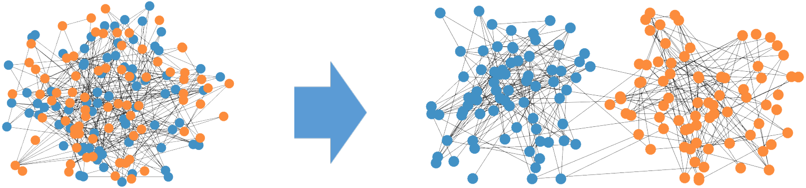
Spectral methods: ℓ_∞ perturbation theory

Cong Ma

University of Chicago, Autumn 2021

Revisit stochastic block model



$\mathcal{G}$

- Community membership vector

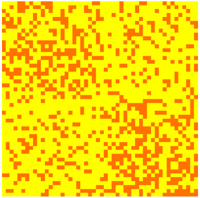
$$x_1^\star = \cdots = x_{n/2}^\star = 1;\ x_{n/2+1}^\star = \cdots = x_n^\star = -1$$

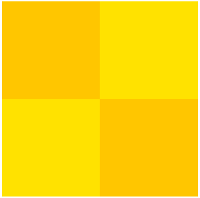
- observe a graph $\mathcal{G}$ (assuming $p > q$)

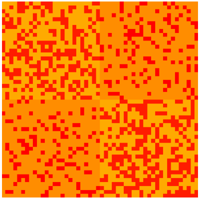
$$(i,j) \in \mathcal{G} \text{ with prob. } \begin{cases} p, & \text{if } x_i = x_j \\ q, & \text{else} \end{cases}$$

- **Goal:** recover community memberships $\pm \boldsymbol{x}^\star$

Revisit spectral clustering







$$\boldsymbol{A} \quad = \quad \underbrace{\mathbb{E}[\boldsymbol{A}]}_{\text{rank 2}} \quad + \quad \boldsymbol{A} - \mathbb{E}\left[\boldsymbol{A}\right]$$

1. computing the leading eigenvector $\boldsymbol{u} = [u_i]_{1 \le i \le n}$ of $\boldsymbol{A} - \frac{p+q}{2}\mathbf{1}\mathbf{1}^\top$
2. rounding: output $x_i = \begin{cases} 1, & \text{if } u_i \ge 0 \\ -1, & \text{if } u_i < 0 \end{cases}$

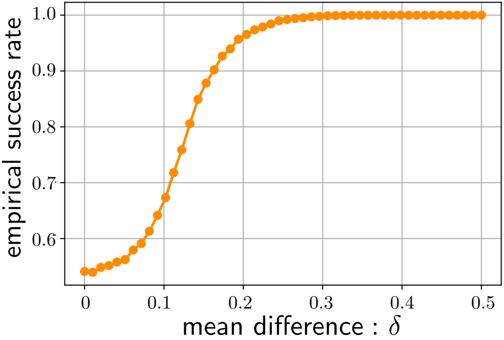
Almost exact recovery

$$\frac{p-q}{\sqrt{p}} \gg \sqrt{\frac{\log n}{n}} \quad \Longrightarrow \quad \text{almost exact recovery}$$

- Almost exact recovery means

$$\min\left\{\frac{1}{n}\sum_{i=1}^{n}\mathbb{1}\{x_i \neq x_i^\star\}, \frac{1}{n}\sum_{i=1}^{n}\mathbb{1}\{x_i \neq -x_i^\star\}\right\} = o(1)$$

Empirical performance of spectral clustering



ℓ_2 perturbation theory alone cannot explain exact recovery guarantees

— call for fine-grained analysis

Reverse engineering

Spectral clustering uses signs of $\boldsymbol{u}$ to cluster nodes

Reverse engineering

Spectral clustering uses signs of $\boldsymbol{u}$ to cluster nodes

$$\Uparrow$$

It achieves exact recovery iff $u_i u_i^\star > 0$ for all i

Reverse engineering

Spectral clustering uses signs of $\boldsymbol{u}$ to cluster nodes

$$\Uparrow$$

It achieves exact recovery iff $u_i u_i^\star > 0$ for all i

$$\Uparrow$$

A sufficient condition is* $\|\boldsymbol{u} - \boldsymbol{u}^\star\|_\infty < 1/\sqrt{n}$

Reverse engineering

Spectral clustering uses signs of $\boldsymbol{u}$ to cluster nodes

$\Uparrow$

It achieves exact recovery iff $u_i u_i^\star > 0$ for all i

$\Uparrow$

A sufficient condition is* $\|\boldsymbol{u} - \boldsymbol{u}^\star\|_\infty < 1/\sqrt{n}$

$\Uparrow$

Need ℓ_∞ perturbation theory

Outline

- An illustrative example: rank-1 matrix denoising
- General ℓ_∞ perturbation theory
- Application: exact recovery in community detection
- Application: entrywise error in matrix completion

Setup and algorithm

- Groundtruth: $\boldsymbol{M}^\star = \lambda^\star \boldsymbol{u}^\star \boldsymbol{u}^{\star\top} \in \mathbb{R}^{n\times n}$, with $\lambda^\star > 0$
- Observation: $\boldsymbol{M} = \boldsymbol{M}^\star + \boldsymbol{E}$, where $\boldsymbol{E}$ is symmetric, and its upper triangular part comprises of i.i.d. $\mathcal{N}(0, \sigma^2)$ entries
- Estimate $\boldsymbol{u}^\star$ using $\boldsymbol{u}$, leading eigenvector of $\boldsymbol{M}$
- Goal: characterize entrywise errror

$$\mathsf{dist}_\infty(\boldsymbol{u}, \boldsymbol{u}^\star) := \min\left\{\|\boldsymbol{u} - \boldsymbol{u}^\star\|_\infty, \|\boldsymbol{u} + \boldsymbol{u}^\star\|_\infty\right\}$$

ℓ_2 guarantees

We start with characterizing noise size

Lemma 6.1

Assume symmetric Gaussian noise model. With high prob., one has

$$\|\boldsymbol{E}\| \leq 5\sigma\sqrt{n}$$

This in conjunction with Davis-Kahan's $\sin\boldsymbol{\Theta}$ theorem leads to:

$$\mathsf{dist}(\boldsymbol{u}, \boldsymbol{u}^\star) \leq \frac{2\|\boldsymbol{E}\|}{\lambda^\star} \leq \frac{10\sigma\sqrt{n}}{\lambda^\star},$$

as long as $\sigma\sqrt{n} \leq \frac{1-1/\sqrt{2}}{5}\lambda^\star$ so that $\|\boldsymbol{E}\| \leq (1 - 1/\sqrt{2})\lambda^\star$

— *implies* $\mathsf{dist}_\infty\big(\boldsymbol{u}, \boldsymbol{u}^\star\big) \leq \mathsf{dist}\big(\boldsymbol{u}, \boldsymbol{u}^\star\big) \lesssim \frac{\sigma\sqrt{n}}{\lambda}$

Incoherence

Definition 6.2

Fix a unit vector $\boldsymbol{u}^\star \in \mathbb{R}^n$. Define its incoherence to be

$$\mu := n\|\boldsymbol{u}^\star\|_\infty^2$$

- Range of possible values of μ: $1 \leq \mu \leq n$
- Two extremes: $\boldsymbol{u}^\star = \boldsymbol{e}_1$, and $\boldsymbol{u}^\star = (1/\sqrt{n}) \cdot \mathbf{1}_n$
- Small μ indicates energy of eigenvector is spread across different entries
- Consider SBM and random Gaussian vectors

ℓ_∞ guarantees for matrix denoising

Theorem 6.3

Suppose that $\sigma\sqrt{n} \leq c_0\lambda^\star$ *for some sufficiently small constant* $c_0 > 0$. *Then whp., we have*

$$\mathsf{dist}_\infty(\boldsymbol{u}, \boldsymbol{u}^\star) \lesssim \frac{\sigma(\sqrt{\log n} + \sqrt{\mu})}{\lambda^\star}$$

- When $\mu \lesssim \log n$ (i.e., no entries are significantly larger than average), our bound reads

$$\mathsf{dist}_\infty(\boldsymbol{u}, \boldsymbol{u}^\star) \lesssim \frac{\sigma\sqrt{\log n}}{\lambda^\star}$$

- Much tighter than ℓ_2 bound: $\sqrt{n/\log n}$ times smaller

Technical hurdle: dependency

We would like to understand u_l. Since $\boldsymbol{u}$ is eigenvector of $\boldsymbol{M}$, we have

$$\boldsymbol{M}\boldsymbol{u} = \lambda \boldsymbol{u},$$

which yields

$$u_l = \frac{1}{\lambda}[\boldsymbol{M}]_{l,:}\boldsymbol{u} = \frac{1}{\lambda}[\boldsymbol{M}^\star + \boldsymbol{E}]_{l,:}\boldsymbol{u}$$

$\boldsymbol{u}$ is dependent on $\boldsymbol{E}$; analyzing $[\boldsymbol{M}^\star + \boldsymbol{E}]_{l,:}\boldsymbol{u}$ is challenging

—*how to deal with such dependency*

An independent proxy

Recall our focus is

$$[\boldsymbol{M}^{\star} + \boldsymbol{E}]_{l,:}\boldsymbol{u}$$

Suppose we have a proxy $\boldsymbol{u}^{(l)}$ which is independent of $[\boldsymbol{E}]_{l,:}$, then

$$[\boldsymbol{M}^{\star} + \boldsymbol{E}]_{l,:}\boldsymbol{u} = [\boldsymbol{M}^{\star} + \boldsymbol{E}]_{l,:}\boldsymbol{u}^{(l)} + [\boldsymbol{M}^{\star} + \boldsymbol{E}]_{l,:}\left(\boldsymbol{u} - \boldsymbol{u}^{(l)}\right)$$

- Independence between $\boldsymbol{u}^{(l)}$ and $[\boldsymbol{E}]_{l,:}$
- Proximity between $\boldsymbol{u}^{(l)}$ and $\boldsymbol{u}$

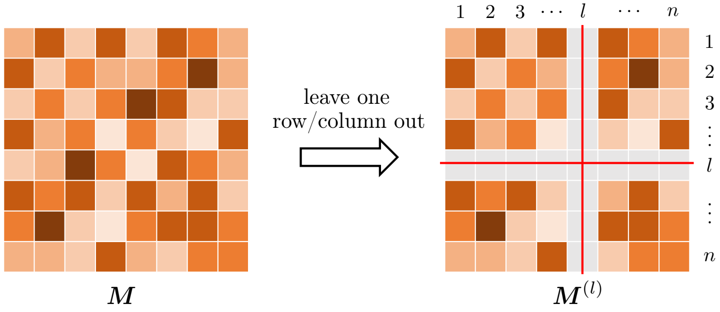
Leave-one-out estimates

For each $1 \le l \le n$, construct an auxiliary matrix $\boldsymbol{M}^{(l)}$

$$\boldsymbol{M}^{(l)} := \lambda^\star \boldsymbol{u}^\star \boldsymbol{u}^{\star\top} + \boldsymbol{E}^{(l)},$$

where the noise matrix $\boldsymbol{E}^{(l)}$ is generated according to

$$E_{i,j}^{(l)} := \begin{cases} E_{i,j}, & \text{if } i \neq l \text{ and } j \neq l, \\ 0, & \text{else.} \end{cases}$$



Leave-one-out estimates (cont.)

For each $1 \le l \le n$, construct an auxiliary matrix $\boldsymbol{M}^{(l)}$

$$\boldsymbol{M}^{(l)} := \lambda^{\star}\boldsymbol{u}^{\star}\boldsymbol{u}^{\star\top} + \boldsymbol{E}^{(l)},$$

where the noise matrix $\boldsymbol{E}^{(l)}$ is generated according to

$$E_{i,j}^{(l)} := \begin{cases} E_{i,j}, & \text{if } i \neq l \text{ and } j \neq l, \\ 0, & \text{else.} \end{cases}$$

Let $\lambda^{(l)}$ and $\boldsymbol{u}^{(l)}$ denote respectively leading eigenvalue and leading eigenvector of $\boldsymbol{M}^{(l)}$

—$\boldsymbol{u}^{(l)}$ *is independent of* $[\boldsymbol{E}]_{l,:}$

Intuition

- Since $\boldsymbol{u}^{(l)}$ is obtained by dropping only a tiny fraction of data, we expect $\boldsymbol{u}^{(l)}$ to be extremely close to $\boldsymbol{u}$, i.e., $\boldsymbol{u} \approx \pm \boldsymbol{u}^{(l)}$
- By construction,

$$
\begin{aligned}
u_l^{(l)} &= \frac{1}{\lambda^{(l)}} \boldsymbol{M}_{l,\cdot}^{(l)} \boldsymbol{u}^{(l)} = \frac{1}{\lambda^{(l)}} \boldsymbol{M}_{l,\cdot}^{\star} \boldsymbol{u}^{(l)} = \frac{\lambda^{\star}}{\lambda^{(l)}} u_l^{\star} \boldsymbol{u}^{\star\top} \boldsymbol{u}^{(l)} \\
&\approx \pm u_l^{\star}.
\end{aligned}
$$

Proof of Theorem 6.3

What we have learned from ℓ_2 analysis

$$\|\boldsymbol{E}\| \leq 5\sigma\sqrt{n} \qquad \|\boldsymbol{E}^{(l)}\| \leq \|\boldsymbol{E}\| \leq 5\sigma\sqrt{n}$$

$$\mathsf{dist}(\boldsymbol{u}, \boldsymbol{u}^\star) \leq \frac{10\sigma\sqrt{n}}{\lambda^\star} \qquad \mathsf{dist}(\boldsymbol{u}^{(l)}, \boldsymbol{u}^\star) \leq \frac{10\sigma\sqrt{n}}{\lambda^\star}$$

$$|\lambda - \lambda^\star| \leq 5\sigma\sqrt{n} \qquad |\lambda^{(l)} - \lambda^\star| \leq 5\sigma\sqrt{n}$$

$$\max_{j:j\geq 2} |\lambda_j(\boldsymbol{M})| \leq 5\sigma\sqrt{n} \qquad \max_{j:j\geq 2} |\lambda_j(\boldsymbol{M}^{(l)})| \leq 5\sigma\sqrt{n}$$

Addressing ambiguity

Assume WLOG,

$$\|\boldsymbol{u} - \boldsymbol{u}^\star\|_2 = \mathsf{dist}(\boldsymbol{u}, \boldsymbol{u}^\star),$$
$$\big\|\boldsymbol{u}^{(l)} - \boldsymbol{u}^\star\big\|_2 = \mathsf{dist}(\boldsymbol{u}^{(l)}, \boldsymbol{u}^\star), \quad 1 \le l \le n$$

A useful byproduct: if $20\sigma\sqrt{n} < \lambda^\star$, then one necessarily has

$$\big\|\boldsymbol{u} - \boldsymbol{u}^{(l)}\big\|_2 = \mathsf{dist}\big(\boldsymbol{u}, \boldsymbol{u}^{(l)}\big), \qquad 1 \le l \le n$$

—*check this*

Bounding $\|\boldsymbol{u} - \boldsymbol{u}^{(l)}\|_2$

Key: view $\boldsymbol{M}$ as perturbation of $\boldsymbol{M}^{(l)}$; apply "sharper" version of Davis-Kahan

$$\left\|\boldsymbol{u} - \boldsymbol{u}^{(l)}\right\|_2 \leq \frac{2\|(\boldsymbol{M} - \boldsymbol{M}^{(l)})\boldsymbol{u}^{(l)}\|_2}{\lambda^{(l)} - \max_{j\geq 2}\left|\lambda_j(\boldsymbol{M}^{(l)})\right|} \leq \frac{4\|(\boldsymbol{M} - \boldsymbol{M}^{(l)})\boldsymbol{u}^{(l)}\|_2}{\lambda^\star}$$

as long as

$$\|\boldsymbol{M} - \boldsymbol{M}^{(l)}\| \leq (1 - 1/\sqrt{2})\Big(\lambda^{(l)} - \max_{j\geq 2}\left|\lambda_j(\boldsymbol{M}^{(l)})\right|\Big),$$

$$\lambda^{(l)} - \max_{j\geq 2}\left|\lambda_j(\boldsymbol{M}^{(l)})\right| \geq \lambda^\star/2$$

Bounding $\|(\boldsymbol{M} - \boldsymbol{M}^{(l)})\boldsymbol{u}^{(l)}\|_2$

By design,

$$(\boldsymbol{M} - \boldsymbol{M}^{(l)})\boldsymbol{u}^{(l)} = \boldsymbol{e}_l \boldsymbol{E}_{l,\cdot}\boldsymbol{u}^{(l)} + u_l^{(l)}(\boldsymbol{E}_{\cdot,l} - E_{l,l}\boldsymbol{e}_l),$$

which together with triangle inequality yields

$$\begin{aligned}
\|(\boldsymbol{M} - \boldsymbol{M}^{(l)})\boldsymbol{u}^{(l)}\|_2 &\leq \left|\boldsymbol{E}_{l,\cdot}\boldsymbol{u}^{(l)}\right| + \left\|\boldsymbol{E}_{\cdot,l}\right\|_2 \cdot \left|u_l^{(l)}\right| \\
&\leq 5\sigma\sqrt{\log n} + \left\|\boldsymbol{E}_{\cdot,l}\right\|_2 (|u_l| + \left\|\boldsymbol{u} - \boldsymbol{u}^{(l)}\right\|_\infty) \\
&\leq 5\sigma\sqrt{\log n} + 5\sigma\sqrt{n}\|\boldsymbol{u}\|_\infty + 5\sigma\sqrt{n}\left\|\boldsymbol{u} - \boldsymbol{u}^{(l)}\right\|_2
\end{aligned}$$

Bounding $\|\boldsymbol{u} - \boldsymbol{u}^{(l)}\|_2$ (cont.)

Combining previous bounds, we arrive at

$$\begin{aligned}\big\|\boldsymbol{u} - \boldsymbol{u}^{(l)}\big\|_2 &\leq \frac{20\sigma\sqrt{\log n} + 20\sigma\sqrt{n}\|\boldsymbol{u}\|_\infty + 20\sigma\sqrt{n}\big\|\boldsymbol{u} - \boldsymbol{u}^{(l)}\big\|_2}{\lambda^\star} \\ &\leq \frac{20\sigma\sqrt{\log n} + 20\sigma\sqrt{n}\|\boldsymbol{u}\|_\infty}{\lambda^\star} + \frac{1}{2}\big\|\boldsymbol{u} - \boldsymbol{u}^{(l)}\big\|_2,\end{aligned}$$

provided that $40\sigma\sqrt{n} \leq \lambda^\star$
Rearranging terms and taking the union bound, we demonstrate that whp.,

$$\big\|\boldsymbol{u} - \boldsymbol{u}^{(l)}\big\|_2 \leq \frac{40\sigma\sqrt{\log n} + 40\sigma\sqrt{n}\|\boldsymbol{u}\|_\infty}{\lambda^\star} \qquad 1 \leq l \leq n$$

Analyzing leave-one-out iterates

Recall that

$$u_l^{(l)} = \frac{1}{\lambda^{(l)}} \boldsymbol{M}_{l,\cdot}^{(l)} \boldsymbol{u}^{(l)} = \frac{1}{\lambda^{(l)}} \boldsymbol{M}_{l,\cdot}^{\star} \boldsymbol{u}^{(l)} = \frac{\lambda^\star}{\lambda^{(l)}} u_l^\star \boldsymbol{u}^{\star\top} \boldsymbol{u}^{(l)}$$

This implies

$$\begin{aligned} u_l^{(l)} - u_l^\star &= u_l^\star \Big(\frac{\lambda^\star}{\lambda^{(l)}} \boldsymbol{u}^{\star\top} \boldsymbol{u}^{(l)} - \boldsymbol{u}^{\star\top} \boldsymbol{u}^\star \Big) \\ &= u_l^\star \Big(\frac{\lambda^\star - \lambda^{(l)}}{\lambda^{(l)}} \boldsymbol{u}^{\star\top} \boldsymbol{u}^{(l)} \Big) + u_l^\star \boldsymbol{u}^{\star\top} \big(\boldsymbol{u}^{(l)} - \boldsymbol{u}^\star \big) \end{aligned}$$

Analyzing leave-one-out iterates (cont.)

Triangle inequality gives

$$\begin{aligned}
\left|u_l^{(l)} - u_l^{\star}\right| &\leq \left|u_l^{\star}\right| \cdot \frac{\left|\lambda^{\star} - \lambda^{(l)}\right|}{\left|\lambda^{(l)}\right|} \cdot \|\boldsymbol{u}^{\star}\|_2 \cdot \|\boldsymbol{u}^{(l)}\|_2 \\
&\quad + \left|u_l^{\star}\right| \cdot \|\boldsymbol{u}^{\star}\|_2 \cdot \left\|\boldsymbol{u}^{(l)} - \boldsymbol{u}^{\star}\right\|_2 \\
&\leq \left|u_l^{\star}\right| \cdot \frac{10\sigma\sqrt{n}}{\lambda^{\star}} + \left|u_l^{\star}\right| \cdot \frac{10\sigma\sqrt{n}}{\lambda^{\star}} \\
&\leq \frac{20\sigma\sqrt{n}}{\lambda^{\star}} \|\boldsymbol{u}^{\star}\|_{\infty}
\end{aligned}$$

Putting pieces together

Now we come to conclude that

$$\left\|\boldsymbol{u}-\boldsymbol{u}^{\star}\right\|_{\infty}=\max_{l}\left|u_{l}-u_{l}^{\star}\right| \leq \max_{l}\left\{\left|u_{l}^{(l)}-u_{l}^{\star}\right|+\left\|\boldsymbol{u}-\boldsymbol{u}^{(l)}\right\|_{2}\right\}$$

$$\leq \frac{20\sigma\sqrt{n}}{\lambda^{\star}}\left\|\boldsymbol{u}^{\star}\right\|_{\infty}+\frac{40\sigma\sqrt{\log n}+40\sigma\sqrt{n}\|\boldsymbol{u}\|_{\infty}}{\lambda^{\star}}$$

One more triangle inequality gives

$$\left\|\boldsymbol{u}-\boldsymbol{u}^{\star}\right\|_{\infty} \leq \frac{40\sigma\sqrt{\log n}+60\sigma\sqrt{n}\left\|\boldsymbol{u}^{\star}\right\|_{\infty}}{\lambda^{\star}}+\frac{1}{2}\left\|\boldsymbol{u}-\boldsymbol{u}^{\star}\right\|_{\infty},$$

provided that $80\sigma\sqrt{n} \leq \lambda^{\star}$. Rearranging terms yields

$$\left\|\boldsymbol{u}-\boldsymbol{u}^{\star}\right\|_{\infty} \leq \frac{80\sigma\sqrt{\log n}+120\sigma\sqrt{n}\left\|\boldsymbol{u}^{\star}\right\|_{\infty}}{\lambda^{\star}}=\frac{80\sigma\sqrt{\log n}+120\sigma\sqrt{\mu}}{\lambda^{\star}},$$

where the last identity results from the definition of μ

General ℓ_∞ perturbation theory

—rank-1 case

Setup and notation

Groundtruth: consider a rank-1 psd matrix $\boldsymbol{M}^\star = \lambda^\star \boldsymbol{u}^\star \boldsymbol{u}^{\star\top} \in \mathbb{R}^{n\times n}$

Incoherence: define

$$\mu := n\|\boldsymbol{u}^\star\|_\infty^2 \qquad (1 \le \mu \le n)$$

Observations:

$$\boldsymbol{M} = \boldsymbol{M}^\star + \boldsymbol{E} \in \mathbb{R}^{n\times n}$$

with $\boldsymbol{E}$ a symmetric noise matrix

Noise assumptions

The entries in the lower triangular part of $\boldsymbol{E} = [E_{i,j}]_{1 \le i,j \le n}$ are independently generated obeying

$$\mathbb{E}[E_{i,j}] = 0, \quad \mathbb{E}[E_{i,j}^2] \le \sigma^2, \quad |E_{i,j}| \le B, \quad \text{for all } i \ge j$$

Further, assume that

$$c_{\mathsf{b}} := \frac{B}{\sigma\sqrt{n/(\mu \log n)}} = O(1)$$

ℓ_∞ perturbation theory

Theorem 6.4

With high prob, there exists $z \in \{1, -1\}$ *such that*

$$\|z\boldsymbol{u} - \boldsymbol{u}^\star\|_\infty \lesssim \frac{\sigma\sqrt{\mu} + \sigma\sqrt{\log n}}{\lambda^\star}, \tag{6.3a}$$

$$\Big\|z\boldsymbol{u} - \frac{1}{\lambda^\star}\boldsymbol{M}\boldsymbol{u}^\star\Big\|_\infty \lesssim \frac{\sigma\sqrt{\mu}}{\lambda^\star} + \frac{\sigma^2\sqrt{n\log n} + \sigma B\sqrt{\mu\log^3 n}}{(\lambda^\star)^2} \tag{6.3b}$$

provided that $\sigma\sqrt{n\log n} \le c_\sigma\lambda^\star$ *for some sufficiently small constant* $c_\sigma > 0$.

First-order expansion

$$\boldsymbol{u} = \frac{\boldsymbol{M}\boldsymbol{u}}{\lambda} \approx \frac{\boldsymbol{M}\boldsymbol{u}^\star}{\lambda^\star} \approx \frac{\boldsymbol{M}^\star\boldsymbol{u}^\star}{\lambda^\star} = \boldsymbol{u}^\star$$

- first approximation is much tighter than the second approximation

Application: exact recovery in community detection

Exact recovery of SBM

We consider the case when (why?)

$$p = \frac{\alpha \log n}{n}, \quad \text{and} \quad q = \frac{\beta \log n}{n}$$

Theorem 6.5

Fix any constant $\varepsilon > 0$. Suppose $\alpha > \beta > 0$ are sufficiently large, and*

$$\left(\sqrt{\alpha} - \sqrt{\beta}\right)^2 \geq 2\left(1 + \varepsilon\right). \tag{6.4}$$

With probability $1 - o(1)$, spectral clustering achieves exact recovery.

Optimality of spectral method

When

$$\left(\sqrt{\alpha} - \sqrt{\beta}\right)^2 \leq 2\,(1+\varepsilon)\,,$$

no method whatsoever can achieve exact recovery
Hellinger distance?

Fine-grained analysis of spectral clustering

Consider "ground-truth" matrix

$$\boldsymbol{M}^\star := \mathbb{E}[\boldsymbol{A}] - \frac{p+q}{2}\mathbf{1}\mathbf{1}^\top = \frac{p-q}{2}\left[\begin{array}{c}\mathbf{1}\\ -\mathbf{1}\end{array}\right]\left[\begin{array}{cc}\mathbf{1}^\top & -\mathbf{1}^\top\end{array}\right],$$

which obeys

$$\lambda_1(\boldsymbol{M}^\star) := \frac{(p-q)n}{2}, \quad \text{and} \quad \boldsymbol{u}^\star := \frac{1}{\sqrt{n}}\left[\begin{array}{c}\mathbf{1}_{n/2}\\ -\mathbf{1}_{n/2}\end{array}\right].$$

These imply

$$\begin{aligned}\lambda^\star &= \tfrac{n(p-q)}{2}\\ \mu &= 1\\ B &= 1\\ \sigma^2 &\leq \max\{p,q\} = p\end{aligned}$$

Invoke ℓ_∞ perturbation theory

ℓ_∞ perturbation bound (6.3b) yields

$$
\begin{aligned}
\left\| z\lambda^\star \boldsymbol{u} - \boldsymbol{M}\boldsymbol{u}^\star \right\|_\infty &\lesssim \sigma + \frac{\sigma^2 \sqrt{n \log n}}{\lambda^\star} + \frac{\sigma B \, \log^{3/2} n}{\lambda^\star} \\
&\leq C\Big(\sqrt{p} + \frac{p\sqrt{\log n}}{\sqrt{n}(p-q)} + \frac{\sqrt{p}\log^{3/2} n}{n(p-q)} \Big) =: \Delta
\end{aligned}
$$

for some constant $C > 0$

it boils down to controlling the entrywise behavior of $\boldsymbol{M}\boldsymbol{u}^\star$

Bounding entries in $\boldsymbol{M}\boldsymbol{u}^\star$

Lemma 6.6

Suppose that

$$\left(\sqrt{p}-\sqrt{q}\right)^2 \geq (1+\varepsilon)\,\frac{2\log n}{n} \tag{6.5}$$

for some quantity $\varepsilon > 0$*. Then with probability exceeding* $1-o(1)$*, one has*

$$\boldsymbol{M}_{l,\cdot}\boldsymbol{u}^\star \geq \frac{2\eta}{a-b} \;\; \text{for all } l \leq \frac{n}{2} \quad \text{and} \quad \boldsymbol{M}_{l,\cdot}\boldsymbol{u}^\star \leq -\frac{2\eta}{a-b} \;\; \text{for all } l > \frac{n}{2},$$

where $\eta > 0$ *obeys* $(\sqrt{a}-\sqrt{b})^2 - \eta\log(a/b) > 2$.

Key message: entries in $\boldsymbol{M}\boldsymbol{u}^\star$ are bounded away from 0 with correct sign

Completing the picture

On one hand

$$\boldsymbol{M}_{l,\cdot}\boldsymbol{u}^\star \geq \varepsilon_0 \text{ for all } l \leq \frac{n}{2} \quad \text{and} \quad \boldsymbol{M}_{l,\cdot}\boldsymbol{u}^\star \leq -\varepsilon_0 \text{ for all } l > \frac{n}{2}.$$

On the other hand

$$\left\| z\lambda^\star \boldsymbol{u} - \boldsymbol{M}\boldsymbol{u}^\star \right\|_\infty \leq \Delta$$

In sum, if one can show

$$\varepsilon_0 > \Delta$$

then it follows that

$$z u_l u_l^\star > 0 \quad \text{for all } 1 \leq l \leq n \quad \Longrightarrow \quad \text{exact recovery}$$

Application: entrywise error in matrix completion