Lecture 10: The Kadison-Singer Problem for Strongly Rayleigh Measures

Lecturer: R Zhang *November 3*

Motivation and Strongly Rayleigh Measures

In the last lecture, we saw a proof of KS_2 using the following spectral bound on random vectors:

Theorem 3.1. *If* $\epsilon > 0$ *and* $v_1, ..., v_m$ *are independent random vectors in* $\mathbb{R}^d$ *with finite support where,*

$$\sum_{i=1}^{m} \mathbb{E}[v_i v_i^T] = I$$

$$\forall i : \mathbb{E}\|v_i\|^2 \leq \epsilon$$

then

$$\mathbb{P}\left[\left\|\sum_{i=1}^{m} v_i v_i^T\right\| \leq (1+\sqrt{\epsilon})^2\right] > 0$$

In this talk, we will see an extension of the above theorem to a more general distribution of random vectors, of which independent distributions are a special case.

Definition 3.2 (Probability Generating Function)**.** Let $\mu : 2^{[m]} \to [0,1]$ be a probability measure on the set of subsets of $[m] = \{1, 2, ..., m\}$. The generating polynomial of this distribution is multi-affine of the variables $z = \{z_1, ..., z_m\}$:

$$g_\mu(z) = \sum_{S \subseteq [m]} \mu(S) \cdot z^S$$

where $z^S = \prod_{i \in S} z_i$.

Definition 3.3. A probability distribution is homogenous if its generating polynomial as defined above is homogenous (if the degree of all its terms is the same)

Definition 3.4. A probability distribution is Strongly Rayleigh if its generating polynomial is Real Stable (it has no roots in the closure of the upper half plane $\mathcal{H}^m$)

Definition 3.5. The marginal probability of element i with respect to the distribution μ is

$$\mathbb{P}_\mu[i \in S] = \sum_{S \subseteq [m], i \in S} \mu(S) = \partial_{z_i} g_\mu(z)|_{z_1 = ... = z_m = 1}$$

Theorem 3.6 (Main)**.** *Let μ be a homogenous Strongly Rayleigh probability measure on $[m]$ with marginal probabilities $\forall i : \mathbb{P}[i \in S] \leq \epsilon_1$. Let $v_1, ..., v_m \in \mathbb{R}^d$ be an isotropic set of vectors, $\sum_{i=1}^{m} v_i v_i^T = I$ and $\forall i : \mathbb{E}\|v_i\|^2 \leq \epsilon_2$. Then*

$$\mathbb{P}_\mu \left[\left\| \sum_{i=1}^{m} v_i v_i^T \right\| \leq 4(\epsilon_1 + \epsilon_2) + 2(\epsilon_1 + \epsilon_2)^2 = O(\epsilon_1 + \epsilon_2) \right] > 0$$

This technical theorem will be used to find a thin-basis to prove the existence of a spectrally thin tree, which could be used for an improved approximation algorithm to the asymmetric Traveling Salesman Problem.

Claim 3.7 (Thin Tree Conjecture)**.** *If the graph G is k-edge connected, there exists a spanning tree (T) of G: $\forall S \subseteq V : |E(S, \overline{S}) \cap T| \leq O(\frac{1}{k})|E(S, \overline{S})|$*

Mixed Characteristic Polynomial and Interlacing

Like the previous applications, in order to bound a particular assignment of strongly Rayleigh vectors, we will first bound identify the Expected characteristic polynomial, then use Interlacing.

Theorem 3.8. *If g_μ is a homogenous Strongly Rayleigh polynomial of degree d_μ,*

$$x^{d_\mu - d} \mathbb{E}_\mu \chi \left[\sum_{i \in S} 2 v_i v_i^T \right] (x^2) = \prod_{i=1}^{m} (1 - \partial_{z_i}^2) \left(g_\mu(x\vec{1} + z) \cdot \det \left(xI + \sum_{i=1}^{m} z_i v_i v_i^T \right) \right) \Bigg|_{z_1 = \ldots = z_m = 0}$$

Proof.

Lemma 3.9 (Cauchy-Binet)**.** *For vectors $v_1, ..., v_m \in \mathbb{R}^m$ and scalars $z_1, ..., z_m$*

$$\det \left(xI + \sum_{i=1}^{m} z_i v_i v_i^T \right) = \sum_{k=0}^{m} x^{m-k} \sum_{S \subseteq [m], |S| = k} z^S \sigma_k (\sum_{i \in S} v_i v_i^T)$$

where $\sigma_k(M)$ is the sum of all principal $k \times k$ minors of M.

Fix $S \subseteq [m]$ such that $|S| = k$. The coefficient of z^{2S} comes from only the terms with z^S in both g_μ and $\det$, which is

$$\left(\prod_{i \in S} \partial_{z_i}^2 \right) \left(g_\mu(x\vec{1} + z) \cdot \det \left(xI + \sum_{i=1}^{m} z_i v_i v_i^T \right) \right) \Bigg|_{z=0}$$

Since both g_μ and $\det$ are multi-affine in z, we can use the product rule

$$= 2^k \cdot \left(\left(\prod_{i \in S} \partial_{z_i} \right) g_\mu(x\vec{1} + z) \right) \Bigg|_{z=0} \cdot \left(\left(\prod_{i \in S} \partial_{z_i} \right) \det \left(xI + \sum_{i=1}^{m} z_i v_i v_i^T \right) \right) \Bigg|_{z=0}$$

$$= 2^k \cdot x^{d_\mu - k} \mathbb{P}_\mu[S \subseteq T] \cdot x^{d-k} \sigma_k \left(\sum_{i \in S} z_i v_i v_i^T \right)$$

by the fact that g_μ is homogenous and by the Cauchy-BInet formula.

Applying the above identities to all subsets $S \subseteq [m]$
$\prod_{i=1}^{m}(1-\partial_{z_i}^2)\left(g_\mu(x\vec{1}+z)\cdot\det\left(xI+\sum_{i=1}^{m} z_i v_i v_i^T\right)\right)\Big|_{z_1=\ldots=z_m=0}$

$$= \sum_{k=0}^{m}(-1)^k \sum_{S\subseteq[m],|S|=k}\left(\prod_{i\in S}\partial_{z_i}^2\right)\left(g_\mu(x\vec{1}+z)\cdot\det\left(xI+\sum_{i=1}^{m} z_i v_i v_i^T\right)\right)\Bigg|_{z=0}$$

$$= \sum_{k=0}^{d} x^{d_\mu-d}(-1)^k(2)^k(x^2)^{d-k}\sum_{S\subseteq[m],|S|=k}\mathbb{P}_\mu[S\subseteq T]\cdot\sigma_k\left(\sum_{i\in S} z_i v_i v_i^T\right)$$

$$= x^{d_\mu-d}\mathbb{E}_\mu\chi\left[\sum_{i\in S}2v_i v_i^T\right](x^2)$$

again by the Cauchy-Binet formula. □

The following properties have been shown previously:

Theorem 3.10. *For any set of PSD matrices A_i, the following polynomial is real stable:*

$$\det\left(\sum_{i=1}^{m} z_i A_i\right)$$

Lemma 3.11. $(1+c\partial_{z_i})$ *is a stability preserving operation.*

Corollary 3.12. $(1-\partial_{z_i}^2)=(1-\partial_{z_i})(1+\partial_{z_i})$ *is also stability preserving.*

Corollary 3.13. *The mixed characteristic polynomial of a strongly Rayleigh measure is real-rooted*

Proof. $\det\left(xI+\sum_{i=1}^{m} z_i v_i v_i^T\right)$ is a sum of a diagonal and $z_i\cdot$(rank-one matrices), so this term is real stable. By definition g_μ is real stable, so the product of these two $g_\mu\cdot\det$ is real stable. And $(1-\partial_{z_i}^2)F|_{z_i=0}$ is a stability preserving operation, so by the previous theorem, the mixed characteristic polynomial differs only by $x^{d_\mu-d}$ so by evaluating at $z=0$, the mixed characteristic polynomial is univariate and real-rooted. □

Theorem 3.14.

$$q_S(x)=\mu(S)\chi\left[\sum_{i\in S}2v_i v_i^T\right](x^2)$$

The polynomials $\{q_S\}_S$ taken over partial assignments of vectors form an interlacing family.

Proof.

$$q_{s_1,\ldots,s_k}(x)=\sum_{\{i|s_i=1\}\subseteq S\subseteq[m]} q_S(x)$$

It is enough to inductively show that $\lambda q_{s_1,\ldots,s_k,1}(x)+(1-\lambda)q_{s_1,\ldots,s_k,0}(x)$ is real rooted for any partial assignment. But these are just mixed characteristic polynomials over conditioned measures:

$$g_{\mu_{s_1,\ldots,s_k,1}}(z)=\frac{z_{k+1}\cdot\partial_{z_{k+1}}g_{\mu_{s_1,\ldots,s_k}}(z)}{\partial_{z_{k+1}}g_{\mu_{s_1,\ldots,s_k}}|_{z=1}}$$

$$g_{\mu_{s_1,\ldots,s_k,0}}(z)=\frac{g_{\mu_{s_1,\ldots,s_k}}(z)|_{z_{k+1}=0}}{g_{\mu_{s_1,\ldots,s_k}}|_{z_{k+1}=0,z_{i\neq k+1}=1}}$$

Because the denominator is a normalizing constant, and differentiation and specialization preserve stability, these new conditioned measures are also strongly Rayleigh. So by the previous theorem, $\lambda q_{s_1,\ldots,s_k,1}(x)+(1-\lambda)q_{s_1,\ldots,s_k,0}(x)$ is real rooted. □

Extension of the Multivariate Barrier Argument

In order to carry out the barrier argument, we need to define multivariate versions of what it means to be above the largest root of a polynomial, and a directional barrier function.

Definition 3.15. $z\in\mathbb{R}^m$ is above all the roots ($:=Ab_p$) of a multivariate polynomial $p(z_1,\ldots,z_m)$ if for all $t\in\mathbb{R}^m_{>0}:p(z+t)>0$.

Definition 3.16. The barrier functions in direction i of a real stable polynomial p is defined for all z above the roots of p:

$$\Phi_p^i(z):=\frac{\partial_{z_i}p(z)}{p(z)}=\partial_{z_i}\log p(z)$$

$$\Psi_p^i(z):=\frac{\partial_{z_i}^2 p(z)}{p(z)}$$

For a univariate restriction $q_{z,i}(t)=p(z_1,\ldots,z_{i-1},t,z_{i+1},z_m)$ with real roots $\lambda_1,\ldots,\lambda_r$, the directional derivative can be written more simply as:

$$\Phi_p^i(z)=\frac{q'_{z,i}(z_i)}{q_{z,i}(z_i)}=\sum_{j=1}^r\frac{1}{z_i-\lambda_j}$$

$$\Psi_p^i(z)=\frac{q''_{z,i}(z_i)}{q_{z,i}(z_i)}=\sum_{1\leq j<k\leq r}\frac{2}{(z_i-\lambda_j)(z_i-\lambda_k)}$$

Theorem 3.17. *Given vectors $v_1,\ldots,v_m\in\mathbb{R}^d$ and a homogenous strongly Rayleigh probability measure μ, such that the marginal probability $P_\mu[i\in S]\leq\epsilon_1$ for all i, such that the vectors are in isotropic position $\sum_{i=1}^m v_iv_i^T=I$, and $\|v_i\|^2\leq\epsilon_2$, the largest root of the mixed characteristic polynomial is $\leq 4(2\epsilon+\epsilon^2)$ where $\epsilon=\epsilon_1=\epsilon_2$.*

Lemma 3.18. *For real stable p and $z\in Ab_p$, the following monotonicity and convexity property holds for all directions i,j and $\delta\geq 0$:*

$$\Phi_p^i(z_1,\ldots,z_j+\delta,\ldots,z_m)\leq\Phi_p^i(z)$$

$$\Phi_p^i(z_1,\ldots,z_j+\delta,\ldots,z_m)\leq\Phi_p^i(z)=\delta\partial_{z_j}\Phi_p^i(z_1,\ldots,z_j+\delta,\ldots,z_m)$$

Unlike in the previous lecture, we need the norm to be bounded away from 1, not $1+\epsilon$. Fortunately, we are applying $(1-\partial_{z_i})$, which should move the boundary up by $(1+\delta_1)$, and then $(1+\partial_{z_i})$, which intuitively should move the boundary back down by $(1+\delta_2)$. This is what allows for the necessary smaller bound.

Lemma 3.19. *If p is real stable and $z \in Ab_p$ is such that $\Phi_p^i(z){<}1$, then $z \in Ab_{p-\partial_{z_i}^2 p}$*

Lemma 3.20. *For real stable p and $z \in Ab_p$, if for $\delta{>}0$,*

$$\frac{2}{\delta}\Phi_p^j(z) + \Phi_p^j(z)^2 \leq 1$$

then for all directions i,

$$\Phi^i_{(1-\partial_{z_j}^2)p}(z_1, ..., z_j + \delta, ..., z_m) \leq \Phi_p^i(z)$$

Proof of Main Theorem. Let $\epsilon = \epsilon_1 + \epsilon_2$. By Theorem, the largest root of the mixed characteristic polynomial of a strongly Rayleigh measure is $\leq 2\sqrt{2\epsilon+\epsilon^2}$. By Theorem, $\{q_S\}_{S|\mu S>0}$ forms an interlacing family. So by Theorem, there exists some set $S \subseteq [m]$ with $\mu(S){>}0$ such that

$$\lambda_{max}\left[\det\left(x^2 I - \sum_{i\in S} 2v_iv_i^T\right)\right] \leq \lambda_{max}\left[\mathbb{E}\chi\right] \leq 2\sqrt{2\epsilon+\epsilon^2}$$

Therefore, for this set S

$$\left\|\sum_{i\in S} v_iv_i^T\right\| = \frac{1}{2}\left\|\sum_{i\in S} 2v_iv_i^T\right\| \leq \frac{1}{2}(2\sqrt{2\epsilon+\epsilon^2})^2 \leq 4\epsilon + 2\epsilon^2$$

□