ФКН ВШЭ
Основы Матричных Вычислений
Весенний семестр 2025

Домашняя работа 1
(теория)

Работу выполнил:

Назмиев Айрат, группа 1

Задача 1

Пусть задан вектор $u \in \mathbb{C}^n\colon \|u\|_2 = 1$. Найдите все $\alpha \in \mathbb{C}$, для которых $A = I - \alpha uu^*$ является: 1) эрмитовой 2) косоэрмитовой 3) унитарной 4) нормальной. Для пункта 3) также нарисуйте найденные α на комплексной плоскости.

□ Найдем все $\alpha \in \mathbb{C}$ в каждом пункте. Обозначим $\alpha = a + ib,\ a, b \in \mathbb{R}$.

1. A — эрмитова: $A^* = A$.

   $A^* = I - \overline{\alpha}uu^* = I - \alpha uu^* = A$, из условия $\|u\|_2 = 1$, значит $uu^* \neq 0$.

   $\alpha = \overline{\alpha} \iff \alpha \in \mathbb{R}$.

2. A — косоэрмитова: $A^* = -A$.

   $A^* = I - \overline{\alpha}uu^* = -I + \alpha uu^* = -A$.

   $2I = (\alpha + \overline{\alpha})uu^*$.

   I имеет ранг n, тогда как $\operatorname{rank} uu^* = 1$ $(u \neq 0)$, в этом случае $uu^* = u^*u = 1$. Значит, равенство может выполнятся только при $n = 1$, в скалярном случае. Тогда также должно быть выполнено $2 = \alpha + \overline{\alpha} = 2a$. Получаем, что для $n \geqslant 1$ таких α не существует (A не может быть косоэрмитовой), для $n = 1$: $\alpha = 1 + ib,\ \forall b \in \mathbb{R}$.

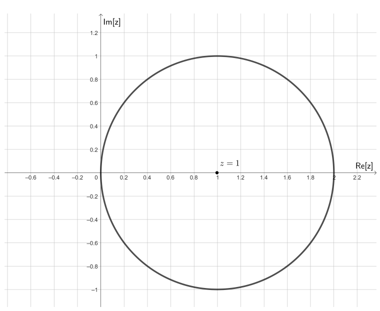
3. A — унитарная: $AA^* = I$.

   $AA^* = (I - \alpha uu^*)(I - \overline{\alpha}uu^*) = I - (\alpha + \overline{\alpha})uu^* + |\alpha|^2 uu^*uu^* = [u^*u = \|u\|_2 = 1] = I - (\alpha + \overline{\alpha})uu^* + |\alpha|^2 uu^* = I$.

   $|\alpha|^2 uu^* = (\alpha + \overline{\alpha})uu^*$.

   $a^2 + b^2 = 2a,\ (a - 1)^2 + b^2 = 1$.

   Таким образом, $\alpha \in \{z \in \mathbb{C} :\ |z - 1| = 1\}$. Изобразим множество на комплексной плоскости: окружность радиуса 1 с центром в $z = 1$.



4. A — нормальная, $AA^* = A^*A$.

   $AA^* = I - (\alpha + \overline{\alpha})uu^* + |\alpha|^2 uu^*$; $A^*A = I - (\alpha + \overline{\alpha})uu^* + |\alpha|^2 uu^*$.

   Можно видеть, что $AA^* = A^*A,\ \forall \alpha \in \mathbb{C}$. ■

Задача 2

1. Докажите, что для любой косоэрмитовой $K \in \mathbb{C}^{n\times n}$ матрица:

$$Q = (I - K)^{-1}(I + K) \tag{1}$$

   будет унитарной.

2. Найдите множество всех унитарных матриц, которые представимы в виде (1).

□ Для начала покажем, что матрица $I - K$ является обратимой. Известно, что собственные значения косоэрмитовых матриц имеет вид $\lambda(K) \in i\mathbb{R}$, то есть являются чисто мнимыми, в том числе могут равняться нулю. Пусть $K = U\Lambda U^{-1} = U\Lambda U^*$ — ее собственное разложение в ОНБ комплексного векторного пространства (косоэрмитовы матрицы нормальные). Тогда матрицу можно расписать в следующем виде: $I - K = UU^* - U\Lambda U^* = U(I - \Lambda)U^*$, то есть $I - K$ имеет собственные значения вида $\lambda(I - K) = 1 - \lambda(K) \neq 0$. Таким образом, все собственные значения данной матрицы ненулевые, а значит, $I - K$ обратима.

Покажем, что матрица вида (1) является унитарной, для этого требуется доказать $QQ^* = I$. Для этого воспользуемся известными свойствами: $(AB)^* = B^*A^*$, $(A^{-1})^* = (A^*)^{-1}$, свойством $K^* = -K$. Получим:

$$Q^* = (I - K)(I + K)^{-1},$$
$$QQ^* = (I - K)^{-1}(I + K)(I - K)(I + K)^{-1}.$$

Докажем простой факт: матрицы вида $I - A$ и $I + A$ коммутируют $\forall A \in \mathbb{C}$. Действительно:

$$(I + A)(I - A) = I - A^2 = (I - A)(I + A).$$

Поэтому можно переставить в QQ^* второй и третий множитель местами:

$$QQ^* = (I - K)^{-1}(I - K)(I + K)(I + K)^{-1} = I,$$

что и требовалось.

Было доказано, что для любой косоэрмитовой $K \in \mathbb{C}^{n\times n}$ формула (1) задает унитарную Q.

Теперь найдем множество унитарных матриц, представимых в виде (1).

Известно, что в общем случае собственные числа косоэрмитовых матриц чисто мнимые, а ортогональных — являются комплексными и равны единице по модулю.

Пусть v — собственный вектор K: $Kv = \lambda v$, $\lambda \in i\mathbb{R}$. Тогда:

$$Qv = (I - K)^{-1}(1 + \lambda)v. \tag{2}$$

Известно, что собственные векторы матрицы и обратной матрицы совпадают, при этом соответствующим данному вектору собственным значением обратной матрицы будет обратное собственное число исходной матрицы:

$$Ax = \lambda x,\ x = \lambda A^{-1}x,\ \lambda^{-1}x = A^{-1}x.$$

Значит, если v - собственный для K, то он собственный как для $(I - K)$:

$$(I - K)v = (1 - \lambda)v,$$

так и для $(I - K)^{-1}$:

$$(I - K)^{-1}v = \frac{1}{1 - \lambda}v.$$

Таким образом, (2) можно записать в виде:

$$Qv = \frac{1+\lambda}{1-\lambda}v = \mu v, \tag{3}$$

что означает, Q имеет те же собственные векторы v, что и K, которым соответствует собственное число $\mu = (1+\lambda)(1-\lambda)^{-1}$.

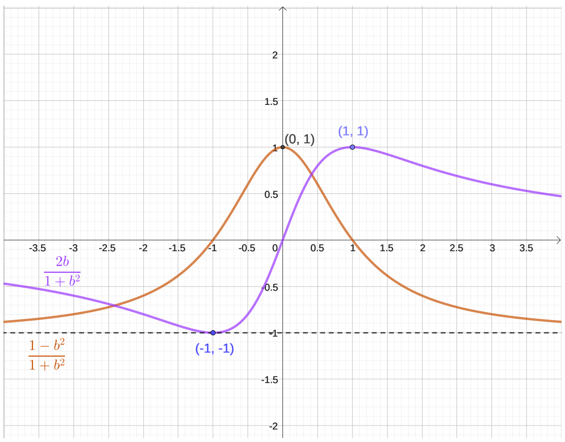
Введем обозначение $\lambda = ib,\ b \in \mathbb{R}$. Определим, какие значения может принимать μ:

$$\mu = \frac{1+ib}{1-ib} = \frac{(1-b^2)+2ib}{1+b^2}. \tag{4}$$

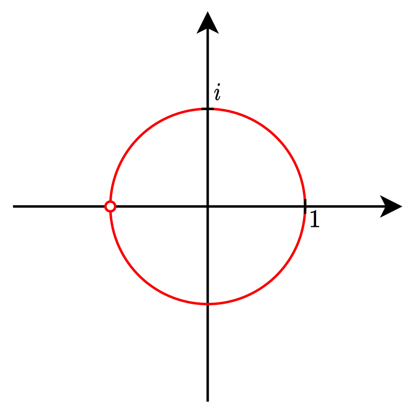
Можно легко убедится, что $|\mu| = 1$, то есть собственные значения лежат на единичной окружности:

$$|\mu| = \left|\frac{(1-b^2)+2ib}{1+b^2}\right| = \sqrt{\frac{(1-b^2)^2+(2b)^2}{(1+b^2)^2}} = \sqrt{\frac{1-2b^2+b^4+4b^2}{(1+b^2)^2}} = 1.$$

Введем обозначение $\mu = e^{i\theta} = \cos(\theta) + i\sin(\theta)$. Из (4) получаем, что $\cos(\theta) = (1-b^2)(1+b^2)^{-1}$, $\sin(\theta) = 2b(1+b^2)^{-1}$. Зададимся вопросом, какие значения может принимать μ на единичной окружности.



(a) Графики $\frac{1-b^2}{1+b^2}$ и $\frac{2b}{1+b^2}$



(b) μ на комплексной плоскости

Можно легко показать тот факт, что в параметризации (4) невозможно задать точку $\mu = -1$. Доказывается это средствами матанализа, исследуя нули, экстремумы, участки монотонности и асимптоты функций $x(b) = (1-b^2)(1+b^2)^{-1}$ и $y(b) = 2b(1+b^2)^{-1}$, но выкладки оказываются крайне громоздкими, поэтому здесь не приводятся. Вместо этого приведены графики функций и полученное множество допустимых значений μ на $\mathbb{C}$. Можно видеть, что при изменении $b \in \mathbb{R}$ достигаются все точки на единичной окружности, $(x(b), y(b))$, $x^2(b) + y^2(b) = 1$, кроме $(-1, 0)$: она не достигается и является пределом при $b \to \pm\infty$. То есть $\mu \in \{z \in \mathbb{C} :\ |z| = 1,\ z \neq -1\}$.

Вместо описанного выше наброска доказательства ГМТ μ можно заметить, что (4) является преобразованием Кэли в скалярном случае, которое отображает мнимую ось на комплексной плоскости в единичную окружность, при этом $\pm i\infty$ отображается в -1. Кстати, в задаче изначально рассматривается более общее отображение Кэли: для комплексных матриц, в приложении к параметризации унитарных матриц через косоэрмитовы.

Таким образом, вид (1) описывает множество всех унитарных матриц $Q \in \mathbb{C}^{n\times n}$, которые не имеют собственного значения -1. ■

Задача 3

Докажите, что

$$\|A\|_{1\to 2} \leqslant \|A\|_2 \leqslant \sqrt{n}\|A\|_{1\to 2}, \quad \forall A \in \mathbb{C}^{m\times n}.$$

Указание: В случае использования констант эквивалентности векторных норм необходимо обосновывать это значение констант.

□ Докажем, что $\|x\|_2 \leqslant \|x\|_1 \leqslant \sqrt{n}\|x\|_2$, для этого воспользуемся полуаддитивностью квадратного корня и неравенством КБШ:

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2} \leqslant \sum_{i=1}^{n} \sqrt{|x_i|^2} = \sum_{i=1}^{n} |x_i| = \|x\|_1 = \sum_{i=1}^{n} |x_i| \cdot 1 \leqslant \sqrt{\sum_{i=1}^{n} |x_i|^2} \cdot \sqrt{\sum_{i=1}^{n} 1^2} = \sqrt{n}\|x\|_2.$$

Выпишем определения операторных норм:

$$\|A\|_{1\to 2} = \sup_{\|x\|_1=1} \|Ax\|_2, \quad \|A\|_2 = \sup_{\|x\|_2=1} \|Ax\|_2$$

Далее докажем первое неравенство в условии задачи: $\|A\|_{1\to 2} \leqslant \|A\|_2$. Из определения операторной нормы ($\forall x \in \mathbb{C}^n$, $\|x\|_1 = 1$ выполнено $\|x\|_2 \leqslant \|x\|_1 = 1$) и доказанных неравенств на векторные нормы следует, что:

$$\|Ax\|_2 \leqslant \|A\|_2\|x\|_2 \leqslant \|A\|_2\|x\|_1 = \|A\|_2 \cdot 1 \leqslant \|A\|_2.$$

Взяв супремум по $\|x\|_1$ на единичной сфере, можно получить:

$$\|A\|_{1\to 2} = \sup_{\|x\|_1=1} \|Ax\|_2 \leqslant \|A\|_2,$$

что и требовалось. Перейдем к доказательству второго неравенства, $\|A\|_2 \leqslant \sqrt{n}\|A\|_{1\to 2}$. Вновь сошлемся на доказанные неравенства на векторные нормы ($\forall x \in \mathbb{C}^n$, $\|x\|_2 = 1$ выполнено $\|x\|_1 \leqslant \sqrt{n}\|x\|_2 = \sqrt{n}$) и определения операторных норм и запишем:

$$\|Ax\|_2 \leqslant \|A\|_{1\to 2}\|x\|_1 \leqslant \|A\|_{1\to 2}\sqrt{n}\|x\|_2 = \sqrt{n}\|A\|_{1\to 2}$$

Взяв супремум по $\|x\|_2$ на единичной сфере, можно получить:

$$\|A\|_2 = \sup_{\|x\|_2=1} \|Ax\|_2 \leqslant \sqrt{n}\|A\|_{1\to 2},$$

оба неравенства доказаны. ■

Задача 4

Обозначим $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ и $A_n = \begin{bmatrix} 0 & 1 \\ 1/n & 0 \end{bmatrix}, n \in \mathbb{N}$.

1. Обоснуйте сходимость $A_n \to A$, $n \to \infty$ исходя из определения сходимости с помощью норм.
2. Найдите собственные разложения $A_n = S_n\Lambda_n S_n^{-1}$ и проверьте существование пределов для каждой из S_n,Λ_n и S_n^{-1}. Почему не у всех из этих матриц существует предел?
3. Найдите разложения Шура $A_n = U_n T_n U_n^{-1}$ и проверьте существование пределов для каждой из U_n,T_n и U_n^{-1}.

□ Докажем сходимость $A_n \to A$, $n \to \infty$. Так как все матричные нормы в конечномерных пространствах эквивалентны, поэтому для удобства будем показывать сходимости по норме Фробениуса:

$$\|A - A_n\|_F = \frac{1}{n} \to 0,\ n \to \infty,$$

то есть A_n сходится к A.

Найдем собственные разложение A и A_n.

$$\det(A - \lambda I) = \lambda^2 = 0, \quad \lambda_1 = \lambda_2 = 0, \quad (A - 0 \cdot I)v = 0, \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} v = 0, \quad v = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

можно видеть, что у матрицы A только один линейно независимый собственный вектор (а не 2), то есть у A не имеет собственного разложения. При этом ЖНФ A имеет вид:

$$A = PJP^{-1}, \quad J = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad P = P^{-1} = I$$

Собственное разложение A_n:

$$\det(A_n - \lambda I) = \lambda^2 - \frac{1}{n} = 0, \quad \lambda_{n1} = \frac{1}{\sqrt{n}}, \quad \lambda_{n2} = -\frac{1}{\sqrt{n}}, \quad (A - 0 \cdot I)v = 0, \quad v_{n1} = \begin{bmatrix} \sqrt{n} \\ 1 \end{bmatrix}, \quad v_{n2} = \begin{bmatrix} -\sqrt{n} \\ 1 \end{bmatrix};$$

$$\Lambda_n = \begin{bmatrix} \frac{1}{\sqrt{n}} & 0 \\ 0 & -\frac{1}{\sqrt{n}} \end{bmatrix}, \quad S_n = \begin{bmatrix} \sqrt{n} & -\sqrt{n} \\ 1 & 1 \end{bmatrix}, \quad S_n^{-1} = \begin{bmatrix} \frac{1}{2\sqrt{n}} & \frac{1}{2} \\ -\frac{1}{2\sqrt{n}} & \frac{1}{2} \end{bmatrix}$$

Можно легко видеть (например, рассмотрев разницу матриц в норме Фробениуса), что последовательности Λ_n и S_n^{-1} имеют предел, а S_n — нет (так как некоторые элементы матрицы стремятся к бесконечности):

$$\Lambda_n \xrightarrow[n\to\infty]{} 0, \quad S_n^{-1} \xrightarrow[n\to\infty]{} \begin{bmatrix} 0 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix}.$$

При этом $A_n = S_n \Lambda_n S_n^{-1}$ для любого *конечного* n. Расходимость некоторых матриц (для выбранных собственных векторов — расходимость S_n) связана с тем, что предел последовательности A_n является вырожденной матрицей A, для которой не существует собственного разложения. При этом важно заметить, что сходимость собственных значений имеет место, λ_{n1} и λ_{n2} сходятся к 0.

Теперь найдем разложения Шура матриц A и A_n.

Для A легко проверить, что можно записать разложение Шура, совпадающее с найденной выше ЖНФ:

$$A = UTU^*, \quad T = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad U = U^* = I.$$

Найдем разложение Шура для A_n. Нормируем найденный собственный вектор v_1, найдем для него ортогональный вектор и также его нормируем (вектора сформируют унитарную матрицу), далее найдем верхнетруегольную матрицу T_n:

$$v_{n1} = \begin{bmatrix} \sqrt{n} \\ 1 \end{bmatrix}, \quad v_{n2} = \begin{bmatrix} -1 \\ \sqrt{n} \end{bmatrix}, \quad u_{ni} = \frac{v_{ni}}{\|v_{ni}\|_2},\ i = 1, 2,$$

$$U_n = \begin{bmatrix} \frac{\sqrt{n}}{\sqrt{n+1}} & -\frac{1}{\sqrt{n+1}} \\ \frac{1}{\sqrt{n+1}} & \frac{\sqrt{n}}{\sqrt{n+1}} \end{bmatrix}, \quad U_n^{-1} = U_n^*, \quad T_n = U_n^* A_n U_n = \begin{bmatrix} \frac{1}{\sqrt{n}} & \frac{n-1}{n} \\ 0 & -\frac{1}{\sqrt{n}} \end{bmatrix}$$

$$U_n, U_n^* \xrightarrow[n\to\infty]{} I = U, \quad T_n \xrightarrow[n\to\infty]{} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = T.$$

Легко видеть, что полученные матрицы имеют пределы (например, это можно показать через фробениусовскую норму), при этом они совпадают с разложением Шура матрицы A, к которой стремится A_n. Таким образом можно заключить, что все матрицы в разложении Шура A_n сходятся к соответствующим матрицам в разложении предельной матрицы A. ■

Начиная со следующей задачи решения задач будут приведены в письменном виде.

№5

Докажем, что $A \in \mathbb{C}^{n \times n}$ – нормальная является унитарной $\Longleftrightarrow$ все её с.з. λ_i, $i = \overline{1,n}$: $|\lambda_i| = 1$

□ • $A \in \mathbb{C}^{n \times n}$ – нормальная, если $A^*A = AA^*$, нормальные матрицы унитарно диагонализуемы

• $A \in \mathbb{C}^{n \times n}$ – унитарная, если $A^*A = AA^* = I$ (частный случай нормальной)

$\Longrightarrow$ Пусть λ – с.з. A, v – с.в., соотв. λ, т.е. $Av = \lambda v$

Рассмотрим также $v^*A^* = \bar{\lambda} v^*$. Умножим $Av = \lambda v$ слева на данное выражение:

$v^*A^*Av = |\lambda|^2 v^*v$.

Т.к. A – унитарная, $A^*A = I$, то есть:

$v^*v = |\lambda|^2 v^*v$.

Т.к. по определению $v \neq 0 \Rightarrow v^*v = \|v\|_2^2 \neq 0$ получаем:

$|\lambda|^2 = 1 \Rightarrow |\lambda| = 1$

Таким образом, $\forall$ с.з. унитарной матрицы A по модулю $= 1$

$\Longleftarrow$ Пусть A – нормальная матрица и $\forall$ её с.з. λ_i, $i = \overline{1,n}$: $|\lambda_i| = 1$.

Т.к. A – нормальная $\Rightarrow$ унитарно диагонализуема, т.е.:

$A = U\Lambda U^*$; $UU^* = I$, $\Lambda = \operatorname{diag}(\lambda_1 \ldots \lambda_n)$

Заметим $A^* = U\Lambda^*U^*$; $\Lambda^* = \operatorname{diag}(\bar{\lambda}_1 \ldots \bar{\lambda}_n)$

Найдём AA^*:

$AA^* = U\Lambda U^* U\Lambda^*U^* = U\Lambda\Lambda^*U^* = \left[\Lambda\Lambda^* = \operatorname{diag}(|\lambda_1|^2 \ldots |\lambda_n|^2) = I\right] =$

$= UIU^* = I$

Значит, A – унитарная ■

№6

а) Докажите, что для $\forall A \in \mathbb{C}^{m\times n}$, $m \geq n$, справедливо:

$$\|A\|_2 \leq \|A\|_F \leq \sqrt{n}\,\|A\|_2$$

б) Найдите все матрицы $A \in \mathbb{C}^{n\times n}$: $\|A\|_F = \sqrt{n}\,\|A\|_2$

□ Найдем явный вид матричных норм $\|\cdot\|_F$ и $\|\cdot\|_2$

$A = U\Sigma V^*$ – SVD ; r – ранг A ; $\Sigma \in \mathbb{C}^{m\times n}$; $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r \geq 0$ $\quad r \leq \min(m,n)$

$$\|A\|_F = \|U\Sigma V^*\|_F = [\|\cdot\|_F \text{ – унитарно-инв.}] = \|\Sigma\|_F = \sqrt{\sum_{i=1}^{r}\sigma_i^2}$$

$$\|A\|_2 = \|U\Sigma V^*\|_F = [\|\cdot\|_2 \text{ – унитарно-инв.}] = \|\Sigma\|_2$$

$$\|\Sigma\|_2 = \sup_{x\neq 0}\frac{\|\Sigma x\|_2}{\|x\|_2} = \sup_{x\neq 0}\frac{\sqrt{\sum_{i=1}^{r}\sigma_i^2|x_i|^2}}{\sqrt{\sum_{i=1}^{n}|x_i|^2}} \leq \sup_{x\neq 0}\frac{\sqrt{\sum_{i=1}^{r}\sigma_i^2|x_i|^2}}{\sqrt{\sum_{i=1}^{r}|x_i|^2}} = \sup_{x\neq 0}\frac{\sigma_1\cdot\sqrt{\sum_{i=1}^{r}|x_i|^2}}{\sqrt{\sum_{i=1}^{r}|x_i|^2}} ⊜$$

$⊜\, \sigma_1$; оценка достигается при $x = e_1$.

а) $\|A\|_2 = \sigma_1 \leq \sqrt{\sum_{i=1}^{r}\sigma_i^2} = \|A\|_F$ $\quad \left(\sigma_1^2 \leq \sum_{i=1}^{r}\sigma_i^2\right)$ – первое нер-во доказано

$$\|A\|_F^2 = \sum_{i=1}^{r}\sigma_i^2 \leq n\cdot\sigma_1^2 = n\cdot\|A\|_2^2 \Rightarrow \|A\|_F \leq \sqrt{n}\cdot\|A\|_2,$$

что и требовалось

б) $A \in \mathbb{C}^{n\times n}$; $\|A\|_F = \sqrt{n}\,\|A\|_2 \to \|A\|_F^2 = n\|A\|_2^2$

$\sum_{i=1}^{r}\sigma_i^2 = n\sigma_1^2$. Т.к. $\sigma_i \leq \sigma_1 \Rightarrow \underline{\sigma_i = \sigma_1 = \sigma\ \forall i}$ $\qquad U, V \in \mathbb{C}^{n\times n}$

Тогда $A = U\Sigma V^* = [\Sigma = \sigma I \in \mathbb{C}^{n\times n}] = \sigma U V^* = \sigma Q$, где $Q \in \mathbb{C}^{n\times n}$ – унитарная как произведение унитарных $(QQ^* = I)$

Таким образом, $\|A\|_F = \sqrt{n}\|A\|_2$ выполняется для матриц $A \in \mathbb{C}^{n\times n}$ вида:

$A = \sigma Q$, $Q \in \mathbb{C}^{n\times n}$ – унитарная, $\sigma \in \mathbb{R}_+$ $(\sigma \geq 0)$ ☒

№7

Дана нормальная матрица $A \in \mathbb{C}^{n \times n}$ и её разложение Шура: $A = UTU^*$

а) Требуется записать SVD матрицы A, используя U и T.

б) Показать, что $\sigma_1(A) = \max\limits_i |\lambda_i(A)|$

в) Привести пример $A \in \mathbb{C}^{2\times 2}$, для которой ~~[illegible]~~ формула в б) неверна

а) Матрица A — нормальная: $AA^* = A^*A$

$A = UTU^*$; $UTU^*UT^*U^* = UTT^*U^* = UT^*U^*UTU^* = UT^*TU^*$

Значит, $TT^* = T^*T$, где T — верхнетреугольная с с.з. λ_i, $i = \overline{1,n}$ на диагонали.

Лемма

• Докажем, что ~~[illegible]~~ верхнетреугольная матрица B коммутирует с $B^* \Leftrightarrow B$ — диагональная

[$(BB^*)_{ij} = \sum\limits_{k=\max(i,j)}^{n} b_{ik}\overline{b_{jk}}$ $\quad (b_{ij} = 0 \ \ \forall i > j)$

$(B^*B)_{ij} = \sum\limits_{k=1}^{\min(i,j)} \overline{b_{ki}}\, b_{kj}$

Рассмотрим диагональные эл-ты матриц:

$(BB^*)_{ii} = \sum\limits_{k=i}^{n} |b_{ik}|^2$

$(B^*B)_{ii} = \sum\limits_{k=1}^{i} |b_{ki}|^2$

$(BB^*)_{ii} = (B^*B)_{ii}$. По индукции докажем, что B — диагональная

$\underline{i=1}$: $\sum\limits_{k=1}^{n} |b_{1k}|^2 = \sum\limits_{k=1}^{1} |b_{k1}|^2 = |b_{11}|^2 \Rightarrow |b_{12}|^2 + \ldots + |b_{1n}|^2 = 0 \Rightarrow b_{1j} = 0$, $j \neq 1$

$\underline{i=m}$: пусть для $\forall i \leq m-1$ выполнено $b_{ij} = 0$, $j \neq i$

$\sum\limits_{k=m}^{n} |b_{mk}|^2 = \sum\limits_{k=1}^{m} |b_{km}|^2 = [\text{по предположению индукции}] = |b_{mm}|^2$

$\Downarrow$

$\cancel{|b_{mm}|^2} + \sum\limits_{k=m+1}^{n} |b_{mk}|^2 = \cancel{|b_{mm}|^2} \Rightarrow b_{mk} = 0$, $n \geq k \geq m+1$

Доказано, что B — диагональная]

Доказанная лемма говорит о том, что $T = \mathrm{diag}(\lambda_1 \ldots \lambda_n) = \Lambda$, то есть разложение Шура нормальной матрицы — диагонализация в унитарном базисе

$A = U\Lambda U^*$

б) $\sigma_i(A) = \sqrt{\lambda_i(AA^*)} = \sqrt{\lambda_i(A^*A)}$

$AA^* = U\Lambda\Lambda^*U^* = U \cdot \mathrm{diag}(|\lambda_1|^2 \ldots |\lambda_n|^2) \cdot U^* \Rightarrow \sigma_i(A) = |\lambda_i(A)|$

а) Таким образом, SVD A имеет вид:

$$A = U \Sigma U^* ; \quad \Sigma = \mathrm{diag}(|\lambda_1| \dots |\lambda_n|)$$

Через $T = \Lambda$ можно записать так:

$$A = U \cdot (T T^*)^{1/2} \cdot U^*$$

, корень диагональной неотриц. определенной матрицы — диагональная матрица с корнями эл-тов исходной матрицы на диагонали

в) $\measuredangle A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$; $\underset{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}}{AA^*} \neq \underset{\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}}{A^* A}$

Как было показано в N4, ~~[illegible]~~ $\lambda_1(A) = \lambda_2(A) = 0$

$$AA^* = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} ; \quad |A - \lambda E| = \lambda(\lambda - 1), \quad \lambda_1(AA^*) = 1, \quad \lambda_2(AA^*) = 0$$

$$\sigma_1(A) = \max_i \sqrt{\lambda_i(AA^*)} = 1 \neq \cancel{\max} \; 0 = \max_i (\lambda_i(A))$$

— условие не выполняется в общем случае

☒

№8

Покажем, что для множителя W из полярного разложения матрицы $A \in \mathbb{R}^{m\times n}$, $m \ge n$ ($A = WH$, $W^TW = I$, $W \in \mathbb{R}^{m\times n}$, $H \succeq 0$, $H \in \mathbb{R}^{n\times n}$) выполняется:

$$\frac{\|A^TA - I\|_2}{\|A\|_2+1} \le \|A-W\|_2 \le \|A^TA-I\|_2$$

Замечание: нер-во даёт соотношение между двумя мерами близости (в смысле $\|\cdot\|_2$) матрицы A к ортогональной: $\|A^TA-I\|_2$ и $\min\limits_{Q:Q^TQ=I}\|A-Q\|_2$

D $A = WH$ — полярное разложение

1) Докажем $\|A-W\|_2 \le \|A^TA-I\|_2$

$\|A-W\|_2 = \|W(H-I)\|_2 = \|H-I\|_2$

$\|A^TA - I\|_2 = \|H^TW^TWH - I\|_2 = \|H^TH-I\|_2$

$H\in\mathbb{R}^{n\times n}$, $H\succeq 0$, $H=H^T$, запишем собственное разложение H:

$H = QDQ^T$, $D = \mathrm{diag}(d_1,\dots,d_n)$, $d_1\ge d_2\ge\dots\ge d_n\ge 0$ — с.з. H

$H^TH = QD^TDQ^T = QD^2Q^T$

- $\|H-I\|_2 = \|QDQ^T-I\|_2 = \|Q(D-I)Q^T\|_2 = \|D-I\|_2$
- $\|H^TH-I\|_2 = \|Q(D^2-I)Q^T\|_2 = \|D^2-I\|_2$

Докажем, что $\|D-I\|_2 \le \|D^2-I\|_2$

т.к. $(D-I)$ и (D^2-I) — диагональные квадратные матрицы, их $\|\cdot\|_2$ — это модуль максимального эл-та на диагонали (оно же является σ_1 (макс. сингулярное число)), то есть $|d_1-1|$ требуется сравнить с $|d_1^2-1|$:

а) $d_1\in[0;1]$: $1-d_1 \vee 1-d_1^2$; $d_1(d_1-1)\le 0 \Rightarrow 1-d_1\le 1-d_1^2$

б) $d_1>1$: $d_1-1 \vee d_1^2-1$; $0\le d_1(d_1-1) \Rightarrow d_1-1\le d_1^2-1$

таким образом, $|d_1-1|\le|d_1^2-1|$. То есть $\|D-I\|_2\le\|D^2-I\|_2$. Это доказывает $\|A-W\|_2\le\|A^TA-I\|_2$

2) Докажем $\dfrac{\|A^TA-I\|_2}{\|A\|_2+1} \le \|A-W\|_2$

$$\underline{\|A^TA-I\|_2} = \|H^TH-I\|_2 = \|D^2-I\|_2 = \|(D-I)(D+I)\|_2 \overset{[\text{субмультипликативность } \|\cdot\|_2]}{\le} \underbrace{\|D-I\|_2}_{\|A-W\|_2}\cdot\|D+I\|_2 \ominus$$

$$\ominus\ \|A-W\|_2\cdot\|D+I\|_2 \le \|A-W\|_2\cdot(\|D\|_2+1) = \underline{\underline{\|A-W\|_2\cdot(\|A\|_2+1)}}$$

$[\|D+I\|_2 = |d_1+1| \overset{[\text{нер-во }\Delta]}{\le} |d_1|+1 = \|D\|_2+1]$

Значит, $\dfrac{\|A^TA-I\|_2}{\|A\|_2+1} \le \|A-W\|_2$ ⊠

Домашняя работа 1
(теория)

№1

Норма $\|\cdot\|^*$ — двойственная к $\|\cdot\|$ на $\mathbb{R}^{m\times n}$, если

$$\|A\|^* = \max_{B:\|B\|=1} \mathrm{Tr}(AB^T)$$

Докажите, что:

а) $(\|A\|^*)^* = \|A\|$

б) $\|\cdot\|$ — унитарно-инв. $\Longleftrightarrow$ $\|\cdot\|^*$ — унитарно-инв

$\square$

б) $(\Rightarrow)$ $\exists$ $\|\cdot\|$ — унитарно-инв; $A, B \in \mathbb{R}^{m\times n}$; $\left.\begin{matrix} U\in\mathbb{R}^{m\times m} \\ V\in\mathbb{R}^{n\times n}\end{matrix}\right\}$ — ортогональные матрицы (унитарные)

$$\|UAV\|^* = \max_{B:\|B\|=1} \mathrm{Tr}(UAVB^T) = \max_{B:\|B\|=1} \mathrm{Tr}(AVB^TU) = \left[\begin{matrix} D = U^TBV^T \\ \|D\| = \|B\| = 1\end{matrix}\right] =$$

$$= \max_{D:\|D\|=1} \mathrm{Tr}(AC^T) = \|A\|^*,$$ т.е. $\|\cdot\|^*$ — унитарно-инв

$(\Leftarrow)$ $\exists$ $\|\cdot\|^*$ — унитарно-инв; воспользуемся $\|A\|^{**} = \|A\|$ из п. а)

$$\|UAV\| = \max_{C:\|C\|^*=1} \mathrm{Tr}(AVC^TU) = \left[\begin{matrix} M = U^TCV^T \\ \|M\| = \|C\| = 1\end{matrix}\right] = \max_{M:\|M\|^*=1} \mathrm{Tr}(AM^T) = \|A\|,$$

т.е. $\|\cdot\|$ — унитарно-инв.

а) Докажем в два этапа: $(\|A\|^*)^* \le \|A\|$, потом $\|A\| \le (\|A\|^*)^*$; $\|A\|^{**} = \max_{C:\|C\|^*=1} \mathrm{Tr}(AC^T)$

I) $\mathrm{Tr}(AC^T) = \|A\| \cdot \mathrm{Tr}\left(\frac{A}{\|A\|}C^T\right) \le \|A\| \cdot \max_{D:\|D\|=1} \mathrm{Tr}(CD^T) = \|A\| \cdot \|C\|^*$

$\Downarrow$

$$\|A\|^{**} = \max_{C:\|C\|^*=1} \mathrm{Tr}(AC^T) \le \|A\|$$

II) От противного, пусть $\exists A_0: \|A_0\|^{**} > \|A_0\|$

$$\|A\|^{**} = \max_{B:\|B\|^*=1} \mathrm{Tr}(AB^T)$$

$\measuredangle$ $B_0 = \frac{A_0}{\|A_0\|}$; $\|B_0\| = 1$

$\measuredangle$ $S = \{B \in \mathbb{R}^{m\times n} : \|B^*\| \le 1\}$ — выпуклое замкнутое мн-во

По предположению:

$\mathrm{Tr}(A_0B^T) \le \|A_0\|^{**} < \|A_0\|$ $\forall B \in S$

Кроме того, в т. B_0: $\mathrm{Tr}(A_0B_0^T) = \frac{\mathrm{Tr}(A_0A_0^T)}{\|A_0\|} = \|A_0\|$

$\Downarrow$

Значит, $B_0 \notin S$

Т.к. S — замкнутое выпуклое мн-во, [теорема ~~...~~ разделяющей гиперплоскости] (*)

утверждает, что $\exists\, \tilde{C} \in \mathbb{R}^{m\times n}$ и $\alpha \in \mathbb{R}$:

(*) теорема из стандартного курса матана или выпуклого анализа, см. [S. Boyd, pp. 46-51] Convex Optimization

$$Tr(\tilde{C}B^T) \le \alpha < Tr(\tilde{C}B_0^T) \quad \forall B \in S$$

Можем отнормировать $\tilde{C}$: $C = \tilde{C}/\alpha$:

$$Tr(CB^T) \le 1 < Tr(CB_0^T) \quad \forall B \in S$$

По определению: $\|B_0\|^* = \max\limits_{D:\|D\|=1} Tr(B_0D^T)$; при этом в нашем случае $\|B_0\|=1$.

Т.к. $\forall$ унитарно-инвариантные нормы субмультипликативны [утверждение от лектора],

заметим:

$$Tr(B_0D^T) \le \|B_0\|\cdot\|D\| \ominus [\|B\|_0=1;\ \|D\|=1] \ominus 1$$

Значит, $\|B_0\|^* \le 1$

С другой стороны, возьмем $D = B_0$ ($B_0 \in \{D: \|D\|=1\}$), получим: $Tr(B_0B_0^T) = \|B_0\|^2 = 1$

Значит, $\|B_0\|^* = 1$ — противоречие, т.к. $B_0 \notin S$ по предположению

$\Downarrow$

$\|A_0\|^{**} \ge \|A_0\|$

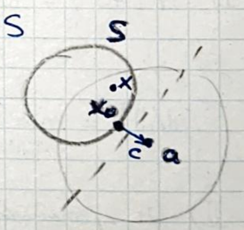
~~По предположению $Tr(A_0B^T) \le \|A_0\|\ \forall B \in S$~~

~~т.к. $B_0 \in S$, получаем $Tr(A_0B_0^T) \le \|A_0\| \Rightarrow$~~

Из (I) и (II) получаем, что $\|A_0\|^{**} = \|A_0\|$ ☒

Приложение — док-во т. ~~...~~ о разделяющей гиперплоскости

$\exists\, S$ — выпуклое и $a \in cl(S) \Rightarrow a$ строго отделима от S



□ 1) ∡ $d(x) = \|x-a\|_2$

2) ∡ шар $B(a,r) = \{t : \|t-a\|_2 \le r\}$

3) ∡ $\bar{X} = S \cap B(a,r)$ — компакт

4) из [т. Вейерштрасса]: $\exists\, x_0$ — т. min на $\bar{X}$, т.е. $d(x) \ge d(x_0)\ \forall x \in \bar{X}$

5) ∡ $c = a - x_0$; покажем, что $\langle c, x - x_0\rangle < 0\ \forall x \in \bar{X}$ (образует тупой угол)

~~Пусть~~ от противного, $\exists\, x_1 \in \bar{X}$: $\langle c, x_1 - x_0\rangle \ge 0$

~~...~~ $x(\theta) = \theta x_1 + (1-\theta)x_0$; ∡ $f(\theta) = d(x(\theta)) = \|x(\theta)-a\|_2^2$

$$f'(0) = 2(x(\theta)-a)^T(x_1-x_0)\Big|_{\theta=0} \overset{[x(0)=x_0]}{=} 2(x_0-a)^T(x_1-x_0) = -2c^T(x_1-x_0) < 0$$

т.е. $\exists\, \theta_0 \in U_\varepsilon(0)$: $d(x(\theta_0)) < d(x_0) \Rightarrow x_0$ — не т. min — противоречие.

6) разделяющая гиперплоскость: $\langle c, x\rangle = \langle c, d\rangle$, где $d = \alpha x_0 + (1-\alpha)a$, $\alpha \in (0;1)$ ☒

№2

Докажем субмультипликативность векторной p-нормы матриц:

$\|A\|_{p,vec} = \|vec(A)\|_p$ при $p \in [1;2]$

[$\forall A \in \mathbb{C}^{m \times n};\ B \in \mathbb{C}^{n \times l}$

- Сначала докажем для $p=1$:

[КБШ] [или $|\sum x| \le \sum |x|$]

$$\|AB\|_{1,vec} = \sum_{i=1}^{m}\sum_{j=1}^{l}\left|\sum_{k=1}^{n} a_{ik}b_{kj}\right| \le \sum_{i=1}^{m}\sum_{j=1}^{l}\sum_{k=1}^{n}|a_{ik}|\cdot|b_{kj}| = \sum_{k=1}^{n}\left(\sum_{i=1}^{m}|a_{ik}|\right)\cdot\left(\sum_{j=1}^{l}|b_{kj}|\right) \le$$

$$\le \left(\sum_{k=1}^{n}\sum_{i=1}^{m}|a_{ik}|\right)\left(\sum_{k=1}^{n}\sum_{j=1}^{l}|b_{kj}|\right) = \|A\|_{1,vec}\cdot\|B\|_{1,vec}$$

- Докажем для $p \in (1;2]$:

Для доказательства потребуется нер-во Гёльдера (доказывалось в курсах математики)

$\forall x, y \in \mathbb{C}^n,\ \forall p, q \ge 1:\ \frac{1}{p}+\frac{1}{q}=1$ $\left(q = \frac{p}{p-1}\right)$ выполняется:

$$\left[|x^T y| \le\right] \sum_{i=1}^{n}|x_i y_i| \le \|x\|_p \cdot \|y\|_q$$

Заметим $\|AB\|_{p,vec}^p$, воспользуемся нер-вом Гёльдера: a_i – i-ая строка A, $i=\overline{1,m}$; b_j – j-ый столбец B, $j=\overline{1,l}$

$$\|AB\|_{p,vec}^p = \sum_{i=1}^{m}\sum_{j=1}^{l}\left|\sum_{k=1}^{n}a_{ik}b_{kj}\right|^p = \sum_{i=1}^{m}\sum_{j=1}^{l}|a_i^T b_j|^p \le \sum_{i=1}^{m}\sum_{j=1}^{l}\|a_i\|_p^p\cdot\|b_j\|_q^p$$

$(i=\overline{1,m};\ j=\overline{1,l})$

Рассмотрим $\forall$ слагаемое и докажем: $|a_i^T b_j|^p \le \|a_i\|_p^p \cdot \|b_j\|_p^p$. Для этого докажем сначала $\forall x \in \mathbb{C}^s$ $\|x\|_q \le \|x\|_p$ при $1 \le p \le q$: $\left(1 \le p \le \frac{p}{p-1};\ p(p-1) \le p \Rightarrow \boxed{1 \le p \le 2}\right)$

[$\|x\|_q^p = \left(\sum_i |x_i|^q\right)^{p/q} \overset{?}{\le} \sum_i |x_i|^p = \|x\|_p^p$

$\forall y = \frac{x}{\|x\|_p};\ \|y\|_p = 1;\ \forall i\ |y_i| \le 1 \Rightarrow \forall i\ |y_i|^q \le |y_i|^p$ (т.к. $1 \le p \le q$)

$\|y\|_q^q = \sum_i |y_i|^q \le \sum_i |y_i|^p = \|y\|_p^p = 1 \Rightarrow \|y\|_q \le 1 \Rightarrow \|x\|_q \le \|x\|_p$]

Таким образом, $|a_i^T b_j|^p \le \|a_i\|_p^p \cdot \|b_j\|_q^p \le \|a_i\|_p^p \cdot \|b_j\|_p^p \quad \forall i,j$

В итоге получаем:

$$\|AB\|_{p,vec}^p \le \sum_{i=1}^{m}\sum_{j=1}^{l}\|a_i\|_p^p\cdot\|b_j\|_p^p = \left(\sum_{i=1}^{m}\|a_i\|_p^p\right)\cdot\left(\sum_{j=1}^{l}\|b_j\|_p^p\right) = \|A\|_{p,vec}^p\cdot\|B\|_{p,vec}^p,$$ то есть

$$\|AB\|_{p,vec} \le \|A\|_{p,vec}\cdot\|B\|_{p,vec}$$ ∎

(*): • $p<1$ — не подходит, т.к. не является нормой (не выполняется нер-во треугольника)

• $p>2$ — на семинарах приводился контрпример (матрицы из единиц)

№3

Решите задачу 8 для произвольной унитарно-инв. нормы $\|\cdot\|$.

Считайте известным неравенство $\|AB\| \leq \|A\|_2 \cdot \|B\|$

□ • Обобщим решение задачи №8. Начнем со 2-го пункта, докажем нер-во: $\dfrac{\|A^T A - I\|}{\|A\|_2 + 1} \leq \|A - W\|$ (Здесь и ниже обозначения совпадают с №8)

Аналогично №8, можем свести док-во к эквивалентному нер-ву:

$\|D^2 - I\| \overset{?}{\leq} (\|D\|_2 + 1) \cdot \|D - I\|$ (обозначения те же, что в №8)

$\|\underbrace{D^2 - I}_{(D-I)(D+I)}\| \leq \|D + I\|_2 \cdot \|D - I\|$ [утв. из условия]

Тогда получим:

$\|D^2 - I\| \leq \|D+I\|_2 \cdot \|D - I\| \overset{?}{\leq} (\|D\|_2 + 1) \cdot \|D - I\|$

$\|D + I\|_2 \overset{?}{\leq} \|D\|_2 + 1$ — было доказано в №8.

Таким образом, $\dfrac{\|A^T A - I\|}{\|A\|_2 + 1} \leq \|A - W\|$.

• Докажем $\|A - W\| \leq \|A^T A - I\|$

Аналогично №8, можем свести док-во к эквивалентному нер-ву:

$\|D - I\| \overset{?}{\leq} \|D^2 - I\|$

$\|D - I\| \overset{?}{\leq} \|(D - I)(D + I)\| \leq \|D + I\|_2 \cdot \|D - I\|$

$1 \overset{?}{\leq} \|D + I\|_2 = d_1 + 1$, но $d_1 \geq 0$

Значит, $1 \leq \|D + I\|_2$, что доказывает $\|A - W\| \leq \|A^T A - I\|$. ▨