Linear Algebra – Lesson 10

Vector spaces, inner product spaces

1. Vector spaces

Today we will properly define the term **vector space**.

How did we start our course? We defined vectors and their two basic operations: vector addition and scalar multiplication. They are what allows for linear combinations. Why does it all boil down to linear combinations? Because linear combinations are everywhere in linear algebra: they appear in $A\bar{x} = \bar{b}$, in the definitions of linear independence, of a basis, of the span of a set of vectors etc.

A vector space V is a set of vectors over a field K (we know two fields, $\mathbb{R}$ and $\mathbb{C}$) with a set of rules, such that the linear combination of any two vectors in V is another vector in V.

More formally, a vector space must be:

- Closed under scalar multiplication
- Closed under vector addition

The presentation shows the 10 axioms that any set V must satisfy in order to be a vector space:

Axioms about the operations

1. $\bar{u} + \bar{v} = \bar{v} + \bar{u}$
2. $(\bar{u} + \bar{v}) + \bar{w} = \bar{u} + (\bar{v} + \bar{w})$
3. $k(\bar{u} + \bar{v}) = k\bar{u} + k\bar{v}$
4. $(k + c)\bar{v} = k\bar{v} + c\bar{v}$
5. $(kc)\bar{v} = k(c\bar{v})$
6. $1\,\bar{v} = \bar{v}$

Axioms about the set

1. $\bar{u} + \bar{v}$ is in V
2. $k\bar{v}$ is in V
3. There exists a unique vector 0 such that $\bar{v} + 0 = \bar{v}$
4. For any $\bar{v} \in V$ there exists a unique vector $-\bar{v}$ such that: $\bar{v} + (-\bar{v}) = 0$

Example – a vector space

$\mathbb{R}^2$ is a vector space, because we can take any linear combination of two vectors, and get a new vector in $\mathbb{R}^2$. So is $\mathbb{C}^3$, for example.

Example – not a vector space

The positive quarter of $\mathbb{R}^2$ is not a vector space. Is it closed under vector addition? Yes. But is it closed under scalar multiplication? Not at all, because $-\bar{v}$ is already outside the positive quarter of $\mathbb{R}^2$.

1.2 Vector subspace

1.2.1 Vector subspace definition

Sometime we can take a **subset** U of a vector space V, and still get a vector space.

Given a vector space V, a subset U of V is a vector subspace if it satisfies the following conditions:

1. U is closed under addition: for any $\bar{v}, \bar{u} \in U$, also $(\bar{v} + \bar{u}) \in U$
2. U is closed under scalar multiplication: for any $\bar{u} \in U$ and scalar k, also $k\bar{u} \in U$.
   Notice that this means that U includes the zero vector

Example – a subspace

The span of $\bar{v} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ – $sp(\bar{v})$ – is a vector subspace in $\mathbb{R}^2$. It includes all scalar multiplications of $\bar{v}$, and therefore it is closed under scalar multiplication and vector addition. Importantly, it also includes the zero vector.

Even the zero vector $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ by itself a vector subspace.

1.2.2 The four fundamental subspaces

We already know the four fundamental subspaces associated with an $m \times n$ matrix:

1. Column space of A: $colsp(A), Im(A)$
$$Sp(\{\bar{c}_1, \bar{c}_2, \dots, \bar{c}_n\})$$
2. Row space of A: $rowsp(A)$
$$Sp(\{\bar{r}_1, \bar{r}_2, \dots, \bar{r}_m\})$$
3. Kernel of A, nullspace of A: $\ker(A)$
$$\{\bar{x} | A\bar{x} = 0\}$$
4. Left nullspace of A: $\ker(A^T)$

$$\{\bar{y} | A^T \bar{y} = 0\} \text{ or equivalently } \{\bar{y} | \bar{y}^T A = 0\}$$

Example

Prove that $\ker(A)$ is a vector subspace.

1.3 Matrices vector space

The axioms that define a vector space allow us to talk about more abstract vector spaces than we had so far. For example, the set of real 2×2 matrices is a vector space. A basis for this space is:

$$M_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, M_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, M_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, M_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

We can represent any 2×2 matrix as a linear combination of these basis "vectors":

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = aM_1 + bM_2 + cM_3 + dM_4$$

It includes the zero vector $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. And we have already defined matrix addition and scalar multiplication for matrices.

2. Polynomial space $P_n(x)$

2.1 Polynomials as vectors

$P_n(x)$ is the set of all polynomials of degree smaller or equal to n:

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

with coefficients a_i in the field K.

$P_n(x)$ is a vector space, with vector addition and scalar multiplication defined as the usual addition and multiplication operations we know. A "vector" in $P_n(x)$ is simply a polynomial (the zero vector here would be just 0).

For example, let $p(x) = 1 + x^2$ and $q(x) = 2 + 2x^4$. We can define linear combinations in this space like that:

$$p(x) + q(x) = 3 + x^2 + 2x^4$$
$$2p(x) = 2 + 2x^2$$

Question: What is the dimension of the vector space $P_n(x)$?
Answer: The set $S = \{1, x, x^2, \dots, x^n\}$ is a basis to $P_n(x)$. We will call these basis vectors $\hat{e}_0, \hat{e}_2, \dots, \hat{e}_n$. Any polynomial of degree less than or equal to n can be expressed as a linear combination of these powers of x (so they span $P_n(x)$) and you can also show that they are linearly independent. Since this basis has $n + 1$ elements in it, the dimension of $P_n(x)$ is $n + 1$.

2.1.1 Mapping polynomials to vectors in $\mathbb{R}^n$

We are still used to thinking about vectors as arrows in space, or as a list of numbers. Notice that in this case of polynomials, we can "translate" between a polynomial and a "regular" vector. Specifically, we can map each polynomial to a vector in $\mathbb{R}^n$:

$$p(x) \rightarrow \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, q(x) \rightarrow \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 2 \end{pmatrix}$$

with the basis vectors we defined above, so:

$$p(x) = \hat{e}_0 + \hat{e}_2$$
$$q(x) = 2\hat{e}_0 + 2\hat{e}_4$$

2.2 Linear operators on polynomials

Back in $\mathbb{R}^2$ we studied linear transformations that took a vector as input and gave another vector as output. We defined a matrix as the coordinate-based representation of a linear transformation (like a rotation-scaling matrix, a shear matrix etc).

What are linear transformations in the case of polynomials? They take as input a polynomial and give as output another polynomial. In this context, when the vectors are actually functions, we usually call the linear transformation a "linear **operator**".

2.2.1 The derivative as a linear operator

The derivative is a linear operator as it satisfies both conditions of linearity:

$$\frac{d}{dx}\big(kf(x)\big) = kf'(x)$$
$$\frac{d}{dx}\big(f(x) + g(x)\big) = f'(x) + g'(x)$$

Since we now have a vector space of polynomials, we should be able to represent this linear operator using a matrix. To build this matrix, we only need to find where each basis vector is mapped to. We denote the derivative operator as $\mathcal{L}$, and the basis vectors as $\hat{e}_0, \hat{e}_1, \dots, \hat{e}_n$. Then:

$$\mathcal{L}(\hat{e}_0) = \frac{d}{dx}(1) = 0 \qquad \mathcal{L}(\hat{e}_1) = \frac{d}{dx}(x) = 1 = \hat{e}_0 \qquad \mathcal{L}(\hat{e}_2) = \frac{d}{dx}(x^2) = 2x = 2\hat{e}_1$$

And in general:

$$\mathcal{L}(\hat{e}_k) = k\hat{e}_{k-1}$$

The matrix representation of the derivative operator is therefore:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ \dots & \dots & \dots & \dots \end{pmatrix}$$

3. Vector space of real functions over $[\mathbf{0}, \mathbf{1}]$

Not only polynomials make up a vector space. We can define a vector space for other functions as well. For example, the space of continuous real functions on the interval $[0,1]$. You can check and see that the axioms of a vector space hold here as well.

Here the vectors are functions ($f(x), g(x)$, …).

Notice that we can't represent these vectors as a list of numbers as we are used to, because this vector space is not finite. In fact, we haven't even defined a proper basis for this space. However, one can define a basis for this space, using sines and cosines. This is the idea behind the Fourier transform, which is discussed in detail in ELSC's Calculus course.

4. Inner product spaces

When we defined a vector space, we only cared about linear combinations. One operation we haven't discussed in this context is the inner product (or dot product). Indeed, different inner products can be defined for the same vector space, provided that they satisfy the required properties.

Definition: An **inner product space** is a vector space V along with an inner product defined on it.

<u>4.1 The inner product for real vector spaces</u>

For a vector space V **over** $\mathbb{R}$, the inner product is a mapping of any pair of vectors $\bar{v}, \bar{w}$ in V to a real scalar. An inner product must satisfy the following properties:

1) **Symmetric**:

$$\langle \bar{v}, \bar{w} \rangle = \langle \bar{w}, \bar{v} \rangle$$

2) **Linear in the first argument**:

$$\langle \lambda_1 \bar{v}_1 + \lambda_2 \bar{v}_2, \bar{w} \rangle = \lambda_1 \langle \bar{v}_1, \bar{w} \rangle + \lambda_2 \langle \bar{v}_2, \bar{w} \rangle$$

This is often presented as two axioms:

I. Additivity: $\langle \bar{v} + \bar{u}, \bar{w} \rangle = \langle \bar{v}, \bar{w} \rangle + \langle \bar{u}, \bar{w} \rangle$

II. Homogeneity: $\langle \lambda \bar{v}, \bar{w} \rangle = \lambda \langle \bar{v}, \bar{w} \rangle$

3) **Positive definite**:

$$0 \leq \langle \bar{v}, \bar{v} \rangle \quad and \quad \langle \bar{v}, \bar{v} \rangle = 0 \;\; iff \;\; \bar{v} = 0$$

Using these properties, you can show that the inner product in a vector space over $\mathbb{R}$ is also linear in the second argument.

4.2 The inner product and similarity measures

An inner product is a useful tool when we want to measure vectors (using norms) or to quantify their similarity. This is done using the distance between them or the angle between them. The latter method is often called **cosine similarity**.

[See the Lesson 10 presentation for a few slides about word2vec and about vector space embedding of brain connectomes]

4.3 The inner product for complex vector spaces

In the previous lesson, we saw that the inner product does **not** satisfy the desired conditions in the case of complex vectors. Using the norm of a vector as a case study, we defined the inner product for complex vectors as[1]:

$$\langle \bar{v}, \bar{w} \rangle = \bar{v}^H \bar{w} = \bar{v}^{*T} \bar{w}$$

In general, an inner product for a vector space V over $\mathbb{C}$ is a mapping of any pair of vectors $\bar{v}, \bar{w}$ in V to a **complex** scalar. A complex inner product must satisfy the following properties:

1) **Conjugate symmetric (also called Hermitian)**:
$$\langle \bar{v}, \bar{w} \rangle = \langle \bar{w}, \bar{v} \rangle^*$$
2) **Conjugate linear in the first argument**:
$$\langle \lambda_1 \bar{v}_1 + \lambda_2 \bar{v}_2, \bar{w} \rangle = \lambda_1^* \langle \bar{v}_1, \bar{w} \rangle + \lambda_2^* \langle \bar{v}_2, \bar{w} \rangle$$
   This is often presented as two axioms:
   I. Additivity: $\langle \bar{v} + \bar{u}, \bar{w} \rangle = \langle \bar{v}, \bar{w} \rangle + \langle \bar{u}, \bar{w} \rangle$
   II. Conjugate homogeneity: $\langle \lambda \bar{v}, \bar{w} \rangle = \lambda^* \langle \bar{v}, \bar{w} \rangle$

3) **Positive definite**:
$$0 \leq \langle \bar{v}, \bar{v} \rangle \quad and \quad \langle \bar{v}, \bar{v} \rangle = 0 \;\; iff \;\; \bar{v} = 0$$

Notice that in general $\langle \bar{v}, \bar{w} \rangle \in \mathbb{C}$, but $\langle \bar{v}, \bar{v} \rangle \in \mathbb{R}$.

Question: What about the linearity in the second argument?
Answer: We use the first two properties to test for additivity and homogeneity:

$$\langle \bar{v}, \bar{w} + \bar{u} \rangle = \langle \bar{w} + \bar{u}, \bar{v} \rangle^* = \langle \bar{w}, \bar{v} \rangle^* + \langle \bar{u}, \bar{v} \rangle^* = \langle \bar{v}, \bar{w} \rangle + \langle \bar{v}, \bar{u} \rangle$$
$$\langle \bar{v}, k\bar{w} \rangle = \langle k\bar{w}, \bar{v} \rangle^* = (k^* \langle \bar{w}, \bar{v} \rangle)^* = k \langle \bar{w}, \bar{v} \rangle^* = k \langle \bar{v}, \bar{w} \rangle$$

So the complex inner product is linear in the second argument.

[1] Some books define the inner product for complex vectors with the complex conjugate in the second argument rather than the first:

$$\langle \bar{v}, \bar{w} \rangle = \bar{v}^T \bar{w}^*$$

Here I preferred the above definition to fit the definition used in MATLAB using the dot function.

4.2.1 Lengths and angles

The norm of a vector $\bar{v}$ is defined using the complex inner product (as we saw before):

$$|\bar{v}| = \sqrt{\langle \bar{v}, \bar{v} \rangle}$$

The angle between two vectors $\bar{v}$ and $\bar{w}$ is defined using the real part of the complex inner product:

$$|\bar{v}||\bar{w}| \cos(\theta) = \mathcal{R}(\langle \bar{v}, \bar{w} \rangle)$$

4.3 Other inner product spaces

The two inner products we have seen so far were rather similar.

The inner product can also have quite a different form. For example, we mentioned the vector space of real functions over the interval $[0,1]$. In this infinite vector space, the inner product is no longer defined using sums (and the Σ notation). Instead, it is defined using the integral:

$$\langle f, g \rangle = \int_0^1 f(x) g(x) dx$$

For example, you can calculate $\langle f, g \rangle$ for $f(x) = 3x - 5$ and $g(x) = x^2$ and find the $\langle f, g \rangle = -\frac{11}{12}$.

Do the properties of an inner product hold? We will check for two general function f, g define over $[a, b]$.

Symmetric:

$$\langle f, g \rangle = \int_a^b f(x) g(x) dx = \int_a^b g(x) f(x) dx = \langle g, f \rangle$$

Linear in the first argument:

$$\langle k_1 f_1 + k_2 f_2, g \rangle = \int_a^b \big(k_1 f_1(x) + k_2 f_2(x)\big) g(x) dx = \int_a^b \big(k_1 f_1(x) g(x) + k_2 f_2(x) g(x)\big) dx$$
$$= k_1 \int_a^b f(x)_1 g(x) dx + k_2 \int_a^b f(x)_2 g(x) dx = k_1 \langle f_1, g \rangle + k_2 \langle f_2, g \rangle$$

Positive definite:

$$\langle f, f \rangle = \int_a^b f(x) f(x) dx = \int_a^b \big(f(x)\big)^2 dx \geq 0$$

In addition:

$$\langle f, f \rangle = 0 \ \leftrightarrow \ \int_a^b \big(f(x)\big)^2 dx = 0 \ \leftrightarrow \ \big(f(x)\big)^2 = 0 \ \leftrightarrow \ f(x) = 0$$

Example – orthogonal functions

We've seen orthogonal functions before, when we talked about the vector space of polynomials. For real continuous functions over $[-\pi, \pi]$ the following pair of functions are orthogonal:

$$f(x) = \cos(x)$$
$$g(x) = \sin(x)$$

To see this calculate their inner product:

$$\langle f, g\rangle = \int_{-\pi}^{\pi} \cos(x)\sin(x)dx = \int_{-\pi}^{\pi} \frac{1}{2}\sin(2\mathrm{x})dx = \frac{1}{2}\frac{1-\cos(2x)}{2}\Bigg|_{-\pi}^{\pi} = -\frac{1}{4}(\cos(2\pi) - \cos(-2\pi)) = 0$$

In fact, the functions $1, \cos(x), \cos(2x), \cos(3x) \ldots, \sin(x), \sin(2x), \sin(3x) \ldots$ form an orthogonal basis for the space of square integrable functions over $[-\pi, \pi]$ (those of you taking ELSC's calculus course may have already proved this).

These functions are called the harmonic functions, and they are at the heart of Fourier analysis, which is used in almost any application of digital signal processing.