

# Package: cvsem (via r-universe)

September 11, 2024

**Title** SEM Model Comparison with K-Fold Cross-Validation

**Version** 1.0.0

**Description** The goal of 'cvsem' is to provide functions that allow for comparing Structural Equation Models (SEM) using cross-validation. Users can specify multiple SEMs using 'lavaan' syntax. 'cvsem' computes the Kullback Leibler (KL) Divergence between 1) the model implied covariance matrix estimated from the training data and 2) the sample covariance matrix estimated from the test data described in Cudeck, Robert & Browne (1983) <[doi:10.18637/jss.v048.i02](https://doi.org/10.18637/jss.v048.i02)>. The KL Divergence is computed for each of the specified SEMs allowing for the models to be compared based on their prediction errors.

**License** GPL (>=3)

**Encoding** UTF-8

**Roxygen** list(markdown = TRUE)

**RoxygenNote** 7.2.0

**Imports** lavaan, stats, Rdpack

**RdMacros** Rdpack

**Suggests** testthat (>= 3.0.0)

**Config/testthat/edition** 3

**Repository** <https://annawyssocki.r-universe.dev>

**RemoteUrl** <https://github.com/annawyssocki/cvsem>

**RemoteRef** HEAD

**RemoteSha** 905604afbe895d922e69c1bf33e4d23ebc8da120

## Contents

cvsem . . . . .	2
<b>Index</b>	<b>4</b>

**Description**

Do model comparison on SEM models using cross-validation as described in (Cudeck and Browne 1983) and (Browne and Cudeck 1992). Cross-validation is based on the discrepancy between the sample covariance matrix and the model implied matrix. Currently, `cvsem` supports 'KL-Divergence', Frobenius Distance and Generalized Least Squares 'GLS' as discrepancy metrics.

**Usage**

```
cvsem(
  data = NULL,
  Models,
  discrepancyMetric = "KL-Divergence",
  k = 5,
  lavaanFunction = "sem",
  echo = TRUE,
  ...
)
```

**Arguments**

<code>data</code>	Data
<code>Models</code>	A collection of models, specified in lavaan syntax. Provide Models with the <code>cvgather()</code> function.
<code>discrepancyMetric</code>	Specify which discrepancy metric to use (one of 'KL-Divergence', 'FD', 'GLS'). Default is KL Divergence.
<code>k</code>	The number of folds. Default is 5.
<code>lavaanFunction</code>	Specify which lavaan function to use. Default is "sem". Other options are "lavaan" and "cfa"
<code>echo</code>	Provide feedback on progress to user, defaults to TRUE. Set to FALSE to suppress.
<code>...</code>	Not used

**Value**

A list with the prediction error for each model.

**References**

Browne MW, Cudeck R (1992). "Alternative Ways of Assessing Model Fit." *Sociological Methods & Research*, **21**, 230–258.

Cudeck R, Browne MW (1983). "Cross-Validation Of Covariance Structures." *Multivariate Behavioral Research*, **18**, 147–167. doi:10.1207/s15327906mbr1802\_2, [https://www.tandfonline.com/doi/abs/10.1207/s15327906mbr1802\\_2](https://www.tandfonline.com/doi/abs/10.1207/s15327906mbr1802_2).

### Examples

```
example_data <- lavaan::HolzingerSwineford1939
colnames(example_data) <- c("id", "sex", "ageyr", "agemo", 'school', "grade",
"visualPerception", "cubes", "lozenges", "comprehension",
"sentenceCompletion", "wordMeaning", "speededAddition",
"speededCounting", "speededDiscrimination")

model1 <- 'comprehension ~ meaning

      ## Add some latent variables:
      meaning =~ wordMeaning + sentenceCompletion
      speed =~ speededAddition + speededDiscrimination + speededCounting
      speed ~~ meaning'

model2 <- 'comprehension ~ wordMeaning + speededAddition'
model3 <- 'comprehension ~ wordMeaning + speededAddition'

models <- cvgather(model1, model2, model3)

fit <- cvsem( data = example_data, Models = models, k = 10, discrepancyMetric = "KL-Divergence")
```

# Index

cvsem, [2](#)