NORMS, THE DUAL AND CONTINUITY

LARRY SUSANKA

1. The Dual and the Continuity Problem

Suppose $L\colon V \to W$ is linear, where V and W are two vector spaces over field $\mathbb{F}$. We say, synonymously, that L is a (vector space) homomorphism and denote the set of all such with domain V and range W by

$$\mathbf{Hom}_{\mathbb{F}}(\mathbf{V}, \mathbf{W}).$$

This collection of functions is itself a vector space over $\mathbb{F}$.

A homomorphism that is one-to-one and onto its range (that is, an *invertible* homomorphism) is called an **isomorphism**. V and W are called **isomorphic** vector spaces if there is an isomorphism between them, and the language is usually used if we are thinking about **identifying** V and W for some reason.

You shouldn't be too casual about this: any two vector spaces of the same dimension can be identified in this way, but in context it doesn't always make sense to do so.

$Hom_{\mathbb{F}}(V, \mathbb{F})$ is also denoted $\mathbf{V}^*$ and called the **algebraic dual** of V. Members of V^* are called **functionals** (on V).

Suppose $\mathcal{B} = \{\, b_i \mid i \in I \,\}$ is an indexed basis for V.

By this we mean that, as a set, $\mathcal{B}$ is linearly independent and spans V, and also $b_i = b_j$ implies $i = j$.

The cardinality of the index set I is the dimension of V.

Every vector space has an indexed basis. Trivially, a basis could be its own index set. But usually an index set has some kind of order on it, such as the natural numbers.

Each member $v \in V$ can be written as

$$v = \sum_{i \in I} \phi^i(v)\, b_i$$

where the $\phi^i(v)$ are uniquely defined numbers, and only *finitely* many of the $\phi^i(v)$ are nonzero for each v.

Every linear function L as above is completely determined by its effect on a basis:

$$L(v) = F\left(\sum_{i \in I} \phi^i(v)\, b_i\right) = \sum_{i \in I} \phi^i(v)\, F(b_i)$$

Date: May 19, 2013.

and if you are trying to create a linear function you can prescribe the values $F(b_i)$ at will in W and then extend by linearity to the unique linear function defined on all of V with these values on the basis members.

When there is a notion of convergence on each vector space, we will see that (in the infinite dimensional setting) many of these assignments yield homomorphisms which fail to possess the virtue of continuity.

Each ϕ^i is linear on V and so a member of V^*. These functions are called the **coordinate functionals** for basis $\mathcal{B}$. It is important to remember that each individual ϕ^i depends on the entire basis $\mathcal{B}$, and not just on the basis vector b_i.

It is easy to show that $\phi^i = \phi^j$ exactly when $i = j$.

Also, the set $\boldsymbol{\mathcal{B}}^* = \{ \phi^i \mid i \in I \}$ is linearly independent in V^*.

So when I is finite the set $\mathcal{B}^*$ of functionals is an indexed basis for V^*, which has the same dimension as V in that case.

However, in the infinite dimensional case $\mathcal{B}^*$ does not span V^*, an important distinction which we explore with the following example.

Let's consider the case of the subspace P of $\mathbb{R}^{\mathbb{N}}$ consisting of those sequences nonzero at only finitely many integers. We might represent a member p of P by the symbols

$$p = (\ p^0, p^1, p^2, \dots\ )$$

but note that for each such p the listing has a last nonzero member: all entries are zero after a certain point.

P has indexed basis $\mathcal{B} = \{ b_i \mid i \in \mathbb{N} \}$ where $b_i^n = 1$ if $i = n$ and $b_i^n = 0$ if $i \neq n$. In "row" format (it should really be a column, but I plead typographical convenience here) vector b_i has a 1 in the ith slot, and zeroes elsewhere.

$$b_i = (\, 0, \dots, 0, 1, 0, 0, 0, \dots\ )$$

Every $p \in P$ can be written as the *finite* sum (include at most one zero term in the summation)

$$p = (\ p^0, p^1, p^2, \dots\ ) = \sum_{i \in \mathbb{N}} p^i\, b_i.$$

So for each $i \in \mathbb{N}$ the coordinate functional ϕ^i is given by

$$\phi^i(p) = \phi^i(\ (\ p^0, p^1, p^2, \dots\ )\ ) = p^i.$$

ϕ^i "picks off" the ith entry of p.

Thus $p = \sum_{i \in \mathbb{N}} \phi^i(p)\, b_i$ which looks even more compact if you adopt the Einstein summation convention.

So we define $L \colon P \to P$ by

$$L(p) = \sum_{i \in \mathbb{N}} (i+1)\phi^{i+1}(p)\, b_i$$

and the related $M \colon P \to \mathbb{R}$

$$M(p) = \sum_{i \in \mathbb{N}} (i+1)\phi^{i+1}(p).$$

It is easy to show that L and M are linear, so $M \in P^*$.

These two functions are defined in terms of all the members of $\mathcal{B}^*$ except ϕ^0 and *neither one* can be written as a sum involving only a finite number of the ϕ^i.

In particular $\mathcal{B}^*$ is not a basis for the algebraic dual P^*.

This example also allows us to illustrate in a useful way the concepts of isomorphism and identification.

It is clear that the set of polynomial functions defined, say, on the interval between 0 and 10, is a real vector space. The association

$$p^0 + p^1 x + p^2 x^2 + \cdots + p^n x^n \longleftrightarrow (\, p^0, p^1 \ldots , p^n, 0, 0, 0, \ldots\,)$$

is an isomorphism between the polynomial space $Poly$ and our sequence subspace P.

With this identification, the map $\frac{d}{dx}\colon Poly \to Poly$ is associated with $L\colon P \to P$.

The map $M\colon P \to \mathbb{R}$ is associated with $\frac{d}{dx}\Big|_{x=1}\colon Poly \to \mathbb{R}$.

Often we will have a metric on our vector space. During our discussion about metrics we learned about the supremum metric. Applied to P it is given as

$$d(p,q) = Max\left\{\, |p^i - q^i| \;\middle|\; i \in \mathbb{N} \,\right\} \qquad \text{for } p = \sum_{i\in\mathbb{N}} p^i b_i \text{ and } q = \sum_{i\in\mathbb{N}} q^i b_i \text{ in } P.$$

The sequence of vectors $A_n = \frac{b_{n+1}}{n+1}$ converges to the zero vector in this metric.

But $L(A_n) = L\left(\frac{b_{n+1}}{n+1}\right) = \frac{n+1}{n+1} b_n = b_n$ does not converge to the zero vector in P.

And $M(A_n) = 1$ for all n and so does not converge to the number 0 in $\mathbb{R}$.

Differentiation is *not* continuous on polynomials with respect to the metric determined by maximum coefficient size which is, of course, the only obvious notion of polynomial size at hand: two polynomials would seem to be close exactly when all their coefficients are close.

This example shows that in the infinite dimensional case with a metric there can be functionals in the algebraic dual V^* that are not continuous.

The subset of the algebraic dual consisting of the continuous functionals will be denoted $\mathbf{V}'$, the **continuous dual** of V.

This is metric dependent: different metrics give different continuous functionals!

We will explore the issues involved with the problem of continuity and featured in our example above. The rather sparse nature of P (polynomials are hardly the only functions we might want to differentiate) is an issue.

How many of the useful tools of Linear Algebra and $\mathbb{R}^n$ (eigenvalues, eigenvectors and so on) can we recover in some form or another? In particular we need to examine different metrics on P: continuity for us is a metric concept, and changing the definition of convergence using new metrics might help.

We have one final definition which needs to be made somewhere, so I'm putting it here.

The algebraic dual V^* of V is, itself, a vector space. So *it too* has an algebraic dual, $(V^*)^*$.

Define the **evaluation map** $\mathbf{E}\colon \mathbf{V} \to (\mathbf{V}^*)^*$ by

$$E(x)(f) = f(x) \quad \text{for each } x \in V \text{ and } f \in V^*.$$

E is linear, and also one-to-one: that is, $E(x) = E(y)$ exactly when $x = y$. So if E is onto then it is invertible and an isomorphism. In that case, $(V^*)^*$ can be identified with (i.e. it "is") V.

If V is finite dimensional, V and V^* have the same dimension. And it follows that $(V^*)^*$ has the same dimension as does V. So E must be onto. The infinite dimensional case is much more delicate and we will consider the extent to which we can recover this important identification in some form later.

2. The Hahn-Banach Theorem

This section is devoted to the possibility of extending a function with certain properties to a larger domain while preserving those properties.

One reason we want this theorem is that it allows us to conclude that there is a rich stock of continuous functionals whenever the theorem applies.

A nonempty subset S of a real vector space V is called **convex** if

$$tu + (1-t)v \in S \quad \forall t \in [0,1] \text{ and } u, v \in S.$$

In other words, all points on the line segment connecting u and v are in S whenever u and v are in S.

If V is any real vector space we say that a function $P\colon V \to \mathbb{R}$ is **convex** provided

$$P(tu + (1-t)v) \leq tP(u) + (1-t)P(v) \quad \forall t \in [0,1].$$

Geometrically, and in case $V = \mathbb{R}$, this means that the graph of a convex function always lies on or beneath the straight line connecting any two points on the graph. For this reason convex functions are also called **sublinear.**

Any seminorm (see the next section) is convex: a seminorm is the most common source of convex functions.

Theorem 2.1. The Hahn-Banach Theorem

If $Y \overset{\substack{\textit{Real Vector}\\ \textit{Subspace}}}{\subset} X \overset{\substack{\textit{Real Vector}\\ \textit{Space}}}{}$ *and* $P\colon X \to \mathbb{R}$ *is convex*

and $\Lambda \in Y^*$ *satisfies* $\Lambda \leq P|_Y$

then $\exists \Psi \in X^*$ *with* $\Lambda = \Psi|_Y$ *and* $\Psi \leq P$.

Proof. If $w \in X - Y$ (that is, w is in X but *not* in Y) and α, β are positive and we select vectors u and v in Y

$$\begin{aligned}\beta\, \Lambda u + \alpha\, \Lambda v \;&=\; (\alpha+\beta)\,\Lambda\left(\frac{\beta}{\alpha+\beta}u + \frac{\alpha}{\alpha+\beta}v\right)\\ &\leq (\alpha+\beta)\,P\left(\frac{\beta}{\alpha+\beta}(u-\alpha w) + \frac{\alpha}{\alpha+\beta}(v+\beta w)\right)\\ &\leq \beta\,P(u-\alpha w) + \alpha\,P(v+\beta w).\end{aligned}$$

$$\text{So} \quad \frac{1}{\alpha}\,[\ \Lambda u - P(u - \alpha\, w)\ ] \;\leq\; \frac{1}{\beta}\,[\ P(v + \beta\, w) - \Lambda v\ ].$$

The left side does not depend on v or β, while the right is independent of α and u. So there is a real number a with

$$\sup_{\substack{u\in Y\\ \alpha>0}} \frac{1}{\alpha}\,[\ \Lambda u - P(u-\alpha\, w)\ ] \ \le\ a\ \le\ \inf_{\substack{v\in Y\\ \beta>0}} \frac{1}{\beta}\,[\ P(v+\beta\, w) - \Lambda v\ ].$$

Define $\Theta\colon Y\bigoplus \mathbb{R}w \to \mathbb{R}$ by $\Theta(v+rw) = \Lambda v + ra$ for each $r\in\mathbb{R}$ and $v\in Y$. Considering the cases of r positive, negative or zero separately, the definition of a yields

$$\Theta(v+rw) = \Lambda v + ra \le \Lambda v + P(v+rw) - \Lambda v = P(v+rw).$$

So any functional such as Λ above, defined on a proper subspace and dominated by P, can be extended to a larger space while preserving its relationship with P.

Let S be the set of all linear extensions of Λ to subspaces of X which are dominated by P on their domain. Partially order this set of extensions by $\Theta \le \Psi$ if Ψ is an extension of Θ. Chains in S have upper bounds in S and we invoke Zorn's lemma and assert that there is a maximal member Ψ of S. The domain of Ψ is X, else it could be extended by one dimension, contradicting maximality. □

When I first encountered this result, my instinct was to say its proof should be obvious: fatten a basis of Y to a basis of X and simply define Ψ to be 0 on those members of this basis not in Y, extending by linearity. Unfortunately this naive attempt fails: an inspection of the proof shows how a value of 0 on certain vectors may not be possible for Ψ. The point here is that there always *is* a value that is consistent with the other values of this function subject to domination by P.

Corollary* 2.2. *The Hahn-Banach Theorem

If $Y \overset{\text{Complex Vector}}{\underset{\text{Subspace}}{\subset}} X \ {}^{\text{Complex Vector}}_{\text{Space}}$ *and* $P\colon X\to\mathbb{R}$ *satisfies*

$P(\alpha v + \beta u) \le |\alpha| P(v) + |\beta| P(u)$ *if* $u, v\in X$ *and* $|\alpha| + |\beta| = 1$

and if $\Lambda \in Y^*_{\mathbb{C}}$ *satisfies* $|\Lambda| \le P|_Y$

then $\exists \Psi\in X^*_{\mathbb{C}}$ *with* $\Lambda = \Psi|_Y$ *and* $|\Psi| \le P$.

Proof. Let L be the real part of Λ, thought of as a real linear functional. $\forall y\in Y,\ Ly \le |\Lambda y| \le P(y)$. Also, for real positive constants α and β the condition on P in the statement of this corollary reduces to convexity. So Theorem 2.1 applies: $\exists$ real linear $M\colon X\to\mathbb{R}$ extending L and with $Mx \le P(x)\ \forall x\in X$.

Let $\Psi x = Mx - iM(ix)\ \forall x\in X$. Check that $\Psi(ix) = i\Psi x$, so Ψ is a complex linear functional and extends Λ to all of X. It remains only to show that $\Psi\le P$.

Pick $x\in X$. Find angle θ so that $\Psi x = |\Psi x| \mathrm{e}^{i\theta}$.

$$\begin{aligned}\text{Then } |\Psi x| \ = \ (\Psi x)\mathrm{e}^{-i\theta} &= M\left(\mathrm{e}^{-i\theta}x\right) - iM\left(i\mathrm{e}^{-i\theta}x\right)\\ &= M\left(\mathrm{e}^{-i\theta}x\right) \quad \text{(the complex part must be zero)}\\ &\le P\left(\mathrm{e}^{-i\theta}x\right) \ \le \left|\mathrm{e}^{-i\theta}\right| P(x) = P(x).\end{aligned}$$

□

3. Continuity and Normed Spaces

The spaces we consider will be more than just vector spaces: they will all have a metric (and hence concepts of convergence, continuity and compactness) and in this section the metric will come from a norm.

NLS1 If you want to do calculus in your space you must have limits, and the easiest way to talk about limits in a vector space is through the explicit notion of distance provided through a norm.

If V is a vector space over $\mathbb{R}$ or $\mathbb{C}$, a **seminorm** on V is a function

$$\|\cdot\|\colon V \to [0,\infty)$$

with the property that for any number k and vectors v and w

$$\|kv\| = |k|\,\|v\|$$
$$\text{and} \quad \|\, v + w \,\| \le \|v\| + \|w\|$$

The seminorm is called **homogeneous** by virtue of the first line. The second of these is called the **triangle inequality**. The triangle inequality can be tweaked slightly to produce a lower limit for the norm of a sum too.

$$|\; \|v\| - \|w\| \;| \le \|\, v + w \,\| \le \|v\| + \|w\|.$$

A **seminormed linear space**, abbreviated **SNLS**, is a real or complex vector space endowed with a seminorm.

A seminorm satisfies

$$\|\, \alpha v + \beta u \,\| \le |\alpha|\,\|\, v \,\| + |\beta|\,\|\, u \,\| \quad \text{if } \; u, v \in X \text{ and } \alpha, \beta \in \mathbb{F}.$$

So a seminorm is an example (the most important example) of a sublinear function as found in the statement of Theorem 2.1 and its Corollary.

If you add the condition

$$\|v\| = 0 \;\text{ when and only when }\; v = 0$$

the seminorm is called a **norm**.

If G is any member of V^*, the map $|G|\colon V \to [0,\infty)$ given as $|G|(v) = |G(v)|$ is a seminorm, and a common source of them too. This seminorm can never be a norm unless V has dimension 1.

A **normed linear space**, abbreviated **NLS**, is a real or complex vector space endowed with a norm.

NLS2 The **distance** between vectors v and w in a SNLS V is defined by

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

The distance notion is a pseudometric on V, and is a particularly nice one, having the properties

$d(x+z, y+z) = d(x,y)\ \forall x, y, z \in V$ (**translation invariance**)

$d(a\,x, a\,y) = |a|\,d(x,y)\ \forall x, y \in V$ and $a \in \mathbb{F}$ (**homogeneity**)

not required of a general metric or pseudometric.

Whenever a notion of distance is used in an SNLS it is this pseudometric which will be intended.

This pseudometric is a metric exactly when the seminorm is a norm.

NLS3 If a sequence v_0, v_1, $v_2, \ldots$ converges in a SNLS using this pseudometric we say that the sequence **converges in seminorm (or norm).** Sometimes this is also called **strong convergence**, particularly when we have a norm from an inner product. (There is a weaker concept of convergence which is also useful in inner product spaces.) In case there might be confusion about the type of convergence involved, we might indicate intent by

$$v_i \underset{\text{strong}}{\longrightarrow} w.$$

If V is an SNLS the set $\mathbf{N} = \{\, x \in V \mid \|x\| = 0 \,\}$ is a vector subspace of V, sometimes called the set of **null vectors** for the seminorm (not to be confused with vectors from the nullspace of a linear transformation.)

In an SNLS a sequence can converge to more than one point. In fact, if $v_i \to w$ then $v_i \to x$ exactly when $x = w + y$ for some $y \in \mathbf{N}$.

This implies that $\mathbf{N}$ is a closed set.

The seminorm is a norm exactly when $\{\, 0 \,\}$ is closed.

Completeness is a very important property for us. Normed linear spaces which are complete are called **Banach spaces.**

$\mathbb{R}$ and $\mathbb{C}$ are themselves Banach spaces, a critical fact that is used often and assumed without discussion in most first-year calculus classes.

In fact, $\mathbb{R}$ is usually built using the linear order and distance notion in the rational numbers $\mathbb{Q}$ by adjoining elements to $\mathbb{Q}$ corresponding to all the "holes" in $\mathbb{Q}$. So $\mathbb{R}$ is ordered too, and complete by its very definition.

If S is a vector subspace of V then $\overline{S}$ is also a SNLS space, a subspace of V. If V is Banach, so is $\overline{S}$.

NLS4 For V and W both SNLSs, a function $F\colon V \to W$ is continuous at $p \in V$ exactly when

$$\forall \varepsilon > 0\ \exists \delta > 0 \text{ so that if } \|\, p - v \,\| < \delta \text{ then } \|\, F(p) - F(v) \,\| < \varepsilon.$$

In general, δ will depend on both ε and p, but for linear functions it does not. Using linearity, one shows that continuity at *any point*, that is for just one v, implies continuity at *every point*, and the δ chosen for a particular ε does not depend on *which* point.

Linear functions between two SNLSs
which are continuous at a point are uniformly continuous.

In fact if you can find $\delta_1 > 0$ for $\varepsilon = 1$ and $p = 0$ then F is continuous. That is because for any $\varepsilon > 0$ we have

$$\|v\| < \delta_1 \varepsilon \;\Rightarrow\; \frac{\|v\|}{\varepsilon} < \delta_1 \Rightarrow \left\| F\left(\frac{v}{\varepsilon}\right) \right\| < 1 \Rightarrow \frac{1}{\varepsilon}\, \|F(v)\| < 1$$

and so $\|F(v)\| < \varepsilon$. Linearity then allows us to translate the argument away from the origin.

A useful equivalent condition, again using linearity of F, is that if $F^{-1}(B)$ is open for even *one open ball* B then $F^{-1}(B)$ will be open for *every* open subset B contained in W and therefore F will be continuous.

For continuous linear F, the set $\mathcal{K}er(F)$ is closed.

NLS5 If $\mathbf{N}$ is the set of null vectors for SNLS V with seminorm $\| \cdot \|$ we can form the quotient space $V/\mathbf{N}$. The function $\| \cdot \|'$ defined on $V/\mathbf{N}$ by

$$\| \, v + \mathbf{N}, \|' = \| \, v \, \|$$

is a norm on $V/\mathbf{N}$.

Further, if $F^{\text{Linear}} \colon V^{\text{SNLS}} \to W^{\text{SNLS}}$ and if $\mathbf{N} \subset \mathcal{K}er(F)$ the function

$$\widetilde{F} \colon (V/\mathbf{N})^{\text{NLS}} \to W^{\text{SNLS}} \quad \text{given by} \quad \widetilde{F}(v + \mathbf{N}) = F(v)$$

is well-defined and linear. $\widetilde{F}$ is continuous when F is continuous.

NLS6 If V and W are two SNLSs with seminorms $\| \cdot \|_1$ and $\| \cdot \|_2$, respectively, the product vector space $V \times W$ can be given the **product seminorm** defined by

$$\|(u, v)\| = \|u\|_1 + \|v\|_2.$$

This function is, in fact, a seminorm on the product space, and a norm exactly when both $\| \cdot \|_1$ and $\| \cdot \|_2$ are norms.

NLS7 A function $F^{\text{Linear}} \colon V^{\text{NLS}} \to W^{\text{NLS}}$ is an **isometry** provided that $\| \, F(x) \, \| = \| \, x \, \| \; \forall x \in V$. Isometries are continuous.

Also, an isometry is automatically one-to-one, so isometry F implements an isomorphism between V and $\mathcal{R}an(F)$. But this is not just any isomorphism: it preserves the notion of distance provided by the norm on the domain.

Using the restriction of the norm on W to $\mathcal{R}an(F)$, the NLS V and the NLS $\mathcal{R}an(F)$ are identified. If F is onto, we have an identification of V and W with norm structure preserved.

NLS8 For $F^{\text{Linear}} \colon V^{\text{SNLS}} \to W^{\text{SNLS}}$ we define $\|\mathbf{F}\|$ by

$$\|F\| = \sup\{ \, \|F(x)\| \mid x \in S_1(0) \, \}.$$

If $\|F\| < \infty$ we say F is **bounded**, and it turns out that $\|F\|$ is bounded exactly when F is continuous. This is important enough that we enshrine the result in:

A linear functions between two SNLSs
is continuous exactly when it is bounded.

$\mathbf{B(V, W)}$ is defined to be the set of bounded linear functions from V to W. The number $\|F\|$ defined for each $F \in B(V, W)$ is a norm on $B(V, W)$, generally called the **operator norm**.

If the range space W is a Banach space, so is $B(V,W)$.

If there are multiple norms floating around, we might distinguish this one by the notation $\|\cdot\|_{op}$.

NLS9 $A^{\text{Linear}}\colon X^{\substack{\text{Banach}\\\text{Space}}} \to Y^{\substack{\text{Banach}\\\text{Space}}}$ is called **compact** exactly when $\overline{A(S)}$ is a compact subset of Y whenever S is a bounded subset of X. If compact, a function A must be bounded and so compact linear functions from one Banach space to another are continuous.

There is an important equivalent condition, which reduces the question of whether or not a mapping is compact or not to an issue involving sequences.

Suppose (x_i) is a bounded sequence in X. Then compactness of A requires $\{\, A(x_i) \mid i \in \mathbb{N}\,\}$ to have compact closure in Y. This implies that the sequence of image points $(A(x_i))$ has a Cauchy subsequence.

Conversely, if this last condition must necessarily hold for any bounded sequence in X then A is a compact mapping.

If a continuous map $A^{\text{Linear}}\colon X^{\substack{\text{Banach}\\\text{Space}}} \to Y^{\substack{\text{Banach}\\\text{Space}}}$ has finite rank, it is clearly compact. It is not too hard to show that the limit (in operator norm) of continuous finite rank maps is compact. It is also true that if Y is any Hilbert space or any Banach space with a Schauder basis, any compact linear function from Y to Y is the limit of continuous finite rank maps, a desirable feature referred to as the **approximation property** for these compact maps.

NLS10 The collection of bounded linear functionals on an NLS V, the continuous dual $V' = B(V,\mathbb{F})$ of V, is a particularly important Banach space. Recall it is a subspace of the algebraic dual, $V^* = Hom_{\mathbb{F}}(V,\mathbb{F})$.

Suppose x is nonzero in NLS V. The linear transformation Λ defined on $\mathbb{F}x$ by $\Lambda(a\,x) = a\,\|x\|$ satisfies the condition of Theorem 2.1 (or its Corollary) where P is the norm on V. In this one dimensional case we have equality:

$$|\,\Lambda(a\,x)\,| = |a|\,\|x\| = \|a\,x\|.$$

So Λ can be extended to linear Ψ defined on all of V and for which

$$|\Psi(v)| \leq \|v\| \quad \text{for all} \quad v \in V.$$

This means that the operator norm $\|\Psi\|$ cannot exceed 1.

But $\Psi(x/\|x\|) = \Lambda(x/\|x\|) = 1$, So $\|\Psi\| = 1$.

To recap, for each $x \in V$ there is a functional Ψ in V' with

$$\|\Psi\| = 1 \quad \text{and} \quad \Psi x = \|x\|.$$

NLS11 Since V' is a NLS, it too has a an algebraic dual $(V')^*$ and continuous dual $V'' \subset (V')^*$.

Recall the evaluation map

$$E\colon V \to (V^*)^* \text{ given by } E(x)(f) = f(x).$$

Every member of $(V^*)^*$ produces a member of $(V')^*$ by restriction and $\widetilde{E}(x)$ defined to be $E(x)|_{V'}$ is such a member.

And if $f \in V'$ and $\|f\| = 1$ then $\left|\widetilde{E}(x)(f)\right| = |f(x)| \leq \|f\|\,\|x\| = \|x\|$.

So each $\widetilde{E}(x)$ is bounded as a function from Banach space V' with operator norm to $\mathbb{F}$. In other words,

$$\widetilde{E}\colon V \to V''.$$

But by the last result of NLS10 there is a functional $\Psi \in V'$ with operator norm 1 for which $|\Psi(x)| = \|x\|$. This means that $\left|\widetilde{E}(x)\right|$ actually attains its maximum value, $\|x\|$, on the members of V' with operator norm 1.

Even more, this means that

$$\widetilde{E}\colon V \to V'' \quad \text{is an isometry.}$$

So the image of $\widetilde{E}$ with operator norm and V itself with its norm are not only isomorphic as vector spaces but are interchangeable in any calculation involving norms as well.

Spaces for which $\widetilde{E}$ is onto V'' are very important, and are called **reflexive**. We will have occasion to refer to this property later. Hilbert spaces are reflexive.

4. Schauder Bases in a Banach Space

The Baire category theorem implies that no complete infinite dimensional space can have a countable basis. However, if we can't have a countable *basis*, we can do almost as well with a Schauder basis, defined below, which uses concepts of limit and continuity provided by a norm to get most of what a true basis provides in the finite dimensional setting.

The ideas to follow make sense in more general settings but we will confine consideration here to Banach spaces.

A **Schauder basis** for Banach X is a countable *ordered* set of vectors $v_0, v_1, \ldots$ for which every member x of X can be written in a *unique* way as

$$x = \sum_{n=0}^{\infty} a^n(x)\, v_n$$

for $a^n(x) \in \mathbb{F}$. The uniqueness refers to the values of the **coordinate functionals** a^n, which are therefore linear, and the convergence of the sequence of partial sums is in norm:

$$\sum_{n=0}^{k} a^n(x)\, v_n \underset{\text{strong}}{\longrightarrow} x \quad \forall x \in X.$$

When referring to a sequence of vectors in vector space X the notation $\mathbf{v} = (\mathbf{v_n}) \subset X$ will be used. Thus, for Schauder basis as above we have the paired sequences

$$\mathbf{v} \subset X \quad \text{with unique associated coordinate functionals} \quad \mathbf{a} \subset X^*.$$

Any Banach space with a Schauder basis is separable, so there are Banach spaces without Schauder bases. In fact, there are Banach spaces for which no infinite dimensional subspace has a Schauder basis. Still, Banach spaces that have these bases are common in practice.

Uniqueness of coefficients implies that 0 is not among the vectors in a Schauder basis, and in fact the vectors in a Schauder basis must constitute a linearly independent list of vectors.

If $(\mathbf{v_n})$ is a Schauder basis, so is $(\mathbf{v_n}/\|\mathbf{v_n}\|)$ and the latter Schauder basis is called **normalized.**

Normalized or not, the members of the sequence of linear functionals $(\mathbf{a^n})$ are all continuous. In fact, (though it takes a bit of work to prove) their norms satisfy

$$1 \leq \|a^n\| \, \|v_n\| \leq K \quad \forall n \in \mathbb{N}$$

for a positive constant K that will vary with the basis.

So for Schauder basis $\mathbf{v} \subset X$ we actually have $\mathbf{a} \subset X'$, not just $\mathbf{a} \subset X^*$.

For each k the functional $P^k = \sum_{n=0}^{k} a^n \, v_n$ is bounded and linear from X onto the finite dimensional subspace of x spanned by $v_0, \ldots, v_k$. It is a **projection** onto that subspace: $P^k \circ P^k = P^k$ and, more generally, if $i \leq j$ then $P^i \circ P^j = P^j \circ P^i = P^i$. The sequence of projections converges to the identity operator in operator norm.

This sequence is used to show that a compact operator K is the limit of finite rank operators in a Banach space with a Schauder basis: examine the sequence $P^i \circ K$.

A Schauder basis is called **unconditional** if, for any $x \in X$ the series obtained by any permutation of the terms in the series representation for x in this Schauder basis also converges to x.

A Schauder basis is called **bounded** if there are positive constants A and B for which

$$A \leq \|v_n\| \leq B \quad \forall n \in \mathbb{N}.$$

Since each vector in a Schauder basis $\mathbf{v}$ can be conceived of as (that is, *it is*) a member of X'' the possibility arises that $\mathbf{a}$ could be a Schauder basis for the Banach space $\overline{Span(\mathbf{a})}$ with operator norm, with coordinate functionals $\mathbf{v}$ (the domain of each vector in $\mathbf{v}$ restricted, of course, to $\overline{Span(\mathbf{a})}$.)

This is, in fact, the case. Even more, we have the following.

Theorem* 4.1. *The Dual Basis Theorem

Suppose $\mathbf{v} \overset{Schauder}{\underset{Basis}{\subset}} X^{Banach}$ *with associated coordinate functionals* $\mathbf{a} \subset X'$.

Then $\mathbf{a} \overset{Schauder}{\underset{Basis}{\subset}} \overline{Span(\mathbf{a})}$ *with associated coordinate functionals* $\mathbf{v}$.

Moreover, if $\mathbf{v}$ *is unconditional or bounded, so is* $\mathbf{a}$.

Proof. See the literature. □

The most interesting case, of course, is when X is reflexive. Then $X' = \overline{Span(\mathbf{a})}$ and so we have a correspondence between Schauder bases with their coefficient sequences for X and those for X'.

5. The Banach Adjoint

Suppose $F \in B\left(V^{\text{Banach}}, W^{\text{Banach}}\right)$.

Define for each $w' \in W'$ the member $F^*(w') \in V^*$ given by

$$F^*(w')(v) = w'(F(v)).$$

Note that $\|F^*(w')(v)\| \leq \|w'\|\,\|F\|\,\|v\|$.

So, in fact

$$\|F^*(w')\| \leq \|w'\|\,\|F\|$$

and this means $F^*(w')$ is actually in V', not just V^*.

Looking again we see that $F^* \in B\left(W', V'\right)$ and $\|F^*\| \leq \|F\|$.

Application of the Hahn-Banach theorem proves more: that $\|F^*\| = \|F\|$.

It is also a fact that if F is an isometry, so is F^*.

The map $F^* \colon W' \to V'$ is called the **Banach adjoint** of F and the **Banach adjoint operator**

$$^* \;:\; B\left(V, W\right) \to B\left(W', V'\right) \quad \text{is an isometry.}$$

Finally, if V and W are reflexive, we note that $F^{**} = F$, and the adjoint operator is an isomorphism.