8 Lecture 8: Matrix Norms

8.1 Subordinate matrix norms

Definition 8.1. Given any norm $\|\cdot\|$ on $\mathbb{R}^n$, there is a *subordinate matrix norm* on $\mathbb{R}^{n\times n}$ defined by

$$\|A\| = \max_{v \in \mathbb{R}^n_\star} \frac{\|Av\|}{\|v\|}, \tag{8.2}$$

where $A \in \mathbb{R}^{n\times n}$ and $\mathbb{R}^n_\star = \mathbb{R}^n/\{\mathbf{0}\}$.

You might wonder why we define a matrix norm like this. The reason is that we like to think of A as an *operator* on $\mathbb{R}^n$: if $v \in \mathbb{R}^n$ then $Av \in \mathbb{R}^n$. So rather than the norm giving us information about the "size" of the entries of a matrix, it tells us how much the matrix can change the size of a vector.

The formula for a subordinate matrix norm in Definition 8.1, is sensible, but not much use to if we actually want to compute, say, $\|A\|_1$, $\|A\|_\infty$ or $\|A\|_2$. For a given A, we'd have to calculate $\|Av\|/\|v\|$ for *all* v. And there is rather a lot of them. Fortunately, there are some easier ways of computing the more important norms. We'll see that

- The ∞-norm of a matrix is just largest absolute-value row sum.
- The 1-norm of a matrix is just largest absolute-value column sum.
- The 2-norm of the matrix A is the square root of the largest eigenvalue of A^TA.

8.1.1 The max-norm and 1-norm on $\mathbb{R}^{n\times n}$

Theorem 8.2. *For any $A \in \mathbb{R}^{n\times n}$ the subordinate matrix norm associated with $\|\cdot\|_\infty$ on $\mathbb{R}^n$ can be computed by*

$$\|A\|_\infty = \max_{i=1,\dots n} \sum_{j=1}^{n} |a_{ij}|.$$

A similar result holds for the 1-norm, the proof of which is left as an exercise:

Theorem 8.3.

$$\|A\|_1 = \max_{j=1,\dots,n} \sum_{i=1}^{n} |a_{i,j}|. \tag{8.3}$$

Computing the 2-norm of a matrix is a little harder that computing the 1- or ∞-norms. However, later we'll need estimates not just for $\|A\|$, but also $\|A^{-1}\|$. And, unlike the 1- and ∞-norms, we can estimate $\|A^{-1}\|_2$ without explicitly forming A^{-1}.

8.2 Eigenvalues

We begin by recalling some important facts about eigenvalues and eigenvectors.

Definition 8.4. Let $A \in \mathbb{R}^{n\times n}$. We call $\lambda \in \mathbb{C}$ an *eigenvalue* of A if there is a non-zero vector $x \in \mathbb{C}^n$ such that

$$Ax = \lambda x.$$

We call any such x an *eigenvector associated with* A.

Some properties of eigenvalues:

(i) If A is a real symmetric matrix (i.e., $A = A^T$), its eigenvalues and eigenvectors are all real-valued.

(ii) If λ is an eigenvalue of A, the $1/\lambda$ is an eigenvalue of A^{-1}.

(iii) If x is an eigenvector associated with the eigenvalue λ then so too is ηx for any non-zero scalar η.

(iv) An eigenvector may be *normalised* as
$\|x\|_2^2 = x^Tx = 1$.

(v) There are n eigenvectors λ_1, λ_n, ..., λ_n associated with the real symmetric matrix A. Let $\boldsymbol{x}^{(1)}$, $\boldsymbol{x}^{(2)}$, ..., $\boldsymbol{x}^{(n)}$ be the associated normalised eigenvectors. Then the eigenvectors are linearly independent and so form a basis for $\mathbb{R}^n$. That is, any vector $\nu \in \mathbb{R}^n$ can be written as a linear combination:

$$\boldsymbol{v} = \sum_{i=1}^{n} \alpha_i \boldsymbol{x}^{(i)}.$$

(vi) Furthermore, these eigenvectors are *orthogonal* and *orthonormal*:

$$\left(x^{(i)}\right)^T x^{(j)} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

8.3 Singular values

The *singular values* of a matrix A are the square roots of the eigenvalues of $A^T A$. They play a very important role in matrix analysis and in areas of applied linear algebra, such as image and text processing. Our interest here is in their relationship to $\|A\|_2$.

Lemma 8.5. *For any matrix A, the eigenvalues of $A^T A$ are real and non-negative.*

9 Lecture 9: More matrix norms

9.1 Singular values, again

At the end of Lecture 7 we proved that if we take any real, square matrix, A, and set $B = A^T A$, then the eigenvalues of B are non-negative real numbers. A consequence of this is that the square root of the eigenvalues of $A^T A$ are also non-negative and real. Furthermore, part of the above proof involved showing that, if $\left(A^T A\right)\boldsymbol{x} = \lambda x$, then

$$\sqrt{\lambda} = \frac{\|A\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2}.$$

This at the very least tells us that

$$\|A\|_2 := \max_{\boldsymbol{x} \in \mathbb{R}^n_\star} \frac{\|A\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2} \geqslant \max_{i=1,\dots,n} \sqrt{\lambda_i}.$$

With a bit more work, we can show that if $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$ are the the eigenvalues of $B = A^T A$, then

$$\|A\|_2 = \sqrt{\lambda_n}.$$

Theorem 9.1. *Let $A \in \mathbb{R}^{n \times n}$. Let $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$, be the eigenvalues of $B = A^T A$. Then*

$$\|A\|_2 = \max_{i=1,\dots,n} \sqrt{\lambda_i} = \sqrt{\lambda_n},$$

9.2 Consistency of matrix norms

It should be clear from (8.2) that, if $\|\cdot\|$ is a subordinate matrix norm then, for *any* $\boldsymbol{u} \in \mathbb{R}^n$, $A \in \mathbb{R}^{n \times n}$,

$$\|A\boldsymbol{u}\| \leqslant \|A\|\|\boldsymbol{u}\|.$$

It is an important result: we'll need it tomorrow. There is an analogous statement for the product of two matrices:

Definition 9.2. A matrix norm $\|\cdot\|$ is *consistent* if

$$\|AB\| \leqslant \|A\|\|B\|, \qquad \text{for all } A, B \in \mathbb{R}^{n \times n}.$$

Theorem 9.3. *Any subordinate matrix norm is consistent.*