Overlap Gap Property for Submatrix Recovery

Subhabrata Sen

Department of Statistics
Harvard University

Joint work with David Gamarnik and Aukosh Jagannath

July 2, 2020

Statistical/Computational Tradeoffs

Intriguing Phenomena

- In many high-dimensional problems, inference might be information theoretically possible, but computationally intractable!
- Believed to be widespread — e.g. Community Detection, Planted Clique, sPCA, tensor recovery, regression etc.
- Accumulating evidence using diverse approaches:
    - Avg. case reductions
    - State Evolution Analysis
    - Sum of Squares lower bounds and the low-degree likelihood method
    - Query Lower bounds
    - ...

Planted submatrix recovery

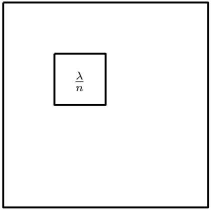
Setup

- We observe $M = \mathbb{R}^{n\times n}$, $M = M^{\mathrm{T}}$.

$$M = \frac{\lambda}{n} vv^{\mathrm{T}} + W.$$

- $v \in \{0,1\}^n$, $\sum_{i=1}^n v_i = n\rho$, $\rho \in (0,1)$.

- $W = (W_{ij})$, $\{W_{ij} : i \leq j\}$ ind. $W_{ij} \sim \mathcal{N}(0, \frac{1}{n})$ $i < j$.



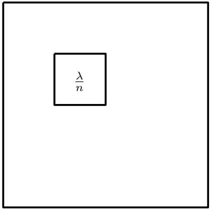
Setup

- We observe $M = \mathbb{R}^{n \times n}$, $M = M^{\mathrm{T}}$.

$$M = \frac{\lambda}{n} v v^{\mathrm{T}} + W.$$

- $v \in \{0,1\}^n$, $\sum_{i=1}^n v_i = n\rho$,
  $\rho \in (0,1)$.

- $W = (W_{ij})$, $\{W_{ij} : i \leq j\}$ ind.
  $W_{ij} \sim \mathcal{N}(0, \frac{1}{n})$ $i < j$.



Today: $\rho \to 0$, following $n \to \infty$.

Natural Questions

- Can we detect the presence of this planted submatrix?
- Can we recover the submatrix?
- Can we recover it *efficiently* ?

Natural Questions

- Can we detect the presence of this planted submatrix?
- Can we recover the submatrix?
- Can we recover it *efficiently* ?

History of the problem

- Deshpande-Montanari '14 — Bayesian formulation, low rank matrix recovery, AMP optimal for some $\rho > \rho_c$.
- Lesieur-Krzakala-Zdeborova '15, '17 — Replica analysis of Mutual Information, State evolution analysis of AMP.

Likelihood Landscape Analysis

The MLE

$$\max \langle x, Mx \rangle$$
$$\text{subject to } x \in \{0,1\}^n, \sum_i x_i = n\rho.$$

The MLE

$$\max \langle x, Mx \rangle$$
$$\text{subject to}\, x \in \{0,1\}^n, \sum_i x_i = n\rho.$$

Definition (Reliable Recovery)

An estimator $\hat{v} := \hat{v}(M) \in \{0,1\}^n$ *reliably recovers* the vector v if there exists $c > 0$ such that w.h.p.

$$\langle \hat{v}, v \rangle \geq cn\rho.$$

Reliable Recovery

Theorem (Gamarnik, Jagannath, S. '19)

If $\lambda = o\left(\sqrt{\frac{1}{\rho}\log\frac{1}{\rho}}\right)$*, the MLE does not recover the support reliably. On the other hand, if* $\lambda > (2+\varepsilon)\sqrt{\frac{1}{\rho}\log\frac{1}{\rho}}$ *for some* $\varepsilon > 0$*, there exists* $c := c(\varepsilon) > 0$ *such that* $\langle \hat{v}_{\mathrm{MLE}}, v\rangle > cn\rho$ *w.h.p.*

A naive Spectral Algorithm

(i) Let $\tilde{v}$ be the top eigenvector of M with $\|\tilde{v}\|_2^2 = n\rho$. Set $S = \{i \in [n] : \tilde{v}_i \geq \frac{\delta}{2}\}$.

(ii) If $|S| < n\rho$, sample $(n\rho - |S|)$ elements from S^c and augment to S. Define the new set $\tilde{S}$.

(iii) If $|S| > n\rho$, sample $n\rho$ elements randomly from S to construct $\tilde{S}$.

(iv) Set $\hat{v} = \mathbf{1}_S$.

A naive Spectral Algorithm (cont'd)

Theorem

For any $\varepsilon > 0$ and $\lambda > (1+\varepsilon)\frac{1}{\rho}$, there exists $\delta := \delta(\varepsilon) > 0$ and $c := c(\varepsilon, \delta) > 0$ such that w.h.p. $\langle \hat{v}, v \rangle \geq cn\rho$.

What happens for intermediate ρ?

The Restricted Likelihood

$$E_n(q : \rho, \lambda) := \frac{1}{n} \max\langle x, Mx\rangle$$
$$\text{subject to } \langle x, v\rangle \approx nq, x \in \{0,1\}^n, \sum_i x_i = n\rho.$$

The Restricted Likelihood

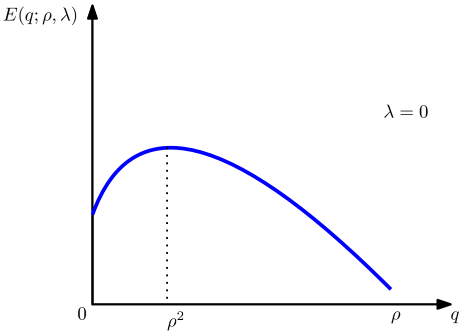
$$E_n(q : \rho, \lambda) := \frac{1}{n} \max\langle x, Mx\rangle$$
$$\text{subject to } \langle x, v\rangle \approx nq, x \in \{0,1\}^n, \sum_i x_i = n\rho.$$

Theorem (Gamarnik, Jagannath, S. '19)

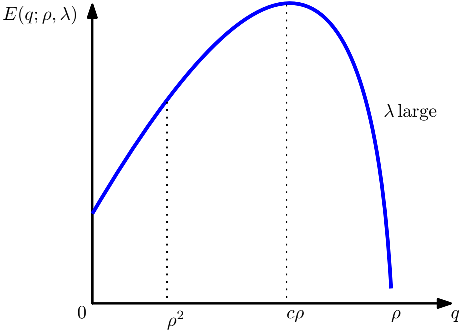
For any $q \in [0, \rho]$, there exists $E(\cdot; \rho, \lambda)$ such that

$$E_n(q; \rho, \lambda) \to E(q; \rho, \lambda).$$

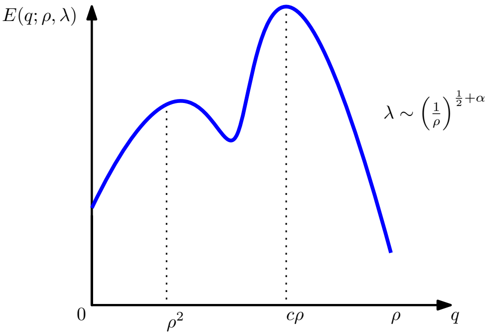
An Overlap Gap Property



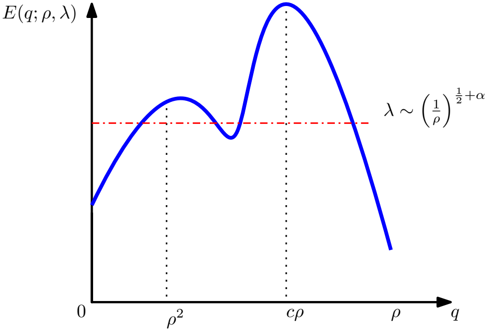
An Overlap Gap Property



An Overlap Gap Property



An Overlap Gap Property



Consequences for Markov Chains

(i) Fix $\beta > 0$, and let $\pi_\beta(x) \propto \exp(\beta\langle x, Mx\rangle)$.

(ii) Let $\{X_t : t \geq 0\}$ be a *local* Markov Chain, reversible with respect to π_β.

Theorem (Gamarnik, Jagannath, S. '19 (informal))

Let $\lambda \sim \left(\frac{1}{\rho}\right)^{\frac{1}{2}+\alpha}$, $\alpha > 0$ *sufficiently small. Define*

$$A = \left\{x \in \{0,1\}^n : \sum_i x_i = n\rho, an < \sum_i x_i v_i < bn\right\}.$$

For $\rho > 0$ *sufficiently small, there exists* $a < \rho^2 < b = o(\rho)$ *such that*

$$\int_A Q_x(\tau_{A^c} \leq T) \mathrm{d}\pi_\beta(x) \leq T\exp(-\Theta(N)).$$

Glimpse of Proof

- A Parisi type formula for the restricted likelihood.
- Overlap Gap Property based on variational analysis of ground state formulae for spin glasses.
- Reliable recovery thresholds based on moment based arguments.

Subsequent Developments

- Barbier, Macris '19, Barbier, Macris, Rush '20 — Mutual information in the regime $\rho_n \to 0$.

- Ben Arous, Wein and Zadik (COLT 2020) — Extended analysis for $k \times k$ planted matrices with $k = o(n)$.

THANK YOU!