

Introduction to Vectors

In the last note, we talked about systems of linear equations and tomography. Now, we are going to talk about a very useful way of conceptualizing things.

What is a vector? Vectors can be used to help solve a system of equations on your math homework, but they are also used to represent a multitude of things. For example, in the tomography example, the first column of the matrix describing light absorption is a *vector* that represents the light absorbed by each bottle in that stack. To begin talking about vectors, let's actually first define what a scalar is.

Definition 2.1 (Scalar): A **scalar** is a number. In mathematics and physics, they can be used to describe magnitude or used to scale things (e.g. halve every element of a vector. Another example is flipping all the signs which we will interpret as multiplying by -1).

Remark. Often, vectors are represented as letters in boldface (\mathbf{x}), or with a small arrow on top (\vec{x}). In these lecture notes we will use the latter arrow notation, and the size of the vector will be given (for example, $\vec{x} \in \mathbb{R}^3$).

Suppose we are given a collection of n real numbers, x_1, x_2, \dots, x_n . We can represent this collection as a single point in an n -dimensional space, denoted:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad (1)$$

Example 2.1 (3-D Vector): For $n = 3$, we could have $x_1 = -1, x_2 = 3.5, x_3 = 0$, and $\vec{x} = \begin{bmatrix} -1 \\ 3.5 \\ 0 \end{bmatrix}$. This vector represents a point in 3-D space.

Definition 2.2 (Vector): $\vec{x} \in \mathbb{R}^n$, where $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, is called a **vector**. In this case, we say it is a vector in \mathbb{R}^n . Each x_i (for i between 1 and n) is called a **component**, or **element**, of the vector. The **size** of a vector is the number of components it contains. Two vectors \vec{x} and \vec{y} are said to be **equal**, $\vec{x} = \vec{y}$, if they have the same size, and $x_i = y_i$ for all i .

Example 2.2 (Different kinds of vectors): Vectors can be much more general than the popular $\vec{x} \in \mathbb{R}^n$. For example,

(a) $\vec{x} \in \mathbb{C}^n$: Here, x_i is a complex number.

(b) $\vec{x} \in \mathbb{R}^{m \times n}$: Here, \vec{x} is an $m \times n$ array with real entries, we can also think of it as an array with the rows stacked on top of each other.

$$\vec{x} = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix} \Rightarrow \begin{bmatrix} x_{11} \\ \vdots \\ x_{mn} \end{bmatrix} \quad (2)$$

(c) Functions are also vectors! For example, x could be a function over the integers. We usually write $x(t)$ as “ x as a function of t ”, where $t \in \mathbb{Z}$ and $x(0), x(1)$, etc. are components of the vector:

$$\vec{x} = \begin{bmatrix} \vdots \\ x(-1) \\ x(0) \\ x(1) \\ \vdots \end{bmatrix} \quad (3)$$

Note that if the function is over all integers, the vector representing it has an infinite number of components! We call such vectors *infinite-dimensional*, and they will be come in handy when we learn signal processing. But for the most part, in this course when we say “vector” we will mean a finite-dimensional vector.

One way to use vectors is to represent the *state* of something.

Definition 2.3 (State): The minimum information you need to completely characterize a system at a given point in time, without any need for more information about the past of the system.

State is a really powerful concept because it lets us separate the past from the future. The state completely captures the present — and the past can only affect the future through the present.

Example 2.3 (Quadrotor): The 3D position, angle, velocity, and angular velocity of a quadrotor (Figure 1) at a particular time can be represented as - you guessed it - a vector $\vec{X} \in \mathbb{R}^{12}$.

Example 2.4 (Color): This is relevant to the lab we will do! The vector

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \vec{x} \in \mathbb{R}^3 \quad (4)$$

can represent a color, with its components giving the red, green, and blue (RGB) intensity values. Typically \vec{x} will fall in a range that reflects the range of the color (0 - 255).

However, the reality of color is quite subtle and involves psychophysics as well as ideas that are related to ideas of sampling that we will talk about in 16B.

Example 2.5 (Position): $\vec{x} \in \mathbb{R}^n$ can represent the sample values of a quantity at n time points. Imagine a car moving along a line. The positions at time t_1, t_2, \dots, t_n can be represented in a vector:

$$\vec{x} = \begin{bmatrix} x_{t_1} \\ x_{t_2} \\ \vdots \\ x_{t_n} \end{bmatrix} \quad (5)$$

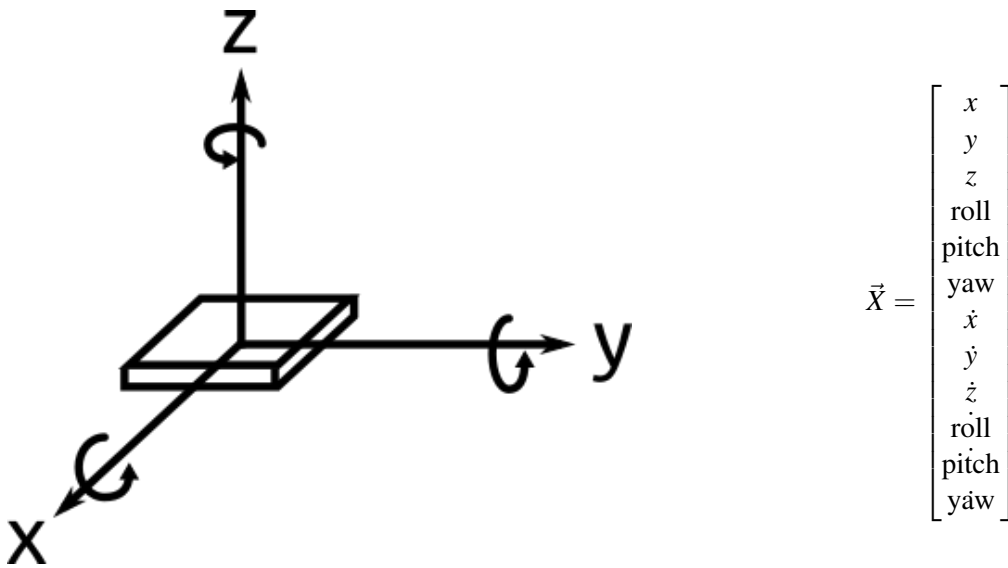


Figure 1: A quadrotor and the vector containing the information needed to determine its state. The dot above the components represents the derivative with respect to time.

Here, x_{t_i} represents the position at time t_i .

Example 2.6 (Image): A black and white image of $m \times n$ pixels can be represented by a vector of length mn , with the components giving greyscale levels at the pixel location.

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} \Rightarrow \begin{bmatrix} x_{11} \\ \vdots \\ x_{1n} \\ x_{21} \\ \vdots \\ x_{2n} \\ \vdots \\ \vdots \\ x_{m1} \\ \vdots \\ x_{mn} \end{bmatrix}$$

What about a color image? The image could be stored in a vector in \mathbb{R}^{3mn} , because each of the mn pixels has values for red, green, and blue.

What about a color video? This could be stored in a vector in $\mathbb{R}^{3mn(\text{number of frames in the video})}$.

Special Vectors

Definition 2.4 (Zero Vector): A **zero vector** is a vector with all the components equal to zero, usually just represented as $\vec{0}$. You can usually tell the size of the zero vector $\vec{0}$ from the context: if $\vec{x} \in \mathbb{R}^n$ is added to $\vec{0}$, then $\vec{0}$ must also be in \mathbb{R}^n .

Definition 2.5 (Standard Unit Vector): A **standard unit vector** is a vector with all components equal to 0 except for one element, which is equal to 1. A standard unit vector where the i th position is equal to 1 is

written as \vec{e}_i . We can denote the 3 standard unit vectors in \mathbb{R}^3 as:

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \vec{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (6)$$

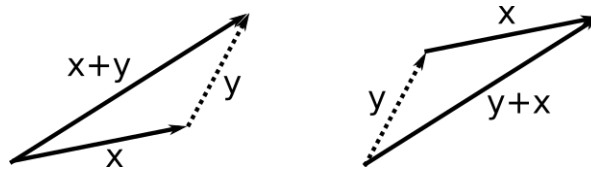
When talking about standard unit vectors in the context of states, we might also use the word “pure” to refer to such states. This is because they only have one kind of thing in them. Other states are mixtures of pure states.

Vector Addition

Two vectors of the same size and in the same space (e.g. complex numbers, real numbers, etc.) can be added together by adding their corresponding components, for example:

$$\begin{bmatrix} -1 \\ 3.5 \\ 0 \end{bmatrix} + \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} -1+2 \\ 3.5-1 \\ 0+3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2.5 \\ 3 \end{bmatrix} \quad (7)$$

In \mathbb{R}^n , you place the first vector’s tail at the origin, and attach the second vector’s tail to the first vector’s head. The origin to the second vector’s head is the resulting vector.



Properties of Vector Addition: for $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$

commutative	$\vec{x} + \vec{y} = \vec{y} + \vec{x}$
associative	$(\vec{x} + \vec{y}) + \vec{z} = \vec{x} + (\vec{y} + \vec{z})$
zero	$\vec{x} + \vec{0} = \vec{x}$
additive inverse	$\vec{x} + (-\vec{x}) = \vec{0}$

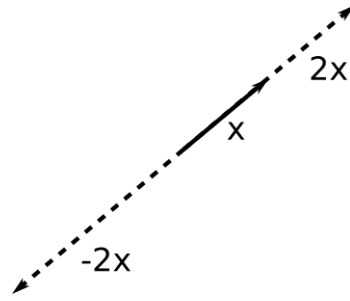
Scalar Multiplication

We can multiply a vector by a number, called a **scalar**, that is in the same space as the vector (e.g. also in the reals or in the complex numbers). Just multiply each of the components of the vector by the scalar:

$$(-3) \begin{bmatrix} -1 \\ 3.5 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ -10.5 \\ 0 \end{bmatrix}$$

In general, this looks like

$$\alpha \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{bmatrix}$$



Example 2.7 (Negative Vector): $-\vec{x}$, where $\vec{x} \in \mathbb{R}^n$, is just $-1 \times \vec{x}$ where -1 is a scalar.

Example 2.8 (Zero Vector): $0\vec{x} = \vec{0}$

Properties of Scalar Multiplication

$$\begin{array}{ll} \text{associative} & (\alpha\beta)\vec{x} = \alpha(\beta\vec{x}) \\ \text{distributive} & (\alpha + \beta)\vec{x} = \alpha\vec{x} + \beta\vec{x} \\ \text{identity} & 1\vec{x} = \vec{x} \end{array} \quad (8)$$

Vector Transpose

The transpose of a vector is crucial for vector and matrix calculations. We write the transpose of vector \vec{x} as \vec{x}^T . If $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, then $\vec{x}^T = [x_1 \ \cdots \ x_n]$. This is sometimes used so that writing down long column vectors doesn't take up vertical space.

The transpose of a column vector is a row vector. The transpose of a row vector is a column vector.