LECTURE 1

Vector space E is $\mathbb{R}$ or $\mathbb{C}$.

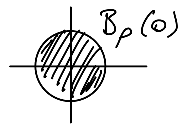
<u>Def</u>: (Norm on E) Function $\|\cdot\| : E \to [0,\infty)$ s.t. $\forall x,y \in E$,

- $\|x\| \geq 0$ and $\|x\| = 0 \iff x = 0$;
- $\|\lambda x\| = |\lambda| \, \|x\|$;
- $\|x+y\| \leq \|x\| + \|y\|$.

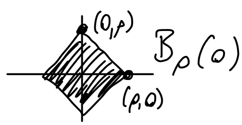
<u>Def</u>: $(E, \|\cdot\|)$ = normed vector space.

<u>Example</u>: on $\mathbb{R}^n$,

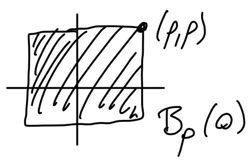
- ℓ^2 norm: $\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}$, $x = \sum_i x_i e_i$.



- ℓ^1 norm: $\|x\|_1 = \sum_{i=1}^{n} |x_i|$



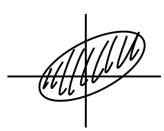
- ℓ^∞ norm: $\|x\|_\infty = \max_i \{x_i\}$



- ℓ^p norm: $\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$, $1 \leq p < \infty$.

- if A is symmetric positive semi-definite

$$\|x\|_A = \sqrt{x^T A x}$$



Claim: $\ell^2, \ell^1, \ell^\infty$ are all equivalent (i.e., if a set is open w.r.t. one of them then it's open w.r.t. the other).

Def: If E is a vector space, $\|\cdot\|_a$ and $\|\cdot\|_b$ are equivalent if $\exists\, c_1, c_2 \in \mathbb{R}_{>0}$ so that $\forall x \in E$,

$$\|x\|_a \leq c_1 \|x\|_b \quad \text{and} \quad \|x\|_b \leq c_2 \|x\|_a .$$

$C = c_2$

$c = 1/c_1$

Can also define it w/ $c, C > 0$ s.t.

$$c\|x\|_a \leq \|x\|_b \leq C\|x\|_a \quad \forall x \in E.$$

$$\Longleftrightarrow \quad \frac{1}{c_2}\|x\|_b \leq \|x\|_a \leq c_1\|x\|_a \quad \forall x \in E.$$

Ex: On $\mathbb{R}^n$

$$1\,\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty, \quad \forall x \in E.$$

Thm: If E is a finite dimensional vector space and $\|\cdot\|_a$, $\|\cdot\|_b$ are norms on E, then they are equivalent.

LECTURE 2

MATRIX NORMS

$M_n(\mathbb{R})$ = $n \times n$ matrices w/ $\mathbb{R}$-entries.

Def: (Matrix norm) A matrix norm on $M_n(\mathbb{C})$ is a norm which is also submultiplicative i.e., $\|AB\| \leq \|A\| \, \|B\|$.

Cor: If $\|\cdot\|$ is a matrix norm $\|Id\| \geq 1$.

$\underset{\circlearrowleft}{\|Id\|} = \|Id^2\| \leq \|Id\| \, \|Id\|$

Cor: If P is a projection (i.e., $P^2 = P$), then $\|P\| \geq 1$.

Def: (Spectral Radius) $\rho(A) := \max_i \{|\lambda_i| \in Spec(A)\}$

$= \max(abs(eig(A)))$ ← MATLAB notation

Prop: $\rho(A) \leq \|A\|$, $A \in M_n(\mathbb{C})$.

Pf: Let $\lambda_0 \in eig(A)$ be such that $|\lambda_0| = \rho(A)$. Let $u \in \mathbb{C}^n$ be an eigenvector $Au = \lambda_0 u$. Then

$$|\lambda_0| \, \|uu^T\| = \|\lambda_0 uu^T\| = \|A \underbrace{uu^T}_{\in M_n(\mathbb{C})}\| \leq \|A\| \, \|uu^T\|$$

So, $|\lambda_0| \leq \|A\|$ i.e., $\rho(A) \leq \|A\|$. ∎

Note: $u^T u \in \mathbb{R}$.

Def: (Frobenius Norm) For $A \in M_n(\mathbb{R})$ (or $\mathbb{C}$),

$$\|A\|_F := \sqrt{\sum_{i,j} |a_{ij}|^2} = \sqrt{tr(A^*A)} = \sqrt{tr(AA^*)}$$

$\boxed{\|Id\|_F = \sqrt{n}}$

Thm: $\|\cdot\|_F$ on $M_n(\mathbb{C})$,

(i) $\|AB\|_F \leq \|A\|_F \|B\|_F$ (i.e., $\|\cdot\|_F$ is a matrix norm)

(ii) For any unitary matrices U, V,

$$\|A\|_F = \|VA\|_F = \|AU\|_F = \|VAU\|_F.$$

(iii) For all $A \in M_n(\mathbb{C})$

$$\sqrt{\rho(A^*A)} \leq \|A\|_F \leq \sqrt{n}\sqrt{\rho(A^*A)}.$$

Pf: (i) $\|AB\|_F^2 = \sum_{i,j} |(AB)_{ij}|^2 = \sum_{i,j} \left| \sum_k A_{ik} B_{kj} \right|^2$

$$= \sum_{i,j} \left(\sum_k |A_{ik}| |B_{kj}| \right)^2 \leq \|A\| \|B\|$$

Cauchy-Schwarz

(ii) $\|A\|_F = \sqrt{tr A^*A} = \sqrt{tr A^*V^*VA}$

$\qquad\qquad\qquad\qquad\qquad\qquad \|$

$\|VA\|_F = \sqrt{tr((VA)^*(VA))}$ and so on.

(iii) $\|A\|_F = \sqrt{\operatorname{tr} A^*A}$ and A^*A is symmetric nonnegative definite

so $\operatorname{eig}(A^*A) \in [0, \infty)$. $\leftarrow$ $(A^*A)u = \lambda u$

$$u^*A^*Au = \lambda u^*u$$

$$(Au)^*(Au) = \lambda u^*u$$

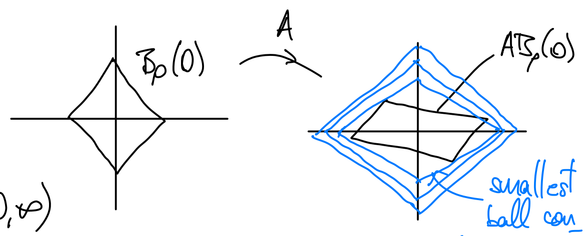
$$\underbrace{\|Au\|_2^2}_{\geq 0} = \lambda \underbrace{\|u^*u\|_2^2}_{\geq 0} \quad \text{i.e., } \lambda \geq 0.$$

Now,

$\operatorname{tr} A^*A = \sum$ eigenvals.

$\Rightarrow \underbrace{\text{largest eigenval}}_{\rho(A^*A)} \leq \operatorname{tr}(A^*A) \leq n \cdot \underbrace{\text{largest eigenval}}_{\rho(A^*A)}$. (Take sqrt) ■

<u>OPERATOR NORMS</u> (OR SUBORDINATE MATRIX NORMS)



<u>Prop</u>: For every norm $\|\cdot\|$ on $\mathbb{C}^n$ (or $\mathbb{R}^n$) and for every $A \in M_n(\mathbb{C})$ (or $M_n(\mathbb{R})$), $\exists C_A \in [0, \infty)$ such that $\|Au\| \leq C_A \|u\|$ $\forall u \in \mathbb{C}^n$.

<u>Pf</u>: Choose basis for $\mathbb{C}^n$ so that $u = \sum_{i=1}^n u_i e_i$ and $Au = \sum_{i=1}^n u_i Ae_i$. Then

$\tilde{C}_A = \max_i \|Ae_i\|$

$$\|Au\| = \left\| \sum_{i=1}^n u_i Ae_i \right\| \leq \sum_{i=1}^n |u_i| \, \|Ae_i\| \leq \tilde{C}_A \|u\|_1$$

($\exists C_1$ s.t. $\|\cdot\|_1 \leq C_1 \|\cdot\|$) equivalence of norms in finite dims. $\rightarrow \leq \tilde{C}_A C_1 \|u\|$ ■

$$\iff \sup_{u\neq 0} \frac{\|Au\|}{\|u\|} \text{ exists and is finite}$$

$$\|$$

$$\iff \sup_{\|u\|=1} \|Au\| =: \|A\|_{op} .$$

Def: (Operator Norm)

$$\|A\|_{op} := \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\|=1} \|Ax\| = \inf_x \{C : \|Ax\| \leq C\|x\|\}$$

Cl: $\|AB\|_{op} \leq \|A\|_{op}\|B\|_{op}$

Pf: $\|ABx\| \leq \|A\|_{op} \|Bx\| \leq \|A\|_{op} \|B\|_{op} \|x\|$

$\implies$ For $x\neq 0$, $\dfrac{\|ABx\|}{\|x\|} \leq \|A\|_{op}\|B\|_{op}$ i.e., $\sup_{x\neq 0} \dfrac{\|ABx\|}{\|x\|} \leq \|A\|_{op}\|B\|_{op}$

$\square$

Claim ℓ^1: $\|A\|_{op}$ for $\|\cdot\|_1$ is

$$\|A\|_{op} = \max_j \{\|A_1\|_1, \dots, \|A_n\|_1\} \quad A = \begin{pmatrix} | & & | \\ A_1 & \cdots & A_n \\ | & & | \end{pmatrix}$$

$$= \max_j \left\{ \sum_i |a_{ij}| \right\}$$

Pf: $\|Ax\|_1 = \sum_{i=1}^n |(Ax)_i| = \sum_{i=1}^n \left| \sum_{j=1}^n a_{ij} x_j \right| \leq \sum_{i=1}^n \sum_{j=1}^n |a_{ij}||x_j|$

$$\le \sum_j \sum_i |a_{ij}||x_j| = \sum_j |x_j| \underbrace{\sum_i |a_{ij}|}_{\|A_j\|_1}$$

$$\le \max_j \|A_j\|_1 \sum_j |x_j| = \max_j \|A_j\|_1 \|x\|_1 .$$

Thus: $\|A\|_{op} \le \max_j \|A_j\|_1$.

On the other hand (want $\ge$ to get $=$). Note that for

$$\hat{x}_j = \begin{cases} 1, & j = j_0 \\ 0, & \text{else} \end{cases}, \quad \|\hat{x}\|_1 = 1 \text{ and } A\hat{x} = A_{j_0}$$

so $\|A\hat{x}\|_1 = \|A_{j_0}\| = \max_j \dots \Rightarrow \|A\|_{op} \ge \max.$ $\blacksquare$

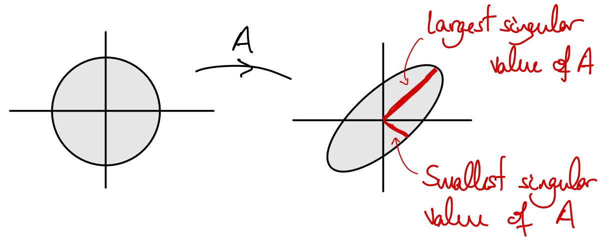
Claim $^{\ell^\infty}$: $\|A\|_{op}$ for $\|\cdot\|_\infty$ is

$$\|A\|_{op} = \max\{\ell^1 \text{ norms of rows of } A\}$$
$$= \max_i \sum_{j=1}^{n} |a_{ij}| .$$

Pf: $\|A\|_{op} \le \max \leftarrow \Delta$ ineq.

$\|A\|_{op} \ge \max \leftarrow$ test fct that sees sign. $\square$

Claim: $\|A\|_{op}$ for $\|\cdot\|_2$ (ℓ^2) is



$$\|A\|_{op} = \text{largest singular value}$$
$$= \sqrt{\rho(A^*A)} = \sqrt{\rho(AA^*)}.$$

A^*A symm. nonneg. definite
$\Rightarrow \operatorname{eig}(A^*A) \subset [0,\infty)$
$\rightsquigarrow \sigma_n \le \sigma_{n-1} \le \cdots \le \sigma_2 \le \sigma_1$
$SVD(A) = \{\sqrt{\sigma_1}, \sqrt{\sigma_2}, \ldots, \sqrt{\sigma_n}\}.$

$$\|A^*\|_{op} = \|A\|_{op}$$

$$\|UAV\|_{op} = \|A\|_{op}$$

If A is normal (i.e., $A^*A = AA^*$) then $\|A\|_{op} = \rho(A)$.

EXAMPLE: For ODE $\begin{cases} \dot{x} = Ax \\ x(0) = x_0 \end{cases} \rightsquigarrow x(t) = e^{tA}x_0$

$$\boxed{\text{Prop: Let } A \in M_n(\mathbb{C}) \text{ (or } M_n(\mathbb{R})\text{), then } \sum_{k=0}^{\infty} \frac{A^k}{k!} \text{ converges absolutely for any induced norm on } M_n(\mathbb{C}).}$$

Pf: $S_n := \sum_{k=0}^{n} \frac{A^k}{k!}$. $\|S_n\|_{op} = \left\| \sum_{k=0}^{n} \frac{A^k}{k!} \right\|_{op} \le \sum_{k=0}^{n} \frac{1}{k!} \|A^k\|_{op}$

$$\overset{\text{submultiplic.}}{\le} \sum_{k=1}^{n} \frac{1}{k!} \|A\|_{op}^k \le e^{\|A\|_{op}}$$

$\Rightarrow$ Series converges absolutely. Define this limit as $e^A := \sum_{k=0}^{\infty} \frac{A^k}{k!}$.

LECTURE 3

ITERATIVE SOLVERS OF LINEAR SYSTEMS

$Au = b$, A square & invertible.

e.g.: $(A \mid b)$ then Gaussian elimination

If A is $n \times n$ then Gaussian elimination is $O(n^3)$

e.g.: LU decomposition

$$\underbrace{L_n \cdots L_2 L_1}_{\text{lower triangular from Gaussian elimination}} A = U$$

upper triangular

$A = LU$ ← upper triang.

$(\text{lower } \Delta)^{-1}$

$A\vec{u} = \vec{b}$

$LU\vec{u} = \vec{b}$

Step 1: solve $L\vec{y} = \vec{b}$

Step 2: solve $U\vec{u} = \vec{y}$

e.g.: QR-decomposition: $A = QR$

orthogonal; upper triang.

Step 1: solve $Q\vec{y} = \vec{b}$

Step 2: solve $R\vec{u} = \vec{y}$

ITERATIVE METHOD: Seek $\{u_k\}_{k=0}^{\infty}$ so that $u_k \longrightarrow \tilde{u}$ and $\tilde{u}$ solves $A\tilde{u} = b$.

JACOBI METHOD: $A \in M_3(\mathbb{R})$, $A\tilde{u} = b$

$$\begin{cases} a_{11}\tilde{u}_1 + a_{12}\tilde{u}_2 + a_{13}\tilde{u}_3 = b_1 \\ a_{21}\tilde{u}_1 + a_{22}\tilde{u}_2 + a_{23}\tilde{u}_3 = b_2 \\ a_{31}\tilde{u}_1 + a_{32}\tilde{u}_2 + a_{33}\tilde{u}_3 = b_3 \end{cases}$$

Given $\vec{u}^k$,

$$\begin{aligned} a_{11} u_1^{k} &= b_1 - a_{12} u_2^k - a_{13} u_3^k \\ a_{22} u_2^{k-1} &= b_2 - a_{21} u_1^k - a_{23} u_3^k \\ a_{33} u_3^{k+1} &= b_3 - a_{31} u_1^k - a_{32} u_2^k \end{aligned} \quad , \ \operatorname{diag} A \neq 0$$

$$A = D - E - F$$

diag(A)

zero on and above diag

zero on and below diagonal

$$(A - E - F)\tilde{u} = b$$

$$D\vec{u}^{k+1} = (E+F)\vec{u}^k + \vec{b}$$

pen and paper only ?!

$$\vec{u}^{k+1} = D^{-1}(E+F)\vec{u}^k + D^{-1}\vec{b}$$

IN COMPUTER: AVOID COMPUTING INVERSES ?!

GAUSS - SEIDEL METHOD:

$$\begin{aligned} a_{11} u_1^{k+1} &= b_1 - a_{12} u_2^k - a_{13} u_3^k \\ a_{22} u_2^{k+1} &= b_2 - a_{21} u_1^{k+1} - a_{23} u_3^k \\ a_{33} u_3^{k+1} &= b_3 - a_{31} u_1^{k+1} - a_{32} u_2^{k+1} \end{aligned}$$

Organize:

$$\begin{pmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \vec{u}^{k+1} = \vec{b} - \begin{pmatrix} 0 & a_{12} & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 0 & 0 \end{pmatrix} \vec{u}^k$$

$$(D-E)\vec{u}^{k+1} = F\vec{u}^{k} + \vec{b}$$

Pen and paper $\left[\ \vec{u}^{k+1} = (D-E)^{-1}F\vec{u}^{k} + (D-E)^{-1}\vec{b}\right.$.

CONVERGENCE

$\vec{u}^{k+1} = B\vec{u}^{k} + \vec{c}$ when does $\{u^k\}$ converge?

If it does converge, then it should converge to $\tilde{u} = B\tilde{u} + \vec{c}$.

Assume $(Id - B)$ is invertible $\Rightarrow$ $\tilde{u}$ is unique.

Let $\vec{e}^{k} := \vec{u}^{k} - \tilde{u}$

$$\begin{aligned} e^{k+1} &= \vec{u}^{k+1} - \tilde{u} \\ &= (B\vec{u}^{k} + \vec{c}) - \tilde{u} \\ &= B\vec{u}^{k} - (\tilde{u} - \vec{c}) \\ &= B\vec{u}^{k} - B\tilde{u} \\ &= B\vec{e}^{k} \quad \xrightarrow{\text{by induction}} \quad \boxed{\vec{e}^{k+1} = B^{k+1}\vec{e}^{0}} \end{aligned}$$

Thm: $B \in M_n(\mathbb{C})$. TFAE

(1) $B^k \to 0$

(2) $\forall v \in \mathbb{C}^n$, $B^k v \longrightarrow 0$

(3) $\rho(B) < 1$

(4) $\exists$ subordinate matrix norm s.t. $\|B\|_{op} < 1$

Pf: (1) $\Rightarrow$ (2) Choose $\|B^k v\| \le \|B^k\|_{op} \|\vec{v}\|$

if $B^k \to 0 \Rightarrow \|B^k\|_{op} \to 0 \Rightarrow \|B^k \vec{v}\| \to 0$

$\Rightarrow B^k \vec{v} \to 0$.

(2) $\Rightarrow$ (3) Assume not. If $\exists \lambda$ s.t. $|\lambda| \ge 1$ use the eigenvec. to violate 1.

LECTURE 4

<u>GOAL</u>: Solve $Au = b$, A $n \times n$ invertible.

Iterated problem: $(I-B)u = c$, $I-B$ invertible

where $\tilde{u}$ solves both $\tilde{u} = B\tilde{u} + c$

$$u^{k+1} = Bu^k + c \quad \text{iteration starts w/ guess } u^0.$$

Let $e^k = u^k - \tilde{u}$. $\quad u^k \to \tilde{u}$ as $k \to \infty \Leftrightarrow e^k \to 0$ as $k \to \infty$

Induction: $e^k = B^k e^0$.

<u>Q</u>: For what types of B does $e^k = B^k e^0 \to 0$? How fast?

Thm: Let $B \in M_n(\mathbb{C})$. TFAE

(1) $B^k \longrightarrow 0$ as $k \to \infty$

(2) $B^k v \longrightarrow 0 \quad \forall v \in \mathbb{C}^n$

(3) $\rho(B) < 1$

(4) $\exists$ subordinate norm s.t. $\|B\|_{op} < 1$.

Pf: $(3) \Rightarrow (4)$ choose $\varepsilon > 0$ s.t. $\rho(B) + \varepsilon < 1$. Then $\exists$ a subordinate matrix norm s.t. $\|B\|_{op} \leq \rho(B) + \varepsilon$.

$(4) \Rightarrow (1)$ $\|B^k\|_{op} \leq \|B\|_{op}^k \leq (\underbrace{\rho(B) + \varepsilon}_{<1})^k$

$\Rightarrow \|B^k\|_{op} \longrightarrow 0$

$\Longrightarrow B^k \longrightarrow 0$

$\blacksquare$

Claim: Given $A \in M_n(\mathbb{C})$ and $\varepsilon > 0$, $\exists$ a subordinate norm s.t. $\|A\|_{op} \leq \rho(A) + \varepsilon$.

Pf: Write A in Jordan canonical form $A \sim \begin{pmatrix} * & 1 & & & 0 \\ & * & 0 & & \\ & & * & 1 & \\ & & & * & 1 \\ 0 & & & & * & 0 \cdots \\ & & & & & * \end{pmatrix}$ since $A \in M_n(\mathbb{C})$.

Jordan's Canonical form thm

Now, $\exists$ invertible matrix U and upper triangular matrix T w/ $\operatorname{diag}(T) = \operatorname{eig}(A)$ and $A = UTU^{-1}$. i.e., $U^{-1}AU = T$

$$T = \begin{pmatrix} \lambda_1 & t_{12} & t_{13} & \cdots \\ & \lambda_2 & & * \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}, \quad T_\delta = \begin{pmatrix} \lambda_1 & \delta t_{12} & \delta^2 t_{13} & \cdots & \delta^{n-1} t_{1n} \\ & \lambda_2 & \delta t_{23} & \cdots & \delta^{n-2} t_{2n} \\ & & \ddots & & * \\ 0 & & & & \lambda_n \end{pmatrix}$$

$$\|T_\delta\|_{op,\infty} = \max_i \{\|\text{row}\|_i\} \le \rho(T) + \varepsilon \overset{\substack{T \text{ is similar to } A \\ \downarrow}}{=} \rho(A) + \varepsilon$$

Let $D_\delta := \operatorname{diag}(1, \delta, \delta^2, \text{—}, \delta^{n-1})$

$$\underbrace{D_\delta^{-1}}_{\substack{\text{Multiplies} \\ i\text{th row by } \delta^{1-i}}} \underbrace{T D_\delta}_{\text{Multiplies } j \text{ column by } \delta^{j-1}}$$

i.e., $(D_\delta^{-1} T D_\delta)_{ij} = \delta^{1-i} t_{ij} \delta^{j-1} = t_{ij} \delta^{j-i}$

Thus $D_\delta^{-1} T D_\delta = T_\delta$.

So
$$U^{-1} A U = T$$
$$D_\delta^{-1} U^{-1} A U D_\delta = D_\delta^{-1} T D_\delta$$
$$(U D_\delta)^{-1} A (U D_\delta) = D_\delta^{-1} T D_\delta$$

Define $\|\cdot\| : \mathbb{C}^n \to [0,\infty)$ by $\|\vec{v}\| := \|(U D_\delta)^{-1} \vec{v}\|_\infty$.

Norm on $\mathbb{C}^n$ and induces $\|B\|_{op} = \|(U D_\delta)^{-1} B (U D_\delta)\|_{op,\infty}$

on the set $M_n(\mathbb{C})$. Then

$\|A\|_{op} = \|D_\delta^{-1} T D_\delta\|_{op,\infty} \le \rho(A) + \varepsilon$ if δ is small enough b/c T is upper triangular.

Check: $|\cdot|$ is a norm on $\mathbb{C}^n$.

$\|\cdot\|_\infty$ is a norm

(i) $\|v\| = \|(UD_\delta)^{-1}v\|_\infty = 0 \iff (UD_\delta)^{-1}v = 0$

U invertible $\iff v = 0$

$\|\cdot\|_\infty$ is norm

(ii) $\|u+v\| = \|(\ )^{-1}(u+v)\|_\infty \leq \|(\ )^{-1}u\|_\infty + \|(\ )^{-1}v\|_\infty$.

(iii) $\|\mu v\| = \|(\ )^{-1}(\mu v)\|_\infty = |\mu|\,\|(\ )^{-1}v\|_\infty$.

Check: $\|B\|_{op} \overset{\text{df}}{=} \sup_{v\neq 0} \dfrac{\|Bv\|}{\|v\|} = \sup_{v\neq 0} \dfrac{\|(\ )^{-1}Bv\|_\infty}{\|(\ )^{-1}v\|_\infty}$

$(\ )^{-1}$ is invertible

$w := (\ )^{-1}v$

$= \sup_{w\neq 0} \dfrac{\|(\ )^{-1}B(\ )w\|_\infty}{\|w\|_\infty}$

$\overset{\text{def}}{=} \|(\ )^{-1}B(\ )\|_{op,\infty}$.

Prop: $B \in M_n(\mathbb{C})$ and $\|\cdot\|$ matrix norm. Then $$\lim_{k\to\infty} \left(\|B^k\|\right)^{1/k} = \rho(B)$$

Pf: Already know $\rho(B) \leq \|B\|$.

i.e., $\rho(B^k) \leq \|B^k\|$

i.e., $[\rho(B)]^k \leq \|B^k\| \Rightarrow \rho(B) \leq \|B^k\|^{1/k}$.

WTS: $\|B^k\|^{1/k} \leq \rho(B) + \varepsilon$ for $k \geq N_\varepsilon$, N_ε large enough.

Take $\varepsilon \to 0$ then done by Squeeze Thm.

Given $\varepsilon > 0$, take $B_\delta := \dfrac{B}{\rho(B)+\varepsilon}$.

$$\Rightarrow \rho(B_\delta) = \frac{\rho(B)}{\rho(B)+\varepsilon} < 1$$

$$\Rightarrow (B_\delta)^k \longrightarrow 0 \text{ as } k \to \infty \text{ by Thm above.}$$

i.e., $\|(B_\delta)^k\| \leq 1$ for $k \geq N_\varepsilon$

i.e., $\|B^k\| \leq (\rho(B)+\varepsilon)^k$. ∎

Rates of Convergence

Thm: $\|\cdot\|$ vector norm. Say $\tilde{u}$ solves $\tilde{u} = B\tilde{u} + c$ where $Id - B$ is invertible. Define $\{u^k\}$ iteratively by

$$u^{k+1} = Bu^k + c$$

Then

$$\lim_{k\to\infty} \left(\sup_{\|u^0 - \tilde{u}\| = 1} \|u^k - \tilde{u}\|^{1/k} \right) = \rho(B)$$

Obs:
$$\sup_{u^0 \neq \tilde{u}} \left(\frac{\|u^k - \tilde{u}\|}{\|u^0 - \tilde{u}\|} \right)^{1/k} \stackrel{\text{def}}{=} \sup_{u^0 \neq \tilde{u}} \left(\frac{\|e^k\|}{\|e^0\|} \right)^{1/k}$$

$$= \sup_{u^0 \neq \tilde{u}} \left(\frac{\|B^k e^0\|}{\|e^0\|} \right)^{1/k}$$

$$= \sup_{u^0 \neq \tilde{u}} \left\| B^k \frac{e^0}{\|e^0\|} \right\|^{1/k}$$

$$= \sup_{\|e^0\| = 1} \|B^k e^0\|^{1/k}$$

LECTURE 5

SINGULAR VALUE DECOMPOSITION

Seek $\vec{u}_1, \vec{u}_2$, $\vec{u}_1 \perp \vec{u}_2$, s.t. $A\vec{u}_1 = \sigma_1 \vec{v}_1$, $A\vec{u}_2 = \sigma_2 \vec{v}_2$, $\vec{v}_1 \perp \vec{v}_2$

$\Updownarrow$

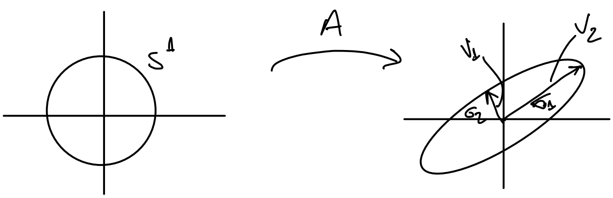
orthogonal diagonal

$$AU = V\Sigma \longleftrightarrow A = V\Sigma U^T$$

Σ – diagonal entries ≥ 0

Goal: Find orthogonal vectors that get mapped to orthogonal vectors.

Equivalently,



optimize $\|Ax\|_2^2$
$x \in S^1 \Leftrightarrow \|x\|_2^2 = 1$

Lagrange Multipliers

$$f(x) = \|Ax\|_2^2$$
$$g(x) = \|x\|_2^2 - 1$$

Seek λ, x_0 s.t.
$$\nabla f(x_0) = \lambda \nabla g(x_0)$$
$$g(x_0) = 0.$$

$$\langle Ax, Ax \rangle = \langle A^TAx, x \rangle \rightsquigarrow \nabla f(x) = 2A^TAx$$

i.e., seek λ, x_0 s.t. $2A^TAx_0 = \lambda \cdot 2x_0$
$$g(x_0) = 0$$

- Consider $f: E \to F$ linear transf. between $\mathbb{R}$-inner product spaces of $\dim E = \dim F = n$.
- Adjoint $f^*: F \to E$ $\quad \langle f(u), v\rangle = \langle u, f^*(v)\rangle$
  $\forall u \in E, v \in F$.

<u>Proposition</u>: Eigenvalues of $f^* \circ f: E \to E$ and $f \circ f^*: F \to F$ are ≥ 0

Rank-Nullity: $\dim \operatorname{im} f + \dim \ker f = \dim E$.

<u>Proposition</u>: $\ker(f) \oplus \operatorname{im}(f^*) = E$

<u>Proposition</u>:
$$\begin{aligned}\dim(\operatorname{im} f) &= \dim(\operatorname{im} f^*)\\ &= \dim(\operatorname{im}(f^* f))\\ &= \dim(\operatorname{im}(f f^*))\end{aligned}$$

<u>Proposition</u>: E, F finite dimensional Euclidean spaces, $f: E \to F$,

$\ker f = \ker f^* f$

$\ker f^* = \ker f f^*$

$\ker f = (\operatorname{im} f^*)^\perp$

$\ker f^* = (\operatorname{im} f)^\perp$

$\dim \operatorname{im}(f^* f) = \dim \operatorname{im} f^*$

$\operatorname{rank} f = \operatorname{rank} f^* = \operatorname{rank} f^* f = \operatorname{rank} f f^*.$

LECTURE 6

SVD FOR SQUARE MATRICES

Thm: $A \in M_n(\mathbb{R})$ then $\exists$ orthogonal matrices U and V and a diagonal matrix Σ such that

$$A = U \Sigma V^T,$$

$$\Sigma = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \end{pmatrix}, \quad \{\sigma_i\}_{i=1}^r \text{ singular values of } A.$$

$r := \operatorname{rank} A.$

$\sigma_{r+1}, \dots, \sigma_n = 0$

i.e., positive square roots of the nonzero eigenvalues of A^TA and AA^T.

$$U = \begin{pmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{pmatrix}, \quad \{u_i\}_{i=1}^n \text{ eigenvectors of } A^TA$$

$$V = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix}, \quad \{v_i\}_{i=1}^n \text{ eigenvectors of } AA^T.$$

Note: if A is symmetric then it's not automatic that

$A = UDU^T \not\Rightarrow D = \Sigma$ because A symm $\not\Rightarrow$ spec $A \geq 0$.

Note: SVD also works for complex matrices.

$$A = \begin{pmatrix} \frac{1}{2}+i & -1/2 \\ -1/2 & \frac{1}{2}+i \end{pmatrix} = U \Sigma V^* \qquad U = \begin{pmatrix} -1/2 - i/2 & -i/\sqrt{2} \\ 1/2 + i/2 & -i/\sqrt{2} \end{pmatrix}$$

$$\text{spec } A = \{i, 1+i\} \qquad \Sigma = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & 1 \end{pmatrix}$$

$$V = \begin{pmatrix} -1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}$$

Polar Form for Matrices

$A \in M_n(\mathbb{R}) \Rightarrow \exists R$ orthogonal

$\exists S$ symmetric positive semi-definite

s.t. $A = RS$. $\longleftrightarrow$ Analogous to $\mathbb{C} \ni z = \underbrace{e^{i\theta}}_{\updownarrow \, R} \underbrace{\rho}_{\updownarrow \, S}$

- Given an SVD, construct polar form

$$A = V \Sigma U^T = \underbrace{V U^T}_{\text{orthogonal}} \underbrace{U \Sigma U^T}_{\text{symm., pos. semi-def.}}$$

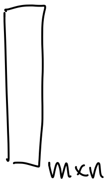
- Given a polar form construct SVD:

$$A = RS = RU\Sigma U^T = V\Sigma U^T$$

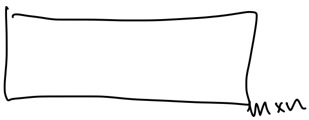
symm. pos. semi-def.

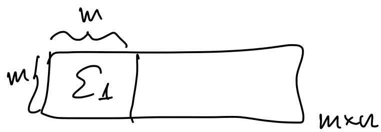
$$V := RU$$

NOTE: If A is not square then

$A =$ 

$\Rightarrow \Sigma =$ 

$A =$ 

$\Rightarrow \Sigma =$ 

NOTE: $A = V\Sigma U^T = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix} \Sigma U^T$

$$= \begin{pmatrix} | & & | & | & & | \\ \sigma_1 v_1 & \cdots & \sigma_r v_r & 0 & \cdots & 0 \\ | & & | & | & & | \end{pmatrix} U^T$$

$$= \sum_{i=1}^{r} \sigma_i v_i u_i^T$$ = sum of rank-1 matrices!

Least Squares & Pseudo-Inverses

Goal: Solve $Ax = b$ with A being $m \times n$, $b \in \mathbb{R}^m$, $x \in \mathbb{R}^n$

$\exists$ solution $x \in \mathbb{R}^n \iff b \in \operatorname{im} A$

But if $\operatorname{im} A \subsetneq \mathbb{R}^m$ what to do if $b \notin \operatorname{im} A$?

Seek: $x \in \mathbb{R}^n$ that minimizes the distance

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|^2 .$$

Write $b = \underbrace{b_0}_{\in \operatorname{im} A} + \underbrace{b_1}_{\in (\operatorname{im} A)^\perp}$ and solve $Ax_0 = b_0$

Seek x_0 s.t. $2A^T(Ax_0 - b) = 0$

Residual $Ax_0 - b$ is $\perp$ to $\operatorname{im} A$

$$A^TAx_0 = A^Tb$$

$$A^TAx = A^Tb$$

$\operatorname{rank}(A^TA) = \operatorname{rank} A \Rightarrow$ if columns of A are lin indep., then $\exists!$ solution

- If $\operatorname{rank} A < n$ then A^TA is not invertible and there are ∞-many solutions to $A^TAx = A^Tb$.

Goal 2.0: Seek solutions of shortest "length".

Thm: $\inf_{x \in \mathbb{R}^n} \|Ax-b\|_2^2 = \inf_{y \in \operatorname{im} A} \|y-b\|_2^2$

Def: Say A is $m\times n$ and $\operatorname{rank} A = r$.

$A = VDU^T$ where $D = \begin{pmatrix} \Lambda_{r\times r} & 0_{r\times n-r} \\ 0_{m-r\times r} & 0_{m-r\times n-r} \end{pmatrix}_{m\times n}$

$\Lambda := \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r)$ is invertible.

$D^+ := \begin{pmatrix} \Lambda^{-1}_{r\times r} & 0_{r\times m-r} \\ 0_{m-r\times r} & 0_{n-r\times m-r} \end{pmatrix}_{n\times m}$

$$\begin{aligned} AA^+ &= VDU^TUD^+V^T \\ &= VDD^+V^T \\ &= \left(\begin{array}{c|c} Id_{r\times r} & 0 \\ \hline 0 & 0 \end{array}\right)_{m\times m} \end{aligned}$$

$A^+ := UD^+V^T$ is the <u>pseudo</u>-<u>inverse</u> of A.

Thm: Least squares solution of smallest norm of the linear system $Ax=b$ where A is $m\times n$ is given by

$$x^+ := A^+b = UD^+V^Tb$$

Pf: Case 1: $A = D$. Then $V = U = \mathrm{Id}$.

$$A = \left(\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_r & \\ \hline & 0 & & 0 \end{array}\right) \qquad \operatorname{im} A = \operatorname{span}(e_1, \dots, e_r)$$

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \|Ax - b\|_2^2 = \min_{x \in \mathbb{R}^n} \|Ax - \operatorname{proj}_{\operatorname{im} A} b\|^2$$

$$= \min_{x \in \mathbb{R}^n} \left\| Dx - \begin{pmatrix} b_1 \\ \vdots \\ b_r \\ 0 \end{pmatrix} \right\|^2$$

$$= \min_{x \in \mathbb{R}^n} \left\| \begin{pmatrix} \sigma_1 x_1 \\ \vdots \\ \sigma_r x_r \\ 0 \end{pmatrix} - \begin{pmatrix} b_1 \\ \vdots \\ b_r \\ 0 \end{pmatrix} \right\|^2$$

$$(x^+)_i = \frac{b_i}{\sigma_i}, \quad 1 \leq i \leq r.$$

$$\text{minimize } \|x\|_2^2 \rightsquigarrow x = \begin{pmatrix} \sigma_1 x_1 \\ \vdots \\ \sigma_r x_r \\ x_{r+1} \\ \vdots \\ x_n \end{pmatrix} \quad \text{set } x_i \overset{!}{=} 0 \text{ for } r+1 \leq i \leq n$$

$$\Longrightarrow x^+ = D^+ b = A^+ b.$$

Case 2: $A \neq D$. Then SVD $\rightsquigarrow A = VDU^T$

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_2 = \min_{x \in \mathbb{R}^n} \|VDU^T x - b\|_2$$

$$= \min_{x \in \mathbb{R}^n} \|VDU^T x - VV^T b\|_2$$

$$= \min_{x \in \mathbb{R}^n} \|V(DU^T x - V^T b)\|_2$$

orthogonal matrices don't change the ℓ^2-norm

$$= \min_{x \in \mathbb{R}^n} \| D U^T x - V^T b \|_2$$

$y := U^T x$

$$= \min_{y \in \mathbb{R}^n} \| D y - V^T b \|_2$$

By Case 1, shortest $y \in \mathbb{R}^n$ that minimizes $\| Dy - V^T b \|_2$ is $y^+ := D^+ (V^T b)$. $y = U^T x \rightsquigarrow x = Uy$.

But $\|x\|_2 = \|y\|_2$ b/c U is orthogonal. Thus:

$$x^+ = U y^+ = U D^+ V^T b = A^+ b$$

is the solution we want. $\blacksquare$

<u>Obs</u>: Easy to find A^+ using Python.

Properties of Pseudo-Inverse

Proposition: If A has full rank, then

$$m \geq n \Rightarrow A^+ = (A^T A)^{-1} A^T \quad \text{and} \quad A^+ A = Id$$

$$m \leq n \Rightarrow A^+ = A^T (A^T A)^{-1} \quad \text{and} \quad A A^+ = Id.$$

- NOTE: In general, $A^+ A$ and $A A^+$ are symmetric

$$A = V \Sigma U^T \qquad A^+ = U \Sigma^+ V^T$$

$$A A^+ = V \Sigma \Sigma^+ U^T \qquad \Sigma \Sigma^+ = \begin{pmatrix} Id_{r \times r} & 0 \\ 0 & 0_{m-r \times m-r} \end{pmatrix}$$

$$A^+ A = U \Sigma^+ \Sigma U^T \qquad \Sigma^+ \Sigma = \begin{pmatrix} Id_{r \times r} & 0 \\ 0 & 0_{n-r \times n-r} \end{pmatrix}.$$

$\Rightarrow A A^+$ is $m \times m$ and

$$A A^+ = V \begin{pmatrix} Id_{r \times r} & 0 \\ 0 & 0 \end{pmatrix} V^T$$

$$\Rightarrow A A^+ A = V \begin{pmatrix} Id_{r \times r} & 0 \\ 0 & 0 \end{pmatrix} V^T V \Sigma U^T$$

$$= V \begin{pmatrix} Id_{r\times r} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \Lambda_{r\times r} & 0 \\ 0 & 0 \end{pmatrix} U^T$$

$$= V \begin{pmatrix} \Lambda_{r\times r} & 0 \\ 0 & 0 \end{pmatrix} U^T$$

$$= A.$$

P projection op. on $\mathbb{R}^m$

$\Rightarrow \mathbb{R}^m = (\operatorname{im} P) \oplus (\operatorname{im} P)^\perp$

$x \in \mathbb{R}^m$; $x = \underbrace{Px}_{\operatorname{im} P} - \underbrace{(x - Px)}_{\ker P}$

So: $\quad AA^+A = A \implies AA^+AA^+ = AA^+$

i.e., $(AA^+)^2 = AA^+$

i.e., AA^+ is an <u>orthogonal</u> <u>projection</u>

Pf: claim $\operatorname{im} AA^+ = \operatorname{im} A$ and $\ker A^+A = \ker A$

$$\boxed{\begin{array}{l}\underline{\text{Thm}}\text{: } AA^+ \text{ is the orthogonal projection onto } \operatorname{im} A; \\ \text{and } A^+A \text{ is the orthogonal projection onto } \operatorname{im} A^T = (\ker A)^\perp.\end{array}}$$

<u>Thm</u>: The pseudo-inverse A^+ is the unique $n\times m$ matrix that satisfy:

$$\left.\begin{aligned} AA^+A &= A \\ A^+AA^+ &= A^+ \\ (AA^+)^T &= AA^+ \\ (A^+A)^T &= A^+A \end{aligned}\right\}$$

Moore–Penrose Inverse

LECTURE 7

- Suppose A is $m \times n$, $r := \operatorname{rank} A$

$$A = \underbrace{V}_{m \times m \text{ orthogonal}} \underbrace{D}_{\begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix}} \underbrace{U^T}_{n \times n \text{ orthogonal}}$$

Λ is $r \times r$ diagonal

- When A^TA is <u>not</u> full rank, then the LEAST SQUARE PROBLEM $A^TAx = A^Tb$ does <u>not</u> have a unique solution.

- <u>Seek</u>: solution of $Ax = b$ that $\begin{cases} \text{minimizes } \|Ax-b\|_2^2 \\ \text{minimizes } \|x\|_2^2 \end{cases}$

Pseudo-Inverse $\underbrace{A^+}_{n \times m} = UD^+V^T$, $\quad D^+ := \begin{pmatrix} \Lambda^{-1} & 0 \\ 0 & 0 \end{pmatrix}$

LECTURE 8

ECKART - YOUNG THEOREM

Goal: Show that SVD is the best rank-r approximation to an $m \times n$ matrix A with rank $A = r$.

Obs: Need to choose a norm

$$A = V \Sigma U^T$$

$$= \begin{pmatrix} | & & | \\ v_1 & \cdots & v_m \\ | & & | \end{pmatrix} \begin{pmatrix} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \sigma_r & \\ 0 & & & 0 \end{pmatrix} \begin{pmatrix} — & u_1^T & — \\ & \vdots & \\ — & u_n^T & — \end{pmatrix}$$

$$= \sum_{i=1}^{n} \underset{\substack{\vee \\ 0}}{\sigma_i(A)} \underbrace{v_i u_i^T}_{\text{rank} = 1}$$

Set $A_k := \sum_{i=1}^{k} \sigma_i(A) v_i u_i^T \rightsquigarrow \text{rank} A_k = k$.

Thm: (Eckart - Young) Set

$$\mathcal{B} := \{\text{set of all rank-}k \ m \times n \text{ matrices w/ } \mathbb{R}\text{-coeffs}\}$$

Then
$$\|A - A_k\| \leq \|A - B\| \quad \forall B \in \mathcal{B}.$$

- ℓ^2-norm and Frobenius norm

Rmk: ℓ^2-norm (induced) $\|A - A_k\|_{op,2} \leq \|A - B\|_{op,2} \quad \forall B \in \mathcal{B}$

Frobenius norm (not induced) $\|A - A_k\|_F \leq \|A - B\|_F \quad \forall B \in \mathcal{B}$

$$\|C\|_F := \sqrt{\sum_i \sum_j |c_{ij}|^2}$$

Pf: (ℓ^2-norm) Cl: $\|A - A_k\|_2 = \min_{B \in \mathcal{B}} \|A - B\|_2$

$$A - A_k = V \begin{pmatrix} \boxed{\Sigma_r} & 0 \\ 0 & 0 \end{pmatrix} U^T - V \begin{pmatrix} \boxed{\Sigma_k} & 0 \\ 0 & 0 \end{pmatrix} U^T$$

$$= V \begin{pmatrix} 0 & & 0 \\ & \boxed{\begin{matrix} \sigma_{k+1} & & \\ & \ddots & \\ & & \sigma_r \end{matrix}} & \\ 0 & & 0 \end{pmatrix} U^T .$$

$$\|A - A_k\|_2 = \left\| V \begin{pmatrix} 0 & & 0 \\ & \boxed{\begin{matrix} \sigma_{k+1} & & \\ & \ddots & \\ & & \sigma_r \end{matrix}} & \\ 0 & & 0 \end{pmatrix} U^T \right\|_2$$

$|\cdot\|_{op,2}$ is unitarily invariant

$$= \left\| \begin{pmatrix} 0 & & 0 \\ & \boxed{\begin{matrix} \sigma_{k+1} & & \\ & \ddots & \\ & & \sigma_r \end{matrix}} & \\ 0 & & 0 \end{pmatrix} \right\|_2$$

$$= \max_{k+1 \leq i \leq r} |\sigma_i| = \sigma_{k+1}$$

<u>Want</u>: $\|A-B\|_2 \geq \sigma_{k+1}$. Suffices to find $w \in \mathbb{R}^n$, $\|w\|=1$, s.t. $\|(A-B)w\|_2 \geq \sigma_{k+1}$

$$\|A-B\|_2 \overset{\text{def}}{=} \sup_{\|w\|=1} \|(A-B)w\|_2$$

$$\Longrightarrow \|A-B\|_2 \geq \sigma_{k+1} = \|A-A_k\|_2$$

$$\Longrightarrow \inf_{B \in \mathcal{B}} \|A-B\|_2 \geq \|A-A_k\|_2$$

$\Longrightarrow A_k \in \mathcal{B}$ so we can go from "inf" to "min".

<u>Note</u>: $A = V\Sigma U^T$, $w \in \mathbb{R}^n$ gives

$$w = \sum_{i=1}^n w_i U_i$$

$$U^T w = \begin{pmatrix} - u_1^T - \\ \vdots \\ - u_n^T - \end{pmatrix} w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$$

$$\Sigma U^T w = \begin{pmatrix} \sigma_1 w_1 \\ \vdots \\ \sigma_r w_r \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\Longrightarrow \|Aw\|_2 = \left\| V \begin{pmatrix} \sigma_1 w_1 \\ \vdots \\ \sigma_r w_r \\ 0 \end{pmatrix} \right\|_2 = \left\| \begin{pmatrix} \sigma_1 w_1 \\ \vdots \\ \sigma_r w_r \\ 0 \end{pmatrix} \right\|_2 = \sqrt{\sum_{i=1}^r (\sigma_i w_i)^2}$$

Take $w \in \operatorname{span}(u_1, -, u_{k+1})$

$$= \sqrt{\sum_{i=1}^{k+1} \sigma_i^2 w_i^2} \leq \sigma_{k+1} \sqrt{\sum_{i=1}^{r} w_i^2} = \sigma_{k+1} \cdot 1$$

$\dim \operatorname{span}(u_1, -, u_{k+1}) = k+1$

$\dim \ker B = n-k$

<u>Want</u>: $w \in \ker B$

$k+1+n-k > n$

$\Rightarrow \operatorname{span}(u_1, -, u_k) \cap \ker B \neq \{0\}$

$\Rightarrow$ Take w in that intersection and scale it so it has $\|w\|_2 = 1$.

Then $\|A-B\|_2 \geq \|(A-B)w\|_2$

$$= \|Aw\|_2 \geq \sigma_{k+1} = \|A - A_k\|_2 \quad \forall B \in \mathcal{B}.$$

(Frobenius norm) <u>Cl</u>: $\|A - A_k\|_F = \min_{B \in \mathcal{B}} \|A - B\|_F$

<u>Note</u>: $\|C\|_F = \sqrt{tr(C^*C)} = \sqrt{tr(CC^*)} = \sqrt{\sum \text{squares of sing. values of } C}$

$$\|A - A_k\|_F = \left\| V(\Sigma_r - \Sigma_k) U^T \right\|_F$$

$$= \sqrt{tr\left(\left(V(\Sigma_r - \Sigma_k) U^T\right)\left(V(\Sigma_r - \Sigma_k) U^T\right)^T \right.}$$

$$= \sqrt{tr\left(V(\Sigma_r - \Sigma_k)(\Sigma_r - \Sigma_k)^T V^T\right)}$$

cyclic tr.

$$= \sqrt{tr\left(V^T V(\Sigma_r - \Sigma_k)(\Sigma_r - \Sigma_k)^T\right)}$$

$$= \sqrt{tr\left((\Sigma_r - \Sigma_k)(\Sigma_r - \Sigma_k)^T\right)}$$

$$= \sqrt{\sum_{i=k+1}^{r} \sigma_i(A)^2}$$

<u>Note</u>: By the same logic, if $A-B$ is rank-$\tilde{r}$, then

$$\|A-B\|_F = \sqrt{\sum_{i=1}^{\tilde{r}} \sigma_i(A-B)^2}$$

<u>Lemma</u>: if A, B are $m \times n$, $m \geq n$, and rank $B = k$ then

$\sigma_{i+k}(A) \leq \sigma_i(A-B)$ for $1 \leq i \leq n-k$.

$$\Downarrow$$

$$\|A-B\|_F = \sqrt{\sum_{i=1}^{\tilde{r}} \sigma_i(A-B)^2} \geq \sqrt{\sum_{i=1}^{\min\{\tilde{r}, n-k\}} \sigma_i(A-B)^2}$$

$$\geq \sqrt{\sum_{i=1}^{\min\{\tilde{r}, n-k\}} \sigma_{i+k}(A)^2}$$

$$= \sqrt{\sum_{i=k+1}^{\min\{\tilde{r}+k, n\}} \sigma_i(A)^2}$$

rank$(C+D) \leq$ rank C + rank D

$$\geq \sqrt{\sum_{i=k+1}^{r} \sigma_i(A)^2} = \|A - A_k\|_F .$$

$\blacksquare$

RAYLEIGH - RIESZ

- Suppose A is (square) symmetric and real $n \times n$.

$$A = UDU^T$$

diagonal w/ real eigenvals. of A

$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$.

<u>Def</u>: (Rayleigh Ratio) $R_A : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$

$$R_A(x) := \frac{x^T A x}{x^T x}.$$

NOTE: Fix an eigenvec. v w/ eigenval λ

$$R_A(v) = \frac{v^T A v}{v^T v} = \frac{\lambda v^T v}{v^T v} = \lambda.$$

Lemma: For A symmetric, $\lambda_n = \max_{x \in \mathbb{R}^n \setminus \{0\}} R_A(x)$

Pf: x_0 eigenvec. corresp. λ_n $\rightsquigarrow$ $R_A(x_0) = \lambda_n$

$$\lambda_n = R_A(x_0) \le \max_{x \in \mathbb{R}^n \setminus \{0\}} R_A(x)$$

A symmetric $\Rightarrow$ $A = UDU^T$ and column(U) spans $\mathbb{R}^n$

$x \in \mathbb{R}^n \setminus \{0\}$ $\rightsquigarrow$ $x = \sum_{i=1}^n x_i u_i$

$$x^T x = \sum_{i=1}^n x_i^2$$

$$Ax = \sum_{i=1}^n x_i \lambda_i u_i$$

$$x^T A x = \sum_{i=1}^n x_i^2 \lambda_i$$

$$R_A(x) = \frac{\sum_{i=1}^n \lambda_i x_i^2}{\sum_{i=1}^n x_i^2} \le \frac{\sum_{i=1}^n \lambda_n x_i^2}{\sum_{i=1}^n x_i^2} = \lambda_n.$$

$$\lambda_n \le \sup_{x \in \mathbb{R}^n \setminus \{0\}} R_A \le \lambda_n \quad \rightsquigarrow \quad \lambda_n = \sup R_A(x)$$

used eigenvec as test vector

$$\lambda_n = \max R_A(x)$$

$$\lambda_1 = \min_{\mathbb{R}^n \setminus \{0\}} R_A(x)$$

$$V_k := \operatorname{span}\{u_1, \dots, u_k\}$$

$$\max_{x \in V_k \setminus \{0\}} R_A(x) = \lambda_k$$

$$\min_{x \in V_{k-1}^{\perp} \setminus \{0\}} R_A(x) = \lambda_k = \min_{x \in \operatorname{span}\{u_k, \dots, u_n\} \setminus \{0\}} R_A(x)$$

$\blacksquare$

EIGENVALUE INTERLACING

Def: $A \in M_n(\mathbb{R})$, $B \in M_m(\mathbb{R})$, $m \le n$, then the eigenvals of B interlace the eigenvalues of A iff

$$\lambda_i(A) \le \lambda_i(B) \le \lambda_{i+(n-m)}(A) \quad \forall i \in \{1, \dots, m\}.$$

Thm: If $A \in M_n(\mathbb{R})$ is symmetric and R is $n \times m$ real orthogonal (i.e., $R^T R = \mathrm{Id}$). Set $B := R^T A R$, then the eigenvals of B interlace the eigenvals of A.

Pf: A symm $\Rightarrow$ B symm

orthonormal set of eigenvecs $\{u_1, \dots, u_n\}$

$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$

orthonormal set of eigenvectors $\{v_1, \dots, v_m\}$

$$\lambda_j(B) = \sup_{\mathrm{span}\{v_1,\dots,v_j\}\setminus\{0\}} \frac{v^T B v}{v^T v} = \sup_{\mathrm{span}\{v_1,\dots,v_j\}\setminus\{0\}} \frac{v^T R^T A R v}{v^T v}$$

$$= \sup_{\mathrm{span}\{v_1,\dots,v_j\}\setminus\{0\}} \frac{(Rv)^T A (Rv)}{v^T v}$$

$R^T R = \mathrm{Id}$

$$= \sup_{\mathrm{span}\{v_1,\dots,v_j\}\setminus\{0\}} \frac{(Rv)^T A (Rv)}{(Rv)^T (Rv)}$$

LECTURE 9

Rayleigh Ratio: $A \in M_n(\mathbb{R})$ symmetric

$$R_A : \mathbb{R}^n \setminus \{0\} \longrightarrow \mathbb{R}, \quad R_A(x) := \frac{x^T A x}{x^T x}.$$

$$A \text{ symmetric} \Rightarrow A = UDU^T, \quad U = \begin{pmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{pmatrix}$$

$$\lambda_1(A) \le \cdots \le \lambda_n(A).$$

$$\lambda_i(A) = \max_{0 \ne x \in \operatorname{span}\{u_1, \dots, u_i\}} R_A(x)$$

$$= \min_{0 \ne x \in (\operatorname{span}\{u_1, \dots, u_{i-1}\})^\perp} R_A(x)$$

Interlacing Theorem: If $A \in M_n(\mathbb{R})$ is symmetric and $R^T R = Id$ and $B = R^T A R$ is $m \times m$, $m \le n$, then

$$\lambda_i(A) \le \lambda_i(B) \le \lambda_{i+(n-m)}(A).$$

Pf: B is symmetric.

Eigenvalues of A: $\lambda_1(A) \le \cdots \le \lambda_n(A)$

Eigenvectors of A: $\{u_1, \dots, u_n\}$

orthonormal families

$$A = UDU^T, \quad B = V\tilde{D}V^T$$

Eigenvalues of B: $\lambda_1(B) \leq \cdots \leq \lambda_m(B)$

Eigenvectors of B: $\{v_1, \ldots, v_m\}$

- Suffices to show $\lambda_i(A) \leq \lambda_i(B)$ since $\tilde{A} := -A$, $\tilde{B} := -B$ then $\lambda_i(\tilde{A}) \leq \lambda_i(\tilde{B})$.

$$\lambda_1(\tilde{A}) \leq \lambda_2(\tilde{A}) \leq \cdots \leq \lambda_n(\tilde{A})$$

$$\lambda_1(\tilde{B}) \leq \lambda_2(\tilde{B}) \leq \cdots \leq \lambda_m(\tilde{B})$$

$$\lambda_1(\tilde{A}) = -\lambda_n(A)$$

$$\lambda_1(\tilde{B}) = -\lambda_m(B)$$

$$-\lambda_n(A) \leq -\lambda_{n-1}(A) \leq \cdots \leq -\lambda_1(A)$$

$$-\lambda_m(B) \leq -\lambda_{m-1}(B) \leq \cdots \leq -\lambda_1(B)$$

$$\lambda_i(\tilde{A}) \leq \lambda_i(\tilde{B})$$

$$-\lambda_{n-i+1}(A) \leq -\lambda_{m-i+1}(B)$$

$$\lambda_{m-i+1}(B) \leq \lambda_{n-i+1}(A)$$

$j = m-i+1$

$j+(n-m) = n-i+1$

$$\lambda_j(B) \leq \lambda_{j+(n-m)}(A) \text{ as needed.}$$

Claim: $\lambda_i(A) \leq \lambda_i(B)$.

Goal: find test vector $\tilde{x}$ s.t. $\lambda_i(A) \overset{(2)}{\leq} R_B(\tilde{x}) \overset{(1)}{\leq} \lambda_i(B)$.

(1) $\forall\ \overset{\circ}{x} \in \operatorname{span}\{v_1, \dots, v_i\}$,

$$R_B(x) \le \max_{0 \ne x \in \operatorname{span}\{v_1, \dots, v_i\}} R_B(x) = \lambda_i(B) \checkmark$$

(2)

$$R_B(x) \overset{\text{def}}{=} \frac{x^T B x}{x^T x} = \frac{x^T R^T A R x}{x^T x} = \frac{(Rx)^T A (Rx)}{x^T x}$$

$$\overset{R^T R = Id}{=} \frac{(Rx)^T A (Rx)}{x^T R^T R x} = \frac{(Rx)^T A (Rx)}{(Rx)^T (Rx)} = R_A(Rx)$$

i.e., $R_B(x) = R_A(Rx)$.

Want: $\tilde{x} \in \operatorname{span}\{v_1, \dots, v_i\}$ and $R\tilde{x} \in \left(\operatorname{span}\{u_1, \dots, u_{i-1}\}\right)^\perp$

Then $R_B(\tilde{x}) \le \lambda_i(B)$

$\lambda_i(A) \le R_A(R\tilde{x})$.

$$\begin{aligned} Rx \in \operatorname{span}(\{u_1, \dots, u_{i-1}\})^\perp &\iff (Rx, u_j) = 0 \quad 1 \le j \le i-1 \\ &\iff \langle x, R^T u_j \rangle = 0 \quad 1 \le j \le i-1 \\ &\iff x \in \left(R^T \operatorname{span}\{u_1, \dots, u_{i-1}\}\right)^\perp \end{aligned}$$

$\dim \operatorname{span}\{u_1, \dots, u_{i-1}\} = i-1$

$\dim R^T \operatorname{span}\{u_1, \dots, u_{i-1}\} \le i-1$

$\dim \left(R^T \operatorname{span}\{u_1, \dots, u_{i-1}\}\right)^\perp \ge n-i+1$

$i + (n-i+1) > n \implies$ non zero intersection with $\operatorname{span}\{v_1, \dots, v_i\}$ $\leftarrow \dim = i$

$\blacksquare$

Thm: If $A, B \in M_n(\mathbb{R})$ both symmetric, then

$$|\lambda_i(A) - \lambda_i(B)| \leq \rho(A-B) \leq \|A-B\|_{op,2}.$$

Def: $V_k :=$ set of k-dimensional subspaces of $\mathbb{R}^n$.

$$W := \operatorname{span}\{u_1, \dots, u_k\} \in V_k$$

$$\max_{0 \neq x \in W} R_A(x) = \lambda_k(A)$$

$$\max_{0 \neq x \in \operatorname{span}\{u_2, \dots, u_{k+1}\}} R_A(x) = \lambda_{k+1}(A)$$

Thm: Let $W := \operatorname{span}\{u_1, \dots, u_k\} \in Gr_k(\mathbb{R}^n)$

$$\min_{W \in Gr_k(\mathbb{R}^n)} \max_{0 \neq x \in W} R_A(x) = \lambda_k(A)$$

$$\max_{W \in Gr_{n-k+1}(\mathbb{R}^n)} \min_{0 \neq x \in W} R_A(x) = \lambda_k(A).$$

LECTURE 10

Thm: $A \in M_n(\mathbb{R})$ symmetric with eigenvals $\lambda_1(A) \leq \cdots \leq \lambda_n(A)$.

$$\lambda_k = \min_{W \in Gr_k(\mathbb{R}^n)} \max_{0 \neq x \in W} R_A(x) = \max_{W \in Gr_{n-k+1}(\mathbb{R}^n)} \min_{0 \neq x \in W} R_A(x).$$

$$A = UDU^T, \quad U = \begin{pmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{pmatrix}, \quad \text{span}\{u_1, -, u_k\} \in Gr_k(\mathbb{R}^n).$$

$(\text{span}\{u_1, -, u_{k-1}\})^\perp \leftarrow (n-k+1)\text{-dim}$

Pf: $W := \text{span}\{u_1, -, u_k\} \in Gr_k(\mathbb{R}^n)$

$$\max_{0 \neq x \in W} R_A(x) = \lambda_k(A)$$

A symm. $\leftrightarrow$ A self-adjoint

$\cap$

$M_n(\mathbb{R})$

$$\min_{W \in Gr_k(\mathbb{R}^n)} \max_{0 \neq x \in W} R_A(x) \leq \lambda_k(A)$$

WTS: $\forall W \in Gr_k(\mathbb{R}^n)$, $\max_{0 \neq x \in W} R_A(x) \geq \lambda_k(A)$ b/c then

$$\inf_{W \in Gr_k(\mathbb{R}^n)} \cdots \geq \lambda_k(A).$$

Cl: $\forall W \in Gr_k(\mathbb{R}^n)$, $\exists x_0 \in W$ s.t. $R_A(x_0) \geq \lambda_k$.

Take $x_0 \in \left(\mathrm{span}\{u_1, \dots, u_{k-1}\}\right)^{\perp}$ $= \mathrm{span}\{u_k, u_{k+1}, \dots, u_n\}$ s.t. $x_0 \in W$, $x_0 \neq 0$.

Note: $\dim W = k$ and $\dim\left(\mathrm{span}\{u_1, \dots, u_{k-1}\}\right)^{\perp} = n-k+1$ and both live in $\mathbb{R}^n$ $\Rightarrow$ non-trivial intersection

$\Rightarrow$ choose x_0 in this intersection.

$\blacksquare$

Thm: Let $A, B \in M_n(\mathbb{R})$ be symmetric w/ eigenvals.

$$\lambda_1(A) \leq \dots \leq \lambda_n(A)$$

$$\lambda_1(B) \leq \dots \leq \lambda_n(B)$$

Then $|\lambda_k(A) - \lambda_k(B)| \leq \rho(A-B) \leq \|A-B\|_{op,2}$

$\forall k = 1, \dots, n$.

Pf: $A = UDU^T$, $U = \begin{pmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{pmatrix}$.

Set $V := \mathrm{span}\{u_1, \dots, u_k\} \in Gr_k(\mathbb{R}^n)$

As before, $\lambda_k(B) = \min\limits_{W \in Gr_k(\mathbb{R}^n)} \max\limits_{0 \neq x \in W} R_B(x)$

$$\leq \max_{0 \neq x \in V} \frac{x^T B x}{x^T x}$$

$B = A - (B-A)$ $\longrightarrow$

$x^T B x = x^T A x - x^T(B-A)x$

$$\leq \underbrace{\max_{0 \neq x \in V} \frac{x^T A x}{x^T x}}_{\lambda_k(A)} + \max_{0 \neq x \in V} \frac{x^T (B-A) x}{x^T x}$$

$$= \lambda_k(A) + \max_{0 \neq x \in V} \frac{x^T (B-A) x}{x^T x}$$

$$\leq \lambda_k(A) + \max_{0 \neq x \in \mathbb{R}^n} \frac{x^T (B-A) x}{x^T x}$$

$$= \lambda_k(A) + \max_i \lambda_i(B-A)$$

$\max_i |\lambda_i(A-B)| = \max_i |\lambda_i(B-A)|$

$$\leq \lambda_k + \rho(A-B)$$

Thm as before $\longrightarrow$

$$\leq \lambda_k + \|A-B\|_{op,2}$$

Upshot: $\lambda_k(B) \leq \lambda_k(A) + \rho(A-B)$

repeat arg. for B and A switched $\longrightarrow$ $\lambda_k(A) \leq \lambda_k(B) - \rho(A-B)$

$$\Rightarrow |\lambda_k(A) - \lambda_k(B)| \leq \rho(A-B) \leq \|A-B\|_2.$$

$\blacksquare$

Cor: Eigenvalues are sensitive to perturbations ?

Thm: $A, B \in M_n(\mathbb{R})$ symmetric

$$\left.\begin{aligned} \lambda_k(A+B) &\le \lambda_k(A) + \lambda_n(B) \\ \lambda_k(A+B) &\le \lambda_k(B) + \lambda_n(A) \end{aligned}\right\} \text{min max}$$

$$\left.\begin{aligned} \lambda_1(A) + \lambda_k(B) &\le \lambda_k(A-B) \\ \lambda_1(B) + \lambda_k(A) &\le \lambda_k(A+B). \end{aligned}\right\} \text{max min}$$

Corollary: λ_1 and λ_n are concave/convex.

Pf: $\lambda_1(\alpha A + (1-\alpha)B) \qquad \alpha \in (0,1)$

$$\begin{aligned} &\ge \lambda_1(\alpha A) + \lambda_1((1-\alpha)B) \\ &= \alpha \lambda_1(A) + (1-\alpha)\lambda_1(B) \end{aligned}$$

Similarly: $\lambda_n(\alpha A + (1-\alpha)B) \le \alpha \lambda_n(A) + (1-\alpha)\lambda_n(B)$.

$\blacksquare$

Pf: (of Thm) Let $W \in Gr_k(\mathbb{R}^n)$.

$$\max_{0 \ne x \in W} R_{A+B}(x) = \max_{0\ne x \in W} \left(\frac{x^T A x}{x^T x} + \frac{x^T B x}{x^T x} \right)$$

$$\le \max_{0 \ne x \in W} \frac{x^T A x}{x^T x} + \max_{0 \ne x \in W} \frac{x^T B x}{x^T x}$$

crude

$$\leq \max_{0 \neq x \in W} \frac{x^T A x}{x^T x} + \underbrace{\max_{0 \neq x \in \mathbb{R}^n} \frac{x^T B x}{x^T x}}_{= \lambda_n(B)}$$

bound

$$= \max_{0 \neq x \in W} R_A(x) + \lambda_n(B).$$

Taking min on both sides, get result. ∎

Corollary: (Monotonicity) $A, B \in M_n(\mathbb{R})$ symmetric and B positive-definite.

$$\lambda_k(A) \leq \lambda_k(A+B) \quad \forall k.$$

Pf: $\lambda_1(B) + \lambda_k(A) \leq \lambda_k(A+B)$ (by the previous Cor.)

B pos. def. $\Rightarrow \lambda_i(B) > 0 \quad \forall i$

$$\lambda_k(A) < \lambda_k(A+B).$$

□

Lemma: If A and B are $m \times n$, $m \geq n$, and rank $B = k$, then

$$\sigma_{i+k}(A) \leq \sigma_i(A-B) \qquad 1 \leq i \leq n-k.$$

Pf: (Sketch) $\tilde{A} := \left(\begin{array}{c|c} 0_{m\times m} & A_{m\times n} \\ \hline A^T_{n\times m} & 0_{n\times n} \end{array}\right)_{(m+n)\times(m+n)}$ symmetric.

$A = V\Sigma U^T$

Relation between eigenvals. of $\tilde{A}$ and A.

$AA^T = V\Sigma^2 V^T$

$A^TA = U\Sigma^2 U^T$

Claim: $\sigma_k(A) = \lambda_{m+n-k+1}(\tilde{A})$

$\sigma_k(A-B) = \lambda_{m+n-k+1}(\tilde{A}-\tilde{B})$

$AU = V\Sigma$

$A^TV = U\Sigma$

Claim: Use Weyl's lemma to get

$$\sigma_{i+j-1}(A) \leq \sigma_i(A-B) + \sigma_j(B)$$

Conclude by taking $j = k+1$.

□

Weyl's lemma: C_1, C_2 are Hermitian, then

$$\lambda_{j+k-n}(C_1+C_2) \leq \lambda_j(C_1) + \lambda_k(C_2)$$

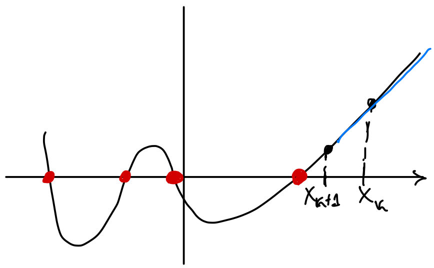
for $1 \leq j \leq n$, $1 \leq k \leq n$, $j+k \geq n+1$.

LECTURE 11

NEWTON'S METHOD. $G: \mathbb{R}^n \to \mathbb{R}^m$

seek points $x \in \mathbb{R}^n$ where $G(x) = 0$.

1-dim: $f: \mathbb{R} \to \mathbb{R}$ seek pts $x \in \mathbb{R}$ s.t. $f(x) = 0$



tangent line: $y - f(x_k) = f'(x_k)(x - x_k)$

Seek x_{k+1} s.t. $y = 0$

$$0 - f(x_k) = f'(x_k)(x_{k+1} - x_k)$$

$$\Rightarrow \boxed{x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}}$$

Hope that $f' \neq 0$

Thm: (Banach Fixed Pt. or Contraction Mapping)

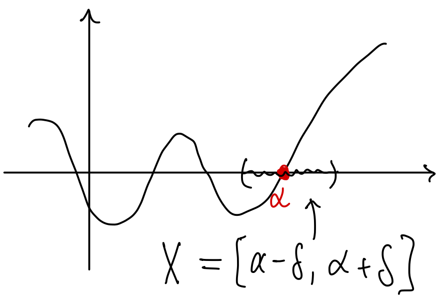
Let (X, d) be a complete metric space. Let $T: X \to X$ be a contraction (i.e., $\exists\, q \in [0,1)$ s.t. $d(Tx, Ty) \leq q\, d(x,y)$ $\forall x, y \in X$). Then $\exists!$ fixed pt. of T in X.

Pf: Goal: find fixed pt. of $g(x) = x - \frac{f(x)}{f'(x)}$.

since we want pts where $f = 0$

$$g(x) = x = x - \frac{f(x)}{f'(x)} \rightsquigarrow f(x) = 0.$$

Suppose $f \in C^2$ and that $f'(\alpha) \neq 0$.



Contraction: By MVT

$$g(x) - g(y) = g'(c)(x-y)$$

$$|g(x) - g(y)| = |g'(c)|\,|x-y|$$

Want: $|g'| < 1$ on $[\alpha-\delta, \alpha+\delta]$.

$$g'(x) = \frac{f(x) f''(x)}{[f'(x)]^2}.$$

Choose $\varepsilon_1 > 0$ s.t. $f' \neq 0$ on $[\alpha - \varepsilon_1, \alpha + \varepsilon_1]$.

$f''(x)$ is continuous on $[\alpha - \varepsilon_1, \alpha + \varepsilon_1]$

$\Rightarrow |f''| \leq M$ for some M

$f'(x)$ is continuous on $[\alpha - \varepsilon_1, \alpha + \varepsilon_1]$

$\Rightarrow |f'| \geq m > 0$

So, on $[\alpha - \varepsilon_1, \alpha + \varepsilon_1]$,

$$|g'(x)| \leq \frac{|f(x)|\, M}{m^2}$$

f continuous, $f(\alpha) = 0$. Given $w \in [0,1)$, $\exists \delta > 0$

s.t. $|x - \alpha| < \delta \Rightarrow |f(x) - f(\alpha)| < w \frac{m^2}{M}$

So,

$$|x - \alpha| < \delta \Rightarrow |f(x)| \leq \frac{m^2 w}{M} \Rightarrow |g'| < 1$$

on X

$\varepsilon := \min\{\varepsilon_1, \delta\}$, $X := [\alpha - \varepsilon, \alpha + \varepsilon]$.

Done by BFPT. $\blacksquare$

<u>NOTE</u>: $|e^{k+1}| \leq \frac{M}{2m} |e^k|^2$

$|e^k| \leq w^k |e^0|$.

LECTURE 12

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

- Assume $f(\alpha) = 0 \Rightarrow f'(\alpha) \neq 0$
- $f' \neq 0$ in nbhd of $\alpha \Rightarrow \left|\frac{1}{f'(x)}\right| < \frac{1}{m}$ in nbhd of α.
- f continuous in nbhd of α $\Rightarrow |f(x)| < M$ in nbhd of α

MULTIDIMENSIONAL NEWTON METHOD : $f : \mathbb{R}^d \to \mathbb{R}^d$ seek $\alpha \in \mathbb{R}^d$ s.t. $f(\alpha) = 0$.

$$x_{k+1} = x_k - \left[Df(x_k)\right]^{-1} f(x_k)$$

Assume:
- At α, $f(\alpha) = 0 \Rightarrow Df(\alpha)$ is invertible
- 1^{st} partials are continuous near α
  $\Rightarrow Df(x)$ is invertible near α
  $\Rightarrow \|Df(x)^{-1}\|_{op,2} < 1$ near α
- 2^{nd} partials are continuous near α.

Components of f: $f_i(x)$. Note: $f(\alpha) = 0 \Leftrightarrow f_i(\alpha) = 0 \;\forall i$.

Taylor Series: $f_i(\alpha) = f_i(x_k) + Df_i(x_k)(\alpha - x_k)$
$+ \frac{1}{2}(\alpha - x_k)^T \text{Hess}(\xi_k^{(i)})(\alpha - x_k)$,

$:: V_i$

$\xi_k^{(i)} :=$ segment connecting α to x_k.

$$0 = f(x_k) + Df(x_k)(\alpha - x_k) + v$$

$$0 = [Df(x_k)]^{-1} f(x_k) + (\alpha - x_k) + [Df(x_k)]^{-1} v$$

$$0 = (x_k - x_{k+1}) + (\alpha - x_k) + [Df(x_k)]^{-1} v.$$

Thus: $x_{k+1} - \alpha = [Df(x_k)]^{-1} v$

$$\Rightarrow \boxed{e_{k+1} = [Df(x_k)]^{-1} v}$$

$Df(\alpha)$ invertible and derivative is cont.

$\Rightarrow \forall x \in \overline{B_\varepsilon(\alpha)}, \quad \|Df(x)^{-1}\|_{op,2} \leq \frac{1}{m}$ for some $m > 0$.

$$\|x_k - \alpha\|_2 = \|[Df(x_k)]^{-1} v\|_2 \leq \|Df(x_k)^{-1}\|_{op,2} \|v\|_2$$

$$\leq \frac{1}{m} \|v\|_2 \text{ if } x_k \text{ is close to } \alpha$$

Then

$$\|v\|_2 = \sqrt{\sum_i v_i^2} \leq \|v\|_\infty \sqrt{d} = \sqrt{d} \max_i |v_i|.$$

$$|v_i| = \frac{1}{2} \left| (\alpha - x_k)^T \operatorname{Hess}(\xi_k^{(i)}) (\alpha - x_k) \right|$$

$$\leq \frac{1}{2}\|e_k\|_2 \, \|\mathrm{Hess}(\xi_k^{(i)})\, e_k\|_2$$

$$\leq \frac{1}{2}\underbrace{\|\mathrm{Hess}(\xi_k^{(i)})\|_2}_{\text{max sing. val.}}\|e_k\|_2^2$$

Hess is symm $\Rightarrow$ max sing val = max eigenval.

$$= \frac{1}{2}\underbrace{\rho(\mathrm{Hess}(\xi_k^{(i)}))}_{\leq M^2 d^2}\|e_k\|_2^2$$

Thm: $f : \Omega \to \mathbb{R}^d$, $\Omega \subset \mathbb{R}^d$ open. Assume $\exists\, \alpha \in \Omega$ s.t. $f(\alpha) = 0$, $Df(\alpha)$ invertible, and that $\exists\, \varepsilon_0 > 0$ s.t. all 2^{nd} derivatives of f are continuous on $\overline{B_{\varepsilon_0}(\alpha)}$.

If $\exists\, \varepsilon > 0$ s.t. $x_0 \in \overline{B_\varepsilon(\alpha)}$, then Newton's method converges to α and $\exists\, C > 0$ s.t.

$$\|e_{k+1}\|_2 \leq C\,\|e_k\|_2 \quad \forall\, k \geq 0.$$

LECTURE 13

Minimize $Q := \frac{1}{2} x^T A x - x^T b$ subject to $\underbrace{B^T x = f}_{n \text{ constraints}}$

$A \in M_m(\mathbb{R})$

Assume A is symmetric throughout $\Rightarrow A = U\underset{\text{Diagonal}}{D}U^T$

$$x^T A x = (U^T x)^T D (U^T x)$$

- If A has a negative eigenvalue $\Rightarrow \nexists$ minimizer.
- If A is symmetric + eig $(A) > 0 \Rightarrow Q(x)$ has a unique global minimizer.
- If A is symmetric + eig $(A) \geq 0 \Rightarrow$ whether or not there's a minimizer will depend on b

$$Q(x) = \frac{1}{2} x^T A x - \underbrace{x^T b}_{\substack{\text{can cause problems in} \\ \text{the zero eigendirection}}}$$

Def: A is symmetric pos. def. if A is symmetric and eig$A > 0$.

A symm pos. def. $\iff A \succ 0$

Def: A is symmetric pos. semidef. if A is symm. and $\operatorname{eig} A \geq 0$
$\iff A \succeq 0$.

$\boxed{\begin{array}{l}\text{Thm: If } Q(x) = \frac{1}{2} x^T A x - x^T b \text{ and } A \succ 0 \text{ then} \\ \text{there exists a unique global minimizer given by} \\ \qquad x_0 = A^{-1} b, \quad Q_{min} = -\frac{1}{2} b^T A^{-1} b.\end{array}}$

Idea: Compute derivative of Q if it's diff'ble

$$\lim_{\varepsilon \to 0} \frac{Q(x+\varepsilon w) - Q(x)}{\varepsilon}.$$

Pf: Look at $Q(x) - Q(x_0)$ and write x as
$x_0 + w$, $w \in \mathbb{R}^m$.
Look at $Q(x_0 + w) - Q(x_0)$. ...

$\square$

Thm: If $\nabla F(x_0) = 0$ and if $\operatorname{Hess}(x_0) \succ 0$ then, close to x, can write

$$F(x) - F(x_0) = \frac{1}{2} (x - x_0)^T \operatorname{Hess}(\xi) (x - x_0).$$

→ w/out symmetry assumption.

REMARK: Only eig $A > 0 \;\not\Rightarrow\; \exists$ global minimizer.

$$A = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}, \quad \text{eig}(A) > 0$$

$$x^T A x = (x_1 \;\; x_2)\, A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \dots = (x_1 - x_2)^2 + 2x_1 x_2 + \alpha x_1 x_2$$

$$= (x_1 - x_2)^2 + (\alpha + 2)\, x_1 x_2$$

Choose $\alpha < -2 \implies$ no minimizer.

QUADRATIC MINIMIZATION w/ LINEAR CONSTRAINTS

PRIMAL PROBLEM: minimize $Q(x) = \frac{1}{2} x^T A x - x^T b$
subject to $B^T x = f$.

$$\begin{cases} A \;\; m \times m, \; A \succ 0 \\ B \;\; m \times n, \; n \le m, \; \text{rank } B = n \\ f \in \mathbb{R}^n \end{cases}$$

LAGRANGE:
$$\mathcal{L}(x, \lambda) = \underbrace{\frac{1}{2} x^T A^{-1} x - b^T x}_{Q(x)} + \underbrace{\lambda^T}_{\substack{\lambda \in \mathbb{R}^n \\ n \text{ Lagrange multipliers}}} \underbrace{(B^T x - f)}_{\text{constraints}}$$

Goal: find critical points of $\mathcal{L}$ in $\mathbb{R}^{n+m}$.

Hope: critical pt solves the constrained problem.

$$\begin{cases} \nabla_x \mathcal{L} = A^{-1}x - b + B\lambda \\ \nabla_y \mathcal{L} = B^T x - f \end{cases} \quad \longleftrightarrow \quad \begin{pmatrix} A^{-1} & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} b \\ f \end{pmatrix}$$

$$\begin{cases} x_0 = A(b - B\lambda) \\ \lambda_0 = (\underbrace{B^T A B}_{\text{invertible}})^{-1} (B^T A b - f) \end{cases}$$

and $\succ 0$ b/c rank $B = n$.

Thm:
$$\begin{cases} \text{minimize} \quad Q(x) := \frac{1}{2} x^T A x - x^T b \\ \text{subject to} \quad B^T x = f \end{cases}$$

has a unique solution (x_0, λ_0) given by the solution to the system

$$\begin{pmatrix} A^{-1} & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} b \\ f \end{pmatrix}$$

Moreover, λ_0 is the <u>unique</u> maximizer of $-G(\lambda)$, where

$$G(\lambda) := \frac{1}{2}(B\lambda - b)^T A (B\lambda - b) + \lambda^T f.$$

$\mathcal{L}(x,\lambda) = Q(x) + \lambda(B^T x - f)$ quadratic in x; minimize w.r.t. $x \to x_0(\lambda)$

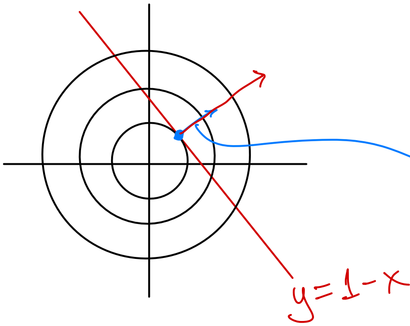
$\rightsquigarrow Q(x_0(\lambda)) = G(\lambda)$

DUAL PROBLEM: Maximize $-G(\lambda)$ over $\mathbb{R}^n$.

Ex: $Q(x,y) = x^2 + y^2$ min. Q over $x+y=1$

$y = 1-x$

$q(x) = x^2 + (1-x)^2$



antiparallel to normal of constraint.

Pf: (Thm) Consider $Q(x) + G(\lambda)$ and assume x is admissible

$$Q(x) + G(\lambda) = \dots = \frac{1}{2}\left(A^{-1}x + B\lambda - b\right)^T A \left(A^{-1}x + B\lambda - b\right)$$

$A \succ 0 \Rightarrow x$ admissible gives $Q(x) + G(\lambda) \geq 0$

with "=" only if $A^{-1}x + B\lambda - b = 0$

i.e., $Q(x) \geq -G(\lambda) \quad \forall \lambda \in \mathbb{R}^n$

$$\underbrace{\min_{x \text{ admissible}} Q(x)} \geq -G(\lambda) \quad \forall \lambda \in \mathbb{R}^n$$

primal problem

$$Q(x_0) = \min_{x \text{ admissible}} Q(x) \geq \max_{\lambda \in \mathbb{R}^n} -G(\lambda) = -G(\lambda_0)$$

$\square$

$\underline{\underline{A \succeq 0}}$:

$$Q(x) = \frac{1}{2} x^T A x - x^T b = \frac{1}{2} x_1^2 - b_1 x_1 - b_2 x_2$$

$$\Rightarrow \begin{cases} \text{No minimizer if } b_2 \neq 0 \\ \exists \text{ minimizer if } b_2 = 0 \end{cases}$$

$x_0 := \begin{pmatrix} b_1 \\ * \end{pmatrix}$ minimizer

i.e., have ∞-many minimizers

i.e., $b \in \operatorname{im} A \Rightarrow \exists$ ∞-many minimizers if $\operatorname{rank} A < n$

$b \notin \operatorname{im} A \Rightarrow \nexists$ minimizer

Thm: Suppose A is symmetric, rank $A = r$. Then,

$Q(x) := \frac{1}{2} x^T A x - x^T b$ has a minimum

if and only if $\underbrace{A \succeq 0 \text{ and } (Id - AA^+) b = 0}_{\text{same as } b \in \text{im } A}$

- if $A \succeq 0$ and $(Id - AA^+) b = 0$ and

$$A = U^T \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix} U,$$

then the optimal value is attained at

$$x = A^+ b + U^T \begin{pmatrix} 0 \\ z \end{pmatrix}, \quad z \in \mathbb{R}^n.$$

Obs: $A \succeq 0$, $n \times n$

A^+ pseudo-inverse

AA^+ projects a vector in $\mathbb{R}^n$ onto im A.

LECTURE 14

GRADIENT DESCENT

$J : V \to \mathbb{R}$ want to find u_* s.t. $J(u_*) = \min_{\text{close to } u} J(u)$

J strictly convex $\Rightarrow \exists$ unique minimizer

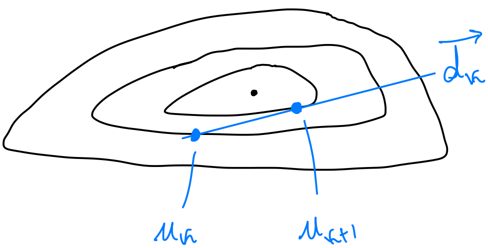
$\{u_k\}$ hope $u_k \to u_*$

Idea: Go downhill $u_{k+1} = u_k + \rho\, d_k$, $\rho > 0$
d_k well-chosen

If $\nabla J(u_k)$ exists and $\underbrace{\langle \nabla J(u_k), d_k \rangle}$ $< 0 \Rightarrow$ decrease

Most negative when $d_k = -\nabla J(u_k)$

- How to choose d_k? Natural to take $d_k = -\nabla J(u_k)$
- How to choose step size?
    - Choose fixed step size
    - Exact line search
    - Backtracking
    - Choose some seq. $\{\rho_k\}$ and hope for the best.



$$\phi(\rho) := J(u_k + \rho d_k)$$

$$\phi'(\rho) := \langle \nabla J(u_k + \rho d_k), d_k \rangle$$

$$\phi'(0) < 0 \quad \text{by choice of } d_k.$$

$\phi'(\rho)$ seek ρ s.t. $\nabla J(u_k + \rho d_k) \perp d_k$.

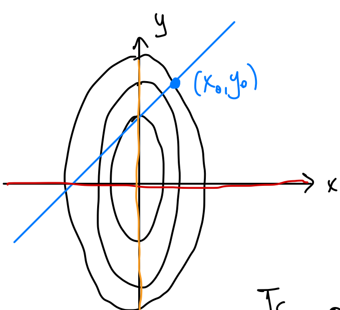
Obs: If J is quadratic, then ∇J is linear so can explicitly solve for ρ, otherwise Newton.

WANT: $J(u_k + \rho d_k) \leq J(u_k) - \frac{1}{2}\langle \nabla J(u_k), d_k \rangle$ $\leftarrow$ "BACKTRACKING"

Ex: $J(x,y) = \alpha x^2 + \beta y^2 \rightsquigarrow \nabla J = \begin{pmatrix} 2\alpha x \\ 2\beta y \end{pmatrix}$

Gradient Descent: $d_k = -2\begin{pmatrix} \alpha x_k \\ \beta y_k \end{pmatrix}$

Exact line Search:



line $2\alpha x_0 (y - y_0) = 2\beta y_0 (x - x_0)$ (*)

if $(x_0, y_0) = (x_0, 0)$ (*)

if $(x_0, y_0) = (0, y_0)$ (*)

Is origin on line otherwise?

$$2\alpha x_0(-y_0) = 2\beta y_0(-x_0) \rightsquigarrow \text{yes if } \alpha=\beta.$$

Rescale x & y : $J(x,y) = \frac{1}{2}(x^2+by^2)$, $b<1$.

$\nabla J = \begin{pmatrix} x \\ by \end{pmatrix}$. Exact line search: seek t s.t.

$$\nabla J(x_k - tx_k, y_k - ty_k) \perp \nabla J(x_k, y_k)$$

linear in t.

$\Rightarrow$ Solve for t... get $t = \dfrac{x_k^2 + b^2 y_k^2}{x_k^2 + b^3 y_k^2}$

$$x_{k+1} = \frac{(b-1)b^2 x_k y_k^2}{x_k^2 + b^3 y_k^2}, \quad y_{k+1} = \frac{(1-b)x_k^2 y_k}{x_k^2 + b^3 y_k^2}.$$

Start at $(x_0, y_0) = (b, 1)$

$$x_k = b\left(\frac{b-1}{b+1}\right)^k, \quad y_k = \left(\frac{1-b}{1+b}\right)^k$$

$$J(x_k, y_k) = \left(\frac{1-b}{1+b}\right)^{2k} J(x_0,y_0) \approx (1-4b)^k J(x_0,y_0)$$

$$b = \frac{1}{10} \Rightarrow \left(\frac{1-b}{1+b}\right)^2 \approx 0.67$$

$b = \frac{1}{100} \Longrightarrow \left(\frac{1-b}{1+b}\right)^2 \approx 0.96$ close to 1 slow descent

Upshot: Gradient descent <u>not</u> <u>so</u> <u>great</u> <u>for</u> nice <u>problems</u>.

POLYAK METHOD : 2 parameters s, β Initial state : (u_0, z_0) and (u_1, z_1)

$$u_{k+1} = u_k - s\left(\nabla J(u_k) + \beta z_{k-1}\right) \qquad \partial_s\,(u_2, z_2)$$

How to choose s, β? Set $z_k := \nabla J(u_k) + \beta u_{k-1}$

$$J(u) = \frac{1}{2} u^T S u, \quad \underset{\text{Hess}(J)}{S} = \begin{pmatrix} 1 & 0 \\ 0 & b \end{pmatrix}, \quad \text{eig } S = \{1, b\}.$$

Assume we start w/ eigenvec. of S $\begin{cases} u_0 = c_0 \vec{q} \\ z_0 = d_0 \vec{q} \end{cases}$

$$\begin{cases} u_1 = u_0 - s z_0 = c_0 \vec{q} - s d_0 \vec{q} \\ z_1 - S u_1 = \beta z_0 = \beta d_0 \vec{q} \end{cases}$$

$$\begin{cases} u_1 = c_1 \vec{q} = c_0 \vec{q} - s d_0 \vec{q} \\ d_1 \vec{q} - \lambda c_1 \vec{q} = \beta d_0 \vec{q} \end{cases}$$

$$\begin{pmatrix} 1 & 0 \\ -\lambda & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ d_1 \end{pmatrix} = \begin{pmatrix} 1 & -s \\ 0 & \beta \end{pmatrix} \begin{pmatrix} c_0 \\ d_0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 \\ -\lambda & 1 \end{pmatrix} \begin{pmatrix} c_{k+1} \\ d_{k+1} \end{pmatrix} = \begin{pmatrix} 1 & -s \\ 0 & \beta \end{pmatrix} \begin{pmatrix} c_k \\ d_k \end{pmatrix}$$

$$\begin{pmatrix} c_{k+1} \\ d_{k+1} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix} \begin{pmatrix} 1 & -s \\ 0 & \beta \end{pmatrix}} \begin{pmatrix} c_k \\ d_k \end{pmatrix}$$

$=: R_{s,\beta}(\lambda) \to$ Eigenvals.
$e_1(\lambda, s, b), e_2(\lambda, s, b)$

For every eigenval. of S: $\mathrm{eig}(S) \subset [\lambda_{min}, \lambda_{max}]$

Seek β, s to minimize $\{|e_1(\lambda, s, b)|, |e_2(\lambda, s, b)|\}$

Rate of Decay of $J(u_k, y_k) \approx (1 - 4\sqrt{b})^k J(x_0, y_0)$

$\lambda_{min} > 0$ strictly convex $\quad J(x_k, y_k) = \left(\frac{1-b}{1+b}\right)^{2k} J(x_0, y_0)$

$$\simeq (1 - 4b)^k J(x_0, y_0)$$

Upshot: Polyak faster than Gradient descent.

GRADIENT DESCENT WITH FIXED STEP SIZE

$$u_{k+1} = u_k - \eta \nabla J(u_0)$$

- Assume ∇J is Lipschitz: $\|\nabla J(u) - \nabla J(w)\|_2 \leq L \|u - w\|_2$

Obs: minimize J over $\mathbb{R}^n$ if J is reasonable (i.e., coercive) then

$$\{u : J(u) \leq J(u_0)\} \text{ is closed \& bdd}$$

if ∇J is diffable on closed & bdd $\Rightarrow$ Lipschitz.

Thm: ∇J is Lipschitz on convex domain Ω, then

$$J(u) \leq J(v) + \langle \nabla J(u), u-v \rangle + \frac{L}{2}\|v-u\|_2^2 .$$

Fundamental Thm of Calculus

Pf:
$$J(u) - J(v) = \int_0^1 \frac{d}{dt} J(v + t(u-v))\, dt$$

$$= \int_0^1 \langle \nabla J(v + t(u-v)), t(u-v) \rangle\, dt$$

$$- \int_0^1 \langle \nabla J(v), u-v \rangle\, dt$$

$$+ \int_0^1 \langle \nabla J(v), u-v \rangle\, dt$$

$$\leq \int_0^1 \|\nabla J(v + t(u-v)) - \nabla J(v)\| \|u-v\|\, dt$$

$$+ \langle J(v), u-v \rangle$$

$$\leq \int_0^1 L \|t(u-v)\| \|u-v\|\, dt + \langle \nabla J(v), u-v \rangle$$

$$= L\|u-v\|^2 \underbrace{\int_0^1 t\, dt}_{=\frac{1}{2}} + \langle \nabla J(v), u-v \rangle$$

$\square$

Cl: Gradient descent converges $\to$ If ∇J is Lipschitz w/ Lipschitz constant L then for any $0 \leq \eta \leq 1/2$ we have $J(u_{k+1}) \leq J(u_k) - \frac{\eta}{2} \|\nabla J(u_k)\|^2$

$\Rightarrow \{J(u_k)\}$ decreasing

Pf: Use $u = u_{k+1}$ in previous thm.
$v = u_k$

and use grad. descent: $u_{k+1} = u_k - \eta \nabla J(u_k)$

$$\iff u_{k+1} - u_k = -\eta \nabla J(u_k).$$

$\square$

Upshot: if $\eta \leq \frac{1}{2}$, then by the above we know,

$$J(u_{k+1}) \leq J(u_k) - \frac{\eta}{2} \|\nabla J(u_k)\|^2.$$

Now, if there is a lower bound J_*, then

$$\iff \|\nabla J(u_k)\|^2 \leq \frac{2}{\eta}\left[J(u_k) - J_*\right]$$

$$\iff \sum_{i=0}^{N-1} \|\nabla J(u_i)\|^2 \leq \frac{2}{\eta}\left(J(u_0) - J_*\right)$$

So at last one $\|\nabla J(u_i)\|^2$ has to be small: namely $\leq \frac{2}{\eta N}\left(J(u_0) - J_*\right)$.

LECTURE 15

CONVERGENCE OF GRADIENT DESCENT w/ FIXED STEPSIZE η

$J: \mathbb{R}^n \to \mathbb{R}$ seek a minimizer $u_{k+1} = u_k - \eta \nabla J(u_k)$

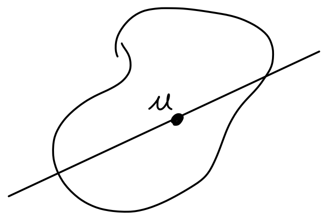
"OBJECTIVE FUNCTION"

(1) Assume ∇J is L-Lipschitz in $\mathbb{R}^n$.

$\Rightarrow$ Given $J(u)$ and $\nabla J(u)$ for some $u \in \mathbb{R}^n$, then we can bound J on a line through that pt. u by:

$$J(v) \leq J(u) + \langle \nabla J(u), v-u \rangle + \frac{L}{2}\|v-u\|_2^2$$

consequ. of being L-Lipschitz



(2) Gradient descent with fixed stepsize η.

- if $\eta \leq \frac{1}{L}$, then

$$J(u_{k+1}) \leq J(u_k) - \frac{L}{2}\|\nabla J(u_k)\|_2^2$$

needed that the direction of descent

is $-\eta \nabla J(u_k)$.

(3) Assume $\inf_{v \in \mathbb{R}^N} J(v) =: J_*$ exists. This implies that

$$\sum_{k=0}^{N-1} \|\nabla J(u_k)\|_2^2 \leq \frac{2}{\eta}\left[J(u_0) - J_*\right]$$

Thus:

- $\|\nabla J(u_k)\|_2^2 \leq \frac{2}{\eta}\left[J(u_0) - J_*\right]$

- If you've taken N steps, $\exists$ at least one k with $\|\nabla J(u_k)\|_2^2 \leq \frac{1}{N}\frac{2}{\eta}\left[J(u_0) - J_*\right]$.

(4) Assume J is convex if there exists some u_* s.t. $J(u_*) = J_*$ lower bound, can better control gradients as:

minimizer

$$J(u_k) - J(u_*) \leq \underbrace{\frac{1}{k}}_{} \frac{1}{2\eta}\|u_0 - u_*\|_2^2 .$$

rate of convergence $= \frac{1}{k}$

$\Rightarrow$ J's converge but not necessarily the iterates.

Q: Can get faster rate of convergence?

(5) Additional condition on J : Polyak - Łojasiewicz

$J: \mathbb{R}^n \to \mathbb{R}$ has lower bound J_* $\left(\inf_{\mathbb{R}^n} J(v)\right)$

satisfies the P-Ł condition iff $\exists\, \mu > 0$ s.t.

$$\frac{1}{2} \|\nabla J(v)\|_2^2 \geq \mu \left[J(v) - J_* \right].$$

Note: Strong convexity $\Rightarrow$ PŁ condition $\not\Rightarrow$ strong convexity

PŁ condition $\not\Rightarrow$ convexity.

With the PŁ condition, we get

$$J(u_k) - J_* \leq \left(1 - \frac{\mu}{L}\right)^k \left[J(u_0) - J_* \right]$$

and $\exists\, u_*$ and $u_k \to u_*$ and $J(u_*) = J_*$

Moreover,

$$\|u_k - u_*\|_2^2 \leq \frac{8\eta L}{\mu^2} \left(1 - \frac{\mu}{L}\right)^{k-1} \left[J(u_0) - J_* \right].$$

rate of convergence of iterates

$\boxed{\text{* linear convergence}}$

LECTURE 16

CONJUGATE GRADIENT

$$J(u) = \frac{1}{2} u^T A u + b^T u, \quad A \succ 0 \implies \boxed{\nabla J(u) = Au - b}$$

- Fast way of solving $Ax = b$ for large problems if A is $n \times n$ then reach exact solution in $n-1$ steps $u_0, u_1, \ldots, u_{n-1}$

iterative method

solves exactly

- At each step determine direction of d_k multiplier ρ_k using $u_{k+1} = u_k - \rho_k d_k$
- Choose u_{k+1} by minimizing $J(u)$ over affine subspace $u_k + G_k$ where $G_k := \text{span}\{\nabla J(u_0), \ldots, \nabla J(u_k)\}$.

↳ Terminates when $G_k = \mathbb{R}^n$

Conjugate Gradient Method

1) Minimize J over $u_0 + \text{span}\{\nabla J(u_0)\}$

$V_e + \text{span}\{Au_0\}$

Denote minimizer: u'

2) Minimize J over $u_1 + \text{span}\{\nabla J(u_0), \nabla J(u_1)\}$

$$u' + \text{span}\{Au_0, Au_1\}$$

Wanting $u_2 = u_1 + \begin{pmatrix} | & | \\ Au_0 & Au_1 \\ | & | \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$

If, after k steps you haven't converged,

- $\{\nabla J(u_0), \_, \nabla J(u_k)\}$ are mutually orthogonal and nonzero $\Rightarrow G_k$ is $k+1$ dimensional

- $u_{k+1} = u_k - \rho_k d_k \Rightarrow u_{k+1} - u_k = -\rho_k d_k$

  and $\{d_0, \_, d_{k-1}\}$ mutually orthogonal w.r.t. $\langle \cdot, \cdot \rangle_A := \langle A\cdot, \cdot \rangle$

  $A \succ 0 \Rightarrow \{d_0, \_, d_{k-1}\}$ are lin. indep. $\langle Ad_i, d_j \rangle = 0 \ \forall i \neq j$.

- Explicit formula for d_k's and ρ_k's:

  $d_1 = \nabla J(u_0) + (\text{smth})\, d_0$

  $d_2 = \nabla J(u_1) + (\text{smth})\, d_1$, $\rho_k = \dfrac{\langle \nabla J(u_k), d_k \rangle}{\langle A d_k, d_k \rangle}$

◢

Prop: Let $J(u) := \frac{1}{2} u^T A u - b^T u$, $A \succeq 0$ so that $\nabla J(u) = Au - b$. Assume $\nabla J(u_i) \neq 0 \ \forall i = 0, \_, k$. Then the problem

$$\min_{v \in u_k + G_k} J(v)$$

has a unique solution and $\langle \nabla J(u_{k+1}), \nabla J(u_i) \rangle = 0$

Pf: HW04 Q1.

$\square$

Claim: If the process hasn't terminated by step k, then

$$\langle A(u_{i+1} - u_i),\ u_{j+1} - u_j \rangle = 0 \quad \forall i \neq j,\ 1 \leq i, j \leq k$$

and these differences are lin. indep..

Note: $\nabla J(u+w) = \nabla J(u) + Aw.$

Thm: Assume Conjugate Gradient hasn't terminated by step k. Let

$$d_\ell := \nabla J(u_\ell) + \sum_{i=0}^{\ell-1} \lambda_i^\ell \nabla J(u_i)$$

Then $\lambda_i^k = \dfrac{\|\nabla J(u_k)\|^2}{\|\nabla J(u_i)\|^2} \quad 0 \leq i \leq k-1$

and $d_0 = \nabla J(u_0)$

$$d_\ell = \nabla J(u_\ell) + \frac{\|\nabla J(u_\ell)\|^2}{\|\nabla J(u_{\ell-1})\|^2} d_{\ell-1} \qquad 0 \leq \ell \leq k.$$

Implement: Generate random u_0

Compute ρ_0 (see next for how to do this)

$u_1 = u_0 - \rho_0 d_0$

$\nabla J(u_1)$

$$d_1 = \nabla J(u_1) + \frac{\|\nabla J(u_1)\|^2}{\|\nabla J(u_0)\|^2} d_0$$

compute ρ_1

$u_2 = u_1 - \rho_1 d_1$

$\nabla J(u_2)$

$$d_2 = \nabla J(u_2) - \frac{\|\nabla J(u_2)\|^2}{\|\nabla J(u_1)\|^2} d_1$$

$\vdots$

LECTURE 17

CONSTRAINED OPTIMIZATION

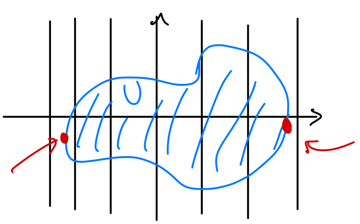
$$\begin{cases} \min J(u) \\ \text{subject to } u \in U \end{cases}$$ constraint set

$J : \Omega \to \mathbb{R}$, Ω open normed vector space

$U \subset \Omega$

Ex: $J(x,y) = x$

J has a relative local min. at u w.r.t. U.

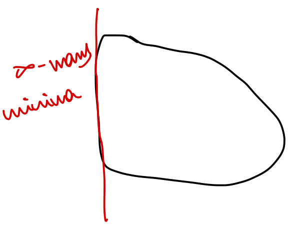


relative local minimum of J at u

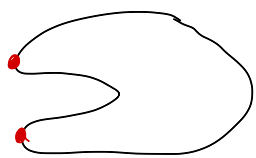
Def: $J : \Omega \to \mathbb{R}$, w/ Ω is an open set in a normed vec. space, and let $U \subset \Omega$. Then J has a local min. at u w.r.t. U if $\exists$ open set $W \subset \Omega$ and $u \in W$ s.t.

$$J(u) \leq J(w) \quad \forall w \in U \cap W.$$

U convex



U not convex



How to specify U? $\phi : U \to \mathbb{R}$

$$U := \{x \in \Omega : \phi_i(x) \overset{\text{equality}}{=} 0 \ \forall i\}$$

or: $U := \{x \in \Omega : \phi_i(x) \underset{\text{inequality}}{\leq} 0 \ \forall i\}$ ← inequality most general b/c can say $\phi(x) = 0 \iff \phi(x) \leq 0$ and $-\phi(x) \leq 0$

Ex: $Ax = b$

$$A = \begin{pmatrix} - a_1 - \\ \vdots \\ - a_n - \end{pmatrix}, \quad \phi_i(x) = a_i x - b_i = 0.$$

4.2 in book

<u>Thm</u>: (Lagrange's Method) Let $J : \Omega \to \mathbb{R}$, $\Omega \underset{open}{\subset} \mathbb{R}^n$, have a relative local extremum at $u \in U \subset \Omega$ and assume J is differentiable at u. Let the constraint fcts $\phi_i : \Omega \to \mathbb{R}$ be C^1 and s.t. $\{\nabla \phi_1(u), \ldots, \nabla \phi_m(u)\}$, $m < n$, are linearly independent. Set

$$L : \Omega \times \mathbb{R}^m \longrightarrow \mathbb{R}$$

$$L(x, \lambda) := J(x) + \lambda_1 \phi(x) + \cdots + \lambda_m \phi_m(x).$$

Then, $\exists \lambda_1, \ldots, \lambda_m$ s.t.

$$\nabla J(u) + \lambda_1 \nabla \phi_1(u) + \cdots + \lambda_m \nabla \phi_m(u) = 0.$$

And $\begin{cases} \nabla_x L = 0 \longrightarrow \text{critical pt} \\ \nabla_\lambda L = 0 \longrightarrow x \in U = \{x \in \Omega : \phi_i(x) = 0,\ i = 1, \ldots, m\}. \end{cases}$

<u>NOTE</u>: Analog of this Thm for inequality constraint "$\phi_i(x) \leq 0$" has the so called KKT condition.

LECTURE 18

CONE OF FEASIBLE DIRECTIONS $C(u)$

← does not have to be convex

$\min J(u)$ $u \in U$, where U is defined by equality constraints.

In this case, can introduce the Lagrangian $\mathcal{L}(x,\lambda)$

$$\mathcal{L}(u,\lambda) = J(u) + \lambda_1 \phi_1(u) + \cdots + \lambda_m \phi_m(u)$$

Thm: If u is a local min. of $\mathcal{L}$ w.r.t. U, then

$$\nabla J(u) + \lambda_1 \nabla\phi_1(u) + \cdots + \lambda_m \nabla\phi_m(u) = 0$$

has a solution $\{\lambda_i\} \subset \mathbb{R}$ with $\lambda_i \geq 0$ $\forall i$.

$$\Updownarrow$$

Trying to solve $A\vec{\lambda} = \vec{b}$ so that $\vec{\lambda} \geq 0$

New cone $C^*(u)$ = intersection of half-spaces

Farkas - Minkowski $\Rightarrow$ KKT conditions

$$\boxed{\text{Thm: } C(u) \subset C^*(u)}$$

"qualified" points $\rightsquigarrow$ If u is a qual. pt. $\Rightarrow C(u) = C^*(u)$.

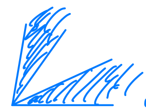
↑ But $C(u) \subsetneq C^*(u)$ is possible

Def: Given a $\mathbb{R}$-vector space V, a nonempty set $C \subset V$ is a <u>cone with apex 0</u> if $\forall v \in C$ we have $\lambda v \in C$ $\forall \lambda > 0$.

If $u \in V$ then $u + C$ is a cone with vertex u if C has a cone with vertex zero.

Note: (i) 0 doesn't have to be in C

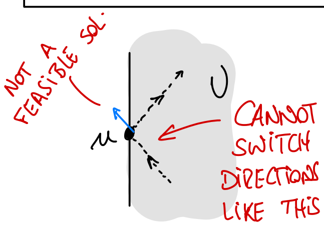
(ii) C doesn't have to be convex $\longrightarrow$

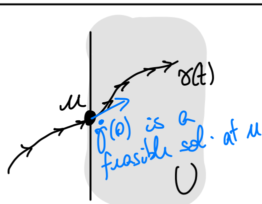


is a cone

(iii) C doesn't have to have interior

Def: Let V be a normed space and $U \subset V$ a nonempty subset. For any $u \in U$, <u>the cone of feasible directions at u</u>, denoted $C(u)$, is the union of $\vec{0}$ and the set of nonzero vectors $w \in V$ s.t. $\exists$ a seq. $\{u_k\} \subset U$, $u_k \neq u$, and $u_k \to u$, and
$$\frac{u_k - u}{\|u_k - u\|} \longrightarrow \frac{w}{\|w\|}.$$





$\gamma(0) = u$

$\gamma(t) \in U$, $0 \leq t \ll 1$

$u_k \overset{\text{def}}{=} \gamma\left(\frac{1}{k}\right)$

$\lim_{k \to \infty} \gamma\left(\frac{1}{k}\right) = \gamma(0) = 0$

$$u_k - u = \gamma\left(\tfrac{1}{k}\right) - \gamma(0) = \dot{\gamma}(\xi),$$
$$0 < \xi < \tfrac{1}{k} \text{ by MVT.}$$

<u>NOTE</u>: If $\dfrac{u_k - u}{\|u_k - u\|} \longrightarrow \dfrac{w}{\|w\|}$

$\Rightarrow \forall \varepsilon > 0 \ \exists N$ s.t. for $k > N$,
$$\frac{u_k - u}{\|u_k - u\|} \in B_\varepsilon\left(\frac{w}{\|w\|}\right).$$

i.e. $\dfrac{u_k - u}{\|u_k - u\|} = \dfrac{w}{\|w\|} + \delta_k$ for some $\|\delta_k\| < \varepsilon$.

$$\Rightarrow \boxed{u_k = u + \|u_k - u\|\left(\frac{w}{\|w\|}\right) + \delta_k \|u_k - u\| \quad \forall k > N.}$$

<u>Proposition</u>: Let $U \subset V$, V normed vec. space, $u \in U$. Then

(i) $C(u)$ is closed

(ii) Let $J : \Omega \to \mathbb{R}$, Ω open set, $U \subset \Omega$. If J has a local minimum at u w.r.t. U and if $\nabla J(u)$ is defined then
$$\langle \nabla J(u), v - u \rangle \geq 0 \quad \forall v - u \in C(u).$$

$\eta_k \to 0$ as $k \to \infty$

<u>Obs</u>: J differentiable at u:

$$0 \leq J(u_k) - J(u) = \langle \nabla J(u), u_k - u \rangle + \| u_k - u \| \eta_k$$

u local min.

$$= \frac{\| u_k - u \|}{\| w \|} \left[\langle \nabla J(u), w \rangle + \| w \| \left(\langle \nabla J(u), \eta_k \rangle + \varepsilon_k \right) \right].$$

- i^{th} constraint is $\begin{cases} \underline{\text{active}} \text{ at } u \text{ if } \phi_i(u) = 0 \\ \underline{\text{inactive}} \text{ at } u \text{ if } \phi_i(u) < 0 \end{cases}$

- Set $I(u) := \{ \text{indices that are active at } u \}$

<u>Def</u>: Assume ϕ_i is differentiable $\forall i \in I(u)$. Then the constraints are <u>qualified</u> at u if

- the constraints are all <u>affine</u> (i.e., ϕ_i is <u>affine</u> $\forall i \in I(u)$) and $\exists w \in V$, $w \neq 0$, s.t. $\langle \nabla \phi_i(w), w \rangle \leq 0 \quad \forall i \in I(u)$.

- the ϕ_i are <u>not affine</u> and $\exists w \in V$, $w \neq 0$, s.t. $\langle \nabla \phi_i(w), w \rangle < 0$.

Def: $C(u) := \{ \text{feasible solutions} \}$

$$\underbrace{C^*(u)}_{\text{intersection of half-spaces for all active constraints.}} := \{ w \in V : \langle \nabla \phi_i(u), w \rangle \leq 0 \quad \forall\, i \in I(u) \}$$

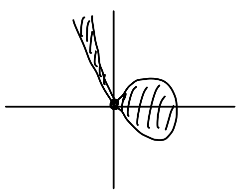
LECTURE 19

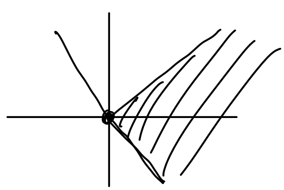
Goal: Find $\begin{cases} \min J(u) \\ \text{subject to } \phi_i(u) \leq 0, \quad i = 1, \dots, m, \end{cases}$

$J : \Omega \to \mathbb{R}$, $\Omega \subset \mathbb{R}^n$ open and nonempty,

$U := \{ u \in \Omega : \phi_i(u) \leq 0, \quad i = 1, \dots, m \}$.

Cone of Feasible Directions: $C(u)$





$\nabla \phi_i(u) = 0$

Def: $C^*(u) := \{ v : \langle \nabla \phi_i(u), v \rangle \leq 0 \quad \forall i \in I(u) \}$

= intersection of $\leq m$ half-spaces

= a convex cone

QUALIFIED CONSTRAINTS: For active, nonaffine constraints, $\exists\, w \neq 0$ s.t. $\langle \nabla \phi_i(u), w \rangle < 0$. The $C^*(u)$ has interior.

Note: If constraints are qualified at u then $C(u) = C^*(u)$

intersection of half-spaces

$\Rightarrow$ KKT $\qquad \min J(u)$

$\qquad\qquad$ s.t. $\phi_i(u) \leq 0$, $i = 1, \ldots, m$.

1) Seek all pts on bdry where constraints not qualified
2) Seek all solutions of KKT

Test for loc. min

Thm: If J is differentiable at u and if u is a local minimum w.r.t. U, then

$$\langle \nabla J(u), v \rangle \geq 0 \quad \forall v \in C(u)$$

If constraints are qualified, $C(u) = C^*(u)$, i.e.,

$$\langle \nabla J(u), v \rangle \geq 0 \quad \forall v \in C^*(u).$$

$\mathrm{cone}(u, v, w) := \{ \lambda_1 u + \lambda_2 v + \lambda_3 w : \lambda_1, \lambda_2, \lambda_3 \geq 0 \}$

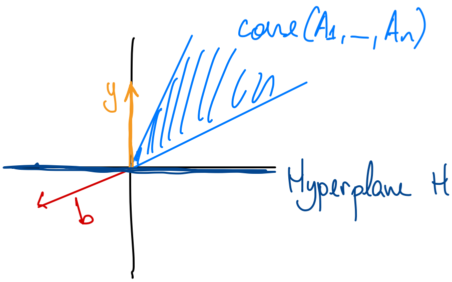
FARKAS LEMMA: Let $A \in M_{m\times n}(\mathbb{R})$ and $b \in \mathbb{R}^m$, then $\nexists\, x,\ x \geq 0$ so that

$$Ax = b \iff \exists\, y \text{ s.t. } y^T A \geq 0 \text{ and } y^T b < 0$$

Note: For $x \geq 0$, $Ax = b \Rightarrow b = \sum_{i=1}^{n} x_i A_i \quad x_i \geq 0$ $\quad \leftarrow (A_i)_{i=1}^{n}$ = cols. of A

$$\Rightarrow b \in \text{cone}(A_1, \ldots, A_n)$$

$$b \notin \text{cone}(A_1, \ldots, A_n) \iff \exists\, y \in \mathbb{R}^m \text{ s.t. } \langle y, A_i \rangle \geq 0 \text{ and } \langle y, b \rangle < 0$$



$b \notin H$

b is on one side and cone$(A_1, \ldots, A_n)$ is on the other.

COUNTERPOSITIVE FARKAS LEMMA: If

$$\langle v, A_i \rangle \geq 0 \quad \forall i \Rightarrow \langle v, b \rangle \geq 0$$

then $b \in \text{cone}(A_1, \ldots, A_n)$, i.e.,

$$b = \sum_{i=1}^{n} \lambda_i A_i, \quad \lambda_i \geq 0.$$

Thm: (14.5) $\phi_i : \Omega \to \mathbb{R}$ m constraints

$J : \Omega \to \mathbb{R}$

$U = \{ \phi_i \leq 0 \quad i = 1, \dots, m \}$

Assume local min. of J w.r.t. U is at u and let $I(u)$ be the active constraint indices.

Assume J is differentiable at u,

Assume ϕ_i are differentiable at u for $i \in I(u)$,

Assume ϕ_i are continuous at u for $i \notin I(u)$.

Assume constraints are qualified at u.

Then: $\exists \lambda_i \geq 0$ such that

$$\nabla J(u) = \sum_{i \in I(u)} \lambda_i \left(-\nabla \phi_i (u) \right)$$

i.e.,

$$\boxed{\nabla J(u) + \sum_{i \in I(u)} \lambda_i \nabla \phi_i (u) = \vec{0}}$$ KKT CONDITIONS

LECTURE 20

ALTERNATIVE FORMS OF KKT

Standard KKT Conditions: $\exists \{\lambda_i\}_{i\in I(u)}$, $\lambda_i \geq 0$, s.t.

$$\nabla J(u) + \sum_{i\in I(u)} \lambda_i \nabla \phi_i(u) = \vec{0}$$

$$\Longleftrightarrow \quad \underbrace{\nabla J(u) + \sum_{i=1}^{m} \lambda_i \nabla\phi_i(u) = 0}_{\text{"1st KKT condition"}} \quad \text{and} \quad \underbrace{\sum_{i=1}^{m} \lambda_i \phi_i = 0}_{\text{"2nd KKT condition"}}$$

Doesn't need to know the active indices

$$\Longleftrightarrow \quad \nabla J(u) + \sum_{i=1}^{m} \lambda_i \nabla\phi_i(u) = \vec{0}, \quad \lambda_i \geq 0,$$

and $\lambda_i \phi_i(u) = 0$ for $i = 1, \ldots, m$.

SLATER'S CONDITION (alternative to qualified constraints)

Book has many mistakes here!

- Easier to check than qualified const..

(Recall)

Def: If ϕ_i are differentiable at u $\forall i \in I(u)$ then the constraints are <u>qualified at u</u> if either

(1) all ϕ_i are <u>affine</u> $\forall i \in I(u)$ or

(2) $\exists w \in V, w \neq 0,$ s.t.

(i) $\langle \nabla \phi_i(u), w \rangle \leq 0 \quad \forall i \in I(u)$

(ii) If ϕ_i is not affine then $\langle \nabla \phi_i(u), w \rangle < 0$.

(ϕ affine $\Leftrightarrow \exists a \in \mathbb{R}^n, b \in \mathbb{R},$ s.t. $\phi(v) = a^T v + b$)

<u>Def</u>: (SLATER'S CONDITION) Let $U \subset \Omega \underset{\text{open}}{\subseteq} V$ (inn. prod. space) be

$U := \{ v : \phi_i(u) \leq 0 \quad \forall i = 1, \_, m \}$.

Assume the ϕ_i are all convex (in the line sense since we make no differentiability assumptions).

Then, either

(a) $U \neq \emptyset$ and all constraints ϕ_i are affine, or

(b) $\exists v \in U$ s.t.

(i) $\phi_i(v) \leq 0$

(ii) if ϕ_i is not affine then $\phi_i(v) < 0$.

<u>NOTE</u>: • Slater requires all ϕ_i to be convex (which reduces the # of problems...)

• Slater is not specific to one u only (it's true for whole constraint set)

• Slater doesn't require differentiability.

<u>Thm</u>: Let $\phi_i : \Omega \to \mathbb{R}$, $\Omega \subset \mathbb{R}^n$ open nonempty, ϕ_i are all convex for $i = 1, \dots, m$. Assume Ω is <u>convex</u>, and let $J : \Omega \to \mathbb{R}$,

$$U := \{ x : \phi_i(x) \le 0, \; i = 1, \dots, m \}.$$

Let $u \in U$ and say J, ϕ_i are all differentiable at u.

(1) If J has a local min. at u w.r.t. U and if Slater's condition holds, then KKT holds at u.

(2) If the restriction of J to U is convex, and if $\exists \lambda \in \mathbb{R}^m$, $\lambda \geq 0$, s.t. KKT holds at u, then J has a global min. at u w.r.t. U.

Pf: (1) Show Slater $\Rightarrow$ Qualified.

$\square$

SPECIAL CASE: ALL CONSTRAINTS ARE AFFINE

$Ax \leq b$, A $m \times n$, $b \in \mathbb{R}^m$

$$A = \begin{pmatrix} - a_1 - \\ \vdots \\ - a_m - \end{pmatrix}$$

$$\phi_i(x) = a_i x - b \Rightarrow \nabla \phi_i(u) = a_i \quad \forall u$$

$$U := \{ x : \phi_i(x) \leq 0 \} = \{ x : Ax \leq b \}$$

KKT: $$\nabla J(u) + \sum_{i=1}^{n} \lambda_i \nabla \phi_i(u) = 0$$

$$\nabla J(u) + \sum_{i=1}^{n} \lambda_i a_i = 0$$

$\Longleftrightarrow \exists \lambda \geq 0$ s.t.

$$\nabla J(u) + A^T \lambda = 0$$

PROBLEM: $\begin{cases} \text{Minimize} & \frac{1}{2} x^T P x + q^T x + r, \quad P \succeq 0 \\ \text{subject to} & Ax = b \end{cases}$

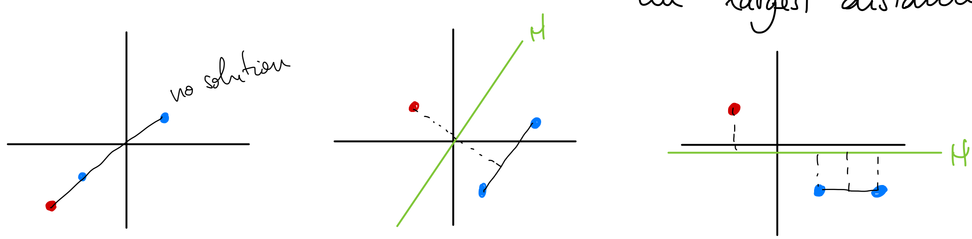
KKT conditions $\Rightarrow$ KKT matrix (HW5)

HARD MARGIN SUPPORT VECTOR MACHINE

$\begin{cases} p \text{ blue points } \{u_i\}_{i=1}^{p} \subset \mathbb{R}^n \\ q \text{ red points } \{v_j\}_{j=1}^{q} \subset \mathbb{R}^n \end{cases}$

Goal: Find a plane that separates them.

Goal: Find the plane that separates them with the largest distance.



Assume the pts. are all separable.

$\exists\, w \in \mathbb{R}^n$ and $b \in \mathbb{R}$ s.t.

open constraint set $\Rightarrow$ not friend

$\begin{cases} w^T u_i - b > 0 & i = 1, \dots, p \\ w^T v_j - b < 0 & j = 1, \dots, q \end{cases}$ $\rightsquigarrow$ Optimize for all w, b.

In general: $H = \{x : w^T x - b = 0\}$

$$\operatorname{dist}(u, H) = \frac{|w^T u - b|}{\|w\|}$$

Rescale w and b w/out changing H: use $\alpha \neq 0$, then

$$(\alpha w)^T x - \alpha b = 0 \iff w^T x - b = 0 .$$

$p+q$ inequalities $\longrightarrow$ smallest LHS for $i=1,\dots,p$
biggest LHS for $i=1,\dots,q$

$$\begin{cases} (w')^T u_i - b' \geq 1 & \forall\, i=1,\dots,p \\ (w')^T v_j - b' \leq -1 & \forall\, j=1,\dots,q \end{cases} \qquad (*)$$

$$\mathrm{dist}(u_i, H) = \frac{|(w')^T u_i - b'|}{\|w'\|} \geq \frac{1}{\|w'\|}$$

$$\mathrm{dist}(v_j, H) = \frac{|(w')^T v_j - b'|}{\|w'\|} \geq \frac{1}{\|w'\|} .$$

$$\delta := \min \{ \mathrm{dist}(u_i, H), \mathrm{dist}(v_j, H) \} \geq \frac{1}{\|w'\|} .$$

Goal: Find H that satisfies $(*)$ and maximizes δ

$\implies$ minimize $\|w'\|$ subject to $(*)$

Process: Given $\{u_i\}, \{v_j\}$, find w, b.

More precisely

$$\boxed{\begin{aligned} &\text{minimize} \quad \frac{1}{2}\|w\|^2 =: J(w) \\ &\text{subject to} \quad w^T u_i - b \geq 1 \qquad i = 1, \dots, p \\ &\qquad\qquad\quad -w^T v_j + b \geq 1 \qquad j = 1, \dots, q \end{aligned}}$$

<u>Note</u>: If can solve KKT then have optimal solution.

<u>EXAMPLE</u>: 2 blue pts. $u_1, u_2 \in \mathbb{R}^2 \rightsquigarrow u_1 = \begin{pmatrix} u_{11} \\ u_{12} \end{pmatrix}, u_2 = \begin{pmatrix} u_{21} \\ u_{22} \end{pmatrix}$

2 red pts. $v_1, v_2 \in \mathbb{R}^2 \rightsquigarrow v_1 = \begin{pmatrix} v_{11} \\ v_{12} \end{pmatrix}, v_2 = \begin{pmatrix} v_{21} \\ v_{22} \end{pmatrix}$

Constraints can be written as:

$$\underbrace{\begin{bmatrix} -u_{11} & -u_{12} & 1 \\ -u_{21} & -u_{22} & 1 \\ v_{11} & v_{12} & -1 \\ v_{21} & v_{22} & -1 \end{bmatrix}}_{=:C} \begin{bmatrix} w_1 \\ w_2 \\ b \end{bmatrix} \leq \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \end{bmatrix}$$

$$\Leftrightarrow C \begin{pmatrix} w_1 \\ w_2 \\ b \end{pmatrix} \leq -\vec{1}.$$

$$J(w, b) = \frac{1}{2}\|w\|^2 = \frac{1}{2}(w_1^2 + w_2^2)$$

$$\nabla J(w,b) + C^T \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad \text{KKT}$$

$$\begin{pmatrix} w_1 \\ w_2 \\ 0 \end{pmatrix} + \begin{pmatrix} & \cdots & \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \mu_1 \\ \mu_2 \end{pmatrix} = \vec{0}, \quad \begin{matrix} \lambda \succeq 0 \\ \mu \succeq 0 \end{matrix}$$

<u>KKT Eq 1</u>: $w = \lambda_1 u_1 + \lambda_2 u_2 - \mu_1 v_1 - \mu_2 v_2$

<u>KKT Eq 2</u>: $\lambda_1 + \lambda_2 - \mu_1 - \mu_2 = 0$

<u>Constraint</u>: $C \begin{pmatrix} w \\ b \end{pmatrix} \leq -\vec{1}$

$$w = X \begin{pmatrix} \lambda \\ \mu \end{pmatrix}, \quad X = \begin{pmatrix} u_{11} & u_{21} & v_{11} & v_{21} \\ u_{12} & u_{22} & v_{21} & v_{22} \end{pmatrix}$$

$$D \begin{pmatrix} \lambda \\ \mu \\ b \end{pmatrix} = \begin{pmatrix} w \\ b \end{pmatrix}, \quad D = \left(\begin{array}{c|c} X & \begin{matrix} 0 \\ 0 \end{matrix} \\ \hline 0 & 1 \end{array}\right)$$

LECTURE 21

Primal Problem vs Dual Problem

(ineq. constraints
min. problem)

PRIMAL PROBLEM: $J, \phi_i : \Omega \to \mathbb{R}$, $\Omega \subset \mathbb{R}^n$ open.

$$(PP) \begin{cases} \text{minimize } J(v) \\ \text{subject to } \phi_i(v) \leq 0, \quad i = 1, \dots, m \end{cases}$$

Lagrangian $\mathcal{L} : \Omega \times \mathbb{R}^m_+ \longrightarrow \mathbb{R}$

$$\mathcal{L}(v, \mu) = J(v) + \sum_{i=1}^{m} \mu_i \phi_i(v), \quad \mu_i \in \mathbb{R}^m_+ \ (\mu \geq 0)$$

Fix $\mu > 0$ and solve $\min_{v \in \Omega} \mathcal{L}(v, \mu) =: G(\mu)$

G is called the DUAL FUNCTION

Goal: Find v_μ s.t. $\mathcal{L}(v_\mu, \mu) = G(\mu)$.

Solve:

$$(DP) \begin{cases} \max_{\mu \geq 0} G(\mu) \end{cases} \qquad \text{(DUAL PROBLEM)}$$

Hope v_{μ^*} solves the primal problem. NOTE: G concave.

Saddle Points of Lagrangian

Def: $(u,\lambda) \in \Omega \times \mathbb{R}^m_+$ is a saddle point of the Lagrangian $\mathcal{L}$ iff

$$\sup_{\mu \in \mathbb{R}^m_+} \mathcal{L}(u,\mu) = \mathcal{L}(u,\lambda) = \inf_{v \in \Omega} \mathcal{L}(v,\lambda).$$

Thm: (Identify saddle pts of $\mathcal{L}$) If (u,λ) is a saddle pt. of $\mathcal{L}$ then

$$\sup_{\mu \in \mathbb{R}^m_+} \inf_{v \in \Omega} \mathcal{L}(v,\mu) = \mathcal{L}(u,\lambda) = \inf_{v \in \Omega} \sup_{\mu \in \mathbb{R}^m_+} \mathcal{L}(v,\mu).$$

Pf: $(\leq)$ Take $(w,\rho) \in \Omega \times \mathbb{R}^m_+$ s.t.

$$\inf_{v \in \Omega} \mathcal{L}(v,\rho) \leq \mathcal{L}(w,\rho) \leq \sup_{\mu \in \mathbb{R}^m_+} \mathcal{L}(w,\mu)$$

i.e., $$\underbrace{\inf_{v \in \Omega} \mathcal{L}(v,\rho)}_{\text{indep of } w} \leq \sup_{\mu \in \mathbb{R}^m_+} \mathcal{L}(w,\mu)$$

i.e., $$\inf_{v \in \Omega} \mathcal{L}(v,\rho) \leq \inf_{w \in \Omega} \sup_{\mu \in \mathbb{R}^m_+} \mathcal{L}(w,\mu)$$

$\sup_{\rho \in \mathbb{R}}$ " " $\leq$ " "

($\geq$) Use that (u, λ) is saddle pt.

$\blacksquare$

> **Thm:** (14.15)
>
> (1) If (u, λ) is saddle pt. of $\mathcal{L}$ then u is a solution of (PP) and $J(u) = \mathcal{L}(u, \lambda)$
>
> (2) If J, ϕ_i are convex, Slater's condition is satisfied, J differentiable at u, and u solves (PP), then $\exists \lambda \in \mathbb{R}^m_+$ s.t. (u, λ) is a saddle pt. of $\mathcal{L}$.

Pf: (1) Assume (u, λ) is saddle pt.

- Show u is admissible $u \in U$
- show $J(u) = \mathcal{L}(u, \lambda)$
- $J(u) \leq J(v) \quad \forall v \in U$

$\Rightarrow u$ solves (PP)

$$\mathcal{L}(u, \mu) \leq \mathcal{L}(u, \lambda) \quad \forall \lambda \in \mathbb{R}^m_+$$

i.e.,
$$\cancel{J(u)} + \sum_i \mu_i \phi_i(u) \leq \cancel{J(u)} + \sum_i \lambda_i \phi_i(u)$$

i.e., $$\sum_i \mu_i \phi_i(u) \leq \sum_i \lambda_i \phi_i(u)$$

$\Rightarrow \quad \phi_i(u) \leq 0 \quad \forall i$ (could contradict this if $\phi_j > 0$ for some j by making μ_j very big)

$\Rightarrow \quad u \in U$ admissible

If $\mu = 0$, $$0 \leq \sum_i \lambda_i \phi_i(u) \leq 0$$

i.e., $$\sum_i \lambda_i \phi_i(u) = 0$$

$$\Rightarrow \quad J(u) = \mathcal{L}(u, \lambda).$$

Now

saddle

$$J(u) = \mathcal{L}(u, \lambda) \leq \mathcal{L}(v, \lambda) \quad \forall v \in U$$
$$= J(v) + \underbrace{\sum \lambda_i \phi_i(v)}_{\leq 0 \text{ if } v \text{ admissible}} \leq J(v) \quad \forall v.$$

(2) Show these $\Rightarrow$ KKT and then show KKT $\Rightarrow \exists \lambda \in \mathbb{R}^m_+ \ldots$

■

NOTE: G is concave b/c

$$G((1-\theta)\hat{\mu}+\theta\tilde{\mu}) \overset{\text{def}}{=} \inf_{v\in\Omega} \mathcal{L}(v) + \sum_i \left[(1-\theta)\hat{\mu}_i + \theta\tilde{\mu}_i\right]\phi_i(v)$$

$$= \inf_{v\in\Omega}\left[(1-\theta)\,\mathcal{L}(v,\hat{\mu}) + \theta\,\mathcal{L}(v,\tilde{\mu})\right]$$

$$\geq \left[\inf_{v\in\Omega}(1-\theta)\,\mathcal{L}(v,\hat{\mu})\right] + \left[\inf_{v\in\Omega}\theta\,\mathcal{L}(v,\tilde{\mu})\right]$$

$$\geq (1-\theta)\,G(\hat{\mu}) + \theta\,G(\tilde{\mu}).$$

//

EXAMPLE: (Linear Program)

$$\begin{cases} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \\ & -x \leq 0 \end{cases}$$

$x \in \mathbb{R}^n$
$c \in \mathbb{R}^n$
$A \in M_{m\times n}(\mathbb{R})$
$b \in \mathbb{R}^m$

Lagrangian for this is:

$$\mathcal{L}(x,\mu,\nu) = c^T x + \mu^T(Ax-b) - \nu^T x, \quad \mu \in \mathbb{R}^m_+, \ \nu \in \mathbb{R}^n_+$$

$$G(\mu,\nu) = \inf_{x\in\mathbb{R}^n} (c + A^T\mu - \nu)^T x - b^T\mu.$$

$$= \begin{cases} -\infty & \text{if } \quad c + A^T\mu - \nu \neq 0 \\ -b^T\mu & \text{if } \quad c + A^T\mu - \nu = 0 \end{cases}$$

(DP) $\quad \max\limits_{\substack{\mu \in \mathbb{R}^m_+ \\ \nu \in \mathbb{R}^n_+}} G(\mu, \nu)$

$\max\limits_{\mu \in \mathbb{R}^m_+} -b^T\mu \quad ?$

$G(\mu,\nu) = -b^T\mu$

$\iff c + A^T\mu - \nu = 0$

$\iff c + A^T\mu \geq 0$

Can rewrite (DP) as:

$$\begin{cases} \max\limits_{\mu \in \mathbb{R}^m_+} -b^T\mu \\ \text{subject to } \quad A^T\mu + c \geq 0 \\ \qquad\qquad\qquad \mu \geq 0 \end{cases}$$

"Hidden" Constraint