Lecture 03

Systems of Linear Equations I: Forms and Elimination

Instructor: Rouyu Sun

香港中文大學(深圳) The Chinese University of Hong Kong, Shenzhen | 数据科学学院 School of Data Science

Recall

In the last lectures …

- Definition of norm and dot (inner product)
- Calculation of vector norms and inner products
- Real-world examples

Today

Today … System of **Linear Equations!**

After this lecture, you should be able to

1. Write the 4 forms of **systems of linear equations**
2. Write various forms of **matrix-vector product**
3. Understand the idea of **elimination for solving systems**

Reading Material: Logic and Proof

Difficulties in Linear Algebra, partially due to:

- lack of training of proving
- lack of LOGIC.

Material to check:

https://www.math.toronto.edu/preparing-for-calculus/ 3_logic/we_2_if_then.html

Part I System of Linear Equations

Linear System of Equations: Preliminary School Example

Problem (Chicken-Rabbit Problem 鸡兔同笼)

There are 16 heads and 36 feet in a cage.
How many chickens and how many rabbits are there?

Assumption: Each chicken has 1 head and 2 feet, and each rabbit has 1 head and 4 feet.

Linear Equations

Definition (Linear Equations)

A **linear equation** is the equation of the form

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b$$

where $a_1, a_2, \ldots, a_n, b$ are real numbers and $x_1, x_2, \ldots, x_n$ are variables

Write it in the vector form:

Exercise

Exercise (Linear Equations)

Are the following linear equations?

1. $-x_1 + 4x_4 = 2x_2 + 3x_3$

x_1, x_2, x_3, x_4 are variables

Exercise

Exercise (Linear Equations)

Are the following linear equations?

1. $-x_1 + 4x_4 = 2x_2 + 3x_3$ — x_1, x_2, x_3, x_4 are variables

2. $-x_1x_4 = 2x_2 + 3x_3$ — x_1, x_2, x_3, x_4 are variables

Exercise

Exercise (Linear Equations)

Are the following linear equations?

1. $-x_1 + 4x_4 = 2x_2 + 3x_3$ x_1, x_2, x_3, x_4 are variables

2. $-x_1x_4 = 2x_2 + 3x_3$ x_1, x_2, x_3, x_4 are variables

3. $a_1x_1 + a_2x_2 = b$ x_1, x_2, a_1, a_2, b are variables

System of Linear Equations

Definition (System of Linear Equations)

An $m \times n$ **system of linear equations** is a *collection* of m linear equations with n variables

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

where all a_{ij} and b_i are real numbers and $x_1, x_2, \ldots, x_n$ are variables

(A system of linear equations can be called a **linear system** for short)

Exercise

Exercise (Linear Equations)

Is the following a linear system?

$$-a_{11} + 4a_{12} = 2a_{13} + 3a_{14}$$

$$5a_{11} + 3a_{12} = a_{13} + 3a_{14}$$

Where $a_{11}, a_{12}, a_{13}, a_{14}$ are variables

Exercise

Exercise (Linear Equations)

Is the following a linear system?

$$-a_{11} + 4a_{12} = 2a_{13} + 3a_{14}$$

$$5a_{11} + 3a_{12} = a_{13} + 3a_{14}$$

Where $a_{11}, a_{12}, a_{13}, a_{14}$ are variables

It is critical to know what are the **variables (unknowns)**!

(such as the weight vectors in our movie preference example)

What is New in this Course?

You learned these in middle school (or even primary school).

What more can we study? What will be new?

1. Practice.

   You can solve system of equations in 2 variables.
   What about 5 variables? What about 100 variables?
   What is a general method to solve an any-variable system? (For computers)

What is New in this Course?

You learned these in middle school (or even primary school).

What more can we study? What will be new?

1. Practice.

   You can solve system of equations in 2 variables.
   What about 5 variables? What about 100 variables?
   What is a general method to solve an any-variable system? (For computers)

2. Theory.

   Does your general method always work? Can you prove it?

First step: To answer these questions, we will rewrite system of equations with vectors and matrices.

Row-Vector Form of System

(F1) scalar form

$$x_1 + x_2 = 14$$
$$2x_1 + 4x_2 = 36$$

$$\xrightarrow{\text{row-reduction}}$$

(F2) Row vector form

•

=

•

dot product

Row-Vector Form of System

(F1) scalar form

$$x_1 + x_2 = 14$$
$$2x_1 + 4x_2 = 36$$

$$\xrightarrow{\text{row-reduction}}$$

(F2) Row vector form

•

=

•

dot product

Can we make the form even simpler？

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad b = \begin{bmatrix} 14 \\ 36 \end{bmatrix}$$

Define $\triangleq$, then the system becomes

(F0) matrix form:

Column-Vector Form

$$x_1 + x_2 = 14$$
$$2x_1 + 4x_2 = 36$$

Another way of writing the equations: Column-vector form

(F3) Column vector form

Four Forms of Linear Systems

Scalar form

$$\begin{cases} x_1 + x_2 = 16, \\ 2x_1 + 4x_2 = 36. \end{cases}$$

Row-vector form

(Unknown vector satisfies n linear equations simultaneously)

Column-vector form

(Unknown combination of columns produces vector b)

Matrix form

(Given matrix times unknown vector produces b)

Four Forms and Matrix-Vector Product

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad b = \begin{bmatrix} 14 \\ 36 \end{bmatrix}$$

$$\begin{cases} x_1 + x_2 = 16, \\ 2x_1 + 4x_2 = 36. \end{cases} \longleftrightarrow Ax = b$$

Define Ax =

$$\begin{cases} [1, 1] \cdot [x_1, x_2] = 16, \\ [2, 4] \cdot [x_1, x_2] = 36. \end{cases} \longleftrightarrow Ax = b$$

Define Ax =

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} x_1 + \begin{bmatrix} 1 \\ 4 \end{bmatrix} x_2 = \begin{bmatrix} 16 \\ 36 \end{bmatrix}. \longleftrightarrow Ax = b$$

Define Ax =

Three Definitions of Matrix-Vector Product

Ignore equations for a while. Summarize the last page.

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Definition 1: $A\mathbf{x}$ =

Definition 2: $A\mathbf{x}$ =

Definition 3: $A\mathbf{x}$ =

Claim: Three definitions are equivalent.

Next, we extend these definitions to general matrix and vectors..

Part II Matrix-Vector Product & Four Forms of Linear Systems

Textbook v5: Sec. 1.3 (only first half) and Sec. 2.1

Matrix Definition

Definition 4.1 (Matrix)
An $m \times n$ **matrix** A is a rectangular array of numbers with m *rows* and n *columns* in the following form:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} =: (a_{ij})_{m\times n} \in \mathbb{R}^{m\times n}.$$

where all a_{ij} are scalars.

Def 4.2 (Dimension):

For an × **matrix,** the dimension of A is × (reads “m by n”).
For an $m \times 1$ vector, the dimension of the vector is m, or $m \times 1$.

The dimension of $A = \begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix}$ is

Matrix: Example and non-example

Often use square bracket

$$\begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0.5 & -3.7 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 & -2.3 & 0.1 \\ 1.3 & 4 & -0.1 & 0 \\ 4.1 & -1 & 0 & 1.7 \end{bmatrix}$$

Can also use round bracket

$$\begin{pmatrix} 1 & 1 \\ 2 & 4 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0.5 & -3.7 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 & -2.3 & 0.1 \\ 1.3 & 4 & -0.1 & 0 \\ 4.1 & -1 & 0 & 1.7 \end{pmatrix}$$

Non-examples:

Yang-hui triangle

round-table

Matrix Notation

Matrices are often denoted by $A, B, C, \ldots$

For the matrix $A = (a_{ij})_{1 \le i \le m, 1 \le j \le n}$, a_{ij} is called the (i, j)-th entry (element) of A.

When $m = n$, A is called a *square matrix*; a *rectangular matrix o.w.*

When all entries are zeros, A is called a *zero matrix.*

Matrix v.s. Vector v.s. Scalar

- When $m = 1$, A is a row vector
- When $n = 1$, A is a column vector

Matrix: general $\times$

Vector: $m \times 1$ or $1 \times n$ matrix

Scalar: 1×1 matrix

Remark: In python, scalar and 1×1 matrix are different!

e.g. scalar 3.5 v.s. 1×1 matrix [[3.5]]

(Easily cause bug if you don't know this!)

Matrix Conventions

- Column of a matrix $$a_j = \begin{bmatrix} A_{1j} \\ \vdots \\ A_{mj} \end{bmatrix}$$

- Row of a matrix $a^{(i)} = \begin{bmatrix} A_{i1} & \cdots & A_{in} \end{bmatrix}$

The skill to identify matrices' columns and matrices' rows is important for the future, but often ignoredl!

Matrix v.s. Column Vectors: Example of 2 by 3 Matrix

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

$$a_1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \quad a_2 = \begin{bmatrix} 2 \\ 5 \end{bmatrix}, \quad a_3 = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

Express A with its columns

Incorrect answer:

OK but imperfect answer:

Correct answer:

Matrix v.s. Column Vectors: General

Write matrix in terms of column vectors.

Suppose the columns of A are:

Then A=

Observation:

Matrix can be written as a with each

being a column vector.

i.e., Matrix = ______________________ of __________________________

(This understanding will be formalized when talking about block matrix)

Matrix v.s. Row Vectors: Example of 2 by 3 Matrix

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

$$a^{(1)} = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}, \quad a^{(2)} = \begin{bmatrix} 4 & 5 & 6 \end{bmatrix}$$

Then A=

i.e., Matrix = ________________ of ____________________

Definition of Matrix-Vector Product

Definition 3.1 (Matrix-vector product definition)

Let $u = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ and A be an $m \times n$ matrix $A = (a_{ij})$,

where $i = 1, \ldots, m$; $j = 1, \ldots, n$,

The matrix-vector product Au is an $m \times 1$ column vector:

$$Au = \begin{bmatrix} \sum_{j=1}^{n} a_{1j}u_j \\ \sum_{j=1}^{n} a_{2j}u_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj}u_j \end{bmatrix}.$$

Remark: Dimension checking

A u = v

? $n \times 1$ $m \times 1$

Example:

Special Case: Row * Column Vectors

$$a = \begin{bmatrix} a_1 & a_2 \end{bmatrix} \qquad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Can we multiply them?

Tentative answer 1: Yes, by Definition of inner product:

PB: inner product of row and column vectors are not well defined.

Tentative answer 2: a is a vector, but also a 1x2 matrix.

view it as matrix-vector product.

Reading: GPT may mislead you if you are not expert

GPT's definition:

Definition 6.1 (Matrix-vector product Def-1)

A is $m\times n$ matrix, $\vec{b}$ is $n\times 1$ vector,

Suppose $A = \begin{bmatrix} \vec{a}^{(1)} \\ \vdots \\ \vec{a}^{(m)} \end{bmatrix}$, $A\cdot b \triangleq \begin{bmatrix} \vec{a}^{(1)}\vec{b} \\ \vdots \\ \vec{a}^{(m)}\vec{b} \end{bmatrix} \in R^{m\times 1}$

transform to formal writing

$$A\mathbf{b} = \begin{bmatrix} \mathbf{a}^{(1)}\mathbf{b} \\ \mathbf{a}^{(2)}\mathbf{b} \\ \vdots \\ \mathbf{a}^{(m)}\mathbf{b} \end{bmatrix} \in \mathbb{R}^{m\times 1},$$

where each element of the resulting vector $A\mathbf{b}$ is obtained by taking the dot product of the corresponding row vector $\mathbf{a}^{(i)}$ of A with the vector $\mathbf{b}$.

GPT4 adds "dot product" in definition though I did not ask it to do so.

Reading: GPT may mislead you if you are not expert

GPT's definition:

Definition 6.1 (Matrix-vector product Def-1)

A is $m\times n$ matrix, $\vec{b}$ is $n\times 1$ vector,

Suppose $A = \begin{bmatrix} \vec{a}^{(1)} \\ \vdots \\ \vec{a}^{(m)} \end{bmatrix}$, $A\cdot b \triangleq \begin{bmatrix} \vec{a}^{(1)}\vec{b} \\ \vdots \\ \vec{a}^{(m)}\vec{b} \end{bmatrix} \in \mathbb{R}^{m\times 1}$

transform to formal writing

$$A\mathbf{b} = \begin{bmatrix} \mathbf{a}^{(1)}\mathbf{b} \\ \mathbf{a}^{(2)}\mathbf{b} \\ \vdots \\ \mathbf{a}^{(m)}\mathbf{b} \end{bmatrix} \in \mathbb{R}^{m\times 1},$$

where each element of the resulting vector $A\mathbf{b}$ is obtained by taking the dot product of the corresponding row vector $\mathbf{a}^{(i)}$ of A with the vector $\mathbf{b}$.

GPT4 adds "dot product" in definition though I did not ask it to do so.

After an expert points out the issue:

is this really a dot product? Here a^{(i)} is a row vector!

You are absolutely correct to point out the discrepancy: $\mathbf{a}^{(i)}$ is indeed a **row vector**, and the operation between $\mathbf{a}^{(i)}$ (a row vector) and $\mathbf{b}$ (a column vector) is **matrix-vector multiplication**, not the traditional dot product of two vectors in $\mathbb{R}^n$.

The correct description of the operation is that the product $\mathbf{a}^{(i)}\mathbf{b}$ involves a matrix multiplication between a $1\times n$ row vector $\mathbf{a}^{(i)}$ and an $n\times 1$ column vector $\mathbf{b}$, yielding a scalar. This operation is sometimes loosely referred to as a dot product, but it's more accurately a matrix-vector multiplication.

Here is a revised and more accurate version of the definition:

This operation is sometimes loosely referred to as a dot product, but it's more accurately a matrix-vector multiplication.

Relation of Inner Product and Matrix-Vector Product

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

Inner product: $\langle x, y \rangle$ **=**

Matrix-vector product of y and x: Invalid.

Relation of Inner Product and Matrix-Vector Product

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

Inner product: $\langle x, y \rangle$ **=**

Matrix-vector product of y and x: Invalid.

Matrix-vector product of $x^\top$ **and** y**:**

$$x^\top y =$$

Two definitions match: $$\langle x, y \rangle = x^T y$$

This is why we can denote the inner product as $x^\top y$

Other Forms of Matrix-Vector Product: Example

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Form 1: $$Ax = \begin{bmatrix} x_1 + x_2 \\ 2x_1 + 4x_2 \end{bmatrix}$$

Form 2: $$Ax = \begin{bmatrix} (1,1) \cdot (x_1, x_2) \\ (2,4) \cdot (x_1, x_2) \end{bmatrix}$$

Row-form of A is so $A\mathbf{x}$ =

Form 3: $$Ax = \begin{bmatrix} 1 \\ 2 \end{bmatrix} x_1 + \begin{bmatrix} 1 \\ 4 \end{bmatrix} x_2$$

Column-form of A is so $A\mathbf{x}$ =

Row-form of Matrix-Vector Product

Property 1 (Row-form of Matrix-vector product)

Let A be an $m \times n$ matrix and $\mathbf{b}$ be an $n \times 1$ column vector. Suppose the matrix A can be represented

$$A = \begin{bmatrix} \mathbf{a}^{(1)} \\ \mathbf{a}^{(2)} \\ \vdots \\ \mathbf{a}^{(m)} \end{bmatrix}$$

Then

$$A\mathbf{b} = \begin{bmatrix} \mathbf{a}^{(1)}\mathbf{b} \\ \mathbf{a}^{(2)}\mathbf{b} \\ \vdots \\ \mathbf{a}^{(m)}\mathbf{b} \end{bmatrix}$$

Column-form of Matrix-Vector Product

Property 2 (Column-form of Matrix-vector product)

Let A be an $m \times n$ matrix and $\mathbf{w}$ be an $n \times 1$ column vector. Suppose A is represented as:

$$A = [\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_n],$$

Then $$A\mathbf{w} = w_1\mathbf{a}_1 + w_2\mathbf{a}_2 + \cdots + w_n\mathbf{a}_n$$

Aw is a of columns.

Very important observation!

More important than the row-form, in the future parts of the course!

More Examples

$$A = \begin{bmatrix} 1 & 4 & 2 & 3 & 5 \\ -2 & 1 & 3 & 0 & -1 \\ 0 & 7 & -1 & -2 & 4 \end{bmatrix}, \ \mathbf{u} = \begin{bmatrix} 1 \\ -2 \\ 0 \\ 5 \\ -1 \end{bmatrix},$$

$$A\,\mathbf{u} = \begin{bmatrix} 1*1+4*(-2)+2*0+3*5+5*(-1) \\ (-2)*1+1*(-2)+3*0+0*5+(-1)*(-1) \\ 0*1+7*(-2)+(-1)*0+(-2)*5+4*(-1) \end{bmatrix}$$

$$A\,\mathbf{u} = \begin{bmatrix} 3 \\ -3 \\ -28 \end{bmatrix}$$

Or: $$A\,\mathbf{u} = 1 \cdot \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} + (-2) \cdot \begin{bmatrix} 4 \\ 1 \\ 7 \end{bmatrix} + 0 \cdot \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix} + 5 \cdot \begin{bmatrix} 3 \\ 0 \\ -2 \end{bmatrix} + (-1) \cdot \begin{bmatrix} 5 \\ -1 \\ 4 \end{bmatrix}$$

$$= \begin{bmatrix} 3 \\ -3 \\ -28 \end{bmatrix}$$

Part III Idea of Elimination

Partly from Sec. 2.2

Solve System of Linear Equations

Example (Solving a 2×2 system)

$$
\begin{aligned}
x - 2y &= 1 \\
3x + 2y &= 11
\end{aligned}
$$

System of Linear Equations

Definition (informal) (Pivot)

First non-zero in the row that does the elimination

More General …

Example (Solving a 3×3 system)

$$\begin{aligned} 2x + 4y - 2z &= 2 \\ 4x + 9y - 3z &= 8 \\ -2x - 3y + 7z &= 10 \end{aligned}$$

Key Idea: Elimination.

Idea of Elimination

How to solve n by n system?

First, eliminate one variable by subtracting one equation from others.

Second, solve the remaining (n-1) by (n-1) system.

Continue the process until getting __ variable and __ equation.

Similar to “prove by induction” (归纳法证明).

Matrix and Linear Systems

Definition (Coefficient Matrix)

Given a linear system,

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

The **coefficient matrix** of the system is an $m \times n$ matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} =: (a_{ij})_{m\times n}$$

Matrix and Linear Systems

Definition (Coefficient Matrix)

Given a linear system,

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

?

The **coefficient matrix** of the system is an $m \times n$ matrix

$$
A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} =: (a_{ij})_{m\times n}
$$

Augmented Matrix

Definition (Augmented matrix)

Given a linear system,

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

the corresponding augmented matrix is

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

Exercise

Consider
$$
\begin{aligned}
2x_1 + 3x_2 &= 3 \\
x_1 - x_2 &= 4.
\end{aligned}
$$

What is the coefficient matrix?

What is the augmented matrix?

Summary Today

Today, we have learned:

- Formulation of systems of linear equations
- Matrix-vector product and four forms of linear system
- Idea of elimination