

Analysis of repeated observations using ASReml and Echidna.

Arthur Gilmour Arthur.Gilmour@cargovale.com.au Dec 2020

Introduction

These notes were generated as a discussion with a PhD student at AGBU, Peter. They are now made available for others interested in fitting models to repeated observations. Both ASReml and Echidna may be used to fit these models.

Repeated observations come in two basic forms, regular and irregular.

Regular repeated measures

Regular repeated measures are taken on all subjects at regular intervals, say monthly, or at particular stages, say birth, weaning, yearling, mature.

Irregular repeated measures are more opportunistic, measures are taken when convenient with respect to the individual but at different times/ages across individuals.

Regular repeated measures can be analysed with an overall mean trend (mean for each time point) and an antedependence variance structure across times. For example, in some data on growth of Thai chickens, the 11677 birds were weighed at day old, and monthly until six months. The mean weights were

Age	1	28	56	84	112	140	168
Weight	30.9±3.4	218.9±56.7	642±139	1098±210	1486±306	1810±406	2124±470

The birds were raised in 88 cohorts. Presumably these were at different times and so would have had individual cohort growth patterns. We could fit these as Age.Cohort set of effects with Age a factor, or as Cohort and Cohort.Age effects where Age is a covariate (say -3,-2,-1,0,1,2,3). I have done the latter treating intercept and slope as random. That is, random regressions about the overall response pattern. One could extend this to a quadratic response.

The pedigree of the birds was known and is summarised in ASReml as

12229 identities in the pedigree over 5 generations.

For first parent labelled Sire, second labelled Dam
Sire Sire_of_Sire Dam_of_Sire Dam Sire_of_Dam Dam_of_Dam
487 178 246 1470 266 518

We can use an ANTEDEPENDENCE covariance structure for the genetic effects. The ANTEDEPENDENCE 1 structure recognises that the size at time t+1 is related to the size at time t and growth in the interval. We will use the standard UNSTRUCTURED variance matrix for the Residual as it picks up as much variation as possible.

This leads us to the following ASReml results:

- - - Results from analysis of WT1 WT28 WT56 WT84 WT112 WT140 WT168 - - -

18 LogL= 7899.80 S2= 1.0000 66112 df

Covariance/Variance/Correlation Matrix US Residual
4.304 0.07640 0.07124 0.07515 0.08214 0.08851 0.08906
4.565 829.4 0.3779 0.1658 -0.008651 -0.06857 -0.1123

```

10.82    797.1    5363.    0.6720    0.5365    0.4611    0.4164
20.65    632.2    6517.    17540    0.7711    0.6809    0.6426
34.71    -50.75    8003.    20800    41490    0.8592    0.7998
51.36    -552.3    9445.    25220    48950    78240    0.8798
61.26    -1072.    10110    28210    54020    81590    109900
Covariance/Variance/Correlation Matrix ANTE ante1(Tr).nrm(aid)
5.974    0.3416    0.2840    0.2083    0.1986    0.1962    0.1920
16.54    392.4    0.8314    0.6097    0.5814    0.5743    0.5621
44.14    1047.    4042.    0.7333    0.6993    0.6908    0.6761
60.88    1444.    5576.    14300    0.9536    0.9420    0.9220
87.64    2079.    8027.    20590    32600    0.9879    0.9669
101.2    2402.    9272.    23790    37660    44570    0.9788
119.2    2829.    10920    28010    44350    52500    64540
Covariance/Variance/Correlation Matrix US hys
0.1976E+05    1.000
6597.    2203.

```

```

Wald F statistics
Source of Variation    NumDF    F-inc
17 Trait                7        4591.51
6 sex                   1        23.90
19 sex.LinA            1        2503.34

```

The heritabilities calculated from these matrices are

Age	1	28	56	84	112	140	168
Weight	0.58±0.??	0.32±0.??	0.43±0.??	0.45±0.??	0.44±0.??	0.36±0.??	0.37±0.??

The genetic correlations are high among the last 4 times. And drop with increasing time interval is modelled in the antedependence structure,

There are 3 other traits I have labelled TA (205 ±25), TB (199±21) and TC (36.9±4.8). Following is an analysis of these traits with Dayold weight and 3 month weight. Here I have used US for both strata because the antedependence logic does not apply.

```

14 LogL=-4917.53    S2= 1.0000    25574 df

```

- - - Results from analysis of WT1 WT84 TA TB TC - - -

```

Covariance/Variance/Correlation Matrix US Residual
4.311    0.6349E-01    -0.4292E-01    -0.8485E-01    -0.2192
16.33    0.1535E+05    0.5237    -0.3197E-01    0.4435E-01
-1.566    1140.    308.7    0.1935    0.1729
-2.877    -64.68    55.54    266.7    0.3086
-1.803    21.77    12.04    19.97    15.70
Covariance/Variance/Correlation Matrix US us(Tr).nrm(aid)
5.953    0.1826    0.4390    0.1903    0.6817
59.22    0.1766E+05    0.3876    -0.5786    -0.1069
18.42    886.2    295.9    0.3567    0.4752
6.005    -994.7    79.37    167.3    0.5772
3.344    -28.58    16.44    15.01    4.043
Covariance/Variance/Correlation Matrix US hys
1.321    0.7134
0.3030    0.1366

```

```

Wald F statistics
Source of Variation    NumDF    F-inc
17 Trait                5        16813.46

```

6 sex

1

1.06

Irregular repeated measures: Random regression

In this case there is no neat pattern to the repeated observations but we can envisage the measures follow some kind of curve with respect to an underlying age.

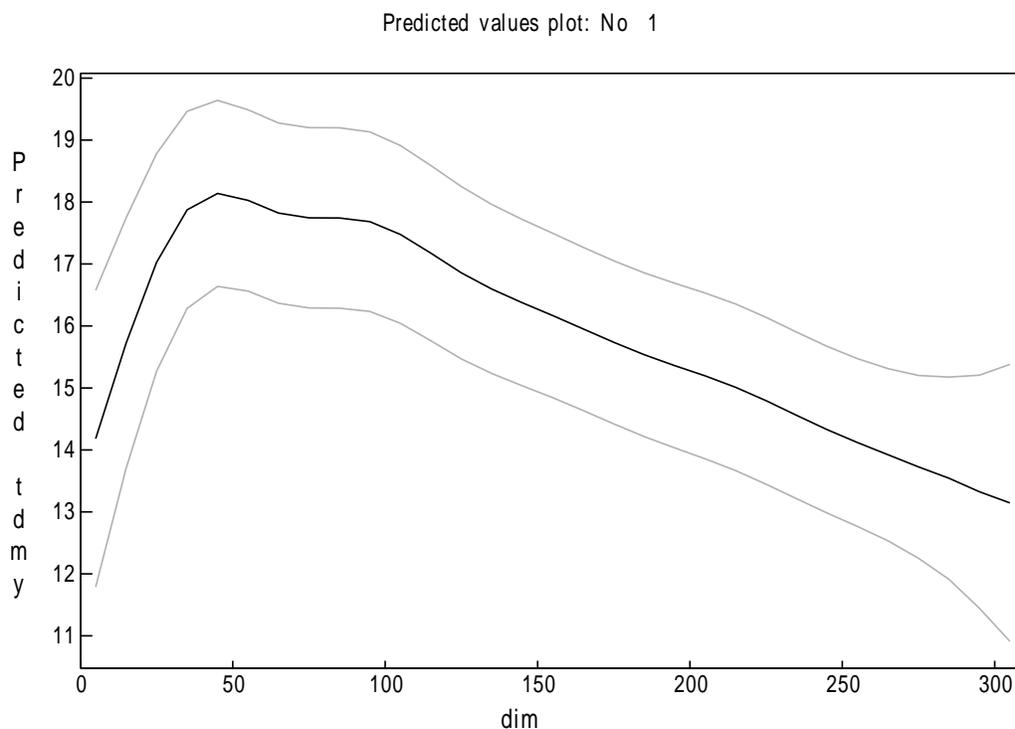
One approach might be split the ages into classes and apply the Regular method to the classes.

However, you should not have more classes than the average number of measures per subject. This approach can work for a variable like milk yield typically measured each month and classified by Days in Milk, even though the actual DIM values are more or less evenly distributed. How this does not work in general.

The idea of a random regression model is that we model the basic common response curve and then model random deviations of the individuals from the common curve. We discuss this model in relation to some milk yield data.

Many studies have shown that milk yield rises quickly early in lactation and drops slowly for a period, then more rapidly at the end. This shape is not easily modelled as a low order polynomial but can be modelled as a cubic smoothing spline.

The full data set contains milk yields from several lactations but first consider the 10 or so milk yields taken at monthly intervals for the first lactation of 2484 cows. The variances as well as the mean change over time. A typical lactation curve is shown in the figure.



For the test data, the period variances were

Period	5-15	16-30	31-60	61-90	-120	-150	-180	-210	-240	-270	-305
	9.2	7.0	7.6	7.0	5.9	6.1	5.7	5.5	6.0	5.5	7.0

These results come from fitting the model

```
tdmy ~ mu dim CG !r spl(dim,11) us(leg(dim,3)).nrm(aid) us(leg(dim,3)).ide(aid) gg  
residual at(SEC).units  
predict aid 466 dim 5 15 : 305 !PLOT
```

given that the data is organised as 1 record per milk yield with covariables dim and SEC (defining the 11 DIM classes). Note that cows will only have 1 record in the interval 5-30 days so different cows appear in the 5-15 group than in the 16-30 group.

The cubic spline is modelled by the model terms `mu dim !r spl(dim,11)` and is built by adding the random curvature components (`spl(dim,11)`) to an underlying straight line (`mu dim`). The variance component for the curvature component (`spl(dim,11)`) controls just how much curvature appears.

The cows are mainly crossbreds and the model term `gg` contains proportions of the base breeds.

The ASReml results:

```
1 LogL=-911.947 S2= 1.0000 23386 df  
- - - Results from analysis of tdmy - - -  
Akaike Information Criterion 81889.89 (assuming 33 parameters).  
Bayesian Information Criterion 82155.87
```

Model_Term	Sigma	Sigma	Sigma/SE	% C
------------	-------	-------	----------	-----

spl(dim,11)	IDV_V	9	2.54025	2.54025	2.05	0 P
gg	IDV_V	74	3.16957	3.16957	2.06	0 P
Residual_1	SCA_V	1300	9.18989	9.18989	14.01	0 P
Residual_2	SCA_V	1159	7.01477	7.01477	13.82	0 P
Residual_3	SCA_V	2368	7.55529	7.55529	25.62	0 P
Residual_4	SCA_V	2451	6.98987	6.98987	27.01	0 P
Residual_5	SCA_V	2424	5.92574	5.92574	25.94	0 P
Residual_6	SCA_V	2441	6.06780	6.06780	27.64	0 P
Residual_7	SCA_V	2444	5.72172	5.72172	27.48	0 P
Residual_8	SCA_V	2370	5.50560	5.50560	26.34	0 P
Residual_9	SCA_V	2307	6.03564	6.03564	26.24	0 P
Residual_10	SCA_V	2214	5.52521	5.52521	23.34	0 P
Residual_11	SCA_V	2292	7.04258	7.04258	20.12	0 P

us(leg(dim,3)).nrm(aid) 124960 effects

leg(dim,3)	US_V	1	1	24.4151	24.4151	11.08	0 P
leg(dim,3)	US_C	2	1	-0.153850	-0.153850	-0.26	0 P
leg(dim,3)	US_V	2	2	1.34127	1.34127	4.70	0 P
leg(dim,3)	US_C	3	1	-2.61239	-2.61239	-5.21	0 P
leg(dim,3)	US_C	3	2	-0.792845E-01	-0.792845E-01	-0.48	0 P
leg(dim,3)	US_V	3	3	0.829146	0.829146	4.28	0 P
leg(dim,3)	US_C	4	1	1.66090	1.66090	5.05	0 P
leg(dim,3)	US_C	4	2	0.102657	0.102657	0.99	0 P
leg(dim,3)	US_C	4	3	-0.195414	-0.195414	-2.28	0 P
leg(dim,3)	US_V	4	4	0.132987	0.132987	1.90	0 P

aid NRM 31240

us(leg(dim,3)).ide(aid) 124960 effects

leg(dim,3)	US_V	1	1	10.5392	10.5392	6.58	0 P
leg(dim,3)	US_C	2	1	0.454458	0.454458	0.94	0 P
leg(dim,3)	US_V	2	2	2.76552	2.76552	9.85	0 P
leg(dim,3)	US_C	3	1	-0.733449	-0.733449	-1.77	0 P
leg(dim,3)	US_C	3	2	0.232973	0.232973	1.41	0 P
leg(dim,3)	US_V	3	3	1.54225	1.54225	7.89	0 P
leg(dim,3)	US_C	4	1	-0.349937	-0.349937	-1.20	0 P
leg(dim,3)	US_C	4	2	-0.441040	-0.441040	-3.83	0 P
leg(dim,3)	US_C	4	3	-0.188553	-0.188553	-1.96	0 P
leg(dim,3)	US_V	4	4	0.652034	0.652034	6.92	0 P

Covariance/Variance/Correlation Matrix US us(leg(dim,3)).nrm(a)

24.42	-0.2689E-01	-0.5806	0.9217
-0.1539	1.341	-0.7518E-01	0.2431
-2.612	-0.7928E-01	0.8291	-0.5885
1.661	0.1027	-0.1954	0.1330

Covariance/Variance/Correlation Matrix US us(leg(dim,3)).ide(a)

10.54	0.8418E-01	-0.1819	-0.1335
0.4545	2.766	0.1128	-0.3284
-0.7335	0.2330	1.542	-0.1880
-0.3499	-0.4410	-0.1886	0.6520

Wald F statistics

Source of Variation	NumDF	F-inc
14 mu	1	1356.58
8 dim	1	599.59

Now this is a random regression model. We have genetic relationships among the cows (aid = animal ID). The model term `us(leg(dim,3)).nrm(aid)` fits the genetic effects modelling random cubic deviations from the common curve. The model term `us(leg(dim,3)).ide(aid)` fits the covariance between the repeated observations, also modelled as cubic deviations.

The CG term is contemporary groups. The 2484 cows represent 383 herds (average 8 cows in each).

Now I do not know whether these cows are subject to seasonal breeding. If not, then the lactation curve for spring calving cows could be different from autumn calving cows. I suspect these are tropical cattle but whether there are seasonal effects should be considered.

The `leg(dim,3)` transforms the 10 or so dim values to cubic Legendre polynomial coefficients. The actual coefficients are reported for each dim value in the .res file. The predict statement above nominated DIM values of 5, 15 : 305. The coefficients for these values are:

5.00000	0.70711	-1.22474	1.58114	-1.87083
15.00000	0.70711	-1.14310	1.27545	-1.18347
25.00000	0.70711	-1.06145	0.99085	-0.61252
35.00000	0.70711	-0.97980	0.72732	-0.14967
45.00000	0.70711	-0.89815	0.48488	0.21341
55.00000	0.70711	-0.81650	0.26352	0.48503
65.00000	0.70711	-0.73485	0.06325	0.67350
75.00000	0.70711	-0.65320	-0.11595	0.78713
85.00000	0.70711	-0.57155	-0.27406	0.83425
95.00000	0.70711	-0.48990	-0.41110	0.82316
...				
255.00000	0.70711	0.81650	0.26352	-0.48503
265.00000	0.70711	0.89815	0.48488	-0.21341
275.00000	0.70711	0.97980	0.72732	0.14967
285.00000	0.70711	1.06145	0.99085	0.61252
295.00000	0.70711	1.14310	1.27545	1.18347
305.00000	0.70711	1.22474	1.58114	1.87083

Now, fitting this model can be difficult because we like to have to estimate the covariances between coefficients as well as the variances, and the ASReml iteration updates may not result in a positive definite matrix.

We end up with a 4x4 variance matrix for genetic effects, and for error covariance effects. Note that the matrix is reported as correlations in the top right half (above the main diagonal) but in using the matrix, it must be symmetric based on the covariances below the diagonal.

Now this wasn't the first model I fitted. First I just fitted linear random regressions:

```
!PART 1
tdmy ~ mu dim CG !r spl(dim,11) us(leg(dim,1)).nrm(aid) us(leg(dim,1)).ide(aid) gg
residual at(SEC).units
predict aid 466 dim 5 15 : 305 !PLOT
```

And obtained a LogLikelihood LogL=-1628.45

Then I fitted quadratic random regressions:

```
!PART 2
tdmy ~ mu dim CG !r spl(dim,11) us(leg(dim,2)).nrm(aid) us(leg(dim,2)).ide(aid) gg
residual at(SEC).units
predict aid 466 dim 5 15 : 305 !PLOT
```

And obtained a LogLikelihood LogL=-1057.55

The likelihood from the cubic random regression was LogL=-911.947

The likelihood would probably further increase if we went to a quartic regression but this is basically because the lactation curve is not really polynomial in nature.

To obtain an estimate of heritability for any particular day, we need first to calculate the variance components for that day.

e.g. from CSS3.res

```
...
30.00000 0.70711 -1.02062 0.85645 -0.36810
...
90.00000 0.70711 -0.53072 -0.34522 0.83546
...
240.00000 0.70711 0.69402 -0.02899 -0.73915
```

The residual variances at these 3 DIM are 7.01476, 6.98986, 6.03564

The variance matrices are

```
Covariance/Variance/Correlation Matrix US us(leg(dim,3)).nrm(aid)
24.42 -0.2689E-01 -0.5806 0.9217
-0.1539 1.341 -0.7518E-01 0.2431
-2.612 -0.7929E-01 0.8291 -0.5885
1.661 0.1027 -0.1954 0.1330
Covariance/Variance/Correlation Matrix US us(leg(dim,3)).ide(a
10.54 0.8418E-01 -0.1819 -0.1335
0.4545 2.766 0.1128 -0.3284
-0.7335 0.2330 1.542 -0.1880
-0.3499 -0.4410 -0.1886 0.6520
```

Putting this in C3.R

```
L30 =c( 0.70711, -1.02062, 0.85645, -0.36810)
L90 =c( 0.70711, -0.53072, -0.34522, 0.83546)
L240 =c( 0.70711, 0.69402, -0.02899, -0.73915)
E=c( 7.01476, 6.98986, 6.03564)
NRM=matrix(c(
24.42, -0.02689, -0.5806, 0.9217,
-0.1539, 1.341, -0.07518, 0.2431,
-2.612, -0.07929, 0.8291, -0.5885,
1.661, 0.1027, -0.1954, 0.1330),4,4)
IDE=matrix(c(
10.54, 0.08418, -0.1819, -0.1335,
0.4545, 2.766, 0.1128, -0.3284,
-0.7335, 0.2330, 1.542, -0.1880,
-0.3499, -0.4410, -0.1886, 0.6520),4,4)
NRM[1,2:4]->NRM[2:4,1]
```

```

NRM[2,3:4]->NRM[3:4,2]
NRM[3,4]->NRM[4,3]
IDE[1,2:4]->IDE[2:4,1]
IDE[2,3:4]->IDE[3:4,2]
IDE[3,4]->IDE[4,3]

```

```

A=c( t(L30) %*% NRM %*% L30, t(L90) %*% NRM %*% L90,t(L240) %*% NRM %*% L240 )
PE= c( t(L30) %*% IDE %*% L30, t(L90) %*% IDE %*% L90,t(L240) %*% IDE %*% L240 )
h2 = A/(A+PE+E)

```

we can use R to perform the calculations.

```

> source("C3.txt")
> h2
[1] 0.4277400 0.5376540 0.4361185

```

The student seeking to analyse this data was interested in getting a genetic covariance/correlation between milk yield (DMY) at particular values of DIM and AFC (Age at first calving). AFC is not a repeated measure.

The method of analysis was to expand the data file by adding the AFC records as SECTION 12 and creating a new factor TRAIT coded 1 for DMY and 2 for AFC; the AFC values being placed in the same data field as the DMY records.

The model specification is

```

!PART 3
res ~ SEC +
!r !{ at(TR,1).leg(dim,3).aid + at(TR,2).aid us(5).nrm(aid) !} +
!{ at(TR,1).leg(dim,3).ide(aid) at(TR,2).ide(aid) us(5).ide(aid) !} +
us(TRAIT).gg !f at(TR,1).CG at(TR,2).CGB
residual at(SEC).units

```

The code `!{ at(TR,1).leg(dim,3).aid at(TR,2).aid us(5).nrm(aid) !}` specifies 2 model terms and then the variance structure that is put across the model terms. The effect of `!{ !}` is to prevent ASReml reordering the terms and messing up the association. We need this kind of structure for both the genetic components and the residual components. However, this model is over-specified because the residual variance for AFC is specified twice (as the last element in `us(5).ide(aid)` and as `at(SEC,12).units`). It also has very poor parameter starting values. So, start the job running and when it has parsed the model, it writes a .tsv file. Interrupt the job and update as many of the parameter values as you can from the results from fitting PART 2 (cubic Legendre). Also, set the parameter for `at(SEC,12).units` to 1 and set the parameter constraint code to F (fixed). Then restart the run using `!CONTINUE 2`

The 5x5 us matrices pertain to the 4 Legendre polynomials and AFC. To get the covariance between say day 30 MY and AFC, a similar process is needed as described above. Let

```

L30A =matrix(c( 0.70711, -1.02062, 0.85645, -0.36810, 0.0, 0.0, 0.0, 0.0, 0.0, 1.0),5,2)
And calculate t(L30A) %*% NRM %*% L30A

```

This gives a 2x2 matrix with variances and the covariance.

18 June

I will do a few other things before working on the LOW data you sent but note
Analysis of res

Source of Variation	Wald F statistics			F-inc	P-inc
	NumDF	DenDF			
mu	1			32682.24	
CGB	518			4.36	
Model_Term	Order	Gamma	Sigma	Z_ratio	%C
gg	74	0.834754E-06	0.778500E-03	0.00	0 B
nrm(aid)	31240	0.798680	744.858	2.03	0 P
ide(aid)	31240	0.619612E-06	0.577857E-03	0.00	0 B
Residual_units	1016	1.00000	932.611	2.81	

The number of contemporary groups seems excessive (takes up over half the DF) for AFC.
For this analysis I scaled the AFC values by 0.25

In the other analysis, CGB was not declared a FACTOR. Maybe it should have been, in which case the variance components will be wrong there.

xfa1(leg(dim,4)).nrm(aid) 187440 effects							
leg(dim,4)	XFA_V	0	1	0.00000	0.00000	0.00	0 F
leg(dim,4)	XFA_V	0	2	1.30175	1.30175	4.90	0 P
leg(dim,4)	XFA_V	0	3	0.216441	0.216441	1.92	0 P
leg(dim,4)	XFA_V	0	4	0.817112E-01	0.817112E-01	0.98	0 P
leg(dim,4)	XFA_V	0	5	0.00000	0.00000	0.00	0 F
leg(dim,4)	XFA_L	1	1	-1.58336	-1.58336	-5.97	0 P
leg(dim,4)	XFA_L	1	2	-0.632362E-01	-0.632362E-01	-0.42	0 P
leg(dim,4)	XFA_L	1	3	0.862410E-01	0.862410E-01	1.11	0 P
leg(dim,4)	XFA_L	1	4	0.155947	0.155947	1.64	0 P
leg(dim,4)	XFA_L	1	5	0.546649E-01	0.546649E-01	1.10	0 P
aid	NRM	31240					

xfa1(leg(dim,4)).ide(pe) 6174 effects							
leg(dim,4)	XFA_V	0	1	2.82678	2.82678	2.68	0 P
leg(dim,4)	XFA_V	0	2	0.361102	0.361102	1.47	0 P
leg(dim,4)	XFA_V	0	3	0.398617	0.398617	3.46	0 P
leg(dim,4)	XFA_V	0	4	0.164632	0.164632	1.77	0 P
leg(dim,4)	XFA_V	0	5	0.205445	0.205445	6.72	0 P
leg(dim,4)	XFA_L	1	1	-1.37671	-1.37671	-4.32	0 P
leg(dim,4)	XFA_L	1	2	0.486370	0.486370	4.34	0 P
leg(dim,4)	XFA_L	1	3	0.219622	0.219622	3.62	0 P
leg(dim,4)	XFA_L	1	4	-0.323878	-0.323878	-3.74	0 P
leg(dim,4)	XFA_L	1	5	0.769527E-01	0.769527E-01	1.73	0 P

Covariance/Variance/Correlation Matrix XFA xfa1(leg(dim,4)).nrm

2.506	0.5532E-01	-0.1821	-0.4800	-1.000	-1.000
0.1001	1.306	-0.1007E-01	-0.2655E-01	-0.5532E-01	-0.5532E-01
-0.1364	-0.5446E-02	0.2238	0.8739E-01	0.1821	0.1821
-0.2474	-0.9877E-02	0.1346E-01	0.1060	0.4800	0.4800
-0.8638E-01	-0.3449E-02	0.4700E-02	0.8524E-02	0.2977E-02	1.000
-1.583	-0.6322E-01	0.8615E-01	0.1562	0.5456E-01	1.000

Covariance/Variance/Correlation Matrix XFA xfa1(leg(dim,4)).ide

4.723	-0.3985	-0.2082	0.3955	-0.1062	-0.6338
-0.6696	0.5977	0.2066	-0.3924	0.1054	0.6288
-0.3025	0.1068	0.4469	-0.2050	0.5504E-01	0.3285
0.4463	-0.1575	-0.7116E-01	0.2696	-0.1046	-0.6240
-0.1061	0.3745E-01	0.1692E-01	-0.2496E-01	0.2114	0.1676
-1.377	0.4861	0.2196	-0.3240	0.7703E-01	1.000

How can I use these estimates to estimate additive variance for different days?

The leading 5x5 block of the 2 6x6 matrices is equivalent to the 5x5 matrix you would have attempted to estimate using us() (say A). So, take the 5 coefficients for the date of interest from the res file (say c) and form c'Ac to get the variance for that day.

This has now been facilitated in VPREDICT commands K and M

V NRM nrm

V IDE ide

K coeff 0.70711, -1.02062, 0.85645, -0.36810, 0.0

K coeff 0.0, 0.0, 0.0, 0.0, 1.0

M Add30 1:2 NRM

Further discussion: Please explain the relationship between the random regression terms and the spline terms.

Genetic variance and PE variance relates to differences between animals. Those differences are centred on a mean value. We have used a cubic smoothing spline to model the mean curve, that is the non genetic effect.

The fact we use a cubic spline is not connected with the degree of the legendre polynomials used for the between animal differences random regression.

The overall lactation curve has a complex shape and is not easily represented by a simple algebraic function. Some have taken that approach (Wilmot curve) but by using a spline, we allow the data to define the overall curve. While DIM values cover the whole range 5:305, we actually only have 11 measures per animal at regular (monthly) intervals so we use 12 knot points.

The spline is constructed by a fixed set of covariates. With 12 knot points, we have 12 covariates being a fixed intercept and slope and 10 random curvature components. The random covariates are given the

same variance which determines the degree of curvature. If the variance component is zero, we just have a straight line.

If all the measures were taken at 11 fixed DIM values, we could in fact use the mean of the values for each date for the common curve. However, it may not be smooth. The spline covariates result in a smooth curve.

Now individuals vary around the common curve. Part of that variation is independent noise (the cow may be sick or injured or some other random thing has happened to influence a particular measurement on a particular day). This is the residual component. We are however interested in systematic differences. These are modelled by the cubic regression across the cows' deviations from the common (spline) curve. Thus a high yielding cow should tend to be high on all 11 measures, a low cow low at each. Some will start high and finish low, or vice versus.

Ultimately I suppose you are interested in the total milk yield over say 300 days. This would involve the integration of the area under the curve. Since we only have monthly measures, it is enough to sum the individual monthly predictions. The sum consists of the component due to the average curve and the bit due to the deviations. Say we have predictions at day 10 40 70 and 100. Using linear interpolation the area is $(P_{10} + P_{40})/2 * 30 + (P_{40} + P_{70})/2 * 30 + (P_{70} + P_{100})/2 * 30 = (P_{10}/2 + P_{40} + P_{70} + P_{100}/2) * 30$. Unfortunately, I do not have a neat way of calculating this in ASReml. It would not be too hard to calculate in R using the appropriate Legendre polynomials and the coefficients in the .sln file. But almost impossible to get a standard error on the prediction.