Linear Algebra

Prof. Asim Tewari
IIT Bombay

Vector, matrix and tensor

Vector

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{n \times 1}$$

n ↑ rows, 1 ↑ Column

Matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & & & \\ a_{31} & \cdots & & \\ a_{m1} & \cdots & & a_{mn} \end{bmatrix} = [a_{ij}]_{m \times n}$$

$m \times n$

↑ rows ↑ Colums.

Vector, matrix and tensor

A_{ijk} a tensor of rank 3.

$x_i \Rightarrow$ vector

a_{ij} is a matrix

Transpose, Addition, Subtraction and Scalar Multiplication

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad A^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$$

$$(A^T)_{ij} = A_{ji}$$

$$(A_{m \times n} + B_{m \times n})_{ij} = a_{ij} + b_{ij}$$

$$(c \underset{m \times n}{A})_{ij} = c\, a_{ij}$$

Matrix Multiplication

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

$$c_{ij} = \sum_k a_{ik} \times b_{kj} = a_{ik} b_{kj}$$

row × Column.

$$\boxed{AB \neq BA}$$

- $A(B+D) = AB + AD$
- $ABD = A(BD) = (AB)D$

vector Dot product.

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{n\times 1} ; \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_{n\times 1}$$

Dot product

$$x \cdot y = \underbrace{x^T}_{1\times n} \underset{n\times 1}{y} = \underbrace{\quad}_{1\times 1}$$

scalar

$$x^T y = y^T x$$

Einstein Summation Notation

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

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

$m \quad p \qquad m \qquad p$

- Summation is performed over repeated index
- No indices appear more than two times in the equation
- Indices which is summed over is called dummy indices appear only in one side of equation
- Indices which appear on both sides of the equation is free indices.

$$a_{ij}\, b_{jk} = \sum_{j} a_{ij}\, b_{jk}$$

$$a_{ii} = \sum_{i=1}^{d} a_{ii} = a_{11} + a_{22} + \cdots + a_{dd}$$

Matrix Multiplication

Dot Product : $x \cdot y = \underbrace{[x_1, x_2 \cdots x_n]}_{x^T} \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}_{y} = \underbrace{x_i y_i}_{\text{scalar}}$

Show that:

$\underbrace{(x^T y)}_{x_i y_i} = \underbrace{(y^T x)}_{y_j x_j}$

Show that:

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

$(A^T)_{ij} = A_{ji}$

Einstein

$(a_{ij} b_{jk})_{ik}$

Square matrix, main diagonal, trace

$A_{m\times n}$ if $m = n \Rightarrow$ sq. matrix

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & \\ \vdots & & & \\ a_{n1} & \cdots & & a_{nn} \end{bmatrix}$$

Main diagonal

Trace = Σ of all elements of the main diagonal

$$= a_{ii} = \sum_{i=1}^{n} a_{ii}$$

Identity and Inverse Matrices

Identity matrix

$I_n \in \mathbb{R}^{n \times n}$ s.t.

$\forall x \in \mathbb{R}^n, \; I_n x = x$

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

Inverse Matrix of matrix A is A^{-1} if

$AA^{-1} = I_n$

$Ax = b \Rightarrow \underbrace{A^{-1}A}_{I_n} x = A^{-1}b \Rightarrow \underbrace{I_n x}_{x} = A^{-1}b$

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

Linear Dependence and Span

Fixed

$Ax = b$ For A^{-1} to exist, this should have exactly one sol.

given given

If the above eqn. has 2 sol, then it should have ∞ sol.

If x, y are two sol; then

$$z = \alpha x + (1-\alpha) y$$ is also a sol.

for any real α.

$$\underset{m \times n}{A} \underset{n \times 1}{x} = \underbrace{A_{:j} x_j}_{\text{Linear Combination}}$$

$$(Ax)_i = A_{ij} x_j = \sum_j A_{ij} x_j$$

$$Ax = A_{:j} x_j$$

Linear Dependence and Span

$$\underset{m\times n}{A}\ \underset{n\times 1}{x} = A_{:j}\, x_j = \sum_{j=1}^{n} v^{(j)} x_j$$

Linear Combination of columns.

$\{v^{(1)}, \ldots, v^{(n)}\}$, $v^{(1)}$: $m\times 1$

Set of all possible points obtained by the linear combination of the vectors is call span.

$Ax = b$ has a sol if 'b' is a part of the span of columns of A $\{v^{(1)}, v^{(2)}, \ldots, v^{(n)}\}$

Linear Dependence and Span

$$\underbrace{A}_{\text{given}}\ \overset{\text{Find}}{x} = \underbrace{b}_{\text{given}}$$

For A^{-1} to exist, this should have exactly one sol.

If the above eqn. has 2 sol, then it should have ∞ sol.

If x, y are two sol; then

$z = \alpha x + (1-\alpha) y$ is also a sol. for any real α.

$$\underset{m\times n}{A}\ \underset{n\times 1}{x} = \underbrace{A_{:j}\, x_j}_{\text{Linear Combination}}$$

$$(Ax)_i = A_{ij}\, x_j = \sum_j A_{ij}\, x_j$$

$$Ax = A_{:j}\, x_j$$

Linear Dependence and Span

For $\underset{m\times n}{A}\ \underset{n\times 1}{x} = \underset{m\times 1}{b}$ to have a sol. 'b' should be a part of span of A.

$b \in R^m$

1. $\Rightarrow$ Column space of A to be $\mathbb{R}^m$

   $n \geq m$ } necessary but not sufficient.

2. Linear independence in the column vectors of A.

A has to be square matrix and have L.I. Column vectors.

$A^{-1}(Ax) = A^{-1}b \Rightarrow \boxed{x = A^{-1}b}$

Norms

$$\|x\|_p = \left[\sum_i |x_i|^p\right]^{1/p}$$

E.g. $\|x\|_2 = \sqrt{\sum x_i^2}$

Norm is a function f that satisfies

- $f(x) = 0 \Rightarrow x = 0$
- $f(x+y) \leq f(x) + f(y)$
- $\forall \alpha \in R ;\ f(\alpha x) = |\alpha| f(x)$

Symmetric matrix, unit vector and orthogonal

$A = A^T$

Unit Vector : $\|x\|_2 = 1$

Two vectors are orthogonal if $x^T y = 0$

Orthonormal

Orthogonal matrix : Columns are mutually orthonormal.

$AA^T = I$

Row wise also orthonormal

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

General Inner Products

Let V be a vector space and $\Omega : V \times V \to R$ be a bilinear mapping that takes two vectors and maps them onto a real number. Then

- Ω is called symmetric if $\Omega(x, y) = \Omega(y, x)$ for all $x, y \in V$, i.e., the order of the arguments does not matter.
- Ω is called positive definite if

  $$\forall x \in V \setminus \{0\} : \Omega(x, x) > 0 \; , \; \Omega(0, 0) = 0$$

- A positive definite, symmetric bilinear mapping $\Omega : V \times V \to R$ is called an inner product on V . We typically write $\langle x, y\rangle$ instead of $\Omega(x, y)$.
- The pair $(V, \langle\cdot, \cdot\rangle)$ is called an inner product space or (real) vector space with inner product. If we use the dot product definition, we call $(V, \langle\cdot, \cdot\rangle)$ a Euclidean vector space.

Inner Product of Functions

$$\langle u, v \rangle := \int_a^b u(x)v(x)dx$$

$$u : \mathbb{R} \to \mathbb{R} \text{ and } v : \mathbb{R} \to \mathbb{R}$$

Kronecker delta

In 3 D space

$$\delta_{ii} = 3$$

$$\delta_{ij}\delta_{jk} = \delta_{ik}$$

$$\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

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

$$\delta_{nm} = \frac{1}{N}\sum_{k=1}^{N} e^{2\pi i \frac{k}{N}(n-m)}$$

Permutation tensor, also called the Levi-Civita tensor or isotropic tensor

$$\epsilon_{ijk} = \begin{cases} 0, & \text{if any two labels are the same} \\ 1, & \text{if } i,j,k \text{ is an even permutation of } 1,2,3 \\ -1, & \text{if } i,j,k \text{ is an odd permutation of } 1,2,3 \end{cases}$$

1, 2, 3 | 2, 3, 1 | 3, 1, 2

$$\epsilon_{ijk}\epsilon_{imn} = \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}$$

Determinant

$$\det(\mathbf{A}) = \varepsilon_{ijk} a_{1i} a_{2j} a_{3k}$$

$$= \sum_i \sum_j \sum_k \epsilon_{ijk} a_{1i} a_{2j} a_{3k}$$

Vector cross product

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ a^1 & a^2 & a^3 \\ b^1 & b^2 & b^3 \end{vmatrix} = \varepsilon_{ijk} \mathbf{e}_i a^j b^k$$

Partial Differentiation and Gradients

$$\frac{\partial f}{\partial x_1} = \lim_{h \to 0} \frac{f(x_1 + h, x_2, \ldots, x_n) - f(x)}{h}$$

$$\vdots$$

$$\frac{\partial f}{\partial x_n} = \lim_{h \to 0} \frac{f(x_1, \ldots, x_{n-1}, x_n + h) - f(x)}{h}$$

$$\nabla_x f = \mathrm{grad} f = \frac{\mathrm{d}f}{\mathrm{d}x} = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} & \frac{\partial f(x)}{\partial x_2} & \cdots & \frac{\partial f(x)}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{1 \times n}$$

Gradients of Vector-Valued Functions

$$f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{bmatrix} \in \mathbb{R}^m .$$

$$\frac{\partial f}{\partial x_i} = \begin{bmatrix} \frac{\partial f_1}{\partial x_i} \\ \vdots \\ \frac{\partial f_m}{\partial x_i} \end{bmatrix} = \begin{bmatrix} \lim_{h \to 0} \frac{f_1(x_1, \ldots, x_{i-1}, x_i + h, x_{i+1}, \ldots, x_n) - f_1(x)}{h} \\ \vdots \\ \lim_{h \to 0} \frac{f_m(x_1, \ldots, x_{i-1}, x_i + h, x_{i+1}, \ldots, x_n) - f_m(x)}{h} \end{bmatrix} \in \mathbb{R}^m$$

Gradients of Vector-Valued Functions

$$\boldsymbol{f}(\boldsymbol{x}) = \begin{bmatrix} f_1(\boldsymbol{x}) \\ \vdots \\ f_m(\boldsymbol{x}) \end{bmatrix} \in \mathbb{R}^m.$$

$$\frac{\partial \boldsymbol{f}}{\partial x_i} = \begin{bmatrix} \frac{\partial f_1}{\partial x_i} \\ \vdots \\ \frac{\partial f_m}{\partial x_i} \end{bmatrix} = \begin{bmatrix} \lim_{h \to 0} \frac{f_1(x_1, \ldots, x_{i-1}, x_i + h, x_{i+1}, \ldots, x_n) - f_1(\boldsymbol{x})}{h} \\ \vdots \\ \lim_{h \to 0} \frac{f_m(x_1, \ldots, x_{i-1}, x_i + h, x_{i+1}, \ldots, x_n) - f_m(\boldsymbol{x})}{h} \end{bmatrix} \in \mathbb{R}^m$$

$$\begin{aligned} \frac{\mathrm{d}\boldsymbol{f}(\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}} &= \begin{bmatrix} \boxed{\frac{\partial \boldsymbol{f}(\boldsymbol{x})}{\partial x_1}} & \cdots & \boxed{\frac{\partial \boldsymbol{f}(\boldsymbol{x})}{\partial x_n}} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial f_1(\boldsymbol{x})}{\partial x_1} & \cdots & \frac{\partial f_1(\boldsymbol{x})}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m(\boldsymbol{x})}{\partial x_1} & \cdots & \frac{\partial f_m(\boldsymbol{x})}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{m \times n} \end{aligned}$$

Gradients of Vector-Valued Functions

$$J = \nabla_x f = \frac{\mathrm{d}f(x)}{\mathrm{d}x} = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} & \cdots & \frac{\partial f(x)}{\partial x_n} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\partial f_1(x)}{\partial x_1} & \cdots & \frac{\partial f_1(x)}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m(x)}{\partial x_1} & \cdots & \frac{\partial f_m(x)}{\partial x_n} \end{bmatrix},$$

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad J(i,j) = \frac{\partial f_i}{\partial x_j}.$$