Matrices

Suppose you have the following system of linear equations:

\[
\begin{align*}
2x + 8y + 4z &= 2 \\
2x + 5y + z &= 5 \\
4x + 10y - z &= 1
\end{align*}
\]

To solve the system for \((x, y, z)\), all you really need are the coefficients of \(x, y, z\) and the right-hand sides of the equations.

All the necessary information can be stored as:

\[
\begin{bmatrix}
2 & 8 & 4 & 2 \\
2 & 5 & 1 & 5 \\
4 & 10 & -1 & 1
\end{bmatrix}
\]

This array of numbers is called a “matrix.”

A matrix is simply a mathematical notation used to store all of the information pertaining to a system of linear equations.
The matrix

\[
\begin{bmatrix}
2 & 8 & 4 & 2 \\
2 & 5 & 1 & 5 \\
4 & 10 & -1 & 1
\end{bmatrix}
\]

has 3 rows and 4 columns and is thus referred to as a $3 \times 4$ (“3 by 4”) matrix (rows are across and columns are up and down).

In general, matrix sizes are referred to as $m \times n$, where

\[
m = \text{# of rows} \\
n = \text{# of columns}
\]

Matrix entries are often labelled using double subscripts

\[
\begin{bmatrix}
a_{11} & a_{12} & a_{13} & a_{14} \\
a_{21} & a_{22} & a_{23} & a_{24} \\
a_{31} & a_{32} & a_{33} & a_{34}
\end{bmatrix}
\]

where entry $a_{ij}$ is located in the $i$'th row and $j$'th column.

If $m = n$, the matrix is square.

If $A_{nn}$ is a square matrix, then the entries, $a_{11}, a_{22}, \ldots, a_{nn}$ form the main diagonal of $A$. 
Types of square matrices include:

- diagonal, where all entries off the main diagonal are 0, or $a_{ij} = 0$ whenever $i \neq j$.
- upper triangular, where all entries below the main diagonal are 0, or $a_{ij} = 0$ whenever $i > j$.
- lower triangular, where all entries above the main diagonal are 0, or $a_{ij} = 0$ whenever $i < j$.

A matrix with only 1 column is called a column vector, or simply a vector.

for example,

\[
\begin{bmatrix}
1 \\
2 \\
9 \\
1
\end{bmatrix}
\]

is a vector in $\mathbb{R}^4$.

A matrix with only 1 row is called a row vector.

for example,

\[
\begin{bmatrix}
1 & 5 & 5 & 3 & 7
\end{bmatrix}
\]

is a row vector in $\mathbb{R}^5$.

Note that the $n$ columns of an $m \times n$ matrix are vectors in $\mathbb{R}^m$, and the $m$ rows in an $m \times n$ matrix are row vectors in $\mathbb{R}^n$. 
Solving Systems of Linear Equations

Again consider the system:

\[
\begin{align*}
2x + 8y + 4z &= 2 \\
2x + 5y + z &= 5 \\
4x + 10y - z &= 1
\end{align*}
\]

The **coefficient matrix** of this system is:

\[
\begin{bmatrix}
2 & 8 & 4 \\
2 & 5 & 1 \\
4 & 10 & -1
\end{bmatrix}
\]

The **augmented matrix** of this system is:

\[
\begin{bmatrix}
2 & 8 & 4 & : & 2 \\
2 & 5 & 1 & : & 5 \\
4 & 10 & -1 & : & 1
\end{bmatrix}
\]

To solve the system, work with the augmented matrix.

The idea of an “equation” in the linear system is replaced by the idea of a “row” in the augmented matrix.
To solve a system, the general goal is to reduce its augmented matrix to the form
\[
\begin{bmatrix}
  1 & 0 & 0 & : & a \\
  0 & 1 & 0 & : & b \\
  0 & 0 & 1 & : & c \\
\end{bmatrix}
\]

This is the augmented matrix that corresponds to the linear system
\[
\begin{align*}
1x + 0y + 0z &= a \\
0x + 1y + 0z &= b \\
0x + 0y + 1z &= c
\end{align*}
\]

which is simply,
\[
\begin{align*}
x &= a \\
y &= b \\
z &= c
\end{align*}
\]

We will use Gauss-Jordan Elimination (or simply “elimination”) to reduce a general augmented matrix to its reduced form, which is called Reduced Row-Echelon Form (rref). The reduced form shown above is just one type of rref, not all matrices will reduce to that particular rref. (rref will be discussed in more detail later)

Gauss-Jordan elimination is motivated by the fact that the following two matrix operations do not change the system:

1. multiplying/dividing a row by a scalar
2. replacing one row with the sum (that row + a multiple of another row)
**Gauss-Jordan Elimination**

1. Write the system in augmented matrix form

2. Start at the top entry of the first nonzero column of the augmented matrix (cursor entry)

   Repeat the following steps 3-6 until you run out of rows or columns

3. If the cursor entry is 0, swap the cursor row with some row below to make the cursor entry nonzero. If all the entries below the cursor are also 0, move the cursor one position to the right.

4. Divide the cursor row by the cursor entry (to make the cursor entry 1).

5. Eliminate all other entries in the cursor column by adding the appropriate multiples of the cursor row to the other rows.

6. Move the cursor down one position and to the right one position.

After these steps, the augmented matrix will be in Reduced Row-Echelon Form (rref).

**Reduced Row-Echelon Form of a Matrix**

A matrix is in **reduced row-echelon form** if it satisfies all of the following:

a. If a row has nonzero entries, then the first nonzero entry is a 1, called the **leading 1** in that row.

b. If a column contains a leading 1, then all other entries in that column are 0.

c. If a row contains a leading 1, then the row above contains a leading 1 further to the left.
e.g., consider our system of linear equations and remember our goal of getting to

\[
\begin{bmatrix}
1 & 0 & 0 & : & a \\
0 & 1 & 0 & : & b \\
0 & 0 & 1 & : & c \\
\end{bmatrix}
\]

the elimination steps would proceed as follows:

\[
\begin{bmatrix}
2 & 8 & 4 & : & 2 \\
2 & 5 & 1 & : & 5 \\
4 & 10 & -1 & : & 1 \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 4 & 2 & : & 1 \\
2 & 5 & 1 & : & 5 \\
4 & 10 & -1 & : & 1 \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 4 & 2 & : & 1 \\
0 & -3 & -3 & : & 3 \\
0 & -6 & -9 & : & -3 \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 4 & 2 & : & 1 \\
0 & 1 & 1 & : & -1 \\
0 & -6 & -9 & : & -3 \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 0 & -2 & : & 5 \\
0 & 1 & 1 & : & -1 \\
0 & 0 & -3 & : & -9 \\
\end{bmatrix}
\]
\[
\begin{array}{ccc}
\downarrow \\
\begin{bmatrix}
1 & 0 & -2 : 5 \\
0 & 1 & 1 : -1 \\
0 & 0 & 1 : 3 \\
\end{bmatrix}
\downarrow \\
\begin{bmatrix}
1 & 0 & 0 : 11 \\
0 & 1 & 0 : -4 \\
0 & 0 & 1 : 3 \\
\end{bmatrix}
\end{array}
\]

this is the form we wanted, which directly gives the solution vector

\[
\begin{bmatrix}
x \\
y \\
z \\
\end{bmatrix} = \begin{bmatrix}
11 \\
-4 \\
3 \\
\end{bmatrix}
\]

can check this by substituting the \(x, y, z\) solution values into the original three equations.