Package 'SurvSPro'

June 19, 2026

Type Package

Title Survival Prediction with Spatially Adjusted Protein Summaries

Version 0.1.0

Maintainer Seungjun Ahn <seungjun.ahn@mountsinai.org>

Description A survival prediction framework using spatially adjusted protein summaries from spatial proteomics data, including imaging mass cytometry data. Cell-level protein intensities are modeled with spatial spline regression to estimate spatially adjusted mean expression and residual variance. Methodological details are described in Ahn et al. (2026) <doi:10.64898/2026.06.08.730964>.

License GPL-3

Encoding UTF-8

Depends R (>= 3.5.0)

Imports dplyr, mgcv, survival, sp

Suggests testthat (>= 3.0.0)

RoxygenNote 7.3.2

LazyData true

NeedsCompilation no

Author Seungjun Ahn [cre, aut] (ORCID:
<https://orcid.org/0000-0002-4816-8924>),
Eun Jeong Oh [aut] (ORCID: <https://orcid.org/0000-0001-8949-6564>),
Diddier Prada [ctb],
Ali Shojaie [ctb]

Repository CRAN

Date/Publication 2026-06-19 11:30:07 UTC

Contents

`cells_example_df` *Simulated spatial proteomics dataset.*

Description

A simulated data containing patient ID, spatial coordinates (u, v), and protein intensity values for a given protein.

Usage

```
cells_example_df
```

Format

An object of class `data.frame` with 50000 rows and 4 columns.

Details

The simulated spatial proteomics dataset includes 100 patients with their spatial coordinates and protein intensity

Source

Simulated using code in ‘inst/scripts/cells_example_df.R‘

`fit_spatial_cox` *fit_spatial_cox*

Description

Fits a Cox proportional hazards model for time-to-event outcomes using regresses clinical covariates and spatially adjusted protein summaries generated by `gam_features()`.

Usage

```
fit_spatial_cox(
  surv_df,
  features_df,
  pid = "patient_id",
  time = "time",
  status = "status",
  clin_cols = c("z1", "z2", "z3"),
  sp_cols = c("mu_sp", "tau_sp")
)
```

Arguments

`surv_df`	survival data frame
`features_df`	output from gam_features() function
`pid`	variable name of patient ID
`time`	variable name of the survival time
`status`	variable name of the event indicator
`clin_cols`	vector of clinical covariate names
`sp_cols`	spatial feature names from gam_features() function

Value

A fitted `coxph` object. The model includes standardized clinical covariates and spatially adjusted protein summaries as predictors of the survival outcome.

Examples

```
# cells_example_df: contains pid, coordinates (u, v), and intensity of a given protein
data(cells_example_df)
data(surv_example_df)

features_df = gam_features(cells_df   = cells_example_df,
                           pid        = "patient_id",
                           coord_u    = "u",
                           coord_v    = "v",
                           intensity  = "intensity",
                           grid_side  = 60,
                           k          = 20)

fit = fit_spatial_cox(surv_df     = surv_example_df,
                      features_df = features_df,
                      pid         = "patient_id",
                      time        = "time",
                      status      = "status",
                      clin_cols   = c("z1", "z2", "z3"))
summary(fit) ## To obtain coefficients, hazard ratios, and p-values
```

`gam_features` *gam_features*

Description

Captures spatial trends in cell-level protein expression and extracts spatially adjusted protein summaries, including spatially adjusted mean expression and residual variance reflecting cell-to-cell variability unexplained by spatial effects.

Usage

```
gam_features(
  cells_df,
  pid = "patient_id",
  coord_u = "u",
  coord_v = "v",
  intensity = "intensity",
  grid_side = 60,
  k = 20
)
```

Arguments

`cells_df`	data frame containing cell-level data
`pid`	variable name of the patient ID
`coord_u`	variable name of the u-axis coordinate
`coord_v`	variable name of the v-axis coordinate
`intensity`	variable name of the intensity for a given protein
`grid_side`	number of grid points along each of the u and v axes
`k`	basis dimension for the GAM smooth term

Value

A data frame with one row per patient and columns: `patient_id`, `mu_sp`, and `tau_sp`. Here, `mu_sp` is the spatially adjusted mean expression and `tau_sp` is the residual variance from the fitted spatial model.

Examples

```
#  cells_example_df: contains pid, coordinates (u, v), and intensity of a given protein
data(cells_example_df)

features_df = gam_features(cells_df   = cells_example_df,
                           pid        = "patient_id",
                           coord_u    = "u",
                           coord_v    = "v",
                           intensity  = "intensity",
                           grid_side  = 60,
                           k          = 20)
```

`surv_example_df` *Simulated patient-level survival data*

Description

A simulated dataset containing patient ID, three clinical covariates, survival time, and an event indicator (i.e, censoring variable).

Usage

```
surv_example_df
```

Format

An object of class `data.frame` with 100 rows and 6 columns.

Details

This simulated dataset includes 100 patients and is used with spatial proteomics features generated from `cells_example_df`.

Source

Simulated using code in ‘inst/scripts/surv_example_df.R‘

Index