Package ‘CompositionalSR’

February 25, 2026

Type Package

Title Spatial Regression Models with Compositional Data

Version 1.1

Date 2026-02-25

Author Michail Tsagris [aut, cre]

Maintainer Michail Tsagris <mtsagris@uoc.gr>

Depends R (>= 4.0)

Imports blockCV, Compositional, doParallel, foreach, gslnls, minpack.lm, parallel, Rfast, sf, stats, utils

Suggests Rfast2

Description Spatial regression models with compositional responses using the alpha--transformation. Relevant papers include: Tsagris M. (2025), <doi:10.48550/arXiv.2510.12663>, Tsagris M. (2015), <https://soche.cl/chjs/volumes/06/02/Tsagris(2015).pdf>, Tsagris M.T., Preston S. and Wood A.T.A. (2011), <doi:10.48550/arXiv.1106.1451>.

LazyData true

NeedsCompilation no

Repository CRAN

Date/Publication 2026-02-25 10:50:13 UTC

Contents

`CompositionalSR-package` *Spatial Regression Models with Compositional Data*

Description

Spatial regression models with compositional responses using the α–transformation. The models includes are the α-regression (not spatial), the α-spatially lagged X (α-SLX) model and the geographically weighted α-regression (GWαR) model.

Details

Package:	CompositionalSR
Type:	Package
Version:	1.1
Date:	2026-02-25
License:	GPL-2

Maintainers

Michail Tsagris <mtsagris@uoc.gr>

Author(s)

Michail Tsagris <`mtsagris@uoc.gr`>.

References

Tsagris M. (2025). The α–regression for compositional data: a unified framework for standard, spatially-lagged, spatial autoregressive and geographically-weighted regression models. https://arxiv.org/pdf/2510.12663

Tsagris M. (2015). Regression analysis with compositional data containing zero values. Chilean Journal of Statistics, 6(2): 47-57. https://arxiv.org/pdf/1508.01913v1.pdf

Tsagris M.T., Preston S. and Wood A.T.A. (2011). A data-based power transformation for compositional data. In Proceedings of the 4th Compositional Data Analysis Workshop, Girona, Spain. https://arxiv.org/pdf/1106.1451.pdf

---

`Computation of the contiguity matrix W`

Computation of the contiguity matrix W

Description

Computation of the contiguity matrix W.

Usage

```
contiguity(coords, k = 10)
```

Arguments

`coords`	A matrix with the coordinates of the locations. The first column is the latitude and the second is the longitude.
`k`	The number of nearest neighbours to consider for the contiguity matrix.

Value

The contiguity matrix W. A square matrix with row standardised values (the elements of each row sum to 1).

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <`mtsagris@uoc.gr`>.

See Also

`alfa.slx`, `cv.alfaslx`, `me.aslx`

Examples

```
data(fadn)
W <- contiguity(fadn[, 1:2])
```

fadn *FADN dataset*

Description

A matrix with 11 columns. The first two are the locations (latitude and longitude), the next five contain the compositional data (percentages of cultivated area of five crops), Y1.1: cereals, Y2.1: cotton, Y3.1: tree crops, Y4.1: other annual crops and pasture and Y5.1: grapes and wine. The next four columns contain the covariates, G1: Human Influence Index, G2: soil pH, G3: topsoil organic carbon content and G7: erosion.

Usage

```
fadn
```

Format

A matrix with 168 rows and 11 columns.

Source

Clark and Dixon (2021), available at https://github.com/nick3703/Chicago-Data.

References

Clark, N. J. and P. M. Dixon (2021). A class of spatially correlated self-exciting statistical models. *Spatial Statistics*, 43, 1–18.

See Also

alfa.slx, gwar, alfa.reg

Examples

```
data(fadn)
y <- fadn[, 3:7]
x <- fadn[, 8:11]
mod <- alfa.reg(y, x, a = 0.1)
```

`K-fold cross-validtion for the alpha-regression`
K-fold cross-validtion the α-regression

Description

K-fold cross-validtion the α-regression.

Usage

```
cv.alfareg(y, x, a = seq(0.1, 1, by = 0.1), nfolds = 10,
folds = NULL, nc = 1, seed = NULL)
```

Arguments

`y`	A matrix with compositional data. zero values are allowed.
`x`	A matrix with the continuous predictor variables or a data frame including categorical predictor variables.
`a`	The value of the power transformation, it has to be between -1 and 1. If zero values are present it has to be greater than 0. If $\alpha = 0$ the isometric log-ratio transformation is applied.
`nfolds`	The number of folds to split the data.
`folds`	If you have the list with the folds supply it here. You can also leave it NULL and it will create folds.
`nc`	The number of cores to use. IF you have a multicore computer it is advisable to use more than 1. It makes the procedure faster. It is advisable to use it if you have many observations and or many variables, otherwise it will slow down th process.
`seed`	You can specify your own seed number here or leave it NULL.

Details

Tuning the value of α in the α-regression takes place using K-fold cross-validation.

Value

A list including:

`runtime`	The runtime required by the cross-validation.
`perf`	A vector with the average Kullback-Leibler divergence, for every value of α.
`opt`	A vector with the minimum Kullback-Leibler divergence and the optimal value of α.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Tsagris M. (2025). The α–regression for compositional data: a unified framework for standard, spatially-lagged, spatial autoregressive and geographically-weighted regression models. https://arxiv.org/pdf/2510.12663

Tsagris M. (2015). Regression analysis with compositional data containing zero values. Chilean Journal of Statistics, 6(2): 47-57. https://arxiv.org/pdf/1508.01913v1.pdf

Tsagris M.T., Preston S. and Wood A.T.A. (2011). A data-based power transformation for compositional data. In Proceedings of the 4th Compositional Data Analysis Workshop, Girona, Spain. https://arxiv.org/pdf/1106.1451.pdf

See Also

`alfa.reg`, `cv.alfaslx`, `cv.gwar`, `me.ar`

Examples

```
data(fadn)
y <- fadn[, 3:7]
x <- fadn[, 8]
mod <- cv.alfareg(y, x, a = c(0.5, 1))
```

`Leave-one-out cross-validation for the GWalphaR model`

Leave-one-out cross-validation for the GWαR model

Description

Leave-one-out cross-validation for the GWαR model

Usage

```
cv.gwar(y, x, a = c(0.1, 0.25, 0.5, 0.75, 1), coords, h,
nfolds = 10, size = 1000, folds = NULL)
```

Arguments

`y`	A matrix with compositional data. zero values are allowed.
`x`	A matrix with the continuous predictor variables or a data frame including categorical predictor variables.
`a`	The value of the power transformation, it has to be between -1 and 1. If zero values are present it has to be greater than 0. If $\alpha = 0$ the isometric log-ratio transformation is applied.

`coords`	A matrix with the coordinates of the locations. The first column is the latitude and the second is the longitude.
`h`	A vector with bandwith values.
`nfolds`	The number of folds to split the data.
`size`	A numeric value of the specified range by which blocks are created and training/testing data are separated. This distance should be in metres. If you have big regions you should consider increasing this number. For more information see the package *blockCV*.
`folds`	If you have the list with the folds supply it here. You can also leave it NULL and it will create folds.

Details

The 10-fold spatial cross-validation protocol is applied to choose the optimal values of α and h.

Value

A list including:

`runtime`	The runtime required by the cross-validation.
`perf`	A vector with the average Kullback-Leibler divergence, for every value of α.
`opt`	A vector with the minimum Kullback-Leibler divergance, the optimal value of α and h.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <`mtsagris@uoc.gr`>.

References

Tsagris M. (2025). The α–regression for compositional data: a unified framework for standard, spatially-lagged, spatial autoregressive and geographically-weighted regression models. https://arxiv.org/pdf/2510.12663

Tsagris M. (2015). Regression analysis with compositional data containing zero values. Chilean Journal of Statistics, 6(2): 47-57. https://arxiv.org/pdf/1508.01913v1.pdf

Tsagris M.T., Preston S. and Wood A.T.A. (2011). A data-based power transformation for compositional data. In Proceedings of the 4th Compositional Data Analysis Workshop, Girona, Spain. https://arxiv.org/pdf/1106.1451.pdf

See Also

`gwar`, `me.gwar` `cv.alfaslx`

Examples

```
data(fadn)
coords <- fadn[, 1:2]
y <- fadn[, 3:7]
x <- fadn[, 8]
mod <- gwar(y, x, a = 1, coords, h = 0.001)
```

Marginal effects for the alpha-regression model

Marginal effects for the α-regression model

Description

Marginal effects for the α-regression model.

Usage

```
me.ar(be, mu, x, cov_be = NULL)
```

Arguments

be	A matrix with the beta regression coefficients of the α-regression model.
mu	The fitted values of the α-regression.
x	A matrix with the continuous predictor variables or a data frame. Categorical predictor variables are not suited here.
cov_be	The covariance matrix of the beta regression coefficients. If you pass this argument, then the standard error of the average marginal effects will be returned.

Details

The marginal effects of the α-regression model are computed.

Value

A list including:

me	An array with the marginal effects of each component for each predictor variable.
ame	The average marginal effects of each component for each predictor variable.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Tsagris M. (2025). The α–regression for compositional data: a unified framework for standard, spatially-lagged, spatial autoregressive and geographically-weighted regression models. https://arxiv.org/pdf/2510.12663

Tsagris M. (2015). Regression analysis with compositional data containing zero values. Chilean Journal of Statistics, 6(2): 47-57. https://arxiv.org/pdf/1508.01913v1.pdf

Tsagris M.T., Preston S. and Wood A.T.A. (2011). A data-based power transformation for compositional data. In Proceedings of the 4th Compositional Data Analysis Workshop, Girona, Spain. https://arxiv.org/pdf/1106.1451.pdf

See Also

`me.aslx`, `me.gwar`, `alfa.reg`

Examples

```
data(fadn)
y <- fadn[, 3:7]
x <- fadn[, 8]
mod <- alfa.reg(y, x, 0.2, xnew = x)
me <- me.ar(mod$be, mod$est, x)
```

`Marginal effects for the alpha-SAR model`

Marginal effects for the α-SAR model

Description

Marginal effects for the α-SAR model.

Usage

```
me.asar(be, rho, mu, x, coords, k, cov_theta = NULL)
```

Arguments

`be`	A matrix with the beta coefficients of the α-SAR model.
`rho`	The spatial auto-regressive coefficient ρ of the α-SAR model.
`mu`	The fitted values of the α-SAR model.
`x`	A matrix with the continuous predictor variables or a data frame. Categorical predictor variables are not suited here.
`coords`	A matrix with the coordinates of the locations. The first column is the latitude and the second is the longitude.
`k`	The number of nearest neighbours to consider for the contiguity matrix.
`cov_theta`	The covariance matrix of the beta and gamma regression coefficients. If you pass this argument, then the standard error of the average marginal effects will be returned.

Details

The marginal effects of the α-SAR model are computed.

Value

A list including:

- `me.dir` An array with the direct marginal effects of each component for each predictor variable.
- `me.indir` An array with the indirect marginal effects of each component for each predictor variable.
- `me.total` An array with the total marginal effects of each component for each predictor variable.
- `ame.dir` An array with the average direct marginal effects of each component for each predictor variable.
- `ame.indir` An array with the average indirect marginal effects of each component for each predictor variable.
- `ame.total` An array with the aerage total marginal effects of each component for each predictor variable.
- `se.amedir` An array with the standard errors of the average direct marginal effects of each component for each predictor variable. This is returned if you supply the covariance matrix cov_theta.
- `se.ameindir` An array with the standard errors of the average indirect marginal effects of each component for each predictor variable. This is returned if you supply the covariance matrix cov_theta.
- `se.ametotal` An array with the standard errors of the average total marginal effects of each component for each predictor variable. This is returned if you supply the covariance matrix cov_theta.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Tsagris M. (2025). The α–regression for compositional data: a unified framework for standard, spatially-lagged, spatial autoregressive and geographically-weighted regression models. https://arxiv.org/pdf/2510.12663

Tsagris M. (2015). Regression analysis with compositional data containing zero values. Chilean Journal of Statistics, 6(2): 47-57. https://arxiv.org/pdf/1508.01913v1.pdf

Tsagris M.T., Preston S. and Wood A.T.A. (2011). A data-based power transformation for compositional data. In Proceedings of the 4th Compositional Data Analysis Workshop, Girona, Spain. https://arxiv.org/pdf/1106.1451.pdf

See Also

`me.ar`, `me.aslx`, `me.gwar`

Examples

```
data(fadn)
coords <- fadn[, 1:2]
y <- fadn[, 3:7]
x <- fadn[, 8]
mod <- alfa.sar(y, x, a = 0.5, coords, k = 8)
me <- me.asar(mod$be, mod$rho, mod$est, x, coords, k = 6)
```

Marginal effects for the alpha-SLX model
Marginal effects for the α-SLX model

Description

Marginal effects for the α-SLX model.

Usage

```
me.aslx(be, gama, mu, x, coords, k = 10, cov_theta = NULL)
```

Arguments

`be`	A matrix with the beta coefficients of the α-SLX model.
`gama`	A matrix with the gamma coefficients of the α-SLX model.
`mu`	The fitted values of the α-SLX model.
`x`	A matrix with the continuous predictor variables or a data frame. Categorical predictor variables are not suited here.
`coords`	A matrix with the coordinates of the locations. The first column is the latitude and the second is the longitude.
`k`	The number of nearest neighbours to consider for the contiguity matrix.
`cov_theta`	The covariance matrix of the beta and gamma regression coefficients. If you pass this argument, then the standard error of the average marginal effects will be returned.

Details

The marginal effects of the α-SLX model are computed.

Value

A list including:

`me.dir`	An array with the direct marginal effects of each component for each predictor variable.
`me.indir`	An array with the indirect marginal effects of each component for each predictor variable.

me.total	An array with the total marginal effects of each component for each predictor variable.
ame.dir	An array with the average direct marginal effects of each component for each predictor variable.
ame.indir	An array with the average indirect marginal effects of each component for each predictor variable.
ame.total	An array with the aerage total marginal effects of each component for each predictor variable.
se.amedir	An array with the standard errors of the average direct marginal effects of each component for each predictor variable. This is returned if you supply the covariance matrix cov_theta.
se.ameindir	An array with the standard errors of the average indirect marginal effects of each component for each predictor variable. This is returned if you supply the covariance matrix cov_theta.
se.ametotal	An array with the standard errors of the average total marginal effects of each component for each predictor variable. This is returned if you supply the covariance matrix cov_theta.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Tsagris M. (2025). The α–regression for compositional data: a unified framework for standard, spatially-lagged, spatial autoregressive and geographically-weighted regression models. https://arxiv.org/pdf/2510.12663

Tsagris M. (2015). Regression analysis with compositional data containing zero values. Chilean Journal of Statistics, 6(2): 47-57. https://arxiv.org/pdf/1508.01913v1.pdf

Tsagris M.T., Preston S. and Wood A.T.A. (2011). A data-based power transformation for compositional data. In Proceedings of the 4th Compositional Data Analysis Workshop, Girona, Spain. https://arxiv.org/pdf/1106.1451.pdf

See Also

me.gwar, me.ar, alfa.slx

Examples

```
data(fadn)
coords <- fadn[, 1:2]
y <- fadn[, 3:7]
x <- fadn[, 8]
mod <- alfa.slx(y, x, a = 0.5, coords, k = 10, xnew = x, coordsnew = coords)
me <- me.aslx(mod$be, mod$gama, mod$est, x, coords, k = 10)
```

`Marginal effects for the GWalphaR model`
Marginal effects for the GWαR model

Description

Marginal effects for the GWαR model.

Usage

```
me.gwar(be, mu, x)
```

Arguments

`be`	A matrix with the beta regression coefficients of the α-regression model.
`mu`	The fitted values of the α-regression.
`x`	A matrix with the continuous predictor variables or a data frame. Categorical predictor variables are not suited here.

Details

The location-specific marginal effects for the GWαR model are computed.

Value

A list including:

`me`	An array with the location-specific marginal effects of each component for each predictor variable.
`ame`	The average location-specific marginal effects of each component for each predictor variable.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <`mtsagris@uoc.gr`>.

References

Tsagris M. (2025). The α–regression for compositional data: a unified framework for standard, spatially-lagged, spatial autoregressive and geographically-weighted regression models. https://arxiv.org/pdf/2510.12663

Tsagris M. (2015). Regression analysis with compositional data containing zero values. Chilean Journal of Statistics, 6(2): 47-57. https://arxiv.org/pdf/1508.01913v1.pdf

Tsagris M.T., Preston S. and Wood A.T.A. (2011). A data-based power transformation for compositional data. In Proceedings of the 4th Compositional Data Analysis Workshop, Girona, Spain. https://arxiv.org/pdf/1106.1451.pdf

See Also

gwar, me.aslx, me.ar

Examples

```
data(fadn)
coords <- fadn[, 1:2]
y <- fadn[, 3:7]
x <- fadn[, 8]
mod <- gwar(y, x, a = 1, coords, h = 0.001)
me <- me.gwar(mod$be, mod$est, x)
```

Prediction with the GWalphaR model

Prediction with the GWαR model

Description

Prediction with GWαR model.

Usage

```
gwar.pred(y, x, a, coords, h, xnew, coordsnew)
```

Arguments

y	A matrix with the compositional data.
x	A matrix with the continuous predictor variables or a data frame including categorical predictor variables.
a	A vector with values for the power transformation, it has to be between -1 and 1.
coords	A matrix with the coordinates of the locations. The first column is the latitude and the second is the longitude.
h	A vector with bandwith values.
xnew	The new data.
coordsnew	A matrix with the coordinates of the new locations. The first column is the latitude and the second is the longitude.

Details

The α-transformation is applied to the compositional data first and then the GWαR model is applied and predictions are given for each observation.

Value

A list including:

`runtime` The time required by the regression.

`est` A list with the fitted values, for each combination of α and h.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Tsagris M. (2025). The α–regression for compositional data: a unified framework for standard, spatially-lagged, spatial autoregressive and geographically-weighted regression models. https://arxiv.org/pdf/2510.12663

Tsagris M. (2015). Regression analysis with compositional data containing zero values. Chilean Journal of Statistics, 6(2): 47-57. https://arxiv.org/pdf/1508.01913v1.pdf

Tsagris M.T., Preston S. and Wood A.T.A. (2011). A data-based power transformation for compositional data. In Proceedings of the 4th Compositional Data Analysis Workshop, Girona, Spain. https://arxiv.org/pdf/1106.1451.pdf

See Also

`cv.gwar`, `me.gwar`, `alfa.slx`, `alfa.reg`

Examples

```
data(fadn)
coords <- fadn[-c(1:10), 1:2]
y <- fadn[-c(1:10), 3:7]
x <- fadn[-c(1:10), 8]
xnew <- fadn[1:10, 8]
coordsnew <- fadn[1:10, 1:2]
mod <- gwar.pred(y, x, a = c(0.25, 0.5, 1), coords,
h = c(0.002, 0.006), xnew = xnew, coordsnew = coordsnew)
```

`Regression with compositional data using the alpha-transformation`
Regression with compositional data using the α-transformation

---

Description

Regression with compositional data using the α-transformation.

Usage

```
alfa.reg(y, x, a, covb = FALSE, xnew = NULL, yb = NULL)
alfa.reg2(y, x, a, xnew = NULL, ncores = 1)
alfa.reg3(y, x, a = c(-1, 1), xnew = NULL)
```

Arguments

`y`	A matrix with the compositional data.
`x`	A matrix with the continuous predictor variables or a data frame including categorical predictor variables.
`a`	The value of the power transformation, it has to be between -1 and 1. If zero values are present it has to be greater than 0. If $\alpha = 0$ the isometric log-ratio transformation is applied and the solution exists in a closed form, since it the classical mutivariate regression. For the alfa.reg2() this should be a vector of α values and the function call repeatedly the alfa.reg() function. For the alfa.reg3() function it should be a vector with two values, the endpoints of the interval of α. This function searches for the optimal vaue of α that minimizes the Kullback-Leibler between the observed and fitted compositions. Using the `optimize` function it searches for the optimal value of α. Instead of choosing the value of α using `cv.alfareg` (that uses cross-validation) one can select it this way.
`covb`	Do you want the covariance matrix of the regression coefficients to be returned? If TRUE, this will slow down the process, as it is computed numerically.
`xnew`	If you have new data use it, otherwise leave it NULL.
`ncores`	The number of cores to use for parallel computations.
`yb`	If you have already transformed the data using the α-transformation with the same α as given in the argument "a", put it here. Othewrise leave it NULL. This is intended to be used in the function `cv.alfareg` in order to speed up the process. The time difference in that function is small for small samples. But, if you have a few thousands and or a few more components, there will be bigger differences.

Details

The α-transformation is applied to the compositional data first and then multivariate regression is applied. This involves numerical optimisation. The alfa.reg2() function accepts a vector with many values of α, while the the alfa.reg3() function searches for the value of α that minimizes the Kulback-Leibler divergence between the observed and the fitted compositional values. The functions are highly optimized.

Value

For the alfa.reg() function a list including:

`runtime`	The time required by the regression.
`be`	The beta coefficients.
`covbe`	The covariance matrix if covb was set to TRUE, otherwise NULL.

`dev` The sum of the squared residuals, as produced by the function **minpack.lm::nls.lm()**.

`est` The fitted values for xnew if xnew is not NULL.

For the alfa.reg2() function a list with the time required by all regressions and the regression coefficients and the fitted values for each value of α.

For the alfa.reg3() function a list with the previous elements plus an output "alfa", the optimal value of α.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <`mtsagris@uoc.gr`>.

References

Tsagris M. (2025). The α–regression for compositional data: a unified framework for standard, spatially-lagged, spatial autoregressive and geographically-weighted regression models. https://arxiv.org/pdf/2510.12663

Tsagris M. (2015). Regression analysis with compositional data containing zero values. Chilean Journal of Statistics, 6(2): 47-57. https://arxiv.org/pdf/1508.01913v1.pdf

Tsagris M.T., Preston S. and Wood A.T.A. (2011). A data-based power transformation for compositional data. In Proceedings of the 4th Compositional Data Analysis Workshop, Girona, Spain. https://arxiv.org/pdf/1106.1451.pdf

Mardia K.V., Kent J.T., and Bibby J.M. (1979). Multivariate analysis. Academic press.

Aitchison J. (1986). The statistical analysis of compositional data. Chapman & Hall.

See Also

`cv.alfareg`, `alfareg.nr`, `alfa.slx`

Examples

```
data(fadn)
y <- fadn[, 3:7]
x <- fadn[, 8]
mod <- alfa.reg(y, x, 0.2)
```

`Robust regression with compositional data using the alpha-transformation`
Regression with compositional data using the α-transformation

---

Description

Regression with compositional data using the α-transformation.

Usage

```
rob.alfareg(y, x, a, loss = "welsh", xnew = NULL, yb = NULL)
```

Arguments

y	A matrix with the compositional data.
x	A matrix with the continuous predictor variables or a data frame including categorical predictor variables.
a	The value of the power transformation, it has to be between -1 and 1. If zero values are present it has to be greater than 0. If $\alpha = 0$ the isometric log-ratio transformation is applied and the solution exists in a closed form, since it the classical mutivariate regression.
loss	The loss function to use. One of these available options, "barron", "bisquare", "welsh", "optimal", "hampel", "ggw", or "lqq". For more information see the package *gslnls*.
xnew	If you have new data use it, otherwise leave it NULL.
yb	If you have already transformed the data using the α-transformation with the same α as given in the argument "a", put it here. Othewrise leave it NULL.

Details

The α-transformation is applied to the compositional data first and then robust multivariate regression is applied. This involves numerical optimisation.

Value

A list including:

runtime	The time required by the regression.
be	The beta coefficients.
est	The fitted values for xnew if xnew is not NULL.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Tsagris M. (2025). The α–regression for compositional data: a unified framework for standard, spatially-lagged, spatial autoregressive and geographically-weighted regression models. https://arxiv.org/pdf/2510.12663

Tsagris M. (2015). Regression analysis with compositional data containing zero values. Chilean Journal of Statistics, 6(2): 47-57. https://arxiv.org/pdf/1508.01913v1.pdf

Tsagris M.T., Preston S. and Wood A.T.A. (2011). A data-based power transformation for compositional data. In Proceedings of the 4th Compositional Data Analysis Workshop, Girona, Spain. https://arxiv.org/pdf/1106.1451.pdf

Mardia K.V., Kent J.T., and Bibby J.M. (1979). Multivariate analysis. Academic press.

Aitchison J. (1986). The statistical analysis of compositional data. Chapman & Hall.

See Also

alfa.reg, alfareg.nr, alfa.slx

Examples

```
data(fadn)
y <- fadn[, 3:7]
x <- fadn[, 8]
mod <- rob.alfareg(y, x, 0.2)
```

Spatial K-fold cross-validation for the alpha-SAR model

Spatial K-fold cross-validation for the α-SAR model

Description

Spatial K-fold cross-validation for the α-SAR model

Usage

```
cv.alfasar(y, x, a = seq(0.1, 1, by = 0.1), coords, k = 2:15,
nfolds = 10, size = 1000, folds = NULL)
```

Arguments

y	A matrix with compositional data. zero values are allowed.
x	A matrix with the continuous predictor variables or a data frame including categorical predictor variables.
a	The value of the power transformation, it has to be between -1 and 1. If zero values are present it has to be greater than 0. If $\alpha = 0$ the isometric log-ratio transformation is applied.
coords	A matrix with the coordinates of the locations. The first column is the latitude and the second is the longitude.
k	A vector with the nearest neighbours to consider for the contiguity matrix.
nfolds	The number of folds to split the data.
size	A numeric value of the specified range by which blocks are created and training/testing data are separated. This distance should be in metres. If you have big regions you should consider increasing this number. For more information see the package *blockCV*.
folds	If you have the list with the folds supply it here. You can also leave it NULL and it will create folds.

Details

The 10-fold spatial cross-validation protocol is applied to choose the optimal values of α and k.

Value

A list including:

`runtime`	The runtime required by the cross-validation.
`perf`	A vector with the average Kullback-Leibler divergence, for every value of α.
`opt`	A vector with the minimum Kullback-Leibler divergence, the optimal value of α and k.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Tsagris M. (2025). The α–regression for compositional data: a unified framework for standard, spatially-lagged, spatial autoregressive and geographically-weighted regression models. https://arxiv.org/pdf/2510.12663

Tsagris M. (2015). Regression analysis with compositional data containing zero values. Chilean Journal of Statistics, 6(2): 47-57. https://arxiv.org/pdf/1508.01913v1.pdf

Tsagris M.T., Preston S. and Wood A.T.A. (2011). A data-based power transformation for compositional data. In Proceedings of the 4th Compositional Data Analysis Workshop, Girona, Spain. https://arxiv.org/pdf/1106.1451.pdf

See Also

`alfa.slx`, `cv.gwar` `cv.alfareg`

Examples

```
data(fadn)
coords <- fadn[1:50, 1:2]
y <- fadn[1:50, 3:7]
x <- fadn[1:50, 8]
```

`Spatial K-fold cross-validation for the alpha-SLX model`
Spatial K-fold cross-validation for the α-SLX model

Description

Spatial K-fold cross-validation for the α-SLX model

Usage

```
cv.alfaslx(y, x, a = seq(0.1, 1, by = 0.1), coords, k = 2:15,
nfolds = 10, size = 1000, folds = NULL)
```

Arguments

`y`	A matrix with compositional data. zero values are allowed.
`x`	A matrix with the continuous predictor variables or a data frame including categorical predictor variables.
`a`	The value of the power transformation, it has to be between -1 and 1. If zero values are present it has to be greater than 0. If $\alpha = 0$ the isometric log-ratio transformation is applied.
`coords`	A matrix with the coordinates of the locations. The first column is the latitude and the second is the longitude.
`k`	A vector with the nearest neighbours to consider for the contiguity matrix.
`nfolds`	The number of folds to split the data.
`size`	A numeric value of the specified range by which blocks are created and training/testing data are separated. This distance should be in metres. If you have big regions you should consider increasing this number. For more information see the package *blockCV*.
`folds`	If you have the list with the folds supply it here. You can also leave it NULL and it will create folds.

Details

The 10-fold spatial cross-validation protocol is applied to choose the optimal values of α and k.

Value

A list including:

`runtime`	The runtime required by the cross-validation.
`perf`	A vector with the average Kullback-Leibler divergence, for every value of α.
`opt`	A vector with the minimum Kullback-Leibler divergence, the optimal value of α and k.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Tsagris M. (2025). The α–regression for compositional data: a unified framework for standard, spatially-lagged, spatial autoregressive and geographically-weighted regression models. https://arxiv.org/pdf/2510.12663

Tsagris M. (2015). Regression analysis with compositional data containing zero values. Chilean Journal of Statistics, 6(2): 47-57. https://arxiv.org/pdf/1508.01913v1.pdf

Tsagris M.T., Preston S. and Wood A.T.A. (2011). A data-based power transformation for compositional data. In Proceedings of the 4th Compositional Data Analysis Workshop, Girona, Spain. https://arxiv.org/pdf/1106.1451.pdf

See Also

alfa.slx, cv.gwar cv.alfareg

Examples

```
data(fadn)
coords <- fadn[1:60, 1:2]
y <- fadn[1:60, 3:7]
x <- fadn[1:60, 8]
mod <- cv.alfaslx(y, x, a = 0.5, coords, k = 2)
```

Spatial k-folds *Spatial k-folds*

Description

Spatial k-folds.

Usage

```
spat.folds(coords, nfolds = 10, size = 1000)
```

Arguments

coords	A matrix with the coordinates of the locations. The first column is the latitude and the second is the longitude.
nfolds	The number of spatial folds to create.
size	A numeric value of the specified range by which blocks are created and training/testing data are separated. This distance should be in metres. If you have big regions you should consider increasing this number. For more information see the package *blockCV*.

Details

Folds of data are created based on their coordinates. For more information see the package **blockCV**.

Value

A list with nfolds elements. Each elements contains a list with two elements, the first is the indices of the training set and the second contains the indices of the test set.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

See Also

`cv.alfaslx`, `me.aslx`, `gwar`, `alfa.reg`

Examples

```
data(fadn)
coords <- fadn[1:100, 1:2]
folds <- spat.folds(coords, nfolds = 5)
```

`The alpha-regression using Newton-Raphson`

The alpha-regression using Newton-Raphson

Description

The $alpha$-regression using Newton-Raphson.

Usage

```
alfareg.nr(y, x, alpha = 1, beta_init = NULL, max_iter = 100,
tol = 1e-6, line_search = TRUE)
```

Arguments

`y`	A matrix with the compositional data.
`x`	A matrix with the continuous predictor variables or a data frame including categorical predictor variables.
`alpha`	The value of the power transformation, it has to be between -1 and 1. If zero values are present it has to be greater than 0.
`beta_init`	A vector of initial parameters (optional). This is then transformed into a matrix.
`max_iter`	The maximum number of iterations for the Newton-Raphson algorithm.
`tol`	The tolerance value to terminate the Newton-Raphson algorithm.
`line_search`	Do you want to perform line search? The default value is TRUE.

Details

The α-transformation is applied to the compositional data first and then multivariate regression is applied. This involves numerical optimisation.

Value

A list including:

`runtime`	The time required by the regression.
`iters`	The iterations of the Newton-Raphson algorithm
`be`	The beta coefficients.
`objective`	The sum of the squared residuals.
`est`	The fitted values.
`covb`	The covariance matrix of the beta coefficients, or NULL if it is singular.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Tsagris M. (2025). The α–regression for compositional data: a unified framework for standard, spatially-lagged, spatial autoregressive and geographically-weighted regression models. https://arxiv.org/pdf/2510.12663

Tsagris M. (2015). Regression analysis with compositional data containing zero values. Chilean Journal of Statistics, 6(2): 47-57. https://arxiv.org/pdf/1508.01913v1.pdf

Tsagris M.T., Preston S. and Wood A.T.A. (2011). A data-based power transformation for compositional data. In Proceedings of the 4th Compositional Data Analysis Workshop, Girona, Spain. https://arxiv.org/pdf/1106.1451.pdf

Mardia K.V., Kent J.T., and Bibby J.M. (1979). Multivariate analysis. Academic press.

Aitchison J. (1986). The statistical analysis of compositional data. Chapman & Hall.

See Also

`alfa.reg`, `cv.alfareg`, `alfa.slx`

Examples

```
data(fadn)
y <- fadn[, 3:7]
x <- fadn[, 8]
mod <- alfareg.nr(y, x, a = 0.2)
mod2 <- alfa.reg(y, x, 0.2)
```

The alpha-SAR model *The α-SAR model*

Description

The α-SAR model.

Usage

```
alfa.sar(y, x, a, coords, k = 10, covb = FALSE, xnew = NULL, coordsnew, yb = NULL)
```

Arguments

`y`	A matrix with the compositional data.
`x`	A matrix with the continuous predictor variables or a data frame including categorical predictor variables.
`a`	The value of the power transformation, it has to be between -1 and 1. If zero values are present it has to be greater than 0. If $\alpha = 0$ the isometric log-ratio transformation is applied and the solution exists in a closed form, since it the classical mutivariate regression.
`coords`	A matrix with the coordinates of the locations. The first column is the latitude and the second is the longitude.
`k`	The number of nearest neighbours to consider for the contiguity matrix.
`covb`	Do you want the covariance matrix of the spatial autoregressive parameter and the regression coefficients to be returned? If TRUE, this will slow down the process, as it is computed numerically.
`xnew`	If you have new data use it, otherwise leave it NULL.
`coordsnew`	A matrix with the coordinates of the new locations. The first column is the latitude and the second is the longitude. If you do not have new data to make predictions leave this NULL.
`yb`	If you have already transformed the data using the α-transformation with the same α as given in the argument "a", put it here. Othewrise leave it NULL.

Details

The α-transformation is applied to the compositional data first and the spatial autocorrelation (SAR) model is applied. The function performs a grid search searching for the range of good values of ρ and then uses that as starting value.

Value

A list including:

`runtime`	The time required by the regression.
`be`	The beta coefficients.

gama	The gamma coefficients.
covbe	The covariance matrix if covb was set to TRUE, otherwise NULL.
dev	The sum of the squared residuals, as produced by the function **minpack.lm::nls.lm()**.
est	The fitted values for xnew if xnew and coordsnew are not NULL.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Tsagris M. (2025). The α–regression for compositional data: a unified framework for standard, spatially-lagged, spatial autoregressive and geographically-weighted regression models. https://arxiv.org/pdf/2510.12663

Tsagris M. (2015). Regression analysis with compositional data containing zero values. Chilean Journal of Statistics, 6(2): 47-57. https://arxiv.org/pdf/1508.01913v1.pdf

Tsagris M.T., Preston S. and Wood A.T.A. (2011). A data-based power transformation for compositional data. In Proceedings of the 4th Compositional Data Analysis Workshop, Girona, Spain. https://arxiv.org/pdf/1106.1451.pdf

See Also

cv.alfaslx, me.aslx, gwar, alfa.reg

Examples

```
data(fadn)
coords <- fadn[1:50, 1:2]
y <- fadn[1:50, 3:7]
x <- fadn[1:50, 8]
```

The alpha-SLX model *The α-SLX model*

Description

The α-SLX model.

Usage

```
alfa.slx(y, x, a, coords, k = 10, covb = FALSE, xnew = NULL, coordsnew, yb = NULL)
alfa.slx2(y, x, a, coords, k = 2:15, xnew = NULL, coordsnew, yb = NULL)
```

Arguments

`y`	A matrix with the compositional data.
`x`	A matrix with the continuous predictor variables or a data frame including categorical predictor variables.
`a`	The value of the power transformation, it has to be between -1 and 1. If zero values are present it has to be greater than 0. If $\alpha = 0$ the isometric log-ratio transformation is applied and the solution exists in a closed form, since it the classical mutivariate regression.
`coords`	A matrix with the coordinates of the locations. The first column is the latitude and the second is the longitude.
`k`	The number of nearest neighbours to consider for the contiguity matrix. For the alfa.slx2() this should be a vector.
`covb`	Do you want the covariance matrix of the regression coefficients to be returned? If TRUE, this will slow down the process, as it is computed numerically.
`xnew`	If you have new data use it, otherwise leave it NULL.
`coordsnew`	A matrix with the coordinates of the new locations. The first column is the latitude and the second is the longitude. If you do not have new data to make predictions leave this NULL.
`yb`	If you have already transformed the data using the α-transformation with the same α as given in the argument "a", put it here. Othewrise leave it NULL. This is intended to be used in the function `cv.alfareg` in order to speed up the process. The time difference in that function is small for small samples. But, if you have a few thousands and or a few more components, there will be bigger differences.

Details

The α-transformation is applied to the compositional data first and then the spatially lagged X (SLX) model is applied.

Value

For the alfa.slx() a list including:

`runtime`	The time required by the regression.
`be`	The beta coefficients.
`gama`	The gamma coefficients.
`covbe`	The covariance matrix if covb was set to TRUE, otherwise NULL.
`dev`	The sum of the squared residuals, as produced by the function **minpack.lm::nls.lm()**.
`est`	The fitted values for xnew if xnew and coordsnew are not NULL.

For the alfa.slx2() a list including:

`runtime`	The time required by the regression.
`be`	A list with the beta coefficients for each value of k.

gama	A list with the gamma coefficients.
dev	A vector with the sum of the squared residuals, as produced by the function **minpack.lm::nls.lm()**. The positions of the vector are the ones defined by the argument k that is a vector.
est	A list with the fitted values for the xnew and coordsnew, for each value of k.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Tsagris M. (2025). The α–regression for compositional data: a unified framework for standard, spatially-lagged, spatial autoregressive and geographically-weighted regression models. https://arxiv.org/pdf/2510.12663

Tsagris M. (2015). Regression analysis with compositional data containing zero values. Chilean Journal of Statistics, 6(2): 47-57. https://arxiv.org/pdf/1508.01913v1.pdf

Tsagris M.T., Preston S. and Wood A.T.A. (2011). A data-based power transformation for compositional data. In Proceedings of the 4th Compositional Data Analysis Workshop, Girona, Spain. https://arxiv.org/pdf/1106.1451.pdf

See Also

cv.alfaslx, me.aslx, gwar, alfa.reg

Examples

```
data(fadn)
coords <- fadn[, 1:2]
y <- fadn[, 3:7]
x <- fadn[, 8]
mod <- alfa.slx(y, x, a = 0.5, coords, k = 10)
```

The gradient vector of the alpha-regression model at each observation

The gradient vector of the α-regression model at each observation

---

Description

The gradient vector of the α-regression model at each observation.

Usage

```
ar.grads(y, x, a, be)
```

Arguments

`y`	A matrix with the compositional data.
`x`	A matrix with the continuous predictor variables or a data frame including categorical predictor variables.
`a`	The value of the power transformation, it has to be between -1 and 1. If zero values are present it has to be greater than 0.
`be`	The regression coefficients of the α-SAR model.

Details

The gradient vector of the α-regression model is computed at each observation.

Value

A matrix with the gradient vector computed at each observation.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Tsagris M. (2025). The α–regression for compositional data: a unified framework for standard, spatially-lagged, spatial autoregressive and geographically-weighted regression models. https://arxiv.org/pdf/2510.12663

Tsagris M. (2015). Regression analysis with compositional data containing zero values. Chilean Journal of Statistics, 6(2): 47-57. https://arxiv.org/pdf/1508.01913v1.pdf

Tsagris M.T., Preston S. and Wood A.T.A. (2011). A data-based power transformation for compositional data. In Proceedings of the 4th Compositional Data Analysis Workshop, Girona, Spain. https://arxiv.org/pdf/1106.1451.pdf

See Also

`alfa.sar`, `cv.alfasar`, `alfa.reg`

Examples

```
data(fadn)
coords <- fadn[, 1:2]
y <- fadn[, 3:7]
x <- fadn[, 8:10]
mod <- alfa.reg(y, x, 0.5)
grads <- ar.grads(y, x, a = 0.5, mod$be)
colSums(grads)
```

The gradient vector of the alpha-SAR model at each observation
The gradient vector of the α-SAR model at each observation

Description

The gradient vector of the α-SAR model at each observation.

Usage

```
asar.grads(y, x, a, rho, be, coords, k)
```

Arguments

y	A matrix with the compositional data.
x	A matrix with the continuous predictor variables or a data frame including categorical predictor variables.
a	The value of the power transformation, it has to be between -1 and 1. If zero values are present it has to be greater than 0.
rho	The spatial autocorrelation parameter ρ.
be	The regression coefficients of the α-SAR model.
coords	A matrix with the coordinates of the locations. The first column is the latitude and the second is the longitude.
k	The number of nearest neighbours to consider for the contiguity matrix.

Details

The gradient vector of the α-SAR model is computed at each observation.

Value

A matrix with the gradient vector computed at each observation.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Tsagris M. (2025). The α–regression for compositional data: a unified framework for standard, spatially-lagged, spatial autoregressive and geographically-weighted regression models. https://arxiv.org/pdf/2510.12663

Tsagris M. (2015). Regression analysis with compositional data containing zero values. Chilean Journal of Statistics, 6(2): 47-57. https://arxiv.org/pdf/1508.01913v1.pdf

Tsagris M.T., Preston S. and Wood A.T.A. (2011). A data-based power transformation for compositional data. In Proceedings of the 4th Compositional Data Analysis Workshop, Girona, Spain. https://arxiv.org/pdf/1106.1451.pdf

See Also

alfa.sar, cv.alfasar, alfa.reg

Examples

```
data(fadn)
coords <- fadn[, 1:2]
y <- fadn[, 3:7]
x <- fadn[, 8:10]

be <- matrix( c( 12.72191991, 0.04300266, -1.78301001, -3.02074120, -23.54785921,
0.06771573, 2.71969599, 1.89312564, 5.38640736, 0.05179626, -1.21336879, 0.40175088,
-1.98258721, 0.06815682, -0.64458883, 0.95470802 ), ncol = 4 )
```

The gradient vector of the alpha-SLX model at each observation

The gradient vector of the α-SLX model at each observation

Description

The gradient vector of the α-SLX model at each observation.

Usage

```
aslx.grads(y, x, a, be, gama, coords, k = 10)
```

Arguments

y	A matrix with the compositional data.
x	A matrix with the continuous predictor variables or a data frame including categorical predictor variables.
a	The value of the power transformation, it has to be between -1 and 1. If zero values are present it has to be greater than 0.
be	The regression coefficients of the α-SLX model.
gama	The gamma coefficients of the α-SLX model.
coords	A matrix with the coordinates of the locations. The first column is the latitude and the second is the longitude.
k	The number of nearest neighbours to consider for the contiguity matrix.

Details

The gradient vector of the α-SLX model is computed at each observation.

Value

A matrix with the gradient vector computed at each observation.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <mtsagris@uoc.gr>.

References

Tsagris M. (2025). The α–regression for compositional data: a unified framework for standard, spatially-lagged, spatial autoregressive and geographically-weighted regression models. https://arxiv.org/pdf/2510.12663

Tsagris M. (2015). Regression analysis with compositional data containing zero values. Chilean Journal of Statistics, 6(2): 47-57. https://arxiv.org/pdf/1508.01913v1.pdf

Tsagris M.T., Preston S. and Wood A.T.A. (2011). A data-based power transformation for compositional data. In Proceedings of the 4th Compositional Data Analysis Workshop, Girona, Spain. https://arxiv.org/pdf/1106.1451.pdf

See Also

alfa.sar, cv.alfasar, alfa.reg

Examples

```
data(fadn)
coords <- fadn[, 1:2]
y <- fadn[, 3:7]
x <- fadn[, 8]
mod <- alfa.slx(y, x, a = 0.5, coords, k = 10)
grads <- aslx.grads(y, x, a = 0.5, mod$be, mod$gama, coords, k = 10)
colSums(grads)
```

The GWaR model — *The GWαR model*

Description

The GWαR model.

Usage

```
gwar(y, x, a, coords, h, yb = NULL, nc = 1)
```

Arguments

y	A matrix with the compositional data.
x	A matrix with the continuous predictor variables or a data frame including categorical predictor variables.
a	The value of the power transformation, it has to be between -1 and 1.

`coords`	A matrix with the coordinates of the locations. The first column is the latitude and the second is the longitude.
`h`	The bandwith value.
`yb`	If you have already transformed the data using the α-transformation with the same α as given in the argument "a", put it here. Othewrise leave it NULL.
`nc`	The number of cores to use. IF you have a multicore computer it is advisable to use more than 1. It makes the procedure faster. It is advisable to use it if you have many observations and or many variables, otherwise it will slow down th process.

Details

The α-transformation is applied to the compositional data first and then the GWαR model is applied.

Value

A list including:

`runtime`	The time required by the regression.
`be`	The beta coefficients.
`est`	The fitted values.

Author(s)

Michail Tsagris.

R implementation and documentation: Michail Tsagris <`mtsagris@uoc.gr`>.

References

Tsagris M. (2025). The α–regression for compositional data: a unified framework for standard, spatially-lagged, spatial autoregressive and geographically-weighted regression models. https://arxiv.org/pdf/2510.12663

Tsagris M. (2015). Regression analysis with compositional data containing zero values. Chilean Journal of Statistics, 6(2): 47-57. https://arxiv.org/pdf/1508.01913v1.pdf

Tsagris M.T., Preston S. and Wood A.T.A. (2011). A data-based power transformation for compositional data. In Proceedings of the 4th Compositional Data Analysis Workshop, Girona, Spain. https://arxiv.org/pdf/1106.1451.pdf

See Also

`cv.gwar`, `me.gwar`, `alfa.slx`, `alfa.reg`

Examples

```
data(fadn)
coords <- fadn[, 1:2]
y <- fadn[, 3:7]
x <- fadn[, 8]
mod <- gwar(y, x, a = 1, coords, h = 0.001)
```

Index