2. Working with Matrices

Last time, we discussed that a system is called homogeneous if all of the constant terms are 0. Notice that $x_1 = x_2 = \cdots = x_n = 0$ is a solution always to a homogeneous system of linear equations.

Definition 1 (Trivial Solution)

We call the solution $x_1 = x_2 = \cdots = x_n = 0$ to the homogeneous equation the "trivial solution". Any other solutions are called nontrivial.

Definition 2 (Square Matrix, Order and Main Diagonal)

A square matrix has the same number of rows and columns. That number is called the order of that matrix. The entries whose row and column indices are equal make up the main diagonal.

We shall now define several operations on matrices.
First, we can add matrices of the same size by adding elementwise.

Example 1

Let $A = \begin{bmatrix} 2 & 1 \\ 5 & -4 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$. Then $A + B = \begin{bmatrix} 4 & 1 \\ 5 & -3 \end{bmatrix}$.

We can subtract matrices using the same procedure. We can also multiply a matrix by a scalar (number):

Example 2

With A as in Example 1, we get $4A = \begin{bmatrix} 8 & 4 \\ 20 & -16 \end{bmatrix}$.

Definition 3 (Transpose)

The transpose of a matrix A is defined by $(A^\top)_{ij} = (A)_{ji}$. That is, to get the transpose of a matrix, we swap rows and columns.

Note that the transpose is defined for a matrix of any size, as can be seen in the following example:

Example 3

Let $C = \begin{bmatrix} 3 & 2 & 1 \\ 1 & 4 & 2 \end{bmatrix}$. Then $C^\top = \begin{bmatrix} 3 & 1 \\ 2 & 4 \\ 1 & 2 \end{bmatrix}$.

Definition 4

The trace of a square matrix is the sum of the entries on the main diagonal.

Note that by this definition, the trace is only defined for a square matrix. Also note that the trace is a function which takes as input a square matrix and outputs a number (not a matrix!).

Example 4

With A as in Examples 1 and 2, $\mathrm{tr}(A) = 2 - 4 = -2$.

We can also multiply matrices! This is a bit different from what you might expect – we cannot just multiply elementwise. Instead, we have a procedure to follow. We multiply the rows of the first matrix by the columns of the second matrix. We multiply elementwise within those rows/columns, and then we add the results. This procedure allows us to form a new product matrix. However, this adds an important note: it is only possible to multiply two matrices where the number of columns of the first matrix is equal to the number of rows of the second matrix. So if you are multiplying a $k \times \ell$ matrix by a $m \times n$ matrix, you need $\ell = m$. Further, the resulting product matrix will be $k \times n$. Let's do an example:

Example 5

$$\begin{bmatrix} -3 & 6 \\ 1 & -2 \\ 7 & 4 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 9 & 0 & -9 \\ -3 & 0 & 3 \\ 15 & 18 & 21 \\ 7 & 5 & 3 \end{bmatrix}.$$

Definition 5 (Identity Matrix)

The identity matrix I is defined as $(I)_{ii} = 1$ and $(I)_{ij} = 0$ if $i \neq j$. That is, the identity matrix has ones along the main diagonal and zeros everywhere else. It must be square.

Definition 6 (Inverse)

A matrix inverse is a matrix A^{-1} such that $AA^{-1} = A^{-1}A = I$. Note that not all matrices have inverses! Matrices that have an inverse are called invertible; matrices which don't have an inverse are called singular.

We have a formula for the inverse of a 2×2 matrix. It is as follows:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Finally, here are some extra facts about matrices:

1. $(A^{-1})^{-1} = A$
2. $(A^\top)^\top = A$
3. $(A \pm B)^\top = A^\top \pm B^\top$
4. $(A^\top)^{-1} = (A^{-1})^\top$
5. $(kA)^{-1} = \frac{1}{k}A^{-1}$
6. $(kA)^\top = kA^\top$
7. $(AB)^\top = B^\top A^\top$