Chapter 2: Solving Linear Equations

2.3. Elimination Using Matrices

As we saw in the presentation, we can use "elimination" to make a system of linear equations into an "upper triangular system" that is easy to solve, and then we can use "back-substitution" to solve it. But as we saw, there are cases where there is no solution or infinitely many of them. First we introduce some terms to help simplify and describe linear systems in terms of matrices and vectors, and then we use them to investigate when elimination fails. The basic concepts are so simple that we can name them in the course of an example:

Example. The linear system

$$\begin{array}{ccccccc} x & + & 2y & + & 3z & = & 8 \\ x & + & 3y & & & = & 7 \\ x & & & + & 2z & = & 3 \end{array}$$

has

$$\textit{coefficient matrix} \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 0 \\ 1 & 0 & 2 \end{bmatrix} \qquad \text{and } \textit{augmented matrix} \begin{bmatrix} 1 & 2 & 3 & 8 \\ 1 & 3 & 0 & 7 \\ 1 & 0 & 2 & 3 \end{bmatrix}.$$

(A matrix like the augmented matrix may have a marker of some sort between columns to help the eye keep things separate; e.g., the augmented matrix may have a marker between the variables and the constant terms on the other sides of the equals sign:

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 8 \\ 1 & 3 & 0 & 7 \\ 1 & 0 & 2 & 3 \end{array}\right].)$$

The three operations on the equations in the linear system amount in the context of matrices to *"elementary row operations"*:

- subtract from one row a (scalar) multiple of another row,
- reverse two rows, or
- multiply a row by a nonzero scalar.

These operations change the original augmented matrix into *"row equivalent"* ones that correspond to systems with the same solutions. For example, we might subtract from the second row 1 times the first row, and subtract from the third row 1 times the first row, giving the augmented matrix

$$\begin{bmatrix} 1 & 2 & 3 & 8 \\ 0 & 1 & -3 & -1 \\ 0 & -2 & -1 & -5 \end{bmatrix}.$$

Then we could subtract from the third row -2 times the second row (i.e., add to the third row 2 times the second row), giving the augmented matrix

$$\begin{bmatrix} 1 & 2 & 3 & 8 \\ 0 & 1 & -3 & -1 \\ 0 & 0 & -7 & -7 \end{bmatrix}.$$

The coefficient matrix is now upper triangular, so we can turn the augmented matrix back into a system and use back substitution to find the solution: $-7z = -7$, so $z = 1$; then $y - 3z = -1$, so $y = -1 + 3(1) = 2$; and $x + 2y + 3z = 8$, so $x = 8 - 2(2) - 3(1) = 1$.

2.4. Rules for Matrix Operations

We want to interpret the example in a different way, and for this we need to introduce *"matrix multiplication."* If matrix A is, say, $m \times n$ (i.e., has m rows and n columns) and B is $n \times p$, then the product matrix AB is the $m \times p$ matrix in which the entry in the i-th row and j-th column is the dot product of i-th row of A with the j-th column of B:

Example:
$$\begin{bmatrix} 1 & -2 & 3 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 4 & -2 \\ 2 & -1 \end{bmatrix}$$
$$= \begin{bmatrix} 1(1) - 2(4) + 3(2) & 1(3) - 2(-2) + 3(-1) \\ 2(1) + 0(4) + 1(-2) & 2(3) + 0(-2) + 1(-1) \end{bmatrix}$$
$$= \begin{bmatrix} -1 & 4 \\ 0 & 5 \end{bmatrix} .$$

We will collect some other properties of matrix multiplication later, but for right now, note that in most cases, $AB \neq BA$: Of course, if the sizes don't match, one of these may make sense while the other does not. But even if they do match, the *commutative law* doesn't hold:

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

Fortunately, though, the *associative law* does hold: For all matrices A, B, C of the proper dimensions,

$$A(BC) = (AB)C .$$

For each positive integer n, there is a square $n \times n$ matrix I (or I_n) with the property that 1 has among numbers; i.e., $IA = A$ and $AI = A$ for any matrix A (of the proper size). That matrix has 1's down the "main diagonal" — upper left to lower right — and 0's everywhere else:

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

Though we don't need it at this point, the text, and therefore we, note that there is another operation on matrices — provided they have the same dimensions: addition. It is as simple as it can be: add the corresponding entries:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} + \begin{bmatrix} -1 & 2 & 2 \\ 3 & -3 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 4 & 5 \\ 7 & 2 & 5 \end{bmatrix} .$$

It is clear that matrix addition is *commutative* (unlike multiplication), and *associative:* For all A, B, C of the proper dimensions,

$$A + B = B + A \qquad\qquad A + (B + C) = (A + B) + C .$$

And you can check that matrix products *distribute* over matrix sums (from both sides):

$$A(B + C) = AB + AC \qquad\qquad (B + C)A = BA + CA .$$

Example. Question: Which matrices $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ "commute with" $E = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$? Such an A satisfies $AE = EA$, i.e.,

$$\begin{bmatrix} a+2b & b \\ c+2d & d \end{bmatrix} = \begin{bmatrix} a & b \\ c+2a & d+2b \end{bmatrix},$$

so $a + 2b = a$, $b = b$, $c + 2d = c + 2a$ and $d = d + 2b$. From these equations, we can see that $b = 0$ and $d = a$; but there are no restrictions on c, and a only has to agree with d. So the A that commute with this E have the form $\begin{bmatrix} a & 0 \\ c & a \end{bmatrix}$ for any a and c.

With these operations, we see that we can do *matrix algebra* with matrices as objects. In some cases we may have to worry about whether the matrices have the right dimensions, but if we restrict to square matrices of the same size, we can add and multiply at will.

2.5. Inverse Matrices

If A is square, then there may be another matrix A^{-1} with the property that $A^{-1}A = I$. If so, that A^{-1} is called the *inverse* of A, and we also have $AA^{-1} = I$. (If A is not square, say 2×3, then it is possible for A to have a "one-sided inverse", but the other is impossible: we may have $AB = I_2$ for some B, but we cannot get $BA = I_3$, or indeed $CA = I_3$ for any C. Why not? Well, such a C must only have 2 columns, and those columns can only span a plane in 3-space, and the columns of CA must lie in that plane — but the three columns of I_3 don't all lie in the same plane.)

Fact (and definition): If we do an elementary row operation (say ϕ) to an identity matrix, the resulting matrix E, an *"elementary matrix,"* has the property that, for any matrix A (of the proper number of rows), the product EA is the result of doing ϕ to A.

So we can think of the steps of elimination as multiplications on the left by elementary matrices:

Example. (Continued):

$$\begin{bmatrix} 1 & 2 & 3 & 8 \\ 1 & 3 & 0 & 7 \\ 1 & 0 & 2 & 3 \end{bmatrix} \qquad [A : \boldsymbol{b}]$$

$$\xrightarrow{\boxed{2}-\boxed{1}} \begin{bmatrix} 1 & 2 & 3 & 8 \\ 0 & 1 & -3 & -1 \\ 1 & 0 & 2 & 3 \end{bmatrix} \qquad [E_{21}A : E_{21}\boldsymbol{b}] \text{ where } E_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{\boxed{3}-\boxed{1}} \begin{bmatrix} 1 & 2 & 3 & 8 \\ 0 & 1 & -3 & -1 \\ 0 & -2 & -1 & -5 \end{bmatrix} \qquad [E_{31}E_{21}A : E_{31}E_{21}\boldsymbol{b}] \text{ where } E_{31} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{\boxed{3}+2\boxed{2}} \begin{bmatrix} 1 & 2 & 3 & 8 \\ 0 & 1 & -3 & -1 \\ 0 & 0 & -7 & -7 \end{bmatrix} \qquad [E_{32}E_{31}E_{21}A : E_{32}E_{31}E_{21}\boldsymbol{b}] \text{ where } E_{21} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}$$

As we did above, from the last augmented matrix, we can read off the solution: We get $-7z = -7$, so $z = 1$; and from $y - 3z = -1$ we get $y = -1 + 3(1) = 2$; and finally from $x + 2y + 3z = 8$ we get $x = 8 - 2(2) - 3(1) = 1$.

Example. Solving

$$\begin{array}{rcrcrcr} x & - & y & + & z & = & 0 \\ 2x & + & 2y & + & z & = & 1 \\ 3x & + & 2y & + & 2z & = & 2 \end{array} \quad :$$

$$\begin{bmatrix} 1 & -1 & 1 & 0 \\ 2 & -2 & 1 & -3 \\ 3 & 2 & 2 & 2 \end{bmatrix} \qquad [A : \boldsymbol{b}]$$

$$\xrightarrow{\boxed{2}-2\boxed{1}} \begin{bmatrix} 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 3 & 2 & 2 & 2 \end{bmatrix} \qquad [E_{21}A : E_{21}\boldsymbol{b}] \text{ where } E_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{\boxed{3}-3\boxed{1}} \begin{bmatrix} 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & -3 \\ 0 & 5 & -1 & 2 \end{bmatrix} \qquad [E_{31}E_{21}A : E_{31}E_{21}\boldsymbol{b}] \text{ where } E_{31} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{\boxed{3}\leftrightarrow\boxed{2}} \begin{bmatrix} 1 & -1 & 1 & 0 \\ 0 & 5 & -1 & 2 \\ 0 & 0 & -1 & -3 \end{bmatrix} \qquad [P_{23}E_{31}E_{21}A : P_{23}E_{31}E_{21}\boldsymbol{b}] \text{ where } P_{23} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The elementary matrix P_{23} is an example of a *permutation matrix*, with one 1 in each row and column and the rest 0's; it is called that because premultiplication (i.e., multiplication on the left) by one of these rearranges (i.e., "permutes") the rows.

Anyway, the coefficient matrix is now upper triangular, so we can read off the solution: $-z = -3$ gives $z = 3$; then $5y - z = 2$ gives $y = (2+3)/5 = 1$; and $x - y + z = 0$ gives $x = 1 - 3 = -2$.

Let's note one more fact about inverse matrices: For invertible matrices A, B of the same size,

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

(It works with any number of factors: $(ABCD)^{-1} = D^{-1}C^{-1}B^{-1}A^{-1}$, as you can easily check by breaking the product into products of two matrices.) Why? Well, multiply $A^{-1}B^{-1}$ by AB on either side, and you realize that A^{-1} isn't next to A, nor is B^{-1} next to B; so without the commutative law, it's not clear that the product is the identity. But if we reverse the order, it comes out right:

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

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

So is there a good way to decide whether a given square matrix has an inverse, and to find that inverse? Before we show how we can answer the question(s) in general, let's do the 2×2 case, which has a formula that we can remember. (This is a special case of the "classical adjoint" result to which Ahmad has already referred, but I don't want to introduce that yet.) We have already talked about the *"determinant"* of a 2×2 matrix, a number associated with the matrix:

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc \ .$$

To get the inverse of this matrix, we reverse the main diagonal entries, negate the off-diagonal entries, and divide by the determinant:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \ .$$

We can easily check that this works:

$$\frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \frac{1}{ad-bc}\begin{bmatrix} da-bc & db-bd \\ -ca+ac & -cb+ad \end{bmatrix}$$
$$= \frac{1}{ad-bc}\begin{bmatrix} ad-bc & 0 \\ 0 & ad-bc \end{bmatrix} = I \ .$$

What if the determinant is 0? We have seen that this is exactly when the second row in this matrix is a multiple of the first, i.e., this matrix is singular, and has no inverse.

Example.

$$\begin{bmatrix} 1 & -2 \\ 3 & 2 \end{bmatrix}^{-1} = \frac{1}{8}\begin{bmatrix} 2 & 2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} \\ -\frac{3}{8} & \frac{1}{8} \end{bmatrix} \ .$$

Example. (from the text): Verify that if $M = A - UW^{-1}V$, then

$$M^{-1} = A^{-1} + A^{-1}U(W - VA^{-1}U)^{-1}VA^{-1} \ .$$

Because the "if" refers to W^{-1}, we can assume that W is invertible, but it doesn't seem entirely fair to assume that A and $(W - VA^{-1}U$ are invertible. But we will pretend that they are, and go on from there. We must show that at least one of the products of M with the claimed M^{-1} is the identity, with M either on the right or the left. So let's try MM^{-1}. For the sake of sanity, I'm going to set

$$B = (W - VA^{-1}U)^{-1}$$

Okay, let's see if it works:

$$\begin{aligned}
&(A - UW^{-1}V)(A^{-1} + A^{-1}U(W - VA^{-1}U)^{-1}VA^{-1}) && \\
&\quad = (A - UW^{-1}V)(A^{-1} + A^{-1}UBVA^{-1}) && \text{incorporating } B \\
&\quad = (A - UW^{-1}V)A^{-1} + (A - UW^{-1}V)A^{-1}UBVA^{-1} && \text{one distributive law} \\
&\quad = AA^{-1} - UW^{-1}VA^{-1} && \\
&\qquad + AA^{-1}UBVA^{-1} - UW^{-1}VA^{-1}UBVA^{-1} && \text{the other distributive law} \\
&\quad = I - UW^{-1}VA^{-1} + UBVA^{-1} - UW^{-1}VA^{-1}UBVA^{-1} && AA^{-1} = I \\
&\quad = I - U[W^{-1} - B + W^{-1}VA^{-1}UB]VA^{-1} && \text{two distributive laws, in reverse} \\
&\quad = I - U[W^{-1} - \{I + W^{-1}VA^{-1}U\}B]VA^{-1} && \text{distributive law in reverse} \\
&\quad = I - U[W^{-1} - \{W^{-1}W - W^{-1}VA^{-1}U\}B]VA^{-1} && W^{-1}W = I \\
&\quad = I - U[W^{-1} - W^{-1}\{W - VA^{-1}U\}B]VA^{-1} && \text{distributive law} \\
&\quad = I - U[W^{-1} - W^{-1}I]VA^{-1} && \text{definition of } B \\
&\quad = I - U[W^{-1} - W^{-1}]VA^{-1} && W^{-1}I = W^{-1} \\
&\quad = I - UOVA^{-1} = I && O \text{ is matrix of 0's}
\end{aligned}$$

Aha! It came out right.

Matrices, especially square ones, can be divided into rectangular pieces, and then we can treat the pieces as if they were numbers, just remembering that multiplication isn't commutative. As long as the pieces on the main diagonal are square, the rest of the sizes seem to work out for multiplication and addition as needed. The results are called *"block matrices."*

Example. Suppose

$$M = \begin{bmatrix} I_3 & A \\ O & B \end{bmatrix} \ ,$$

where I_3 is the $3{\times}3$ identity matrix, B is $2{\times}2$ invertible, O is a matrix of 0's and A is anything that fits. Is M invertible?

First comment: To "fit", the A must be a $3{\times}2$ matrix, and O must be a $2{\times}3$ matrix of 0's. Now if the entries in M were numbers instead of matrices, we could reverse the entries on the main diagonal, take the negatives of the other two entries, and divide by the determinant B. But the determinant $I_3B - AC$ is meaningless — the dimensions are wrong in the first product — and we can't divide by matrices anyway. It does, though, give a block matrix $\begin{bmatrix} I_3 & -A/B \\ O & B^{-1} \end{bmatrix}$ that we can try out; we only have to figure out how to interpret A/B — is it $B^{-1}A$ or AB^{-1}? The first of these is meaningless, because the dimensions are wrong, so let's try the second:

$$\begin{bmatrix} I_3 & A \\ O & B \end{bmatrix} \begin{bmatrix} I_3 & -AB^{-1} \\ O & B^{-1} \end{bmatrix} = \begin{bmatrix} I_3 + AO & -AB^{-1} + AB^{-1} \\ O & BB^{-1} \end{bmatrix} = \begin{bmatrix} I_3 & \hat{O} \\ O & I_2 \end{bmatrix} = I_5$$

where $\hat{O}$ means the $3{\times}2$ matrix of 0's — notice that AO is the $3{\times}3$ matrix of 0's. At any rate, we see that M is invertible.