Package ‘HDNRA’

May 7, 2026

Type Package

Title High-Dimensional Location Testing with Normal-Reference Approaches

Version 2.1.0

Date 2026-04-27

Author Pengfei Wang [aut, cre],
Shuqi Luo [aut],
Tianming Zhu [aut],
Bu Zhou [aut]

Maintainer Pengfei Wang `<nie23.wp8738@e.ntu.edu.sg>`

Description Provides inverse-free high-dimensional location tests for two-sample and general linear hypothesis testing (GLHT) problems under equal or unequal covariance structures. The package implements classical normal-approximation procedures, scale-invariant procedures, normal-reference procedures based on covariance-matched Gaussian companions, and F-type normal-reference calibrations for heteroscedastic Behrens-Fisher and GLHT settings. Implemented two-sample normal-approximation and scale-invariant procedures include Bai and Saranadasa (1996) <`https://www.jstor.org/stable/24306018`>, Chen and Qin (2010) <doi:10.1214/09-aos716>, Srivastava and Du (2008) <doi:10.1016/j.jmva.2006.11.002>, and Srivastava et al. (2013) <doi:10.1016/j.jmva.2012.08.014>. Implemented two-sample normal-reference procedures include Zhang, Guo, Zhou and Cheng (2020) <doi:10.1080/01621459.2019.1604366>, Zhang, Zhou, Guo and Zhu (2021) <doi:10.1016/j.jspi.2020.11.008>, Zhang, Zhu and Zhang (2020) <doi:10.1016/j.ecosta.2019.12.002>, Zhang, Zhu and Zhang (2023) <doi:10.1080/02664763.2020.1834516>, Zhang and Zhu (2022) <doi:10.1080/10485252.2021.2015768>, Zhang and Zhu (2022) <doi:10.1007/s42519-021-00232-w>, and Zhu, Wang and Zhang (2023) <doi:10.1007/s00180-023-01433-6>. Implemented GLHT normal-approximation procedures include Fujikoshi et al. (2004) <doi:10.14490/jjss.34.19>, Srivastava and Fujikoshi (2006) <doi:10.1016/j.jmva.2005.08.010>, Yamada and Srivastava (2012) <doi:10.1080/03610926.2011.581786>, Schott (2007) <doi:10.1016/j.jmva.2006.11.007>, and Zhou, Guo and Zhang (2017) <doi:10.1016/j.jspi.2017.03.005>. Implemented GLHT normal-reference procedures include Zhang, Guo and Zhou (2017) <doi:10.1016/j.jmva.2017.01.002>, Zhang, Zhou and Guo (2022) <doi:10.1016/j.jmva.2021.104816>, Zhu, Zhang and Zhang (2022) <doi:10.5705/ss.202020.0362>, Zhu and Zhang (2022) <doi:10.1007/s00180-021-01110-6>, Zhang and Zhu (2022) <doi:10.1016/j.csda.2021.107385>, and Cao et al. (2024) <doi:10.1007/s00362-024-01530-8>. The package also includes the random-integration normal-approximation GLHT procedure of Li et al. (2025) <doi:10.1007/s00362-024-01624-3>. A package-level overview is given in Wang, Zhu and Zhang (2026) <doi:10.1016/j.csda.2025.108269>.

License GPL (>= 3)

URL `https://github.com/nie23wp8738/HDNRA`, `https://nie23wp8738.github.io/HDNRA/`

BugReports `https://github.com/nie23wp8738/HDNRA/issues`

Encoding UTF-8

RoxygenNote 7.3.2

SystemRequirements OpenMP

LinkingTo Rcpp, RcppArmadillo

Imports expm, Rcpp, Rdpack, readr, stats, utils

Suggests devtools, dplyr, knitr, rmarkdown, spelling, testthat (>= 3.0.0), tidyr

RdMacros Rdpack

Depends R (>= 4.0.0)

LazyData true

Language en-US

Config/testthat/edition 3

NeedsCompilation yes

Repository CRAN

Date/Publication 2026-04-29 07:00:08 UTC

Contents

`BS1996.TS.NABT` *Normal-approximation-based test for two-sample problem proposed by Bai and Saranadasa (1996)*

Description

Bai and Saranadasa (1996)'s test for testing equality of two-sample high-dimensional mean vectors with assuming that two covariance matrices are the same.

Usage

```
BS1996.TS.NABT(y1, y2)
```

Arguments

`y1`	The data matrix ($n_1 \times p$) from the first population. Each row represents a p-dimensional observation.
`y2`	The data matrix ($n_2 \times p$) from the second population. Each row represents a p-dimensional observation.

Details

Suppose we have two independent high-dimensional samples:

$$\boldsymbol{y}_{i1}, \ldots, \boldsymbol{y}_{in_i}, \text{ are i.i.d. with } \mathrm{E}(\boldsymbol{y}_{i1}) = \boldsymbol{\mu}_i, \ \mathrm{Cov}(\boldsymbol{y}_{i1}) = \boldsymbol{\Sigma}, i = 1, 2.$$

The primary object is to test

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \text{ versus } H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2.$$

Bai and Saranadasa (1996) proposed the following centralised L^2-norm-based test statistic:

$$T_{BS} = \frac{n_1 n_2}{n} \|\bar{\boldsymbol{y}}_1 - \bar{\boldsymbol{y}}_2\|^2 - \mathrm{tr}(\hat{\boldsymbol{\Sigma}}),$$

where $\bar{\boldsymbol{y}}_i, i = 1, 2$ are the sample mean vectors and $\hat{\boldsymbol{\Sigma}}$ is the pooled sample covariance matrix. They showed that under the null hypothesis, T_{BS} is asymptotically normally distributed.

Value

A list of class `"NRtest"` containing the results of the hypothesis test. See the help file for `NRtest.object` for details.

References

Bai Z, Saranadasa H (1996). "Effect of high dimension: by an example of a two sample problem." *Statistica Sinica*, 311–329. `https://www.jstor.org/stable/24306018`.

Examples

```
library("HDNRA")
data("COVID19")
dim(COVID19)
group1 <- as.matrix(COVID19[c(2:19, 82:87), ]) ## healthy group
group2 <- as.matrix(COVID19[-c(1:19, 82:87), ]) ## COVID-19 patients
BS1996.TS.NABT(group1,group2)
```

`CCXH2024.GLHTBF.2cNRT` *Normal-reference-test with two-cumulant (2-c) matched χ^2-approximation for GLHT problem proposed by Cao et al. (2024)*

Description

Implements the scale-invariant test of Cao et al. (2024) for high-dimensional linear hypotheses of k-sample mean vectors under heteroscedastic covariance structures.

Usage

```
CCXH2024.GLHTBF.2cNRT(Y, B, n, p, alpha = 0.05)
```

Arguments

`Y`	A list of k data matrices. The ith element represents the data matrix $(n_i \times p)$ from the ith population with each row representing a p-dimensional observation.
`B`	A vector of k known scalars $(B_1, \dots, B_k)$ specifying the linear combination of mean vectors.
`n`	A vector of k sample sizes. The ith element represents the sample size of group i, n_i.
`p`	The dimension of data.
`alpha`	Significance level used to report the critical value (default 0.05). P-value does not depend on alpha.

Details

Suppose we have k independent high-dimensional samples

$$\boldsymbol{Y}_{i1}, \dots, \boldsymbol{Y}_{in_i} \text{ are i.i.d. with } \mathrm{E}(\boldsymbol{Y}_{i1}) = \boldsymbol{\mu}_i,\ \mathrm{Cov}(\boldsymbol{Y}_{i1}) = \boldsymbol{\Sigma}_i,\ i = 1, \dots, k,$$

where the covariance matrices $\boldsymbol{\Sigma}_i$ may differ across groups.

It is of interest to test the k-sample linear hypothesis

$$H_0 : \sum_{i=1}^{k} B_i \boldsymbol{\mu}_i = \mathbf{0} \quad \text{vs.} \quad H_1 : \sum_{i=1}^{k} B_i \boldsymbol{\mu}_i \neq \mathbf{0}.$$

Cao et al. (2024) proposed the following scale-invariant test statistic:

$$T = p^{-1} \boldsymbol{Y}^\top \boldsymbol{D}_\sigma^{-1} \boldsymbol{Y}, \quad \boldsymbol{Y} = \sqrt{n} \sum_{i=1}^{k} B_i \bar{\boldsymbol{Y}}_i, \quad n = \sum_{i=1}^{k} n_i,$$

where $\bar{\boldsymbol{Y}}_i$ is the sample mean vector of group i and $\boldsymbol{D}_\sigma$ is the diagonal matrix formed from a pooled covariance estimator. They showed that under the null hypothesis, T can be approximated by a Welch–Satterthwaite chi-square reference law $\chi^2_{\nu^*}/\nu^*$, where ν^* is an adjusted degrees-of-freedom parameter.

Value

A list of class `"NRtest"` containing the results of the hypothesis test.

References

Cao M, Cheng Z, Xu K, He D (2024). “A scale-invariant test for linear hypothesis of means in high dimensions.” *Statistical Papers*, **65**(6), 3477–3497.

Examples

```
library("HDNRA")
data("corneal")
```

```
# corneal: 150 x p, split into 4 groups (n_i x p)
group1 <- as.matrix(corneal[1:43,  ])      # normal
group2 <- as.matrix(corneal[44:57, ])      # unilateral suspect
group3 <- as.matrix(corneal[58:78, ])      # suspect map
group4 <- as.matrix(corneal[79:150,])      # clinical keratoconus

Y <- list(group1, group2, group3, group4)
n <- c(nrow(group1), nrow(group2), nrow(group3), nrow(group4))
p <- ncol(group1)

# Example linear combination (single contrast)
B <- c(-2, 1, 2, -1)

CCXH2024.GLHTBF.2cNRT(Y, B, n, p, alpha = 0.05)
```

`corneal` *HDNRA_data corneal*

Description

This dataset was acquired during a keratoconus study, a collaborative project involving Ms.Nancy Tripoli and Dr.Kenneth L.Cohen of Department of Ophthalmology at the University of North Carolina, Chapel Hill. The fitted feature vectors for the complete corneal surface dataset collectively into a feature matrix with dimensions of 150 × 2000.

Usage

```
data(corneal)
```

Format

'corneal':

A data frame with 150 observations on the following 4 groups.

normal group1 row 1 to row 43 in total 43 rows of the feature matrix correspond to observations from the normal group

unilateral suspect group2 row 44 to row 57 in total 14 rows of the feature matrix correspond to observations from the unilateral suspect group

suspect map group3 row 58 to row 78 in total 21 of the feature matrix correspond to observations from the suspect map group

clinical keratoconus group4 row 79 to row 150 in total 72 of the feature matrix correspond to observations from the clinical keratoconus group

References

Smaga Ł, Zhang J (2019). "Linear hypothesis testing with functional data." *Technometrics*, **61**(1), 99–110. doi:10.1080/00401706.2018.1456976.

Examples

```
library(HDNRA)
data(corneal)
dim(corneal)
group1 <- as.matrix(corneal[1:43, ]) ## normal group
dim(group1)
group2 <- as.matrix(corneal[44:57, ]) ## unilateral suspect group
dim(group2)
group3 <- as.matrix(corneal[58:78, ]) ## suspect map group
dim(group3)
group4 <- as.matrix(corneal[79:150, ]) ## clinical keratoconus group
dim(group4)
```

`COVID19` *HDNRA_data COVID19*

Description

A COVID19 data set from NCBI with ID GSE152641. The data set profiled peripheral blood from 24 healthy controls and 62 prospectively enrolled patients with community-acquired lower respiratory tract infection by SARS-COV-2 within the first 24 hours of hospital admission using RNA sequencing.

Usage

```
data(COVID19)
```

Format

'COVID19':

A data frame with 86 observations on the following 2 groups.

healthy group1 row 2 to row 19, and row 82 to 87, in total 24 healthy controls

patients group2 row 20 to 81, in total 62 prospectively enrolled patients

References

Thair SA, He YD, Hasin-Brumshtein Y, Sakaram S, Pandya R, Toh J, Rawling D, Remmel M, Coyle S, Dalekos GN, others (2021). “Transcriptomic similarities and differences in host response between SARS-CoV-2 and other viral infections.” *Iscience*, **24**(1). doi:10.1016/j.isci.2020.101947.

Examples

```
library(HDNRA)
data(COVID19)
dim(COVID19)
group1 <- as.matrix(COVID19[c(2:19, 82:87), ]) ## healthy group
dim(group1)
group2 <- as.matrix(COVID19[-c(1:19, 82:87), ]) ## COVID-19 patients
dim(group2)
```

`CQ2010.TSBF.NABT` *Normal-approximation-based test for two-sample BF problem proposed by Chen and Qin (2010)*

Description

Chen and Qin (2010)'s test for testing equality of two-sample high-dimensional mean vectors without assuming that two covariance matrices are the same.

Usage

```
CQ2010.TSBF.NABT(y1, y2)
```

Arguments

`y1`	The data matrix ($n_1 \times p$) from the first population. Each row represents a p-dimensional observation.
`y2`	The data matrix ($n_2 \times p$) from the second population. Each row represents a p-dimensional observation.

Details

Suppose we have two independent high-dimensional samples:

$$\boldsymbol{y}_{i1},\ldots,\boldsymbol{y}_{in_i}, \text{ are i.i.d. with } \mathrm{E}(\boldsymbol{y}_{i1})=\boldsymbol{\mu}_i,\ \mathrm{Cov}(\boldsymbol{y}_{i1})=\boldsymbol{\Sigma}_i, i=1,2.$$

The primary object is to test

$$H_0:\boldsymbol{\mu}_1=\boldsymbol{\mu}_2 \text{ versus } H_1:\boldsymbol{\mu}_1\neq\boldsymbol{\mu}_2.$$

Chen and Qin (2010) proposed the following test statistic:

$$T_{CQ}=\frac{\sum_{i\neq j}^{n_1}\boldsymbol{y}_{1i}^\top\boldsymbol{y}_{1j}}{n_1(n_1-1)}+\frac{\sum_{i\neq j}^{n_2}\boldsymbol{y}_{2i}^\top\boldsymbol{y}_{2j}}{n_2(n_2-1)}-2\frac{\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\boldsymbol{y}_{1i}^\top\boldsymbol{y}_{2j}}{n_1n_2}.$$

They showed that under the null hypothesis, T_{CQ} is asymptotically normally distributed.

Value

A list of class `"NRtest"` containing the results of the hypothesis test. See the help file for `NRtest.object` for details.

References

Chen SX, Qin Y (2010). "A two-sample test for high-dimensional data with applications to gene-set testing." *The Annals of Statistics*, **38**(2). doi:10.1214/09aos716.

Examples

```
library("HDNRA")
data("COVID19")
dim(COVID19)
group1 <- as.matrix(COVID19[c(2:19, 82:87), ]) ## healthy group
group2 <- as.matrix(COVID19[-c(1:19, 82:87), ]) ## COVID-19 patients
CQ2010.TSBF.NABT(group1,group2)
```

`FHW2004.GLHT.NABT` *Normal-approximation-based test for GLHT problem proposed by Fujikoshi et al. (2004)*

Description

Fujikoshi et al. (2004)'s test for general linear hypothesis testing (GLHT) problem for high-dimensional data with assuming that underlying covariance matrices are the same.

Usage

```
FHW2004.GLHT.NABT(Y,X,C,n,p)
```

Arguments

Y	A list of k data matrices. The ith element represents the data matrix ($n_i \times p$) from the ith population with each row representing a p-dimensional observation.
X	A known $n \times k$ full-rank design matrix with $\mathrm{rank}(\boldsymbol{X}) = k < n$.
C	A known matrix of size $q \times k$ with $\mathrm{rank}(\boldsymbol{C}) = q < k$.
n	A vector of k sample sizes. The ith element represents the sample size of group i, n_i.
p	The dimension of data.

Details

A high-dimensional linear regression model can be expressed as

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\Theta} + \boldsymbol{\epsilon},$$

where Θ is a $k \times p$ unknown parameter matrix and $\boldsymbol{\epsilon}$ is an $n \times p$ error matrix.

It is of interest to test the following GLHT problem

$$H_0 : \boldsymbol{C}\boldsymbol{\Theta} = \boldsymbol{0}, \quad \text{vs.} \quad H_1 : \boldsymbol{C}\boldsymbol{\Theta} \neq \boldsymbol{0}.$$

Fujikoshi et al. (2004) proposed the following test statistic:

$$T_{FHW} = \sqrt{p}\left[(n-k)\frac{\mathrm{tr}(\boldsymbol{S}_h)}{\mathrm{tr}(\boldsymbol{S}_e)} - q\right],$$

where $\boldsymbol{S}_h$ and $\boldsymbol{S}_e$ are the matrices of sums of squares and products due to the hypothesis and the error, respecitively.

They showed that under the null hypothesis, T_{FHW} is asymptotically normally distributed.

Value

A list of class `"NRtest"` containing the results of the hypothesis test. See the help file for `NRtest.object` for details.

References

Fujikoshi Y, Himeno T, Wakaki H (2004). “Asymptotic results of a high dimensional MANOVA test and power comparison when the dimension is large compared to the sample size.” *Journal of the Japan Statistical Society*, **34**(1), 19–26. doi:10.14490/jjss.34.19.

Examples

```
library("HDNRA")
data("corneal")
dim(corneal)
group1 <- as.matrix(corneal[1:43, ]) ## normal group
group2 <- as.matrix(corneal[44:57, ]) ## unilateral suspect group
group3 <- as.matrix(corneal[58:78, ]) ## suspect map group
group4 <- as.matrix(corneal[79:150, ]) ## clinical keratoconus group
p <- dim(corneal)[2]
Y <- list()
k <- 4
Y[[1]] <- group1
Y[[2]] <- group2
Y[[3]] <- group3
Y[[4]] <- group4
n <- c(nrow(Y[[1]]),nrow(Y[[2]]),nrow(Y[[3]]),nrow(Y[[4]]))
X <- matrix(c(rep(1,n[1]),rep(0,sum(n)),rep(1,n[2]), rep(0,sum(n)),
            rep(1,n[3]),rep(0,sum(n)),rep(1,n[4])),ncol=k,nrow=sum(n))
q <- k-1
C <- cbind(diag(q),-rep(1,q))
FHW2004.GLHT.NABT(Y,X,C,n,p)
```

`LHNB2025.GLHTBF.NABT` *Normal-approximation-based test for k-sample linear hypothesis via random integration proposed by Li et al. (2025)*

Description

Li et al. (2025)’s test for general linear hypothesis testing (GLHT) problem for high-dimensional data under heteroscedasticity.

Usage

```
LHNB2025.GLHTBF.NABT(Y, B, O, A, n, p)
```

Arguments

`Y`	A list of k data matrices. The ith element represents the data matrix $(n_i \times p)$ from the ith population with each row representing a p-dimensional observation.
`B`	A vector of k coefficients $(B_1, \ldots, B_k)$ specifying the linear combination of group mean vectors.
`O`	A length-p vector used to form $\Omega = \mathrm{diag}(O_1^2, \ldots, O_p^2)$.
`A`	A length-p vector used in $W = \Omega + AA^\top$.
`n`	A vector of k sample sizes. The ith element represents the sample size of group i, n_i.
`p`	The dimension of data.

Details

Suppose we have k independent high-dimensional samples

$$\boldsymbol{Y}_{i1}, \ldots, \boldsymbol{Y}_{in_i} \text{ are i.i.d. with } \mathrm{E}(\boldsymbol{Y}_{i1}) = \boldsymbol{\mu}_i,\ \mathrm{Cov}(\boldsymbol{Y}_{i1}) = \boldsymbol{\Sigma}_i,\ i = 1, \ldots, k,$$

where the covariance matrices $\boldsymbol{\Sigma}_i$ may differ across groups.

It is of interest to test the k-sample linear hypothesis

$$H_0: \sum_{i=1}^{k} B_i \boldsymbol{\mu}_i = \mathbf{0} \quad \text{vs.} \quad H_1: \sum_{i=1}^{k} B_i \boldsymbol{\mu}_i \neq \mathbf{0}.$$

Li et al. (2025) proposed a random-integration-based U-statistic T_n (Eq. (5) in the paper), constructed using the weight matrix $\boldsymbol{W} = \boldsymbol{\Omega} + \boldsymbol{A}\boldsymbol{A}^\top$ with $\boldsymbol{\Omega} = \mathrm{diag}(O_1^2, \ldots, O_p^2)$. They showed that the standardized statistic $Z = T_n/\sqrt{\hat{\sigma}^2}$ is approximated by $N(0, 1)$ under H_0.

A recommended default choice of tuning parameters is of the form $A_1 = \cdots = A_p = \sqrt{5}\,p^{-3/8}$ and $O_k = \sqrt{\epsilon\left(1 + \frac{2k}{3p}\right)}$, $k = 1, \ldots, p$.

Value

A list of class `"NRtest"` containing the results of the hypothesis test.

References

Li J, Hong S, Niu Z, Bai Z (2025). “Test for high-dimensional linear hypothesis of mean vectors via random integration.” *Statistical Papers*, **66**(1), 8.

Examples

```
library("HDNRA")
data("corneal")

# corneal: 150 x p, split into 4 groups (n_i x p)
group1 <- as.matrix(corneal[1:43,  ])      # normal
group2 <- as.matrix(corneal[44:57, ])      # unilateral suspect
```

```
group3 <- as.matrix(corneal[58:78, ])      # suspect map
group4 <- as.matrix(corneal[79:150,])      # clinical keratoconus

Y <- list(group2, group3, group4)
n <- c(nrow(group2), nrow(group3), nrow(group4))
p <- ncol(group2)

# One linear combination (example): B = (4, -1.5, -2.5)
B <- c(4, -1.5, -2.5)

# Paper-style tuning parameters (example with eps = 2)
A <- rep(sqrt(5) * p^(-3/8), p)
O <- sqrt(2) * (1 + 2*(1:p)/(3*p))

LHNB2025.GLHTBF.NABT(Y, B, O, A, n, p)
```

`NRtest.object` *S3 Class "NRtest"*

Description

The "`NRtest`" objects provide a comprehensive summary of hypothesis test outcomes, including test statistics, p-values, parameter estimates, and confidence intervals, if applicable.

Usage

```
NRtest.object(
  statistic,
  p.value,
  method,
  null.value,
  alternative,
  parameter = NULL,
  sample.size = NULL,
  sample.dimension = NULL,
  estimation.method = NULL,
  data.name = NULL,
  ...
)
```

Arguments

`statistic`	Numeric scalar containing the value of the test statistic, with a `names` attribute indicating the name of the test statistic.
`p.value`	Numeric scalar containing the p-value for the test.

`method`	Character string giving the name of the test.
`null.value`	Character string indicating the null hypothesis.
`alternative`	Character string indicating the alternative hypothesis.
`parameter`	Numeric vector containing the estimated approximation parameter(s) associated with the approximation method. This vector has a `names` attribute describing its element(s).
`sample.size`	Numeric vector containing the number of observations in each group used for the hypothesis test.
`sample.dimension`	Numeric scalar containing the dimension of the dataset used for the hypothesis test.
`estimation.method`	Character string giving the name of the approximation approach used to approximate the null distribution of the test statistic.
`data.name`	Character string describing the data set used in the hypothesis test.
`...`	Additional optional arguments.

Details

A class of objects returned by high-dimensional hypothesis testing functions in the **HDNRA** package, designed to encapsulate detailed results from statistical hypothesis tests. These objects are structured similarly to **htest** objects in the package **EnvStats** but are tailored to the needs of the **HDNRA** package.

Value

An object of class `"NRtest"` containing both required and optional components depending on the specifics of the hypothesis test, shown as follows:

Required Components

These components must be present in every `"NRtest"` object:

`statistic` Must e present.

`p.value` Must e present.

`null.value` Must e present.

`alternative` Must e present.

`method` Must e present.

Optional Components

These components are included depending on the specifics of the hypothesis test performed:

`parameter` May be present.

`sample.size` May be present.

`sample.dimension` May be present.

`estimation.method` May be present.

`data.name` May be present.

Methods

The class has the following methods:

`print.NRtest` Printing the contents of the NRtest object in a human-readable form.

Examples

```
# Example 1: Using Bai and Saranadasa (1996)'s test (two-sample problem)
NRtest.obj1 <- NRtest.object(
  statistic = c(”T[BS]” = 2.208),
  p.value = 0.0136,
  method = ”Bai and Saranadasa (1996)'s test”,
  data.name = ”group1 and group2”,
  null.value = c(”Difference between two mean vectors is o”),
  alternative = ”Difference between two mean vectors is not 0”,
  parameter = NULL,
  sample.size = c(n1 = 24, n2 = 26),
  sample.dimension = 20460,
  estimation.method = ”Normal approximation”
)
print(NRtest.obj1)

# Example 2: Using Fujikoshi et al. (2004)'s test (GLHT problem)
NRtest.obj2 <- NRtest.object(
  statistic = c(”T[FHW]” = 6.4015),
  p.value = 0,
  method = ”Fujikoshi et al. (2004)'s test”,
  data.name = ”Y”,
  null.value  = ”The general linear hypothesis is true”,
  alternative = ”The general linear hypothesis is not true”,
  parameter = NULL,
  sample.size = c(n1 = 43, n2 = 14, n3 = 21, n4 = 72),
  sample.dimension = 2000,
  estimation.method = ”Normal approximation”
)
print(NRtest.obj2)
```

`print.NRtest` *Print Method for S3 Class "NRtest"*

Description

Prints the details of the **NRtest** object in a user-friendly manner. This method provides a clear and concise presentation of the test results contained within the **NRtest** object, including all relevant statistical metrics and test details.

Usage

```
## S3 method for class \pkg{NRtest}
## S3 method for class 'NRtest'
print(x, ...)
```

Arguments

`x`	an **NRtest** object.
`...`	further arguments passed to or from other methods.

Details

The `print.NRtest` function formats and presents the contents of the **NRtest** object, which includes statistical test results and related parameters. This function is designed to provide a user-friendly display of the object's contents, making it easier to understand the results of the analysis.

Value

Invisibly returns the input `x`.

Author(s)

Pengfei Wang <`nie23.wp8738@e.ntu.edu.sg`>

See Also

`NRtest.object`

`S2007.ks.NABT` *Normal-approximation-based test for one-way MANOVA problem proposed by Schott (2007)*

Description

Schott, J. R. (2007)'s test for one-way MANOVA problem for high-dimensional data with assuming that underlying covariance matrices are the same.

Usage

```
S2007.ks.NABT(Y, n, p)
```

Arguments

`Y`	A list of k data matrices. The ith element represents the data matrix $(n_i \times p)$ from the ith population with each row representing a p-dimensional observation.
`n`	A vector of k sample sizes. The ith element represents the sample size of group i, n_i.
`p`	The dimension of data.

Details

Suppose we have the following k independent high-dimensional samples:

$$\boldsymbol{y}_{i1},\ldots,\boldsymbol{y}_{in_i}, \text{ are i.i.d. with } \mathrm{E}(\boldsymbol{y}_{i1})=\boldsymbol{\mu}_i,\ \mathrm{Cov}(\boldsymbol{y}_{i1})=\boldsymbol{\Sigma}, i=1,\ldots,k.$$

It is of interest to test the following one-way MANOVA problem:

$$H_0:\boldsymbol{\mu}_1=\cdots=\boldsymbol{\mu}_k,\quad \text{vs.}\ H_1: H_0 \text{ is not ture}.$$

Schott (2007) proposed the following test statistic:

$$T_S=[\mathrm{tr}(\boldsymbol{H})/h-\mathrm{tr}(\boldsymbol{E})/e]/\sqrt{N-1},$$

where $\boldsymbol{H}=\sum_{i=1}^k n_i(\bar{\boldsymbol{y}}_i-\bar{\boldsymbol{y}})(\bar{\boldsymbol{y}}_i-\bar{\boldsymbol{y}})^\top$, $\boldsymbol{E}=\sum_{i=1}^k\sum_{j=1}^{n_i}(\boldsymbol{y}_{ij}-\bar{\boldsymbol{y}}_i)(\boldsymbol{y}_{ij}-\bar{\boldsymbol{y}}_i)^\top$, $h=k-1$, and $e=N-k$, with $N=n_1+\cdots+n_k$. They showed that under the null hypothesis, T_S is asymptotically normally distributed.

Value

A list of class `"NRtest"` containing the results of the hypothesis test. See the help file for `NRtest.object` for details.

References

Schott JR (2007). “Some high-dimensional tests for a one-way MANOVA.” *Journal of Multivariate Analysis*, **98**(9), 1825–1839. doi:10.1016/j.jmva.2006.11.007.

Examples

```
library("HDNRA")
data("corneal")
dim(corneal)
group1 <- as.matrix(corneal[1:43, ]) ## normal group
group2 <- as.matrix(corneal[44:57, ]) ## unilateral suspect group
group3 <- as.matrix(corneal[58:78, ]) ## suspect map group
group4 <- as.matrix(corneal[79:150, ]) ## clinical keratoconus group
p <- dim(corneal)[2]
Y <- list()
Y[[1]] <- group1
Y[[2]] <- group2
Y[[3]] <- group3
Y[[4]] <- group4
n <- c(nrow(Y[[1]]),nrow(Y[[2]]),nrow(Y[[3]]),nrow(Y[[4]]))
S2007.ks.NABT(Y, n, p)
```

`SD2008.TS.NABT` *Normal-approximation-based test for two-sample problem proposed by Srivastava and Du (2008)*

Description

Srivastava and Du (2008)'s test for testing equality of two-sample high-dimensional mean vectors with assuming that two covariance matrices are the same.

Usage

```
SD2008.TS.NABT(y1, y2)
```

Arguments

`y1`	The data matrix ($n_1 \times p$) from the first population. Each row represents a p-dimensional observation.
`y2`	The data matrix ($n_2 \times p$) from the second population. Each row represents a p-dimensional observation.

Details

Suppose we have two independent high-dimensional samples:

$$\boldsymbol{y}_{i1}, \ldots, \boldsymbol{y}_{in_i}, \text{ are i.i.d. with } \mathrm{E}(\boldsymbol{y}_{i1}) = \boldsymbol{\mu}_i, \ \mathrm{Cov}(\boldsymbol{y}_{i1}) = \boldsymbol{\Sigma}, i = 1, 2.$$

The primary object is to test

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \text{ versus } H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2.$$

Srivastava and Du (2008) proposed the following test statistic:

$$T_{SD} = \frac{n^{-1} n_1 n_2 (\bar{\boldsymbol{y}}_1 - \bar{\boldsymbol{y}}_2)^\top \boldsymbol{D}_S^{-1} (\bar{\boldsymbol{y}}_1 - \bar{\boldsymbol{y}}_2) - \frac{(n-2)p}{n-4}}{\sqrt{2 \left[\mathrm{tr}(\boldsymbol{R}^2) - \frac{p^2}{n-2} \right] c_{p,n}}},$$

where $\bar{\boldsymbol{y}}_i, i = 1, 2$ are the sample mean vectors, $\boldsymbol{D}_S$ is the diagonal matrix of sample variance, $\boldsymbol{R}$ is the sample correlation matrix and $c_{p,n}$ is the adjustment coefficient proposed by Srivastava and Du (2008). They showed that under the null hypothesis, T_{SD} is asymptotically normally distributed.

Value

A list of class `"NRtest"` containing the results of the hypothesis test. See the help file for `NRtest.object` for details.

References

Srivastava MS, Du M (2008). "A test for the mean vector with fewer observations than the dimension." *Journal of Multivariate Analysis*, **99**(3), 386–402. doi:10.1016/j.jmva.2006.11.002.

Examples

```
library("HDNRA")
data("COVID19")
dim(COVID19)
group1 <- as.matrix(COVID19[c(2:19, 82:87), ]) ## healthy group
group2 <- as.matrix(COVID19[-c(1:19, 82:87), ]) ## COVID-19 patients
SD2008.TS.NABT(group1,group2)
```

`SF2006.GLHT.NABT` *Normal-approximation-based test for GLHT problem proposed by Srivastava and Fujikoshi (2006)*

Description

Srivastava and Fujikoshi (2006)'s test for general linear hypothesis testing (GLHT) problem for high-dimensional data with assuming that underlying covariance matrices are the same.

Usage

```
SF2006.GLHT.NABT(Y,X,C,n,p)
```

Arguments

`Y`	A list of k data matrices. The ith element represents the data matrix $(n_i \times p)$ from the ith population with each row representing a p-dimensional observation.
`X`	A known $n \times k$ full-rank design matrix with $\mathrm{rank}(\boldsymbol{X}) = k < n$.
`C`	A known matrix of size $q \times k$ with $\mathrm{rank}(\boldsymbol{C}) = q < k$.
`n`	A vector of k sample sizes. The ith element represents the sample size of group i, n_i.
`p`	The dimension of data.

Details

A high-dimensional linear regression model can be expressed as

$$\boldsymbol{Y} = \boldsymbol{X\Theta} + \boldsymbol{\epsilon},$$

where Θ is a $k \times p$ unknown parameter matrix and $\boldsymbol{\epsilon}$ is an $n \times p$ error matrix.

It is of interest to test the following GLHT problem

$$H_0 : \boldsymbol{C\Theta} = \boldsymbol{0}, \quad \text{vs.} \quad H_1 : \boldsymbol{C\Theta} \neq \boldsymbol{0}.$$

Srivastava and Fujikoshi (2006) proposed the following test statistic:

$$T_{SF} = \left[2q\hat{a}_2(1 + (n-k)^{-1}q)\right]^{-1/2} \left[\frac{\mathrm{tr}(\boldsymbol{B})}{\sqrt{p}} - \frac{q}{\sqrt{n-k}} \frac{\mathrm{tr}(\boldsymbol{W})}{\sqrt{(n-k)p}}\right].$$

where $\boldsymbol{W}$ and $\boldsymbol{B}$ are the matrix of sum of squares and products due to error and the error, respectively, and $\hat{a}_2 = [\mathrm{tr}(\boldsymbol{W}^2) - \mathrm{tr}^2(\boldsymbol{W})/(n-k)]/[(n-k-1)(n-k+2)p]$. They showed that under the null hypothesis, T_{SF} is asymptotically normally distributed.

Value

A list of class `"NRtest"` containing the results of the hypothesis test. See the help file for `NRtest.object` for details.

References

Srivastava MS, Fujikoshi Y (2006). "Multivariate analysis of variance with fewer observations than the dimension." *Journal of Multivariate Analysis*, **97**(9), 1927–1940. doi:10.1016/j.jmva.2005.08.010.

Examples

```
library("HDNRA")
data("corneal")
dim(corneal)
group1 <- as.matrix(corneal[1:43, ]) ## normal group
group2 <- as.matrix(corneal[44:57, ]) ## unilateral suspect group
group3 <- as.matrix(corneal[58:78, ]) ## suspect map group
group4 <- as.matrix(corneal[79:150, ]) ## clinical keratoconus group
p <- dim(corneal)[2]
Y <- list()
k <- 4
Y[[1]] <- group1
Y[[2]] <- group2
Y[[3]] <- group3
Y[[4]] <- group4
n <- c(nrow(Y[[1]]),nrow(Y[[2]]),nrow(Y[[3]]),nrow(Y[[4]]))
X <- matrix(c(rep(1,n[1]),rep(0,sum(n)),rep(1,n[2]), rep(0,sum(n)),
            rep(1,n[3]),rep(0,sum(n)),rep(1,n[4])),ncol=k,nrow=sum(n))
q <- k-1
C <- cbind(diag(q),-rep(1,q))
SF2006.GLHT.NABT(Y,X,C,n,p)
```

`SKK2013.TSBF.NABT` *Normal-approximation-based test for two-sample BF problem proposed by Srivastava et al. (2013)*

Description

Srivastava et al. (2013)'s test for testing equality of two-sample high-dimensional mean vectors without assuming that two covariance matrices are the same.

Usage

```
SKK2013.TSBF.NABT(y1, y2)
```

Arguments

`y1`	The data matrix ($n_1 \times p$) from the first population. Each row represents a p-dimensional observation.
`y2`	The data matrix ($n_2 \times p$ from the second population. Each row represents a p-dimensional observation.

Details

Suppose we have two independent high-dimensional samples:

$$\boldsymbol{y}_{i1}, \ldots, \boldsymbol{y}_{in_i}, \text{ are i. i. d. with } \mathrm{E}(\boldsymbol{y}_{i1}) = \boldsymbol{\mu}_i, \ \mathrm{Cov}(\boldsymbol{y}_{i1}) = \boldsymbol{\Sigma}_i, i = 1, 2.$$

The primary object is to test

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \text{ versus } H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2.$$

Srivastava et al. (2013) proposed the following test statistic:

$$T_{SKK} = \frac{(\bar{\boldsymbol{y}}_1 - \bar{\boldsymbol{y}}_2)^\top \hat{\boldsymbol{D}}^{-1} (\bar{\boldsymbol{y}}_1 - \bar{\boldsymbol{y}}_2) - p}{\sqrt{2\widehat{\mathrm{Var}}(\hat{q}_n) c_{p,n}}},$$

where $\bar{\boldsymbol{y}}_i, i = 1, 2$ are the sample mean vectors, $\hat{\boldsymbol{D}} = \hat{\boldsymbol{D}}_1/n_1 + \hat{\boldsymbol{D}}_2/n_2$ with $\hat{\boldsymbol{D}}_i, i = 1, 2$ being the diagonal matrices consisting of only the diagonal elements of the sample covariance matrices. $\widehat{\mathrm{Var}}(\hat{q}_n)$ is given by equation (1.18) in Srivastava et al. (2013), and $c_{p,n}$ is the adjustment coefficient proposed by Srivastava et al. (2013). They showed that under the null hypothesis, T_{SKK} is asymptotically normally distributed.

Value

A list of class `"NRtest"` containing the results of the hypothesis test. See the help file for `NRtest.object` for details.

References

Srivastava MS, Katayama S, Kano Y (2013). “A two sample test in high dimensional data.” *Journal of Multivariate Analysis*, **114**, 349–358. doi:10.1016/j.jmva.2012.08.014.

Examples

```
library("HDNRA")
data("COVID19")
dim(COVID19)
group1 <- as.matrix(COVID19[c(2:19, 82:87), ]) ## healthy group
group2 <- as.matrix(COVID19[-c(1:19, 82:87), ]) ## COVID-19 patients
SKK2013.TSBF.NABT(group1,group2)
```

`WZ2026.GLHTBF.2cNRT` *F-approximation-based F-type test for GLHT problem under heteroscedasticity*

Description

An F-type normal reference test for the high-dimensional general linear hypothesis testing (GLHT) problem under heteroscedasticity. The null distribution is approximated by an F distribution using Welch–Satterthwaite (W–S) chi-square approximations.

Usage

```
WZ2026.GLHTBF.2cNRT(Y, G, n, p)
```

Arguments

`Y`	A list of k data matrices. The ith element represents the data matrix ($n_i \times p$) from the ith population with each row representing a p-dimensional observation.
`G`	A known full-rank coefficient matrix ($q \times k$) with $\mathrm{rank}(\boldsymbol{G}) < k$.
`n`	A vector of k sample sizes. The ith element represents the sample size of group i, n_i.
`p`	The dimension of data.

Details

The test statistic is of F-type form

$$F_{n,p} = \frac{\|\boldsymbol{C}\hat{\boldsymbol{\mu}}\|^2}{\mathrm{tr}(\widehat{\Omega}_n)}.$$

The degrees of freedom are estimated by matching the first two cumulants via W–S approximation, resulting in an $F_{\hat{d}_1,\hat{d}_2}$ reference distribution.

Value

A list of class `"NRtest"` containing the results of the hypothesis test.

References

Wang, P. and Zhu, T. (preprint). An F-type Test for Heteroscedastic General Linear Hypothesis Testing Problem for High Dimensional Data: a Normal Reference Approach.

Examples

```
library("HDNRA")
data("corneal")
dim(corneal)
group1 <- as.matrix(corneal[1:43, ]) ## normal group
group2 <- as.matrix(corneal[44:57, ]) ## unilateral suspect group
group3 <- as.matrix(corneal[58:78, ]) ## suspect map group
group4 <- as.matrix(corneal[79:150, ]) ## clinical keratoconus group
p <- dim(corneal)[2]
Y <- list()
k <- 4
Y[[1]] <- group1
Y[[2]] <- group2
Y[[3]] <- group3
Y[[4]] <- group4
n <- c(nrow(Y[[1]]),nrow(Y[[2]]),nrow(Y[[3]]),nrow(Y[[4]]))
G <- cbind(diag(k-1),rep(-1,k-1))
WZ2026.GLHTBF.2cNRT(Y,G,n,p)
```

`YS2012.GLHT.NABT` *Normal-approximation-based test for GLHT problem proposed by Yamada and Srivastava (2012)*

Description

Yamada and Srivastava (2012)'test for general linear hypothesis testing (GLHT) problem for high-dimensional data with assuming that underlying covariance matrices are the same.

Usage

```
YS2012.GLHT.NABT(Y,X,C,n,p)
```

Arguments

`Y`	A list of k data matrices. The ith element represents the data matrix $(n_i \times p)$ from the ith population with each row representing a p-dimensional observation.
`X`	A known $n \times k$ full-rank design matrix with $\mathrm{rank}(\boldsymbol{X}) = k < n$.
`C`	A known matrix of size $q \times k$ with $\mathrm{rank}(\boldsymbol{C}) = q < k$.
`n`	A vector of k sample sizes. The ith element represents the sample size of group i, n_i.
`p`	The dimension of data.

Details

A high-dimensional linear regression model can be expressed as

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\Theta} + \boldsymbol{\epsilon},$$

where Θ is a $k \times p$ unknown parameter matrix and $\boldsymbol{\epsilon}$ is an $n \times p$ error matrix.

It is of interest to test the following GLHT problem

$$H_0 : \boldsymbol{C\Theta} = \boldsymbol{0}, \quad \text{vs. } H_1 : \boldsymbol{C\Theta} \neq \boldsymbol{0}.$$

Yamada and Srivastava (2012) proposed the following test statistic:

$$T_{YS} = \frac{(n-k)\,\mathrm{tr}(\boldsymbol{S}_h \boldsymbol{D}_{\boldsymbol{S}_e}^{-1}) - (n-k)pq/(n-k-2)}{\sqrt{2q[\mathrm{tr}(\boldsymbol{R}^2) - p^2/(n-k)]c_{p,n}}},$$

where $\boldsymbol{S}_h$ and $\boldsymbol{S}_e$ are the variation matrices due to the hypothesis and error, respectively, and $\boldsymbol{D}_{\boldsymbol{S}_e}$ and $\boldsymbol{R}$ are diagonal matrix with the diagonal elements of $\boldsymbol{S}_e$ and the sample correlation matrix, respectively. $c_{p,n}$ is the adjustment coefficient proposed by Yamada and Srivastava (2012). They showed that under the null hypothesis, T_{YS} is asymptotically normally distributed.

Value

A list of class `"NRtest"` containing the results of the hypothesis test. See the help file for `NRtest.object` for details.

References

Yamada T, Srivastava MS (2012). “A test for multivariate analysis of variance in high dimension.” *Communications in Statistics-Theory and Methods*, **41**(13-14), 2602–2615. doi:10.1080/03610926.2011.581786.

Examples

```
library("HDNRA")
data("corneal")
dim(corneal)
group1 <- as.matrix(corneal[1:43, ]) ## normal group
group2 <- as.matrix(corneal[44:57, ]) ## unilateral suspect group
group3 <- as.matrix(corneal[58:78, ]) ## suspect map group
group4 <- as.matrix(corneal[79:150, ]) ## clinical keratoconus group
p <- dim(corneal)[2]
Y <- list()
k <- 4
Y[[1]] <- group1
Y[[2]] <- group2
Y[[3]] <- group3
Y[[4]] <- group4
n <- c(nrow(Y[[1]]),nrow(Y[[2]]),nrow(Y[[3]]),nrow(Y[[4]]))
X <- matrix(c(rep(1,n[1]),rep(0,sum(n)),rep(1,n[2]), rep(0,sum(n)),rep(1,n[3]),
            rep(0,sum(n)),rep(1,n[4])),ncol=k,nrow=sum(n))
```

```
q <- k-1
C <- cbind(diag(q),-rep(1,q))
YS2012.GLHT.NABT(Y,X,C,n,p)
```

`ZGZ2017.GLHT.2cNRT` *Normal-reference-test with two-cumulant (2-c) matched χ^2-approximation for GLHT problem proposed Zhang et al. (2017)*

Description

Zhang et al. (2017)'s test for general linear hypothesis testing (GLHT) problem for high-dimensional data with assuming that underlying covariance matrices are the same.

Usage

```
ZGZ2017.GLHT.2cNRT(Y,G,n,p)
```

Arguments

`Y`	A list of k data matrices. The ith element represents the data matrix $(n_i \times p)$ from the ith population with each row representing a p-dimensional observation.
`G`	A known full-rank coefficient matrix $(q \times k)$ with $\mathrm{rank}(\boldsymbol{G}) < k$.
`n`	A vector of k sample sizes. The ith element represents the sample size of group i, n_i.
`p`	The dimension of data.

Details

Suppose we have the following k independent high-dimensional samples:

$$\boldsymbol{y}_{i1},\ldots,\boldsymbol{y}_{in_i}, \text{ are i.i.d. with } \mathrm{E}(\boldsymbol{y}_{i1})=\boldsymbol{\mu}_i,\ \mathrm{Cov}(\boldsymbol{y}_{i1})=\boldsymbol{\Sigma},\ i=1,\ldots,k.$$

It is of interest to test the following GLHT problem:

$$H_0: \boldsymbol{G}\boldsymbol{M}=\boldsymbol{0}, \quad \text{vs.} \quad H_1: \boldsymbol{G}\boldsymbol{M}\neq\boldsymbol{0},$$

where $\boldsymbol{M}=(\boldsymbol{\mu}_1,\ldots,\boldsymbol{\mu}_k)^\top$ is a $k\times p$ matrix collecting k mean vectors and $\boldsymbol{G}: q\times k$ is a known full-rank coefficient matrix with $\mathrm{rank}(\boldsymbol{G})<k$.

Zhang et al. (2017) proposed the following test statistic:

$$T_{ZGZ}=\|\boldsymbol{C}\hat{\boldsymbol{\mu}}\|^2,$$

where $\boldsymbol{C}=[(\boldsymbol{G}\boldsymbol{D}\boldsymbol{G}^\top)^{-1/2}\boldsymbol{G}]\otimes\boldsymbol{I}_p$, and $\hat{\boldsymbol{\mu}}=(\bar{\boldsymbol{y}}_1^\top,\ldots,\bar{\boldsymbol{y}}_k^\top)^\top$, with $\bar{\boldsymbol{y}}_i, i=1,\ldots,k$ being the sample mean vectors and $\boldsymbol{D}=\mathrm{diag}(1/n_1,\ldots,1/n_k)$.

They showed that under the null hypothesis, T_{ZGZ} and a chi-squared-type mixture have the same normal or non-normal limiting distribution.

Value

A list of class `"NRtest"` containing the results of the hypothesis test. See the help file for `NRtest.object` for details.

References

Zhang J, Guo J, Zhou B (2017). "Linear hypothesis testing in high-dimensional one-way MANOVA." *Journal of Multivariate Analysis*, **155**, 200–216. doi:10.1016/j.jmva.2017.01.002.

Examples

```
library("HDNRA")
data("corneal")
dim(corneal)
group1 <- as.matrix(corneal[1:43, ]) ## normal group
group2 <- as.matrix(corneal[44:57, ]) ## unilateral suspect group
group3 <- as.matrix(corneal[58:78, ]) ## suspect map group
group4 <- as.matrix(corneal[79:150, ]) ## clinical keratoconus group
p <- dim(corneal)[2]
Y <- list()
k <- 4
Y[[1]] <- group1
Y[[2]] <- group2
Y[[3]] <- group3
Y[[4]] <- group4
n <- c(nrow(Y[[1]]),nrow(Y[[2]]),nrow(Y[[3]]),nrow(Y[[4]]))
G <- cbind(diag(k-1),rep(-1,k-1))
ZGZ2017.GLHT.2cNRT(Y,G,n,p)
```

`ZGZ2017.GLHTBF.NABT` *Normal-approximation-based test for GLHT problem under heteroscedasticity proposed by Zhou et al. (2017)*

Description

Zhou et al. (2017)'s test for general linear hypothesis testing (GLHT) problem for high-dimensional data under heteroscedasticity.

Usage

```
ZGZ2017.GLHTBF.NABT(Y,G,n,p)
```

Arguments

`Y`	A list of k data matrices. The ith element represents the data matrix $(n_i \times p)$ from the ith population with each row representing a p-dimensional observation.
`G`	A known full-rank coefficient matrix $(q \times k)$ with $\mathrm{rank}(\boldsymbol{G}) < k$.

n	A vector of k sample sizes. The ith element represents the sample size of group i, n_i.
p	The dimension of data.

Details

Suppose we have the following k independent high-dimensional samples:

$$\boldsymbol{y}_{i1}, \ldots, \boldsymbol{y}_{in_i}, \text{ are i. i. d. with } \mathrm{E}(\boldsymbol{y}_{i1}) = \boldsymbol{\mu}_i, \ \mathrm{Cov}(\boldsymbol{y}_{i1}) = \boldsymbol{\Sigma}_i, i = 1, \ldots, k.$$

It is of interest to test the following GLHT problem:

$$H_0 : \boldsymbol{GM} = \boldsymbol{0}, \quad \text{vs. } H_1 : \boldsymbol{GM} \neq \boldsymbol{0},$$

where $\boldsymbol{M} = (\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_k)^\top$ is a $k \times p$ matrix collecting k mean vectors and $\boldsymbol{G} : q \times k$ is a known full-rank coefficient matrix with $\mathrm{rank}(\boldsymbol{G}) < k$.

Let $\bar{\boldsymbol{y}}_i, i = 1, \ldots, k$ be the sample mean vectors and $\hat{\boldsymbol{\Sigma}}_i, i = 1, \ldots, k$ be the sample covariance matrices.

Zhou et al. (2017) proposed the following U-statistic based test statistic:

$$T_{ZGZ} = \|\boldsymbol{C}\hat{\boldsymbol{\mu}}\|^2 - \sum_{i=1}^{k} h_{ii} \operatorname{tr}(\hat{\boldsymbol{\Sigma}}_i)/n_i,$$

where $\boldsymbol{C} = [(\boldsymbol{GDG}^\top)^{-1/2}\boldsymbol{G}] \otimes \boldsymbol{I}_p$, $\boldsymbol{D} = \mathrm{diag}(1/n_1, \ldots, 1/n_k)$, and h_{ij} is the (i,j)th entry of the $k \times k$ matrix $\boldsymbol{H} = \boldsymbol{G}^\top(\boldsymbol{GDG}^\top)^{-1}\boldsymbol{G}$.

They showed that under the null hypothesis, T_{ZGZ} is asymptotically normally distributed.

Value

A list of class `"NRtest"` containing the results of the hypothesis test. See the help file for `NRtest.object` for details.

References

Zhou B, Guo J, Zhang J (2017). “High-dimensional general linear hypothesis testing under heteroscedasticity.” *Journal of Statistical Planning and Inference*, **188**, 36–54. doi:10.1016/j.jspi.2017.03.005.

Examples

```
library("HDNRA")
data("corneal")
dim(corneal)
group1 <- as.matrix(corneal[1:43, ]) ## normal group
group2 <- as.matrix(corneal[44:57, ]) ## unilateral suspect group
group3 <- as.matrix(corneal[58:78, ]) ## suspect map group
group4 <- as.matrix(corneal[79:150, ]) ## clinical keratoconus group
p <- dim(corneal)[2]
Y <- list()
k <- 4
```

```
Y[[1]] <- group1
Y[[2]] <- group2
Y[[3]] <- group3
Y[[4]] <- group4
n <- c(nrow(Y[[1]]),nrow(Y[[2]]),nrow(Y[[3]]),nrow(Y[[4]]))
G <- cbind(diag(k-1),rep(-1,k-1))
ZGZ2017.GLHTBF.NABT(Y,G,n,p)
```

`ZGZC2020.TS.2cNRT` *Normal-reference-test with two-cumulant (2-c) matched χ^2-approximation for two-sample problem proposed by Zhang et al. (2020)*

Description

Zhang et al. (2020)'s test for testing equality of two-sample high-dimensional mean vectors with assuming that two covariance matrices are the same.

Usage

```
ZGZC2020.TS.2cNRT(y1, y2)
```

Arguments

`y1`	The data matrix ($n_1 \times p$) from the first population. Each row represents a p-dimensional observation.
`y2`	The data matrix ($n_2 \times p$) from the second population. Each row represents a p-dimensional observation.

Details

Suppose we have two independent high-dimensional samples:

$$\boldsymbol{y}_{i1},\ldots,\boldsymbol{y}_{in_i}, \text{ are i.i.d. with } \mathrm{E}(\boldsymbol{y}_{i1})=\boldsymbol{\mu}_i,\ \mathrm{Cov}(\boldsymbol{y}_{i1})=\boldsymbol{\Sigma}, i=1,2.$$

The primary object is to test

$$H_0:\boldsymbol{\mu}_1=\boldsymbol{\mu}_2 \text{ versus } H_1:\boldsymbol{\mu}_1\neq\boldsymbol{\mu}_2.$$

Zhang et al.(2020) proposed the following test statistic:

$$T_{ZGZC}=\frac{n_1n_2}{n}\|\bar{\boldsymbol{y}}_1-\bar{\boldsymbol{y}}_2\|^2,$$

where $\bar{\boldsymbol{y}}_i, i=1,2$ are the sample mean vectors. They showed that under the null hypothesis, T_{ZGZC} and a chi-squared-type mixture have the same normal or non-normal limiting distribution.

Value

A list of class `"NRtest"` containing the results of the hypothesis test. See the help file for `NRtest.object` for details.

References

Zhang J, Guo J, Zhou B, Cheng M (2020). "A simple two-sample test in high dimensions based on L 2-norm." *Journal of the American Statistical Association*, **115**(530), 1011–1027. doi:10.1080/01621459.2019.1604366.

Examples

```
library("HDNRA")
data("COVID19")
dim(COVID19)
group1 <- as.matrix(COVID19[c(2:19, 82:87), ]) ## healthy group
group2 <- as.matrix(COVID19[-c(1:19, 82:87), ]) ## COVID-19 patients
ZGZC2020.TS.2cNRT(group1, group2)
```

`ZWZ2023.TSBF.2cNRT` *Normal-reference-test with two-cumulant (2-c) matched χ^2-approximation for two-sample BF problem proposed by Zhu et al. (2023)*

Description

Zhu et al. (2023)'s test for testing equality of two-sample high-dimensional mean vectors without assuming that two covariance matrices are the same.

Usage

```
ZWZ2023.TSBF.2cNRT(y1, y2)
```

Arguments

`y1`	The data matrix ($n_1 \times p$) from the first population. Each row represents a p-dimensional observation.
`y2`	The data matrix ($n_2 \times p$) from the second population. Each row represents a p-dimensional observation.

Details

Suppose we have two independent high-dimensional samples:

$$\boldsymbol{y}_{i1}, \ldots, \boldsymbol{y}_{in_i}, \text{ are i.i.d. with } \mathrm{E}(\boldsymbol{y}_{i1}) = \boldsymbol{\mu}_i, \ \mathrm{Cov}(\boldsymbol{y}_{i1}) = \boldsymbol{\Sigma}_i, \ i = 1, 2.$$

The primary object is to test

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \text{ versus } H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2.$$

Zhu et al. (2023) proposed the following test statistic:

$$T_{ZWZ} = \frac{n_1 n_2 n^{-1} \|\bar{\boldsymbol{y}}_1 - \bar{\boldsymbol{y}}_2\|^2}{\operatorname{tr}(\hat{\boldsymbol{\Omega}}_n)},$$

where $\bar{\boldsymbol{y}}_i, i = 1, 2$ are the sample mean vectors and $\hat{\boldsymbol{\Omega}}_n$ is the estimator of $\operatorname{Cov}[(n_1 n_2/n)^{1/2}(\bar{\boldsymbol{y}}_1 - \bar{\boldsymbol{y}}_2)]$. They showed that under the null hypothesis, T_{ZWZ} and an F-type mixture have the same normal or non-normal limiting distribution.

Value

A list of class `"NRtest"` containing the results of the hypothesis test. See the help file for `NRtest.object` for details.

References

Zhu T, Wang P, Zhang J (2023). “Two-sample Behrens–Fisher problems for high-dimensional data: a normal reference F-type test.” *Computational Statistics*, 1–24. doi:10.1007/s00180023014336.

Examples

```
library("HDNRA")
data("COVID19")
dim(COVID19)
group1 <- as.matrix(COVID19[c(2:19, 82:87), ]) ## healthy group
group2 <- as.matrix(COVID19[-c(1:19, 82:87), ]) ## COVID-19 patients
ZWZ2023.TSBF.2cNRT(group1, group2)
```

`ZZ2022.GLHT.3cNRT` *Normal-reference-test with three-cumulant (3-c) matched χ^2-approximation for GLHT problem proposed by Zhu and Zhang (2022)*

Description

Zhu and Zhang (2022)’s test for general linear hypothesis testing (GLHT) problem for high-dimensional data with assuming that underlying covariance matrices are the same.

Usage

```
ZZ2022.GLHT.3cNRT(Y,G,n,p)
```

Arguments

`Y`	A list of k data matrices. The ith element represents the data matrix $(n_i \times p)$ from the ith population with each row representing a p-dimensional observation.
`G`	A known full-rank coefficient matrix $(q \times k)$ with $\operatorname{rank}(\boldsymbol{G}) < k$.
`n`	A vector of k sample sizes. The ith element represents the sample size of group i, n_i.
`p`	The dimension of data.

Details

Suppose we have the following k independent high-dimensional samples:

$$\boldsymbol{y}_{i1}, \ldots, \boldsymbol{y}_{in_i}, \text{ are i. i. d. with } \mathrm{E}(\boldsymbol{y}_{i1}) = \boldsymbol{\mu}_i,\ \mathrm{Cov}(\boldsymbol{y}_{i1}) = \boldsymbol{\Sigma},\ i = 1, \ldots, k.$$

It is of interest to test the following GLHT problem:

$$H_0 : \boldsymbol{GM} = \boldsymbol{0}, \quad \text{vs.} \quad H_1 : \boldsymbol{GM} \neq \boldsymbol{0},$$

where $\boldsymbol{M} = (\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_k)^\top$ is a $k \times p$ matrix collecting k mean vectors and $\boldsymbol{G} : q \times k$ is a known full-rank coefficient matrix with $\operatorname{rank}(\boldsymbol{G}) < k$.

Zhu and Zhang (2022) proposed the following test statistic:

$$T_{ZZ} = \|\boldsymbol{C}\hat{\boldsymbol{\mu}}\|^2 - q \operatorname{tr}(\hat{\boldsymbol{\Sigma}}),$$

where $\boldsymbol{C} = [(\boldsymbol{GDG}^\top)^{-1/2}\boldsymbol{G}] \otimes \boldsymbol{I}_p$, and $\hat{\boldsymbol{\mu}} = (\bar{\boldsymbol{y}}_1^\top, \ldots, \bar{\boldsymbol{y}}_k^\top)^\top$, with $\bar{\boldsymbol{y}}_i, i = 1, \ldots, k$ being the sample mean vectors and $\hat{\boldsymbol{\Sigma}}$ being the usual pooled sample covariance matrix of the k samples.

They showed that under the null hypothesis, T_{ZZ} and a chi-squared-type mixture have the same normal or non-normal limiting distribution.

Value

A list of class `"NRtest"` containing the results of the hypothesis test. See the help file for `NRtest.object` for details.

References

Zhu T, Zhang J (2022). “Linear hypothesis testing in high-dimensional one-way MANOVA: a new normal reference approach.” *Computational Statistics*, **37**(1), 1–27. doi:10.1007/s0018002101110-6.

Examples

```
library("HDNRA")
data("corneal")
dim(corneal)
group1 <- as.matrix(corneal[1:43, ]) ## normal group
group2 <- as.matrix(corneal[44:57, ]) ## unilateral suspect group
group3 <- as.matrix(corneal[58:78, ]) ## suspect map group
group4 <- as.matrix(corneal[79:150, ]) ## clinical keratoconus group
```

```
p <- dim(corneal)[2]
Y <- list()
k <- 4
Y[[1]] <- group1
Y[[2]] <- group2
Y[[3]] <- group3
Y[[4]] <- group4
n <- c(nrow(Y[[1]]),nrow(Y[[2]]),nrow(Y[[3]]),nrow(Y[[4]]))
G <- cbind(diag(k-1),rep(-1,k-1))
ZZ2022.GLHT.3cNRT(Y,G,n,p)
```

`ZZ2022.GLHTBF.3cNRT` *Normal-reference-test with three-cumulant (3-c) matched χ^2-approximation for GLHT problem under heteroscedasticity proposed by Zhang and Zhu (2022)*

Description

Zhang and Zhu (2022)'s test for general linear hypothesis testing (GLHT) problem for high-dimensional data under heteroscedasticity.

Usage

```
ZZ2022.GLHTBF.3cNRT(Y,G,n,p)
```

Arguments

`Y`	A list of k data matrices. The ith element represents the data matrix ($n_i \times p$) from the ith population with each row representing a p-dimensional observation.
`G`	A known full-rank coefficient matrix ($q \times k$) with $\operatorname{rank}(\boldsymbol{G}) < k$.
`n`	A vector of k sample sizes. The ith element represents the sample size of group i, n_i.
`p`	The dimension of data.

Details

Suppose we have the following k independent high-dimensional samples:

$$\boldsymbol{y}_{i1},\ldots,\boldsymbol{y}_{in_i}, \text{ are i.i.d. with } \mathrm{E}(\boldsymbol{y}_{i1}) = \boldsymbol{\mu}_i,\ \mathrm{Cov}(\boldsymbol{y}_{i1}) = \boldsymbol{\Sigma}_i, i = 1,\ldots,k.$$

It is of interest to test the following GLHT problem:

$$H_0 : \boldsymbol{GM} = \boldsymbol{0}, \quad \text{vs. } H_1 : \boldsymbol{GM} \neq \boldsymbol{0},$$

where $\boldsymbol{M} = (\boldsymbol{\mu}_1,\ldots,\boldsymbol{\mu}_k)^\top$ is a $k \times p$ matrix collecting k mean vectors and $\boldsymbol{G} : q \times k$ is a known full-rank coefficient matrix with $\operatorname{rank}(\boldsymbol{G}) < k$.

Let $\bar{\boldsymbol{y}}_i, i = 1, \dots, k$ be the sample mean vectors and $\hat{\boldsymbol{\Sigma}}_i, i = 1, \dots, k$ be the sample covariance matrices.

Zhang and Zhu (2022) proposed the following U-statistic based test statistic:

$$T_{ZZ} = \|\boldsymbol{C}\hat{\boldsymbol{\mu}}\|^2 - \sum_{i=1}^{k} h_{ii} \operatorname{tr}(\hat{\boldsymbol{\Sigma}}_i)/n_i,$$

where $\boldsymbol{C} = [(\boldsymbol{G}\boldsymbol{D}\boldsymbol{G}^\top)^{-1/2}\boldsymbol{G}] \otimes \boldsymbol{I}_p$, $\boldsymbol{D} = \operatorname{diag}(1/n_1, \dots, 1/n_k)$, and h_{ij} is the (i, j)th entry of the $k \times k$ matrix $\boldsymbol{H} = \boldsymbol{G}^\top(\boldsymbol{G}\boldsymbol{D}\boldsymbol{G}^\top)^{-1}\boldsymbol{G}$.

Value

A list of class `"NRtest"` containing the results of the hypothesis test. See the help file for `NRtest.object` for details.

References

Zhang J, Zhu T (2022). “A new normal reference test for linear hypothesis testing in high-dimensional heteroscedastic one-way MANOVA.” *Computational Statistics & Data Analysis*, **168**, 107385. doi:10.1016/j.csda.2021.107385.

Examples

```
library("HDNRA")
data("corneal")
dim(corneal)
group1 <- as.matrix(corneal[1:43, ]) ## normal group
group2 <- as.matrix(corneal[44:57, ]) ## unilateral suspect group
group3 <- as.matrix(corneal[58:78, ]) ## suspect map group
group4 <- as.matrix(corneal[79:150, ]) ## clinical keratoconus group
p <- dim(corneal)[2]
Y <- list()
k <- 4
Y[[1]] <- group1
Y[[2]] <- group2
Y[[3]] <- group3
Y[[4]] <- group4
n <- c(nrow(Y[[1]]),nrow(Y[[2]]),nrow(Y[[3]]),nrow(Y[[4]]))
G <- cbind(diag(k-1),rep(-1,k-1))
ZZ2022.GLHTBF.3cNRT(Y,G,n,p)
```

`ZZ2022.TS.3cNRT` *Normal-reference-test with three-cumulant (3-c) matched χ^2-approximation for two-sample problem proposed by Zhang and Zhu (2022)*

Description

Zhang and Zhu (2022)'s test for testing equality of two-sample high-dimensional mean vectors with assuming that two covariance matrices are the same.

Usage

```
ZZ2022.TS.3cNRT(y1, y2)
```

Arguments

`y1`	The data matrix ($n_1 \times p$) from the first population. Each row represents a p-dimensional observation.
`y2`	The data matrix ($n_2 \times p$) from the second population. Each row represents a p-dimensional observation.

Details

Suppose we have two independent high-dimensional samples:

$$\boldsymbol{y}_{i1}, \ldots, \boldsymbol{y}_{in_i}, \text{ are i.i.d. with } \mathrm{E}(\boldsymbol{y}_{i1}) = \boldsymbol{\mu}_i, \ \mathrm{Cov}(\boldsymbol{y}_{i1}) = \boldsymbol{\Sigma}, i = 1, 2.$$

The primary object is to test

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \text{ versus } H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2.$$

Zhang et al.(2022) proposed the following test statistic:

$$T_{ZZ} = \frac{n_1 n_2}{n} \|\bar{\boldsymbol{y}}_1 - \bar{\boldsymbol{y}}_2\|^2 - \mathrm{tr}(\hat{\boldsymbol{\Sigma}}),$$

where $\bar{\boldsymbol{y}}_i, i = 1, 2$ are the sample mean vectors and $\hat{\boldsymbol{\Sigma}}$ is the pooled sample covariance matrix. They showed that under the null hypothesis, T_{ZZ} and a chi-squared-type mixture have the same normal or non-normal limiting distribution.

Value

A list of class `"NRtest"` containing the results of the hypothesis test. See the help file for `NRtest.object` for details.

References

Zhang J, Zhu T (2022). "A revisit to Bai–Saranadasa's two-sample test." *Journal of Nonparametric Statistics*, **34**(1), 58–76. doi:10.1080/10485252.2021.2015768.

Examples

```
library("HDNRA")
data("COVID19")
dim(COVID19)
group1 <- as.matrix(COVID19[c(2:19, 82:87), ]) ## healthy group
group2 <- as.matrix(COVID19[-c(1:19, 82:87), ]) ## COVID-19 patients
ZZ2022.TS.3cNRT(group1, group2)
```

`ZZ2022.TSBF.3cNRT` *Normal-reference-test with three-cumulant (3-c) matched χ^2-approximation for two-sample BF problem proposed by Zhang and Zhu (2022)*

Description

Zhang and Zhu (2022)'s test for testing equality of two-sample high-dimensional mean vectors without assuming that two covariance matrices are the same.

Usage

```
ZZ2022.TSBF.3cNRT(y1, y2)
```

Arguments

`y1`	The data matrix ($n_1 \times p$) from the first population. Each row represents a p-dimensional observation.
`y2`	The data matrix ($n_2 \times p$) from the second population. Each row represents a p-dimensional observation.

Details

Suppose we have two independent high-dimensional samples:

$$\boldsymbol{y}_{i1}, \ldots, \boldsymbol{y}_{in_i}, \text{ are i.i.d. with } \mathrm{E}(\boldsymbol{y}_{i1}) = \boldsymbol{\mu}_i, \ \mathrm{Cov}(\boldsymbol{y}_{i1}) = \boldsymbol{\Sigma}_i, i = 1, 2.$$

The primary object is to test

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \text{ versus } H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2.$$

Zhang and Zhu (2022) proposed the following test statistic:

$$T_{ZZ} = \|\bar{\boldsymbol{y}}_1 - \bar{\boldsymbol{y}}_2\|^2 - \mathrm{tr}(\hat{\boldsymbol{\Omega}}_n),$$

where $\bar{\boldsymbol{y}}_i, i = 1, 2$ are the sample mean vectors and $\hat{\boldsymbol{\Omega}}_n$ is the estimator of $\mathrm{Cov}(\bar{\boldsymbol{y}}_1 - \bar{\boldsymbol{y}}_2)$. They showed that under the null hypothesis, T_{ZZ} and a chi-squared-type mixture have the same normal or non-normal limiting distribution.

Value

A list of class `"NRtest"` containing the results of the hypothesis test. See the help file for `NRtest.object` for details.

References

Zhang J, Zhu T (2022). "A further study on Chen-Qin's test for two-sample Behrens–Fisher problems for high-dimensional data." *Journal of Statistical Theory and Practice*, **16**(1), 1. doi:10.1007/s4251902100232w.

Examples

```
library("HDNRA")
data("COVID19")
dim(COVID19)
group1 <- as.matrix(COVID19[c(2:19, 82:87), ]) ## healthy group
group2 <- as.matrix(COVID19[-c(1:19, 82:87), ]) ## COVID-19 patients
ZZ2022.TSBF.3cNRT(group1, group2)
```

`ZZG2022.GLHTBF.2cNRT` *Normal-reference-test with two-cumulant (2-c) matched χ^2-approximation for GLHT problem under heteroscedasticity proposed by Zhang et al. (2022)*

Description

Zhang et al. (2022)'s test for general linear hypothesis testing (GLHT) problem for high-dimensional data under heteroscedasticity.

Usage

```
ZZG2022.GLHTBF.2cNRT(Y,G,n,p)
```

Arguments

`Y`	A list of k data matrices. The ith element represents the data matrix ($n_i \times p$) from the ith population with each row representing a p-dimensional observation.
`G`	A known full-rank coefficient matrix ($q \times k$) with $\mathrm{rank}(\boldsymbol{G}) < k$.
`n`	A vector of k sample sizes. The ith element represents the sample size of group i, n_i.
`p`	The dimension of data.

Details

Suppose we have the following k independent high-dimensional samples:

$$\boldsymbol{y}_{i1},\ldots,\boldsymbol{y}_{in_i}, \text{ are i. i. d. with } \mathrm{E}(\boldsymbol{y}_{i1}) = \boldsymbol{\mu}_i,\ \mathrm{Cov}(\boldsymbol{y}_{i1}) = \boldsymbol{\Sigma}_i, i = 1,\ldots,k.$$

It is of interest to test the following GLHT problem:

$$H_0 : \boldsymbol{GM} = \boldsymbol{0}, \quad \text{vs. } H_1 : \boldsymbol{GM} \neq \boldsymbol{0},$$

where $\boldsymbol{M} = (\boldsymbol{\mu}_1,\ldots,\boldsymbol{\mu}_k)^\top$ is a $k \times p$ matrix collecting k mean vectors and $\boldsymbol{G} : q \times k$ is a known full-rank coefficient matrix with $\mathrm{rank}(\boldsymbol{G}) < k$.

Zhang et al. (2022) proposed the following test statistic:

$$T_{ZZG} = \|\boldsymbol{C}\hat{\boldsymbol{\mu}}\|^2,$$

where $\boldsymbol{C} = [(\boldsymbol{G}\boldsymbol{D}\boldsymbol{G}^\top)^{-1/2}\boldsymbol{G}] \otimes \boldsymbol{I}_p$ with $\boldsymbol{D} = \text{diag}(1/n_1, \ldots, 1/n_k)$, and $\hat{\boldsymbol{\mu}} = (\bar{\boldsymbol{y}}_1^\top, \ldots, \bar{\boldsymbol{y}}_k^\top)^\top$ with $\bar{\boldsymbol{y}}_i, i = 1, \ldots, k$ being the sample mean vectors.

They showed that under the null hypothesis, T_{ZZG} and a chi-squared-type mixture have the same normal or non-normal limiting distribution.

Value

A list of class `"NRtest"` containing the results of the hypothesis test. See the help file for `NRtest.object` for details.

References

Zhang J, Zhou B, Guo J (2022). "Linear hypothesis testing in high-dimensional heteroscedastic one-way MANOVA: A normal reference L^2-norm based test." *Journal of Multivariate Analysis*, **187**, 104816. doi:10.1016/j.jmva.2021.104816.

Examples

```
library("HDNRA")
data("corneal")
dim(corneal)
group1 <- as.matrix(corneal[1:43, ]) ## normal group
group2 <- as.matrix(corneal[44:57, ]) ## unilateral suspect group
group3 <- as.matrix(corneal[58:78, ]) ## suspect map group
group4 <- as.matrix(corneal[79:150, ]) ## clinical keratoconus group
p <- dim(corneal)[2]
Y <- list()
k <- 4
Y[[1]] <- group1
Y[[2]] <- group2
Y[[3]] <- group3
Y[[4]] <- group4
n <- c(nrow(Y[[1]]),nrow(Y[[2]]),nrow(Y[[3]]),nrow(Y[[4]]))
G <- cbind(diag(k-1),rep(-1,k-1))
ZZG2022.GLHTBF.2cNRT(Y,G,n,p)
```

`ZZGZ2021.TSBF.2cNRT` *Normal-reference-test with two-cumulant (2-c) matched χ^2-approximation for two-sample BF problem proposed by Zhang et al. (2021)*

Description

Zhang et al. (2021)'s test for testing equality of two-sample high-dimensional mean vectors without assuming that two covariance matrices are the same.

Usage

```
ZZGZ2021.TSBF.2cNRT(y1, y2)
```

Arguments

`y1`	The data matrix ($n_1 \times p$) from the first population. Each row represents a p-dimensional observation.
`y2`	The data matrix ($n_2 \times p$) from the second population. Each row represents a p-dimensional observation.

Details

Suppose we have two independent high-dimensional samples:

$$\boldsymbol{y}_{i1}, \ldots, \boldsymbol{y}_{in_i}, \text{ are i. i. d. with } \mathrm{E}(\boldsymbol{y}_{i1}) = \boldsymbol{\mu}_i, \ \mathrm{Cov}(\boldsymbol{y}_{i1}) = \boldsymbol{\Sigma}_i, i = 1, 2.$$

The primary object is to test

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \text{ versus } H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2.$$

Zhang et al.(2021) proposed the following test statistic:

$$T_{ZZGZ} = \frac{n_1 n_2}{n} \|\bar{\boldsymbol{y}}_1 - \bar{\boldsymbol{y}}_2\|^2,$$

where $\bar{\boldsymbol{y}}_i, i = 1, 2$ are the sample mean vectors. They showed that under the null hypothesis, T_{ZZGZ} and a chi-squared-type mixture have the same normal or non-normal limiting distribution.

Value

A list of class `"NRtest"` containing the results of the hypothesis test. See the help file for `NRtest.object` for details.

References

Zhang J, Zhou B, Guo J, Zhu T (2021). "Two-sample Behrens-Fisher problems for high-dimensional data: A normal reference approach." *Journal of Statistical Planning and Inference*, **213**, 142–161. doi:10.1016/j.jspi.2020.11.008.

Examples

```
library("HDNRA")
data("COVID19")
dim(COVID19)
group1 <- as.matrix(COVID19[c(2:19, 82:87), ]) ## healthy group
group2 <- as.matrix(COVID19[-c(1:19, 82:87), ]) ## COVID-19 patients
ZZGZ2021.TSBF.2cNRT(group1, group2)
```

`ZZZ2020.TS.2cNRT` *Normal-reference-test with two-cumulant (2-c) matched χ^2-approximation for two-sample problem proposed by Zhang et al. (2020)*

Description

Zhang et al. (2020)'s test for testing equality of two-sample high-dimensional mean vectors with assuming that two covariance matrices are the same.

Usage

```
ZZZ2020.TS.2cNRT(y1, y2)
```

Arguments

`y1`	The data matrix ($n_1 \times p$) from the first population. Each row represents a p-dimensional observation.
`y2`	The data matrix ($n_2 \times p$) from the second population. Each row represents a p-dimensional observation.

Details

Suppose we have two independent high-dimensional samples:

$$\boldsymbol{y}_{i1}, \ldots, \boldsymbol{y}_{in_i}, \text{ are i.i.d. with } \mathrm{E}(\boldsymbol{y}_{i1}) = \boldsymbol{\mu}_i,\ \mathrm{Cov}(\boldsymbol{y}_{i1}) = \boldsymbol{\Sigma}, i = 1, 2.$$

The primary object is to test

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \text{ versus } H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2.$$

Zhang et al.(2020) proposed the following test statistic:

$$T_{ZZZ} = \frac{n_1 n_2}{np}(\bar{\boldsymbol{y}}_1 - \bar{\boldsymbol{y}}_2)^\top \hat{\boldsymbol{D}}^{-1}(\bar{\boldsymbol{y}}_1 - \bar{\boldsymbol{y}}_2),$$

where $\bar{\boldsymbol{y}}_i, i = 1, 2$ are the sample mean vectors, $\hat{\boldsymbol{D}}$ is the diagonal matrix of sample covariance matrix. They showed that under the null hypothesis, T_{ZZZ} and a chi-squared-type mixture have the same limiting distribution.

Value

A list of class `"NRtest"` containing the results of the hypothesis test. See the help file for `NRtest.object` for details.

References

Zhang L, Zhu T, Zhang J (2020). "A simple scale-invariant two-sample test for high-dimensional data." *Econometrics and Statistics*, **14**, 131–144. doi:10.1016/j.ecosta.2019.12.002.

Examples

```
library("HDNRA")
data("COVID19")
dim(COVID19)
group1 <- as.matrix(COVID19[c(2:19, 82:87), ]) ## healthy group
group2 <- as.matrix(COVID19[-c(1:19, 82:87), ]) ## COVID-19 patients
ZZZ2020.TS.2cNRT(group1,group2)
```

`ZZZ2022.GLHT.2cNRT` *Normal-reference-test with two-cumulant (2-c) matched χ^2-approximation for GLHT problem proposed by Zhu et al. (2022)*

Description

Zhu et al. (2022)'s test for general linear hypothesis testing (GLHT) problem for high-dimensional data with assuming that underlying covariance matrices are the same.

Usage

```
ZZZ2022.GLHT.2cNRT(Y,X,C,n,p)
```

Arguments

`Y`	A list of k data matrices. The ith element represents the data matrix $(n_i \times p)$ from the ith population with each row representing a p-dimensional observation.
`X`	A known $n \times k$ full-rank design matrix with $\mathrm{rank}(\boldsymbol{X}) = k < n$.
`C`	A known matrix of size $q \times k$ with $\mathrm{rank}(\boldsymbol{C}) = q < k$.
`n`	A vector of k sample sizes. The ith element represents the sample size of group i, n_i.
`p`	The dimension of data.

Details

A high-dimensional linear regression model can be expressed as

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\Theta} + \boldsymbol{\epsilon},$$

where Θ is a $k \times p$ unknown parameter matrix and $\boldsymbol{\epsilon}$ is an $n \times p$ error matrix.

It is of interest to test the following GLHT problem

$$H_0 : \boldsymbol{C}\boldsymbol{\Theta} = \boldsymbol{0}, \quad \text{vs. } H_1 : \boldsymbol{C}\boldsymbol{\Theta} \neq \boldsymbol{0}.$$

Zhu et al. (2022) proposed the following test statistic:

$$T_{ZZZ} = \frac{(n-k-2)}{(n-k)pq} \operatorname{tr}(\boldsymbol{S}_h \boldsymbol{D}^{-1}),$$

where $\boldsymbol{S}_h$ and $\boldsymbol{S}_e$ are the variation matrices due to the hypothesis and error, respectively, and $\boldsymbol{D}$ is the diagonal matrix with the diagonal elements of $\boldsymbol{S}_e/(n-k)$. They showed that under the null hypothesis, T_{ZZZ} and a chi-squared-type mixture have the same limiting distribution.

Value

A list of class `"NRtest"` containing the results of the hypothesis test. See the help file for `NRtest.object` for details.

References

Zhu T, Zhang L, Zhang J (2023). "Hypothesis Testing in High-Dimensional Linear Regression: A Normal Reference Scale-Invariant Test." *Statistica Sinica*. doi:10.5705/ss.202020.0362.

Examples

```
library("HDNRA")
data("corneal")
dim(corneal)
group1 <- as.matrix(corneal[1:43, ]) ## normal group
group2 <- as.matrix(corneal[44:57, ]) ## unilateral suspect group
group3 <- as.matrix(corneal[58:78, ]) ## suspect map group
group4 <- as.matrix(corneal[79:150, ]) ## clinical keratoconus group
p <- dim(corneal)[2]
Y <- list()
k <- 4
Y[[1]] <- group1
Y[[2]] <- group2
Y[[3]] <- group3
Y[[4]] <- group4
n <- c(nrow(Y[[1]]),nrow(Y[[2]]),nrow(Y[[3]]),nrow(Y[[4]]))
X <- matrix(c(rep(1,n[1]),rep(0,sum(n)),rep(1,n[2]), rep(0,sum(n)),
           rep(1,n[3]),rep(0,sum(n)),rep(1,n[4])),ncol=k,nrow=sum(n))
q <- k-1
C <- cbind(diag(q),-rep(1,q))
ZZZ2022.GLHT.2cNRT(Y,X,C,n,p)
```

`ZZZ2023.TSBF.2cNRT` *Normal-reference-test with two-cumulant (2-c) matched χ^2-approximation for two-sample BF problem proposed by Zhang et al. (2023)*

Description

Zhang et al. (2023)'s test for testing equality of two-sample high-dimensional mean vectors without assuming that two covariance matrices are the same.

Usage

```
ZZZ2023.TSBF.2cNRT(y1, y2, cutoff)
```

Arguments

`y1`	The data matrix ($n_1 \times p$) from the first population. Each row represents a p-dimensional observation.
`y2`	The data matrix ($n_2 \times p$) from the second population. Each row represents a p-dimensional observation.
`cutoff`	An empirical criterion for applying the adjustment coefficient

Details

Suppose we have two independent high-dimensional samples:

$$\boldsymbol{y}_{i1}, \ldots, \boldsymbol{y}_{in_i}, \text{ are i.i.d. with } \mathrm{E}(\boldsymbol{y}_{i1}) = \boldsymbol{\mu}_i, \ \mathrm{Cov}(\boldsymbol{y}_{i1}) = \boldsymbol{\Sigma}_i, i = 1, 2.$$

The primary object is to test

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \text{ versus } H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2.$$

Zhang et al.(2023) proposed the following test statistic:

$$T_{ZZZ} = \frac{n_1 n_2}{np} (\bar{\boldsymbol{y}}_1 - \bar{\boldsymbol{y}}_2)^\top \hat{\boldsymbol{D}}_n^{-1} (\bar{\boldsymbol{y}}_1 - \bar{\boldsymbol{y}}_2),$$

where $\bar{\boldsymbol{y}}_i, i = 1, 2$ are the sample mean vectors, and $\hat{\boldsymbol{D}}_n = \mathrm{diag}(\hat{\boldsymbol{\Sigma}}_1/n + \hat{\boldsymbol{\Sigma}}_2/n)$ with $n = n_1 + n_2$. They showed that under the null hypothesis, T_{ZZZ} and a chi-squared-type mixture have the same limiting distribution.

Value

A list of class `"NRtest"` containing the results of the hypothesis test. See the help file for `NRtest.object` for details.

References

Zhang L, Zhu T, Zhang J (2023). “Two-sample Behrens–Fisher problems for high-dimensional data: a normal reference scale-invariant test.” *Journal of Applied Statistics*, **50**(3), 456–476. doi:10.1080/02664763.2020.1834516.

Examples

```
library("HDNRA")
data("COVID19")
dim(COVID19)
group1 <- as.matrix(COVID19[c(2:19, 82:87), ]) ## healthy group
group2 <- as.matrix(COVID19[-c(1:19, 82:87), ]) ## COVID-19 patients
ZZZ2023.TSBF.2cNRT(group1,group2,cutoff=1.2)
```

Index