Repeated Measures with `proc mixed`

In a repeated measures research design, also called *within-subjects* or *longitudinal*, the dependent variable is measured on more than one occasion for each case (there are *n* cases). Predictor variables can be measured repeatedly over time, too. Sometimes the subjects experience a different experimental treatment every time the dependent variable is measured. This is typical of laboratory studies. Other times, the variable of interest is just tracked over time. This is typical of longitudinal studies in epidemiology. But in general, you do not expect observations coming from the same individual to be independent. The dependence has to be modeled somehow.

In the world of proc mixed, dependent variables are normally distributed. Observations coming from the same case have a multivariate normal distribution, with the same covariance matrix $\boldsymbol{\Sigma}$ for each case.

Proc mixed allows you to specify various structures for the covariance matrix. Some of these appear on the following page.

Variance Components: type = vc

$$\Sigma = \begin{vmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{vmatrix}$$

Compound Symmetry: type = cs

$$\Sigma = \begin{vmatrix} \sigma^2 + \sigma_1 & \sigma_1 & \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma^2 + \sigma_1 & \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma_1 & \sigma^2 + \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma_1 & \sigma_1 & \sigma^2 + \sigma_1 \end{vmatrix}$$

Unknown: type = un

$$\Sigma = \begin{vmatrix} \sigma_1^2 & \sigma_{1,2} & \sigma_{1,3} & \sigma_{1,4} \\ \sigma_{1,2} & \sigma_2^2 & \sigma_{2,3} & \sigma_{2,4} \\ \sigma_{1,3} & \sigma_{2,3} & \sigma_3^2 & \sigma_{3,4} \\ \sigma_{1,4} & \sigma_{2,4} & \sigma_{3,4} & \sigma_4^2 \end{vmatrix}$$

Banded: type =

$$\Sigma = \begin{vmatrix} \sigma_1^2 & \sigma_5 & 0 & 0 \\ \sigma_5 & \sigma_2^2 & \sigma_6 & 0 \\ 0 & \sigma_6 & \sigma_3^2 & \sigma_7 \\ 0 & 0 & \sigma_7 & \sigma_4^2 \end{vmatrix}$$

First order autoregressive: type = ar(1)

$$\Sigma = \sigma^2 \begin{vmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{vmatrix}$$

There are more, including Toeplitz, banded Toeplitz & spatial (covariance is a function of Euclidian distance).

In this example done by the Consulting Service in 2005, the dependent variable was hunger. There was a treatment Group, which is a between-subjects factor, and two within-subjects factors. The within-subjects factors are Drink (with values 1, 2 and 3: maybe representing different amounts) and time (values are 0, 15, 30 and 60). So hunger was measured for each subject on 4x3=12 occasions. The original raw data file had a *multivariate setup*, meaning that it had the standard row-by-column format, with one row of data for each subject. The variables on each row were something like this

```
Subject
Group
Hunger1  (with Drink=1 and Time=0)
Hunger2  (with Drink=1 and Time=15)
Hunger3  (with Drink=1 and Time=30)
Hunger4  (with Drink=1 and Time=60)
Hunger5  (with Drink=2 and Time=0)
Hunger6  (with Drink=2 and Time=15)
Hunger7  (with Drink=2 and Time=30)
Hunger8  (with Drink=2 and Time=60)
Hunger9  (with Drink=3 and Time=0)
Hunger10 (with Drink=3 and Time=15)
Hunger11 (with Drink=3 and Time=30)
Hunger12 (with Drink=3 and Time=60)
```

Unfortunately, proc glm requires a setup like this (though of course more than one line of data per case is okay), but proc mixed can't use it. So the raw data were read into a SAS data set, and then a new SAS data set was created in which each observation of the dependent variable resides on a separate case. Yes, there are 12*n* cases. I'll show you how it was done at the end. The new data set has a structure shown on the following page.

The data set `hungermixed` looks like this:

Subj	Group	Drink	Time	Hunger
1	1	1	0	6.3
1	1	1	15	4.2
1	1	1	30	7.7
1	1	1	60	2.8
1	1	2	0	4.2
1	1	2	15	6.0
1	1	2	30	7.1
1	1	2	60	8.3
1	1	3	0	9.5
1	1	3	15	4.7
1	1	3	30	3.3
1	1	3	60	5.5
2	1	1	0	6.3
2	1	1	15	7.8
2	1	1	30	5.9
2	1	1	60	8.1
2	1	2	0	9.0
2	1	2	15	1.3
2	1	2	30	2.6
2	1	2	60	4.5
2	1	3	0	6.9
2	1	3	15	2.7
2	1	3	30	7.5
2	1	3	60	8.8
		...		
40	2	1	0	4.6
40	2	1	15	3.3
40	2	1	30	7.8
40	2	1	60	8.1
40	2	2	0	2.0
40	2	2	15	1.7
40	2	2	30	4.4
40	2	2	60	8.8
40	2	3	0	3.2
40	2	3	15	5.7
40	2	3	30	6.5
40	2	3	60	9.0

Here's the code:

```
proc mixed data=hungermixed;
  class group subj drink time;
  model hunger=  group|drink|time / outp=resids;
  repeated/type=ar(1) sub=subj;
```

Here's how the data set hungermixed was created. First, the original data were read into a data set called hungerglm. Then, still in the data step,

```
data hungermixed;
  set hungerglm;
  drink = 1; time = 0; hunger = hunger1; output;
  drink = 1; time = 15; hunger = hunger2; output;
  drink = 1; time = 30; hunger = hunger3; output;
  drink = 1; time = 60; hunger = hunger4; output;
  drink = 2; time = 0; hunger = hunger5; output;
  drink = 2; time = 15; hunger = hunger6; output;
  drink = 2; time = 30; hunger = hunger7; output;
  drink = 2; time = 60; hunger = hunger8; output;
  drink = 3; time = 0; hunger = hunger9; output;
  drink = 3; time = 15; hunger = hunger10; output;
  drink = 3; time = 30; hunger = hunger11; output;
  drink = 3; time = 60; hunger = hunger12; output;
  drop hunger1-hunger12;
  /* I'm dropping the old variables because I don't need them
     anymore.  */
```

Noise

Here is part of noise.dat

```
   1  2.5  1  2  1  4  50.7
   1  2.5  1  2  2  1  27.4
   1  2.5  1  2  3  3  39.1
   1  2.5  1  2  4  2  37.5
   1  2.5  1  2  5  5  35.4
   2  1.9  1  2  1  3  40.3
   2  1.9  1  2  2  1  30.1
   2  1.9  1  2  3  5  38.9
   2  1.9  1  2  4  2  31.9
   2  1.9  1  2  5  4  31.6
```

```
/**************** snoise.sas ***********************/
options linesize=79 noovp pagesize=250 formdlim='_';
title 'Repeated measures on Noise data:  Try Spatial Autocorrelation';
proc format;      value sexfmt    0 = 'Male'  1 = 'Female' ;

data loud;
     infile 'noise.dat';  /* Univariate data read */
     input ident  interest  sex  age  noise  time  discrim ;
     format sex sexfmt.;
     label interest = 'Interest in topic (politics)'
           time     = 'Order of presenting noise level';

/* As of SAS version 9.1, eight spatial autocorrelation structures were
available. The one that reduces to ar(1) is called spatial power. The last
argument in parentheses is a LIST of variables representing the spatial
co-ordinates. Spatial autocorrelation dies out exponentially with Euclidean
distance. */


proc mixed;
     title2 'Spatial power on order = ar(1) on actual time';
     class age sex noise;
     model discrim = age|sex|noise;
     repeated / type = sp(pow)(time) subject = ident r;

/* Note: This is very promising for unequally spaced repeated measures data,
         too! */
```

Pliner's Eating Study

O.K., remember the analysis you were going to do for us-a repeated measures with different covariates for the two different occasions. Well, here's a dataset.

So, here's the study. We're interested in the extent to which people conform in terms of amount eaten to the behavior of a confederate who is pre-programmed to eat a lot or a little. Actually, we already know that that has a huge effect on how much people eat. People eat a lot with the big eater and a little with the small eater. So, here's the spin. What if you eat with a model in Session 1 and then a day or so later you eat in the same situation alone--without a model? Will the "norm" established in Session 1 affect people's behavior in Session 2-i.e., will they internalize the norm? Now, here's an additional spin. What if people eat alone in Session 1 and maybe establish their own "norm." If you then put them with a big- or small-eating confederate in Session 2, will they be immune to the confederate's norm because they have already established their own? In other words, for some Ss the sessions occur in an alone-confederate order and for others, the sessions occur in a confederate-alone order. P.S. if this sounds like the old Sherif study, it's not a co-incidence.

Basically, we've already analyzed the data and we find more or less that a) the norm established in Session 1 carries over to Session 2, even in the absence of the confederate, and b) eating alone in Session 1 doesn't make you immune to the effect of the confederate in Session 2.

One slight annoyance is that there were some differences (nearly significant) between conditions in terms of how "full" Ss said they were before the sessions began and in terms of the number of hours of food deprivation prior to coming to the sessions. So, that's what we want to use as covariates-with each session having as its covariate the person's particular level of fullness or deprivation for that session. Is this clear?

So, I'm including a dataset (an Excel file) for you with the relevant variables as follows:

cdnOrder: the order in which the sessions occurred (0 = alone-confederate; 1 = confederate-alone)

cdnNorm: whether in the confed condition, the confed ate a little or a lot (0= a little; 1 = a lot)

amtEatC: amount eaten in the session in which the confederate was present

amtEatA: amount eaten in the session in which the subject ate alone

fullC: rated fullness in the session in which the confederate was present

fullA: rated fullness in the session in which the subject ate alone

deptimeC: hours of food deprivation prior to the session in which the confederate was present

deptimeA: hours of food deprivation prior to the session in which the subject ate alone

So, basically, we're talking about split plot analyses with cdnOrder and cdnNorm as between-subjects independent variables and confederate/alone as the within-subjects variable (amtEatC and amtEatA as the dvs) and fullC/fullA as the covariates.

And the same thing with deptimeC/deptimeA as covariates.

Does this all make sense? Let me know if it doesn't.

Patty

```
/dos/brunner/consult/pliner > less RepeatPliner.data

ID,cdnNorm,cdnOrder,ResEater,amtEatC,amtEatA,fullc,fulla,deptimec,deptimea
8,0.0,0.0,1.0,16.00,11.00,2.00,4.00,2.30,4.30
102,0.0,0.0,1.0,12.00,12.00,1.00,1.00,17.25,5.25
106,0.0,0.0,2.0,12.00,13.00,1.00,1.00,2.90,3.40
12,0.0,0.0,1.0,11.00,11.00,1.00,1.00,4.75,7.25
14,0.0,0.0,1.0,10.00,7.00,1.00,1.00,18.12,5.12


/* eatrepeat.sas */
options linesize=79 pagesize=500 noovp formdlim='_';
title 'Social facilitation in eating: ANCOVA with time-varying covariates';

proc format;
     value ofmt 0 = 'AloneConfederate' 1 = 'ConfederateAlone';
     value amtfmt 0 = 'A little'  1 = 'A lot';


data hungry;
     infile 'RepeatPliner.data' firstobs=2 delimiter=',';
     input ID cdnNorm cdnOrder ResEater amtEatC amtEatA fullC fullA deptimeC
           deptimeA;
     label
           cdnOrder = 'Order'
           cdnNorm  = 'Amount eaten by Confederate'
           amtEatC  = 'Amount eaten with confederate present'
           amtEatA  = 'Amount eaten with confederate absent'
           fullC    = 'Rated fullness before eating, confederate present'
           fullA    = 'Rated fullness before eating, confederate absent'
           deptimeC = 'Deprivation time, confederate present'
           deptimeA = 'Deprivation time, confederate absent';
      format cdnOrder ofmt.;
      format  cdnNorm amtfmt.;

proc means;
     title2 'Describe Quantitative Variables';
     var amtEatC amtEatA fullc fulla deptimec deptimea;

proc freq;
     title2 'Frequency distributions and crosstabs';
     tables cdnOrder * (cdnNorm ResEater) / norow nocol nopercent;
     tables cdnNorm*ResEater / norow nocol nopercent;

proc freq;
     tables fullc fulla;
```

```
data fressen; /* Separate cases for confederate present, absent */
     set hungry;
     confed = 'Present';
          amtEat = amtEatC; full = fullC; deptime = deptimeC; output;
     confed = 'Absent';
          amtEat = amtEatA; full = fullA; deptime = deptimeA; output;
     keep ID cdnNorm cdnOrder ResEater confed amtEat full deptime;
     label
           cdnOrder = 'Order'
           cdnNorm  = 'Amount eaten by Confederate'
           amtEat   = 'Amount eaten'
           full     = 'Rated fullness before eating'
           deptime  = 'Food Deprivation time';
      format cdnOrder ofmt.;
      format  cdnNorm amtfmt.;

proc print;

proc corr;
     var deptime full amtEat;

proc mixed;
     title2 'Covariates: Deprivation time and rated fullness';
     class cdnOrder cdnNorm confed;
     model amtEat = deptime full cdnOrder|cdnNorm|confed;
     repeated / type=un subject=id r;
     lsmeans cdnOrder|cdnNorm|confed;

proc mixed;
     title2 'Covariate: Rated fullness';
     class cdnOrder cdnNorm confed;
     model amtEat = full cdnOrder|cdnNorm|confed;
     repeated / type=un subject=id r;
     lsmeans cdnOrder|cdnNorm|confed;

proc mixed;
     title2 'Covariate: Deprivation time';
     class cdnOrder cdnNorm confed;
     model amtEat = deptime cdnOrder|cdnNorm|confed;
     repeated / type=un subject=id r;
     lsmeans cdnOrder|cdnNorm|confed;
```

Now just part of the list file: Just the proc mixed output for covariate = fullness

The Mixed Procedure

Model Information

Data Set	WORK.FRESSEN
Dependent Variable	amtEat
Covariance Structure	Unstructured
Subject Effect	ID
Estimation Method	REML
Residual Variance Method	None
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Between-Within

Class Level Information

Class	Levels	Values
cdnOrder	2	AloneConfederate ConfederateAlone
cdnNorm	2	A little A lot
confed	2	Absent Present

Dimensions

Covariance Parameters	3
Columns in X	28
Columns in Z	0
Subjects	74
Max Obs Per Subject	2

Number of Observations

Number of Observations Read	148
Number of Observations Used	148
Number of Observations Not Used	0

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	874.01315426	
1	2	818.31545055	0.00000004
2	1	818.31544035	0.00000000

Convergence criteria met.

Estimated R Matrix
for Subject 1

Row	Col1	Col2
1	24.5588	19.1109
2	19.1109	26.9232

Covariance Parameter Estimates

Cov Parm	Subject	Estimate
UN(1,1)	ID	24.5588
UN(2,1)	ID	19.1109
UN(2,2)	ID	26.9232

Fit Statistics

-2 Res Log Likelihood	818.3
AIC (smaller is better)	824.3
AICC (smaller is better)	824.5
BIC (smaller is better)	831.2

Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
2	55.70	<.0001

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
full	1	70	4.36	0.0404
cdnOrder	1	70	0.01	0.9386
cdnNorm	1	70	10.39	0.0019
cdnOrder*cdnNorm	1	70	0.42	0.5206
confed	1	70	4.99	0.0286
cdnOrder*confed	1	70	5.60	0.0208
cdnNorm*confed	1	70	14.40	0.0003
cdnOrd*cdnNor*confed	1	70	2.04	0.1578

Least Squares Means

Effect	confed	Order	Amount eaten by Confederate	Estimate
cdnOrder		AloneConfederate		9.7728
cdnOrder		ConfederateAlone		9.8581
cdnNorm			A little	8.0165
cdnNorm			A lot	11.6143
cdnOrder*cdnNorm		AloneConfederate	A little	8.3307
cdnOrder*cdnNorm		AloneConfederate	A lot	11.2148
cdnOrder*cdnNorm		ConfederateAlone	A little	7.7024
cdnOrder*cdnNorm		ConfederateAlone	A lot	12.0139
confed	Absent			10.2993
confed	Present			9.3316
cdnOrder*confed	Absent	AloneConfederate		9.7550
cdnOrder*confed	Present	AloneConfederate		9.7905
cdnOrder*confed	Absent	ConfederateAlone		10.8436
cdnOrder*confed	Present	ConfederateAlone		8.8726
cdnNorm*confed	Absent		A little	9.3153
cdnNorm*confed	Present		A little	6.7177
cdnNorm*confed	Absent		A lot	11.2833
cdnNorm*confed	Present		A lot	11.9454
cdnOrd*cdnNor*confed	Absent	AloneConfederate	A little	9.4318
cdnOrd*cdnNor*confed	Present	AloneConfederate	A little	7.2295
cdnOrd*cdnNor*confed	Absent	AloneConfederate	A lot	10.0781
cdnOrd*cdnNor*confed	Present	AloneConfederate	A lot	12.3515
cdnOrd*cdnNor*confed	Absent	ConfederateAlone	A little	9.1988
cdnOrd*cdnNor*confed	Present	ConfederateAlone	A little	6.2059
cdnOrd*cdnNor*confed	Absent	ConfederateAlone	A lot	12.4884
cdnOrd*cdnNor*confed	Present	ConfederateAlone	A lot	11.5394

Least Squares Means

Effect	confed	Order	Amount eaten by Confederate	Standard Error	DF
cdnOrder		AloneConfederate		0.7907	70
cdnOrder		ConfederateAlone		0.7694	70
cdnNorm			A little	0.7859	70
cdnNorm			A lot	0.7832	70
cdnOrder*cdnNorm		AloneConfederate	A little	1.1567	70
cdnOrder*cdnNorm		AloneConfederate	A lot	1.0915	70
cdnOrder*cdnNorm		ConfederateAlone	A little	1.0596	70
cdnOrder*cdnNorm		ConfederateAlone	A lot	1.1178	70
confed	Absent			0.6060	70
confed	Present			0.5786	70
cdnOrder*confed	Absent	AloneConfederate		0.8682	70
cdnOrder*confed	Present	AloneConfederate		0.8277	70
cdnOrder*confed	Absent	ConfederateAlone		0.8435	70
cdnOrder*confed	Present	ConfederateAlone		0.8069	70
cdnNorm*confed	Absent		A little	0.8716	70
cdnNorm*confed	Present		A little	0.8175	70
cdnNorm*confed	Absent		A lot	0.8565	70
cdnNorm*confed	Present		A lot	0.8202	70

cdnOrd*cdnNor*confed	Absent	AloneConfederate	A little	1.2802	70
cdnOrd*cdnNor*confed	Present	AloneConfederate	A little	1.2026	70
cdnOrd*cdnNor*confed	Absent	AloneConfederate	A lot	1.1958	70
cdnOrd*cdnNor*confed	Present	AloneConfederate	A lot	1.1410	70
cdnOrd*cdnNor*confed	Absent	ConfederateAlone	A little	1.1641	70
cdnOrd*cdnNor*confed	Present	ConfederateAlone	A little	1.1083	70
cdnOrd*cdnNor*confed	Absent	ConfederateAlone	A lot	1.2234	70
cdnOrd*cdnNor*confed	Present	ConfederateAlone	A lot	1.1713	70

Least Squares Means

Effect	confed	Order	Amount eaten by Confederate	t Value	Pr > |t|
cdnOrder		AloneConfederate		12.36	<.0001
cdnOrder		ConfederateAlone		12.81	<.0001
cdnNorm			A little	10.20	<.0001
cdnNorm			A lot	14.83	<.0001
cdnOrder*cdnNorm		AloneConfederate	A little	7.20	<.0001
cdnOrder*cdnNorm		AloneConfederate	A lot	10.27	<.0001
cdnOrder*cdnNorm		ConfederateAlone	A little	7.27	<.0001
cdnOrder*cdnNorm		ConfederateAlone	A lot	10.75	<.0001
confed	Absent			16.99	<.0001
confed	Present			16.13	<.0001
cdnOrder*confed	Absent	AloneConfederate		11.24	<.0001
cdnOrder*confed	Present	AloneConfederate		11.83	<.0001
cdnOrder*confed	Absent	ConfederateAlone		12.86	<.0001
cdnOrder*confed	Present	ConfederateAlone		11.00	<.0001
cdnNorm*confed	Absent		A little	10.69	<.0001
cdnNorm*confed	Present		A little	8.22	<.0001
cdnNorm*confed	Absent		A lot	13.17	<.0001
cdnNorm*confed	Present		A lot	14.56	<.0001
cdnOrd*cdnNor*confed	Absent	AloneConfederate	A little	7.37	<.0001
cdnOrd*cdnNor*confed	Present	AloneConfederate	A little	6.01	<.0001
cdnOrd*cdnNor*confed	Absent	AloneConfederate	A lot	8.43	<.0001
cdnOrd*cdnNor*confed	Present	AloneConfederate	A lot	10.83	<.0001
cdnOrd*cdnNor*confed	Absent	ConfederateAlone	A little	7.90	<.0001
cdnOrd*cdnNor*confed	Present	ConfederateAlone	A little	5.60	<.0001
cdnOrd*cdnNor*confed	Absent	ConfederateAlone	A lot	10.21	<.0001
cdnOrd*cdnNor*confed	Present	ConfederateAlone	A lot	9.85	<.0001