Extending non-commutative Schatten norms to product spaces

Christopher King

Thursday, June 4, 2009

1 Schatten norm

The Schatten norm is a non commutative extension of the l_p norm. Let M_n be the set of $n \times n$ complex valued matrices; given $A \in M_n$, let A^* denote its adjoint.
Recall the following facts:

- a matrix $P \in M_n$ is called *positive semidefinite* (notation: $P \geqslant 0$), if $P = P^*$ and $\langle x, Px \rangle \geqslant 0 \; \forall x \in \mathbb{C}^n$;
- if $P \geqslant 0$, then $P = UDU^*$ where $UU^* = U^*U = I$ and $D = diag(\lambda_1, \ldots, \lambda_n)$ with $\lambda_i \geqslant 0$;
- if $P \geqslant 0$, then $P^p := UD^pU^*$ where $D^p := diag(\lambda_1^p, \ldots, \lambda_n^p)$.

Given $A \in M_n$, we define the absolute value of A by setting $|A| := (A^*A)^{1/2}$, which makes sense because $A^*A \geqslant 0$. For any $A \in M_n$ and real number $p \geqslant 1$, we define a Schatten norm by

$$\|A\|_p := (\operatorname{tr}|A|^p)^{1/2} = (\operatorname{tr}(A^*A)^{p/2})^{1/p}.$$

The Schatten norms satisfy the following properties for any pair of real numbers $p, p' \geqslant 1$ such that $1/p + 1/p' = 1$:

1. triangle inequality: $\|A + B\|_p \leqslant \|A\|_p + \|B\|_p$;
2. Hölder's inequality: $|\operatorname{tr}(AB)| \leqslant \|A\|_p \|B\|_{p'}$;
3. duality:
$$\|A\|_p = \sup_B \{|\operatorname{tr}(AB)| : \|B\|_{p'} \leqslant 1\};$$

4. for any diagonal matrix $D = diag(s_1, \ldots, s_n)$,

$$\|D\|_p = \left(\sum_{i=1}^{n} |s_i|^p \right)^{1/p}.$$

The set M_n with the norm $\|\cdot\|_p$ is a Banach space; we denote it by S_p. It is a non commutative version of l_p. Its dual is given by $(S_p)^* \cong S_{p'}$. Moreover by 4) above, l_p is a subalgebra of the diagonal matrices.

2 Product algebra

We now consider the tensor product of two matrix algebras. There is an isomorphism

$$M_n \otimes M_m \cong M_{nm}.$$

For example, if $n = 2$, an element in $M_2 \otimes M_m$ is a $2m \times 2m$ matrix that can be represented by

$$\left(\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right)$$

with $A, B, C, D \in M_m$. If M_n has a norm with index q and M_m has a norm with index p, how can we give a norm on the tensor product? Let us consider the commutative case first. Let $D = diag(s_{ij}) \in M_n \otimes M_m$, $i = 1, \ldots, n$, $j = 1, \ldots, m$. We can regard D as a diagonal block matrix

$$D = \begin{pmatrix} D_1 & & & \\ & D_2 & & \\ & & \ddots & \\ & & & D_n \end{pmatrix}$$

where $D_1, \ldots, D_n \in M_m$ are all diagonal matrices. If we apply the $\|\cdot\|_p$ to each matrix D_i, we obtain the $n \times n$ diagonal matrix

$$\begin{pmatrix} \|D_1\|_p & & & \\ & \|D_2\|_p & & \\ & & \ddots & \\ & & & \|D_n\|_p \end{pmatrix}$$

to which we can then apply the $\|\cdot\|_q$ norm. Ultimately we have

$$\|D\|_{(q,p)} = \left(\sum_{i=1}^{n} \left(\sum_{j=1}^{m} |s_{ij}|^p \right)^{q/p} \right)^{1/q}.$$

3 Extending to $M_n \otimes M_m$

The goal is to extend the norm introduced in the previous section to the whole algebra $M_n \otimes M_m$. To achieve this goal we use a tool called partial traces.
The partial traces are operators

$$\begin{aligned} \operatorname{tr}_1 : M_n \otimes M_m &\longrightarrow M_m \\ \operatorname{tr}_2 : M_n \otimes M_m &\longrightarrow M_n \end{aligned}$$

such that $\forall B \in M_n \otimes M_m$:

$$\begin{aligned} \forall A \in M_m,\ \operatorname{tr}_2(A \operatorname{tr}_1 B) &= \operatorname{tr}((I \otimes A)B), \\ \forall A \in M_n,\ \operatorname{tr}_1(A \operatorname{tr}_2 B) &= \operatorname{tr}((A \otimes I)B). \end{aligned}$$

For example, if $n = 2$,

$$\begin{aligned} \operatorname{tr}_2 \begin{pmatrix} A & B \\ C & D \end{pmatrix} &= \begin{pmatrix} \operatorname{tr} A & \operatorname{tr} B \\ \operatorname{tr} C & \operatorname{tr} D \end{pmatrix}, \\ \operatorname{tr}_1 \begin{pmatrix} A & B \\ C & D \end{pmatrix} &= A + D. \end{aligned}$$

Now, to extend the norm, we replace sums with partial traces. For $A \in M_n \otimes M_m$ we get

$$(\operatorname{tr}_1(\operatorname{tr}_2 |A|^p)^{q/p})^{1/q}.$$

To be a norm, the expression above should satisfy the triangle inequality but it does not. Since the expression is homogeneous of degree one, requiring that it satisfies the triangle inequality is equivalent to saying it should be a convex operator in A. But $A \mapsto |A|$ is not convex. For example, if

$$A = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

then we have

$$|A| = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \qquad |B| = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

and

$$|A + B| = \sqrt{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

but the matrix $|A| + |B| - |A + B|$ is not positive definite.

4 A partial fix (Carlen, Lieb 2008)

Define a function

$$\psi_{p,q}(A) := (\mathrm{tr}_1(\mathrm{tr}_2\,|A|^p)^{q/p})^{1/q}$$

for any positive matrix A (i.e. $A = A^* \geqslant 0$).

Theorem. a) For $1 \leqslant p \leqslant 2$, $q \geqslant 1$, $\psi_{p,q}$ is convex;

b) for $0 \leqslant p \leqslant q \leqslant 1$, $\psi_{p,q}$ is concave;

c) for $p > 2$, $\psi_{p,q}$ is neither convex nor concave.

This result gives a norm on positive definite matrices for $1 \leqslant p \leqslant 2$. To define a norm for the full algebra, set for any self adjoint matrix X (i.e. $X = X^*$):

$$|||X||| := \inf_{\substack{A,B \geqslant 0: \\ X = A - B}} \{\psi_{p,q}(A) + \psi_{p,q}(B)\}$$

Lemma. $|||\cdot|||$ satisfies the triangle inequality.

The proof goes as follows: if $X = A - B$, $Y = C - D$, with $A, B, C, D \geqslant 0$, then $X + Y = A + C - (B + D)$ hence

$$\begin{aligned} |||X + Y||| &\leqslant \psi_{p,q}(A + C) + \psi_{p,q}(B + D) \leqslant \\ &\leqslant \psi_{p,q}(A) + \psi_{p,q}(C) + \psi_{p,q}(B) + \psi_{p,q}(D) \leqslant \\ &\leqslant |||X||| + \epsilon + |||Y||| + \epsilon \end{aligned}$$

and since this holds for any ϵ, the claim follows.
Finally, for a general matrix $A \in M_n \otimes M_m$, we consider the matrix

$$\begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix} \in M_n \otimes M_{2m}$$

and then set

$$\|A\|_{CL} := \frac{1}{2} \left|\left|\left| \begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix} \right|\right|\right| .$$

$\|\cdot\|_{CL}$ is a norm on $M_n \otimes M_m$; moreover, when restricted to diagonal matrices, it is the same as applying the p norm followed by the q norm. Still it requires $1 \leqslant p \leqslant 2$ and other properties like Hölder's inequality do not follow immediately. In other words, it works but it is unsatisfactory.

5 Haagerup norm on $M_n \otimes M_m$

Define

$$\|Y\|_h = \inf\left\{\sum_i \|A_i\|_q \|B_i\|_p : Y = \sum_i A_i \otimes B_i\right\}.$$

It is easy to check that this is a norm. Let us check the triangle inequality holds. Let

$$X = \sum_i A_i \otimes B_i, \qquad Y = \sum_j C_j \otimes D_j,$$

then we have

$$X + Y = \sum_i A_i \otimes B_i + \sum_j C_j \otimes D_j$$

and this implies

$$\begin{aligned}\|X+Y\|_h &\leqslant \sum_i \|A_i\|_q \|B_i\|_p + \sum_j \|C_j\|_q \|D_j\|_p \leqslant \\ &\leqslant \|X\|_h + \epsilon + \|Y\|_h + \epsilon.\end{aligned}$$

When $p = q$, we would like $\|\cdot\|_h$ to reduce to the Schatten norm but it does not. It is easy to see that, for $p = q$, $\|Y\|_p \leqslant \|Y\|_h$. In fact, if

$$Y = \sum_i A_i \otimes B_i,$$

then

$$\|Y\|_p \leqslant \sum_i \|A_i \otimes B_i\|_p \leqslant \|Y\|_h + \epsilon.$$

To get equality, we would need $A_i \otimes B_i$ to be proportional to Y for every index i, which is impossible when Y is a product.

6 Solution: a "souped up" Haagerup norm

Assume $1 \leqslant q \leqslant p \leqslant \infty$ and let r be such that $1/r = 1/q - 1/p$. Define

$$\|Y\|_{(q,p)} := \inf\left\{\sum_i \|A_i\|_{2r} \|Z_i\|_p \|B_i\|_{2r} : Y = \sum_i (A_i \otimes I) Z_i (B_i \otimes I)\right\}.$$

If $p \leqslant q$, define

$$\|Y\|_{(p,q)} := \sup_{A,B}\left\{\frac{\|(A \otimes I) Y (B \otimes I)\|_q}{\|A\|_{2r} \|B\|_{2r}}\right\}.$$

Theorem. Let $Y, W \in M_n \otimes M_m$ and p', q' be such that $1/p + 1/p' = 1$ and $1/q + 1/q' = 1$. Then

a) triangle inequality: $\|Y + W\| \leqslant \|Y\| + \|W\|$;

b) Hölder's inequality: $|\operatorname{tr}(YW)| \leqslant \|Y\|_{(p,q)} \|W\|_{(p',q')}$;

c) duality:

$$\|Y\|_{(p,q)} = \sup\left\{|\operatorname{tr}(YW)| : \|W\|_{(p',q')} \leqslant 1\right\}$$
$$\|Y\|_{(q,p)} = \sup\left\{|\operatorname{tr}(YW)| : \|W\|_{(q',p')} \leqslant 1\right\};$$

d) $\|Y_1 \otimes Y_2\|_{(p,q)} = \|Y_1\|_p \|Y_2\|_q$;

e) $\|Y\|_{(p,p)} = \|Y\|_p$;

f) if Y is diagonal, then $\|Y\|_{(p,q)} = (\operatorname{tr}_1 (\operatorname{tr}_2 |Y|^q)^{p/q})^{1/p}$.