Optimization - Summary of Background Material.

① Linear Algebra ② Analysis ③ Calculus.

Vector

$$x \in \mathbb{R}^n \quad \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{C}^n, \quad A \text{ matrix} \quad A \in \mathbb{R}^{m \times n}$$

rows

Cols

Pos definite matrix → $x^T A x > 0 \quad \forall x \neq 0$

$$[\quad]\begin{bmatrix} & \end{bmatrix}\begin{bmatrix} \\ \end{bmatrix}$$

Pos. semidefinite matrix $x^T A x \geq 0 \quad \forall x \neq 0$

↳ Gram matrix $\{v_i\}_{i=1}^k$, $G_{ij} = v_i^T v_j$

$G \to X^T X$ where $X = \begin{bmatrix} v_1 & v_2 & \cdots & v_k \\ | & | & & | \end{bmatrix} \begin{bmatrix} \\ \end{bmatrix}$

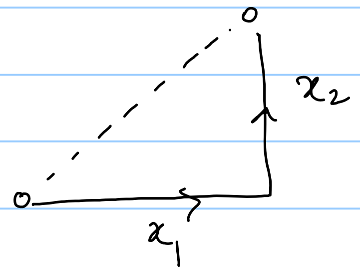
Take $y \in \mathbb{R}^n$ $\quad y^T G y = y^T X^T X y = \|X y\|^2 = \left[\sum_i (v_i y_i)\right]^2 \geq 0$

when $\{v_i\}$ are lin indepn $\rightarrow$ > 0 P.D

" " " " depen ≥ 0 P.S.D

Vector Norms

$$\|x\|_p = \left[\sum_i |x_i|^p\right]^{1/p}, \quad 1 \leq p < \infty, \quad \text{Most common, } p = 2.$$



$p = 1 \rightarrow L_1$ norm

$p = \infty, \quad \|x\|_\infty = \max_i |x_i|$

a) $\|x\| = 0 \Rightarrow x = 0$

b) $\|\alpha x\| = |\alpha| \, \|x\| \quad \alpha \in \mathbb{R}$

(c) $\|x + z\| \leq \|x\| + \|z\|$

$\forall x, z \in \mathbb{R}^n$

$p < 1$

Matrix norms $\longrightarrow$ (1) Consistent matrix norm $\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$

(2) Frobenius norm $\|A\|_F = \left[\sum_{i,j} |A_{ij}|^2 \right]^{1/2}$

$\|A\|_2$ = largest eigen value of $(A^T A)^{1/2}$

$= \sigma_1$ (largest singular value)

Condition number, $K(A) = \|A\| \, \|A^{-1}\|$ (for any norm)

$Ax = b$

fundamental subspaces of a matrix. $A \in \mathbb{R}^{m \times n}$

Row space $\in \mathbb{R}^n$	Col. space $\in \mathbb{R}^m$
$\{x \mid x = A^T y,\ y \in \mathbb{R}^m\}$	$\{y \mid y = Ax,\ x \in \mathbb{R}^n\}$
Nul space $\in \mathbb{R}^n$	Left Null space
$\{x \mid Ax = 0\}$	$\{y \mid A^T y = 0\}$
$= \mathbb{R}^n$	$= \mathbb{R}^m$

$\perp$ $\perp$

Eigen decomposition:

↳ $Ax = \lambda x, \quad \lambda \in \mathbb{R}, \mathbb{C}$

A is square.

↳ A is symmetric → Eig values are real, λ_i's.

Eig vectors are q_i

Eig decomp $A = \sum \lambda_i q_i q_i^T$ outer product

Rank 1

$= \boxed{Q \Lambda Q^T}$

$\begin{bmatrix} q_1 & \cdots & q_n \\ | & & | \end{bmatrix}$ $\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$ orthogonal matrix

A is pos. def. $\implies \lambda_i$'s > 0

SVD Singular Value Decomp square, orthogonal

$A = U \Sigma V^T$

largest

$S = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \end{bmatrix}$

$\begin{bmatrix} S \\ 0 \end{bmatrix}$ A is tall

$\begin{bmatrix} S \end{bmatrix}$ A is sq.

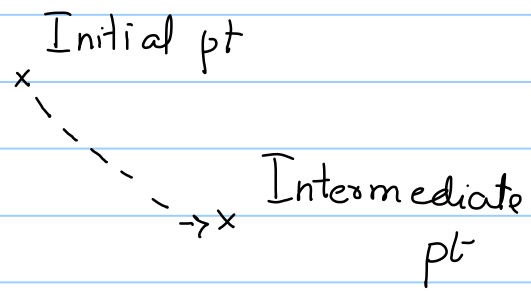
$\begin{bmatrix} S & 0 \end{bmatrix}$ A is fat

$\kappa(A) = \sigma_1 / \sigma_n$ Cond(κ) $1 \leq \kappa < \infty$

↳ Square A: $\text{trace}(A) = \sum_i A_{ii} = \sum_i \lambda_i$

$\det(A) = \prod \lambda_i$

— x —



(2) Analysis

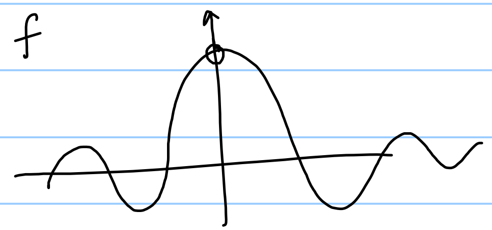
Sequence of vectors → $\{x_i\}$ of points in $\mathbb{R}^n$

↳ Convergence: A seq is said to converge to a point x $\lim_{k \to \infty} x_k \to x$; if for any $\varepsilon > 0$, there is an index K, such that $\underline{\| x_k - x \|_2 \leq \varepsilon}$ $\forall k > K$.

↳ A sequence is called a Cauchy sequence if $\varepsilon > 0$ there exists an integer K such that $\| x_k - x_\ell \| \leq \varepsilon$ for all $k, \ell \geq K$.

↳ Scalar sequences $\{t_k\} \in \mathbb{R}$

$$f(x) = \frac{\sin x}{x}, \quad x \neq 0$$



Let S be a nonempty subset of $\mathbb{R}$

a) An upper bound of the set S is $u \in \mathbb{R}$ s.t. $y \leq u \ \forall \ y \in S$

b) Least upper bound $\leftrightarrow$ Supremum (if it exists)

c) If UB $\in S$ $\rightarrow$ we call it the maximum

↳ Convergence → How to quantify. 'Q' → quotient

1) Q-linear : $\dfrac{\| x_{k+1} - x^* \|}{\| x_k - x^* \|} \leq r$ for all large k and $\boxed{r \in (0,1)}$

2) Q-Superlinear : $\lim_{k\to\infty} \dfrac{\| x_{k+1} - x^* \|}{\| x_k - x^* \|} = 0$

3) Q - quadratic : $\dfrac{\| x_{k+1} - x^* \|}{\| x_k - x^* \|^2} \leq M$, $M > 0$

e.g. What is the rate of convergence of $x_k = \frac{1}{k!} \to x^* = 0$

$$\frac{|x_{k+1}|}{|x_k|} = \frac{1}{k+1} \quad \text{as } k \to \infty, \ \to 0 \Rightarrow \text{superlinear}$$

$$\frac{|x_{k+1}|}{|x_k|^2} = \frac{k!}{k+1} \to \infty \quad \text{as } k \Uparrow$$