Системы линейных алгебраических уравнений
Часть 1. Матричные нормы. Обусловленность

Скалько Юрий Иванович
Цыбулин Иван

Постановка задачи

Дана квадратная матрица $\mathbf{A}$ и столбец $\mathbf{b}$

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}, \quad \det \mathbf{A} \neq 0$$

Решить систему линейных алгебраических уравнений

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

Вырожденные матрицы

Хорошо известно, что не все системы с вырожденными матрицами $\mathbf{A}$ имеют решение. Даже если некоторая система с вырожденной $\mathbf{A}$ имеет решение, достаточно немного изменить правую часть, чтобы нарушить условие совместности.

Поскольку все вычисления производятся с некоторой точностью, то на практике матрица никогда не бывает строго вырожденной ($\det \mathbf{A} = 0$).

Система, близкая к вырожденной

Рассмотрим пример двух немного отличающихся систем

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 4.0001 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \qquad \mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 4.0001 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} 1 \\ 2.0001 \end{pmatrix}$$

Решения этих систем отличаются уже существенно

$$\mathbf{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \mathbf{x} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Причина существенного различия в том, что матрица $\mathbf{A}$ «близка» к вырожденной матрице

$$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$$

Векторные нормы

В пространстве векторов $\mathbb{R}^n$ можно ввести много разных норм. Самыми распространенными в вычислительной математике являются следующие три нормы:

$$\|\mathbf{x}\|_\infty = \max_i |x_i|$$

$$\|\mathbf{x}\|_{\ell_1} = \sum_i |x_i|$$

$$\|\mathbf{x}\|_E = \sqrt{(\mathbf{x}, \mathbf{x})} = \sqrt{\sum_i x_i^2}$$

Векторные нормы

В пространстве векторов $\mathbb{R}^n$ можно ввести много разных норм. Самыми распространенными в вычислительной математике являются следующие три нормы:

$$\|\mathbf{x}\|_\infty = \max_i |x_i|$$

$$\|\mathbf{x}\|_{\ell_1} = \sum_i |x_i|$$

$$\|\mathbf{x}\|_E = \sqrt{(\mathbf{x}, \mathbf{x})} = \sqrt{\sum_i x_i^2}$$

Например, для $\mathbf{x} = \begin{pmatrix} 3 \\ -4 \end{pmatrix}$

$$\|\mathbf{x}\|_\infty = 4, \quad \|\mathbf{x}\|_{\ell_1} = 7, \quad \|\mathbf{x}\|_E = 5$$

Подчиненные матричные нормы

Определим норму матрицы $\|\mathbf{A}\|$ таким образом, чтобы для любого $\mathbf{x}$ было справедливо

$$\|\mathbf{y}\| \leqslant \|\mathbf{A}\|\|\mathbf{x}\|, \qquad \mathbf{y} = \mathbf{A}\mathbf{x},$$

причем $\|\mathbf{A}\|$ — минимальное такое число.

Определение

Норма матрицы $\|\mathbf{A}\|$ называется подчиненной векторной норме $\|\mathbf{x}\|$, если

$$\|\mathbf{A}\| = \sup_{\|\mathbf{x}\|=1} \|\mathbf{A}\mathbf{x}\|.$$

В левой части равенства стоит матричная норма $\mathbf{A}$, а в правой — векторные нормы $\mathbf{x}$ и $\mathbf{A}\mathbf{x}$

Различные векторные нормы порождают различные матричные нормы.

Матричная норма $\|\cdot\|_\infty$

Точка $\mathbf{x} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ с нормой $\|\mathbf{x}\|_\infty = 1$ переходит в точку

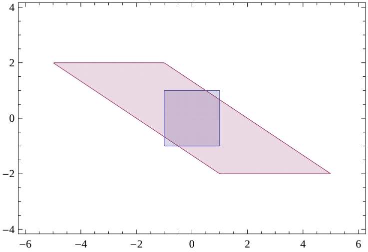
$$\mathbf{y} = \mathbf{A}\mathbf{x} = \begin{pmatrix} 5 \\ -2 \end{pmatrix}$$

с нормой $\|\mathbf{y}\|_\infty = 5$.
Для остальных точек $\|\mathbf{x}\|_\infty = 1$

$$\|\mathbf{A}\mathbf{x}\|_\infty \leqslant 5$$

$$\|\mathbf{A}\|_\infty = 5$$

$$\mathbf{A} = \begin{pmatrix} 2 & -3 \\ 0 & 2 \end{pmatrix}$$



В общем случае

$$\|\mathbf{A}\|_\infty = \max_i \sum_j |a_{ij}|$$

Матричная норма $\|\cdot\|_{\ell_1}$

Точка $\mathbf{x} = \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$ с нормой $\|\mathbf{x}\|_{\ell_1} = 1$ переходит в точку

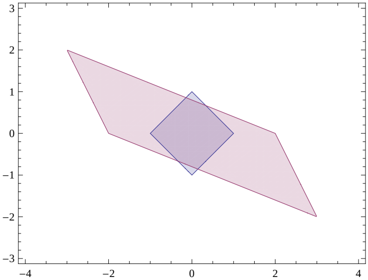
$$\mathbf{y} = \mathbf{A}\mathbf{x} = \begin{pmatrix}-3\\2\end{pmatrix}$$

с нормой $\|\mathbf{y}\|_{\ell_1} = 5$.
Для остальных точек $\|\mathbf{x}\|_{\ell_1} = 1$

$$\|\mathbf{A}\mathbf{x}\|_{\ell_1} \leqslant 5$$

$$\|\mathbf{A}\|_{\ell_1} = 5$$

$$\mathbf{A} = \begin{pmatrix}2 & -3\\0 & 2\end{pmatrix}$$



В общем случае

$$\|\mathbf{A}\|_{\ell_1} = \max_j \sum_i |a_{ij}| = \|\mathbf{A}^\mathsf{T}\|_\infty$$

Матричная норма $\|\cdot\|_E$

Точка $\mathbf{x} = \frac{1}{\sqrt{5}}\binom{-1}{2}$ с нормой $\|\mathbf{x}\|_E = 1$ переходит в точку

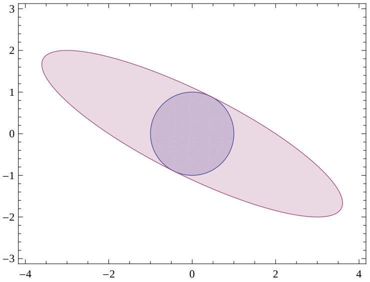
$$\mathbf{A} = \begin{pmatrix} 2 & -3 \\ 0 & 2 \end{pmatrix}$$

$$\mathbf{y} = \mathbf{A}\mathbf{x} = \frac{1}{\sqrt{5}}\begin{pmatrix} -8 \\ 4 \end{pmatrix}$$

с нормой $\|\mathbf{y}\|_E = 4$.
Для остальных точек $\|\mathbf{x}\|_E = 1$

$$\|\mathbf{A}\mathbf{x}\|_E \leqslant 4$$

$$\|\mathbf{A}\|_E = 4$$



В общем случае

$$\|\mathbf{A}\|_E = \sigma_{\max}(\mathbf{A}) = \sqrt{\max \lambda(\mathbf{A}^\mathsf{T}\mathbf{A})}$$

Матричная норма $\|\cdot\|_E$

$$\mathbf{A} = \begin{pmatrix} 2 & -3 \\ 0 & 2 \end{pmatrix}, \qquad \mathbf{A}^\mathsf{T}\mathbf{A} = \begin{pmatrix} 4 & -6 \\ -6 & 13 \end{pmatrix}$$

Собственные числа матрицы $\mathbf{A}^\mathsf{T}\mathbf{A}$: $\lambda_1 = 1, \lambda_2 = 16$, ортонормированные собственные вектора

$$\mathbf{x}_1 = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \mathbf{x}_2 = \frac{1}{\sqrt{5}}\begin{pmatrix} -1 \\ 2 \end{pmatrix},$$

$$\|\mathbf{A}\|_E = \sqrt{16} = 4$$

Особенно просто норма $\|\cdot\|_E$ вычисляется в случае $\mathbf{A} = \mathbf{A}^\mathsf{T}$

$$\|\mathbf{A}\|_E = \sqrt{\max \lambda(\mathbf{A}^\mathsf{T}\mathbf{A})} = \sqrt{\max \lambda(\mathbf{A}^2)} = \max |\lambda(\mathbf{A})|$$

Число обусловленности системы

Рассмотрим систему

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

и «близкую» к ней систему

$$\mathbf{A}(\mathbf{x} + \delta\mathbf{x}) = \mathbf{b} + \delta\mathbf{b}$$

Оценим относительную погрешность решения

$$\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \leqslant \frac{\|\mathbf{A}^{-1}\|\|\delta\mathbf{b}\|}{\|\mathbf{x}\|} = \frac{\|\mathbf{A}^{-1}\|\|\delta\mathbf{b}\|}{\|\mathbf{A}^{-1}\mathbf{b}\|} = \frac{\|\mathbf{A}^{-1}\|\|\mathbf{b}\|}{\|\mathbf{A}^{-1}\mathbf{b}\|}\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|}$$

Обозначим $\nu(\mathbf{A}, \mathbf{b}) = \frac{\|\mathbf{A}^{-1}\|\|\mathbf{b}\|}{\|\mathbf{A}^{-1}\mathbf{b}\|}$

$$\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \leqslant \nu(\mathbf{A}, \mathbf{b})\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|}$$

Назовем $\nu(\mathbf{A}, \mathbf{b})$ числом обусловленности системы $\mathbf{A}\mathbf{x} = \mathbf{b}$.

Свойства числа обусловленности

$$\frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|} \leqslant \nu(\mathbf{A}, \mathbf{b}) \frac{\|\delta \mathbf{b}\|}{\|\mathbf{b}\|}, \quad \nu(\mathbf{A}, \mathbf{b}) = \frac{\|\mathbf{A}^{-1}\| \|\mathbf{b}\|}{\|\mathbf{A}^{-1}\mathbf{b}\|}$$

Поскольку $\|\mathbf{A}^{-1}\mathbf{b}\| \leqslant \|\mathbf{A}^{-1}\| \|\mathbf{b}\|$

$$\nu(\mathbf{A}, \mathbf{b}) \geq 1,$$

причем существует такая правая часть $\mathbf{b}$, что $\nu(\mathbf{A}, \mathbf{b}) = 1$. Это именно то значение $\mathbf{b}$, при котором

$$\|\mathbf{A}^{-1}\mathbf{b}\| = \|\mathbf{A}^{-1}\| \|\mathbf{b}\|$$

При этом относительная погрешность решения не превосходит относительную погрешность правой части

Произвольная правая часть

Посмотрим, насколько большой может быть величина $\nu(\mathbf{A}, \mathbf{b})$ в зависимости от $\mathbf{b}$. С одной стороны,

$$\nu(\mathbf{A}, \mathbf{b}) = \frac{\|\mathbf{A}^{-1}\|\|\mathbf{b}\|}{\|\mathbf{A}^{-1}\mathbf{b}\|} \geq 1$$

Введем

$$\mu(\mathbf{A}) = \max_{\mathbf{b}} \nu(\mathbf{A}, \mathbf{b})$$

$$\mu(\mathbf{A}) = \|\mathbf{A}^{-1}\| \max_{\mathbf{b}} \frac{\|\mathbf{b}\|}{\|\mathbf{A}^{-1}\mathbf{b}\|} = \|\mathbf{A}^{-1}\| \max_{\mathbf{x}} \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|} = \|\mathbf{A}^{-1}\|\|\mathbf{A}\|$$

Таким образом, для произвольной правой части

$$\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \leqslant \mu(\mathbf{A}) \frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|}, \quad \mu(\mathbf{A}) = \|\mathbf{A}\|\|\mathbf{A}^{-1}\|$$

Число обусловленности матрицы

Число $\mu(\mathbf{A})$ называется обусловленностью матрицы $\mathbf{A}$ и показывает, насколько матрица «близка» к вырожденной. Поскольку

$$\mu(\mathbf{A}) \geq \nu(\mathbf{A}, \mathbf{b}) \geq 1,$$

относительная погрешность решения никогда не меньше относительной погрешности правой части. Более того, для любой матрицы $\mathbf{A}$ можно найти такие $\mathbf{b}$ и $\delta\mathbf{b}$, что относительная погрешность решения будет ровно в $\mu(\mathbf{A})$ раз больше относительной погрешности правой части.

Эта погрешность связана с самой задачей решения СЛАУ, а не с конкретным методом ее решения, и ни один численный метод не может решить эту задачу точнее. Поэтому данная погрешность будет *неустранимой*

Спасибо за внимание!

Цыбулин Иван
e-mail: tsybulin@crec.mipt.ru