Package 'AGCA4extremes'

July 17, 2026

Title Anchored Geodesic Component Analysis for Extremes

Version 0.1.0

Description Implements anchored geodesic component analysis for multivariate extremes. The package provides Pareto and rank-Pareto marginal standardization, top-k angular extraction, anchored eigensolutions, reconstruction, diagnostics, nonparametric bootstrap tools, and plotting methods for benchmark-relative angular variation.

License GPL-3

SystemRequirements C++17

URL https://github.com/a91quaini/AGCA4extremes

BugReports https://github.com/a91quaini/AGCA4extremes/issues

Encoding UTF-8

Language en-US

Depends R (>= 4.1.0)

Imports graphics, Rcpp, stats

LinkingTo Rcpp, RcppArmadillo

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

Config/roxygen2/version 8.0.0

NeedsCompilation yes

Author Alberto Quaini [aut, cre]

Maintainer Alberto Quaini <alberto91quaini@gmail.com>

Repository CRAN

Date/Publication 2026-07-17 13:00:02 UTC

Contents

`agca` *Fit anchored geodesic component analysis*

Description

`agca()` is the main user interface. It standardizes margins, extracts large-radius observations, resolves the anchor, and fits anchored geodesic components to the resulting angular directions.

Usage

```
agca(
  x,
  k = NULL,
  threshold = NULL,
  margin = c("rank_pareto", "pareto", "none"),
  cdf = NULL,
```

```
  anchor = "canonical",
  p = NULL,
  decluster = NULL,
  bootstrap = NULL,
  keep_data = FALSE,
  ties_method = "average",
  seed = NULL
)
```

Arguments

x	Numeric matrix or data frame. Larger values are treated as more extreme in each margin.
k	Number of largest radial observations to retain. Specify either k or threshold.
threshold	Radial threshold for selecting extremes.
margin	Marginal standardization method. The default "rank_pareto" uses empirical ranks. Use "pareto" with cdf, or "none" for already standardized observations.
cdf	Optional CDF function or list of CDF functions for margin = "pareto".
anchor	"canonical", "frechet", "principal", or a numeric anchor vector.
p	Working reconstruction rank. Defaults to the full tangent rank.
decluster	Optional. Use NULL for no declustering, TRUE for runs declustering with run = 1, a nonnegative integer run length, or a list with component run.
bootstrap	Optional integer number of bootstrap resamples.
keep_data	Logical. If TRUE, store the standardized data matrix in the returned object.
ties_method	Tie method used by rank_pareto().
seed	Optional random seed used when bootstrap is supplied.

Value

An object of class "agca_fit".

Examples

```
data(agca_10d_simulation)
x <- agca_10d_simulation[paste0("X", 1:10)]
fit <- agca(x, k = 100, p = 3)
fit
agca_rank_summary(fit)
```

agca_10d_simulation *Simulated 10-dimensional AGCA example*

Description

A package-owned simulated heavy-tailed sample from the 10-dimensional design used in the AGCA paper. Variables `X1`–`X8` share a low-dimensional logistic-block extremal mechanism. Variables `X9` and `X10` contain independent Pareto sources, creating near-axis extreme regimes alongside the shared low-rank angular structure.

Usage

```
data(agca_10d_simulation)
```

Format

A data frame with 10,000 rows and 11 columns. Columns `X1`, ..., `X10` are positive heavy-tailed observations. Column `regime` is a latent factor identifying the dominant source for the observation: shared low-rank, axis 9, or axis 10.

Source

Simulated by `data-raw/simulate_data.R` using `simulate_agca_10d()`.

agca_fit_directions *Fit AGCA to angular directions*

Description

Fit AGCA to angular directions

Usage

```
agca_fit_directions(g, anchor = "canonical", p = NULL, normalize = TRUE)
```

Arguments

`g`	Matrix of angular directions.
`anchor`	`"canonical"`, `"frechet"`, `"principal"`, or a numeric anchor vector.
`p`	Working reconstruction rank. Defaults to the full tangent rank.
`normalize`	Logical. If `TRUE`, rows of `g` are normalized first.

Value

An object of class `"agca_fit"`.

agca_rank_summary *AGCA rank summary*

Description

AGCA rank summary

Usage

```
agca_rank_summary(fit)
```

Arguments

fit	An object returned by agca() or agca_fit_directions().

Value

A data frame with rank, residual risk, and variation explained.

agca_reconstruct *Reconstruct angular directions from leading AGCA components*

Description

Reconstruct angular directions from leading AGCA components

Usage

```
agca_reconstruct(fit, p = fit$p)
```

Arguments

fit	An object returned by agca() or agca_fit_directions().
p	Reconstruction rank.

Value

A matrix of reconstructed angular directions.

agca_residual_risk *Residual risk by AGCA rank*

Description

Residual risk by AGCA rank

Usage

```
agca_residual_risk(fit, max_rank = length(fit$eigenvalues))
```

Arguments

fit	An object returned by agca() or agca_fit_directions().
max_rank	Maximum rank to report.

Value

A numeric vector indexed by ranks 0:max_rank.

agca_standardize *Marginal standardization for AGCA*

Description

Marginal standardization for AGCA

Usage

```
agca_standardize(
  x,
  margin = c("rank_pareto", "pareto", "none"),
  cdf = NULL,
  ties_method = "average"
)
```

Arguments

x	A numeric matrix or data frame.
margin	Standardization method: "rank_pareto" (default), "pareto" for supplied CDFs, or "none" for already standardized data.
cdf	Optional CDF function or list of CDF functions for margin = "pareto".
ties_method	Tie method used by rank_pareto().

Value

A numeric matrix.

agca_variation_explained *Anchored variation explained*

Description

Anchored variation explained

Usage

```
agca_variation_explained(fit)
```

Arguments

fit	An object returned by agca() or agca_fit_directions().

Value

Cumulative anchored variation explained by each rank.

anchored_departures *Anchored departures*

Description

Computes the anchor coordinate and tangent departure $u_\mu(g) = (I - \mu\mu^\top)g$.

Usage

```
anchored_departures(g, mu, normalize = TRUE)
```

Arguments

g	A numeric matrix of directions.
mu	Anchor direction.
normalize	Logical. If TRUE, rows of g are normalized first.

Value

A list with normalized directions, anchor, anchor coordinates, and anchored departures.

anchor_sensitivity *Anchor sensitivity diagnostics*

Description

Anchor sensitivity diagnostics

Usage

```
anchor_sensitivity(x, k, anchors = c("canonical", "frechet", "principal"), ...)
```

Arguments

x	Numeric matrix or data frame.
k	Number of top radial observations.
anchors	Character vector of anchors to compare.
...	Additional arguments passed to agca().

Value

A data frame of rank summaries across anchors.

angular_functional_error *Functional approximation errors from angular reconstruction*

Description

Computes mean angular functional values on the fitted and reconstructed directions for a collection of portfolio weights.

Usage

```
angular_functional_error(fit, weights, ranks = fit$p, power = 1, cap = Inf)
```

Arguments

fit	An object returned by agca() or agca_fit_directions().
weights	A numeric vector or matrix. Rows are portfolios.
ranks	Integer vector of AGCA ranks.
power	Power applied to positive portfolio exposures.
cap	Optional finite cap applied to the powered exposure.

Value

A data frame with original, reconstructed, and relative errors.

bootstrap_agca *Nonparametric AGCA bootstrap*

Description

Resamples fitted angular directions and recomputes AGCA diagnostics.

Usage

```
bootstrap_agca(
  fit,
  B = 199L,
  ranks = NULL,
  fixed_anchor = TRUE,
  anchor = "canonical",
  seed = NULL
)
```

Arguments

fit	An object returned by agca() or agca_fit_directions().
B	Number of bootstrap resamples.
ranks	Integer ranks to summarize.
fixed_anchor	Logical. If TRUE, keep the fitted anchor fixed. Otherwise refit the requested anchor type.
anchor	Anchor used when fixed_anchor = FALSE.
seed	Optional random seed.

Value

An object of class "agca_bootstrap".

canonical_anchor *Canonical anchor*

Description

Canonical anchor

Usage

```
canonical_anchor(d)
```

Arguments

d	Ambient dimension.

Value

The balanced direction $d^{-1/2}(1, \ldots, 1)$.

---

decluster_runs *Runs declustering for radial extremes*

Description

Exceedances are split into clusters separated by more than run consecutive non-exceedances. The representative of each cluster is the observation with the largest radius.

Usage

```
decluster_runs(x, k = NULL, threshold = NULL, run = 1L)
```

Arguments

x	A numeric matrix of standardized observations.
k, threshold	Top-k count or radial threshold used to define exceedances.
run	Nonnegative run length.

Value

A list like tail_directions(), with one index per cluster.

---

frechet_anchor *Spherical Frechet anchor*

Description

Computes a Karcher-mean approximation to the spherical Frechet mean.

Usage

```
frechet_anchor(g, normalize = TRUE, max_iter = 100L, tol = 1e-10)
```

Arguments

g	Matrix of angular directions.
normalize	Logical. If TRUE, rows are normalized first.
max_iter	Maximum number of iterations.
tol	Convergence tolerance for the tangent update norm.

Value

A unit vector.

`normalize_rows` *Normalize matrix rows*

Description

Normalize matrix rows

Usage

```
normalize_rows(x)
```

Arguments

`x`	A numeric matrix with nonzero rows.

Value

A numeric matrix whose rows have Euclidean norm one.

`pareto_from_cdf` *Pareto marginal standardization from supplied CDFs*

Description

Pareto marginal standardization from supplied CDFs

Usage

```
pareto_from_cdf(x, cdf, eps = 1e-12)
```

Arguments

`x`	A numeric matrix or data frame.
`cdf`	A function applied to every margin, or a list of one CDF function per margin. Each CDF must return values in `[0, 1]`.
`eps`	Tail clipping constant used to avoid zero and infinite values.

Value

A numeric matrix with standard Pareto margins.

plot.agca_bootstrap *Plot AGCA bootstrap summaries*

Description

Plot AGCA bootstrap summaries

Usage

```
## S3 method for class 'agca_bootstrap'
plot(x, statistic = c("variation_explained", "residual_risk"), ...)
```

Arguments

x	An object returned by bootstrap_agca().
statistic	Statistic to plot.
...	Additional graphical arguments.

Value

Invisibly returns x.

plot.agca_fit *Plot AGCA output*

Description

Plot AGCA output

Usage

```
## S3 method for class 'agca_fit'
plot(x, type = c("variation", "scree", "scores", "loadings"), p = 1L, ...)
```

Arguments

x	An object returned by agca() or agca_fit_directions().
type	Plot type: eigenvalue scree plot, cumulative variation, first two scores, or loadings.
p	Component index used for the loadings plot.
...	Additional graphical arguments.

Value

Invisibly returns x.

principal_anchor *Principal angular anchor*

Description

Principal angular anchor

Usage

```
principal_anchor(g, normalize = TRUE)
```

Arguments

g	Matrix of angular directions.
normalize	Logical. If TRUE, rows are normalized first.

Value

The leading eigenvector of the angular second-moment matrix, oriented to have positive sum.

rank_pareto *Rank-Pareto marginal standardization*

Description

Transforms each margin to empirical standard Pareto scores using $(n+1)/(n+1-rank)$. Larger observations are treated as more extreme.

Usage

```
rank_pareto(x, ties_method = "average")
```

Arguments

x	A numeric matrix or data frame.
ties_method	Tie method passed to base::rank().

Value

A numeric matrix with standard Pareto-like margins.

row_norms *Row Euclidean norms*

Description

Row Euclidean norms

Usage

```
row_norms(x)
```

Arguments

x	A numeric matrix.

Value

A numeric vector containing one Euclidean norm per row.

simulate_agca_10d *Simulate the 10-dimensional AGCA example design*

Description

Generates the 10-dimensional heavy-tailed design used as the package example. Variables X1–X8 share a low-dimensional logistic-block extremal mechanism. Variables X9 and X10 contain independent Pareto sources, so selected extremes include near-axis regimes in addition to the shared low-rank angular structure.

Usage

```
simulate_agca_10d(
  n = 10000L,
  seed = NULL,
  theta = 0.45,
  tau = 0.25,
  axis9_scale = 1,
  axis10_scale = 1
)
```

Arguments

n	Number of observations.
seed	Optional random seed.
theta	Logistic dependence parameter in (0, 1).
tau	Nonnegative finite-threshold noise scale.
axis9_scale, axis10_scale	Positive scales for the independent Pareto sources in variables X9 and X10.

Value

A data frame with variables X1, ..., X10 and a factor regime giving the dominant latent source for each observation.

Examples

```
x <- simulate_agca_10d(n = 500, seed = 1)
fit <- agca(x[paste0("X", 1:10)], k = 75, p = 3)
agca_rank_summary(fit)
```

---

sphere_distance *Spherical geodesic distance*

Description

Computes great-circle distances on the unit sphere. If y has one row and x has several rows, the single direction in y is recycled.

Usage

```
sphere_distance(x, y, normalize = TRUE)
```

Arguments

x, y	Numeric matrices with the same number of columns.
normalize	Logical. If TRUE, rows are normalized before distances are computed.

Value

A numeric vector of geodesic distances in radians.

summary.agca_bootstrap *Summarize AGCA bootstrap output*

Description

Summarize AGCA bootstrap output

Usage

```
## S3 method for class 'agca_bootstrap'
summary(object, probs = c(0.025, 0.5, 0.975), ...)
```

Arguments

object	An object returned by bootstrap_agca().
probs	Quantile probabilities.
...	Unused.

Value

A data frame of bootstrap summaries by rank.

tail_directions *Top-k angular directions*

Description

Top-k angular directions

Usage

```
tail_directions(x, k = NULL, threshold = NULL)
```

Arguments

x	A numeric matrix of standardized observations.
k	Number of largest radii to retain. Specify either k or threshold.
threshold	Radial threshold. Observations with radius greater than threshold are retained.

Value

A list containing angular directions, radii, selected indices, and the threshold.

`threshold_stability` *Threshold stability diagnostics*

Description

Threshold stability diagnostics

Usage

```
threshold_stability(x, k, ...)
```

Arguments

`x`	Numeric matrix or data frame.
`k`	Integer vector of top-k values.
`...`	Additional arguments passed to `agca()`.

Value

A data frame of rank summaries across thresholds.

Index