Math 123, Week 2:
Matrix Operations, Inverses

Section 1: Matrices

We have introduced ourselves to the grid-like coefficient matrix when performing Gaussian elimination. We now formally define general matrices and several useful matrix operations, such as additional, multiplication, and a division-like operation known as inversion.

**Definition 1**

The grid structure $A = [a_{ij}]$, $i = 1, \ldots, m$, $j = 1, \ldots, n$, is called a matrix of **dimension** $m \times n$ (or simply an $m$-by-$n$ matrix). Explicitly, we have

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

We will adopt the convention of using capital letters to denote matrices (e.g. $A, B, C$, etc.) and will assume that the matrix entries $a_{ij}$ are either real-valued or integer-valued coefficients. It should be noted that the operations we will define extend easily to matrices with complex-valued entries and, in fact, there are a surprising number of applications of such matrices (e.g. quantum physics!).

The easiest (and most frustrating) way to mess up a matrix problem is to accidentally interchange the columns and the rows. By convention, the first index of the dimension corresponds to the **number of rows** and the second index corresponds to the **number of columns**. For example, a 5-by-3 matrix is a matrix with five rows and three columns. If we confuse it for a matrix with three rows and five columns, we will probably not be able to perform the required operations. A matrix entry $a_{ij}$ will always corre-
spond to the element in the \( i^{th} \) row and \( j^{th} \) column. Fortunately, this “row
then column” convention is adopted universally throughout the sciences and
mathematics.

We will be interested in defining operations on matrices. In many ways,
the operations we will define for matrices will be the same as those defined
for numbers. That is to say, we are interested in such things are addition
(e.g. \( A + B \)), subtraction (e.g. \( A - B \)), multiplication (e.g. \( A \cdot B \)), exponen-
tiation (e.g. \( A^2 \)), etc., of matrices. This is not a coincidence, but there are
subtleties of which we will have to be aware.

The most basic operation we can perform is to take the transpose, which
is defined in the following way.

**Definition 2**

Let \( A \) be an \( m \times n \) matrix with entries \( A = [a_{ij}] \). Then the transpose
of \( A \) is the \( n \times m \) matrix denoted \( A^T \) and has entries
\( A^T = [a_{ji}] \).

When determining the transpose, we simply switch the indices of the entries.
For instance, an entry in the 5\(^{th} \) row and 1\(^{st} \) column of \( A \) will be in the 1\(^{st} \) row and 5\(^{th} \) column of \( A^T \).

**Example 1**

Determine the transpose of the matrix:

\[
A = \begin{bmatrix}
2 & -1 \\
0 & 5 \\
1 & -1
\end{bmatrix}
\]

**Solution:** By the definition, we have

\[
A^T = \begin{bmatrix}
2 & 0 & 1 \\
-1 & 5 & -1
\end{bmatrix}.
\]

It should be obvious from both the definition and the example that, if \( A \) is
an \( m \)-by-\( n \) matrix, then \( A^T \) will be a \( n \)-by-\( m \) matrix. It is *very important*
to remember this fact since dimension matters when performing the matrix
operations we are about to define.
Section 2: Matrix Addition

The first and most basic algebraic operation we can define for matrices is matrix addition.

**Definition 3**

Let $A$ and $B$ be two $m$-by-$n$ matrices with entries $A = [a_{ij}]$, $B = [b_{ij}]$. Then the matrix $A + B$ is defined to be the $m \times n$ matrix with entries $A + B = [a_{ij} + b_{ij}]$.

**Example 2**

Given the following matrices, find $A + B$:

\[
A = \begin{bmatrix} 2 & -3 & 0 \\ 1 & 0 & -1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 7 & -1 \end{bmatrix}.
\]

**Solution:** It is just a matter of performing component-wise addition. We have

\[
A + B = \begin{bmatrix} (2) + (1) & (-3) + (1) & (0) + (1) \\ (1) + (0) & (0) + (7) & (-1) + (-1) \end{bmatrix} = \begin{bmatrix} 3 & -2 & 1 \\ 1 & 7 & -2 \end{bmatrix}.
\]

The only really important note to make about matrix addition is that the matrices being added must have exactly the same dimensions. We are not allowed, for instance, to add a 2-by-6 matrix to a 3-by-5 matrix, or even a 2-by-6 matrix to a 2-by-5 matrix. Passing this test, however, the procedure is exactly as easy as component-wise addition.

Section 3: Matrix Multiplication

To motivate the discussion of matrix multiplication, let’s consider multiplying a matrix by a scalar value. We have the following definition.
**Definition 4**

Let \( A \) be an \( m \times n \) matrix and \( c \in \mathbb{R} \) be a real number. Then the matrix \( c \cdot A \) is defined to be the matrix with entries \( c \cdot A = [c \cdot a_{ij}] \). This operation is called **scalar multiplication**.

**Example 3**

Given the following matrix \( A \), find \( 2 \cdot A \) and \( (-1) \cdot A \):

\[
A = \begin{bmatrix} 2 & -3 & 0 \\ 1 & 0 & -1 \end{bmatrix}
\]

**Solution:** We can easily evaluate

\[
2 \cdot A = 2 \cdot \begin{bmatrix} 2 & -3 & 0 \\ 1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 4 & -6 & 0 \\ 2 & 0 & -2 \end{bmatrix}
\]

and

\[
(-1) \cdot A = (-1) \cdot \begin{bmatrix} 2 & -3 & 0 \\ 1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} -2 & 3 & 0 \\ -1 & 0 & 1 \end{bmatrix}.
\]

Now consider multiplying two matrices \( A \) and \( B \) together. We are tempted to think the operation should be carried out component-wise since this is what we did with matrix addition and scalar multiple. It turns out, however, that this is not the standard, or most useful, definition. (Although component-wise multiplication is defined, and is called the **Hadamard product**). We will just not have any applications of that operation in this course.

We will use the following definition.

**Definition 5**

Let \( A \) and \( B \) be matrices with dimension \( m \)-by-\( p \) and \( p \)-by-\( n \) respectively. Then the matrix \( C = A \cdot B \) (or simply \( AB \)) is the matrix with entries

\[
c_{ij} = \sum_{k=1}^{p} a_{ik}b_{kj}.
\]
This operation is called **matrix multiplication**.

**Note:** The more common intuition of matrix multiplication is that we multiply the elements of the \(i^{th}\) row of \(A\) by the elements of the \(j^{th}\) column of \(B\), add the results, and place the final product in the \(i^{th}\) row and \(j^{th}\) column of the new matrix.

### Example 4

For the following matrices \(A\) and \(B\), determine the products \(A \cdot B\) and \(B \cdot A\):

\[
A = \begin{bmatrix} -2 & 1 & 0 & 1 \\ 1 & 1 & -2 & -1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 3 & 5 \\ 2 & 0 \\ 1 & -3 \\ 1 & 1 \end{bmatrix}.
\]

**Solution:** The matrix \(A \cdot B\) is given by

\[
A \cdot B = \begin{bmatrix} -2 & 1 & 0 & 1 \\ 1 & 1 & -2 & -1 \end{bmatrix} \begin{bmatrix} 3 & 5 \\ 2 & 0 \\ 1 & -3 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -3 & -9 \\ 2 & 10 \end{bmatrix}.
\]

while the matrix \(B \cdot A\) is given by:

\[
B \cdot A = \begin{bmatrix} 3 & 5 \\ 2 & 0 \\ 1 & -3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -2 & 1 & 0 & 1 \\ 1 & 1 & -2 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 8 & -10 & -2 \\ -4 & 2 & 0 & 2 \\ -5 & -2 & 6 & 4 \\ -1 & 2 & -2 & 0 \end{bmatrix}.
\]
This example raises a few very important warnings about matrix multiplication. First of all, the dimension is very important. We need to have the interior dimension be the same in order for the multiplication of matrices to be defined. We can also see that, even when multiplication is defined, order matters. In general, matrix multiplication is not commutative, that is \( A \cdot B \neq B \cdot A \). This is a significant difference between multiplication of numbers, where \( a \cdot b = b \cdot a \) held trivially!

We would take some time to clarify which common results from algebra extend to matrices. We will omit the proofs, but they can be found in multiple source (including Wikipedia).

**Theorem 1**

Consider matrices \( A, B, \) and \( C, \) and a constant \( k \in \mathbb{R} \). Then, whenever the following operations are defined, we have:

1. **Associativity:** \((A \cdot B) \cdot C = A \cdot (B \cdot C)\)
2. **Distributivity:** \(A \cdot (B + C) = A \cdot B + A \cdot C\) and \((A + B) \cdot C = A \cdot C + B \cdot C\)
3. **Scalar Multiplication:** \(k \cdot A \cdot B = (k \cdot A) \cdot B = A \cdot (k \cdot B)\)
4. **Transpose:** \((A \cdot B)^T = B^T \cdot A^T\)

**Remark:** Notice that, even without computing anything, we can determine whether multiplication of two matrices \( A \) and \( B \) is allowed, and what dimension the resulting matrix is. For an \( m \times p \) matrix \( A \) and \( p \times n \) matrix \( B \), we have that \( A \cdot B \) is defined and is an \( m \times n \) matrix. We know the operation is defined because the interior dimensions match, and the dimension of the resulting matrix is given by simply removing the interior dimensions.

**Section 4: Vectors**

Matrices may seem like a foreign concept at first glance, but we have all probably already seen matrices in other contexts. In particular, matrices
with only a single relevant dimension (i.e. an \( n \times 1 \) or \( 1 \times n \) matrix) commonly go by another name: they are called **vectors**.

**Definition 6**

A \( 1 \times n \) matrix is called an \( n \)-dimensional **row vector** and an \( n \times 1 \) matrix is called an \( n \)-dimensional **column vector**.

Vectors are commonly denoted with boldface (e.g. \( \mathbf{v}, \mathbf{w} \), etc.) or with a small arrow overtop (e.g. \( \vec{v}, \vec{w} \), etc.). In the notes, we will use \( \mathbf{v} \) while in lecture we will use \( \vec{v} \) (due to the difficulty in writing bold-face with whiteboard markers!). Since vectors only have one relevant dimension, they will be indexed with a single index (e.g. \( \mathbf{v} = (v_1, v_2, \ldots, v_n) \)). Unless otherwise stated, we will assume a vector \( \mathbf{v} \) is a column vector by default. Note that, if \( \mathbf{v} \) is a column vector, then \( \mathbf{v}^T \) is a row vector, and vice-versa.

We will encounter vectors in a number of contexts throughout the remainder of this course but we will not study them in depth until a few weeks from now. It is worth noting now, however, that the matrix operations we have defined for matrices are also valid for vectors. In fact, we may have seen a few of them already in earlier algebra, calculus, computer science, and/or physics courses.

**Definition 7**

Given \( n \)-dimensional vectors \( \mathbf{v} = (v_1, v_2, \ldots, v_n) \) and \( \mathbf{w} = (v_1, v_2, \ldots, v_n) \), and a constant \( c \in \mathbb{R} \), then we define the following operations:

1. **vector addition**: \( \mathbf{v} + \mathbf{w} = (v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n) \);
2. **vector scalar multiplication**: \( c \cdot \mathbf{v} = (c \cdot v_1, c \cdot v_2, \ldots, c \cdot v_n) \); and
3. **dot product**: \( \mathbf{v} \cdot \mathbf{w} = \sum_{i=1}^{n} v_i \cdot w_i \).

Notice that these are exactly the same operations we have previous defined for matrices. At first glance, the **dot product** may appear different than anything we have seen so far, but it can always be written as the matrix multiplication by taking one or the other of \( \mathbf{v} \) or \( \mathbf{w} \) to be its transpose (i.e. writing \( \mathbf{v}^T \cdot \mathbf{w} \) or \( \mathbf{v} \cdot \mathbf{w}^T \)). The only difference after that is that we have dropped the matrix notation for the end result. (Since multiplying a 1-by-\( n \)
matrix by an \( n \)-by-1 matrix produces a 1-by-1 matrix, we may just consider it to be a scalar.)

**Example 5**

Suppose we have \( \mathbf{v} = (1, 2, -1) \), \( \mathbf{w} = (0, -1, 3) \), and \( c = 5 \). Determine \( \mathbf{v} + \mathbf{w} \), \( c \cdot \mathbf{v} \), \( c \cdot \mathbf{w} \), and \( \mathbf{v} \cdot \mathbf{w} \).

**Solution:** By the definitions, we have

\[
\begin{align*}
\mathbf{v} + \mathbf{w} &= (1, 2, -1) + (0, -1, 3) = (1, 1, 2), \\
5 \cdot \mathbf{v} &= 5 \cdot (1, 2, -1) = (5(1), 5(2), 5(-1)) = (5, 10, -5), \\
5 \cdot \mathbf{w} &= 5 \cdot (0, -1, 3) = (5(0), 5(-1), 5(3)) = (0, -5, 15), \\
\mathbf{v} \cdot \mathbf{w} &= (1, 2, -1) \cdot (0, -1, 3) = (1)(0) + (2)(-1) + (-1)(3) = -5.
\end{align*}
\]

Notice that vector addition and scalar multiplication produce vectors of the same dimension as the original vectors, but that the dot product produces a scalar.

**Section 5: Linear Systems**

An immediate application of vectors and matrix multiplication comes from reconsidering linear systems. Define the \( m \)-by-\( n \) matrix \( A = [a_{ij}] \) and the vectors \( \mathbf{x} = (x_1, x_2, \ldots, x_n) \) and \( \mathbf{b} = (b_1, b_2, \ldots, b_m) \). Then we have that the matrix expression \( A \cdot \mathbf{x} = \mathbf{b} \) expands to

\[
\begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{bmatrix} =
\begin{bmatrix}
b_1 \\
b_2 \\
\vdots \\
b_m
\end{bmatrix}.
\]

Expanding this row by row gives the familiar linear system of \( m \) equa-
tions in \( n \) unknowns:

\[
\begin{align*}
    a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
    a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
    \vdots \\
    a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m.
\end{align*}
\]

We will prefer the form \( A \cdot \mathbf{x} = \mathbf{b} \) to the explicit form above for a number of reasons. The most apparent is that it is a more compact form; however, we will see shortly that there is also theory which will allows us a significantly more in-depth study of such equations.

**Example 6**

Rewrite the linear system

\[
\begin{align*}
    x - 2y + z &= 4 \\
    -3x + y + 2z &= 13 \\
    x - z &= -6.
\end{align*}
\]

in the matrix form \( A \cdot \mathbf{x} = \mathbf{b} \).

**Solution:** We have

\[
\begin{bmatrix}
    1 & -2 & 1 \\
    -3 & 1 & 2 \\
    1 & 0 & -1
\end{bmatrix}
\begin{bmatrix}
    x \\
    y \\
    z
\end{bmatrix}
=
\begin{bmatrix}
    4 \\
    13 \\
    -6
\end{bmatrix}.
\]

Notice that we must account for zero coefficient in the third equation (1) \( x + (0)y + (-1)z = -6 \) even though we would not typically write it in the expanded form.

**Section 6: Special Matrices**

There are several matrices which will turn out to have particularly nice properties. The first is a matrix which is particularly well-behaved under the **matrix addition** operator.
Definition 8
The **zero matrix** of dimension \( m \times n \) is denoted \( \mathbf{0} \) and is defined as the \( n \times m \) matrix with zeroes in every entry.

The key feature of the zero matrix is that it does not change a matrix upon matrix addition. That is to say \( A + \mathbf{0} = \mathbf{0} = \mathbf{0} + A \) for all matrices \( A \) where matrix addition is defined.

Next, we define a square matrix which is particularly well behaved under the **matrix multiplication** operation.

Definition 9
The **identity matrix** of dimension \( n \) is denoted \( \mathbf{I} \) and is defined as the \( n \times n \) matrix with entries

\[
a_{ij} = \begin{cases} 
1, & \text{if } i = j \\
0, & \text{if } i \neq j.
\end{cases}
\]

In other words, it is the matrix with ones along the main diagonal and zeroes elsewhere:

\[
\mathbf{I} = \begin{bmatrix}
1 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1
\end{bmatrix}.
\]

Example 7
For the following matrix \( A \), evaluate \( \mathbf{I} \cdot A \) and \( A \cdot \mathbf{I} \) (where \( \mathbf{I} \) is the identity matrix of appropriate dimensions):

\[
A = \begin{bmatrix}
2 & -1 & -2 \\
0 & 3 & 1
\end{bmatrix}.
\]

**Solution:** Since \( A \) is a \( 2 \times 3 \) matrix, the operation \( \mathbf{I} \cdot A \) requires the \( 2 \times 2 \) identity matrix while the operation \( A \cdot \mathbf{I} \) requires a \( 3 \times 3 \) matrix.
We have
\[
I \cdot A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & -2 \\ 0 & 3 & 1 \end{bmatrix} = \begin{bmatrix} (1)(2) + (0)(-1) & (1)(-1) + (0)(3) & (1)(-2) + (0)(1) \\ (0)(2) + (1)(0) & (0)(-1) + (1)(3) & (0)(-2) + (1)(1) \end{bmatrix} = \begin{bmatrix} 2 & -1 & -2 \\ 0 & 3 & 1 \end{bmatrix}
\]
and
\[
A \cdot I = \begin{bmatrix} 2 & -1 & -2 \\ 0 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} (2)(1) + (-1)(0) + (-2)(0) & (2)(-1) + (-1)(1) + (-2)(0) & (2)(0) + (-1)(0) + (-2)(1) \\ (0)(1) + (3)(0) + (1)(0) & (0)(-1) + (3)(1) + (1)(0) & (0)(0) + (3)(0) + (1)(1) \end{bmatrix} = \begin{bmatrix} 2 & -1 & -2 \\ 0 & 3 & 1 \end{bmatrix}.
\]

In other words, the identity matrix has the property that it does not change a matrix upon matrix multiplication. For this reason, it can be thought of as the matrix multiplication equivalent of the number 1 (since \(1 \cdot x = x \cdot 1 = x\) for all numbers \(x \in \mathbb{R}\)).

**Section 7: Matrix Inverses**

Reconsider the linear system
\[
A \cdot x = b
\]
where \(A\) is a \(n \times n\) matrix and \(x\) and \(b\) are \(n\)-dimensional vectors. Notice that a solution to this expression amounts to solving for \(x\); that is, rearranging the matrix expression to the form \(x = \cdots\).

In order to consider how we might accomplish this using matrix operation, we consider the following algebraic equation:
\[
a \cdot x = b.
\]
It is easy to solve for \(x\) in this expression. We simply divide both the left-hand and right-hand sides by \(a\) to get
\[
x = \frac{b}{a}.
\]
When we consider the linear system $A\mathbf{x} = \mathbf{b}$, we quickly realize that the solution is to divide by $A$. Since $A$ is a matrix, not a number, this is no trivial task. After all, we have already seen that multiplication of matrices is more complicated than simply performing the operations component-wise. Division will be similarly more complicated.

The trick is to introduce the inverse of a matrix. By analogy, consider the following (somewhat pedantic) sequence of steps to attain the solution for the $a \cdot x = b$:

$$ax = b \implies \left(\frac{1}{a}\right) \cdot a \cdot x = \left(\frac{1}{a}\right) \cdot b \implies 1 \cdot x = \left(\frac{1}{a}\right) \cdot b \implies x = \frac{b}{a}.$$ 

When considering algebraic expressions, writing all this in place of a simple division operation is grown-inducing. We will see, however, that when we perform matrix algebra we will have no choice but to follow this detailed logic since there is no explicitly-defined division operation. Rather, we will need to find a matrix $A^{-1}$ which satisfies the relationship $A^{-1} \cdot A = I$. If we could find such a matrix, then we would have

$$A \cdot \mathbf{x} = \mathbf{b} \implies A^{-1} \cdot A \cdot \mathbf{x} = A^{-1} \cdot \mathbf{b} \implies I \cdot \mathbf{x} = A^{-1} \cdot \mathbf{b} \implies \mathbf{x} = A^{-1} \cdot \mathbf{b}.$$ 

That is, we could solve for $\mathbf{x}$ by isolating it in the matrix expression.

**Example 8**

Show that the matrix

$$A^{-1} = \begin{bmatrix} -1 & 1 & 2 \\ -2 & 2 & 3 \\ 1 & 0 & -1 \end{bmatrix}$$

is the inverse of

$$A = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 1 & 1 \\ 2 & -1 & 0 \end{bmatrix}.$$ 

Use this fact to find the solution of the following linear system:

$$\begin{align*}
2x - y + z &= 1 \\
-x + y + z &= 0 \\
2x - y &= -1.
\end{align*} \quad (2)$$

**Solution:** To check that $A^{-1}$ is the inverse of $A$, we need to check
$A^{-1} \cdot A$. Explicitly, we have that

$$A^{-1} \cdot A = \begin{bmatrix}
-1 & 1 & 2 \\
-2 & 2 & 3 \\
1 & 0 & -1
\end{bmatrix} \begin{bmatrix}
2 & -1 & 1 \\
-1 & 1 & 1 \\
2 & -1 & 0
\end{bmatrix} = \begin{bmatrix}
-1(2) + (1)(-1) + (2)(2) & (-1)(-1) + (1)(1) + (2)(-1) & (-1)(1) + (1)(1) + (2)(0) \\
-2(2) + (2)(-1) + (3)(2) & (-2)(-1) + (2)(1) + (3)(-1) & (-2)(1) + (2)(1) + (3)(0) \\
1(2) + (0)(-1) + (-1)(2) & (1)(-1) + (0)(1) + (-1)(-1) & (1)(1) + (0)(1) + (-1)(0)
\end{bmatrix} = \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}.$$  

It follows that $A^{-1}$ is truly the inverse of $A$. Now notice that we can write the linear system of equations (2) as

$$\begin{bmatrix}
2 & -1 & 1 \\
-1 & 1 & 1 \\
2 & -1 & 0
\end{bmatrix} \begin{bmatrix}
x_1 \\
x_2 \\
x_3
\end{bmatrix} = \begin{bmatrix}
1 \\
0 \\
-1
\end{bmatrix}.$$  

By the previous sequence of matrix operations, we have that the solution must be

$$A \cdot x = b \implies x = A^{-1} \cdot b$$  

so that

$$\begin{bmatrix}
x_1 \\
x_2 \\
x_3
\end{bmatrix} = \begin{bmatrix}
-1 & 1 & 2 \\
-2 & 2 & 3 \\
1 & 0 & -1
\end{bmatrix} \begin{bmatrix}
1 \\
0 \\
-1
\end{bmatrix} = \begin{bmatrix}
-3 \\
-5 \\
2
\end{bmatrix}.$$  

It follows that $x_1 = -3$, $x_2 = -5$, and $x_3 = 2$. It can be verified through Gaussian elimination that this is the correct solution.

It is important to remember that matrix multiplication in general is not commutative. In particular, in general we have that $A^{-1} \cdot b \neq b \cdot A^{-1}$.

We will have to be careful to be consist when multiplying matrices on the left or the right of a given expression.

The inverse $A^{-1}$ allows us to perform the algebraic operation of division for matrices. We should pause, however, to give a formal definition and consider some properties and theorems resulting from this definition. The following is perhaps the most important property of a square matrix.

**Definition 10**

Suppose $A$ is an $n \times n$ matrix. Then $A$ will be called **invertible** if there exists an $n \times n$ matrix $B$ for which $B \cdot A = A \cdot B = I$. Such a matrix $B$ will be called the **inverse** of $A$ and will commonly be denoted $A^{-1}$. 

---

13
We should have suspected that inverses do not always exist. After all, we are not always able to solve \( a \cdot x = b \) since we cannot divide by \( a = 0 \). This exception is trivial in the algebra of real numbers, but significantly more subtle when considering matrix algebra.

We also have the following properties of matrix inverses, which allows us to definitively say that a matrix \( A^{-1} \) is “the” inverse of a matrix \( A \). The properties are simple enough to prove that will do so.

**Theorem 2**
Consider an \( n \times n \) matrix \( A \). Then:

1. If an inverse \( A^{-1} \) exists, it is unique.
2. If \( A^{-1} \) is the inverse of \( A \), then \( A \) is the inverse of \( A^{-1} \).

**Proof**

*Property 1.* Suppose there are \( B \) and \( C \) so that \( B \cdot A = I \) and \( A \cdot C = I \). Basic operations allow us to write

\[
B = B \cdot I = B \cdot (A \cdot C) = (B \cdot A) \cdot C = I \cdot C = C.
\]

It follows that the left inverse and right inverse are the same matrix, and that it is unique.

*Property 2.* This follows immediately from \( A^{-1} \cdot A = A \cdot A^{-1} = I \).

Notice that we have not proved that the existence of a left inverse (i.e. \( A^{-1} \cdot A = I \)) implies the existence of a right inverse (i.e. \( A \cdot A^{-1} = I \)). We will consider the detail further when we consider determinants.

**Section 8: Computation of Inverses**

It is easy to verify that two given matrices \( A \) and \( B \) are inverses of one another. We also know that inverses are unique, and can be multiplied on either side of the original matrix. Unfortunately, none of this information gives us any information on how to find matrix inverses.
We start consideration of this case by considering the simplest non-trivial case: 2-by-2 matrices. So, given a matrix

\[
A = \begin{bmatrix}
  a & b \\
  c & d \\
\end{bmatrix},
\]

let’s try to find a matrix \( B \) such that \( A \cdot B = I \) (we can write this as \( B \cdot A = I \), but the other order will be more convenient). We want to solve

\[
\begin{bmatrix}
  a & b \\
  c & d \\
\end{bmatrix} \begin{bmatrix}
  b_1 & b_2 \\
  b_3 & b_4 \\
\end{bmatrix} = \begin{bmatrix}
  1 & 0 \\
  0 & 1 \\
\end{bmatrix}.
\]

This system may look daunting, but there is a very important simplification we can make. If we consider the columns of \( B \) and \( I \) separately, this is equivalent to the two system of equations

\[
\begin{align*}
  \begin{bmatrix}
    a & b \\
    c & d \\
  \end{bmatrix} \begin{bmatrix}
    b_1 \\
    b_3 \\
  \end{bmatrix} &= \begin{bmatrix}
    1 \\
    0 \\
  \end{bmatrix} \\
  \begin{bmatrix}
    a & b \\
    c & d \\
  \end{bmatrix} \begin{bmatrix}
    b_2 \\
    b_4 \\
  \end{bmatrix} &= \begin{bmatrix}
    0 \\
    1 \\
  \end{bmatrix}.
\end{align*}
\]

These matrix equations are immediate recognizable as the matrix form of a linear system of equations of two equations in two unknowns (\( b_1 \) and \( b_3 \), and \( b_2 \) and \( b_4 \), respectively). In other words, we can solve \( b_1, b_2, b_3, \) and \( b_4 \) (i.e. we can solve for \( B \)) by Gaussian elimination! The corresponding coefficient matrices are

\[
\begin{bmatrix}
  a & b & 1 \\
  c & d & 0 \\
\end{bmatrix} \quad \text{and} \quad \begin{bmatrix}
  a & b & 0 \\
  c & d & 1 \\
\end{bmatrix}.
\]

Solving the first system will solve for \( b_1 \) and \( b_3 \) while the second system will solve for \( b_2 \) and \( b_4 \).

We could perform these operations directly in this form, but there is another simplification we can make. In order to achieve the row-reduced echelon form, we would be duplicating the same process on the left-hand side. This is because the left-hand sides are identical. We can avoid duplication of arithmetic and solve both linear systems simultaneously by writing

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

When we perform row reduction, we will have to remember that the first column to the right of the line solves for \( b_1 \) and \( b_3 \), while the second column to the right of the line solves for \( b_2 \) and \( b_4 \).

Fortunately, we can perform this row reduction operation directly. Consider the following example.
Example 9

Determine the inverse of the following matrix:

$$A = \begin{bmatrix} 3 & -2 \\ -1 & 4 \end{bmatrix}.$$ 

Solution: We want to find a matrix $B$ such that $A \cdot B = I$, so that we want to solve

$$\begin{bmatrix} 3 & -2 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$ 

The coefficient matrix for this system is

$$\begin{bmatrix} 3 & -2 & 1 & 0 \\ -1 & 4 & 0 & 1 \end{bmatrix} \xrightarrow{\text{swap}} \begin{bmatrix} -1 & 4 & 0 & 1 \\ 3 & -2 & 1 & 0 \end{bmatrix}.$$ 

$$R_1' = -R_1 \quad \xrightarrow{R_2' = R_2 - 3R_1} \quad \begin{bmatrix} 1 & -4 & 0 & -1 \\ 0 & 1 & \frac{1}{10} & \frac{3}{10} \end{bmatrix}.$$ 

$$R_2' = (1/10)R_2 \quad \xrightarrow{R_1' = R_1 + 4R_2} \quad \begin{bmatrix} 1 & 0 & \frac{2}{5} & \frac{1}{5} \\ 0 & 1 & \frac{1}{10} & \frac{3}{10} \end{bmatrix}.$$ 

It follows that the inverse is

$$A^{-1} = B = \begin{bmatrix} \frac{2}{5} & \frac{1}{10} \\ \frac{1}{10} & \frac{3}{10} \end{bmatrix}.$$ 

It can checked directly that this inverse is correct (and works on both the left and right). We should be slightly distressed, however, by the amount of work that took. It turns out that, for $2 \times 2$ matrices, the general formula for inverses is well-known and relatively compact to write down. We have the following result.

Theorem 3

Consider an invertible $2 \times 2$ matrix of the form

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$
Then the inverse $A^{-1}$ has the form
\[
A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.
\]

**Proof**
Omitted! Try it for yourself. You can either perform the Guassian elimination method (the hard way!) or directly verify the result (the easy way!).

**Example 10**
Use Theorem 3 to determine the inverse of
\[
A = \begin{bmatrix} 3 & -2 \\ -1 & 4 \end{bmatrix}.
\]

**Solution:** We have
\[
A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{(3)(4) - (-2)(-1)} \begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} \frac{2}{5} & \frac{1}{5} \\ \frac{1}{10} & \frac{3}{10} \end{bmatrix}.
\]
This agrees with our earlier result.

In addition to allow us to compute $2 \times 2$ inverses, Theorem 3 allows us to easily identify the condition for the invertibility of a $2 \times 2$ matrix. In order for the matrix to be invertible we need to have $ad - bc \neq 0$; otherwise, we divide by zero and the expression is undefined. We will examine and exploit this condition more significantly when we consider determinants of matrices.

Unfortunately, the closed form representation for inverses of the form given in Theorem 3 is not attainable for higher dimension. For such matrices, we will have to resort to the Guassian elimination method. Consider the following example, which motivated our study of inverses.
Example 11

Determine the inverse of the following matrix:

\[
A = \begin{bmatrix}
2 & -1 & 1 \\
-1 & 1 & 1 \\
2 & -1 & 0
\end{bmatrix}.
\]

Solution: We set up the linear system \(A \cdot B = I\), which in expanded form is

\[
\begin{bmatrix}
2 & -1 & 1 \\
-1 & 1 & 1 \\
2 & -1 & 0
\end{bmatrix}
\begin{bmatrix}
b_1 \\
b_4 \\
b_7
\end{bmatrix}
= \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}.
\]

We can solve for \(b_1, \ldots, b_9\) by performing Gaussian elimination on the following coefficient matrix:

\[
\begin{bmatrix}
2 & -1 & 1 & 1 & 0 & 0 \\
-1 & 1 & 1 & 0 & 1 & 0 \\
2 & -1 & 0 & 0 & 0 & 1
\end{bmatrix}
\]

\[
R'_1 = -R_2 \\
R'_2 = R_1 \\
R'_3 = R_2 - 2R_1 \\
R'_4 = R_3 - 2R_1 \\
R'_5 = R_1 + R_2 \\
R'_6 = R_3 - R_2 \\
R'_7 = -R_3 \\
R'_8 = R_1 - 2R_3 \\
R'_9 = R_2 - 3R_3
\]

Following the logic from before, we can read off the inverse matrix as
what is left on the right-hand side. We have

\[ A^{-1} = B = \begin{bmatrix} -1 & 1 & 2 \\ -2 & 2 & 3 \\ 1 & 0 & -1 \end{bmatrix} \]

which we have already verified is the correct inverse.

Suggested Problems

1. Consider the following matrices:

\[ A = \begin{bmatrix} 1 & 3 \\ -1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 2 & -1 \\ 0 & 5 & 7 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 1 \\ 1 & -1 \\ 2 & 0 \end{bmatrix}. \]

Compute the following: \( A \cdot B, \quad B + C^T, \quad 2 \cdot B \cdot C, \quad C \cdot A \cdot B, \quad B \cdot A \cdot C. \)

If the operation is not defined, explain why not.

2. Determine the inverses of the following matrices:

\[ A = \begin{bmatrix} 3 & -4 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} -2 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & 2 & -1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & -3 \\ 1 & -1 & 5 \\ 0 & 1 & -1 \end{bmatrix}. \]

3. Use the matrix inversion method to solve the following system of equations:

\[
\begin{align*}
2x + y + 3z &= 1 \\
-2x + y - 2z &= 0 \\
x + y + 2z &= 3.
\end{align*}
\]

4. Determine the value(s) of \( k \) for which the following system has an infinite number of solutions:

\[
\begin{bmatrix} -2 & 1 & 3 \\ -2 & 2 & k - 2 \\ 1 & -2 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ -1 \end{bmatrix}.
\]