MAS109: Introduction to Linear Algebra **Fall 2019**

Lecture 2: September 5

Lecturer: Seonjeong Park (psjeong@kaist.ac.kr)

Text Book: *Contemporary Linear Algebra, H. Anton and R.C. Busby, John Wiley & Sons, Inc.*

Review

In this course, we will study various ways to find a solution set of a linear system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{2.1}$$

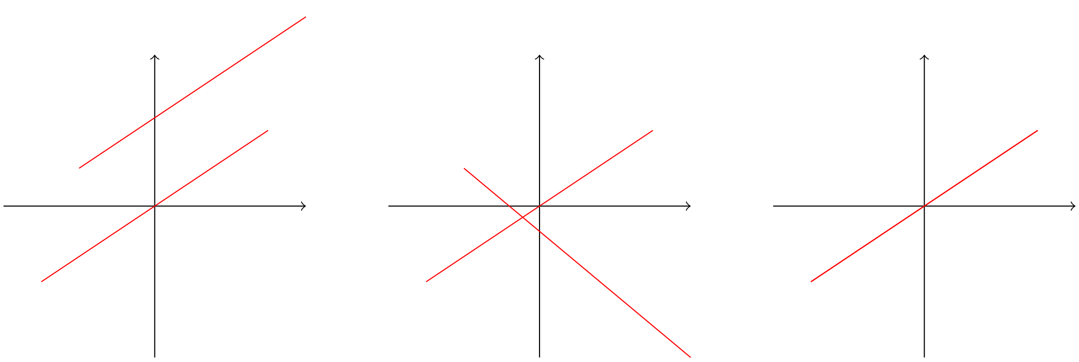
Note that if $b_1 = \cdots = b_m = 0$, then (2.1) is called a *homogeneous linear system*.

Definition 2.1 *A linear system is* consistent *(respectively, inconsistent) if it has at least one solution (respectively, if it has no solution).*

We will prove the following theorem in Section 3.5.

Theorem 2.2 *Every system of linear equations has zero, one, or infinitely many solutions; there are no other possibilities.*

For example, when $n = 2$, there are three possible cases.



(a) parallel (b) Intersecting at a point (c) coincident

The *augmented matrix* of the system (2.1) is

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_n \end{bmatrix}. \tag{2.2}$$

How can we solve a linear system?

1. Find the augmented matrix A of the linear system.
2. By using the elementary row operations, reduce A to the reduced row echelon form (respectively, row echelon form).
3. Then we can obtain a general solution of the system.

We call this algorithm *Gauss-Jordan elimination* (respectively, *Gaussian elimination*).

Quiz. Consider the following matrices.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

1. How many matrices are in reduced row echelon form?
2. How many matrices are in row echelon form?

How can we find the solution set using Gaussian elimination?

Consider the linear system corresponding to the matrix $\begin{bmatrix} 1 & 1 & 2 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$. Then we have

$$\begin{cases} x + y + 2z = 1 \\ y + z = 1 \\ z = 1. \end{cases}$$

Substituting $z = 1$ into $y + z = 1$, we get $y = 0$. Substituting $y = 0$ and $z = 1$ into $x+y+2z = 1$, we get $x = -1$. This procedure is called the "back substitution". Comment: In fact, we did not cover the above example on Gaussian elimination. Please read it individually.

Now we consider a homogeneous linear system:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases} \tag{2.3}$$

Note that every homogeneous linear system is consistent since

$$x_1 = 0, x_2 = 0, \ldots, x_n = 0$$

is a solution of (2.3). This is called the *trivial solution*. All other solutions, if any, are called *nontrivial solutions*. If (2.3) has a nontrivial solution

$$x_1 = s_1, \ldots, x_n = s_n,$$

then

$$x_1 = ks_1, \ldots, x_n = ks_n \qquad \forall k \neq 0$$

is also a solution of (2.3). Therefore, we get the following.

Theorem 2.3 *A homogeneous linear system has only the trivial solution or infinitely many solutions; there are no other possibilities.*

The following two theorems do not hold in general, but they are always true for homogeneous linear systems.

Theorem 2.4 *If the reduced row echelon form has r nonzero rows, then the system has $n-r$ free variables.*

Theorem 2.5 *If $m < n$, then the system has infinitely many solutions.*

Remark (Comment: a further explanation on the above theorems.) Let $[A \mid \mathbf{b}]$ be the augmented matrix of a linear system with m equations and n unkowns. Let R be the reduced row echelon form of the matrix $[A \mid \mathbf{b}]$. If the linear system is homogeneous, then $\mathbf{b} = \mathbf{0}$ and the number of leading 1's in R is at most $\min\{m, n\}$. Hence, if $m < n$ and R has r nonzero rows, then r is the number of leading 1's and $r \leq m$.

Theorem 2.4 in the above is not true in general if $\mathbf{b} \neq \mathbf{0}$. For example, if we consider the system

$$\begin{cases} x + y + z = 2 \\ x + y + z = 3 \end{cases} . \tag{2.4}$$

Then the augmented matrix corresponding the above system is $\begin{bmatrix} 1 & 1 & 1 & 2 \\ 1 & 1 & 1 & 3 \end{bmatrix}$. Hence the reduced row echelon form of the above matrix is $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. Therefore, system (2.4) has three unknowns and 2 leading variables. However, (2.4) is inconsistent.

Suggested problems: (§2.2) 12, 26, 41, 46, D8.

2.1 Operations on Matrices

For an $m \times n$ matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \tag{2.5}$$

if $m = n$, we call A a square matrix of order n. For simplicity, we write $A = [a_{i,j}]_{m\times n}$ or $A = [a_{i,j}]$. We also use the symbol $(A)_{ij}$ to stand for the entry in row i and column j of a matrix A.

Definition 2.6 *For two matrices A and B,*

1. *$A = B$ if and only if A and B have the same size and their corresponding entries are equal.*

2. *when A and B have the same size, we define the* sum *and the* difference*:*

$$(A \pm B)_{ij} = (A)_{ij} \pm (B)_{ij}.$$

3. *for a scalar c, the* product *cA is defined by $(cA)_{ij} = c(A)_{ij}$.*

4. *For a matrix $A = [a_{ij}]_{m\times n}$, we define a row vector and a column vector of A as follows:*

   - *for $1 \le i \le m$, the ith row vector of A is*

$$\mathbf{r}_i(A) = \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix}.$$

   - *for $1 \le j \le n$, the jth column vector of A is*

$$\mathbf{c}_j(A) = \begin{bmatrix} a_{1j} & a_{2j} & \cdots & a_{mj} \end{bmatrix}.$$

Now we define the product of two matrices. Let

$$A := \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \text{and} \quad \mathbf{x} := \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

We define the *product* $A\mathbf{x}$ as the linear combination of the column vectors of A that has the entries of $\mathbf{x}$ as coefficients. i.e.,

$$A\mathbf{x} = x_1\mathbf{c}_1(A) + x_2\mathbf{c}_2(A) + \cdots + x_n\mathbf{c}_n(A) = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}.$$

Then we can see that a linear system can be written as a single matrix equation as follows:

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

We call A the *coefficient matrix* of the system.

Note that a linear system (2.1) **is consistent if and only if** $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$ **can be written as a linear combination of the column vectors of** A**.**

Theorem 2.7 (Linearity) *Let A be an $m \times n$ matrix. For any column vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ and a scalar $c \in \mathbb{R}$, we have*

$$A(c\mathbf{u} + \mathbf{v}) = c(A\mathbf{u}) + A\mathbf{v}.$$

Proof: Put

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}, \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}.$$

Then $c\mathbf{u}+\mathbf{v} = \begin{bmatrix} cu_1+v_1 \\ cu_2+v_2 \\ \vdots \\ cu_n+v_n \end{bmatrix}$. Thus we have

$$\begin{aligned} A(c\mathbf{u}+\mathbf{v}) &= \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} cu_1+v_1 \\ cu_2+v_2 \\ \vdots \\ cu_n+v_n \end{bmatrix} \\ &= (cu_1+v_1)\mathbf{a}_1 + \cdots + (cu_n+v_n)\mathbf{a}_n \\ &= c(u_1\mathbf{a}_1 + \cdots + u_n\mathbf{a}_n) + (v_1\mathbf{a}_1 + \cdots + v_n\mathbf{a}_n) \\ &= c(A\mathbf{u}) + A\mathbf{v}. \end{aligned}$$

■

How can we define the product AB?

Let A be an $m\times s$ matrix and B an $s\times n$ matrix. Then the matrix B defines a map $\mathbb{R}^n \to \mathbb{R}^s$, $\mathbf{x} \mapsto B\mathbf{x}$, and A defines a map $\mathbb{R}^s \to \mathbb{R}^m$, $\mathbf{y} \mapsto A\mathbf{y}$. Then the product AB will be defined to satisfy the associativity condition:

$$(AB)\mathbf{x} = A(B\mathbf{x}) \text{ for } \mathbf{x} \in \mathbb{R}^n.$$

Let $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_s \end{bmatrix}$ and $B = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{bmatrix}$. Then

$$A(B\mathbf{x}) = A(x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n) = x_1A\mathbf{b}_1 + \cdots x_nA\mathbf{b}_n.$$

Hence we define the product AB as follows:

Definition 2.8 *Let A be an $m\times s$ matrix, B an $s\times n$ matrix, and $B = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{bmatrix}$. Then*

$$AB = \begin{bmatrix} A\mathbf{b}_1 & \cdots & A\mathbf{b}_n \end{bmatrix}.$$

Theorem 2.9 (The Row-Column Rule or Dot Product Rule) *Let A be an $m\times s$ matrix and B an $s\times n$ matrix. Then*

$$(AB)_{ij} = \mathbf{r}_i(A)\mathbf{c}_j(B) = \mathbf{r}_i(A)\cdot\mathbf{c}_j(B).$$

Proof: Note that $(AB)_{ij} = (A\mathbf{b}_j)_i$. Set $A = \begin{bmatrix}\mathbf{a}_1 & \cdots & \mathbf{a}_s\end{bmatrix}$ and $\mathbf{b}_j = \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{sj}\end{bmatrix}$. Then we have

$$A\mathbf{b}_j = b_{1j}\mathbf{a}_1 + \cdots + b_{sj}\mathbf{a}_s = b_{1j}\begin{bmatrix} a_{11} \\ \vdots \\ a_{i1} \\ \vdots \\ a_{m1}\end{bmatrix} + \cdots + b_{sj}\begin{bmatrix} a_{1s} \\ \vdots \\ a_{is} \\ \vdots \\ a_{ms}\end{bmatrix}.$$

Therefore,

$$(A\mathbf{b}_j)_i = a_{i1}b_{1j} + a_{is}b_{2j} + \cdots + a_{is}b_{sj} = \begin{bmatrix} a_{i1} & \cdots & a_{is}\end{bmatrix}\begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{sj}\end{bmatrix} = \mathbf{r}_i(A)\mathbf{c}_j(B).$$

■

Example 2.10 *Let*

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \qquad \textit{and} \qquad B = \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix}.$$

Then

$$AB = \begin{bmatrix} 12 & 27 & 30 & 13 \\ 8 & -4 & 26 & 12 \end{bmatrix}$$

Theorem 2.11

$$\begin{aligned}
\mathbf{c}_j(AB) &= A\mathbf{b}_j = A\mathbf{c}_j(B) = b_{1j}\mathbf{c}_1(A) + \cdots + b_{sj}\mathbf{c}_s(A) \\
\mathbf{r}_i(AB) &= \begin{bmatrix}\mathbf{r}_i(A)\mathbf{c}_1(B) & \cdots & \mathbf{r}_i(A)\mathbf{c}_n(B)\end{bmatrix} = \mathbf{r}_i(A)B \\
&= \begin{bmatrix} a_{i1} & \cdots & a_{is}\end{bmatrix}\begin{bmatrix}\mathbf{r}_1(B) \\ \vdots \\ \mathbf{r}_s(B)\end{bmatrix} \\
&= a_{i1}\mathbf{r}_1(B) + \cdots + a_{is}\mathbf{r}_s(B)
\end{aligned}$$

Definition 2.12 *Let A be an $m \times n$ matrix. Then the* transpose of A, *denoted by A^T, is the $n \times m$ matrix such that $(A^T)_{ij} = (A)_{ji}$.*

Definition 2.13 *For a square matrix A, the* trace of A, *denoted by* $\mathrm{tr}(A)$, *is the sum of the entries on the main diagonals.*

Example 2.14 *If* $A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix}$, *then* $A^T = \begin{bmatrix} 1 & 2 \\ 2 & 6 \\ 4 & 0 \end{bmatrix}$.

Example 2.15 *If* $A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}$, *then* $\mathrm{tr}(A) = 0$.

Exercise. For an $m \times n$ matrix A, show that

$$\mathrm{tr}(AA^T) = \mathrm{tr}(A^T A) = \sum_{\substack{1 \le i \le m \\ 1 \le j \le n}} a_{ij}^2.$$

Definition 2.16 *For two column vectors* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$,

- $\mathbf{u}^T\mathbf{v}$*: the* matrix inner product *of* $\mathbf{u}$ *with* $\mathbf{v}$
- $\mathbf{u}\mathbf{v}^T$*: the* matrix outer product *of* $\mathbf{u}$ *with* $\mathbf{v}$

Then one can also check that $\mathrm{tr}(\mathbf{u}\mathbf{v}^T) = \mathrm{tr}(\mathbf{v}\mathbf{u}^T) = \mathbf{u} \cdot \mathbf{v}$.