LECTURE 1 I. Inverse matrices

We return now to the problem of solving linear equations. Recall that we are trying to find $\vec{x}$ such that

$$A\vec{x} = \vec{y}.$$

Recall: there is a matrix I such that

$$I\vec{x} = \vec{x}$$

for all $\vec{x} \in \mathbb{R}^n$. It follows that

$$IA = A$$

for all $n \times n$ matrices A.

For the rest of the day, suppose A is an $n \times n$ matrix.

Definition: Let A be an $n \times n$ matrix. We say a matrix B is an *inverse* for A if $AB = BA = I$.

Notation: If B is an inverse for A, we write $B = A^{-1}$, and we say A is *invertible.* Note that it is also true that $B^{-1} = A$ (check!)

Theorem (Invertibility Theorem I): Suppose A has an inverse A^{-1}. Then the equation

$$A\vec{x} = \vec{y}$$

has a unique solution, namely

$$\vec{x} = A^{-1}\vec{y}.$$

Remark: The converse of this theorem is also true, but we can't prove it yet. We'll do it later in the week when we discuss linear transformations.

Proof. We have

$$A\vec{x} = \vec{y}$$

from which it follows that

$$A^{-1}A\vec{x} = A^{-1}\vec{y},$$

and by definition of inverse,

$$I\vec{x} = A^{-1}\vec{y},$$

and by definition of the identity,

$$\vec{x} = A^{-1}\vec{y}.$$

So if there's a solution, this is it. To verify that this *is* a solution, check

$$A(A^{-1}\vec{y}) = (AA^{-1})\vec{y} = I\vec{y} = \vec{y}$$

which was the desired result. $\square$

Theorem: Suppose B is inverse to A. Then B is the *only* inverse to A.

Proof. Suppose B and B' were two different inverses to A. Then we would have

$$BAB' = (BA)B' = IB' = B'$$

but on the other hand

$$BAB' = B(AB') = BI = B$$

so $B = B'$. $\square$

Theorem: Suppose A and B are both invertible. Then AB is invertible, and its inverse is $B^{-1}A^{-1}$.

Proof. We observe that $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$. Left as exercise to check the other direction. $\square$

Not all matrices have inverses. For instance **Example:** $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ does not have an inverse. Because the equation

$$A\vec{x} = \vec{0}$$

has solutions $\begin{bmatrix} 0 \\ y \end{bmatrix}$ for *any* y; in particular, the solution is not unique, so if A were invertible it would violate the theorem above.

Some large classes of matrices have inverses.

Example: Let D be a *diagonal* matrix, i. e. one which has all its nonzero entries on the diagonal. Suppose that D has no zeroes on the diagonal. Then D is invertible.

In fact, its inverse can be written in a very simple form. For example:

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 4 \end{bmatrix}^{-1} = \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1/5 & 0 \\ 0 & 0 & 1/4 \end{bmatrix}.$$

Exercise: check this fact.

Example: Let $A = \begin{bmatrix} 1 & 1 & 1 \\ -3 & 1 & 0 \\ 1 & 2 & -8 \end{bmatrix}$. I claim that A has an inverse: namely,

$$A^{-1} = \begin{bmatrix} -1/7 & 5/14 & 1/14 \\ 13/14 & -9/28 & 1/28 \\ 3/14 & -1/28 & -3/28 \end{bmatrix}$$

To verify this, one needs only check that $AA^{-1} = A^{-1}A = I$. But how did I *find* the inverse?

II. Gaussian elimination and the inverse

In order to get this down, we begin with a definition.

Definition: An *elementary matrix* is an $n \times n$ matrix with 1's on the diagonal and exactly one non-zero entry off the diagonal.

Example:

$$E = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Observation: $E\vec{x}$ is the vector obtained by adding $-3\times$ the first coordinate of $\vec{x}$ to the second coordinate. Thus, if A is a $3 \times p$ matrix, EA is the matrix obtained from A by adding $-3\times$ the first row to the second row. Many of you observed this in the groupwork on Friday.

Note that E is invertible; its inverse is

$$E^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Exercise: check this.

Theorem Every elementary matrix is invertible. Proof left as an exercise. Use the form of E^{-1} above as a hint.

So we have

$$\begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 3 & 0 & 2 \\ 1 & 2 & -8 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & -1 \\ 1 & 2 & -8 \end{bmatrix}$$

Write

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}, E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1/3 & 1 \end{bmatrix}$$

Then $E_3 E_2 E_1 A$ is the matrix obtained from A by subtracting $3\times$ first row from second row, $1\times$ first row from third row, and adding $1/3\times$ second row to third row. But that's exactly what we did when we did Gaussian elimination. In other words, we have

$$E_3 E_2 E_1 A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & -1 \\ 0 & 0 & -28/3 \end{bmatrix}$$

Write U for that latter, upper triangular matrix. Now we can write

$$E_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 3/28 \\ 0 & 0 & 1 \end{bmatrix}$$

so that

$$E_4 U = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & 0 \\ 0 & 0 & -28/3 \end{bmatrix}$$

And likewise, E_5 and E_6 will be elementary matrices chosen to eliminate the other columns.

Finally, one gets

$$E_6 E_5 E_4 E_3 E_2 E_1 A = D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -28/3 \end{bmatrix}.$$

But we know that D is invertible, by the theorems. So can write

$$D^{-1} E_6 E_5 E_4 E_3 E_2 E_1 A = I.$$

Warning: We haven't really shown that $D^{-1}E_6E_5E_4E_3E_2E_1 = A^{-1}$, because we haven't shown that you get I when you multiply in the other order! See Strang for a way around this problem, or (better) look forward to our proof of the following theorem:

Theorem: Suppose A, B are square matrices with $AB = I$. Then $BA = I$.

LECTURE 2

I. Subspaces of $\mathbb{R}^n$.

This week we're really going to get into the basic theoretical notions of linear algebra–vector spaces and linear transformations. Let's jump right in.

Definition: A *subspace* of $\mathbb{R}^n$ is a subset $V \subset \mathbb{R}^n$ such that:

- If $\vec{v}$ and $\vec{w}$ are in V, then so is

$$\vec{v} + \vec{w}$$

- If $\vec{v}$ is in V, then so is $a\vec{v}$ for any $a \in \mathbb{R}$.

Example: Let $V \in \mathbb{R}^3$ be the set of vectors of the form

$$\begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix}.$$

Then $\vec{v} = \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} y \\ 0 \\ 0 \end{bmatrix}$ are elements of V, we have

$$\vec{v} + \vec{w} = \begin{bmatrix} x+y \\ 0 \\ 0 \end{bmatrix} \in V$$

and

$$a\vec{v} = \begin{bmatrix} ax \\ 0 \\ 0 \end{bmatrix} \in V.$$

Example: Let $V \in \mathbb{R}^3$ be the set of vectors of the form

$$\begin{bmatrix} x \\ 0 \\ 1 \end{bmatrix}.$$

This set is *not* a subspace. Because take, for example,

$$\vec{v} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \vec{w} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Then $\vec{v}$ and $\vec{w}$ are in V, but their sum $\vec{v} + \vec{w}$ is not. So, by definition, V is not a subspace.

2 minute contemplation: With a neighbor, think of all the subspaces of $\mathbb{R}^3$ you can.

Regroup, discuss.

Remark: a subspace is an example of a more general notion of a *vector space*, which section 2.1 of Strang discusses but which we, for the moment, shall not.

Definition: Let $\vec{v}_1, \ldots, \vec{v}_n$ be a set of vectors in $\mathbb{R}^m$. Then the **span** of $\vec{v}_1, \ldots, \vec{v}_n$ is the set of vectors

$$x_1\vec{v}_1 + x_2\vec{v}_2 + \ldots + x_n\vec{v}_n.$$

Theorem: The span of $\vec{v}_1, \ldots, \vec{v}_n$ is a subspace of $\mathbb{R}^m$.

Proof. Let V be the span of $\vec{v}_1, \ldots, \vec{v}_n$. Let $\vec{v}, \vec{w} \in V$. Then

$$\vec{v} = x_1\vec{v}_1 + x_2\vec{v}_2 + \ldots + x_n\vec{v}_n$$

and

$$\vec{w} = y_1\vec{v}_1 + y_2\vec{v}_2 + \ldots + y_n\vec{v}_n$$

for some choices of the x_i and y_i. So

$$\vec{v} + \vec{w} = (x_1 + y_1)\vec{v}_1 + \ldots + (x_n + y_n)\vec{v}_n$$

which is also in V. The check of the other condition is left as an exercise. □

Example: The span of $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$ is the set of all vectors of the form

$$x_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix};$$

as we discussed, this is a plane.

Key example of span: Once again, consider the equation $A\vec{x} = \vec{y}$. Can we solve this equation? Well, observe that the set of all

$$A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

is just the set of all

$$x_1\vec{v}_1 + x_2\vec{v}_2 + \ldots + x_n\vec{v}_n$$

as the x_i range over $\mathbb{R}$; that is, *$A\vec{x} = \vec{y}$ has a solution if and only if $\vec{y}$ is in the span of the columns of A.* (see 2A in Strang).

Definition: The *column space* of A is the span of the columns of A. It is a subspace of $\mathbb{R}^m$.

Definition: The *nullspace* of A is the set of vectors $\vec{x} \in \mathbb{R}^n$ such that $A\vec{x} = 0$. (Exercise: check that the nullspace is a subspace of $\mathbb{R}^n$.–see Strang, p. 68)

II. Linear transformations

Let $f : \mathbb{R}^n \to R^m$ be a function. For instance, a function from $\mathbb{R}^3$ to R^3 might be

$$f\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x^2 \\ y^2 \\ z^2 \end{bmatrix}$$

or a function from $\mathbb{R}$ to $\mathbb{R}^2$ might be

$$f(x) = \begin{bmatrix} x \\ sqrt1 + x^2 \end{bmatrix}$$

In multivariable calculus, we studied the general properties of functions of this type–or at least those which were nice enough to have derivatives. Now we are going to restrict ourselves to a very special class of function, which turns out to be quite important.

Definition: A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is called a *linear transformation* if (and only if)

- $f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y})$
- $f(a\vec{x}) = af(\vec{x})$ for all $a \in f(x)$.

Examples:

0. The functions above are *not* linear transformations. For instance, note that in the first case

$$f\left(\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}\right) = \begin{bmatrix} x_1^2 + x_2^2 \\ y_1^2 + y_2^2 \\ z_1^2 + z_2^2 \end{bmatrix} \neq \begin{bmatrix} x_1^2 \\ y_1^2 \\ z_1^2 \end{bmatrix} + \begin{bmatrix} x_2^2 \\ y_2^2 \\ z_2^2 \end{bmatrix}.$$

1. The identity transformation $i : \mathbb{R}^n \to \mathbb{R}^n$ given by

$$f(\vec{x}) = \vec{x}$$

2. The zero transformation $z : \mathbb{R}^n \to \mathbb{R}^m$ given by

$$z(\vec{x}) = \vec{0}$$

(Possibly skip example 3)

3. The "averaging" transformation $a : \mathbb{R}^3 \to \mathbb{R}^3$ given by

$$a(\begin{bmatrix} x \\ y \\ z \end{bmatrix}) = \begin{bmatrix} (x+y+z)/3 \\ (x+y+z)/3 \\ (x+y+z)/3 \end{bmatrix}$$

Observe: the set of vectors such that $a\vec{x} = \vec{0}$ is $vecx : x + y + z = 0$.

What is the set of vectors $\vec{y}$ such that $a(\vec{x}) = \vec{y}$ has a solution? Maybe take a one-minute contemplation here?

Geometrically speaking: a projects $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ onto the line $t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

4. Given any $m \times n$ matrix A, the function $T_A : \mathbb{R}^n \to \mathbb{R}^m$ defined by

$$T_A(\vec{x}) = A\vec{x}$$

is linear. Proof:

$$T_A(\vec{x} + \vec{y}) = A(\vec{x} + \vec{y}) = A\vec{x} + A\vec{y} = T_A(\vec{x}) + T_A(\vec{y})$$

and

$$T_A(a\vec{x}) = Aa\vec{x} = aA\vec{x} = aT_A(\vec{x})$$

Note that 3 is an example of 4. We can write

$$\begin{bmatrix} (x+y+z)/3 \\ (x+y+z)/3 \\ (x+y+z)/3 \end{bmatrix} = x \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} + y \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} + z \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

So $a = T_A$ where $A = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$.

(11 Feb 2000: I got to this point with only 10 minutes left, so briefly talked about the electrical network example but didn't break into groups.)

Now we're going to do a groupwork. We draw a picture of a 4-node 6-edge network. (Triangular is best). Point out that we can describe the state of such a network (that is, the charge on each vertex) by a vector in $\mathbb{R}^4$. The potential differences (that is, the potential difference on each edge) can be described by a vector in $\mathbb{R}^6$. For example, if the potentials on the nodes are $3, 4, 5, 6$, the potential differences on the edges are $-2, 1, 1, -3, -2, -1$.

GROUPWORK: Linear transformations in an electrical network

Suppose $\vec{x}$ is the vector of potentials. Define a function

$$C : \mathbb{R}^4 \to \mathbb{R}^6$$

by the rule that $C(\vec{x})$ is the vector of potential differences corresponding to the vector of potentials $\vec{x}$.

Example:

$$C\left(\begin{bmatrix} 3 \\ 4 \\ 5 \\ 6 \end{bmatrix}\right) = \begin{bmatrix} -2 \\ 1 \\ 1 \\ -3 \\ -2 \\ -1 \end{bmatrix}$$

- What is $C(\vec{x})$ for an arbitrary vector of potentials $\vec{x} = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}$?
- Show that C is a linear transformation.
- Is there a matrix A such that $C = T_A$? In other words, is there a matrix A such that

  $$A\vec{x} = C(\vec{x})$$

  for all $\vec{x} \in \mathbb{R}^4$?
- For which $\vec{x} \in \mathbb{R}^4$ do we have $C(\vec{x}) = \vec{0}$?
- For which $\vec{y} \in \mathbb{R}^6$ does $C(\vec{x}) = \vec{y}$ have a solution? In other words, which $\vec{y}$ are possible vectors of potential differences?

That should take 20-25 minutes. Regroup, have students present solutions. Let them know that there's a dicussion of this kind of system in Strang, 2.5 (optional reading). Observe that we have not discussed the transpose this week–leave it to Strang.

If there's time, I want to prove the following theorem, which may not be surprising by now.

Theorem. Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation. Then there exists an $m \times n$ matrix A such that

$$T(\vec{x}) = A\vec{x}$$

for all $\vec{x}$ in $\mathbb{R}^m$.

In other words, "all linear transformations come from matrices." And the study of matrices is one and the same as the study of linear transformations.

Proof. For each i, let $\vec{e_i}$ be the matrix with 1 in the i'th place and 0 elsewhere; i. e.

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \end{bmatrix}, \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ vdots \end{bmatrix}, \dots$$

Now given a linear transformation T, define

$$A = \begin{bmatrix} | & | & | \\ T(\vec{e}_1) & \dots & T(\vec{e}_n) \\ | & | & | \end{bmatrix}.$$

We claim this is the desired matrix. Check:

$$\begin{aligned} A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} &= x_1 T(\vec{e}_1) + x_2 T(\vec{e}_2) + \dots + x_n T(\vec{e}_n) \\ &= T(x_1\vec{e}_1) + T(x_2\vec{e}_2) + \dots + T(x_n\vec{e}_n) \\ &= T\left(\begin{bmatrix} x_1 \\ 0 \\ \vdots \end{bmatrix}\right) + \dots + T\left(\begin{bmatrix} \vdots \\ 0 \\ x_n \end{bmatrix}\right) \\ &= T\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right). \end{aligned}$$

We did it! □

Using this fact, we can prove a better version of our invertibility theorem from Monday.

Theorem (Invertibility Theorem II): Suppose A is an $n \times n$ matrix such that $A\vec{x} = \vec{y}$ has a unique solution for all $vecy \in \mathbb{R}^n$. Then A is invertible.

Proof. We need to prove that there is a matrix B such that $AB = BA = I$.

Define a function $U : \mathbb{R}^n \to \mathbb{R}^n$ by the rule:

$U(\vec{y})$ is the unique $\vec{x}$ such that $A\vec{x} = \vec{y}$.

Claim: U is a linear transformation.

To check: $U(\vec{y} + \vec{z}) = U(\vec{y}) + U(\vec{z})$.

But what is $U(\vec{y} + \vec{z})$? It is the unique vector $\vec{w}$ such that $A\vec{w} = \vec{y} + \vec{z}$. But note that

$$A[U(\vec{y}) + U(\vec{z})] = AU(\vec{y}) + AU(\vec{z}) = \vec{y} + \vec{z}.$$

So $U(\vec{y} + \vec{z})$ can be nothing other than $U(\vec{y}) + U(\vec{z})$.

Proof of good behavior under scalar multiplication left to reader.

Now, by the **Theorem** above, there exists an $n \times n$ matrix B such that $U(\vec{y}) = B\vec{y}$ for all $\vec{y} \in \mathbb{R}^n$. Claim: $BA = I$. Proof. Let $\vec{x}$ be some vector in $\mathbb{R}^n$.

$$(BA)\vec{x} = B(A\vec{x}) = B(T(\vec{x})) = U(T(\vec{x}))$$

which is the unique vector $\vec{w}$ such that $T(\vec{w}) = T(\vec{x})$. That is, it is $\vec{x}$! So we've shown $(BA)\vec{x} = \vec{x}$ for all $\vec{x} \in \mathbb{R}^n$; in other words, $BA = I$.

The proof that $AB = I$ is very similar and is left as an exercise. □