MULTIPLE REGRESSION AND PATH ANALYSIS

Introduction

Path analysis and multiple regression go hand in hand (almost). Also, it is easier to learn about multivariate regression using path analysis than using algebra. We will start with an intuitive approach and later develop the algebraic notation.

Consider the following SAS statements on the INTEREST data:

```
LIBNAME carey '~carey/p7291dir';
TITLE path Analysis and Multivariate Multiple Regression;
PROC REG DATA=carey.interest CORR;
     VAR lawyer architct educ vocab geometry;
     MODEL lawyer architct = educ vocab geometry / STB;
     MTEST / PRINT;
```

This performs a multiple regression on two dependent variables, vocational interest in becoming a lawyer (LAWYER) and vocational interest in becoming an (ARCHITCT). The independent variables are education (EDUC) and two tests of cognitive ability, vocabulary (VOCAB) and geometry (GEOMETRY). (Technically, it would be preferable to include other demographics such as gender and age in this analysis, but these variables are ignored here to keep matters simple.) The SAS program used for this example may be found on

~carey/p7291dir/pathreg1.sas

and the output, which is attached to this handout, may be found on

~carey/p7291dir/pathreg1.out

The coefficients for path analysis may be expressed in either of two metrics. The first metric is called *unstandardized*, and it uses the measurement scale of the original variables. Here, paths are unstandardized regression coefficients, covariances link the independent variables, and the purpose is to explain variance and covariance. The second metric is called *standardized*. Literally, this is the result of a path analysis or regression performed on all variables that have been transformed into standardized variables (i.e., with means of 0 and standard deviations of 1.0). In standardized units, the path coefficients equal the standardized regression coefficients (i.e., the weights), and the purpose is to explain the proportions of variance and the correlations among variables. The following gives path analysis information using standardized units.

To construct a path diagram, we require two pieces of information. The first piece is the correlation matrix among the variables. This may be obtained by performing PROC CORR on the variables or by specifying the CORR option on the PROC REG command. The second piece is the vector of standardized regression coefficients. We got this by specifying the STB (for STandardized Beta) on the MODEL subcommand for PROC REG.

Begin the path analysis by writing down the independent variables and connecting each

pair with a double headed arrow. Write the dependent variable. From each independent variable, draw a straight, single headed arrow shooting into the dependent variable. Finally, make a notation for a residual variable and draw an arrow from it into the dependent variable. Figure 1 shows this path diagram for LAWYER. The residual in Figure 1 is denoted as U_L.

Figure 1. Setting up a path diagram for multiple regression.
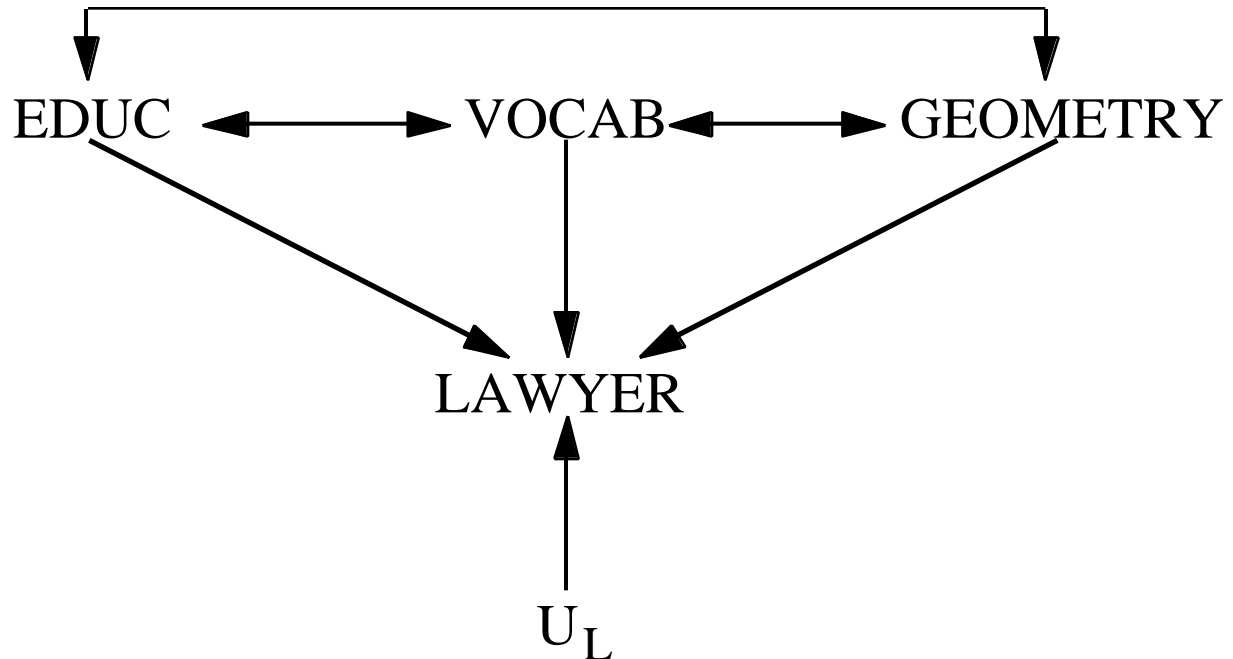
The second step is to place numbers on the arrows. On the double headed arrows write down the correlations between the independent variables. For example, the correlation between EDUC and VOCAB is .5182, so that number is written on the double headed arrow between EDUC and VOCAB. On the straight arrows, place the *standardized* (not the unstandardized) regression coefficients. These standardized regression coefficients are referred to as *path coefficients*. Finally, take the square root of $(1 - R^2)$ and place this value on the arrow going from the residual to the dependent variable. Performing all these operations gives Figure 2.

Figure 2. Path model for variable LAWYER.
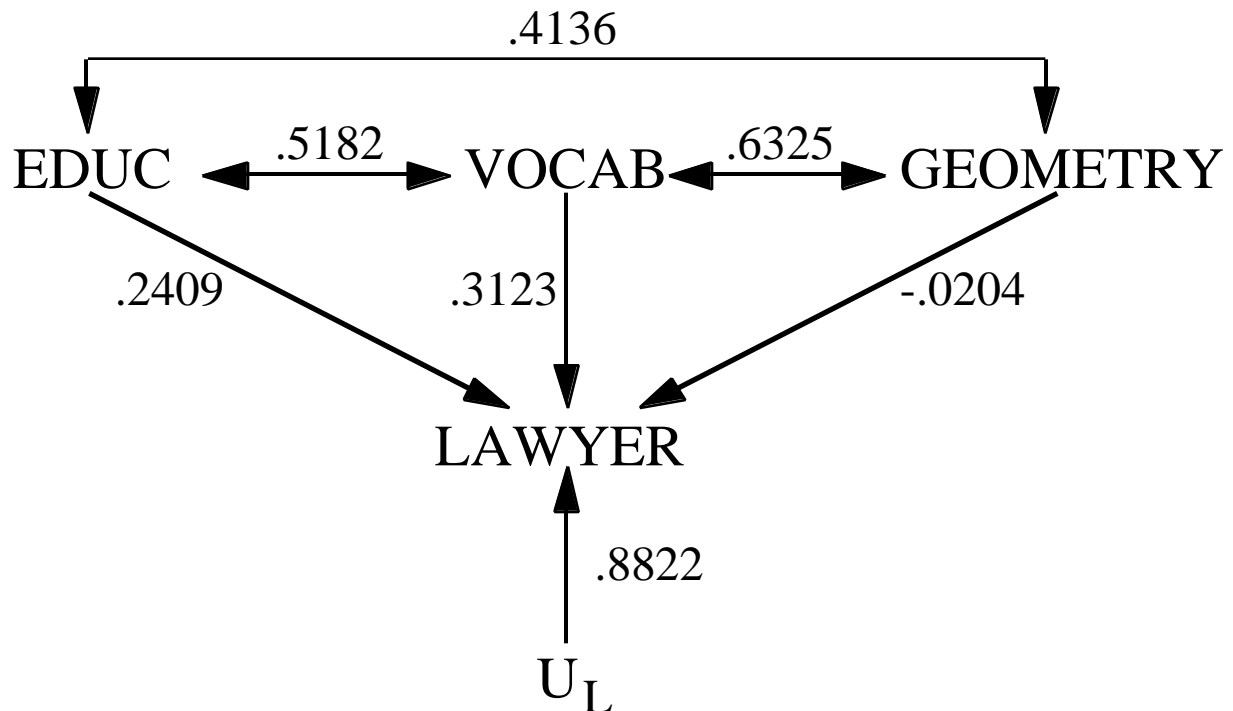
The utility of path analysis here is to *decompose the sources of a correlation* between an independent variable and a dependent variable. That is, we can use the path diagram to uncover why education, say, is correlated with (or predicts) interest in a legal profession.

Let's consider the relationship between LAWYER and EDUC. First, education has a *direct* effect on LAWYER[1]. This is depicted by the straight arrow going into LAWYER from EDUC . The magnitude of this effect is quantified by the standardized regression coefficient, or .2409. Second, education has two *indirect* effects. The first indirect effect arises because education is correlated with vocabulary and vocabulary directly predicts LAWYER. This is depicted by the pathway starting from EDUC, going into VOCAB, and then exiting from VOCAB directly into LAWYER. This indirect effect is quantified by the *product* of these two paths. Thus, the indirect effect of EDUC going through VOCAB equals .5182(.3123) = .1618. The second indirect effects reflects the correlation between EDUC and GEOMETRY and the direct effect of GEOMETRY on LAWYER. This is depicted by the pathway from EDUC to GEOMETRY and then the direct arrow from GEOMETRY to LAWYER. The magnitude of this indirect effect is .4136*(-.0206) = -.0085.

_____

[1] The term "effect" is used in a noncausal or predictive sense. Statistics themselves cannot determine causal relationships although they may aid in uncovering causal associations. Issues of experimental design and previous empirical research must always be considered.

Applied to multiple regression, the primary rule of path analysis states that *the correlation between an independent and a dependent variable is the sum of the direct effect and all indirect effects*. Thus, the correlation between EDUC and LAWYER equals .2409 + 5182(.3123) + .4136*(-.0206) = .2409 + .1618 + (-.0085) = .3942. Now look at the observed correlation between these two variables. You can verify that it, in fact, equals .3942 (within rounding error, of course). The advantage of examining the correlation between EDUC and LAWYER in this way is that one can compare the direct with the indirect effects. In this case, EDUC predicts interest in a legal career more strongly in a direct way (.2409) than it does in an indirect way (.1618 - .0085 = .1533).

Going through the same procedure for VOCAB and LAWYER gives a direct effect of .3123, an indirect effect through EDUC of .5182(.2409) = .1248, and an indirect effect through GEOMETRY of .6325(-.0204) = -.0129. Once again, the direct effect (.3123) is larger than the indirect effects (.1248 - .0129 = .1119).

For GEOMETRY, the direct effect is -.0204, the indirect effect through EDUC is .4136(.2409) = .0996 and the indirect effect through VOCAB is .6325(.3123) = .1975. Thus, the correlation between GEOMETRY and LAWYER -.0204 + .0996 + .1975 = .2767. Note here that the total indirect effect (.0996 + .1975 = .2971) is stronger than the direct effect (-.0204). Thus, even though the observed correlation between GEOMETRY and LAWYER is significant, the path interpretation suggests that the correlation arises because GEOMETRY is correlated with other variables that have direct effects upon LAWYER, not because geometry itself directly predicts LAWYER.

Figure 3 gives the path model for ARCHITCT. You should be able to calculate the direct and the indirect effects of the independent variables for this case.

Figure 3. Path model for variable ARCHTCT.

.4136

EDUC ←.5182→ VOCAB ←.6325→ GEOMETRY

.2506   .1484   .007

ARCHTCT

.9348

U_A
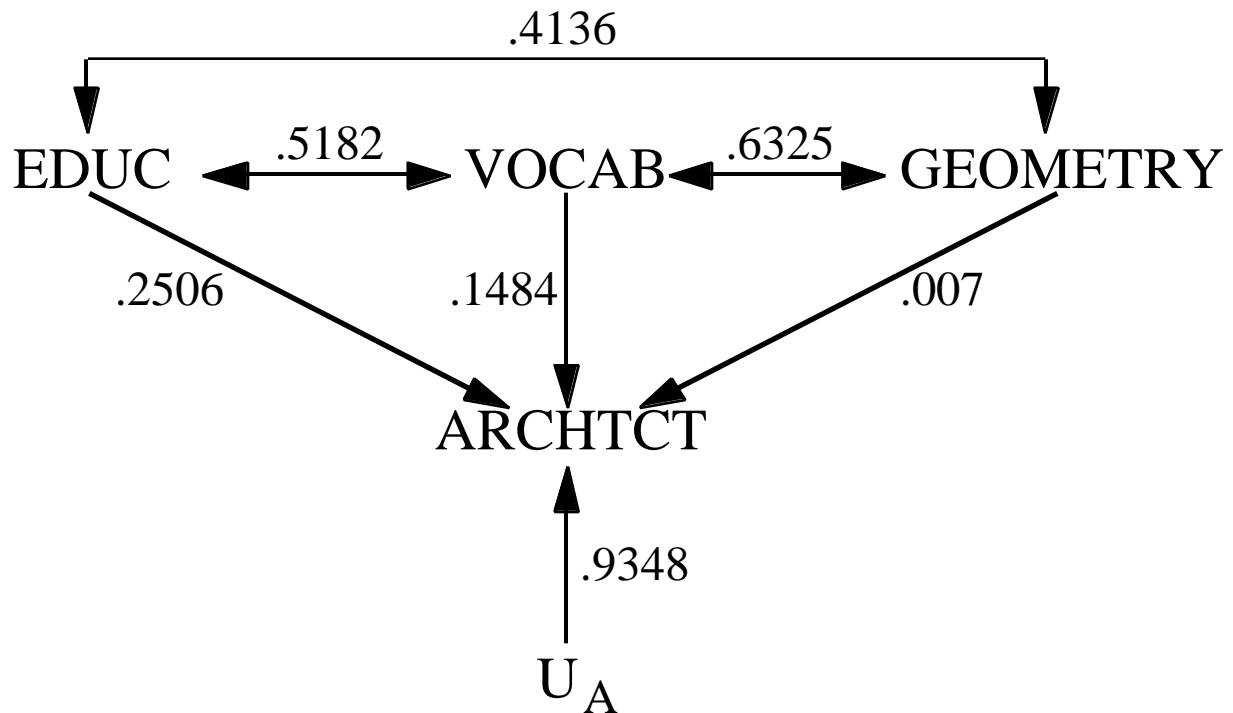
Formal Tracing Rules

Although we "intuited" the rules for path analysis above, there are formal tracing rules for calculating the correlations for a path diagram. First, pick a variable to start. It can be either the independent variable or the dependent variable. Then trace a route to the other variable, multiplying the coefficients when you go through two or more paths. Add together the results of these tracings for all the unique pathways. There are two exclusionary rules: (1) if you enter a variable on an arrowhead, you cannot exit on an arrowhead. Therefore, tracing from EDUC to VOCAB and then from VOCAB to GEOMETRY is illegal because we entered VOCAB on an arrowhead and exited VOCAB on an arrowhead. But tracing from LAWYER to EDUC and then from EDUC to VOCAB is legal because we did not enter on EDUC on an arrowhead. (2) in any single pathway, you cannot go through the same variable twice. This rule is mentioned here for completeness. One will not encounter cases in a multiple regression path model where one could go through the same variable twice.

Multivariate Multiple Regression & Path Analysis

An astute person who examines the significance and values of the standardized beta weights and the correlations will quickly realize that interpretation through path analysis and interpretation of these weights give the same substantive conclusions. The chief advantage of path analysis is seen when there are two or more dependent variables. Technically, this is referred to as *multivariate* multiple regression. Here path analysis *decomposes the sources of the correlations among the dependent variables*. For the present example, we use path analysis to

answer the following question: Why are interests in being a lawyer and an architect correlated?

The first task is to draw the path diagram. This is actually very easy if the diagrams for the single dependent variables have already been drawn. The multivariate regression diagram is almost the two single diagrams "superimposed" on one another. The correlations among the independent variables remain the same and path coefficient that goes from any independent variable into any dependent variable is simply the standardized regression coefficient for the independent variable from the single multiple regression. The only other thing to do is to draw double headed arrows connecting each pair of residuals in the multivariate diagram.

Figure 4 shows the multivariate diagram for the example. Note how the path coefficients for LAWYER are the same as those in Figure 2 and the coefficients for ARCHTCT are the same as those in Figure 3. Note also that the residual for LAWYER is correlated with the residual for ARCHTCT.

Figure 4. Multivariate regression between LAWYER and ARCHTCT expressed as a path diagram.
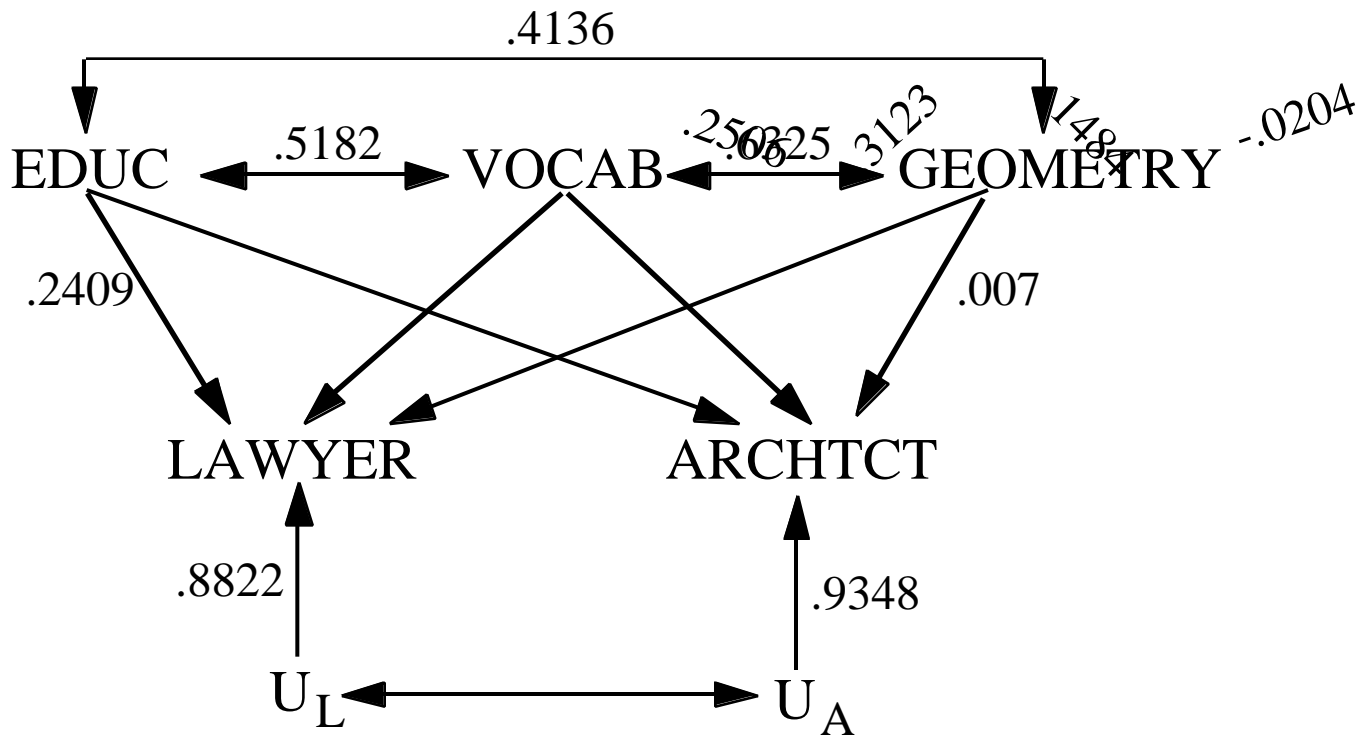
The only missing quantity in Figure 4 is the correlation between U_L and U_A or between the residual for LAWYER and ARCHTCT. This correlation may be calculated by hand using the printout from the multivariate test. To compute this test, simply add the following statement in the regression procedure after the model statement:

```
MTEST / PRINT;
```

The multivariate F statistics will test the null hypothesis that all regression coefficients are 0 for

all dependent variables. The PRINT option will print out the residual sums of squares and cross products matrix, labeled "E, the Error Matrix" on the output, and the hypothesis sums of squares and cross products matrix, labeled "H, the Hypothesis Matrix." To calculate the correlation between the residuals for the ith and jth variable, take element ($i,,j$) in the E matrix and divide it by the square root of (ith diagonal element multiplied by jth diagonal element). For the current problem, the E matrix is

$$\begin{matrix} 209.2821 & 81.8160 \\ 81.8160 & 221.2480 \end{matrix}$$

Thus, the correlation between the residuals for LAWYER and ARCHTCT is

$$\frac{81.816}{\sqrt{(209.2821)(221.2480)}} = .3802$$

(From your knowledge of a SSCP matrix, a covariance matrix, and a correlation matrix, you should be able to prove why this equals the correlation.)

Now that the diagram is explained, we want to quantify the reasons for the correlation between LAWYER and ARCHTCT. We can do this using the tracing rules, starting with LAWYER and ending in ARCHTCT. There are three direct pathways, one for each independent variable. The direct pathway for EDUC goes from LAWYER to EDUC and then directly from EDUC to ARCHTCT. The magnitude of this path is .2409(.2506) = .0604. This quantifies how much the direct effects of education are responsible for the correlation between LAWYER and ARCHTCT. The second direct effect goes through VOCAB and equals .3123(.1484) = .0463. The third and final direct effect goes though GEOMETRY and is -.0204(.007) = .0001, a trivial amount.

Next we want to calculate all the indirect ways in which the two dependent variables may be correlated. Consider first the indirect effects that pass through the two independent variables EDUC and VOCAB. The first way that these two variables influence the correlation between LAWYER and ARCHTCT comes about because education directly influences LAWYER, education is correlated with vocabulary ability, and vocabulary ability directly predicts ARCHTCT. This pathway starts in LAWYER, goes to EDUC, then to VOCAB, and then to ARCHTCT. The pathway here is .2409(.5182)(.1484) = .0185. There is another pathway reflecting a similar indirect mechanism--VOCAB has a direct effect on LAWYER, VOCAB is correlated with EDUC and EDUC has a direct effect on ARCHTCT. This path starts at LAWYER, goes to VOCAB, thence to EDUC, and finally to ARCHTCT, or .3123(.5182)(.2506) = .0406. Thus, the results of the indirect effects of EDUC and VOCAB on the correlation between LAWYER and ARCHTCT is the sum of these two pathways, or .0185 + .0406 = .0591. We would continue with this logic until we have exhausted all the indirect pathways between every pair of independent variables. These pathways are given in Table 1.

The final way to trace paths from LAWYER to ARCHTCT is through residuals. The magnitude of this is .8822(.3802)(.9348) =.3135. It is always illustrative to compare the amount that the independent variables contribute to the correlation with the amount that the residuals contribute. Theoretically, if the independent variables explained a large percentage of the

correlation between LAWYER and ARCHTCT, then the residual correlation should go towards 0. Once again, it is a sobering experience in data analysis to see how much we cannot predict behavioral traits.

Table 1. Decomposition of the correlation between LAWYER and ARCHTCT.

Type	Source	Amount	Total
Direct	EDUC	.2409(.2506) = .0604	
	VOCAB	.3123(.1484) = .0463	
	GEOMETRY	-.0204(.007) = .0001	
	Total Direct		.1068
Indirect	EDUC & VOCAB	.2409(.5182)(.1484) + .3123(.5182)(.2506) = .0591	
	EDUC & GEOMETRY	.2409(.4136)(.007) + (-.0204)(.4136)(.2506) = -.0014	
	VOCAB & GEOMETRY	.3123(.6325)(.007) + (-.0204)(.6325)(.1484) =-.0005	
	Total Indirect		.0572
Residual		.8822(.3802)(.9348) =.3135	.3135
Total			.4775

Decomposing the Variance of a Dependent Variable

The rules of path analysis may also be used to decompose the variance of a single dependent variable. For example, reconsider the prediction of LAWYER as given in Figure 2. To calculate the variance of LAWYER, simply follow the rules of path analysis using LAWYER as the first variance and LAWYER as the second variable. That is, we start with LAWYER and trace all paths that end at LAWYER.

One path is from LAWYER to EDUC and thence back to LAWYER. The amount of this pathway is .2409*.2409 = .0580. Another pathway starts with LAWYER goes to EDUC , then to VOCAB, and then to LAWYER. The quantity here is .2409*.5182*.3123 = .0390. Table 2 shows the various pathways and the quantities associated with those pathways for explaining the variance of LAWYER.

Table 2. Decomposition of the variance of LAWYER.

Pathway	Quantity	Total
LAWYER to EDUC to LAWYER	.2409*.2409=	.0580
LAWYER to VOCAB to LAWYER	.3123*.3123=	.0975
LAWYER to GEOMETRY to LAWYER	-.0204*-.0204=	.0004
LAWYER to EDUC to VERBAL to LAWYER	.2409*.5182*.3123=	.0390
LAWYER to VERBAL to EDUC to LAWYER	.3123*.5182*.2409=	.0390
LAWYER to EDUC to GEOMETRY to LAWYER	.2409*.4136*-.0204=	-.0020
LAWYER to GEOMETRY to EDUC to LAWYER	-.0204*.4136*.2409=	-.0020
LAWYER to VOCAB to GEOMETRY to LAWYER	.3123*.6325*-.0204=	-.0040
LAWYER to GEOMETRY to VOCAB to LAWYER	-.0204*.6325*.3123=	-.0040
LAWYER to U_L to LAWYER	.8822*.8822=	.7782

Note that the pathway LAWYER to EDUC to VOCAB to LAWYER is different from the pathway LAWYER to VOCAB to EDUC to LAWYER, even though both result in the same quantity. As an analogy, we may consider a trip between three cities. A trip originating in Wachington DC, then going to Philadelphia, then to New York City, and finally returning to DC is really a different trip from that starting in DC but then going to New York, then to Philadelphia, and back DC, even though both trips have the same mileage.

Algebra and the Tracing Rules

The tracing rules may also be expressed in algebra. These rules are given here for the sake of completeness; it is not necessary to know them to perform path analysis. Let Y denote a dependent variable, X an independent variable,   a standardized regression coefficient, and   a correlation coefficient. Then the correlation between the dependent variable and the ith

independent variable will be

$$(Y, X_i) = \sum_{j=1}^{p} {}_{Y,X_j}(X_j, X_i).$$

The summation here is taken over the *p* independent variables.

For example, the correlation between Education and Lawyer, would be

$${}_{LE} = {}_{LE} + {}_{LV}{}_{VE} + {}_{LG}{}_{GE}.$$

Substituting the numbers for the algebraic quantities gives

$$.3943 = .2409 + (3123)(.5182) + (-.0206)(.4136).$$

Carrying out the multiplications we have

$$.3943 = .2409 + .1618 - .0085.$$

(Note that the answer is not exact because of minor rounding error.) As in our earlier discussion, each of these quantities has a meaning because each number quantifies the extent to which Education is correlated with Lawyer through the different paths. The algebraic notion, the numerical quantities, and the meaning of the algebra and the numerical quantities are given below in Table 3.

Table 3. Why is Education correlated with Lawyer?		
Algebra	**Numerical Quantity**	**Meaning**
EL	.2409	Because Education predicts lawyer directly.
EV VL	.1618	Because Education is correlated with Vocabulary and Vocabulary predicts Lawyer directly.
G GL	-.0085	Because Education is correlated with Geometry and Geometry predicts Lawyer directly.
Total	.3942	

This general algebraic rule may be expressed very parsimoniously in matrix notation. Let ${}_{xy}$ denote a column vector of the correlations between the independent variables and the dependent variable, let $\mathbf{R_{XX}}$ denote the correlation matrix among all the independent variables, and let ${}_{xy}$ denote a column vector containing the standardized partial regression weights. Then the correlation between the dependent variable and the independent variables will equal

$${}_{Xy} = \mathbf{R}_{XX}\,{}_{Xy}.$$

For the current example, this matrix equation expressed in terms of its elements would look like this

$$\begin{matrix} {}_{EL} \\ {}_{VL} \\ {}_{GL} \end{matrix} = \begin{matrix} 1.0 & {}_{EV} & {}_{EG} \\ {}_{VE} & 1.0 & {}_{VG} \\ {}_{GE} & {}_{GV} & 1.0 \end{matrix} \begin{matrix} {}_{EL} \\ {}_{VL} \\ {}_{GL} \end{matrix}.$$

In a multivariate multiple regression the correlation between the ith dependent variable and the jth dependent variable may be expressed as

$$(Y_i, Y_j) = \sum_{k=1}^{P}\sum_{l=1}^{P} {}_{Y_i,X_k}(X_k, X_l)\,{}_{Y_j,X_l} + u_i u_j\,(U_i, U_j).$$

Once again, this equation can be expressed in matrix algebra. Let $\mathbf{R_{XX}}$ denote the correlation matrix among the independent variables; $\mathbf{R_{YY}}$, the correlation matrix among the dependent variables; $\mathbf{R_{XY}}$, the correlation matrix between the independent variables (rows) and the dependent variables (columns), and $\mathbf{R_{UU}}$ the correlation matrix among the residuals. Let ${}_{XY}$ denote the matrix of standardized regression coefficients; the coefficients for the first dependent variable form the first column in ${}_{XY}$, those for the second dependent variable form the second column, etc. Finally, let $\mathbf{U}$ denote a diagonal matrix containing the path coefficients from the residual to the dependent variable along the diagonal. That is, the ith diagonal element equals

$$u_i = \sqrt{1 - R_i^2}$$

Then,

$$\mathbf{R}_{XY} = \mathbf{R}_{XX}\,{}_{XY}.$$

In terms of matrix elements for the current problem, this equation would be written as

$$
\begin{bmatrix} EL & EA \\ VL & VA \\ GL & GA \end{bmatrix}
=
\begin{bmatrix} 1.0 & EV & EG \\ VE & 1.0 & VG \\ GE & GV & 1.0 \end{bmatrix}
\begin{bmatrix} EL & EA \\ VL & VA \\ GL & GA \end{bmatrix} .
$$

Finally, the matrix equation for the correlation among the dependent variables is

$$\mathbf{R}_{YY} = {}_{XY}\mathbf{R}_{XX}\,{}_{XY} + \mathbf{U}\mathbf{R}_{UU}\mathbf{U}.$$

Expressing this equation in terms of the matrix elements for the present example gives

$$
\begin{bmatrix} 1.0 & LA \\ AL & 1.0 \end{bmatrix}
=
$$

$$
\begin{bmatrix} EL & VL & GL \\ EA & VA & GA \end{bmatrix}
\begin{bmatrix} 1.0 & EV & EG \\ VE & 1.0 & VG \\ GE & GV & 1.0 \end{bmatrix}
\begin{bmatrix} EL & EA \\ VL & VA \\ GL & GA \end{bmatrix}
+
\begin{bmatrix} u_L & 0 \\ 0 & u_A \end{bmatrix}
\begin{bmatrix} 1.0 & u_L u_A \\ u_A u_L & 1.0 \end{bmatrix}
\begin{bmatrix} u_L & 0 \\ 0 & u_A \end{bmatrix}
$$

An interesting exercise, left to the reader, is to use the algebra of multiple regression to prove that the diagonals of $\mathbf{R_{YY}}$ will always equal 1.0.

```
NOTE: AUTOEXEC processing completed.

1     /* -----------------------------------------------------
2         Psychology 7291: Multivariate Statistics
3         File: ~carey/p7291dir/pathreg1.sas
4
5         Example of the statements for multivariate
6         multiple regression and path analysis. This
7         program has the results given in the handout
8         PATH ANALYSIS & REGRESSION
9        ----------------------------------------------------- */
10
11    /*
12         The LIBNAME statement gives a short (8 letters or less)
13         abbreviation for a path or directory or, in MAC terms,
14         a folder. The statement below assigns the name "carey"
15         to the directory ~carey/p7291dir
16    */
17    LIBNAME carey '~carey/p7291dir';
NOTE: Libref CAREY was successfully assigned as follows:
      Engine:        V609
      Physical Name: /home/facstaff/carey/p7291dir
18
19    /*
20         A sas statement of the form:
21              LIBNAME.MEMBER
22         instructs sas to access the disk directory
23         specified by LIBNAME and the SAS data set called
24         MEMBER.
25         In the PROC REG statement below, the statement is
26              DATA=CAREY.INTEREST
27         instructs SAS to use the SAS data set
28         named "interest" which is located on the
29         disk area specified by "carey", i.e., ~carey/p7291dir
30    */
31
32    /* To perform a path analysis, make certain to use the
33       CORR option on the PROC REG statement to print out
34       the correlation matrix for the variables. Also make
35       certain to use the STB option on the MODEL statement
36       to get standardized beta coefficients */
37
38    TITLE 'Path Analysis and Multivariate Multiple Regression';
39    PROC REG DATA=carey.interest CORR;
40      MODEL lawyer archtct = educ vocab geometry   / STB;
41      MTEST / PRINT;
42    run;

NOTE: 250 observations read.
NOTE: 250 observations used in computations.
```

```
                  Path Analysis and Multivariate Multiple Regression
 1
                                          11:06 Tuesday, February 6,
1996

                                 Correlation
CORR          EDUC           VOCAB        GEOMETRY
EDUC         1.0000          0.5182          0.4136   education in years
VOCAB        0.5182          1.0000          0.6325   cognitive: vocabulary test
GEOMETRY     0.4136          0.6325          1.0000   cognitive: geometry
LAWYER       0.3943          0.4242          0.2767   Interests: lawyer
ARCHTCT      0.3304          0.2827          0.2045   Interests: architect

        CORR          LAWYER         ARCHTCT
        EDUC          0.3943          0.3304   education in years
        VOCAB         0.4242          0.2827   cognitive: vocabulary test
        GEOMETRY      0.2767          0.2045   cognitive: geometry
        LAWYER        1.0000          0.4773   Interests: lawyer
        ARCHTCT       0.4773          1.0000   Interests: architect
```

Model: MODEL1
Dependent Variable: LAWYER   Interests: lawyer

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	3	59.65504	19.88501	23.374	0.0001
Error	246	209.28213	0.85074		
C Total	249	268.93717			

```
        Root MSE       0.92236    R-square       0.2218
        Dep Mean      -0.01096    Adj R-sq       0.2123
        C.V.      -8415.65585
```

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	-1.946186	0.52488697	-3.708	0.0003
EDUC	1	0.155090	0.04269582	3.632	0.0003
VOCAB	1	0.325128	0.08115635	4.006	0.0001
GEOMETRY	1	-0.020564	0.07362127	-0.279	0.7802

Variable	DF	Standardized Estimate	Variable Label
INTERCEP	1	0.00000000	Intercept
EDUC	1	0.24090497	education in years
VOCAB	1	0.31232007	cognitive: vocabulary test
GEOMETRY	1	-0.02045512	cognitive: geometry

```
                    Path Analysis and Multivariate Multiple Regression
  3
                                                       11:06 Tuesday, February 6,
1996

Dependent Variable: ARCHTCT      Interests: architect

                              Analysis of Variance

                              Sum of          Mean
            Source       DF     Squares        Square      F Value       Prob>F

            Model         3    31.95809      10.65270       11.844       0.0001
            Error       246   221.24795       0.89938
            C Total     249   253.20604

                Root MSE        0.94836    R-square        0.1262
                Dep Mean        0.10920    Adj R-sq        0.1156
                C.V.          868.45924

                              Parameter Estimates

                         Parameter        Standard     T for H0:
            Variable  DF    Estimate          Error    Parameter=0     Prob > |T|

            INTERCEP   1   -1.831357     0.53968376        -3.393         0.0008
            EDUC       1    0.156556     0.04389943         3.566         0.0004
            VOCAB      1    0.149919     0.08344418         1.797         0.0736
            GEOMETRY   1    0.006847     0.07569668         0.090         0.9280

                         Standardized  Variable
            Variable  DF     Estimate     Label

            INTERCEP   1   0.00000000  Intercept
            EDUC       1   0.25062320  education in years
            VOCAB      1   0.14841971  cognitive: vocabulary test
            GEOMETRY   1   0.00701897  cognitive: geometry
```

```
                     Path Analysis and Multivariate Multiple Regression
 4
                                             11:06 Tuesday, February 6,
1996
```

Multivariate Test:

```
                              E, the Error Matrix


                         209.28213143        81.815974611
                         81.815974611        221.24795443


                           H, the Hypothesis Matrix


                         59.655038175        42.723933389
                         42.723933389        31.958085574
```

Multivariate Statistics and F Approximations

```
                         S=2      M=0      N=121.5
```

Statistic	Value	F	Num DF	Den DF	Pr > F
Wilks' Lambda	0.75322602	12.4317	6	490	0.0001
Pillai's Trace	0.24831671	11.6242	6	492	0.0001
Hotelling-Lawley Trace	0.32557459	13.2400	6	488	0.0001
Roy's Greatest Root	0.31915719	26.1709	3	246	0.0001

```
        NOTE: F Statistic for Roy's Greatest Root is an upper bound.
             NOTE: F Statistic for Wilks' Lambda is exact.
```