# Inverses and Elementary Matrices

Matrix inversion gives a method for solving *some* systems of equations. Suppose

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

is a system of n linear equations in n variables. Write

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

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 $A^{-1}$ :

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

I've solved for the vectors x of unknowns.

Since not every matrix has an inverse, it's important to know:

- When a matrix has an inverse.
- How to find the inverse, if there is one.

I'll discuss these questions in this section.

**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.)

In the pictures below, the elements that are not shown are the same as those in the identity matrix.

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

interchanges rows i and j.

$$\begin{array}{c} \\ i \\ \\ \end{array}\!\!\begin{array}{c} \\ \vdots \\ \\ \end{array}\!\!\begin{bmatrix} \cdots & a & \cdots \end{bmatrix}$$

multiplies row i by a.

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

replaces row i with row i plus a times row j.

Their inverses are the elementary matrices

$$\begin{array}{c} \\ i \\ \\ j \\ \\ \end{array}\!\!\begin{array}{c} \\ \vdots \\ \\ \vdots \\ \\ \end{array}\!\!\begin{bmatrix} \cdots & 0 & \cdots & 1 & \cdots \\ & \vdots & & \vdots & \\ \cdots & 1 & \cdots & 0 & \cdots \\ & \vdots & & \vdots & \end{bmatrix}, \qquad \begin{array}{c} \\ i \\ \\ \end{array}\!\!\begin{array}{c} \\ \vdots \\ \\ \end{array}\!\!\begin{bmatrix} \cdots & \frac{1}{a} & \cdots \end{bmatrix}, \qquad \begin{array}{c} \\ i \\ \\ j \\ \\ \end{array}\!\!\begin{array}{c} \\ \vdots \\ \\ \vdots \\ \\ \end{array}\!\!\begin{bmatrix} \cdots & 1 & \cdots & -a & \cdots \\ & \vdots & & \vdots & \\ \cdots & 0 & \cdots & 1 & \cdots \\ & \vdots & & \vdots & \end{bmatrix},$$

respectively. Multiplication by the first matrix swaps rows i and j. Multiplication by the second matrix divides row i by a. Multiplication by the third matrix subtracts a times row j from row i. These operations are the inverses of the operations implemented by the original matrices.

---

**Example.** 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

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

Multiplying on the left by

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{swaps row 2 and row 3.}$$

The inverse

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{swaps row 2 and row 3.}$$

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.

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

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

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

**Proof.** I have to show three things:

1. (Reflexivity) Every matrix is row equivalent to itself.
1. (Symmetry) If A row reduces to B, then B row reduces to A.
1. (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. Write

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

where $E_1$ , ... $E_n$ are elementary matrices. Then

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

Since the inverse of an elementary matrix is an elementary matrix, it follows that B row reduces to A.

It remains to prove (c). If A row reduces to B and B row reduces to C, then there are elementary matrices $E_1$ , ..., $E_m$ , $F_1$ , ..., $F_n$ such that

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

Then

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

so A row reduces to C.

Therefore, row equivalence is an equivalence relation. $\square$

**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.

**Notation.** If A is a square matrix, then

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

(where $A^n$ for $n < 0$ only makes sense if A is invertible.

The usual rules for powers hold:

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

---

**Example.** Let

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

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 $2 \times 2$ 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,

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

---

**Proposition.**

1. If A and B are invertible $n \times n$ matrices, then

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

1. If A is invertible, then $(A^T)^{-1} = (A^{-1})^T$ .

**Proof.** (a) The inverse of $AB$ is the *thing* which, when multiplied by $AB$ , gives the identity I. Now

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

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

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

(b) The inverse of $A^T$ is the *thing* which, when multiplied by $A^T$ , give the identity I. Now

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

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

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

---

**Example.** Solve the following matrix equation for X, assuming that A and B are invertible:

$$A^2XBA = AB.$$

$$\begin{aligned} A^2XBA &= AB \\ A^{-2}A^2XBA &= A^{-2}AB \\ XBA &= A^{-1}B \\ XBAA^{-1} &= A^{-1}BA^{-1} \\ XB &= A^{-1}BA^{-1} \\ XBB^{-1} &= A^{-1}BA^{-1}B^{-1} \\ X &= A^{-1}BA^{-1}B^{-1} \end{aligned}$$

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. The reason I have to be careful is that in general, $MN \neq NM$ --- matrix multiplication is not commutative. □

---

**Example.** Note that $(A+B)^{-1} \neq A^{-1} + B^{-1}$. In fact, if A and B are invertible, $A+B$ need not be invertible. For example, if

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

then A and B are invertible --- each is its own inverse.

But

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

which is not invertible. □

---

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

1. A is row equivalent to I.
1. A is a product of elementary matrices.
1. A is invertible.

1. The system

$$Ax = 0$$

\item{} has only the trivial solution $x = 0$ .

1. For any n-dimensional vector b, the system

$$Ax = b$$

\item{} has a unique solution.

**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 $E_1, \ldots, E_p$ 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

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

Ignoring the last column (which never changes), this means there is a sequence of row operations $E_1$ , ..., $E_n$ 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 $b = \langle b_1, \ldots b_n \rangle$ be an arbitrary n-dimensional vector. Then

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

Thus, $z = \langle b'_1, \ldots b'_n \rangle$ 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, $z = \langle b'_1, \ldots b'_n \rangle$ 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). □

---

**Example.** ( **Writing an invertible matrix as a product of elementary matrices**) 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.

For example, suppose

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

Row reduce A to I:

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

$$\begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \underset{r_1 \to r_1 + 2r_2}{\to} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Represent each row operation as an elementary matrix:

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

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

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

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*.

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

Now solve 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.

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

---

**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:

$$\begin{aligned} ABy &= A \cdot 0 \\ Iy &= 0 \\ y &= 0 \end{aligned}$$

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

**Remark.** This result shows that if you're checking that two square matrices A and B are inverses by multiplying to get the identity, you only need to check $AB = I$ --- $BA = I$ is then automatic. □

**Algorithm.** The proof provides an algorithm for inverting a matrix A.

If $E_1, \ldots, E_p$ are elementary matrices which row reduce A to I, then

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

Then

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

That is, the row operations which reduce A to the identity also transform the identity 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**

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

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 on the right into the inverse.

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

$$\left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 1 & 0 \\ 0 & 0 & -1 & -1 & -1 & 1 \end{array}\right] \xrightarrow[r_1 \to r_1 - 2r_2]{} \left[\begin{array}{ccc|ccc} 1 & 0 & -5 & -1 & -2 & 0 \\ 0 & 1 & 2 & 1 & 1 & 0 \\ 0 & 0 & -1 & -1 & -1 & 1 \end{array}\right] \xrightarrow[r_3 \to -r_3]{}$$

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

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 4 & 3 & -5 \\ 0 & 1 & 0 & -1 & -1 & 2 \\ 0 & 0 & 1 & 1 & 1 & -1 \end{array}\right].$$

Thus,

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

---

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

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

$$\begin{bmatrix} 1 & 0 & 2 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \underset{r_2 \to r_2 - r_1}{\to} \begin{bmatrix} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 2 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \underset{r_3 \to r_3 - 2r_1}{\to}$$

$$\begin{bmatrix} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 2 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \underset{r_3 \to r_3 - r_2}{\to} \begin{bmatrix} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 2 & 1 \end{bmatrix} \underset{r_1 \to r_1 - 2r_3}{\to}$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 2 & 1 \\ 0 & 1 & 2 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 2 & 1 \end{bmatrix} \underset{r_2 \to r_2 - 2r_3}{\to} \begin{bmatrix} 1 & 0 & 0 & 0 & 2 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 & 2 & 1 \end{bmatrix}$$

Therefore,

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

---

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

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

1. 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.

Suppose then that there is more than one solution. Let $x_1$ and $x_2$ be distinct solutions to $Ax = b$ , so

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

Note that

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

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

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

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

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

Now $t - t' \neq 0$ implies $x_1 = x_2$ , a contradiction. Therefore, $t - t' = 0$ , so $t = t'$ .

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 than the number of solutions will be some power of $p^n$.) □

---

**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.

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, ....) □

---

Send comments about this page to: Bruce.Ikenaga@millersville.edu.

Bruce Ikenaga's Home Page