

Inverses and Elementary Matrices

Matrix inversion gives a method for solving *some* systems of equations. Suppose we have a system of n linear equations in n variables:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned}$$

Let

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

The system can then be written in matrix form:

$$Ax = b.$$

(One reason for using matrix notation is that it saves writing!) If A has an inverse A^{-1} , I can multiply both sides by A^{-1} :

$$\begin{aligned} A^{-1}Ax &= A^{-1}b \\ I \cdot x &= A^{-1}b \\ x &= A^{-1}b \end{aligned}$$

I've solved for the vector x of variables.

Not every matrix has an inverse — an obvious example is the zero matrix, but here's a nonzero non-invertible matrix over the real numbers:

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

Suppose this matrix had an inverse $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then

$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{so} \quad \begin{bmatrix} a+c & b+d \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

But equating entries in row 2, column 2 gives the contradiction $0 = 1$. Hence, the original matrix does not have an inverse.

If we want to know whether a matrix has an inverse, we could try to do what we did in this example — set up equations and see if we can solve them. But you can see that it could be pretty tedious if the matrix was large or the entries were messy. And we saw earlier that you can solve systems of linear equations using row reduction.

In this section, we'll see how you can use row reduction to determine whether a matrix has an inverse — and, if it does, how to find the inverse. We'll begin by explaining the connection between elementary row operations and matrices.

Definition. An **elementary matrix** is a matrix which represents an elementary row operation. “Represents” means that multiplying on the left by the elementary matrix performs the row operation.

Here are the elementary matrices that represent our three types of row operations. In the pictures below, the elements that are not shown are the same as those in the identity matrix. In particular, all of the

elements that are not on the main diagonal are 0, and all the main diagonal entries — except those shown — are 1.

Multiplying by this matrix swaps rows i and j :

$$\begin{array}{c} \\ \\ \\ i \\ j \end{array} \begin{array}{c} \\ \\ \\ \\ \\ \end{array} \begin{array}{c} i \quad j \\ \left[\begin{array}{cccccc} 1 & 0 & & & & \\ 0 & 1 & & & & \\ & & \ddots & & & \\ & & & 0 & \cdots & 1 \\ & & & \vdots & & \vdots \\ & & & 1 & \cdots & 0 \\ & & & & & \ddots \\ & & & & & & 1 & 0 \\ & & & & & & 0 & 1 \end{array} \right] \end{array}$$

The “ i ” and “ j ” on the borders of the matrix label rows and columns so you can see where the elements are.

This is the same as the identity matrix, except that rows i and j have been swapped. In fact, you obtain this matrix by applying the row operation (“swap rows i and j ”) to the identity matrix. *This is true for our other elementary matrices.*

Multiplying by this matrix multiplies row i by the number c :

$$\begin{array}{c} \\ \\ \\ i \end{array} \begin{array}{c} \\ \\ \\ \\ \end{array} \begin{array}{c} i \\ \left[\begin{array}{cccccc} 1 & 0 & & & & \\ 0 & 1 & & & & \\ & & \ddots & & & \\ & & & c & & \\ & & & & \ddots & \\ & & & & & 1 & 0 \\ & & & & & 0 & 1 \end{array} \right] \end{array}$$

This is the same as the identity matrix, except that row i has been multiplied by c . *Note that this is only a valid operation if the number c has a multiplicative inverse.* For instance, if we’re working over the real numbers, c can be any *nonzero* number.

Multiplying by this matrix replaces row i with row i plus c times row j .

$$\begin{array}{c} \\ \\ \\ i \\ j \end{array} \begin{array}{c} \\ \\ \\ \\ \\ \end{array} \begin{array}{c} i \quad j \\ \left[\begin{array}{cccccc} 1 & 0 & & & & \\ 0 & 1 & & & & \\ & & \ddots & & & \\ & & & 1 & \cdots & c \\ & & & \vdots & & \vdots \\ & & & 0 & \cdots & 1 \\ & & & & & \ddots \\ & & & & & & 1 & 0 \\ & & & & & & 0 & 1 \end{array} \right] \end{array}$$

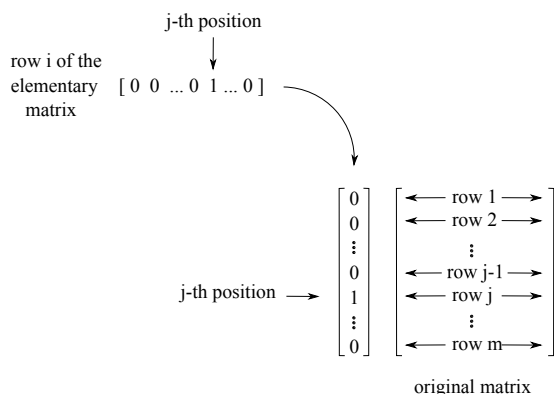
To get this matrix, apply the operation “add c times row j to row i ” to the identity matrix.

While we could give formal proofs that these matrices do what we want — we would have to write formulas for elements of the matrices, then use the definition of matrix multiplication — I don’t think the

proofs would be very enlightening. It's good, however, to visualize the multiplications for yourself to see why these matrices work: Take rows of the elementary matrix and picture them multiplying rows of the original matrix. For example, consider the elementary matrix that swaps row i and row j .

$$\begin{array}{c}
 \begin{array}{cc} & i & j \\ \begin{array}{c} i \\ j \end{array} & \begin{bmatrix} 1 & 0 & & & \\ 0 & 1 & & & \\ & & \ddots & & \\ & & & 0 & \cdots & 1 \\ & & & \vdots & & \vdots \\ & & & 1 & \cdots & 0 \\ & & & & & \ddots \\ & & & & & & 1 & 0 \\ & & & & & & 0 & 1 \end{bmatrix} \end{array}
 \end{array}$$

When you multiply the original matrix by row FOO of this matrix, you get row FOO of the product. So multiplying the original matrix by first row of this matrix gives the first row of the product, and so on. Let's look at what happens when you multiply the original matrix by row i of this matrix.



Row i has 0's everywhere except for a 1 in the j^{th} position. So when it multiplies the original matrix, all the rows of the original matrix get multiplied by 0, except for the j^{th} row, which is multiplied by 1. The net result is the j^{th} row of the original matrix. Thus, the i^{th} row of the product is the j^{th} row of the original matrix.

If you picture this process one row at a time, you'll see that the original matrix is replaced with the same matrix with the i and j rows swapped.

Let's try some examples.

This elementary matrix should swap rows 2 and 3 in a 3×3 matrix:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Notice that it's the identity matrix with rows 2 and 3 swapped.

Multiply a 3×3 matrix by it on the left:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} a & b & c \\ g & h & i \\ d & e & f \end{bmatrix}.$$

Rows 2 and 3 were swapped — it worked!

This elementary matrix should multiply row 2 of a 2×2 matrix by 13:

$$\begin{bmatrix} 1 & 0 \\ 0 & 13 \end{bmatrix}$$

Notice that it's the identity matrix with row 2 multiplied by 13. (We'll assume that we're in a number system where 13 is invertible.)

Multiply a 2×2 matrix by it on the left:

$$\begin{bmatrix} 1 & 0 \\ 0 & 13 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ 13c & 13d \end{bmatrix}.$$

Row 2 of the original matrix was multiplied by 13.

This elementary matrix should add 5 times row 1 to row 3:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 5 & 0 & 1 \end{bmatrix}$$

Notice that it's the identity matrix with 5 times row 1 added to row 3.

Multiply a 3×3 matrix by it on the left:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 5 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g+5a & h+5b & i+5c \end{bmatrix}.$$

You can see that 5 times row 1 was added to row 3.

The *inverses* of these matrices are, not surprisingly, the elementary matrices which represent the *inverse* row operations. The **inverse** of a row operation is the row operation which undoes the original row operation. Let's look at the three operations in turn.

The inverse of swapping rows i and j is swapping rows i and j — to undo swapping two things, you swap the two things back! So the inverse of the “swap row i and j ” elementary matrix is the same matrix:

$$\begin{matrix} & & & i & & j & & \\ & & & & & & & \\ & & & & & & & \\ i & & & & & & & \\ & & & & & & & \\ j & & & & & & & \end{matrix} \begin{bmatrix} 1 & 0 & & & & & \\ 0 & 1 & & & & & \\ & & \ddots & & & & \\ & & & 0 & \cdots & 1 & \\ & & & \vdots & & \vdots & \\ & & & 1 & \cdots & 0 & \\ & & & & & & \ddots & \\ & & & & & & & 1 & 0 \\ & & & & & & & 0 & 1 \end{bmatrix}$$

The inverse of multiplying row i by c is dividing row i by c . To account for the fact that a number system might not have “division”, I'll say “multiplying row i by c^{-1} ”. Just take the original “multiply row i by c ” elementary matrix above and replace c with c^{-1} :

$$\begin{matrix} & & & i & & & & \\ & & & & & & & \\ & & & & & & & \\ i & & & & & & & \\ & & & & & & & \\ & & & & & & & \end{matrix} \begin{bmatrix} 1 & 0 & & & & & \\ 0 & 1 & & & & & \\ & & \ddots & & & & \\ & & & c^{-1} & & & \\ & & & & \ddots & & \\ & & & & & 1 & 0 \\ & & & & & 0 & 1 \end{bmatrix}$$

The inverse of adding c times row j to row i is subtracting c times row j from row i . To write down the inverse, I just replace c with $-c$ in the matrix for “row i goes to row i plus c times row j ”:

$$\begin{matrix} & & & i & & j & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ i & & & & & & & \\ & & & & & & & \\ j & & & & & & & \end{matrix} \begin{bmatrix} 1 & 0 & & & & & & \\ 0 & 1 & & & & & & \\ & & \ddots & & & & & \\ & & & 1 & \cdots & -c & & \\ & & & \vdots & & \vdots & & \\ & & & 0 & \cdots & 1 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 & 0 \\ & & & & & & & 0 & 1 \end{bmatrix}$$

Example. In each case, tell what row operation is performed by multiplying on the left by the elementary matrix. Then find the inverse of the elementary matrix, and tell what row operation is performed by multiplying on the left by the inverse. (All the matrices are real number matrices.)

(a) $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(b) $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 17 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(a) Multiplying on the left by

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{adds 2 times row 3 to row 1.}$$

The inverse

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{subtracts 2 times row 3 from row 1.}$$

(b) Multiplying on the left by

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{swaps row 1 and row 2.}$$

The inverse

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{swaps row 1 and row 1.}$$

(c) Multiplying on the left by

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 17 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{multiplies row 2 by 17.}$$

The inverse

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 17 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{17} & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{divides row 2 by 17.} \quad \square$$

Definition. Matrices A and B are **row equivalent** if A can be transformed to B by a finite sequence of elementary row operations.

Of course, row equivalent matrices must have the same dimensions.

Remark. Since row operations may be performed by multiplying by elementary matrices, A and B are row equivalent if and only if there are elementary matrices E_1, \dots, E_n such that

$$E_1 \cdots E_n A = B.$$

Lemma. Row equivalence is an **equivalence relation**.

Proof. Let's recall what it means (in this situation) to be an **equivalence relation**. I have to show three things:

- (a) (Reflexivity) Every matrix is row equivalent to itself.
 - (b) (Symmetry) If A row reduces to B , then B row reduces to A .
 - (c) (Transitivity) If A row reduces to B and B row reduces to C , then A row reduces to C .
- (a) is obvious, since I can row reduce a matrix to itself by performing the identity row operation.

For (b), suppose A row reduces to B . Then there are elementary matrices E_1, \dots, E_n such that

$$E_1 \cdots E_n A = B.$$

Hence,

$$A = E_n^{-1} \cdots E_1^{-1} B.$$

Since the inverse of an elementary matrix is an elementary matrix, each E_i^{-1} is an elementary matrix. This equation gives a sequence of row operations which row reduces B to A .

To prove (c), suppose A row reduces to B and B row reduces to C . Then there are elementary matrices E_1, \dots, E_m and F_1, \dots, F_n such that

$$E_1 \cdots E_m A = B \quad \text{and} \quad F_1 \cdots F_n B = C.$$

Hence,

$$F_1 \cdots F_n E_1 \cdots E_m A = C.$$

This equation gives a sequence of row operations which row reduces A to C .
Therefore, row equivalence is an equivalence relation. \square

Let's recall the definition of invertibility and the inverse of a matrix.

Definition. An $n \times n$ matrix A is **invertible** if there is an $n \times n$ matrix B such that $AB = BA = I$, where I is the $n \times n$ identity matrix.

In this case, B is the **inverse** of A (or A is the inverse of B), and we write A^{-1} for B (or B^{-1} for A).

We use the usual notation for integer powers of a square matrix.

Notation. If A is a square matrix, then

$$A^n = \begin{cases} \overbrace{A \cdot A \cdots A}^{n \text{ times}} & \text{if } n > 0 \\ I & \text{if } n = 0 \\ \overbrace{A^{-1} \cdot A^{-1} \cdots A^{-1}}^{n \text{ times}} & \text{if } n < 0 \end{cases}.$$

Note that A^n for $n < 0$ only makes sense if A is invertible.

The usual rules for powers hold:

(a) $A^m A^n = A^{m+n}.$

(b) $(A^m)^n = A^{mn}.$

The proofs involve induction and taking cases. I don't think they're that enlightening, so I will skip them.

Example. Consider the real matrix

$$A = \begin{bmatrix} 3 & 1 \\ 2 & 0 \end{bmatrix}.$$

Compute A^2 and A^{-2} .

$$A^2 = A \cdot A = \begin{bmatrix} 3 & 1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 11 & 3 \\ 6 & 2 \end{bmatrix}.$$

Using the formula for the inverse of a 2×2 matrix,

$$A^{-1} = \frac{1}{3 \cdot 0 - 2 \cdot 1} \begin{bmatrix} 0 & -1 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 0.5 \\ 1 & -1.5 \end{bmatrix}.$$

Therefore,

$$A^{-2} = A^{-1} \cdot A^{-1} = \begin{bmatrix} 0 & 0.5 \\ 1 & -1.5 \end{bmatrix} \begin{bmatrix} 0 & 0.5 \\ 1 & -1.5 \end{bmatrix} = \begin{bmatrix} 0.5 & -0.75 \\ -1.5 & 2.75 \end{bmatrix}. \quad \square$$

Remember that matrix multiplication isn't commutative, so AB is not necessarily equal to BA . So what is $(AB)^2$? Since X^2 is shorthand for $X \cdot X$, and in this case $X = AB$, the best we can say is that

$$(AB)^2 = ABAB.$$

Example. Give a specific example of two 2×2 real matrices A and B for which

$$(AB)^2 \neq A^2 B^2.$$

How should I construct this counterexample? On the one hand, I want to avoid matrices which are "too special", because I might accidentally get a case where the equation holds. (The example in this problem just shows that the equation " $(AB)^2 = A^2 B^2$ " isn't *always* true; this doesn't mean that it's *never* true.) For instance, I should not take A to be the identity matrix or the zero matrix (in which case the equation would actually be true).

On the other hand, I'd like to keep the matrices simple — first, so that I don't struggle to do the computation, and second, so that a reader can *easily* see that the computation works. For instance, this would be a bad idea:

$$A = \begin{bmatrix} -1.378 & \pi \\ \frac{\sqrt{2}}{171} & 1013 \end{bmatrix}.$$

When you're trying to construct a counterexample, try to keep these ideas in mind. In the end, you may have to make a few trials to get a suitable counterexample — you will not know whether something works without trying.

I will take

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Following the ideas above, I tried to make matrices which were simple without being too special. Then

$$A^2 = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \quad \text{and} \quad B^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{so} \quad A^2 B^2 = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}.$$

On the other hand,

$$AB = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{so} \quad (AB)^2 = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}.$$

So for these two matrices, $A^2 B^2 \neq (AB)^2$. \square

Proposition.

(a) If A and B are invertible $n \times n$ matrices, then AB is invertible, and its inverse is given by

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

(b) If A is invertible, then A^T is invertible, and its inverse is given by

$$(A^T)^{-1} = (A^{-1})^T.$$

Proof. (a) Remember that AB is not necessarily equal to BA , since matrix multiplication is not necessarily commutative.

I have

$$\begin{aligned} (B^{-1}A^{-1})(AB) &= B^{-1}IB = B^{-1}B = I, \\ (AB)(B^{-1}A^{-1}) &= AIA^{-1} = AA^{-1} = I. \end{aligned}$$

Since $B^{-1}A^{-1}$ gives the identity when multiplied by AB , it means that $B^{-1}A^{-1}$ must be the inverse of AB — that is, $(AB)^{-1} = B^{-1}A^{-1}$.

(b) I have

$$A^T(A^{-1})^T = (A^{-1}A)^T = I^T = I \quad \text{and} \quad (A^{-1})^T A^T = (AA^{-1})^T = I^T = I.$$

Since $(A^{-1})^T$ gives the identity when multiplied by A^T , it means that $(A^{-1})^T$ must be the inverse of A^T — that is, $(A^T)^{-1} = (A^{-1})^T$. \square

Remark. Look over the proofs of the two parts of the last proposition and be sure you understand why the computations *proved* the things that were to be proved. The idea is that the inverse of a matrix is defined by a *property*, not by *appearance*. By analogy, it is like the difference between the set of mathematicians (a set defined by a *property*) and the set of people with purple hair (a set defined by *appearance*).

A matrix C is the inverse of a matrix D if it has the *property* that multiplying C by D (in both orders) gives the identity I . So to check whether a matrix C *really is* the inverse of D , you multiply C by D (in both orders) any see whether you get I . \square

Example. Suppose that A and B are $n \times n$ invertible matrices. Simplify the following expression:

$$(AB)^{-2} A (BA)^2 A.$$

Note: $(AB)^{-2}$ is not $A^{-2}B^{-2}$, nor is it $B^{-2}A^{-2}$.

$$\begin{aligned}
(AB)^{-2}A(BA)^2A &= [(AB)^2]^{-1}A(BA)^2A \\
&= (ABAB)^{-1}A(BABA)A \\
&= B^{-1}A^{-1}B^{-1}A^{-1}ABABAA \quad \square \\
&= A^2
\end{aligned}$$

Example. (Solving a matrix equation) Solve the following matrix equation for X , assuming that A and B are invertible:

$$A^2 X B A = A B.$$

$$\begin{aligned}
A^2 X B A &= A B \\
A^{-2} A^2 X B A &= A^{-2} A B \\
X B A &= A^{-1} B \\
X B A A^{-1} &= A^{-1} B A^{-1} \\
X B &= A^{-1} B A^{-1} \\
X B B^{-1} &= A^{-1} B A^{-1} B^{-1} \\
X &= A^{-1} B A^{-1} B^{-1}
\end{aligned}$$

Notice that I can multiply both sides of a matrix equation by the same thing, but I must multiply *on the same side* of both sides. So when I multiplied by A^{-2} , I had to put A^{-2} on the *left side* of both sides of the equation.

Once again, the reason I have to be careful is that in general, $MN \neq NM$ — matrix multiplication is not commutative. \square

Example. Give a specific example of two invertible 2×2 real matrices A and B for which

$$(A + B)^{-1} \neq A^{-1} + B^{-1}.$$

One way to get a counterexample is to choose A and B so that $A + B$ isn't invertible. For instance,

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Then A and B happen to be equal to their inverses:

$$A^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad B^{-1} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}.$$

But

$$A + B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

The zero matrix is not invertible, because $0 \cdot C = 0$ for any matrix C — so for no matrix C can $0 \cdot C = I$.

But this feels like “cheating”, because the left side $(A + B)^{-1}$ of the equation isn't defined. Okay — can we find two matrices A and B for which $(A + B)^{-1} \neq A^{-1} + B^{-1}$ *and* both sides of the equation are defined? Thus, we need A , B , and $A + B$ to all be invertible.

I'll use

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Then

$$A^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \quad \text{and} \quad B^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

(You can find the inverses using the formula for the inverse of a 2×2 matrix which I gave when I discussed matrix arithmetic. You can also find the inverses using row reduction.)

Thus,

$$A^{-1} + B^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

On the other hand,

$$A + B = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \quad \text{so} \quad (A + B)^{-1} = \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix}.$$

Thus, $(A + B)^{-1} \neq A^{-1} + B^{-1}$ even though both sides of the equation are defined. \square

The next result connects several of the ideas we've looked at: Row reduction, elementary matrices, invertibility, and solving systems of equations.

Theorem. Let A be an $n \times n$ matrix. The following are equivalent:

- (a) A is row equivalent to I .
- (b) A is a product of elementary matrices.
- (c) A is invertible.
- (d) The only solution to the following system is the vector $x = 0$:

$$Ax = 0.$$

- (e) For any n -dimensional vector b , the following system has a unique solution:

$$Ax = b.$$

Proof. When you are trying to prove several statements are *equivalent*, you must prove that if you assume any one of the statements, you can prove any of the others. I can do this here by proving that (a) implies (b), (b) implies (c), (c) implies (d), (d) implies (e), and (e) implies (a).

(a) \Rightarrow (b): Let E_1, \dots, E_p be elementary matrices which row reduce A to I :

$$E_1 \cdots E_p A = I.$$

Then

$$A = E_p^{-1} \cdots E_1^{-1}.$$

Since the inverse of an elementary matrix is an elementary matrix, A is a product of elementary matrices.

(b) \Rightarrow (c): Write A as a product of elementary matrices:

$$A = F_1 \cdots F_q.$$

Now

$$F_1 \cdots F_q \cdot F_q^{-1} \cdots F_1^{-1} = I,$$

$$F_q^{-1} \cdots F_1^{-1} \cdot F_1 \cdots F_q = I.$$

Hence,

$$A^{-1} = F_q^{-1} \cdots F_1^{-1}.$$

(c) \Rightarrow (d): Suppose A is invertible. The system $Ax = 0$ has at least one solution, namely $x = 0$. Moreover, if y is any other solution, then

$$Ay = 0, \quad \text{so} \quad A^{-1}Ay = A^{-1}0, \quad \text{or} \quad y = 0.$$

That is, 0 is the one and only solution to the system.

(d) \Rightarrow (e): Suppose the only solution to $Ax = 0$ is $x = 0$. If $A = (a_{ij})$, this means that row reducing the augmented matrix

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 \end{bmatrix} \quad \text{produces} \quad \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}.$$

Ignoring the last column (which never changes), this means there is a sequence of row operations E_1, \dots, E_n which reduces A to the identity I — that is, A is row equivalent to I . (I've actually proved (d) \Rightarrow (a) at this point.)

Let $b = \langle b_1, \dots, b_n \rangle$ be an arbitrary n -dimensional vector. Then

$$E_1 \cdots E_n \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 & b'_1 \\ 0 & 1 & \cdots & 0 & b'_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & b'_n \end{bmatrix}.$$

Thus, $z = \langle b'_1, \dots, b'_n \rangle$ is a solution.

Suppose y is another solution to $Ax = b$. Then

$$A(y - z) = Ay - Az = b - b = 0.$$

Therefore, $y - z$ is a solution to $Ax = 0$. But the only solution to $Ax = 0$ is 0 , so $y - z = 0$, or $y = z$. Thus, $z = \langle b'_1, \dots, b'_n \rangle$ is the unique solution to $Ax = b$.

(e) \Rightarrow (a): Suppose $Ax = b$ has a unique solution for every b . As a special case, $Ax = 0$ has a unique solution (namely $x = 0$). But arguing as I did in (d) \Rightarrow (e), I can show that A row reduces to I , and that is (a). \square

Remark. We'll see that there are many other conditions that are equivalent to a matrix being invertible. For instance, when we study **determinants**, we'll find that a matrix is invertible if and only if its determinant is invertible as a number.

If A is invertible, the theorem implies that A can be written as a product of elementary matrices. To do this, row reduce A to the identity, keeping track of the row operations you're using. Write each row operation as an elementary matrix, and express the row reduction as a matrix multiplication. Finally, solve the resulting equation for A .

Example. (Writing an invertible matrix as a product of elementary matrices) Express the following real matrix as a product of elementary matrices:

$$A = \begin{bmatrix} 2 & -4 \\ -2 & 3 \end{bmatrix}.$$

First, row reduce A to I :

$$\begin{bmatrix} 2 & -4 \\ -2 & 3 \end{bmatrix} \xrightarrow{r_1 \rightarrow r_1/2} \begin{bmatrix} 1 & -2 \\ -2 & 3 \end{bmatrix} \xrightarrow{r_2 \rightarrow r_2 + 2r_1}$$

$$\begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \xrightarrow{r_1 \rightarrow r_1 + 2r_2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Next, represent each row operation as an elementary matrix:

$$r_1 \rightarrow \frac{1}{2}r_1 \quad \text{corresponds to} \quad \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix},$$

$$r_2 \rightarrow r_2 + 2r_1 \quad \text{corresponds to} \quad \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix},$$

$$r_1 \rightarrow r_1 + 2r_2 \quad \text{corresponds to} \quad \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}.$$

Using the elementary matrices, write the row reduction as a matrix multiplication. A must be multiplied on the left by the elementary matrices *in the order in which the operations were performed*.

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix} \cdot A = I.$$

Solve the last equation for A , being careful to get the inverses in the right order:

$$\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix} \cdot A = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}^{-1} \cdot I,$$

$$A = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}^{-1}.$$

Finally, write each inverse as an elementary matrix.

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}.$$

You can check your answer by multiplying the matrices on the right. \square

For two $n \times n$ matrices A and B to be inverses, I must have

$$AB = I \quad \text{and} \quad BA = I.$$

Since in general AB need not equal BA , it seems as though I must check that *both* equations hold in order to show that A and B are inverses. It turns out that, thanks to the earlier theorem, we only need to check that $AB = I$.

Corollary. If A and B are $n \times n$ matrices and $AB = I$, then $A = B^{-1}$ and $BA = I$.

Proof. Suppose A and B are $n \times n$ matrices and $AB = I$. The system $Bx = 0$ certainly has $x = 0$ as a solution. I'll show it's the only solution.

Suppose y is *another* solution, so

$$By = 0.$$

Multiply both sides by A and simplify:

$$AB y = A \cdot 0$$

$$I y = 0$$

$$y = 0$$

Thus, 0 is a solution, and it's the only solution.

Thus, B satisfies condition (d) of the Theorem. Since the five conditions are equivalent, B also satisfies condition (c), so B is invertible. Let B^{-1} be the inverse of B . Then

$$\begin{aligned} AB &= I \\ ABB^{-1} &= IB^{-1} \\ AI &= B^{-1} \\ A &= B^{-1} \end{aligned}$$

This proves the first part of the Corollary. Finally,

$$BA = BB^{-1} = I.$$

This finishes the proof. \square

The proof of the theorem gives an algorithm for inverting a matrix A .

If A is invertible, there are elementary matrices E_1, \dots, E_p which row reduce A to I :

$$E_1 \cdots E_p A = I.$$

But this equation says that $E_1 \cdots E_p$ is the inverse of A , since multiplying A by $E_1 \cdots E_p$ gives the identity. Thus,

$$A^{-1} = E_1 \cdots E_p = E_1 \cdots E_p \cdot I.$$

We can interpret the last expression $E_1 \cdots E_p \cdot I$ as applying the row operations for E_1, \dots, E_p to the identity matrix I . And the same row operations row reduce A to I . So form an augmented matrix by placing the identity matrix next to A :

$$\left[\begin{array}{ccc|ccc} & & & & & \\ & & & & & \\ \mathbf{A} & & & \mathbf{I} & & \\ & & & & & \\ & & & & & \end{array} \right]$$

augmented matrix

Row reduce the augmented matrix. The left-hand block A will row reduce to the identity; at the same time, the right-hand block I will be transformed into A^{-1} .

Example. Invert the following matrix over \mathbb{R} :

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

Form the augmented matrix by putting the 3×3 identity matrix on the right of the original matrix:

$$\left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ -1 & -1 & 3 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array} \right].$$

Next, row reduce the augmented matrix. The row operations are entirely determined by the block on the left, which is the original matrix. The row operations turn the left block into the identity, while simultaneously turning the identity in the right block into the inverse.

$$\left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ -1 & -1 & 3 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array} \right] \xrightarrow{r_2 \rightarrow r_2 + r_1} \left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array} \right] \xrightarrow{r_3 \rightarrow r_3 - r_2}$$

$$\begin{aligned}
& \left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 1 & 0 \\ 0 & 0 & -1 & -1 & -1 & 1 \end{array} \right] \xrightarrow{r_1 \rightarrow r_1 - 2r_2} \left[\begin{array}{ccc|ccc} 1 & 0 & -5 & -1 & -2 & 0 \\ 0 & 1 & 2 & 1 & 1 & 0 \\ 0 & 0 & -1 & -1 & -1 & 1 \end{array} \right] \xrightarrow{r_3 \rightarrow -r_3} \\
& \left[\begin{array}{ccc|ccc} 1 & 0 & -5 & -1 & -2 & 0 \\ 0 & 1 & 2 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & -1 \end{array} \right] \xrightarrow{r_1 \rightarrow r_1 + 5r_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 4 & 3 & -5 \\ 0 & 1 & 2 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & -1 \end{array} \right] \xrightarrow{r_2 \rightarrow r_2 - 2r_3} \\
& \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 4 & 3 & -5 \\ 0 & 1 & 0 & -1 & -1 & 2 \\ 0 & 0 & 1 & 1 & 1 & -1 \end{array} \right].
\end{aligned}$$

Thus,

$$\begin{bmatrix} 1 & 2 & -1 \\ -1 & -1 & 3 \\ 0 & 1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 4 & 3 & -5 \\ -1 & -1 & 2 \\ 1 & 1 & -1 \end{bmatrix}. \quad \square$$

In the future, I won't draw the vertical bar between the two blocks; you can draw it if it helps you keep the computation organized.

Example. (Inverting a matrix over \mathbb{Z}_p) Find the inverse of the following matrix over \mathbb{Z}_3 :

$$\begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix}.$$

Form the augmented matrix and row reduce:

$$\begin{aligned}
& \left[\begin{array}{cccccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 & 1 \end{array} \right] \xrightarrow{r_2 \rightarrow r_2 - r_1} \left[\begin{array}{cccccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 2 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 & 1 \end{array} \right] \xrightarrow{r_3 \rightarrow r_3 - 2r_1} \\
& \left[\begin{array}{cccccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 2 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{array} \right] \xrightarrow{r_3 \rightarrow r_3 - r_2} \left[\begin{array}{cccccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 2 & 1 \end{array} \right] \xrightarrow{r_1 \rightarrow r_1 - 2r_3} \\
& \left[\begin{array}{cccccc} 1 & 0 & 0 & 0 & 2 & 1 \\ 0 & 1 & 2 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 2 & 1 \end{array} \right] \xrightarrow{r_2 \rightarrow r_2 - 2r_3} \left[\begin{array}{cccccc} 1 & 0 & 0 & 0 & 2 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 & 2 & 1 \end{array} \right]
\end{aligned}$$

Therefore,

$$\begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 2 & 1 \\ 1 & 0 & 1 \\ 2 & 2 & 1 \end{bmatrix}. \quad \square$$

The theorem also tells us about the number of solutions to a system of linear equations.

Proposition. Let F be a field, and let $Ax = b$ be a system of linear equations over F . Then:

(a) If F is infinite, then the system has either no solutions, exactly one solution, or infinitely many solutions.

(b) If F is a finite field with p^n elements, where p is prime and $n \geq 1$, then the system has either no solutions, exactly one solution, or at least p^n solutions.

Proof. Suppose the system has more than one solution. I must show that there are infinitely many solutions if F is infinite, or at least p^n solutions if F is a finite field with p^n elements.

Since there is more than one solution, there are at least two different solutions. So let x_1 and x_2 be two different solutions to $Ax = b$:

$$Ax_1 = b \quad \text{and} \quad Ax_2 = b.$$

Subtracting the equations gives

$$A(x_1 - x_2) = Ax_1 - Ax_2 = b - b = 0.$$

Since $x_1 - x_2 \neq 0$, $x_1 - x_2$ is a nontrivial solution to the system $Ax = 0$. Now if $t \in F$,

$$A(x_1 + t(x_1 - x_2)) = Ax_1 + t \cdot A(x_1 - x_2) = b + 0 = b.$$

Thus, $x_1 + t(x_1 - x_2)$ is a solution to $Ax = b$. Moreover, the only way two solutions of the form $x_1 + t(x_1 - x_2)$ can be the same is if they have the same t . For

$$x_1 + t(x_1 - x_2) = x_1 + t'(x_1 - x_2) \quad \text{gives} \quad t(x_1 - x_2) = t'(x_1 - x_2).$$

Since $x_1 \neq x_2$, I have $x_1 - x_2 \neq 0$. So I can divide both sides by $x_1 - x_2$, and I get $t = t'$.

Thus, different t 's give different $x_1 + t(x_1 - x_2)$'s, each of which is a solution to the system.

If F has infinitely many elements, there are infinitely many possibilities for t , so there are infinitely many solutions.

If F has p^n elements, there are p^n possibilities for t , so there are at least p^n solutions. (Note that there may be solutions which are not of the form $x_1 + t(x_1 - x_2)$, so there may be more than p^n solutions. In fact, I'll be able to show later that the number of solutions will be some power of p^n .) \square

For example, since \mathbb{R} is an infinite field, a system of linear equations over \mathbb{R} has no solutions, exactly one solution, or infinitely many solutions.

On the other hand, since \mathbb{Z}_3 is a field with 3 elements, a system of linear equations over \mathbb{Z}_3 has no solutions, exactly one solution, or at least 3 solutions. (And I'll see later that if there's more than one solution, then there might be 3 solutions, 9 solutions, 27 solutions, ...)