Math 1600A Lecture 17, Section 2, 18 Oct 2013

Announcements:

Read Section 3.5 for Monday. We aren't covering 3.4. Work through recommended homework questions.

Tutorials: Quiz 4 next week covers Section 3.3 and maybe some of 3.5 (I'll tell you on Monday).

Office hour: Monday, 1:30-2:30, MC103B.

Help Centers: Monday-Friday 2:30-6:30 in MC 106.

Partial review of Lecture 16:

Definition: An **inverse** of an $n \times n$ matrix A is an $n \times n$ matrix A' such that

$$AA' = I \quad \text{and} \quad A'A = I.$$

If such an A' exists, we say that A is **invertible**.

Theorem 3.6: If A is an invertible matrix, then its inverse is unique.

Because of this, we write A^{-1} for **the** inverse of A, when A is invertible. We do *not* write $\frac{1}{A}$.

Example: If $A = \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix}$, then $A^{-1} = \begin{bmatrix} 7 & -2 \\ -3 & 1 \end{bmatrix}$ is the inverse of A.

But the zero matrix and the matrix $B = \begin{bmatrix} -1 & 3 \\ 2 & -6 \end{bmatrix}$ are not invertible.

Theorem 3.7: If A is an invertible matrix $n \times n$ matrix, then the system $A\vec{x} = \vec{b}$ has the unique solution $\vec{x} = A^{-1}\vec{b}$ for any $\vec{b}$ in $\mathbb{R}^n$.

Remark: This is **not** in general an efficient way to solve a system.

Theorem 3.8: The matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible if and only if $ad - bc \neq 0$.

When this is the case,

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

We call $ad - bc$ the **determinant** of A, and write it $\det A$.
It *determines* whether or not A is invertible, and also shows up in the formula for A^{-1}.

Properties of Invertible Matrices

Theorem 3.9: Assume A and B are invertible matrices of the same size. Then:
a. A^{-1} is invertible and $(A^{-1})^{-1} = A$
b. If c is a non-zero scalar, then cA is invertible and $(cA)^{-1} = \frac{1}{c}A^{-1}$
c. AB is invertible and $(AB)^{-1} = B^{-1}A^{-1} \quad \text{(socks and shoes rule)}$
d. A^T is invertible and $(A^T)^{-1} = (A^{-1})^T$
e. A^n is invertible for all nonnegative integers n and $(A^n)^{-1} = (A^{-1})^n$

To verify these, in every case you just check that the matrix shown is an inverse.

Remark: Property (c) is the most important, and generalizes to more than two matrices, e.g. $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$.

Remark: For n a positive integer, we define A^{-n} to be $(A^{-1})^n = (A^n)^{-1}$. Then $A^nA^{-n} = I = A^0$, and more generally $A^rA^s = A^{r+s}$ for all integers r and s.

Remark: There is no formula for $(A+B)^{-1}$. In fact, $A+B$ might not be invertible, even if A and B are.

We can use these properties to solve a matrix equation for an unknown matrix.

New material

Challenge problem:

Can you find a 2×3 matrix A and a 3×2 matrix A' such that $AA' = I_2$ and $A'A = I_3$?

There's no problem getting $AA' = I_2$. (Find an example.)

But it's not possible to have $A'A = I_3$ with the given sizes.
Suppose we did have $A'A = I_3$ with A a 2×3 matrix.
Consider the homogenous system

$$A\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Since $\operatorname{rank} A \leq 2$ and there are three variables, this system must have infinitely many solutions. But

$$A\vec{x} = \vec{0} \quad \Longrightarrow \quad A'A\vec{x} = A'\vec{0} \quad \Longrightarrow \quad \vec{x} = \vec{0},$$

so there is only one solution. This is a contradiction.

More generally, unless A is square, you can't find a matrix A' that makes both $AA' = I$ and $A'A = I$ true.

The fundamental theorem of invertible matrices:

Very important! Will be used repeatedly, and expanded later.

Theorem 3.12: Let A be an $n \times n$ matrix. The following are equivalent:
a. A is invertible.
b. $A\vec{x} = \vec{b}$ has a unique solution for every $\vec{b} \in \mathbb{R}^n$.
c. $A\vec{x} = \vec{0}$ has only the trivial (zero) solution.
d. The reduced row echelon form of A is I_n.

Proof: We have seen that (a) $\Longrightarrow$ (b) in Theorem 3.7 above.

We'll use our past work on solving systems to show that (b) $\Longrightarrow$ (c) $\Longrightarrow$ (d) $\Longrightarrow$ (
will prove that (b), (c) and (d) are equivalent.

We will only partially explain why (b) implies (a).

(b) $\Longrightarrow$ (c): If $A\vec{x} = \vec{b}$ has a unique solution for every $\vec{b}$, then it's true when $\vec{b}$ happ
the zero vector.

(c) $\Longrightarrow$ (d): Suppose that $A\vec{x} = \vec{0}$ has only the trivial solution.
That means that the rank of A must be n.
So in reduced row echelon form, every row must have a leading 1.
The only $n \times n$ matrix in reduced row echelon form with n leading 1's is the identity r

(d) $\Longrightarrow$ (b): If the reduced row echelon form of A is I_n, then the augmented matrix $[$
reduces to $[I_n \mid \vec{c}\,]$, from which you can read off the unique solution $\vec{x} = \vec{c}$.

(b) $\implies$ (a) (partly): Assume $A\vec{x} = \vec{b}$ has a solution for every $\vec{b}$.
That means we can find $\vec{x}_1, \ldots, \vec{x}_n$ such that $A\vec{x}_i = \vec{e}_i$ for each i.
If we let $B = [\,\vec{x}_1 \mid \cdots \mid \vec{x}_n\,]$ be the matrix with the $\vec{x}_i$'s as columns, then

$$AB = A[\,\vec{x}_1 \mid \cdots \mid \vec{x}_n\,] = [A\vec{x}_1 \mid \cdots \mid A\vec{x}_n\,] = [\,\vec{e}_1 \mid \cdots \mid \vec{e}_n\,] = I_n.$$

So we have found a *right* inverse for A.
It turns out that $BA = I_n$ as well, but this is harder to see. □

Note: We have omitted (e) from the theorem, since we aren't covering elementary matrices. They are used to prove the other half of (b) $\implies$ (a).

We will see many important applications of Theorem 3.12. For now, we illustrate one theoretical application and one computational application.

Theorem 3.13: Let A be a square matrix. If B is a square matrix such that either $AB = I$ or $BA = I$, then A is invertible and $B = A^{-1}$.

Proof: If $BA = I$, then the system $A\vec{x} = \vec{0}$ has only the trivial solution, as we saw in the challenge problem. So (c) is true. Therefore (a) is true, i.e. A is invertible. Then:

$$B = BI = BAA^{-1} = IA^{-1} = A^{-1}. \text{ (The uniqueness argument again!)}$$

This is very useful! It means you only need to check multiplication in one order to know you have an inverse.

Gauss-Jordan method for computing the inverse

Theorem 3.14: Let A be a square matrix. If a sequence of row operations reduces A to I, then the **same** sequence of row operations transforms I into A^{-1}.

Why does this work? It's the combination of our arguments that (d) $\implies$ (b) and (b) $\implies$ (a). If we row reduce $[A \mid \vec{e}_i\,]$ to $[I \mid \vec{c}_i\,]$, then $A\vec{c}_i = \vec{e}_i$. So if B is the matrix whose columns are the $\vec{c}_i$'s, then $AB = I$. So, by Theorem 3.14, $B = A^{-1}$.

The trick is to notice that we can solve all of the systems $A\vec{x} = \vec{e}_i$ **at once** by row reducing $[A \mid I\,]$. The matrix on the right will be exactly B!

Example on whiteboard: Find the inverse of $A = \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix}$.

Illustrate proof of Theorem 3.14.

Example on whiteboard: Find the inverse of $A = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 3 \\ 1 & -2 & 5 \end{bmatrix}$.

Example on whiteboard: Find the inverse of $B = \begin{bmatrix} -1 & 3 \\ 2 & -6 \end{bmatrix}$.

So now we have a general purpose method for determining whether a matrix A is invertible, and finding the inverse:

1. Form the $n \times 2n$ matrix $[A \mid I\,]$.

2. Use row operations to get it into reduced row echelon form.

3. If a zero row appears in the left-hand portion, then A is not invertible.

4. Otherwise, A will turn into I, and the right hand portion is A^{-1}.

The trend continues: when given a problem to solve in linear algebra, we usually find a way to solve it using row reduction!

We aren't covering inverse matrices over $\mathbb{Z}_m$.

.