2. Preliminaries

Def 2.1 Singular values

$$s_1(A) \geq s_2(A) \geq \cdots \geq s_n(A) \geq 0$$

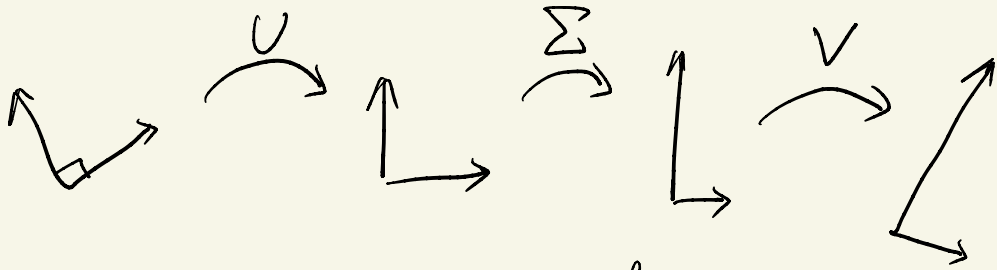
all square roots of eigenvalues of $A^T A$

Observation:

$s_{max}(A)$ and $s_{min}(A)$ are the smallest $M \in \mathbb{R}$ and largest $m \in \mathbb{R}$ s.t.

$$m\|x\|_2 \leq \|Ax\|_2 \leq M\|x\|_2 \quad \forall x \in \mathbb{R}^n$$

$A = V^* \Sigma U$



orthogonal preserve l_2 norm but not l_p norm

- For flat matrices A, $\underline{S_{min}(A) = 0}$

- geometric interpretation:

Extreme singular values control the distance of the Euc geo with the action of A

Def 2.2 (Spectral Norm)

(operator norm)

$$\|A\| = \|A\|_{\ell_2^n \to \ell_2^n} = \sup_{x \in \mathbb{R}^n \setminus \{0\}} \frac{\|Ax\|_2}{\|x\|_2}$$

$$= \sup_{x \in S^{n-1}} \|Ax\|_2$$

The $s_{max}(A) = \|A\|$,

$$s_{min}(A) = \frac{1}{\|A^\dagger\|}$$

if A is injective

pseudoinverse

Note: $\|A\| = \|s\|_\infty$ where $s = (s_1, s_2, \cdots, s_n)$

Def 2.3 (schalten norm)

Let $A \in \mathbb{R}^{N \times n}$ or $\mathbb{C}^{N \times n}$

with singular values $(s_1, \cdots, s_n) = s$

then for $1 \leq p \leq \infty$ the schalten-p norm

is defined as

$$\|A\|_{sp} := \|s\|_p$$

the Schalten 2 norm is $\| \ \|_F$

2.2 Nets:

Recall:

$$s_{max}(A) = \|A\| = \sup_{x \in S^{n-1}} \|Ax\|_2$$

supermum over infinitely many x

To analyze the distribution of $s_{max}(A)$ for a random matrix A, we need to discretie this expression

Def 2.4 (covering, ε-net)

...

compactness (cover $\to$ finite cover)

X compact $\Leftrightarrow$ $N(x, \varepsilon)$ finite

Lemma 2.5 (covering numbers of the sphere)

the sphere is not the ball!

Consider the vector space $\mathbb{R}^n$ equipped with the norm $\|\cdot\|$ and let

← arbitrary norm

$S = \{x \in \mathbb{R}^n : \|x\| = 1\}$ be the associated with sphere

Then for every $\varepsilon > 0$ one has

$$\mathcal{N}(S, \varepsilon) \leq \left(1 + \frac{2}{\varepsilon}\right)^n$$

(the shape of sphere and cover change simtaneously)

Fix $\varepsilon > 0$ and choose $\mathcal{N}_\varepsilon$ to the a maximal ε-separated subset of S, that is, $\mathcal{N}_\varepsilon$ is such that $\|x - y\| \geq \varepsilon$ for all $x, y \in \mathcal{N}_\varepsilon$, $x \neq y$, no subset S containing $\mathcal{N}_\varepsilon$ has this property.

(can construct by instantly adding points)

No additional point can be added with that property

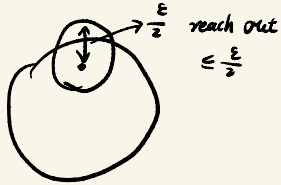
$\Rightarrow$ every point in S has distance $< \varepsilon$ to the nearest point in N_ε

N_ε is ε-net

claims: balls of radius $\frac{\varepsilon}{2}$ centered at the points in N_ε are disjoint.

All such balls lie in $(1+\frac{\varepsilon}{2})B$

where $B = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$



comparing volumes:

$|N_\varepsilon|$ balls of radius $\frac{\varepsilon}{2}$ are contained

in one ball of radius $(1+\frac{\varepsilon}{2})$

$$|N_\varepsilon| \, \mathrm{Vol}(\tfrac{\varepsilon}{2}B) \le \mathrm{vol}\left((1+\tfrac{\varepsilon}{2})B\right)$$

Using $\mathrm{vol}(rB) = r^n \, \mathrm{vol}(B)$

$$\Rightarrow |N_\varepsilon| \left(\frac{\varepsilon}{2}\right)^n \mathrm{vol}(B) \le \left(1+\frac{\varepsilon}{2}\right)^n \mathrm{vol}(B)$$

Goal: Nets can help estimating spectral norms

Prob: spectral norms are defined as a supremum over an infinite set of points

→ How can we make sure we do not miss the mini / maximizer?

Idea: Estimate action of A on ball points in $\mathcal{N}_\varepsilon$ guarantee to all points on the sphere is a perturbation argument

Lemma 7.6

Let $A \in \mathbb{R}^{N \times n}$ and let $\mathcal{N}_\varepsilon$ are finite ε-net of S^{n-1} for some $\varepsilon \in (0, 1)$.

Then

$$\max_{x \in N_\varepsilon} \|Ax\|_2 \le \|A\| \le (1-\varepsilon)^{-1} \max_{x \in N_\varepsilon} \|Ax\|_2$$

pf: Lower bound follows from def

upper bound:

By computation, there exists

$x_0 \in S^{n-1}$ s.t. $\|Ax_0\|_2 = \|A\|$

Choose $y \in N_\varepsilon$ with an appro

x_0 an $\|x_0 - y\|_2 \le \varepsilon$

Lemma 2.7 Let A be a symmetric $n \times n$ matrix and let $\mathcal{N}_\varepsilon$ be an ε-net of S^{n-1} for some $\varepsilon \in (0,1)$

Then $\|A\| = \sup_{x \in S^{n-1}} |\langle Ax, x \rangle|$

$$\leq (1-2\varepsilon)^{-1} \sup_{x \in \mathcal{N}_\varepsilon} |\langle Ax, x \rangle|$$

$A = UDU^T$,

U: ortho

$|\langle Ax, x \rangle| = |\langle UDU^T x, x \rangle|$

$= |\langle DU^T x, U^T x \rangle|$

$\downarrow$

$|\langle Dx, x \rangle|$