$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

$$[1 \quad x \quad x^2 \quad \cdots \quad x^n]$$

$$[a_0 \quad a_1 \quad \cdots \quad a_n]$$

$$X \quad \{1, 2, 3, \ldots, n\}$$

$$P_i = \text{Prob}(X = i)$$

$$1^T P \qquad = 1$$

$$1^T P \longrightarrow [P_1 \quad P_2 \quad \cdots \quad P_n]$$

$$[1 \quad 1 \quad 1 \quad \cdots \quad 1]$$

$$f(x)$$

$$f(x_0 + \epsilon) = f(x_0) + f'(x_0)\epsilon + \frac{1}{2} f''(x_0)\epsilon^2 + \frac{1}{3!} f'''(x_0)\epsilon^3 + \cdots$$

$$f(x_0 + \epsilon) \approx f(x_0) + f'(x_0)\epsilon$$

first order Taylor series.

$$f(x_1, x_2, x_3, \dots, x_n)$$

$$f(x_1+\epsilon_1, x_2+\epsilon_2, \dots, x_n+\epsilon_n)$$

$$\approx f(x_1, x_2, \dots, x_n) + \epsilon_1 \frac{\partial f}{\partial x_1} + \epsilon_2 \frac{\partial f}{\partial x_2} + \dots + \epsilon_n \frac{\partial f}{\partial x_n}$$

$$= f(x_1, x_2 \dots, x_n) + \epsilon^T \nabla f$$

$$\epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}, \quad \nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \underline{x}^T x = [x_1 \ x_2 \ \dots \ x_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$\|\underline{x} + \underline{y}\| \qquad \|x\| + \|y\|$$

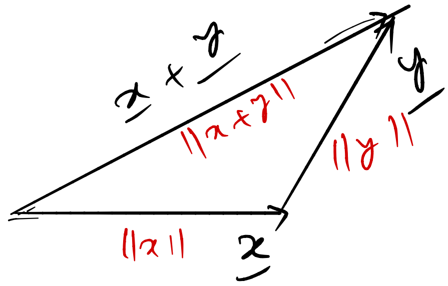
$$\boxed{\|\underline{x} + \underline{y}\|^2 = \|x\|^2 + \|y\|^2 + 2x^T y}$$

Excer:

$$(\|\underline{x}\| + \|\underline{y}\|)^2 = \|x\|^2 + \|y\|^2 + 2\|x\|\|y\|$$

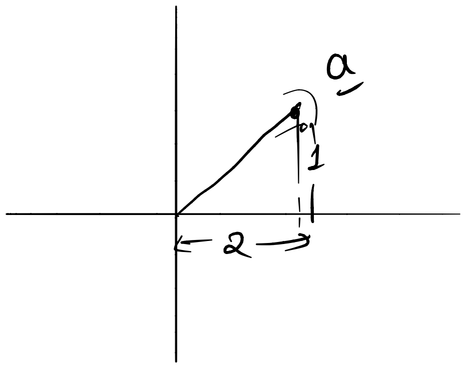
$$\|\underline{x}+\underline{y}\| \leq \|\underline{x}\|+\|\underline{y}\|$$

Triangle inequality

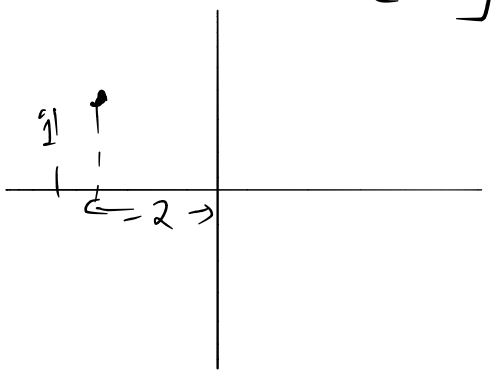


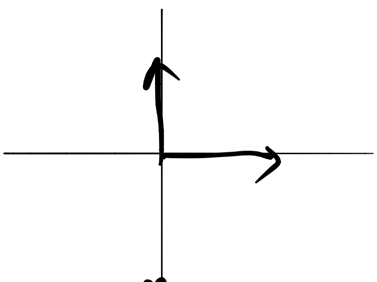
$$\|\underline{x}\| = 0$$

$$\underline{x} = \begin{bmatrix} 1 \\ 0 \\ i \end{bmatrix}$$



$$\underline{a} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

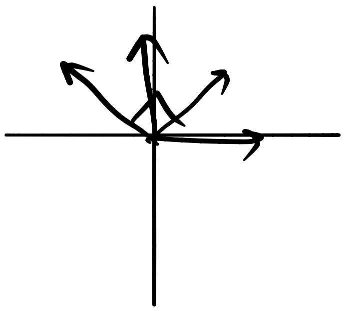




in $\mathbb{R}^n$

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \vdots \end{bmatrix}$$

$$- - - \; e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

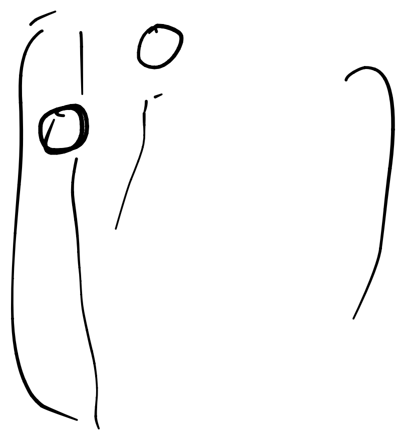


Shopping list.

$$\underline{a} = \begin{array}{c} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \end{array} \begin{array}{l} \leftarrow \text{apples} \\ \leftarrow \text{bana} \\ \leftarrow \text{orangs} \end{array}$$

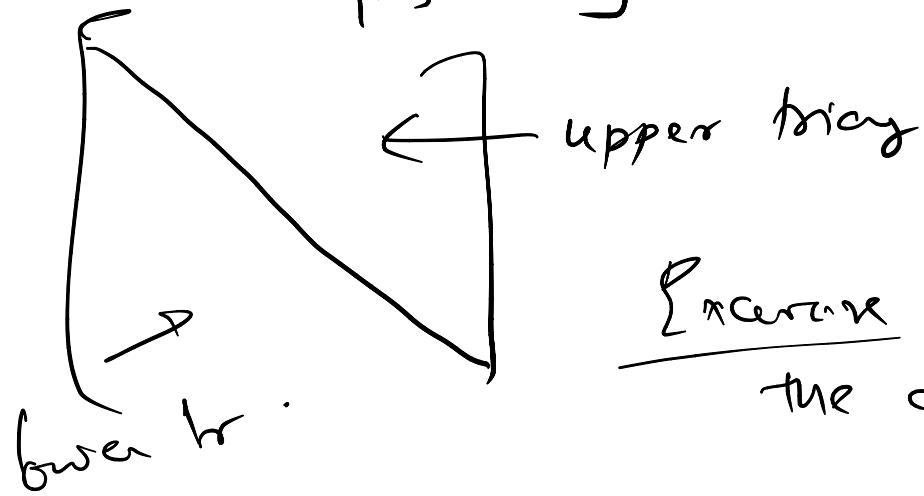
$\underline{a}^T \underline{b} = 0$?

$$\begin{bmatrix} A_{m\times n} & B_{m\times p} \\ C_{k\times n} & D_{k\times p} \end{bmatrix}$$



$n \times n$

$$\begin{bmatrix} 1 & 1.5 \\ 1.5 & 2 \end{bmatrix}$$



Excercise: Find the degrees of freedom.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \ddots & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 \\ \end{bmatrix}$$

$$\underbrace{[x_1 \ x_2 \ \cdots \ x_n]}_{x^T} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \sum x_i y_i$$

$$\underset{1\times 1}{b} \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}_{1\times n}$$

$$= \begin{bmatrix} ba_1 & ba_2 & \cdots & ba_n \end{bmatrix}_{1\times n}$$

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}_{n\times 1} b_{1\times 1} = \begin{bmatrix} a_1 b \\ a_2 b \\ \vdots \\ a_n b \end{bmatrix}_{n\times 1}$$

Outer product.

$\underline{x}$ $\underline{y}$ $\quad \underline{x}^T \underline{y}$

outer product $\underline{x}\,\underline{y}^T$

$n\times 1 \quad 1\times n$

$n\times n$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \begin{bmatrix} y_1 & y_2 \end{bmatrix} = \begin{bmatrix} x_1 y_1 & x_1 y_2 \\ x_2 y_1 & x_2 y_2 \end{bmatrix}$$

$$\mathbb{R}^m \longleftarrow \boxed{A} \longleftarrow \mathbb{R}^n \qquad A : m \times n.$$

$A\underline{x} = 0$ for any $\underline{x}$

$A = 0$.

Columns of V are orthogonal to each other

& unit norm.

$$\underbrace{V^T V = I}\ ; \qquad \underbrace{V V^T = I}$$

$$V = [v_1\ v_2\ \cdots\ v_n]$$

$$V^T V = \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{bmatrix} [v_1\ v_2\ \cdots\ v_n]$$

$$= \begin{bmatrix} v_1^T v_1 & v_1^T v_2 & v_1^T v_3 & \cdots \\ v_2^T v_1 & v_2^T v_2 & v_2^T v_3 & \cdots \\ \vdots & & \ddots & \\ \vdots & & & \ddots \end{bmatrix}$$

$$\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$
e_1 e_2

$$V = \begin{bmatrix}2 & 0\\ 0 & 3\end{bmatrix}$$

$$V^T V = \begin{bmatrix}4 & 0\\ 0 & 9\end{bmatrix}$$

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

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

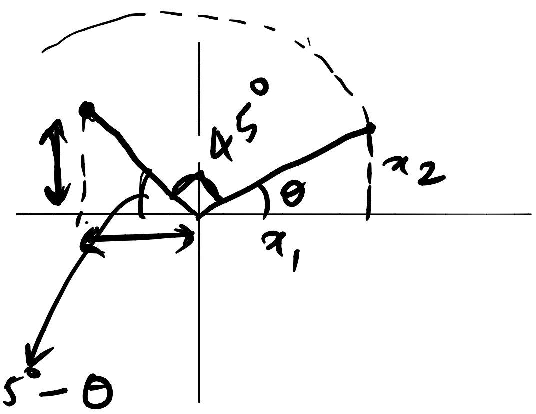
$$\begin{bmatrix}-2 & 0\\ 0 & -2\end{bmatrix}$$

$$\begin{bmatrix}1/\sqrt{2} & -1/\sqrt{2}\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{bmatrix}$$

rotating by $45°$

$$\begin{bmatrix}1/\sqrt{2} & -1/\sqrt{2}\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} = \left.\begin{bmatrix}\frac{x_1 - x_2}{\sqrt{2}}\\ \frac{x_1 + x_2}{\sqrt{2}}\end{bmatrix}\right\} \checkmark$$

Excercise



Excercise: Rotation by θ is achieved by

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

If V is an orthogonal matrix

$$\|V\underline{x}\|_2 = \|\underline{x}\|_2$$

$$\|V\underline{x}\|_2^2 = (V\underline{x})^T(V\underline{x})$$

$$= \underline{x}^T V^T V \underline{x}$$

$$= \underline{x}^T I \underline{x} = \underline{x}^T \underline{x}$$

$$= \|\underline{x}\|^2$$

$$\|\underline{y}\|^2 = \underline{y}^T \underline{y}$$

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

$$\underset{m\times n}{A} \quad \underset{n\times k}{B}$$

$$\underset{k\times n}{B^T} \quad \underset{n\times m}{A^T}$$

For any matrix A, A^TA is the Gram matrix of A.

$$\underline{x} \xleftarrow[BA\underline{x}]{} \boxed{B} \xleftarrow[A\underline{x}]{} \boxed{A} \leftarrow \underline{x}$$

$$\boxed{BA = I} \longrightarrow (AB = I)$$

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_1 + x_2 \end{bmatrix}$$

$$A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$A \begin{bmatrix} 2 \\ 2 \end{bmatrix} = 2 \cdot A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 0$$

$$\left.\begin{aligned} x_1 + x_2 &= 5 \\ x_1 - x_2 &= 10 \end{aligned}\right\} \quad x_1 = 7.5$$

$$x_1 = 10$$

Q1: P and Q are $n \times n$ orthogonal matrices.

is PQ orthogonal?.

$$(PQ)^T PQ = Q^T P^T P Q = I$$

$$PQ\underline{x} \leftarrow \boxed{P} \leftarrow \boxed{Q} \leftarrow \underline{x}$$

Q2: all the entries in PQ are between -1 and 1.

Q3

$$\underset{(n+1)\times(n+1)}{D} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} a_1 \\ \vdots \\ (n-1)a_{n-1} \\ na_n \\ 0 \end{bmatrix}$$

mimics polynomial derivative

$$\boxed{D^{n+1} = ?}$$

$$\begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} d_1 x_1 \\ d_2 x_2 \end{bmatrix}$$

Hint: write down D entrywise.

$$D \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} a_1 \\ 0 \end{bmatrix}$$

$$D = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \checkmark$$

D: polynomial derivative.

$$(\quad) \leftarrow \boxed{D} \leftarrow \underbrace{(a_0 \; a_1 \; a_2 \; \cdots \; a_n)}$$

$$\underbrace{\sum_{k=1}^{n} k a_k x^{k-1}} \quad \xleftarrow{\;D\;} \quad \underbrace{\sum_{k=0}^{n} a_k x^k}$$

x^n

$a_n \leftarrow x^n$

Q4: $A = \begin{bmatrix} \cos \pi/5 & -\sin \pi/5 \\ \sin \pi/5 & \cos \pi/5 \end{bmatrix}$

$A^{10} = I$

$A^5 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$

$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

$A\underline{x} = \lambda \underline{x}$

If $\lambda = 0$,

$A\underline{x} = \underline{0}$

$D\underline{a} = \begin{bmatrix} a_1 \\ 2a_2 \\ 3a_3 \\ \vdots \\ n a_n \\ 0 \end{bmatrix}$ $\leftarrow$ derivative of polynomial.

eigen vectors of D correspond to constants.

$\begin{bmatrix} 1 & -1/2 \\ -1/2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \end{bmatrix}$

$x_2 = 2[1-\lambda]x_1$

$x_2 = 2[1-\frac{1}{2}]x_1$

$= x_1$

$x_1 - x_2/2 = \lambda x_1$

$-\frac{1}{2}x_1 + x_2 = \lambda x_2$

$-\frac{1}{2}x_1 + 2(1-\lambda)x_1 = \lambda 2[1-\lambda]x_1$

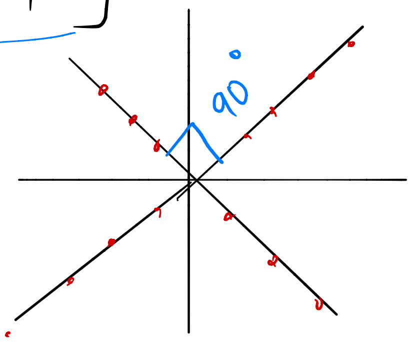
$\lambda = \frac{3}{2}, \frac{1}{2}.$

$2 \cdot \cancel{6} = 3 \cdot \cancel{6}$

Eigenspaces

If $\underline{x}$ is an eigen vector, so is $\alpha\underline{x}$

Any vector of the type $\begin{bmatrix} x_1 \\ x_1 \end{bmatrix}$ is an eigen vector of $\begin{bmatrix} 1 & -1/2 \\ -1/2 & 1 \end{bmatrix}$ with eigen value $1/2$.



$\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ ✓ $\begin{bmatrix} -1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$

For $\lambda = 3/2$ we have $x_1 = -x_2$

$\begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$ ✓, $\begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

If A is symmetric, eigen vectors corresponding to distinct eigen values are orthogonal.

Spectral decomposition:

If $A \in \mathbb{R}^{n \times n}$ is symmetric, then A has n eigen values $\lambda_1, \lambda_2, \ldots, \lambda_n$ and n eigen vectors $v_1, v_2, \ldots, v_n$; such that $(v_1, v_2, \ldots, v_n)$ forms an orthogonal set.

$$Av_1 = \lambda_1 v_1, \quad Av_2 = \lambda_2 v_2, \quad \ldots \quad Av_n = \lambda_n v_n$$

$$A\underbrace{[v_1 \; v_2 \; \ldots \; v_n]}_{V} = [Av_1 \; Av_2 \; \ldots \; Av_n]$$

$$= [\lambda_1 v_1 \; \lambda_2 v_2 \; \ldots \; \lambda_n v_n]$$

$$= \underbrace{[v_1 \; v_2 \; \ldots \; v_n]}_{V} \underbrace{\begin{bmatrix} \lambda_1 & 0 & 0 & \ldots \\ 0 & \lambda_2 & 0 & \ldots \\ 0 & 0 & \lambda_3 & \ldots \end{bmatrix}}_{D}$$

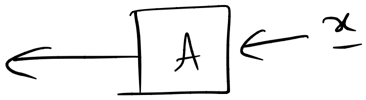
$AVV^T = VDV^T$

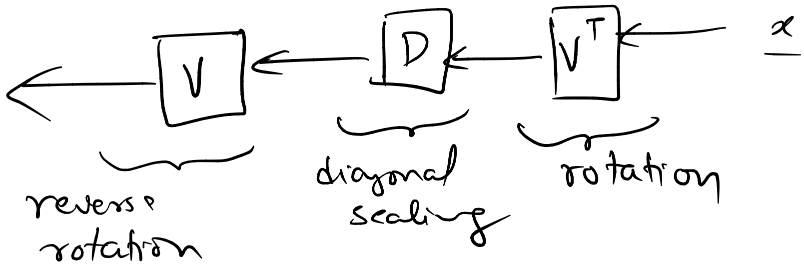
$\Rightarrow A = VDV^T$

Any symmetric matrix

$$A = VDV^T$$

- the columns of V are eigenvectors
- D is a diagonal matrix with diagonal entries as eigenvalues
- V is orthogonal.





Are the eigen vectors / eigen values unique? .

$$A = VDV^T$$

What are V, D for I ?

$D = I$, V can be any orthogonal matrix

$$V\underbrace{D}_{I}V^T = VV^T$$

$$= I$$

$$\det(A) = \prod_{i=1}^{n} \lambda_i \quad \text{(for symmetric matrices)}$$

$$\sum_{i=1}^{n} \lambda_i = \text{trace}(A) = \text{sum of diagonal elements.}$$

$$\boxed{\text{trace}(AB) = \text{trace}(BA)}$$

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \qquad \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{12} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$a_{11}b_{11} + \underline{a_{12}b_{21}} + \underline{a_{21}b_{12}} + \underline{a_{22}b_{22}}$$

$$\mathrm{tr}(A) = \mathrm{tr}(\underbrace{V}_{A}\underbrace{DV^T}_{B})$$

$$= \mathrm{tr}(DV^TV) = \mathrm{tr}(D) = \sum_{i=1}^{n} \lambda_i$$

$$\det(AB) = \det(A)\det(B)$$

$$\det(A) = \det(VDV^T) = \det(V)\det(D)\det(V^T)$$

$$= \det(D)$$

$$= \prod_{i=1}^{n} \lambda_i .$$

$VV^T = I$

$\det V \det V^T = 1$

Least squares

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

A is tall $m \times n$ $\quad m \geq n$.

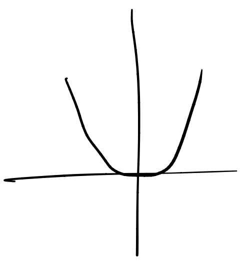
Find $\underline{x}$ such that $\underbrace{A\underline{x} - \underline{b}}_{\text{residuals}}$ is "small".

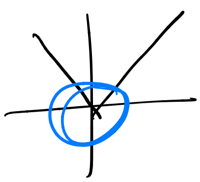
Find $\underline{x}$ such that $\|A\underline{x} - \underline{b}\|_2^2$ is small,

Cost f^n

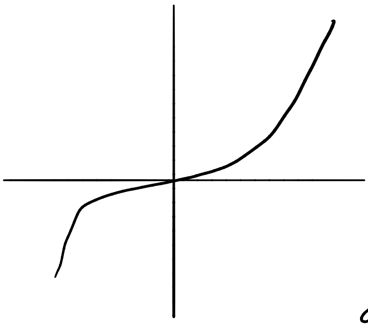
$$\begin{bmatrix} a_1^T x \\ a_2^T x \\ \vdots \\ a_n^T x \end{bmatrix} - \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \longrightarrow \sum_{i=1}^{n} (a_i^T x - b_i)^2$$

$$\underline{x} = \boxed{} \qquad \sum_{i=1}^{n} |a_i^T x - b_i|$$





Find $\underline{x}$ such that $\|A\underline{x} - \underline{b}\|^2$ is minimized.



$f(x) = x^3$

$f'(x) = 3x^2$

"Saddle points"

at $x = 0$, $f'(0) = 0$.

Convex function

$f(x) = \|x\|$, $\|x\|^2$ are convex.

Take first derivative of $\|A\underline{x} - \underline{b}\|^2$ and set to zero.

$A: m \times n$

$b: m \times 1$

$\underline{x}: n \times 1$

$$\|y\|_2^2 = y^T y$$

$$\|A\underline{x} - \underline{b}\|^2 = (A\underline{x} - \underline{b})^T (A\underline{x} - \underline{b})$$

$$= (x^T A^T - b^T)(Ax - b)$$

$$f(x) = x^T A^T A x - b^T A x - x^T A^T b + b^T b$$

$$\underbrace{(b^T A x)^T}_{1 \times 1} = \underbrace{x^T A^T b}$$

$$= x^T A^T A x - 2 b^T A x + b^T b$$

ax^2 $\quad$ cx

Exercise: $\nabla f(x) = 2A^T A x - 2A^T b$ $\quad$ $2ax$

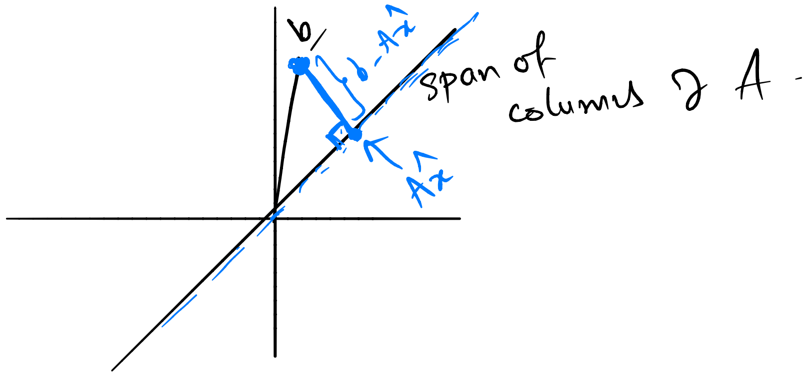
$$\nabla f(x) = 0$$

$$\Rightarrow A^T A x = A^T b$$

Assume $\underbrace{A^T A}_{n \times n}$ is invertible, $\boxed{x = (A^T A)^{-1} A^T b}$

distance between $\underline{b}$ and Ax $\quad \|A\underline{x} - \underline{b}\|^2$

span of column of a.



$$\underbrace{A\underline{x}}_{\text{span}} = \underline{b}$$

$\underline{b} - A\hat{x}$ is perpendicular to the span.

$$(b - A\hat{x})^T A x = 0 \quad \text{for any } x.$$

$$\Rightarrow (b - A\hat{x})^T A = 0$$

$$\Rightarrow A^T A x = A^T b$$

$$\Rightarrow x = (A^T A)^{-1} A^T b.$$