Package ‘DiceDesign’

February 19, 2015

Type Package
Title Designs of Computer Experiments
Version 1.6
Date 2014-05-06
Author Jessica Franco, Delphine Dupuy, Olivier Roustant, Guillaume Damblin and Bertrand Iooss.
Maintainer Celine Helbert <Celine.Helbert@ec-lyon.fr>
Description Space-Filling Designs and Uniformity Criteria.
License GPL-3
Suggests rgl, randtoolbox, lattice
Encoding latin1
URL http://dice.emse.fr/
NeedsCompilation yes
Repository CRAN
Date/Publication 2014-05-09 23:59:42

R topics documented:

DiceDesign-package .................................................. 2
coverage .............................................................. 4
discrepancyCriteria .............................................. 5
discrepESE_LHS .................................................... 7
discrepSA_LHS .................................................... 9
dmaxDesign .......................................................... 11
factDesign ........................................................... 12
lhsDesign ............................................................ 13
maximinESE_LHS ................................................... 14
maximinSA_LHS .................................................... 16
meshRatio ........................................................... 18
mindist .............................................................. 19
mstCriteria .......................................................... 20
Description

Space-Filling Designs (SFD) and space-filling criteria (distance-based and uniformity-based).

Details

<table>
<thead>
<tr>
<th>Package:</th>
<th>DiceDesign</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type:</td>
<td>Package</td>
</tr>
<tr>
<td>Version:</td>
<td>1.2</td>
</tr>
<tr>
<td>Date:</td>
<td>2013-26-12</td>
</tr>
<tr>
<td>License:</td>
<td>GPL-3</td>
</tr>
</tbody>
</table>

This package provides tools to create some specific Space-Filling Design (SFD) and to test their quality:

- Latin Hypercube Design (randomized or centered)
- Strauss SFD and Maximum entropy SFD, WSP design
- Optimal (low-discrepancy and maximin) Latin Hypercube Design by simulated annealing and genetic algorithms,
- Discrepancies criteria, distance measures,
- Minimal spanning tree criteria,
- Radial scanning statistic

Note

Part of this work was conducted within the frame of the DICE (Deep Inside Computer Experiments) Consortium between ARMINES, Renault, EDF, IRSN, ONERA and TOTAL S.A. (http://dice.emse.fr/).

In this package only Faure’s sequence is implemented. Note that the randtoolbox package provides the following quasi random sequences: the Sobol sequence, the Halton (hence Van Der Corput)
sequence and the Torus sequence (also known as Kronecker sequence). Note also that the \texttt{lhs} package provides other types of algorithms to compute optimized LHS.

**Author(s)**

J. Franco, D. Dupuy, O. Roustant, G. Damblin and B. Iooss. Thanks to A. Jourdan for discussions about OA131.

(maintainer: Celine Helbert <Celine.Helbert@ec-lyon.fr>)

**References**


**Examples**

```r
# *******************************************************
# Designs of experiments
# *******************************************************

# A maximum entropy design with 20 points in [0,1]^2
p <- dmaxDesign(20,2,0.9,200)
plot(p$design,xlim=c(0,1),ylim=c(0,1))

# *******************************************************
# Criteria: L2-discrepancy
# *******************************************************

dp <- discrepancyCriteria(p$design,type=c('L2','C2'))
# Coverage measure
covp <- coverage(p$design)

# *******************************************************
# Criteria: Minimal Spanning Tree
# *******************************************************
mstCriteria(p$design,plot2d=TRUE)

# *******************************************************
# Radial scanning statistic: Detection of defects of Sobol designs
# *******************************************************

# requires randtoolbox package
library(randtoolbox)
```
coverage

Description

Compute the coverage measure

Usage

coverage(design)

Arguments

design a matrix (or a data.frame) representing the design of experiments representing the design of experiments in the unit cube $[0,1]^d$. If this last condition is not fulfilled, a transformation into $[0,1]^d$ is applied before the computation of the criteria.

Details

The coverage criterion is defined by

$$coverage = \frac{1}{\bar{\gamma}} \left[ \frac{1}{n} \sum_{i=1}^{n} (\gamma_i - \bar{\gamma})^2 \right]^{1/2}$$

where $\gamma_i$ is the minimal distance between the point $x_i$ and the other points of the design and $\bar{\gamma}$ is the mean of the $\gamma_i$.

Note that for a regular mesh, cov=0. Then, a small value of cov means that the design is close to a regular grid.
discrepancyCriteria

Value
A real number equal to the value of the coverage criterion for the design.

Author(s)
J. Franco

References

See Also
other distance criteria like meshRatio, phiP and mindist.
discrepancy measures provided by discrepancyCriteria.

Examples

dimension <- 2
n <- 40
X <- matrix(runif(n*dimension),n,dimension)
 coverage(X)

discrepancyCriteria   Discrepancy measure

Description
Compute discrepancy criteria.

Usage
discrepancyCriteria(design,type='all')

Arguments
design a matrix (or a data.frame) corresponding to the design of experiments. The discrepancy criteria are computed for a design in the unit cube [0,1]^d. If this condition is not satisfied the design is automatically rescaled.
type type of discrepancies (single value or vector) to be computed:

'all'   all type of discrepancies (default)
'C2'    centered L2-discrepancy
'L2'    L2-discrepancy
'L2star' L2star-discrepancy
'M2'    modified L2-discrepancy
The discrepancy measures how far a given distribution of points deviates from a perfectly uniform one. Different L2 discrepancies are available in DiceDesign. For example, if we denote by $\text{Vol}(J)$ the volume of a subset $J$ of $[0; 1]^d$ and $A(X; J)$ the number of points of $X$ falling in $J$, the $L_2$ discrepancy is:

$$D_{L_2}(X) = \left[ \int_{[0,1]^d} \left( \frac{A(X,J_{a,b})}{n} - \text{Vol}(J_{a,b}) \right)^2 \, dadb \right]^{1/2}$$

where $a = (a_1;...;a_d)'$, $b = (b_1;...;b_d)'$ and $J_{a,b} = [a_1;b_1) \times ... \times [a_d;b_d)$. The other L2 discrepancies are defined according to the same principle with different form from the subset $J$. Among all the possibilities, discrepancyCriteria implements only the L2 discrepancies because it can be expressed analytically even for high dimension.

Centered L2-discrepancy is computed using the analytical expression done by Hickernell (1998). The user will refer to Pleming and Manteufel (2005) to have more details about the wrap around discrepancy.

**Value**

A list containing the L2-discrepancies of the design.

**Author(s)**

J. Franco, D. Dupuy & B. Iooss

**References**


**See Also**

distance criteria(*coverage, meshRatio, mindist* and *phiP*)
Examples

```r
dimension <- 2
n <- 40
X <- matrix(runif(n*dimension),n,dimension)
discrepancyCriteria(X)
```

**Description**

The objective is to produce low-discrepancy LHS. ESE is a powerful genetic algorithm to produce space-filling designs. It has been adapted here to main discrepancy criteria.

**Usage**

```r
discrepESE_LHS(design,T0=0.005*discrepancyCriteria(design,type='C2')[[1]],
            inner_it=100,J=50,it=2,criterion='C2')
```

**Arguments**

- `design`: a matrix (or a data.frame) corresponding to the design of experiments.
- `T0`: The initial temperature of the ESE algorithm.
- `inner_it`: The number of iterations for inner loop.
- `J`: The number of new proposed LHS inside the inner loop.
- `it`: The number of iterations for outer loop.
- `criterion`: The criterion to be optimized. One can choose three different L2-discrepancies: the C2 (centered) discrepancy ("C2"), the L2-star discrepancy ("L2star") and the W2 (wrap-around) discrepancy ("W2").

**Details**

This function implements a stochastic algorithm (ESE) to produce optimized LHS. It is based on Jin et al works (2005). Here, it has been adapted to some discrepancy criteria taking into account new ideas about the revaluations of discrepancy value after a LHS elementary perturbation (in order to avoid computing all terms in the discrepancy formulas).

**Value**

A list containing:

- `InitialDesign`: the starting design.
- `T0`: the initial temperature of the ESE algorithm.
- `inner_it`: the number of iterations for inner loop.
J  the number of new proposed LHS inside the inner loop
it  the number of iterations for outer loop
criterion  the criterion to be optimized
design  the matrix of the final design (low-discrepancy LHS)
critValues  vector of criterion values along the iterations
tempValues  vector of temperature values along the iterations
probaValues  vector of acceptation probability values along the iterations

**Author(s)**
G. Damblin & B. Iooss

**References**


**See Also**

Latin Hypercube Sample(*lhsDesign*), discrepancy criteria(*discrepancyCriteria*), geometric criterion(*mindistphi*), optimization(*maximinSA_LHS*, *maximise_LHS*, *discrepSA_LHS*)

**Examples**

```r
## Not run:
dimension <- 2
n <- 10
X <- lhsDesign(n,dimension)$design
Xopt <- discrepESE_LHS(X,T0=0.005*discrepancyCriteria(X,type='C2')[[1]],inner_it=100,J=50,iter=2)
plot(Xopt$design)
plot(Xopt$critValues,type="l")

## End(Not run)
```
discrepSA_LHS

Simulated annealing (SA) routine for Latin Hypercube Sample (LHS) optimization via L2-discrepancy criteria

Description

The objective is to produce low-discrepancy LHS. SA is an efficient algorithm to produce space-filling designs. It has been adapted here to main discrepancy criteria.

Usage

`discrepSA_LHS(design, T0=10, c=0.95, it=2000, criterion="C2", profile="GEOM", Imax=100)`

Arguments

design: a matrix (or a data.frame) corresponding to the design of experiments
T0: The initial temperature
c: A constant parameter regulating how the temperature goes down
it: The number of iterations
criterion: The criterion to be optimized. One can choose three different L2-discrepancies: the C2 (centered) discrepancy ("C2"), the L2-star discrepancy ("L2star") and the W2 (wrap-around) discrepancy ("W2")
profile: The temperature down-profile, purely geometric called "GEOM", geometrical according to the Morris algorithm called "GEOM_MORRIS" or purely linear called "LINEAR"
Imax: A parameter given only if you choose the Morris down-profile. It adjusts the number of iterations without improvement before a new elementary perturbation

Details

This function implements a classical routine to produce optimized LHS. It is based on the work of Morris and Mitchell (1995). They have proposed a SA version for LHS optimization according to mindist criterion. Here, it has been adapted to some discrepancy criteria taking in account new ideas about the reevaluations of a discrepancy value after a LHS elementary perturbation (in order to avoid computing all terms in the discrepancy formulas).

Value

A list containing:

- InitialDesign: the starting design
- T0: the initial temperature of the SA algorithm
- c: the constant parameter regulating how the temperature goes down
- it: the number of iterations
criterion the criterion to be optimized
profile the temperature down-profile
Imax The parameter given in the Morris down-profile
design the matrix of the final design (low-discrepancy LHS)
critValues vector of criterion values along the iterations
tempValues vector of temperature values along the iterations
probaValues vector of acceptation probability values along the iterations

Author(s)
G.Damblin & B. Iooss

References


See Also
Latin Hypercube Sample(lhsDesign), discrepancy criteria(discrepancyCriteria), geometric criterion (mindistphiP), optimization (maximINSA_LHS,maximinESE_LHS,discrepESE_LHS)

Examples

dimension <- 2
n <- 10
X <- lhsDesign(n,dimension)$design
Xopt <- discrepSA_LHS(X,T0=10,c=0.99,it=2000,criterion="C2")
plot(Xopt$design)
plot(Xopt$critValues,type="l")
## Not run:
Xopt <- discrepSA_LHS(X,T0=10,c=0.99,it=1000,criterion="C2",profile="GEOM_MORRIS")
## End(Not run)
**dmaxDesign**

**Maximum Entropy Designs**

**Description**

Space-Filling Designs with \( n \) experiments based on covariance matrix in \([0,1]^d\).

**Usage**

\[
dmaxDesign(n, \text{dimension}, \text{range}, \text{niter\_max}=1000, \text{seed}=\text{NULL})
\]

**Arguments**

- \( n \) number of experiments
- \( \text{dimension} \) number of variables
- \( \text{range} \) range of variogram
- \( \text{niter\_max} \) number of iterations
- \( \text{seed} \) seed used to generate uniform design

**Details**

Maximum entropy design is a kind of optimal design based on Shannon’s definition of entropy as the amount of information. Originally, maximum entropy sampling was proposed by Shewry and Wynn (1987). The goal of the design is to maximize the entropy defined as the determinant of the correlation matrix using a Fedorov-Mitchell exchange algorithm.

The spatial correlation matrix is defined by \( C = (\rho_{ij}) \):

\[
\rho_{ij} = \begin{cases} 
1 - \gamma (h_{ij}) & \text{if } h_{ij} \leq a, \\
0 & \text{if } h_{ij} > a,
\end{cases}
\]

where \( h_{ij} \) is the distance between \( x_i \) and \( x_j \), \( a \) denotes the range of the variogram and \( \gamma \) is a spherical variogram:

\[
\gamma(h) = 1.5 \frac{h}{a} - 0.5 \left( \frac{h}{a} \right)^3 \text{ for } h \leq a
\]

**Value**

A list with components:

- \( n \) the number of points
- \( \text{design} \) the design of experiments
- \( \text{dimension} \) the number of variables
- \( \text{range} \) the range of the variogram
- \( \text{niter\_mx} \) the number of iterations
The `factDesign` function creates a factorial design with \( n = \text{pow}(\text{levels}, \text{dimension}) \) experiments in \([0,1]^d\).

**Description**

Create a factorial design with \( n = \text{pow}(\text{levels}, \text{dimension}) \) experiments in \([0,1]^d\).

**Usage**

```
factDesign(dimension, levels)
```

**Arguments**

- `dimension`: an integer given the number of input variables
- `levels`: an integer given the number of levels

**Details**

It is possible to take a different number of levels for any factor. In this case, the argument `levels` should be a vector.
Value

factDesign returns a list containing all the input arguments detailed before, plus the following components:

- \( n \) the number of experiments
- \( \text{design} \) the design of experiments

Author(s)

G. Pujol and J. Franco

Examples

```r
## First example
g <- factDesign(2,7)
plot(g$design,xlim=c(0,1),ylim=c(0,1))
## Second example
g <- factDesign(2,c(2,7))
plot(g$design,xlim=c(0,1),ylim=c(0,1))
```

\section*{lhsDesign Latin Hypercube Designs}

Description

Simple (random) Latin Hypercube Design (randomized or centered) with \( n \) experiments in \([0,1]^d\).

Usage

\[ \text{lhsDesign}(n, \text{dimension}, \text{randomized=TRUE, seed=NULL}) \]

Arguments

- \( n \) number of experiments
- \( \text{dimension} \) number of variables
- \( \text{randomized} \) TRUE for randomized LHS; FALSE for centered LHS
- \( \text{seed} \) seed used to generate the random permutations and perturbations

Details

This program builds a Latin Hypercube Design (LHD), also called a Latin Hypercube Sample (LHS), on the space \([0,1]^d\) (with uniform probability measures). LHD aims at ensuring that each variable has its whole range well scanned: the range of each variable is divided into \( n \) equally probable strata. Each stratum of each variable contains only one point of the LHD. Centered LHD is obtained by choosing for each point the center of the corresponding case, while randomized LHD is obtained by adding random perturbations inside each point case.

Once the sample is generated, the uniform sample from a column can be transformed to any distribution by using the quantile functions.
Value

A list with components:

- `n` the number of points
- `dimension` the number of variables
- `design` the design of experiments
- `randomized` the type of LHD
- `seed` the value of the seed

Author(s)

B. Iooss

References

McKay M., Conover W. and Beckman R. (1979) A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, Technometrics, 21, 2, 239-245.


See Also

LHD optimization (maximinsa_lhs, discrepsa_lhs, maximinESE_LHS, discrepESE_LHS)

Examples

```r
n <- 20
dimension <- 2
out <- lhsDesign(n, dimension)
```

<table>
<thead>
<tr>
<th>maximinESE_LHS</th>
<th>Enhanced Stochastic Evolutionnary (ESE) algorithm for Latin Hypercube Sample (LHS) optimization via phiP criteria</th>
</tr>
</thead>
</table>

Description

The objective is to produce maximin LHS. ESE is a powerful genetic algorithm allowing to produce space-filling designs.

Usage

```r
maximinESE_LHS(design,T0=0.005*phiP(design,p=50),inner_it=100,J=50, it=1,p=50)
```
Arguments

design      a matrix (or a data.frame) corresponding to the design of experiments.
T0          The initial temperature of the ESE algorithm
inner_it    The number of iterations for inner loop
J           The number of new proposed LHS inside the inner loop
it          The number of iterations for outer loop
p           power required in phiP criterion

Details

This function implements a stochastic algorithm (ESE) to produce optimized LHS. It is based on Jin et al works (2005).

Value

A list containing:

InitialDesign  the starting design
T0            the initial temperature of the ESE algorithm
inner_it      the number of iterations for inner loop
J             the number of new proposed LHS inside the inner loop
it            the number of iterations for outer loop
p             power required in phiP criterion
design        the matrix of the final design (maximin LHS)
critValues    vector of criterion values along the iterations
tempValues    vector of temperature values along the iterations
probaValues   vector of acceptance probability values along the iterations

Author(s)

G.Damblin & B. Iooss

References


maximinSA_LHS

Simulated annealing (SA) routine for Latin Hypercube Sample (LHS) optimization via phiP criteria

Description
The objective is to produce maximin LHS. SA is an efficient algorithm to produce space-filling designs.

Usage
maximinSA_LHS(design,T0=10,c=0.95,it=2000,p=50,profile="GEOM",Imax=100)

Arguments
- **design**: a matrix (or a data.frame) corresponding to the design of experiments
- **T0**: The initial temperature of the SA algorithm
- **c**: A constant parameter regulating how the temperature goes down
- **it**: The number of iterations
- **p**: power required in phiP criterion
- **profile**: The temperature down-profile, purely geometric called "GEOM", geometrical according to the Morris algorithm called "GEOM_MORRIS" or purely linear called "LINEAR"
- **Imax**: A parameter given only if you choose the Morris down-profile. It adjusts the number of iterations without improvement before a new elementary perturbation

Details
This function implements a classical routine to produce optimized LHS. It is based on the work of Morris and Mitchell (1995). They have proposed a SA version for LHS optimization according to mindist criterion. Here, it has been adapted to the phiP criterion. It has been shown (Pronzato and Muller, 2012, Damblin et al., 2013) that optimizing phiP is more efficient to produce maximin designs than optimizing mindist. When \( p \) tends to infinity, optimizing a design with phi\(_p\) is equivalent to optimizing a design with mindist.
Value

A list containing:

- `InitialDesign`: the starting design
- `T0`: the initial temperature of the SA algorithm
- `c`: the constant parameter regulating how the temperature goes down
- `it`: the number of iterations
- `p`: power required in phiP criterion
- `profile`: the temperature down-profile
- `Imax`: The parameter given in the Morris down-profile
- `design`: the matrix of the final design (maximin LHS)
- `critValues`: vector of criterion values along the iterations
- `tempValues`: vector of temperature values along the iterations
- `probaValues`: vector of acceptation probability values along the iterations

Author(s)

G. Damblin & B. Iooss

References


See Also

Latin Hypercube Sample(`lhsDesign`), discrepancy criteria(`discrepancyCriteria`), geometric criterion (`mindistphiP`), optimization (`discrepSA_LHS,maximinESE_LHS,discrepESE_LHS`)

Examples

dimension <- 2
n <- 10
X <- lhsDesign(n,dimension)$design
Xopt <- maximinSA_LHS(X,T0=10,c=0.99,it=2000)
plot(Xopt$design)
plot(Xopt$critValues,type="l")
plot(Xopt$tempValues,type="l")
## Not run:
Xopt <- maximinSA_LHS(X,T0=10,c=0.99,it=1000,profile="GEOM_MORRIS")
## End(Not run)

table

table

meshRatio | MeshRatio measure

Description

The meshRatio criterion is the ratio between the maximum and the minimum distance between two points of the experimental design.

Usage

meshRatio(design)

Arguments

design: a matrix (or a data.frame) representing the design of experiments in the unit cube [0,1]^d. If this last condition is not fulfilled, a transformation into [0,1]^d is applied before the computation of the criteria.

Details

The meshRatio criterion is defined by

\[
\text{meshRatio} = \frac{\max_{1 \leq i \leq n} \gamma_i}{\min_{1 \leq i \leq n} \gamma_i}
\]

where \(\gamma_i\) denotes the minimal distance between the point \(x_i\) and the other points of the design. Note that for a regular mesh, meshRatio=1.

Value

A real number equal to the value of the meshRatio criterion for the design.

Author(s)

J. Franco

References


See Also

other distance criteria like meshRatio, phiP and mindist; discrepancy measures provided by discrepancyCriteria.
Examples

```r
dimension <- 2
n <- 40
X <- matrix(runif(n*dimension),n,dimension)
meshRatio(X)
```

<table>
<thead>
<tr>
<th>mindist</th>
<th>Mindist measure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Description

Compute the mindist criterion (also called maximin)

Usage

```r
mindist(design)
```

Arguments

- `design`: a matrix (or a data.frame) representing the design of experiments in the unit cube $[0,1]^d$. If this last condition is not fulfilled, a transformation into $[0,1]^d$ is applied before the computation of the criteria.

Details

The mindist criterion is defined by

$$mindist = \min_{x_i \in X} (\gamma_i)$$

where $\gamma_i$ is the minimal distance between the point $x_i$ and the other points $x_k$ of the design.

A higher value corresponds to a more regular scattering of design points.

Value

A real number equal to the value of the mindist criterion for the design.

Author(s)

- J. Franco

References


See Also

other distance criteria like meshRatio and phiP, discrepancy measures provided by discrepancyCriteria.

Examples

dimension <- 2
n <- 40
X <- matrix(runif(n*dimension),n,dimension)
mdist(X)

mstCriteria Deriving the MST criteria

Description

Compute both the mean and the standard deviation of the Minimal Spanning Tree (MST)

Usage

mstCriteria(design,plot2d="FALSE")

Arguments

design a matrix (or a data.frame) corresponding to the design of experiments.

plot2d an argument for visualizing the mst of a 2d design

Details

In our context, a MST is a tree whose the sum of the lengths of the edges is minimal. Even if unicity does not hold, the overall length is stable. The mean and the standard deviation of the lengths of the edges are usually derived to analyze the geometric profile of the design. A large mean and a small standard deviation characterize a so-called quasi-periodic design.

Value

A list containing two components:

tree a list containing the MST: each component of it contains a vector with all vertices which are connected with the experiment corresponding to the number of the components

stats vector with both the mean and the standard deviation values of the lengths of the edges

Author(s)

G.Damblin & B.Iooss
OA131

References


Examples

dimension <- 2
n <- 40
X <- matrix(runif(n*dimension),n,dimension)
mstCriteria(X,plot2d=TRUE)

---

OA131

A 3D orthogonal array of strength 2

Description

A 3-dimensional linear orthogonal array (OA) of strength 2 with 49 points. The design points are equally spaced into 2 dimensional coordinate planes. However by construction, such OAs satisfy a linear relation, here: \( x_1 + 3x_2 + x_3 = 0 \mod 7 \). As a consequence, the design points are contained in parallel planes orthogonal to \((1,3,1)\). Actually, they are also contained in parallel planes orthogonal to other directions, as \((2,-1,2)\) or \((3,2,3)\), since the congruence relation leads to \(2x_1 - x_2 + 2x_3 = 0 \mod 7\) or \(3x_1 + 2x_2 + 3x_3 = 0 \mod 7\). For instance, they are contained in 4 parallel planes orthogonal to \((2,-1,2)\).

Usage

data(OA131)

Format

A data frame with 49 observations on the following 3 variables.

- \(x_1\) first coordinate
- \(x_2\) second coordinate
- \(x_3\) third coordinate
Source

Examples

data(OA131)

# centering and reducing to [0,1]^3
OA <- (OA131 + 0.5)/7

pairs(OA, xlim=c(0,1), ylim=c(0,1))

## Not run: library(lattice)
cloud(x3~x1+x2, data=OA, xlim=c(0,1), ylim=c(0,1), zlim=c(0,1),
    screen = list(z = 50, x = -70, y = 0))
## End(Not run)

OA131_scrambled A scrambled 3D orthogonal array of strength 2

Description
This design is obtained by adding a uniform noise to each coordinate of the orthogonal array OA131.

Usage
data(OA131_scrambled)

Format
A data frame with 49 observations on the following 3 variables.

x1 first coordinate
x2 second coordinate
x3 third coordinate

Source

Examples

data(OA131)
data(OA131_scrambled)
pairs(OA131, xlim=c(0,1), ylim=c(0,1))
pairs(OA131_scrambled, xlim=c(0,1), ylim=c(0,1))
The $\phi_p$ criterion is defined by the $L_p$ norm of the sum of the inverses of the design inter-point euclidean distances:

$$\phi_p = \left[ \sum_{i,j=1...N, i<j} d_{ij}^{-p} \right]^{\frac{1}{p}}$$

A higher value corresponds to a more regular scaterring of design points.

When $p$ tends to infinity, optimizing a design with $\phi_p$ is equivalent to optimizing a design with mindist.
See Also
general criterion \(\text{mindist}\)

Examples
dimension <- 2
n <- 40
X <- matrix(runif(n*dimension), n, dimension)
phiP(X)

\texttt{rss2d}

\textit{2D graphical tool for defect detection of Space-Filling Designs.}

Description
For a 2-dimensional design, the 2D radial scanning statistic (RSS) scans angularly the domain. In each direction, it compares the distribution of projected points to their theoretical distribution under the assumption that all design points are drawn from uniform distribution. For a d-dimensional design, all pairs of dimensions are scanned. The RSS detects the defects of low discrepancy sequences or orthogonal arrays, and can be used for selecting space-filling designs.

Usage
\texttt{rss2d\(\text{design, lower, upper, gof.test.type=}&\text{"greenwood",
  gof.test.stat=NULL, transform=NULL, n.angle=360, graphics=1,
  trace=TRUE, lines.lwd = 1, lines.lty = &\text{"dotted"}, \ldots\)}}

Arguments
design a matrix or data.frame containing the d-dimensional design of experiments. The row no. i contains the values of the d input variables corresponding to simulation no. i
lower the domain lower boundaries.
upper the domain upper boundaries.
gof.test.type an optional character indicating the kind of statistical test to be used to test the goodness-of-fit of the design projections to their theoretical distribution. Several tests are available, see \texttt{unif.test.statistic}. Default is "greenwood".
gof.test.stat an optional number equal to the goodness-of-fit statistic at level 5%. Default is the modified test statistic for fully specified distribution (see details below).
transform an optional character indicating what type of transformation should be applied before testing uniformity. Only one choice available "spacings", that lead to over-detection. Default - and recommended - is NULL.
n.angle an optional number indicating the number of angles used. Default is 360 corresponding to a 0.5-degree discretization step. Note that the RSS curve is continuous.
graphics an optional integer indicating whether a graph should be produced. If negative, no graph is produced. If superior to 2, the RSS curve only is plotted in the worst 2D coordinate subspace (corr. to the worst value of statistic). If 1 (default), the design is also added, with its projections onto the worst (oblique) axis.

trace an optional boolean. Turn it to FALSE if you want no verbosity.

lines.lwd optional number specifying the width of the straight lines involved in the graphical outputs (axis and projections)

lines.lty optional character string specifying the type of the straight lines involved in the graphical outputs (axis and projections)

... optional graphical parameters of plot function, to draw the RSS curve.

Value

da list with components:

global.stat a matrix containing the values of the global statistic (equal to the maximum of statistic values over the RSS curve) for all pairs of dimensions.
worst.case the worst pair of dimensions, that is the one that gives the worst value of global.stat.
worst.dir the worst direction, that is the one that gives the worst value of the global statistic in the coordinate plane defined by worst.case.
stat a vector of length n.angle containing the statistic values for each angle, in the coordinate plane defined by worst.case.
ageangle a vector of length n.angle containing the corresponding angles used.
curve a (2*n.angle)x2 matrix containing the discretized RSS curve.
gof.test.stat the threshold at significance level 0.05 for the specified goodness-of-fit statistic. It is equal to the radius of the circle superimposed on the RSS figure.

Author(s)

O. Roustant

References


See Also

unif.test.statistic, unif.test.quantile, rss3d
Examples

# Detection of defects of Sobol designs

# requires randtoolbox package
library(randtoolbox)

# in 2D
rss <- rss3d(design=sobol(n=20, dim=2), lower=c(0,0), upper=c(1,1), type="l",
             col="red")

# in 8D.
# All pairs of dimensions are tried to detect the worst defect
# (according to the specified goodness-of-fit statistic).
d <- 8
n <- 10*d
rss <- rss3d(design=sobol(n=n, dim=d), lower=rep(0,d), upper=rep(1,d), type="l",
             col="red")

# avoid this defect with scrambling?
# 1. Faure-Tezuka scrambling (type "sobol" for more details and options)
rss <- rss3d(design=sobol(n=n, dim=d, scrambling=2), lower=rep(0,d),
             upper=rep(1,d), type="l", col="red")
# 2. Owen scrambling
rss <- rss3d(design=sobol(n=n, dim=d, scrambling=1), lower=rep(0,d),
             upper=rep(1,d), type="l", col="red")

---

rss3d 3D graphical tool for defect detection of Space-Filling Designs.

Description

For a 3-dimensional design, the 3D radial scanning statistic (RSS) scans angularly the domain. In each direction, it compares the distribution of projected points to their theoretical distribution under the assumption that all design points are drawn from uniform distribution. For a d-dimensional design, all triplets of dimensions are scanned. The RSS detects the defects of low discrepancy sequences or orthogonal arrays, and can be used for selecting space-filling designs.

Usage

rss3d(design, lower, upper, gof.test.type = "greenwood",
      gof.test.stat = NULL, transform = NULL, n.angle = 60,
      graphics = 1, trace = TRUE)

Arguments

design a matrix or data.frame containing the d-dimensional design of experiments. The row no. i contains the values of the d input variables corresponding to simulation no. i
lower
the domain lower boundaries.
upper
the domain upper boundaries.
gof.test.type
an optional character indicating the kind of statistical test to be used to test the
goodness-of-fit of the design projections to their theoretical distribution. Several
tests are available, see unif.test.statistic. Default is "greenwood".
gof.test.stat
an optional number equal to the goodness-of-fit statistic at level 5%. Default is
the modified test statistic for fully specified distribution (see details below).
transform
an optional character indicating what type of transformation should be applied
before testing uniformity. Only one choice available "spacings", that lead to
over-detection. Default - and recommended - is NULL.
n.angle
an optional number indicating the number of angles used. Default is 60 corre-
sponding to a 3-degree discretization step. Note that the RSS surface is contin-
uous.
graphics
an optional integer indicating whether a graph should be produced. If negative,
no graph is produced. Otherwise (default), the design is plotted in the worst 3D
coordinate subspace (corr. to the worst value of statistic), with its projections
onto the worst (oblique) axis.
trace
an optional boolean. Turn it to FALSE if you want no verbosity.

Details
The RSS surface is continuous. However for computational purposes, a discretization is used. The
default discretization step is tunable with n.angle.

Value
a list with components:

global.stat an array containing the values of the global statistic (equal to the maximum of
statistic values over the RSS surface) for all triplets of dimensions.
print.out the same as global.stat, but with a user-friendly printing.
worst.case the worst triplet of dimensions, that is the one that gives the worst value of
global.stat.
worst.dir the worst direction, that is the one that gives the worst value of the statistic in
the coordinate 3D subspace defined by worst.case.
stat a matrix of size n.angle*n.angle containing the statistic values for each angles
(spherical coordinates).
angle a matrix of size n.angle*n.angle containing the corresponding angles used
(spherical coordinates).
gof.test.stat the threshold at significance level 0.05 for the specified goodness-of-fit statistic.

Author(s)
O. Roustant
References


See Also

unif.test.statistic, unif.test.quantile, rss2d

Examples

# An orthogonal array in 3D
data(OA131)

# centering the design points of this 7-levels design
OA <- (OA131 + 0.5)/7

# 2D projections onto coordinate axis
pairs(OA, xlim=c(0,1), ylim=c(0,1))

# Now let us look at the 3D properties with the 3D RSS (requires the rgl package)
rss <- rss3d(OA, lower=c(0,0,0), upper=c(1,1,1))
# The worst direction detected is nearly proportional to (2,-1,2)
# (type "?OA131" for explanations about this linear orthogonal array)
print(rss$worst.dir)

# Now, scramble this design
# X <- (OA131 + matrix(runif(49*3, 49, 3))/7
# or load the design obtained this way
data(OA131_scrambled)
OA2 <- OA131_scrambled

# no feature is detected by the 2D RSS:
rss <- rss2d(OA2, lower=c(0,0,0), upper=c(1,1,1))
# 4 clusters are detected by the 3D RSS:
rss <- rss3d(OA2, lower=c(0,0,0), upper=c(1,1,1))

# Defect detection of 8D Sobol sequences
# All triplets of dimensions are tried to detect the worst defect
# (according to the specified goodness-of-fit statistic).
# requires randtoolbox library to generate the Sobol sequence
## Not run:
library(randtoolbox)
d <- 8
n <- 10*d
rss <- rss3d(design=sobol(n=n, dim=d), lower=rep(0,d), upper=rep(1,d))
## End(Not run)
runif.faure

Low discrepancy sequence: Faure

Description

Generate a Faure sequence with \( n \) experiments in \([0,1]^d\).

Usage

\texttt{runif.faure(n, dimension)}

Arguments

- \texttt{n} the number of experiments
- \texttt{dimension} the number of variables (<100)

Details

A quasirandom or low discrepancy sequence, such as the Faure, Halton, Hammersley, Niederreiter or Sobol sequences, is "less random" than a pseudorandom number sequence, but more useful for such tasks as approximation of integrals in higher dimensions, and in global optimization. This is because low discrepancy sequences tend to sample space "more uniformly" than random numbers. see \texttt{randtoolbox} or \texttt{fOptions} packages for other low discrepancy sequences.

Value

\texttt{runif.halton} returns a list containing all the input arguments detailed before, plus the following component:

- \texttt{design} the design of experiments

Author(s)

J. Franco

References

Faure H. (1982) \textit{Discrepance de suites associees a un systeme de numeration (en dimension s)}, Acta Arith. 41, 337-351

Examples

\begin{verbatim}
f <- runif.faure(20, 2)
plot(f$design, xlim=c(0,1), ylim=c(0,1))
\end{verbatim}
Description

Space-Filling Designs based on Strauss process

Usage

straussDesign(n,dimension,RND,alpha=0.5,repulsion=0.001,NMC=1000,
  constraints1D=0,repulsion1D=0.0001,seed=NULL)

Arguments

n the number of experiments
dimension the number of input variables
RND a real number which represents the radius of interaction
alpha the potential power (default, fixed at 0.5)
repulsion the repulsion parameter in the unit cube (gamma)
NMC the number of McMC iterations (this number must be large to converge)
constraints1D 1 to impose 1D projection constraints, 0 otherwise
repulsion1D the repulsion parameter in 1D
seed seed for the uniform generation of number

Details

Strauss designs are Space-Filling designs initially defined from Strauss process:

$$\pi(X) = k \gamma^s(X)$$

where $s(X)$ is is the number of pairs of points $(x^i, x^j)$ of the design $X = (x^1, \ldots, x^n)$ that are separated by a distance no greater than the radius of interaction RND, $k$ is the normalizing constant and $\gamma$ is the repulsion parameter. This distribution corresponds to the particular case alpha=0.

For the general case, a stochastic simulation is used to construct a Markov chain which converges to a spatial density of points $\pi(X)$ described by the Strauss-Gibbs potential. In practice, the Metropolis-Hastings algorithm is implemented to simulate a distribution of points which converges to the stationary law:

$$\pi(X) \propto \exp(-U(X))$$

with a potential $U$ defined by:

$$U(X) = \beta \sum_{1 \leq i < j \leq n} \varphi(\|x^i - x^j\|)$$

where $\beta = -\ln \gamma$, $\varphi(h) = \left(1 - \frac{h}{RND}\right)^\alpha$ if $h \leq RND$ and 0 otherwise.
The input parameters of `straussDesign` function can be interpreted as follows:

- `rnd` is used to compute the number of pairs of points of the design separated by a distance no more than `rnd`. A point is said "in interaction" with another if the spheres of radius `RND/2` centered on these points intersect.
- `alpha` is the potential power $\alpha$. The case $\alpha=0$ corresponds to Strauss process (0-1 potential).
- `repulsion` is equal to the $\gamma$ parameter of the Strauss process. Note that $\gamma$ belongs to $]0,1]$.
- `constraints1D` allows to specify some constraints into the margin. If `constraints1D==1`, two repulsion parameters are needed: one for the all space (`repulsion`) and the other for the 1D projection (`repulsion1D`). Default values are `repulsion=0.001` and `repulsion1D=0.001`. Note that the value of the radius of interaction in the one-dimensional axis is not an input parameter and is automatically fixed at $0.75/n$.

**Value**

A list containing:

- `n` the number of experiments
- `dimension` the number $d$ of variables
- `design_init` the initial distribution of $n$ points $[0,1]^d$
- `radius` the radius of interaction
- `alpha` the potential power $\alpha$
- `repulsion` the repulsion parameter $\gamma$
- `NMC` the number of iterations McMC
- `constraints1D` an integer indicating if constraints on the factorial axis are imposed. If its value is different from zero, a component `repulsion1D` containing the value of the repulsion parameter $\gamma$ in dimension 1 is added at the list.
- `design` the design of experiments in $[0,1]^d$
- `seed` the seed corresponding to the design

**Author(s)**

J. Franco

**References**


**Examples**

```r
# Strauss-Gibbs designs in dimension 2 (n=20 points)
S1 <- straussDesign(n=20, dimension=2, RND=0.2)
plot(S1$design, xlim=c(0,1), ylim=c(0,1))
theta <- seq(0,2*pi,by =2*pi/(100 - 1))
for(i in 1:S1$n){
```
lines(S1$design[i,1]+S1$radius/2*cos(theta),
S1$design[i,2]+S1$radius/2*sin(theta),col='red')
}

# 2D-Strauss design
S2 <- straussDesign(n=20,dimension=2, RND=0.2, NMC=200,
constrained=1D=0, alpha=0, repulsion=0.01)

plot(S2$design,xlim=c(0,1),ylim=c(0,1))

# 2D-Strauss designs with constraints on the axis
S3 <- straussDesign(n=20,dimension=2, RND=0.18, NMC=200,
constrained=1D=1, alpha=0.5, repulsion=0.1, repulsion1D=0.01)

plot(S3$design,xlim=c(0,1),ylim=c(0,1))
rug(S3$design[,1],side=1)
rug(S3$design[,2],side=2)

---

unif.test.quantile  

Quantile of some uniformity tests

Description

Computes the quantile of a uniformity test at a given significance level (see available tests and levels below).

Usage

unif.test.quantile(type, n, alpha)

Arguments

type

a character indicating which test is used. The choices are the following: "greenwood", "qm" (for Quesenberry-Miller), "ks" (Kolmogorov-Smirnov), "cvm" (Cramer-Von Mises) and "V" (D+ + D- from Kolmogorov-Smirnov).

n

an integer equal to the sample size.

alpha

a real number equal to significance level. At present stage, only four values are available: 0.1, 0.05, 0.025 and 0.01.

Details

Modified statistics are used. For alpha = 0.05, the quantile is (see D Agostino and Stephens, 1986, section 4.4.): 1.358/(sqrt(n) + 0.12 + 0.11/sqrt(n)) for Kolmogorov-Smirnov and 0.461/(1+1/n) + 0.4/n - 0.6/n^2 for Cramer-von Mises. When the design size is < 20, the corrected value seems to be a good approximation, but the non asymptotical value should be preferred.
Value
A real number equal to the quantile of the specified test at significance level \( \alpha \) for \( n \) observations.

Author(s)
O. Roustant

References

See Also
unif.test.statistic, rss2d, rss3d

---

unif.test.statistic

Statistic of some uniformity tests

Description
Computes the statistic of a uniformity test (see available tests below).

Usage
unif.test.statistic(x, type, transform=NULL)

Arguments

- \( x \) a vector containing the sample values.
- \( \text{type} \) a character indicating which test is used. The choices are the following: "greenwood", "qm" (for Quesenberry-Miller), "ks" (Kolmogorov-Smirnov), "cvm" (Cramer-Von Mises) and "V" (\( D+ + D- \) from Kolmogorov-Smirnov).
- \( \text{transform} \) an optional character indicating what type of transformation should be applied before testing uniformity. Default is NULL.

Value
A real number equal to the statistic of the specified test.

Author(s)
O. Roustant

References
Description

The WSP (Wooton, Sergent, Phan-Tan-Luu) algorithm is an iterative algorithm based on suppression of some experiments from an initial design in each step. WSP leads to a space filling design.

Usage

wspDesign(design, dmin)

Arguments

design a matrix (or a data.frame) corresponding to the design of experiments.
dmin a minimum bound for mindist value of the final design

Details

WSP enables to create a design D which is such that mindist(D)>dmin. However, it cannot assess the number of experiments. Similarly to straussDesign function, WSP is a powerful algorithm to construct space filling designs in high dimension.

Value

A list containing:

- InitialDesign the starting design
- dmin minimum bound for mindist value of the final design
- design the matrix of the final design

Author(s)

G.Damblin & B.Iooss

References

Examples

dimension <- 2
n <- 100
X <- matrix(runif(n*dimension),n,dimension)
m=wspDesign(X,0.1)
plot(m$design)
Index

*Topic datasets
  OA131, 21
  OA131_scrambled, 22

*Topic design
  coverage, 4
discrepancyCriteria, 5
discrepESE_LHS, 7
discrepSA_LHS, 9
dmaxDesign, 11
factDesign, 12
lhsDesign, 13
maximinESE_LHS, 14
maximinSA_LHS, 16
meshRatio, 18
mindist, 19
mstCriteria, 20
phip, 5
rssRd, 24
rssSd, 25
runif.faure, 29
straussDesign, 30
unif.test.quantile, 25
unif.test.statistic, 24

*Topic package
  DiceDesign-package, 2
  coverage, 4, 6
  DiceDesign (DiceDesign-package), 2
  DiceDesign-package, 2
discrepancyCriteria, 5, 8, 10, 16–18, 20
discrepESE_LHS, 7, 10, 14, 16, 17
discrepSA_LHS, 8, 9, 14, 16, 17
dmaxDesign, 11
factDesign, 12
lhsDesign, 8, 10, 13, 16, 17
maximinESE_LHS, 8, 10, 14, 17
maximinSA_LHS, 8, 10, 14, 16, 16
meshRatio, 5, 6, 18, 18, 20
mindist, 5, 6, 8, 10, 16–18, 19, 24
mstCriteria, 20
OA131, 21
OA131_scrambled, 22
phip, 5, 6, 8, 10, 16–18, 20, 23
rss2d, 24, 28, 33, 34
rss3d, 25, 26, 33
runif.faure, 29
straussDesign, 30
unif.test.quantile, 25, 28, 32, 34
unif.test.statistic, 24, 25, 27, 28, 33, 33
wspDesign, 34