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

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

Матрицы вырожденные и “почти” вырожденные

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 ($\det \mathbf{A} = 0$).

Матрицы вырожденные и “почти” вырожденные

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

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

Матрицы вырожденные и “почти” вырожденные

Взяв другую матрицу **A** получаем

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 5.0001 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \qquad \mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 5.0001 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} 1 \\ 2.0001 \end{pmatrix}$$

Решения этих систем довольно близки

$$\mathbf{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \mathbf{x} = \begin{pmatrix} 0.998002000 \\ 0.000999001 \end{pmatrix}$$

Но все же, отличие систем в четвертом знаке после запятой привело к различию решений уже в третьем.

Векторные нормы

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

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

$$\|\mathbf{x}\|_1 = \sum_i |x_i|$$

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

Векторные нормы

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

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

$$\|\mathbf{x}\|_1 = \sum_i |x_i|$$

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

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

$$\|\mathbf{x}\|_\infty = 4, \quad \|\mathbf{x}\|_1 = 7, \quad \|\mathbf{x}\|_E = 5$$

Матричные нормы

Основное нетривиальное свойство векторной нормы - это неравенство треугольника

$$\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$$

Можно по аналогии ввести понятие *матричной нормы*, со следующим свойством: если $\mathbf{y} = \mathbf{A}\mathbf{x}$, то

$$\|\mathbf{y}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{x}\| \qquad (1)$$

здесь $\|\mathbf{x}\|$, $\|\mathbf{y}\|$ — векторные нормы, а $\|\mathbf{A}\|$ — матричная норма.

Определение

Матричная норма $\|\mathbf{A}\|$ — это наименьшее число, при котором справедливо (1)

Матричные нормы

Можно дать и другое (эквивалентное) определение матричной нормы:

Определение

Матричная норма $\|\mathbf{A}\|$ — это наибольшее значение $\|\mathbf{y}\| = \|\mathbf{Ax}\|$, когда $\mathbf{x}$ пробегает все значения $\|\mathbf{x}\| = 1$

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

Заметим, что в определении матричной нормы участвует векторная норма $\|\cdot\|$. Разные векторные нормы будут порождать разные матричные нормы. Эти матричные нормы называются *индуцированными* или *присоединенными к соответствующей векторной норме*

Матричная норма $\|\cdot\|_\infty$

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

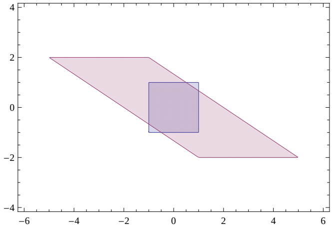
Точка $\mathbf{x} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ с нормой $\|\mathbf{x}\|_\infty = 1$ переходит в точку

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

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

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

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



В общем случае

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

Матричная норма $\|\cdot\|_1$

Точка $\mathbf{x} = \binom{0}{1}$ с нормой $\|\mathbf{x}\|_1 = 1$ переходит в точку

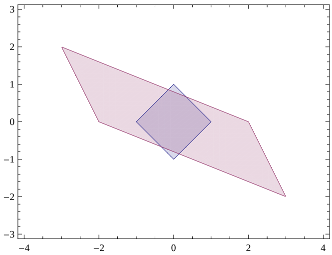
$$\mathbf{y} = \mathbf{A}\mathbf{x} = \begin{pmatrix} -3 \\ 2 \end{pmatrix}$$

с нормой $\|\mathbf{y}\|_1 = 5$.
Для остальных точек $\|\mathbf{x}\|_1 = 1$

$$\|\mathbf{A}\mathbf{x}\|_1 \le 5$$

$$\|\mathbf{A}\|_1 = 5$$

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



В общем случае

$$\|\mathbf{A}\|_1 = \max_j \sum_i |a_{ij}| = \|\mathbf{A}^T\|_\infty$$

Матричная норма $\|\cdot\|_E$

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

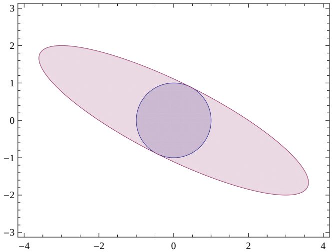
$$\mathbf{y} = \mathbf{A}\mathbf{x} = \frac{1}{\sqrt{5}}\begin{pmatrix}-8\\4\end{pmatrix}$$

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

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

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

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



В общем случае

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

Матричная норма $\|\cdot\|_E$

Докажем, что $\|\mathbf{A}\|_E = \sqrt{\max \lambda(\mathbf{A}^T\mathbf{A})}$.

$$\|\mathbf{Ax}\|_E = \sqrt{(\mathbf{Ax}, \mathbf{Ax})} = \sqrt{\mathbf{x}^T\mathbf{A}^T\mathbf{Ax}} = \sqrt{(\mathbf{x}, \mathbf{A}^T\mathbf{Ax})}$$

Матричная норма $\|\cdot\|_E$

Докажем, что $\|\mathbf{A}\|_E = \sqrt{\max \lambda(\mathbf{A}^T\mathbf{A})}$.

$$\|\mathbf{Ax}\|_E = \sqrt{(\mathbf{Ax}, \mathbf{Ax})} = \sqrt{\mathbf{x}^T\mathbf{A}^T\mathbf{Ax}} = \sqrt{(\mathbf{x}, \mathbf{A}^T\mathbf{Ax})}$$

Матрица $\mathbf{A}^T\mathbf{A}$ симметрична и положительно определена. Значит у нее существует ортонормированный базис из собственных векторов, причем все собственные числа неотрицательны

$$\mathbf{A}^T\mathbf{Ax}_k = \lambda_k\mathbf{x}_k, \quad \lambda_k \geq 0$$

Матричная норма $\|\cdot\|_E$

$$\mathbf{A}^T\mathbf{A}\mathbf{x}_k = \lambda_k \mathbf{x}_k, \quad \lambda_k \geq 0$$

Разложим $\mathbf{x}$ по этому базису $\{\mathbf{x}_k\}$

$$\mathbf{x} = \sum_k \alpha_k \mathbf{x}_k$$

$$\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{A}\sum_k \alpha_k \mathbf{x}_k = \sum_k \alpha_k \lambda_k \mathbf{x}_k$$

Учитывая, что базис ортонормированный $(\mathbf{x}_k, \mathbf{x}_m) = \delta_{km}$

$$\left(\mathbf{x}, \mathbf{A}^T\mathbf{A}\mathbf{x}\right) = \sum_k \lambda_k \alpha_k^2 \leq \max_k \lambda_k \sum_k \alpha_k^2 = (\mathbf{x}, \mathbf{x}) \max_k \lambda_k$$

Возвращаясь к норме $\mathbf{A}\mathbf{x}$

$$\|\mathbf{A}\mathbf{x}\|_E = \sqrt{(\mathbf{x}, \mathbf{A}^T\mathbf{A}\mathbf{x})} \leq \sqrt{\max \lambda(\mathbf{A}^T\mathbf{A})}\sqrt{(\mathbf{x}, \mathbf{x})} = \sqrt{\max \lambda(\mathbf{A}^T\mathbf{A})}$$

Матричная норма $\|\cdot\|_E$

$$\|\mathbf{A}\mathbf{x}\|_E = \sqrt{(\mathbf{x}, \mathbf{A}^T\mathbf{A}\mathbf{x})} \leq \sqrt{\max\lambda(\mathbf{A}^T\mathbf{A})}\sqrt{(\mathbf{x}, \mathbf{x})} = \sqrt{\max\lambda(\mathbf{A}^T\mathbf{A})}$$

Покажем, что норма $\|\mathbf{A}\|_E$ не может быть меньше $\sqrt{\max\lambda(\mathbf{A}^T\mathbf{A})}$. Действительно, норма должна удовлетворять

$$\|\mathbf{A}\|_E = \|\mathbf{A}\|_E\|\mathbf{x}_k\|_E \geq \|\mathbf{A}\mathbf{x}_k\|_E = \sqrt{(\mathbf{x}_k, \mathbf{A}^T\mathbf{A}\mathbf{x}_k)} = \sqrt{\lambda_k}$$

Таким образом,

$$\sqrt{\lambda_k} \leq \|\mathbf{A}\|_E \leq \sqrt{\max_k \lambda_k}$$

Последнее возможно только при

$$\|\mathbf{A}\|_E = \sqrt{\max_k \lambda_k}$$

Матричная норма $\|\cdot\|_E$

Еще раз вернемся к матрице

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

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

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

Фиксированная правая часть

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

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

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

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

Фиксированная правая часть

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

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

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

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, какой 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

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

Число обусловленности

Число $\mu(\mathbf{A})$ называется обусловленностью матрицы **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 **A** можно найти такие **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 называется *неустранимой*

Спасибо за внимание!

tsybulin@crec.mipt.ru