Multiplying matrices

The definition of matrix multiplication that we use turns out to be the right one for a lot of reasons.

One reason is that it lets us solve simultaneous linear equations in several unknowns with a procedure that looks a lot like the one-variable procedure.

To solve $ax = b$, we multiply both sides by $a^{-1}$ to get $x = a^{-1}b$ (also known as $b/a$).

To solve

\[
ax + by = e \\
 cx + dy = f
\]

we write it as

\[
\begin{pmatrix}
a & b \\
c & d
\end{pmatrix}
\begin{pmatrix}
x \\
y
\end{pmatrix}
= \begin{pmatrix}
e \\
f
\end{pmatrix}
\]

or more compactly as $Ax = b$, where $x$ is the vector

\[
\begin{pmatrix}
x \\
y
\end{pmatrix}
\]

and $b$ is the vector

\[
\begin{pmatrix}
e \\
f
\end{pmatrix}
\]

and $A$ is the matrix

\[
\begin{pmatrix}
a & b \\
c & d
\end{pmatrix}
\] .

The way we solve the linear system $Ax = b$ is by “left-multiplying” the LHS and the RHS by the inverse matrix $A^{-1}$ (which we are assuming exists); then

\[ Ax = b \]

implies

\[ A^{-1}Ax = A^{-1}b \]

and the left hand side simplifies to $Ix$, which is just $x$, so we get

\[ x = A^{-1}b, \]
analogous to the formula $x = a^{-1}b$ we got in the single-variable case. (Note however that we do not write $A^{-1}$ as $1/A$ and we do not write $A^{-1}b$ as $b/A$.)

The determinant of

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is $ad - bc$, and it has the property that for any two 2-by-2 matrices $A, B$,

$$\det(AB) = \det(A) \det(B).$$

In fact, this formula is valid for $n$-by-$n$ matrices, for all $n$. In the case where $A$ is invertible and we set $B = A^{-1}$, we get

$$\det(AA^{-1}) = \det(A) \det(A^{-1}).$$

The left hand side simplifies to $\det(I)$, which is 1. So

$$1 = \det(A) \det(A^{-1}),$$

which implies

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

In particular, if $A$ is invertible, its determinant must be non-zero. In fact:

**Theorem:** The square matrix $A$ has an inverse if and only if $\det(A)$ is not equal to 0.

**Theorem:** If $A$ and $B$ are invertible square matrices of the same order, then

$$(AB)(B^{-1}A^{-1}) = I = (B^{-1}A^{-1})(AB),$$

implying that $AB$ is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.

**Proof:** Use associativity.

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$

and

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I.$$