1 Vectors

The classical definition of a vector is quantity that has both a direction and a magnitude. In this class, that definition still holds, but often we will be thinking of a vector as representing the location of a point in a Euclidean coordinate system. So the point $\mathbf{v} = (4, 3)$ could also be represented as the vector

$$\mathbf{v} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}.$$

Here we're using bold font to denote vectors. If we drew an arrow from the origin (0,0) to (4,3), that would represent our vector $\mathbf{v}$. Its direction is the direction of the arrow, and its magnitude $|\mathbf{v}|$ is the length of the arrow, in this case 5. In general, if

$$\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}, \quad \text{then} \quad |\mathbf{v}| = \sqrt{x^2 + y^2}.$$

1.1 Vector addition

To add two (or more) vectors together, we can add their x components and y components separately, which has a very visual representation, shown in Figure 1. In this example, adding the components separately gives us

$$\mathbf{v}_1 + \mathbf{v}_2 = \begin{bmatrix} 8 \\ 3 \end{bmatrix} + \begin{bmatrix} 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 8+1 \\ 3+5 \end{bmatrix} = \begin{bmatrix} 9 \\ 8 \end{bmatrix},$$

which is exactly what we obtain from placing the vectors head to tail in the figure.

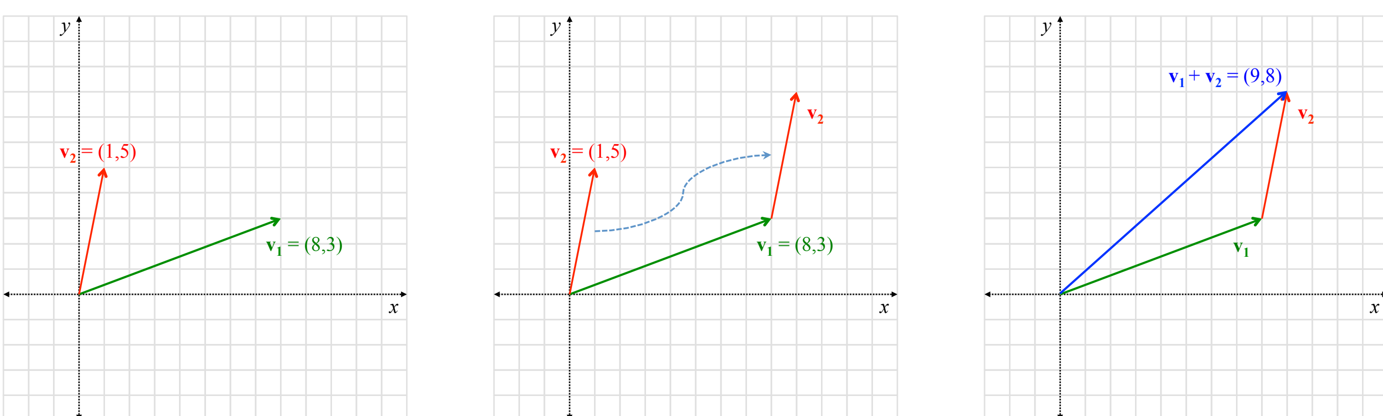

Figure 1: Addition of two vectors $\mathbf{v}_1 = (8, 3)$ in green and $\mathbf{v}_2 = (1, 5)$ in red. If we place the vectors head to tail, the resulting vector is their addition, which is the same as adding the x and y components separately. Either way we get that $\mathbf{v}_1 + \mathbf{v}_2 = (9, 8)$, represented by the blue vector.

We will also be using non-2D vectors, but the same notation and algebra still applies. Vector notation displays the components as a column, which is important for matrix-vector multiplication later on. However, to save space we can also use transpose (switch rows and columns) notation:

$$\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix}^{\mathrm{T}}$$

2 Matrices

In this class, a matrix is an array of numbers, arranged in rows and columns. For us matrices will often represent a dataset with examples on the rows and features on the columns. It is important to note that a vector is also a matrix. Usually we denote matrices by a bold capital letter, for example

$$\mathbf{A} = \begin{bmatrix} 4 & 1 & 10 \\ -3 & 2 & 0 \end{bmatrix}$$

The dimension of $\mathbf{A}$ is ordered "rows $\times$ columns", so in this case it is 2×3.

2.1 Matrix addition

Matrices can be added together component-wise just like vectors, provided they are of the *same* dimension. In fact, we can think of each column of a matrix as being a vector (i.e. point). So the following addition of matrices $\mathbf{A}$ and $\mathbf{B}$ can be thought of as adding $[a \quad c]^{\mathrm{T}} + [e \quad g]^{\mathrm{T}}$ (first column of each matrix) and $[b \quad d]^{\mathrm{T}} + [f \quad h]^{\mathrm{T}}$ (second column of each matrix):

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix}.$$

2.2 Matrix multiplication

We can also multiply matrices, but this is not as straightforward as addition. Matrix multiplication is defined in a particular way, and can be thought of as an extension of the dot product. Using the matrices from our addition example, we can build up $\mathbf{AB}$ one entry at a time:

$$\begin{bmatrix} \text{a} & \text{b} \\ c & d \end{bmatrix} \begin{bmatrix} \text{e} & f \\ \text{g} & h \end{bmatrix} = \begin{bmatrix} \text{ae + bg} & \\ & \end{bmatrix} \qquad \begin{bmatrix} \text{a} & \text{b} \\ c & d \end{bmatrix} \begin{bmatrix} e & \text{f} \\ g & \text{h} \end{bmatrix} = \begin{bmatrix} ae+bg & \text{af + bh} \\ & \end{bmatrix}$$

$$\begin{bmatrix} a & b \\ \text{c} & \text{d} \end{bmatrix} \begin{bmatrix} \text{e} & f \\ \text{g} & h \end{bmatrix} = \begin{bmatrix} ae+bg & af+bh \\ \text{ce + dg} & \end{bmatrix} \qquad \begin{bmatrix} a & b \\ \text{c} & \text{d} \end{bmatrix} \begin{bmatrix} e & \text{f} \\ g & \text{h} \end{bmatrix} = \begin{bmatrix} ae+bg & af+bh \\ ce+dg & \text{cf + dh} \end{bmatrix}$$

Therefore overall:

$$\mathbf{AB} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{bmatrix}.$$

Here we multiplied two 2×2 matrices, but, unlike addition, we can multiply matrices that do not have exactly the same dimension. However, the number of columns of the first matrix *must* match the number of rows of the second matrix. So we could multiply a 4×2 matrix and a 2×3 matrix to obtain a 4×3 matrix. One way to remember this is that the "inner" dimensions must match (2 in this example), and the "outer" dimensions will be the dimensions of the resulting matrix. Note that swapping the multiplication order in this case would not be valid (we cannot multiply a 2×3 matrix and a 4×2 matrix in that order).