REMLF90 Manual

April 2, 1998 - May 9, 2002

REMLF90 is a program to estimate variance components for a variety of models. It can cover the same models as BLUPF90, which is described elsewhere. REMLF90 uses the EM algorithm with acceleration. It is slow but also reliable (except for programming errors).

REMLF90 uses the same parameter file as BLUPF90. The only specifics to REMLF90 is that covariances that cannot be estimated (e.g., there is not a record with two traits observed simultaneously, or one effect is missing from one trait) should be set to 0. (Co)variances that are set to 0 or become 0 are never recomputed again.

Whenever the variance-covariance matrices are not positive-definite, they are made positive definite by increasing values of the smallest eigenvalues.

The program is still in development and output contains many irrelevant lines that are used for problem resolution. In the future, the output will be refined and speed improved. AIREMLF90 by S. Tsuruta, which uses the Average Information rather than the EM algorithm converges sometimes 50 times faster but does not converge for all models.

When priors for the variance components are way off, the convergence criterion may temporarily increase. This does not indicate divergence as the likelihood is increasing all the time. When the computed values are sufficiently close to converged ones, the convergence criterion becomes smaller again.

So far, a few problems have been observed:

1. The estimates may go out of bounds at extrapolation for complicated models. Solution: compile the program with more conservative criterion for extrapolation. This may reduce convergence for simpler problems.

2. If starting values for variances are much smaller that correct final estimates, the convergence is extremely slow, may not occur at all, or it may be to wrong estimates. Solution: use starting variances too large than too small.

3. The above problem seem more important with random regression models, estimates may be incorrect if starting variances are much smaller than true estimates; in such cases each starting value results in different estimates. When starting variances are as large or much larger than the “true” estimates, the convergence is always to the “true” estimates. Starting values of(co)variances do not seem to be critical. This could be more problem of EM than of REMLF90.

4. When variances are very large, residual variances may be estimated as zero. In this case, scale the trait down or decrease the value of working zero (denseop_tol) in denseop.f90.

Example 1

```
[ignacy@nce]$ cat exesa
# Two trait example from Esa Mantysaari's paper in JABG106:409.
DATAFILE
esadat
NUMBER_OF_TRAITS
2
NUMBER_OF_EFFECTS
2
OBSERVATION(S)
3 4
WEIGHT(S)

EFFECTS: POSITIONS_IN_DATAFILE NUMBER_OF_LEVELS TYPE_OF_EFFECT  [EFFECT NESTED]
1 1 3 cross
2 2 3 cross
RANDOM_RESIDUAL VALUES
160 75
75 140
RANDOM_GROUP
2
RANDOM_TYPE
diagonal
FILE

(CO)VARIANCES
12 9
9 10


[ignacy@nce]$ cat esadat
1 1 109 131
1 2 115 140
1 1  99   0
2 1 105 121
2 2 109 130
2 1 120 150
2 2 121 148
2 1  98 132
2 2 110   0
2 1 111 148
2 2 132 155
2 1 117 129
2 2 105   0
2 1 101 122
2 1  87  0
3 2 111   0
3 3 127 151
3 2 132 153
3 3 120 149
3 2 117 141
3 3 129 160
3 2 129 149
3 3 145 159
3 2 125 155
3 3 137 149
3 2 119 137
3 3 139 167
3 3 131 144
3 3 125   0
3 3 111   0

[ignacy@nce]$ remlf90
 name of parameter file?exesa
 Parameter file:             exesa
 Data file:                  esadat
 Number of Traits             2
 Number of Effects            2
 Position of Observations      3  4
 Position of Weight (1)        0
 Value of Missing Trait/Observation          0

EFFECTS
 #  type                position (2)        levels   [positions for nested]
 1  cross-classified     1  1                3
 2  cross-classified     2  2                3
```

```
Residual (co)variance Matrix
  160.000    75.000
   75.000   140.000

Random Effect(s)   2
Type of Random Effect:      diagonal
trait   effect    (CO)VARIANCES
 1       2      12.000     9.000
 2       2       9.000    10.000

REMARKS
 (1) Weight position 0 means no weights utilized
 (2) Effect positions of 0 for some effects and traits means that such
     effects are missing for specified traits

Data record length =   4
original G
 12.000      9.0000
 9.0000      10.000
inverted G
 .25641     -.23077
-.23077      .30769
read   30  records in   3.000000E-02  s,    42  nonzeroes
finished peds in   4.000000E-02  s,    42  nonzeroes
rank=  12
 In round   1  convergence=  7.024953752662561E-004
new r
 112.77      76.818
 76.818      113.43
original G
 12.170      9.3442
 9.3442      10.183
inverted G
 .27810     -.25518
-.25518      .33236
 In round   2  convergence=  6.321248114363380E-003
new r
 106.42      81.331
 81.331      114.71
original G
 13.027      10.342
 10.342      10.850
inverted G
 .31558     -.30080
-.30080      .37888
 In round   3  convergence=  6.741428259295574E-003
new r
 105.29      83.650
 83.650      117.40
original G
 14.041      11.436
 11.436      11.610
inverted G
 .36018     -.35477
-.35477      .43558
.........................
.........................
 In round   91  convergence=  1.010214272952340E-008
new r
 100.92      81.188
 81.188      114.86
original G
 35.260      32.518
 32.518      30.143
inverted G
 5.5188     -5.9536
-5.9536      6.4558
 In round   92  convergence=  9.743411597834022E-009
STOP
```

Example 2

```
[ignacy@nce]$ cat exmm1

# Two trait example from canonical transformation program MTDFS
DATAFILE
mmdat
NUMBER_OF_TRAITS
2
NUMBER_OF_EFFECTS
2
OBSERVATION(S)
3 4
WEIGHT(S)

EFFECTS: POSITIONS_IN_DATAFILE NUMBER_OF_LEVELS TYPE_OF_EFFECT  [EFFECT NESTED]
1 1 1 cross
2 2 5 cross
RANDOM_RESIDUAL VALUES
    181.53422   -224.92828
   -224.92828   1994.85849
RANDOM_GROUP
2
RANDOM_TYPE
add_an_upg
FILE
mmrel
(CO)VARIANCES
   64.55265    268.70723
    268.70723   1168.14129

[ignacy@nce]$ cat mmdat
1 1    -5.31  58.60  -65.16    -3.470      .486     -.352 1
1 1     -.65  31.56  147.92      .394     1.191      .756 2
1 2   -10.91  13.08 -105.54    -3.029     -.612     -.419 1
1 2    15.47 -11.57  120.34     3.731      .843      .427 2
1 2     6.18  76.62  148.69     -.298     2.132      .594 1
1 2    32.77 -27.93  118.57     6.554     1.086      .181 2
1 3    -5.03 -29.80  -47.34     -.115     -.858     -.135 1
1 3     1.30 -29.72    7.02     1.297     -.411      .060 2
1 3   -25.60 -14.43 -108.65    -4.011    -1.513     -.176 1
1 3   -28.90 -47.86 -341.21    -5.769    -3.252    -1.322 2
1 3    -1.27 -26.29  147.47     2.354      .230      .845 1
1 3    -6.75 -85.94 -207.11     -.078    -2.576     -.884 2
1 4     6.82 -44.82   10.35     2.589     -.476      .016 1
1 4   -15.20 125.94  192.79    -4.364     2.502     1.081 2
1 4   -16.24  13.58  -60.09    -3.253     -.549     -.096 1
1 4   -11.93  86.25   58.00    -3.986     1.319      .367 2
1 4    18.13  11.71  160.25     3.687     1.489      .568 1
1 4    -8.14 -22.23  -27.12     -.571     -.733      .009 2
1 4   -14.50  16.60  -69.41    -3.233     -.492     -.177 1
1 4     3.09  32.86  201.49     1.415     1.577      .984 1

 [ignacy@nce]$ cat mmrel
1 5 3 2
2 5 5 3
3 5 5 3
4 5 5 3
[ignacy@nce]$ remlf90
 name of parameter file?exmm1

 Parameter file:             exmm1
 Data file:                  mmdat
 Number of Traits             2
 Number of Effects            2
 Position of Observations      3  4
 Position of Weight (1)        0
 Value of Missing Trait/Observation          0

EFFECTS
 #  type                position (2)        levels   [positions for nested]
 1  cross-classified     1  1                1
 2  cross-classified     2  2                5

 Residual (co)variance Matrix
   181.534  -224.928
  -224.928  1994.858
```

```
 Random Effect(s)   2
 Type of Random Effect:      additive animal with unknown parent groups
 Pedigree File:              mmrel
 trait   effect    (CO)VARIANCES
  1       2      64.553   268.707
  2       2     268.707  1168.141

 REMARKS
  (1) Weight position 0 means no weights utilized
  (2) Effect positions of 0 for some effects and traits means that such
      effects are missing for specified traits

 Data record length =   4
 original G
  64.553      268.71
  268.71      1168.1
 inverted G
  .36470    -0.83891E-01
-0.83891E-01 0.20154E-01
 read   20  records in   5.000000E-02  s,    31  nonzeroes
  read   4  additive pedigrees
 finished peds in   .100000  s,    54  nonzeroes
 rank=  10
  In round   1  convergence=  2.043287539467460E-016
 STOP
```

In this example, the final estimates were put into the parameter file and therefore convergence occurred in one round.