FORTRAN PROGRAMS FOR SAD MODELS

Programs are written in fortran 90. They have been successfully compiled using f95 compiler on unix system. No options are needed for compilation.

The programs are free for research but should be cited in publications (David et al. JAS 2015, GSE 2018 or Zenodo DOI). Use them at your own risk.

The sadmultiuni.f90 combines the saduni.f90 and SADmulti.f90 programs (zenodo https://zenodo.org/record/192036) which is useful for single trait analysis, multiple trait analysis or single trait analysis with correlated random effects (direct and indirect genetic effects for instance).

- **The sadmultiuni.f90 is for versions 1 to 3 of ASReml**
- **The sadmultiuni4.f90 is for ASReml4**

For analyzing longitudinal traits using a SAD model you need:

- 1 datafile
- 1 pedigree file if there are genetic effects in your model
- 1 ".as" file (description of the model)
- As many "paraX" files as the number of independent random effects in the model (having a SAD modeling).

1. EXPLANATION FOR A SINGLE TRAIT MODEL

Details in : David, I., Ruesche, J., Drouilhet, L., Garreau, H., & Gilbert, H. (2015). Genetic modeling of feed intake. Journal of animal science, 93(3), 965-977

For instance, the model is $\boldsymbol{y}(t_j) = \boldsymbol{\mu}(t_j) + \boldsymbol{u}(t_j) + \boldsymbol{p}(t_j)$

With, $\boldsymbol{\mu}(t_j)$ the fixed effects at time t_j, $\boldsymbol{u}(t_j)$ the vector of genetic effects at time t_j, $\boldsymbol{p}(t_j)$ the vector of pseudo-permanent environmental effects at time t_j. For a given random effect (*p* for instance) with antedependence of order α, degree β_i for the i^{th} antedependence parameter (θ_i) and innovation variance of degree γ :

$$\boldsymbol{p}(t_j) = \sum_{s=1}^{\alpha} \theta_{sj}\, \boldsymbol{p}(t_{j-s}) + \boldsymbol{e}(t_j),$$

$$\theta_{sj} = \sum_{k=0}^{\beta_s} a_{sk} t_j^k,$$

$$\boldsymbol{e}(t_j) \sim N(0, I\sigma^2_{e,j}), \sigma^2_{e,j} = \exp\left(\sum_{k=0}^{\gamma} b_k t_j^k\right)$$

We define a single trait SAD model by "SAD $\alpha\beta_1\beta_2\ldots\beta_\alpha\gamma$" for each random term

A SAD111 in then:

$$\boldsymbol{p}(t_j) = \theta_{1j}\, \boldsymbol{p}(t_{j-1}) + \boldsymbol{e}(t_j),$$

$$\theta_{1j} = a_{10} + a_{11} t_j,$$

$$\boldsymbol{e}(t_j) \sim N(0, I\sigma^2_{e,j}), \sigma^2_{e,j} = \exp(b_0 + b_1 t_j)$$

1.1. Datafile

In the datafile, in addition to phenotype, fixed effects, animal ID columns…, there must be a column for the variable that indicates the time of measurement (note: for simplicity, we consider the analysis of longitudinal data in this tutorial, so we use the term "time" to define the variable that identifies the different repeated measurements, but one can use SAD model to analyze repeated measurements across different ordered environments for instance). If some records are missing, there is no need to perform data augmentation (i.e. to have the same number of observations per animal).

It is mandatory to sort the data by time-animal. If you do not, check carefully that the time vector constructed by ASreml (same order as in the dataset) is in the same order as the one used for the SAD model (the latter appears on the screen when ASreml is running= figures after "time vector initial scale")

The time vector used in the SAD model is transformed on a [-1,1] scale (values appear on the screen after "time vector [-1,1] scale")

Datafile example: successive measurements on 4 days: 1, 3, 8, and 11

time	animal	factor	phenotype
1	1	1	43
1	2	2	21
1	3	1	25
1	4	1	14
3	2	2	52
3	3	2	62
3	4	3	42
8	1	3	51
8	2	1	42
8	3	2	12
8	4	1	45
11	1	3	52
11	2	2	14

1.2. Pedigree file (if needed)

As recommended by ASReml.

1.3. .as file

Because a pseudo-permanent environmental effect (*p*) is included in the model, there is no classical residual in this model. **However, even if theoretically the residual variance should be fixed to 0, our experience is that a too low value of the residual variance/other variances in the model induces convergence to abnormal parameter values. Thus, we recommend fixing the residual variance to a value around 100 times lower than the other variances of the model.**

If you do not include a pseudo-permanent environmental effect that follows a SAD in your model or if the SAD model for the pseudo-permanent effect is simple, then **you can consider a classical residual and you do not have to fix its variance**. However, you have to keep the !S2==1 option when specifying the model. *Ex1 in the examples part of the tutorial provides an example of SAD model with and without a true residual.*

Example of ".as" file for a SAD111 for pseudo-permanent environmental and genetic effects

```
unisad
 time !A #needed
 animal !P
 factor !A
 phenotype
Pedigree
Datafile !OWN sadmultiuni_exe #link to the fortran executable
Phenotype ~ mu time factor !r time.animal time.ide(animal)
1 1 2
0 0 IDV 0.001 !GF !S2==1  #to fix residual variance to a very low value : 0.001 ⚠ NOT TOO LOW
time.animal 2
4 time OWN4 0.1 0.1 0.9 0.9 !TCCCC #SAD111 for u: 4 parameters a10,a11,b0,b1 , parameter file: para
animal
time.ide(animal) 2
4 time OWN4 0.1 0.1 0.9 0.9 !TCCCC !F2 #SAD111 for p: 4 parameters; + link to para2
ide(animal)
```

For each random effect following a SAD model, the layout of the line indicating the SAD is as follows:
Number of time points -name of the variable indicating the time points-OWNx a b .. !Tyy !Fz with:

- x = the number of parameters to estimate (4 for a SAD111)
- a, b, ...= initial values for the parameters in this order:
  initial values for antedependence parameters (a_{sk} of θ_s starting with parameters of θ_1 then θ_2 and so on) , then initial values for the innovation variance parameters (b_k of σ_e^2)
- y: parameter types: **In the SAD model, we advise to write !TCCCCC**….(as many C as parameters)
- z is a number between 1 and 9 that forms the link to the parameter file , if omitted, the parameter file is "para", if z = 2, the parameter file is "para2", if z = 3, the parameter file is "para3" and so on. Each random effect must have its own parameter file i.e. even if model is the same for 2 random terms, one has parameter file "para" and the second parameter file "para2".

1.4. Para file:

The parameter files provide information about the specification of the antedependence and innovation variance (i.e. order, degree…). Their names must be "para", "para2" or …"para9"

Example of para file for a SAD111 model (comments in blue, **in bold: unmodifiable words**)

MODEL
SAD
ANTEDEP_ORDER #order of the antedependence α
1
MODEL_ANTEDEP #degree of the polynomial functions for the antedependence parameters $\beta_1\ \beta_2\ \ldots\beta_\alpha$
1 *# as many figures as the antedep order* α
MODEL_INNOVATION_VARIANCE #degree of the polynomial function for the innovation variance γ
1
DATA_FILE #name of the datafile
Datafile
COLUMN_FACTOR # column number in datafile of the variable that indicates the time
1

1.5. Results (most important in .asr and .gdg files)

- The **.asr** file contains the output of the model for the parameter estimates.

Example of .asr file for SAD111 for genetic and permanent environmental effects

…….
- - - Results from analysis of ADG - - -

Source	Model	terms	Gamma	Component	Comp/SE	% C	
Residual	identity	12384	0.100000E-03	0.100000E-03	0.00	0 F	*#fixed residual*
time.animal	OWNxx	4	-0.126282	-0.126282	-0.45	0 U	*# $\hat{a}_{10}$ for genetic effect u*
time.animal	OWNxx	4	0.193263	0.193263	1.73	0 U	*# $\hat{a}_{11}$ for genetic effect u*
time.animal	OWNxx	4	2.88308	2.88308	11.08	0 U	*# $\hat{b}_0$ for genetic effect u*
time.animal	OWNxx	4	-0.632185	-0.632185	-3.59	0 U	*# $\hat{b}_1$ for genetic effect u*
time.ide(ani	OWNxx	4	0.238707	0.238707	0.54	0 U	*# $\hat{a}_{10}$ for effect p*
time.ide(ani	OWNxx	4	-0.239293	-0.239293	-1.71	0 U	*# $\hat{a}_{11}$ for effect p*
time.ide(ani	OWNxx	4	1.14631	1.14631	3.36	0 U	*# $\hat{b}_0$ for effect p*
time.ide(ani	OWNxx	4	0.240279	0.240279	1.88	0 U	*# $\hat{b}_1$ for effect p*

……

- The **.gdg** file contains the inverse of the variance-covariance matrix (named B) for the random effect sorted by time (i.e the vector $\left[p(t_0)\quad p(t_1)\quad \mathrm{L}\quad p(t_n)\right]$) and its derivatives with respect to each parameter .

If "para" is used, the matrices are in asfilenameB.gdg
If "para2" is used, the matrices are in asfilenameC.gdg
If "para3" is used, the matrices are in asfilenameD.gdg…

Example of .gdg file for a SAD111

```
  4   4   1  -1 #number of parameters, number of time points, option, option
0.2750838D+00-.1186720D+000.2958084D+010.2086888D+00 #parameter estimates
  1   1 0.4218808934E-01 #value in position [1 :1] of B-1
  2   1 -0.1290789340E-02 #value in [2 :1] of B-1
  2   2 0.3438416123E-01
  3   2 0.2246697433E-02
  3   3 0.2865794115E-01
  4   3 0.4497389309E-02
  4   4 0.2253153920E-01 #value in [4 :4]  of B-1
  1   1 0.2581578679E-02 #value in [1 :1] of the derivative of B-1relative to the first parameter  a10
  2   1 -0.3420232981E-01
  2   2 -0.4493394867E-02
  3   2 -0.2776024304E-01
  3   3 -0.8994778618E-02
  4   3 -0.2253153920E-01
  4   4  0.000000000
  1   1 0.5163157359E-02 #value in [1 :1] of the derivative of B-1relative to the second parameter a11
  2   1 -0.6840465963E-01
  2   2 -0.1348018367E-01
  3   2 -0.8328072727E-01
  3   3 -0.3597911447E-01
  4   3 -0.9012615681E-01
  4   4  0.000000000
  1   1 -0.4218808562E-01 #value in [1 :1] of the derivative of B-1relative to the third parameter  b0
  2   1 0.1290789223E-02
  2   2 -0.3438415751E-01
  3   2 -0.2246697433E-02
  3   3 -0.2865794115E-01
  4   3 -0.4497389309E-02
  4   4 -0.2253153920E-01
  1   1 -0.4223679751E-01 #value in cell 1 :1 of the derivative of B-1relative  to the fourth parameter b1
  2   1 0.2581578447E-02
  2   2 -0.6895013899E-01
  3   2 -0.6740091834E-02
  3   3 -0.8687151968E-01
  4   3 -0.1798955724E-01
  4   4 -0.9012615681E-01
```

To obtain the co(variance) matrix B for each random effects (time*time covariance matrix), you have 3 options:

- you use parameter estimates in the .asr file to compute the L and D^{-1} matrices (see articles for details) and $B=\left(L'D^{-1}L\right)^{-1}$. If you choose this option, don't forget that the parameter estimates are for the time vector on the [-1,1] scale.
- you use the .gdg file to obtain values of the B^{-1} matrix and you calculate its inverse.
- you use "matrixSAD.f90" program that reads the .gdg file and computes the inverse of B^{-1} for you.

2. EXPLANATION FOR A MULTIPLE TRAIT MODEL

Details in :David, I., Garreau, H., Balmisse, E., Billon, Y., & Canario, L. (2017). Multiple-trait structured antedependence model to study the relationship between litter size and birth weight in pigs and rabbits. Genetics Selection Evolution, 49(1), 11.

Let's consider several longitudinal traits *i*. The model for each trait *i* is of the form:

$$\boldsymbol{y}_i\left(t_j\right)=\boldsymbol{\mu}_i\left(t_j\right)+\boldsymbol{u}_i\left(t_j\right)+\boldsymbol{p}_i\left(t_j\right)$$

For two traits $\boldsymbol{y}_1,\boldsymbol{y}_2$, the general form of the multiple-trait SAD model of order α,α' (for the antedependence) and η,η' (for the cross-antedependence) for a given random effect $\boldsymbol{p}$ can be written as (for $j>\max(\alpha,\alpha',\eta,\eta')$):

$$\boldsymbol{p}_1\left(t_j\right)=\sum_{s=1}^{\alpha}\theta_{sj}\,\boldsymbol{p}_1\left(t_{j-s}\right)+\sum_{s=c}^{\eta}\delta_{sj}\,\boldsymbol{p}_2\left(t_{j-s}\right)+\boldsymbol{e}\left(t_j\right)$$

$$\boldsymbol{p}_2\left(t_j\right)=\sum_{s=1}^{\alpha'}\theta'_{sj}\,\boldsymbol{p}_2\left(t_{j-s}\right)+\sum_{s=c'}^{\eta'}\delta'_{sj}\,\boldsymbol{p}_1\left(t_{j-s}\right)+\boldsymbol{\varepsilon}\left(t_j\right)$$

With

$$\theta_{sj}=\sum_{k=0}^{\beta_s}a_{sk}t_j^k,\ \theta'_{sj}=\sum_{k=0}^{\beta'_s}a'_{sk}\,t_j^k$$

$$\delta_{sj}=\sum_{k=0}^{\omega_s}\kappa_{sk}t_j^k,\ \delta'_{sj}=\sum_{k=0}^{\omega'_s}\kappa'_{sk}\,t_j^k,$$

$$\sigma^2_{e,j}=\exp\left(\sum_{k=0}^{\gamma}b_k t_j^k\right),\ \sigma^2_{\varepsilon,j}=\exp\left(\sum_{k=0}^{\gamma'}b'_k\,t_j^k\right)$$

Where, as for the single trait model, θ_{sj},θ'_{sj} are the s[th] antedependence parameters for time j for traits 1 and 2, respectively. δ_{sj},δ'_{sj} are the $(s-c+1)^{\text{th}}$ (or $(s-c'+1)^{\text{th}}$) cross-antedependence parameters for time j for traits 1 and 2, respectively. It should be noted that, conversely to the antedependence relationship that starts at time t_{j-1}, the cross antedependence relationships show greater flexibility and start at time t_{j-c} $(t_{j-c'})$ with c (c') greater or equal to 0. The $\boldsymbol{e}\left(t_j\right),\boldsymbol{\varepsilon}\left(t_j\right)$ parameters are normally distributed random effects with mean 0 and innovation variance $\sigma^2_{e,j},\sigma^2_{\varepsilon,j}$, respectively. The $\boldsymbol{e},\boldsymbol{\varepsilon}$ parameters are assumed to be independent, except if $c>0$ and $c'>0$ when a correlation between the two can be considered for the first time point.

The multiple-trait SAD model is then defined for two traits by the order of the antedependence for each trait (α, α'), the starting points (c, c') and the order of the cross-antedependence ($\eta - c + 1, \eta' - c' + 1$), the degree of the polynomial for each (cross-)antedependence parameter (β_1 to β_α, β'_1 to $\beta'_{\alpha'}$ for the antedependence ϖ_1 to $\varpi_{\eta-c+1}$, ϖ'_1 to $\varpi'_{\eta'-c'+1}$ for the cross antedependence) and the degree of the polynomial for the innovation variance of each trait (γ, γ'), as well as an indicator of the presence of an initial correlation between $\boldsymbol{e}(t_0), \boldsymbol{\varepsilon}(t_0)$ or not.

For instance, for 2 traits, a SAD111 for the antedependence for p_1, a SAD111 for the antedependence for p_2, one way cross antedependence (in this example, we assume that value of p_1 has an effect on p_2 but not the reverse) of order 1, degree 1 , and a starting point c'=0 (total: 10 parameters) the model is:

$$\boldsymbol{p}_1(t_j) = \theta_{1j}\boldsymbol{p}_1(t_{j-1}) + \boldsymbol{e}(t_j)$$
$$\boldsymbol{p}_2(t_j) = \theta'_{1j}\,\boldsymbol{p}_2(t_{j-1}) + \delta'_{0j}\,\boldsymbol{p}_1(t_j) + \boldsymbol{\varepsilon}(t_j)$$

eq1

$$\theta_{1j} = a_{10} + a_{11}t_j, \quad \theta'_{1j} = a'_{10} + a'_{11}t_j$$
$$\delta'_{0j} = \kappa'_{00} + \kappa'_{01}t_j$$
$$\sigma^2_{e,j} = \exp(b_0 + b_1 t_j), \quad \sigma^2_{\varepsilon,j} = \exp(b'_0 + b'_1 t_j)$$

2.1. Datafile

The same time vector must be used for the different traits. It means that the different traits must be measured at the same time. However, if some records are missing, there is no need to perform data augmentation (i.e. to have the same number of observations per trait and per animal) as long as there are records for each time point within each trait.

The datafile must contain two columns to describe the time (time, time2). Time corresponds to the "real" time. Time2 is a vector used to organize the time for each trait. For n time points and k traits it should be numbered from 1 to n*k (this vector is needed to define the size of the B matrix of each random effects and to sort the random effects by trait and time). **Data must be sorted by time2-animal**.

The time vector used in the SAD model is transformed on a [-1,1] scale (values appear on the screen after "time vector [-1,1] scale")

Datafile example: successive measurements on 4 days: 1, 3, 8, and 11, and 2 traits

time	Time2	animal	factor	character	phenotype
1	1	1	1	1	43

1	1	2	2	1	21
1	1	3	1	1	25
1	1	4	1	1	14
3	2	2	2	1	52
3	2	3	2	1	62
3	2	4	3	1	42
8	3	1	3	1	51
8	3	2	1	1	42
8	3	3	2	1	12
8	3	4	1	1	45
11	4	1	3	1	52
11	4	2	2	1	14
1	5	1	1	2	1
1	5	2	3	2	5
1	5	3	1	2	6
1	5	4	2	2	2
3	6	1	3	2	3
3	6	3	2	2	4
3	6	4	2	2	5
8	7	1	2	2	2
8	7	3	1	2	1
11	8	1	1	2	2
11	8	2	3	2	5

2.2. Pedigree

As recommended by ASReml.

2.3. File.as

As for the file.as for single trait analysis, because a pseudo-permanent environmental effect is included (*p*) in the model, there is no classical residual in this model. **However, even if theoretically the residual variance should be fixed to 0, our experience is that a too low value of the residual variance/other variances induces convergence to abnormal parameter values. Thus we recommend fixing the residual variance to a value that is around 100 times lower than the other variances of the model.**

If you do not include a pseudo-permanent environmental effect that follows a SAD in your model or if the SAD model for the pseudo-permanent effect is simple, then **you can consider a classical residual and you do not have to fix its variance**. However, you have to keep the !S2==1 option when specifying the model. *Ex1 in the examples part of the tutorial provides an example of SAD model with and without a true residual.*

Example of .as for a 2 traits analysis, SAD111 for $\mathbf{u}_1$ and $\mathbf{u}_2$, $\mathbf{u}_1$ and $\mathbf{u}_2$ are independent (total: 8 parameters). For **p**, model described in eq1

```
multisad
 time !A #needed
 time2 !A #needed
 animal !P
 factor !A
 character !A
 phenotype
Pedigree
Datafile !OWN sadmultiuni_exe #link to the fortran executable
Phenotype ~ character.time character. factor !r time2.animal time2.ide(animal)
1 1 2
0 0 IDV 0.001 !GF !S2==1  #to fix residual variance to a very low value : 0.001 ⚠ NOT TOO LOW
time2.animal 2
8 time2 OWN8 0.1 0.1 0.1 0.1 0.9 0.9 0.9 0.9 !TCCCCCCCC #SAD111 for u1 and u2 : 8 parameters
a10,a11,a'10,a'11,b0,b1,b'0,b'1 , parameter file: para
Animal
time2.ide(animal) 2
8 time2 OWN10 0.6 0.1 0.8 0.4 0.5 0.1 1 2 -0.1 0.1 !TCCCCCCCCCC !F2 #SAD111 for antedependence
p1,p2, one way cross antedependence order 1, degree 1 : 10 parameters
a10,a11,a'10,a'11,κ'00,κ'01,b0,b1,b'0,b'1 , parameter file: para2
ide(anim)
```

For each random effect following a multiple-trait SAD model, the layout of the line indicating the SAD is as follows: Number of levels name of the variable OWNx a b .. !Tyy !Fz with:

- x = the number of parameters to estimate

- a, b, ... initial values for the parameters in this order (for 3 traits):
  - **initial values for antedependence** parameters trait 1, antedependence parameters trait2…
  - **initial values for cross-antedependence** parameters in lower triangular order (in one direction and in the other direction) i.e.
    Trait1 has an influence on trait2 (cross antedependence δ') (one direction),
    Trait2 has an influence on trait1 (cross antedependence δ) (other direction),
    trait1➔trait3 cross- antedependence parameters,
    trait3➔trait1 cross- antedependence parameters,
    trait2➔trait3 cross- antedependence parameters,
    trait3➔trait2 cross- antedependence parameters

    order of the parameters for cross antedependence

	Trait1	Trait2	Trait3
Trait 1			
Trait 2	1 (Tr1 ➔Tr2) 2 (Tr2 ➔Tr1)		
Trait3	3 (Tr1 ➔Tr3) 4 (Tr3 ➔Tr1)	5 (Tr2 ➔Tr3) 6 (Tr3 ➔Tr2)	

- **initial values for the innovation variance** trait 1, for the innovation variance trait 2, for the innovation variance trait3
- **initial correlation value** (if needed)

○ y: parameter types: we advise to write !TCCCCC….

○ z is a number between 1 and 9 that forms the link with the parameter file , if omitted, the parameter file is "para", if z = 2, the parameter file is "para2", if z = 3, the parameter file is "para3" and so on.

2.4. Parameter file

The parameter files provide information about the specification of the (cross-)antedependence and innovation variance (i.e. order, degree…).

Example of parameter file for a random term that follows model in eq1 (antedependence SAD111 for both p_1 and p_2, cross-antedepence in one direction (p_1 has an influence on p_2) of order 1, degree 1 with c'=0, no initial correlation between e and ε (comments in blue, **in bold: unmodifiable words**)

NBCHARACTER
2 *# number of traits (called "trait" in the following formula)*
MODEL
SAD *# model type (do not use anything else)*
ANTEDEP_ORDER
1 1 # order antedependence: α α'
CROSS_ANTEDEP_ORDER *#it is not really the order it is=(order cross antedependence + begin_time) in triangular inferior in one direction and in the other direction 1->2,2->1,(2traits) + 1->3,3->1,2->3,3->2 (3traits) …*
1 0 # $\eta'+1$, $\eta+1$ *…as many figures as possible cross-antedependence (trait*(trait-1)), 0 if no cross antedependence* **line 1**
BEGIN_CROSS_ANTEDEP_ORDER #i.e. "begin time"
0 0 # c' c as many figures as possible cross-antedependence *(trait*(trait-1)) , 0 if no cross antedependence* **line 2**
MODEL_ANTEDEP *#degree of the polynomial function for each antedependence parameters trait1, trait2…*
1 1 # β_1, β'_1 as many figures as the sum of the antedependence orders
MODEL_CROSS_ANTEDEP *#degree of the polynomial function for each cross-antedependence parameters, write nothing if there is no cross-antedependence*
1 # ϖ'_0 (as many figures as the sum of cross-antedep order ($\sum(\text{line } 1 - \text{line } 2)$)
MODEL_INNOVATION_VARIANCE_COVARIANCE *#degree of the polynomial function for each innovation variance trait1, trait2…*
0 0 # γ, γ' as many figures as the number of traits
CORRINIT *#0 = no initial correlation between error terms, 1 = initial correlation at time t0 (in triangular inferior order)*
0 #as many figures as trait*(trait-1)/2
DATA_FILE #*same as in .as*
datafile
COLUMN_FACTOR *#the number of the column containing "time" (real time points)*
1

2.5. Results

- The **.asr** file contains the output of the model for the parameter estimates.

Example of .asr file for antedependence SAD111 for independent u_1 and u_2. Antedependence SAD111 for both p_1 and p_2, cross-antedepence in one direction (p_1 has an influence on p_2) of order 1, degree 1 with c'=0, no initial correlation between e and ε (see equation 1)

……. Source	Model	terms	Gamma	Component	Comp/SE	% C	
Residual	identity	12384	0.100000E-03	0.100000E-03	0.00	0 F	*#fixed residual*
time.animal	OWNxx	8	-0.126282	-0.126282	-0.45	0 U	*# $\hat{a}_{10}$ for genetic effect u1*
time.animal	OWNxx	8	0.193263	0.193263	1.73	0 U	*# $\hat{a}_{11}$ for genetic effect u1*
time.animal	OWNxx	8	-0.226282	-0.226282	-1.45	0 U	*# $\hat{a}'_{10}$ for genetic effect u2*
time.animal	OWNxx	8	0.93263	0.93263	1.39	0 U	*# $\hat{a}'_{11}$ for genetic effect u2*
time.animal	OWNxx	8	2.88308	2.88308	11.08	0 U	*# $\hat{b}_0$ for genetic effect u1*
time.animal	OWNxx	8	-0.632185	-0.632185	-3.59	0 U	*# $\hat{b}_1$ for genetic effect u1*
time.animal	OWNxx	8	2.30811	2.30811	1.08	0 U	*# $\hat{b}'_0$ for genetic effect u2*
time.animal	OWNxx	8	-0.32185	-0.32185	-1.59	0 U	*# $\hat{b}'_1$ for genetic effect u2*
time.ide(ani	OWNxx	10	0.238707	0.238707	0.54	0 U	*# $\hat{a}_{10}$ for effect p1*
time.ide(ani	OWNxx	10	-0.239293	-0.239293	-1.71	0 U	*# $\hat{a}_{11}$ for effect p1*
time.ide(ani	OWNxx	10	0.38707	0.38707	5.4	0 U	*# $\hat{a}'_{10}$ for effect p2*
time.ide(ani	OWNxx	10	-0.39293	-0.39293	-7.1	0 U	*# $\hat{a}'_{11}$ for effect p2*
time.ide(ani	OWNxx	10	0.707	0.707	1.54	0 U	*# $\hat{\kappa}'_{00}$ for p1 ➔p2*
time.ide(ani	OWNxx	10	-1.239293	-1.239293	-1.1	0 U	*# $\hat{\kappa}'_{01}$ for p1 ➔p2*
time.ide(ani	OWNxx	10	1.14631	1.14631	3.36	0 U	*# $\hat{b}_0$ for effect p1*
time.ide(ani	OWNxx	10	0.240279	0.240279	1.88	0 U	*# $\hat{b}_1$ for effect p1*
time.ide(ani	OWNxx	10	1.4631	1.4631	3.6	0 U	*# $\hat{b}'_0$ for effect p2*
time.ide(ani	OWNxx	10	0.40279	0.40279	0.88	0 U	*# $\hat{b}'_1$ for effect p2*

- The **.gdg** file contains the inverse of the variance-covariance matrix for each random term. The matrix is sorted by trait-time i.e. corresponds to the matrix for the vector $\left[p_1(t_0) \quad p_1(t_1) \quad \mathrm{K} \quad p_1(t_n) \quad p_2(t_0) \quad p_2(t_1) \quad \mathrm{K} \quad p_2(t_n)\right]$

If "para" is used, the matrix is in asfilenameB.gdg
If "para2" is used, the matrix is in asfilenameC.gdg
If "para3" is used, the matrix is in asfilenameD.gdg…

To obtain the co(variance) matrix for each random effects (time x trait*time x trait covariance matrix, named B), you have 3 options:

- you use parameter estimates in the .asr file to compute the L and D^{-1} matrices (see articles for details) and $B=\left(L'D^{-1}L\right)^{-1}$. If you choose this option, don't forget that the parameter estimates are for the time vector on the [-1,1] scale.
- you use the .gdg file to obtain values of the B^{-1} matrix and you calculate its inverse.
- you use "matrixSAD.f90" program that reads the .gdg file and computes the inverse of B^{-1} for you.

3. EXPLANATION FOR A SINGLE TRAIT MODEL WITH CORRELATED RANDOM EFFECTS

Details in :David I., Sanchez J-P., Piles M., 2018. Longitudinal analysis of direct and indirect effects on average daily gain in rabbits using a structured antedependence model . GSE

This program is useful to analyze trait with direct and indirect genetic effects. For instance, suppose that animals are raised in pen of two, the model for animal *i* living with animal *l* is

$$y_i(t_j) = \mu_i(t_j) + d_i(t_j) + s_l(t_j) + p_i(t_j)$$

With, $\boldsymbol{\mu}(t_j)$ the fixed effect at time t_j, $\boldsymbol{d}(t_j)$ the vector of direct genetic effects at time t_j, $\boldsymbol{s}(t_j)$ the vector of indirect genetic effects at time t_j, $\boldsymbol{p}(t_j)$ the vector of pseudo-permanent environmental effect at time t_j.

Specification of the SAD model for $\boldsymbol{p}$ is the same as in the single trait SAD model.

Specification of the SAD model for correlated *s* and *d* is performed using the multiple trait SAD model with constraint (d and s are two correlated terms as the random terms of different traits in a multiple trait model that are also correlated). Constraints that need to be done on the multiple trait SAD model to consider correlated random effects within trait are: cross-antedependence in one direction only (the one you want), the order of the cross antedependence is 1 and c (or c' if you chose the other direction) is 0 (consequently, no initial correlation). The correlated-effects SAD model of antedependence order α, α' can then be written as (for $j > \max(\alpha, \alpha')$):

$$\boldsymbol{d}(t_j) = \sum_{s=1}^{\alpha} \theta_{sj} \boldsymbol{d}(t_{j-s}) + \delta_j \boldsymbol{s}(t_j) + \boldsymbol{e}_d(t_j)$$

$$\boldsymbol{s}(t_j) = \sum_{s=1}^{\alpha'} \theta'_{sj}\, \boldsymbol{s}(t_{j-s}) + \boldsymbol{e}_s(t_j)$$

$$\theta_{sj} = \sum_{k=0}^{\beta_s} a_{sk} t_j^k,\ \theta'_{sj} = \sum_{k=0}^{\beta'_s} a'_{sk}\, t_j^k$$

with $\delta_j = \sum_{k=0}^{\omega} \kappa_k t_j^k$

$$\sigma^2_{ed,j} = \exp\left(\sum_{k=0}^{\gamma} b_k t_j^k\right),\ \sigma^2_{es,j} = \exp\left(\sum_{k=0}^{\gamma'} b'_k\, t_j^k\right)$$

Thus a correlated effects SAD model is defined (2 random effects) by: the order of the antedependence for each random effect (α, α'), the degree of the polynomial for the antedependence parameter (β_1 to β_α, β'_1 to $\beta'_{\alpha'}$), the degree of the polynomial for the cross-antedependence parameter (ϖ_1 or ϖ'_1) and the degree of the polynomial for the innovation variance of each random effect (γ, γ').

For instance, a correlated effects SAD model, SAD111 for antedependence effect 1, SAD111 for antedependence effect 2 and a cross antedependence of degree 1 is:

$$\boldsymbol{d}\left(t_j\right)=\theta_{1j}\boldsymbol{d}\left(t_{j-1}\right)+\delta_j\boldsymbol{s}\left(t_j\right)+\boldsymbol{e}_d\left(t_j\right)$$
$$\boldsymbol{s}\left(t_j\right)=\theta'_{1j}\,\boldsymbol{s}\left(t_{j-1}\right)+\boldsymbol{e}_s\left(t_j\right)$$

eq. 2

$$\theta_{1j}=a_{10}+a_{11}t_j,\quad \theta'_{1j}=a'_{10}+a'_{11}t_j$$
$$\delta_j=\kappa_0+\kappa_1 t_j$$
$$\sigma^2_{ed,j}=\exp(b_0+b_1t_j),\quad \sigma^2_{es,j}=\exp(b'_0+b'_1t_j)$$

3.1. Datafile

The datafile is the same as for the single trait analysis with an additional column containing the ID of the co-mate. **It is mandatory to sort the data by time-animal. If you do not, check carefully that the time vector constructed by ASreml (same order as in the dataset) is in the same order as the one used for the SAD model (the latter appears on the screen when ASreml is running= figures after "time vector")**.

Datafile example for correlated SAD model: successive measurements on 4 days: 1, 3, 8, and 11

time	animal	Co-mate	factor	phenotype
1	1	2	1	43
1	2	3	2	21
1	3	4	1	25
1	4	1	1	14
3	2	3	2	52
3	3	4	2	62
3	4	1	3	42
8	1	2	3	51
8	2	3	1	42
8	3	4	2	12
8	4	1	1	45
11	1	2	3	52
11	2	3	2	14

3.2. Pedigree

As recommended by ASReml.

3.3. File.as

As for the file.as for single trait analysis and multiple trait analysis, because a pseudo-permanent environmental effect is included (*p*) in the model, there is no classical

residual in this model. **However, even if theoretically the residual variance should be fixed to 0, our experience is that a too low value of the residual variance/other variances induces convergence to abnormal parameter values. Thus we recommend fixing the residual variance to a value that is around 100 times lower than the other variances of the model.**

If you do not include a pseudo-permanent environmental effect that follows a SAD in your model or if the SAD model for the pseudo-permanent effect is simple, then you can consider a classical residual and you do not have to fix its variance. However, you have to keep the !S2==1 option when specifying the model. *Ex1 in the examples part of the tutorial provides an example of SAD model with and without a true residual.*

Example of .as for a correlated effects SAD model. SAD111 for d, s and p , cross antedependence between d and s of degree 1 (total: 14 parameters)

```
sadcorrelated
 time !A #needed
 animal !P
 comate !P
 factor !A
 character !A
 phenotype
Pedigree
Datafile !OWN sadmultiuni_exe #link to the fortran executable
Phenotype ~ character.time character. factor !r ![ time.animal time.comate !]  time.ide(animal) #one G matrix (8*8) will be used for time.animal time.comate
1 1 2
0 0 IDV 0.001 !GF !S2==1  #to fix residual variance to a very low value : 0.001 ⚠ NOT TOO LOW
time.animal 2
8 0 OWN8 OWN10 0.6 0.1 0.8 0.4 0.5 0.1 1 2 -0.1 0.1 !TCCCCCCCCCC  #SAD111 for antedependence s and d, one way cross antedependence order 1 (constraint), degree 1, 10 parameters
 a10,a11,a'10,a'11,κ00,κ01,b0,b1,b'0,b'1  file: para
animal
time.ide(animal) 2
4 time OWN4 0.1 0.1 0.9 0.9 !TCCCC !F2 #SAD111 for p 4 parameters  a10,a11,b0,b1 , parameter file: para2
ide(anim)
```

For each random effect following a SAD model, the layout of the line indicating the SAD is as for the single and multiple trait cases.

3.4. Parameter file

The parameter files provide information about the specification of the (cross-)antedependence and innovation variance (i.e. order, degree…).

Example of parameter file for genetic correlated effects (s,d) described in eq2 (antedependence SAD111 for both s and d, cross-antedepence in one direction (s has an influence on d) of degree 1 with c=0, (comments in blue, **in bold: unmodifiable words**)

NBCHARACTER
2 *# number of traits (called "trait" in the following formula)*
MODEL
SAD *# model type (do not use anything else)*
ANTEDEP_ORDER
1 1 # order antedependence: α α'
CROSS_ANTEDEP_ORDER *#for 2 correlated effects, 1 0 or 0 1 (one direction or the other, not both)…*
0 1 # η'+1 , η +1…*as many figures as possible cross-antedependence (*trait*(trait-1)), if no cross antedependence:* 0 **line 1**
BEGIN_CROSS_ANTEDEP_ORDER #i.e. "begin time"
0 0 #a line with trait*(trait-1) "0" **line 2**
MODEL_ANTEDEP *#degree of the polynomial function for each antedependence parameters trait1, trait2…*
1 1 # β_1, β'_1 as many figures as the sum of the antedependence orders
MODEL_CROSS_ANTEDEP *#degree of the polynomial function for the cross-antedependence parameters, write nothing if there is no cross-antedependence*
1 # ϖ_0 (as many figures as the sum of cross-antedep order ($\sum(\text{line } 1 - \text{line } 2)$)
MODEL_INNOVATION_VARIANCE_COVARIANCE *#degree of the polynomial function for each innovation variance trait1, trait2…*
1 1 # γ, γ' as many figures as the number of traits
CORRINIT *#for correlated effects 0 = no initial correlation between error terms*
0 #as many "0" as trait*(trait-1)/2
DATA_FILE #*same as in .as*
datafile
COLUMN_FACTOR *#the number of the column containing "time" (real time points)*
1

3.5. Results

- The **.asr** file contains the output of the model for the parameter estimates (see .1.4 and 2.4 for details)
- The **.gdg** file contains the covariance matrix (inverse and derivatives). For the correlated random effects, the B matrix is given for the vector: [d1,d2,d3…s1,s2,s3] corresponding to the order in which correlated effects appear in the line specifying model (i.e. Phenotype ~ character.time character. factor !r **![time.animal (=d) time.comate (=s) !]** time.ide(animal))

4. EXPLANATION TO COMPUTE B MATRIX USING matrixSAD_exe

This program uses the results in the .gdg file to calculate the variance covariance matrix B and the variance correlation matrix for each random term. For the executable to work properly, the parameter estimates have to be on the same line in the .gdg file. If it is not the case, put them on the same line before running matrixSAD_exe
Type matrixSAD_exe on the command line
The program will ask you the name of the .gdg file
Then B^{-1}, B and the corresponding correlation matrix (variance on the diagonal, correlation elsewhere) will appear on the screen. B (var-covariance) and B (var-correlation) will be written in a "resumat" file also.

5. EXAMPLES

5.1. Examples with .as, datafiles, parafiles available

5.1.1. Ex1.as example

Study of average daily gain (simulated data) for 960 animals raised in pen of 8, 5 measurements per animal (week 1 to 5).

The datafile **dataex1** is of the form:

#simulation	#generation	Animal ID	Comate1 ID	…	Comate7 ID	pen	week	ADG	group
1	1	2279	2283	…	2940	1	1	31.49	1
1	1	2279	2283	…	2940	1	2	31.72	1

The pedigree file is **pedex1**

5.1.1.1. Single trait analysis

- The model is:

$\boldsymbol{ADG}(t_i) = week_i + \boldsymbol{u}_i + \boldsymbol{p}_i + \boldsymbol{g}_i$

$\boldsymbol{u}$ the direct genetic effects for week *i* that follows a SAD111 (t_i=i because week=[1,2,3,4,5])

$\boldsymbol{u}_1 = \boldsymbol{e}_{u,1}$, $\boldsymbol{u}_i = (a_{u0} + a_{u1} * i)\boldsymbol{u}_{i-1} + \boldsymbol{e}_{u,i}$ and $\sigma^2_{e_{u,i}} = \exp(b_{u0} + b_{u1} * i)$

$\boldsymbol{p}$ the pseudo-permanent environmental effects for week i that follows a SAD111:

$\boldsymbol{p}_1 = \boldsymbol{e}_{p,1}$, $\boldsymbol{p}_i = (a_{p0} + a_{p1} * i)\boldsymbol{p}_{i-1} + \boldsymbol{e}_{p,i}$ and $\sigma^2_{e_{p,i}} = \exp(b_{p0} + b_{p1} * i)$

$\boldsymbol{g}$ the group effects for week i that follows a SAD111 :

$\boldsymbol{g}_1 = \boldsymbol{e}_{g,1}$, $\boldsymbol{g}_i = (a_{g0} + a_{g1} * i)\boldsymbol{g}_{i-1} + \boldsymbol{e}_{g,i}$ and $\sigma^2_{e_{g,i}} = \exp(b_{g0} + b_{g1} * i)$

- The .as is:

See ex1.as PATH1

- The parafiles are :

Parafiles are the same (with different names) for all the random effects because they follow the same SAD model

See “para1ex1”, “para2ex1” and “para3ex1” (to use them, rename to “para”, “para2” and “para3”)

- Results:

LogL=-2246

Source	Model	terms	Gamma	Component	Comp/SE	% C
Residual	identity	4800	0.100000E-02	0.100000E-02	0.00	0 F
week.groupe	OWNxx	5	-0.318555E-01	-0.318555E-01	-0.35	0 U
week.groupe	OWNxx	5	0.127130	0.127130	0.94	0 U
week.groupe	OWNxx	5	2.53568	2.53568	27.50	0 U
week.groupe	OWNxx	5	0.163560	0.163560	1.36	0 U
week.aniID	OWNxx	5	0.925040	0.925040	26.55	0 U
week.aniID	OWNxx	5	-0.211101	-0.211101	-3.49	0 U
week.aniID	OWNxx	5	1.88091	1.88091	7.56	0 U
week.aniID	OWNxx	5	-1.38197	-1.38197	-5.56	0 U
week.ide(aniID)	OWNxx	5	0.427733E-01	0.427733E-01	1.15	0 U
week.ide(aniID)	OWNxx	5	0.246649E-01	0.246649E-01	0.44	0 U
week.ide(aniID)	OWNxx	5	3.64844	3.64844	93.74	0 U
week.ide(aniID)	OWNxx	5	0.406783	0.406783	7.92	0 U

Estimated correlation (variance on the diagonal) matrix for week.group:

	Week1	Week 2	Week 3	Week 4	Week 5
Week 1	10.72	-0.09	0.00	0.00	0.00
Week 2	-0.09	11.73	-0.03	0.00	0.00
Week 3	0.00	-0.03	12.64	0.03	0.00
Week 4	0.00	0.00	0.03	13.71	0.09
Week 5	0.00	0.00	0.00	0.09	14.99

Estimated correlation (variance on the diagonal) matrix for week.aniID

26.12	0.82	0.76	0.72	0.68
0.82	40.84	0.92	0.87	0.83
0.76	0.92	41.52	0.95	0.90
0.72	0.87	0.95	31.17	0.95
0.68	0.83	0.90	0.95	17.54

Estimated correlation (variance on the diagonal) matrix for week.ide(aniID)

25.57	0.03	0.00	0.00	0.00
0.03	31.37	0.04	0.00	0.00
0.00	0.04	38.47	0.05	0.00
0.00	0.00	0.05	47.19	0.06
0.00	0.00	0.00	0.06	57.90

5.1.1.2. Single trait analysis – version with "true"residual

- The model is:

$\boldsymbol{ADG}(t_i) = week_i + \boldsymbol{u}_i + \boldsymbol{p}_i + \boldsymbol{g}_i + \boldsymbol{\varepsilon}_i$

$\boldsymbol{u}$ the direct genetic effects for week *i* that follows a SAD111 (t_i=i because week=[1,2,3,4,5])

$\boldsymbol{u}_1 = \boldsymbol{e}_{u,1}$, $\boldsymbol{u}_i = (a_{u0} + a_{u1} * i)\boldsymbol{u}_{i-1} + \boldsymbol{e}_{u,i}$ and $\sigma^2_{e_{u,i}} = \exp(b_{u0} + b_{u1} * i)$

$\boldsymbol{p}$ the pseudo-permanent environmental effects for week i that follows a SAD111:

$\boldsymbol{p}_1 = \boldsymbol{e}_{p,1}$, $\boldsymbol{p}_i = (a_{p0} + a_{p1} * i)\boldsymbol{p}_{i-1} + \boldsymbol{e}_{p,i}$ and $\sigma^2_{e_{p,i}} = \exp(b_{p0} + b_{p1} * i)$

$\boldsymbol{g}$ the group effects for week i that follows a SAD111 :

$\boldsymbol{g}_1 = \boldsymbol{e}_{g,1}$, $\boldsymbol{g}_i = (a_{g0} + a_{g1} * i)\boldsymbol{g}_{i-1} + \boldsymbol{e}_{g,i}$ and $\sigma^2_{e_{g,i}} = \exp(b_{g0} + b_{g1} * i)$

$\boldsymbol{\varepsilon}$ the residual (homogeneous with time) $\boldsymbol{\varepsilon} \sim \boldsymbol{N}(0, I\sigma^2_{\varepsilon})$

- The .as is:

See ex1.as PATH2

- The parafiles are :

Parafiles are the same (with different names) for all the random effects because they follow the same SAD model

See "para1ex1", "para2ex1" and "para3ex1" (to use them, rename to "para", "para2" and "para3")

- Results:

logL=-2246

Source	Model	terms	Gamma	Component	Comp/SE	% C
Residual	identity	4800	10.4926	10.4926	0.55	0 U
week.groupe	OWNxx	5	-0.316541E-01	-0.316541E-01	-0.35	0 U
week.groupe	OWNxx	5	0.128340	0.128340	0.95	0 U
week.groupe	OWNxx	5	2.53406	2.53406	27.45	0 U
week.groupe	OWNxx	5	0.164622	0.164622	1.36	0 U
week.aniID	OWNxx	5	0.930556	0.930556	24.93	0 U
week.aniID	OWNxx	5	-0.222345	-0.222345	-3.33	0 U
week.aniID	OWNxx	5	1.89759	1.89759	7.69	0 U
week.aniID	OWNxx	5	-1.34528	-1.34528	-5.28	0 U
week.ide(aniID)	OWNxx	5	0.616740E-01	0.616740E-01	0.77	0 U
week.ide(aniID)	OWNxx	5	0.200493E-01	0.200493E-01	0.23	0 U
week.ide(aniID)	OWNxx	5	3.30647	3.30647	4.40	0 U
week.ide(aniID)	OWNxx	5	0.555973	0.555973	1.48	0 U

✓ Estimated correlation (variance on the diagonal) matrix for week.group:

Similar matrix as the previous model without true residual

	Week1	Week 2	Week 3	Week 4	Week 5
Week 1	10.69	-0.09	0.00	0.00	0.00
Week 2	-0.09	11.71	-0.03	0.00	0.00
Week 3	0.00	-0.03	12.62	0.03	0.00
Week 4	0.00	0.00	0.03	13.70	0.09
Week 5	0.00	0.00	0.00	0.09	14.99

✓ Estimated correlation (variance on the diagonal) matrix for week.aniID

Similar matrix as the previous model without true residual

25.61	0.82	0.76	0.71	0.68
0.82	40.86	0.92	0.87	0.82
0.76	0.92	42.06	0.94	0.90
0.71	0.87	0.94	31.65	0.95
0.68	0.82	0.90	0.95	17.61

✓ Estimated correlation (variance on the diagonal) matrix for week.ide(aniID)

15.63	0.04	0.00	0.00	0.00
0.04	20.69	0.05	0.00	0.00
0.00	0.05	27.34	0.06	0.00
0.00	0.00	0.06	36.15	0.07
0.00	0.00	0.00	0.07	47.80

Variances on the diagonal are the same as the one obtained without residual minus the value of the residual variance

✓ Estimated correlation (variance on the diagonal) matrix for week.ide(aniID)+residual

Similar matrix as the one obtained for week.ide(aniID) in the model without residual

26.12	0.03	0.00	0.00	0.00
0.03	31.18	0.04	0.00	0.00
0.00	0.04	37.84	0.05	0.00
0.00	0.00	0.05	46.64	0.06
0.00	0.00	0.00	0.06	58.29

Conclusion: including a true residual gave the same variance component estimates, same logL but with an additional parameter to estimate. The model without residual is thus preferable.

5.1.1.3. Single trait analysis with correlated random effects

This time we include the social effects in the model.

- The model is:

$$\boldsymbol{ADG}(t_i) = week_i + \boldsymbol{u}_i + \boldsymbol{p}_i + \sum_1^7 \boldsymbol{s}_i + \boldsymbol{g}_i$$

$\boldsymbol{p}, \boldsymbol{g}$ are the same as in the previous exemple.

$\boldsymbol{u}, \boldsymbol{s}$ are the correlated direct and social genetic effects for week i that both follows a SAD111, the cross antedependence is of degree 1.

$$\boldsymbol{u}_1 = \boldsymbol{e}_{u,1}, \quad \boldsymbol{u}_i = (a_{u0} + a_{u1} * i)\boldsymbol{u}_{i-1} + \boldsymbol{e}_{u,i}$$
$$\boldsymbol{s}_1 = (\kappa_0 + \kappa_1 * i)\boldsymbol{u}_1 + \boldsymbol{e}_{s,1}, \quad \boldsymbol{s}_i = (\kappa_0 + \kappa_1 * i)\boldsymbol{u}_i + (a_{s0} + a_{s1} * i)\boldsymbol{s}_{i-1} + \boldsymbol{e}_{s,i}$$

and

$$\sigma^2_{e_{u,i}} = \exp(b_{u0} + b_{u1} * i)$$
$$\sigma^2_{e_{s,i}} = \exp(b_{s0} + b_{s1} * i)$$

- The .as is:

See ex1.as, PATH 3

- The parafiles :

Parafiles are the same as in the previous example for p and g, for u and s it is “para4ex1” :

(to use them, rename to “para”, “para3”, “para4”)

- Results:

```
Residual             identity  4800  0.100000E-02  0.100000E-02   0.00   0 F
 week.groupe          OWNxx        5  0.170995E-01  0.170995E-01   0.18   0 U
 week.groupe          OWNxx        5 -0.297556E-01 -0.297556E-01  -0.20   0 U
 week.groupe          OWNxx        5   2.38097       2.38097      21.37   0 U
 week.groupe          OWNxx        5  0.166689      0.166689       1.04   0 U
 week.aniID           OWNxx       10  0.938876      0.938876      24.59   0 U
 week.aniID           OWNxx       10 -0.310897     -0.310897      -4.60   0 U
 week.aniID           OWNxx       10  0.182829      0.182829       1.20   0 U
 week.aniID           OWNxx       10   1.45198       1.45198       5.43   0 U
 week.aniID           OWNxx       10 -0.831765E-01 -0.831765E-01  -4.37   0 U
 week.aniID           OWNxx       10  0.128416      0.128416       3.73   0 U
 week.aniID           OWNxx       10   1.90384       1.90384       8.42   0 U
 week.aniID           OWNxx       10  -1.06136      -1.06136      -4.67   0 U
 week.aniID           OWNxx       10  -3.61676      -3.61676      -1.85   0 U
 week.aniID           OWNxx       10  -1.88848      -1.88848      -0.84   0 U
 week.ide(aniID)      OWNxx        5  0.438980E-01  0.438980E-01   1.14   0 U
 week.ide(aniID)      OWNxx        5  0.629396E-02  0.629396E-02   0.11   0 U
 week.ide(aniID)      OWNxx        5   3.62167       3.62167      89.44   0 U
 week.ide(aniID)      OWNxx        5  0.421749      0.421749       7.91   0 U
```

- ✓ Estimated correlation (variance on the diagonal) matrix week.group

9.16	0.03	0.00	0.00	0.00
0.03	9.96	0.02	0.00	0.00
0.00	0.02	10.82	0.00	0.00
0.00	0.00	0.00	11.75	-0.01
0.00	0.00	0.00	-0.01	12.78

- ✓ Estimated genetic correlation (variance on the diagonal) matrix

Direct genetic effects					Social genetic effects				
Week 1	Week 2	Week 3	Week 4	Week 5	Week 1	Week 2	Week 3	Week 4	Week 5
19.40	0.82	0.74	0.68	0.62	-0.91	-0.32	-0.68	-0.68	-0.66
0.82	34.64	0.91	0.84	0.76	-0.75	-0.71	-0.89	-0.88	-0.87
0.74	0.91	37.25	0.92	0.84	-0.68	-0.64	-0.95	-0.95	-0.93
0.68	0.84	0.92	26.81	0.91	-0.62	-0.59	-0.88	-0.90	-0.86
0.62	0.76	0.84	0.91	12.89	-0.57	-0.54	-0.80	-0.82	-0.75
-0.91	-0.75	-0.68	-0.62	-0.57	1.04	0.14	0.59	0.60	0.58
-0.32	-0.71	-0.64	-0.59	-0.54	0.14	0.41	0.72	0.70	0.71
-0.68	-0.89	-0.95	-0.88	-0.80	0.59	0.72	0.37	0.98	0.98
-0.68	-0.88	-0.95	-0.90	-0.82	0.60	0.70	0.98	0.43	0.99
-0.66	-0.87	-0.93	-0.86	-0.75	0.58	0.71	0.98	0.99	0.89

- ✓ Estimated pseudo permanent environmental correlation (variance on the diagonal) matrix

24.53	0.04	0.00	0.00	0.00
0.04	30.33	0.04	0.00	0.00
0.00	0.04	37.46	0.04	0.00
0.00	0.00	0.04	46.26	0.05
0.00	0.00	0.00	0.05	57.14

5.1.2. Ex2 exemple: multiple trait analysis

We consider 2 traits, 6 measurements per trait (time t_0 to t_5: 1,2,3,4,5,10).note: simulated data without any subjective interpretation.

The datafile **dataex2** is as follow (factor: 10 levels):

animal ID	permanent ID	factor	week	week2	character	phenotype
35	1	2	1	1	1	2.1621369
35	1	5	2	2	1	-0.24649745
35	1	9	3	3	1	3.97515901
35	1	5	4	4	1	2.48136414
35	1	8	5	5	1	0.8077682
35	1	5	10	6	1	1.82469341
36	2	2	1	1	1	6.35244978
36	2	6	2	2	1	3.40466406
36	2	8	3	3	1	4.05081936
36	2	7	4	4	1	1.94876435
36	2	5	5	5	1	8.20519782
36	2	10	10	6	1	3.6706812
…	…	…	…	…	…	…
35	1	2	1	7	2	1.33369744
35	1	5	2	8	2	-1.05635543
35	1	9	3	9	2	1.32508293
35	1	5	4	10	2	1.58862245
35	1	8	5	11	2	3.75727085
35	1	5	10	12	2	7.28011657
36	2	2	1	7	2	3.68366083
36	2	6	2	8	2	2.79504194
36	2	8	3	9	2	0.7385331
36	2	7	4	10	2	2.52987967
36	2	5	5	11	2	1.07048596
36	2	10	10	12	2	-2.46401598

The pedigree is **ped2ex**

- The model is:

$$\boldsymbol{y}_1(t_i) = factor_1 + \boldsymbol{u}_{1i} + \boldsymbol{p}_{1i}$$
$$\boldsymbol{y}_2(t_i) = factor_2 + \boldsymbol{u}_{2i} + \boldsymbol{p}_{2i}$$

With :

SAD100 for u_1 and u_2, a one way cross antedependence of order 1 with c=1 and an initial correlation:

$$\boldsymbol{u}_{1,t_0} = \boldsymbol{e}_{u1,0},\quad \boldsymbol{u}_{1,i} = \left(a_{u1}\right)\boldsymbol{u}_{1,t_{i-1}} + \left(\kappa_{u1,0} + \kappa_{u1,1}t_i\right)\boldsymbol{u}_{2,t_{i-1}} + \boldsymbol{e}_{u1,i}$$
$$\boldsymbol{u}_{2,t_0} = \boldsymbol{e}_{u2,0},\quad \boldsymbol{u}_{2,i} = \left(a_{u2}\right)\boldsymbol{u}_{2,t_{i-1}} + \boldsymbol{e}_{u2,i}$$
with $\sigma^2_{e_{u1,i}} = \exp\left(b_{u1}\right)$, $\sigma^2_{e_{u2,i}} = \exp\left(b_{u2}\right)$, $\operatorname{cov}(e_{u1,0}, e_{u2,0}) = \text{constant}$

7 parameters

SAD100 for p_1, SAD101 for p_2, p_1 and p_2 are independent:

$$\boldsymbol{p}_{1,t_0} = \boldsymbol{e}_{p1,0},\quad \boldsymbol{p}_{1,i} = \left(a_{p1}\right)\boldsymbol{p}_{1,t_{i-1}} + \boldsymbol{e}_{p1,i}$$
$$\boldsymbol{p}_{2,t_0} = \boldsymbol{e}_{p2,0},\quad \boldsymbol{p}_{2,i} = \left(a_{p2}\right)\boldsymbol{p}_{2,t_{i-1}} + \boldsymbol{e}_{p2,i}$$
with $\sigma^2_{e_{p1,i}} = \exp\left(b_{p1}\right)$, $\sigma^2_{e_{p2,i}} = \exp\left(b_{p2,0} + b_{p2,1}t_i\right)$

5 parameters

- <u>The .as is:</u>

See ex2.as

- <u>The parafiles are:</u>

See paraex2 para2ex2 (to use them, rename to para para2)

- <u>Results:</u>

```
Source               Model  terms      Gamma      Component    Comp/SE   % C
 Residual           identity  2160  0.100000E-02  0.100000E-02    0.00    0 F
 week2.aniID        OWNxx       12  0.862414       0.862414       2.48    0 U
 week2.aniID        OWNxx       12  0.632747       0.632747       4.32    0 U
 week2.aniID        OWNxx       12  0.722647       0.722647       1.02    0 U
 week2.aniID        OWNxx       12   1.56582        1.56582       1.43    0 U
 week2.aniID        OWNxx       12 -0.405226      -0.405226      -0.40    0 U
 week2.aniID        OWNxx       12  -2.03149       -2.03149      -0.85    0 U
 week2.aniID        OWNxx       12   3.35849        3.35849       1.56    0 U
 week2.perm         OWNxx       12  0.491626       0.491626      13.98    0 U
 week2.perm         OWNxx       12  0.312969       0.312969       9.91    0 U
 week2.perm         OWNxx       12   3.02890        3.02890      58.09    0 U
 week2.perm         OWNxx       12   3.63538        3.63538      73.68    0 U
 week2.perm         OWNxx       12 -0.366708      -0.366708      -5.30    0 U
```

✓ estimated genetic correlation (variance on the diagonal) matrix

Trait 1						Trait 2					
T0	t1	T2	T3	T4	T5	T0	T1	T2	T3	T4	T5
0.67	-0.61	-0.60	-0.36	-0.03	0.46	0.99	0.97	0.92	0.83	0.69	0.51
-0.61	1.13	0.79	0.59	0.32	-0.12	-0.63	-0.62	-0.59	-0.53	-0.44	-0.33
-0.60	0.79	1.87	0.78	0.48	-0.03	-0.61	-0.61	-0.58	-0.52	-0.43	-0.32
-0.36	0.59	0.78	1.79	0.74	0.25	-0.37	-0.36	-0.33	-0.30	-0.25	-0.18
-0.03	0.32	0.48	0.74	1.90	0.61	-0.03	-0.02	0.02	0.08	0.07	0.05
0.46	-0.12	-0.03	0.25	0.61	4.45	0.46	0.48	0.53	0.62	0.73	0.55
0.99	-0.63	-0.61	-0.37	-0.03	0.46	7.64	0.98	0.93	0.84	0.69	0.52
0.97	-0.62	-0.61	-0.36	-0.02	0.48	0.98	3.20	0.95	0.86	0.71	0.53
0.92	-0.59	-0.58	-0.33	0.02	0.53	0.93	0.95	1.41	0.90	0.74	0.55
0.83	-0.53	-0.52	-0.30	0.08	0.62	0.84	0.86	0.90	0.70	0.82	0.62
0.69	-0.44	-0.43	-0.25	0.07	0.73	0.69	0.71	0.74	0.82	0.41	0.75
0.51	-0.33	-0.32	-0.18	0.05	0.55	0.52	0.53	0.55	0.62	0.75	0.30

✓ estimated pseudo-permanent correlation (variance on the diagonal) matrix

20.67	0.44	0.21	0.10	0.05	0.03	0.00	0.00	0.00	0.00	0.00	0.00
0.44	25.67	0.48	0.23	0.12	0.06	0.00	0.00	0.00	0.00	0.00	0.00
0.21	0.48	26.88	0.49	0.24	0.12	0.00	0.00	0.00	0.00	0.00	0.00
0.10	0.23	0.49	27.17	0.49	0.24	0.00	0.00	0.00	0.00	0.00	0.00
0.05	0.12	0.24	0.49	27.24	0.49	0.00	0.00	0.00	0.00	0.00	0.00
0.03	0.06	0.12	0.24	0.49	27.26	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	54.71	0.31	0.10	0.03	0.01	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.31	55.79	0.32	0.11	0.03	0.01
0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.32	51.95	0.33	0.11	0.04
0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.11	0.33	47.93	0.33	0.12
0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.03	0.11	0.33	44.19	0.38
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.04	0.12	0.38	30.60

5.2. Other examples

5.2.1. Single trait analysis, 5 time points (t_1,t_2,t_3,t_4,t_5)

Model $\boldsymbol{y}(t_j) = \boldsymbol{\mu}(t_j) + \boldsymbol{u}(t_j) + \boldsymbol{p}(t_j)$

- structure for genetic effects **u** (initial values in brackets)

SAD 1-10 (0.6 -0.1 , 2)
total: 3 parameters

$$u(t_j) = (a_{u0} + a_{u1}t_j)u(t_{j-1}) + e_u(t_j), \qquad \sigma^2_{e,j} = \exp(b_{u0})$$

- structure for permanent effects

SAD 2-021 (0.66 0.7-0.11 0.01, 2.1 -0.1)
total: 6 parameters

$$p(t_j) = (a_{p,10})p(t_{j-1}) + (a_{p,20} + a_{p,21}t_j + a_{p,22}t_j^2)p(t_{j-2}) + e_p(t_j), \qquad \sigma^2_{ep,j} = \exp(b_{p0} + b_{p1}t_j)$$

- *file.as, the model*

(in blue: antedependence parameters, green: innovation variance parameters)

```
....
Phenotype ~ fixed effects !r time.animal time.ide(animal)
1 1 2
0 0 IDV 0.001 !GF !S2==1  #to fix residual variance to a very low value : 0.001
time.animal 2
5 time OWN3 0.6 -0.1 2 !TCCC
Animal
time.ide(animal) 2
5 time OWN6 0.66 0.7 -0.11 0.01 2 -0.1 !TCCCCCC !F2
ide(animal)
```

- para *(for the genetic effects)* :

```
MODEL
SAD
ANTEDEP_ORDER #order of the antedependence α
1
MODEL_ANTEDEP #degree of the polynomial functions for the antedependence parameters β1 β2 β3...βα
1
MODEL_INNOVATION_VARIANCE #degree of the polynomial function for the innovation variance γ
0
DATA_FILE
Datafile
COLUMN_FACTOR
1
```

- Para2 *(for the permanent effects)* :

```
MODEL
SAD
ANTEDEP_ORDER #order of the antedependence α
2
```

MODEL_ANTEDEP #degree of the polynomial functions for the antedependence parameters $\beta_1\ \beta_2\ \beta_3 \ldots \beta_\alpha$
0 2
MODEL_INNOVATION_VARIANCE #degree of the polynomial function for the innovation variance γ
1
DATA_FILE
Datafile
COLUMN_FACTOR
1

5.2.2. Multiple trait analysis: 2 traits, 5 time points

Model

$$\boldsymbol{y}_1(t_j) = \boldsymbol{\mu}_1(t_j) + \boldsymbol{u}_1(t_j) + \boldsymbol{p}_1(t_j)$$
$$\boldsymbol{y}_2(t_j) = \boldsymbol{\mu}_2(t_j) + \boldsymbol{u}_2(t_j) + \boldsymbol{p}_2(t_j)$$

- structure for genetic effects (initial values in brackets)

trait1: antedependence SAD 1-10 (0.6 -0.1 , 2)
trait2: antedependence SAD1-11 (0.66 -0.11, 1 -0.1)
(4 parameters antedependence, 3 parameters innovation variances)

crossantedependence 1->2: SAD1-0, starting point:2, initial correlation between error terms (2 parameters: 0.8, 0.5)
crossantedependence 2->1: nothing

Total: 9 parameters

$$u_1(t_j) = (a_{u1,0} + a_{u1,1}t_j)u_1(t_{j-1}) + e_{u1}(t_j) \qquad \sigma^2_{eu1,j} = \exp(b_{u1,0})$$
$$u_2(t_j) = (a_{u2,0} + a_{u2,1}t_j)u_2(t_{j-1}) + (\kappa_{u2,0})u_1(t_{j-2}) + e_{u2}(t_j)' \qquad \sigma^2_{eu2,j} = \exp(b_{u2,0} + b_{u2,1}t_j)$$

And $corr(e_{u1}(t_0), e_{u2}(t_0)) = \text{constant}$

- structure for permanent effects

trait1: antedependence SAD 1-21 (0.6 -0.1 0.01, 2 -0.1)
trait2: antedependence SAD1-21 (0.66 -0.11 0.11, 1 -0.11)

crossantedependence 1->2: SAD1-0, starting point:0 (0.8)
crossantedependence 2->1: nothing

total: 11 parameters

$$p_1(t_j) = (a_{p1,0} + a_{p1,1}t_j + a_{p1,1}t_j^2)p_1(t_{j-1}) + e_{p1}(t_j) \qquad \sigma^2_{ep1,j} = \exp(b_{p1,0} + b_{p1,1}t_j)$$
$$p_2(t_j) = (a_{p2,0} + a_{p2,1}t_j + a_{p2,1}t_j^2)p_2(t_{j-1}) + (\kappa_{p2,0})p_1(t_j) + e_{p2}(t_j)' \qquad \sigma^2_{ep2,j} = \exp(b_{p2,0} + b_{p2,1}t_j)$$

- *file.as, the model*

(in blue: antedependence parameters, red:cross antedependence parameters, green: innovation variance parameters, black:initial correlation)

```
Phenotype ~ fixed effects !r time2.animal time2.ide(animal)
1 1 2
0 0 IDV 0.001 !GF !S2==1  #to fix residual variance to a very low value : 0.001
time2.animal 2
10 time2 OWN9 0.6 -0.1 0.66 -0.11 0.8 2 1 -0.1 0.5 !TCCCCCCCCC
Animal
```

```
time2.ide(animal) 2
10 time2 OWN11 0.6 -0.1 0.01 0.66 -0.11 0.11 0.8 2 -0.1 1 -0.11 !TCCCCCCCCCCC !F2
ide(animal)
```

- *Para (for the genetic effects)*

```
NBCHARACTER
2
MODEL
SAD
ANTEDEP_ORDER
1 1
CROSS_ANTEDEP_ORDER
3 0
BEGIN_CROSS_ANTEDEP_ORDER
2 0
MODEL_ANTEDEP
1 1
MODEL_CROSS_ANTEDEP
0
MODEL_INNOVATION_VARIANCE_COVARIANCE
0 1
CORRINIT
1
DATA_FILE #same as in .as
data
COLUMN_FACTOR
1
```

- *para2 (for the permanent effects)*

```
NBCHARACTER
2
MODEL
SAD
ANTEDEP_ORDER
1 1
CROSS_ANTEDEP_ORDER
1 0
BEGIN_CROSS_ANTEDEP_ORDER
0 0
MODEL_ANTEDEP
2 2
MODEL_CROSS_ANTEDEP
0
MODEL_INNOVATION_VARIANCE_COVARIANCE
1 1
CORRINIT
0
DATA_FILE #same as in .as
data
COLUMN_FACTOR
1
```

5.2.3. Multiple trait analysis: 3 traits, 10 time points

Model

$$\begin{aligned} y_1(t_j) &= \mu_1(t_j) + u_1(t_j) + p_1(t_j) \\ y_2(t_j) &= \mu_2(t_j) + u_2(t_j) + p_2(t_j) \\ y_3(t_j) &= \mu_3(t_j) + u_3(t_j) + p_3(t_j) \end{aligned}$$

Note: it is a very complex model to provide an example; it may not converge.

- structure for genetic effects (initial values in brackets)

trait1: antedependence SAD 1-10 (0.6 -0.1 , 2)
trait2: antedependence SAD1-11 (0.66 -0.11, 1 -0.1)
trait3:antedependence: SAD2-100 (0.7 -0.1, 0.5, 2.1)
(7 parameters antedependence, 4 parameters innovation variances)

crossantedependence trait 1->2: no relationship, no initial correlation between error terms
crossantedependence trait 2->1: no relationship
crossantedependence traits 1->3: recursive relationship (trait1 has an effect on trait3 but not the reverse) of order 1, starting point: c'=0, degree 1. no initial correlation between error terms (0.8 -0.2)
(2 parameters)
crossantedependence traits 3->1: no relationship
cross antedependence traits 2->3 no relationship, no initial correlation between error terms
cross antedependence traits 3->2 no relationship, no initial correlation between error terms

total: 13 parameters

$$\begin{aligned} u_1(t_j) &= (a_{u1,0} + a_{u1,1}t_j)u_1(t_{j-1}) + e_{u1}(t_j), \\ u_2(t_j) &= (a_{u2,0} + a_{u2,1}t_j)u_2(t_{j-1}) + e_{u2}(t_j), \\ u_3(t_j) &= (a_{u3,10} + a_{u3,11}t_j)u_3(t_{j-1}) + (a_{u3,20})u_3(t_{j-2}) + (\kappa_{u3,0} + \kappa_{u3,1}t_j)u_1(t_j) + e_{u3}(t_j), \end{aligned}$$

$$\begin{aligned} \sigma^2_{eu1,j} &= \exp(b_{u1,0}) \\ \sigma^2_{eu2,j} &= \exp(b_{u2,0} + b_{u2,1}t_j) \\ \sigma^2_{eu3,j} &= \exp(b_{u3,0}) \end{aligned}$$

- structure for permanent effects

trait1: antedependence SAD 1-21 (0.6 -0.1 0.01, 2 -0.1)
trait2: antedependence SAD1-21 (0.66 -0.11 0.11, 1 -0.11)
trait3:antedependence: SAD1-11 (0.7 -0.2, 2 -0.2)

crossantedependence 1->2: SAD1-0, starting point:0 (0.8)
crossantedependence 2->1: -

crossantedependence 1->3: SAD1-0, starting point:1, initial correlation between error terms (0.3, 0.5)
crossantedependence 3->1: SAD1-0, starting point:1 (0.2)
crossantedependence 2->3: SAD1-1, starting point:0 (0.5 -0.2)
crossantedependence 3->2: SAD1-1, starting point:0 (0.6 -0.1)

total: 22 parameters

$$p_1(t_j) = (a_{p1,0} + a_{p1,1}t_j + a_{p1,1}t_j^2)\,p_1(t_{j-1}) + (\omega_{p1,0})\,p_3(t_{j-1}) + e_{p1}(t_j)$$
$$p_2(t_j) = (a_{p2,0} + a_{p2,1}t_j + a_{p2,1}t_j^2)\,p_2(t_{j-1}) + (\omega_{p2,0})\,p_1(t_j) + (\omega'_{p2,0} + \omega'_{p2,1}t_j)\,p_3(t_j) + e_{p2}(t_j)$$
$$p_3(t_j) = (a_{p3,0} + a_{p3,1}t_j)\,p_3(t_{j-1}) + (\omega_{p3,0})\,p_1(t_{j-1}) + (\omega'_{p3,0} + \omega'_{p3,1}t_j)\,p_2(t_j) + e_{p3}(t_j)$$
$$\sigma^2_{ep1,j} = \exp(b_{p1,0} + b_{p1,1}t_j)$$
$$\sigma^2_{ep2,j} = \exp(b_{p2,0} + b_{p2,1}t_j)$$
$$\sigma^2_{ep3,j} = \exp(b_{p3,0} + b_{p3,1}t_j)$$
$$corr(e_{p1}(t_0), e_{p3}(t_0)) = \text{constant}$$

- *file.as, the model*

(in blue: antedependence parameters, red:cross antedependence parameters, green: innovation variance parameters, black:initial correlation)

```
Phenotype ~ fixed effects !r time2.animal time2.ide(animal)
1 1 2
0 0 IDV 0.00001 !GF !S2==1  #to fix residual variance to a very low value : 0.00001
time2.animal 2
30 time2 OWN13 0.6 -0.1 0.66 -0.11 0.7 -0.1 0.5 0.8 -0.2 2 1 -0.1 2.1 !TCCCCCCCCCCCCC
Animal
time2.ide(animal) 2
30 time2 OWN22 0.6 -0.1 0.01 0.66 -0.11 0.11 0.7 -0.2 0.8 0.3 0.2 0.5 -0.2 0.6 -0.1 2 -0.1 1 -0.11 2 -0.2
0.5 !TCCCCCCCCCCCC !F2
ide(animal)
```

- *Para (for the genetic effects)*

```
NBCHARACTER
3
MODEL
SAD
ANTEDEP_ORDER
1 1 2
CROSS_ANTEDEP_ORDER
0 0 1 0 0 0
BEGIN_CROSS_ANTEDEP_ORDER
0 0 0 0 0 0
MODEL_ANTEDEP
1 1 1 0
MODEL_CROSS_ANTEDEP
1
```

MODEL_INNOVATION_VARIANCE_COVARIANCE
0 1 0
CORRINIT
0 0 0
DATA_FILE #same as in .as
data
COLUMN_FACTOR
1

- *para2 (for the permanent effects)*

NBCHARACTER
3
MODEL
SAD
ANTEDEP_ORDER
1 1 1
CROSS_ANTEDEP_ORDER
1 0 2 2 1 1
BEGIN_CROSS_ANTEDEP_ORDER
0 0 1 1 0 0
MODEL_ANTEDEP
2 2 1
MODEL_CROSS_ANTEDEP
0 0 0 1 1
MODEL_INNOVATION_VARIANCE_COVARIANCE
1 1 1
CORRINIT
0 1 0
DATA_FILE #same as in .as
data
COLUMN_FACTOR
1