Effects of Collinearity

Collinearity can be interpreted theoretically and empirically

Theoretical:

If $E[Y|X_1, X_2]$ linear in $X_1 X_2$, is $E(Y|X_1)$ linear in X_1?
It so, how are the linear relationships related?

Empirically:

How does including or excluding a variable change the OLS coef for another?

Theoretical collinearity:

Suppose $E[Y|X_1, X_2] = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2$

$$y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \epsilon \quad , \quad E[\epsilon | X_1, X_2] = 0$$

Suppose we fit $y \sim X_1$; i.e. $Y = \beta_0 + \beta_1 X_1 + \epsilon$. Is this incorrect?

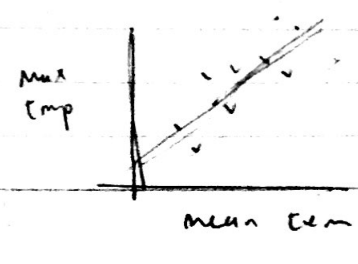
Case 0: $\alpha_2 = 0$, then $E[Y|X_1, X_2] = E[Y|X_1] = \alpha_0 + \alpha_1 X_1$: correct

Case 1: $\alpha_2 \neq 0$, X_2 is a deterministic experimental condition; incorrect

Case 2: X_2 is a randomly sampled quantity or random experimental condition : possibly correct

Suppose $E[X_2 | X_1] \approx \gamma_0 + \gamma_1 X_1$

Example: X_1 = Mean temperature
X_2 = Max temperature



In this case,

$$Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \epsilon$$

$$E[Y|X_1] = \alpha_0 + \alpha_1 X_1 + E[\alpha_2 X_2 | X_1] + E[\epsilon | X_1]$$

$$= \alpha_0 + \alpha_1 X_1 + \alpha_2 [\gamma_0 + \gamma_1 X_1] + 0$$

$$= (\alpha_0 + \alpha_2 \gamma_0) + (\alpha_1 + \alpha_2 \gamma_1) X_1$$

$$= \beta_0 + \beta_1 X_1$$

✓ Linear model is correct.

<u>Weather Example Revisited</u>:

X_1 = mean temp , assume true values α_1 small, neg.
X_2 = max temp α_2 large, positive
γ_1 large positive

then $\beta_1 = \underbrace{\alpha_1}_{\text{small} -} + \underbrace{\alpha_2 \gamma_1}_{\text{large} +}$ = large, positive.

$\Rightarrow$ true <u>marginal</u> <u>effect</u> of X_1 is large positive

true <u>conditional</u> <u>effect</u> of X_1 is small negative.

<u>Marginal</u>: Model is "marginalized" over possible values of X_2

$$E[Y|X_1] = \int E[Y|X_1, X_2]\, p(X_2|X_1)\, dX_2$$

<u>Conditional</u>: model is <u>conditional</u> on each value of X_2.

In this case, * Both models are correct!!

\# Their interpretations are different

\# "effects" of a variable depend on the context.

Thought question:

Suppose $E[Y|X_1, X_2] = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2$

$E[X_2|X_1] = \gamma_0 + \gamma_1 X_1$

$\Rightarrow E[Y|X_1] = (\alpha_0 + \alpha_2\gamma_0) + (\underline{\alpha_1} + \alpha_2\gamma_1) X_1$

$= \beta_0 + \beta_1 X_1$

Conditional effect of X_1 is $\underline{\alpha_1}$

What is range of marginal effect $\underline{\beta_1}$ as γ_1 ranges from $-\infty$, to ∞?

What is relationship between α_1, β_1 when $\gamma_1 = 0$?

Empirical collinearity: given data $\underset{\sim}{y}, \underset{\sim}{x}_1, \underset{\sim}{x}_2$

fit $y \sim x_1 + x_2 \Rightarrow \hat{y} = \hat{\alpha}_0 + \hat{\alpha}_1 x_1 + \hat{\alpha}_2 x_2$

$y \sim x_1 \Rightarrow \hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1$

Questions: what is relation between $\hat{\alpha}_1, \hat{\beta}_1$?
How does $\underset{\sim}{x}_2$ affect $\hat{\alpha}_1, \hat{\beta}_1$?

OLS Estimates: for convenience, assume centered $\underset{\sim}{x}_1, \underset{\sim}{x}_2$

SLR: $\hat{\beta}_1 = \frac{S_{X_1Y}}{S_{X_1X_1}} = \frac{\Sigma(x_{i1} - \bar{x}_1)(y_i - \bar{y})}{\Sigma(x_{i1} - \bar{x})^2} = \frac{\Sigma x_{i1}(y_i - \bar{y})}{\Sigma x_{i1}^2}$

$= \frac{\Sigma x_{i1} y_i}{\Sigma x_{i1}^2} = \frac{\underset{\sim}{x}_1 \cdot \underset{\sim}{y}}{\underset{\sim}{x} \cdot \underset{\sim}{x}}$

$= \frac{x_1^T y}{x_1^T x_1}$

MLR: $\hat{\underset{\sim}{\alpha}} = (X^T X)^{-1} X^T y$; $X = \begin{pmatrix} 1 & x_1 & x_2 \\ \vdots & | & | \\ 1 & & \end{pmatrix}$

$$X^T y = \begin{pmatrix} 1 \cdots 1 \\ x_1^T — \\ x_2^T — \end{pmatrix} \begin{pmatrix} y \\ | \end{pmatrix} = \begin{pmatrix} n\bar{y} \\ x_1^T y \\ x_2^T y \end{pmatrix}$$

$$X^T X = \begin{pmatrix} 1 — \\ x_1^T — \\ x_2^T — \end{pmatrix} \begin{pmatrix} 1 & x_1 & x_2 \\ | & | & | \end{pmatrix} = \begin{pmatrix} n & 0 & 0 \\ 0 & x_1^T x_1 & x_1^T x_2 \\ 0 & x_2^T x_1 & x_2^T x_2 \end{pmatrix}$$

$$(X^T X)^{-1} = \begin{bmatrix} 1/n & 0 & 0 \\ 0 & \boxed{} & \\ 0 & & \end{bmatrix} \to \begin{pmatrix} x_2^T x_2 & -x_1^T x_2 \\ -x_1^T x_2 & x_1^T x_1 \end{pmatrix} \times \frac{1}{|x_1|^2 |x_2|^2 - |x_1^T x_2|^2}$$

$$(X^T X)^{-1} X^T y = \begin{pmatrix} \hat{\alpha}_0 \\ \hat{\alpha}_1 \\ \hat{\alpha}_2 \end{pmatrix} \qquad \hat{\alpha}_0 = \frac{1}{n} n\bar{y} = \bar{y}$$

$$\hat{\alpha}_1 = \frac{1}{|x_1|^2 |x_2|^2 - |x_1^T x_2|^2} \times \left((x_2^T x_2)\, x_1^T y - (x_1^T x_2)\, x_2^T y \right)$$

$$= \frac{x_1^T y}{x_1^T x_1} \times \frac{(x_1^T x_1)(x_2^T x_2)}{(x_1^T x_1)(x_2^T x_2) - (x_1^T x_2)^2} - \frac{x_2^T y}{x_2^T x_2} \; \frac{(x_2^T x_2)(x_1^T x_2)}{(x_1^T x_1)(x_2^T x_2) - (x_1^T x_2)^2}$$

$$= \hat{\beta}_1 \times w_1 - \text{“}\hat{\beta}_2\text{''} \left(w_2 \right)$$

slope when only x_2 is in model

what happens when : ① $x_1^T x_2 = 0$
② $x_1^T x_2 > 0$
③ $x_1^T x_2 < 0$

Example: reconsider ... example.