Math 5616H, Spring 2006 **A note on several "matrix facts"**

To prove that the map $A \mapsto A^{-1}$ is continuous on the set Ω of invertible $n \times n$ matrices we began with this idea:

(1) $$A^{-1} - B^{-1} = A^{-1}(I - AB^{-1}) = A^{-1}(B - A)B^{-1}. \text{ Now let } \|A - B\| \to 0.$$

But there is a problem: we don't know that $\left\|B^{-1}\right\|$ remains bounded as $B \to A$ with respect to the matrix norm.

One way to deal with this issue is to use a series to represent B^{-1}. As usual we look for a starting-point in series of numbers. If we know that $|1 - z| < 1$ then $z \neq 0$ and also the series $\sum_0^\infty (1-z)^n$ converges absolutely. Then

$$z\sum_0^\infty (1-z)^n = -(1-z)\sum_0^\infty (1-z)^n + \sum_0^\infty (1-z)^n = -\sum_0^\infty (1-z)^{n+1} + \sum_0^\infty (1-z)^n = 1. \text{ I.e., } \sum_0^\infty (1-z)^n = \frac{1}{z}.$$

Let's work with a certain matrix in place of z as though we knew evertything worked, to show that $\left\|B^{-1}\right\|$ remains bounded as $B \to A$. Then we can check that our steps work in the matrix context.

We modify (1) to start:

(2) $$A^{-1}(A - B) = (I - A^{-1}B) \text{ and } \left\|I - A^{-1}B\right\| = \left\|A^{-1}(A-B)\right\| \leq \left\|A^{-1}\right\| \|A - B\|.$$

Let us require that $\|A - B\| < \dfrac{1}{2\left\|A^{-1}\right\|}$, so that $\left\|I - A^{-1}B\right\| < 1/2$. We'll write $C := I - A^{-1}B$; $\|C\| < 1/2$. Now we define

$$D := \sum_0^\infty C^n \text{ and notice that } (I - C)\sum_0^\infty C^n = \sum_0^\infty C^n - \sum_0^\infty C^{n+1} = \sum_0^\infty C^n - \sum_1^\infty C^n = I.$$

Thus $D = (I - C)^{-1} = (A^{-1}B)^{-1} = B^{-1}A$. This will be useful because then $B^{-1} = DA^{-1}$ and we *can* find a good bound on $\|D\|$:

$$\|D\| = \left\|\sum_0^\infty C^n\right\| \leq \sum_0^\infty \|C^n\| \leq \sum_0^\infty \|C\|^n < \sum_0^\infty (1/2)^n = 1.$$

Therefore, if $\|A - B\| < \dfrac{1}{2\left\|A^{-1}\right\|}$,

(3) $$\left\|B^{-1}\right\| = \left\|DA^{-1}\right\| \leq \|D\| \left\|A^{-1}\right\| < \left\|A^{-1}\right\|.$$

We can now use (3) in (1) to see that

(4) $$\left\|A^{-1} - B^{-1}\right\| = \left\|A^{-1}(B - A)B^{-1}\right\| \leq \left\|A^{-1}\right\| \|B - A\| \left\|B^{-1}\right\| < \left\|A^{-1}\right\|^2 \|B - A\| \to 0 \text{ as } B \to A.$$

Some points we need to cover

- What does $\sum_{k=0}^\infty C_k$ mean when the C_k are $m \times n$ matrices?
- Is it really true that $D\sum_{k=0}^\infty C_k = \sum_{k=0}^\infty DC_k$ if D is an $\ell \times m$ matrix?
- Is it really true that $\|\sum_{k=0}^\infty C_k\| \leq \sum_{k=0}^\infty \|C_k\|$?

Some useful norm estimates for $m \times n$ matrices

(5) **Theorem:** *If $A = (a_{ij})$ is an $m \times n$ matrix then*

- $\max_{i,j} |a_{ij}| \leq \|A\|$
- $\|A\| \leq \left(\sum_{i,j} |a_{ij}|^2\right)^{1/2} \leq \sqrt{mn} \cdot \max_{i,j} |a_{ij}|$.

Proof: We already knew the first estimate (i.e., inequality) in the second line. The second estimate is true because each $|a_{k\ell}|^2 \leq \max_{i,j} |a_{ij}|^2$, there are $m \cdot n$ terms in the sum and we take the square root. Note that the indices i run from 1 to m and the indices j run from 1 to n.

To prove the first estimate we apply the inequality $\|ABC\| \leq \|A\| \|B\| \|C\|$ to the (matrix-product!) equation

(for arbitrary admissible i and j) $a_{ij} = e_i^T A e_j$: $|a_{ij}| = \|(a_{ij})\|_{1\times 1} \leq \left\|e_i^T\right\|_{1\times m} \|A\|_{m\times n} \|e_j\|_{n\times 1} = \|A\|.$

This completes the proof of Theorem (5).

- **What $\sum_{k=0}^{\infty} C_k$ means when the C_k are $m \times n$ matrices**

First, a series *is* its sequence of partial sums, so we need to decide what we mean by the limit of a sequence $\{S_\ell\}$ of $m \times n$ matrices. We can regard $m \times n$ matrices as vectors in $\mathbb{R}^{m \cdot n}$, so we decide that we can simply take the limits entry-wise, requiring that $(S_\ell)_{ij} \to S_{ij}$ for all admissible pairs (i, j) of indices. But some may decide that they want the definition to be expressed in terms of matrix norms, that is, that $S_\ell \to S$ means $\|S_\ell - S\| \to 0$. By Theorem (5) either definition is equivalent to the other.

The notion of Cauchy sequences of matrices is important because it is useful.

Now we decide that $\sum_{k=0}^{\infty} C_k$ means $\lim_{\ell \to \infty} S_\ell$, where $S_\ell := \sum_{k=0}^{\ell} C_k$ denotes a partial sum of the series, *provided that* the limit exists. otherwise, $\sum_{k=0}^{\infty} C_k$ is just a symbol with no meaning as a matrix.

There is one more (still equivalent) way of looking at $\sum_{k=0}^{\infty} C_k$:

$$\sum_{k=0}^{\infty} C_k = \left(\sum_{k=0}^{\infty} (C_k)_{ij} \right),$$

that is, as a matrix of series. Then each partial sum would be a matrix of partial sums of series, one for each admissible pair (i, j).

- **It is true that $D \sum_{k=0}^{\infty} C_k = \sum_{k=0}^{\infty} DC_k$ if D is an $\ell \times m$ matrix and the series converges.**

The equation is true for partial sums and differences of them. Then, with $K < L$

$$\sum_{k=K}^{L} DC_k = D \sum_{k=K}^{L} C_k, \quad \text{so} \quad \left\| \sum_{k=K}^{L} DC_k \right\| \leq \|D\| \left\| \sum_{k=K}^{L} C_k \right\| \to 0 \ \text{ as } \ K \to \infty,$$

so that the second series converges to D times the limit of the original series.

Similarly we could multiply on the right by an $n \times p$ matrix.

- **It is true that $\|\sum_{k=0}^{\infty} C_k\| \leq \sum_{k=0}^{\infty} \|C_k\|$ if the matrix series converges**

If $\sum_{k=0}^{\infty} C_k$ does not converge, $\|\sum_{k=0}^{\infty} C_k\|$ is undefined. The triangle inequality for norms holds, so by induction the inequality holds for partial sums. By taking limits, the inequality holds if the series converges, even if the series of norms diverges.

We notice that if the series on the right converges, then the matrix series converges.