Check the course website

`http://www.math.ust.hk/~emarberg/Math2121/`

for the syllabus and other course details.

1 Notation

Today's lecture corresponds to Section 1.1 in the textbook. See the book for a more detailed discussion!

Throughout, we'll using the following notation:

- $\mathbb{C}$ denotes the complex numbers $a + b\sqrt{-1}$.
- $\mathbb{R}$ denotes the real numbers.
- $\mathbb{Q}$ denotes the rational numbers p/q.
- $\mathbb{Z}$ denotes the integers $\{\dots, -2, -1, 0, 1, 2, \dots\}$.
- $\mathbb{N}$ denotes the nonnegative integers $\{0, 1, 2, \dots\}$.

Ellipsis ("…") notation: we write $a_1, a_2, \dots, a_7$ instead of the full list $a_1, a_2, a_3, a_4, a_5, a_6, a_7$.

We use the same convention to write $a_1, a_2, \dots, a_n$ even when n is a variable integer.

2 Systems of linear equations

Let $x_1, x_2, \dots, x_n$ be variables, where $n \geq 1$ is some integer.

Let $a_1, a_2, \dots, a_n, b$ be numbers in $\mathbb{R}$ (or $\mathbb{C}$).

(We'll usually work with real numbers, but nothing is any harder with complex numbers.)

Unlike in calculus, where our favorite variables are x, y, z, in linear algebra we prefer $x_1, x_2, x_3, \dots$. Reason: we want to go beyond 3 dimensions.

Definition. We refer to

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

as a *linear equation* in the variables $x_1, x_2, \dots, x_n$.

Notation. Another way of writing this equation is $\sum_{i=1}^n a_ix_i = b$.

The symbol "$\sum$" is the Greek letter sigma, for "sum."

There are many other equivalent ways of writing the same equation. For example:

$$\begin{aligned} a_1x_1 + a_2x_2 + \cdots + a_nx_n - b &= 0 \\ b &= a_1x_1 + a_2x_2 + \cdots + a_nx_n \\ a_1x_1 + a_3x_3 + a_5x_5 + \dots &= b - a_2x_2 - a_4x_4 - \dots \end{aligned}$$

We treat all of these equations as the same object.

A *system of linear equations* or *linear system* is a list of linear equations.

Example.

$$\begin{aligned} 2x_1 - x_2 + \sqrt{3}x_3 &= 8 \\ x_1 - 4x_4 &= 8 \\ x_2 &= 0 \end{aligned}$$

is a linear system in the variables x_1, x_2, x_3.

Definition. A *solution* of a linear system in variables $x_1, x_2, \ldots, x_n$ is a list of n numbers $(s_1, s_2, \ldots, s_n)$, with the property that if we plug in $x_1 = s_1$, $x_2 = s_2$, $\ldots$, $x_n = s_n$ in our equations, we get all true statements.

Note: if our system contains any false equations like "$0 = 1$", then it cannot have any solutions.

Two linear systems are *equivalent* if they have the same set of solutions.

Example. How many solutions can a linear system have?

1. The system
$$\begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 3x_2 &= 3 \end{aligned}$$
has one solution $(s_1, s_2) = (3, 2)$.
2. But the system
$$\begin{aligned} x_1 - 2x_2 &= -1 \\ 3x_1 - 6x_2 &= -3 \end{aligned}$$
has many solutions: $(s_1, s_2) = (1, 1)$ or $(3, 2)$ or $(5, 3)$ or $\ldots$.
3. Whereas the system
$$\begin{aligned} x_1 - 2x_2 &= -1 \\ x_1 - 2x_2 &= 0 \end{aligned}$$
has no solutions.

Theorem. A linear system in two variables x_1 and x_2 has either 0, 1, or ∞ solutions.

Remark. The symbol "∞" is pronounced "infinity." Saying a linear system has ∞ solutions is bad style, since ∞ isn't a number. When we say this, we really mean: "does not have finitely many solutions."

Proof by geometry. A solution to one equation $ax_1 + bx_2 = c$ represents a point on a line after we identify the pair of numbers (x_1, x_2) with a point in the Cartesian plane.

A solution to a system of 2-variable linear equations represents a point where the lines corresponding to the equations all intersect.

But a collection of lines all intersect either at 0 points (they don't have a common intersection), 1 point (the unique point of intersection) or at infinitely many points (in the case when the lines are all *the same line*, though they might come from different equations). □

Proof by algebra. Suppose the linear system has two different solutions (s_1, s_2) and (r_1, r_2).

Define $\lambda_1 = s_1 - r_1$ and $\lambda_2 = s_2 - r_2$.

The symbol "λ" is the Greek letter lambda.

If $ax_1 + bx_2 = c$ was one of the equations in our system, then by definition $as_1 + bs_2 = c$ and $ar_1 + br_2 = c$.

Taking the difference of these equations gives $a(s_1 - r_1) + b(s_2 - r_2) = 0$. In other words, $a\lambda_1 + b\lambda_2 = 0$.

It follows that $a(s_1 + z\lambda_1) + b(s_2 + z\lambda_2) = as_1 + bs_2 = c$ for all z.

This works for all the equations in our system.

Therefore $(s_1 + z\lambda_1, s_2 + z\lambda_2)$ is a new solution to our system, for any choice of z.

So the system has infinitely many solutions. □

A linear system is *consistent* if it has one or infinitely many solutions, and *inconsistent* if it has zero solutions. Both the algebraic and geometric proofs generalize to any number of variables. (Think about how to do this!) Therefore:

Theorem. A linear system in n variables is either consistent or inconsistent, i.e., has 0, 1, or infinitely many solutions.

3 Matrices

A *matrix* is just a rectangular array of numbers, like these ones:

$$\begin{bmatrix} 1 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 5 & 3 \\ 2 & \pi \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 7 & 6 & 4 & 3 \\ 2 & 1 & 1 & 0 \end{bmatrix}.$$

We denote a general matrix by

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \end{bmatrix}$$

Here "A_{23}" is pronounced "A, two, three". This matrix is 3-by-4: it has 3 rows and 4 columns.

Say that a matrix A is m-by-n or $m \times n$ if has m rows and n columns.

We usually write A_{ij} (pronounced "A, i, j") for the entry in the ith row and jth column of the matrix.

Matrices are useful as a compact way of writing a linear system.

Consider the linear system

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ 5x_1 - 5x_3 &= 10 \end{aligned}$$

Define the *coefficient matrix* of this system to be

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ 5 & 0 & -1 \end{bmatrix}$$

In other words, the matrix A where A_{ij} is the coefficient of x_j in the ith equation.

The *augmented matrix* of the system is

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -1 & 10 \end{bmatrix}.$$

Exercise: how would you generalize this definition to any linear system?

4 Solving linear systems

We solve linear systems by adding equations together to cancel variables.

Example. To solve

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ 5x_1 - 5x_3 &= 10 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -1 & 10 \end{bmatrix}$$

we first add -5 time eq. 1 to eq. 3 to get

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ 10x_2 - 10x_3 &= 10 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & 10 & -10 & 10 \end{bmatrix}.$$

We then multiply eq. 2 by 1/2 to get

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ x_2 - 4x_3 &= 4 \\ 10x_2 - 10x_3 &= 10 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 10 & -10 & 10 \end{bmatrix}.$$

We then add -10 times eq. 2 to eq. 3:

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ x_2 - 4x_3 &= 4 \\ 30x_3 &= -30 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 30 & -30 \end{bmatrix}.$$

Multiple eq. 3 by 1/30:

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

The argument matrix of the last system if *triangular*: all entries in positions (i, j) with $i > j$ are zero.

(Remember: i is the row, j is the column.)

We can easily solve for x_1, x_2, x_3 from a triangular system, working from the bottom up:

- The last equation $\boxed{x_3 = -1}$ is already as simple as possible.
- Substitute into second equation: $x_2 - 4x_3 = x_1 - 4(-1) = 4 \Rightarrow \boxed{x_2 = 0}$.
- Substitute into first equation: $x_1 - 2x_2 + x_1 = x_1 - 2(0) + (-1) = 0 \Rightarrow \boxed{x_1 = 1}$.

Definition. In solving this system of equations, we performed the following *(elementary) row operations* on the augmented matrix of the system:

1. Replacement: replace one row by the sum of itself and a multiple of another row.
2. Scaling: multiple all entries in a row by a *nonzero* number.
3. Interchange: swap two rows.

Note: we "add" rows by add the corresponding entries:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix} + \sqrt{7}\begin{bmatrix} 0 & 8 & 4 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 2+8\sqrt{7} & 3+4\sqrt{7} & 4+6\sqrt{7} \end{bmatrix}.$$

Two matrices are *row equivalent* if one can be transformed to the other by a sequence of row operations.

Note that each row operation is reversible.

Theorem. If the augmented matrices of two linear systems are row equivalent, then the systems are equivalent (i.e., have same solutions).

Proof. Here's the idea, minus the details: check that performing one row operation does not change whether a given $(s_1, s_2, \ldots, s_n)$ is a solution to the linear system. $\square$

Given a linear system with augmented matrix A, suppose we perform row operations to A until we get a matrix T with the property that whenever $T_{ij} \neq 0$ but $T_{i1} = T_{i2} = \cdots = T_{i,j-1} = 0$, it holds that $T_{i+1,j} = T_{i+2,j} = \cdots = T_{m,j} = 0$.

This means: if T_{ij} is the first nonzero entry in the ith row of T going left to right, then T_{ij} is the last nonzero entry in the jth column of T going top to bottom. For example:

$$T = \begin{bmatrix} 1 & 6 & 8 & 9 & 0 \\ 0 & 0 & 3 & 2 & 1 \\ 0 & 0 & 0 & 4 & 2 \end{bmatrix} \qquad \text{or} \qquad T = \begin{bmatrix} 1 & 6 & 8 & 9 & 0 \\ 0 & 0 & 3 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

From T in this form, we can easily determine if the system we started out with is consistent or inconsistent.

If T is the left matrix, the system is consistent: we have

$$x_4 = 4, \qquad 3x_3 + 2x_4 = 1, \qquad \text{and} \qquad x - 1 + 6x_2 + 8x_3 + 9x_4 = 0.$$

Exercise: find a solution!

If T is the right matrix, the system is inconsistent: it includes the equation $0 = 2$, from the last row.

In general, a linear system is inconsistent if and only if its augmented matrix can be transformed by row operations to a matrix with a row of the form

$$\begin{bmatrix} 0 & 0 & \ldots & 0 & q \end{bmatrix}$$

where $q \neq 0$. We'll prove this next time, after introducing the course's most important algorithm, *row reduction to echelon form*.