

Specialised analyses with WOMBAT

Karin Meyer

May 27, 2009

Specialised analyses with WOMBAT

This page provides some hints for non-standard analyses.

Fitting a factor-analytic model

WOMBAT can accommodate a factor-analytic model for the structure of covariance matrices of random effects. While there is no explicit option for this structure, it is easily fitted indirectly.

Factor analysis implies a latent model which models a random effect (\mathbf{r}) for a set of q traits as the sum of a vector of m common effects (\mathbf{c}), weighted by a matrix of factor loadings (\mathbf{F}), and a vector of q specific effects (\mathbf{s}). i.e. $\mathbf{r} = \mathbf{F} \mathbf{c} + \mathbf{s}$. The common effects are assumed to be *iid* distributed with variances of unity, while the specific effects are assumed to be uncorrelated but have heterogeneous variances. This gives $\text{Var}(\mathbf{r}) = \mathbf{F} \mathbf{F}' + \mathbf{D}$, with \mathbf{D} the the diagonal matrix of specific variances.

WOMBAT allows for the first part, i.e. has an option which allows $\text{Var}(\mathbf{r})$ to be estimated as $\mathbf{F} \mathbf{F}'$ with reduced rank m . In addition, WOMBAT incorporates an option for diagonal covariance matrices. Hence a factor-analytic model is readily fitted by simply fitting the corresponding random effect twice: In the first instance with reduced rank covariance matrix, representing the common factors, and in the second instance with diagonal covariance matrix, representing the specific effects.

WOMBAT only allows a particular column in the data file to be used once, i.e. to be mapped to a single effect in the model of analysis. Hence, fitting a factor-analytic structure for a particular effect requires the corresponding code (column) in the data file to be replicated ! If WOMBAT encounters two random effects, listed successively in the .par file, with the names only differing by the first letter (e.g. animal and bnimal in Example 5) and with reduced rank and diagonal covariance matrices, respectively, it will assume that a factor-analytic model has been intended and automatically write out summary statistics for the sum of the two matrices.

```
COM Example parameter file for factor analytic model
```

```
ANAL MUV PC 6
```

```
PEDS ../carcped.dat
```

```
DATA ../carcsix.dat GRP
```

```
TRNOS 1 2 3 4 5 6
```

```
traitno 6
```

```
animal 0
```

```
bnimal 0
```

```
cg 299
```

```
NAMES cwt cema cimf crby cp8 crib
```

```
END
```

```
MODEL
```

```
FIX cg
```

```
RAN animal nrm
```

```
RAN bnimal nrm
```

```
tr cwt 1
```

```
tr cema 2
```

```
tr cimf 3
```

```
tr crby 4
```

```

tr cp8 5
tr crib 6
END MOD

VAR animal 6 1
177 29 3.1 -11 15 8
9.8 0.03 -0.2 1.2 1.6
0.91 -0.76 0.9 0.3
3.62 -2.9 -1.96
4.95 2.2
1.95

VAR banimal 6 6 DIAG
50 3 0.3 0.5 1.1 0.8
VAR residual 6 6
292 39 3.2 -5.9 22.4 11.
36 0.8 -0.9 4.7 1.2
1 -0.06 0.15 0.32
0.73 -0.55 -0.45
6.9 2.0
5.1

```

Dealing with 'clones'

In some instances we have clones, i.e. subjects which are genetically identical. For instance, monozygotic twins are naturally occurring clones. If these individuals have different identity codes and the same parents, WOMBAT will simply treat them as ordinary full-sibs. Alternatively, if all members of the clone are given the same identity, records on all members of the clone are treated as if they were repeated observations for the same individual. That is not correct either - hence special action needs to be taken.

To model clones correctly, we need to fit the **same genetic** effects for all members of the clone, but allow for **different environmental** effects for different individuals. WOMBAT has no in-built option to do this automatically, but can easily be tricked to do so:

- Give all members of the clone the same genetic identity.
- For data with repeated records per trait, assign a different code for permanent environmental effects to each clone member.
- WOMBAT requires the data file to be sorted according to traits within individual. However, it does not require identities to be in ascending or descending order the program simply checks for a change in identity to determine where the records for a new individual start. Hence, the 'trick' is to arrange the data file so that members of the same clone do **not** occur successively (while still keeping records taken on the same individual together and, for multivariate analyses, sorted in ascending trait number). This will cause WOMBAT to assign different residual effects to different individuals and thus model the genetic and environmental covariance structure correctly.

Example: Individuals "1","2" and "3" belong to a clone given the identity "100",
"4", "5" and "6" do not belong to a clone

Trait	Genetic_Code	Record	Individual
1	100	10	1
2	100	99	1
1	4	12	4
2	4	88	4
1	100	9	2
2	100	94	2
1	5	13	5
2	5	87	5
1	100	11	3
2	100	77	3

1 6 7 6

WOMBAT is likely to refuse running your analysis on grounds that there are too few "repeated" records per individual or that the data file is not sorted correctly. To override these 'safety catches', you need to add the special option "clones", i.e. your parameter file should have a block

```
SPECIAL
  CLONES
END SPECIAL
```

Alternatively, if animals are almost but not 100% identical, the inverse of the numerator relationship matrix can be constructed as described by Oikawa & Yasuda (2009) and fed into WOMBAT as a *.giv file.

Semiparametric regression

<http://agbu.une.edu.au/~kmeyer/dokuwiki/>