Learning objectives:

- ☐ calculate the dot product and cross product of two vectors,
- ☐ perform operations on matrices and vectors such as addition and multiplication,
- ☐ prove identities that involve matrices using induction.

1 Vectors

A **vector** is an element of some abstract space, called a **vector space** where certain *axioms* need to be satisfied (see below). When thinking about vectors, it might make sense to think about an arrow in $2d$, which has x and y components. The $2d$ plane is the vector space, and any arrow that we draw in the plane would be a vector in that space. Vectors are often written with an arrow overhead (e.g. $\vec{v}$) but will often be written in boldface when typed (e.g. $\mathbf{v}$). We will typically represent vectors as an array of numbers stacked in a column.

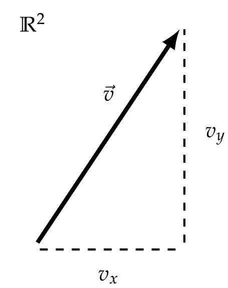

Figure 1: A two-dimensional vector $\vec{v} \in \mathbb{R}^2$ with x-component v_x and y-component v_y.

1.1 Axioms of a vector space

Let $\mathcal{V}$ be a vector space and let $\vec{v}_1,\ \vec{v}_2,\ \vec{v}_3,\ \dots$ satisfy

- $\vec{v}_1 + \vec{v}_2 = \vec{v}_2 + \vec{v}_1 \in \mathcal{V}$.
- $\vec{v}_1 + (\vec{v}_2 + \vec{v}_3) = (\vec{v}_1 + \vec{v}_2) + \vec{v}_3$.
- $\mathcal{V}$ contains the null element:
  $\vec{v} + \vec{0} = \vec{v},\ \forall \vec{v} \in \mathcal{V}$.
- Each element of $\mathcal{V}$ has an inverse with respect to addition:
  $\vec{v} + (-\vec{v}) = \vec{0},\ \forall \vec{v} \in \mathcal{V}$.

Furthermore, let $\alpha,\ \beta,\ \gamma,\ \dots$ be scalars. Then if

- $\alpha\vec{v} \in \mathcal{V}$,
- $(\alpha + \beta)\vec{v} = \alpha\vec{v} + \beta\vec{v}$,
- $\alpha(\vec{v}_1 + \vec{v}_2) = \alpha\vec{v}_1 + \alpha\vec{v}_2$,
- $\alpha\beta\vec{v} = \alpha(\beta\vec{v})$,
- $1\vec{v} = \vec{v}$,

then $\mathcal{V}$ is a vector space with the elements $\vec{v}_1,\ \vec{v}_2,\ \vec{v}_3,\ \dots$ being called vectors.

How do I add two vectors together?

When adding two vectors together, add the *components* of each vector to form a new vector whose components is the sum of the components of the two vectors you are adding. Also, when multiplying a vector by a scalar, multiply each component by the scalar to form the components of the new vector. See Example 1.2 for some practice.

1.2 Linear independence, basis and dimension (optional)

Consider a set of n vectors $\vec{v}_1,\ \vec{v}_2,\ \vec{v}_3,\ \ldots,\ \vec{v}_n$ (all of which are not the null vector $\vec{0}$), as well as the expression

$$c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 + \cdots + c_n\vec{v}_n = \vec{0}, \tag{1}$$

for some constants $c_1,\ c_2,\ c_3,\ \ldots,\ c_n \in \mathbb{R}$. If one of the scalars is nonzero, say c_1, then we can rewrite this as

$$\vec{v}_1 = -\frac{c_2}{c_1}\vec{v}_2 - \frac{c_3}{c_1}\vec{v}_3 - \cdots - \frac{c_n}{c_1}\vec{v}_n.$$

Since v_1 can be rewritten in terms of the other vectors, then the set of vectors are *linearly dependent*. That is, v_1 depends on the other vectors (since it can be written as a weighted sum of the other ones). However, if the only solution to Eq. 1 is $c_1 = c_2 = c_3 = \cdots = c_n = 0$, then $\vec{v}_1,\ \vec{v}_2,\ \vec{v}_3,\ \ldots,\ \vec{v}_n$ are called *linearly independent*. In this case, $\vec{v}_1,\ \vec{v}_2,\ \vec{v}_3,\ \ldots,\ \vec{v}_n$ form a *basis* for a vector space $\mathcal{V}$ and any vector $\vec{v} \in \mathcal{V}$ can be written as

$$\vec{v} = a_1\vec{v}_1 + a_2\vec{v}_2 + a_3\vec{v}_3 + \cdots + a_n\vec{v}_n,$$

for some constants $a_1,\ a_2,\ a_3,\ \ldots,\ a_n$ (some of which can be zero). The *dimension* of $\mathcal{V}$ is the number of linearly independent vectors that are needed to form an arbitrary vector $\vec{v} \in \mathcal{V}$. In this case, $\dim(\mathcal{V}) = n$ since there are n linearly independent vectors.

This set of linear independent vectors forms a *basis* for $\mathcal{V}$. We also say that the set of vectors $\vec{v}_1,\ \vec{v}_2,\ \vec{v}_3,\ \ldots,\ \vec{v}_n$ *spans* the vector space $\mathcal{V}$.

Is the basis unique?

Great question, Piper! Actually the basis is not unique and we can transform from one basis to another using a transformation (which involves a matrix - see below). For example, consider two vectors in two-dimensional space, $\mathbb{R}^2$: $\vec{v}_1 = [1, 0]^t$ and $\vec{v}_2 = [0, 1]^t$. I can apply a transformation that rotates these two vector by some arbitrary angle θ. Suppose this angle is $\theta = 45°$. Our rotated vectors could then be $\vec{u}_1 = [\sqrt{2}, \sqrt{2}]^t$ and $\vec{u}_2 = [-\sqrt{2}, \sqrt{2}]^t$. Sometimes, it may be easier to work in one basis over another, and we can perform a change of bases.

Example 1:

Let $\vec{u} = [2, 3]^t$, $\vec{v} = [-2, 1]^t$ and $\vec{w} = [-4, 2]^t$. Don't worry about the ^{t} for now (it just means these are column vectors) - see the next page for the *transpose*.

(a) Calculate $\alpha\vec{u} + \beta\vec{v}$ where $\alpha = 3$ and $\beta = 5$.

(b) Consider the space with the vectors $\vec{u},\ \vec{v},\ \vec{w}$. What is the dimension of this space? Are $\vec{u},\ \vec{v},\ \vec{w}$ linearly independent?

Solution:

(a) The weighted sum is: $3[2, 3]^t + 5[-2, 1]^t = [6, 9]^t + [-10, 5]^t = [-4, 14]^t$.

(b) Notice that $\vec{w} = 2\vec{v}$, therefore, $\vec{w}$ is linearly dependent on $\vec{v}$. The dimension of the vector space with vectors $\vec{u},\ \vec{v},\ \vec{w}$ is 2.

1.3 Inner (dot) product between vectors in $\mathbb{R}^n$

The **inner (dot) product** between two vectors $\vec{u} \in \mathbb{R}^n$ and $\vec{v} \in \mathbb{R}^n$, denoted by $\vec{u} \cdot \vec{v}$ or $\vec{u}^t\vec{v}$ is the *sum of all the multiplied components of* $\vec{u}$ *and* $\vec{v}$. Specifically,

$$\vec{u} \cdot \vec{v} = \vec{u}^t\vec{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n.$$

When the inner product between two vectors is zero, then the vectors are *orthogonal* or *perpendicular* to each other. Geometrically, if we glued the tails of two orthogonal vectors together, then there would be an angle of 90° between them. The inner product of a vector with itself is useful for calculating the **magnitude**, also called the **norm**, of the vector. In the geometric interpretation of vectors in Figure 1, the magnitude (norm) corresponds to the *length* of the arrow. For a general vector $\vec{u} \in \mathbb{R}^n$, the norm is denoted by $||\vec{u}||$ and is calculated as

$$||\vec{u}|| = \sqrt{\vec{u} \cdot \vec{u}} = \sqrt{\vec{u}^t\vec{u}} = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}.$$

What is that superscript t?

The superscript t on a vector, for example $\vec{u}^t$, denotes the *transpose*. The elements of a vector can be thought to be stacked vertically, like this

$$\vec{u} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

The transpose of $\vec{u}$ converts this column into a row: $\vec{u}^t = [2\ 3]$.

1.4 Cross product between vectors in $\mathbb{R}^3$

When you want to calculate the normal of a surface, like the surface defining Mike Wazowski, then we will need the *cross product*. The **cross product** of two vectors $\vec{u}$ and $\vec{v}$ creates a new vector $\vec{w}$ which is orthogonal to both $\vec{u}$ and $\vec{v}$. The cross product between two vectors $\vec{u} = [u_x, u_y, u_z]$ and $\vec{v} = [v_x, v_y, v_z]$ is denoted by $\vec{u} \times \vec{v}$ and is calculated as

$$\vec{u} \times \vec{v} = [u_yv_z - u_zv_y,\ u_zv_x - u_xv_z,\ u_xv_y - u_yv_x]$$

2 Matrices

A matrix is a rectangular array of numbers. It is handy if you want to describe *relationships* between things (like in graph theory), transformations on vectors (like in computer graphics), or if you want to solve some system of linear equations (which pretty much comes up anytime you have a set of linear equations and unknowns – this happens a lot!).

We will only consider square matrices here, in which the number of rows is equal to the number of columns. We will also only look at either 2-by-2 (2×2) or 3-by-3 (3×3) matrices, but the concepts extend to any n-by-n matrix. The *entries* of a matrix A are often subscripted with the row and column of that particular entry. For example, a 2×2 matrix A and 3×3 matrix B would respectively be denoted as

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}, \quad B = \begin{bmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{2,1} & b_{2,2} & b_{2,3} \\ b_{3,1} & b_{3,2} & b_{3,3} \end{bmatrix}.$$

3 Operations

3.1 Transpose

Just like when we took the transpose of a vector (a column) and obtained a row vector, the transpose of a matrix involves exchanging rows for columns and vice versa. For example,

$$B = \begin{bmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{2,1} & b_{2,2} & b_{2,3} \\ b_{3,1} & b_{3,2} & b_{3,3} \end{bmatrix}^t = \begin{bmatrix} b_{1,1} & b_{2,1} & b_{3,1} \\ b_{1,2} & b_{2,2} & b_{3,2} \\ b_{1,3} & b_{2,3} & b_{3,3} \end{bmatrix}.$$

Matrices that satisfy $B^t = B$ are called *symmetric*.

3.2 Multiplication

We'll look at two types of multiplication here. The first is the multiplication of a matrix with a vector. Matrix-vector multiplication is useful if you want to transform a vector, say with a scaling, translation, or rotation, which you may often encounter in computer graphics. For example, suppose you want to rotate a model of Mike Wazowski: you would construct a transformation matrix which represents the rotation about the center of Mike, and then do a matrix-vector multiplication for all the points on the surface of the geometry.

What if the sizes are different?

The sizes of whatever vectors or matrices you are multiplying need to match! In other words, the number of components in a vector should be equal to the number of columns in your matrix. For square matrices, both matrices need to have the same number of rows and columns.

Multiplying a $n \times n$ matrix with a vector (with n components) results in a vector with n components. Each of these components is the result of the inner product of a row of the matrix with the vector we are multiplying with. For example, multiplying a 3×3 matrix B with a three-dimensional vector $\vec{u}$ gives a new three-dimensional vector $\vec{v}$ as follows:

$$B\vec{u} = \begin{bmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{2,1} & b_{2,2} & b_{2,3} \\ b_{3,1} & b_{3,2} & b_{3,3} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} b_{1,1}u_1 + b_{1,2}u_2 + b_{1,3}u_3 \\ b_{2,1}u_1 + b_{2,2}u_2 + b_{2,3}u_3 \\ b_{3,1}u_1 + b_{3,2}u_2 + b_{3,3}u_3 \end{bmatrix} = \begin{bmatrix} \vec{b}_{r_1}^t \\ \vec{b}_{r_2}^t \\ \vec{b}_{r_3}^t \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} \vec{b}_{r_1}^t \vec{u} \\ \vec{b}_{r_2}^t \vec{u} \\ \vec{b}_{r_3}^t \vec{u} \end{bmatrix}$$

where $\vec{b}_{r_1}$, $\vec{b}_{r_2}$, $\vec{b}_{r_3}$ are vectors containing the entries in the rows of the matrix B (note that I have written these to be column vectors even though they are the rows of B, hence the transpose). This can also be written as

$$B\vec{u} = \begin{bmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{2,1} & b_{2,2} & b_{2,3} \\ b_{3,1} & b_{3,2} & b_{3,3} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = u_1 \begin{bmatrix} b_{1,1} \\ b_{2,1} \\ b_{3,1} \end{bmatrix} + u_2 \begin{bmatrix} b_{1,2} \\ b_{2,2} \\ b_{3,2} \end{bmatrix} + u_3 \begin{bmatrix} b_{1,3} \\ b_{2,3} \\ b_{3,3} \end{bmatrix} = u_1\vec{b}_{c_1} + u_2\vec{b}_{c_2} + u_3\vec{b}_{c_3}.$$

where $\vec{b}_{c_1}$, $\vec{b}_{c_2}$ and $\vec{b}_{c_3}$ are vectors containing the entries of the columns of B. Note that $B\vec{u}$ is in the space spanned by the columns of B.

Example 2:
Say we have two equations and two unknowns (x and y) as follows

$$\begin{aligned} 2x + 3y &= 8, \\ -4x + 5y &= 9. \end{aligned} \tag{2}$$

This system of equations can be represented as a matrix-vector product:

$$\underbrace{\begin{bmatrix} 2 & 3 \\ -4 & 5 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x \\ y \end{bmatrix}}_{\vec{u}} = \underbrace{\begin{bmatrix} 8 \\ 9 \end{bmatrix}}_{\vec{v}}.$$

With these definitions of A, $\vec{u}$ and $\vec{v}$, we can write this system of equations more concisely as $A\vec{u} = \vec{v}$. Solving for x and y then amounts to *inverting* A in order to obtain $\vec{u}$: $\vec{u} = A^{-1}\vec{v}$. We will not look at techniques for inverting matrices in this course.

It doesn't take much to extend our matrix-vector product method to multiplying two matrices together. In fact, multiplying two $n \times n$ matrices, A and B, results in a new $n \times n$ matrix, C, where each column corresponds to the matrix-vector product of the first matrix A with the columns of B.

We will restrict our attention to multiplying 2×2 matrices in this course. For 2×2 matrices A and B, this looks like:

$$C = AB = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{bmatrix} = \begin{bmatrix} (a_{1,1}b_{1,1} + a_{1,2}b_{2,1}) & (a_{1,1}b_{1,2} + a_{1,2}b_{2,2}) \\ (a_{2,1}b_{1,1} + a_{2,2}b_{2,1}) & (a_{2,1}b_{1,2} + a_{2,2}b_{2,2}) \end{bmatrix} = \begin{bmatrix} c_{1,1} & c_{1,2} \\ c_{2,1} & c_{2,2} \end{bmatrix}$$

Example 3:
Given the matrices A and B below, calculate AB and BA.

$$A = \begin{bmatrix} 3 & 5 \\ -1 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 8 & -4 \\ 3 & 1 \end{bmatrix}.$$

Solution:

$$AB = \begin{bmatrix} 3 & 5 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 8 & -4 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 39 & -7 \\ -2 & 6 \end{bmatrix}, \quad BA = \begin{bmatrix} 8 & -4 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & 5 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 28 & 32 \\ 8 & 17 \end{bmatrix}.$$

Note that these are **not** equal!

4 Proofs involving matrices

Let's now combine our knowledge of matrices and do a proof by induction!

Example 4:
Prove, using a proof by induction, that $A^n = B(n)$ for all $n \in \mathbb{N}$, where

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}, \quad B(n) = \begin{bmatrix} 1 & 1-2^n \\ 0 & 2^n \end{bmatrix}.$$

Proof. We use a proof by induction. Let the induction hypothesis be the predicate $p(n) = A^n = B(n)$ for the matrices A and $B(n)$ given above.

Base case: For $n = 1$, we have

$$A^1 = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}^1 = \begin{bmatrix} 1 & 1-2^1 \\ 0 & 2^1 \end{bmatrix} = B(1).$$

Inductive step: Assume $p(n)$ is true. That is, $A^n = B$ for some positive integer n. We will prove that $p(n+1)$ is true, meaning $A^{n+1} = B(n+1)$. Starting with A^{n+1} gives

$$\begin{aligned} A^{n+1} = A^n A &= \begin{bmatrix} 1 & 1-2^n \\ 0 & 2^n \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}, && \text{by } p(n) \\ &= \begin{bmatrix} 1 & (-1+2-2^{n+1}) \\ 0 & (2 \cdot 2^n) \end{bmatrix}, && \text{multiplying } A^n \text{ by } A \\ &= \begin{bmatrix} 1 & 1-2^{n+1}) \\ 0 & 2^{n+1} \end{bmatrix}, && \text{verifying } p(n+1). \end{aligned}$$

By induction on n, $p(n)$ is true.

□