Linear Algebra

Aiphabet - aiphabet.org

1 Vectors

A **vector space** is a set V along with addition and scalar multiplication operations on V. Elements of a vector space are called vectors. A **vector** is a quantity that has a direction and magnitude, which we will see shortly.
Now, we must first define the elementary operations that we can perform on vectors. Addition and scalar multiplication is performed component-wise:

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix}$$

$$\lambda \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \lambda y_1 \end{bmatrix}$$

Note:
A common notation for indicating that a quantity is vector-valued is by including a small arrow above the variable, such as $\vec{u}$, to indicate that the quantity has a direction.

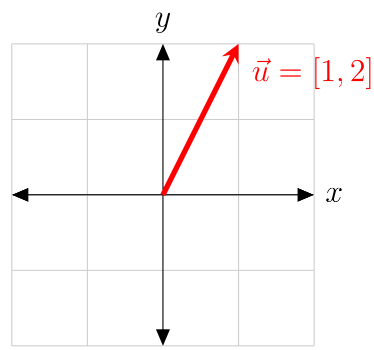
- **Vector addition**:
  1. Commutative: $\vec{v} + \vec{u} = \vec{u} + \vec{v}$
  2. Associative: $(\vec{v} + \vec{u}) + \vec{w} = \vec{v} + (\vec{u} + \vec{w})$
  3. Identity: $\vec{v} + \vec{0} = \vec{v}$
- **Scalar multiplication**:
  1. Associative: $(ab)\vec{v} = a(b\vec{v})$
  2. Identity: $\vec{1}\vec{v} = \vec{v}$
  3. A *positive-valued* scalar stretches the vector in its current direction, with the magnitude of the scalar dictating how much to “pull"
  4. A *negative-valued* scalar stretches the vector by the same magnitude, just in the *opposite* direction
- **Magnitude**: $|\vec{v}| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$ for $\vec{v} \in \mathbb{R}^n$

1.1 Visualizing Vectors

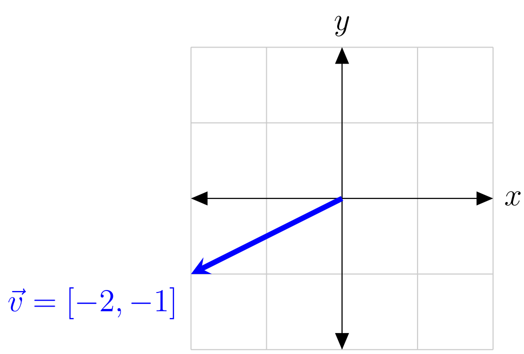
A vector contains information about the x and y components, namely,

$$\begin{bmatrix} v_x \\ v_y \end{bmatrix}$$

So say we wanted to visualize the vector $\vec{u} = \left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$. We therefore have to move one unit in the positive direction in the x axis (right) and two units in the positive direction in the y axis (up), like so:



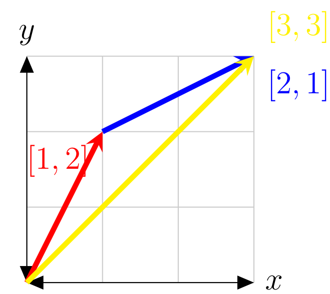
Now say you wanted to visualize another vector, $\vec{v} = \left[\begin{smallmatrix} -2 \\ -1 \end{smallmatrix}\right]$. Therefore, we need to move two units in the negative direction in the x axis (left) and one unit in the negative direction in the y axis (down):



1.2 Vector Addition

Now we will explore how to visualize the addition of vectors. In order to add together the following two vectors in the coordinate space, we will use the traditional “tip-to-tail" method. This involves orienting the first vector you wish to add at the origin, and then placing the “tail" of the second vector at the “tip" of the first, and orienting the second vector from there. Then we create a triangle from the tail of the first vector to the tip of the second, which expresses the total sum of the vectors and is aptly named the **resultant**.
The figure below illustrates the addition of the following two vectors, yielding the resultant vector $\left[\begin{smallmatrix} 3 \\ 3 \end{smallmatrix}\right]$.

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$



Note:
Vector subtracting is performed in a similar fashion except that we orient the vector we wish to add in the opposite direction.

2 Matrices

A **matrix** is an array $\mathbf{A}$ of m rows and n columns. The size of a matrix $\mathbf{A}$ can be denoted by the notation $m \times n$, which indicates a matrix with m *rows* and n *columns*, and we can denote a specific entry of $\mathbf{A}$ through the indexes i and j, where i is the row index and j is the column index. Hence, we can formalize the elements of a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ as $a_{ij} \in \mathbb{R}$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n$.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}$$

2.1 Matrix Operations

We can now define the elementary operation that we can perform on vectors. Suppose we wanted to add together two matrices, $\mathbf{A} + \mathbf{B}$. This can be done in an element-wise fashion as follows:

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} a_{11} + b_{11} & \ldots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & \ldots & a_{2n} + b_{2n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \ldots & a_{mn} + b_{mn} \end{bmatrix}$$

Note:
Matrices must be of the *same* dimension for the addition operation to be compatible, that is, both $\mathbf{A}$ and $\mathbf{B}$ must have the same number of *rows* and *columns*.

To multiply $\mathbf{A}$ by some real-valued scalar λ, we similarly multiply the scalar element-wise:

$$\lambda\mathbf{A} = \begin{bmatrix} \lambda a_{11} & \dots & \lambda a_{1n} \\ \lambda a_{21} & \dots & \lambda a_{2n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \dots & \lambda a_{mn} \end{bmatrix}$$

Scalar multiplication:

1. Associative: $(\lambda\gamma)\mathbf{A} = \lambda(\gamma\mathbf{A})$ for $\lambda, \gamma \in \mathbb{R}$ and $\mathbf{A} \in \mathbb{R}^{m\times n}$
2. Distributive: $(\lambda + \gamma)\mathbf{A} = \lambda A + \gamma A$ for $\lambda, \gamma \in \mathbb{R}$ and $\mathbf{A} \in \mathbb{R}^{m\times n}$

In order to multiply together two matrices, we must define a new operation called the **dot product**. Suppose we wish to multiply together matrices $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times k}$. Their product, $\mathbf{C} \in \mathbb{R}^{m\times k}$, is formally defined as follows:

$$c_{ij} = \sum_{l=1}^{n} a_{il} b_{lj}$$

for $i = 1, \dots, m$ and $j = 1, \dots, k$. Informally, this notation simply takes the ith row of matrix $\mathbf{A}$ and multiplies it with the jth column of matrix $\mathbf{B}$, and add all these entries together to fill the ijth entry of the resulting matrix $\mathbf{C}$.

Note:

Matrices must have the *same inner* dimension for the dot product operation to be compatible, that is, the number of columns of $\mathbf{A}$ must be equal to the number of rows of $\mathbf{B}$. We can see this easily in our example above where matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ has n columns and matrix $\mathbf{B} \in \mathbb{R}^{n\times k}$ has n rows. The resulting matrix $\mathbf{C} \in \mathbb{R}^{m\times k}$ has dimensions equal to the outer left and outer right dimensions of $\mathbf{A}$ and $\mathbf{B}$, respectively.

Matrices also have a multiplicative identity defined as a matrix with ones along the diagonal and zeros everywhere else. Below is the square $(n \times n)$ identity matrix:

$$\mathbb{I}_n = \begin{bmatrix} 1 & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 1 \end{bmatrix}$$

Now we can proceed to define some more properties of matrix multiplication:
Matrix multiplication:

1. Associative: $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$ for $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times p}$, $\mathbf{C} \in \mathbb{R}^{p\times q}$
2. Distributive: $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$ for $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times p}$, $\mathbf{C} \in \mathbb{R}^{p\times q}$

3. Identity: $\mathbb{I}_m\mathbf{A} = \mathbf{A}\mathbb{I}_n = \mathbf{A}$ for $\mathbf{A} \in \mathbb{R}^{m\times n}$

Note:

Matrices are generally *not commutative*, that is, $\mathbf{AB} \neq \mathbf{BA}$!

2.2 Inverses and Determinants

Similar to how you can "undo" a product by multiplying by an integer's reciprocal to get back 1, such as $4 \times \frac{1}{4} = 1$, we can do a very similar thing with matrices by finding the **inverse** of a matrix. Suppose we have the square matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$. Then we call the square matrix $\mathbf{B} \in \mathbb{R}^{n\times n}$ the inverse of $\mathbf{A}$ if the following holds:

$$\mathbf{AB} = \mathbb{I}_n = \mathbf{BA}$$

We can denote the inverse of $\mathbf{A}$ as $\mathbf{A}^{-1}$. Unfortunately, not every matrix has an inverse. Firstly, you may have noted that we required our matrix $\mathbf{A}$ to be square, otherwise an inverse would not exist! Secondly, a property of the matrix called the **determinant** must be non-zero in order for the inverse to exist. If the determinant of a matrix $\mathbf{A}$ is zero, we say that $\mathbf{A}$ is *noninvertible* or *singular*, and therefore has no inverse. Otherwise, if $\mathbf{A}$ has an inverse, we call it *invertible* or *nonsingular*.

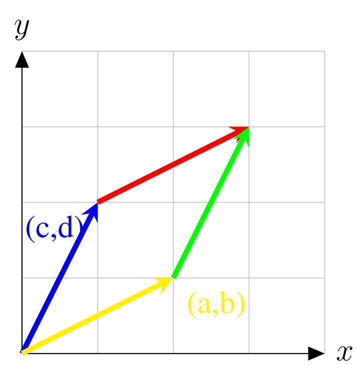
At this point, you may be wondering what is this magical property called the *determinant*. Despite the fancy name, computing the determinant of a matrix is relatively straightforward for 2×2 matrices. Suppose we have the following matrix:

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

The determinant of $\mathbf{A}$, which we denote by $\det(A)$ is computed as follows:

$$\det(A) = ad - cb$$

If we think about the determinant geometrically, it is computing the area of the parallelogram spanned by the column vectors of $\mathbf{A}$:



Another important property of a matrix is its **transpose**. The transpose of a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ is found by simply flipping the position of its rows by the position of its columns, which can formally be expressed as $b_{ij} = a_{ji}$, after transposing matrix $\mathbf{A}$ into matrix $\mathbf{B}$, $\mathbf{A}^\top = \mathbf{B}$. So if we were to take the transpose of matrix $\mathbf{A}$ below:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

We simply “reflect" the entries across the diagonal to flip the positions between row and column elements:

$$\mathbf{A}^\top = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$$

Below we present some useful properties of inverses and transposes:
Inverse Properties:

- $\mathbf{A}\mathbf{A}^{-1} = \mathbb{I} = \mathbf{A}^{-1}\mathbf{A}$
- $(\mathbf{A}\mathbf{B})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$

Transpose Properties:

- $(\mathbf{A}^\top)^\top = \mathbf{A}$
- $(\mathbf{A}\mathbf{B})^\top = \mathbf{B}^\top\mathbf{A}^\top$
- $(\mathbf{A} + \mathbf{B})^\top = \mathbf{A}^\top + \mathbf{B}^\top$

Note:

The inverse cannot be “distributed" across addition, so $(\mathbf{A} + \mathbf{B})^{-1} \neq \mathbf{A}^{-1} + \mathbf{B}^{-1}$.

3 Span and Linear Independence

The vector space V of elements $x_1, \ldots, x_n \in V$ are called **linearly dependent** if there exists scalars $\lambda \in \mathbb{R}$ that are **not all zero** such that:

$$\sum_{i=1}^{n} \lambda_i x_i = \lambda_1 x_1 + \ldots \lambda_n x_n = \vec{0}$$

If only the trivial linear combination of $\lambda_1 = \cdots = \lambda_n = 0$ holds such that $\sum_{i=1}^{n} \lambda_i x_i = \vec{0}$, then the vectors $x_1, \ldots, x_n$ are **linearly independent**.

We call the **span** of a set of vectors $X = \{x_1, x_2, \dots, x_n\}$ all linear combinations of these vectors, of the form $\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 + \cdots + \lambda_n x_n$

Example: Consider the following set of vectors:

$$X = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}$$

Is this is a linearly independent set?

Observe that we can choose $\lambda_1 = 2$, $\lambda_2 = -1$, and $\lambda_3 = -1$ such that:

$$2 \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} + (-1) \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} + (-1) \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Therefore X is *not* a linearly independent set and is, therefore, *linearly dependent*.

4 Linear Systems

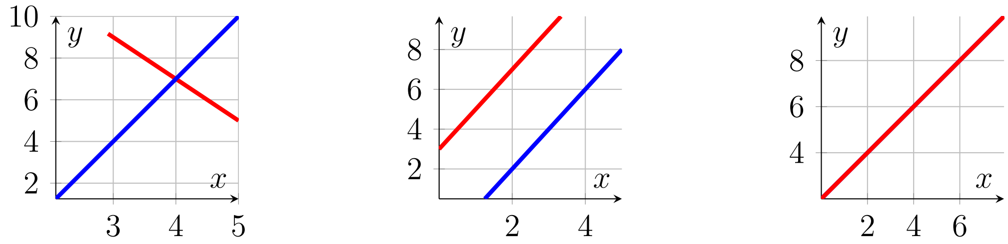
Suppose we have a set of n unknowns with a set of constraints. We can solve for this set of unknowns by defining a **system of linear equations**, for which we can express the general form as follows:

$$\begin{gathered} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = b_m \end{gathered}$$

Where $x_1, \dots, x_n \in \mathbb{R}^n$ is the set of n unknowns and $a_{ij} \in \mathbb{R}$ and $b_i \in \mathbb{R}$ are constants. Every n-tuple $(x_1, \dots, x_n) \in \mathbb{R}^n$ that satisfies the system above is a **solution** of the linear system.

There are three types of linear systems, depending on the number of unknowns relative to the number of equations in the system.

- **Infinite solutions** (under-determined): More unknowns that knowns
- **No solution** (over-determined): More knowns than unknowns
- **Unique solution**: same number of knowns and unknowns



(a) Unique Solution **(b)** No Solution **(c)** Infinite Solutions

5 Solving Systems of Linear Equations

In order to solve the system of linear equations, first we collect the coefficients a_{ij} of each term and the corresponding solutions to each linear equation and express it as an **augmented matrix**, as follows:

$$
\begin{aligned}
x_1 + x_2 + x_3 &= 3 \\
x_1 - x_2 + 2x_3 &= 2 \\
x_2 + x_3 &= 2
\end{aligned}
$$

We can translate the system into matrix format by keeping the coefficients of each term:

$$
\left[\begin{array}{ccc|c}
1 & 1 & 1 & 3 \\
1 & -1 & 2 & 2 \\
0 & 1 & 1 & 2
\end{array}\right]
$$

In order to solve the system of equations, we use the following rules to reduce the augmented matrix into **row-echelon form**.

1. Interchange any two rows
2. Multiply row by a non-zero constant
3. Add a multiple of one row to another

The final goal is to convert the matrix into an upper-triangular form:

$$
\left[\begin{array}{cccc|c}
1 & * & * & * & * \\
0 & 1 & * & * & * \\
0 & 0 & 1 & * & * \\
0 & 0 & 0 & 1 & *
\end{array}\right]
$$

Example: Solve the following linear system:

$$
\begin{aligned}
&x_1 - 2x_2 + x_3 = 0 \\
&2x_2 - 8x_3 = 8 \\
&5x_1 - 5x_3 = 10
\end{aligned}
$$

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{array}\right]$$

Now, to express that we wish to subtract fives times the first row $(\mathbf{r_1})$ to the third row $(\mathbf{r_3})$ and replace that with row 3, using rules 2 and 3, we can denote this as follows:

$\mathbf{r_3 \leftarrow r_3 + (-5)r_1}$:

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & 10 & -10 & 10 \end{array}\right]$$

$\mathbf{r_3 \leftarrow r_3 + (-5)r_2}$:

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & 0 & 30 & -30 \end{array}\right]$$

$\mathbf{r_2 \leftarrow \frac{1}{2}r_2}$:

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 30 & -30 \end{array}\right]$$

$\mathbf{r_3 \leftarrow \frac{1}{30}r_3}$:

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & -1 \end{array}\right]$$

$\mathbf{r_2 \leftarrow r_2 + 4r_3}$:

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{array}\right]$$

$\mathbf{r_1 \leftarrow r_1 + (-1)r_3}$:

$$\left[\begin{array}{ccc|c} 1 & -2 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{array}\right]$$

$\mathbf{r_1 \leftarrow r_1 + 2r_2}$:

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{array}\right]$$

So in the end, we get the following solution to our set of unknowns: $x_1 = 1, x_2 = 0, x_3 = -1$.

There are cases in which the final row-echelon form has an upper-triangular form does not yield a single solution. Such a scenario would fall under either the under-determined or the over-determined case:

Infinite Solutions (under-determined):

$$\left[\begin{array}{cccc|c} 1 & * & * & * & * \\ 0 & 1 & * & * & * \\ 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

No Solutions (over-determined):

$$\left[\begin{array}{cccc|c} 1 & * & * & * & * \\ 0 & 1 & * & * & * \\ 0 & 0 & 1 & 0 & * \\ 0 & 0 & 0 & 0 & * \end{array}\right]$$