1. Возмущения и обратимость матрицы

1. Пусть $A \in M_n$ — обратимая матрица, т. е. $|A| \neq 0$. Пусть, далее, $B \in M_n$. Возникает вопрос, при каких условиях на B матрица $A+B$ будет также обратимой? Поскольку $A+B = A(I + A^{-1}B)$, то для существования матрицы, обратной к $A+B$, очевидно, необходимо и достаточно, чтобы спектр матрицы $A^{-1}B$ не содержал -1[1)]. Отсюда вытекают следующие практически важные достаточные условия обратимости матрицы $A+B$ ($\|A\|$) :

1) матрица $A+B$ обратима, если $\rho(A^{-1}B) < 1$;
2) матрица $A+B$ обратима, если $\|A^{-1}B\| < 1$;
3) матрица $A+B$ обратима, если $\|A^{-1}\|\|B\| < 1$,

Здесь
$$\rho(B) = \max_{1\leqslant k\leqslant n} |\lambda_k(B)|$$
есть *спектральный радиус* матрицы B. Третье условие часто записывают так:
$$\mathrm{cond}(A)(\|B\|/\|A\|) < 1, \tag{1.1}$$
где $\mathrm{cond}(A) = \|A^{-1}\|\|A\|$. Это число называют числом обусловленности матрицы A. Ясно, что $\mathrm{cond}(A) \geqslant 1$, так как $1 = \|I\| = \|A^{-1}A\| \leqslant \|A^{-1}\|\,\|A\| = \mathrm{cond}(A)$.

Условие (1.1) можно интерпретировать следующим образом: матрица $A+B$ обратима, если относительное возмущение матрицы A, т. е. $\|B\|/\|A\|$, мало по сравнению с ее числом обусловленности.

2. Пример. Пусть $A = \{a_{ij}\}_{i,j=1}^n$ — квадратная матрица с диагональным преобладанием (по строкам), т.е.
$$|a_{ii}| > \sum_{1\leqslant j\leqslant n,\ j\neq i} |a_{ij}| \quad \forall\, i = 1, 2, \ldots, n. \tag{1.2}$$
Покажем, что она невырождена (в прошлом семестре было дано доказательство того, что все ее главные миноры, а, следовательно, и ее определитель, отличны от нуля). Пусть $D = \mathrm{diag}(a_{11}, a_{22}, \ldots, a_{nn})$. Вследствие условия (1.2) матрица D невырождена. Запишем матрицу A в виде $A = D + (A - D)$. Вновь используя условие (1.2), получим, что $\|D^{-1}(A-D)\|_\infty < 1$, значит выполнено условие 2, и матрица A невырождена.

Поскольку определители матриц A и A^T совпадают, то матрица с диагональным преобладанием по столбцам также невырождена.

[1)] под спектром матрицы понимается множество ее собственных чисел

Теорема 1. *Пусть матрицы A и $\tilde{A} = A + B$ обратимы. Тогда*

$$\frac{\|A^{-1} - \tilde{A}^{-1}\|}{\|\tilde{A}^{-1}\|} \leqslant \|A^{-1}B\|. \tag{1.3}$$

Если $\|A^{-1}B\| < 1$, *то*

$$\|\tilde{A}^{-1}\| \leqslant \frac{\|A^{-1}\|}{1 - \|A^{-1}B\|}, \tag{1.4}$$

$$\frac{\|A^{-1} - \tilde{A}^{-1}\|}{\|A^{-1}\|} \leqslant \frac{\|A^{-1}B\|}{1 - \|A^{-1}B\|}. \tag{1.5}$$

Доказательство. По условию теоремы $I = (A + B)\tilde{A}^{-1}$, следовательно, $A^{-1} = (I + A^{-1}B)\tilde{A}^{-1}$, поэтому $A^{-1} - \tilde{A}^{-1} = A^{-1}B\tilde{A}^{-1}$. Отсюда, очевидно, следует (1.6). Далее, $\tilde{A}^{-1} = A^{-1} - A^{-1}B\tilde{A}^{-1}$, значит, $\|\tilde{A}^{-1}\| \leqslant \|A^{-1}\| + \|A^{-1}B\|\|\tilde{A}^{-1}\|$, откуда вытекает (1.4). Наконец, (1.5) — очевидное следствие (1.6), (1.4). $\square$

Из теоремы 1 непосредственно вытекает

Следствие 1. *Пусть матрицы A и $\tilde{A} = A+B$ обратимы. Тогда*

$$\frac{\|A^{-1} - \tilde{A}^{-1}\|}{\|\tilde{A}^{-1}\|} \leqslant \operatorname{cond}(A)(\|B\|/\|A\|). \tag{1.6}$$

Если $\operatorname{cond}(A)(\|B\|/\|A\|) < 1$, *то*

$$\|\tilde{A}^{-1}\| \leqslant \frac{\|A^{-1}\|}{1 - \operatorname{cond}(A)(\|B\|/\|A\|)}, \tag{1.7}$$

$$\frac{\|A^{-1} - \tilde{A}^{-1}\|}{\|A^{-1}\|} \leqslant \frac{\operatorname{cond}(A)(\|B\|/\|A\|)}{1 - \operatorname{cond}(A)(\|B\|/\|A\|)|}. \tag{1.8}$$

2. Устойчивость систем линейных уравнений

1. В этом параграфе норма матрицы считается подчиненной норме вектора. Следующая теорема устанавливает связь относительного возмущениям матрицы системы и ее правой части с относительным возмущением решения. Главную роль в получаемых здесь оценках играет число обусловленности матрицы системы уравнений.

Теорема 1. *Пусть матрица A обратима, матрица ΔA такова, что $\|A^{-1}\Delta A\| < 1$, вектор x — решение системы уравнений*

$$Ax = b, \tag{2.1}$$

вектор $\tilde{x}$ — решение системы уравнений

$$\tilde{A}\tilde{x} = \tilde{b}, \quad \tilde{b} = b + \Delta b, \quad \tilde{A} = A + \Delta A. \tag{2.2}$$

Тогда

$$\frac{\|x - \tilde{x}\|}{\|x\|} \leqslant \frac{\operatorname{cond}(A)}{1 - \|A^{-1}\Delta A\|} \left(\frac{\|\Delta b\|}{\|b\|} + \frac{\|\Delta A\|}{\|A\|} \right). \tag{2.3}$$

Если дополнительно потребовать, чтобы выполнялось условие

$$\|A^{-1}\|\|\Delta A\| < 1,$$

то

$$\frac{\|x - \tilde{x}\|}{\|x\|} \leqslant \frac{\operatorname{cond}(A)}{1 - \operatorname{cond}(A)(\|\Delta A\|/\|A\|)} \left(\frac{\|\Delta b\|}{\|b\|} + \frac{\|\Delta A\|}{\|A\|} \right). \tag{2.4}$$

Доказательство. По условию теоремы матрицы A^{-1} и $\tilde{A}^{-1}$ существуют, поэтому $x = A^{-1}b$, $\tilde{x} = \tilde{A}^{-1}(b + \Delta b)$, следовательно, $\tilde{x} - x = \tilde{A}^{-1}\Delta b + (\tilde{A}^{-1} - A^{-1})b$, и

$$\|x - \tilde{x}\| \leqslant \|\tilde{A}^{-1}\|\|\Delta b\| + \|\tilde{A}^{-1} - A^{-1}\|\|b\|,$$

откуда, используя (1.4), (1.5) и неравенство $\|b\| \leqslant \|A\|\|x\|$, после элементарных преобразований получим (2.3). Оценка (2.4) есть очевидное следствие (2.3). □

Оценка возмущения упрощается, когда $\Delta A = 0$.

Следствие 1. *Пусть $Ax = b$, $\quad A\tilde{x} = \tilde{b}$, $\quad \tilde{b} = b + \Delta b$. Тогда*

$$\frac{\|x - \tilde{x}\|}{\|x\|} \leqslant \operatorname{cond}(A) \frac{\|\Delta b\|}{\|b\|}. \tag{2.5}$$

Левая часть в (2.3) характеризует относительную величину возмущения решения, отношение $\|\Delta b\|/\|b\|$ — относительную величину возмущения правой части СЛАУ, отношение $\|\Delta A\|/\|A\|$ — относительную величину возмущения матрицы A. Если считать, что $\|A^{-1}\Delta A\| \ll 1$, то из (2.3) следует, что оценка относительной величины возмущения решения определяется числом обусловленности матрицы и относительной величиной возмущения данных СЛАУ.

В случае $\mathrm{cond}(A) \gg 1$ систему уравнений, а также матрицу A называют плохо обусловленной. В этом случае, как это следует из оценки (2.3), погрешность решения системы уравнений может оказаться неприемлемо большой. Понятие приемлемости или неприемлемости погрешности определяется постановкой задачи.

2. Пример плохо обусловленной системы. Рассмотрим СЛАУ, в которой

$$A = \begin{pmatrix} 1 & -1 & -1 & -1 & \vdots & -1 & -1 \\ 0 & 1 & -1 & -1 & \vdots & -1 & -1 \\ 0 & 0 & 1 & -1 & \vdots & -1 & -1 \\ 0 & 0 & 0 & 1 & \vdots & -1 & -1 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \\ \dots \\ -1 \\ 1 \end{pmatrix}.$$

Легко вычислить, что

$$A = \begin{pmatrix} 1 & 2^0 & 2^1 & 2^2 & \vdots & 2^{n-3} & 2^{n-2} \\ 0 & 1 & 2^0 & 2^1 & \vdots & 2^{n-4} & 2^{n-3} \\ 0 & 0 & 1 & 2^0 & \vdots & 2^{n-5} & 2^{n-4} \\ 0 & 0 & 0 & 1 & \vdots & 2^{n-6} & 2^{n-5} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 1 & 2^0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \Gamma$$

Следовательно, $\mathrm{cond}_\infty(A) = n\,2^{n-1}$, т.к.

$$\|A\|_\infty = n, \quad \|A^{n-1}\|_\infty = 1 + 2^0 + 2^2 1 + \ldots + 2^{n-2} = 2^{n-1}.$$

При больших n матрица плохо обусловлена. Например, при $n = 50$ имеем $\mathrm{cond}_\infty(A) \approx 2.8 \cdot 10^{16}$. Насколько это число велико?

Система $Ax = b$ имеет единственное решение $x = (0, 0, \ldots, 0, 1)^T$, $\|x\|_\infty = 1$. Пусть правая часть системы содержит погрешность $\Delta b = (0, 0, \ldots, 0, \varepsilon)$, $\varepsilon > 0$, $\Delta A = 0$. Если $\varepsilon = 2.2 \cdot 10^{-16}$, что соответствует относительной точности представления чисел в ЭВМ в плавающей арифметике типа double, то оценка (2.5) принимает вид

$$\frac{\|x - \tilde{x}\|_\infty}{\|x\|_\infty} \leqslant 6.2, \tag{2.6}$$

из которой не следует малость погрешности решения. Проверим, насколько эта оценка точна.

Система $A\tilde{x} = \tilde{b}$ легко решается и, если $\tilde{x} = x + \Delta x$, то

$$\Delta x_n = \varepsilon, \quad \Delta x_{n-1} = \varepsilon, \quad , \Delta x_{n-2} = 2\varepsilon, \quad \ldots, \Delta x_1 = 2^{n-2}\varepsilon,$$

и левая часть оценки (2.5) равна

$$\frac{\|x - \tilde{x}\|_\infty}{\|x\|_\infty} = 2^{n-2}\varepsilon \approx 0.062.$$

Таким образом, при числе обусловленности матрицы порядка 10^{16} мы получили решение с относительной погрешностью $6 \cdot 10^{-2}$. Получить решение с большей точностью невозможно.

3. Пусть некоторым способом найден вектор $\tilde{x}$, который мы считаем приближением к решению уравнения (2.1). Наша цель — оценить погрешность $\|x - \tilde{x}\|$ через норму невязки $\|A\tilde{x} - b\|$. Введем используемую в дальнейшем вспомогательную величину. Пусть матрица A обратима, $x \neq 0$, $Ax = b$. Положим $\eta = \|A\|\|x\|/\|b\|$. Очевидно, что $\eta \geqslant 1$, и поскольку $\|x\| \leqslant \|A^{-1}\|\|b\|$, то $\eta \leqslant \|A\|\|A^{-1}\| = \mathrm{cond}(A)$. Для $\tilde{x} \in \mathbb{C}^n$ положим $r = A\tilde{x} - b$. Тогда $x - \tilde{x} = A^{-1}r$,

$$\|x - \tilde{x}\| \leqslant \|A^{-1}\|\|r\|.$$

Поэтому

$$\|x - \tilde{x}\|/\|x\| \leqslant (\mathrm{cond}(A)/\eta)\|r\|/\|b\|, \tag{2.7}$$

и как следствие

$$\|x - \tilde{x}\|/\|x\| \leqslant \mathrm{cond}(A)\|r\|/\|b\|. \tag{2.8}$$

Оценка (2.7) показывает, что чем ближе величина η к величине $\mathrm{cond}(A)$, тем лучше относительная погрешность оценивается относительной невязкой приближенного решения.

4. О точности оценки (2.8). Рассмотрим СЛАУ, в которой

$$A = \begin{pmatrix} 1 & -1 \\ 1 & \varepsilon - 1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Ее точное решение $x = (1, 0)^T$, и пусть $\tilde{x} = (1 + \varepsilon^{-1/2}, \varepsilon^{-1/2})^T$ есть приближение к этому решению. Тогда $r = (0, \varepsilon^{1/2})^T$, вектор ошибки $x - \tilde{x} = -(\varepsilon^{-1/2}, \varepsilon^{-1/2})^T$. Поэтому

a) $\|r\|/\|b\| = O(\varepsilon^{1/2}) \to 0$ при $\varepsilon \to 0$;

b) относительная ошибка $\|x - \tilde{x}\|/\|x\| = O(\varepsilon^{-1/2}) \to \infty$ при $\varepsilon \to 0$.

Как видим, малость невязки не гарантирует точность приближения к решению. Это объясняется тем, что $\mathrm{cond}(A) = O(\varepsilon^{-1}) \to \infty$ при $\varepsilon \to 0$. Обе части оценки (2.8) имеют одинаковый порядок $O(\varepsilon^{-1/2})$, что характеризует точность этой оценки.

5. Свойства числа обусловленности Установите самостоятельно следующие свойства числа обусловленности:

1) $\operatorname{cond}(A) \geqslant 1$;
2) $\operatorname{cond}(A) = \operatorname{cond}(A^{-1})$;
3) $\operatorname{cond}(AB) \leqslant \operatorname{cond}(A)\operatorname{cond}(B)$;
4) Если $A = A^*$, $\lambda_{\max}(A)$ ($\lambda_{\min}(A)$) — максимальное (минимальное) по модулю собственное значение A, то $\operatorname{cond}(A) = \left|\dfrac{\lambda_{\max}(A)}{\lambda_{\min}(A)}\right|$.
5) Пусть $\operatorname{cond}(A)$ есть число обусловленности относительно любой матричной нормы. Тогда $\operatorname{cond}(A) \geqslant \left|\dfrac{\lambda_{\max}(A)}{\lambda_{\min}(A)}\right|$ (воспользуйтесь оценкой $\rho(A) \leqslant \|A\|$).