

# Implementation of Pairwise Fitting Technique for Analyzing Multivariate Longitudinal Data in SAS

Madan Gopal Kundu, Indiana University Purdue University at Indianapolis, Indianapolis, IN

## ABSTRACT

Mixed models are widely used in the literature for the analysis of single outcome variable measured repeatedly over time. Multivariate longitudinal data arise when a set of different responses on the same unit are measured repeatedly over time. A joint modeling of such kind of data is necessary to quantify, firstly, the relationship between evolutions of different responses and, secondly, the evolution of relationship between different response variables over time. An associated problem with joint modeling is as the number of response variable goes up, the convergence issue becomes more and more severe. To resolve this computational complexity, a pairwise fitting approach has been proposed in the literature (Fieus and Verbeke, 2006, 2007). Purpose of this article is to describe implementation of pairwise fitting approach and estimation of correlations among response variables in SAS.

**Keywords:** Joint modeling, Mixed model, Multivariate longitudinal data, Pairwise fitting approach, Correlation between evolutions, Marginal correlation.

## 1. INTRODUCTION

Mixed models are widely used in the literature for the analysis of single outcome variable, measured repeatedly over time. We can denote this standard situation as the analysis of univariate longitudinal data. Multivariate longitudinal data arise when a set of different responses on the same unit are measured repeatedly over time. In such a case a mixed model can be used for each response variable, separately. However, the above strategy is not useful to answer the following research questions:

Q1: How the evolution of one response is related to the evolution of another response

Q2: How the association between responses evolves over time

To answer the above questions a joint modeling strategy is needed.

This paper has been organized in a way first to describe the statistical detail and then its implementation in SAS. In section 2, A very brief discussion on univariate longitudinal modeling has been given. In section 3, joint model for multivariate longitudinal data is presented along with its advantage and issues in fitting. Pairwise fitting approach for joint modeling is described in section 4. In section 5, implementation of pairwise fitting approach in SAS has been described in step by step. Joint modeling through pairwise fitting and estimation of pairwise correlations using sas macro %*allpairs* has been illustrated in section 6. Throughout this paper the response variables are assumed as continuous one.

## 2. UNIVARIATE LONGITUDINAL ANALYSIS

Under the usual univariate longitudinal set-up we use the following mixed model

$$Y_i(t) = \mathbf{X}_i(t)^T \boldsymbol{\beta} + \mathbf{Z}_i(t)^T \mathbf{b}_i + \varepsilon_i(t)$$

where,

$Y_i(t)$  : Measurement of univariate response in  $i^{\text{th}}$  patient at time  $t$

$\mathbf{X}_i(t)$  : Vector of fixed covariate for  $i^{\text{th}}$  subject at time  $t$  (of dimension  $d$ )

$\mathbf{Z}_i(t)$ : Vector of random covariate for  $i^{\text{th}}$  subject at time  $t$  (of dimension  $q$ )

$\boldsymbol{\beta}$  : Vector of unknown parameters associated with fixed covariate (of dimension  $d$ )

$\mathbf{b}_i$  : Vector of unknown parameters associated with random covariate for  $i^{\text{th}}$  subject (of dimension  $q$ )  $\sim \text{MVN}(\mathbf{0}, \mathbf{D})$

$\varepsilon_i(t)$  : Random error component

Further,

- $\mathbf{Z}_i(t)$  is subset of  $\mathbf{X}_i(t)$
- $\boldsymbol{\varepsilon}_i = [\varepsilon_i(t_1), \varepsilon_i(t_2), \dots, \varepsilon_i(t_{n_i})]^T \sim \text{MVN}(\mathbf{0}, \mathbf{R}_i)$
- $\boldsymbol{\varepsilon}_i$  is independent of  $\mathbf{b}_i$

Note that it is usual practice to assume  $\mathbf{R}_i$  as  $\mathbf{I}_{n_i}$  which implies conditional independence. In other words, given  $\mathbf{b}_i$ , the observations on response variable are assumed to be independent.

### 3. EXTENSION TO MULTIVARIATE CASE

Now under Multivariate set-up more than one response variables are observed in each occasion. Hence, we measure vector of responses,  $\mathbf{Y}_i(t)$ , at each occasion and thus we can use the following model:

$$\mathbf{Y}_i(t) = \mathbf{X}_i(t)^T \boldsymbol{\beta} + \mathbf{Z}_i(t)^T \mathbf{b}_i + \boldsymbol{\varepsilon}_i(t)$$

$$\boldsymbol{\varepsilon}_i = [\boldsymbol{\varepsilon}_i(t_1), \boldsymbol{\varepsilon}_i(t_2), \dots, \boldsymbol{\varepsilon}_i(t_{n_i})]^T \sim \text{MVN}(\mathbf{0}, \mathbf{R}_i)$$

$$\mathbf{b}_i \sim \text{MVN}(\mathbf{0}, \mathbf{D})$$

$$\text{Cov}(\mathbf{b}_i, \boldsymbol{\varepsilon}_i) = \mathbf{0}$$

Note that here the assumption of conditional independence does not hold because, given  $\mathbf{b}_i$ , the observations are not totally independent. Given  $\mathbf{b}_i$ , the observations measured at same occasion on same individual might be correlated. As a result,

$$\mathbf{R}_i = \mathbf{I}_{n_i} \otimes \boldsymbol{\Sigma}_{m \times m}$$

where  $\boldsymbol{\Sigma}_{m \times m}$  is the variance covariance matrix of  $m$  response variables conditional on  $\mathbf{b}_i$ .

Let's look at the problem in its simplest form when we have only two continuous response variables  $Y_1$  and  $Y_2$  are measured over time for a number of subjects. Each of the variables is described using the linear mixed-effects model:

$$\begin{aligned} Y_{1i}(t) &= \mu_1(t) + a_{1i} + b_{1i}t + \varepsilon_{1i}(t) \\ Y_{2i}(t) &= \mu_2(t) + a_{2i} + b_{2i}t + \varepsilon_{2i}(t) \end{aligned}$$

where  $\mu_1(t)$  and  $\mu_2(t)$  refer to the population means at time  $t$ . We assume that random effects are jointly distributed as follows

$$\begin{bmatrix} a_{1i} \\ b_{1i} \\ a_{2i} \\ b_{2i} \end{bmatrix} \sim N(\mathbf{0}, \mathbf{D})$$

where,  $\mathbf{D}$ , the covariance matrix of the random effects, has the following structure:

$$\begin{bmatrix} \sigma_{a_1}^2 & \sigma_{a_1 a_2} & \sigma_{a_1 b_1} & \sigma_{a_1 b_2} \\ & \sigma_{a_2}^2 & \sigma_{a_2 b_1} & \sigma_{a_2 b_2} \\ & & \sigma_{b_1}^2 & \sigma_{b_1 b_2} \\ & & & \sigma_{b_2}^2 \end{bmatrix}$$

The error components are uncorrelated and not associated with the random effects

$$\begin{bmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ 0 & \sigma_2^2 \end{bmatrix}\right)$$

Clearly, (Q1) the correlation between the evolution of  $Y_1$  and  $Y_2$  is given by:

$$r_E = \frac{\text{Cov}(b_1, b_2)}{\sqrt{\text{Var}(b_1) \times \text{Var}(b_2)}} = \frac{\sigma_{b_1 b_2}}{\sqrt{\sigma_{b_1}^2 \times \sigma_{b_2}^2}}$$

And (Q2) the marginal correlation between  $Y_1$  and  $Y_2$  at time  $t$  is given as:

$$r_M(t) = \frac{\text{Cov}(Y_{1i}(t), Y_{2i}(t))}{\sqrt{\text{Var}(Y_{1i}(t)) \times \text{Var}(Y_{2i}(t))}} = \frac{\sigma_{a_1 a_2} + t\sigma_{a_1 b_2} + t\sigma_{a_2 b_1} + t^2\sigma_{b_1 b_2} + \sigma_{12}}{\sqrt{(\sigma_{a_1}^2 + 2t\sigma_{a_1 b_1} + t^2\sigma_{b_1}^2 + \sigma_1^2) \times (\sigma_{a_2}^2 + 2t\sigma_{a_2 b_2} + t^2\sigma_{b_2}^2 + \sigma_2^2)}}$$

It is not difficult to comprehend that as the number of response variables (or the dimension of multivariate response) increases, the number of covariance parameter increases exponentially and the problem of estimation of covariance parameters becomes more and more difficult.

If we have  $m$  response variables and 2 random effects (random slope and intercept) for each response variables, then we have  $2m$  random effects. If we assume that random effects follow  $\text{MVN}(\mathbf{0}, \mathbf{D})$  then  $\mathbf{D}$  will have  $\binom{2m}{2} + 2m$  covariance parameters and  $\mathbf{R}$  will contain covariance  $\binom{m}{2} + m$  unknown parameters. Therefore, together  $\mathbf{D}$  and  $\mathbf{R}$  will have  $\binom{2m}{2} + \binom{m}{2} + 3m$  covariance parameters. For example, when  $m=14$ , this quantity will be 511.

### 4. PAIRWISE FITTING APPROACH

To resolve the computational complexity of high dimensional joint random-effects models, the dimensionality of the problem needs to be reduced. One possible strategy might be fitting first all pairwise bivariate models separately, instead of maximizing

the likelihood of the full joint multivariate model (Fieuws and Verbeke, 2006, 2007). Assuming the full joint model is correct, all possible pairwise models are correct. This approach is equivalent to maximizing a pseudo-likelihood function (Besag, 1975) of the following form

$$pl(\boldsymbol{\theta}) = l(\mathbf{Y}_1, \mathbf{Y}_2 | \boldsymbol{\theta}_{1,2}) l(\mathbf{Y}_1, \mathbf{Y}_3 | \boldsymbol{\theta}_{1,3}) \dots l(\mathbf{Y}_{m-1}, \mathbf{Y}_m | \boldsymbol{\theta}_{m-1,m}) = \prod_{r=1}^{m-1} \prod_{s=r+1}^m l(\mathbf{Y}_r, \mathbf{Y}_s | \boldsymbol{\theta}_{r,s})$$

So the pseudo-log likelihood can be written as

$$pll(\boldsymbol{\theta}) = \sum_{r=1}^{m-1} \sum_{s=r+1}^m ll(\mathbf{Y}_r, \mathbf{Y}_s | \boldsymbol{\theta}_{r,s})$$

where,  $l(\mathbf{Y}_r, \mathbf{Y}_s | \boldsymbol{\theta}_{r,s})$ ,  $ll(\mathbf{Y}_r, \mathbf{Y}_s | \boldsymbol{\theta}_{r,s})$  and  $\boldsymbol{\theta}_{r,s}$  represent likelihood, log likelihood and the vector of all parameters in the bivariate joint mixed model corresponding to the  $r^{\text{th}}$  and  $s^{\text{th}}$  response variable, respectively.

We have  $P$  possible pairs from  $m$  response variables where  $P=m(m-1)/2$ . We can write the  $pll(\boldsymbol{\theta})$  in following way as well

$$pll(\boldsymbol{\theta}) = \sum_{p=1}^P ll(\mathbf{Y}_p | \boldsymbol{\theta}_p)$$

where,  $\mathbf{Y}_p$  contains all the observations in  $p^{\text{th}}$  pair. Similarly,  $\boldsymbol{\theta}_p$  contains all the parameters of  $p^{\text{th}}$  pair.

Let,  $\boldsymbol{\theta}^*$  be the vector containing all parameters (fixed effect parameters as well as covariance parameters) of full multivariate joint mixed model and  $\boldsymbol{\theta}$  be the stacked vector combining all pair-specific parameter vectors  $\boldsymbol{\theta}_p$  ( $p=1, \dots, P$ ). Also assume that  $\hat{\boldsymbol{\theta}}_p$  is the maximizer of  $l(\mathbf{Y}_p | \boldsymbol{\theta}_p)$ . Then  $\hat{\boldsymbol{\theta}}$ , the stacked vector combining all  $\hat{\boldsymbol{\theta}}_p$ , would maximize the  $pll(\boldsymbol{\theta})$ .

Under regularity conditions, the asymptotic distribution of  $\hat{\boldsymbol{\theta}}$  (maximizer of pseudo likelihood) is given as (Geys, Molenberghs & Ryan, 1999):

$$\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \sim MVN(\mathbf{0}, \mathbf{J}^{-1} \mathbf{K} \mathbf{J}^{-1})$$

where,

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_1 & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{J}_P \end{bmatrix} \quad \mathbf{K} = \begin{bmatrix} \mathbf{K}_{11} & \dots & \mathbf{K}_{1P} \\ \vdots & \ddots & \vdots \\ \mathbf{K}_{P1} & \dots & \mathbf{K}_{PP} \end{bmatrix}$$

where,

$$\mathbf{J}_p = -\frac{1}{N} \sum_{i=1}^N E \left( \frac{\partial^2 ll_i(\mathbf{Y}_p | \boldsymbol{\theta}_p)}{\partial \boldsymbol{\theta}_p \partial \boldsymbol{\theta}_p^T} \right)$$

$$\mathbf{K}_{pp'} = \frac{1}{N} \sum_{i=1}^N E \left( \frac{\partial ll_i(\mathbf{Y}_p | \boldsymbol{\theta}_p)}{\partial \boldsymbol{\theta}_p} \frac{\partial ll_i(\mathbf{Y}_{p'} | \boldsymbol{\theta}_{p'})}{\partial \boldsymbol{\theta}_{p'}} \right)$$

Let,

$$\hat{\mathbf{V}}_{i,p} = \mathbf{Z}_{i,p} \hat{\mathbf{D}}_p \mathbf{Z}_{i,p}^T + \hat{\mathbf{R}}_{i,p} \quad \hat{\mathbf{W}}_{i,p} = \hat{\mathbf{V}}_{i,p}^{-1}$$

Then,  $\mathbf{J}_p$  and  $\mathbf{K}_{pp'}$  are estimated as follows

$$\hat{\mathbf{J}}_p = \frac{1}{N} \sum_{i=1}^N \mathbf{X}_{i,p}^T \hat{\mathbf{W}}_{i,p} \mathbf{X}_{i,p}$$

$$\hat{\mathbf{K}}_{pp'} = \frac{1}{N} \sum_{i=1}^N \mathbf{X}_{i,p}^T \hat{\mathbf{W}}_{i,p} \mathbf{e}_{i,p} (\mathbf{X}_{i,p'}^T \hat{\mathbf{W}}_{i,p'} \mathbf{e}_{i,p'})^T$$

where,  $\mathbf{X}_{i,p}$ ,  $\mathbf{Z}_{i,p}$  and  $\mathbf{e}_{i,p}$  are contribution due to  $i^{\text{th}}$  subject in  $\mathbf{X}_p$  (design matrix pertaining to fixed factor),  $\mathbf{Z}_p$  (design matrix pertaining to random factor) and  $\mathbf{e}_p$  (residual vector).  $N$  indicates the number of subjects.

If we write

$$\mathbf{H}_p = \sum_{i=1}^N \mathbf{X}_{i,p}^T \hat{\mathbf{W}}_{i,p} \mathbf{X}_{i,p} \text{ and } \mathbf{G}_p = [\mathbf{X}_{1,p}^T \hat{\mathbf{W}}_{1,p} \mathbf{e}_{1,p} \quad \dots \quad \mathbf{X}_{N,p}^T \hat{\mathbf{W}}_{N,p} \mathbf{e}_{N,p}]$$

and also assume,

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{H}_P \end{bmatrix} \quad \mathbf{G} = \begin{bmatrix} \mathbf{G}_1 \\ \vdots \\ \mathbf{G}_P \end{bmatrix}$$

Then,

$$\hat{\mathbf{J}} = \frac{1}{N} \mathbf{H} \quad \text{and} \quad \hat{\mathbf{K}} = \frac{1}{N} \mathbf{G} \mathbf{G}^T$$

It is important to note that the parameter vectors  $\boldsymbol{\theta}^*$  and  $\boldsymbol{\theta}$  are not equivalent (see step 7 - step 10 in section 5). Indeed, some parameters in  $\boldsymbol{\theta}^*$  will have a singular counterpart in  $\boldsymbol{\theta}$ , for example, the covariance between random effects from two different

outcomes. Other elements in  $\Theta^*$  will have a multiple counterpart in  $\Theta$ , for example, the covariance between random effects from the same outcome. In later case, a single estimate is obtained by averaging all corresponding ML estimates in  $\hat{\Theta}$ .

We can find out a matrix  $\mathbf{A}$  such that  $\Theta^* = \mathbf{A}\Theta$ . Then,

$$\sqrt{N}(\hat{\Theta}^* - \Theta^*) \sim \text{MVN}(\mathbf{0}, \mathbf{A}\Sigma\mathbf{A}^T),$$

where,  $\Sigma = \mathbf{J}^{-1}\mathbf{K}\mathbf{J}^{-1}$  and  $\hat{\Theta}^* = \mathbf{A}\hat{\Theta}$

## 5. IMPLEMENTATION IN SAS

Here implementation of pairwise fitting approach (as described in section 4) in SAS has been explained assuming 3 response variables (say Y1, Y2 and Y3) and 2 predictor variables (say, X1 and X2) in addition to time variable (say, TIMEVAR). The random part of the model includes random intercept and random slope for TIMEVAR.

### STEP 1: DETERMINE THE POSSIBLE PAIRS OF RESPONSE VARIABLES AND CREATE PAIR-SPECIFIC DATASETS

Since we have 3 response variables, we can have 3 possible pairs, namely, (Y1, Y2), (Y1, Y3) and (Y2, Y3). Let consider the 1<sup>st</sup> pair (Y1, Y2). We need to create a dataset (say, data1) with the following variables:

- **Patid:** Patient or subject id
- **Y12:** all the entries of Y1 followed by all the entries of Y2. If there are n1 observations in Y1 and n2 in Y2 then Y12 will have n1+n2 observations.
- **Outcome\_num:** 1 (2) if observation in Y12 belongs to Y1 (Y2).
- **Timevar\_1:** if outcome\_num=1 then Timevar\_1=Timevar, otherwise Timevar\_1=0
- **Timevar\_2:** if outcome\_num=2 then Timevar\_2=Timevar, otherwise Timevar\_2=0
- **int\_1:** if outcome\_num=1 then int\_1=1, otherwise int\_1=0
- **int\_2:** if outcome\_num=2 then int\_2=1, otherwise int\_2=0
- **X1\_1:** if outcome\_num=1 then X1\_1=X1, otherwise X1\_1=0
- **X1\_2:** if outcome\_num=2 then X1\_2=X1, otherwise X1\_2=0
- **X2\_1:** if outcome\_num=1 then X2\_1=X2, otherwise X2\_1=0
- **X2\_2:** if outcome\_num=2 then X2\_2=X2, otherwise X2\_2=0

*Important note: For every patient, if there is an observation with Outcome\_num=1 at given value of timevar, then there need to be a row with Outcome\_num=2 (In case there is missing observation, use dot(.) to represent missingness in Y12) or vice-versa. This should be ensured for correct calculation of J and K in next steps.*

In a similar way, we need to create dataset data2 and data3 for pair (Y1, Y3) and (Y2, Y3), respectively. We can create dataset data2 (just replace Y12 by Y13 and Y2 by Y3 in the above) and dataset data3 for pair (Y2, Y3) (replace Y12 by Y23, Y1 by Y2 and Y2 by Y3 in the above) in similar way as instructed above.

### STEP 2: FIT BIVARIATE NORMAL MODEL FOR EACH PAIR

We can fit bivariate normal (BVN) model either using PROC MIXED or Proc NLMIXED. When response variable is continuous, either PROC MIXED or PROC NLMIXED can be used. However, PROC MIXED may be preferred because we need to provide initial parameter estimate in PROC NLMIXED. Following code can be used to fit BVN model for the 1<sup>st</sup> pair (Y1, Y2).

```
proc mixed data=data1 covtest noclprint;
  class patid timevar outcome_num ;
  model Y12=int_1 int_2 timevar_1 timevar_2 X1_1 X1_2 X2_1 X2_2
        /noint solution outpm=resid_1;
  random int_1 int_2 timevar_1 timevar_2/subject=patid type=un;
  repeated outcome_num/subject=patid*timevar type=un;
  ods output covparms=cov_1 solutionF=Fixed_1;
```

run;

Suppose it returns the following output (partial)

The Mixed Procedure						
Covariance Parameter Estimates						
Cov Parm	Subject	Standard Estimate	Z Error	Value	Pr Z	
UN(1,1)	patid	0.04838	0.006903	7.01	<.0001	
UN(2,1)	patid	0.2496	0.04027	6.20	<.0001	
UN(2,2)	patid	1.8470	0.3293	5.61	<.0001	
UN(3,1)	patid	-0.00227	0.001210	-1.88	0.0603	
UN(3,2)	patid	-0.01231	0.007868	-1.56	0.1178	
UN(3,3)	patid	0.000713	0.000294	2.42	0.0077	

UN(4,1)	patid	-0.01127	0.009597	-1.17	0.2404
UN(4,2)	patid	-0.02576	0.06806	-0.38	0.7051
UN(4,3)	patid	0.004793	0.002193	2.19	0.0288
UN(4,4)	patid	0.05864	0.02358	2.49	0.0064
UN(1,1)	patid*timevar	0.02911	0.001863	15.63	<.0001
UN(2,1)	patid*timevar	0.07433	0.01125	6.61	<.0001
UN(2,2)	patid*timevar	1.4938	0.1088	13.73	<.0001

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr >  t
Int_1	1.3072	0.07771	292	16.82	<.0001
Int_2	12.2925	0.5539	284	22.19	<.0001
Timevar_1	-0.01184	0.003994	234	-2.96	0.0033
Timevar_2	-0.2013	0.03337	225	-6.03	<.0001
X1_1	-0.00173	0.001598	670	-1.08	0.2787
X1_2	-0.05062	0.01143	670	-4.43	<.0001
X2_1	0.06824	0.03844	670	1.78	0.0763
X2_2	0.1563	0.2761	670	0.57	0.5716

So here,

$$\hat{\theta}_1 = \begin{bmatrix} \text{Int}_1 \\ \text{Int}_2 \\ \text{Timevar}_1 \\ \text{Timevar}_2 \\ \text{X1}_1 \\ \text{X1}_2 \\ \text{X2}_1 \\ \text{X2}_2 \end{bmatrix} = \begin{bmatrix} 1.3072 \\ 12.2925 \\ -0.0118 \\ -0.2013 \\ -0.0017 \\ -0.0506 \\ 0.0682 \\ 0.1563 \end{bmatrix} \quad \hat{D}_1 = \text{Cov} \begin{bmatrix} \text{ri of Y1} \\ \text{ri of Y2} \\ \text{rs of Y1} \\ \text{rs of Y2} \end{bmatrix} = \begin{bmatrix} 0.048 & 0.250 & -0.002 & -0.011 \\ & 1.847 & -0.012 & -0.026 \\ & & 0.001 & 0.005 \\ & & & 0.059 \end{bmatrix}, \text{cov} \begin{bmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{bmatrix} = \begin{bmatrix} 0.029 & 0.074 \\ & 1.494 \end{bmatrix}$$

Similarly, we can obtain  $(\hat{\theta}_2, \hat{D}_2$  and  $\hat{R}_2)$  after running PROC MIXED in data2 and  $(\hat{\theta}_3, \hat{D}_3$  and  $\hat{R}_3)$  from data3. Let we have,

$$\hat{\theta}_2 = \begin{bmatrix} \text{Int}_1 \\ \text{Int}_3 \\ \text{Timevar}_1 \\ \text{Timevar}_3 \\ \text{X1}_1 \\ \text{X1}_3 \\ \text{X2}_1 \\ \text{X2}_3 \end{bmatrix} = \begin{bmatrix} 1.3027 \\ 4.6399 \\ -0.0087 \\ -0.0438 \\ -0.0017 \\ -0.0063 \\ 0.0639 \\ 0.4017 \end{bmatrix} \quad \hat{D}_2 = \text{Cov} \begin{bmatrix} \text{ri of Y1} \\ \text{ri of Y3} \\ \text{rs of Y1} \\ \text{rs of Y3} \end{bmatrix} = \begin{bmatrix} 0.050 & 0.158 & -0.003 & -0.005 \\ & 0.654 & -0.013 & -0.026 \\ & & 0.001 & 0.002 \\ & & & 0.009 \end{bmatrix}, \text{cov} \begin{bmatrix} \varepsilon_{1i} \\ \varepsilon_{3i} \end{bmatrix} = \begin{bmatrix} 0.029 & 0.047 \\ & 0.447 \end{bmatrix}$$

$$\hat{\theta}_3 = \begin{bmatrix} \text{Int}_2 \\ \text{Int}_3 \\ \text{Timevar}_2 \\ \text{Timevar}_3 \\ \text{X1}_2 \\ \text{X1}_3 \\ \text{X2}_2 \\ \text{X2}_3 \end{bmatrix} = \begin{bmatrix} 12.3390 \\ 4.5710 \\ -0.2220 \\ -0.0506 \\ -0.0504 \\ -0.0046 \\ 0.1964 \\ 0.4188 \end{bmatrix} \quad \hat{D}_3 = \text{Cov} \begin{bmatrix} \text{ri of Y2} \\ \text{ri of Y3} \\ \text{rs of Y2} \\ \text{rs of Y3} \end{bmatrix} = \begin{bmatrix} 1.835 & 0.780 & -0.044 & -0.029 \\ & 0.635 & -0.045 & -0.021 \\ & & 0.069 & 0.009 \\ & & & 0.007 \end{bmatrix}, \text{cov} \begin{bmatrix} \varepsilon_{2i} \\ \varepsilon_{3i} \end{bmatrix} = \begin{bmatrix} 1.459 & 0.166 \\ & 0.449 \end{bmatrix}$$

where *ri* and *rs* stands for random intercept and random slope, respectively.

### STEP 3: CALCULATE H AND G FOR EACH PAIR

Following code can be used to calculate H and G matrix for the pair (Y1, Y2).

```
proc iml symsize=10000 worksize=10000;
  free H; free G;

  /*Between subject covariance matrix*/
  D={ 0.048 0.250 -0.002 -0.011,
      0.250 1.847 -0.012 -0.026,
      -0.002 -0.012 0.001 0.005,
```

```

-0.011 -0.026 0.005 0.059 };

/*Within subject covariance matrix*/
R={ 0.029    0.074,
    0.074    1.494};
/*Extracting Y, X, Z and residuals*/
use resid_1;
read all var{patid} into id;
read all var{Y12} into Y;
read all var{resid} into resid;
read all var{ int_1 int_2 timevar_1 timevar_2 X1_1 X1_2 X2_1 X2_2} into X;
read all var{ int_1 int_2 timevar_1 timevar_2} into Z;
close resid_1;
numobs=nrow(X);

/*Generate vsize that contains number of observations for each subject*/
count=1;
free vsize;
do i=1 to (numobs-1);
    if id[i]=id[i+1] then
        do; count=count+1; end;
    else if id[i]^=id[i+1] then
        do; vsize=vsize//count; count=1; end;
    if i=numobs-1 then vsize=vsize//count;
end;
nsubjects=nrow(vsize);
p=ncol(X);

H=J(p,p,0);
do i=1 to nsubjects;
    if i=1 then pnt=1;

    /*Z, X, Y and resid matrix for i-th subject*/
    Zi=Z[pnt:pnt+vsize[i]-1,];
    Xi=X[pnt:pnt+vsize[i]-1,];
    yi=y[pnt:pnt+vsize[i]-1];
    resid_i=resid[pnt:pnt+vsize[i]-1];

    /*Generates Ri */
    ni=nrow(yi);
    I_ni=diag(J(ni/2,1,1));
    Ri=I_ni@R;

    /*Check for missing observation in Y and X
    Vector pr_Y: contains the list of non-missing observations in Y
    Vector pr_X: contains the list of non-missing observations in X.
    (If any of the covariate value is missing - it will be considered as missing
    in general)*/
    pr_Y=loc(resid_i^=.);
    tloc=ncol(Xi);
    fnd=(Xi^=.);
    fnd2=fnd[,+];
    pr_X=loc(fnd2=tloc);
    if (ncol(pr_Y)>0 & ncol(pr_X)>0) then
        do;
            Present=xsect(pr_Y, pr_X);
            Zi=Zi[Present,];
            Xi=Xi[Present,];
            yi=yi[Present,];
            Resid_i=Resid_i[Present,];
            Ri=Ri[Present,Present];
            Vi=Zi*D*t(Zi)+Ri; /*Vi*/
            Wi=ginv(Vi); /*Wi*/
            H_i=t(Xi)*Wi*Xi; /*Contribution to H from i-th subject*/
            H=H+H_i; /*Accumulated Hessian matrix till i-th individual*/
        end;
    end;
end;

```

```

                G_ik=t(Xi)*Wi*Residi; /*Contribution to G from i-th subject*/
                G=G||G_ik;           /*Accumulated G matrix till i-th individual*/
            end;
        else G=G||J(p,1,0);
        pnt=pnt+vsize[i];
    end;
create J_1 from H; append from H;
create G_1 from G; append from G;
quit;

```

Resultant H\_1 and G\_1 dataset are the **H** and **G** matrix for pair 1. We will call them **H<sub>1</sub>** and **G<sub>1</sub>**. In our case **H<sub>1</sub>** will be a matrix of order (8X8). (Why 8? Because we have 8 fixed parameters to estimate for each pair).

9631.37	-891.58	23938.42	-1910.03	456738.19	-42309.15	1410.04	-129.72
-891.58	189.03	-2015.23	375.09	-42309.15	8966.39	-129.72	27.28
23938.42	-2015.23	145568.15	-9688.88	1134228.64	-95704.15	3541.55	-303.88
-1910.03	375.09	-9688.88	1975.93	-90637.37	17817.42	-287.86	57.21
456738.19	-42309.15	1134228.64	-90637.37	22349023.99	-2071133.04	67838.54	-6272.72
-42309.15	8966.39	-95704.15	17817.42	-2071133.04	438808.64	-6272.72	1317.91
1410.04	-129.72	3541.55	-287.86	67838.54	-6272.72	1410.04	-129.72
-129.72	27.28	-303.88	57.21	-6272.72	1317.91	-129.72	27.28

**G<sub>1</sub>** is an matrix of order (8XN), where *N* is the number of subjects. Similarly we need to find out **H<sub>2</sub>**, **G<sub>2</sub>** (from pair 2) and **H<sub>3</sub>** and **G<sub>3</sub>** (from pair 3) using the above code. Just we need to update D and R matrix in the code for each pair.

#### STEP 4: COMBINE ALL J AND K OBTAINED FOR ALL PAIRS

Once we have (**H<sub>1</sub>**, **G<sub>1</sub>**), (**H<sub>2</sub>**, **G<sub>2</sub>**) and (**H<sub>3</sub>**, **G<sub>3</sub>**), we need to combine **H<sub>1</sub>**, **H<sub>2</sub>** and **H<sub>3</sub>** into **H** and **G<sub>1</sub>**, **G<sub>2</sub>** and **G<sub>3</sub>** into **G**. Combined **H** and **K** matrix will be

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{H}_3 \end{bmatrix}_{24 \times 24} \qquad \mathbf{G} = \begin{bmatrix} \mathbf{G}_1 \\ \dots \\ \mathbf{G}_2 \\ \dots \\ \mathbf{G}_3 \end{bmatrix}_{24 \times N}$$

Following code serve this purpose.

```

proc iml;
use H_1; read all into H_1; close H_1;
use H_2; read all into H_2; close H_2;
use H_3; read all into H_3; close H_3;

H_comb=block(H_1, H_2, H_3);
create H from H_comb;
append from H_comb;

use G_1; read all into G_1; close G_1;
use G_2; read all into G_2; close G_2;
use G_3; read all into G_3; close G_3;

G_comb=G_1//G_2//G_3;
create G from G_comb; append from G_comb;
quit;

```

#### STEP 5: FIND OUT ESTIMATE J AND K

Once we have the H and G, we can find out estimate of J and K as follows

$$\hat{\mathbf{j}} = \frac{1}{N} \mathbf{H} \qquad \text{and} \qquad \hat{\mathbf{K}} = \frac{1}{N} \mathbf{G} \mathbf{G}^T$$

Following code serve this purpose.

```

proc iml;
use G; read all into G; close G;
nsubjects=ncol(G);
K_1=G*t(G);

```

```

K=K_1#1/nsubjects;
create K from K; append from K;

use H; read all into H; close H;
J=H#1/nsubjects;
create J from J; append from J;

```

**quit;**

In our case, both **J** and **K** will be of order (24X24).

**STEP 6: ESTIMATE  $\Sigma = J^{-1}KJ^{-1}$  AND  $\Sigma_0 = \frac{1}{N}\Sigma$**

```

proc iml;
  use J; read all into J; close J;
  use K; read all into K; close K;
  nsubjects=<specify the number of subjects>;
  Sigma=inv(J)*K*inv(J);
  Sigma0=Sigma#1/nsubjects;
  create Sigma0 from Sigma0; append from Sigma0;
quit;

```

Again, both  $\Sigma$  and  $\Sigma_0$  will be of order (24X24).

**STEP 7: COMBINE THE FIXED PARAMETER ESTIMATES FOR EACH PAIR TO OBTAIN  $\hat{\Theta}$**

We obtained the fixed parameter estimates for each pair in Step 2. Now we need to combine them to obtain  $\hat{\Theta}$ .

```

proc iml;
  Theta_est1={1.3072, 12.2925, -0.01184, -0.2013, -0.00173, -0.05062, 0.06824, 0.1563};
  Theta_est2={1.3027, 4.6399, -0.00874, -0.04383, -0.00174, -0.00628, 0.0638, 0.4017};
  Theta_est3={12.3390, 4.5710, -0.2220, -0.05057, -0.05038, -0.00457, 0.1964, 0.4188};
  Theta_est=Theta_est1//Theta_est2//Theta_est3;
  create Theta_est from Theta_est;
  append from Theta_est;
quit;

```

For the ongoing example, the combined fixed parameter estimates should look like as follows and it will be stored in the dataset Theta\_est.

$$\hat{\Theta} = \begin{bmatrix} \hat{\Theta}_1 \\ \dots \\ \hat{\Theta}_2 \\ \dots \\ \hat{\Theta}_3 \end{bmatrix} = \begin{array}{l} \text{Int}_1 \\ \text{Int}_2 \\ \text{Timevar}_1 \\ \text{Timevar}_2 \\ \text{X1}_1 \\ \text{X1}_2 \\ \text{X2}_1 \\ \text{X2}_2 \\ \dots \dots \dots \\ \text{Int}_1 \\ \text{Int}_3 \\ \text{Timevar}_1 \\ \text{Timevar}_3 \\ \text{X1}_1 \\ \text{X1}_3 \\ \text{X2}_1 \\ \text{X2}_3 \\ \dots \dots \dots \\ \text{Int}_2 \\ \text{Int}_3 \\ \text{Timevar}_2 \\ \text{Timevar}_3 \\ \text{X1}_2 \\ \text{X1}_3 \\ \text{X2}_2 \\ \text{X2}_3 \end{array} = \begin{array}{l} 1.3072 \\ 12.2925 \\ -0.0118 \\ -0.2013 \\ -0.0017 \\ -0.0506 \\ 0.0682 \\ 0.1563 \\ \dots \dots \dots \\ 1.3027 \\ 4.6399 \\ -0.0087 \\ -0.0438 \\ -0.0017 \\ -0.0063 \\ 0.0638 \\ 0.4017 \\ \dots \dots \dots \\ 12.3390 \\ 4.5710 \\ -0.2220 \\ -0.0506 \\ -0.0504 \\ -0.0046 \\ 0.1964 \\ 0.4188 \end{array}$$

**Step 8: Determine the appropriate A matrix**

$$\hat{\theta}^* = \begin{bmatrix} Int\_1 \\ Int\_2 \\ Int\_3 \\ Timevar\_1 \\ Timevar\_2 \\ Timevar\_3 \\ X1\_1 \\ X1\_2 \\ X1\_3 \\ X2\_1 \\ X2\_2 \\ X2\_3 \end{bmatrix} = A\hat{\theta}$$

Here **A** is

$$\begin{bmatrix} 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 \end{bmatrix}$$

### STEP 9: CALCULATE, $\hat{\theta}^* = A\hat{\theta}$ AND $\Sigma\Lambda^T$

We can accomplish step 9 executing the following code.

```
proc iml;
A={0.5 0 0 0 0 0 0 0.5 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0,
  0 0.5 0 0 0 0 0 0 0 0 0 0 0 0.5 0 0 0 0 0 0 0 0 0,
  0 0 0 0 0 0 0 0.5 0 0 0 0 0 0 0 0.5 0 0 0 0 0 0 0,
  0 0 0.5 0 0 0 0 0 0 0.5 0 0 0 0 0 0 0 0 0 0 0 0 0,
  0 0 0 0.5 0 0 0 0 0 0 0 0 0 0 0.5 0 0 0 0 0 0 0 0,
  0 0 0 0 0 0 0 0 0.5 0 0 0 0 0 0 0.5 0 0 0 0 0 0 0,
  0 0 0 0 0.5 0 0 0 0 0 0.5 0 0 0 0 0 0 0 0 0 0 0 0,
  0 0 0 0 0 0 0.5 0 0 0 0 0.5 0 0 0 0 0 0 0 0 0 0 0,
  0 0 0 0 0 0 0 0.5 0 0 0 0 0 0.5 0 0 0 0 0 0 0 0 0,
  0 0 0 0 0 0 0 0 0.5 0 0 0 0 0 0 0 0.5 0 0 0 0 0 0,
  0 0 0 0 0 0 0 0 0 0.5 0 0 0 0 0 0 0 0 0.5 0 0 0,
  0 0 0 0 0 0 0 0 0 0 0.5 0 0 0 0 0 0 0 0 0.5};
use Sigma0; read all into Sigma0; close Sigma0;
use Theta_est; read all into Theta_est; close Theta_est;
Theta_est_star=A*Theta_est;
temp=diag(A*Sigma0*t(A));
stderr=J(12,1,0);
do i=1 to 12;
  stderr[i,1]=sqrt(temp[i,i]);
end;
create Theta_est_star from Theta_est_star; append from Theta_est_star;
create Stderr from Stderr; append from Stderr;
quit;
```

The dataset Theta\_est\_star and Stderr will contain the fixed parameter estimates with their standard errors.

### STEP 10: OBTAIN $\hat{D}$ AND $\hat{R}$

Here we need to combine  $\hat{D}_1$ ,  $\hat{D}_2$  and  $\hat{D}_3$  into  $\hat{D}$  and obtain  $\hat{R}$  from  $\hat{R}_1$ ,  $\hat{R}_2$  and  $\hat{R}_3$

$$\hat{D} = \text{Cov} \begin{bmatrix} \text{ri ofY1} \\ \text{rs ofY1} \\ \text{ri ofY2} \\ \text{rs ofY2} \\ \text{ri ofY3} \\ \text{rs ofY3} \end{bmatrix} = \begin{bmatrix} 0.049 & -0.0025 & 0.250 & -0.011 & 0.158 & -0.005 \\ & 0.001 & -0.012 & 0.005 & -0.013 & 0.002 \\ & & 1.841 & -0.035 & 0.780 & -0.029 \\ & & & 0.064 & -0.045 & 0.009 \\ & & & & 0.644 & -0.024 \\ & & & & & 0.008 \end{bmatrix}$$

$$\hat{R} = \text{Cov} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix} = \begin{bmatrix} 0.029 & 0.074 & 0.047 \\ & 1.4765 & 0.166 \\ & & 0.448 \end{bmatrix}$$

where  $ri$  and  $rs$  stands for random intercept and random slope, respectively. Note that two estimates of variance of random intercept of Y1 are available – 0.048 from  $\hat{D}_1$  and 0.50 from  $\hat{D}_2$ . The average of this two estimates have been included in  $\hat{D}$  as a combined estimates of variance of random intercept of Y1. Other entries in  $\hat{D}$  and  $\hat{R}$  have been obtained in similar way.

#### Step 11: Calculation of Correlation among evolutions and marginal correlation

Once we have estimated  $D$  and  $R$  matrix, we are ready to calculate pairwise correlation between evolutions and Marginal correlation at any time point. Note that marginal correlation varies with time. We can estimate pairwise correlation between evolutions and marginal correlation at time point 2 using the following code. We can calculate marginal correlation at any time point making appropriate change in  $T$  matrix in the following code.

```
proc iml;
  D = { 0.049 -0.0025 0.158 -0.005 0.250 -0.011,
        -0.0025 0.001 -0.013 0.002 -0.012 0.005,
        0.158 -0.013 0.644 -0.024 0.780 -0.045,
        -0.005 0.002 -0.024 0.008 -0.029 0.009,
        0.250 -0.012 0.780 -0.029 1.841 -0.035,
        -0.011 0.005 -0.045 0.009 -0.035 0.064};

  D_slope={0.001 0.002 0.005,
           0.002 0.008 0.009,
           0.005 0.009 0.064};

  R={0.029 0.047 0.074,
     0.047 0.448 0.166,
     0.074 0.166 1.476};

  T = {1 5 0 0 0 0,
       0 0 1 5 0 0,
       0 0 0 0 1 5}; /*Here 2 is used to calculate marginal correlation at time 2*/

  D_marg=T*D*t(D) + R;

  /*Association between slopes*/
  corr_bet_evol=j(nrow(D_slope), nrow(D_slope), 0);
  do i=1 to nrow(D_slope);
    do j=1 to ncol(D_slope);
      corr_bet_evol[i,j]=D_slope[i,j]/sqrt(D_slope[i,i]*D_slope[j,j]);
    end;
  end;
  print corr_bet_evol;

  /*Marginal correlation at time 2*/
  Marg_corr=j(nrow(D_marg), nrow(D_marg), 0);
  do i=1 to nrow(D_marg);
    do j=1 to ncol(D_marg);
      Marg_corr[i,j]=D_marg[i,j]/sqrt(D_marg[i,i]*D_marg[j,j]);
    end;
  end;
  print Marg_corr;
run;
quit;
```

## 7. GENERALIZED MACRO

Macro `%jointpair`, written by Steffen Fieuws, implementing pairwise approach to multivariate longitudinal data modeling can be obtained from <http://med.kuleuven.be/biostat/software/software.htm#MLMMpw>. We present here a macro `%allpairs` which extends the functionality of the `%jointpair` macro. The major contributions of this paper include the ability to fit the models with some of the measure outcomes missing, comparison of the model AIC values and pairwise correlations (as explained in step 11 of previous section) of the outcomes. The first contribution enables the model fitting in the common longitudinal data situation where outcome missingness is a norm rather than the exception. AIC comparison, implemented via the `randomno1` key parameter, lets us select the best error covariance structure for each pair in an automatic way. The macro is not included in the paper because of space constraints, but can be obtained by contacting with the authors (see the contact information below) or from the webpage <http://mgkundu.webs.com/>. The use of `%allpairs` has been illustrated below.

Example 1: `%allpairs(data=MyData,`

```

respvars=%str(Y1 Y2 Y3),
fixed=%str(Timevar X1 X2),
randomno2=%str(1 0 1),
timevar=Timevar,
time4mc=2);

```

Example 2: `%allpairs`(data=MyData,  
 respvars=%str(Y1 Y2 Y3),  
 fixed=%str(Timevar X1 X2),  
 randomno1=1,  
 timevar=Timevar,  
 time4mc=2);

Here one need to specify following parameters in the above macro:

data	Specify the name of the input dataset
respvars	Specify all the response variables, separated by space
fixed	Specify all the fixed variables, separated by space
randomno1	Specify 1 if the decision on random slope or intercept need to be made for each pair based on AIC and convergence criteria.
randomno2	specify string of 1 and 0, separated by space. The number of elements should be equal with number of response variables. Specify '1' if random slope of corresponding response variable is of interest. Specify zero to consider random intercept only.
timevar	Specify the time variable.
time4mc	Time for marginal correlation

Note that we need to specify either `randomno1` or `randomno2`. If both of them are specified, then `random2` will be ignored. To specify `random2` we need to know beforehand, for each response variable, whether random slope or random intercept works better than the other. For example, in Example 1, following covariance models will be used:

(Y1, Y2)	random intercept of Y1, random slope(week) of Y1 and random intercept of Y2
(Y1, Y3)	random intercept of Y1, random slope(week) of Y1, random intercept of Y3 and random slope(week) of Y3
(Y2, Y3)	random intercept of Y2, random intercept of Y3 and random slope(week) of Y3

However, there might be case where we don't have this information and we want to make the decision on random slope or intercept based on AIC. In that case we need to specify `randomno1=1` as in Example 2. In this case, for each pair, following 4 models with different covariance structures will be considered.

	Covariance structure
Model1	both Intercept and slope of both the variables in the pair.
Model2	both Intercept and slope of 1st variable, but only intercept for 2nd variable.
Model3	both Intercept and slope of 2nd variable, but only intercept for 1st variable.
Model4	only Intercept of both the variables in the pair.

Best model will be chosen according to AIC value, satisfaction of the convergence criterion and positive definiteness of the estimated covariance matrices. Note that if for any particular pair we failed to obtain a model that converges and produce positive definite covariance matrix then that pair will be ignored.

The macro produces the outputs in the form of Marginal correlation at the requested time point, Correlation among evolutions, **D** matrix for random effects ( $\mathbf{b} \sim N(\mathbf{0}, \mathbf{D})$ ), Error variances ( $\mathbf{epsilon} \sim N(\mathbf{0}, \mathbf{R})$ ) and Fixed parameter estimates with standard error. Further details on the macro come with the `%allpairs` macro itself.

## 8. CONCLUSION

Joint modeling of longitudinal multivariate responses is necessary to explore the association between response variables. An usual problem with the joint modeling is failing to convergence because of large number of association parameter to estimate. Pairwise fitting approach (using pseudolikelihood) is helpful in overcome this shortcoming. Both `%jointpair` and `%allpairs` implements joint modeling through pairwise fitting approach. The macro `%allpairs` has its advantages like it takes care of missing observations and also can choose correct covariance model for each pair based on AIC value, satisfaction of convergence criterion and positive definiteness of the estimated covariance matrices. On execution, `%allpairs` also returns the estimates of correlation between evolutions and marginal correlation for each pair of response variables.

## 9. REFERENCE

- Besag J. E. (1975). Statistical analysis of non-lattice data. *The Statistician*, **24**: 179 – 195.
- Fieuws S. and Verbeke G. (2004). Joint modeling of multivariate longitudinal profiles: pitfalls of the random effects approach. *Statistics in Medicine*, **23**: 3093 – 3104.
- Fieuws S. and Verbeke G. (2006). Pairwise fitting of mixed models for the joint modeling of multivariate longitudinal profiles. *Biometrics*, **62**: 424 – 431.
- Fieuws S., Verbeke G. and Molenberghs G. (2007). Random-effects models for multivariate repeated measures. *Statistical Methods in Medical Research*, **16**: 387 – 397.
- Geys H., Molenberghs G. and Ryan L. M. (1999). Pseudolikelihood modeling of multivariate outcomes in developmental toxicology. *Journal of the American Statistical Association*, **94(447)**: 734 – 745.
- Laird N. M. and Ware J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, **38**: 963–974.
- SAS institute Inc. (1999), SAS/Stat User's Guide, Version 8, Cary, NC: *SAS Institute Inc.*

## CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Madan Gopal Kundu  
Division of Biostatistics (3<sup>rd</sup> floor)  
410 W 10<sup>th</sup> St, Indianapolis, IN 46202  
E-mail: [mgkundu@iupui.edu](mailto:mgkundu@iupui.edu); [madan4331@yahoo.co.in](mailto:madan4331@yahoo.co.in)  
Web: <http://mgkundu.webs.com/>

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.