

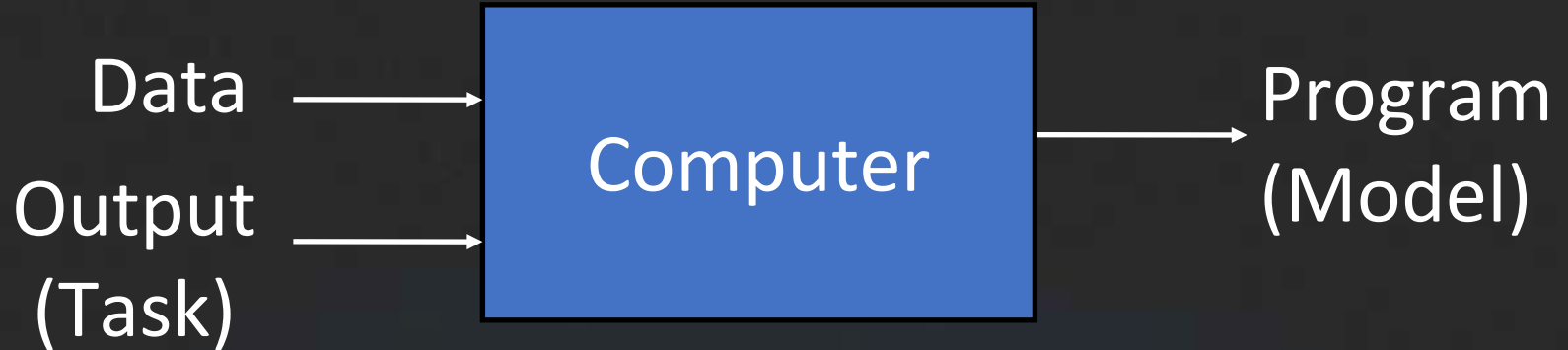
Linear Algebra

Textbook Sec. 2.1 - 2.3

Rahul Shome

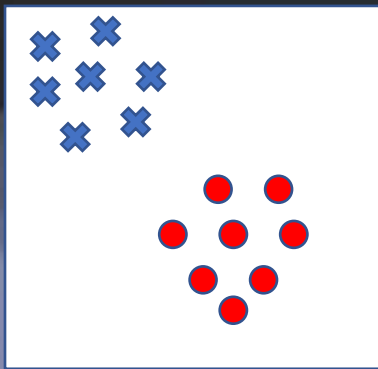
based on content by Jo Ciucă
Australian National University

Quick review

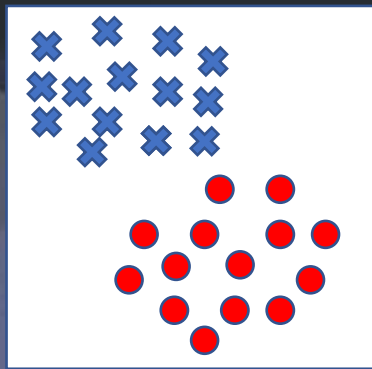


Quick review

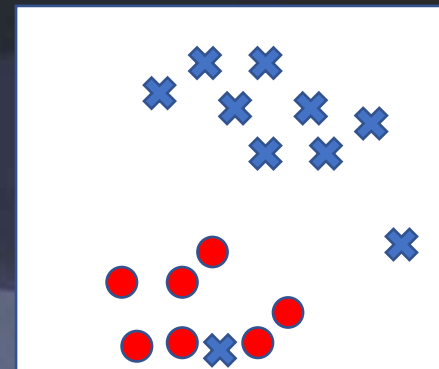
- **Training** is the process of making the system able to learn.
- A model that explains a certain situation well may fail in another situation.
 - The training set and test set come from the same distribution (in-distribution vs. out-of-distribution)
 - Before applying a model, check the assumptions!



Training data



Test data



Quick review

Supervised

Input: Data \mathbf{X} and label y

Goal: Learn how to map \mathbf{X} to y

Examples: Regression, classification

Unsupervised

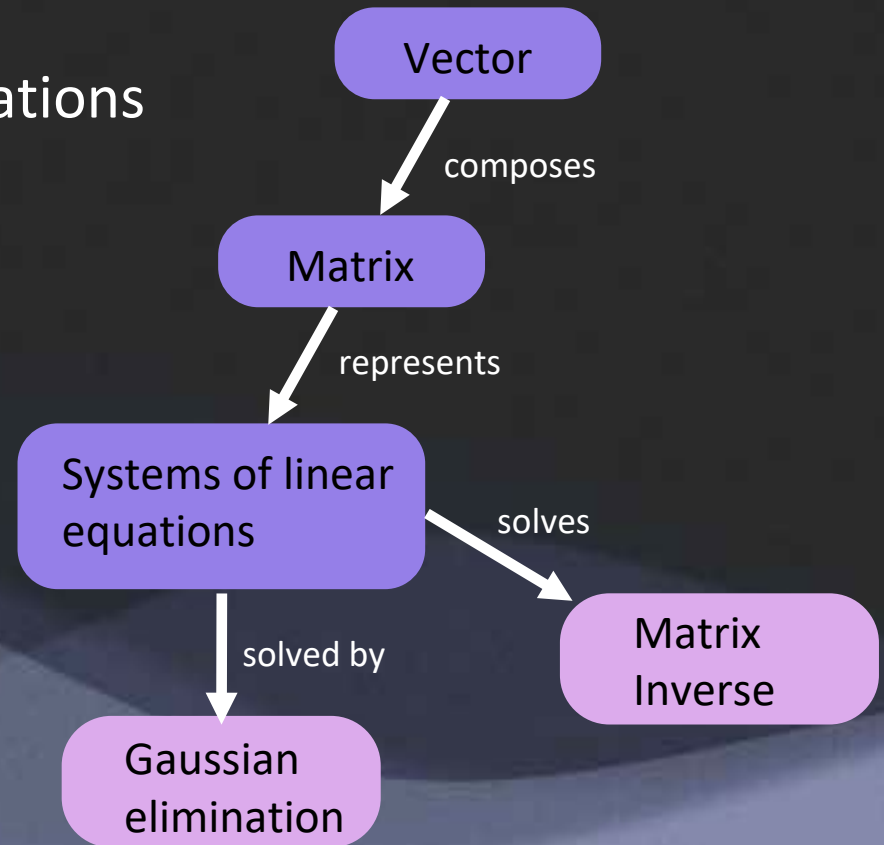
Input: Data \mathbf{X} , no label

Goal: Learn underlying structure of data

Examples: Clustering, dimensionality reduction

Outline

- Base concepts: vectors and systems of linear equations
- Matrices
- Solving systems of linear equations



2.1 Base concepts

Linear Algebra

- Algebra comes from the Arabic word 'al-jabr' (restoration/completion) introduced by the Persian astronomer and mathematician Muhammad ibn Musa al-Khwarizmi in the 9th century
- Linear Algebra: “the study of vectors and certain rules to manipulate vectors”
- Add vectors \rightarrow vector ; scale vector \rightarrow vector
- The “mathematics of data” since data is represented by vectors and matrices

Vectors as elements of \mathbb{R}^n

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

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

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

$$\mathbf{a} + \mathbf{b} = \mathbf{c} \in \mathbb{R}^n$$
$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \begin{bmatrix} -2 \\ 5 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 7 \\ 3 \end{bmatrix}$$

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

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

2.1 Systems of Linear Equations

You only have 5 AUD to buy pens, notebooks or snacks.

Because it's Canberra, we have Coles, Supabarn, Aldi...

Coles: one pen + 2 notebooks + no snack = 10 AUD

Supabarn: 2 pens + 1 notebook + 1 snack = 15 AUD

Aldi: no pen + 2 notebooks + 1 snack looking sadder than usual = 7 AUD

What can you buy?

2.1 Systems of Linear Equations

- Example 1: Does it have a solution?

$$x_1 + x_2 + x_3 = 0 \quad (1)$$

$$x_1 + x_2 + 2x_3 = 2 \quad (2)$$

$$+3x_3 = 6 \quad (3)$$

2.1 Systems of Linear Equations

- Example 1: Does it have a solution?

$$x_1 + x_2 + x_3 = 0 \quad (1)$$

$$x_1 + x_2 + 2x_3 = 2 \quad (2)$$

$$+3x_3 = 6 \quad (3)$$

Yes, it has infinitely many solutions.

$$\begin{array}{l} x_3 = 2 \\ x_1 + x_2 = -2 \end{array}$$

2.1 Systems of Linear Equations

- Example 2: Does it have a solution?

$$-x_1 + x_2 + 3x_3 = 3 \quad (1)$$

$$x_1 + x_2 + 2x_3 = 2 \quad (2)$$

$$2x_2 + 5x_3 = 1 \quad (3)$$

3 unknowns

x_1 x_2 x_3

2.1 Systems of Linear Equations

- Example 2: Does it have a solution?

$$-x_1 + x_2 + 3x_3 = 3 \quad (1)$$

$$x_1 + x_2 + 2x_3 = 2 \quad (2)$$

$$2x_2 + 5x_3 = 1 \quad (3)$$

3 unknowns

x_1 x_2 x_3

No.

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

2.2 Matrices

2.2 Matrices

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

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

i th row, j th 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**.

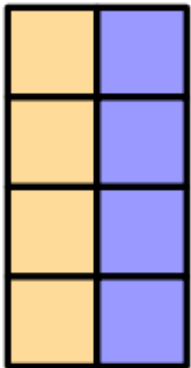
2.2 Matrices

#rows, #cols

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

Space

$$A \in \mathbb{R}^{4 \times 2}$$



re-shape

$$a \in \mathbb{R}^8$$



image

reshape



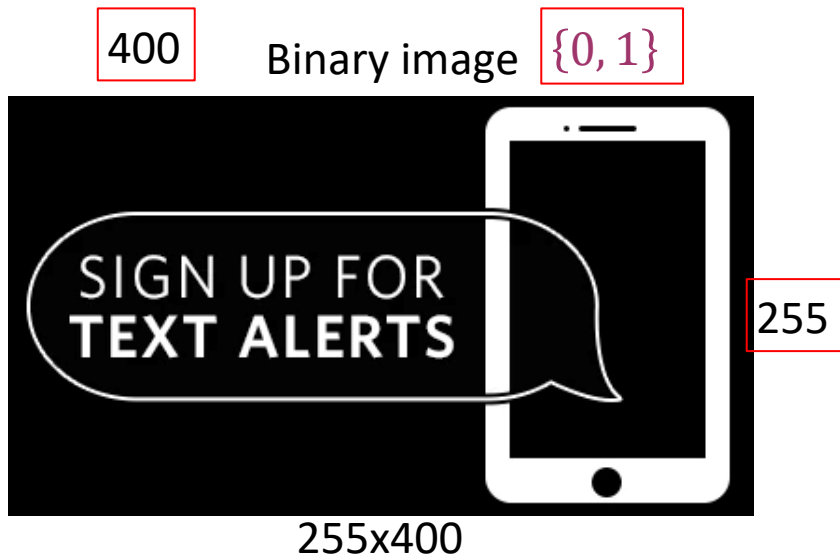
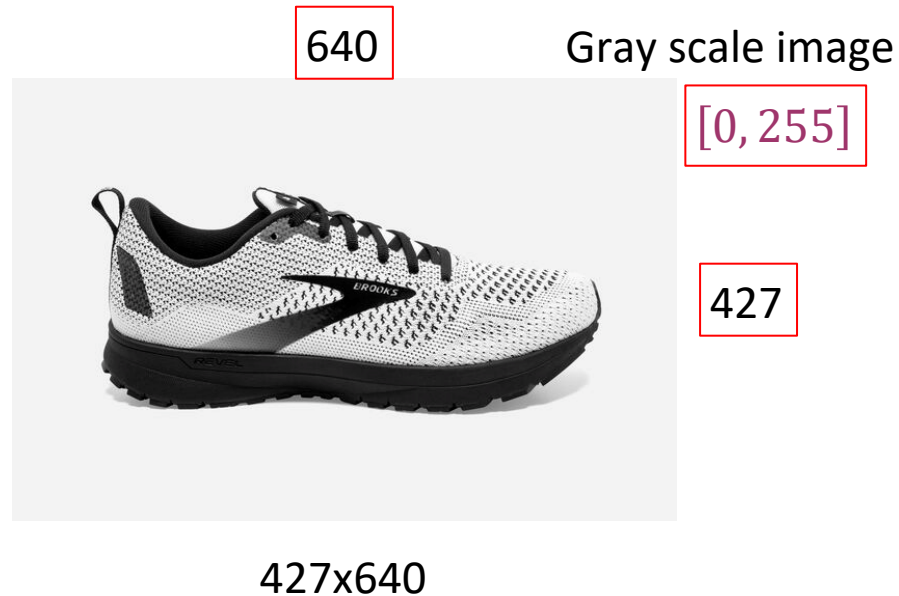
feature

classifier

Shoe

Perhaps the simplest feature representation

Matrix - example



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

3x6

6

2.2.1 Matrix Addition and Multiplication

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

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

- Example

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

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

2.2.1 Matrix Addition and Multiplication

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

$$c_{ij} = \sum_{l=1}^n a_{il}b_{lj}, \quad i = 1, \dots, m. \quad j = 1, \dots, k, \quad c_{ij} \neq a_{ij}b_{ij}$$

- To compute element c_{ij} we multiply the elements of the i th row of \mathbf{A} with the j th column of \mathbf{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:

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

- Identity Matrix

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

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

2.2.1 The Properties of Matrix Multiplication

- Associativity

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

- Distributivity

$$\begin{aligned}\forall A, B \in \mathbb{R}^{m \times n}, C, D \in \mathbb{R}^{n \times p}: (A + B)C &= AC + BC \\ A(C + D) &= AC + AD\end{aligned}$$

- Multiplication with the identity matrix:

$$\begin{aligned}\forall A \in \mathbb{R}^{m \times n}: I_m A &= A I_n = A \\ I_m &\neq I_n \text{ for } m \neq n.\end{aligned}$$

2.2.2 Inverse and Transpose

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

- Example

$$AB = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

$$B = A^{-1}$$

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

The matrices are inverse to each other, because

$$AB = I_2 = BA$$

2.2.2 Inverse and Transpose

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

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

- Important properties of inverses and transposes:

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

$$(A^T)^T = A$$

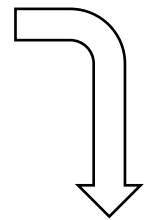
$$(A + B)^T = A^T + B^T$$

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

$$AA^{-1} = I = A^{-1}A$$

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

$$(AB)^T = B^T A^T$$



$$(AB) \cdot (B^{-1}A^{-1}) = AIA^{-1} = AA^{-1} = I$$

2.2.2 Inverse and Transpose

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

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

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

$$A + B \stackrel{?}{=} (A + B)^T \quad \begin{bmatrix} 1 & -2 \\ -2 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 3 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix}$$

- The product of two symmetric matrices 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 $\mathbf{A} \in \mathbb{R}^{m \times n}$. Then $\lambda \mathbf{A} = \mathbf{K}$, where $k_{ij} = \lambda a_{ij}$

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

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

2.2.3 Multiplication by a Scalar

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

- Associativity

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

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

- Transpose

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

- Distributivity

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

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

Matrix Algebra

$$1) A(B+C) = AB + AC$$

$$2) AB \neq BA \text{ in general}$$

$$3) (AB)C = A(BC)$$

$$4) AI = IA = A$$

$$5) AO = OA = O$$

6) AB may not be possible

7) $AB = O$ does not imply
 $A = O$ or $B = O$

Number Algebra

$$a(b+c) = ab + ac$$

$$ab = ba$$

$$a(bc) = (ab)c$$

$$1. a = a \cdot 1 = a$$

$$0 \cdot a = a \cdot 0 = 0$$

ab is always possible

$$ab = 0 \rightarrow a = 0 \text{ or } b = 0$$

2.2.4 Compact Representations of Systems of Linear Equations

- Consider the system of linear equations,

$$2x_1 + 3x_2 + 5x_3 = 1$$

$$4x_1 - 2x_2 - 7x_3 = 8$$

$$9x_1 + 5x_2 - 3x_3 = 2$$

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

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

2.3 Solving systems of linear equations

2.3 Solving Systems of Linear Equations

- Now we have a general form of an equation system

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

- 2.3.1 Particular and General Solution

Step 1. Find a **particular solution** to $A\mathbf{x} = \mathbf{b}$

Step 2. Find **all solutions** to $A\mathbf{x} = \mathbf{0}$

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

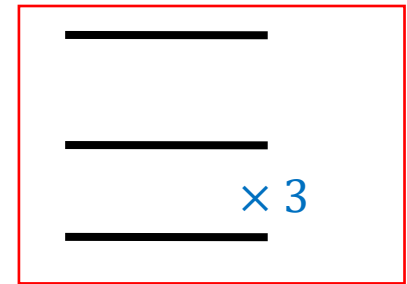
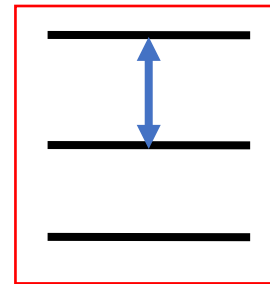
We use **Gaussian elimination** to solve the equation system.

2.3.2 Elementary Transformations

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

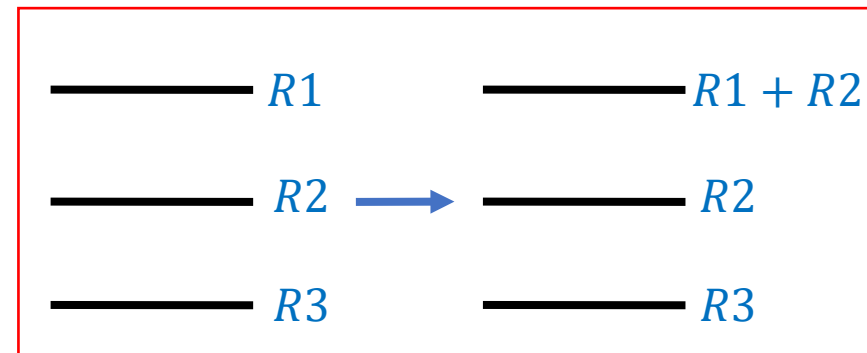
- Elementary transformations include:

- Exchange of two equations



- Multiplication of an equation (row) with a constant $\lambda \in \mathbb{R} \setminus \{0\}$

- Addition of two equations (rows)



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

$$\begin{pmatrix} * & * & \dots & \dots & \dots \\ 0 & * & \dots & \dots & \dots \\ 0 & 0 & * & \dots & \dots \\ 0 & 0 & 0 & * & \dots \\ 0 & 0 & 0 & 0 & \dots \end{pmatrix}$$

REF

$$\begin{pmatrix} 1 & 0 & 0 & 0 & * & \dots & \dots \\ 0 & 1 & 0 & 0 & * & \dots & \dots \\ 0 & 0 & 1 & 0 & * & \dots & \dots \\ 0 & 0 & 0 & 1 & * & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

RREF

- Row Echelon Form
- All rows with 0s must be at the bottom
- You don't have to have all 0 rows, but if they are they must be at the bottom
- Staircase pattern of the 1st non-zero entries of each row (pivots)
- A pivot is always strictly to the right of the pivot of the row above it

From Lorenzo A. Sadun's teaching video

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

Gaussian Elimination - example

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

augmented matrix

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



R1+R3 -> R3

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 9 \\ 0 & 1 & 3 & 3 \\ 0 & 1 & -3 & 11 \end{array} \right] \xrightarrow{\text{R3-R2} \rightarrow \text{R3}}$$

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 9 \\ 0 & 1 & 3 & 3 \\ 0 & 0 & -6 & 8 \end{array} \right]$$

Gaussian Elimination - example

- Seek all solutions to the following system of equations

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

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

$$\begin{array}{l} \text{R2}-2\text{R1} \rightarrow \text{R2} \\ \xrightarrow{\hspace{1cm}} \\ \text{R1}+\text{R3} \rightarrow \text{R3} \end{array} \begin{bmatrix} 1 & 2 & -1 & 3 & 2 \\ 0 & -1 & 0 & -1 & -3 \\ 0 & 0 & 0 & 2 & 6 \end{bmatrix} \text{ row-echelon form (REF)}$$

How to find the general solution to $Ax = b$

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

Gaussian elimination \rightsquigarrow

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

$x_1 \quad x_2 \quad x_3 \quad x_4 \quad \hat{b}$

Step 1. Find a particular solution to $Ax = b$

Step 2. Find all solutions to $Ax = 0$

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

Step 1: Finding a particular solution to $Ax = b$

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

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

$$\begin{matrix} x_1 & x_2 & x_3 & x_4 & \hat{b} \\ & & & & \mathbf{b} \end{matrix}$$

x_3 : free variable

$x_1 \ x_2 \ x_4$: basic variables

$$0 + 0 + 0 + 2x_4 = 6$$



$$x_4 = 3$$

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



$$x_2 = 0$$

$$\begin{aligned} x_1 + 2x_2 - x_3 + 3x_4 &= 2 \\ x_1 + 0 - 0 + 9 &= 2 \end{aligned}$$



$$x_1 = -7$$

A particular solution:

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

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

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

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

$x_1 \quad x_2 \quad x_3 \quad \boxed{x_4} \quad \mathbf{0}$

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

After setting $x_3 = 1$, we have

$$\begin{aligned} 0 - x_2 + 0 - x_4 &= 0, \\ x_1 + 2x_2 - 1 + 3x_4 &= 0 \end{aligned}$$

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

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

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

Another
example

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

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

$$\begin{array}{l} \text{R2}-4\text{R1} \rightarrow \text{R2} \\ \text{R3}+2\text{R1} \rightarrow \text{R3} \\ \text{R4}-\text{R1} \rightarrow \text{R4} \end{array} \xrightarrow{\quad} \begin{bmatrix} 1 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & -3 & 2 \\ 0 & 0 & 0 & -3 & 6 & -3 \\ 0 & 0 & -1 & -2 & 3 & a \end{bmatrix} \xrightarrow{\text{R4}-\text{R2} \rightarrow \text{R4}} \begin{bmatrix} 1 & -2 & 1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & -3 & 2 \\ 0 & 0 & 0 & -3 & 6 & -3 \\ 0 & 0 & 0 & -3 & 6 & a-2 \end{bmatrix}$$

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

a must equal -1 for this equation system to have solutions

Finding a particular solution to $A\mathbf{x} = \mathbf{b}$

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

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

$$x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5$$

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

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

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

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

A particular solution:

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

Find all solutions to $A\mathbf{x} = \mathbf{0}$

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

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

$x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5$

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

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

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

How to find the general solution to $Ax = b$

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

Step 1:

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

Step 2:

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

General solution:

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

Proof

- Given $A \in \mathbb{R}^{m \times n}$ with $m < n$, then $Ax = 0$ has infinitely many solutions
- Proof
- This system always has at least one solution since homogeneous
 - Consider $A0 = 0$ always holds
- Matrix A brought in row echelon form contains at most m pivots.

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

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

Proof

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

- Proof

proof by contradiction

- Let's assume the system $Ax = b$ has two solutions p and q .

- We have

$$Ap = b$$

$$Aq = b$$

- Consider

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.

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

$$Av = A(p + t(q - p)) = Ap + t(Aq - Ap) = b + t(b - b) = b$$

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

Calculating the Inverse with Gaussian Elimination

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

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

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

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

- First, write down the augmented matrix

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

A I_n

Calculating the Inverse with Gaussian Elimination

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

$$A = \left[\begin{array}{cccc|cccc} 1 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{array} \right] \xrightarrow{\text{REF}} \dots \xrightarrow{\text{REF}} A = \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -1 & 2 & -2 & 2 \\ 0 & 1 & 0 & 0 & 1 & -1 & 2 & -2 \\ 0 & 0 & 1 & 0 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 1 & -1 & 0 & -1 & 2 \end{array} \right] A^{-1}$$

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

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

Checking invertibility by calculating Reduced Row-echelon form - example

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

$$\xrightarrow{R3-R2 \rightarrow R3} \begin{bmatrix} 2 & -2 & 4 & -2 \\ 0 & 3 & 6 & 9 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{\substack{\text{Multiplication} \\ \text{by scalar}}} \begin{bmatrix} 1 & -1 & 2 & -1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

REF

$$\xrightarrow{\substack{R1+R3 \rightarrow R1 \\ R2-3R3 \rightarrow R2}} \begin{bmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{R1+R2 \rightarrow R1} \begin{bmatrix} 1 & 0 & 4 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \neq I_4$$

This matrix is not invertible

RREF

Moore-Penrose pseudo-inverse

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

$$Ax = b \Leftrightarrow A^T Ax = A^T b \Leftrightarrow x = (A^T A)^{-1} A^T b$$

- $(A^T A)^{-1} A^T$ is the Moore-Penrose pseudo-inverse of A

Check your understanding

- Which of the following are correct?
 - (A) A vector, when multiplied by a scale, is still a vector.
 - (B) For a system of linear equations with n variables, it is possible that none of them are free variables.
 - (C) For a system of linear equations with n variables, the maximum number of pivots in the REF is $n - 1$.
 - (D) A matrix, when added by an identity matrix, stays as is.
 - (E) We can use matrix transpose in Gaussian Elimination.
 - (F) Two arbitrary matrices can be multiplied.
 - (G) Two arbitrary matrices can be added.
 - (H) A greyscale image with black borders is not a matrix.

Check your understanding

- Let A , B , C be 2×2 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
 - $(AB)I_2$
 - $(BA)I_2$

Next lecture: Textbook Sec. 2.4 - 2.7

