

Package ‘mlmmm’

October 13, 2022

Version 0.3-1.2

Date 2010-07-07

Title ML estimation under multivariate linear mixed models with missing values

Author Recai Yucel <ryucel@albany.edu>.

Maintainer Recai Yucel <ryucel@albany.edu>

Description Computational strategies for multivariate linear mixed-effects models with missing values, Schafer and Yucel (2002), Journal of Computational and Graphical Statistics, 11, 421-442.

License GPL-2

URL <http://www.albany.edu/~ry491374>

Repository CRAN

Date/Publication 2010-07-07 06:09:45

NeedsCompilation yes

R topics documented:

adg	2
mlmmm.em	2
mlmmmbd.em	5
pain	6
Index	7

adg	<i>Average Daily Gain Data</i>
-----	--------------------------------

Description

Data example used here come from "SAS System for Mixed Models" by Little et. al. (1996) and it is an example data within SASmixed library. This dataset was collected to investigate average daily gains (ADG) of steers fed for 160 days. The treatments are four diets consisting of a base ration and three levels of a medicated feed additive added to the base ration. The objective of the experiment was to determine the optimal level of feed additive to maximize the average daily gain. The steers were housed in barns, the blocking factor, where each barn held four steers and the steers were individually fed. For the purposes of illustration we imposed missing values on the response values, following a missing completely at random mechanism.

Usage

```
data(adg)
```

Format

A data frame containing 32 rows and 7 columns(response variables: first two columns which are average daily gain in grams, initial weight on the log scale ; predictors next 7 rows (a column of 1 and 3 treatment dummies) and subject indicator is last column)

Source

Littell, R.C., Milliken, G.A., Stroup, W.W., and Wolfinger, R.D. (1996). SAS System for Mixed Models, SAS Institute, Cary, NC.

mlmmm.em	<i>ML estimation via EM-algorithm under multivariate linear mixed models with missing values</i>
----------	--

Description

ML estimation via hybrid of EM and Fisher scoring algorithm under the multivariate linear mixed models with missing values described by Schafer and Yucel (2002), Yucel (2007). This function will typically be used to produce maximum likelihood estimation of the unknown parameters under the model

$$y_i = X_i \beta + Z_i b_i + e_i, i=1, \dots, m,$$

where

y_i = ($n_i \times r$) matrix of incomplete multivariate data for subject or cluster i ;

X_i = ($n_i \times p$) matrix of covariates;

Z_i = ($n_i \times q$) matrix of covariates;

β = ($p \times r$) matrix of coefficients common to the population (fixed effects);

b_i = ($q \times r$) matrix of coefficients specific to subject or cluster i (random effects); and

e_i = ($n_i \times r$) matrix of residual errors.

The matrix b_i , when stacked into a single column, is assumed to be normally distributed with mean zero and unstructured covariance matrix ψ , and the rows of e_i are assumed to be independently normal with mean zero and unstructured covariance matrix σ . Missing values may appear in y_i in any pattern.

In most applications of this model, the first columns of X_i and Z_i will be constant (one) and Z_i will contain a subset of the columns of X_i .

Usage

```
mlmmm.em(y, subj, pred, xcol, zcol, start, maxits=200, eps=0.0001)
```

Arguments

<code>y</code>	matrix of responses. This is simply the individual y_i matrices stacked upon one another. Each column of y corresponds to a response variable. Each row of y corresponds to a single subject-occasion, or to a single subject within a cluster. Missing values (NA) may occur in any pattern.
<code>subj</code>	vector of length <code>nrow(y)</code> giving the subject (or cluster) indicators i for the rows of y . For example, suppose that y is in fact <code>rbind(y1,y2,y3,y4)</code> where <code>nrow(y1)=2</code> , <code>nrow(y2)=3</code> , <code>nrow(y3)=2</code> , and <code>nrow(y4)=7</code> . Then <code>subj</code> should be <code>c(1,1,2,2,2,3,3,4,4,4,4,4,4,4)</code> .
<code>pred</code>	matrix of covariates used to predict y . This should have the same number of rows as y . The first column will typically be constant (one), and the remaining columns correspond to other variables appearing in X_i and Z_i .
<code>xcol</code>	vector of integers indicating which columns of <code>pred</code> will be used in X_i . That is, <code>pred[,xcol]</code> is the X_i matrices (stacked upon one another).
<code>zcol</code>	vector of integers indicating which columns of <code>pred</code> will be used in Z_i . That is, <code>pred[,zcol]</code> is the Z_i matrices (stacked upon one another).
<code>start</code>	optional list of quantities to specify the initial estimates of the parameters for the EM. If "start" is omitted then <code>mlmmm.em()</code> chooses its own initial values.
<code>maxits</code>	maximum number of cycles of EM to be performed. The algorithm runs to convergence or until "maxits" iterations, whichever comes first.
<code>eps</code>	convergence criterion. The algorithm is considered to have converged if the relative differences in all parameters from one iteration to the next are less than <code>eps</code> —that is, if <code>all(abs(new-old)<eps*abs(old))</code> .

Details

The EM algorithm used in `mlmmm.pan()` is described in detail by Schafer and Yucel (2002) and Yucel (2007).

Value

A list containing the following elements:

beta	A matrix containing the final value of the estimate of the fixed effects. The first column corresponds to the estimates for the first column of y , the second column corresponds to the estimates of the second column of y , and so on.
Sigma	A matrix containing the final value of the estimate of the variance covariance matrix of the vectorized residual matrix term.
Psi	A matrix containing the final value of the estimate of the variance covariance matrix of the (vectorized) random-effects matrix.
eb	A matrix (of dimensions $r \times q$ by m) containing the empirical bayes estimates of the random-effects b_i .
varb	An array of dimensions $r \times q \times r \times r \times m$, containing the variance covariance matrix of the random-effects.
xtwxinv	Variance-covariance matrix of the estimate of fixed estimates.
converged	An indicator showing whether the algorithm converged or not.
iter	Number of iterations to convergence.
npatt	Number of distinct missingness patterns, not counting the ones missing all variables making the response matrix.
pstfin	A matrix of dimensions npatt by r , indicating the number of rows with the underlying missingness pattern.
iposn	A vector showing the row numbers of y , which belong to missingness patterns showed in pstfin.
patt	A vector of n denoting the missingness patterns of the rows of y .
rmat	A matrix showing the distinct missingness patterns, excluding the rows that are completely missing.
logll	A vector of expected loglikelihood values at each iteration.
logoll	A vector of observed loglikelihood values at each iteration.
clock	How much time (in seconds) mlmmm.em took to converge.

Author(s)

Recai M. Yucel, Division of Biostatistics and Epidemiology, University of Massachusetts-Amherst <yucel@schoolph.umass.edu>.

References

- Schafer, J.L. and Yucel, R.M. (2002) Computational strategies for multivariate linear mixed-effects models with missing values. *Journal of the Computational and Graphical Statistics*, Volume 11, Number 2, 437–457.
- Yucel, R.M. (2007) R mlmmm package: Fitting multivariate linear mixed-effects models with missing values

Examples

```
## Not run:
# For a detailed example, see the file "mlmmmmex.s" distributed
# with this function. Here is a simple example of how mlmmm.em()
# might be used to produce Ml estimates.
library(mlmmm)
data(adg)
y<-cbind(adg$y.1,adg$y.2)
colnames(y)=c("adg","initwt")
subj=adg$subj
# see the relationship between avd and intwt which are jointly modeled
library(lattice)
xyplot(y[,1]~log(y[,2]) | subj, ylab="Average Daily Gain",xlab="Initial Weight")
# below adg$subj is the block or barn
subj<-adg$subj
pred <- cbind(adg$pred.int,adg$pred.dummy1,adg$pred.dummy2,adg$pred.dummy3)
xcol<-1:4
zcol<-1
unst.psi.result <- mlmmm.em(y,subj,pred,xcol,zcol,maxits=200,eps=0.0001)

## End(Not run)
```

mlmmmbd.em

ML estimation under multivariate linear mixed models with block-diagonal covariance matrix and missing values

Description

Implementation of em.pan() that restricts the covariance matrix for the random effects to be block-diagonal. This function is identical to pan() in every way except that psi is now characterized by a set of r matrices of dimension $q \times q$.

Usage

```
mlmmmbd.em(y, subj, pred, xcol, zcol, start, maxits=100, eps=0.01)
```

Arguments

y	See description for mlmmm.em().
subj	See description for mlmmm.em().
pred	See description for mlmmm.em().
xcol	See description for mlmmm.em().
zcol	See description for mlmmm.em().
start	Same as for em.pan() except that the starting value for psi have new dimensions: $(q \times q \times r)$
maxits	See description for mlmmm.em().
eps	See description for mlmmm.em().

Value

A list with the same components as that from `em.pan()`, with a minor difference: the dimension of "psi" is now $(q \times q \times r)$.

References

Schafer, J.L. and Yucel, R.M. (2002) Computational strategies for multivariate linear mixed-effects models with missing values, *Journal of Computational and Graphical Statistics*, 11, 421-442.

pain

The longitudinal study on pain in children

Description

This data consist of up to four observations on 64 children 8 to 10. The response is the length of time in seconds that the child can tolerate keeping his or her arm in very cold water, a proxy measure of pain tolerance. Missing data exist due to absenteeism, broken arms, or other reasons. Two measures were taken during a first visit followed by two more measures during a second visit after a two-week gap.

Usage

```
data(pain)
```

Format

A data frame containing 256 rows and 10 columns(response variables: first two rows; predictors next 7 rows and subject indicator is last column)

Source

Weiss, R.E. (2005) *Modeling Longitudinal Data*. New York, Springer.

Index

* **datasets**

adg, [2](#)

pain, [6](#)

* **models**

mlmm.em, [2](#)

mlmmbd.em, [5](#)

adg, [2](#)

mlmm.em, [2](#)

mlmmbd.em, [5](#)

pain, [6](#)