Norm Compression Inequalities for Block Partitioned Matrices

Institute for
Mathematical Sciences

Koenraad M.R. Audenaert

Quantum Information at
Imperial College
London

December 16, 2006

Block Partitioned Matrices

- Consider two spaces V_1 and V_2.
- Single out orthogonal subspaces in V_1 and V_2, and partition V_1 and V_2 as direct sums

$$V_1 = \bigoplus_j V_{1j}, \quad V_2 = \bigoplus_i V_{2i}$$

- Linear mappings between V_1 and V_2 can be represented by Block Matrices

$$A = [A_{ij}] = \begin{pmatrix} A_{11} & A_{12} & \dots \\ A_{21} & A_{22} & \dots \\ \vdots & \vdots & \ddots \end{pmatrix},$$

where A_{ij} are matrices, representing mappings from V_{1j} to V_{2i}.

Matrix Trace Norms

- Among the zoo of norms that can be defined for matrices, the Schatten q-norms (a.k.a. Trace Norms) stand out:
- For $q \geq 1$, $||A||_q := (\mathrm{Tr}(|A|^q))^{1/q}$.
- The matrix absolute value $|A|$ is defined as $|A| = (A^*A)^{1/2}$.
- Non-commutative generalisation of ℓ_q-norms for vectors.
- Depend only on singular values of A: $||A||_q = (\sum_i \sigma_i(A)^q)^{1/q}$

Norm Compression

- Suppose I have a block partitioned matrix $A = [A_{ij}]$

Norm Compression

- Suppose I have a block partitioned matrix $A = [A_{ij}]$
- ...and I give you the Schatten q-norms of all the blocks: $||A_{ij}||_q$

Norm Compression

- Suppose I have a block partitioned matrix $A = [A_{ij}]$
- ...and I give you the Schatten q-norms of all the blocks: $||A_{ij}||_q$
- Let's call that the Norm Compression of A:

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots \\ A_{21} & A_{22} & \dots \\ \vdots & \vdots & \ddots \end{pmatrix} \quad \mapsto \quad \begin{pmatrix} ||A_{11}||_q & ||A_{12}||_q & \dots \\ ||A_{21}||_q & ||A_{22}||_q & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

Norm Compression

- Suppose I have a block partitioned matrix $A = [A_{ij}]$
- ...and I give you the Schatten q-norms of all the blocks: $||A_{ij}||_q$
- Let's call that the Norm Compression of A:

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots \\ A_{21} & A_{22} & \dots \\ \vdots & \vdots & \ddots \end{pmatrix} \quad \mapsto \quad \begin{pmatrix} ||A_{11}||_q & ||A_{12}||_q & \dots \\ ||A_{21}||_q & ||A_{22}||_q & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

- Q: What can you then tell me about $||A||_q$?

Norm Compression

- Suppose I have a block partitioned matrix $A = [A_{ij}]$
- ...and I give you the Schatten q-norms of all the blocks: $||A_{ij}||_q$
- Let's call that the Norm Compression of A:

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots \\ A_{21} & A_{22} & \dots \\ \vdots & \vdots & \ddots \end{pmatrix} \quad \mapsto \quad \begin{pmatrix} ||A_{11}||_q & ||A_{12}||_q & \dots \\ ||A_{21}||_q & ||A_{22}||_q & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

- Q: What can you then tell me about $||A||_q$?
- A: You could give me upper and lower bounds on $||A||_q$ using Norm Compression Inequalities (NCI's)!

Why is this interesting?

- Quantum mechanics (my main motivation).
  - State of a particle: vector in a Hilbert space $\mathcal{H}$.
  - State of N particles: vector in $\mathcal{H} = \mathcal{H}_1 \otimes \ldots \otimes \mathcal{H}_N$.
  - State vector need not be a tensor product: Entanglement.
  - If particles are in different labs, observables are local, and only local basis transformations are physically meaningful.
  - Partitioning of $\mathcal{H}$ comes out naturally.
  - Trace norms directly related to entropy (and friends).
- Numerical Analysis: approximating norms of really quite big matrices.
- Computer Science: bounds on “formula size” (Kolmogorov complexity).

But is this feasible?

- At first sight it seems that partitioning messes up the matrix.
- Why would there be any simple relation between, say, the spectra of the blocks and the spectrum of the full matrix?
- Surprising fact: NCI's do yield reasonable (=short) answers.

Easiest case: the Frobenius norm

- The situation is very simple for the Frobenius norm

$$||A||_2^2 = \mathrm{Tr}(A^*A) = \sum_{ij} |a_{ij}|^2.$$

- Indeed, for any partitioning $A = [A_{ij}]$,

$$||A||_2^2 = \sum_{i,j} ||A_{ij}||_2^2.$$

That is: equality!

- But what about other q-norms?

Pinching Inequality

- Let A be symmetrically partitioned: $A = [A_{ij}]$
- The Pinching of A is obtained by setting all off-diagonal blocks to 0:

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots \\ A_{21} & A_{22} & \dots \\ \vdots & \vdots & \ddots \end{pmatrix} \quad \mapsto \quad \begin{pmatrix} A_{11} & 0 & \dots \\ 0 & A_{22} & \dots \\ \vdots & \vdots & \ddots \end{pmatrix} = \bigoplus_i A_{ii}.$$

- The Pinching Inequality says that any unitarily invariant norm is reduced under pinchings:

$$|||A||| \geq |||\bigoplus_i A_{ii}|||.$$

- For Schatten q-norms, this is an NCI:

$$||A||_q^q \geq ||\bigoplus_i A_{ii}||_q^q = \sum_i ||A_{ii}||_q^q.$$

An Upper Bound

- There is a complementary NCI for *positive semidefinite* (PSD) symmetrically partitioned block matrices:

$$||A||_q \leq \sum_i ||A_{ii}||_q.$$

- Ref: Horn and Johnson's “Topics in Matrix Analysis”, p. 217 *Problem* 22.

An Upper Bound

- There is a complementary NCI for *positive semidefinite* (PSD) symmetrically partitioned block matrices:

$$||A||_q \leq \sum_i ||A_{ii}||_q.$$

- Ref: Horn and Johnson's "Topics in Matrix Analysis", p. 217 *Problem* 22.
- When A is not PSD, there is no NCI upper bound in terms of only the diagonal block norms $||A_{ii}||_q$.
- Reason: if one of the off-diagonal block norms goes off to ∞, so does $||A||_q$.
- That can't happen when A is PSD because then $||A_{ij}||_q^2 \leq ||A_{ii}||_q \, ||A_{jj}||_q$.

NCI's exploiting all blocks

- For a general $d \times d$ partitioned matrix $A = [A_{ij}]$, with $1 \leq i, j \leq d$, one has, for $1 \leq q \leq 2$:

$$d^{q-2} \sum_{i,j} ||A_{ij}||_q^q \leq ||A||_q^q \leq \sum_{i,j} ||A_{ij}||_q^q.$$

- The reversed inequalities hold for $q \geq 2$.
- The split-up in two cases ($q < 2$, $q > 2$) is a common phenomenon; likely to happen when an NCI reduces to equality for $q = 2$.
- Ref: Bhatia and Kittaneh, Math. Ann. **287**, 719–726 (1990).
- Note: ℓ_q norm of the *vectorised* norm compression $||(||A_{11}||_q, ||A_{12}||_q, \ldots)||_q$.

An NCI for 2X2 PSD matrices

- What about having the norm compression in matrix form?
- C. King (Commun. Math. Phys. 242, 531–545 (2003)) found an NCI for PSD 2×2 block matrices:

$$\left\| \begin{pmatrix} B & C \\ C^* & D \end{pmatrix} \right\|_q \geq \left\| \begin{pmatrix} ||B||_q & ||C||_q \\ ||C||_q & ||D||_q \end{pmatrix} \right\|_q, \quad 1 \leq q \leq 2,$$

  while the reversed inequality holds for $q \geq 2$.

Strong Sharpness

- In the setting of NCI's, it makes sense to ask for the *sharpest possible* bounds.
- Strongly sharp NCI's can be saturated for any allowed choice of the "constituent quantities" of the bound.
- Both the Pinching and H&J inequality are strongly sharp: for any possible choice of $a_i \geq 0$, there is an A such that $||A_{ii}||_q = a_i$ and equality holds in the NCI.
- King's bound is also strongly sharp. Take positive, scalar blocks.
- Bhatia and Kittaneh's "vectorised" bounds are *not* strongly sharp.

A Companion for King's Inequality

- Recall,

$$\left\|\begin{pmatrix} B & C \\ C^* & D \end{pmatrix}\right\|_q \geq \left\|\begin{pmatrix} ||B||_q & ||C||_q \\ ||C||_q & ||D||_q \end{pmatrix}\right\|_q, \quad 1 \leq q \leq 2,$$

  while the reversed inequality holds for $q \geq 2$.

A Companion for King’s Inequality

- Recall,

$$\left\| \begin{pmatrix} B & C \\ C^* & D \end{pmatrix} \right\|_q \geq \left\| \begin{pmatrix} ||B||_q & ||C||_q \\ ||C||_q & ||D||_q \end{pmatrix} \right\|_q, \quad 1 \leq q \leq 2,$$

  while the reversed inequality holds for $q \geq 2$.

- This invites the question: What about an *upper bound* for $1 \leq q \leq 2$, and a lower bound for $q \geq 2$?

A Companion for King’s Inequality

- Recall,

$$\left|\left|\begin{pmatrix} B & C \\ C^* & D \end{pmatrix}\right|\right|_q \geq \left|\left|\begin{pmatrix} ||B||_q & ||C||_q \\ ||C||_q & ||D||_q \end{pmatrix}\right|\right|_q, \quad 1 \leq q \leq 2,$$

while the reversed inequality holds for $q \geq 2$.

- This invites the question: What about an *upper bound* for $1 \leq q \leq 2$, and a lower bound for $q \geq 2$?

- The answer is (KA05):

$$\left|\left|\begin{pmatrix} B & C \\ C^* & D \end{pmatrix}\right|\right|_q^q \leq ||B||_q^q + ||D||_q^q + (2^q - 2)||C||_q^q,$$

and the reversed inequality for $q \geq 2$.

That strange factor...

- Factor $(2^q - 2)$ looks rather weird, but...
- For $q = 1$ and $q = 2$ we know

$$\left\| \begin{pmatrix} B & C \\ C^* & D \end{pmatrix} \right\|_1 = ||B||_1 + ||D||_1$$

$$\left\| \begin{pmatrix} B & C \\ C^* & D \end{pmatrix} \right\|_2^2 = ||B||_2^2 + ||D||_2^2 + 2||C||_2^2.$$

- The factor $2^q - 2$ on $||C||_q^q$ interpolates between 0 and 2.
- One can make even more sense out of it by reformulating the inequality.
- W.l.o.g. one can take $C = C^* \geq 0$ (via polar decomposition).

Reformulation

- Now note the following:

$$\operatorname{Tr}\begin{pmatrix} C & C \\ C & C \end{pmatrix}^q = 2^q \operatorname{Tr} C^q, \quad \operatorname{Tr}\begin{pmatrix} C & 0 \\ 0 & C \end{pmatrix}^q = 2 \operatorname{Tr} C^q.$$

- The inequality can thus be rewritten as

$$\operatorname{Tr}\begin{pmatrix} B & C \\ C & D \end{pmatrix}^q - \operatorname{Tr}\begin{pmatrix} B & 0 \\ 0 & D \end{pmatrix}^q \leq \operatorname{Tr}\begin{pmatrix} C & C \\ C & C \end{pmatrix}^q - \operatorname{Tr}\begin{pmatrix} C & 0 \\ 0 & C \end{pmatrix}^q.$$

- Both sides are non-negative, since $\begin{pmatrix} B & 0 \\ 0 & D \end{pmatrix}$ is a pinching of $\begin{pmatrix} B & C \\ C & D \end{pmatrix}$.
- The left-hand side is a measure of the norm decrease caused by this pinching.
- The inequality says that, when fixing C and constraining B and D to keep A positive, this norm decrease is maximal when $B = D = C$.

Strong Sharpness

- For scalar blocks, my bound is *not* strongly sharp, because King's bound is an equality in that case.
- My bound is strongly sharp when going to blocks of size at least 2×2, provided $||C||_q \leq ||B||_q, ||D||_q$.
- By a theorem of Horn and Mathias, positivity of A implies $||C||_q^2 \leq ||B||_q||D||_q$.
- For $||B||_q \leq ||C||_q \leq (||B||_q||D||_q)^{1/2}$ one can find a better bound (the form is known, the proof isn't).

Proof of the inequality

- Very intricate and rather long proof (Is there an easier one???)
- Does not fit into the margin of this page
- See LAA **413**, 155-176 (2006).

- But let’s now discuss some unknown results...

Hanner's Inequality

- Hanner's inequality for ℓ_q function spaces (f and g are functions), $1 \leq q \leq 2$:

$$||f+g||_q^q + ||f-g||_q^q \geq (||f||_q + ||g||_q)^q + |\,||f||_q - ||g||_q|^q,$$

while for $2 \leq q$, the inequality is reversed.

- Can f and g be replaced by matrices A, B, and ℓ_q-norm by Schatten q-norm?

$$||A+B||_q^q + ||A-B||_q^q \geq (||A||_q + ||B||_q)^q + |\,||A||_q - ||B||_q|^q?$$

- This can be rewritten as an NCI.

Hanner's Inequality

- Original form:

$$||A+B||_q^q + ||A-B||_q^q \geq (||A||_q + ||B||_q)^q + |\,||A||_q - ||B||_q|^q.$$

- Apply $||X \oplus Y||_q^q = ||X||_q^q + ||Y||_q^q$:

$$\left|\left|\begin{pmatrix} A+B & 0 \\ 0 & A-B \end{pmatrix}\right|\right|_q^q \geq \left|\left|\begin{pmatrix} ||A||_q + ||B||_q & 0 \\ 0 & ||A||_q - ||B||_q \end{pmatrix}\right|\right|_q^q.$$

- Apply unitary conjugation with Hadamard matrix $U = \begin{pmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{I} & -\mathbf{I} \end{pmatrix} / \sqrt{2}$:

$$\left|\left|\begin{pmatrix} A & B \\ B & A \end{pmatrix}\right|\right|_q \geq \left|\left|\begin{pmatrix} ||A||_q & ||B||_q \\ ||B||_q & ||A||_q \end{pmatrix}\right|\right|_q.$$

Hanner's Inequality

- Proven in the following instances (Ball, Carlen and Lieb):

  1. $A+B$ and $A-B$ are positive semidefinite.
  2. General A and B: only for $1 \leq q \leq 4/3$, $q=2$, and $4 \leq q \leq \infty$.

- Red-Herring Removal Procedure: Generalise!

Big Fat Conjecture

- Let T be a general matrix partitioned in $2 \times N$ blocks:

  $$T = \begin{pmatrix} A_1 & A_2 & \cdots & A_N \\ B_1 & B_2 & \cdots & B_N \end{pmatrix}.$$

  Then I conjecture the following NCI, for $q \geq 2$:

  $$||T||_q \leq \left|\left| \begin{pmatrix} ||A_1||_q & ||A_2||_q & \cdots & ||A_N||_q \\ ||B_1||_q & ||B_2||_q & \cdots & ||B_N||_q \end{pmatrix} \right|\right|_q$$

  while for $1 \leq q \leq 2$ the ordering of the inequality is reversed.

- Hanner's inequality: $N = 2$, $B_2 = A_1$, $B_1 = A_2$.

- King's NCI: $N = 2$, $T \geq 0$.

Special cases

- Proven for:
  - All B_k zero
  - All A_k and B_k are rank 1
  - $A_k = \alpha_k X$ and $B_k = \beta_k Y$, for some scalars α_k, β_k and some matrices X, Y.
  - $q \geq 4$; follows easily from King's NCI.
- All A_k and B_k diagonal: would follow from a strange convexity statement: the function

$$T \mapsto \mathrm{Tr}\left|T^{\circ\frac{1}{q}}\right|^q$$

defined over $M_{2,N}(\mathbb{R}^+)$ is convex for $1 \leq q \leq 2$ and concave for $q \geq 2$.

- Proven by King and Nathanson for $N = 2$ (LAA **389**, 77–93 (2004)).

Conclusion

Norm Compression Inequalities are:

- beautiful
- useful
- challenging