

ECE 586: Vector Space Methods

Lecture 10: Frequently Asked Questions

Henry D. Pfister
Duke University

September 25, 2020

1 Vector Spaces

Groups and rings were mentioned briefly. Can vector spaces be defined for groups and rings as well? A group is defined by a set and a single binary operation. For example, the addition operation in a field satisfies exactly the group axioms. Since there is only one operation, there is no real concept of a vector space over a group.

A ring is a set with two binary operations (e.g., addition and multiplication). The difference from a field is that multiplication is not required to be commutative or invertible. For example, consider the set of square matrices over a field with standard addition and multiplication. The concept of a vector space over a ring is called a *module*. Many properties are similar to vector spaces if the ring is nice enough (e.g., if it is a principal ideal domain) but there are also some differences.

I should note that modules are not commonly used (as far as I know) in engineering mathematics. I have only used them once in my research.

For slide 2, I don't quite understand the idea of a general vector space of functions. This is a vector space V where each element of the space is a function that maps X to Y . To make the construction work, the set Y must also be equipped with a vector space structure (so that function evaluations can be added together).

For example, consider the case where $X = [-2, -1] \cup [1, 2]$ and $Y = \mathbb{R}^2$. Then, each element of V is a function $f: [-2, -1] \cup [1, 2] \rightarrow \mathbb{R}^2$. One function in this space is $f(x) = (x^2, e^x)$.

Notice that X is not even closed under addition because $-1 + 1 = 0 \notin X$. Thus, we say that X is a non-empty set because no other structure is required. Still, V is a vector space over \mathbb{R} because we can add functions pointwise and also define scalar multiplication by real numbers.

2 Subspaces

Just to have a concrete example: Using cartesian coordinates, would the \mathbb{R}^2 xy plane be a subspace of the \mathbb{R}^3 xyz volume? If we consider the x-y plane in the \mathbb{R}^3 vector space, then it is a subspace of the \mathbb{R}^3 vector space since the x-y plane can be interpreted as the span of $\{(1, 0, 0), (0, 1, 0)\}$. But if we consider the x-y plane as a vector space of \mathbb{R}^2 , then it is not a subspace of the \mathbb{R}^3 vector space since \mathbb{R}^2 spanned from $\{(1, 0), (0, 1)\}$ is not a subset of \mathbb{R}^3 spanned from $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$.

Subspaces are "closed" under vector addition and scalar multiplication. does this relate to the definition of a closed set we learned earlier? This is a very good question and the answer is no. In the first case, the term closed refers to an idea for binary operations and, in the second case, closed refers to an idea in topology. I don't know of any formal connection between these two concepts.

- A binary operation \circ , which maps $x, y \in X$ to $x \circ y \in X$, is called closed on $A \subseteq X$ if $x \circ y \in A$ for all $x, y \in A$.
- For a metric space (X, d) , a subset $A \subseteq X$ is called closed if it contains all its limit points.

On slide 3, does the lemma require that subset W has the same scalar field F as V ? Yes.

When proving the example on slide 3, would it be sufficient to say that one could use elementary row operations on $A\underline{v}$ and still satisfy the equation $A\underline{v} = \underline{0}$, and thereby construct the $s\underline{w}_1 + \underline{w}_2$ in W requirement for $A\underline{v} = \underline{0}$ to be a subspace? I'm not sure I understand the question completely. In the first part, I assume that you are referring to elementary row operations on A . By doing this, you would be changing A in way that preserves the set of vectors mapped to $\underline{0}$. Thus, elementary row operations show that many matrices have exactly the same null space.

On the other hand, I do not see how elementary row operations can be used to prove that V is a subspace. Instead, one can apply the lemma in slide 3. For every pair $\underline{v}_1, \underline{v}_2 \in V$, we have $A\underline{v}_1 = \underline{0}$ and $A\underline{v}_2 = \underline{0}$. For every scalar $s \in F$, $A(s\underline{v}_1 + \underline{v}_2) = sA\underline{v}_1 + A\underline{v}_2 = \underline{0}$. So $s\underline{v}_1 + \underline{v}_2 \in V$. From the lemma, $V \subseteq F^{n \times 1}$ is a subspace of $F^{n \times 1}$.

3 Linear Dependence and Independence

If all columns of a matrix are linearly independent, does it mean that all rows are also linearly independent? For a $m \times n$ matrix, if $m = n$, the linear independence of columns implies the linear independence of rows. More generally, the number of linearly independent columns always equals the number of linearly independent rows. But, if $m > n$, the m row vectors of length n cannot be linearly independent.

4 Basis

For an invertible matrix, the columns form a basis, and the rows form another basis. Are these two bases the same? If the matrix is symmetric, the two bases are the same. Otherwise, they are bases for the same space (e.g., F^n) but the two ordered lists of basis vectors are not identical.

Does a vector space have only one unique basis? No. For a vector space with dimension n , any linearly independent set with size n is a Hamel basis. In this sense, a vector space can have infinitely many distinct bases.

Could you explain more on the Hamel basis and how it is different from the case in real vectors? The term "Hamel" is a name used to distinguish this definition from a later one that is more appropriate for infinite dimensional spaces. For finite dimensional spaces, you can drop the term Hamel and just say basis. For a vector space with dimension n , any linearly independent set with size n is a basis for the space.

What is the difference between a Hamel basis and a standard basis? The standard basis of a vector space is the set of vectors whose coordinates are all zero, except one that equals 1. The standard basis is orthonormal whereas a Hamel basis only requires linear independence of the basis. For example, $(1, 0)$ and $(1, 1)$ form a Hamel basis for \mathbb{R}^2 .

Given a finite-dimensional vector space V and set of vectors W , what algorithm can determine if W is a basis of V ? One only needs to check the size of W and that the vectors in W are linearly independent. For the space $V = F^n$, one can form a matrix where each row is a vector in W and then use Gaussian elimination to put the matrix in reduced row-echelon form. If the resulting matrix does not have an all-zero row, then W forms a basis for V .

5 Dimension

Does the $\dim(V)$ equal the number of columns of a matrix in a vector space? In general, one cannot say without knowing how V is related to the matrix being discussed. In most examples in the

notes, we assume that $V = F^n$ and consider matrices of size $m \times n$. In this case, we have $\dim(V) = n$ and also A has n columns. More generally, if $T: V \rightarrow W$ is a linear transform and A is coordinate matrix for T , then the number of columns in A equals $\dim(V)$.

In what context, do infinite dimensional vector spaces arise? In engineering, the most common examples are functions (e.g., the set of all $f: [0, 1] \rightarrow \mathbb{R}$) and sequences (e.g., the set of all $x_1, x_2, \dots \in \mathbb{R}$).

Is it possible that an infinite-dimensional matrix is invertible? Yes, an “infinite-dimensional matrix A ” is an array $a_{ij} \in \mathbb{R}$ for $i, j \in \mathbb{N}$. First, one needs to generalize our notion of matrix multiplication to use an infinite sum with convergence defined using the standard metric space of the real numbers. Then, the standard definition of invertibility still applies: An infinite-dimensional matrix A is invertible if there is an infinite-dimensional matrix B such that $AB = BA = I$ (where all implied limits must exist). There is one new element: a finite-dimensional matrix is invertible if it is non-singular (e.g., it does not map any non-zero vector to $\underline{0}$) but this is not sufficient for infinite-dimensional matrices.

6 Other questions

What is the preferred notation for vectors? In most other courses, I’ve used bold non-italic lower case letters. This is a good question without a good answer. First, I use the underline convention in this class so that I can use the same notation on the board during lecture. I do not use it in my papers and talks. Using bold non-italic fonts for vectors (lowercase) and matrices (uppercase) is quite common.

In papers and presentations, my choice depends mainly on the convention in the field and the full set of objects that are discussed. For example, in probability, random variables are typically uppercase italics while realizations are lowercase italics. But, some fields also use uppercase letters for matrices. Thus, if you have vectors and matrices which are both random and non-random, it can be hard to simultaneously satisfy all the conventions.

In general, any reasonable notation is fine as long as it is defined and it is consistent.

Will we establish a link between random variables and vector spaces? Yes, we will see later that one can use vector space methods to unify deterministic (e.g., linear least-squares) estimation and stochastic (e.g., linear minimum mean-squared error estimation).

Is there a formal notation for saying “ V is a vector space over F ”? I see that you used epsilon to denote this for a matrix A (which is a set of vectors), so I’m guessing that epsilon is an appropriate notation for these cases. There is a formal notation for defining the set V (e.g., $V = \mathbb{C}^{10}$) but the scalar field and operations are often implicit. For membership, I think you are referring to the notation $A \in F^{m \times n}$ which strictly means that A is a member of the set of $m \times n$ arrays with elements in F . After one has shown (or stated) that $F^{m \times n}$ is a vector space over F (assuming standard matrix addition and scalar multiplication), then A can be treated as a member of that vector space. People typically say things like “Consider the vector space $V = \mathbb{C}^{10}$ ” and the reader is expected to infer the standard operations from other clues in the paper.

What is the advantage of defining vectors as belonging to vector spaces rather than the usual definition of vectors taught in an introductory linear algebra course? Many of the same definitions are used in an introductory linear algebra course but the focus there is often on the basic rules and visualizing 2D/3D space.

Our goal is focus on the general case so that we can see connections between disparate problems and discuss things like the vector space of real random variables.

The early lecture slides give a very general definition of vector addition and scalar multiplication. In practice, is there ever a vector space which operates under different addition rules than we are used to? In most cases, the addition and scalar multiplication operations will be defined by a field or another vector space. But, one can artificially construct strange vector spaces. Working over \mathbb{R} , consider $V = (-1, 1)$ with $u + v = \tanh(\tanh^{-1}(u) + \tanh^{-1}(v))$ and $sv = \tanh(s \tanh^{-1}(v))$.

Why does a vector space consists of a scalar field F and a vector set V ? What it means in graphical terms? A vector is graphically illustrated by a direction and a length. A “scalar” multiplies a vector by scaling or changing its length. For real vectors, this picture matches our experience in the 3D world. For complex vectors, we cannot really picture this exactly but things happen similarly. For a finite field, many of the concepts break down but the math still works.

Are the vector dot product and cross product defined by vector space? No, vector spaces only provide the bare minimum structure: vector addition and scalar multiplication. Shortly, we will add a measure of length (called a norm) to get a normed vector space. Then, we will add an inner product (i.e., dot product) to get an inner product space. Those two concepts provide pretty much everything we need to solve many problems. To get an abstract version of the cross product, one needs something called an exterior algebra (which we will not discuss).