Section 3.9. Matrix Norm

Note. We define several matrix norms, some similar to vector norms and some reflecting how multiplication by a matrix affects the norm of a vector. We use matrix norms to discuss the convergence of sequences and series of matrices.

Definition. Consider the vector space $\mathbb{R}^{n\times m}$ of all $n\times m$ matrices with real entries. A *matrix norm* on $\mathbb{R}^{n\times m}$ is a mapping $\|\cdot\| : \mathbb{R}^{n\times m} \to \mathbb{R}$ such that for all $A, B \in \mathbb{R}^{n\times m}$ and $a \in \mathbb{R}$:

(1) Nonnegativity and Mapping of the Identity: If $A \neq 0$ then $\|A\| > 0$ and $\|0\| = 0$.

(2) Relation of Real Scalar Multiplication to Real Multiplication: $\|aA\| = |a|\|A\|$.

(3) Triangle Inequality: $\|A + B\| \leq \|A\| + \|B\|$.

(4) Consistence Property: $\|AB\| \leq \|A\|\|B\|$.

A matrix norm $\|\cdot\|$ is *orthogonally invariant* if for A and B orthogonally similar we have $\|A\| = \|B\|$. The consistence property is commonly called the "submultiplicative property."

Note. We briefly denote the norm of a vector as $\|\cdot\|_v$ and the norm of a matrix as $\|\cdot\|_M$.

Definition. The *matrix norm* on $\mathbb{R}^{n\times m}$ is $\|A\|_M = \sup_{x\neq 0} \frac{\|Ax\|_v}{\|x\|_v}$ where $A \in \mathbb{R}^{n\times m}$ and $x \in \mathbb{R}^m$. This is also called the matrix norm *induced* by the vector norm $\|\cdot\|_v$, or the *operator norm*.

Note. Gentle uses "max" in place of "sup" but this cannot be done since $\mathbb{R}^m$ contains an infinite number of nonzero vectors. In Exercise 3.22 (modified) you are to show that the matrix norm actually is a norm.

Theorem 3.9.1. The vector norm and its induced matrix norm satisfy:

(1) $\|Ax\| \leq \|A\|\|x\|$.

(2) $\|A\| = \sup_{\|x\|=1} \|Ax\|$.

Note. The proof of Theorem 3.9.1 is to be given in Exercise 3.23. In $\mathbb{R}^m$, $\{x \mid \|x\| = 1\}$ is a compact set and so it is correct to define $\|A\| = \max_{\|x\|=1} \|Ax\|$. However, in infinite dimensional vector spaces, the set $\{x \mid \|x\| = 1\}$ is not compact and in that case "sup" is necessary. (Implicit in this observation is the fact that $\|\cdot\| : \mathbb{R}^m \to \mathbb{R}$ is continuous.) For related ideas in a more general (and potentially infinite dimensional) setting, see my online notes for Fundamentals of Functional Analysis (MATH 5740) on 2.4. Bounded Linear Operators.

Definition. If we use the ℓ^p norm on $\mathbb{R}^n$ and $\mathbb{R}^m$ (see Section 2.1), then the induced matrix norm is the ℓ^p *matrix norm* $\|A\|_p = \max_{\|x\|_p=1} \|Ax\|_p$.

Theorem 3.9.2. For $n \times m$ matrix $A = [a_{ij}]$, the ℓ^1 norm satisfies $\|A\|_1 = \max_{1\leq j\leq m} \{\sum_{i=1}^n |a_{ij}|\}$ and so it is also called the *column-sum norm*. The ℓ^∞ norm satisfies $\|A\|_\infty = \max_{1\leq i\leq n} \left\{\sum_{j=1}^m |a_{ij}|\right\}$ and so it is also called the *row-sum norm*.

Note. Theorem 3.9.2 immediately implies that for any $A \in \mathbb{R}^{n\times m}$, we have $\|A^T\|_\infty = \|A\|_1$. If A is symmetric then $\|A\|_1 = \|A\|_\infty$.

Theorem 3.9.3. The ℓ^2 matrix norm and spectral radius are related as:

$$\|A\|_2 = \sqrt{\rho(A^TA)}.$$

Note. The proof of Theorem 3.9.3 is to be given in Exercise 3.24.

Note. In Exercise 3.25(a), it is shown that for orthogonal Q that $\|Qx\|_2 = \|x\|_2$; that is, the matrix transformation $Q : \mathbb{R}^m \to \mathbb{R}^n$ is an "isometry" and orthogonal matrices are examples of "isometric matrices." Notice that this gives $\|Q\|_2 = 1$. More generally, if A and B are orthogonally similar then (by the Consistency Property) $\|A\|_2 = \|B\|_2$.

Definition. For $A \in \mathbb{R}^{n\times m}$, the *Frobenius norm* is

$$\|A\|_F = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m}(a_{ij})^2}$$

(also called the *Euclidean matrix norm*).

Note. The fact that the Frobenius norm satisfies properties (1), (2), (3) of the definition of matrix norm follows from the observation that for any $A \in \mathbb{R}^{n\times m}$ there is a vector $v \in \mathbb{R}^{nm}$ (and conversely) with $\|A\|_2 = \|v\|_2$ and that $\|\cdot\|_2$ is a vector norm. Proof of the Consistency Property is to be given in Exercise 3.27.

Note. Recall that for A and B $n \times m$ matrices with the columns of A as $a_1, a_2, \ldots, a_m$ and the columns of B as $b_1, b_2, \ldots, b_m$, we have the inner product

$$\langle A, B \rangle = \sum_{j=1}^{m} a_j^T b_j = \sum_{j=1}^{m} \langle a_j, b_j \rangle.$$

So with $A = [a_{ij}]$ and $B = [b_{ij}]$, $\langle A, B \rangle = \sum_{j=1}^{m} \sum_{i=1}^{n} a_{ij} b_{ij}$ and so $\langle A, A \rangle = \sum_{i=1}^{n} \sum_{j=1}^{m} (a_{ij})^2 = \|A\|_F^2$. Also, by Theorem 3.2.8(5), $\langle A, B \rangle = \mathrm{tr}(A^T B)$, so we also have $\|A\|_F = \sqrt{\langle A, A \rangle}$ and hence

$$\|A\|_F = \sqrt{\mathrm{tr}(A^T A)} = \sqrt{\langle A, A \rangle}.$$

So the Frobenius norm is the norm induced by the matrix inner product (see page 74 of the text). Clearly from the definition of Frobenius norm we have $\|A^T\|_F = \|A\|_F$ (since the entries of A and A^T are collectively the same).

Theorem 3.9.4. If square matrices A and B are orthogonally similar then $\|A\|_F = \|B\|_F$.

Note. Since the Frobenius norm is induced by the matrix inner product and the spectral decomposition of A is $A = \sum_{i=1}^{r} d_i u_i v_i^T$ where $\langle u_i v_i^T, u_j v_j^T \rangle = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$ and $d_i = \langle A, u_i v_i^T \rangle$ then $\|A\|_F^2 = \sum_{i=1}^{r} d_i^2$ where the d_i are the singular values of A. This is Parseval's identity for the Frobenius norm (see Section 2.2 and Theorem 3.8.17).

Note 3.9.A. In Theorem 2.1.10 we showed that all vector norms on a given finite dimensional vector space are equivalent. In Exercise 3.28 it is to be shown that any two matrix norms induced by vector norms are equivalent. In fact, Gentle states that all matrix norms are equivalent (see page 133).

Theorem 3.9.5. Let A be an $n \times m$ real matrix. Then

$$\begin{aligned} \|A\|_\infty &\le \sqrt{m}\|A\|_F \\ \|A\|_F &\le \sqrt{\min\{n,m\}}\|A\|_2 \\ \|A\|_2 &\le \sqrt{m}\|A\|_1 \\ \|A\|_1 &\le \sqrt{n}\|A\|_2 \\ \|A\|_2 &\le \|A\|_F \\ \|A\|_F &\le \sqrt{n}\|A\|_\infty \end{aligned}$$

and each inequality is sharp (that is, there is a matrix A for which the inequality reduces to equality).

Note. The proof of Theorem 3.9.5 is to be given in Exercise 2.30.

Theorem 3.9.6. For any matrix norm $\|\cdot\|$ and any square matrix A, $\rho(A) \le \|A\|$.

Note. For square A, $\|A\|_1 = \max_{1 \le j \le n} \{\sum_{i=1}^n |a_{ij}|\}$ and $\|A\|_\infty = \max_{1 \le i \le n} \left\{\sum_{j=1}^m |a_{ij}|\right\}$. So by Theorem 3.9.6 we see that the spectral radius $\rho(A)$ is less than or equal to both the largest sum of absolute values of the elements in any row or column.

Note. Since we have a matrix norm, we can use it to define limits of sequences of matrices (of the same size). Since all matrix norms are equivalent by Note 3.9.A we can express the definition in terms of a "generic" matrix norm $\|\cdot\|$.

Definition. A sequence of matrices of the same size, $A_1, A_2, \ldots$ *converges* to matrix A if for all $\varepsilon > 0$ there is $N \in \mathbb{N}$ such that for all $n \geq N$ we have $\|A - A_n\| < \varepsilon$.

Theorem 3.9.7. Let A be a square matrix. Then $\lim_{k\to\infty} A^k = 0$ if and only if $\rho(A) < 1$.

Note. If fact, we can express the spectral radius in terms of a limit of the matrix norms as follows.

Theorem 3.9.8. For square matrix A, $\lim_{k\to\infty} \|A^k\|^{1/k} = \rho(A)$.

Theorem 3.9.9. Let A be an $n \times n$ matrix with $\|A\| < 1$. Then

$$I + \lim_{k\to\infty} \left(\sum_{n=1}^{k} A^n \right) = (I - A)^{-1}.$$

Definition. A square matrix A such that $A^k = 0$ and $A^{k-1} \neq 0$ for some $k \in \mathbb{N}$ is *nilpotent of index k*.

Theorem 3.9.10. Suppose A is an $n \times n$ matrix which is nilpotent of order $k \in \mathbb{N}$. Then:

(1) $\rho(A) = 0$ (so all eigenvalues of A are 0), and

(2) $\text{tr}(A) = 0$, and

(3) $\text{rank}(A) \leq n - 1$.

Note. The proof of Theorem 3.9.10 is to be given in Exercise 3.33.

Revised: 7/2/2020