

Lecture 5-1: Regression with Two Regressors

Goal: Introduce mechanics of multiple regression with two regressors.

- Motivation for multiple regression
 - Richer description and prediction
 - Control for confounders
 - Model nonlinearity and interactions
 - Parameters, estimators, estimates
- Two cases
 - Binary + continuous
 - Continuous + continuous
- Mechanics of multiple regression
 - Derivation and Properties of OLS
 - Analysis of Variance
 - Partialling out interpretation

Worksheet

1. (*Interpreting a Binary Covariate in a Two-Regressor Model*) Suppose

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2, \quad (\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2) = (78.3, -10.6, -21.5),$$

where y is executive corruption (0–100), x_1 is high court independence (0–4), and $x_2 \in \{0, 1\}$ indicates democracy.

- For a fixed level of court independence x_1 , compute the predicted difference $\hat{y}(x_2 = 1) - \hat{y}(x_2 = 0)$ and interpret.
 - For a fixed regime type, compute the predicted change when x_1 increases by one unit.
 - Verify that the slope on court independence is the same in both groups and explain why.
2. (*Regression Plane and Counterfactual Prediction*) With two continuous covariates, suppose

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2, \quad (\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2) = (166.4, -13.95, -22.0),$$

where y is executive corruption (0–100), x_1 is high court independence (0–4), and x_2 is $\log_{10}$ GDP per capita. (a) Show that for any two units A and B ,

$$\hat{y}_A - \hat{y}_B = \hat{\beta}_1(x_{1A} - x_{1B}) + \hat{\beta}_2(x_{2A} - x_{2B}).$$

- Using $(x_1, x_2) = (2.96, 4.48)$ for Chile and $(0.21, 4.38)$ for China, compute the predicted levels and their difference. (c) Explain the “partial effect” meaning of $\hat{\beta}_1$ and $\hat{\beta}_2$.
3. (*Properties of the fitted values and residuals*) Show that
- $\sum_{i=1}^n \hat{u}_i = 0 \Rightarrow \bar{y} = \hat{\bar{y}}$.
 - The point $(\bar{y}, \bar{x}, \bar{z})$ is always on the regression plane.
 - $Cov(x, \hat{u}) = Cov(z, \hat{u}) = 0 \Rightarrow Cov(\hat{y}, \hat{u}) = 0$.
4. (*ANOVA Decomposition, R^2 , and Adjusted R^2*) Let $\hat{y}_i$ be fitted values from OLS with intercept and two regressors, and $\hat{u}_i = y_i - \hat{y}_i$. Prove that

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n \hat{u}_i^2 \quad (\text{SST} = \text{SSE} + \text{SSR}).$$

Then explain why $R^2 = \text{SSE}/\text{SST}$ is weakly increasing when a regressor is added.