GAUSSIAN ELIMINATION AND ITS APPLICATIONS

GLENN SUN
UCLA MATH CIRCLE ADVANCED 1

FEBRUARY 5, 2023

1 Introduction

From your school classes, you are likely familiar with how to solve some simple systems of equations. For example:

Problem 1. Solve the following systems of equations. (They may have more than one solution, in which case you should describe them, or they may have no solutions.)

1. $\begin{cases} x + 2y = -3 \\ 3x - y = 5 \end{cases}$

2. $\begin{cases} 2x - 3y = 4 \\ -4x + 6x = 2 \end{cases}$

3. $\begin{cases} 2x + y + 2z = 2 \\ 4x - 2y = 4 \\ -4x + y - z = -4 \end{cases}$

Solving systems of equations is important for many problems (you'll see a few in section 3), but as you saw, even 3 equations with 3 variables can be a little challenging sometimes. What if you need to solve 4 equations with 4 variables, or 5 or 10? What about the millions of equations that describe the latest AI models such as ChatGPT? We need an *algorithm* (a methodical way) to solve equations, so computers can help us do the work. This field of math is called *linear algebra*.

2 The Gaussian elimination algorithm

First, we will make our lives easier by writing less. We will fix an order for the variables, and then stop writing them. We will also draw one vertical line instead of writing many = signs. A grid of numbers is called a *matrix*. [1]

$$\begin{cases} 2x + y + 2z = 2 \\ 4x - 2y = 4 \\ -4x + y - z = -4 \end{cases} \longrightarrow \left[\begin{array}{ccc|c} 2 & 1 & 2 & 2 \\ 4 & -2 & 0 & 4 \\ -4 & 1 & -1 & -4 \end{array}\right]$$

The goal of the algorithm is to reach something that looks like:

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & * \\ 0 & 1 & 0 & * \\ 0 & 0 & 1 & * \end{array}\right] \longleftrightarrow \begin{cases} x = * \\ y = * \\ z = * \end{cases}$$

Of course, this is impossible if there are no solutions or more than one solution, but we'll see how close we can get. The algorithm relies on three observations:

1. One can switch two rows, and the solutions of the system will not change.
2. One can multiply (or divide) an entire row by any non-zero constant, and the solutions of the system will not change.
3. One can add (or subtract) a multiple of one row to (or from) any other row, and the solutions of the system will not change.

Problem 2. Briefly check that the above 3 observations are correct (especially observation #3).

Now, we use the above observations repeatedly to give an example of how to use the algorithm. Hopefully, the general case will be clear from the example.

[1] Some other sources call a matrix with a line an *augmented matrix*, we will not make this distinction.

Problem 3. In each step below, identify which of the 3 observations was used to make the deduction.

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

To check your understanding of the algorithm, also answer the following:

1. What would be the next steps if the last row was not all zeros?
2. Which observation did we not use? Can you think of a situation in which you might need to use it?
3. Interpret the final matrix above to get a description of all solutions to the system.

Problem 4. Use the Gaussian elimination algorithm to solve the other two problems from the introduction.

1. $\begin{cases} x + 2y = -3 \\ 3x - y = 5 \end{cases}$

2. $\begin{cases} 2x - 3y = 4 \\ -4x + 6x = 2 \end{cases}$

Most graphing calculators can perform this algorithm. The function is often called `rref`.

3 Applications

For every question in this section, you are *not required* to do the tedious calculation. (Of course, you can if you want.) Set up the matrix, ask your instructor to give you the output of the algorithm (or use your calculator), and interpret the output of the algorithm in context.

Problem 5. You may be familiar with the fact that $n + 1$ points uniquely determine a polynomial of degree at most n. Rephrase the following as a linear algebra problem: Find the cubic polynomial passing through the points $(-1, -2)$, $(0, 1)$, $(1, 0)$, and $(2, 1)$.

Problem 6. In nutrition studies, one would like to create a balanced diet using a variety of unbalanced foods, for example, for nutritional drinks like Soylent and Huel. Translate to linear algebra the following problem: Using the following fictional foods (shown is the nutrition per serving), find a combination that is 50% carbohydrates, 30% protein, and 20% fat. (One can imagine extending this to finer categories like fiber or other nutrients like sodium and vitamins.)

Food	Carbs	Protein	Fat
Tofu	20	60	20
Doritos	30	0	30
Rice	60	30	0

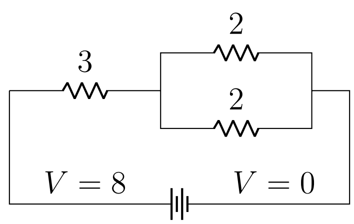
Problem 7. In chemistry, one often has to balance chemical equations (*stoichiometry*), i.e. given some molecules as reactants and products, figure out in what ratios these molecules are used and produced. In a molecule like H_2O, there are 2 atoms of hydrogen and 1 atom of oxygen. The number of atoms of each type must be the same on both sides of the equation. Rephrase the following stoichiometry problem as a linear algebra problem.

$$\_\, C_6H_{12}O_6 + \_\, O_2 \longrightarrow \_\, CO_2 + \_\, H_2O$$

Problem 8. (Harder) In physics and electrical engineering, a resistive electrical circuit consists of a battery (4 lines symbol) and resistors (zig-zag symbol) connected by wires (lines). One can measure the *voltage* V (a directionless quantity) and *current* I (a directed quantity) on any point of a wire. Voltage is constant on pieces of wire delimited by resistors and batteries, whereas current is constant on pieces of wire delimited by nodes (junctions of 3 or more wires). The current and voltage at any other point can be determined by the following rules:

1. (Ohm's law) Between the endpoints of any resistor with resistance R, we have $\Delta V = IR$. (ΔV means the difference in voltage between the endpoints, and the current I is constant between the endpoints, directed towards lower voltage.)
2. (Kirchoff's node law) At every junction, the directed sum of current is 0. ("Current in equals current out.")

Find the voltage and current at every point in the following circuit by transforming it into a linear algebra problem.



4 Vectors, multiplication, and inverses

There is a nice and concise way to represent what we have been doing so far. Given a matrix such as

$$\left[\begin{array}{ccc|c} 2 & 1 & 2 & 2 \\ 4 & -2 & 0 & 4 \\ -4 & 1 & -1 & -4 \end{array}\right],$$

call the square left side A (a matrix) and the right side b (a *vector*). For our purposes, a vector is a matrix that has only one column. When we solve this system, we also introduce the following notation: if x is the vector with entries x_1, x_2, and x_3, we say that we are looking solutions to $Ax = b$, or in other words

$$\begin{bmatrix} 2 & 1 & 2 \\ 4 & -2 & 0 \\ -4 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ -4 \end{bmatrix}.$$

Of course, this notation is a bit funny: one might expect to write the x_i's horizontally, since by the above we mean $2x_1 + x_2 + 2x_3 = 2$, etc., but we will see in minute why this is natural. Also, despite the fact that both addition and multiplication are happening here, we will just call this kind of operation *matrix-vector multiplication*. You should think of the word "multiplication" as just meaning "special operation" here.

Keep in mind: matrix vector multiplication is only defined with a matrix on the left and an (appropriately sized) vector on the right.

Problem 9. Compute the following matrix-vector products.

1. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix}$

2. $\begin{bmatrix} -2 & 1 & 0 \\ 0 & 3 & 2 \\ -1 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$

Problem 10. In the following, if it is tedious to calculate for $n \times n$ matrices, feel free to use 2×2 matrices and then just say how it generalizes.

1. What is the matrix I_n such that for all vectors x with n entries, $I_n x = x$? (This is called the identity matrix.)

2. It makes sense to add vectors, as long as they have the same number of entries: just add each component. Find the matrix C, such that $Cx = Ax + Bx$ for all x. The first step is provided.

   (We define the matrix sum $A + B$ to be C, so that we get the distributive law.)

   $$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} =$$

3. Find the matrix C, such that $Cx = A(Bx)$ for all x. The first step is provided.

   (We define the matrix product AB to be C, so that we get the associative law. Do you see now why it was natural to write x as a column?)

   $$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \left(\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) =$$

Problem 11. An $n \times n$ matrix A is *invertible* if there exists an $n \times n$ matrix A^{-1} such that $AA^{-1} = I_n$ and $A^{-1}A = I_n$. We call A^{-1} the inverse.

1. Using the fact that the system $2x - 3y = 4$, $-4x + 6x = 2$ has no solution, explain why the matrix
$$A = \begin{bmatrix} 2 & -3 \\ -4 & 6 \end{bmatrix}$$
is not invertible.

2. (Harder) Devise a method to use Gaussian elimination to compute the inverse of a matrix, if it exists. Then compute the inverse of
$$\begin{bmatrix} 1 & 3 \\ -2 & -5 \end{bmatrix}.$$

3. Explain why a matrix A is invertible if and only if for every vector b, the system $Ax = b$ has exactly one solution.

Problem 12. When talking about algorithms, it is important to know how fast algorithms are so that we can pick the best one, and we measure speed by counting the number of operations (addition/multiplication) used. In the below problems, feel free to use rough estimates.

1. Show that using Gaussian elimination to compute solutions to $Ax = b$ takes approximately some constant times n^3 operations.

2. Show that using Gaussian elimination to compute the inverse of a matrix takes approximately some constant times n^3 operations.

3. Show that matrix-vector multiplication takes approximately some constant times n^2 operations.

4. Given a list of n vectors $b^{(1)}, \ldots, b^{(n)}$, we wish to solve all of the systems $Ax^{(i)} = b^{(i)}$. Using the fastest method that you can think of, how many operations will this take?