Calculations of $\mathbf{b}_{LS}$ without (much) Matrix Algebra

$$\underset{\hat{\alpha},\hat{\beta}_1,\ldots,\hat{\beta}_k}{Min} \sum_i e_i^2 \equiv \underset{\hat{\boldsymbol{\beta}}}{Min}\ \underset{\sim}{\mathbf{e}}'\underset{\sim}{\mathbf{e}} \equiv \underset{\hat{\boldsymbol{\beta}}}{Min}\ \left(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}\right)'\left(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}\right)$$

$$\Rightarrow \frac{\partial \underset{\sim}{\mathbf{e}}'\underset{\sim}{\mathbf{e}}}{\partial \hat{\boldsymbol{\beta}}} = 0 \Rightarrow \mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}^* = \mathbf{X}'\mathbf{y}$$

now, notice that, for **X** 3x3, for example, we have:

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} \underset{\sim}{\mathbf{x}}_1'\underset{\sim}{\mathbf{x}}_1 & \underset{\sim}{\mathbf{x}}_1'\underset{\sim}{\mathbf{x}}_2 & \underset{\sim}{\mathbf{x}}_1'\underset{\sim}{\mathbf{x}}_3 \\ \underset{\sim}{\mathbf{x}}_2'\underset{\sim}{\mathbf{x}}_1 & \underset{\sim}{\mathbf{x}}_2'\underset{\sim}{\mathbf{x}}_2 & \underset{\sim}{\mathbf{x}}_2'\underset{\sim}{\mathbf{x}}_3 \\ \underset{\sim}{\mathbf{x}}_3'\underset{\sim}{\mathbf{x}}_1 & \underset{\sim}{\mathbf{x}}_3'\underset{\sim}{\mathbf{x}}_2 & \underset{\sim}{\mathbf{x}}_3'\underset{\sim}{\mathbf{x}}_3 \end{bmatrix},$$ so, we can rewrite:

$$\begin{bmatrix} \underset{\sim}{\mathbf{x}}_1'\underset{\sim}{\mathbf{x}}_1 & \underset{\sim}{\mathbf{x}}_1'\underset{\sim}{\mathbf{x}}_2 & \underset{\sim}{\mathbf{x}}_1'\underset{\sim}{\mathbf{x}}_3 \\ \underset{\sim}{\mathbf{x}}_2'\underset{\sim}{\mathbf{x}}_1 & \underset{\sim}{\mathbf{x}}_2'\underset{\sim}{\mathbf{x}}_2 & \underset{\sim}{\mathbf{x}}_2'\underset{\sim}{\mathbf{x}}_3 \\ \underset{\sim}{\mathbf{x}}_3'\underset{\sim}{\mathbf{x}}_1 & \underset{\sim}{\mathbf{x}}_3'\underset{\sim}{\mathbf{x}}_2 & \underset{\sim}{\mathbf{x}}_3'\underset{\sim}{\mathbf{x}}_3 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1^* \\ \hat{\beta}_2^* \\ \hat{\beta}_3^* \end{bmatrix} = \begin{bmatrix} \underset{\sim}{\mathbf{x}}_1'\mathbf{y} \\ \underset{\sim}{\mathbf{x}}_2'\mathbf{y} \\ \underset{\sim}{\mathbf{x}}_3'\mathbf{y} \end{bmatrix},$$ which in scalar notation is:

where superscript * means these are the optimums.

(1) $$\mathbf{x}_1'\mathbf{x}_1\hat{\beta}_1^* + \mathbf{x}_1'\mathbf{x}_2\hat{\beta}_2^* + \mathbf{x}_1'\mathbf{x}_3\hat{\beta}_3^* = \mathbf{x}_1'\mathbf{y}$$

(2) $$\mathbf{x}_2'\mathbf{x}_1\hat{\beta}_1^* + \mathbf{x}_2'\mathbf{x}_2\hat{\beta}_2^* + \mathbf{x}_2'\mathbf{x}_3\hat{\beta}_3^* = \mathbf{x}_2'\mathbf{y}$$

(3) $$\mathbf{x}_3'\mathbf{x}_1\hat{\beta}_1^* + \mathbf{x}_3'\mathbf{x}_2\hat{\beta}_2^* + \mathbf{x}_3'\mathbf{x}_3\hat{\beta}_3^* = \mathbf{x}_3'\mathbf{y}$$

Now, usually first column (variable) is $\underset{\sim}{\mathbf{x}}_1 = [\cdots \quad 1 \quad \cdots]$, which implies:

$\underset{\sim}{\mathbf{x}}_1'\underset{\sim}{\mathbf{x}}_1 = \sum_{i=1}^{n}(1\times 1) = n$, and $\underset{\sim}{\mathbf{x}}_1'\underset{\sim}{\mathbf{x}}_2 = \sum_{i=1}^{n}\left(1\times \underset{\sim}{\mathbf{x}}_{2_i}\right) = \sum_{i=1}^{n} x_2$ and $\underset{\sim}{\mathbf{x}}_1'\underset{\sim}{\mathbf{x}}_3 = \sum_{i=1}^{n}\left(1\times \underset{\sim}{\mathbf{x}}_{3_i}\right) = \sum_{i=1}^{n} x_3$

Using this, we can simplify (1),(2),(3) as…

(4) $$n\hat{\beta}_1^* + \hat{\beta}_2^*\sum x_2 + \hat{\beta}_3^*\sum x_3 = \sum y$$

(5) $$\hat{\beta}_1^*\sum x_2 + \mathbf{x}_2'\mathbf{x}_2\hat{\beta}_2^* + \mathbf{x}_2'\mathbf{x}_3\hat{\beta}_3^* = \mathbf{x}_2'\mathbf{y}$$

(6) $$\hat{\beta}_1^*\sum x_3 + \mathbf{x}_3'\mathbf{x}_2\hat{\beta}_2^* + \mathbf{x}_3'\mathbf{x}_3\hat{\beta}_3^* = \mathbf{x}_3'\mathbf{y}$$

Now, let's finish converting to scalar notation:

(7) $$n\hat{\beta}_1^* + \hat{\beta}_2^*\sum x_2 + \hat{\beta}_3^*\sum x_3 = \sum y$$

(8) $$\hat{\beta}_1^*\sum x_2 + \hat{\beta}_2^*\sum x_2^2 + \hat{\beta}_3^*\sum x_2 x_3 = \sum x_2 y$$

(9) $$\hat{\beta}_1^*\sum x_3 + \hat{\beta}_2^*\sum x_2 x_3 + \hat{\beta}_3^*\sum x_3^2 = \sum x_3 y$$

Now, divide (7) through by n to get the solution for $\hat{\beta}_1$:

(10) $$\hat{\beta}_1^* = \bar{y} - \hat{\beta}_2^*\bar{x}_2 - \hat{\beta}_3^*\bar{x}_3$$

UPSHOTS:

The intercept, $\hat{\beta}_1^*$*, is the conditional (sample) mean of* y.

The regression line passes through the (multidimensional) mean of the data (which is called the centroid *of the data).*

Now, substitute (10) into (8) or (9), let's use (8). ⇒ (11)

substituting:

$$\left(\bar{y}-\hat{\beta}_2^*\bar{x}_2-\hat{\beta}_3^*\bar{x}_3\right)\sum x_2+\hat{\beta}_2^*\sum x_2^2+\hat{\beta}_3^*\sum x_2x_3=\sum x_2y$$

$$\text{(multiplying)}\Rightarrow\left(\bar{y}\sum x_2-\hat{\beta}_2^*\bar{x}_2\sum x_2-\hat{\beta}_3^*\bar{x}_3\sum x_2\right)+\hat{\beta}_2^*\sum x_2^2+\hat{\beta}_3^*\sum x_2x_3=\sum x_2y$$

$$\left(\text{since }\sum x_2=n\bar{x}_2:\right)\Rightarrow\left(\bar{y}n\bar{x}_2-\hat{\beta}_2^*\bar{x}_2n\bar{x}_2-\hat{\beta}_3^*\bar{x}_3n\bar{x}_2\right)+\hat{\beta}_2^*\sum x_2^2+\hat{\beta}_3^*\sum x_2x_3=\sum x_2y$$

$$\text{(rearranging)}\Rightarrow\hat{\beta}_2^*\left(\sum x_2^2-n\bar{x}_2^2\right)+\hat{\beta}_3^*\left(\sum x_2x_3-n\bar{x}_3\bar{x}_2\right)=\sum x_2y\text{-}n\bar{x}_2\bar{y}$$

$$\Rightarrow\hat{\beta}_2^*\times S_{x_2}^2+\hat{\beta}_3^*\times S_{x_2x_3}=S_{x_2y}$$

...by definitions of sample variations & covariations

$$\text{(solving for }\hat{\beta}_2^*)\Rightarrow\hat{\beta}_2^*=\frac{S_{x_2y}-\hat{\beta}_3^*\times S_{x_2x_3}}{S_{x_2}^2}$$

Which can also be rewritten (note the change in subscripting notation!):

$$\Rightarrow\hat{\beta}_{2.3}^*=\frac{S_{x_2y}}{S_{x_2}^2}-\frac{\hat{\beta}_{3.2}\times S_{x_2x_3}}{S_{x_2}^2}=\hat{\beta}_2-\hat{\beta}_{3.2}^*\frac{S_{x_2x_3}}{S_{x_2}^2}=\hat{\beta}_2-\hat{\beta}_{3.2}^*\times\hat{\beta}_{x_3x_2}$$

And, by analogy: (12) $\hat{\beta}_{3.2}^*=\hat{\beta}_3-\hat{\beta}_{2.3}^*\times\hat{\beta}_{x_2x_3}$.

The notation ${}^*_{2.3}$, etc., serves to underscore that these are the optimal estimates of the conditional or partial coefficients on x_2 from a regression including, i.e., controlling for, x_3, and *vice versa.* The notation ${}_{x_2x_3}$, etc., means from a regression of x_2 on x_3. Finally, the lack of a * on the bivariate coefficients is to stress that they are from bivariate (unconditional) regression equations, not from this multivariate regression.

Working now from the more-elaborate expressions (1)-(3) from your original notes:

(1) $$b_1^* = \overline{y} - b_2^* \overline{x}_2 - b_3^* \overline{x}_3$$

(2) $$b_2^* = \frac{\left(\Sigma x_2^* y^*\right)\left(\Sigma (x_3^*)^2\right) - \left(\Sigma x_3^* y^*\right)\left(\Sigma x_2^* x_3^*\right)}{\left(\Sigma (x_2^*)^2\right)\left(\Sigma (x_3^*)^2\right) - \left(\Sigma x_2^* x_3^*\right)^2} = \frac{S_{x_2 y} S_{x_3}^2 - S_{x_3 y} S_{x_2 x_3}}{S_{x_2}^2 S_{x_3}^2 - \left(S_{x_2 x_3}\right)^2}$$

(3) $$b_3^* = \frac{\left(\Sigma x_3^* y^*\right)\left(\Sigma (x_2^*)^2\right) - \left(\Sigma x_2^* y^*\right)\left(\Sigma x_2^* x_3^*\right)}{\left(\Sigma (x_3^*)^2\right)\left(\Sigma (x_2^*)^2\right) - \left(\Sigma x_2^* x_3^*\right)^2} = \frac{S_{x_3 y} S_{x_2}^2 - S_{x_2 y} S_{x_2 x_3}}{S_{x_2}^2 S_{x_3}^2 - \left(S_{x_2 x_3}\right)^2}$$

First, the notational shift here from $\hat{\beta}$ to b signifies nothing; it serves here only to match the notation from the original notes for the versions of the expressions for the same quantities from which we are beginning.

To get to these expressions from (11) and (12), you must solve out the right-hand-side $\hat{\beta}$'s in those expressions.

Doing so, you get:

$$\hat{\beta}^*_{2.3} = \frac{S_{x_2y}}{S^2_{x_2}} - \frac{\hat{\beta}^*_{3.2} \times S_{x_3x_2}}{S^2_{x_2}} = \frac{S_{x_2y}}{S^2_{x_2}} - \frac{\left(\hat{\beta}_3 - \hat{\beta}^*_{2.3} \times \hat{\beta}_{x_2x_3}\right) \times S_{x_3x_2}}{S^2_{x_2}}$$

$$\Rightarrow S^2_{x_2} \times \hat{\beta}^*_{2.3} = S_{x_2y} - \left(\hat{\beta}_3 - \hat{\beta}^*_{2.3} \times \hat{\beta}_{x_2x_3}\right) \times S_{x_3x_2}$$

$$\Rightarrow S^2_{x_2} \times \hat{\beta}^*_{2.3} = S_{x_2y} - \left(\frac{S_{x_3y}}{S^2_{x_3}} - \hat{\beta}^*_{2.3} \times \frac{S_{x_2x_3}}{S^2_{x_3}}\right) \times S_{x_3x_2}$$

$$\Rightarrow S^2_{x_2} \times \hat{\beta}^*_{2.3} = S_{x_2y} - \left(\frac{S_{x_3y} S_{x_3x_2}}{S^2_{x_3}} - \hat{\beta}^*_{2.3} \times \frac{\left(S_{x_2x_3}\right)\left(S_{x_3x_2}\right)}{S^2_{x_3}}\right)$$

$$\Rightarrow S^2_{x_3} S^2_{x_2} \times \hat{\beta}^*_{2.3} = S^2_{x_3} S_{x_2y} - S_{x_3y} S_{x_2x_3} + \hat{\beta}^*_{2.3} \times \left(S_{x_2x_3}\right)^2$$

$$\Rightarrow \left\{S^2_{x_3} S^2_{x_2} - \left(S_{x_2x_3}\right)^2\right\} \times \hat{\beta}^*_{2.3} = S^2_{x_3} S_{x_2y} - S_{x_3y} S_{x_2x_3}$$

$$\Rightarrow \hat{\beta}^*_{2.3} = \frac{S^2_{x_3} S_{x_2y} - S_{x_3y} S_{x_2x_3}}{S^2_{x_3} S^2_{x_2} - \left(S_{x_2x_3}\right)^2}$$

Which was (2) as in the original notes. Now, recall:

$$b_2 = \frac{Cov(X_2, Y)}{Var(X_2)} = \frac{S_{x_2y}}{S^2_{x_2}}$$

So, ÷ numerator & denom of (2) & (3) by $S_{x_2}^2$ and $S_{x_3}^2$:

$$\hat{\beta}_{2.3}^* = \frac{S_{x_3}^2 S_{x_2 y} \big/ S_{x_3}^2 S_{x_2}^2 - S_{x_3 y} S_{x_2 x_3} \big/ S_{x_3}^2 S_{x_2}^2}{S_{x_2}^2 S_{x_3}^2 \big/ S_{x_3}^2 S_{x_2}^2 - \left(S_{x_2 x_3}\right)^2 \big/ S_{x_3}^2 S_{x_2}^2} = \frac{S_{x_2 y} \big/ S_{x_2}^2 - S_{x_3 y} S_{x_2 x_3} \big/ S_{x_3}^2 S_{x_2}^2}{1 - r_{x_2 x_3}^2}$$

$$\left.\begin{aligned} (2')\quad b_{2.3}^* &= \frac{b_2 - b_3 b_{x_3 x_2}}{1 - r_{x_2 x_3}^2} \\ (3')\quad b_{3.2}^* &= \frac{b_3 - b_2 b_{x_2 x_3}}{1 - r_{x_2 x_3}^2} \end{aligned}\right\} \text{ where } \begin{cases} b_{2.3} \text{ is coeff reg } y \text{ on } x_2 \text{ in presence } x_3; \\ b_{x_2 x_3} \text{ is coeff reg } x_2 \text{ on } x_3; \\ \text{and } b_2 \text{ is bivariate reg coeff}. \end{cases}$$

(We return to the original *b* notation here at the end to reflect from where we began these calculations.)

<u>UPSHOT</u>: whether by (11) & (12) or by (2′) & (3′), the multivariate, conditional coefficients equal the bivariate coefficients if either (i) x_2 ***&*** x_3 ***do not covary or (ii)*** x_3 ***does not affect*** y ***.***

You might notice here, though, that the expressions for $\hat{\beta}_{2.3}^*$ & $b_{2.3}^*$ ***<u>seem</u>*** to differ & contradict, even though they are the same entity: the partial coefficient on x_2:

$$b_{2.3}^* = \frac{b_2 - b_3 b_{x_3 x_2}}{1 - r_{x_2 x_3}^2} \;\; vs. \;\; \hat{\beta}_{2.3}^* = \hat{\beta}_2 - \hat{\beta}_{3.2}^* \times \hat{\beta}_{x_3 x_2}$$

Latter says multivariate, partial coefficient equals the bivariate (unconditional) coefficient if either (*i*) x_2 & x_3 do not covary or (*ii*) x_3 does not affect y. The former seems to agree regarding (*i*) but not (*ii*). Even if x_3 does not affect y, non-zero $r^2_{x_2x_3}$ could still imply $b^*_{2.3} \neq b_2$. The seeming contradiction, though, is actually not one. Note the different second (numerator) term, b_3 *vs.* $\hat{\beta}^*_{3.2}$; the different denominators, $1 - r^2_{x_2x_3}$ *vs.* 1, exactly cancel this difference in the numerators.

So, what the apparent difference in the expressions does is ***clarify and underscore how we should understand the verbal descriptions of the conditions under which unconditional and partial coefficients equate***. Namely:

(*i*) x_2 & x_3 do not covary

means they do not covary, <u>unconditionally</u>; whereas:

(*ii*) x_3 does not affect y

means x_3 does not covary with y, <u>conditionally</u>; i.e., the partial coefficient on x_3 is (would be) zero; i.e., x_3 does (would) not matter, controlling for x_2.

Showing the equivalence of the two expressions:

$$b^*_{2.3} = \frac{b_2 - b_3 b_{x_3 x_2}}{1 - r^2_{x_2 x_3}} = \hat{\beta}^*_{2.3} = \hat{\beta}_2 - \hat{\beta}^*_{3.2} \times \hat{\beta}_{x_3 x_2}$$

...first unify the notation:

$$\hat{\beta}^*_{2.3} = \frac{\hat{\beta}_2 - \hat{\beta}_3 \hat{\beta}_{x_3 x_2}}{1 - r^2_{x_2 x_3}} = \hat{\beta}_2 - \hat{\beta}^*_{3.2} \times \hat{\beta}_{x_3 x_2}$$

...then substitute analogous formula for $\hat{\beta}^*_{3.2}$:

$$\hat{\beta}^*_{2.3} = \hat{\beta}_2 - \left(\hat{\beta}_3 - \hat{\beta}^*_{2.3} \times \hat{\beta}_{x_2 x_3} \right) \times \hat{\beta}_{x_3 x_2}$$

...multiply through:

$$\hat{\beta}^*_{2.3} = \hat{\beta}_2 - \hat{\beta}_3 \times \hat{\beta}_{x_3 x_2} + \hat{\beta}^*_{2.3} \times \hat{\beta}_{x_2 x_3} \times \hat{\beta}_{x_3 x_2}$$

...gather terms:

$$\hat{\beta}^*_{2.3} \left(1 - \hat{\beta}_{x_2 x_3} \times \hat{\beta}_{x_3 x_2} \right) = \hat{\beta}_2 - \hat{\beta}_3 \times \hat{\beta}_{x_3 x_2}$$

...solve & simplify:

$$\hat{\beta}^*_{2.3} = \frac{\hat{\beta}_2 - \hat{\beta}_3 \times \hat{\beta}_{x_3 x_2}}{1 - \frac{Cov(x_2, x_3)}{Var(x_3)} \times \frac{Cov(x_3, x_2)}{Var(x_2)}} = \frac{\hat{\beta}_2 - \hat{\beta}_3 \times \hat{\beta}_{x_3 x_2}}{1 - \frac{\left[Cov(x_2, x_3)\right]^2}{V(x_3) V(x_2)}} = \frac{\hat{\beta}_2 - \hat{\beta}_3 \times \hat{\beta}_{x_3 x_2}}{1 - r^2_{x_2, x_3}}$$