Chapter 11

Spectral Rounding

In the spectral sparsification problem in Chapter 9 and Chapter 10, we are given $v_1, v_2, \ldots, v_m \in \mathbb{R}^n$, and the goal is to find a "reweighting" $s_1, \ldots, s_m$ with few nonzeros so that $\sum_{i=1}^m s_i v_i v_i^T \approx \sum_{i=1}^m v_i v_i^T$.

In this chapter, we consider the following spectral rounding problem where the goal is to find an "integral reweighting" that approximates the input.

Definition 11.1 (Spectral Rounding). *Given $v_1, v_2, \ldots, v_m \in \mathbb{R}^n$ and scalars $x_1, \ldots, x_m \in \mathbb{R}_{\geq 0}$, find integer scalars $z_1, \ldots, z_m \in \mathbb{Z}_{\geq 0}$ such that*

$$\sum_{i=1}^m x_i v_i v_i^T \approx \sum_{i=1}^m z_i v_i v_i^T.$$

More generally, we are also given k linear constraints in a matrix $A \in \mathbb{R}_{\geq 0}^{k \times m}$ and are required to find integer scalars $z_1, \ldots, z_m \in \mathbb{Z}_{\geq 0}$ that also satisfies

$$A\vec{x} \approx A\vec{z}.$$

Note that there is no requirement on the number of nonzeros in $\vec{z}$ as in the spectral sparsification problem, rather the requirement is on the integrality of $\vec{z}$.

The motivation of this problem is from designing approximation algorithms for some discrete optimization problems, where we should think of $\vec{x}$ as an optimal fractional solution to some convex relaxation of a combinatorial problem, and our goal is to find an integer solution $\vec{z}$ that is almost as good as $\vec{x}$. The additional linear constraints can be used to incorporate the objective value of the solutions, and/or some other constraints such as upper and lower bound on the size of the solutions. We will see two concrete applications in the next section.

This problem in its strongest form is as general as the Kadison-Singer problem that we will study in the second part of the course. In this chapter, we consider a simpler setting where the two-sided approximation requirement is replaced by a one-sided approximation requirement. The main result that we will study is by Allen-Zhu, Li, Singh and Wang [ALSW17], who formulated the following minimum eigenvalue problem and used it to design approximation algorithms for experimental design problems.

Theorem 11.2 (Minimum Eigenvalue Problem [ALSW17]). *Given $v_1, \ldots, v_m \in \mathbb{R}^n$ and scalars $x_1, \ldots, x_m \in \mathbb{R}_{\geq 0}$ satisfying*

$$\sum_{i=1}^m x_i v_i v_i^T = I_n \quad \textit{and} \quad \sum_{i=1}^m x_i = k,$$

there is a polynomial time algorithm to find integer scalars $z_1, \ldots, z_m \in \mathbb{Z}_{\geq 0}$ *satisfying*

$$\sum_{i=1}^m z_i v_i v_i^T \succcurlyeq \left(1 - \sqrt{\frac{n-1}{k}}\right)^2 \cdot I_n \quad \textit{and} \quad \sum_{i=1}^m z_i = k.$$

To see its connection to the spectral rounding problem in Definition 11.1, first we apply the same reduction as in Lemma 9.11 to reduce to the case when $\sum_{i=1}^m x_i v_i v_i^T = I_n$. Then the two-sided approximation requirement in Definition 11.1 becomes $(1-\epsilon)I_n \preccurlyeq \sum_{i=1}^m z_i v_i v_i^T \preccurlyeq (1+\epsilon)I_n$ for an ϵ as small as possible. In Theorem 11.2, the two-sided requirement is replaced by the one-sided requirement $\sum_{i=1}^m z_i v_i v_i^T \succcurlyeq (1-\epsilon)I_n$. And there is one linear "cardinality/budget" constraint $\sum_{i=1}^m x_i \approx \sum_{i=1}^m z_i$ to satisfy, without which the problem is trivial.

The proof of Theorem 11.2 in [ALSW17] is based on the regret minimization framework developed for spectral sparsification by Allen-Zhu, Liao, and Orecchia [ALO15]. It will take quite some time to introduce this framework properly and we will not do so, but we will briefly describe their framework at the end of this chapter.

Instead, we will present a new proof of Theorem 11.2, following the (informal) polynomial perspective for spectral sparsification from [BSS14] that we described in the beginning of Chapter 10. I hope this proof serves better as a bridge to connect to the second part of the course, starting next chapter.

11.1 Applications

Before we see the proof, let's first see some applications of Theorem 11.2, which is useful in designing approximation algorithms for choosing a good subset of points/vectors/edges.

Experimental Design

In experimental design problems, we are given vectors $v_1, \ldots, v_m \in \mathbb{R}^n$ and a parameter $k \geq n$, and the goal is to choose a (multi-)subset S of k vectors so that $\sum_{i \in S} v_i v_i^T$ optimizes some objective function. The most popular and well-studied objective functions are:

- D-design: Maximizing $\left(\det\left(\sum_{i\in S} v_i v_i^T\right)\right)^{\frac{1}{n}}$.
- A-design: Minimizing $\operatorname{Tr}\left(\left(\sum_{i\in S} v_i v_i^T\right)^{-1}\right)$.
- E-design: Maximizing $\lambda_{\min}\left(\sum_{i\in S} v_i v_i^T\right)$.

These problems of choosing a representative subset of vectors have a wide range of applications (see [ALSW21, LZ21]), but these are all NP-hard. To design approximation algorithms, we consider the following natural convex programming relaxations for D/A/E-design.

$$\begin{aligned}
\max \quad & \left(\det\left(\sum_{i=1}^n x_i \cdot v_i v_i^T\right)\right)^{\frac{1}{n}} \Big/ \ \min \ \operatorname{Tr}\left(\sum_{i=1}^n x_i \cdot v_i v_i^T\right)^{-1} \Big/ \ \max \ \lambda_{\min}\left(\sum_{i=1}^n x_i \cdot v_i v_i^T\right) \\
\text{s.t.} \quad & \sum_{i=1}^m x_i \leq k. \\
& x_i \geq 0, \qquad \text{for } 1 \leq i \leq n.
\end{aligned}$$

After we computed an approximately optimal solution x in polynomial time, we can apply the transformation as in Lemma 9.11 to reduce to the case where $\sum_{i=1}^m x_i v_i v_i^T = I_n$. Then we can apply Theorem 11.2 to obtain an integral solution z, and then apply the reverse transformation in Lemma 9.11 to see that z has the following performance guarantee.

Problem 11.3 (Experimental Design). *Prove that Theorem 11.2 can be used to obtain a $(1 \pm \epsilon)$-approximation algorithm for D/A/E-design when $k \gtrsim n/\epsilon^2$.*

This approach is used in [ALSW21, LZ21] to provide a unifying algorithmic framework for designing the best known approximation algorithms for a large class of experimental design problems. We will discuss some ideas of these work in the end of this chapter.

Network Design

The general setting of network design is to find a minimum cost subgraph satisfying certain requirements. The most well-studied problem is the survivable network design problem, where the requirement is to have at least a specified number $f_{u,v}$ of edge-disjoint paths between every pair of vertices u, v. Linear programming is the default approach in designing approximation algorithms for network design problems. It is observed in [LZ20] that spectral techniques can also be used for survivable network design problems, as well as to incorporate additional spectral constraints. For example, consider the following convex relaxation:

$$
\begin{aligned}
\min_x \ & \sum_{e \in E} c_e x_e \\
& \sum_{e \in \delta(S)} x_e \geq \max_{u \in S, v \notin S} \{f_{u,v}\} && \forall S \subseteq V && \text{(connectivity constraints)} \\
& \lambda_2(L_x) \geq \lambda && && \text{(algebraic connectivity constraint)} \\
& 0 \leq x_e \leq 1 && \forall e \in E && \text{(capacity constraints)}
\end{aligned}
$$

where c_e is the given cost of an edge $e \in E$, and L_x is the Laplacian matrix where each edge e has weight x_e. The algebraic connectivity constraint can be used to lower bound the edge expansion of the solution.

Exercise 11.4 (Second Laplacian Eigenvalue and Edge Expansion). *Let $G = (V, E)$ be an undirected graph. Prove that*

$$\lambda_2(L_G) \leq 2 \min_{0 \leq |S| \leq |V|/2} \frac{|\delta(S)|}{|S|}.$$

Without the algebraic connectivity constraint, the above is a linear program and there is an elegant iterative rounding 2-approximation algorithm by Jain to solve the problem. With the algebraic connectivity constraint, the above becomes a convex program and it was not known how to handle both connectivity constraints and the algebraic connectivity constraint simultaneously. The observation in [LZ20] is that the one-sided spectral rounding result in Theorem 11.2 can be adapted to design an approximation algorithm for this problem.

Exercise 11.5 (Spectral Rounding for Network Design). *Let $x \in [0, 1]^m$ be a fractional solution to the above convex program and L_x be its Laplacian matrix. Show that if $z \in \{0, 1\}^m$ is an integral solution satisfying $L_z \succcurlyeq L_x$, then z satisfies the connectivity constraints and the algebraic connectivity constraint simultaneously.*

In [LZ20], Theorem 11.2 is extended to find $z \in \{0,1\}^m$ satisfying $L_z \succcurlyeq L_x$ with

$$\sum_{e \in E} c_e z_e \leq (1 + O(\epsilon)) \cdot \sum_{e \in E} c_e x_e + O\left(\frac{n \cdot c_{\max}}{\epsilon}\right)$$

for any $0 < \epsilon < 1/4$, where $c_{\max} := \max_{e \in E}\{c_e\}$ is the maximum cost of an edge. This spectral rounding approach enlarges the set of constraints that one could handle in designing approximation algorithms for network design problems. We will discuss some technical ideas in the end of this chapter.

11.2 Barrier Method with Polynomials

The goal of this section is to present a proof of Theorem 11.2 using the polynomial perspective from [BSS14]. First, we will rephrase the barrier functions in Definition 10.3 in terms of polynomials. Then, we will present the plan following the intuition in the beginning of Chapter 10. Finally, we will proceed with the analysis and introduce some ideas about interlacing polynomials.

Soft-Max and Soft-Min of Polynomials

Recall the ϕ-soft-max and ϕ-soft-min in Definition 10.5 using the barrier functions in Definition 10.3. The strategy in the deterministic greedy Algorithm 7 in Chapter 10 is to fix ϕ_u, ϕ_l and then prove that $\phi_u\text{-}\max(A_t) \leq \phi_u\text{-}\max(A_{t-1}) + \delta_u$ and $\phi_l\text{-}\min(A_t) \geq \phi_l\text{-}\min(A_{t-1}) + \delta_l$ for all $t \geq 1$.

There are natural interpretations of the barrier functions from the polynomial perspective.

Remark 11.6 (Soft-Max and Soft-Min of Polynomials). *Let $p_A(y) = \det(yI - A)$ be the characteristic polynomial of A. Note that*

$$\phi\text{-}\max(p_A) \;:=\; \phi\text{-}\max(A) \;=\; \max\left\{u \;\middle|\; \Phi^A(u) = \frac{p_A'(u)}{p_A(u)} = \phi\right\} \;=\; \lambda_{\max}\Big(p_A - \frac{1}{\phi}p_A'\Big)$$

and

$$\phi\text{-}\min(p_A) \;:=\; \phi\text{-}\min(A) \;=\; \min\left\{l \;\middle|\; \Phi_A(l) = -\frac{p_A'(l)}{p_A(l)} = \phi\right\} \;=\; \lambda_{\min}\Big(p_A + \frac{1}{\phi}p_A'\Big).$$

So, using the ϕ-soft-min to lower bound $\lambda_{\min}(p)$ can be understood as using the minimum root a related polynomial $p + \frac{1}{\phi}p'$ to lower bound $\lambda_{\min}(p)$. Actually, slightly more can be said.

Exercise 11.7 (Soft-Min). *Let A be a real symmetric matrix. Show that $\lambda_{\min}(p_A) \geq \phi\text{-}\min(p_A) + \frac{1}{\phi}$ for any $\phi > 0$.*

Proof Plan

Given $v_1, \ldots, v_m \in \mathbb{R}^n$ and scalars $x_1, \ldots, x_m \in \mathbb{R}_{\geq 0}$ such that $\sum_{i=1}^m x_i v_i v_i^T = I_n$ and $\sum_{i=1}^m x_i = k$, our goal is to find $z_1, \ldots, z_m \in \mathbb{Z}_{\geq 0}$ with $\sum_{i=1}^m z_i \leq k$ and $\lambda_{\min}\big(\sum_{i=1}^m z_i v_i v_i^T\big)$ as large as possible.

Initially, we start with A_0 being the $n \times n$ zero matrix. In each iteration $1 \leq t \leq k$, we would like to find a vector $v \in \{v_1, \ldots, v_m\}$ and set $A_t = A_{t-1} + vv^T$. This will ensure that $A_t = \sum_{i=1}^m z_i v_i v_i^T$ for integers $z_1, \ldots, z_m$ for any $t \geq 0$.

As in Section 10.1, we consider the expected characteristic polynomial when we add a random vector with probability proportional to x_i. The following lemma is by the same calculation as in Section 10.1.

Exercise 11.8 (Expected Rank-One Update). *The expected characteristic polynomial after we add a vector v_i with probability x_i/k is*

$$\mathbb{E}[p_{A+vv^T}] := \sum_{i=1}^{m} \frac{x_i}{k} \cdot p_{A+v_i v_i^T} = p_A - \frac{1}{k} p'_A.$$

Now, instead of considering the roots of $(1 - \frac{1}{k}\partial)^k x^n$ after k iterations as in Section 10.1, we would like to use ϕ- min to show that (i) the "expected progress" after one iteration is good and (ii) there is a vector v which achieves this expected progress. Concretely, the plan is to prove that there exists a vector v with

$$\phi\text{- min}(p_{A+vv^T}) \geq \phi\text{- min}\left(\mathbb{E}_x[p_{A+vv^T}]\right) = \phi\text{- min}\left(p_A - \frac{1}{k} \cdot p'_A\right) \geq \phi\text{- min}(p_A) + \frac{1}{k+\phi}. \tag{11.1}$$

The equality is from Exercise 11.8. We will prove the last inequality in the next subsection, and then the first inequality in the subsection after. Assume the two inequalities in Equation 11.1 always hold. Then, by induction, after k iterations,

$$\lambda_{\min}(p_{A_k}) \geq \frac{1}{\phi} + \phi\text{- min}(p_{A_k}) \geq \frac{1}{\phi} + \phi\text{- min}(p_{A_0}) + \frac{k}{k+\phi} = -\frac{n-1}{\phi} + \frac{k}{k+\phi},$$

where the first inequality is from Exercise 11.7. Some calculations show that choosing

$$\phi = \frac{(n-1)k}{\sqrt{(n-1)k} - (n-1)} \quad \Longrightarrow \quad \lambda_{\min}\left(\sum_{i \in S} v_i v_i^T\right) \geq \left(1 - \sqrt{\frac{n-1}{k}}\right)^2.$$

This proves Theorem 11.2. It remains to prove the two inequalities in Equation 11.1 in the following two subsections.

Shifting Lower Barrier

It turns out that the techniques developed for the barrier functions in Chapter 10 can be used to bound the maximum and minimum root of a real-rooted polynomial as well. The following lemma is from Lemma 4.3 of [MSS21], proving the last inequality in Equation 11.1. The proof is very similar to the proofs in Lemma 10.8 and Lemma 10.10 for the lower barrier function, just rephrased in the language of polynomials.

Lemma 11.9 (Lower Barrier Shift [MSS21]). *If p is real-rooted and $s, \phi > 0$, then $p - sp'$ is real rooted and*

$$\phi\text{- min}\left(p - sp'\right) \geq \phi\text{- min}(p) + \frac{1}{\frac{1}{s} + \phi}.$$

Proof. It is well known and we will prove it in the next chapter that $p - sp'$ is real-rooted if p is. Let $l = \phi\text{- min}(p)$ such that l is the minimum value with $\Phi_p(l) = \phi$. Let

$$\delta := \frac{1}{\frac{1}{s} + \phi}.$$

To prove the lemma, we will prove that (i) $l+\delta < \lambda_{\min}(p-sp')$ and (ii) $\Phi_{p-sp'}(l+\delta) \leq \phi$, and this would imply that $\phi\text{-}\min(p-sp') \geq l+\delta = \phi\text{-}\min(p)+\delta$.

For (i), we claim that

$$\lambda_{\min}(p-sp') \geq \lambda_{\min}(p) \geq \phi\text{-}\min(p) + \frac{1}{\phi} > l+\delta.$$

The second inequality is from Exercise 11.7 and the third inequality is from the definition of δ. To see the first inequality, note that $p(y)$ and $-s \cdot p'(y)$ with $s > 0$ have the same sign for all $y < \lambda_{\min}(p)$, and thus any $y < \lambda_{\min}(p)$ cannot be a root of $p(y) - s \cdot p'(y)$, which implies that $\lambda_{\min}(p-sp') \geq \lambda_{\min}(p)$.

For (ii), we write $\Phi_{p-sp'}$ in terms of $\Phi_p = -p'/p$ so that

$$\Phi_{p-sp'} = -\frac{(p-sp')'}{p-sp'} = -\frac{((1+s\Phi_p)p)'}{(1+s\Phi_p)p} = -\frac{p'}{p} - \frac{s\Phi_p'}{1+s\Phi_p} = \Phi_p - \frac{\Phi_p'}{\frac{1}{s}+\Phi_p},$$

whenever all the quantities are finite, which happens everywhere except at the roots of p and $p-sp'$. Since $l+\delta$ is below the roots of p and $p-sp'$, it follows that

$$\Phi_{p-sp'}(l+\delta) = \Phi_p(l+\delta) - \frac{\Phi_p'(l+\delta)}{\frac{1}{s}+\Phi_p(l+\delta)} = \Phi_p(l) + \underbrace{(\Phi_p(l+\delta)-\Phi_p(l))}_{\text{loss}} - \underbrace{\frac{\Phi_p'(l+\delta)}{\frac{1}{s}+\Phi_p(l+\delta)}}_{\text{gain}}.$$

Therefore,

$$\Phi_{p-sp'}(l+\delta) \leq \Phi_p(l) = \phi \quad\Longleftrightarrow\quad \Phi_p(l+\delta) - \Phi_p(l) \leq \frac{\Phi_p'(l+\delta)}{\frac{1}{s}+\Phi_p(l+\delta)}. \tag{11.2}$$

As in Lemma 10.10, using convexity of $\Phi_p(l)$ will get us close but not enough; see Remark 11.10. So we need to work a bit harder as in Lemma 10.10. Using $\frac{1}{s} = \frac{1}{\delta} - \phi$ and rearranging, the condition in Equation 11.2 is equivalent to

$$\big(\Phi_p(1+\delta) - \Phi_p(l)\big)^2 \leq \Phi_p'(l+\delta) - \frac{1}{\delta}\big(\Phi_p(1+\delta) - \Phi_p(l)\big).$$

This is exactly what Claim 10.11 proved, which completes the proof of (ii) that $\Phi_{p-sp'}(l+\delta) \leq \phi$. $\square$

Remark 11.10 (Convexity Not Enough). *Using convexity $\Phi_p(l+\delta) - \Phi_p(l) \leq \delta \cdot \Phi_p'(l+\delta)$ in Exercise 10.4, the condition in Equation 11.2 holds if*

$$\delta \cdot \Phi_p'(l+\delta) \leq \frac{\Phi_p'(l+\delta)}{\frac{1}{s}+\Phi_p(l+\delta)} \quad\Longleftrightarrow\quad \delta \leq \frac{1}{\frac{1}{s}+\Phi_p(l+\delta)},$$

but we cannot conclude that δ being $1/\big(\frac{1}{s}+\Phi_p(l)\big) = 1/\big(\frac{1}{s}+\phi\big)$ suffices to maintain the nonincreasing potential. This is the same situation as in Lemma 10.10.

Common Interlacing

Now we would like to prove the first inequality in Equation 11.1. In general, given real-rooted polynomials $p_1, \ldots, p_m$ and a convex combination $q := \sum_{i=1}^m \mu_i p_i$ of them, there may not be any useful relations between the roots of $p_1, \ldots, p_m$ and q. For example, q may not have real roots even if $p_1, \ldots, p_m$ are all real-rooted, so we cannot hope to prove inequalities such as $\max_{1 \le i \le m}\{\lambda_{\min}(p_i)\} \ge \lambda_{\min}(q)$. We will discuss more in the next chapter.

An important observation of Marcus, Spielman, and Srivastava is that if $p_1, \ldots, p_m$ have a "common interlacing", then we can relate the roots of $p_1, \ldots, p_m$ and the roots of q and prove the inequality $\max_{1 \le i \le m}\{\lambda_{\min}(p_i)\} \ge \lambda_{\min}(q)$ and more. We will introduce interlacing polynomials properly in the next chapter. After reading the next chapter, it will be a good exercise to prove that the polynomials

$$p_1 := p_{A+v_1v_1^T} + \frac{1}{\phi}p'_{A+v_1v_1^T}, \quad \ldots, \quad p_m := p_{A+v_mv_m^T} + \frac{1}{\phi}p'_{A+v_mv_m^T}$$

have a common interlacing, to establish the first inequality in Equation 11.1.

Problem 11.11 (Interlacing Property of ϕ-min)**.** *If $p_1, \ldots, p_m$ are real-rooted polynomials that have a common interlacing, then for any expected polynomial $q = \sum_{i=1}^m \mu_i p_i$ with $\sum_{i=1}^m \mu_i = 1$ and $\mu_i \ge 0$ for $1 \le i \le m$, there exists $i \in [m]$ with*

$$\phi\text{-}\min(p_i) \ge \phi\text{-}\min(q).$$

Assuming Problem 11.11, we have completed the proof of Theorem 11.2 using polynomials.

11.3 Regret Minimization

The original proof of Theorem 11.2 is based on the regret minimization framework developed in [ALO15]. The general idea in regret minimization is to find some distributions of experts that are almost as good as the best experts. I won't be able to introduce regret minimization properly, but let me try to give an informal high level idea of the regret minimization framework in the specific setting of spectral rounding.

In the minimum eigenvalue problem in Theorem 11.2, the objective is to find a (multi-)subset S with large $\lambda_{\min}\big(\sum_{i \in S} v_i v_i^T\big)$, or equivalently a (multi-)subset S such that $x^T\big(\sum_{i\in S} v_i v_i^T\big)x$ is large for all vectors $x \in \mathbb{R}^n$ on the unit sphere. In this problem, we think of each direction x on the unit sphere as an expert. In each iteration t, given the current solution A_t, the regret minimization framework would maintain a "smart" probability distribution μ_t on the unit sphere, which puts higher probability on x if $x^T A_t x$ is small and a lower probability on x if $x^T A_t x$ is large. In words, the probability distribution puts more focus on the directions that the current solution A_t has not covered well. The distribution is summarized succinctly by a density matrix $P_t = \int xx^T d\mu_t$. This density matrix guides us naturally to add a vector v_t that maximizes the inner product $\langle v_i v_i^T, P_t\rangle$ to A_t, to cover the directions that are not covered well. The analysis in the regret minimization framework proves that if the probability distributions μ_t are smart, then the "regret"

$$\sum_{t\ge1}\langle v_t v_t^T, P_t\rangle - \min_{x\in\mathbb{R}^n:\|x\|=1}\sum_{t\ge1}\langle v_t v_t^T, xx^T\rangle$$

of using P_t over time instead of focusing on the worst directions (or best experts) is small. So, if we could always find a vector v_t in each iteration with a large inner product with P_t, which we can because of the isotropy condition, then we can conclude that $\sum_t\langle v_t v_t^T, P_t\rangle$ is large and hence $\min_x \sum_{t\ge1}\langle v_t v_t^T, xx^T\rangle = \min_x x^T\big(\sum_{t\ge1} v_t v_t^T\big)x = \lambda_{\min}\big(\sum_{t\ge1} v_t v_t^T\big)$ is also large.

A versatile and commonly used approach to maintain the distributions is by the multiplicative weight update method. If we use it for spectral sparsification, then we can recover the $O(n \log n/\epsilon^2)$ result by Spielman and Srivastava in Theorem 9.9.

The insight in [ALO15] is that the barrier functions used by Batson, Spielman and Srivastava in Definition 10.3 can be interpreted as a new way of updating the probability distributions, by setting $P_t = (\phi A_t - l_t I)^{-2}$ where A_t is the current solution and l_t is chosen such that $P_t \succcurlyeq 0$ and $\operatorname{Tr}(P_t) = 1$. This can be integrated into the regret minimization framework to give the following regret bound:

$$\lambda_{\min}\left(\sum_{t=1}^{\tau} v_t v_t^T\right) \geq \underbrace{-\frac{2\sqrt{n}}{\phi}}_{\text{initial lower bound}} + \sum_{t=1}^{\tau} \underbrace{\frac{\langle v_t v_t^T, P_t\rangle}{1 + \phi\langle v_t v_t^T, P_t^{1/2}\rangle}}_{\text{increase of } \phi\text{-soft-min}} \tag{11.3}$$

where we should interpret ϕ as the same parameter in soft-min, the negative term is the initial lower bound in the barrier method, and each term in the summation as the increase of the ϕ-soft-min. Using the isotropy condition $\sum_{i=1}^m v_i v_i^T = I_n$, it is possible to show that there always exists a vector with a large increase of ϕ-soft-min to prove Theorem 11.2.

Local Search

In applications of spectral rounding or the minimum eigenvalue problem, often we are given $x \in [0,1]^m$ and we would like to find $z \in \{0,1\}^m$ instead of just $z \in \mathbb{Z}^m$. This is called the "without repetition" setting, where each vector can be chosen at most once, which is a more general setting than the "with repetition" setting, where each vector can be chosen more than once. In [ALSW17], it was shown that the same greedy algorithm can only achieve a constant factor approximation algorithm, not the $(1 \pm \epsilon)$-approximation algorithm in Problem 11.3 when $k \geq n/\epsilon^2$.

An interesting new idea in [ALSW21] is to analyze a local search algorithm, where we start from an arbitrary subset S_0 with k vectors, and in each iteration $t \geq 1$ we find a pair $i_{t-1} \in S_{t-1}$ and $j_{t-1} \notin S_{t-1}$ and set $S_t := S_{t-1} - i_{t-1} + j_{t-1}$. This guarantees that each vector is chosen at most once. They developed the following rank-two update formula for regret minimization:

$$\lambda_{\min}\left(\sum_{l \in S_t} v_l v_l^T\right) \geq -\frac{2\sqrt{n}}{\phi} + \sum_{t=1}^{\tau}\left(\frac{\langle v_{j_t} v_{j_t}^T, P_t\rangle}{1 + 2\phi\langle v_{j_t} v_{j_t}^T, P_t^{1/2}\rangle} - \frac{\langle v_{i_t} v_{i_t}^T, P_t\rangle}{1 - 2\phi\langle v_{i_t} v_{i_t}^T, P_t^{1/2}\rangle}\right), \tag{11.4}$$

and used it to get the same result as in Theorem 11.2 in the more challenging without repetition setting. It would be interesting to recover this result using the polynomial approach in Section 11.2. See Problem 11.14 for a possible starting point.

Randomized Local Search

The local search approach in [ALSW21] is extended in [LZ20] to handle linear constraints as described in Definition 11.1. The idea is to randomly choose a vector v_{i_t} to remove from S_{t-1} with probability proportional to $(1 - x_i) \cdot \big(1 + 2\alpha\langle v_{j_t} v_{j_t}^T, P_t^{1/2}\rangle\big)$, where x_i is the fractional value of the i-th vector and the other term is in the denominator in Equation 11.4, and similarly choose a vector v_{j_t} to add to S_{t-1} with probability proportional $x_j \cdot \big(1 - 2\phi\langle v_{i_t} v_{i_t}^T, P_t^{1/2}\rangle\big)$. Informally, the terms $(1 - x_i)$ and x_j are to ensure that the linear constraints are approximately preserved, and the terms from the denominators in Equation 11.4 are to ensure that the minimum eigenvalue is improving. The proof is by showing that these sums are concentrated around their expected values.

In [LZ21], it was shown that this randomized local search approach gives the best known algorithms for experimental design problems, where for D/A-design the randomized local search algorithm achieves a $(1 \pm \epsilon)$-approximation when $k \gtrsim n/\epsilon$, better than the requirement $k \gtrsim n/\epsilon^2$ for E-design.

Again, it would be interesting to recover these results using the polynomial approach in Section 11.2.

Two-Sided Spectral Rounding

We will study the two-sided spectral rounding problem in Definition 11.1 in the second part of the course.

11.4 Problems

Problem 11.12 (Total Effective Resistance)**.** *Let L_G be the Laplcian matrix of a graph $G = (V, E)$. Recall that $\mathit{Reff}_G(u, v) = (\chi_u - \chi_v)^T L_G^\dagger (\chi_u - \chi_v)$ is the effective resistance between vertices $u, v \in V$. Show that $|V| \cdot \mathrm{Tr}(L_G^\dagger) = \frac{1}{2}\sum_{u,v \in V} \mathit{Reff}_G(u, v)$. Use this fact with Theorem 11.2 to obtain an approximation algorithm for minimizing the total effective resistance subject to the constraint that the (multi-)subgraph has at most k edges.*

Problem 11.13 (Upper Barrier Shift)**.** *Prove the following analog of Lemma 11.9 for the upper barrier function. If p has real roots and $s, \phi > 0$, then $p - sp'$ is real-rooted and*

$$\phi\text{-}\max\left(p - sp'\right) \leq \phi\text{-}\max(p) + \frac{1}{\frac{1}{s} - \phi}.$$

Problem 11.14 (Expected Polynomial After Removal)**.** *This problem might be helpful in obtaining the bound in Theorem 11.2 in the more challenging without repetition setting using the polynomial approach. Suppose the current solution A has k vectors say $v_1, \ldots, v_k \in \mathbb{R}^n$. Show that the expected characteristic polynomial after removing a uniformly random vector is*

$$\mathbb{E}\left[p_{A - vv^T}(y)\right] := \frac{1}{k}\sum_{i=1}^{k} p_{A - v_i v_i^T}(y) = \left(1 - \frac{n}{k}\right) \cdot p_A(y) + \frac{y}{k} \cdot p'_A(y).$$

Question 11.15 (Improved Approximation Ratio when $k = n$)**.** *Is it possible to improve the $\Theta(1/n^2)$ minimum eigenvalue bound in Theorem 11.2 when $k = n$?*

11.5 References

[ALO15] Zeyuan Allen-Zhu, Zhenyu Liao, and Lorenzo Orecchia. Spectral sparsification and regret minimization beyond matrix multiplicative updates. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015, Portland, OR, USA, June 14-17, 2015*, pages 237–245. ACM, 2015. 105, 108, 113, 114

[ALSW17] Zeyuan Allen-Zhu, Yuanzhi Li, Aarti Singh, and Yining Wang. Near-optimal design of experiments via regret minimization. In *Proceedings of the 34th International Conference on Machine Learning, ICML 2017, Sydney, NSW, Australia, 6-11 August 2017*, volume 70 of *Proceedings of Machine Learning Research*, pages 126–135. PMLR, 2017. 107, 108, 114

[ALSW21] Zeyuan Allen-Zhu, Yuanzhi Li, Aarti Singh, and Yining Wang. Near-optimal discrete optimization for experimental design: a regret minimization approach. *Math. Program.*, 186(1):439–478, 2021. 108, 109, 114

[BSS14] Joshua D. Batson, Daniel A. Spielman, and Nikhil Srivastava. Twice-ramanujan sparsifiers. *SIAM Rev.*, 56(2):315–334, 2014. 97, 98, 99, 100, 103, 108, 110

[LZ20] Lap Chi Lau and Hong Zhou. A spectral approach to network design. In *Proccedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, STOC 2020, Chicago, IL, USA, June 22-26, 2020*, pages 826–839. ACM, 2020. 109, 110, 114

[LZ21] Lap Chi Lau and Hong Zhou. A local search framework for experimental design. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms, SODA 2021, Virtual Conference, January 10 - 13, 2021*, pages 1039–1058. SIAM, 2021. 108, 109, 115

[MSS21] Adam W Marcus, Daniel A Spielman, and Nikhil Srivastava. Interlacing families iii: Sharper restricted invertibility estimates. *Israel Journal of Mathematics*, pages 1–28, 2021. 104, 111