Package: growthrate (via r-universe)

August 30, 2024

Type Package

Version 1.3

Date 2014-08-13

Title Bayesian reconstruction of growth velocity

Author Sara Lopez-Pintado and Ian W. McKeague

Maintainer Ian W. McKeague <im2131@columbia.edu>

Description A nonparametric empirical Bayes method for recovering gradients (or growth velocities) from observations of smooth functions (e.g., growth curves) at isolated time points.

Depends Matrix, clime, mvtnorm

License GPL-3 | file LICENSE

NeedsCompilation no

Repository CRAN

Date/Publication 2014-08-13 18:59:22

Contents

growthrate-package — *Recovery of gradients from sparsely observed functional data*

Description

A nonparametric empirical Bayes method for recovering gradients (or growth velocities) from observations of smooth functions (e.g., growth curves) at isolated time points.

References

Lopez-Pintado, S. and McKeague, I. W. (2013). *Recovering gradients from sparsely observed functional data.* Biometrics 69, 396-404 (2013). http://www.columbia.edu/~im2131/ps/growthrate-package-reference.pdf

cv.growth — *cross validation error*

Description

Computes the cross validation error resulting from the removal of the data at a given interior observation time as a function of the infinitessimal standard deviation `sigma` on a grid of `k` equispaced values in the interval [`a`, `b`].

Usage

```
cv.growth(data, tobs, d, K, a, b, r)
```

Arguments

`data`	Input matrix of size N (subjects) times n (observation times). Each column contains the heights (of all subjects) at a given observation time, each row contains the heights (at the observation times) for a given subject.
`tobs`	Row vector of n observation times (in increasing order, same for each subject).
`d`	Number of points on the fine time-grid (between the first and last observation times in `tobs`) at which the posterior means and covariances are computed.
`K`	Number of points on the grid for `sigma`.
`a`	Minimum value for `sigma` (positive).
`b`	Maximum value for `sigma` (strictly larger than `a`).
`r`	Index of the interior observation time in `tobs` removed in the cross-validation.

Details

The data for the `r`th observation time (for a given $1 < \texttt{r} < n$) are removed and the mean squared error of the reconstructed data at that time point computed over the grid for `sigma`.

Value

`sigmavec`	`K`-vector of `sigma`-values considered in the cross validation.
`CVer`	Cross validation error at each value of `sigma` in `sigmavec`.

Author(s)

Sara Lopez-Pintado and Ian W. McKeague

Maintainer: Ian W. McKeague <im2131@columbia.edu>

References

Lopez-Pintado, S. and McKeague, I. W. (2013). *Recovering gradients from sparsely observed functional data.* Biometrics 69, 396-404 (2013). http://www.columbia.edu/~im2131/ps/growthrate-package-reference.pdf

Examples

```
## Not run:
    ## example using the height data provided in the package
    ## there are 7 observation times (age in years):
    ## WARNING: cv.growth is time-consuming. This example uses only part of the data.
    data(height_data);
    ht = height_data[1:100,];
    tobs=c(0,1/3,2/3,1,3,4,7);
    cvg=cv.growth(ht, tobs, 100, 21, 1, 5, 2);

    ## Plot of the cross validation error as a function of sigma:
    plot(cvg$sigmavec, cvg$CVer, xlab="Sigma", ylab="Cross validation error");

    ## Value of sigma that minimizes the cross validation error:
    sigmaopt=cvg$sigmavec[which(cvg$CVer==min(cvg$CVer))];

## End(Not run)
```

`diffquo` *converts height data into difference quotients*

Description

Internal function. Calculates the crude one-sided difference quotients and the second-order difference quotients based on the height data and the observed time points.

Usage

```
diffquo(data, tobs)
```

Arguments

`data`	As in `growth`.
`tobs`	As in `growth`.

Value

`YI`	An N times (n-1) matrix with rows given by the values of the one-sided difference quotients $y_i, i = 1 \ldots, n-1$ for the N subjects.
`Xtilda`	An n by N matrix with columns given by the values of the second-order difference quotients **y** for the N subjects.

`growth` *recovery of growth velocities*

Description

Computes the mean function and covariance kernel (over a fine grid of equispaced time points) of the posterior growth velocity for each subject, based on growth data (e.g., heights) at fixed observation times.

Usage

```
growth(data, tobs, sigma, d)
```

Arguments

`data`	Input matrix of size N (subjects) times n (observation times). Each column contains the heights (of all subjects) at a given observation time, each row contains the heights (at the observation times) for a given subject.
`tobs`	Row vector of n observation times (in increasing order, same for each subject).
`sigma`	A positive scalar representing the infinitessimal standard deviation of the tied-down Brownian motion in the prior. Can be selected by cross-validation.
`d`	Number of time points on the fine grid.

Details

The Bayesian reconstruction implemented here uses a prior growth velocity model that is specified by a general multivariate normal distribution at the n fixed observation times, and a tied-down Brownian motion (having infinitessimal standard deviation specified by `sigma`) between the observation times.

The prior mean and prior precision matrix at the observation times are estimated using the data on N subjects. Clime (constrained L1 minimization) provides the estimate of the prior precision matrix, with the clime constraint parameter lambda selected by 5-fold cross validation using the likelihood loss function.

Value

muhatcurve
: Posterior means of the growth velocities (for each subject) on the fine grid `tgrid`. An N (subjects) times `d` matrix.

Khat
: Posterior covariance kernel of the growth velocities on the fine grid `tgrid`. A `d` times `d` matrix (the same for every subject).

tgrid
: The fine grid of `d` equispaced-time points (between the first and last observation times in `tobs`) at which the posterior means and covariances are computed.

Author(s)

Sara Lopez-Pintado and Ian W. McKeague

Maintainer: Ian W. McKeague <im2131@columbia.edu>

References

Lopez-Pintado, S. and McKeague, I. W. (2013). *Recovering gradients from sparsely observed functional data.* Biometrics 69, 396-404 (2013). http://www.columbia.edu/~im2131/ps/growthrate-package-reference.pdf

Examples

```
## Not run:
## example using the height data provided in the package
## there are 7 observation times (age in years):
 data(height_data);
 tobs=c(0,1/3,2/3,1,3,4,7);
 d=200;
 sigma=1;
 g=growth(height_data,tobs,sigma,d);

## Plot of the posterior mean and credible interval for a specific individual
 indiv=1;
## posterior standard deviation (same for all subjects):
 postsd=sqrt(diag(g$Khat));
 plot(g$tgrid,g$muhatcurve[indiv,],type='l',
   xlab="Age (years)",ylab="Growth velocity (cms/year)");
 lines(g$tgrid,g$muhatcurve[indiv,]);
 lines(g$tgrid,g$muhatcurve[indiv,]+2*postsd,lty=2);
 lines(g$tgrid,g$muhatcurve[indiv,]-2*postsd,lty=2);

## Plot of a draw from the posterior growth velocity for a specific individual:
 draw=rmvnorm(n=1, mean=g$muhatcurve[indiv,], sigma=g$Khat, method="chol");
 plot(g$tgrid,draw,type='l',xlab="Age (years)",ylab="Growth
 velocity (cms/year)");

## End(Not run)
```

`height_data` *height data*

Description

The heights of 532 girls at birth, four, eight, and twelve months, and three, four, and seven years of age.

Usage

```
data(height_data)
```

Format

A matrix with 532 observations on 7 variables (height at each age).

References

Lopez-Pintado, S. and McKeague, I. W. (2013). *Recovering gradients from sparsely observed functional data.* Biometrics 69, 396-404 (2013). http://www.columbia.edu/~im2131/ps/growthrate-package-reference.pdf

Examples

```
data(height_data)
```

`Khatf` *posterior covariance kernel*

Description

Internal function. Finds the posterior covariance kernel on the fine grid `tgrid` (same for every subject); see Section 2 of the referenced article for further details.

Usage

```
Khatf(tgrid, tobs, sigma, Sigmahat)
```

Arguments

`tgrid`	The fine grid of `d` equispaced-time points (between the first and last observation times in `tobs`) at which the posterior means and covariances are computed.
`tobs`	Row vector of `n` observation times (in increasing order, same for each subject).
`sigma`	Infinitessimal standard deviation of the tied-down Brownian motion in the prior.
`Sigmahat`	Posterior covariance at the observation times `tobs` (same for every subject).

`new.growth` *recovery of growth velocity for a new subject*

Description

Computes the posterior mean and covariance kernel for a new subject having data at observation times `newtobs` different from `tobs` (apart from the first and the last). `growth` needs to be run first.

Usage

```
new.growth(newdata, newtobs, sigma, d, muhatcurve, Khat, tgrid)
```

Arguments

`newdata`	Row vector of p heights for the new subject.
`newtobs`	Row vector of p observation times for the new subject (in increasing order; must include the first and last time points in `tobs`).
`sigma`	Infinitessimal standard deviation of the Brownian motion prior (same as in `growth`).
`d`	Number of time points on the fine grid.
`muhatcurve`	Output from `growth`.
`Khat`	Output from `growth`.
`tgrid`	The fine grid (output from `growth`).

Value

`muhatcurvenew`	Posterior mean (on `tgrid`) for the new subject.
`Khatnew`	Posterior covariance kernel (on `tgrid`) for the new subject.

Author(s)

Sara Lopez-Pintado and Ian W. McKeague

Maintainer: Ian W. McKeague <im2131@columbia.edu>

References

Lopez-Pintado, S. and McKeague, I. W. (2013). *Recovering gradients from sparsely observed functional data.* Biometrics 69, 396-404 (2013). http://www.columbia.edu/~im2131/ps/growthrate-package-reference.pdf

Examples

```
## Not run:
## example using the height data provided in the package
##    (after first running growth to obtain the output g):
## suppose a new subject has 5 observation times (including 0 and 7)
 data(height_data);
 tobs=c(0,1/3,2/3,1,3,4,7);
 d=200;
 sigma=1;
 g=growth(height_data,tobs,sigma,d);

 newtobs=c(0, 2, 3, 5, 7);
 newdata=t(as.vector(c(50,70,87,100,115)));
 ng=new.growth(newdata,newtobs,sigma,d,g$muhatcurve,g$Khat,g$tgrid);

## plot of the posterior mean growth velocity for the new subject:
 plot(g$tgrid,ng$muhatcurvenew,xlab="Age (years)",ylab="Growth
 velocity (cms/year)");

## End(Not run)
```

`posteriordistribcurve` *posterior mean*

Description

Internal function. Finds the posterior mean on the fine grid `tgrid` for every subject.

Usage

```
posteriordistribcurve(muhatmatrix, Sigmahat, sigma, tobs, d, YI)
```

Arguments

`muhatmatrix`	Posterior mean at the observation times `tobs` obtained from `posteriorobs`.
`Sigmahat`	Posterior covariance at the observation times `tobs` obtained from `posteriorobs`.
`sigma`	Infinitessimal standard deviation of the tied-down Brownian motion in the prior.
`tobs`	Row vector of n observation times (in increasing order, same for each subject).
`d`	Number of time points on the fine grid.
`YI`	An N times (n-1) matrix with rows given by the values of $y_i, i = 1 \ldots, n-1$ for the N subjects.

`posteriorobs` *posterior mean and covariance at observation times*

Description

Internal function. Finds the posterior mean and covariance at the observation times `tobs`, for every subject; see Section 2 of the referenced article for further details.

Usage

```
posteriorobs(Sigma0inv, sigma, muprior, Xtilda, tobs, YI)
```

Arguments

`Sigma0inv`	Prior precision matrix at the observation times `tobs`.
`sigma`	Infinitessimal standard deviation of the tied-down Brownian motion in the prior.
`muprior`	Prior mean at the observation times `tobs`.
`Xtilda`	An n by N matrix with columns given by the values of **y** for the N subjects.
`tobs`	Row vector of n observation times (in increasing order, same for each subject).
`YI`	An N times (n-1) matrix with rows given by the values of $y_i, i = 1\ldots, n-1$ for the N subjects.

`priormeanobs` *prior mean*

Description

Internal function. Finds the sample mean of the n-vector **y** for use in specifying the prior mean $\boldsymbol{\mu}_0$; see Section 2 of the referenced article for further details.

Usage

```
priormeanobs(YI, tobs)
```

Arguments

`YI`	The N times (n-1) matrix with rows being the values of $y_i, i = 1\ldots, n-1$ for the N subjects
`tobs`	Row vector of n observation times (in increasing order, same for each subject).

Index