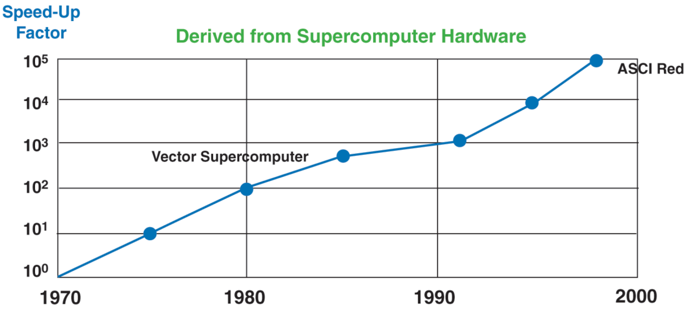
Вычислительная математика.

Нормы векторов и матриц, сингулярное разложение.

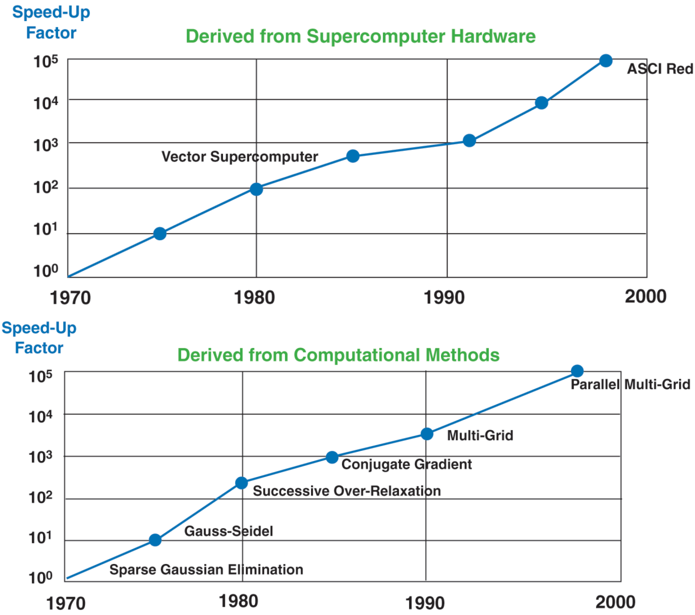
МФТИ

Развитие методов вычислительной ЛА

Развитие методов вычислительной ЛА



Развитие методов вычислительной ЛА



10 важнейших алгоритмов 20-го века

Список, опубликованный в журнале *Computing in Science and Engineering* в 2001 году, содержит 4 алгоритма вычислительной ЛА:

10 важнейших алгоритмов 20-го века

Список, опубликованный в журнале *Computing in Science and Engineering* в 2001 году, содержит 4 алгоритма вычислительной ЛА:

- Вычислительные методы, основанные на разложениях матриц

10 важнейших алгоритмов 20-го века

Список, опубликованный в журнале *Computing in Science and Engineering* в 2001 году, содержит 4 алгоритма вычислительной ЛА:

- Вычислительные методы, основанные на разложениях матриц
- QR алгоритм для вычисления собственных значений

10 важнейших алгоритмов 20-го века

Список, опубликованный в журнале *Computing in Science and Engineering* в 2001 году, содержит 4 алгоритма вычислительной ЛА:

- Вычислительные методы, основанные на разложениях матриц
- QR алгоритм для вычисления собственных значений
- Итерационные методы подпространств Крылова

10 важнейших алгоритмов 20-го века

Список, опубликованный в журнале *Computing in Science and Engineering* в 2001 году, содержит 4 алгоритма вычислительной ЛА:

- Вычислительные методы, основанные на разложениях матриц
- QR алгоритм для вычисления собственных значений
- Итерационные методы подпространств Крылова
- Быстрое преобразование Фурье

Векторные нормированные пространства

Векторные нормированные пространства

- $Ax = b$, y – приближенное решение.
  Как оценить ошибку $E(x, y)$?

Векторные нормированные пространства

- $Ax = b$, y – приближенное решение.
  Как оценить ошибку $E(x, y)$?
- $x \in V$, $\|x\| : V \to \mathbb{R}$ называется *нормой*, если

Векторные нормированные пространства

- $Ax = b$, y – приближенное решение.
  Как оценить ошибку $E(x, y)$?
- $x \in V$, $\|x\| : V \to \mathbb{R}$ называется *нормой*, если
  1. $\|x\| \geq 0 \; \forall x \in V$; $\|x\| = 0 \Leftrightarrow x = 0$

Векторные нормированные пространства

- $Ax = b$, y – приближенное решение.
  Как оценить ошибку $E(x, y)$?
- $x \in V$, $\|x\| : V \to \mathbb{R}$ называется *нормой*, если
  1. $\|x\| \geq 0 \; \forall x \in V$; $\|x\| = 0 \Leftrightarrow x = 0$
  2. $\|\alpha x\| = |\alpha| \|x\|$, $\forall \alpha \in \mathbb{R}$ (или $\mathbb{C}$)

Векторные нормированные пространства

- $Ax = b$, y – приближенное решение.
  Как оценить ошибку $E(x, y)$?
- $x \in V$, $\|x\| : V \to \mathbb{R}$ называется *нормой*, если
  1. $\|x\| \geq 0 \, \forall x \in V$; $\|x\| = 0 \Leftrightarrow x = 0$
  2. $\|\alpha x\| = |\alpha| \|x\|$, $\forall \alpha \in \mathbb{R}$ (или $\mathbb{C}$)
  3. $\|x + y\| \leq \|x\| + \|y\|$ (неравенство треугольника)

Векторные нормированные пространства

- $Ax = b$, y – приближенное решение.
  Как оценить ошибку $E(x, y)$?
- $x \in V$, $\|x\| : V \to \mathbb{R}$ называется *нормой*, если
  1. $\|x\| \geq 0 \; \forall x \in V$; $\|x\| = 0 \Leftrightarrow x = 0$
  2. $\|\alpha x\| = |\alpha| \|x\|$, $\forall \alpha \in \mathbb{R}$ (или $\mathbb{C}$)
  3. $\|x + y\| \leq \|x\| + \|y\|$ (неравенство треугольника)

Определение 2.1 (полное пространство)

Векторное пространство называется *полным* или *банаховым*, если любая фундаментальная последовательность в нём сходится.

p - нормы

- В $\mathbb{C}^n$ (или $\mathbb{R}^n$) $\forall p > 1$ нормой является функция

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

p - нормы

- В $\mathbb{C}^n$ (или $\mathbb{R}^n$) $\forall p > 1$ нормой является функция

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$$

- К p-нормам относят также

$$||x||_\infty = \lim_{p\to\infty} ||x||_p = \max_i |x_i|, \quad \|x\|_1 = \sum_i |x_i|$$

- В $\mathbb{C}^n$ (или $\mathbb{R}^n$) $\forall p > 1$ нормой является функция

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$$

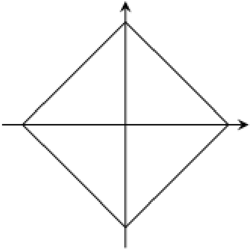
- К p-нормам относят также

$$||x||_\infty = \lim_{p\to\infty} ||x||_p = \max_i |x_i|, \quad \|x\|_1 = \sum_i |x_i|$$

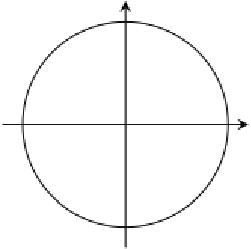
- Неравенство Гёльдера: $\frac{1}{p} + \frac{1}{q} = 1, \; p, q \in [1, \infty]$

$$\sum_{i=1}^{n} \left|x_i y_i\right| \leq \|x\|_p \|y\|_q, \quad \|y\|_q = \max_{x\neq 0} \frac{|x^T y|}{\|x\|_p}$$

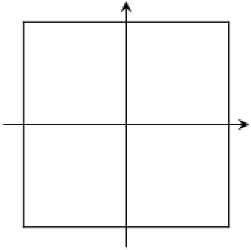
p - нормы, единичные сферы



$p = 1$



$p = 2$



$p = \infty$

Эквивалентность норм

Определение 2.2 (эквивалентные нормы)

Нормы $||\cdot||_{(1)}$ и $||\cdot||_{(2)}$ называются эквивалентными, если $\exists c_1, c_2$: $\forall x \in V$ $c_1||x||_{(1)} \leq ||x||_{(2)} \leq c_2||x||_{(1)}$

Эквивалентность норм

Определение 2.2 (эквивалентные нормы)

Нормы $||\cdot||_{(1)}$ и $||\cdot||_{(2)}$ называются эквивалентными, если $\exists c_1, c_2 \colon \forall x \in V \; c_1||x||_{(1)} \leq ||x||_{(2)} \leq c_2||x||_{(1)}$

Теорема 2.1 Эквивалентность норм

Любые две нормы в конечномерном векторном пространстве эквивалентны.

Эквивалентность норм

Определение 2.2 (эквивалентные нормы)

Нормы $||\cdot||_{(1)}$ и $||\cdot||_{(2)}$ называются эквивалентными, если $\exists c_1, c_2$: $\forall x \in V\ c_1||x||_{(1)} \leq ||x||_{(2)} \leq c_2||x||_{(1)}$

Теорема 2.1 Эквивалентность норм

Любые две нормы в конечномерном векторном пространстве эквивалентны.

Схема доказательства:

- Единичная сфера $S_n = \{x \in \mathbb{R}^n : \|x\|_2 = 1\}$ компактна относительно 2-нормы
- Любая норма $\|\cdot\|_*$ непрерывна относительно 2-нормы
- По теореме Вейерштрасса непрерывная на компакте функция является ограниченной: $\|x\|_* \leq c_2, \|x\|_*^{-1} \leq c_1^{-1}\ \forall x \in S_n$

Скалярное произведение

Пусть на векторном пространстве V для каждой пары векторов x, y определено число (x, y). $(\cdot, \cdot)$ называется *скалярным произведением*, если

Скалярное произведение

Пусть на векторном пространстве V для каждой пары векторов x, y определено число (x, y). $(\cdot, \cdot)$ называется *скалярным произведением*, если

1. $(x, x) \geq 0$, $\forall x$, $(x, x) = 0 \Leftrightarrow x = 0$

Скалярное произведение

Пусть на векторном пространстве V для каждой пары векторов x, y определено число (x, y). $(\cdot, \cdot)$ называется *скалярным произведением*, если

1. $(x, x) \geq 0$, $\forall x$, $(x, x) = 0 \Leftrightarrow x = 0$
2. $(x, y) = \overline{(y, x)}$

Скалярное произведение

Пусть на векторном пространстве V для каждой пары векторов x, y определено число (x, y). $(\cdot, \cdot)$ называется *скалярным произведением*, если

1. $(x, x) \geq 0$, $\forall x$, $(x, x) = 0 \Leftrightarrow x = 0$
2. $(x, y) = \overline{(y, x)}$
3. $(\alpha x, y) = \alpha(x, y)$, α - число

Скалярное произведение

Пусть на векторном пространстве V для каждой пары векторов x, y определено число (x, y). $(\cdot, \cdot)$ называется *скалярным произведением*, если

1. $(x, x) \geq 0$, $\forall x$, $(x, x) = 0 \Leftrightarrow x = 0$
2. $(x, y) = \overline{(y, x)}$
3. $(\alpha x, y) = \alpha(x, y)$, α - число
4. $(x + y, z) = (x, z) + (y, z)$

Скалярное произведение

Пусть на векторном пространстве V для каждой пары векторов x, y определено число (x, y). $(\cdot, \cdot)$ называется *скалярным произведением*, если

1. $(x, x) \geq 0$, $\forall x$, $(x, x) = 0 \Leftrightarrow x = 0$
2. $(x, y) = \overline{(y, x)}$
3. $(\alpha x, y) = \alpha(x, y)$, α - число
4. $(x + y, z) = (x, z) + (y, z)$

- Стандартное с.п. в $\mathbb{C}^n$:

$$(x, y) = \sum_k x_k \overline{y}_k, \ \|x\|_2 = (x, x)^{1/2} = \left(\sum_k |x_k|^2 \right)^{1/2}$$

Скалярное произведение

Пусть на векторном пространстве V для каждой пары векторов x, y определено число (x, y). $(\cdot, \cdot)$ называется *скалярным произведением*, если

1. $(x, x) \geq 0$, $\forall x$, $(x, x) = 0 \Leftrightarrow x = 0$
2. $(x, y) = \overline{(y, x)}$
3. $(\alpha x, y) = \alpha(x, y)$, α - число
4. $(x + y, z) = (x, z) + (y, z)$

- Стандартное с.п. в $\mathbb{C}^n$:

$$(x, y) = \sum_k x_k \overline{y}_k, \ \|x\|_2 = (x, x)^{1/2} = \left(\sum_k |x_k|^2 \right)^{1/2}$$

- $|(x, y)| \leq (x, x)^{1/2} (y, y)^{1/2}$ (неравенство К-Б-Ш)

Основные обозначения (1)

Основные обозначения (1)

Вектор-столбец: $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

Основные обозначения (1)

Вектор-столбец: $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

Матрица $A = \left[\begin{array}{c|c|c} a_1 & \ldots & a_n \end{array}\right] = \begin{bmatrix} s_1^T \\ \hline \ldots \\ \hline s_m^T \end{bmatrix}$

Основные обозначения (1)

Вектор-столбец: $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

Матрица $A = \left[\begin{array}{c|c|c} a_1 & \dots & a_n \end{array}\right] = \left[\begin{array}{c} s_1^T \\ \hline \dots \\ \hline s_m^T \end{array}\right]$

$$b = \left[\begin{array}{c|c|c} a_1 & \dots & a_n \end{array}\right] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} a_1 \end{bmatrix} + \dots + x_n \begin{bmatrix} a_n \end{bmatrix}$$

Основные обозначения (2)

$$c^T = [x_1, \ldots, x_m] \left[\begin{array}{c} s_1^T \\ \hline \ldots \\ \hline s_m^T \end{array}\right] = x_1[\quad s_1^T \quad] + \ldots + x_m[\quad s_m^T \quad]$$

Основные обозначения (2)

$$c^T = [x_1, \dots, x_m] \left[\begin{array}{c} s_1^T \\ \hline \dots \\ \hline s_m^T \end{array}\right] = x_1 [\quad s_1^T \quad] + \ldots + x_m [\quad s_m^T \quad]$$

$A^* = \overline{A^T}$

Основные обозначения (2)

$$c^T = [x_1, \ldots, x_m] \begin{bmatrix} s_1^T \\ \hline \ldots \\ \hline s_m^T \end{bmatrix} = x_1[\quad s_1^T \quad] + \ldots + x_m[\quad s_m^T \quad]$$

$$A^* = \overline{A^T}\ ,\ (x, y) = \sum_k x_k \overline{y_k} = x^T \overline{y} = y^* x \underset{\mathbb{R}^n}{=} y^T x = x^T y$$

Основные обозначения (2)

$$c^T = [x_1, \ldots, x_m] \left[\begin{array}{c} s_1^T \\ \hline \ldots \\ \hline s_m^T \end{array}\right] = x_1[\quad s_1^T \quad] + \ldots + x_m[\quad s_m^T \quad]$$

$$A^* = \overline{A^T}\ ,\ (x, y) = \sum_k x_k \overline{y_k} = x^T \overline{y} = y^* x \underset{\mathbb{R}^n}{=} y^T x = x^T y$$

$$AB = \left[\begin{array}{c} s_1^T \\ \hline \ldots \\ \hline s_n^T \end{array}\right] \left[\begin{array}{c|c|c} b_1 & \ldots & b_n \end{array}\right] = \begin{bmatrix} s_1^T b_1 & & s_1^T b_n \\ & \ddots & \\ s_n^T b_1 & & s_n^T b_n \end{bmatrix}$$

Основные обозначения (2)

$$c^T = [x_1, \ldots, x_m] \begin{bmatrix} s_1^T \\ \hline \ldots \\ \hline s_m^T \end{bmatrix} = x_1 [\quad s_1^T \quad] + \ldots + x_m [\quad s_m^T \quad]$$

$$A^* = \overline{A^T}\ ,\ (x, y) = \sum_k x_k \overline{y_k} = x^T \overline{y} = y^* x \underset{\mathbb{R}^n}{=} y^T x = x^T y$$

$$AB = \begin{bmatrix} s_1^T \\ \hline \ldots \\ \hline s_n^T \end{bmatrix} \left[\begin{array}{c|c|c} b_1 & \ldots & b_n \end{array} \right] = \begin{bmatrix} s_1^T b_1 & & s_1^T b_n \\ & \ddots & \\ s_n^T b_1 & & s_n^T b_n \end{bmatrix}$$

$$AB = \left[\begin{array}{c|c|c} a_1 & \ldots & a_n \end{array} \right] \begin{bmatrix} q_1^T \\ \hline \ldots \\ \hline q_n^T \end{bmatrix} = a_1 q_1^T + \ldots + a_n q_n^T$$

Основные обозначения (3)

Собственные числа и векторы ($A \in \mathbb{C}^{n\times n}$):

$$Au_k = \lambda_k u_k,\ \det(A - \lambda I) = 0 \tag{1}$$

(2)

Основные обозначения (3)

Собственные числа и векторы ($A \in \mathbb{C}^{n\times n}$):

$$Au_k = \lambda_k u_k, \ \det(A - \lambda I) = 0 \tag{1}$$

$$U = [u_1| \dots |u_n], \Lambda = \mathsf{diag}(\lambda_1, \dots, \lambda_n) \Rightarrow AU = U\Lambda \tag{2}$$

Основные обозначения (3)

Собственные числа и векторы ($A \in \mathbb{C}^{n\times n}$):

$$Au_k = \lambda_k u_k, \ \det(A - \lambda I) = 0 \tag{1}$$

$$U = [u_1|\dots|u_n], \Lambda = \mathsf{diag}(\lambda_1, \dots, \lambda_n) \Rightarrow AU = U\Lambda \tag{2}$$

Блочное умножение матриц:

$$A = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right], B = \left[\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right] \Rightarrow$$

$$C = AB = \left[\begin{array}{c|c} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ \hline A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{array}\right]$$

Матричные нормы

Матричные нормы

- $\mathbb{C}^{m\times n}$ – векторное пространство $\Rightarrow$ можно взять любую векторную норму

Матричные нормы

- $\mathbb{C}^{m \times n}$ – векторное пространство $\Rightarrow$ можно взять любую векторную норму
- *Матричная норма* удовлетворяет дополнительному свойству

Матричные нормы

- $\mathbb{C}^{m\times n}$ – векторное пространство $\Rightarrow$ можно взять любую векторную норму
- *Матричная норма* удовлетворяет дополнительному свойству

Определение 2.3 Матричная норма

Пусть каждой матрице A поставлено в соответствие число $||A||$. $||\cdot||$ называется матричной нормой, если:

Матричные нормы

- $\mathbb{C}^{m\times n}$ – векторное пространство $\Rightarrow$ можно взять любую векторную норму
- *Матричная норма* удовлетворяет дополнительному свойству

Определение 2.3 Матричная норма

Пусть каждой матрице A поставлено в соответствие число $||A||$. $||\cdot||$ называется матричной нормой, если:

1. на $\mathbb{C}^{m\times n}$ $||\cdot||$ является векторной нормой

Матричные нормы

- $\mathbb{C}^{m\times n}$ – векторное пространство $\Rightarrow$ можно взять любую векторную норму
- *Матричная норма* удовлетворяет дополнительному свойству

Определение 2.3 Матричная норма

Пусть каждой матрице A поставлено в соответствие число $||A||$. $||\cdot||$ называется матричной нормой, если:

1. на $\mathbb{C}^{m\times n}$ $||\cdot||$ является векторной нормой
2. для любых матриц A и B допускающих умножение выполняется:
   $||AB|| \leq ||A||\ ||B||$ (субмультипликативность)

Пример: норма Фробениуса

$$||A||_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2}$$

Доказательство субмультипликативности:

Пример: норма Фробениуса

$$||A||_F = \left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2\right)^{1/2}$$

Доказательство субмультипликативности:

$$A = [a_1|, \dots, |a_n] \in \mathbb{C}^{m\times n},\ B = \begin{bmatrix} b_1^T \\ \hline \dots \\ \hline b_n^T \end{bmatrix} \quad AB = \sum_{i=1}^{n} a_i b_i^T$$

Пример: норма Фробениуса

$$||A||_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2}$$

Доказательство субмультипликативности:

$$A = [a_1|, \dots, |a_n] \in \mathbb{C}^{m \times n}, \; B = \begin{bmatrix} b_1^T \\ \hline \dots \\ \hline b_n^T \end{bmatrix} \quad AB = \sum_{i=1}^{n} a_i b_i^T$$

$$\|AB\|_F = \| \sum_{i=1}^{n} a_i b_i^T \|_F$$

Пример: норма Фробениуса

$$||A||_F = \left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2\right)^{1/2}$$

Доказательство субмультипликативности:

$$A = [a_1|, \dots, |a_n] \in \mathbb{C}^{m\times n},\ B = \begin{bmatrix} b_1^T \\ \hline \dots \\ \hline b_n^T \end{bmatrix} \quad AB = \sum_{i=1}^{n} a_i b_i^T$$

$$\|AB\|_F = \|\sum_{i=1}^{n} a_i b_i^T\|_F \leq \sum_i \|a_i b_i^T\|_F$$

Пример: норма Фробениуса

$$||A||_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2}$$

Доказательство субмультипликативности:

$$A = [a_1|, \dots, |a_n] \in \mathbb{C}^{m \times n}, \ B = \begin{bmatrix} b_1^T \\ \hline \dots \\ \hline b_n^T \end{bmatrix} \quad AB = \sum_{i=1}^{n} a_i b_i^T$$

$$\|AB\|_F = \| \sum_{i=1}^{n} a_i b_i^T \|_F \leq \sum_{i} \|a_i b_i^T\|_F = \sum_{i} \left(\sum_{k,l} |a_{ki} b_{il}|^2 \right)^{1/2}$$

Пример: норма Фробениуса

$$||A||_F = \left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2\right)^{1/2}$$

Доказательство субмультипликативности:

$$A = [a_1|, \dots, |a_n] \in \mathbb{C}^{m\times n},\ B = \begin{bmatrix} b_1^T \\ \hline \dots \\ \hline b_n^T \end{bmatrix} \quad AB = \sum_{i=1}^{n} a_i b_i^T$$

$$\|AB\|_F = \|\sum_{i=1}^{n} a_i b_i^T\|_F \leq \sum_i \|a_i b_i^T\|_F = \sum_i \left(\sum_{k,l}|a_{ki}b_{il}|^2\right)^{1/2}$$

$$= \sum_i \|a_i\|_2 \|b_i\|_2$$

Пример: норма Фробениуса

$$||A||_F = \left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2\right)^{1/2}$$

Доказательство субмультипликативности:

$$A = [a_1|,\dots,|a_n] \in \mathbb{C}^{m\times n},\ B = \left[\begin{array}{c} b_1^T \\ \hline \dots \\ \hline b_n^T \end{array}\right] \quad AB = \sum_{i=1}^{n} a_i b_i^T$$

$$\|AB\|_F = \|\sum_{i=1}^{n} a_i b_i^T\|_F \le \sum_i \|a_i b_i^T\|_F = \sum_i \left(\sum_{k,l}|a_{ki}b_{il}|^2\right)^{1/2}$$

$$= \sum_i \|a_i\|_2\|b_i\|_2 \le \left(\sum_i \|a_i\|_2^2\right)^{1/2}\left(\sum_i \|b_i\|_2^2\right)^{1/2}$$

Пример: норма Фробениуса

$$||A||_F = \left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2\right)^{1/2}$$

Доказательство субмультипликативности:

$$A = [a_1|,\ldots,|a_n] \in \mathbb{C}^{m\times n},\ B = \begin{bmatrix} b_1^T \\ \hline \ldots \\ \hline b_n^T \end{bmatrix} \quad AB = \sum_{i=1}^{n} a_i b_i^T$$

$$\|AB\|_F = \|\sum_{i=1}^{n} a_i b_i^T\|_F \leq \sum_i \|a_i b_i^T\|_F = \sum_i \left(\sum_{k,l}|a_{ki}b_{il}|^2\right)^{1/2}$$

$$= \sum_i \|a_i\|_2\|b_i\|_2 \leq \left(\sum_i \|a_i\|_2^2\right)^{1/2}\left(\sum_i \|b_i\|_2^2\right)^{1/2} = \|A\|_F\|B\|_F$$

Пример: нарушение свойства субмультипликативности

- $\|A\| = \max\limits_{i,j} |a_{ij}|$

Пример: нарушение свойства субмультипликативности

- $\|A\| = \max\limits_{i,j} |a_{ij}|$
-
$$A = B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \; AB = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$$
$$\|AB\| = 2 > \|A\|\|B\| = 1 \cdot 1 = 1$$

Операторные нормы

- Линейный оператор – отображение $P: \ V_1 \to V_2$

Операторные нормы

- Линейный оператор – отображение $P: \ V_1 \to V_2$

Определение 2.4 (норма оператора)

Нормой оператора $P: \ V_1 \to V_2$ называется

$$||P|| = \sup_{x \neq 0} \frac{||Px||_{(2)}}{||x||_{(1)}} = \sup_{\|x\|_{(1)}=1} ||Px||_{(2)}$$

Операторные нормы

- Линейный оператор – отображение $P: \ V_1 \to V_2$

Определение 2.4 (норма оператора)

Нормой оператора $P: \ V_1 \to V_2$ называется

$$||P|| = \sup_{x \neq 0} \frac{||Px||_{(2)}}{||x||_{(1)}} = \sup_{\|x\|_{(1)}=1} ||Px||_{(2)}$$

- $||A|| = \max\limits_{x \neq 0} \dfrac{||Ax||_{(2)}}{||x||_{(1)}} = \max\limits_{\|x\|_{(1)}=1} ||Ax||_{(2)}$

Операторные нормы

- Линейный оператор – отображение $P: \ V_1 \to V_2$

Определение 2.4 (норма оператора)

Нормой оператора $P: \ V_1 \to V_2$ называется

$$||P|| = \sup_{x \neq 0} \frac{||Px||_{(2)}}{||x||_{(1)}} = \sup_{\|x\|_{(1)}=1} ||Px||_{(2)}$$

- $$||A|| = \max_{x \neq 0} \frac{||Ax||_{(2)}}{||x||_{(1)}} = \max_{\|x\|_{(1)}=1} ||Ax||_{(2)}$$

$$\|AB\| = \max_{x \neq 0, Bx \neq 0} \frac{\|ABx\|_{(3)}}{\|x\|_{(1)}}$$

Операторные нормы

- Линейный оператор – отображение $P: V_1 \to V_2$

Определение 2.4 (норма оператора)

Нормой оператора $P: V_1 \to V_2$ называется

$$||P|| = \sup_{x \neq 0} \frac{||Px||_{(2)}}{||x||_{(1)}} = \sup_{\|x\|_{(1)}=1} ||Px||_{(2)}$$

- $$||A|| = \max_{x \neq 0} \frac{||Ax||_{(2)}}{||x||_{(1)}} = \max_{\|x\|_{(1)}=1} ||Ax||_{(2)}$$

$$\|AB\| = \max_{x \neq 0, Bx \neq 0} \frac{\|ABx\|_{(3)}}{\|x\|_{(1)}} = \max_{x \neq 0, Bx \neq 0} \frac{\|ABx\|_{(3)}}{\|Bx\|_{(2)}} \frac{\|Bx\|_{(2)}}{\|x\|_{(1)}}$$

Операторные нормы

- Линейный оператор – отображение $P:\ V_1 \to V_2$

Определение 2.4 (норма оператора)

Нормой оператора $P:\ V_1 \to V_2$ называется

$$||P|| = \sup_{x \neq 0} \frac{||Px||_{(2)}}{||x||_{(1)}} = \sup_{\|x\|_{(1)}=1} ||Px||_{(2)}$$

- $$||A|| = \max_{x \neq 0} \frac{||Ax||_{(2)}}{||x||_{(1)}} = \max_{\|x\|_{(1)}=1} ||Ax||_{(2)}$$

$$\|AB\| = \max_{x\neq 0, Bx\neq 0} \frac{\|ABx\|_{(3)}}{\|x\|_{(1)}} = \max_{x\neq 0, Bx\neq 0} \frac{\|ABx\|_{(3)}}{\|Bx\|_{(2)}} \frac{\|Bx\|_{(2)}}{\|x\|_{(1)}}$$
$$\leq \max_{x\neq 0, Bx\neq 0} \frac{\|A(Bx)\|_{(3)}}{\|Bx\|_{(2)}} \max_{x\neq 0} \frac{\|Bx\|_{(2)}}{\|x\|_{(1)}}$$

Операторные нормы

- Линейный оператор – отображение $P: V_1 \to V_2$

Определение 2.4 (норма оператора)

Нормой оператора $P: V_1 \to V_2$ называется

$$||P|| = \sup_{x \neq 0} \frac{||Px||_{(2)}}{||x||_{(1)}} = \sup_{\|x\|_{(1)}=1} ||Px||_{(2)}$$

- $||A|| = \max\limits_{x \neq 0} \dfrac{||Ax||_{(2)}}{||x||_{(1)}} = \max\limits_{\|x\|_{(1)}=1} ||Ax||_{(2)}$

$$\|AB\| = \max_{x \neq 0, Bx \neq 0} \frac{\|ABx\|_{(3)}}{\|x\|_{(1)}} = \max_{x \neq 0, Bx \neq 0} \frac{\|ABx\|_{(3)}}{\|Bx\|_{(2)}} \frac{\|Bx\|_{(2)}}{\|x\|_{(1)}}$$
$$\leq \max_{x \neq 0, Bx \neq 0} \frac{\|A(Bx)\|_{(3)}}{\|Bx\|_{(2)}} \max_{x \neq 0} \frac{\|Bx\|_{(2)}}{\|x\|_{(1)}} \leq \|A\|\|B\|$$

Матричные нормы, порожденные p-нормами

Матричные нормы, порожденные p-нормами

- $||A||_1 = \max\limits_{1 \le j \le n} \sum\limits_{i=1}^{m} |a_{ij}|$
  (максимальная сумма модулей по столбцу)

Матричные нормы, порожденные p-нормами

- $$||A||_1 = \max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}|$$
  (максимальная сумма модулей по столбцу)
- $$||A||_\infty = \max_{1\leq i\leq m} \sum_{j=1}^{n} |a_{ij}|$$
  (максимальная сумма модулей по строке)

Матричные нормы, порожденные p-нормами

- $||A||_1 = \max\limits_{1\le j\le n} \sum\limits_{i=1}^{m} |a_{ij}|$
  (максимальная сумма модулей по столбцу)
- $||A||_\infty = \max\limits_{1\le i\le m} \sum\limits_{j=1}^{n} |a_{ij}|$
  (максимальная сумма модулей по строке)
- $||A||_2 = \max\limits_{k} \sigma_k$
  (максимальное сингулярное число A, $\sigma_k = \sqrt{\lambda_k(A^*A)}$).

Изометричные матрицы

Изометричные матрицы

Определение 2.5 (изометричные матрицы)

Матрица $Q \in \mathbb{C}^{n\times n}$ *изометрична* относительно нормы $\|\cdot\|$ на $\mathbb{C}^n$, если $\|Qx\| = \|x\| \; \forall x \in \mathbb{C}^n$

Определение 2.5 (изометричные матрицы)

Матрица $Q \in \mathbb{C}^{n \times n}$ *изометрична* относительно нормы $\|\cdot\|$ на $\mathbb{C}^n$, если $\|Qx\| = \|x\| \; \forall x \in \mathbb{C}^n$

- $p, q \in [1, \infty] \quad 1/p + 1/q = 1.$

Изометричные матрицы

Определение 2.5 (изометричные матрицы)

Матрица $Q \in \mathbb{C}^{n\times n}$ *изометрична* относительно нормы $\|\cdot\|$ на $\mathbb{C}^n$, если $\|Qx\| = \|x\| \ \forall x \in \mathbb{C}^n$

- $p, q \in [1, \infty] \quad 1/p + 1/q = 1.$
- $\|Qx\|_p = \|x\|_p \Rightarrow p$-норма столбцов равна 1.

Изометричные матрицы

Определение 2.5 (изометричные матрицы)

Матрица $Q \in \mathbb{C}^{n \times n}$ *изометрична* относительно нормы $\|\cdot\|$ на $\mathbb{C}^n$, если $\|Qx\| = \|x\| \ \forall x \in \mathbb{C}^n$

- $p, q \in [1, \infty] \quad 1/p + 1/q = 1.$
- $\|Qx\|_p = \|x\|_p \Rightarrow p$-норма столбцов равна 1.
- Неравенство Гёльдера $\left|\sum_{i=1}^n x_i y_i\right| \leq \|x\|_p \|y\|_q$:

Изометричные матрицы

Определение 2.5 (изометричные матрицы)

Матрица $Q \in \mathbb{C}^{n\times n}$ *изометрична* относительно нормы $\|\cdot\|$ на $\mathbb{C}^n$, если $\|Qx\| = \|x\|\ \forall x \in \mathbb{C}^n$

- $p, q \in [1, \infty] \quad 1/p + 1/q = 1.$
- $\|Qx\|_p = \|x\|_p \Rightarrow p$-норма столбцов равна 1.
- Неравенство Гёльдера $\left|\sum_{i=1}^n x_i y_i\right| \le \|x\|_p \|y\|_q$:

$$\|Q^T y\|_q \underset{\text{нер-во Г.}}{=} \max_{x\neq 0} \frac{|y^T Q x|}{\|x\|_p}$$

Изометричные матрицы

Определение 2.5 (изометричные матрицы)

Матрица $Q \in \mathbb{C}^{n \times n}$ *изометрична* относительно нормы $\|\cdot\|$ на $\mathbb{C}^n$, если $\|Qx\| = \|x\| \ \forall x \in \mathbb{C}^n$

- $p, q \in [1, \infty] \quad 1/p + 1/q = 1.$
- $\|Qx\|_p = \|x\|_p \Rightarrow p$-норма столбцов равна 1.
- Неравенство Гёльдера $\left|\sum_{i=1}^n x_i y_i\right| \le \|x\|_p \|y\|_q$:

$$\|Q^T y\|_q \underset{\text{нер-во Г.}}{=} \max_{x \neq 0} \frac{|y^T Q x|}{\|x\|_p} = \max_{x \neq 0} \frac{|y^T Q x|}{\|Qx\|_p} =$$

Изометричные матрицы

Определение 2.5 (изометричные матрицы)

Матрица $Q \in \mathbb{C}^{n \times n}$ *изометрична* относительно нормы $\|\cdot\|$ на $\mathbb{C}^n$, если $\|Qx\| = \|x\| \ \forall x \in \mathbb{C}^n$

- $p, q \in [1, \infty] \quad 1/p + 1/q = 1.$
- $\|Qx\|_p = \|x\|_p \Rightarrow p$-норма столбцов равна 1.
- Неравенство Гёльдера $\left|\sum_{i=1}^n x_i y_i\right| \le \|x\|_p \|y\|_q$:

$$\|Q^T y\|_q \underset{\text{нер-во Г.}}{=} \max_{x \neq 0} \frac{|y^T Q x|}{\|x\|_p} = \max_{x \neq 0} \frac{|y^T Q x|}{\|Qx\|_p} = \\ = \max_{z \neq 0} \frac{|y^T z|}{\|z\|_p}$$

Изометричные матрицы

Определение 2.5 (изометричные матрицы)

Матрица $Q \in \mathbb{C}^{n\times n}$ *изометрична* относительно нормы $\|\cdot\|$ на $\mathbb{C}^n$, если $\|Qx\| = \|x\| \ \forall x \in \mathbb{C}^n$

- $p, q \in [1, \infty] \quad 1/p + 1/q = 1.$
- $\|Qx\|_p = \|x\|_p \Rightarrow p$-норма столбцов равна 1.
- Неравенство Гёльдера $\left|\sum_{i=1}^n x_i y_i\right| \le \|x\|_p \|y\|_q$:

$$\|Q^T y\|_q \underset{\text{нер-во Г.}}{=} \max_{x\neq 0} \frac{|y^T Q x|}{\|x\|_p} = \max_{x\neq 0} \frac{|y^T Q x|}{\|Qx\|_p} = $$

$$= \max_{z\neq 0} \frac{|y^T z|}{\|z\|_p} \underset{\text{нер-во Г.}}{=} \|y\|_q$$

Изометричные матрицы

Определение 2.5 (изометричные матрицы)

Матрица $Q \in \mathbb{C}^{n\times n}$ *изометрична* относительно нормы $\|\cdot\|$ на $\mathbb{C}^n$, если $\|Qx\| = \|x\| \ \forall x \in \mathbb{C}^n$

- $p, q \in [1, \infty] \quad 1/p + 1/q = 1.$
- $\|Qx\|_p = \|x\|_p \Rightarrow p$-норма столбцов равна 1.
- Неравенство Гёльдера $\left|\sum_{i=1}^n x_i y_i\right| \le \|x\|_p \|y\|_q$:

$$\|Q^T y\|_q \underset{\text{нер-во Г.}}{=} \max_{x\neq 0} \frac{|y^T Q x|}{\|x\|_p} = \max_{x\neq 0} \frac{|y^T Q x|}{\|Qx\|_p} = \max_{z\neq 0} \frac{|y^T z|}{\|z\|_p} \underset{\text{нер-во Г.}}{=} \|y\|_q$$

q-норма строк Q равна 1.

Изометричные матрицы, $p = 2$

Изометричные матрицы, $p = 2$

- Пусть Q (и Q^T) сохраняют 2-норму:

Изометричные матрицы, $p = 2$

- Пусть Q (и Q^T) сохраняют 2-норму:

$$\|Q^T \overline{q}_k\|_2^2 = \|\overline{q}_k\|_2^2 = \|q_k\|_2^2 = 1$$

Изометричные матрицы, $p = 2$

- Пусть Q (и Q^T) сохраняют 2-норму:

$$\|Q^T \overline{q}_k\|_2^2 = \|\overline{q}_k\|_2^2 = \|q_k\|_2^2 = 1 = \left\| \begin{bmatrix} q_1^T \overline{q_k} \\ \cdots \\ q_n^T \overline{q_k} \end{bmatrix} \right\|_2^2 =$$

$$= \sum_{i=1}^{n} |(q_k, q_i)|^2 = 1 + \sum_{i=1, i \neq k}^{n} |(q_k, q_i)|^2$$

Изометричные матрицы, $p = 2$

- Пусть Q (и Q^T) сохраняют 2-норму:

$$\|Q^T \overline{q}_k\|_2^2 = \|\overline{q}_k\|_2^2 = \|q_k\|_2^2 = 1 = \left\| \begin{bmatrix} q_1^T \overline{q_k} \\ \cdots \\ q_n^T \overline{q_k} \end{bmatrix} \right\|_2^2 =$$

$$= \sum_{i=1}^{n} |(q_k, q_i)|^2 = 1 + \sum_{i=1, i \neq k}^{n} |(q_k, q_i)|^2$$

- $\Rightarrow (q_i, q_j) = \delta_{ij} \Leftrightarrow Q^* Q = I$

Изометричные матрицы, $p = 2$

- Пусть Q (и Q^T) сохраняют 2-норму:

$$\|Q^T \overline{q}_k\|_2^2 = \|\overline{q}_k\|_2^2 = \|q_k\|_2^2 = 1 = \left\| \begin{bmatrix} q_1^T \overline{q_k} \\ \cdots \\ q_n^T \overline{q_k} \end{bmatrix} \right\|_2^2 =$$

$$= \sum_{i=1}^{n} |(q_k, q_i)|^2 = 1 + \sum_{i=1, i \neq k}^{n} |(q_k, q_i)|^2$$

- $\Rightarrow (q_i, q_j) = \delta_{ij} \Leftrightarrow Q^* Q = I$
- аналогично $QQ^* = I$, $Q^{-1} = Q^*$ (Q^T в $\mathbb{R}^{n \times n}$)

Изометричные матрицы, $p = 2$

- Пусть Q (и Q^T) сохраняют 2-норму:

$$\|Q^T\overline{q}_k\|_2^2 = \|\overline{q}_k\|_2^2 = \|q_k\|_2^2 = 1 = \left\| \begin{bmatrix} q_1^T\overline{q_k} \\ \cdots \\ q_n^T\overline{q_k} \end{bmatrix} \right\|_2^2 =$$

$$= \sum_{i=1}^{n} |(q_k, q_i)|^2 = 1 + \sum_{i=1, i\neq k}^{n} |(q_k, q_i)|^2$$

- $\Rightarrow (q_i, q_j) = \delta_{ij} \Leftrightarrow Q^*Q = I$
- аналогично $QQ^* = I$, $Q^{-1} = Q^*$ (Q^T в $\mathbb{R}^{n\times n}$)

Определение 2.6 (унитарная матрица)

Матрица $Q \in \mathbb{C}^{n\times n}$: $Q^{-1} = Q^*$ называется *унитарной*.

Унитарные матрицы

Унитарные матрицы

- Унитарные матрицы и только они сохраняют 2-норму

Унитарные матрицы

- Унитарные матрицы и только они сохраняют 2-норму
- Обратная к унитарной матрице – унитарная матрица:

$$Q^{-1} = Q^* \Rightarrow (Q^{-1})^{-1} = (Q^*)^{-1} = (Q^{-1})^*$$

Унитарные матрицы

- Унитарные матрицы и только они сохраняют 2-норму
- Обратная к унитарной матрице – унитарная матрица:

$$Q^{-1} = Q^* \Rightarrow (Q^{-1})^{-1} = (Q^*)^{-1} = (Q^{-1})^*$$

- Произведение унитарных матриц - унитарная матрица

$$(Q_1 Q_2)^{-1} = Q_2^{-1} Q_1^{-1} = Q_2^* Q_1^* = (Q_1 Q_2)^*$$

Унитарные матрицы

- Унитарные матрицы и только они сохраняют 2-норму
- Обратная к унитарной матрице – унитарная матрица:

$$Q^{-1} = Q^* \Rightarrow (Q^{-1})^{-1} = (Q^*)^{-1} = (Q^{-1})^*$$

- Произведение унитарных матриц - унитарная матрица

$$(Q_1 Q_2)^{-1} = Q_2^{-1} Q_1^{-1} = Q_2^* Q_1^* = (Q_1 Q_2)^*$$

- В действительном случае $Q \in \mathbb{R}^{n \times n}$: $Q^{-1} = Q^T$ называется *ортогональной*.

Теорема Шура

Теорема 2.2 (разложение Шура)

$\forall\ A \in \mathbb{C}^{n \times n}$ с собственными значениями $\lambda_1, \dots, \lambda_n$ $\exists$ унитарная Q такая, что:

$$Q^* A Q = T = \begin{bmatrix} \lambda_1 & \times & \cdots & \times \\ 0 & \lambda_2 & \dots & \times \\ \vdots & \ddots & \ddots & \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}, \quad A = QTQ^*$$

Доказательство теоремы Шура

Доказательство теоремы Шура

1. Пусть $Av_1 = \lambda_1 v_1$, $\|v_1\|_2 = 1$

Доказательство теоремы Шура

1. Пусть $Av_1 = \lambda_1 v_1$, $\|v_1\|_2 = 1$
2. Дополним v_1 до ортонормированного базиса:
   $V_1 = [v_1|v_2|\dots|v_n]$ – унитарная.

Доказательство теоремы Шура

1. Пусть $Av_1 = \lambda_1 v_1$, $\|v_1\|_2 = 1$
2. Дополним v_1 до ортонормированного базиса:
   $V_1 = [v_1|v_2|\dots|v_n]$ – унитарная.
3. $$V_1^* A V_1 = \begin{bmatrix} v_1^* \\ v_2^* \\ \vdots \\ v_n^* \end{bmatrix} [\lambda v_1 | A v_2 | \dots] = \begin{bmatrix} \lambda_1 & \times & \dots & \times \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{bmatrix}$$

Доказательство теоремы Шура

1. Пусть $Av_1 = \lambda_1 v_1$, $\|v_1\|_2 = 1$
2. Дополним v_1 до ортонормированного базиса: $V_1 = [v_1|v_2|\dots|v_n]$ – унитарная.
3. $$V_1^* A V_1 = \begin{bmatrix} v_1^* \\ v_2^* \\ \vdots \\ v_n^* \end{bmatrix} \left[\lambda v_1 | A v_2 | \dots \right] = \begin{bmatrix} \lambda_1 & \times & \dots & \times \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{bmatrix}$$
4. $\det(A - \lambda I) =$

Доказательство теоремы Шура

1. Пусть $Av_1 = \lambda_1 v_1$, $\|v_1\|_2 = 1$
2. Дополним v_1 до ортонормированного базиса:
   $V_1 = [v_1|v_2|\dots|v_n]$ – унитарная.
3. $$V_1^* A V_1 = \begin{bmatrix} v_1^* \\ v_2^* \\ \vdots \\ v_n^* \end{bmatrix} [\lambda v_1 | A v_2 | \dots] = \begin{bmatrix} \lambda_1 & \times & \dots & \times \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{bmatrix}$$
4. $\det(A - \lambda I) = \det(V_1^*) \det(A - \lambda I) \det(V_1) =$

Доказательство теоремы Шура

1. Пусть $Av_1 = \lambda_1 v_1$, $\|v_1\|_2 = 1$
2. Дополним v_1 до ортонормированного базиса: $V_1 = [v_1|v_2|\dots|v_n]$ – унитарная.
3. $$V_1^* A V_1 = \begin{bmatrix} v_1^* \\ v_2^* \\ \vdots \\ v_n^* \end{bmatrix} [\lambda v_1 | A v_2 | \dots] = \begin{bmatrix} \lambda_1 & \times & \dots & \times \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{bmatrix}$$
4. $\det(A - \lambda I) = \det(V_1^*)\det(A - \lambda I)\det(V_1) = \det(V_1^* A V_1 - \lambda I) =$

Доказательство теоремы Шура

1. Пусть $Av_1 = \lambda_1 v_1$, $\|v_1\|_2 = 1$
2. Дополним v_1 до ортонормированного базиса: $V_1 = [v_1|v_2|\dots|v_n]$ – унитарная.
3.

$$V_1^* A V_1 = \begin{bmatrix} v_1^* \\ v_2^* \\ \vdots \\ v_n^* \end{bmatrix} \left[\lambda v_1 | A v_2 | \dots \right] = \begin{bmatrix} \lambda_1 & \times & \dots & \times \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{bmatrix}$$

4. $\det(A - \lambda I) = \det(V_1^*)\det(A - \lambda I)\det(V_1) = \det(V_1^* A V_1 - \lambda I) = (\lambda_1 - \lambda)\det(A_1 - \lambda I_{n-1})$

Доказательство теоремы Шура

1. Пусть $Av_1 = \lambda_1 v_1$, $\|v_1\|_2 = 1$
2. Дополним v_1 до ортонормированного базиса: $V_1 = [v_1|v_2|\dots|v_n]$ – унитарная.
3. $$V_1^* A V_1 = \begin{bmatrix} v_1^* \\ v_2^* \\ \vdots \\ v_n^* \end{bmatrix} [\lambda v_1 | A v_2 | \dots] = \begin{bmatrix} \lambda_1 & \times & \dots & \times \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{bmatrix}$$
4. $\det(A - \lambda I) = \det(V_1^*)\det(A - \lambda I)\det(V_1) = \det(V_1^* A V_1 - \lambda I) = (\lambda_1 - \lambda)\det(A_1 - \lambda I_{n-1})$
5. Далее – по индукции для A_1 $\square$

Нормальные матрицы

Нормальные матрицы

Определение 2.7 (нормальная матрица)

Матрица $A \in \mathbb{C}^{n \times n}$ называется *нормальной*, если $A^* A = A A^*$

Нормальные матрицы

Определение 2.7 (нормальная матрица)

Матрица $A \in \mathbb{C}^{n\times n}$ называется *нормальной*, если $A^*A = AA^*$

Примеры нормальных матриц:

Нормальные матрицы

Определение 2.7 (нормальная матрица)

Матрица $A \in \mathbb{C}^{n\times n}$ называется *нормальной*, если $A^*A = AA^*$

Примеры нормальных матриц:

- Эрмитовы (самосопряженные) матрицы: $H^* = H$

Нормальные матрицы

Определение 2.7 (нормальная матрица)

Матрица $A \in \mathbb{C}^{n \times n}$ называется *нормальной*, если $A^*A = AA^*$

Примеры нормальных матриц:

- Эрмитовы (самосопряженные) матрицы: $H^* = H$
- Унитарные матрицы: $U^* = U^{-1}$

Нормальные матрицы

Определение 2.7 (нормальная матрица)

Матрица $A \in \mathbb{C}^{n\times n}$ называется *нормальной*, если $A^*A = AA^*$

Примеры нормальных матриц:

- Эрмитовы (самосопряженные) матрицы: $H^* = H$
- Унитарные матрицы: $U^* = U^{-1}$

Теорема 2.3 (спектральное разложение нормальной матрицы)

Матрица $A \in \mathbb{C}^{n\times n}$ нормальная $\iff$ в $\mathbb{C}^n$ $\exists$ ортонормированный базис из ее собственных векторов:

$$AU = U\Lambda,\; U^*AU = \Lambda,\; A = U\Lambda U^*$$

Доказательство теоремы ($\Leftarrow$)

- ($\Leftarrow$) $U = [u_1| \dots |u_n]$ – матрица с.в.

$$AU = U\Lambda,\ U^* = U^{-1} \Rightarrow A = U\Lambda U^*$$

Доказательство теоремы $(\Leftarrow)$

- $(\Leftarrow)$ $U = [u_1| \dots |u_n]$ – матрица с.в.

$$AU = U\Lambda,\ U^* = U^{-1} \Rightarrow A = U\Lambda U^*$$

-

$$A^*A = (U\Lambda U^*)^* \, (U\Lambda U^*)$$

Доказательство теоремы ($\Leftarrow$)

- ($\Leftarrow$) $U = [u_1| \dots |u_n]$ – матрица с.в.

$$AU = U\Lambda, \ U^* = U^{-1} \Rightarrow A = U\Lambda U^*$$

-

$$A^*A = (U\Lambda U^*)^* \ (U\Lambda U^*) = U\Lambda^* \underline{U^*U}\Lambda U^*$$

Доказательство теоремы ($\Leftarrow$)

- ($\Leftarrow$) $U = [u_1|\dots|u_n]$ – матрица с.в.

$$AU = U\Lambda,\ U^* = U^{-1} \Rightarrow A = U\Lambda U^*$$

-

$$\begin{aligned} A^*A &= (U\Lambda U^*)^*\ (U\Lambda U^*) = U\Lambda^* \underline{U^*U} \Lambda U^* \\ &= U\Lambda^*\Lambda U^* \end{aligned}$$

Доказательство теоремы ($\Leftarrow$)

- ($\Leftarrow$) $U = [u_1| \dots |u_n]$ – матрица с.в.

$$AU = U\Lambda,\ U^* = U^{-1} \Rightarrow A = U\Lambda U^*$$

-

$$\begin{aligned} A^*A &= (U\Lambda U^*)^* \, (U\Lambda U^*) = U\Lambda^* \underline{U^*U}\Lambda U^* \\ &= U\Lambda^*\Lambda U^* = U|\Lambda|^2 U^* \end{aligned}$$

Доказательство теоремы ($\Leftarrow$)

- ($\Leftarrow$) $U = [u_1| \dots |u_n]$ – матрица с.в.

$$AU = U\Lambda,\ U^* = U^{-1} \Rightarrow A = U\Lambda U^*$$

- $$\begin{aligned} A^*A &= (U\Lambda U^*)^* \, (U\Lambda U^*) = U\Lambda^* \underline{U^*U}\Lambda U^* \\ &= U\Lambda^*\Lambda U^* = U|\Lambda|^2 U^* \\ AA^* &= (U\Lambda U^*) \, (U\Lambda U^*)^* = U\Lambda \underline{U^*U}\Lambda^* U^* \\ &= U\Lambda\Lambda^* U^* = U|\Lambda|^2 U^* \ \square \end{aligned}$$

Доказательство теоремы ($\Rightarrow$)

- по теореме Шура $A = UTU^*$

Доказательство теоремы ($\Rightarrow$)

- по теореме Шура $A = UTU^*$
- $A^*A = AA^* \Rightarrow UT^*TU^* = UTT^*U^* \Rightarrow T^*T = TT^*$

Доказательство теоремы ($\Rightarrow$)

- по теореме Шура $A = UTU^*$
- $A^*A = AA^* \Rightarrow UT^*TU^* = UTT^*U^* \Rightarrow T^*T = TT^*$
- $\underbrace{\begin{bmatrix} \lambda_1 & t_{12} & \dots & t_{1n} \\ 0 & \ddots & & \\ \vdots & & & \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}}_{T} \quad \underbrace{\begin{bmatrix} \bar{\lambda}_1 & 0 & \dots & 0 \\ \bar{t}_{12} & \ddots & & \\ \vdots & & & \\ \bar{t}_{1n} & 0 & \dots & \bar{\lambda}_n \end{bmatrix}}_{T^*}$

Доказательство теоремы ($\Rightarrow$)

- по теореме Шура $A = UTU^*$
- $A^*A = AA^* \Rightarrow UT^*TU^* = UTT^*U^* \Rightarrow T^*T = TT^*$
- $$\underbrace{\begin{bmatrix} \lambda_1 & t_{12} & \dots & t_{1n} \\ 0 & \ddots & & \\ \vdots & & & \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}}_{T} \quad \underbrace{\begin{bmatrix} \overline{\lambda}_1 & 0 & \dots & 0 \\ \bar{t}_{12} & \ddots & & \\ \vdots & & & \\ \bar{t}_{1n} & 0 & \dots & \overline{\lambda}_n \end{bmatrix}}_{T^*}$$
- $[T^*T]_{11} =$

Доказательство теоремы ($\Rightarrow$)

- по теореме Шура $A = UTU^*$
- $A^*A = AA^* \Rightarrow UT^*TU^* = UTT^*U^* \Rightarrow T^*T = TT^*$
- $$\underbrace{\begin{bmatrix} \lambda_1 & t_{12} & \dots & t_{1n} \\ 0 & \ddots & & \\ \vdots & & & \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}}_{T} \quad \underbrace{\begin{bmatrix} \bar{\lambda}_1 & 0 & \dots & 0 \\ \bar{t}_{12} & \ddots & & \\ \vdots & & & \\ \bar{t}_{1n} & 0 & \dots & \bar{\lambda}_n \end{bmatrix}}_{T^*}$$
- $[T^*T]_{11} = |\lambda_1|^2$

Доказательство теоремы ($\Rightarrow$)

- по теореме Шура $A = UTU^*$
- $A^*A = AA^* \Rightarrow UT^*TU^* = UTT^*U^* \Rightarrow T^*T = TT^*$
- $$\underbrace{\begin{bmatrix} \lambda_1 & t_{12} & \dots & t_{1n} \\ 0 & \ddots & & \\ \vdots & & & \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}}_{T} \quad \underbrace{\begin{bmatrix} \bar{\lambda}_1 & 0 & \dots & 0 \\ \bar{t}_{12} & \ddots & & \\ \vdots & & & \\ \bar{t}_{1n} & 0 & \dots & \bar{\lambda}_n \end{bmatrix}}_{T^*}$$
- $[T^*T]_{11} = |\lambda_1|^2 = [TT^*]_{11} =$

Доказательство теоремы ($\Rightarrow$)

- по теореме Шура $A = UTU^*$
- $A^*A = AA^* \Rightarrow UT^*TU^* = UTT^*U^* \Rightarrow T^*T = TT^*$
- $$\underbrace{\begin{bmatrix} \lambda_1 & t_{12} & \dots & t_{1n} \\ 0 & \ddots & & \\ \vdots & & & \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}}_{T} \quad \underbrace{\begin{bmatrix} \bar{\lambda}_1 & 0 & \dots & 0 \\ \bar{t}_{12} & \ddots & & \\ \vdots & & & \\ \bar{t}_{1n} & 0 & \dots & \bar{\lambda}_n \end{bmatrix}}_{T^*}$$
- $[T^*T]_{11} = |\lambda_1|^2 = [TT^*]_{11} = |\lambda_1|^2 + |t_{12}|^2 + \dots |t_{1n}|^2$

Доказательство теоремы ($\Rightarrow$)

- по теореме Шура $A = UTU^*$
- $A^*A = AA^* \Rightarrow UT^*TU^* = UTT^*U^* \Rightarrow T^*T = TT^*$
- $$\underbrace{\begin{bmatrix} \lambda_1 & t_{12} & \dots & t_{1n} \\ 0 & \ddots & & \\ \vdots & & & \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}}_{T} \quad \underbrace{\begin{bmatrix} \bar{\lambda}_1 & 0 & \dots & 0 \\ \bar{t}_{12} & \ddots & & \\ \vdots & & & \\ \bar{t}_{1n} & 0 & \dots & \bar{\lambda}_n \end{bmatrix}}_{T^*}$$
- $[T^*T]_{11} = |\lambda_1|^2 = [TT^*]_{11} = |\lambda_1|^2 + |t_{12}|^2 + \dots |t_{1n}|^2$
- Внедиагональные элементы равны 0: $T = \Lambda$

Доказательство теоремы ($\Rightarrow$)

- по теореме Шура $A = UTU^*$
- $A^*A = AA^* \Rightarrow UT^*TU^* = UTT^*U^* \Rightarrow T^*T = TT^*$
- $\underbrace{\begin{bmatrix} \lambda_1 & t_{12} & \dots & t_{1n} \\ 0 & \ddots & & \\ \vdots & & & \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}}_{T} \quad \underbrace{\begin{bmatrix} \bar{\lambda}_1 & 0 & \dots & 0 \\ \bar{t}_{12} & \ddots & & \\ \vdots & & & \\ \bar{t}_{1n} & 0 & \dots & \bar{\lambda}_n \end{bmatrix}}_{T^*}$
- $[T^*T]_{11} = |\lambda_1|^2 = [TT^*]_{11} = |\lambda_1|^2 + |t_{12}|^2 + \dots |t_{1n}|^2$
- Внедиагональные элементы равны 0: $T = \Lambda$
- $AU = U\Lambda$ – столбцы унитарной матрицы U являются собственными векторами A $\square$.

Свойства нормальных матриц

- Степени нормальных матриц:

$$A^2 = U\Lambda U^* U\Lambda U^* = U\Lambda^2 U^*$$

Свойства нормальных матриц

- Степени нормальных матриц:

$$A^2 = U\Lambda U^* U\Lambda U^* = U\Lambda^2 U^*$$

- Нормальная матрица является эрмитовой $\Leftrightarrow$ $\lambda_k \in \mathbb{R},\ k = \overline{1, n}$

Свойства нормальных матриц

- Степени нормальных матриц:

$$A^2 = U\Lambda U^* U\Lambda U^* = U\Lambda^2 U^*$$

- Нормальная матрица является эрмитовой $\Leftrightarrow$ $\lambda_k \in \mathbb{R},\ k = \overline{1,n}$
- Нормальная матрица является унитарной $\Leftrightarrow$ $|\lambda_k| = 1,\ k = \overline{1,n}$

Знакоопределенные матрицы

Знакоопределенные матрицы

- Для любой эрмитовой матрицы $H = H^*$, $\forall x \in \mathbb{C}^n$

$$(Hx, x) = (x, H^*x) = (x, Hx) = \overline{(Hx, x)} \Rightarrow (Hx, x) \in \mathbb{R}$$

Знакоопределенные матрицы

- Для любой эрмитовой матрицы $H = H^*$, $\forall x \in \mathbb{C}^n$

$$(Hx, x) = (x, H^*x) = (x, Hx) = \overline{(Hx, x)} \Rightarrow (Hx, x) \in \mathbb{R}$$

Определение 2.8 (положительно определенная м-ца)

Эрмитова матрица A называется *положительно определенной* (*полуопределенной*), если

$$\forall x \in \mathbb{C}^n, x \neq 0 : \ (Ax, x) = x^* Ax > 0 \ (\geq 0)$$

Знакоопределенные матрицы

- Для любой эрмитовой матрицы $H = H^*$, $\forall x \in \mathbb{C}^n$

$$(Hx, x) = (x, H^*x) = (x, Hx) = \overline{(Hx, x)} \Rightarrow (Hx, x) \in \mathbb{R}$$

Определение 2.8 (положительно определенная м-ца)

Эрмитова матрица A называется *положительно определенной* (*полуопределенной*), если

$$\forall x \in \mathbb{C}^n, x \neq 0 : \ (Ax, x) = x^* Ax > 0 \ (\geq 0)$$

- $A > 0 \ (\geq 0) \iff \lambda_k > 0(\geq 0), \ k = \overline{1, n}$

Знакоопределенные матрицы

- Для любой эрмитовой матрицы $H = H^*$, $\forall x \in \mathbb{C}^n$

$$(Hx, x) = (x, H^*x) = (x, Hx) = \overline{(Hx, x)} \Rightarrow (Hx, x) \in \mathbb{R}$$

Определение 2.8 (положительно определенная м-ца)

Эрмитова матрица A называется *положительно определенной* (*полуопределенной*), если

$$\forall x \in \mathbb{C}^n, x \neq 0 : \ (Ax, x) = x^* A x > 0 \ (\geq 0)$$

- $A > 0 \ (\geq 0) \iff \lambda_k > 0 (\geq 0), \ k = \overline{1, n}$
- Можно ввести понятие положительной определенности *для всех* $A \in C^{n \times n}$: $Re\,(Ax, x) > 0 (\geq 0)$

Сингулярное разложение (SVD)(1)

Сингулярное разложение (SVD)(1)

- $\forall A \in C^{m\times n}$ эрмитова матрица A^*A положительно полуопределенная:
  $(A^*Ax, x) = (Ax, Ax) = \|Ax\|_2^2 \geq 0$

Сингулярное разложение (SVD)(1)

- $\forall A \in C^{m \times n}$ эрмитова матрица A^*A положительно полуопределенная: $(A^*Ax, x) = (Ax, Ax) = \|Ax\|_2^2 \geq 0$
- По теореме о диагонализации нормальной матрицы: $\exists V = [v_1| \ldots |v_n] \in \mathbb{C}^{n \times n}$:

$$V^*(A^*A)V = diag(\sigma_1^2, \ldots, \sigma_n^2),\ \sigma_1 \geq \ldots \geq \sigma_n \geq 0$$

Сингулярное разложение (SVD)(1)

- $\forall A \in C^{m\times n}$ эрмитова матрица A^*A положительно полуопределенная:
  $(A^*Ax, x) = (Ax, Ax) = \|Ax\|_2^2 \geq 0$
- По теореме о диагонализации нормальной матрицы: $\exists V = [v_1|\ldots|v_n] \in \mathbb{C}^{n\times n}$:

$$V^*(A^*A)V = diag(\sigma_1^2, \ldots, \sigma_n^2),\ \sigma_1 \geq \ldots \geq \sigma_n \geq 0$$

- Обозначим r – число ненулевых σ_k, $r \leq \min(n, m)$ возьмём $V_r = [v_1|, \ldots, |v_r]$, $\Sigma_r = diag(\sigma_1, \ldots, \sigma_r)$

Сингулярное разложение (SVD)(1)

- $\forall A \in C^{m\times n}$ эрмитова матрица A^*A положительно полуопределенная:
  $(A^*Ax, x) = (Ax, Ax) = \|Ax\|_2^2 \geq 0$
- По теореме о диагонализации нормальной матрицы: $\exists V = [v_1|\ldots|v_n] \in \mathbb{C}^{n\times n}$:

$$V^*(A^*A)V = diag(\sigma_1^2, \ldots, \sigma_n^2),\ \sigma_1 \geq \ldots \geq \sigma_n \geq 0$$

- Обозначим r – число ненулевых σ_k, $r \leq \min(n, m)$ возьмём $V_r = [v_1|, \ldots, |v_r]$, $\Sigma_r = diag(\sigma_1, \ldots, \sigma_r)$

$$V_r^*A^*AV_r = \Sigma_r^2 \Rightarrow (\Sigma_r^{-1}V_r^*A^*)(AV_r\Sigma_r^{-1}) = I_{r\times r}$$

Сингулярное разложение (SVD)(1)

- $\forall A \in C^{m\times n}$ эрмитова матрица A^*A положительно полуопределенная:
$(A^*Ax, x) = (Ax, Ax) = \|Ax\|_2^2 \geq 0$
- По теореме о диагонализации нормальной матрицы: $\exists V = [v_1|\ldots|v_n] \in \mathbb{C}^{n\times n}$:

$$V^*(A^*A)V = diag(\sigma_1^2, \ldots, \sigma_n^2),\ \sigma_1 \geq \ldots \geq \sigma_n \geq 0$$

- Обозначим r – число ненулевых σ_k, $r \leq \min(n, m)$ возьмём $V_r = [v_1|, \ldots, |v_r]$, $\Sigma_r = diag(\sigma_1, \ldots, \sigma_r)$

$$V_r^*A^*AV_r = \Sigma_r^2 \Rightarrow (\Sigma_r^{-1}V_r^*A^*)(AV_r\Sigma_r^{-1}) = I_{r\times r}$$

- $U_r = AV_r\Sigma_r^{-1} \in \mathbb{C}^{m\times r}$:
$U_r^*U_r = I_{r\times r} \Rightarrow$ столбцы U_r - ортонормированные.

Сингулярное разложение (SVD)(2)

- $U_r = AV_r\Sigma_r^{-1} \in \mathbb{C}^{m\times r}$:
  $U_r^* U_r = I_{r\times r} \Rightarrow$ столбцы U_r - ортонормированные.

Сингулярное разложение (SVD)(2)

- $U_r = AV_r\Sigma_r^{-1} \in \mathbb{C}^{m\times r}$:
  $U_r^* U_r = I_{r\times r} \Rightarrow$ столбцы U_r - ортонормированные.
- При $1 \le i \le r$: $Av_i = \sigma_i u_i$

Сингулярное разложение (SVD)(2)

- $U_r = AV_r\Sigma_r^{-1} \in \mathbb{C}^{m\times r}$:
  $U_r^* U_r = I_{r\times r} \Rightarrow$ столбцы U_r - ортонормированные.
- При $1 \le i \le r$: $Av_i = \sigma_i u_i$
- При $i > r$:
  $A^* A v_i = 0 \Rightarrow (v_i^* A^*)(Av_i) = \|Av_i\|_2^2 = 0 \Rightarrow Av_i = 0$.

Сингулярное разложение (SVD)(2)

- $U_r = AV_r\Sigma_r^{-1} \in \mathbb{C}^{m\times r}$:
  $U_r^* U_r = I_{r\times r} \Rightarrow$ столбцы U_r - ортонормированные.
- При $1 \le i \le r$: $Av_i = \sigma_i u_i$
- При $i > r$:
  $A^*Av_i = 0 \Rightarrow (v_i^* A^*)(Av_i) = \|Av_i\|_2^2 = 0 \Rightarrow Av_i = 0.$
- Достроим U_r до унитарной матрицы $U \in \mathbb{C}^{m\times m}$:

$$AV = U \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \Rightarrow A = U\Sigma_{m\times n}V^*$$

Сингулярное разложение (SVD)(2)

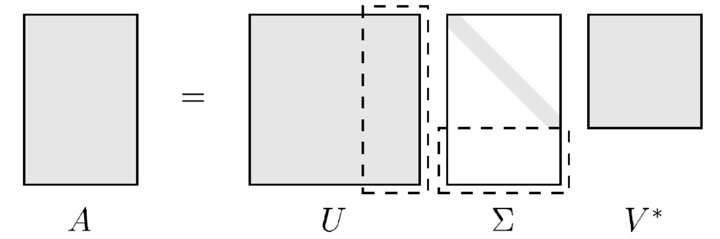
- $U_r = AV_r\Sigma_r^{-1} \in \mathbb{C}^{m\times r}$:
  $U_r^* U_r = I_{r\times r} \Rightarrow$ столбцы U_r - ортонормированные.
- При $1 \le i \le r$: $Av_i = \sigma_i u_i$
- При $i > r$:
  $A^*Av_i = 0 \Rightarrow (v_i^* A^*)(Av_i) = \|Av_i\|_2^2 = 0 \Rightarrow Av_i = 0.$
- Достроим U_r до унитарной матрицы $U \in \mathbb{C}^{m\times m}$:

$$AV = U \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \Rightarrow A = U\Sigma_{m\times n}V^*$$

- Т.к. умножение на невырожденные матрицы не меняет ранг, то $rank(A) = rank(\Sigma) = r$

Теорема 2.4 (сингулярное разложение)

$\forall A \in \mathbb{C}^{m\times n}$ ранга r существует разложение:



где $U \in \mathbb{C}^{m\times m}$, $V \in \mathbb{C}^{n\times n}$ - унитарные; $\Sigma \in \mathbb{C}^{m\times n}$:

$$\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & 0 \\ & & & \end{bmatrix}, \ \sigma_k > 0, k = 1, \dots, r$$

Сингулярное разложение, следствия

- Можно задать матрицу $(m+n+1)r$ параметрами

Сингулярное разложение, следствия

- Можно задать матрицу $(m+n+1)r$ параметрами
- Разделение «переменных» (индексов)

$$A = U\Sigma V^* = \sum_{k=1}^{r} \sigma_k u_k v_k^* \Rightarrow A(i,j) = \sum_{k=1}^{r} \sigma_k u_k(i) v_k^*(j)$$

Сингулярное разложение, следствия

- Можно задать матрицу $(m+n+1)r$ параметрами
- Разделение «переменных» (индексов)

$$A = U\Sigma V^* = \sum_{k=1}^{r} \sigma_k u_k v_k^* \Rightarrow A(i,j) = \sum_{k=1}^{r} \sigma_k u_k(i) v_k^*(j)$$

- $Ax = \sum_{k=1}^{r} \sigma_k u_k (v_k^* x)$ – $O((n+m)r)$ операций

Сингулярное разложение, следствия

- Можно задать матрицу $(m+n+1)r$ параметрами
- Разделение «переменных» (индексов)

$$A = U\Sigma V^* = \sum_{k=1}^{r} \sigma_k u_k v_k^* \Rightarrow A(i,j) = \sum_{k=1}^{r} \sigma_k u_k(i) v_k^*(j)$$

- $Ax = \sum_{k=1}^{r} \sigma_k u_k (v_k^* x)$ – $O((n+m)r)$ операций
- 4 основные подпространства, связанные с матрицей A:

$$\begin{aligned}
\ker(A) &= \mathcal{L}\{v_{r+1}, \dots, v_n\} \\
\operatorname{im}(A) &= \mathcal{L}\{u_1, \dots, u_r\} \\
\ker(A^*) &= \mathcal{L}\{u_{r+1}, \dots, u_n\} \\
\operatorname{im}(A^*) &= \mathcal{L}\{v_1, \dots, v_r\}
\end{aligned}$$

SVD и приближение малого ранга

Теорема 2.5 (Эккарта-Янга)

Пусть $k < \mathrm{rank}(A) = R$, $A_k = U_k \Sigma_k V_k^*$. Тогда

$$\min_{\mathrm{rank}(B)=k} \|A - B\|_2 = \|A - A_k\|_2 = \sigma_{k+1}$$

$$\min_{\mathrm{rank}(B)=k} \|A - B\|_F = \|A - A_k\|_F = \sqrt{\sigma_{k+1}^2 + \cdots + \sigma_R^2}$$

- A_k можно задать $(m+n+1)k$ параметрами
- Применение $(k \ll r)$:
  - ✓ сжатие данных
  - ✓ обработка сигналов
  - ✓ итерационные методы обращения матриц
  - ✓ метод главных компонент
  - ✓ латентно-семантический анализ

Примеры применения SVD

Заключение

Заключение

- Векторные и матричные нормы

Заключение

- Векторные и матричные нормы
- Унитарные, эрмитовы, нормальные, знакоопределенные матрицы

Заключение

- Векторные и матричные нормы
- Унитарные, эрмитовы, нормальные, знакоопределенные матрицы
- Разложение Шура, сингулярное разложение