

Package ‘lpcade’

July 31, 2014

Type Package

Title Local PCA dimension estimators

Version 1.0

Date 2012-17-20

Author Kerstin Johnsson, Lund University

Maintainer Kerstin Johnsson <johnsson@maths.lth.se>

Description

Provides functions to do manifold dimension estimation with various versions of local PCA.

Depends manifgen

License GPL (>=2)

URL <http://www.maths.lu.se/staff/kerstin-johnsson/research/manifold-dimension-estimation/>

LazyLoad yes

R topics documented:

bs	1
lpca	2

Index 4

bs	<i>Dimension Estimation with Optimally Topology Perserving Maps</i>
----	---

Description

Intrinsic dimension estimation with the metod proposed in Bruske and Sommer (1998). A graph called optimally topology perserving map (OTPM) is constructed and on this local PCA is made with the Fukunaga-Olsen criterion to determine which eigenvalues that are significant.

Usage

```
bs(data, N, alpha = .05)
```

Arguments

data	a data set for which dimension should be estimated.
N	the number of the nodes in the OTPM.
alpha	the significance level for the Fukunaga-Olsen method.

Value

A matrix with two columns

de	The dimension estimate at each point.
nbr.nb	The number of neighbouring nodes used for the dimension estimate at each point.

Author(s)

Kerstin Johnsson, Lund university

References

Bruske, J. and Sommer, G. (1998) Intrinsic dimensionality estimation with optimally topology perserving maps. *IEEE Trans. on Pattern Anal. and Mach. Intell.*, **20**(5), 572-575.

See Also

[lpca](#)

Examples

```
library(manifgen)
data <- hball(1000, 5)
bs(data, 400)
```

lpca

Local Dimension Estimation with PCA

Description

Estimates manifold dimension using the largest singular values of the covariance matrix.

Usage

```
lpca(data, ver, alphaF0 = .05, alphaFan = 10, betaFan = .8, PFan = .95,
      ngap = 5, maxdim = min(dim(data)), verbose = TRUE)
```

Arguments

data	a local data set for which dimension should be estimated.
ver	possible values: 'FO', 'fan', 'maxgap', 'cal'. 'cal' is often very slow.
alphaFO	only for ver = 'FO'. An eigenvalue is considered significant if it is larger than alpha times the largest eigenvalue.
alphaFan	only for ver = 'Fan'. The alpha parameter (large gap threshold).
betaFan	only for ver = 'Fan'. The beta parameter (total covariance threshold).
PFan	only for ver = 'Fan'. Total covariance in non-noise.
ngap	only for ver = cal. The number of gaps for which likelihood should be calculated.
maxdim	only for ver = 'cal'. The maximal manifold dimension of the data.
verbose	should information about the process be printed out?

Details

Version 'FO' is the method by Fukunaga-Olsen, 'fan' is the method by Fan et al., 'maxgap' returns the largest relative gap in the singular values, 'cal' is a calibrated version.

Value

A list with two components

de	the dimension estimate and
gap.size	if ver != cal; the size of the gap in singular values corresponding to the estimated dimension
likelihood	if ver = cal; the likelihood of the estimated dimension.

Author(s)

Kerstin Johnsson, Lund University

Examples

```
library(manifgen)
data <- cuthplane(100, 4, 10, .05)
lpca(data, 'fan')
lpca(data, 'FO')
lpca(data, 'maxgap')
```

Index

bs, 1

lpca, 2, 2