1111: Linear Algebra I

Dr. Vladimir Dotsenko (Vlad)

Lecture 5

Systems of linear equations

Geometrically, we are quite used to the fact that if we take two planes in 3D which are not parallel, their intersection is a line. With our new algebraic approach, this means that if we take a system of two equations

$$\begin{cases} A_1x + B_1y + C_1z = D_1, \\ A_2x + B_2y + C_2z = D_2, \end{cases}$$

for which the triples (A_1, B_1, C_1) and (A_2, B_2, C_2) are not proportional, then the solution set of this system can be described parametrically

$$\begin{cases} x = x_0 + at, \\ y = y_0 + bt, \\ z = z_0 + ct. \end{cases}$$

Systems of linear equations

A linear equation with unknowns $x_1, \ldots, x_n$ is an equation of the form

$$A_1x_1 + A_2x_2 + \cdots + A_nx_n = B,$$

where A_1, ..., A_n, and B are known numbers.
We shall develop a method for solving systems of m simultaneous linear equations with n unknowns

$$\begin{cases} A_{1,1}x_1 + A_{1,2}x_2 + \cdots + A_{1,n}x_n = B_1, \\ A_{2,1}x_1 + A_{2,2}x_2 + \cdots + A_{2,n}x_n = B_2, \\ \qquad\qquad\qquad\qquad\qquad\qquad \ldots \\ A_{m,1}x_1 + A_{m,2}x_2 + \cdots + A_{m,n}x_n = B_m, \end{cases}$$

where $A_{i,j}(1 \le i \le m, 1 \le j \le n)$, and $B_i(1 \le i \le m)$ are known numbers.
To save space, we shall often write A_{ij} instead of $A_{i,j}$, implicitly assuming the comma between i and j (and of course taking care to never multiply i by j in this context!)

Gauss–Jordan elimination

The most common technique for solving simultaneous systems of linear equations is *Gauss–Jordan elimination*. Anyone who ever tried to solve a system of linear equations probably did something of that sort, carefully eliminating one variable after another. We shall formulate this recipe in the form of an algorithm, that is a sequence of instructions that a person (or a computer) can perform mechanically, ending up with a solution to the given system.

For convenience, we shall not carry around the symbols representing the unknowns, and will encode the given system of linear equations by $m \times (n+1)$-matrix of coefficients

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} & B_1 \\ A_{21} & A_{22} & \cdots & A_{2n} & B_2 \\ \vdots & \dots & \ddots & \dots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} & B_m \end{pmatrix}$$

From equations to matrices

For example, if we consider the system of equations

$$\begin{cases} x_1 + 2x_2 + x_3 + x_4 + x_5 = 1, \\ -3x_1 - 6x_2 - 2x_3 - x_5 = -3, \\ 2x_1 + 4x_2 + 2x_3 + x_4 + 3x_5 = -3, \end{cases}$$

then the corresponding matrix is

$$A = \begin{pmatrix} 1 & 2 & 1 & 1 & 1 & 1 \\ -3 & -6 & -2 & 0 & -1 & -3 \\ 2 & 4 & 2 & 1 & 3 & -3 \end{pmatrix}$$

(note the zero entry that indicates that x_4 is not present in the second equation).

Elementary row operations

We define *elementary row operations* on matrices to be the following moves that transform a matrix into another matrix with the same number of rows and columns:

- *Swapping rows*: literally swap the row i and the row j for some $i \neq j$, keep all other rows (except for these two) intact.
- *Re-scaling rows*: multiply all entries in the row i by a nonzero number c, keep all other rows (except for the row i) intact.
- *Combining rows*: for some $i \neq j$, add to the row i the row j multiplied by some number c, keep all other rows (except for the row i) intact.

Let us remark that elementary row operations are clearly reversible: if the matrix B is obtained from the matrix A by elementary row operations, then the matrix A can be recovered back. Indeed, each individial row operation is manifestly reversible.

From the equations viewpoint, elementary row operations are simplest transformations that do not change the set of solutions, so we may hope to use them to simplify the system enough to be easily solved.

Gauss–Jordan elimination

Gauss–Jordan elimination is a certain pre-processing of a matrix by means of elementary row operations.

- **Step 1:** Find the smallest k for which $A_{ik} \neq 0$ for at least one i, that is, the smallest k for which the k^{th} column of the matrix A has a nonzero entry (this means x_k actually appears in at least one equation). Pick one such i, swap the first row of A with the i^{th} one, and pass the new matrix A to Step 2.
- **Step 2:** If $m = 1$, terminate. Otherwise, take the smallest number k for which $A_{1k} \neq 0$. For each $j = 2, \ldots, m$, subtract from the j^{th} row of A the first row multiplied by $\frac{A_{jk}}{A_{1k}}$. Divide the first row by A_{1k}. Finally, temporarily set aside the first row, and pass the matrix consisting of the last $m - 1$ rows as the new matrix A to Step 1.

Gauss–Jordan elimination

Once the procedure that we just described terminates, let us assemble together all the rows set aside along the way. The matrix A thus formed satisfies the following properties:

- For each row of A, either all entries of the row are equal to zero, or the first non-zero entry is equal to 1. (In this case we shall call that entry the *pivot* of that row).
- For each pivot of A, all entries in the same column below that pivot are equal to zero, as well as all entries below to the left.

A matrix satisfying these two conditions is said to be in *row echelon form*.

Gauss–Jordan elimination

To complete pre-processing, for each row s of A that has nonzero entries, we do the following: for each $r < s$, subtract from the row r the row s multiplied by A_{rt}, where t is the position of the pivot in the row s.
As a result, the matrix A obtained after this is done satisfies the following properties:

- For each row of A, either all entries of the row are equal to zero, or the first non-zero entry is equal to 1. (In this case we shall call that entry the *pivot* of that row).
- For each pivot of A, all other entries in the same column are equal to zero, as well as all entries below to the left.

A matrix satisfying these two conditions is said to be in *reduced row echelon form*. We proved an important theoretical result: every matrix can be transformed into a matrix in reduced row echelon form using elementary row operations.

Gauss–Jordan elimination: an example

Back to our example, we take the matrix we obtained and start transforming (writing down row operations to make it easy to check afterwards):

$$\begin{pmatrix} 1 & 2 & 1 & 1 & 1 & 1 \\ -3 & -6 & -2 & 0 & -1 & -3 \\ 2 & 4 & 2 & 1 & 3 & -3 \end{pmatrix} \overset{(2)+3(1),(3)-2(1)}{\mapsto}$$

$$\begin{pmatrix} 1 & 2 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 3 & 2 & 0 \\ 0 & 0 & 0 & -1 & 1 & -5 \end{pmatrix} \overset{(1)+(3),(2)+3(3),-1\times(3)}{\mapsto}$$

$$\begin{pmatrix} 1 & 2 & 1 & 0 & 2 & -4 \\ 0 & 0 & 1 & 0 & 5 & -15 \\ 0 & 0 & 0 & 1 & -1 & 5 \end{pmatrix} \overset{(1)-(2)}{\mapsto}$$

$$\begin{pmatrix} 1 & 2 & 0 & 0 & -3 & 11 \\ 0 & 0 & 1 & 0 & 5 & -15 \\ 0 & 0 & 0 & 1 & -1 & 5 \end{pmatrix}$$

Gauss–Jordan elimination

Now we can use our results to solve systems of linear equations. We restore the unknowns, and look at the resulting system of equations. This system can be investigated as follows.

If the last non-zero equation reads $0 = 1$, the system is clearly inconsistent.

If the pivot of last non-zero equation is a coefficient of some unknown, the system is consistent, and all solutions are easy to describe. For that, we shall separate unknowns into two groups, the *principal* (*pivotal*) unknowns, that is unknowns for which the coefficient in one of the equations is the pivot of that equation, and all the other ones, that we call *free* unknowns.

Once we assign arbitrary numeric values to free unknowns, each of the equations gives us the unique value of its pivotal unknown which makes the system consistent. Thus, we described the solution set in a parametric form using free unknowns as parameters.

Gauss–Jordan elimination: an example

Continuing with our example, the matrix

$$A = \begin{pmatrix} 1 & 2 & 0 & 0 & -3 & 11 \\ 0 & 0 & 1 & 0 & 5 & -15 \\ 0 & 0 & 0 & 1 & -1 & 5 \end{pmatrix}$$

is in reduced row echelon form. The corresponding system of equations is

$$\begin{cases} x_1 + 2x_2 - 3x_5 = 11, \\ x_3 + 5x_5 = -15, \\ x_4 - x_5 = 5. \end{cases}$$

The pivotal unknowns are x_1, x_3, and x_4, and the free unknowns are x_2 and x_5. Assigning arbitrary parameters $x_2 := t_2$ and $x_5 := t_5$ to the free unknowns, we obtain the following description of the solution set:

$$x_1 = 11 - 2t_2 + 3t_5, x_2 = t_2, x_3 = -15 - 5t_5, x_4 = 5 + t_5, x_5 = t_5.$$

where t_2 and t_5 are arbitrary numbers.