MAT 473
Spring, 2005

NOTES ON MATRIX NORMS

We will let $M_{m\times n}$ denote the set of all $m \times n$ real matrices. We can think of $M_{m\times n}$ as being a copy of $\mathbf{R}^{mn}$, so it has a Euclidean norm. We denote this norm by $\|\cdot\|_2$.

Definition. Let $A \in M_{m\times n}$. The *2-norm* of A is given by

$$\|A\|_2 = \Big(\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2\Big)^{1/2}.$$

There are many other (useful) norms that can be put on M_n. The most important norm is called the *operator norm*, which we write as $\|\cdot\|$:

Definition. Let $A \in M_{m\times n}$. The *operator norm* of A is given by

$$\|A\| = \max_{\|x\|=1} \|Ax\|.$$

(Note that since the unit sphere in $\mathbf{R}^n$ is compact, and the linear map defined by A is continuous, the maximum in the definition exists by the extreme value theorem.)

Examples. Let I_n denote the $n \times n$ identity matrix.

$$\|I_n\|_2 = \sqrt{n}$$
$$\|I_n\| = 1.$$

Lemma. $\|\cdot\|$ is a norm on $M_{m\times n}$.

Proof. (i) It is clear that $\|A\| \geq 0$.

(ii) If $\|A\| = 0$, then $Ax = 0$ for all $x \in \mathbf{R}^n$ with $\|x\| = 1$. But then $Ax = 0$ for all $x \in \mathbf{R}^n$, so that $A = 0$.

(iii) $\|cA\| = \max_{\|x\|=1} \|cAx\| = \max_{\|x\|=1} |c|\,\|Ax\| = |c| \max_{\|x\|=1} \|Ax\| = |c|\,\|A\|.$

(iv) Let $\|x\| = 1$. Then

$$\big\|(A+B)x\big\| = \big\|Ax + Bx\big\| \leq \|Ax\| + \|Bx\| \leq \|A\| + \|B\|.$$

This is true for all x with $\|x\| = 1$. Therefore

$$\|A + B\| = \max_{\|x\|=1} \big\|(A + B)x\big\| \le \|A\| + \|B\|. \qquad \blacksquare$$

Properties of the Operator Norm

1. For all $x \in \mathbf{R}^n$, $\|Ax\| \le \|A\|\,\|x\|$. Moreover, if $C > 0$ is a constant such that $\|Ax\| \le C\|x\|$ for all $x \in \mathbf{R}^n$, then $\|A\| \le C$.

Proof. First, it is clear that $\|Ax\| \le \|A\|\,\|x\|$ if $x = 0$. So suppose that $x \ne 0$. Let $y = x/\|x\|$. Then $\|y\| = 1$, so by the definition of operator norm we have $\|Ay\| \le \|A\|$. Multiplying both sides by $\|x\|$ gives $\|Ax\| \le \|A\|\,\|x\|$.

Next let C be as in the statement. Then in particular, if $\|x\| = 1$, we have $\|Ax\| \le C\|x\| = C$. Then if we consider the maximum over all x of norm 1 we see that $\|A\| \le C$. $\blacksquare$

2. $\|A\| \le \|A\|_2$.

Proof. Let $\|x\| = 1$. Realizing Ax as a linear combination of the columns of A, we have

$$\begin{aligned}
\|Ax\| &= \Big\|\sum_{j=1}^{n} x_j A_{:j}\Big\| \\
&\le \sum_{j=1}^{n} |x_j|\,\|A_{:j}\| \\
&\le \Big(\sum_{j=1}^{n} x_j^2\Big)^{1/2} \Big(\sum_{j=1}^{n} \|A_{:j}\|^2\Big)^{1/2}, \quad \text{by Cauchy-Schwarz,} \\
&= 1 \cdot \Big(\sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}^2\Big)^{1/2} \\
&= \|A\|_2.
\end{aligned}$$

It follows that $\|A\| \le \|A\|_2$. $\blacksquare$

3. If $A \in M_{m\times n}$, then $\|A\|_2 \le \sqrt{mn}\,\|A\|$.

Proof. We first estimate the entries of A:

$$|a_{ij}| = \big|(Ae_j) \cdot e_i\big| \le \|Ae_j\|\,\|e_i\| = \|Ae_j\| \le \|A\|\,\|e_j\| = \|A\|.$$

Therefore,

$$\|A\|_2 = \Big(\sum_{i,j} a_{ij}^2\Big)^{1/2} \le \Big(\sum_{i,j} \|A\|^2\Big)^{1/2} = \Big(mn\|A\|^2\Big)^{1/2} = \sqrt{mn}\|A\|. \qquad \blacksquare$$

Corollary. If $\|\cdot\|$ is used to define open balls in $M_{m\times n}$, we get the same collection of open sets as when $\|\cdot\|_2$ is used.

Corollary. A function $G : \mathbf{R}^k \to M_{m\times n}$ is continuous for $\|\cdot\|$ if and only if it is continuous for $\|\cdot\|_2$.

Corollary. Let $f : \mathbf{R}^n \to \mathbf{R}^m$ be of class C^1. Then $f' : \mathbf{R}^n \to M_{m\times n}$ is continuous for $\|\cdot\|$.

Proof.

$$\begin{aligned}\|f'(x) - f'(y)\|_2^2 &= \sum_{i=1}^{m}\sum_{j=1}^{n}\big(D_j f_i(x) - D_j f_i(y)\big)^2 \\ &\to 0, \qquad \text{as } y \to x,\end{aligned}$$

since $D_j f_i$ are continuous functions on $\mathbf{R}^n$. Therefore f' is continuous for $\|\cdot\|_2$ on $M_{m\times n}$. By the previous corollary, f' is continuous for $\|\cdot\|$. ∎

4. Let $A \in M_{m\times n}$ and $B \in M_{n\times p}$. Then $\|AB\| \le \|A\|\,\|B\|$.

Proof. Let $x \in \mathbf{R}^n$. Then $\|ABx\| \le \|A\|\,\|Bx\| \le \|A\|\,\|B\|\,\|x\|$. By the second part of 1 above, we have $\|AB\| \le \|A\|\,\|B\|$. ∎

5. Let $A \in M_n$. If $\|A - I\| < 1$, then A is invertible.

Proof. Let $x \neq 0$. Then

$$\begin{aligned}\|Ax\| &= \big\|x + (A-I)x\big\| \\ &\ge \|x\| - \big\|(A-I)x\big\| \\ &\ge \|x\| - \|A-I\|\,\|x\| \\ &> \|x\| - \|x\| \\ &= 0.\end{aligned}$$

Therefore A is one-to-one. Hence A is invertible. ∎

6. The invertible $n \times n$ matrices form an open subset of M_n (for $\|\cdot\|$, and hence also for $\|\cdot\|_2$). More precisely, if A is invertible, and if $\|A - B\| < 1/\|A^{-1}\|$, then B is also invertible. (In other words, the open ball (with respect to the operator norm) centered at A with radius $1/\|A^{-1}\|$ is entirely contained in the set of invertible matrices.)

Proof. Let $\|A - B\| < 1/\|A^{-1}\|$. Then

$$\begin{aligned} \|I - A^{-1}B\| &= \|A^{-1}A - A^{-1}B\| \\ &= \|A^{-1}(A - B)\| \\ &\leq \|A^{-1}\| \, \|A - B\|, \quad \text{by 4, above,} \\ &< 1. \end{aligned}$$

Therefore $A^{-1}B$ is invertible, by property 5. But then $B = A(A^{-1}B)$ is the product of two invertible matrices, and hence is invertible. ∎

The following companion to property 1 could have been proved at the same time. It is usually described by saying that a matrix with trivial nullspace is *bounded below*.

7. If $A \in M_{m \times n}$ has trivial nullspace, then there is $c > 0$ such that for all $x \in \mathbf{R}^n$,

$$\|Ax\| \geq c \, \|x\|.$$

Proof. If A has trivial nullspace, then $Ae_1, \ldots, Ae_n$ are linearly independent in $\mathbf{R}^m$. Let $f_1, \ldots, f_{m-n}$ be vectors in $\mathbf{R}^m$ such that $Ae_1, \ldots, Ae_n, f_1, \ldots, f_{m-n}$ form a basis for $\mathbf{R}^m$. Define $B : \mathbf{R}^m \to \mathbf{R}^n$ by $B(Ae_i) = e_i$ for $1 \leq i \leq n$, and $Bf_j = 0$ for $1 \leq j \leq m - n$. Then $BA = I_n$. For any $x \neq 0$ in $\mathbf{R}^n$ we get $\|x\| = \|BAx\| \leq \|B\| \cdot \|Ax\|$. Multiply by $c = 1/\|B\|$ to get $\|Ax\| \geq c\|x\|$. If $x = 0$ the inequality is clearly true. ∎