WEEK 1: REVIEW OF LINEAR ALGEBRA, CALCULUS AND ANALYSIS. PART I.

1. Properties of the Complex Exponential

This material in this section comes mainly from the first section in [3].

The exponential function, denoted $\exp(x)$ or e^x, is widely considered the single most important function in mathematics. Not only are great many other functions—trigonometric functions, power functions, logarithm, Gaussian, and gamma function, for starters—defined in terms of it, but the exponential itself plays a central role in harmonic analysis, the study of differential equations, and many other topics. The material in this course involves Fourier analysis, therefore, familiarity with the exponential's basic properties is a must.

Definition 1. We define the function $\exp : \mathbb{C} \to \mathbb{C}$ by

$$\exp z = \sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \frac{z^3}{3!} + \frac{z^4}{4!} + ...$$

As the series converges absolutely for every $z \in \mathbb{C}$, there is no ambiguity in the definition. The shorthand $\exp z = e^z$ is common. However, math students should keep in mind that exp id not *defined* to be e^z. Indeed, if b is not an integer, then a^b is defined for real and complex a as

$$a^b = \exp(b \ln a),$$

so that without Definition 1, e^z would have no meaning.

We will assume without proof the basic algebraic properties of the exponential function, in particular $\exp(z + w) = (\exp z)(\exp w)$ for all $z, w \in \mathbb{C}$.

Proposition 2. *The function* exp *is complex differentiable at every $z \in \mathbb{C}$, and $(\exp z)' = \exp z$. This means that the limit*

$$\lim_{w \to z} \frac{f(w) - f(z)}{w - z}$$

exists for every $z \in \mathbb{C}$ and is equal to $\exp z$.

Proof. Although the argument can be made rigorous (see any book on complex analysis) we will proceed with a formal argument, assuming among other things that the complex derivative of a polynomial takes the same form as its real derivative. Since the series converges absolutely for all z, we differentiate

$$\exp z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

term by term to get that

$$\frac{d}{dz}(\exp z) = \sum_{n=1}^{\infty} \frac{z^{n-1}}{(n-1)!},$$

which, we see after reindexing by $n' = n - 1$, is the original series. □

Corollary 3. *The exponential function is infinitely differentiable, and all of its derivatives are equal to itself.*

Definition 4. A function with a complex derivative at z is called *analytic* or *holomorphic* at z, and a function that is analytic on the complex plane is called *entire*.

Complex differentiability turns out to be much stronger requirement than real differentiability. For example, if a function has a complex derivative on an open set then that derivative is always continuous, and in fact, is automatically differentiable. In other words, a function with a single complex derivative automatically has infinitely many of them. The same result is not even remotely true for real derivatives. The subject of complex analysis consists primarily of the study of analytic functions.

Now, we focus on the restriction of the exponential function to the imaginary axis; that is, we look at the function from $\mathbb{R}$ to $\mathbb{C}$ defined by $\exp ix,\ x \in \mathbb{R}$. We will refer to this function as the complex exponential, although the reader should be aware that this terminology is not standardized.

Proposition 5. (Euler's formula) *Let* $x \in \mathbb{R}$. *Then* $\exp(ix) = \cos x + i \sin x$.

Proof. Exercise (see Assignment 1). □

Corollary 6. *For any real* x, $|\exp(ix)| = 1$.

Proof. Using Euler's formula and the definition of $|\cdot|$ on the complex plane, $|z|^2 = (\operatorname{Re} z)^2 + (\operatorname{Im} z)^2$, this is simply the Pythagorean identity. □

Once can see from Euler's formula that as x increases, $\exp(ix)$ simply precesses counterclockwise around the unit circle, coming back around to 1 whenever $x = 2\pi n$ for $n \in \mathbb{Z}$.

After having had exposure only to the rapidly increasing real exponential, the boundedness of the complex exponential can take some getting used to, but it is important to remember it, since otherwise the Fourier Transform could not exist.

Remark 7. It is common to see polar notation used for complex numbers, where the real coordinates r and θ are defined as they would be if the point $z \in \mathbb{C}$ were viewed as a point in $\mathbb{R}^2$. By Euler's formula, this polar representation is equivalent to letting $z = re^{i\theta}$. This, combined with the algebraic identities regarding the exponential function, can be useful in proving algebraic properties of the complex numbers that would be difficult to prove using the rectangular coordinates. For example:

Proposition 8. *For every* $z \in \mathbb{C}, n \in \mathbb{N}$ *with* $k \leq n$, *there exist* $w_1, w_2, ..., w_{n-1} \in \mathbb{C}$ *such that* $w^n = z$ *and* $i \neq j \Rightarrow w_i \neq w_j$. *In other words, each complex number has a full set of (distinct)* n^{th} *roots.*

Proof. We assume basic trigonometric identities. Let $z \in \mathbb{C}$. We can express z as $z = re^{i\theta_0}$ for some $\theta_0 \in \mathbb{R}$. Note that if we assume $\theta_0 \in [0, 2\pi)$, then $z = re^{i(\theta_0 + 2\pi k)}$ for any $k \in \mathbb{Z}$. Let $\theta_k = \theta_0 + 2\pi k$. For each $0 \leq k \leq n-1$, let $z_k = r^{1/n}e^{i(\theta_k/n)}$. It is easy to check that all the z_k's are distinct and that for each k, $(z_k)^n = z$. □

This last proposition is important for theory of inner product spaces, which we will use often in this course. We leave its (easy) proof to the reader.

Proposition 9. *For* $z \in \mathbb{C}$, $z\bar{z} = |z|^2$.

2. Vector Spaces, Bases, and Dimension

This material in this section is adapted primarily from [2].

Many—in fact, most—of the operations which we will study in our survey of medical imaging are linear; that is, their domains are vector spaces, and they commute with addition and scalar multiplication. In order to make use of this useful property, we need a solid understanding of the abstract idea of a vector space. This section of notes introduces abstract vector spaces, and in order to give the reader a taste of abstract linear algebra we include some proofs.

In an elementary linear algebra course, and before that in precalculus and calculus III, you were probably introduced to the idea of vectors in the context of ordered n-tuples of real numbers, and to linear operators in terms of matrices. In our exploration of medical imaging, we will encounter many other vector spaces, and most of them will be vector spaces over the complex numbers rather than the real numbers. In order to give the general definition of a vector space, we first must define a field.

Definition 10. A *field* is a set $\mathbb{F}$, equipped with two commutative and associative binary operations $+$ and $\cdot$ such that

(a) $\mathbb{F}$ is closed under $*$ and $+$, that is, if $a, b \in \mathbb{F}$, then $a \cdot b \in \mathbb{F}$ and $a + b \in \mathbb{F}$.
(b) Both operations are commutative and associative.
(c) Both $+$ and $*$ have identities, denoted 0 and 1 respectively,
(d) Every element $x \in \mathbb{F}$ has an additive inverse, written $-x$, and every nonzero element has a multiplicative inverse, denoted x^{-1} or $1/x$.
(e) Multiplication distributes over addition; that is, for all $a, b, c \in \mathbb{F}$, $a(b + c) = ab + ac$.

In practice, the $\cdot$ is often omitted, and $a \cdot b$ is written ab.

Example 11. By far, the most common fields in analysis are the rational numbers, the real numbers, and the complex numbers. For prime numbers p, $\mathbb{Z}_p$ is also a field under modular addition and multiplication. The simplest example is $\mathbb{Z}_2$, the set $\{0, 1\}$ endowed with the operations $0 + 0 = 1 + 1 = 0 * 0 = 0 * 1 = 0$, $1 * 1 = 0 + 1 = 1$. It is easy to check that this is a field.

We can now define vector spaces over a field:

Definition 12. A *vector space* over a field $\mathbb{F}$ is a set V endowed with a vector addition $+$ on $V \times V$ and a scalar multiplication $\cdot$ on $\mathbb{F} \times V$ such that

i. V is closed under vector addition and scalar multiplication.
ii. Addition is commutative and associative.
iii. There is an additive identity, a 0 vector, and every $v \in V$ has an additive inverse, denoted $-v$
iv. Multiplication by the scalar identity is associative, that is for $a, b \in \mathbb{F}$ and $v \in V$, $a(bv) = (ab)v$, where the multiplication at right is multiplication within the field $\mathbb{F}$, rather than scalar multiplication in V.
v. Scalar multiplication respects the multiplicative identity of $\mathbb{F}$, that is, $1 \cdot v = v$ for all $v \in V$.
vi. Scalar multiplication distributes over addition in both $\mathbb{F}$ and V, that is, for $a, b \in \mathbb{F}$ and $u, v \in V$, $a * v + b * v = (a + b) * v$ and $au + av = a(u + v)$.

As with fields, the $*$ is often omitted in practice.

A set with one element can always be made into a vector space, called 'the trivial vector space', by treating that one element as the zero vector. A field $\mathbb{F}$ is always a vector space over itself, as is the set of ordered n-tuples of numbers from $\mathbb{F}$, endowed with pointwise addition and scalar multiplication (as in the case of $\mathbb{R}^n$ and $\mathbb{C}^n$). In fact, given an arbitrary set S as the domain, the set of all functions from S to F is a vector space under the operations of pointwise addition and pointwise scalar multiplication. There are many other function spaces besides the rather unrestrictive space of *all* functions $S \to \mathbb{F}$. You may recall from advanced calculus that sums and scalar multiples of continuous/Riemann integrable real-valued functions functions on an interval are continuous/Riemann integrable, thus the real continuous/Riemann integrable functions on an interval constitute real vector spaces. We will encounter many more function spaces in this course.

Proposition 13. *If V is a vector space over $\mathbb{F}$, then for every $v \in V$, $0 \cdot v = 0$, where the zero on the left is in $\mathbb{F}$ and the zero on the right is in V.*

Proof. By the distributive property, we have $v = 1v = (0 + 1)v = 0v + 1v = 0v + v$, then $0v$ is the additive identity in V, and hence $0v = 0$. □

Definition 14. Let W be a subset of a vector space V over $\mathbb{F}$. W is called a *subspace* of V if W is a vector space under the operations induced by V, that is, W is closed under addition and scalar multiplication.

Example 15. The set containing the zero vector is always a subspace of V, as is the set of scalar multiples of any single vector in V.

Definition 16. Let V be a vector space over $\mathbb{F}$, and let $v_1, ..., v_N \in V$. A vector $u \in V$ is a *linear combination* of the v_j's if u has the form

$$u = \sum_{j=1}^{N} a_j v_j,$$

where $a_j \in \mathbb{F}$ for $j = 1, ..., N$.

By convention, when we write a vector as a linear combination of other vectors, we assume without explicitly saying that $v_j \neq v_k$ whenever $j \neq k$, in order to avoid cluttering out theorems and proofs with words like 'distinct' and 'unique.'

Definition 17. Let V be a vector space over a field F, and let S be a set of vectors in V. Then S is said to be *linearly independent* if for any finite collection $v_1, ..., v_N \in S$ and $a_1, ..., a_N \in \mathbb{F}$,

$$\sum_{j=1}^{N} a_j v_j = 0 \iff a_j = 0 \quad \text{for all} \quad j = 1, ..., N.$$

Note that no set containing the zero vector can ever be linearly independent.

Lemma 18. *Let $v_1, ..., v_N$ be a set of linearly independent vectors in a vector space V over $\mathbb{F}$. Then*

$$\sum_{j=1}^{N} a_j v_j = \sum_{j=1}^{N} b_j v_j,$$

if and only if $a_j = b_j$ for all j.

Proof. By the distributive property, the equality above is equivalent to

$$\sum_{j=1}^{N} (a_j - b_j) v_j = 0,$$

and by linear independence and Proposition 13, the above equality holds if and only if $a_j - b_j = 0$ for all j. □

Corollary 19. *Let S be a (possibly infinite) set of linearly independent vectors in a vector space V over $\mathbb{F}$. Let $v = \sum_{j=1}^{N} a_j v_j$ and $u = \sum_{j=1}^{M} b_j u_j$ be linear combinations of elements in S. Then $u = v$ if and only if we have that $v_j = u_k$ only if $a_j = b_k$, and also that whenever $a_j \neq 0$, then $v_j = w_k$ for some k.*

Proof. If the v_j's and the w_k's are not the same, add the missing terms to each linear combination with coefficients of zero and reorder terms in order to make them the same, so that we may assume without loss of generality that $M = N$ and that $w_j = v_j$ for all j. The result follows by Lemma 18. □

Definition 20. Let S be a set of vectors from a vector space V over $\mathbb{F}$. The *span* of S is the set of vectors

$$\{v \in V : v \text{ can be written as a linear combination of vectors in } S\}.$$

By convention, the span of the empty set is $\{0\}$.

Note that the span of S is always a subspace of V.

Definition 21. Let V be a vector space over $\mathbb{F}$. A set B of vectors in V is called a *basis* for V if B is linearly independent and spans W.

We remark that the definition of a basis given above is more restrictive than that used by some authors. In a normed vector space (which we will define later, if you have not previously seen the definition), some authors would allow an infinite linearly independent set to be a basis of V if its span were merely dense in V. When working with infinite-dimensional vector spaces, in our work on Fourier Series, instead we use the related idea of complete orthonormal sets in infinite- dimensional inner product spaces. For now, we will confine our attention to spaces with finite bases.

The most common example of a basis is the standard Euclidean basis in $\mathbb{R}^n$ or $\mathbb{C}^n$, is the set of vectors which are zero in all but one component, and are 1 in that component.

Lemma 22. *Let V be a vector space over $\mathbb{F}$, let W be a subspace of V (possibly V itself), and let $B = \{v_1, ..., v_N\} \subset W$ be a finite set. Then B is a basis for W if and only if for every $v \in W$,*

$$v = \sum_{j=1}^{N} a_j v_j,$$

where the a_j's are determined uniquely by v.

Proof. By definition, B spans W if and only if each $v \in W$ can be written as a linear combination of the v_j's, and by Lemma 18, B is linearly independent if and only if the coefficients of all linear combinations of vectors in B are unique. □

Theorem 23. *Let W be a subspace of a vector space V over $\mathbb{F}$. Let B_1 and B_2 be bases for W, and suppose that B_1 contains n elements. Then B_2 also contains n elements.*

Proof. Exercise (see Assignment 1). □

Definition 24. Let V be a vector space, and suppose that V has a finite basis B. The *dimension* of V is defined to be the number of elements in B. If there is no finite basis for V, then we say that V has infinite dimension.

Note that we needed Theorem 23 to guarantee that one could not find a vector space with two bases of unequal size. Otherwise, the idea of dimension would not be well-defined.

Theorem 25. *Let V be an n-dimensional vector space, and $v_1, ..., v_n$ be distinct vectors in V. Then $v_1, ..., v_n$ are linearly independent if and only if they span V.*

Proof. Exercise (see Assignment 1). □

Definition 26. Let V be a vector space over $\mathbb{F}$, and let $B = \{v_1, ..., v_n\}$ be a basis for V. For any $v \in V$, there are unique coefficients $a_j \in \mathbb{F},\ j = 1, ..., n$ such that $v = \sum_{j=1}^n a_j v_j$. We define $[v]_B \in \mathbb{F}^n$ by

$$[v]_B = (a_1, a_2, ..., a_n),$$

and we call a_j the jth *component of v with respect to B.*

Note that the vector $[v]_B$ is not the same as the vector v; they are completely different mathematical objects if $V \neq \mathbb{F}^n$, and even if $V = \mathbb{F}^n$ it need not be true that $[v]_B = v$. On the other hand, if B is the standard basis for $\mathbb{F}^n$, then it will be true that $[v]_B = v$.

3. Metrics, Norms, and Inner Products

Metric spaces and normed spaces are among the most useful mathematical objects in analysis. Inner products allow for even more precise analysis on certain normed spaces by generalizing the idea of perpendicularity. You have almost certainly been exposed to all these ideas before, although possibly only to real inner products. But we need to give the abstract definitions, because in our study of medical imaging, we will encounter many normed and inner product spaces besides the familiar example of n-tuples of real or complex numbers, most notably series and functions. This material comes from a combination of [3], [2], [1], and [4].

3.1. Metric Spaces.

Definition 27. Let X be a set. A function $d : X \times X \to \mathbb{R}$ is called a *metric* on X and (X, d) is called a *metric space* if d satisfies the following properties:

(i) For any $x, y \in X,\ d(x, y) \geq 0$, and $d(x, y) = 0$ iff $x = y$. (positive definiteness)
(ii) For any $x, y \in X, d(x, y) = d(y, x)$. (symmetry)
(iii) For any $x, y, z \in X, d(x, z) \leq d(x, y) + d(y, z)$. (triangle inequality)

The metric with which you are probably most familiar is the distance function $|\cdot|$ on $\mathbb{R}$, defined by $d(x, y) = |x - y|$. This is the metric one uses when studying convergence and limits in Advanced Calculus. You are probably also familiar from precalculus with the Euclidean distance function on $\mathbb{R}^n$, defined by

$$d(\vec{x}, \vec{y}) = \left(\sum_{i=1}^n (x_i - y_i)^2 \right)^{1/2}.$$

(Here, x_i and y_i refer to the components of the vectors $\vec{x}$ and $\vec{y}$.) This is the most common metric on $\mathbb{R}^n$, but not the only one, as we shall see. In fact, the discrete metric defined in the next example defines another metric on $\mathbb{R}^n$, albeit one of very limited usefulness.

Really, the idea of a metric is a generalization of the geometric notion of distance in $\mathbb{R}$, $\mathbb{R}^2$, and $\mathbb{R}^3$. The axioms that metrics satisfy ensure that they meet our most basic expectations of a distance function.

Example 28. The simplest metric, which can be put on *any* set S, is called the discrete metric $d : S \times S \to \mathbb{R}$, defined for any $x, y \in S$ by

$$d(x, y) = \begin{cases} 0 & \text{if} \quad x = y \\ 1 & \text{if} \quad x \neq y \end{cases}$$

It is easy to verify that the discrete metric satisfies the properties of a metric on S :

- *Symmetry:* Clearly if $d(x,y) = 0$, then $x = y$, which means $d(y,x) = 0$. Similarly, if $d(x,y) = 1$, then $x \neq y$, thus, $d(y,x) = 1$.
- *Positive definiteness:* Positive definiteness comes from constructions, since $d(x,y)$ takes the value of 0 whenever x and y are equal, and the value $1 \geq 0$ otherwise.
- *Triangle Inequality:* Suppose $x = z$. Then $d(x,z) = 0$, therefore $d(x,z) \leq d(x,y) + d(y,z)$ follows immediately since the norm cannot take negative values. If $x \neq z$, then either $x \neq y$ or $y \neq z$, since $x = y$ and $y = z$ would lead to $x = z$. Thus, $d(x,z) = 1 \leq d(x,y) + d(y,z)$.

Proposition 29. (Reverse Triangle Inequality) *Let (X,d) be a metric space. Then any $x,y \in X$,*

$$|d(x,z) - d(z,y)| \leq d(x,y).$$

Proof. From the Triangle Inequality, we have

$$d(x,z) \leq d(x,y) + d(y,z) \Rightarrow d(x,z) - d(y,z) \leq d(x,y),$$

and similarly

$$d(y,z) \leq d(y,x) + d(x,z) \Rightarrow$$
$$d(y,z) - d(x,z) = -(d(x,z) - d(y,z)) \leq d(y,x) = d(x,y).$$

Taken together, these two inequalities give us the desired result. □

We conclude our look at metric spaces with a result which may seem obvious, but has great importance, because in applications we often work only with a subset of a metric space.

Proposition 30. *Let (X,d) be a metric space, and let $Y \subset X$. Let $d|_Y$ denote the restriction of d to $Y \times Y$. Show that $(Y, d|_Y)$ is a metric space.*

Proof. Let $x,y,z \in Y$. Since $Y \subset X$, then $x,y,z \in X$, thus the metric d satisfies all required properties for x,y,z. Since $d|_Y = d$ on $Y \times Y$, d_Y also satisfies all the properties of a metric for x,y,z. Since $x,y,z \in Y$ were arbitrary, the statement follows. □

3.2. Normed Spaces.

Many of the most useful metric spaces are also vector spaces, and their metrics behave well with respect to vector space operations of addition and scalar multiplication. These are metrics which come from norms. Before we show that normed spaces are vector spaces, we must define a norm.

Definition 31. Let V be a vector space over a field (generally $\mathbb{R}$ or $\mathbb{C}$). A function $\|\cdot\| : V \to \mathbb{R}$ is called a *norm* on V and $(V, \|\cdot\|)$ is referred to as a *normed vector space* if $\|\cdot\|$ satisfies the following properties:

(i) For any $x \in V$, $\|x\| \geq 0$ and $\|x\| = 0$ only if $x = 0$. (positive definiteness)
(ii) For $\alpha \in F$, $\|\alpha x\| = |\alpha| \, \|x\|$. (positive homogeneity)
(iii) For any $x,y \in V$, $\|x + y\| \leq \|x\| + \|y\|$. (triangle inequality)

Like metrics, norms satisfy a reverse triangle inequality:

Proposition 32. *Let $(V, \|\cdot\|)$ be a normed space. Then for any $x,y \in V$,*

$$|\|x\| - \|y\|| \leq \|x - y\|.$$

Proof. Using the triangle inequality, we have that

$$\|x\| \le \|x-y\| + \|y\| \Rightarrow \|x\| - \|y\| \le \|x-y\|,$$

and

$$\|y\| \le \|y-x\| + \|x\| \Rightarrow$$
$$\|y\| - \|x\| \le |-1|\,\|x-y\| = \|x-y\|.$$

Together, these two inequalities give us our desired result. □

Like metrics, norms generalize the idea of distance, specifically distance from the zero vector. But norms tell us more than metrics; in order to have a norm, a space must at least be a linear space. As we saw with the discrete metric, a metric space need not meet these requirements. An important quality of norms is that they extend to define a metric on the normed space.

Proposition 33. *Let* $(V, \|\cdot\|)$ *be a normed space. Define* $d : V \times V \to \mathbb{R}$ *by* $d(x,y) = \|x-y\|$. *Then* (V,d) *is a metric space.*

Proof. Exercise (see Assignment 1). □

Example 34. Several norms should already be familiar. The most basic is the absolute value function on $\mathbb{R}$. The Euclidean norm on $\mathbb{R}^n$ is another common norm, defined by

$$\|\vec{x}\| = \left(\sum_{j=1}^{n} (x_j)^2 \right)^{1/2}.$$

The absolute value function on $\mathbb{C}$ is another important norm, in which we treat $z = x + iy$ as a vector (x,y) in $\mathbb{R}^2$ and use the Euclidean norm on (x,y) to define $|z|$.

Example 35. Define the *max norm* or *uniform norm* $\|\cdot\|_u : \mathbb{R}^n \to \mathbb{R}$ by $\|\cdot\|_u(\vec{x}) = \max_{1\le j\le n} |x_j|$. Then $\|\cdot\|_u$ is a norm on $\mathbb{R}^n$. [1]

Proof.

- *Positive Definiteness:* Let $\vec{x} \in \mathbb{R}^n$. It is clear that $\max_j |x_j| \ge 0$, since every term $|x_j|$ is. Moreover, if $\vec{x} = 0$, then by definition each $x_j = 0$, hence, $\sum_{j=1}^{n} |x_j| = 0$, and this holds only if $\vec{x} = 0$.
- *Positive Homogeneity:* Let $\vec{x} \in \mathbb{R}^n$, and let $a \in \mathbb{R}$. Then for each $1 \le j \le n$, $|ax_j| = |a|\,|x_j|$. Suppose that the largest component of $\vec{x}$ was $|x_k|$. It is clear that $|x_k|$ is still the largest after scaling by $|a|$, and thus, $\|a\vec{x}\|_u = |a|\,\|\vec{x}\|_u$.
- *Triangle Inequality:* Let $\vec{x}, \vec{y} \in \mathbb{R}^n$, and let $\vec{z} = \vec{x} + \vec{y}$. Suppose the largest component of $\vec{z}$ is $|z_k|$. We know $z_k = x_k + y_k$, and by the triangle inequality on $\mathbb{R}$, $\|\vec{z}\|_u = |z_k| \le |x_k| + |y_k|$. By definition of the norm, $|x_k| \le \|\vec{x}\|_u$ and $|y_k| \le \|\vec{y}\|_u$, thus $\|\vec{z}\|_u \le \|\vec{x}\|_u + \|\vec{y}\|_u$.

□

The same proof obviously shows that the max norm is a norm on $\mathbb{C}^n$.

[1] This uniform norm can be defined similarly on sequences and bounded functions, as we will see later

Example 36. A common collection of norms on $\mathbb{C}^n$ is that of the p-norms. For $p \geq 1$, the norm $\|\cdot\|_p : \mathbb{R}^n \to \mathbb{R}$, is defined by

$$\|\vec{x}\|_p = \left(\sum_{j=1}^{n} |x_j|^p \right)^{1/p}.$$

Note that the ordinary Euclidean distance function is just the special case $p = 2$. Except in the easy case of $p = 1$, it is nontrivial to show that the triangle inequality holds for the p-norms, although since the 2-norm is induced by an inner product, we will be able to show after the next section that the 2-norm does satisfy the triangle inequality.

Remark 37. The max norm and p-norms described in this section on $\mathbb{C}^n$ have important analogues on function spaces given by integration rather than summation (technically, they are actually special cases of the more general norms). The rigorous study of these spaces, L^p theory, requires measure theory and can be found in graduate real analysis courses, although we will look at some important facts about L^p spaces later in this course, as they are important for harmonic analysis.

3.3. Inner Product Spaces.

Many of the most useful normed spaces have, in addition to a convenient way of looking at distance, a sense of direction; that is, given two vectors we can determine, in some sense, the extent to which they point the same way. The tool that allows us to do this is the inner product.

Definition 38. Let H be a real or complex vector space[2] A function $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{C}$ is called an *inner product* on H and $(H, \langle \cdot, \cdot \rangle)$ is referred to as a (complex) [3] *inner product space* if $\langle \cdot, \cdot \rangle$ has the following properties:

(i) For any $x, y \in H$, $\langle x, y \rangle = \overline{\langle y, x \rangle}$. (conjugate symmetry)
(ii) For any $x, y, z \in H$ and $a, b \in \mathbb{C}$, $\langle ax + by, z \rangle = a \langle x, z \rangle + b \langle y, z \rangle$. (linearity in first parameter)
(iii) If $x \neq 0$, then $\langle x, x \rangle > 0$. (nonnegativity)

Example 39. We define an inner product on $\mathbb{C}^n$ by

$$\langle \vec{x}, \vec{y} \rangle = \vec{x} \cdot \vec{y} = \sum_{j=1}^{n} x_i \bar{y}_i.$$

We check that this Euclidean inner product is, in fact, an inner product:

Proof.

- *Symmetry:* Let $\vec{x}, \vec{y} \in \mathbb{C}^n$. Then, since multiplication of complex numbers is commutative, and the product of the conjugates of two complex numbers is the conjugate of their product, $\vec{x} \cdot \vec{y}$ is defined by the same sum as $\overline{\vec{y} \cdot \vec{x}}$, so the two must be equal.
- *Linearity in the first parameter:* Let $\vec{x}, \vec{y}, \vec{z} \in \mathbb{C}^n$, and $a, b \in \mathbb{C}$. When we look at the sum defining $(a\vec{x} + b\vec{y}) \cdot \vec{z}$, we see that for each $1 \leq j \leq n$, $(ax_j + by_j)\bar{z}_j = a(x_j \bar{z}_j) + b(y_j \bar{z}_j)$. Since each term in the finite sum is linear, so is the sum itself, hence, $(a\vec{x} + b\vec{y}) \cdot \vec{z} = a\vec{x} \cdot \vec{z} + b\vec{y} \cdot \vec{z}$.
- *Nonnegativity* Suppose $\vec{x} \neq 0$. Then at least one component $x_k \neq 0$, thus, $x_k \bar{x}_k = |x_k|^2 > 0$. Moreover, for each $1 \leq j \leq n$, we know that $x_j \bar{x}_j = |x_j|^2$ is nonnegative. Taken together, these two facts imply that the sum defining $\vec{x} \cdot \vec{x}$ is positive.

□

[2]By 'complex vector space' and 'real vector space,' we mean a vector space over the field $\mathbb{C}$ or the field $\mathbb{R}$, respectively.

[3]if the vector space is real, then we can similarly define a real inner product by only allowing scalar multiples in $\mathbb{R}$ and regarding conjugate symmetry as regular symmetry

Proposition 40. *Let H be a complex inner product space.*

i. *If $x = 0$, then $\langle x, x\rangle = 0$.*
ii. *Linearity and complex symmetry imply complex linearity in the second parameter, that is,*

$$\langle x, ay + bz\rangle = \bar{a}\,\langle x, y\rangle + \bar{b}\,\langle x, z\rangle\,.$$

Proof. i. Suppose $x = 0$. Choose any vector $y \in V$. Then $x = 0y$, so by linearity

$$\langle x, x\rangle = 0\,\langle y, x\rangle = 0.$$

ii.

$$\begin{aligned}\langle x, ay + bz\rangle &= \overline{\langle ay + bz, x\rangle}\\ &= \overline{a\,\langle y, x\rangle + b\,\langle z, x\rangle}\\ &= \bar{a}\overline{\langle y, x\rangle} + \bar{b}\overline{\langle z, x\rangle}\\ &= \bar{a}\,\langle x, y\rangle + \bar{b}\,\langle x, z\rangle\,.\end{aligned}$$

□

Just as the norm on a normed vector space defines a metric on the space, so does an inner product give us a norm.

Definition 41. Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space, and define $\|\cdot\| : V \to \mathbb{R}$ by $\|x\| = \sqrt{\langle x, x\rangle}$.

In order to verify that $\|\cdot\|$ actually satisfies the properties of a norm, we need the important and useful Cauchy-Schwartz inequality:

Theorem 42. (Cauchy-Schwartz Inequality) *Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space and let $\|\cdot\| : V \to \mathbb{R}$ be defined as above. Then for any $x, y \in V$,*

$$|\langle x, y\rangle| \le \|x\|\,\|y\|\,.$$

Proof. Exercise (see Assignment 1). □

Proposition 43. *For an inner product space $(V, \langle\cdot,\cdot\rangle)$, if we define $\|\cdot\|$ as above, then $\|\cdot\|$ is a norm on V.*

Proof. Exercise (see Assignment 1). □

Remark 44. It follows that the Euclidean norm is, in fact, a norm on $\mathbb{R}^n$.

References

[1] Charles Epstein. *Introduction to the Mathematics of Medical Imaging.* Prentice Hall, Upper Saddle River, 2nd edition, 2007.
[2] Michael W. Frazier. *An Introduction to Wavelets Through Linear Algebra.* Springer, New York, 1999.
[3] Walter Rudin. *Real and Complex Analysis.* McGraw-Hill, New York, 3rd edition, 1987.
[4] Walter Rudin. *Real Analysis: Modern Techniques and Their Applications.* John Wiley and Sons, New York, 2nd edition, 1999.