# Lecture 1: Solving a system of linear equations

January 3, 2024

# Lecture 1: Solving a system of linear equations

1. What does it mean?
2. Why do we care?
3. What do we do?
4. Why did that work?

# What does it mean?

This is a system of linear equations.

$$\begin{array}{rcrcrcl} x & + & 2y & - & z & = & 5 \\ 3x & + & y & - & 2z & = & 9 \\ -x & + & 4y & + & 2z & = & 0 \end{array}$$

The $x, y, z$ are the variables.
Linear means each variable appears with a number as a coefficient.
Things like $x^2 + y = 17$ or $\frac{\sqrt{x}}{1+xy} = 4$ are equations but they are not linear.

A solution of this system is an assignment of (numerical) values to $x, y, z$ that simultaneously satisfies all the equations. For example $(x, y, z) = (2, 1, -1)$ is a solution (this is easily checked).

Is it the only solution? How could we know?

"$(x, y, z) = (2, 1, -1)$" means $x = 2, y = 1, z = -1$.

# What does it mean?

Here is another system of linear equations.

$$\begin{array}{rcrcrcl} x & + & 2y & - & z & = & 5 \\ 3x & + & y & - & 2z & = & 9 \\ 6x & + & 7y & - & 5z & = & 24 \end{array}$$

$(x, y, z) = (2, 1, -1)$ is a solution of this system. (Check this!)
So is $(x, y, z) = (5, 2, 4)$.
So is $(x, y, z) = (-1, 0, -6)$.
So is $(x, y, z) = (47, 8, 34)$.

What else? There are many more. What do these four have in common? (This is not obvious).

> To solve a system of linear equations means to identify all solutions and describe them in a clear way.

# Why do we care?

- Because we can do it. There are very few types of equations (or systems of equations) that mathematics can solve exactly. We can't solve polynomial equations or differential equations for example, except in very special cases. Systems of linear equations are "easy" in the sense that we have methods that always identify a solution. This is already a reason to care.
- For this reason, many numerical and computational methods for approximating solutions to differential equations involve translating the problem to a system of linear equations. These situations arise often in mathematical modelling (of physical phenomena like fluid flow, or finance, economics, ecology etc) where some initial or boundary information might be known.
- Systems of linear equations arise very frequently in applied contexts, and in situations that involve interpreting observed data, and in optimization problems.

# A few (baking) examples

1. A bakery makes cakes (500g), loaves (500g) and biscuits (in 500g packets), all using the same ingredients (flour, sugar, eggs) in different ratios. Their supply chain was interrupted this week and they have limited stocks of ingredients. How much of each product should they make in order to use their supplies as fully as possible?
2. In their spicy cookies, the bakery uses a mixture of spices that has 60% cinnamon, 20% nutmeg and 20% ginger. Spice deliveries have been hit by the supply chain problems, and the bakery has run out of all three. They do have two tubs of mixed spice. The first has 40% cinnamon, 40% nutmeg and 20% ginger, and the second has 70% cinnamon, 15% nutmeg and 15% ginger. Can they reconstruct their usual mixture by combining supplies from the two tubs? If so, how?

Exercise How can these problems be represented as linear systems?

# What do we do to solve a linear system?

First, we rewrite the system of equations as a matrix. A matrix is a rectangular array of numbers.

$$\begin{array}{rcrcrcr} x & + & 2y & - & z & = & 5 \\ 3x & + & y & - & 2z & = & 9 \\ -x & + & 4y & + & 2z & = & 0 \end{array} \leftrightarrow \left[\begin{array}{rrr|r} 1 & 2 & -1 & 5 \\ 3 & 1 & -2 & 9 \\ -1 & 4 & 2 & 0 \end{array}\right] \begin{array}{l} \text{Equation 1} \\ \text{Equation 2} \\ \text{Equation 3} \end{array}$$

This matrix is called the augmented matrix of the system of equations. Its rows correspond to the equations in the system, and its columns respectively to the variables $x, y, z$, and to the right hand side (RHS).

Once we know about this correspondence, we can reconstruct the equations from the matrix.

# What do we do to solve a linear system?

We work with the augmented matrix, applying operations of three types to convert it to a simpler form. At each step, we can do one of the following elementary row operations (EROs)

1. Add a scalar[1] multiple of one row to another row.
2. Multiply all entries of a row by the same non-zero scalar.
3. Swap two rows.

| ERO | Matrix | System |
|---|---|---|
| | $\begin{bmatrix} 1 & 2 & -1 & 5 \\ 3 & 1 & -2 & 9 \\ -1 & 4 & 2 & 0 \end{bmatrix}$ | $\begin{array}{rcrcrcr} x & + & 2y & - & z & = & 5 \\ 3x & + & y & - & 2z & = & 9 \\ -x & + & 4y & + & 2z & = & 0 \end{array}$ |
| $R2 \to R2 + (-3)R1$<br>$R3 \to R3 + R1$ | $\begin{bmatrix} 1 & 2 & -1 & 5 \\ 0 & -5 & 1 & -6 \\ 0 & 6 & 1 & 5 \end{bmatrix}$ | $\begin{array}{rcrcrcr} x & + & 2y & - & z & = & 5 \\ & + & -5y & + & z & = & -6 \\ & & 6y & + & z & = & 5 \end{array}$ |

[1] "scalar" means number

| ERO | Matrix | System |
|---|---|---|
| $R2 \to R2 - 3R1$ | $\begin{bmatrix} 1 & 2 & -1 & 5 \\ 0 & -5 & 1 & -6 \\ 0 & 6 & 1 & 5 \end{bmatrix}$ | $\begin{array}{rcrcrcr} x & + & 2y & - & z & = & 5 \\ & - & 5y & + & z & = & -6 \\ & & 6y & + & z & = & 5 \end{array}$ |
| $R2 \to R2 + R3$ | $\begin{bmatrix} 1 & 2 & -1 & 5 \\ 0 & 1 & 2 & -1 \\ 0 & 6 & 1 & 5 \end{bmatrix}$ | $\begin{array}{rcrcrcr} x & + & 2y & - & z & = & 5 \\ & & y & + & 2z & = & -1 \\ & & 6y & + & z & = & 5 \end{array}$ |
| $R3 \to R3 - 6R2$ | $\begin{bmatrix} 1 & 2 & -1 & 5 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & -11 & 11 \end{bmatrix}$ | $\begin{array}{rcrcrcr} x & + & 2y & - & z & = & 5 \\ & & y & + & 2z & = & -1 \\ & & & - & 11z & = & 11 \end{array}$ |
| $R3 \times \left(-\frac{1}{11}\right)$ | $\begin{bmatrix} 1 & 2 & -1 & 5 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix}$ | $\begin{array}{rcrcrcr} x & + & 2y & - & z & = & 5 \\ & & y & + & 2z & = & -1 \\ & & & & z & = & -1 \end{array}$ |

| ERO | Matrix | System |
|---|---|---|
| $R1 \to R1 - 2R2$ | $\left[\begin{array}{ccc\|c} 1 & 0 & -5 & 7 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 1 & -1 \end{array}\right]$ | $\begin{array}{rcrcrcr} x & & & - & 5z & = & 7 \\ & & y & + & 2z & = & -1 \\ & & & & z & = & -1 \end{array}$ |
| $R1 \to R1 + 5R3$<br>$R2 \to R2 - 2R3$ | $\left[\begin{array}{ccc\|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{array}\right]$ | $\begin{array}{rcrcrcr} x & & & & & = & 2 \\ & & y & & & = & 1 \\ & & & & z & = & -1 \end{array}$ |

Conclusion The only solution of the system is $(x, y, x) = (2, 1, -1)$ (or $x = 2, y = 1, z = -1$ if you prefer to write it this way).

# Why did that work?

Every ERO changes the system, by rewriting one equation.
But the new system that has exactly the same solutions as the original.

Suppose that a particular choice of values of $x, y, z$ satisfies two different linear equations. Then it also satisfies

- the equation that we get by adding these two together;
- any equation that we get from either of them by multiplying by a non-zero scalar.

So every solution of the original system is a solution of every subsequent system, including the "simplified" last one.

Also, every ERO is reversible by an ERO, so the original system can be reconstructed from the final one (think about this important point). This means that every solution of the finl system is a solution of the orginal one - we didn't introduce new solutions along the way.