Linear Algebra & Differential Equations

Week 1 Notes

Elif Uskuplu

If you see any mistake, please email me (euskuplu@usc.edu).

---

1 Week 1

1.1 Synopsis of the course

This course is about the algebra and analysis of linear equations like

$$\begin{aligned} 2x + y &= 5 \\ x + 3y &= 10. \end{aligned} \tag{1}$$

During the lectures, we will build many tools to solve such equations and advanced ones. How such a course is related to differential equations? When such equations involve derivatives, for example,

$$\begin{aligned} x' - 6y &= 0 \\ y' - x - y &= 0, \end{aligned}$$

where x, y are functions, our tools, we build during linear algebra part, will solve these differential equations.

We can regard the coefficients in the system (1) as a rectangular array of numbers:

$$\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \tag{2}$$

we call this array a *matrix*. We will learn basics and fundamentals of matrices.

We can also regard the system (1) as a linear combinations of *vectors*:

$$x \begin{bmatrix} 2 \\ 1 \end{bmatrix} + y \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 5 \\ 10 \end{bmatrix}. \tag{3}$$

We will understand *vector spaces* to build the tools for solving linear equations.

A simple approach to solve (1) would be taking double of the second row and subtracting the first row from it to get $5y = 5$ so $y = 1$, and hence $x = 2$. This simple technique will be applied by *reducing operations* on the matrix (2), and we will see more advanced versions of reducing operations during the course.

When a system of equations is too complex, we expect to simplify it. This means that we should be able to *turn a system to another* which is done by *linear transformations*, another major concept in the course.

Another way to solve linear equations involves playing with the matrix form like (2). We'll learn how (and when) to make *inverse of a matrix* and how to *decompose* a matrix into simpler ones.

1.2 Matrices

Definition 1.1. *An $m \times n$* ***matrix*** *is a rectangular array of numbers in m rows and n columns.*

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}_{m \times n} \tag{4}$$

The number $m \times n$ is called ***the size*** *(or dimension) of A. When $m = 1$, A is called a* ***row vector****, and when $n = 1$, A is called a* ***column vector****.*

Taking a_{ij} as the entry at ith row and jth column, we can write $A = [a_{ij}]$ in short. In (4), if $\mathbf{r}_i$ is the ith row for $1 \leq i \leq m$, and $\mathbf{c}_j$ is the jth column for $1 \leq j \leq n$, we can also write A as below. The first is *row representation*, and the second is *column representation*.

$$A = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_m \end{bmatrix} = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \dots & \mathbf{c}_n \end{bmatrix}$$

Remark. If $A = [a_{ij}]$ is an $m \times n$ matrix and $B = [b_{ij}]$ is an $r \times s$ matrix, then $A = B$ if and only if $m = r$, $n = s$, and $a_{ij} = b_{ij}$ for each index. In other words, equal matrices have the same sizes and the same entries at each index.

Definition 1.2. *The* ***transpose*** *of a matrix is a flipped version of the original matrix. We can transpose a matrix by switching its rows with its columns. If $A = [a_{ij}]$ is an $m \times n$ matrix, its transpose, denoted by A^T, is an $n \times m$ matrix such that $A^T = [c_{ij}]$ where $c_{ij} = a_{ji}$.*

Example.

$$A = \begin{bmatrix} 4 & 3 & 7 \\ 12 & 0 & 5 \end{bmatrix} \quad A^T = \begin{bmatrix} 4 & 12 \\ 3 & 0 \\ 7 & 5 \end{bmatrix}$$

Example. The transpose of a row vector is a column vector, and vice versa.

Definition 1.3. *If A is a matrix of size $n \times n$, namely, # of rows = # of columns, then A is a* ***square*** *matrix. The following notions make sense* <u>*only*</u> *for such a square matrices. Let $A = [a_{ij}]$ be a square matrix of size $n \times n$:*

- *The elements $\{a_{11}, a_{22}, \ldots, a_{nn}\}$ form the* ***diagonal*** *of A.*
- *The* ***trace*** *of A is the sum of entries on the diagonal, namely, $Tr(A) = a_{11} + a_{22} + \ldots + a_{nn}$.*
- *If $a_{ij} = 0$ for $i < j$, then A is called* ***lower triangular matrix****.*
- *If $a_{ij} = 0$ for $i > j$, then A is called* ***upper triangular matrix****.*
- *If $a_{ij} = 0$ for $i = j$, then A is called* ***diagonal matrix****.*
- *If $A^T = A$, then A is called* ***symmetric matrix.***
- *If $A^T = -A$, then A is called* ***anti(skew)-symmetric matrix****. Here we mean $-A = [-a_{ij}]$.*

Example. The following are lower triangular, upper triangular, diagonal, symmetric, and anti-symmetric, respectively:

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 5 & 0 \\ 7 & 10 & 11 \end{bmatrix}, \begin{bmatrix} 8 & 2 & 3 \\ 0 & 5 & 4 \\ 0 & 0 & 9 \end{bmatrix}, \begin{bmatrix} 42 & 0 & 0 \\ 0 & 17 & 0 \\ 0 & 0 & 100 \end{bmatrix}, \begin{bmatrix} 1 & 4 & 5 \\ 4 & 2 & 6 \\ 5 & 6 & 3 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 3 \\ -2 & -3 & 0 \end{bmatrix}$$

Remark. The diagonal entries of an anti-symmetric matrix cannot be nonzero.

Definition 1.4. *The* ***zero matrix*** *of size $m \times n$ is a matrix all of whose entries are zero. It is also called* ***null matrix****. We denote it by $\mathbf{0}_{m\times n}$.*

1.3 Matrix addition and scalar multiplication

Definition 1.5. *Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be two $m \times n$ matrices, and s be a scalar(number). We define the addition of A and B, and the scalar multiple of A as follows:*

$$A + B := [a_{ij} + b_{ij}] \qquad\qquad sA := [s(a_{ij})]$$

Then we also have subtraction defined as $A - B := A + (-1)B = [a_{ij} - b_{ij}]$.

If two matrices have different sizes, we **cannot** add them. These operations have the following properties:

- $A + B = B + A$,
- $A + (B + C) = (A + B) + C$,
- $A + \mathbf{0} = A$, $1A = A$,

- $s(A+B) = sA + sB$,
- $(s+t)A = sA + tA$,
- $s(tA) = (st)A = (ts)A = t(sA)$.

Proof. Exercise. Just use $A = [a_{ij}]$ notation for each case. $\square$

1.4 Matrix multiplication

Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be two $m \times n$ matrices. You may think that the product matrix AB is obtained by like the addition, namely, the componentwise multiplication $AB = [a_{ij}b_{ij}]$. Although this is a valid operation, this is not the multiplication operation we deal with. The reason will become clear after we cover *linear transformations*. Let's define the right way of multiplication. We define it in three steps:

Definition 1.6. *(Step 1) Let* $A = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix}$ *be a row vector and* $B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$ *be a column vector. Then the* ***dot product*** *of* A, B *is a scalar (number), denoted by* $A \cdot B$*, defined by*

$$A \cdot B := a_1b_1 + a_2b_2 + \ldots + a_nb_n.$$

Example. $\begin{bmatrix} 2 & 4 & 5 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} = 6 + 4 + 0 = 10.$

Definition 1.7. *(Step 2) Let* $A = \begin{bmatrix} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \\ \vdots \\ \boldsymbol{r}_m \end{bmatrix}$ *be an* $m \times n$ *matrix, namely,* $\boldsymbol{r}_i$*'s are* n*-row vectors. Let* $B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$ *be a column vector. Then the product of* A *and* B *is an* m*-column vector (*$m \times 1$ *matrix) defined by*

$$AB := \begin{bmatrix} \boldsymbol{r}_1 \cdot B \\ \boldsymbol{r}_2 \cdot B \\ \vdots \\ \boldsymbol{r}_m \cdot B \end{bmatrix}.$$

Example. $\begin{bmatrix} 2 & 4 & 7 \\ 5 & 12 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 2 & 4 & 7 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix} \\ \begin{bmatrix} 5 & 12 & 1 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 6+24+14 \\ 15+72+2 \end{bmatrix} = \begin{bmatrix} 44 \\ 89 \end{bmatrix}.$

Remark. In Definition 3.2, if we write $A = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \dots & \mathbf{c}_n \end{bmatrix}$ where $\mathbf{c}_i$'s are m-column vectors, and $B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$, we have

$$AB = b1\mathbf{c}_1 + b_2\mathbf{c}_2 + \ldots + b_n\mathbf{c}_m.$$

Example. $\begin{bmatrix} 2 & 4 & 7 \\ 5 & 12 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix} = 3 \begin{bmatrix} 2 \\ 5 \end{bmatrix} + 6 \begin{bmatrix} 4 \\ 12 \end{bmatrix} + 2 \begin{bmatrix} 7 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 15 \end{bmatrix} + \begin{bmatrix} 24 \\ 72 \end{bmatrix} + \begin{bmatrix} 14 \\ 2 \end{bmatrix} = \begin{bmatrix} 44 \\ 89 \end{bmatrix}.$

Definition 1.8. *(Step 3) Let $A = \begin{bmatrix} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \\ \vdots \\ \boldsymbol{r}_m \end{bmatrix}$ be an $m \times n$ matrix and $B = \begin{bmatrix} \boldsymbol{c}_1 & \boldsymbol{c}_2 & \dots & \boldsymbol{c}_p \end{bmatrix}$ be an $n \times p$ matrix. Then the product of A and B is an $m \times p$ matrix defined by $AB := [c_{ij}]$ such that $c_{ij} := \boldsymbol{r}_i \cdot \boldsymbol{c}_j$ for $1 \le i \le m$ and $1 \le j \le p$. So in the full expansion, we have*

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{np} \end{bmatrix}$$
$$= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + \ldots + a_{1n}b_{n1} & \dots & a_{11}b_{1p} + a_{12}b_{2p} + \ldots + a_{1n}b_{np} \\ a_{21}b_{11} + a_{22}b_{21} + \ldots + a_{2n}b_{n1} & \dots & a_{21}b_{1p} + a_{22}b_{2p} + \ldots + a_{2n}b_{np} \\ \vdots & \ddots & \vdots \\ a_{m1}b_{11} + a_{m2}b_{21} + \ldots + a_{mn}b_{n1} & \dots & a_{m1}b_{1p} + a_{m2}b_{2p} + \ldots + a_{mn}b_{np} \end{bmatrix}$$

.

In other words, the entries of AB is given by $c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$ for $1 \le i \le m$ and $1 \le j \le p$.

Remark. In matrix multiplication, BE AWARE of the dimensions. In order to multiply A and B, we must have

# of columns of A = # of rows of B.

Example. $\begin{bmatrix} 2 & 4 & 7 \\ 5 & 12 & 1 \end{bmatrix} \begin{bmatrix} 3 & 3 \\ 7 & 1 \\ 5 & 9 \end{bmatrix} = \begin{bmatrix} 6+28+35 & 6+4+63 \\ 15+84+5 & 15+12+9 \end{bmatrix} = \begin{bmatrix} 69 & 73 \\ 104 & 36 \end{bmatrix}$.

Definition 1.9. *The* ***identity matrix****, denoted by I_n, is an $n \times n$ matrix such that all diagonal entries are 1, the other entries are 0.*

The following are properties of multiplication operation:

- $A(BC) = (AB)C$,
- $A(B + C) = AB + AC$,
- $(A + B)C = AC + BC$,
- $AB \neq BA$ in general !!!
- $A_{m\times n}I_n = A$,
- $I_m A_{m\times n} = A$.

Also, the following are the properties of transpose with operations:

- $(A^T)^T = A$,
- $(A + B)^T = A^T + B^T$,
- $(AB)^T = B^T A^T$.

Proof. We'll prove some of them, the others are exercises.

- $(A^T)^T = A$

  Let $A = [a_{ij}]$ be an $m \times n$ matrix, then A^T has size $n \times m$, so $(A^T)^T$ has dimension $m \times n$. Thus, $(A^T)^T$ and A have the same dimensions. Also, $A^T = [(a_{ij})^T]$, and hence $(A^T)^T = [((a_{ij})^T)^T]$. Observe that $((a_{ij})^T)^T = (a_{ji})^T = a_{ji}$. Thus, $(A^T)^T$ and A have the same entries.

- $(AB)^T = B^T A^T$

  Let $A = [a_{ij}]$ be an $m \times n$ matrix, and $B = [b_{ij}]$ be an $n \times p$ matrix. Then AB is an $m \times p$ matrix, so $(AB)^T$ is a $p \times m$ matrix. On the other hand, B^T is a $p \times n$ matrix and A^T is an $n \times m$ matrix, so $B^T A^T$ is a $p \times m$ matrix. Thus, $(AB)^T$ and $B^T A^T$ have the same dimensions.

  Let $1 \leq i \leq p$ and $1 \leq j \leq m$. We want to show that the ijth entries of $(AB)^T$ and $B^T A^T$ are equal. Now, we have the following:

  1. $((AB)^T)_{ij} = (AB)_{ji} = \sum_{k=1}^{n} a_{jk} b_{ki}$

2. $(B^TA^T)_{ij} = \sum_{k=1}^{n}(B^T)_{ik}(A^T)_{kj} = \sum_{k=1}^{n}(B)_{ki}(A)_{jk} = \sum_{k=1}^{n} b_{ki}a_{jk}$

Clearly, last expressions on both are the same. So we have $(AB)^T = B^TA^T$.

□

Proposition 1.1. *The product of two lower (upper) triangular matrices is a lower (upper) triangular matrix.*

Proof. Exercise. This is Theorem 2.2.24, as presented in the textbook. While the proof is available for reading, it is advisable to attempt solving it on your own first. □

1.5 Linear equations

Definition 1.10. *An $m \times n$ **system of linear equations** is the list of m equations with n variables:*

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

*We call a_{ij}'s **coefficients**, x_i's **variables (unknowns)**, and b_i's **constants**. If $b_i = 0$ for all i, then the system is called **homogeneous**. Otherwise, the system is **nonhomogeneous**.*

*We say an n-tuple $(c_1, c_2, \ldots, c_n)$ is a **solution** for the system if this tuple satisfies each equations. If a system has at least one solution, it is called **consistent**. Otherwise, it is **inconsistent**.*

Example. The following system is consistent because $(x = 1, y = 3)$ gives a solution.

$$\begin{aligned}
2x + y &= 5 \\
-3x + 6y &= 15
\end{aligned}$$

Remark. We can write such a system using matrices. The **matrix of coefficients** of the system is given by

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}.$$

The **augmented matrix** of the system is given by

$$A^{\#} = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right].$$

Then we write such a system as follows, $Ax = b$ where:

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

This is called **vector formulation** of a system of linear equations.

During the lectures, we'll answer such questions:

- Does a system have a solution?
- If yes, then how many solutions are there?
- How we determine all the solutions?

For those who are interested in the geometry behind small cases, the following videos provide, maybe too cartoonish :), but good references:

* Visualizing Linear Equations in Three Variables

`https://www.youtube.com/watch?v=Wm27Y6hxbRs`

* Types of Linear Systems in Three Variables

`https://www.youtube.com/watch?v=WAzUwzV1F3g`

1.6 Questions from the discussion sessions

The questions 2.1.10, 11, 18, 20, 22-27 from the textbook[1]

Some parts of the question 2.2.3 from the textbook

QUIZ : Posted on Blackboard and the course webpage.

[1]Differential Equations and Linear Algebra by Stephen Goode & Scott Annin