0. BANACH SPACE BACKGROUND

Normed linear spaces, Hilbert spaces

We assume that the reader is familiar with the notions normed linear space, Banach space, inner product space, Hilbert space, and with the really basic facts about such spaces. Here we give a brief summary of the results that are particularly relevant to our purposes. Proofs are given only when it cannot be confidently asserted that they are to be found in any elementary text on the subject. At the same time, we establish some notation.

We use the same notation $\| \ \|$ for the norms in the various spaces considered, except when it is necessary to distinguish different norms. The (closed) unit ball in a normed linear space X (denoted by U_X) is the set $\{x \in X : \|x\| \leqslant 1\}$.

The scalar field may be either $\mathbb{R}$ or $\mathbb{C}$. Most results will apply to both cases simultaneously, or with minor modifications for the complex case. Exceptions to this will be pointed out.

0.1. Every Hilbert space has a (finite or infinite) orthonormal basis (b_j). For each element x,

$$x = \sum_j \langle x, b_j \rangle \, b_j \ ,$$

$$\|x\|^2 = \sum_j |\langle x, b_j \rangle|^2 \ .$$

(This means in the sense of "summation" when (b_j) is uncountable, but our main interest is in the finite-dimensional case.)

Operators

A linear operator T (from one normed linear space to another) is continuous if and only if there exists M such that $\|Tx\| \leqslant M\|x\|$ for all $x \in X$.

We will use the word "operator" to mean "continuous linear operator". The (linear) space of operators from X to Y will be denoted by L(X,Y), and we write L(X) for L(X,X). "Operator norm" is defined on L(X,Y) by:

$$\|T\| = \sup\{ \|Tx\| : x \in U_X\} .$$

This is a norm, and $\|TS\| \leqslant \|T\|.\|S\|$.

An operator T is an isometry if $\|Tx\| = \|x\|$ for all x. Spaces X,Y are said to be isometric if there is an isometry of X onto Y. An operator T is an isomorphism if it is bijective and T, T^{-1} are both continuous. Spaces X,Y are said to be isomorphic if there is an isomorphism of X onto Y. The Banach-Mazur distance between X and Y is then defined to be

$$d(X,Y) = \inf \{ \|T\|.\|T^{-1}\| : T \text{ an isomorphism of } X \text{ onto } Y\}.$$

Though not truly a "distance" (or metric), this is a measure of the similarity of X and Y. Clearly, $d(X,Y) \geqslant 1$, with equality when X is isometric to Y. Also, $d(X,Z) \leqslant d(X,Y)\, d(Y,Z)$.

We say that T is an M-open operator of X onto Y if, given $y \in Y$, there exists $x \in X$ with $Tx = y$ and $\|x\| \leqslant M\|y\|$.

Duality and the Hahn-Banach theorem

Operators mapping into the scalar field ($\mathbb{R}$ or $\mathbb{C}$) are called linear functionals. The space of all continuous linear functionals on X, with operator norm, is the dual space X*.

The Hahn-Banach theorem is the basic theorem on extension (and existence) of linear functionals. There are two versions, as follows. A real-valued function p (on a real linear space X) is sublinear if $p(\lambda x) = \lambda p(x)$ and $p(x+y) \leqslant p(x) + p(y)$ for all $x,y \in X$ and $\lambda \geqslant 0$.

0.2 Theorem. (i) Let X be a real linear space, X_1 a linear subspace and p a sublinear real function on X. Let f_1 be a linear functional defined on X_1, with $f_1(x) \leqslant p(x)$ for all $x \in E$. Then there is a linear functional f on X that extends f_1 and satisfies $f(x) \leqslant p(x)$ for all $x \in X$.

(ii) Let X be a normed linear space (real or complex), X_1 a linear subspace. Let f_1 be a continuous linear functional defined on X_1. Then there is a linear functional f on X that extends f_1 and satisfies $\|f\| = \|f_1\|$.

0.3 Corollary. Let x_0 be an element of a normed linear space X. Then there is an element f of U_{X^*} such that $f(x_0) = \|x_0\|$.

0.4 Corollary. Any normed linear space X embeds isometrically into its second dual X^{**}, under the mapping J defined by : $(Jx)(f) = f(x)$ for $f \in X^*$. (If J maps onto X^{**}, then X is said to be reflexive).

A norming subset of U_{X^*} is a subset K such that $\|x\| = \sup\{|f(x)| : f \in K\}$ for all $x \in X$.

For T in L(X,Y), the adjoint (or dual) operator T^* in $L(Y^*,X^*)$ is defined by : $(T^*g)(x) = g(Tx)$ for $g \in Y^*$, $x \in X$.

0.5. $\|T^*\| = \|T\|$. Hence if X is isomorphic to Y, then $d(X^*,Y^*) \leqslant d(X,Y)$, with equality if X,Y are reflexive.

Some particular spaces

We denote by $\ell_\infty(S)$ the set of all bounded functions (real or complex, according to context) on a set S, with norm defined by : $\|x\| = \sup\{|x(s)| : s \in S\}$. We write simply ℓ_∞ for $\ell_\infty(\mathbb{N})$, and ℓ_∞^n for $\ell_\infty(S)$ when S $= \{1,2, \dots ,n\}$; of course, ℓ_∞^n is simply $\mathbb{R}^n$ or $\mathbb{C}^n$ with the above norm. (We shall normally regard elements of $\mathbb{R}^n$, $\mathbb{C}^n$ as functions on $\{1,2, \dots ,n\}$, hence we use the notation x(j) for the jth term).

When K is a compact topological space, we denote by C(K) the space of all continuous real (or complex) functions on K, with norm defined as for $\ell_\infty(S)$.

The symbol ℓ_p^n (for $p \geqslant 1$) denotes $\mathbb{R}^n$ (or $\mathbb{C}^n$) with norm :

$$\|x\|_p = \Big(\sum_i |x(i)|^p \Big)^{1/p} .$$

In particular, $\|x\|_1 = \sum_i |x(i)|$. Further, ℓ_p denotes the space of all infinite sequences x for which $\|x\|_p$ (defined in the same way) is finite. We shall distinguish "real ℓ_p^n " and "complex ℓ_p^n " when it matters.

The norm of ℓ_2^n (or ℓ_2) is derived from the "natural" inner product:

$$\langle x, y \rangle = \sum_i x(i) \overline{y(i)} .$$

Every n-dimensional Hilbert space is isometric to ℓ_2^n .

Cambridge University Press
978-0-521-34937-6 - Summing and Nuclear Norms in Banach Space Theory
G. J. O. Jameson
Excerpt
More information

We use the notation e_j for the sequence (finite or infinite) having 1 in place j and 0 elsewhere.

The identity in $\mathbb{R}^n$ (or $\mathbb{C}^n$), regarded as an operator from ℓ_p^n to ℓ_q^n, will be denoted by $I_{p,q}^{(n)}$.

0.6. $\|I_{1,2}^{(n)}\| = \|I_{2,\infty}^{(n)}\| = \sqrt{n}$.

Proof. Easy, except $\|x\|_1 \leqslant \sqrt{n}\,\|x\|_2$. This follows from $\Sigma(|x(i)| - c)^2 \geqslant 0$, with $c = \frac{1}{n}\|x\|_1$.

0.7. The duals of ℓ_1^n, ℓ_2^n, ℓ_∞^n are isometric to ℓ_∞^n, ℓ_2^n, ℓ_1^n respectively (and the dual of ℓ_p^n to $\ell_{p'}^n$, where $\frac{1}{p} + \frac{1}{p'} = 1$). In each case, the functional corresponding to an element y is f_y, where $f_y(x) = \sum_i x(i)y(i)$.

The spaces $L_p(\mu)$ will occasionally be mentioned in examples, but nothing of importance in this book depends on measure theory.

Finite-dimensional spaces

0.8 Theorem. Every linear mapping defined on a finite-dimensional normed linear space is continuous. Consequently all n-dimensional normed linear spaces (over the same field) are isomorphic.

0.9 Corollary. If X is finite-dimensional, then U_X is compact.

If dim X = n, then by elementary algebra, dim X* = n. Hence X is isometric to X** and d(X*,Y*) = d(X,Y). In fact, if $\{b_1, \dots ,b_n\}$ is a basis of X, then the dual basis of X* is $\{f_1, \dots ,f_n\}$, where the f_i are defined by : $f_i(x_j) = \delta_{ij}$. Clearly if $\|b_i\| = 1$, then $\|f_i\| \geqslant 1$.

0.10 Theorem. Let X be a n-dimensional normed linear space. Then there exists a basis $\{b_1, \dots ,b_n\}$ of X, with dual basis $\{f_1, \dots ,f_n\}$, such that $\|b_i\| = \|f_i\| = 1$ for all i. (Such a basis is called an Auerbach basis).

Proof. Take any basis $\{a_1, \dots ,a_n\}$ of X, and let T be the corresponding isomorphism of X onto $\mathbb{R}^n$ (or $\mathbb{C}^n$). Given elements $x_1, \dots ,x_n$ of X, let $D(x_1, \dots ,x_n)$ be the determinant of the matrix with columns $Tx_1, \dots ,Tx_n$. Then D is a continuous function on X^n, since it is formed by taking sums and products of coordinate functionals. Hence D attains its maximum absolute value on the compact set $(U_X)^n$, say at $(b_1, \dots, b_n)$. Write $D(b_1, \dots, b_n) = \mu$, and define

$$f_i(x) = \frac{1}{\mu} D(b_1, \dots, x, \dots, b_n)$$

(in which b_i is replaced by x). Then $f_i(b_i) = 1$, and $|f_i(x)| \leqslant 1$ for $x \in U_X$. By the elementary properties of determinants, f_i is linear and $f_i(b_j) = 0$ for $i \neq j$.

0.11 Corollary. If dim X = n, then $d(X, \ell_\infty^n) \leqslant n$, $d(X, \ell_2^n) \leqslant n$. If dim X = dim Y = n, then $d(X,Y) \leqslant n^2$.

Proof. With $\{b_j\}$ as in 0.10, let $Tx = \Sigma f_i(x)e_i \in \ell_\infty^n$. We have $\max |f_i(x)| \leqslant \|x\| \leqslant \Sigma |f_i(x)|$, hence $\|Tx\| \leqslant \|x\| \leqslant n\|Tx\|$.

Some statements about infinite-dimensional spaces are really statements about their finite-dimensional subspaces. This motivates the following definition. If X,Y are normed linear spaces, we say that Y is finitely represented in X if, given any finite-dimensional subspace Y_1 of Y and $\varepsilon > 0$, there is a subspace X_1 of X such that $d(X_1, Y_1) < 1 + \varepsilon$. This says that all the finite-dimensional subspaces of Y are "nearly isometric" to subspaces of X.

Embedding in $\ell_\infty(S)$

0.12 Proposition. Every normed linear space is isometric to a subspace of $\ell_\infty(S)$ for some set S, and to a subspace of C(K) for some compact space K.

Proof. Let $S = U_{X^*}$. Given x in X, define Jx in $\ell_\infty(S)$ by : $(Jx)(f) = f(x)$. It follows from 0.3 that $\|Jx\| = \|x\|$. The same construction proves the second statement, since U_{X^*} is compact in the weak-star topology, and Jx is continuous with respect to this topology. (Familiarity with the weak-star topology is not really needed for the purposes of this book).

For the first statement in 0.12, it is clearly enough for S to be a norming subset of U_{X^*}.

An important variation of this for finite-dimensional spaces is :

0.13 Proposition. Let $\dim X = n$ and $\varepsilon > 0$. Then there exist N and a subspace X_0 of ℓ_∞^N such that $d(X, X_0) \leqslant 1 + \varepsilon$.

Proof. The set $S_{X^*} = \{f \in X^* : \|f\| = 1\}$ is totally bounded, so contains elements $f_1, \dots, f_N$ such that, given any $f \in S_{X^*}$, we have $\|f - f_i\| \leqslant \varepsilon$ for some i. For $x \in X$, let Jx be the element $[f_1(x), \dots, f_N(x)]$ of ℓ_∞^N. Using 0.3, we have

$$(1 - \varepsilon)\|x\| \leqslant \|Jx\| \leqslant \|x\| .$$

A variant of this gives precise isometric embedding into a close copy of ℓ_∞^N.

0.14. Let $\dim X = n$ and $\varepsilon > 0$. Then there exist N and a space Y such that $d(Y, \ell_\infty^N) \leqslant 1 + \varepsilon$ and X is isometric to a subspace of Y.

Proof. Let $T : X \to \ell_\infty^N$ be such that $\|x\| \leqslant \|Tx\| \leqslant (1+\varepsilon)\|x\|$ for all x in X. Let U_∞ be the unit ball of ℓ_∞^N, and let Y be $\mathbb{R}^N$ with the norm defined by taking as unit ball the convex cover of $U_\infty \cup T(U_X)$. One verifies easily that the conditions hold.

Extensions and projections

Let X be a subspace of a normed linear space Y. A projection of Y onto X is an operator $P : Y \to X$ such that $Px = x$ for all $x \in X$. If there is such a projection, then X is said to be complemented in Y; it must then be a closed subspace, since $X = \ker(I - P)$.

0.15 Proposition. If Y is a Hilbert space, X a closed subspace, then there is a projection (the "orthogonal" projection) P of Y onto X with $\|P\| = 1$. The kernel of P is $X^\perp$. If X is finite-dimensional, then P is given by:

$$Py = \sum_i \langle y, b_i \rangle b_i$$

where $\{b_i\}$ is an orthonormal basis of X.

A normed linear space X is said to be injective if there is a real number λ such that the following holds: given any normed linear space E, a subspace E_1 and T_1 in $L(E_1, X)$, there is an extension T in $L(E, X)$ with

$\|T\| \leq \lambda\|T_1\|$. We define $\lambda(X)$ to be the infimum of such λ.

0.16. If X is isomorphic to Y, then $\lambda(Y) \leq d(X,Y)\,\lambda(X)$.

Proof. Elementary.

0.17 Proposition. For any set S, $\lambda[\ell_\infty(S)] = 1$.

Proof. Let E_1 be a subspace of E, and let T_1 be an operator from E_1 to $\ell_\infty(S)$. For each $s \in S$, define $f_s \in E_1^*$ by $f_s(e) = (T_1e_1)(s)$. Then $\|f_s\| \leq \|T_1\|$. By the Hahn-Banach theorem, f_s can be extended to $g_s \in E^*$ with $\|g_s\| \leq \|T_1\|$. For $e \in E$, define $(Te)(s) = g_s(e)$.

0.18 Corollary. If $\dim X = n$, then $\lambda(X) \leq d(X, \ell_\infty^n) \leq n$.

Proof. By 0.16, 0.17, 0.11.

0.19 Proposition. The following statements (for a given space X) are equivalent:

(i) $\lambda(X) \leq \lambda$,

(ii) if X is isometric to a subspace X_0 of a space Y, and $\varepsilon > 0$, then there is a projection P of Y onto X_0 with $\|P\| \leq (1+\varepsilon)\lambda$,

(iii) for some set S, X is isometric to a subspace X_0 of $\ell_\infty(S)$, and for every $\varepsilon > 0$, there is a projection P of $\ell_\infty(S)$ onto X_0 with $\|P\| \leq (1+\varepsilon)\lambda$.

Proof. (i) implies (ii). We have $\lambda(X_0) = \lambda$. The projection P is obtained by extending I_{X_0}.

(ii) implies (iii), clearly.

(iii) implies (i). Let E_1 be a subspace of E, and T_1 an element of $L(E_1, X_0)$. By 0.16, there is an extension $\bar{T}$ in $L(E, \ell_\infty(S))$ with $\|\bar{T}\| = \|T_1\|$. With P as in (iii), let $T = P\bar{T}$. Then T is in $L(E, X_0)$, extends T_1 and has $\|T\| \leq (1+\varepsilon)\,\lambda\|T_1\|$. Hence $\lambda(X) = \lambda(X_0) \leq (1+\varepsilon)\lambda$.

Because of the equivalence with (ii), $\lambda(X)$ is called the projection constant of X. We will see that in fact for n-dimensional X, both $\lambda(X)$ and $d(X, \ell_2^n)$ are not greater than $\sqrt{n}$ (compare 0.11 and 0.18). We shall also describe the evaluation of the projection constants of ℓ_1^n and ℓ_2^n.

Orderings; linear lattices

The (real) spaces $\mathbb{R}^n$, ℓ_p, $\ell_\infty(S)$, $C(S)$, $L_p(\mu)$ all have a natural partial ordering defined "pointwise" : $x \leqslant y$ means $x(s) \leqslant y(s)$ for each s.

In general, a real linear space is said to be a linear lattice (or Riesz space) if it has a partial ordering $\leqslant$ such that :

(i) if $x \leqslant y$ and $y \leqslant x$, then $x = y$;

(ii) if $x \leqslant y$, then $x + z \leqslant y + z$ for all z ;

(iii) if $x \geqslant 0$ and $\lambda \geqslant 0$, then $\lambda x \geqslant 0$;

(iv) any two elements have a supremum.

The supremum of x and -x is denoted by $|x|$. The above examples are clearly linear lattices, and $|x|$ is the function given by : $|x|(s) = |x(s)|$. Note that $|x| \leqslant y$ is equivalent to $-y \leqslant x \leqslant y$.

A norm on a linear lattice is a lattice norm if :

(i) $0 \leqslant x \leqslant y$ implies $\|x\| \leqslant \|y\|$,

and (ii) $\|\,|x|\,\| = \|x\|$ for all x.

The space is then called a normed lattice. The above examples are all normed lattices.

A linear mapping T between linear lattices is positive if $Tx \geqslant 0$ whenever $x \geqslant 0$. This definition applies in particular to linear functionals, thereby giving a partial ordering of the dual space. The functional f_y on $\mathbb{R}^n$ defined by $f_y(x) = \Sigma x(i)y(i)$ is positive if and only if $y \geqslant 0$.

The spaces ℓ_1^k, ℓ_1, $L_1(\mu)$ have the special property that $\|\Sigma x_i\| = \Sigma\|x_i\|$ for positive elements (in general, normed lattices with this property are called L-spaces).

The above terminology will be used where appropriate, but we do not assume any knowledge of the general theory of normed lattices.

$\mathcal{L}_1$ and $\mathcal{L}_\infty$ spaces

These notions are the key to the extension of certain results to the infinite-dimensional case. However, they can be omitted without serious loss.

We will say that a Banach space X is an $\mathcal{L}_p$-space if for every finite-dimensional subspace E of X and $\varepsilon > 0$, there is a finite-dimensional subspace F such that $E \subseteq F \subseteq X$ and $d(F, \ell_p^N) \leqslant 1 + \varepsilon$, where $N = \dim F$.

(This is not quite the usual terminology; according to this, X is an "$\mathcal{L}_{p,\lambda}$-space" if we have $d(F,\ell_p^N) \leqslant \lambda$; hence our definition equates to an "$\mathcal{L}_{p,1+\varepsilon}$-space for every $\varepsilon > 0$".)

We only need the fact the certain naturally arising spaces are $\mathcal{L}_\infty$ or $\mathcal{L}_1$ spaces. The next lemma is useful for this purpose.

0.20 Lemma. The following is sufficient for X to be an $\mathcal{L}_p$-space. Given $b_1, \dots, b_n \in X$ and $\varepsilon > 0$, there is a finite-dimensional subspace F of X such that $d(F, \ell_p^N) \leq 1+\varepsilon$ (where N = dim F) and $\text{dist}(b_i, F) < \varepsilon$ for each i.

Proof. Given E, let $\{b_i\}$ be an Auerbach basis of E. . Let F be as stated, with ε replaced by $\varepsilon' = \varepsilon/2n^2$. For each i, take $f_i \in F$ with $\|b_i - f_i\| < \varepsilon'$. Essentially, we modify F by replacing the b_i's by the f_i's. Define $J : E \to F$ by $Jb_i = f_i$. One verifies easily that $\|b - Jb\| \leq n\varepsilon'\|b\|$ for $b \in B$. We may assume $n\varepsilon' \leq \frac{1}{2}$: then $\|b - Jb\| \leq 2n\varepsilon'\|Jb\|$. Let P be a projection of F onto J(B) with $\|P\| \leq n$ (see 0.18), and define T on F by :

$$Tx = J^{-1}Px + (I - P)x .$$

Then T(F) contains B, and

$$Tx - x = J^{-1}Px - Px = b - Jb ,$$

where $b = J^{-1}Px$. Hence $\|Tx - x\| \leq 2n\varepsilon'\|Px\| \leq 2n^2\varepsilon'\|x\| = \varepsilon\|x\|$, from which $\|T\|. \|T^{-1}\| \leq (1+\varepsilon)/(1-\varepsilon)$.

0.21. ℓ_1 is an $\mathcal{L}_1$-space, and c_0 (the space of sequences tending to 0) is an $\mathcal{L}_\infty$-space.

Proof. Let $E_N = \text{lin}(e_1, \dots, e_N)$. This is isometric to ℓ_1^N, ℓ_∞^N in the two cases. Given elements $b_1, \dots, b_n$ and $\varepsilon > 0$, there exists N such that $\text{dist}(b_i, E_N) < \varepsilon$ for each i.

0.22. If μ is a positive measure on a measure space, then $L_1(\mu)$ is an $\mathcal{L}_1$-space, $L_\infty(\mu)$ an $\mathcal{L}_\infty$-space.

Proof. In both cases, simple functions are dense. Hence if elements $b_1, \dots, b_n$ and $\varepsilon > 0$ are given, there are disjoint measurable sets $A_1, \dots, A_N$ (with finite measure in the case of $L_1(\mu)$) such that $\text{dist}(b_i, F) \leq \varepsilon$ for each i, where F is the subspace spanned by the characteristic functions of the A_i. It is easily seen that F is isometric to ℓ_1^N, ℓ_∞^N in the two cases.

In particular, $\ell_\infty(S)$ is an $\mathcal{L}_\infty$-space. By 0.12, it is clear that the property of being an $\mathcal{L}_\infty$-space is not inherited by subspaces.

For readers with sufficient grounding in General Topology, we show also that C(K) is an $\mathcal{L}_\infty$-space. We use the fact that if $\{G_1, \ldots, G_N\}$ is an open covering of a compact, Hausdorff space K, then there exist non-negative continuous functions g_i such that $g_1 + \ldots + g_N = 1$ and $g_i(s) = 0$ for s not in G_i (a "partition of unity"). The covering might as well be chosen so that each G_i contains a point s_i not in the other G_j : clearly, we then have $g_i(s_i) = 1$.

<u>0.23 Lemma.</u> Under these conditions, $\text{lin}(g_1, \ldots, g_N)$ is isometric to ℓ_∞^N .

Proof. Let $g = \sum_i \lambda_i g_i$. Evaluation at s_i shows that $\|g\| \geq |\lambda_i|$ for each i. Conversely, if $M = \max |\lambda_i|$, then

$$|g(s)| \leq M \sum_i g_i(s) = M$$

for all s.

<u>0.24 Proposition.</u> If K is a compact, Hausdorff space, then C(K) is an $\mathcal{L}_\infty$-space.

Proof. Take elements $f_1, \ldots, f_n$ and $\varepsilon > 0$. Let M be an integer such that $\|f_i\| \leq M\varepsilon$ for each i, and write the numbers $r\varepsilon$ $(r = -M, \ldots, M-1, M)$ as $\lambda_1, \ldots, \lambda_k$. For $(r_1, \ldots, r_n)$ in $\{1, \ldots, k\}^n$, let

$$G(r_1, \ldots, r_n) = \{s : |f_i(x) - \lambda_{r_i}| < \varepsilon \text{ for each } i\} .$$

These sets form an open covering: remove any that are empty or contained in the others, and let B be the set of remaining $(r_1, \ldots, r_n)$. Choose corresponding functions $g_{r_1,\ldots,r_n}$ to form a partition of unity. We must show that each f_i is close to the linear subspace spanned by these functions. We do this for f_1. Let

$$h_1 = \Sigma \{\lambda_{r_1}, \ldots, r_n - (r_1, \ldots, r_n) \in B\} .$$

Choose $s \in K$. There exists p such that $\lambda_p \leq f_1(s) < \lambda_p + \varepsilon$. Then s belongs to $G(r_1, \ldots, r_n)$ only for $r_1 = p$ and $r_1 = p + 1$ (in which case $\lambda_{r_1} = \lambda_p + \varepsilon$). Hence

$$h_1(s) = \lambda_p \Sigma g_{p, r_2, \ldots, r_n}(s) + (\lambda_p + \varepsilon) \Sigma g_{p+1, r_2, \ldots, r_n}(s)$$