Package 'RandomProjectionTest'

Type Package

Title Two-Sample Test in High Dimensions using Random Projection

Version 0.1.4

Description Performs the random projection test (Lopes et al., (2011) <doi:10.48550/arXiv.1108.2401>) for the one-sample and two-sample hypothesis testing problem for equality of means in the high dimensional setting. We are interested in detecting the mean vector in the one-sample problem or the difference between mean vectors in the two-sample problem.

License GPL-3

Encoding UTF-8

Imports MASS, stats, glue

RoxygenNote 7.3.1

NeedsCompilation no

Author Juan Ortiz Author [aut, cre, cph, rev]

Maintainer Juan Ortiz Author <juan.ortiz1alonso@gmail.com>

Repository CRAN

Date/Publication 2024-06-04 09:44:51 UTC

R topics documented:

---

random_projection_test

Two-Sample Test in High Dimensions using Random Projection

Description

This function performs the random projection test (Lopes et al., (2011) <arXiv:1108.2401>) for the one-sample and two-sample hypothesis testing problem for equality of means in the high dimensional setting. We are interested in detecting the mean vector in the one-sample problem or the difference between mean vectors in the two-sample problem.

Usage

```
random_projection_test(X, Y = NULL, mu0 = NULL, proj_dimension = NULL)
```

Arguments

X	The n1-by-p observation matrix with numeric column variables.
Y	An optional n2-by-p observation matrix with numeric column variables. If NULL, one-sample test is conducted on X; otherwise, a two-sample test is conducted on X and Y.
mu0	The null hypothesis vector to be tested. If NULL, the default value is the 0 vector of lenght p.
proj_dimension	Dimension where to project the given samples. If NULL, the default value is floor(n/2), where n=n1 if Y=NULL or n=n1+n2-2 if not, as in Lopes et al.

Details

Since the matrix used to project the data into a lower-dimension subset is a random matrix, obtaining the exactly same p-values in two repetitions is not likely. However, power function has been proved to perform adequately in the vast majority of settings.

Value

statistic	Value of the test's statistic T_k^2.
p_value	The p-value of the test.
degrees_freedom	
	The degrees of freedom used for the F distribution, returns list(k, n-k+1).
null_value	Returns mu0.
method	Brief description of the test that has been carried out.

Author(s)

Juan Ortiz, <juan.ortiz1alonso@gmail.com>

References

Lopes, M. E., Jacob, L. J. & Wainwright, M. J. (2011). *A More Powerful Two-Sample Test in High Dimensions using Random Projection*. <arXiv:1108.2401>.

Examples

```
set.seed(10086)
# One-sample test
n1=30; p=120
X = matrix(rnorm(n1*p), nrow = n1, ncol = p)
res1 = random_projection_test(X)

# Two-sample test
n2=65
Y = matrix(rnorm(n2*p), nrow = n2, ncol = p)
res2 = random_projection_test(X, Y)

# Specify a null hypothesis vector
res3 = random_projection_test(X, Y, mu0 = rep(0.1, p))

# Choose a projection dimension manually, will work worse than previous example
res4 = random_projection_test(X, Y, mu0 = rep(0.1, p), proj_dimension = 4)
```

Index