

Package ‘WPKDE’

May 26, 2025

Type Package

Title Weighted Piecewise Kernel Density Estimation

Version 1.0

Date 2025-05-22

Description Weighted Piecewise Kernel Density Estimation for large data.

Depends R ($\geq 4.3.0$), Rcpp, plotly, RANN

License GPL (≥ 2)

Encoding UTF-8

LazyData true

NeedsCompilation yes

RoxygenNote 7.3.2

LinkingTo Rcpp

Author Xiaotong Liu [aut, cre],
Kunyu Ye [aut],
Siyao Wang [aut],
Xudong Liu [aut],
Tianwei Yu [aut, ths]

Maintainer Xiaotong Liu <xiaotongliu@link.cuhk.edu.cn>

Repository CRAN

Date/Publication 2025-05-26 06:00:07 UTC

Contents

findPeak	2
kdeC	2
plot_peak_2d	3
plot_peak_3d	4
r	5
Index	6

findPeak	<i>Find peaks using the estimated values</i>
----------	--

Description

Find peaks using the estimated values

Usage

```
findPeak(k, filter, select)
```

Arguments

k	Output of the 'kdeC' function, containing estimated values.
filter	A numeric value used to filter out results with estimated values less than the given 'filter' argument.
select	A numeric value specifying the number of peaks to retain, selecting the K peaks with the largest estimated values.

Value

A three-column matrix ('markMat') where: - Column 1: x-coordinates of the peaks - Column 2: y-coordinates of the peaks - Column 3: Corresponding estimated values of the peaks.

Examples

```
data(r)
k <- kdeC(r$dat, H = c(0.014, 0.014), gridsize = c(330, 330), cutNum = c(1, 1), w = r$z)
m <- findPeak(k, filter = 0, select = 100)
```

kdeC	<i>Two-dimensional fast weighted kernel density estimation</i>
------	--

Description

Two-dimensional fast weighted kernel density estimation

Usage

```
kdeC(x, H, gridsize, cutNum, w)
```

Arguments

x	Data points in the format of an n x 2 matrix.
H	Bandwidth, a vector containing 2 numeric values.
gridsize	Number of points for each direction, a vector containing 2 integer values.
cutNum	Number of pieces to be cut for each direction, a vector containing 2 integer values.
w	Weight, a vector corresponding to parameter 'x'.

Value

A list containing three elements:

estimate	The estimated values of the kernel density.
evalpointsX	The evaluation points along the X direction.
evalpointsY	The evaluation points along the Y direction.

Examples

```
data(r)
k <- kdeC(r$dat, H = c(0.014, 0.014), gridsize = c(330, 330), cutNum = c(1, 1), w = r$z)
```

plot_peak_2d	<i>Plot of the 2D data points with peaks highlighted in green</i>
--------------	---

Description

Plot of the 2D data points with peaks highlighted in green

Usage

```
plot_peak_2d(dat, peaks, x.range = NA, y.range = NA)
```

Arguments

dat	Data points used for kernel density estimation.
peaks	A matrix of detected peaks with x- and y-coordinates.
x.range	(optional) A numeric 2D vector specifying the x-axis range for filtering.
y.range	(optional) A numeric 2D vector specifying the y-axis range for filtering.

Value

A scatter plot of the data points with the detected peaks highlighted in green.

Examples

```
data(r)
k <- kdeC(r$dat, H = c(0.014, 0.014), gridsize = c(330, 330), cutNum = c(1, 1), w = r$z)
m <- findPeak(k, filter = 0, select = 100)
plot_peak_2d(r$dat, m)
```

plot_peak_3d

Plot of the 3D data points with peaks highlighted in green

Description

This function creates an interactive 3D scatter plot of data points and highlights the peaks that are within a specified tolerance distance from any data point.

Usage

```
plot_peak_3d(dat, peaks, x.range = NA, y.range = NA, tol = 1e-05)
```

Arguments

dat	A numeric matrix or data frame with at least three columns representing x, y, and z coordinates of data points.
peaks	A numeric matrix or data frame with at least two columns representing the x and y coordinates of peak candidates.
x.range	A numeric vector of length 2 specifying the x-axis range to include.
y.range	A numeric vector of length 2 specifying the y-axis range to include.
tol	A numeric value specifying the tolerance threshold: only peaks within this Euclidean distance from a data point are retained.

Examples

```
data(r)
k <- kdeC(r$dat, H = c(0.014, 0.014), gridsize = c(330, 330), cutNum = c(1, 1), w = r$z)
m <- findPeak(k, filter = 0, select = 100)
dat <- cbind(r$dat, r$z)
plot_peak_3d(dat, m)
```

r

Simulated 2D Weighted Data Set

Description

This is a simulated dataset containing two-dimensional data points, their corresponding weights, and the true peaks' coordinates.

Usage

```
data(r)
```

Format

A list with 3 components:

dat A data.frame of size 100000 x 2, representing data point coordinates.

m A numeric matrix of true peaks' coordinates.

z A numeric vector of length 100000, representing weights for each data point.

Examples

```
data(r)
```

Index

* **datasets**

r, [5](#)

findPeak, [2](#)

kdeC, [2](#)

plot_peak_2d, [3](#)

plot_peak_3d, [4](#)

r, [5](#)