The AdaptFit Package

February 16, 2008

Version 0.2-1

Title Adaptive Semiparametic Regression

Author Tatyana Krivobokova <Tatyana.Krivobokova@econ.kuleuven.be>

Maintainer Tatyana Krivobokova <Tatyana.Krivobokova@econ.kuleuven.be>

Depends SemiPar, MASS, nlme, cluster

Description Based on the function "spm" of the SemiPar package fits semiparametric regression models with spatially adaptive penalized splines.

License GPL (version 2 or later)

R topics documented:

`AdaptFit-internal` *Internal AdaptFit objects*

Description

Internal functions

Details

These are not to be called by the user

`absent` *Absenteeism data*

Description

The `absent` contains data on the absenteeism of workers of a medium-sized industrial company in southern Germany observed between 01.01.1981 and 31.12.1998.

Usage

```
data(absent)
```

Format

This data frame contains the following columns:

id employee's ID

absent number of days of absenteeism

status 1 for the uncensored last day of absenteeism

date date of the first day of a sick leave

day day of the week of the first day of a sick leave

month month

Source

Kauermann, G and Ortlieb, R. (2004). Temporal pattern in number of staff on sick leave: the effect of downsizing. *Journal of Royal Statistical Society, Series C*, **53**, 353-367.

References

Krivobokova, T., Crainiceanu, C.M. and Kauermann, G. (2007)
Fast Adaptive Penalized Splines
Journal of Computational and Graphical Statistics.

Examples

```
library(AdaptFit)
data(absent)
attach(absent)
```

`asp` *Fit a semiparametric regression model with spatially adaptive penalized splines*

Description

`asp` fits semiparametric regression models using the mixed model representation of penalized splines with spatially adaptive penalties, based on the "spm" function of the SemiPar-package.

Usage

```
asp(form,adap=TRUE,random=NULL,group=NULL,family="gaussian",
                spar.method="REML",omit.missing=NULL,niter=20,niter.var=50,tol=1e-0
```

Arguments

`form`	a formula describing the model to be fit. Note, that an intercept is always included, whether given in the formula or not.
`adap`	TRUE (default) if an adaptive fit should be performed, otherwise the fit is identical to that of function "spm".
`random`	"random= 1" specifies inclusion of a random intercept according to the groups specified by the "group" argument.
`group`	a vector of labels for specifying groups.
`family`	for specification of the type of likelihood model assumed in the fitting. May be "gaussian","binomial" or "poisson".
`spar.method`	method for automatic smoothing parameter selection. May be "REML" (restricted maximum likelihood) or "ML" (maximum likelihood).
`omit.missing`	a logical value indicating whether fields with missing values are to be omitted.
`niter`	a maximum number of iterations for the mean estimation, default is 20.
`niter.var`	a maximum number of iterations for the variance of random effects estimation, default is 50.
`tol`	tolerance for the convergence criterion. Default is 1e-6.
`returnFit`	a logical value indicating whether the fitted object should be returned when the maximum number of iterations is reached without convergence of the algorithm. Default is FALSE.
`weights`	to use only with grouped binary data.
`correlation`	correlation structure of the response; see documentation to "nlme".
`control`	see lmeControl in the documentation to "nlme".

Details

See the SemiPar Users' Manual for details and examples.

Value

A list object of class "spm" containing the fitted model. The components are:

fitted	fitted values.
coef.mean	estimated mean coefficients.
design.matrices	design matrices both for knots und subknots.
x	x values.
knots	knots.
y.cov	estimated covariance matrix of the response.
random.var	estimated covariance matrix of the random effects.
subknots	subknots.
coef.random	estimated spline coefficients of the covariance matrix of the random effects.
var.random.var	estimated variance of the spline coefficients of the covariance matrix of the random effects.
fit	mimics fit object of lme() for family="gaussian" and glmmPQL() for family="binomial" or family="poisson".
info	information about the inputs.
aux	auxiliary information such as variability estimates.

Author(s)

Tatyana Krivobokova ⟨Tatyana.Krivobokova@econ.kuleuven.be⟩

References

Krivobokova, T., Crainiceanu, C.M. and Kauermann, G. (2007)
Fast Adaptive Penalized Splines. *Journal of Computational and Graphical Statistics.*

Ganguli, B. and Wand, M.P. (2005)
SemiPar 1.0 Users' Manual.
http://www.maths.unsw.edu.au/~wand/papers.html

Ruppert, D., Wand, M.P. and Carroll, R.J. (2003)
Semiparametric Regression Cambridge University Press.
http://stat.tamu.edu/~carroll/semiregbook/

See Also

gam (in package 'mgcv') lme (in package 'nlme') glmmPQL (in package 'MASS')

Examples

```
## scatterplot smoothing

x <- 1:1000/1000
mu <- exp(-400*(x-0.6)^2)+5*exp(-500*(x-0.75)^2)/3+2*exp(-500*(x-0.9)^2)
y <- mu+0.5*rnorm(1000)

#fit with default knots
y.fit <- asp(y~f(x))
plot(y.fit)

## one more scatterplot smoothing with specified knots and subknots

x <- 1:400/400
mu <- sqrt(x*(1-x))*sin((2*pi*(1+2^((9-4*6)/5)))/(x+2^((9-4*6)/5)))
y <- mu+0.2*rnorm(400)

kn <- default.knots(x,80)
kn.var <- default.knots(kn,20)

y.fit <- asp(y~f(x,knots=kn,var.knot=kn.var))
plot(y.fit)

## additive models

x1 <- 1:300/300
x2 <- runif(300)
mu1 <- exp(-400*(x1-0.6)^2)+5*exp(-500*(x1-0.75)^2)/3+2*exp(-500*(x1-0.9)^2)
mu2 <- sin(2*pi*x2)
y2 <- mu1+mu2+0.3*rnorm(300)

y2.fit <- asp(y2~f(x1)+f(x2))
par(mfrow=c(2,2))
y21.fit <- asp(y2~f(x1,adap=FALSE)+f(x2)) #switch off adaptive fitting for the first functic
plot(y2.fit)
plot(y21.fit)
par(mfrow=c(1,1))

## spatial smoothing

mu3 <- x1*sin(4*pi*x2)
y3 <- mu3+diff(range(mu3))*rnorm(300)/4

#for the specified knots and subknots use
        # kn <- default.knots.2D(x1,x2,12^2) # !!! interactive function !!!
        # kn.var <- default.knots.2D(kn[,1],kn[,2],5^2)
        # y3.fit <- asp(y3~f(x1,x2,knots=kn,var.knot=kn.var))

## non-normal response

x <- 1:1000/1000
```

```
mu <- exp(-400*(x-0.6)^2)+5*exp(-500*(x-0.75)^2)/3+2*exp(-500*(x-0.9)^2)
y4 <- rbinom(1000,5,1/(1+exp(-mu)))
nn <- rep(5,1000)
y4.fit <- asp(cbind(y4,nn-y4)~f(x),family="binomial")
### same as ###  y4.fit <- asp(y4/nn~f(x),family="binomial",weights=nn)
plot(y4.fit) #plot of systematic component

## correlated errors

y5 <- sin(2*pi*x1)+0.3*arima.sim(300,model=list(ar=0.6))

y5.fit <- asp(y5~f(x1),adap=FALSE,correlation=corAR1())
plot(y5.fit)

#see also SemiPar User Manual

#
# The current version of the SemiPar User Manual is posted on the web-site:
#
#     www.maths.unsw.edu.au/~wand/papers.html
```

Index