Package 'SGDinference'

September 21, 2022

Type Package

Title Inference with Stochastic sub-Gradient Descent

Version 0.0.0.9003

Date 2022-09-17

Author SGDinference-Lab

Maintainer Youngki Shin <yshin12@gmail.com>, Sokbae ``Simon" Lee <sokbae@gmail.com>

Description This package provides R code to apply the S-subGD method. The algorithm is developed for fast and robust inference with large-scale quantile regression.

License MIT + file LICENSE

Imports Rcpp (>= 1.0.5)

LinkingTo Rcpp, RcppArmadillo

RoxygenNote 7.2.1

Encoding UTF-8

Suggests knitr, rmarkdown

VignetteBuilder knitr

R topics documented:

sgdi *Averaged SGD and its Inference via Random Scaling*

Description

Compute the averaged SGD estimator and the variance-covariance matrix via random scaling method.

Usage

```
sgdi(
  x,
  y,
  gamma_0 = 1,
  alpha = 0.667,
  burn = 1,
  model = "lm",
  inference = "rs",
  bt_start = NULL,
  path_output = NULL,
  qt = 0.5,
  studentize = TRUE,
  intercept = TRUE,
  rss_idx = c(1)
)
```

Arguments

x	numeric. (n x p) matrix of regressors. Should not include 1 (the intercept)
y	numeric. (n x 1) vector of a dependent variable.
gamma_0	numeric. A tuning parameter for the learning rate (gamma_0 x t ^ alpha). Default is 1.
alpha	numeric. A tuning parameter for the learning rate (gamma_0 x t ^ alpha). Default is 0.667.
burn	numeric. A tuning parameter for "burn-in" observations. We burn-in up to (burn-1) observations and use observations from (burn) for estimation. Default is 1, i.e. no burn-in.
model	character specifying the model to be used: "lm" (linear mean regression model), "qr" (quantile regression)
inference	character. Specifying the inference method. Default is "rs" (random scaling). "rss" is for ransom scaling subset inference. Then, "rss_indx" should be provided.
bt_start	numeric. (p x 1) vector. User-provided starting value Default is NULL.
path_output	numeric specifying the sequence that print out the output paths
qt	numeric. Quantile. Default is 0.5.

studentize	logical. Studentize regressors. Default is TRUE
intercept	logical. Use the intercept term for regressors. Default is TRUE
rss_idx	numeric. Index of x for random scaling subset inference. Default is 1, the first regressor of x. For example, if we want to infer the 1st, 3rd covariate of x, then set it to be c(1,3).

Value

An object of class "sgdi", which is a list containing the following

beta_hat A (p + 1)-vector of estimated parameter values including the intercept.

V_hat A (p+1)x (p+1) variance-covariance matrix of beta_hat

V_hat_sub A variance-covariance sub-matrix of beta_hat. If the subset size is not provided, it returns 0.

Examples

```
n = 1e05
p = 5
bt0 = rep(5,p)
x = matrix(rnorm(n*(p-1)), n, (p-1))
y = cbind(1,x) %*% bt0 + rnorm(n)
sgdi.out = sgdi(x,y)
```

SGDinference *SGDinference*

Description

ADD DESCRIPTION LATER

Author(s)

SGDinference-Lab

sgdi_boot_qr *Inference of quantile regression with SGD and random scaling*

Description

Compute the averaged SGD estimator of quantile regression and the confidence intervals via random scaling method.

Usage

```
sgdi_boot_qr(
  x,
  y,
  gamma_0 = 1,
  alpha = 0.667,
  burn = 1,
  inference = "boot",
  bt_start = NULL,
  path_output = NULL,
  qt = 0.5,
  studentize = TRUE,
  intercept = TRUE,
  n_boot = 1000
)
```

Arguments

x	numeric. (n x p) matrix of regressors. Should not include 1 (the intercept)
y	numeric
gamma_0	numeric. Tuning parameter for the learning rate: gamma_0 * n^(alpha) (default is 1)
alpha	numeric. Tuning parameter for the learning rate: gamma_0 * n^(alpha) (default is 0.667)
burn	numeric. Number of burn-in observations (default is 1, no burn-in)
inference	character specifying the inference method. Default is "rs" (random scaling)
bt_start	numeric.
path_output	numeric specifying the sequence that print out the output paths
qt	numeric. Quantile. Default is 0.5.
studentize	logical. Studentize regressors. Default is TRUE
intercept	logical. Use the intercept term for regressors. Default is TRUE
n_boot	numeric. The number of boostrap samples. Default is 1000.

Value

#' An object of class "sgdi", which is a list containing the following components:

Examples

```
n = 1e05
p = 5
bt0 = rep(5,p)
x = matrix(rnorm(n*(p-1)), n, (p-1))
y = cbind(1,x) %*% bt0 + rnorm(n)
sgdi.out = sgdi_qr(x,y)
```

---

sgdi_lm *Averaged SGD and its Inference via Random Scaling*

Description

Compute the averaged SGD estimator and the confidence intervals via random scaling method.

Usage

```
sgdi_lm(
  x,
  y,
  gamma_0 = 1,
  alpha = 0.667,
  burn = 1,
  inference = "rs",
  bt_start = NULL,
  path_output = NULL,
  studentize = TRUE,
  intercept = TRUE
)
```

Arguments

x	numeric. (n x p) matrix of regressors. Should not include 1 (the intercept)
y	numeric
gamma_0	numeric
alpha	numeric
burn	numeric
inference	character specifying the inference method. Default is "rs" (random scaling)
bt_start	numeric
path_output	numeric specifying the sequence that print out the output paths
studentize	logical. Studentize regressors. Default is TRUE
intercept	logical. Use the intercept term for regressors. Default is TRUE

Value

#' An object of class "sgdi", which is a list containing the following components:

Examples

```
n = 1e05
p = 5
bt0 = rep(5,p)
x = matrix(rnorm(n*(p-1)), n, (p-1))
y = cbind(1,x) %*% bt0 + rnorm(n)
sgdi.out = sgdi(x,y)
```

---

sgdi_lm_cpp *Leading NA*

Description

Explanation

Usage

```
sgdi_lm_cpp(x, y, burn, gamma_0, alpha, bt_start, inference)
```

Arguments

x	numeric. (n x p) matrix of regressors. Should not include 1 (the intercept)
y	numeric
burn	numeric
gamma_0	numeric
alpha	numeric
bt_start	numeric
inference	character specifying the inference method. Default is "rs" (random scaling)

sgdi_qr *Inference of quantile regression with SGD and random scaling*

Description

Compute the averaged SGD estimator of quantile regression and the confidence intervals via random scaling method.

Usage

```
sgdi_qr(
  x,
  y,
  gamma_0 = 1,
  alpha = 0.667,
  burn = 1,
  inference = "rs",
  bt_start = NULL,
  path_output = NULL,
  qt = 0.5,
  studentize = TRUE,
  intercept = TRUE,
  rss_idx = c(1)
)
```

Arguments

x	numeric. (n x p) matrix of regressors. Should not include 1 (the intercept)
y	numeric
gamma_0	numeric. Tuning parameter for the learning rate: gamma_0 * n^(alpha) (default is 1)
alpha	numeric. Tuning parameter for the learning rate: gamma_0 * n^(alpha) (default is 0.667)
burn	numeric. Number of burn-in observations (default is 1, no burn-in)
inference	character specifying the inference method. Default is "rs" (random scaling).
bt_start	numeric.
path_output	numeric specifying the sequence that print out the output paths
qt	numeric. Quantile. Default is 0.5.
studentize	logical. Studentize regressors. Default is TRUE
intercept	logical. Use the intercept term for regressors. Default is TRUE
rss_idx	numeric. Index of x for random scaling subset inference. Default is 1, the first regressor of x.

Value

#' An object of class "sgdi", which is a list containing the following components:

Examples

```
n = 1e05
p = 5
bt0 = rep(5,p)
x = matrix(rnorm(n*(p-1)), n, (p-1))
y = cbind(1,x) %*% bt0 + rnorm(n)
sgdi.out = sgdi_qr(x,y)
```

---

sgd_qr *Averaged SsubGD estimator for quantile regression*

Description

Compute the averaged SGD estimator for quantile regression. This function computes only the point estimate.

Usage

```
sgd_qr(
  x,
  y,
  gamma_0 = 1,
  alpha = 0.667,
  burn = 1,
  bt_start = NULL,
  path_output = NULL,
  qt = 0.5,
  studentize = TRUE,
  intercept = TRUE
)
```

Arguments

x	numeric. (n x p) matrix of regressors. Should not include 1 (the intercept)
y	numeric. (n x 1) vector of a dependent variable.
gamma_0	numeric. A tuning parameter for the learning rate (gamma_0 x t ^ alpha). Default is 1.
alpha	numeric. A tuning parameter for the learning rate (gamma_0 x t ^ alpha). Default is 0.667.
burn	numeric. A tuning parameter for "burn-in" observations. We burn-in up to (burn-1) observations and use observations from (burn) for estimation. Default is 1, i.e. no burn-in.

bt_start	numeric. (p x 1) vector. User-provided starting value Default is NULL.
path_output	numeric specifying the sequence that print out the output paths
qt	numeric. Quantile. Default is 0.5.
studentize	logical. Studentize regressors. Default is TRUE
intercept	logical. Use the intercept term for regressors. Default is TRUE

Value

An object of class "sgd_qr", which is a list containing the following

beta_hat A (p + 1)-vector of estimated parameter values including the intercept.

beta_hat_path A (p+1) x P matrix. The path of estimated coeffcieitn values, where P is the length of the path

Examples

```
n = 1e05
p = 5
bt0 = rep(5,p)
x = matrix(rnorm(n*(p-1)), n, (p-1))
y = cbind(1,x) %*% bt0 + rnorm(n)
sgd.out = sgd_qr(x,y)
```

Index