Package 'generalCorr'

January 3, 2022

Type Package

Title Generalized Correlations, Causal Paths and Portfolio Selection

Version 1.2.2

Date 2021-12-24

Author Prof. H. D. Vinod, Fordham University, NY.

Maintainer H. D. Vinod <vinod@fordham.edu>

Encoding UTF-8

Depends R (>= 3.0.0), np (>= 0.60), xtable (>= 1.8), meboot (>= 1.4), psych, lattice

Suggests R.rsp

VignetteBuilder R.rsp

Description Since causal paths from data are important for all sciences, the package provides many sophisticated functions. causeSummBlk() and causeSum2Blk() give easy-to-interpret causal paths. Let Z denote control variables and compare two flipped kernel regressions: X=f(Y, Z)+e1 and Y=g(X,Z)+e2. Our criterion Cr1 says that if |e1*Y|>|e2*X| then variation in X is more ``exogenous or independent'' than in Y and causal path is X to Y. Criterion Cr2 requires |e2|<|e1|. These inequalities between many absolute values are quantified by four orders of stochastic dominance. Our third criterion Cr3 for the causal path X to Y requires new generalized partial correlations to satisfy |r*(x|y,z)|< |r*(y|x,z)|. The function parcorVec() reports generalized partials between the first variable and all others. The package provides several R functions including get0outliers() for outlier detection, bigfp() for numerical integration by the trapezoidal rule, stochdom2() for stochastic dominance, pillar3D() for 3D charts, canonRho() for generalized canonical correlations, depMeas() measures nonlinear dependence, and causeSummary(mtx) reports summary of causal paths among matrix columns is easiest to use. Portfolio selection: decileVote(), momentVote(), dif4mtx(), exactSdMtx() can rank several stocks. Several functions whose names begin with 'boot' provide bootstrap
statistical inference including a new bootGcRsq() test for ``Granger-causality'' allowing nonlinear relations. A new tool for evaluation of out-of-sample portfolio performance is outOFsamp(). See six vignettes of the package for theory and usage tips. See Vinod (2019) \doi{10.1080/03610918.2015.1122048}.

License GPL (>= 2)

LazyData true

RoxygenNote 7.1.2

NeedsCompilation no

Repository CRAN

Date/Publication 2022-01-03 17:10:02 UTC

R topics documented:

`absBstdres` *Block version of abs-stdres Absolute values of residuals of kernel regressions of standardized x on standardized y, no control variables.*

Description

1) Standardize the data to force mean zero and variance unity, 2) kernel regress x on y, with the option 'residuals = TRUE' and finally 3) compute the absolute values of residuals.

Usage

```
absBstdres(x, y, blksiz = 10)
```

Arguments

`x`	vector of data on the dependent variable
`y`	data on the regressors which can be a matrix
`blksiz`	block size, default=10, if chosen blksiz >n, where n=rows in matrix then blksiz=n. That is, no blocking is done

Details

The first argument is assumed to be the dependent variable. If `abs_stdres(x,y)` is used, you are regressing x on y (not the usual y on x). The regressors can be a matrix with 2 or more columns. The missing values are suitably ignored by the standardization.

Value

Absolute values of kernel regression residuals are returned after standardizing the data on both sides so that the magnitudes of residuals are comparable between regression of x on y on the one hand and regression of y on x on the other.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: 10.1080/03610918.2015.1122048

Examples

```
## Not run:
set.seed(330)
x=sample(20:50)
y=sample(20:50)
abs_stdres(x,y)

## End(Not run)
```

`absBstdresC` *Block version of Absolute values of residuals of kernel regressions of standardized x on standardized y and control variables.*

Description

1) standardize the data to force mean zero and variance unity, 2) kernel regress x on y and a matrix of control variables, with the option 'residuals = TRUE' and finally 3) compute the absolute values of residuals.

Usage

```
absBstdresC(x, y, ctrl, blksiz = 10)
```

Arguments

`x`	vector of data on the dependent variable
`y`	data on the regressors which can be a matrix
`ctrl`	Data matrix on the control variable(s) beyond causal path issues
`blksiz`	block size, default=10, if chosen blksiz >n, where n=rows in matrix then blksiz=n. That is, no blocking is done

Details

The first argument is assumed to be the dependent variable. If `abs_stdres(x,y)` is used, you are regressing x on y (not the usual y on x). The regressors can be a matrix with two or more columns. The missing values are suitably ignored by the standardization.

Value

Absolute values of kernel regression residuals are returned after standardizing the data on both sides so that the magnitudes of residuals are comparable between regression of x on y on the one hand and regression of y on x on the other.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D.'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: gffn86

See Also

See `abs_stdres`.

Examples

```
## Not run:
set.seed(330)
x=sample(20:50)
y=sample(20:50)
z=sample(21:51)
absBstdresC(x,y,ctrl=z)

## End(Not run)
```

`absBstdrhserC` *Block version abs_stdrhser Absolute residuals kernel regressions of standardized x on y and control variables, Cr1 has abs(Resid*RHS).*

Description

1) standardize the data to force mean zero and variance unity, 2) kernel regress x on y and a matrix of control variables, with the option 'residuals = TRUE' and finally 3) compute the absolute values of residuals.

Usage

```
absBstdrhserC(x, y, ctrl, ycolumn = 1, blksiz = 10)
```

Arguments

x	vector of data on the dependent variable
y	data on the regressors which can be a matrix
ctrl	Data matrix on the control variable(s) beyond causal path issues
ycolumn	if y has more than one column, the column number used when multiplying residuals times this column of y, default=1 or first column of y matrix is used
blksiz	block size, default=10, if chosen blksiz >n, where n=rows in matrix then blksiz=n. That is, no blocking is done

Details

The first argument is assumed to be the dependent variable. If `absBstdrhserC(x,y)` is used, you are regressing x on y (not the usual y on x). The regressors can be a matrix with 2 or more columns. The missing values are suitably ignored by the standardization.

Value

Absolute values of kernel regression residuals are returned after standardizing the data on both sides so that the magnitudes of residuals are comparable between regression of x on y on the one hand and regression of y on x on the other.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, https://doi.org/gffn86

See Also

See `abs_stdres`.

Examples

```
## Not run:
set.seed(330)
x=sample(20:50)
y=sample(20:50)
z=sample(21:51)
absBstdrhserC(x,y,ctrl=z)

## End(Not run)
```

abs_res *Absolute residuals of kernel regression of x on y.*

Description

This internal function calls the `kern` function to implement kernel regression with the option `residuals=TRUE` and returns absolute residuals.

Usage

```
abs_res(x, y)
```

Arguments

x	vector of data on the dependent variable
y	vector of data on the regressor

Details

The first argument is assumed to be the dependent variable. If `abs_res(x,y)` is used, you are regressing x on y (not the usual y on x)

Value

absolute values of kernel regression residuals are returned.

Note

This function is intended for internal use.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

Examples

```
## Not run:
set.seed(330)
x=sample(20:50)
y=sample(20:50)
abs_res(x,y)

## End(Not run)
```

`abs_stdapd` *Absolute values of gradients (apd's) of kernel regressions of x on y when both x and y are standardized.*

Description

1) standardize the data to force mean zero and variance unity, 2) kernel regress x on y, with the option 'gradients = TRUE' and finally 3) compute the absolute values of gradients

Usage

```
abs_stdapd(x, y)
```

Arguments

x	vector of data on the dependent variable
y	data on the regressors which can be a matrix

Details

The first argument is assumed to be the dependent variable. If `abs_stdapd(x,y)` is used, you are regressing x on y (not the usual y on x). The regressors can be a matrix with 2 or more columns. The missing values are suitably ignored by the standardization.

Value

Absolute values of kernel regression gradients are returned after standardizing the data on both sides so that the magnitudes of amorphous partial derivatives (apd's) are comparable between regression of x on y on the one hand and regression of y on x on the other.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

Examples

```
## Not run:
set.seed(330)
x=sample(20:50)
y=sample(20:50)
abs_stdapd(x,y)

## End(Not run)
```

abs_stdapdC *Absolute values of gradients (apd's) of kernel regressions of x on y when both x and y are standardized and control variables are present.*

Description

1) standardize the data to force mean zero and variance unity, 2) kernel regress x on y and a matrix of control variables, with the option 'gradients = TRUE' and finally 3) compute the absolute values of gradients

Usage

```
abs_stdapdC(x, y, ctrl)
```

Arguments

x	vector of data on the dependent variable
y	data on the regressors which can be a matrix
ctrl	Data matrix on the control variable(s) beyond causal path issues

Details

The first argument is assumed to be the dependent variable. If `abs_stdapdC(x,y)` is used, you are regressing x on y (not the usual y on x). The regressors can be a matrix with 2 or more columns. The missing values are suitably ignored by the standardization.

Value

Absolute values of kernel regression gradients are returned after standardizing the data on both sides so that the magnitudes of amorphous partial derivatives (apd's) are comparable between regression of x on y on the one hand and regression of y on x on the other.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

See Also

See `abs_stdapd`.

Examples

```
## Not run:
set.seed(330)
x=sample(20:50)
y=sample(20:50)
z=sample(20:50)
abs_stdapdC(x,y,ctrl=z)
```

```
## End(Not run)
```

`abs_stdres` *Absolute values of residuals of kernel regressions of x on y when both x and y are standardized.*

Description

1) Standardize the data to force mean zero and variance unity, 2) kernel regress x on y, with the option 'residuals = TRUE' and finally 3) compute the absolute values of residuals.

Usage

```
abs_stdres(x, y)
```

Arguments

`x`	vector of data on the dependent variable
`y`	data on the regressors which can be a matrix

Details

The first argument is assumed to be the dependent variable. If `abs_stdres(x,y)` is used, you are regressing x on y (not the usual y on x). The regressors can be a matrix with 2 or more columns. The missing values are suitably ignored by the standardization.

Value

Absolute values of kernel regression residuals are returned after standardizing the data on both sides so that the magnitudes of residuals are comparable between regression of x on y on the one hand and regression of y on x on the other.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: 10.1080/03610918.2015.1122048

Examples

```
## Not run:
set.seed(330)
x=sample(20:50)
y=sample(20:50)
abs_stdres(x,y)

## End(Not run)
```

`abs_stdresC` *Absolute values of residuals of kernel regressions of x on y when both x and y are standardized and control variables are present (C for control presence).*

Description

1) standardize the data to force mean zero and variance unity, 2) kernel regress x on y and a matrix of control variables, with the option 'residuals = TRUE' and finally 3) compute the absolute values of residuals.

Usage

```
abs_stdresC(x, y, ctrl)
```

Arguments

`x`	vector of data on the dependent variable
`y`	data on the regressors which can be a matrix
`ctrl`	Data matrix on the control variable(s) beyond causal path issues

Details

The first argument is assumed to be the dependent variable. If `abs_stdres(x,y)` is used, you are regressing x on y (not the usual y on x). The regressors can be a matrix with two or more columns. The missing values are suitably ignored by the standardization.

Value

Absolute values of kernel regression residuals are returned after standardizing the data on both sides so that the magnitudes of residuals are comparable between regression of x on y on the one hand and regression of y on x on the other.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D.'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: 10.1080/03610918.2015.1122048

See Also

See `abs_stdres`.

Examples

```
## Not run:
set.seed(330)
x=sample(20:50)
y=sample(20:50)
z=sample(21:51)
abs_stdresC(x,y,ctrl=z)

## End(Not run)
```

`abs_stdrhserC` *Absolute residuals kernel regressions of standardized x on y and control variables, Cr1 has abs(RHS*y) not gradients.*

Description

1) standardize the data to force mean zero and variance unity, 2) kernel regress x on y and a matrix of control variables, with the option 'residuals = TRUE' and finally 3) compute the absolute values of residuals.

Usage

```
abs_stdrhserC(x, y, ctrl, ycolumn = 1)
```

Arguments

`x`	vector of data on the dependent variable
`y`	data on the regressors which can be a matrix
`ctrl`	Data matrix on the control variable(s) beyond causal path issues
`ycolumn`	if y has more than one column, the column number used when multiplying residuals times this column of y, default=1 or first column of y matrix is used

Details

The first argument is assumed to be the dependent variable. If `abs_stdrhserC(x,y)` is used, you are regressing x on y (not the usual y on x). The regressors can be a matrix with 2 or more columns. The missing values are suitably ignored by the standardization.

Value

Absolute values of kernel regression residuals are returned after standardizing the data on both sides so that the magnitudes of residuals are comparable between regression of x on y on the one hand and regression of y on x on the other.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: 10.1080/03610918.2015.1122048

See Also

See abs_stdres.

Examples

```
## Not run:
set.seed(330)
x=sample(20:50)
y=sample(20:50)
z=sample(21:51)
abs_stdrhserC(x,y,ctrl=z)

## End(Not run)
```

---

abs_stdrhserr *Absolute values of Hausman-Wu null in kernel regressions of x on y when both x and y are standardized.*

Description

1) standardize the data to force mean zero and variance unity, 2) kernel regress x on y, with the option 'gradients = TRUE' and finally 3) compute the absolute values of Hausman-Wu null hypothesis for testing exogeneity, or E(RHS.regressor*error)=0 where error is approximated by kernel regression residuals

Usage

```
abs_stdrhserr(x, y)
```

Arguments

x vector of data on the dependent variable

y data on the regressors which can be a matrix

Details

The first argument is assumed to be the dependent variable. If `abs_stdrhserr(x,y)` is used, you are regressing x on y (not the usual y on x). The regressors can be a matrix with 2 or more columns. The missing values are suitably ignored by the standardization.

Value

Absolute values of kernel regression RHS*residuals are returned after standardizing the data on both sides so that the magnitudes of Hausman-Wu null values are comparable between regression of x on y on the one hand and flipped regression of y on x on the other.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

Examples

```
## Not run:
set.seed(330)
x=sample(20:50)
y=sample(20:50)
abs_stdrhserr(x,y)

## End(Not run)
```

---

allPairs *Report causal identification for all pairs of variables in a matrix (deprecated function). It is better to choose a target variable and pair it with all others, instead of considering all possible targets.*

Description

This studies all possible (perhaps too many) causal directions in a matrix. It is deprecated because it uses older criterion 1 by caling `abs_stdapd` I recommend using `causeSummary` or its block version `cuseSummBlk`. This uses `abs_stdres`, `comp_portfo2`, etc. and returns a matrix with 7 columns having detailed output. Criterion 1 has been revised as described in Vinod (2019) and is known to work better.

Usage

```
allPairs(mtx, dig = 6, verbo = FALSE, typ = 1, rnam = FALSE)
```

Arguments

mtx	Input matrix with variable names
dig	Digits of accuracy in reporting (=6, default)
verbo	Logical variable, set to 'TRUE' if printing is desired
typ	Causal direction criterion number (typ=1 is default) Criterion 1 (Cr1) compares kernel regression absolute values of gradients. Criterion 2 (Cr2) compares kernel regression absolute values of residuals. Criterion 3 (Cr3) compares kernel regression based r*(x|y) with r*(y|x).
rnam	Logical variable, default `rnam=FALSE` means the user does not want the row names to be (somewhat too cleverly) assigned by the function.

Value

A 7-column matrix called 'outcause' with names of variables X and Y in the first two columns and the name of the 'causal' variable in 3rd col. Remaining four columns report numerical computations of SD1 to SD4, r*(x|y), r*(y|x). Pearson r and p-values for its traditional significance testing.

Note

The cause reported in the third column is identified from the sign of the first SD1 only, ignoring SD2, SD3 and SD4 under both Cr1 and Cr2. It is a good idea to loop a call to this function with typ=1:3. One can print the resulting 'outcause' matrix with the `xtable(outcause)` for the Latex output. A similar deprecated function included in this package, called `some0Pairs`, incorporates all SD1 to SD4 and all three criteria Cr1 rto Cr3 to report a 'sum' of indexes representing the signed number whose sign can more comprehensively help determine the causal direction(s). Since the Cr1 here is revised in later work, this is deprecated.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D.'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, https://doi.org/gffn86

Vinod, H. D. 'New exogeneity tests and causal paths,' Chapter 2 in 'Handbook of Statistics: Conceptual Econometrics Using R', Vol.32, co-editors: H. D. Vinod and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2019, pp. 33-64.

See Also

See Also `somePairs`, `some0Pairs` `causeSummary`

Examples

```
data(mtcars)
options(np.messages=FALSE)
for(j in 1:3){
a1=allPairs(mtcars[,1:3], typ=j)
print(a1)}
```

---

badCol *internal badCol*

Description

intended for internal use

Usage

```
data(badCol)
```

Format

The format is: int 4

---

bigfp *Compute the numerical integration by the trapezoidal rule.*

Description

See page 220 of Vinod (2008) "Hands-on Intermediate Econometrics Using R," for the trapezoidal integration formula needed for stochastic dominance. The book explains pre-multiplication by two large sparse matrices denoted by I_F, I_f. Here we accomplish the same computation without actually creating the large sparse matrices. For example, the I_f is replaced by `cumsum` in this code (unlike the R code in my textbook).

Usage

```
bigfp(d, p)
```

Arguments

d	A vector of consecutive interval lengths, upon combining both data vectors
p	Vector of probabilities of the type 1/2T, 2/2T, 3/2T, etc. to 1.

Value

Returns a result after pre-multiplication by I_F, I_f matrices, without actually creating the large sparse matrices. This is an internal function.

Note

This is an internal function, called by the function `stochdom2`, for comparison of two portfolios in terms of stochastic dominance (SD) of orders 1 to 4. Typical usage is: `sd1b=bigfp(d=dj,p=rhs)` `sd2b=bigfp(d=dj,p=sd1b)` `sd3b=bigfp(d=dj,p=sd2b)` `sd4b=bigfp(d=dj,p=sd3b)`. This produces numerical evaluation vectors for the four orders, SD1 to SD4.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D.', 'Hands-On Intermediate Econometrics Using R' (2008) World Scientific Publishers: Hackensack, NJ. https://www.worldscientific.com/worldscibooks/10.1142/6895

`bootDom12` *bootstrap confidence intervals for (x2-x1) exact SD1 to SD4 stochastic dominance .*

Description

This calls the meboot package to create J=999 replications of portfolio return matrices and compute 95% confidence intervals on x1, x2 and their difference (x2-x1). If the interval on (x2-x1) conta.ins zero the choice between the two can reverse due to sampling variation

Usage

```
bootDom12(x1, x2, confLevel = 95, reps = 999)
```

Arguments

`x1`	a vector of n portfolio returns
`x2`	a vector of n portfolio returns
`confLevel`	confidene level confLevel=95 is default
`reps`	number of bootstrap resamples, default is reps=999

Value

A matrix with six columns. First two Low1 and Upp1 are confidence interval limits for x1. Next two columns have analogous limits for x2. The last but first columns entitled Lowx2mx1 means lower confidence limit for (x2-x1), where m=minus. The last column entitled Uppx2mx1 means upper confidence limit for (x2-x1).

For strong stochastic dominance of x2 over x1 dominance beyond sampling variability, zero should not be inside the confidence interval in the last two columns.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

See Also

see `exactSdMtx`

---

bootGcLC — *Compute vector of n999 nonlinear Granger causality paths*

Description

Maximum entropy bootstrap (meboot) package is used for statistical inference The bootstrap output can be analyzed to estimate an approximate confidence interval on sample-based direction of the causal path. The LC in the function name stands for local constant. Kernel regression np package options regtype="lc" for local constant, and bwmethod="cv.ls" for least squares-based bandwidth selection are fixed.

Usage

```
bootGcLC(x1, x2, px2 = 4, px1 = 4, pwanted = 4, ctrl = 0, n999 = 9)
```

Arguments

`x1`	The data vector x1
`x2`	The data vector x2
`px2`	number of lags of x2 in the data, default px2=4
`px1`	number of lags of x1 in the data default px1=4
`pwanted`	number of lags of both x2 and x1 wanted for Granger causal analysis, default =4
`ctrl`	data matrix having control variable(s) if any
`n999`	Number of bootstrap replications (default=9)

Value

out is n999 X 3 matrix for 3 outputs of GcauseX12 resampled

Note

This computation is computer intensive and generally very slow. It may be better to use this function it at a later stage in the investigation, after a preliminary causal determination is already made. The 3 outputs of GauseX12 are two Rsquares and the difference between after subtracting the second from the first. Col. 1 has (RsqX1onX2) Col.2 has (RsqX2onX1), and Col.3 has dif=(RsqX1onX2 -RsqX2onX1) Note that R-squares are always positive. If dif>0, RsqX1onX2>RsqX2onX1, implying that x2 on RHS performs better that is, x2 –> x1 is the path, or x2 Granger-causes x1. If dif<0, x1 –> x2 holds. If dif is too close to zero, we may have bidirectional causality x1 <–> x2. The proportion of resamples (out of n999) having dif<0 suggests level of confidence in the conclusion x1 –> x2. The proportion of resamples (out of n999) having dif>0 suggests level of confidence in the conclusion x2 –> x1.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, https://doi.org/gffn86

Zheng, S., Shi, N.-Z., and Zhang, Z. (2012). Generalized measures of correlation for asymmetry, nonlinearity, and beyond. Journal of the American Statistical Association, vol. 107, pp. 1239-1252.

Vinod, H. D. and Lopez-de-Lacalle, J. (2009). 'Maximum entropy bootstrap for time series: The meboot R package.' Journal of Statistical Software, Vol. 29(5), pp. 1-19.

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: https://www.ssrn.com/abstract=2982128

See Also

See Also GcRsqX12c.

Examples

```
## Not run:
library(Ecdat);options(np.messages=FALSE);attach(data.frame(MoneyUS))
bootGcLC(y,m,n999=9)

## End(Not run)
## Not run:
library(lmtest); data(ChickEgg);attach(data.frame(ChickEgg))
b2=bootGcLC(x1=chicken,x2=egg,pwanted=3,px1=3,px2=3,n999=99)

## End(Not run)
```

`bootGcRsq` *Compute vector of n999 nonlinear Granger causality paths*

Description

Maximum entropy bootstrap (meboot) package is used for statistical inference The bootstrap output can be analyzed to estimate an approximate confidence interval on sample-based direction of the causal path. Kernel regression np package options regtype="ll" for local linear, and bwmethod="cv.aic" for AIC-based bandwidth selection are fixed.

Usage

```
bootGcRsq(x1, x2, px2 = 4, px1 = 4, pwanted = 4, ctrl = 0, n999 = 9)
```

Arguments

`x1`	The data vector x1
`x2`	The data vector x2
`px2`	number of lags of x2 in the data, default px2=4
`px1`	number of lags of x1 in the data default px1=4
`pwanted`	number of lags of both x2 and x1 wanted for Granger causal analysis, default =4
`ctrl`	data matrix having control variable(s) if any
`n999`	Number of bootstrap replications (default=9)

Value

out is n999 X 3 matrix for 3 outputs of GcauseX12 resampled

Note

This computation is computer intensive and generally very slow. It may be better to use this function it at a later stage in the investigation, after a preliminary causal determination is already made. The 3 outputs of GauseX12 are two Rsquares and the difference between them after subtracting the second from the first. Col. 1 has (RsqX1onX2), Col.2 has (RsqX2onX1), and Col.3 has dif=(RsqX1onX2 -RsqX2onX1) Note that R-squares are always positive. If dif>0, RsqX1onX2>RsqX2onX1, implying that x2 on RHS performs better that is, x2 –> x1 is the causal path. If dif<0, x1 –> x2 holds. If dif is too close to zero, we may have bidirectional causality x1 <–> x2. The proportion of resamples (out of n999) having dif<0 suggests level of confidence in the conclusion x1 –> x2. The proportion of resamples (out of n999) having dif>0 suggests level of confidence in the conclusion x2 –> x1.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, https://doi.org/gffn86

Zheng, S., Shi, N.-Z., and Zhang, Z. (2012). Generalized measures of correlation for asymmetry, nonlinearity, and beyond. Journal of the American Statistical Association, vol. 107, pp. 1239-1252.

Vinod, H. D. and Lopez-de-Lacalle, J. (2009). 'Maximum entropy bootstrap for time series: The meboot R package.' Journal of Statistical Software, Vol. 29(5), pp. 1-19.

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: https://www.ssrn.com/abstract=2982128

See Also

See Also GcRsqX12.

Examples

```
## Not run:
library(Ecdat);options(np.messages=FALSE);attach(data.frame(MoneyUS))
bootGcRsq(y,m,n999=9)

## End(Not run)
## Not run:
library(lmtest); data(ChickEgg);attach(data.frame(ChickEgg))
options(np.messages=FALSE)
b2=bootGcLC(x1=chicken,x2=egg,pwanted=3,px1=3,px2=3,n999=99)
Fn=function(x)quantile(x,prob=c(0.025, 0.975))#confInt
apply(b1,2,Fn)#reports 95 percent confidence interval

## End(Not run)
```

---

bootPair2 — *Compute matrix of n999 rows and p-1 columns of bootstrap 'sum' (scores from Cr1 to Cr3).*

Description

The '2' in the name of the function suggests a second implementation of 'bootPair,' where exact stochastic dominance, decileVote, and momentVote are used. Maximum entropy bootstrap (meboot) package is used for statistical inference using the sum of three signs sg1 to sg3, from the three criteria Cr1 to Cr3, to assess preponderance of evidence in favor of a sign, (+1, 0, -1). The bootstrap output can be analyzed to assess the approximate preponderance of a particular sign which determines the causal direction.

Usage

```
bootPair2(mtx, ctrl = 0, n999 = 9)
```

Arguments

`mtx`	data matrix with two or more columns
`ctrl`	data matrix having control variable(s) if any
`n999`	Number of bootstrap replications (default=9)

Value

Function creates a matrix called 'out'. If the input to the function called `mtx` has p columns, the output `out` of `bootPair2(mtx)` is a matrix of n999 rows and p-1 columns, each containing resampled 'sum' values summarizing the weighted sums associated with all three criteria from the function `silentPair2(mtx)` applied to each bootstrap sample separately.

Note

This computation is computer-intensive and generally very slow. It may be better to use it later in the investigation, after a preliminary causal determination is already made. A positive sign for j-th weighted sum reported in the column 'sum' means that the first variable listed in the argument matrix `mtx` is the 'kernel cause' of the variable in the (j+1)-th column of `mtx`.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, https://doi.org/gffn86

Zheng, S., Shi, N.-Z., and Zhang, Z. (2012). Generalized measures of correlation for asymmetry, nonlinearity, and beyond. Journal of the American Statistical Association, vol. 107, pp. 1239-1252.

Vinod, H. D. and Lopez-de-Lacalle, J. (2009). 'Maximum entropy bootstrap for time series: The meboot R package.' Journal of Statistical Software, Vol. 29(5), pp. 1-19.

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: https://www.ssrn.com/abstract=2982128

Vinod, Hrishikesh D., R Package GeneralCorr Functions for Portfolio Choice (November 11, 2021). Available at SSRN: https://ssrn.com/abstract=3961683

Vinod, Hrishikesh D., Stochastic Dominance Without Tears (January 26, 2021). Available at SSRN: https://ssrn.com/abstract=3773309

See Also

See Also `silentPair2`.

Examples

```
## Not run:
options(np.messages = FALSE)
set.seed(34);x=sample(1:10);y=sample(2:11)
bb=bootPair2(cbind(x,y),n999=29)
apply(bb,2,summary) #gives summary stats for n999 bootstrap sum computations

bb=bootPair2(airquality,n999=999);options(np.messages=FALSE)
apply(bb,2,summary) #gives summary stats for n999 bootstrap sum computations

data('EuroCrime')
attach(EuroCrime)
bootPair2(cbind(crim,off),n999=29)#First col. crim causes officer deployment,
#hence positives signs are most sensible for such call to bootPairs
#note that n999=29 is too small for real problems, chosen for quickness here.

## End(Not run)
```

`bootPairs` *Compute matrix of n999 rows and p-1 columns of bootstrap 'sum' (strength from Cr1 to Cr3).*

Description

Maximum entropy bootstrap (meboot) package is used for statistical inference using the sum of three signs sg1 to sg3 from the three criteria Cr1 to Cr3 to assess preponderance of evidence in favor of a sign. (+1, 0, -1). The bootstrap output can be analyzed to assess approximate preponderance of a particular sign which determines the causal direction.

Usage

```
bootPairs(mtx, ctrl = 0, n999 = 9)
```

Arguments

`mtx`	data matrix with two or more columns
`ctrl`	data matrix having control variable(s) if any
`n999`	Number of bootstrap replications (default=9)

Value

out When `mtx` has p columns, `out` of `bootPairs(mtx)` is a matrix of n999 rows and p-1 columns each containing resampled 'sum' values summarizing the weighted sums associated with all three criteria from the function `silentPairs(mtx)` applied to each bootstrap sample separately.

Note

This computation is computer intensive and generally very slow. It may be better to use it at a later stage in the investigation when a preliminary causal determination is already made. A positive sign for j-th weighted sum reported in the column 'sum' means that the first variable listed in the argument matrix mtx is the 'kernel cause' of the variable in the (j+1)-th column of mtx.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, https://doi.org/gffn86

Zheng, S., Shi, N.-Z., and Zhang, Z. (2012). Generalized measures of correlation for asymmetry, nonlinearity, and beyond. Journal of the American Statistical Association, vol. 107, pp. 1239-1252.

Vinod, H. D. and Lopez-de-Lacalle, J. (2009). 'Maximum entropy bootstrap for time series: The meboot R package.' Journal of Statistical Software, Vol. 29(5), pp. 1-19.

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: https://www.ssrn.com/abstract=2982128

See Also

See Also silentPairs.

Examples

```
## Not run:
options(np.messages = FALSE)
set.seed(34);x=sample(1:10);y=sample(2:11)
bb=bootPairs(cbind(x,y),n999=29)
apply(bb,2,summary) #gives summary stats for n999 bootstrap sum computations

bb=bootPairs(airquality,n999=999);options(np.messages=FALSE)
apply(bb,2,summary) #gives summary stats for n999 bootstrap sum computations

data('EuroCrime')
attach(EuroCrime)
bootPairs(cbind(crim,off),n999=29)#First col. crim causes officer deployment,
#hence positives signs are most sensible for such call to bootPairs
#note that n999=29 is too small for real problems, chosen for quickness here.

## End(Not run)
```

`bootPairs0` *Compute matrix of n999 rows and p-1 columns of bootstrap 'sum' index (strength from older criterion Cr1, with newer Cr2 and Cr3).*

Description

Maximum entropy bootstrap (meboot) package is used for statistical inference using the sum of three signs sg1 to sg3 from the three criteria Cr1 to Cr3 to assess preponderance of evidence in favor of a sign. (+1, 0, -1). The bootstrap output can be analyzed to assess approximate preponderance of a particular sign which determines the causal direction.

Usage

```
bootPairs0(mtx, ctrl = 0, n999 = 9)
```

Arguments

`mtx`	data matrix with two or more columns
`ctrl`	data matrix having control variable(s) if any
`n999`	Number of bootstrap replications (default=9)

Value

out When `mtx` has p columns, `out` of `bootPairs(mtx)` is a matrix of n999 rows and p-1 columns each containing resampled 'sum' values summarizing the weighted sums associated with all three criteria from the function `silentPairs(mtx)` applied to each bootstrap sample separately.

Note

This computation is computer intensive and generally very slow. It may be better to use it at a later stage in the investigation when a preliminary causal determination is already made. A positive sign for j-th weighted sum reported in the column 'sum' means that the first variable listed in the argument matrix `mtx` is the 'kernel cause' of the variable in the (j+1)-th column of `mtx`.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, https://doi.org/gffn86

Zheng, S., Shi, N.-Z., and Zhang, Z. (2012). Generalized measures of correlation for asymmetry, nonlinearity, and beyond. Journal of the American Statistical Association, vol. 107, pp. 1239-1252.

Vinod, H. D. and Lopez-de-Lacalle, J. (2009). 'Maximum entropy bootstrap for time series: The meboot R package.' Journal of Statistical Software, Vol. 29(5), pp. 1-19.

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: https://www.ssrn.com/abstract=2982128

See Also

See Also silentPairs0, bootPairs has the version with later version of Cr1.

Examples

```
## Not run:
options(np.messages = FALSE)
set.seed(34);x=sample(1:10);y=sample(2:11)
bb=bootPairs0(cbind(x,y),n999=29)
apply(bb,2,summary) #gives summary stats for n999 bootstrap sum computations

bb=bootPairs0(airquality,n999=999);options(np.messages=FALSE)
apply(bb,2,summary) #gives summary stats for n999 bootstrap sum computations

data('EuroCrime')
attach(EuroCrime)
bootPairs0(cbind(crim,off),n999=29)#First col. crim causes officer deployment,
#hence positives signs are most sensible for such call to bootPairs
#note that n999=29 is too small for real problems, chosen for quickness here.

## End(Not run)
```

---

bootQuantile — *Compute confidence intervals [quantile(s)] of indexes from bootPairs output*

Description

Begin with the output of bootPairs function, a (n999 by p-1) matrix when there are p columns of data, `bootQuantile` produces a (k by p-1) mtx of quantile(s) of bootstrap ouput assuming that there are k quantiles needed.

Usage

```
bootQuantile(out, probs = c(0.025, 0.975), per100 = TRUE)
```

Arguments

`out`	output from bootPairs with p-1 columns and n999 rows
`probs`	quantile evaluation probabilities. The default is k=2, probs=c(.025,0.975) for a 95 percent confidence interval. Note that there are k=2 quantiles desired for each column with this specification
`per100`	logical (default per100=TRUE) to change the range of 'sum' to [-100, 100] values which are easier to interpret

Value

CI k quantiles evaluated at probs as a matrix with k rows and quantile of pairwise p-1 indexes representing p-1 column pairs (fixing the first column in each pair) This function summarizes the output of of `bootPairs(mtx)` (a n999 by p-1 matrix) each containing resampled 'sum' values summarizing the weighted sums associated with all three criteria from the function `silentPairs(mtx)` applied to each bootstrap sample separately. #'

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, https://doi.org/gffn86

Vinod, H. D. and Lopez-de-Lacalle, J. (2009). 'Maximum entropy bootstrap for time series: The meboot R package.' Journal of Statistical Software, Vol. 29(5), pp. 1-19.

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: https://www.ssrn.com/abstract=2982128

See Also

See Also `silentPairs`.

Examples

```
## Not run:
options(np.messages = FALSE)
set.seed(34);x=sample(1:10);y=sample(2:11)
bb=bootPairs(cbind(x,y),n999=29)
bootQuantile(bb) #gives summary stats for n999 bootstrap sum computations

bb=bootPairs(airquality,n999=999);options(np.messages=FALSE)
bootQuantile(bb,tau=0.476)#signs for n999 bootstrap sum computations

data('EuroCrime')
attach(EuroCrime)
bb=bootPairs(cbind(crim,off),n999=29) #col.1= crim causes off
#hence positive signs are more intuitively meaningful.
#note that n999=29 is too small for real problems, chosen for quickness here.
bootQuantile(bb)# quantile matrix for n999 bootstrap sum computations

## End(Not run)
```

`bootSign`	*Probability of unambiguously correct (+ or -) sign from bootPairs output*

Description

If there are p columns of data, `bootSign` produces a p-1 by 1 vector of probabilities of correct signs assuming that the mean of n999 values has the correct sign and assuming that m of the 'sum' index values inside the range [-tau, tau] are neither positive nor negative but indeterminate or ambiguous (being too close to zero). That is, the denominator of P(+1) or P(-1) is (n999-m) if m signs are too close to zero. Thus it measures the bootstrap success rate in identifying the correct sign, when the sign of the average of n999 bootstraps is assumed to be correct.

Usage

```
bootSign(out, tau = 0.476)
```

Arguments

`out`	output from bootPairs with p-1 columns and n999 rows
`tau`	threshold to determine what value is too close to zero, default tau=0.476 is equivalent to 15 percent threshold for the unanimity index ui

Value

sgn When `mtx` has p columns, `sgn` reports pairwise p-1 signs representing (fixing the first column in each pair) the average sign after averaging the output of of `bootPairs(mtx)` (a n999 by p-1 matrix) each containing resampled 'sum' values summarizing the weighted sums associated with all three criteria from the function `silentPairs(mtx)` applied to each bootstrap sample separately. #'

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, https://doi.org/gffn86

Vinod, H. D. and Lopez-de-Lacalle, J. (2009). 'Maximum entropy bootstrap for time series: The meboot R package.' Journal of Statistical Software, Vol. 29(5), pp. 1-19.

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: https://www.ssrn.com/abstract=2982128

See Also

See Also `silentPairs`, `bootQuantile`, `bootSignPcent`.

Examples

```
## Not run:
options(np.messages = FALSE)
set.seed(34);x=sample(1:10);y=sample(2:11)
bb=bootPairs(cbind(x,y),n999=29)
bootSign(bb,tau=0.476) #gives success rate in n999 bootstrap sum computations

bb=bootPairs(airquality,n999=999);options(np.messages=FALSE)
bootSign(bb,tau=0.476)#signs for n999 bootstrap sum computations

data('EuroCrime');options(np.messages=FALSE)
attach(EuroCrime)
bb=bootPairs(cbind(crim,off),n999=29) #col.1= crim causes off
#hence positive signs are more intuitively meaningful.
#note that n999=29 is too small for real problems, chosen for quickness here.
bootSign(bb,tau=0.476)#gives success rate in n999 bootstrap sum computations

## End(Not run)
```

`bootSignPcent` *Probability of unambiguously correct (+ or -) sign from bootPairs output transformed to percentages.*

Description

If there are p columns of data, `bootSignPcent` produces a p-1 by 1 vector of probabilities of correct signs assuming that the mean of n999 values has the correct sign and assuming that m of the 'ui' index values inside the range [-tau, tau] are neither positive nor negative but indeterminate or ambiguous (being too close to zero). That is, the denominator of P(+1) or P(-1) is (n999-m) if m signs are too close to zero. Thus it measures the bootstrap success rate in identifying the correct sign, when the sign of the average of n999 bootstraps is assumed to be correct.

Usage

```
bootSignPcent(out, tau = 5)
```

Arguments

`out`	output from bootPairs with p-1 columns and n999 rows
`tau`	threshold to determine what value is too close to zero, default tau=5 is 5 percent threshold for the unanimity index ui

Value

sgn When `mtx` has p columns, `sgn` reports pairwise p-1 signs representing (fixing the first column in each pair) the average sign after averaging the output of of `bootPairs(mtx)` (a n999 by p-1 matrix) each containing resampled 'sum' values summarizing the weighted sums associated with all three criteria from the function `silentPairs(mtx)` applied to each bootstrap sample separately. #'

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, https://doi.org/gffn86

Vinod, H. D. and Lopez-de-Lacalle, J. (2009). 'Maximum entropy bootstrap for time series: The meboot R package.' Journal of Statistical Software, Vol. 29(5), pp. 1-19.

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: https://www.ssrn.com/abstract=2982128

See Also

See Also `silentPairs`, `bootQuantile`, `bootSign`.

Examples

```
## Not run:
options(np.messages = FALSE)
set.seed(34);x=sample(1:10);y=sample(2:11)
bb=bootPairs(cbind(x,y),n999=29)
bootSignPcent(bb,tau=5) #gives success rate in n999 bootstrap sum computations

bb=bootPairs(airquality,n999=999);options(np.messages=FALSE)
bootSignPcent(bb,tau=5)#success rate for signs from n999 bootstraps

data('EuroCrime');options(np.messages=FALSE)
attach(EuroCrime)
bb=bootPairs(cbind(crim,off),n999=29) #col.1= crim causes off
#hence positive signs are more intuitively meaningful.
#note that n999=29 is too small for real problems, chosen for quickness here.
bootSignPcent(bb,tau=5)#successful signs from n999 bootstraps

## End(Not run)
```

`bootSummary` *Compute usual summary stats of 'sum' indexes from bootPairs output*

Description

Begin with the output of bootPairs function, a (n999 by p-1) matrix when there are p columns of data, `bootSummary` produces a (6 by p-1) mtx of summary of bootstrap ouput (Min, 1st Qu,Median, Mean, 3rd Qi.,Max)

Usage

```
bootSummary(out, per100 = TRUE)
```

Arguments

`out`	output from bootPairs with p-1 columns and n999 rows in input here
`per100`	logical (default per100=TRUE) to change the range of 'sum' to [-100, 100] values which are easier to interpret

Value

summ summary output from the (n999 by p-1) matrix output of `bootPairs(mtx)` each containing resampled 'sum' values summarizing the weighted sums associated with all three criteria from the function `silentPairs(mtx)` applied to each bootstrap sample separately.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, https://doi.org/gffn86

Vinod, H. D. and Lopez-de-Lacalle, J. (2009). 'Maximum entropy bootstrap for time series: The meboot R package.' Journal of Statistical Software, Vol. 29(5), pp. 1-19.

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: https://www.ssrn.com/abstract=2982128

See Also

See Also `silentPairs`.

Examples

```
## Not run:
options(np.messages = FALSE)
set.seed(34);x=sample(1:10);y=sample(2:11)
bb=bootPairs(cbind(x,y),n999=29)
bootSummary(bb) #gives summary stats for n999 bootstrap sum computations

bb=bootPairs(airquality,n999=999);options(np.messages=FALSE)
bootSummary(bb)#signs for n999 bootstrap sum computations

data('EuroCrime')
attach(EuroCrime)
bb=bootPairs(cbind(crim,off),n999=29) #col.1= crim causes off
#hence positive signs are more intuitively meaningful.
#note that n999=29 is too small for real problems, chosen for quickness here.
bootSummary(bb)#signs for n999 bootstrap sum computations

## End(Not run)
```

`bootSummary2` *Compute usual summary stats of 'sum' index in (-100, 100) from bootPair2*

Description

The '2' in the name of the function suggests a second implementation where exact stochastic dominance, decileVote and momentVote are used. Begin with the output of bootPairs function, a (n999 by p-1) matrix when there are p columns of data, `bootSummary` produces a (6 by p-1) mtx of summary of bootstrap ouput (Min, 1st Qu,Median, Mean, 3rd Qi.,Max)

Usage

```
bootSummary2(out, per100 = TRUE)
```

Arguments

`out`	output from bootPair2 with p-1 columns and n999 rows in input here
`per100`	logical (default per100=TRUE) to change the range of 'sum' to [-100, 100] values which are easier to interpret

Value

summ a summary matrix (n999 by p-1) having usual parameters using the output of `bootPair2(mtx)` Each containing resampled 'sum' values summarizing the weighted sums associated with all three criteria from the function `silentPair2(mtx)` applied to each bootstrap sample separately.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, https://doi.org/gffn86

Vinod, H. D. and Lopez-de-Lacalle, J. (2009). 'Maximum entropy bootstrap for time series: The meboot R package.' Journal of Statistical Software, Vol. 29(5), pp. 1-19.

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: https://www.ssrn.com/abstract=2982128

See Also

See Also `silentPairs`.

Examples

```
## Not run:
options(np.messages = FALSE)
set.seed(34);x=sample(1:10);y=sample(2:11)
bb=bootPair2(cbind(x,y),n999=29)
bootSummary2(bb) #gives summary stats for n999 bootstrap sum computations

bb=bootPair2(airquality,n999=999);options(np.messages=FALSE)
bootSummary2(bb)#signs for n999 bootstrap sum computations

data('EuroCrime')
attach(EuroCrime)
bb=bootPair2(cbind(crim,off),n999=29) #col.1= crim causes off
#hence positive signs are more intuitively meaningful.
#note that n999=29 is too small for real problems, chosen for quickness here.
bootSummary2(bb)#signs for n999 bootstrap sum computations

## End(Not run)
```

`canonRho` *Generalized canonical correlation, estimating alpha, beta, rho.*

Description

This function uses data on two sets of column vectors. LHS set [x1, x2 .. xr] has r=nLHS number of columns with coefficients alpha and the larger RHS set [xr+1, xr+2, .. xp] has nRHS=(p-r) columns and RHS coefficients beta. Must arrange the sets so that the larger set in on RHS with coefficients beta estimated first from an eigenvector of the problem A* beta = rho^2 beta where A* is a partitioning of generalized matrix of (non-symmetric) correlation coefficients.

Usage

```
canonRho(mtx, nLHS = 2, sgn = 1, verbo = FALSE, ridg = c(0, 0))
```

Arguments

`mtx`	Input matrix of generalized correlation coefficients R*
`nLHS`	number of columns in the LHS set, default=2
`sgn`	preferred sign of coefficients default=1 for positive, use sgn= -1 if prior knowledge suggests that negative signs of coefficients are more realistic
`verbo`	logical, verbo=FALSE default means do not print results
`ridg`	two regularization constants added before computing matrix inverses of S11 and S22, respectively, with default=c(0,0). Some suggest ridg=c(0.01,0.01) for stable results

Value

A	eigenvalue computing matrix for Generalized canonical correlations
rho	Generalized canonical correlation coefficient
bet	RHS coefficient vector
alp	LHS coefficient vector

Note

This function calls kern,

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

Vinod, H. D. 'Matrix Algebra Topics in Statistics and Economics Using R', Chapter 4 in 'Handbook of Statistics: Computational Statistics with R', Vol.32, co-editors: M. B. Rao and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2014, pp. 143-176.

Vinod, H. D. 'Canonical ridge and econometrics of joint production,' Journal of Econometrics, vol. 4, 147–166.

Vinod, H. D. 'New exogeneity tests and causal paths,' Chapter 2 in 'Handbook of Statistics: Conceptual Econometrics Using R', Vol.32, co-editors: H. D. Vinod and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2019, pp. 33-64.

See Also

See gmcmtx0.

Examples

```
## Not run:
set.seed(99)
mtx2=matrix(sample(1:25),nrow=5)
g1=gmcmtx0(mtx2)
canonRho(g1,verbo=TRUE)

## End(Not run)#'
```

`causeAllPair` *All Pair Version Kernel (block) causality summary paths from three criteria*

Description

Allowing input matrix of control variables, this function produces a 5 column matrix summarizing the results where the estimated signs of stochastic dominance order values, (+1, 0, -1), are weighted by `wt=c(1.2,1.1,1.05,1)` to compute an overall result for all orders of stochastic dominance by a weighted sum for the criteria Cr1 and Cr2 and added to the Cr3 estimate as: (+1, 0, -1). The final range for the unanimity of sign index is [–100, 100].

Usage

```
causeAllPair(
  mtx,
  nam = colnames(mtx),
  blksiz = 10,
  ctrl = 0,
  dig = 6,
  wt = c(1.2, 1.1, 1.05, 1),
  sumwt = 4
)
```

Arguments

`mtx`	The data matrix with many columns, We consider causal paths among all possible pairs of mtx columns.
`nam`	vector of column names for `mtx`. Default: colnames(mtx)
`blksiz`	block size, default=10, if chosen blksiz >n, where n=rows in matrix then blksiz=n. That is, no blocking is done
`ctrl`	data matrix for designated control variable(s) outside causal paths
`dig`	Number of digits for reporting (default `dig`=6).
`wt`	Allows user to choose a vector of four alternative weights for SD1 to SD4.
`sumwt`	Sum of weights can be changed here =4(default).

Details

The reason for slightly declining weights on the signs from SD1 to SD4 stochastic dominance orders is simply their slightly increasing sampling unreliability due to higher order trapezoidal approximations of integrals of densities involved in definitions of SD1 to SD4. The summary results for all three criteria are reported in one matrix called `out`:

Value

If there are p columns in the input matrix, x1, x2, .., xp, say, there are choose(p,2) or [p*(p-1)/2] possible pairs and as many causal paths. This function returns a matrix of p*(p-1)/2 rows and 5 columns entitled: “cause", “response", “strength", “corr." and “p-value", respectively with self-explanatory titles. The first two columns have names of variables x1 or x(1+j), depending on which is the cause. The ‘strength’ column has absolute value of summary index in range [0,100] providing summary of causal results based on preponderance of evidence from criteria Cr1 to Cr3 from four orders of stochastic dominance, etc. The fourth column ‘corr.’ reports the Pearson correlation coefficient while the fifth column has the p-value for testing the null of zero Pearson coeff. This function merely calls `causeSumNoP` repeatedly to include all pairs. The background function `siPairsBlk` allows for control variables. The output of this function can be sent to ‘xtable’ for a nice Latex table.

Note

The European Crime data has all three criteria correctly suggesting that high crime rate kernel causes the deployment of a large number of police officers. Since Cr1 to Cr3 near unanimously suggest ‘crim’ as the cause of ‘off’, strength index 100 suggests unanimity. `attach(EuroCrime); causeSummary(cbind(crim,off))`

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

Vinod, H. D. ‘Generalized Correlation and Kernel Causality with Applications in Development Economics’ in Communications in Statistics -Simulation and Computation, 2015, https://doi.org/gffn86

Vinod, H. D. ’New exogeneity tests and causal paths,’ Chapter 2 in ’Handbook of Statistics: Conceptual Econometrics Using R’, Vol.32, co-editors: H. D. Vinod and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2019, pp. 33-64.

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: https://www.ssrn.com/abstract=2982128

See Also

See `bootPairs`, `causeSummBlk`

See `someCPairs`

`siPairsBlk`, `causeSummary`

Examples

```
## Not run:
mtx=data.frame(mtcars[,1:3]) #make sure columns of mtx have names
ctrl=data.frame(mtcars[,4:5])
```

```
 causeAllPair(mtx=mtx,ctrl=ctrl)

## End(Not run)

options(np.messages=FALSE)
set.seed(234)
z=runif(10,2,11)# z is independently created
x=sample(1:10)+z/10 #x is somewhat indep and affected by z
y=1+2*x+3*z+rnorm(10)
w=runif(10)
x2=x;x2[4]=NA;y2=y;y2[8]=NA;w2=w;w2[4]=NA
causeAllPair(mtx=cbind(x2,y2), ctrl=cbind(z,w2))
```

`causeSum2Blk` *Block Version 2: Kernel causality summary of causal paths from three criteria*

Description

The '2' in the name of the function suggests a second implementation where exact stochastic dominance, 'decileVote' and 'momentVote' functions are used, Block version allows a new bandwidth (chosen by the np package) while fitting kernel regressions for each block of data. This may not be appropriate in all situations. Block size is flexible. The function develops a unanimity index regarding which regression flip, (y on xi) or (xi on y) is the best. The "cause" is always on the right-hand side of a regression equation, and the superior flip gives the correct sign. The summary of all signs determines the causal direction and unanimity index among three criteria. This is a block version of `causeSummary2()`. While allowing the researcher to keep some variables as controls, or outside the scope of causal path determination (e.g., age or latitude) this function produces detailed causal path information in a 5 column matrix identifying the names of variables, causal path directions, path strengths re-scaled to be in the range [–100, 100], (table reports absolute values of the strength) plus Pearson correlation and its p-value.

The algorithm determines causal path directions from the sign of the strength index and strength index values by comparing three aspects of flipped kernel regressions: [x1 on (x2, x3, .. xp)] and its flipped version [x2 on (x1, x3, .. xp)] We compare (i) formal exogeneity test criterion, (ii) absolute residuals, and (iii) R-squares of the flipped regressions implying three criteria Cr1, to Cr3. The criteria are quantified by new methods using four orders of stochastic dominance, SD1 to SD4. See Vinod (2021) two SSRN papers.

Usage

```
causeSum2Blk(mtx, nam = colnames(mtx), blksiz = 10, ctrl = 0, dig = 6)
```

Arguments

`mtx`	The data matrix with many columns, y the first column is a fixed target, and then it is paired with all other columns, one by one, and still called x for flipping.

`nam`	vector of column names for `mtx`. Default: colnames(mtx)
`blksiz`	block size, default=10, if chosen blksiz >n, where n=rows in the matrix then blksiz=n. That is, no blocking is done
`ctrl`	data matrix for designated control variable(s) outside causal paths
`dig`	The number of digits for reporting (default `dig`=6).

Value

If there are p columns in the input matrix, x1, x2, .., xp, say, and if we keep x1 as a common member of all causal-direction-pairs (x1, x(1+j)) for (j=1, 2, .., p-1) which can be flipped. That is, either x1 is the cause or x(1+j) is the cause in a chosen pair. The control variables are not flipped. The printed output of this function reports the results for p-1 pairs indicating which variable (by name) causes which other variable (also by name). It also prints the strength or signed summary strength index in the range [-100,100]. A positive sign of the strength index means x1 kernel causes x(1+j), whereas negative strength index means x(1+j) kernel causes x1. The function also prints Pearson correlation and its p-value. This function also returns a matrix of p-1 rows and 5 columns entitled: “cause", “response", “strength", “corr." and “p-value", respectively with self-explanatory titles. The first two columns have names of variables x1 or x(1+j), depending on which is the cause. The ‘strength’ column has an absolute value of the summary index in the range [0,100], providing a summary of causal results based on the preponderance of evidence from Cr1 to Cr3 from deciles, moments, from four orders of stochastic dominance. The order of input columns in "mtx" matters. The fourth column, ‘corr.’, reports the Pearson correlation coefficient, while the fifth column has the p-value for testing the null of zero Pearson coefficient. This function calls `siPairsBlk`, allowing for control variables. The output of this function can be sent to ‘xtable’ for a nice Latex table.

Note

The European Crime data has all three criteria correctly suggesting that high crime rate kernel causes the deployment of a large number of police officers. If Cr1 to Cr3 near-unanimously suggest ‘crim’ as the cause of ‘off’, strength index would be near 100 suggesting unanimity. `attach(EuroCrime); causeSum2Blk(cbind(crim,off))`

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

Vinod, H. D. ‘Generalized Correlation and Kernel Causality with Applications in Development Economics’ in Communications in Statistics -Simulation and Computation, 2015, doi: gffn86

Vinod, H. D. ’New exogeneity tests and causal paths,’ Chapter 2 in ’Handbook of Statistics: Conceptual Econometrics Using R’, Vol.32, co-editors: H. D. Vinod and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2019, pp. 33-64.

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: `https://www.ssrn.com/abstract=2982128`

Vinod, Hrishikesh D., R Package GeneralCorr Functions for Portfolio Choice (November 11, 2021). Available at SSRN: https://ssrn.com/abstract=3961683

Vinod, Hrishikesh D., Stochastic Dominance Without Tears (January 26, 2021). Available at SSRN: https://ssrn.com/abstract=3773309

See Also

See `bootPairs`, `causeSummary` has an older version of this function.

See `someCPairs`

`siPair2Blk`, `causeSummary2`

Examples

```
## Not run:
mtx=as.matrix(mtcars[,1:3])
ctrl=as.matrix(mtcars[,4:5])
causeSum2Blk(mtx,ctrl,nam=colnames(mtx))

## End(Not run)
```

`causeSummary` *Kernel causality summary of evidence for causal paths from three criteria*

Description

While allowing the researcher to keep some variables as controls, or outside the scope of causal path determination (e.g., age or latitude) this function produces detailed causal path information in a 5 column matrix identifying the names of variables, causal path directions, path strengths re-scaled to be in the range [–100, 100], (table reports absolute values of the strength) plus Pearson correlation and its p-value.

Usage

```
causeSummary(
  mtx,
  nam = colnames(mtx),
  ctrl = 0,
  dig = 6,
  wt = c(1.2, 1.1, 1.05, 1),
  sumwt = 4
)
```

Arguments

`mtx`	The data matrix with many columns, y the first column is fixed and then paired with all columns, one by one, and still called x for the purpose of flipping.
`nam`	vector of column names for `mtx`. Default: colnames(mtx)
`ctrl`	data matrix for designated control variable(s) outside causal paths
`dig`	Number of digits for reporting (default `dig=6`).
`wt`	Allows user to choose a vector of four alternative weights for SD1 to SD4.
`sumwt`	Sum of weights can be changed here =4(default).

Details

The algorithm determines causal path directions from the sign of the strength index and strength index values by comparing three aspects of flipped kernel regressions: [x1 on (x2, x3, .. xp)] and its flipped version [x2 on (x1, x3, .. xp)] We compare (i) formal exogeneity test criterion, (ii) absolute residuals, and (iii) R-squares of the flipped regressions implying three criteria Cr1, to Cr3. The criteria are quantified by sophisticated methods using four orders of stochastic dominance, SD1 to SD4. We assume slightly declining weights on causal path signs because known reliability ranking. SD1 is better than SD2, better than SD3, better than SD4. The user can optionally change our weights.

Value

If there are p columns in the input matrix, x1, x2, .., xp, say, and if we keep x1 as a common member of all causal direction pairs (x1, x(1+j)) for (j=1, 2, .., p-1) which can be flipped. That is, either x1 is the cause or x(1+j) is the cause in a chosen pair. The control variables are not flipped. The printed output of this function reports the results for p-1 pairs indicating which variable (by name) causes which another variable (also by name). It also prints a signed summary strength index in the range [-100,100]. A positive sign of the strength index means x1 kernel causes x(1+j), whereas negative strength index means x(1+j) kernel causes x1. The function also prints Pearson correlation and its p-value. In short, function returns a matrix of p-1 rows and 5 columns entitled: “cause", “response", “strength", “corr." and “p-value", respectively with self-explanatory titles. The first two columns have names of variables x1 or x(1+j), depending on which is the cause. The ‘strength’ column reports the absolute value of summary index, now in the range [0,100] providing summary of causal results based on preponderance of evidence from Cr1 to Cr3 from four orders of stochastic dominance, etc. The order of input columns matters. The fourth column ‘corr.’ reports the Pearson correlation coefficient while the fifth column has the p-value for testing the null of zero Pearson coeff. This function calls `silentPairs` allowing for control variables. The output of this function can be sent to ‘xtable’ for a nice Latex table.

Note

The European Crime data has all three criteria correctly suggesting that high crime rate kernel causes the deployment of a large number of police officers. Since Cr1 to Cr3 near unanimously suggest ‘crim’ as the cause of ‘off’, strength index 100 suggests unanimity. In portfolio applications of stochastic dominance one wants higher returns. Here we are comparing two probability distributions of absolute residuals for two flipped models. We choose that flip which has smaller absolute residuals or better fit. `attach(EuroCrime); causeSummary(cbind(crim,off))`

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

Vinod, H. D. ‘Generalized Correlation and Kernel Causality with Applications in Development Economics’ in Communications in Statistics -Simulation and Computation, 2015, doi: gffn86

Vinod, H. D. ’New exogeneity tests and causal paths,’ Chapter 2 in ’Handbook of Statistics: Conceptual Econometrics Using R’, Vol.32, co-editors: H. D. Vinod and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2019, pp. 33-64.

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: https://www.ssrn.com/abstract=2982128

See Also

See bootPairs, causeSummary0 has an older version of this function.

See someCPairs

silentPairs

Examples

```
## Not run:
mtx=as.matrix(mtcars[,1:3])
ctrl=as.matrix(mtcars[,4:5])
 causeSummary(mtx,ctrl,nam=colnames(mtx))

## End(Not run)

options(np.messages=FALSE)
set.seed(234)
z=runif(10,2,11)# z is independently created
x=sample(1:10)+z/10 #x is somewhat indep and affected by z
y=1+2*x+3*z+rnorm(10)
w=runif(10)
x2=x;x2[4]=NA;y2=y;y2[8]=NA;w2=w;w2[4]=NA
causeSummary(mtx=cbind(x2,y2), ctrl=cbind(z,w2))
```

`causeSummary0` *Older Kernel causality summary of evidence for causal paths from three criteria*

Description

Allowing input matrix of control variables, this function produces a 5 column matrix summarizing the results where the estimated signs of stochastic dominance order values, (+1, 0, -1), are weighted by `wt=c(1.2,1.1,1.05,1)` to compute an overall result for all orders of stochastic dominance by a weighted sum for the criteria Cr1 and Cr2 and added to the Cr3 estimate as: (+1, 0, -1). The final range for the unanimity of sign index is [–100, 100].

Usage

```
causeSummary0(
  mtx,
  nam = colnames(mtx),
  ctrl = 0,
  dig = 6,
  wt = c(1.2, 1.1, 1.05, 1),
  sumwt = 4
)
```

Arguments

`mtx`	The data matrix with many columns, y the first column is fixed and then paired with all columns, one by one, and still called x for the purpose of flipping.
`nam`	vector of column names for `mtx`. Default: colnames(mtx)
`ctrl`	data matrix for designated control variable(s) outside causal paths
`dig`	Number of digits for reporting (default `dig`=6).
`wt`	Allows user to choose a vector of four alternative weights for SD1 to SD4.
`sumwt`	Sum of weights can be changed here =4(default).

Details

The reason for slightly declining weights on the signs from SD1 to SD4 is simply that the local mean comparisons implicit in SD1 are known to be more reliable than local variance implicit in SD2, local skewness implicit in SD3 and local kurtosis implicit in SD4. The reason for slightly declining sampling unreliability of higher moments is simply that SD4 involves fourth power of the deviations from the mean and SD3 involves 3rd power, etc. The summary results for all three criteria are reported in one matrix called `out`:

Value

If there are p columns in the input matrix, x1, x2, .., xp, say, and if we keep x1 as a common member of all causal direction pairs (x1, x(1+j)) for (j=1, 2, .., p-1) which can be flipped. That is, either x1 is the cause or x(1+j) is the cause in a chosen pair. The control variables are not flipped. The printed output of this function reports the results for p-1 pairs indicating which variable (by name) causes which other variable (also by name). It also prints strength or signed summary strength index in range [-100,100]. A positive sign of the strength index means x1 kernel causes x(1+j), whereas negative strength index means x(1+j) kernel causes x1. The function also prints Pearson correlation and its p-value. This function also returns a matrix of p-1 rows and 5 columns entitled: “cause", “response", “strength", “corr." and “p-value", respectively with self-explanatory titles. The first two columns have names of variables x1 or x(1+j), depending on which is the cause. The ‘strength’ column has absolute value of summary index in range [0,100] providing summary of causal results based on preponderance of evidence from Cr1 to Cr3 from four orders of stochastic dominance, etc. The order of input columns matters. The fourth column ‘corr.’ reports the Pearson correlation coefficient while the fifth column has the p-value for testing the null of zero Pearson coeff. This function calls `silentPairs0` (the older version) allowing for control variables. The output of this function can be sent to ‘xtable’ for a nice Latex table.

Note

The European Crime data has all three criteria correctly suggesting that high crime rate kernel causes the deployment of a large number of police officers. Since Cr1 to Cr3 near unanimously suggest ‘crim’ as the cause of ‘off’, strength index 100 suggests unanimity. `attach(EuroCrime); causeSummary0(cbind(crim,off))`. Both versions give identical result for this example. Old version of Cr1 using gradients was also motivated by the same Hausman-Wu test statistic.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

Vinod, H. D. ‘Generalized Correlation and Kernel Causality with Applications in Development Economics’ in Communications in Statistics -Simulation and Computation, 2015, https://doi.org/gffn86

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: https://www.ssrn.com/abstract=2982128

See Also

See `bootPairs`

See `someCPairs`

`silentPairs`

Examples

```
## Not run:
mtx=as.matrix(mtcars[,1:3])
ctrl=as.matrix(mtcars[,4:5])
 causeSummary0(mtx,ctrl,nam=colnames(mtx))

## End(Not run)

options(np.messages=FALSE)
set.seed(234)
z=runif(10,2,11)# z is independently created
x=sample(1:10)+z/10 #x is somewhat indep and affected by z
```

```
y=1+2*x+3*z+rnorm(10)
w=runif(10)
x2=x;x2[4]=NA;y2=y;y2[8]=NA;w2=w;w2[4]=NA
causeSummary0(mtx=cbind(x2,y2), ctrl=cbind(z,w2))
```

---

causeSummary2 *Kernel causality summary of evidence for causal paths from three criteria using new exact stochastic dominance. The function develops a unanimity index regarding the which flip (y on xi) or (xi on y) is best. Relevant signs determine the causal direction and unanimity index among three criteria. While allowing the researcher to keep some variables as controls, or outside the scope of causal path determination (e.g., age or latitude) this function produces detailed causal path information in a 5 column matrix identifying the names of variables, causal path directions, path strengths re-scaled to be in the range [–100, 100], (table reports absolute values of the strength) plus Pearson correlation and its p-value. The '2' in the name of the function suggests a second implementation where exact stochastic dominance, decileVote and momentVote are used and where we avoid Anderson's trapezoidal approximation.*

Description

The algorithm determines causal path directions from the sign of the strength index and strength index values by comparing three aspects of flipped kernel regressions: [x1 on f(x2, x3, .. xp)] and its flipped version [x2 on f(x1, x3, .. xp)] We compare (i) formal exogeneity test criterion, (ii) absolute residuals, and (iii) R-squares of the flipped regressions implying three criteria Cr1, to Cr3. The criteria are quantified by newer exact methods using four orders of stochastic dominance, SD1 to SD4. See Vinod (2021) SSRN papers. In portfolio applications of stochastic dominance one wants higher values. Here we are comparing two probability distributions of absolute residuals for two flipped models. We choose that flip which has smaller absolute residuals which will have a better fit.

Usage

```
causeSummary2(mtx, nam = colnames(mtx), ctrl = 0, dig = 6)
```

Arguments

mtx	The data matrix with many columns, y the first column is fixed and then paired with all columns, one by one, and still called x for the purpose of flipping.
nam	vector of column names for `mtx`. Default: colnames(mtx)
ctrl	data matrix for designated control variable(s) outside causal paths
dig	Number of digits for reporting (default `dig`=6).

Value

If there are p columns in the input matrix, x1, x2, .., xp, say, and if we keep x1 as a common member of all causal direction pairs (x1, x(1+j)) for (j=1, 2, .., p-1) which can be flipped. That is, either x1 is the cause or x(1+j) is the cause in a chosen pair. The control variables are not flipped. The printed output of this function reports the results for p-1 pairs indicating which variable (by name) causes which another variable (also by name). It also prints a signed summary strength index in the range [-100,100]. A positive sign of the strength index means x1 kernel causes x(1+j), whereas negative strength index means x(1+j) kernel causes x1. The function also prints Pearson correlation and its p-value. In short, function returns a matrix of p-1 rows and 5 columns entitled: "cause", "response", "strength", "corr." and "p-value", respectively with self-explanatory titles. The first two columns have names of variables x1 or x(1+j), depending on which is the cause. The 'strength' column reports the absolute value of summary index, in the range [0,100] providing summary of causal results based on preponderance of evidence from Cr1 to Cr3 from four orders of stochastic dominance, moments, deciles etc. The order of input columns in mtx matters. The fourth column 'corr.' of 'out' reports the Pearson correlation coefficient while the fifth column has the p-value for testing the null of zero Pearson coeff. This function calls `silentPair2` allowing for control variables. The output of this function can be sent to 'xtable' for a nice Latex table.

Note

The European Crime data has all three criteria correctly suggesting that high crime rate kernel causes the deployment of a large number of police officers. Since Cr1 to Cr3 near unanimously suggest 'crim' as the cause of 'off', strength index 100 suggests unanimity among the criteria. `attach(EuroCrime); causeSummary(cbind(crim,off))`

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: gffn86

Vinod, H. D. 'New exogeneity tests and causal paths,' Chapter 2 in 'Handbook of Statistics: Conceptual Econometrics Using R', Vol.32, co-editors: H. D. Vinod and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2019, pp. 33-64.

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: `https://www.ssrn.com/abstract=2982128`

Vinod, Hrishikesh D., R Package GeneralCorr Functions for Portfolio Choice (November 11, 2021). Available at SSRN: https://ssrn.com/abstract=3961683

Vinod, Hrishikesh D., Stochastic Dominance Without Tears (January 26, 2021). Available at SSRN: https://ssrn.com/abstract=3773309

See Also

See `siPair2Blk` for a block version

See `causeSummary` is subject to trapezoidal approximation.

see silentPair2 called by this function.

Examples

```
## Not run:
mtx=as.matrix(mtcars[,1:3])
ctrl=as.matrix(mtcars[,4:5])
 causeSummary2(mtx,ctrl,nam=colnames(mtx))

## End(Not run)

options(np.messages=FALSE)
set.seed(234)
z=runif(10,2,11)# z is independently created
x=sample(1:10)+z/10 #x is somewhat indep and affected by z
y=1+2*x+3*z+rnorm(10)
w=runif(10)
x2=x;x2[4]=NA;y2=y;y2[8]=NA;w2=w;w2[4]=NA
causeSummary2(mtx=cbind(x2,y2), ctrl=cbind(z,w2))
```

---

causeSummBlk *Block Version 2: Kernel causality summary of causal paths from three criteria*

Description

A block version of causeSummary() chooses new bandwidth for every ten (blksiz=10) observations chosen by the 'np' package injecting flexibility. While allowing the researcher to keep some variables as controls, or outside the scope of causal path determination (e.g., age or latitude), this function produces detailed causal path information. The output table is a 5 column matrix identifying the names of variables, causal path directions, path strengths re-scaled to be in the range [–100, 100], (table reports absolute values of the strength) plus Pearson correlation coefficient and its p-value.

Usage

```
causeSummBlk(
  mtx,
  nam = colnames(mtx),
  blksiz = 10,
  ctrl = 0,
  dig = 6,
  wt = c(1.2, 1.1, 1.05, 1),
  sumwt = 4
)
```

Arguments

mtx	The data matrix with many columns, y the first column is a fixed target, and then it is paired with all other columns, one by one, and still called x for the purpose of flipping.
nam	vector of column names for mtx. Default: colnames(mtx)
blksiz	block size, default=10, if chosen blksiz >n, where n=rows in the matrix then blksiz=n. That is, no blocking is done
ctrl	data matrix for designated control variable(s) outside causal paths
dig	The number of digits for reporting (default dig=6).
wt	Allows user to choose a vector of four alternative weights for SD1 to SD4.
sumwt	Sum of weights can be changed here =4(default).

Details

The algorithm determines causal path directions from the sign of the strength index. The strength index magnitudes are computed by comparing three aspects of flipped kernel regressions: [x1 on (x2, x3, .. xp)] and its flipped version [x2 on (x1, x3, .. xp)]. The cause should be on the right-hand side of regression equation. The properties of regression fit determine which flip is superior. We compare (Cr1) formal exogeneity test criterion, (residuals times RHS regressor, where smaller in absolute value is better) (Cr2) absolute residuals, where smaller in absolute value is better, and (Cr3) R-squares of the flipped regressions implying three criteria Cr1, to Cr3. The criteria are quantified by sophisticated methods using four orders of stochastic dominance, SD1 to SD4. We assume slightly declining weights on the sign observed by Cr1 to Cr3. The user can change default weights.

Value

If there are p columns in the input matrix, x1, x2, .., xp, say, and if we keep x1 as a common member of all causal-direction-pairs (x1, x(1+j)) for (j=1, 2, .., p-1) which can be flipped. That is, either x1 is the cause or x(1+j) is the cause in a chosen pair. The control variables are not flipped. The printed output of this function reports the results for p-1 pairs indicating which variable (by name) causes which other variable (also by name). It also prints a strength, or signed summary strength index forced to be in the range [-100,100] for easy interpretation. A positive sign of the strength index means x1 kernel causes x(1+j), whereas negative strength index means x(1+j) kernel causes x1. The function also prints Pearson correlation and its p-value. This function also returns a matrix of p-1 rows and 5 columns entitled: “cause", “response", “strength", “corr." and “p-value", respectively with self-explanatory titles. The first two columns have names of variables x1 or x(1+j), depending on which is the cause. The ‘strength’ column has the absolute value of a summary index in the range [0,100], providing a summary of causal results based on the preponderance of evidence from Cr1 to Cr3 from four orders of stochastic dominance, etc. The order of input columns matters. The fourth column of the output matrix entitled ‘corr.’ reports the Pearson correlation coefficient, while the fifth column of the output matrix has the p-value for testing the null hypothesis of a zero Pearson coefficient. This function calls siPairsBlk, allowing for control variables. The output of this function can be sent to ‘xtable’ for a nice Latex table.

Note

The European Crime data has all three criteria correctly suggesting that a high crime rate kernel causes the deployment of a large number of police officers. Since Cr1 to Cr3 near-unanimously suggest 'crim' as the cause of 'off', a strength index of 100 suggests unanimity. `attach(EuroCrime); causeSummBlk(cbind(crim,off))`

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: gffn86

Vinod, H. D. 'New exogeneity tests and causal paths,' Chapter 2 in 'Handbook of Statistics: Conceptual Econometrics Using R', Vol.32, co-editors: H. D. Vinod and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2019, pp. 33-64.

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: https://www.ssrn.com/abstract=2982128

See Also

See `bootPairs`, `causeSummary` has an older version of this function.

See `someCPairs`

`siPairsBlk`, `causeSummary`

Examples

```
## Not run:
mtx=as.matrix(mtcars[,1:3])
ctrl=as.matrix(mtcars[,4:5])
 causeSummBlk(mtx,ctrl,nam=colnames(mtx))

## End(Not run)

options(np.messages=FALSE)
set.seed(234)
z=runif(10,2,11)# z is independently created
x=sample(1:10)+z/10 #x is somewhat indep and affected by z
y=1+2*x+3*z+rnorm(10)
w=runif(10)
x2=x;x2[4]=NA;y2=y;y2[8]=NA;w2=w;w2[4]=NA
causeSummBlk(mtx=cbind(x2,y2), ctrl=cbind(z,w2))
```

`causeSumNoP` *No print (NoP) version of causeSummBlk summary causal paths from three criteria*

Description

Allowing input matrix of control variables, this function produces a 5 column matrix summarizing the results where the estimated signs of stochastic dominance order values, (+1, 0, -1), are weighted by `wt=c(1.2,1.1,1.05,1)` to compute an overall result for all orders of stochastic dominance by a weighted sum for the criteria Cr1 and Cr2 and added to the Cr3 estimate as: (+1, 0, -1). The final range for the unanimity of sign index is [–100, 100].

Usage

```
causeSumNoP(
  mtx,
  nam = colnames(mtx),
  blksiz = 10,
  ctrl = 0,
  dig = 6,
  wt = c(1.2, 1.1, 1.05, 1),
  sumwt = 4
)
```

Arguments

`mtx`	The data matrix with many columns, y the first column is a fixed target and then it is paired with all other columns, one by one, and still called x for the purpose of flipping.
`nam`	vector of column names for `mtx`. Default: colnames(mtx)
`blksiz`	block size, default=10, if chosen blksiz >n, where n=rows in matrix then blksiz=n. That is, no blocking is done
`ctrl`	data matrix for designated control variable(s) outside causal paths
`dig`	Number of digits for reporting (default `dig`=6).
`wt`	Allows user to choose a vector of four alternative weights for SD1 to SD4.
`sumwt`	Sum of weights can be changed here =4(default).

Details

The reason for slightly declining weights on the signs from SD1 to SD4 is simply that the higher order stochastic dominance numbers are less reliable. The summary results for all three criteria are reported in one matrix called `out` but not printed:

Value

If there are p columns in the input matrix, x1, x2, .., xp, say, and if we keep x1 as a common member of all causal-direction-pairs (x1, x(1+j)) for (j=1, 2, .., p-1) which can be flipped. That is, either x1 is the cause or x(1+j) is the cause in a chosen pair. The control variables are not flipped. The printed output of this function reports the results for p-1 pairs indicating which variable (by name) causes which other variable (also by name). It also prints strength or signed summary strength index in range [-100,100]. A positive sign of the strength index means x1 kernel causes x(1+j), whereas negative strength index means x(1+j) kernel causes x1. The function also prints Pearson correlation and its p-value. This function also returns a matrix of p-1 rows and 5 columns entitled: “cause", “response", “strength", “corr." and “p-value", respectively with self-explanatory titles. The first two columns have names of variables x1 or x(1+j), depending on which is the cause. The ‘strength’ column has absolute value of summary index in range [0,100] providing summary of causal results based on preponderance of evidence from Cr1 to Cr3 from four orders of stochastic dominance, etc. The order of input columns matters. The fourth column ‘corr.’ reports the Pearson correlation coefficient while the fifth column has the p-value for testing the null of zero Pearson coeff. This function calls `siPairsBlk` allowing for control variables. The output of this function can be sent to ‘xtable’ for a nice Latex table.

Note

The European Crime data has all three criteria correctly suggesting that high crime rate kernel causes the deployment of a large number of police officers. Since Cr1 to Cr3 near unanimously suggest ‘crim’ as the cause of ‘off’, strength index 100 suggests unanimity. `attach(EuroCrime); causeSummary(cbind(crim,off))`

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

Vinod, H. D. ‘Generalized Correlation and Kernel Causality with Applications in Development Economics’ in Communications in Statistics -Simulation and Computation, 2015, doi: gffn86

Vinod, H. D. ’New exogeneity tests and causal paths,’ Chapter 2 in ’Handbook of Statistics: Conceptual Econometrics Using R’, Vol.32, co-editors: H. D. Vinod and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2019, pp. 33-64.

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: `https://www.ssrn.com/abstract=2982128`

See Also

See `bootPairs`, `causeSummary0` has an older version of this function.

See `causeAllPair`

`siPairsBlk`, `causeSummary`

Examples

```
## Not run:
mtx=data.frame(mtcars[,1:3])
ctrl=data.frame(mtcars[,4:5])
 causeSumNoP(mtx=mtx,ctrl=ctrl)

## End(Not run)

options(np.messages=FALSE)
set.seed(234)
z=runif(10,2,11)# z is independently created
x=sample(1:10)+z/10 #x is somewhat indep and affected by z
y=1+2*x+3*z+rnorm(10)
w=runif(10)
x2=x;x2[4]=NA;y2=y;y2[8]=NA;w2=w;w2[4]=NA
causeSumNoP(mtx=cbind(x2,y2), ctrl=cbind(z,w2))
```

`cofactor` *Compute cofactor of a matrix based on row r and column c.*

Description

Compute cofactor of a matrix based on row r and column c.

Usage

```
cofactor(x, r, c)
```

Arguments

`x`	matrix whose cofactor is desired to be computed
`r`	row number
`c`	column number

Value

cofactor of x, w.r.t. row r and column c.

Note

needs the function ‘minor” in memory. attaches sign (-1)^(r+c) to the minor.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

See Also

`minor(x,r,c)`

Examples

```
## The function is currently defined as
function (x, r, c)
{
    out = minor(x, r, c) * ((-1)^(r + c))
    return(out)
  }
```

---

compPortfo — *Compares two vectors (portfolios) using momentVote, DecileVote and exactSdMtx functions.*

Description

Given two vectors of portfolio returns this function summarizes their ranks based on moments, deciles and exact measures of stochastic dominance. as explained in Vinod (2021). This algorithm has model selection applications.

Usage

```
compPortfo(xa, xb)
```

Arguments

xa	Data on returns for portfolio A in the form of a T by 1 vector
xb	Data on returns for portfolio B in the form of a T by 1 vector

Value

Returns three numbers which represent signs based differences in ranks (rank=1 for most desirable) measured by [rank(xa)-rank(xb)] using momentVote, decileVote, and exactSdMtx which are weighted averages of four moments, nine deciles and exact measures of stochastic dominance (from ECDFs of four orders, SD1 to SD4) respectively.

Note

There are model-selection applications where two models A and B are compared and one wants to choose the model smaller absolute value of residuals. This function when applied for model-selection will have he inputs xa and xb as absolute residuals. We can compare the entire probability distributions of absolute residuals by moments, deciles or SD1 to SD4. Of course, care must be taken to choose xa or xb depending on which model has smaller absolute residuals. This choice is the exact opposite of portfolio choice application where larger return is more desirable. `silentPair2()` and `siPair2Blk` call this function for model selection application.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D.", "Hands-On Intermediate Econometrics Using R" (2008) World Scientific Publishers: Hackensack, NJ. (Chapter 4) https://www.worldscientific.com/worldscibooks/10.1142/6895

Vinod, Hrishikesh D., R Package GeneralCorr Functions for Portfolio Choice (November 11, 2021). Available at SSRN: https://ssrn.com/abstract=3961683

See Also

exactSdMtx

momentVote

decileVote

Examples

```
set.seed(30)
xa=sample(20:30)#generally lower returns
xb=sample(32:40)# higher returns in xb
gp = compPortfo(xa, xb)#all Av(sdi) positive means xb dominates
##output (1,1,1) means xb dominates xa. xb are larger by consruction
```

---

comp_portfo2 — *Compares two vectors (portfolios) using stochastic dominance of orders 1 to 4.*

Description

Given two vectors of portfolio returns this function calls the internal function wtdpapb to report the simple means of four sophisticated measures of stochastic dominance. as explained in Vinod (2008).

Usage

```
comp_portfo2(xa, xb)
```

Arguments

xa	Data on returns for portfolio A in the form of a T by 1 vector
xb	Data on returns for portfolio B in the form of a T by 1 vector

Value

Returns four numbers which are averages of four sophisticated measures of stochastic dominance measurements called SD1 to SD4.

Note

It is possible to modify this function to report the median or standard deviation or any other descriptive statistic by changing the line in the code 'oumean = apply(outb,2,mean)' toward the end of this function. A trimmed mean may be of interest when outliers are suspected.

require(np)

Make sure that functions wtdpapb, bigfp, stochdom2 are in the memory. and options(np.messages=FALSE)

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D.", "Hands-On Intermediate Econometrics Using R" (2008) World Scientific Publishers: Hackensack, NJ. (Chapter 4) https://www.worldscientific.com/worldscibooks/10.1142/6895

See Also

stochdom2

Examples

```
set.seed(30)
xa=sample(20:30)#generally lower returns
xb=sample(32:40)# higher returns in xb
gp = comp_portfo2(xa, xb)#all Av(sdi) positive means xb dominates
##positive SD1 to SD4 means xb dominates xa as it should
```

---

da *internal da*

Description

intended for internal use only

Usage

```
da
```

`da2Lag` *internal da2Lag*

Description

intended for internal use

Usage

```
data(da2Lag)
```

Format

The format is: int 4

`decileVote` *Function compares nine deciles of stock return distributions.*

Description

The first step computes a minimum reference return and nine deciles. The input x must be a matrix having p columns (with a name for each column) and n rows as in the data. If data are missing for some columns, insert NA's. Thus x has p column of the data matrix ready for comparison and ranking. For example, x has a matrix of stock returns. The output matrix produced by this function also has p columns for each column (i.e., for each stock being compared). The output matrix has nineteen rows. The top nine rows have the magnitudes of deciles. Rows 10 to 18 have respective ranks of the decile magnitudes. The next (19-th) row of the output reports a weighted sum of ranks. Ranking always gives the smallest number 1 to the most desirable outcome. We suggest that a higher portfolio weight be given to the column having smallest rank value (along the 19th line). The 20-th row further ranks the weighted sums of ranks in row 19. Investor should choose the stock (column) representing the smallest rank value along the last (20th) row of the 'out' matrix.

Usage

```
decileVote(mtx, howManySd = 0.1)
```

Arguments

`mtx`	(n X p) matrix of data. For example, returns on p stocks n months
`howManySd`	used to define 'fixmin'= imaginary lowest return defined by going howManySd=default=0.1 maximum of standard deviations of all stocks below the minimum return for all stocks in the data

Value

out is a matrix with p columns (same as in the input matrix) and twenty rows. Top nine rows have 9 deciles, next nine rows have their ranks. The 19-th row of 'out' has a weighted sum of 9 ranks. All columns refer to one stock. The weighted sum for each stock is then ranked. A portfolio manager is assumed to prefer higher return represented by high decile values represented by the column with the largest weighted sum. can give largest weight to the column with the smallest bottom line. The bottom line (20-th) labeled "choice" of the 'out' matrix is defined so that choice =1 suggests the stock deserving the highest weight in the portfolio. The portfolio manager will generally give the lowest weight (=0?) to the stock representing column having number p as the choice number. The manager may want to sell this stock. Another output of the 'decileVote' function is 'fixmin' representing the smallest possible return of all the stocks in the input 'mtx' of returns. It is useful as a reference stock. We compute stochastic dominance numbers for each stock with this imaginary stock yielding fixmin return for all time periods.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

Examples

```
x1=c(1,4,7,2,6)
x2=c(3,4,8,4,7)
decileVote(cbind(x1,x2))
```

---

depMeas — *depMeas Signed measure of nonlinear nonparametric dependence between two vectors.*

Description

An infant may depend on the mother for survival, but not vice versa. Dependence relations need not be symmetric, yet correlation coefficients are symmetric. One way to measure the extent of dependence is to find the max of the absolute values of the two asymmetric correlations using Vinod (2015) definition of generalized (asymmetric) correlation coefficients. It requires a kernel regression of x on y obtained by using the 'np' package and its flipped version (regress y on x). We use a block version of 'gmcmtx0' called 'gmcmtxBlk' to admit several bandwidths for every ten observations (say) blksiz=10 seems to be a good choice.

Usage

```
depMeas(x, y, blksiz = length(x))
```

Arguments

`x`	Vector of data on the first variable
`y`	Vector of data on the second variable
`blksiz`	block size, default blksiz =n, where n=rows in the matrix or no blocking is done

Value

A measure of dependence having the same sign as Pearson correlation. Its magnitude equals the larger of the two generalized correlation coefficients.

Note

This function needs the gmcmtxBlk function, which in turn needs the np package.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, https://doi.org/gffn86

Vinod, H. D. 'Matrix Algebra Topics in Statistics and Economics Using R', Chapter 4 in Handbook of Statistics: Computational Statistics with R, Vol.32, co-editors: M. B. Rao and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2014, pp. 143-176.

Vinod, H. D. (2021) 'Generalized, Partial and Canonical Correlation Coefficients' Computational Economics, 59(1), 1–28.

See Also

See Also gmcmtx0 and gmcmtxBlk

Examples

```
library(generalCorr)
options(np.messages = FALSE)
x=1:20;y=sin(x)
depMeas(x,y,blksiz=20)
```

---

dif4 *order 4 differencing of a time series vector*

Description

This is for momentum traders who focus on growth, acceleration, its gorwth and further acceleration. The diff function of R seems to do recycling of available numbers, not wanted for our purposes.

Usage

```
dif4(x)
```

Arguments

`x`	(n X 1) vector of time series (market returns) with n items each

Value

ou2 matrix having five columns, first for x, the next four columns have diff(x), diff-squared(x), diff-cubed(x) and diff-fourth(x)

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

Examples

```
x=c(2,8,3,5,1,8,19,22,23)
dif4(x)
```

`dif4mtx` *order four differencing of a matrix of time series*

Description

This is for momentum traders who focus on growth, acceleration, its growth and further acceleration. The diff function of R seems to do recycling of available numbers, not wanted for our purposes. Hence, this function is needed in portfolio studies based on time series.

Usage

```
dif4mtx(mtx)
```

Arguments

`mtx`	(n X p) matrix of p time series (market returns) with n items each

Value

out matrix having 12 rows, (data, D1 to D4 and ranks of D1 to D4 The column names of out are those of input matrix mtx.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

Examples

```
x=c(2,8,3,5,1,8,19,22,23)
y=c(3,11,2,6,7,9,20,25,21)
dif4mtx(cbind(x,y))
```

`diff.e0` *Internal diff.e0*

Description

Internal diff.e0

Usage

```
data(diff.e0)
```

`dig` *Internal dig*

Description

Intended for internal use

Usage

```
data(dig)
```

Format

The format digs: int 78

`e0` *internal e0*

Description

intended for internal use only

Usage

```
e0
```

`EuroCrime` *European Crime Data*

Description

This data set refers to crime in European countries during 2008. The sources are World Bank and Eurostat. The crime statistics refers to homicides. It avoids possible reporting bias from the presence of police officers, because homicide reporting in most countries is standardized. Typical usage is: `data(EuroCrime);attach(EuroCrime)`. The secondary source 'quandl.com' was used for collecting these data.

Details

The variables included in the dataset are:

- `Country` Name of the European country
- `crim` Per capita crime rate
- `off` Per capita deployment of police officers

`exactSdMtx` *Exact stochastic dominance computation from areas above ECDF pillars.*

Description

ECDF=empirical cumulative distribution function. The exact computation needs a common reference minimum (refmin) return for computation of dominance orders SD1 to SD4. This function inputs 'mtx' (n X p) matrix data (e.g., monthly returns on p stocks). Its output has four matrices SD1 to SD4, each with dimension (n X p). They measure exact dominance areas between empirical CDF for each column to the ECDF of (x.ref) an artificial stock with minimal return in all time periods. A fifth output matrix 'out' has 4 rows and p columns containing column sums of SD1 to SD4. The 'out' matrix produced by this function is input to `summaryRank` function to indicate the choice of the best column in 'mtx' for investment based on ranks.

Usage

```
exactSdMtx(mtx, howManySd = 0.1)
```

Arguments

`mtx`	(n X p) matrix of data. For example, returns on p stocks n months
`howManySd`	used to define (x.ref)= lowest return number. If the grand minimum of all returns in 'mtx' is dented GrMin, then howManySd equals the number of max(sd) (maximum standard deviation for data columns) below the GrMin used to define (x.ref). Thus, (x.ref)=GrMin-howManySd*max(sd). default howManySd=0.1

Value

five matrices. SD1 to SD4 contain four orders of stochastic dominance areas using the ECDF pillars and a common (x.ref). The fifth "out" matrix is another output having 4 rows for SD1 to SD4 and p columns (p=No. of columns in data matrix mtx) having a summary of ranks using all four, SD1 to SD4.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

Examples

```
x1=c(2,5,6,9,13,18,21)
x2=c(3,6,9,12,14,19,27)
st1=exactSdMtx(cbind(x1,x2))
```

GcRsqX12 *Generalized Granger-Causality. If dif>0, x2 Granger-causes x1.*

Description

The usual Granger-causality assumes linear regressions. This function allows nonlinear nonparametric kernel regressions using a local linear (ll) option. Granger-causality (Gc) is generalized using nonlinear kernel regressions using local linear (ll) option. This functionn computes two R^2 values. (i) R12 or kernel regression R^2 of x1t on its own lags and x2t and its lags. (ii) R21 or kernel regression R^2 of x2t on its own lags and x1t and its lags. (iii) dif=R12-R21, the difference between the two R^2 values. If dif>0 then x2 Granger-causes x1.

Usage

```
GcRsqX12(x1, x2, px1 = 4, px2 = 4, pwanted = 4, ctrl = 0)
```

Arguments

`x1`	The data vector x1
`x2`	The data vector x2
`px1`	The number of lags of x1 in the data default px1=4
`px2`	The number of lags of x2 in the data, default px2=4
`pwanted`	number of lags of both x2 and x1 wanted for Granger causal analysis, default =4
`ctrl`	data matrix for designated control variable(s) outside causal paths default=0 means no control variables are present

Details

Calls GcRsqYX for R-square from kernel regression (local linear version) R^2[x1=f(x1,x2)] choosing GcRsqYX(y=x1, x=x2). It predicts x1 from both x1 and x2 using all information till time (t-1). It also calls GcRsqYX again after flipping x1 and x2. It returns RsqX1onX2, RsqX2onX1 and the difference dif=(RsqX1onX2-RsqX2onX1) If (dif>0) the regression x1=f(x1,x2) is better than the flipped version implying that x1 is more predictable or x2 Granger-causes x1, x2 –> x1, rather than vice versa. The kernel regressions use regtype="ll" for local linear, bwmethod="cv.aic" for AIC-based bandwidth selection.

Value

This function returns 3 numbers: RsqX1onX2, RsqX2onX1 and dif

returns a list of 3 numbers. RsqX1onX2=(Rsquare of kernel regression of X1 on lags of X1 and X2 and its lags), RsqX2onX1= (Rsquare of kernel regression of x2 on own lags of X2 and X1), and the difference between the two Rquares (first minus second) called 'dif.' If dif>0 then x2 Granger-causes x1

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, https://doi.org/gffn86

Vinod, H. D. 'New exogeneity tests and causal paths,' Chapter 2 in 'Handbook of Statistics: Conceptual Econometrics Using R', Vol.32, co-editors: H. D. Vinod and C.R. Rao. New York: North-Holland, Elsevier Science Publishers, 2019, pp. 33-64.

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: https://www.ssrn.com/abstract=2982128

Zheng, S., Shi, N.-Z., Zhang, Z., 2012. Generalized measures of correlation for asymmetry, nonlinearity, and beyond. Journal of the American Statistical Association 107, 1239-1252. -at-note internal routine

See Also

bootGcRsq, causeSummary, GcRsqYX.

Examples

```
## Not run:
library(Ecdat);options(np.messages=FALSE);attach(data.frame(MoneyUS))
GcRsqX12(y,m)

## End(Not run)
```

GcRsqX12c *Generalized Granger-Causality. If dif>0, x2 Granger-causes x1.*

Description

The usual Granger-causality assumes linear regressions. This allows nonlinear nonparametric kernel regressions using a local constat (lc) option. Calls GcRsqYXc for R square from kernel regression. R^2[x1=f(x1,x2)] choosing GcRsqYXc(y=x1, x=x2). The name 'c' in the function refers to local constant option of kernel regressions.' It predicts x1 from both x1 and x2 using all information till time (t-1). It also calls GcRsqYXc again after flipping x1 and x2. It returns RsqX1onX2, RsqX2onX1 and the difference dif=(RsqX1onX2-RsqX2onX1) If (dif>0) the regression x1=f(x1,x2) is better than the flipped version implying that x1 is more predictable or x2 Granger-causes x1 x2 –> x1, rather than vice versa. The kernel regressions use regtype="lc" for local constant, bwmethod="cv.ls" for least squares-based bandwidth selection.

Usage

```
GcRsqX12c(x1, x2, px1 = 4, px2 = 4, pwanted = 4, ctrl = 0)
```

Arguments

x1	The data vector x1
x2	The data vector x2
px1	number of lags of x1 in the data default px1=4
px2	number of lags of x2 in the data, default px2=4
pwanted	number of lags of both x2 and x1 wanted for Granger causal analysis, default =4
ctrl	data matrix for designated control variable(s) outside causal paths default=0 means no control variables are present

Value

This function returns 3 numbers: RsqX1onX2, RsqX2onX1 and dif

returns a list of 3 numbers. RsqX1onX2=(Rsquare of kernel regression of X1 on X1 and X2), RsqX2onX1= (Rsquare of kernel regression of x2 on X2 and X1), and the difference between the two Rquares called dif

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

Vinod, H. D. ‘Generalized Correlation and Kernel Causality with Applications in Development Economics’ in Communications in Statistics -Simulation and Computation, 2015, https://doi.org/gffn86

Vinod, H. D. ’New exogeneity tests and causal paths,’ Chapter 2 in ’Handbook of Statistics: Conceptual Econometrics Using R’, Vol.32, co-editors: H. D. Vinod and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2019, pp. 33-64.

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: https://www.ssrn.com/abstract=2982128

Zheng, S., Shi, N.-Z., Zhang, Z., 2012. Generalized measures of correlation for asymmetry, nonlinearity, and beyond. Journal of the American Statistical Association 107, 1239-1252. -at-note internal routine

See Also

causeSummary

GcRsqYXc

Examples

```
## Not run:
library(Ecdat);options(np.messages=FALSE);attach(data.frame(MoneyUS))
GcRsqX12c(y,m)

## End(Not run)
```

GcRsqYX — *Nonlinear Granger causality between two time series workhorse function.*

Description

Function input is y=LHS=First time series and x=RHS=Second time series. Kernel regression np package options regtype="ll" for local linear, and bwmethod="cv.aic" for AIC-based bandwidth selection are fixed. Denote Rsq=Rsquare=R^2 in nonlinear kernel regression. GcRsqYX(.) computes the following two R^2 values. out[1]=Rsqyyx = R^2 when we regress y on own lags of y and x. out[2]=Rsqyy = R^2 when we regress y on own lags of y alone.

Usage

```
GcRsqYX(y, x, px = 4, py = 4, pwanted = 4, ctrl = 0)
```

Arguments

y	The data vector y for the Left side or dependent or first variable
x	The data vector x for the right side or explanatory or second variable
px	number of lags of x in the data
py	number of lags of y in the data. px=4 for quarterly data
pwanted	number of lags of both x and y wanted for Granger causal analysis
ctrl	data matrix for designated control variable(s) outside causal paths default=0 means no control variables are present

Value

This function returns a set of 2 numbers measuring nonlinear Granger-causality for time series. out[1]=Rsqyyx, out[2]=Rsqyy.

Note

If data are annual or if no quarterly-type structure is present, use this function with pwanted=px=py. For example, the egg or chicken data from lmtest package.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, https://doi.org/gffn86

Vinod, H. D. 'New exogeneity tests and causal paths,' Chapter 2 in 'Handbook of Statistics: Conceptual Econometrics Using R', Vol.32, co-editors: H. D. Vinod and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2019, pp. 33-64.

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: https://www.ssrn.com/abstract=2982128

Zheng, S., Shi, N.-Z., Zhang, Z., 2012. Generalized measures of correlation for asymmetry, nonlinearity, and beyond. Journal of the American Statistical Association 107, 1239-1252.

See Also

GcRsqX12, kern2, kern2ctrl.

Examples

```
## Not run:
library(Ecdat);options(np.messages=FALSE);attach(data.frame(MoneyUS))
GcRsqYX(y,m)
```

```
## End(Not run)
```

GcRsqYXc — *Nonlinear Granger causality between two time series workhorse function.(local constant version)*

Description

Function input is y=LHS=First time series and x=RHS=Second time series. Kernel regression np package options regtype="lc" for local constant, and bwmethod="cv.ls" for least squares-based bandwidth selection are fixed. Denote Rsq=Rsquare=R^2 in nonlinear kernel regression. GcRsqYXc(.) computes the following two R^2 values. out[1]=Rsqyyx = R^2 when we regress y on own lags of y and x. out[2]=Rsqyy = R^2 when we regress y on own lags of y alone.

Usage

```
GcRsqYXc(y, x, px = 4, py = 4, pwanted = 4, ctrl = 0)
```

Arguments

`y`	The data vector y for the Left side or dependent or first variable
`x`	The data vector x for the right side or explanatory or second variable
`px`	number of lags of x in the data
`py`	number of lags of y in the data. px=4 for quarterly data
`pwanted`	number of lags of both x and y wanted for Granger causal analysis
`ctrl`	data matrix for designated control variable(s) outside causal paths default=0 means no control variables are present

Value

This function returns a set of 2 numbers measuring nonlinear Granger-causality for time series. out[1]=Rsqyyx, out[2]=Rsqyy.

Note

If data are annual or if no quarterly-type structure is present, use this function with pwanted=px=py. For example, the egg or chicken data from lmtest package, Thurman W.N. and Fisher M.E. (1988)

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, https://doi.org/gffn86

Vinod, H. D. 'New exogeneity tests and causal paths,' Chapter 2 in 'Handbook of Statistics: Conceptual Econometrics Using R', Vol.32, co-editors: H. D. Vinod and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2019, pp. 33-64.

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: https://www.ssrn.com/abstract=2982128

Zheng, S., Shi, N.-Z., Zhang, Z., 2012. Generalized measures of correlation for asymmetry, nonlinearity, and beyond. Journal of the American Statistical Association 107, 1239-1252.

See Also

`GcRsqX12c`

`kern_ctrl`

Examples

```
## Not run:
library(Ecdat);options(np.messages=FALSE);attach(data.frame(MoneyUS))
GcRsqYXc(y,m)

## End(Not run)
```

`generalCorrInfo` *generalCorr package description:*

Description

This package provides convenient software tools for causal path determinations using Vinod (2014, 2015, 2018, 2021) and is explained in many package vignettes. `causeSummary(mtx)`, `causeSummary2(mtx)`,`causeSum2Blk(` `causeSummBlk` are various versions reporting pair-wise causal path directions and causal strengths. We fit a kernel regression of X1 on (X2, X3,..Xk) and another flipped regression of X2 on (X1, x3, ..Xk). We compare the two fits using three sophisticated criteria called Cr1 to Cr3. We rescale the weighted sum of the quantified three criteria to the [-100, 100] range. The sign of the weighted sum gives the direction of the causal path, and the magnitude of the weighted sum gives the strength of the causal path. A matrix of non-symmetric generalized correlations r*(x|y) is reported by the functions `rstar()` and `gmcmtx0()`. `sudoCoefParcor()` computes pseudo kernel regression coefficients based on generalized partial correlation coefficients (GPCC) `depMeas()` a measure of nonlinear nonparametric dependence between two vectors. `parcorVec()` has generalized partial correlation coefficients, Vinod (2021) `parcorVecH()` has a hybrid version of the above (using HG-PCC). The usual partial correlations r(x,y|z) for regression of y on (x, z) measure the effect of y on x after removing the effect of z, where z can have several variables. Vinod (2021) suggests new generalized partial correlation coefficients (GPCC) using kernel regressions, r*(x,y|z).

Details

The criterion Cr1 uses observable values of standard exogeneity test criterion, namely, (kernel regression residual) times (regressor values) Cr2 computes absolute values kernel regression residuals. The quantification of Cr1 and Cr2 further uses four orders of stochastic dominance measures. Cr3 compares the R-square of the two fits. The package provides additional tools for matrix algebra, such as `cofactor()`, for outlier detection `get0outlier()`, for numerical integration by the trapezoidal rule, stochastic dominance `stochdom2()` and `comp_portfo2()`, etc. The package has a function `pcause()` for bootstrap-based statistical inference and another one for a heuristic t-test called `heurist()`. Pairwise deletion of missing data is done in `napair()`, while triplet-wise deletion is in `naTriplet()` intended for use when control variable(s) are also present. If one has panel data, functions `PanelLag()` and `Panel2Lag()` are relevant. `pillar3D` provides 3-dimensional plots of data that look more like surfaces, than usual plots with vertical pins.

Recent 2020 additions include `canonRho()` for generalized canonical correlations, and many functions for Granger causality between lagged time series including `GcRsqX12()`, `bootGcRsq()` and `GcRsqYXc()`.

Recent 2021 additions include several functions for portfolio choice. `sudoCoefParcor()` for pseudo regression coefficients for kernel regressions. `decileVote()`, `momentVote()`, `exactSdMtx()` for exact computation of stochastic dominance from ECDF areas. The newer stochastic dominance tools are used in `causeSummary2(mtx)`,`causeSum2Blk(mtx)` `dif4mtx()` computes growth, change in growth etc. up-to order 4 differencing of time series.

Note

Six vignettes provided with this package at CRAN describe the theory and usage of the package with examples. Read them using the command: `vignette("generalCorr-vignette")` to read the first vignette. vignettes 2 to 6 can be read by including the vignette number. For example, `vignette("generalCorr-vignette6")` to read the sixth vignette.

References

Vinod, H. D.'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: gffn86

Vinod, H. D. 'Matrix Algebra Topics in Statistics and Economics Using R', Chapter 4 in 'Handbook of Statistics: Computational Statistics with R', Vol.32, co-editors: M. B. Rao and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2014, pp. 143-176.

Zheng, S., Shi, N.-Z., and Zhang, Z. (2012). 'Generalized measures of correlation for asymmetry, nonlinearity, and beyond,' Journal of the American Statistical Association, vol. 107, pp. 1239-1252.

Vinod, H. D. (2021) 'Generalized, Partial and Canonical Correlation Coefficients' Computational Economics, 59(1), 1–28.

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: https://www.ssrn.com/abstract=2982128

Vinod, H. D. 'New exogeneity tests and causal paths,' Chapter 2 in 'Handbook of Statistics: Conceptual Econometrics Using R', Vol.32, co-editors: H. D. Vinod and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2019, pp. 33-64.

`get0outliers` *Function to compute outliers and their count using Tukey method using 1.5 times interquartile range (IQR) to define boundaries.*

Description

Function to compute outliers and their count using Tukey method using 1.5 times interquartile range (IQR) to define boundaries.

Usage

```
get0outliers(x, verbo = TRUE, mult = 1.5)
```

Arguments

`x`	vector of data.
`verbo`	set to TRUE(default) assuming printed details are desired.
`mult`	=1.5(default), the number of times IQR is used in defining outlier boundaries.

Value

`below`	which items are lower than the lower limit
`above`	which items are larger than the upper limit
`low.lim`	the lower boundary for outlier detection
`up.lim`	the upper boundary for outlier detection
`nUP`	count of number of data points above upper boundary
`nLO`	count of number of data points below lower boundary

Note

The function removes the missing data before checking for outliers.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

Examples

```
set.seed(101);x=sample(1:100)[1:15];x[16]=150;x[17]=NA
get0outliers(x)#correctly identifies outlier=150
```

`getSeq` *Two sequences: starting+ending values from n and blocksize (internal use)*

Description

This is an auxiliary function for gmcmtxBlk. It gives sequences of starting and ending values

Usage

```
getSeq(n, blksiz)
```

Arguments

`n`	length of the range
`blksiz`	blocksize

Value

two vectors sqLO and sqUP

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

See Also

gmcmtxBlk

Examples

```
getSeq(n=99, blksiz=10)
```

`gmc0` *internal gmc0*

Description

intended for internal use only

Usage

```
gmc0
```

gmc1 *internal gmc1*

Description

intended for internal use only

Usage

```
gmc1
```

gmcmtx0 *Matrix R* of generalized correlation coefficients captures nonlinearities.*

Description

This function checks for missing data for each pair individually. It then uses the `kern` function to kernel regress x on y, and conversely y on x. It needs the R package 'np' which reports R-squares of each regression. This function reports their square roots after assigning them the observed sign of the Pearson correlation coefficient. Its advantages are: (i) It is asymmetric yielding causal direction information by relaxing the assumption of linearity implicit in usual correlation coefficients. (ii) The r* correlation coefficients are generally larger upon admitting arbitrary nonlinearities. (iii) max(|R*ij|, |R*ji|) measures (nonlinear) dependence. For example, let x=1:20 and y=sin(x). This y has a perfect (100 percent) nonlinear dependence on x and yet Pearson correlation coefficient r(xy) -0.0948372 is near zero and usual confidence interval (-0.516, 0.363) includes zero, implying that it is not different from zero. This shows a miserable failure of traditional r(x,y) to measure dependence when nonlinearities are present. `gmcmtx0(x,y)` will correctly reveal perfect (nonlinear) dependence with generalized correlation coefficient =1.

Usage

```
gmcmtx0(mym, nam = colnames(mym))
```

Arguments

`mym`	A matrix of data on variables in columns
`nam`	Column names of the variables in the data matrix

Value

A non-symmetric R* matrix of generalized correlation coefficients

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D.'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, https://doi.org/gffn86

Vinod, H. D. 'Matrix Algebra Topics in Statistics and Economics Using R', Chapter 4 in 'Handbook of Statistics: Computational Statistics with R', Vol.32, co-editors: M. B. Rao and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2014, pp. 143-176.

Vinod, H. D. 'New exogeneity tests and causal paths,' Chapter 2 in 'Handbook of Statistics: Conceptual Econometrics Using R', Vol.32, co-editors: H. D. Vinod and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2019, pp. 33-64.

Zheng, S., Shi, N.-Z., and Zhang, Z. (2012). 'Generalized measures of correlation for asymmetry, nonlinearity, and beyond,' Journal of the American Statistical Association, vol. 107, pp. 1239-1252.

See Also

See Also as gmcmtxBlk for a more general version using blocking allowing several bandwidths.

Examples

```
gmcmtx0(mtcars[,1:3])

## Not run:
set.seed(34);x=matrix(sample(1:600)[1:99],ncol=3)
colnames(x)=c('V1', 'v2', 'V3')
gmcmtx0(x)
## End(Not run)
```

---

gmcmtxBlk *Matrix R* of generalized correlation coefficients captures nonlinearities using blocks.*

Description

The algorithm uses two auxiliary functions, `getSeq` and `NLhat`. The latter uses the `kern` function to kernel regress x on y, and conversely y on x. It needs the package 'np,' which reports residuals and allows one to compute fitted values (xhat, yhat). Unlike `gmcmtx0`, this function considers blocks of blksiz=10 (default) pairs of data points separately with distinct bandwidths for each block, usually creating superior local fits.

Usage

```
gmcmtxBlk(mym, nam = colnames(mym), blksiz = 10)
```

Arguments

`mym`	A matrix of data on selected variables arranged in columns
`nam`	Column names of the variables in the data matrix
`blksiz`	block size, default=10, if chosen blksiz >n, where n=rows in matrix then blksiz=n. That is, no blocking is done

Details

This function does pairwise checks of missing data for all pairs. Assume that there are n rows in the input matrix 'mym' with some missing rows. If the columns of mym are denoted (X1, X2, ...Xp), we are considering all pairs (Xi, Xj), treated as (x, y), with 'nv' number of valid (non-missing) rows Note that each x and y is an (nv by 1) vector. This function further splits these (x, y) vectors into as many subgroups or blocks as are needed for the nv paired valid data points for the chosen block length (blksiz)

Next, the algorithm strings together various blocks of fitted value vectors (xhat, yhat) also of dimension nv by 1. Now for each pair of Xi Xj (column Xj= cause, row Xi=response, treated as x and y), the algorithm computes R*ij the simple Pearson correlation coefficient between (x, xhat) and as R*ji the correlation coeff. between (y, yhat). Next, it assigns |R*ij| and |R*ji| the observed sign of the Pearson correlation coefficient between x and y.

Its advantages discussed in Vinod (2015, 2019) are: (i) It is asymmetric yielding causal direction information, by relaxing the assumption of linearity implicit in usual correlation coefficients. (ii) The R* correlation coefficients are generally larger upon admitting arbitrary nonlinearities. (iii) max(|R*ij|, |R*ji|) measures (nonlinear) dependence. For example, let x=1:20 and y=sin(x). This y has a perfect (100 percent) nonlinear dependence on x and yet Pearson correlation coefficient r(x y)= -0.0948372 is near zero, and its 95% confidence interval (-0.516, 0.363) includes zero, implying that the population r(x,y) is not significantly different from zero. This example highlights a serious failure of the traditional r(x,y) in measuring dependence between x and y when nonlinearities are present. `gmcmtx0` without blocking does work if x=1:n, and y=f(x)=sin(x) is used with n<20. But for larger n, the fixed bandwidth used by the `kern` function becomes a problem. The block version has additional bandwidths for each block, and hence it correctly quantifies the presence of high dependence even when x=1:n, and y=f(x) are defined for large n and complicated nonlinear functional forms for f(x).

Value

A non-symmetric R* matrix of generalized correlation coefficients

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D.'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, https://doi.org/gffn86

Vinod, H. D. 'Matrix Algebra Topics in Statistics and Economics Using R', Chapter 4 in 'Handbook of Statistics: Computational Statistics with R', Vol.32, co-editors: M. B. Rao and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2014, pp. 143-176.

Vinod, H. D. 'New exogeneity tests and causal paths,' Chapter 2 in 'Handbook of Statistics: Conceptual Econometrics Using R', Vol.32, co-editors: H. D. Vinod and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2019, pp. 33-64.

Zheng, S., Shi, N.-Z., and Zhang, Z. (2012). 'Generalized measures of correlation for asymmetry, nonlinearity, and beyond,' Journal of the American Statistical Association, vol. 107, pp. 1239-1252.

Examples

```
## Not run:
x=1:20; y=sin(x)
gmcmtxBlk(cbind(x,y),blksiz=10)
## End(Not run)
```

`gmcmtxZ` *compute the matrix R* of generalized correlation coefficients.*

Description

This function checks for missing data separately for each pair using `kern` function to kernel regress x on y, and conversely y on x. It needs the library 'np' which reports R-squares of each regression. This function reports their square roots with the sign of the Pearson correlation coefficients. Its appeal is that it is asymmetric yielding causal direction information. It avoids the assumption of linearity implicit in the usual correlation coefficients.

Usage

```
gmcmtxZ(mym, nam = colnames(mym))
```

Arguments

`mym`	A matrix of data on variables in columns
`nam`	Column names of the variables in the data matrix

Value

A non-symmetric R* matrix of generalized correlation coefficients

Note

This allows the user to change `gmcmtx0` and further experiment with my code.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, https://doi.org/gffn86

Examples

```
## Not run:
set.seed(34);x=matrix(sample(1:600)[1:99],ncol=3)
colnames(x)=c('V1', 'v2', 'V3')
gmcmtxZ(x)

## End(Not run)
```

---

gmcxy_np   *Function to compute generalized correlation coefficients r*(x|y) and r*(y|x) from two vectors (not matrices)*

Description

This function uses the 'np' package and assumes that there are no missing data.

Usage

```
gmcxy_np(x, y)
```

Arguments

x	vector of x data
y	vector of y data

Value

corxy	r*(xly) from regressing x on y, where y is the kernel cause.
coryx	r*(ylx) from regressing y on x, where x is the cause.

Note

This is provided if the user want to avoid calling `kern`.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, https://doi.org/gffn86

Vinod, H. D. 'Matrix Algebra Topics in Statistics and Economics Using R,' Chapter 4 in 'Handbook of Statistics: Computational Statistics with R,' Vol.32, co-editors: M. B. Rao and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2014, pp. 143-176.

Examples

```
## Not run:
set.seed(34);x=sample(1:10);y=sample(2:11)
gmcxy_np(x,y)
## End(Not run)
```

---

goodCol *internal goodCol*

Description

intended for internal use only

Usage

```
goodCol
```

---

heurist *Heuristic t test of the difference between two generalized correlations.*

Description

Function to run a heuristic t test of the difference between two generalized correlations.

Usage

```
heurist(rxy, ryx, n)
```

Arguments

rxy	generalized correlation r*(x|y) where y is the kernel cause.
ryx	generalized correlation r*(y|x) where x is the kernel cause.
n	Sample size needed to determine the degrees of freedom for the t test.

Value

Prints the t statistics and p-values.

Note

This function requires Revele's R package called 'psych' in memory. This test is known to be conservative (i.e., often fails to reject the null hypothesis of zero difference between the two generalized correlation coefficients.)

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

Examples

```
set.seed(34);x=sample(1:10);y=sample(2:11)
g1=gmcxy_np(x,y)
n=length(x)
h1=heurist(g1$corxy,g1$coryx,n)
print(h1)
print(h1$t) #t statistic
print(h1$p) #p-value
```

`i` *internal i*

Description

intended for internal use

Usage

```
data(i)
```

Format

The format is: int 78

`ibad` *internal object*

Description

intended for internal use

`ii` *internal ii*

Description

intended for internal use

`j` *internal j*

Description

intended for internal use

Usage

```
data(j)
```

Format

The format is: int 4

`kern` *Kernel regression with options for residuals and gradients.*

Description

Function to run kernel regression with options for residuals and gradients asssuming no missing data.

Usage

```
kern(dep.y, reg.x, tol = 0.1, ftol = 0.1, gradients = FALSE, residuals = FALSE)
```

Arguments

`dep.y`	Data on the dependent (response) variable
`reg.x`	Data on the regressor (stimulus) variables
`tol`	Tolerance on the position of located minima of the cross-validation function (default =0.1)
`ftol`	Fractional tolerance on the value of cross validation function evaluated at local minima (default =0.1)
`gradients`	Make this TRUE if gradients computations are desired
`residuals`	Make this TRUE if residuals are desired

Value

Creates a model object 'mod' containing the entire kernel regression output. Type `names(mod)` to reveal the variety of outputs produced by 'npreg' of the 'np' package. The user can access all of them at will by using the dollar notation of R.

Note

This is a work horse for causal identification.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D.'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: 10.1080/03610918.2015.1122048

See Also

See `kern_ctrl`.

Examples

```
## Not run:
set.seed(34);x=matrix(sample(1:600)[1:50],ncol=2)
require(np); options(np.messages=FALSE)
k1=kern(x[,1],x[,2])
print(k1$R2) #prints the R square of the kernel regression

## End(Not run)
```

`kern2` *Kernel regression version 2 with optional residuals and gradients with regtype="ll" for local linear, bwmethod="cv.aic" for AIC-based bandwidth selection.*

Description

Kernel regression version 2 with optional residuals and gradients with regtype="ll" for local linear, bwmethod="cv.aic" for AIC-based bandwidth selection.

Usage

```
kern2(
  dep.y,
  reg.x,
  tol = 0.1,
  ftol = 0.1,
  gradients = FALSE,
  residuals = FALSE
)
```

Arguments

dep.y	Data on the dependent (response) variable
reg.x	Data on the regressor (stimulus) variables
tol	Tolerance on the position of located minima of the cross-validation function (default =0.1)
ftol	Fractional tolerance on the value of cross validation function evaluated at local minima (default =0.1)
gradients	Make this TRUE if gradients computations are desired
residuals	Make this TRUE if residuals are desired

Value

Creates a model object 'mod' containing the entire kernel regression output. Type `names(mod)` to reveal the variety of outputs produced by 'npreg' of the 'np' package. The user can access all of them at will by using the dollar notation of R.

Note

This is version 2 ("ll","cv.aic") of a work horse for causal identification.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D.'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: 10.1080/03610918.2015.1122048

See Also

See `kern_ctrl`.

Examples

```
## Not run:
set.seed(34);x=matrix(sample(1:600)[1:50],ncol=2)
require(np); options(np.messages=FALSE)
k1=kern(x[,1],x[,2])
print(k1$R2) #prints the R square of the kernel regression

## End(Not run)
```

`kern2ctrl` *Kernel regression with control variables and optional residuals and gradients. version 2 regtype="ll" for local linear, bwmethod="cv.aic" for AIC-based bandwidth selection. It admits control variables.*

Description

Kernel regression with control variables and optional residuals and gradients. version 2 regtype="ll" for local linear, bwmethod="cv.aic" for AIC-based bandwidth selection. It admits control variables.

Usage

```
kern2ctrl(
  dep.y,
  reg.x,
  ctrl,
  tol = 0.1,
  ftol = 0.1,
  gradients = FALSE,
  residuals = FALSE
)
```

Arguments

`dep.y`	Data on the dependent (response) variable
`reg.x`	Data on the regressor (stimulus) variable
`ctrl`	Data matrix on the control variable(s) kept outside the causal paths. A constant vector is not allowed as a control variable.
`tol`	Tolerance on the position of located minima of the cross-validation function (default=0.1)
`ftol`	Fractional tolerance on the value of cross validation function evaluated at local minima (default=0.1)
`gradients`	Set to TRUE if gradients computations are desired
`residuals`	Set to TRUE if residuals are desired

Value

Creates a model object 'mod' containing the entire kernel regression output. If this function is called as `mod=kern_ctrl(x,y,ctrl=z)`, the researcher can simply type `names(mod)` to reveal the large variety of outputs produced by 'npreg' of the 'np' package. The user can access all of them at will using the dollar notation of R.

Note

This is version 2 ("ll","cv.aic") of a work horse for causal identification.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: 10.1080/03610918.2015.1122048

See Also

See `kern`.

Examples

```
## Not run:
set.seed(34);x=matrix(sample(1:600)[1:50],ncol=5)
require(np)
k1=kern_ctrl(x[,1],x[,2],ctrl=x[,4:5])
print(k1$R2) #prints the R square of the kernel regression

## End(Not run)
```

`kern_ctrl` *Kernel regression with control variables and optional residuals and gradients.*

Description

Allowing matrix input of control variables, this function runs kernel regression with options for residuals and gradients.

Usage

```
kern_ctrl(
  dep.y,
  reg.x,
  ctrl,
  tol = 0.1,
  ftol = 0.1,
  gradients = FALSE,
  residuals = FALSE
)
```

Arguments

`dep.y`	Data on the dependent (response) variable
`reg.x`	Data on the regressor (stimulus) variable
`ctrl`	Data matrix on the control variable(s) kept outside the causal paths. A constant vector is not allowed as a control variable.
`tol`	Tolerance on the position of located minima of the cross-validation function (default=0.1)
`ftol`	Fractional tolerance on the value of cross validation function evaluated at local minima (default=0.1)
`gradients`	Set to TRUE if gradients computations are desired
`residuals`	Set to TRUE if residuals are desired

Value

Creates a model object 'mod' containing the entire kernel regression output. If this function is called as `mod=kern_ctrl(x,y,ctrl=z)`, the researcher can simply type `names(mod)` to reveal the large variety of outputs produced by 'npreg' of the 'np' package. The user can access all of them at will using the dollar notation of R.

Note

This is a work horse for causal identification.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: 10.1080/03610918.2015.1122048

See Also

See `kern`.

Examples

```
## Not run:
set.seed(34);x=matrix(sample(1:600)[1:50],ncol=5)
require(np)
k1=kern_ctrl(x[,1],x[,2],ctrl=x[,4:5])
print(k1$R2) #prints the R square of the kernel regression

## End(Not run)
```

---

mag *Approximate overall magnitudes of kernel regression partials dx/dy and dy/dx.*

Description

Uses Vinod (2015) and runs kernel regression of x on y, and also of y on x by using the 'np' package. The function goes on to compute a summary magnitude of the overall approximate partial derivative dx/dy (and dy/dx), after adjusting for units by using an appropriate ratio of standard deviations. Of course, the real partial derivatives of nonlinear functions are generally distinct for each observation.

Usage

```
mag(x, y)
```

Arguments

x	Vector of data on the dependent variable
y	Vector of data on the regressor

Value

vector of two magnitudes of kernel regression partials dx/dy and dy/dx.

Note

This function is intended for use only after the direction of causal path is already determined by various functions in this package (e.g. `somePairs`). For example, if the researcher knows that x causes y, then only dy/dx denoted by dydx is relevant. The other output of the function dxdy is to be ignored. Similarly, only 'dxdy' is relevant if y is known to be the cause of x.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: gffn86

Vinod, H. D. 'Matrix Algebra Topics in Statistics and Economics Using R', Chapter 4 in Handbook of Statistics: Computational Statistics with R, Vol.32, co-editors: M. B. Rao and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2014, pp. 143-176.

See Also

See `mag_ctrl`.

Examples

```
set.seed(123);x=sample(1:10);y=1+2*x+rnorm(10)
mag(x,y)#dxdy approx=.5 and dydx approx=2 will be nice.
```

`mag_ctrl` *After removing control variables, magnitude of effect of x on y, and of y on x.*

Description

Uses Vinod (2015) and runs kernel regressions: `x~ y + ctrl` and `x~ ctrl` to evaluate the 'incremental change' in R-squares. Let (rxy;ctrl) denote the square root of that 'incremental change' after its sign is made the same as that of the Pearson correlation coefficient from `cor(x,y)`). One can interpret (rxy;ctrl) as a generalized partial correlation coefficient when x is regressed on y after removing the effect of control variable(s) in `ctrl`. It is more general than the usual partial correlation coefficient, since this one allows for nonlinear relations among variables. Next, the function computes 'dxdy' obtained by multiplying (rxy;ctrl) by the ratio of standard deviations, `sd(x)/sd(y)`. Now our 'dxdy' approximates the magnitude of the partial derivative (dx/dy) in a causal model where y is the cause and x is the effect. The function also reports entirely analogous 'dydx' obtained by interchanging x and y.

Usage

```
mag_ctrl(x, y, ctrl)
```

Arguments

`x`	Vector of data on the dependent variable.
`y`	Vector of data on the regressor.
`ctrl`	data matrix for designated control variable(s) outside causal paths. A constant vector is not allowed as a control variable.

Value

vector of two magnitudes 'dxdy' (effect when x is regressed on y) and 'dydx' for reverse regression. Both regressions remove the effect of control variable(s).

Note

This function is intended for use only after the causal path direction is already determined by various functions in this package (e.g. `someCPairs`). That is, after the researcher knows whether x causes y or vice versa. The output of this function is a vector of two numbers: (dxdy, dydx), in that order, representing the magnitude of effect of one variable on the other. We expect the researcher to use only 'dxdy' if y is the known cause, or 'dydx' if x is the cause. These approximate overall measures may not be well-defined in some applications, because the real partial derivatives of nonlinear functions are generally distinct for each evaluation point.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: gffn86

Vinod, H. D. 'Matrix Algebra Topics in Statistics and Economics Using R', Chapter 4 in Handbook of Statistics: Computational Statistics with R, Vol.32, co-editors: M. B. Rao and C. R. Rao. New York: North Holland, Elsevier Science Publishers, 2014, pp. 143-176.

See Also

See `mag`

Examples

```
set.seed(123);x=sample(1:10); z=runif(10); y=1+2*x+3*z+rnorm(10)
options(np.messages=FALSE)
mag_ctrl(x,y,z)#dx/dy=0.47 is approximately 0.5, but dy/dx=1.41 is not approx=2,
```

`min.e0` *internal min.e0*

Description

intended for internal use only

Usage

```
min.e0
```

minor	*Function to do compute the minor of a matrix defined by row r and column c.*

Description

Function to do compute the minor of a matrix defined by row r and column c.

Usage

```
minor(x, r, c)
```

Arguments

x	The input matrix
r	The row number
c	The column number

Value

The appropriate 'minor' matrix defined from the input matrix.

Note

This function is needed by the cofactor function.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

Examples

```
## Not run:
 x=matrix(1:20,ncol=4)
minor(x,1,2)
## End(Not run)
```

momentVote *Function compares Pearson Stats and Sharpe Ratio for a matrix of stock returns*

Description

The first step computes mean, std.dev, skewness, kurtosis (kurt),and the Sharpe Ratio (mean/sd) representing risk-adjusted return where sd measures the risk. The input x must be a matrix having p columns (col.names recommended). and n rows as in the data. If data are missing for some columns, insert NA's. Thus x has p column of data matrix ready for comparison and ranking. For example, x has a matrix of stock returns. The output matrix produced by this function has p columns for each data column (i.e. for each stock being compared). The output matrix has twelve rows. Top five rows have the magnitudes of mean, sd, skew, kurt, Sharpe ratios. Output matrix rows 6 to 10 have respective ranks of moment stats. The output 11-th row reports a weighted sum of ranks with following weights mean=1,sd=-1,skew=0.5,kurt=-0.5,Sharpe Ratio=1. User has the option to change the weights. They measure relative importance.

Usage

```
momentVote(mtx, weight = c(1, -1, 0.5, -0.5, 1))
```

Arguments

`mtx`	n by p matrix of data, For example, n stock returns for p stocks. The mtx columns should have some names (ticker symbols)
`weight`	vector of reliability weights. default: mean=1, sd=1, skew=0.5,kurt=0.5,sharpe=1

Details

Since skewness and kurtosis are measured relatively less reliably (have greater sampling variation due to higher powers) their weight is 0.5. Our ranking gives the smallest number 1 to the most desirable outcome. The 11-th line of the output matrix has weighted sum of ranks and we suggest higher portfolio weight be given to the column having smallest value (in the bottom line). The 12-th row of output matrix has 'choice,' where input weights give the number 1 is for the top choice column of data and all other choice numbers. The (p+1)-th column of the output matrix has the chosen weights. The argument weight to the 'momentVote' function allows one to change these weights.

Value

a matrix with same number of columns as in the input matrix x and eleven rows. Top five rows have moment quantities, next five are their ranks the eleventh row has weighted sum of ranks with the input weights (see default) and the 12-th row has choice numbers (choice=1 is best)

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

Examples

```
x1=c(1,4,7,2,6)
x2=c(3,4,8,4,7)
momentVote(cbind(x1,x2))
```

`mtx` *internal mtx*

Description

intended for internal use only

Usage

```
mtx
```

`mtx0` *internal mtx0*

Description

intended for internal use only

Usage

```
mtx0
```

`mtx2` *internal mtx2*

Description

intended for internal use only

Usage

```
mtx2
```

`n` *internal n*

Description

intended for internal use

Usage

```
n
```

Format

The format is: int 78

`nall` *internal nall*

Description

intended for internal use only

Usage

```
nall
```

`nam.badCol` *internal nam.badCol*

Description

intended for internal use only

Usage

```
nam.badCol
```

`nam.goodCol` *internal nam.goodCol*

Description

intended for internal use only

Usage

```
nam.goodCol
```

`nam.mtx0` *internal nam.mtx0*

Description

intended for internal use only

Usage

```
nam.mtx0
```

`napair` *Function to do pairwise deletion of missing rows.*

Description

The aim in pair-wise deletions is to retain the largest number of available data pairs with all non-missing data.

Usage

```
napair(x, y)
```

Arguments

`x`	Vector of x data
`y`	Vector of y data

Value

`newx`	A new vector x after removing pairwise missing data
`newy`	A new vector y after removing pairwise missing data

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

Examples

```
## Not run:
x=sample(1:10);y=sample(1:10);x[2]=NA; y[3]=NA
napair(x,y)
## End(Not run)
```

`naTriple` *Function to do matched deletion of missing rows from x, y and z variable(s).*

Description

The aim in three-way deletions is to retain only the largest number of available data triplets with all non-missing data. This works where naTriplet fails (e.g.parcorVecH()). This is called by parcorHijk

Usage

```
naTriple(x, y, z)
```

Arguments

`x`	Vector of x data
`y`	Vector of y data
`z`	vector or a matrix of additional variable(s)

Value

`newx`	A new vector x after removing triplet-wise missing data
`newy`	A new vector or matrix y after removing triplet-wise missing data
`newz`	A new vector or matrix ctrl after removing triplet-wise missing data

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

See Also

See `napair` `naTriplet`.

Examples

```
## Not run:
x=sample(1:10);y=sample(1:10);x[2]=NA; y[3]=NA
w=sample(2:11)
naTriple(x,y,w)
## End(Not run)
```

`naTriplet` *Function to do matched deletion of missing rows from x, y and control variable(s).*

Description

The aim in three-way deletions is to retain only the largest number of available data triplets with all non-missing data.

Usage

```
naTriplet(x, y, ctrl)
```

Arguments

`x`	Vector of x data
`y`	Vector of y data
`ctrl`	Data matrix on the control variable(s) kept beyond causal path determinations

Value

`newx`	A new vector x after removing triplet-wise missing data
`newy`	A new vector or matrix y after removing triplet-wise missing data
`newctrl`	A new vector or matrix ctrl after removing triplet-wise missing data

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

See Also

See `napair`.

Examples

```
## Not run:
x=sample(1:10);y=sample(1:10);x[2]=NA; y[3]=NA
w=sample(2:11)
naTriplet(x,y,w)
## End(Not run)
```

NLhat *Compute fitted values from kernel regression of x on y and y on x*

Description

This is an auxiliary function for 'gmcmtxBlk.' It uses two numerical vectors (x, y) of same length to create two vectors (xhat, yhat) of fitted values using nonlinear kernel regressions. It uses package 'np' called by `kern` function to kernel regress x on y, and conversely y on x. It uses the option 'residuals=TRUE' of 'kern'

Usage

```
NLhat(x, y)
```

Arguments

x	A column vector of x data
y	A column vector of y data

Value

two vectors named xhat and yhat for fitted values

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

See Also

See Also as `gmcmtxBlk`.

Examples

```
## Not run:
set.seed(34);x=sample(1:15);y=sample(1:15)
NLhat(x,y)
## End(Not run)
```

out1 — *internal out1*

Description

intended for internal use only

Usage

```
out1
```

outOFsamp — *Compare out-of-sample portfolio choice algorithms by a leave-percent-out method.*

Description

This function randomly leaves out 5 percent (‘pctOut’=5 by default) data and finds portfolio choice by seven different portfolio selection algorithms using the data on the remaining 95 percent (say). For example, let the input to `outOFsamp` called ‘mtx’ is a matrix with p columns for p stocks and n returns. Also, let the maximum number of stocks admitted to belong in the portfolio be four, or ‘maxChosen=4’. Now `outOFsamp` function computes the returns earned by the seven portfolio selection algorithms, called "SD1","SD2","SD3","SD4","SDAll4","decile," and "moment," where SDAll4 refers to a weighted sum of SD1 to SD4 algorithms. Each algorithm provides a choice ranking of p stocks with choice values 1,2,3,..,p where stock ranked 1 should get the highest portfolio weight. The `outOFsamp` function then calls the function ‘rank2return’ which uses these rank choice numbers to the selected ‘maxChosen’ stocks. The allocation is linearly declining. For example, it is 1/10, 2/10, 3/10 and 4/10, with the top choice receiving 4/10 of the capital. Each choice of ‘pctOut’ rows of the ‘mtx’ data yields an outOFsamp return for each of the seven portfolio selection algorithms. These outOFsamp return computations are repeated `reps` times. A new random selection of ‘pctOut’ rows (must be 2 or more) of data is made for each repetition. We set reps=20 by default. The low default is set to save processing time in early phases, but we recommend reps=100+. The final choice of stock-picking algorithm out of seven is suggested by the average out-of-sample return over the ‘reps’ repetitions.‘

Usage

```
outOFsamp(mtx, pctOut = 5, reps = 10, seed = 23, maxChosen = 2, verbo = FALSE)
```

Arguments

`mtx`	matrix size n by p of data on n returns from p stocks
`pctOut`	percent of n randomly chosen rows left out as out-of-sample, default=5 percent. One must leave out at least two rows of data

`reps`	number of random repetitions of left-out rows over which we average the out-of-sample performance of a stock-picking algorithm, default reps=20
`seed`	seed for random number generation, default =23
`maxChosen`	number of stocks (out of p) with nonzero weights in the portfolio
`verbo`	logical, TRUE means print details, default=FALSE

Value

a matrix called 'avgRet' with seven columns for seven stock-picking algorithms "SD1","SD2","SD3","SD4","SDAll4","decile "moment," containing out-of-sample average returns for linearly declining allocation in a portfolio. User needs to change rank2return() for alternate portfolio allocations.

Note

The traditional time-series out-of-sample leaves out the last few time periods, and estimates the stock-picking model using part of the data time periods. The pandemic of 2019 has revealed that the traditional out-of-sample would have a severe bias in favor of pessimistic stock-picking algorithms. The traditional method is fundamentally flawed since it is sensitive to the trends (ups and downs) in the out-of-sample period. The method proposed here is free from such biases. The stock-picking algorithm recommended by our outOFsamp() is claimed to be robust against such biases.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

See Also

`rank2return`

Examples

```
## Not run:
x1=c(2,5,6,9,13,18,21,5,11,14,4,7,12,13,6,3,8,1,15,2,10,9)
x2=c(3,6,9,12,14,19,27,9,11,2,3,8,1,6,15,10,13,14,5,7,4,12)
x3=c(2,6,NA,11,13,25,25,11,9,10,12,6,4,3,2,1,7,8,5,15,14,13)
mtx=cbind(x1,x2,x3)
mtx=mtx[complete.cases(mtx),]
os=outOFsamp(mtx,verbo=FALSE,maxChosen=2, reps=3)
apply(os,2,mean)
## End(Not run)
```

p1 — *internal p1*

Description

intended for internal use only

Usage

```
p1
```

Panel2Lag — *Function to compute a vector of 2 lagged values of a variable from panel data.*

Description

The panel data have a set of time series for each entity (e.g. country) arranged such that all time series data for one entity is together. The data for the second entity should be below the entire data for first entity. When a variable is lagged twice, special care is needed to insert NA's for the first two time points (e.g. weeks) for each entity (country).

Usage

```
Panel2Lag(ID, xj)
```

Arguments

ID	Location of the column having time identities (e.g. the week number)
xj	Data on variable to be lagged linked to ID

Value

Vector containing 2 lagged values of xj.

Note

This function is provided for convenient user modifications.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

See Also

A more general function `PanelLag` has examples.

PanelLag *Function for computing a vector of one-lagged values of xj, a variable from panel data.*

Description

Panel data have a set of time series for each entity (e.g. country) arranged such that all time series data for one entity is together, and the data for the second entity should be below the entire data for first entity and so on for entities. In such a data setup, When a variable is lagged once, special care is needed to insert an NA for the first time point in the data (e.g. week) for each entity.

Usage

```
PanelLag(ID, xj, lag = 1)
```

Arguments

`ID`	Location of the column having time identities (e.g. week number).
`xj`	Data vector of variable to be lagged and is linked with the ID.
`lag`	Number of lags desired (lag=1 is the default).

Value

Vector containing one-lagged values of variable xj.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

Examples

```
## Not run:
indiv=gl(6,12,labels=LETTERS[1:6])
#creates A,A,A 12 times B B B also 12 times etc.
set.seed(99);cost=sample(30:90, 72, replace=TRUE)
revenu=sample(50:110, 72, replace=TRUE); month=rep(1:12,6)
df=data.frame(indiv,month,cost,revenu);head(df);tail(df)
L2cost=PanelLag(ID=month,xj=df[,'cost'], lag=2)
head(L2cost)
tail(L2cost)

gmcmtx0(cbind(revenu,cost,L2cost))

gmcxy_np(revenu,cost)

## End(Not run)
```

`parcorBijk` *Block version of generalized partial correlation coefficients between Xi and Xj, after removing the effect of xk, via nonparametric regression residuals.*

Description

This function uses data on two column vectors, xi, xj and a third xk which can be a vector or a matrix, usually of the remaining variables in the model, including control variables, if any. It first removes missing data from all input variables. Then, it computes residuals of kernel regression (xi on xk) and (xj on xk). This is a block version of parcor_ijk.

Usage

```
parcorBijk(xi, xj, xk, blksiz = 10)
```

Arguments

`xi`	Input vector of data for variable xi
`xj`	Input vector of data for variable xj
`xk`	Input data for variables in xk, usually control variables
`blksiz`	block size, default=10, if chosen blksiz >n, where n=rows in matrix then blksiz=n. That is, no blocking is done

Value

`ouij`	Generalized partial correlation Xi with Xj (=cause) after removing xk
`ouji`	Generalized partial correlation Xj with Xi (=cause) after removing xk

allowing for control variables.

Note

This function calls `kern`,

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

See Also

See `parcor_ijk`.

Examples

```
## Not run:
set.seed(34);x=matrix(sample(1:600)[1:99],ncol=3)
options(np.messages=FALSE)
parcorBijk(x[,1], x[,2], x[,3], blksi=10)

## End(Not run)#'
```

`parcorBMany` *Block version reports many generalized partial correlation coefficients allowing control variables.*

Description

This function calls a block version `parcorBijk` of the function which uses original data to compute generalized partial correlations between X_{idep} and X_j where j can be any one of the remaining variables in the input matrix `mtx`. Partial correlations remove the effect of variables X_k other than X_i and X_j. Calculation further allows for the presence of control variable(s) (if any) to remain always outside the input matrix and whose effect is also removed in computing partial correlations.

Usage

```
parcorBMany(mtx, ctrl = 0, dig = 4, idep = 1, blksiz = 10, verbo = FALSE)
```

Arguments

`mtx`	Input data matrix with at least 3 columns.
`ctrl`	Input vector or matrix of data for control variable(s), default is ctrl=0 when control variables are absent
`dig`	The number of digits for reporting (=4, default)
`idep`	The column number of the dependent variable (=1, default)
`blksiz`	block size, default=10, if chosen blksiz >n, where n=rows in matrix then blksiz=n. That is, no blocking is done
`verbo`	Make this TRUE for detailed printing of computational steps

Value

A five column 'out' matrix containing partials. The first column has the name of the `idep` variable. The second column has the name of the j variable, while the third column has partial correlation coefficients r*(i,j | k).The last column reports the absolute difference between two partial correlations.

Note

This function reports all partial correlation coefficients, while avoiding ridge type adjustment.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

Vinod, H. D. 'Generalized Correlations and Instantaneous Causality for Data Pairs Benchmark,' (March 8, 2015) https://www.ssrn.com/abstract=2574891

Vinod, H. D. (2021) 'Generalized, Partial and Canonical Correlation Coefficients' Computational Economics, 59(1), 1–28.

Vinod, H. D. 'Matrix Algebra Topics in Statistics and Economics Using R', Chapter 4 in Handbook of Statistics: Computational Statistics with R, Vol.32, co-editors: M. B. Rao and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2014, pp. 143-176.

See Also

See Also parcor_ijk, parcorMany.

Examples

```
set.seed(234)
z=runif(10,2,11)# z is independently created
x=sample(1:10)+z/10  #x is partly indep and partly affected by z
y=1+2*x+3*z+rnorm(10)# y depends on x and z not vice versa
mtx=cbind(x,y,z)
parcorBMany(mtx, blksiz=10)


## Not run:
set.seed(34);x=matrix(sample(1:600)[1:99],ncol=3)
colnames(x)=c('V1', 'v2', 'V3')
parcorBMany(x, idep=1)

## End(Not run)
```

---

parcorHijk   *Generalized partial correlation coefficients between Xi and Xj, after removing the effect of Xk, via OLS regression residuals.*

Description

This function uses data on two column vectors, xi, xj, and a third set xk, which can be a vector or a matrix. xk usually has the remaining variables in the model, including control variables, if any. This function first removes missing data from all input variables. Then, it computes residuals of OLS (no kernel) regression (xi on xk) and (xj on xk). This hybrid version uses both OLS and then generalized correlation among OLS residuals. This solves the potential problem of having too little information content in kernel regression residuals, since kernel fits are sometimes too close, especially when there are many variables in xk.

Usage

```
parcorHijk(xi, xj, xk)
```

Arguments

xi	Input vector of data for variable xi
xj	Input vector of data for variable xj
xk	Input data for all variables in xk, usually control variables

Value

ouij	Generalized partial correlation Xi with Xj (=cause) after removing xk
ouji	Generalized partial correlation Xj with Xi (=cause) after removing xk

allowing for control variables.

Note

This function calls kern,

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

See Also

See parcor_ijk.

Examples

```
## Not run:
set.seed(34);x=matrix(sample(1:600)[1:99],ncol=3)
options(np.messages=FALSE)
parcorHijk(x[,1], x[,2], x[,3])

## End(Not run)#'
```

---

parcorHijk2 *Generalized partial correlation coefficients between Xi and Xj,*

Description

The 2 in the name of the function means second version. The H in the function name means hybrid. This removes the effect of Xk, via OLS regression residuals. This function uses data on two column vectors, xi, xj, and a third set xk, which can be a vector or a matrix, usually of the remaining variables in the model, including control variables, if any. It first removes missing data from all input variables. Then, it computes residuals of OLS regression (xi on xk) and (xj on xk). The function reports the generalized correlation between two OLS residuals. This hybrid version uses both OLS and then generalized correlation among OLS residuals. This second version works when 'parcorVecH' fails. It is called by the function 'parcorVecH2'.

Usage

```
parcorHijk2(xi, xj, xk)
```

Arguments

`xi`	Input vector of data for variable xi
`xj`	Input vector of data for variable xj
`xk`	Input data for variables in xk, usually control variables

Value

`ouij`	Generalized partial correlation Xi with Xj (=cause) after removing xk
`ouji`	Generalized partial correlation Xj with Xi (=cause) after removing xk

allowing for control variables.

Note

This function calls `kern`,

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

See Also

See `parcor_ijk`.

Examples

```
## Not run:
set.seed(34);x=matrix(sample(1:600)[1:99],ncol=3)
options(np.messages=FALSE)
parcorHijk2(x[,1], x[,2], x[,3])

## End(Not run)#'
```

parcorMany *Report many generalized partial correlation coefficients allowing control variables.*

Description

This function calls parcor_ijk function which uses original data to compute generalized partial correlations between X_{idep} and X_j where j can be any one of the remaining variables in the input matrix mtx. Partial correlations remove the effect of variables x_k other than X_i and X_j. Calculation further allows for the presence of control variable(s) (if any) to remain always outside the input matrix and whose effect is also removed in computing partial correlations.

Usage

```
parcorMany(mtx, ctrl = 0, dig = 4, idep = 1, verbo = FALSE)
```

Arguments

mtx	Input data matrix with at least 3 columns.
ctrl	Input vector or matrix of data for control variable(s), default is ctrl=0 when control variables are absent
dig	The number of digits for reporting (=4, default)
idep	The column number of the first variable (=1, default)
verbo	Make this TRUE for detailed printing of computational steps

Value

A five column 'out' matrix containing partials. The first column has the name of the idep variable. The second column has the name of the j variable, while the third column has partial correlation coefficients r*(i,j | k). The last column reports the absolute difference between two partial correlations.

Note

This function reports all partial correlation coefficients, while avoiding ridge type adjustment.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

Vinod, H. D. 'Generalized Correlations and Instantaneous Causality for Data Pairs Benchmark,' (March 8, 2015) https://www.ssrn.com/abstract=2574891

Vinod, H. D. 'Matrix Algebra Topics in Statistics and Economics Using R', Chapter 4 in Handbook of Statistics: Computational Statistics with R, Vol.32, co-editors: M. B. Rao and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2014, pp. 143-176.

See Also

See Also parcor_ijk.

Examples

```
set.seed(234)
z=runif(10,2,11)# z is independently created
x=sample(1:10)+z/10  #x is partly indep and partly affected by z
y=1+2*x+3*z+rnorm(10)# y depends on x and z not vice versa
mtx=cbind(x,y,z)
parcorMany(mtx)


## Not run:
set.seed(34);x=matrix(sample(1:600)[1:99],ncol=3)
colnames(x)=c('V1', 'v2', 'V3')
parcorMany(x, idep=1)

## End(Not run)
```

---

parcorMtx — *Matrix of generalized partial correlation coefficients, always leaving out control variables, if any.*

Description

This function calls parcor_ijk function which uses original data to compute generalized partial correlations between X_i and X_j where j can be any one of the remaining variables in the input matrix mtx. Partial correlations remove the effect of variables x_k other than X_i and X_j. Calculation further allows for the presence of control variable(s) (if any) to remain always outside the input matrix and whose effect is also removed in computing partial correlations.

Usage

```
parcorMtx(mtx, ctrl = 0, dig = 4, verbo = FALSE)
```

Arguments

mtx	Input data matrix with p columns. p is at least 3 columns.
ctrl	Input vector or matrix of data for control variable(s), default is ctrl=0 when control variables are absent
dig	The number of digits for reporting (=4, default)
verbo	Make this TRUE for detailed printing of computational steps

Value

A p by p ‘out’ matrix containing partials r*(i,j | k). and r*(j,i | k).

Note

We want to get all partial correlation coefficient pairs removing other column effects. Vinod (2018) shows why one needs more than one criterion to decide the causal paths or exogeneity.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

Vinod, H. D. 'Generalized Correlations and Instantaneous Causality for Data Pairs Benchmark,' (March 8, 2015) https://www.ssrn.com/abstract=2574891

Vinod, H. D. 'Matrix Algebra Topics in Statistics and Economics Using R', Chapter 4 in Handbook of Statistics: Computational Statistics with R, Vol.32, co-editors: M. B. Rao and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2014, pp. 143-176.

Vinod, H. D. 'New Exogeneity Tests and Causal Paths,' (June 30, 2018). Available at SSRN: https://www.ssrn.com/abstract=3206096

See Also

See Also parcor_ijk.

Examples

```
set.seed(234)
z=runif(10,2,11)# z is independently created
x=sample(1:10)+z/10  #x is partly indep and partly affected by z
y=1+2*x+3*z+rnorm(10)# y depends on x and z not vice versa
mtx=cbind(x,y,z)
parcorMtx(mtx)


## Not run:
set.seed(34);x=matrix(sample(1:600)[1:99],ncol=3)
colnames(x)=c('V1', 'v2', 'V3')
parcorMtx(x)

## End(Not run)
```

---

parcorSilent *Silently compute generalized (ridge-adjusted) partial correlation coefficients from matrix R*.*

Description

This function calls `parcor_ijkOLD` function which uses a generalized correlation matrix R* as input to compute generalized partial correlations between X_i and X_j where j can be any one of the remaining variables. Computation removes the effect of all other variables in the matrix. It further adjusts the resulting partial correlation coefficients to be in the appropriate [-1,1] range by using an additive constant in the fashion of ridge regression.

Usage

```
parcorSilent(gmc0, dig = 4, idep = 1, verbo = FALSE, incr = 3)
```

Arguments

`gmc0`	This must be a p by p matrix R* of generalized correlation coefficients.
`dig`	The number of digits for reporting (=4, default)
`idep`	The column number of the first variable (=1, default)
`verbo`	Make this TRUE for detailed printing of computational steps
`incr`	incremental constant for iteratively adjusting 'ridgek' where ridgek is the constant times the identity matrix used to make sure that the gmc0 matrix is positive definite. If not, this function iteratively increases the `incr` till relevant partial correlations are within the [-1,1] interval.

Value

A five column 'out' matrix containing partials. The first column has the name of the `idep` variable. The second column has the name of the j variable, while the third column has r*(i,j | k). The 4-th column has r*(j,i | k) (denoted partji), and the 5-th column has rijMrji, that is the difference in absolute values (abs(partij) - abs(partji)).

Note

The ridgek constant created by the function during the first round may not be large enough to make sure that that other pairs of r*(i,j | k) are within the [-1,1] interval. The user may have to choose a suitably larger input `incr` to get all relevant partial correlation coefficients in the correct [-1,1] interval.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

Vinod, H. D. 'Generalized Correlations and Instantaneous Causality for Data Pairs Benchmark,' (March 8, 2015) https://www.ssrn.com/abstract=2574891

Vinod, H. D. 'Matrix Algebra Topics in Statistics and Economics Using R', Chapter 4 in Handbook of Statistics: Computational Statistics with R, Vol.32, co-editors: M. B. Rao and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2014, pp. 143-176.

Vinod, H. D. "A Survey of Ridge Regression and Related Techniques for Improvements over Ordinary Least Squares," Review of Economics and Statistics, Vol. 60, February 1978, pp. 121-131.

See Also

See Also parcor_ijk for a better version using original data as input.

Examples

```
set.seed(234)
z=runif(10,2,11)# z is independently created
x=sample(1:10)+z/10  #x is partly indep and partly affected by z
y=1+2*x+3*z+rnorm(10)# y depends on x and z not vice versa
mtx=cbind(x,y,z)
g1=gmcmtx0(mtx)
parcor_ijkOLD(g1,1,2) # ouji> ouij implies i=x is the cause of j=y
parcor_ridg(g1,idep=1)
parcorSilent(g1,idep=1)


## Not run:
set.seed(34);x=matrix(sample(1:600)[1:99],ncol=3)
colnames(x)=c('V1', 'v2', 'V3')
gm1=gmcmtx0(x)
parcorSilent(gm1, idep=1)

## End(Not run)
```

---

parcorVec — *Vector of generalized partial correlation coefficients (GPCC), always leaving out control variables, if any.*

Description

This function calls `parcor_ijk` function which uses original data to compute generalized partial correlations between X_i, the dependent variable, and X_j which is the current regressor of interest. Note that j can be any one of the remaining variables in the input matrix `mtx`. Partial correlations remove the effect of variables X_k other than X_i and X_j. Calculation merges control variable(s) (if any) into X_k. Let the remainder effect from kernel regressions of X_i on X_k equal the residuals u*(i,k). Analogously define u*(j,k). (asterisk for kernel regressions) Now partial correlation is generalized correlation between u*(i,k) and u*(j,k). Calculation merges control variable(s) (if any) into X_k.

Usage

```
parcorVec(mtx, ctrl = 0, verbo = FALSE, idep = 1)
```

Arguments

`mtx`	Input data matrix with p (> or = 3) columns
`ctrl`	Input vector or matrix of data for control variable(s), default is ctrl=0 when control variables are absent
`verbo`	Make this TRUE for detailed printing of computational steps
`idep`	The column number of the dependent variable (=1, default)

Value

A p by 1 'out' vector containing partials r*(i,j | k).

Note

Generalized Partial Correlation Coefficients (GPCC) allow comparison of the relative contribution of each X_j to the explanation of X_i, because GPCC are scale-free pure numbers

We want to get all partial correlation coefficient pairs removing other column effects. Vinod (2018) shows why one needs more than one criterion to decide the causal paths or exogeneity.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

Vinod, H. D. 'Generalized Correlations and Instantaneous Causality for Data Pairs Benchmark,' (March 8, 2015) https://www.ssrn.com/abstract=2574891

Vinod, H. D. 'Matrix Algebra Topics in Statistics and Economics Using R', Chapter 4 in Handbook of Statistics: Computational Statistics with R, Vol.32, co-editors: M. B. Rao and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2014, pp. 143-176.

Vinod, H. D. 'New Exogeneity Tests and Causal Paths,' (June 30, 2018). Available at SSRN: https://www.ssrn.com/abstract=3206096

Vinod, H. D. (2021) 'Generalized, Partial and Canonical Correlation Coefficients' Computational Economics, 59(1), 1–28.

See Also

See Also `parcor_ijk`.

See Also a hybrid version `parcorVecH`.

Examples

```
set.seed(234)
z=runif(10,2,11)# z is independently created
x=sample(1:10)+z/10  #x is partly indep and partly affected by z
y=1+2*x+3*z+rnorm(10)# y depends on x and z not vice versa
mtx=cbind(x,y,z)
parcorVec(mtx)
```

```
## Not run:
set.seed(34);x=matrix(sample(1:600)[1:99],ncol=3)
colnames(x)=c('V1', 'v2', 'V3')#some names needed
parcorVec(x)

## End(Not run)
```

`parcorVecH` *Vector of hybrid generalized partial correlation coefficients.*

Description

This is a hybrid version of parcorVec subtracting only the linear effects (OLS residuals instead of kernel regression residuals), but using the generalized correlation between the OLS residuals for the last stage of the generalized partial correlation.

Usage

```
parcorVecH(mtx, ctrl = 0, dig = 4, verbo = FALSE, idep = 1)
```

Arguments

`mtx`	Input data matrix with p (> or = 3) columns, the first column must have the dependent variable
`ctrl`	Input vector or matrix of data for control variable(s), default is ctrl=0 when control variables are absent
`dig`	The number of digits for reporting (=4, default)
`verbo`	Make this TRUE for detailed printing of computational steps
`idep`	The column number of the dependent variable (=1, default)

Details

This function calls `parcor_ijk` function, which uses original data to compute generalized partial correlations between X_i, the dependent variable, and X_j, which is the current regressor of interest. Note that j can be any one of the remaining variables in the input matrix `mtx`. Partial correlations remove the effect of variables X_k other than X_i and X_j. Calculation merges control variable(s) (if any) into X_k. Let the remainder effect from OLS regressions of X_i on X_k equal the residuals u(i,k). Analogously define u(j,k). It is a hybrid of OLS and generalized. Finally, partial correlation is generalized (kernel) correlation between u(i,k) and u(j,k).

Value

A p by 1 'out' vector containing hybrid partials r*(i,j | k).

Note

Hybrid Generalized Partial Correlation Coefficients (HGPCC) allow comparison of the relative contribution of each X_j to the explanation of X_i, because HGPCC has scale-free pure numbers.

We want to get all partial correlation coefficient pairs removing other column effects. Vinod (2018) shows why one needs more than one criterion to decide the causal paths or exogeneity.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

Vinod, H. D. 'Generalized Correlations and Instantaneous Causality for Data Pairs Benchmark,' (March 8, 2015) https://www.ssrn.com/abstract=2574891

Vinod, H. D. 'Matrix Algebra Topics in Statistics and Economics Using R', Chapter 4 in Handbook of Statistics: Computational Statistics with R, Vol.32, co-editors: M. B. Rao and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2014, pp. 143-176.

Vinod, H. D. 'New Exogeneity Tests and Causal Paths,' (June 30, 2018). Available at SSRN: https://www.ssrn.com/abstract=3206096

Vinod, H. D. (2021) 'Generalized, Partial and Canonical Correlation Coefficients' Computational Economics, 59(1), 1–28.

See Also

See Also parcor_ijk.

See Also parcorVec.

Examples

```
set.seed(234)
z=runif(10,2,11)# z is independently created
x=sample(1:10)+z/10  #x is partly indep and partly affected by z
y=1+2*x+3*z+rnorm(10)# y depends on x and z not vice versa
mtx=cbind(x,y,z)
parcorVecH(mtx)


## Not run:
set.seed(34);mtx=matrix(sample(1:600)[1:80],ncol=4)
colnames(mtx)=c('V1', 'v2', 'V3', 'V4')
parcorVecH(mtx,verbo=TRUE, idep=2)

## End(Not run)
```

parcorVecH2 *Vector of hybrid generalized partial correlation coefficients.*

Description

This is a second version to be used when 'parcorVecH' fails. (H=hybrid). This hybrid version of parcorVec subtracting only linear effects but using generlized correlation between OLS residuals

Usage

```
parcorVecH2(mtx, dig = 4, verbo = FALSE, idep = 1)
```

Arguments

`mtx`	Input data matrix with p (> or = 3) columns, first column must have the dependent variable
`dig`	The number of digits for reporting (=4, default)
`verbo`	Make this TRUE for detailed printing of computational steps
`idep`	The column number of the dependent variable (=1, default)

Details

This function calls `parcorHijk2` function which uses original data to compute generalized partial correlations between X_i, the dependent variable, and X_j which is the current regressor of interest. Note that j can be any one of the remaining variables in the input matrix `mtx`. Partial correlations remove the effect of variables X_k other than X_i and X_j. Calculation merges control variable(s) (if any) into X_k. Let the remainder effect from OLS regressions of X_i on X_k equal the residuals u(i,k). Analogously define u(j,k). It is a hybrid of OLS and generalized. Finally, partial correlation is generalized (kernel) correlation between u(i,k) and u(j,k).

Value

A p by 1 'out' vector containing hybrid partials r*(i,j | k).

Note

Hybrid Generalized Partial Correlation Coefficients (HGPCC) allow comparison of the relative contribution of each X_j to the explanation of X_i, because HGPCC are scale-free pure numbers.

We want to get all partial correlation coefficient pairs removing other column effects. Vinod (2018) shows why one needs more than one criterion to decide the causal paths or exogeneity.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

Vinod, H. D. 'Generalized Correlations and Instantaneous Causality for Data Pairs Benchmark,' (March 8, 2015) https://www.ssrn.com/abstract=2574891

Vinod, H. D. 'Matrix Algebra Topics in Statistics and Economics Using R', Chapter 4 in Handbook of Statistics: Computational Statistics with R, Vol.32, co-editors: M. B. Rao and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2014, pp. 143-176.

Vinod, H. D. 'New Exogeneity Tests and Causal Paths,' (June 30, 2018). Available at SSRN: https://www.ssrn.com/abstract=3206096

Vinod, H. D. (2021) 'Generalized, Partial and Canonical Correlation Coefficients' Computational Economics, 59(1), 1–28.

See Also

See Also parcor_ijk.

See Also parcorVec.

Examples

```
set.seed(234)
z=runif(10,2,11)# z is independently created
x=sample(1:10)+z/10  #x is partly indep and partly affected by z
y=1+2*x+3*z+rnorm(10)# y depends on x and z not vice versa
mtx=cbind(x,y,z)
parcorVecH2(mtx)


## Not run:
set.seed(34);mtx=matrix(sample(1:600)[1:80],ncol=4)
colnames(mtx)=c('V1', 'v2', 'V3', 'V4')
parcorVecH2(mtx,verbo=TRUE, idep=2)

## End(Not run)
```

---

parcor_ijk — *Generalized partial correlation coefficients between Xi and Xj, after removing the effect of xk, via nonparametric regression residuals.*

Description

This function uses data on two column vectors, xi, xj and a third xk which can be a vector or a matrix, usually of the remaining variables in the model, including control variables, if any. It first removes missing data from all input variables. Then, it computes residuals of kernel regression (xi on xk) and (xj on xk). The function reports the generalized correlation between two kernel residuals. This version avoids ridge type adjustment present in an older version.

Usage

```
parcor_ijk(xi, xj, xk)
```

Arguments

xi	Input vector of data for variable xi
xj	Input vector of data for variable xj
xk	Input data for variables in xk, usually control variables

Value

ouij	Generalized partial correlation Xi with Xj (=cause) after removing xk
ouji	Generalized partial correlation Xj with Xi (=cause) after removing xk

allowing for control variables.

Note

This function calls kern,

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

See Also

See parcor_linear.

Examples

```
## Not run:
set.seed(34);x=matrix(sample(1:600)[1:99],ncol=3)
options(np.messages=FALSE)
parcor_ijk(x[,1], x[,2], x[,3])

## End(Not run)#'
```

---

parcor_ijkOLD

Generalized partial correlation coefficient between Xi and Xj after removing the effect of all others. (older version, deprecated)

Description

This function uses a generalized correlation matrix R* as input to compute generalized partial correlations between X_i and X_j where j can be any one of the remaining variables. Computation removes the effect of all other variables in the matrix. The user is encouraged to remove all known irrelevant rows and columns from the R* matrix before submitting it to this function.

Usage

```
parcor_ijkOLD(x, i, j)
```

Arguments

x	Input a p by p matrix R* of generalized correlation coefficients.
i	A column number identifying the first variable.
j	A column number identifying the second variable.

Value

ouij	Partial correlation Xi with Xj (=cause) after removing all other X's
ouji	Partial correlation Xj with Xi (=cause) after removing all other X's
myk	A list of column numbers whose effect has been removed

Note

This function calls `minor`, and `cofactor` and is called by `parcor_ridge`.

Examples

```
## Not run:
set.seed(34);x=matrix(sample(1:600)[1:99],ncol=3)
colnames(x)=c('V1', 'v2', 'V3')
gm1=gmcmtx0(x)
parcor_ijkOLD(gm1, 2,3)

## End(Not run)#'
```

`parcor_linear` *Partial correlation coefficient between Xi and Xj after removing the linear effect of all others.*

Description

This function uses a symmetric correlation matrix R as input to compute usual partial correlations between X_i and X_j where j can be any one of the remaining variables. Computation removes the effect of all other variables in the matrix. The user is encouraged to remove all known irrelevant rows and columns from the R matrix before submitting it to this function.

Usage

```
parcor_linear(x, i, j)
```

Arguments

x	Input a p by p matrix R of symmetric correlation coefficients.
i	A column number identifying the first variable.
j	A column number identifying the second variable.

Value

ouij	Partial correlation Xi with Xj after removing all other X's
ouji	Partial correlation Xj with Xi after removing all other X's
myk	A list of column numbers whose effect has been removed

Note

This function calls `minor`, and `cofactor`

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

See Also

See `parcor_ijk` for generalized partial correlation coefficients useful for causal path determinations.

Examples

```
## Not run:
set.seed(34);x=matrix(sample(1:600)[1:99],ncol=3)
colnames(x)=c('V1', 'v2', 'V3')
c1=cor(x)
parcor_linear(c1, 2,3)

## End(Not run)
```

---

parcor_ridg — *Compute generalized (ridge-adjusted) partial correlation coefficients from matrix R*. (deprecated)*

Description

This function calls `parcor_ijkOLD` function which uses a generalized correlation matrix R* as input to compute generalized partial correlations between X_i and X_j where j can be any one of the remaining variables. Computation removes the effect of all other variables in the matrix. It further adjusts the resulting partial correlation coefficients to be in the appropriate [-1,1] range by using an additive constant in the fashion of ridge regression.

Usage

```
parcor_ridg(gmc0, dig = 4, idep = 1, verbo = FALSE, incr = 3)
```

Arguments

gmc0	This must be a p by p matrix R* of generalized correlation coefficients.
dig	The number of digits for reporting (=4, default)
idep	The column number of the first variable (=1, default)
verbo	Make this TRUE for detailed printing of computational steps
incr	incremental constant for iteratively adjusting 'ridgek' where ridgek is the constant times the identity matrix used to make sure that the gmc0 matrix is positive definite. If not iteratively increas the `incr` till all partial correlations are within the [-1,1] interval.

Value

A five column 'out' matrix containing partials. The first column has the name of the `idep` variable. The second column has the name of the j variable, while the third column has r*(i,j | k). The 4-th column has r*(j,i | k) (denoted partji), and the 5-th column has rijMrji, that is the difference in absolute values (abs(partij) - abs(partji)).

Note

The ridgek constant created by the function during the first round may not be large enough to make sure that that other pairs of r*(i,j | k) are within the [-1,1] interval. The user may have to choose a suitably larger input `incr` to get all relevant partial correlation coefficients in the correct [-1,1] interval.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

Vinod, H. D. 'Generalized Correlations and Instantaneous Causality for Data Pairs Benchmark,' (March 8, 2015) https://www.ssrn.com/abstract=2574891

Vinod, H. D. 'Matrix Algebra Topics in Statistics and Economics Using R', Chapter 4 in Handbook of Statistics: Computational Statistics with R, Vol.32, co-editors: M. B. Rao and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2014, pp. 143-176.

Vinod, H. D. "A Survey of Ridge Regression and Related Techniques for Improvements over Ordinary Least Squares," Review of Economics and Statistics, Vol. 60, February 1978, pp. 121-131.

See Also

See Also `parcor_ijkOLD`.

Examples

```
set.seed(234)
z=runif(10,2,11)# z is independently created
x=sample(1:10)+z/10  #x is partly indep and partly affected by z
y=1+2*x+3*z+rnorm(10)# y depends on x and z not vice versa
mtx=cbind(x,y,z)
g1=gmcmtx0(mtx)
parcor_ijkOLD(g1,1,2) # ouji> ouij implies i=x is the cause of j=y
parcor_ridg(g1,idep=1)


## Not run:
set.seed(34);x=matrix(sample(1:600)[1:99],ncol=3)
colnames(x)=c('V1', 'v2', 'V3')
gm1=gmcmtx0(x)
parcor_ridg(gm1, idep=1)

## End(Not run)
```

---

pcause *Compute the bootstrap probability of correct causal direction.*

Description

Maximum entropy bootstrap ('meboot') package is used for statistical inference regarding δ which equals GMC(X|Y)-GMC(Y|X) defined by Zheng et al (2012). The bootstrap provides an approximation to chances of correct determination of the causal direction.

Usage

```
pcause(x, y, n999 = 999)
```

Arguments

x	Vector of x data
y	Vector of y data
n999	Number of bootstrap replications (default=999)

Value

P(cause) the bootstrap proportion of correct causal determinations.

Note

'pcause' is computer intensive and generally slow. It is better to use it at a later stage in the investigation when a preliminary causal determination is already made. Its use may slow the exploratory phase. In my experience, if P(cause) is less than 0.55, there is a cause for concern.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: 10.1080/03610918.2015.1122048

Zheng, S., Shi, N.-Z., and Zhang, Z. (2012). Generalized measures of correlation for asymmetry, nonlinearity, and beyond. Journal of the American Statistical Association, vol. 107, pp. 1239-1252.

Vinod, H. D. and Lopez-de-Lacalle, J. (2009). 'Maximum entropy bootstrap for time series: The meboot R package.' Journal of Statistical Software, Vol. 29(5), pp. 1-19.

Examples

```
## Not run:
set.seed(34);x=sample(1:10);y=sample(2:11)
pcause(x,y,n999=29)

data('EuroCrime')
attach(EuroCrime)
pcause(crim,off,n999=29)

## End(Not run)
```

---

pillar3D — *Create a 3D pillar chart to display (x, y, z) data coordinate surface.*

Description

Give data on x, y, z coordinate values of a 3D surface, this function plots them after making pillars near each z value by adding and subtracting small amounts dz. Instead of pins of the height z this creates pillars which better resemble a surface. It uses the wireframe() function of 'lattice' package to do the plotting.

Usage

```
pillar3D(
  z = c(657, 936, 1111, 1201),
  x = c(280, 542, 722, 1168),
  y = c(162, 214, 186, 246),
  drape = TRUE,
  xlab = "y",
  ylab = "x",
  zlab = "z",
  mymain = "Pillar Chart"
)
```

Arguments

`z`	z-coordinate values
`x`	x-coordinate values
`y`	y-coordinate values
`drape`	logical value, default drape=TRUE to give color to heights
`xlab`	default "x" label on the x axis
`ylab`	default "y" label on the y axis
`zlab`	default "z" label on the z axis
`mymain`	default "Pillar Chart" main label on the plot

Details

For additional plotting features type 'pillar3D()' on the R console to get my code and adjust wire-frame() function defaults.

Value

A 3D plot

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

Examples

```
## Not run:
pillar3D())
## End(Not run)
```

`prelec2` *Intermediate weighting function giving Non-Expected Utility theory weights.*

Description

Computes cumulative probabilities and difference between consecutive cumulative probabilities described in Vinod (2008) textbook. This is a simpler version of the version in the book without mapping to non-expected utility theory weights as explained in Vinod (2008).

Usage

```
prelec2(n)
```

Arguments

`n`	A (usually small) integer.

Value

`x`	sequence 1:n
`p`	probabilities p= x[i]/n
`pdif`	consecutive differences p[i] - p[i - 1]

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Hands-On Intermediate Econometrics Using R' (2008) World Scientific Publishers: Hackensack, NJ. https://www.worldscientific.com/worldscibooks/10.1142/6895

Examples

```
## Not run: prelec2(10)
```

`probSign` *Compute probability of positive or negative sign from bootPairs output*

Description

If there are p columns of data, `probSign` produces a p-1 by 1 vector of probabilities of correct signs assuming that the mean of n999 values has the correct sign and assuming that m of the 'sum' index values inside the range [-tau, tau] are neither positive nor negative but indeterminate or ambiguous (being too close to zero). That is, the denominator of P(+1) or P(-1) is (n999-m) if m signs are too close to zero.

Usage

```
probSign(out, tau = 0.476)
```

Arguments

`out`	output from bootPairs with p-1 columns and n999 rows
`tau`	threshold to determine what value is too close to zero, default tau=0.476 is equivalent to 15 percent threshold for the unanimity index ui

Value

sgn When mtx has p columns, sgn reports pairwise p-1 signs representing (fixing the first column in each pair) the average sign after averaging the output of of bootPairs(mtx) (a n999 by p-1 matrix) each containing resampled 'sum' values summarizing the weighted sums associated with all three criteria from the function silentPairs(mtx) applied to each bootstrap sample separately. #'

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: 10.1080/03610918.2015.1122048

Vinod, H. D. and Lopez-de-Lacalle, J. (2009). 'Maximum entropy bootstrap for time series: The meboot R package.' Journal of Statistical Software, Vol. 29(5), pp. 1-19.

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: https://www.ssrn.com/abstract=2982128

See Also

See Also silentPairs.

Examples

```
## Not run:
options(np.messages = FALSE)
set.seed(34);x=sample(1:10);y=sample(2:11)
bb=bootPairs(cbind(x,y),n999=29)
probSign(bb,tau=0.476) #gives summary stats for n999 bootstrap sum computations

bb=bootPairs(airquality,n999=999);options(np.messages=FALSE)
probSign(bb,tau=0.476)#signs for n999 bootstrap sum computations

data('EuroCrime')
attach(EuroCrime)
bb=bootPairs(cbind(crim,off),n999=29) #col.1= crim causes off
#hence positive signs are more intuitively meaningful.
#note that n999=29 is too small for real problems, chosen for quickness here.
probSign(bb,tau=0.476)#signs for n999 bootstrap sum computations

## End(Not run)
```

`rank2return` *Compute the portfolio return knowing the rank of a stock in the input 'mtx'.*

Description

This function computes the return earned knowing the rank of a stock in the input mtx of stock returns. For example, mtx has p=28 Dow Jones stocks over n=169 monthly returns. Portfolio weights are assumed to be linearly declining. If maxChosen=4, the weights are 1/10, 2/10, 3/10 and 4/10, which add up to unity. These portfolio weights are assigned in reverse order in the sense that first chosen stock (choice rank =1) gets portfolio weight=4/10. The function computes return from the stocks by the 'myrank' argument.

Usage

```
rank2return(mtx, myrank, maxChosen = 0, pctChoose = 20, verbo = FALSE)
```

Arguments

`mtx`	a matrix with n rows (number of returns) p columns (number of stocks)
`myrank`	vector of p integers listing the rank of each stock, 1=best
`maxChosen`	number of stocks in the portfolio (with nonzero weights) default=0. When maxChosen=0, we let pctChoose determine the maxChosen
`pctChoose`	percent of p stocks chosen inside the portfolio, default=20
`verbo`	logical if TRUE, print, default=TRUE

Value

average return from the linearly declining portfolio implied by the myrank vector.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

See Also

`outOFsamp`

`rhs.lag2` *internal rhs.lag2*

Description

intended for internal use only

Usage

```
rhs.lag2
```

`rhs1` *internal rhs1*

Description

intended for internal use only

Usage

```
rhs1
```

`ridgek` *internal ridgek*

Description

intended for internal use only

Usage

```
ridgek
```

`rij` *internal rij*

Description

intended for internal use only

Usage

```
rij
```

`rijMrji` *internal rijMrji*

Description

intended for internal use only

Usage

```
rijMrji
```

`rji` *internal rji*

Description

intended for internal use only

Usage

```
rji
```

`rrij` *internal rrij*

Description

intended for internal use only

Usage

```
rrij
```

`rrji` *internal rrji*

Description

intended for internal use only

Usage

```
rrji
```

rstar — *Function to compute generalized correlation coefficients r*(x,y).*

Description

Uses Vinod (2015) definition of generalized (asymmetric) correlation coefficients. It requires kernel regression of x on y obtained by using the 'np' package. It also reports usual Pearson correlation coefficient r and p-value for testing the null hypothesis that (population r)=0.

Usage

```
rstar(x, y)
```

Arguments

x	Vector of data on the dependent variable
y	Vector of data on the regressor

Value

Four objects created by this function are:

corxy	r*x|y or regressing x on y
coryx	r*y|x or regressing y on x
pearson.r	Pearson's product moment correlation coefficient
pv	The p-value for testing the Pearson r

Note

This function needs the kern function which in turn needs the np package.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, https://doi.org/gffn86

Vinod, H. D. 'Matrix Algebra Topics in Statistics and Economics Using R', Chapter 4 in Handbook of Statistics: Computational Statistics with R, Vol.32, co-editors: M. B. Rao and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2014, pp. 143-176.

See Also

See Also gmcmtx0 and gmcmtxBlk.

Examples

```
x=sample(1:30);y=sample(1:30); rstar(x,y)
```

`sales2Lag` *internal sales2Lag*

Description

intended for internal use only

Usage

```
sales2Lag
```

`salesLag` *internal salesLag*

Description

intended for internal use only

Usage

```
salesLag
```

`seed` *internal seed*

Description

intended for internal use only

Usage

```
seed
```

sgn.e0	*internal sgn.e0*

Description

intended for internal use only

Usage

```
sgn.e0
```

silentMtx	*No-print kernel-causality unanimity score matrix with optional control variables*

Description

Allowing input matrix of control variables and missing data, this function produces a p by p matrix summarizing the results, where the estimated signs of stochastic dominance order values (+1, 0, –1) are weighted by `wt=c(1.2,1.1,1.05,1)` to compute an overall result for all orders of stochastic dominance by a weighted sum for the criteria Cr1 and Cr2 and added to the Cr3 estimate as: (+1, 0, –1). Final weighted index is always in the range [–3.175, 3.175]. It is converted to the more intuitive range [–100, 100].

Usage

```
silentMtx(mtx, ctrl = 0, dig = 6, wt = c(1.2, 1.1, 1.05, 1), sumwt = 4)
```

Arguments

`mtx`	The data matrix with p columns. Denote x1 as the first column which is fixed and then paired with all other columns, say: x2, x3, .., xp, one by one for the purpose of flipping with x1. p must be 2 or more
`ctrl`	data matrix for designated control variable(s) outside causal paths
`dig`	Number of digits for reporting (default `dig`=6).
`wt`	Allows user to choose a vector of four alternative weights for SD1 to SD4.
`sumwt`	Sum of weights can be changed here =4(default).

Details

The reason for slightly declining weights on the signs from SD1 to SD4 is simply that the local mean comparisons implicit in SD1 are known to be more reliable than local variance implicit in SD2, local skewness implicit in SD3 and local kurtosis implicit in SD4. Why are higher moment estimates less reliable? The higher power of the deviations from the mean needed in their computations lead to greater sampling variability. The summary results for all three criteria are reported in a vector of numbers internally called `crall`:

Value

With p columns in `mtx` argument to this function, `x1` can be paired with a total of p-1 columns (x2, x3, .., xp). Note we never flip any of the control variables with `x1`. This function produces i=1,2,..,p-1 numbers representing the summary sign, or 'sum' from the signs sg1 to sg3 associated with the three criteria: Cr1, Cr2 and Cr3. Note that sg1 and sg2 themselves are weighted signs using weighted sum of signs from four orders of stochastic dominance. In general, a positive sign in the i-th location of the 'sum' output of this function means that `x1` is the kernel cause while the variable in (i+1)-th column of `mtx` is the 'effect' or 'response' or 'endogenous.' The magnitude represents the strength (unanimity) of the evidence for a particular sign. Conversely a negative sign in the i-th location of the 'sum' output of this function means that that the first variable listed as the input to this function is the 'effect,' while the variable in (i+1)-th column of `mtx` is the exogenous kernel cause. This function is a summary of `someCPairs` allowing for control variables.

Note

The European Crime data has all three criteria correctly suggesting that high crime rate kernel causes the deployment of a large number of police officers. The command `attach(EuroCrime); silentPairs(cbind(crim,off))` returns only one number: 3.175, implying a high unanimity strength. The index 3.175 is the highest. The positive sign of the index suggests that 'crim' variable in the first column of the matrix input to this function kernel causes 'off' in the second column of the matrix argument `mtx` to this function.

Interpretation of the output matrix produced by this function is as follows. A negative index means the variable named in the column kernel-causes the variable named in the row. A positive index means the row name variable kernel-causes the column name variable. The abs(index) measures unanimity by three criteria, Cr1 to Cr3 representing the strength of evidence for the identified causal path.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

H. D. Vinod 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: 10.1080/03610918.2015.1122048

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: `https://www.ssrn.com/abstract=2982128`

See Also

See `silentPairs`.

See `someCPairs`, `some0Pairs`

Examples

```
## Not run:
options(np.messages=FALSE)
colnames(mtcars[2:ncol(mtcars)])
silentMtx(mtcars[,1:3],ctrl=mtcars[,4:5]) # mpg paired with others

## End(Not run)

options(np.messages=FALSE)
set.seed(234)
z=runif(10,2,11)# z is independently created
x=sample(1:10)+z/10 #x is somewhat indep and affected by z
y=1+2*x+3*z+rnorm(10)
w=runif(10)
x2=x;x2[4]=NA;y2=y;y2[8]=NA;w2=w;w2[4]=NA
silentMtx(mtx=cbind(x2,y2), ctrl=cbind(z,w2))
```

`silentMtx0` *Older kernel-causality unanimity score matrix with optional control variables*

Description

Allowing input matrix of control variables and missing data, this function produces a p by p matrix summarizing the results, where the estimated signs of stochastic dominance order values (+1, 0, –1) are weighted by `wt=c(1.2,1.1,1.05,1)` to compute an overall result for all orders of stochastic dominance by a weighted sum for the criteria Cr1 and Cr2 and added to the Cr3 estimate as: (+1, 0, –1). Final weighted index is always in the range [–3.175, 3.175]. It is converted to the more intuitive range [–100, 100].

Usage

```
silentMtx0(mtx, ctrl = 0, dig = 6, wt = c(1.2, 1.1, 1.05, 1), sumwt = 4)
```

Arguments

`mtx`	The data matrix with p columns. Denote x1 as the first column which is fixed and then paired with all other columns, say: x2, x3, .., xp, one by one for the purpose of flipping with x1. p must be 2 or more
`ctrl`	data matrix for designated control variable(s) outside causal paths
`dig`	Number of digits for reporting (default `dig`=6).
`wt`	Allows user to choose a vector of four alternative weights for SD1 to SD4.
`sumwt`	Sum of weights can be changed here =4(default).

Details

The reason for slightly declining weights on the signs from SD1 to SD4 is simply that the local mean comparisons implicit in SD1 are known to be more reliable than local variance implicit in SD2, local skewness implicit in SD3 and local kurtosis implicit in SD4. Why are higher moment estimates less reliable? The higher power of the deviations from the mean needed in their computations lead to greater sampling variability. The summary results for all three criteria are reported in a vector of numbers internally called `crall`:

Value

With p columns in `mtx` argument to this function, `x1` can be paired with a total of p-1 columns (x2, x3, .., xp). Note we never flip any of the control variables with `x1`. This function produces i=1,2,..,p-1 numbers representing the summary sign, or 'sum' from the signs sg1 to sg3 associated with the three criteria: Cr1, Cr2 and Cr3. Note that sg1 and sg2 themselves are weighted signs using weighted sum of signs from four orders of stochastic dominance. In general, a positive sign in the i-th location of the 'sum' output of this function means that `x1` is the kernel cause while the variable in (i+1)-th column of `mtx` is the 'effect' or 'response' or 'endogenous.' The magnitude represents the strength (unanimity) of the evidence for a particular sign. Conversely a negative sign in the i-th location of the 'sum' output of this function means that that the first variable listed as the input to this function is the 'effect,' while the variable in (i+1)-th column of `mtx` is the exogenous kernel cause. This function allows for control variables.

Note

The European Crime data has all three criteria correctly suggesting that high crime rate kernel causes the deployment of a large number of police officers. The command `attach(EuroCrime); silentPairs(cbind(crim,off))` returns only one number: 3.175, implying a high unanimity strength. The index 3.175 is the highest. The positive sign of the index suggests that 'crim' variable in the first column of the matrix input to this function kernel causes 'off' in the second column of the matrix argument `mtx` to this function.

Interpretation of the output matrix produced by this function is as follows. A negative index means the variable named in the column kernel-causes the variable named in the row. A positive index means the row name variable kernel-causes the column name variable. The abs(index) measures unanimity by three criteria, Cr1 to Cr3 representing the strength of evidence for the identified causal path.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

H. D. Vinod 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: gffn86

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: `https://www.ssrn.com/abstract=2982128`

See Also

See silentPairs0 using older Cr1 criterion based on kernel regression local gradients.

See someCPairs, some0Pairs

Examples

```
## Not run:
options(np.messages=FALSE)
colnames(mtcars[2:ncol(mtcars)])
silentMtx0(mtcars[,1:3],ctrl=mtcars[,4:5]) # mpg paired with others

## End(Not run)

## Not run:
options(np.messages=FALSE)
set.seed(234)
z=runif(10,2,11)# z is independently created
x=sample(1:10)+z/10 #x is somewhat indep and affected by z
y=1+2*x+3*z+rnorm(10)
w=runif(10)
x2=x;x2[4]=NA;y2=y;y2[8]=NA;w2=w;w2[4]=NA
silentMtx0(mtx=cbind(x2,y2), ctrl=cbind(z,w2))

## End(Not run)
```

---

silentPair2 *kernel causality (version 2) scores with control variables*

Description

This function uses flipped kernel regressions to decide causal directions. This version 2 avoids Anderson's trapezoidal approximation used in 'silenPairs.' It calls functions: decileVote, momentVote, exactSdMtx, and summaryRank after stochastic dominance is computed. It computes an average of ranks used. The column with the "choice" rank value helps in choosing the flip having the lowest Hausman-Wu (residual times RHS regressor) and secondly the lowest absolute residual. The chosen flipped regression defines the "cause" based on the variable on its right-hand side. In portfolio selection, choice rank 1 has the highest return. Here we want low residuals and low Hausman-Wu value, hence we choose choice=2 as the desirable flip.

The function develops a unanimity index regarding the particular flip (y on xi) or (xi on y) is best. A summary of all relevant signs determines the causal direction and unanimity index among three criteria. The '2' in the name of the function suggests a second implementation where exact stochastic dominance, decileVote, and momentVote algorithms are used.

Usage

```
silentPair2(mtx, ctrl = 0, dig = 6)
```

Arguments

mtx	The data matrix with p columns. Denote x1 as the first column, which is fixed in all rows of the output and then it is paired with all other columns, say: x2, x3, .., xp, one by one for the purpose of flipping with x1. p must be 2 or more
ctrl	data matrix for designated control variable(s) outside causal paths, default is ctrl=0, which means that there are no control variables used.
dig	Number of digits for reporting (default dig=6).

Value

A matrix with p columns in mtx argument to this function, x1 can be paired with a total of p-1 columns (x2, x3, .., xp). Note we never flip any of the control variables with x1. This function produces i=1,2,..,p-1 numbers representing the summary sign, or 'sum' from the signs sg1 to sg3 associated with the three criteria: Cr1, Cr2, and Cr3. Note that sg1 and sg2 themselves are weighted signs using a weighted sum of signs from four orders of stochastic dominance. In general, a positive sign in the i-th location of the 'sum' output of this function means that x1 is the kernel cause while the variable in (i+1)-th column of mtx is the 'effect' or 'response' or 'endogenous.' The magnitude represents the strength (unanimity) of the evidence for a particular sign. Conversely, a negative sign in the i-th location of the 'sum' output of this function means that the first variable listed as the input to this function is the 'effect,' while the variable in (i+1)-th column of mtx is the exogenous kernel cause.

Note

The European Crime data has all three criteria correctly suggesting that a high crime rate kernel causes the deployment of a large number of police officers. The command `attach(EuroCrime); silentPairs(cbind(crim,off))` returns only one number: 3.175, implying the highest unanimity strength index, with the positive sign suggesting 'crim' in the first column kernel causes 'off' in the second column of the argument mtx to this function.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

H. D. Vinod 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: 10.1080/03610918.2015.1122048

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: https://www.ssrn.com/abstract=2982128

See Also

See summaryRank, decileVote

See momentVote, exactSdMtx

Examples

```
## Not run:
options(np.messages=FALSE)
colnames(mtcars[2:ncol(mtcars)])
silentPair2(mtcars[,1:3],ctrl=mtcars[,4:5]) # mpg paired with others

## End(Not run)

options(np.messages=FALSE)
set.seed(234)
z=runif(10,2,11)# z is independently created
x=sample(1:10)+z/10 #x is somewhat indep and affected by z
y=1+2*x+3*z+rnorm(10)
w=runif(10)
x2=x;x2[4]=NA;y2=y;y2[8]=NA;w2=w;w2[4]=NA
silentPair2(mtx=cbind(x2,y2), ctrl=cbind(z,w2))
```

`silentPairs` *No-print kernel causality scores with control variables Hausman-Wu Criterion 1*

Description

Allowing input matrix of control variables and missing data, this function produces a 3 column matrix summarizing the results where the estimated signs of stochastic dominance order values (+1, 0, -1) are weighted by `wt=c(1.2,1.1,1.05,1)` to compute an overall result for all orders of stochastic dominance by a weighted sum for the criteria Cr1 and Cr2 and added to the Cr3 estimate as: (+1, 0, -1), always in the range [–3.175, 3.175].

Usage

```
silentPairs(mtx, ctrl = 0, dig = 6, wt = c(1.2, 1.1, 1.05, 1), sumwt = 4)
```

Arguments

`mtx`	The data matrix with p columns. Denote x1 as the first column which is fixed and then paired with all other columns, say: x2, x3, .., xp, one by one for the purpose of flipping with x1. p must be 2 or more

`ctrl`	data matrix for designated control variable(s) outside causal paths default ctrl=0 which means that there are no control variables used.
`dig`	Number of digits for reporting (default `dig`=6).
`wt`	Allows user to choose a vector of four alternative weights for SD1 to SD4.
`sumwt`	Sum of weights can be changed here =4(default).

Details

The reason for slightly declining weights on the signs from SD1 to SD4 is simply that the local mean comparisons implicit in SD1 are known to be more reliable than local variance implicit in SD2, local skewness implicit in SD3 and local kurtosis implicit in SD4. The source of slightly declining sampling unreliability of higher moments is the higher power of the deviations from the mean needed in their computations. The summary results for all three criteria are reported in a vector of numbers internally called `crall`:

Value

With p columns in `mtx` argument to this function, `x1` can be paired with a total of p-1 columns (x2, x3, .., xp). Note we never flip any of the control variables with `x1`. This function produces i=1,2,..,p-1 numbers representing the summary sign, or 'sum' from the signs sg1 to sg3 associated with the three criteria: Cr1, Cr2 and Cr3. Note that sg1 and sg2 themselves are weighted signs using weighted sum of signs from four orders of stochastic dominance. In general, a positive sign in the i-th location of the 'sum' output of this function means that `x1` is the kernel cause while the variable in (i+1)-th column of `mtx` is the 'effect' or 'response' or 'endogenous.' The magnitude represents the strength (unanimity) of the evidence for a particular sign. Conversely a negative sign in the i-th location of the 'sum' output of this function means that that the first variable listed as the input to this function is the 'effect,' while the variable in (i+1)-th column of `mtx` is the exogenous kernel cause.

Note

The European Crime data has all three criteria correctly suggesting that high crime rate kernel causes the deployment of a large number of police officers. The command `attach(EuroCrime); silentPairs(cbind(crim,off))` returns only one number: 3.175, implying the highest unanimity strength index, with the positive sign suggesting 'crim' in the first column kernel causes 'off' in the second column of the argument `mtx` to this function.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

H. D. Vinod 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: 10.1080/03610918.2015.1122048

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: https://www.ssrn.com/abstract=2982128

See Also

See `bootPairs`, `silentMtx`

See `someCPairs`, `some0Pairs`

Examples

```
## Not run:
options(np.messages=FALSE)
colnames(mtcars[2:ncol(mtcars)])
silentPairs(mtcars[,1:3],ctrl=mtcars[,4:5]) # mpg paired with others

## End(Not run)

options(np.messages=FALSE)
set.seed(234)
z=runif(10,2,11)# z is independently created
x=sample(1:10)+z/10 #x is somewhat indep and affected by z
y=1+2*x+3*z+rnorm(10)
w=runif(10)
x2=x;x2[4]=NA;y2=y;y2[8]=NA;w2=w;w2[4]=NA
silentPairs(mtx=cbind(x2,y2), ctrl=cbind(z,w2))
```

`silentPairs0` *Older version, kernel causality weighted sum allowing control variables*

Description

Allowing input matrix of control variables and missing data, this function produces a 3 column matrix summarizing the results where the estimated signs of stochastic dominance order values (+1, 0, -1) are weighted by `wt=c(1.2,1.1,1.05,1)` to compute an overall result for all orders of stochastic dominance by a weighted sum for the criteria Cr1 and Cr2 and added to the Cr3 estimate as: (+1, 0, -1), always in the range [–3.175, 3.175].

Usage

```
silentPairs0(mtx, ctrl = 0, dig = 6, wt = c(1.2, 1.1, 1.05, 1), sumwt = 4)
```

Arguments

`mtx`	The data matrix with p columns. Denote x1 as the first column which is fixed and then paired with all other columns, say: x2, x3, .., xp, one by one for the purpose of flipping with x1. p must be 2 or more

ctrl	data matrix for designated control variable(s) outside causal paths default ctrl=0 which means that there are no control variables used.
dig	Number of digits for reporting (default dig=6).
wt	Allows user to choose a vector of four alternative weights for SD1 to SD4.
sumwt	Sum of weights can be changed here =4(default).

Details

This uses an older version of the first criterion Cr1 based on absolute values of local gradients of kernel regressions, not absolute Hausman-Wu statistic (RHS variable times kernel residuals). It calls abs_stdapd and abs_stdapdC The reason for slightly declining weights on the signs from SD1 to SD4 is simply that the local mean comparisons implicit in SD1 are known to be more reliable than local variance implicit in SD2, local skewness implicit in SD3 and local kurtosis implicit in SD4. The source of slightly declining sampling unreliability of higher moments is the higher power of the deviations from the mean needed in their computations. The summary results for all three criteria are reported in a vector of numbers internally called crall:

Value

With p columns in mtx argument to this function, x1 can be paired with a total of p-1 columns (x2, x3, .., xp). Note we never flip any of the control variables with x1. This function produces i=1,2,..,p-1 numbers representing the summary sign, or 'sum' from the signs sg1 to sg3 associated with the three criteria: Cr1, Cr2 and Cr3. Note that sg1 and sg2 themselves are weighted signs using weighted sum of signs from four orders of stochastic dominance. In general, a positive sign in the i-th location of the 'sum' output of this function means that x1 is the kernel cause while the variable in (i+1)-th column of mtx is the 'effect' or 'response' or 'endogenous.' The magnitude represents the strength (unanimity) of the evidence for a particular sign. Conversely a negative sign in the i-th location of the 'sum' output of this function means that that the first variable listed as the input to this function is the 'effect,' while the variable in (i+1)-th column of mtx is the exogenous kernel cause. This function is a summary of someCPairs allowing for control variables.

Note

The European Crime data has all three criteria correctly suggesting that high crime rate kernel causes the deployment of a large number of police officers. The command attach(EuroCrime); silentPairs(cbind(crim,off)) returns only one number: 3.175, implying the highest unanimity strength index, with the positive sign suggesting 'crim' in the first column kernel causes 'off' in the second column of the argument mtx to this function.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

H. D. Vinod 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: 10.1080/03610918.2015.1122048

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: https://www.ssrn.com/abstract=2982128

See Also

See bootPairs, silentMtx

See someCPairs, some0Pairs

See silentPairs for newer version using more direct Hausman-Wu exogeneity test statistic.

Examples

```
## Not run:
options(np.messages=FALSE)
colnames(mtcars[2:ncol(mtcars)])
silentPairs0(mtcars[,1:3],ctrl=mtcars[,4:5]) # mpg paired with others

## End(Not run)

options(np.messages=FALSE)
set.seed(234)
z=runif(10,2,11)# z is independently created
x=sample(1:10)+z/10 #x is somewhat indep and affected by z
y=1+2*x+3*z+rnorm(10)
w=runif(10)
x2=x;x2[4]=NA;y2=y;y2[8]=NA;w2=w;w2[4]=NA
silentPairs0(mtx=cbind(x2,y2), ctrl=cbind(z,w2))
```

---

siPair2Blk *Block Version of silentPair2 for causality scores with control variables*

Description

Block version allows a new bandwidth (chosen by the np package) while fitting kernel regressions for each block of data. This may not be appropriate in all situations. Block size is flexible. The function develops a unanimity index regarding the particular flip (y on xi) or (xi on y) is best. Relevant signs determine the causal direction and unanimity index among three criteria. The '2' in the name of the function suggests a second implementation where exact stochastic dominance, decileVote, and momentVote are used. It avoids Anderson's trapezoidal approximation. The summary results for all three criteria are reported in a vector of numbers internally called `crall`.

Usage

```
siPair2Blk(mtx, ctrl = 0, dig = 6, blksiz = 10)
```

Arguments

mtx	The data matrix with p columns. Denote x1 as the first column, which is fixed and then paired with all other columns, say: x2, x3, .., xp, one by one flipping with x1.The number of columns, p, must be 2 or more
ctrl	data matrix for designated control variable(s) outside causal paths. The default ctrl=0 means that there are no control variables used.
dig	Number of digits for reporting (default `dig=6`).
blksiz	block size, default=10, if chosen blksiz >n, where n=rows in the matrix, then blksiz=n. That is, no blocking is done

Value

With p columns in `mtx` argument to this function, `x1` can be paired with a total of p-1 columns (x2, x3, .., xp). Note we never flip any of the control variables with `x1`. This function produces i=1,2,..,p-1 numbers representing the summary sign, or 'sum' from the signs sg1 to sg3 associated with the three criteria: Cr1, Cr2 and Cr3. Note that sg1 and sg2 themselves are weighted signs using the weighted sum of signs from four orders of stochastic dominance. In general, a positive sign in the i-th location of the 'sum' output of this function means that `x1` is the kernel cause while the variable in (i+1)-th column of `mtx` is the 'effect' or 'response' or 'endogenous.' The magnitude represents the strength (unanimity) of the evidence for a particular sign. Conversely, a negative sign in the i-th location of the 'sum' output of this function means that that the first variable listed as the input to this function is the 'effect,' while the variable in (i+1)-th column of `mtx` is the exogenous kernel cause.

Note

The European Crime data has all three criteria correctly suggesting that high crime rate kernel causes the deployment of a large number of police officers. The command `attach(EuroCrime); silentPairs(cbind(crim,off))` returns only one number: 3.175, implying the highest unanimity strength index, with the positive sign suggesting 'crim' in the first column kernel causes 'off' in the second column of the argument `mtx` to this function.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

H. D. Vinod 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: 10.1080/03610918.2015.1122048

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: https://www.ssrn.com/abstract=2982128

See Also

See `bootPairs`, `silentMtx`

See `someCPairs`, `compPortfo`

Examples

```
## Not run:
options(np.messages=FALSE)
colnames(mtcars[2:ncol(mtcars)])
siPair2Blk(mtcars[,1:3],ctrl=mtcars[,4:5]) # mpg paired with others

## End(Not run)

options(np.messages=FALSE)
set.seed(234)
z=runif(10,2,11)# z is independently created
x=sample(1:10)+z/10 #x is somewhat indep and affected by z
y=1+2*x+3*z+rnorm(10)
w=runif(10)
x2=x;x2[4]=NA;y2=y;y2[8]=NA;w2=w;w2[4]=NA
siPair2Blk(mtx=cbind(x2,y2), ctrl=cbind(z,w2))
```

`siPairsBlk` *Block Version of silentPairs for causality scores with control variables*

Description

Allowing input matrix of control variables and missing data, this function produces a 3 column matrix summarizing the results where the estimated signs of stochastic dominance order values (+1, 0, -1) are weighted by `wt=c(1.2,1.1,1.05,1)` to compute an overall result for all orders of stochastic dominance by a weighted sum for the criteria Cr1 and Cr2 and added to the Cr3 estimate as: (+1, 0, -1), always in the range [–3.175, 3.175].

Usage

```
siPairsBlk(
  mtx,
  ctrl = 0,
  dig = 6,
  blksiz = 10,
  wt = c(1.2, 1.1, 1.05, 1),
  sumwt = 4
)
```

Arguments

`mtx` The data matrix with p columns. Denote x1 as the first column which is fixed and then paired with all other columns, say: x2, x3, .., xp, one by one for the purpose of flipping with x1. p must be 2 or more

ctrl	data matrix for designated control variable(s) outside causal paths default ctrl=0 which means that there are no control variables used.
dig	Number of digits for reporting (default dig=6).
blksiz	block size, default=10, if chosen blksiz >n, where n=rows in matrix then blksiz=n. That is, no blocking is done
wt	Allows user to choose a vector of four alternative weights for SD1 to SD4.
sumwt	Sum of weights can be changed here =4(default).

Details

The reason for slightly declining weights on the signs from SD1 to SD4 is simply that the local mean comparisons implicit in SD1 are known to be more reliable than local variance implicit in SD2, local skewness implicit in SD3 and local kurtosis implicit in SD4. The source of slightly declining sampling unreliability of higher moments is the higher power of the deviations from the mean needed in their computations. The summary results for all three criteria are reported in a vector of numbers internally called `crall`:

Value

With p columns in `mtx` argument to this function, `x1` can be paired with a total of p-1 columns (x2, x3, .., xp). Note we never flip any of the control variables with `x1`. This function produces i=1,2,..,p-1 numbers representing the summary sign, or 'sum' from the signs sg1 to sg3 associated with the three criteria: Cr1, Cr2 and Cr3. Note that sg1 and sg2 themselves are weighted signs using weighted sum of signs from four orders of stochastic dominance. In general, a positive sign in the i-th location of the 'sum' output of this function means that `x1` is the kernel cause while the variable in (i+1)-th column of `mtx` is the 'effect' or 'response' or 'endogenous.' The magnitude represents the strength (unanimity) of the evidence for a particular sign. Conversely a negative sign in the i-th location of the 'sum' output of this function means that that the first variable listed as the input to this function is the 'effect,' while the variable in (i+1)-th column of `mtx` is the exogenous kernel cause.

Note

The European Crime data has all three criteria correctly suggesting that high crime rate kernel causes the deployment of a large number of police officers. The command `attach(EuroCrime); silentPairs(cbind(crim,off))` returns only one number: 3.175, implying the highest unanimity strength index, with the positive sign suggesting 'crim' in the first column kernel causes 'off' in the second column of the argument `mtx` to this function.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

H. D. Vinod 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: 10.1080/03610918.2015.1122048

Vinod, H. D. Causal Paths and Exogeneity Tests in Generalcorr Package for Air Pollution and Monetary Policy (June 6, 2017). Available at SSRN: https://www.ssrn.com/abstract=2982128

See Also

See bootPairs, silentMtx

See someCPairs, some0Pairs

Examples

```
## Not run:
options(np.messages=FALSE)
colnames(mtcars[2:ncol(mtcars)])
siPairsBlk(mtcars[,1:3],ctrl=mtcars[,4:5]) # mpg paired with others

## End(Not run)

options(np.messages=FALSE)
set.seed(234)
z=runif(10,2,11)# z is independently created
x=sample(1:10)+z/10 #x is somewhat indep and affected by z
y=1+2*x+3*z+rnorm(10)
w=runif(10)
x2=x;x2[4]=NA;y2=y;y2[8]=NA;w2=w;w2[4]=NA
siPairsBlk(mtx=cbind(x2,y2), ctrl=cbind(z,w2))
```

---

some0Pairs *Function reporting detailed kernel causality results in a 7-column matrix (uses deprecated criterion 1, no longer recommended but may be useful for second and third criterion typ=2,3)*

Description

The seven columns produced by this function summarize the results where the signs of stochastic dominance order values (+1 or -1) are weighted by `wt=c(1.2,1.1,1.05,1)` to compute an overall result for all orders of stochastic dominance by a weighted sum for the criteria Cr1 and Cr2. The weighting is obviously not needed for the third criterion Cr3.

Usage

```
some0Pairs(
  mtx,
  dig = 6,
  verbo = TRUE,
  rnam = FALSE,
```

```
  wt = c(1.2, 1.1, 1.05, 1),
  sumwt = 4
)
```

Arguments

mtx	The data matrix in the first column is paired with all others.
dig	Number of digits for reporting (default `dig=6`).
verbo	Make `verbo= TRUE` for printing detailed steps.
rnam	Make `rnam= TRUE` if cleverly created row-names are desired.
wt	Allows user to choose a vector of four alternative weights for SD1 to SD4.
sumwt	Sum of weights can be changed here =4(default).

Details

The reason for slightly declining weights on the signs from SD1 to SD4 is simply that the local mean comparisons implicit in SD1 are known to be more reliable than local variance implicit in SD2, local skewness implicit in SD3 and local kurtosis implicit in SD4. The source of slightly declining sampling unreliability of higher moments is the higher power of the deviations from the mean needed in their computations. The summary results for all three criteria are reported in one matrix called `outVote`:

typ=1 reports ('Y', 'X', 'Cause', 'SD1apd', 'SD2apd', 'SD3apd', 'SD4apd') naming variables identifying 'cause' and measures of stochastic dominance using absolute values of kernel regression gradients (or amorphous partial derivatives, apd-s) being minimized by the kernel regression algorithm while comparing the kernel regression of X on Y with that of Y on X.

typ=2 reports ('Y', 'X', 'Cause', 'SD1res', 'SD2res', 'SD3res', 'SD4res') and measures of stochastic dominance using absolute values of kernel regression residuals comparing regression of X on Y with that of Y on X.

typ=3 reports ('Y', 'X', 'Cause', 'r*x|y', 'r*y|x', 'r', 'p-val') containing generalized correlation coefficients r*, 'r' refers to. Pearson correlation coefficient p-val is the p-value for testing the significance of 'r'

Value

Prints three matrices detailing results for Cr1, Cr2 and Cr3. It also returns a grand summary matrix called 'outVote' which summarizes all three criteria. In general, a positive sign for weighted sum reported in the column 'sum' means that the first variable listed as the input to this function is the 'kernel cause.' For example, crime 'kernel causes' police officer deployment (not vice versa) is indicated by the positive sign of 'sum' (=3.175) reported for that example included in this package.

Note

The output matrix last column for 'mtcars' example has the sum of the scores by the three criteria combined. If 'sum' is positive, then variable X (mpg) is more likely to have been engineered to kernel cause the response variable Y, rather than vice versa.

The European Crime data has all three criteria correctly suggesting that high crime rate kernel causes the deployment of a large number of police officers.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: 10.1080/03610918.2015.1122048

See Also

See Also somePairs

Examples

```
## Not run:
some0Pairs(mtcars) # first variable is mpg and effect on mpg is of interest

## End(Not run)

## Not run:
data(EuroCrime)
attach(EuroCrime)
some0Pairs(cbind(crim,off))

## End(Not run)
```

---

someCPairs	*Kernel causality computations admitting control variables.*

Description

This function reports a 7-column matrix (has the older version of criterion Cr1). It allows an additional input matrix having control variables. It produces a 7-column matrix summarizing the results, where the signs of stochastic dominance order values (+1 or -1) are weighted by `wt=c(1.2,1.1,1.05,1)` to compute an overall result for all orders of stochastic dominance by a weighted sum for the criteria Cr1 and Cr2. The weighting is obviously not needed for the third criterion Cr3 which compares asymmetric correlation coefficients.

Usage

```
someCPairs(
  mtx,
  ctrl,
  dig = 6,
  verbo = TRUE,
  rnam = FALSE,
  wt = c(1.2, 1.1, 1.05, 1),
  sumwt = 4
)
```

Arguments

`mtx`	The data matrix with many columns where the first column is fixed and then paired with all other columns, one by one.
`ctrl`	data matrix for designated control variable(s) outside causal paths
`dig`	Number of digits for reporting (default `dig`=6).
`verbo`	Make `verbo= TRUE` for printing detailed steps.
`rnam`	Make `rnam= TRUE` if cleverly created rownames are desired.
`wt`	Allows user to choose a vector of four alternative weights for SD1 to SD4.
`sumwt`	Sum of weights can be changed here =4(default).

Details

The reason for slightly declining weights on the signs from SD1 to SD4 is somewhat arbitrary. The summary results for all three criteria are reported in one matrix called `outVote`:

typ=1 reports ('Y', 'X', 'Cause', 'SD1apdC', 'SD2apdC', 'SD3apdC', 'SD4apdC') naming variables identifying 'cause' and measures of stochastic dominance using absolute values of kernel regression gradients (or amorphous partial derivatives, apd-s) being minimized by the kernel regression algorithm while comparing the kernel regression of X on Y with that of Y on X. The letter C in the titles reminds presence of control variable(s).

typ=2 reports ('Y', 'X', 'Cause', 'SD1resC', 'SD2resC', 'SD3resC', 'SD4resC') and measures of stochastic dominance using absolute values of kernel regression residuals comparing regression of X on Y with that of Y on X.

typ=3 reports ('Y', 'X', 'Cause', 'r*x|yC', 'r*y|xC', 'r', 'p-val') containing generalized correlation coefficients r*, 'r' refers to. Pearson correlation coefficient p-val is the p-value for testing the significance of 'r'. The letter C in the titles reminds the presence of control variable(s).

Value

Prints three matrices detailing results for Cr1, Cr2 and Cr3. It also returns a grand summary matrix called 'outVote' which summarizes all three criteria. In general, a positive sign for weighted sum reported in the column 'sum' means that the first variable listed as the input to this function is the 'kernel cause.' This function is an extension of `some0Pairs` to allow for control variables. For example, crime 'kernel causes' police officer deployment (not vice versa) is indicated by the positive sign of 'sum' (=3.175) reported for that example included in this package.

Note

The output matrix last column for 'mtcars' example has the sum of the scores by the three criteria combined. If 'sum' is positive, then variable X (mpg) is more likely to have been engineerd to kernel cause the response variable Y, rather than vice versa.

The European Crime data has all three criteria correctly suggesting that high crime rate kernel causes the deployment of a large number of police officers.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: 10.1080/03610918.2015.1122048

See Also

See Also somePairs, some0Pairs

Examples

```
## Not run:
someCPairs(mtcars[,1:3],ctrl=mtcars[4:5]) # first variable is mpg and effect on mpg is of interest

## End(Not run)

## Not run:
set.seed(234)
z=runif(10,2,11)# z is independently created
x=sample(1:10)+z/10  #x is somewhat indep and affected by z
y=1+2*x+3*z+rnorm(10)
w=runif(10)
x2=x;x2[4]=NA;y2=y;y2[8]=NA;w2=w;w2[4]=NA
someCPairs(cbind(x2,y2), cbind(z,w2)) #yields x2 as correct cause

## End(Not run)
```

someCPairs2 — *Kernel causality computations admitting control variables reporting a 7-column matrix, version 2.*

Description

Second version of `someCPairs` also allows input matrix of control variables, produce 7 column matrix summarizing the results where the signs of stochastic dominance order values (+1 or -1) are weighted by `wt=c(1.2,1.1,1.05,1)` to compute an overall result for all orders of stochastic dominance by a weighted sum for the criteria Cr1 and Cr2. The weighting is obviously not needed for the third criterion Cr3.

Usage

```
someCPairs2(
  mtx,
  ctrl,
  dig = 6,
  verbo = TRUE,
  rnam = FALSE,
  wt = c(1.2, 1.1, 1.05, 1),
  sumwt = 4
)
```

Arguments

`mtx`	The data matrix with many columns where the first column is fixed and then paired with all other columns, one by one.
`ctrl`	data matrix for designated control variable(s) outside causal paths
`dig`	Number of digits for reporting (default `dig`=6).
`verbo`	Make `verbo= TRUE` for printing detailed steps.
`rnam`	Make `rnam= TRUE` if cleverly created rownames are desired.
`wt`	Allows user to choose a vector of four alternative weights for SD1 to SD4.
`sumwt`	Sum of weights can be changed here =4(default).

Details

The reason for slightly declining weights on the signs from SD1 to SD4 is simply that the local mean comparisons implicit in SD1 are known to be more reliable than local variance implicit in SD2, local skewness implicit in SD3 and local kurtosis implicit in SD4. The source of slightly declining sampling unreliability of higher moments is the higher power of the deviations from the mean needed in their computations. The summary results for all three criteria are reported in one matrix called `outVote`:

(typ=1) reports ('Y', 'X', 'Cause', 'SD1.rhserr', 'SD2.rhserr', 'SD3.rhserr', 'SD4.rhserr') naming variables identifying the 'cause' and measures of stochastic dominance using absolute values of kernel regression abs(RHS first regressor*residual) values comparing flipped regressions X on Y versus Y on X. The letter C in the titles reminds presence of control variable(s).

typ=2 reports ('Y', 'X', 'Cause', 'SD1resC', 'SD2resC', 'SD3resC', 'SD4resC') and measures of stochastic dominance using absolute values of kernel regression residuals comparing regression of X on Y with that of Y on X.

typ=3 reports ('Y', 'X', 'Cause', 'r*x|yC', 'r*y|xC', 'r', 'p-val') containing generalized correlation coefficients r*, 'r' refers to. Pearson correlation coefficient p-val is the p-value for testing the significance of 'r'. The letter C in the titles reminds the presence of control variable(s).

Value

Prints three matrices detailing results for Cr1, Cr2 and Cr3. It also returns a grand summary matrix called 'outVote' which summarizes all three criteria. In general, a positive sign for weighted sum reported in the column 'sum' means that the first variable listed as the input to this function is the 'kernel cause.' This function is an extension of `some0Pairs` to allow for control variables. For example, crime 'kernel causes' police officer deployment (not vice versa) is indicated by the positive sign of 'sum' (=3.175) reported for that example included in this package.

Note

The output matrix last column for 'mtcars' example has the sum of the scores by the three criteria combined. If 'sum' is positive, then variable X (mpg) is more likely to have been engineered to kernel cause the response variable Y, rather than vice versa.

The European Crime data has all three criteria correctly suggesting that high crime rate kernel causes the deployment of a large number of police officers.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: 10.1080/03610918.2015.1122048

See Also

See Also `somePairs`, `some0Pairs`

Examples

```
## Not run:
someCPairs2(mtcars[,1:3],ctrl=mtcars[4:5]) # first variable is mpg and effect on mpg is of interest

## End(Not run)
```

```
## Not run:
set.seed(234)
z=runif(10,2,11)# z is independently created
x=sample(1:10)+z/10  #x is somewhat indep and affected by z
y=1+2*x+3*z+rnorm(10)
w=runif(10)
x2=x;x2[4]=NA;y2=y;y2[8]=NA;w2=w;w2[4]=NA
someCPairs2(cbind(x2,y2), cbind(z,w2)) #yields x2 as correct cause

## End(Not run)
```

---

someMagPairs *Summary magnitudes after removing control variables in several pairs where dependent variable is fixed.*

Description

This builds on the function mag_ctrl, where the input matrix mtx has p columns. The first column is present in each of the (p-1) pairs. Its output is a matrix with four columns containing the names of variables and approximate overall estimates of the magnitudes of partial derivatives (dy/dx) and (dx/dy) for a distinct (x,y) pair in a row. The estimated overall derivatives are not always well-defined, because the real partial derivatives of nonlinear functions are generally distinct for each observation point.

Usage

```
someMagPairs(mtx, ctrl, dig = 6, verbo = TRUE)
```

Arguments

mtx	The data matrix with many columns where the first column is fixed and then paired with all other columns, one by one.
ctrl	data matrix for designated control variable(s) outside causal paths. A constant vector is not allowed as a control variable.
dig	Number of digits for reporting (default dig=6).
verbo	Make verbo= TRUE for printing detailed steps.

Details

The function mag_ctrl has kernel regressions: x~ y + ctrl and x~ ctrl to evaluate the‘incremental change’ in R-squares. Let (rxy;ctrl) denote the square root of that ‘incremental change’ after its sign is made the same as that of the Pearson correlation coefficient from cor(x,y)). One can interpret (rxy;ctrl) as a generalized partial correlation coefficient when x is regressed on y after removing the effect of control variable(s) in ctrl. It is more general than the usual partial correlation coefficient, since this one allows for nonlinear relations among variables. Next, the function computes ‘dxdy’ obtained by multiplying (rxy;ctrl) by the ratio of standard deviations, `sd(x)/sd(y)`. Now our 'dxdy' approximates the magnitude of the partial derivative (dx/dy) in a causal model where y is the cause and x is the effect. The function also reports entirely analogous 'dydx' obtained by interchanging x and y.

`someMegPairs` function runs the function `mag_ctrl` on several column pairs in a matrix input `mtx` where the first column is held fixed and all others are changed one by one, reporting two partial derivatives for each row.

Value

Table containing names of Xi and Xj and two magnitudes: (dXidXj, dXjdXi). dXidXj is the magnitude of the effect on Xi when Xi is regressed on Xj (i.e., when Xj is the cause). The analogous dXjdXi is the magnitude when Xj is regressed on Xi.

Note

This function is intended for use only after the causal path direction is already determined by various functions in this package (e.g. `someCPairs`). That is, after the researcher knows whether Xi causes Xj or vice versa. The output of this function is a matrix of 4 columns, where first columns list the names of Xi and Xj and the next two numbers in each row are dXidXj, dXjdXi, respectively, representing the magnitude of effect of one variable on the other.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: 10.1080/03610918.2015.1122048

Vinod, H. D. 'Matrix Algebra Topics in Statistics and Economics Using R', Chapter 4 in Handbook of Statistics: Computational Statistics with R, Vol.32, co-editors: M. B. Rao and C. R. Rao. New York: North Holland, Elsevier Science Publishers, 2014, pp. 143-176.

See Also

See `mag_ctrl`, `someCPairs`

Examples

```
set.seed(34);x=sample(1:10);y=1+2*x+rnorm(10);z=sample(2:11)
w=runif(10)
ss=someMagPairs(cbind(y,x,z),ctrl=w)
```

somePairs *Function reporting kernel causality results as a 7-column matrix.(deprecated)*

Description

This function lets the user choose one of three criteria to determine causal direction by setting `typ` as 1, 2 or 3. This function reports results for only one criterion at a time unlike the function `some0Pairs` which summarizes the resulting causal directions for all criteria with suitable weights. If some variables are 'control' variables, use `someCPairs`, C=control.

Usage

```
somePairs(mtx, dig = 6, verbo = FALSE, typ = 1, rnam = FALSE)
```

Arguments

`mtx`	The data matrix in the first column is paired with all others.
`dig`	Number of digits for reporting (default `dig`=6).
`verbo`	Make `verbo= TRUE` for printing detailed steps.
`typ`	Must be 1 (default), 2 or 3 for the three criteria.
`rnam`	Make `rnam= TRUE` if cleverly created rownames are desired.

Details

(typ=1) reports ('Y', 'X', 'Cause', 'SD1apd', 'SD2apd', 'SD3apd', 'SD4apd') nameing variables identifying 'cause' and measures of stochastic dominance using absolute values of kernel regression gradients comparing regresson of X on Y with that of Y on X.

(typ=2) reports ('Y', 'X', 'Cause', 'SD1res', 'SD2res', 'SD3res', 'SD4res') and measures of stochastic dominance using absolute values of kernel regression residuals comparing regresson of X on Y with that of Y on X.

(typ=3) reports ('Y', 'X', 'Cause', 'r*X|Y', 'r*Y|X', 'r', 'p-val') containing generalized correlation coefficients r*, 'r' refers to the Pearson correlation coefficient and p-val column has the p-values for testing the significance of Pearson's 'r'.

Value

A matrix containing causal identification results for one criterion. The first column of the input `mtx` having p columns is paired with (p-1) other columns The output matrix headings are self-explanatory and distinct for each criterion Cr1 to Cr3.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

H. D. Vinod 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: 10.1080/03610918.2015.1122048

See Also

The related function `some0Pairs` may be more useful, since it reports on all three criteria (by choosing typ=1,2,3) and further summarizes their results by weighting to help choose causal paths.

Examples

```
## Not run:
data(mtcars)
somePairs(mtcars)

## End(Not run)
```

`somePairs2` *Function reporting kernel causality results as a 7-column matrix, version 2.*

Description

This function is an alternative implementation of `somePairs` which also lets the user choose one of three criteria to determine causal direction by setting `typ` as 1, 2 or 3. This function reports results for only one criterion at a time unlike the function `some0Pairs` which summarizes the resulting causal directions for all criteria with suitable weights. If some variables are 'control' variables, use `someCPairs`, where notation C=control.

Usage

```
somePairs2(mtx, dig = 6, verbo = FALSE, typ = 1, rnam = FALSE)
```

Arguments

`mtx`	The data matrix in the first column is paired with all others.
`dig`	Number of digits for reporting (default `dig`=6).
`verbo`	Make `verbo= TRUE` for printing detailed steps.
`typ`	Must be 1 (default), 2 or 3 for the three criteria.
`rnam`	Make `rnam= TRUE` if cleverly created rownames are desired.

Details

(typ=1) reports ('Y', 'X', 'Cause', 'SD1.rhserr', 'SD2.rhserr', 'SD3.rhserr', 'SD4.rhserr') naming variables identifying the 'cause,' using Hausman-Wu criterion. It measures of stochastic dominance using absolute values of kernel regression abs(RHS first regressor*residual), comparing flipped regressions X on Y versus Y on X.

(typ=2) reports ('Y', 'X', 'Cause', 'SD1res', 'SD2res', 'SD3res', 'SD4res') and measures of stochastic dominance using absolute values of kernel regression residuals comparing regression of X on Y with that of Y on X.

(typ=3) reports ('Y', 'X', 'Cause', 'r*X|Y', 'r*Y|X', 'r', 'p-val') containing generalized correlation coefficients r*, 'r' refers to the Pearson correlation coefficient and p-val column has the p-values for testing the significance of Pearson's 'r'.

Value

A matrix containing causal identification results for one criterion. The first column of the input `mtx` having p columns is paired with (p-1) other columns The output matrix headings are self-explanatory and distinct for each criterion Cr1 to Cr3.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

H. D. Vinod 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: 10.1080/03610918.2015.1122048

See Also

The related function `some0Pairs` may be more useful, since it reports on all three criteria (by choosing typ=1,2,3) and further summarizes their results by weighting to help choose causal paths.

Alternative and revised function `somePairs2` implements the Cr1 (first criterion) with a direct estimate of the Hausman-Wu statistic for testing exogeneity.

Examples

```
## Not run:
data(mtcars)
somePairs2(mtcars)

## End(Not run)
```

`sort.abse0` *internal sort.abse0*

Description

intended for internal use only

Usage

```
sort.abse0
```

`sort.e0` *internal sort.e0*

Description

intended for internal use only

Usage

```
sort.e0
```

`sort_matrix` *Sort all columns of matrix x with respect to the j-th column.*

Description

This function can use the sort.list function in R. The reason for using it is that one wants the sort to carry along all columns.

Usage

```
sort_matrix(x, j)
```

Arguments

`x`	An input matrix with several columns
`j`	The column number with reference to which one wants to sort

Value

A sorted matrix

Examples

```
set.seed(30)
x=matrix(sample(1:50),ncol=5)
y=sort_matrix(x,3);y
```

---

stdres *Residuals of kernel regressions of x on y when both x and y are standardized.*

Description

1) Standardize the data to force mean zero and variance unity, 2) kernel regress x on y, with the option 'residuals = TRUE', and finally 3) compute the residuals. The standardization yields comparable residuals.

Usage

```
stdres(x, y)
```

Arguments

x	vector of data on the dependent variable
y	data on the regressors which can be a matrix

Details

The first argument is assumed to be the dependent variable. If `stdres(x,y)` is used, you are regressing x on y (not the usual y on x). The regressors can be a matrix with 2 or more columns. The missing values are suitably ignored by the standardization.

Value

kernel regression residuals are returned after standardizing the data on both sides so that the magnitudes of residuals are comparable between regression of x on y on the one hand, and the flipped regression of y on x on the other.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D. 'Generalized Correlation and Kernel Causality with Applications in Development Economics' in Communications in Statistics -Simulation and Computation, 2015, doi: 10.1080/03610918.2015.1122048

Examples

```
## Not run:
set.seed(330)
x=sample(20:50)
y=sample(20:50)
stdres(x,y)

## End(Not run)
```

`stdz_xy` *Standardize x and y vectors to achieve zero mean and unit variance.*

Description

Standardize x and y vectors to achieve zero mean and unit variance.

Usage

```
stdz_xy(x, y)
```

Arguments

`x`	Vector of data which can have NA's
`y`	Vector of data which can have NA's

Value

`stdx`	standardized values of x
`stdy`	standardized values of y

Note

This works even if there are missing x or y values.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

Examples

```
## Not run:
set.seed(30)
x=sample(20:30)
y=sample(21:31)
stdz_xy(x,y)
## End(Not run)
```

stochdom2	*Compute vectors measuring stochastic dominance of four orders.*

Description

Stochastic dominance originated as a sophisticated comparison of two distributions of stock market returns. The dominating distribution is superior in terms of local mean, variance, skewness and kurtosis respectively, representing dominance orders 1 to 4, without simply computing the four moment summary measures for the entire data. Vinod (2008, sec. 4.3) explains the details. This function uses the output of 'wtdpapb' and Anderson's algorithm subject to trapwzoidal approximation.

Usage

```
stochdom2(dj, wpa, wpb)
```

Arguments

dj	Vector of (unequal) distances of consecutive intervals defined on common support of two probability distributions being compared
wpa	Vector of the first set of (weighted) probabilities
wpb	Vector of the second set of (weighted) probabilities

Value

sd1b	Vector measuring stochastic dominance of order 1, SD1
sd2b	Vector measuring stochastic dominance of order 2, SD2
sd3b	Vector measuring stochastic dominance of order 3, SD3
sd4b	Vector measuring stochastic dominance of order 4, SD4

Note

The input to this function is the output of the function `wtdpapb`.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D.', 'Hands-On Intermediate Econometrics Using R' (2008) World Scientific Publishers: Hackensack, NJ. https://www.worldscientific.com/worldscibooks/10.1142/6895

Vinod, H. D. 'Ranking Mutual Funds Using Unconventional Utility Theory and Stochastic Dominance,' Journal of Empirical Finance Vol. 11(3) 2004, pp. 353-377.

See Also

See Also `wtdpapb`

Examples

```
## Not run:
set.seed(234);x=sample(1:30);y=sample(5:34)
w1=wtdpapb(x,y) #y should dominate x with mostly positive SDs
stochdom2(w1$dj, w1$wpa, w1$wpb)
## End(Not run)
```

---

sudoCoefParcor *Pseudo regression coefficients from generalized partial correlation coefficients, (GPCC).*

Description

This function gets the GPCCs by calling the `parcorVec` function. The pseudo regression coefficient of a kernel regression is then obtained by [GPCC*(sd dep.var)/(sd regressor)], that is, by multiplying the GPCC by the standard deviation (sd) of the dependent variable, and dividing by the sd of the regressor.

Usage

```
sudoCoefParcor(mtx, ctrl = 0, verbo = FALSE, idep = 1)
```

Arguments

`mtx`	Input data matrix with p (> or = 3) columns,
`ctrl`	Input vector or matrix of data for control variable(s), default is ctrl=0, when control variables are absent
`verbo`	Make this TRUE for detailed printing of computational steps
`idep`	The column number of the dependent variable (=1, default)

Value

A p by 1 'out' vector pseudo partial derivatives.

Note

Generalized Partial Correlation Coefficients (GPCC) allow comparison of the relative contribution of each X_j to the explanation of X_i, because GPCC are scale-free. The pseudo regression coefficient are not scale-free since they equal GPCC*(sd dep.var)/(sd regressor)

We want to get all partial correlation coefficient pairs removing other column effects. Vinod (2018) shows why one needs more than one criterion to decide the causal paths or exogeneity.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

Vinod, H. D. 'Generalized Correlations and Instantaneous Causality for Data Pairs Benchmark,' (March 8, 2015) https://www.ssrn.com/abstract=2574891

Vinod, H. D. 'Matrix Algebra Topics in Statistics and Economics Using R', Chapter 4 in Handbook of Statistics: Computational Statistics with R, Vol.32, co-editors: M. B. Rao and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2014, pp. 143-176.

Vinod, H. D. 'New Exogeneity Tests and Causal Paths,' (June 30, 2018). Available at SSRN: https://www.ssrn.com/abstract=3206096

Vinod, H. D. (2021) 'Generalized, Partial and Canonical Correlation Coefficients' Computational Economics, 59(1), 1–28.

See Also

See Also parcor_ijk.

See Also a hybrid version parcorVecH.

Examples

```
set.seed(234)
z=runif(10,2,11)# z is independently created
x=sample(1:10)+z/10  #x is partly indep and partly affected by z
y=1+2*x+3*z+rnorm(10)# y depends on x and z not vice versa
mtx=cbind(x,y,z)
sudoCoefParcor(mtx, idep=2)


## Not run:
set.seed(34);x=matrix(sample(1:600)[1:99],ncol=3)
colnames(x)=c('V1', 'v2', 'V3')#some names needed
sudoCoefParcor(x)

## End(Not run)
```

---

sudoCoefParcorH *Peudo regression coefficients from hybrid generalized partial correlation coefficients (HGPCC).*

Description

This function gets HGPCCs by calling parcorVecH function. Pseudo regression coefficient of a kernel regression is obtained by HGPCC*(sd dep.var)/(sd regressor), that is multiplying the HGPCC by the standard deviation (sd) of the dependent variable and dividing by the sd of the regressor.

Usage

```
sudoCoefParcorH(mtx, ctrl = 0, verbo = FALSE, idep = 1)
```

Arguments

mtx	Input data matrix with p (> or = 3) columns,
ctrl	Input vector or matrix of data for control variable(s), default is ctrl=0 when control variables are absent
verbo	Make this TRUE for detailed printing of computational steps
idep	The column number of the dependent variable (=1, default)

Value

A p by 1 'out' vector pseudo partial derivatives

Note

Hybrid Generalized Partial Correlation Coefficients (HGPCC) allow comparison of the relative contribution of each X_j to the explanation of X_i, because GPCC are scale-free. Hybrid refers to use of OLS residuals. Now pseudo hybrid regr coeff are HGPCC*(sd dep.var)/(sd regressor)

We want to get all partial correlation coefficient pairs removing other column effects. Vinod (2018) shows why one needs more than one criterion to decide the causal paths or exogeneity.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

References

Vinod, H. D. 'Generalized Correlations and Instantaneous Causality for Data Pairs Benchmark,' (March 8, 2015) https://www.ssrn.com/abstract=2574891

Vinod, H. D. 'Matrix Algebra Topics in Statistics and Economics Using R', Chapter 4 in Handbook of Statistics: Computational Statistics with R, Vol.32, co-editors: M. B. Rao and C.R. Rao. New York: North Holland, Elsevier Science Publishers, 2014, pp. 143-176.

Vinod, H. D. 'New Exogeneity Tests and Causal Paths,' (June 30, 2018). Available at SSRN: https://www.ssrn.com/abstract=3206096

Vinod, H. D. (2021) 'Generalized, Partial and Canonical Correlation Coefficients' Computational Economics, 59(1), 1–28.

See Also

See Also parcor_ijk.

See Also a hybrid version parcorVecH.

Examples

```
set.seed(234)
z=runif(10,2,11)# z is independently created
x=sample(1:10)+z/10  #x is partly indep and partly affected by z
y=1+2*x+3*z+rnorm(10)# y depends on x and z not vice versa
mtx=cbind(x,y,z)
```

```
sudoCoefParcor(mtx, idep=2)


## Not run:
set.seed(34);x=matrix(sample(1:600)[1:99],ncol=3)
colnames(x)=c('V1', 'v2', 'V3')#some names needed
sudoCoefParcorH(x)

## End(Not run)
```

---

summaryRank *Compute ranks of rows of matrix and summarize them into a choice suggestion.*

Description

This function allows getting out the choice (of a column representing a stock) from four rows of numbers quantifying the four orders of exact stochastic dominance comparisons. If the last or 10-th row for “choice" has 1 then the stock representing that column is to be chosen. That is it should get the largest (portfolio) weight. If the original matrix row names are SD1 to SD4, the same names are repeated for the extra rows representing their ranks. The row name for “sum of ranks" is sumRanks. Finally, the ranks associated with sumRanks provide the row named choice along the bottom (10-th) row of the output matrix called "out."

Usage

```
summaryRank(mtx)
```

Arguments

mtx	matrix to be ranked by row and summarized

Value

a matrix called ‘out’ having 10 rows and p columns (p=No.of stocks). Row Numbers 1 to 4 have SD1 to SD4 evaluation of areas over ECDFs. There are 6 more rows. Row No.5= SD1 ranks, Row No.6= SD2 ranks, Row No.7= SD3 ranks, Row No.8= SD4 ranks Row No.9= sum of the ranks in earlier four rows for ranks of SD1 to SD4 Row No.10= choice rank based on all four (SD1 to SD4) added together Thus, the tenth row yields choice priority number for each stock (asset) after combining the all four criteria.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

See Also

exactSdMtx

symmze — *Replace asymmetric matrix by max of abs values of [i,j] or [j,i] elements.*

Description

It is useful in symmetrizing the gmcmtx0 matrix containing a non-symmetric generalized correlation matrix.

Usage

```
symmze(mtx)
```

Arguments

mtx	non-symmetric matrix

Value

mtx2	replace [i,j] and [j,i] by the max of absolute values with common sign

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY.

Examples

```
## Not run:
example
mtx=matrix(1:16,nrow=4)
symmze(mtx)

## End(Not run)#'
```

wtdpapb — *Creates input for the stochastic dominance function stochdom2*

Description

Stochastic dominance is a sophisticated comparison of two distributions of stock market returns. The dominating distribution is superior in terms of mean, variance, skewness and kurtosis respectively, representing dominance orders 1 to 4, without directly computing four moments. Vinod(2008) sec. 4.3 explains the details. The ‘wtdpapb’ function creates the input for stochdom2 which in turn computes the stochastic dominance. See Vinod (2004) for details about quantitative stochastic dominance.

Usage

```
wtdpapb(xa, xb)
```

Arguments

xa	Vector of (excess) returns for the first investment option A or values of any random variable being compared to another.
xb	Vector of returns for the second option B

Value

wpa	Weighted vector of probabilities for option A
wpb	Weighted vector of probabilities for option B
dj	Vector of interval widths (distances) when both sets of data are forced on a common support

Note

Function is needed before using stochastic dominance

In Vinod (2008) where the purpose of `wtdpapb` is to map from standard ‘expected utility theory’ weights to more sophisticated ‘non-expected utility theory’ weights using Prelec’s (1998, Econometrica, p. 497) method. These weights are not needed here. Hence we provide the function `prelec2` which does not use Prelec weights at all, thereby simplifying and speeding up the R code provided in Vinod (2008). This function avoids sophisticated ‘non-expected’ utility theory which incorporates commonly observed human behavior favoring loss aversion and other anomalies inconsistent with precepts of the expected utility theory. Such weighting is not needed for our application.

Author(s)

Prof. H. D. Vinod, Economics Dept., Fordham University, NY

References

Vinod, H. D.’, ’Hands-On Intermediate Econometrics Using R’ (2008) World Scientific Publishers: Hackensack, NJ. https://www.worldscientific.com/worldscibooks/10.1142/6895

Vinod, H. D. ’Ranking Mutual Funds Using Unconventional Utility Theory and Stochastic Dominance,’ Journal of Empirical Finance Vol. 11(3) 2004, pp. 353-377.

See Also

See Also `stochdom2`

Examples

```
 ## Not run:
 set.seed(234);x=sample(1:30);y=sample(5:34)
 wtdpapb(x,y)
## End(Not run)
```

Index