Vectors and linear combinations

COMS 3251 Fall 2022 (Daniel Hsu)

1 Vectors

The term "vector" can mean many things. For now, we define an *n-vector* to be an n-tuple of *real numbers* (a.k.a. *scalars*). Here, n is a placeholder for a non-negative integer. (We'll drop the "n-" when it is clear from context.)

An example of a 2-vector is $(3, 2)$: it has 2 *components* (a.k.a. *entries*, *coordinates*), each of which is a scalar. The first component of $(3, 2)$ is 3, and the second component is 2. We are used to thinking about such pairs of numbers as points in the Cartesian plane. We refer to this plane as $\mathbb{R}^2$. The only wrinkle is that 2-vectors are better thought of as "displacements", rather than "points". So instead of just plotting the point $(3, 2)$, we usually draw an "arrow" that starts at the origin $(0, 0)$ and ends at the point $(3, 2)$.

The concept of 2-vectors generalizes to 3-vectors (in 3-dimensional Cartesian space) and also to n-vectors for any natural number n. The n-dimensional Cartesian space is denoted by $\mathbb{R}^n$. An n-vector has n components.

We use two ways to write an n-vector. The first is *list format*: $(-\pi, \sqrt{2}, e^2)$, separating components with commas. The second is *column format*:

$$\begin{bmatrix} -\pi \\ \sqrt{2} \\ e^2 \end{bmatrix},$$

arranging components from top to bottom in a column surrounded by brackets. We use symbols (mostly lowercase Roman letters) to assign names to vectors, e.g., $\mathbf{v} = (3, 2)$.[1] We also use subscripts to refer to components of a named vector: v_i for the ith component of $\mathbf{v}$. Two vectors $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_1, \ldots, v_n)$ are equal, written "$\mathbf{u} = \mathbf{v}$", if all of their corresponding components are equal: $u_i = v_i$ for all $i \in \{1, \ldots, n\}$.

Caution. Subscripts are often used in other ways; for instance, they may be used to refer to a vector in a sequence of vectors $(\mathbf{v}_1, \mathbf{v}_2, \ldots)$. So in these

[1] Various mnemonics are used to help us remember that a particular symbol refers to a vector, including bold letter styles ($\mathbf{v}$) and arrow decorations ($\vec{v}$).

cases, it will be important to use context to understand what, say, "vee three" refers to: the third component of a vector, or the third vector in a sequence of vectors. To help, when v_i is supposed to be the ith component of the n-vector $\mathbf{v}$, we'll introduce the vector $\mathbf{v}$ by writing "$\mathbf{v} = (v_1, \ldots, v_n)$". ∎

2 Vector arithmetic

Because vectors, as we have defined them, are "just" tuples of real numbers, we can generalize some arithmetic operations for real numbers to vectors. These operations have intuitive interpretations when we think of vectors as "displacements" in Cartesian space.

- Add two vectors $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$: The result is the 2-vector

  $$\mathbf{u} + \mathbf{v} \;=\; (u_1 + v_1, u_2 + v_2).$$

  E.g., if $\mathbf{u} = (1, 1)$ and $\mathbf{v} = (3, 2)$, then $\mathbf{u} + \mathbf{v} = (4, 3)$.

  To find where the sum vector $\mathbf{u}+\mathbf{v}$ should point in the Cartesian plane, move the "start" of the vector $\mathbf{v}$ to the "end" of the vector $\mathbf{u}$.

  Switching the roles of $\mathbf{u}$ and $\mathbf{v}$ in this operation gives the same result: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$. Adding vectors in any order gives the same result: $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$.

- Scale vector $\mathbf{v} = (v_1, v_2)$ by a real number c: The result is the 2-vector

  $$c\,\mathbf{v} \;=\; (c\,v_1, c\,v_2).$$

  E.g., if $\mathbf{v} = (3, 2)$ and $c = 2$, then $c\,\mathbf{v} = (6, 4)$.

  In the Cartesian plane, the scaled vector may point in the same direction (if $c > 0$) or in the "opposite" direction (if $c < 0$); it may also be longer (if $|c| > 1$) or shorter (if $|c| < 1$). It could also be the same (if $c = 1$).

The *zero vector* (a.k.a. *origin*) $\mathbf{0} = (0, 0)$ is a special vector.

- Scaling any vector by the real number 0 results in the zero vector $\mathbf{0}$.

- Adding $\mathbf{0}$ to another vector $\mathbf{u}$ yields that same vector $\mathbf{u}$.

Of course, vector arithmetic generalizes in the obvious way to 3-vectors, and also to n-vectors for any natural number n.

3 Linear combinations

A *linear combination* of a (finite) collection of vectors is an expression that scales these vectors by real numbers and then adds up the results. The result of a linear combination is another vector. For instance, if $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$ are 3-vectors, and a and b are real numbers, then the 3-vector $a\mathbf{u} + b\mathbf{v} = (au_1 + bv_1, au_2 + bv_2, au_3 + bv_3)$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

Linear combination is the basic way of combining vectors to produce another vector.

Examples.

1. Consider $n = 5$ "resources" that a community might need: wool, grain, lumber, brick, ore. Suppose the n-vector $\mathbf{v} = (5, 2, 2, 6, 0)$ represents the amount of each resource in a community's inventory. We also use n-vectors to represent the amount of each resource required to complete particular projects: $\mathbf{r} = (0, 0, 1, 1, 0)$ for building a road; $\mathbf{s} = (1, 1, 1, 1, 0)$ for building a settlement. The linear combination $\mathbf{v} - 2\mathbf{r} - \mathbf{s}$ represents the amount of each resource the community has in surplus (or shortfall) after building 2 roads and 1 settlement.

2. Consider a vocabulary consisting of n words in total. A "word count vector" for a document is an n-vector $\mathbf{v} = (v_1, \ldots, v_n)$ in which v_i is the number of times the ith vocabulary word appears in the document. Adding up word count vectors for all documents in a thematically coherent collection gives a histogram of word usage for the theme. See Figure 1.

3. Suppose a 3-second audio signal is sampled at 8000 hertz, meaning that measurements of acoustic pressure are taken at $n = 3 \times 8000$ regularly-spaced times within the 3-second duration. This results in an n-vector $\mathbf{x}$ of acoustic pressure measurements. If two such n-vectors $\mathbf{x}$ and $\mathbf{y}$, corresponding to two different audio signals, are added together, then the result $\mathbf{x} + \mathbf{y}$ represents the superposition of the two audio signals—i.e., what you would hear if both are played back at the same time. A different linear combination $\mathbf{x} + 2\mathbf{y}$ would represent the audio signal in which the second signal is twice as loud as the first. See Figure 2. ∎

basis	0	164
column	36	168
combination	27	64
execution	10	17
independence	48	147
leading	8	10
linear	90	429
matrix	35	512
orthogonal	0	85
space	0	172
vector	47	633
$\vdots$	$\vdots$	$\vdots$

Figure 1: Vocabulary words (left); word count vector for `dependence.tex` (center); sum of word count vectors for several documents (right).

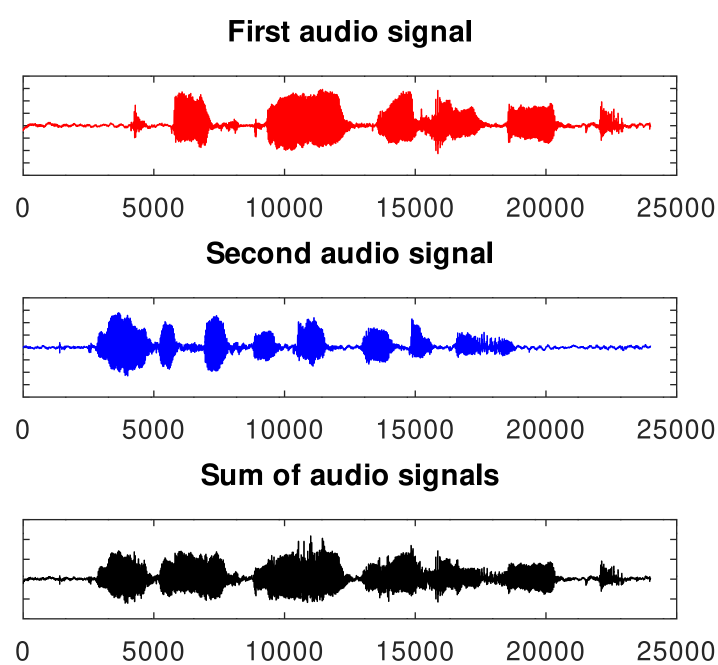

Figure 2: Plots of two 3-second audio signals and their sum.

Linear combinations are also at the heart of systems of linear equations[2]: given a collection of vectors $\mathcal{S} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots\}$, can the vector $\mathbf{w}$ be obtained as a linear combination of vectors from $\mathcal{S}$?

Example. Given $\mathbf{u} = (1, 2, 3)$ and $\mathbf{v} = (3, 4, 5)$, can the vector $\mathbf{w} = (4, 4, 4)$ be obtained as a linear combination of $\mathbf{u}$ and $\mathbf{v}$? In other words, are there real numbers x and y such that the following equation holds?

$$x \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + y \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 4 \end{bmatrix}$$

Equivalent question: Does the following system of linear equations (over the unknown variables x and y) have a solution?

$$\begin{cases} x + 3y &= 4 \\ 2x + 4y &= 4 \\ 3x + 5y &= 4. \end{cases}$$ ∎

Finding solutions to systems of linear equations is perhaps one of the most important motivating applications of linear algebra. There is an important (and ancient!) algorithm called "elimination" for systematically tackling these systems in full generality (for m equations in n unknowns, where both m and n can be very large), which will be studied in much detail later.

In our example, by hook or by crook, we can find the solution $(x, y) = (-2, 2)$, and it turns out this solution is the only one. Note that some systems of equations do not have any solutions, and in other systems, there are infinitely-many solutions! So it is special that this system has a unique solution. In any case, because we found a solution, we have verified that $\mathbf{w}$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

4 Span

The *span* of a collection of vectors is defined to be the set of *all* linear combinations of these vectors. So the question from the previous example can be

[2] A *linear equation* over a collection of variables $\mathcal{V}$ is an equation in which the left-hand side is a linear combination of variables from $\mathcal{V}$, and the right-hand side does not involve any variables from $\mathcal{V}$.

equivalently phrased as: Is $\mathbf{w}$ in the span of $\{\mathbf{u}, \mathbf{v}\}$?

However, we may also ask about the nature of the entire span of a collection of vectors, not just whether a particular vector is in it or not. For instance, for any collection of 3-vectors, the span could be:

- the entirety of $\mathbb{R}^3$ (i.e., all 3-vectors are in the span),
- a particular plane within $\mathbb{R}^3$,
- a particular line within $\mathbb{R}^3$, or
- a particular point within $\mathbb{R}^3$ (and if this is case, the point must be $\mathbf{0}$).

(Note that $\operatorname{span}(\emptyset) = \{\mathbf{0}\}$, because the "empty" sum is equal to 0.) And of course, we can ask the same for collections of m-vectors, where both m and size of the collection may be very large.

Example. The span of $\mathbf{u} = (1, 2, 3)$ and $\mathbf{v} = (3, 4, 5)$ is a plane in $\mathbb{R}^3$, and $\mathbf{w} = (4, 4, 4)$ happens to lie inside that plane. There are other 3-vectors that do not lie inside this plane, such as $(2, 1, 1)$. To see this, we consider the following system of linear equations:

$$\begin{cases} x + 3y &= 2 \\ 2x + 4y &= 1 \\ 3x + 5y &= 1. \end{cases}$$

The elimination algorithm would reveal that this system has no solutions. Without elimination, we make a lucky guess to try adding the first and third equations together, and then subtracting twice the second equation:

$$(x + 3y) + (3x + 5y) - 2 \cdot (2x + 4y) = 2 + 1 - 2 \cdot 1.$$

This simplifies to

$$0\,x + 0\,y = 1,$$

which is impossible (as $0 \neq 1$), so the system of linear equations has no solutions. ∎

5 Solution spaces

Vectors can also be to define linear equations another way. We explain with an example.

The 3-vector $\mathbf{v} = (3, 4, 5)$ defines the single linear equation

$$3x + 4y + 5z \ = \ 0$$

(with x, y, and z as unknown variables). Since the right-hand side of the equation is 0, it is called a *homogeneous equation*, and a solution to this equation is also a 3-vector. But there are many solutions to this equation: $(-4, 3, 0)$ is one, and another one is $(-5, 0, 3)$. In fact, there is an entire (two-dimensional) plane of solutions: all infinitely-many vectors obtained by linear combinations of $(-4, 3, 0)$ and $(-5, 0, 3)$.

Now consider both 3-vectors $\mathbf{u} = (1, 2, 3)$ and $\mathbf{v} = (3, 4, 5)$; together, they define the system of linear equations (again with unknown variables x, y, z):

$$\begin{cases} x + 2y + 3z &= 0 \\ 3x + 4y + 5z &= 0. \end{cases} \tag{1}$$

What is the nature of the solution space? It turns out it is a (one-dimensional) line: all multiples of $(1, -2, 1)$.

It is no coincidence that the solutions to the system of linear equations (1) is a line and the vectors $\mathbf{u}$ and $\mathbf{v}$ span a plane:

$$\underbrace{\text{dim. of solution space}}_{1} + \underbrace{\text{dim. of span}(\{\mathbf{u}, \mathbf{v}\})}_{2} \ = \ \underbrace{\text{number of variables}}_{3}.$$

This is an example of a more general rule that we'll explore later.

6 Linear transformations

A *function* T from a set $\mathcal{X}$ to another set $\mathcal{Y}$ assigns to each element of $\mathcal{X}$ exactly one element of $\mathcal{Y}$. Here, $\mathcal{X}$ is the *input space* (a.k.a. *domain*) of T, and $\mathcal{Y}$ is the *output space* (a.k.a. *target space*, *co-domain*) of T. We "declare" the function T with its input and output spaces by writing "$T \colon \mathcal{X} \to \mathcal{Y}$". Synonyms for "function" include *map* and *transformation*. If $\mathcal{X} = \mathcal{Y}$, then we also use the term *operator*; if $\mathcal{Y} = \mathbb{R}$, then we also use the term *functional* (in noun form).

Consider a transformation $T \colon \mathbb{R}^n \to \mathbb{R}^m$; the inputs to T are n-vectors, and the outputs of T are m-vectors. (Here, n and m are possibly different positive integers.) We say the transformation T is *linear* if T satisfies the following two properties:

Additivity: $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all n-vectors $\mathbf{u}$ and $\mathbf{v}$.

Homogeneity: $T(c\,\mathbf{v}) = c\,T(\mathbf{v})$ for all n-vectors $\mathbf{v}$ and scalars c.

Examples of transformations that are linear.

- $T \colon \mathbb{R}^2 \to \mathbb{R}$ defined by $T(x, y) = 3x + 2y$.
- $T \colon \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(x, y) = (-y, x)$.
- $T \colon \mathbb{R}^5 \to \mathbb{R}$ defined by $T(v_1, \dots, v_5) = v_1 + 1.6v_2 + 1.3v_3 + 1.7v_4 + 1.9v_5$ (inventory value; coefficients specify values of resource types).
- $T \colon \mathbb{R}^2 \to \mathbb{R}^5$ defined by

$$T(x, y) \;=\; \begin{bmatrix} \text{units of wool} & \text{needed for } x \text{ roads \& } y \text{ settlements} \\ \text{units of grain} & \text{needed for } x \text{ roads \& } y \text{ settlements} \\ \text{units of lumber} & \text{needed for } x \text{ roads \& } y \text{ settlements} \\ \text{units of brick} & \text{needed for } x \text{ roads \& } y \text{ settlements} \\ \text{units of ore} & \text{needed for } x \text{ roads \& } y \text{ settlements} \end{bmatrix},$$

  allowing for fractional roads and settlements.
- $T \colon \mathbb{R}^n \to \mathbb{R}^{n/2}$ defined by $T(x_1, x_2, \dots, x_n) = (x_1, x_3, x_5, \dots, x_{n-1})$. This halves the sampling rate of an audio signal (for n even). ∎

Examples of transformations that are not linear.

- $T \colon \mathbb{R} \to \mathbb{R}$ defined by $T(x) = 2x + 1$ (but almost ...).
- $T \colon \mathbb{R}^2 \to \mathbb{R}$ defined by $T(x, y) = \sqrt{x^2 + y^2}$.
- $T \colon \mathbb{R}^n \to \mathbb{R}^n$ defined by $T(x_1, \dots, x_n) = (e^{x_1}, \dots, e^{x_n})$. ∎

Linearity is such a special property that it gives rise to the entire sub-area of mathematics called "linear algebra": the study of linear transformations.

What makes linearity special? Suppose you find a collection of n-vectors $\mathbf{v}_1, \dots, \mathbf{v}_d$ on which two linear transformations T and U agree: $T(\mathbf{v}_i) = U(\mathbf{v}_i)$ for each $i \in \{1, \dots, d\}$. Then, T and U agree on all vectors in the span—i.e., $T(\mathbf{v}) = U(\mathbf{v})$ for all $\mathbf{v} \in \operatorname{span}(\{\mathbf{v}_1, \dots, \mathbf{v}_d\})$. If either T or U was not linear, then such a conclusion would not be justified! Another amazing property: if T is a linear transformation is invertible, then its inverse T^{-1} is also linear.

Many (if not most) transformations that you might encounter in the "real world" are not linear. But in many cases, we can approximate a non-linear transformation by a linear transformation (cf. differential calculus). In these cases, we may instead use (or reason about) the approximating linear transformation instead, at least for inputs where the approximation is known to be accurate.