Semi-Random Models

def: A monotone adversary is given $G \sim SBM$ and can modify it:

(1) add edge (u,v) btwn nodes in same community

(2) delete edge (u,v) btwn nodes in diff. communities

Seems helpful, but can it hurt?

Recall counting common neighbors $(p=\frac{1}{2}, q=\frac{1}{4})$

same $(p^2)(\frac{n}{2}) + \cancel{q^2(\frac{n}{2})} = \frac{5n}{32}$

diff $pqn = \frac{4n}{32}$ delete

after monotone changes, some pairs from same comm. will have fewer common neigh. than pairs from diff. comm.

Following [Feige, Killian] will give robust algorithm via SDPs

SemiDefinite Programming

Primal $\min \langle C, X \rangle \triangleq \sum_{ij} C_{ij} X_{ij}$

s.t. $\langle A_i, X \rangle = b_i \quad \forall i = 1 \text{ to } m$

$X \succeq 0$

Can think of feasible region as

$$X \succeq 0 \iff \forall a,\ a^T X a \geq 0$$

infinite set of linear constraints

Prop: Can be efficiently solved via the Ellipsoid algorithm

Caveat: only get optimal value $\pm\varepsilon$, cannot "jump" to optimal value as in LPs

Duality

The dual is

$$\max\ b^T y$$

$$\text{s.t. } \sum y_i A_i + S = C$$

$$S \succeq 0$$

Lemma [Weak Duality] If X & y are feasible then

$$b^T y \leq \langle C, X \rangle$$

Proof: Since y is feasible, we have

$$\langle C, X \rangle = \langle \sum y_i A_i, X \rangle + \langle S, X \rangle$$

① ②

By feasibility of X, we have

$$① = \sum y_i \langle A_i, X \rangle = b^T y$$

And because $S, X \succeq 0$ we have

$$S = YY^T = \sum_i y_i y_i^T$$

Then we have

$$② = \langle S, X \rangle = \text{Trace}(S^T X)$$

$$= \text{Trace}(Y^T X Y)$$

$$= \sum_i y_i^T X y_i \geq 0 \quad \boxtimes$$

<u>Back to Community Detection</u>

We will solve the following relaxation for minimum bisection

$$\min \frac{|E|}{2} - \sum_{(u,v) \in E} \frac{X_{u,v}}{2}$$

$$\text{s.t.} \quad \sum_{u,v} X_{u,v} = 0$$

$$X_{u,u} = 1 \quad \forall u$$

$$X \succeq 0$$

Notice if $(U, V\setminus U)$ is a bisection we can set $s \in \{\pm 1\}^n$ as

$$s_u = \begin{cases} +1 & \text{if } u \in U \\ -1 & \text{else} \end{cases}$$

and $X = ss^T$.

It is easy to check X is feasible and achieves obj. value $|E(U, V\setminus U)|$

Thm If $(p-q)n \geq C\sqrt{pn\log n}$ then whp over $G \sim SBM$ the value of the SDP is equal to size of its minimum bisection

We will pass to the dual to show there's no better soln than the planted one

$$\max \frac{|E|}{2} \sum_u \frac{y_u}{4}$$

$$\text{s.t. } \quad M \triangleq \underset{\substack{\uparrow \\ \text{adjacency} \\ \text{matrix}}}{-A} - y_0 J - \underset{\substack{\uparrow \\ \text{diagonal} \\ \text{matrix of } y_u\text{'s}}}{Y} \succeq 0$$

Let's guess the solution - can pretend we know the planted bisection

Let O_u = # neighbors of u of diff. community

S_u = # " " same community

Now we set $y_0 = 1$, $y_u = O_u - S_u$

Claim: For this choice of y_u's we have

$$\frac{|E|}{2} + \frac{\sum y_u}{4} = |E(U, V \setminus U)|$$

What remains is to show $M = -A + J - Y \succeq 0$

From our discussion last time, we expect

$$\lambda_1(A) \sim \frac{p+q}{2} n$$

$$\lambda_2(A) \leq 2\sqrt{pn} + (pn)^{1/4 + o(1)}$$

Moreover $\lambda_1(A)$ should be covered by J

And the diagonal entries of Y should behave like

$$\sim \mathrm{Bin}(\tfrac{q}{2}, n) - \mathrm{Bin}(\tfrac{p}{2}, n)$$

Thus we expect

$$\lambda_{min}(-A + J - Y) \gtrsim -\lambda_2(A) - \lambda_{min}(Y)$$

$$\gtrsim \left(\frac{p-q}{2}\right) n \pm c\sqrt{pn}$$

Robustness

Now let's prove it's robust

def. A function h is robust wrt monotone changes if

$$h(G) = b(G) \Rightarrow h(H) = b(H)$$

minimum bisection

obtained from G via monotone changes

Lemma Suppose h satisfies the following conditions

1. $\forall G,\ h(G) \le b(G)$

2. If $H = G + e$ then

$$h(G) \le h(H) \le h(G) + 1$$

then h is robust wrt monotone changes

Proof: Suppose $h(G) = b(G)$ and H is obtained by adding an edge within a community

then $h(G) \overset{2}{\le} h(H) \overset{1}{\le} b(H) = b(G)$

And so $h(H) = b(H)$

Suppose instead H is obtained by deleting an edge btwn communities. Then

$$h(G)-1 \overset{②}{\leq} h(H) \overset{①}{\leq} b(H) = b(G)-1$$

notice G and H
are swapped b/c
we added an edge
to H to form G

And so $h(H) = b(H)$ again. ☒

Now let's verify the SDP satisfies the conditions of the lemma:

① $h(G) \leq b(G)$ because it is a relaxation

② Suppose $H = G + \{u,v\}$. Then

$$Obj' = Obj + \frac{1}{2} - \frac{X_{u,v}}{2}$$

But because $X \succeq 0$ and $X_{u,u} = X_{v,v} = 1$ we must have $|X_{u,v}| \leq 1$

otherwise the 2x2 subdeterminant would be negative

Thus for any feasible X, the objective value can increase by at most one

So we have

Thm [Feige, Kilian] If $(p-q)n \geq C\sqrt{pn\log n}$ the SDP recovers the value of the exact minimum bisection whp in semirandom model