1 Solving Linear Systems

In the first chapter we show how to find the solution set of a system of linear equations.

1.1 Systems of Linear Equations

*A **linear equation** is that the sum of coefficients times variables is some value; for example, $3x - y = 7$. A **linear system** is a collection of linear equations. A **solution** of the system satisfies all the equations. A system is **consistent** if it has at least one solution. The **solution set** is all solutions.*

Consider, for example, a linear equation in the two variables x and y. This represents a line in the plane. If there are two such equations, then the corresponding lines can either intersect, be the same line, or be parallel and different. In the first case the system has a unique solution, in the second case any point on the line is a solution, and in the third case they have no point in common. Here are some examples:

EXAMPLE.

$2x + 3y = 7$	$2x + 3y = 7$	$2x + 3y = 7$
$x - y = 1$	$2x + 3y = 11$	$4x + 6y = 14$
Solution: $x = 2$, $y = 1$	*No Solution*	*Infinite solutions*

We will show that the above behavior captures all possibilities:

Fact 1.1 *There are exactly three possibilities for the solution set of a linear system: no solution, unique solution, or infinitely many solutions.*

A Word About Proofs. Every Fact in this course has a proof. We sometimes give the proof, sometimes sketch the highlights, and sometimes just skip it. Mathematics rests on proof. Proof provides a guarantee that the Fact is true. Proofs use logic, calculation, previous facts, and definitions.

1.2 Triangular Systems

A special case of a linear system is what is called a triangular system.

> *A linear system is **triangular** if the first equation has only one variable, the second equation has only the first variable and another, and so on.*

Triangular linear systems can be solved by an algorithm/process known as **back-substitution**:

> ALGOR BACK SUBSTITUTION.
>
> ≻ *Solve the first equation for its variable.*
>
> ≻ *Substitute the result into the second equation, and solve for its remaining variable.*
>
> ≻ *Repeat.*

EXAMPLE. *Consider the system*

$$\begin{aligned} 2x_1 \qquad\qquad &= 6 \\ x_1 + x_2 \qquad &= 2 \\ -x_1 + 4x_2 + x_3 &= 19 \end{aligned}$$

The first equation implies that $x_1 = 3$. *Substituting this into the second equation implies* $x_2 = -1$. *Substituting both these values into the third equation implies* $x_3 = 26$. *The solution is unique.*

1.3 Matrices

To both represent and solve linear systems, it is convenient to use the notation and ideas of matrices. We will see in the rest of the course that matrices have many many other uses.

> *A **matrix** is a rectangular arrangement of numbers. Its **size** is the number of rows and columns. We use the terminology "$m \times n$ matrix" to mean a matrix with m rows and n columns. (Always rows first.)*

A matrix can represent the left-hand-side of a linear system by recording the coefficients of each variable, one row for each equation. The **augmented** matrix is formed by adding the constants as the last column.

EXAMPLE. *The linear system given in the above example has the augmented matrix*

$$\left[\begin{array}{ccc|c} 2 & 0 & 0 & 6 \\ 1 & 1 & 0 & 2 \\ -1 & 4 & 1 & 19 \end{array}\right]$$

The augmented matrix is a 3×4 *matrix.*

1.4 Elementary Row Operations

Our approach to solving a linear system uses what are called elementary row operations. We apply these to bring the linear system to a simpler-looking system (such as one that is triangular) without changing the solution set.

The three **elementary row operations** we use are: replacement, interchange, and scaling:

- **Replacement**: One can replace a row by the sum of it and a multiple of another row. For example, replace the second row by the sum of it and 3 times the first row. This is often abbreviated to "add 3 times the first row to the second" or "$R2' = R2 + 3R1$" (where the prime means the new value).
- **Interchange**: One can interchange two rows. For example, swap the first and third rows.
- **Scaling**: One can scale a row by a nonzero factor. For example, multiply all entries in the second row by 5.

The crucial point is that these operations do not change the solution set. Hopefully this is obvious for the interchange and the scaling operations. For the replacement operation, note that any solution of the original system remains a solution to the new system. Further, each row operation is reversible. So by the same logic, any solution of the new system remains a solution if we revert to the original system. That is, any solution of the system after the replacement is a solution to the original system. We state this as a fact:

Fact 1.2 *Any row operation preserves the solution set.*

Two matrices are defined to be **row equivalent** if it is possible to get from one matrix to the other by elementary row operations. Note that, since the row operations are reversible, if matrix A is row equivalent to matrix B, then matrix B is also row equivalent to matrix A.

1.5 Echelon Forms

The overall approach to solving a linear system is to perform a series of row operations on the augmented matrix until the solution set is easy to obtain. There are two phases:

- The first phase brings it to a triangular system, or rather a generalization of that;
- The second phase solves for all the variables, or at least as many as one can.

The goal of the first phase of row operations is called echelon form; the goal of the second phase is called reduced row echelon form. These forms are defined by the unaugmented columns.

An unaugmented matrix is defined to be in ***echelon form*** *if:*

(A) for each row that is not completely zero, the leftmost nonzero entry has zeros below it and below its preceding zeroes; and

(B) the all-zero rows (if any) are at the bottom.

Echelon form is a generalization of triangular system, except that we list the equations in decreasing number of variables.

EXAMPLE. *An example of echelon form is:*

$$\begin{bmatrix} -7 & 5 & 4\frac{1}{3} & -10 \\ 0 & 0 & 1 & \frac{1}{5} \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

An unaugmented matrix in echelon form is defined to be in ***reduced row echelon form*** *if furthermore*

(C) the first nonzero entry in every row is a 1 and has zeroes above it.

Example. *An example of reduced row echelon form is:*

$$\begin{bmatrix} 0 & 1 & 5^\star & 0 & 0 & -1^\star \\ 0 & 0 & 0 & 1 & 0 & -\frac{1}{2}^\star \\ 0 & 0 & 0 & 0 & 1 & 43^\star \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

It remains in such form if any of the starred values are changed arbitrarily (even to 0*).*

It can be shown that every matrix is row-equivalent to a *unique* matrix in reduced row echelon form.

1.6 Row Reduction

We now present the **row reduction algorithm**. It is often called **Gaussian Elimination** or **Gauss-Jordan Elimination**.

A ***pivot*** *is the leftmost nonzero entry in a row.*

The first phase of the algorithm proceeds:

ALGOR Achieving Echelon Form.

≻ *Go to first nonzero column (but not the augmented column). If needed, interchange rows so top entry in that column is not zero.*

≻ *Add suitable multiples of the first row to create zeroes below pivot.*

≻ *Repeat, ignoring rows with pivots, until there is no nonzero row.*

Example. *For example, consider the matrix*

$$\left[\begin{array}{ccc|c} -1 & 2 & -1 & 0 \\ 0 & 2 & -8 & 8 \\ 6 & -2 & -4 & 10 \end{array}\right]$$

The first nonzero column is the first column. The top entry is nonzero; so no interchange needed. The -1 *is the pivot. The second entry in the column is* 0*. So no work needed there. The third entry is* 6*. The key point is that if one adds* 6 *times the first row to the third row, in the third row the first entry will become zero. That is, one* *gets*

$$\left[\begin{array}{ccc|c} \textcircled{-1} & 2 & -1 & 0 \\ 0 & 2 & -8 & 8 \\ 6 & -2 & -4 & 10 \end{array}\right] R_3' = R_3 + 6R_1 \quad \rightsquigarrow \quad \left[\begin{array}{ccc|c} -1 & 2 & -1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & 10 & -10 & 10 \end{array}\right]$$

Next we work on the second column. We pretend the first row is not there. The column has a nonzero where we need it in the second row: so 2 is the pivot. We just have to get a zero in the third row. This time the row replacement entails subtracting 5 times the second row from the third row. Thus we get to echelon form:

$$\left[\begin{array}{ccc|c} -1 & 2 & -1 & 0 \\ 0 & \textcircled{2} & -8 & 8 \\ 0 & 10 & -10 & 10 \end{array}\right] R_3' = R_3 - 5R_2 \quad \rightsquigarrow \quad \left[\begin{array}{ccc|c} -1 & 2 & -1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & 0 & 30 & -30 \end{array}\right]$$

We now present the second phase of the **row reduction algorithm**. The algorithm proceeds:

ALGOR ACHIEVING REDUCED ROW ECHELON FORM.

From right to left:

- ≻ *add suitable multiples of the rows to create zeroes above each pivot; and*
- ≻ *make each pivot 1 by scaling.*

(When doing this process by hand, it is sometimes easier to scale first and then zero-out the column, and sometimes it is easier to zero-out the column with the old row. It just depends...)

EXAMPLE. *We earlier reached an echelon form of*

$$\left[\begin{array}{ccc|c} -1 & 2 & -1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & 0 & 30 & -30 \end{array}\right]$$

We start with the third column. The reduction continues with scaling the third row by dividing by 30. This is followed by adding 8 times the new third row to the second row, and adding 1 times the new third row to the first row. Thus we get

$$\left[\begin{array}{ccc|c} -1 & 2 & -1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & 0 & \textcircled{30} & -30 \end{array}\right] \begin{array}{l} R_1' = R_1 + R_3' \\ R_2' = R_2 + 8R_3' \\ R_3' = R_3/30 \end{array} \quad \rightsquigarrow \quad \left[\begin{array}{ccc|c} -1 & 2 & 0 & -1 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{array}\right]$$

There after we do the same for the second column

$$\left[\begin{array}{ccc|c} -1 & 2 & 0 & -1 \\ 0 & \textcircled{2} & 0 & 0 \\ 0 & 0 & 1 & -1 \end{array}\right] \begin{array}{l} R_1' = R_1 - 2R_2' \\ R_2' = R_2/2 \\ \end{array} \quad \rightsquigarrow \quad \left[\begin{array}{ccc|c} -1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{array}\right]$$

and finally, only scaling is needed in the first row to produce reduced row echelon form.

$$\left[\begin{array}{ccc|c} \textcircled{-1} & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{array}\right] \begin{array}{l} R_1' = -R_1 \\ \\ \end{array} \quad \rightsquigarrow \quad \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{array}\right]$$

This system has the unique solution $x_1 = 1$, $x_2 = 0$, *and* $x_3 = -1$.

1.7 The Solution Set of a Linear Systems

We saw earlier that some linear systems have no solution, some a unique solution, and some infinitely many solutions. The next fact gives the rule to determine whether a linear system has a solution or not, that is, whether it is consistent or not:

Fact 1.3 *A linear system is consistent if and only if the echelon form of the augmented matrix has no row that is all zeroes except for a nonzero in the augmented column.*

We skip the full justification of this fact. We note at least that such a row creates a problem; for example, requiring that $0x + 0y = -5$ is impossible.

EXAMPLE. *In the following echelon forms, the first two correspond to systems that are consistent but the third does not.*

$$\left[\begin{array}{ccc|c} 1 & 0 & 2 & 13 \\ 0 & 1 & -5 & 0 \end{array}\right] \qquad \left[\begin{array}{ccc|c} 1 & 0 & 2 & 13 \\ 0 & 1 & -5 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \qquad \left[\begin{array}{ccc|c} 1 & 0 & 2 & 13 \\ 0 & 1 & -5 & 0 \\ 0 & 0 & 0 & 5 \end{array}\right]$$

Consider a matrix that represents a linear system. If the system is consistent, we can read off the solution set from the reduced row echelon form. We define a **basic** variable as one whose column has a pivot; otherwise it is a **free** variable. If every column has a pivot (equivalently every variable is basic), then the solution is unique. But what happens in general?

ALGOR *The solution set is obtained by expressing each basic variable in terms of the free variables; this is called a* ***parametric description****.*

With the parametric description, we can think of the free variables as "parameters": each setting of the free variables yields one solution of the original linear system.

EXAMPLE. *Consider the following row reduced echelon form.*

$$\left[\begin{array}{cccc|c} 1 & 2 & 0 & -3 & -1 \\ 0 & 0 & 1 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

The system is consistent. (There is no bad row.) Assume the variables are x_1, x_2, x_3, x_4*. Then the free variables are* x_2 *and* x_4*. So the system has solution set with parametric description:*

$$\begin{aligned} x_1 &= -1 - 2x_2 + 3x_4 \\ x_3 &= 5 - x_4 \end{aligned}$$

Note that this algorithm proves that Fact 1.1, which we stated on the first page, is true. If a consistent system has a free variable, then there are infinitely many solutions.

Practice

1.1. Solve the following triangular system.

$$\begin{aligned} x - y + z &= 7 \\ 2x - y &= 11 \\ 3y &= 42 \end{aligned}$$

1.2. Consider this matrix

$$B = \begin{bmatrix} 0 & 1 & -3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Is this matrix B in echelon form? In reduced row echelon form? Explain.

1.3. Solve the following systems.

$$\begin{aligned} a + 2b + 3c &= 2 \\ 2a \quad\quad - c &= 8 \\ -3a + 7b + 2c &= -19 \end{aligned} \qquad \begin{aligned} a + 2b + 3c &= 2 \\ 2a \quad\quad - c &= 8 \\ 4a + 4b + 5c &= 11 \end{aligned} \qquad \begin{aligned} a + 2b + 3c &= 2 \\ 2a \quad\quad - c &= 8 \\ 4a + 4b + 5c &= 12 \end{aligned}$$

Solutions to Practice Exercieses

1.1. In order we find $y = 14$, $x = 25/2$, $z = 17/2$

1.2. Yes to both. The first nonzero entry in each row is a 1, and it has zeroes above, below, and lower-left of it.

1.3. $\begin{bmatrix} \textcircled{1} & 2 & 3 & 2 \\ 2 & 0 & -1 & 8 \\ -3 & 7 & 2 & -19 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 2 \\ 0 & \textcircled{-4} & -7 & 4 \\ 0 & 13 & 11 & -13 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 2 \\ 0 & -4 & -7 & 4 \\ 0 & 0 & \textcircled{-\frac{47}{4}} & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & \textcircled{-4} & 0 & 4 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

Solution is $a = 4$, $b = -1$, $c = 0$

$\begin{bmatrix} \textcircled{1} & 2 & 3 & 2 \\ 2 & 0 & -1 & 8 \\ 4 & 4 & 5 & 11 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 2 \\ 0 & \textcircled{-4} & -7 & 4 \\ 0 & -4 & -7 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 2 \\ 0 & -4 & -7 & 4 \\ 0 & 0 & 0 & -1 \end{bmatrix}$

Inconsistent

$\begin{bmatrix} \textcircled{1} & 2 & 3 & 2 \\ 2 & 0 & -1 & 8 \\ 4 & 4 & 5 & 12 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 2 \\ 0 & \textcircled{-4} & -7 & 4 \\ 0 & -4 & -7 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 2 \\ 0 & \textcircled{-4} & -7 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & -\frac{1}{2} & 4 \\ 0 & 1 & \frac{7}{4} & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

Parametric description: $a = 4 + c/2$, $b = -1 - 7c/4$, c free