Package 'densratio'

June 30, 2019

Type Package

Version 0.2.1

Title Density Ratio Estimation

Description Density ratio estimation.
The estimated density ratio function can be used in many applications such as
anomaly detection, change-point detection, covariate shift adaptation.
The implemented methods are uLSIF (Hido et al. (2011) <doi:10.1007/s10115-010-0283-2>),
RuLSIF (Yamada et al. (2011) <doi:10.1162/NECO_a_00442>),
and KLIEP (Sugiyama et al. (2007) <doi:10.1007/s10463-008-0197-x>).

URL https://github.com/hoxo-m/densratio

BugReports https://github.com/hoxo-m/densratio/issues

License MIT + file LICENSE

Imports utils

Suggests graphics, knitr, mvtnorm, rmarkdown, stats, testthat

RoxygenNote 6.1.1

VignetteBuilder knitr

Encoding UTF-8

NeedsCompilation no

Author Koji Makiyama [aut, cre]

Maintainer Koji Makiyama <hoxo.smile@gmail.com>

Repository CRAN

Date/Publication 2019-06-30 17:20:03 UTC

R topics documented:

densratio *Estimate Density Ratio p(x)/q(x)*

Description

Estimate Density Ratio p(x)/q(x)

Usage

```
densratio(x1, x2, method = c("uLSIF", "RuLSIF", "KLIEP"),
  sigma = "auto", lambda = "auto", alpha = 0.1, kernel_num = 100,
  fold = 5, verbose = TRUE)
```

Arguments

x1	numeric vector or matrix. Data from a numerator distribution p(x).
x2	numeric vector or matrix. Data from a denominator distribution q(x).
method	"uLSIF" (default), "RuLSIF", or "KLIEP".
sigma	positive numeric vector. Search range of Gaussian kernel bandwidth.
lambda	positive numeric vector. Search range of regularization parameter for uLSIF and RuLSIF.
alpha	numeric in [0, 1]. Relative parameter for RuLSIF. Default 0.1.
kernel_num	positive integer. Number of kernels.
fold	positive integer. Numer of the folds of cross validation for KLIEP.
verbose	logical (default TRUE).

Value

densratio object that contains a function to compute estimated density ratio.

Examples

```
x1 <- rnorm(200, mean = 1, sd = 1/8)
x2 <- rnorm(200, mean = 1, sd = 1/2)

densratio_obj <- densratio(x1, x2)

new_x <- seq(0, 2, by = 0.05)
estimated_density_ratio <- densratio_obj$compute_density_ratio(new_x)

plot(new_x, estimated_density_ratio, pch=19)
```

KLIEP *Estimate Density Ratio p(x)/q(x) by KLIEP (Kullback-Leibler Importance Estimation Procedure)*

Description

Estimate Density Ratio p(x)/q(x) by KLIEP (Kullback-Leibler Importance Estimation Procedure)

Usage

```
KLIEP(x1, x2, sigma = "auto", kernel_num = 100, fold = 5,
  verbose = TRUE)
```

Arguments

x1	numeric vector or matrix. Data from a numerator distribution p(x).
x2	numeric vector or matrix. Data from a denominator distribution q(x).
sigma	positive numeric vector. Search range of Gaussian kernel bandwidth.
kernel_num	positive integer. Number of kernels.
fold	positive integer. Number of the folds of cross validation.
verbose	logical (default TRUE).

Value

KLIEP object that contains a function to compute estimated density ratio.

RuLSIF *Estimate alpha-Relative Density Ratio p(x)/(alpha p(x) + (1-alpha) q(x)) by RuLSIF (Relative unconstrained Least-Square Importance Fitting)*

Description

Estimate alpha-Relative Density Ratio p(x)/(alpha p(x) + (1-alpha) q(x)) by RuLSIF (Relative unconstrained Least-Square Importance Fitting)

Usage

```
RuLSIF(x1, x2, sigma = 10^seq(-3, 1, length.out = 9),
  lambda = 10^seq(-3, 1, length.out = 9), alpha = 0.1,
  kernel_num = 100, verbose = TRUE)
```

Arguments

x1	numeric vector or matrix. Data from a numerator distribution p(x).
x2	numeric vector or matrix. Data from a denominator distribution q(x).
sigma	positive numeric vector. Search range of Gaussian kernel bandwidth.
lambda	positive numeric vector. Search range of regularization parameter.
alpha	numeric value from 0.0 to 1.0. Relative parameter. Default 0.1.
kernel_num	positive integer. Number of kernels.
verbose	logical. Default TRUE.

Value

RuLSIF object which has 'compute_density_ratio()'.

`squared_euclid_distance` *Compute Squared Euclid Distance*

Description

Compute Squared Euclid Distance

Usage

```
squared_euclid_distance(x, y)
```

Arguments

x	a numeric vector.
y	a numeric vector.

Value

squared Euclid distance

uLSIF *Estimate Density Ratio p(x)/q(x) by uLSIF (unconstrained Least-Square Importance Fitting)*

Description

Estimate Density Ratio p(x)/q(x) by uLSIF (unconstrained Least-Square Importance Fitting)

Usage

```
uLSIF(x1, x2, sigma = 10^seq(-3, 1, length.out = 9),
  lambda = 10^seq(-3, 1, length.out = 9), kernel_num = 100,
  verbose = TRUE)
```

Arguments

x1	numeric vector or matrix. Data from a numerator distribution p(x).
x2	numeric vector or matrix. Data from a denominator distribution q(x).
sigma	positive numeric vector. Search range of Gaussian kernel bandwidth.
lambda	positive numeric vector. Search range of regularization parameter.
kernel_num	positive integer. Number of kernels.
verbose	logical(default TRUE).

Value

uLSIF object that contains a function to compute estimated density ratio.

Index