Linear Algebra

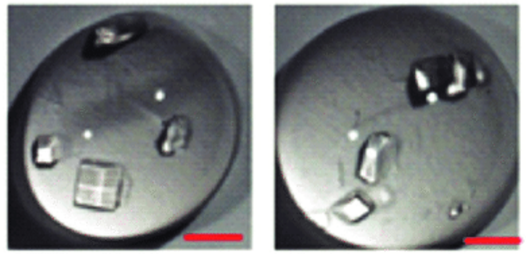
Liang Zheng
Australian National University
liang.zheng@anu.edu.au

AlphaFold2

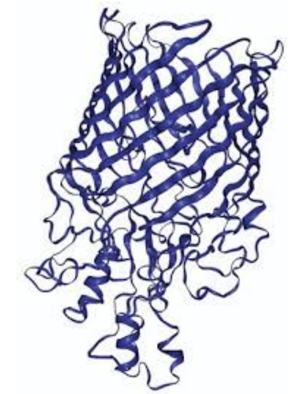
- Protein is a chain of amino acids and has complex 3D structures.

What structural biologists do:

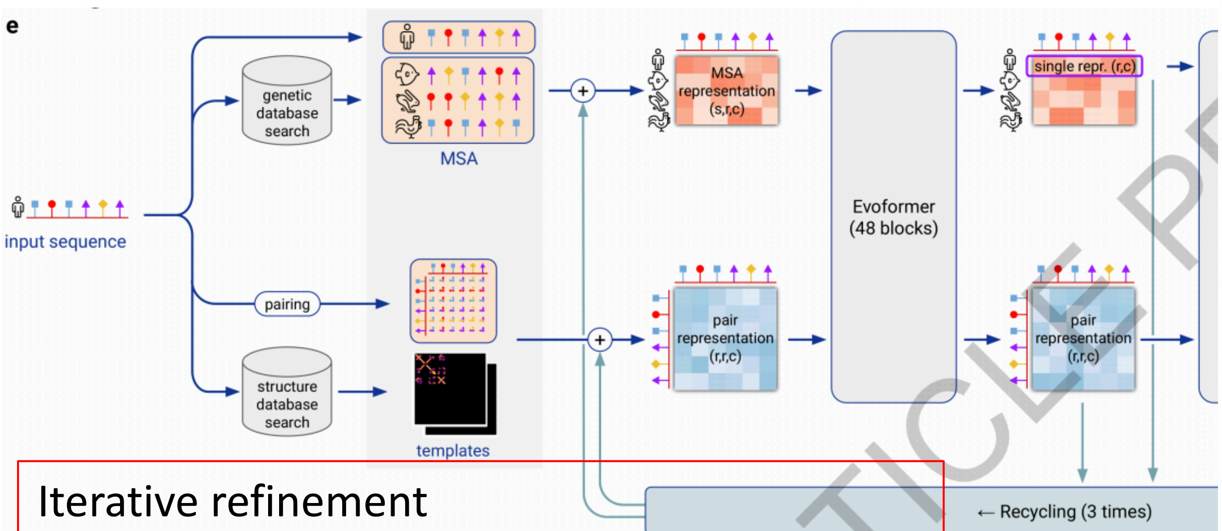
Working hard to find a crystallization condition

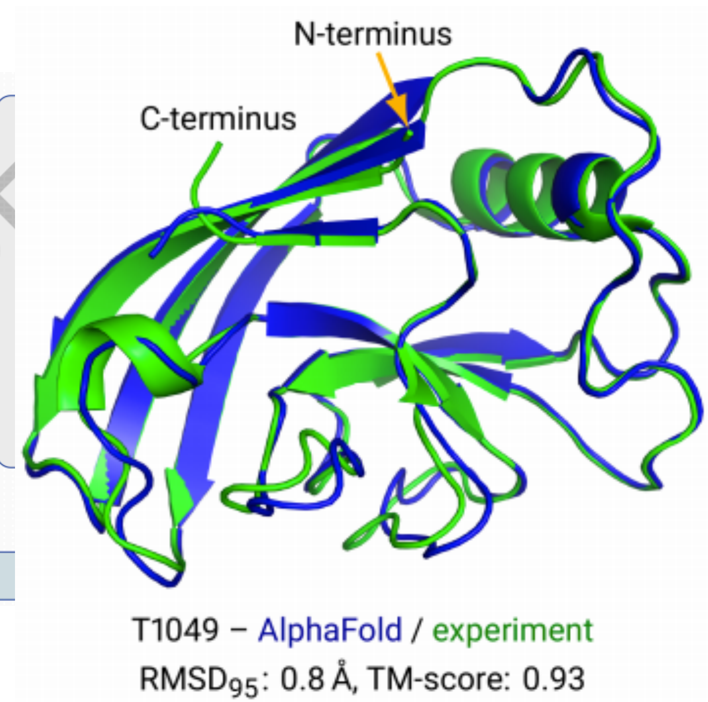


Cryogenic electron microscopy



What AlphaFold2 does:





Iterative refinement
Widely used attention
Noise student training
Noise added to input – robust model training

Jumper et al., Highly accurate protein structure prediction with AlphaFold, Nature, 2021

Vectors

- A simple example of vector, an element of $\mathbb{R}^n$

$$\boldsymbol{x} = \begin{bmatrix} 1 \\ 3 \\ 2 \\ \vdots \\ -1 \end{bmatrix} \in \mathbb{R}^n \qquad \boldsymbol{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \in \mathbb{R}^3$$

- Adding two vectors (component wise) $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$:

$$\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{c} \in \mathbb{R}^n$$

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \begin{bmatrix} -2 \\ 5 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 7 \\ 3 \end{bmatrix}$$

- Multiplying $\boldsymbol{a} \in \mathbb{R}^n$ by $\lambda \in \mathbb{R}$ results in a scaled vector:

$$\lambda \boldsymbol{a} \in \mathbb{R}^n$$

2.1 Systems of Linear Equations

- Examples

$$-x_1 + x_2 + 3x_3 = 3 \qquad (1)$$
$$x_1 + x_2 + 2x_3 = 2 \qquad (2)$$
$$2x_2 + 5x_3 = 1 \qquad (3)$$

3 unknowns

$x_1 \quad x_2 \quad x_3$

Does it have solution? No

Adding the first two equations yields $2x_2 + 5x_3 = 5$.
It contradicts Equation (3)

2.1 Systems of Linear Equations

- Examples

$$
\begin{aligned}
x_1 + x_2 &= 1 \qquad (1)\\
x_1 - x_2 &= 3 \qquad (2)
\end{aligned}
$$

$$
\boxed{\begin{aligned} x_1 &= 2 \\ x_2 &= -1 \end{aligned}}
$$

Does it have solution? Yes, it has a unique solution (2,-1)

$$
\begin{aligned}
x_1 + x_2 + x_3 &= 0 \qquad (1)\\
x_1 + x_2 + 2x_3 &= 2 \qquad (2)\\
+3x_3 &= 6 \qquad (3)
\end{aligned}
$$

$$
\boxed{\begin{aligned} x_3 &= 2 \\ x_1 + x_2 &= -2 \end{aligned}}
$$

Does it have solution? Yes, it has infinitely many solutions

2.2 Matrices

- A rectangular scheme consisting of m rows and n columns:

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \qquad a_{ij} \in \mathbb{R}.$$

ith row, jth column

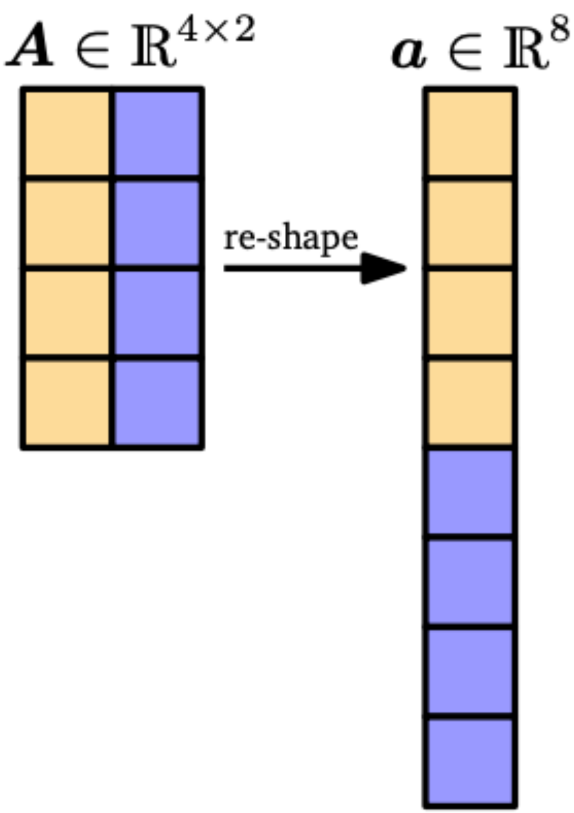
- By convention $(1, n)$-matrices are called rows and $(m, 1)$-matrices are called columns. These special matrices are also called row/column vectors.

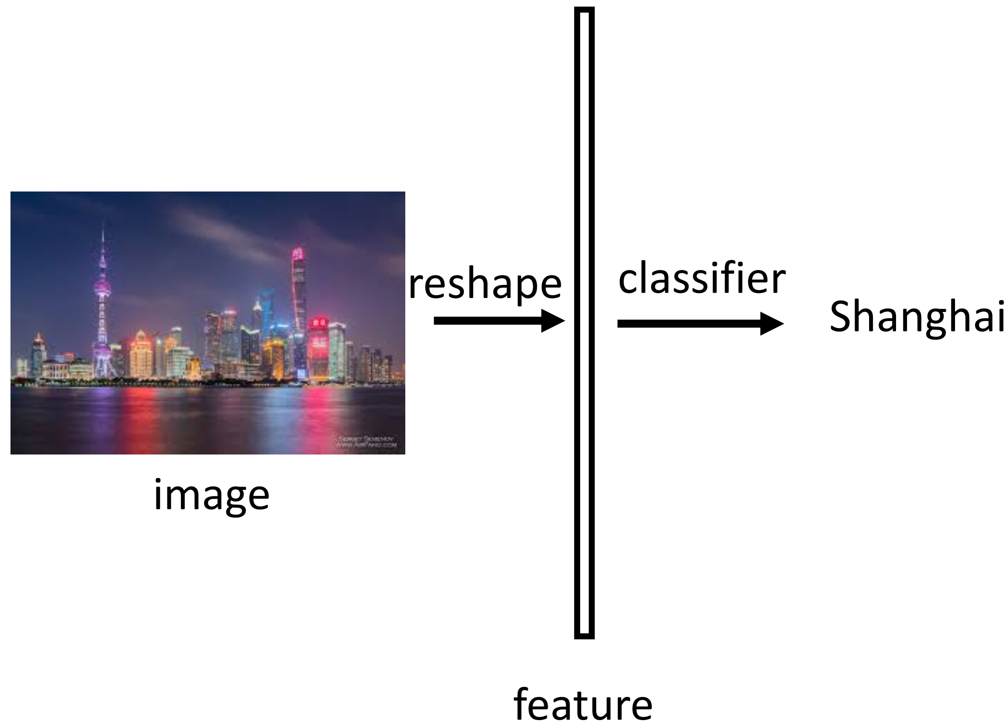
2.2 Matrices

#rows, #cols

- $\mathbb{R}^{m \times n}$ is the set of all real-valued (m, n)-matrices.

Space





Perhaps the simplest feature representation

Matrix - example

RGB image 5466

3244

5466x3244x3

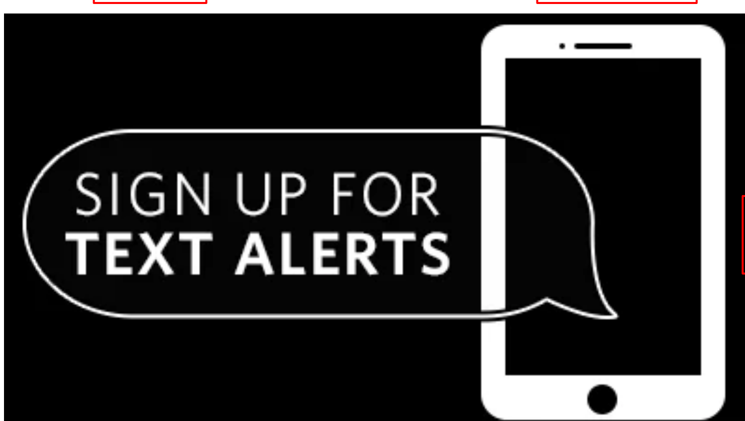
$[0, 255]$

640 Gray scale image

427

640x427

400 Binary image $\{0, 1\}$



255

400x255

Summary statistics table

	TREATMENT					
	A			B		
	N	Mean	SD	N	Mean	SD
LENGTH	50	176.500	5.9083	50	175.640	5.5467
WEIGHT	50	77.680	10.6492	50	76.400	8.4540
Body Mass Index	50	24.918	3.0644	50	24.763	2.4787

3

6

3x6

2.2.1 Matrix Addition and Multiplication

- The sum of two matrices $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{B} \in \mathbb{R}^{m\times n}$ is defined as the element-wise sum,

$$\boldsymbol{A} + \boldsymbol{B} := \begin{bmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{bmatrix} \in \mathbb{R}^{m\times n}$$

- Example

For $\boldsymbol{A} = \begin{bmatrix} 0 & 1 \\ 1 & 2 \\ 3 & -2 \end{bmatrix} \in \mathbb{R}^{3\times 2}$, $\boldsymbol{B} = \begin{bmatrix} -5 & 0 \\ 1 & 1 \\ 0 & -4 \end{bmatrix} \in \mathbb{R}^{3\times 2}$, we obtain

$$\boldsymbol{A} + \boldsymbol{B} = \begin{bmatrix} -5 & 1 \\ 2 & 3 \\ 3 & -6 \end{bmatrix} \in \mathbb{R}^{3\times 2}$$

2.2.1 Matrix Multiplication

- Example

For $\boldsymbol{A} = \begin{bmatrix} -1 & 1 \\ 0 & 2 \\ 1 & 3 \end{bmatrix} \in \mathbb{R}^{3\times 2}$, $\boldsymbol{B} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \end{bmatrix} \in \mathbb{R}^{2\times 3}$, we obtain

$-1\times 1 + 1\times 0$ $\quad$ $-1\times 3 + 1\times 1$

$$\boldsymbol{AB} = \begin{bmatrix} -1 & 1 \\ 0 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \end{bmatrix} = \begin{bmatrix} \boxed{-1} & 0 & \boxed{-2} \\ 0 & 4 & 2 \\ 1 & 8 & 6 \end{bmatrix} \in \mathbb{R}^{3\times 3}$$

$$\boldsymbol{BA} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 0 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 14 \\ 1 & 7 \end{bmatrix} \in \mathbb{R}^{2\times 2}$$

2.2.1 Matrix Addition and Multiplication

- For matrices $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{B} \in \mathbb{R}^{n \times k}$, the element c_{ij} of the product $\boldsymbol{C} = \boldsymbol{AB} \in \mathbb{R}^{m \times k}$ is defined as

$$c_{ij} = \sum_{l=1}^{n} a_{il} b_{lj}\,,$$

$$c_{ij} \neq a_{ij} b_{ij}$$

$$i = 1, \ldots, m.\ j = 1, \ldots, k$$

- To compute element c_{ij} we multiply the elements of the ith row of $\boldsymbol{A}$ with the jth column of $\boldsymbol{B}$ and sum them up.

2.2.1 Matrix Addition and Multiplication

- One property that is unique to matrices is the dimension property. This property has two parts:

$$\underbrace{\boldsymbol{A}}_{n\times k}\ \underbrace{\boldsymbol{B}}_{k\times m} = \underbrace{\boldsymbol{C}}_{n\times m}$$

- Identity Matrix

$$\boldsymbol{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

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

$$\boldsymbol{I}_n := \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \in \mathbb{R}^{n\times n}$$

2.2.1 Matrix Addition and Multiplication

- Properties of matrices
- Associativity

$$\forall \boldsymbol{A} \in \mathbb{R}^{m \times n}, \boldsymbol{B} \in \mathbb{R}^{n \times p}, \boldsymbol{C} \in \mathbb{R}^{p \times q}: (\boldsymbol{A}\boldsymbol{B})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{B}\boldsymbol{C})$$

- Distributivity

$$\forall \boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{m \times n}, \boldsymbol{C}, \boldsymbol{D} \in \mathbb{R}^{n \times p}: (\boldsymbol{A} + \boldsymbol{B})\boldsymbol{C} = \boldsymbol{A}\boldsymbol{C} + \boldsymbol{B}\boldsymbol{C}$$
$$\boldsymbol{A}(\boldsymbol{C} + \boldsymbol{D}) = \boldsymbol{A}\boldsymbol{C} + \boldsymbol{A}\boldsymbol{D}$$

- Multiplication with the identity matrix:

$$\forall \boldsymbol{A} \in \mathbb{R}^{m \times n}: \boldsymbol{I}_m \boldsymbol{A} = \boldsymbol{A}\boldsymbol{I}_n = \boldsymbol{A}$$

$\boldsymbol{I}_m \neq \boldsymbol{I}_n$ for $m \neq n$.

2.2.2 Inverse and Transpose

- Inverse: consider a square matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$. Let matrix $\boldsymbol{B} \in \mathbb{R}^{n \times n}$ have the property that $\boldsymbol{AB} = \boldsymbol{I}_n = \boldsymbol{BA}$. $\boldsymbol{B}$ is called the inverse of $\boldsymbol{A}$ and denoted by $\boldsymbol{A}^{-1}$.

- Example

$$\boxed{\boldsymbol{AB} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \boldsymbol{I}_2}$$

$$\boxed{\boldsymbol{B} = \boldsymbol{A}^{-1}}$$

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

The matrices are inverse to each other, because

$$\boldsymbol{AB} = \boldsymbol{I}_2 = \boldsymbol{BA}$$

2.2.2 Inverse and Transpose

- Transpose: For $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, the matrix $\boldsymbol{B} \in \mathbb{R}^{n \times m}$ with $b_{ij} = a_{ji}$ is called the transpose of $\boldsymbol{A}$. We write $\boldsymbol{B} = \boldsymbol{A}^{\mathrm{T}}$

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 2 \\ -1 & -5 & 6 \\ 0 & 1 & 3 \end{bmatrix}, \boldsymbol{A}^{\mathrm{T}} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & -5 & 1 \\ 2 & 6 & 3 \end{bmatrix}$$

- Important properties of inverses and transposes:

$$(\boldsymbol{A}+\boldsymbol{B})^{-1} \neq \boldsymbol{A}^{-1} + \boldsymbol{B}^{-1}$$

$$\boxed{\frac{1}{a+b} \neq \frac{1}{a} + \frac{1}{b}}$$

$$(\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}} = \boldsymbol{A}$$

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

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

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

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

$$\boxed{(\boldsymbol{A}\boldsymbol{B}) \cdot (\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}) = \boldsymbol{A}\boldsymbol{I}\boldsymbol{A}^{-1} = \boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}}$$

2.2.2 Inverse and Transpose

- Symmetric: A matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ is symmetric if $\boldsymbol{A} = \boldsymbol{A}^{\mathrm{T}}$

$$\boldsymbol{A} = \begin{bmatrix} 1 & -2 \\ -2 & 0 \end{bmatrix} \qquad \boldsymbol{A}^{\mathrm{T}} = \begin{bmatrix} 1 & -2 \\ -2 & 0 \end{bmatrix} \qquad \boxed{\boldsymbol{A} = \boldsymbol{A}^{\mathrm{T}}}$$

- The sum of symmetric matrices $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n \times n}$ is always symmetric.

$$\boldsymbol{A} + \boldsymbol{B} \overset{?}{=} (\boldsymbol{A} + \boldsymbol{B})^{\mathrm{T}}$$

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

- The product of two symmetric matrics is generally not symmetric

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

2.2.3 Multiplication by a Scalar

- A scalar $\lambda \in \mathbb{R}$
- Let $\boldsymbol{A} \in \mathbb{R}^{m \times n}$. Then $\lambda \boldsymbol{A} = \boldsymbol{K}$, where $k_{ij} = \lambda a_{ij}$

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 3 \\ 2 & 0 & -1 \end{bmatrix} \qquad \lambda = 1.5$$

$$\lambda \boldsymbol{A} = \begin{bmatrix} 1.5 & 0 & 4.5 \\ 3 & 0 & -1.5 \end{bmatrix}$$

2.2.3 Multiplication by a Scalar

- For $\lambda, \varphi \in \mathbb{R}$, there following properties hold:
- Associativity

$$(\lambda\varphi)\boldsymbol{C} = \lambda(\varphi\boldsymbol{C}), \quad \boldsymbol{C} \in \mathbb{R}^{m\times n}$$
$$\lambda\,(\boldsymbol{BC}) = (\lambda\boldsymbol{B})\boldsymbol{C} = \boldsymbol{B}(\lambda\boldsymbol{C}) = (\boldsymbol{BC})\lambda, \boldsymbol{B} \in \mathbb{R}^{m\times n}, \boldsymbol{C} \in \mathbb{R}^{n\times k}$$

- Transpose

$$(\lambda\boldsymbol{C})^{\mathrm{T}} = \boldsymbol{C}^{\mathrm{T}}\lambda^{\mathrm{T}} = \boldsymbol{C}^{\mathrm{T}}\lambda = \lambda\boldsymbol{C}^{\mathrm{T}}, \quad \boldsymbol{C} \in \mathbb{R}^{m\times n}$$

- Distributivity

$$(\lambda + \varphi)\boldsymbol{C} = \lambda\boldsymbol{C} + \varphi\boldsymbol{C}, \quad \boldsymbol{C} \in \mathbb{R}^{m\times n}$$
$$\lambda\,(\boldsymbol{B} + \boldsymbol{C}) = \lambda\boldsymbol{B} + \lambda\boldsymbol{C}, \quad \boldsymbol{B}, \boldsymbol{C} \in \mathbb{R}^{m\times n}$$

2.2.4 Compact Representations of Systems of Linear Equations

- Consider the system of linear equations,

$$\begin{aligned} 2x_1 + 3x_2 + 5x_3 &= 1 \\ 4x_1 - 2x_2 - 7x_3 &= 8 \\ 9x_1 + 5x_2 - 3x_3 &= 2 \end{aligned}$$

- Using matrix multiplication, we can write it into a compact form

$$\begin{bmatrix} 2 & 3 & 5 \\ 4 & -2 & -7 \\ 9 & 5 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 8 \\ 2 \end{bmatrix} \qquad \boldsymbol{Ax} = \boldsymbol{b}$$

2.3 Solving Systems of Linear Equations

- Now we have a general form of an equation system

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

- 2.3.1 Particular and General Solution

Step 1. Find a particular solution to $\boldsymbol{Ax} = \boldsymbol{b}$

Step 2. Find all solutions to $\boldsymbol{Ax} = \boldsymbol{0}$

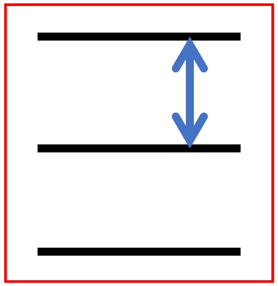
Step 3. Combine the solutions from steps 1. and 2. to the general solution

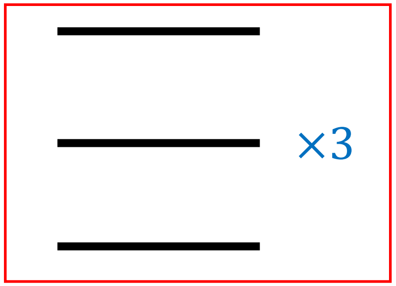
We use Gaussian elimination to solve the equation system

2.3.2 Elementary Transformations

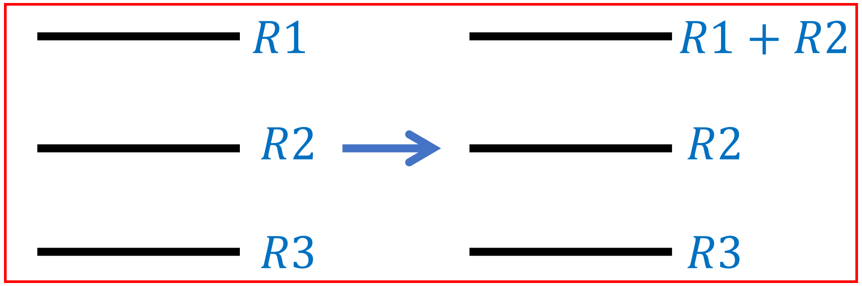
- Elementary transformations keep the solution set the same, but transform the equation system into a simpler form.

- Elementary transformations include:
- Exchange of two equations

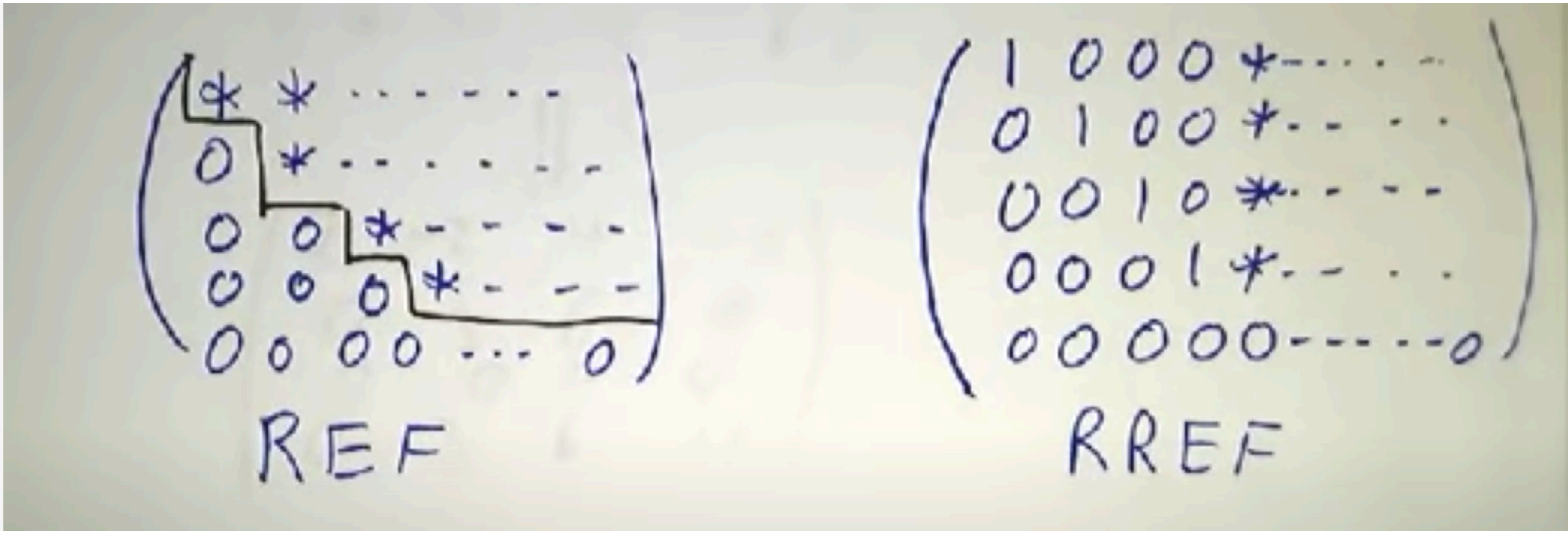




- Multiplication of an equation (row) with a constant $\lambda \in \mathbb{R}\backslash\{0\}$
- Addition of two equations (rows)



Row-echelon form (REF) and reduced row-echelon form (RREF)



From Lorenzo A. Sadun's teaching video

- Row Echelon Form
- All rows with 0s only are at the bottom
- A pivot is always strictly to the right of the pivot of the row above it

- Reduced Row Echelon Form
- Every pivot is 1
- The pivot is the only non-zero entry in its column

Gaussian Elimination - example

$$\begin{array}{rcrcrcr} x_1 & + & x_2 & - & x_3 & = & 9 \\ & & x_2 & + & 3x_3 & = & 3 \\ -x_1 & & & - & 2x_3 & = & 2 \end{array}$$

augmented matrix

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 9 \\ 0 & 1 & 3 & 3 \\ -1 & 0 & -2 & 2 \end{array}\right]$$

$\Big\downarrow$ R1+R3 -> R3

$$\begin{bmatrix} 1 & 1 & -1 & 9 \\ 0 & 1 & 3 & 3 \\ 0 & 1 & -3 & 11 \end{bmatrix} \xrightarrow[\text{R3-R2 -> R3}]{} \begin{bmatrix} 1 & 1 & -1 & 9 \\ 0 & 1 & 3 & 3 \\ 0 & 0 & -6 & 8 \end{bmatrix}$$

Gaussian Elimination - example

- Seek all solutions to the following system of equations

$$\begin{array}{rcrcrcrcl} 2x_1 & + & 3x_2 & - & 2x_3 & + & 5x_4 & = & 1 \\ x_1 & + & 2x_2 & - & x_3 & + & 3x_4 & = & 2 \\ -x_1 & - & 2x_2 & + & x_3 & - & x_4 & = & 4 \end{array}$$

$$\begin{bmatrix} 2 & 3 & -2 & 5 & 1 \\ 1 & 2 & -1 & 3 & 2 \\ -1 & -2 & 1 & -1 & 4 \end{bmatrix} \xrightarrow{\text{Swap R1 and R2}} \begin{bmatrix} 1 & 2 & -1 & 3 & 2 \\ 2 & 3 & -2 & 5 & 1 \\ -1 & -2 & 1 & -1 & 4 \end{bmatrix}$$

$$\xrightarrow[\text{R1+R3 -> R3}]{\text{R2-2R1 -> R2}} \begin{bmatrix} 1 & 2 & -1 & 3 & 2 \\ 0 & -1 & 0 & -1 & -3 \\ 0 & 0 & 0 & 2 & 6 \end{bmatrix}$$

row-echelon form (REF)

How to find the general solution to $\boldsymbol{Ax} = \boldsymbol{b}$

$$\begin{array}{rcrcrcrcr} 2x_1 & + & 3x_2 & - & 2x_3 & + & 5x_4 & = & 1 \\ x_1 & + & 2x_2 & - & x_3 & + & 3x_4 & = & 2 \\ -x_1 & - & 2x_2 & + & x_3 & - & x_4 & = & 4 \end{array}$$

Gaussian elimination $\rightsquigarrow$

$$\begin{array}{ccccc} \left[\begin{array}{ccccc} 1 & 2 & -1 & 3 & 2 \\ 0 & -1 & 0 & -1 & -3 \\ 0 & 0 & 0 & 2 & 6 \end{array}\right] \\ \begin{array}{ccccc} x_1 & x_2 & x_3 & x_4 & \hat{\boldsymbol{b}} \end{array} \end{array}$$

Step 1. Find a particular solution to $\boldsymbol{Ax} = \boldsymbol{b}$
Step 2. Find all solutions to $\boldsymbol{Ax} = \boldsymbol{0}$
Step 3. Combine the solutions from steps 1. and 2. to the general solution

Step 1: Finding a particular solution to $\boldsymbol{Ax} = \boldsymbol{b}$

Let free variables be 0, calculate the value of basic variables

$$\begin{array}{ccccc} & & & & \\ \left[\begin{matrix} 1 & 2 & -1 & 3 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & 0 & 2 \end{matrix}\right. & & & & \left.\begin{matrix} 2 \\ -3 \\ 6 \end{matrix}\right] \end{array}$$

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

$x_1 \quad x_2 \quad \underline{x_3} \quad x_4 \quad \widehat{\boldsymbol{b}}$

x_3: free variable

$x_1\ x_2\ x_4$: basic variables

$0 + 0 + 0 + 2x_4 = 6 \rightarrow x_4 = 3$

$0 - x_2 + 0 - x_4 = -3 \rightarrow x_2 = 0$

$x_1 + 2x_2 - x_3 + 3x_4 = 2$
$x_1 + 0 - 0 + 9 = 2$
$\rightarrow x_1 = -7$

A particular solution:

$$\begin{bmatrix} -7 \\ 0 \\ 0 \\ 3 \end{bmatrix}$$

Step 2: Find all solutions to $\boldsymbol{Ax} = \mathbf{0}$

Let one free variables be 1, and the rest free variables be 0, calculate the value of basic variables

$$
\begin{array}{c}
\begin{bmatrix} 1 & 2 & -1 & 3 & 0 \\ 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 2 & 0 \end{bmatrix} \\
\begin{array}{ccccc} x_1 & x_2 & x_3 & \boxed{x_4} & \mathbf{0} \end{array}
\end{array}
$$

We first immediately get $x_4 = 0$ from Row 3.

After setting $x_3 = 1$, we have
$0 - x_2 + 0 - x_4 = 0$,
$x_1 + 2x_2 - 1 + 3x_4 = 0$

$\longrightarrow$ all solutions to $\boldsymbol{Ax} = \mathbf{0}$: $\left\{ \boldsymbol{x} \in R^4 : \boldsymbol{x} = \lambda \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \lambda \in R \right\}$

Step 3: Combine the solutions from steps 1. and 2. to the general solution

all solutions to $\boldsymbol{Ax} = \boldsymbol{b}$: $\left\{ \boldsymbol{x} \in R^4 : \boldsymbol{x} = \begin{bmatrix} -7 \\ 0 \\ 0 \\ 3 \end{bmatrix} + \lambda \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \lambda \in R \right\}$

Another example

$$\begin{aligned}
-2x_1 + 4x_2 - 2x_3 - x_4 + 4x_5 &= -3\\
4x_1 - 8x_2 + 3x_3 - 3x_4 + x_5 &= 2\\
x_1 - 2x_2 + x_3 - x_4 + x_5 &= 0\\
x_1 - 2x_2 \qquad\quad - 3x_4 + 4x_5 &= a
\end{aligned}$$

$$\begin{bmatrix} -2 & 4 & -2 & -1 & 4 & -3 \\ 4 & -8 & 3 & -3 & 1 & 2 \\ 1 & -2 & 1 & -1 & 1 & 0 \\ 1 & -2 & 0 & -3 & 4 & a \end{bmatrix} \xrightarrow{\text{Swap R1 and R3}} \begin{bmatrix} 1 & -2 & 1 & -1 & 1 & 0 \\ 4 & -8 & 3 & -3 & 1 & 2 \\ -2 & 4 & -2 & -1 & 4 & -3 \\ 1 & -2 & 0 & -3 & 4 & a \end{bmatrix}$$

R2-4R1 -> R2
R3+2R1 -> R3
R4-R1 -> R4

$$\longrightarrow \begin{bmatrix} 1 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & -3 & 2 \\ 0 & 0 & 0 & -3 & 6 & -3 \\ 0 & 0 & -1 & -2 & 3 & a \end{bmatrix} \xrightarrow{\text{R4-R2 -> R4}} \begin{bmatrix} 1 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & -3 & 2 \\ 0 & 0 & 0 & -3 & 6 & -3 \\ 0 & 0 & 0 & -3 & 6 & a-2 \end{bmatrix}$$

$$\xrightarrow{\text{R4-R3 -> R4}} \begin{bmatrix} 1 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & -3 & 2 \\ 0 & 0 & 0 & -3 & 6 & -3 \\ 0 & 0 & 0 & 0 & 0 & a+1 \end{bmatrix}$$

a must equal to ~~1~~ −1 for this equation system to have solutions

Finding a particular solution to $\boldsymbol{Ax} = \boldsymbol{b}$

Let free variables be 0, calculate the value of basic variables

$$
\begin{array}{ccccc}
\left[\begin{array}{ccccc|c}
1 & -2 & 1 & -1 & 1 & 0 \\
0 & 0 & 1 & -1 & 3 & -2 \\
0 & 0 & 0 & 1 & -2 & 1 \\
0 & 0 & 0 & 0 & 0 & 0
\end{array}\right] \\
x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5
\end{array}
$$

It's already in the REF. We let x_2 and x_5 be 0.

$$x_4 - 2x_5 = 1 \longrightarrow x_4 = 1$$

$$x_3 - x_4 + 3x_5 = -2 \longrightarrow x_3 = -1$$

$$x_1 - 2x_2 + x_3 - x_4 + x_5 = 0 \longrightarrow x_1 = 2$$

A particular solution:

$$\begin{bmatrix} 2 \\ 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}$$

Find all solutions to $\boldsymbol{Ax} = \boldsymbol{0}$

- Let one free variables be 1, and the rest free variables be 0, calculate the value of basic variables

$$\begin{array}{c} \begin{bmatrix} 1 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 3 & 0 \\ 0 & 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ \begin{array}{cccccc} x_1 & x_2 & x_3 & x_4 & x_5 & \end{array} \end{array}$$

Let x_2 be 1 and x_5 be 0. We get $\begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

Let x_2 be 0 and x_5 be 1. We get $\begin{bmatrix} 2 \\ 0 \\ -1 \\ 1 \\ 1 \end{bmatrix}$

all solutions to $\boldsymbol{Ax} = \boldsymbol{0}$: $\left\{ \boldsymbol{x} \in R^5 : \lambda_1 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 2 \\ 0 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \lambda_1, \lambda_2 \in R \right\}$

How to find the general solution to $\boldsymbol{Ax} = \boldsymbol{b}$

- Step 3. Combine the solutions from steps 1. and 2. to the general solution

Step 1: $\boxed{\boldsymbol{Ax} = \boldsymbol{0}}$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}$$

Step 2: $\boxed{\boldsymbol{Ax} = \boldsymbol{b}}$

$$\left\{ \boldsymbol{x} \in R^5 : \boldsymbol{x} = \lambda_1 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 2 \\ 0 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \lambda_1, \lambda_2 \in R \right\}$$

General solution:

$$\left\{ \boldsymbol{x} \in R^5 : \boldsymbol{x} = \begin{bmatrix} 2 \\ 0 \\ -1 \\ 1 \\ 0 \end{bmatrix} + \lambda_1 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 2 \\ 0 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \lambda_1, \lambda_2 \in R \right\}$$

Proof

- Given $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ with $m < n$, then $\boldsymbol{Ax} = \mathbf{0}$ has infinitely many solutions
- Proof
- This system always has at least one solution since homogeneous
  - Consider $\boldsymbol{A}\mathbf{0} = \mathbf{0}$ always holds
- Matrix $\boldsymbol{A}$ brought in row echelon form contains at most m pivots.

For example, $\begin{bmatrix} \mathbf{1} & -2 & 0 & 0 & -2 \\ 0 & 0 & \mathbf{1} & 0 & 1 \\ 0 & 0 & 0 & \mathbf{1} & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \in \mathbb{R}^{4 \times 5}$

- There will have $n - m \geq 1$ non-pivot columns, or free variables. It means we can find at least one solution $\boldsymbol{x}^* \neq \mathbf{0}$. Then, $\lambda \boldsymbol{x}^*$, $\lambda \in \mathbb{R}$ are solutions to $\boldsymbol{Ax} = \mathbf{0}$.

Proof

- A system of linear equations $\boldsymbol{Ax} = \boldsymbol{b}$ either has no solutions, a unique solution or infinitely many solutions
- Proof

proof by contradiction

- Let's assume the system $\boldsymbol{Ax} = \boldsymbol{b}$ has two solutions $\boldsymbol{p}$ and $\boldsymbol{q}$.
- We have

$$\boldsymbol{Ap} = \boldsymbol{b} \qquad \boldsymbol{Aq} = \boldsymbol{b}$$

a form of proof that establishes the truth or the validity of a proposition, by showing that assuming the proposition to be false leads to a contradiction.

- Consider

$$\boldsymbol{v} = \boldsymbol{p} + t(\boldsymbol{q} - \boldsymbol{p}), t \in \mathbb{R}$$

- We have

$$\boldsymbol{Av} = \boldsymbol{A}\big(\boldsymbol{p} + t(\boldsymbol{q} - \boldsymbol{p})\big) = \boldsymbol{Ap} + t(\boldsymbol{Aq} - \boldsymbol{Ap}) = \boldsymbol{b} + t(\boldsymbol{b} - \boldsymbol{b}) = \boldsymbol{b}$$

- We thus have infinitely many solutions (by varying t)

Calculating the Inverse with Gaussian Elimination

- To compute the inverse $\boldsymbol{A}^{-1}$ of $\boldsymbol{A} \in \mathbb{R}^{n \times n}$,
- We need to find a matrix $\boldsymbol{X}$ that satisfies $\boldsymbol{AX} = \boldsymbol{I_n}$.
- Then, $\boldsymbol{X} = \boldsymbol{A}^{-1}$.
- We can write this down as a set of simultaneous linear equations $\boldsymbol{AX} = \boldsymbol{I_n}$, where we solve for $\boldsymbol{X} = [\boldsymbol{x_1}|\cdots|\boldsymbol{x_n}]$
- We use the augmented matrix notation and use Gaussian Elimination.

$$[\boldsymbol{A} \mid \boldsymbol{I_n}] \rightsquigarrow \cdots \rightsquigarrow [\boldsymbol{I_n} \mid \boldsymbol{A}^{-1}]$$

Calculating the Inverse with Gaussian Elimination

$$[A \mid I_n] \rightsquigarrow \cdots \rightsquigarrow [I_n \mid A^{-1}]$$

Calculating the Inverse with Gaussian Elimination

- Example: determine the inverse of

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \in R^4$$

- First, write down the augmented matrix

$$\boldsymbol{A} = \left[\begin{array}{cccc|cccc} 1 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\boldsymbol{A} \qquad\qquad\qquad \boldsymbol{I_n}$$

Calculating the Inverse with Gaussian Elimination

- Use Gaussian elimination to bring it into reduced row-echelon form (RREF)

REF

$$A = \left[\begin{array}{cccc|cccc} 1 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \rightsquigarrow \cdots \rightsquigarrow A = \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -1 & 2 & -2 & 2 \\ 0 & 1 & 0 & 0 & 1 & -1 & 2 & -2 \\ 0 & 0 & 1 & 0 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 1 & -1 & 0 & -1 & 2 \end{array}\right]$$

A^{-1}

- The desired inverse is given as its right-hand side

$$A^{-1} = \begin{bmatrix} -1 & 2 & -2 & 2 \\ 1 & -1 & 2 & -2 \\ 1 & -1 & 1 & -1 \\ -1 & 0 & -1 & 2 \end{bmatrix}$$

Calculating Reduced Row-echelon form - example

$$\begin{bmatrix} 2 & -2 & 4 & -2 \\ 2 & 1 & 10 & 7 \\ -4 & 4 & -8 & 4 \\ 4 & -1 & 14 & 6 \end{bmatrix} \xrightarrow[\text{R4-2R1 -> R4}]{\text{R2-R1 -> R2, R3+2R1 -> R3}} \begin{bmatrix} 2 & -2 & 4 & -2 \\ 0 & 3 & 6 & 9 \\ 0 & 0 & 0 & 0 \\ 0 & 3 & 6 & 10 \end{bmatrix} \xrightarrow[\text{R3 and R4}]{\text{Swap}} \begin{bmatrix} 2 & -2 & 4 & -2 \\ 0 & 3 & 6 & 9 \\ 0 & 3 & 6 & 10 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{\text{R3-R2 -> R3}} \underbrace{\begin{bmatrix} 2 & -2 & 4 & -2 \\ 0 & 3 & 6 & 9 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}}_{\boxed{\text{REF}}} \xrightarrow[\text{by scalar}]{\text{Multiplication}} \begin{bmatrix} 1 & -1 & 2 & -1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

This matrix is not invertible

$$\xrightarrow[\text{R2-3R3 -> R2}]{\text{R1+R3 -> R1}} \begin{bmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{\text{R1+R2 -> R1}} \underbrace{\begin{bmatrix} 1 & 0 & 4 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}}_{\boxed{\text{RREF}}} \neq \boldsymbol{I}_4$$

Moore-Penrose pseudo-inverse

- We can calculate $\boldsymbol{A}^{-1}$only when $\boldsymbol{A}$ is a square matrix and is invertible
- Otherwise, under mild conditions, we can use the following pseudo-inverse:

$$\boldsymbol{A}x = \boldsymbol{b} \Leftrightarrow \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}x = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b} \Leftrightarrow x = \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathrm{T}}\,\boldsymbol{b}$$

- $\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathrm{T}}$ is the Moore-Penrose pseudo-inverse of $\boldsymbol{A}$

Check your understanding

- Which of the following are correct?

(A) A vector, when multiplied by a scale, is still a vector. True

(B) For a system of linear equations with n variables, it is possible that none of them are free variables. True

(C) For a system of linear equations with n variables, the maximum number of pivots in the REF is $n-1$. False

(D) A matrix, when added by an identity matrix, stays as is. False

(E) We can use matrix transpose in Gaussian Elimination. False

(F) Two arbitrary matrices can be multiplied False

(G) Two arbitrary matrices can be added. False

(H) An image with black borders is not a matrix.

False

Check your understanding

- Let A, B, C be 2x2 matrices.
- Which of the following are equivalent to A(B+C)?
  - AB+AC
  - BA+CA
  - A(C+B)
  - (B+C)A

- Which of the following expressions are equivalent to $I_2(AB)$?
  - AB
  - BA
  - $(AB)I_2$
  - $(BA)I_2$