Derivation of the OLS Estimator, $\hat{\boldsymbol{\beta}}$, in Matrix Form

What do we know?

Population equation: $y = X\beta + u$

Sample equation: $y = X\hat{\beta} + \hat{u}$

Regression assumptions, MLR.1 – MLR.6

$var - cov\ matrix\ of\ true\ errors = Var(u) = E(uu') = \sigma^2 I_n$

$u \sim Normal(0, \sigma^2 I_n)$

$y \sim Normal(X\beta, \sigma^2 I_n)$

$E(u) = 0\ and\ E(y) = X\beta$

Mission: Minimize the sum of squared residuals (SSR) with respect to $\hat{\beta}$. How do we write SSR in matrix form?

SSR = $\hat{u}'\ \hat{u}$

Thus, our OLS solution will the partial derivative of SSR with respect to $\hat{\beta}$

$$\frac{\partial \hat{u}'\ \hat{u}}{\partial \hat{\beta}}$$

First, rewrite $\hat{u}'\ \hat{u}$ in terms of y, x, and $\hat{\beta}$.

1. $\hat{u}'\ \hat{u} = (y - X\hat{\beta})'(y - X\hat{\beta})$
2. $= (y' - (X\hat{\beta})')(\mathrm{y} - \mathrm{X}\hat{\beta})$
3. $= (y' - \hat{\beta}'X')(\mathrm{y} - \mathrm{X}\hat{\beta})$
4. $= y'y - \hat{\beta}'X'\mathrm{y} - \mathrm{y}'\mathrm{X}\hat{\beta} + \hat{\beta}'\mathrm{X}'\mathrm{X}\hat{\beta}$ [the 2nd and 3rd products are scalars, so $\hat{\beta}'X'\mathrm{y} = \mathrm{y}'\mathrm{X}\hat{\beta}$]
5. $= y'y - (2y'X)\hat{\beta} + \hat{\beta}'(\mathrm{X}'\mathrm{X})\hat{\beta}$ [1st phrase is a constant w.r.t. $\hat{\beta}$; 2nd phrase is linear w.r.t. $\hat{\beta}$; 3rd is the quadratic w.r.t. $\hat{\beta}$]

Find partial derivative of function w.r.t. $\hat{\beta}$:

6. $\frac{\partial \hat{u}'\ \hat{u}}{\partial \hat{\beta}} = 0 - 2X'y + 2X'X\hat{\beta}$ [follows rules of matrix differentiation]

Set this function equal to 0 to find minimum of function; solve for $\hat{\beta}$:

7. $0 = -2X'y + 2X'X\hat{\beta}$
8. $2X'y = 2X'X\hat{\beta}$
9. $X'y = X'X\hat{\beta}$ [pre-multiply both sides by $(X'X)^{-1}$; see Matrix Rules handout]
10. $(X'X)^{-1}X'y = (X'X)^{-1}X'X\hat{\beta}$
11. $(X'X)^{-1}X'y = I\hat{\beta}$

This is the solution for the OLS estimator in matrix form:

12. $\hat{\beta} = (X'X)^{-1}X'y$

Proof that OLS Estimator is Unbiased

Unbiasedness means:

$E(\hat{\beta}) = \beta$

We need to prove that this statement is true, given what we know. Begin by rewriting $E(\hat{\beta})$:

1. $E(\hat{\beta}) = E[(X'X)^{-1}X'y]$
2. $= (X'X)^{-1}X'E(y)$ [Expectation operator moves through non-random variables]
3. $= (X'X)^{-1}X'X\beta$ [We already know that E(y) = $X\beta$]
4. $E(\hat{\beta}) = I\beta$
5. $E(\hat{\beta}) = \beta$