Lecture 2

THE TRIANGLE INEQUALITY

The triangle inequality is surely a basic constituent of mathematics. In this lecture I shall discuss a matrix valued version of it, together with some related facts.

If $A$ is a complex matrix, $n \times n$, define

$$|A| = \text{diag}(s_1, s_2, \ldots, s_n)$$

where $s_1, \ldots, s_n$ are the singular values of $A$. Alternatively, one might define $|A|$ as

$$|A| = (AA^*)^{1/2}, \quad \text{or as } |A| = (A^*A)^{1/2},$$

but I shall generally prefer the diagonal matrix of singular values.

If $A$ and $B$ are Hermitian, define $A \preceq B$ to mean that $B - A$ is positive semidefinite. This is often called the Loewner order.

Now, for arbitrary matrices $A$ and $B$, one would hope that,

$$|A + B| \leq |A| + |B|.$$  

But there are easy counterexamples for this, just take $A$ and $B$ to be $2 \times 2$ and highly non-Hermitian. Nonetheless, there is a matrix version of the triangle inequality in which the role of noncommutativity is plainly visible:

$$|A + B| \leq |U|A|U^* + V|B|V^*$$

for unitary matrices $U$ and $V$ that depend on $A$ and $B$. The discovery of this matrix triangle inequality preceded the exponential function investigations in the previous lecture, and was a strongly motivating factor to initiate those investigations.

Our results, for the triangle inequality and for the exponential function, illustrate a general principle: If your favorite scalar equality or inequality becomes false when written in a matrix valued form, then it is false because you haven’t properly allowed for
noncommutativity. To allow for noncommutativity, insert a similarity or a unitary similarity wherever natural. The resulting inequality is then either true (possibly hard to prove) or there are immediate counterexamples. That is, if a counterexample isn't reasonably at hand, a valid theorem has been discovered, although its proof still has to be found.

The matrix valued triangle inequality also has a case of equality that naturally generalizes the scalar case of equality. For complex scalars, equality holds if and only if the two complex numbers lie on the same ray through the origin. For matrices, equality means the two matrices $A$ and $B$ have polar factorizations with a common unitary factor. (This has to be stated precisely.)

The matrix triangle inequality and its case of equality extends without difficulty to matrices with quaternion elements. The singular value decomposition holds for matrices over the quaternions, the notion of Hermitian makes sense, and everything works, including the case of equality. It is the matrix generalization of the triangle inequality satisfied by the usual norm on quaternions.

Nonetheless, there are some mysteries: What is the role of $\mathbf{U}$ and $\mathbf{V}$? More precisely, how do they depend on $\mathbf{A}$ and $\mathbf{B}$? For the exponential function case, formulas were developed for the $s$ and $t$ in

$$
\begin{align*}
x \cdot y &= sxs^{-1} + tyt^{-1} \\
e \cdot e &= e
\end{align*}
$$

Good formulas are needed for the $\mathbf{U}$ and $\mathbf{V}$ in the matrix triangle inequality. Perhaps one should write $\mathbf{U} = e^0$, $\mathbf{V} = e^\tau$, and look for somehow natural formulas for $\rho$ and $\tau$. I haven't tried to work out this very reasonable idea, although I have had it for a long time.

A key step is in the proof of the matrix triangle inequality is this: Take the eigenvalues of $\frac{1}{2}(\mathbf{A} + \mathbf{A}^*)$ in decreasing order, $\lambda_1 \geq \lambda_2 \geq \ldots$ (these are usually called the real singular values of $\mathbf{A}$.) Then

$$
\lambda_i \leq s_i \quad \text{for all } i.
$$

Thus this is an inequality between the real and the absolute singular values of a matrix. It was discovered by Fan and Hoffman and rediscovered by Fan. Here is the simplest proof, which points to a connection with the Schubert calculus to be discussed in a later
lecture: Let \( x \) be a unit vector in the span of the eigenspaces of \( \frac{1}{2}(A+A^*) \) belonging to \( \lambda_1, \ldots, \lambda_l \) and simultaneously in the span of the eigenspaces of \( AA^* \) for \( s_i, \ldots, s_n \). This vector \( x \) exists because the sum of the dimensions of these two subspaces is \( n+1 \), so their intersection must have a nonzero vector. Then

\[
\lambda_i \leq \langle x^*(A+A^*)x = \text{Re} x^*Ax \\
\leq \|x^*Ax\| \leq \|Ax\| \|x\| = (Ax, Ax)^{1/2} \\
= (A^*Ax, x)^{1/2} \leq s_i.
\]

Knowing that \( \lambda_i \leq s_i \) for all \( i \), it follows that the diagonal form of \( \frac{1}{2}(A+A^*) \) is dominated element by element by the diagonal form \( |A| \). From this one immediately gets that

\[
\frac{1}{2}(A+A^*) \leq U|A|U^*
\]

for some unitary \( U \), dependent on \( A \). Having this fact, it is just a short calculation to deduce the matrix valued triangle inequality: Let \( C = A + B \), and without loss of generality assume \( C = |C| \). Then \( C = A + B \) implies

\[
C = \frac{1}{2}(A+A^*) + \frac{1}{2}(B+B^*) \leq U|A|U^* + V|B|V^*
\]

for suitable unitary \( U \) and \( V \), and we are done.

In number theory, the usual absolute value stands on an equal footing with the \( p \)-adic absolute values, usually called the \( p \)-adic valuations. Here \( p \) is a prime number. To define it, if \( q \) is a rational number, extract the power of \( p \) from numerator or denominator, so that

\[
q = p^e \frac{m}{n}, \quad m \text{ and } n \text{ integers not divisible by } p, \text{ and } e \text{ is an integer.}
\]

Then define the \( p \)-adic absolute value of \( q \), denoted by \( \|q\|_p \), as

\[
\|q\|_p = p^e, \text{ where } p \text{ is a fixed number in } (0,1).
\]
This is a standard definition in number theory. Usually one takes \( p = \frac{1}{p} \). Also define \( \text{l}0\text{l}_p = 0 \). Then this \( p \)-adic absolute value (= \( p \)-adic valuation) satisfies the triangle inequality and is multiplicative, easily verified:

\[
|a+b|_p \leq |a|_p + |b|_p, \quad |ab|_p = |a|_p |b|_p.
\]

The \( p \)-adic numbers are (by definition) the closure of the rational numbers under the topology given by the \( p \)-adic metric. They are genuinely beautiful, a person hasn't really enjoyed life until he/she has proved something in the \( p \)-adic domain. Here is a nice and classical theorem:

In the \( p \)-adic numbers, -1 is a sum of 2 squares if \( p \equiv 3 \pmod{4} \), -1 is a square if \( p \equiv 1 \pmod{4} \), and -1 is a sum of 4 squares if \( p = 2 \). (Shortest representations.)

So the \( p \)-adic numbers aren't ordered, making them mysterious in the eyes of the naive beholder. When using real numbers, it is surprising how comforting the mental image of the real line is. There is no corresponding comfort in the \( p \)-adic domain.

The \( p \)-adic valuation satisfies a strong triangle inequality:

\[
|a+b|_p \leq \max(|a|_p,|b|_p) \leq |a|_p + |b|_p.
\]

The first inequality here is called the nonarchimedean or strong version of the usual triangle inequality, with the inequality between the extreme left and right parts sometimes called the weak triangle inequality.

Can we make a matrix valued version of this scalar triangle inequality, either strong or weak?

A well known fact is that if \( A \) is a matrix of integers, there exist unimodular integral matrices \( U \) and \( V \) such that \( UAV \) is diagonal, with the diagonal elements forming a divisibility chain:

\[
UAV = \text{diag}(s_1, s_2, \ldots, s_n)
\]

with \( s_1 | s_2 | \ldots | s_n \).
This is the famous Smith form theorem. The symbol $|$ here denotes divisibility of integers, and is not to be confused with the absolute value symbol. The diagonal elements informing the divisibility chain are usually called the invariant factors of $A$; they are unique to within unit multipliers. The same diagonalization theorem holds for rational matrices $A$ (but $U$, $V$ are still integral.) Now we define the $p$-adic matrix absolute value of an integral or rational matrix $A$, denoted by $|A|_p$, in this way:

$$|A|_p = \text{diag}(|s_1|_p, \ldots, |s_n|_p).$$

That is, we form the diagonal matrix of the $p$-adic values of the invariant factors of $A$. So, the $p$-adic absolute value of a matrix is a real symmetric matrix, just as the $p$-adic absolute value of a number is a real number.

Now, our question is: Is there a matrix valued version of the scalar weak or strong triangle inequality? For the weak triangle inequality, this would be

$$|A+B|_p \leq |A|_p + |B|_p,$$

where $\leq$ means that the difference of the two sides is semidefinite.

Answer: No, we didn't allow for noncommutativity.

Bringing noncommutativity in according to the standard rule, we get:

**Theorem:** There exist unitary matrices $U$, $V$ such that

$$|A+B|_p \leq U|A|_p U^* + V|B|_p V^*.$$  

(In fact, $U$ and $V$ can be taken to be permutation matrices.)

This is the weak matrix valued $p$-adic triangle law. It is natural also to expect a matrix version of the strong triangle law, and there is one:

**Theorem:** $|A + B|_p \leq \max(|A|_p, |B|_p)$
where max is the $n \times n$ diagonal matrix formed using the $n$ largest of the $2n$ numbers comprising the diagonal elements of $|A|_p$, $|B|_p$. The complete statement incorporating the weak and strong matrix $p$-adic triangle laws then is:

$$|A+B|_p \leq \max(|A|_p,|B|_p) \leq U|A|_pU^* + V|B|_pV^*$$

for $U$, $V$ suitable unitary (permutation) matrices. As always, $\leq$ signifies the Loewner order.

It is more than a little intriguing that the noncommutativity (inclusion of the unitaries $U$, $V$) occurs in the nonarchimedean part of the inequality, and not in the archimedean part. (These terms refer to the fact that the usual absolute value of a rational number is a nonarchimedan valuation, whereas the $p$-adic absolute value is archimedean.)

A question that seems not to be answered is the representability of a $p$-adic matrix as a sum of squares of $p$-adic matrices. In particular, this for $-I$.

Now we turn to operators on Hilbert space. Diagonalization is a delicate matter here, so avoid it by defining $|A|$ as

$$|A| = (AA^*)^{1/2}.$$

The triangle inequality for operators was investigated by Akemann, Pedersen, and Anderson (1982). They found:

**Theorem:** There exist isometries $U$ and $V$ such that

$$|A+B| \leq U|A|U^* + V|B|V^*$$

where $U$ and $V$ depend on $A$ and $B$. But in infinite dimensions isometries may be into, not onto, that is, not be unitaries. If one wants unitaries, the best that could be done was this: For each $\varepsilon > 0$ there exist unitaries $U$ and $V$ such that

$$|A+B| \leq U|A|U^* + V|B|V^* + \varepsilon I.$$
Akemann et al did not produce an example in which the $\varepsilon$ is actually needed. However, the key step in their proof of the triangle inequality was that

$$\frac{1}{2}(A+A^*) \leq |U|A|U^* + \varepsilon I$$

for some unitary $U$ dependent on $A$. (The $\varepsilon$ isn't needed if we just want $U$ to be an isometry). They did produce an example in which the $\varepsilon$ is needed in this inequality if $U$ is to be a unitary. It is also noteworthy that their proof of this inequality was of a purely functional-analytic nature, avoiding all mention of spectra. This perhaps suggests that the kinds of inequalities we are considering are independent of spectral calculations.

In the finite dimensional case, the matrix valued triangle inequality holds for quaternion entried matrices. The proof is the same as in the complex case, making proper allowance for the noncommutativity of the quaternion scalars. There exists a theory of Hilbert spaces over the quaternions, and it's not completely analogous to the Hilbert space theory over the complexes. So it would be worthwhile doing the operator valued triangle inequality for algebras of operators on quaternion Hilbert spaces. It may be just the same, or perhaps it isn't.

There is also p-adic Hilbert space theory, another interesting and fascinating topic, perhaps regarded as a small specialized corner of functional analysis. Nevertheless, it exists, and the p-adic version of the triangle inequality needs to be lifted up to this context.

There is another aspect to this theory. With the aid of the matrix or operator triangle inequality, the following can be proved:

$$\det(I + |A+B|) \leq \det(I+|A|)\det(I+|B|).$$

The best presentation of this is in Akemann et al.

E. Lieb generalized this scalar inequality to a large class of scalar valued functions, immanants, in place of $\det$. Now immanants are generalized matrix functions, that is, are defined using traces of representations, that is, characters. In representation theory one has the fact that two representations are equivalent if and only if they have the same character. I therefore suspect that the matrix triangle inequality is equivalent to a family of scalar inequalities
involving immanants or other scalar functions defined using characters. But this is vague, and I don't really know how to do it. Nevertheless, I think there is something here. The basic and vague idea is that for some scalar functions \( f \), we have \( f(|A+B|) \leq f(|A|) + f(|B|) \), and by character theory this must imply, and be implied by, the matrix triangle inequality.

The matrix valued triangle inequality has been extended by functional analysts beyond the point to which Akemann et al took it. But the extensions are so abstract that I can no longer recognize the theorem to be the one I found. It sometimes is more psychologically satisfying to take the low road, and not the high one.

Now we turn to another aspect. Any decent (scalar valued) norm must have the property that

\[
\|AB\| \leq \|A\| \|B\|
\]

and we wish now to look at matrix valued versions of this inequality, in the form

\[
|AB| \leq |A||B|,
\]

where \( |\cdot| \) is the matrix valued absolute value. Unfortunately, this is false in dimension 2 or more, as might be expected since noncommutativity hasn't been properly allowed for. But there is a true theorem. For \( n \times n \) matrices,

\[
|AB| \leq \frac{1}{n} \sum_{i=1}^{n} U_i |A||B| U_i^*
\]

for unitaries \( U_i \), \( i = 1, \ldots, n \) dependent on \( A \) and \( B \). However, this is a weak theorem, and probably the correct theorem is:

Conjecture: There exist unitaries \( U \) and \( V \) such that

\[
|AB| \leq \frac{1}{2}(U|A||B|U^* + V|A||B|V^*).
\]

This conjecture cries out for an elementary treatment: One would like to square it, choose the \( U \) and \( V \) so that we have an obvious identity,
then take the square root of this identity, using the fact that for positive semidefinite $H$ and $K$,

$$H \preceq K \implies H^{1/2} \preceq K^{1/2}.$$ 

But I have not been able to make this device work. However, I have evolved two nonelementary strategies for attempting to prove this conjecture, but both use powerful tools (and are incomplete.) Here is one of them. The other may be in a later lecture.

Define monotonically sequences of integers,

$$F = (f_1, \ldots, f_L), \; 1 \leq f_1 < \ldots < f_L \leq n,$$

$$G = (g_1, \ldots, g_L), \; 1 \leq g_1 < \ldots < g_L \leq n,$$

$$H = (h_1, \ldots, h_L), \; 1 \leq h_1 < \ldots < h_L \leq n,$$

to be consistent by induction on $n$, as follows.

For $L = 1$, consistency means $f_1 + g_1 = h_1 + 1$.

For $L > 1$, consistency means

$$f_1 + \ldots + f_L + g_1 + \ldots + g_L = h_1 + \ldots + h_L + \frac{L(L+1)}{2}, \quad \text{and}$$

for consistent sequences $P = (p_1, \ldots, p_k)$, $Q = (q_1, \ldots, q_k)$, $R = (r_1, \ldots, r_k)$ with entries from $[1, L]$,

$$\sum_{j \in P} f_j + \sum_{j \in Q} g_j \leq \sum_{j \in R} h_j + \frac{k(k+1)}{2}$$

Then we have these facts:

(i) If $A, B, C$ are $n \times n$ matrices with $C = AB$, having singular values

$$\alpha_1 \geq \ldots \geq \alpha_n, \quad \beta_1 \geq \ldots \geq \beta_n, \quad \gamma_1 \geq \ldots \geq \gamma_n,$$

respectively, then
\[ \gamma_{h_1} \cdots \gamma_{h_L} \leq \alpha_{f_1} \cdots \alpha_{f_L} \beta_{g_1} \cdots \beta_{g_L} \]

for all consistent sequences F,G,H from \([1,n]\). (With equality when \(L = n\).) And (B.V.Lidskii), if

\[ \gamma_{h_1} + \cdots + \gamma_{h_L} \leq \alpha_{f_1} + \cdots + \alpha_{f_L} + \beta_{g_1} + \cdots + \beta_{g_L} \]

for all consistent sequences F,G,H from \([1,n]\), with equality when \(L=n\), then there exist Hermitian matrices \(A, B, C = A + B\) having spectra \(\alpha, \beta, \gamma\).

Granted this, we would have

\[
\gamma_{h_1} \cdots \gamma_{h_t} = \sqrt{\frac{2}{\gamma_{h_1} \cdots \gamma_{h_t}}} 
\]

\[
\leq \sqrt{\alpha_{f_1} \cdots \alpha_{f_t} \beta_{g_1} \cdots \beta_{g_t} \alpha_{g_1} \cdots \alpha_{g_t} \beta_{f_1} \cdots \beta_{f_t}} 
\]

\[
\leq \frac{1}{2} (\alpha_{f_1} \beta_{f_1} + \alpha_{g_1} \beta_{g_1}) \cdots \frac{1}{2} (\alpha_{f_t} \beta_{f_t} + \alpha_{g_t} \beta_{g_t}) 
\]

for \(t = 1, \ldots, L\). By convexity this implies

\[
\gamma_{h_1} + \cdots + \gamma_{h_L} \leq \frac{1}{2} (\alpha_{f_1} \beta_{f_1} + \alpha_{g_1} \beta_{g_1}) + \cdots + \frac{1}{2} (\alpha_{f_L} \beta_{f_L} + \alpha_{g_L} \beta_{g_L}) 
\]

\[
= \frac{1}{2} (\alpha_{f_1} \beta_{f_1} + \cdots + \alpha_{f_L} \beta_{f_L}) + \cdots + \frac{1}{2} (\alpha_{g_1} \beta_{g_1} + \cdots + \alpha_{g_L} \beta_{g_L}) 
\]

Thus we have the consistent inequalities linking the spectrum of \(|AB|, \frac{1}{2} |AB|, \frac{1}{2} |AB|\). If we can increase the left side and/or decrease the right side so as to get equality in the trace condition, then by Lidskii's theorem we will deduce the existence of unitary U and V meeting the desired condition:
\[ |AB| \leq \frac{1}{2} (|A| |B| U^* + V |A| |B| V^*) \]
as desired.

This plan reduces to two steps:

(i) If \( F, G, H \) are consistent \( L \)-tuples, then there are consistent \( (L-1) \)-tuples \( F', G', H' \) with \( H' \leq H, F \geq F', G' \geq G \). (Inequalities term by term);

(ii) A similar statement involving real sequences satisfying the majorization inequalities with consistent tuples as subscripts but without the equality condition on last inequality: increase the real terms on the lower side, decrease the real terms on the upper side, to preserve all inequalities and get the desired last equality.

It turns out both parts involve the same steps, if the combinatorial part (i) can be done, so can the seemingly different real part. So the proof reduces solely to (i).

Now, in the low dimensional cases, (i) can be done by direct calculation. However, the highly recursive nature of the defining relations for consistent sequences has made the proof hard to do in general. So, one resorts to another tactic.

The tactic is that consistent sequences are generally believed to be described by Young tableau. (More on these in a later lecture.) One is then brought into the combinatorics of tableaux, and immediately led into most intriguing combinatorial problems involving them, which seem different from any presently occurring in the literature. It would carry us too far afield to describe these. The proof can now be reduced to two combinatorial steps in the properties of tableau. I have succeeded in doing one of the two, and am working on the other.

Now, we go back to the \( p \)-adic case. For the \( p \)-adic norm on rational numbers \( a, b \), we have

\[ |ab|_p = |a|_p |b|_p. \]

So, for the matrix valued \( p \)-adic norm, we ask whether

\[ |AB|_p \leq |A|_p |B|_p. \]
But this is false. However, there is a true theorem:

$$|AB|_p \leq \frac{1}{n} \sum_{i=1}^{n} U_i |A|_p |B|_p U_i^*$$

for unitaries $U_1, ..., U_n$ dependent on $A$ and $B$. However, one suspects this theorem is too weak, and that the correct statement is:

$$|AB|_p \leq \frac{1}{2} (|A|_p |B|_p U^* + |V|_p |B|_p V^*)$$

for unitaries $U$ and $V$ dependent on $a$ and $B$. To prove this, one can use the same technique outlined for the real matrix absolute norm inequality. These are divisibility relations satisfied by the invariant factors of a product of matrices, the inequalities describing the spectrum of a sum of Hermitian matrices, plus a bit of convexity. The previous inequalities for the singular values of a product were

$$\gamma_{h_1} \cdots \gamma_{h_L} \leq \alpha_{f_1} \cdots \alpha_{f_L} \beta_{g_1} \cdots \beta_{g_L}$$

for all consistent sequences $F, G, H$. For the invariant factors of a product of integral matrices, these becomes divisibility relations, namely (under a suitable numbering of invariant factors)

$$\gamma_{h_1} \cdots \gamma_{h_L} \mid \alpha_{f_1} \cdots \alpha_{f_L} \beta_{g_1} \cdots \beta_{g_L}$$

where the $\mid$ signifies divisibility, this to hold for all consistent sequences. (The invariant factors are integers, so divisibility makes sense.) These relations will be described in the next lecture.

However, the proof of this matrix valued $p$-adic norm inequality is still not complete, because the steps described for the real matrix norm inequality are still incomplete. However, the proof can be completed in low dimensional cases.

No one has yet considered the matrix valued norm inequalities in infinite dimensional spaces. This seems a very natural problem to look at.
There are other matrix valued inequalities that one can consider, the Cauchy-Schwartz inequality for scalars being a most natural candidate. For scalars, the Cauchy-Schwartz inequality implies the triangle inequality. A good matrix valued version of Cauchy-Schwartz should therefore imply the matrix valued triangle inequality. Although matrix valued versions of Cauchy have been attempted, by me, and perhaps by others, none meets this very natural test. This problem therefore is open.

Another natural question is that for the matrix valued norms, the norm of a submatrix should be smaller than the norm of the matrix containing it. This takes the following form: If

\[
Z = \begin{pmatrix}
Z_{11} & Z_{12} \\
Z_{21} & Z_{22}
\end{pmatrix}
\]

is a partitioned matrix, then we have:

\[
\left| \begin{pmatrix} Z_{11} & 0 \\ 0 & 0 \end{pmatrix} \right| \leq |U| \left| \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix} \right| |U^*|
\]

(note the unfortunately small matrix absolute value symbols \(| |\)) for a unitary \(U\) dependent on \(Z\).

As is invariably the case in these inequalities, the unitary \(U\) cannot be omitted.

The p-adic version of this has not been checked, but probably is true. The Hilbert space version hasn't been looked at, either. Nor has the quaternion version, but it probably is true.

Another theorem involving submatrices is this: Let \(Z = [Z_{ij}]\) be partitioned, \(i, j = i, \ldots, m\). Then, for any \(k \leq m\),

\[
|\text{diag}(Z_{11}, \ldots, Z_{kk}, 0, \ldots, 0)| \leq \frac{1}{k} \sum_{i=1}^{k} U_i |Z| U_i^*
\]

for unitaries \(U_i\) dependent on \(Z\). Although, generally in inequalities of this type the number of terms on the right side is usually one or two only (in the sharpest theorems), in the present case, \(k\) terms are needed, in that there are examples in which it is impossible to write

\[
\text{diag}(Z_{11}, \ldots, Z_{kk}, 0, \ldots, 0)
\]
as any convex combination to fewer than \( k \) terms of the form

\[ U_i Z_i U_i^* \]

for any choice of the unitaries \( U_i \).

The conclusion evolving from this lecture is that the triangle inequality, perhaps the oldest inequality known, is today full of mysteries sufficient to challenge capable mathematicians.

The conclusion from the first two lectures is: scalar equalities and inequalities continue to be valid in algebras of matrices, provided they are properly formulated to exhibit noncommutativity.

End of Lecture 2