Package: ScreeNOT (via r-universe)

November 5, 2024

Type Package

Title 'ScreeNOT': MSE-Optimal Singular Value Thresholding in Correlated Noise

Version 0.1.0

Description Optimal hard thresholding of singular values. The procedure adaptively estimates the best singular value threshold under unknown noise characteristics. The threshold chosen by 'ScreeNOT' is optimal (asymptotically, in the sense of minimum Frobenius error) under the the so-called ``Spiked model'' of a low-rank matrix observed in additive noise. In contrast to previous works, the noise is not assumed to be i.i.d. or white; it can have an essentially arbitrary and unknown correlation structure, across either rows, columns or both. 'ScreeNOT' is proposed to practitioners as a mathematically solid alternative to Cattell's ever-popular but vague Scree Plot heuristic from 1966. If you use this package, please cite our paper: David L. Donoho, Matan Gavish and Elad Romanov (2023). ``ScreeNOT: Exact MSE-optimal singular value thresholding in correlated noise.'' Annals of Statistics, 2023 (To appear). <arXiv:2009.12297>.

License MIT + file LICENSE

Encoding UTF-8

RoxygenNote 7.2.3

NeedsCompilation no

Author Elad Romanov [aut, cre]

Maintainer Elad Romanov <elad.romanov@gmail.com>

Date/Publication 2023-02-27 08:22:36 UTC

Repository https://eladromanov.r-universe.dev

RemoteUrl https://github.com/cran/ScreeNOT

RemoteRef HEAD

RemoteSha 435eb483bafa612c54f64360d474c3a1e2e0ceaa

Contents

---

adaptiveHardThresholding

Adaptive hard thresholding

Description

Performs optimal adaptive hard thresholding on the input matrix Y.

Usage

```
adaptiveHardThresholding(Y, k, strategy = "i")
```

Arguments

Y	A data matrix, on whose singular values thresholding should be performed.
k	An upper bound (potentially loose) on the latent signal rank. That is, the procedure assumes that there are AT MOST k informative principal components of Y.
strategy	Method for reconstructing the noise bulk (optional). Can be one of the following: '0': tranpsort to zero; 'w': winsorization; 'i': imputation (default option).

Value

Xest	An estimate of the low-rank signal. That is: the matrix obtained by thresholding the singular values of Y.
Topt	The hard threshold computed by the procedure. To wit, the procedure retains the i-th PC of Y if and only if the corresponding singular value, y_i, satisfies $y_i > Topt$.
r	The number of "relevant" components: r = rank(Xest).

Author(s)

Elad Romanov

References

David L. Donoho, Matan Gavish and Elad Romanov. "ScreeNOT: Exact MSE-optimal singular value thresholding in correlated noise." Annals of Statistics (2023). https://github.com/eladromanov/ScreeNOT

Examples

```
Y <- matrix(rnorm(1e6)/sqrt(1e3),nrow=1e3)
    # Y is a 1000x1000 i.i.d. Gaussian matrix
val <- ScreeNOT::adaptiveHardThresholding(Y, 10)
    # Runs the ScreeNOT procedure, with an upper bound k=10
cat('Computed threshold: ', val$Topt)
    # The adaptively computed threshold
cat('Known optimal threshold: ', 4/sqrt(3))
    # The known optimal threshold for this noise bulk
```

Index