F Test: Unusual Model Comparison

- Response variable Y and 3 predictor variables X_1, X_2, and X_3.
  - The full model is $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \epsilon$, and compute $SSE(X_1, X_2, X_3)$.
- Test

$$H_0 : \beta_1 = \beta_2 \text{ v.s. } H_a : \beta_1 \neq \beta_2.$$

  - The reduced model is
  - The test statistic is
  - Under H_0, $F^* \sim$
  - The decision rule is to

Partial Sums of Squares

For $X_1, \ldots, X_{p-1}$ in general, we may summarize the decomposition of SSR into partial sums of squares in an ANOVA table:

Source	SS	df
Regression	$SSR(X_1, X_2, \ldots, X_{p-1})$	$p-1$
X_1	$SSR(X_1 \mid X_2, X_3, \ldots, X_{p-1})$	1
X_2	$SSR(X_2 \mid X_1, X_3, \ldots, X_{p-1})$	1
$\ldots$	$\ldots$	
X_{p-1}	$SSR(X_{p-1} \mid X_1, X_2, \ldots, X_{p-2})$	1
Error	$SSE(X_1, X_2, \ldots, X_{p-1})$	$n-p$
Total	$SSTO$	$n-1$

Partial Sums of Squares

- The results are independent of the order of the predictor variables.
- The partial sums of squares in this ANOVA table are called Type III SS in SAS.
- These provide $p - 1$ individual hypothesis tests (Beware of multiple comparisons problem) and has backward flavor.
- These are the default output of most regression programs such as `lm()` in `R`.
  - Note: For extra SS, use functions like `anova()` to produce the ANOVA table from `lm()` results.

Coefficient of Partial Determination

- Coefficient of partial determination: the marginal contribution of one predictor variable when all others are already in the model.
- For example, with 3 predictor variables, the coefficients of partial determination are

$$
\begin{aligned}
R^2_{Y1|23} &= \frac{SSR(X_1|X_2, X_3)}{SSE(X_2, X_3)} \\
R^2_{Y2|13} &= \frac{SSR(X_2|X_1, X_3)}{SSE(X_1, X_3)} \\
R^2_{Y3|12} &= \frac{SSR(X_3|X_1, X_2)}{SSE(X_1, X_2)}
\end{aligned}
$$

- These are rarely used.