The AVP Theorem

Theorem: *(AVP or Frisch-Waugh-Lovell)* Consider the regression of Y on two blocks of predictors, X_1 and X_2, where the partitioned matrix, $[X_1 X_2]$ is of full column rank. Suppose

$$Y = X_1\hat{\beta}_1 + X_2\hat{\beta}_2 + e, \text{ with } e \perp \text{span}(X_1, X_2)$$

Then the (residual of Y regressed on X_2) regressed on the (residual of X_1 on X_2) has

1. least-squares regression coefficient $\hat{\beta}_1$,
2. least-squares residual equal to e, and
3. $SSE = e'e$.

$$+4. \quad Var_{\underbrace{AVP}_{\text{from AVP}}}(\hat{\underset{\sim}{\beta}}_{X_1}) = Var_{\underbrace{MR}_{\text{from multiple regression}}}(\hat{\underset{\sim}{\beta}}_{X_1})$$

Proof:

Proof: The residual of Y in the regression on X_2 is obtained by pre-multiplying Y by

$$Q_2 = I - P_2 = I - X_2(X_2'X_2)^{-1}X_2'$$

and similarly for X_1. We obtain

$$Q_2 Y = Q_2 X_1 \hat{\beta}_1 + Q_2 X_2 \hat{\beta}_2 + Q_2 e$$

We are just premultiplying

$\underset{\sim}{Y} = X_1 \underset{\sim}{\hat{\beta}}_1 + X_2 \underset{\sim}{\hat{\beta}}_2 + \underset{\sim}{e}$ by Q_2

So what happens?

$Q_2 X_2 \overset{?}{=}$

$Q_2 \underset{\sim}{e} \overset{?}{=}$

So we get $\boxed{Q_2 \underset{\sim}{Y}} = \boxed{Q_2 X_1} \underset{\sim}{\hat{\beta}}_1 + \underset{\sim}{e}$

So if $\underset{\sim}{e} \perp Q_2X_1$, then $\hat{\beta}_1$ is the L-S reg. coef. by the projection theorem.

Now $(Q_2X_1)'\underset{\sim}{e} = X_1'Q_2'\underset{\sim}{e}$

$= X_1'Q_2\underset{\sim}{e}$ since $Q_2 = Q_2'$

$= X_1'\underset{\sim}{e}$ since $\underset{\sim}{e} \perp X_2$

$= \underset{\sim}{0}$ since $\underset{\sim}{e}$ is also $\perp X_1$

(it was the residual from regressing on X_1 and X_2)

$\therefore \hat{\underset{\sim}{\beta}}_1$ is the L-S reg. coef. regressing $Q_2\underset{\sim}{Y}$ on Q_2X_1

now for the Variance :

$$\begin{aligned}\mathrm{Var}_{AVP}(\hat{\beta}_1) &= \sigma^2((Q_2X_1)'(Q_2X_1))^{-1}\\ &= \sigma^2(X_1'Q_2X_1)^{-1}\end{aligned}$$

and, using the Schur complement for the inverse of a partitioned matrix, we get:

$$\mathrm{Var}_{MR}(\hat{\beta}) = \sigma^2\begin{pmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2\end{pmatrix}^{-1}$$

and

$$\begin{aligned}\mathrm{Var}_{MR}(\hat{\beta}_1) &= \sigma^2(X_1'X_1 - X_1'X_2(X_2'X_2)^{-1}X_2'X_1)^{-1}\\ &= \sigma^2(X_1'(I - X_2(X_2'X_2)^{-1}X_2')X_1)^{-1}\\ &= \sigma^2(X_1'Q_2X_1)^{-1}\\ &= \mathrm{Var}_{AVP}(\hat{\beta}_1)\end{aligned}$$

Corollary: Regressing Y on X & Z_1

and " " " " & Z_2

produces the same $\hat{\beta}_1$ if $\text{span}(Z_1) = \text{span}(Z_2)$