21-241 MATRICES AND LINEAR TRANSFORMATIONS SUMMER 1 2012 COURSE NOTES DAY 5

PAUL MCKENNEY

Definition 1. A linear equation with a constant term of zero, ie of the form

$$a_1x_1 + \cdots + a_nx_n = 0$$

is called *homogeneous*. A system of linear equations is called *homogeneous* if each of its equations is homogeneous.

Note that a homogeneous system always has at least one solution, namely the zero vector.

Lemma 1. *Let A be an $m \times n$ matrix. If A is left-invertible then the* only *solution to the homogeneous system $Ax = 0$ is the zero vector.*

Proof. Let B be a left-inverse for A. (Why is it incorrect to talk about A^{-1} here?) Then if s is any solution to the homogeneous system $Ax = 0$, we have

$$s = Is = (BA)s = B(As) = B0 = 0$$

So in fact, 0 is the only solution. □

The proof of the following lemma will have to wait for a later date. I'll give a sketch, and an example of its use.

Lemma 2. *Suppose a system $Ax = b$, where A is $m \times n$, has at least one solution, $p \in \mathbb{R}^n$. Let k be the number of free variables in some echelon form of the system $Ax = b$. Then there exist solutions $h_1, \ldots, h_k \in \mathbb{R}^n$ to the homogeneous equation $Ax = 0$, such that*

$$\{s \in \mathbb{R}^n \mid As = b\} = \{p + c_1h_1 + \cdots + c_kh_k \mid c_1, \ldots, c_k \in \mathbb{R}\}$$

Proof sketch. The main idea here is that if $h \in \mathbb{R}^n$ is any solution to the homogeneous system $Ax = 0$, then

$$A(p + h) = Ap + Ah = Ap + 0 = Ap = b$$

and then so is $p + h$. Moreover if any $s \in \mathbb{R}^n$ is a solution to $Ax = b$, then

$$A(s - p) = As - Ap = b - b = 0$$

and so $s - p$ is a solution to the homogeneous system $Ax = 0$. This tells us that a vector $s \in \mathbb{R}^n$ is a solution to $Ax = b$ if and only if it's of the form $p + h$, where h is a solution to the homogeneous system $Ax = 0$. To see that there are fixed vectors $h_1, \ldots, h_k \in \mathbb{R}^n$ which generate the solutions to $Ax = 0$ will require more work; it essentially falls out of back-substitution. $\square$

Example. The system

$$\begin{aligned} x \quad\quad + \; z + \; w &= 2 \\ 2x - y \quad\quad + \; w &= 1 \\ x + y + 3z + 2w &= 5 \end{aligned}$$

has a solution $\begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$. What about the homogeneous system? It reduces to

$$\begin{aligned} x \;+\; z + w &= 0 \\ y + 2z + w &= 0 \\ 0 &= 0 \end{aligned}$$

Then the solution set to the homogeneous system is

$$\left\{ z \begin{pmatrix} -1 \\ -2 \\ 1 \\ 0 \end{pmatrix} + w \begin{pmatrix} -1 \\ -1 \\ 0 \\ 1 \end{pmatrix} \;\middle|\; z, w \in \mathbb{R} \right\}$$

and so the solution set to the original system of equations is

$$\left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} + z \begin{pmatrix} -1 \\ -2 \\ 1 \\ 0 \end{pmatrix} + w \begin{pmatrix} -1 \\ -1 \\ 0 \\ 1 \end{pmatrix} \;\middle|\; z, w \in \mathbb{R} \right\}$$

Lemma 3. *Suppose R is an invertible, $n \times n$ matrix in reduced row echelon form. Then $R = I$.*

Proof. Consider the homogeneous system $Rx = 0$. If there were any free variables in this system, then there would be a nonzero vector $h \in \mathbb{R}^n$ such that $Rh = 0$. But as R is invertible, by the lemma above this can't happen. So there can't be any free variables in this system. Since R is in reduced row echelon form, this means exactly that $R = I_n$. $\square$

Lemma 4. *If A and B are square, invertible matrices of the same size, then AB is also invertible. Moreover, $(AB)^{-1} = B^{-1}A^{-1}$.*

Proof. We have

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

and

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

□

Corollary 1. *If $A_1, \ldots, A_k$ are square, invertible matrices of the same size, then $A_1A_2\cdots A_k$ is also invertible, and $(A_1\cdots A_k)^{-1} = A_k^{-1}\cdots A_1^{-1}$.*

Proof. By induction on k. This proof is routine, but I'll include it here just to give an example of such a routine. When $k = 1$ the statement is trivial. So suppose it holds for k. To prove it for $k+1$, let $A_1, \ldots, A_k, A_{k+1}$ be square, invertible matrices of the same size. Then $A_1\cdots A_k$ is invertible and $(A_1\cdots A_k)^{-1} = A_k^{-1}\cdots A_1^{-1}$ by the induction hypothesis. Hence by the above lemma, $(A_1\cdots A_k)A_{k+1}$ is invertible, and

$$((A_1\cdots A_k)A_{k+1})^{-1} = A_{k+1}^{-1}(A_1\cdots A_k)^{-1} = A_{k+1}^{-1}(A_k^{-1}\cdots A_1^{-1})$$

Hence the statement is proven for $k+1$, and by induction it holds for all k. □

Lemma 5. *If A is an invertible matrix, then so is A^{-1}, and $(A^{-1})^{-1} = A$.*

Proof. We have

$$AA^{-1} = I \qquad A^{-1}A = I$$

simply by definition of A^{-1}. Then A is both a left and right inverse for A^{-1}. □

We've already essentially seen the following result, but I'll repeat the proof for clarity.

Lemma 6. *If E is an elementary matrix, then E is invertible, and E^{-1} is the elementary matrix which implements the row operation that reverses that which E implements.*

Proof. Suppose E is $m \times m$. Let F be the $(m \times m)$ elementary matrix which implements the reverse of the row operation E implements. (I'm not calling it E^{-1} yet because that would be presumptuous.) Then $(EF)A = A$ and $(FE)A = A$ for all $m \times n$ matrices A, for all n. In particular, with $A = I_m$, we get

$$EF = (EF)I_m = I_m \qquad FE = (FE)I_m = I_m$$

□

Theorem 1. *Let A be a square matrix. Then A is invertible if and only if A is row-equivalent to I.*

Proof. Let R be the reduced row echelon form of A. Since we can obtain R from A by row operations, there are elementary matrices $E_1, \ldots, E_k$ such that

$$R = E_kE_{k-1}\cdots E_1A$$

As we've seen before, elementary matrices are invertible. Hence by the above, R is invertible, and square. Then R is actually I.

If A is row-equivalent to I, then there are elementary matrices $E_1, \dots, E_k$ such that

$$E_k E_{k-1} \cdots E_1 A = I$$

Then

$$A = IA = (E_k E_{k-1} \cdots E_1)^{-1}(E_k E_{k-1} \cdots E_1)A = (E_k E_{k-1} \cdots E_1)^{-1} I = (E_k E_{k-1} \cdots E_1)^{-1}$$

□

The proof of the theorem above actually gives us a way of computing the inverse of a square matrix. I'll summarize it in this fact.

Fact 1. Let A be an invertible matrix, and let $E_1, \dots, E_k$ be elementary matrices reducing A to I, ie, such that

$$E_k E_{k-1} \cdots E_1 A = I$$

Then $A^{-1} = E_k E_{k-1} \cdots E_1 I$. In terms of row-operations, to find the inverse of A, we start with I and apply the same row operations we used to reduce A to I.

Typically, when finding the inverse of A, one performs row operations on A and I in parallel, as in the following example.

Example. Let $A = \left(\begin{smallmatrix} 2 & 7 \\ 1 & 4 \end{smallmatrix}\right)$. Then we compute A^{-1} using the following.

$$\begin{array}{rcc}
 & \begin{pmatrix} 2 & 7 \\ 1 & 4 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
\rho_1 \leftrightarrow \rho_2 & \begin{pmatrix} 1 & 4 \\ 2 & 7 \end{pmatrix} & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
\rho_2 \to \rho_2 - 2\rho_1 & \begin{pmatrix} 1 & 4 \\ 0 & -1 \end{pmatrix} & \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix} \\
\rho_1 \to \rho_1 + 4\rho_1 & \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} & \begin{pmatrix} 4 & -7 \\ 1 & -2 \end{pmatrix} \\
\rho_2 \to -\rho_2 & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 4 & -7 \\ -1 & 2 \end{pmatrix}
\end{array}$$

Now to check, we compute;

$$\begin{pmatrix} 2 & 7 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 4 & -7 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

So it worked!