Package ‘Jacquard’

January 20, 2025

Type Package

Title Estimation of Jacquard's Genetic Identity Coefficients

Version 1.0.2

Date 2024-09-17

Author Jan Graffelman [aut, cre]

Maintainer Jan Graffelman <jan.graffelman@upc.edu>

Depends R (>= 3.3.0), Rsolnp

Description Contains procedures to estimate the nine condensed Jacquard genetic identity coefficients (Jacquard, 1974) <doi:10.1007/978-3-642-88415-3> by constrained least squares (Graffelman et al., 2024) <doi:10.1101/2024.03.25.586682> and by the method of moments (Csuros, 2014) <doi:10.1016/j.tpb.2013.11.001>. These procedures require previous estimation of the
allele frequencies. Functions are supplied that estimate relationship parameters that derive from the Jacquard
coefficients, such as individual inbreeding coefficients and kinship coefficients.

License GPL (>= 2)

URL https://www.r-project.org, http://www-eio.upc.edu/~jan/

Suggests knitr, rmarkdown

VignetteBuilder knitr, rmarkdown

NeedsCompilation no

Repository CRAN

Date/Publication 2024-09-17 16:20:50 UTC

Contents

`BoxplotDelta` *Boxplots of Jacquard Coefficients*

Description

Function `BoxplotDelta` produces boxplots of the Jacquard coefficients from a list structure containing nine matrices with the pairwise coefficients. The diagonals of J1 and J7 are plotted in separate boxplots.

Usage

```
BoxplotDelta(J, ind.sub = 1:nrow(J[[1]]), ...)
```

Arguments

`J`	The list structure with nine fields consisting of nine matrices
`ind.sub`	Index for subsetting the individuals
`...`	Additional arguments to pass on to `boxplot`

Value

NULL

Author(s)

Jan Graffelman (jan.graffelman@upc.edu)

Examples

```
data(DeltaSimulatedPedigree)
BoxplotDelta(DeltaSimulatedPedigree)
```

`BoxplotTheta` *Boxplots of Relateness parameters*

Description

Function `BoxplotTheta` makes boxplots of relatedess parameters from a list object containing estimates of pairwise relatedness parameters.

Usage

```
BoxplotTheta(KS, ind.sub = 1:nrow(KS[[1]]), ...)
```

Arguments

`KS`	A list object with four fields of matrices: kinship, inbreeding, T3 and T4.
`ind.sub`	Index for subsetting the individuals
`...`	Additional arguments passed on to the function `boxplot`.

Details

For individual inbreeding coefficients, two boxplots are made, the first based on the diagonal of T2, the second on the row means of T2. The remaining boxplots (kinship, T3 and T4) are all pairwise, and exclude self-pairs.

Value

NULL

Author(s)

Jan Graffelman (jan.graffelman@upc.edu)

References

Csuros, M. (2014) Non-identifiability of identity coefficients at biallelic loci. Theoretical Population Biology 92, pp. 22-29. doi:10.1016/j.tpb.2013.11.001.

Graffelman, J., Weir, B.S. and Goudet, J. (2024) Estimation of Jacquard's genetic identity coefficients with bi-allelic variants by constrained least-squares. Preprint at bioRxiv doi:10.1101/2024.03.25.586682.

See Also

`CalculateTheta`

Examples

```
data(DeltaSimulatedPedigree)
Theta <- CalculateTheta(DeltaSimulatedPedigree)
BoxplotTheta(Theta)
```

`CalculateMom` *Calculate Moment Estimators for Relatedness Parameters*

Description

Function `CalculateMom` computes moment estimators for a set of relatedness parameters (kinship, inbreeding, least one IBD out of three and T4) using the genotype data and the allele frequencies.

Usage

```
CalculateMom(Xgen, mafvec, ind.sub = 1:nrow(Xgen), verbose = TRUE)
```

Arguments

`Xgen`	Genotype data coded in (0,1,2) format
`mafvec`	A vector of minor allele frequencies
`ind.sub`	Index for subsetting individuals
`verbose`	Print output on the progress of the algorithm if `verbose=TRUE`.

Value

A list object with four fields:

`T1`	The pairwise coancestry or kinship coefficients (symmetric)
`T2`	The pairwise inbreeding coefficients (non-symmetric)
`T3`	The pairwise Least one IBD out of three (symmetric)
`T4`	T4 (skew-symmetric)

Author(s)

Jan Graffelman (jan.graffelman@upc.edu)

References

Csuros, M. (2014) Non-identifiability of identity coefficients at biallelic loci. Theoretical Population Biology 92, pp. 22-29. doi:10.1016/j.tpb.2013.11.001.

See Also

`BoxplotTheta`

Examples

```
data(SimulatedPedigree)
Xgen <- as.matrix(SimulatedPedigree[,6:ncol(SimulatedPedigree)])
mafvec <- mafvector(Xgen)
Theta.mom <- CalculateMom(Xgen[1:10,],mafvec)
```

`CalculateTheta` *Calculate Relatedness Paramenters from Jacquard Coefficients tA Capitalized Title (ideally limited to 65 characters)*

Description

Function `CalculateTheta` calculates five identifiable relatedness parameters from a list structure containing nine matrices of Jacquard coefficients.

Usage

```
CalculateTheta(J, ind.sub = 1:nrow(J[[1]]))
```

Arguments

`J`	A list structure with nine matrices of pairwise Jacquard coefficients
`ind.sub`	Index for subsetting the individuals

Details

`CalculateTheta` produces four matrices according to the expressions in Graffelman et al. (2024).

Value

A list object with four fields:

`T1`	The pairwise coancestry or kinship coefficients (symmetric)
`T2`	The pairwise inbreeding coefficients (non-symmetric)
`T3`	The pairwise Least one IBD out of three (symmetric)
`T4`	T4 (skew-symmetric)

Author(s)

Jan Graffelman (jan.graffelman@upc.edu)

References

Csuros, M. (2014) Non-identifiability of identity coefficients at biallelic loci. Theoretical Population Biology 92, pp. 22-29. doi:10.1016/j.tpb.2013.11.001.

Graffelman, J., Weir, B.S. and Goudet, J. (2024) Estimation of Jacquard's genetic identity coefficients with bi-allelic variants by constrained least-squares. Preprint at bioRxiv doi:10.1101/2024.03.25.586682.

See Also

`Jacquard.cls`

Examples

```
data(DeltaSimulatedPedigree)
Theta <- CalculateTheta(DeltaSimulatedPedigree)
```

`DeltaPair` *Extract estimates of Jacquard coefficients for a given pair*

Description

Function `DeltaPair` extracts from the list object of all pairwise Jacquard coefficients the set of coefficients of a given pair (i,j).

Usage

```
DeltaPair(Delta, i, j, digits = 7)
```

Arguments

`Delta`	A list with nine matrices of pairswise coefficients.
`i`	Index of the first individual.
`j`	Index of the second individual.
`digits`	Number of digits to which the coefficients are rounded.

Value

A vector with nine elements

Author(s)

Jan Graffelman (jan.graffelman@upc.edu)

See Also

`PairwiseList`,`Jacquard.cls`

Examples

```
data(DeltaSimulatedPedigree)
DeltaPair(DeltaSimulatedPedigree,1,2)
```

`DeltaSimulatedPedigree` *Estimates of Jacquard Coefficients for Simulated SNP data according to a Pedigree*

Description

A list object containing nine matrices with pairwise Jacquard coefficients

Usage

```
data("DeltaSimulatedPedigree")
```

Format

A list with nine fields.

Details

`DeltaSimulatedPedigree` can be generated by applying function `Jacquard.cls` to the data in `SimulatedPedigree`.

References

Graffelman, J., Weir, B.S. and Goudet, J. (2024) Estimation of Jacquard's genetic identity coefficients with bi-allelic variants by constrained least-squares. Preprint at bioRxiv doi:10.1101/2024.03.25.586682.

Examples

```
data(DeltaSimulatedPedigree)
```

`GTC` *Joint Genotype Counts according to a Pedigree*

Description

Contains a list object with nine fields, consisting of nine lower triangular matrices with the joint genotype counts for 111 individuals.

Usage

```
data("GTC")
```

Format

A list with nine fields.

References

Goudet, J. (2022) JGTeach: JG Teaching material. R package version 0.1.9. https://github.com/jgx65

Graffelman, J., Weir, B.S. and Goudet, J. (2024) Estimation of Jacquard's genetic identity coefficients with bi-allelic variants by constrained least-squares. Preprint at bioRxiv doi:10.1101/2024.03.25.586682.

Examples

```
data(GTC)
```

`Jacquard.cls` *Estimation of Jacquard coefficients by Constrained Least Squares*

Description

Function `Jacquard.cls` estimates the nine condensed Jacquard coefficients of a pair of individuals using their joint genotype probabilities and the allele frequencies using constrained least squares.

Usage

```
Jacquard.cls(Xlist, mafvec = NULL, eps = 1e-06, delta.init = runif(9), Mavg = NULL,
             inner.iter = 1000, outer.iter = 1000, verbose = TRUE)
```

Arguments

`Xlist`	A list object with nine fields containing the matrices with joint genotype counts.
`mafvec`	A vector with the minor allele frequencies for all genetic variants.
`eps`	Tolerance criterion for the solver (`eps=1e-06` by default).
`delta.init`	Initial vector of estimates for the nine condensed Jacquard coefficients.
`Mavg`	A nine by nine matrix of conditional probabilities, allele frequency dependent. This matrix is calculated by `Jacquard.cls` if it is not specified.
`inner.iter`	Maximum number of inner iterations for the solver (1000 by default).
`outer.iter`	Maximum number of outer iterations for the solver (1000 by default).
`verbose`	Print output on the progress of the algorithm if `verbose=TRUE`.

Details

Function `Jacquard.cls` relies on the solver `solnp` from the Rsolnp package.

Value

A list object with fields:

`delta`	A list with nine matrices of estimates of pairwise Jacquard coefficients.
`convergence`	A matrix with the convergence status for each pair (0 = converged; 1 = not converged).

Author(s)

Jan Graffelman (jan.graffelman@upc.edu)

References

Graffelman, J., Weir, B.S. and Goudet, J. (2024) Estimation of Jacquard's genetic identity coefficients with bi-allelic variants by constrained least-squares. Preprint at bioRxiv doi:10.1101/2024.03.25.586682.

Ghalanos, A. and Theussl, S. (2015) Rsolnp: General Non-linear Optimization Using Augmented Lagrange Multiplier Method. R package version 1.16. https://cran.r-project.org/package=Rsolnp

See Also

`JointGenotypeCounts`

Examples

```
data(SimulatedPedigree)
Xgen <- as.matrix(SimulatedPedigree[,6:ncol(SimulatedPedigree)])

data(GTC)

mafvec <- mafvector(Xgen)

ii <- 1:3

GTCsubset <- list(length = 9)
for (k in 1:9) {
  GTCsubset[[k]] <- matrix(numeric(3^2), ncol = 3)
  GTCsubset[[k]] <- GTC[[k]][ii,ii]
}

output <- Jacquard.cls(GTCsubset,mafvec=mafvec,
                       eps=1e-06)
Delta.cls <- output$delta
print(Delta.cls)
print(output$convergence)

# A particular estimate of a Jacquard coefficient for a particular pair can
# be extracted from Delta.cls
#
# E.g., Delta\_9 of the first pair of individuals (1,2) can be extracted by
#

D9_12 <- Delta.cls[[9]][1,2]
```

`JointGenotypeCounts` *Calculate joint pairwise Genotype Counts*

Description

Function `JointGenotypeCounts` counts for each pair of individuals in the database their nine joint genotype counts

Usage

```
JointGenotypeCounts(X.gen, one.is.minor = TRUE)
```

Arguments

`X.gen`	A matrix with genotype data coded in (0,1,2) format
`one.is.minor`	If `TRUE` the genotype data represent the count of the minor allele for each marker, if `FALSE` the major allele.

Value

A list object with nine fields containing:

`f0000`	Matrix of (0/0,0/0) counts for all pairs
`f1111`	Matrix of (1/1,1/1) counts for all pairs
`f1101`	Matrix of (1/1,0/1) counts for all pairs
`f0111`	Matrix of (0/1,1/1) counts for all pairs
`f0101`	Matrix of (0/1,0/1) counts for all pairs
`f1100`	Matrix of (1/1,0/0) counts for all pairs
`f0011`	Matrix of (0/0,1/1) counts for all pairs
`f0100`	Matrix of (0/1,0/0) counts for all pairs
`f0001`	Matrix of (0/0,0/1) counts for all pairs

Author(s)

Jan Graffelman (jan.graffelman@upc.edu)

References

Graffelman, J., Weir, B.S. and Goudet, J. (2024) Estimation of Jacquard's genetic identity coefficients with bi-allelic variants by constrained least-squares. Preprint at bioRxiv. doi:10.1101/2024.03.25.586682

Examples

```
data(SimulatedPedigree)
JC <- JointGenotypeCounts(SimulatedPedigree[1:3,1:100])
print(JC)
```

`mafvector` *Calculate the Minor Allele Frequency (MAF) for a matrix with (0,1,2) genotype data*

Description

Function `mafvector` calculates genotype counts columnwise and determines the minor allele frequency for each column.

Usage

```
mafvector(X)
```

Arguments

`X`	A matrix of with (0,1,2) genotype data, individuals in rows, markers in columns.

Details

`mafvector` calculates the frequency of the minor allele irrespective of the coding; i.e., irrespective of whether the genotype data represent major or minor allele counts. Missing values are discarded for the calculation of the MAF.

Value

a vector

Author(s)

Jan Graffelman (jan.graffelman@upc.edu)

Examples

```
data(SimulatedPedigree)
p <- mafvector(SimulatedPedigree[,1:10])
print(p)
```

`MakeM` *Create matrix of conditional joint genotype probabilities*

Description

Function `MakeM` creates the matrix of conditional joint genotype probabilities for biallelic markers for a given allele frequency.

Usage

```
MakeM(p)
```

Arguments

`p`	the allele frequency

Value

a 9 by 9 matrix

Author(s)

Jan Graffelman (jan.graffelman@upc.edu)

References

Csuros, M. (2014) Non-identifiability of identity coefficients at biallelic loci. Theoretical Population Biology 92, pp. 22–29. doi: 10.1016/j.tpb.2013.11.001

Graffelman, J., Weir, B.S. and Goudet, J. (2024) Estimation of Jacquard's genetic identity coefficients with bi-allelic variants by constrained least-squares. Under review.

Examples

```
set.seed(123)
p <- runif(1)
M <- MakeM(p)
```

`PairwiseList` *Create a Pairwise List of Jacquard Coefficients*

Description

Function `PairwiseList` processes the list structure of Jacquard coefficients and converts it into to table of pairs of individuauls with their Jacquard coefficients.

Usage

```
PairwiseList(X, digits = 3)
```

Arguments

`X`	A list structure of nine matrices of Jacquard coefficients.
`digits`	Jacquard coefficients are rounded to the given number of digits.

Value

A matrix

Author(s)

Jan Graffelman (jan.graffelman@upc.edu)

See Also

`Jacquard.cls`

Examples

```
data(DeltaSimulatedPedigree)

ii <- 1:3

SubSet <- list(length = 9)
for (k in 1:9) {
  SubSet[[k]] <- matrix(numeric(3^2), ncol = 3)
  SubSet[[k]] <- DeltaSimulatedPedigree[[k]][ii,ii]
}

List <- PairwiseList(SubSet)

print(List)
```

`SimulatedPedigree` *Simulated SNP data according to a Pedigree*

Description

A matrix containing 111 individuals (rows), pedigree information (first five columns) and 20.000 single nucleotide polymorphisms (remaining columns) coded in (0,1,2) format and simulated according to a pedigree of 20 unrelated founders with six posterior generations.

Usage

```
data("SimulatedPedigree")
```

Format

A data frame containing 111 rows and 20.005 columns.

Details

The SNP data was generated using the JGTeach package. The genotype counts in the object SimulatedPedigree represent the counts of the minor allele.

References

Goudet, J. (2022) JGTeach: JG Teaching material. R package version 0.1.9. https://github.com/jgx65

Graffelman, J., Weir, B.S. and Goudet, J. (2024) Estimation of Jacquard's genetic identity coefficients with bi-allelic variants by constrained least-squares. Preprint at bioRxiv doi:10.1101/2024.03.25.586682.

Examples

```
data(SimulatedPedigree)
```

Index