Package 'LHD'

June 2, 2020

Type Package

Title Latin Hypercube Designs (LHDs) Algorithms

Version 1.1.0

Author Hongzhi Wang, Qian Xiao, Abhyuday Mandal

Maintainer Hongzhi Wang <hw34508@uga.edu>

Description Contains different algorithms for efficient Latin Hypercube Designs (LHDs) with flexible sizes. Our package is comprehensive since it is capable of generating nearly orthogonal LHDs, maximin distance LHDs, and maximum projection LHDs. Documentation for each algorithm includes useful information and explanation along with corresponding references. This package is particularly useful in the area of Design and Analysis of Experiments (DAE). More specifically, design of computer experiments.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 7.1.0

Suggests testthat, knitr, rmarkdown, devtools

Imports stats

VignetteBuilder knitr

NeedsCompilation no

Repository CRAN

Date/Publication 2020-06-02 04:30:02 UTC

R topics documented:

`AvgAbsCor` *Calculate the Average Absolute Correlation*

Description

`AvgAbsCor` returns the average absolute correlation of an LHD

Usage

```
AvgAbsCor(X)
```

Arguments

`X`	A Matrix.

Details

- `X` stands for the design matrix.
- The average absolute correlation formula is given by the Note Section below.

Value

If all inputs are logical, then the output will be a positive number indicating average absolute correlation.

Note

`average absolute correlation = \frac{2 \sum_{i=1}^{k-1} \sum_{j=i+1}^{k}|q_{ij}|}{k(k-1)}`

References

Georgiou, S. D. (2009) Orthogonal Latin hypercube designs from generalized orthogonal designs. *Journal of Statistical Planning and Inference*, **139**, 1530-1540.

Examples

```
#create a toy LHD with 5 rows and 3 columns
toy=rLHD(n=5,k=3);toy

#Calculate the average absolute correlation of toy
AvgAbsCor(X=toy)
```

---

dij *Calculate the Inter-site Distance*

Description

dij returns the inter-site distance of two design points of an LHD

Usage

```
dij(X, i, j, q = 1)
```

Arguments

X	A Matrix.
i	A positive integer.
j	A positive integer.
q	The default is set to be 1, and it could be either 1 or 2.

Details

- X stands for the design matrix.
- i stands for the i^th row of X.
- j stands for the j^th row of X.
- Both i and j should be in [1,nrow(X)] and they should not be equal to each other.
- If q is 1 (the default setting), dij is the rectangular distance. If q is 2, dij is the Euclidean distance.

Value

If all inputs are logical, then the output will be a positive number indicating the distance.

Examples

```
#create a toy LHD with 5 rows and 3 columns
toy=rLHD(n=5,k=3);toy

#Calculate the inter-site distance of the 2nd and the 4th row of toy (with default q)
dij(X=toy,i=2,j=4)

#Calculate the inter-site distance of the 2nd and the 4th row of toy (with q=2)
dij(X=toy,i=2,j=4,q=2)
```

---

exchange *Exchange two random elements*

Description

exchange returns a new design matrix after two randomly selected elements are switched from a user-defined column.

Usage

```
exchange(X, j)
```

Arguments

X	A Matrix.
j	A positive integer.

Details

- X stands for the design matrix.
- j is the j^th column of the design matrix, and it should be in [1,ncol(X)].

Value

If all inputs are logical, then the output will be a new design matrix after the exchange.

Examples

```
#create a toy LHD with 5 rows and 3 columns
toy=rLHD(n=5,k=3);toy

#Choose the first column of toy and exchange two randomly selected elements.
toy_new=exchange(X=toy,j=1)
toy;toy_new
```

GA *Genetic Algorithm for LHD*

Description

GA returns an LHD matrix generated by genetic algorithm (GA)

Usage

```
GA(n, k, m = 10, N = 10, pmut = 1/(k - 1), OC = "phi_p", p = 15, q = 1)
```

Arguments

n	A positive integer.
k	A positive integer.
m	A positive even integer.
N	A positive integer.
pmut	A probability.
OC	An optimality criterion.
p	A positive integer.
q	The default is set to be 1, and it could be either 1 or 2.

Details

- n stands for the number of rows (or run size).
- k stands for the number of columns (or the number of factors).
- m stands for the number of population and it must be an even number. The default is set to be 10.
- N stands for the number of iterations. The default is set to be 10.
- pmut stands for the probability of mutation. The default is set to be 1/(k - 1).
- OC stands for the optimality criterion, the default setting is "phi_p", and it could be one of the following: "phi_p", "AvgAbsCor", "MaxAbsCor", "MaxProCriterion".
- p is the parameter in the phi_p formula, and p is prefered to be large. The default is set to be 15.
- If q is 1 (the default setting), dij is the rectangular distance. If q is 2, dij is the Euclidean distance.

Value

If all inputs are logical, then the output will be a n by k LHD.

References

Liefvendahl, M., and Stocki, R. (2006) A study on algorithms for optimization of Latin hypercubes. *Journal of Statistical Planning and Inference*, **136**, 3231-3247.

Examples

```
#generate a 5 by 3 maximin distance LHD with the default setting
try=GA(n=5,k=3)
try
phi_p(try)   #calculate the phi_p of "try".

#Another example
#generate a 8 by 4 nearly orthogonal LHD
try2=GA(n=8,k=4,OC="AvgAbsCor")
try2
AvgAbsCor(try2)  #calculate the average absolute correlation.
```

LaPSO *Particle Swarm Optimization for LHD*

Description

LaPSO returns an LHD matrix generated by particle swarm optimization algorithm (PSO)

Usage

```
LaPSO(
  n,
  k,
  m = 10,
  N = 10,
  SameNumP = 0,
  SameNumG = n/4,
  p0 = 1/(k - 1),
  OC = "phi_p",
  p = 15,
  q = 1
)
```

Arguments

n	A positive integer.
k	A positive integer.
m	A positive integer.
N	A positive integer.
SameNumP	A non-negative integer.

SameNumG	A non-negative integer.
p0	A probability.
OC	An optimality criterion.
p	A positive integer.
q	The default is set to be 1, and it could be either 1 or 2.

Details

- n stands for the number of rows (or run size).
- k stands for the number of columns (or the number of factors).
- m stands for the number of particles. The default is set to be 10.
- N stands for the number of iterations. The default is set to be 10.
- SameNumP stands for how many elements in current column of current particle LHD should be the same as corresponding Personal Best. SameNumP=0, 1, 2, ..., n, and 0 means to skip the "exchange". The default is set to be 0.
- SameNumG stands for how many elements in current column of current particle LHD should be the same as corresponding Global Best. SameNumP=0, 1, 2, ..., n, and 0 means to skip the "exchange". The default is set to be n/4.
- SameNumP and SameNumG cannot be 0 at the same time.
- p0 stands the probability of exchange two randomly selected elements in current column of current particle LHD. The default is set to be 1/(k - 1).
- OC stands for the optimality criterion, the default setting is "phi_p", and it could be one of the following: "phi_p", "AvgAbsCor", "MaxAbsCor", "MaxProCriterion".
- p is the parameter in the phi_p formula, and p is prefered to be large. The default is set to be 15.
- If q is 1 (the default setting), dij is the rectangular distance. If q is 2, dij is the Euclidean distance.

Value

If all inputs are logical, then the output will be a n by k LHD.

Note

Here are some general suggestions about the parameters:

- SameNumP is approximately n/2 when SameNumG is 0.
- SameNumG is approximately n/4 when SameNumP is 0.
- p0 * (k - 1) = 1 or 2 is often sufficient. So p0 = 1/(k - 1) or 2/(k - 1).

References

Chen, R.-B., Hsieh, D.-N., Hung, Y., and Wang, W. (2013) Optimizing Latin hypercube designs by particle swarm. *Stat. Comput.*, **23**, 663-676.

Examples

```
#generate a 5 by 3 maximin distance LHD with the default setting
try=LaPSO(n=5,k=3)
try
phi_p(try)   #calculate the phi_p of "try".

#Another example
#generate a 8 by 4 nearly orthogonal LHD
try2=LaPSO(n=8,k=4,OC="AvgAbsCor")
try2
AvgAbsCor(try2)  #calculate the average absolute correlation.
```

`MaxAbsCor` *Calculate the Maximum Absolute Correlation*

Description

`MaxAbsCor` returns the maximum absolute correlation of an LHD

Usage

```
MaxAbsCor(X)
```

Arguments

`X`	A Matrix.

Details

- `X` stands for the design matrix.
- The maximum absolute correlation formula is given by the Note Section below.

Value

If all inputs are logical, then the output will be a positive number indicating maximum absolute correlation.

Note

`maximum absolute correlation` = `max_{ij} |q_{ij}|`

References

Georgiou, S. D. (2009) Orthogonal Latin hypercube designs from generalized orthogonal designs. *Journal of Statistical Planning and Inference*, **139**, 1530-1540.

Examples

```
#create a toy LHD with 5 rows and 3 columns
toy=rLHD(n=5,k=3);toy

#Calculate the maximum absolute correlation of toy
MaxAbsCor(X=toy)
```

MaxProCriterion *Calculate the Maximum Projection Criterion*

Description

MaxProCriterion returns the maximum projection criterion of an LHD

Usage

```
MaxProCriterion(X)
```

Arguments

X	A Matrix.

Details

- X stands for the design matrix.
- The maximum projection criterion formula is given by the Note Section below.

Value

If all inputs are logical, then the output will be a positive number indicating maximum projection criterion.

Note

maximum projection criterion = \Bigg{ \frac{1}{{n \choose 2}} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{1}{\Pi_{l=1}^{k}(x_{il}-x_{jl})^2} \Bigg}^{1/k}

References

Joseph, V. R., Gul, E., and Ba, S. (2015) Maximum projection designs for computer experiments. *Biometrika*, **102**, 371-380.

Examples

```
#create a toy LHD with 5 rows and 3 columns
toy=rLHD(n=5,k=3);toy

#Calculate the maximum projection criterion of toy
MaxProCriterion(X=toy)
```

OA2LHD *Transfer an Orthogonal Array (OA) into an LHD*

Description

OA2LHD transfers an OA into an LHD with corresponding size

Usage

```
OA2LHD(OA)
```

Arguments

OA	A Matrix.

Details

- OA stands for the input orthogonal array matrix.

Value

If the input is logical, then the output will be an LHD whose sizes are the same as input OA.

Note

The assumption is that the elements of OAs must be positive.

References

Tang, B. (1993) Orthogonal-array-based latin hypercubes. *Journal of the Americal Statistical Association*, **88**, 1392-1397.

Examples

```
#create an OA(9,2,3,2)
OA=matrix(c(rep(1:3,each=3),rep(1:3,times=3)),ncol=2,nrow=9,byrow = FALSE);OA

#Transfer the "OA" above into a LHD according to Tang (1993)
tryOA=OA2LHD(OA)
OA;tryOA
```

OASA *Orthogonal-Array-Based Simulated Annealing*

Description

OASA returns an LHD matrix generated by orthogonal-array-based simulated annealing algorithm (OASA)

Usage

```
OASA(
  OA,
  N = 10,
  T0 = 10,
  rate = 0.1,
  Tmin = 1,
  Imax = 5,
  OC = "phi_p",
  p = 15,
  q = 1
)
```

Arguments

OA	A Matrix.
N	A positive integer.
T0	A positive number.
rate	A positive percentage.
Tmin	A positive number.
Imax	A positive integer.
OC	An optimality criterion.
p	A positive integer.
q	The default is set to be 1, and it could be either 1 or 2.

Details

- OA stands for the input orthogonal array matrix.
- N stands for the number of iterations. The default is set to be 10.
- T0 stands for the user-defined initial temperature. The default is set to be 10.
- rate stands for temperature decrease rate, and it should be in (0,1). For example, rate=0.25 means the temperature decreases by 25% each time. The default is set to be 10%.
- Tmin stands for the minimium temperature allowed. When current temperature becomes smaller or equal to Tmin, the stopping criterion for current loop is met. The default is set to be 1.

- Imax stands for the maximum perturbations the algorithm will try without improvements before temperature is reduced. For the computation complexity consideration, Imax is recommended to be smaller or equal to 5, which is the default setting.
- OC stands for the optimality criterion, the default setting is "phi_p", and it could be one of the following: "phi_p", "AvgAbsCor", "MaxAbsCor", "MaxProCriterion".
- p is the parameter in the phi_p formula, and p is prefered to be large. The default is set to be 15.
- If q is 1 (the default setting), dij is the rectangular distance. If q is 2, dij is the Euclidean distance.

Value

If all inputs are logical, then the output will be an LHD whose sizes are the same as input OA.

Note

The assumption is that the elements of OAs must be positive.

References

Leary, S., Bhaskar, A., and Keane, A. (2003) Optimal orthogonal-array-based latin hypercubes. *Journal of Applied Statistics*, **30**, 585-598.

Examples

```
#create an OA(9,2,3,2)
OA=matrix(c(rep(1:3,each=3),rep(1:3,times=3)),ncol=2,nrow=9,byrow = FALSE);OA

#Use above "OA" as the input OA to generate a 9 by 2 maximin distance LHD
#with the default setting
tryOASA=OASA(OA=OA)
tryOASA
phi_p(tryOASA)   #calculate the phi_p of "tryOASA".

#Another example
#generate a 9 by 2 nearly orthogonal LHD
tryOASA2=OASA(OA=OA,OC="MaxAbsCor")
tryOASA2
MaxAbsCor(tryOASA2)  #calculate the maximum absolute correlation.
```

---

phi_p *Calculate the phi_p Criterion*

Description

phi_p returns the phi_p criterion of an LHD

Usage

```
phi_p(X, p = 15, q = 1)
```

Arguments

X	A Matrix.
p	A positive integer. The default is set to be 15.
q	The default is set to be 1, and it could be either 1 or 2.

Details

- X stands for the design matrix.
- p is the parameter in the phi_p formula (see Note Section below), and p is prefered to be large.
- If q is 1 (the default setting), dij is the rectangular distance. If q is 2, dij is the Euclidean distance.

Value

If all inputs are logical, then the output will be a positive number indicating phi_p.

Note

\phi_p = (\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}dij^{-p})^{1/p}

References

Jin, R., Chen, W., and Sudjianto, A. (2005) An efficient algorithm for constructing optimal design of computer experiments. *Journal of Statistical Planning and Inference*, **134**, 268-287.

Examples

```
#create a toy LHD with 5 rows and 3 columns
toy=rLHD(n=5,k=3);toy

#Calculate the phi_p criterion of toy with default setting
phi_p(X=toy)

#Calculate the phi_p criterion of toy with p=50 and q=2
phi_p(X=toy,p=50,q=2)
```

rLHD *Generate a random Latin Hypercube Design (LHD)*

Description

rLHD returns an Latin hypercube design matrix with user-defined dimention

Usage

```
rLHD(n, k)
```

Arguments

n	A positive integer.
k	A positive integer.

Details

- n stands for the number of rows (or run size).
- k stands for the number of columns (or the number of factors).

Value

If all inputs are positive integer, then the output will be a n by k design matrix.

Examples

```
#create a toy LHD with 5 rows and 3 columns
toy=rLHD(n=5,k=3);toy

#another example with 9 rows and 2 columns
rLHD(9,2)
```

SA *Simulated Annealing for LHD*

Description

SA returns an LHD matrix generated by simulated annealing algorithm (SA)

Usage

```
SA(
  n,
  k,
  N = 10,
  T0 = 10,
  rate = 0.1,
  Tmin = 1,
  Imax = 5,
  OC = "phi_p",
  p = 15,
  q = 1
)
```

Arguments

n	A positive integer.
k	A positive integer.
N	A positive integer.
T0	A positive number.
rate	A positive percentage.
Tmin	A positive number.
Imax	A positive integer.
OC	An optimality criterion.
p	A positive integer.
q	The default is set to be 1, and it could be either 1 or 2.

Details

- n stands for the number of rows (or run size).
- k stands for the number of columns (or the number of factors).
- N stands for the number of iterations. The default is set to be 10.
- T0 stands for the user-defined initial temperature. The default is set to be 10.
- rate stands for temperature decrease rate, and it should be in (0,1). For example, rate=0.25 means the temperature decreases by 25% each time. The default is set to be 10%.
- Tmin stands for the minimium temperature allowed. When current temperature becomes smaller or equal to Tmin, the stopping criterion for current loop is met. The default is set to be 1.
- Imax stands for the maximum perturbations the algorithm will try without improvements before temperature is reduced. For the computation complexity consideration, Imax is recommended to be smaller or equal to 5, which is the default setting.
- OC stands for the optimality criterion, the default setting is "phi_p", and it could be one of the following: "phi_p", "AvgAbsCor", "MaxAbsCor", "MaxProCriterion".

- p is the parameter in the phi_p formula, and p is prefered to be large. The default is set to be 15.
- If q is 1 (the default setting), dij is the rectangular distance. If q is 2, dij is the Euclidean distance.

Value

If all inputs are logical, then the output will be a n by k LHD.

References

Morris, M.D., and Mitchell, T.J. (1995) Exploratory designs for computer experiments. *Journal of Statistical Planning and Inference*, **43**, 381-402.

Examples

```
#generate a 5 by 3 maximin distance LHD with the default setting
try=SA(n=5,k=3)
try
phi_p(try)   #calculate the phi_p of "try".

#Another example
#generate a 8 by 4 nearly orthogonal LHD
try2=SA(n=8,k=4,OC="AvgAbsCor")
try2
AvgAbsCor(try2)  #calculate the average absolute correlation.
```

SA2008 *Simulated Annealing for LHD with Multi-objective Optimization Approach*

Description

SA2008 returns an LHD matrix generated by simulated annealing algorithm with multi-objective optimization approach.

Usage

```
SA2008(
  n,
  k,
  N = 10,
  T0 = 10,
  rate = 0.1,
  Tmin = 1,
  Imax = 5,
  OC = "phi_p",
  p = 15,
  q = 1
)
```

Arguments

n	A positive integer.
k	A positive integer.
N	A positive integer.
T0	A positive number.
rate	A positive percentage.
Tmin	A positive number.
Imax	A positive integer.
OC	An optimality criterion.
p	A positive integer.
q	The default is set to be 1, and it could be either 1 or 2.

Details

- n stands for the number of rows (or run size).
- k stands for the number of columns (or the number of factors).
- N stands for the number of iterations. The default is set to be 10.
- T0 stands for the user-defined initial temperature. The default is set to be 10.
- rate stands for temperature decrease rate, and it should be in (0,1). For example, rate=0.25 means the temperature decreases by 25% each time. The default is set to be 10%.
- Tmin stands for the minimium temperature allowed. When current temperature becomes smaller or equal to Tmin, the stopping criterion for current loop is met. The default is set to be 1.
- Imax stands for the maximum perturbations the algorithm will try without improvements before temperature is reduced. For the computation complexity consideration, Imax is recommended to be smaller or equal to 5, which is the default setting.
- OC stands for the optimality criterion, the default setting is "phi_p", and it could be one of the following: "phi_p", "AvgAbsCor", "MaxAbsCor", "MaxProCriterion".
- p is the parameter in the phi_p formula, and p is prefered to be large. The default is set to be 15.
- If q is 1 (the default setting), dij is the rectangular distance. If q is 2, dij is the Euclidean distance.

Value

If all inputs are logical, then the output will be a n by k LHD.

Note

This modified simulated annealing algorithm reduces columnwise correlations and maximizes minimum distance between design points simultaneously, with a cost of more computational complexity.

References

Joseph, V.R., and Hung, Y. (2008) Orthogonal-maximin Latin hypercube designs. *Statistica Sinica*, **18**, 171-186.

Examples

```
#generate a 5 by 3 maximin distance LHD with the default setting
try=SA2008(n=5,k=3)
try
phi_p(try)   #calculate the phi_p of "try".

#Another example
#generate a 8 by 4 nearly orthogonal LHD
try2=SA2008(n=8,k=4,OC="AvgAbsCor")
try2
AvgAbsCor(try2)  #calculate the average absolute correlation.
```

SLHD *Sliced Latin Hypercube Design (SLHD)*

Description

SLHD returns an LHD matrix generated by "improved two-stage algorithm" from Ba et al. (2015).

Usage

```
SLHD(
  n,
  k,
  t = 1,
  N = 10,
  T0 = 10,
  rate = 0.1,
  Tmin = 1,
  Imax = 3,
  OC = "phi_p",
  p = 15,
  q = 1,
  stage2 = FALSE
)
```

Arguments

n	A positive integer.
k	A positive integer.
t	A positive integer.
N	A positive integer.

T0	A positive number.
rate	A positive percentage.
Tmin	A positive number.
Imax	A positive integer.
OC	An optimality criterion.
p	A positive integer.
q	The default is set to be 1, and it could be either 1 or 2.
stage2	The default is set to be FALSE, and it could be either FALSE or TRUE.

Details

- n stands for the number of rows (or run size).
- k stands for the number of columns (or the number of factors).
- t stands for the number of slices. n/t must be an integer, that is, n is divisible by t. t must not exceed k when n is 9 or larger, and t must be smaller than k when n is smaller than 9. Otherwise, the funtion will never stop. The default is set to be 1.
- N stands for the number of iterations. The default is set to be 10.
- T0 stands for the user-defined initial temperature. The default is set to be 10.
- rate stands for temperature decrease rate, and it should be in (0,1). For example, rate=0.25 means the temperature decreases by 25% each time. The default is set to be 10%.
- Tmin stands for the minimium temperature allowed. When current temperature becomes smaller or equal to Tmin, the stopping criterion for current loop is met. The default is set to be 1.
- Imax stands for the maximum perturbations the algorithm will try without improvements before temperature is reduced. For the computation complexity consideration, Imax is recommended to be smaller or equal to 3, which is the default setting.
- OC stands for the optimality criterion, the default setting is "phi_p", and it could be one of the following: "phi_p", "AvgAbsCor", "MaxAbsCor", "MaxProCriterion".
- p is the parameter in the phi_p formula, and p is prefered to be large. The default is set to be 15.
- If q is 1 (the default setting), dij is the rectangular distance. If q is 2, dij is the Euclidean distance.
- If stage2 is FALSE (the default setting), SLHD will only implement the first stage of the algorithm. If stage2 is TRUE, SLHD will implement the whole algorithm.

Value

If all inputs are logical, then the output will be a n by k LHD.

Note

As mentioned from the original paper, the first stage plays a much more important role since it optimizes the slice level. More resources should be given to the first stage if computational budgets are limited. Let m=n/t, where m is the number of rows for each slice, if (m)^k » n, the second stage becomes optional. That is the reason why we add a stage2 parameter to let users decide if the second stage is needed.

References

Ba, S., Myers, W.R., and Brenneman, W.A. (2015) Optimal Sliced Latin Hypercube Designs. *Technometrics*, **57**, 479-487.

Examples

```
#generate a 5 by 3 maximin distance LHD with the default setting
trySLHD1=SLHD(n=5,k=3)
trySLHD1
phi_p(trySLHD1)   #calculate the phi_p of "trySLHD1".

#generate a 5 by 3 maximin distance LHD with stage II
#let stage2=TRUE and other input are the same as above
trySLHD2=SLHD(n=5,k=3,stage2=TRUE)
trySLHD2
phi_p(trySLHD2)   #calculate the phi_p of "trySLHD2".

#Another example
#generate a 8 by 4 nearly orthogonal LHD
trySLHD3=SLHD(n=8,k=4,OC="AvgAbsCor",stage2=TRUE)
trySLHD3
AvgAbsCor(trySLHD3)  #calculate the average absolute correlation.
```

Index