Linear Algebra Refresher Part 1

An introduction of Vector and Matrix

Instructors

Mousum Dutta
Chief Data Scientist, Spotle.ai
IIT Kharagpur

Dr Soumen Nandi
Research Scientist
Indian Statistical Institute

Topics To Be Covered:

- **A** Scalars, Vectors and Matrices
- **B** Operations on Vectors and Matrices
- **C** Special Matrices
- **D** Determinants and Inverse of Matrix
- **E** Solving Simultaneous Equations

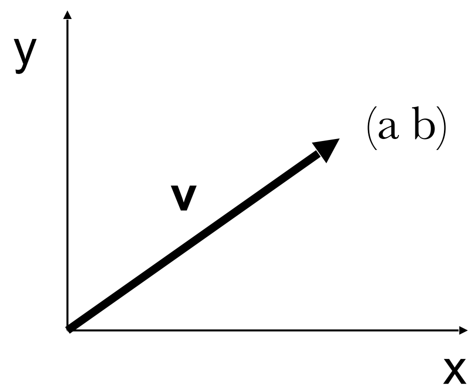
A Scalars

Any real number, or any quantity that can be measured using a single real number. Temperature, length, and mass are all scalars. A scalar has magnitude but no direction.

A What is a Vector?

- Column of numbers e.g. height, weight and age of a person.
- Think of a vector as a directed line segment in N-dimensions! (has “length” and “direction”).

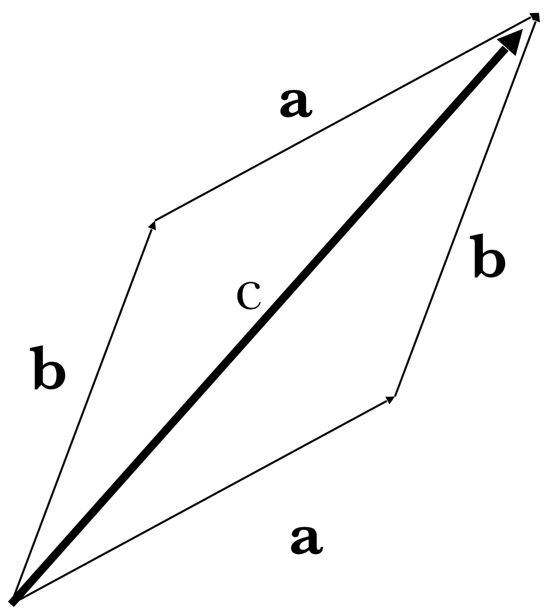
$$\boldsymbol{x} = \begin{pmatrix} height \\ weight \\ age \end{pmatrix}$$



$$\boldsymbol{x}_n = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_n \end{pmatrix}$$

B Vector Addition

$$\mathbf{x} + \mathbf{y} = (x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$$

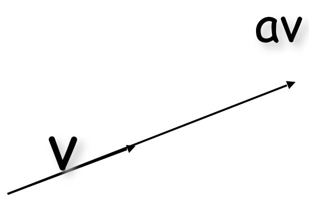


$$\mathbf{a} + \mathbf{b} = \mathbf{c}$$

The Statement of Parallelogram law of vector addition is, If two vectors are considered to be the adjacent sides of a parallelogram, then the resultant of two vectors is given by the vector that is a diagonal passing through the point of contact of two vectors.

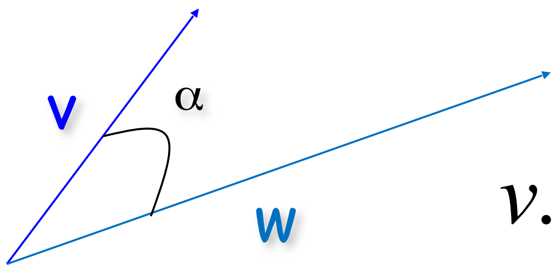
B Scalar Product

$$a\mathbf{v} = a(x_1, x_2) = (ax_1, ax_2)$$



- Multiplying a vector by scalar changes its magnitude but keeps the direction fixed.
- This operation is also known as scaling.

Inner (dot) Product



$$v.w = (x_1, x_2).(y_1, y_2) = x_1 y_1 + x_2 y_2$$

$$v.w = (x_1, x_2).(y_1, y_2) = \| v \| \cdot \| w \| \cos\alpha$$

$$v.w = 0 \Leftrightarrow v \perp w$$

- When two vectors are orthogonal (or perpendicular to each other) their inner product is zero.
- The L_2 norm of a vector **v** is the square root of self dot product. That is square root of (**v.v**) = square root of $(x_1^2 + x_2^2)$.

A What is a Matrix?

- A matrix is a set of elements, organized into rows and columns.
- A matrix is an array of vectors.

rows

columns

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$\begin{array}{ccc} person1 & person2 & person3 \\ \left(\begin{matrix} height \\ weight \\ age \end{matrix}\right. & \begin{matrix} height \\ weight \\ age \end{matrix} & \left.\begin{matrix} height \\ weight \\ age \end{matrix}\right) \end{array}$$

A matrix of order rxc means there are r rows and c columns.
Vector is a n x 1 matrix.

Matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 4 & 1 \\ 6 & 7 & 4 \end{bmatrix} \qquad \mathbf{C} = \begin{bmatrix} 1 & 4 \\ 2 & 7 \\ 3 & 8 \end{bmatrix} \qquad \mathbf{D} = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix}$$

Square (3 x 3)

Rectangular (3 x 2)

$d_{i\,j}$: i^{th} row, j^{th} column

B Matrix Operations

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix}$$

Just add elements

Commutative: **A+B=B+A**
Associative: (**A**+**B**)+**C**=**A**+(**B**+**C**)

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} - \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a-e & b-f \\ c-g & d-h \end{bmatrix}$$

Just subtract elements

B Matrix Operations

- Scalar multiplication

$$\alpha A = \alpha \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} \alpha a & \alpha b & \alpha c \\ \alpha d & \alpha e & \alpha f \\ \alpha g & \alpha h & \alpha i \end{pmatrix}$$

- Transpose

$$A^T = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}^T = \begin{pmatrix} a & d & g \\ b & e & h \\ c & f & i \end{pmatrix}$$

$(i,j)^{th}$ element of a matrix becomes $(j,i)^{th}$ element of it's transpose matrix.

A matrix is said to be symmetric matrix if it is equal to its own transpose.

B Matrix Operations

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{bmatrix}$$

Multiply each row by each column

$$\begin{bmatrix} l_{11} & l_{12} & l_{13} \\ l_{21} & l_{22} & l_{23} \\ l_{31} & l_{32} & l_{33} \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \cdot \begin{bmatrix} n_{11} & n_{12} & n_{13} \\ n_{21} & n_{22} & n_{23} \\ n_{31} & n_{32} & n_{33} \end{bmatrix}$$

$$l_{12} = m_{11}n_{12} + m_{12}n_{22} + m_{13}n_{32}$$

$$C^{m \times p} = A^{m \times n} \times \ B^{n \times p}$$

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} \times b_{k,j}$$

B Matrix multiplication

- Matrix multiplication is NOT commutative.
  $AB \neq BA$
- Matrix multiplication is associative.
  $A(BC)=(AB)C$
- Matrix multiplication is distributive.
  $A(B+C)=AB+AC$
  $(A+B)C=AC+BC$

B Vector Products

Two vectors:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

Inner product = scalar

Inner product X^TY is a scalar
(1xn) (nx1) = (1x1)

$$\mathbf{x}^T\mathbf{y} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = x_1y_1 + x_2y_2 + x_3y_3 = \sum_{i=1}^{3} x_iy_i$$

Outer product = matrix

$$\mathrm{xy}^\mathrm{T} = \begin{bmatrix} \mathrm{x}_1 \\ \mathrm{x}_2 \\ \mathrm{x}_3 \end{bmatrix} \begin{bmatrix} \mathrm{y}_1 & \mathrm{y}_2 & \mathrm{y}_3 \end{bmatrix} = \begin{bmatrix} \mathrm{x}_1\mathrm{y}_1 & \mathrm{x}_1\mathrm{y}_2 & \mathrm{x}_1\mathrm{y}_3 \\ \mathrm{x}_2\mathrm{y}_1 & \mathrm{x}_2\mathrm{y}_2 & \mathrm{x}_2\mathrm{y}_3 \\ \mathrm{x}_3\mathrm{y}_1 & \mathrm{x}_3\mathrm{y}_2 & \mathrm{x}_3\mathrm{y}_3 \end{bmatrix}$$

Outer product XY^T is a matrix
(nx1) (1xn) = (nxn)

C Identity matrix

Consider the 3x3 matrix:

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

For any nxn matrix $\boldsymbol{A}$, we have $\boldsymbol{A}\,\boldsymbol{I}_n = \boldsymbol{I}_n\,\boldsymbol{A} = \boldsymbol{A}$

For any nxm matrix $\boldsymbol{A}$, we have $\boldsymbol{I}_n\,\boldsymbol{A} = \boldsymbol{A}$, and $\boldsymbol{A}\,\boldsymbol{I}_m = \boldsymbol{A}$

Worked example
$A\,I_3 = A$
for a 3x3 matrix:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \text{X} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1+0+0 & 0+2+0 & 0+0+3 \\ 4+0+0 & 0+5+0 & 0+0+6 \\ 7+0+0 & 0+8+0 & 0+0+9 \end{bmatrix}$$

C Special types of matrix

- Null matrix

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

- Diagonal matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

- Lower triangular matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 8 & 7 & 4 \end{pmatrix}$$

- Upper triangular matrix

$$\begin{pmatrix} 1 & 2 & 8 \\ 0 & 3 & 7 \\ 0 & 0 & 4 \end{pmatrix}$$

- Orthogonal matrix
    - An orthogonal matrix is a square matrix whose rows and columns are orthogonal and they have unit length.

$$\begin{pmatrix} 0 & -0.8 & -0.6 \\ 0.8 & -0.36 & 0.48 \\ 0.6 & -0.48 & -0.64 \end{pmatrix}$$

D Matrix inverse

- A square matrix *A* is called nonsingular or invertible if there exists a matrix *B* such that:

$$AB = BA = I_n$$

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

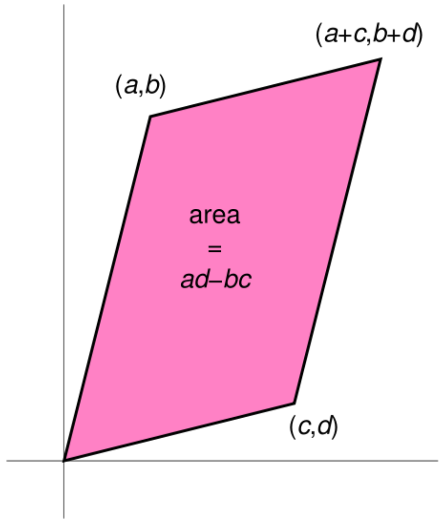
- A common notation for the inverse of a matrix *A* is A^{-1}. So:

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

- The inverse matrix is unique when it exists.
- So if *A* is invertible, then A^T is also invertible and then $(A^T)^{-1} = (A^{-1})^T$

D Determinant

- Given a square matrix A, its determinant is a real number associated with the matrix.
- The determinant of A is written as det(A) or |A|.



For a 2x2 matrix, the definition is

$$\det(A) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

The area of the parallelogram is the absolute value of the determinant of the matrix formed by the vectors representing the parallelogram's sides.

D Determinant: 2x2 examples

$$\det \begin{pmatrix} 1 & 1 \\ 3 & 4 \end{pmatrix} = \begin{vmatrix} 1 & 1 \\ 3 & 4 \end{vmatrix} = (1)(4) - (1)(3) = 1$$

$$\det \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = \begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = (1)(4) - (2)(2) = 0$$

D Determinant

- To define det(A) for larger matrices, we will need the definition of a minor M_{ij}.
- The minor M_{ij} of a matrix A is the matrix formed by removing the i^{th} row and the j^{th} column of A.

M_{11} : remove row 1, col 1

$$\underline{A} = \begin{pmatrix} 1 & 1 & -2 \\ -1 & 2 & 3 \\ 2 & 7 & 0 \end{pmatrix} \qquad M_{11} = \begin{pmatrix} 2 & 3 \\ 7 & 0 \end{pmatrix}$$

M_{12} : remove row 1, col 2

$$\underline{\underline{A}} = \begin{pmatrix} 1 & 1 & -2 \\ -1 & 2 & 3 \\ 2 & 7 & 0 \end{pmatrix} \qquad M_{12} = \begin{pmatrix} -1 & 3 \\ 2 & 0 \end{pmatrix}$$

D Determinant

For a matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

- Its determinant is given by

$$|A| = a_{11}|M_{11}| - a_{12}|M_{12}| + a_{13}|M_{13}|$$

- From the formula for a 2x2 matrix:

$$|M_{11}| = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} = a_{22}a_{33} - a_{23}a_{32}$$

D Determinant

$$\underline{A} = \begin{pmatrix} 1 & 1 & -2 \\ -1 & 2 & 3 \\ 2 & 7 & 0 \end{pmatrix}$$

$$|A| = 1 \times |M_{11}| - 1 \times |M_{12}| + (-2) \times |M_{13}|$$

$$|A| = 1 \times \begin{vmatrix} 2 & 3 \\ 7 & 0 \end{vmatrix} - 1 \times \begin{vmatrix} -1 & 3 \\ 2 & 0 \end{vmatrix} + (-2) \times \begin{vmatrix} -1 & 2 \\ 2 & 7 \end{vmatrix}$$

$$= 1 \times (-21) - 1 \times (-6) + (-2) \times (-11) = 7$$

D A general formula for determinants

- For a nxn matrix $A=(a_{ij})$ the **co-factors** of A are defined by

$$C_{ij} = (-1)^{i+j}|M_{ij}|$$

- The determinant of A is given by the formula

$$|A| = \sum_{i=1}^{n} a_{ij} C_{ij} \quad \text{for any } j=1,2,\ldots,n$$

D Matrix Inverse

$$A^{-1} = \frac{1}{det(A)} \begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \cdots & \cdots & \cdots & \cdot\cdot \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{pmatrix}^T$$

E Solving simultaneous equations

- For one linear equation *ax=b* where the unknown is *x* and *a* and *b* are constants,
- 3 possibilities:

If $\boxed{a \neq 0}$ then $x = \dfrac{b}{a} \equiv a^{-1}b$ thus there is single solution

If $\boxed{a = 0\,,\ b = 0}$ then the equation $ax = b$ becomes $0 = 0$ and any value of x will do

If $\boxed{a = 0\,,\ b \neq 0}$ then $ax = b$ becomes $0 = b$ which is a contradiction

E Solving simultaneous equations

Let's use solution $x = a^{-1}b$ from the single equation to solve

For example

$$2x_1 + 3x_2 = 5$$

$$x_1 - 2x_2 = -1$$

In matrix notation,

$$\begin{pmatrix} 2 & 3 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ -1 \end{pmatrix}$$

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

$$\boldsymbol{x} = A^{-1}\boldsymbol{b}$$

E Solving simultaneous equations

$$A^{-1} = \frac{1}{\det(A)} adjoint(A)$$

$$x = \frac{1}{\det\begin{pmatrix} 2 & 3 \\ 1 & -2 \end{pmatrix}} \begin{pmatrix} -2 & -1 \\ -3 & 2 \end{pmatrix}^T \begin{pmatrix} 5 \\ -1 \end{pmatrix}$$

$$x = \frac{1}{-7} \begin{pmatrix} -2 & -3 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 5 \\ -1 \end{pmatrix} = \frac{1}{-7} \begin{pmatrix} -10+3 \\ -5-2 \end{pmatrix}$$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{-7} \begin{pmatrix} -7 \\ -7 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} => \left[\begin{matrix} x_1 = 1 \\ x_2 = 1 \end{matrix}\right.$$

Spotle.ai