Лекция
25 марта 2020г

РЕШЕНИЕ ПЛОХО ОБУСЛОВЛЕННЫХ СИСТЕМ ЛИНЕЙНЫХ УРАВНЕНИЙ

Лектор: Лазарева Галина Геннадьевна

Плохо обусловленные матрицы

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ a_{n1} & \dots & a_{nn} \end{pmatrix},$$

$$\begin{array}{llll} a_{11}x_1 + a_{12}x_2 + & \dots & +a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + & \dots & +a_{2n}x_n = b_2, \\ . \quad . \quad . \quad . \quad . \quad . & \dots & . \quad . \quad . \quad . \\ . \quad . \quad . \quad . \quad . \quad . & \dots & . \quad . \quad . \quad . \\ a_{n1}x_1 + a_{n2}x_2 + & \dots & +a_{nn}x_n = b_n \end{array}$$

Ax = b

Плохо обусловленные матрицы

При $\det A \neq 0,$
система уравнений **Ax=b**
имеет единственное решение.
Численные методы позволяют получать приближенные решения системы .
Рассмотрим вопрос устойчивости решения **x** по отношению к погрешностям вектора **b**.

Плохо обусловленные матрицы

Пусть в конечномерном линейном пространстве $\mathbb{E}^n$ введена норма элемента $\mathbf{x}$ пространства $\mathbb{E}^n$. Определим норму $\|A\|$ матрицы A:

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

Величина

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

называется **числом обусловленности** матрицы A.

При $\nu(A) \gg 1$

матрица A и система уравнений Ax=b называются **плохо обусловленными.**

Плохо обусловленные системы

Величина

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

называется **числом обусловленности** матрицы A.

$$\nu(A) \gg 1$$

плохо обусловленные системы – это системы, близкие к вырожденным. Матрица вырожденной системы содержит линейно зависимые строки. Таким образом, матрицы, содержащие близкие строки, близки к вырожденным, то есть плохо обусловлены. Такая ситуация возникает, например, при дискретизации интегральных уравнений

Псевдорешение

$$\|A\mathbf{x} - \mathbf{b}\| = 0,$$

$$\mathbf{x} : \min_{\mathbf{y}} \|A\mathbf{y} - \mathbf{b}\|,$$

минимум, очевидно, достигается на решении $\mathbf{x}$

$$\mathbf{x} : \min_{\mathbf{y}} \|A\mathbf{y} - \mathbf{b}\|^2.$$

Плохо обусловленные матрицы

число обусловленности велико если минимальное собственное значение матрицы A^*A мало, то есть – близко к нулю. Если матрица имеет нулевое собственное значение , то уравнение $A^*A\mathbf{x} = \mathbf{0}$ имеет нетривиальное решение. При этом $det A^*A = 0$, а, следовательно, $det A = 0$, то есть матрица A вырождена. Таким образом, плохо обусловленные системы – это системы, близкие к вырожденным. Матрица вырожденной системы содержит линейно зависимые строки. Таким образом, матрицы, содержащие близкие строки, близки к вырожденным, то есть плохо обусловлены. Такая ситуация возникает, например, при дискретизации интегральных уравнений

Решение плохо обусловленных систем уравнений

$Ax + A\Delta x = b + \Delta b.$

Пусть система $Ax = b$ плохо обусловлена и пусть теперь правая часть уравнения задана с погрешностью, а именно, вместо точной правой части $\mathbf{b}$ задан вектор $\mathbf{b}^\delta$ такой, что:

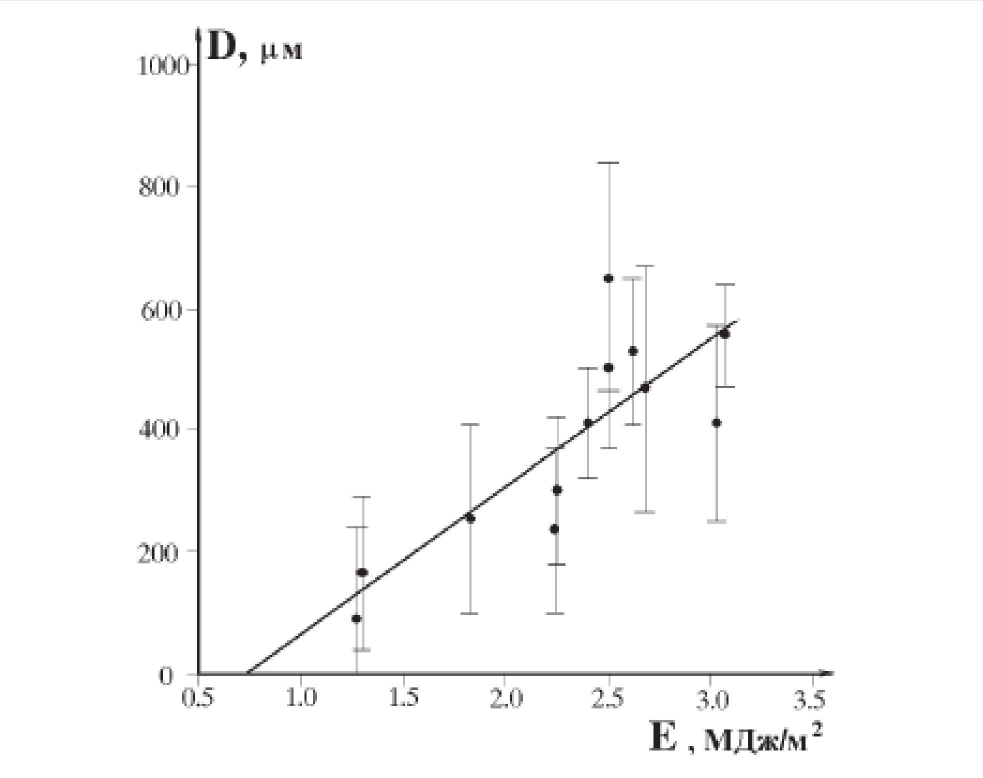
$$\frac{\|\Delta x\|}{\|x\|} \leqslant \|A\| \cdot \|A^{-1}\| \frac{\|\Delta b\|}{\|b\|}.$$

$$\|\mathbf{b}^\delta - \mathbf{b}\| = \delta.$$

Решая систему $Ax = b$ с приближенной правой частью $\mathbf{b}^\delta$, получаем приближенное решение системы. Как следует из оценки , для $\Delta\mathbf{b} = \mathbf{b}^\delta - \mathbf{b}$ даже при $\nu(A) \gg 1$ имеет место сходимость приближенного решения к точному при $\delta \to 0$, но при этом малым фиксированным значениям δ могут соответствовать существенные погрешности в решении системы $Ax = b$.

Решение плохо обусловленных систем уравнений

понятие нормального решения



$$\mathbf{x} : \min_{\mathbf{y}} \|\mathbf{y}\|^2 \quad \text{при} \quad \|A\mathbf{y} - \mathbf{b}\|^2 = 0.$$

Решение плохо обусловленных систем уравнений

ПРИМЕР

понятие нормального решения

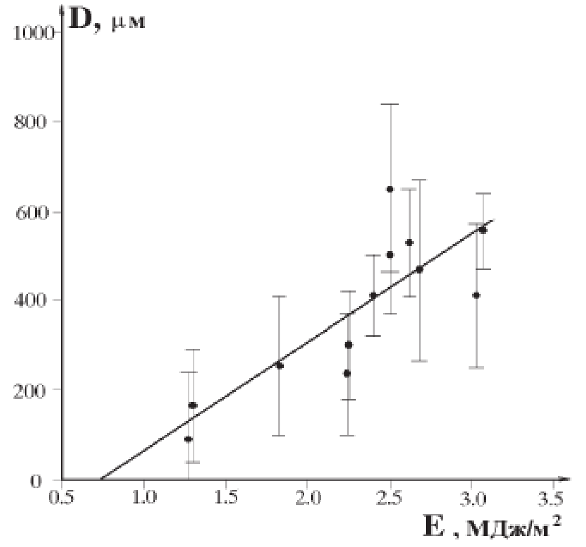


Рис. 9. Пример построения линейной зависимости методом наименьших квадратов

Основные формулы метода наименьших квадратов

$$y = a + bx,$$

Разумно предположить, что наилучшим приближением к "истинной" кривой будет такая прямая, для которой сумма

$$S = \sum_{i=1}^{n} (y_i - bx_i - a)^2$$

была бы минимальной. Здесь выражение в скобках — расстояние i-й экспериментальной точки от прямой с параметрами a и b.

На практике часто возникает необходимость найти функциональную зависимость между величинами *x* и *y*, которые получены в результате эксперимента. Часто вид эмпирической зависимости известен, но числовые параметры неизвестны.

ПРИМЕР

Известно, что минимум (или максимум) функции можно найти приравняв нулю производную по варьируемому параметру, т.е. искомые значения параметров должны удовлетворять соотношениям:

$$\frac{\partial S}{\partial b} = \sum_{i=1}^{n} (-2x_i)(y_i - bx_i - a) = 0,$$

$$\frac{\partial S}{\partial a} = \sum_{i=1}^{n} (-2)(y_i - bx_i - a) = 0.$$

Просуммировав и обратив внимание, что $\sum x_i/n$ и $\sum y_i/n$ есть просто $\overline{x}$ и $\overline{y}$ - координаты "центра тяжести" экспериментальных точек, получим

$$b\sum x_i^2 + a\sum x_i = \sum x_i y_i,$$

$$b\overline{x} + a = \overline{y}.$$

Из следует, что "наилучшая" кривая проходит через центр тяжести $(\overline{x}, \overline{y})$, т.е.

$$a = \overline{y} - b\overline{x},$$

а значение b вычисляется по формуле:

$$b = \frac{\sum x_i y_i - \overline{y}\sum x_i}{\sum x_i^2 - \overline{x}\sum x_i} = \frac{\overline{xy} - \overline{x}\cdot\overline{y}}{\overline{x^2} - \overline{x}^2}.$$

Решение плохо обусловленных систем уравнений

ПРИМЕР

Основные формулы метода наименьших квадратов

Таким образом, мы получим аналитическое выражение прямой, наилучшим образом аппроксимирующей экспериментальные точки. Поскольку число точек конечно, прямая, полученная из другой серии измерений, будет идти несколько по другому. Следуя идеологии, изложенной нами ранее, можно, используя значения s_{y_i} для точек y_i, вычислить среднеквадратичную ошибку для параметров b и a и представить окончательный ответ в виде

$$y = (b \pm \Delta b)x + (a \pm \Delta a),$$

$$\text{где } (\Delta b)^2 = \frac{\sum (y_i - bx_i - a)^2}{(n-2)(\sum x_i^2 - n\overline{x}^2)},$$

$$(\Delta a)^2 = \left(\frac{\sum x_i^2}{n}\right) \cdot (\Delta b)^2.$$

Пример: зависимость силы тока от напряжения на резисторе

ПРИМЕР

Результаты измерений силы тока и напряжения

I, A	1	2	3	4	5
U, B	2,9	6,1	9,2	11,8	16,0

Необходимо подобрать такую формулу $U = f\,(I)$, чтобы она наиболее удачно отражала зависимость между силой тока I и напряжением U.
Закон Ома устанавливает эту зависимость в виде $U = R\,I$.
Это линейная зависимость.
Какова при этом величина сопротивления R?

Алгоритм 1:

1. определить значение R для каждого из N измерений,
2. найти среднее значение сопротивления,
3. Найти погрешность такого *косвенного* измерения сопротивления.

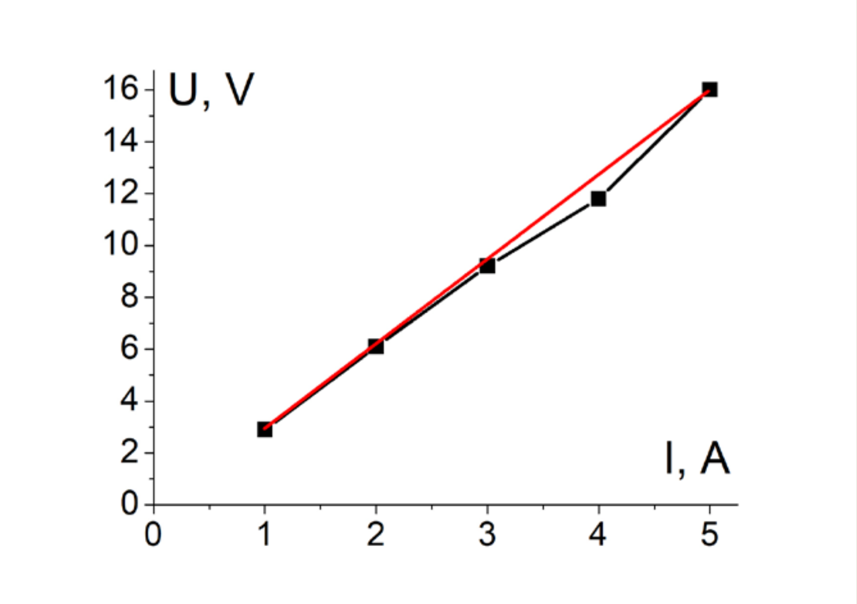
Алгоритм 2:

1. применить метод наименьших квадратов,
2. Найти погрешность

Решение плохо обусловленных систем уравнений

ПРИМЕР

понятие нормального решения



Результаты измерений силы тока и напряжения

I, A	1	2	3	4	5
U, В	2,9	6,1	9,2	11,8	16,0

Необходимо подобрать такую формулу $U = f(I)$, чтобы она наиболее удачно отражала зависимость между силой тока I и напряжением U.
Закон Ома устанавливает эту зависимость в виде $U = R\,I$.
Это линейная зависимость.
Какова при этом величина сопротивления R?

Плохо обусловленные матрицы

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ a_{n1} & \dots & a_{nn} \end{pmatrix},$$

$$\begin{array}{llll} a_{11}x_1 + a_{12}x_2 + & \dots & +a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + & \dots & +a_{2n}x_n = b_2, \\ . \quad . \quad . \quad . \quad . \quad . & \dots & . \quad . \quad . \quad . \\ . \quad . \quad . \quad . \quad . \quad . & \dots & . \quad . \quad . \quad . \\ a_{n1}x_1 + a_{n2}x_2 + & \dots & +a_{nn}x_n = b_n \end{array}$$

$$Ax = b$$

Плохо обусловленные матрицы

Пусть в конечномерном линейном пространстве $\mathbb{E}^n$ введена норма элемента $\mathbf{x}$ пространства $\mathbb{E}^n$. Определим норму $\|A\|$ матрицы A:

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

Отсюда следует, что $\|A\| \geqslant \|A\mathbf{x}\|/\|\mathbf{x}\|$, или

$$\|A\mathbf{x}\| \leqslant \|A\|\|\mathbf{x}\|.$$

Если матрица A есть произведение матриц: $A = BC$, то

$$\|A\mathbf{x}\| = \|BC\mathbf{x}\| \leqslant \|B\|\|C\mathbf{x}\| \leqslant \|B\|\|C\|\|\mathbf{x}\|.$$

Разделив на $\|\mathbf{x}\| \neq 0$, получим

$$\frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \leqslant \|B\|\|C\|,$$

откуда следует неравенство

$$\|A\| = \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \leqslant \|B\|\|C\|.$$

Плохо обусловленные матрицы

погрешность $\Delta \mathbf{b} \in E^n$. $A(\mathbf{x} + \Delta\mathbf{x}) = \mathbf{b} + \Delta\mathbf{b}$,

где $\mathbf{x} + \Delta\mathbf{x}$ – приближенное решение системы $A\mathbf{x} = \mathbf{b}$ Отсюда в силу линейности операции умножения матрицы на вектор

$$A\mathbf{x} + A\Delta\mathbf{x} = \mathbf{b} + \Delta\mathbf{b}.$$

Так как $A\mathbf{x} = \mathbf{b}$, то $A\Delta\mathbf{x} = \Delta\mathbf{b}$, или $\Delta\mathbf{x} = A^{-1}\Delta\mathbf{b}$ и, таким образом, в силу $\|A\mathbf{x}\| \leqslant \|A\|\|\mathbf{x}\|$

$$\|\Delta\mathbf{x}\| \leqslant \|A^{-1}\| \cdot \|\Delta\mathbf{b}\|.$$

Применяя неравенство $\|A\mathbf{x}\| \leqslant \|A\|\|\mathbf{x}\|$ получим

$$A\mathbf{x} = \mathbf{b} \longrightarrow \|\mathbf{b}\| \leqslant \|A\| \cdot \|\mathbf{x}\|.$$

Перемножая последние два неравенства и разделив результат на $\|\mathbf{x}\|\|\mathbf{b}\|$ получаем оценку

$$\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} \leqslant \|A\| \cdot \|A^{-1}\| \frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|}.$$

называется **числом обусловленности** матрицы A.

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

Плохо обусловленные матрицы

$$\|E\| = \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|E\mathbf{x}\|}{\|\mathbf{x}\|} = \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{x}\|}{\|\mathbf{x}\|} = 1.$$

Так как $E = A^{-1}A$, то используя неравенство $\|A\| = \sup\limits_{\mathbf{x}\neq 0} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \leqslant \|B\|\|C\|.$

$$\nu(A) = \|A\|\|A^{-1}\| \geqslant \|A^{-1}A\| = \|E\| = 1.$$

Очевидно, приближенное решение $\mathbf{x} + \Delta\mathbf{x}$ системы $A\mathbf{x} = \mathbf{b}$ тем лучше, чем меньше относительная погрешность $\|\Delta\mathbf{x}\|/\|\mathbf{x}\|$ приближенного решения, и, как следует из оценки ( ), при данной относительной погрешности правой части $\|\Delta\mathbf{b}\|/\|\mathbf{b}\|$, — чем меньше число обусловленности $\nu(A)$.

$$\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} \leqslant \|A\| \cdot \|A^{-1}\| \frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|}.$$

Плохо обусловленные матрицы

Получим выражение нормы матрицы A (и таким образом, и числа обусловленности $\nu(A)$ матрицы) через собственные значения матрицы A.

Введем скалярное произведение двух векторов $\mathbf{x}$ и $\mathbf{y}$ в унитарном пространстве E^n

$$(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} x_i \bar{y}_i, \quad \mathbf{x}, \mathbf{y} \in E^n$$

и норму вектора $\mathbf{x}$

$$\|\mathbf{x}\| = \sqrt{(\mathbf{x}, \mathbf{x})}.$$

Плохо обусловленные матрицы

$$\|\mathbf{x}\| = \sqrt{(\mathbf{x}, \mathbf{x})}.$$

Согласно определению нормы матрицы $\|A\| = \sup\limits_{\mathbf{x}\neq 0} \dfrac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}$. пространстве E^n (конечномерном линейном пространстве со скалярным произведением) для нормы матрицы получаем

$$\|A\| = \sup_{\mathbf{x}\neq 0} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} = \sup_{\mathbf{x}\neq 0} \frac{\sqrt{(A\mathbf{x}, A\mathbf{x})}}{\sqrt{(\mathbf{x}, \mathbf{x})}} = \\ = \sqrt{\sup_{\mathbf{x}\neq 0} \frac{(A\mathbf{x}, A\mathbf{x})}{(\mathbf{x}, \mathbf{x})}} = \sqrt{\sup_{\mathbf{x}\neq 0} \frac{(A^* A\mathbf{x}, \mathbf{x})}{(\mathbf{x}, \mathbf{x})}},$$

где A^*– матрица, сопряженная к матрице A.

Плохо обусловленные матрицы

Получим выражение элементов сопряжённой матрицы A^* через элементы матрицы A. Пользуясь определением скалярного произведения, меняя порядок суммирования, получаем

$$(A\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} \left(\sum_{k=1}^{n} a_{ik} x_k \right) \overline{y_i} = \sum_{k=1}^{n} \sum_{i=1}^{n} a_{ik} x_k \overline{y_i}.$$

Меняя порядок индексов $a_{ik} = a^T_{ki}$, где a^T_{ki} – коэффициенты транспонированной матрицы, и учитывая, что $a^T_{ki} = \overline{\overline{a^T_{ki}}}$, получаем

$$\begin{aligned}(A\mathbf{x}, \mathbf{y}) &= \sum_{k=1}^{n} x_k \sum_{i=1}^{n} a_{ik} \overline{y_i} = \sum_{k=1}^{n} x_k \sum_{i=1}^{n} \overline{\overline{a^T_{ki}}}\, \overline{y_i} = \\ &= \sum_{k=1}^{n} x_k \overline{\sum_{i=1}^{n} \overline{a^T_{ki}} y_i} = \sum_{k=1}^{n} x_k \overline{\sum_{i=1}^{n} a^*_{ki} y_i} = (\mathbf{x}, A^*\mathbf{y}).\end{aligned}$$

Плохо обусловленные матрицы

Мы обозначили $a^*_{ki} = \overline{a^T_{ki}}$ и, таким образом, для сопряженной матрицы получаем $A^* = \overline{A^T}$. Очевидно, если коэффициенты матрицы – действительные числа, то сопряженная матрица – это просто транспонированная матрица.

Заметим, что матрица A^*A самосопряженная:

$$(A^*A)^* = \overline{(\overline{A^T}A)^T} = (A^T\overline{A})^T = \overline{A^T}(A^T)^T = A^*A.$$

Здесь мы воспользовались тем свойством операции транспонирования, что если $B = CD$, то $B^T = D^T C^T$. В самом деле

$$b^T_{ij} = b_{ji} = \sum_{k=1}^{n} c_{jk} d_{ki} = \sum_{k=1}^{n} d_{ki} c_{jk} = \sum_{k=1}^{n} d^T_{ik} c^T_{kj}.$$

Плохо обусловленные матрицы

У всякой матрицы A имеется n собственных значений (с учетом кратности, в частности, может быть n собственных значений одной кратности) и n собственных и присоединенных векторов. Всякая же самосопряженная матрица B $(B = B^*)$ имеет n попарно ортогональных собственных векторов [18, стр. 137]. Пусть $\mathbf{e}_k$ – собственные вектора матрицы B, а $\lambda_{B,k}$ – соответствующие собственные значения (среди которых могут быть кратные), то есть

$$B\mathbf{e}_k = \lambda_{B,k}\mathbf{e}_k, \quad k = 1, 2, ..., n.$$

Очевидно, собственные вектора находятся с точностью до множителя, поэтому можно считать их нормированными. Отсюда следует, что система $\{\mathbf{e}_k\}_{k=1}^{k=n}$ собственных векторов матрицы B образует ортонормированный базис в E^n.

[18] *Ильин В.А., Позняк Э.Г.* Линейная алгебра. М.:Наука, 1974. 296 с.

Плохо обусловленные матрицы

У всякой самосопряженной матрицы B собственные значения действительны

$$\lambda_{B,k} = \lambda_{B,k}(\mathbf{e}_k, \mathbf{e}_k) = (B\mathbf{e}_k, \mathbf{e}_k) = (\mathbf{e}_k, B\mathbf{e}_k) = \overline{\lambda_{B,k}}(\mathbf{e}_k, \mathbf{e}_k) = \overline{\lambda_{B,k}}.$$

Теперь для любого ненулевого вектора $\mathbf{x} \in E^n$, имея в виду разложение по базису

$$\mathbf{x} = \sum_{k=1}^{n} \alpha_k \mathbf{e}_k,$$

получаем оценку

$$(B\mathbf{x}, \mathbf{x}) = \Big(\sum_{k=1}^{n} \lambda_{B,k} \alpha_k \mathbf{e}_k, \sum_{k=1}^{n} \alpha_k \mathbf{e}_k\Big) =$$

$$\sum_{k=1}^{n} \lambda_{B,k} |\alpha_k|^2 \leqslant \max_k \lambda_{B,k} (\mathbf{x}, \mathbf{x}).$$

Плохо обусловленные матрицы

Разделив на $(\mathbf{x},\mathbf{x})$ и считая, что $\max\limits_{1\leqslant k\leqslant n} \lambda_{B,k} = \lambda_{B,m}$, получим

$$\frac{(B\mathbf{x},\mathbf{x})}{(\mathbf{x},\mathbf{x})} \leqslant \max_k \lambda_{B,k} = \lambda_{B,m},$$

или

$$\sup_{\mathbf{x}\neq 0} \frac{(B\mathbf{x},\mathbf{x})}{(\mathbf{x},\mathbf{x})} \leqslant \max_k \lambda_{B,k} = \lambda_{B,m}.$$

Имея в виду, что величина $\lambda_{B,m}$ достигается на векторе $\mathbf{e}_m$:

$$\frac{(B\mathbf{e}_m,\mathbf{e}_m)}{(\mathbf{e}_m,\mathbf{e}_m)} = \lambda_{B,m},$$

получаем, что для самосопряженной матрицы B

$$\sup_{\mathbf{x}\neq 0} \frac{(B\mathbf{x},\mathbf{x})}{(\mathbf{x},\mathbf{x})} = \lambda_{B,m} = \max_k \lambda_{B,k}.$$

Плохо обусловленные матрицы

У самосопряженной матрицы $B = A^*A$ все собственные значения $\lambda_{A^*A,k}$ положительны (матрица A^*A положительна). В самом деле,

$$\lambda_{A^*A,k} = \lambda_{A^*A,k}(\mathbf{e}_k, \mathbf{e}_k) = (A^*A\mathbf{e}_k, \mathbf{e}_k) = (A\mathbf{e}_k, A\mathbf{e}_k) \geqslant 0.$$

Кроме того, $\lambda = 0$ не является собственным значением матрицы A^*A, так как в противном случае система уравнений

$$A^*A\mathbf{x} = 0$$

имеет ненулевое решение ($det A^*A = det A^* det A = det \bar{A}^T det A = det \bar{A}\, det A = |det A|^2 \neq 0$ и решение только тривиальное). Следовательно, все собственные значения матрицы A^*A положительны

$$\lambda_{A^*A,k} > 0, \quad k = 1, ..., n.$$

Плохо обусловленные матрицы

$$\|A\| = \sup_{\mathbf{x}\neq 0} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} = \sup_{\mathbf{x}\neq 0} \frac{\sqrt{(A\mathbf{x}, A\mathbf{x})}}{\sqrt{(\mathbf{x}, \mathbf{x})}} = \sqrt{\sup_{\mathbf{x}\neq 0} \frac{(A\mathbf{x}, A\mathbf{x})}{(\mathbf{x}, \mathbf{x})}} = \sqrt{\sup_{\mathbf{x}\neq 0} \frac{(A^*A\mathbf{x}, \mathbf{x})}{(\mathbf{x}, \mathbf{x})}},$$

$$\sup_{\mathbf{x}\neq 0} \frac{(B\mathbf{x}, \mathbf{x})}{(\mathbf{x}, \mathbf{x})} = \lambda_{B,m} = \max_k \lambda_{B,k}.$$

при $B = A^*A$ и $\lambda_{A^*A,k} > 0$ норму матрицы $\|A\|$ выражаем через максимальное собственное значение матрицы A^*A:

$$\|A\| = \sqrt{\sup_{\mathbf{x}\neq 0} \frac{(A^*A\mathbf{x}, \mathbf{x})}{(\mathbf{x}, \mathbf{x})}} = \sqrt{\max_i \lambda_{A^*A,i}}.$$

В случае самосопряженной матрицы A (то есть – $A = A^*$)

$$\lambda_{A^*A,i} = \lambda_{A^2,i} = (\lambda_{A,i})^2$$

и норма матрицы $\|A\|$ при этом выражается через максимальное по модулю собственное значение матрицы A:

$$\|A\| = \sqrt{\max_i (\lambda_{A,i})^2} = \max_i |\lambda_{A,i}|.$$

Плохо обусловленные матрицы

$\sup\limits_{\mathbf{x}\neq 0}\dfrac{(B\mathbf{x},\mathbf{x})}{(\mathbf{x},\mathbf{x})} = \lambda_{B,m} = \max\limits_k \lambda_{B,k}$. для самосопряженной положительной матрицы $B = A^*A$ получаем (из $\lambda_{A^*A,k} > 0$ следует, что $\min\limits_i \lambda_{A^*A,i} > 0$)

$$\sup_{\mathbf{x}\neq 0}\frac{(\mathbf{x},\mathbf{x})}{(B\mathbf{x},\mathbf{x})} = \frac{1}{\min \lambda_{A^*A,i}}$$

и, таким образом,

$$||A^{-1}|| = \sqrt{\sup_{\mathbf{b}\neq 0}\frac{(A^{-1}\mathbf{b},A^{-1}\mathbf{b})}{(\mathbf{b},\mathbf{b})}} = \sqrt{\sup_{\mathbf{x}\neq 0(det A\neq 0)}\frac{(\mathbf{x},\mathbf{x})}{(A\mathbf{x},A\mathbf{x})}} =$$

$$= \sqrt{\sup_{\mathbf{x}\neq 0}\frac{(\mathbf{x},\mathbf{x})}{(A^*A\mathbf{x},\mathbf{x})}} = \sqrt{\frac{1}{\min\limits_i \lambda_{A^*A,i}}}.$$

Плохо обусловленные матрицы

$$\|A\| = \sqrt{\sup_{x \neq 0} \frac{(A^* A\mathbf{x}, \mathbf{x})}{(\mathbf{x}, \mathbf{x})}} = \sqrt{\max_i \lambda_{A^* A, i}}$$

$$\|A^{-1}\| = \sqrt{\sup_{b \neq 0} \frac{(A^{-1}\mathbf{b}, A^{-1}\mathbf{b})}{(\mathbf{b}, \mathbf{b})}} = \sqrt{\frac{1}{\min_i \lambda_{A^* A, i}}}$$

Теперь для числа обусловленности $\lambda_{A^* A, i}$ с использованием получаем

$$\nu(A) = \sqrt{\frac{\max_i \lambda_{A^* A, i}}{\min_i \lambda_{A^* A, i}}} \geqslant 1.$$

Для самосопряженной матрицы $A = A^*$ отсюда, используя $\lambda_{A^* A, i} = \lambda_{A^2, i} = (\lambda_{A,i})^2$ получаем

$$\nu(A) = \sqrt{\frac{\max_i {\lambda_{A,i}}^2}{\min_i {\lambda_{A,i}}^2}} = \frac{\max_i |\lambda_{A,i}|}{\min_i |\lambda_{A,i}|} \geqslant 1.$$

При $\nu(A) \gg 1$ матрица A и система (1.1) называются **плохо обусловленными**.

Плохо обусловленные матрицы

При $\det A \neq 0,$
система уравнений **Ax=b**
имеет единственное решение.
Численные методы позволяют получать приближенные решения системы .
Рассмотрим вопрос устойчивости решения **x** по отношению к погрешностям вектора **b**.

Плохо обусловленные матрицы

Пусть в конечномерном линейном пространстве $\mathbb{E}^n$ введена норма элемента $\mathbf{x}$ пространства $\mathbb{E}^n$. Определим норму $\|A\|$ матрицы A:

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

Величина

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

называется **числом обусловленности** матрицы A.

При $\nu(A) \gg 1$
матрица A и система уравнений Ax=b называются **плохо обусловленными.**

Плохо обусловленные системы

Величина

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

называется **числом обусловленности** матрицы A.

$$\nu(A) \gg 1$$

плохо обусловленные системы – это системы, близкие к вырожденным. Матрица вырожденной системы содержит линейно зависимые строки. Таким образом, матрицы, содержащие близкие строки, близки к вырожденным, то есть плохо обусловлены. Такая ситуация возникает, например, при дискретизации интегральных уравнений

Псевдорешение

$$\|A\mathbf{x} - \mathbf{b}\| = 0,$$

$$\mathbf{x} : \min_{\mathbf{y}} \|A\mathbf{y} - \mathbf{b}\|,$$

минимум, очевидно, достигается на решении $\mathbf{x}$

$$\mathbf{x} : \min_{\mathbf{y}} \|A\mathbf{y} - \mathbf{b}\|^2.$$

Плохо обусловленные матрицы

число обусловленности велико если минимальное собственное значение матрицы A^*A мало, то есть – близко к нулю. Если матрица имеет нулевое собственное значение , то уравнение $A^*A\mathbf{x} = \mathbf{0}$ имеет нетривиальное решение. При этом $det A^*A = 0$, а, следовательно, $det A = 0$, то есть матрица A вырождена. Таким образом, плохо обусловленные системы – это системы, близкие к вырожденным. Матрица вырожденной системы содержит линейно зависимые строки. Таким образом, матрицы, содержащие близкие строки, близки к вырожденным, то есть плохо обусловлены. Такая ситуация возникает, например, при дискретизации интегральных уравнений

Решение плохо обусловленных систем уравнений

$$\boxed{A\mathbf{x} + A\Delta\mathbf{x} = \mathbf{b} + \Delta\mathbf{b}.}$$

Пусть система $A\mathbf{x} = \mathbf{b}$ плохо обусловлена и пусть теперь правая часть уравнения задана с погрешностью, а именно, вместо точной правой части $\mathbf{b}$ задан вектор $\mathbf{b}^\delta$ такой, что:

$$\boxed{\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} \leqslant \|A\| \cdot \|A^{-1}\| \frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|}.}$$

$$\|\mathbf{b}^\delta - \mathbf{b}\| = \delta.$$

Решая систему $A\mathbf{x} = \mathbf{b}$ с приближенной правой частью $\mathbf{b}^\delta$, получаем приближенное решение системы. Как следует из оценки , для $\Delta\mathbf{b} = \mathbf{b}^\delta - \mathbf{b}$ даже при $\nu(A) \gg 1$ имеет место сходимость приближенного решения к точному при $\delta \to 0$, но при этом малым фиксированным значениям δ могут соответствовать существенные погрешности в решении системы $A\mathbf{x} = \mathbf{b}$.

Решение плохо обусловленных систем уравнений

$$\mathbf{x} : \min_{\mathbf{y}} \|\mathbf{y}\|^2 \quad \text{при} \quad \|A\mathbf{y} - \mathbf{b}^\delta\|^2 = \delta^2,$$

Точка экстремума $\mathbf{x}^*$ определяется как точка минимума функции Лагранжа

$$\mathbf{x}^* : \min_{\mathbf{y}} \left(\|\mathbf{y}\|^2 + \beta(\|A\mathbf{y} - \mathbf{b}^\delta\|^2 - \delta^2) \right),$$

причем множитель Лагранжа β определяется из условия

$$\beta : \|A\mathbf{x}^* - \mathbf{b}^\delta\| = \delta.$$

Решение плохо обусловленных систем уравнений

$$\mathbf{x}^* : \min_{\mathbf{y}} \left(\|\mathbf{y}\|^2 + \beta(\|A\mathbf{y} - \mathbf{b}^\delta\|^2 - \delta^2) \right),$$

убрать слагаемое $-\beta\delta^2$ и заменить $\beta = 1/\alpha$

$$\boxed{\mathbf{x}^* : \min_{\mathbf{y}} \left(\|A\mathbf{y} - \mathbf{b}^\delta\|^2 + \alpha\|\mathbf{y}\|^2 \right), \ \alpha > 0,}$$

Минимизируемая функция называется **функционалом Тихонова**.

где параметр α определяется из условия, называемого **условием невязки**

$$\alpha : \|A\mathbf{x}^* - \mathbf{b}^\delta\| = \delta.$$

Решение плохо обусловленных систем уравнений

для производной

$$\frac{\partial}{\partial y_j}\left(\|A\mathbf{y}-\mathbf{b}^\delta\|^2+\alpha\|\mathbf{y}\|^2\right)=$$

$$=2\sum_{l=1}^{n}\left(\sum_{i=1}^{n}A^T_{ji}A_{il}\right)y_l-2\sum_{i=1}^{n}A^T_{ji}b^\delta_i+2\alpha y_j=$$

$$=2\sum_{l=1}^{n}\left(\sum_{i=1}^{n}A^T_{ji}A_{il}\right)y_l-2\sum_{i=1}^{n}A^T_{ji}b^\delta_i+2\alpha\sum_{l=1}^{n}\delta_{jl}y_l,\quad j=1,\ldots,n,$$

δ_{jl} – символ Кронекера

$$A^*A\mathbf{y}+\alpha E\mathbf{y}=A^*\mathbf{b}^\delta,$$

$$(A^*A+\underline{\alpha E})\mathbf{y}=A^*\mathbf{b}^\delta.$$

$$A^*A\mathbf{y}=A^*\mathbf{b}.$$

$\alpha=0$

Решение плохо обусловленных систем уравнений

Для второго дифференциала, если $\alpha > 0$, получаем положительно определенную форму так как матрица $A^*A + \alpha E$ положительна.

$$2\sum_{k=1}^{n}\sum_{l=1}^{n}\left(\sum_{i=1}^{n} A_{ki}^{T} A_{il} + \alpha\delta_{kl}\right) dy_k dy_l = ((A^*A + \alpha E)d\mathbf{y}, d\mathbf{y}) > 0,$$

Поэтому решение $(A^*A + \alpha E)\mathbf{y} = A^*\mathbf{b}^\delta$. дает точку минимума функции

$$\mathbf{x}^* : \min_{\mathbf{y}} \left(\|A\mathbf{y} - \mathbf{b}^\delta\|^2 + \alpha\|\mathbf{y}\|^2\right), \ \alpha > 0,$$

Решение плохо обусловленных систем уравнений

При $\delta = 0$ и $\alpha > 0$

ает точного решения

$\|\mathbf{b}^\delta - \mathbf{b}\| = \delta.$

при согласовании параметра α с погрешностью δ сходится к точному решению $A\mathbf{x} = \mathbf{b}$ при $\delta \to 0.$

Метод регуляризации Тихонова

$$(A^*A + \alpha E)\mathbf{y} = A^*\mathbf{b}^\delta.$$

где параметр α определяется из условия, называемого **условием невязки**

$$\alpha : \|A\mathbf{x}^* - \mathbf{b}^\delta\| = \delta.$$

параметр регуляризации.

Решение плохо обусловленных систем уравнений

$$\nu(A) = \sqrt{\frac{\max\limits_i \lambda_{A^*A,i}}{\min\limits_i \lambda_{A^*A,i}}} \geqslant 1.$$

Для каждого фиксированного значения погрешности δ эффект регуляризации, в частности, состоит в том, что минимальное собственное значение матрицы A^*A (в знаменателе увеличивается на величину параметра α, в то время как максимальное собственное значение, увеличенное на ту же величину α, меняется незначительно.

Решение плохо обусловленных систем уравнений

В самом деле, если λ – собственное значение матрицы A^*A, соответствующее собственному вектору $\mathbf{e}$:

$$(A^*A - \lambda E)\mathbf{e} = \mathbf{0},$$

то отсюда следует, что величина $\lambda + \alpha$ является собственным значением матрицы $A^*A + \alpha E$:

$$((A^*A + \alpha E) - (\lambda + \alpha)E)\mathbf{e} = \mathbf{0}.$$

Таким образом, даже небольшие значения параметра α могут существенно уменьшить число обусловленности матрицы.

Решение плохо обусловленных систем уравнений

ПРИМЕР

$$\begin{pmatrix} 1 & 1.005 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

Определитель системы отличен от нуля и ее единственное решение

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

Решение плохо обусловленных систем уравнений

ПРИМЕР

Изменим «незначительно» правую часть системы и будем решать систему уравнений $A\mathbf{x} = \mathbf{b}^\delta$

$$\begin{pmatrix} 1 & 1.005 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2,005 \end{pmatrix}.$$

Очевидно, абсолютная погрешность правой части системы равна

$$\|\mathbf{b}^\delta - \mathbf{b}\| = \delta = 0.005,$$

а величина относительной погрешности правой части системы составляет

$$\frac{\|\Delta \mathbf{b}\|}{\|\mathbf{b}\|} = \frac{\|\mathbf{b}^\delta - \mathbf{b}\|}{\|\mathbf{b}\|} = \frac{\sqrt{(2-2)^2 + (2.005-2)^2}}{\sqrt{2^2 + 2^2}} \approx 0.0017 = 0,17\%.$$

Решение плохо обусловленных систем уравнений

ПРИМЕР

$$\begin{pmatrix} 1 & 1.005 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2,005 \end{pmatrix}.$$

приближенным решением является вектор

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3.005 \\ -1 \end{pmatrix}$$

В чем причина такого результата?

который далек от точного решения исходной системы с невозмущенной правой частью. В самом деле, величина относительной погрешности приближенного решения системы составляет

$$\frac{\|\Delta \mathbf{x}\|}{\|\mathbf{x}\|} = \frac{\sqrt{(3.005-2)^2 + (-1-0)^2}}{\sqrt{2^2+0^2}} \approx 0.7089 = 70,89\%$$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

Решение плохо обусловленных систем уравнений

ПРИМЕР

Найдем число обусловленности матрицы системы. Для этого необходимо найти собственные значения матрицы A^*A. Вычислим матрицу A^*A (значение элемента a_{22} приведено округленно):

$$A^*A = \begin{pmatrix} 2. & 2.005 \\ 2.005 & 2.010 \end{pmatrix}.$$

Характеристическое (квадратное относительно λ) уравнение для этой матрицы, которое имеет вид $det(A^*A - \lambda E) = 0$, или в данном случае

$$\begin{vmatrix} 2. - \lambda & 2.005 \\ 2.005 & 2.010 - \lambda \end{vmatrix} = 0,$$

дает корни $\lambda_{max} \approx 4.010$, $\lambda_{min} \approx 1.248 \cdot 10^{-5}$ и тогда для числа обусловленности из (1.18) получаем

число обусловленности много больше единицы

$$\nu(A) = \sqrt{\frac{\lambda_{max}}{\lambda_{min}}} \approx \sqrt{\frac{4.010}{1.248 \cdot 10^{-5}}} \approx 566.7 \gg 1.$$

Решение плохо обусловленных систем уравнений

ПРИМЕР

$$\frac{\|\Delta \mathbf{x}\|}{\|\mathbf{x}\|} \leqslant \nu(A)\frac{\|\Delta \mathbf{b}\|}{\|\mathbf{b}\|} \approx 566.7 \cdot 0.0017 \approx 1.002 = 100.2\%,$$

Заметим, что на практике при решении возмущенной системы точная правая часть $\mathbf{b}$ нам не известна

Решение плохо обусловленных систем уравнений

ПРИМЕР

$$\begin{pmatrix} 1 & 1.005 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

$$(A^* A + \alpha E)\mathbf{x} = A^* \mathbf{b}^\delta$$

$$\begin{pmatrix} 2. + \alpha & 2.005 \\ 2.005 & 2.010 + \alpha \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4.005 \\ 4.015 \end{pmatrix}$$

Решение плохо обусловленных систем уравнений

ПРИМЕР

Приведем таблицу значений невязки $\|A\mathbf{x}_\alpha - \mathbf{b}^\delta\|$ и координат приближенного решения $\mathbf{x}_\alpha$ в зависимости от значений параметра регуляризации α

α	x_1	x_2	невязка
0.0	3.005	−1.000	0.00000
$3.0 \cdot 10^{-6}$	2.353	−0.350	0.00230
$6.0 \cdot 10^{-6}$	2.021	−0.019	0.00347
$6.25 \cdot 10^{-6}$	2.001	0.002	0.00354
$9.0 \cdot 10^{-6}$	1.820	0.182	0.00418
$12.0 \cdot 10^{-6}$	1.685	0.317	0.00466
$15.0 \cdot 10^{-6}$	1.588	0.414	0.00500
$18.0 \cdot 10^{-6}$	1.515	0.486	0.00525

$\boxed{\alpha = 0}$

Решение плохо обусловленных систем уравнений

ПРИМЕР

α	x_1	x_2	невязка
$12.0 \cdot 10^{-6}$	1.685	0.317	0.00466
$15.0 \cdot 10^{-6}$	1.588	0.414	0.00500
$18.0 \cdot 10^{-6}$	1.515	0.486	0.00525

Величина невязки, приблизительно равная погрешности правой части $\delta = 0.005$, соответствует величине параметра $\alpha = 15.0 \cdot 10^{-6}$. Соответствующее приближенное решение $x_1 = 1.588$ $x_2 = 0.414$. Относительная погрешность этого приближенного решения составляет

$$\frac{\|\Delta \mathbf{x}\|}{\|\mathbf{x}\|} = \frac{\sqrt{(1.588-2)^2 + (0.414-0)^2}}{\sqrt{2^2+0^2}} \approx 0.2920 = 29.2\%.$$

Решение плохо обусловленных систем уравнений

ПРИМЕР

α	x_1	x_2	невязка
$6.0 \cdot 10^{-6}$	2.021	−0.019	0.00347
$6.25 \cdot 10^{-6}$	2.001	0.002	0.00354
$9.0 \cdot 10^{-6}$	1.820	0.182	0.00418

Отметим, что однопараметрическое $\mathbf{x}_\alpha$ семейство приближенных решений, приведенных в таблице, содержит и более точное

приближенное решение $x_1 = 2.000612$ $x_2 = 0.00187$, которое соответствует величине параметра регуляризации $\alpha = 6.25 \cdot 10^{-6}$. Относительная погрешность этого приближенного решения составляет

$$\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} = \frac{\sqrt{(2.000612 - 2)^2 + (0.00187 - 0)^2}}{\sqrt{2^2 + 0^2}} \approx 0.0010 = 0.1\%.$$