CSC4007 Advanced Machine Learning

Lesson 02: Review on Linear Algebra

by Vien Ngo
EEECS / ECIT / DSSC

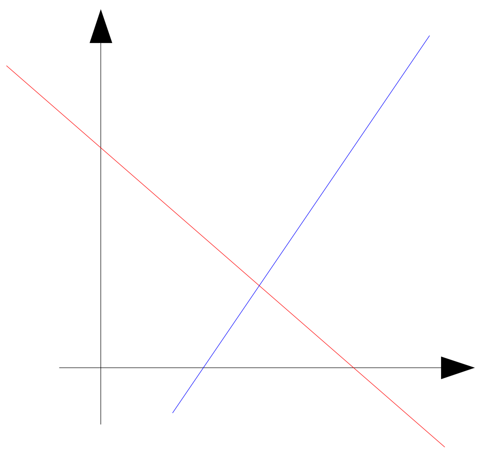
What is linear algebra?

- Linear algebra is the branch of mathematics concerning linear equations such as

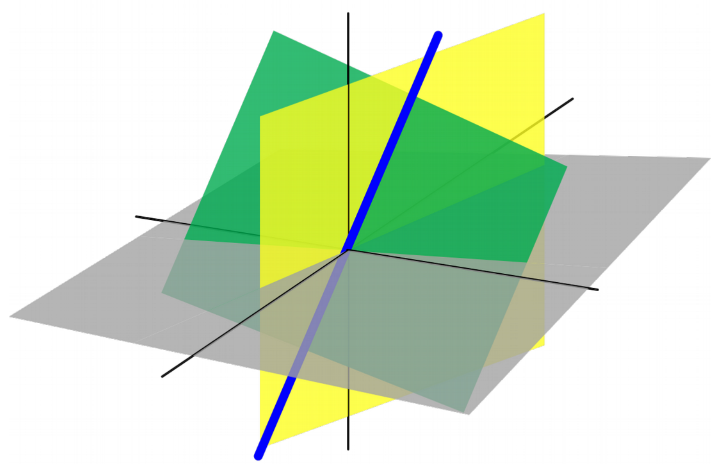
$$a_1 x_1 + a_2 x_2 + \ldots + a_n x_n = b$$

- Linear algebra is fundamental to geometry, for defining objects such as lines, planes, rotations

$$a_1 x_1 + a_2 x_2 = b$$



$$a_1 x_1 + a_2 x_2 + a_3 x_3 = b$$



Why do we need to know it?

- Linear Algebra is used throughout engineering
- Essential for understanding ML algorithms
  - Describe systems: the input space X and output space Y in ML.
  - The space of functions $f : X \mapsto Y$

A combination of linear transformations, i.e. translation, rotation

- Here we discuss:

– Concepts of linear algebra needed for ML

(Some materials are from Sargur N. Srihari srihari@cedar.buffalo.edu)

Outline

- Scalars, Vectors
- Vector Operations
- Matrices
- Multiplying Matrices and Vectors
- Identity and Inverse Matrices
- The Determinant

Scalars

- Single number
  - In contrast to other objects in linear algebra, which are usually arrays of numbers
- Represented in lower-case italic *x*
  - They can be real-valued or be integers

    E.g., let $x \in \mathbb{R}$ be the slope of the line

    - Defining a real-valued scalar

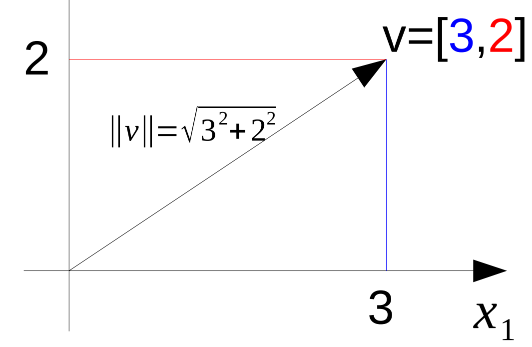
    E.g., let $n \in \mathbb{N}$ be the number of units

    - Defining a natural number scalar

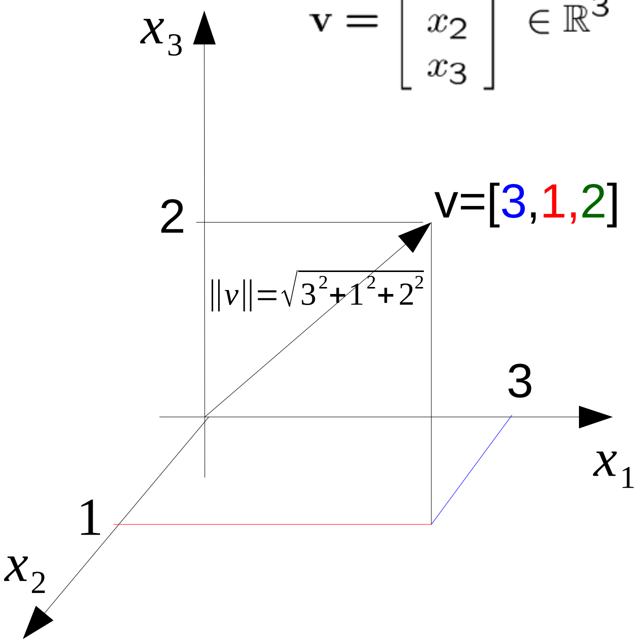
Example: 2D and 3D Vectors

2D,3D Vectors

$$\mathbf{v} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$$



$$\mathbf{v} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$$



Magnitude:

$$\|v\| = \sqrt{x_1^2 + x_2^2}$$

$$\|v\| = \sqrt{x_1^2 + x_2^2 + x_3^2}$$

If $\|\mathbf{v}\| = 1$, $\mathbf{v}$ is a UNIT vector

Vector

- An array of numbers arranged in order
- Each no. identified by an index
- Written as x and , defined as a column vector
  - its elements are in italics lower case, subscripted
  - By convention (in this module): x means a column vector

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

row vector $$x=[x_1, x_2, \cdots, x_n]$$

- If each element is in R then x is in $\mathbb{R}^n$
- We can think of vectors as points in space
  - Each element gives coordinate along an axis

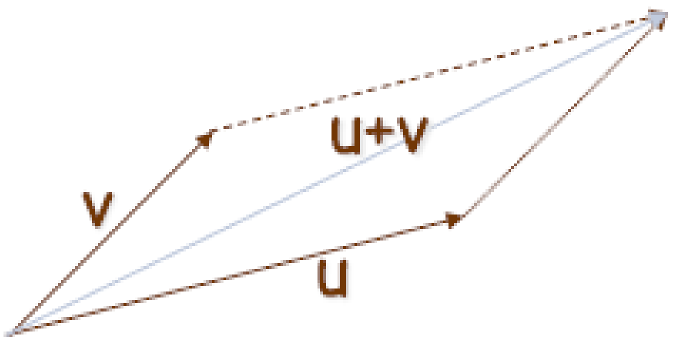
Outline

- Scalars, Vectors
- **Vector Operations**
- Matrices
- Multiplying Matrices and Vectors
- Identity and Inverse Matrices
- The Determinant

Vector Addition

- Addition in 2D

$$u + v = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \end{bmatrix}$$



e.g.: $u+v=\begin{bmatrix} 2 \\ 5 \end{bmatrix}+\begin{bmatrix} 4 \\ 6 \end{bmatrix}=\begin{bmatrix} 2+4 \\ 5+6 \end{bmatrix}=\begin{bmatrix} 6 \\ 11 \end{bmatrix}$

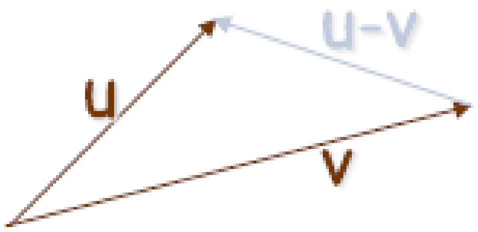
- Addition in *n*-dim

$$u+v=\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}+\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}=\begin{bmatrix} u_1+v_1 \\ u_2+v_2 \\ \vdots \\ u_n+v_n \end{bmatrix}$$

Vector Subtraction

- Subtraction in 2D

$$u - v = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} - \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} u_1 - v_1 \\ u_2 - v_2 \end{bmatrix}$$



e.g.: $u-v=\begin{bmatrix} 2 \\ 5 \end{bmatrix}-\begin{bmatrix} 4 \\ 6 \end{bmatrix}=\begin{bmatrix} 2-4 \\ 5-6 \end{bmatrix}=\begin{bmatrix} -2 \\ -1 \end{bmatrix}$

- Subtraction in *n*-dim

$$u-v=\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}-\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}=\begin{bmatrix} u_1-v_1 \\ u_2-v_2 \\ \vdots \\ u_n-v_n \end{bmatrix}$$

Scalar product

$a \in \mathbb{R}$ is a scalar

- Product in 2D

$$av = a\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} av_1 \\ av_2 \end{bmatrix}$$

e.g.: $4 \times v = 4 \times \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 8 \\ 20 \end{bmatrix}$

- Product in *n*-dim

$$a\,v = \begin{bmatrix} a\,v_1 \\ a\,v_2 \\ \vdots \\ a\,v_n \end{bmatrix}$$

Vector Transpose

- Transpose:
  - Transform a column vector to a row vector
  - Transform a row vector to a column vector

If x is a column vector

$$x^T = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}^T = [x_1, x_2, \cdots, x_n]$$

If x is a row vector

$$x^T = [x_1, x_2, \cdots, x_n]^T = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

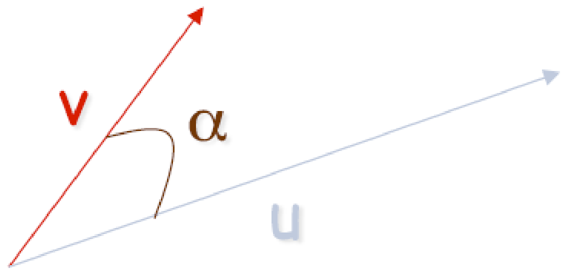
e.g.: $u^T = \begin{bmatrix} 2 \\ 5 \end{bmatrix}^T = [2,5]$

e.g.: $u^T = [3,4]^T = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$

Inner (dot) Product

- Inner product in 2D and 3D

$$u^T v = [u_1, u_2] \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = u_1 v_1 + u_2 v_2$$



The inner product is a SCALAR!

$$u^T v = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \|u\|\|v\| \cos \alpha$$

$$u^T v = 0 \leftrightarrow u \perp v$$

e.g.: $u^T v = \begin{bmatrix} 4 \\ 2 \end{bmatrix}^T \begin{bmatrix} 3 \\ 5 \end{bmatrix} = 4*3+2*5=22$

$$[4,2]\begin{bmatrix} 3 \\ 5 \end{bmatrix}$$

$$u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \qquad v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

$$\langle u, v \rangle \doteq u^T v = u_1 v_1 + u_2 v_2 + u_3 v_3$$

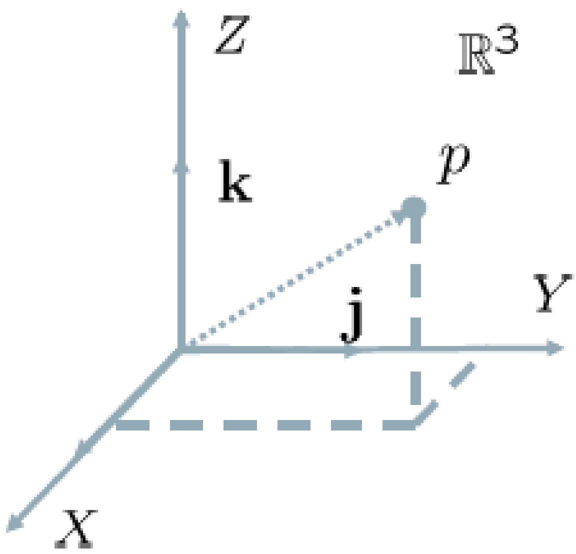
$$\|u\| = \sqrt{u^T u} = \sqrt{u_1^2 + u_2^2 + u_3^2}$$

Orthonormal Basis

- Orthonormal basis in 3D

Standard base vectors:

$$\mathbf{i} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \mathbf{j} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \mathbf{k} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$



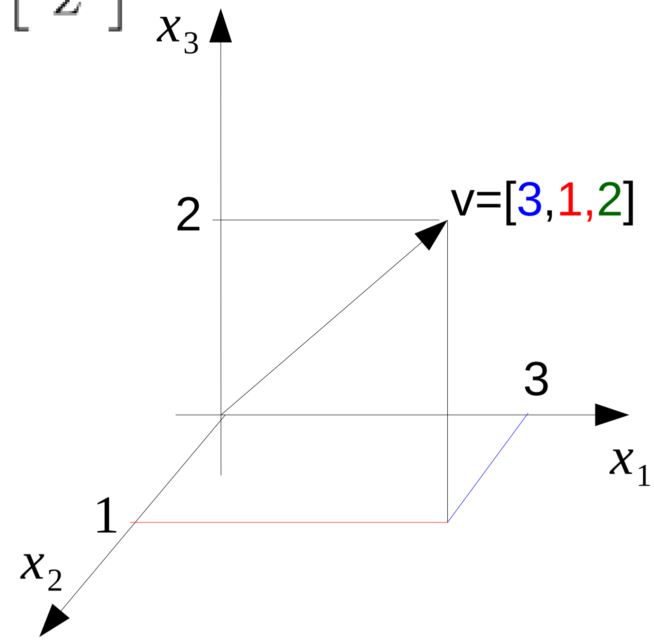
Coordinates of a point p in space:

$$\boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \in \mathbb{R}^3 \qquad \boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = X.\mathbf{i} + Y.\mathbf{j} + Z.\mathbf{k}$$

Orthonormal Basis

- Orthonormal basis in 3D

$$\boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \in \mathbb{R}^3 \qquad \boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = X.\mathbf{i} + Y.\mathbf{j} + Z.\mathbf{k}$$



$$v = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} = 3\overset{\mathrm{i}}{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}} + 1\overset{\mathrm{j}}{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}} + 2\overset{\mathrm{k}}{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}} = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$$

Orthonormal Basis

- Orthonormal basis in *n*-dim

$$i_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad i_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \qquad \cdots \qquad i_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

coordinates of a point *u* in *n*-dim space:

$$u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = u_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + u_2 \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \ldots + u_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

Outline

- Scalars, Vectors
- Vector Operations
- **Matrices**
- Multiplying Matrices and Vectors
- Identity and Inverse Matrices
- The Determinant

Matrices

- 2-D array of numbers
  - So each element identified by two indices

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

e.g. $A = \begin{bmatrix} 2 & 5 \\ 3 & 8 \end{bmatrix}$

- $n \times m$ matrix

$$A_{n\times m} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ & \cdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

n rows

m columns

e.g. $A = \begin{bmatrix} 2 & 5 & 4 \\ 1 & 3 & 8 \\ 11 & 23 & 18 \end{bmatrix}$

Matrix Addition and Subtraction

$$A_{n\times m} = \begin{bmatrix} a_{11} & a12 & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ & \ldots & & \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{bmatrix}$$

Sum:

$$C_{n\times m} = A_{n\times m} + B_{n\times m}$$

$$c_{ij} = a_{ij} + b_{ij}$$

A and B must have the same dimensions

Example:

Addition

$$\begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix} + \begin{bmatrix} 6 & 2 \\ 1 & 5 \end{bmatrix} = \begin{bmatrix} 8 & 7 \\ 4 & 6 \end{bmatrix}$$

Subtraction

$$\begin{bmatrix} 5 & 3 \\ 0 & 4 \end{bmatrix} - \begin{bmatrix} 3 & 2 \\ 1 & 10 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ -1 & -6 \end{bmatrix}$$

Matrix Product

Product:

$$C_{n\times p} = A_{n\times m} B_{m\times p}$$

A and B must have compatible dimensions

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

$$A_{n\times n} B_{n\times n} \neq B_{n\times n} A_{n\times n}$$

Examples:

$$\begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 6 & 2 \\ 1 & 5 \end{bmatrix} = \begin{bmatrix} 17 & 29 \\ 19 & 11 \end{bmatrix}$$

Matrix Product

Product:

$$C_{n\times p} = A_{n\times m} B_{m\times p}$$

A and B must have compatible dimensions

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

$$A_{n\times n} B_{n\times n} \neq B_{n\times n} A_{n\times n}$$

Examples:

$$\begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 6 & 2 \\ 1 & 5 \end{bmatrix} = \begin{array}{c} C_{11} \\ \begin{bmatrix} 17 & 29 \\ 19 & 11 \end{bmatrix} \end{array}$$

2*6+ 5*1 = 17 = C_{11}

Matrix Product

Product:

$$C_{n\times p} = A_{n\times m} B_{m\times p}$$

A and B must have compatible dimensions

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

$$A_{n\times n} B_{n\times n} \neq B_{n\times n} A_{n\times n}$$

Examples:

$$\begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 6 & 2 \\ 1 & 5 \end{bmatrix} = \begin{array}{cc} & C_{12} \\ \end{array} \begin{bmatrix} 17 & 29 \\ 19 & 11 \end{bmatrix}$$

2*2+ 5*5 = 29 = C_{12}

Matrix Product

Product:

$$C_{n\times p} = A_{n\times m} B_{m\times p}$$

A and B must have compatible dimensions

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

$$A_{n\times n} B_{n\times n} \neq B_{n\times n} A_{n\times n}$$

Examples:

$$\begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 6 & 2 \\ 1 & 5 \end{bmatrix} = \begin{bmatrix} 17 & 29 \\ 19 & 11 \end{bmatrix}$$

C_{21}

3*6+ 1*1 = 19 = C_{21}

Matrix Product

Product:

$$C_{n\times p} = A_{n\times m} B_{m\times p}$$

A and B must have compatible dimensions

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

$$A_{n\times n} B_{n\times n} \neq B_{n\times n} A_{n\times n}$$

Examples:

$$\begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 6 & 2 \\ 1 & 5 \end{bmatrix} = \begin{bmatrix} 17 & 29 \\ 19 & 11 \end{bmatrix}$$

C_{22}

3*2+ 1*5 = 11 = C_{22}

Matrix Product

- Working example

$$\begin{bmatrix} 6 & 2 \\ 1 & 5 \end{bmatrix} \cdot \begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} ?? & ?? \\ ?? & ?? \end{bmatrix}$$

Matrix Product

- Working example

$$\begin{bmatrix} 6 & 2 \\ 1 & 5 \end{bmatrix} \cdot \begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} ?? & ?? \\ ?? & ?? \end{bmatrix}$$

Solution: $\begin{bmatrix} 6 & 2 \\ 1 & 5 \end{bmatrix} \cdot \begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix}$ **=** $\begin{bmatrix} 18 & 32 \\ 17 & 10 \end{bmatrix}$

Matrix Transpose

- The transpose of a matrix A is denoted as A^T
  - Defined as the mirror image across a diagonal line

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \\ A_{3,1} & A_{3,2} & A_{3,3} \end{bmatrix} \Rightarrow A^T = \begin{bmatrix} A_{1,1} & A_{2,1} & A_{3,1} \\ A_{1,2} & A_{2,2} & A_{3,2} \\ A_{1,3} & A_{2,3} & A_{3,3} \end{bmatrix}$$

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \\ A_{3,1} & A_{3,2} \end{bmatrix} \Rightarrow A^T = \begin{bmatrix} A_{1,1} & A_{2,1} & A_{3,1} \\ A_{1,2} & A_{2,2} & A_{3,2} \end{bmatrix}$$

Transpose:

$$C_{m\times n} = A^T{}_{n\times m} \qquad (A+B)^T = A^T + B^T$$

$$c_{ij} = a_{ji} \qquad (AB)^T = B^T A^T$$

If $A^T = A$ A is symmetric

e.g. $\rightarrow \begin{bmatrix} 2 & 6 & 5 \\ 6 & 1 & 4 \\ 5 & 4 & 0 \end{bmatrix}$

Examples:

$$\begin{bmatrix} 6 & 2 \\ 1 & 5 \end{bmatrix}^T = \begin{bmatrix} 6 & 1 \\ 2 & 5 \end{bmatrix} \qquad \begin{bmatrix} 6 & 2 \\ 1 & 5 \\ 3 & 8 \end{bmatrix}^T = \begin{bmatrix} 6 & 1 & 3 \\ 2 & 5 & 8 \end{bmatrix}$$

Matrix Determinant

- Determinant of a **square** matrix A is a mapping to **a scalar**
- Denoted as **det(A)** or **|A|**

$$\det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

$$\det\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

Example:

$$\det\begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix} = 2 - 15 = -13$$

Identity Matrix

- Identity matrix does not change value of vector when we multiply the vector by identity matrix

$$AI = A$$

  - Denote identity matrix that preserves n-dimensional vectors as I_n
  - Formally $I_n \in \mathbb{R}^{n \times n}$ and $\forall \mathbf{x} \in \mathbb{R}^n, I_n \mathbf{x} = \mathbf{x}$
  - Example of I_3 $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Matrix Inversion

- Matrix inversion is a powerful tool to analytically solve **Ax=b**
- A must be square. Matrix inversion is denoted as A^{-1}

Inverse:

$$A_{n\times n}A^{-1}{}_{n\times n} = A^{-1}{}_{n\times n}A_{n\times n} = I$$

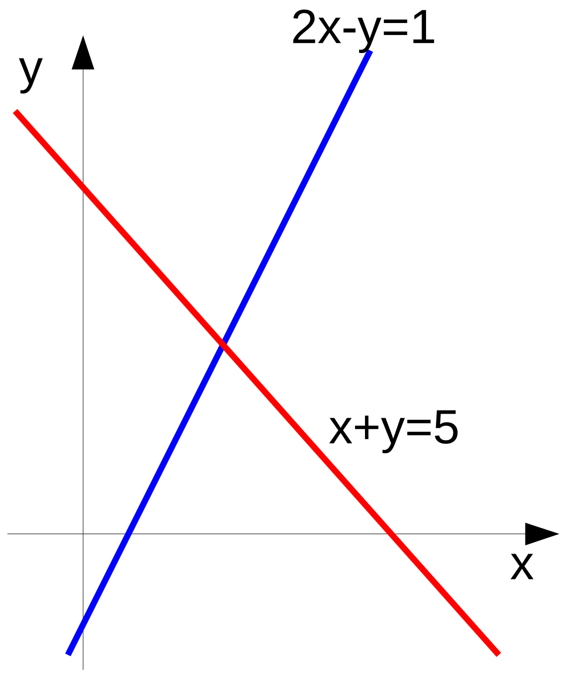
→ e.g. $5\times\frac{1}{5}=1$ so $5^{-1}=\frac{1}{5}$

e.g. 2D matrix $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} = \frac{1}{a_{11}a_{22} - a_{21}a_{12}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$

Example: $\begin{bmatrix} 6 & 2 \\ 1 & 5 \end{bmatrix}^{-1} = \frac{1}{28} \begin{bmatrix} 5 & -2 \\ -1 & 6 \end{bmatrix}$

$$\begin{bmatrix} 6 & 2 \\ 1 & 5 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 6 & 2 \\ 1 & 5 \end{bmatrix} = \frac{1}{28} \begin{bmatrix} 5 & -2 \\ -1 & 6 \end{bmatrix} \cdot \begin{bmatrix} 6 & 2 \\ 1 & 5 \end{bmatrix} = \frac{1}{28} \begin{bmatrix} 28 & 0 \\ 0 & 28 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Geometric Interpretation



Lines in 2D space - row solution
Equations are considered isolation

$$\begin{aligned} 2x - y &= 1 \\ x + y &= 5 \end{aligned}$$

Linear combination of vectors in 2D
Column solution

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix} x + \begin{bmatrix} -1 \\ 1 \end{bmatrix} y = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

We already know how to multiply the vector by scalar

Geometric Interpretation

In 3D

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

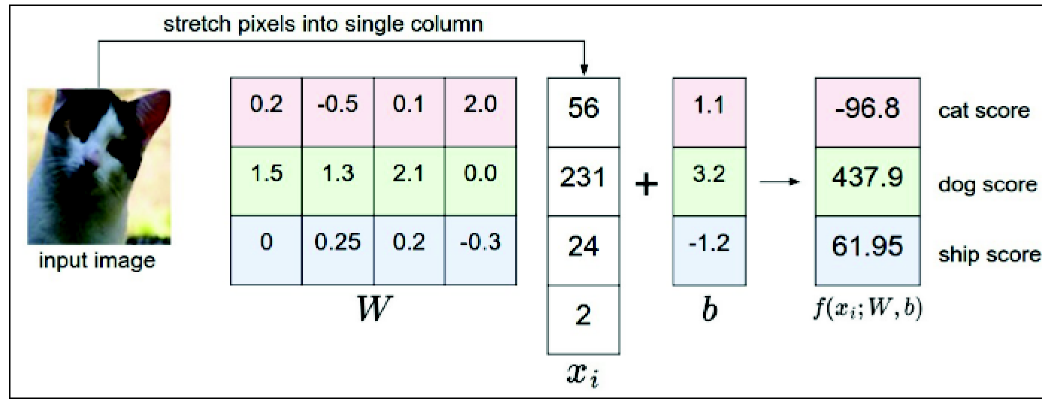
When is RHS a linear combination of LHS

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

Example of flow of matrix operation in ML

Vector $\boldsymbol{x}$ is converted into vector $\boldsymbol{y}$ by multiplying $\boldsymbol{x}$ by a matrix $\boldsymbol{W}$

A linear classifier $\boldsymbol{y} = \boldsymbol{W}\boldsymbol{x}^{\mathrm{T}} + \boldsymbol{b}$



A linear classifier with bias eliminated $\boldsymbol{y} = \boldsymbol{W}\boldsymbol{x}^{\mathrm{T}}$