Infinite Dimensional Vector Spaces: How do you get there and what do you do when you're there?

Robert Gardner

East Tennessee State University
Department of Mathematics
Institute of Mathematical and Physical Sciences

Department of Physics Seminar
February 9, 1998

I. VECTOR SPACES - INTRODUCTION

Note. A vector space consists of two things: *scalars* and *vectors*. We take as our scalar field either the real numbers $\mathbf{R}$ or the complex numbers $\mathbf{C}$.

Example. $n-$dimensional Euclidean space $\mathbf{R}^n$ is a vector space. With $n = 2$ or 3, this yields the familiar idea of vectors as "arrows" which represent position, velocity, or acceleration in introductory physics and engineering classes. In general, elements of $\mathbf{R}^n$ look like $x = (x_1, x_2, \ldots, x_n)$ where each $x_i \in \mathbf{R}$ (we take the scalar field to be $\mathbf{R}$).

Example. $\mathbf{C}^n = \{(z_1, z_2, \ldots, z_n) \mid z_i \in \mathbf{C}\}$ forms a complex vector space (we take the scalar field to be $\mathbf{C}$). Notice that the "arrows" interpretation is more difficult here (at least for $n > 1$).

II. VECTOR SPACES - DIMENSION

Example. The collection of all polynomials of degree n or less forms a vector space of dimension $n+1$ (we can take real or complex polynomials) denoted $\mathcal{P}_n$.

Notice. There is a "natural relationship" between $\mathcal{P}_n$ and $\mathbf{R}^{n+1}$. For Example, we can associate with the polynomial $p(x) = a_0 + a_1x + a_2x^2 \in \mathcal{P}_2$ the element $(a_0, a_1, a_2) \in \mathbf{R}^3$.

Definition. The *span* of a finite set of vectors

$$\{x_1, x_2, \ldots, x_k\}$$

is the collection of all possible linear combinations of the vectors:

$$\text{span}(\{x_1, x_2, \ldots, x_k\}) = \{\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_k x_k \mid$$

$$\alpha_1, \alpha_2, \ldots, \alpha_k \in F(\text{the scalar field})\}.$$

Definition. A set of vectors $\mathcal{B} \subset V$ is a *basis of* V if $\mathcal{B}$ is linearly independent and span $\mathcal{B} = V$. If a vector space has a finite basis, it is *finite dimensional.* Otherwise, it is *infinite dimensional.*

Note. The vector space $\mathbf{R}^3$ has as a basis

$$\{(1,0,0),(0,1,0),(0,0,1)\}.$$

The vector space $\mathcal{P}_2$ has as a basis $\{1, x, x^2\}$. In general, a basis for $\mathbf{R}^n$ is

$$\{(1,0,0,0,\ldots,0),(0,1,0,0,\ldots,0),(0,0,1,0,\ldots,0),$$
$$\ldots,(0,0,0,0,\ldots,0,1)\}.$$

Note. The set of all sequences

$$\{(x_1, x_2, \ldots) \mid x_i \in \mathbf{R}\}$$

forms an infinite dimensional vector space with basis

$$\{(1,0,0,0,\ldots),(0,1,0,0,\ldots),(0,0,1,0,\ldots)\}.$$

(...we have not defined the term "basis" for an infinite dimensional vector space, though!)

III. VECTOR SPACES - ISOMORPHISM

Note/Definition. An *isomorphism* between vector spaces V_1 and V_2, both over the scalar field F, is a function π which maps the vectors of V_1 to the vectors of V_2 such that the operations of vector addition and scalar multiplication are preserved. We say V_1 is *isomorphic* to V_2, denoted $V_1 \cong V_2$. For example, $\mathbf{R}^3 \cong \mathcal{P}_2$. An isomorphism between $\mathbf{R}^3$ and $\mathcal{P}_2$ is the mapping $\pi : \mathbf{R}^3 \to \mathcal{P}_2$ defined as $\pi((a_0, a_1, a_2)) = a_0 + a_1 x + a_2 x^2$.

Theorem. "Fundamental Theorem of Linear Algebra"

An n dimensional vector space over the field $\mathbf{R}$ (or F in general) is isomorphic to $\mathbf{R}^n$ (or F^n in general).

IV. VECTOR SPACES - NORMS

Definition. A real function $\|\cdot\|$ on a vector space H is a *norm* if

(a) $\|x\| \geq 0$ for all $x \in V$ and $\|x\| = 0$ if and only if $x = 0$

(b) $\|\lambda x\| = |\lambda|\|x\|$ for all $x \in V$ and $\lambda \in F$

(c) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in V$ (triangle inequality).

Note. If $\|\cdot\|$ is a norm on a vector space, then $d(x, y) = \|x - y\|$ defines a *metric* on the vector space with which we can measure distance.

Example. A norm on $\mathbf{R}^n$ is

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

where $x = (x_1, x_2, \ldots, x_n) \in \mathbf{R}^n$. This is the *Euclidean norm* and can be used to define the *Euclidean metric* on $\mathbf{R}^n$. Notice that for $n = 1$ this is simply absolute value.

Example. A norm on $\mathbf{C}^n$ is

$$\|z\| = \sqrt{|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2}$$

where $z = (z_1, z_2, \ldots, z_n) \in \mathbf{C}^n$. With $n = 1$ this is just the familiar modulus of a complex number.

V. BANACH SPACES - COMPLETENESS

Note. We now need to explore the difficult subject of completeness. We do so very informally.

Geometric Note. When you hear the term "complete," think "no holes." The rational numbers

$$\mathbf{Q} = \{p/q | \, p, q \in \mathbf{Z}, q \neq 0\}$$

is not a complete vector space (here we take the scalar field to be $\mathbf{Q}$ itself) since the sequence

$$\{1, 1.4, 1.41, 1.414, \ldots\}$$

is Cauchy but does not converge *in this space* (since the limit is $\sqrt{2}$). In some sense, $\mathbf{Q}$ is not complete since it has holes! In particular, it has a hole at $\sqrt{2}$.

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

Note. The real numbers are complete (in fact, this is part of the definiton of $\mathbf{R}$) and so form a Banach space. More generally, $\mathbf{R}^n$ and $\mathbf{C}^n$ form Banach spaces.

Example. The vector space of all square summable sequences of complex numbers

$$l^2 = \left\{ (z_1, z_2, \ldots) \mid z_i \in \mathbf{C} \text{ and } \sum_{i=1}^{\infty} |z_i|^2 < \infty \right\}$$

with the norm

$$\|(z_1, z_2, \ldots)\| = \left(\sum_{i=1}^{\infty} |z_i|^2 \right)^{1/2}$$

is a (very fundamental) Banach space. This is a somewhat difficult result and it is not even clear that this space is closed under addition.

VI. INNER PRODUCTS

Note. We are ultimately interested in generalizing the idea of dot product in $\mathbf{R}^n$ (or $\mathbf{C}^n$) to the setting of infinite dimensional spaces.

Definition. Let V be a vector space over the field of scalars $\mathbf{C}$. A mapping

$$(\cdot,\cdot) : V \times V \to \mathbf{C}$$

is an *inner product* in V if for any $x, y, z \in V$ and $\alpha, \beta \in \mathbf{C}$, the following hold:

(a) $(x, y) = \overline{(y, x)}$ (the bar represents complex conjugate),

(b) $(\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z)$,

(c) $(x, x) \geq 0$ and $(x, x) = 0$ implies $x = 0$.

A vector space with an inner product is an *inner product space* (or *pre-Hilbert space*).

Example. An inner product can be put on the real vector space $\mathbf{R}^n$ as follows: for $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$, define

$$(x, y) = x \cdot y = \sum_{i=1}^{n} x_i y_i.$$

Example. An inner product can be put on the complex vector space $\mathbf{C}^n$ as follows: for $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$, define

$$(x, y) = \sum_{i=1}^{n} x_i \overline{y_i}.$$

Example. An inner product can be put on the vector space l^2 as follows: for $x = (x_1, x_2, \ldots)$ and $y = (y_1, y_2, \ldots)$ define

$$(x, y) = \sum_{i=1}^{\infty} x_i \overline{y_i}.$$

Notice. In each of the three inner product spaces above, the inner product can be used to define a norm: $\|x\| = \sqrt{(x, x)}$. In fact, in each case the norm determined by the inner product is the norm on the vector space we mentioned when these spaces were originally introduced.

Example. The space $L^2([a,b])$ of all square (Lebesgue) integrable functions on the real interval $[a,b]$:

$$L^2([a,b]) = \left\{ f : [a,b] \to \mathbf{C} \;\middle|\; \int_a^b |f(x)|^2 \, dx < \infty \right\}$$

has as an inner product defined by

$$(f,g) = \int_a^b f(x)\overline{g(x)} \, dx.$$

Definition/Theorem. An inner product space has a *norm* $\|\cdot\|$ induced by the inner product as follows: $\|x\| = \sqrt{(x,x)}$.

Note. Since an inner product space necessarily has a norm, it is of interest to know if this normed space is a Banach space (i.e. if it is complete).

Definition. Two vectors x and y in an inner product space are *orthogonal* if $(x, y) = 0$.

Theorem (Pythagorean Formula.)
If x and y are orthogonal vectors in an inner product space, then

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

In $\mathbf{R}^2$, this is simply the Pythagorean Theorem.

VII. HILBERT SPACES

Definition. A complete inner product space is a *Hilbert space.*

Note. We have the following general inclusions:

Hilbert spaces $\subset$ Banach spaces $\subset$ vector spaces

Example. We have already seen several examples of Hilbert spaces. Some of these are:

(a) $\mathbf{C}^n = \{(z_1, z_2, \ldots, c_n) \mid z_i \in \mathbf{C}\}$.

(b) $l^2 = \left\{(z_1, z_2, \ldots) \mid z_i \in \mathbf{C} \text{ and } \sum_{i=1}^{\infty} |z_i|^2 < \infty\right\}$.

(c) $L^2([a,b]) = \left\{f : [a,b] \to \mathbf{C} \;\middle|\; \int_a^b |f(x)|^2\, dx < \infty\right\}$.

To establish that these are in fact Hilbert spaces, the only difficult part is the establishment of completeness.

VII.1 Bases in Hilbert Spaces

Definition. A *Schauder basis* (or simply *basis*) of an infinite dimensional Banach space is a set of vectors $\{x_1, x_2, \ldots\}$ such that for any vector x in the Banach space, there is a unique sequence of scalars $\{\alpha_1, \alpha_2, \ldots\}$ such that $x = \sum_{i=1}^{\infty} \alpha_i x_i$.

Note. Not every Banach space has a basis. We are interested in Hilbert spaces which have bases.

VII.2 Seperable Hilbert Spaces

Note. We now need a few "mathy" definitions.

Definition. A set is *countable* if a complete "listing" of the set can be made.

Examples.

The natural numbers are countable: $\{1, 2, 3, \ldots\}$.

The integers are countable: $\{0, 1, -1, 2, -2, \ldots\}$.

Surprisingly, the rational numbers are countable (even though they are very different from the integers *topologically*).

The real numbers are not countable!

Definition. Suppose X is a normed space. A set D is *dense* in X if every open set in X includes an element of D.

Example. The rational numbers are dense in the real numbers. The integers are not dense in the reals.

Definition. A Hilbert space with a countable dense subset is *seperable.*

Note. Since $\mathbf{Q}$ is countable and dense in $\mathbf{R}$, then $\mathbf{R}$ forms a seperable Hilbert space (in fact, any finite dimensional Hilbert space is seperable - and remember, a finite dimensional Hilbert space/vector space is isomorphic to either $\mathbf{R}^n$ or $\mathbf{C}^n$ depending on the scalar field).

Definition. A subset X of a Hilbert space is an *orthonormal set* if $\|x\| = 1$ for all $x \in X$ and $(x, y) = 0$ (that is, x and y are orthogonal) for all $x, y \in X$.

Theorem. A Hilbert space is seperable if and only if it has an orthonormal basis.

VII.3 Classification of Seperable Hilbert Spaces

Definition. A Hilbert space H_1 is *isomorphic* to a Hilbert space H_2 if there exists a one-to-one linear mapping T from H_1 onto H_2 such that $(T(x), T(y)) = (x, y)$ for every $x, y \in H_1$.

Note. Now for the BIG RESULT! Recall that the "Fundamental Theorem of Linear Algebra" tells you what a finite dimensional vector space "looks like."

Theorem (Riesz-Fisher Theorem).
An infinite dimensional Hilbert space with scalar field $\mathbf{C}$ (that is, a seperable Hilbert space that is not isomorphic to some $\mathbf{C}^n$) is isomorphic to

$$l^2 = \left\{(z_1, z_2, \ldots) \mid z_i \in \mathbf{C} \text{ and } \sum_{i=1}^{\infty} |z_i|^2 < \infty\right\}.$$

Note. An orthonormal basis for l^2 is

$$\{e_1, e_2, e_3, \ldots\} = \{(1,0,0,0,\ldots),(0,1,0,0,\ldots),$$
$$(0,0,1,0,\ldots),\ldots\}.$$

VIII. An Application

Definition. Suppose a flexible, elastic homogeneous string of length l lies along the $x-$axis from $x = 0$ to $x = l$. Displace the string by an amount $\Phi(x)$ at position x (initial position) and give the string a velocity $\Psi(x)$ at position x (initial velocity). Let $u(x,t)$ represent the displacement of the string at position x and time t. Then u satisfies

$$u_{tt} = c^2 u_{xx}$$

where $c = \sqrt{T/\rho}$ is the *wave speed*, T is the tension in the string (a constant), and ρ is the density of the string (mass per unit length). Therefore the *wave equation* with initial and boundary conditions is

$$\begin{aligned}
\frac{\partial^2 u}{\partial t^2} &= c^2 \frac{\partial^2 u}{\partial x^2} \text{ for } x \in (0,l) \\
u(0,t) &= u(l,t) = 0 \text{ (boundary conditions)} \\
u(x,0) &= \Phi(x) \text{ (initial position)} \\
\frac{\partial u}{\partial t}(x,0) &= \Psi(x) \text{ (initial velocity).}
\end{aligned}$$

Note. We look for a solution of the wave equation of the form $u(x,t) = X(x)T(t)$ (the method of seperation of variables). Plugging this into the wave equation yields:

$$X(x)T''(t) = c^2X''(x)T(t)$$

$$\text{or} \quad \frac{-T''}{c^2T} = -\frac{X''}{X} = \lambda.$$

Notice that λ is constant since $\dfrac{\partial\lambda}{\partial x} = \dfrac{\partial\lambda}{\partial t} = 0$. In fact, $\lambda > 0$. Let $\lambda = \beta^2$. Then we have the pair of ODEs:

$$X'' + \beta^2X = 0$$

$$\text{and} \quad T'' + c^2\beta^2T = 0.$$

These second order ODEs have solutions of the form

$$\begin{aligned} X(x) &= C\cos(\beta x) + D\sin(\beta x) \\ T(t) &= A\cos(\beta ct) + B\sin(\beta ct) \end{aligned}$$

respectively where A, B, C, and D are arbitrary constants.

Imposing the boundary conditions:

$$\begin{aligned} X(0) &= C \equiv 0 \\ X(l) &= D\sin(\beta l) \equiv 0, \end{aligned}$$

we see that $C = 0$ and $\beta l = n\pi$ for some integer n and so

$$\lambda_n = \beta^2 = \left(\frac{n\pi}{l}\right)^2 \text{ where } n = 1, 2, 3, \ldots..$$

Therefore there are an infinite number of solutions of the wave equation which satisfy the boundary conditions, each of the form:

$$u(x,t) = \left(A_n \cos\left(\frac{n\pi ct}{l}\right) + B_n \sin\left(\frac{n\pi ct}{l}\right)\right) \sin\left(\frac{n\pi x}{l}\right)$$

where $n = 1, 2, 3 \ldots.$ Summing these solutions we get

$$u(x,t) = \sum_{n=1}^{\infty} \left(A_n \cos\left(\frac{n\pi ct}{l}\right) + B_n \sin\left(\frac{n\pi ct}{l}\right)\right) \sin\left(\frac{n\pi x}{l}\right)$$

is THE solution to the wave equation provided the initial conditions are satisfied.

This leads us to require that

$$\begin{aligned} u(x,0) &= \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi x}{l}\right) = \Phi(x) \text{ and} \\ u_t(x,0) &= \sum_{n=1}^{\infty} \frac{n\pi c}{l} B_n \sin\left(\frac{n\pi x}{l}\right) = \Psi(x). \end{aligned}$$

Therefore, we can solve the wave equation provided that we can represent the initial conditions $\Phi(x)$ and $\Psi(x)$ in the above form.

Theorem. The set

$$\left\{\frac{2}{l}\sin\left(\frac{n\pi x}{l}\right) \mid n = 1, 2, 3 \ldots\right\}$$

is an orthonormal basis for

$$L^2([0,l]) = \left\{f : [0,l] \to \mathbf{R} \mid \int_0^l |f(x)|^2\, dx < \infty\right\}.$$

Note. We see that the previous theorem tells us that we can find a unique solution to the wave equation with initial and boundary conditions as long as $\Phi, \Psi \in L^2([0,l])$. That is, provided that Φ and Ψ lie in the unique infinite dimensional vector space.

Theorem. With f as above,

$$A_n = \int_0^l f(x) \left\{ \frac{2}{l} \sin\left(\frac{n\pi x}{l}\right) \right\} dx.$$

These A_n are called the *Fourier coefficients* of f.

Notice. There is a very geometric interpretation of the A_ns as *components* of f "in the direction"

$$\frac{2}{l} \sin\left(\frac{n\pi x}{l}\right).$$

Each A_n is simply the inner product ("dot product" if you like) of f with the nth basis "vector."