Package 'TSdisaggregation'

January 20, 2025

Title High-Dimensional Temporal Disaggregation

Version 2.0.0

Description First - Generates (potentially high-dimensional) high-frequency and low-frequency series for simulation studies in temporal disaggregation; Second - a toolkit utilizing temporal disaggregation and benchmarking techniques with a low-dimensional matrix of indicator series previously proposed in Dagum and Cholette (2006, ISBN:978-0-387-35439-2) ; and Third - novel techniques proposed by Mosley, Gibberd and Eckley (2021) <arXiv:2108.05783> for disaggregating low-frequency series in the presence of high-dimensional indicator matrices.

Imports Rdpack, zoo, lars, Matrix, withr

RdMacros Rdpack

License GPL (>= 3)

Encoding UTF-8

RoxygenNote 7.1.2

NeedsCompilation no

Author Luke Mosley [aut, cre],
Kaveh S. Nobari [aut] (<https://orcid.org/0000-0002-4053-0781>)

Maintainer Luke Mosley <l.mosley@lancaster.ac.uk>

Repository CRAN

Date/Publication 2022-05-18 19:00:05 UTC

Contents

`chowlin` *Function to do Chow-Lin temporal disaggregation from Chow and Lin (1971) and Litterman.*

Description

Used in disaggregation.R to find estimates given the optimal rho parameter.

Usage

```
chowlin(Y, X, rho, aggMat, aggRatio, litterman = FALSE)
```

Arguments

`Y`	The low-frequency response series (n_l x 1 matrix).
`X`	The high-frequency indicator series (n x p matrix).
`rho`	The AR(1) residual parameter (strictly between -1 and 1).
`aggMat`	Aggregation matrix according to 'first', 'sum', 'average', 'last' (default is 'sum').
`aggRatio`	Aggregation ratio e.g. 4 for annual-to-quarterly, 3 for quarterly-to-monthly (default is 4).
`litterman`	TRUE to use litterman vcov. FALSE for Chow-Lin vcov. Default is FALSE.

Value

y Estimated high-frequency response series (n x 1 matrix).

betaHat Estimated coefficient vector (p x 1 matrix).

u_l Estimated aggregate residual series (n_l x 1 matrix).

References

Chow GC, Lin A (1971). "Best linear unbiased interpolation, distribution, and extrapolation of time series by related series." *The review of Economics and Statistics*, 372–375.

`chowlin_likelihood` *Likelihood function from Chow-Lin or Litterman temporal disaggregation.*

Description

Used in disaggregation.R to find estimates of the optimal rho parameter.

Usage

```
chowlin_likelihood(Y, X, vcov)
```

Arguments

Y	The low-frequency response series (n_l x 1 matrix).
X	The aggregated high-frequency indicator series (n_l x p matrix).
vcov	Aggregated variance-covariance matrix of Chow-Lin or Litterman residuals.

References

There are no references for Rd macro `\insertAllCites` on this help page.

---

disaggregate *Temporal Disaggregation Methods*

Description

This function contains the traditional standard-dimensional temporal disaggregation methods proposed by Denton (1971), Dagum and Cholette (2006), Chow and Lin (1971), Fernandez (1981) and Litterman (1983), and the high-dimensional methods of Mosley et al. (2021).

Usage

```
disaggregate(
  Y,
  X = matrix(data = rep(1, times = nrow(Y)), nrow = nrow(Y)),
  aggMat = "sum",
  aggRatio = 4,
  method = "Chow-Lin",
  Denton = "first"
)
```

Arguments

Y	The low-frequency response series (n_l x 1 matrix).
X	The high-frequency indicator series (n x p matrix).
aggMat	Aggregation matrix according to 'first', 'sum', 'average', 'last' (default is 'sum').
aggRatio	Aggregation ratio e.g. 4 for annual-to-quarterly, 3 for quarterly-to-monthly (default is 4).
method	Disaggregation method using 'Denton', 'Denton-Cholette', 'Chow-Lin', 'Fernandez', 'Litterman', 'spTD' or 'adaptive-spTD' (default is 'Chow-Lin').
Denton	Type of differencing for Denton method: 'absolute', 'first', 'second' and 'proportional' (default is 'first').

Details

Takes in a n_l x 1 low-frequency series to be disaggregated Y and a n x p high-frequency matrix of p indicator series X. If n > n_l x aggRatio where aggRatio is the aggregation ration (e.g. aggRatio = 4 if annual-to-quarterly disagg or aggRatio = 3 if quarterly-to-monthly disagg) then extrapolation is done to extrapolate up to n.

Value

y_Est Estimated high-frequency response series (n x 1 matrix).

beta_Est Estimated coefficient vector (p x 1 matrix).

rho_Est Estimated residual AR(1) autocorrelation parameter.

ul_Est Estimated aggregate residual series (n_l x 1 matrix).

References

Chow GC, Lin A (1971). "Best linear unbiased interpolation, distribution, and extrapolation of time series by related series." *The review of Economics and Statistics*, 372–375.

Dagum EB, Cholette PA (2006). *Benchmarking, temporal distribution, and reconciliation methods for time series*, volume 186. Springer Science \& Business Media.

Denton FT (1971). "Adjustment of monthly or quarterly series to annual totals: an approach based on quadratic minimization." *Journal of the american statistical association*, **66**(333), 99–102.

Fernandez RB (1981). "A methodological note on the estimation of time series." *The Review of Economics and Statistics*, **63**(3), 471–476.

Litterman RB (1983). "A random walk, Markov model for the distribution of time series." *Journal of Business \& Economic Statistics*, **1**(2), 169–173.

Mosley L, Eckley I, Gibberd A (2021). "Sparse Temporal Disaggregation." *arXiv preprint arXiv:2108.05783*.

Examples

```
data = TempDisaggDGP(n_l=25,n=100,p=10,rho=0.5)
X = data$X_Gen
Y = data$Y_Gen
fit_chowlin = disaggregate(Y=Y,X=X,method='Chow-Lin')
y_hat = fit_chowlin$y_Est
```

`sptd` *Function to do sparse temporal disaggregation from Mosley et al. (2021).*

Description

Used in disaggregation.R to find estimates given the optimal rho parameter.

Usage

```
sptd(Y, X, rho, aggMat, aggRatio, adaptive = FALSE)
```

Arguments

Y	The low-frequency response series (n_l x 1 matrix).
X	The high-frequency indicator series (n x p matrix).
rho	The AR(1) residual parameter (strictly between -1 and 1).
aggMat	Aggregation matrix according to 'first', 'sum', 'average', 'last' (default is 'sum').
aggRatio	Aggregation ratio e.g. 4 for annual-to-quarterly, 3 for quarterly-to-monthly (default is 4).
adaptive	TRUE to use adaptive lasso penalty. FALSE for lasso penalty. Default is FALSE.

Value

y Estimated high-frequency response series (n x 1 matrix).

betaHat Estimated coefficient vector (p x 1 matrix).

u_l Estimated aggregate residual series (n_l x 1 matrix).

References

Mosley L, Eckley I, Gibberd A (2021). "Sparse Temporal Disaggregation." *arXiv preprint arXiv:2108.05783*.

---

sptd_BIC *Function to calculate the BIC score from sparse temporal disaggregation.*

Description

Used in disaggregation.R to find estimates of the optimal rho parameter.

Usage

```
sptd_BIC(Y, X, vcov)
```

Arguments

Y	The low-frequency response series (n_l x 1 matrix).
X	The aggregated high-frequency indicator series (n_l x p matrix).
vcov	Aggregated variance-covariance matrix of AR(1) residuals.

References

There are no references for Rd macro `\insertAllCites` on this help page.

`TempDisaggDGP` *High and Low-Frequency Data Generating Processes*

Description

This function generates the high-frequency $n \times 1$ response vector y, according to $y = X\beta + \epsilon$, where X is an $n \times p$ matrix of indicator series, and the $p \times 1$ coefficient vector may be sparse. The low-frequency $n_l \times 1$ vector Y can be generated by pre-multiplying an aggregation matrix $n_l \times n$ matrix, such that the sum, the average, the last or the first value of y equates the corresponding Y observation. The parameter aggRatio is the specified aggregation ratio between the low and high frequency series, e.g. aggRatio = 4 for annual-to-quarterly and aggRatio = 3 for quarterly-to-monthly. If $n > aggRatio \times n_l$, then the last $n - aggRatio \times n_l$ columns of the aggregation matrix are 0 such that Y is only observed up to n_l. For a comprehensive review, see Dagum and Cholette (2006).

Usage

```
TempDisaggDGP(
  n_l,
  n,
  aggRatio = 4,
  p = 1,
  beta = 1,
  sparsity = 1,
  method = "Chow-Lin",
  aggMat = "sum",
  rho = 0,
  mean_X = 0,
  sd_X = 1,
  sd_e = 1,
  simul = FALSE,
  setSeed = 42
)
```

Arguments

`n_l`	Size of the low frequency series.
`n`	Size of the high frequency series.
`aggRatio`	aggregation ratio (default is 4)
`p`	The number of high-frequency indicator series to include.
`beta`	The positive and negative beta elements for the coefficient vector.
`sparsity`	Sparsity percentage of the coefficient vector.
`method`	DGP of residuals, either 'Denton', 'Denton-Cholette', 'Chow-Lin', 'Fernandez', 'Litterman'.
`aggMat`	Aggregation matrix according to 'first', 'sum', 'average', 'last'.

rho	The residual autocorrelation coefficient. Default is 0.
mean_X	Mean of the design matrix. Default is 0.
sd_X	Standard deviation of the design matrix. Default is 1.
sd_e	Standard deviation of the errors. Default is 1.
simul	When 'TRUE' the design matrix and the coefficient vector are fixed.
setSeed	The seed used when 'simul' is set to 'TRUE'.

Value

y_Gen Generated high-frequency response series.

Y_Gen Generated low-frequency response series.

X_Gen Generated high-frequency indicator series.

Beta_Gen Generated coefficient vector.

e_Gen Generated high-frequency residual series.

References

Dagum EB, Cholette PA (2006). *Benchmarking, temporal distribution, and reconciliation methods for time series*, volume 186. Springer Science \& Business Media.

Examples

```
data = TempDisaggDGP(n_l=25, n=100, aggRatio=4,p=10, rho=0.5)
X = data$X_Gen
Y = data$Y_Gen
```

Index