(*August 30, 2005*)

Normed and Banach Spaces

Paul Garrett garrett@math.umn.edu http://www.math.umn.edu/~garrett/

We have seen that many interesting spaces of functions have natural structures of Banach spaces: $C^o(K)$ for K compact, as well as $C^k[a,b]$, and generalizations of these.

Banach spaces are less special than Hilbert spaces, but still sufficiently simple that their fundamental properties can be explained readily. Several standard results which are true in greater generality have simpler and more transparent proofs in this setting.

The Banach-Steinhaus (uniform boundedness) theorem and the open mapping theorem are significantly more substantial than the first results here, since they invoke the Baire category theorem. The Hahn-Banach theorem is non-trivial, but does *not* use completeness.

Finally, as made clear in work of Gelfand, of Grothendieck, and of many others, many subtler sorts of topological vectorspaces are expressible as *limits* of Banach spaces, making clear that Banach spaces play an even more central role than would be apparent from many conventional elementary functional analysis texts. We will pursue this later.

---

1. Basic Definitions

A complex vectorspace[1] V with a real-valued function

$$| \, | : V \longrightarrow \mathbf{R}$$

so that

$$\begin{array}{cc} |x+y| \le |x| + |y| & \text{(triangle inequality)} \\ |\alpha x| = |\alpha||x| & (\alpha \text{ complex, } x \in V) \\ |x| = 0 \longrightarrow x = 0 & \text{(positivity)} \end{array}$$

is a **normed complex vectorspace**, or simply **normed space**. Because of the triangle inequality, the function

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

is a *metric*. When the space V is *complete* with respect to this metric, V is a **Banach space**.

Because of the Cauchy-Schwarz-Bunyakowsky inequality, pre-Hilbert spaces are normed spaces, and Hilbert spaces are Banach spaces. But there are certainly many Banach spaces which are *not* Hilbert spaces.

---

2. Normed spaces of linear maps

[1] There are occasions where one wants the scalars to be $\mathbf{R}$ rather than $\mathbf{C}$. In fact, for many purposes, the scalars need not be the real or complex numbers, need not be locally compact, and need not even be commutative. Mostly these other possibilities will not concern us.

There is a *natural norm* on the collection of all continuous (k-)linear maps $T : X \longrightarrow Y$ from one normed space X to another one Y (over common scalars k).

Let $\mathrm{Hom}^o(X,Y)$ denote[2] the collection of all continuous k-linear maps from the normed k-vectorspace X to the normed k-vectorspace Y. We use the same notation $|\,|$ for the norms on both X and Y, since context should make clear which is meant.

For a k-linear (not necessarily continuous) map $T : X \longrightarrow Y$ from one normed space to another the **uniform norm** is

$$|T| = |T|_{\text{uniform}} = \sup_{|x| \le 1} |Tx|$$

where we allow the value $+\infty$. Such a linear map T is called **bounded** if $|T| < +\infty$. There are several fairly obvious variants of the expression for the uniform norm:

$$|T| = \sup_{|x| \le 1} |Tx| = \sup_{|x| < 1} |Tx| = \sup_{|x| \ne 0} \frac{|Tx|}{|x|}$$

Proposition: For a k-linear map $T : X \longrightarrow Y$ from one normed space to another, the following conditions are equivalent:
- T is *continuous.*
- T is *continuous at one point.*
- T is *bounded.*

Proof: First, let's show that continuity at a point x_o implies continuity everywhere. Take another point x_1. Given $\varepsilon > 0$, take $\delta > 0$ so that $|x - x_o| < \delta$ implies $|Tx - Tx_o| < \varepsilon$. Then for $|x' - x_1| < \delta$

$$|(x' + x_o - x_1) - x_o| < \delta$$

Invoking the linearity of T,

$$|Tx' - Tx_1| = |T(x' + x_o - x_1) - Tx_o| < \varepsilon$$

which is the desired continuity at x_1.

Now suppose that T is continuous at 0. For $\varepsilon > 0$ there is $\delta > 0$ so that $|x| < \delta$ implies $|Tx| < \varepsilon$. For given $x \ne 0$,

$$\left|\frac{\delta}{2|x|} x\right| < \delta$$

and so

$$\left|T \frac{\delta}{2|x|} x\right| < \varepsilon$$

Multiplying out and using the linearity,

$$|Tx| < \frac{2\varepsilon}{\delta} |x|$$

giving the boundedness.

Finally, prove that boundedness implies continuity at 0. Suppose there is C such that $|Tx| < C|x|$ for all x. Then, given $\varepsilon > 0$, for $|x| < \varepsilon/C$

$$|Tx| < C|x| < C \cdot \frac{\varepsilon}{C} = \varepsilon$$

which is continuity at 0. ///

[2] Another traditional notation for the collection of continuous linear maps from X to Y is $B(X,Y)$, where B stands for *bounded*.

The space $\mathrm{Hom}^o(X,Y)$ of continuous linear maps from one normed space X to another normed space Y has a natural structure of k-vectorspace by

$$(\alpha T)(x) = \alpha \cdot (Tx)$$

$$(S+T)x = Sx + Tx$$

for $\alpha \in k$, $S,T \in \mathrm{Hom}^o(X,Y)$, and $x \in X$.

Proposition: With the uniform norm, the space $\mathrm{Hom}^o(X,Y)$ of continuous linear operators from a normed space X to a *Banach* space Y is *complete*, even if X itself is not.

Proof: Let $\{T_i\}$ be a Cauchy sequence of continuous linear maps $T : X \longrightarrow Y$. Try defining the limit operator T simply by

$$Tx = \lim_i Tx_i$$

First, we should check that this limit exists. Given $\varepsilon > 0$, take i_o large enough so that for $i,j > i_o$ we have $|T_i - T_j| < \varepsilon$. Then, by the definition of the norm on operators,

$$|T_i x - T_j x| < |x|\varepsilon$$

Thus, the sequence of values $T_i x$ is Cauchy in Y, so has a limit in Y, call it Tx.

We need to prove that $x \longrightarrow Tx$ is *continuous* and *linear*. The arguments required are inevitable. Given $c \in \mathbf{C}$ and $x \in X$, for given $\varepsilon > 0$ choose index i so that for $j > i$ both $|Tx - T_j x| < \varepsilon$ and $|Tcx - T_j cx| < \varepsilon$. Then

$$|Tcx - cTx| \le |Tcx - T_j cx| + |cT_j x - cTx| = |Tcx - T_j cx| + |c| \cdot |T_j x - Tx| < (1+|c|)\varepsilon$$

This is true for every ε, so $Tcx = cTx$. Similarly, given $x, x' \in X$, for $\varepsilon > 0$ choose an index i so that for $j > i$ $|Tx - T_j x| < \varepsilon$ and $|Ty - T_j y| < \varepsilon$ and $|T(x+y) - T_j(x+y)| < \varepsilon$. Then

$$|T(x+y) - Tx - Ty| \le |T(x+y) - T_j(x+y)| + |T_j x - Tx| + |T_j y - Ty| < 3\varepsilon$$

Again, this holds for every ε, so it must be that $T(x+y) = Tx + Ty$.

For continuity: it suffices to show that T is *bounded*. Choose an index i_o so that for $i,j \ge i_o$

$$|T_i - T_j| \le 1$$

This is possible since the sequence of operators is Cauchy. Then for such i,j

$$|T_i - T_j x| \le |x|$$

for all x. Thus, for $i \ge i_o$

$$|T_i x| \le |(T_i - T_{i_o})x| + |T_{i_o} x| \le |x|(1 + |T_{i_o}|)$$

Then, taking a limsup,

$$\limsup_i |T_i x| \le |x|(1 + |T_{i_o}|)$$

This implies that T is bounded, and so is continuous. ///

3. Functionals, duals of normed spaces

In this section we consider an important special case of continuous linear maps between normed spaces, namely continuous linear maps from Banach spaces to the *scalars*. All the assertions here are special cases of those for continuous linear maps to more general Banach spaces, but do deserve special attention.

Let X be a normed vectorspace with norm $|\,|$. A *continuous* k-linear map

$$\lambda : X \longrightarrow k$$

is often called a **(continuous linear) functional** on X. Let

$$X^* = \mathrm{Hom}(X, k)$$

denote the collection of all such functionals.

For *any* k-linear map $\lambda : X \longrightarrow k$ of a normed k-vectorspace to k, define the **norm** $|\lambda|$ by

$$|\lambda| = \sup_{|x| \le 1} |\lambda x|$$

where $|\lambda x|$ is the absolute value of the value $\lambda x \in k$. We explicitly allow the value $+\infty$. Such a linear map λ is called **bounded** if $|\lambda| < +\infty$.

Proposition: For a k-linear map $\lambda : X \longrightarrow k$ from a normed space X to k, the following conditions are equivalent:
- The map λ is *continuous.*
- The map λ is *continuous at one point.*
- The map λ is *bounded.*

The **dual space**

$$X^* = \mathrm{Hom}^o(X, \mathbf{C})$$

of X is the collection of *continuous* linear functionals on X. This dual space has a natural structure of k-vectorspace by

$$(\alpha\lambda)(x) = \alpha \cdot (\lambda x)$$

$$(\lambda + \mu)x = \lambda x + \mu x$$

for $\alpha \in k$, $\lambda, \mu \in X^*$, and $x \in X$. It is easy to check that the norm

$$|\lambda| = \sup_{|x| \le 1} |\lambda x|$$

really is a norm on X^*, in that it meets the conditions

- *Positivity:* $|\lambda| \ge 0$ with equality only if $\lambda = 0$.
- *Homogeneity:* $|\alpha\lambda| = |\alpha| \cdot |\lambda|$ for $\alpha \in k$ and $\lambda \in X^*$. As a special case of the discussion of the uniform norm on linear maps, we have

Corollary: The dual space X^* of a normed space X, with the natural norm, is a Banach space. That is, with respect to the natural norm on continuous functionals, it is *complete.* ///

---

4. Banach-Steinhaus (uniform boundedness) theorem

Now we come to some non-trivial results, *non-trivial* in the sense that they use the *Baire category theorem.*

Theorem: *(Banach-Steinhaus/uniform boundedness)* For each α in an index set A, let $T_\alpha : X \longrightarrow Y$ be a continuous linear map from a Banach space X to a normed space Y. Then *either* there is a uniform bound $M < \infty$ so that $|T_\alpha| \le M$ for all $\alpha \in A$, or there is $x \in X$ such that

$$\sup_{\alpha \in A} |T_\alpha x| = +\infty$$

In fact, there is a dense G_δ of such x.

Proof: Let $p(x) = \sup_\alpha |T_\alpha x|$. Being the sup of continuous functions, p is *lower semi-continuous*: for each integer n, the set $U_n = \{x : p(x) > n\}$ is *open*.

On one hand, if each of these is dense in X, then by Baire category the intersection is dense, and by definition is a dense (hence, non-empty) G_δ. On that set p is $+\infty$.

If some one of the U_n should fail to be dense, then by definition there would be a ball B of radius $r > 0$ about some point x_o which failed to meet U_n. Then for $|x - x_o| < r$ and for all α

$$|T_\alpha(x - x_o)| \le |T_\alpha x| + |T_\alpha x_o| \le 2n$$

Thus, as $x - x_o$ varies over the open ball of radius r the vector $x' = (x - x_o)/r$ varies over the open ball of radius 1, and we have

$$|T_\alpha x'| = |T_\alpha \frac{(x - x_o)}{r}| \le 2n/r$$

From this it follows that $|T_\alpha| \le 2n/r$, which is the uniform boundedness. ///

---

5. Open mapping theorem

The open mapping theorem is also non-trivial, insofar as it invokes the Baire category theorem.

Theorem: *(open mapping)* Let $T : X \longrightarrow Y$ be a continuous linear *surjective* map of Banach spaces. Then there is $\delta > 0$ such that for all $y \in Y$ with $|y| < \delta$ there is $x \in X$ with $|x| \le 1$ such that $Tx = y$. In particular, T is an *open map*.

Corollary: If $T : X \longrightarrow Y$ is a *bijective* continuous linear map of Banach spaces, then T is a homeomorphism (so is an *isomorphism*). ///

Proof: Note that the in the corollary the non-trivial point is that T is open, which is the point of the theorem. The linearity of the inverse is easy.

For every $y \in Y$ there is $x \in X$ so that $Tx = y$. For some integer n we have $n > |x|$, so Y is the union of the sets $TB(n)$, where

$$B(n) = \{x \in X : |x| < n\}$$

are the usual open balls. By Baire category, the *closure* of some one of the sets $TB(n)$ contains a non-empty open ball

$$V = \{y \in Y : |y - y_o| < r\}$$

for some $r > 0$ and $y_o \in Y$. Since we are in a metric space, the conclusion is that every point of V occurs as the limit of a Cauchy sequence consisting of elements from $TB(n)$.

Certainly

$$\{y \in Y : |y| < r\} \subset \{y_1 - y_2 : y_1, y_2 \in V\}$$

Thus, every point in the ball B'_r of radius r centered at 0 in Y is the sum of two limits of Cauchy sequences from $TB(n)$. Thus, surely every point in B'_r is the limit of a single Cauchy sequence from the image $TB(2n)$ of the open ball $B(2n)$ of twice the radius. That is, the *closure* of $TB(2n)$ contains the ball B'_r.

Using the linearity of T, the *closure* of $TB(\rho)$ contains the ball $B'_{r\rho/2n}$ in Y.

Then, given $|y| < 1$, choose $x_1 \in B(2n/r)$ so that $|y - Tx_1| < \varepsilon$. Then choose $x_2 \in B(\varepsilon \cdot \frac{2n}{r})$ so that

$$|(y - Tx_1) - Tx_2| < \varepsilon/2$$

Choose $x_3 \in B(\frac{\varepsilon}{2} \cdot \frac{2n}{r})$ so that

$$|(y - Tx_1 - Tx_2) - Tx_3| < \varepsilon/2^2$$

Choose $x_4 \in B(\frac{\varepsilon}{2^2} \cdot \frac{2n}{r})$ so that

$$|(y - Tx_1 - Tx_2 - Tx_3) - Tx_4| < \varepsilon/2^3$$

and so on. Then the sequence

$$x_1,\ x_1 + x_2,\ x_1 + x_2 + x_3,\ \ldots$$

is Cauchy in X. Since X is complete, the limit x of this sequence exists in X, and $Tx = y$. We find that

$$x \in B(2n/r) + B(\varepsilon\frac{2n}{r}) + B(\frac{\varepsilon}{2} \cdot \frac{2n}{r}) + B(\frac{\varepsilon}{2^2} \cdot \frac{2n}{r}) + \ldots \subset B((1+\varepsilon)\frac{2n}{r})$$

This holds for every $\varepsilon > 0$. That is, for every $\varepsilon > 0$

$$TB((1+\varepsilon)\frac{2n}{r}) \supset \{y \in Y : |y| < 1\}$$

Using linearity, for every $\varepsilon > 0$,

$$TB(2n/r) \supset \{y \in Y : |y| < (1+\varepsilon)^{-1}\}$$

Now

$$\bigcup_{\varepsilon} > 0\{y \in Y : |y| < (1+\varepsilon)^{-1}\} = \{y \in Y : |y| < 1\}$$

Thus, we have proven that

$$TB(2n/r) \supset \{y \in Y : |y| < 1\}$$

or, by linearity, that

$$TB(1)\{y \in Y : |y| < r/2n\}$$

as desired. ///

---

6. Hahn-Banach Theorem

For this result, *completeness* is not used at all. Rather, the salient issue is *convexity*, and we definitely need the scalars to be either **R** or **C**. Indeed, the Hahn-Banach theorem seems to be a result about *real* vectorspaces. Note that a **C**-vectorspace may immediately be considered as a **R**-vectorspace simply by forgetting some of the structure.

For Y a vector subspace of X, and for $S : Y \longrightarrow Z$ a linear map to another vectorspace Z, say that a linear map $T : X \longrightarrow Z$ is an **extension** of S to X if the restriction $T|_Y$ of T to Y is S.

Theorem: *(Hahn-Banach)* Let X be a normed vectorspace with scalars **R** or **C**. Let Y be a subspace. Let λ be a bounded linear functional on Y. Then there is an extension Λ of λ to X such that

$$|\Lambda| = |\lambda|$$

Corollary: Given $x \neq y$ both in a normed space X, neither a scalar multiple of the other, there is a continuous linear functional λ on X so that $\lambda x = 1$ while $\lambda y = 0$. ///

Corollary: Let Y be a closed subspace of a normed space X, and take $x_o \notin Y$. Then there is a continuous linear functional λ on X which is 0 on Y, has $|\lambda| = 1$, and $\lambda(x_o) = |x_o|$. ///

Proof: We first treat the case that the scalars are $\mathbf{R}$, and then reduce the complex case to this.

The first and most critical part is to see how to extend a linear functional by just one dimension. That is, for given $x_o \notin Y$ make an extension λ' of λ to $Y' = Y + \mathbf{R}x_o$. Every vector in Y' has a unique expression as $y + cx_o$ with $c \in \mathbf{R}$, so we can define functionals by

$$\mu(y + cx_o) = \lambda y + c\ell$$

for any $\ell \in \mathbf{R}$. The issue is to choose ℓ so that $|\mu| = |\lambda|$.

Certainly $\lambda = 0$ is extendable by $\Lambda = 0$, so we only consider the case that $|\lambda| \neq 0$. In that case we can divide by $|\lambda|$ so as to suppose without loss of generality that $|\lambda| = 1$.

The condition $|\mu| = |\lambda|$ is

$$|\lambda y + c\ell| \leq |y + cx_o|$$

for every $\ell \in \mathbf{R}$ and for every $y \in Y$. We have simplified to the situation that we know this *does* hold for $c = 0$. So for $c \neq 0$, divide through by $|c|$ and replace $y \in Y$ by cy, so that the condition becomes

$$|\lambda y + \ell| \leq |y + x_o|$$

Replacing y by $-y$, the condition is that

$$|\ell - \lambda y| \leq |y - x_o|$$

for every $y \in Y$.

View the collection of all these conditions as a family of conditions upon ℓ. For given $y \in Y$, the condition is that

$$\lambda y - |y - x_o| \leq \ell \leq \lambda y + |y - x_o|$$

For these conditions to have a common solution ℓ, it is exactly necessary that *every* lower bound be less than *every* upper bound. To see that this is so, start from

$$\lambda y_1 - \lambda y_2 = \lambda(y_1 - y_2) \leq |\lambda(y_1 - y_2)| \leq |y_1 - y_2| \leq |y_1 - x_o| + |y_2 - x_o|$$

by the triangle inequality. Then, subtracting $|y_1 - x_o|$ from both sides and adding λy_2 to both sides, we have

$$\lambda y_1 - |y_1 - x_o| \leq \lambda y_2 + |y_2 - x_o|$$

as desired. That is, we have proven the existence of at least one extension from Y to $Y' = Y + \mathbf{R}x_o$ (and with the same norm).

Now invoke an equivalent of the Axiom of Choice to prove that we can extend to the *whole* space (while preserving the norm). Consider the set of pairs (Z, ζ) where Z is a subspace containing Y and ζ is a continuous linear functional on Z extending λ and with $|\zeta| \leq 1$. *Order* these by writing

$$(Z, \zeta) \leq (Z', \zeta')$$

when $Z \subset Z'$ and ζ' extends ζ. For a totally ordered collection (Z_α, ζ_α) of such,

$$Z' = \bigcup_\alpha Z_\alpha$$

is a subspace of X. In general, of course, the union of a family of subspaces would not be a subspace, but these are *nested*.

Then we obtain a continuous linear functional ζ' on this union Z', extending λ and with $|\zeta'| \leq 1$, as follows. The idea is that any *finite* batch of elements already occur inside some Z_α. Given $z \in Z'$, let α be any index large enough so that $z \in Z_\alpha$. Then put

$$\zeta'(z) = \zeta_\alpha(z)$$

Since the family is totally ordered, the choice of α does not matter so long as it is sufficiently large. Certainly for $c \in \mathbf{R}$ we have

$$\zeta'(cz) = \zeta_\alpha(cz) = c\zeta_\alpha(z) = c\zeta'(z)$$

And, given z_1 and z_2 choose α large enough so that *both* z_1 and z_2 are in Z_α. Then

$$\zeta'(z_1 + z_2) = \zeta_\alpha(z_1 + z_2) = \zeta_\alpha(z_1) + \zeta_\alpha(z_2) = \zeta'(z_1) + \zeta'(z_2)$$

proving linearity.

Thus, there is a *maximal* pair (Z', ζ'). Then the earlier argument shows that Z' must be all of X, since otherwise we could construct a further extension, contradicting the maximality. *This completes the proof for the case that the scalars are the real numbers.*

Now reduce the complex case to the real case. First, there is a trick: let λ_o be a merely *real*-linear real-valued functional, and let

$$\lambda x = \lambda_o(x) - i\lambda(ix)$$

Then λ is *complex*-linear, and has the *same norm* as λ_o. (We postpone the proof of this little fact until the end). In particular, when

$$\lambda_o(x) = \Re\lambda(x) = \frac{\lambda x + \overline{\lambda x)}}{2}$$

is the real part of λ then one can see that we recover λ itself by this formula.

Then, given λ on a complex subspace, we take its real part λ_o, a real-linear functional. Extend λ_o to a real-linear functional Λ_o with the same norm. Then let

$$\Lambda x = \Lambda_o(x) - i\Lambda(ix)$$

This is the desired extension, proving the theorem in the complex case by reducing it to the real case.

Last, consider the construction

$$\lambda x = \lambda_o(x) - i\lambda(ix)$$

Since $\lambda_o(x + y) = \lambda_o x + \lambda_o y$ it follows that λ also has this additivity property. Now let a, b be real, and consider $\lambda((a + bi)x)$. We have

$$\lambda((a+bi)x) = \lambda_o((a+bi)x) - i\lambda_o(i(a+bi)x) = \lambda_o(ax) + \lambda_o(ibx) - i\lambda_o(iax) - i\lambda_o(-bx)$$
$$= a\lambda_o x + b\lambda_o(ix) - ia\lambda_o(ix) + ib\lambda_o x = (a+bi)\lambda_o x - i(a+bi)\lambda_o(ix) = (a+bi)\lambda(x)$$

This gives the linearity.

Regarding the norm: since λ_o is real-valued, always

$$|\lambda_o(x)| \leq \sqrt{\lambda_o(x)^2 + \lambda_o(ix)^2} = |\lambda x|$$

On the other hand, given x there is a complex number μ of absolute value 1 so that $\mu\lambda(x) = |\lambda x|$. And note that

$$\lambda_o(x) = \lambda(x) + \overline{\lambda(x)}2$$

Then

$$|\lambda(x)| = \mu\lambda(x) = \lambda(\mu x) = \lambda_o(\mu x) - i\lambda_o(i\mu x)$$

Since the left-hand side is real, and since λ_o is real-valued, it must be that $\lambda_o(\mu x) = 0$. Thus,

$$|\lambda(x)| = \lambda_o(\mu x)$$

Since $|\mu x| = |x|$, we have equality of norms of the functionals λ_o and λ. This completes the justification of the little trick used to reduce the complex case to the real case. ///

.