Лекции по Численным методам

Лектор: Тараканов Александр

Веса 2024

Содержание

1 Прямые методы решения систем линейных уравнений

Постановка задачи

Рассмотрим систему уравнений:

$$\begin{cases} A_{11}x_1 + \ldots A_{1n}x_n = y_1 \\ \ldots \\ A_{n1}x_1 + \ldots + A_{nn}x_n = y_n \end{cases} \tag{1}$$

Система уравнений (1) называется линейной системой уравнений (СЛУ). A_{ij} — матрица коэффициентов, y_i — правая часть и x_i — неизвестные

Про СЛУ естественно говорить на языке линейных пространств и операторов: пусть задан линеный оператор $A : X \longrightarrow Y$, вектор $y \in Y$ и требуется найти его прообраз $x \in X : Ax = y$

Предположения

В общем случае при решении СЛУ возможны несколько случаев:

- Решение существует и единственно $\ker A = 0$ и $y \in \operatorname{Im} A$
- Решение существует, но не единственно $\ker A \neq 0$ и $y \in \operatorname{Im} A$
- Решения не существует $y \notin \operatorname{Im} A$

В данном курсе мы будем рассматривать только частный случай: $\ker A = 0$ и $\operatorname{Im} A = Y$ или $\operatorname{coker} A = 0$

Иными словами: будем считать, что X и Y — линейные пространства одинаковой размерности n, матрица СЛУ A является несингулярной (квадратной и невырожденной)

В такой постановке задача $Ax = y$ определена корректно и решение существует и единственно для любой правой части

Далее в курсе будут разбираться различные алгоритмы поиска решения в зависимости от свойств матрицы A

1.1 Нижние и верхние треугольные матрицы

Определение 1. Матрица L называется *нижней треугольной* матрицей, если $L_{ij} = 0$, если $j > i$

Определение 2. Матрица U называется *верхней треугольной* матрицей, если $U_{ij} = 0$, если $j < i$

Нижние и верхние треугольные матрицы обладают следующим важным свойством

Утверждение 3. *Пусть A и B нижние (верхние) треугольные матрицы, то и матрица $C = AB$ является нижней (верхней) треугольной матрицей*

Доказательство. Для доказательства вычислим значение матричного элемента C_{ij} при $j > i$ для нижних треугольных матриц:

$$C_{ij} = \sum_{k=1}^{n} A_{ik} \cdot B_{kj} = \sum_{k=1}^{i} A_{ik} \cdot B_{kj} + \sum_{k=i+1}^{n} A_{ik} \cdot B_{kj} = \sum_{k=1}^{i} A_{ik} \cdot 0 + \sum_{k=i+1}^{n} 0 \cdot B_{kj} = 0$$

$\square$

Решение системы с нижней треугольной матрицей

Рассмотрим СЛУ с нижней треугольной матрицей L:

$$\begin{cases} L_{11}x_1 & = y_1 \\ L_{21}x_1 + L_{22}x_2 & = y_2 \\ \ldots \\ L_{n1}x_1 + \ldots + L_{nn}x_n & = y_n \end{cases}$$

Тогда решение можно найти, исключая неизвестные:

$$
\begin{aligned}
&x_1 = y_1/L_{11} \\
&x_2 = (y_2 - L_{21}x_1)/L_{22} \\
&\dots \\
&x_i = \left(y_i - \sum_{j=1}^{i-1} L_{ij}x_j \right) /L_{ii} \\
&\dots
\end{aligned}
$$

Решение системы с верхней треугольной матрицей

Рассмотрим СЛУ с верхней треугольной матрицей U:

$$
\begin{cases}
U_{11}x_1 + \ldots + U_{1n}x_n & = y_1 \\
\dots \\
U_{(n-1)(n-1)}x_{n-1} + \ldots + U_{(n-1)n}x_n & = y_{n-1} \\
U_{nn}x_1 & = y_n
\end{cases}
$$

Тогда решение можно найти, исключая неизвестные:

$$
\begin{aligned}
&x_n = y_n/U_{nn} \\
&x_{n-1} = (y_{n-1} - U_{(n-1)n}x_{n-1})/U_{(n-1)(n-1)} \\
&\dots \\
&x_i = \left(y_i - \sum_{j=i+1}^{n} U_{ij}x_j \right) /U_{ii} \\
&\dots
\end{aligned}
$$

1.2 Решение системы уравнений с помощью LU разложения

Пусть есть СЛУ $Ax = y$

Идея алгоритма состоит в следующем:

- Представить матрицу A в виде произведения: $A = LU$, где L — нижняя треугольная матрица, U — верхняя треугольная
- Решить систему $Lz = y$
- Решиить систему $Ux = z$

Доказательство. Пусть x — решение исходной системы. Тогда

$$Ax = LUx = L(Ux) = Lz = y$$

□

Замечание

- В общем случае LU разложение не определено однозначно: если D — невырожденная диагональная матрица, то можно построить другое LU разложение по уже имеющемуся:

  $$A = LU = LDD^{-1}U = (LD)(D^{-1}U) = L'U'$$

  Данную неопределенность можно решить, зафиксировав, что $L_{ii} = 1$ или $U_{ii} = 1$

Алгоритм построения LU разложения

Рассмотрим случай, когда диагональ верхней треугольной матрицы $U_{ii} = 1$. Будем вычислять элементы матриц L и U построчно:

- Первая строка: $L_{11}U_{11} = A_{11}$ и $U_{11} = 1$, поэтому $L_{11} = A_{11}/U_{11}$ (перемножили 1-ую строку и 1-ый столбец). Далее вычислим U_{1i}: (перемножаем 1-ую строку и i-ый столбец)

$$L_{11}U_{1i} = A_{1i} \iff U_{1i} = A_{1i}/L_{11}$$

- Вторая строка: $L_{21}U_{11} = A_{21}$ и $U_{11} = 1$, поэтому $L_{21} = A_{21}/U_{11}$ (перемножили 1-ую строку и 2-ый столбец). Далее вычислим L_{22}: (перемножим 2-ую строку и 2-ый столбец)

$$L_{21}U_{12} + L_{22}U_{22} = A_{22} \iff L_{22} = (A_{22} - L_{21}U_{12})/U_{22}$$

  Теперь вычислим U_{2i}: (перемножим 2-ую строку и i-ый столбец)

$$L_{21}U_{12} + L_{22}U_{2i} = A_{2i} \iff U_{2i} = (A_{2i} - L_{21}U_{12})/L_{22}$$

  И так далее

- В итоге получаем, что вычислить i-ую строку матрицы L, можно следующим образом: ($j \leqslant i$)

$$L_{ij} = \left(A_{ij} - \sum_{k=1}^{j-1} L_{ik}U_{kj}\right)/U_{jj}$$

  А i-ая строка матрицы U вычисляется так: ($j > i$)

$$U_{ij} = \left(A_{ij} - \sum_{k=1}^{i-1} L_{ik}U_{kj}\right)/L_{ii}$$

Замечание Если мы задаем диагональные элементы нижней треугольной матрицы L, то задача сводится к уже решенной:

$$A^\top = L'U' \iff A = (U')^\top(L')^\top = LU$$

Данный алгоритм позволяет найти LU разложение для матрицы A единственным образом, если зафиксировать диагональные элементы матрицы U или L. Однако если на m-ом шаге окажется, что $L_{mm} = 0$, то алгоритм позволяет найти только лишь LU разложение главного минора порядка m матрицы A:

$$[A]_m = [L]_m[U]_m,$$

где $[L]_m$ и $[U]_m$ — нижняя и верхняя треугольные матрицы, вычисленные за m шагов

В общем случае, данный алгоритм не гарантирует сходимости к LU разложению для матрицы A, однако существует класс матриц A, для которых алгоритм корректно находит LU разложение

Класс матриц

Необходимо проверить, есть ли нули на диагонали матрицы L, тогда вышеописанный алгоритм будет работать корректно

Утверждение 4. *Если* $\forall m \in \{1, \dots n\} : \det[A]_m \neq 0$*, то* $L_{mm} \neq 0$

Доказательство. Докажем от противного. Зафиксируем $m \in \{1, \dots, n\}$ и $\det[A]_m \neq 0$. Пусть $L_{ii} \neq 0$ при $i < m$ и $L_{mm} = 0$. Тогда верно, что

$$[A]_m = [L]_m[U]_m \implies \det[A]_m = \det[L]_m \cdot \det[U]_m = \prod_{i=1}^{m} L_{ii} \cdot \prod_{i=1}^{m} U_{ii} = 0$$

Получили противоречие, что определитель главного минора $[A]_m$ не равен 0 □

В общем случае данное утверждение сложно проверить. Если матрица является симметричной положительно определенной (SPD), то тогда оно выполнено по критерию Сильвестра

Помимо SPD матриц, часто встречаются матрицы со следующим особым свойством

Определение 5. Матрица A называется *матрицей с диагональным преобладанием*, если $\forall\, i \in \{1, \ldots, n\}$

$$|A_{ii}| - \sum_{j \neq i} |A_{ij}| > 0$$

Видно, что главный минор такой матрицы также является матрицей с диагональным преобладанием. Докажем следующее утверждение

Утверждение 6. *Матрица с диагональным преобладанием является несингулярной*

Доказательство. Докажем от противного. Пусть матрица вырожденная. Тогда $\exists\, x \in \mathbb{R}^n : Ax = 0$. Найдем максимальный по модулю элемент в векторе $x : |x_i| = \max\limits_j |x_j|$ и рассмотрим i-ую строку Ax:

$$0 = \left|\sum_j A_{ij} x_j\right| = |x_i| \cdot \left|\sum_j A_{ij} \frac{x_j}{|x_i|}\right| = |x_i| \cdot \left|A_{ii} + \sum_{j \neq i} A_{ij} \cdot \frac{x_j}{|x_i|}\right| \geqslant |x_i| \cdot \left|A_{ii} - \sum_{j \neq i} |A_{ij}| \cdot \frac{|x_j|}{|x_i|}\right| > 0$$

Предпоследнее неравенство верно по обратному неравенства треугольника. Последнее неравенство верно, так как матрица A является матрицей с диагональным преобладанием, а отношение $|x_j|/|x_i| < 1$ при $j \neq i$.

Получили противоречие □

1.3 Решение системы уравнений с помощью разложения Холецкого

Пусть есть СЛУ $Ax = y$ и матрица A является SPD матрицей. Тогда существует специальный вид (причем единственный) LU разложения — разложение Холецкого:

$$A = LL^\top,$$

где L — нижняя треугольная матрица с положительными элементами на диагонали

Утверждение 7. *Существование разложения Холецкого*

Доказательство. Пусть задано какое-то LU разложение: $A = LU$ (существование его доказывалось ранее). Тогда верно следующее:

$$LU = A = A^\top = U^\top L^\top$$

Домножим слева равенство на L^{-1}:

$$U = L^{-1} U^\top L^\top$$

Теперь домножим справа равенство на $(L^\top)^{-1}$:

$$U (L^\top)^{-1} = L^{-1} U^\top = D$$

Получим, что слева у нас верхняя треугольная матрица, а справа нижняя треугольная матрица, поэтому и справа, и слева диагональная матрица D

Рассмотрим исходное LU разложение:

$$A = LU = L \cdot (DL^\top) = LD^{1/2} D^{1/2} L^\top = (LD^{1/2})(LD^{1/2})^\top = L' L'^\top$$

Причем диагональные элементы D могут быть только положительными, так как A является SPD матрицей и L — матрица перехода от одного базиса к другому

□

Утверждение 8. *Единственность разложения Холецкого*

Доказательство. Единственность доказывается вместе с построением алгоритма вычисления, аналогичному для LU разложения

- $L_{11}L_{11} = A_{11} \iff L_{11} = \sqrt{A_{11}}$
- i-ая строка при $j < i$:

$$\sum_{k=1}^{j-1} L_{ik}L_{kj} + L_{ij}L_{jj} = A_{ij} \iff L_{ij} = \left(A_{ij} - \sum_{k=1}^{j-1} L_{ik}L_{kj}\right)/L_{jj}$$

- Диагональные элементы L_{ii}:

$$L_{ii} = \sqrt{A_{ii} - \sum_{k=1}^{j-1} L_{ik}L_{kj}}$$

$\square$

1.4 Решение систем уравнений с помощью алгоритма Гаусса

Пусть есть СЛУ $Ax = y$

Будем решать ее алгоритмом Гаусса: сначала применять прямой алгоритм Гаусса, потом обратный. Данный алгоритм является одним из способов построения LU разложения

Во время прямого алгоритма Гаусса мы будем приводить матрицу A к ступенчатому виду, выполняя операции над строками:

- 1 тип: $i \mapsto i \cdot \lambda, \lambda \neq 0$
- 2 тип: $i \leftrightarrow j$
- 3 тип: $i \mapsto i + j \cdot \lambda$

Во время обратного хода мы теми же действиями будем приводить матрицу A к улучшенному ступенчатому виду, чтобы главные коэффициенты были равны 1

Каждой операции над строками однозначно сопоставляется умножение слева на матрицу

Сложность алгоритма Гаусса составляет $O(n^3)$, где n — разномерность матрицы A. Поэтому данные алгоритмы не используются для решения СЛУ больших размерностей

2 Норма и анализ сходимости

Определение 9. Пусть задано линейное векторное пространство V над полем $\mathbb{R}$.
Функцию $\|\cdot\| : V \longrightarrow \mathbb{R}$ будем называть *нормой*, если выполнены следующие свойства:

- $\forall\, x \in V : \|x\| \geqslant 0$
- $\|x\| = 0 \iff x = 0$
- $\|\alpha x\| = |\alpha| \cdot \|x\|$
- $\forall\, x, y \in V : \|x + y\| \leqslant \|x\| + \|y\|$ — неравенство треугольника

Пространство V с нормой $\|\cdot\|$ называется *нормированным пространством*

2.1 Векторные l_p нормы

Важным примером норм является l_p норма. Пусть $V = \mathbb{R}^n$. Тогда для $x \in V$, который будем записывать в виде вектор-столбца $x = [x_1, \dots, x_n]^\top$, определим l_p норму:

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

Существуют особые виды l_p нормы:

- $p = 1 : \|x\|_1 = |x_1| + \ldots + |x_n|$
- $p = 2 : \|x\|_2 = \sqrt{|x_1|^2 + \ldots + |x_n|^2}$
- $p = \infty : \|x\|_\infty = \max\limits_i |x_i|$

Утверждение 10. *Докажем, что приведенные функции являются нормами*

Доказательство. Разберем каждый случай по отдельности:

1. Случай $p = 1$

   - $\|x\|_1 \geqslant 0$ очевидно. Пусть $\|x\| = 0$. Тогда $|x_1| + \ldots + |x_n| = 0 \iff x_1 = \ldots = x_n = 0 \iff x = 0$. В обратную сторону очевидно
   - $\|\alpha x\|_1 = |\alpha x_1| + \ldots + |\alpha x_n| = |\alpha| \cdot (|x_1| + \ldots + |x_n|) = |\alpha| \cdot \|x\|_1$
   - Зафиксируем $x, y \in V$. Тогда

   $$\|x + y\|_1 = \sum_{i=1}^{n} |x_i + y_i| \leqslant \sum_{i=1}^{n} |x_i| + \sum_{i=1}^{n} |y_i| = \|x\|_1 + \|y\|_1$$

2. Случай $p = 2$

   - Первые три свойства проверяются аналогично. Докажем только неравенство треугольника. Зафиксируем $x, y \in V$ и воспользуемся неравенством Коши-Буняковского:

   $$\|x+y\|_2^2 = \sum_{i=1}^{n} |x_i+y_i|^2 = \sum_{i=1}^{n} |x_i|^2 + \sum_{i=1}^{n} |y_i|^2 + 2\sum_{i=1}^{n} |x_i| \cdot |y_i| \leqslant \|x\|_2^2 + \|y\|_2^2 + 2\|x\|_2 \cdot \|y\|_2 = (\|x\|_2 + \|y\|_2)^2$$

3. Случай $p \in \mathbb{R} : p > 1$

   - Первые три свойства проверяются аналогично. Докажем неравенство треугольника

- Предположим, что $x, y \in V : x \neq 0, y \neq 0$ и $\forall\, i \in \{1, \dots, n\} : x_i \geqslant 0$ без ограничения общности рассуждений.
  Зафиксируем x и будем искать максимум $f(y) = \|x + y\|_p$ по y при условии, что $\|y\|_p = C = \text{const}$
  Из курса математического анализа известно, что непрерывная функция на компакте достигает своего максимума. Пусть $f(y)$ достигает максимума в точке $y^\star$
  Тогда запишем уравнение касательной плоскости к поверхности $\|y\|_p = C$:
  ($dy = [dy_1, \dots, dy_n]^\top$ — вектор приращений)

$$\sum_{i=1}^{n} \frac{\partial}{\partial y_i} \|y\|_p^p \, dy_i = \sum_{i=1}^{n} p|y_i|^{p-1} \, dy_i = 0 = \left\langle \nabla \|y\|_p^p, \, dy \right\rangle \tag{2}$$

- Так как $y^\star$ — точка экстремума функции, то найдем производную $f(y)^p$:

$$\frac{d}{dy} f(y^\star)^p = \sum_{i=1}^{n} p|x_i + y^\star|^{p-1} \, dy_i = 0 = \left\langle \nabla f(y^\star)^p, \, dy \right\rangle \tag{3}$$

- Из (2) в точке $y^\star$ и (3) следует, что векторы

$$\nabla \|y^\star\|_p^p = \left[|y_1^\star|^{p-1}, \dots, |y_n^\star|^{p-1} \right]^\top$$

  и

$$\nabla f(y^\star)^p = \left[|x_1 + y_1^\star|^{p-1}, \dots, |x_n + y_n^\star|^{p-1} \right]^\top$$

  перпендикулярны вектору приращений dy, а значит коллинеарны:

$$|x_i + y_i^\star| = \lambda |y_i^\star|$$

- Так как $y^\star$ — точка максимума, то в знаки x_i и $y_i^\star$ должны совпадать. То есть $y_i = kx_i$ для некоторого $k > 0$, которое можно найти следующим образом:

$$k = \frac{\|y\|_p}{\|x\|_p} = \frac{C}{\|x\|_p}$$

$$\|x + y\|_p \leqslant \|x + y^\star\|_p = \|x + kx\|_p = \|x\|_p + \|kx\|_p = \|x\|_p + \|y\|_p$$

4. Случай $p = \infty$

- Рассмотрим следующий предел:

$$\lim_{p \to \infty} \|x\|_p$$

- Пусть дан вектор $x \in V : \|x\|_\infty = |x_k|$. Тогда

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

- Отсюда и из $n^{1/p} \longrightarrow 1$ при $p \longrightarrow \infty$ видно, что $\|x\|_\infty = \lim\limits_{p \to \infty} \|x\|_p$
- Теперь мы можем доказать, что $\|x\|_\infty$ является нормой по предельным переходам

$\square$

Определение 11. Пусть задано линейное векторное пространство V над полем $\mathbb{R}$.
Функцию $\rho : V \times V \longrightarrow \mathbb{R}$ будем называть *метрикой*, если выполнены следующие свойства:

- $\forall\, x, y \in V : \rho(x, y) \geqslant 0$
- $\rho(x, y) = 0 \iff x = y$

- $\forall\, x, y \in V : \rho(x, y) = \rho(y, x)$
- $\forall\, x, y, z \in V : \rho(x + z) \leqslant \rho(x, y) + \rho(y, z)$ — неравенство треугольника

Пространство V с метрикой ρ называется *метрическим пространством*

Заметим, что любая норма на линейном пространстве задает метрику:

$$\rho(x, y) = \|x - y\|$$

2.2 Сходимость по норме

Определение 12. Две нормы $\|\cdot\|_a$ и $\|\cdot\|_b$ (необязательно l_p нормы) на нормированном пространстве V называются эквивалентными, если $\exists\, C_1, C_2 > 0 : \forall\, x \in V$

$$C_1 \cdot \|x\|_a \leqslant \|x\|_b \leqslant C_2 \cdot \|x\|_b$$

Известно, что на конечномерных пространствах все нормы являются эквивалентными. Будем говорить, что последовательность векторов $\{x_k\}$ сходится к x по норме, если $\|x_k - x\| \longrightarrow 0$ при $k \longrightarrow \infty$. Так как все нормы являются эквивалентными, то для исследования сходимости можно использовать любую норму. Также для конечномерных пространств верно, что из покоордиантной сходимости следует сходимость по норме и наоборот

2.3 Норма линейного оператора

Определение 13. Пусть задано нормированное пространство V с нормой $\|x\|$. Пусть задан линейный оператор $A : V \longrightarrow V$. Определим норму линейного оператора следующим образом:

$$\|A\| = \sup_{\|x\| \neq 0} \frac{\|Ax\|}{\|x\|}$$

Утверждение 14. *Докажем, что это действительно норма, и перечислим свойства*

- Первые три свойства нормы выполняются
- Проверим неравенство треугольника:

$$\|A + B\| = \sup_{\|x\| \neq 0} \frac{\|Ax + Bx\|}{\|x\|} \leqslant \sup_{\|x\| \neq 0} \left(\frac{\|Ax\|}{\|x\|} + \frac{\|Bx\|}{\|x\|} \right) \leqslant \sup_{\|x\| \neq 0} \frac{\|Ax\|}{\|x\|} + \sup_{\|x\| \neq 0} \frac{\|Bx\|}{\|x\|} = \|A\| + \|B\|$$

- Видно из определения нормы, что

$$\|Ax\| \leqslant \|A\| \cdot \|x\|$$

- Оценим сверху норму композиции операторов BA:

$$\|BAx\| = \|B(Ax)\| \leqslant \|B\| \cdot \|Ax\| \leqslant \|B\| \cdot \|A\| \cdot \|x\|$$

  Поэтому $\|BA\| \leqslant \|B\| \cdot \|A\|$

- Из линейности оператора следует, что

$$\sup_{\|x\| \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\| \neq 0} \left\| A \cdot \left(\frac{x}{\|x\|} \right) \right\| = \sup_{\|x\| = 1} \|Ax\|$$

  По этой причине норма любого линейного оператора на конечномерном линейном пространстве с l_p нормой существует и конечна

Примеры

- Норма диагонального оператора $D_{n\times n} = D$ с помощью l_p нормы:

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

  Пусть $D_{kk} = \max_i |D_{ii}|$. Тогда

$$\left(\sum_{i=1}^{n} |D_{ii}|^p \cdot |x_i|^p\right)^{1/p} \leqslant \left(\sum_{i=1}^{n} |D_{kk}|^p \cdot |x_i|^p\right)^{1/p} = |D_{kk}| \left(\sum_{i=1}^{n} |x_i|\right)^{1/p}$$

  Получается, что

$$\sup_{\|x\|=1} \|Dx\|_p \leqslant \max_i |D_{ii}|$$

  Пример, на котором достигается максимум легко построить: возьмем $x = [0, \dots, 1, \dots, 0]^\top$ — ненулевая координата только на k-ой позиции

- Рассмотрим l_2 норму на конечномерном линейном пространстве V. Пусть есть некоторый оператор A. Известно, что любой оператор можно разложить в виде композиции поворотов, отражений и растяжений вдоль осей с положительными коэффициентами — SVD:

$$A = U_1 D U_2,$$

  где U_1, U_2 — ортогональные матрицы, которые сохраняют расстояние $\|U_1 x\| = \|U_2 x\| = \|x\|$

$$\|A\| = \sup_{\|x\|=1} \|Ax\| = \sup_{\|x\|=1} \|U_1 D U_2 x\| = \sup_{\|x\|=1} \|D U_2 x\| = \sup_{\|x\|=1} \|Dx\| = \max_i D_{ii}$$

- Рассмотрим l_2 норму на конечномерном линейном пространстве V. Пусть есть некоторый самосопряженный оператор A (заданный SPD матрицей). Тогда представим его в следующем виде:

$$A = UDU^\top,$$

  где U — ортогональная матрица, D — диагональная матрица из собственный значений λ_i. Аналогично предыдущему пункту:

$$\|A\| = \max_i \lambda_i$$

- Рассмотрим l_∞ норму на конечномерном линейном пространстве V и произвольный оператор A

$$\|A\| = \sup_{\|x\|=1} \|Ax\|_\infty = \max_i \left|\sum_j A_{ij} x_j\right| = \left|\sum_j A_{kj} x_j\right|$$

  Заметим, что A_{kj} и x_j должны быть одного знака, иначе можно поменять знак координаты j вектора x на противоположный и значение увеличится, что противоречит максимальности выражения. Так как $\|x\|_\infty = 1$, тогда есть координата $|x_m| = 1$. Тогда заменим все x_j на 1, от этого норма не изменится. В итоге получили, что

$$\|A\| = \max_i \left|\sum_j A_{ij}\right|$$

2.4 Число обусловленности

Определение 15. Пусть на нормированном конечномерном пространстве V задан невырожденный линейный оператор $A : V \longrightarrow V$. *Числом обусловленности* линейного оператора будем называть следующее выражение:

$$\kappa(A) = \|A\| \cdot \|A^{-1}\|$$

Видно, что $\kappa(A) = \kappa(A^{-1})$

Утверждение 16. *Число обусловленности $\kappa(A) \geqslant 1$*

Доказательство.

$$\|x\| = \|A^{-1}Ax\| \leqslant \|A^{-1}\| \cdot \|Ax\| \leqslant \|A^{-1}\| \cdot \|A\| \cdot \|x\| = \kappa(A) \cdot \|x\|$$

□

Утверждение 17. *Пусть A — самосопряженный линейный оператор и задана l_2 норма. Тогда число обусловленности равно*

$$\kappa(A) = \frac{\max_i \lambda_i}{\min_i \lambda_i},$$

где λ_i — собственное значение матрицы A

Доказательство. Докажем, что

$$\kappa(A) = \frac{\sup_{\|x\|=1}\|Ax\|}{\inf_{\|x\|=1}\|Ax\|}$$

- Так как A — самосопряженный оператор, то $A = U^\top DU$ и $A^{-1} = U^\top D^{-1}U$
- Найдем $\|A^{-1}\|$

$$\|A^{-1}\| = \sup_{\|x\|\neq 0} \frac{\|A^{-1}x\|}{\|x\|} = \sup_{\|y\|\neq 0} \frac{\|A^{-1}Ay\|}{\|Ay\|} = \sup_{\|y\|\neq 0} \frac{\|y\|}{\|Ay\|} = \frac{1}{\inf_{\|x\|=1}\|Ax\|}$$

- Отсюда и так как задана l_2 норма получаем нужное равенство через собственные значения

□

В общем случае, число обусловленности показывает, насколько матрица близка к сингулярной: чем больше число обусловленности, тем ближе к сингулярности

2.5 Число обусловленности и устойчивость решения системы уравнений

Рассмотрим СЛУ $Ax = b$. Допустим, что правая часть b известна с точностью до ошибок Δb. Тогда мы решаем систему $Ax^\star = b + \Delta b$

Пусть погрешность тогда равна $\Delta x = x^\star - x$. Оценим относительную ошибку:

$$\frac{\|\Delta x\|}{\|x\|} = \frac{\|A^{-1}\Delta b\|}{\|x\|} \leqslant \|A^{-1}\| \cdot \frac{\|\Delta b\|}{\|b\|} \cdot \frac{\|b\|}{\|x\|} = \|A^{-1}\| \cdot \frac{\|\Delta b\|}{\|b\|} \cdot \frac{\|Ax\|}{\|x\|} \leqslant \|A^{-1}\| \cdot \|A\| \cdot \frac{\|\Delta b\|}{\|b\|} = \kappa(A) \cdot \frac{\|\Delta b\|}{\|b\|}$$

Рассмотрев $A^{-1}b = x$ аналогично можно получить следующую оценку:

$$\frac{1}{\kappa(A)} \cdot \frac{\|\Delta b\|}{\|b\|} \leqslant \frac{\|\Delta x\|}{\|x\|} \leqslant \kappa(A) \cdot \frac{\|\Delta b\|}{\|b\|}$$

СЛУ, где матрица A имеет большое число обусловленности, могут иметь неустойчивые решения, которые могут сильно отличаться от аналитического решения