Chapter 2

Linear Spaces

Contents

In systems analysis, **linear spaces** are ubiquitous.

Why? Linear systems/models are easier to analyze; many systems, particularly in signal processing, are deliberately designed to be linear; linear models are a useful starting point (approximation) for more complicated nonlinear cases.

Formal definitions of a vector space use the concept of a **field** of scalars, so we first review that.

Field of Scalars (from Applied Linear Algebra, Noble and Daniel, 2nd ed.)

A **field of scalars** $\mathcal{F}$ is a collection of elements $\alpha, \beta, \gamma, \ldots$ along with an "addition" and a "multiplication" operator.

To every pair of scalars α, β in $\mathcal{F}$, there must correspond a scalar $\alpha + \beta$ in $\mathcal{F}$, called the **sum** of α and β, such that

- Addition is commutative: $\alpha + \beta = \beta + \alpha$
- Addition is associative: $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$
- There exists a unique element $0 \in \mathcal{F}$, called **zero**, for which $\alpha + 0 = \alpha$, $\forall \alpha \in \mathcal{F}$
- For every $\alpha \in \mathcal{F}$, there corresponds a unique scalar $(-\alpha) \in \mathcal{F}$ for which $\alpha + (-\alpha) = 0$.

To every pair of scalars α, β in $\mathcal{F}$, there must correspond a scalar $\alpha\beta$ in $\mathcal{F}$, called the **product** of α and β, such that

- Multiplication is commutative: $\alpha\beta = \beta\alpha$
- Multiplication is associative: $\alpha(\beta\gamma) = (\alpha\beta)\gamma$
- Multiplication distributes over addition: $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$
- There exists a unique element $1 \in \mathcal{F}$, called **one**, or **unity**, or the **identity** element, for which $1\alpha = \alpha$, $\forall \alpha \in \mathcal{F}$
- For every nonzero $\alpha \in \mathcal{F}$, there corresponds a unique scalar $\alpha^{-1} \in \mathcal{F}$, called the **inverse** of α for which $\alpha\alpha^{-1} = 1$.

Simple facts for fields:

- $0 + 0 = 0$ (use $\alpha = 0$ in the definition of 0)
- $-0 = 0$ *Proof.* For any α, by the associative property $(\alpha + 0) + (-0) = \alpha + (0 + (-0))$ hence $\alpha + (-0) = \alpha$. Hence, since the zero element is unique, $-0 = 0$.

Example. The set of rational numbers $\mathbb{Q}$ (with the usual definition of addition and multiplication) is a field.

The only fields that we will need are the field of real numbers $\mathbb{R}$ and the field of complex numbers $\mathbb{C}$.
Therefore, hereafter we will use $\mathcal{F}$ when describing results that hold for either $\mathbb{R}$ or $\mathbb{C}$.

Vector Spaces

In simple words, a vector space is a space that is closed under vector addition and under scalar multiplication.

Definition. A **vector space** or **linear space** consists of the following four entities.

1. A field $\mathcal{F}$ of scalars.
2. A set $\mathcal{X}$ of elements called **vectors**.
3. An operation called **vector addition** that associates a **sum** $\boldsymbol{x} + \boldsymbol{y} \in \mathcal{X}$ with each pair of vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ such that
   - Addition is commutative: $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$
   - Addition is associative: $\boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z}) = (\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z}$
   - There exists an element $\mathbf{0} \in \mathcal{X}$, called the **zero vector**, for which $\boldsymbol{x} + \mathbf{0} = \boldsymbol{x}$, $\forall \boldsymbol{x} \in \mathcal{X}$
   - For every $\boldsymbol{x} \in \mathcal{X}$, there corresponds a unique vector $(-\boldsymbol{x}) \in \mathcal{X}$ for which $\boldsymbol{x} + (-\boldsymbol{x}) = \mathbf{0}$.
4. An operation called **multiplication by a scalar** that associates with each scalar $\alpha \in \mathcal{F}$ and vector $\boldsymbol{x} \in \mathcal{X}$ a vector $\alpha\boldsymbol{x} \in \mathcal{X}$, called the **product** of α and $\boldsymbol{x}$, such that:
   - Associative: $\alpha(\beta\boldsymbol{x}) = (\alpha\beta)\boldsymbol{x}$
   - Distributive $\alpha(\boldsymbol{x} + \boldsymbol{y}) = \alpha\boldsymbol{x} + \alpha\boldsymbol{y}$
   - Distributive $(\alpha + \beta)\boldsymbol{x} = \alpha\boldsymbol{x} + \beta\boldsymbol{x}$
   - If 1 is the identify element of $\mathcal{F}$, then $1\boldsymbol{x} = \boldsymbol{x}$. $\forall \boldsymbol{x} \in \mathcal{X}$.
   - $0\boldsymbol{x} = \mathbf{0}$ for any $\boldsymbol{x} \in \mathcal{X}$.

The requirement that $\boldsymbol{x} + \boldsymbol{y} \in \mathcal{X}$ and $\alpha\boldsymbol{x} \in \mathcal{X}$ is sometimes called the **closure property**.

Simple facts for vector spaces:

- $\mathbf{0}$ is unique.
- $(-1)\boldsymbol{x} = -\boldsymbol{x}$ for $\boldsymbol{x} \in \mathcal{X}$.
  Proof. $\boldsymbol{x} + (-1)\boldsymbol{x} = 1\boldsymbol{x} + (-1)\boldsymbol{x} = (1 + (-1))\boldsymbol{x} = 0\boldsymbol{x} = \mathbf{0}$.
- $\alpha\mathbf{0} = \mathbf{0}$ for $\alpha \in \mathcal{F}$.
  Proof. $\alpha\mathbf{0} = \alpha\mathbf{0} + \mathbf{0} = \alpha\mathbf{0} + [\alpha\mathbf{0} + (-\alpha\mathbf{0})] = [\alpha\mathbf{0} + \alpha\mathbf{0}] + (-\alpha\mathbf{0}) = \alpha(\mathbf{0} + \mathbf{0}) + (-\alpha\mathbf{0}) = \alpha\mathbf{0} + (-\alpha\mathbf{0}) = \mathbf{0}$.
- $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{x} + \boldsymbol{z}$ implies $\boldsymbol{y} = \boldsymbol{z}$ (cancellation law)
- $\alpha\boldsymbol{x} = \alpha\boldsymbol{y}$ and $\alpha \neq 0$ implies $\boldsymbol{x} = \boldsymbol{y}$ (cancellation law)
- $\alpha\boldsymbol{x} = \beta\boldsymbol{x}$ and $\boldsymbol{x} \neq \mathbf{0}$ implies $\alpha = \beta$ (cancellation law)
- $\alpha(\boldsymbol{x} - \boldsymbol{y}) = \alpha\boldsymbol{x} - \alpha\boldsymbol{y}$ (distributive law)
- $(\alpha - \beta)\boldsymbol{x} = \alpha\boldsymbol{x} - \beta\boldsymbol{x}$ (distributive law)
- $-\alpha\boldsymbol{x} = \alpha(-\boldsymbol{x}) = -(\alpha\boldsymbol{x})$

What are some examples? (Linear algebra classes focus on finite-dimensional examples.)

Important Vector Spaces

- **Euclidean space** or n-**tuple space**: $\mathcal{X} = \mathbb{R}^n$. If $\boldsymbol{x} \in \mathcal{X}$, then $\boldsymbol{x} = (a_1, a_2, \ldots, a_n)$ where $a_i \in \mathbb{R}$ and we use ordinary addition and multiplication: $\boldsymbol{x} + \boldsymbol{y} = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)$ and $\alpha \boldsymbol{x} = (\alpha a_1, \ldots, \alpha a_n)$.
  (As a special case, the set of real numbers $\mathbb{R}$ (with ordinary addition and multiplication) is a trivial vector space.)
- $\mathcal{X} = \mathcal{L}_1\left[\mathbb{R}^2\right]$. The set of functions $f : \mathbb{R}^2 \to \mathbb{R}$ that are **absolutely (Lebesgue) integrable**: $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |f(x,y)| \, dx \, dy < \infty$, with the usual pointwise definition of addition and scalar multiplication.
  To show show that $f, g \in \mathcal{L}_1\left[\mathbb{R}^2\right]$ implies $f + g \in \mathcal{L}_1\left[\mathbb{R}^2\right]$, one can apply the triangle inequality: $|f + g| \leq |f| + |g|$.
- The set of functions on the plane $\mathbb{R}^2$ that are zero outside of the unit square.
- The set of solutions to a homogeneous linear system of equations $A\boldsymbol{x} = \boldsymbol{0}$.
- $C[a, b]$: the space of real-valued, continuous functions defined on the interval $[a, b]$.
- The space of **band-limited** signals.
- Many more in Luenberger...

Example. For $1 \leq p < \infty$, define the following infinite-dimensional[1] space:

$$\mathcal{X} = R_p[a,b] = \left\{ f : [a,b] \to \mathbb{R} : f \text{ is Riemann integrable and } \int_0^1 |f(t)|^p \, dt < \infty \right\},$$

(with the usual pointwise definitions of addition of functions and multiplication of functions by a scalar). To show that this space is a vector space, the only nontrivial work is verifying **closure**.

Clearly if $f \in \mathcal{X}$ then $\alpha f \in \mathcal{X}$ since $\int_0^1 |\alpha f(t)|^p \, dt = |\alpha|^p \int_0^1 |f(t)|^p \, dt < \infty$, so $\mathcal{X}$ is closed under scalar multiplication.

To show that $f + g \in \mathcal{X}$ if $f, g \in \mathcal{X}$, *i.e.*, closure under addition, requires a bit more work. Note that since $|a+b| \leq 2 \max\{|a|, |b|\}$, it follows for $p \geq 1$ that

$$|a + b|^p \leq 2^p \max\{|a|^p, |b|^p\} \leq 2^p[|a|^p + |b|^p].$$

Hence if $f, g \in \mathcal{X}$:

$$\int_0^1 |f(t) + g(t)|^p \, dt \leq \int_0^1 [2^p |f(t)|^p + 2^p |f(t)|^p] \, dt \leq 2^p \int_0^1 |f(t)|^p \, dt + 2^p \int_0^1 |f(t)|^p \, dt < \infty + \infty = \infty.$$

showing closure under addition.

Example. Can $\mathcal{X} = (0, \infty)$ with $\mathcal{F} = \mathbb{R}$ be a vector space? **??**

Cartesian product

We can make a "larger" vector space from two vector spaces $\mathcal{X}$ and $\mathcal{Y}$ (having a common field $\mathcal{F}$) by forming the **Cartesian product** of $\mathcal{X}$ and $\mathcal{Y}$, denoted $\mathcal{X} \times \mathcal{Y}$, which is the collection of ordered pairs $(\boldsymbol{x}, \boldsymbol{y})$ where $\boldsymbol{x} \in \mathcal{X}$ and $\boldsymbol{y} \in \mathcal{Y}$.

$$\mathcal{X} \times \mathcal{Y} = \{(\boldsymbol{x}, \boldsymbol{y}) : \boldsymbol{x} \in \mathcal{X} \text{ and } \boldsymbol{y} \in \mathcal{Y}\}.$$

To be a vector space we must define vector addition and scalar multiplication operations, which we define component-wise:

- $(\boldsymbol{x}_1, \boldsymbol{y}_1) + (\boldsymbol{x}_2, \boldsymbol{y}_2) = (\boldsymbol{x}_1 + \boldsymbol{x}_2, \boldsymbol{y}_1 + \boldsymbol{y}_2)$
- $\alpha(\boldsymbol{x}, \boldsymbol{y}) = (\alpha \boldsymbol{x}, \alpha \boldsymbol{y}), \ \forall \alpha \in \mathcal{F}$.

Fact. With these definitions the Cartesian product of two vector spaces is indeed a vector space.
The above definition generalizes easily to higher-order combinations.

Example. $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$.

[1] A definition of dimension is forthcoming...

2.3

Subspaces

A (nonempty) subset S of a vector space $\mathcal{X}$ is called a **subspace** of $\mathcal{X}$ if S, when endowed with the addition and scalar multiplication operations defined for $\mathcal{X}$, is a vector space, *i.e.*, $\alpha \boldsymbol{x} + \beta \boldsymbol{y} \in S$ whenever $\boldsymbol{x}, \boldsymbol{y} \in S$ and $\alpha, \beta \in \mathcal{F}$.

Example. The subset of $\mathcal{X} = R_2[-1, 1]$ consisting of symmetric functions ($f(-t) = f(t)$) is a subspace of $\mathcal{X}$. It is clearly closed under addition and scalar multiplication.

What are the four types of subspaces of $\mathbb{R}^3$? **??**

Intuition: think of a subspace like a line or plane (or hyperplane) through the origin.

Properties of subspaces
- $\mathbf{0} \in S$
- $\{\mathbf{0}\}$ is a subspace of $\mathcal{X}$
- $\mathcal{X}$ is a subspace of $\mathcal{X}$
- A subspace not equal to the entire space $\mathcal{X}$ is called a **proper subspace**
- If M and N are subspaces of a vector space $\mathcal{X}$, then the intersection $M \cap N$ is also a subspace of $\mathcal{X}$. *Proof.* ***see text***
  Think: intersection of planes (through the origin) in 3d.
- Typically the **union** of two subspaces is *not* a subspace.
  Think: union of planes (through the origin) in 3d.

Although unions usually fail, we can combine two subspaces by an appropriate sum, defined next.

Sum of subsets

Definition. If S and T are two *subsets* of a vector space, then the **sum** of those subsets, denoted $S + T$ is defined by

$$S + T = \{\boldsymbol{x} = \boldsymbol{s} + \boldsymbol{t} : \boldsymbol{s} \in S, \ \boldsymbol{t} \in T\}.$$

Example. What is the sum of a plane and a line (both through origin) in $\mathbb{R}^3$? **??**

Example. Consider $\mathcal{X} = \mathbb{R}^2$, with $S = \{(x, 0) : x \in [0, 1]\}$ and $T = \{(0, y) : y \in [0, 1]\}$. Then $S + T$ is the unit square.

Proposition. If M and N are subspaces of a vector space $\mathcal{X}$, then the sum $M + N$ is a subspace of $\mathcal{X}$.

Proof. ***see text***

Does the previous example illustrate this proposition? **??**

Example. Let $\mathcal{X} = \{f : f(t) = a \sin(t + \phi) \text{ for } a, \phi \in \mathbb{R}\}$ (with the usual definitions of addition and scalar multiplication[2]). Then $M = \{f : f(t) = a \sin(t) \text{ for } a \in \mathbb{R}\}$ and $N = \{f : f(t) = a \cos(t) \text{ for } a \in \mathbb{R}\}$ are both (proper) subspaces of $\mathcal{X}$.

What is $M + N$? **??**

[2] It is time to stop saying this. From now on we leave it implicit whenever this is clear, which it usually is.

Linear combinations

Definition. A *finite* sum $\sum_{i=1}^n \alpha_i \boldsymbol{x}_i$ for $\boldsymbol{x}_i \in \mathcal{X}$ and $\alpha_i \in \mathcal{F}$ is called a **linear combination**.

(The associative property of vector addition allows us to write such a sum without parentheses.)

Depending on where the $\boldsymbol{x}_i$'s originated we get various properties of linear combinations.

- $\mathcal{X}$: If $\boldsymbol{x}_i \in \mathcal{X},\ i = 1, \ldots, n$, then $\sum_{i=1}^n \alpha_i \boldsymbol{x}_i \in \mathcal{X}$. This is shown easily by induction from the definition of a vector space.
- M: If $\boldsymbol{x}_i \in M,\ i = 1, \ldots, n$, where M is a subspace, then $\sum_{i=1}^n \alpha_i \boldsymbol{x}_i \in M$ by induction from the definition of a subspace. Any linear combination of vectors from a subspace is also in the subspace.
- S: What if we take linear combinations from a *subset* rather than a *subspace*?

Definition. If S is a *subset* of a vector space $\mathcal{X}$, then the **subspace generated** by S is the subspace of linear combinations drawn from S, defined by

$$[S] = \left\{ \boldsymbol{x} \in \mathcal{X} : \boldsymbol{x} = \sum_{i=1}^n \alpha_i \boldsymbol{x}_i,\ \text{for } \boldsymbol{x}_i \in S,\ \alpha_i \in \mathcal{F},\ \text{ and } n \in \mathbb{N} \right\}.$$

- $[S]$ is also called the **span** or the **linear hull** of S.
- $[S]$ is indeed a subspace of $\mathcal{X}$ since a linear combinations of linear combinations is itself a linear combination.
- $[S]$ is the smallest subspace of $\mathcal{X}$ containing S, *i.e.*, if M is a subspace of $\mathcal{X}$ that contains S, then $[S] \subseteq M$.
- If M is a subspace of $\mathcal{X}$, then $[M] =$ **??**
- Clearly $S \subseteq [S]$
- Note that only *finite* sums are involved, as in all linear combinations.

Example. For $\mathcal{X} = \mathbb{R}^3$, what is $[S]$ when S consists of a line through the origin plus any point not on that line?
??

Intuition: a subspace of a general vector space generalizes the notion of a line or plane through the origin of Euclidean 3D space.

What about lines or planes that are *not* through the origin?

Linear varieties

skip for now. Not needed in Ch. 2 problems. Wait until *3.10*.

Definition. A subset V of a vector space $\mathcal{X}$ is called a **linear variety** iff $V = \boldsymbol{x}_0 + M$ for some $\boldsymbol{x}_0 \in \mathcal{X}$ and some subspace M of $\mathcal{X}$. Another term used is **affine subspace**.

Linear varieties arise in certain minimum norm problems.

2.5

Linear independence

Often we need to quantify how "big" a subspace is.

Definition. A vector $\boldsymbol{x}$ is called **linearly dependent** on a set S of vectors iff $\boldsymbol{x} \in [S]$, *i.e.*, $\boldsymbol{x}$ is in the **span** of S.

Otherwise, if $\boldsymbol{x} \notin [S]$, then $\boldsymbol{x}$ is called **linearly independent** of S.

Definition. A set S of vectors is called a **linearly independent set** if each vector in the set is linearly independent of the remaining vectors in the set, *i.e.*,

$$\forall \boldsymbol{x} \in S, \quad \boldsymbol{x} \notin [S - \{\boldsymbol{x}\}] .$$

Remark. S may be uncountable, but testing whether $\boldsymbol{x} \in [S - \{\boldsymbol{x}\}]$ requires consideration only of finite sums, by the definition of linear combinations.

Example. This illustrates that S can be uncountable!

Let $\mathcal{X} = \{f : [0,1] \to \mathbb{R}\}$ and define $g_s(t) = \begin{cases} 1, & t = s \\ 0, & \text{otherwise.} \end{cases}$

Then $S = \{g_s : s \in [0,1]\}$ is a linearly independent subset of $\mathcal{X}$.

> **Theorem.** *A finite set of vectors $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ is linearly independent iff $\sum_{i=1}^n \alpha_i \boldsymbol{x}_i = \mathbf{0}$ implies that $\alpha_i = 0, \ \forall i$.*

(We are skipping proofs that are found in basic linear algebra books.)

> **Corollary**. If a finite set $S = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ is linearly independent and $\boldsymbol{y} \in [S]$, then $\boldsymbol{y}$ has a *unique* expansion $\boldsymbol{y} = \sum_{i=1}^n \alpha_i \boldsymbol{x}_i$ for some $\alpha_1, \ldots, \alpha_n \in \mathcal{F}$.

Basis and Dimension

Definition. A set S is called a **basis** or **Hamel basis** [3, p 183] for $\mathcal{X}$ iff S is linearly independent and $[S] = \mathcal{X}$.

Luenberger [4, p. 20] says "finite set" but Naylor [3, p. 183] and Maddox [2, p. 74] do not. Let us agree to use the above definition, rather than Luenberger's.

Note that $\mathcal{X}$ can (and usually will) have more than one basis!

Definition. If $\mathcal{X}$ has a basis S that is a finite set, then we call $\mathcal{X}$ **finite dimensional**.
Otherwise, if no such finite S exists, we call $\mathcal{X}$ **infinite dimensional**.

Definition. A space with a basis consisting of n elements is called an n**-dimensional space**.
This terminology is acceptable thanks to the following result.

> **Theorem.** *Any two bases for a finite-dimensional vector space contain the same number of elements.*

Most, but not all, properties of (more easily understood) finite-dimensional spaces generalize to infinite-dimensional spaces.

Example. $P_n = \{\text{polynomials of degree} \leq n\}$

A basis is $\{1, t, t^2, \ldots, t^n\}$, which has dimension n. Why linearly independent? **??**
Another basis is the **Legendre polynomials**: $d^k/dt^k (t^2 - 1)^k, \ k = 1, \ldots, n$.

Exercise. $C[0,1]$ is infinite dimensional. Hint: prove by counter example considering P_{n+1}.

A basis is a generalization of the usual concept of a **coordinate system**.

Fact. [3, p 184] If S is linearly independent set in a vector space $\mathcal{X}$, then there exists a basis B for $\mathcal{X}$ such that $S \subseteq B$.
Thus, every vector space has a Hamel basis, (since the empty set is linearly independent).

However, "Hamel basis is not the only concept of basis that arises in analysis. There are concepts of basis that involve topological as well as linear structure. ... In applications involving infinite-dimensional spaces, a useful basis, if one even exists, is usually something other than a Hamel basis. For example, a complete orthonormal set is far more useful in an infinite-dimensional Hilbert space than a Hamel basis” [3, p. 183]. (More on this later!)

We often need sets with less rigid structure than subspaces but that nevertheless still have some structure, so we digress a bit here.

2.4

Convexity

Luenberger describes convexity as “the fundamental algebraic concept of vector space.” (p. 25)

According to Oxford English dictionary, an **algebra** is “a calculus of symbols combining according to certain defined laws.”

Definition. A set $\mathcal{K}$ in a vector space is called **convex** iff for any $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{K}$, $\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y} \in \mathcal{K}$ for all $\alpha \in [0,1]$.

Geometrically: for any two points in a convex set, the “line segment” between them is also in the set.

Properties of convex sets
- Subspaces and linear varieties are convex.
- $\{\mathbf{0}\}$ is “vacuously” convex.
- For $\alpha \in \mathcal{F}$, $\alpha\mathcal{K} \triangleq \{\boldsymbol{x} = \alpha\boldsymbol{k} : \boldsymbol{k} \in \mathcal{K}\}$ is convex if $\mathcal{K}$ is convex (**magnification** or **minification** of a set)
- $\mathcal{K}_1 + \mathcal{K}_2$ is convex if $\mathcal{K}_1$ and $\mathcal{K}_2$ are convex sets in a common vector space.
- If $\mathcal{C}$ is a collection of convex sets (in a common vector space), then $\bigcap_{\mathcal{K}\in\mathcal{C}} \mathcal{K}$ is a convex set. (Important for **POCS** methods.)

Convex hull

Definition. The **convex cover** or **convex hull** of a set S in a vector space is the smallest convex set containing S, denoted $\operatorname{co}(S)$. Equivalently, the convex hull of S is the intersection of all convex sets containing S:

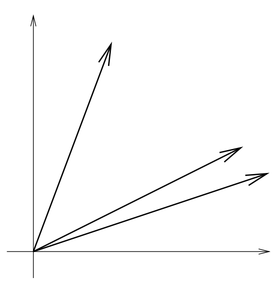
$$\operatorname{co}(S) = \bigcap_{\{cK \text{ convex} \,:\, S \subseteq \mathcal{K}\}} \mathcal{K}.$$

(Picture in 2D of blob and its convex hull)

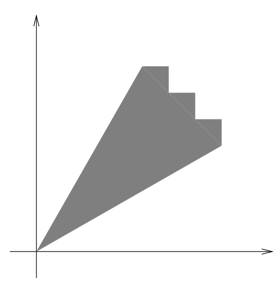
Problem 2.4 gives a more constructive form for $\operatorname{co}(S)$.

Cone

Definition. A set C in a vector space is called a **cone with vertex at the origin** if $\boldsymbol{x} \in C$ implies that $\alpha\boldsymbol{x} \in C$ for all $\alpha \in [0, \infty)$.



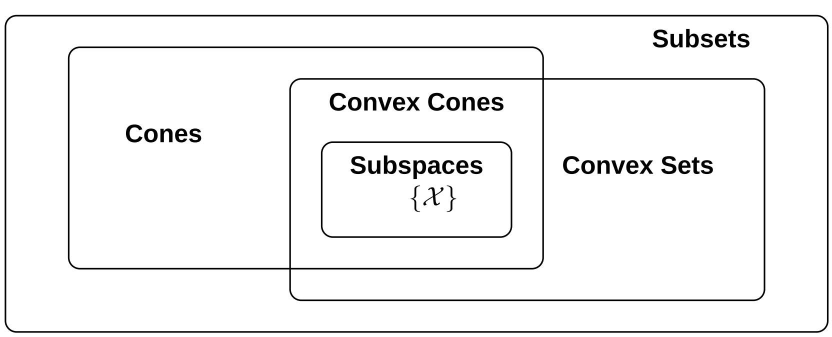
Cone



Convex Cone

Example. The space of nonnegative continuous (real) functions is a **convex cone** in the vector space of continuous functions.

Relationships between cones, convex sets, and subspaces



Exercise. An arbitrary union of subspaces (in a common VS) is a cone.

2.6

Normed linear spaces

We've gone about as far as we can with just the basic axioms of a vector space. Fortunately, most of the vector spaces of interest have additional structure: a norm.

Definition. A **norm** on a vector space $\mathcal{X}$ is a function $\|\cdot\| : \mathcal{X} \to \mathbb{R}$ that satisfies the following for every $\boldsymbol{x} \in \mathcal{X}$.

- $\|\boldsymbol{x}\| \geq 0$ (**nonnegativity**)
- $\|\boldsymbol{x}\| = 0$ iff $\boldsymbol{x} = \mathbf{0}$ (**positive definiteness**)
- $\|\alpha\boldsymbol{x}\| = |\alpha|\,\|\boldsymbol{x}\|$, $\forall \alpha \in \mathcal{F}$ (**scaling property**) (**homogeneity**)
- $\|\boldsymbol{x}+\boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$, $\forall \boldsymbol{y} \in \mathcal{X}$ (**triangle inequality**)

Notice that the absolute value function $|\alpha|$ appears above. It is here where our "restriction" to the fields $\mathbb{R}$ and $\mathbb{C}$ enters[3].

Definition. $(\mathcal{X}, \|\cdot\|)$ is a **normed vector space** or **normed linear space** or **normed linear vector space** or just **normed space**[4].

Clearly a norm generalizes the usual notion of **length**.

The following lemma arises remarkably frequently in proofs.

Lemma. In a normed space, $\|\boldsymbol{x}\| - \|\boldsymbol{y}\| \leq \|\boldsymbol{x}-\boldsymbol{y}\|$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$, *i.e.*, $|\,\|\boldsymbol{x}\| - \|\boldsymbol{y}\|\,| \leq \|\boldsymbol{x}-\boldsymbol{y}\|$.
Proof. $\|\boldsymbol{x}\| = \|\boldsymbol{x}-\boldsymbol{y}+\boldsymbol{y}\| \leq \|\boldsymbol{x}-\boldsymbol{y}\| + \|\boldsymbol{y}\|$.

Example. **Euclidean** n-space: $\mathbb{E}^n \triangleq \left(\mathbb{R}^n, \|\boldsymbol{x}\| = \sqrt{\sum_{i=1}^n x_i^2}\right)$

Example. Euclidean n-space with a weighted norm: $\left(\mathbb{R}^n, \|\boldsymbol{x}\| = \sqrt{\sum_{i=1}^n w_i x_i^2}\right)$ for $w_i > 0$.

Example. A different norm for $\mathbb{R}^n$: $(\mathbb{R}^n, \|\boldsymbol{x}\| = \max_k |x_k|)$
Proof. Clearly this norm is nonnegative, and is zero only if $|x_k| = 0\ \forall k$, *i.e.*, if $\boldsymbol{x} = \mathbf{0}$. And $\max_k |\alpha x_k| = |\alpha| \max_k |x_k|$.
For the triangle inequality, we note that $\max_k |x_k + y_k| \leq \max_k [|x_k| + |y_k|] \leq \max_k |x_k| + \max_k |y_k|$.

Those are all finite-dimensional examples.

Example. The vector space of continuous functions on the interval $[a, b]$ with the norm $\|f\|_2 = \sqrt{\int_a^b |f(t)|^2 \,\mathrm{d}t}$.

What if we replaced "continuous" with "square integrable?"
(see footnote p. 32 regarding equivalence classes of functions that are equal a.e.)

Example. $C[a, b]$. The space of continuous functions on the interval $[a, b]$ with the norm $\|f\|_\infty = \max_{a \leq t \leq b} |f(t)|$.
Later we will see a reason why this "max" norm can be preferable.

Example. Space of real $m \times n$ matrices, with $\|A\| = \sqrt{\mathrm{tr}(AA^T)}$.

Problems involving spaces of matrices arise fairly frequently in systems analysis.

$\mathbb{R}^2$ and $\mathbb{R}$

For $\mathbb{R}^2$ there are three particularly important norms. For $\boldsymbol{x} = (a, b)$ consider:

$\sqrt{a^2+b^2}$	Euclidean or $\mathbb{E}^2$
$\lvert a\rvert + \lvert b\rvert$	ℓ_1 or **city block**
$\max\{\lvert a\rvert, \lvert b\rvert\}$	ℓ_∞

Are there others? Yes, *e.g.*, weighted versions: $\sqrt{w_1 a^2 + w_2 b^2}$, $w_k > 0$.

What about $\mathbb{R}$? One norm is $\|x\| = |x|$. Are there others?
If $g(x)$ is a norm on $\mathbb{R}$, then one condition is that $g(\alpha x) = |\alpha| g(x), \forall \alpha \in \mathbb{R}$.
Thus, choosing $x = 1$ we must have $g(\alpha) = |\alpha| g(1)$, where $g(1) > 0$ since $1 \neq 0$.

Thus all norms on $\mathbb{R}$ have the form $\|x\| = w|x|$ for some $w > 0$.
All such norms are **equivalent**, in a sense to be defined in HW, so there is no point in considering anything but the case $w = 1$.
So we usually just speak of $\mathbb{R}$, rather than $\mathbb{E}^1 = (\mathbb{R}, \|\cdot\| = |\cdot|)$.

---

[3] Probably it is possible to define normed spaces over other fields for which one can define a suitable $|\cdot|$ function that satisfies:
(i) $|\alpha| = 0 \iff \alpha = 0$, (ii) $|\alpha\beta| = |\alpha|\,|\beta|$, (iii) $|\alpha+\beta| \leq |\alpha| + |\beta|$. But would such spaces be useful?

[4] The term normed space should suffice because the presence of a vector space is implied in the axioms of a norm, since vector addition, scalar multiplication, and the zero vector are all part of those axioms.

2.10

The ℓ_p and $\mathcal{L}_p$ spaces

These normed spaces are ubiquitous in the engineering literature, and represent perhaps the most important examples of (infinite dimensional) normed vector spaces.

Definition. Let $p \in [1, \infty)$. The space ℓ_p consists of all sequences of scalars $a_1, a_2, \ldots$ for which $\sum_{i=1}^{\infty} |a_i|^p < \infty$. The norm of a vector $\boldsymbol{x} \in \ell_p$ is defined by

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^{\infty} |a_i|^p \right)^{1/p} < \infty.$$

Definition. The space ℓ_∞ consists of **bounded sequences**. The norm of a vector $\boldsymbol{x} = \{a_i\}$ in ℓ_∞ is defined by

$$\|\boldsymbol{x}\|_\infty = \sup_i |a_i|.$$

Before referring to these spaces as normed spaces, we must confirm that each of the functionals defined above is indeed a **norm**.
- $\|\alpha \boldsymbol{x}\|_p = |\alpha| \, \|\boldsymbol{x}\|_p$ is trivial to verify
- $\|\boldsymbol{x}\|_p > 0$ unless $\boldsymbol{x} = \mathbf{0}$.
- What about the triangle inequality? The holds due to the **Minkowski inequality**, which in turn follows from the Hölder inequality.

Theorem. *(The Hölder inequality)*
If $p \in [1, \infty)$ and $q \in [1, \infty)$ satisfy $1/p + 1/q = 1$, and if $\boldsymbol{x} = (a_1, a_2, \ldots) \in \ell_p$ and $\boldsymbol{y} = (b_1, b_2, \ldots) \in \ell_q$, then

$$\sum_{i=1}^{\infty} |a_i b_i| \le \|\boldsymbol{x}\|_p \, \|\boldsymbol{y}\|_q \,.$$

Moreover, equality holds iff either of the following two conditions hold:
- *either $\boldsymbol{x}$ or $\boldsymbol{y}$ equal $\mathbf{0}$, or*
- *both $\boldsymbol{x}$ and $\boldsymbol{y}$ are nonzero and* $\left(\frac{|a_i|}{\|\boldsymbol{x}\|_p} \right)^{1/q} = \left(\frac{|b_i|}{\|\boldsymbol{y}\|_q} \right)^{1/p}, \; \forall i.$

See errata: Luenberger's equality condition omits the cases where x or y equal 0.

Proof. ***See text.***

The special case $p = q = 2$ is particularly important, and is known as the **Cauchy-Schwarz inequality**:

$$\sum_i |a_i b_i| \le \sqrt{\sum_i |a_i|^2} \sqrt{\sum_i |b_i|^2} \text{ and hence } \left| \sum_i a_i b_i \right| \le \sqrt{\sum_i |a_i|^2} \sqrt{\sum_i |b_i|^2}.$$

Theorem. *(The Minkowski inequality)*
If $\boldsymbol{x}, \boldsymbol{y} \in \ell_p, \; p \in [1, \infty]$ then so is $\boldsymbol{x} + \boldsymbol{y}$, and $\|\boldsymbol{x} + \boldsymbol{y}\|_p \le \|\boldsymbol{x}\|_p + \|\boldsymbol{y}\|_p$.
For $p \in [1, \infty)$, equality holds iff $\boldsymbol{x}$ and $\boldsymbol{y}$ are ***linearly dependent****.*
See errata: Luenberger's condition for equality is incomplete since it omits the cases where x or y equal 0.

Proof. ***See text.*** (It uses the Hölder inequality.)

At one point in the proof, we have the inequality $\left(\sum_{i=1}^n |a_i + b_i|^p\right)^{1/p} \le \left(\sum_{i=1}^n |a_i|^p\right)^{1/p} + \left(\sum_{i=1}^n |b_i|^p\right)^{1/p}$.
Taking the limit as $n \to \infty$ on the RHS (which increases monotonically with n) yields $\left(\sum_{i=1}^n |a_i + b_i|^p\right)^{1/p} \le \|\boldsymbol{x}\|_p + \|\boldsymbol{y}\|_p$.
Then taking the limit as $n \to \infty$ on the LHS yields $\|\boldsymbol{x} + \boldsymbol{y}\|_p \le \|\boldsymbol{x}\|_p + \|\boldsymbol{y}\|_p$.
This technique of first working with one side and then the other arises frequently.

The trouble with Riemann integration

The $\mathcal{L}_p$ space, defined below, involves Lebesgue integration. Why? Because the set of Riemann integrable functions is not as general as we would like. For example, consider the function $f(t) = \begin{cases} 1, & t \in \mathbb{Q} \\ 0, & \text{otherwise.} \end{cases}$ This function is not Riemann integrable, but it is Lebesgue integrable. (Its Lebesgue integral is zero since the function is nonzero only on a set of Lebesgue measure zero.) When a function is Riemann integrable, its Lebesgue integral will equal its Riemann integral.

Considering Lebesgue integrable functions will more than general enough for any engineering problems.

The $\mathcal{L}_p$ space

Definition. Let $p \in [1, \infty)$. The space $\mathcal{L}_p[a, b]$ consists of all real-valued **measurable** functions $\boldsymbol{x}$ on the interval $[a, b]$ for which $|x(t)|^p$ is Lebesgue integrable. The norm on this space is defined as

$$\|\boldsymbol{x}\|_p = \left(\int_a^b |x(t)|^p \, \mathrm{d}t \right)^{1/p}.$$

However, there is a subtle caveat with this space. There are functions that are nonzero on a set of measure zero for which $\|\boldsymbol{x}\|_p = 0$. So to consider this space to be a normed space, we must treat functions that are equal almost everywhere (a.e.) as being equivalent.

In other words, a vector in $\mathcal{L}_p$ is really an **equivalence class** of measurable functions that are all equal almost everywhere.

This makes the definition of $\mathcal{L}_\infty[a, b]$ a bit more subtle than ℓ_∞. In particular, we cannot define $\|\boldsymbol{x}\|_\infty$ to be simply the obvious choice "$\sup_{a \le t \le b} |x(t)|$," because that value will be different for different functions in the equivalence class. Instead, we define

$$\|\boldsymbol{x}\|_\infty = \text{essential supremum of } |x(t)| = \underset{y(t)=x(t) \text{ a.e.}}{\text{infimum}} \sup |y(t)| = \text{ess sup } |x(t)| .$$

See [4, p. 33] for an example.

If $\boldsymbol{x} \in \mathcal{L}_p[a, b]$ and $\boldsymbol{y} \in \mathcal{L}_q[a, b]$ with $p, q > 1$ and $1/p + 1/q = 1$, then the Hölder inequality is $\int_a^b |x(t)y(t)| \, \mathrm{d}t \le \|\boldsymbol{x}\|_p \|\boldsymbol{y}\|_q$.

Similarly, $\|\boldsymbol{x} + \boldsymbol{y}\|_p \le \|\boldsymbol{x}\|_p + \|\boldsymbol{y}\|_p$ is the **Minkowski inequality** for $\mathcal{L}_p$. The proofs are very similar.

Topological concepts

According to the Oxford English dictionary, **topology** is the "branch of mathematics concerned with those properties of figures and surfaces which are independent of size and shape and are unchanged by any deformation that is continuous, neither creating new points nor fusing existing ones; hence, with those of abstract spaces that are invariant under homeomorphic transformations."

Anyway, we have been systematically generalizing notions from geometry to more general settings, but one important concept we have yet to generalize is that of **distance**.

The concept of distance is central to any discussion of **optimization**, since many optimization problems involve finding the element (within a set) that is the *closest* to some given point outside that set.

Definition. If $\boldsymbol{x}$ and $\boldsymbol{y}$ are two points in a normed space $(\mathcal{X}, \|\cdot\|)$, then the **distance**[5] between $\boldsymbol{x}$ and $\boldsymbol{y}$ is defined by

$$d(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\|.$$

More generally, if S is a nonempty subset of $\mathcal{X}$, then the **distance** between a point $\boldsymbol{x} \in \mathcal{X}$ and the set S is defined by

$$d(\boldsymbol{x}, S) = \inf_{\boldsymbol{y} \in S} \|\boldsymbol{x} - \boldsymbol{y}\|.$$

Example. In $\mathbb{R}$, $d(1, (3, 4]) = 2$, for the usual norm $\|x\| = |x|$. ***(Picture)***
Note that there is no "closest point," a complication that will require careful attention later.

Example. In $\mathbb{R}^2$, what is the distance between $(0, 0)$ and $\{(a, b) : 2 \leq a \leq 3,\ 1 \leq b \leq 2\}$? (Trick question) **??**

How would you define the distance between two sets S and T? **??**

Properties of distance d. (We will use these later.)

Lemma 2.1 $|d(\boldsymbol{x}, S) - d(\boldsymbol{y}, S)| \leq \|\boldsymbol{x} - \boldsymbol{y}\|$. ***(Picture)***

Proof. $d(\boldsymbol{x}, S) = \inf_{\boldsymbol{z} \in S} \|\boldsymbol{x} - \boldsymbol{z}\| = \inf_{\boldsymbol{z} \in S} \|(\boldsymbol{x} - \boldsymbol{y}) - (\boldsymbol{z} - \boldsymbol{y})\| \geq \inf_{\boldsymbol{z} \in S} \|\boldsymbol{z} - \boldsymbol{y}\| - \|\boldsymbol{x} - \boldsymbol{y}\| = d(\boldsymbol{y}, S) - \|\boldsymbol{x} - \boldsymbol{y}\|$. Now rearrange. □

Lemma 2.2 *For any two subsets U and V of a normed space $\mathcal{X}$, $d(U, V) \leq d(\boldsymbol{x}, U) + d(\boldsymbol{x}, V)$, $\forall \boldsymbol{x} \in \mathcal{X}$.* ***(Picture)***

Proof. $d(U, V) = \inf_{\boldsymbol{u} \in U} d(\boldsymbol{u}, V) \leq \inf_{\boldsymbol{u} \in U} d(\boldsymbol{x}, V) + \|\boldsymbol{x} - \boldsymbol{u}\| = d(\boldsymbol{x}, V) + d(\boldsymbol{x}, U)$. □

Lemma. Let $(\mathcal{X}, \|\cdot\|)$ be a normed space, $M \subseteq \mathcal{X}$ a subspace, and $\boldsymbol{x} \in \mathcal{X}$ a point. Then $\forall \alpha \in \mathcal{F}$: $d(\alpha \boldsymbol{x}, M) = |\alpha|\, d(\boldsymbol{x}, M)$.
Proof. Trivial for $\alpha = 0$.
For $\alpha \neq 0$: $d(\alpha \boldsymbol{x}, M) = \inf_{\boldsymbol{y} \in M} \|\alpha \boldsymbol{x} - \boldsymbol{y}\| = \inf_{\boldsymbol{z} \in M} \|\alpha \boldsymbol{x} - \alpha \boldsymbol{z}\| = |\alpha| \inf_{\boldsymbol{z} \in M} \|\boldsymbol{x} - \boldsymbol{z}\| = |\alpha|\, d(\boldsymbol{x}, M)$ ***(Picture)*** . □

Preview of optimization

Many problems in optimization can be expressed as follows.
Given $\boldsymbol{x}$ in a normed space $(\mathcal{X}, \|\cdot\|)$, and a subset S in $\mathcal{X}$, find "the" vector $\boldsymbol{s} \in S$ that minimizes $\|\boldsymbol{x} - \boldsymbol{s}\|$.

What questions should we ask about such problems?
- Is there any best $\boldsymbol{s}$? I.e., does there exist $\boldsymbol{s}^\star \in S$ s.t. $\|\boldsymbol{x} - \boldsymbol{s}^\star\| = d(\boldsymbol{x}, S)$?
- If so, is $\boldsymbol{s}^\star$ unique?
- How is $\boldsymbol{s}^\star$ characterized? (Better yet would be an explicit formula for $\boldsymbol{s}^\star$.)

One purpose of some of the material that follows is to answer these questions. (But be prepared for a challenge since some aspects of these questions remain open problems!) We will return to these questions after introducing Hilbert spaces in *Ch. 3*.

[5] This type of distance is a special case of the more general concept of a distance function that is called a **metric** if it satisfies:
- $d(\boldsymbol{x}, \boldsymbol{x}) = 0$
- $d(\boldsymbol{x}, \boldsymbol{y}) > 0$ if $\boldsymbol{x} \neq \boldsymbol{y}$
- $d(\boldsymbol{x}, \boldsymbol{y}) = d(\boldsymbol{y}, \boldsymbol{x})$
- $d(\boldsymbol{x}, \boldsymbol{y}) \leq d(\boldsymbol{x}, \boldsymbol{z}) + d(\boldsymbol{z}, \boldsymbol{y})$ (**triangle inequality**)

Exercise. Verify that a metric is a more general concept than a norm, *i.e.*, find a metric $d(\boldsymbol{x}, \boldsymbol{y})$ that is not of the form $d(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\|$ for any norm $\|\cdot\|$.
Hint: a metric need not satisfy the scaling property. **??**

2.7

Open sets

Looking ahead: in optimization we often use iterative algorithms that generate sequences, and we need to know when a sequence converges to a limit that belongs to the same set as the sequence itself. Such questions are related to whether a set is closed or not, so we need concepts of open and closed sets.

Definition. The **open sphere** centered at $\boldsymbol{x}$ of radius ε is defined by $\boxed{S(\boldsymbol{x}, \varepsilon) \triangleq \{\boldsymbol{y} \in \mathcal{X} : \|\boldsymbol{x} - \boldsymbol{y}\| < \varepsilon\}.}$
Also called the **open ball**, which is perhaps more descriptive since a sphere is often considered to be a surface, whereas a ball might be more often considered to be a solid.

Example. Consider $(\mathcal{X} = C[0,1], \|\cdot\|_\infty)$ and $\boldsymbol{x} = \boldsymbol{0}$.
Then $S(\boldsymbol{x}, 2) = \left\{\text{continuous functions } f : [0,1] \to \mathbb{R} : \max_{t\in[0,1]} |f(t)| < 2\right\}$. ***(Picture of continuous functions between*** ± 2***.)***

Clearly this definition of a "sphere" generalizes the 3-space notion of a sphere!
Now we need a definition for "inside," another generalization.

Definition. Let P be a subset of a normed space $(\mathcal{X}, \|\cdot\|)$.
A point $\boldsymbol{p} \in P$ is called an **interior point** of P iff $\exists \varepsilon > 0$ s.t. $S(\boldsymbol{p}, \varepsilon) \subseteq P$. ***(Picture in 2D with*** $\odot$ ***in a generic set.)***

Example. Let $\mathcal{X} = \left\{(a,b) \in \mathbb{R}^2 : a = b\right\}$ (the line at 45° in the plane) with $\|\boldsymbol{x}\| = |a|$ if $\boldsymbol{x} = (a, a)$.
Is $\boldsymbol{x} = (1,1)$ an interior point of $\mathcal{X}$, *i.e.*, does there exist an open ball that is a subset of $\mathcal{X}$? **??**

Suppose we change to $\mathcal{X} = \mathbb{R}^2$ with the usual Euclidean norm and $S = \left\{(a,b) \in \mathbb{R}^2 : a = b\right\}$.
Is $\boldsymbol{x} = (1,1)$ an interior point of S? **??**

Which points within an open sphere are interior points? **??**

Definition. The **interior** of a set P, denoted $\operatorname{Int}(P)$, is the collection of all interior points of P.

$$\operatorname{Int}(P) \triangleq \{\boldsymbol{p} \in P : \boldsymbol{p} \text{ is an interior point of } P\} = \{\boldsymbol{p} \in P : \exists \varepsilon > 0 \text{ s.t. } S(\boldsymbol{p}, \varepsilon) \subseteq P\}.$$

Definition. A set P is called **open** iff $P = \operatorname{Int}(P)$.

Remark. $\operatorname{Int}(P) \subseteq P$, so to show a set P is open we must show $P \subseteq \operatorname{Int}(P)$, *i.e.*, show that $\boldsymbol{x} \in P \Longrightarrow \exists \varepsilon > 0$ s.t. $S(\boldsymbol{x}, \varepsilon) \subset P$.

Examples of open sets.
- $\emptyset$ and $\mathcal{X}$ are open sets
- open spheres are open (exercise)
- $\operatorname{Int}(\operatorname{Int}(P)) = \operatorname{Int}(P)$
- $\operatorname{Int}(P)$ is open
- $P \times Q$ is open in $\mathcal{X} \times \mathcal{Y}$ if P is open in $\mathcal{X}$ and Q is open in $\mathcal{Y}$ (problem 2.16 p. 44)
- $(1,2) \times (0,1) \subset \mathbb{E}^2$ is open

Proposition. $\boldsymbol{x} \in \operatorname{Int}(P) \iff d(\boldsymbol{x}, \widetilde{P}) > 0$. ***(Picture)***

Proof. $\boldsymbol{x} \in \operatorname{Int}(P) \Longrightarrow \exists \varepsilon > 0$ s.t. $S(\boldsymbol{x}, \varepsilon) \subseteq P$, so $d(\boldsymbol{x}, \widetilde{P}) = \inf_{\boldsymbol{y}\in\widetilde{P}} \|\boldsymbol{x} - \boldsymbol{y}\| = \inf_{\boldsymbol{y}\notin P} \|\boldsymbol{x} - \boldsymbol{y}\| \geq \varepsilon$,
since $\boldsymbol{y} \notin P \Longrightarrow \boldsymbol{y} \notin S(\boldsymbol{x}, \varepsilon) \Longrightarrow \|\boldsymbol{y} - \boldsymbol{x}\| \geq \varepsilon$.

We can prove the **converse** by proving its **contrapositive**[6].
Suppose $\boldsymbol{x} \notin \operatorname{Int}(P)$. Then $\forall \varepsilon > 0$, $S(\boldsymbol{x}, \varepsilon) \cap \widetilde{P} \neq \emptyset$, so $\exists \boldsymbol{y} \in \widetilde{P}$ s.t. $\|\boldsymbol{x} - \boldsymbol{y}\| < \varepsilon$. Thus $d(\boldsymbol{x}, \widetilde{P}) = \inf_{\boldsymbol{y}\in\widetilde{P}} \|\boldsymbol{x} - \boldsymbol{y}\| = 0$.

Alternative direct proof. $\delta = d(\boldsymbol{x}, \widetilde{P}) > 0 \Longrightarrow \|\boldsymbol{z} - \boldsymbol{x}\| \geq \delta > 0, \forall \boldsymbol{z} \in \widetilde{P}$. But $\boldsymbol{y} \in S(\boldsymbol{x}, \delta) \Longrightarrow \|\boldsymbol{y} - \boldsymbol{x}\| < \delta \Longrightarrow \boldsymbol{y} \notin \widetilde{P} \Longrightarrow \boldsymbol{y} \in P$. So $S(\boldsymbol{x}, \delta) \subseteq P$ and hence $\boldsymbol{x} \in \operatorname{Int}(P)$. □

Fact. Similarly: $\boldsymbol{x} \in \operatorname{Int}\left(\widetilde{P}\right) \iff d(\boldsymbol{x}, P) > 0$, since $\widetilde{\widetilde{P}} = P$.

[6] We have shown that $A \Longrightarrow B$ and we want to show the converse $B \Longrightarrow A$, the contrapositive of which is: not A $\Longrightarrow$ not B.

Closed sets

Definition. A point $\boldsymbol{x} \in \mathcal{X}$ is called a **closure point** (or a **cluster point** or an **adherent point**) of a set P iff

$$\forall \varepsilon > 0, \; \exists \boldsymbol{p} \in P \text{ s.t. } \|\boldsymbol{x} - \boldsymbol{p}\| < \varepsilon.$$

Fact. If P is nonempty, then $\boldsymbol{x}$ is a closure point of P iff $d(\boldsymbol{x}, P) = 0$.

Definition. The collection of all closure points of a set P is called the **closure** of P and is denoted $\overline{P}$.

Example. $P = \{1, 1/2, 1/3, \ldots, 1/10\} \Longrightarrow \overline{P} = P$.
Example. $P = [0, 1) \Longrightarrow \overline{P} = [0, 1]$.
Example. $P = \{1/n : n \in \mathbb{N}\} \Longrightarrow \overline{P} = P \cup \{0\}$.
Example. $\overline{\mathbb{Q}} = \mathbb{R}$.

Properties.
- $\overline{P} = \{\boldsymbol{x} \in \mathcal{X} : d(\boldsymbol{x}, P) = 0\}$ if P is nonempty. (If P is empty, then so is $\overline{P}$.)
- $P \subseteq \overline{P}$
- $\overline{\overline{P}} = \overline{P}$

Definition. A set P is called **closed** iff $P = \overline{P}$.

Remark. Since $P \subseteq \overline{P}$, to show a set P is closed we must show $\overline{P} \subseteq P$, *i.e.*, if $\boldsymbol{p}$ is any closure point of P, then $\boldsymbol{p} \in P$.

Examples of closed sets.
- $\emptyset$ is closed (!)
- $\mathcal{X}$ is closed (!)
- $\{\boldsymbol{x}\}$ is closed
- $\{\boldsymbol{y} : d(\boldsymbol{x}, \boldsymbol{y}) \leq 1\}$ is closed

Proposition.
- P is open $\Longrightarrow \widetilde{P}$ is closed
- P is closed $\Longrightarrow \widetilde{P}$ is open

Proof. Suppose P is open. $\boldsymbol{x} \in P \Longrightarrow \exists \varepsilon > 0$ s.t. $S(\boldsymbol{x}, \varepsilon) \subseteq P \Longrightarrow d(\boldsymbol{x}, \widetilde{P}) \geq \varepsilon > 0 \Longrightarrow \boldsymbol{x} \notin \overline{\widetilde{P}}$.
Thus the contrapositive is $\boldsymbol{x} \in \overline{\widetilde{P}} \Longrightarrow \boldsymbol{x} \notin P$ *i.e.*, $\overline{\widetilde{P}} \subseteq \widetilde{P}$. Hence $\widetilde{P}$ is closed.

Suppose P is closed. $\boldsymbol{x} \in \widetilde{P} \Longrightarrow \boldsymbol{x} \notin P = \overline{P} \Longrightarrow \varepsilon \triangleq d(\boldsymbol{x}, P) > 0$. Thus $S(\boldsymbol{x}, \varepsilon) \subset \widetilde{P}$ so $\boldsymbol{x} \in \operatorname{Int}\left(\widetilde{P}\right)$.
Thus $\widetilde{P} \subseteq \operatorname{Int}\left(\widetilde{P}\right)$ and we conclude $\widetilde{P}$ is open. ***(See Luenberger errata for p. 25.)*** □

Remark. We could also state the proposition simply as "P is open $\iff$ $\widetilde{P}$ is closed."

Proposition.
- The intersection of a *finite number* of open sets is open.
- The union of a *finite number* of closed sets is closed.
- The intersection of an *arbitrary number* of closed sets is closed.
- The union of an *arbitrary number* of open sets is open.

Proof. **Exercise.**

Example. Consider the open intervals $S_n = (-1/n, 1/n) \subset \mathbb{R}$, $n = 1, 2, \ldots$.
Then $\cup_{n=1}^{\infty} S_n = (-1, 1)$, which is open. And $\cap_{n=1}^{N} S_n = S_N$, which is open. What is $\cap_{n=1}^{\infty} S_n$? **??**

Now we consider open and closedness in the context of the special types of sets we considered previously: convex sets and subspaces.

Proposition. If C is a convex set in a normed space, then $\overline{C}$ and $\mathrm{Int}(C)$ are convex.

Proof. ***see text***

Open and closed subspaces

Fact. If M is a subspace of $\mathcal{X}$ and M is open, then $M = \mathcal{X}$ [3, p. 229].

Fact. If M is a subspace of $\mathcal{X}$, then $\overline{M}$ is a subspace of $\mathcal{X}$ [3, p. 229].

In general, subspaces are not necessarily closed in infinite dimensional normed spaces.

Example. The subspace M of continuous functions in $\mathcal{L}_2$.
Because the rect function is in $\mathcal{L}_2$ and is a closure point of M but is not in M. ***(Picture)*** .

Example. The subspace M of finite-length sequences in ℓ_2.
Because the infinite geometric series $(1, 1/2, 1/4, 1/8, \ldots)$ is in ℓ_2, and is a closure point of M, but is not in M.

Caution! We have now seen one of our first examples of a situation where our intuition from $\mathbb{E}^3$ does not generalize to general normed spaces! We have been thinking of subspaces as being like planes or hyperplanes. Yet even though planes are closed subsets of $\mathbb{E}^3$, subspaces are not necessarily closed in general.

Remark. We had to use infinite-dimensional examples above because (as we will show soon), finite-dimensional subspaces are closed.

Closed sets and distances

Lemma. In a normed space, let U and V be disjoint subsets. If V is closed, then $d(\boldsymbol{u}, V) > 0, \forall \boldsymbol{u} \in U$.
Proof. Pick any $\boldsymbol{u} \in U$ and suppose $d(\boldsymbol{u}, V) = 0$. Then $\boldsymbol{u} \in \overline{V} = V$ since V is closed.
But $\boldsymbol{u} \in V$ contradicts the assumption that U and V are disjoint. □

Lemma 2.3 *In a normed space,* $d(\boldsymbol{y}, S) > 0 \Longrightarrow d(\boldsymbol{y}, \overline{S}) > 0$.

??

Bounded sets

Definition. A set S in a normed space $(\mathcal{X}, \|\cdot\|)$ is called **bounded** iff $\exists M < \infty$ such that $\|\boldsymbol{x}\| \leq M,\ \forall \boldsymbol{x} \in S$.

Are closed sets bounded? **??**

Bounded sets can be open, or closed, or neither.

Sequences

Definition. A **sequence** is a set of vectors indexed by the natural numbers $\mathbb{N}$, *e.g.*, $\{\boldsymbol{x}_n\} = \{\boldsymbol{x}_n : n \in \mathbb{N}\}$.
Formally, a sequence is a mapping from $\mathbb{N}$ to some vector space $\mathcal{X}$.

Definition. If $\{\boldsymbol{x}_n\}$ is a sequence, and $n_1 < n_2 < \cdots$, then $\{\boldsymbol{x}_{n_i}\}$ is called a **subsequence** of $\{\boldsymbol{x}_n\}$.

Notation. $\{\boldsymbol{x}_n\} \in \mathcal{X}$ iff $\boldsymbol{x}_n \in \mathcal{X},\ \forall n$

2.8

Convergence of sequences

Definition. In a normed space, we say a sequence of vectors $\{\boldsymbol{x}_n\}$ **converges** to a vector $\boldsymbol{x}$ iff the sequence of real numbers $\|\boldsymbol{x}_n - \boldsymbol{x}\|$ converges to zero, in which case we write $\boldsymbol{x}_n \to \boldsymbol{x}$ or $\lim_{n\to\infty} \boldsymbol{x}_n = \boldsymbol{x}$.
In other words, considering the definition of convergence of real numbers:

$$\begin{aligned} \boldsymbol{x}_n \to \boldsymbol{x} \quad &\iff \quad \|\boldsymbol{x}_n - \boldsymbol{x}\| \to 0 \\ &\iff \quad \forall \varepsilon > 0,\ \exists N_\varepsilon < \infty \text{ s.t. } n \geq N_\varepsilon \Longrightarrow \|\boldsymbol{x}_n - \boldsymbol{x}\| < \varepsilon \\ &\iff \quad \forall \varepsilon > 0,\ \exists N_\varepsilon < \infty \text{ s.t. } n \geq N_\varepsilon \Longrightarrow \boldsymbol{x}_n \in S(\boldsymbol{x}, \varepsilon). \end{aligned}$$

Example. (Infinite dimensional, of course.) Consider $\mathcal{L}_2[0,1]$ with $x(t) = 1$ and $x_n(t) = 1 + t/n$.
Then $\boldsymbol{x}_n \to \boldsymbol{x}$ because $\|\boldsymbol{x} - \boldsymbol{x}_n\|_2 = \sqrt{\int_0^1 |x(t) - x_n(t)|^2 \,\mathrm{d}t} = \sqrt{\int_0^1 |t/n|^2 \,\mathrm{d}t} = \frac{1}{\sqrt{3}\,n} \to 0$ as $n \to \infty$.

Example. (Infinite dimensional, of course.) Consider ℓ_p with $\boldsymbol{x} = (1, 1/2, 1/3, \ldots)$ and $\boldsymbol{x}_n = (1, 1/2, \ldots, 1/n, 0, 0, \ldots)$.
Then $\|\boldsymbol{x}_n - \boldsymbol{x}\|_p = \left(\sum_{k=n+1}^{\infty} 1/k^p\right)^{1/p}$. That power series is convergent (and hence goes to zero as $n \to \infty$) for $p > 1$, but diverges for $p = 1$. So we can say "$\boldsymbol{x}_n \to \boldsymbol{x}$" in ℓ_2, for example, but we cannot say that in ℓ_1!

So all norms are not equivalent in general, unlike in finite-dimensional vector spaces.

Proposition. The limit of a convergent sequence is unique.

Proof. Suppose $\boldsymbol{x}_n \to \boldsymbol{x}$ and $\boldsymbol{x}_n \to \boldsymbol{y}$. Then $\forall n$:

$$\|\boldsymbol{x} - \boldsymbol{y}\| = \|\boldsymbol{x} - \boldsymbol{x}_n + \boldsymbol{x}_n - \boldsymbol{y}\| \leq \|\boldsymbol{x} - \boldsymbol{x}_n\| + \|\boldsymbol{x}_n - \boldsymbol{y}\|$$

which $\to 0$ as $n \to \infty$. Thus $\|\boldsymbol{x} - \boldsymbol{y}\| = 0$ so $\boldsymbol{x} - \boldsymbol{y} = \boldsymbol{0}$ so $\boldsymbol{x} = \boldsymbol{y}$. □

Proposition. $\boldsymbol{x}_n \to \boldsymbol{x} \Longrightarrow \|\boldsymbol{x}_n\| \to \|\boldsymbol{x}\|$.

Proof. As shown earlier, $|\,\|\boldsymbol{x}\| - \|\boldsymbol{y}\|\,| \leq \|\boldsymbol{x} - \boldsymbol{y}\|$. Thus, $|\|\boldsymbol{x}_n\| - \|\boldsymbol{x}\|| \leq \|\boldsymbol{x}_n - \boldsymbol{x}\| \to 0$ if $\boldsymbol{x}_n \to \boldsymbol{x}$, so $\|\boldsymbol{x}_n\| \to \|\boldsymbol{x}\|$. □

Proposition. If $\boldsymbol{x}_n \to \boldsymbol{x}$, then $\sup_n \|\boldsymbol{x}_n\| < \infty$, *i.e.*, convergent sequences are **bounded**.

Proof. $\exists N_1$ s.t. $\|\boldsymbol{x}_n - \boldsymbol{x}\| < 1,\ \forall n > N_1$. Thus $\|\boldsymbol{x}_n\| = \|\boldsymbol{x}_n - \boldsymbol{x} + \boldsymbol{x}\| \leq \|\boldsymbol{x}_n - \boldsymbol{x}\| + \|\boldsymbol{x}\| < 1 + \|\boldsymbol{x}\|$,
so $\sup_n \|\boldsymbol{x}_n\| \leq \max\{1 + \|\boldsymbol{x}\|, \|\boldsymbol{x}_1\|, \ldots, \|\boldsymbol{x}_{N_1}\|\} < \infty$. ***(Picture)*** □

The concept of a closed set is closely connected with limits of sequences.

Definition. If S is a subset of $\mathcal{X}$, we call $\boldsymbol{x} \in \mathcal{X}$ a **limit point** of S iff there is a sequence of elements of S that converges to $\boldsymbol{x}$.

Proposition. $\boldsymbol{x}$ is a limit point of S iff $\boldsymbol{x} \in \overline{S}$. *Thus limit points and cluster points are equivalent!*

Proof. If $\boldsymbol{x}$ is a limit point, then $\exists \{\boldsymbol{x}_n\} \in S$ s.t. $\boldsymbol{x}_n \to \boldsymbol{x}$. Thus $d(\boldsymbol{x}, \{\boldsymbol{x}_n\}_{n=1}^\infty) = 0$, so[7] $d(\boldsymbol{x}, S) = 0$. Hence $\boldsymbol{x} \in \overline{S}$.
Suppose $\boldsymbol{x} \in \overline{S}$. Then $\forall \varepsilon > 0,\ \exists \boldsymbol{y} \in S$ s.t. $\|\boldsymbol{x} - \boldsymbol{y}\| < \varepsilon$. Choose $\varepsilon = 1/n$ and identify $\boldsymbol{x}_n$ with the corresponding $\boldsymbol{y}$.
Since $\|\boldsymbol{x}_n - \boldsymbol{x}\| < 1/n$, we see $\boldsymbol{x}_n \to \boldsymbol{x}$, so $\boldsymbol{x}$ is a limit point of S. □

Corollary. A set is closed iff it contains its limit points.

Summary
- distance
- open sphere
- interior point & interior
- open set = its interior
- closure point: $d(x, P) = 0$, & closure
- closed set = its closure
- bounded set
- convergence sequences
- limit point of a set (exists convergent sequence to it) = cluster point
- closed sets contain their limit points

Series (A special kind of sequence.)

Definition. An **infinite series** of the form $\sum_{i=1}^\infty \boldsymbol{x}_i$ is said to **converge** to $\boldsymbol{x}$ in a normed space iff the sequence of partial sums $\boldsymbol{s}_n = \sum_{i=1}^n \boldsymbol{x}_i$ converges to $\boldsymbol{x}$, in which case we write $\boldsymbol{x} = \sum_{i=1}^\infty \boldsymbol{x}_i$ as short hand for $\boldsymbol{x} = \lim_{n\to\infty} \sum_{i=1}^n \boldsymbol{x}_i$.

Caution. Sometimes as a lazy short hand one might write $\sum_{i=1}^\infty 1/i = \infty$. Saying that $\sum_{i=1}^\infty \alpha_i = \infty$ is shorthand for saying $\forall M < \infty,\ \exists N \in \mathbb{N}$ s.t. $n > N \Longrightarrow \sum_{i=1}^n \alpha_i > M$.
Since $\infty \notin \mathbb{R}$, it is truly an abuse of notation to write $\sum_{n=1}^\infty 1/n = \infty$, because in the above definition of convergence of an infinite series, it is implied that the limit $\boldsymbol{x}$ is an element of $\mathcal{X}$.

Example. In ℓ_2, consider $\boldsymbol{x} = (1, 1/2, 1/3, \ldots)$ and $\boldsymbol{x}_n = (0, \ldots, 0, 1/n, 0, \ldots)$ where $1/n$ is in the nth element.
Then $\sum_{i=1}^\infty \boldsymbol{x}_i$ converges to $\boldsymbol{x}$. (But it does not converge in ℓ_1.)

[7] $A \subseteq B \Longrightarrow d(\boldsymbol{x}, B) \le d(\boldsymbol{x}, A)$. ***(Picture)***

2.11

Cauchy sequences

Often in convergence analysis it is easier to examine $\|\boldsymbol{x}_n - \boldsymbol{x}_m\|$ than $\|\boldsymbol{x}_n - \boldsymbol{x}_\star\|$, especially if we have not yet shown that a limit $\boldsymbol{x}_\star$ even exists.

Definition. A sequence $\{\boldsymbol{x}_n\}$ in a normed space is called a **Cauchy sequence** iff

$$\|\boldsymbol{x}_n - \boldsymbol{x}_m\| \to 0 \text{ as } n, m \to \infty.$$

In other words,

$$\forall \varepsilon > 0,\ \exists N > 0 \text{ s.t. } n, m > N \implies \|\boldsymbol{x}_n - \boldsymbol{x}_m\| < \varepsilon.$$

Fact. In a normed space, every convergent sequence is a Cauchy sequence, since if $\boldsymbol{x}_n \to \boldsymbol{x}$, then

$$\|\boldsymbol{x}_n - \boldsymbol{x}_m\| \leq \|\boldsymbol{x}_n - \boldsymbol{x}\| + \|\boldsymbol{x}_m - \boldsymbol{x}\| \to 0.$$

Note the repeated use of the triangle inequality in proofs. In a normed space, it is about all we have to work with!

The *converse* of the above fact is *not* true in general: Cauchy sequences need not converge in general normed spaces. This is a bit unfortunate since the converse is what we would really like to use usually!

Example. Consider the sequence of functions $f_n(t) = 1 - e^{-nt}$ in the normed space

$$\left(\mathcal{X} = \{f : [0,1] \to \mathbb{R} : f \text{ continuous}\}, \|f\| = \|f\|_1 = \int_0^1 |f(t)|\, \mathrm{d}t\right), \textbf{\textit{(Picture)}}\ .$$

$\|f_n - f_m\| = \int_0^1 |e^{-nt} - e^{-mt}|\, \mathrm{d}t = |(1 - e^{-n})/n - (1 - e^{-m})/m| \to 0$ as $n, m \to \infty$. So $\{f_n\}$ is Cauchy.
But the "apparent limit" of f_n is a step function, which is not continuous, and hence not an element of $\mathcal{X}$.

How can we "fix" this problem?

- Broaden the vector space $\mathcal{X}$ to $\mathcal{L}_1[0,1]$ (*i.e.*, drop the continuity requirement).
- Replace the norm $\|f\|_1$ with $\|f\|_\infty$.

One can show for this example that $\|f_n - f_m\|_\infty$ does not approach zero as "$n, m \to \infty$."
(For any fixed m, $\|f_n - f_m\|_\infty \to 1$ as $n \to \infty$.)

This "incompleteness" can also arise even in subspaces of ℓ_2.

Example. $\mathcal{X}$ = set of infinite sequence of reals with only finitely many nonzero terms:

$$\mathcal{X} = \{(a_1, \ldots, a_k, 0, 0, \ldots) : k \geq 1,\ a_i \in \mathbb{R}\}$$

with the ℓ_p norm $\|\boldsymbol{x}\|_p$. Now consider the sequence with elements $\boldsymbol{x}_n = (1, 1/2, 1/4, \ldots, 1/2^n, 0, 0, \ldots)$.
Since $\|\boldsymbol{x}_n - \boldsymbol{x}_m\|_p = \left(\sum_{k=\min\{n+1,m+1\}}^{\max\{n,m\}} \left(\frac{1}{2^k}\right)^p\right)^{1/p} \leq \left(\frac{1}{2^{p \min\{m,n\}}} \frac{1}{1-1/2^p}\right)^{1/p} \to 0$ as $n, m \to \infty$, $\{\boldsymbol{x}_n\}$ is Cauchy. But there is no $\boldsymbol{x} \in \mathcal{X}$ to which $\{\boldsymbol{x}_n\}$ converges.

In some sense, the problem is that $\mathcal{X}$ has "holes" in it. Broadening the space is the natural solution.
Is there an alternate norm that would make this vector space $\mathcal{X}$ complete? (I doubt it, can you show otherwise?)

2.11

Banach spaces

Often we prefer to use normed spaces that are free of the pathologies described in the preceding examples, so we name them.

Definition. A normed space $(\mathcal{X}, \|\cdot\|)$ is called **complete** iff every Cauchy sequence in $\mathcal{X}$ has a limit in $\mathcal{X}$ (and hence converges).

Definition. A complete normed space is called a **Banach space**.

Examples of Banach spaces

- $\left(\mathbb{R}^n, \|\cdot\|_p\right)$, for $p \in [1, \infty]$
- $C[a, b]$ with $\|f\|_\infty = \sup_{t\in[a,b]} |f(t)|$
- ℓ_p for $p \in [1, \infty]$ with usual $\|\boldsymbol{x}\|_p$
- $\mathcal{L}_p[a, b]$ for $p \in [1, \infty]$ with usual $\|f\|_p$

Is finding a suitable Banach space usually difficult? Fortunately not, since every normed space has a **completion**.

Theorem. *If $(\mathcal{X}, \|\cdot\|_\mathcal{X})$ is a normed space, then $\exists \left(\mathcal{Y}, \|\cdot\|_\mathcal{Y}\right)$ that is a Banach space (called a* ***completion*** *of $\mathcal{X}$) with*
- *$\mathcal{X}$ is a subspace of $\mathcal{Y}$*
- *$\overline{\mathcal{X}} = \mathcal{Y}$*
- *$\boldsymbol{x} \in \mathcal{X} \Longrightarrow \|\boldsymbol{x}\|_\mathcal{X} = \|\boldsymbol{x}\|_\mathcal{Y}$.*

*Moreover, $\mathcal{Y}$ is essentially unique (*i.e.*, all the completions of $\mathcal{X}$ are isometric with one another [3, p. 121]).*

Example. $\mathcal{L}_1[0, 1]$ is the completion of $\left(\mathcal{X} = \{f : [0, 1] \to \mathbb{R} : f \text{ continuous}\}, \|f\| = \|f\|_1 = \int_0^1 |f(t)| \,\mathrm{d}t\right)$.

Example. $\mathcal{L}_p[a, b]$ is the completion of $R_p[a, b]$ for $p \in [1, \infty]$.

Showing such results for $\mathcal{L}_p$ requires measure theory.

Showing completeness of $C[a, b]$ uses the following fact that is a key result from Math 451.

Theorem. $(\mathbb{R}, |\cdot|)$ *is complete,* i.e., *if $\{\alpha_n\} \subset \mathbb{R}$ is Cauchy, then $\exists \alpha \in \mathbb{R}$ s.t. $\alpha_n \to \alpha$.*

There exist bounded functions that are Lebesgue integrable but not Riemann integrable, *e.g.* the indicator function on the rationals [3, p. 561]. And in fact $R_1[a, b]$ is not complete [3, p. 564].

2.12

Complete subsets

Definition. A subset S of a normed space is **complete** iff every Cauchy sequence from the subset converges to a limit within S.

Example. Any finite set is complete. (The next theorems give more interesting cases.)

Theorem. *In a normed space, any complete subset is closed.*

Proof. If $P \subseteq \mathcal{X}$ is complete, then every Cauchy sequence in P has a limit in P.
Thus all convergent sequences (which are of course Cauchy) have limits in P. Thus P is closed. □

Theorem. *In a Banach space, a subset is complete if and only if it is closed.*

Proof. The "only if" direction follows from the preceding theorem.
Suppose P is closed and $\{\boldsymbol{x}_n\}$ is a Cauchy sequence in P (and hence in $\mathcal{X}$). Since $\mathcal{X}$ is Banach, $\exists \boldsymbol{x} \in \mathcal{X}$ such that $\boldsymbol{x}_n \to \boldsymbol{x}$.
Hence $\boldsymbol{x}$ is a limit point of P. Since P is closed, $\boldsymbol{x} \in P$. Thus P is complete. □

Exercise. In a normed space, intersections of arbitrarily many complete subsets are complete.

What about unions of complete subsets? **??** **??**

There is an asymmetry between the two preceding theorems. We assumed a normed space to show complete set $\Longrightarrow$ closed set, but we assumed a Banach space to show closed set $\Longrightarrow$ complete set.

Is the stronger assumption (of a Banach space) truly necessary to show closed set $\Longrightarrow$ complete set? **??**

Ok, but in some (incomplete) $\mathcal{X}$ is there a closed *proper* subset that is incomplete? **??**

In the preceding theorem we assumed we are already working in a Banach space. What if we only have an ordinary normed space? Are there complete subsets of it? The next theorem shows that the answer can be yes, at least for finite-dimensional subspaces.

Recall that if M is a subspace of a normed space $\mathcal{X}$, then $d(\alpha \boldsymbol{x}, M) = |\alpha|\, d(\boldsymbol{x}, M)$ for all $\boldsymbol{x} \in \mathcal{X}$ and $\alpha \in \mathcal{F}$.

Theorem. *Any finite-dimensional subspace of a normed space is complete (and hence closed).*

Proof. By induction on the dimension of the subspace.

For a 1D subspace: $M = \{\alpha \boldsymbol{e} : \alpha \in \mathcal{F}\}$, where $\boldsymbol{e} \in \mathcal{X}$ is a fixed basis vector.
Any Cauchy sequence $\{\boldsymbol{x}_n\}$ in this subspace has elements of the form $\boldsymbol{x}_n = \alpha_n \boldsymbol{e}$. Since $\|\boldsymbol{x}_n - \boldsymbol{x}_m\| = |\alpha_n - \alpha_m|\, \|\boldsymbol{e}\|$, the sequence of reals $\{\alpha_n\}$ is also Cauchy and hence convergent (to some limit α) since $\mathbb{E}^1 = (\mathbb{R}, |\cdot|)$ is complete.
Thus, $\|\boldsymbol{x}_n - \alpha \boldsymbol{e}\| = |\alpha_n - \alpha|\, \|\boldsymbol{e}\| \to 0$, so $\boldsymbol{x}_n \to \boldsymbol{x} = \alpha \boldsymbol{e} \in M$. Hence, any 1D subspace of a normed space is complete.

Now assume the theorem is true for subspaces of dimension $N-1$.
Suppose M is a N-dimensional subspace of a normed space $\mathcal{X}$. We must show that M is complete.

Let $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_N$ denote a basis for M. For $k = 1, \ldots, N$, define $M_k = [\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{k-1}, \boldsymbol{e}_{k+1}, \ldots, \boldsymbol{e}_N\}]$ and $\delta_k = d(\boldsymbol{e}_k, M_k)$.

Claim: $\delta_k > 0$.
Suppose $\delta_k = 0$. Then since M_k is a $N-1$ dimensional space it is complete by assumption and hence closed, so $d(\boldsymbol{e}_k, M_k) = 0$ would imply $\boldsymbol{e}_k \in M_k$, contradicting the linear independence of $\{\boldsymbol{e}_k\}$.

Suppose $\{\boldsymbol{x}_n\} \in M$ is Cauchy. Each $\boldsymbol{x}_n$ has a (unique) representation $\boldsymbol{x}_n = \sum_{k=1}^N \lambda_k^n \boldsymbol{e}_k$, so for each $k \in \{1, \ldots, N\}$:

$$\begin{aligned} \|\boldsymbol{x}_n - \boldsymbol{x}_m\| &= \left\| \sum_{k=1}^N (\lambda_k^n - \lambda_k^m) \boldsymbol{e}_k \right\| = \Bigg\| (\lambda_k^n - \lambda_k^m) \boldsymbol{e}_k - \underbrace{\sum_{j \neq k}^N (\lambda_j^m - \lambda_j^n) \boldsymbol{e}_j}_{\in M_k} \Bigg\| \geq d((\lambda_k^n - \lambda_k^m) \boldsymbol{e}_k, M_k) \\ &= |\lambda_k^n - \lambda_k^m|\, d(\boldsymbol{e}_k, M_k) = |\lambda_k^n - \lambda_k^m|\, \delta_k. \end{aligned}$$

Since $\|\boldsymbol{x}_n - \boldsymbol{x}_m\| \to 0$ for a Cauchy sequence, we see that $|\lambda_k^n - \lambda_k^m| \to 0$ since $\delta_k > 0$. Thus $\{\lambda_k^n\}$ is Cauchy and hence (by the completeness of $\mathbb{R}$) converges to some limit λ_k. Defining $\boldsymbol{x} = \sum_{k=1}^N \lambda_k \boldsymbol{e}_k \in M$ we find that $\boldsymbol{x}_n \to \boldsymbol{x}$ since

$$\|\boldsymbol{x}_n - \boldsymbol{x}\| = \left\| \sum_{k=1}^N (\lambda_k^n - \lambda_k) \boldsymbol{e}_k \right\| \leq \sum_{k=1}^N |\lambda_k^n - \lambda_k|\, \|\boldsymbol{e}_k\| \to 0.$$

Thus any such Cauchy sequence in M converges to a limit in M, so M is complete. □

Note this proof's immediate use of the basis to represent any $\boldsymbol{x} \in M$. Proofs about finite-dimensional spaces often start this way.

Corollary.
- Any finite-dimensional subspace of a normed space is closed.
- Any finite-dimensional normed space is complete (and closed).

Exercise. [3, p. 218]. A (Hamel) basis for a Banach space is either finite or uncountably infinite.

2.9

Transformations

In systems theory, we analyze many systems and mathematical operations that *transform* one signal into another.

Definition. Let $\mathcal{X}$ and $\mathcal{Y}$ be two vector spaces (over a common field $\mathcal{F}$), and let D be a subset of $\mathcal{X}$.
A rule "T" that assigns a single element $\boldsymbol{y} \in \mathcal{Y}$ to each $\boldsymbol{x} \in D$ is called a **transformation** from $\mathcal{X}$ to $\mathcal{Y}$ with **domain** D.
We write $\boldsymbol{y} = T(\boldsymbol{x})$ and $T : \mathcal{X} \to \mathcal{Y}$. or $T : D \to \mathcal{Y}$.

See *p. 27* or c1 review notes for **one-to-one** and **onto**.

Definition. A transformation from a vector space $\mathcal{X}$ (over a scalar field $\mathcal{F}$) into $\mathcal{F}$ is called a **functional** on $\mathcal{X}$.

Example. A norm is a functional.

Linear transformations

Definition. A transformation $T : \mathcal{X} \to \mathcal{Y}$ (where $\mathcal{X}$ and $\mathcal{Y}$ are vector spaces over a common field) is called **linear** iff

$$T(\alpha \boldsymbol{x}_1 + \beta \boldsymbol{x}_2) = \alpha\, T(\boldsymbol{x}_1) + \beta\, T(\boldsymbol{x}_2),\ \forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{X},\ \forall \alpha, \beta \in \mathcal{F}.$$

Definition. If T is a linear transformation from $\mathcal{X}$ into $\mathcal{X}$ itself, then we say T is a **linear operator**.

However, the terminology distinguishing linear transformations from linear operators is not universal, and the two terms are often used interchangeably.

Example. Let $\mathcal{F} = \mathbb{R}$ and let $\mathcal{X} = \mathcal{Y}$ be the space of continuous functions on $[0, 1]$.
Define the linear transformation T by: $F = T(f)$ iff $F(t) = \int_0^t f(\tau)\, d\tau$.
Integration (with suitable limits) is a linear transformation.

Simple fact for linear transformations:
- $T(\mathbf{0}) = \mathbf{0}$. *Proof.* $T(\mathbf{0}) = T(0\mathbf{0}) = 0\, T(\mathbf{0}) = \mathbf{0}$. This is called the "zero in, zero out" property in linear systems theory.

Caution! (From Linear Systems by T. Kailath.) By induction it follows that $T(\sum_{i=1}^n \alpha_i \boldsymbol{x}_i) = \sum_{i=1}^n \alpha_i\, T(\boldsymbol{x}_i)$ for any finite n, but the above *does not* imply in general that linearity holds for infinite summations or integrals.
Further assumptions about "smoothness" or "regularity" or "continuity" of T are needed for that.

This point is always glossed over in introductory signals and systems courses, where infinite sums (and integrals) are routinely "passed through" linear systems according to the superposition property, with no attempt to verify the validity of such exchanges.

Continuity

To define such continuity, we restrict attention to normed spaces.
Note that a transformation can be continuous for one pair of normed spaces but discontinuous for another pair, *e.g.* [3, p. 63,65].

Definition. A transformation T from a *normed space* $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ into a *normed space* $\left(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}}\right)$ is called **continuous** at $\boldsymbol{x}_0 \in \mathcal{X}$ iff

$$\forall \varepsilon > 0,\ \exists \delta = \delta(\boldsymbol{x}_0, \varepsilon) > 0 \text{ s.t. } \|\boldsymbol{x} - \boldsymbol{x}_0\|_{\mathcal{X}} < \delta \Longrightarrow \|T(\boldsymbol{x}) - T(\boldsymbol{x}_0)\|_{\mathcal{Y}} < \varepsilon.$$

Definition. If T is continuous at all $\boldsymbol{x}_0 \in \mathcal{X}$, then we simply call T **continuous**.

Example. $f : \mathcal{X} \to \mathbb{R}$ defined by $f(\boldsymbol{x}) = d(\boldsymbol{x}, S)$ is continuous for any set S in a normed space $\mathcal{X}$, since $|d(\boldsymbol{x}, S) - d(\boldsymbol{y}, S)| \leq \|\boldsymbol{x} - \boldsymbol{y}\|$.

Definition. A transformation T from a *normed space* $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ into a *normed space* $\left(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}}\right)$ is called **uniformly continuous** iff

$$\forall \varepsilon > 0,\ \exists \delta = \delta(\varepsilon) > 0 \text{ s.t. } \|\boldsymbol{x} - \boldsymbol{z}\|_{\mathcal{X}} < \delta \Longrightarrow \|T(\boldsymbol{x}) - T(\boldsymbol{z})\|_{\mathcal{Y}} < \varepsilon.$$

Here δ depends only on ε. Sometimes $\delta(\varepsilon)$ is called the **modulus of continuity** of T.

Example. Is $\|\cdot\|$ uniformly continuous? **??**

We often want to exchange transformations and limits; the following proposition shows that continuity is the key condition.

Proposition. A transformation T from a *normed space* $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ into a *normed space* $\left(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}}\right)$ is **continuous** at $\boldsymbol{x} \in \mathcal{X}$ iff

$$\boldsymbol{x}_n \to \boldsymbol{x} \Longrightarrow T(\boldsymbol{x}_n) \to T(\boldsymbol{x}) \text{ (for any such sequence } \{\boldsymbol{x}_n\}),\ \textit{i.e.}, \lim_{n\to\infty} T(\boldsymbol{x}_n) = T\left(\lim_{n\to\infty} \boldsymbol{x}_n\right).$$

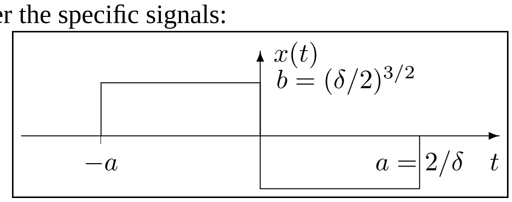
Proof. ***see text***

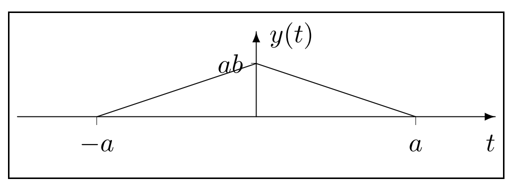
Exercise. A linear transformation from a finite-dimensional normed space to any normed space is continuous.

Thus any "counter-example" showing a linear transformation that is not continuous will be infinite dimensional, *e.g.*, the integrator system example [3, p.63].

Example. Consider $\mathcal{X} = \mathcal{Y} = \mathcal{L}_2[\mathbb{R}]$ and the linear operator T corresponding to the integral transformation: $y(t) = \int_{-\infty}^{t} x(\tau)\,\mathrm{d}\tau$.
We show that T is not continuous at $\boldsymbol{x}_0 = \boldsymbol{0}$. (And in fact is discontinuous everywhere!)

Consider the specific signals:





Then $\|\boldsymbol{x} - \boldsymbol{0}\| = \sqrt{\int |x(t)|^2 \,\mathrm{d}t} = \sqrt{2ab^2} = \delta/\sqrt{2} < \delta$ but $\|T(\boldsymbol{x}) - T(\boldsymbol{0})\| = \|\boldsymbol{y} - \boldsymbol{0}\| = \sqrt{\frac{2}{3}b^2a^3} = \sqrt{2/3}$.

An example of a continuous (linear) operator in an infinite dimensional vector space is the discrete-time convolution operator.

Example. Consider $\mathcal{X} = \mathcal{Y} = \ell_\infty$ and the transformation $\boldsymbol{y} = T(\boldsymbol{x}) \iff y_n = \sum_{k=-\infty}^{\infty} h_{n-k} x_k$, where $h \in \ell_1$.
Note that ℓ_1 corresponds to BIBO stability!
Exercise. Show that $\|T(\boldsymbol{x})\|_\infty \leq \|h\|_1 \|\boldsymbol{x}\|_\infty$. (This also ensures that $T(\boldsymbol{x})$ is well defined.)
Thus, since T is linear, if $\boldsymbol{x}, \boldsymbol{z} \in \ell_\infty$ then $\|T(\boldsymbol{x}) - T(\boldsymbol{z})\|_\infty = \|T(\boldsymbol{x} - \boldsymbol{z})\|_\infty \leq \|h\|_1 \|\boldsymbol{x} - \boldsymbol{z}\|_\infty$, so T is (uniformly) continuous, with $\delta(\varepsilon) = \varepsilon / \|h\|_1$. So BIBO LTI systems are uniformly continuous, and thus for such systems we can freely exchange limits and sums. (And in fact we would need to make such exchanges to derive rigorously convolution properties like the commutative property.)
We will talk much more about such **bounded linear operators** in *Ch. 6*.

2.13

Compactness

Optimization is about maximizing a functional over some set, or more precisely, usually about finding the maximizer (within some set) of a functional. When does a functional achieve its maximum? We wish that the answer were "if the set is closed and bounded," but unfortunately that is incorrect in general. The concept of a **compact set** helps answer this question in general.

Definition. A subset K of a normed space $(\mathcal{X}, \|\cdot\|)$ is called **compact** or **sequentially compact** iff every sequence $\{\boldsymbol{x}_n\}$ in K has some subsequence $\{\boldsymbol{x}_{n_i}\}$ that converges to a limit $\boldsymbol{x} \in K$.

(This is not quite the definition of compactness typically used first in a real analysis course, but it is more convenient for our purposes. One can first define a compact set in terms of set coverings, and then prove that a metric space is compact if and only if it is sequentially compact, *e.g.*, [2, p. 63].)

As usual, having defined a new concept, we now attempt to relate it to previously defined concepts.

Can a cone (other than $\{\mathbf{0}\}$) be compact? **??**

Exercise. In a normed space, arbitrary intersections of compact subsets are compact.

Exercise. What about unions? **??** **??**

Lemma. If $\boldsymbol{x}_n \to \boldsymbol{x}$ and $\{\boldsymbol{x}_{n_i}\}$ is a subsequence of $\{\boldsymbol{x}_n\}$, then $\lim_{i\to\infty} \boldsymbol{x}_{n_i} = \boldsymbol{x}$.
Proof. Pick any $\varepsilon > 0$. $\boldsymbol{x}_n \to \boldsymbol{x} \implies \exists N$ s.t. $n > N \implies \|\boldsymbol{x}_n - \boldsymbol{x}\| < \varepsilon$.
Choose I such that $i > I \implies n_i > N$. Then $i > I \implies \|\boldsymbol{x}_{n_i} - \boldsymbol{x}\| < \varepsilon$. □

Proposition. A compact subset K of a normed space $(\mathcal{X}, \|\cdot\|)$ is complete, closed, (and bounded).

Proof.
- Claim: K is complete, because Cauchy sequences will have limits in a compact set K.
  Let $\{\boldsymbol{x}_n\}$ be any Cauchy in K, then $\forall \varepsilon > 0,\ \exists M$ s.t. $n, m > M \implies \|\boldsymbol{x}_n - \boldsymbol{x}_m\| < \varepsilon/2$.
  Since K is compact, $\exists \{n_i\}$ s.t. $\boldsymbol{x}_{n_i} \to \boldsymbol{x} \in K$, *i.e.*, $\forall \varepsilon > 0, \exists I$ s.t. $i > I \implies \|\boldsymbol{x}_{n_i} - \boldsymbol{x}\| < \varepsilon/2$.
  For any $\varepsilon > 0$, let $N = \max\{M, n_I\}$.
  For $n > N$ and $i > I$ we have $\|\boldsymbol{x}_n - \boldsymbol{x}\| \le \|\boldsymbol{x}_n - \boldsymbol{x}_{n_i}\| + \|\boldsymbol{x}_{n_i} - \boldsymbol{x}\| < \varepsilon/2 + \varepsilon/2 = \varepsilon$. Thus $\boldsymbol{x}_n \to \boldsymbol{x}$.
- To show K is closed, we show that $\boldsymbol{x}_n \to \boldsymbol{x} \implies \boldsymbol{x} \in K$ for $\{\boldsymbol{x}_n\} \in K$.
  $\{\boldsymbol{x}_n\} \in K \implies \exists \boldsymbol{x}_{n_i} \to \boldsymbol{y} \in K$. But $\boldsymbol{x}_n \to \boldsymbol{x} \implies \boldsymbol{x}_{n_i} \to \boldsymbol{x}$, so $\boldsymbol{x} = \boldsymbol{y} \in K$. Thus K is closed.
- Before we show boundedness, we need to prove the Weierstrass theorem! □

So is the reverse true? Are closed and bounded sets compact?
The following theorem shows that in general the answer is "yes" in finite-dimensional normed spaces. (But not in general.)

Theorem. *In a normed space $(\mathcal{X}, \|\cdot\|)$, the following are equivalent.*
(a) $\mathcal{X}$ is finite dimensional.
(b) Every closed and bounded subset is compact.
(c) The closed unit ball $B_1(\mathbf{0}) = \{\boldsymbol{x} \in \mathcal{X} : \|\boldsymbol{x}\| \le 1\}$ is compact [3, p. 269].

Proof.
(a) $\implies$ (b) is the Heine-Borel theorem (see Math 451...)
(b) $\implies$ (c) is obvious
(c) $\implies$ (a) will be a homework problem

This theorem illustrates why infinite-dimensional spaces are more challenging than finite-dimensional spaces. The fact that every closed and bounded subset is compact can be very useful in finite dimensional problems, and we do not have this tool in general infinite-dimensional cases.

By having a notion of compactness, we can make statements (such as the Weierstrass theorem below) that apply to both infinite and finite-dimensional spaces, whereas those statements would not hold if we merely assumed that the sets were closed and bounded.

Analogy: we write sup when we think max. Here, we write compact when we think "closed and bounded."

Summary

In any normed space:
- compact $\Longrightarrow$ complete $\Longrightarrow$ closed
- compact $\Longrightarrow$ closed and bounded

In a Banach space: closed $\Longrightarrow$ complete
In a finite-dimensional normed space: closed and bounded $\Longrightarrow$ compact

Examples

As has been / will be shown, any compact set is closed, bounded, and complete, and that the reverse is true in finite-dimensional spaces. But what about infinite-dimensional spaces?

Example.

Consider any of the Banach spaces ℓ_p, for $p \in [1, \infty]$ and the subset $B = \{\boldsymbol{x} \in \mathcal{X} : \|\boldsymbol{x}\| \leq 1\}$. This is a closed set, and hence it is complete (since $\mathcal{X}$ is complete). Furthermore, B is clearly bounded.

We now proceed to exhibit a sequence $\{\boldsymbol{x}_n\}$ in B that has no convergent subsequences.

Let $\boldsymbol{x}_n$ denote the sequence in ℓ_p whose nth element is unity and all other elements are zero. Clearly $\{\boldsymbol{x}_n\} \in B$.
Yet $\|\boldsymbol{x}_n - \boldsymbol{x}_m\|_p = 1$, so it is impossible for this $\{\boldsymbol{x}_n\}$ to have any convergent subsequences since any such convergent subsequence would need to be Cauchy which cannot occur when $\|\boldsymbol{x}_n - \boldsymbol{x}_m\|_p = 1, \forall n, m \in \mathbb{N}$.

So the set B in ℓ_p is closed, bounded, and complete, but not compact.

One might begin to wonder then: are there any compact sets in infinite-dimensional spaces?

Example. Any finite set in an infinite-dimensional space is compact.

But that is not very interesting since the set $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ is just a closed and bounded subset of the finite-dimensional subspace $[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n]$, so of course it is compact.

Example. In any normed space $\mathcal{X}$, suppose $\{\boldsymbol{x}_n\}$ converges to $\boldsymbol{x} \in \mathcal{X}$. Then the set $\{\boldsymbol{x}\} \cup \cup_n \{\boldsymbol{x}_n\}$ is compact.

This still seems like a rather contrived and limited compact set.

Are there any interesting (*e.g.*, nonempty) *convex* compact sets in infinite-dimensional spaces?

Conjecture: the following (convex) set is (?) compact:

$$\left\{\boldsymbol{x} \in \ell_1 : \sum_{i=1}^{\infty} 2^i |x_i| \leq 1\right\}. \tag{2-1}$$

Note: the following set is *not* compact

$$\left\{\boldsymbol{x} \in \ell_1 : \|\boldsymbol{x}\|_1 = \sum_{i=1}^{\infty} |x_i| \leq 1\right\},$$

as shown above, so the question is whether adding the "2^i" part is enough of a change to make a compact set.

Extra credit (=30 homework points) to anyone who can show that (2-1) is or is not compact, or who can give a different example of a nontrivial convex, compact set in an infinite dimensional normed space.

Compactness, closedness, and distances

Lemma 2.4 *In a normed space, closed subsets of compact subsets are themselves compact.*

??

Lemma 2.5 *In a normed space, any compact set contains the closures of all of its subsets.*

??

Lemma 2.6 *If U and V are compact, disjoint subsets of a normed space, then $d(U, V) > 0$.* ***(Picture)***

Proof. Suppose $d(U, V) = 0$. Then there exists $\{\boldsymbol{x}_n\} \in U$ and $\{\boldsymbol{y}_n\} \in V$ such that $\|\boldsymbol{x}_n - \boldsymbol{y}_n\| \to 0$.
Since U is compact, there is a subsequence $\{\boldsymbol{x}_{n_i}\}$ that converges to some $\boldsymbol{x} \in U$.
Now $\|\boldsymbol{y}_{n_i} - \boldsymbol{x}\| \leq \|\boldsymbol{y}_{n_i} - \boldsymbol{x}_{n_i}\| + \|\boldsymbol{x}_{n_i} - \boldsymbol{x}\| \to 0$ as $i \to \infty$, so $\boldsymbol{y}_{n_i} \to \boldsymbol{x} \in U$.
But since V is compact, it is also closed, so the limit of $\{\boldsymbol{y}_{n_i}\}$ must lie in V, contradicting the disjointedness of U and V. □

. . . .

Would it suffice for U and V to be closed?

No. Consider the sets $U = \{(x, y) \in \mathbb{R}^2 : y > 0,\ x \geq 1/y\}$ and $V = \{(x, y) \in \mathbb{R}^2 : y < 0,\ x \leq -1/y\}$. ***(Picture)***
These sets are closed and disjoint, yet $d(U, V) = 0$.

Can you find a 1D example? **??**

. . . .

Would it suffice for U and V to be *closed* and *bounded*?

(We need not look in finite-dimensional normed spaces for counter-examples since there all closed and bounded sets are compact.)

For a counter-example, we turn to the (incomplete) normed space $\left\{\mathcal{X} = \{\text{finite-length real sequences}\}, \|\cdot\|_p\right\}$, where $\|\cdot\|_p$ is the usual ℓ_p norm. Define $U = \cup_{n=1}^{\infty} \boldsymbol{x}_n$ and $V = \cup_{n=1}^{\infty} \boldsymbol{y}_n$, where $\boldsymbol{x}_n = (1, 1/2, \ldots, 1/2^n, 0, 0, \ldots)$ and $\boldsymbol{y}_n = \boldsymbol{x}_n + (1/n, 0, 0, \ldots)$. Then U and V are disjoint, closed (in this incomplete $\mathcal{X}$), and bounded (for any $p \in [1, \infty]$) yet $d(U, V) = 0$.

So it seems that we do need a stronger condition than "closed and bounded" in Lemma 2.6.

. . . .

Would "complete and bounded" suffice? Or do we need compactness?

Upper semicontinuous functions

Definition. A (real) functional f defined on a normed space $(\mathcal{X}, \|\cdot\|)$ is called **upper semicontinuous** at $\boldsymbol{x} \in \mathcal{X}$ iff

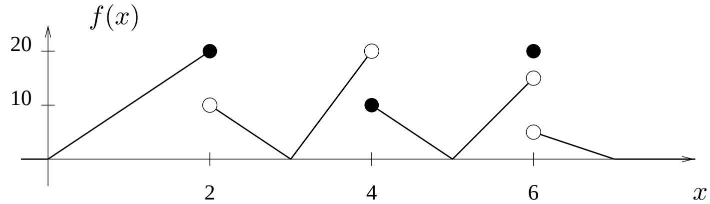
$$\forall \varepsilon > 0,\ \exists \delta > 0 \text{ s.t. } \forall \boldsymbol{y} \in \mathcal{X},\ \|\boldsymbol{y} - \boldsymbol{x}\| < \delta \Longrightarrow f(\boldsymbol{y}) < f(\boldsymbol{x}) + \varepsilon.$$

We call f **lower semicontinuous** if at $\boldsymbol{x}$ iff $-f$ is upper semicontinuous at $\boldsymbol{x}$.
Fact. f is **continuous** iff f is both upper and lower semicontinuous.

Equivalently, one can show that f is u.s.c. at $\boldsymbol{x}$ iff $f(\boldsymbol{x}) \geq \limsup_{\boldsymbol{y} \to \boldsymbol{x}} f(\boldsymbol{y})$.

$S = \limsup_{\boldsymbol{y} \to \boldsymbol{x}} f(\boldsymbol{y})$ iff $\forall \varepsilon > 0,\ \exists \delta > 0$ s.t. $\|\boldsymbol{y} - \boldsymbol{x}\| < \delta \Longrightarrow f(\boldsymbol{y}) < S + \varepsilon$, and $\forall \varepsilon > 0,\ \forall \delta > 0,\ \exists \boldsymbol{y}$ s.t. $\|\boldsymbol{y} - \boldsymbol{x}\| < \delta$ and $f(\boldsymbol{y}) > S - \varepsilon$.
Loosely speaking, it is like a least upper bound "in the limit."

Example. Consider the following "almost piecewise continuous" real-valued function.



What is $\limsup_{\boldsymbol{y} \to 2} f(\boldsymbol{y})$? We get the biggest limit approaching from the left, so $\limsup_{\boldsymbol{y} \to 2} f(\boldsymbol{y}) = 20$. So f is u.s.c. at $x = 2$.
Similarly, $\limsup_{\boldsymbol{y} \to 4} f(\boldsymbol{y}) = 20$, so f is not u.s.c. at $x = 4$. But $\limsup_{\boldsymbol{y} \to 6} f(\boldsymbol{y}) = 20$, so f is u.s.c. at $x = 6$.

This function is u.s.c. everywhere except **??** and lower semicontinuous everywhere except **??**

Does this function achieve a maximum on [3,5]? **??** What about on [1,3]? **??** How about on [0,1)? **??**

Theorem. *(Weierstrass)*
An upper semicontinuous (real) functional f on a compact subset K of a normed space $(\mathcal{X}, \|\cdot\|)$ (i) is bounded on K, and (ii) achieves a maximum on K.

Proof. Let $M = \sup_{\boldsymbol{x} \in K} f(x) = \begin{cases} \inf \{g \in \mathbb{R} : g \geq f(x),\ \forall \boldsymbol{x} \in K\}, & \text{if } f \text{ is bounded above on } K \\ \infty, & \text{otherwise.} \end{cases}$
(At this point we do not know if M is finite or not.)
By definition of supremum, $\exists$ a sequence $\{\boldsymbol{x}_n\} \in K$ such $f(\boldsymbol{x}_n) \to M$. (This is true even if $M = \infty$.)
However, the definition of supremum alone does not ensure that $\boldsymbol{x}_n$ converges.

Since K is compact, $\exists$ a convergent subsequence $\boldsymbol{x}_{n_i} \to \boldsymbol{x}_\star \in K$, for some $\boldsymbol{x}_\star$.
Since subsequences of convergent sequences have the same limit, $\lim_{i \to \infty} f(\boldsymbol{x}_{n_i}) = M$, considering $\{f(\boldsymbol{x}_n)\}$ as a sequence in $\mathbb{R}$.

Since f is upper semicontinuous, $f(\boldsymbol{x}_\star) \geq \limsup_{i \to \infty} f(\boldsymbol{x}_{n_i}) = M$.
Since $f(\boldsymbol{x})$ is real and hence finite, $f(\boldsymbol{x}_\star) \geq M$ implies M must be finite. $\therefore$ f is bounded on K.

On the other hand, by definition of supremum, since $\boldsymbol{x}_\star \in K$, $M \geq f(\boldsymbol{x}_\star)$.
So we conclude $f(\boldsymbol{x}_\star) = M$, meaning that $\boldsymbol{x}_\star$ achieves the maximum of f on K. □

Corollary. A real-valued, **continuous** functional f on a compact subset K of a normed space $(\mathcal{X}, \|\cdot\|)$ achieves its maximum and minimum on K.

Example. On a normed space $(\mathcal{X}, \|\cdot\|)$, the function $f : \mathcal{X} \to \mathbb{R}$ defined by $f(\boldsymbol{x}) = \|\boldsymbol{x}\|$ is continuous.
Proof. Recall $|\ \|\boldsymbol{x}\| - \|\boldsymbol{y}\|\ | \leq \|\boldsymbol{x} - \boldsymbol{y}\|$, so $\boldsymbol{x} \to \boldsymbol{y} \Longrightarrow \|\boldsymbol{x} - \boldsymbol{y}\| \to 0 \Longrightarrow |f(\boldsymbol{x}) - f(\boldsymbol{y})| \to 0$.

An inf. dim. example like splines would be nice here...

Corollary (to Weierstrass theorem)
A compact subset K of a normed space is (complete, closed, and) bounded.

Proof. Since $\|\cdot\|$ is continuous, by the Weierstrass theorem $\exists \boldsymbol{y} \in K$ s.t. $M \triangleq \|\boldsymbol{y}\| = \sup_{\boldsymbol{x} \in K} \|\boldsymbol{x}\|$.
Hence $\|\boldsymbol{x}\| \leq M < \infty$, $\forall \boldsymbol{x} \in K$, *i.e.*, K is bounded. □

2.14

Quotient Spaces

skip

2.15

Denseness

One last topological concept.

Definition. A subset D of a normed space $(\mathcal{X}, \|\cdot\|)$ is called **dense** iff any of the following equivalent conditions hold.
- $\forall \boldsymbol{x} \in \mathcal{X}$ and $\forall \varepsilon > 0,\ \exists \boldsymbol{y} \in D$ s.t. $\|\boldsymbol{x} - \boldsymbol{y}\| < \varepsilon,$ *i.e.,* $d(\boldsymbol{x}, D) = 0$.
- $\overline{D} = \mathcal{X}$.
- $\forall \boldsymbol{x} \in \mathcal{X},\ \exists \{\boldsymbol{y}_n\} \in D$ s.t. $\boldsymbol{y}_n \to \boldsymbol{x}$

Example. The set of rational numbers is dense in the real line.

Separability

Definition. A normed space $\mathcal{X}$ is called **separable** iff it contains a **countable** dense set, *i.e.,* $D = \bigcup_{n=1}^{\infty} \{\boldsymbol{y}_n\}$.

Example. Euclidean space $\mathbb{E}^n$ is separable. The collection of vectors $\boldsymbol{x} = (a_1, \ldots, a_n)$ with rational components is countable and dense in $\mathbb{E}^n$.

Example. ℓ_p is separable for $p \in [1, \infty)$
If $D_n = \{(r_1, \ldots, r_n, 0, 0, \ldots) : r_k \in \mathbb{Q}\}$, then $D = \cup_{n=1}^{\infty} D_n$ is dense in ℓ_p.
See text for proof.

Example. $\mathcal{L}_p$ is separable for $p \in [1, \infty)$
If $D_n = \{\sum_{k=1}^{n} r_k 1_{I_k}(t) : r_k \in \mathbb{Q},\ I_k = (a_k, b_k),\ a_k, b_k \in \mathbb{Q}\}$ for $n \in \mathbb{N}$, then $D = \cup_{n=1}^{\infty} D_n$ is dense in $\mathcal{L}_p$.
Alternatively, if $D_n = \{f : [a, b] \to \mathbb{R} : f$ is piecewise linear with $n \in \mathbb{N}$ breakpoints at $r_k \in \mathbb{Q}$ and $f(r_k) \in \mathbb{Q}\}$,
then $D = \cup_{n=1}^{\infty} D_n$ is dense in $\mathcal{L}_p$.

Example. $C[a, b]$ is separable. The set of all polynomials with rational coefficients is countable and dense in $C[a, b]$.
If $D_n = \{r_0 + r_1 t + \cdots + r_n t^n : r_k \in \mathbb{Q}\}$, then $D = \cup_{n=1}^{\infty} D_n$ is dense in $C[a, b]$.

Example. ℓ_∞ and $\mathcal{L}_\infty$ are not separable (Problem 2.21)

Schauder basis

Definition. In a normed space, $\{\boldsymbol{x}_n\} \in \mathcal{X}$ is a **Schauder basis** for $\mathcal{X}$ iff for each $\boldsymbol{x} \in \mathcal{X}$, there exists a unique sequence $\{\lambda_n\}$ such that $\boldsymbol{x} = \sum_{n=1}^{\infty} \lambda_n \boldsymbol{x}_n$ [5] [2, p. 98].

Theorem. *If a normed space has a Schauder basis, then it is separable [2, p. 100].*

What about the converse?
The famous Banach conjecture that every separable Banach space has a Schauder basis was shown to be incorrect by Elfon in 1973 [2, p. 100]. This could be considered surprising since separable, complete, normed spaces should be about as "nice as they come."

. . . .

So is the concept of separability of limited use to use? No, thanks the following key result.

Fact. A Hilbert space has a countable orthonormal basis iff it is separable [3, p. 314].
(Any countable orthonormal basis is a Schauder basis.)

Summary

Geometrical concepts and their generalizations.

point	$\boldsymbol{x} \in \mathcal{X}$
line	$\{\alpha \boldsymbol{x} : \alpha \in \mathbb{R}\}$
plane	$\{\sum_i \alpha_i \boldsymbol{x}_i : \alpha_i \in \mathbb{R}\}$
cone	$\boldsymbol{x} \in C \Longrightarrow \alpha \boldsymbol{x} \in C$ for $\alpha \geq 0$
length	$\|\boldsymbol{x}\|$
sphere	$\{\boldsymbol{x} \in \mathcal{X} : \|\boldsymbol{x}\| < \varepsilon\}$
distance	$\|\boldsymbol{x} - \boldsymbol{y}\|$
coordinate system	basis

Hierarchy of spaces: vector space $\supset$ normed space $\supset$ Banach space

Some principal results

- A subset of a Banach space is complete iff it is closed.
- Any finite-dimensional subspace of a normed space is complete (and hence closed).
- A real-valued, continuous functional on a compact subset of a normed space achieves its maximum and minimum on that subset.
- Any closed and bounded subset of a finite dimensional normed space is compact.

Bibliography

[1] P. Enflo. A counterexample to the approximation problem in Banach spaces. *Acta Math*, 130:309–17, 1973.

[2] I. J. Maddox. *Elements of functional analysis*. Cambridge, 2 edition, 1988.

[3] A. W. Naylor and G. R. Sell. *Linear operator theory in engineering and science*. Springer-Verlag, New York, 2 edition, 1982.

[4] D. G. Luenberger. *Optimization by vector space methods*. Wiley, New York, 1969.

[5] J. Schauder. Zur theorie stetiger abbildungen in funktionenrumen. *Math. Zeitsch.*, 26:47–65, 1927.

[6] P. P. Vaidyanathan. Generalizations of the sampling theorem: Seven decades after Nyquist. *IEEE Tr. Circ. Sys. I, Fundamental theory and applications*, 48(9):1094–109, September 2001.

[7] F. Deutsch. The convexity of Chebyshev sets in Hilbert space. In A. Yanushauskas Th. M. Rassias, H. M. Srivastava, editor, *Topics in polynomials of one and several variables and their applications*, pages 143–50. World Sci. Publishing, River Edge, NJ, 1993.

[8] M. Jiang. On Johnson's example of a nonconvex Chebyshev set. *J. Approx. Theory*, 74(2):152–8, August 1993.

[9] P. L. Combettes and H. J. Trussell. Method of successive projections for finding a common point of sets in metric spaces. *J. Optim. Theory Appl.*, 67(3):487–507, December 1990.