

Package ‘binsegRcpp’

September 14, 2020

Type Package

Title Efficient Implementation of Binary Segmentation

Version 2020.9.3

Author Toby Dylan Hocking

Maintainer Toby Dylan Hocking <toby.hocking@r-project.org>

Description Standard template library
containers are used to implement an efficient binary segmentation
algorithm, which is log-linear on average and quadratic in the
worst case.

License GPL-3

LinkingTo Rcpp

URL <https://github.com/tdhock/binseg>

BugReports <https://github.com/tdhock/binseg/issues>

Imports data.table, Rcpp

Suggests ggplot2, testthat

NeedsCompilation yes

Repository CRAN

Date/Publication 2020-09-14 08:30:07 UTC

R topics documented:

| | |
|------------------------------|----------|
| binsegRcpp-package | 2 |
| binseg_normal | 2 |
| coef.binseg_normal | 4 |
| plot.binseg_normal | 5 |
| Index | 6 |

binsegRcpp-package *A short title line describing what the package does*

Description

A more detailed description of what the package does. A length of about one to five lines is recommended.

Details

This section should provide a more detailed overview of how to use the package, including the most important functions.

Author(s)

Your Name, email optional.

Maintainer: Your Name <your@email.com>

References

This optional section can contain literature or other references for background information.

See Also

Optional links to other man pages

Examples

```
## Not run:
## Optional simple examples of the most important functions
## These can be in \dontrun{} and \donttest{} blocks.

## End(Not run)
```

binseg_normal *Binary segmentation, normal change in mean*

Description

Efficient implementation of binary segmentation for change in mean, max normal likelihood = min square loss. Output includes columns which can be used to compute parameters for a single model in log-linear time.

Usage

```
binseg_normal(data.vec,
              max.segments = length(data.vec))
```

Arguments

`data.vec` Vector of numeric data to segment.
`max.segments` Maximum number of segments to compute, default=length(data.vec).

Value

data.table with a row for each model and columns

| | |
|--------------------------------|--|
| <code>segments</code> | number of parameters |
| <code>loss</code> | square loss |
| <code>end</code> | index of last data point per segment |
| <code>before.mean</code> | mean before changepoint |
| <code>after.mean</code> | mean after changepoint |
| <code>before.size</code> | number of data before changepoint |
| <code>after.size</code> | number of data after changepoint |
| <code>invalidates.index</code> | |
| | index of model parameter no longer used after this changepoint is used |
| <code>invalidates.after</code> | |
| | idem |

Author(s)

Toby Dylan Hocking

Examples

```
x <- c(0.1, 0, 1, 1.1, 0.1, 0)
## Compute full path of binary segmentation models from 1 to 6
## segments.
(models.dt <- binsegRcpp::binseg_normal(x))

## Plot loss values using base graphics.
plot(models.dt)

## Same loss values using ggplot2.
if(require("ggplot2")){
  ggplot()+
    geom_point(aes(
      segments, loss),
      data=models.dt)
}

## Compute data table of segments to plot.
(segs.dt <- coef(models.dt, 2:4))

## Plot data, segments, changepoints.
if(require("ggplot2")){
```

```

ggplot()+
  theme_bw()+
  theme(panel.spacing=grid::unit(0, "lines"))+
  facet_grid(segments ~ ., labeller=label_both)+
  geom_vline(aes(
    xintercept=start-0.5),
    color="green",
    data=segs.dt[1<start])+
  geom_segment(aes(
    start-0.5, mean,
    xend=end+0.5, yend=mean),
    data=segs.dt,
    color="green")+
  xlab("Position/index")+
  ylab("Data/mean value")+
  geom_point(aes(
    pos, x),
    data=data.frame(x, pos=seq_along(x)))
}

```

coef.binseg_normal *coef binseg normal*

Description

Compute a data table of segment start/end/mean values for all models given by segments.

Usage

```

## S3 method for class 'binseg_normal'
coef(object,
      segments = 1:min(nrow(object),
                       10), ...)

```

Arguments

| | |
|----------|--|
| object | data.table from binseg_normal . |
| segments | integer vector, model sizes in number of segments. |
| ... | ignored. |

Value

data.table with one row for each segment.

Author(s)

Toby Dylan Hocking

`plot.binseg_normal` *plot binseg normal*

Description

Plot loss values from binary segmentation.

Usage

```
## S3 method for class 'binseg_normal'  
plot(x,  
     ...)
```

Arguments

`x` data.table from [binseg_normal](#).
`...` ignored.

Author(s)

Toby Dylan Hocking

Index

* package

binsegRcpp-package, 2

binseg_normal, 2, 4, 5

binsegRcpp (binsegRcpp-package), 2

binsegRcpp-package, 2

coef.binseg_normal, 4

plot.binseg_normal, 5