Maximum Likelihood Equal-Variances Ratio Test and Estimation of Moderator Effects On Correlation with SAS

Michael Smithson
Department of Psychology, The Australian National University

(email: Michael.Smithson@anu.edu.au)

SAS NLMIXED syntax components

The model in Smithson (2012) can be run in SAS under its NLMIXED procedure. SAS requires three main components for these models:

1. Formulas for the submodels,
2. A formula for the negative log-likelihood kernel, which is the loss-function to be minimized, and
3. Names and starting-values for the model parameters.

Formulas for the submodels and negative log-likelihood kernel

The dependent variable must be the variable listed in the `model Y ~ general(ll)` subcommand. The submodels are identified by the red font in the generic code shown below. The first two are the "nuisance parameter" submodels for the means of X and Y. The remaining three are the submodels for the standard deviations and the correlation. The negative log-likelihood kernel formula is identified by the **bold** text.

```
title ' ';
data [NameofData];
input [variable list];
cards;
 [data goes here]
;
run;
;
proc nlmixed data = [NameofData] tech = trureg hess corr itdetails;
*Starting values;
parms MNX =  , MNY =  , DX0 =, DY0 =, DR0 =;
title ' ';
*This is the model;
MUX = MNX;
MUY = MNY;
DEVX = X - MUX;
DEVY = Y - MUY;
DX = DX0;
DY = DY0;
SX = exp(DX);
SY = exp(DY);
RHO = (exp(DR0)-1)/(exp(DR0)+1);
ll = ((SX**2)*(DEVY**2) - 2*RHO*SX*SY*DEVX*DEVY +
(SY**2)*(DEVX**2))/(2*(RHO**2-1)*(SX**2)*(SY**2)) - log(2*3.145927) -
0.5*(log(1 - RHO**2) + log(SX**2) + log(SY**2));
model Y ~ general(ll);
run;
```

Moderator and HeV effects may be added to the submodels in a straightforward way. If our moderator variable is Z, then a moderator effect on σ_y with coefficient DY1 would be represented in the submodel for the standard deviation of Y as follows:

```
DY = DY0 + DY1*Z;
```

If we also add a first-order moderation effect with coefficient DR1 then the submodel for the correlation would look like this:

```
RHO = (exp(DR0+DR1*Z)-1)/(exp(DR0+DR1*Z)+1);
```

Model parameters and starting-values

It is important to get good starting values for these models. If Z is categorical, the best way to generate starting-values for the parameters is to use method of moments estimation. There are functions for doing this in R, available from the same webpage linked to this document. Alternatively, you can do so using SAS or a spreadsheet program such as Excel.

Equal variance ratios tests

The model syntax described above is readily adapted to an EVR test by comparing models. The null model is estimated by assigning the same term to the moderator effect of Z on σ_x and σ_y. The syntax fragment below does this via the coefficient labeled D1.

```
DX = DX0 + D1*Z;
DY = DY0 + D1*Z;
```

The alternative model assigns separate coefficients to Z in the DX and DY submodels:

```
DX = DX0 + DX1*Z;
DY = DY0 + DY1*Z;
```

Two-category moderator example

This is a two-category moderator example in which the null hypothesis of EVR is true. It uses an artificial data-set sampled from two bivariate normal distributions, associated with one category of a binary moderator variable *Z* which takes values -1 and +1. For the first moderator category $\sigma_x^2 = 1$ and $\sigma_y^2 = 2$, while for the second category $\sigma_x^2 = 4$ and $\sigma_y^2 = 8$. Thus, the population variance ratio in both moderator categories is 1/2. The population covariances are 1 and the means are 0 for both categories. The file with SAS code and data is EVR_SAScode.txt. There are 500 observations in each category. In the first category, s_x^2 = 0.883 and s_y^2 = 1.958, so s_x^2/s_y^2 = 0.451. In the second category, s_x^2 = 4.068 and s_y^2 = 7.515, so s_x^2/s_y^2 = 0.541.

First, we fit a saturated model permitting unequal variance ratios:

```
title 'no EVR';
data eqdat;
input X Y Z;
cards;
0.620217031 1.325695514 -1
… [rest of data here]
2.416855609 -4.904633331      1
;
run;
;
proc nlmixed data = eqdat tech = trureg hess corr itdetails;
*Starting values;
parms MNX =  .01, MNY =  0.001, DX0 = 0.01, DY0 = 0.36, DR0 = 1.0, DX1 =
0.01, DY1 = 0.01, DR1 = -0.3;
title 'EVR false';
*This is the model;
MUX = MNX;
MUY = MNY;
DEVX = X - MUX;
DEVY = Y - MUY;
DX = DX0 + DX1*Z;
DY = DY0 + DY1*Z;
SX = exp(DX);
```

```
SY = exp(DY);
RHO = (exp(DR0 + DR1*Z)-1)/(exp(DR0 + DR1*Z)+1);
ll = ((SX**2)*(DEVY**2) - 2*RHO*SX*SY*DEVX*DEVY +
(SY**2)*(DEVX**2))/(2*(RHO**2-1)*(SX**2)*(SY**2)) - log(2*3.145927) -
0.5*(log(1 - RHO**2) + log(SX**2) + log(SY**2));
model Y ~ general(ll);
run;
```

The relevant output is shown in the tables below.

Fit Statistics

-2 Log Likelihood	7255.7
AIC (smaller is better)	7271.7
AICC (smaller is better)	7271.9
BIC (smaller is better)	7311.0

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper	Gradient
MNX	0.01905	0.03680	1000	0.52	0.6048	0.05	-0.05317	0.09127	2.36E-11
MNY	0.02558	0.05433	1000	0.47	0.6379	0.05	-0.08103	0.1322	2.87E-11
DX0	0.3187	0.02236	1000	14.25	<.0001	0.05	0.2748	0.3626	-1.65E-9
DY0	0.6716	0.02236	1000	30.03	<.0001	0.05	0.6277	0.7154	-1.51E-9
DR0	1.0969	0.06325	1000	17.34	<.0001	0.05	0.9728	1.2210	1.452E-9
DX1	0.3820	0.02236	1000	17.08	<.0001	0.05	0.3381	0.4258	-1.84E-9
DY1	0.3363	0.02237	1000	15.04	<.0001	0.05	0.2924	0.3802	-1.76E-9
DR1	-0.7905	0.06326	1000	-12.50	<.0001	0.05	-0.9146	-0.6664	1.443E-9

Correlation Matrix of Parameter Estimates

Row	Parameter	MNX	MNY	DX0	DY0	DR0	DX1	DY1	DR1
1	MNX	1.0000	0.6783	-0.00718	-0.00149	-0.00056	0.01827	0.01849	0.02010
2	MNY	0.6783	1.0000	-0.00487	-0.00219	-0.00024	0.01239	0.02726	0.01671
3	DX0	-0.00718	-0.00487	1.0000	0.2831	0.3143	-0.00013	-0.2600	-0.2069
4	DY0	-0.00149	-0.00219	0.2831	1.0000	0.3143	-0.2600	-0.00006	-0.2068
5	DR0	-0.00056	-0.00024	0.3143	0.3143	1.0000	-0.2068	-0.2067	-0.00001
6	DX1	0.01827	0.01239	-0.00013	-0.2600	-0.2068	1.0000	0.2833	0.3145
7	DY1	0.01849	0.02726	-0.2600	-0.00006	-0.2067	0.2833	1.0000	0.3146
8	DR1	0.02010	0.01671	-0.2069	-0.2068	-0.00001	0.3145	0.3146	1.0000

The next model restricts the ratios to be equal:

```
proc nlmixed data = eqdat tech = trureg hess corr itdetails;
*Starting values;
parms MNX =  .01, MNY =  0.001, DX0 = 0.01, DY0 = 0.36, DR0 = 1.0, D1 =
0.01, DR1 = -0.3;
title 'EVR true';
*This is the model;
MUX = MNX;
MUY = MNY;
DEVX = X - MUX;
DEVY = Y - MUY;
DX = DX0 + D1*Z;
DY = DY0 + D1*Z;
SX = exp(DX);
SY = exp(DY);
RHO = (exp(DR0 + DR1*Z)-1)/(exp(DR0 + DR1*Z)+1);
ll = ((SX**2)*(DEVY**2) - 2*RHO*SX*SY*DEVX*DEVY +
(SY**2)*(DEVX**2))/(2*(RHO**2-1)*(SX**2)*(SY**2)) - log(2*3.145927) -
0.5*(log(1 - RHO**2) + log(SX**2) + log(SY**2));
model Y ~ general(ll);
run;
```

The relevant output is shown below:

```
                     Fit Statistics

                          -2 Log Likelihood              7258.6
                          AIC (smaller is better)        7272.6
                          AICC (smaller is better)       7272.7
                          BIC (smaller is better)        7307.0
```

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper	Gradient
MNX	0.01905	0.03720	1000	0.51	0.6087	0.05	-0.05395	0.09206	2.19E-11
MNY	0.02670	0.05383	1000	0.50	0.6200	0.05	-0.07894	0.1323	2.73E-11
DX0	0.3110	0.02172	1000	14.32	<.0001	0.05	0.2684	0.3536	-1.78E-9
DY0	0.6804	0.02196	1000	30.99	<.0001	0.05	0.6373	0.7235	-1.42E-9
DR0	1.0959	0.06325	1000	17.33	<.0001	0.05	0.9718	1.2201	1.477E-9
D1	0.3595	0.01792	1000	20.06	<.0001	0.05	0.3244	0.3947	-3.47E-9
DR1	-0.7901	0.06327	1000	-12.49	<.0001	0.05	-0.9143	-0.6659	1.362E-9

The chi-square difference is 7258.6-7255.7 = 2.9, close to the corresponding Mplus (v. 6.12) SEM chi-square value of 2.920.
The estimated variance ratio is
(exp(DX0 – DY0))^2 = (exp(-.369))^2 = 0.478,
agreeing with the Mplus estimate of .478.
Correlation estimates from the SAS output:
r_1 = (exp(1.096-.790)-1)/(exp(1.096-.790)+1) = .152
r_2 = (exp(1.096+.790)-1)/(exp(1.096+.790)+1)= .737
The Mplus estimates are .151 and .737, so again agreement is very close.

SEM example from Smithson (2012)
The data-set for this example is as described in Smithson (2012), and the file is semex_SAScode.txt. Starting with SEMs for regression coefficients, we begin with a model that allows moderator effects on the variances and correlation with no restrictions on the parameters.

```
proc nlmixed data = semex tech = trureg hess corr itdetails;
*Starting values;
parms MNX =  .01, MNY =  0.001, DX0 = 0.01, DY0 = 0.4, DR0 = 1.0, DX1 =
0.01, DY1 = 0.01, DR1 = -0.3;
title 'EVR false';
*This is the model;
MUX = MNX;
MUY = MNY;
DEVX = X - MUX;
DEVY = Y - MUY;
DX = DX0 + DX1*Z;
DY = DY0 + DY1*Z;
SX = exp(DX);
SY = exp(DY);
RHO = (exp(DR0 + DR1*Z)-1)/(exp(DR0 + DR1*Z)+1);
ll = ((SX**2)*(DEVY**2) - 2*RHO*SX*SY*DEVX*DEVY +
(SY**2)*(DEVX**2))/(2*(RHO**2-1)*(SX**2)*(SY**2)) - log(2*3.145927)
- 0.5*(log(1 - RHO**2) + log(SX**2) + log(SY**2));
model Y ~ general(ll);
run;
```

The relevant output is:

```
                 Fit Statistics
-2 Log Likelihood                  3681.4
AIC (smaller is better)            3697.4
AICC (smaller is better)           3697.6
BIC (smaller is better)            3732.5
```

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper	Gradient
MNX	0.002212	0.03903	600	0.06	0.9548	0.05	-0.07443	0.07885	-3.52E-7
MNY	-0.00541	0.05378	600	-0.10	0.9199	0.05	-0.1110	0.1002	-1.11E-6
DX0	-0.03282	0.02887	600	-1.14	0.2560	0.05	-0.08952	0.02387	-5.94E-6
DY0	0.3810	0.02887	600	13.20	<.0001	0.05	0.3243	0.4377	-0.00002
DR0	0.9978	0.08165	600	12.22	<.0001	0.05	0.8374	1.1581	-2.89E-6
DX1	-0.01744	0.02889	600	-0.60	0.5464	0.05	-0.07418	0.03931	5.217E-6
DY1	0.3168	0.02890	600	10.96	<.0001	0.05	0.2600	0.3735	0.000015
DR1	0.006576	0.08172	600	0.08	0.9359	0.05	-0.1539	0.1671	2.598E-6

We now restrict the variance ratios to be equal:

```
proc nlmixed data = semex tech = trureg hess corr itdetails;
*Starting values;
parms MNX =  .01, MNY =  0.001, DX0 = 0.01, DY0 = 0.36, DR0 = 1.0, D1 =
0.01, DR1 = -0.3;
title 'EVR true';
*This is the model;
MUX = MNX;
MUY = MNY;
DEVX = X - MUX;
DEVY = Y - MUY;
DX = DX0 + D1*Z;
DY = DY0 + D1*Z;
SX = exp(DX);
SY = exp(DY);
RHO = (exp(DR0 + DR1*Z)-1)/(exp(DR0 + DR1*Z)+1);
ll = ((SX**2)*(DEVY**2) - 2*RHO*SX*SY*DEVX*DEVY +
(SY**2)*(DEVX**2))/(2*(RHO**2-1)*(SX**2)*(SY**2)) - log(2*3.145927)
- 0.5*(log(1 - RHO**2) + log(SX**2) + log(SY**2));
model Y ~ general(ll);
run;
```

The relevant output is:

```
                 Fit Statistics
-2 Log Likelihood                  3763.7
AIC (smaller is better)            3777.7
AICC (smaller is better)           3777.8
BIC (smaller is better)            3808.4
```

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper	Gradient
MNX	0.007365	0.03975	600	0.19	0.8531	0.05	-0.07069	0.08542	-1.61E-8
MNY	-0.01962	0.06019	600	-0.33	0.7445	0.05	-0.1378	0.09858	-1.89E-8
DX0	-0.00596	0.02989	600	-0.20	0.8419	0.05	-0.06466	0.05273	-2.24E-6
DY0	0.4089	0.02991	600	13.67	<.0001	0.05	0.3501	0.4676	-1.32E-6
DR0	0.9363	0.08165	600	11.47	<.0001	0.05	0.7760	1.0967	-7.19E-7
D1	0.1494	0.02321	600	6.44	<.0001	0.05	0.1038	0.1950	3.132E-6
DR1	0.006478	0.08288	600	0.08	0.9377	0.05	-0.1563	0.1693	6.234E-7

The chi-square difference is 3763.7-3681.4 = 82.3, quite close to Mplus' 82.246.

We can reject the EVR hypothesis.

We now estimate a model with HoV for *X* and moderator effects for the variance of *Y* and the correlation:

```
proc nlmixed data = semex tech = trureg hess corr itdetails;
*Starting values;
parms MNX =  .01, MNY =  0.001, DX0 = 0.01, DY0 = 0.36, DR0 = 1.0, DY1 =
0.01, DR1 = -0.3;
title 'HoV for X';
*This is the model;
MUX = MNX;
MUY = MNY;
DEVX = X - MUX;
DEVY = Y - MUY;
DX = DX0;
DY = DY0 + DY1*Z;
SX = exp(DX);
SY = exp(DY);
RHO = (exp(DR0 + DR1*Z)-1)/(exp(DR0 + DR1*Z)+1);
ll = ((SX**2)*(DEVY**2) - 2*RHO*SX*SY*DEVX*DEVY +
(SY**2)*(DEVX**2))/(2*(RHO**2-1)*(SX**2)*(SY**2)) - log(2*3.145927)

- 0.5*(log(1 - RHO**2) + log(SX**2) + log(SY**2));
model Y ~ general(ll);
run;
```

The relevant output is:

Fit Statistics

-2 Log Likelihood	3681.7
AIC (smaller is better)	3695.7
AICC (smaller is better)	3695.9
BIC (smaller is better)	3726.5

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper	Gradient
MNX	0.003227	0.03901	600	0.08	0.9341	0.05	-0.07339	0.07984	-1.69E-7
MNY	-0.00480	0.05378	600	-0.09	0.9290	0.05	-0.1104	0.1008	-2.95E-7
DX0	-0.03251	0.02887	600	-1.13	0.2605	0.05	-0.08921	0.02418	-9.63E-6
DY0	0.3811	0.02888	600	13.20	<.0001	0.05	0.3244	0.4378	-0.00001
DR0	0.9983	0.08168	600	12.22	<.0001	0.05	0.8379	1.1587	-3.9E-6
DY1	0.3205	0.02825	600	11.35	<.0001	0.05	0.2650	0.3760	2.565E-6
DR1	0.02272	0.07724	600	0.29	0.7688	0.05	-0.1290	0.1744	7.31E-7

The chi-square difference is 3681.7-3681.4 = 0.3, fairly close to Mplus' 0.370.
We retain the HoV for *X* hypothesis.

Finally, we estimate a model with HoV for *X* and equal correlations:

```
proc nlmixed data = semex tech = trureg hess corr itdetails;
*Starting values;
parms MNX =  0.001, MNY =  0.001, DX0 = -0.03, DY0 = 0.38, DR0 = 1.0, DY1 =
0.32;
title 'HoV X and equal corr';
*This is the model;
MUX = MNX;
MUY = MNY;
DEVX = X - MUX;
DEVY = Y - MUY;
DX = DX0;
```

```
DY = DY0 + DY1*Z;
SX = exp(DX);
SY = exp(DY);
RHO = (exp(DR0)-1)/(exp(DR0)+1);
ll = ((SX**2)*(DEVY**2) - 2*RHO*SX*SY*DEVX*DEVY +
(SY**2)*(DEVX**2))/(2*(RHO**2-1)*(SX**2)*(SY**2)) - log(2*3.145927)

- 0.5*(log(1 - RHO**2) + log(SX**2) + log(SY**2));
model Y ~ general(ll);
run;
```

The relevant output is:

Fit Statistics

-2 Log Likelihood	3681.8
AIC (smaller is better)	3693.8
AICC (smaller is better)	3694.0
BIC (smaller is better)	3720.2

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper	Gradient
MNX	0.002940	0.03907	600	0.08	0.9400	0.05	-0.07379	0.07966	-4.89E-7
MNY	-0.00518	0.05376	600	-0.10	0.9233	0.05	-0.1107	0.1004	9.425E-7
DX0	-0.03252	0.02887	600	-1.13	0.2605	0.05	-0.08921	0.02418	-3.33E-7
DY0	0.3810	0.02887	600	13.20	<.0001	0.05	0.3243	0.4377	-1.85E-6
DR0	0.9975	0.08165	600	12.22	<.0001	0.05	0.8371	1.1579	5.751E-7
DY1	0.3182	0.02712	600	11.73	<.0001	0.05	0.2649	0.3714	1.342E-6

The chi-square difference is 3681.8-3681.7 = 0.1, close to Mplus' 0.083. We conclude that there is no moderation of correlations. The estimate of the correlation is
r = (exp(0.9975)-1)/(exp(0.9975)+1) = .461,
which is close to the Mplus estimate of .460.

Four-category moderator example

Now we consider a four-category moderator example, with EVR for the first two categories and for the second two, but not for both pairs of categories. An artificial data-set has been sampled from four bivariate normal distributions, associated with a four-category moderator variable Z, with 300 observations in each category. The variance ratio for the first two moderator categories equals 1/2, whereas for the third and fourth categories the ratio is 1/6. The first two categories' correlations also are identical. The data-set is named fourgroup_SAScode.txt. An appropriate design matrix for testing the four-groups model is shown below.

	σ_y dummy vars.			σ_x dummy vars.		
	z2	z3	z4	z2	z3	z4
Group 1	0	0	0	0	0	0
Group 2	1	0	0	1	0	0
Group 3	0	1	0	0	1	0
Group 4	0	1	1	0	1	1

The models to be compared are:

Mod1: $\log(\sigma_y) = \delta_{y0} + \delta_2 + \delta_{y3} + \delta_4$

$\log(\sigma_x) = \delta_{x0} + \delta_2 + \delta_{x3} + \delta_4$

$\text{logit}(\rho) = \delta_{r0} + \delta_{r3} + \delta_{r4}$

Mod2: $\log(\sigma_y) = \delta_{y0} + \delta_{y2} + \delta_{y3} + \delta_{y4}$
$\log(\sigma_x) = \delta_{x0} + \delta_{x2} + \delta_{x3} + \delta_{x4}$
$\text{logit}(\rho) = \delta_{r0} + \delta_{r3} + \delta_{r4}$

Mod3: $\log(\sigma_y) = \delta_{y0} + \delta_{y2} + \delta_{y3} + \delta_{y4}$
$\log(\sigma_x) = \delta_{x0} + \delta_{x2} + \delta_{x3} + \delta_{x4}$
$\text{logit}(\rho) = \delta_{r0} + \delta_{r2} + \delta_{r3} + \delta_{r4}$

Mod1 includes both EVR hypotheses and the equal-correlations hypothesis. Mod2 relaxes the EVR hypothesis and Mod3 relaxes both that and the equal-correlations hypothesis. Syntax for Mod1 is shown below. The other models' syntax can be developed by modifying this syntax in the obvious ways.

```
title 'four groups';
data fourgp;
input X Y group Z2 Z3 Z4;
cards;
-0.479855311      -0.272689233      1     0     0     0
… [rest of the data here]
0.664409246 0.838487134 4     0     1     1
;
run;
;
proc nlmixed data = fourgp tech = trureg hess corr itdetails;
*Starting values;
parms MNX =  .01, MNY =  0.001, DX0 = 0.03, DY0 = .4, DR0 = 1.8, D2 = .18,
DX3 = -.02, DY3 = -.5, DR3 = .5, D4 = 0.3, DR4 = -1.0;
title 'four group EVR model';
*This is the model;
DEVX = X - MNX;
DEVY = Y - MNY;
SX = exp(DX0 + D2*Z2 + DX3*Z3 + D4*Z4);
SY = exp(DY0 + D2*Z2 + DY3*Z3 + D4*Z4);
RHO = (exp(DR0 + DR3*Z3 + DR4*Z4)-1)/(exp(DR0 + DR3*Z3 + DR4*Z4)+1);
ll = ((SX**2)*(DEVY**2) - 2*RHO*SX*SY*DEVX*DEVY +
(SY**2)*(DEVX**2))/(2*(RHO**2-1)*(SX**2)*(SY**2)) - log(2*3.145927) -
0.5*(log(1 - RHO**2) + log(SX**2) + log(SY**2));
model Y ~ general(ll);
run;
```

The table below shows -LL and chi-square differences for the three models. The df differences between Mod1 and Mod2 , Mod1 and Mod3, and Mod2 and Mod3 are 2, 3, and 1 respectively. Thus, none of the model comparisons yield significant differences in model fit, so Mod1 is retained as the most parsimonious model with acceptable fit.

	-2LL	Mod1 vs:	Mod2 vs:
Mod1	8050.5		
Mod2	8050.2	0.3	
Mod3	8048.3	2.2*	1.9

* This is similar to the corresponding SEM chi-square 2.25 from Mplus.

The Mod1 estimates closely match their counterparts from Mplus. The variance ratio estimate for the first two moderator categories is $(\exp(.026-.367))^2 = 0.5056$, agreeing with the Mplus estimate of 0.506. The variance ratio for the last two categories is $(\exp(.026-.367-.021-.508))^2$

= 0.1755, agreeing with the Mplus estimate of 0.176. The table below compares the remaining parameter estimates from SAS and Mplus in each of the four moderator categories, demonstrating close agreement between the packages.

Categ.	Package	s_y	s_x	r
1	Mplus	1.441	1.025	.734
	SAS	1.444	1.026	.734
2	Mplus	1.731	1.231	.734
	SAS	1.732	1.231	.734
3	Mplus	2.395	1.005	.833
	SAS	2.399	1.005	.832
4	Mplus	3.321	1.393	.495
	SAS	3.330	1.395	.496

Reference

Smithson, M. (2012) *A simple statistic for comparing moderation of slopes and correlations*.