The Kadison-Singer Conjecture

John E. M$^{\rm c}$Carthy

April 18, 2006

1 Set-up

We are given a large integer N, and a fixed basis $\{e_i : 1 \leq i \leq N\}$ for the space $\mathcal{H} = \mathbb{C}^N$. We are also given an N-by-N matrix H that has all its diagonal entries zero, and has norm one.

If $\sigma \subseteq \{1, 2, \ldots, N\}$, let P_σ be the projection onto the space $\mathcal{H}_\sigma := \vee\{e_i : i \in \sigma\}$. Call such a projection *basic*, and the range of a basic projection a *basic subspace*.

For each constant $0 < \gamma < 1$, we want to know

$$K(H, \gamma) \ := \ \max\{|\sigma| \ : \ \|P_\sigma H P_\sigma\| \leq \gamma\} \tag{1.1}$$

and

$$\begin{aligned} \Pi(H, \gamma) \ := \ \min\{k \quad : \quad & \text{there is a partition } \sigma_1, \ldots, \sigma_k \text{ of } \{1, \ldots, N\} \\ & \text{such that } \forall\ i, \quad \|P_{\sigma_i} H P_{\sigma_i}\| \leq \gamma \ \}. \end{aligned} \tag{1.2}$$

The Kadison-Singer conjecture [3] is equivalent to the assertion that for each γ, there is a universal bound on $\Pi(H, \gamma)$ that is independent of H and N.

For the Kadison-Singer conjecture to hold, it is clearly necessary that there exists some constant $c_\gamma > 0$ such that

$$K(H, \gamma) \ \geq \ c_\gamma N. \tag{1.3}$$

In [1], K. Berman, H. Halpern, V. Kaftal and G. Weiss proved that there exists a self-adjoint Hadamard matrix H (a unitary all of whose entries have modulus $1/\sqrt{N}$) such that, if you drop the diagonal and renormalize,

$$\Pi(H,\gamma) \ \approx \ \gamma^{-2}.$$

In [2], J. Bourgain and L. Tzafiri proved that there exists a universal constant C (of order 10^{-7}) such that

$$K(H,\gamma) \ \geq \ C\gamma^2 \qquad \forall\ H. \tag{1.4}$$

2 An approach to $K(\gamma, H)$

For now, let us assume that H is self-adjoint, and try to find a large σ so that the one-sided inequality

$$P_\sigma H P_\sigma \ \leq \ \gamma I \tag{2.1}$$

holds. (Later, we hope to extend the method to satisfying (2.1) simultaneously with several operators; if they are H and $-H$, we would get a bound on $\|H\|$).

Assume that σ of size k has been chosen; after relabelling the indices, let us assume that $\sigma = \{1, \dots, k\}$. Let P_k denote P_σ, and assume that

$$P_k H P_k \ \leq \ C_k I_k.$$

We wish to choose i in $\{k+1, \dots, N\}$ so that, if $\sigma' \ = \ \sigma \cup \{i\}$ and $P_{k+1} := P_{\sigma'}$, then

$$P_{k+1} H P_{k+1} \ < \ C_{k+1} I_{k+1}, \tag{2.2}$$

with $C_{k+1} = C_k + \varepsilon_k$ as small as possible.

Let u^i be the k-vector $P_k H e_i$.

Lemma 2.3 *Inequality (2.2) holds, with $\varepsilon_k \geq 0$, if and only if*

$$C_{k+1} \ > \ \langle (C_{k+1} - P_k H P_k)^{-1} u^i, u^i \rangle.$$

PROOF: This is just the condition for something in block matrix form to be positive definite. □

Let the eigenvalues of $P_k H P_k$ be $\mu_1^k \geq \mu_2^k \geq \ldots \geq \mu_k^k$, and let the corresponding orthonormal eigenvectors be w_j. If ε is so small that (2.2) fails for every i, with $C_{k+1} = C_k + \varepsilon$, then

$$\begin{aligned}
(N-k)\,(C_k+\varepsilon) &< \sum_{i=k+1}^{N} \langle (C_k+\varepsilon-P_kHP_k)^{-1}u^i, u^i\rangle \\
&= \sum_{i=k+1}^{N}\sum_{j,l=1}^{k} \langle (C_k+\varepsilon-P_kHP_k)^{-1}\langle u^i, w_j\rangle w_j, \langle u^i, w_l\rangle w_l\rangle \\
&= \sum_{j=1}^{k}\sum_{i=k+1}^{N} |\langle u^i, w_j\rangle|^2 \frac{1}{C_k+\varepsilon-\mu_j} \\
&= \sum_{j=1}^{k} \frac{\|(I-P_k)Hw_j\|^2}{C_k+\varepsilon-\mu_j} \\
&\leq \sum_{j=1}^{k} \frac{1}{C_k+\varepsilon-\mu_j}.
\end{aligned}$$

Therefore we have

Lemma 2.4 *The smallest attainable ε_k is less than or equal to that value of ε where the decreasing function*

$$\sum_{j=1}^{k} \frac{1}{C_k+\varepsilon-\mu_j^k}$$

equals $(N-k)C_k$.

The very crude estimate $\mu_j^k \leq C_k$ for each j would give the bound

$$\varepsilon_k \;\leq\; \frac{k}{(N-k)C_k},$$

but this is not sharp enough.

3 Weyl's inequality

Weyl's inequality for self-adjoint n-by-n matrices A, B and D with $A+B = D$ is that

$$\mu_j(D) \;\le\; \mu_{j-m+1}(A) \;+\; \mu_m(B) \tag{3.1}$$

for $1 \le m \le j \le n$, where $\mu_j(A)$ is the j^{th} largest eigenvalue of A.

If we apply this to $P_{k+1}HP_{k+1}$, which we think of the sum of the matrix

$$A_k \;=\; \left(\begin{array}{cccc|c} & & & & 0 \\ & P_kHP_k & & & 0 \\ & & & & \vdots \\ \hline 0 & 0 & \cdots & & 0 \end{array}\right)$$

with the rank two, trace zero, matrix

$$B \;=\; u^i \otimes e_{k+1} \;+\; e_{k+1} \otimes u^i,$$

and let $m = 2$, we get that, for $2 \le j \le k+1$,

$$\begin{aligned} \mu_j^{k+1}(P_{k+1}HP_{k+1}) \quad &\le \quad \mu_{j-1}(A_k) \\ &\le \quad \max(\mu_{j-1}^k, 0). \end{aligned} \tag{3.2}$$

Applying (3.2) inductively to Lemma 2.4, we get

Lemma 3.3

$$(N-k)C_{k+1} \;\le\; \sum_{j=1}^{k} \frac{1}{C_{k+1} - C_j}. \tag{3.4}$$

Equation (3.4) gives, I think, that

$$C_k \;=\; O(k \log(k)/N)^{1/2},$$

which is logarithmically worse than (1.4).

4 Tightening it up

One can redo the arguments in Section 3 and Lemma 2.4 to keep track of all the eigenvalues at once. Indeed, typically (ie unless the first component of u with respect to the basis $w_1, \ldots, w_k$ is 0) the second eigenvalue at the $k+1^{\mathrm{st}}$ stage will be strictly smaller than the largest eigenvalue at the k^{th} stage.

Indeed, look at $C - A_{k+1}$, for some constant C, and take the Schur complement for the $(k+1)$-by-$(k+1)$ slot.

$$C - A_{k+1} \;=\; \left(\begin{array}{cccc|c} C-\mu_1 & 0 & \cdots & 0 & -u_1 \\ 0 & C-\mu_2 & \cdots & 0 & -u_2 \\ 0 & 0 & \cdots & C-\mu_k & \vdots \\ \hline -\overline{u_1} & -\overline{u_2} & \cdots & & C \end{array}\right),$$

where we suppress the superscripts and write u_j for $\langle u^i, w_j \rangle$.

This gives that the signature (number of positive minus number of negative eigenvalues) for $C - A_{k+1}$ is the same as the sum for the k-by-k matrix $C - A_k$ and the 1-by-1 matrix

$$C - \sum_{j=1}^{k} \frac{|u_j|^2}{C - \mu_j}. \tag{4.1}$$

As C varies, looking at the sign changes in (4.1) one gets that the new eigenvalues are the intercepts of the function

$$F(x) \;=\; \sum_{j=1}^{k} \frac{|u_j|^2}{x - \mu_j^k}$$

with the function $G(x) = x$. (If $u_j \neq 0$, one doesn't get a sign change in the signature at μ_j).

Looking at a picture of the graphs of F and G, if the old (k^{th} stage) eigenvalues with an extra 0 added are

$$\mu_1' \geq \mu_2' \geq \ldots \geq \mu_{k+1}'$$

and the new ones are $\nu_1 \geq \ldots \geq \nu_{k+1}$, then

$$|\nu_j| \ \geq \ |\mu_j| \qquad \forall\ j.$$

So as the Hilbert-Schmidt norm squared of the matrix equals the sum of the squares of the eigenvalues, writing $\nu_j = \mu_j + \varepsilon_j$, we get

$$2\sum|\varepsilon_j||\mu_j| + \sum\varepsilon_j^2 \ = \ 2\sum|u_j|^2. \tag{4.2}$$

The game then is to show that the solutions of $F(x) = x$ don't grow to fast as k increases. If the largest root goes up a lot, then by (4.2) the others can't change much. So at the next stage the largest root does not increase much (draw picture of $F(x)$ to see this).

5 Case $u_j = \frac{1}{\sqrt{N}}$

In this case, the solutions of $F(x) = x$ are the roots of the polynomial q_{k+1}, where, after rescaling by $y = \sqrt{N}x$, we have $q_1(y) = y$, and

$$q_{k+1}(y) \ = \ yq_k(y) - q_k'(y).$$

One then wants the largest root of q_k to be $O(\sqrt{k})$. The q_k's turn out to be approximately the Hermite polynomials, for which indeed the largest root is $O(\sqrt{k})$.

6 Possible Refinements

1. There are several analogies with orthogonal polynomials; if one believes they are profound, one could be led to the conjecture that

$$|\mu_j^k| \ = \ O\left(\sqrt{\frac{k}{N}}\cos\left[\frac{j-1}{k-1}\pi\right]\right).$$

2. The estimate preceding Lemma 2.4 is sharp only if all the terms are the same — if some u^i's actually contribute a lot to the sum, others will contribute less. Can this be exploited?

3. It seems one cannot rule out occasional large jumps. For example, suppose N is a power of 2, and H is the tensor product of

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

with itself may times. (The diagonal entries of H are not 0, but $\pm 1/\sqrt{N}$; this doesn't really matter). If you choose the indices in order, then at the $k = 2^l$ stage, you have 2^{l-1} eigenvalues of size $\sqrt{\frac{k}{N}}$, and the other half are $-\sqrt{\frac{k}{N}}$. At the $k+1$ stage, one eigenvalue leaps to $\sqrt{\frac{2^{l+1}}{N}}$, a large increase. But for each successive step up to the 2^{l+1} stage, you get one more eigenvalue with size $\sqrt{\frac{2^{l+1}}{N}}$, without any increase in the largest eigenvalue. By the time you reach $2k$, the overall norm of the compression has gone up only by a factor of $\sqrt{2}$, but it did it all in one step.

References

[1] K. Berman, H. Halpern, V. Kaftal, and G. Weiss. Matrix norm inequalities and the relative Dixmier property. *Integral Equations and Operator Theory*, 1:28–48, 1988.

[2] J. Bourgain and L. Tzafiri. On a problem of kadison and singer. *J. Reine Angew. Math.*, 420:1–43, 1991.

[3] R. Kadison and I. Singer. Extensions of pure states. *Amer. J. Math.*, 81:547–564, 1959.