Lecture #3: Solving Linear Equations

usually, if you have a linear system you want to solve, can just use existing software

still, it is important to understand how you could do it by hand.

Today: How to make a linear system simpler while preserving its solution(s)

**Gaussian Elimination:** Adding / subtracting rows from each other

Let's do an example, given the linear system:

\[
\begin{align*}
    x - y + 2z &= 1 \\
    -2x + 2y - 3z &= -1 \\
    -3x - y + 2z &= -3
\end{align*}
\]

First put it in matrix-vector form
\[
\begin{bmatrix}
1 & -1 & 2 \\
-2 & 2 & -3 \\
-3 & -1 & 2
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z
\end{bmatrix}
= 
\begin{bmatrix}
1 \\
-1 \\
-3
\end{bmatrix}
\]

def: the augmented matrix \([A \ b]\)

i.e. what we get from appending \(b\) to the end

\[
\begin{bmatrix}
1 & -1 & 2 \\
-2 & 2 & -3 \\
-3 & -1 & 2
\end{bmatrix}
\begin{bmatrix}
1 \\
-1 \\
-3
\end{bmatrix}
\]

helpful to keep track of what the r.h.s. is

def: the pivot \(\square\) in a row is leftmost non-zero

Goal: Make the pivots go from left to right, strictly

Step #1: Move the pivot of 2\(^{nd}\) & 3\(^{rd}\) rows to the right by adding/subtracting 1\(^{st}\) row
e.g. \[
\begin{align*}
\underline{r_2} &= \begin{bmatrix} -2 & 2 & -3 & -1 \end{bmatrix} \\
+ 2\underline{r_1} &= 2\times\begin{bmatrix} 1 & -1 & 2 & 1 \end{bmatrix} \\
\underline{r_2}' &= \begin{bmatrix} 0 & 0 & 1 & 1 \end{bmatrix}
\end{align*}
\]

Poll: what multiple of \( r_1 \) should we add to \( r_3 \) to move its pivot to the right?

(a) 2  (b) -2  (c) 3  (d) -1

The new (augmented) matrix is

\[
\begin{bmatrix}
1 & -1 & 2 & 1 \\
0 & 0 & 1 & 1 \\
0 & -4 & 8 & 0
\end{bmatrix}
\]

Now you might be wondering: Why are we allowed to do these operations?

Because they preserve the set of all solutions
Let's think about this in our example. First, what is the new linear system?

\[
\begin{align*}
x - y + 2z &= 1 \\
\quad z &= 1 \\
-4y + 8z &= 0
\end{align*}
\]

**Note:** the augmented matrix is just bookkeeping.

Let's think about just one operation.

**Claim:** \(x, y, z\) satisfy:

\[
\begin{align*}
\begin{cases}
x - y + 2z &= 1 \\
-2x + 2y - 3z &= -1
\end{cases}
\end{align*}
\]  \(\text{(1)}\)

iff they satisfy

\[
\begin{align*}
\begin{cases}
x - y + 2z &= 1 \\
\quad z &= 1
\end{cases}
\end{align*}
\]  \(\text{(2)}\)
Let's actually prove this (always good to try to convince yourself of key facts)

\[x,y,z \quad x,y,z\]

Proof: \((1) \Rightarrow (2)\): This is true b/c we're just taking two equations \((r_1 \& r_2)\) we satisfy and adding them to get a new equation that we must also satisfy \((r'_2)\) \[\left[ r_1 \& r_2 \Rightarrow r_1 \& r_2 \right] \wedge (r'_2) \]

The more interesting direction is \((2) \Rightarrow (1)\): i.e. we are not creating new solutions when we replace \(r_2\) with \(r'_2\)

The key is: the steps are invertible

If we have a solution that satisfies \(r_1\) and \(r'_2\), we can derive \(r_2\) from \(r_1\) and \(r'_2\), so it's also satisfied.
The new linear system definitely looks simpler. But when should we stop?

**def.** An augment matrix is in row echelon form if all pivots are non-zero and go from left to right.

**Q2:** Is the new augmented matrix in ref?

No, and we'll need a new operation to fix it.

We can **swap rows**: (swap the 2\text{nd} and 3\text{rd} rows)

\[
\begin{bmatrix}
1 & -1 & 2 & 17 \\
0 & -4 & 8 & 0 \\
0 & 0 & 1 & 1
\end{bmatrix}
\]

upper triangular

**Q3:** why are we allowed this operation?

Again, swap on the rows exactly preserves the set of all solutions.
Now that we’re in ref finding a solution is easy by **back substitution**

\[ z = 1 \quad \text{(from 3rd equation)} \]

\[-4y + 8(i) = 0 \Rightarrow y = 2 \quad \text{(from 2nd equation)} \]

\[ x - (2i) + 2(i) = 1 \Rightarrow x = 1 \quad \text{(from 1st eqn)} \]

Now let’s connect this back to earlier lectures. In particular, I claimed

\[
\begin{bmatrix}
2 & -3 \\
1 & 1
\end{bmatrix}
\begin{bmatrix}
x \\
y
\end{bmatrix} =
\begin{bmatrix}
b_1 \\
b_2
\end{bmatrix}
\]

has a solution for any \( b_1 \) & \( b_2 \).

Q4: How does Gaussian elimination reveal this fact?

Let’s put it in ref:
\[
\begin{align*}
\mathbf{r}_2 &= \begin{bmatrix} 1 & 1 & b_2 \end{bmatrix} \\
-\frac{1}{2} \mathbf{r}_1 &= -\frac{1}{2} \begin{bmatrix} 2 & -3 & b_1 \end{bmatrix} \\
\mathbf{r}_2' &= \begin{bmatrix} 0 & \frac{5}{2} & b_2 - \frac{b_1}{2} \end{bmatrix}
\end{align*}
\]

So our equivalent linear system is:

\[
\begin{bmatrix} 2 & -3 \\ 0 & \frac{5}{2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 - \frac{b_1}{2} \end{bmatrix}
\]

For any \( b_1 \) and \( b_2 \) you give me, I can back-solve to find \( x \) and \( y \).

**Q5:** Is the solution always unique?

Yes, because the ops preserve the set of solutions, and back-substitution finds here is unique.

This is an example where understanding how to do things by hand can help — gives you insights even if you’re interested in a family of linear systems.
Some linear systems do not have a unique solution

e.g. \[
\begin{bmatrix}
2 & -3 & 0 \\
1 & 1 & 2 \\
3 & -2 & 2
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z
\end{bmatrix} =
\begin{bmatrix}
1 \\
1 \\
2
\end{bmatrix}
\]

what would its ref look like?

\[
\begin{bmatrix}
12 & -3 & 0 \\
0 & \frac{5}{2} & 2 \\
0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z
\end{bmatrix} =
\begin{bmatrix}
1 \\
\frac{1}{2} \\
0
\end{bmatrix}
\]

How does back substitution tell you all the solutions?

\[0 = 0 \quad \text{any choice of } z \text{ is fine!}\]

\[\frac{5}{2} y + 2 z = \frac{1}{2}\]

\[2x - 3y = 1\]

Aha! The set of solutions looks like a line in 3-d
def: A matrix where you get a row of all zeros in the matrix is called singular.

We saw above that square systems with a singular matrix A can have many solutions.

But they can also have no solutions.

\[
\begin{bmatrix}
2 & -3 & 0 \\
1 & 1 & 2 \\
3 & -2 & 2
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z
\end{bmatrix}
= 
\begin{bmatrix}
1 \\
1 \\
3
\end{bmatrix}
\]

Again, we are revisiting an example from before with new tools.

What does its ref look like?

\[
\begin{bmatrix}
2 & -3 & 0 \\
0 & \frac{5}{2} & 2 \\
0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z
\end{bmatrix}
= 
\begin{bmatrix}
1 \\
\frac{1}{2} \\
\frac{1}{3}
\end{bmatrix}
\]
Q6: How does this tell us there is no solution?

Again, via back substitution

\[ 0 = 1 \quad (3^{rd} \text{ equation}) \]

All the mileage we got out of rref is just the beginning:

Much easier to get geometric insights about a linear system by first putting it in a convenient normal form
On a related note:

**Gauss-Jordan Elimination:** Do even more work to make the linear system even simpler.

\[
\text{ref } \sim \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z \\
w
\end{bmatrix}
= \begin{bmatrix}
b_1 \\
b_2 \\
b_3 \\
b_4
\end{bmatrix}
\]

Every pivot is 1 and above it are only 0s (rref)

**Q2:** What does back substitution do?

\[z = b_3, \ y = b_2, \ x = b_1\]

You can't always get the identity matrix, sometimes have to settle for things like

\[
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 2 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z \\
w
\end{bmatrix}
= \begin{bmatrix}
b_1 \\
b_2 \\
b_3 \\
b_4
\end{bmatrix}
\]
Again, we just use row ops (add/subtract/swap).

In the remaining time, let's do an example, particularly of how to model problems you might need to solve as a linear system.

Ex: Electrical Circuit

We know the current across a resistor is

\[ \frac{v_1 - v_2}{R} = i \]

so we can write down a linear system that describes voltages \( \Rightarrow \) currents

\[
\begin{bmatrix}
1 & -1 & 0 & 0 & 0 & 0 \\
-1 & 3 & -1 & 0 & -1 & 0 \\
\vdots & \vdots & \ddots & \ddots & \ddots & \ddots \\
\end{bmatrix}
\begin{bmatrix}
v_1 \\
v_2 \\
\vdots \\
v_6 \\
\end{bmatrix}
= 
\begin{bmatrix}
\dot{v}_1 \\
\dot{v}_2 \\
\vdots \\
\dot{v}_6 \\
\end{bmatrix}
\]
where $i_1$ = net current out of junction #1

But we know what the current is supposed to be:

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & -1 \end{bmatrix} \text{ at all junctions except 18-6,}
\text{ net current should be zero.}$$

Hence solving linear system $\implies$ fully describing the behavior of the circuit.

End note: (for your psed) If $A$ is an $m \times n$ matrix, its transpose denoted $A^T$ is an $n \times m$ matrix with

$$(A^T)_{i,j} = A_{j,i}$$

e.g. $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}^T = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$

useful way to reshape a matrix.