Introduction to Functional Analysis

George Voutsadakis[1]

[1]Mathematics and Computer Science
Lake Superior State University

LSSU Math 500

Subsection 1

Vector Space

Vector Spaces

- Let K be a field that may be $\mathbb{R}$ or $\mathbb{C}$. The elements of K are called **scalars**.

Definition (Vector Space)

A **vector space** (or **linear space**) **over a field** K is a nonempty set X of elements $x, y, \ldots$ (called **vectors**) together with two algebraic operations. These operations are called **vector addition** and **multiplication of vectors by scalars**, that is, by elements of K.
Vector addition associates with every ordered pair (x, y) of vectors a vector $x + y$, called the **sum** of x and y, in such a way that the following properties hold:

- Vector addition is **commutative** and **associative**: for all vectors we have $x + y = y + x$ and $x + (y + z) = (x + y) + z$;
- There exists a vector 0, called the **zero vector**, and for every vector x, there exists a vector $-x$, such that for all vectors we have $x + 0 = x$ and $x + (-x) = 0$.

Definition of Vector Spaces (Cont'd)

Definition (Vector Space) (Cont'd)

Multiplication by scalars associates with every vector x and scalar α a vector αx (also written $x\alpha$), called the **product** of α and x, in such a way that for all vectors x, y and scalars α, β, we have:

- $\alpha(\beta x) = (\alpha\beta)x$ and $1x = x$;
- the **distributive laws**: $\alpha(x + y) = \alpha x + \alpha y$ and $(\alpha + \beta)x = \alpha x + \beta x$.

- Vector addition is a mapping $X \times X \to X$ whereas multiplication by scalars is a mapping $K \times X \to X$.
- K is called the **scalar field** (or **coefficient field**) of the vector space X, and X is called a **real vector space** if $K = \mathbb{R}$ and a **complex vector space** if $K = \mathbb{C}$.
- The use of 0 for the scalar 0 as well as for the zero vector should cause no confusion, in general. If desirable for clarity, we can denote the zero vector by θ.

Properties of Vector Spaces

- Let X be a vector space over a field K.
  Then, for all vectors x and scalars α,
  (a) $0x = \theta$;
  (b) $\alpha\theta = \theta$;
  (c) $(-1)x = -x$.

(a) We have

$$0x + 0x = (0+0)x = 0x \Rightarrow 0x = \theta.$$

(b) We have

$$\alpha x = \alpha(x+\theta) = \alpha x + \alpha\theta \Rightarrow \alpha\theta = \theta.$$

(c) We have

$$x + (-1)x = 1x + (-1)x = (1+(-1))x = 0x = \theta \Rightarrow (-1)x = -x.$$

Vector Spaces $\mathbb{R}^n$ and $\mathbb{C}^n$

- (**Space** $\mathbb{R}^n$) This is the Euclidean space with underlying set being the set of all n-tuples of real numbers, $x = (\xi_1, \ldots, \xi_n)$, $y = (\eta_1, \ldots, \eta_n)$, etc.

  This is a real vector space with the two algebraic operations defined in the usual fashion

$$\begin{aligned} x + y &= (\xi_1 + \eta_1, \ldots, \xi_n + \eta_n) \\ \alpha x &= (\alpha \xi_1, \ldots, \alpha \xi_n), \quad \alpha \in \mathbb{R}. \end{aligned}$$

- (**Space** $\mathbb{C}^n$) This space consists of all ordered n-tuples of complex numbers $x = (\xi_1, \ldots, \xi_n)$, $y = (\eta_1, \ldots, \eta_n)$, etc..

  It is a complex vector space with the algebraic operations defined as in the previous example, where now $\alpha \in \mathbb{C}$.

Spaces of Functions

- (**Space** $C[a,b]$) The points of this space are continuous real-valued functions on $[a,b]$. The set of all these functions forms a real vector space with the algebraic operations defined in the usual way:

$$\begin{aligned}(x+y)(t) &= x(t)+y(t)\\ (\alpha x)(t) &= \alpha x(t), \quad \alpha \in \mathbb{R}.\end{aligned}$$

  In fact, $x+y$ and αx are continuous real-valued functions defined on $[a,b]$ if x and y are such functions and α is real.
- Other important vector spaces of functions are:
  (a) the vector space $B(A)$;
  (b) the vector space of all differentiable functions on $\mathbb{R}$;
  (c) the vector space of all real-valued functions on $[a,b]$ which are integrable in some sense.

Sequence Spaces

- (**Space** ℓ^2) This space is a vector space with the algebraic operations defined as usual in connection with sequences:

$$\begin{array}{rcl}(\xi_1,\xi_2,\ldots)+(\eta_1,\eta_2,\ldots) & = & (\xi_1+\eta_1,\xi_2+\eta_2,\ldots)\\ \alpha(\xi_1,\xi_2,\ldots) & = & (\alpha\xi_1,\alpha\xi_2,\ldots).\end{array}$$

  In fact, $x=(\xi_j)\in\ell^2$ and $y=(\eta_j)\in\ell^2$ implies $x+y\in\ell^2$, as follows readily from the Minkowski inequality; also $\alpha x\in\ell^2$.
- Other vector spaces whose points are sequences are ℓ^∞, ℓ^p, where $1\le p<+\infty$, and s.

Subspaces

- A **subspace** of a vector space X is a nonempty subset Y of X, such that for all $y_1, y_2 \in Y$ and all scalars α, β, we have
$$\alpha y_1 + \beta y_2 \in Y.$$
- Hence, Y is itself a vector space, the two algebraic operations being those induced from X.
- A special subspace of X is the **improper subspace** $Y = X$.
- Another special subspace of any vector space X is $Y = \{0\}$.
- Every other subspace of X ($\neq X, \{0\}$) is called **proper**.

Spans

- A **linear combination** of vectors $x_1,\ldots,x_m$ of a vector space X is an expression of the form

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_m x_m,$$

  where the coefficients $\alpha_1,\ldots,\alpha_m$ are any scalars.
- For any nonempty subset $M \subseteq X$ the set of all linear combinations of vectors of M is called the **span** of M, written $\operatorname{span} M$.
- $\operatorname{span} M$ is a subspace of X, and we say that it is **spanned** or **generated** by M.

Linear Dependence and Independence

Definition (Linear Independence, Linear Dependence)

Let M be a given set of vectors $x_1,\ldots,x_r$, $r \geq 1$, in a vector space X and consider the equation $\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_r x_r = 0$, where $\alpha_1,\ldots,\alpha_r$ are scalars. Clearly, the equation holds for $\alpha_I = \alpha_2 = \cdots = \alpha_r = 0$.

- If this is the only r-tuple of scalars for which it holds, the set M is said to be **linearly independent**.
- M is said to be **linearly dependent** if M is not linearly independent, i.e., if the equation also holds for some r-tuple of scalars, not all zero.

An arbitrary subset M of X is said to be **linearly independent** if every nonempty finite subset of M is linearly independent.
M is said to be **linearly dependent** if M is not linearly independent.

A Consequence of Linear Dependence

Proposition

In a vector space X, $M = \{x_1, \ldots, x_r\}$ is linearly dependent if and only if at least one vector of M can be written as a linear combination of the others.

- Suppose x_i can be written as a linear combination of the other vectors, i.e., there exist scalars $\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_r$, such that

$$x_i = \alpha_1 x_1 + \cdots + \alpha_{i-1} x_{i-1} + \alpha_{i+1} x_{i+1} + \cdots + \alpha_r x_r.$$

Setting $\alpha_i = -1$, we get $\sum_{j=1}^r \alpha_j x_j = 0$, with not all α_j equal to 0. Hence, by definition, M is linearly dependent.

Conversely, if M is linearly dependent, there exist scalars a_j, $j = 1, \ldots, r$, not all 0, such that $\sum_{j=1}^r \alpha_j x_j = 0$. Suppose that $\alpha_i \neq 0$, for some $1 \leq i \leq r$. Then we have

$$x_i = -\frac{\alpha_1}{\alpha_i} x_1 - \cdots - \frac{\alpha_{i-1}}{\alpha_i} x_{i-1} - \frac{\alpha_{i+1}}{\alpha_i} x_{i+1} - \cdots - \frac{\alpha_r}{\alpha_i} x_r.$$

Dimension

Definition (Finite and Infinite Dimensional Vector Spaces)

A vector space X is said to be **finite dimensional** if there is a positive integer n such that X contains a linearly independent set of n vectors, whereas any set of $n+1$ or more vectors of X is linearly dependent.
n is called the **dimension** of X, written $n = \dim X$.
By definition, $X = \{0\}$ is **finite dimensional** and $\dim X = 0$.
If X is not finite dimensional, it is said to be **infinite dimensional**.

- In analysis, infinite dimensional vector spaces are of greater interest than finite dimensional ones:
  - $C[a,b]$ and ℓ^2 are infinite dimensional;
  - $\mathbb{R}^n$ and $\mathbb{C}^n$ are n-dimensional.

Bases, Canonical Bases and Hamel Bases

- If $\dim X = n$, a linearly independent n-tuple of vectors of X is called a **basis** for X (or a **basis** in X).

  If $\{e_1, \ldots, e_n\}$ is a basis for X, every $x \in X$ has a unique representation as a linear combination of the basis vectors $x = \alpha_1 e_1 + \cdots + \alpha_n e_n$.

  Example: A basis for $\mathbb{R}^n$ is $e_1 = (1, 0, 0, \ldots, 0)$, $e_2 = (0, 1, 0, \ldots, 0)$, ..., $e_n = (0, 0, 0, \ldots, 1)$.

  This is sometimes called the **canonical basis** for $\mathbb{R}^n$.

- If X is any vector space, not necessarily finite dimensional, and B is a linearly independent subset of X which spans X, then B is called a **basis** (or **Hamel basis**) for X.

  If B is a basis for X, then every nonzero $x \in X$ has a unique representation as a linear combination of (finitely many!) elements of B with nonzero scalars as coefficients.

Existence of Bases

Theorem

Every vector space $X \neq \{0\}$ has a basis.

- Let X be a vector space over some field K.
  Consider the collection P of all linearly independent subset of X ordered by inclusion

$$P = \{S \subseteq X : S \text{ is linearly independent}\}.$$

  To apply Zorn's lemma, suppose that $C = \{S_i\}_{i \in I}$ is a chain in P.
  We show that $M = \bigcup_{i \in I} S_i$ is an upper bound of C in P.
  That it is an upper bound of C is immediate by definition.
  So we check that M is a linearly independent subset of X.
  If not, there are vectors $s_1, \ldots, s_n$, where $s_k \in S_{i_k}$ for some S_{i_k}, and scalars $\alpha_1, \ldots, \alpha_n$ not all 0, such that $\alpha_1 s_1 + \cdots + \alpha_n s_n = 0$.
  As C is totally ordered, one of $S_{i_1}, \ldots, S_{i_n}$, say S, contains the others.

Existence of Bases (Cont'd)

- So every vector s_i belongs to some S in C.

  This says there is a non-trivial dependence relation among vectors in S, contradicting that S is linearly independent ($S \in C \subseteq P$).

  Therefore, M is a linearly independent subsets of X, i.e., $M \in P$.

  By Zorn's lemma, P contains a maximal element, say B.

  It suffices to show that B spans X.

  If not, then there is some $x \in X$, such that $x \notin \text{span}(B)$.

  This says that $B \cup \{x\}$ is a linearly independent subset of X.

  Since $B \subsetneqq B \cup \{x\}$, this contradicts the maximality of B.

Steinitz Exchange Lemma

Theorem (Steinitz Exchange Lemma)

Let Y and Z be finite subsets of a vector space X. If Y is linearly independent and Z spans X, then:

- $|Y| \leq |Z|$;
- There exists $Z' \subseteq Z$, with $|Z'| = |Z| - |Y|$, such that $Y \cup Z'$ spans X.

- Let $Y = \{y_1, \ldots, y_m\}$ and $Z = \{z_1, \ldots, z_n\}$.

  We show by induction on $k = 0, 1, \ldots, m$, that $k \leq n$ and $\{y_1, \ldots, y_k, z_{k+1}, \ldots, z_n\}$ spans X (here, the z_j may have been reordered and the reordering depends on k).

  For $k = 0$, there are no y_i and, by hypothesis, $\{z_1, \ldots, z_n\}$ spans X.

  Suppose that the conclusion holds for some $k < m$.

Steinitz Exchange Lemma (Cont'd)

- Since $\{y_1,\ldots,y_k,z_{k+1},\ldots,z_n\}$ spans X, there exist $\alpha_1,\ldots,\alpha_n$, such that $y_{k+1}=\sum_{j=1}^k \alpha_j y_j+\sum_{k+1}^n \alpha_j z_j$.

  As $\{y_1,\ldots,y_{k+1}\}$ are linearly independent, one of $\{\alpha_{k+1},\ldots,\alpha_n\}$ must be nonzero.

  This implies $k+1\le n$.

  By reordering, assume $\alpha_{k+1}\neq 0$.

  Then, we get $z_{k+1}=\frac{1}{\alpha_{k+1}}(y_{k+1}-\sum_{j=1}^k \alpha_j y_j-\sum_{j=k+2}^n \alpha_j z_j)$.

  Hence, z_{k+1} is in the span of $\{y_1,\ldots,y_{k+1},z_{k+2},\ldots,z_n\}$.

  Thus, by the induction hypothesis, since the latter span includes z_{k+1}, it must be X.

  Therefore, the span of $\{y_1,\ldots,y_{k+1},z_{k+2},\ldots,z_n\}$ is also X.

Dimensions

Theorem

All bases for a given (finite or infinite dimensional) vector space X have the same cardinal number, called the **dimension** of X.

- Suppose $\{x_i : i \in I\}$ is linearly independent and that $\{y_j : j \in J\}$ is a generating set.

  If $|J|$ is finite, the conclusion follows by Steinitz's Lemma.

  Suppose $|J|$ is infinite. As the case where $|I|$ is finite leaves nothing to prove, assume that $|I|$ is also infinite.

  Towards obtaining a contradiction, suppose $|I| > |J|$.

  By Zorn's Lemma, every linearly independent set is contained in a maximal one, which is necessarily a basis.

  So, without loss of generality, we may assume that $\{x_i : i \in I\}$ is a basis.

Dimensions (Cont'd)

- Thus, for every $j \in J$, there exist scalars α_{ij}, such that $y_j = \sum_{i \in E_j} \alpha_{ij} x_i$, where E_j is a finite subset of I.

  Since J is infinite, $|\bigcup_{j \in J} E_j| = |J|$.

  Hence, by hypothesis, $|\bigcup_{j \in J} E_j| < |I|$.

  We conclude that there exists $i_0 \in I$, $i_0 \notin \bigcup_{j \in J} E_j$.

  As $\{y_j : j \in J\}$ is a generating set, x_{i_0} may be expressed as a finite linear combination of y_j's.

  By what was shown above, each y_j may, in turn, be expressed as a finite linear combination of x_i's other that x_{i_0}.

  Therefore, x_{i_0} is linearly dependent on the other x_i's, which contradicts the hypothesis.

Dimension of Subspaces

Theorem (Dimension of a Subspace)

Let X be an n-dimensional vector space. Then any proper subspace Y of X has dimension less than n.

- If $n = 0$, then $X = \{0\}$ and has no proper subspace.

  If $\dim Y = 0$, then $Y = \{0\}$, and $X \neq Y$ implies $\dim X \geq 1$. Clearly, $\dim Y \leq \dim X = n$. If $\dim Y$ were n, then Y would have a basis of n elements, which would also be a basis for X since $\dim X = n$. So $X = Y$. This shows that any linearly independent set of vectors in Y must have fewer than n elements. Hence $\dim Y < n$.

Subsection 2

Normed Space, Banach Space

Norms and Induced Metrics

Definition (Norms and Induced Metrics)

A **norm** on a (real or complex) vector space X is a real-valued function on X whose value at an $x \in X$ is denoted by $\|x\|$ and which has the properties, for all $x, y \in X$ and α a scalar:

(N1) $\|x\| \geq 0$;

(N2) $\|x\| = 0$ if and only if $x = 0$;

(N3) $\|\alpha x\| = |\alpha| \|x\|$;

(N4) $\|x + y\| \leq \|x\| + \|y\|$ (**Triangle Inequality**);

A norm on X defines a metric d on X which is given, for all $x, y \in X$, by

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

It is called the **metric induced by the norm**.

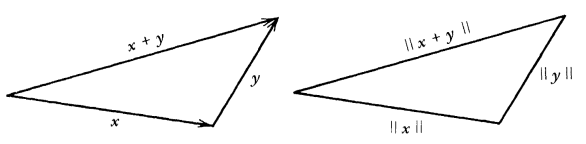
Normed Spaces and Banach Spaces

Definition (Normed Space, Banach Space)

A **normed space** X is a vector space with a norm $\|\cdot\|$ defined on it.
A **Banach space** is a complete normed space (complete in the metric defined by the norm).
The normed space is denoted by $(X, \|\cdot\|)$ or simply by X.

Remarks on the Properties

- The defining properties (N1) to (N4) of a norm are suggested and motivated by the length $\|x\|$ of a vector x in elementary vector algebra.
  In this case we can write $\|x\| = |x|$.
- In fact, (N1) and (N2) state that all vectors have positive lengths except the zero vector which has length zero.
- (N3) means that when a vector is multiplied by a scalar, its length is multiplied by the absolute value of the scalar.
- (N4) means that the length of one side of a triangle cannot exceed the sum of the lengths of the two other sides.



- From (N1) to (N4), we see that $d(x,y)$ does define a metric.
  Hence, normed spaces and Banach spaces are metric spaces.

Continuity of the Norm

- Note that, by (N3), $\|x-y\| = \|y-x\|$.
- By (N4), we get:
  - $\|y\| - \|x\| \le \|y-x\|$;
  - $-\|y-x\| \le \|y\| - \|x\|$.

  These give $-\|y-x\| \le \|y\| - \|x\| \le \|y-x\|$.

  or, equivalently,
  $$|\|y\| - \|x\|| \le \|y-x\|.$$
- This formula implies the **continuity property** of the norm:

  The norm is continuous, i.e., $x \mapsto \|x\|$ is a continuous mapping of $(X, \|\cdot\|)$ into $\mathbb{R}$.

Examples

- (**Euclidean Space** $\mathbb{R}^n$ and **Unitary Space** $\mathbb{C}^n$) These spaces are Banach spaces with norm defined by

$$\|x\| = \left(\sum_{j=1}^{n} |\xi_j|^2\right)^{1/2} = \sqrt{|\xi_1|^2 + \cdots + |\xi_n|^2}.$$

In fact, $\mathbb{R}^n$ and $\mathbb{C}^n$ are complete, and the norm yields the metric encountered before:

$$d(x,y) = \|x - y\| = \sqrt{|\xi_1 - \eta_1|^2 + \cdots + |\xi_n - \eta_n|^2}.$$

- Note, in particular, that in $\mathbb{R}^3$ we have

$$\|x\| = |x| = \sqrt{\xi_1^2 + \xi_2^2 + \xi_3^2}.$$

Thus, the norm does indeed generalize the elementary notion of the length $|x|$ of a vector.

The Spaces ℓ^p and ℓ^∞

- (**Space** ℓ^p) This space is a Banach space with norm given by

$$\|x\| = \left(\sum_{j=1}^{\infty} |\xi_j|^p\right)^{1/p}.$$

In fact, this norm induces the metric

$$d(x,y) = \|x-y\| = \left(\sum_{j=1}^{\infty} |\xi_j - \eta_j|^p\right)^{1/p}.$$

Completeness has been proved.

- (**Space** ℓ^∞) This space is a Banach space since its metric is obtained from the norm defined by

$$\|x\| = \sum_j |\xi_j|.$$

Completeness has also been shown.

Space $C[a,b]$

- (**Space** $C[a,b]$) This space is a Banach space with norm given by

$$\|x\| = \max_{t\in J} |x(t)|,$$

where $J = [a,b]$.
Completeness has been proved.

- (**Incomplete Normed Spaces**) From the incomplete metric spaces that we studied before, we may readily obtain incomplete normed spaces.
For instance, the metric $d(x,y) = \int_0^1 |x(t) - y(t)| dt$ is induced by the norm

$$\|x\| = \int_0^1 |x(t)| dt.$$

- We saw that every incomplete metric space may be completed.
It turns out that it is also possible to extend the operations of a vector space and the norm to the completion, thereby completing an incomplete normed vector space.

An Incomplete Normed Space and its Completion $L^2[a,b]$

- The vector space of all continuous real-valued functions on $[a,b]$ forms a normed space X with norm defined by

$$\|x\| = \left(\int_a^b x(t)^2 dt\right)^{1/2}.$$

  This space is not complete.

- For instance, consider $[a,b] = [0,1]$.

  The sequence

$$x_m(t) = \begin{cases} 0, & \text{if } 0 \le x \le \frac{1}{2} \\ m(x - \frac{1}{2}), & \text{if } \frac{1}{2} \le x \le \frac{1}{2} + \frac{1}{m} \\ 1, & \text{if } x \ge \frac{1}{2} + \frac{1}{m} \end{cases}$$

  is Cauchy in X.

An Incomplete Normed Space and its Completion (Cont'd)

- In fact, for $n > m$, we obtain

$$\begin{array}{rcl}
\|x_n - x_m\|^2 & = & \int_0^1 [x_n(t) - x_m(t)]^2 dt \\
 & = & \int_0^{1/n} (nx - mx)^2 dx + \int_{1/n}^{1/m} (1 - mx)^2 dx \\
 & = & (n-m)^2 \frac{1}{3} x^3 \mid_0^{1/n} - \frac{1}{3m} (1 - mx)^3 \mid_{1/n}^{1/m} \\
 & = & \frac{(n-m)^2}{3n^3} + \frac{1}{3m} (1 - \frac{m}{n})^3 = \frac{(n-m)^3}{3n^3} + \frac{(n-m)^3}{3mn^3} \\
 & = & \frac{m(n-m)^2 + (n-m)^3}{3mn^3} = \frac{(n-m)^2(m+n-m)}{3mn^2} \\
 & = & \frac{(n-m)^2}{3mn^2} < \frac{1}{3m} - \frac{1}{3n}.
\end{array}$$

  This Cauchy sequence does not converge.

- The space X can be completed. The completion is denoted $L^2[a,b]$.

  This is a Banach space: In fact, the norm on X and the operations of vector space can be extended to the completion of X, as we will see in the next section.

The Spaces $L^p[a,b]$

- More generally, for any fixed real number $p \geq 1$, the Banach space $L^p[a,b]$ is the completion of the normed space which consists of all continuous real-valued functions on $[a,b]$, as before, and the norm defined by
$$\|x\|_p = \left(\int_a^b |x(t)|^p dt\right)^{1/p}.$$
The subscript p is supposed to remind us that this norm depends on the choice of p, which is kept fixed.
- The space $L^p[a,b]$ can also be obtained in a direct way by the use of the Lebesgue integral and Lebesgue measurable functions x on $[a,b]$, such that the Lebesgue integral of $|x|^p$ over $[a,b]$ exists and is finite.

  The elements of $L^p[a,b]$ are equivalence classes of those functions, where x is equivalent to y if the Lebesgue integral of $|x-y|^p$ over $[a,b]$ is zero.

Properties of Metrics Induced by Norms

Lemma (Properties of Metrics Induced by Norms)

A metric d induced by a norm on a normed space X satisfies, for all $x, y, a \in X$ and all scalars α:

(a) $d(x+a, y+a) = d(x,y)$; (**Translation Invariance**)

(b) $d(\alpha x, \alpha y) = |\alpha| d(x,y)$.

- We have $d(x+a, y+a) = \|x+a-(y+a)\| = \|x-y\| = d(x,y)$ and $d(\alpha x, \alpha y) = \|\alpha x - \alpha y\| = |\alpha| \|x-y\| = |\alpha| d(x,y)$.
- Not every metric on a vector space can be obtained from a norm: s is a vector space, but its metric d defined by

$$d(x,y) = \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|\xi_j - \eta_j|}{1 + |\xi_j - \eta_j|}$$

cannot be obtained from a norm.

This may immediately be seen from the preceding lemma.

Subsection 3

Further Properties of Normed Spaces

Subspaces of Normed Spaces and of Banach Spaces

- By definition, a **subspace** Y of a normed space X is a subspace of X considered as a vector space, with the norm obtained by restricting the norm on X to the subset Y.

  This norm on Y is said to be **induced** by the norm on X.
- If Y is closed in X, then Y is called a **closed subspace** of X.
- A **subspace** Y of a Banach space X is a subspace of X considered as a normed space. We do not require Y to be complete.

  In this connection, we get by a previous theorem,

Theorem (Subspace of a Banach Space)

A subspace Y of a Banach space X is complete if and only if the set Y is ciosed in X.

Convergent Sequences and Cauchy Sequences

- Convergence of sequences and related concepts in normed spaces follow readily from the corresponding definitions for metric spaces, given the metric
$$d(x,y)=\|x-y\|.$$

(i) A sequence (x_n) in a normed space X is **convergent** if X contains an x such that
$$\lim_{n\to\infty}\|x_n-x\|=0.$$
Then we write $x_n\to x$ and call x the **limit** of (x_n).

(ii) A sequence (x_n) in a normed space X is **Cauchy** if for every $\varepsilon>0$, there is an N, such that
$$\|x_m-x_n\|<\varepsilon,\quad \text{for all } m,n>N.$$

Infinite Series, Convergence and Absolute Convergence

- If (x_k) is a sequence in a normed space X, we can associate with (x_k) the sequence (s_n) of **partial sums** $s_n = x_1 + x_2 + \cdots + x_n$, $n = 1, 2, \ldots$.
- If (s_n) is convergent, say, $s_n \to s$, that is, $\|s_n - s\| \to 0$, then the **infinite series** or, briefly, **series**

$$\sum_{k=1}^{\infty} x_k = x_1 + x_2 + \cdots$$

  is said to **converge** or to be **convergent**. s is called the **sum** of the series and we write

$$s = \sum_{k=1}^{\infty} x_k = x_1 + x_2 + \cdots.$$

- If $\|x_1\| + \|x_2\| + \cdots$ converges, the series $\sum_{k=1}^{\infty} x_k$ is said to be **absolutely convergent**.

  Warning: In a normed space X, absolute convergence implies convergence if and only if X is complete.

Schauder Basis

- If a normed space X contains a sequence (e_n) with the property that, for every $x \in X$, there is a unique sequence of scalars (α_n) such that

$$\|x - (\alpha_1 e_1 + \cdots + \alpha_n e_n)\| \stackrel{n\to\infty}{\longrightarrow} 0,$$

then (e_n) is called a **Schauder basis** (or **basis**) for X.

- The series $\sum_{k=1}^{\infty} \alpha_k e_k$ which has the sum x is then called the **expansion** of x with respect to (e_n), and we write

$$x = \sum_{k=1}^{\infty} \alpha_k e_k.$$

Example: ℓ^p has a Schauder basis, namely (e_n), where $e_n = (\delta_{nj})$, that is, e_n is the sequence whose n-th term is 1 and all other terms are zero; thus, $e_1 = (1,0,0,0,\ldots)$, $e_2 = (0,1,0,0,\ldots)$, $e_3 = (0,0,1,0,\ldots)$, etc.

Schauder Bases and Separability

- If a normed space X has a Schauder basis, then X is separable.
  Suppose $\{x_i\}_{i=1}^{\infty}$ is a Schauder basis, with $\|x_i\| = 1$, for all i.
  We show that an arbitrary $x \in X$ can be ε-approximated by elements drawn from a countable set in X.
  By the basis property, there exist $n \in \mathbb{N}$ and α_i, $i \le n$, such that
  $$\|x - \sum_{i=1}^{n} \alpha_i x_i\| < \frac{\varepsilon}{2}.$$
  For every i, there exists β_i (in $\mathbb{Q}$ or $\mathbb{Q} + i\mathbb{Q}$), such that $|\alpha_i - \beta_i| < \frac{\varepsilon}{2^{i+1}}$.
  Therefore,
  $$\begin{array}{lll} \|x - \sum_{i=1}^{n} \beta_i x_i\| & < & \|x - \sum_{i=1}^{n} \alpha_i x_i\| + \|\sum_{i=1}^{n} \alpha_i x_i - \sum_{i=1}^{n} \beta_i x_i\| \\ & < & \frac{\varepsilon}{2} + \sum_{i=1}^{\infty} \frac{\varepsilon}{2^{i+1}} < \varepsilon. \end{array}$$
- Almost all known separable Banach spaces possess a Schauder basis.
- However, there exist separable Banach spaces that do not have have a Schauder basis (Enflo 1973).

Completion Theorem

Theorem (Completion)

Let $X = (X, \|\cdot\|)$ be a normed space. Then there is a Banach space $\widehat{X}$ and an isometry A from X onto a subspace W of $\widehat{X}$ which is dense in $\widehat{X}$. The space $\widehat{X}$ is unique, except for isometries.

- We know there exists a complete metric space $\widehat{X} = (\widehat{X}, \widehat{d})$ and an isometry $A : X \to W = A(X)$, where W is dense in $\widehat{X}$ and $\widehat{X}$ is unique, except for isometries.

  Consequently, to prove the present theorem, we must make $\widehat{X}$ into a vector space and then introduce on $\widehat{X}$ a suitable norm.

  To define on $\widehat{X}$ the two algebraic operations of a vector space, we consider any $\widehat{x}, \widehat{y} \in \widehat{X}$ and any representatives $(x_n) \in \widehat{x}$ and $(y_n) \in \widehat{y}$.

  $\widehat{x}$ and $\widehat{y}$ are equivalence classes of Cauchy sequences in X.

Completion Theorem (Addition)

- Consider the sequence (z_n), with $z_n = x_n + y_n$.

  (z_n) is Cauchy:

  $$\|z_n - z_m\| = \|x_n + y_n - (x_m + y_m)\| \leq \|x_n - x_m\| + \|y_n - y_m\|.$$

  Define the sum $\widehat{x} + \widehat{y} := \widehat{z}$, the equivalence class of (z_n).

  The definition is independent of representatives:

  Suppose $(x_n) \sim (x'_n)$ and $(y_n) \sim (y'_n)$. Then

  $$\|x_n + y_n - (x'_n + y'_n)\| \leq \|x_n - x'_n\| + \|y_n - y'_n\|,$$

  whence $(x_n + y_n) \sim (x'_n + y'_n)$.

Completion Theorem (Remaining Structure)

- Similarly, we define the product $\alpha\hat{x} \in \hat{X}$ of a scalar α and $\hat{x}$ to be the equivalence class for which (αx_n) is a representative.

  This definition is independent of the representative of $\hat{x}$.

  The zero element of X is the equivalence class containing all Cauchy sequences which converge to zero.

  It is not difficult to see that those two algebraic operations have all the properties required by the definition, so that $\hat{X}$ is a vector space.

  From the definition it follows that on W the operations of vector space induced from $\hat{X}$ agree with those induced from X by means of A.

Completion Theorem (The Norm)

- A induces on W a norm $\|\cdot\|_1$, whose value at every $\widehat{y} = Ax \in W$ is

  $$\|\widehat{y}\|_1 = \|x\|.$$

  The corresponding metric on W is the restriction of $\widehat{d}$ to W since A is isometric.

  We can extend the norm $\|\cdot\|_1$ to X by setting, for every $\widehat{x} \in \widehat{X}$,

  $$\|x\|_2 = \widehat{d}(\widehat{0}, \widehat{x}).$$

  It is obvious that $\|\cdot\|_2$ satisfies (N1) and (N2).

  The other two axioms (N3) and (N4) follow from those for $\|\cdot\|_1$ by a limit process.

Subsection 4

Finite Dimensional Normed Spaces and Subspaces

Linear Combinations

Lemma (Linear Combinations)

Let $\{x_1,\ldots,x_n\}$ be a linearly independent set of vectors in a normed space X (of any dimension). Then, there is a number $c>0$, such that, for every choice of scalars $\alpha_1,\ldots,\alpha_n$, we have

$$\|\alpha_1 x_1+\cdots+\alpha_n x_n\|\geq c(|\alpha_1|+\cdots+|\alpha_n|),\quad c>0.$$

- Let $s=|\alpha_1|+\cdots+|\alpha_n|$. If $s=0$, all α_j are zero, so the relation holds for any c. Let $s>0$. Then the inequality is equivalent to that obtained by dividing by s and writing $\beta_j=\frac{\alpha_j}{s}$, i.e., $\|\beta_1 x_1+\cdots+\beta_n x_n\|\geq c$, with $\sum_{j=1}^n|\beta_j|=1$. Hence it suffices to prove the existence of a $c>0$, such that the latter holds for every n-tuple $\beta_1,\ldots,\beta_n$, with $\sum|\beta_j|=1$. Suppose this is false. Then there exists a sequence (y_m) of vectors $y_m=\beta_1^{(m)}x_1+\cdots+\beta_n^{(m)}x_n$, with $\sum_{j=1}^n|\beta_j^{(m)}|=1$, such that $\|y_m\|\stackrel{m\to\infty}{\to}0$. Since $\sum|\beta_j^{(m)}|=1$, we have $|\beta_j^{(m)}|\leq 1$. Hence for each fixed j, the sequence $(\beta_j^{(m)})=(\beta_j^{(1)},\beta_j^{(2)},\ldots)$ is bounded.

Linear Combinations (Cont'd)

- The sequence $(\beta_j^{(m)}) = (\beta_j^{(1)}, \beta_j^{(2)}, \ldots)$ is bounded. Consequently, by the Bolzano-Weierstrass theorem, $(\beta_1^{(m)})$ has a convergent subsequence. Let β_1 denote the limit of that subsequence, and let $(y_{1,m})$ denote the corresponding subsequence of (y_m). By the same argument, $(y_{1,m})$ has a subsequence $(y_{2,m})$ for which the corresponding subsequence of scalars $(\beta_2^{(m)})$ converges. Let β_2 denote the limit. Continuing in this way, after n steps we obtain a subsequence $(y_{n,m}) = (y_{n,1}, y_{n,2}, \ldots)$ of (y_m) with terms of the form $y_{n,m} = \sum_{j=1}^n \gamma_j^{(m)} x_j$, with $\sum_{j=1}^n |\gamma_j^{(m)}| = 1$, and $\gamma_j^{(m)}$ satisfying $\gamma_j^{(m)} \stackrel{m\to\infty}{\longrightarrow} \beta_j$. Hence, $y_{n,m} \stackrel{m\to\infty}{\longrightarrow} y = \sum_{j=1}^n \beta_j x_j$, where $\sum |\beta_j| = 1$, so that not all β_j can be zero. Since $\{x_1, \ldots, x_n\}$ is a linearly independent set, we thus have $y \neq 0$. On the other hand, $y_{n,m} \to y$ implies $\|y_{n,m}\| \to \|y\|$, by the continuity of the norm. Since $\|y_m\| \to 0$ by assumption and $(y_{n,m})$ is a subsequence of (y_m), we must have $\|y_{n,m}\| \to 0$. Hence $\|y\| = 0$, so that $y = 0$. This contradicts $y \neq 0$.

Completeness

Theorem (Completeness)

Every finite dimensional subspace Y of a normed space X is complete. In particular, every finite dimensional normed space is complete.

- Consider an arbitrary Cauchy sequence (y_m) in Y. We show that it is convergent in Y; the limit will be denoted by y. Let $\dim Y = n$ and $\{e_1, \ldots, e_n\}$ any basis for Y. Then each y_m has a unique representation of the form $y_m = \alpha_1^{(m)} e_1 + \cdots + \alpha_n^{(m)} e_n$. Since (y_m) is Cauchy, for every $\varepsilon > 0$, there is an N, such that $\|y_m - y_r\| < \varepsilon$ when $m, r > N$. From this and the lemma we have, for some $c > 0$, $\varepsilon > \|y_m - y_r\| = \|\sum_{j=1}^n (\alpha_j^{(m)} - \alpha_j^{(r)}) e_j\| \geq c \sum_{j=1}^n |\alpha_j^{(m)} - \alpha_j^{(r)}|$, where $m, r > N$. Division by $c > 0$ gives $|\alpha_j^{(m)} - \alpha_j^{(r)}| \leq \sum_{j=1}^n |\alpha_j^{(m)} - \alpha_j^{(r)}| < \frac{\varepsilon}{c}$, $m, r > N$. Thus, each of the n sequences $(\alpha_j^{(m)}) = (\alpha_j^{(1)}, \alpha_j^{(2)}, \ldots)$, $j = 1, \ldots, n$, is Cauchy in $\mathbb{R}$ or $\mathbb{C}$. Hence it converges to a limit a_j.

Completeness (Cont'd)

- Using these n limits $\alpha_1,\ldots,\alpha_n$, we define $y = \alpha_1 e_1 + \cdots + \alpha_n e_n$. Clearly, $y \in Y$. Furthermore,

$$\|y_m - y\| = \|\sum_{j=1}^{n}(\alpha_j^{(m)} - \alpha_j)e_j\| \le \sum_{j=1}^{n}|\alpha_j^{(m)} - \alpha_j|\|e_j\|.$$

On the right, $\alpha_j^{(m)} \to \alpha_j$. Hence $\|y_m - y\| \to 0$, that is, $y_m \to y$. This shows that (y_m) is convergent in Y.

Since (y_m) was an arbitrary Cauchy sequence in Y, this proves that Y is complete.

Closedness

Theorem (Closedness)

Every finite dimensional subspace Y of a normed space X is closed in X.

- A subspace of a complete metric space is complete iff it is closed.
- Infinite dimensional subspaces need not be closed.

  Example: Let $X = C[0,1]$ and $Y = \text{span}(x_0, x_1, \ldots)$, where $x_j(t) = t^j$, so that Y is the set of all polynomials. Y is not closed in X.

Equivalent Norms

Definition (Equivalent Norms)

A norm $\|\cdot\|$ on a vector space X is said to be **equivalent** to a norm $\|\cdot\|_0$ on X if there are positive numbers a and b, such that, for all $x \in X$, we have

$$a\|x\|_0 \leq \|x\| \leq b\|x\|_0.$$

Claim: Equivalent norms on X define the same topology for X.

This follows from the definition and the fact that every nonempty open set is a union of open balls.

- The Cauchy sequences in $(X, \|\cdot\|)$ and $(X, \|\cdot\|_0)$ are the same.

Equivalent Norms in Finite Dimensions

Theorem (Equivalent Norms)

On a finite dimensional vector space X, any norm $\|\cdot\|$ is equivalent to any other norm $\|\cdot\|_0$.

- Let $\dim X = n$ and $\{e_1, \ldots, e_n\}$ any basis for X. Then, every $x \in X$ has a unique representation $x = \alpha_1 e_1 + \cdots + \alpha_n e_n$. By a preceding lemma, there is a positive constant c, such that $\|x\| \geq c(|\alpha_1| + \cdots + |\alpha_n|)$. On the other hand, the triangle inequality gives

$$\|x\|_0 \leq \sum_{j=1}^{n} |\alpha_j| \|e_j\|_0 \leq k \sum_{j=1}^{n} |\alpha_j|, \quad k = \max_j \|e_j\|_0.$$

  Together, $a\|x\|_0 \leq \|x\|$, where $a = \frac{c}{k} > 0$. The other inequality is obtained by interchanging the roles of $\|\cdot\|$ and $\|\cdot\|_0$ in the argument.
- This theorem implies that convergence or divergence of a sequence in a finite dimensional vector space does not depend on the particular choice of a norm on that space.

Subsection 5

Compactness and Finite Dimension

Compactness

Definition (Compactness)

A metric space X is said to be **compact** if every sequence in X has a convergent subsequence.
A subset M of X is said to be **compact** if M is compact considered as a subspace of X, that is, if every sequence in M has a convergent subsequence whose limit is an element of M.

Lemma (Compactness)

A compact subset M of a metric space is closed and bounded.

- For every $x \in \overline{M}$, there is a sequence (x_n) in M such that $x_n \to x$. Since M is compact, $x \in M$. Hence M is closed because $x \in \overline{M}$ was arbitrary.
  M is bounded: If not, it would contain an unbounded sequence (y_n), such that $d(y_n, b) > n$, where b is any fixed element. This sequence could not have a convergent subsequence since a convergent subsequence must be bounded.

On the Converse of the Compactness Lemma

Claim: The converse of this lemma is in general false.

Consider the sequence (e_n) in ℓ^2, where $e_n = (\delta_{nj})$ has the n-th term 1 and all other terms 0. This sequence is bounded since $\|e_n\| = 1$. Its terms constitute a point set which is closed because it has no point of accumulation. For the same reason, that point set is not compact.

Compactness in Finite Dimensional Normed Spaces

Theorem (Compactness)

In a finite dimensional normed space X, any subset $M \subseteq X$ is compact if and only if M is closed and bounded.

- We know compactness implies closedness and boundedness. Conversely, let M be closed and bounded. Let $\dim X = n$ and $\{e_1, \ldots, e_n\}$ a basis for X. We consider any sequence (x_m) in M. Each x_m has a representation $x_m = \xi_1^{(m)} e_1 + \cdots + \xi_n^{(m)} e_n$. Since M is bounded, so is (x_m), say, $\|x_m\| \le k$, for all m. By a preceding lemma, $k \ge \|x_m\| = \|\sum_{j=1}^n \xi_j^{(m)} e_j\| \ge c \sum_{j=1}^n |\xi_j^{(m)}|$ where $c > 0$. Hence, the sequence of numbers $(\xi_j^{(m)})$, j fixed, is bounded. By the Bolzano - Weierstrass theorem, it has a point of accumulation $\xi_j, 1 \le j \le n$. As in the proof of the preceding lemma, (x_m) has a subsequence (z_m) which converges to $z = \sum \xi_j e_j$. Since M is closed, $z \in M$. Thus, an arbitrary sequence (x_m) in M has a subsequence which converges in M.

Compactness in $\mathbb{R}^n$

- Our discussion shows the following:
  In $\mathbb{R}^n$ (or in any other finite dimensional normed space) the compact subsets are precisely the closed and bounded subsets, so that this property (closedness and boundedness) can be used for defining compactness.

  However, this can no longer be done in the case of an infinite dimensional normed space.

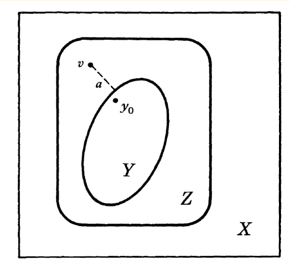
Riesz's Lemma

Riesz's Lemma

Let Y and Z be subspaces of a normed space X (of any dimension), and suppose that Y is closed and is a proper subset of Z. Then for every real number θ in the interval $(0,1)$, there is a $z \in Z$, such that $\|z\| = 1$, $\|z - y\| \geq \theta$, for all $y \in Y$.

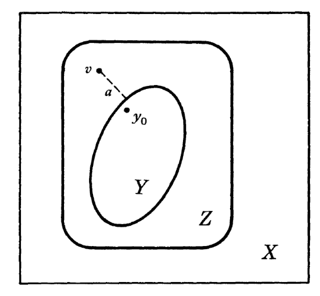
- We consider any $v \in Z - Y$ and denote its distance from Y by a: $a = \inf_{y \in Y} \|v - y\|$.

  Clearly, $a > 0$ since Y is closed. We now take any $\theta \in (0,1)$. By the definition of an infimum, there is a $y_0 \in Y$, such that $a \leq \|v - y_0\| \leq \frac{a}{\theta}$, ($\frac{a}{\theta} > a$, since $0 < \theta < 1$).



Riesz's Lemma (Cont'd)

- Let $z = c(v - y_0)$ where $c = \frac{1}{\|v-y_0\|}$. Then $\|z\| = 1$. We show that $\|z - y\| \geq \theta$, for every $y \in Y$. We have



$$\begin{array}{rcl} \|z-y\| & = & \|c(v-y_0)-y\| \\ & = & c\|v-y_0-c^{-1}y\| \\ & = & c\|v-y_1\|, \end{array}$$

where $y_1 = y_0 + c^{-1}y$.

The form of y_1 shows that $y_1 \in Y$. Hence $\|v - y_1\| \geq a$, by the definition of a. Writing c out, we obtain

$$\|z-y\| = c\|v-y_1\| \geq ca = \frac{a}{\|v-y_0\|} \geq \frac{a}{a/\theta} = \theta.$$

Since $y \in Y$ was arbitrary, this completes the proof.

Closedness of the Unit Ball Implies Finite Dimensionality

- In a finite dimensional normed space the closed unit ball is compact.

Theorem (Finite Dimension)

If a normed space X has the property that the closed unit ball $M = \{x : \|x\| \leq 1\}$ is compact, then X is finite dimensional.

- Assume that M is compact but $\dim X = \infty$. We choose any x_1 of norm 1. This x_1 generates a one dimensional subspace x_1 of X, which is closed and is a proper subspace of X since $\dim X = \infty$. By Riesz's Lemma, there is an $x_2 \in X$ of norm 1, such that $\|x_2 - x_1\| \geq \theta = \frac{1}{2}$. The elements x_1, x_2 generate a two dimensional proper closed subspace X_2 of X. By Riesz's Lemma, there is an x_3 of norm 1 such that for all $x \in X_2$, we have $\|x_3 - x\| \geq \frac{1}{2}$. In particular, $\|x_3 - x_1\| \geq \frac{1}{2}$, $\|x_3 - x_2\| \geq \frac{1}{2}$. Proceeding by induction, we obtain a sequence (x_n) of elements $x_n \in M$ such that $\|x_m - x_n\| \geq \frac{1}{2}$. Obviously, (x_n) cannot have a convergent subsequence. This contradicts the compactness of M.

Continuity and Compactness

- Compact sets are important since they are "well-behaved": they have several basic properties similar to those of finite sets and not shared by noncompact sets.

Theorem (Continuous Mapping)

Let X and Y be metric spaces and $T: X \to Y$ a continuous mapping. Then the image of a compact subset M of X under T is compact.

- By the definition of compactness, it suffices to show that every sequence (y_n) in the image $T(M) \subseteq Y$ contains a subsequence which converges in $T(M)$. Since $y_n \in T(M)$, we have $y_n = Tx_n$, for some $x_n \in M$. Since M is compact, (x_n) contains a subsequence (x_{n_k}) which converges in M. The image of (x_{n_k}) is a subsequence of (y_n) which converges in $T(M)$ because T is continuous. Hence, $T(M)$ is compact.

Maximum and Minimum Value Theorem

- The Continuous Mapping Theorem shows that the following property, well-known from calculus for continuous functions, carries over to metric spaces:

Corollary (Maximum and Minimum)

A continuous mapping T of a compact subset M of a metric space X into $\mathbb{R}$ assumes a maximum and a minimum at some points of M.

- $T(M) \subseteq \mathbb{R}$ is compact and closed and bounded by preceding lemmas. So $\inf T(M) \in T(M)$ and $\sup T(M) \in T(M)$. The inverse images of these two points consist of points of M at which Tx is minimum or maximum, respectively.

Subsection 6

Linear Operators

Linear Operators

Definition (Linear Operator)

A **linear operator** T is an operator such that:

(i) the domain $\mathscr{D}(T)$ of T is a vector space and the range $\mathscr{R}(T)$ lies in a vector space over the same field;

(ii) for all $x, y \in \mathscr{D}(T)$ and scalars α,

$$T(x+y) = Tx + Ty, \qquad T(\alpha x) = \alpha Tx.$$

- We write Tx instead of $T(x)$.
- $\mathscr{D}(T)$ denotes the domain of T.
- $\mathscr{R}(T)$ denotes the range of T.
- $\mathscr{N}(T)$ denotes the **null space** of T: this is the set of all $x \in \mathscr{D}(T)$, such that $Tx = 0$.

Operators and Arrows

- Let $\mathscr{D}(T) \subseteq X$ and $\mathscr{R}(T) \subseteq Y$, where X and Y are vector spaces, both real or both complex.
- Then T is an operator **from** (or mapping **of**) $\mathscr{D}(T)$ **onto** $\mathscr{R}(T)$, written $T : \mathscr{D}(T) \to \mathscr{R}(T)$, or from $\mathscr{D}(T)$ **into** Y, written $T : \mathscr{D}(T) \to Y$.
- If $\mathscr{D}(T)$ is the whole space X, then - and only then - we write $T : X \to Y$.

The Homomorphism Property

- The equations

$$T(x+y)=Tx+Ty, \qquad T(\alpha x)=\alpha Tx \tag{1}$$

are equivalent to

$$T(\alpha x+\beta y)=\alpha Tx+\beta Ty.$$

- By taking $\alpha=0$, we obtain

$$T0=0.$$

- The equations (1) express the fact that a linear operator T is a **homomorphism** of a vector space (its domain) into another vector space, i.e., T preserves the two operations of vector space:
  - On the left we first apply a vector space operation (addition or multiplication by scalars) and then map the resulting vector into Y;
  - On the right we first map x and y into Y and then perform the vector space operations in Y;

The outcome is the same.

Examples

- (**Identity Operator**) The identity operator $I_X : X \to X$ is defined by $I_X x = x$, for all $x \in X$. We also write simply I for I_X; thus, $Ix = x$.
- (**Zero Operator**) The zero operator $0 : X \to Y$ is defined by $0x = 0$, for all $x \in X$.
- (**Differentiation**) Let X be the vector space of all polynomials on $[a, b]$. We may define a linear operator T on X by setting

$$Tx(t) = x'(t),$$

  for every $x \in X$, where the prime denotes differentiation with respect to t. This operator T maps X onto itself.
- (**Integration**) A linear operator T from $C[a, b]$ into itself can be defined by

$$Tx(t) = \int_a^t x(\tau) d\tau, \quad \tau \in [a, b].$$

Examples (Cont'd)

- (**Multiplication by** t) Another linear operator from $C[a,b]$ into itself is defined by
$$Tx(t) = tx(t).$$
- (**Elementary Vector Algebra**) The **cross product** with one factor kept fixed defines a linear operator $T_1 : \mathbb{R}^3 \to \mathbb{R}^3$. Similarly, the **dot product** with one fixed factor defines a linear operator $T_2 : \mathbb{R}^3 \to \mathbb{R}$, say,
$$T_2 x = x \cdot a = \xi_1 \alpha_1 + \xi_2 \alpha_2 + \xi_3 \alpha_3,$$
where $a = (\alpha_j) \in \mathbb{R}^3$ is fixed.

Matrices

- A real matrix $A = (\alpha_{jk})$ with r rows and n columns defines an operator $T : \mathbb{R}^n \to \mathbb{R}^r$ by means of $y = Ax$, where $x = (\xi_j)$ has n components and $y = (\eta_j)$ has r components and both vectors are written as column vectors because of the usual convention of matrix multiplication;

  writing $y = Ax$ out, we have

$$\begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_r \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \dots & \vdots \\ \alpha_{r1} & \alpha_{r2} & \cdots & \alpha_{rn} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \vdots \\ \xi_n \end{bmatrix}.$$

  T is linear because matrix multiplication is a linear operation.

  If A were complex, it would define a linear operator from $\mathbb{C}^n$ into $\mathbb{C}^r$.

Range and Null Space

Theorem (Range and Null Space)

Let T be a linear operator. Then:

(a) The range $\mathscr{R}(T)$ is a vector space.

(b) If $\dim\mathscr{D}(T) = n < \infty$, then $\dim\mathscr{R}(T) \leq n$.

(c) The null space $\mathscr{N}(T)$ is a vector space.

(a) We take any $y_1, y_2 \in \mathscr{R}(T)$ and show that $\alpha y_1 + \beta y_2 \in \mathscr{R}(T)$, for any scalars α, β. Since $y_1, y_2 \in \mathscr{R}(T)$, we have $y_1 = Tx_1$, $y_2 = Tx_2$, for some $x_1, x_2 \in \mathscr{D}(T)$, and $\alpha x_1 + \beta x_2 \in \mathscr{D}(T)$ because $\mathscr{D}(T)$ is a vector space. The linearity of T yields $T(\alpha x_1 + \beta x_2) = \alpha Tx_1 + \beta Tx_2 = \alpha y_1 + \beta y_2$. Hence, $\alpha y_1 + \beta y_2 \in \mathscr{R}(T)$. Since $y_1, y_2 \in \mathscr{R}(T)$ were arbitrary and so were the scalars, this proves that $\mathscr{R}(T)$ is a vector space.

Range and Null Space (Cont'd)

(b) We choose $n+1$ arbitrary elements $y_1,\ldots,y_{n+1}$ of $\mathscr{R}(T)$. Then we have $y_1 = Tx_1,\ldots,y_{n+1} = Tx_{n+1}$, for some $x_1,\ldots,x_{n+1}$ in $\mathscr{D}(T)$. Since $\dim\mathscr{D}(T) = n$, this set $\{x_1,\ldots,x_{n+1}\}$ must be linearly dependent. Hence $\alpha_1 x_1 + \cdots + \alpha_{n+1} x_{n+1} = 0$, for some scalars $\alpha_1,\ldots,\alpha_{n+1}$ not all zero. Since T is linear and $T0 = 0$, application of T on both sides gives $T(\alpha_1 x_1 + \cdots \alpha_{n+1} x_{n+1}) = \alpha_1 y_1 + \cdots + \alpha_{n+1} y_{n+1} = 0$. This shows that $\{y_1,\ldots,y_{n+1}\}$ is a linearly dependent set because the α_j's are not all zero. Remembering that this subset of $\mathscr{R}(T)$ was chosen in an arbitrary fashion, we conclude that $\mathscr{R}(T)$ has no linearly independent subsets of $n+1$ or more elements. By the definition this means that $\dim\mathscr{R}(T) \le n$.

(c) We take any $x_1, x_2 \in \mathscr{N}(T)$. Then $Tx_1 = Tx_2 = 0$. Since T is linear, for any scalars α, β, we have $T(\alpha x_1 + \beta x_2) = \alpha T x_1 + \beta T x_2 = 0$. This shows that $\alpha x_1 + \beta x_2 \in \mathscr{N}(T)$. Hence $\mathscr{N}(T)$ is a vector space.

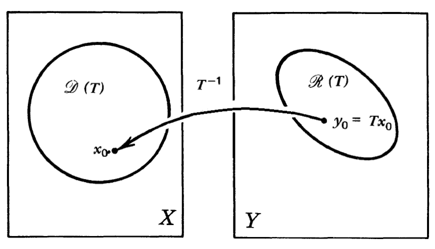
- (b) shows that linear operators preserve linear dependence.

Injective or One-to-one Mappings and Inverses

- A mapping $T:\mathscr{D}(T)\to Y$ is said to be **injective** or **one-to-one** if different points in the domain have different images, i.e., if for any $x_1,x_2\in\mathscr{D}(T)$, $x_1\neq x_2$ implies $Tx_1\neq Tx_2$.
- Equivalently, $Tx_1=Tx_2$ implies $x_1=x_2$.
- In this case there exists the mapping

$$T^{-1}:\mathscr{R}(T)\to\mathscr{D}(T);y_0\mapsto x_0,$$

which maps every $y_0\in\mathscr{R}(T)$ onto that $x_0\in\mathscr{D}(T)$ for which $Tx_0=y_0$. The mapping T^{-1} is called the **inverse** of T.



- Clearly, for all $x\in\mathscr{D}(T)$, $T^{-1}Tx=x$ and, for all $y\in\mathscr{R}(T)$, $TT^{-1}y=y$.

Inverse Operator

Theorem (Inverse Operator)

Let X, Y be vector spaces, both real or both complex. Let $T : \mathscr{D}(T) \to Y$ be a linear operator with domain $\mathscr{D}(T) \subseteq X$ and range $\mathscr{R}(T) \subseteq Y$. Then:

(a) The inverse $T^{-1} : \mathscr{R}(T) \to \mathscr{D}(T)$ exists if and only if $Tx = 0$ implies $x = 0$.

(b) If T^{-1} exists, it is a linear operator.

(c) If $\dim \mathscr{D}(T) = n < \infty$ and T^{-1} exists, then $\dim \mathscr{R}(T) = \dim \mathscr{D}(T)$.

(a) Suppose that $Tx = 0$ implies $x = 0$. Let $Tx_1 = Tx_2$. Since T is linear, $T(x_1 - x_2) = Tx_1 - Tx_2 = 0$. So $x_1 - x_2 = 0$ by the hypothesis. Hence $Tx_1 = Tx_2$ implies $x_1 = x_2$, and T^{-1} exists.

Conversely, if T^{-1} exists, then $Tx_1 = 0 = T0$ implies $x_1 = T^{-1}Tx_1 = T^{-1}T0 = 0$.

Inverse Operator (Cont'd)

(b) Assume that T^{-1} exists. The domain of T^{-1} is $\mathscr{R}(T)$ and is a vector space by a preceding theorem. We consider any $x_1, x_2 \in \mathscr{D}(T)$ and their images $y_1 = Tx_1$ and $y_2 = Tx_2$. Then $x_1 = T^{-1}y_1$ and $x_2 = T^{-1}y_2$. T is linear, so that, for any scalars α and β, we have

$$\alpha y_1 + \beta y_2 = \alpha Tx_1 + \beta Tx_2 = T(\alpha x_1 + \beta x_2).$$

Since $x_i = T^{-1}y_i$, we get

$$T^{-1}(\alpha y_1 + \beta y_2) = \alpha x_1 + \beta x_2 = \alpha T^{-1}y_1 + \beta T^{-1}y_2$$

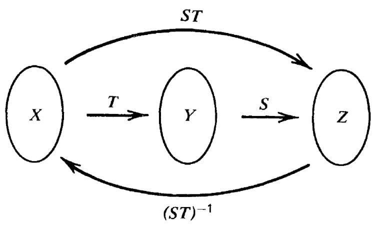
proving that T^{-1} is linear.

(c) We have $\dim \mathscr{R}(T) \leq \dim \mathscr{D}(T)$. Also, $\dim \mathscr{D}(T) \leq \dim \mathscr{R}(T)$ by the same theorem applied to T^{-1}.

Inverse of Product

Lemma (Inverse of Product)

Let $T : X \to Y$ and $S : Y \to Z$ be bijective linear operators, where X, Y, Z are vector spaces. Then the inverse $(ST)^{-1} : Z \to X$ of the product (the composite) ST exists, and $(ST)^{-1} = T^{-1}S^{-1}$.



- The operator $ST : X \to Z$ is bijective, so that $(ST)^{-1}$ exists. We thus have $ST(ST)^{-1} = I_Z$, where I_Z is the identity operator on Z. Applying S^{-1} and using $S^{-1}S = I_Y$ (the identity operator on Y), we obtain $T(ST)^{-1} = S^{-1}ST(ST)^{-1} = S^{-1}I_Z = S^{-1}$. Applying T^{-1} and using $T^{-1}T = I_X$, we obtain the desired result $(ST)^{-1} = T^{-1}T(ST)^{-1} = T^{-1}S^{-1}$.

Subsection 7

Bounded and Continuous Linear Operators

Bounded Linear Operators

Definition (Bounded Linear Operator)

Let X and Y be normed spaces and $T:\mathscr{D}(T)\to Y$ a linear operator, where $\mathscr{D}(T)\subseteq X$. The operator T is said to be **bounded** if there is a real number c such that for all $x\in\mathscr{D}(T)$, $\|Tx\|\le c\|x\|$.

- In the defining inequality, the norm on the left is that on Y, and the norm on the right is that on X.
- The formula shows that a bounded linear operator maps bounded sets in $\mathscr{D}(T)$ onto bounded sets in Y.

The Norm of a Bounded Linear Operator

- Consider the relationship $\|Tx\| \leq c\|x\|$.
  By division, $\frac{\|Tx\|}{\|x\|} \leq c$ showing that, for c to satisfy $\|Tx\| \leq c\|x\|$, for all nonzero $x \in \mathscr{D}(T)$, it must be at least as big as the supremum of the expression on the left taken over $\mathscr{D}(T) - \{0\}$.
  Thus, the smallest possible c for which $\|Tx\| \leq c\|x\|$ is the supremum.
- This quantity is denoted by $\|T\|$:
$$\|T\| = \sup_{\substack{x \in \mathscr{D}(T) \\ x \neq 0}} \frac{\|Tx\|}{\|x\|}.$$
  $\|T\|$ is called the **norm** of the operator T.
- If $\mathscr{D}(T) = \{0\}$, we define $\|T\| = 0$.
- Note that
$$\|Tx\| \leq \|T\|\|x\|.$$

The Norm Lemma

Lemma (Norm)

Let T be a bounded linear operator. Then:

(a) An alternative formula for the norm of T is

$$\|T\| = \sup_{\substack{x \in \mathscr{D}(T) \\ \|x\|=1}} \|Tx\|.$$

(b) The norm defined by $\|T\| = \sup_{\substack{x \in \mathscr{D}(T) \\ x \neq 0}} \frac{\|Tx\|}{\|x\|}$ satisfies (N1)-(N4).

(a) We write $\|x\| = a$ and set $y = \frac{1}{a}x$, where $x \neq 0$. Then $\|y\| = \frac{\|x\|}{a} = 1$.
Since T is linear,
$\|T\| = \sup_{\substack{x \in \mathscr{D}(T) \\ x \neq 0}} \frac{1}{a}\|Tx\| = \sup_{\substack{x \in \mathscr{D}(T) \\ x \neq 0}} \|T(\frac{1}{a}x)\| = \sup_{\substack{y \in \mathscr{D}(T) \\ \|y\|=1}} \|Ty\|$.
Writing x for y on the right, we have the desired equation.

The Norm Lemma (Cont'd)

(b) (N1) is obvious, and so is $\|0\| = 0$. From $\|T\| = 0$, we have $Tx = 0$, for all $x \in \mathscr{D}(T)$, so that $T = 0$. Hence (N2) holds.

Furthermore, (N3) is obtained from the following, for $x \in \mathscr{D}(T)$:

$$\sup_{\|x\|=1} \|\alpha Tx\| = \sup_{\|x\|=1} |\alpha| \|Tx\| = |\alpha| \sup_{\|x\|=1} \|Tx\|.$$

Finally, (N4) follows from the following, for $x \in \mathscr{D}(T)$:

$$\begin{aligned} \sup_{\|x\|=1} \|(T_1 + T_2)x\| &= \sup_{\|x\|=1} \|T_1 x + T_2 x\| \\ &\leq \sup_{\|x\|=1} \|T_1 x\| + \sup_{\|x\|=1} \|T_2 x\|. \end{aligned}$$

Examples

- (**Identity Operator**) The identity operator $I : X \to X$ on a normed space $X \neq \{0\}$ is bounded and has norm $\|I\| = 1$.
- (**Zero Operator**) The zero operator $0 : X \to Y$ on a normed space X is bounded and has norm $\|0\| = 0$.
- (**Differentiation Operator**) Let X be the normed space of all polynomials on $J = [0,1]$ with norm given $\|x\| = \max|x(t)|, t \in J$. A differentiation operator T is defined on X by $Tx(t) = x'(t)$, where the prime denotes differentiation with respect to t. This operator is linear but not bounded. Indeed, let $x_n(t) = t^n$, where $n \in \mathbb{N}$. Then $\|x_n\| = 1$ and $Tx_n(t) = x_n'(t) = nt^{n-1}$, so that $\|Tx_n\| = n$ and $\frac{\|Tx_n\|}{\|x_n\|} = n$. Since $n \in \mathbb{N}$ is arbitrary, this shows that there is no fixed number c, such that $\frac{\|Tx_n\|}{\|x_n\|} \leq c$. Thus, T is not bounded.

The Integral Operator

- We can define an integral operator $T : C[0,1] \to C[0,1]$ by $y = Tx$, where $y(t) = \int_0^1 k(t,\tau)x(\tau)d\tau$. Here k is a given function, which is called the **kernel** of T and is assumed to be continuous on the closed square $G = J \times J$ in the $t\tau$-plane, where $J = [0,1]$. This operator is linear. T is bounded: To prove this, note that:
  - The continuity of k on the closed square implies that k is bounded, say, $|k(t,\tau)| \le k_0$, for all $(t,\tau) \in G$, where k_0 is a real number;
  - $|x(t)| \le \max_{t \in J} |x(t)| = \|x\|$.

  Hence,

$$\begin{aligned} \|y\| &= \|Tx\| = \max_{t\in J} \left|\int_0^1 k(t,\tau)x(\tau)d\tau\right| \\ &\le \max_{t\in J} \int_0^1 |k(t,\tau)||x(\tau)|d\tau \\ &\le k_0\|x\|. \end{aligned}$$

  The result is $\|Tx\| \le k_0\|x\|$. This is the required inequality with $c = k_0$. Hence T is bounded.

Matrices

- A real matrix $A=(\alpha_{jk})$ with r rows and n columns defines an operator $T:\mathbb{R}^n\to\mathbb{R}^r$ by means of $y=Ax$, where $x=(\xi_j)$ and $y=(\eta_j)$ are column vectors with n and r components, respectively.

  In terms of components, we get $\eta_j=\sum_{k=1}^n\alpha_{jk}\xi_k, j=1,\ldots,r$. T is linear because matrix multiplication is a linear operation. T is bounded:

  To prove this, recall that the norm on $\mathbb{R}^n$ is $\|x\|=(\sum_{m=1}^n\xi_m^2)^{1/2}$. Similarly for $y\in\mathbb{R}^r$. By definition and the Cauchy-Schwarz inequality,

$$\begin{aligned}\|Tx\|^2 &= \sum_{j=1}^r\eta_j^2=\sum_{j=1}^r\left[\sum_{k=1}^n\alpha_{jk}\xi_k\right]^2\\ &\leq \sum_{j=1}^r\left[(\sum_{k=1}^n\alpha_{jk}^2)^{1/2}(\sum_{m=1}^n\xi_m^2)^{1/2}\right]^2\\ &= \|x\|^2\sum_{j=1}^r\sum_{k=1}^n\alpha_{jk}^2.\end{aligned}$$

  Note that the last double sum does not depend on x. So we can write our result in the form $\|Tx\|^2\leq c^2\|x\|^2$, where $c^2=\sum_{j=1}^r\sum_{k=1}^n\alpha_{jk}^2$. Thus, T is bounded.

Finite Dimension and Boundedness

Theorem (Finite Dimension)

If a normed space X is finite dimensional, then every linear operator on X is bounded.

- Let $\dim X = n$ and $\{e_1, \ldots, e_n\}$ a basis for X. We take any $x = \sum \xi_j e_j$ and consider any linear operator T on X. Since T is linear,

$$\|Tx\| = \|\sum \xi_j Te_j\| \leq \sum |\xi_j| \|Te_j\| \leq \max_k \|Te_k\| \sum |\xi_j|.$$

From the last sum, applying the lemma on linear combinations, we obtain

$$\sum |\xi_j| \leq \frac{1}{c} \|\sum \xi_j e_j\| = \frac{1}{c} \|x\|.$$

Together, $\|Tx\| \leq \gamma |x\|$, where $\gamma = \frac{1}{c} \max_k \|Te_k\|$. Thus, T is bounded.

Continuity and Boundedness

- Let $T : \mathscr{D}(T) \to Y$ be any operator, not necessarily linear, where $\mathscr{D}(T) \subseteq X$ and X and Y are normed spaces. The operator T is **continuous at an** $x_0 \in \mathscr{D}(T)$ if for every $\varepsilon > 0$, there is a $\delta > 0$, such that $\|Tx - Tx_0\| < \varepsilon$, for all $x \in \mathscr{D}(T)$ satisfying $\|x - x_0\| < \delta$. T is **continuous** if T is continuous at every $x \in \mathscr{D}(T)$.

Theorem (Continuity and Boundedness)

Let $T : \mathscr{D}(T) \to Y$ be a linear operator, where $\mathscr{D}(T) \subseteq X$ and X, Y are normed spaces. Then:

(a) T is continuous if and only if T is bounded.

(b) If T is continuous at a single point, it is continuous.

(a) For $T = 0$ the statement is trivial. Let $T \neq 0$. Then $\|T\| \neq 0$. We assume T to be bounded and consider any $x_0 \in \mathscr{D}(T)$. Let any $\varepsilon > 0$ be given.

Continuity and Boundedness (Cont'd)

- Since T is linear, for every $x \in \mathscr{D}(T)$, such that $\|x - x_0\| < \delta = \frac{\varepsilon}{\|T\|}$,

$$\|Tx - Tx_0\| = \|T(x - x_0)\| \leq \|T\|\|x - x_0\| < \|T\|\delta = \varepsilon.$$

Since $x_0 \in \mathscr{D}(T)$ was arbitrary, this shows that T is continuous.
Conversely, assume that T is continuous at an arbitrary $x_0 \in \mathscr{D}(T)$. Then, given any $\varepsilon > 0$, there is a $\delta > 0$, such that $\|Tx - Tx_0\| \leq \varepsilon$, for all $x \in \mathscr{D}(T)$, such that $\|x - x_0\| \leq \delta$. We now take any $y \neq 0$ in $\mathscr{D}(T)$ and set $x = x_0 + \frac{\delta}{\|y\|}y$. Then $x - x_0 = \frac{\delta}{\|y\|}y$. Hence $\|x - x_0\| = \delta$, so that we may use continuity together with linearity:

$$\|Tx - Tx_0\| = \|T(x - x_0)\| = \left\|T\left(\frac{\delta}{\|y\|}y\right)\right\| = \frac{\delta}{\|y\|}\|Ty\|.$$

Hence $\frac{\delta}{\|y\|}\|Ty\| \leq \varepsilon$. Thus, $\|Ty\| \leq \frac{\varepsilon}{\delta}\|y\|$. This can be written $\|Ty\| \leq c\|y\|$, where $c = \frac{\varepsilon}{\delta}$. This shows that T is bounded.

(b) Continuity of T at a point implies boundedness of T by the second part of the proof of (a). This, in turn, implies continuity of T by (a).

Continuity and Null Space

Corollary (Continuity, Null Space)

Let T be a bounded linear operator. Then:

(a) $x_n \to x$ (where $x_n, x \in \mathscr{D}(T)$) implies $Tx_n \to Tx$.

(b) The null space $\mathscr{N}(T)$ is closed.

(a) This follows from from $\|Tx\| \leq \|T\| \|x\|$ because, as $n \to \infty$,

$$\|Tx_n - Tx\| = \|T(x_n - x)\| \leq \|T\| \|x_n - x\| \to 0.$$

(b) For every $x \in \overline{\mathscr{N}(T)}$, there is a sequence (x_n) in $\mathscr{N}(T)$, such that $x_n \to x$. Hence $Tx_n \to Tx$ by part (a) of this Corollary. Also $Tx = 0$ since $Tx_n = 0$. So $x \in \mathscr{N}(T)$. Since $x \in \overline{\mathscr{N}(T)}$ was arbitrary, $\mathscr{N}(T)$ is closed.

Additional Properties

- The range of a bounded linear operator may not be closed:
  - The operator $T : \ell^\infty \to \ell^\infty$ defined by $y = (\eta_j) = Tx$, $\eta_j = \frac{\xi_j}{j}$, $x = (\xi_j)$ is linear and bounded.
  - The range $\mathscr{R}(T)$ is not closed in Y.
- The formula

$$\|T_1 T_2\| \le \|T_1\| \|T_2\|, \quad \|T^n\| \le \|T\|^n, \ n \in \mathbb{N},$$

is valid for bounded linear operators $T_2 : X \to Y, T_1 : Y \to Z$ and $T : X \to X$, where X, Y, Z are normed spaces.

Equality, Restriction and Extension of Operators

- Two operators T_1 and T_2 are defined to be **equal**, written $T_1 = T_2$, if they have the same domain $\mathscr{D}(T_1) = \mathscr{D}(T_2)$ and if $T_1 x = T_2 x$, for all $x \in \mathscr{D}(T_1) = \mathscr{D}(T_2)$.
- The **restriction** of an operator $T : \mathscr{D}(T) \to Y$ to a subset $B \subseteq \mathscr{D}(T)$ is denoted by $T|_B$ and is the operator defined by $T|_B : B \to Y$, $T|_B x = Tx$, for all $x \in B$.
- An **extension** of T to a set $M \supseteq \mathscr{D}(T)$ is an operator $\widetilde{T} : M \to Y$, such that $\widetilde{T}\big|_{\mathscr{D}(T)} = T$, i.e., $\widetilde{T}x = Tx$, for all $x \in \mathscr{D}(T)$.

  Hence T is the restriction of $\widetilde{T}$ to $\mathscr{D}(T)$.

Bounded Linear Extension Theorem

Theorem (Bounded Linear Extension)

Let $T:\mathscr{D}(T)\to Y$ be a bounded linear operator, where $\mathscr{D}(T)$ lies in a normed space X and Y is a Banach space. Then T has an extension $\widetilde{T}:\overline{\mathscr{D}(T)}\to Y$, where $\widetilde{T}$ is a bounded linear operator of norm $\|\widetilde{T}\|=\|T\|$.

- We consider any $x\in\overline{\mathscr{D}(T)}$. There is a sequence (x_n) in $\mathscr{D}(T)$, such that $x_n\to x$. Since T is linear and bounded, we have

  $$\|Tx_n-Tx_m\|=\|T(x_n-x_m)\|\le\|T\|\|x_n-x_m\|.$$

  Since (x_n) converges, (Tx_n) is Cauchy. Since Y is complete, (Tx_n) converges, say, $Tx_n\to y\in Y$. Define $\widetilde{T}$ by $\widetilde{T}x=y$.
  The definition is independent of the choice of a sequence in $\mathscr{D}(T)$ converging to x: Suppose that $x_n\to x$ and $z_n\to x$. Then $v_m\to x$, where (v_m) is $(x_1,z_1,x_2,z_2,\ldots)$. Hence (Tv_m) converges and the two subsequences (Tx_n) and (Tz_n) of (Tv_m) must have the same limit. Thus, T is uniquely defined at every $x\in\mathscr{D}(T)$.

Bounded Linear Extension Theorem

- Clearly, $\widetilde{T}$ is linear and $\widetilde{T}x = Tx$, for every $x \in \mathscr{D}(T)$. So $\widetilde{T}$ is an extension of T.

  We now use $\|Tx_n\| \leq \|T\|\|x_n\|$ and let $n \to \infty$. Then $Tx_n \to y = \widetilde{T}x$. Since $x \to \|x\|$ defines a continuous mapping, $\|\widetilde{T}x\| \leq \|T\|\|x\|$. Hence, $\widetilde{T}$ is bounded and $\|\widetilde{T}\| \leq \|T\|$.

  Of course, $\|\widetilde{T}\| \geq \|T\|$ because the norm, being defined by a supremum, cannot decrease in an extension.

  Together we have $\|\widetilde{T}\| = \|T\|$.

Subsection 8

Linear Functionals

Linear Functionals

- A **functional** is an operator whose range lies on the real line $\mathbb{R}$ or in the complex plane $\mathbb{C}$.
- We denote functionals by lowercase letters $f, g, h, \ldots$, the domain of f by $\mathscr{D}(f)$, the range by $\mathscr{R}(f)$ and the value of f at an $x \in \mathscr{D}(f)$ by $f(x)$, with parentheses.
- Functionals are operators, so that previous definitions apply.

Definition (Linear Functional)

A **linear functional** f is a linear operator with domain in a vector space X and range in the scalar field K of X. Thus, $f : \mathscr{D}(f) \to K$, where $K = \mathbb{R}$ if X is real and $K = \mathbb{C}$ if X is complex.

Bounded Linear Functionals

Definition (Bounded Linear Functional)

A **bounded linear functional** f is a bounded linear operator with range in the scalar field of the normed space X in which the domain $\mathscr{D}(f)$ lies. Thus, there exists a real number c, such that, for all $x \in \mathscr{D}(f)$, $|f(x)| \leq c\|x\|$. Furthermore, the **norm** of f is $\|f\| = \sup_{\substack{x \in \mathscr{D}(f) \\ x \neq 0}} \frac{|f(x)|}{\|x\|}$ or $\|f\| = \sup_{\substack{x \in \mathscr{D}(f) \\ \|x\|=1}} |f(x)|$.

- Thus, we get $|f(x)| \leq \|f\|\|x\|$.

Theorem (Continuity and Boundedness)

A linear functional f with domain $\mathscr{D}(f)$ in a normed space is continuous if and only if f is bounded.

Examples

- (**Norm**) The norm $\|\cdot\|: X \to \mathbb{R}$ on a normed space $(X, \|\cdot\|)$ is a functional on X which is not linear.
- (**Dot Product**) The familiar dot product with one factor kept fixed defines a functional $f: \mathbb{R}^3 \to \mathbb{R}$ by means of

$$f(x) = x \cdot a = \xi_1 \alpha_1 + \xi_2 \alpha_2 + \xi_3 \alpha_3,$$

where $a = (\alpha_j) \in \mathbb{R}^3$ is fixed.

f is linear.

f is bounded: In fact, $|f(x)| = |x \cdot a| \le \|x\| \|a\|$, so that $\|f\| \le \|a\|$ follows if we take the supremum over all x of norm one. On the other hand, by taking $x = a$, we obtain $\|f\| \ge \frac{|f(a)|}{\|a\|} = \frac{\|a\|^2}{\|a\|} = \|a\|$. Hence the norm of f is $\|f\| = \|a\|$.

Definite Integral

- The **definite integral** is a number, if considered for a single function.
- If we consider the integral for all functions in a certain function space, it becomes a functional on that space, say f.

  As a space let us choose $C[a,b]$. Then f is defined by

  $$f(x) = \int_a^b x(t)dt, \quad x \in C[a,b].$$

  f is linear.

  We prove that f is bounded and has norm $\|f\| = b - a$: In fact, writing $J = [a,b]$ and remembering the norm on $C[a,b]$, we obtain

  $$|f(x)| = \left|\int_a^b x(t)dt\right| \leq (b-a)\max_{t \in J}|x(t)| = (b-a)\|x\|.$$

  Taking the supremum over all x of norm 1, we obtain $\|f\| \leq b - a$.

  To get $\|f\| \geq b - a$, we choose $x = x_0 = 1$ and note that $\|x_0\| = 1$: $\|f\| \geq \frac{|f(x_0)|}{\|x_0\|} = \int_a^b dt = b - a$.

The Space $C[a,b]$

- Another functional on $C[a,b]$ is obtained if we choose a fixed $t_0 \in J = [a,b]$ and set

$$f_1(x) = x(t_0), \quad x \in C[a,b].$$

f_1 is linear.

f_1 is bounded and has norm $\|f_1\| = 1$: In fact, we have $|f_1(x)| = |x(t_0)| \leq \|x\|$. This implies $\|f_1\| \leq 1$.

On the other hand, for $x_0 = 1$, we have $\|x_0\| = 1$, whence $\|f_1\| \geq |f_1(x_0)| = 1$.

The Space ℓ^2

- We can obtain a linear functional f on the Hilbert space ℓ^2 by choosing a fixed $a = (\alpha_j) \in \ell^2$ and setting

$$f(x) = \sum_{j=1}^{\infty} \xi_j \alpha_j,$$

where $x = (\xi_j) \in \ell^2$.

This series converges absolutely and f is bounded, since the Cauchy-Schwarz inequality gives

$$|f(x)| = \left| \sum_{j=1}^{\infty} \xi_j \alpha_j \right| \leq \sum_{j=1}^{\infty} |\xi_j \alpha_j| \leq \sqrt{\sum_{j=1}^{\infty} |\xi_j|^2} \sqrt{\sum_{j=1}^{\infty} |\alpha_j|^2} = \|x\| \|a\|.$$

The Algebraic Dual Space

- The set of all linear functionals defined on a vector space X can itself be made into a vector space.
- This space is denoted by X^* and is called the **algebraic dual space** of X.
- Its algebraic operations of vector space are defined in a natural way:
  - The **sum** $f_1 + f_2$ of two functionals f_1 and f_2 is the functional s whose value at every $x \in X$ is

$$s(x) = (f_1 + f_2)(x) = f_1(x) + f_2(x);$$

  - The **product** αf of a scalar α and a functional f is the functional p whose value at $x \in X$ is

$$p(x) = (\alpha f)(x) = \alpha f(x).$$

Note that this agrees with the usual way of adding functions and multiplying them by constants.

The Second Algebraic Dual Space

- We may also consider the algebraic dual $(X^*)^*$ of X^*, whose elements are the linear functionals defined on X^*.
- We denote $(X^*)^*$ by X^{**} and call it the **second algebraic dual space** of X.
- Define the notations:

Space	General Element	Value at a Point
X	x	–
X^*	f	$f(x)$
X^{**}	g	$g(f)$

- A $g \in X^{**}$ can be obtained by choosing a fixed $x \in X$ and setting

$$g(f) = g_x(f) = f(x), \quad (x \in X \text{ fixed}, f \in X^* \text{ varable}).$$

- g_x is linear:
$g_x(\alpha f_1 + \beta f_2) = (\alpha f_1 + \beta f_2)(x) = \alpha f_1(x) + \beta f_2(x) = \alpha g_x(f_1) + \beta g_x(f_2)$.
Hence g_x is an element of X^{**}, by the definition of X^{**}.

The Canonical Mapping

- To each $x \in X$ there corresponds a $g_x \in X^{**}$.
  This defines a mapping

$$C : X \mapsto X^{**}; \quad x \mapsto g_x.$$

  C is called the **canonical mapping** of X into X^{**}.

  C is linear since its domain is a vector space and we have

$$\begin{array}{rcl} (C(\alpha x + \beta y))(f) & = & g_{\alpha x + \beta y}(f) \\ & = & f(\alpha x + \beta y) \\ & = & \alpha f(x) + \beta f(y) \\ & = & \alpha g_x(f) + \beta g_y(f) \\ & = & \alpha (Cx)(f) + \beta (Cy)(f). \end{array}$$

  C is also called the **canonical embedding** of X into X^{**}.

Isomorphism of Vector Spaces

- An **isomorphism** T **of a vector space** X onto a vector space $\tilde{X}$ over the same field is a bijective mapping which preserves the two algebraic operations of vector space:

  i.e., for all $x, y \in X$ and scalars α,

  $$T(x+y) = Tx + Ty, \quad T(\alpha x) = \alpha Tx,$$

  that is, $T : X \to \tilde{X}$ is a bijective linear operator.

  $\tilde{X}$ is then called **isomorphic** with X, and X and $\tilde{X}$ are called **isomorphic vector spaces**.

Embeddability and Algebraic Reflexivity

- It can be shown that the canonical mapping C is injective.

  Since C is linear, it is an isomorphism of X onto the range $\mathscr{R}(C) \subseteq X^{**}$.
- If X is isomorphic with a subspace of a vector space Y, we say that X is **embeddable** in Y.

  Hence X is embeddable in X^{**}, and C is also called the **canonical embedding** of X into X^{**}.
- If C is surjective (hence bijective), so that $\mathscr{R}(C) = X^{**}$, then X is said to be **algebraically reflexive**.

  We will show next that, if X is finite dimensional, then X is algebraically reflexive.

Subsection 9

Linear Operators on Finite Dimensional Spaces

Determining Linear Transformations by Action on Bases

- Let X and Y be finite dimensional vector spaces over the same field and $T : X \to Y$ a linear operator.
- We choose a basis $E = \{e_1, \ldots, e_n\}$ for X and a basis $B = \{b_1, \ldots, b_r\}$ for Y, with the vectors arranged in a definite order which we keep fixed.
- Then every $x \in X$ has a unique representation $x = \xi_1 e_1 + \cdots + \xi_n e_n$.
- Since T is linear, x has the image

$$y = Tx = T\Big(\sum_{k=1}^{n} \xi_k e_k\Big) = \sum_{k=1}^{n} \xi_k Te_k.$$

Lemma

T is uniquely determined if the images $y_k = Te_k$ of the n basis vectors $e_1, \ldots, e_n$ are prescribed.

Expression of the Image Under a Linear Transformation

- Since y and $y_k = Te_k$ are in Y, they have unique representations of the form
$$y = \sum_{j=1}^{r} \eta_j b_j, \qquad Te_k = \sum_{j=1}^{r} \tau_{jk} b_j.$$
Substitution into $y = \sum_{k=1}^{n} \xi_k Te_k$ gives
$$y = \sum_{j=1}^{r} \eta_j b_j = \sum_{k=1}^{n} \xi_k Te_k = \sum_{k=1}^{n} \xi_k \sum_{j=1}^{r} \tau_{jk} b_j = \sum_{j=1}^{r} (\sum_{k=1}^{n} \tau_{jk} \xi_k) b_j.$$
Since the b_j's form a linearly independent set, the coefficients of each b_j on the left and on the right must be the same: $\eta_j = \sum_{k=1}^{n} \tau_{jk} \xi_k$, $j = 1, \ldots, r$.

Lemma

The image $y = Tx = \sum \eta_j b_j$ of $x = \sum \xi_k e_k$ can be obtained from $\eta_j = \sum_{k=1}^{n} \tau_{jk} \xi_k$, where $Te_k = \sum_{j=1}^{r} \tau_{jk} b_j$.

Matrix Representation

- The coefficients in $\eta_j = \sum_{k=1}^n \tau_{jk}\xi_k$ form a matrix $T_{EB} = (\tau_{jk})$ with r rows and n columns.
- If a basis E for X and a basis B for Y are given, with the elements of E and B arranged in some definite order, then the matrix T_{EB} is uniquely determined by the linear operator T.
- We say that the matrix T_{EB} **represents** the operator T with respect to those bases.
- By introducing the column vectors $\widetilde{x} = (\xi_k)$ and $\widetilde{y} = (\eta_j)$ we can write $\widetilde{y} = T_{EB}\widetilde{x}$.
- Similarly, $Te = T_{EB}^\top b$, where Te is the column vector with components $Te_1, \ldots, Te_n$ (which are themselves vectors) and b is the column vector with components $b_1, \ldots, b_r$ and we have to use the transpose $T_{EB}^\top$ of T_{EB} because we sum the first subscript.
- Conversely, any matrix with r rows and n columns determines a linear operator which it represents with respect to given bases for X and Y.

Linear Functionals on Finite Dimensional Spaces

- Consider linear functionals on X, where $\dim X = n$ and $\{e_1, \ldots, e_n\}$ is a basis for X, as before.
- These functionals constitute the algebraic dual space X^* of X.
- For every such functional f and every $x = \sum \xi_j e_j \in X$, we have

$$f(x) = f(\sum_{j=1}^{n} \xi_j e_j) = \sum_{j=1}^{n} \xi_j f(e_j) = \sum_{j=1}^{n} \xi_j \alpha_j,$$

  where $\alpha_j = f(e_j)$, $j = 1, \ldots, n$.
- So f is uniquely determined by its values α_j at the n basis vectors of X.

The Converse

- Every n-tuple of scalars $\alpha_1, \ldots, \alpha_n$ determines a linear functional on X by $f(x) = \sum_{j=1}^n \xi_j \alpha_j$, where $\alpha_j = f(e_j)$, $j = 1, \ldots, n$.
- In particular, considering the n-tuples

$$(1,0,0,\ldots,0,0),(0,1,0,\ldots,0,0),\ldots,(0,0,0,\ldots,0,1)$$

we get n functionals, denoted by $f_1, \ldots, f_n$, with values

$$f_k(e_j) = \delta_{jk} = \begin{cases} 0, & \text{if } j \neq k \\ 1, & \text{if } j = k \end{cases}.$$

- f_k has the value 1 at the k-th basis vector and 0 at the $n-1$ other basis vectors.
- δ_{jk} is called the **Kronecker delta**.
- $\{f_1, \ldots, f_n\}$ is called the **dual basis** of the basis $\{e_1, \ldots, e_n\}$ for X.

Dimension of X^*

Theorem (Dimension of X^*)

Let X be an n-dimensional vector space and $E = \{e_1, \ldots, e_n\}$ a basis for X. Then $F = \{f_1, \ldots, f_n\}$ given by $f_k(e_j) = \delta_{jk}$ is a basis for the algebraic dual X^* of X, and $\dim X^* = \dim X = n$.

- F is a linearly independent set: Suppose $\sum_{k=1}^n \beta_k f_k(x) = 0$. For $x = e_j$, $\sum_{k=1}^n \beta_k f_k(e_j) = \sum_{k=1}^n \beta_k \delta_{jk} = \beta_j = 0$. So all the β_k's are zero.

  We show that every $f \in X^*$ can be represented as a linear combination of the elements of F in a unique way: We write $f(e_j) = \alpha_j$. Then $f(x) = \sum_{j=1}^n \xi_j \alpha_j$, for every $x \in X$. Also $f_j(x) = f_j(\xi_1 e_1 + \cdots + \xi_n e_n) = \xi_j$. Together, $f(x) = \sum_{j=1}^n \alpha_j f_j(x)$. Hence the unique representation of the arbitrary linear functional f on X in terms of the functionals $f_1, \ldots, f_n$ is $f = \alpha_1 f_1 + \cdots \alpha_n f_n$.

Zero Vector

Lemma (Zero Vector)

Let X be a finite dimensional vector space. If $x_0 \in X$ has the property that $f(x_0) = 0$, for all $f \in X^*$, then $x_0 = 0$.

- Let $\{e_1, \ldots, e_n\}$ be a basis for X and $x_0 = \sum_{j=1}^{n} \xi_{0j} e_j$. Then

$$f(x_0) = \sum_{j=1}^{n} \xi_{0j} \alpha_j.$$

By assumption this is zero for every $f \in X^*$, i.e., for every choice of $\alpha_1, \ldots, \alpha_n$. Hence, all ξ_{0j} must be zero.

Algebraic Reflexivity

Theorem (Algebraic Reflexivity)

A finite dimensional vector space is algebraically reflexive.

- The canonical mapping $C : X \to X^{**}$ is linear.

  If $Cx_0 = 0$, we have, by the definition of C, for all $f \in X^*$,

$$(Cx_0)(f) = g_{x_0}(f) = f(x_0) = 0.$$

  By the preceding lemma, $x_0 = 0$. Hence, the mapping C has an inverse $C^{-1} : \mathscr{R}(C) \to X$, where $\mathscr{R}(C)$ is the range of C.

  We also have $\dim \mathscr{R}(C) = \dim X$. Now $\dim X^{**} = \dim X^* = \dim X$. Together, they yield $\dim \mathscr{R}(C) = \dim X^{**}$. Hence $\mathscr{R}(C) = X^{**}$ because $\mathscr{R}(C)$ is a vector space and a proper subspace of X^{**} has dimension less than $\dim X^{**}$.

  By definition, this proves algebraic reflexivity.

Subsection 10

Normed Spaces of Operators and Dual Space

The Space $B(X,Y)$

- We consider two arbitrary normed spaces X and Y (both real or both complex) and the set $B(X,Y)$ consisting of all bounded linear operators from X into Y.
- $B(X,Y)$ becomes a vector space if we define:
  - The sum $T_1 + T_2$ of two operators $T_1, T_2 \in B(X,Y)$ in a natural way by
  $$(T_1 + T_2)x = T_1 x + T_2 x;$$
  - The product αT of $T \in B(X,Y)$ and a scalar α by
  $$(\alpha T)x = \alpha Tx.$$

Theorem (Space $B(X,Y)$)

The vector space $B(X,Y)$ of all bounded linear operators from a normed space X into a normed space Y is itself a normed space with norm defined by
$$\|T\| = \sum_{\substack{x \in X \\ x \neq 0}} \frac{\|Tx\|}{\|x\|} = \sup_{\substack{x \in X \\ \|x\|=1}} \|Tx\|.$$

Completeness

Theorem (Completeness)

If Y is a Banach space, then $B(X,Y)$ is a Banach space.

- Consider an arbitrary Cauchy sequence (T_n) in $B(X,Y)$. We show that (T_n) converges to an operator $T \in B(X,Y)$. Since (T_n) is Cauchy, for every $\varepsilon > 0$, there is an N, such that $\|T_n - T_m\| < \varepsilon$, $m,n > N$. For all $x \in X$ and $m,n > N$, we thus obtain

$$\|T_n x - T_m x\| = \|(T_n - T_m)x\| \leq \|T_n - T_m\|\|x\| < \varepsilon\|x\|.$$

Now for any fixed x and given $\widetilde{\varepsilon}$ we may choose $\varepsilon = \varepsilon_x$ so that $\varepsilon_x\|x\| < \widetilde{\varepsilon}$. Then we have $\|T_n x - T_m x\| < \widetilde{\varepsilon}$ and see that $(T_n x)$ is Cauchy in Y. Since Y is complete, $(T_n x)$ converges, say, $T_n x \to y$. Clearly, the limit $y \in Y$ depends on the choice of $x \in X$. This defines an operator $T : X \to Y$, where $y = Tx$.

Completeness (Cont'd)

- The operator T is linear:

$$\lim T_n(\alpha x+\beta z)=\lim(\alpha T_n x+\beta T_n z)=\alpha\lim T_n x+\beta\lim T_n z.$$

- We prove that T is bounded and $T_n\to T$, i.e., $\|T_n-T\|\to 0$.
  - Since $\|T_n x-T_m x\|\leq\varepsilon\|x\|$, for every $m>N$ and $T_m x\to Tx$, we may let $m\to\infty$. Using the continuity of the norm, we then obtain, for every $n>N$ and all $x\in X$,

    $$\|T_n x-Tx\|=\|T_n x-\lim_{m\to\infty}T_m x\|=\lim_{m\to\infty}\|T_n x-T_m x\|\leq\varepsilon\|x\|.$$

    This shows that (T_n-T), with $n>N$, is a bounded linear operator. Since T_n is bounded, $T=T_n-(T_n-T)$ is bounded, i.e., $T\in B(X,Y)$.
  - Furthermore, if in $\|T_n x-Tx\|\leq\varepsilon\|x\|$, we take the supremum over all x of norm 1, we obtain $\|T_n-T\|\leq\varepsilon$, $n>N$. Hence $\|T_n-T\|\to 0$.

The Dual Space X'

Definition (Dual space X')

Let X be a normed space. Then the set of all bounded linear functionals on X constitutes a normed space with norm defined by

$$\|f\| = \sup_{\substack{x \in X \\ x \neq 0}} \frac{|f(x)|}{\|x\|} = \sup_{\substack{x \in X \\ \|x\|=1}} |f(x)|$$

which is called the **dual space** of X and is denoted by X'.

- Since a linear functional on X maps X into $\mathbb{R}$ or $\mathbb{C}$ (the scalar field of X), and since $\mathbb{R}$ or $\mathbb{C}$, taken with the usual metric, is complete, we see that X' is $B(X, Y)$, with the complete space $Y = \mathbb{R}$ or $\mathbb{C}$.

Theorem (Dual Space)

The dual space X' of a normed space X is a Banach space (whether or not X is).

Isomorphisms of Normed Spaces

- An **isomorphism** of a normed space X onto a normed space $\widetilde{X}$ is a bijective linear operator $T : X \to \widetilde{X}$ which preserves the norm, that is, for all $x \in X$, $\|Tx\| = \|x\|$ (hence, T is isometric).

  X is then called **isomorphic** with $\widetilde{X}$.

  X and $\widetilde{X}$ are called **isomorphic normed spaces**.
- From an abstract point of view, X and $\widetilde{X}$ are then identical, the isomorphism merely amounting to renaming of the elements (attaching a "tag" T to each point).

The Space $\mathbb{R}^n$

- The dual space of $\mathbb{R}^n$ is $\mathbb{R}^n$.

  We have $\mathbb{R}^{n\prime} = \mathbb{R}^{n*}$, and every $f \in \mathbb{R}^{n*}$ has a representation $f(x) = \sum \xi_k \gamma_k$, $\gamma_k = f(e_k)$. By the Cauchy-Schwarz inequality,

  $$|f(x)| \leq \sum |\xi_k \gamma_k| \leq (\sum \xi_j^2)^{1/2} (\sum \gamma_k^2)^{1/2} = \|x\| (\sum \gamma_k^2)^{1/2}.$$

  Taking the supremum over all x of norm 1 we obtain $\|f\| \leq (\sum \gamma_k^2)^{1/2}$. Since for $x = (\gamma_1, \ldots, \gamma_n)$, equality is achieved in the Cauchy-Schwarz inequality, we must have $\|f\| = (\sum_{k=1}^n \gamma_k^2)^{1/2}$. This proves that the norm of f is the Euclidean norm, and $\|f\| = \|c\|$, where $c = (\gamma_k) \in \mathbb{R}^n$. Hence the mapping of $\mathbb{R}^{n\prime}$, onto $\mathbb{R}^n$ defined by

  $$f \mapsto c = (\gamma_k), \ \gamma_k = f(e_k),$$

  is norm preserving. Since it is linear and bijective, it is an isomorphism.

The Space ℓ^1

- The dual space of ℓ^1 is ℓ^∞.

  A Schauder basis for ℓ^1 is (e_k), where $e_k = (\delta_{kj})$ has 1 in the k-th place and zeros elsewhere. Then, every $x \in \ell^1$ has a unique representation $x = \sum_{k=1}^\infty \xi_k e_k$. Consider any $f \in \ell^{1\prime}$, where $\ell^{1\prime}$ is the dual space of ℓ^1. Since f is linear and bounded, $f(x) = \sum_{k=1}^\infty \xi_k \gamma_k$, $\gamma_k = f(e_k)$, where the numbers $\gamma_k = f(e_k)$ are uniquely determined by f. Also, $\|e_k\| = 1$ and $|\gamma_k| = |f(e_k)| \le \|f\| \|e_k\| = \|f\|$, $\sup_k |\gamma_k| \le \|f\|$. Hence $(\gamma_k) \in \ell^\infty$.

  On the other hand, for every $b = (\beta_k) \in \ell^\infty$, we can obtain a corresponding bounded linear functional g on ℓ^1. In fact, we may define g on ℓ^1 by $g(x) = \sum_{k=1}^\infty \xi_k \beta_k$, where $x = (\xi_k) \in \ell^1$. Then g is linear. Boundedness follows from

$$|g(x)| \le \sum |\xi_k \beta_k| \le \sup_j |\beta_j| \sum |\xi_k| = \|x\| \sup_j |\beta_j|.$$

  Hence $g \in \ell^{1\prime}$.

The Space ℓ^1 (Cont'd)

- We finally show that the norm of f is the norm on the space ℓ^∞. From $f(x) = \sum_{k=1}^{\infty} \xi_k \gamma_k$, $\gamma_k = f(e_k)$, we have
$$|f(x)| = |\sum \xi_k \gamma_k| \leq \sup_j |\gamma_j| \sum |\xi_k| = \|x\| \sup_j |\gamma_j|.$$
Taking the supremum over all x of norm 1, we see that $\|f\| \leq \sup_j |\gamma_j|$. From this, it follows $\|f\| = \sup_j |\gamma_j|$, which is the norm on ℓ^∞.
Hence this formula can be written $\|f\| = \|c\|_\infty$, where $c = (\gamma_j) \in \ell^\infty$.
It shows that the bijective linear mapping of $\ell^{1\prime}$ onto ℓ^∞ defined by $f \mapsto c = (\gamma_j)$ is an isomorphism.

The Space ℓ^p

- The dual space of ℓ^p is ℓ^q; here, $1 < p < +\infty$ and q is the conjugate of p, that is, $\frac{1}{p} + \frac{1}{q} = 1$.

  A Schauder basis for ℓ^p is (e_k), where $e_k = (\delta_{kj})$. Then, every $x \in \ell^p$ has a unique representation $x = \sum_{k=1}^{\infty} \xi_k e_k$. We consider any $f \in \ell^{p\prime}$, where $\ell^{p\prime}$ is the dual space of ℓ^p. Since f is linear and bounded, $f(x) = \sum_{k=1}^{\infty} \xi_k \gamma_k$, $\gamma_k = f(e_k)$. Let q be the conjugate of p. Consider $x_n = (\xi_k^{(n)})$ with $\xi_k^{(n)} = \begin{cases} \frac{|\gamma_k|^q}{\gamma_k}, & \text{if } k \leq n \text{ and } \gamma_k \neq 0 \\ 0, & \text{if } k > n \text{ or } \gamma_k = 0 \end{cases}$. Now we get $f(x_n) = \sum_{k=1}^{\infty} \xi_k^{(n)} \gamma_k = \sum_{k=1}^{n} |\gamma_k|^q$. We also have, using $(q-1)p = q$, $f(x_n) \leq \|f\| \|x_n\| = \|f\| (\sum |\xi_k^{(n)}|^p)^{1/p} = \|f\| (\sum |\gamma_k|^{(q-1)p})^{1/p} = \|f\| (\sum |\gamma_k|^q)^{1/p}$. Together, $f(x_n) = \sum |\gamma_k|^q \leq \|f\| (\sum |\gamma_k|^q)^{1/p}$. Dividing by the last factor and using $1 - \frac{1}{p} = \frac{1}{q}$, we get $(\sum_{k=1}^{n} |\gamma_k|^q)^{1-1/p} = (\sum_{k=1}^{n} |\gamma_k|^q)^{1/q} \leq \|f\|$. Since n is arbitrary, letting $n \to \infty$, we obtain $(\sum_{k=1}^{\infty} |\gamma_k|^q)^{1/q} \leq \|f\|$. Hence $(\gamma_k) \in \ell^q$.

The Space ℓ^p (Cont'd)

- Conversely, for any $b = (\beta_k) \in \ell^q$ we can get a corresponding bounded linear functional g on ℓ^p. In fact, we may define g on ℓ^p by setting $g(x) = \sum_{k=1}^{\infty} \xi_k \beta_k$, where $x = (\xi_k) \in \ell^p$. Then g is linear, and boundedness follows from the Hölder inequality. Hence $g \in \ell^{p\prime}$.

  We finally prove that the norm of f is the norm on the space ℓ^q. From the Hölder inequality we have

$$|f(x)| = |\sum \xi_k \gamma_k| \leq (\sum |\xi_k|^p)^{1/p} (\sum |\gamma_k|^q)^{1/q} = \|x\| (\sum |\gamma_k|^q)^{1/q}.$$

  Hence by taking the supremum over all x of norm 1 we obtain $\|f\| \leq (\sum |\gamma_k|^q)^{1/q}$. The equality sign must hold, that is, $\|f\| = (\sum_{k=1}^{\infty} |\gamma_k|^q)^{1/q}$. This can be written $\|f\| = \|c\|_q$, where $c = (\gamma_k) \in \ell^q$ and $\gamma_k = f(e_k)$.

  The mapping of $\ell^{p\prime}$ onto ℓ^q defined by $f \mapsto c$ is linear and bijective. Since it is norm preserving, it is an isomorphism.

Comments on Duals and Double Duals

- In applications it is frequently quite useful to know the general form of bounded linear functionals on spaces of practical importance:
  - We gave general representations of bounded linear functionals on $\mathbb{R}^n$, ℓ^1 and ℓ^p with $p > 1$.
  - The space $C[a, b]$ will be considered later, after establishing the Hahn-Banach Theorem.
- Furthermore, it is worthwhile to consider $X'' = (X')'$, the second dual space of X.

  We have to postpone this discussion for after having developed suitable tools for obtaining substantial results in that direction.