Package 'sensitivity'

April 19, 2009

Version 1.4-0

Date 2008-07-14	
Title Sensitivity Analysis	
Author Gilles Pujol	
Maintainer Bertrand Iooss deca.fr>	
Depends R (>= $2.7.0$), boot	
Suggests rgl	
Description A collection of functions for factor screening and global sensitivity analysis of model output.	
License CeCILL-2	
Repository CRAN	
Date/Publication 2008-07-15 15:47:55	
sobol	3 4 6 8 10 12 14 15 17 18
Index	19

2 sensitivity-package

sensitivity-package

Sensitivity Analysis

Description

Methods and functions for global sensitivity analysis.

Details

The sensitivity package implements some global sensitivity analysis methods:

- Linear regression coefficients: SRC and SRRC (src), PCC and PRCC (pcc).
- Morris's "OAT" elementary effects screening method (morris).
- Bettonvil's sequential bifurcations (sb).
- Monte Carlo estimation of Sobol' indices: Sobol's scheme (1993) to compute the indices given by the variance decomposition up to a specified order (sobol), and Saltelli's scheme (2002) to compute first order and total indices with a reduced cost (sobol2002).
- Estimation of the Sobol' first order and total indices with Saltelli's so-called "extended-FAST" method (fast99).

Moreover, some utilities are provided: standard test-cases (testmodels) and template file generation (template.replace).

Model managing

The **sensitivity** package works either on R models than on external models (such as executables).

R models must be functions or objects that have a predict method, such as lm objects. Models defined as functions will be called once with an expression of the form y < -f(X) where X is the design of experiments, i.e. a data.frame with p columns (the input factors) and n lines (each, an experiment), and y is the vector of length n of the model responses (we say that such functions are vectorized).

If the model is external to R, for instance a computational code, it must be analyzed with the decoupled approach, see decoupling. This approach can also be used on R models that doesn't fit the specifications.

References

A. Saltelli, K. Chan and E. M. Scott eds, 2000, Sensitivity Analysis, Wiley.

decoupling 3

-1 -		- 7			_
ae	cour	$^{\perp}$	_	11	a

Decoupling Simulations and Estimations

Description

tell and ask are S3 generic methods for decoupling simulations and sensitivity measures estimations. In general, they are not used by the end-user for a simple R model, but rather for an external computational code. Most of the sensitivity analyses objects of this package overload tell, whereas ask is overloaded for iterative methods only.

Usage

```
tell(x, y = NULL, ...) ask(x, ...)
```

Arguments

x a typed list storing the state of the sensitivity study (parameters, data, estimates), as returned by sensitivity analyses objects constructors, such as src, morris, etc.

y a vector of model responses.

... additional arguments, depending on the method used.

Details

When a sensitivity analysis method is called with no model (i.e. argument model = NULL), it generates an incomplete object x that stores the design of experiments (field X), allowing the user to launch "by hand" the corresponding simulations. The method tell allows to pass these simulation results to the incomplete object x, thereafter estimating the sensitivity measures.

When the method is iterative, the data to simulate are not stored in the sensitivity analysis object x, but generated at each iteration with the ask method; see for example sb.

Value

tell doesn't return anything. It computes the sensitivity measures, and stores them in the list x. Side effect: tell modifies its argument x.

ask returns the set of data to simulate.

4 fast99

fast99

Extended Fourier Amplitude Sensitivity Test

Description

fast 99 implements the so-called "extended-FAST" method (Saltelli et al. 1999). This method allows the estimation of first order and total Sobol' indices for all the factors (alltogether 2p indices, where p is the number of factors) at a total cost of $n \times p$ simulations.

Usage

Arguments

model	a function, or a model with a ${\tt predict}$ method, defining the model to analyze.
factors	an integer giving the number of factors, or a vector of character strings giving their names.
n	an integer giving the sample size, i.e. the length of the discretization of the s-space (see Cukier et al.).
М	an integer specifying the interference parameter, i.e. the number of harmonics to sum in the Fourier series decomposition (see Cukier et al.).
omega	a vector giving the set of frequencies, one frequency for each factor (see details below).
d	a vector of quantile functions names corresponding to wanted factors distributions (see details below).
q.arg	a list of quantile functions parameters (see details below).
х	a list of class "fast99" storing the state of the sensitivity study (parameters, data, estimates).
У	a vector of model responses.
ylim	y-coordinate plotting limits.
• • •	any other arguments for ${\tt model}$ which are passed unchanged each time it is called.

fast99 5

Details

If not given, the set of frequencies omega is taken from Saltelli et al. The first frequency of the vector omega is assigned to each factor X_i in turn (corresponding to the estimation of Sobol' indices S_i and S_{T_i}), other frequencies being assigned to the remaining factors.

If the arguments q and q.args are not given, the factors are taken uniformly distributed on [0,1]. The argument q must be list of character strings, giving the names of the quantile functions (one for each factor), such as qunif, qnorm... It can also be a single character string, meaning same distribution for all. The argument q.arg must be a list of lists, each one being additional parameters for the corresponding quantile function. For example, the parameters of the quantile function qunif could be list (min=1, max=2), giving an uniform distribution on [1,2]. If q is a single character string, then q.arg must be a single list (rather than a list of one list).

Value

fast 99 returns a list of class "fast 99", containing all the input arguments detailed before, plus the following components:

call	the matched call.
X	a data.frame containing the factors sample values.
У	a vector of model responses.
V	the estimation of variance.
D1	the estimations of Variances of the Conditional Expectations (VCE) with respect to each factor.
Dt	the estimations of VCE with respect to each factor complementary set of factors ("all but X_i ").

References

A. Saltelli, S. Tarantola and K. Chan, 1999, A quantitative, model independent method for global sensitivity analysis of model output, Technometrics, 41, 39–56.

R. I. Cukier, H. B. Levine and K. E. Schuler, 1978, *Nonlinear sensitivity analysis of multiparameter model systems*. J. Comput. Phys., 26, 1–42.

See Also

```
decoupling
```

6 morris

morris

Morris's Elementary Effects Screening Method

Description

morris implements the Morris's elementary effects screening method (Morris 1992). This method, based on design of experiments, allows to identify the few important factors at a cost of $r \times (p+1)$ simulations (where p is the number of factors). This implementation includes some improvements of the original method: space-filling optimization of the design (Campolongo et al. 2007) and simplex-based design (Pujol 2008).

Usage

Arguments

model

a function, or a model with a predict method, defining the model to analyze.

factors

an integer giving the number of factors, or a vector of character strings giving their names.

r

either an integer giving the number of repetitions of the design, i.e. the number of elementary effect computed per factor, or a vector of two integers c(r1, r2) for the space-filling improvement (Campolongo et al.). In this case, r1 is the wanted design size, and r2 (> r1) is the size of the (bigger) population in which is extracted the design (this can throw a warning, see below).

design

a list specifying the design type and its parameters:

- type = "oat" for Morris's OAT design (Morris 1992), with the parameters:
 - levels: either an integer specifying the number of levels of the design, or a vector of integers for different values for each factor.
 - grid.jump: either an integer specifying the number of levels that are increased/decreased for computing the elementary effects, or a vector of integers for different values for each factor. If not given, it is set to grid.jump = 1. Notice that this default value of one does not follow Morris's recommendation of levels/2.
- type = "simplex" for simplex-based design (Pujol 2008), with the parameter:

morris 7

	 scale.factor: a numeric value, the homothety factor of the (isometric) simplexes. Edges equal one with a scale factor of one.
binf	either an integer, specifying the minimum value for the factors, or a vector for different values for each factor.
bsup	either an integer, specifying the maximum value for the factors, or a vector for different values for each factor.
scale	logical. If ${\tt TRUE},$ the input and output data are scaled before computing the elementary effects.
X	a list of class "morris" storing the state of the screening study (parameters, data, estimates).
У	a vector of model responses.
identify	logical. If TRUE, the user selects with the mouse the factors to label on the (μ^*,σ) graph (see identify).
• • •	any other arguments for model which are passed unchanged each time it is called.
alpha	a vector of three values between 0.0 (fully transparent) and 1.0 (opaque) (see rgl.material). The first value is for the cone, the second for the planes.
sphere.size	a numeric value, the scale factor for displaying the spheres.

Details

```
plot2d draws the (\mu^*, \sigma) graph.
```

plot3d.morris draws the (μ, μ^*, σ) graph (requires the **rgl** package). On this graph, the points are in a domain bounded by a cone and two planes (application of the Cauchy-Schwarz inequality).

Value

morris returns a list of class "morris", containing all the input argument detailed before, plus the following components:

call	the matched call.
X	\boldsymbol{a} data.frame containing the design of experiments.
У	a vector of model responses.
ee	a $r \times p$ matrix of elementary effects for all the factors.

Notice that the statitics of interest $(\mu, \mu^* \text{ and } \sigma)$ are not stored. They can be printed by the print method, but to extract numerical values, one has to compute them with the following instructions:

```
mu <- apply(x$ee, 2, mean)
mu.star <- apply(x$ee, 2, function(x) mean(abs(x)))
sigma <- apply(x$ee, 2, sd)</pre>
```

Warning messages

"keeping r' repetitions out of r" when generating the design of experiments, identical repetitions are removed, leading to a lower number than requested.

8 pcc

References

- M. D. Morris, 1991, Factorial sampling plans for preliminary computational experiments, Technometrics, 33, 161–174.
- F. Campolongo, J. Cariboni and A. Saltelli, 2007, *An effective screening design for sensitivity*, Environmental Modelling & Software, 22, 1509–1518.
- G. Pujol (2008), *Simplex-based screening designs for estimating metamodels*, submitted to Reliability Engineering and System Safety.

See Also

```
decoupling
```

Examples

рсс

Partial Correlation Coefficients

Description

pcc computes the Partial Correlation Coefficients (PCC), or Partial Rank Correlation Coefficients (PRCC), which are sensitivity indices based on linear (resp. monotonic) assumptions, in the case of (linearly) correlated factors.

Usage

```
pcc(X, y, rank = FALSE, nboot = 0, conf = 0.95)
## S3 method for class 'pcc':
print(x, ...)
## S3 method for class 'pcc':
plot(x, ylim = c(-1,1), ...)
```

Arguments

X	a data frame (or object coercible by as.data.frame) containing the design of experiments (model input variables).
У	a vector containing the responses corresponding to the design of experiments (model output variables).
rank	logical. If TRUE, the analysis is done on the ranks.
nboot	the number of bootstrap replicates.

pcc 9

conf	the confidence level of the bootstrap confidence intervals.
Х	the object returned by pcc.
ylim	the y-coordinate limits of the plot.
	arguments to be passed to methods, such as graphical parameters (see par).

Value

pcc returns a list of class "pcc", containing the following components:

call	the matched call.
PCC	a data frame containing the estimations of the PCC indices, bias and confidence intervals (if $\mathtt{rank} = \mathtt{TRUE}$).
PRCC	a data frame containing the estimations of the PRCC indices, bias and confidence intervals (if rank = TRUE).

References

A. Saltelli, K. Chan and E. M. Scott eds, 2000, Sensitivity Analysis, Wiley.

See Also

src

10 sb

b Sequential Bifurcations

Description

sb implements the Sequential Bifurcations screening method (Bettonvil and Kleijnen 1996). **This** is an alpha version that might strongly evolve in the future.

Usage

```
sb(p, sign = rep("+", p), interaction = FALSE)
## S3 method for class 'sb':
ask(x, i = NULL, ...)
## S3 method for class 'sb':
tell(x, y, ...)
## S3 method for class 'sb':
print(x, ...)
## S3 method for class 'sb':
plot(x, ...)
```

Arguments

р	number of factors.
sign	a vector fo length ${\tt p}$ filled with "+" and "-", giving the (assumed) signs of the factors effects.
interaction	a boolean, ${\tt TRUE}$ if the model is supposed to be with interactions, ${\tt FALSE}$ otherwise.
х	a list of class " ${\tt sb}$ " storing the state of the screening study at the current iteration.
У	a vector of model responses.
i	an integer, used to force a wanted bifurcation instead of that proposed by the algorithm.
	not used.

Details

The model without interaction is

$$Y = \beta_0 + \sum_{i=1}^p \beta_i X_i$$

while the model with interactions is

$$Y = \beta_0 + \sum_{i=1}^p \beta_i X_i + \sum_{1 \le i < j \le p} \gamma_{ij} X_i X_j$$

sb

sb 11

In both cases, the factors are assumed to be uniformly distributed on [-1,1]. This is a difference with Bettonvil et al. where the factors vary across [0,1] in the former case, while [-1,1] in the latter

Another difference with Bettonvil et al. is that in the current implementation, the groups are splitted right in the middle.

Value

sb returns a list of class "sb", containing all the input arguments detailed before, plus the following components:

```
    the vector of bifurcations.
    the vector of observations.
    the vector of mirror observations (model with interactions only).
```

The groups effects can be displayed with the print method.

References

B. Bettonvil and J. P. C. Kleijnen, 1996, Searching for important factors in simulation models with many factors: sequential bifurcations, European Journal of Operational Research, 96, 180–194.

```
# a model with interactions
p <- 50
beta <- numeric(length = p)</pre>
beta[1:5] \leftarrow runif(n = 5, min = 10, max = 50)
beta[6:p] \leftarrow runif(n = p - 5, min = 0, max = 0.3)
beta <- sample(beta)</pre>
gamma \leftarrow matrix(data = runif(n = p^2, min = 0, max = 0.1), nrow = p, ncol = p)
gamma[lower.tri(gamma, diag = TRUE)] <- 0</pre>
gamma[1,2] <- 5
gamma[5, 9] < -12
f \leftarrow function(x) \{ return(sum(x * beta) + (x %*% gamma %*% x)) \}
# 10 iterations of SB
sa <- sb(p, interaction = TRUE)
for (i in 1 : 10) {
  x < -ask(sa)
  y <- list()
  for (i in names(x)) {
    y[[i]] <- f(x[[i]])
  tell(sa, y)
}
print(sa)
plot(sa)
```

12 sobol

sobol

Monte Carlo Estimation of Sobol' Indices

Description

sobol implements the Monte Carlo estimation of the Sobol' sensitivity indices. This method allows the estimation of the indices of the variance decomposition, sometimes referred to as functional ANOVA decomposition, up to a given order, at a total cost of $(N+1) \times n$ where N is the number of indices to estimate. This function allows also the estimation of the so-called subset indices, i.e. the first-order indices with respect to single multidimensional inputs.

Usage

```
sobol(model = NULL, X1, X2, order = 1, nboot = 0, conf = 0.95, ...)
## S3 method for class 'sobol':
tell(x, y = NULL, return.var = NULL, ...)
## S3 method for class 'sobol':
print(x, ...)
## S3 method for class 'sobol':
plot(x, ylim = c(0, 1), ...)
```

Arguments

model	a function, or a model with a predict method, defining the model to analyze.
X1	the first random sample.
X2	the second random sample.
order	either an integer, the maximum order in the ANOVA decomposition (all indices up to this order will be computed), or a list of numeric vectors, the multidimensional compounds of the wanted subset indices.
nboot	the number of bootstrap replicates.
conf	the confidence level for bootstrap confidence intervals.
х	a list of class " $\verb sobol $ " storing the state of the sensitivity study (parameters, data, estimates).
У	a vector of model responses.
return.var	a vector of character strings giving further internal variables names to store in the output object \mathbf{x} .
ylim	y-coordinate plotting limits.
•••	any other arguments for $model$ which are passed unchanged each time it is called.

sobol 13

Value

sobol returns a list of class "sobol", containing all the input arguments detailed before, plus the following components:

call	the matched call.
X	a data.frame containing the design of experiments.
У	a vector of model responses.
V	the estimations of Variances of the Conditional Expectations (VCE) with respect to one factor or one group of factors.
D	the estimations of the terms of the ANOVA decomposition (not for subset indices).
S	the estimations of the Sobol' sensitivity indices (not for subset indices).

Users can ask more ouput variables with the argument return.var (for example, bootstrap outputs V.boot, D.boot and S.boot).

References

I. M. Sobol, 1993, *Sensitivity analysis for non-linear mathematical model*, Math. Modelling Comput. Exp., 1, 407–414.

See Also

```
sobo12002
```

```
# Test case : the non-monotonic Sobol g-function

# The method of sobol requires 2 samples
# (there are 8 factors, all following the uniform distribution on [0,1])
n <- 1000
X1 <- data.frame(matrix(runif(8 * n), nrow = n))
X2 <- data.frame(matrix(runif(8 * n), nrow = n))

# sensitivity analysis
x <- sobol(model = sobol.fun, X1 = X1, X2 = X2, order = 2, nboot = 100)
print(x)
#plot(x)</pre>
```

14 sobol2002

sobol2002

Monte Carlo Estimation of Sobol' Indices (scheme by Saltelli 2002)

Description

sobol2002 implements the Monte Carlo estimation of the Sobol' indices for both first-order and total indices at the same time (alltogether 2p indices), at a total cost of $(p+2) \times n$ model evaluations.

Usage

```
sobol2002(model = NULL, X1, X2, nboot = 0, conf = 0.95, ...)
## S3 method for class 'sobol2002':
tell(x, y = NULL, return.var = NULL, ...)
## S3 method for class 'sobol2002':
print(x, ...)
## S3 method for class 'sobol2002':
plot(x, ylim = c(0, 1), ...)
```

Arguments

model	a function, or a model with a predict method, defining the model to analyze.
X1	the first random sample.
X2	the second random sample.
nboot	the number of bootstrap replicates.
conf	the confidence level for bootstrap confidence intervals.
Х	a list of class "sobol" storing the state of the sensitivity study (parameters, data, estimates).
У	a vector of model responses.
return.var	a vector of character strings giving further internal variables names to store in the output object \mathbf{x} .
ylim	y-coordinate plotting limits.
• • •	any other arguments for \ensuremath{model} which are passed unchanged each time it is called

Value

sobol2002 returns a list of class "sobol2002", containing all the input arguments detailed before, plus the following components:

call	the matched call.
X	a data.frame containing the design of experiments.
V	the response used

src 15

V	the estimations of Variances of the Conditional Expectations (VCE) with respect to each factor and also with respect to the complementary set of each factor ("all but X_i ").
S	the estimations of the Sobol' first-order indices.
T	the estimations of the Sobol' total sensitivity indices.

Users can ask more ouput variables with the argument return.var (for example, bootstrap outputs V.boot, S.boot and T.boot).

References

A. Saltelli, 2002, *Making best use of model evaluations to compute sensitivity indices*, Computer Physics Communication, 145, 580–297.

Examples

```
# Test case : the non-monotonic Sobol g-function
# The method of sobol requires 2 samples
# There are 8 factors, all following the uniform distribution
# on [0,1]

n <- 1000
X1 <- data.frame(matrix(runif(8 * n), nrow = n))
X2 <- data.frame(matrix(runif(8 * n), nrow = n))
# sensitivity analysis

x <- sobol2002(model = sobol.fun, X1, X2, nboot = 100)
print(x)
plot(x)</pre>
```

Standardized Regression Coefficients

Description

src

src computes the Standardized Regression Coefficients (SRC), or the Standardized Rank Regression Coefficients (SRRC), which are sensitivity indices based on linear or monotonic assumptions in the case of independent factors.

Usage

```
src(X, y, rank = FALSE, nboot = 0, conf = 0.95)
## S3 method for class 'src':
print(x, ...)
## S3 method for class 'src':
plot(x, ylim = c(-1,1), ...)
```

16 src

Arguments

X	a data frame (or object coercible by as.data.frame) containing the design of experiments (model input variables).
У	a vector containing the responses corresponding to the design of experiments (model output variables).
rank	logical. If TRUE, the analysis is done on the ranks.
nboot	the number of bootstrap replicates.
conf	the confidence level of the bootstrap confidence intervals.
X	the object returned by src.
ylim	the y-coordinate limits of the plot.
	arguments to be passed to methods, such as graphical parameters (see par).

Value

src returns a list of class "src", containing the following components:

call	the matched call.
SRC	a data frame containing the estimations of the SRC indices, bias and confidence intervals (if rank = FALSE).
SRRC	a data frame containing the estimations of the SRRC indices, bias and confidence intervals (if rank = TRUE).

References

A. Saltelli, K. Chan and E. M. Scott eds, 2000, Sensitivity Analysis, Wiley.

See Also

рсс

template.replace 17

```
x <- src(X, y, nboot = 100)
print(x)
plot(x)</pre>
```

template.replace

Replace Values in a Template Text

Description

template.replace replaces keys within special markups with values in a so-called template file. Pieces of R code can be put into the markups of the template file, and are evaluated during the replacement.

Usage

Arguments

```
text vector of character strings, the template text.
replacement the list values to replace in text.
eval boolean, TRUE if the code within code.pattern has to be evaluated, FALSE otherwise.
key.pattern custom pattern for key replacement (see below)
code.pattern custom pattern for code replacement (see below)
```

Details

In most cases, a computational code reads its inputs from a text file. A template file is like an input file, but where some missing values, identified with generic keys, will be replaced by specific values.

By default, the keys are enclosed into markups of the form \$ (KEY).

Code to be interpreted with R can be put in the template text. Pieces of code must be enclosed into markups of the form $@\{CODE\}$. This is useful for example for formating the key values (see example). For interpreting the code, set eval = TRUE.

Users can define custom patterns. These patterns must be perl-compatible regular expressions (see regexpr. The default ones are:

```
key.pattern = "\\$\\(KEY\\)"
code.pattern = "@\\{CODE\\}"
```

Note that special characters have to be escaped both (one for perl, one for R).

18 testmodels

Examples

testmodels

Test Models for Sensitivity Analysis

Description

These functions are standard testcase for sensitivity analysis benchmarks: The g-function of Sobol', the function of Ishigami and the function of Morris (see Saltelli et al. 2000, section 2.9)

Usage

```
sobol.fun(X)
ishigami.fun(X)
morris.fun(X)
```

Arguments

Χ

a matrix (or data.frame) containing the input sample.

Value

A vector of function responses.

References

A. Saltelli, K. Chan and E. M. Scott eds, 2000, Sensitivity Analysis, Wiley.

Index

*Topic IO	print.pcc(pcc),8
template.replace, 16	print.sb (sb) ,9
*Topic design	print.sobol(sobol),11
fast99,3	print.sobol2002(sobol2002), 13
morris,5	print.src(src),14
sb, 9	
sobol, 11	regexpr, 16
sobol2002, 13	. 2.2.0
*Topic methods	sb, 2, 3, 9
decoupling, 2	sensitivity
*Topic misc	(sensitivity-package), 1
testmodels, 17	sensitivity-package, 1
*Topic package	sobol, 2, 11
sensitivity-package, 1	sobol.fun(testmodels), 17
*Topic regression	sobol2002, 2, 12, 13
pcc, 8	src, 2, 3, 9, 14
src, 14	tell (decoupling), 2
	tell.fast99(fast99),3
ask (decoupling), 2	tell.morris(morris),5
ask.sb(<i>sb</i>),9	tell.sb (sb) , 9
decoupling, 2, 2, 5, 7	tell.sobol(sobol), 11
aecoupiing, 2, 2, 3, 7	tell.sobol2002(sobol2002), 13
fast99,2,3	template.replace, 2, 16
, , .	testmodels, 2, 17
ishigami.fun(testmodels),17	
morris, 2, 3, 5	
morris.fun(testmodels),17	
pcc, 2, 8, 15	
plot.fast99 (fast99), 3	
plot.morris(morris),5	
plot.moiiis (moiiis), 5 plot.pcc(pcc), 8	
plot.sb(<i>sb</i>), 9	
plot.sobol(sobol), 11	
plot.sobol2002 (sobol2002), 13	
plot.src(<i>src</i>), 14	
plot3d.morris (morris), 5	
print.fast99(<i>fast99</i>),3	
print.morris(morris),5	
P	