EE951: Linear Algebra **eMasters**

Chapter 1: Vectors

Author: Ketan Rajawat

Contents

1.1 Introduction

A vector is typically represented as an ordered finite list of numbers. Each number in the list is called a component or coordinate of the vector. The order of the components is significant as it determines the vector's position and direction in space. For example, a vector with three components might be denoted as follows:

$$\mathbf{v} = \begin{bmatrix} 1.2 \\ 2.1 \\ -1.8 \end{bmatrix}$$

The vector is enclosed within square brackets or curved brackets to distinguish it from other mathematical entities. The symbol $\mathbf{v}$ represents the vector and its *elements* (also called *entries* or *components*) are the values in the array.

Note that vectors can also be represented as horizontal arrays, called row vectors, using square brackets or curved brackets. For instance:

$$\mathbf{v} = \begin{bmatrix} 1.2 & 2.1 & -1.8 \end{bmatrix}$$

The choice of vertical or horizontal representation depends on convenience and the context in which the vectors are used. By default, we will always consider column vectors.

We often use symbols to denote vectors. In the example above, the ith element of the vector $\mathbf{v}$ is denoted as v_i, where the subscript i is an integer index that runs from 1 to 3. The elements in a vector are also referred to as scalars. In most applications, we will focus on vectors where the scalars are real numbers. In such cases, we denote them as real vectors. However, there are instances where other types of scalars, such as complex numbers, may arise. In those situations, we label the vector as a complex vector. In practice, the elements can represent various quantities, such as spatial coordinates, forces, velocities, or any other measurable quantities in a given context.

The *size* or *dimension* of the vector is the number of elements it contains. In the example above, the size $\mathbf{v}$ is 3. A vector of size n is called a n-vector. The set of all real numbers is denoted by $\mathbb{R}$. Likewise, the set of all real n-vectors is represented as $\mathbb{R}^n$. Therefore, stating that $\mathbf{v} \in \mathbb{R}^n$ implies that vector $\mathbf{v}$ has real entries and has size n.

1.2 Block or Stacked Vectors

Block vectors, also known as stacked vectors, provide a useful way to organize and represent multiple vectors together. By stacking individual vectors vertically, we create a block vector that encapsulates several vectors as its components. This concept is particularly helpful when dealing with systems of equations or combining vectors with different properties or dimensions.

To illustrate this idea, let us consider an example involving two vectors, $\mathbf{u}$ and $\mathbf{v}$:

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

We can combine these vectors into a single block vector, denoted as $\mathbf{w}$, by stacking them vertically:

$$\mathbf{w} = \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

Here, $\mathbf{w}$ is a block vector formed by stacking $\mathbf{u}$ above $\mathbf{v}$. By doing so, we can treat $\mathbf{w}$ as a single entity, allowing us to perform operations on multiple vectors simultaneously.

Block vectors can have more than two blocks and blocks with different sizes. For instance, consider three vectors, $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$:

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

We can form a block vector $\mathbf{d}$ by stacking these vectors vertically:

$$\mathbf{d} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \\ \mathbf{c} \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ b_1 \\ b_2 \\ c_1 \\ c_2 \end{bmatrix}$$

Block vectors offer a convenient way to organize and manipulate collections of vectors, facilitating operations such as addition, scalar multiplication, and solving systems of equations.

By using block vectors, we can combine and treat multiple vectors as a single entity, enabling us to express complex relationships and structures more efficiently. This concept becomes especially powerful when working with large systems or dealing with diverse sets of vectors with varying dimensions or properties.

1.3 Subvectors

We refer to vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ as subvectors of $\mathbf{d}$, each having sizes 3, 2, and 2, respectively. This notion of subvectors allows us to extract and work with specific portions of a block vector. To denote subvectors, we employ colon notation, which indicates a range of indices. For example, suppose we have a vector $\mathbf{a}$ with k components:

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_k \end{bmatrix}$$

To denote a subvector that includes elements from index i to index j (inclusive) for $j \geq i$, we use $\mathbf{a}_{i:j}$. This notation signifies the subvector of $\mathbf{a}$ consisting of $j - i + 1$ elements, i.e.,

$$\mathbf{a}_{i:j} = \begin{bmatrix} a_i \\ a_{i+1} \\ \vdots \\ a_j \end{bmatrix}. \tag{1.1}$$

It allows us to isolate and manipulate specific sections of the block vector conveniently.

For example, if we consider the block vector $\mathbf{w}$ from the earlier illustration:

$$\mathbf{w} = \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

We can express subvectors of $\mathbf{w}$ using colon notation. Suppose we want to extract the subvector $\mathbf{b}$ consisting of elements u_2, u_3, and v_1 from $\mathbf{w}$. We can write it as $\mathbf{b} = \mathbf{w}_{2:4}$, where the indices 2 and 4 represent the start and end positions, respectively.

The ability to work with subvectors allows us to isolate relevant components of a block vector for computations or analysis, providing flexibility and convenience when dealing with complex systems and multi-dimensional data.

Remark 1. Some programming languages, such as C, C++, Java, Python, and many others (but not MATLAB), adopt 0-based indexing, where the first element of a vector or array is referenced with an index of 0. For example, in these languages, the elements of a vector $\mathbf{a}$ would be accessed as $a[0]$, $a[1]$, $a[2]$, and so on. It is crucial to be aware of this distinction when translating mathematical algorithms or equations into code or when working with code that involves indexing operations.

1.4 Zero Vector

The zero vector, denoted as $\mathbf{0}_n$, is a special vector with all its components equal to zero. The subscript n represents the size or dimension of the vector. For example, the zero vector $\mathbf{0}_3 \in \mathbb{R}^3$ would be represented as:

$$\mathbf{0}_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The zero vector serves as an additive identity in vector operations. When added to any vector $\mathbf{v}$, it does not change the vector:

$$\mathbf{v} + \mathbf{0}_n = \mathbf{v}$$

The size or dimension of the zero vector, n, can sometimes be inferred from the context. If the size of the vector is clear from the surrounding discussion or equations, the subscript n may be omitted. For instance, if we are working in $\mathbb{R}^3$ and mention $\mathbf{0}$ without specifying the subscript, it is understood to be $\mathbf{0}_3$. Likewise, if $\mathbf{v} \in \mathbb{R}^n$, then the expression

$$\mathbf{v} + \mathbf{0} \tag{1.2}$$

refers to $\mathbf{0}_n$. This convention simplifies notation when the size can be deduced from the context.

1.5 Standard Unit Vectors

A unit vector is a vector that has a length or magnitude of 1. Unit vectors are particularly useful for indicating direction and can be employed to describe vectors in a normalized form.

Let's consider a concrete example in three dimensions. In $\mathbb{R}^3$, we have three principal unit vectors, often denoted as $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$, which correspond to the standard Cartesian coordinate axes.

The unit vector $\mathbf{i}$ points in the direction of the positive x-axis, $\mathbf{j}$ points in the direction of the positive y-axis, and $\mathbf{k}$ points in the direction of the positive z-axis. These vectors have magnitudes of 1 and can be represented as follows:

$$\mathbf{i} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{j} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{k} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

The unit vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ are also known by alternative names such as the basis vectors, standard unit vectors, or coordinate vectors. They form a basis for the three-dimensional Cartesian coordinate system and can be used to express any vector in $\mathbb{R}^3$ as a linear combination of these unit vectors.

In the general case of n dimensions, standard unit vectors are denoted as $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$, where each unit vector $\mathbf{e}_i$ corresponds to the i-th dimension and given by

$$(\mathbf{e}_i)_j = \begin{cases} 1 & j = i \\ 0 & j \neq i \end{cases} \tag{1.3}$$

for all $1 \leq i, j \leq n$.

1.6 All-One Vector

The all-one vector or the ones vector is a special vector that consists of all its elements equal to 1. It is commonly denoted as $\mathbf{1}_n$ to indicate its size or dimension, where the subscript n represents the number of elements in the vector.

For example, in $\mathbb{R}^3$, the ones vector $\mathbf{1}_3$ can be represented as:

$$\mathbf{1}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

The ones vector is often used in various mathematical operations and computations. It can serve as a convenient tool for summing or scaling other vectors.

1.7 Sparse Vectors

Sparse vectors are those that contain several zeros. The sparsity pattern of a vector is set of indices of its non-zero entries. The number of non-zero entries of a vector $\mathbf{x}$ is denoted by $\mathrm{nnz}(\mathbf{x})$. For instance, $\mathrm{nnz}(\mathbf{e}_i) = 1$ and $\mathrm{nnz}(\mathbf{0}) = 0$.

1.8 Applications of Vectors

1.8.1 Location

A location vector represents the position of an object or point in space relative to a reference point or coordinate system. By assigning a vector to a specific location, we can specify both the magnitude and direction of the position. For example, in two-dimensional space, we can represent a point with coordinates (x, y) as a position vector:

$$\mathbf{r} = \begin{bmatrix} x \\ y \end{bmatrix}$$

This location vector describes the displacement from the origin (0, 0) to the point (x, y). It encapsulates the spatial information necessary to identify the object's location within the coordinate system.

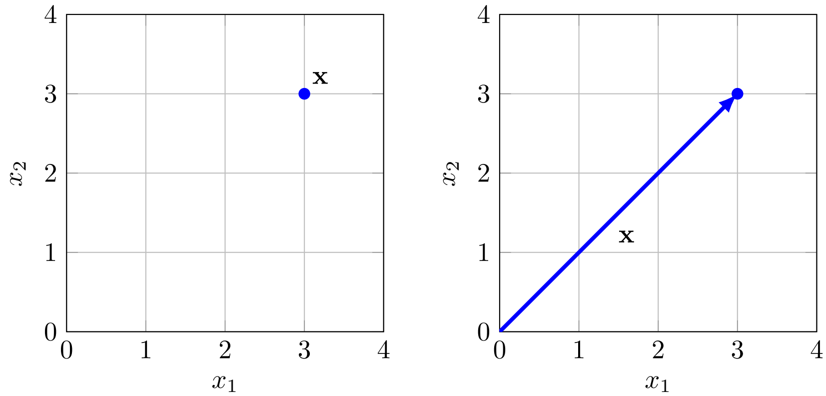


Figure 1.1: Left: Vector **x** specifies the position (shown as a dot) with coordinates $(3, 3)$. Vector **x** represents a displacement from $(0, 0)$ by 3 units along x_1 and 3 units along x_2.

1.8.2 Displacement

Displacement vectors, on the other hand, convey the change in position of an object or point. They represent the difference between two location vectors, indicating the movement from one position to another. For instance, consider two position vectors representing the initial and final locations of an object:

$$\mathbf{r}_1 = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \quad \text{and} \quad \mathbf{r}_2 = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$$

The displacement vector, denoted as $\Delta\mathbf{r}$, is obtained by subtracting the initial position vector from the final position vector:

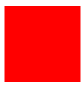
$$\Delta\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1 = \begin{bmatrix} x_2 - x_1 \\ y_2 - y_1 \end{bmatrix}$$

This displacement vector represents the change in position and provides information about the direction and magnitude of the movement.

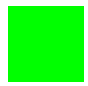
1.8.3 Color

Vectors can also be employed to denote colors in various color models. One commonly used color model is the RGB (Red, Green, Blue) model, which defines colors by specifying the intensity of red, green, and blue components. By combining different intensities of these primary colors, we can create a wide range of colors.

In the RGB model, each color component is typically represented by an integer value ranging from 0 to 255 or a normalized value between 0 and 1, indicating the intensity or brightness of that particular color. To represent a color using a vector, we can use a three-dimensional vector where each element corresponds to the intensity of the red, green, and blue components, respectively.



RGB Vector: $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$



RGB Vector: $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

RGB Vector: $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

We can create various other colors by adjusting the intensity values of the red, green, and blue components. For example, to represent the color yellow, we can use the RGB vector $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$:

Using RGB vectors, we can precisely specify and manipulate colors, making them a valuable tool in various applications, including computer graphics, image processing, and web design.

1.8.4 Image

A monochrome image can be represented as a vector by considering each pixel in the image as an element of the vector. The elements of the vector correspond to the grayscale intensity value of the pixels. By arranging these intensity values sequentially, we obtain a vector representation of the image.

The size of the vector is determined by the resolution of the image. For instance, if we have a monochrome image with a resolution of 800 pixels in width and 600 pixels in height, the corresponding vector would have a size of 480,000 (800×600) elements. Each element of the vector represents the grayscale intensity of a specific pixel, ranging from 0 (black) to 255 (white), with intermediate values representing various shades of gray.

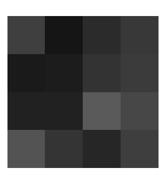


Figure 1.2: Example of a 4×4 grayscale image.

A similar idea can be used to represent color images by specifying the RGB values of each pixel. For example, if we have a color image with the same resolution as the previous monochrome image (800×600), the corresponding vector would have a size of 1,440,000 ($800 \times 600 \times 3$) elements. In a video, if each frame is represented as an vector, the stacked vector can be used to represent the entire video.

1.8.5 Word counts

Word counts from a piece of text can be represented as a vector. This representation allows us to analyze and compare texts using mathematical operations. Each element of the vector corresponds to a specific word, and its value represents the frequency or count of that word in the text.

To illustrate this concept, let us consider a simple example. Suppose we have two texts: Text A and Text B. We want to represent the word counts of these texts as vectors. Let us assume we have a vocabulary of three words: "apple," "banana," and "orange."

Text A: "I like apple and banana."

Text B: "She prefers banana and orange."

To represent the word counts of these texts, we can create vectors where each element corresponds to a word in the vocabulary. The value of each element indicates the frequency of that word in the text. Using the given example, the vectors representing the word counts would be:

Text A: $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ (One "apple," one "banana," and zero "orange.")

Text B: $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ (Zero "apple," one "banana," and one "orange.")

By representing the word counts as vectors, we can perform various mathematical operations on them. For instance, we can measure the similarity between texts using techniques like dot product, cosine similarity, or Euclidean distance. These are widely used in natural language processing, information retrieval, and text mining.

The concept of representing word counts as vectors relates to the concept of a histogram. In a histogram, the frequency of each word is represented by the height of a corresponding bar. Similarly, in our word count vectors, the value of each element represents the frequency or count of a specific word. Therefore, the word count vector can be seen as a form of histogram, where the frequencies are organized in a structured manner for mathematical analysis.

1.8.6 One-hot coding

One-hot coding is a binary representation where each element in a vector corresponds to a unique category or label, and only one element is "hot" or "on" (represented by 1), while all other elements are "cold" or "off" (represented by 0).

One-hot coding is often used when dealing with categorical data, such as class labels, attributes, or features that have distinct categories. It allows us to represent categorical variables as vectors in a format suitable for mathematical computations. Let us consider a simple example to illustrate one-hot coding. Suppose we have a dataset of fruits, and we want to represent their types: "apple," "banana," and "orange." We can create one-hot vectors where each element corresponds to a specific fruit type. If the dataset contains three fruits, the one-hot vectors representing their types would be:

Apple: $\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$ (The first element is 1 to indicate it is an apple, while the other elements are 0.)

Banana: $\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$ (The second element is 1 to indicate it is a banana, while the other elements are 0.)

Orange: $\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$ (The third element is 1 to indicate it is an orange, while the other elements are 0.)

By using one-hot coding, we can easily represent categorical data as vectors, where each element captures the presence or absence of a particular category. One-hot coding is particularly useful when applying machine learning algorithms or performing mathematical operations on categorical variables. Furthermore, one-hot coding is akin to creating "indicator variables" or "dummy variables" in statistics. It allows us to convert categorical information into a numerical format that can be used for analysis and modeling.

1.8.7 Other Examples

Other examples of vectors arising in various applications include resources, portfolio, time-series, proportions, cash flow, and features.