High Dimensional Data Analysis With Dependency and Under Limited Memory

Yiming Sun

Cornell University

Based on joint work with
Madeleine Udell (Cornell), Sumanta Basu(Cornell)
Yang Guo (UW Madison), Charlene Luo (Columbia)
Joel Tropp (Caltech), Amy Kuceyeski(Cornell University)
Yige Li(Harvard University)

September 8, 2019

Outline

Multivariate Spectral Density Estimation Under Weak Sparsity

Low Rank Tucker Approximation of a Tensor from Streaming Data

A Connection of These Two

Motivation

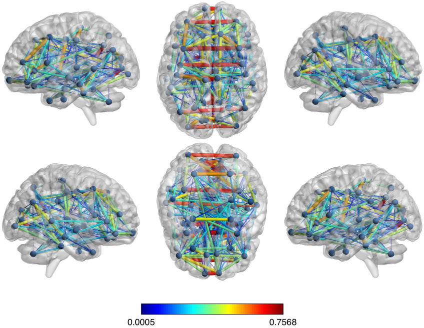

Figure: Interactions Between Regions in Brain

Weakly & Strongly Stationary Time

Definition (Weak Stationarity)

p-variate time series X is weakly stationary, if $\mathbb{E}X_t = \mathbb{E}X_s$ for any t, s and $\Gamma(\ell) := \mathbb{E}X_t X_{t-\ell}^\top$ only depends on the lag ℓ.

Definition (Strong Stationarity)

p-variate time series X is strongly stationary, if for any sequence $t_1, \cdots, t_n$, $X_{t_1} \cdots X_{t_n}$ has the same distribution of $X_{t_1+\tau} \cdots X_{t_n+\tau}$ for any integer τ,

Gaussian Process

Definition (Gaussian Process)

p-variate time series X is Gaussian process if for any sequence $t_1, \cdots, t_n$, $X_{t_1} \cdots X_{t_n}$ are jointly Gaussian distributed.

For Gaussian process, weak stationarity is equivalent to strong stationarity.

Spectral Density

Given a weakly stationary p-variate time series X, the spectral density at frequency $\omega \in [-\pi, \pi)$ is defined

$$f(\omega) = \sum_{\ell=-\infty}^{\infty} \Gamma(\ell) e^{-i\omega\ell}$$

where $\Gamma(\ell) = \mathbb{E} X_0 X_{-\ell}^{\top}$. X_t is independent with X_s, $t \neq s$ iff $f_{rs}(\omega) = 0$ for any ω.

Thresholding Estimator Under Weak Sparsity- A Example

Suppose that we have n observation of p-variate Gaussian distribution as follows.

$$y_i \overset{i.i.d}{\sim} \mathcal{N}\left(\mu, \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & \sigma_p \end{bmatrix}\right),$$

$i = 1, \cdots, n$.

A Example

The maximum likelihood estimator for μ_j is $\bar{y}_j = \frac{1}{n}\sum_{i=1}^{n} y_{ij}$.
Does not Work Well Under Weak Sparsity

$$\mu \in \left\{ \mu \in \mathbb{R}^p, \sum_{j=1}^{p} |\mu_j|^q \leq s(p) \right\}.$$

for some $0 \leq q < 1$.

Solution : Thresholding

Suppose $\sigma_i \leq B$, define element-wise thresholding operator

$$T_\lambda(x) = \begin{cases} x & |x| \geq \lambda \\ 0 & \text{else} \end{cases}$$

. Hard thresholding estimator $T_\lambda(\bar{y}_j)$ can be shown asymptotically consistent under weak sparsity where we set

$$\lambda \propto B\sqrt{\frac{\log p}{n}}.$$

Outline

Multivariate Spectral Density Estimation Under Weak Sparsity

Low Rank Tucker Approximation of a Tensor from Streaming Data

A Connection of These Two

Outline

Multivariate Spectral Density Estimation Under Weak Sparsity

Low Rank Tucker Approximation of a Tensor from Streaming Data

A Connection of These Two