Setting-up and solving a simple (and not so simple) RRM with MiX99

Timo Pitkänen

MiX99 course: test-day models and single step genomic prediction

COURSE DAY 1, April 10th, 2025

Contents

Introduction to MiX99 and test-day models

- Simple 1 trait test-day model:
  - CLIM, data structure, parameter files, pedigree, covariable file, residual classes

Multitrait test-day models

- 1 lactation 3 trait test-day model:
  - Ordering of random effects and variance components.
  - Residual covariance between observations.
- Full 3 lactations 3 traits test-day model using trait groups:
  - Trait groups.
  - DATASORT, WITHINBLOCKORDER, sorting data and pedigree

Simple 1 trait test-day model

© Luke

- Let’s consider test-day model for 1 lactation milk animal *a*, dim *d*, herd test-day *h*

Fixed Lactation curve
3rd order Legendre
polynomial

Random
HTD

Random PE
2nd order +
Wilmink

Breeding value
2nd order +
Wilmink

$$milk_{adh} = \sum_{i=1}^{4} \boldsymbol{l_i(d)\beta_i} + \boldsymbol{HTD_h} + \sum_{i=1}^{4} \boldsymbol{c_i(d)pe_{ai}} + \sum_{i=1}^{4} \boldsymbol{c_i(d)u_{ai}} + \boldsymbol{\epsilon_{adh}}$$

Residual

Variance components

```
HTD
 0.413

Residual
 1.104
```

```
PE
 3.028 0.033  0.088 -2.249
 0.033 1.796  0.197  1.837
 0.088 0.197  1.057 -2.521
-2.249 1.837 -2.521 17.058
```

```
GE
 2.552 0.161 -0.227 -0.812
 0.161 0.958  0.084  0.526
-0.227 0.084  0.200  -0.120
-0.812 0.526 -0.120  2.390
```

© Luke

CLIM for the model

```
TITLE   MT random regression TDM

DATAFILE  lact1.dat
INTEGER   HERD ANI LACT HTD HC2Y C2Y AGE DCC DDRY DIM ResClass
REAL      MILK PROTEIN FAT
MISSING   -64.0

PEDFILE   TD9.ped

PEDIGREE G  am

TABLEFILE LG3_W004.cov
TABLEINDEX DIM

PARFILE   milk1.par

TMPDIR  tmpMiX

MODEL

MILK    =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)

RANDOM   HTD  PE G

#        1    2  3
```

```
# Title for the model. For the reader of CLIM file.

# DATAFILE location.
# Integer columns must be first, then REAL columns.
# Only first 32 characters of the names are used!!
# Missing value is -64. Data file can contain extra columns.

# Genetic effect is named G. Animal model pedigree without
# unknown parent groups is used.
# UPG can be specified as am+p, and random am+p 0.333

# Values for Legendre polynomials and wilmk for each test-day are
# stored to separate file.
# Integer variable DIM is a key between data and tablefile.

# PARFILE location.

# TMPDIR defines the directory for temporary files.

# Model definition.

# Model has fixed lactation curve and
# 4 random effects: 1=HTD ,2=PE, 3=G, and 4= Residual
```

Data, covariable, and parameter table files

© Luke

```
head lact1.dat

2669 457 1 128191 266903 3 7 1 5 20 1 15.7575 11.8857 9.8591
2669 457 1 128192 266903 3 7 1 5 48 3 20.0062 15.5429 11.3455
2669 457 1 128193 266903 3 7 1 5 75 5 17.2425 12.6143 9.2864
2669 457 1 128194 266903 3 7 1 5 97 6 16.7475 12.6286 10.4727
2669 457 1 128195 266903 3 7 1 5 123 7 16.1287 12.5286 9.6136
2669 457 1 128196 266903 3 7 1 5 158 7 13.1175 10.2 8.2091
2669 457 1 128197 266903 3 7 1 5 181 8 14.4994 11.9286 9.6955
2669 457 1 128198 266903 3 7 1 5 222 9 15.51 13.1143 9.2455
2669 457 1 128199 266903 3 7 1 5 250 9 14.3756 12.1571 9.4091
2669 457 1 128200 266903 3 7 1 5 291 10 12.4987 10.5714 9.6409

head LG3_W004.cov
#DIM      t1         t2         t3          t4         t5
   8    1.00000   -1.22474    1.58114   -1.87083    0.72615
   9    1.00000   -1.21788    1.55464   -1.80838    0.69768
  10    1.00000   -1.21102    1.52829   -1.74681    0.67032
  11    1.00000   -1.20416    1.50209   -1.68611    0.64404
  12    1.00000   -1.19730    1.47603   -1.62628    0.61878
  13    1.00000   -1.19044    1.45013   -1.56731    0.59452
  14    1.00000   -1.18358    1.42438   -1.50919    0.57121
  15    1.00000   -1.17672    1.39877   -1.45193    0.54881
  16    1.00000   -1.16985    1.37331   -1.39551    0.52729
  17    1.00000   -1.16299    1.34800   -1.33993    0.50662
```

Parfile format is:

Effect number, i, j, val.

Only lower or upper part of the covariance matrix is needed.

First column in covariable table is the key column.

```
cat milk1.par
# HTD
1 1 1 0.41327678
# PE
2 1 1 3.0284761
2 1 2 0.032571576
2 2 2 1.7962894
2 1 3 0.088109073
2 2 3 0.19685252
2 3 3 1.0572924
2 1 4 -2.2488402
2 2 4 1.8372373
2 3 4 -2.5214496
2 4 4 17.057958
# GEN
3 1 1 2.5521561
3 1 2 0.16126933
3 2 2 0.95818471
3 1 3 -0.22728278
3 2 3 0.083704976
3 3 3 0.20033431
3 1 4 -0.81232366
3 2 4 0.5259268
3 3 4 -0.11959052
3 4 4 2.3896462
#RESID
4 1 1 1.1035821
```

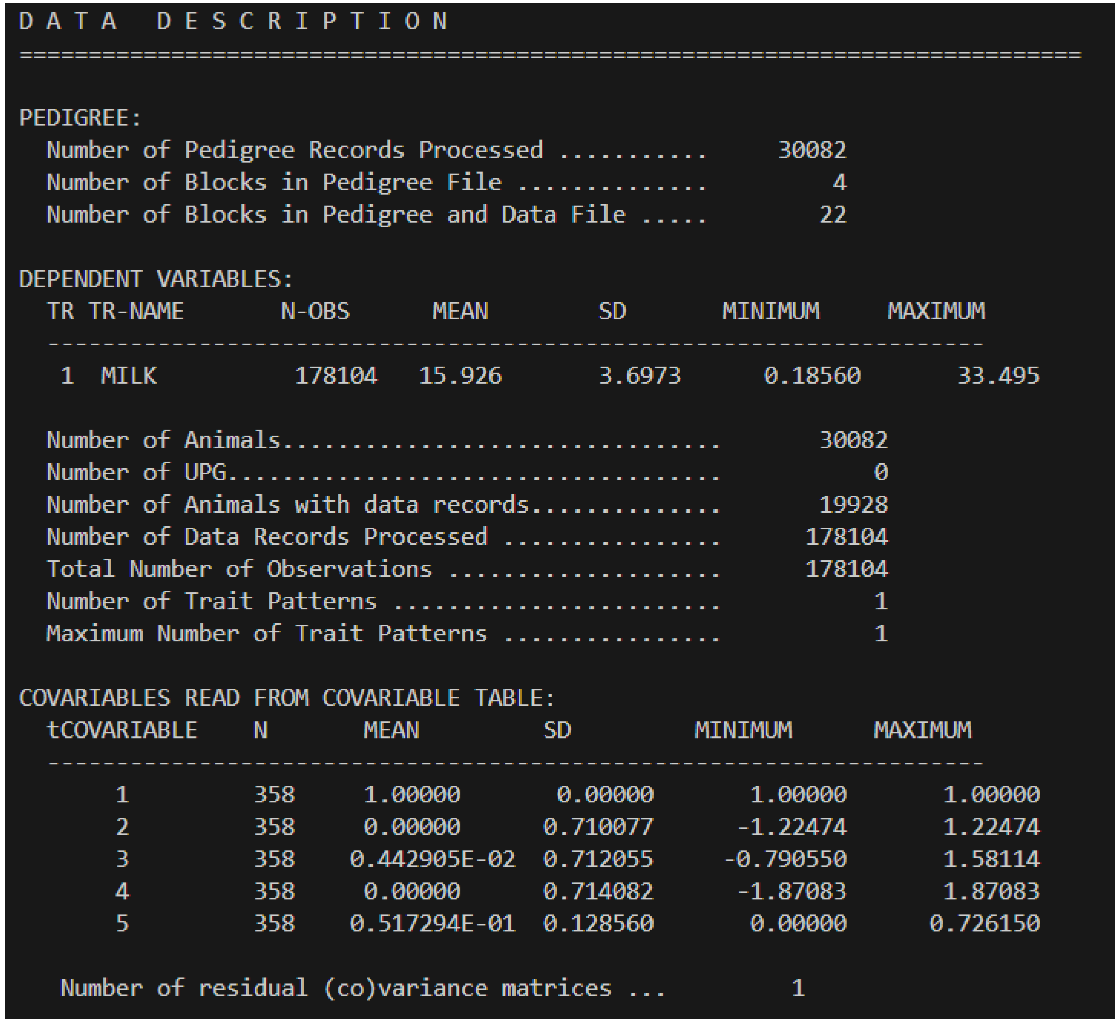
© Luke

Running preprosessor

- Run preprosessor

```
$ mix99i milk1.clm
```

- Mix99i prints model information to `MiX99.lst` file and standard output.
- After successful execution, `OK_mix99i` file is created.
- It can be used in pipelines to check for errors.
- It contains timing information.
- **ALWAYS check** at least DATA DESCRIPTION from `Mix99.lst`!!
- Also check the standard output!



```
D A T A   D E S C R I P T I O N
==============================================================================

PEDIGREE:
  Number of Pedigree Records Processed ...........      30082
  Number of Blocks in Pedigree File ...............          4
  Number of Blocks in Pedigree and Data File .....          22

DEPENDENT VARIABLES:
  TR TR-NAME       N-OBS      MEAN        SD       MINIMUM     MAXIMUM
  ----------------------------------------------------------------------
   1  MILK          178104    15.926       3.6973      0.18560       33.495

  Number of Animals.................................        30082
  Number of UPG.....................................            0
  Number of Animals with data records...............        19928
  Number of Data Records Processed .................       178104
  Total Number of Observations .....................       178104
  Number of Trait Patterns .........................            1
  Maximum Number of Trait Patterns .................            1

COVARIABLES READ FROM COVARIABLE TABLE:
  tCOVARIABLE     N      MEAN          SD         MINIMUM      MAXIMUM
  ----------------------------------------------------------------------
       1         358     1.00000       0.00000       1.00000       1.00000
       2         358     0.00000      0.710077      -1.22474       1.22474
       3         358     0.442905E-02  0.712055     -0.790550       1.58114
       4         358     0.00000      0.714082      -1.87083       1.87083
       5         358     0.517294E-01  0.128560       0.00000      0.726150

   Number of residual (co)variance matrices ...         1
```

© Luke

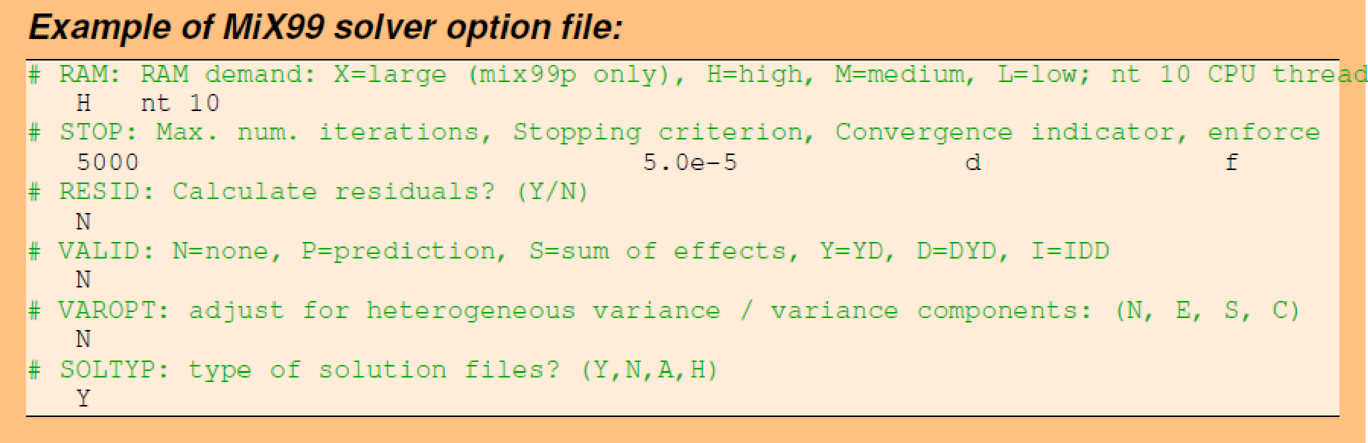
Running solver

- Run solver with default settings and produce default solution files

```
$ mix99s -s
```

Or give own settings:

```
$ mix99s < solver.slv
```



Example of MiX99 solver option file:

```
# RAM: RAM demand: X=large (mix99p only), H=high, M=medium, L=low; nt 10 CPU thread
   H   nt 10
# STOP: Max. num. iterations, Stopping criterion, Convergence indicator, enforce
   5000                                 5.0e-5                 d                f
# RESID: Calculate residuals? (Y/N)
   N
# VALID: N=none, P=prediction, S=sum of effects, Y=YD, D=DYD, I=IDD
   N
# VAROPT: adjust for heterogeneous variance / variance components: (N, E, S, C)
   N
# SOLTYP: type of solution files? (Y,N,A,H)
   Y
```

- After successful execution, `OK_mix99s` file is created
- It contains timing and other technical information.
- The other information is written to the standard output.
- Iteration statistics are also written to `Conlog`

© Luke

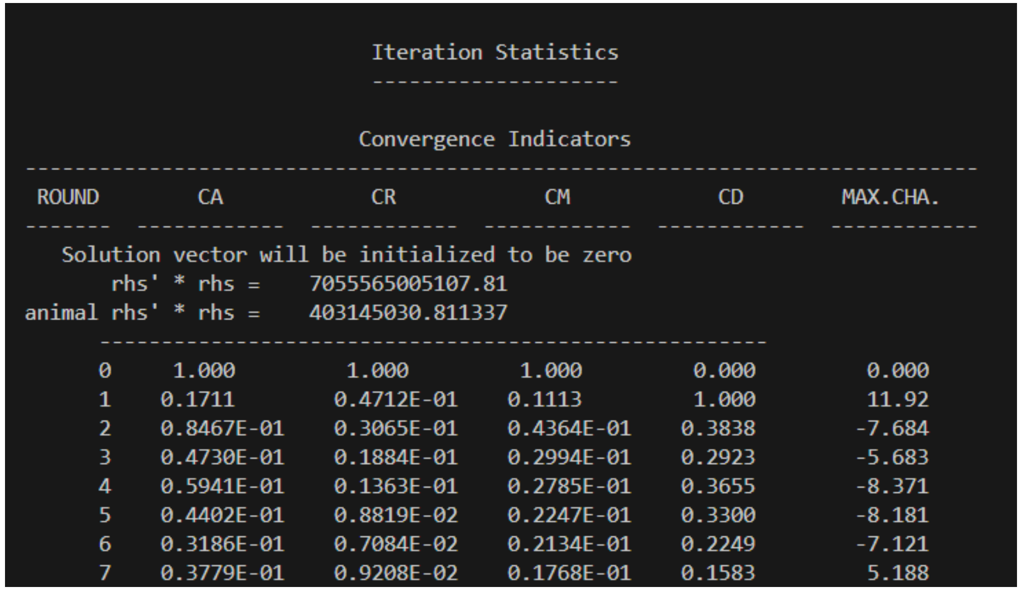
Output of the solver

- Solver output has the following parts
  - version information,
  - preconditioner information,
  - solver parameters,
  - iteration statistics,
  - solutions,
  - description of solution files.

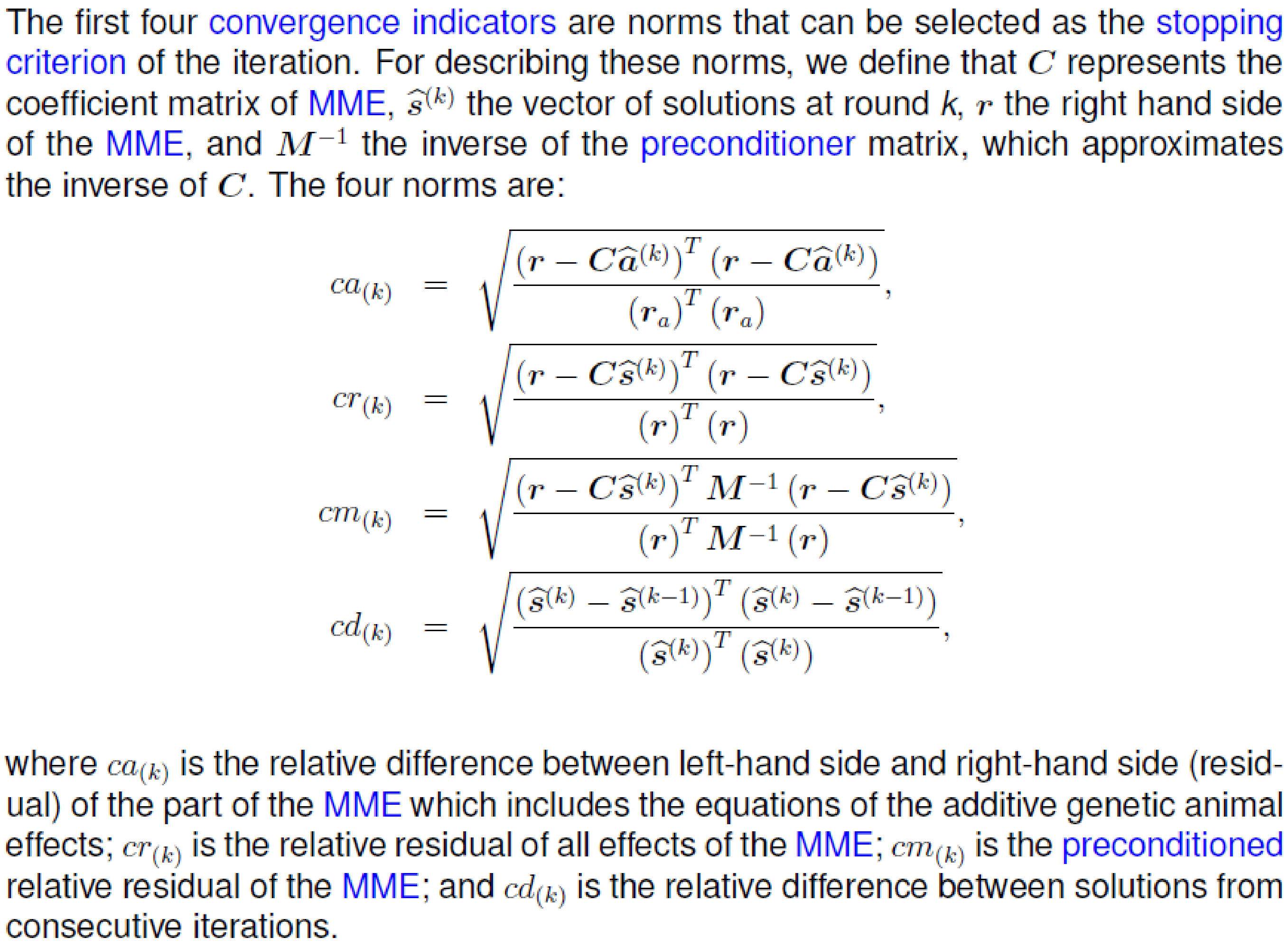
© Luke

Convergence

PCG method solves a system of equations: $\mathbf{M}^{-1}\mathbf{C}\mathbf{s} = \mathbf{M}^{-1}\mathbf{r}$,



```
                         Iteration Statistics
                         --------------------

                        Convergence Indicators
 ------------------------------------------------------------------------------
  ROUND        CA            CR            CM            CD          MAX.CHA.
 -------  ------------  ------------  ------------  ------------  ------------
    Solution vector will be initialized to be zero
        rhs' * rhs =     7055565005107.81
 animal rhs' * rhs =     403145030.811337
      ----------------------------------------------------------
        0     1.000         1.000         1.000         0.000         0.000
        1    0.1711        0.4712E-01    0.1113         1.000         11.92
        2    0.8467E-01    0.3065E-01    0.4364E-01    0.3838        -7.684
        3    0.4730E-01    0.1884E-01    0.2994E-01    0.2923        -5.683
        4    0.5941E-01    0.1363E-01    0.2785E-01    0.3655        -8.371
        5    0.4402E-01    0.8819E-02    0.2247E-01    0.3300        -8.181
        6    0.3186E-01    0.7084E-02    0.2134E-01    0.2249        -7.121
        7    0.3779E-01    0.9208E-02    0.1768E-01    0.1583         5.188
```



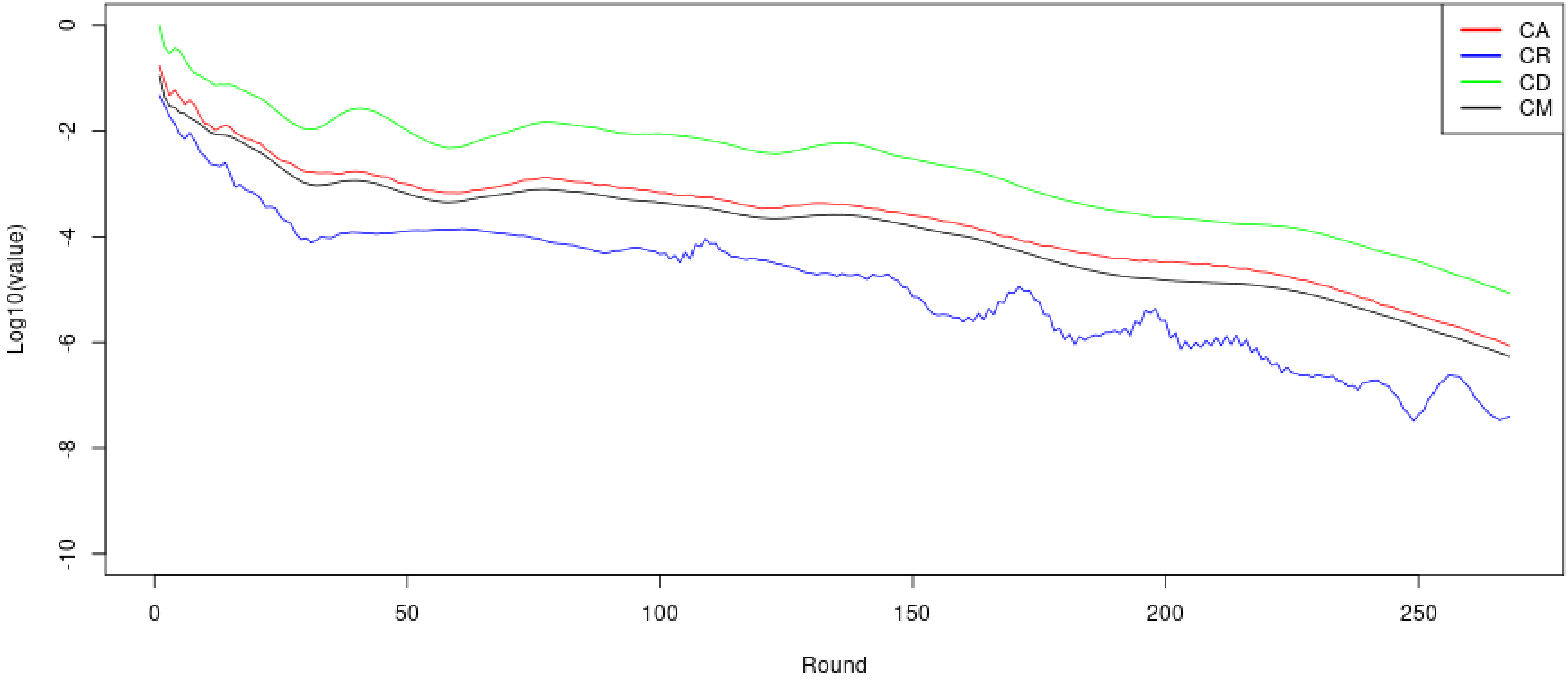
The first four convergence indicators are norms that can be selected as the stopping criterion of the iteration. For describing these norms, we define that C represents the coefficient matrix of MME, $\widehat{s}^{(k)}$ the vector of solutions at round k, r the right hand side of the MME, and M^{-1} the inverse of the preconditioner matrix, which approximates the inverse of C. The four norms are:

$$
\begin{aligned}
ca_{(k)} &= \sqrt{\frac{\left(r - C\widehat{a}^{(k)}\right)^{T}\left(r - C\widehat{a}^{(k)}\right)}{\left(r_a\right)^{T}\left(r_a\right)}},\\
cr_{(k)} &= \sqrt{\frac{\left(r - C\widehat{s}^{(k)}\right)^{T}\left(r - C\widehat{s}^{(k)}\right)}{\left(r\right)^{T}\left(r\right)}},\\
cm_{(k)} &= \sqrt{\frac{\left(r - C\widehat{s}^{(k)}\right)^{T} M^{-1}\left(r - C\widehat{s}^{(k)}\right)}{\left(r\right)^{T} M^{-1}\left(r\right)}},\\
cd_{(k)} &= \sqrt{\frac{\left(\widehat{s}^{(k)} - \widehat{s}^{(k-1)}\right)^{T}\left(\widehat{s}^{(k)} - \widehat{s}^{(k-1)}\right)}{\left(\widehat{s}^{(k)}\right)^{T}\left(\widehat{s}^{(k)}\right)}},
\end{aligned}
$$

where $ca_{(k)}$ is the relative difference between left-hand side and right-hand side (residual) of the part of the MME which includes the equations of the additive genetic animal effects; $cr_{(k)}$ is the relative residual of all effects of the MME; $cm_{(k)}$ is the preconditioned relative residual of the MME; and $cd_{(k)}$ is the relative difference between solutions from consecutive iterations.

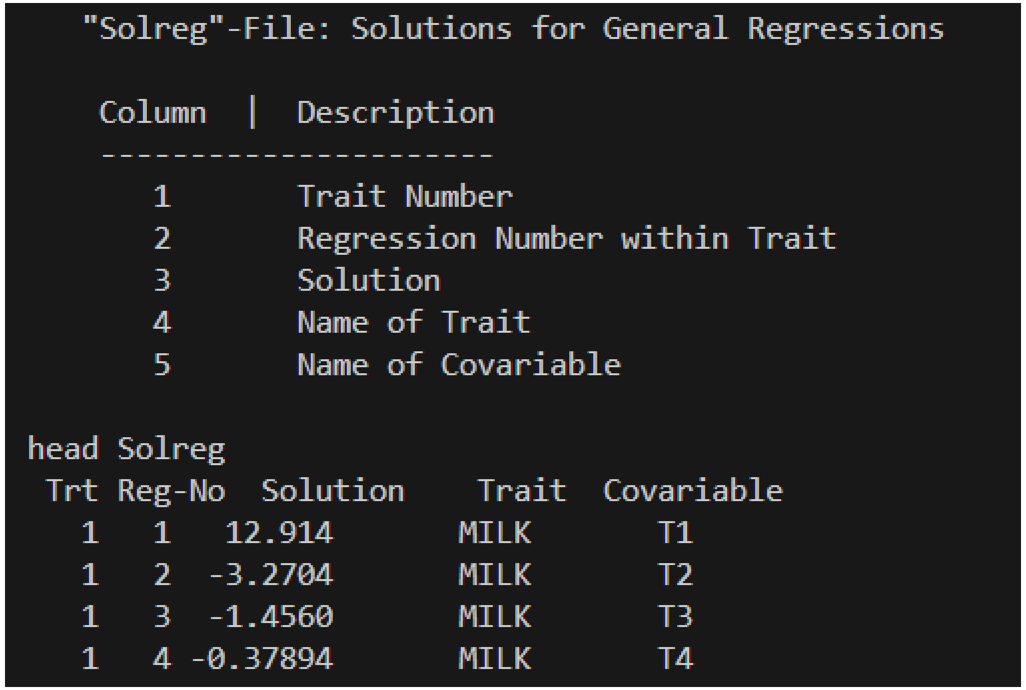
© Luke

Convergence indicators

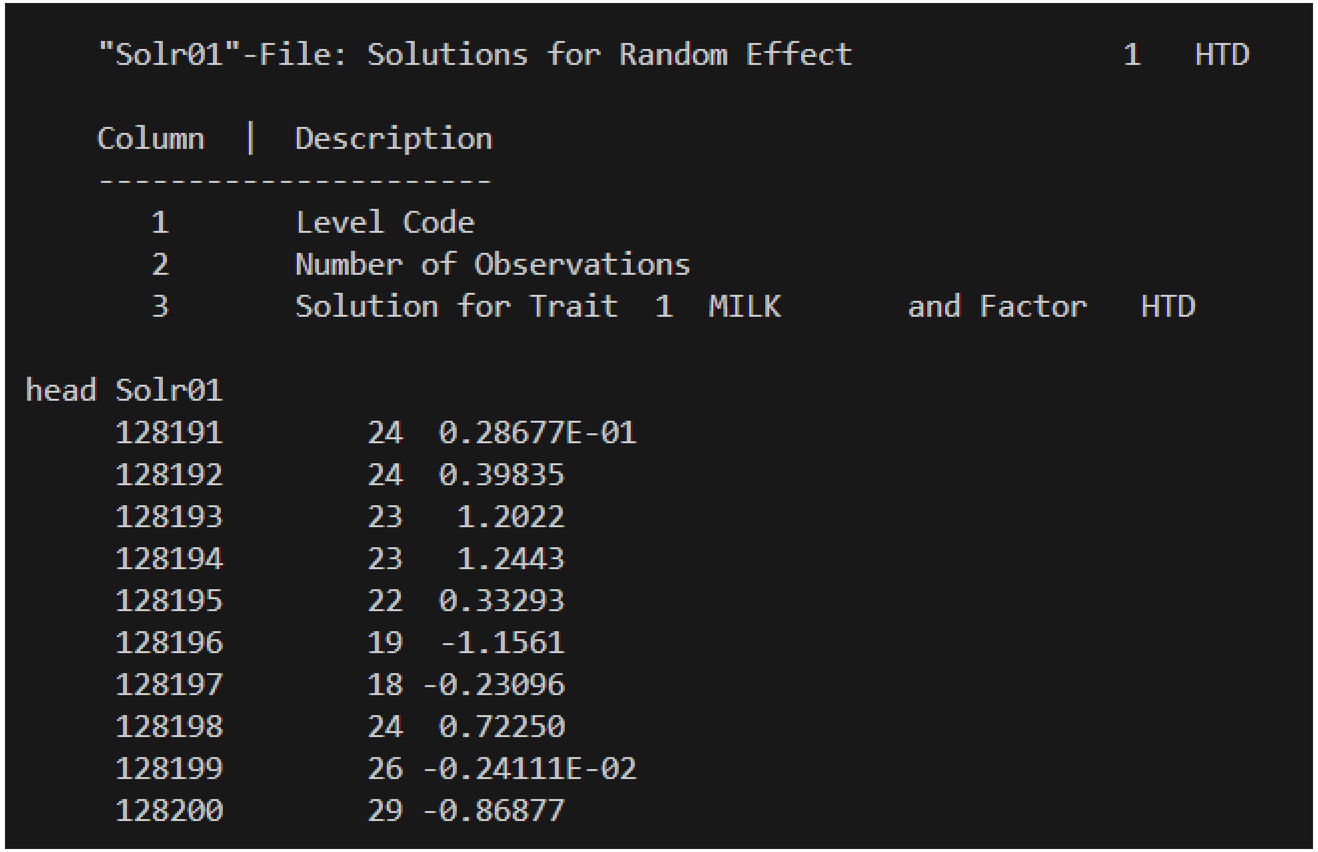


© Luke

Solution files: Solreg(Lactcurve) and Solr01 (HTD)



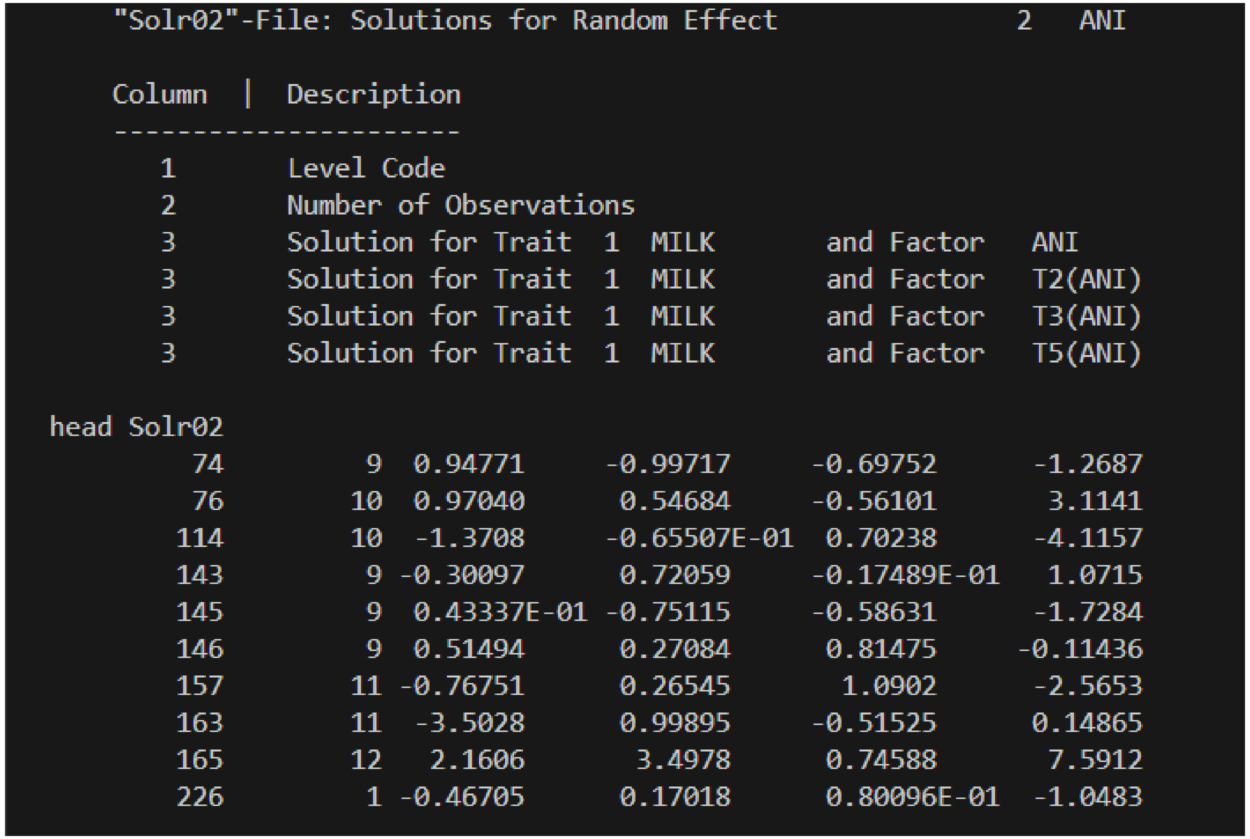
```
    "Solreg"-File: Solutions for General Regressions

     Column  |  Description
     ----------------------
        1       Trait Number
        2       Regression Number within Trait
        3       Solution
        4       Name of Trait
        5       Name of Covariable

head Solreg
 Trt Reg-No  Solution    Trait  Covariable
   1   1   12.914       MILK       T1
   1   2  -3.2704       MILK       T2
   1   3  -1.4560       MILK       T3
   1   4 -0.37894       MILK       T4
```



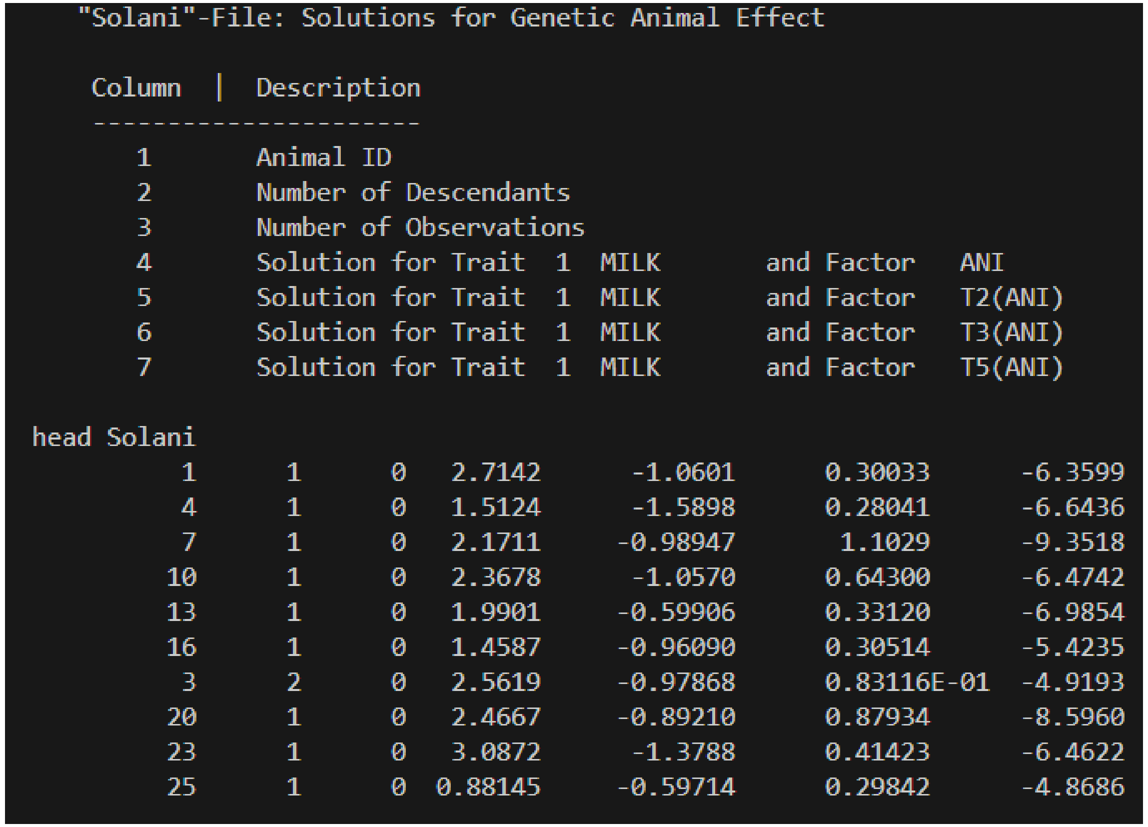
```
    "Solr01"-File: Solutions for Random Effect             1   HTD

     Column  |  Description
     ----------------------
        1       Level Code
        2       Number of Observations
        3       Solution for Trait  1  MILK       and Factor   HTD

head Solr01
     128191        24  0.28677E-01
     128192        24  0.39835
     128193        23   1.2022
     128194        23   1.2443
     128195        22  0.33293
     128196        19  -1.1561
     128197        18 -0.23096
     128198        24  0.72250
     128199        26 -0.24111E-02
     128200        29 -0.86877
```

© Luke

Solution files Solr02 (PE) and Solani (EBV)



```
     "Solr02"-File: Solutions for Random Effect            2   ANI

     Column  |  Description
     ----------------------
        1       Level Code
        2       Number of Observations
        3       Solution for Trait  1  MILK       and Factor   ANI
        3       Solution for Trait  1  MILK       and Factor   T2(ANI)
        3       Solution for Trait  1  MILK       and Factor   T3(ANI)
        3       Solution for Trait  1  MILK       and Factor   T5(ANI)

 head Solr02
         74        9  0.94771     -0.99717     -0.69752      -1.2687
         76       10  0.97040      0.54684     -0.56101       3.1141
        114       10  -1.3708     -0.65507E-01  0.70238      -4.1157
        143        9 -0.30097      0.72059     -0.17489E-01   1.0715
        145        9  0.43337E-01 -0.75115     -0.58631      -1.7284
        146        9  0.51494      0.27084      0.81475     -0.11436
        157       11 -0.76751      0.26545       1.0902      -2.5653
        163       11  -3.5028      0.99895     -0.51525      0.14865
        165       12   2.1606       3.4978      0.74588       7.5912
        226        1 -0.46705      0.17018      0.80096E-01  -1.0483
```



```
    "Solani"-File: Solutions for Genetic Animal Effect

     Column  |  Description
     ----------------------
        1       Animal ID
        2       Number of Descendants
        3       Number of Observations
        4       Solution for Trait  1  MILK       and Factor   ANI
        5       Solution for Trait  1  MILK       and Factor   T2(ANI)
        6       Solution for Trait  1  MILK       and Factor   T3(ANI)
        7       Solution for Trait  1  MILK       and Factor   T5(ANI)

 head Solani
         1      1      0   2.7142      -1.0601      0.30033      -6.3599
         4      1      0   1.5124      -1.5898      0.28041      -6.6436
         7      1      0   2.1711     -0.98947       1.1029      -9.3518
        10      1      0   2.3678      -1.0570      0.64300      -6.4742
        13      1      0   1.9901     -0.59906      0.33120      -6.9854
        16      1      0   1.4587     -0.96090      0.30514      -5.4235
         3      2      0   2.5619     -0.97868      0.83116E-01  -4.9193
        20      1      0   2.4667     -0.89210      0.87934      -8.5960
        23      1      0   3.0872      -1.3788      0.41423      -6.4622
        25      1      0  0.88145     -0.59714      0.29842      -4.8686
```

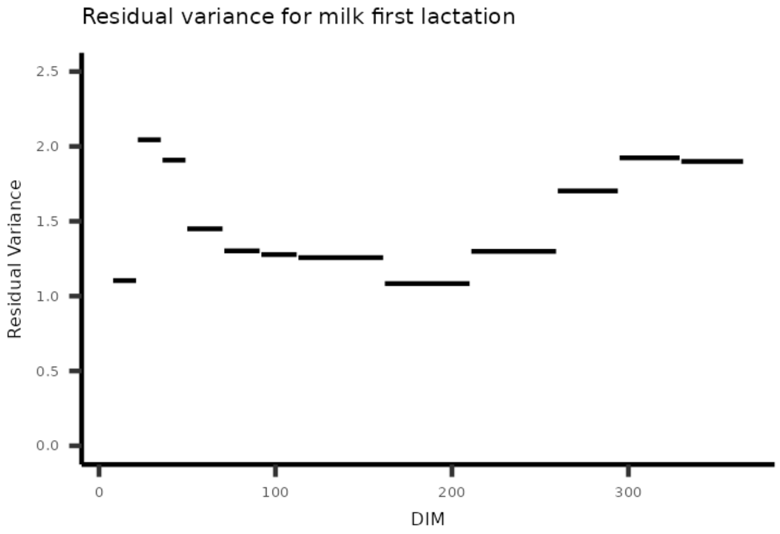
© Luke

Add residual variance classes to model

```
TITLE   MT random regression TDM
DATAFILE  lact1.dat
INTEGER   HERD ANI LACT HTD HC2Y C2Y AGE DCC DDRY DIM ResClass
REAL      MILK PROTEIN FAT
MISSING   -64.0


PEDFILE   TD9.ped
PEDIGREE G  am

TABLEFILE LG3_W004.cov
TABLEINDEX DIM

RESIDUAL ResClass
RESIDFILE milk1.res

PARFILE   milk1.par

RANDOM   HTD  PE G

TMPDIR  tmpMiX

MODEL

MILK    =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
```

- Integer column **ResClass** specifies which residual class observation belongs to
- Residual variances for classes are given in own parameter file (milk1.res) with columns:
  class number, i,j, value

```
cat milk1.res
1  1 1 1.1035821
2  1 1 2.044039
3  1 1 1.9081354
4  1 1 1.4494774
5  1 1 1.3019862
6  1 1 1.2774864
7  1 1 1.2569414
8  1 1 1.0833885
9  1 1 1.2988478
10 1 1 1.7024491
11 1 1 1.9234934
12 1 1 1.8995175
```



© Luke

Multiple traits random regression

```
TITLE   MT random regression TDM
DATAFILE  lact1.dat
INTEGER   HERD ANI LACT HTD HC2Y C2Y AGE DCC DDRY DIM ResClass
REAL      MILK PROTEIN FAT
MISSING   -64.0


PEDFILE   TD9.ped
PEDIGREE G  am

TABLEFILE LG3_W004.cov
TABLEINDEX DIM

RESIDUAL ResClass
RESIDFILE lactation1.res

PARFILE   lactation1.par

RANDOM   HTD  PE G

TMPDIR  tmpMiX

MODEL

MILK    =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
PROTEIN =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
FAT     =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
```

- Model rows for protein and fat are just copy paste from milk.
- The model effects for each trait are the same.
- Sizes of variance component matrices are

  HTD = 3x3, PE=12x12, G=12x12.
- The tricky part is the structure of covariance matrices for PE and G.

© Luke

Covariance matrices for 1 first lactation model: HTD

- HTD effect has only three random variables for each herd test-day:

$$h_{htd} = \begin{bmatrix} h_{htd;milk} \\ h_{htd;prot} \\ h_{htd;fat} \end{bmatrix},$$

$$var(h_{htd}) = \begin{bmatrix} 0.41 & 0.31 & 0.20 \\ 0.31 & 0.30 & 0.18 \\ 0.20 & 0.18 & 0.26 \end{bmatrix}$$

```
MILK    =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
PROTEIN =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
FAT     =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
```

© Luke

Covariance matrices for 1 first lactation model: PE and G

- PE and G have 4 random regression coefficient for each trait.
- MiX99 numbers random variables first by row, then by column.
- Variable numbers are in subscript.
- Ordering of random variables is the same for PE and G.
- Both effects require 78 covariance parameters!

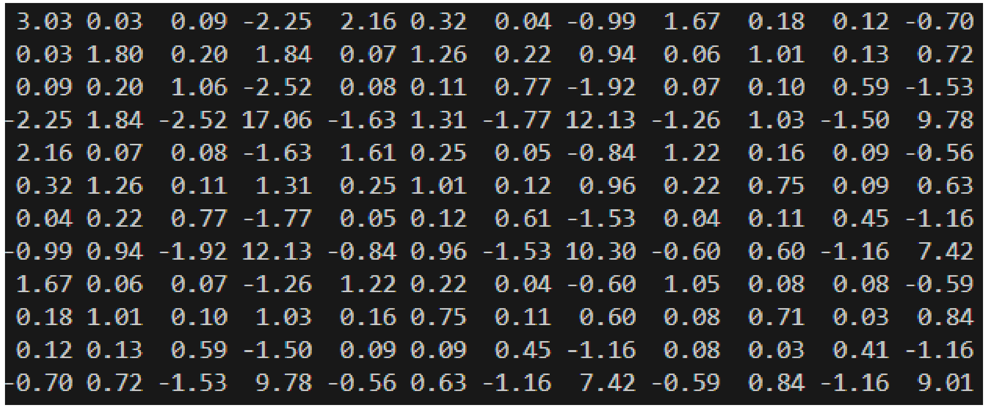
$$
pe_i = \begin{bmatrix} p1_i; milk \\ p1_i; prot \\ p1_i; fat \\ pt2_i; milk \\ pt2_i; prot \\ pt2_i; fat \\ pt3_i; milk \\ pt3_i; prot \\ pt3_i; fat \\ pt5_i; milk \\ pt5_i; prot \\ pt5_i; fat \end{bmatrix}
$$

$$
\begin{array}{lllll}
\texttt{\textbf{MILK}} & \texttt{\textbf{=}} & \texttt{LACcurve(t1 t2 t3 t4) HTD} & \texttt{\textbf{PE(}}\mathbf{1_{1}\ t2_{4}\ t3_{7}\ t5_{10}}\texttt{\textbf{ |ANI)}} & \texttt{\textbf{G(}}\mathbf{1_{1}\ t2_{4}\ t3_{7}\ t5_{10}}\texttt{\textbf{ |ANI)}} \\
\texttt{\textbf{PROTEIN}} & \texttt{\textbf{=}} & \texttt{LACcurve(t1 t2 t3 t4) HTD} & \texttt{\textbf{PE(}}\mathbf{1_{2}\ t2_{5}\ t3_{8}\ t5_{11}}\texttt{\textbf{ |ANI)}} & \texttt{\textbf{G(}}\mathbf{1_{2}\ t2_{5}\ t3_{8}\ t5_{11}}\texttt{\textbf{ |ANI)}} \\
\texttt{\textbf{FAT}} & \texttt{\textbf{=}} & \texttt{LACcurve(t1 t2 t3 t4) HTD} & \texttt{\textbf{PE(}}\mathbf{1_{3}\ t2_{6}\ t3_{9}\ t5_{12}}\texttt{\textbf{ |ANI)}} & \texttt{\textbf{G(}}\mathbf{1_{3}\ t2_{6}\ t3_{9}\ t5_{12}}\texttt{\textbf{ |ANI)}}
\end{array}
$$

© Luke

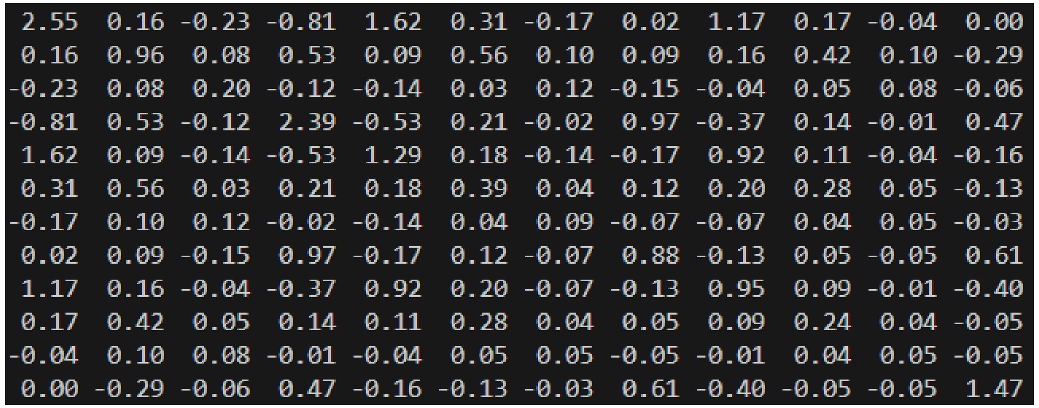
Covariance matrices for 1 first lactation model: PE and G

- PE covariance matrix



```
 3.03 0.03  0.09 -2.25  2.16 0.32  0.04 -0.99  1.67  0.18  0.12 -0.70
 0.03 1.80  0.20  1.84  0.07 1.26  0.22  0.94  0.06  1.01  0.13  0.72
 0.09 0.20  1.06 -2.52  0.08 0.11  0.77 -1.92  0.07  0.10  0.59 -1.53
-2.25 1.84 -2.52 17.06 -1.63 1.31 -1.77 12.13 -1.26  1.03 -1.50  9.78
 2.16 0.07  0.08 -1.63  1.61 0.25  0.05 -0.84  1.22  0.16  0.09 -0.56
 0.32 1.26  0.11  1.31  0.25 1.01  0.12  0.96  0.22  0.75  0.09  0.63
 0.04 0.22  0.77 -1.77  0.05 0.12  0.61 -1.53  0.04  0.11  0.45 -1.16
-0.99 0.94 -1.92 12.13 -0.84 0.96 -1.53 10.30 -0.60  0.60 -1.16  7.42
 1.67 0.06  0.07 -1.26  1.22 0.22  0.04 -0.60  1.05  0.08  0.08 -0.59
 0.18 1.01  0.10  1.03  0.16 0.75  0.11  0.60  0.08  0.71  0.03  0.84
 0.12 0.13  0.59 -1.50  0.09 0.09  0.45 -1.16  0.08  0.03  0.41 -1.16
-0.70 0.72 -1.53  9.78 -0.56 0.63 -1.16  7.42 -0.59  0.84 -1.16  9.01
```

- G covariance matrix



```
 2.55  0.16 -0.23 -0.81  1.62  0.31 -0.17  0.02  1.17  0.17 -0.04  0.00
 0.16  0.96  0.08  0.53  0.09  0.56  0.10  0.09  0.16  0.42  0.10 -0.29
-0.23  0.08  0.20 -0.12 -0.14  0.03  0.12 -0.15 -0.04  0.05  0.08 -0.06
-0.81  0.53 -0.12  2.39 -0.53  0.21 -0.02  0.97 -0.37  0.14 -0.01  0.47
 1.62  0.09 -0.14 -0.53  1.29  0.18 -0.14 -0.17  0.92  0.11 -0.04 -0.16
 0.31  0.56  0.03  0.21  0.18  0.39  0.04  0.12  0.20  0.28  0.05 -0.13
-0.17  0.10  0.12 -0.02 -0.14  0.04  0.09 -0.07 -0.07  0.04  0.05 -0.03
 0.02  0.09 -0.15  0.97 -0.17  0.12 -0.07  0.88 -0.13  0.05 -0.05  0.61
 1.17  0.16 -0.04 -0.37  0.92  0.20 -0.07 -0.13  0.95  0.09 -0.01 -0.40
 0.17  0.42  0.05  0.14  0.11  0.28  0.04  0.05  0.09  0.24  0.04 -0.05
-0.04  0.10  0.08 -0.01 -0.04  0.05  0.05 -0.05 -0.01  0.04  0.05 -0.05
 0.00 -0.29 -0.06  0.47 -0.16 -0.13 -0.03  0.61 -0.40 -0.05 -0.05  1.47
```

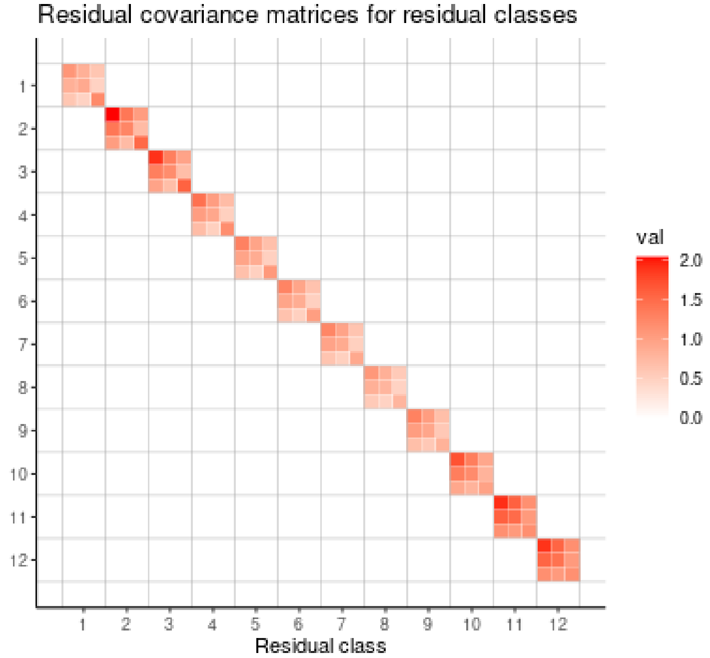
MILK = LACcurve(t1 t2 t3 t4) HTD **PE(1_1 $t2_4$ $t3_7$ $t5_{10}$ |ANI)** **G(1_1 $t2_4$ $t3_7$ $t5_{10}$ |ANI)**
PROTEIN = LACcurve(t1 t2 t3 t4) HTD **PE(1_2 $t2_5$ $t3_8$ $t5_{11}$ |ANI)** **G(1_2 $t2_5$ $t3_8$ $t5_{11}$ |ANI)**
FAT = LACcurve(t1 t2 t3 t4) HTD **PE(1_3 $t2_6$ $t3_9$ $t5_{12}$ |ANI)** **G(1_3 $t2_6$ $t3_9$ $t5_{12}$ |ANI)**

Residual covariance in multitrait models

- Residual covariance between traits is applied to observations in the same record (dataline).
- Residual covariance between observations in different data lines is assumed to be 0.

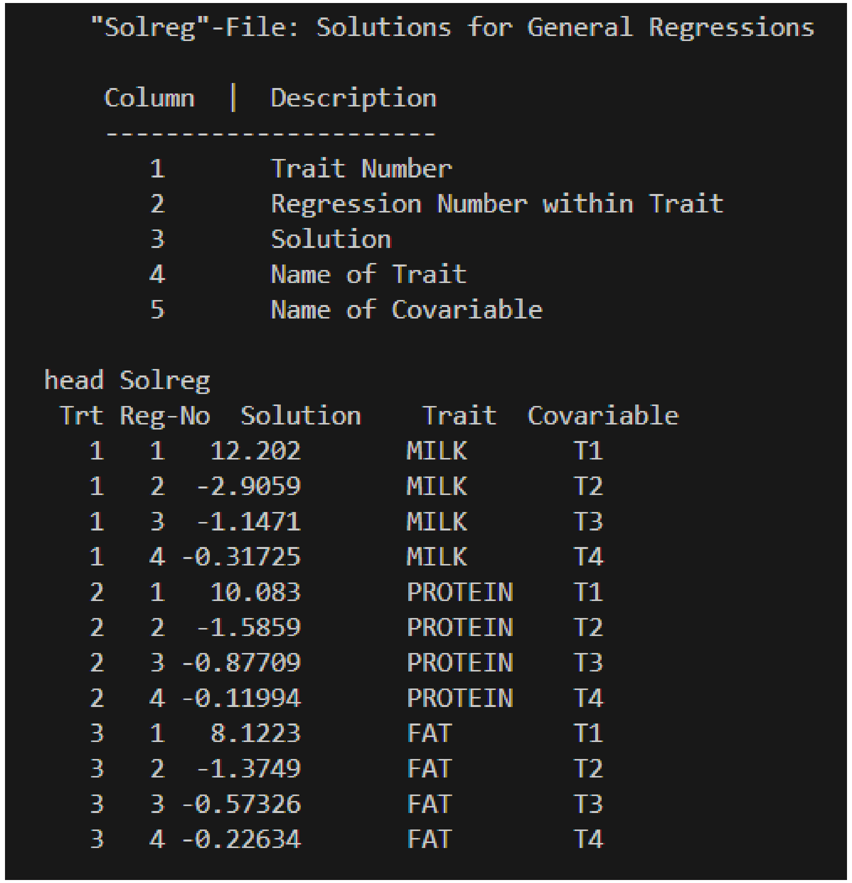
- If the dependency of residuals for different DIMs is needed, it can be modelled using PE like random animal effect.
- However, it is challenging and may cause convergence issues.

- In this model 12 residual classes are used
- Each residual class has 3x3 covariance matrix.
- In total, 12 * 6 = 72 covariance parameters for residual

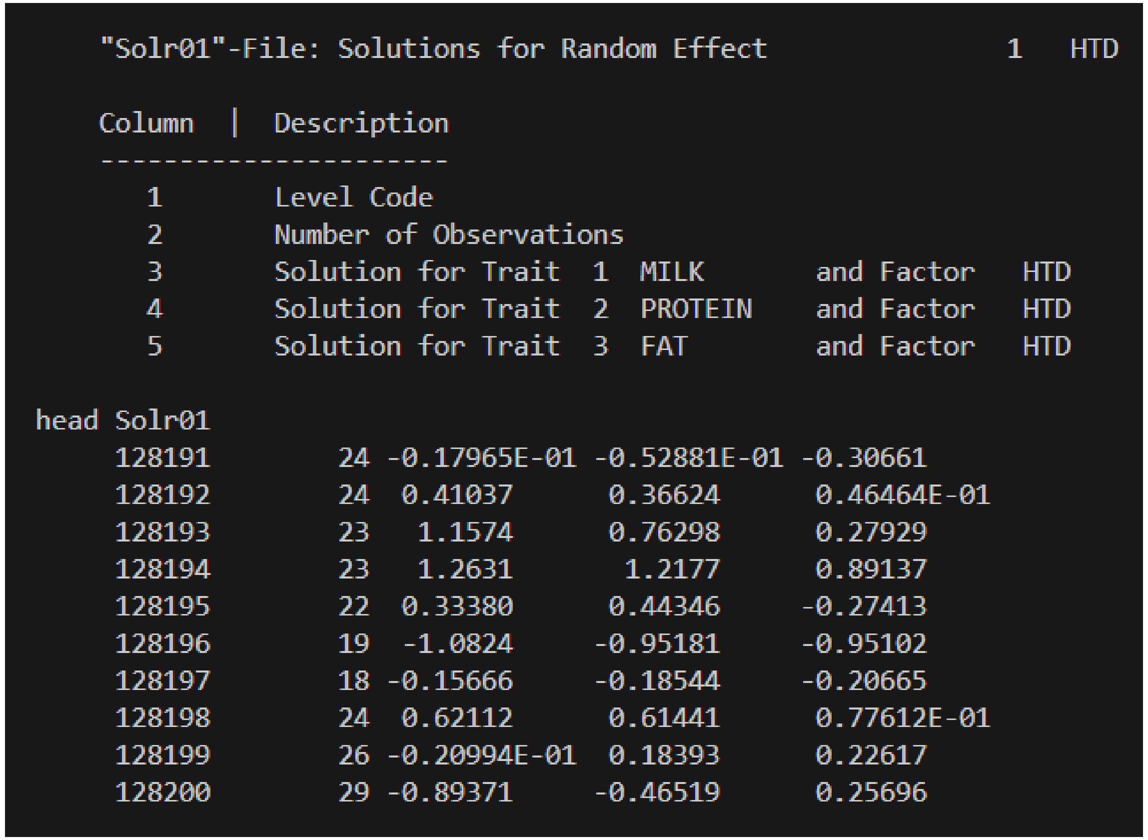


© Luke

Solution files for first lactation model

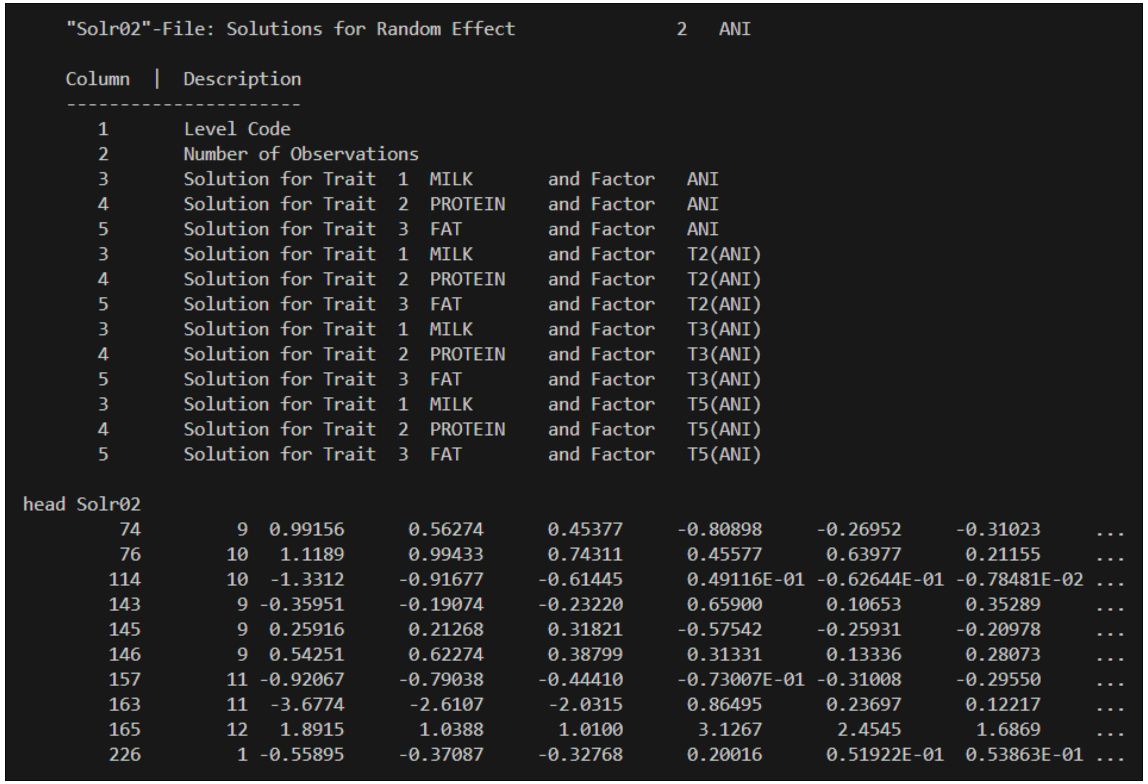


```
    "Solreg"-File: Solutions for General Regressions

     Column  |  Description
     ----------------------
        1       Trait Number
        2       Regression Number within Trait
        3       Solution
        4       Name of Trait
        5       Name of Covariable

head Solreg
 Trt Reg-No  Solution    Trait  Covariable
   1   1   12.202       MILK       T1
   1   2  -2.9059       MILK       T2
   1   3  -1.1471       MILK       T3
   1   4 -0.31725       MILK       T4
   2   1   10.083       PROTEIN    T1
   2   2  -1.5859       PROTEIN    T2
   2   3 -0.87709       PROTEIN    T3
   2   4 -0.11994       PROTEIN    T4
   3   1   8.1223       FAT        T1
   3   2  -1.3749       FAT        T2
   3   3 -0.57326       FAT        T3
   3   4 -0.22634       FAT        T4
```

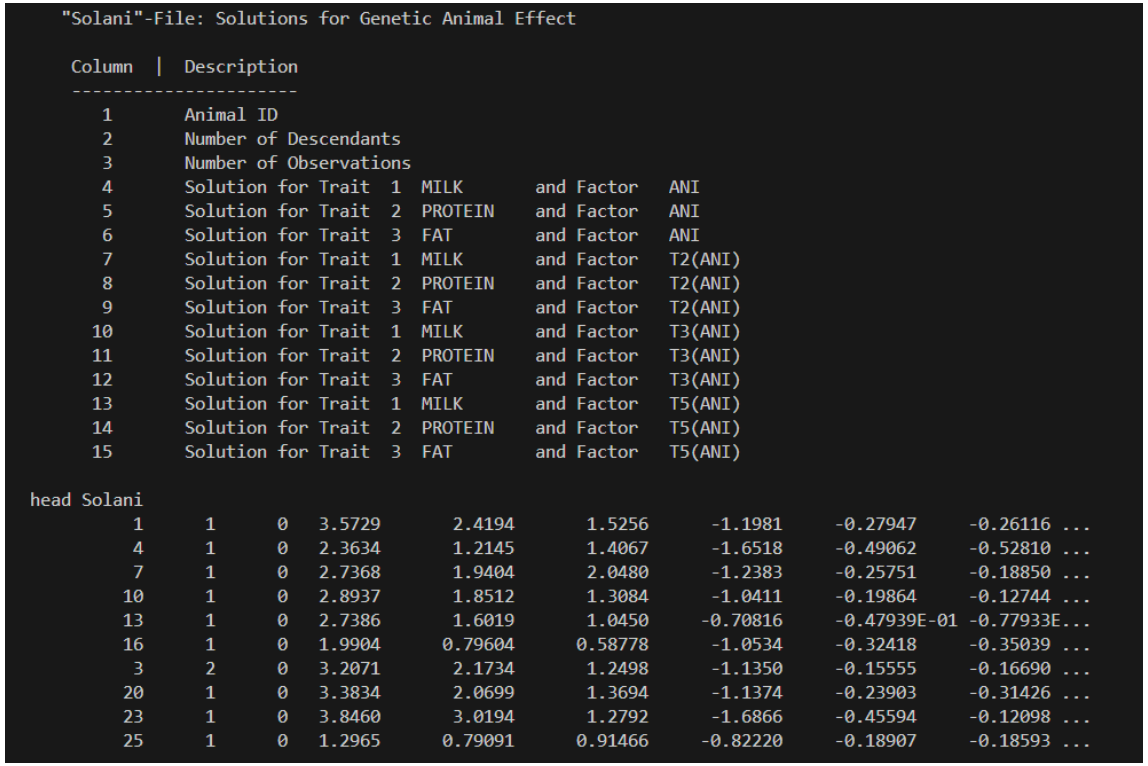


```
    "Solr01"-File: Solutions for Random Effect             1   HTD

    Column  |  Description
    ----------------------
       1       Level Code
       2       Number of Observations
       3       Solution for Trait  1  MILK       and Factor   HTD
       4       Solution for Trait  2  PROTEIN    and Factor   HTD
       5       Solution for Trait  3  FAT        and Factor   HTD

head Solr01
     128191        24 -0.17965E-01 -0.52881E-01 -0.30661
     128192        24  0.41037       0.36624      0.46464E-01
     128193        23   1.1574       0.76298      0.27929
     128194        23   1.2631        1.2177      0.89137
     128195        22  0.33380       0.44346     -0.27413
     128196        19  -1.0824      -0.95181     -0.95102
     128197        18 -0.15666      -0.18544     -0.20665
     128198        24  0.62112       0.61441      0.77612E-01
     128199        26 -0.20994E-01  0.18393      0.22617
     128200        29 -0.89371      -0.46519      0.25696
```

Solution files for first lactation model



```
    "Solr02"-File: Solutions for Random Effect             2   ANI

    Column  |  Description
    ----------------------
       1       Level Code
       2       Number of Observations
       3       Solution for Trait  1  MILK       and Factor   ANI
       4       Solution for Trait  2  PROTEIN    and Factor   ANI
       5       Solution for Trait  3  FAT        and Factor   ANI
       3       Solution for Trait  1  MILK       and Factor   T2(ANI)
       4       Solution for Trait  2  PROTEIN    and Factor   T2(ANI)
       5       Solution for Trait  3  FAT        and Factor   T2(ANI)
       3       Solution for Trait  1  MILK       and Factor   T3(ANI)
       4       Solution for Trait  2  PROTEIN    and Factor   T3(ANI)
       5       Solution for Trait  3  FAT        and Factor   T3(ANI)
       3       Solution for Trait  1  MILK       and Factor   T5(ANI)
       4       Solution for Trait  2  PROTEIN    and Factor   T5(ANI)
       5       Solution for Trait  3  FAT        and Factor   T5(ANI)

head Solr02
          74         9  0.99156       0.56274       0.45377      -0.80898      -0.26952      -0.31023      ...
          76        10   1.1189       0.99433       0.74311       0.45577       0.63977       0.21155      ...
         114        10  -1.3312      -0.91677      -0.61445       0.49116E-01 -0.62644E-01 -0.78481E-02 ...
         143         9 -0.35951      -0.19074      -0.23220       0.65900       0.10653       0.35289      ...
         145         9  0.25916       0.21268       0.31821      -0.57542      -0.25931      -0.20978      ...
         146         9  0.54251       0.62274       0.38799       0.31331       0.13336       0.28073      ...
         157        11 -0.92067      -0.79038      -0.44410      -0.73007E-01 -0.31008      -0.29550      ...
         163        11  -3.6774       -2.6107       -2.0315       0.86495       0.23697       0.12217      ...
         165        12   1.8915        1.0388        1.0100        3.1267        2.4545        1.6869      ...
         226         1 -0.55895      -0.37087      -0.32768       0.20016       0.51922E-01  0.53863E-01 ...
```



```
    "Solani"-File: Solutions for Genetic Animal Effect

     Column  |  Description
     ----------------------
        1       Animal ID
        2       Number of Descendants
        3       Number of Observations
        4       Solution for Trait  1  MILK       and Factor   ANI
        5       Solution for Trait  2  PROTEIN    and Factor   ANI
        6       Solution for Trait  3  FAT        and Factor   ANI
        7       Solution for Trait  1  MILK       and Factor   T2(ANI)
        8       Solution for Trait  2  PROTEIN    and Factor   T2(ANI)
        9       Solution for Trait  3  FAT        and Factor   T2(ANI)
       10       Solution for Trait  1  MILK       and Factor   T3(ANI)
       11       Solution for Trait  2  PROTEIN    and Factor   T3(ANI)
       12       Solution for Trait  3  FAT        and Factor   T3(ANI)
       13       Solution for Trait  1  MILK       and Factor   T5(ANI)
       14       Solution for Trait  2  PROTEIN    and Factor   T5(ANI)
       15       Solution for Trait  3  FAT        and Factor   T5(ANI)

head Solani
           1       1       0   3.5729        2.4194        1.5256       -1.1981       -0.27947      -0.26116 ...
           4       1       0   2.3634        1.2145        1.4067       -1.6518       -0.49062      -0.52810 ...
           7       1       0   2.7368        1.9404        2.0480       -1.2383       -0.25751      -0.18850 ...
          10       1       0   2.8937        1.8512        1.3084       -1.0411       -0.19864      -0.12744 ...
          13       1       0   2.7386        1.6019        1.0450      -0.70816      -0.47939E-01 -0.77933E...
          16       1       0   1.9904       0.79604       0.58778      -1.0534       -0.32418      -0.35039 ...
           3       2       0   3.2071        2.1734        1.2498       -1.1350       -0.15555      -0.16690 ...
          20       1       0   3.3834        2.0699        1.3694       -1.1374       -0.23903      -0.31426 ...
          23       1       0   3.8460        3.0194        1.2792       -1.6866       -0.45594      -0.12098 ...
          25       1       0   1.2965       0.79091       0.91466      -0.82220      -0.18907      -0.18593 ...
```

© Luke

Trait groups

- The previous models had only first lactation traits
- Test-day models can have more than 1 lactation.
- Usually milk, protein and fat yields for each lactation are considered as own traits
- Trait group feature in MiX99 uses one integer column in the data to specify which trait group the record is from.
- In the model line, trait group number is within parentheses after trait number.
- In general, effects and traits can vary between trait groups.

Test-day model for 3 lactations

© Luke

- Integer column **LACT** specifies the lactation
- In this model, each trait has the same effect structure.
- No need to define lactation specific columns for observations!

```
INTEGER    HERD ANI LACT HTD HC2Y C2Y AGE DCC DDRY DIM ResClass
REAL       MILK PROTEIN FAT


RANDOM   HTD  PE G
TRAITGROUP LACT

MODEL
MILK(1)    =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
PROTEIN(1) =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
FAT(1)     =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
MILK(2)    =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
PROTEIN(2) =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
FAT(2)     =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
MILK(3)    =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
PROTEIN(3) =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
FAT(3)     =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)



head lact3.dat
2669 457 1 128191 266903 3 7 1 5 20 1 15.7575 11.8857 9.8591
2669 457 1 128192 266903 3 7 1 5 48 3 20.0062 15.5429 11.3455
2669 457 1 128193 266903 3 7 1 5 75 5 17.2425 12.6143 9.2864
2669 457 1 128194 266903 3 7 1 5 97 6 16.7475 12.6286 10.4727
2669 457 1 128195 266903 3 7 1 5 123 7 16.1287 12.5286 9.6136
2669 457 1 128196 266903 3 7 1 5 158 7 13.1175 10.2 8.2091
2669 457 1 128197 266903 3 7 1 5 181 8 14.4994 11.9286 9.6955
```

© Luke

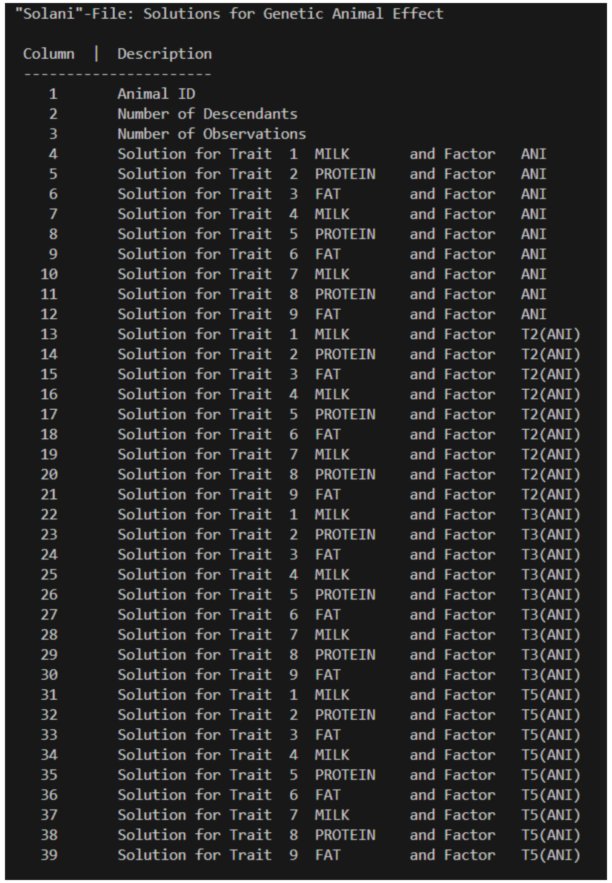
Test-day model for 3 lactations: variance components

- Trait groups don’t have direct effect to covariance matrices.
- The order of the model lines defines the order of variance components.
- The sizes of covariance matrices:
  - HTD = 9x9
  - PE = 36x36
  - G = 36x36
  - Residual has 12 classes. Each class has 9x9 covariance matrix. Covariance between lactations 0.
- In total 1593 variance components!

MILK(1) = .. HTD_1 PE(1_1 $t2_{10}$ $t3_{19}$ $t5_{28}$ |ANI) G(1_1 $t2_{10}$ $t3_{19}$ $t5_{28}$ |ANI)
PROTEIN(1) = .. HTD_2 PE(1_2 $t2_{11}$ $t3_{20}$ $t5_{29}$ |ANI) G(1_2 $t2_{11}$ $t3_{20}$ $t5_{29}$ |ANI)
FAT(1) = .. HTD_3 PE(1_3 $t2_{12}$ $t3_{21}$ $t5_{30}$ |ANI) G(1_3 $t2_{12}$ $t3_{21}$ $t5_{30}$ |ANI)
MILK(2) = .. HTD_4 PE(1_4 $t2_{13}$ $t3_{22}$ $t5_{31}$ |ANI) G(1_4 $t2_{13}$ $t3_{22}$ $t5_{31}$ |ANI)
PROTEIN(2) = .. HTD_5 PE(1_5 $t2_{14}$ $t3_{23}$ $t5_{32}$ |ANI) G(1_5 $t2_{14}$ $t3_{23}$ $t5_{32}$ |ANI)
FAT(2) = .. HTD_6 PE(1_6 $t2_{15}$ $t3_{24}$ $t5_{33}$ |ANI) G(1_6 $t2_{15}$ $t3_{24}$ $t5_{33}$ |ANI)
MILK(3) = .. HTD_7 PE(1_7 $t2_{16}$ $t3_{25}$ $t5_{34}$ |ANI) G(1_7 $t2_{16}$ $t3_{25}$ $t5_{34}$ |ANI)
PROTEIN(3) = .. HTD_8 PE(1_8 $t2_{17}$ $t3_{26}$ $t5_{35}$ |ANI) G(1_8 $t2_{17}$ $t3_{26}$ $t5_{35}$ |ANI)
FAT1(3) = .. HTD_9 PE(1_9 $t2_{18}$ $t3_{27}$ $t5_{36}$ |ANI) G(1_9 $t2_{18}$ $t3_{27}$ $t5_{36}$ |ANI)

Genetic solutions for 9 trait model

© Luke



```
"Solani"-File: Solutions for Genetic Animal Effect

 Column  |  Description
 ---------------------
    1       Animal ID
    2       Number of Descendants
    3       Number of Observations
    4       Solution for Trait  1  MILK       and Factor   ANI
    5       Solution for Trait  2  PROTEIN    and Factor   ANI
    6       Solution for Trait  3  FAT        and Factor   ANI
    7       Solution for Trait  4  MILK       and Factor   ANI
    8       Solution for Trait  5  PROTEIN    and Factor   ANI
    9       Solution for Trait  6  FAT        and Factor   ANI
   10       Solution for Trait  7  MILK       and Factor   ANI
   11       Solution for Trait  8  PROTEIN    and Factor   ANI
   12       Solution for Trait  9  FAT        and Factor   ANI
   13       Solution for Trait  1  MILK       and Factor   T2(ANI)
   14       Solution for Trait  2  PROTEIN    and Factor   T2(ANI)
   15       Solution for Trait  3  FAT        and Factor   T2(ANI)
   16       Solution for Trait  4  MILK       and Factor   T2(ANI)
   17       Solution for Trait  5  PROTEIN    and Factor   T2(ANI)
   18       Solution for Trait  6  FAT        and Factor   T2(ANI)
   19       Solution for Trait  7  MILK       and Factor   T2(ANI)
   20       Solution for Trait  8  PROTEIN    and Factor   T2(ANI)
   21       Solution for Trait  9  FAT        and Factor   T2(ANI)
   22       Solution for Trait  1  MILK       and Factor   T3(ANI)
   23       Solution for Trait  2  PROTEIN    and Factor   T3(ANI)
   24       Solution for Trait  3  FAT        and Factor   T3(ANI)
   25       Solution for Trait  4  MILK       and Factor   T3(ANI)
   26       Solution for Trait  5  PROTEIN    and Factor   T3(ANI)
   27       Solution for Trait  6  FAT        and Factor   T3(ANI)
   28       Solution for Trait  7  MILK       and Factor   T3(ANI)
   29       Solution for Trait  8  PROTEIN    and Factor   T3(ANI)
   30       Solution for Trait  9  FAT        and Factor   T3(ANI)
   31       Solution for Trait  1  MILK       and Factor   T5(ANI)
   32       Solution for Trait  2  PROTEIN    and Factor   T5(ANI)
   33       Solution for Trait  3  FAT        and Factor   T5(ANI)
   34       Solution for Trait  4  MILK       and Factor   T5(ANI)
   35       Solution for Trait  5  PROTEIN    and Factor   T5(ANI)
   36       Solution for Trait  6  FAT        and Factor   T5(ANI)
   37       Solution for Trait  7  MILK       and Factor   T5(ANI)
   38       Solution for Trait  8  PROTEIN    and Factor   T5(ANI)
   39       Solution for Trait  9  FAT        and Factor   T5(ANI)
```

- Solani file has 39 columns
- When using the columns in Solani, care must be taken to ensure that the calculations are correct

© Luke

DATASORT

```
INTEGER   HERD ANI LACT HTD HC2Y C2Y AGE DCC DDRY DIM ResClass
REAL      MILK PROTEIN FAT

RANDOM   HTD  PE G
TRAITGROUP LACT

MODEL
MILK(1)     =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
PROTEIN(1) =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
FAT(1)      =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
MILK(2)     =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
PROTEIN(2) =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
FAT(2)      =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
MILK(3)     =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
PROTEIN(3) =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
FAT(3)      =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)

DATASORT  BLOCK=HERD PEDIGREECODE=ANI

WITHINBLOCKORDER G PE HTD
```

- Ordering data AND pedigree by block can speed up solving the model
- Natural block in test-day models are herd – code
- Pedigree must contain block code for each animal (also for pedigree animals) as 4th column.
- Integer variable for block code and pedigree code is given in DATASORT BLOCK command

© Luke

WITHINBLOCKORDER

```
INTEGER    HERD ANI LACT HTD HC2Y C2Y AGE DCC DDRY DIM ResClass
REAL       MILK PROTEIN FAT

RANDOM   HTD  PE G
TRAITGROUP LACT

MODEL
MILK(1)    =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
PROTEIN(1) =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
FAT(1)     =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
MILK(2)    =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
PROTEIN(2) =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
FAT(2)     =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
MILK(3)    =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
PROTEIN(3) =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)
FAT(3)     =  LACcurve(t1 t2 t3 t4) HTD PE(1 t2 t3 t5|ANI)  G(1 t2 t3 t5|ANI)

DATASORT  BLOCK=HERD PEDIGREECODE=ANI

WITHINBLOCKORDER G PE HTD
```

- When BLOCK code is given, some of the effects can be classified as “within block” effects using WITHINBLOCKORDER statement.
- Within block effect levels usually have interaction with block variable, like HERD x TEST-DAY, or whose observations mainly belong to one block, like animal effects .
- Genetic effect must be always the first within block effect.
- WITHINBLOCKORDER is needed for R2 calculation.

© Luke

Sorting data and pedigree for DATASORT

```
INTEGER    HERD ANI LACT HTD HC2Y C2Y AGE DCC DDRY DIM ResClass
REAL       MILK PROTEIN FAT

TRAITGROUP LACT

DATASORT  BLOCK=HERD PEDIGREECODE=ANI

WITHINBLOCKORDER G PE HTD
```

- MiX99 checks if the data is in correct order

DATA

```
#               BLOCK    animal   tgrp    DIM
 >sort  -n -k1,1  -k2,2 -k3,3 -k10,10  lact3.dat > lact3.s
```

Pedigree

```
#               BLOCK    ANIMAL
>sort  -n -k4,4  -k1,1   TD9.ped
```

Conclusion

© Luke

This lecture briefly covered the following topics:

- Setting up single trait random regression model
- multiple trait model without trait groups
- multiple trait model using trait groups
- parameter files for these models
- format of solution files

Questions?