Newsom
USP 655 SEM
Winter 2008

Wright's Rules of Tracing

Sewall Wright (1918, 1934) developed a method of estimating causal path coefficients by decomposing the correlations among a set of variables. He articulated a set of rules for examining a path diagram that would allow for this mathematical decomposition.[1] The correlation of any two variables in a path diagram can be expressed as the sum of coefficients that connect the two variables.

1) No loops are allowed. In tracing from one variable to another, you cannot pass through the same variable twice.

2) No going forward and then backward. Once you have traveled along a path forward, you cannot travel backward across the path. However, going backward and then forward is possible.

3) Only one curved arrow is allowed in tracing from the first variable to the last variable.

References

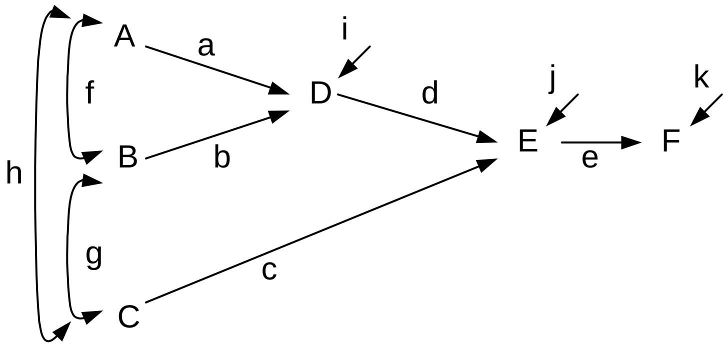
Wright, S. (1918). On the Nature of Size Factors. *Genetics, 3*, 367-74.

Wright, S. (1934). The Method of Path Coefficients. *Annals of Mathematical Statistics, 5,* 161-215.

---

[1] Note that these rules are often worded differently by different authors and I provide a paraphrased and reorganized version.

Tracing Example[2]



In the above diagram, lower case letters represent values of the coefficients. For curved arrows, they represent correlations; for straight arrows, they represent regression coefficients; and for the short arrows, they represent the value of the error or disturbance.

By using Wright's tracing rules, one can obtain the value of the correlation between any two variables, by adding up the values of the coefficients that link them--provided the rules are followed. Every route between the two variables is traced. The coefficients that make up each route are multiplied. If there are multiple routes that link the two variables, products for each route are added together.

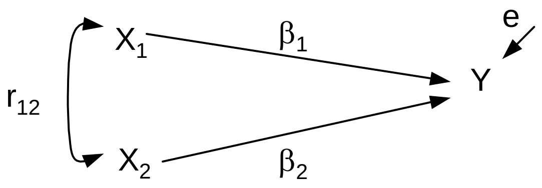
- To find the correlation between B and D (or r_{BD}), two routes are possible: b and fa. So, the value of r_{BD} is equal to b + fa. If b=.3, f=.2, and a=.4, $r_{BD} = .3 + (.2)(.4) = .38$.
- r_{CD} = gb + ha.
- r_{AB} is simply equal to f, because the curved arrow represents a correlation.
- r_{AE} = ad + fbd + hc

---

[2] This example is adapted from Loehlin, J.C. (1992). *Latent variable models* (2nd Ed.). Hillsdale, NJ: Erlbaum.

Decomposing the Correlation Matrix

Although this seems like just a fun little game, it turns out to be immensely useful. Instead of computing the correlations between two variables, one can work backward from the correlations to derive the path coefficients. Taking a simple path diagram representing a two variable regression model, and assuming some values for the correlations between our three variables, we can derive the path coefficients.



Assume that $r_{12} = .50$, $r_{1Y} = .65$, and $r_{2Y} = .70$.

From our tracing rules, we know $r_{1Y} = \beta_1 + r_{12}\beta_2$ and $r_{2Y} = \beta_2 + r_{12}\beta_1$. We can plug in the known values and solve for β_1 and β_2.

$$.65 = \beta_1 + .50\beta_2$$
$$.70 = \beta_2 + .50\beta_1$$

by rearranging, substituting, and solving for β_1 and then β_2, we get $\beta_1 = .4$ and $\beta_2 = .5$. These coefficients are the standardized regression coefficients, because we started with a correlation matrix (i.e., standardized variables).

The disturbance term, *e*, is the amount of unaccounted for variance in *Y* and is equal to $\sqrt{1-R^2}$. We know from regression analysis that $R^2 = \beta_1^2 + \beta_2^2 + \beta_1\beta_2 r_{12}$. So, $R^2 = .4^2 + .5^2 + .5(.4)(.5) = .61$. The disturbance term then equals $e = \sqrt{1-.61} = .62$

What we have just done is decompose the correlation matrix into unique values for the coefficients that are implied by the model we specified.