Linear Algebra

soupless#2543

June 5, 2024

Contents

1 Introduction to Linear Systems 1

2 Matrix operations, part 2 4

3 Matrix operations, part 3 8

Lecture 1: Introduction to Linear Systems

Example : Consider the following system of equations:

$$\left.\begin{aligned} x_1 - x_2 &= 3 \\ x_1 + x_2 &= 4 \end{aligned}\right\}$$

We can see that $x_1 = 3.5$ and $x_2 = 0.5$ solves the system.

Note 1.1

System of equations with a solution are *consistent*.

Example : Consider another system of equations:

$$\left.\begin{aligned} 2x_1 - 5x_2 &= 8 \\ 4x_1 - 10x_2 &= 9 \end{aligned}\right\}$$

Observe that both sides of the first equation can be multiplied by 2 to get $4x_1 - 10x_2 = 16$ which, together with the second equation, yields $16 = 9$. This is clearly a contradiction, hence, there are no solutions to this system.

Note 1.2

System of equations with no solution are called *inconsistent*.

Definition 1.1 (System of equations)

A system of m linear equations in n unknowns $x_1, x_2, \ldots, x_n$ is a set of m equations each with n unknowns. It is usually written as

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

where a_{ij} is the coefficient of x_j in the ith equation.

Definition 1.2 (Solution)

A solution to $(*)$ is a sequence of n numbers $(s_1, s_2, \ldots, s_n)$ such that each equation in $(*)$ is satisfied if $x_i = s_i$ for each $i = 1, 2, \ldots, n$.

Definition 1.3 (Trivial solution)

The solution set $(0, 0, \ldots, 0)$ is called the trivial solution.

Definition 1.4 (Homogeneous system)

If a system $(*)$ has $b_i = 0$ for each $i = 1, 2, \ldots, n$, then $(*)$ is called a homogeneous system.

Note 1.3

Homogeneous systems are always consistent since it is satisfied by the trivial solution.

Definition 1.5 (Equivalent systems)

Two systems are equivalent if they have the same number of equations and unknowns, and they have exactly the same set of solutions.

Definition 1.6 (Elimination method)

This method is used to solve systems of linear equations. The possible steps are as follows:

- Interchange the ith and jth equation.
- Multiply an equation by a nonzero constant.
- Replace the ith equation by c times the jth equation plus the ith equation.

Definition 1.7 (Matrix)

An $m \times n$ matrix is a rectangular array of mn real (or complex) numbers arranged in m horizontal rows and n vertical columns:

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

where the ith row of A is

$$\begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix}$$

and the jth row of A is

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}.$$

Definition 1.8 (Row and column of a matrix)

We define $\text{row}_i(A)$ as the ith row of A, that is,

$$\text{row}_i(A) = \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix}$$

and $\text{col}_j(A)$ as the jth column of A, that is,

$$\text{col}_j(A) = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}.$$

Definition 1.9 (Matrix size)

An $m \times n$ matrix A is said to have a size of $m \times n$. If $m = n$, then A is a square matrix of order n.

Definition 1.10 (Main diagonal)

In a square matrix of order n A, the numbers a_{ii} for $i = 1, 2, \ldots, n$ form the main diagonal of A. It is usually written as $(a_{11}, a_{22}, \ldots, a_{nn})$.

Definition 1.11 (n-vector)

An n-vector K is an $n \times 1$ matrix.

Definition 1.12 (Equality of matrices)

Two matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ are said to be equal if they have the same size, that is, they are both $m \times n$, and $a_{ij} = b_{ij}$ for each $1 \leq i \leq m$ and $1 \leq j \leq n$.

Definition 1.13 (Matrix addition)

If $A = [a_{ij}]$ and $B = [b_{ij}]$ are both $m \times n$ matrices, then the sum $A + B$ is the $m \times n$ matrix $C = [c_{ij}]$ such that $c_{ij} = a_{ij} + b_{ij}$ for all i, j.

Example: Let $A = \begin{bmatrix} 1 & 0 & 3 \\ 2 & 1 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 5 & -1 & -3 \\ 4 & 2 & -4 \end{bmatrix}$. Then, $A + B = \begin{bmatrix} 6 & -1 & 0 \\ 6 & 3 & 0 \end{bmatrix}$.

Definition 1.14 (Scalar multiplication)

If $A = [a_{ij}]$ is an $m \times n$ matrix and r is a real number, then the scalar multiple of A by r, denoted by rA, is the $m \times n$ matrix $C = [c_{ij}]$ such that $c_{ij} = ra_{ij}$ for all i, j.

Example:

$$3 \begin{bmatrix} 1 & 2 & 6 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 6 & 18 \\ 3 & 3 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 8 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 8 \end{bmatrix} = 8 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Definition 1.15 (Matrix subtraction)

Let A and B be $m \times n$ matrices. The difference of A and B, denoted by $A - B$, is defined by $A + (-1)B$.

Lecture 2: Matrix operations, part 2

Definition 2.1 (Transpose)

Let $A = [a_{ij}]$ be an $m \times n$ matrix. The transpose of A, denoted by A^T, is the $n \times m$ matrix $C = [c_{ij}]$ such that $c_{ij} = a_{ji}$ for all i, j.

Example:

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 4 & 1 & 5 \end{bmatrix} \implies A^T = \begin{bmatrix} 1 & 4 \\ 0 & 1 \\ 2 & 5 \end{bmatrix}$$

$$B = \begin{bmatrix} 3 & 6 & 9 & 12 \end{bmatrix} \implies B^T = \begin{bmatrix} 3 \\ 6 \\ 9 \\ 12 \end{bmatrix}.$$

Definition 2.2 (Set of n-vectors)

We define $\mathbb{R}^n$ as the set of n-vectors over R, that is,

$$\mathbb{R}^n = \left\{ \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \middle| \; u_i \in \mathbb{R} \text{ for } 1 \leq i \leq n \right\}$$

Definition 2.3 (Dot product)

Let $u, v \in \mathbb{R}^n$. The dot product of u and v, denoted by $u \cdot v$, is defined by

$$c = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

or more compactly,

$$c = \sum_{i=1}^{n} u_i v_i.$$

Example: Let $u = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix}$ and $v = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$. Then,

$$\begin{aligned} u \cdot v &= (-1)(1) + (1)(-1) + (-1)(1) + (1)(-1) \\ u \cdot v &= -4. \end{aligned}$$

Example: Let $u = \begin{bmatrix} \sin\theta \\ \cos\theta \end{bmatrix}$ for some real number θ. Then,

$$u \cdot u = \sin^2\theta + \cos^2\theta$$

$$u \cdot u = 1$$

Theorem 2.1

Let $u \in \mathbb{R}^n$. Then, $u \cdot u \geq 0$.

Proof: Let $u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$. Then, $u \cdot u = u_1^2 + u_2^2 + \cdots + u_n^2$ Since u_i is a real number, then $u_i^2 \geq 0$ for all i. Then, $u_1^2 + u_2^2 + \cdots + u_n^2 \geq 0$, hence $u \cdot u \geq 0$. ■

Definition 2.4 (Matrix multiplication)

Let $A = [a_{ij}]$ be an $m \times p$ matrix, and let $B = [b_{ij}]$ be a $p \times n$ matrix. The product of A and B, denoted by AB, is the $m \times n$ matrix $C = [c_{ij}]$ where

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}.$$

A more concise way to keep it in mind is to set the (i, j) entry of AB as $\text{row}_i(A)^T \cdot \text{col}_j(B)$.

Example: Let $A = \begin{bmatrix} 1 & 1 & -2 \\ 0 & 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 2 \\ 0 & 4 \\ 1 & 6 \end{bmatrix}$. We can see that A is of size 2×3 and B is of size 3×2. Hence, AB will be of size 2×2. Then,

$$AB = \begin{bmatrix} 1(0) + (-1)(0) + (-2)(1) & 1(2) + (-1)(4) + (-2)(6) \\ 0(0) + 3(0) + 4(1) & 0(2) + 3(4) + 4(6) \end{bmatrix}$$
$$AB = \begin{bmatrix} -2 & -6 \\ 4 & 36 \end{bmatrix}$$

Example:

1. If $u = \begin{bmatrix} x \\ 3 \\ 4 \end{bmatrix}$ and $u \cdot u = 50$, find the values of x.

$$\begin{aligned} u \cdot u &= 50 \\ x^2 + 3^2 + 4^2 &= 50 \\ x^2 + 9 + 16 &= 50 \\ x^2 &= 25 \\ x &= \pm 5 \end{aligned}$$

2. Let $A = \begin{bmatrix} 1 & 2 & x \\ 3 & -1 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} y \\ x \\ 1 \end{bmatrix}$. If $AB = \begin{bmatrix} 6 \\ 8 \end{bmatrix}$, find x and y.

$$\begin{aligned}
AB &= \begin{bmatrix} 6 \\ 8 \end{bmatrix} \\
\begin{bmatrix} 1 & 2 & x \\ 3 & -1 & 2 \end{bmatrix} \begin{bmatrix} y \\ x \\ 1 \end{bmatrix} &= \begin{bmatrix} 6 \\ 8 \end{bmatrix} \\
\begin{bmatrix} y + 2x + x \\ 3y - x + 2 \end{bmatrix} &= \begin{bmatrix} 6 \\ 8 \end{bmatrix} \\
\begin{bmatrix} y + 3x \\ 3y - x + 2 \end{bmatrix} &= \begin{bmatrix} 6 \\ 8 \end{bmatrix}
\end{aligned}$$

We get the system of linear equations

$$\begin{cases} y + 3x & = 6 \\ 3y - x + 2 & = 8 \end{cases}$$

We solve this by elimination method:

$$\begin{aligned}
3y - x + 2 &= 8 \\
3y - x &= 6 \\
9y - 3x &= 18 \\
(y + 3x) + (9y - 3x) &= 6 + 18 \\
10y &= 24 \\
y &= \frac{12}{5}
\end{aligned}$$

$$\begin{aligned}
y + 3x &= 6 \\
\frac{12}{5} + 3x &= 6 \\
\frac{4}{5} + x &= 2 \\
x &= \frac{6}{5}
\end{aligned}$$

Hence, $x = \frac{6}{5}$ and $y = \frac{12}{5}$.

3. Let $A = \begin{bmatrix} 4 & 1 & -7 \\ 1 & 1 & -2 \\ -3 & 4 & 6 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & -5 \\ 1 & 7 \\ 0 & 8 \end{bmatrix}$. Find AB.

$$\begin{aligned}
AB &= \begin{bmatrix} 4 & 1 & -7 \\ 1 & 1 & -2 \\ -3 & 4 & 6 \end{bmatrix} \begin{bmatrix} 1 & -5 \\ 1 & 7 \\ 0 & 8 \end{bmatrix} \\
AB &= \begin{bmatrix} 4(1) + (-1)(1) + 7(0) & 4(-5) + (-1)(7) + 7(8) \\ 1(1) + 1(1) + (-2)(0) & 1(-5) + 1(7) + (-2)(8) \\ (-3)(1) + 4(1) + 6(0) & (-3)(-5) + 4(7) + 6(8) \end{bmatrix} \\
AB &= \begin{bmatrix} 3 & 29 \\ 2 & -14 \\ 1 & 0 \end{bmatrix}
\end{aligned}$$

4. Let $A = [a_{ij}]$ be a 10×10 matrix where $a_{ij} = i$ for $i = 1, 2, \ldots, 10$. Find the $(8, 9)$ entry of AA.
   Let x be the $(8, 9)$ entry of AA. Then,

$$x = \text{row}_8(A)^T \cdot \text{col}_9(A)$$

$$x = \begin{bmatrix} 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 \end{bmatrix}^T \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \end{bmatrix}$$

$$x = 8(1) + 8(2) + 8(3) + 8(4) + 8(5) + 8(6) + 8(7) + 8(8) + 8(9) + 8(10)$$

$$x = 440$$

Note 2.1

We use the notation A^k to mean $\underbrace{AA \cdots A}_{k \text{ times}}$, and is only valid if A is square.

Lecture 3: Matrix operations, part 3

Recall that a system of m linear equations of n unknowns $(x_1, x_2, \ldots, x_n)$ is written as

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

Let

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

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

be the coefficient matrix, unknown matrix, and the constant matrix, respectively. Then, we can write the system as $A\mathbf{x} = b$.

We can also write the system in the form

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

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

Example: Consider the system

$$\left.\begin{aligned} 10x_1 - 2x_2 + x_3 + x_4 = 5 \\ 7x_1 + x_2 - x_4 = 7 \\ 2x_1 + 9x_2 + 4x_3 = -1 \end{aligned}\right\}$$

In matrix form, we can write this as

$$\begin{bmatrix} 10 & -2 & 1 & 1 \\ 7 & 1 & 0 & -1 \\ 2 & 9 & 4 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \\ -1 \end{bmatrix}.$$

Definition 3.1 (Augmented matrix)

For a system of m linear equations with n unknowns, we can implicitly drop the unknown matrix and have a representation for the system by an augmented matrix, that is, we use only the coefficient and constant matrices and write it in the form

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

Definition 3.2 (Trace)

Let A be a square matrix of order n. The trace of A, denoted by $\mathrm{Tr}(A)$, is defined as

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

Theorem 3.1

Let A and B be square matrices of order n, and r be a real number. Then,

- $\mathrm{Tr}(rA) = r\,\mathrm{Tr}(A)$,
- $\mathrm{Tr}(A + B) = \mathrm{Tr}(A) + \mathrm{Tr}(B)$,
- $\mathrm{Tr}(A) = \mathrm{Tr}(A^T)$,
- $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

Proof: Let $C = rA$. Then, $c_{ij} = ra_{ij}$ for all i, j. Also,

$$\begin{aligned} rA &= C \\ \mathrm{Tr}(rA) &= \mathrm{Tr}(C) \\ \mathrm{Tr}(rA) &= \sum_{i=1}^{n} c_{ii} \\ \mathrm{Tr}(rA) &= \sum_{i=1}^{n} ra_{ii} \\ \mathrm{Tr}(rA) &= r\sum_{i=1}^{n} a_{ii} \\ \mathrm{Tr}(rA) &= r\,\mathrm{Tr}(A) \end{aligned}$$

∎

Proof: It is clear that the (i,j) entry of $A+B$ is $a_{ij}+b_{ij}$. Then,

$$\mathrm{Tr}(A+B) = \sum_{i=1}^{n}(a_{ii}+b_{ii})$$
$$\mathrm{Tr}(A+B) = \sum_{i=1}^{n} a_{ii} + \sum_{i=1}^{n} b_{ii}$$
$$\mathrm{Tr}(A+B) = \mathrm{Tr}(A) + \mathrm{Tr}(B)$$

■

Proof: The (i,j) entry of A is the (j,i) entry of A^T. Hence, the (i,i) entry of A is the (i,i) entry of A^T, which is the same.

$$\mathrm{Tr}(A^T) = \sum_{i=1}^{n} a_{ii}^T$$
$$\mathrm{Tr}(A^T) = \sum_{i=1}^{n} a_{ii}$$
$$\mathrm{Tr}(A^T) = \mathrm{Tr}(A)$$

■

Proof: We first prove that
$$\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} = \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}$$
for finite natural numbers m,n and for real numbers a_{ij}. We interpret this as having the $m\times n$ matrix A, and we get the sum of all the entries. One way is to get the sum of each row, then get the sum of these sums. This would be $\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}$. Another way would be to get the sum of each column, then get the sum of these sums, which would be $\sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}$. Since we are adding real numbers, addition should be commutative. Hence, $\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} = \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}$.

Then,

$$\mathrm{Tr}(AB) = \sum_{i=1}^{n}(AB)_{ii}$$
$$\mathrm{Tr}(AB) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ji}$$
$$\mathrm{Tr}(AB) = \sum_{j=1}^{n}\sum_{i=1}^{n} b_{ji}a_{ij}$$
$$\mathrm{Tr}(AB) = \sum_{j=1}^{n}(BA)_{jj}$$
$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$$

■