System of Linear Equations

Chia-Ping Chen

Professor
Department of Computer Science and Engineering
National Sun Yat-sen University

Linear Algebra

Outline

- System of linear equations
- Gauss elimination
- Matrix algebra
- Gauss elimination with matrix
- Matrix inverse and transpose
- Differential equations with linear algebra

NOTATION

- Scalar: a, b, x
- Vector: $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{x}$
- Matrix: $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{X}$
- Set: $\mathbb{A}, \mathbb{B}, \mathbb{X}$

Systems of Linear Equations

Example (system of linear equations)

Consider a system of linear equations

$$\begin{cases} 1x & + & 2y & = & 3 & \cdots\cdots\cdots\cdots ① \\ 4x & + & 5y & = & 6 & \cdots\cdots\cdots\cdots ② \end{cases}$$

The numbers of unknowns and equations are critical.

- There are 2 equations, ① and ②
- There are 2 unknowns, x and y

The system can be solved via elimination and substitution.

- elimination

$$② - 4 \times ① \;\rightarrow\; -3y = -6 \;\rightarrow\; y = 2$$

- substitution

$$① \xrightarrow{y=2} x = -1$$

ROW PICTURE

- row = equation = line
- solve (x, y) for intersection point of the lines

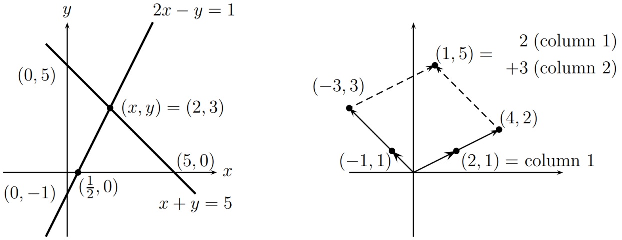
COLUMN PICTURE

- column = vector
- solve x, y for right combination of the vectors

Example (row picture and column picture)

$$\text{row picture: } \begin{cases} 2x - y = 1 \\ x + y = 5 \end{cases}$$

$$\text{column picture: } x\begin{bmatrix} 2 \\ 1 \end{bmatrix} + y\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$



(a) Lines meet at $x = 2$, $y = 3$ (b) Columns combine with 2 and 3

Figure 1.2: Row picture (two lines) and column picture (combine columns).

Definition (types of linear equation systems)

Consider a system of linear equations with m equations and n unknowns. The system is

- square if $m = n$
- under-determined if $m < n$
- over-determined if $m > n$

Theorem (solution of a square system)

Consider a square system of linear equations. Exactly one of the following cases is true.

- *no solution*
- *unique solution*
- *infinite solutions*

Definition (non-singular/singular systems)

Consider a square system of linear equations.

- It is non-singular if its solution is unique
- It is singular if it has no solution or infinite solutions

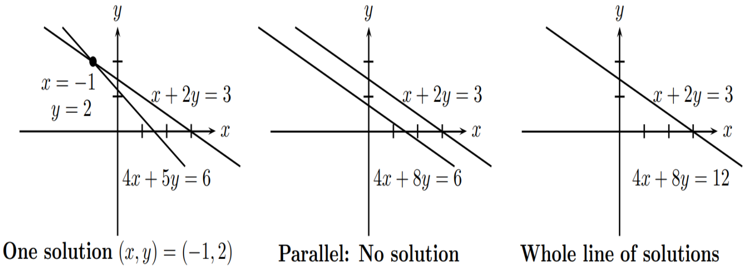


Figure 1.1: The example has one solution. Singular cases have none or too many.

Example (square system of linear equations)

$$\begin{cases} 2u + v + w = 5 \\ 4u - 6v = -2 \\ -2u + 7v + 2w = 9 \end{cases}$$

- It has a solution (verify)

$$u = 1,\ v = 1,\ w = 2$$

- It is non-singular

- row picture

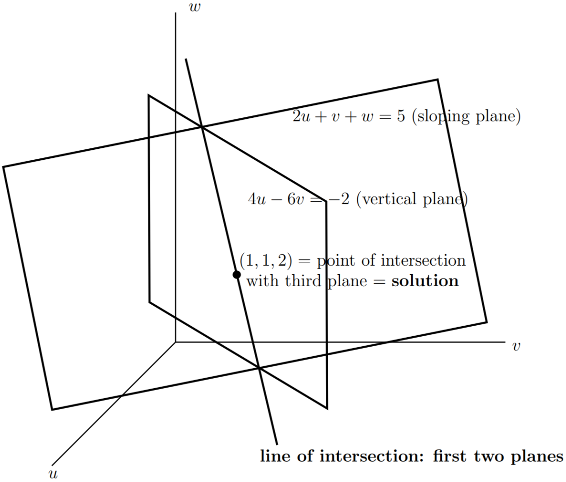
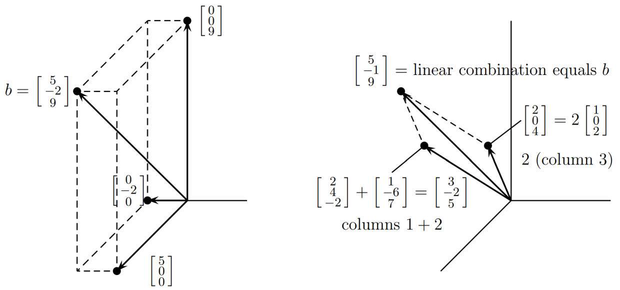


Figure 1.3: The row picture: three intersecting planes from three linear equations.

- It corresponds to vector equation

$$u\begin{bmatrix}2\\4\\-2\end{bmatrix}+v\begin{bmatrix}1\\-6\\7\end{bmatrix}+w\begin{bmatrix}1\\0\\2\end{bmatrix}=\begin{bmatrix}5\\-2\\9\end{bmatrix}$$

- column picture



(a) Add vectors along axes

(b) Add columns $1 + 2 + (3 + 3)$

Figure 1.4: The column picture: linear combination of columns equals b.

Gauss Elimination

BASIC IDEAS OF GAUSS METHOD

We apply Gauss method to solve a system of linear equations.

- Suppose n equations with n unknowns are to be solved
- Unknowns are eliminated from the equations systematically
- Elimination downsizes the system of equations: for $k = n \ldots 2$, derive a square system of $(k-1)$ unknowns from a square system of k unknowns
- Substitution finds the unknowns: for $k' = 1 \ldots n$, substitute the values of $(k'-1)$ unknowns to an equation with k' unknowns and then solve the k'th unknown

Example (Gauss method)

$$\begin{cases} 2u + v + w = 5 & \cdots\cdots\cdots\cdots ① \\ 4u - 6v = -2 & \cdots\cdots\cdots\cdots ② \\ -2u + 7v + 2w = 9 & \cdots\cdots\cdots\cdots ③ \end{cases}$$

$$\xrightarrow{\text{elimination}} \begin{cases} 2u + v + w = 5 & \cdots\cdots\cdots\cdots ① \\ -8v - 2w = -12 & \cdots\cdots\cdots\cdots ②' \\ 8v + 3w = 14 & \cdots\cdots\cdots\cdots ③' \end{cases}$$

$$\xrightarrow{\text{elimination}} \begin{cases} 2u + v + w = 5 & \cdots\cdots\cdots\cdots ① \\ -8v - 2w = -12 & \cdots\cdots\cdots\cdots ②' \\ w = 2 & \cdots\cdots\cdots\cdots ③'' \end{cases}$$

$$\xrightarrow{③''} w = 2 \xrightarrow{\text{substitution } ②'} v = 1 \xrightarrow{\text{substitution } ①} u = 1$$

n equations in n unknowns

(eli.) $\downarrow$ $\uparrow$ (sub.)

$(n-1)$ equations in $(n-1)$ unknowns

(eli.) $\downarrow$ $\uparrow$ (sub.)

$\vdots$

(eli.) $\downarrow$ $\uparrow$ (sub.)

1 equation in 1 unknown

ELIMINATION STEP

Consider 2 linear equations, both including unknown u.

- Decide the ratio of coefficients of u in the equations
- Multiply the first equation by the ratio
- Subtract the result from the second equation
- Now u is removed from the second equation
- We call the above procedure an elimination step

Suppose there are k unknowns (including u).

- The elimination step for u requires $(k+2)$ multiplies
- 1 multiply to decide the ratio
- $(k+1)$ multiplies to multiply the first equation

ELIMINATION RUN

Consider a square system of linear equations with unknown u.

- Suppose we want to eliminate u
- The first equation is repeatedly multiplied and subtracted to eliminate u's in the other equations
- We call the above procedure the elimination run for u

Suppose the system has k unknowns and k equations.

- An elimination run for u has $(k-1)$ elimination steps
- It produces a sub-system with $(k-1)$ unknowns and $(k-1)$ equations (where u is removed)

ELIMINATION PART

Consider a square system with n unknowns.

- The elimination part consists of $(n-1)$ elimination runs
- The system is converted to a *triangular* system
- The triangular system is easy to solve

Substitution Step

A substitution step for u does 2 things.

- Substitute values of the unknowns other than u
- Find the value of u

Suppose the equation has k unknowns (including u).

- A substitution step for u requires k multiplies
- $(k-1)$ multiplies to substitute the values of the unknowns besides u
- 1 multiply to find the value of u

SUBSTITUTION PART

Consider a square system with n unknowns.

- The elimination part of Gauss method converts it to a triangular system
- The substitution part of Gauss method solves it in n substitution steps

Pivots and Multipliers

Let $\mathcal{S}$ be a square system of linear equations and u be an unknown in $\mathcal{S}$. Consider an elimination run for u.

- A non-zero coefficient of u is chosen as pivot
- Given the pivot, the ratio of a coefficient of u in another equation of $\mathcal{S}$ over pivot is a multiplier

unknown	pivot (in eq.)	multipliers
u	2 (①)	2, -1
v	-8 (②$'$)	-1
w	1 (③$''$)	-

Complexity of Gauss elimination

Let $\mathcal{S}$ be a square system with n unknowns. The numbers of multiplies in Gauss method are as follows.

- elimination part

$$\sum_{k=2}^{n} \overbrace{(k-1)(k+2)}^{\text{an elimination run}} = O(n^3)$$

- substitution part

$$\sum_{k'=1}^{n} \overbrace{k'}^{\text{a substitution step}} = O(n^2)$$

Matrix Algebra

Definition (matrix)

A matrix consists of numbers. It can be seen as

- 2-D array of elements (entries)
- list of row vectors
- list of column vectors

Let $\boldsymbol{A}$ be a matrix.

- The position of an element is specified by 2 subscripts
- The order (size) of $\boldsymbol{A}$ is $m \times n$ if it has m rows n columns
- $\boldsymbol{A}$ can be represented in various forms

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_{1:} \\ \vdots \\ \boldsymbol{a}_{m:} \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1 & \dots & \boldsymbol{a}_n \end{bmatrix}$$
$$= \{a_{ij}\}_{m \times n}$$

MATRIX ADDITION

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be matrices.

- If $\boldsymbol{A}$ and $\boldsymbol{B}$ have different orders, matrix addition $\boldsymbol{A}+\boldsymbol{B}$ is undefined
- When they have the same orders, $\boldsymbol{A}+\boldsymbol{B}$ is defined by

$$\begin{aligned}\boldsymbol{A}+\boldsymbol{B} &= \{(a_{ij}+b_{ij})\}_{m\times n} \\ &= \begin{bmatrix} a_{11}+b_{11} & \dots & a_{1n}+b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & \dots & a_{mn}+b_{mn} \end{bmatrix}\end{aligned}$$

Matrix Multiplication

Let $\boldsymbol{A}$ be of order $m \times l$ and $\boldsymbol{B}$ be of order $p \times n$.

- If $l \neq p$, matrix multiplication $\boldsymbol{AB}$ is undefined
- When $l = p$, $\boldsymbol{AB}$ is a matrix of order $m \times n$ with

$$
\begin{aligned}
\boldsymbol{AB} &= \left\{ \sum_{k=1}^{l} a_{ik} b_{kj} \right\}_{m \times n} \\
&= \begin{bmatrix} \sum_{k=1}^{l} a_{1k} b_{k1} & \dots & \sum_{k=1}^{l} a_{1k} b_{kn} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^{l} a_{mk} b_{k1} & \dots & \sum_{k=1}^{l} a_{mk} b_{kn} \end{bmatrix}
\end{aligned}
$$

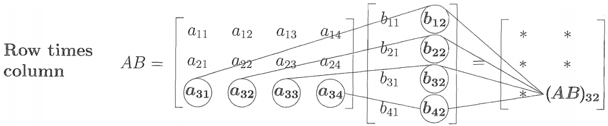


Figure 1.7: A 3 by 4 matrix A times a 4 by 2 matrix B is a 3 by 2 matrix AB.

For $\boldsymbol{AB}$ to be defined

- size of a row vector in $\boldsymbol{A}$ = size of a column vector in $\boldsymbol{B}$
- number of columns in $\boldsymbol{A}$ = number of rows in $\boldsymbol{B}$

SPECIAL CASE OF $m = n = 1$

Let $\boldsymbol{A} = \boldsymbol{a}^T = \begin{bmatrix} a_1 \dots a_l \end{bmatrix}$ and $\boldsymbol{B} = \boldsymbol{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_l \end{bmatrix}$.

- $\boldsymbol{a}^T\boldsymbol{b}$ is

$$\boldsymbol{a}^T\boldsymbol{b} = \begin{bmatrix} a_1 \dots a_l \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_l \end{bmatrix} = \begin{bmatrix} a_1 b_1 + \cdots + a_l b_l \end{bmatrix}$$

- For example

$$\begin{bmatrix} 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 + 0 + 0 \end{bmatrix} = \begin{bmatrix} 2 \end{bmatrix}$$

SPECIAL CASE OF $n = 1$

Let $\boldsymbol{B} = \boldsymbol{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_l \end{bmatrix}$.

- $\boldsymbol{Ab}$ is a combination of the columns of $\boldsymbol{A}$

$$\boldsymbol{Ab} = \begin{bmatrix} \sum_k a_{1k} b_k \\ \vdots \\ \sum_k a_{mk} b_k \end{bmatrix} = \sum_k \begin{bmatrix} a_{1k} b_k \\ \vdots \\ a_{mk} b_k \end{bmatrix} = \sum_k \begin{bmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{bmatrix} b_k$$

$$= \boldsymbol{a}_1 \, b_1 + \cdots + \boldsymbol{a}_l \, b_l = b_1 \, \boldsymbol{a}_1 + \cdots + b_l \, \boldsymbol{a}_l$$

- For example

$$\begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + 0 \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

Special case of $m = 1$

Let $\boldsymbol{A} = \boldsymbol{a}^T = \begin{bmatrix} a_1 \dots a_l \end{bmatrix}$.

- $\boldsymbol{a}^T \boldsymbol{B}$ is a combination of the rows of $\boldsymbol{B}$

$$\begin{aligned}
\boldsymbol{a}^T \boldsymbol{B} &= \begin{bmatrix} \sum_k a_k b_{k1} & \dots & \sum_k a_k b_{kn} \end{bmatrix} \\
&= \sum_k \begin{bmatrix} a_k b_{k1} & \dots & a_k b_{kn} \end{bmatrix} \\
&= \sum_k a_k \begin{bmatrix} b_{k1} & \dots & b_{kn} \end{bmatrix} \\
&= a_1 \boldsymbol{b}_{1:} + \cdots + a_l \boldsymbol{b}_{l:}
\end{aligned}$$

- For example

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

SPECIAL CASE OF $l = 1$

Let $\boldsymbol{A} = \boldsymbol{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix}$ and $\boldsymbol{B} = \boldsymbol{b}^T = \begin{bmatrix} b_1 \dots b_n \end{bmatrix}$.

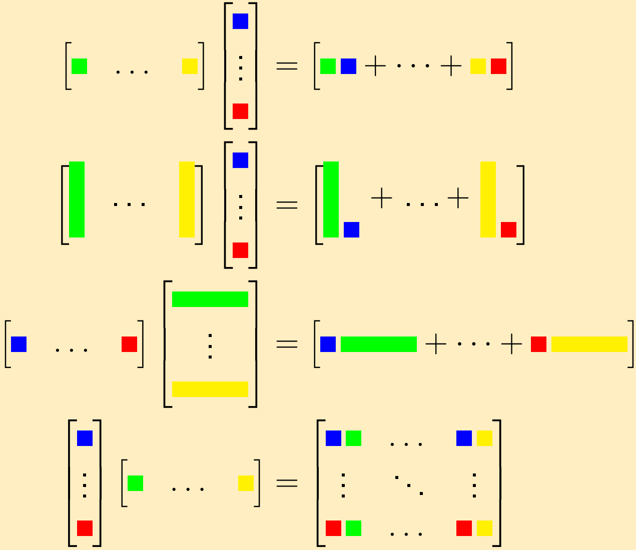
- $\boldsymbol{a}\boldsymbol{b}^T$ is

$$\boldsymbol{a}\boldsymbol{b}^T = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} \begin{bmatrix} b_1 \dots b_n \end{bmatrix} = \begin{bmatrix} a_1 b_1 & \dots & a_1 b_n \\ \vdots & \ddots & \vdots \\ a_m b_1 & \dots & a_m b_n \end{bmatrix}$$

- For example

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

DIAGRAMS OF SPECIAL CASES



COMPUTATION OF MATRIX MULTIPLICATION

Matrix multiplication can be computed by the following ways.

- element by element
- matrix by matrix
- row by row
- column by column
- any feasible block matrix multiplication*

ELEMENT BY ELEMENT COMPUTATION

$(\boldsymbol{AB})_{ij}$ is the product of $\boldsymbol{a}_{i:}$ and $\boldsymbol{b}_j$.

$$
\begin{aligned}
\boldsymbol{AB} &= \begin{bmatrix} \sum_{k=1}^{l} a_{1k}b_{k1} & \dots & \sum_{k=1}^{l} a_{1k}b_{kn} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^{l} a_{mk}b_{k1} & \dots & \sum_{k=1}^{l} a_{mk}b_{kn} \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{a}_{1:}\boldsymbol{b}_1 & \dots & \boldsymbol{a}_{1:}\boldsymbol{b}_n \\ \vdots & \boldsymbol{a}_{i:}\boldsymbol{b}_j & \vdots \\ \boldsymbol{a}_{m:}\boldsymbol{b}_1 & \dots & \boldsymbol{a}_{m:}\boldsymbol{b}_n \end{bmatrix}
\end{aligned}
$$

MATRIX BY MATRIX COMPUTATION

$$\boldsymbol{AB} = \begin{bmatrix} \sum_{k=1}^{l} a_{1k}b_{k1} & \dots & \sum_{k=1}^{l} a_{1k}b_{kn} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^{l} a_{mk}b_{k1} & \dots & \sum_{k=1}^{l} a_{mk}b_{kn} \end{bmatrix}$$

$$= \sum_{k=1}^{l} \begin{bmatrix} a_{1k}b_{k1} & \dots & a_{1k}b_{kn} \\ \vdots & \ddots & \vdots \\ a_{mk}b_{k1} & \dots & a_{mk}b_{kn} \end{bmatrix}$$

$$= \sum_{k=1}^{l} \begin{bmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{bmatrix} \begin{bmatrix} b_{k1} & \dots & b_{kn} \end{bmatrix}$$

$$= \sum_{k=1}^{l} \boldsymbol{a}_k \boldsymbol{b}_{k:}$$

Column by column computation

$(\boldsymbol{AB})_j$ is the product of $\boldsymbol{A}$ and $\boldsymbol{b}_j$.

$$\boldsymbol{AB} = \begin{bmatrix} \sum_{k=1}^{l} a_{1k}b_{k1} & \dots & \sum_{k=1}^{l} a_{1k}b_{kn} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^{l} a_{mk}b_{k1} & \dots & \sum_{k=1}^{l} a_{mk}b_{kn} \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{k=1}^{l} \begin{bmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{bmatrix} b_{k1} \cdots \sum_{k=1}^{l} \begin{bmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{bmatrix} b_{kn} \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{k=1}^{l} \boldsymbol{a}_k b_{k1} & \dots & \sum_{k=1}^{l} \boldsymbol{a}_k b_{kn} \end{bmatrix}$$

$$= \begin{bmatrix} \boldsymbol{Ab}_1 & \dots & \boldsymbol{Ab}_j & \dots & \boldsymbol{Ab}_n \end{bmatrix}$$

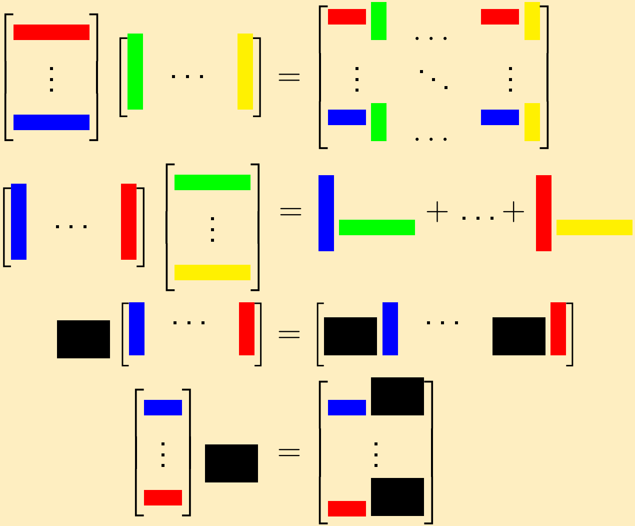
ROW BY ROW COMPUTATION

$(\boldsymbol{AB})_{i:}$ is the product of $\boldsymbol{a}_{i:}$ and $\boldsymbol{B}$.

$$\boldsymbol{AB} = \begin{bmatrix} \sum_{k=1}^{l} a_{1k}b_{k1} & \dots & \sum_{k=1}^{l} a_{1k}b_{kn} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^{l} a_{mk}b_{k1} & \dots & \sum_{k=1}^{l} a_{mk}b_{kn} \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{k=1}^{l} a_{1k} \begin{bmatrix} b_{k1} & \dots & b_{kn} \end{bmatrix} \\ \vdots \\ \sum_{k=1}^{l} a_{mk} \begin{bmatrix} b_{k1} & \dots & b_{kn} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{l} a_{1k}\boldsymbol{b}_{k:} \\ \vdots \\ \sum_{k=1}^{l} a_{mk}\boldsymbol{b}_{k:} \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_{1:}\boldsymbol{B} \\ \vdots \\ \boldsymbol{a}_{i:}\boldsymbol{B} \\ \vdots \\ \boldsymbol{a}_{m:}\boldsymbol{B} \end{bmatrix}$$

DIAGRAMS OF MATRIX MULTIPLICATION



These are special cases of feasible block matrix multiplication.

Example (computation of matrix multiplication)

$$\boldsymbol{AB} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 2 & 0 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & 3 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 7 & 7 \\ -2 & 10 \end{bmatrix}$$

$$\overset{\text{ele.}}{=} \begin{bmatrix} 1\cdot -1 + 2\cdot 1 + 3\cdot 2 & 1\cdot 1 + 2\cdot 3 + 3\cdot 0 \\ 4\cdot -1 + 2\cdot 1 + 0\cdot 2 & 4\cdot 1 + 2\cdot 3 + 0\cdot 0 \end{bmatrix}$$

$$\overset{\text{mat.}}{=} \begin{bmatrix} 1 \\ 4 \end{bmatrix} \begin{bmatrix} -1 & 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 2 \end{bmatrix} \begin{bmatrix} 1 & 3 \end{bmatrix} + \begin{bmatrix} 3 \\ 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \end{bmatrix}$$

$$\overset{\text{col.}}{=} \left[-1 \begin{bmatrix} 1 \\ 4 \end{bmatrix} + 1 \begin{bmatrix} 2 \\ 2 \end{bmatrix} + 2 \begin{bmatrix} 3 \\ 0 \end{bmatrix} \quad 1 \begin{bmatrix} 1 \\ 4 \end{bmatrix} + 3 \begin{bmatrix} 2 \\ 2 \end{bmatrix} + 0 \begin{bmatrix} 3 \\ 0 \end{bmatrix} \right]$$

$$\overset{\text{row}}{=} \begin{bmatrix} 1 \begin{bmatrix} -1 & 1 \end{bmatrix} + 2 \begin{bmatrix} 1 & 3 \end{bmatrix} + 3 \begin{bmatrix} 2 & 0 \end{bmatrix} \\ 4 \begin{bmatrix} -1 & 1 \end{bmatrix} + 2 \begin{bmatrix} 1 & 3 \end{bmatrix} + 0 \begin{bmatrix} 2 & 0 \end{bmatrix} \end{bmatrix}$$

Definition (block matrix*)

A block matrix is a matrix of blocks.

- Each block is a matrix
- A block row is a row of blocks
- A block column is a column of blocks
- The number of block rows is block row order
- The number of block columns is block column order
- The blocks in a block row have the same row order
- The blocks in a block column have the same column order

BLOCK MATRIX AS MATRIX*

Let $\boldsymbol{A}$ be a block matrix.

- $\boldsymbol{A}$ is a matrix
- The row order of $\boldsymbol{A}$ is the sum of the row orders of the blocks in a block column
- The column order of $\boldsymbol{A}$ is the sum of the column orders of the blocks in a block row

For example, let $\boldsymbol{A}$ be a block matrix with 2 block rows and 3 block columns. Let m_i be the row order of the blocks of $\boldsymbol{A}$ in block row i and n_j be the column order of the blocks of $\boldsymbol{A}$ in block column j. Then $\boldsymbol{A}$ is a matrix is of order $m \times n$ where

$$m = m_1 + m_2, \ n = n_1 + n_2 + n_3$$

MATRIX AS BLOCK MATRIX*

Let $\boldsymbol{A}$ be a matrix is of order $m \times n$.

- Partition the rows of $\boldsymbol{A}$ into M block rows
- Partition the columns of $\boldsymbol{A}$ into N block columns
- We have a block matrix of order $M \times N$ (a block matrix with M block rows and N block columns)

Block Matrix Multiplication*

Let $\boldsymbol{A}$ be a block matrix of order $M \times L$ and $\boldsymbol{B}$ be a block matrix of order $P \times N$.

- If $L \neq P$, block matrix multiplication $\boldsymbol{AB}$ is undefined
- When $L = P$, $\boldsymbol{AB}$ is a block matrix of order $M \times N$ with

$$\boldsymbol{AB} = \begin{bmatrix} \sum_{k=1}^{L} \boldsymbol{A}_{1k}\boldsymbol{B}_{k1} & \dots & \sum_{k=1}^{L} \boldsymbol{A}_{1k}\boldsymbol{B}_{kN} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^{L} \boldsymbol{A}_{Mk}\boldsymbol{B}_{k1} & \dots & \sum_{k=1}^{L} \boldsymbol{A}_{Mk}\boldsymbol{B}_{kN} \end{bmatrix}$$

Simply **treat blocks as elements**. For example

$$\begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_{11}\boldsymbol{B}_{11} + \boldsymbol{A}_{12}\boldsymbol{B}_{21} & \boldsymbol{A}_{11}\boldsymbol{B}_{12} + \boldsymbol{A}_{12}\boldsymbol{B}_{22} \\ \boldsymbol{A}_{21}\boldsymbol{B}_{11} + \boldsymbol{A}_{22}\boldsymbol{B}_{21} & \boldsymbol{A}_{21}\boldsymbol{B}_{12} + \boldsymbol{A}_{22}\boldsymbol{B}_{22} \end{bmatrix}$$

Theorem (general matrix multiplication*)

Let $\boldsymbol{A}$ be $m \times l$ and $\boldsymbol{B}$ be $l \times n$. The multiplication $\boldsymbol{AB}$ can be carried out in any feasible block matrix multiplication.

Suppose $\boldsymbol{A}$ is partitioned to be a block matrix of order $M^{\boldsymbol{A}} \times N^{\boldsymbol{A}}$, and $\boldsymbol{B}$ is partitioned to be a block matrix of order $M^{\boldsymbol{B}} \times N^{\boldsymbol{B}}$. Feasible block matrix multiplication requires

1. block row order of $\boldsymbol{A}$ matches block column order of $\boldsymbol{B}$

$$N^{\boldsymbol{A}} = M^{\boldsymbol{B}} = L$$

2. column orders of the blocks of $\boldsymbol{A}$ matches row orders of the blocks of $\boldsymbol{B}$

$$n_l^{\boldsymbol{A}} = m_l^{\boldsymbol{B}}, \; l = 1, \ldots, L$$

Example (block matrix multiplication)

$$\boldsymbol{AB} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 2 & 0 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & 3 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & 3 \end{bmatrix} + \begin{bmatrix} 3 \\ 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 7 & 7 \\ -2 & 10 \end{bmatrix}$$

$$\boldsymbol{BA} = \begin{bmatrix} -1 & 1 \\ 1 & 3 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 2 & 0 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 4 & 2 \end{bmatrix} & \begin{bmatrix} -1 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 3 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 4 & 2 \end{bmatrix} & \begin{bmatrix} 2 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 0 \end{bmatrix} \end{bmatrix}$$

$$= \begin{bmatrix} 3 & 0 & -3 \\ 13 & 8 & 3 \\ 2 & 4 & 6 \end{bmatrix}$$

PROPERTIES OF MATRIX MULTIPLICATION

- always associative

$$(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC})$$

- always distributive

$$\boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C}) = \boldsymbol{AB}+\boldsymbol{AC}$$
$$(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{C} = \boldsymbol{AC}+\boldsymbol{BC}$$

- normally not commutative

$$\boldsymbol{AB} \neq \boldsymbol{BA}$$

Example (matrix multiplication)

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

$$\boldsymbol{AB} = \boldsymbol{BA}?$$

$$\boldsymbol{AC} = \boldsymbol{CA}?$$

Gauss Elimination with Matrix

Representation of a system of linear equations

A system of linear equations of m equations and n unknowns

$$\left\{\begin{array}{ccccccc} a_{11}\,x_1 & + & \dots & + & a_{1n}\,x_n & = & b_1 \\ & & \vdots & & & & \\ a_{m1}\,x_1 & + & \dots & + & a_{mn}\,x_n & = & b_m \end{array}\right.$$

can be represented by $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ where

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}, \ \boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \ \boldsymbol{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

Example (matrix representation)

The system of linear equations

$$\begin{cases} 2u + v + w = 5 \\ 4u - 6v = -2 \\ -2u + 7v + 2w = 9 \end{cases}$$

can be represented by

$$\overbrace{\begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix}}^{\boldsymbol{A}} \overbrace{\begin{bmatrix} u \\ v \\ w \end{bmatrix}}^{\boldsymbol{x}} = \overbrace{\begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix}}^{\boldsymbol{b}}$$

MATRIX AND VECTORS IN LINEAR EQUATION SYSTEM

Let $\boldsymbol{Ax} = \boldsymbol{b}$ be a system of linear equations.

- $\boldsymbol{A}$ is the coefficient matrix
- $\boldsymbol{x}$ is the unknown vector
- $\boldsymbol{b}$ is the right side
- The augmented matrix is

$$\left[\, \boldsymbol{A} \,|\, \boldsymbol{b} \,\right]$$

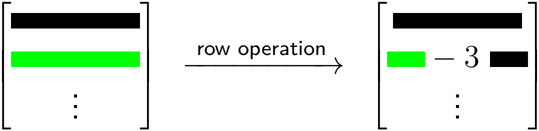
Definition (row operation)

Let $\boldsymbol{A}$ be a matrix. A row operation on $\boldsymbol{A}$ subtracts a multiple of one row of $\boldsymbol{A}$ from another row of $\boldsymbol{A}$.

Example. Subtracting the triple of row 1 from row 2

$$\boldsymbol{a}_{2:} \leftarrow \boldsymbol{a}_{2:} - 3\boldsymbol{a}_{1:}$$

is a row operation. It means



Only one row is changed in a row operation (here row 2).

ELIMINATION STEP = ROW OPERATION

Let $\boldsymbol{Ax} = \boldsymbol{b}$ be a system of linear equations. An elimination step is equivalent to a row operation on augmented matrix $\left[\, \boldsymbol{A} \,\middle|\, \boldsymbol{b} \,\right]$.

For example

$$\left\{\begin{array}{rcr} -8v - 2w & = & -12 \\ 8v + 3w & = & 14 \end{array}\right. \xrightarrow{\text{elimination step}} \left\{\begin{array}{rcr} -8v - 2w & = & -12 \\ w & = & 2 \end{array}\right.$$

is equivalent to

$$\left[\begin{array}{rr|r} -8 & -2 & -12 \\ 8 & 3 & 14 \end{array}\right] \xrightarrow{\text{row operation}} \left[\begin{array}{rr|r} -8 & -2 & -12 \\ 0 & 1 & 2 \end{array}\right]$$

ELIMINATION PART = SEQUENCE OF ROW OPERATION

Let $\boldsymbol{Ax} = \boldsymbol{b}$ be a system of linear equations to solve. The elimination part in Gauss method is equivalent to a sequence of row operation on $\begin{bmatrix} \boldsymbol{A} \mid \boldsymbol{b} \end{bmatrix}$.

$$\overbrace{\begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix}\begin{bmatrix} u \\ v \\ w \end{bmatrix}}^{\boldsymbol{Ax}} = \overbrace{\begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix}}^{\boldsymbol{b}} \xrightarrow{\text{eli.}} \overbrace{\begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} u \\ v \\ w \end{bmatrix}}^{\boldsymbol{Ux}} = \overbrace{\begin{bmatrix} 5 \\ -12 \\ 2 \end{bmatrix}}^{\boldsymbol{c}}$$

$$\overbrace{\left[\begin{array}{ccc|c} 2 & 1 & 1 & 5 \\ 4 & -6 & 0 & -2 \\ -2 & 7 & 2 & 9 \end{array}\right]}^{[\boldsymbol{A} \mid \boldsymbol{b}]} \xrightarrow{\text{3 row operation}} \overbrace{\left[\begin{array}{ccc|c} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ 0 & 0 & 1 & 2 \end{array}\right]}^{[\boldsymbol{U} \mid \boldsymbol{c}]}$$

Definition (identity matrix)

An identity matrix is a square matrix in which every diagonal element is 1 and every off-diagonal element is 0.

The identity matrix of order $n \times n$ is denoted by $\boldsymbol{I}_n$. For example

$$\boldsymbol{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Definition (triangular matrix)

A triangular matrix is a square matrix in which every off-diagonal element is 0 on one side of the diagonal line.

- An upper-triangular matrix has 0s below the diagonal line
- A lower-triangular matrix has 0s above the diagonal line
- A unit triangular matrix has 1s on the diagonal line
- A diagonal matrix is lower-triangular and upper-triangular

Definition (elementary matrix)

An elementary matrix is a square matrix.

- Every diagonal element is 1
- Exactly one off-diagonal element is non-zero

An elementary matrix is specified by its size, the position and the value of the non-zero off-diagonal element. That is

$$\boldsymbol{E} = \{e_{ij} = -m\}_{n\times n}$$

For example

$$\boldsymbol{E} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \longleftrightarrow \boldsymbol{E} = \{e_{21} = -2\}_{3\times 3}$$

ROW OPERATION = ELEMENTARY MATRIX

Let $\boldsymbol{A}$ be a matrix.

- A row operation on $\boldsymbol{A}$ is equivalent to multiplication of an elementary matrix from left
- Specifically, left multiplication of $\boldsymbol{E} = \{e_{ij} = -m\}$ corresponds to subtracting m times of row j from row i
- For example

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_{1:} \\ \boldsymbol{a}_{2:} \\ \boldsymbol{a}_{3:} \end{bmatrix} \xrightarrow{\text{row operation}} \begin{bmatrix} \boldsymbol{a}_{1:} \\ \boldsymbol{a}_{2:} - 2\boldsymbol{a}_{1:} \\ \boldsymbol{a}_{3:} \end{bmatrix}$$

$$\boldsymbol{E}\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{a}_{1:} \\ \boldsymbol{a}_{2:} \\ \boldsymbol{a}_{3:} \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_{1:} \\ \boldsymbol{a}_{2:} - 2\boldsymbol{a}_{1:} \\ \boldsymbol{a}_{3:} \end{bmatrix}$$

Elimination = Sequence of Elementary Matrices

Let $\boldsymbol{Ax} = \boldsymbol{b}$ be a system of linear equations. The elimination part of Gauss method is equivalent to multiplying $[\boldsymbol{A} \,|\, \boldsymbol{b}]$ by a sequence of elementary matrices from left.

By Gauss method

$$\overbrace{\left[\begin{array}{rrr|r} 2 & 1 & 1 & 5 \\ 4 & -6 & 0 & -2 \\ -2 & 7 & 2 & 9 \end{array}\right]}^{[\boldsymbol{A}\,|\,\boldsymbol{b}]} \xrightarrow{\text{elimination part}} \overbrace{\left[\begin{array}{rrr|r} 2 & 1 & 1 & 5 \\ 0 & -8 & -2 & -12 \\ 0 & 0 & 1 & 2 \end{array}\right]}^{[\boldsymbol{U}\,|\,\boldsymbol{c}]}$$

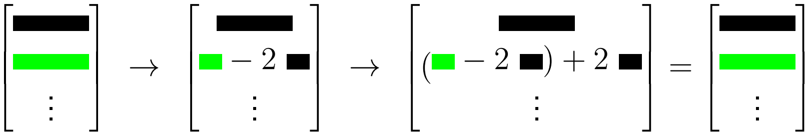
This is equivalent to $\boldsymbol{GFE}\,[\boldsymbol{A} \,|\, \boldsymbol{b}] = [\boldsymbol{U} \,|\, \boldsymbol{c}]$ with

$$\boldsymbol{E} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{F} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{G} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

REVERSE OF ROW OPERATION: ALSO ROW OPERATION

Let $\boldsymbol{A}$ be a matrix. A row operation on $\boldsymbol{A}$ can be reversed by another row operation.

Consider the row operation of subtracting twice the first row from the second row. Its reverse is adding twice the first row to the second row. Both are row operation.



In terms of elementary matrix, we have

$$\boldsymbol{E} = \{e_{21} = -2\}_{n\times n} \xrightarrow{\text{reverse}} \reflectbox{$\boldsymbol{E}$} = \{e_{21} = 2\}_{n\times n}$$

REVERSE OF A SEQUENCE OF ROW OPERATION

Let $\boldsymbol{Ax} = \boldsymbol{b}$ be a non-singular system of linear equations which is reduced to $\boldsymbol{Ux} = \boldsymbol{c}$ by elimination.

- The row operations converting $\boldsymbol{A}$ to $\boldsymbol{U}$ can be reversed
- The reverse is multiplication of $\boldsymbol{U}$ by a sequence of elementary matrices from left
- It is a matrix with the multipliers as elements

In the current example, we have

$$\boldsymbol{GFEA} = \boldsymbol{U} \Rightarrow \boldsymbol{FEA} = \rotatebox[origin=c]{180}{$\boldsymbol{G}$}\boldsymbol{U} \Rightarrow \boldsymbol{EA} = \rotatebox[origin=c]{180}{$\boldsymbol{F}$}\rotatebox[origin=c]{180}{$\boldsymbol{G}$}\boldsymbol{U} \Rightarrow \boldsymbol{A} = \rotatebox[origin=c]{180}{$\boldsymbol{E}$}\rotatebox[origin=c]{180}{$\boldsymbol{F}$}\rotatebox[origin=c]{180}{$\boldsymbol{G}$}\boldsymbol{U}$$

Thus, the reverse matrix is

$$\rotatebox[origin=c]{180}{$\boldsymbol{E}$}\rotatebox[origin=c]{180}{$\boldsymbol{F}$}\rotatebox[origin=c]{180}{$\boldsymbol{G}$} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}$$

TRIANGULAR FACTORIZATION: LU DECOMPOSITION

Let $\boldsymbol{Ax} = \boldsymbol{b}$ be a non-singular system of linear equations which is reduced to $\boldsymbol{Ux} = \boldsymbol{c}$ by elimination. Then

- $\boldsymbol{A} = \boldsymbol{LU}$
- $\boldsymbol{L}$ is lower-triangular containing the multipliers
- $\boldsymbol{U}$ is upper-triangular with the pivots as diagonal elements

$$\begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix} = \begin{bmatrix} 1 & & \\ m_{21} & 1 & \\ m_{31} & m_{32} & 1 \end{bmatrix} \begin{bmatrix} p_1 & * & * \\ & p_2 & * \\ & & p_3 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix}$$

TRIANGULAR FACTORIZATION: LDU DECOMPOSITION

Let $\boldsymbol{Ax} = \boldsymbol{b}$ be a non-singular system of linear equations which is reduced to $\boldsymbol{Ux} = \boldsymbol{c}$ by elimination. Then

- $\boldsymbol{A} = \boldsymbol{LDU'}$
- $\boldsymbol{L}$ is unit lower-triangular containing the multipliers
- $\boldsymbol{D}$ is diagonal with the pivots
- $\boldsymbol{U'}$ is unit upper-triangular

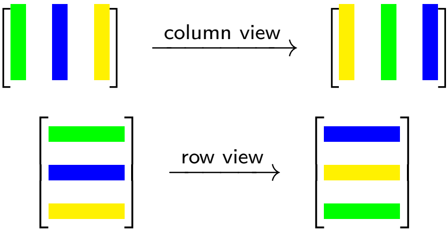
LDU decomposition is achieved by converting $\boldsymbol{LU}$ to $\boldsymbol{LDU'}$. For example

$$\begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & -8 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & \frac{1}{4} \\ 0 & 0 & 1 \end{bmatrix}$$

Definition (permutation matrix)

A permutation matrix results from row permutation(s) or column permutation(s) of an identity matrix.

$$\boldsymbol{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{\text{permutation}} \boldsymbol{P} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$



Note $\boldsymbol{P}$ has exactly a 1 in each row and in each column.

Multiplication by permutation matrix

Let $\boldsymbol{P}$ be a permutation matrix and $\boldsymbol{A}$ be square.

- $\boldsymbol{P}\boldsymbol{A}$ is row permutation of $\boldsymbol{A}$
- $\boldsymbol{A}\boldsymbol{P}$ is column permutation of $\boldsymbol{A}$

$$\boldsymbol{P}\,\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{a}_{1:} \\ \boldsymbol{a}_{2:} \\ \boldsymbol{a}_{3:} \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_{1:} \\ \boldsymbol{a}_{3:} \\ \boldsymbol{a}_{2:} \end{bmatrix}$$

$$\boldsymbol{A}\,\boldsymbol{P} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_3 & \boldsymbol{a}_2 \end{bmatrix}$$

LU decomposition with permutation

Let $\boldsymbol{Ax} = \boldsymbol{b}$ be a non-singular system of linear equations which is reduced to $\boldsymbol{Ux} = \boldsymbol{c}$ by elimination and row exchanges. Then

$$\boldsymbol{PA} = \boldsymbol{LU}$$

where

- $\boldsymbol{P}$ is a permutation matrix
- $\boldsymbol{L}$ is a lower-triangular matrix
- $\boldsymbol{U}$ is an upper-triangular matrix

Matrix Inverse and Transpose

Definition (matrix inverse)

Let $\boldsymbol{A}$ be a square matrix of order $n \times n$. A matrix $\boldsymbol{B}$ is an inverse of $\boldsymbol{A}$ if

$$\boldsymbol{AB} = \boldsymbol{I}_n = \boldsymbol{BA}$$

A matrix is invertible if it has an inverse matrix.

Theorem (uniqueness of matrix inverse)

Let $\boldsymbol{A}$ be an invertible matrix.

- *The inverse matrix of $\boldsymbol{A}$ is unique*
- *It is denoted by $\boldsymbol{A}^{-1}$*
- *The inverse matrix of $\boldsymbol{A}^{-1}$ is $\boldsymbol{A}$*

Proof. Let $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ be inverses of $\boldsymbol{A}$.

$$\begin{aligned}\boldsymbol{B}_1\boldsymbol{A}\boldsymbol{B}_2 = \boldsymbol{B}_1\boldsymbol{A}\boldsymbol{B}_2 &\Rightarrow (\boldsymbol{B}_1\boldsymbol{A})\,\boldsymbol{B}_2 = \boldsymbol{B}_1\,(\boldsymbol{A}\boldsymbol{B}_2)\\ &\Rightarrow \boldsymbol{B}_2 = \boldsymbol{B}_1\end{aligned}$$

Theorem (invertible and non-singular)

A matrix is invertible if and only if it is non-singular.

Proof.

- Suppose $\boldsymbol{A}$ is invertible. Consider system $\boldsymbol{Ax} = \boldsymbol{b}$. Multiplying both sides by $\boldsymbol{A}^{-1}$, we obtain $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$, which is a unique solution. So $\boldsymbol{A}$ is non-singular.
- Suppose $\boldsymbol{A}$ is non-singular. Consider systems

$$\boldsymbol{Ax} = \boldsymbol{i}_1, \ldots, \boldsymbol{Ax} = \boldsymbol{i}_n$$

  where $\boldsymbol{i}_1, \ldots, \boldsymbol{i}_n$ are standard unit vectors. Construct matrix $\boldsymbol{B}$ using the solutions as columns. Then $\boldsymbol{AB} = \boldsymbol{I}_n$. So $\boldsymbol{B}$ is an inverse of $\boldsymbol{A}$, and $\boldsymbol{A}$ is invertible.

Inverse of Matrix Product

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be invertible and $\boldsymbol{AB}$ not be undefined. Then

$$(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$$

Proof.

$$\begin{aligned}(\boldsymbol{AB})\left(\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}\right) &= \boldsymbol{ABB}^{-1}\boldsymbol{A}^{-1} \\ &= \boldsymbol{A}\left(\boldsymbol{BB}^{-1}\right)\boldsymbol{A}^{-1} \\ &= \boldsymbol{AI}_n\boldsymbol{A}^{-1} \\ &= \boldsymbol{AA}^{-1} \\ &= \boldsymbol{I}_n\end{aligned}$$

So

$$(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$$

Inverse of a triangular matrix

Let $\boldsymbol{A}$ be an invertible matrix.

- If $\boldsymbol{A}$ is lower-triangular, $\boldsymbol{A}^{-1}$ is lower-triangular
- If $\boldsymbol{A}$ is upper-triangular, $\boldsymbol{A}^{-1}$ is upper-triangular
- If $\boldsymbol{A}$ is diagonal, $\boldsymbol{A}^{-1}$ is diagonal

This can be shown by constructing $\boldsymbol{A}^{-1}$, one row by one row. Details are omitted.

Inverse of a unit triangular matrix

Let $\boldsymbol{A}$ be an invertible matrix.

- If $\boldsymbol{A}$ is unit lower-triangular, $\boldsymbol{A}^{-1}$ is unit lower-triangular
- If $\boldsymbol{A}$ is unit upper-triangular, $\boldsymbol{A}^{-1}$ is unit upper-triangular

This can be shown by row-by-row construction of $\boldsymbol{A}^{-1}$. Details are omitted.

Column-by-column computation of matrix inverse

Let $\boldsymbol{A}$ be an invertible matrix of order $n \times n$.

- $\boldsymbol{A}^{-1}$ exists and

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}_n$$

- Explicitly

$$\boldsymbol{A}\overbrace{\begin{bmatrix} \boldsymbol{u} & \ldots & \boldsymbol{z} \end{bmatrix}}^{\boldsymbol{A}^{-1}} = \begin{bmatrix} \boldsymbol{A}\boldsymbol{u} & \ldots & \boldsymbol{A}\boldsymbol{z} \end{bmatrix} = \overbrace{\begin{bmatrix} \boldsymbol{i}_1 & \ldots & \boldsymbol{i}_n \end{bmatrix}}^{\boldsymbol{I}_n}$$

- Columns $\boldsymbol{u}, \ldots, \boldsymbol{z}$ of $\boldsymbol{A}^{-1}$ can be found by solving systems

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{i}_1, \ \ldots, \ \boldsymbol{A}\boldsymbol{x} = \boldsymbol{i}_n$$

Gauss-Jordan method

Let $\boldsymbol{A}$ be invertible. The Gauss-Jordan method

- finds the columns of $\boldsymbol{A}^{-1}$ simultaneously
- is a sequence of row operations and scalar multiplications

Suppose

$$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$$

Gauss-Jordan method

$$\left[\boldsymbol{A} \,|\, \boldsymbol{I}\right] \xrightarrow{\text{row operations}} \left[\boldsymbol{U} \,|\, \boldsymbol{L}^{-1}\right]$$

$$\xrightarrow{\text{row operations}^*} \left[\boldsymbol{I} \,|\, \boldsymbol{U}^{-1}\boldsymbol{L}^{-1}\right] = \left[\boldsymbol{I} \,|\, \boldsymbol{A}^{-1}\right]$$

Example (Gauss-Jordan method)

$$\boldsymbol{A} = \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix}$$

$$\overbrace{\left[\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 4 & -6 & 0 & 0 & 1 & 0 \\ -2 & 7 & 2 & 0 & 0 & 1 \end{array}\right]}^{[\boldsymbol{A} \mid \boldsymbol{I}]} \xrightarrow{\text{row ops}} \overbrace{\left[\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 0 & -8 & -2 & -2 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right]}^{[\boldsymbol{U} \mid \boldsymbol{L}^{-1}]}$$

$$\xrightarrow{\text{row ops}^*} \underbrace{\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{3}{4} & -\frac{5}{16} & -\frac{3}{8} \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{3}{8} & -\frac{1}{4} \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right]}_{[\boldsymbol{I} \mid \boldsymbol{U}^{-1}\boldsymbol{L}^{-1}]} = [\,\boldsymbol{I} \mid \boldsymbol{A}^{-1}]$$

Gauss-Jordan with row exchange

Let $\boldsymbol{A}$ be a non-singular matrix whose LU decomposition requires permutation

$$\boldsymbol{PA} = \boldsymbol{LU}$$

The Gauss-Jordan method still finds the inverse of $\boldsymbol{A}$

$$
\begin{aligned}
\left[\boldsymbol{A} \,|\, \boldsymbol{I}\right] &\xrightarrow{\text{row exchanges}} \left[\boldsymbol{P}\,\boldsymbol{A} \,|\, \boldsymbol{P}\right] \\
&\xrightarrow{\text{row operation}^*} \left[\boldsymbol{I} \,|\, \boldsymbol{U}^{-1}\,\boldsymbol{L}^{-1}\,\boldsymbol{P}\right] \\
&= \left[\boldsymbol{I} \,|\, \boldsymbol{A}^{-1}\,\boldsymbol{P}^{-1}\,\boldsymbol{P}\right] \\
&= \left[\boldsymbol{I} \,|\, \boldsymbol{A}^{-1}\right]
\end{aligned}
$$

Definition (matrix transpose)

The transpose of a matrix $\boldsymbol{A}$ is denoted by $\boldsymbol{A}^T$. It is defined by

$$\left(\boldsymbol{A}^T\right)_{ji} = (\boldsymbol{A})_{ij}$$

- Taking transpose flips a matrix with respect to its diagonal
- The transpose of an $m \times n$ matrix is an $n \times m$ matrix
- For example

$$\begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \end{bmatrix}^T = \begin{bmatrix} 2 & 4 \\ 1 & -6 \\ 1 & 0 \end{bmatrix}$$

- The transpose of a lower-triangular matrix is upper-triangular
- The transpose of a diagonal matrix is diagonal

Transpose of Matrix Product

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be matrices and $\boldsymbol{AB}$ not be undefined. Then

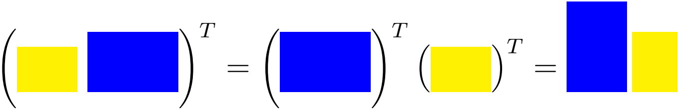
$$(\boldsymbol{AB})^T = \boldsymbol{B}^T\boldsymbol{A}^T$$

Proof.

$$\begin{aligned}\left((\boldsymbol{AB})^T\right)_{ij} &= (\boldsymbol{AB})_{ji} = \sum_{k=1}^{l} a_{jk}b_{ki} = \sum_{k=1}^{l}\left(\boldsymbol{B}^T\right)_{ik}\left(\boldsymbol{A}^T\right)_{kj}\\ &= \left(\boldsymbol{B}^T\boldsymbol{A}^T\right)_{ij}\end{aligned}$$

Thus

$$(\boldsymbol{AB})^T = \boldsymbol{B}^T\boldsymbol{A}^T$$



Transpose and Inverse

Let $\boldsymbol{A}$ be a square matrix of order $n \times n$. Then

$$\left(\boldsymbol{A}^{-1}\right)^T = \left(\boldsymbol{A}^T\right)^{-1} = \boldsymbol{A}^{-T}$$

$$\begin{aligned} \boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}_n \;\Rightarrow\;& \left(\boldsymbol{A}\boldsymbol{A}^{-1}\right)^T = \boldsymbol{I}_n \\ \Rightarrow\;& \left(\boldsymbol{A}^{-1}\right)^T \boldsymbol{A}^T = \boldsymbol{I}_n \\ \Rightarrow\;& \left(\boldsymbol{A}^{-1}\right)^T = \left(\boldsymbol{A}^T\right)^{-1} \end{aligned}$$

Definition (symmetric matrix)

Let $\boldsymbol{A}$ be a matrix. $\boldsymbol{A}$ is symmetric if $\boldsymbol{A}^T = \boldsymbol{A}$.

- A symmetric matrix must be square
- The product of a matrix and its transpose is symmetric

$$\left(\boldsymbol{R}^T\boldsymbol{R}\right)^T = \boldsymbol{R}^T\left(\boldsymbol{R}^T\right)^T = \boldsymbol{R}^T\boldsymbol{R}$$

$$\left(\boldsymbol{R}\boldsymbol{R}^T\right)^T = \left(\boldsymbol{R}^T\right)^T\boldsymbol{R}^T = \boldsymbol{R}\boldsymbol{R}^T$$

Differential Equations with Linear Algebra

Example (differential equations in physics)

- Newton's second law of motion

$$\boldsymbol{f} = m\boldsymbol{a} = m\frac{d^2\boldsymbol{r}}{dt^2}$$

- Maxwell's equations

$$\begin{cases} \nabla \cdot \boldsymbol{E} = \frac{\rho}{\epsilon_0} \\ \nabla \cdot \boldsymbol{B} = 0 \\ \nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} \\ \nabla \times \boldsymbol{B} = \mu_0(\boldsymbol{J} + \epsilon_0 \frac{\partial \boldsymbol{E}}{\partial t}) \end{cases}$$

- Schrödinger equation

$$i\hbar\frac{d\boldsymbol{\psi}}{dt} = \boldsymbol{H}\boldsymbol{\psi}$$

Ordinary Differential Equation

Consider an ordinary differential equation (abbr. ODE)

$$-\frac{d^2u(x)}{dx^2} = f(x), \quad 0 \le x \le 1$$

with boundary condition

$$u(0) = 0,\ u(1) = 0$$

- $u(x)$ is an unknown function
- $f(x)$ is a given function

DISCRETIZATION

We solve discretized version of ODE for approximate solution.

- only look at n discrete points

$$x_i = ih = i\left(\frac{1}{n+1}\right),\ i = 1, \ldots, n$$

- large $n \quad \Rightarrow \quad$ small $h \quad \Rightarrow \quad$ fine granularity
- introduce unknowns

$$u_i = u(x_i),\ i = 1, \ldots, n$$

- denote values

$$f_i = f(x_i),\ i = 1, \ldots, n$$

APPROXIMATION FOR DERIVATIVES

1st-order

$$\left.\frac{du}{dx}\right|_{x=x_i} = \lim_{\Delta x \to 0} \frac{u(x_i + \Delta x) - u(x_i)}{\Delta x}$$

$$\approx \frac{u(x_{i+1}) - u(x_i)}{x_{i+1} - x_i}$$

$$\approx \frac{u_{i+1} - u_i}{h}$$

2nd-order

$$\left.\frac{d^2u}{dx^2}\right|_{x=x_i} = \frac{d}{dx}\left(\frac{du}{dx}\right)\Bigg|_{x=x_i}$$

$$\approx \frac{\frac{du}{dx}(x_i) - \frac{du}{dx}(x_{i-1})}{h}$$

$$\approx \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2}$$

Recall the ODE

$$-\frac{d^2u(x)}{dx^2} = f(x), \quad 0 \le x \le 1,\ u(0) = 0,\ u(1) = 0$$

For the discretized version, we have a system of linear equations.

- n unknowns: $u(x)$ at the discrete points

$$u_1, \ldots, u_n$$

- n equations: approximate derivatives at the discrete points

$$\begin{aligned} -\frac{d^2u(x_i)}{dx^2} &= f(x_i) \\ \Rightarrow \quad -\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} &= f_i \\ \Rightarrow \quad -u_{i+1} + 2u_i - u_{i-1} &= h^2 f_i \end{aligned}$$

Example (Solve ODE with linear algebra)

Suppose $f(x) = 6x$ and $n = 5$. We have

$$h = \frac{1}{n+1} = \frac{1}{6}, x_i = ih = \frac{i}{6}, u_i = u(x_i), f_i = f(x_i) = 6x_i = i$$

- The first equation $(i = 1)$ is

$$-u_2 + 2u_1 - u_0 = h^2 f_1$$

- The system is

$$\begin{bmatrix} 2 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \end{bmatrix} = h^2 \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \end{bmatrix} = \frac{1}{36} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}$$