Informal finite element course:
introduction to vector spaces, inner products and norms.

Dr David Ham

November 15, 2013

1 Real vector spaces

Abstraction is core to mathematics: mathematical objects are defined by the core properties which they might have. When we think of vectors, we think of points in $\mathbb{R}^n$, the space of n-vectors whose components are real numbers. However thinking more abstractly, we can define vectors as any objects which behave like vectors under the appropriate set of operations.

Definition 1.1. A real[1] vector space, V, is a set equipped with two operations, vector addition and multiplication by a (real) scalar such that for any $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$, and for any $\alpha, \beta \in \mathbb{R}$:

1. *Multiplicative closure*: $\alpha\mathbf{x} \in V$.
2. *Additive closure*: $\mathbf{x} + \mathbf{y} \in V$.
3. *Commutativity*: $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$.
4. *Additive associativity*: $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$.
5. *Multiplicative associativity* : $\alpha(\beta\mathbf{x}) = (\alpha\beta)\mathbf{x}$.
6. *Distributivity of scalar sums* : $(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$.
7. *Distributivity of vector sums* : $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$.
8. *Additive identity*: There exists $\mathbf{0} \in V$ such that for any $\mathbf{x} \in V$, $\mathbf{0} + \mathbf{x} = \mathbf{x}$.
9. *Multiplicative identity*: $1\mathbf{x} = \mathbf{x}$.

Informally, this means that vectors can be added to each other, and multiplied by scalars, and these operations combine together in the expected ways. The most obvious missing property is any way of multiplying vectors together. In fact, this is not a required property of vector spaces, although we shall primarily be interested in vector spaces for which some form of multiplication exists.

[1] more generally, it is possible to define vector spaces for which the scalars are drawn from another field, such as the complex numbers.

1.1 Examples of vector spaces

Example 1.1.1 ($\mathbb{R}^n$)**.** The most obvious example of a real vector space is the space of n-vectors whose components are real numbers.

Example 1.1.2 (Functions over Ω)**.** Consider the set of functions from some domain Ω to $\mathbb{R}$:

$$V = \{f : \Omega \to \mathbb{R}\} \tag{1}$$

In this case each such function is a vector with the operations of addition and multiplication considered pointwise. That is to say, if $f \in V$ and $\alpha \in \mathbb{R}$ then $g = \alpha f$ is the function given by:

$$g(\mathbf{x}) = \alpha f(\mathbf{x}) \quad \forall \mathbf{x} \in \Omega \tag{2}$$

Similarly, if $f, g \in V$ then $h = f + g$ is given by:

$$h(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x}) \quad \forall \mathbf{x} \in \Omega \tag{3}$$

The zero vector in V is the zero function, given by $\mathbf{0}(\mathbf{x}) = 0\ \forall \mathbf{x} \in \Omega$.

Definition 1.2 (function space)**.** A *function space* is a vector space whose vectors are functions.

2 Subspaces

An absolutely critical part of the finite element method is working with function spaces contained in other function spaces.

Definition 2.1 (subspace)**.** A subset, S of a vector space V equipped with the same addition and multiplication by a scalar operations as V is a *subspace* if it is closed under addition and multiplication by a scalar.

Lemma 2.2. A subspace of a vector space is itself a vector space.

This is simple to prove by checking that the subspace still fulfils all of the conditions of a vector space.

2.1 Subspace (counter)-examples

Example 2.1.1 (Lines through the origin)**.** The set of points in $\mathbb{R}^2$ such that $\alpha x + \beta y = 0$ for some fixed α and β is a subspace.

Example 2.1.2 (Lines *not* through the origin)**.** The set of points $\alpha x + \beta y = \gamma$ for some fixed α and β and some fixed $\gamma \neq 0$ is *not* a subspace of $\mathbb{R}^2$.

Example 2.1.3 (Continuous functions)**.** The set of continuous functions over some domain Ω is a subspace of all of the functions over that domain.

Example 2.1.4 (Positive functions)**.** The set of functions over some domain Ω such that $f(\mathbf{x}) > 0 \ \forall \mathbf{x} \in \Omega$ is *not* a subspace of all of the functions over Ω.

Definition 2.3 ($C^n(\Omega)$)**.** The set $C^n(\Omega)$ is the set of functions $f : \Omega \to \mathbb{R}$ such that the n-th partial derivatives of f exist and are continuous.

Hence $C^0(\Omega)$ consists of the continuous functions on Ω, $C^1(\Omega)$ consists of all of the functions whose first derivative is continuous, and so on. In particular, the space $C^\infty(\Omega)$ consists of all functions for which an unlimited number of continuous derivatives exist.

Example 2.1.5 ($C^n(\Omega)$)**.** $C^n(\Omega)$ is a subspace of $C^m(\Omega)$ for all $m < n$.

3 Inner products

We noted earlier that a vector space need not be equipped with any form of product between vectors. However we already know that the function spaces $\mathbb{R}^n$ come equipped with the dot product:

$$a \cdot b = \sum_{i=1}^{n} a_i b_i \quad \forall a, b \in \mathbb{R}^n \tag{4}$$

Just as mathematicians generalise the concept of vector to anything with the properties of a vector, the concept of a dot product can also be generalised.

Definition 3.1 (Inner product)**.** An inner product $\langle \cdot, \cdot \rangle$ on a real vector space V is a function $V \times V \to \mathbb{R}$. With the following properties for any vectors $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$, and scalars α and β:

1. *Commutativity* $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$.
2. *Bilinearity* $\langle \alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z} \rangle = \alpha\langle \mathbf{x}, \mathbf{z} \rangle + \beta\langle \mathbf{y}, \mathbf{z} \rangle$.
3. *Positivity* $\langle \mathbf{x}, \mathbf{x} \rangle \geqslant 0$. Additionally $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ if and only if $\mathbf{x} = 0$.

An important property of vector spaces equipped with inner products is orthogonality:

Definition 3.2 (orthogonality)**.** Two vectors $\mathbf{x}$ and $\mathbf{y}$ are orthogonal with respect to an inner product $\langle \cdot, \cdot \rangle$ if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.

3.1 Inner product examples

Example 3.1.1 (the dot product)**.** Unsurprisingly, the dot product on $\mathbb{R}^n$ is an inner product. This product is also referred to as the l_2 inner product.

Example 3.1.2 (scalar multiplication)**.** On the vector space $\mathbb{R}$, the dot product just reduces to multiplication of scalars, so multiplication is an inner product on $\mathbb{R}$

Example 3.1.3 (the L_2 inner product in one dimension)**.** For square-integrable real-valued functions on a closed interval $[a, b]$, the L_2 inner product is given by:

$$\langle f, g \rangle_{L_2} = \int_a^b fg \mathrm{d}x \tag{5}$$

Example 3.1.4 (the L_2 inner product in higher dimensions)**.** The L_2 inner product generalises to vector-valued functions on domains of any dimension ($\Omega \subset \mathbb{R}^n \to \mathbb{R}^m$) by using the dot product to form the integral. For any square-integrable functions f, g:

$$\langle f, g \rangle_{L_2} = \int_\Omega f \cdot g \mathrm{d}\mathbf{x} \tag{6}$$

The preceding examples use the space of square-integrable functions, which we define here for completeness:

Definition 3.3 (square-integrable functions, $\mathrm{L}_2(\Omega)$)**.** The vector space of square-integrable functions $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ is the set of functions for which:

$$\int_\Omega f \cdot f \mathrm{d}\mathbf{x} \tag{7}$$

exists and is finite.

The precise limits of $\mathrm{L}_2(\Omega)$ is a more complex question than we will address here, but it is sufficient to know that all "well-behaved" functions we will encounter in the finite element method are in $\mathrm{L}_2(\Omega)$.

4 Norms

Another property which we might be interested in for a vector is its magnitude. For a vector in $\mathbb{R}^n$ we are used to the Euclidean norm, or l_2 norm given by:

$$\|\mathbf{x}\| = \sqrt{\sum_{i=1}^n \mathbf{x}_i^2} = \sqrt{\mathbf{x} \cdot \mathbf{x}} \quad \forall \mathbf{x} \in \mathbb{R}^n \tag{8}$$

Like the inner product, there is a generalised concept of a magnitude function on a vector space:

Definition 4.1 (norm)**.** A *norm* on $\|\cdot\|$ on a real vector space V is a function $V \to \mathbb{R}$ with the following properties for any vectors $\mathbf{x}$ and $\mathbf{y}$, and scalar α:

1. *Positivity* $\|\mathbf{x}\| \geqslant 0$. Additionally $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = 0$.

2. *Linear scaling* $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$.
3. *Triangle inequality* $\|\mathbf{x}+\mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$.

The properties of norms are quite closely related to those for inner products, so it is not surprising that there is a relationship between inner products and norms:

Theorem 4.2 (induced norms)**.** For every inner product $\langle\cdot,\cdot\rangle$ on a vector space V there is a norm on that vector space given by:

$$\|\mathbf{x}\| = \sqrt{\langle\mathbf{x},\mathbf{x}\rangle} \tag{9}$$

This norm is referred to as the norm *induced* by the inner product.

We will also state without proof the following important result:

Theorem 4.3 (Cauchy-Schwarz inequality)**.** If V is an inner product space then:

$$\langle\mathbf{x},\mathbf{y}\rangle \leqslant \|\mathbf{x}\|\|\mathbf{y}\| \tag{10}$$

where the norm is the norm induced by the inner product.

4.1 Induced norm examples

Definition 4.4. l_2 norm The Euclidean, or l_2 norm on $\mathbb{R}^n$ is the norm induced by the dot product.

Definition 4.5. L_2 norm The L_2 norm on $L_2(\Omega)$ is the norm induced by the L_2 inner product. It is given by:

$$\|f\|_{L_2} = \sqrt{\int_\Omega f(\mathbf{x})^2 \mathrm{d}\mathbf{x}} \tag{11}$$

5 Inner product spaces and Hilbert spaces

Definition 5.1 (Inner product space)**.** A vector space equipped with a particular inner product is an *inner product space*.

Definition 5.2 (Hilbert spaces)**.** An inner product space which is complete in the norm induced by the inner product is called a *Hilbert space*.

Completeness is beyond the scope of this course, however in a non-precise sense, it means that the limit of any convergent sequence of vectors is also in the vector space. Loosely speaking, this means that calculus "works" in Hilbert spaces. Consequently they are very important to the finite element method.

6 Bases for vector spaces

A core feature of a vector space is closure under linear combinations. That is for a vector space V:

$$\sum_i \alpha_i \mathbf{x}_i \in V, \quad \forall \mathbf{x}_i \in V, \forall \alpha_i \in \mathbb{R} \tag{12}$$

This leads us to define a very useful concept:

Definition 6.1 (span)**.** The *span* of a set of vectors is the set of all linear combinations of those vectors.

An immediate consequence of this is the following:

Lemma 6.2. The span of a set of vectors in a vector space V is a subspace of V.

If we have a set of vectors, and the subspace they span, it is very useful to know if this set is minimal: do we need all the vectors to be in the set in order to span that subspace, or is there a smaller subset which would do.

Definition 6.3 (linear independence)**.** A set of vectors $\mathbf{x}_i \in V$ is linearly independent if none of the vectors in the set lies in the span of the other vectors. That is to say:

$$\mathbf{x}_j \neq \sum_{i \neq j} \alpha_i \mathbf{x}_i, \quad \forall \alpha_i \in \mathbb{R} \tag{13}$$

This leads us to the core definition of this part:

Definition 6.4 (basis)**.** A *basis* for a vector space is a linearly independent set of vectors which span the whole vector space.

That is to say, a basis is a minimal set of vectors $\{\phi_i\} \in V$ such that:

$$\forall \mathbf{x} \in V, \exists x_i \in \mathbb{R} \text{ such that } \mathbf{x} = \sum x_i \phi_i \tag{14}$$

It is *not* the case that a unique basis exists for any given function space. Indeed, every function space has an infinite number of bases. However the cardinality of the basis, that is to say the number of vectors in the basis set, is unique to a function space.

Definition 6.5 (dimension)**.** The dimension of a vector space is the cardinality of its basis.

6.1 Basis examples

Example 6.1.1 ($\mathbb{R}^n$)**.** Clearly the canonical vector space is $\mathbb{R}^n$. In this case, the most commonly used basis vectors are aligned with the coordinate axes. In $\mathbb{R}^3$ these are conventionally denoted **i**, **j** and **k**, and we know we can write any $\mathbf{x} \in \mathbb{R}^3$ as:

$$\mathbf{x} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \tag{15}$$

of course we usually just write:

$$\mathbf{x} = \begin{bmatrix} i \\ j \\ k \end{bmatrix} \tag{16}$$

When we express a vector as a list of numbers, we have implicitly chosen a basis and the list of numbers is the list of coefficients of the basis functions.