# Systems of Linear Equations

This week we will learn about:

- Systems of linear equations,
- Elementary row operations and Gaussian elimination, and
- The (reduced) row echelon form of a matrix.

Extra reading and watching:

- Section 2.1 in the textbook
- Lecture videos 17, 18, 19, 20, and 21 on YouTube
- System of linear equations at Wikipedia
- Gaussian elimination at Wikipedia

Extra textbook problems:

⋆ 2.1.1, 2.1.2, 2.1.4, 2.1.5

⋆⋆ 2.1.7–2.1.9, 2.1.11, 2.1.15–2.1.17, 2.1.25, 2.1.26

⋆⋆⋆ 2.1.18, 2.1.23, 2.1.24, 2.1.27–2.1.29

☠ none this week

# (Systems of) Linear Equations

Much of linear algebra is about solving and manipulating the simplest types of equations that exist—linear equations:

## Definition 5.1 — Linear Equations

A **linear equation** in $n$ variables $x_1, x_2, \ldots, x_n$ is an equation that can be written in the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b,$$

where $a_1, a_2, \ldots, a_n$ and $b$ are constants.

**Example.** Examples of linear and non-linear equations.

$2x + 5y = 17$ ✓  $\frac{3}{5}x = 7y + 2.3$ ✓

$3x + \sqrt{2}y + 4z = \sin(1)$ ✓

numbers!

$3x + 2\sqrt{y} + 4\sin(z) = 0$ ✗

cannot apply weird functions to variables

$2xy - 3x/z = -7$

cannot multiply or divide variables by each other

The point is that an equation is linear if each variable is only multiplied by a constant: variables cannot be multiplied by other variables, they can only be raised to the first power, and they cannot have other functions applied to them.

You (hopefully) learned how to manipulate linear equations quite some time ago, and then you "ramped up" to non-linear equations (like $x^2 = 2$ or $2^x = 8$). In this course, we instead "ramp up" in a different direction: we work with multiple linear equations simultaneously.

## Definition 5.2 — Systems of Linear Equations

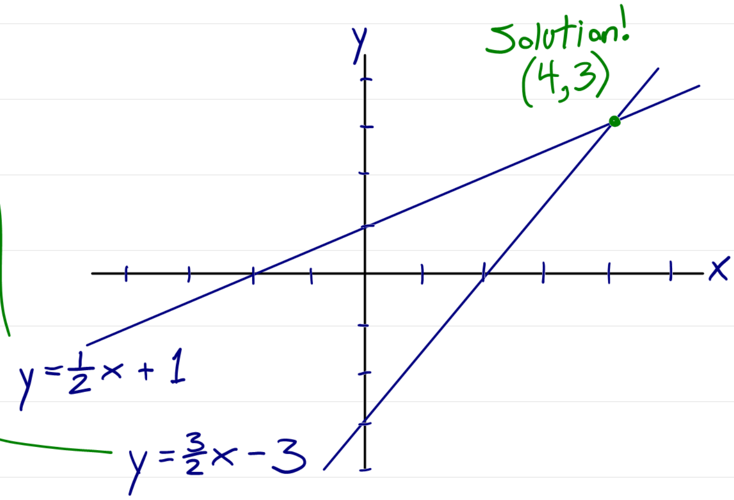
A **system of linear equations** (or a **linear system**) is a finite set of linear equations, each with the same variables $x_1, x_2, \ldots, x_n$.

Some more terminology:

- A **solution** of a system of linear equations is a vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ whose entries satisfy *all* of the linear equations in the system.
- The **solution set** of a system of linear equations is the set of *all* solutions of the system.

**Example.** Solving a linear system geometrically.

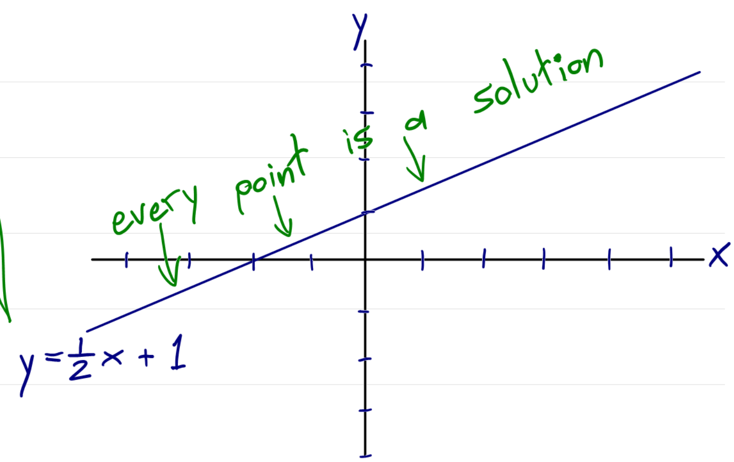
$$x - 2y = -2$$
$$3x - 2y = 6$$




Solution is unique (non-parallel lines intersect once).

**Example.** Two more (weirder!) systems of linear equations.

$$x - 2y = -2$$
$$3x - 6y = -6$$




Infinitely many solutions.

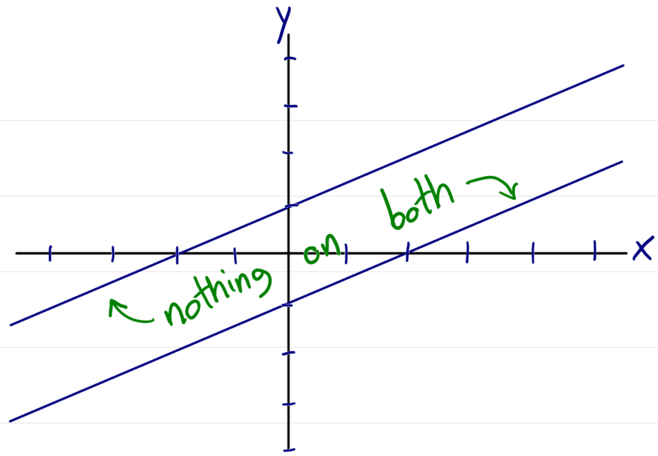
$x - 2y = -2$

$3x - 6y = 6$

No solutions (no intersection).

$y = \frac{1}{2}x + 1$

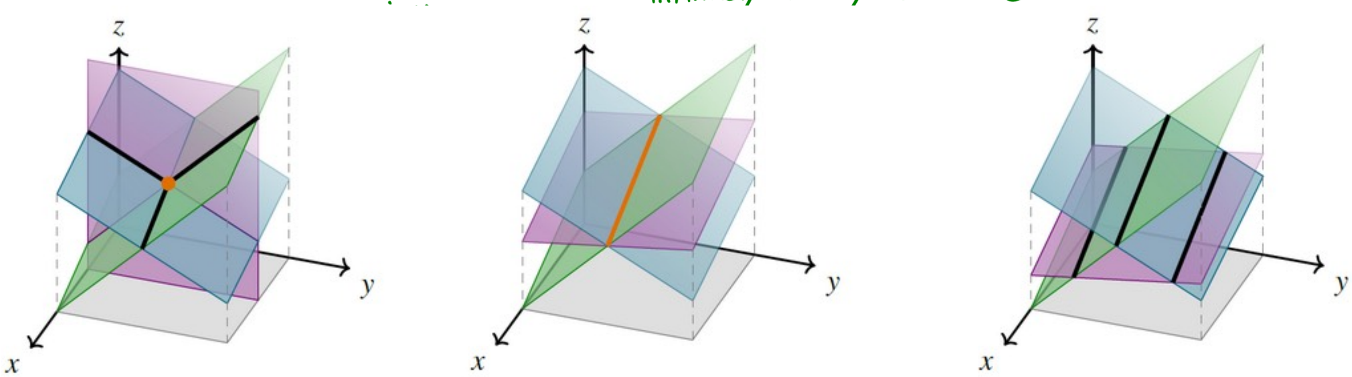
$y = \frac{1}{2}x - 1$




The above examples show that systems of linear equations can have no solutions, exactly one solution, or infinitely many solutions. We will show shortly that these are the only possibilities.

Note that we can also visualize systems of linear equations with 3 variables in 3 dimensions, but it's a bit tougher:

infinitely many solutions




# Matrix Equations

One of the primary uses of matrices is that they give us a way of working with linear systems much more compactly and cleanly. In particular, any system of linear equations...

$$
\begin{aligned}
a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n &= b_1 \\
a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n &= b_2 \\
&\vdots \\
a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n &= b_m
\end{aligned}
$$

can be rewritten as a single matrix equation:

$$A\vec{x} = \vec{b}, \quad \text{where}$$

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \text{and} \quad \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

**Example.** Write the following system of linear equations as a single matrix equation:

$$\begin{aligned} x - 2y &= -2 \\ 3x - 2y &= 6 \end{aligned} \quad \Longleftrightarrow \quad \begin{bmatrix} 1 & -2 \\ 3 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -2 \\ 6 \end{bmatrix}$$

(from earlier)

The advantage of writing linear systems in this way (beyond the fact that it requires less writing) is that we can now make use of the various properties of matrices and matrix multiplication that we already know to help us understand linear systems a bit better. For example, we can now prove the observation that we made earlier: every linear system has either 0, 1, or infinitely many solutions.

**Theorem 5.1 — Trichotomy for Linear Systems**

Every system of linear equations has either

**a)** no solutions,

**b)** exactly one solution, or

**c)** infinitely many solutions.

*Proof.* We just need to show that if a linear system has at least two different solutions, then it has infinitely many solutions.

Suppose there are two solutions, $\vec{x}_1$ and $\vec{x}_2$. That is, $A\vec{x}_1 = \vec{b}$ and $A\vec{x}_2 = \vec{b}$, and $\vec{x}_1 \neq \vec{x}_2$.

Then for all $c \in \mathbb{R}$ we have
$A(c\vec{x}_1 + (1-c)\vec{x}_2) = cA\vec{x}_1 + (1-c)A\vec{x}_2 = c\vec{b} + (1-c)\vec{b} = \vec{b}$.

$\therefore$ $c\vec{x}_1 + (1-c)\vec{x}_2$ is a solution for all $c \in \mathbb{R}$, so there are infinitely many solutions. ∎

When a system of linear equations has at least one solution (i.e., in cases (b) and (c) of the theorem), it is called **consistent**. If it has no solutions (i.e., in case (a) of the theorem), it is called **inconsistent**.

## Solving Linear Systems

Let's now discuss how we might find the solutions of a system of linear equations. If the linear system has a certain special form, then solving it is fairly intuitive.

**Example.** Solve the following system of linear equations:

$$\begin{aligned} x + 3y - 2z &= 5 \\ 2y - 6z &= 4 \\ 3z &= 6 \end{aligned}$$

- $3z = 6 \Rightarrow z = 2$
- Plug $z=2$ into two other equations to get $x + 3y = 9$, $2y = 16$
- $2y = 16 \Rightarrow y = 8$
- Plug $y = 8$ into $x + 3y = 9$ to get $x = -15$.

$\therefore$ $(x,y,z) = (-15, 8, 2)$ is the unique solution.

The procedure that we used to solve the previous example is called **back substitution**, and it worked because of the "triangular" nature of the equations. We were able to easily solve for $z$, which we then could plug into the second equation and easily solve for $y$, which we could plug into the first equation and easily solve for $x$.

So let's try to put *every* system of equations into this triangular form! We start by

eliminating $x$ from equations 2 and 3, and then eliminating $y$ from equation 3.
(Then to solve, do what we did before.)

To reduce the amount of writing we have to do when solving the linear system $A\mathbf{x} = \mathbf{b}$, we typically use the block matrix $[\,A \mid \mathbf{b}\,]$.

**Example.** Solve the following (much uglier) system of linear equations:

$$\begin{aligned} ①\quad & x + 3y - 2z = 5 \\ ②\quad & 3x + 5y + 6z = 7 \\ ③\quad & 2x + 4y + 3z = 8 \end{aligned}$$

$R_2$ means "row 2"

$$\left[\begin{array}{ccc|c} 1 & 3 & -2 & 5 \\ 3 & 5 & 6 & 7 \\ 2 & 4 & 3 & 8 \end{array}\right]$$

$$\begin{aligned} & x + 3y - 2z = 5 \\ ②-3①:\quad & -4y + 12z = -8 \\ ③-2①:\quad & -2y + 7z = -2 \end{aligned}$$

$$\xrightarrow[R_3 - 2R_1]{R_2 - 3R_1} \left[\begin{array}{ccc|c} 1 & 3 & -2 & 5 \\ 0 & -4 & 12 & -8 \\ 0 & -2 & 7 & -2 \end{array}\right]$$

$$\begin{aligned} & x + 3y - 2z = 5 \\ \tfrac{-1}{4}②:\quad & y - 3z = 2 \\ & -2y + 7z = -2 \end{aligned}$$

$$\xrightarrow{\frac{-1}{4}R_2} \left[\begin{array}{ccc|c} 1 & 3 & -2 & 5 \\ 0 & 1 & -3 & 2 \\ 0 & -2 & 7 & -2 \end{array}\right]$$

not absolutely necessary, but makes the next step easier

$$\begin{aligned} x + 3y - 2z &= 5 \\ y - 3z &= 2 \\ z &= 2 \end{aligned}$$

③+2②:

$$\xrightarrow{R_3+2R_2} \left[\begin{array}{ccc|c} 1 & 3 & -2 & 5 \\ 0 & 1 & -3 & 2 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

The linear system is now upper triangular and can be solved by back substitution:
②: $y-6=2 \Rightarrow y=8$, and ①: $x+24-4=5 \Rightarrow x=-15$.
$\therefore (x,y,z)=(-15,8,2)$ is the unique solution.

There were three basic types of operations that we performed on the matrix when solving the previous system of linear equations. These are called the **elementary row operations**:

**a)** Adding a multiple of a row to another row ($R_i + cR_j$).

$$\left[\begin{array}{cc|c} 1 & 2 & 2 \\ 3 & 4 & 7 \end{array}\right] \xrightarrow{R_2-3R_1} \left[\begin{array}{cc|c} 1 & 2 & 2 \\ 0 & -2 & 1 \end{array}\right]$$

**b)** Multiplying a row by a non-zero constant ($cR_i$).

$$\left[\begin{array}{cc|c} 2 & 4 & 10 \\ 1 & 3 & -2 \end{array}\right] \xrightarrow{\frac{1}{2}R_1} \left[\begin{array}{cc|c} 1 & 2 & 5 \\ 1 & 3 & -2 \end{array}\right]$$

**c)** Interchanging rows ($R_i \leftrightarrow R_j$).

$$\left[\begin{array}{cc|c} 0 & 3 & 5 \\ 1 & 2 & 3 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{cc|c} 1 & 2 & 3 \\ 0 & 3 & 5 \end{array}\right]$$

These are the only operations we will ever need to solve a linear system!

As mentioned before, our goal when solving these systems of equations is to first make the matrix "triangular." We now make this a bit more precise.

**Definition 5.3 — (Reduced) Row Echelon Form**

A matrix is in **row echelon form** if it satisfies both of these properties:

**a)** All rows consisting entirely of zeros are below the non-zero rows.

**b)** In each non-zero row, the first non-zero entry (called the **leading entry**) is to the left of any leading entries below it.

If the matrix also satisfies the following additional constraints, then it is in **reduced row echelon form** (**RREF**):

**c)** The leading entry in each non-zero row is 1.

**d)** Each leading 1 is the only non-zero entry in its column.

**Example.** Some matrices that are and are not in (reduced) row echelon form.

$$\left[\begin{array}{cccc|c} 3 & 2 & 1 & 4 & 6 \\ 1 & 3 & 1 & 1 & 0 \\ 0 & 2 & -2 & 0 & 3 \\ 0 & 0 & 1 & 0 & 1 \end{array}\right] \qquad \left[\begin{array}{cccc|c} 0 & 2 & 1 & 4 & 6 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

not in REF

yes, in REF, but not RREF

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

yes, in RREF (and thus REF too)

To solve a system of linear equations, we use elementary row operations to bring it into row echelon form. Once it is in this form, we can easily solve it via back-substitution.

Alternatively, we can use elementary row operations to bring a matrix all the way into reduced row echelon form. Once an augmented matrix is in this form, the solutions of the associated linear system can be read directly from the entries of the matrix.

**Example.** Find the solutions of the systems of equations represented by the following augmented matrices:

$$\left[\begin{array}{cc|c} 1 & -2 & -2 \\ 3 & -2 & 6 \end{array}\right] \xrightarrow{R_2 - 3R_1} \left[\begin{array}{cc|c} 1 & -2 & -2 \\ 0 & 4 & 12 \end{array}\right], \text{ which is in REF.}$$

(Eq. 2): $4y = 12$, so $y = 3$. (Eq. 1): $x - 6 = -2$, so $x = 4$.

$\therefore$ $(x,y) = (4,3)$ is the unique solution.

---

$$\left[\begin{array}{cc|c} 1 & 0 & -6 \\ 0 & 1 & 5 \end{array}\right], \text{ which is in RREF.}$$

(Eq. 2): $y = 5$, (Eq. 1): $x = -6$. Done.

The process of using elementary row operations to bring a matrix into a row echelon form is called **row reduction**. The process of using row reduction to find a row echelon form, and then back substitution to solve the system of linear equations, is called **Gaussian elimination**.

**Example.** Use Gaussian elimination to solve the following system of linear equations:

$$\begin{aligned} x + 2y - 4z &= -4 \\ 2x + 4y \phantom{{}+3z} &= 0 \\ -x + \phantom{2}y + 3z &= 6 \end{aligned}$$

$$\left[\begin{array}{ccc|c} 1 & 2 & -4 & -4 \\ 2 & 4 & 0 & 0 \\ -1 & 1 & 3 & 6 \end{array}\right] \xrightarrow[R_3 + R_1]{R_2 - 2R_1} \left[\begin{array}{ccc|c} 1 & 2 & -4 & -4 \\ 0 & 0 & 8 & 8 \\ 0 & 3 & -1 & 2 \end{array}\right]$$

$$R_2 \leftrightarrow R_3 \quad \left[\begin{array}{ccc|c} 1 & 2 & -4 & -4 \\ 0 & 3 & -1 & 2 \\ 0 & 0 & 8 & 8 \end{array}\right],$$ which is in REF.

Back substitution: (Eq. 3): $8z = 8$, so $z = 1$.
(Eq. 2): $3y - 1 = 2$, so $y = 1$. (Eq. 1): $x + 2 - 4 = -4$, so $x = -2$.
$\therefore \quad (x, y, z) = (-2, 1, 1)$.

Some notes about row echelon form and elementary row operations are in order:

- The elementary row operations are reversible: if there is an elementary row operation that transforms $A$ into $B$, then there is an elementary row operation that transforms $B$ into $A$.
- Is the row echelon form of a matrix **unique** or **not unique**?
- Two matrices are called **row equivalent** if one can be converted to the other via elementary row operations.

The process of using row reduction to find a *reduced* row echelon form, and hence solve the system of linear equations, is called **Gauss–Jordan elimination**.

**Example.** Use Gauss–Jordan elimination to solve the following linear system:

$$\begin{aligned} x + 2y - 4z &= -4 \\ 2x + 4y \phantom{{}+3z} &= 0 \\ -x + \phantom{2}y + 3z &= 6 \end{aligned}$$

$$\left[\begin{array}{ccc|c} 1 & 2 & -4 & -4 \\ 2 & 4 & 0 & 0 \\ -1 & 1 & 3 & 6 \end{array}\right] \xrightarrow{\text{row operations}} \left[\begin{array}{ccc|c} 1 & 2 & -4 & -4 \\ 0 & 3 & -1 & 2 \\ 0 & 0 & 8 & 8 \end{array}\right] \text{(REF)}$$

$$\xrightarrow{\substack{\frac{1}{3}R_2 \\ \frac{1}{8}R_3}} \left[\begin{array}{ccc|c} 1 & 2 & -4 & -4 \\ 0 & 1 & -\frac{1}{3} & \frac{2}{3} \\ 0 & 0 & 1 & 1 \end{array}\right]$$

$$\xrightarrow[R_2+\frac{1}{3}R_3]{R_1+4R_3} \left[\begin{array}{ccc|c} 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right] \xrightarrow{R_1-2R_2} \left[\begin{array}{ccc|c} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

RREF! Just read off the solution:

$(x,y,z) = (-2,1,1)$

Some notes about reduced row echelon form and Gauss–Jordan elimination are in order:

- Neither Gaussian elimination nor Gauss–Jordan elimination is a "better" method than the other. Which one you use is typically just based on personal preference.
- Is the reduced row echelon form of a matrix **unique** or **not unique**?
- To check if two matrices are row equivalent, check whether or not they have the same reduced row echelon form.

# Free Variables and Systems Without Unique Solutions

Recall that systems of linear equations do not always have a unique solution: they might have no solutions or infinitely many solutions. Identifying systems with no solutions is intuitive enough...

**Example.** Solve the following system of linear equations:

$$\begin{aligned} x + 2y - 2z &= -4 \\ 2x + 4y + \phantom{2}z &= 0 \\ x + 2y + 7z &= 2 \end{aligned}$$

$$\left[\begin{array}{ccc|c} 1 & 2 & -2 & -4 \\ 2 & 4 & 1 & 0 \\ 1 & 2 & 7 & 2 \end{array}\right] \xrightarrow[R_3-R_1]{R_2-2R_1} \left[\begin{array}{ccc|c} 1 & 2 & -2 & -4 \\ 0 & 0 & 5 & 8 \\ 0 & 0 & 9 & 6 \end{array}\right]$$

$$\xrightarrow{R_3 - \frac{9}{5}R_2} \left[\begin{array}{ccc|c} 1 & 2 & -2 & -4 \\ 0 & 0 & 5 & 8 \\ 0 & 0 & 0 & -42/5 \end{array}\right]$$

Row 3 means $0x + 0y + 0z = -42/5$, which is nonsense.

$\therefore$ There are no solutions.

The behaviour in the previous example is what happens in general: a linear system has no solutions if and only if the row echelon forms of its augmented matrix $[\,A \mid \mathbf{b}\,]$ have a row consisting of zeros in the left ($A$) block and a non-zero entry in the right ($\mathbf{b}$) block.

Things are somewhat more complicated when a system of equations has infinitely many solutions, though. After all, how can we even *describe* all of the solutions in this case? We illustrate the method with a couple more examples:

**Example.** Solve the following system of linear equations:

$$\begin{aligned} v - 2w \qquad\qquad + 2z &= 3 \\ x \qquad - 3z &= 7 \\ y + \ z &= 4 \end{aligned}$$

$$\begin{array}{cccccc} v & w & x & y & z & \\ \end{array}$$
$$\left[\begin{array}{ccccc|c} 1 & -2 & 0 & 0 & 2 & 3 \\ 0 & 0 & 1 & 0 & -3 & 7 \\ 0 & 0 & 0 & 1 & 1 & 4 \end{array}\right]$$

Already in RREF! Can rearrange to get $v, x, y$ in terms of $w$ and $z$, but $w$ and $z$ can be anything (infinitely many solutions).

<u>Leading variable:</u> Corresponds to a column with a leading entry $(v, x, y)$.

<u>Free variable:</u> The rest ($w$ and $z$).

$$(v, w, x, y, z) = (3 + 2w - 2z,\ w,\ 7 + 3z,\ 4 - z,\ z) \qquad (w, z \text{ free}).$$

**Example.** Solve the following system of linear equations:

$$\begin{aligned} w - \phantom{2}x - y + 2z &= 1 \\ 2w - 2x - y + 3z &= 3 \\ -w + \phantom{2}x - y \phantom{+3z} &= -3 \end{aligned} \qquad \left[\begin{array}{cccc|c} 1 & -1 & -1 & 2 & 1 \\ 2 & -2 & -1 & 3 & 3 \\ -1 & 1 & -1 & 0 & -3 \end{array}\right]$$

$$\xrightarrow[R_3+R_1]{R_2-2R_1} \left[\begin{array}{cccc|c} 1 & -1 & -1 & 2 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & -2 & 2 & -2 \end{array}\right]$$

$$\xrightarrow[R_3+2R_2]{R_1+R_2} \begin{array}{c} \begin{array}{cccc} w & x & y & z \end{array} \\ \left[\begin{array}{cccc|c} \boxed{1} & -1 & 0 & 1 & 2 \\ 0 & 0 & \boxed{1} & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \end{array}$$

This is now in RREF!

Leading: $w$ and $y$

Free: $x$ and $z$

(Eq. 1): $w - x + z = 2$, so $w = 2 + x - z$

(Eq. 2): $y - z = 1$, so $y = 1 + z$.

$\therefore\ (w, x, y, z) = (2 + x - z, x, 1 + z, z)$ ($x, z$ free)

Again, the behaviour in the previous example is completely general: variables corresponding to columns that have a leading entry in the row echelon form are called **leading variables**, and we write these variables in terms of the non-leading variables (called **free variables**).

Each free variable corresponds to one "dimension" or "degree of freedom" in the solution set. For example, if there is one free variable then the solution set is a line, if there are two then it is a plane, and so on.

In previous example, it was a plane (2-dimensional).