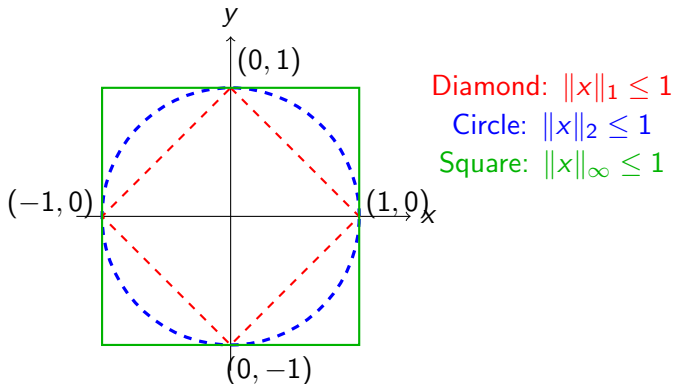


MA385 Part 4: Linear Algebra 2

4.1: Vector and Matrix Norms

Dr Niall Madden

November 2025



1. Outline Section 4.1

- | | | | |
|---|--|---|---|
| 1 | Introduction | 4 | Computing Matrix Norms |
| 2 | Three vector norms | 5 | The max-norm on $\mathbb{R}^{n \times n}$ |
| | ■ $\ \cdot\ _\infty$ is a norm on \mathbb{R}^n | 6 | Computing $\ A\ _2$ |
| | ■ $\ \cdot\ _2$ is a norm on \mathbb{R}^n | | ■ Eigenvalues |
| 3 | Matrix Norms | 7 | Exercises |

For more, see Section 2.7 of Suli and Mayers:

<https://ebookcentral.proquest.com/lib/nuig/reader.action?docID=221072&ppg=51&c=UERG>

2. Introduction

This is the final section of MA385. It is kinda a direct continuation of Section 3 – and much of the material is from the same chapter as Section 3 in the text-book (though we'll also take some material from Chapter 5).

At its heart, is the task of bounding the eigenvalues and singular values of a matrix. Our motivation comes from doing an error analysis for LU -factorization. However, the applications are far more general than that.

2. Introduction

But for now, we'll just note that all computer implementations of algorithms that involve floating-point numbers (roughly, finite decimal approximations of real numbers) contain errors due to round-off error.

It transpires that computer implementations of LU -factorization, and related methods, lead to these round-off errors being greatly magnified: and we want to understand why.

2. Introduction

You might remember from earlier sections of the course that we had to assume functions were well-behaved in the sense that

$$\frac{|f(x) - f(y)|}{|x - y|} \leq L,$$

for some number L , so that our numerical schemes (e.g., fixed point iteration, Euler's method, etc) would work. If a function *doesn't* satisfy a condition like this, we say it is “ill-conditioned”. One of the consequences is that a small error in the inputs gives a large error in the outputs.

We'd like to be able to express similar ideas about matrices: that $A(u - v) = Au - Av$ is not too “large” compared to $u - v$. To do this we used the notion of a “norm” to describing the relative sizes of the vectors u and Au .

3. Three vector norms

When we want to consider the size of a real number, without regard to sign, we use the *absolute value*. Important properties of this function are:

1. $|x| \geq 0$ for all x .
2. $|x| = 0$ if and only if $x = 0$.
3. $|\lambda x| = |\lambda||x|$.
4. $|x + y| \leq |x| + |y|$ (triangle inequality).

This notion can be extended to vectors and matrices.

3. Three vector norms

Definition 4.1.1

Let \mathbb{R}^n be all the vectors of length n of real numbers. The function $\|\cdot\|$ is called a **norm** of \mathbb{R}^n if, for all $u, v \in \mathbb{R}^n$

1. $\|v\| \geq 0$,
2. $\|v\| = 0$ if and only if $v = 0$.
3. $\|\lambda v\| = |\lambda| \|v\|$ for any $\lambda \in \mathbb{R}$,
4. $\|u + v\| \leq \|u\| + \|v\|$ (triangle inequality).

Norms on vectors in \mathbb{R}^n quantify the *size* of the vector. But there are different ways of doing this...

3. Three vector norms

Definition 4.1.2

Let $\mathbf{v} \in \mathbb{R}^n$: $\mathbf{v} = (v_1, v_2, \dots, v_{n-1}, v_n)^T$.

(i) The 1-norm (a.k.a. the *Taxi cab* norm) is

$$\|\mathbf{v}\|_1 = \sum_{i=1}^n |v_i|.$$

(ii) The 2-norm (a.k.a. the *Euclidean* norm)

$$\|\mathbf{v}\|_2 = \left(\sum_{i=1}^n v_i^2 \right)^{1/2}.$$

Note, if \mathbf{v} is a vector in \mathbb{R}^n , then

$$\mathbf{v}^T \mathbf{v} = v_1^2 + v_2^2 + \dots + v_n^2 = \|\mathbf{v}\|_2^2.$$

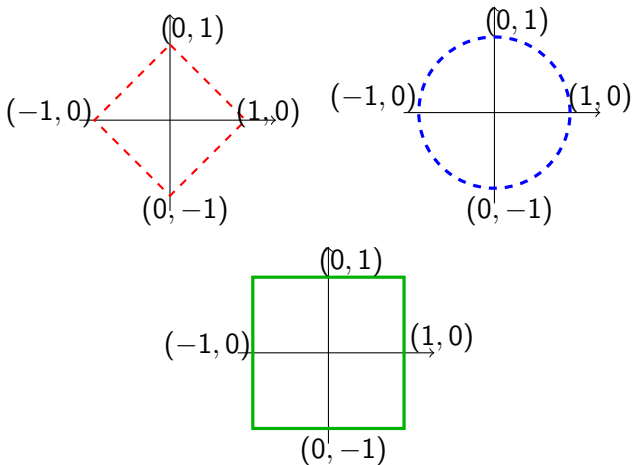
(iii) The ∞ -norm (a.k.a. the *max-norm*) $\|\mathbf{v}\|_\infty = \max_{i=1}^n |v_i|$.

3. Three vector norms

Example: Compute the 1-, 2- and ∞ -norm of $v = (-2, 4, -4)$

3. Three vector norms

The unit balls in \mathbb{R}^2 given by $\|\cdot\|_1$ (top left),
 $\|x\|_2 = \sqrt{x_1^2 + x_2^2} = 1$ (top right), and $\|\cdot\|_\infty$.



3. Three vector norms

It is easy to show that $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are norms (see next slide).

And it is not hard to show that $\|\cdot\|_2$ satisfies conditions (1), (2) and (3) of Definition 4.1.1.

It takes a little bit of effort to show that $\|\cdot\|_2$ satisfies the triangle inequality; so we'll do that with care.

3. Three vector norms

$\|\cdot\|_\infty$ is a norm on \mathbb{R}^n

3. Three vector norms

$\|\cdot\|_2$ is a norm on \mathbb{R}^n

As mentioned, it takes a little effort to show that $\|\cdot\|_2$ is indeed a norm on \mathbb{R}^2 ; in particular to show that it satisfies the triangle inequality, we need the Cauchy-Schwarz inequality.

Lemma (Cauchy-Schwarz)

$$\left| \sum_{i=1}^n u_i v_i \right| \leq \|u\|_2 \|v\|_2, \quad \forall u, v \in \mathbb{R}^n.$$

Idea: $0 \leq \|\lambda u + v\|_2^2$.

3. Three vector norms

$\|\cdot\|_2$ is a norm on \mathbb{R}^n

Now can now apply Cauchy-Schwartz to show that

$$\|u + v\|_2 \leq \|u\|_2 + \|v\|_2.$$

This is because

$$\begin{aligned}\|u + v\|_2^2 &= (u + v)^T(u + v) \\ &= u^T u + 2u^T v + v^T v \\ &\leq u^T u + 2|u^T v| + v^T v \quad (\text{by the triangle-inequality}) \\ &\leq u^T u + 2\|u\|\|v\| + v^T v \quad (\text{by Cauchy-Schwarz}) \\ &= (\|u\| + \|v\|)^2.\end{aligned}$$

It follows directly that

Corollary

$\|\cdot\|_2$ is a norm.

4. Matrix Norms

Definition 4.1.3

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

$$\|A\| = \max_{v \in \mathbb{R}_*^n} \frac{\|Av\|}{\|v\|}, \quad (1)$$

where $A \in \mathbb{R}^{n \times n}$ and $\mathbb{R}_*^n = \mathbb{R}^n / \{0\}$.

We define a matrix norm like this because we think of A as an *operator* on \mathbb{R}^n : if $v \in \mathbb{R}^n$ then $Av \in \mathbb{R}^n$. So the norm of A gives us information on how much the matrix can change the size of a vector.

5. Computing Matrix Norms

It is not obvious from the above definition how to calculate the norm of a given matrix. We'll see that

- ▶ The ∞ -norm of a matrix is also the largest absolute-value row sum.
- ▶ The 1-norm of a matrix is also the largest absolute-value column sum.
- ▶ The 2-norm of the matrix A is the square root of the largest eigenvalue of $A^T A$.

6. The max-norm on $\mathbb{R}^{n \times n}$

Theorem 4.1.1

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

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

6. The max-norm on $\mathbb{R}^{n \times n}$

6. The max-norm on $\mathbb{R}^{n \times n}$

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

Theorem 4.1.2

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

$$\|A\|_1 = \max_{j=1, \dots, n} \sum_{i=1}^n |a_{i,j}|. \quad (2)$$

Computing the 2-norm of a matrix is a little harder than 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} .

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

Definition 4.1.4

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* of A associated with λ .

- (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 \mathbf{x} is an eigenvector associated with the eigenvalue λ then so too is $\eta\mathbf{x}$ for any non-zero scalar η .
- (iv) An eigenvector may be *normalised* as $\|\mathbf{x}\|_2^2 = \mathbf{x}^T \mathbf{x} = 1$.

- (v) There are n eigenvectors $\lambda_1, \lambda_2, \dots, \lambda_n$ associated with the real symmetric matrix A . Let $x^{(1)}, x^{(2)}, \dots, x^{(n)}$ be the associated normalised eigenvectors. The eigenvectors are linearly independent and so form a basis for \mathbb{R}^n . That is, any vector $v \in \mathbb{R}^n$ can be written as a linear combination:

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

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

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

Here is a useful consequence of (v) and (vi), which we will use repeatedly.

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, applied linear algebra, and statistics (principal component analysis).

Our interest here is in their relationship to $\|A\|_2$.

But first we'll prove a theorem about certain matrices (so called, “normal matrices”).

Theorem 4.1.3

For any matrix $A \in \mathbb{R}^{n \times n}$, the eigenvalues of $A^T A$ are real and non-negative.

Part of the above proof involved showing that, if $(A^T A)x = \lambda x$, then

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

This at the very least tells us that

$$\|A\|_2 := \max_{x \in \mathbb{R}_*^n} \frac{\|Ax\|_2}{\|x\|_2} \geq \max_{i=1,\dots,n} \sqrt{\lambda_i}.$$

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

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

Theorem 4.1.4

Let $A \in \mathbb{R}^{n \times n}$. Let $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$, be the eigenvalues of $B = A^T A$. Then

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

Here is the main idea. For full details, see the text-book.

8. Exercises

Exercise 4.1.1 (★)

Show that, for any vector $x \in \mathbb{R}^n$, $\|x\|_\infty \leq \|x\|_2$ and $\|x\|_2^2 \leq \|x\|_1 \|x\|_\infty$. For each of these inequalities, give an example for which the equality holds. Deduce that $\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1$.

Exercise 4.1.2

Show that if $x \in \mathbb{R}^n$, then $\|x\|_1 \leq n\|x\|_\infty$ and that $\|x\|_2 \leq \sqrt{n}\|x\|_\infty$.

Exercise 4.1.3

Show that, for any subordinate matrix norm on $\mathbb{R}^{n \times n}$, the norm of the identity matrix is 1.

8. Exercises

Exercise 4.1.4 (★)

Prove that

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

Hint: Suppose that

$$\sum_{i=1}^n |a_{ij}| \leq C, \quad j = 1, 2, \dots, n,$$

show that for *any* vector $x \in \mathbb{R}^n$

$$\sum_{i=1}^n |(Ax)_i| \leq C \|x\|_1.$$

Now find a vector x such that $\sum_{i=1}^n |(Ax)_i| = C \|x\|_1$. Now deduce the result.