Package: VBJM (via r-universe)

October 5, 2024

Type Package

Title Variational Inference for Joint Model

Version 0.1.0

Author Jiehuan Sun [aut, cre]

Maintainer Jiehuan Sun <jiehuan.sun@gmail.com>

Description The shared random effects joint model is one of the most
widely used approaches to study the associations between
longitudinal biomarkers and a survival outcome and make dynamic
risk predictions using the longitudinally measured biomarkers.
One major limitation of joint models is that they could be
computationally expensive for complex models where the number
of the shared random effects is large. This package can be used
to fit complex multivariate joint models using our newly
developed algorithm Jieqi Tu and Jiehuan Sun (2023)
<doi:10.1002/sim.9619>, which is based on Gaussian variational
approximate inference and is computationally efficient.

Depends R (>= 3.6.0)

Imports Rcpp (>= 1.0.0), survival(>= 3.2), statmod(>= 1.4), pracma(>=
2.2), Matrix(>= 1.3)

LinkingTo Rcpp, RcppArmadillo, RcppEnsmallen

License GPL-2

Encoding UTF-8

LazyData true

RoxygenNote 7.2.1

NeedsCompilation yes

Repository CRAN

Date/Publication 2023-09-02 08:20:02 UTC

Contents

---

control_list   *control_list*

Description

This list contains a list of parameters specifying the joint model.

Details

- ID_name the variable name for the patient ID in both longitudinal data and survival data.

- item_name the variable name for the longitudinal outcomes in the longitudinal data.

- value_name the variable name for the longitudinal measurements in the longitudinal data.

- time_name the variable name for the measurement timepoints in the longitudinal data.

- fix_cov a set of variables names indicating the covariates of fixed-effects in the longitudinal submodel. If NULL, not baseline covariates are included.

- random_cov a set of variables names indicating the covariates of random-effects in the longitudinal submodel. If NULL, not baseline covariates are included.

- FUN a function specifying the time-related basis functions in the longitudinal submodel.

- ran_time_ind a vector of integers specifying which time-related basis functions are also included with random-effects in the longitudinal submodel.

- surv_time_name the variable name for the survival time in the survival data.

- surv_status_name the variable name for the censoring indicator in the survival data.

- surv_cov a set of variables names specifying the baseline covariates in the survival submodel.

- n_points an integer indicating the numebr of nodes being used in the Gaussian quadrature.

Author(s)

Jiehuan Sun <jiehuan.sun@gmail.com>

LongData   *Simulated Longtidunal Data*

Description

This dataset contains longitudinal outcomes.

Usage

```
data(VBJMdata)
```

Format

A data frame with 48700 rows and 4 variables

Details

- ID patient ID
- item types of longitudinal outcome
- years measurement timepoints
- value measurements

Author(s)

Jiehuan Sun <jiehuan.sun@gmail.com>

SurvData   *Simulated Survival Data*

Description

This dataset contains survival outcome.

Usage

```
data(VBJMdata)
```

Format

A data frame with 100 rows and 4 variables

Details

- ID patient ID

- fstat censoring indicator

- ftime survival time

- x baseline covariates

Author(s)

Jiehuan Sun <jiehuan.sun@gmail.com>

VBJM_fit   *The function to fit VBJM.*

Description

The function is used to fit joint models using variational inference algorithm.

Usage

```
VBJM_fit(
  LongData = NULL,
  SurvData = NULL,
  marker.name = NULL,
  control_list = NULL,
  maxiter = 100,
  eps = 1e-04
)
```

Arguments

LongData	a data frame containing the longitudinal data (see [LongData](#)).
SurvData	a data frame containing the survival data (see [SurvData](#)).
marker.name	a vector indicating which set of longitudinal biomarkers to be analyzed. If NULL, all biomarkers in LongData will be used.
control_list	a list of parameters specifying the joint model (see [control_list](#)).
maxiter	the maximum number of iterations.
eps	threshold for convergence.

Value

return a data frame with estimates, standard errors, and 95% CIs for each of the following parameters, where VAR indicates the corresponding variable name.

VAR_alpha	the parameters for the effects of biomarkers in the survival submodel, where VAR indicates the names for the biomarkers.
Weibull_shape	the shape parameter in the Weibull baseline hazard in the survival submodel.
Weibull_scale	the scale parameter in the Weibull baseline hazard in the survival submodel.
Surv_gamma_VAR	the parameters for the effects of baseline covariates in the survival submodel.
VAR_fix	the parameters for the fixed-effects in the longitudinal submodel.

References

Jieqi Tu and Jiehuan Sun (2023). "Gaussian variational approximate inference for joint models of longitudinal biomarkers and a survival outcome". Statistics in Medicine, 42(3), 316-330.

Examples

```
data(VBJMdata)
flex_time_fun <- function(x=NULL){
    xx = matrix(x, ncol = 1)
    colnames(xx) = c("year_l")
    xx
}
ran_time_ind = 1 ## random time-trend effects
control_list = list(
  ID_name = "ID", item_name = "item",
  value_name = "value",  time_name = "years",
  fix_cov = NULL, random_cov = NULL,
  FUN = flex_time_fun, ran_time_ind=ran_time_ind,
  surv_time_name = "ftime",  surv_status_name = "fstat",
  surv_cov = "x", n_points = 5
)

## takes about one minute.
res = VBJM_fit(LongData=LongData, SurvData=SurvData,
               control_list=control_list)
```

Index