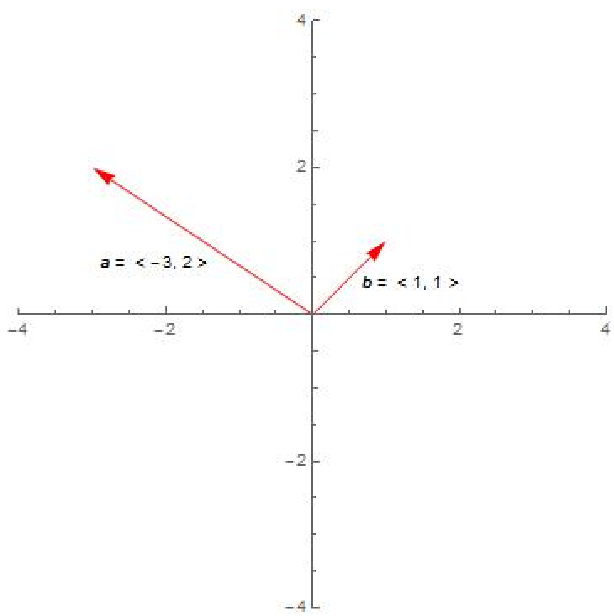
Introduction

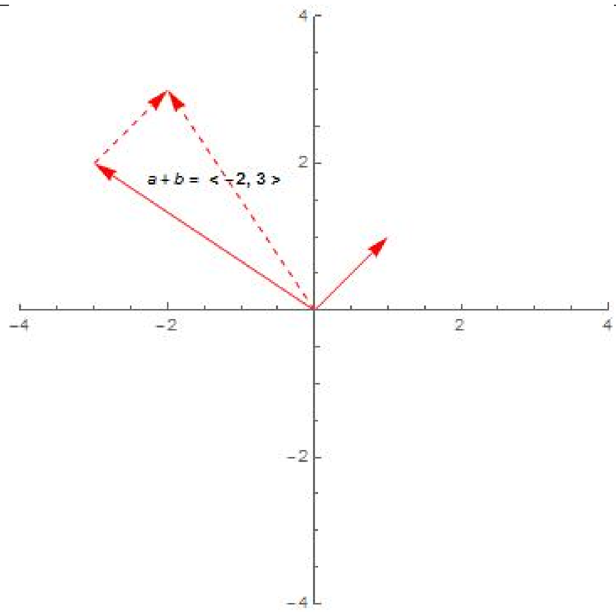
As indicated by the title of this course, our main topic of study will be vector spaces. Very informally, a vector space is a collection of objects called *vectors*, along with two operations: a way to add vectors, and a way to scale them. You actually already know several examples of *vector spaces*: the simplest one is $\mathbb{R}^2$, the xy plane.

Example 1: $\mathbb{R}^2$

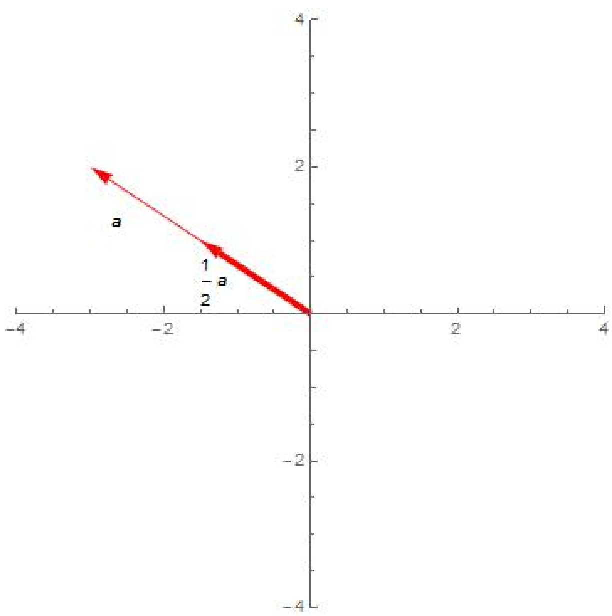
We can think of points in the xy plane as the tips of vectors whose tails start at the origin: for example, the points $(-3, 2)$ and $(1, 1)$ in the xy plane can be thought of as the vectors $\mathbf{a} = \langle -3, 2\rangle$ and $\mathbf{b} = \langle 1, 1\rangle$, graphed below:



Their sum $\mathbf{a} + \mathbf{b}$ is the vector $\mathbf{a} + \mathbf{b} = \langle -2, 3\rangle$ indicated below:



In addition, we can scale these vectors; $\frac{1}{2}\mathbf{a}$ is indicated below:



It turns out that there are many "spaces" whose structure and behaviour is practically identical to that which we have observed in $\mathbb{R}^2$. Below we record another simple example of a vector space.

Example 2: $P_1(\mathbb{R})$

The space $P_1(\mathbb{R})$ consists of all real valued polynomials of degree up to 1, i.e. all polynomials of the form

$$ax + b,$$

where a and b are real numbers. We think of each polynomial of this form as a vector; for example, let $\mathbf{f}$ and $\mathbf{g}$ be the vectors so that

$$f(x) = -3x + 2 \text{ and } g(x) = x + 1.$$

As with $\mathbb{R}^2$, we have two operations on this space: we can add vectors using normal polynomial addition, so the vector $\mathbf{f}+\mathbf{g}$ is the polynomial so that

$$(f + g)(x) = -2x + 3.$$

We can scale these vectors using normal multiplication; for example, $\frac{1}{2}\mathbf{f}$ is the vector so that

$$\frac{1}{2}f(x) = -\frac{3}{2}x + 1.$$

You may have notice that the example of addition in P_1 is eerily similar to that presented in $\mathbb{R}^2$–indeed, we saw that

$$\langle -3, 2\rangle + \langle 1, 1\rangle = \langle -2, 3\rangle \qquad \text{in } \mathbb{R}^2,$$

and

$$(-3x + 2) + (x + 1) = -2x + 3 \qquad \text{in } P_1.$$

Aside from the "decorations" of angle brackets $\langle\cdot,\cdot\rangle$ or variables x, these two operations seem to work exactly the same way.

This is no coincidence. Throughout this course, we will see many examples of spaces that appear different on the surface, but which do in fact share many important properties. Thus it would be convenient to have a method to encode the essential data about these spaces so that we can analyze them without the distraction of the particular decoration in the space.

It turns out that there is a lovely tool for doing just that: matrices. Thus in the first unit of this course, we will spend a great deal of time discussing matrices. Later in the course, we will see how we can use matrices to

1. encode the essential data about spaces such as $\mathbb{R}^2$ and $P_1(\mathbb{R})$;
2. analyze the structure of the spaces; and
3. understand how various spaces relate to each other.

Matrices

Definition. An $m \times n$ matrix is an array of mn numbers arranged in m (horizontal) rows and n (vertical) columns; such a matrix has form

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix},$$

where each a_{ij}, referred to as the ij entry of the matrix, is the number in the ith row and jth column. Unless otherwise specified, entries of a matrix may be real or complex, and the word *number* is used to refer to both.

For example, the matrix

$$\begin{pmatrix} i & 3 & 2 & 3 \\ 1 & 1 & 5 & 1 \\ -2 & 1 & 3-i & 0 \end{pmatrix}$$

is a 3×4 matrix, since it has 3 rows and 4 columns. The $2,3$ entry is 5, which we denote by $a_{23} = 5$.

We will generally indicate a particular matrix using uppercase letters, and entries of the matrix using the lowercase version of the same letter. Writing $A = [a_{ij}]$ is a shorthand way to refer to the matrix A, whose i, j entry is a_{ij}.

Two matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ are *equal* if each of their entries is equal, i.e. $a_{ij} = b_{ij}$ for each i and j.

A Few Special Matrices

There are a few special matrices which have helpful properties, which we introduce below.

Row and Column Matrices

A *row matrix* or row vector is a matrix consisting of a single row; thus it is a $1 \times n$ matrix, and must have form

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \end{pmatrix};$$

since it has only a single row, we may drop the first indices of the entries of A and write

$$A = \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix}.$$

A *column matrix* or column vector is a matrix consisting of a single column. It must be an $m \times 1$ matrix, with form

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

It is occasionally useful to rewrite an $m \times n$ matrix in terms of its row or column vectors. For example, if

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 4 & 5 \end{pmatrix},$$

we can think of A as being made up of the row vectors

$$\mathbf{r}_1 = \begin{pmatrix} 1 & 2 & 0 \end{pmatrix} \text{ and } \mathbf{r}_2 = \begin{pmatrix} 3 & 4 & 5 \end{pmatrix}.$$

Then we may rewrite A as

$$A = \begin{pmatrix} - & \mathbf{r}_1 & - \\ - & \mathbf{r}_2 & - \end{pmatrix}.$$

Bold face indicates that $\mathbf{r}_1$ and $\mathbf{r}_2$ are indeed vectors, and not numbers.

Similarly, if we think of the columns of A as the column matrices

$$\mathbf{c}_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \ \mathbf{c}_2 = \begin{pmatrix} 2 \\ 4 \end{pmatrix}, \text{ and } \mathbf{c}_3 = \begin{pmatrix} 0 \\ 5 \end{pmatrix},$$

then we may rewrite A as

$$A = \begin{pmatrix} | & | & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3 \\ | & | & | \end{pmatrix}.$$

Square Matrices

A *square matrix* is a matrix with the same number of rows as columns; for example, the matrix

$$A = \begin{pmatrix} 3 & -2 & 0 & 0 \\ 2 & 12 & -1 & 14 \\ 0 & -10 & 4 & 5 \\ -8 & 0 & 2 & 2 \end{pmatrix}$$

is square since it has four rows and four columns.

The *main diagonal* of a square matrix consists of the entries of the matrix that have the same row and column number. So the entries a_{11} in the first row and first column, a_{22} in the second row and second column, etc., are all on the main diagonal of a square matrix. The main diagonal of the matrix A above is highlighted here in red:

$$A = \begin{pmatrix} \color{red}{3} & -2 & 0 & 0 \\ 2 & \color{red}{12} & -1 & 14 \\ 0 & -10 & \color{red}{4} & 5 \\ -8 & 0 & 2 & \color{red}{2} \end{pmatrix}.$$

The 0 Matrix

The $\mathbf{0}$ matrix is the $m \times n$ matrix each of whose entries is 0; below is an example of a 2×3 $\mathbf{0}$ matrix:

$$\mathbf{0} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Be careful to distinguish between a zero *matrix* and the *number* 0; the former is a matrix all of whose entries are 0s, while the latter is simply a scalar. I will type $\mathbf{0}$ to refer to the *matrix* whose entries are 0s.

The Identity Matrix

The $n \times n$ *identity matrix* I_n is the (square) $n \times n$ matrix whose iith entry is 1, and whose ijth entry is 0 if $i \neq j$:

$$I = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

In other words, I is the square matrix which has 1s along its main diagonal, and 0s everywhere else. A few specific examples are below:

$$I_1 = \begin{pmatrix} 1 \end{pmatrix}, \quad I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I_5 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

The Transpose and Conjugate Transpose of a Matrix

There are many ways to use a given matrix to build a new, related matrix; one such way is by taking the transpose, as described below:

Definition. Given an $m \times n$ matrix $A = [a_{ij}]$, its *transpose*, denoted $A^\top$, is the $n \times m$ matrix

$$A^\top = [a_{ji}],$$

whose (i, j)th entry is just the (j, i)th entry of A.

In a sense, $A^\top$ just switches the columns and rows of A; row 1 of A is column 1 of $A^\top$, row 2 of A is column 2 of $A^\top$, etc.

For example, given the 2×3 matrix

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 4 & 5 \end{pmatrix}.$$

$A^\top$ is the 3×2 matrix given by

$$A^\top = \begin{pmatrix} 1 & 3 \\ 2 & 4 \\ 0 & 5 \end{pmatrix}.$$

Notice that, say, the $(2,3)$ entry of A is 5, which is the $(3,2)$ entry of $A^\top$.

If A has complex entries, it is interesting to consider the *conjugate transpose* of A:

Definition. Given an $m \times n$ matrix $A = [a_{ij}]$, its *conjugate transpose*, denoted A^*, is the $n \times m$ matrix

$$A^\top = [\overline{a_{ji}}],$$

whose (i,j)th entry is the complex conjugate of the (j,i)th entry of A.

As an example, let

$$A = \begin{pmatrix} i & 1-i & 3 \\ 0 & 1 & -i \\ 2i & 1 & 2+2i \end{pmatrix};$$

then the conjugate transpose of A is given by

$$A^* = \begin{pmatrix} -i & 0 & -2i \\ 1+i & 1 & 1 \\ 3 & i & 2-2i \end{pmatrix},$$

whereas the *transpose* of A is

$$A^\top = \begin{pmatrix} i & 0 & 2i \\ 1-i & 1 & 1 \\ 3 & i & 2+-2i \end{pmatrix}.$$

Key Point. If A is a matrix all of whose entries are real numbers, then $\overline{a_{ij}} = a_{ij}$ for each entry; thus

$$A^* = A^\top.$$

Trace of a Square Matrix

In this course, we will spend a great deal of time analyzing the characteristics of specific matrices. The first characteristic that we will study is called the *trace*:

Definition. The *trace* of a square matrix A, denoted $\mathrm{tr}\,(A)$, is the sum of its main diagonal entries.

Example. Calculate the trace of the matrix

$$A = \begin{pmatrix} 3 & -2 & 0 & 0 \\ 2 & 12 & -1 & 14 \\ 0 & -10 & 4 & 5 \\ -8 & 0 & 2 & 2 \end{pmatrix}.$$

Earlier, we pointed out that the main diagonal of A consists of the entries highlighted in red below:

$$A = \begin{pmatrix} \color{red}{3} & -2 & 0 & 0 \\ 2 & \color{red}{12} & -1 & 14 \\ 0 & -10 & \color{red}{4} & 5 \\ -8 & 0 & 2 & \color{red}{2} \end{pmatrix}.$$

So

$$\mathrm{tr}\,(A) = 3 + 12 + 4 + 2 = 21.$$

While the trace of a matrix might seem like a useless quantity to calculate, it turns out that this quantity actually gives us a some interesting data about the matrix, which we must wait to discuss until we have learned more about the properties of square matrices.