Package 'doct'

July 6, 2020

Type Package

Title Decisions Optimized in Continuous Time

Version 1.0

Date 2020-07-05

Author William Hua <whua4@jhu.edu>, Yanxun Xu <yanxun.xu@jhu.edu>

Maintainer William Hua <whua4@jhu.edu>

Description Our package implements methods from Personalized Dynamic Treatment Regimes in Continuous Time: A Bayesian Joint Model for Optimizing Clinical Decisions with Timing. This paper builds a generative model for sequence of medical interventions---which are discrete events in continuous time---with a marked temporal point process (MTPP) where the mark is the assigned treatment or dosage. This clinical action model is then embedded into a Bayesian joint framework where the other components model clinical observations including medical longitudinal measurements and patient survival. Moreover, we propose a policy gradient method to learn the personalized optimal clinical decision that maximizes patient survival by interacting the MTPP with the model on clinical observations while accounting for uncertainties in clinical observations.

License CC0

Imports Rcpp (>= 1.0.4.6), MASS, pracma, mvtnorm, LaplacesDemon, RcppEigen, RcppNumerical, RcppArmadillo, stats, GoFKernel, survival

LinkingTo Rcpp, RcppArmadillo, RcppEigen, RcppNumerical

RoxygenNote 7.1.0

Encoding UTF-8

Depends R (>= 2.10)

Archs x64

R topics documented:

doct-package *Decisions Optimized in Continuous Time*

Description

Our package implements methods from Personalized Dynamic Treatment Regimes in Continuous Time: A Bayesian Joint Model for Optimizing Clinical Decisions with Timing. This paper builds a generative model for sequence of medical interventions—which are discrete events in continuous time—with a marked temporal point process (MTPP) where the mark is the assigned treatment or dosage. This clinical action model is then embedded into a Bayesian joint framework where the other components model clinical observations including medical longitudinal measurements and patient survival. Moreover, we propose a policy gradient method to learn the personalized optimal clinical decision that maximizes patient survival by interacting the MTPP with the model on clinical observations while accounting for uncertainties in clinical observations.

Details

The main functions in the package are mcmc_joint and SGD_run. mcmc_joint runs the Bayesian parameter estimation and collects the MCMC samples to be used SGD_run, which optimizes the dosage and visitation scheduling policy. The other functions in this package are for analyzing the results of these two main functions. For example, DIC_calculation outputs the DIC for simulated data analysis. This package accompanies our statistical methodology paper.

Our joint model has three submodels. For patient i, the visitation intensity function and dosing submodel is expressed below for $t \in (t_{i,j}, t_{i,j+1}]$:

$$\lambda_i(t) = \exp(\mu) + \ \alpha_{i,j}(t - t_{i,j})^{\kappa-1} e^{-\gamma(t-t_{i,j})} \frac{\gamma^\kappa}{\Gamma(\kappa)}$$

$$\alpha_{i,j} = \frac{\xi}{1 + \exp((1, y_{i,j})\boldsymbol{\beta_\alpha})} \quad \frac{\kappa - 1}{\gamma} = \exp(\nu_1), \ \ \kappa = \exp(\nu_2) + 1$$

$$d_{i,j} = (1, y_{i,j}, \boldsymbol{x}_i)\boldsymbol{\beta}_d + \epsilon_d, \ \ \epsilon_d \sim \text{Normal}(0, \sigma_d^2)$$

The longitudinal submodel for the log creatinine levels is represented below, with the following likelihood function:

$$y_i(t) = y_i^*(t) + \epsilon_l = \boldsymbol{z}_i(t)\boldsymbol{\beta}_l + \boldsymbol{r}_i(t)\boldsymbol{b}_i + \epsilon_l, \ \ \epsilon_l \sim \text{Normal}(0, \sigma_l^2), \ \ \boldsymbol{b}_i \sim \text{Normal}(\mathbf{0}, \boldsymbol{\Sigma_b})$$

$$\boldsymbol{z}_i(t) = (1, d_i(t), \boldsymbol{x}_i, t, t^2), \ \ \boldsymbol{r}_i(t) = (1, d_i(t), t),$$

The survival submodel is represented using the hazard function:

$$h_i(t) = \exp\Big(- (\beta_{s1} y_i^*(t) + \beta_{s2} d_i(t) + \beta_{s3} \text{Tox}_i(t) + \beta_{s4} \alpha_i(t) + h_0)\Big) s t^{s-1}$$

Author(s)

William Hua <whua4@jhu.edu>, Yanxun Xu <yanxun.xu@jhu.edu>

Maintainer: William Hua <whua4@jhu.edu>

`Generate_Simulated_Data` *Generate Simulated Data*

Description

Generate Simulated Data

Usage

```
Generate_Simulated_Data(
  Npat,
  beta_l = c(5.3, 0.1, 0.3, 0.4, 0.25, -1 * 10^(-4), 3 * 10^(-8)),
  sigma2_l = 0.1^2,
  beta_d = c(1, 0.2, 0.15, 0.2, 0.15),
  sigma2_d = 0.3^2,
  nu_1 = 2.5,
  nu_2 = 1.5,
  mu = -4.8,
  beta_s1 = 1,
  beta_s2 = 0.9,
  beta_s3 = -0.75,
  beta_s4 = -5,
  h0 = 5,
  s = 1.05
)
```

Arguments

`Npat`	Number of patients to simulate.
`beta_l`	Linear coefficient parameter in longitudinal submodel
`sigma2_l`	Error parameter in longitudinal submodel
`beta_d`	Linear coefficient parameter in dosing submodel
`sigma2_d`	Error parameter in dosing submodel
`nu_1`	Visitation intensity peak parameter in vistation submodel
`nu_2`	Visitation intensity shape parameter in vistation submodel
`mu`	Baseline visitation intensity parameter in vistation submodel
`beta_s1`	Creatinine-associated parameter in survival submodel
`beta_s2`	Dosage-associated parameter in survival submodel
`beta_s3`	Toxicity-associated parameter in survival submodel
`beta_s4`	Visitation-associated parameter in survival submodel
`h0`	Baseline hazard parameter in survival submodel
`s`	Shape parameter in survival submodel

Value

List of data. X0_inds is the baseline data for each patient, D is the dosage, Y is the longitudinal process, and Ts are the visit times, id are the patient id's (ranging from 1 to the total number of patients), censor and surv_time are the censoring and survival times, and Npat is the total number of patients.

Examples

```
## Not run:
#Simulate Data
seed=203
set.seed(seed)
data_all<-Generate_Simulated_Data(10)

## End(Not run)
```

`mcmc_joint` *MCMC function for joint longitudinal, survival, and dosage and visitation model*

Description

MCMC function for joint longitudinal, survival, and dosage and visitation model

Usage

```
mcmc_joint(data_all, mcmc_settings, seed)
```

Arguments

`data_all`	A list of complete data. X0_inds is the baseline data for each patient, D is the dosage, Y is the longitudinal process, and Ts are the visit times, id are the patient id's (ranging from 1 to the total number of patients), censor and surv_time are the censoring and survival times, and Npat is the total number of patients.
`mcmc_settings`	A list for MCMC setup. burn.in is the number of total iterations to burn, ndisplay is number of iterations per which the display message will appear, and Niter is the total number of iterations (including burn-in iterations).
`seed`	Seed for running MCMC

Value

`beta_l`	Linear coefficient parameter in longitudinal submodel
`sigma2_l`	Error parameter in longitudinal submodel
`beta_d`	Linear coefficient parameter in dosing submodel
`sigma2_d`	Error parameter in dosing submodel
`nu_1`	Visitation intensity peak parameter in vistation submodel
`nu_2`	Visitation intensity shape parameter in vistation submodel
`mu`	Baseline visitation intensity parameter in vistation submodel
`beta_s1`	Creatinine-associated parameter in survival submodel

beta_s2	Dosage-associated parameter in survival submodel
beta_s3	Toxicity-associated parameter in survival submodel
beta_s4	Visitation-associated parameter in survival submodel
h0	Baseline hazard parameter in survival submodel
shape	Shape parameter in survival submodel

Examples

```
## Not run:
data("simulated_data")
############
#Run MCMC
mcmc_settings=NULL
mcmc_settings$Niter=100
mcmc_settings$burn.in=10
mcmc_settings$ndisplay=10
mcmc_settings$peak_dist='gamma'
thin=10
post_thin_iters<-seq(mcmc_settings$burn.in+thin,mcmc_settings$Niter,thin)
seed=203
set.seed(seed)
mcmc_out_joint<-mcmc_joint(simulated_data,mcmc_settings,seed)


## End(Not run)
```

---

mcmc_separate *MCMC function for model with separate longitudinal and survival processes (SLS model)*

Description

MCMC function for model with separate longitudinal and survival processes (SLS model)

Usage

```
mcmc_separate(data_all, mcmc_settings, seed)
```

Arguments

data_all	A list of complete data. X0_inds is the baseline data for each patient, D is the dosage, Y is the longitudinal process, and Ts are the visit times, id are the patient id's (ranging from 1 to the total number of patients), censor and surv_time are the censoring and survival times, and Npat is the total number of patients.
mcmc_settings	A list for MCMC setup. burn.in is the number of total iterations to burn, ndisplay is number of iterations per which the display message will appear, and Niter is the total number of iterations (including burn-in iterations).
seed	Seed for running MCMC

Value

beta_l	Linear coefficient parameter in longitudinal submodel
sigma2_l	Error parameter in longitudinal submodel
beta_d	Linear coefficient parameter in dosing submodel
sigma2_d	Error parameter in dosing submodel
nu_1	Visitation intensity peak parameter in vistation submodel
nu_2	Visitation intensity shape parameter in vistation submodel
mu	Baseline visitation intensity parameter in vistation submodel
beta_s1	Creatinine-associated parameter in survival submodel
beta_s2	Dosage-associated parameter in survival submodel
beta_s3	Toxicity-associated parameter in survival submodel
beta_s4	Visitation-associated parameter in survival submodel
h0	Baseline hazard parameter in survival submodel
shape	Shape parameter in survival submodel

Examples

```
## Not run:
data("simulated_data")
############
#Run MCMC
mcmc_settings=NULL
mcmc_settings$Niter=100
mcmc_settings$burn.in=10
mcmc_settings$ndisplay=10
mcmc_settings$peak_dist='gamma'
thin=10
post_thin_iters<-seq(mcmc_settings$burn.in+thin,mcmc_settings$Niter,thin)
seed=203
set.seed(seed)
mcmc_out_separate<-mcmc_separate(simulated_data,mcmc_settings,seed)


## End(Not run)
```

SGD_run *Function for calculating optimal dosage and visit parameters for a specific patient using stochastic gradient descent (SGD)*

Description

Function for calculating optimal dosage and visit parameters for a specific patient using stochastic gradient descent (SGD)

Usage

```
SGD_run(mcmc, data_all, id_i, mcmc_inds, seed, Niter_SGD = 1000)
```

Arguments

mcmc	output from mcmc_joint or mcmc_separate
data_all	A list of complete data. X0_inds is the baseline data for each patient, D is the dosage, Y is the longitudinal process, and Ts are the visit times, id are the patient id's (ranging from 1 to the total number of patients). censor and death_time are the censoring and survival times. Npat is the total number of patients.
id_i	ID of the patient to optimize
mcmc_inds	Vector of post-thinning MCMC iterations indices
seed	Seed.
Niter_SGD	Number of SGD iterations.

Value

A list containing the dosage and visit parameters across all SGD iterations. beta_d is the dosage linear coefficient parameter vector, sigma2_d is the dosage error parameter, mu is the baseline visitation intensity, nu_1 is the visitation intensity peak, and nu_2 is the visitation intensity shape. The optimal value for beta_d is opt_beta_d and similarly for the other four parameters.

Examples

```
## Not run:
data("simulated_data")
############
#Run MCMC
mcmc_settings=NULL
mcmc_settings$Niter=100
mcmc_settings$burn.in=10
mcmc_settings$ndisplay=10
mcmc_settings$peak_dist='gamma'
thin=10
post_thin_iters<-seq(mcmc_settings$burn.in+thin,mcmc_settings$Niter,thin)
seed=203
set.seed(seed)
mcmc_out_joint<-mcmc_joint(simulated_data,mcmc_settings,seed)
############
#Run SGD
id_i=5
SGD_out<-SGD_run(mcmc_out_joint,simulated_data,id_i,post_thin_iters,seed,Niter_SGD=10)


## End(Not run)
```

---

simulated_data *Simulated dataset for 500 patients.*

Description

Simulated dataset for 500 patients.

Usage

```
simulated_data
```

Format

List of data. X0_inds is the baseline data for each patient, D is the dosage, Y is the longitudinal process, and Ts are the visit times, id are the patient id's (ranging from 1 to the total number of patients), censor and surv_time are the censoring and survival times, and Npat is the total number of patients.

Index