# Linear Algebra

## Vectors in $\mathbb{R}^n$

$x \in \mathbb{R}^n$

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

## Matrices in $\mathbb{R}^{n \times m}$

$A \in \mathbb{R}^{n \times m}$

$$A = \begin{array}{c c} & \text{column } j \\ \text{row } i & \begin{bmatrix} a_{11} & \cdots & & a_{1m} \\ \vdots & & a_{ij} & \vdots \\ a_{n1} & \cdots & & a_{nm} \end{bmatrix} \end{array}$$

# Vector inner products.

$x, y \in \mathbb{R}^n$

$$x \cdot y = \langle x, y \rangle = \sum_{i=1}^{n} x_i y_i = x^T y$$

$$x^T = [x_1 \cdots x_n]$$

$$[x_1 \cdots x_n] \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

# Vector Norms

inner product induced norm

$$\|x\|_2^2 = x^T x$$ ← square

$$\|x\|_2 = (x \cdot x)^{1/2}$$

$L_2$ or Euclidean norm

$$\|x\|_p = \left( \sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

$L_p$-norm $\quad p \geq 0$

Vector
Norm

$\|x\| \geq 0$ with $\|x\| = 0$ iff $x = 0$

$\|ax\| = a\|x\|$

scalar

$\|x+y\| \leq \|x\| + \|y\|$ Triangle inequality

Lp norms

Matrix vector multiplication

$A \in \mathbb{R}^{n \times m}$ $\quad x \in \mathbb{R}^m$

$Ax \in \mathbb{R}^n$

$$\begin{bmatrix} a_{11} & \sim & a_{1m} \\ \wr & & \wr \\ a_{n1} & \sim & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ | \\ x_m \end{bmatrix} = \begin{bmatrix} (a_{11}x_1 + \sim + a_{1m}x_m) \\ \wr \\ (a_{n1}x_1 + \sim + a_{nm}x_m) \end{bmatrix}$$

vector-norm induced matrix norm

$$\|A\|_p = \sup_{x} \frac{\|Ax\|_p}{\|x\|_p}$$

$$= \sup_{\|x\|_p = 1} \|Ax\|_p$$

supremum (sup): the largest value

$$\sup_{\|x\|=1} \|Ax\|$$

the largest value that $\|Ax\|$ attains while $\|x\|=1$

infemum (inf)

$$\inf_{\|x\|=1} \|Ax\|.$$

# Linear Algebra

## Vectors in $\mathbb{R}^n$

$x \in \mathbb{R}^n$

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

## Matrices in $\mathbb{R}^{n \times m}$

$A \in \mathbb{R}^{n \times m}$

$$A = \begin{array}{cc} & \text{column } j \\ \text{row } i & \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & a_{ij} & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \end{array}$$

Vector inner products.

$x, y \in \mathbb{R}^n$

$$x \cdot y = \langle x, y \rangle = \sum_{i=1}^{n} x_i y_i = x^T y$$

$x^T = [x_1 \sim x_n]$

$[x_1 \sim x_n] \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$

## Vector Norms

inner product induced norm

$\|x\|_2^2 = x^T x$ ← square

$\|x\|_2 = (x \cdot x)^{1/2}$

$L_2$ or Euclidean norm

$$\|x\|_p = \left( \sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

$L_p$-norm $\quad p \geq 0$

## Vector Norm

$\|x\| \geq 0$ with $\|x\| = 0$ iff $x = 0$

$\|a x\| = a \|x\|$

scalar

$\|x + y\| \leq \|x\| + \|y\|$ Triangle inequality

$L_p$ norms

Matrix vector multiplication

$A \in \mathbb{R}^{n \times m}$ $\quad x \in \mathbb{R}^m$

$Ax \in \mathbb{R}^n$

$$\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} (a_{11}x_1 + \cdots + a_{1m}x_m) \\ \vdots \\ (a_{n1}x_1 + \cdots + a_{nm}x_m) \end{bmatrix}$$

vector-norm induced matrix norm

$$\|A\|_p = \sup_{x} \frac{\|Ax\|_p}{\|x\|_p}$$

$$= \sup_{\|x\|_p = 1} \|Ax\|_p$$

supremum (sup): the largest value

$$\sup_{\|x\|=1} \|Ax\|$$

the largest value that $\|Ax\|$ attains while $\|x\|=1$

infemum (inf)

$$\inf_{\|x\|=1} \|Ax\|.$$

Assume $A \in \mathbb{R}^{n \times n}$ (square) and $A^T = A$ (symmetric)

Eigenvalue/vector pairs

$$A = \begin{bmatrix} a_{11} & & \\ & & a_{ji} \\ & a_{ij} & \\ & & a_{nn} \end{bmatrix}$$

$a_{ij} = a_{ji}$

$\forall (i,j)$

for all

* $A v_i = \lambda_i v_i \quad (\lambda_i, v_i)$

$\{(\lambda_1, v_1), \ldots, (\lambda_n, v_n)\}$

$$\begin{bmatrix} 1 & 4 & 5 \\ 4 & 2 & 6 \\ 5 & 6 & 3 \end{bmatrix}$$

$$v_i^T v_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases} \quad \text{(convention)}$$

$\{v_1, \ldots, v_n\}$ is an orthonormal basis for $\mathbb{R}^n$

$$A\begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix} = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

$$A \underbrace{\cancel{V V^T}}_{I} = V \quad \Lambda \quad V^T \iff \boxed{A = V\Lambda V^T}$$

orthogonal matrix

$$V^{-1} = V^T \qquad \boxed{VV^T = V^TV = I}$$

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

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

$$A = V \Lambda V^T = \sum_{i=1}^{n} \lambda_i v_i v_i^T \qquad A \in \mathbb{R}^{n \times n} \quad A^T = A$$

Singular Value Decomposition (SVD)

$A \in \mathbb{R}^{n \times m}$

$$A = \sum_{i=1}^{n} \sigma_i (u_i v_i^T)$$

rank-1 matrix

$\boxed{n < m}$

$U = [u_1 \dots u_n] \in \mathbb{R}^{n \times n}$ orthogonal

$V = [v_1 \dots v_m] \in \mathbb{R}^{m \times m}$ orthogonal

$$S = \left[\begin{array}{ccc:c} \sigma_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & \sigma_n & \end{array}\right] \in \mathbb{R}^{n \times m}$$

$n < m$

diagonal

$$AV = U S \cancel{V^T V}$$

$n \times m \quad n \times n \quad n \times m \quad m \times m$

$\boxed{AV = US}$

$$\star\ \mathrm{tr}(ABC) = \mathrm{tr}(CAB) = \mathrm{tr}(BCA)$$

$$\mathrm{tr}\left(E\left[x^T \Sigma x\right]\right) = E\left[\mathrm{tr}(x^T \Sigma x)\right] = E\left[\mathrm{tr}(x x^T \Sigma)\right]$$

$$= \mathrm{tr}\left[E\left[x x^T \Sigma\right]\right]$$

$$= \mathrm{tr}\left[E\left[x x^T\right] \Sigma\right] = \mathrm{tr}(\Sigma^2)$$

$\Sigma = \mathrm{cov}(x)$

$$\mathrm{trace}(A) = \sum_{i=1}^{n} a_{ii}$$

$$A = \begin{bmatrix} a_{11} & & X \\ & \ddots & \\ X & & a_{nn} \end{bmatrix}$$

$E[x] = 0$

$\Sigma = E[x x^T]$

$A \in \mathbb{R}^{n \times n}$

# Derivatives

$w, x \in \mathbb{R}^n$

Ex) $f(x, w, b) = w^T x + b = w_1 x_1 + w_2 x_2 + \dots + w_n x_n + b$

$0 + x_2 + 0 + \dots 0 + 0$

$$\frac{\partial f(x, w, b)}{\partial w} = \left[ \frac{\partial f}{\partial w_1} \quad \frac{\partial f}{\partial w_2} \quad \dots \quad \frac{\partial f}{\partial w_n} \right] = [x_1 \; x_2 \; \dots \; x_n] = x^T$$

$$\frac{\partial f(x, w, b)}{\partial b} = 1$$

Ex

$$f(x; A, b, c) = x^T A x + b^T x + c$$

$b, x \in \mathbb{R}^n$
$A \in \mathbb{R}^{n \times n}$
$c \in \mathbb{R}$

$$\frac{\partial f}{\partial c} = 1 \qquad \frac{\partial f}{\partial b} = x^T \qquad \frac{\partial f}{\partial a_{ij}} = x_i x_j$$

$$A = A^T$$

$$\frac{\partial f}{\partial x} = \frac{\partial (x^T A x)}{\partial x} + \frac{\partial (b^T x)}{\partial x} + \frac{\partial c}{\partial x}$$

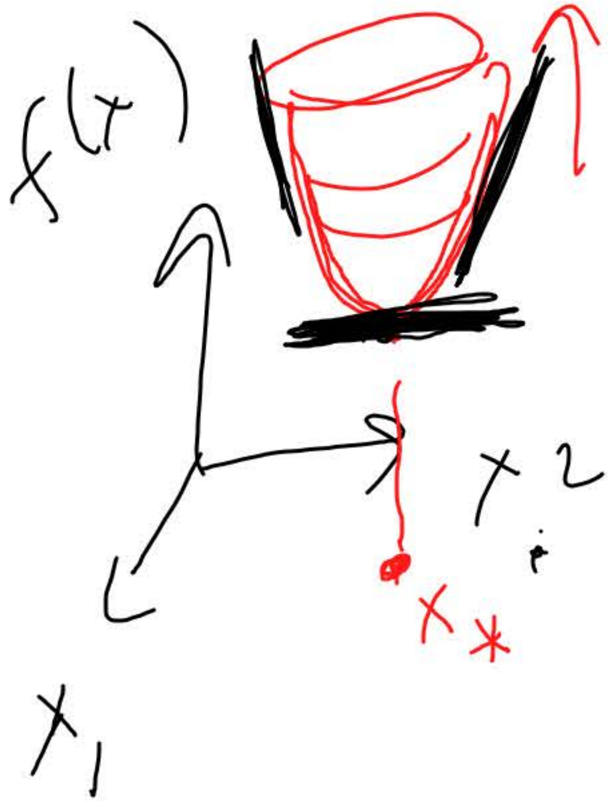
column

$$= \left(2 x^T A + b^T + 0\right)^T$$

$$2Ax + b$$

Ex) $f(x) = x^T A x + b^T x + c$ $\quad \left\{ A^T = A \in \mathbb{R}^{n\times n},\ b, x \in \mathbb{R}^n,\ c \in \mathbb{R} \right.$

$$x_* = \arg\min_{x\in\mathbb{R}^n} f(x)$$




$$\min_x f(x)$$

$$\frac{\partial f}{\partial x^T} = 2Ax + b + 0 \Big|_{x=x_*} \equiv 0$$

$$2Ax_* + b = 0$$

$$\hookrightarrow x_* = -\frac{1}{2}A^{-1}b$$

$A \succ 0$ positive definite

Defn $A \succ 0$ iff $x^T A x \geq 0 \ \forall x \neq 0$

$\Updownarrow$

$A \succ 0$ iff $\lambda_A \geq 0$ all e.values

positive semi-definite

Ex

$$\min_x \frac{x^TAx}{x^TBx} \qquad f(x)$$

$$\frac{\partial f}{\partial x^T} = \frac{\frac{\partial}{\partial x^T}(x^TAx)\,(x^TBx) - (x^TAx)\frac{\partial}{\partial x^T}(x^TBx)}{(x^TBx)^2}$$

$$0 = (x^TBx)\cancel{2}Ax - (x^TAx)\cancel{2}Bx$$

$$Ax = \left(\frac{x^TAx}{x^TBx}\right)Bx$$

$A^{n\times n}, B^{n\times n}$ sp. sym. real pos. definite

$B^{-1}$ $B^{-1}$

$$Av = \lambda Bv$$

$(\lambda, v)$ is a generalized e. value/vector for $(A, B)$

$$\boxed{(B^{-1}A)v = \lambda v}$$