

Package ‘sem’

January 2, 2012

Version 2.1-1

Date 2010/11/20

Title Structural Equation Models

Author John Fox <jfox@mcmaster.ca> and Jarrett Byrnes <byrnes@nceas.ucsb.edu>, with contributions from Michael Culbertson, Michael Friendly, Adam Kramer, and Georges Monette.

Maintainer John Fox <jfox@mcmaster.ca>

Depends R (>= 2.14.0), stats, MASS, matrixcalc

Suggests boot, tcltk, polycor

LazyLoad yes

LazyData yes

ByteCompile yes

Description This package contains functions for fitting general linear structural equation models (with observed and unobserved variables) using the RAM approach, and for fitting structural equations in observed-variable models by two-stage least squares.

License GPL (>= 2)

URL <http://www.r-project.org>, <http://socserv.socsci.mcmaster.ca/jfox/>

Repository CRAN

Repository/R-Forge/Project sem

Repository/R-Forge/Revision 80

Date/Publication 2011-11-23 16:04:51

R topics documented:

| | |
|--------------------------|-----------|
| Bollen | 2 |
| bootSem | 3 |
| CNES | 6 |
| effects.sem | 7 |
| fscores | 9 |
| information.criteria | 11 |
| Klein | 12 |
| Kmenta | 13 |
| ML.methods | 14 |
| modIndices | 15 |
| objective.functions | 18 |
| optimizers | 19 |
| pathDiagram | 20 |
| ram | 23 |
| rawMoments | 24 |
| readMoments | 25 |
| residuals.sem | 26 |
| sem | 29 |
| sem-deprecated | 47 |
| specifyModel | 48 |
| standardizedCoefficients | 54 |
| tsls | 56 |
| Index | 59 |

Bollen

*Bollen's Data on Industrialization and Political Democracy***Description**

This data set includes four measures of democracy at two points in time, 1960 and 1965, and three measures of industrialization in 1960, for 75 developing countries.

Usage

Bollen

Format

A data frame with 75 observations on the following 11 variables.

- y1 freedom of the press, 1960
- y2 freedom of political opposition, 1960
- y3 fairness of elections, 1960
- y4 effectiveness of elected legislature, 1960

y5 freedom of the press, 1965
y6 freedom of political opposition, 1965
y7 fairness of elections, 1965
y8 effectiveness of elected legislature, 1965
x1 GNP per capita, 1960
x2 energy consumption per capita, 1960
x3 percentage of labor force in industry, 1960

Details

Variables y1 through y4 are intended to be indicators of the latent variable *political democracy in 1960*; y5 through y8 indicators of *political democracy in 1965*; and x1 through x3 indicators of *industrialization in 1960*.

Source

personal communication from Ken Bollen.

References

Bollen, K. A. (1989) *Structural Equations With Latent Variables*. Wiley.

| | |
|---------|--|
| bootSem | <i>Bootstrap a Structural Equation Model</i> |
|---------|--|

Description

Bootstraps a structural equation model in an sem object (as returned by the [sem](#) function).

Usage

```
bootSem(model, R=100, cov=cov, data=model$data, ...)

## S3 method for class 'bootsem'
print(x, digits=getOption("digits"), ...)

## S3 method for class 'bootsem'
summary(object,
  type=c("perc", "bca", "norm", "basic", "none"), level=0.95, ...)
```

Arguments

| | |
|-----------|--|
| model | an sem object, produced by the sem function. |
| R | the number of bootstrap replications; the default is 100, which should be enough for computing standard errors, but not confidence intervals (except for the normal-theory intervals). |
| cov | a function to compute the input covariance or moment matrix; the default is <code>cov</code> . Use <code>cor</code> if the model is fit to the correlation matrix. The function <code>hetcor</code> in the <code>polycor</code> package will compute product-moment, polychoric, and polyserial correlations among mixed continuous and ordinal variables (see the example below for an illustration). |
| data | a data frame or numeric matrix containing the data to which the model was fit; note that the original observations are required, not just the covariance matrix of the observed variables in the model. The default is the data set stored in the sem object, which will be present only if the model was fit to a data set rather than a covariance or moment matrix. |
| x, object | an object of class <code>bootsem</code> . |
| digits | controls the number of digits to print. |
| type | type of bootstrapped confidence intervals to compute; the default is "perc" (percentile); see <code>boot.ci</code> for details. |
| level | level for confidence intervals; default is 0.95. |
| ... | in <code>bootSem</code> , arguments to be passed to <code>sem</code> ; otherwise ignored. |

Details

`bootSem` implements the nonparametric bootstrap, assuming an independent random sample. Convergence failures in the bootstrap resamples are discarded (and a warning printed); 10 consecutive convergence failures result in an error. You can use the `boot` function in the `boot` package for more complex sampling schemes and additional options.

Bootstrapping is implemented by resampling the observations in `data`, recalculating the input covariance matrix with `cov`, and refitting the model with `sem`, using the parameter estimates from the original sample as start-values.

Warning: the bootstrapping process can be very time-consuming.

Value

`bootSem` returns an object of class `bootsem`, which inherits from class `boot`, supported by the `boot` package. The returned object contains the following components:

| | |
|-----------|---|
| t0 | the estimated parameters in the model fit to the original data set. |
| t | a matrix containing the bootstrapped estimates, one bootstrap replication per row. |
| data | the data frame containing the data to which the model was fit. |
| seed | the value of <code>.Random.seed</code> when <code>bootSem</code> was called. |
| statistic | the function used to produce the bootstrap replications; this is always the local function <code>refit</code> from <code>bootSem</code> . |

| | |
|--------|--|
| sim | always set to "ordinary"; see the documentation for the boot function. |
| stype | always set to "i"; see the documentation for the boot function. |
| call | the call of the bootSem function. |
| strata | always a vector of ones. |

Author(s)

John Fox <jfox@mcmaster.ca>

References

Davison, A. C., and Hinkley, D. V. (1997) *Bootstrap Methods and their Application*. Cambridge.
 Efron, B., and Tibshirani, R. J. (1993) *An Introduction to the Bootstrap*. Chapman and Hall.

See Also

[boot](#), [sem](#)

Examples

```
## Not run:

# A simple confirmatory factor-analysis model using polychoric correlations.
# The polycor package is required for the hetcor function.

library(polycor)

# The following function returns correlations computed by hetcor,
# and is used for the bootstrapping.

hcor <- function(data) hetcor(data, std.err=FALSE)$correlations

model.cnes <- specifyModel()
F -> MBSA2, lam1
F -> MBSA7, lam2
F -> MBSA8, lam3
F -> MBSA9, lam4
F <-> F, NA, 1
MBSA2 <-> MBSA2, the1
MBSA7 <-> MBSA7, the2
MBSA8 <-> MBSA8, the3
MBSA9 <-> MBSA9, the4

R.cnes <- hcor(CNES)

sem.cnes <- sem(model.cnes, R.cnes, N=1529)
summary(sem.cnes)

# Note: this can take a couple of minutes:

system.time(boot.cnes <- bootSem(sem.cnes, R=100, cov=hcor, data=CNES))
```

```
summary(boot.cnes, type="norm")
# cf., standard errors to those computed by summary(sem.cnes)

## End(Not run)
```

CNES

Variables from the 1997 Canadian National Election Study

Description

These variables are from the mailback questionnaire to the 1997 Canadian National Election Study, and are intended to tap attitude towards “traditional values.”

Usage

CNES

Format

A data frame with 1529 observations on the following 4 variables.

MBSA2 an ordered factor with levels StronglyDisagree, Disagree, Agree, and StronglyAgree, in response to the statement, “We should be more tolerant of people who choose to live according to their own standards, even if they are very different from our own.”

MBSA7 an ordered factor with levels StronglyDisagree, Disagree, Agree, and StronglyAgree, in response to the statement, “Newer lifestyles are contributing to the breakdown of our society.”

MBSA8 an ordered factor with levels StronglyDisagree, Disagree, Agree, and StronglyAgree, in response to the statement, “The world is always changing and we should adapt our view of moral behaviour to these changes.”

MBSA9 an ordered factor with levels StronglyDisagree, Disagree, Agree, and StronglyAgree, in response to the statement, “This country would have many fewer problems if there were more emphasis on traditional family values.”

Source

York University Institute for Social Research.

Description

The sem method for the standard generic function effects computes total, direct, and indirect effects for a fitted structural equation model according to the method described in Fox (1980).

Usage

```
## S3 method for class 'sem'
effects(object, ...)

## S3 method for class 'semeffects'
print(x, digits = getOption("digits"), ...)
```

Arguments

| | |
|--------|---|
| object | a fitted structural-equation model object produced by the sem function. |
| x | an object of class semeffects, produced by effect.sem. |
| digits | digits to print. |
| ... | not used. |

Value

effect.sem returns an object of class semeffects with Total, Direct, and Indirect elements.

Author(s)

John Fox <jfox@mcmaster.ca>

References

Fox, J. (1980) Effect analysis in structural equation models: Extensions and simplified methods of computation. *Sociological Methods and Research* **9**, 3–28.

See Also

[sem](#)

Examples

```
## Not run:

# These examples are from Fox (1980)

# In the first pair of examples, readMoments() and specifyModel() read from the
# input stream. These examples cannot be executed via example() but can be entered
```

```

# at the command prompt. The Blau and Duncan example is repeated using file input;
# this example can be executed via example().

# The recursive Blau and Duncan basic stratification model:
# x1 is father's education, x2 father's SES, y3 respondent's education,
# y4 SES of respondent's first job, y5 respondent's SES in 1962

R.bd <- readMoments(names=c("x1", "x2", "y3", "y4", "y5"))
1
.516 1
.453 .438 1
.332 .417 .538 1
.322 .405 .596 .541 1

mod.bd <- specifyModel()
y3 <- x1, gam31
y3 <- x2, gam32
y4 <- x2, gam42
y4 <- y3, beta43
y5 <- x2, gam52
y5 <- y3, beta53
y5 <- y4, beta54

sem.bd <- sem(mod.bd, R.bd, N=20700, fixed.x=c("x1", "x2"))
summary(sem.bd)
effects(sem.bd)

# The nonrecursive Duncan, Haller, and Portes peer-influences model for observed variables:

R.DHP <- readMoments(diag=FALSE, names=c("R0ccAsp", "REdAsp", "F0ccAsp",
"FEdAsp", "RParAsp", "RIQ", "RSES", "FSES", "FIQ", "FParAsp"))
.6247
.3269 .3669
.4216 .3275 .6404
.2137 .2742 .1124 .0839
.4105 .4043 .2903 .2598 .1839
.3240 .4047 .3054 .2786 .0489 .2220
.2930 .2407 .4105 .3607 .0186 .1861 .2707
.2995 .2863 .5191 .5007 .0782 .3355 .2302 .2950
.0760 .0702 .2784 .1988 .1147 .1021 .0931 -.0438 .2087

model.dhp <- specifyModel()
RIQ -> R0ccAsp, gam51, NA
RSES -> R0ccAsp, gam52, NA
FSES -> F0ccAsp, gam63, NA
FIQ -> F0ccAsp, gam64, NA
F0ccAsp -> R0ccAsp, beta56, NA
R0ccAsp -> F0ccAsp, beta65, NA
R0ccAsp <-> R0ccAsp, ps55, NA
F0ccAsp <-> F0ccAsp, ps66, NA
R0ccAsp <-> F0ccAsp, ps56, NA

```

```

# Note: The following generates a warning because not all of the variables
#       in the correlation matrix are used
sem.dhp <- sem(model.dhp, R.DHP, 329,
              fixed.x=c('RIQ', 'RSES', 'FSES', 'FIQ'))
summary(sem.dhp)
effects(sem.dhp)

## End(Not run)

## the following example may be executed via example()

etc <- file.path(.path.package(package="sem")[1], "etc") # path to data and model files

# The recursive Blau and Duncan basic stratification model:
# x1 is father's education, x2 father's SES, y3 respondent's education,
# y4 SES of respondent's first job, y5 respondent's SES in 1962

(R.bd <- readMoments(file=file.path(etc, "R-Blau-Duncan.txt"),
                    names=c("x1", "x2", "y3", "y4", "y5")))
(mod.bd <- specifyModel(file=file.path(etc, "model-Blau-Duncan.txt")))
sem.bd <- sem(mod.bd, R.bd, N=20700, fixed.x=c("x1", "x2"))
summary(sem.bd)
effects(sem.bd)

```

fscores

Factor Scores for Latent Variables

Description

Calculate factor scores or factor-score coefficients for the latent variables in a structural-equation model.

Usage

```

## S3 method for class 'sem'
fscores(model, data=model$data, center=TRUE, scale=FALSE, ...)

```

Arguments

| | |
|--------|---|
| model | an object of class "sem", produced by the sem function. |
| data | an optional numeric data frame or matrix containing the observed variables in the model; if not NULL, the estimated factor scores are returned; if NULL, the factor-score <i>coefficients</i> are returned. The default is the data element of model, which is non-NULL if the model was fit to a data set rather than a covariance or moment matrix. |
| center | if TRUE, the default, the means of the observed variables are subtracted prior to computing factor scores. One would normally use this option if the model is estimated from a covariance or correlation matrix among the observed variables. |

scale if TRUE, the possibly centered variables are divided by their root-mean-squares; the default is FALSE. One would normally use this option if the model is estimated from a correlation matrix among the observed variables. Centering and scaling are performed by the `scale` function.

... arguments to pass down.

Details

Factor-score coefficients are computed by the “regression” method as $B = C^{-1}C^*$, where C is the model-implied covariance or moment matrix among the observed variables and C^* is the matrix of model-implied covariances or moments between the observed and latent variables.

Value

Either a matrix of estimated factor scores (if the data argument is supplied) or a matrix of factor-score coefficients (otherwise).

Author(s)

John Fox <jfox@mcmaster.ca>

References

Bollen, K. A. (1989) *Structural Equations With Latent Variables*. Wiley.

See Also

[sem](#), [scale](#)

Examples

```
# In the first example, readMoments() and specifyModel() read from the
# input stream. This example cannot be executed via example() but can be entered
# at the command prompt. The example is repeated using file input;
# this example can be executed via example().
## Not run:
```

```
S.wh <- readMoments(names=c('Anomia67', 'Powerless67', 'Anomia71',
                           'Powerless71', 'Education', 'SEI'))
```

```
11.834
 6.947  9.364
 6.819  5.091 12.532
 4.783  5.028  7.495  9.986
-3.839 -3.889 -3.841 -3.625  9.610
-21.899 -18.831 -21.748 -18.775 35.522 450.288
```

```
# This model in the SAS manual for PROC CALIS
```

```
model.wh.1 <- specifyModel()
  Alienation67 -> Anomia67,    NA,    1
```

```

Alienation67 -> Powerless67, NA, 0.833
Alienation71 -> Anomia71, NA, 1
Alienation71 -> Powerless71, NA, 0.833
SES -> Education, NA, 1
SES -> SEI, lamb, NA
SES -> Alienation67, gam1, NA
Alienation67 -> Alienation71, beta, NA
SES -> Alienation71, gam2, NA
Anomia67 <-> Anomia67, the1, NA
Anomia71 <-> Anomia71, the1, NA
Powerless67 <-> Powerless67, the2, NA
Powerless71 <-> Powerless71, the2, NA
Education <-> Education, the3, NA
SEI <-> SEI, the4, NA
Anomia67 <-> Anomia71, the5, NA
Powerless67 <-> Powerless71, the5, NA
Alienation67 <-> Alienation67, psi1, NA
Alienation71 <-> Alienation71, psi2, NA
SES <-> SES, phi, NA

```

```
sem.wh.1 <- sem(model.wh.1, S.wh, 932)
```

```
fscores(sem.wh.1)
```

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

```
# The following example can be executed via example():
```

```
etc <- file.path(.path.package(package="sem")[1], "etc") # path to data and model files
```

```
(S.wh <- readMoments(file=file.path(etc, "S-Wheaton.txt"),
```

```
names=c('Anomia67','Powerless67','Anomia71',
```

```
'Powerless71','Education','SEI'))
```

```
(model.wh.1 <- specifyModel(file=file.path(etc, "model-Wheaton-1.txt")))
```

```
(sem.wh.1 <- sem(model.wh.1, S.wh, 932))
```

```
fscores(sem.wh.1)
```

information.criteria *Additional Information Criteria*

Description

These are generic functions for computing, respectively, the AICc (second-order corrected Akaike Information Criterion) and CAIC (consistent Akaike Information Criterion).

Usage

```
AICc(object, ...)
```

```
CAIC(object, ...)
```

Arguments

`object` an object for which an appropriate AICc or CAIC method exists.
`...` possible additional arguments for methods.

Author(s)

Jarrett Byrnes and John Fox <jfox@mcmaster.ca>

References

Burnham, K. P., and Anderson, D. R. (1998) *Model Selection and Inference: A Practical Information-Theoretical Approach*. Springer.
 Bozdogan, H. (1987) Model selection and Akaike's information criterion (AIC). *Psychometrika* bold52, 345–370.

See Also

[AICc.objectiveML](#), [CAIC.objectiveML](#)

Klein

Klein's Data on the U. S. Economy

Description

Data for Klein's (1950) simple econometric model of the U. S. economy.
 The Klein data frame has 22 rows and 10 columns.

Usage

Klein

Format

This data frame contains the following columns:

Year 1921–1941

C consumption.

P private profits.

Wp private wages.

I investment.

K.lag capital stock, lagged one year.

X equilibrium demand.

Wg government wages.

G government non-wage spending.

T indirect business taxes and net exports.

Source

Greene, W. H. (1993) *Econometric Analysis, Second Edition*. Macmillan.

References

Klein, L. (1950) *Economic Fluctuations in the United States 1921–1941*. Wiley.

Examples

```
Klein$P.lag <- c(NA, Klein$P[-22])
Klein$X.lag <- c(NA, Klein$X[-22])

summary(tsls(C ~ P + P.lag + I(Wp + Wg),
  instruments=~1 + G + T + Wg + I(Year - 1931) + K.lag + P.lag + X.lag,
  data=Klein))

summary(tsls(I ~ P + P.lag + K.lag,
  instruments=~1 + G + T + Wg + I(Year - 1931) + K.lag + P.lag + X.lag,
  data=Klein))

summary(tsls(Wp ~ X + X.lag + I(Year - 1931),
  instruments=~1 + G + T + Wg + I(Year - 1931) + K.lag + P.lag + X.lag,
  data=Klein))
```

Kmenta

Partly Artificial Data on the U. S. Economy

Description

These are partly contrived data from Kmenta (1986), constructed to illustrate estimation of a simultaneous-equation model.

The Kmenta data frame has 20 rows and 5 columns.

Usage

Kmenta

Format

This data frame contains the following columns:

Q food consumption per capita.

P ratio of food prices to general consumer prices.

D disposable income in constant dollars.

F ratio of preceding year's prices received by farmers to general consumer prices.

A time in years.

Details

The exogenous variables D, F, and A are based on real data; the endogenous variables P and Q were generated by simulation.

Source

Kmenta, J. (1986) *Elements of Econometrics, Second Edition*, Macmillan.

| | |
|------------|---|
| ML.methods | <i>Methods for sem Objects Fit Using the objectiveML and objectiveGLS Objective Functions</i> |
|------------|---|

Description

These functions are for objects fit by `sem` using the `objectiveML` (multivariate-normal full-information maximum-likelihood) and `objectiveGLS` (generalized least squares) objective functions.

Usage

```
## S3 method for class 'objectiveML'
anova(object, model.2, robust=FALSE, ...)

## S3 method for class 'objectiveML'
logLik(object, ...)
## S3 method for class 'objectiveML'
deviance(object, ...)
## S3 method for class 'objectiveML'
AIC(object, ..., k)
## S3 method for class 'objectiveML'
AICc(object, ...)
## S3 method for class 'objectiveML'
BIC(object, ...)
## S3 method for class 'objectiveML'
CAIC(object, ...)

## S3 method for class 'objectiveML'
print(x, ...)
## S3 method for class 'objectiveGLS'
print(x, ...)
## S3 method for class 'objectiveML'
summary(object, digits=5, conf.level=.90, robust=FALSE, analytic.se=object$t <= 100, ...)
## S3 method for class 'objectiveGLS'
summary(object, ...)
```

Arguments

object, model.2, x
 an object inheriting from class objectiveML or objectiveGLS.

robust
 if TRUE, compute robust standard errors or test.

k, ...
 ignored.

digits
 digits to be printed.

conf.level
 level for confidence interval for the RMSEA index (default is .9).

analytic.se
 use analytic (as opposed to numeric) coefficient standard errors; default is TRUE
 is there are no more than 100 parameters in the model and FALSE otherwise.

Author(s)

John Fox <jfox@mcmaster.ca> and Jarrett Byrnes

References

See [sem](#).

See Also

[sem](#), [objective.functions](#), [modIndices.objectiveML](#)

 modIndices

Modification Indices for Structural Equation Models

Description

mod.indices calculates modification indices (score tests) and estimated parameter changes for the fixed and constrained parameters in a structural equation model fit by multinormal maximum likelihood.

Usage

```
## S3 method for class 'objectiveML'
modIndices(model, ...)

## S3 method for class 'modIndices'
print(x, n.largest=5, ...)

## S3 method for class 'modIndices'
summary(object, round=2,
  print.matrices=c("both", "par.change", "mod.indices"), ...)
```

Arguments

| | |
|----------------|--|
| model | an object of class <code>sem</code> , produced by the <code>sem</code> function. |
| object, x | an object of class <code>modIndices</code> , produced by the <code>modIndices</code> function. |
| n.largest | number of modification indices to print in each of the A and P matrices of the RAM model. |
| round | number of places to the right of the decimal point in printing modification indices. |
| print.matrices | which matrices to print: estimated changes in the fixed parameters, modification indices, or both (the default). |
| ... | arguments to be passed down. |

Details

Modification indices are one-df chi-square score (“Lagrange-multiplier”) test statistics for the fixed and constrained parameters in a structural equation model. They may be regarded as an estimate of the improvement in the likelihood-ratio chi-square statistic for the model if the corresponding parameter is respecified as a free parameter. The `modIndices` function also estimates the change in the value of a fixed or constrained parameter if the parameter is respecified as free. When several parameters are set equal, modification indices and estimated changes are given for all but the first. Modification indices and estimated parameter changes for currently free parameters are given as NA.

The method employed is described in Saris, Satorra, and Sorbom (1987) and Sorbom (1989).

Value

`modIndices` returns an object of class `modIndices` with the following elements:

| | |
|--------------------|---|
| <code>mod.A</code> | modification indices for the elements of the A matrix. |
| <code>mod.P</code> | modification indices for the elements of the P matrix. |
| <code>par.A</code> | estimated parameter changes for the elements of the A matrix. |
| <code>par.P</code> | estimated parameter changes for the elements of the P matrix. |

Author(s)

John Fox <jfox@mcmaster.ca> and Michael Culbertson

References

Sarris, W. E., Satorra, A., and Sorbom, D. (1987) The detection and correction of specification errors in structural equation models. Pp. 105–129 in Clogg, C. C. (ed.), *Sociological Methodology 1987*. American Sociological Association.

Sorbom, D. (1989) Model modification. *Psychometrika* **54**, 371–384.

See Also

[sem](#)

Examples

```

# In the first example, readMoments() and specifyModel() read from the
# input stream. This example cannot be executed via example() but can be entered
# at the command prompt. The example is repeated using file input;
# this example can be executed via example().
## Not run:
# This example is adapted from the SAS manual

S.wh <- readMoments(names=c('Anomia67','Powerless67','Anomia71',
                           'Powerless71','Education','SEI'))

11.834
 6.947   9.364
 6.819   5.091  12.532
 4.783   5.028   7.495   9.986
-3.839  -3.889  -3.841  -3.625   9.610
-21.899 -18.831 -21.748 -18.775  35.522  450.288

model.wh <- specifyModel()
  Alienation67  -> Anomia67,      NA,  1
  Alienation67  -> Powerless67,   NA,  0.833
  Alienation71  -> Anomia71,      NA,  1
  Alienation71  -> Powerless71,   NA,  0.833
  SES           -> Education,     NA,  1
  SES           -> SEI,           lamb, NA
  SES           -> Alienation67,  gam1, NA
  Alienation67  -> Alienation71,  beta, NA
  SES           -> Alienation71,  gam2, NA
  Anomia67      <-> Anomia67,      the1, NA
  Anomia71      <-> Anomia71,      the1, NA
  Powerless67   <-> Powerless67,   the2, NA
  Powerless71   <-> Powerless71,   the2, NA
  Education     <-> Education,     the3, NA
  SEI           <-> SEI,           the4, NA
  Anomia67      <-> Anomia71,      the5, NA
  Powerless67   <-> Powerless71,   the5, NA
  Alienation67  <-> Alienation67,  psi1, NA
  Alienation71  <-> Alienation71,  psi2, NA
  SES           <-> SES,           phi,  NA

sem.wh <- sem(model.wh, S.wh, 932)
modIndices(sem.wh)

## End(Not run)

# The following example can be executed via example():

etc <- file.path(.path.package(package="sem")[1], "etc") # path to data and model files

(S.wh <- readMoments(file=file.path(etc, "S-Wheaton.txt"),
names=c('Anomia67','Powerless67','Anomia71',
        'Powerless71','Education','SEI'))))
(model.wh <- specifyModel(file=file.path(etc, "model-Wheaton-1.txt"))))

```

```
(sem.wh <- sem(model.wh, S.wh, 932))  
modIndices(sem.wh)
```

objective.functions *sem Objective-Function Builders*

Description

These functions return objective functions suitable for use with optimizers called by [sem](#). The user would not normally call these functions directly, but rather supply one of them in the `objective` argument to `sem`. Users may also write their own objective functions.

Usage

```
objectiveML(gradient=TRUE)  
objectiveGLS(gradient=FALSE)
```

Arguments

`gradient` If TRUE, the object that's returned includes a function for computing an analytic gradient; there is at present no analytic gradient available for `objectiveGLS`.

Value

These functions return an object of class "semObjective", with up to two elements:

`objective` an objective function.
`gradient` a gradient function.

Author(s)

John Fox <jfox@mcmaster.ca>

References

See [sem](#).

See Also

[sem](#), [optimizers](#)

 optimizers

sem Optimizers

Description

These functions call optimizers (`nlm`, `optim`, or `nlminb`), to fit structural equation models, and are called by the `sem` function. The user would not normally call these functions directly, but rather supply one of them in the `optimizer` argument to `sem`. Users may also write their own optimizer functions.

Usage

```
optimizerNlm(start, objective=objectiveML, gradient=TRUE,
maxiter, debug, par.size, model.description, warn, ...)
```

```
optimizerOptim(start, objective=objectiveML, gradient=TRUE,
maxiter, debug, par.size, model.description, warn, method="CG", ...)
```

```
optimizerNlminb(start, objective=objectiveML, gradient=TRUE, maxiter,
debug, par.size, model.description, warn, ...)
```

Arguments

| | |
|--------------------------------|---|
| <code>start</code> | a vector of start values for the parameters. |
| <code>objective</code> | the objective function to be optimized; see objective functions . |
| <code>gradient</code> | TRUE if an analytic gradient is to be used (if one is available). |
| <code>maxiter</code> | the maximum number of iterations allowed. |
| <code>debug</code> | TRUE to show the iteration history and other available information about the optimization. |
| <code>par.size</code> | "startvalues" to have the optimizer scale the problem according to the magnitudes of the start values (ignored by <code>optimizerNlminb</code>). |
| <code>model.description</code> | a list with elements describing the structural-equation model (see the code for details). |
| <code>warn</code> | if FALSE, suppress warnings during the optimization. |
| <code>method</code> | the method to be employed by the <code>optim</code> optimizer; the default is "CG" (conjugate-gradient). |
| <code>...</code> | additional arguments for the <code>nlm</code> , <code>optim</code> , or <code>nlminb</code> optimizer. |

Value

An object of class "semResult", with elements:

| | |
|--------------------------|--|
| <code>convergence</code> | TRUE if the optimization apparently converged. |
|--------------------------|--|

| | |
|------------|--|
| iterations | the number of iterations required. |
| par | the vector of parameter estimates. |
| vcov | the estimated covariance matrix of the parameter estimates, based on a numeric Hessian; not supplied by optimizerNlminb. |
| criterion | the optimized value of the objective function. |
| C | the model-implied covariance or moment matrix at the parameter estimates. |
| A | the estimated A matrix. |
| P | the estimated P matrix. |

Author(s)

John Fox <jfox@mcmaster.ca>

See Also

[sem](#), [objective.functions](#), [nlm](#), [optim](#), [nlminb](#)

pathDiagram

Draw Path Diagram

Description

pathDiagram creates a description of the path diagram for a structural-equation-model object to be processed by the graph-drawing program *dot*, which can be called automatically; see Koutsofios and North (2002) and <http://www.graphviz.org/>. To obtain graphics output directly, the *dot* program must be on the system search path.

Usage

```
## S3 method for class 'sem'
pathDiagram(model, file, min.rank=NULL, max.rank=NULL,
  same.rank=NULL, variables=model$var.names, parameters=rownames(model$ram),
  ignore.double=TRUE, edge.labels=c("names", "values", "both"),
  size=c(8, 8), node.font=c("Helvetica", 14),
  edge.font=c("Helvetica", 10), rank.direction=c("LR", "TB"),
  digits=2, standardize=FALSE, output.type=c("graphics", "dot"),
  graphics.fmt="pdf", dot.options=NULL, ...)
```

Arguments

| | |
|-------|---|
| model | a structural-equation-model object produced by sem. |
| ... | ignored. |

| | |
|----------------|---|
| file | a file name, given <i>without</i> an extension, to which to write the <i>dot</i> description of the path diagram and, by default, to use that to draw a graph of the model; if not specified, the <i>dot</i> description is written to standard output (normally the R Console). Note that by default (see the output . type argument below), <i>two</i> files are written: a ".dot" file and a graphics file of type specified by the graphics .fmt argument (below); file may include a path specification. |
| min.rank | a character string listing names of variables to be assigned minimum rank (order) in the graph; the names should be separated by commas. |
| max.rank | a character string listing names of variables to be assigned maximum rank in the graph; the names should be separated by commas. |
| same.rank | a character string or vector of character strings of variables to be assigned equivalent rank in the graph; names in each string should be separated by commas. |
| variables | variable names; defaults to the variable names in model. If specified, the variable names should be in the same order as in model. |
| parameters | parameter names; defaults to the parameter names in model. If specified, the parameter names should be in the same order as in model. |
| ignore.double | if TRUE, the default, double-headed arrows, representing variances and covariances, are not graphed. |
| edge.labels | "names" to label arrows with parameter names; "values" to label arrows with parameter estimates, or "both". |
| size | the size of the graph, in inches. |
| node.font | font name and point-size for printing variable names. |
| edge.font | font name and point-size for printing arrow names or values. |
| rank.direction | draw graph left-to-right, "LR", the default, or top-to-bottom, "TB". |
| digits | number of digits after the decimal point (default, 2) to which to round parameter estimates. |
| standardize | if TRUE, display standardized coefficients; default is FALSE. |
| output.type | if "graphics", the default, <i>both</i> a ".dot" file and a graphics file will be created. |
| graphics.fmt | a graphics format recognized by the <i>dot</i> program; the default is "pdf"; graphics .fmt is also used for the extension of the graphics file that is created. |
| dot.options | options to be passed to the <i>dot</i> program, given as a character string. |

Value

NULL: pathDiagram is used for its side-effect, producing a graph description of the model.

Author(s)

John Fox <jfox@mcmaster.ca>, Adam Kramer, and Michael Friendly

References

Koutsofios, E., and North, S. C. (2002) Drawing graphs with *dot*. <http://www.graphviz.org/Documentation/dotguide.pdf>.

See Also[sem](#)**Examples**

```

## Not run:
# The Duncan, Haller, and Portes Peer-Influences Model

R.DHP <- readMoments(diag=FALSE, names=c('R0ccAsp', 'REdAsp', 'F0ccAsp',
    'FEdAsp', 'RParAsp', 'RIQ', 'RSES', 'FSES', 'FIQ', 'FParAsp'))
.6247
.3269 .3669
.4216 .3275 .6404
.2137 .2742 .1124 .0839
.4105 .4043 .2903 .2598 .1839
.3240 .4047 .3054 .2786 .0489 .2220
.2930 .2407 .4105 .3607 .0186 .1861 .2707
.2995 .2863 .5191 .5007 .0782 .3355 .2302 .2950
.0760 .0702 .2784 .1988 .1147 .1021 .0931 -.0438 .2087

model.dhp <- specifyModel()
RParAsp -> RGenAsp, gam11, NA
RIQ -> RGenAsp, gam12, NA
RSES -> RGenAsp, gam13, NA
FSES -> RGenAsp, gam14, NA
RSES -> FGenAsp, gam23, NA
FSES -> FGenAsp, gam24, NA
FIQ -> FGenAsp, gam25, NA
FParAsp -> FGenAsp, gam26, NA
FGenAsp -> RGenAsp, beta12, NA
RGenAsp -> FGenAsp, beta21, NA
RGenAsp -> R0ccAsp, NA, 1
RGenAsp -> REdAsp, lam21, NA
FGenAsp -> F0ccAsp, NA, 1
FGenAsp -> FEdAsp, lam42, NA
RGenAsp <-> RGenAsp, ps11, NA
FGenAsp <-> FGenAsp, ps22, NA
RGenAsp <-> FGenAsp, ps12, NA
R0ccAsp <-> R0ccAsp, theta1, NA
REdAsp <-> REdAsp, theta2, NA
F0ccAsp <-> F0ccAsp, theta3, NA
FEdAsp <-> FEdAsp, theta4, NA

sem.dhp <- sem(model.dhp, R.DHP, 329,
    fixed.x=c('RParAsp', 'RIQ', 'RSES', 'FSES', 'FIQ', 'FParAsp'))

pathDiagram(sem.dhp, min.rank='RIQ, RSES, RParAsp, FParAsp, FSES, FIQ',
    max.rank='R0ccAsp, REdAsp, FEdAsp, F0ccAsp')

## End(Not run)

```

| | |
|-----|---|
| ram | <i>RAM Matrix for a Structural-Equation Model</i> |
|-----|---|

Description

Print the labelled RAM definition matrix for a structural-equation model fit by sem.

Usage

```
ram(object, digits=5, startvalues=FALSE)
```

Arguments

| | |
|-------------|---|
| object | an object of class sem returned by the sem function. |
| digits | number of digits for printed output. |
| startvalues | if TRUE, start values for parameters are printed; otherwise, the parameter estimates are printed; the default is FALSE. |

Value

A data frame containing the labelled RAM definition matrix, which is normally just printed.

Author(s)

John Fox <jfox@mcmaster.ca>

See Also

[sem](#)

Examples

```
# ----- assumes that Duncan, Haller and Portes peer-influences model
# ----- has been fit and is in sem.dhp

## Not run:
ram(sem.dhp)

## End(Not run)
```

rawMoments

*Compute Raw Moments Matrix***Description**

Computes the “uncorrected” sum-of-squares-and-products matrix divided by the number of observations.

Usage

```
## S3 method for class 'formula'
rawMoments(formula, data, subset, na.action,
            contrasts=NULL, ...)

## Default S3 method:
rawMoments(object, na.rm=FALSE, ...)

cov2raw(cov, mean, N, sd)

## S3 method for class 'rawmoments'
print(x, ...)
```

Arguments

| | |
|-----------|---|
| object | a one-sided model formula or an object coercible to a numeric matrix. |
| formula | a one-sided model formula specifying the model matrix for which raw moments are to be computed; note that a constant is included if it is not explicitly suppressed by putting -1 in the formula. |
| data | an optional data frame containing the variables in the formula. By default the variables are taken from the environment from which rawMoments is called. |
| subset | an optional vector specifying a subset of observations to be used in computing moments. |
| na.action | a function that indicates what should happen when the data contain NAs. The default is set by the na.action option. |
| contrasts | an optional list. See the contrasts.arg argument of model.matrix.default . |
| na.rm | if TRUE, any data rows with missing data will be removed. |
| cov | a covariance or correlation matrix. |
| mean | a vector of means. |
| N | the number of observations on which the covariances or correlations are based. |
| sd | an optional vector of standard deviations, to be given if cov is a correlation matrix. |
| x | an object of class rawmoments to print. |
| ... | arguments passed down. |

Value

rawMoments and cov2raw return an object of class rawmoments, which is simply a matrix with an attribute "N" that contains the number of observations on which the moments are based.

Author(s)

John Fox <jfox@mcmaster.ca>

See Also

[sem](#)

Examples

```
# the following are all equivalent (with the exception of the name of the intercept):

rawMoments(cbind(1, Kmenta))

rawMoments(~ Q + P + D + F + A, data=Kmenta)

Cov <- with(Kmenta, cov(cbind(Q, P, D, F, A)))
cov2raw(Cov, colMeans(Kmenta), nrow(Kmenta))
```

readMoments

Input a Covariance, Correlation, or Raw Moment Matrix

Description

This functions makes it simpler to input covariance, correlation, and raw-moment matrices to be analyzed by the [sem](#) function. The matrix is input in lower-triangular form on as many lines as is convenient, omitting the above-diagonal elements. The elements on the diagonal may also optionally be omitted, in which case they are taken to be 1.

Usage

```
readMoments(file="", diag=TRUE,
            names=as.character(paste("X", 1:n, sep = "")))
```

Arguments

| | |
|-------|--|
| file | The (quoted) file from which to read the model specification, including the path to the file if it is not in the current directory. If "" (the default), then the specification is read from the standard input stream, and is terminated by a blank line. |
| diag | If TRUE (the default), then the input matrix includes diagonal elements. |
| names | a character vector containing the names of the variables, to label the rows and columns of the moment matrix. |

Value

Returns a lower-triangular matrix (i.e., with zeroes above the main diagonal) suitable for input to sem.

Author(s)

John Fox <jfox@mcmaster.ca>

See Also

[sem](#)

Examples

```
## Not run:
R.DHP <- readMoments(diag=FALSE, names=c("R0ccAsp", "REdAsp", "F0ccAsp",
    "FEdAsp", "RParAsp", "RIQ", "RSES", "FSES", "FIQ", "FParAsp"))
.6247
.3269 .3669
.4216 .3275 .6404
.2137 .2742 .1124 .0839
.4105 .4043 .2903 .2598 .1839
.3240 .4047 .3054 .2786 .0489 .2220
.2930 .2407 .4105 .3607 .0186 .1861 .2707
.2995 .2863 .5191 .5007 .0782 .3355 .2302 .2950
.0760 .0702 .2784 .1988 .1147 .1021 .0931 -.0438 .2087

R.DHP

## End(Not run)
```

residuals.sem

Residual Covariances for a Structural Equation Model

Description

These functions compute residual covariances, variance-standardized residual covariances, and normalized residual covariances for the observed variables in a structural-equation model fit by sem.

Usage

```
## S3 method for class 'sem'
residuals(object, ...)

## S3 method for class 'sem'
standardizedResiduals(object, ...)

## S3 method for class 'objectiveML'
```

```
normalizedResiduals(object, ...)

## S3 method for class 'objectiveGLS'
normalizedResiduals(object, ...)
```

Arguments

`object` an object inheriting from class `sem` returned by the `sem` function.
`...` not for the user.

Details

Residuals are defined as $S - C$, where S is the sample covariance matrix of the observed variables and C is the model-reproduced covariance matrix. The *standardized* residual covariance for a pair of variables divides the residual covariance by the product of the sample standard deviations of the two variables, $(s_{ij} - c_{ij}) / (s_{ii}s_{jj})^{1/2}$. The *normalized* residual is given by

$$\frac{s_{ij} - c_{ij}}{[(c_{ii}c_{jj} - c_{ij}^2)/N^*]^{1/2}}$$

where N^* is the number of observations minus one if the model is fit to a covariance matrix, or the number of observations if it is fit to a raw moment matrix.

Value

Each function returns a matrix of residuals.

Author(s)

John Fox <jfox@mcmaster.ca>

References

Bollen, K. A. (1989) *Structural Equations With Latent Variables*. Wiley.

See Also

[sem](#)

Examples

```
# In the first example, readMoments() and specifyModel() read from the
# input stream. This example cannot be executed via example() but can be entered
# at the command prompt. The example is repeated using file input;
# this example can be executed via example().
## Not run:
# Duncan, Haller, and Portes peer-influences model

R.DHP <- readMoments(diag=FALSE, names=c("R0ccAsp", "REdAsp", "FOccAsp",
    "FEdAsp", "RParAsp", "RIQ", "RSES", "FSES", "FIQ", "FParAsp"))
.6247
```

```
.3269 .3669
.4216 .3275 .6404
.2137 .2742 .1124 .0839
.4105 .4043 .2903 .2598 .1839
.3240 .4047 .3054 .2786 .0489 .2220
.2930 .2407 .4105 .3607 .0186 .1861 .2707
.2995 .2863 .5191 .5007 .0782 .3355 .2302 .2950
.0760 .0702 .2784 .1988 .1147 .1021 .0931 -.0438 .2087
```

```
model.dhp <- specifyModel()
  RParAsp -> RGenAsp, gam11, NA
  RIQ     -> RGenAsp, gam12, NA
  RSES    -> RGenAsp, gam13, NA
  FSES    -> RGenAsp, gam14, NA
  RSES    -> FGenAsp, gam23, NA
  FSES    -> FGenAsp, gam24, NA
  FIQ     -> FGenAsp, gam25, NA
  FParAsp -> FGenAsp, gam26, NA
  FGenAsp -> RGenAsp, beta12, NA
  RGenAsp -> FGenAsp, beta21, NA
  RGenAsp -> ROccAsp, NA, 1
  RGenAsp -> REdAsp, lam21, NA
  FGenAsp -> FOccAsp, NA, 1
  FGenAsp -> FEdAsp, lam42, NA
  RGenAsp <-> RGenAsp, ps11, NA
  FGenAsp <-> FGenAsp, ps22, NA
  RGenAsp <-> FGenAsp, ps12, NA
  ROccAsp <-> ROccAsp, theta1, NA
  REdAsp  <-> REdAsp, theta2, NA
  FOccAsp <-> FOccAsp, theta3, NA
  FEdAsp  <-> FEdAsp, theta4, NA

sem.dhp <- sem(model.dhp, R.DHP, 329,
  fixed.x=c('RParAsp', 'RIQ', 'RSES', 'FSES', 'FIQ', 'FParAsp'))
residuals(sem.dhp)
normalizedResiduals(sem.dhp)
standardizedResiduals(sem.dhp) # same as residuals because model is fit to correlations

## End(Not run)
# The following example can be executed via example():

etc <- file.path(.path.package(package="sem")[1], "etc") # path to data and model files

(R.DHP <- readMoments(file=file.path(etc, "R-DHP.txt"),
  diag=FALSE, names=c("ROccAsp", "REdAsp", "FOccAsp",
    "FEdAsp", "RParAsp", "RIQ", "RSES", "FSES", "FIQ", "FParAsp")))
(model.dhp <- specifyModel(file=file.path(etc, "model-DHP.txt")))
(sem.dhp <- sem(model.dhp, R.DHP, 329,
  fixed.x=c('RParAsp', 'RIQ', 'RSES', 'FSES', 'FIQ', 'FParAsp')))

residuals(sem.dhp)

normalizedResiduals(sem.dhp)
```

```
standardizedResiduals(sem.dhp) # same as residuals because model is fit to correlations
```

 sem

General Structural Equation Models

Description

sem fits general structural equation models (with both observed and unobserved variables). Observed variables are also called *indicators* or *manifest variables*; unobserved variables are also called *factors* or *latent variables*. Normally, the generic function (sem) is called directly with a semmod first argument, invoking the sem.semmod method, which in turn sets up a call to the sem.default method; thus, the user may wish to specify arguments accepted by the semmod and default methods.

Usage

```
## S3 method for class 'semmod'
sem(model, S, N, data, raw=FALSE, obs.variables=rownames(S),
fixed.x=NULL, formula= ~ ., robust=!missing(data), debug=FALSE, ...)

## Default S3 method:
sem(model, S, N, data=NULL, raw=FALSE, param.names,
var.names, fixed.x=NULL, robust=!is.null(data), semmod=NULL, debug=FALSE,
analytic.gradient=TRUE, warn=FALSE, maxiter=500,
par.size=c("ones", "startvalues"), start.tol=1E-6,
optimizer=optimizerNlm, objective=objectiveML, ...)

startvalues(S, ram, debug=FALSE, tol=1E-6)

## S3 method for class 'sem'
coef(object, standardized=FALSE, ...)
## S3 method for class 'sem'
vcov(object, robust=FALSE,
analytic=inherits(object, "objectiveML") && object$t <= 100, ...)
## S3 method for class 'sem'
df.residual(object, ...)
```

Arguments

| | |
|-------|--|
| model | RAM specification, which is a simple encoding of the path diagram for the model. The model may be given either in symbolic form (as a semmod object, as returned by the specifyModel , specifyEquations , or cfa function, or as a character matrix), invoking sem.semmod, which calls sem.default after setting up the model, or (less conveniently) in numeric form, invoking sem.default directly, which is not recommended (see Details below). |
|-------|--|

| | |
|----------------------------|--|
| S | covariance matrix among observed variables; may be input as a symmetric matrix, or as a lower- or upper-triangular matrix. S may also be a raw (i.e., “uncorrected”) moment matrix — that is, a sum-of-squares-and-products matrix divided by N. This form of input is useful for fitting models with intercepts, in which case the moment matrix should include the mean square and cross-products for a unit variable all of whose entries are 1; of course, the raw mean square for the unit variable is 1. Raw-moment matrices may be computed by rawMoments . If the <code>ram</code> argument is given in symbolic form, then the observed-variable covariance or raw-moment matrix may contain variables that do not appear in the model, in which case a warning is printed. As an alternative to specifying S the user may supply a data frame containing the data for the model (see the argument <code>data</code>). |
| N | number of observations on which the covariance matrix is based. |
| <code>data</code> | As a generally preferable alternative to specifying S and N, the user may supply a data frame containing the data to which the model is to be fit. |
| <code>raw</code> | TRUE if S is a raw moment matrix or if a raw moment matrix — as opposed to a covariance matrix — is to be computed from data; the default is FALSE. |
| <code>obs.variables</code> | names of observed variables, by default taken from the row names of the covariance or moment matrix S, which may be given directly or generated according to the <code>data</code> and <code>formula</code> arguments. |
| <code>fixed.x</code> | names (if the <code>ram</code> matrix is given in symbolic form) or indices (if it is in numeric form) of fixed exogenous variables. Specifying these obviates the necessity of having to fix the variances and covariances among these variables (and produces correct degrees of freedom for the model <code>chisquare</code>). |
| <code>formula</code> | a one-sided formula, to be applied to data to generate the variables for which covariances or raw moments are computed. The default formula is <code>~.</code> , i.e., all of the variables in the data, including an implied intercept; if a covariance matrix is to be computed, the constant is suppressed. |
| <code>robust</code> | In <code>sem</code> : if TRUE, then quantities are calculated that can be used to compute robust estimates of coefficient standard errors and robust tests when the model is fit by multinormal maximum likelihood; the default is TRUE when the <code>data</code> argument is TRUE, and this option is only available when the <code>data</code> argument is given. In <code>vcov</code> : if TRUE, return a robust coefficient covariance matrix (if object contains the requisite information). |
| <code>semmod</code> | a <code>semmod</code> object containing the description of the model; optional, and normally supplied not directly by the user but via the <code>semmod</code> method for <code>sem</code> . |
| <code>debug</code> | if TRUE, some information is printed to help you debug the symbolic model specification; for example, if a variable name is misspelled, <code>sem</code> will assume that the variable is a (new) latent variable. Information about the optimization will also be printed, but details will vary with the optimizer employed. The default is FALSE. |
| <code>...</code> | arguments to be passed down, including from <code>sem.default</code> to the optimizer. |
| <code>param.names</code> | names of the t free parameters, given in their numerical order; default names are <code>Param1, ..., Paramt</code> . Note: Should not be specified when the model is given in symbolic form. |

| | |
|---|---|
| <code>var.names</code> | names of the m entries of the v vector (typically the observed and latent variables — see below), given in their numerical order; default names are <code>Var1, ..., Varm</code> . Note: Should not be specified when the model is given in symbolic form. |
| <code>analytic.gradient</code> | if TRUE (the default), then analytic first derivatives are used in the maximization of the likelihood if the optimizer employed will accept them; otherwise numeric derivatives are used, again if the optimizer will compute them. |
| <code>warn</code> | if TRUE, warnings produced by the optimization function will be printed. This should generally not be necessary, since <code>sem</code> prints its own warning, and saves information about convergence. The default is FALSE. |
| <code>maxiter</code> | the maximum number of iterations for the optimization of the objective function, to be passed to the optimizer. |
| <code>par.size</code> | the anticipated size of the free parameters; if "ones", a vector of ones is used; if "startvalues", taken from the start values. You can try changing this argument if you encounter convergence problems. The default is "startvalues" if the largest input variance is at least 100 times the smallest, and "ones" otherwise. Whether this argument is actually used depends upon the optimizer employed. |
| <code>start.tol</code> , <code>tol</code> | if the magnitude of an automatic start value is less than <code>start.tol</code> , then it is set to <code>start.tol</code> ; defaults to 1E-6. |
| <code>optimizer</code> | a function to be used to minimize the objective function; the default is <code>optimizerNlm</code> , which employs the standard R <code>nlm</code> optimizer. Alternatives are <code>optimizerOptim</code> , which employs <code>optim</code> , and <code>optimizerNlminb</code> , which uses <code>nlminb</code> — or the user can supply an optimizer. |
| <code>objective</code> | An objective function to be minimized, sometimes called a "fit" function in the SEM literature. The default is <code>objectiveML</code> , which produces maximum-likelihood estimates assuming multinormality. An alternative is <code>objectiveGLS</code> , which produced generalized least squares estimates — or the user can supply an objective function to be minimized. |
| <code>ram</code> | numeric RAM matrix. |
| <code>object</code> | an object of class "sem", returned by <code>sem</code> . |
| <code>standardized</code> | if TRUE, return standardized coefficients. |
| <code>analytic</code> | return an analytic (as opposed to numeric) estimate of the coefficient covariance matrix; at present only available for the <code>objectiveML</code> objective function. The default is FALSE for this objective function if there are no more than 100 parameters and FALSE otherwise. |

Details

The model is set up using either RAM ("reticular action model" – don't ask!) notation – a simple format for specifying general structural equation models by coding the "arrows" in the path diagram for the model (see, e.g., McArdle and McDonald, 1984) – typically using the `specifyModel` function; in equation format using the `specifyEquations` function; or, for a simple confirmatory factor analysis model, via the `cfa` function. In any case, the model is represented internally in RAM format.

The variables in the v vector in the model (typically, the observed and unobserved variables, but not error variables) are numbered from 1 to m . the RAM matrix contains one row for each (free or

constrained) parameter of the model, and may be specified either in symbolic format or in numeric format.

A symbolic ram matrix consists of three columns, as follows:

- 1. Arrow specification:** This is a simple formula, of the form "A -> B" or, equivalently, "B <- A" for a regression coefficient (i.e., a single-headed or directional arrow); "A <-> A" for a variance or "A <-> B" for a covariance (i.e., a double-headed or bidirectional arrow). Here, A and B are variable names in the model. If a name does not correspond to an observed variable, then it is assumed to be a latent variable. Spaces can appear freely in an arrow specification, and there can be any number of hyphens in the arrows, including zero: Thus, e.g., "A->B", "A --> B", and "A>B" are all legitimate and equivalent.
- 2. Parameter name:** The name of the regression coefficient, variance, or covariance specified by the arrow. Assigning the same name to two or more arrows results in an equality constraint. Specifying the parameter name as NA produces a fixed parameter.
- 3. Value:** start value for a free parameter or value of a fixed parameter. If given as NA, sem will compute the start value.

It is simplest to construct the RAM matrix with the `specifyModel`, `specifyEquations`, or `cfa` function, all of which return an object of class `semmod`, and also incorporate some model-specification convenience shortcuts. This process is illustrated in the examples below.

A numeric ram matrix consists of five columns, as follows:

- 1. Number of arrow heads:** 1 (directed arrow) or 2 (covariance).
- 2. Arrow to:** index of the variable at the head of a directional arrow, or at one end of a bidirectional arrow. Observed variables should be assigned the numbers 1 to n , where n is the number of rows/columns in the covariance matrix S , with the indices corresponding to the variables' positions in S . Variable indices above n represent latent variables.
- 3. Arrow from:** the index of the variable at the tail of a directional arrow, or at the other end of a bidirectional arrow.
- 4. Parameter number:** free parameters are numbered from 1 to t , but do not necessarily appear in consecutive order. Fixed parameters are given the number 0. Equality constraints are specified by assigning two or more parameters the same number.
- 5. Value:** start value for a free parameter, or value of a fixed parameter. If given as NA, the program will compute a start value, by a slight modification of the method described by McDonald and Hartmann (1992). *Note:* In some circumstances, some start values are selected randomly; this might produce small differences in the parameter estimates when the program is rerun.

The numeric ram matrix is normally generated automatically, not specified directly by the user.

For `specifyEquations`, each input line is either a regression equation or the specification of a variance or covariance. Regression equations are of the form

$$y = \text{par1} * x_1 + \text{par2} * x_2 + \dots + \text{park} * x_k$$

where y and the x s are variables in the model (either observed or latent), and the pars are parameters. If a parameter is given as a numeric value (e.g., 1) then it is treated as fixed. Note that no "error" variable is included in the equation; "error variances" are specified via either the `covs` argument, via $V(y) = \text{par}$ (see immediately below), or are added automatically to the model when, as by default, `endog.variances=TRUE`.

Variances are specified in the form $V(\text{var}) = \text{par}$ and covariances in the form $C(\text{var1}, \text{var2}) = \text{par}$, where the vars are variables (observed or unobserved) in the model. The symbols V and C may be in either lower- or upper-case. If par is a numeric value (e.g., 1) then it is treated as fixed. In conformity with the RAM model, a variance or covariance for an endogenous variable in the model is an “error” variance or covariance.

To set a start value for a free parameter, enclose the numeric start value in parentheses after the parameter name, as `parameter(value)`.

`sem` fits the model by calling the optimizer specified in the `optimizer` argument to minimize the objective function specified in the `objective` argument. If the optimization fails to converge, a warning message is printed.

The RAM formulation of the general structural equation model is given by the basic equation

$$v = Av + u$$

where v and u are vectors of random variables (observed or unobserved), and the parameter matrix A contains regression coefficients, symbolized by single-headed arrows in a path diagram. Another parameter matrix,

$$P = E(uu')$$

contains covariances among the elements of u (assuming that the elements of u have zero means). Usually v contains endogenous and exogenous observed and unobserved variables, but not error variables (see the examples below).

The `startvalues` function may be called directly, but is usually called by `sem.default`.

Value

`sem` returns an object of class `c(objective, "sem")`, where `objective` is the name of the objective function that was optimized (e.g., "objectiveML"), with the following elements:

| | |
|------------------------|--|
| <code>var.names</code> | vector of variable names. |
| <code>ram</code> | RAM matrix, including any rows generated for covariances among fixed exogenous variables; column 5 includes computed start values. |
| <code>S</code> | observed covariance matrix. |
| <code>J</code> | RAM selection matrix, J , which picks out observed variables. |
| <code>n.fix</code> | number of fixed exogenous variables. |
| <code>n</code> | number of observed variables. |
| <code>N</code> | number of observations. |
| <code>m</code> | number of variables (observed plus unobserved). |
| <code>t</code> | number of free parameters. |
| <code>raw</code> | TRUE if the model is fit to a raw moment matrix, FALSE otherwise. |
| <code>data</code> | the observed-variable data matrix, or NULL if data are not supplied. |
| <code>semmod</code> | the <code>semmod</code> specification object for the model, if one was supplied; otherwise NULL. |
| <code>optimizer</code> | the optimizer function. |
| <code>objective</code> | the objective function. |

| | |
|-------------|--|
| coeff | estimates of free parameters. |
| vcov | estimated asymptotic covariance matrix of parameter estimates, based on a numeric Hessian, if supplied by the optimizer; otherwise NULL. |
| par.posn | indices of free parameters. |
| convergence | TRUE or FALSE, depending upon whether the optimization apparently converged. |
| iterations | number of iterations performed. |
| criterion | value of the objective function at the minimum. |
| C | model-reproduced covariance matrix. |
| A | RAM <i>A</i> matrix. |
| P | RAM <i>P</i> matrix. |
| adj.obj | robust adjusted value of the objective function; NULL if robust is FALSE. |
| robust.vcov | robust estimated coefficient covariance matrix; NULL if robust is FALSE. |

Warning

A common error is to fail to specify variance or covariance terms in the model, which are denoted by double-headed arrows, <->.

In general, every observed or latent variable in the model should be associated with a variance or error variance. This may be a free parameter to estimate or a fixed constant (as in the case of a latent exogenous variable for which you wish to fix the variance, e.g., to 1). Again in general, there will be an *error variance* associated with each endogenous variable in the model (i.e., each variable to which at least one single-headed arrow points — including observed indicators of latent variables), and a *variance* associated with each exogenous variable (i.e., each variable that appears only at the tail of single-headed arrows, never at the head).

To my knowledge, the only *apparent* exception to this rule is for observed variables that are declared to be fixed exogenous variables. In this case, the program generates the necessary (fixed-constant) variances and covariances automatically.

If there are missing variances, a warning message will be printed, and estimation will almost surely fail in some manner. Missing variances might well indicate that there are missing covariances too, but it is not possible to deduce this in a mechanical manner. The `specifyModel` function will by default supply error-variance parameters if these are missing.

Author(s)

John Fox <jfox@mcmaster.ca> and Jarrett Byrnes

References

- Fox, J. (2006) Structural equation modeling with the sem package in R. *Structural Equation Modeling* **13**:465–486.
- Bollen, K. A. (1989) *Structural Equations With Latent Variables*. Wiley.
- Bollen, K. A. and Long, J. S. (eds.) *Testing Structural Equation Models*, Sage.
- McArdle, J. J. and Epstein, D. (1987) Latent growth curves within developmental structural equation models. *Child Development* **58**, 110–133.

McArdle, J. J. and McDonald, R. P. (1984) Some algebraic properties of the reticular action model. *British Journal of Mathematical and Statistical Psychology* **37**, 234–251.

McDonald, R. P. and Hartmann, W. M. (1992) A procedure for obtaining initial values of parameters in the RAM model. *Multivariate Behavioral Research* **27**, 57–76.

Raftery, A. E. (1993) Bayesian model selection in structural equation models. In Bollen, K. A. and Long, J. S. (eds.) *Testing Structural Equation Models*, Sage.

Raftery, A. E. (1995) Bayesian model selection in social research (with discussion). *Sociological Methodology* **25**, 111–196.

Satorra, A. (2000) Scaled and adjusted restricted tests in multi-sample analysis of moment structures. pp. 233–247 in Heijmans, R.D.H., Pollock, D.S.G. & Satorra, A. (eds.) *Innovations in Multivariate Statistical Analysis. A Festschrift for Heinz Neudecker*, Kluwer.

See Also

[rawMoments](#), [startvalues](#), [objectiveML](#), [objectiveGLS](#), [optimizerNlm](#), [optimizerOptim](#), [optimizerNlminb](#), [nlm](#), [optim](#), [nlminb](#), [specifyModel](#), [specifyEquations](#), [cfa](#)

Examples

```
# Note: The first set of examples can't be run via example() because the default file
# argument of specifyModel() and readMoments() requires that the model
# specification and covariances, correlations, or raw moments be entered
# at the command prompt. The examples can be copied and run in the R console,
# however. See ?specifyModel and ?readMoments for further information.
# These examples are repeated below using file input to specifyModel() and
# readMoments(). The second version of the examples may be executed through example().
```

```
## Not run:
```

```
# ----- Duncan, Haller and Portes peer-influences model -----
# A nonrecursive SEM with unobserved endogenous variables and fixed exogenous variables
```

```
R.DHP <- readMoments(diag=FALSE, names=c("ROccAsp", "REdAsp", "FOccAsp",
    "FEdAsp", "RParAsp", "RIQ", "RSES", "FSSES", "FIQ", "FParAsp"))
```

```
.6247
.3269 .3669
.4216 .3275 .6404
.2137 .2742 .1124 .0839
.4105 .4043 .2903 .2598 .1839
.3240 .4047 .3054 .2786 .0489 .2220
.2930 .2407 .4105 .3607 .0186 .1861 .2707
.2995 .2863 .5191 .5007 .0782 .3355 .2302 .2950
.0760 .0702 .2784 .1988 .1147 .1021 .0931 -.0438 .2087
```

```
# Fit the model using a symbolic ram specification
```

```
model.dhp <- specifyModel()
  RParAsp -> RGenAsp, gam11, NA
  RIQ     -> RGenAsp, gam12, NA
```

```

RSES    -> RGenAsp, gam13, NA
FSES    -> RGenAsp, gam14, NA
RSES    -> FGenAsp, gam23, NA
FSES    -> FGenAsp, gam24, NA
FIQ     -> FGenAsp, gam25, NA
FParAsp -> FGenAsp, gam26, NA
FGenAsp -> RGenAsp, beta12, NA
RGenAsp -> FGenAsp, beta21, NA
RGenAsp -> ROccAsp, NA,    1
RGenAsp -> REdAsp, lam21, NA
FGenAsp -> FOccAsp, NA,    1
FGenAsp -> FEdAsp, lam42, NA
RGenAsp <-> RGenAsp, ps11, NA
FGenAsp <-> FGenAsp, ps22, NA
RGenAsp <-> FGenAsp, ps12, NA
ROccAsp <-> ROccAsp, theta1, NA
REdAsp  <-> REdAsp, theta2, NA
FOccAsp <-> FOccAsp, theta3, NA
FEdAsp  <-> FEdAsp, theta4, NA

# an equivalent specification, allowing specifyModel() to generate
# variance parameters for endogenous variables (and suppressing the
# unnecessary NAs):

model.dhp <- specifyModel()
RParAsp  -> RGenAsp, gam11
RIQ      -> RGenAsp, gam12
RSES     -> RGenAsp, gam13
FSES     -> RGenAsp, gam14
RSES     -> FGenAsp, gam23
FSES     -> FGenAsp, gam24
FIQ      -> FGenAsp, gam25
FParAsp  -> FGenAsp, gam26
FGenAsp  -> RGenAsp, beta12
RGenAsp  -> FGenAsp, beta21
RGenAsp  -> ROccAsp, NA,    1
RGenAsp  -> REdAsp, lam21
FGenAsp  -> FOccAsp, NA,    1
FGenAsp  -> FEdAsp, lam42
RGenAsp  <-> FGenAsp, ps12

# Another equivalent specification, telling specifyModel to add paths for
# variances and covariance of RGenAsp and FGenAsp:

model.dhp <- specifyModel(covs="RGenAsp, FGenAsp")
RParAsp  -> RGenAsp, gam11
RIQ      -> RGenAsp, gam12
RSES     -> RGenAsp, gam13
FSES     -> RGenAsp, gam14
RSES     -> FGenAsp, gam23
FSES     -> FGenAsp, gam24
FIQ      -> FGenAsp, gam25
FParAsp  -> FGenAsp, gam26

```

```

FGenAsp -> RGenAsp, beta12
RGenAsp -> FGenAsp, beta21
RGenAsp -> ROccAsp, NA, 1
RGenAsp -> REdAsp, lam21
FGenAsp -> FOccAsp, NA, 1
FGenAsp -> FEdAsp, lam42

# Yet another equivalent specification using specifyEquations():

model.dhp.1 <- specifyEquations(covs="RGenAsp, FGenAsp")
RGenAsp = gam11*RParAsp + gam12*RIQ + gam13*RSES + gam14*FSSES + beta12*FGenAsp
FGenAsp = gam23*RSES + gam24*FSSES + gam25*FIQ + gam26*FParAsp + beta21*RGenAsp
ROccAsp = 1*RGenAsp
REdAsp = lam21(1)*RGenAsp # to illustrate setting start values
FOccAsp = 1*FGenAsp
FEdAsp = lam42(1)*FGenAsp

sem.dhp.1 <- sem(model.dhp, R.DHP, 329,
  fixed.x=c('RParAsp', 'RIQ', 'RSES', 'FSSES', 'FIQ', 'FParAsp'))
summary(sem.dhp.1)

# Fit the model using a numerical ram specification (not recommended!)

ram.dhp <- matrix(c(
#      heads  to    from  param  start
      1,      1,   11,    0,     1,
      1,      2,   11,    1,    NA, # lam21
      1,      3,   12,    0,     1,
      1,      4,   12,    2,    NA, # lam42
      1,     11,    5,    3,    NA, # gam11
      1,     11,    6,    4,    NA, # gam12
      1,     11,    7,    5,    NA, # gam13
      1,     11,    8,    6,    NA, # gam14
      1,     12,    7,    7,    NA, # gam23
      1,     12,    8,    8,    NA, # gam24
      1,     12,    9,    9,    NA, # gam25
      1,     12,   10,   10,    NA, # gam26
      1,     11,   12,   11,    NA, # beta12
      1,     12,   11,   12,    NA, # beta21
      2,      1,    1,   13,    NA, # theta1
      2,      2,    2,   14,    NA, # theta2
      2,      3,    3,   15,    NA, # theta3
      2,      4,    4,   16,    NA, # theta4
      2,     11,   11,   17,    NA, # psi11
      2,     12,   12,   18,    NA, # psi22
      2,     11,   12,   19,    NA, # psi12
    ), ncol=5, byrow=TRUE)

params.dhp <- c('lam21', 'lam42', 'gam11', 'gam12', 'gam13', 'gam14',
  'gam23', 'gam24', 'gam25', 'gam26',
  'beta12', 'beta21', 'theta1', 'theta2', 'theta3', 'theta4',
  'psi11', 'psi22', 'psi12')

```

```

vars.dhp <- c('ROccAsp', 'REdAsp', 'FOccAsp', 'FEdAsp', 'RParAsp', 'RIQ',
             'RSES', 'FSSES', 'FIQ', 'FParAsp', 'RGenAsp', 'FGenAsp')

sem.dhp.2 <- sem(ram.dhp, R.DHP, 329, param.names=params.dhp, var.names=vars.dhp,
fixed.x=5:10)
summary(sem.dhp.2)

# ----- Wheaton et al. alienation data -----

S.wh <- readMoments(names=c('Anomia67','Powerless67','Anomia71',
                          'Powerless71','Education','SEI'))

11.834
 6.947   9.364
 6.819   5.091  12.532
 4.783   5.028   7.495   9.986
-3.839  -3.889  -3.841  -3.625   9.610
-21.899 -18.831 -21.748 -18.775  35.522  450.288

# This is the model in the SAS manual for PROC CALIS: A Recursive SEM with
# latent endogenous and exogenous variables.
# Curiously, both factor loadings for two of the latent variables are fixed.

model.wh.1 <- specifyModel()
  Alienation67  -> Anomia67,      NA,      1
  Alienation67  -> Powerless67,  NA,      0.833
  Alienation71  -> Anomia71,     NA,      1
  Alienation71  -> Powerless71,  NA,      0.833
  SES           -> Education,    NA,      1
  SES           -> SEI,          lamb,    NA
  SES           -> Alienation67, gam1,    NA
  Alienation67  -> Alienation71, beta,    NA
  SES           -> Alienation71, gam2,    NA
  Anomia67      <-> Anomia67,     the1,   NA
  Anomia71      <-> Anomia71,     the1,   NA
  Powerless67   <-> Powerless67,  the2,   NA
  Powerless71   <-> Powerless71,  the2,   NA
  Education     <-> Education,    the3,   NA
  SEI           <-> SEI,          the4,   NA
  Anomia67      <-> Anomia71,     the5,   NA
  Powerless67   <-> Powerless71,  the5,   NA
  Alienation67  <-> Alienation67, psi1,   NA
  Alienation71  <-> Alienation71, psi2,   NA
  SES           <-> SES,          phi,    NA

sem.wh.1 <- sem(model.wh.1, S.wh, 932)
summary(sem.wh.1)

# The same model in equation format:

model.wh.1 <- specifyEquations()
Anomia67 = 1*Alienation67

```

```

Powerless67 = 0.833*Alienation67
Anomia71 = 1*Alienation71
Powerless71 = 0.833*Alienation71
Education = 1*SES
SEI = lamb*SES
Alienation67 = gam1*SES
Alienation71 = gam2*SES + beta*Alienation67
V(Anomia67) = the1
V(Anomia71) = the1
V(Powerless67) = the2
V(Powerless71) = the2
V(SES) = phi
C(Anomia67, Anomia71) = the5
C(Powerless67, Powerless71) = the5

# The same model, but treating one loading for each latent variable as free
# (and equal to each other).

model.wh.2 <- specifyModel()
  Alienation67  -> Anomia67,      NA,      1
  Alienation67  -> Powerless67, lamby,   NA
  Alienation71  -> Anomia71,     NA,      1
  Alienation71  -> Powerless71, lamby,   NA
  SES           -> Education,   NA,      1
  SES           -> SEI,         lambx,   NA
  SES           -> Alienation67, gam1,   NA
  Alienation67  -> Alienation71, beta,   NA
  SES           -> Alienation71, gam2,   NA
  Anomia67      <-> Anomia67,    the1,   NA
  Anomia71      <-> Anomia71,    the1,   NA
  Powerless67   <-> Powerless67, the2,   NA
  Powerless71   <-> Powerless71, the2,   NA
  Education     <-> Education,   the3,   NA
  SEI           <-> SEI,         the4,   NA
  Anomia67      <-> Anomia71,    the5,   NA
  Powerless67   <-> Powerless71, the5,   NA
  Alienation67  <-> Alienation67, psi1,   NA
  Alienation71  <-> Alienation71, psi2,   NA
  SES           <-> SES,         phi,    NA

sem.wh.2 <- sem(model.wh.2, S.wh, 932)
summary(sem.wh.2)

# And again, in equation format:

model.wh <- specifyEquations()
Anomia67 = 1*Alienation67
Powerless67 = lamby*Alienation67
Anomia71 = 1*Alienation71
Powerless71 = lamby*Alienation71
Education = 1*SES
SEI = lambx*SES

```

```

Alienation67 = gam1*SES
Alienation71 = gam2*SES + beta*Alienation67
V(Anomia67) = the1
V(Anomia71) = the1
V(Powerless67) = the2
V(Powerless71) = the2
V(SES) = phi
C(Anomia67, Anomia71) = the5
C(Powerless67, Powerless71) = the5

# Compare the two models by a likelihood-ratio test:

anova(sem.wh.1, sem.wh.2)

# ----- Thurstone data -----
# Second-order confirmatory factor analysis, from the SAS manual for PROC CALIS

R.thur <- readMoments(diag=FALSE, names=c('Sentences', 'Vocabulary',
    'Sent.Completion', 'First.Letters', '4.Letter.Words', 'Suffixes',
    'Letter.Series', 'Pedigrees', 'Letter.Group'))
.828
.776 .779
.439 .493 .46
.432 .464 .425 .674
.447 .489 .443 .59 .541
.447 .432 .401 .381 .402 .288
.541 .537 .534 .35 .367 .32 .555
.38 .358 .359 .424 .446 .325 .598 .452

model.thur <- specifyModel()
F1 -> Sentences, lam11
F1 -> Vocabulary, lam21
F1 -> Sent.Completion, lam31
F2 -> First.Letters, lam41
F2 -> 4.Letter.Words, lam52
F2 -> Suffixes, lam62
F3 -> Letter.Series, lam73
F3 -> Pedigrees, lam83
F3 -> Letter.Group, lam93
F4 -> F1, gam1
F4 -> F2, gam2
F4 -> F3, gam3
F1 <-> F1, NA, 1
F2 <-> F2, NA, 1
F3 <-> F3, NA, 1
F4 <-> F4, NA, 1

sem.thur <- sem(model.thur, R.thur, 213)
summary(sem.thur)

# The model in equation format:

```

```

model.thur <- specifyEquations()
Sentences = lam11*F1
Vocabulary = lam21*F1
Sent.Completion = lam31*F1
First.Letters = lam42*F2
4.Letter.Words = lam52*F2
Suffixes = lam62*F2
Letter.Series = lam73*F3
Pedigrees = lam83*F3
Letter.Group = lam93*F3
F1 = gam1*F4
F2 = gam2*F4
F3 = gam3*F4
V(F1) = 1
V(F2) = 1
V(F3) = 1
V(F4) = 1

#----- Kerchoff/Kenney path analysis -----
# An observed-variable recursive SEM from the LISREL manual

R.kerch <- readMoments(diag=FALSE, names=c('Intelligence', 'Siblings',
                                           'FatherEd', 'FatherOcc', 'Grades', 'EducExp', 'OccupAsp'))

-.100
.277 -.152
.250 -.108 .611
.572 -.105 .294 .248
.489 -.213 .446 .410 .597
.335 -.153 .303 .331 .478 .651

model.kerch <- specifyModel()
Intelligence -> Grades,      gam51
Siblings -> Grades,        gam52
FatherEd -> Grades,        gam53
FatherOcc -> Grades,       gam54
Intelligence -> EducExp,    gam61
Siblings -> EducExp,       gam62
FatherEd -> EducExp,       gam63
FatherOcc -> EducExp,      gam64
Grades -> EducExp,         beta65
Intelligence -> OccupAsp,   gam71
Siblings -> OccupAsp,      gam72
FatherEd -> OccupAsp,     gam73
FatherOcc -> OccupAsp,    gam74
Grades -> OccupAsp,       beta75
EducExp -> OccupAsp,      beta76

sem.kerch <- sem(model.kerch, R.kerch, 737, fixed.x=c('Intelligence', 'Siblings',
                                                       'FatherEd', 'FatherOcc'))
summary(sem.kerch)

```

```

# The model in equation format:

model.kerch <- specifyEquations()
Grades = gam51*Intelligence + gam52*Siblings + gam53*FatherEd + gam54*FatherOcc
EducExp = gam61*Intelligence + gam62*Siblings + gam63*FatherEd + gam64*FatherOcc + beta65*Grades
OccupAsp = gam71*Intelligence + gam72*Siblings + gam73*FatherEd + gam74*FatherOcc + beta75*Grades + beta76*EducExp

#----- McArdle/Epstein latent-growth-curve model -----
# This model, from McArdle and Epstein (1987, p.118), illustrates the use of a
# raw moment matrix to fit a model with an intercept. (The example was suggested
# by Mike Stoolmiller.)

M.McArdle <- readMoments(names=c('WISC1', 'WISC2', 'WISC3', 'WISC4', 'UNIT'))
  365.661
  503.175    719.905
  675.656    958.479    1303.392
  890.680    1265.846    1712.475    2278.257
   18.034     25.819     35.255     46.593     1.000

mod.McArdle <- specifyModel()
  C -> WISC1, NA, 6.07
  C -> WISC2, B2, NA
  C -> WISC3, B3, NA
  C -> WISC4, B4, NA
  UNIT -> C, Mc, NA
  C <-> C, Vc, NA,
  WISC1 <-> WISC1, Vd, NA
  WISC2 <-> WISC2, Vd, NA
  WISC3 <-> WISC3, Vd, NA
  WISC4 <-> WISC4, Vd, NA

sem.McArdle <- sem(mod.McArdle, M.McArdle, 204, fixed.x="UNIT", raw=TRUE)
summary(sem.McArdle)

# The model in equation format:

mod.McArdle <- specifyEquations()
WISC1 = 6.07*C
WISC2 = B2*C
WISC3 = B3*C
WISC4 = b4*C
C = Mc*UNIT
v(C) = Vc
v(WISC1) = Vd
v(WISC2) = Vd
v(WISC3) = Vd
v(WISC4) = Vd

#----- Bollen industrialization and democracy example -----
# This model, from Bollen (1989, Ch. 8), illustrates the use in sem() of a
# case-by-variable data (see ?Bollen) set rather than a covariance or moment matrix

```

```

model.bollen <- specifyModel()
Demo60 -> y1, NA, 1
Demo60 -> y2, lam2,
Demo60 -> y3, lam3,
Demo60 -> y4, lam4,
Demo65 -> y5, NA, 1
Demo65 -> y6, lam2,
Demo65 -> y7, lam3,
Demo65 -> y8, lam4,
Indust -> x1, NA, 1
Indust -> x2, lam6,
Indust -> x3, lam7,
y1 <-> y5, theta15
y2 <-> y4, theta24
y2 <-> y6, theta26
y3 <-> y7, theta37
y4 <-> y8, theta48
y6 <-> y8, theta68
Indust -> Demo60, gamma11,
Indust -> Demo65, gamma21,
Demo60 -> Demo65, beta21,
Indust <-> Indust, phi

sem.bollen <- sem(model.bollen, data=Bollen)
summary(sem.bollen)
summary(sem.bollen, robust=TRUE) # robust SEs and tests
summary(sem.bollen, analytic.se=FALSE) # uses numeric rather than analytic Hessian

sem.bollen.gls <- sem(model.bollen, data=Bollen, objective=objectiveGLS) # GLS rather than ML estimator
summary(sem.bollen.gls)

# The model in equation format:

model.bollen <- specifyEquations()
y1 = 1*Demo60
y2 = lam2*Demo60
y3 = lam3*Demo60
y4 = lam4*Demo60
y5 = 1*Demo65
y6 = lam2*Demo65
y7 = lam3*Demo65
y8 = lam4*Demo65
x1 = 1*Indust
x2 = lam6*Indust
x3 = lam7*Indust
c(y1, y5) = theta15
c(y2, y4) = theta24
c(y2, y6) = theta26
c(y3, y7) = theta37
c(y4, y8) = theta48
c(y6, y8) = theta68
Demo60 = gamma11*Indust

```

```

Demo65 = gamma21*Indust + beta21*Demo60
v(Indust) = phi

# ----- A simple CFA model for the Thurstone mental tests data -----

R.thur <- readMoments(diag=FALSE, names=c('Sentences','Vocabulary',
    'Sent.Completion','First.Letters','4.Letter.Words','Suffixes',
    'Letter.Series','Pedigrees', 'Letter.Group'))

.828
.776 .779
.439 .493 .46
.432 .464 .425 .674
.447 .489 .443 .59 .541
.447 .432 .401 .381 .402 .288
.541 .537 .534 .35 .367 .32 .555
.38 .358 .359 .424 .446 .325 .598 .452

# (1) in CFA format:

mod.cfa.thur.c <- cfa()
FA: Sentences, Vocabulary, Sent.Completion
FB: First.Letters, 4.Letter.Words, Suffixes
FC: Letter.Series, Pedigrees, Letter.Group

cfa.thur.c <- sem(mod.cfa.thur.c, R.thur, 213)
summary(cfa.thur.c)

# (2) in equation format:

mod.cfa.thur.e <- specifyEquations(covs="F1, F2, F3")
Sentences = lam11*F1
Vocabulary = lam21*F1
Sent.Completion = lam31*F1
First.Letters = lam42*F2
4.Letter.Words = lam52*F2
Suffixes = lam62*F2
Letter.Series = lam73*F3
Pedigrees = lam83*F3
Letter.Group = lam93*F3
V(F1) = 1
V(F2) = 1
V(F3) = 1

cfa.thur.e <- sem(mod.cfa.thur.e, R.thur, 213)
summary(cfa.thur.e)

# (3) in path format:

mod.cfa.thur.p <- specifyModel(covs="F1, F2, F3")
F1 -> Sentences, lam11
F1 -> Vocabulary, lam21
F1 -> Sent.Completion, lam31

```

```

F2 -> First.Letters,          lam41
F2 -> 4.Letter.Words,        lam52
F2 -> Suffixes,              lam62
F3 -> Letter.Series,         lam73
F3 -> Pedigrees,             lam83
F3 -> Letter.Group,          lam93
F1 <-> F1,                   NA,      1
F2 <-> F2,                   NA,      1
F3 <-> F3,                   NA,      1

cfa.thur.p <- sem(mod.cfa.thur.p, R.thur, 213)
summary(cfa.thur.p)

## End(Not run)

## =====

# The following examples use file input and may be executed via example():

etc <- file.path(.path.package(package="sem")[1], "etc") # path to data and model files

# ----- Duncan, Haller and Portes peer-influences model -----
# A nonrecursive SEM with unobserved endogenous variables and fixed exogenous variables

(R.DHP <- readMoments(file=file.path(etc, "R-DHP.txt"),
  diag=FALSE, names=c("ROccAsp", "REdAsp", "FOccAsp",
    "FEdAsp", "RParAsp", "RIQ", "RSES", "FSSES", "FIQ", "FParAsp")))
(model.dhp <- specifyModel(file=file.path(etc, "model-DHP.txt")))
sem.dhp.1 <- sem(model.dhp, R.DHP, 329,
  fixed.x=c('RParAsp', 'RIQ', 'RSES', 'FSSES', 'FIQ', 'FParAsp'))
summary(sem.dhp.1)

# ----- Wheaton et al. alienation data -----

(S.wh <- readMoments(file=file.path(etc, "S-Wheaton.txt"),
  names=c('Anomia67', 'Powerless67', 'Anomia71',
    'Powerless71', 'Education', 'SEI')))

# This is the model in the SAS manual for PROC CALIS: A Recursive SEM with
# latent endogenous and exogenous variables.
# Curiously, both factor loadings for two of the latent variables are fixed.

(model.wh.1 <- specifyModel(file=file.path(etc, "model-Wheaton-1.txt")))
sem.wh.1 <- sem(model.wh.1, S.wh, 932)
summary(sem.wh.1)

# The same model, but treating one loading for each latent variable as free
# (and equal to each other).

(model.wh.2 <- specifyModel(file=file.path(etc, "model-Wheaton-2.txt")))
sem.wh.2 <- sem(model.wh.2, S.wh, 932)
summary(sem.wh.2)

```

```

# Compare the two models by a likelihood-ratio test:

anova(sem.wh.1, sem.wh.2)

# ----- Thurstone data -----

# Second-order confirmatory factor analysis, from the SAS manual for PROC CALIS

(R.thur <- readMoments(file=file.path(etc, "R-Thurstone.txt"),
diag=FALSE, names=c('Sentences', 'Vocabulary',
                    'Sent.Completion', 'First.Letters', '4.Letter.Words', 'Suffixes',
                    'Letter.Series', 'Pedigrees', 'Letter.Group')))
(model.thur <- specifyModel(file=file.path(etc, "model-Thurstone.txt")))
sem.thur <- sem(model.thur, R.thur, 213)
summary(sem.thur)

#----- Kerchoff/Kenney path analysis -----

# An observed-variable recursive SEM from the LISREL manual

(R.kerch <- readMoments(file=file.path(etc, "R-Kerchoff.txt"),
diag=FALSE, names=c('Intelligence', 'Siblings',
                    'FatherEd', 'FatherOcc', 'Grades', 'EducExp', 'OccupAsp')))
(model.kerch <- specifyModel(file=file.path(etc, "model-Kerchoff.txt")))
sem.kerch <- sem(model.kerch, R.kerch, 737, fixed.x=c('Intelligence', 'Siblings',
                    'FatherEd', 'FatherOcc'))
summary(sem.kerch)

#----- McArdle/Epstein latent-growth-curve model -----

# This model, from McArdle and Epstein (1987, p.118), illustrates the use of a
# raw moment matrix to fit a model with an intercept. (The example was suggested
# by Mike Stoolmiller.)

(M.McArdle <- readMoments(file=file.path(etc, "M-McArdle.txt"),
names=c('WISC1', 'WISC2', 'WISC3', 'WISC4', 'UNIT')))
(mod.McArdle <- specifyModel(file=file.path(etc, "model-McArdle.txt")))
sem.McArdle <- sem(mod.McArdle, M.McArdle, 204, fixed.x="UNIT", raw=TRUE)
summary(sem.McArdle)

#----- Bollen industrialization and democracy example -----

# This model, from Bollen (1989, Ch. 8), illustrates the use in sem() of a
# case-by-variable data set (see ?Bollen) rather than a covariance or moment matrix

(model.bollen <- specifyModel(file=file.path(etc, "model-Bollen.txt")))
sem.bollen <- sem(model.bollen, data=Bollen)
summary(sem.bollen)

```

```
summary(sem.bollen, robust=TRUE) # robust SEs and tests
summary(sem.bollen, analytic.se=FALSE) # uses numeric rather than analytic Hessian

sem.bollen.gls <- sem(model.bollen, data=Bollen, objective=objectiveGLS) # GLS rather than ML estimator
summary(sem.bollen.gls)
```

sem-deprecated

Deprecated Functions in the sem Package

Description

These functions are provided for compatibility with older versions of the **sem** package only, and may be removed eventually. Although an effort has been made to insure backwards-compatibility, commands that worked in versions of the **sem** package prior to version 2.0-0 will not necessarily work in version 2.0-0 and beyond, or may not work in the same manner.

Usage

```
boot.sem(...)
mod.indices(...)
normalized.residuals(...)
path.diagram(...)
raw.moments(...)
read.moments(...)
specify.model(...)
standardized.coefficients(...)
standardized.residuals(...)
std.coef(...)
```

Arguments

... pass arguments down to replacements for deprecated functions.

Details

`boot.sem` is now a synonym for the `bootSem` function.

`mod.indices` is now a synonym for `modIndices`.

`normalized.residuals` is now a synonym for `normalizedResiduals`.

`path.diagram` is now a synonym for `pathDiagram`.

`raw.moments` is now a synonym for `rawMoments`.

`read.moments` is now a synonym for `readMoments`.

`specify.model` is now a synonym for `specifyModel`.

`standardized.coefficients` and `std.coef` are now synonyms for the `standardizedCoefficients` and `stdCoef` functions.

`standardized.residuals` is now a synonym for `standardizedResiduals`.

specifyModel

*Specify a Structural Equation Model***Description**

Create the RAM specification of a structural equation model.

Usage

```
specifyModel(file="", exog.variances=FALSE, endog.variances=TRUE, covs, quiet=FALSE)
specifyEquations(file="", ...)
cfa(file="", covs=paste(factors, collapse=","), reference.indicators=FALSE, ...)
classifyVariables(model)
removeRedundantPaths(model, warn=TRUE)
## S3 method for class 'semmod'
combineModels(..., warn=TRUE)
## S3 method for class 'semmod'
update(object, file = "", ...)
## S3 method for class 'semmod'
print(x, ...)
```

Arguments

| | |
|-----------------------------------|--|
| <code>file</code> | The (quoted) file from which to read the model specification, including the path to the file if it is not in the current directory. If "" (the default), then the specification is read from the standard input stream, and is terminated by a blank line. |
| <code>exog.variances</code> | If TRUE (the default is FALSE), free variance parameters are added for the exogenous variables that lack them. |
| <code>endog.variances</code> | If TRUE (the default), free error-variance parameters are added for the endogenous variables that lack them. |
| <code>covs</code> | optional: a character vector of one or more elements, with each element giving a string of variable names, separated by commas. Variances and covariances among all variables in each such string are added to the model. For confirmatory factor analysis models specified via <code>cfa</code> , <code>covs</code> defaults to all of the factors in the model, thus specifying all covariances among these factors. |
| <code>reference.indicators</code> | if FALSE, the default, variances of factors are set to 1 by <code>cfa</code> ; if TRUE, variances of factors are free parameters to estimate from the data, and instead the first factor loading for each factor is set to 1 to identify the model. |

| | |
|------------------|--|
| quiet | if FALSE, the default, then the number of input lines is reported. |
| x, model, object | An object of class semmod, as produced by specifyModel. |
| warn | print a warning if redundant paths are detected. |
| ... | for specifyEquations and cfa, arguments (such as covs, in the case of specifyEquations) to be passed to specifyModel; for combineModels, sem objects; ignored in the update and print methods. |

Details

The principal functions for model specification are `specifyModel`, to specify a model in RAM (path) format via single- and double-headed arrows; `specifyEquations`, to specify a model in equation format, which is then translated by the function into RAM format; and `cfa`, for compact specification of simple confirmatory factor analysis models.

`specifyModel`:

Each line of the RAM specification for `specifyModel` consists of three (unquoted) entries, separated by commas:

- 1. Arrow specification:** This is a simple formula, of the form $A \rightarrow B$ or, equivalently, $B \leftarrow A$ for a regression coefficient (i.e., a single-headed or directional arrow); $A \leftrightarrow A$ for a variance or $A \leftrightarrow B$ for a covariance (i.e., a double-headed or bidirectional arrow). Here, A and B are variable names in the model. If a name does not correspond to an observed variable, then it is assumed to be a latent variable. Spaces can appear freely in an arrow specification, and there can be any number of hyphens in the arrows, including zero: Thus, e.g., $A \rightarrow B$, $A \rightarrow B$, and $A \rightarrow B$ are all legitimate and equivalent.
- 2. Parameter name:** The name of the regression coefficient, variance, or covariance specified by the arrow. Assigning the same name to two or more arrows results in an equality constraint. Specifying the parameter name as NA produces a fixed parameter.
- 3. Value:** start value for a free parameter or value of a fixed parameter. If given as NA (or simply omitted), sem will compute the start value.

Lines may end in a comment following #.

`specifyEquations`:

For `specifyEquations`, each input line is either a regression equation or the specification of a variance or covariance. Regression equations are of the form

$$y = \text{par1} * x_1 + \text{par2} * x_2 + \dots + \text{park} * x_k$$

where y and the xs are variables in the model (either observed or latent), and the pars are parameters. If a parameter is given as a numeric value (e.g., 1) then it is treated as fixed. Note that no “error” variable is included in the equation; “error variances” are specified via either the covs argument, via $V(y) = \text{par}$ (see immediately below), or are added automatically to the model when, as by default, `endog.variances=TRUE`.

Variances are specified in the form $V(\text{var}) = \text{par}$ and covariances in the form $C(\text{var1}, \text{var2}) = \text{par}$, where the vars are variables (observed or unobserved) in the model. The symbols V and C may be in either lower- or upper-case. If par is a numeric value (e.g., 1) then it is treated as fixed. In conformity with the RAM model, a variance or covariance for an endogenous variable in the model is an “error” variance or covariance.

To set a start value for a free parameter, enclose the numeric start value in parentheses after the parameter name, as `parameter(value)`.

`cfa`:

For `cfa`, each input line includes the names of the variables, separated by commas, that load on the corresponding factor; the name of the factor is given optionally at the beginning of the line, followed by a colon. If necessary, the variables that load on a factor may be continued across two or more input lines; in this case, each such line but the last must end in a comma. A variable may load on more than one factor (as long as the resulting model is identified, of course), but each factor may appear in only one input line (or set of input lines, if the variable list is continued onto the next line). If the argument `reference.indicators=FALSE`, the default, `cfa` will fix the variance of each factor to 1, and by default include covariances (i.e., correlations) among all pairs of factors. Alternatively, if `reference.indicators=TRUE`, then the factor variances are free parameters to be estimated from the data, and the first loading for each factor is set to 1 to identify the model. These two approaches produce equivalent models, with the same fit to the data, but alternative parametrizations. Specifying the argument `covs=NULL` implicitly fixes the factor intercorrelations to 0.

See [sem](#) and the examples for further details on model specification.

Other Functions:

`classifyVariables` classifies the variables in a model as endogenous or exogenous.

`combineModels` and `removeRedundantPaths` take `semmod` objects as arguments and do what their names imply.

The file input argument to the `update` method for `semmod` objects, which by default comes from standard input, is a set of update directives, one per line. There are three kinds of directives. In each case the directive begins with the directive name, followed by one or more fields separated by commas.

- 1. delete:** Remove a path from the model. Example: `delete, RSES -> FGenAsp`
- 2. add:** Add a path to the model. Example: `add, RSES -> FGenAsp, gam14, NA`
- 3. replace:** Replace every occurrence of the first string with the second in the variables and parameters of the model. This directive may be used, for example, to change one variable to another or to rename a parameter. Example: `replace, gam, gamma`, substitutes the string "gamma" for "gam" wherever the latter appears, presumably in parameter names.

Value

`specifyModel`, `specifyEquations`, `cfa`, `removeRedundantPaths`, `combineModels`, and `update` return an object of class `semmod`, suitable as input for [sem](#).

`classifyVariables` returns a list with two character vectors: `endogenous`, containing the names of endogenous variables in the model; and `exogenous`, containing the names of exogenous variables.

Author(s)

John Fox <jfox@mcmaster.ca> and Jarrett Byrnes

See Also

[sem](#)

Examples

```
# Note: These examples can't be run via example() because the default file
# argument of specifyModel() requires that the model specification be entered
# at the command prompt. The examples can be copied and run in the R console,
# however.
```

```
## Not run:
model.dhp <- specifyModel()
RParAsp -> RGenAsp, gam11, NA
RIQ      -> RGenAsp, gam12, NA
RSES     -> RGenAsp, gam13, NA
FSES     -> RGenAsp, gam14, NA
RSES     -> FGenAsp, gam23, NA
FSES     -> FGenAsp, gam24, NA
FIQ      -> FGenAsp, gam25, NA
FParAsp  -> FGenAsp, gam26, NA
FGenAsp  -> RGenAsp, beta12, NA
RGenAsp  -> FGenAsp, beta21, NA
RGenAsp  -> ROccAsp, NA, 1
RGenAsp  -> REdAsp, lam21, NA
FGenAsp  -> FOccAsp, NA, 1
FGenAsp  -> FEdAsp, lam42, NA
RGenAsp <-> RGenAsp, ps11, NA
FGenAsp <-> FGenAsp, ps22, NA
RGenAsp <-> FGenAsp, ps12, NA
ROccAsp <-> ROccAsp, theta1, NA
REdAsp <-> REdAsp, theta2, NA
FOccAsp <-> FOccAsp, theta3, NA
FEdAsp <-> FEdAsp, theta4, NA
```

```
model.dhp
```

```
# an equivalent specification, allowing specifyModel() to generate
# variance parameters for endogenous variables (and suppressing
# the unnecessary trailing NAs):
```

```
model.dhp <- specifyModel()
RParAsp -> RGenAsp, gam11
RIQ      -> RGenAsp, gam12
RSES     -> RGenAsp, gam13
FSES     -> RGenAsp, gam14
RSES     -> FGenAsp, gam23
FSES     -> FGenAsp, gam24
FIQ      -> FGenAsp, gam25
FParAsp  -> FGenAsp, gam26
FGenAsp  -> RGenAsp, beta12
RGenAsp  -> FGenAsp, beta21
RGenAsp  -> ROccAsp, NA, 1
RGenAsp  -> REdAsp, lam21
FGenAsp  -> FOccAsp, NA, 1
FGenAsp  -> FEdAsp, lam42
```

```

RGenAsp <-> FGenAsp, ps12

model.dhp

# Another equivalent specification, telling specifyModel to add paths for
# variances and covariance of RGenAsp and FGenAsp:

model.dhp <- specifyModel(covs="RGenAsp, FGenAsp")
RParAsp -> RGenAsp, gam11
RIQ      -> RGenAsp, gam12
RSES     -> RGenAsp, gam13
FSES     -> RGenAsp, gam14
RSES     -> FGenAsp, gam23
FSES     -> FGenAsp, gam24
FIQ      -> FGenAsp, gam25
FParAsp  -> FGenAsp, gam26
FGenAsp  -> RGenAsp, beta12
RGenAsp  -> FGenAsp, beta21
RGenAsp  -> ROccAsp, NA,      1
RGenAsp  -> REdAsp, lam21
FGenAsp  -> FOccAsp, NA,      1
FGenAsp  -> FEdAsp, lam42

model.dhp

# The same model in equation format:

model.dhp.1 <- specifyEquations(covs="RGenAsp, FGenAsp")
RGenAsp = gam11*RParAsp + gam12*RIQ + gam13*RSES + gam14*FSES + beta12*FGenAsp
FGenAsp = gam23*RSES + gam24*FSES + gam25*FIQ + gam26*FParAsp + beta21*RGenAsp
ROccAsp = 1*RGenAsp
REdAsp = lam21(1)*RGenAsp # to illustrate setting start values
FOccAsp = 1*FGenAsp
FEdAsp = lam42(1)*FGenAsp

model.dhp

classifyVariables(model.dhp)

# updating the model to impose equality constraints
# and to rename the latent variables and gamma parameters

model.dhp.eq <- update(model.dhp)
delete, RSES -> FGenAsp
delete, FSES -> FGenAsp
delete, FIQ  -> FGenAsp
delete, FParAsp -> FGenAsp
delete, RGenAsp -> FGenAsp
add, RSES      -> FGenAsp, gam14, NA
add, FSES      -> FGenAsp, gam13, NA
add, FIQ       -> FGenAsp, gam12, NA
add, FParAsp   -> FGenAsp, gam26, NA
add, RGenAsp   -> FGenAsp, beta12, NA

```

```

replace, gam, gamma
replace, Gen, General

model.dhp.eq

# A three-factor CFA model for the Thurstone mental-tests data, specified three equivalent ways:

R.thur <- readMoments(diag=FALSE, names=c('Sentences', 'Vocabulary',
    'Sent.Completion', 'First.Letters', '4.Letter.Words', 'Suffixes',
    'Letter.Series', 'Pedigrees', 'Letter.Group'))

.828
.776 .779
.439 .493 .46
.432 .464 .425 .674
.447 .489 .443 .59 .541
.447 .432 .401 .381 .402 .288
.541 .537 .534 .35 .367 .32 .555
.38 .358 .359 .424 .446 .325 .598 .452

# (1a) in CFA format:

mod.cfa.thur.c <- cfa()
FA: Sentences, Vocabulary, Sent.Completion
FB: First.Letters, 4.Letter.Words, Suffixes
FC: Letter.Series, Pedigrees, Letter.Group

cfa.thur.c <- sem(mod.cfa.thur.c, R.thur, 213)
summary(cfa.thur.c)

# (1b) in CFA format, using reference indicators:

mod.cfa.thur.r <- cfa(reference.indicators=TRUE)
FA: Sentences, Vocabulary, Sent.Completion
FB: First.Letters, 4.Letter.Words, Suffixes
FC: Letter.Series, Pedigrees, Letter.Group

cfa.thur.r <- sem(mod.cfa.thur.r, R.thur, 213)
summary(cfa.thur.r)

# (2) in equation format:

mod.cfa.thur.e <- specifyEquations(covs="F1, F2, F3")
Sentences = lam11*F1
Vocabulary = lam21*F1
Sent.Completion = lam31*F1
First.Letters = lam42*F2
4.Letter.Words = lam52*F2
Suffixes = lam62*F2
Letter.Series = lam73*F3
Pedigrees = lam83*F3
Letter.Group = lam93*F3
V(F1) = 1
V(F2) = 1

```

```

V(F3) = 1

cfa.thur.e <- sem(mod.cfa.thur.e, R.thur, 213)
summary(cfa.thur.e)

# (3) in path format:

mod.cfa.thur.p <- specifyModel(covs="F1, F2, F3")
F1 -> Sentences,           lam11
F1 -> Vocabulary,         lam21
F1 -> Sent.Completion,     lam31
F2 -> First.Letters,       lam41
F2 -> 4.Letter.Words,     lam52
F2 -> Suffixes,           lam62
F3 -> Letter.Series,       lam73
F3 -> Pedigrees,          lam83
F3 -> Letter.Group,        lam93
F1 <-> F1,                 NA,      1
F2 <-> F2,                 NA,      1
F3 <-> F3,                 NA,      1

cfa.thur.p <- sem(mod.cfa.thur.p, R.thur, 213)
summary(cfa.thur.p)

## End(Not run)

```

standardizedCoefficients

Standardized Coefficients for Structural Equation Models

Description

These functions calculate standardized regression coefficients for structural equation models. The function `stdCoef` is simply an abbreviation for `standardizedCoefficients`.

Usage

```
standardizedCoefficients(object, digits = 5, oneheaded=TRUE, twoheaded=TRUE)
```

```
stdCoef(...)
```

Arguments

| | |
|------------------------|--|
| <code>object</code> | an object of class <code>sem</code> returned by the <code>sem</code> function. |
| <code>digits</code> | number of digits for printed output. |
| <code>oneheaded</code> | standardize path coefficients? Default is TRUE. |
| <code>twoheaded</code> | standardize variances and covariances? Default is TRUE. |
| <code>...</code> | arguments to pass to <code>standardized.coefficients</code> . |

Value

Returns a data frame with the coefficients, labelled both by parameter names and by arrows in the path diagram for the model.

Author(s)

John Fox <jfox@mcmaster.ca> and Adam Kramer

References

Bollen, K. A. (1989) *Structural Equations With Latent Variables*. Wiley.

See Also

[sem](#)

Examples

```
# In the first example, readMoments() and specifyModel() read from the
# input stream. This example cannot be executed via example() but can be entered
# at the command prompt. The example is repeated using file input;
# this example can be executed via example().
## Not run:
```

```
# Duncan, Haller, and Portes peer-influences model
```

```
R.DHP <- readMoments(diag=FALSE, names=c("ROccAsp", "REdAsp", "FOccAsp",
    "FEdAsp", "RParAsp", "RIQ", "RSES", "FSES", "FIQ", "FParAsp"))
```

```
.6247
.3269 .3669
.4216 .3275 .6404
.2137 .2742 .1124 .0839
.4105 .4043 .2903 .2598 .1839
.3240 .4047 .3054 .2786 .0489 .2220
.2930 .2407 .4105 .3607 .0186 .1861 .2707
.2995 .2863 .5191 .5007 .0782 .3355 .2302 .2950
.0760 .0702 .2784 .1988 .1147 .1021 .0931 -.0438 .2087
```

```
model.dhp <- specifyModel()
  RParAsp -> RGenAsp, gam11, NA
  RIQ      -> RGenAsp, gam12, NA
  RSES     -> RGenAsp, gam13, NA
  FSES     -> RGenAsp, gam14, NA
  RSES     -> FGenAsp, gam23, NA
  FSES     -> FGenAsp, gam24, NA
  FIQ      -> FGenAsp, gam25, NA
  FParAsp -> FGenAsp, gam26, NA
  FGenAsp -> RGenAsp, beta12, NA
  RGenAsp -> FGenAsp, beta21, NA
  RGenAsp -> ROccAsp, NA, 1
  RGenAsp -> REdAsp, lam21, NA
  FGenAsp -> FOccAsp, NA, 1
  FGenAsp -> FEdAsp, lam42, NA
```

```

RGenAsp <-> RGenAsp, ps11, NA
FGenAsp <-> FGenAsp, ps22, NA
RGenAsp <-> FGenAsp, ps12, NA
ROccAsp <-> ROccAsp, theta1, NA
REdAsp <-> REdAsp, theta2, NA
FOccAsp <-> FOccAsp, theta3, NA
FEdAsp <-> FEdAsp, theta4, NA

sem.dhp <- sem(model.dhp, R.DHP, 329,
  fixed.x=c('RParAsp', 'RIQ', 'RSES', 'FSES', 'FIQ', 'FParAsp'))
standardizedCoefficients(sem.dhp)

## End(Not run)
# The following example can be executed via example():

etc <- file.path(.path.package(package="sem")[1], "etc") # path to data and model files

(R.DHP <- readMoments(file=file.path(etc, "R-DHP.txt"),
  diag=FALSE, names=c("ROccAsp", "REdAsp", "FOccAsp",
    "FEdAsp", "RParAsp", "RIQ", "RSES", "FSES", "FIQ", "FParAsp")))
(model.dhp <- specifyModel(file=file.path(etc, "model-DHP.txt")))
(sem.dhp <- sem(model.dhp, R.DHP, 329,
  fixed.x=c('RParAsp', 'RIQ', 'RSES', 'FSES', 'FIQ', 'FParAsp')))
standardizedCoefficients(sem.dhp)

```

 tsls

Two-Stage Least Squares

Description

Fits a regression equation, such as an equation in a structural-equation model, by two-stage least squares. This is equivalent to direct instrumental-variables estimation when the number of instruments is equal to the number of predictors.

Usage

```

## S3 method for class 'formula'
tsls(formula, instruments, data, subset, na.action, contrasts=NULL, ...)
## Default S3 method:
tsls(y, X, Z, names=NULL, ...)

## S3 method for class 'tsls'
print(x, ...)
## S3 method for class 'tsls'
summary(object, digits=4, ...)
## S3 method for class 'tsls'
anova(object, model.2, s2, dfe, ...)

## S3 method for class 'tsls'

```

```
fitted(object, ...)
## S3 method for class 'tsls'
residuals(object, ...)
## S3 method for class 'tsls'
coef(object, ...)
## S3 method for class 'tsls'
vcov(object, ...)
```

Arguments

| | |
|--------------------|---|
| formula | model formula for structural equation to be estimated; a regression constant is implied if not explicitly omitted. |
| instruments | one-sided model formula specifying instrumental variables. |
| data | an optional data frame containing the variables in the model. By default the variables are taken from the environment from which <code>tsls</code> is called. |
| subset | an optional vector specifying a subset of observations to be used in fitting the model. |
| na.action | a function that indicates what should happen when the data contain NAs. The default is set by the <code>na.action</code> option. |
| contrasts | an optional list. See the <code>contrasts.arg</code> argument of model.matrix.default . |
| y | Response-variable vector. |
| X | Matrix of predictors, including a constant (if one is in the model). |
| Z | Matrix of instrumental variables, including a constant (if one is in the model). |
| names | optional character vector of names for the columns of the X matrix. |
| x, object, model.2 | objects of class <code>tsls</code> returned by <code>tsls.formula</code> , for anova containing nested models to be compared by an incremental F -test. One model should be nested in the other; the order of models is immaterial. |
| s2 | an optional estimate of error variance for the denominator of the F -test. If missing, the error-variance estimate is taken from the larger model. |
| dfe | optional error degrees of freedom, to be specified when an estimate of error variance is given. |
| digits | number of digits for summary output. |
| ... | arguments to be passed down. |

Value

`tsls.formula` returns an object of class `tsls`, with the following components:

| | |
|--------------|--|
| n | number of observations. |
| p | number of parameters. |
| coefficients | parameter estimates. |
| V | estimated covariance matrix of coefficients. |
| s | residual standard error. |

| | |
|---------------|--|
| residuals | vector of residuals. |
| response | vector of response values. |
| X | model matrix. |
| Z | instrumental-variables matrix. |
| response.name | name of response variable, or expression evaluating to response. |
| formula | model formula. |
| instruments | one-sided formula for instrumental variables. |

Author(s)

John Fox <jfox@mcmaster.ca>

References

- Fox, J. (1979) Simultaneous equation models and two-stage least-squares. In Schuessler, K. F. (ed.) *Sociological Methodology 1979*, Jossey-Bass.
- Greene, W. H. (1993) *Econometric Analysis, Second Edition*, Macmillan.

See Also

[sem](#)

Examples

```
summary(tsls(Q ~ P + D, ~ D + F + A, data=Kmenta)) # demand equation
summary(tsls(Q ~ P + F + A, ~ D + F + A, data=Kmenta)) # supply equation
anova(tsls(Q ~ P + F + A, ~ D + F + A, data=Kmenta),
      tsls(Q ~ 1, ~ D + F + A, data=Kmenta))
```

Index

- *Topic **datasets**
 - Bollen, 2
 - CNES, 6
 - Klein, 12
 - Kmenta, 13
- *Topic **dplot**
 - pathDiagram, 20
- *Topic **htest**
 - bootSem, 3
- *Topic **manip**
 - fscores, 9
 - rawMoments, 24
 - readMoments, 25
- *Topic **models**
 - bootSem, 3
 - effects.sem, 7
 - fscores, 9
 - information.criteria, 11
 - ML.methods, 14
 - modIndices, 15
 - objective.functions, 18
 - optimizers, 19
 - pathDiagram, 20
 - ram, 23
 - residuals.sem, 26
 - sem, 29
 - specifyModel, 48
 - standardizedCoefficients, 54
 - tsls, 56
- AIC.objectiveML (ML.methods), 14
- AICc (information.criteria), 11
- AICc.objectiveML, 12
- AICc.objectiveML (ML.methods), 14
- anova.objectiveML (ML.methods), 14
- anova.tsls (tsls), 56
- BIC.objectiveML (ML.methods), 14
- Bollen, 2
- boot, 4, 5
- boot.ci, 4
- boot.sem (sem-deprecated), 47
- bootSem, 3, 47
- CAIC (information.criteria), 11
- CAIC.objectiveML, 12
- CAIC.objectiveML (ML.methods), 14
- cfa, 29, 31, 32, 35
- cfa (specifyModel), 48
- classifyVariables (specifyModel), 48
- CNES, 6
- coef.sem (sem), 29
- coef.tsls (tsls), 56
- combineModels (specifyModel), 48
- cor, 4
- cov, 4
- cov2raw (rawMoments), 24
- deviance.objectiveML (ML.methods), 14
- df.residual.sem (sem), 29
- effects.sem, 7
- fitted.tsls (tsls), 56
- fscores, 9
- GLS.methods (ML.methods), 14
- hetcor, 4
- information.criteria, 11
- Klein, 12
- Kmenta, 13
- logLik.objectiveML (ML.methods), 14
- ML.methods, 14
- mod.indices (sem-deprecated), 47
- model.matrix.default, 24, 57
- modIndices, 15, 47

- modIndices.objectiveML, 15
- nlm, 19, 20, 31, 35
- nlminb, 19, 20, 31, 35
- normalized.residuals (sem-deprecated), 47
- normalizedResiduals, 47
- normalizedResiduals (residuals.sem), 26
- objective.functions, 15, 18, 19, 20
- objectiveGLS, 14, 31, 35
- objectiveGLS (objective.functions), 18
- objectiveML, 14, 31, 35
- objectiveML (objective.functions), 18
- optim, 19, 20, 31, 35
- optimizerNlm, 31, 35
- optimizerNlm (optimizers), 19
- optimizerNlminb, 31, 35
- optimizerNlminb (optimizers), 19
- optimizerOptim, 31, 35
- optimizerOptim (optimizers), 19
- optimizers, 18, 19
- path.diagram (sem-deprecated), 47
- pathDiagram, 20, 47
- print.bootsem (bootSem), 3
- print.modIndices (modIndices), 15
- print.objectiveGLS (ML.methods), 14
- print.objectiveML (ML.methods), 14
- print.rawmoments (rawMoments), 24
- print.semeffects (effects.sem), 7
- print.semmod (specifyModel), 48
- print.summary.bootsem (bootSem), 3
- print.summary.objectiveML (ML.methods), 14
- print.tsls (tsls), 56
- ram, 23
- raw.moments (sem-deprecated), 47
- rawMoments, 24, 30, 35, 47
- read.moments (sem-deprecated), 47
- readMoments, 25, 47
- removeRedundantPaths (specifyModel), 48
- residuals.sem, 26
- residuals.tsls (tsls), 56
- scale, 10
- sem, 3–5, 7, 9, 10, 14–16, 18–20, 22, 23, 25–27, 29, 50, 54, 55, 58
- sem-deprecated, 47
- specify.model (sem-deprecated), 47
- specifyEquations, 29, 31, 32, 35
- specifyEquations (specifyModel), 48
- specifyModel, 29, 31, 32, 34, 35, 47, 48
- standardized.coefficients (sem-deprecated), 47
- standardized.residuals (sem-deprecated), 47
- standardizedCoefficients, 47, 54
- standardizedResiduals, 47
- standardizedResiduals (residuals.sem), 26
- startvalues, 35
- startvalues (sem), 29
- std.coef (sem-deprecated), 47
- stdCoef, 47
- stdCoef (standardizedCoefficients), 54
- summary.bootsem (bootSem), 3
- summary.modIndices (modIndices), 15
- summary.objectiveGLS (ML.methods), 14
- summary.objectiveML (ML.methods), 14
- summary.tsls (tsls), 56
- tsls, 56
- update.semmod (specifyModel), 48
- vcov.sem (sem), 29
- vcov.tsls (tsls), 56