EE290 Course Notes

Feynman Liang*
Department of Statistics, UC Berkeley

Last updated: September 21, 2019

Contents

1 9/5/2019

1.1 Results from random matrix theory

Today we consider random matrices $Z = (Z_{ij}) \in \mathbb{R}^{n \times n}$. IID matrix ensemble is when $Z_{ij} \sim P$ are drawn IID, and the Gaussian Orthogonal Ensemble (GOE) has $Z_{ii} \sim N(0,2)$ and $Z_{ij} = Z_{ji} \sim N(0,1)$ for $i \neq j$.

By convention, normalize and center so $\mathbb{E}Z_{ij} = 0$ and $\mathbb{E}Z_{ij}^2 = 1$.

Intuition: $\|Z\|_{op} \leq C\sqrt{n}$ with high probability.

Consider Gaussian orthogonal ensemble matrix: $Z_{ij} \sim N(0,1)$ and $Z_{ii} \sim N(0,2)$. View $Z = \begin{bmatrix} Z_1, \dots, Z_n \end{bmatrix}$ with $Z_i \sim N(0, I_n)$. Then

$$\mathbb{E}\|Z_1\|_2^2 = \mathbb{E}[\sum_{i=1}^n Z_{i1}^2] = n \tag{1}$$

$$Z_1^\top Z_2 = \sum_{i=1}^n Z_{i1} Z_{i2} \tag{2}$$

$$\mathbb{E}Z_1^\top Z_2 = 0 \tag{3}$$

$$\mathbb{E}(Z_1^\top Z_2)^2 = n \tag{4}$$

$$|Z_1^\top Z_2| \sim \sqrt{n} \tag{5}$$

$$\frac{Z_1^\top Z_2}{\|Z_1\|\|Z_2\|} \sim \frac{1}{\sqrt{n}} \tag{6}$$

Theorem 1 (*Latała et al. (2006)*)

$$\sup_i \sum_{j=1}^n \mathbb{E}|Z_{ij}|^2 \leq k^2 n \tag{7}$$

$$\sup_j \sum_{i=1}^n \mathbb{E}|Z_{ij}|^2 \leq k^2 n \tag{8}$$

*feynman@berkeley.edu

Fourth moment bound

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbb{E}|Z_{ij}|^4 \leq k^4 n^2 \tag{9}$$

Then $\mathbb{E}\|Z\|_{op} = O(k\sqrt{n})$

1.2 Gaussian Orthogonal Ensemble

$\|Z\|_{op} = \sigma_{max} = \max_{\|v\|=1} v^\top Zv$

For any fixed $v \in S^{n-1}$, we have a Gaussian tail bound

$$v^\top Zv = \sum_i Z_{ii}v_i + \sum_{i<j} 2Z_{ij}v_iv_j \tag{10}$$

$$= N(0, \sum_i v_i^4 + \sum_{i<j} 4v_i^2v_j^2) \tag{11}$$

$$\Pr(|v^\top Zv| > t) \leq 2e^{-t^2/4} \tag{12}$$

Using an ϵ-net, can find a set of vectors V_ϵ such that

$$\max_{v\in V_\epsilon}|v^\top Zv| \geq (1-2\epsilon)\max_{|v|=1}|z^\top Zv| \geq (1-2\epsilon)t \tag{13}$$

Then by a union bound

$$\Pr[\|Z\|_{op} \geq t] \leq \Pr[\max_{v\in V_\epsilon}|v^\top Zv| \geq (1-2\epsilon)t] \tag{14}$$

$$\leq \sum_{v\in V_\epsilon}\Pr[|v^\top Zv| \geq (1-2\epsilon)t] \tag{15}$$

$$\leq 2|V|e^{-\frac{1}{4}(1-2\epsilon)^2t^2} \leq \delta \tag{16}$$

If $|V| \leq c^n$, then

$$e^{c(n-ct^2)} \leq e^{\log\delta} \tag{17}$$

$$\log\frac{1}{\delta} \leq ct^2 - n \implies t \geq \sqrt{n + \log\frac{1}{\delta}} \tag{18}$$

Intuition: when dealing with infinite dimensional maximization (Rayleigh quotient for eigenvalue problem), can pass to ϵ-net for cardinality bboud.

Definition 2 (*Covering*)

$V \subset S^{n-1}$ is called an ϵ-net if $\forall u \in S^{n-1}$, $\exists v \in V$ such that $\|u - v\|_2 \leq \epsilon$.

Theorem 3

ϵ-net yields Eq. (13)

Definition 4 (*Packing*)

For $A \subset \mathbb{R}^d$, $V = \{v_i\}_{i=1}^n \subset A$ is an ϵ-packing if $\forall i \neq jJ$, $\|v_i - v_j\|_2 \geq \epsilon$.

Theorem 5

Maximal ϵ-packing is an ϵ-net.

Hence, we can lower bound the packing number (size of largest packing) by the covering number (size of the smallest covering). The following result gives an (obvious?) upper bound:

Lemma 6 (*Volume ratio*)

For any ϵ-packing $V \subset A$,

$$|V| \leq \frac{Vol(A + \frac{\epsilon}{2}B)}{Vol(\frac{\epsilon}{2}B)} \tag{19}$$

where $B = \{x : \|x\|_2 \leq 1\}$.

Why is the diagonal not important? Let $A = \mathrm{diag}(Z)$. Then we have

$$\|Z - A\|_{op} \leq \|Z\|_{op} + \|A\|_{op} \tag{20}$$

$$\max_{x \in S^{n-1}} \|Ax\| = \max_i |Z_{ii}| = O(\sqrt{2 \log n}) \tag{21}$$

So the diagonal term $\|A\|_{op}$ is an order of magnitude smaller that $\|Z\|_{op}$.

Example 7 (*Planted clique*)

Let $G \sim G(1/2, n, k)$. In other words, generate an Erdös-Renyi random graph from $G(n, 1/2)$ and then randomly choose a set $K \subset [n]$ connect together to form a clique.

Goal: find K given G.

Theorem 8 (*Alon et al. (1998)*)

For any c, $k = c\sqrt{n}$, then exists polytime algorithm such that it returns $\hat{K}$ with $P(\hat{K} = K) \to 1$.

Let the adjacency matrix $A_{ij} = \begin{cases} 1 & (i,j) \in K \\ \mathrm{Bern}(1/2) & i \notin K \text{ or } j \notin K, i \neq j \\ 0 & i = j \end{cases}$ and define $W_{ij} = \begin{cases} 2A_{ij} - 1 & i \neq j \\ 0 & i = j \end{cases}$

1. Find top eigenvector u of W
2. Let $\tilde{K}$ index the k largest coordinates $|u_i|$
3. Thresholding

$$\hat{K} = \left\{ v \in [n] : d_{\tilde{K}}(v) \geq \frac{3k}{4} \right\} \tag{22}$$

$$d_{\tilde{K}}(v) = \sum_{j \in \tilde{K}} \mathbb{1}\{(j,v) \text{ connected}\} \tag{23}$$

Goal: show $|\tilde{K} \cap K| \geq (1-\epsilon)k$ whp.

Note that $\mathbb{E}[W] =: 1_k 1_k^\top - \mathrm{diag}(1_k)$ consists of 1s in $K \times K$ and 0 everywhere else. Let

$$W^* = 1_k 1_k^\top \tag{24}$$

$$v = \frac{1}{\sqrt{k}} 1_k \tag{25}$$

$$\tag{26}$$

Notice thresholding over v exactly recovers K, so we want the top eigenvector u of W to be close to v. By Davis-Kahan,

$$\min_{s \in \{\pm 1\}} \|u + sv\|_2 \leq \frac{\|W - W^*\|_{op}}{\lambda_1(W^*) - \lambda_2(W^*)} \tag{27}$$

Note $\lambda_1(W^*) = k$. Suppose extrema attained at $s = -1$, then

$$\|W - W^*\|_{op} \leq \|W - \mathbb{E}W\| + \underbrace{\|\mathbb{E}W - W^*\|}_{=\|\,\mathrm{diag}\, 1_k\|=1} \leq c\sqrt{n} + 1 \tag{28}$$

By Weyl's inequality

$$|\lambda_2(W)| = |\lambda_2(W^*) - \lambda_2(W)| \le \|W^* - W\|_{op} \le c\sqrt{n} + 1 \tag{29}$$

Finally

$$\|u - v\|_2 \le \frac{c\sqrt{n} + 1}{c\sqrt{n} - (c\sqrt{n} + 1)} \le \epsilon \tag{30}$$

NOTE: when you have bounded fourth moments, the rate is always $n^{-1/2}$! Deep result.

2 9/10/2019

Recall the planted clique from Alon et al. (1998): $G \sim G(1/2, n, k)$ is a random graph on $V = [n]$ with some fully connected clique $K \subset [n]$ of cardinality $|K| = k$.

The adjacency matrix

$$A_{ij} = \begin{cases} 1 & \text{if } i, j \in K \\ \text{Bern}(1/2) & i \neq j \text{ ow} \end{cases} \tag{31}$$

Let

$$W_{ij} = \begin{cases} 2A_{ij} & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases} \tag{32}$$

Algorithm 1 of Alon et al. (1998):

1. Find top eigenvector of W, say u
2. Let $\tilde{K}$ index the largest k coordinates $|u_i|$
3. Define $\hat{K} = \{v \in V : d_{\tilde{K}}(v) \ge \frac{3k}{4}\}$

Theorem 9 (*Alon et al. (1998)*)

Algorithm 1 finds $\hat{K}$ such that $\Pr[\hat{K} = K] \to 1$ as $n \to \infty$ if $k \ge c\sqrt{n}$ for sufficiently large c.

Proof. Note that $\mathbb{E}A$ is:

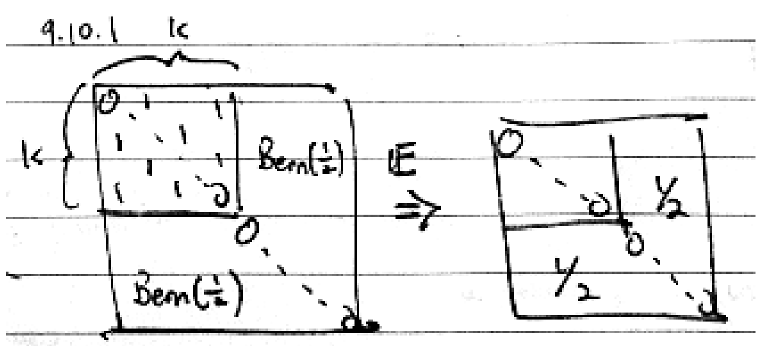

Figure 1: $\mathbb{E}A$ has ones in the upper $k \times k$ block, 0 on the diagonal, and 1/2 everywhere else

From this, we can easily see that the $\mathbb{E}W$ is:

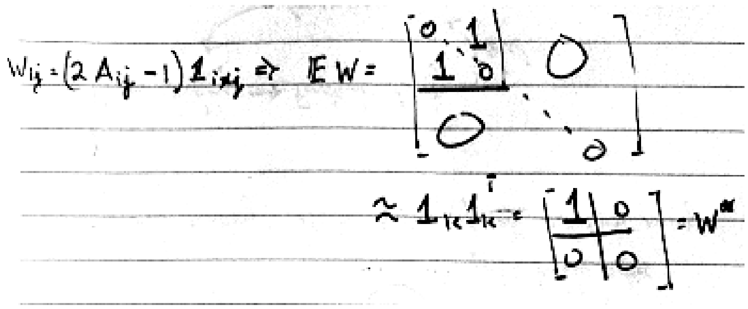

Figure 2: $\mathbb{E}W$ differs from $W^* = 1_k 1_k^\top$ only in the upper k diagonal

Note $\mathbb{E}W = 1_K 1_K^\top - \text{diag}(1_K) \approx 1_K 1_K^\top = W^*$, which is good because we have seen that "differenes in the diagonal are asymptotically negligible."

reference for this? 9-5 lecture

Goal: show $|\tilde{K} \cap K| \geq (1-\varepsilon)k$ whp, $\varepsilon = \varepsilon(c)$.

We first show the top eigenvector of W^* is close to u (the top eigenvector of W). Let $v = \frac{1}{\sqrt{k}} 1_K$ be the top eigenvector of W^*. Note $\lambda_1(W^*) = k$. By Davis-Kahan

$$\min_{s \in \{\pm 1\}} \|u + sv\|_2 \leq \frac{\|W - W^*\|_2}{\lambda_1(w^*) - \lambda_2(w)} \tag{33}$$

Note

$$\|W - W^*\| \leq \|W - \mathbb{E}W\| + \|\mathbb{E}W - W^*\| \leq c\sqrt{n} + 1 \tag{34}$$

Also $\lambda_1(W^*) = k$ and

$$|\lambda_2(W)| \leq |\lambda_2(W^*) - \lambda_2(W) \leq \|W^* - W\| \tag{35}$$

So by Weyl's inequality

$$\min_{s \in \{\pm 1\}} \|u + sv\|_2 \leq \frac{c\sqrt{n} + 1}{k - (c\sqrt{n} + 1)} \tag{36}$$

$$\leq \frac{c\sqrt{n} + 1}{c\sqrt{n} - c\sqrt{n} + 1} \leq \varepsilon \tag{37}$$

Aside: Davis-Kahan to get bound between difference of eigenvectors in 2-norm. Open problem to control others.

Next, if $|K| = k = |\tilde{K}|$ then $|K \setminus \tilde{K}| = |\tilde{K} \setminus K|$.

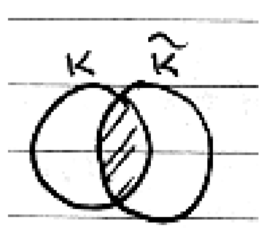

Figure 3: $|K| = |\tilde{K}| \implies |K \setminus \tilde{K}| = |\tilde{K} \setminus K|$ follows from elementary set theory

By definition of v

$$\varepsilon^2 \geq \|u - v\|_2^2 = \sum_{i \in K} (u_i - \frac{1}{\sqrt{k}})^2 + \sum_{i \notin K} u_i^2 \tag{38}$$

Lemma 10

If all $|u_i| \le \frac{1}{2\sqrt{k}}$ for $i \notin \tilde{K}$, then

$$\varepsilon^2 \ge \sum_{i\in K\setminus\tilde{K}} (\frac{1}{\sqrt{k}} - u_i)^2 \ge \sum_{i\in K\setminus\tilde{K}} \frac{1}{4k} \tag{39}$$

This implies $|K \setminus \tilde{K}| \le 4\varepsilon^2 k$.

Lemma 11

If the condition of the previous lemma does not hold, then $\exists i \in \tilde{K}$ with $|u_i| \ge \frac{1}{2\sqrt{k}}$. Then in fact $|u_i| \ge \frac{1}{2\sqrt{k}}$ for all $i \in \tilde{K}$ since

$$\varepsilon^2 \ge \sum_{i\in\tilde{K}\setminus K} u_i^2 \ge \sum_{i\in\tilde{K}\setminus K} (\frac{1}{2\sqrt{k}})^2 = \sum_{i\in\tilde{K}\setminus K} \frac{1}{4k} \tag{40}$$

Hence $|\tilde{K} \setminus K| \le 4\varepsilon^2 k$

So we have achieved our goal.

To finish the proof, first assume $\|u - v\|_2 \le \varepsilon$. For $a \in K$,

$$d_{\tilde{K}}(a) \ge d_{\tilde{K}\cap K}(a) = |\tilde{K} \cap K| - 1 \ge (1 - \varepsilon')k \tag{41}$$

so for $a \in K$, we will get $a \in \hat{K}$.

Now if $a \notin K$,

$$d_{\tilde{K}}(a) \le \underbrace{d_K(a)}_{\sim\text{Binom}(k,1/2)} + \underbrace{|\tilde{K} \setminus K|}_{\le\varepsilon' k} \approx \frac{k}{2} \pm c\sqrt{k} \tag{42}$$

where $\approx$ means concentration. To be concrete,

$$\begin{aligned}
\Pr[\hat{K} \ne K] &\le \Pr[\|u - v\|_2 \ge t] + \Pr[\exists a \notin K : d_K(a) \ge (\frac{3}{4} - \varepsilon')k] && (43)\\
&\le \Pr[\|W - \mathbb{E}W\| \ge c\sqrt{n}] + (n - k)\Pr[B(k, 1/2) \ge (\frac{3}{4} - \varepsilon)k] && (44)\\
&\le ce^{-c'n} + (n - k) && (45)
\end{aligned}$$

Where above we used the multiplicative version of Chernoff bound (useful in combinatorial statistics):

Lemma 12 (*Multiplicative Chernoff Bound*) {lem:mult-che

$$\Pr[X \ge (1+\delta)\mu] \le \begin{cases} e^{-\delta^2\mu/3} & \delta \in [0,1] \\ e^{-\delta\mu/3} & \delta \ge 1 \end{cases} \tag{46}$$

$$\Pr[X \le (1-\delta)\mu] \le e^{-\delta^2\mu/2} \tag{47}$$

As $n \to \infty$, we see that $\Pr[\hat{K} = K] \to 1$. □

Lemma 12 is self-normalizing: let $X = \sum_{i=1}^n X_i$ with X_i independent binary and $\mu = \mathbb{E}X$. Note that after applying, the RHS does not depend on n Verify

AKS Algorithm 2: This algorithm is designed to handle the case when k is not big enough (recall algorithm 1 requires $k \ge c\sqrt{n}$). Search over all S with $|S| = C(c) = 2\log_2 \frac{10}{c} + 2$. For each S:

1. Define $N^*(S) = \{v \in V : v \sim a, \forall a \in S\} \setminus S$
2. Run Algorithm 1 on the induced subgraph (which has distribution $G(1/2, N^*(S), K - S)$), return $Q_S \cup S$
3. Output if $Q_S \cup S$ is a k-clique

Intuition: Suppose $k = 0$ so there's no clique. Then $|N^*(S)| \sim B(n - s, 2^{-s}) \approx \frac{n-s}{2^s}$ so the total number of nodes is much smaller (by order of 2^{-s}). However, the number of clique nodes in $N^*(S)$ is still relatively large, $\geq k - s$. Solving the critical equation (also for algorithm 1) Track htis down

$$k - s \geq C\sqrt{\frac{n}{2^s}} \tag{48}$$

yields the expression for $C(c)$.

Theorem 13

As long as $k \geq (2 + \varepsilon)\log_2 n$, then exhaustive search finds k with probability $\to 1$.

Proof. Exhaustive search will always find the clique, but it may return a clique that we didn't plant. So we need to guarantee there is no clique of size $(2 + \varepsilon)\log_2 n$ in G whp.

For $S \subset [n]$, $|S| = k$,

$$\Pr[S \text{ is clique}] = \frac{1}{2^{\binom{k}{2}}} \tag{49}$$

$$\Pr[\exists S \subset [n] : S \text{ is clique}] \leq \binom{n}{k}\frac{1}{2^{\binom{k}{2}}} \leq (n2^{-(k-1)/2})^k \to 0 \tag{50}$$

$$\tag{51}$$

as $n \to \infty$ ($k = (2 + \varepsilon)\log_2 n$). □

3 9/12/2019

3.1 Planted cliques and semidefinite programming

Recall the matrix W from before, which has 1s in the top $k \times k$ block, zero on the diagonal, and Rad(1/2) RVs elsewhere.

Recall the spectral method:

$$\hat{u}_{spec} = \operatorname{argmax}_{\substack{u \in \mathbb{R}^n \\ \|u\|^2 = k}} u^\top W u \tag{52}$$

This needs a cleaning step, which we analyzed previously.

How did they come up with this algorithm? Can we get more insight by analyzing htis method in a more principled framework? Yes, through maximum likelihood!

Consider an alterantive model where within clique we have connection probability p (instead of 1) and other connections with probability q (instead of 1/2), where $p \gg q$.

$$\hat{u}_{MLE} = \operatorname{argmax}_{\substack{u \in \{0,1\}^n \\ \sum_i u_i = k}} u^\top W u \tag{53}$$

From this, we see that the spectral method is a continuous relaxation of the MLE integer program. To make this more precise, consider the SDP

$$\hat{X}_{spec} = \operatorname{argmax}_{\substack{X \succeq 0 \\ \operatorname{Tr} X = k}} \langle W, X \rangle \tag{54}$$

If we let $X = uu^\top$, then we automatically have $X \succeq 0$ and additionally we have $\operatorname{Tr} X = \|u\|_2^2$. Thus, the feasible set of Eq. (52) is the same as Eq. (54).

How do we know the optima of Eq. (54) is attained at a rank 1 matrix $X = uu^\top$? Since $X = \sum_i \lambda_i u_i u_i^\top$ ($\lambda_i \geq 0$) and optima are attained at extremal points, by linearity of $\langle W, X\rangle$ we can put all of the weight on a single λ_i corresponding to the top eigenvector of W.

How can we get Eq. (54) closer to Eq. (53)? Since Eq. (53) is more constrained, we can consider adding more constraints:

$$\begin{aligned}
\tilde{X}_{MLE} &= \operatorname{argmax}_X \langle W, X\rangle && (55)\\
\text{s.t. } & X \succeq 0 && (56)\\
& \operatorname{Tr} X = k && (57)\\
& 0 \leq X \leq J \quad \text{entrywise} && (58)\\
& \langle X, J\rangle = k^2 && (59)\\
& \operatorname{rank}(X) = 1 && (60)
\end{aligned}$$

where $J = 11^\top$.

The solution $X = uu^\top$ where $u \in \{0,1\}^n$, where u indexes the clique.

Conversely, we need to show that the feasible set coincides with Eq. (53). If $X \succeq 0$ and rank $X = 1$, then we can always write $X = uu^\top$. The trace constraint now reads $k = \operatorname{Tr} X = \sum_i u_i^2$. The third constraint becomes $\langle X, J\rangle = k^2 \implies (\sum_i u_i)^2 = k^2$.

Proposition 14

The optima of Eq. (55) must satisfy: $u_i \in [-1,1]$, $\sum u_i^2 = k$, $(\sum_i u_i)^2 = k^2$, $\{u_i\} \in \{0,1\}^n$ *or* $\{u_i\} \in \{0,-1\}^n$.

In fact, the solution is $u = 1_k$ *or* $u = -1_k$.

The linear constraints in Eq. (55) are fine, but the rank constraints are difficult. Here is an easier candidate SDP:

$$\begin{aligned}
\hat{X}_{SDP} &= \operatorname{argmax}_X \langle W, X\rangle && (61)\\
\text{s.t. } & X \preceq 0 && (62)\\
& X \geq 0 && (63)\\
& \operatorname{Tr} X = k && (64)\\
& \langle X, J\rangle = k^2 && (65)
\end{aligned}$$

Notice we have dropped the rank constraint as well as the upper entrywise bound.

Theorem 15

$\exists c > 0$ *such that for* $k \geq c\sqrt{n}$*, Eq. (61) has unique maximizer* $X^* = 1_k 1_k^\top$ *with high probability.*

Proof. We first show X^* is a maximizer.

$$\langle W, X^*\rangle = 1_k^\top W 1_k = k^2 - k \tag{66}$$

$$\langle W, X\rangle = \langle W + I, X\rangle - \operatorname{Tr} X \tag{67}$$

$$\operatorname{Tr}(I - X) = \operatorname{Tr} X \leq \langle J, X\rangle - \operatorname{Tr}(X) \tag{68}$$

$$\underbrace{W + I \leq J}_{X \geq 0} \implies \langle J, X\rangle \geq \langle W + I, X\rangle \tag{69}$$

$$\therefore \operatorname{Tr}(I - X) = \operatorname{Tr} X \leq k^2 - k \tag{70}$$

The harder part is uniqueness. We will develop a general technique called dual certificate / KKT condition. Write the Lagrangian for the optimization problem. Introduce dual variables $S \succeq 0$, $B \geq 0$, $\eta \in \mathbb{R}$, $\lambda \in \mathbb{R}$ and

$$\mathcal{L}(X, S, B, \eta, \lambda) = \langle W, X\rangle + \langle S, X\rangle + \langle B, X\rangle + \eta\left(k \operatorname{Tr}(X) + \lambda(k^2 - \langle X, J\rangle\right) \tag{71}$$

Notice

$$\max_{X \text{ feas}} \langle W, X\rangle = \max_X \min_{S,B,\eta,\lambda} \mathcal{L} \tag{72}$$

as desired. Since $\mathcal{L}$ is linear, by Sion's minimax theorem we have

$$\max_X \min_{S,B,\eta,\lambda} \mathcal{L} = \min_{S,B,\eta,\lambda} \max_X \mathcal{L} \tag{73}$$

Note $\langle S, X\rangle = \operatorname{Tr}(S^{1/2}XS^{1/2}) \geq 0$ is non-negative. $\langle B, X\rangle$ is also trivially non-negative.

Lemma 16

The following conditions imply X^* is the unique maximizer: {lem:x-star-u

1. *Stationarity: $W + S + B - \eta I - \lambda J = 0$ (can't improve any more)*
2. *Primal/dual feasibility*
3. *Complementary slackness: $\langle S, X^*\rangle = 0$ and $\langle B, X^*\rangle = 0$.*
4. *Uniqueness: $\lambda_{n-1}(S) > 0$ (second smallest eigenvalue of S)*

The first three conditions are the "KKT conditions." Together, they guarantee X is a maximizer.

Proof of Lemma 16. X^* **is a maximizer**: for feasible variables

$$\begin{aligned}
\langle W, X\rangle &\leq \mathcal{L}(X, S, B, \eta, \lambda) && \text{feasible} && (74)\\
&= \mathcal{L}(X^*, S, B, \eta, \lambda) && \text{stationarity} && (75)\\
&= \langle W, X^*\rangle && \text{comp. slackness} && (76)
\end{aligned}$$

Uniqueness: Suppose X' satisfies $\langle W, X'\rangle = \langle W, X^*\rangle$. Then $\langle S, X'\rangle = 0$, and $\langle S, X^*\rangle = 0 \implies 1_k^\top S 1_k = 0 \implies S1_k = 0$. In other words, 1_k is an eignevector with eigenvalue 0 for S. But condition (4) means that 1_k is the only eigenvector with eigenvalue 0, hence $X' = cX^*$ for some $c \in \mathbb{R}$. But by the constrant $\operatorname{Tr} X = k$, we must have $X' = X^*$. □

Hence, if we can find (S, B, η, λ) satisfying Lemma 16, then we have a certificate that X^* is the unique maximizer.

But how can we find this certificate? It's hard in general, but in this case we have an explicit construction.

$$B \geq 0, \quad \eta \in \mathbb{R}, \quad \lambda \in \mathbb{R} \tag{77}$$

$$S = \eta I + \lambda J - B - W \succeq 0 \tag{78}$$

$$S1_k = 0, \quad \langle B, X^*\rangle = 0, \quad \lambda_{n-1}(S) > 0 \tag{79}$$

$$S1_k = 0 \implies \eta I_k + \lambda k 1 = B1_k + W1_k \tag{80}$$

$X^* = 1_k 1_k^\top$. Since we want $\langle B, X^*\rangle = 0$, we want $B_{ij} = 0$ for $(i, j) \in K \times K$. This implies that $(B1_k)i = 0$ for $i \in K$. Let $y = W1_k$.

ith entry, $i \in K$, of Eq. (79) implies $\eta + k\lambda = (B1_k)_i + y_i = k - 1$. Then, choose $\eta = k - 1 - k\lambda$

Now for $i \notin K$, Eq. (79) implies $\lambda k = (B1_k)_i + y_i$. Construct $B = 1_k b^\top + b1_k^\top$ for some $b \in \mathbb{R}^n$ such that $b_i = 0$ for $i \in K$. Then $B1_k = kb$.

Fig 9.12.1

$b_i = \lambda - \frac{y_i}{k}$ for all $i \notin k$. Check $B \geq 0 \implies b_i \geq 0$. Since $\lambda \geq \frac{y_i}{k}$ for all $i \in K$, $\lambda \geq \max_{i \notin K} \frac{y_i}{k}$. $y_i = W1_k$ which is a sum of Rad(1/2) RVs, so by concentration for some $\lambda \geq c$ this is satisfied whp.

For the last part, we need to show $x^\top S x > 0$ for all x such that $x^\top 1_k = 0$. The exact formula for S is

$$
\begin{aligned}
S &= \eta + \underbrace{\lambda x^\top J x}_{\geq O(\sqrt{n})} - \underbrace{x^\top B x}_{=0} - \underbrace{x^\top W x}_{\geq O(\sqrt{n})} && (81)\\
&\geq \frac{k}{2} - 1 - x^\top \mathbb{E}[W]x - \|W - \mathbb{E}W\|_{op} && (82)\\
&\geq 0 \qquad \text{for suff large } k && (83)
\end{aligned}
$$

□

Bibliography

Alon, N., M. Krivelevich, and B. Sudakov
1998. Finding a large hidden clique in a random graph. *Random Structures & Algorithms*, 13(3-4):457–466.

Latała, R. et al.
2006. Estimates of moments and tails of gaussian chaoses. *The Annals of Probability*, 34(6):2315–2331.