1

Asymptotic Covariance Matrix

In this primer, I show you how to locate the covariance matrix for coefficients (also known as the asymptotic covariance matrix) on the output of different software. I use as my running example an RET for a program to increase discretionary income by teaching people concepts related to financial literacy. Discretionary income is after-tax income a family has after basic living expenses are covered. The program addressed two topics (1) budgeting, and (2) the ins and outs of using credit cards. The program designers felt that educating people about each of these topics would lead to increases in discretionary income. There was a control and treatment condition (scored 0 and 1, respectively). Each mediator was measured by a knowledge test ranging from 0 to 100. A score of 90 means that 90% of the items were answered correctly, a score of 80 means 80% of the items were answered correctly, and so on. The knowledge measures were obtained at baseline and again at the immediate posttest. The outcome was an index of the monthly discretionary income measured three months after program completion. A baseline measure of monthly discretionary income also was obtained. The sample size was 400. This was a low-income population whose overall annual income was uniformly close to $20,000. The influence diagram describing the logic model is in Figure 1.

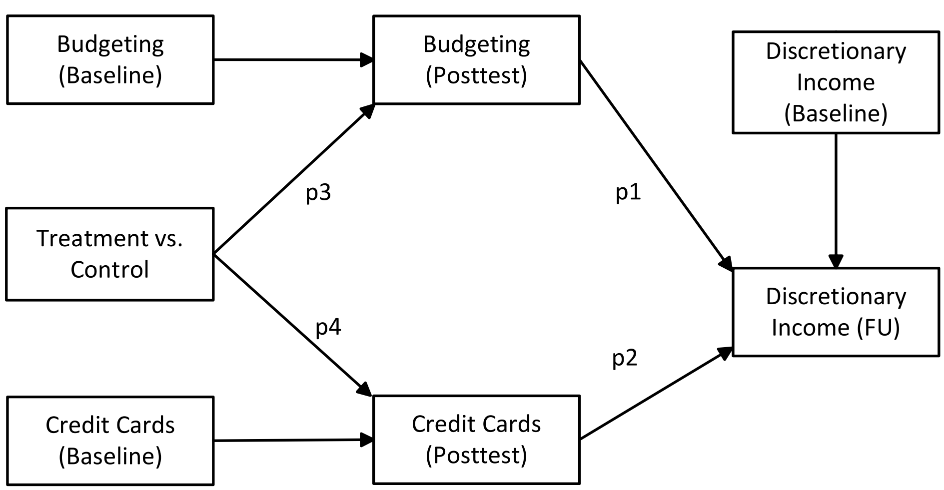

FIGURE 1. Discretionary income example

I illustrate the case where I want to calculate Monte Carlo confidence intervals for the total effect. The total effect is defined by the expression p1*p3 + p2*p4 using the path notation in Figure 1. Given this, I need to isolate a lower triangle asymptotic covariance matrix (ACM) for input into the program on my website. The columns of the matrix are p1, p2, p3 and p4, as are the rows.

For ways of isolating the ACM for different model forms and in software packages not covered here, see http://www.quantpsy.org/interact/acov.htm

MPLUS

Here is the program code for the current example, which also shows the names of the variables I used in Mplus to identify the variables:

```
TITLE: Discretionary income examples ;
DATA: FILE IS c:\mplus\ret\newchap8\bayes3M.txt ;
VARIABLE:
    NAMES ARE id treat income1 budget1 credit1
      cbudget1 ccredit1 cincome1 budget2 credit2 income3 ;
    USEVARIABLES ARE treat budget1 credit1 income1  !baseline variables
      budget2 credit2 income3 ;   !posttreatment variables
    MISSING ARE ALL (-9999) ;
ANALYSIS:
    ESTIMATOR = MLR ;
    MODEL:
    income3 ON budget2 credit2 income1 ;   !regresses Y onto mediators and cov
    budget2 ON treat budget1 ;  !regresses mediator onto treatment and cov
    credit2 ON treat credit1 ;  !regresses mediator onto treatment and cov
OUTPUT:
      SAMP STANDARDIZED(STDYX) MOD(ALL 4) RESIDUAL
      CINTERVAL TECH4 TECH1 TECH3 ;
```

All programming is routine except for the `OUTPUT` line. I added the options `TECH1` and `TECH3`. I need these additions to obtain the desired entries for the ACM.

The `TECHNICAL 1 OUTPUT` provides an ordered list of model parameters that are estimated. The list is organized into different matrices that have Greek headings. The two matrices that are of primary interest to us are the `BETA` matrix and the `GAMMA` matrix. The variables in the column of the `BETA` matrix are possible causal determinants; the variables in the rows are possible outcomes. If a 0 occurs in the cell where the row and column intersect, then the referenced parameter was not estimated. If a positive integer appears, then the referenced parameter was estimated and it is referred to later by that number. The `GAMMA` matrix has the same format but it focuses on a different subset of variables and is only produced for weighted least squares modeling. It does not appear in the current case. I boldfaced the relevant cells for the path coefficients p1, p2, p3 and p4 to make them easier for you to see, but this will not occur in practice.

```
TECHNICAL 1 OUTPUT

           NU
              BUDGET2       CREDIT2       INCOME3       TREAT         CBUDGET1
              ________      ________      ________      ________      ________
                  0             0             0             0             0
           NU
              CCREDIT1      CINCOME1
              ________      ________
                  0             0

           LAMBDA
              BUDGET2       CREDIT2       INCOME3       TREAT         CBUDGET1
              ________      ________      ________      ________      ________
 BUDGET2            0             0             0             0             0
 CREDIT2            0             0             0             0             0
 INCOME3            0             0             0             0             0
 TREAT              0             0             0             0             0
 CBUDGET1           0             0             0             0             0
 CCREDIT1           0             0             0             0             0
 CINCOME1           0             0             0             0             0

           LAMBDA
              CCREDIT1      CINCOME1
              ________      ________
 BUDGET2            0             0
 CREDIT2            0             0
 INCOME3            0             0
 TREAT              0             0
 CBUDGET1           0             0
 CCREDIT1           0             0
 CINCOME1           0             0

           THETA
              BUDGET2       CREDIT2       INCOME3       TREAT         CBUDGET1
              ________      ________      ________      ________      ________
 BUDGET2            0
 CREDIT2            0             0
 INCOME3            0             0             0
 TREAT              0             0             0             0
 CBUDGET1           0             0             0             0             0
 CCREDIT1           0             0             0             0             0
 CINCOME1           0             0             0             0             0
```

THETA

	CCREDIT1	CINCOME1
CCREDIT1	0	
CINCOME1	0	0

ALPHA

BUDGET2	CREDIT2	INCOME3	TREAT	CBUDGET1
1	2	3	0	0

ALPHA

CCREDIT1	CINCOME1
0	0

BETA

	BUDGET2	CREDIT2	INCOME3	TREAT	CBUDGET1
BUDGET2	0	0	0	**4**	5
CREDIT2	0	0	0	**6**	0
INCOME3	**8**	**9**	0	0	0
TREAT	0	0	0	0	0
CBUDGET1	0	0	0	0	0
CCREDIT1	0	0	0	0	0
CINCOME1	0	0	0	0	0

BETA

	CCREDIT1	CINCOME1
BUDGET2	0	0
CREDIT2	7	0
INCOME3	0	10
TREAT	0	0
CBUDGET1	0	0
CCREDIT1	0	0
CINCOME1	0	0

PSI

	BUDGET2	CREDIT2	INCOME3	TREAT	CBUDGET1
BUDGET2	11				
CREDIT2	0	12			
INCOME3	0	0	13		
TREAT	0	0	0	0	
CBUDGET1	0	0	0	0	0
CCREDIT1	0	0	0	0	0
CINCOME1	0	0	0	0	0

```
          PSI
             CCREDIT1      CINCOME1
             ________      ________
 CCREDIT1           0
 CINCOME1           0             0
```

To use the program on Monte Carlo confidence intervals, I need to locate the parameter numbers for p1, p2, p3 and p4 in Figure 1. In the `BETA` matrix, they are:

p1 = `BUDGET2` → `INCOME3` = parameter number 8
p2 = `CREDIT2` → `INCOME3` = parameter number 9
p3 = `TREAT` → `BUDGET2` = parameter number 4
p4 = `TREAT` → `CREDIT2` = parameter number 6

Next I go to the output in the section called `TECHNICAL 3 OUTPUT`. Here is the output:

ESTIMATED COVARIANCE MATRIX FOR PARAMETER ESTIMATES

	1	2	3	4	5
1	0.225505D+00				
2	-0.855223D-02	0.253947D+00			
3	0.203298D+01	0.297062D-01	0.831604D+04		
4	-0.384073D-01	-0.239781D-05	0.475768D-01	**0.548990D-01**	
5	-0.517527D-01	0.225024D-02	-0.434244D+00	0.296279D-02	0.134975D-01
6	-0.546799D-02	-0.391533D-01	0.120153D+01	**0.655900D-03**	0.124756D-02
7	0.381422D-02	-0.602238D-01	-0.940821D-01	-0.239754D-03	-0.951734D-03
8	0.522197D-01	0.142926D+00	-0.364193D+02	**-0.401068D-01**	-0.306255D-02
9	0.636498D-01	-0.113771D+00	-0.886961D+01	**-0.158563D-01**	-0.203911D-01
10	-0.127269D-01	-0.605928D-03	-0.401965D+02	0.589787D-03	0.278425D-02
11	0.486719D-02	0.101548D-02	0.142411D+01	-0.492881D-02	-0.806491D-03
12	-0.420115D-02	0.252779D-02	0.216630D+01	-0.331603D-02	0.124098D-02
13	-0.401327D+02	0.145221D+03	-0.110502D+05	0.506373D+01	0.629385D+01

ESTIMATED COVARIANCE MATRIX FOR PARAMETER ESTIMATES

	6	7	8	9	10
6	**0.541482D-01**				
7	0.272425D-02	0.161266D-01			
8	**-0.284563D-01**	-0.316030D-01	**0.141128D+02**		
9	**-0.756705D-02**	0.264801D-01	**-0.294761D+01**	**0.111349D+02**	
10	-0.437027D-02	0.413324D-03	-0.962850D-01	-0.130563D+00	0.207242D+00
11	0.238494D-02	-0.139425D-02	-0.102785D+00	-0.107613D-01	-0.373338D-02
12	0.688524D-02	-0.294879D-02	-0.214674D-01	0.400748D-01	-0.125012D-01
13	0.466598D+02	-0.440414D+02	-0.843155D+03	-0.484537D+03	0.101474D+03

```
       ESTIMATED COVARIANCE MATRIX FOR PARAMETER ESTIMATES
             11             12             13
          ________       ________       ________
 11     0.158290D+00
 12    -0.969607D-02  0.140607D+00
 13    -0.228664D+02 -0.584311D+02  0.912023D+07
```

The columns and rows are the parameter numbers from the `TECH1` output. The entries are the relevant variance or covariance of the ACM using double precision notation, hence the letter D towards the end of each entry. A D followed by a negative number means to move the decimal point to the left by the number of digits indicated (e.g. `-0.969607D-02` is `-.00969607`. A D followed by a positive number means to move the decimal point to the right by the number of digits indicated (e.g. `-0.228664D+02` is `-22.8664`. I boldface the values I need to extract. Here is the lower triangle (plus the diagonal) of the matrix that I want:

	p1	p2	p3	p4
p1	14.113			
p2	-2.948	11.135		
p3	-0.040	-0.016	0.055	
p4	-0.028	-0.008	0.001	0.054

I input this matrix into the Monte Carlo confidence interval program on a per column basis.

Lavaan and MIIVsem

Here is the lavaan syntax for the model and the ACM:

```
XXX<-read.table('c:/ret/incomeR.dat',header=TRUE)  # get the data
XXX[XXX== -9999]<-NA                               # set missing data to NA
attach(XXX)              # bring forward the data file called XXX
library(lavaan)          # load the lavaan library
# specify the model
model<-'
budget2 ~ treat + budget1
credit2 ~ treat + credit1
income3 ~ income1 + budget2 + credit2
'                        # end model specification
fit <- sem(model, data=XXX,estimator='MLR',mimic='mplus')    # fit the data
inter<-unclass(vcov(fit))  # show the ACM to many decimals
inter[upper.tri(inter)] <- NA  # change upper triangle to NA
```

The final line of the code will produce the ACM. Here is the output (sorry for the small font but I want to fit the output in as it appears in R):

	budget2~treat	budget2~budget1	credit2~treat	credit2~credit1	income3~income1	income3~budget2	income3~credit2
budget2~treat	**0.0548989941**	NA	NA	NA	NA	NA	NA
budget2~budget1	0.0029627849	0.0134975347	NA	NA	NA	NA	NA
credit2~treat	**0.0006559065**	0.0012475610	**0.054148149**	NA	NA	NA	NA
credit2~credit1	-0.0002397591	-0.0009517356	0.002724250	1.612663e-02	NA	NA	NA
income3~income1	0.0005897681	0.0027842926	-0.004370204	4.132466e-04	2.072424e-01	NA	NA
income3~budget2	**-0.0401052778**	-0.0030609691	**-0.028457172**	-3.160277e-02	-9.628607e-02	**14.11285273**	NA
income3~credit2	**-0.0158540733**	-0.0203915280	**-0.007567998**	2.647894e-02	-1.305628e-01	**-2.94761239**	**11.13493118**
budget2~~budget2	-0.0049283223	-0.0008064760	0.002384940	-1.394250e-03	-3.733524e-03	-0.10277391	-0.01076429
credit2~~credit2	-0.0033160579	0.0012409864	0.006885200	-2.948753e-03	-1.250108e-02	-0.02146709	0.04007075
income3~~income3	5.0637199626	6.2936944686	46.660436182	-4.404179e+01	1.014815e+02	-843.27853258	-484.46584490
budget2~1	-0.0384072843	-0.0517527564	-0.005467985	3.814222e-03	-1.272708e-02	0.05221378	0.06364975
credit2~1	-0.0000023931	0.0022502407	-0.039153325	-6.022377e-02	-6.056827e-04	0.14292493	-0.11376634
income3~1	0.0475645299	-0.4342575981	1.201524311	-9.406277e-02	-4.019649e+01	-36.41909001	-8.86949636

	budget2~~budget2	credit2~~credit2	income3~~income3	budget2~1	credit2~1	income3~1
budget2~treat	NA	NA	NA	NA	NA	NA
budget2~budget1	NA	NA	NA	NA	NA	NA
credit2~treat	NA	NA	NA	NA	NA	NA
credit2~credit1	NA	NA	NA	NA	NA	NA
income3~income1	NA	NA	NA	NA	NA	NA
income3~budget2	NA	NA	NA	NA	NA	NA
income3~credit2	NA	NA	NA	NA	NA	NA
budget2~~budget2	0.158289268	NA	NA	NA	NA	NA
credit2~~credit2	-0.009696037	0.140606671	NA	NA	NA	NA
income3~~income3	-22.866301569	-58.431481492	9120415.75178	NA	NA	NA
budget2~1	0.004866879	-0.004201139	-40.13222	0.225504788	NA	NA
credit2~1	0.001015489	0.002527683	145.22256	-0.008552229	0.25394655	NA
income3~1	1.424096699	2.166291964	-11051.36174	2.033049345	0.02963523	8316.033

I boldface the entries that are relevant to the ACM for p1, p2, p3, and p4. The entries use double precision notation but with an e instead of a D. Here is the lower triangle (plus the diagonal) of the matrix that I want:

	p1	p2	p3	p4
p1	14.113			
p2	-2.948	11.135		
p3	-0.040	-0.016	0.055	
p4	-0.028	-0.008	0.001	0.054

I input this matrix into the Monte Carlo confidence interval program on a per column basis.

Here is the syntax for MIIV-SEM

```
XXX<-read.table('c:/ret/incomeR.dat',header=TRUE)  # get the data
XXX[XXX== -9999]<-NA                               # set missing data to NA
attach(XXX)              # bring forward the data file called XXX
library(MIIVsem)         # load the lavaan library
# specify the model
model<-'
budget2 ~ treat + budget1
credit2 ~ treat + credit1
```

```
income3 ~ income1 + budget2 + credit2
'                        # end model specification
fit <- miive(model, data=XXX)     # fit the data
inter<-fit$coefCov
inter[upper.tri(inter)] <- NA     # change upper triangle to NA
```

The final line of the code will produce the ACM. Here is the output (sorry for the small font but I want to fit the output in as it appears in R):

	budget2~1	budget2~treat	budget2~budget1	credit2~1	credit2~treat	credit2~credit1	income3~1	income3~income1
budget2~1	0.21813688	NA	NA	NA	NA	NA	NA	NA
budget2~treat	-0.03418529	**0.054628917**	NA	NA	NA	NA	NA	NA
budget2~budget1	-0.05006153	0.001692185	0.01316348	NA	NA	NA	NA	NA
credit2~1	0.00000000	0.000000000	0.00000000	0.20862227	NA	NA	NA	NA
credit2~treat	0.00000000	**0.000000000**	0.00000000	-0.02821457	**0.053957550**	NA	NA	NA
credit2~credit1	0.00000000	0.000000000	0.00000000	-0.04893388	0.000188485	0.01322173	NA	NA
income3~1	0.00000000	0.000000000	0.00000000	0.00000000	0.000000000	0.00000000	10337.48937	NA
income3~income1	0.00000000	0.000000000	0.00000000	0.00000000	0.000000000	0.00000000	-50.09902	0.255486643
income3~budget2	0.00000000	**0.000000000**	0.00000000	0.00000000	**0.000000000**	0.00000000	-26.41036	-0.120223557
income3~credit2	0.00000000	**0.000000000**	0.00000000	0.00000000	**0.000000000**	0.00000000	-37.53747	-0.007199023

	income3~budget2	income3~credit2
budget2~1	NA	NA
budget2~treat	NA	NA
budget2~budget1	NA	NA
credit2~1	NA	NA
credit2~treat	NA	NA
credit2~credit1	NA	NA
income3~1	NA	NA
income3~income1	NA	NA
income3~budget2	**14.224680**	NA
income3~credit2	**-3.774098**	**12.59338**

I boldface the entries that are relevant to the ACM for p1, p2, p3, and p4. The entries use double precision notation but with an e instead of a D. Here is the lower triangle (plus the diagonal) of the matrix that I want:

	p1	p2	p3	p4
p1	14.225			
p2	-3.774	12.593		
p3	0	0	0.055	
p4	0	0	0	0.054

The zero covariances reflect the fact that this is a form of LISEM in which only p1 and p2 were estimated in the same equation. I input this matrix into the Monte Carlo confidence interval program on a per column basis.

SPSS, STATA, R

To use OLS-based LISEM in a standard software package, you will conduct three separate regression equations as I discuss in Chapter 8:

$$Income_{t3} = a_1 + b_1\ Budget_{t2} + b_2\ Credit_{t2} + b_3\ Income_{t1} + d_3 \quad [1]$$

$$Budget_{t2} = a_2 + b_4\ TREAT + b_5\ Budget_{t1} + d_1 \quad [2]$$

$$Credit_{t2} = a_3 + b_6\ TREAT + b_7\ Credit_{t1} + d_2 \quad [3]$$

For SPSS, in the "Linear Regression" window, click on the "Statistics" button. In the section for "Regression Coefficients," check the box for "Covariance matrix." The ACM will appear clearly labeled on the output. If two coefficients in the original equations were estimated in different equations, set their covariance in the ACM to 0 (like MIIVsem above).

For STATA, the regressions can be conducted using the regress command and the post-estimation command estat vce can be used to obtain the variance-covariance matrix of the predictors.

For R, here are the commands for Equation 1:

```
reg<-lm(income3~budget2+credit2+income1)  # do the regression
summary(reg) # show the results f the regression
vcov(reg)    # calculate and show the ACM
```