Domácí práce na LAF II.

Michal Grňo

16. 6. 2019

1 Cvičení 5 – rozšiřující úloha 1

1.1 Zadání

Nechť $A \in \mathbb{C}^{m \times n}$. Číslo $\|A\|_2 := \max_{x \in \mathbb{R}^n} \frac{\|Ax\|}{\|x\|}$ se nazývá *spektrální norma* matice A. Normy ve zlomku na pravé straně definice jsou normy indukované standardním skalárním součinem na $\mathbb{C}^m$ a $\mathbb{C}^n$. Ukažte, že

1. spektrální norma matice A je rovna její nejvyšší singulární hodnotě σ_1.
2. spektrální norma je norma na prostoru $\mathbb{C}^{m \times n}$, ale není indukovaná skalárním součinem.
3. tzv. *Frobeniova norma* na $\mathbb{C}^{m \times n}$ indukovaná skalárním součinem $\langle A, B \rangle := \mathrm{Tr}\ (A^+ B)$ se pomocí singulárních hodnot vyjádří jako $\|A\|_F = \sqrt{{\sigma_1}^2 + ... + {\sigma_r}^2}$.
4. Frobeniova i spektrální norma splňují nerovnost $\|AB\| \le \|A\| \, \|B\|$ pro libovolné matice $A \in \mathbb{C}^{m \times n}$, $B \in \mathbb{C}^{n \times p}$.

1.2 Řešení

1.2.1 $\|A\|_2 = \sigma_1$

Označme si singulární rozklad $A = U \Sigma V^+$ a normovaný vektor $\widehat{x} = \frac{x}{\|x\|}$. Pro pohodlnost si ještě zavedeme označení $\widehat{\mathbb{R}}^n = \left\{ x \in \mathbb{R}^n \;\middle|\; \|x\| = 1 \right\}$. Definici spektrální normy můžeme rozepsat jako:

$$\|A\|_2 = \max_{x \in \mathbb{R}^n} \frac{\|Ax\|}{\|x\|} = \max_{x \in \mathbb{R}^n} \frac{\|U \Sigma V^+ x\|}{\|x\|} = \max_{\widehat{x} \in \widehat{\mathbb{R}}^n} \frac{\left\| U \Sigma V^+ \, \|x\| \, \widehat{x} \right\|}{\|x\|} = \max_{\widehat{x} \in \widehat{\mathbb{R}}^n} \left\| U \Sigma V^+ \widehat{x} \right\|.$$

Nyní využijeme toho, že matice U a V jsou unitární, platí tedy $\|Ux\| = \|x\|$ a $\widehat{\mathbb{R}}^n \xrightarrow{V} \widehat{\mathbb{R}}^n$.

$$\|A\|_2 = \max_{\widehat{x} \in \widehat{\mathbb{R}}^n} \left\| U \Sigma V^+ \widehat{x} \right\| = \max_{\widehat{x} \in \widehat{\mathbb{R}}^n} \|U \Sigma \widehat{x}\| = \max_{\widehat{x} \in \widehat{\mathbb{R}}^n} \|\Sigma \widehat{x}\|$$

$$= \max_{\widehat{x} \in \widehat{\mathbb{R}}^r} \sqrt{\sigma_1 \widehat{x}_1 + ... \sigma_r \widehat{x}_r} = \max_{i \in \{1, \dots, r\}} \sigma_i = \sigma_1.$$

1.2.2 $\|A\|_2$ je norma

Nyní dokážeme, že $\|A\|_2$ je norma. Norma na $\mathbb{C}^{m \times n}$ musí splňovat tyto tři podmínky:

1. $\|A\| = 0 \iff A = 0$
2. $\|aA\| = |a| \; \|A\| \quad \forall a \in \mathbb{C}$
3. $\|A + B\| \le \|A\| + \|B\|$

První podmínku lze rozepsat jako

$$\|A\|_2 = 0 \iff \|\Sigma\|_2 = 0 \iff \Sigma = 0 \iff A = 0.$$

Protože je Σ pouze vyjádřením A vůči jiným ON bázím, je zřejmé, že musí ekvivalence platit.

V důkazu druhé podmínky použijeme vyjádření komplexního čísla v polárních souřadnicích $a = |a| \exp \arg a$. Singulární rozklad škálované matice označíme $aA = U'\Sigma'V'^{+}$. Dostaneme:

$$aA = |a| \, (\exp \arg a) \; U\Sigma V^{+} = U'\Sigma'V'^{+}$$

Aby rovnost platila, musíme členy $|a|$ a $\exp \arg a$ „rozdělit“ mezi matice U', Σ' a V'^{+}. Je zřejmé, že unitární matice U, V nemůžeme škálovat faktorem $|a|$, oproti tomu matice Σ musí zůstat reálná kladná, proto ji nemůžeme násobit komplexní jednotkou $\exp \arg a$. Bude tedy platit:

$$U' = bU, \quad V' = cV, \quad \text{kde } b\overline{c} = \exp \arg a,$$

$$\Sigma' = |a|\Sigma \implies \sigma_1' = |a|\sigma_1 \implies \|aA\|_2 = |a| \; \|A\|_2 \, .$$

A nakonec důkaz třetí podmínky:

$$\begin{aligned} \|A + B\|_2 &= \max_{\widehat{x} \in \widehat{R}^n} \; \|(A + B)\widehat{x}\| \\ &= \max_{\widehat{x} \in \widehat{R}^n} \; \|A\widehat{x} + B\widehat{x}\| \\ &\leq \max_{\widehat{x} \in \widehat{R}^n} \; (\|A\widehat{x}\| + \|B\widehat{x}\|) \\ &\leq \max_{\widehat{x} \in \widehat{R}^n} \|A\widehat{x}\| + \max_{\widehat{x} \in \widehat{R}^n} \|B\widehat{x}\| \\ &= \|A\|_2 + \|B\|_2 \, . \end{aligned}$$

1.2.3 $\|A\|_2$ není norma indukovaná skalárním součinem

Mějme skalární součin $\langle \cdot, \cdot \rangle$ obecného vektorového prostoru, jím indukovaná norma je definována jako $\|x\|^2 = \langle x, x \rangle$. Vybereme-li dva libovolné vektory x, y z tohoto prostoru, zřejmě pro ně musí platit

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + \langle x, y \rangle + \langle y, x \rangle,$$

$$\|x - y\|^2 = \|x\|^2 + \|y\|^2 - \langle x, y \rangle - \langle y, x \rangle.$$

Sečtením těchto rovnic získáme tzv. *rovnoběžníkový zákon* – nutnou podmínku pro existenci skalárního součinu, který by indukoval normu $\|\cdot\|$:

$$\|x + y\|^2 + \|x - y\|^2 = 2\,\|x\|^2 + 2\,\|y\|^2 \, .$$

My chceme dokázat, že neexistuje skalární součin, který by indukoval spektrální normu na $\mathbb{C}^{m \times n}$. Budeme předpokládát, že $m, n \geq 2$, pro menší matice provedeme diskusi na konci. Protože musí rovnoběžníkový zákon v případě indukované normy platit pro všechny prvky $\mathbb{C}^{m \times n}$, stačí nám ukázat, že neplatí na 4-rozměrném podprostoru $\mathbb{C}^{2 \times 2}$. Prvky, pro které budeme neplatnost dokazovat, budou blokové matice (spodní indexy značí rozměry bloků):

$$\widetilde{A} = \begin{pmatrix} A_{2\times 2} & 0_{2\times(n-2)} \\ 0_{(m-2)\times 2} & 0_{(m-2)\times(n-2)} \end{pmatrix}, \qquad \widetilde{B} = \begin{pmatrix} B_{2\times 2} & 0_{2\times(n-2)} \\ 0_{(m-2)\times 2} & 0_{(m-2)\times(n-2)} \end{pmatrix}.$$

Zde se nám bude hodit mezivýsledek procvičovací úlohy 3:

Mějme matici $A \in \mathbb{C}^{m\times n}$ a blokovou matici $\widetilde{A} \in \mathbb{C}^{(m+p)\times(n+q)}$ takovou, že:

$$\widetilde{A} = \begin{pmatrix} A_{m\times n} & 0_{m\times q} \\ 0_{p\times n} & 0_{p\times q} \end{pmatrix}.$$

Označíme-li si nyní jejich singulární rozklady $A = U\Sigma V^+$, $\widetilde{A} = \widetilde{U}\widetilde{\Sigma}\widetilde{V}^+$, bude platit:

$$\widetilde{\Sigma} = \begin{pmatrix} \Sigma_{m\times n} & 0_{m\times q} \\ 0_{p\times n} & 0_{p\times q} \end{pmatrix}.$$

Z toho je zřejmé, že $\|\widetilde{A}\|_2 = \|A\|_2$, stačí tedy najít matice 2×2, které nesplňují rovnoběžníkový zákon, a dokážeme tím jeho neplatnost pro $\mathbb{C}^{m\times n}$. Takovými maticemi jsou například:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

$$\|A\|_2 = \left\| \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\|_2 = 1,$$

$$\|B\|_2 = \left\| \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\|_2 = 1,$$

$$\|A+B\|_2 = \left\| \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right\|_2 = 2,$$

$$\|A-B\|_2 = \left\| \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \right\|_2 = 2,$$

$$\|A+B\|_2{}^2 + \|A-B\|_2{}^2 = 2^2 + 2^2 = 8,$$

$$2\|A\|_2{}^2 + 2\|A\|_2{}^2 = 2\cdot 1^2 + 2\cdot 1^2 = 4,$$

$$8 \neq 4 \implies \lightning.$$

Dokázali jsme, že $\|\cdot\|_2$ není norma indukovaná skalárním součinem na $\mathbb{C}^{m\times n}$, kde $m, n \geq 2$. Vraťme se ještě nakrátko k prostorům menší dimenze. V $\mathbb{C}^{1\times 1} = \mathbb{C}$ je singulární rozklad

$$z = \underbrace{1}_{U}\,\underbrace{|z|}_{\Sigma}\,\underbrace{\exp\arg z}_{V^+}.$$

Zřejmě tedy $\|z\|_2 = |z|$, což je norma indukovaná skalárním součinem $\langle a, b\rangle = \bar{a}b$. Pro prostor sloupcových vektorů $\mathbb{C}^{m\times 1} = \mathbb{C}^m$ lze ukázat, že jejich jediné singulární číslo je rovno normě indukované standardním skalárním součinem, obdobně je tomu u řádkových vektorů $\mathbb{C}^{1\times m}$.

1.2.4 Vztah $\|A\|_F$ a singulárních čísel

Frobeniova norma je definována jako $\|A\|_F = \sqrt{\operatorname{Tr} A^+A}$. Z teorie víme, že matice A^+A je hermitovská, tedy i ortogonálně diagonalizovatelná:

$$A^+A = V\operatorname{diag}(\lambda_1, \lambda_2, \ldots)V^+.$$

Víme, že transformace typu $R^{-1}AR$ zachovávají charakteristický polynom matice A, a tedy i její stopu, bude proto

$$\text{Tr } A^+A = \lambda_1 + \lambda_2 + ...$$

Vlastní čísla matice A^+A jsou ovšem čtverce singulárních hodnot matice A, platí tedy vztah:

$$\|A\|_F = \sqrt{\text{Tr } A^+A} = \sqrt{\lambda_1 + \lambda_2 + ...} = \sqrt{{\sigma_1}^2 + {\sigma_2}^2 + ...}$$

1.2.5 Submultiplikativita norem

Začneme Frobeniovou normou:

$$\begin{aligned}
{\|AB\|_F}^2 &= \text{Tr } ((AB)^+AB) \\
&= \text{Tr } (B^+A^+AB) \\
&= \text{Tr } (A^+A\ BB^+) \\
&\leq \text{Tr } (A^+A) \text{ Tr } (BB^+) \\
&= \text{Tr } (A^+A) \text{ Tr } (BB^+) \\
&= {\|A\|_F}^2 \ {\|B\|_F}^2,
\end{aligned}$$

kde na 3. a 5. řádku byla využita vlastnost $\text{Tr } (AB...C) = \text{Tr } (B...CA)$ a na 4. řádku $\text{Tr } (AB) \leq \text{Tr } A \text{ Tr } B$ pro pozitivně semidefinitní A, B (vyplývá z Cauchy-Schwarzovy nerovnosti). Protože je norma nezáporná, platí po odmocnění rovnice

$$\|AB\|_F \leq \|A\|_F \|B\|_F .$$

Z definice spektrální normy vyplývá:

$$\begin{aligned}
&\|A\|_2 = \max_{x \in \mathbb{R}^n} \frac{\|Ax\|}{\|x\|} \\
\implies\ &\|A\|_2 \geq \frac{\|Ax\|}{\|x\|} \quad \forall x \in \mathbb{R}^n \\
\implies\ &\|A\|_2 \|x\| \geq \|Ax\| \quad \forall x \in \mathbb{R}^n \setminus \{0\}.
\end{aligned}$$

S použitím této znalosti můžeme vyjádřit:

$$\|ABx\| = \|A(Bx)\| \leq \|A\|_2 \|Bx\| \leq \|A\|_2 \|B\|_2 \|x\| \quad \forall x \in \mathbb{R}^n \setminus \{0\}$$

$$\frac{\|ABx\|}{\|x\|} \leq \|A\|_2 \|B\|_2 \quad \forall x \in \mathbb{R}^n \setminus \{0\}$$

Vybereme-li nyní takové $x_0 \in \mathbb{R}^n \setminus \{0\}$, aby pro něj byl výraz na levé straně maximální, bude platit:

$$\|AB\|_2 = \max_{x \in \mathbb{R}^n} \frac{\|ABx\|}{\|x\|} = \frac{\|ABx_0\|}{\|x_0\|} \leq \|A\|_2 \|B\|_2 .$$