Mate1009 Algebra — Lecture 3

Mārtiņš Kālis
Notes partially adapted from
K. R. Mathews "Elementary linear algebra"

24 February 2021

1 Systems of linear equations

1.5 Homogeneous systems

A system of homogeneous linear equations is a system in the form (notice the zeros on the right-hand side):

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0
\end{aligned}$$

Such a system is always consistent as $x_1 = x_2 = \cdots = x_n = 0$ is a solution. This solution is called the *trivial* solution. Any other solution is called a *non–trivial* solution.

For example the homogeneous system

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

has only the trivial solution, whereas the homogeneous system

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

has the complete solution $x = -z$, $y = 0$, z arbitrary. In particular, taking $z = 1$ gives the non–trivial solution $x = -1, y = 0, z = 1$.

There is simple but fundamental theorem concerning homogeneous systems.

Theorem 1.5.1. *A homogeneous system of m linear equations in n variables always has a non–trivial solution if $m < n$.*

Proof. Suppose that $m < n$ and that the coefficient matrix of the system is row–equivalent to B, a matrix in reduced row–echelon form. Let r be the number of non–zero rows in B. Then $r \leq m < n$ and hence $n - r > 0$ and so the number $n - r$ of arbitrary variables is in fact positive. Taking one of these unknowns to be 1 gives a non–trivial solution. □

Remark. Let two systems of homogeneous equations in n variables have coefficient matrices A and B, respectively. If each row of B is a linear combination of the rows of A (i.e. a sum of multiples of the rows of A) and each row of A is a linear combination of the rows of B, then it is easy to prove that the two systems have identical solutions. The converse is true, but is not easy to prove. Similarly if A and B have the same reduced row–echelon form, apart from possibly zero rows, then the two systems have identical solutions and conversely.

There is a similar situation in the case of two systems of linear equations (not necessarily homogeneous), with the proviso that in the statement of the converse, the extra condition that both the systems are consistent, is needed.

Problem 1.5.1. Solve the homogeneous system

$$\begin{array}{lcl} -3x_1 + x_2 + x_3 + x_4 & = & 0 \\ x_1 - 3x_2 + x_3 + x_4 & = & 0 \\ x_1 + x_2 - 3x_3 + x_4 & = & 0 \\ x_1 + x_2 + x_3 - 3x_4 & = & 0 \end{array}$$

Problem 1.5.2. Solve the homogeneous system

$$\begin{array}{lcl} 3x_1 + x_2 + x_3 + x_4 & = & 0 \\ 5x_1 - x_2 + x_3 - x_4 & = & 0 \end{array}$$

2 Matrices

2.1 Matrix arithmetic

A matrix is a rectangular array of (rational, real, complex, etc.) numbers. Unless otherwise specified, in this course you can assume these are real numbers. The symbol $M_{m\times n}$ denotes the collection of all $m \times n$ matrices. Matrices will usually be denoted by capital letters and the equation $A = [a_{ij}]$ means that the element in the i–th row and j–th column of the matrix A equals a_{ij}. It is also occasionally convenient to write $a_{ij} = (A)_{ij} = A_{ij}$.

Example 2.1.1. The formula $a_{ij} = 1/(i + j)$ for $1 \leq i \leq 3, 1 \leq j \leq 4$ defines a 3×4 matrix $A = [a_{ij}]$, namely

$$A = \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{bmatrix}$$

Definition 2.1.1 (Equality of matrices)**.** Matrices A and B are said to be equal if A and B have the same size and their corresponding elements are equal; i.e., $A, B \in M_{m\times n}$ and $A = [a_{ij}], B = [b_{ij}]$, with $a_{ij} = b_{ij}$ for $1 \leq i \leq m, 1 \leq j \leq n$.

Definition 2.1.2 (Addition of matrices)**.** Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be of the same size. Then $A + B$ is the matrix obtained by adding the corresponding elements of A and B; that is

$$A + B = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}].$$

Definition 2.1.3 (Scalar multiple of a matrix)**.** Let $A = [a_{ij}]$ and t — a scalar, e.g. $t \in \mathbb{R}$. Then tA is the matrix obtained by multiplying all elements of A by t; that is

$$tA = t[a_{ij}] = [ta_{ij}].$$

Definition 2.1.4 (Additive inverse of a matrix)**.** Let $A = [a_{ij}]$. Then $-A$ is the matrix obtained by replacing the elements of A by their additive inverses; that is

$$-A = -[a_{ij}] = [-a_{ij}].$$

Definition 2.1.5 (Subtraction of matrices)**.** Matrix subtraction is defined for two matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ of the same size, in the usual way; that is

$$A - B = [a_{ij}] - [b_{ij}] = [a_{ij} - b_{ij}].$$

Definition 2.1.6 (The zero matrix)**.** For each m, n the matrix in $M_{m\times n}$, all of whose elements are zero, is called the zero matrix (of size $m \times n$) and is denoted by the symbol 0.

The matrix operations of addition, scalar multiplication, additive inverse and subtraction satisfy the usual laws of arithmetic. (In what follows, s and t will be arbitrary scalars and A, B, C are matrices of the same size.)

1. $(A + B) + C = A + (B + C)$;
2. $A + B = B + A$;
3. $0 + A = A$;
4. $A + (-A) = 0$;
5. $(s + t)A = sA + tA, (s - t)A = sA - tA$;
6. $t(A + B) = tA + tB, t(A - B) = tA - tB$;
7. $s(tA) = (st)A$;
8. $1A = A, 0A = 0, (-1)A = -A$;
9. $tA = 0 \Rightarrow t = 0$ or $A = 0$.

Definition 2.1.7. Let $A = [a_{ij}]$ be a matrix of size $m \times n$ and $B = [b_{jk}]$ be a matrix of size $n \times p$; (that is the number of columns of A equals the number of rows of B). Then AB is the $m \times p$ matrix $C = [c_{ik}]$ whose (i, k)–th element is defined by the formula

$$c_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk} = a_{i1}b_{1k} + \cdots + a_{in}b_{nk}.$$

Example 2.1.2.

1. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 1\cdot 5 + 2\cdot 7 & 1\cdot 6 + 2\cdot 8 \\ 3\cdot 5 + 4\cdot 7 & 3\cdot 6 + 4\cdot 8 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}.$

2. $\begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 23 & 34 \\ 31 & 46 \end{bmatrix} \neq \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}.$

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

4. $\begin{bmatrix} 3 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 11 \end{bmatrix}$ (this will often be simplified to $\begin{bmatrix} 11 \end{bmatrix} = 11$.

5. $\begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0.$

Matrix multiplication obeys many of the familiar laws of arithmetic apart from the commutative law.

1. $(AB)C = A(BC)$if $A \in M_{m\times n}, B \in M_{n\times p}, C \in M_{p\times q}$;
2. $t(AB) = (tA)B = A(tB), A(-B) = (-A)B = -(AB)$;
3. $(A+B)C = AC + BC$ if $A, B \in M_{m\times n}, C \in M_{n\times p}$;
4. $D(A+B) = DA + DB$ if $A, B \in M_{m\times n}, D \in M_{p\times m}$.

We prove the associative law only: First observe that $(AB)C$ and $A(BC)$ are both of size $m \times q$. Let $A = [a_{ij}], B = [b_{jk}], C = [c_{kl}]$. Then

$$\begin{aligned} ((AB)C)_{il} &= \sum_{k=1}^{p} (AB)_{ik}c_{kl} = \sum_{k=1}^{p}\left(\sum_{j=1}^{n} a_{ij}b_{jk}\right)c_{kl} \\ &= \sum_{k=1}^{p}\sum_{j=1}^{n} a_{ij}b_{jk}c_{kl} \end{aligned}$$

Similarly

$$(A(BC))_{il} = \sum_{k=1}^{n}\sum_{k=1}^{p} a_{ij}b_{jk}c_{kl}$$

The double summations are equal, since $\sum_{j=1}^{n}\sum_{k=1}^{p} d_{jk}$ and $\sum_{k=1}^{p}\sum_{j=1}^{n} d_{jk}$ represent the sum of the np elements of the rectangular array $[d_{jk}]$ by rows and by columns respectively. Thus $((AB)C)_{il} = (A(BC))_{il}$ for $1 \leq i \leq m, 1 \leq l \leq q$, hence $(AB)C = A(BC)$.

The system of m linear equations in n variables

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

is equivalent to a single matrix equation

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

that is $AX = B$, where $A = [a_{ij}]$ is the *coefficient matrix* of the system, $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is the *vector of variables* and $B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$ is the *vector of constants*. Sometimes X and B are written as $\vec{x}$ and $\vec{b}$ or $\mathbf{x}$ and $\mathbf{b}$ to differentiate them form multi-column matrices and scalars. Column vectors such as $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and $\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$ are sometimes written as $[x_1, x_2, \dots, x_n]$ and $[b_1, b_2, \dots, b_m]$ or $\begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T$ and $\begin{bmatrix} b_1 & b_2 & \dots & b_m \end{bmatrix}^T$ in text, where T stands for "transpose".

Another useful matrix equation equivalent to the above system of linear equations is

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \dots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \dots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Example 2.1.3. The system

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

is equivalent to the matrix equation

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

and to the equation

$$x \begin{bmatrix} 1 \\ 1 \end{bmatrix} + y \begin{bmatrix} 1 \\ -1 \end{bmatrix} + z \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Such an approach allows us to save work when solving $A\mathbf{x} = \mathbf{b}$ for a fixed matrix / system A and various values for $\mathbf{b}$. You can see how this is implemented in Python on git.

Problem 2.1.1. Let A, B, C, D be matrices defined by

$$A = \begin{bmatrix} 3 & 0 \\ -1 & 2 \\ 1 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 5 & 2 \\ -1 & 1 & 0 \\ -4 & 1 & 3 \end{bmatrix}, C = \begin{bmatrix} -3 & -1 \\ 2 & 1 \\ 4 & 3 \end{bmatrix}, D = \begin{bmatrix} 4 & -1 \\ 2 & 0 \end{bmatrix}$$

.

Which of the following matrices are defined? Compute those matrices which are defined.

$$A + B, A + C, AB, BA, CD, DC, D^2.$$