Linear Algebra 1: Lecture 1

Irena Penev

Winter 2022/2023

1 Systems of linear equations

A *linear equation* in the variables $x_1, \dots, x_m$ is an equation that can be written in the form

$$a_1x_1 + \dots + a_mx_m = b,$$

where b and the coefficients $a_1, \dots, a_n$ are elements of some "field." We have not studied fields yet, but here are some examples of fields (these are the fields that we will use in our examples):

1. the field of real numbers $\mathbb{R}$;
2. the field of complex numbers $\mathbb{C}$;
3. the field $\mathbb{Z}_p$, where p is a prime number.

Fields always come equipped with two operations: addition and multiplication. Importantly, the numbers $a_1, \dots, a_m, b$ must all come from the same field.

For example, the equation

$$x_1 - 3(x_2 - x_1) = 7x_3 - 4,$$

with coefficients understood to be in $\mathbb{R}$, is a linear equation because it can be algebraically rearranged to have the following form:

$$4x_1 - 3x_2 - 7x_3 = -4,$$

which is obviously a linear equation.

On the other hand, equations

$$x_1^3 + x_2 = 17 \quad \text{and} \quad x_1 - \sqrt{x_2} = 5$$

are not linear because of x_1^3 and $\sqrt{x_2}$.

A *system of linear equations*, or a *linear system*, is a collection of one of more linear equations involving the same variables, say $x_1, \dots, x_m$ (and with coefficients coming from the same field). For example, the following is a linear system (here, the coefficients are assumed to be in $\mathbb{R}$):[1]

$$\begin{array}{rcrcrcrcr} 2x_1 & + & 7x_2 & & & - & \pi x_4 & = & -\sqrt{3} \\ & & -3x_2 & + & 17x_3 & - & 3x_4 & = & 2 \\ x_1 & + & x_2 & - & 2x_3 & + & 7x_4 & = & \frac{11}{2} \end{array}$$

A *solution* of a linear system in variables $x_1, \ldots, x_m$ is a list $s_1, \ldots, s_m$ of numbers (from the same field as the coefficients of the system) such that make each equation becomes a true statement when $s_1, \ldots, s_m$ are substituted for $x_1, \ldots, x_m$, respectively.

Example 1.1. *Consider the linear system*

$$\begin{array}{rcrcrcr} x_1 & + & 2x_2 & - & x_3 & = & 9 \\ & & 2x_2 & + & 3x_3 & = & 16 \\ x_1 & + & x_2 & - & x_3 & = & 4 \end{array}$$

with coefficients in $\mathbb{R}$*. Then*

$$\begin{array}{rcl} x_1 & = & 1 \\ x_2 & = & 5 \\ x_3 & = & 2 \end{array}$$

is a solution of the system above.

Example 1.2. *Consider the linear system*

$$\begin{array}{rcrcr} x_1 & + & x_2 & = & 0 \\ 2x_1 & + & x_2 & = & 1 \end{array}$$

with coefficients in $\mathbb{Z}_3$*. Then*

$$\begin{array}{rcl} x_1 & = & 1 \\ x_2 & = & 2 \end{array}$$

is a solution of the system above.

The *set of solutions* or *solution set* of a linear system is the set of all solutions of that system. Our goal is to describe a procedure for finding the solution set of any linear system.

A linear system may have no solutions, one solution, or more than one solution. A system that has at least one solution is called *consistent*; a system that has no solutions is *inconsistent*.

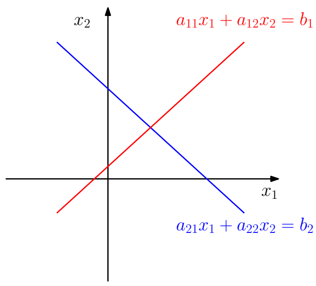
[1] Typographically, we normally arrange equations in our system so that the terms involving the same variable are in the same column.

Consider, for instance, a system of two linear equations in two variables, with coefficients in $\mathbb{R}$:

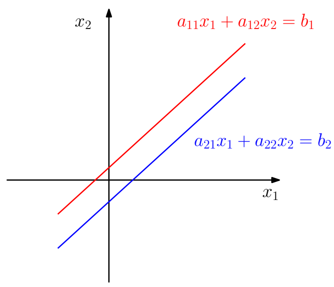
$$\begin{array}{rcrcl} a_{11}x_1 & + & a_{12}x_1 & = & b_1 \\ a_{21}x_1 & + & a_{22}x_1 & = & b_2 \end{array}$$

Each of the two equations defines a line in the plane. There are three possibilities for these two lines:

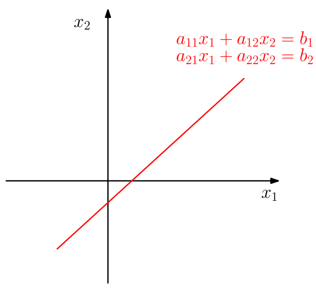
1. the two lines may intersect in one point (in this case, the system has a unique solution, and in particular, it is consistent);



2. the two lines may be distinct, parallel lines (in this case, the system has no solutions, i.e. it is inconsistent);



3. the two lines may be identical (in this case, the system has infinitely many solutions, and in particular, the system is consistent).[2]



[2]Note that the two lines may be identical even if the two equations are different. For instance, $x_1 + x_2 = 1$ and $2x_1 + 2x_2 = 2$ define the same line.

On the other hand, suppose we have a system of two linear equations in three variables (with coefficients in $\mathbb{R}$).

$$\begin{array}{ccccccc} a_{11}x_1 & + & a_{12}x_1 & + & a_{13}x_3 & = & b_1 \\ a_{21}x_1 & + & a_{22}x_1 & + & a_{23}x_3 & = & b_2 \end{array}$$

Each of those equations defines a plane in the three-dimensional Euclidean space. Those two planes may intersect in line (in which case the system has infinitely many solutions, and in particular, the system is consistent); or the two planes may be distinct and parallel (in which case, the system has no solutions, i.e. it is inconsistent); or the two planes may be identical (in which case the system has infinitely many solutions, and in particular, the system is consistent).

2 Matrices

A *matrix* is a rectangular array of numbers. An $n \times m$ *matrix* (read "n by m matrix") is a matrix with n rows and m columns. Consider, for example, the following matrices.

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 3 & 4 \end{bmatrix} \quad B = \begin{bmatrix} 3 & 1 \\ 2 & 5 \\ 1 & 0 \end{bmatrix} \quad C = \begin{bmatrix} 3 & 3 & 2 \\ 1 & -1 & -5 \\ -2 & 2 & 3 \end{bmatrix}$$

A is a 2×3 matrix, B is a 3×2 matrix, and C is a 3×3 matrix. A *square matrix* is one that has the same number of rows and columns. So, C is a square matrix, but A and B are not square matrices.

2.1 The augmented matrix and the coefficient matrix of a linear system

The information from a linear system can be recorded in a matrix called the *augmented matrix* as follows. Suppose we are given a system of n linear equations in m variables, as follows.

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1m}x_m & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2m}x_m & = & b_2 \\ & & & & & & & \vdots & \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \dots & + & a_{nm}x_m & = & b_n \end{array}$$

This system can be fully "encoded" in an $n \times (m+1)$ matrix, called its *augmented matrix*, as follows:

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1m} & b_1 \\ a_{21} & a_{22} & \dots & a_{2m} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} & b_n \end{array}\right]$$

The dotted vertical line is optional, but visually, it helps separate the coefficients to the left of the equality sign from the numbers to the right of the equality sign.

Example 2.1. *Find the augmented matrix of following linear system (with coefficients in* $\mathbb{R}$*):*

$$\begin{array}{rcrcrcr} 3x_1 & + & 2x_2 & + & 5x_3 & = & 7 \\ & & 3x_2 & - & x_3 & = & 0 \end{array}$$

Solution.

$$\left[\begin{array}{rrr:r} 3 & 2 & 5 & 7 \\ 0 & 3 & -1 & 0 \end{array}\right]$$

□

Example 2.2. *Find the augmented matrix of the following linear system (with coefficients in* $\mathbb{Z}_3$*):*

$$\begin{array}{rcl} 2x_1 + x_3 + 2 & = & x_2 \\ x_2 + x_3 & = & 2x_1 \end{array}$$

Solution. We first algebraically rearrange the system above to get it into standard form:[3]

$$\begin{array}{rcrcrcr} 2x_1 & + & 2x_2 & + & x_3 & = & 1 \\ x_1 & + & x_2 & + & x_3 & = & 0 \end{array}$$

We can now read the augmented matrix of the system:

$$\left[\begin{array}{rrr:r} 2 & 2 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{array}\right]$$

□

3 Elementary row operations

Two linear systems (with the same variables) are *equivalent* if they gave exactly the same solution set. Now, suppose we are given a system of linear

[3]Remember: We are working in $\mathbb{Z}_3$! We manipulated the first equation as follows.

1. First, we added $2x_2$ to both sides of the equation to obtain $2x_1 + 2x_2 + x_3 + 2 = 0$.
   - We are using the fact that $x_2 + 2x_2 = (1 + 2)x_2 = 0x_2 = 0$.
2. Then, we added 1 to both sides of the equation to obtain $2x_1 + 2x_2 + x_3 = 1$.
   - We are using the fact that $2 + 1 = 0$.

We manipulated the second equation by adding x_1 to both sides to obtain $x_1 + x_2 + x_3 = 0$. (We used the fact that $2x_1 + x_2 = (2 + 1)x_2 = 0x_2 = 0$.)

equations.

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1m}x_m & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2m}x_m & = & b_2 \\ & & & & & & & \vdots & \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \dots & + & a_{nm}x_m & = & b_n \end{array}$$

We would like to manipulate this system in a way that allows us to "read off" the solution set of the system. There are three basic ways that we can manipulate the system in a way that does not change the solution set (i.e. in a way that produces an equivalent linear system). These are the following.

1. Swap two equations.

   - For example, we can swap the first and third equation in the system on the left to obtain the one on the right.

$$\begin{array}{rcrcrcr} x_1 & + & 3x_2 & - & 2x_3 & = & -1 \\ \frac{1}{2}x_1 & & & + & 2x_3 & = & 0 \\ x_1 & + & x_2 & + & 2x_3 & = & 2 \end{array} \longrightarrow \begin{array}{rcrcrcr} x_1 & + & x_2 & + & 2x_3 & = & 2 \\ \frac{1}{2}x_1 & & & + & 2x_3 & = & 0 \\ x_1 & + & 3x_2 & - & 2x_3 & = & -1 \end{array}$$

2. Multiply one equation by a **non-zero** scalar.[4]

   - For example, we can multiply the second equation by 2:

$$\begin{array}{rcrcrcr} x_1 & + & x_2 & + & 2x_3 & = & 2 \\ \frac{1}{2}x_1 & & & + & 2x_3 & = & 0 \\ x_1 & + & 3x_2 & - & 2x_3 & = & -1 \end{array} \longrightarrow \begin{array}{rcrcrcr} x_1 & + & x_2 & + & 2x_3 & = & 2 \\ x_1 & & & + & 4x_3 & = & 0 \\ x_1 & + & 3x_2 & - & 2x_3 & = & -1 \end{array}$$

3. Add a scalar multiple of some equation to another equation.[5]

   - For example, we can add (-1) times the second equation to the third equation:

$$\begin{array}{rcrcrcr} x_1 & + & x_2 & + & 2x_3 & = & 2 \\ x_1 & & & + & 4x_3 & = & 0 \\ x_1 & + & 3x_2 & - & 2x_3 & = & -1 \end{array} \longrightarrow \begin{array}{rcrcrcr} x_1 & + & x_2 & + & 2x_3 & = & 2 \\ x_1 & & & + & 4x_3 & = & 0 \\ & & 3x_2 & - & 6x_3 & = & -1 \end{array}$$

Instead of manipulating systems of linear equations in this way, we can manipulate augmented matrices. There are three types of "elementary row operations," as follows.

1. Swap two rows.

[4] The scalar is supposed to be from the same field as the coefficients of the linear system.
[5] The scalar is supposed to be from the same field as the coefficients of the linear system.

- For example, we can swap the first and third row in the matrix on the left to obtain the matrix on the right.

$$\left[\begin{array}{ccc|c} 1 & 3 & -2 & -1 \\ \frac{1}{2} & 0 & 2 & 0 \\ 1 & 1 & 2 & 2 \end{array}\right] \overset{R_1 \leftrightarrow R_3}{\sim} \left[\begin{array}{ccc|c} 1 & 1 & 2 & 2 \\ \frac{1}{2} & 0 & 2 & 0 \\ 1 & 3 & -2 & -1 \end{array}\right]$$

2. Multiply one row by a **non-zero** scalar.[6]

   - For instance, we can multiply the second row by 2.

$$\left[\begin{array}{ccc|c} 1 & 1 & 2 & 2 \\ \frac{1}{2} & 0 & 2 & 0 \\ 1 & 3 & -2 & -1 \end{array}\right] \overset{R_2 \to 2R_2}{\sim} \left[\begin{array}{ccc|c} 1 & 1 & 2 & 2 \\ 1 & 0 & 4 & 0 \\ 1 & 3 & -2 & -1 \end{array}\right]$$

3. Add a scalar multiple of some row to another row.[7]

   - For example, we can add (-1) times the second row to the third row:

$$\left[\begin{array}{ccc|c} 1 & 1 & 2 & 2 \\ 1 & 0 & 4 & 0 \\ 1 & 3 & -2 & -1 \end{array}\right] \overset{R_3 \to R_3 - R_2}{\sim} \left[\begin{array}{ccc|c} 1 & 1 & 2 & 2 \\ 1 & 0 & 4 & 0 \\ 0 & 3 & -6 & -1 \end{array}\right]$$

If one matrix can be obtained from another via some (possibly null) sequence of elementary row operations, then the two matrices are *row equivalent*. If matrices A and B are row equivalent, then we write $A \sim B$. It is easy to see that row equivalence is an equivalence class (on the set of matrices of the same size, with entries from the same field), that is, that the following hold:

- $A \sim A$
- if $A \sim B$, then $B \sim A$
- if $A \sim B$ and $B \sim C$, then $A \sim C$.

Clearly, if two matrices are row equivalent, then they encode equivalent linear systems.

4 Row reduction

A *zero row* of a matrix is a row in which all entries are zero. A *non-zero row* of a matrix is a row that has at least one non-zero entry. The *leading entry* of a non-zero row of a matrix is the left-most non-zero entry of that row.

A matrix is in *row echelon form* (or simply *echelon form*), abbreviated *REF*, if it satisfies the following two conditions:

[6]The scalar is supposed to be from the same field as the entries of the matrix.
[7]The scalar is supposed to be from the same field as the entries of the matrix.

1. all non-zero rows are above any zero rows;
2. each leading entry of a row is in a column strictly to the right of the leading entry of the row above it.[8]

If, in addition, the matrix satisfies the following two conditions, then it is in *reduced row echelon form* (or simply *reduced echelon form*), abbreviated RREF:

3. the leading entry in each non-zero row is 1;
4. each leading 1 is the only non-zero entry in its column.

Schematically, a matrix in row echelon form looks like this (here, $\blacksquare$ represents a non-zero number, and $*$ represents any number):

$$\begin{bmatrix} 0 & \blacksquare & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Schematically, a matrix in **reduced** echelon form looks like this ($*$ represents any number):

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

Theorem 4.1. *Every matrix (with entries in some field $\mathbb{F}$) is row equivalent to a unique matrix in reduced row echelon form.*

Proof. Omitted. □

Notation: For a matrix A, we denote by $RREF(A)$ the unique matrix in reduced row echelon form that is row equivalent to A. As we shall see, if a matrix is in reduced row echelon form, then it is particularly easy to "read off" the solution set of the linear system that it encodes (as an augmented matrix).

We begin with a definition. A *pivot column* of a matrix in reduced echelon form is any column that contains a leading 1 of some row; a *pivot position* of a matrix in reduced row echelon form is the position of a leading 1. The

[8] So, all entries in a column below a leading entry are zeros.

matrix below is in reduced echelon form. Its pivot positions are boxed. The pivot columns (four of them) are the ones with the boxed entries.

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

The *pivot columns* of any matrix A are the columns that correspond to the pivot columns of $RREF(A)$. Similarly, the pivot positions of A are the positions of A that correspond to the pivot positions of $RREF(A)$. For example, if A is any matrix that is row equivalent to the matrix shown right above, then the pivot columns of A are the second, fourth, fifth, and eighth column of A. The pivot positions of such a matrix A are the ones that correspond to the boxed 1's above.

Remark: While each matrix A is row equivalent to exactly one matrix in **reduced** row echelon form, A may be row equivalent to many (possibly, infinitely many) matrices in row echelon form. However, all such matrices (i.e. all matrices in row echelon form that A is row equivalent to) have the same "shape," in the sense that, when the matrix is represented with symbols $\blacksquare$, $*$, and 0 (as explaned in the definition of a matrix in row echelon form), we get the same picture.

4.1 The row reduction algorithm

We now describe an algorithm that transforms any matrix into a row-equivalent matrix that is in row echelon form. The algorithm has two parts: the "forward phase" and the "backward phase." The forward phase transforms the matrix into one in row echelon form. The backward phase transforms a matrix in row echelon form into one in **reduced** echelon form. The forward phase of the row reduction algorithm is also called "Gaussian elimination." The entire row reduction algorithm (with both the forward and the backward phase) is also called the "Gauss-Jordan elimination." The algorithm is as follows.

Forward phase:

1. Begin with the leftmost nonzero column. This is a pivot column. The pivot position is at the top.

2. Select a nonzero entry in the pivot column as a pivot. If necessary, interchange rows to move this entry into the pivot position.

3. Use row addition operations to create zeros in all positions below the pivot.

4. Cover (or ignore) the row containing the pivot position and cover all rows, if any, above it. Apply steps 1-3 to the submatrix that remains. Repeat the process until there are no more nonzero rows to modify.

Backward phase:

5. Beginning with the rightmost pivot and working upward and to the left, create zeros above each pivot. If a pivot is not 1, make it 1 by a scaling operation.

A couple of implementations (examples) of the row reduction algorithm are shown below. In each case, we first apply the forward phase, and then the backward phase. In the forward phase, we use a horizontal dotted line as a visual aid: it separates the rows that have already been processed (those are the ones above the dotted line) from the ones that have not yet been processed (those are the ones below the dotted line). Numbers in red designate the pivot column we have identified (as per step 1 or step 5) and are currently processing.

Example 4.2. *Apply the row reduction algorithm to the matrix A below (with entries understood to be in $\mathbb{R}$) in order to transform it into a matrix in reduced row echelon form.*

$$A = \begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix}$$

Solution. We first apply the forward phase of the algorithm in order to transform the matrix into one in row echelon form, as follows.

$$\begin{aligned} A &= \begin{bmatrix} \color{red}{0} & 3 & -6 & 6 & 4 & -5 \\ \color{red}{3} & -7 & 8 & -5 & 8 & 9 \\ \color{red}{3} & -9 & 12 & -9 & 6 & 15 \end{bmatrix} \\ &\overset{R_1 \leftrightarrow R_3}{\sim} \begin{bmatrix} \color{red}{3} & -9 & 12 & -9 & 6 & 15 \\ \color{red}{3} & -7 & 8 & -5 & 8 & 9 \\ \color{red}{0} & 3 & -6 & 6 & 4 & -5 \end{bmatrix} \\ &\overset{R_2 \to R_2 - R_1}{\sim} \begin{bmatrix} 3 & \color{red}{-9} & 12 & -9 & 6 & 15 \\ \hdashline 0 & \color{red}{2} & -4 & 4 & 2 & -6 \\ 0 & \color{red}{3} & -6 & 6 & 4 & -5 \end{bmatrix} \\ &\overset{R_3 \to R_3 - \frac{3}{2} R_2}{\sim} \begin{bmatrix} 3 & -9 & 12 & -9 & \color{red}{6} & 15 \\ 0 & 2 & -4 & 4 & \color{red}{2} & -6 \\ \hdashline 0 & 0 & 0 & 0 & \color{red}{1} & 4 \end{bmatrix} \end{aligned}$$

The forward part of our row reduction algorithm is now complete: our matrix is in row echelon form. It remains to complete the backward part, in order to transform the matrix into one in **reduced** row echelon form.

$$
\begin{array}{cccl}
A & \sim & \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} & \text{by the above (forward part)} \\
 & \overset{R_1 \to R_1 - 6R_3}{\underset{R_2 \to R_2 - 2R_3}{\sim}} & \begin{bmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 2 & -4 & 4 & 0 & -14 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} & \\
 & \overset{R_2 \to \frac{1}{2}R_2}{\sim} & \begin{bmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} & \\
 & \overset{R_1 \to R_1 + 9R_2}{\sim} & \begin{bmatrix} 3 & 0 & -6 & 9 & 0 & -72 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} & \\
 & \overset{R_1 \to \frac{1}{3}R_1}{\sim} & \begin{bmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} &
\end{array}
$$

The backward part of row reduction is now complete: our matrix is in reduced row echelon form. Thus,

$$
RREF(A) = \begin{bmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}.
$$

□

Example 4.3. *Apply the row reduction algorithm to the matrix B below (with entries understood to be in $\mathbb{Z}_3$) in order to transform it into a matrix in reduced row echelon form.*

$$
B = \begin{bmatrix} 0 & 1 & 1 & 0 & 2 \\ 2 & 1 & 0 & 1 & 1 \\ 2 & 1 & 1 & 1 & 1 \\ 1 & 0 & 2 & 2 & 1 \end{bmatrix}
$$

Solution. We first apply the forward phase of the algorithm in order to transform the matrix into one in row echelon form, as follows.

$$
\begin{aligned}
B &= \begin{bmatrix} 0 & 1 & 1 & 0 & 2 \\ 2 & 1 & 0 & 1 & 1 \\ 2 & 1 & 1 & 1 & 1 \\ 1 & 0 & 2 & 2 & 1 \end{bmatrix} \\
&\overset{R_1 \leftrightarrow R_4}{\sim} \begin{bmatrix} 1 & 0 & 2 & 2 & 1 \\ 2 & 1 & 0 & 1 & 1 \\ 2 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 2 \end{bmatrix} \\
&\overset{\substack{R_2 \to R_2 + R_1 \\ R_3 \to R_3 + R_1}}{\sim} \begin{bmatrix} 1 & 0 & 2 & 2 & 1 \\ \hdashline 0 & 1 & 2 & 0 & 2 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 1 & 1 & 0 & 2 \end{bmatrix} \\
&\overset{\substack{R_3 \to R_3 + 2R_2 \\ R_4 \to R_4 + 2R_2}}{\sim} \begin{bmatrix} 1 & 0 & 2 & 2 & 1 \\ 0 & 1 & 2 & 0 & 2 \\ \hdashline 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \end{bmatrix} \\
&\overset{R_4 \to R_4 + R_3}{\sim} \begin{bmatrix} 1 & 0 & 2 & 2 & 1 \\ 0 & 1 & 2 & 0 & 2 \\ 0 & 0 & 1 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 \end{bmatrix}
\end{aligned}
$$

The forward part of our row reduction algorithm is now complete: our matrix is in row echelon form. It remains to complete the backward part, in order to transform the matrix into one in **reduced** row echelon form.

$$
\begin{aligned}
B &\sim \begin{bmatrix} 1 & 0 & 2 & 2 & 1 \\ 0 & 1 & 2 & 0 & 2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} && \text{by the above (forward part)} \\
&\overset{\substack{R_1 \to R_1 + R_3 \\ R_2 \to R_2 + R_3}}{\sim} \begin{bmatrix} 1 & 0 & 0 & 2 & 1 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}
\end{aligned}
$$

The backward part of row reduction is now complete: our matrix is in reduced row echelon form. Thus,

$$RREF(B) \quad = \quad \begin{bmatrix} 1 & 0 & 0 & 2 & 1 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

□

5 Solving linear systems

To find the solution set of a linear system, we proceed as follows. First, we form the augmented matrix of our linear system, and using row reduction, we find the reduced row echelon form of that matrix. Then, we "translate" this matrix (in reduced row echelon form) into the linear system that it encodes. The linear system that we obtain is equivalent to the one that we started with, that is, the two systems have the same solution set. We now read off the solution set as follows.

1. If the last column of the augmented matrix (the one behind the dotted vertical line) is a pivot column, then the system is inconsistent, i.e. it has no solutions.

   - For example, suppose that by row reduction, we obtained the following matrix (say, with coefficients in $\mathbb{R}$).

     $$\left[\begin{array}{ccc:c} 1 & 0 & -1 & 0 \\ 0 & 1 & 5 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

     This matrix encodes the following linear system:

     $$\begin{array}{ccccccc} x_1 & & & - & x_3 & = & 0 \\ & & x_2 & + & 5x_3 & = & 0 \\ & & & & 0 & = & 1 \\ & & & & 0 & = & 0 \end{array}$$

     Because of the equation "$0 = 1$," the system is inconsistent (i.e. it has no solutions).

2. If the last column of the augmented matrix (the one behind the dotted vertical line) is **not** a pivot column, but all the other columns **are** pivot columns, then the system has a unique solution.

- For example, suppose that by row reduction, we obtained the following matrix (say, with coefficients in $\mathbb{R}$).

$$\left[\begin{array}{ccc:c} 1 & 0 & 0 & -5 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

This matrix encodes the following linear system:

$$\begin{array}{ccccc} x_1 & & & = & -5 \\ & x_2 & & = & 0 \\ & & x_3 & = & 3 \\ & & 0 & = & 0 \end{array}$$

This system is consistent and has a unique solution, which we can immediately read off, as follows.

$$\begin{array}{rcr} x_1 & = & -5 \\ x_2 & = & 0 \\ x_3 & = & 3 \end{array}$$

3. If the last column of the augmented matrix (the one behind the dotted vertical line) is **not** a pivot column, and at least one of the other columns is also **not** a pivot column, then the system has more than one solution, which we read off as follows. The variables corresponding the **non-pivot** columns (we call these variables *free variables*) may take **any** value; these values (**parameters**) are denotes by letters such as r, s, t. The variables corresponding to the pivot columns are not free, and we express them in terms of our parameters. This form of solution is called the *parametric form of the solution.*

   - For example, suppose that by row reduction, we obtained the following matrix (say, with coefficients in $\mathbb{R}$).

   $$\left[\begin{array}{ccccc:c} 1 & 2 & 0 & 5 & 6 & 0 \\ 0 & 0 & 1 & -1 & 7 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

   This matrix encodes the following linear system:

   $$\begin{array}{ccccccccccc} x_1 & + & 2x_2 & & & + & 5x_4 & + & 6x_5 & = & 0 \\ & & & & x_3 & - & x_4 & + & 7x_5 & = & -3 \\ & & & & & & & & 0 & = & 0 \end{array}$$

   The system is consistent and has more than one solution. The variables variables x_2, x_4, x_5 are free (because the pivot columns of the augmented matrix are columns $2, 4, 5$). The remaining variables are not free. We now read off the solutions as follows:

$$\begin{array}{lcll} x_1 & = & -2r - 5s - 6t & \\ x_2 & = & r & \\ x_3 & = & s - 7t - 3 & \\ x_4 & = & s & \\ x_5 & = & t & \qquad\qquad \text{where } r, s, t \in \mathbb{R} \end{array}$$

Remark: Do not forget to specify which field your parameters come from! Here, we have "$r, s, t \in \mathbb{R}$" because the coefficients of our system were in $\mathbb{R}$.

Example 5.1. *Find the solution set of the following system of linear equations (with coefficients in* $\mathbb{R}$*).*

$$\begin{array}{rcrcrcrcrcr} & & 3x_2 & - & 6x_3 & + & 6x_4 & + & 4x_5 & = & -5 \\ 3x_1 & - & 7x_2 & + & 8x_3 & - & 5x_4 & + & 8x_5 & = & 9 \\ 3x_1 & - & 9x_2 & + & 12x_3 & - & 9x_4 & + & 6x_5 & = & 15 \end{array}$$

Solution. The augmented matrix of this linear system is the matrix A below.

$$A = \left[\begin{array}{rrrrr:r} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{array}\right]$$

But this is precisely the matrix from Example 4.2. The reduced row echelon form of this matrix is

$$RREF(A) = \left[\begin{array}{rrrrr:r} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array}\right].$$

$RREF(A)$ encodes the linear system below.

$$\begin{array}{rcrcrcrcr} x_1 & & & - & 2x_3 & + & 3x_4 & & = & -24 \\ & & x_2 & - & 2x_3 & + & 2x_4 & & = & -7 \\ & & & & & & & x_5 & = & 4 \end{array}$$

This system is consistent and has two free variables (namely, x_3 and x_4). We read off the solutions as follows:

$$\begin{array}{lcll} x_1 & = & 2s - 3t - 24 & \\ x_2 & = & 2s - 2t - 7 & \\ x_3 & = & s & \\ x_4 & = & t & \\ x_5 & = & 4 & \qquad\qquad \text{where } s, t \in \mathbb{R} \end{array}$$

□

Example 5.2. *Find the solution set of the following system of linear equations (with coefficients in $\mathbb{Z}_3$).*

$$\begin{array}{ccccccccc} & & x_2 & + & x_3 & & & = & 2 \\ 2x_1 & + & x_2 & & & + & x_4 & = & 1 \\ 2x_1 & + & x_2 & + & x_3 & + & x_4 & = & 1 \\ x_1 & & & + & 2x_3 & + & 2x_4 & = & 1 \end{array}$$

Solution. The augmented matrix of this linear system is the matrix B below.

$$B = \left[\begin{array}{cccc|c} 0 & 1 & 1 & 0 & 2 \\ 2 & 1 & 0 & 1 & 1 \\ 2 & 1 & 1 & 1 & 1 \\ 1 & 0 & 2 & 2 & 1 \end{array}\right]$$

But this is precisely the matrix from Example 4.3. The reduced row echelon form of this matrix is

$$RREF(B) = \left[\begin{array}{cccc|c} 1 & 0 & 0 & 2 & 1 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

$RREF(B)$ encodes the linear system below.

$$\begin{array}{ccccccc} x_1 & & & + & 2x_4 & = & 1 \\ & x_2 & & & & = & 2 \\ & & x_3 & & & = & 0 \\ & & & & 0 & = & 0 \end{array}$$

This system is consistent and has one free variable (namely, x_4). We read off the solutions as follows:

$$\begin{array}{lcl} x_1 & = & t+1 \\ x_2 & = & 2 \\ x_3 & = & 0 \\ x_4 & = & t \qquad \text{where } t \in \mathbb{Z}_3 \end{array}$$

Remark: To get x_1, we computed $x_1 = -2x_4 + 1 = x_4 + 1 = t + 1$, where we used the fact that $-2 \equiv 1 \pmod 3$. □

5.1 Specifying the number of solutions

An inconsistent linear system has zero solutions. A consistent system may have a unique solution (i.e. exactly one solution), or it may have more than one solution. A consistent system has more than one solution if, at the end of our calculation, we get at least one free variable. Each free variable can take an arbitrary value from the field in question. So, if our coefficients are in $\mathbb{R}$ or $\mathbb{C}$, and the system is consistent with at least one free variable, then the system has infinitely many solutions. If the coefficients are in $\mathbb{Z}_p$ for some prime number p, and our linear system is consistent with exactly k free variables, then the number of solutions is precisely p^k.

For instance, the system from Example 5.1 has infinitely many solutions. On the other hand, the system from Example 5.2 has three solutions (one for each of the three possible values of t).