

Package ‘dkDNA’

June 1, 2015

Type Package

Title Diffusion Kernels on a Set of Genotypes

Version 0.1.1

Date 2015-05-31

Author Gota Morota and Masanori Koyama

Maintainer Gota Morota <morota@unl.edu>

Description

Compute diffusion kernels on DNA polymorphisms, including SNP and bi-allelic genotypes.

License GPL-2

URL <http://morotalab.org/>

ByteCompile true

NeedsCompilation yes

Repository CRAN

Date/Publication 2015-06-01 01:05:22

R topics documented:

dkDNA-package	2
hypercube	2
snpgrid	3
snphamming	4

Index	6
--------------	----------

 dkDNA-package

Diffusion kernels on a set of genotypes

Description

Compute diffusion kernels on DNA polymorphisms, including SNP and binary genotypes.

Author(s)

Gota Morota <morota@unl.edu> and Masanori Koyama <koyama.masanori@gmail.com>

References

Kondor R and Lafferty J: (2002) Diffusion Kernels on Graphs and Other Discrete Input Spaces. ICML.

Morota G, Koyama M, Rosa GJM, Weigel KA, and Gianola D. (2013). Predicting complex traits using a diffusion kernel on genetic markers with an application to dairy cattle and wheat data. *Genetics Selection Evolution*. 45:17.

 hypercube

Diffusion kernels on bi-allelic genotypes

Description

This function construct a diffusion kernel on a p -dimensional hypercube, where each genotype takes on two possible configurations. This graph is obtained by the p -Cartesian graph product of a complete graph K_2 . It contains 2^p vertices corresponding to sequences of genotypes, and two vertices are adjacent if and only if just one SNP locus differs.

Usage

```
hypercube(X, theta)
```

Arguments

X	A genotype matrix of n individuals with p bi-allelic genotypes ($n \times p$).
theta	The rate of diffusion.

Value

Diffusion kernel matrix of size $n \times n$. This can be viewed as a covariance among individuals given the diffusion rate.

Author(s)

Gota Morota <morota@unl.edu> and Masanori Koyama <koyama.masanori@gmail.com>

References

Kondor R and Lafferty J: (2002) Diffusion Kernels on Graphs and Other Discrete Input Spaces. ICML.

Morota G, Koyama M, Rosa GJM, Weigel KA, and Gianola D. (2013). Predicting complex traits using a diffusion kernel on genetic markers with an application to dairy cattle and wheat data. Genetics Selection Evolution. 45:17.

Examples

```
# set a seed
set.seed(4321)

# create a genotype matrix of 5 individuals with 10 bi-allelic genotypes
X <- matrix(sample(c(0,1), 50, prob=c(0.6,0.4), replace=TRUE), ncol=10)

# set the rate of diffusion equal to 1
theta <- 1

# compute a hypercube kernel
hypercube(X, theta)
```

snpgrid

Diffusion kernels on SNP genotypes

Description

This function constructs a diffusion kernel on a p -dimensional SNP grid graph, where each genotype takes on three possible configurations, namely 0 ('aa'), 1 ('Aa'), and 2 ('AA'). This graph is obtained by the p -Cartesian graph product of a path graph 0-1-2. It contains 3^p vertices corresponding to sequences of genotypes, and two vertices are adjacent if and only if just one SNP locus differs by 1.

Usage

```
snpgrid(X, theta)
```

Arguments

X	A genotype matrix of n individuals with p SNPs ($n \times p$).
theta	The rate of diffusion.

Value

Diffusion kernel matrix of size $n \times n$. This can be viewed as a covariance among individuals given the diffusion rate.

Author(s)

Gota Morota <morota@unl.edu> and Masanori Koyama <koyama.masanori@gmail.com>

References

Kondor R and Lafferty J: (2002) Diffusion Kernels on Graphs and Other Discrete Input Spaces. ICML.

Morota G, Koyama M, Rosa GJM, Weigel KA, and Gianola D. (2013). Predicting complex traits using a diffusion kernel on genetic markers with an application to dairy cattle and wheat data. Genetics Selection Evolution. 45:17.

See Also

[snphamming](#)

Examples

```
# set a seed
set.seed(4321)

# create a genotype matrix of 5 individuals with 10 SNPs
X <- matrix(sample(c(0,1,2), 50, prob=c(0.35, 0.3, 0.35), replace=TRUE),
ncol=10)

# set the rate of diffusion equal to 1
theta <- 1

# compute a SNP grid kernel
snpgrid(X, theta)
```

snphamming

Diffusion kernels on SNP genotypes

Description

This function construct a diffusion kernel on a p -dimensional SNP hamming graph, where each genotype takes on three possible configurations, namely 0 ('aa'), 1 ('Aa'), and 2 ('AA'). This graph is obtained by the p -Cartesian graph product of a complete graph K_3 . It contains 3^p vertices corresponding to sequences of genotypes, and two vertices are adjacent if and only if just one SNP locus differs.

Usage

```
snphamming(X, theta)
```

Arguments

X A genotype matrix of n individuals with p SNPs ($n \times p$).

theta The rate of diffusion.

Value

Diffusion kernel matrix of size $n \times n$. This can be viewed as a covariance among individuals given the diffusion rate.

Author(s)

Gota Morota <morota@unl.edu> and Masanori Koyama <koyama.masanori@gmail.com>

References

Kondor R and Lafferty J: (2002) Diffusion Kernels on Graphs and Other Discrete Input Spaces. ICML.

Morota G, Koyama M, Rosa GJM, Weigel KA, and Gianola D. (2013). Predicting complex traits using a diffusion kernel on genetic markers with an application to dairy cattle and wheat data. Genetics Selection Evolution. 45:17.

See Also

[snpgrid](#)

Examples

```
# set a seed
set.seed(4321)

# create a genotype matrix of 5 individuals with 10 SNPs
X <- matrix(sample(c(0,1,2), 50, prob=c(0.35, 0.3, 0.35), replace=TRUE),
ncol=10)

# set the rate of diffusion equal to 1
theta <- 1

# compute a SNP hamming kernel
snphamming(X, theta)
```

Index

*Topic **graphs**

hypercube, [2](#)

snpgrid, [3](#)

snphamming, [4](#)

*Topic **nonparametric**

hypercube, [2](#)

snpgrid, [3](#)

snphamming, [4](#)

*Topic **package**

dkDNA-package, [2](#)

dkDNA (dkDNA-package), [2](#)

dkDNA-package, [2](#)

hypercube, [2](#)

snpgrid, [3](#), [5](#)

snphamming, [4](#), [4](#)