Package 'speccalt'

February 20, 2015

Version 0.1.1

Date 2013-09-16

Author Pierrick Bruneau

Maintainer Pierrick Bruneau <pbruneau@gmail.com>

Title Alternative spectral clustering, with automatic estimation of k.

Description Alternative to the kernlab::specc function. Includes a spectral clustering implementation, a locally adapted kernel function akin to what is already proposed in kernlab, and an optional procedure that automatically estimates the optimal number of clusters. Several sample data sets are also included.

Depends R (>= 2.10.0)

Imports

LazyLoad yes

LazyData yes

License LGPL-3

NeedsCompilation no

Repository CRAN

Date/Publication 2013-09-16 14:28:13

R topics documented:

`bartlett` *bartlett*

Description

Adaptation of the Bartlett statistical test for equal variances, to a procedure of detection of the optimal number of clusters, in the context of spectral clustering.

Usage

```
bartlett(eigvals, thres=0.95, maxk=20)
```

Arguments

`eigvals`	A vector of eigenvalues, as returned by the eigendecomposition of the normalized Laplacian (see examples), i.e. in decreasing order.
`thres`	1 - alpha significance threshold, defaults to 95
`maxk`	upper bound to the returned value.

Value

The computed optimal number of clusters.

Author(s)

Pierrick Bruneau

Examples

```
# normalized Laplacian computation
kern <- local.rbfdot(synth1)
nelts <- dim(kern)[1]
diag(kern) <- 0
deg <- sapply(1:nelts, function(i) {
return(sum(kern[i,]))
})
deg <- diag(deg)
L <- diag(nelts) - solve(deg)
```

```
eig <- eigen(L)

optK <- bartlett(eig$values)$k
```

`irisdat` *irisdat*

Description

matrix of 150 elements on 4 features, extracted from iris standard data.frame (4 first variables). See reference.

Format

The format is: num [1:150, 1:4] 5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...

References

Fisher, R. A. (1936) _The use of multiple measurements in taxonomic problems_, Annals of Eugenics, Volume 7, Part II, Pages 179-188.

Examples

```
plot(irisdat)
```

`irislab` *irislab*

Description

Ground truth labels for elements in irisdat.

Format

The format is: num [1:150] 2 2 2 2 2 2 2 2 2 2 ...

Examples

```
plot(irisdat, col=irislab)
```

isodat *isolet*

Description

matrix of 1500 elements on 617 features. Vowels taken from the isolet UCI data set.

Format

The format is: num [1:1500, 1:617] -0.4394 -0.4348 -0.3928 -0.5494 -0.4550 -0.4804 -0.3656 -0.5154 -0.1836 -0.2844 ...

Examples

```
plot(isodat)
```

isolab *isolab*

Description

Ground truth labels for elements in isodat.

Format

The format is: num [1:1500] 2 2 3 3 4 4 5 5 6 6 ...

Examples

```
plot(isodat, col=isolab)
```

lab1 *lab1*

Description

Ground truth labels for elements in synth1.

Format

The format is: num [1:298] 4 4 4 4 4 4 4 4 4 4 ...

Examples

```
plot(synth1, col=lab1)
```

lab2 *lab2*

Description

Ground truth labels for elements in synth1.

Format

The format is: num [1:302] 4 4 4 4 4 4 4 4 4 4 ...

Examples

```
plot(synth2, col=lab2)
```

lab3 *lab3*

Description

Ground truth labels for elements in synth1.

Format

The format is: num [1:265] 4 4 4 4 4 4 4 4 4 4 ...

Examples

```
plot(synth3, col=lab3)
```

lab4 *lab4*

Description

Ground truth labels for elements in synth4.

Format

The format is: num [1:621] 4 4 4 4 4 4 4 4 4 4 ...

Examples

```
plot(synth4, col=lab4)
```

lab5 *lab5*

Description

Ground truth labels for elements in synth5.

Format

The format is: num [1:511] 4 4 4 4 4 4 4 4 4 4 ...

Examples

```
plot(synth5, col=lab5)
```

lab6 *lab6*

Description

Ground truth labels for elements in synth6.

Format

The format is: num [1:237] 4 4 4 4 4 4 4 4 4 4 ...

Examples

```
plot(synth6, col=lab6)
```

local.rbfdot *local.rbfdot*

Description

Locally adapted RBF kernel function. Taken and adapted from kernlab.

Usage

```
local.rbfdot(data, n=5)
```

Arguments

data	A data set matrix, with N elements as rows.
n	Parameter for median computation, defaults to 5.

Value

A NxN matric of RBF kernel values.

Author(s)

Pierrick Bruneau

Examples

```
kern <- local.rbfdot(synth1)
```

---

speccalt *speccalt*

Description

Spectral clustering algorithm, with optional automatic estimation of the optimal number of clusters.

Usage

```
speccalt(kern, k=NA, maxk=20)
```

Arguments

kern	A kernel matrix object, as returned e.g. by local.rbfdot.
k	Parameterized number of clusters. If NA, the automatic estimation procedure (bartlett) is used.
maxk	If k is NA, an upper bound for the automatic estimation. Defaults to 20.

Value

A numeric vector of cluster labels. Clusters are indexed from 1 to k.

Author(s)

Pierrick Bruneau

See Also

local.rbfdot bartlett

Examples

```
kern <- local.rbfdot(synth2)
clust <- speccalt(kern) # with automatic estimation
clust2 <- speccalt(kern, 4) # without automation
```

synth1 *synth1*

Description

matrix of 298 elements on 2 features. Synthetic data set, see reference.

Format

The format is: num [1:298, 1:2] 0.4550112 0.4475910 0.4357260 0.4640844 0.4460929 0.4528965 0.4542374 0.4690467 0.4396205 0.4610803 ...

References

Zelnik-Manor, L. and Perona, P. (2004). Self-tuning spectral clustering. Neural Information Processing Symposium.

Examples

```
plot(synth1)
```

synth2 *synth2*

Description

matrix of 302 elements on 2 features. Synthetic data set, see reference.

Format

The format is: num [1:302, 1:2] 0.6988742 0.9332699 -0.1523060 0.5511859 0.8133928 0.5597358 0.7980163 0.1351562 0.2319774 0.9182794 ...

References

Zelnik-Manor, L. and Perona, P. (2004). Self-tuning spectral clustering. Neural Information Processing Symposium.

Examples

```
plot(synth2)
```

synth3 *synth3*

Description

matrix of 265 elements on 2 features. Synthetic data set, see reference.

Format

The format is: num [1:265, 1:2] 0.1666711 0.1630728 0.1680026 0.1471314 0.1814486 0.1667588 0.1567425 0.1613646 0.1677744 0.1515392 ...

References

Zelnik-Manor, L. and Perona, P. (2004). Self-tuning spectral clustering. Neural Information Processing Symposium.

Examples

```
plot(synth3)
```

synth4 *synth4*

Description

matrix of 621 elements on 2 features. Synthetic data set, see reference.

Format

The format is: num [1:621, 1:2] 1.084640271 0.414354482 0.628829245 0.246931608 0.160063275 0.078559001 0.002920947 0.103797589 1.039537189 0.045987087 ...

References

Zelnik-Manor, L. and Perona, P. (2004). Self-tuning spectral clustering. Neural Information Processing Symposium.

Examples

```
plot(synth4)
```

synth5 *synth5*

Description

matrix of 511 elements on 2 features. Synthetic data set, see reference.

Format

The format is: num [1:511, 1:2] 0.09884874 0.10504572 0.11124270 0.11743968 0.12363666 0.12983365 0.13603063 0.14222761 0.14842459 0.15462157 ...

References

Zelnik-Manor, L. and Perona, P. (2004). Self-tuning spectral clustering. Neural Information Processing Symposium.

Examples

```
plot(synth5)
```

synth6 *synth6*

Description

matrix of 237 elements on 2 features. Synthetic data set, see reference.

Format

The format is: num [1:237, 1:2] 0.7479543 0.7539890 0.7473104 0.7202754 0.7134525 0.6835628 0.6653153 0.6258207 0.6194835 0.5892482 ...

References

Zelnik-Manor, L. and Perona, P. (2004). Self-tuning spectral clustering. Neural Information Processing Symposium.

Examples

```
plot(synth6)
```

Index