Package 'ctgimme'

August 28, 2026

Title Continuous-Time Subgrouping with GIMME

Version 0.1.0

Description Estimates group-, subgroup-, and individual-level dynamic structures from multivariate intensive longitudinal data using continuous-time state-space models. The subgrouping procedure combines iterative shared-path searches with recurrent-evidence feature screening and partitioning around medoids. The continuous-time group iterative multiple model estimation method is described in Park et al. (2025) <doi:10.1080/10705511.2024.2429544>.

URL https://github.com/JPark93/ctgimme

BugReports https://github.com/JPark93/ctgimme/issues

License Apache License (== 2.0)

Encoding UTF-8

Language en-US

Depends R (>= 4.1.0)

Imports cluster, gtools, matrixStats, methods, OpenMx, parallel, ps, qgraph, stats, utils

Suggests expm, igraph, nloptr, testthat (>= 3.2.0)

Config/testthat/edition 3

Config/roxygen2/version 8.1.0

NeedsCompilation no

Author Jonathan J. Park [aut, cre, cph] (ORCID: <https://orcid.org/0000-0002-6636-3504>), Nathan Xin Mills [aut, ctb]

Maintainer Jonathan J. Park <imJPark@UCDavis.edu>

Repository CRAN

Date/Publication 2026-08-28 14:30:15 UTC

Contents

`ctgimme` *Estimate Continuous-Time Group, Subgroup, and Individual Networks*

Description

Fits continuous-time state-space models for multiple individuals and uses iterative modification-index searches to identify paths shared at the group, subgroup, and individual levels. When data-driven subgroup detection is enabled, the default method uses recurrent-evidence feature screening followed by partitioning around medoids (PAM).

Usage

```
ctgimme(
  varnames = NULL,
  dataframe = NULL,
  id = NULL,
  time = NULL,
  cores = 1,
  directory = NULL,
  sig.thrsh = 0.55,
  sub.sig.thrsh = 1,
  Galpha = 0.05,
  ben.hoch = TRUE,
  S.Galpha = 0.05,
  Ialpha = 0.01,
  ME.var = 1e-08,
  PE.var = NULL,
  subgroup.model = FALSE,
  time.intervals = c(1),
  keep.intermediate = FALSE,
  conduct = TRUE,
  subgroup.method = c("pam", "legacy"),
  max.subgroups = NULL,
  scale.data = FALSE,
  ME.free = FALSE,
  PE.free = FALSE,
  verbose = TRUE
)
```

Arguments

varnames	Nonempty character vector naming distinct process-variable columns in dataframe. These columns must be numeric. Their order defines the row and column order of the drift and noise matrices and the order of vector-valued ME.free and PE.free specifications. This argument is required.
dataframe	Data frame in long format, with one row per measurement occasion and columns named by id, time, and varnames. Identifier and time values must be nonmissing, time values and observed process values must be finite and numeric, and rows must already be ordered by time within subject. Missing process measurements are passed to OpenMx's raw-data likelihood. This argument is required.
id	Character scalar naming the subject-identifier column in dataframe. Each distinct value defines one subject. Values are converted to character for matching, membership labels, and intermediate filenames, so use nonmissing identifiers that are safe in filenames and remain distinct when compared without case. This argument is required.
time	Character scalar naming the numeric observation-time column in dataframe. Its units determine the scale of continuous-time parameters fit to the observed-time data. The multisubject subgroup model preserves the supplied elapsed intervals independently within each subject and rebases each person's first observation to local time zero. These values also define the units of time.intervals. The selected column is copied internally to a reserved column named Time, so Time should not name a process variable. This argument is required.
cores	Positive integer giving the requested number of parallel PSOCK workers. The default is 1, which runs without a worker pool. Requests are reduced only to the number of subjects; the package imposes no additional worker maximum. Users are responsible for requesting a value supported by their machine or computing allocation. One pool is reused across fitting stages, and OpenMx uses one thread in each R process.
directory	Character scalar giving the output directory for model, modification-index, plot, membership, and diagnostic artifacts. It is created recursively when necessary. Existing files with package-defined names may be replaced or cleaned up, so a new or dedicated directory is recommended. This argument is required.
sig.thrsh	Numeric scalar in [0, 1] giving the minimum support proportion for group-level paths. A candidate is added when at least this proportion of subjects with usable modification-index results support it at the current Galpha threshold; the proportion also governs low-support pruning. No shared path is added unless more than half of the requested subjects have usable modification-index results. The default is 0.55.
sub.sig.thrsh	Numeric scalar in [0, 1] with two roles. Values below 1 enable the detector selected by subgroup.method and set the support proportion for paths added and retained within each recovered membership group. The default, 1, bypasses data-driven detection and assigns every subject to membership group 1; it does not skip the subsequent within-group shared-path search or individual-model fitting.
Galpha	Numeric scalar between zero and one giving the base significance level for modification-index evidence during group-level path additions. The default is

0.05; ben.hoch determines whether it is used directly or as the start of an adapted threshold sequence.

ben.hoch Logical scalar. If TRUE (the default), use the adapted Benjamini-Hochberg threshold sequence during group-, subgroup-, and individual-level path additions. If FALSE, use the corresponding Galpha, S.Galpha, or Ialpha value at every addition step. This option does not control the separate, fixed BH correction used to screen recurrent features for PAM subgrouping.

S.Galpha Numeric scalar between zero and one giving the base significance level for modification-index evidence during subgroup-level path additions. The default is 0.05; its stepwise use is controlled by ben.hoch.

Ialpha Numeric scalar between zero and one giving the base significance level for individual-level path additions and the direct significance cutoff for individual-level pruning. The default is 0.01; ben.hoch controls its stepwise use during additions.

ME.var Finite, nonnegative measurement-error variances, supplied as a numeric scalar or a diagonal numeric matrix with one row and column per process variable. A scalar is repeated down the diagonal. The default is 1e-8. Entries selected by ME.free are starting values and must be at least 1e-8; other entries remain fixed. These values populate the OpenMx R matrix.

PE.var Finite, nonnegative process-noise (innovation) variances, supplied as NULL, a numeric scalar, or a diagonal numeric matrix with one row and column per process variable. A scalar is repeated down the diagonal; the default, NULL, uses an identity matrix. Entries selected by PE.free are starting values and must be at least 1e-5; other entries remain fixed. These values populate the OpenMx Q matrix.

subgroup.model Logical scalar. If TRUE, fit one additional joint continuous-time model for each recovered membership group using that group's final drift structure. Every member contributes an independent state-space likelihood and fresh filter initialization. The subgroup has one shared A/Q/R specification, whose free entries are optimized jointly once. The fit writes a parameter plot, the requested discrete-time transition plots, and one subgroup-specific fitted-model RDS file. No subject series are chained or separated by synthetic rows. The default is FALSE; subgroup structural searches and individual-model fits are still performed. This option requires the suggested package expm.

time.intervals Nonempty numeric vector of nonnegative elapsed-time values in the same units as time. When subgroup.model = TRUE, every value delta produces a plot of the fitted discrete-time transition matrix exp(A * delta) for each membership group. The default is 1; this argument is otherwise ignored.

keep.intermediate
Logical scalar controlling end-of-run cleanup. If FALSE (the default), temporary per-subject Model_<id>.RDS and MI_<id>.RDS files from the group- and subgroup-level shared-path searches are removed after a successful run. If TRUE, the final set of those files is retained; iterations may still replace earlier versions. Final individual models, fitted subgroup models, and summary artifacts are retained under either setting.

conduct Logical scalar used only when subgroup.method = "legacy" and subgroup detection is enabled. If TRUE (the default), nloptr selects a quantile cutoff for the

subject-similarity graph using the legacy conductance objective before Walktrap clustering. If FALSE, Walktrap uses the unoptimized similarity weights after shifting their minimum to zero. TRUE requires the suggested package nloptr. This argument is ignored by the default "pam" method and when sub.sig.thrsh = 1.

subgroup.method

Character scalar selecting the data-driven subgroup detector. The default, "pam", screens signed, non-group path evidence for recurrence, computes mean Manhattan distances, and chooses a partitioning-around-medoids solution by average silhouette width. "legacy" constructs a weighted subject-similarity graph and applies Walktrap community detection; it requires the suggested package igraph. PAM's recurrence screen uses a fixed 0.05 subject-level evidence threshold and retains features with a BH-adjusted recurrence value no greater than 0.05. The choice is inactive when sub.sig.thrsh = 1.

max.subgroups NULL or a single integer. When PAM subgrouping is active, an integer at least 2 is required and gives the requested maximum number of subgroups. Candidate counts run from two through the smaller of max.subgroups and one less than the number of subjects with usable subgroup features. This argument is ignored for "legacy" subgrouping and when sub.sig.thrsh = 1.

scale.data Logical scalar. If TRUE, center and standardize every process variable separately within each subject before all estimation stages; within-subject constant columns are centered to zero. The default is FALSE, which preserves the supplied scale.

ME.free Logical scalar, vector with one entry per process variable, or square logical matrix with one row and column per process variable, selecting the diagonal measurement-error variances to estimate. A scalar is repeated down the diagonal; a matrix must have FALSE off the diagonal. The default is FALSE, which fixes all measurement-error variances at ME.var.

PE.free Logical scalar, vector with one entry per process variable, or square logical matrix with one row and column per process variable, selecting the diagonal process-noise variances to estimate. A scalar is repeated down the diagonal; a matrix must have FALSE off the diagonal. The default is FALSE, which fixes all process-noise variances at PE.var.

verbose Logical scalar controlling informational progress output. If TRUE (the default), stage messages and optimizer reporting from fits in the main R process are shown; informational PSOCK worker output is not forwarded by the parallel backend. If FALSE, package messages and main-process optimizer reports are suppressed. Worker warnings are relayed through the main process regardless of verbose. Warnings and errors remain visible and can be handled with standard R condition tools such as suppressWarnings() and tryCatch().

Details

Estimation is computationally intensive and writes intermediate and final artifacts to directory. Every successfully completed run writes GStruc.RDS, subgroup_detection.rds, subgroup_membership.csv, and Subgroups Plot.png, as well as a subgroup structure file for each membership group and a shared individual-model directory. When usable PAM distances or legacy similarities are available,

the PNG is a two-dimensional subject-distance map: closer nodes have more similar individual model evidence, nearest neighbors are connected, and colors and enclosing hulls indicate membership. Subjects lacking a diagnostic distance are noted but omitted from the map. A complete membership roster is plotted when no usable pairwise distance is available.

With `subgroup.model = TRUE`, each subgroup directory `Models/Subgroup <g>/` additionally contains `Subgroup <g> Params.png`, one `Subgroup <g> Delta_t = <delta>.png` transition plot for every value in `time.intervals`, and `Subgroup_<g>Model.RDS`. The RDS contains one fitted OpenMx model with shared subgroup parameters and one internal likelihood block per person; those blocks are not separately estimated subject models. The transition plots display `exp(A * delta)`, the discrete-time form of the fitted continuous-time drift at the requested elapsed time.

PAM subgrouping requires at least three usable subject model and modification-index pairs plus at least one eligible off-diagonal, non-group path that passes its fixed recurrence screen. It chooses among candidate counts from two through `max.subgroups`, bounded above by the number of usable subjects minus one, using average silhouette width. Legacy subgrouping requires at least two usable subject model and modification-index pairs. If these requirements are not met, all subjects are assigned to membership group 1. When only some subjects have usable artifacts, either detector assigns the remaining subjects to group 1 in the complete membership output. The method-native memberships can cover only the usable subjects: PAM's `$clustering` vector and the legacy communities object omit unusable subjects, although returned PAM objects also receive the complete `$membership` vector described above.

`ME.free` and `PE.free` affect diagonal variances only; measurement-error and process-noise covariances remain fixed to zero. Estimating both variance sets generally requires more within-person information than estimating process noise while fixing measurement error. Estimated values are stored in the `R` and `Q` matrices, respectively, of saved OpenMx models. The `verbose` argument controls package progress and main-process OpenMx reporting; the parallel backend does not forward raw worker output. Internal fits do not use the interactive progress callback, which avoids `imxReportProgress` lookup failures reported in some RStudio/OpenMx installations.

Value

If subgroup detection is disabled or falls back to one membership group, an invisible list with elements `message`, `G.DRIFT`, `membership`, `directory`, `clustering`, and `subgroup.detection`. Thus every successful run returns an object from which the complete membership can be extracted. With enabled subgroup detection, a successful PAM run returns a `cluster::pam` object augmented with a complete, ID-aligned `membership` element; a successful legacy run returns an `igraph` communities object. Both successful clustering objects carry the complete ID-aligned vector in the `ctgimme.membership` attribute and the output paths in the `ctgimme.membership.artifacts` attribute. In every successfully completed case, the complete membership is also written to `subgroup_membership.csv` and saved in `subgroup_detection.rds`.

References

Park, J. J., Fisher, Z. F., Hunter, M. D., Shenk, C., Russell, M., Molenaar, P. C. M., & Chow, S.-M. (2025). Unsupervised model construction in continuous-time. *Structural Equation Modeling: A Multidisciplinary Journal, 32*(3), 377–399. doi:10.1080/10705511.2024.2429544

See Also

ctgimme_demo() for a fast, estimation-free demonstration of the subgroup selection step.

Examples

```
quick <- ctgimme_demo()
table(quick$membership)

fit_membership <- local({
  set.seed(123)
  n_time <- 12L
  observations <- data.frame(
    id = rep(c("S1", "S2"), each = n_time),
    time = rep(0:(n_time - 1L), 2L),
    x = c(
      stats::arima.sim(model = list(ar = 0.6), n = n_time),
      stats::arima.sim(model = list(ar = 0.6), n = n_time)
    )
  )
  example_output <- tempfile("ctgimme-example-")
  on.exit(
    unlink(example_output, recursive = TRUE, expand = FALSE),
    add = TRUE
  )
  fit <- ctgimme(
    varnames = "x",
    dataframe = observations,
    id = "id",
    time = "time",
    directory = example_output,
    ME.var = 0.05,
    PE.var = 1,
    verbose = FALSE
  )
  fit$membership
})
fit_membership
```

---

ctgimme_demo *Run a Fast ctgimme Subgrouping Demonstration*

Description

Demonstrates the recurrent-evidence screening, distance calculation, and PAM subgroup selection used by ctgimme() without fitting OpenMx models. The deterministic example is deliberately small and completes in well under a second.

Usage

```
ctgimme_demo()
```

Value

A list with the synthetic signed path-score `features`, the feature-selection diagnostics, the mean Manhattan `distance` matrix, candidate silhouette widths in `candidates`, and the named subgroup `membership`.

See Also

`ctgimme()`

Examples

```
demo_result <- ctgimme_demo()
demo_result$membership
demo_result$candidates
```

Index