Package 'mcll'

February 20, 2015

Type Package

Title Monte Carlo Local Likelihood Estimation

Version 1.2

Date 2014-3-1

Author Minjeong Jeon, Cari Kaufman, and Sophia Rabe-Hesketh

Maintainer Minjeong Jeon<jeon.117@osu.edu>

Depends R (>= 2.13.0),statmod, locfit

Description Maximum likelihood estimation using a Monte Carlo local likelihood (MCLL) method

License GPL (>= 2)

LazyLoad yes

NeedsCompilation no

Repository CRAN

Date/Publication 2014-03-01 20:25:56

R topics documented:

`mcll-package` *Monte Carlo local likelihood estimation*

Description

This package provides estimates of parameters and standard errors using a Monte Carlo local likelihood method. The pacakge `mcll` is based on Jeon, Kaufman, and Rabe-Hesketh (2014).

Details

Package:	mcll
Type:	Package
Version:	1.2
Date:	2014-3-1
License:	GPL

Author(s)

Minjeong Jeon <jeon.117@osu.edu>, Cari Kaufman <cgk@berkeley.edu>, and Sophia Rabe-Hesketh <sophiarh@berkeley.edu>

Maintainer: Minjeong Jeon <jeon.117@osu.edu>

References

Jeon, M., Kaufman, C., and Rabe-Hesketh, S. (2014). Monte Carlo local likelihood for approximate MLE of complex models. Under revision.

`mcll_est` *Parameter estimation using MCLL*

Description

Parameter estimation using Monte Carlo local likelihood

Usage

```
mcll_est(data, prior.func, alp=0.7,
        method="BFGS",  lower = -Inf, upper = Inf,
        control=list(), use.locfit=TRUE,
        con.manual=list(method="BFGS", lower = -Inf, upper = Inf,
        control=list()) )
```

Arguments

`data`	posterior samples of model parameters. a `matrix` or `data.frame` of size $m \times p$ (m: sample size, p: dimension of the parameters).
`prior.func`	a prior function. An argument should be a vector of parameter values and a return value should be the log prior density for those parameter values.
`alp`	a real value between 0 and 1. α takes a value between 0 and 1, which is the nearest neighbor bandwidth with the kth smallest distance d where $k = \lfloor n\alpha \rfloor$ and $d(x, x_i) = \lvert x - x_i \rvert$ with the sample size n

method	an optimization method to be used in maximizing the approximation to the unnormalized log-likelihood. Options from optim are Nelder-Mead, BFGS, CG, L-BFGS-B, and SANN.
lower, upper	bounds on the variables for the L-BFGS-B method in optim.
control	a list of control parameters. See control options for optim.
use.locfit	logical. If TRUE, locfit is used to compute a local likelihood density estimate. If FALSE, a code from the mcll package is used. locfit is typically faster but sometimes fails for high-dimensional parameter spaces.
con.manual	a list. An optimization method for finding the polynomial coefficients, lower and upper bounds on the variables for the L-BFGS-B method, and a list of control parameters when use.locfit = FALSE. See control options for optim.

Details

Nested maximizations in Step 2 in the Monte Carlo local likelihood estimation. It makes use of the R package locfit and the R function optim. The posterior samples should be on the real line (e.g., variance parameters should be on the log-scale). The prior distributions (provided as a form of prior.func) should be the same as those used for obtaining the posterior samples of the model parameters. For details, see Section 2 in Jeon et al. (2012).

Value

mcll_est returns a list of the following components,

par	parameter estimates on the original scale.
value	value of the function corresponding to par. This is an unnormalized log-likelihood from the MCLL algorithm. One can use this to compute the Bayes factor. For details, see Appendix of Jeon et al. (2012).
counts	a two-element integer vector giving the number of calls to function and gradient, respectively.
convergence	an integer code. 0 indicates successful completion. For possible error codes, see the document for optim.
message	a character string giving any additional information returned by optim, or NULL.

Author(s)

Minjeong Jeon <jeon.117@osu.edu>

References

Jeon, M., Kaufman, C., and Rabe-Hesketh, S. (2014). Monte Carlo local likelihood for approximate MLE of complex models. Under revision.

Loader, C. (2012). locfit: Local regression, likelihood, and density estimation. Downloadable from http://cran.r-project.org/web/packages/locfit/index.html.

See Also

mcll_se

Examples

```
## example

# posterior samples
data(samp)

# prior function
prior.func <- function(vec.t) {
    sum(dnorm(vec.t, m= c(0,0,0,0, -0.9870405, -0.9870405) ,
                sd=c(100,100,100,100, 1/0.766672, 1/0.766672) , log=TRUE))
}

## parameter estimation
run1 <- system.time(
    result1 <- mcll_est(data=samp, prior.func= prior.func, alp=0.7,
        method = "BFGS", control= list(maxit=10000))
)

# result1$par
#            b0        b1         b2       b3       tau0     tau1
#[1,] 0.9275766 -2.871686 -0.6488625 3.589313 0.08118962 0.148478
```

`mcll_se` *Standard error estimation using MCLL*

Description

Standard error estimation using Monte Carlo local likelihood

Usage

```
mcll_se(data, par, H.prior, alp=0.7,
        method="Nelder-Mead",  lower = -Inf, upper = Inf, control=list() )
```

Arguments

`data`	posterior samples of model parameters. a `matrix` or `data.frame` of size $m \times p$ (m: sample size, p: dimension of the parameters).
`par`	MCLL parameter estimates on the original scale.
`H.prior`	Hessian matrix of the prior evaluated at `par`.
`alp`	a real value between 0 and 1. α takes a value between 0 and 1, which is the nearest neighbor bandwidth with the kth smallest distance d where $k = \lfloor n\alpha \rfloor$ and $d(x, x_i) = \lvert x - x_i \rvert$ with the sample size n.
`method`	an optimization method to be used to find the coefficients of the polynomial approximation to the log-posterior at the MCLL estimates `par`. Options from `optim` are Nelder-Mead, BFGS, CG, L-BFGS-B, and SANN.

lower, upper	bounds on the variables for the L-BFGS-B method in optim.
control	a list of control parameters. See control options for optim.

Details

Standard error estimation in the Monte Carlo local likelihood method. For details, see Section 3 in Jeon et al. (2012). The posterior samples and paramter values should be on the real line (e.g., variance parameters should be in the log-scale).

Value

mcll_se returns a vector containing standard error estimates for the MCLL parameter estimates par.

Author(s)

Minjeong Jeon <jeon.117@osu.edu>

References

Jeon, M., Kaufman, C., and Rabe-Hesketh, S. (2014). Monte Carlo local likelihood for approximate MLE of complex models. Under revision.

See Also

mcll_est

Examples

```
## example

# data preparation
data(samp)

# prior function
prior.func <- function(vec.t) {
    sum(dnorm(vec.t, m= c(0,0,0,0, -0.9870405, -0.9870405) ,
                sd=c(100,100,100,100, 1/0.766672, 1/0.766672) , log=TRUE))
}

## parameter estimation
run1 <- system.time(
    result1 <- mcll_est(data=samp, prior.func= prior.func, alp=0.7,
        method = "BFGS", control= list(maxit=10000) )
)

par <- result1$par # original scale

## standard error estimation

# H.prior: analytical solution
```

```
p.var = c(100,100,100,100, 1/0.766672, 1/0.766672)^2
H.prior <- -diag(1/p.var)

# H.prior: numerical solution
# library(numDeriv)
# H.prior <- hessian(prior.func, par)

# SE estimation (NOT RUN)
#run2 <- system.time(
#    se <- mcll_se(data=samp, par=par, H.prior = H.prior,  alp=0.7,
#        method= "Nelder-Mead" , control=list(maxit=20000))
#)

#se
#       b0        b1        b2        b3      tau0      tau1
#0.4057844 0.5640063 0.4907643 0.6663096 0.3022842 0.2999727
```

`salamander` *Salamander mating dataset from McCullagh and Nelder (1989)*

Description

The salamader dataset is from an experiment conducted at the University of Chicago in 1986 to study the extent to which mountain dusky salamanders from different populations would interbred. More detailed description of the data is given in McCullagh and Nelder (1989, Section 14.5).

The dataset contains 6 columns and 360 rows.

y whether a mating was successful, Yes=1 and No=0.

female identification number of the female salamander.

male identification number of the male salamander.

group group number; 1,2,3,4,5,6

experiment experiment number; 1,2,3.

rbm type of the male salamander; Rough Butt=1 and White Side=0

rbf type of the female salamander; Rough Butt=1 and White Side=0

wsm type of the male salamander; White Side=1 and Rough Butt=0

wsf type of the female salamander; White Side=1 and Rough Butt=0

ww interaction between female White Side and male White side; 1: a White Side female was crossed with a White Side male

Source

McCullagh P. and Nelder, J. A. 1989. *Generalized Linear Models*, Section 14.5, Chapman and Hall/CRC.

Examples

```
data(salamander)
str(salamander)
```

`samp` *Posterior samples of model parameters for the salamander mating model*

Description

These are 3,000 posterior samples of the six model parameters (four fixed effects and two log standard deviation parameters). For the model and priors, see Section 4 in Jeon et al. (2012).

Format

A data matrix with 3,000 posterior samples for the six parameters.

References

Jeon, M., Kaufman, C., and Rabe-Hesketh, S. (2012). Monte Carlo local likelihood for approximate MLE of complex models. Submitted for publication.

Examples

```
data(samp)
str(samp)
```

Index