# Math 341
# Notes on the inverse of a matrix

This is a slightly extended version of part of my lecture from Wednesday, March 4, 2015.

First we need a couple of lemmas.

**Lemma 1.** *If $A$ is an $m \times n$ matrix, then $A = 0$ if and only if $Ax = 0$ for all $x \in F^n$.*

*Proof.* Let $e_1, \ldots, e_n$ denote the standard basis vectors for $F^n$. Of course if $A = 0$, then $Ax = 0$ for all $x \in F^n$. Conversely, suppose $Ax = 0$ for all $x \in F^n$. Then, since $Ae_i$ is the $i^{\text{th}}$ column of $A$ for each standard basis vector $e_i$, we immediately conclude that each column of $A$ is the zero vector, i.e. every entry of $A$ is zero, i.e. $A = 0$. □

**Lemma 2.** *If $A$ and $B$ are $m \times n$ matrices, then $A = B$ if and only if $Ax = Bx$ for all $x \in F^n$.*

*Proof.* If $A = B$, then clearly $Ax = Bx$ for all $x \in F^n$. Conversely, suppose $Ax = Bx$ for all $x \in F^n$. Then, using basic matrix properties, we have $(A - B)x = Ax - Bx = 0$, for all $x \in F^n$. By Lemma 1, we conclude that $A - B = 0$, which means that $A = B$. □

You should interpret Lemma 2 as saying that two matrices being equal is the same as saying they do the same thing to any vector. In other words, $A = B$ if and only if $L_A = L_B$.

Now we can prove something useful about inverses. Recall that $I_n$ denotes the $n \times n$ identity matrix. Obviously, for any vector $x \in F^n$, we have $I_n x = x$. (This is why it's called the identity matrix.)

**Theorem 3.** *Let $A$ and $B$ be $n \times n$ matrices. If $AB = I_n$, then $BA = I_n$.*

*Proof.* Assume that $AB = I_n$. If we can show that $(BA)y = y$ for all $y$ in $F^n$, then we will have shown that $BA = I_n$, by Lemma 2.

Let $\{u_1, \ldots, u_n\}$ be a basis for $F^n$ (like the standard basis, for example). We want to prove that $\{Bu_1, \ldots, Bu_n\}$ is also a basis for $F^n$. It's enough to prove they are linearly independent (why)?

Suppose

$$c_1(Bu_1) + \cdots + c_n(Bu_n) = 0$$

for some scalars $c_1, \ldots, c_n$. Each side of this equation is a vector in $F^n$. We multiply both sides by $A$ on the left, yielding

$$A(c_1(Bu_1)) + \cdots + A(c_n(Bu_n)) = A0,$$

which can be simplified using basic properties of matrices to yield

$$c_1((AB)u_1 + \cdots + c_n((AB)u_n) = 0.$$

Since $AB = I_n$, we know $(AB)u_i = u_i$ for all $i$, so we have

$$c_1u_1 + \cdots c_nu_n = 0,$$

which means that $c_1 = c_2 = \cdots = c_n = 0$, since the vectors $u_1, \ldots u_n$ are linearly independent. We conclude that the vectors $Bu_1, \ldots Bu_n$ are linearly independent as well, which means they form a basis for $F^n$.

If $y$ is an arbitrary vector in $F^n$, then $y$ is therefore in the span of $\{Bu_1, \ldots, Bu_n\}$, which means that there are scalars $a_1, \ldots, a_n$ such that

$$a_1(Bu_1) + \cdots + a_n(Bu_n) = y,$$

which means that $B(a_1u_1 + \cdots + a_nu_n) = y$. Let $x = a_1u_1 + \cdots + a_nu_n$, so that $Bx = y$. Now we have

$$(BA)y = (BA)(Bx) = B((AB)x) = B(I_nx) = Bx = y.$$

This completes our proof.

Note: this proof went into more detail than the one I gave in class. You might want to compare the two when reading them for understanding.

Why is this theorem interesting? In the book, the following definition is given for an invertible $n \times n$ matrix:

**Definition 1.** An $n \times n$ matrix $A$ is *invertible* if there is an $n \times n$ matrix $B$ such that $AB = BA = I_n$.

The above theorem shows that $AB = I_n$ implies that $BA = I_n$, and vice versa. With this knowledge, you don't have to consider both $AB$ and $BA$ in showing that a matrix is invertible, or demonstrating that the inverse of $A$ is $B$. This is why the definition $I$ gave in class is actually equivalent to the one in the book. The definition I gave was:

**Definition 2.** An $n \times n$ matrix $A$ is *invertible* if there is an $n \times n$ matrix $B$ such that $AB = I_n$.

□