$Y = X_1\beta_1 + \varepsilon$ all the regressors are in X_1

Hence $\hat{\beta}_{1_{OLS}} = (X_1'X_1)^{-1}X_1'Y$

the impact of one unit increase in X_1 on Y

$Y = X_1\beta_1 + X_2\beta_2 + \varepsilon$

X_2 → controls

How does β_1 changes

$\hat{\beta}_1 = (X_1' M_{X_2} X_1)^{-1} X_1' M_{X_2} Y$

↓

estimated impact of the increase of β_1 ceteris paribus

When adding controls does not have any impact on β_1

↓

probably when the control variable is a linear combination of the column of X

↓

$X_{k+1} \in$ column space (X)

Influential observation

↳ has the power to change the slope of the regression line significantly

↳ similar to an outlier, but outliers have statistical significance

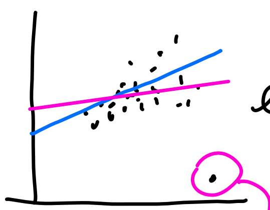
Leave-one-out regression

↓

we leave one obs. out and see what happens to the regression line

↓

to understand if the result drastically depends on some observation



the residual is very large and the value of X is very high

↓

it is far right and low in the graph

this observation has the power to change significantly the β

$e^{(i)}$ is the extraction vector, is a vector made of zeros except a 1 in the ith position

$$\underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{N \times 1}$$

$e^{(i)}X = x_i$ $(1 \times k)$

↓

the ith row of the matrix X

$e^{(i)\prime}Y = Y_i$ → i obtain the ith Y applying $e^{(i)}$

Y_i is a scalar, x_i is a row vector

$$M_e = I_N - e\underbrace{(e'e)}_{\text{is } 1}^{-1}e' = I_N - ee'$$

→ Hence this is the Identity matrix with a zero on (i,i) on the main diagonal

ee': $N \times N$ matrix full of zeros except a 1 on (i,i), which is on the main diagonal

$$\begin{bmatrix} 1 & 0 & \cdots & \\ 0 & 1 & & \\ 0 & & 0 & \\ 0 & & & \ddots \end{bmatrix}$$

of the regression

$$y = e^{(i)}y + \varepsilon$$

$$M_e y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ 0 \\ y_{i+1} \\ \vdots \\ y_N \end{bmatrix}$$ → zero vector in i-th position

so $e^{(i)\prime}y$ is y_i (scalar)

$Xe^{(i)}$ is x_i (vector)

$M_e y$ is $\begin{bmatrix} y_1 \\ y_2 \\ 0 \\ \vdots \\ y_N \end{bmatrix}$

$$y = X\beta + \varepsilon$$

is the central regression of interest

$$y = X\beta + \alpha e^{(i)} + \varepsilon$$

when $\alpha = 0$, the two models are the same, there is no influence in excluding the observation

I want to capture with α the variation of y not explained by X but by $e^{(i)}$

$$Z = [X, e]$$

$N \times (k+1)$

Horizontal concatenation of X and e

$y = P_Z y + M_Z y$ by construction this is y

$\text{rank}(Z) = k+1$, has full rank

$$y = \underbrace{X\hat{\beta} + \hat{\alpha} e}_{P_Z y} + \overbrace{M_Z y}^{\hat{\varepsilon}}$$

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

while the other $\hat{\beta}$ is different and we call it $\hat{\beta}^{(i)}$

we are partialling out the impact of the ith information

$$P_x y = P_x X\hat{\beta}^{(i)} + \hat{\alpha} P_x e + P_x M_x y \quad \leftarrow \quad y = X\beta + \alpha e + \varepsilon$$

multiplying both members by P_x

$$X\hat{\beta} = X\hat{\beta}^{(i)} + \hat{\alpha} P_x e + 0$$

→ since X belongs to ? and P_x is orthogonal to M_x

without obs. — with observation

$$\underbrace{X\left(\hat{\beta}^{(i)} - \hat{\beta}\right)}_{\text{How much the regression line changes when we add the ith observation}} = -\hat{\alpha} P_x e$$

What $\hat{\alpha}$ is?

Obtained by the FWL th.

$$y = X\beta + \alpha e^{(i)} + \varepsilon$$

$$\hat{\alpha} = \underbrace{(e' M_x e)^{-1}}\underbrace{e' M_x y}$$

both scalars

we treat X as the controls

FWL

$\hat{\alpha}$ is a scalar

$$\boxed{\frac{e' M_x y}{e' M_x e}}$$

→ this is a ratio, the larger $\hat{\alpha}$, the more the impact on the regression of the ith obs.

vector of residuals

$$\frac{e'\hat{\varepsilon}}{e' M_x e}, \quad e'\hat{\varepsilon} = \hat{\varepsilon}^{(i)}$$

→ just returns the i-th residual $\hat{\varepsilon}^{(i)}$

$$e'(I_N - P_x)e$$

→ leveraged points

$e' I_N e$ is just one

$e' P_x e$ → h_{ii}, element in position (i,i) in the projection matrix

$$\boxed{\frac{\hat{\varepsilon}^{(i)}}{1 - h_{ii}}}$$

$$[0 \cdots 1 \cdots 0]\begin{bmatrix} \vdots \\ \vdots \end{bmatrix}_{N\times N}$$

$0 < h_{ii} < 1$

this property has to do with the property that the not eigenvalue of P_X is one

$h_{ii} \geq \frac{1}{N}$ when the intercept is included in the model

$\bar{h} = \frac{K}{N}$

rank of P_X is K, and trace = K

this property has to do with intersects

datapoints with high leverage are those with h_{ii} closer to one → a also great relative to the other ones

this makes the ratio very high

leverage because they have the potential to generate a significant change in the slope

potentially, because it also depends by how large the numerator is

$$X(\hat{\beta}^{(i)} - \hat{\beta}) = -\frac{\hat{\varepsilon}_i}{1-h_{ii}} P_X e$$

one has to look at $\frac{\hat{\varepsilon}^{(i)}}{1-h_{ii}}$

$$\cancel{(X'X)^{-1}X'}X(\hat{\beta}^{(i)} - \hat{\beta}) = -\frac{\hat{\varepsilon}_i}{1-h_{ii}} \cancel{P_X}(X'X)^{-1}X'e$$

$$(\hat{\beta}^{(i)} - \hat{\beta}) = -\frac{\hat{\varepsilon}_i}{1-h_{ii}} x_i(X'X)^{-1}x_i'$$

ith row transposed as a column

i can multiply this by x_i to obtain a scalar and not a vector

$x_i = e'X$

$x_i' = X'e$

$e'P_X e$

$$-\frac{\hat{\varepsilon}_i h_{ii}}{1-h_{ii}}$$

$\hat{\beta}_{OLS}$ is an estimator of the unknown β.

- is the estimator unbiased?

$\hat{\theta}$ is an estimator of θ

if $E[\hat{\theta}] = \theta$ then $\hat{\theta}$ is an unbiased estimator of θ

$$E[\hat{\beta} | X] = \beta$$ i want to verify this condition to say that $\hat{\beta}$ is conditionally unbiased

$$E\left[(X'X)^{-1}X'y \mid X\right] = (X'X)^{-1}X' \underbrace{E[y|X]}_{X\beta^0}$$

$$(X'X)^{-1}X'X\beta^0 = \beta^0$$

Hence
$$E[\hat{\beta}|X] = \beta^0$$

or $E[y|X]$

$$= E[X\beta^0 + \varepsilon | X]$$

$$= E[X\beta|X] + \underbrace{E[\varepsilon|X]}_{\text{exogeneity assumption}}$$

$$= X\beta^0 + 0$$

Hence we have the same result as before

$$(\hat{\beta} - \hat{\beta}^{(i)}) = \frac{\hat{\varepsilon}_i h_{ii}}{1 - h_{ii}}$$

$$= \beta_M + (\beta_{MS} - \beta_M)$$

$$\beta_{MS} = \beta_M + (\underbrace{\beta_{MS}}_{\hat{\beta}_{MS}} - \underbrace{\beta_M}_{\beta^{(i)}})$$

$$x\underbrace{(\hat{\beta}^{(i)} - \hat{\beta})}_{\beta_M} = -\frac{\hat{\varepsilon}_i}{1-h_{ii}} P_x e$$

-1_s

$$(\hat{\beta} - \beta^{(i)}) = \frac{\hat{\varepsilon}_i P_x e}{1-h_{ii}}$$

$$x(\hat{\beta} - \beta^{(i)}) \frac{\quad}{h_{ii}-1}$$

$$(\hat{\beta} - \hat{\beta}^{(i)}) = \frac{\hat{\varepsilon}_i h_{ii}}{1-h_{ii}}$$

$$\frac{(y_s - x_s'\beta_M)\, e' P_x e}{1 - e' P_x e} = \frac{(y_s - x_s'\beta_M)\, \underbrace{e'X}_{x_s}(X'X)^{-1}\underbrace{X'e}_{x_s'}}{1 - e'X(X'X)^{-1}X'e}$$

$$\frac{(y_s - x_s'\beta_M)\, x_s (X'X)^{-1} x_s'}{1 - x_s (X'X)^{-1} x_s'}$$