LECTURE 7: ELEMENTARY MATRICES AND MATRIX INVERSES

MA1111: LINEAR ALGEBRA I, MICHAELMAS 2016

1. Elementary Matrices

We say that M is an **elementary matrix** if it is obtained from the identity matrix I_n by one elementary row operation. For example, the following are all elementary matrices:

$$\begin{pmatrix} \pi & 0 \\ 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Fact. *Multiplying a matrix M on the left by an elementary matrix E performs the corresponding elementary row operation on M.*

Example. *If*

$$E = \begin{pmatrix} \pi & 0 \\ 0 & 1 \end{pmatrix},$$

then for any matrix $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, we have

$$EM = \begin{pmatrix} \pi \cdot a + 0 \cdot c & \pi \cdot b + 0 \cdot d \\ 0 \cdot a + 1 \cdot c & 0 \cdot b + 1 \cdot d \end{pmatrix} = \begin{pmatrix} \pi a & \pi b \\ c & d \end{pmatrix}.$$

This would have worked for any $2 \times n$ matrix, not only 2×2 matrices, however we will now be mostly interested in square matrices, which is why we have chosen to highlight this case.

2. Invertible matrices

We have seen above that solving general systems of linear equations can be related to the problem of solving the matrix equation $Ax = b$ for a fixed matrix A, a fixed vector b, and an unknown vector x. We have seen how to solve this equation using Gauss-Jordan elimination, but we want to discuss another method which is inspired by encoding the system in this way. What if we were to solve the same system just over the real numbers (which does correspond to the case when we have one equation in one variable, after all)? Well, the answer would be simple! If we solve $\alpha x = \beta$ for fixed real numbers α, β, then if $\alpha = 0$, there are either no solutions x or every real x is a solution (depending on whether $b = 0$ or not), but if $\alpha \neq 0$, we would just multiply by α^{-1} to get $x = \alpha^{-1}\beta$.

Date: October 14, 2016.

This seems like a trivial observation, but the analogous idea for more general matrices (not just 1×1 matrices) is more interesting. We are so accustomed to the idea that α^{-1} exists for any non-zero α that we rarely think about it. We may then ask when there is something like an inverse for matrices. This is made precise in the following definition.

Definition. A matrix A is **invertible** if there exists a matrix B such that

$$AB = BA = I_n.$$

In this case, we also say that B is an inverse of A.

Note that all invertible matrices are necessarily square (they have the same number of rows and columns).

Lemma. *If A has an inverse, than it is unique.*

Proof. Suppose that B and C are **both** inverses of A. We aim to show $B = C$. Indeed,

$$C = I_N C = (BA)C = BAC = B(AC) = BI_N = B.$$

□

Note that we have used both equations of the shape $AB = I_n$ and $BA = I_n$. In general settings, only requiring one of these equations to hold (so-called one-sided inverses) does not imply uniqueness of the inverse.

Because of this lemma, we can safely refer to any inverse of A (if it exists) as **the** inverse of A, and we denote it by A^{-1}. Now, clearly not everything has an inverse. Indeed, multiplication of 1×1 matrices is the same as multiplication of real numbers, and we already know that 0 isn't an inverse. But the problem runs deeper. Indeed, we saw that $AB = 0$ for some matrices A, B of which neither is equal to zero. If this equation holds and A^{-1} existed, then we could multiply both sides by it and obtain $B = A^{-1}0 = 0$, which is false by assumption. Thus, even in the case of non-zero square matrices of higher dimensions, there can be non-invertible matrices.

Here is one more very useful fact, which follows by a simple trick. In fact, it has a famous name, and is fondly referred to as the **socks and shoes property.**

Lemma. *If A and B are both invertible square matrices of size $n \times n$, then*

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

Proof. Since inverses are unique, it suffices to check that the right hand side is an inverse of AB. Indeed,

$$ABB^{-1}A^{-1} = AA^{-1} = I_n,$$

and

$$B^{-1}A^{-1}AB = B^{-1}B = I_n.$$

□

We return to applications to systems of linear equations. Clearly, if the system is represented by $Ax = b$, then if A is invertible, we can solve the system as $x = A^{-1}b$, and performing this multiplication gives the unique solution to the equation.

Example. *The matrix*

$$A = \begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix}$$

has inverse (check!)

$$A^{-1} = \begin{pmatrix} -7 & 3 \\ 5 & -2 \end{pmatrix}.$$

Now, the system of equations

$$\begin{cases} 2a + 3b = 4 \\ 5a + 7b = 1 \end{cases}$$

corresponds to the equation $Ax = b$ with $x = \begin{pmatrix} a \\ b \end{pmatrix}$, $A = \begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix}$, and $b = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$, so the solution is given by

$$x = A^{-1}b = \begin{pmatrix} -7 & 3 \\ 5 & -2 \end{pmatrix} \begin{pmatrix} 4 \\ 1 \end{pmatrix} = \begin{pmatrix} -25 \\ 18 \end{pmatrix}.$$

That is, the unique solution is $a = -25$, $b = 18$.

There are two critical questions. Firstly, when do inverses exist, and when they do, how do we compute them? One answer to the first question is the following, along with several other useful properties, is given as follows.

Theorem. *For any $n \times n$ matrix A, the following are equivalent.*

1. *Every vector $b \in \mathbb{R}^n$ is a linear combination of the columns of A (i.e., the column space of A is all of $\mathbb{R}^n$).*
2. *The system $Ax = b$ has a* **unique** *solution for all vectors b.*
3. *The system $Ax = b$ has a unique solution for* **some** *vector b.*
4. *The RREF of A is I_n.*
5. *A is a product of elementary matrices.*
6. *A is invertible.*

Proof. (2) $\Longrightarrow$ (3) is obvious; just pick $b = 0$, for example.

(3) $\Longrightarrow$ (4): We saw in our Gauss-Jordan algorithm before that there is a unique solution only if every variable is pivotal. This means that the RREF of $(A|b)$ is of the form $(I_n|*)$, which by forgetting about the $|b$ shows that the RREF of A is I_n.

(4) $\Longrightarrow$ (5): To any elementary row operation E there is a corresponding row operation E^{-1} which "undoes" the effect. It is clear that the corresponding matrices are inverses. Hence, every elementary matrix is invertible. Moreover, by using the socks and shoes property, we see that any product of invertible matrices is **invertible, so that**

a product of elementary matrices is. Moreover, this shows that the inverse of this product is itself a product of elementary matrices.

Now, if the RREF of A is I_n, then this precisely means that there are elementary matrices $E_1, \dots, E_m$ such that $E_1 E_2 \dots E_m A = I_n$. Multiplying both sides by the inverse of $E_1 E_2 \dots E_m$ shows that A is a product of elementary matrices.

(5) $\implies$ (6): The argument in the last step shows this.

(6) $\implies$ (2): If A is invertible, then the unique solution of $Ax = b$ is $x = A^{-1}b$.

(1) $\iff$ (2): We showed before that b is in the column space of A if and only if $Ax = b$ has a solution. Now we must discuss the uniqueness condition. If the solution isn't unique for some b, then then the system has infinitely many solutions, in which case the RREF of A has a row of all zeros. In this case, though, we can choose some b for which we get a pivot in the last column in the RREF of $(A|b)$, in which cases there are no solutions for some b, a contradiction.

□

3. Computing inverses

We now come to the second big question about inverses: how do we compute them? Using the last theorem, if A is invertible, then the RREF of A is I_n, so that for some elementary matrices $E_1, \dots, E_m$, we have $(E_1 \dots E_m)A = I_n$. Hence, the inverse of A is $E_1 \dots E_m$. We can compute this matrix by performing the same operation on I_n. That is, we have the following algorithm.

Algorithm. *If A is invertible, we can find its inverse by starting with the augmented matrix $(A|I_n)$ and row-reducing A to I_n. When we do this, we obtain the matrix $(I_n|A^{-1})$.*

Example. *Let's compute the inverse in the example above, for $A = \begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix}$. In this case, we form the matrix*

$$\left(\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ 5 & 7 & 0 & 1 \end{array}\right).$$

and row reduce

$$\left(\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ 5 & 7 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{cc|cc} 1 & \frac{3}{2} & \frac{1}{2} & 0 \\ 5 & 7 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{cc|cc} 1 & \frac{3}{2} & \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & -\frac{5}{2} & 1 \end{array}\right)$$
$$\longrightarrow \left(\begin{array}{cc|cc} 1 & \frac{3}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 5 & -2 \end{array}\right) \longrightarrow \left(\begin{array}{cc|cc} 1 & 0 & -7 & 3 \\ 0 & 1 & 5 & -2 \end{array}\right),$$

from which we read off that

$$A^{-1} = \begin{pmatrix} -7 & 3 \\ 5 & -2 \end{pmatrix},$$

which agrees with the claim above (and is easily checked directly).