Hodnosť matice AA^T

Úloha 1. Dokážte, že $h(A^T A) = h(A)$ pre ľubovoľnú maticu typu $n \times n$ nad $\mathbb{R}$.

Budem sa snažiť napísať riešenia so všetkými detailami – určite to tu píšem podstatne detailnejšie, než som to očakával od Vás na písomke. (Napríklad ak by ste sa na skúškovej písomke odvolali v niektorej časti na bonusovú úlohu z poslednej písomky na cvičeniach, bral by som to ako dostatočné zdôvodnenie.)

Poznámka 1. Vieme, že $h(A) = h(A^T)$. Je úplne jedno, či budeme dokazovať $h(A^T A) = h(A)$, alebo $h(A) = h(AA^T)$. Stačí si všimnúť, že ak dokážeme jedno z týchto tvrdení a použijeme ho pre maticu A^T (namiesto A), tak dostaneme druhé z nich.

Poznámka 2. Nerovnosť $h(A^T A) \leq h(A)$ vyplýva z všeobecne platnej nerovnosti $h(BA) \leq h(A)$. Túto úlohu ste dostali ako jeden z bonusových príkladoch na cvičeniach. Môžete sa pozrieť aj na to, ktoré kroky v jednotlivých dôkazoch pripomínajú niečo, čo sa dá použiť v tejto úlohe. Pozri aj lemu 3, ku ktorej som stručne naznačil dôkaz.

Toto tvrdenie neplatí pre matice nad poľom $\mathbb{C}$, ako ukazuje nasledujúci príklad. Čiže ak nájdeme nejaký dôkaz, určite musí využívať nejakú vlastnosť poľa $\mathbb{R}$, ktorú pole $\mathbb{C}$ nemá.

$$\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Budeme potrebovať tento výsledok:

Lema 1. *Nech $\vec{\alpha} \in \mathbb{R}^n$. Ak $\vec{\alpha}\vec{\alpha}^T = 0$, tak $\vec{\alpha} = 0$.*

Dôkaz. Nech $\vec{\alpha} = (a_1, \dots, a_n)$. Potom $\vec{\alpha}\vec{\alpha}^T = \sum_{k=1}^n a_k^2$. Ak platí

$$a_1^2 + a_2^2 + \dots + a_n^2 = 0,$$

znamená to, že sme dostali nulu ako súčet nezáporných reálnych čísel, teda každé z nich musí byť 0. Dostávame teda $a_1^2 = a_2^2 = \dots = a_k^2 = 0$, čo znamená, že $a_1 = a_2 = \dots = a_k = 0$ a $\vec{\alpha} = (0, 0, \dots, 0) = \vec{0}$. □

Prvý prístup

V tejto časti uvedieme dve riešenia, ktoré sú vlastne len preformulovaním toho istého riešenia v dvoch rôznych jazykoch. Podľa vášho vkusu si môžete vybrať, či sa vám viac hodí pozerať sa na úlohu cez homogénne sústavy alebo cez jadrá lineárnych zobrazení.

Riešenie úlohy 1. Maticiam A a AA^T zodpovedajú lineárne zobrazenia $f_A, f_{AA^T} : \mathbb{R}^n \to \mathbb{R}^n$.

Pre jadro zobrazenia $f \colon V \to W$ platí $d(\operatorname{Ker} f) = d(W) - d(\operatorname{Im} f)$, pričom $d(\operatorname{Im} f_A) = h(A)$. (Rozmyslite si prečo – malo by to ísť pomerne ľahko priamo z definície matice zobrazenia a z definície hodnosti.) V našom prípade máme teda $d(\operatorname{Ker} f_A) = n - h(A)$ a $d(\operatorname{Ker} f_{AA^T}) = n - h(AA^T)$.

Ukážeme, že $\operatorname{Ker} f_A = \operatorname{Ker} f_{AA^T}$, t.j. $\vec{\alpha} \in \operatorname{Ker} f_A \Leftrightarrow \vec{\alpha} \in \operatorname{Ker} f_{AA^T}$.

$\boxed{\Rightarrow}$ $\vec{\alpha} \in \operatorname{Ker} f_A \Rightarrow \vec{\alpha}A = \vec{0} \Rightarrow \vec{\alpha}AA^T = \vec{0}A^T = \vec{0} \Rightarrow \vec{\alpha} \in \operatorname{Ker} f_{AA^T}$.

$\boxed{\Leftarrow}$ $\vec{\alpha} \in \operatorname{Ker} f_{AA^T} \Rightarrow \vec{\alpha}AA^T = \vec{0} \Rightarrow \vec{\alpha}AA^T\vec{\alpha}^T = \vec{0}\vec{\alpha}^T = 0 \Rightarrow (\vec{\alpha}A)(\vec{\alpha}A)^T = 0 \Rightarrow$ (Lema 1) $\vec{\alpha}A = 0 \Rightarrow \alpha \in \operatorname{Ker} f_A$

Pretože $\operatorname{Ker} f_A = \operatorname{Ker} f_{AA^T}$, tieto podpriestory majú rovnakú dimenziu, teda

$$n - h(A) = d(\operatorname{Ker} f_A) = d(\operatorname{Ker} f_{AA^T}) = n - h(AA^T).$$

Z poslednej rovnosti vyplýva
$$h(A) = h(AA^T).$$

$\square$

Riešenie úlohy 1. Uvažujme homogénne systémy zodpovedajúce maticiam A a AA^T, podpriestory riešení označme S_A a S_{A^TA}, t.j.

$$S_A = \{\vec{\alpha} \in \mathbb{R}^n; A\vec{\alpha}^T = \vec{0}^T\}$$
$$S_{A^TA} = \{\vec{\alpha} \in \mathbb{R}^n; A^TA\vec{\alpha}^T = \vec{0}^T\}$$

Ukážeme, že $S_A = S_{A^TA}$, t.j. $\vec{\alpha} \in S_A \Leftrightarrow \vec{\alpha} \in S_{A^TA}$.

$\boxed{\Rightarrow}$ $\vec{\alpha} \in S_A \Rightarrow A\vec{\alpha}^T = \vec{0}^T \Rightarrow A^TA\vec{\alpha}^T = A^T\vec{0}^T = \vec{0}^T \Rightarrow \vec{\alpha} \in S_{A^TA}$

$\boxed{\Leftarrow}$ $\vec{\alpha} \in S_{A^TA} \Rightarrow A^TA\vec{\alpha}^T = \vec{0}^T \Rightarrow \vec{\alpha}A^TA\vec{\alpha}^T = \vec{\alpha}\vec{0}^T = 0 \Rightarrow \vec{\alpha}A^T(\vec{\alpha}A^T)^T = 0 \Rightarrow$ (podľa lemy 1) $\vec{\alpha}A^T = \vec{0}$ $\Rightarrow A\vec{\alpha}^T = \vec{0}^T \Rightarrow \vec{\alpha} \in S_{A^TA}$

Máme teda $S_A = S_{A^TA}$. Teda dimenzie týchto podpriestorov sú rovnaké. Dimenziu priestoru riešení homogénnej sústavy však vieme vyjadriť pomocou hodnosti matice, teda máme

$$n - h(A) = d(S_A) = d(S_{A^TA}) = n - h(A^TA),$$

z čoho vyplýva
$$h(A) = h(A^TA).$$

$\square$

Druhý prístup

Ďalšie riešenie, ktoré uvedieme, si človek ľahšie všimne, ak má prax v práci so skalárnym súčinom. (Všimnite si, že $\vec{\alpha}\vec{\beta}^T$ je presne skalárny súčin vektorov $\vec{\alpha}$ a $\vec{\beta}$ tak, ako ho poznáte zo strednej školy.) Na tejto prednáške sme s ním nepracovali, čiže skôr som očakával predchádzajúce riešenie pomocou sústav/jadier.

Opäť popíšeme v podstate to isté riešenie dvoma spôsobmi. V oboch prípadoch sa nám bude hodiť nasledujúce pozorovanie:

Lema 2. *Nech B je matica typu $k \times n$ nad $\mathbb{R}$. Ak $h(B) = k$, tak $h(BB^T) = k$. (Inak povedané: Ak sú riadky matice B lineárne nezávislé, tak sú nezávislé i riadky matice BB^T.)*

Dôkaz. Označme riadky matice B ako $\vec{\alpha}_1, \dots, \vec{\alpha}_k$, t.j.

$$B = \begin{pmatrix} \vec{\alpha}_1 \\ \vdots \\ \vec{\alpha}_k \end{pmatrix} \qquad B^T = \begin{pmatrix} \vec{\alpha}_1^T & \dots & \vec{\alpha}_k^T \end{pmatrix}.$$

Pre súčin týchto matíc potom dostaneme

$$BB^T = \begin{pmatrix} \vec{\alpha}_1\vec{\alpha}_1^T & \vec{\alpha}_1\vec{\alpha}_2^T & \dots & \vec{\alpha}_1\vec{\alpha}_k^T \\ \vec{\alpha}_2\vec{\alpha}_1^T & \vec{\alpha}_2\vec{\alpha}_2^T & \dots & \vec{\alpha}_2\vec{\alpha}_k^T \\ \dots & \dots & \dots & \dots \\ \vec{\alpha}_k\vec{\alpha}_1^T & \vec{\alpha}_k\vec{\alpha}_2^T & \dots & \vec{\alpha}_k\vec{\alpha}_k^T \end{pmatrix}$$

$\square$

O vektoroch $\vec{\alpha}_1, \dots, \vec{\alpha}_k$ vieme, že sú lineárne nezávislé. Pokúsme sa využiť tento fakt na dôkaz, že aj riadky matice BB^T sú lineárne nezávislé.

Ak by sme ako lineárnu kombináciu riadkov matice BB^T dostali nulový vektor, znamená to, že

$$\begin{gathered}
c_1\vec{\alpha}_1\vec{\alpha}_1^T + c_2\vec{\alpha}_2\vec{\alpha}_1^T + \cdots + c_k\vec{\alpha}_k\vec{\alpha}_1^T = 0\\
c_1\vec{\alpha}_1\vec{\alpha}_2^T + c_2\vec{\alpha}_2\vec{\alpha}_2^T + \cdots + c_k\vec{\alpha}_k\vec{\alpha}_2^T = 0\\
\dots\\
c_1\vec{\alpha}_1\vec{\alpha}_k^T + c_2\vec{\alpha}_2\vec{\alpha}_k^T + \cdots + c_k\vec{\alpha}_k\vec{\alpha}_k^T = 0
\end{gathered}$$

Vynásobme jednotlivé riadky postupne číslami $c_1, \dots, c_k$. Dostaneme:

$$\begin{gathered}
c_1\vec{\alpha}_1(c_1\vec{\alpha}_1)^T + c_2\vec{\alpha}_2(c_1\vec{\alpha}_1)^T + \cdots + c_k\vec{\alpha}_k(c_1\vec{\alpha}_1)^T = 0\\
c_1\vec{\alpha}_1(c_2\vec{\alpha}_2)^T + c_2\vec{\alpha}_2(c_2\vec{\alpha}_2)^T + \cdots + c_k\vec{\alpha}_k(c_2\vec{\alpha}_2)^T = 0\\
\dots\\
c_1\vec{\alpha}_1(c_k\vec{\alpha}_k)^T + c_2\vec{\alpha}_2(c_k\vec{\alpha}_k)^T + \cdots + c_k\vec{\alpha}_k(c_k\vec{\alpha}_k)^T = 0
\end{gathered}$$

Na základe týchto rovností už vieme ukázať, že pre vektor $\vec{\beta} = c_1\vec{\alpha}_1 + c_2\vec{\alpha}_2 + \cdots + c_k\vec{\alpha}_k$ platí $\vec{\beta}\vec{\beta}^T = 0$. Po roznásobení totiž dostaneme presne

$$\begin{aligned}
\vec{\beta}\vec{\beta}^T &= (c_1\vec{\alpha}_1 + c_2\vec{\alpha}_2 + \cdots + c_k\vec{\alpha}_k)(c_1\vec{\alpha}_1 + c_2\vec{\alpha}_2 + \cdots + c_k\vec{\alpha}_k)^T\\
&= (c_1\vec{\alpha}_1 + c_2\vec{\alpha}_2 + \cdots + c_k\vec{\alpha}_k)((c_1\vec{\alpha}_1)^T + (c_2\vec{\alpha}_2)^T + \cdots + (c_k\vec{\alpha}_k)^T)\\
&= c_1\vec{\alpha}_1(c_1\vec{\alpha}_1)^T + c_2\vec{\alpha}_2(c_1\vec{\alpha}_1)^T + \cdots + c_k\vec{\alpha}_k(c_1\vec{\alpha}_1)^T +\\
&+ c_1\vec{\alpha}_1(c_2\vec{\alpha}_2)^T + c_2\vec{\alpha}_2(c_2\vec{\alpha}_2)^T + \cdots + c_k\vec{\alpha}_k(c_2\vec{\alpha}_2)^T +\\
&\dots\\
&+ c_1\vec{\alpha}_1(c_k\vec{\alpha}_k)^T + c_2\vec{\alpha}_2(c_k\vec{\alpha}_k)^T + \cdots + c_k\vec{\alpha}_k(c_k\vec{\alpha}_k)^T\\
&= 0
\end{aligned}$$

Podľa lemy 1 to ale znamená, že $\vec{\beta} = c_1\vec{\alpha}_1 + c_2\vec{\alpha}_2 + \cdots + c_k\vec{\alpha}_k = \vec{0}$ a z lineárnej nezávislosti vektorov $\vec{\alpha}_1, \dots, \vec{\alpha}_n$ máme potom $c_1 = c_2 = \cdots = c_k = 0$.

Teraz by sme už chceli uviesť prvý spôsob, ako môžeme využiť lemu 2 na riešenie úlohy 1. Ako pomocné tvrdenie budeme potrebovať

Lema 3. *$h(AB) \le h(B)$ a $h(AB) \le h(A)$ platí pre ľubovoľné matice, ktoré sa dajú násobiť.*

Opäť si môžeme všimnúť, že ak dokážeme jednu z uvedených nerovností, tak druhú dostaneme transponovaním a použitím $h(A) = h(A^T)$. Keďže túto úlohu poznáte z cvičení (bonusová úloha na druhej písomke), tak len stručne naznačím niektoré z možných dôkazov.

Dôkaz. Riadky matice AB sú lineárne kombinácie riadkov matice B. To znamená, že $V_{AB} \subseteq V_B$ a pre dimenzie týchto priestorov platí $d(V_{AB}) \le d(V_B)$. Posledná nerovnosť je len iný zápis nerovnosti $h(AB) \le h(B)$. □

Dôkaz. $B\vec{\alpha}^T = \vec{0}^T \Rightarrow AB\vec{\alpha}^T = \vec{0}^T$

To nám dáva inklúziu medzi množinami riešení homogénnych sústav s maticami B a AB, a teda aj nerovnosť medzi dimenziami

$$n - h(B) \le n - h(AB).$$

□

Riešenie úlohy 1. Nech $h(A) = k$. Teda medzi riadkami matice A je určite k lineárne nezávislých riadkov. Ak tieto riadky dáme do matice B, dostaneme tak podmaticu matice A. Súčasne BB^T je podmatica matice AA^T, a táto podmatica má hodnosť k. Z toho vyplýva, že $h(AA^T) \ge h(A)$. (Matica AA^T obsahuje podmaticu hodnosti k, celá matica teda má hodnosť aspoň k.)

Nerovnosť $h(AA^T) \le h(A)$ môžeme zdôvodniť napríklad na základe lemy 3. □

V nasledujúcom riešení využijeme ten fakt, že násobenie regulárnou maticou nemení hodnosť.

Lema 4. *Nech A je ľubovoľná matica, B a C sú regulárne matice takých rozmerov, že uvedené súčiny majú zmysel. Potom*

$$h(AB) = h(A) = h(CA).$$

Opäť stačí dokazovať jednu z rovností, druhá vyjde transponovaním.

Dôkaz. Násobenie maticou B sprava znamená, že sme riadky matice A zobrazili izomorfizmom f_B. Izomorfizmus nemení dimenziu podpriestoru. □

Dôkaz. Dvakrát využijeme lemu 3. Máme

$$h(AB) \leq h(A).$$

Súčasne môžeme napísať $A = AB.B^{-1}$ a použitím lemy 3 pre matice AB a B^{-1} dostaneme

$$h(A) = h(AB.B^{-1}) \leq h(AB).$$

□

Riešenie úlohy 1. Maticu A môžeme upraviť na redukovaný trojuholníkový tvar pomocou riadkových operácií. Ak riadkové operácie prepíšeme pomocou matíc, tak máme $E_n \dots E_1 A = R$, čiže $A = {E_1}^{-1} \dots {E_n}^{-1} R$. Označme $E := {E_1}^{-1} \dots {E_n}^{-1}$. Dostali sme

$$A = ER,$$

kde R je redukovaná trojuholníková matica a E je regulárna matica. Samozrejme, potom aj matica E^T musí byť regulárna, teda podľa lemy 4 máme[1]

$$h(AA^T) = h(ERR^T E^T) = h(RR^T).$$

Súčasne vieme, že $h(A) = h(R)$. Teda naozaj nám už stačí dokázať, že $h(RR^T) = h(R)$.

Označme $h(R) = k$. Redukovaná trojuholníková matica R pozostáva z k nenulových a $n - k$ nulových riadkov. Označme podmaticu pozostávajúcu z nenulových riadkov matice R ako B. Teda $R = \left(\begin{smallmatrix} B \\ 0 \end{smallmatrix}\right)$ a

$$RR^T = \begin{pmatrix} B \\ 0 \end{pmatrix} \begin{pmatrix} B^T & 0 \end{pmatrix} = \begin{pmatrix} BB^T & 0 \\ 0 & 0 \end{pmatrix}.$$

Nulové riadky a stĺpce neovplyvnia hodnosť matice, teda $h(RR^T) = h(BB^T)$. Riadky matice B sú lineárne nezávislé – sú to presne nenulové riadky RTM. Teda podľa lemy 2 platí $h(BB^T) = k$. □

[1] Toto by sme vedeli zdôvodniť aj bez lemy 4, keď by sme si uvedomili, že násobenie maticou E^T sprava zodpovedá stĺpcovým úpravám, stĺpcové ani riadkové operácie nemenia hodnosť.