Lecture 13: Systems of Linear Algebraic Equations

Row reduction, REF and RREF (§3.2)

APPM2360: Introduction to Differential Equations with Linear Algebra

1 Setting the stage

Recall the two kinds of *linear* problem we care about:

(1) $\mathrm{L}[y] = f(t)$ (differential equations), (2) $\mathbf{L}\vec{x} = \vec{b}$ (algebraic equations).

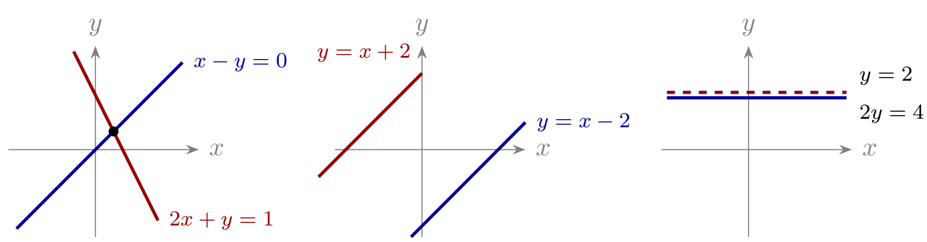
These share many similarities. **Today:** tools for solving (2). **Later:** *Sike!* — the same tools also work for (1).

Suppose we have two linear equations

$$\begin{aligned} 2x + y &= 1 \\ x - y &= 0 \end{aligned}$$

Are there solutions? How many?

1.1 In 2D we can just plot



one solution (lines cross) | no solutions (parallel, distinct) | ∞ many solutions (same line)

(In the third picture the two lines have been drawn slightly apart only so that both are visible — they are in fact the *same* line.)

If the lines are parallel, there are *either* ∞ many solutions *or* 0 solutions.

1.2 Equivalent systems

What about

$$\begin{aligned} 10y - 4x &= 2 \\ -5y + 2x &= -1 \end{aligned}\ ?$$

Multiply the second equation by -2 to see that it is *the same equation*:

$$\begin{aligned} 10y - 4x &= 2 \\ 10y - 4x &= 2 \end{aligned}.$$

The two lines coincide, so there are infinitely many solutions. Note that this is an example of an *equivalent system*: whenever we algebraically manipulate one or more equations in this way to write down another set of equations that *looks* different but is mathematically equivalent, that is an "equivalent system".

Compare this with

$$\begin{array}{r} [\,10y - 4x = 2\,] \\ -[\,10y - 4x = 3\,] \\ \hline 0y - 0x = -1 \end{array} \qquad \Longrightarrow \qquad 0 = -1 \quad ✗$$

which is a contradiction: there are **no solutions** (i.e., choices of x and y values) that can make these equations simultaneously true.

2 Matrices and row operations

Generally, we seek to understand **solutions / existence / uniqueness** of systems of algebraic equations, using matrices and "row operations".

Example 2.1 (A 3×3 system)**.**

$$\begin{aligned} x + 2y + z &= 2 \\ 2x + 6y + z &= 7 \\ x + y + 4z &= 3 \end{aligned}$$

First write this as $\mathbf{A}\vec{x} = \vec{b}$:

$$\underbrace{\begin{bmatrix} 1 & 2 & 1 \\ 2 & 6 & 1 \\ 1 & 1 & 4 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} x \\ y \\ z \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} 2 \\ 7 \\ 3 \end{bmatrix}}_{\vec{b}}.$$

Shorthand — the **augmented matrix**:

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 2 & 6 & 1 & 7 \\ 1 & 1 & 4 & 3 \end{array}\right] \begin{array}{l} \leftarrow \text{1st row } R_1 \\ \leftarrow R_2 \\ \leftarrow R_3 \end{array}$$

2.1 The three row operations

We may manipulate the equations in one of the following ways (R_i^* denotes the row *after* the operation, R_i the row *before*):

(i)	Multiply row i by $c \neq 0$:	$R_i^* = c\,R_i$
(ii)	Exchange rows i and j:	$R_i \leftrightarrow R_j$, or: $R_i^* = R_j,\ R_j^* = R_i$
(iii)	Add $c\times$ row j to row i:	$R_i^* = R_i + c\,R_j$

2.2 How to use these row operations: one strategy

When using row operations on augmented matrices like this, our goal is to write down a sequence of equivalent systems until we end up with something that looks like an identity matrix to the left of the vertical bar. Here's one strategy that will always work:

- Working left to right, get 0's <u>below</u> the diagonal.
- Multiply by scalars where necessary to make the diagonal entries = 1.
- Working left to right, get 0's above the diagonal.

3 Worked example: full row reduction

Example 3.1 (Row-reducing the 3×3 system)**.**

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 2 & 6 & 1 & 7 \\ 1 & 1 & 4 & 3 \end{array}\right] \qquad \begin{array}{l} R_2^* = R_2 - 2R_1 \\ R_3^* = R_3 - R_1 \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & 2 & -1 & 3 \\ 0 & -1 & 3 & 1 \end{array}\right] \qquad \begin{array}{l} R_2^* = \frac{1}{2}R_2 \\ R_3^* = R_3 + \frac{1}{2}R_2 \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & 1 & -\frac{1}{2} & \frac{3}{2} \\ 0 & 0 & \frac{5}{2} & \frac{5}{2} \end{array}\right] \qquad R_3^* = \frac{2}{5}R_3$$

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 2 \\ 0 & 1 & -\frac{1}{2} & \frac{3}{2} \\ 0 & 0 & 1 & 1 \end{array}\right] \qquad \begin{array}{l} R_1^* = R_1 - 2R_2 \\ R_2^* = R_2 + \frac{1}{2}R_3 \end{array}$$

This is **REF** (defined below) — we could stop here and back-substitute by saying $z = 1$, $y - (1/2)z = 3/2 \Rightarrow y = 3/2 + 1/2$, etc. But let's keep going for now.

$$\left[\begin{array}{ccc|c} 1 & 0 & 2 & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{array}\right] \qquad R_1^* = R_1 - 2R_3$$

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & -3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{array}\right] \qquad \textbf{this is RREF.}$$

3.1 REF and RREF

Row echelon form (REF):

(i) Zero rows (if any) are on the bottom.

(ii) The leftmost nonzero entry of every nonzero row is 1 — call it the **pivot**, or **leading 1**.

(iii) Each pivot is to the right of any pivots above it.

Reduced row echelon form (RREF): additionally,

(iv) Each pivot is the only nonzero entry in its column (its **pivot column**).

Recall that the RREF above is equivalent to

$$\begin{aligned} x + 0y + 0z &= -3 \\ 0x + y + 0z &= 2 \\ 0x + 0y + z &= 1 \end{aligned} \qquad \Longrightarrow \qquad \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}.$$

Here we can define the **rank** of a matrix in terms of its REF/RREF:

$$\#\text{ pivots } = \#\text{ columns } = \text{ rank},$$

which is why the solution is unique.

To add in lecture

Define REF/RREF here. Also define $\vec{x} = \mathbf{A}^{-1}\vec{b}$, together with

$$(\mathbf{A}^{-1})^{-1} = \mathbf{A}, \qquad (\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}, \qquad (\mathbf{A}^{\mathsf{T}})^{-1} = (\mathbf{A}^{-1})^{\mathsf{T}},$$

$$(\mathbf{A}+\mathbf{B})^{-1} \neq \mathbf{A}^{-1} + \mathbf{B}^{-1}.$$

STUDENTS READING THIS: take a quick look at pages 146 – 148 (up to and including example 2) in section 3.3 of the book; you need that for this week's homework.

4 What does this mean? What have we done?

Go back to the original system and plug in $x = -3$, $y = 2$, $z = 1$:

$$\begin{aligned} -3 + 2 \cdot 2 + 1 = 2 \quad &\longrightarrow \quad 2 = 2 \checkmark \\ 2 \cdot (-3) + 6 \cdot 2 + 1 = 7 \quad &\longrightarrow \quad 7 = 7 \checkmark \\ -3 + 2 + 4 \cdot 1 = 3 \quad &\longrightarrow \quad 3 = 3 \checkmark \end{aligned}$$

> This vector solves <u>each</u> equivalent system — every augmented matrix in the chain above represents the same solution set.

5 A system with infinitely many solutions

this is a lot like §3.2, Example 6, p. 139 in the book!

Example 5.1 (Underdetermined system)**.**

$$\begin{aligned} x - 3y + z &= 2 \\ 2x - 3y + 5z &= 4 \\ x - 2y + 2z &= 2 \end{aligned} \qquad \Longleftrightarrow \qquad \left[\begin{array}{ccc|c} 1 & -3 & 1 & 2 \\ 2 & -3 & 5 & 4 \\ 1 & -2 & 2 & 2 \end{array}\right]$$

Row-reducing to RREF gives

$$\Longleftrightarrow \qquad \left[\begin{array}{ccc|c} 1 & 0 & 4 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \longleftarrow \textbf{zero row}$$

The zero row means:

- z is unspecified — a **free variable**, like the "$+c$" before;
- the system is **under-determined**;
- there are ∞ many solutions.

Reading the rows from the bottom up:

$$\begin{aligned} &R_3 \iff 0 = 0, \\ &R_2 \iff y + z = 0 \iff y = -z, \\ &R_1 \iff x + 4z = 2 \iff x = 2 - 4z. \end{aligned}$$

So z can be any real number, but for a given z, the possible values of x and y are constrained by the above equations. So if there was an additional bit of information given (an initial condition, or something) that constrained z to a particular value, then we would immediately be able to solve for x and y. This is a lot like when we did non-homogeneous linear differential equations, and had $y = y_p + y_h$ where y_h was $Ce^{\int p dt}$. In general, that C means there are infinitely many possible solutions $y(t)$, but if we were given some initial condition that forced C to be some specific value, then there's now only one particular solution $y(t)$.

That analogy to chapter 2 is no coincidence. We can write the above general solution in the language of the non-homogeneous principle of solutions as follows:

On your own: verify that this is equivalent to the preceding expressions above, where we had $R_3 \Leftrightarrow 0 = 0$, etc.

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} + z \begin{bmatrix} -4 \\ -1 \\ 1 \end{bmatrix} = \vec{x}_p + \vec{x}_h.$$

Remark. Example 2.1 was also a non-homogeneous system of algebraic equations (i.e., it was of the form $\mathbf{A}\vec{x} = \vec{b}$ rather than $\mathbf{A}\vec{x} = \vec{0}$) and therefore the solution there can also be written as $\vec{x} = \vec{x}_p + \vec{x}_h$, except $\vec{x}_h$ was simply $\vec{0}$, hence why there was only 1 solution.

Note. If you get a row like

$$\left[\begin{array}{ccc|c} 0 & 0 & 0 & k \end{array}\right] \quad \text{with } k \neq 0,$$

then "$0 = k$" is a contradiction $\Longrightarrow$ **no solutions**; the system is **inconsistent**.