Vectors, Functions and Random Variables

How to define the scalar product and othogonality of vectors, functions and random variables?

Troels Pedersen

October 7, 2010

1 Abstract

We give the definition of a vector space and discuss in what sense functions and random variables may be called vectors. We also discuss how to take inner products of functions or random variables.

2 Introduction

In high-school, you might have heard the informal definition of a vector:

> A vector is a number with a direction. It can be written as a pair parentheses embracing two (or three) real numbers.

If not paying enough attention to the formal definition of a vector space in the first course on linear algebra, you might think that the only thing that happened in that course was the this definition was extended from the two or three dimensional case to the n-dimensional case (or even to complex vectors). You probably struggled for some time to first accept that a vector could have more that three dimensions since the notion of direction was difficult to grasp. This problem is actually caused, not by the definition of higher dimensional vectors, but by the informal definition of what a vector is. In the following we will give a new definition of what a vector is and see that functions (in this sense) are vectors as well.

3 Definition of a vector space

Contrary to the very direct high-school definition of a vector, we first define the concept of a vector space and then vectors as elements in the vector space.

Let F be a field (such as the real numbers or complex numbers), whose elements will be called scalars. A vector space over the field F is a set V together with two operations:

Vector addition: $V \times V \to V$ denoted $v + w$, where $v, w \in V$

Scalar multiplication: $F \times V \to V$ denoted av, where $a \in F$ and $v \in V$,

satisfying eight axioms:

Axiom 1: Vector addition is associative:
For all $u, v, w \in V$, we have $u + (v + w) = (u + v) + w$.

Axiom 2: Vector addition is commutative:
For all $v, w \in V$, we have $v + w = w + v$.

Axiom 3: Vector addition has an identity element:
There exists an element $0 \in V$, called the zero vector, such that $v + 0 = v$ for all $v \in V$.

Axiom 4: Vector addition has an inverse element:
For all $v \in V$, there exists an element $w \in V$, called the additive inverse of v, such that $v + w = 0$.

Axiom 5: Distributivity holds for scalar multiplication over vector addition:
For all $a \in F$ and $v, w \in V$, we have $a(v + w) = av + aw$.

Axiom 6: Distributivity holds for scalar multiplication over field addition:
For all $a, b \in F$ and $v \in V$, we have $(a + b)v = av + bv$.

Axiom 7: Scalar multiplication is compatible with multiplication in the field of scalars:
For all $a, b \in F$ and $v \in V$, we have $a(bv) = (ab)v$.

Axiom 8: Scalar multiplication has an identity element:
For all $v \in V$, we have $1v = v$, where 1 denotes the multiplicative identity in F.

It follows formally from the definition of vector addition that a vector space is "closed under vector addition" meaning that adding two vectors in the vector space gives a vector in the vector space. Similarly it follows from the definition of scalar multiplication that the vector space is "closed under scalar multiplication". Indeed these two properties are very useful when one needs to determine if a *subset of a vector space is itself a vector space* (this is a well-known exercise in the first-year course in linear algebra).

Notice that we said nothing about what the elements of V (the vector space) really are. The objects in V *could* be complex numbers in between braces, (i.e. members of $\mathbb{C}^n$), but it is not necessary. In that way the formal definition above differs from the high-school definition from before.

So what do we need to do when we are provided a collection of objects and are asked if these form a vector space? We simply need to check two things:

1. Is the concepts of vector addition and scalar multiplication well-defined?

2. Do the eight axioms of a vector space hold?

Next, we think of two examples of sets with and show that they are in fact vector spaces.

3.1 Example: Real Functions

Let us consider the set V of all real-valued functions defined on the real domain:

$$V = \{\text{all functions } v : \mathbb{R} \to \mathbb{R}\},$$

where the sum $u = v + w$ of two functions[1] v, w in V, is defined pointwise as $u(x) = v(x) + w(x)$ for all $x \in \mathbb{R}$. A function $v \in V$ multiplied by a real scalar $a \in \mathbb{R}$, denoted by $u = av$, is defined pointwise as $u(x) = a \cdot v(x)$, for all $x \in \mathbb{R}$.

Clearly, vector addition and scalar multiplication are defined in this case. Hence we only have to check if V fulfil the eight axioms:

1. For $u, v, w \in V$, $u(x) + \big(v(x) + w(x)\big) = \big(u(x) + v(x)\big) + w(x)$, for all $x \in \mathbb{R}$, and thus $u + (v + w) = (u + v) + w$.

2. For $u, v \in V$, we have $u(x) + v(x) = v(x) + u(x)$, for all $x \in \mathbb{R}$. It follows that $v + w = w + v$.

3. Define the zero function as $z(x) = 0$, for all $x \in \mathbb{R}$. Since z is a real function it is a member of V. Then for $v \in V$ it is seen that $v(x) + z(x) = v(x) + 0 = v(x)$ for all $x \in \mathbb{R}$ and therefore $v + z = v$.

4. For any $v \in V$ we can define a function w such that $w(x) = -v(x)$ for all $x \in \mathbb{R}$. It is clear that $w \in V$. It follows that $v(x) + w(x) = v(x) - v(x) = 0$, for all $x \in \mathbb{R}$.

5. For all $a \in \mathbb{R}$ and $v, w \in V$, we have $a(v(x) + w(x)) = av(x) + aw(x)$ for all $x \in \mathbb{R}$. Thus $a(v + w) = av + aw$.

[1]Notice that we make the distinction between the *function* v and the number $v(x)$. The notation is standard in the field of mathematics, but conflicts with the ‘engineering’ notation where $v(x)$ denotes a function v of x.

6. For all $a, b \in \mathbb{R}$ and $v \in V$, we have $(a+b)v(x) = av(x) + bv(x)$ for all $x \in \mathbb{R}$. Thus $(a+b)v = av + bv$.

7. For all $a, b \in \mathbb{R}$ and $v \in V$, we have $a(bv(x)) = (ab)v(x)$ for all $x \in \mathbb{R}$. Thus $a(bv) = (ab)v$.

8. For all $v \in V$, we have $1v(x) = v(x)$ for all $x \in \mathbb{R}$, therefore $1v = v$.

As can be seen, after proving a few of the axioms it gets a boring tedious standard routine — it is certainly not something you would like to go through too often. Therefore, it is often a nice procedure to show that the set of vectors is a subspace of a "larger" vector space. It then suffices to show that the *subset* is closed under vector addition and scalar multiplication.

3.2 Example: Linear Combination of Real Functions

Consider the set of functions $S \subset V$ given as[2]

$$S = \left\{ \sum_{k=1}^{N} a_k v_k \ : \ a_1, \ldots, a_N \in \mathbb{R}, v_1, \ldots, v_N \in V \right\}$$

We see that S is the set of all linear combinations of the vectors (functions) $v \in V$. Now, is S a vector space?

We already now that V is a vector space. Therefore, as explained above, it suffices to test if S is closed under vector addition and scalar multiplication. Since any element s of S is a linear combination of $v_1, \ldots, v_N$ so is the sum of any two elements of S. Thus it is closed for vector addition. Furthermore, since $s \in S$ is a linear combination of $v_1, \ldots, v_N$ so is as for $a \in \mathbb{R}$. Thus $as \in S$.

3.3 Example: Square Integrable Real Functions

Consider the set of functions where we can integrate the sqaured value:

$$U = \{ f \in V : \int_{-\infty}^{\infty} f^2(x) \mathrm{d}x < \infty \}.$$

Is U a vector space? Clearly, $U \subset V$, and thus we need only to verify that U, is closed closed under addition and scalar multiplication. For $v, w \in U$ and $a \in \mathbb{R}$:

[2] We might also write the set S as $S = \mathrm{span}\{v_1, v_2, \ldots, v_N\}$. Basically it means that we take all possible linear combinations of the vectors $v_1, \ldots, v_N$.

1. Multiplication with a constant a: We have to check that for $v \in U$ then av is square integrable, which can be seen by

$$\int (av(x))^2 \mathrm{d}x = a^2 \int v(x)^2 \mathrm{d}x < \infty$$

2. Closed under addition: We have to show that for $v, w \in U$, $v + w$ is still square integrable, i.e., we must show that

$$\int (v(x) + w(x))^2 \mathrm{d}x < \infty.$$

We do this using the inequality[3]: $(a+b)^2 \leq 2(a^2 + b^2)$. Consequently,

$$\int (v(x) + w(x))^2 \mathrm{d}x \leq 2\Big(\int v(x)^2 \mathrm{d}x + \int w(x)^2 \mathrm{d}x \Big) < \infty,$$

and thus the sum of two square integrable functions is also square integrable which shows that U is closed under addition.

Thus U is a sub-space of V and is therefore a vector space by itself.

Comparing the two vector spaces U and S we see a remarkable difference: Due to the definition of S, this space has at most N linearly independent basis vectors, and is therefore finite dimensional. On the contrary, U is infinite dimensional, which can be verified by considering an infinite series of square integrable functions:

$$v_k(x) = \begin{cases} \frac{1}{\sqrt{k}}, & |x| < \frac{k}{2}, \\ 0 & \text{otherwise} \end{cases}, \quad k = 1, 2, 3, \ldots$$

Since for any k, $\int v_k(x)^2 \mathrm{d}x = 1$ the vector v_k is a member of U. It is also apparent that v_k is linearly independent of $v_{k-1}, \ldots, v_1$. So a basis for U includes infinitely many vectors, and the dimension of U is therefore infinite! Can you think of another infinite series of functions that are linearly independent and square integrable?

3.4 Example: Random Variables

It is sometimes useful to view a random variable as a vector, that is, as a member of a vector space. As an example we consider the vector space of all real-valued random variables:

$$W = \{\text{all real-valued random variables}\}.$$

[3] To prove this inequality, consider first $0 \leq (a-b)^2 = a^2 + b^2 - 2ab$ and thus we have $2ab \leq a^2 + b^2$. Inserting this inequality in $(a+b)^2 = a^2 + b^2 + 2ab$ we obtain $(a+b)^2 \leq 2(a^2 + b^2)$.

Thus the random variable $X \sim \mathcal{N}(0,1)$ is an element of the vector space W. We can define addition of two elements X, Y in W denoted by $Z = X + Y$. Similarly, multiplication of a random variable X by a scalar a is defined as the random variable $Z = aX$. With these two definitions it can be seen that all eight axioms are fulfilled. (Verify this claim for your self!)

What is the number of dimensions of W?

3.5 Example: Random Variables with Finite Second Moment

Most random variables of interest (there are certainly exceptions) have finite second moment (mean square). Let

$$\tilde{W} = \{X \in W : E[X^2] < \infty\}$$

where $E[\cdot]$ denotes the expectation operator. So is $\tilde{W}$ a vector space? The answer is "yes". It is easy to check that $\tilde{W}$ is a sub-space of W by checking that it is closed under addition and scalar multiplication:

1. Scalar multiplication: $E[(aX)^2] = a^2 E[X^2] < \infty$.
2. Closed under addition: Using the trick $(a+b)^2 \le 2(a^2 + b^2)$ again, we see that $E[(X+Y)^2] \le 2(E[X^2] + E[Y^2]) < \infty$.

We shall see in the next section how to define the inner product of vectors in $\tilde{W}$, i.e., for random variables with finite second moment.

4 Inner Product

You are probably familiar with the inner product (or scalar product) of two vectors v, w in $\mathbb{R}^n$:

$$\langle v, w \rangle = \sum_{k=1}^{n} v_k w_k,$$

where $v_1, \ldots, v_n$ are the entries of v and similarly for w. The first time you saw this definition was in high-school (at least for $n \le 3$).

This definition is OK for vectors in $\mathbb{R}^n$, but for vectors in $\mathbb{C}^n$ we use another:

$$\langle v, w \rangle = \sum_{k=1}^{n} v_k^* w_k.$$

The difference here, is the asterix denoting complex conjugation. So why do we need the asterix? There must be something more to it!

Formally, the inner product (or scalar product) of two vectors in V is a function $\langle \cdot, \cdot \rangle : V \times V \to F$ which satisfies three axioms for all $u, v, w \in V$ and all scalars $a \in F$:

1. Conjugate symmetry: $\langle v, w\rangle = \langle w, v\rangle^*$
2. Linearity: $\langle av, w\rangle = a\langle w, v\rangle$, and $\langle v+u, w\rangle = \langle v, w\rangle + \langle u, w\rangle$.
3. Positive-definiteness: $\langle v, v\rangle \geq 0$ with equality if and only if $v = 0$.

So why the astirisk? Obviously, Axiom 1 dictates it. However, if you do not include it, Axiom 3 does not make sense for complex vectors. Axiom 3 ensures that the 'length' of a vector ends up being a real, positive, number — if your inner product does not obey Axiom 3, you cannot meaningfully speak about the length of a vector.

If $\langle v, w\rangle = 0$ we say that v and w are *orthogonal*, denoted by $v \perp w$. For vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$, orthogonal vectors can be said to have an 90° angle in between them. This interpretation, however, becomes cumbersome, already when considering complex vectors of dimension two and is not as relevant for more advanced vector fields. Thus in such cases we simply stick to the definition that vectors are orthogonal if (and only if) their inner product is zero.

4.1 Example: Inner Product of Square Integrable Functions

Now, we would like to define an inner product for functions. For technical reasons, however, this cannot be done done any real-valued function. We therefore limit the discussion to squared integrable functions, i.e. to members of the vector space U defined earlier. For vectors in U, we can define the inner product as

$$\langle v, w\rangle = \int_{-\infty}^{\infty} v(x)w(x)\mathrm{d}x.$$

To show that this is actually an inner product, we have to check the three axioms (we suppress the integration limits):

1. Conjugate symmetry:
$$\langle v, w\rangle = \int v(x)w(x)\mathrm{d}x = \int w(x)v(x)\mathrm{d}x = \langle w, w\rangle.$$
Since v and w are real-valued functions, $\langle v, w\rangle$ is also real and the conjugate symmetry then follows by $\langle w, v\rangle = \langle v, w\rangle^*$.

2. Linearity:
$$\langle av, w\rangle = \int av(x)w(x)\mathrm{d}x = a\int av(x)w(x)\mathrm{d}x = a\langle v, w\rangle.$$
$$\begin{aligned}\langle v+u, w\rangle &= \int (v(x)+u(x))w(x)\mathrm{d}x \\ &= \int v(x)w(x)\mathrm{d}x + \int u(x)w(x)\mathrm{d}x = \langle v, w\rangle + \langle u, w\rangle.\end{aligned}$$

3. Positive-definiteness:

$$\langle v, v \rangle = \int v(x)v(x)\mathrm{d}x = \int v(x)^2 \mathrm{d}x \geq 0.$$

It follows trivially that $\langle 0, 0 \rangle = 0$. Contrarily, if $\langle v, v \rangle = 0$, and since $v(x)^2 \geq 0$ for all x, the norm $\langle v, v \rangle$ is positive unless[4] $v = 0$.

This is not the only choice of an inner product for real-valued functions. Can you give an example of another inner product which fulfills the three axioms?

4.2 Example: Inner Product of Real-Valued Random Variables With Finite Second Moment

We consider $\tilde{W}$. Let us define the inner product as the correlation:

$$\langle X, Y \rangle = \mathrm{E}[XY]$$

where $E[\cdot]$ denotes the expectation operator. We need to check that the three axioms are satisfied:

1. Conjugate symmetry: $\langle X, Y \rangle = \mathrm{E}[XY] = \mathrm{E}[YX] = \langle Y, X \rangle = \langle Y, X \rangle^*$.
2. Linearity: $\langle aX, Y \rangle = E[aXY] = aE[XY]$, and $\langle X + Z, Y \rangle = E[(X + Z)Y] = E[XY] + E[ZY]$.
3. Positive-defininiteness: $\langle X, X \rangle = E[X^2] \geq 0$ with equality if and only if $X = 0$.

So, the correlation is truly an inner product. You can also define the inner product for complex random variables (what will change in the above definition?).

[4] This is actually true only for Riemann integrable functions. So if you are familiar with other types of integrals, please take U to be the set of all real-valued and Riemann square integrable functions.