Package: fastFGEE (via r-universe)

October 21, 2026

Title Fast Functional Generalized Estimating Equations via a One-Step Estimator

Version 0.2.2

Description Fits functional generalized estimating equations for longitudinal functional outcomes and covariates using a one-step estimator that is fast even for large cluster sizes or large numbers of clusters. The package supports quasi-likelihoods derived from a range of distributions, with substantial simulations run for quasi-likelihoods derived from Gaussian, binomial, Poisson, negative binomial, Gamma and beta families. It supports common link functions and several working correlation structures. An optimized engine constructs cluster score and sensitivity statistics in one pass, provides coefficient-space Gaussian cross-validation, analytic-gradient fast cluster cross-validation, and an experimental sandwich-scaled working restricted quasi-likelihood selector. Internal compiled routines provide symmetric positive-definite Cholesky solves and exact tridiagonal precision operations for irregularly sampled continuous-time AR(1) working correlations. Uncertainty quantification is based on sandwich variance estimators and studentized wild cluster bootstrap procedures for cluster-robust pointwise intervals and optional simultaneous bands. The package implements methods described in Loewinger et al. (2025) <https://pmc.ncbi.nlm.nih.gov/articles/PMC12306803/>.

License GPL (>= 3)

Encoding UTF-8

RoxygenNote 7.3.3

Depends R (>= 4.2)

Imports data.table, ggplot2, gridExtra, MASS, Matrix, mgcv, Rcpp, refund (>= 0.1-40)

Suggests knitr, rmarkdown, RcppArmadillo, SuperGauss, testthat (>= 3.0.0)

VignetteBuilder knitr

URL https://github.com/gloewing/fastFGEE

BugReports https://github.com/gloewing/fastFGEE/issues

LinkingTo Rcpp

NeedsCompilation yes

Config/testthat/edition 3

Author Gabriel Loewinger [aut, cre] (ORCID:
<https://orcid.org/0000-0002-0755-8520>)

Maintainer Gabriel Loewinger <gloewinger@gmail.com>

Repository https://cran.r-universe.dev

Date/Publication 2026-09-21 10:10:02 UTC

RemoteUrl https://github.com/cran/fastFGEE

RemoteRef HEAD

RemoteSha ed22ecde8ce9fbdfbe7ed1ce5a36663734655328

Contents

d 2
fgee 3
fgee.plot 7

Index **9**

`d` *Simulated longitudinal functional example data*

Description

A simulated dataset used in examples and testing for **fastFGEE**. The object `d` is a data frame containing a functional response stored in an `AsIs` matrix column together with a cluster identifier, two scalar covariates, and a longitudinal time variable.

Usage

```
data(d)
```

Format

A data frame with 5 variables:

`Y` An `AsIs` matrix-valued column containing the functional response. In the included example, the matrix has 100 columns named `Y_1` to `Y_100`.

`ID` Cluster identifier.

`X1` First scalar covariate.

`X2` Second scalar covariate.

`time` Longitudinal time variable.

Details

This dataset is intended for package examples, vignettes, and quick testing of `fgee`. It represents a simulated binary-response setting on a common functional grid.

Source

Simulated for the package examples.

`fgee` *Fit a One-Step Functional Generalized Estimating Equation*

Description

Fits longitudinal functional-response generalized estimating equations using the validated one-step estimator.

Usage

```
fgee(
  formula,
  data,
  cluster,
  family,
  corr_fn = "ar1",
  corr_long = "ar1",
  time = NULL,
  long.dir = TRUE,
  var.type = "sandwich",
  pffr.mod = NULL,
  knots = NULL,
  bs = "bs",
  cv = "fastkfold",
  cv.grid = NULL,
  rho.smooth = FALSE,
  rho.pool = c("fn", "none", "long", "both"),
  joint.CI = "wild",
  linpred_method = c("accumulate", "matrix"),
  clip_mu = 0,
  m.pffr = c(2, 1),
  check_alignment = TRUE,
  boot.samps = 3000,
  sp.method = c("auto", "fastk_staged", "fastk_grad", "fastk_grad_fast",
    "sandwich_qreml", "qreml_fastk"),
  working.retain = c("auto", "scores", "aggregate", "full"),
  corr.solver = c("auto", "exact", "supergauss"),
  fastk.K = 10L,
  fastk.seed = 1L,
```

```
  fastk.memory = c("balanced", "speed", "lowmem"),
  fastk.start = c("qreml", "robust", "compact"),
  fastk.kernel = fgee_fastk_kernel_ok(),
  qreml.phi.method = c("penalized", "all", "fixed"),
  qreml.phi.fixed = NULL,
  qreml.phi.weight = 1,
  qreml.phi.clip = c(1, 8),
  keep.tuning.workspace = FALSE,
  verbose.tuning = TRUE,
  keep.data = TRUE,
  keep.initial.fit = TRUE,
  keep.working.stats = TRUE,
  ...
)
```

Arguments

formula	A model formula whose left-hand side is a functional response.
data	A data frame containing the variables in formula.
cluster	Name of the cluster identifier column.
family	A family object or family name understood by refund::pffr(). Substantial simulation evidence exists for quasi-likelihoods derived from Gaussian, binomial, Poisson, negative binomial, Gamma and beta families; other families supported by refund::pffr() are likely to work but have not been examined as thoroughly. For negative binomial outcomes use mgcv::nb(), and for proportion outcomes use mgcv::betar(); MASS::negative.binomial() is not accepted and raises an informative error.
corr_fn	Working correlation in the functional direction: "independent", "exchangeable", "ar1", or "fpca".
corr_long	Working correlation in the longitudinal direction: "independent", "exchangeable", or "ar1".
time	Optional name of the longitudinal ordering variable.
long.dir	Logical retained for backwards compatibility.
var.type	Variance estimator: "sandwich", "fastboot", or "boot".
pffr.mod	Optional fitted refund::pffr() object used as the initial estimator.
knots	Number of spline knots for the initial pffr() fit.
bs	Basis type passed to refund::pffr().
cv	Cross-validation mode used for smoothing-parameter selection. The public one-step interface currently supports only "fastkfold".
cv.grid	Optional grid or staged grids of smoothing parameters.
rho.smooth	Logical; smooth pointwise working-correlation estimates.
rho.pool	Controls pooling of the working-correlation parameter when only one direction is correlated. "fn" (default) estimates a single pooled functional correlation when corr_long = "independent"; "long" pools the longitudinal correlation

when corr_fn = "independent"; "both" pools either; "none" lets the surviving parameter vary over the opposite index, reproducing the behaviour of earlier versions. When both directions are correlated the working correlation is separable and both parameters are pooled regardless of this setting.

joint.CI — Controls interval construction. "wild" requests the studentized wild-cluster procedure and its optional simultaneous band. A requested band is a simultaneous-inference procedure, not a guarantee of nominal finite-sample coverage.

linpred_method — Method used to form linear predictors.

clip_mu — Lower bound used for numerical stabilization of fitted means.

m.pffr — Penalty-order specification passed to refund::pffr().

check_alignment — Logical; check alignment between the wide data and the long representation produced by pffr().

boot.samps — Number of bootstrap replicates when applicable.

sp.method — Smoothing-parameter selector. "auto" uses staged fastK for Gaussian identity-link models and "fastk_grad_fast" otherwise. Explicit choices are "fastk_staged", "fastk_grad", "fastk_grad_fast", "sandwich_qreml", and "qreml_fastk". The last is retained for reproducibility but normally reaches the same exact fastK solution by a slower route.

working.retain — Retention profile for compact working statistics: "auto", "scores", "aggregate", or "full".

corr.solver — Correlation inverse backend. "auto" uses direct package operators; "supergauss" forces the optional Toeplitz backend.

fastk.K — Number of cluster folds for optimized fastK tuning.

fastk.seed — Seed used to construct optimized fastK folds.

fastk.memory — Non-Gaussian fastK workspace. "balanced" is recommended and is the only layout eligible for the compiled loss/gradient kernel. "lowmem" is deprecated because validation found no peak-memory advantage and it cannot use the compiled kernel.

fastk.start — Initialization strategy for analytic-gradient fastK.

fastk.kernel — Logical; allow the compiled loss/gradient kernel when supported.

qreml.phi.method — Information-scaling diagnostic used by sandwich qREML.

qreml.phi.fixed — Fixed information scale when qreml.phi.method = "fixed".

qreml.phi.weight — Weight applied to the information mismatch before clipping.

qreml.phi.clip — Lower and upper safeguards for the qREML information scale.

keep.tuning.workspace — Logical; retain large tuning workspaces for debugging.

verbose.tuning — Logical; print optimized tuning progress.

keep.data — Logical; retain long-format working data in the fitted object.

`keep.initial.fit`
: Logical; retain the initial `pffr()` fit.

`keep.working.stats`
: Logical; retain compact initial and final working statistics.

`...`
: Reserved to detect removed estimator controls. Unknown entries are rejected rather than silently ignored.

Details

Public calls always use one coefficient update and the optimized working-statistics engine. FastK selectors use the one-step fast cluster cross-validation construction; `"sandwich_qreml"` remains an explicit opt-in selector rather than a cross-validation method. Historical exact-GLS, legacy-engine, pffr-only, and fully iterated controls remain internal for regression testing and method development and are not accepted by `fgee()`.

Pointwise intervals and simultaneous bands answer different inferential questions. A simultaneous band uses a whole-curve maximum statistic in the asymptotic or resampling procedure. Nominal finite-sample simultaneous coverage is not guaranteed, particularly with few or high-leverage independent clusters or unstable nuisance and correlation estimates.

Value

An object of class `"fgee1step"`. Important components include `beta`, `vb`, `model`, `lambda`, and `tuning`.

References

Loewinger, G., Levis, A. W., Cui, E., and Pereira, F. (2025). Fast Penalized Generalized Estimating Equations for Large Longitudinal Functional Datasets.

See Also

`fgee.plot`

Examples

```
## Not run:
data("d", package = "fastFGEE")
fit <- fgee(
  Y ~ X1 + X2,
  data = d,
  cluster = "ID",
  family = binomial(link = "logit"),
  time = "time",
  corr_long = "exchangeable",
  corr_fn = "independent"
)
fgee.plot(fit)

## End(Not run)
```

`fgee.plot` *Plot coefficient estimates from a fitted* `fastFGEE` *model*

Description

Produces coefficient plots with pointwise intervals and optional simultaneous confidence bands when available. A displayed simultaneous band represents the fitted procedure's whole-curve calibration; it is not a guarantee of nominal finite-sample simultaneous coverage, especially with few independent clusters.

Usage

```
fgee.plot(
  fit,
  num_row = NULL,
  xlab = "Functional Domain",
  title_names = NULL,
  ylim = NULL,
  align_x = NULL,
  x_rescale = 1,
  y_val_lim = 1.1,
  y_scal_orig = 0.05,
  return = FALSE,
  terms.plot = TRUE,
  all.terms = TRUE,
  int.uncertainty = FALSE
)
```

Arguments

`fit`	A fitted object returned by `fgee`.
`num_row`	Number of rows used when arranging plots.
`xlab`	X-axis label.
`title_names`	Optional replacement titles for coefficient plots.
`ylim`	Optional y-axis limits.
`align_x`	Optional value used to re-center the x-axis.
`x_rescale`	Optional x-axis rescaling factor.
`y_val_lim`	Expansion factor for the upper y-axis limit.
`y_scal_orig`	Expansion factor for the lower y-axis limit.
`return`	Logical; if `TRUE`, return the plotting data instead of only drawing the plots.
`terms.plot`	Logical; retained for backwards compatibility.
`all.terms`	Logical; retained for backwards compatibility.
`int.uncertainty`	Logical; retained for backwards compatibility.

Value

Invisibly returns a list of plotting data frames when `return = TRUE`; otherwise draws plots.

Index

∗ **datasets**
d, 2

d, 2

fgee, 3, *3*, *7*
fgee.plot, *6*, 7