

Глава 3

Решение задач линейной алгебры

§1 Нормы векторов и матриц

К задачам линейной алгебры, которые будут рассматриваться в этой главе, относятся решение систем линейных алгебраических уравнений, обращение матриц, нахождение собственных чисел и собственных векторов матриц. Введем некоторые обозначения. \mathbb{C}^n — множество (линейное пространство) n -мерных векторов с комплексными компонентами. Для векторов $x, y \in \mathbb{C}^n$ обычным обозначением их компонент будет $x = (\xi_1, \xi_2, \dots, \xi_n)$, $y = (\eta_1, \eta_2, \dots, \eta_n)$. Обозначением скалярного произведения векторов служит $(x, y) = \sum_{k=1}^n \xi_k \bar{\eta}_k$ (\bar{z} — число, комплексно сопряженное числу z). Единичную матрицу будем обозначать через E и введем еще одно стандартное обозначение: $e_k = (0, \dots, 0, 1, 0, \dots, 0)$, где единица — это k -ая компонента вектора e_k .

Для измерения и оценки погрешностей, возникающих при решении задач линейной алгебры, используется понятие нормы.

Определение. Векторной *нормой* называется заданная на \mathbb{C}^n вещественная функция $\nu(x)$, обозначаемая обычно $\nu(x) = \|x\|$ и обладающая свойствами:

- 1) для всех векторов $x \in \mathbb{C}^n$ $\|x\| \geq 0$ и $\|x\| = 0$ тогда и только тогда, когда $x = 0$ (нулевой вектор),
- 2) для любых векторов $x \in \mathbb{C}^n$ и чисел $\lambda \in \mathbb{C}$ $\|\lambda x\| = |\lambda| \|x\|$,
- 3) для любых $x, y \in \mathbb{C}^n$ $\|x + y\| \leq \|x\| + \|y\|$ (неравенство треугольника).

Сразу же отметим очевидное следствие 3):

- 4) Для любых векторов $x, y \in \mathbb{C}^n$ $|\|x\| - \|y\|| \leq \|x - y\|$. Действительно, считая для определенности $\|x\| \geq \|y\|$, имеем $\|x\| \leq \|x - y\| + \|y\|$.

Наиболее употребительными являются следующие нормы, имеющие специальные обозначения*:

$$\|x\|_2 = \sqrt{\sum_{k=1}^n |\xi_k|^2} = \sqrt{(x, x)} \quad \text{— евклидова норма, длина вектора,}$$

$$\|x\|_\infty = \max |\xi_k|,$$

$$\|x\|_1 = \sum_{k=1}^n |\xi_k|.$$

Проверка всех свойств 1)-3) для двух последних из этих норм очевидна. Неравенство треугольника для первой — это известное неравенство Минковского.

В силу неравенства Коши - Буняковского $|(x, y)| \leq \|x\|_2 \|y\|_2$. Кроме того, очевидно неравенство $|(x, y)| \leq \|x\|_1 \|y\|_\infty$.

Приведем еще один важный пример нормы.

*Эти обозначения связаны с тем, что перечисляемые ниже нормы являются частными случаями *норм Гельдера*: $\|x\|_p = (\sum_{k=1}^n |\xi_k|^p)^{1/p}$ ($p \geq 1$).

Пример. Пусть φ_k ($k = 1, \dots, n$) комплекснозначные непрерывные на $[a, b]$ линейно независимые функции. Тогда

$$\|x\| = \left\| \sum_{k=1}^n \xi_k \varphi_k \right\|_C = \max_{t \in [a, b]} \left| \sum_{k=1}^n \xi_k \varphi_k(t) \right|$$

есть норма в \mathbb{C}^n . Свойства 1)-3) для этой нормы легко проверяются.

Теорема 1 (об эквивалентности норм). Все нормы в \mathbb{C}^n эквивалентны. Это значит, что для любых двух норм $\|\cdot\|'$ и $\|\cdot\|''$ найдутся такие постоянные c' и c'' , что для всех векторов $x \in \mathbb{C}^n$ выполняются неравенства

$$\|x\|' \leq c' \|x\|'', \quad \|x\|'' \leq c'' \|x\|'.$$

Доказательство. Поскольку понятие эквивалентности норм, определенное в формулировке теоремы, транзитивно (т.е. если две нормы эквивалентны третьей, то они эквивалентны и между собой), то можно считать, что одна из норм есть $\|\cdot\|_2$, а другая $\|\cdot\|$ произвольна. Тогда

$$\|x\| = \left\| \sum_{k=1}^n \xi_k e_k \right\| \leq \sum_{k=1}^n |\xi_k| \|e_k\| \leq c' \|x\|_2, \quad \text{где } c' = \sqrt{\sum_{k=1}^n \|e_k\|^2}.$$

В частности, $|\|x\| - \|y\|| \leq \|x - y\| \leq c' \|x - y\|_2$, так что $\|\cdot\|$ — непрерывная функция. На сфере $S = \{x \mid \|x\|_2 = 1\}$ она достигает своего минимального значения, отличного от нуля, т.к. на S в нуль не обращается, так что $\min_{x \in S} \|x\| = m > 0$. Для любого $x \neq 0$ положим $y = x/\|x\|_2$; тогда $\|x\| = \|x\|_2 \|y\| \geq m \|x\|_2$, так что для любого x $\|x\|_2 \leq \frac{1}{m} \|x\|$. ■

Если использовать приведенный выше пример нормы, то из теоремы 1 немедленно получим

Следствие 1. При заданных n и $[a, b]$ найдется такая постоянная $m > 0$, что для любого полинома $P_n \in \mathbb{P}_n$ ($P_n(t) = a_0 + \dots + a_n t^n$) выполняется неравенство $\|P_n\|_{C[a, b]} \geq m \sqrt{\sum_{k=1}^n a_k^2}$.

Это следствие формулировалось (но не доказывалось) в §1 главы 1 в виде леммы.

Пусть даны последовательность векторов $x_s = (\xi_1^s, \dots, \xi_n^s)$ и вектор x .

Теорема 2. Эквивалентны утверждения:

А. При всех k при $s \rightarrow \infty$ $\xi_k^s \rightarrow \xi_k$,

Б. $\|x_s - x\| \rightarrow 0$,

В. для всех $y \in \mathbb{C}^n$ $(x_s, y) \rightarrow (x, y)$.

Доказательство. Если утверждение Б выполняется для какой-то одной нормы, то в силу теоремы 1 оно выполняется и для всех остальных. Доказательство теоремы складывается из трех пунктов.

1) А \Rightarrow Б. Это очевидно, если, например, рассматриваемая норма евклидова.

2) Б \Rightarrow В. Это вытекает из неравенства $|(x_s, y) - (x, y)| \leq \|x_s - x\|_2 \|y\|_2$.

3) В \Rightarrow А. Положим $y = e_k$ ($k = 1, 2, \dots, n$). Тогда $\xi_k^s = (x_s, y)$, $\xi_k = (x, y)$ и $\xi_k^s \rightarrow \xi_k$. ■

Если выполнено хоть одно из требований А-В (а значит, и два другие), то говорят, что последовательность векторов x_s сходится к вектору x ($x_s \rightarrow x$). Сходимость в смысле А называется покомпонентной, в смысле В — сходимостью по норме, в смысле В — слабой сходимостью векторов. Таким образом в пространстве \mathbb{C}^n все эти три вида сходимости означают одно и то же.

Перейдем теперь к рассмотрению матриц. Множество квадратных матриц порядка n с комплексными элементами будем обозначать через \mathbb{M}_n , а элементы матрицы A — через a_{kj} :

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

Единичную матрицу, как уже говорилось, будем обозначать буквой E . Заданная на \mathbb{M}_n вещественная функция $\|A\|$ называется *нормой*, если она удовлетворяет требованиям:

1) для всех $A \in \mathbb{M}_n$ $\|A\| \geq 0$ и $\|A\| = 0$ в том и только том случае, если $A = 0$, т.е. равны нулю все элементы этой матрицы,

2) для всех $\lambda \in \mathbb{C}$ и $A \in \mathbb{M}_n$ $\|\lambda A\| = |\lambda| \|A\|$,

3) для всех $A, B \in \mathbb{M}_n$ $\|A + B\| \leq \|A\| + \|B\|$.

Как и для векторов, из 3) следует

4) $|\|A\| - \|B\|| \leq \|A - B\|$.

Каждой матрице $A \in \mathbb{M}_n$ можно сопоставить некоторый вектор порядка n^2 по правилу $x_A = (a_{11}, \dots, a_{1n}, a_{21}, \dots, a_{nn})$. Это соотношение между \mathbb{M}_n и \mathbb{C}^{n^2} линейно и взаимно-однозначно, причем $\|x_A\| = \|A\|$ есть векторная норма в \mathbb{C}^{n^2} . Таким образом, норму в \mathbb{M}_n можно рассматривать как норму в \mathbb{C}^{n^2} . Поэтому в силу теоремы 1 все нормы в \mathbb{M}_n эквивалентны.

Рассмотрим последовательность матриц

$$A_s = \begin{pmatrix} a_{11}^s & \dots & a_{1n}^s \\ \dots & \dots & \dots \\ a_{n1}^s & \dots & a_{nn}^s \end{pmatrix}$$

и матрицу A .

Теорема 3. Эквивалентны утверждения:

А. при всех k, j $a_{kj}^s \rightarrow a_{kj}$,

Б. $\|A_s - A\| \rightarrow 0$,

В. для всех $x \in \mathbb{C}^n$ $A_s x \rightarrow Ax$,

Г. для всех $x, y \in \mathbb{C}^n$ $(A_s x, y) \rightarrow (Ax, y)$.

Доказательство. Как и в теореме 2, норма в утверждении Б произвольна, и доказательство теоремы сводится к доказательству четырех соотношений:

$A \Leftrightarrow Б$ — это следует из теоремы 2 и указанного выше соответствия между нормами в \mathbb{M}_n и \mathbb{C}^{n^2} .

$A \Rightarrow В$. Для $x_s = A_s x$ и $y = Ax$ имеем $\xi_k^s = \sum a_{kj}^s \xi_j \rightarrow \sum a_{kj} \xi_j = \eta_k$.

$В \Rightarrow Г$. Это соотношение очевидно ввиду пункта В из теоремы 2.

$Г \Rightarrow А$. Достаточно рассмотреть случай векторов $x = e_j$, $y = e_k$, для которого $(A_s x, y) = a_{kj}^s$, $(Ax, y) = a_{kj}$ ■

Определение сходимости последовательности матриц дается так же, как для векторов.

Введем теперь класс наиболее употребительных матричных норм.

Определение. Пусть $\|\cdot\|$ некоторая норма в \mathbb{C}^n . Матричная норма

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

называется *операторной* нормой, порожденной соответствующей векторной.

То, что определяемая формулой (1) функция действительно есть норма, проверяется элементарно. Если $\|\cdot\|$ — операторная норма, то помимо 1)-4) она обладает еще легко проверяемыми свойствами.

5) Для всех векторов x $\|Ax\| \leq \|A\| \|x\|$ (векторная норма $\|x\|$ здесь именно та, которая порождает операторную матричную) и существует такой вектор x_0 , что $\|x_0\| = 1$, $\|Ax_0\| = \|A\|$. Последнее утверждение следует из ограниченности и замкнутости (компактности) множества $S = \{x \in \mathbb{C}^n \mid \|x\| = 1\}$ и непрерывности нормы.

6) Для единичной матрицы $\|E\| = 1$.

7) Операторная норма произведения матриц не превосходит произведения их норм: $\|AB\| \leq \|A\| \|B\|$.

8) Любое собственное число λ матрицы A не превосходит по модулю любой ее операторной нормы. Действительно, для соответствующего собственного вектора x имеем $|\lambda| \|x\| = \|Ax\| \leq \|A\| \|x\|$.

Операторные нормы матрицы, порожденные введенными выше векторными $\|\cdot\|_2$, $\|\cdot\|_\infty$ и $\|\cdot\|_1$, помечаются теми же значками.

Теорема 4. Справедливы равенства:

$$\|A\|_\infty = \max_k \sum_j |a_{kj}|, \quad (2)$$

$$\|A\|_1 = \max_j \sum_k |a_{kj}|, \quad (3)$$

$$\|A\|_2 = \sqrt{\Lambda}, \quad (4)$$

где Λ — максимальное собственное число матрицы* $A^* A$.

Доказательство. 1) Обозначим правую часть (2) через \varkappa и для произвольно выбранного вектора x положим $y = Ax$. Тогда

$$|\eta_k| = \left| \sum_j a_{kj} \xi_j \right| \leq \varkappa \|x\|_\infty,$$

Корни из собственных чисел матрицы $A^ A$ называются *сингулярными числами* матрицы A , так что $\|A\|_2$ есть максимальное сингулярное число матрицы A .

так что $\|Ax\|_\infty = \|y\|_\infty \leq \varkappa\|x\|_\infty$ и $\|A\|_\infty \leq \varkappa$. Пусть максимум в правой части равенства (2) достигается при $k = k_0$: $\varkappa = \sum_j |a_{k_0j}|$. Положим* $\xi_j = \overline{\text{sign } a_{k_0j}}$, $y = Ax$. Тогда $\|x\|_\infty = 1$, $\eta_{k_0} = \varkappa$, $\|Ax\|_\infty = \|y\|_\infty \geq \varkappa$ и потому $\|A\|_\infty \geq \varkappa$.

2) Пусть \varkappa — правая часть (3), $Ax = y$. Тогда

$$\|y\|_1 = \sum_k \left| \sum_j a_{kj} \xi_j \right| \leq \sum_j |\xi_j| \sum_k |a_{kj}| \leq \varkappa \|x\|_1$$

и потому $\|A\|_1 \leq \varkappa$. Пусть $\varkappa = \sum_k |a_{kj_0}|$. Тогда для вектора $x = e_{j_0}$ имеем $\|x\|_1 = 1$, $\|Ax\|_1 = \varkappa$, и $\|A\|_1 \geq \varkappa$.

3) Матрица A^*A самосопряженная (эрмитова). Все ее собственные числа вещественны, неотрицательны, и она имеет полную ортогональную систему собственных векторов. Пусть $\Lambda_1 \geq \Lambda_2 \geq \dots \geq \Lambda_n$ все ее собственные числа ($\Lambda = \Lambda_1$) и z_1, \dots, z_n — соответствующие собственные векторы, такие что $(z_k, z_j) = \delta_{kj}$. Произвольный вектор x разложим по этим векторам: $x = \sum \alpha_k z_k$. Тогда $\|x\|_2^2 = (x, x) = \sum |\alpha_k|^2$ и

$$\|Ax\|_2^2 = (Ax, Ax) = (x, A^*Ax) = \sum \Lambda_k |\alpha_k|^2 \leq \Lambda \|x\|_2^2.$$

Так что $\|A\|_2 \leq \sqrt{\Lambda}$. В то же время $\|Az_1\|_2^2 = (Az_1, Az_1) = \Lambda(z_1, z_1) = \Lambda\|z_1\|_2^2$ и $\|A\|_2 \geq \sqrt{\Lambda}$. ■

Следствие 2. Если матрица A самосопряженная, то $\|A\|_2 = \max |\lambda_k|$, где λ_k — собственные числа самой матрицы A .

Доказательство. В рассматриваемом случае $A^*A = A^2$, и собственные числа этой матрицы суть квадраты собственных чисел матрицы A . ■

Вычисление нормы $\|A\|_2$ обычно требует значительной затраты труда, и поэтому представляют интерес ее оценки:

Теорема 5. Для $\|A\|_2$ справедливы оценки:

$$\|A\|_2 \leq \sqrt{\sum_k \sum_j |a_{kj}|^2}, \quad \|A\|_2 \leq \sqrt{\|A\|_\infty \|A\|_1}.$$

Доказательство. 1) Для $y = Ax$ имеем

$$\|y\|_2^2 = \sum_k \left| \sum_j a_{kj} \xi_j \right|^2 \leq \sum_k \left(\sum_j |a_{kj}|^2 \right) \left(\sum_j |\xi_j|^2 \right) = \|x\|_2^2 \sum_k \sum_j |a_{kj}|^2.$$

$$2) \|A\|_2^2 = \Lambda \leq \|A^*A\|_1 \leq \|A\|_1 \|A^*\|_1 = \|A\|_1 \|A\|_\infty. \quad \blacksquare$$

Замечание. Правая часть первой из доказанных оценок есть некоторая матричная норма. Она называется нормой Фробениуса.

*Для комплексного числа α

$$\text{sign } \alpha = \begin{cases} 0 & \text{при } \alpha = 0 \\ \alpha/|\alpha| & \text{при } \alpha \neq 0. \end{cases}$$

Тем самым $\overline{\alpha \text{sign } \alpha} = |\alpha|$.

[* З а м е ч а н и е. Многие из изложенных выше результатов — это частные случаи известных в функциональном анализе результатов. В частности, операторная норма матрицы — это норма матрицы как линейного оператора в пространстве \mathbb{C}^n , в котором задана некоторая норма.*]

Задача 1. Доказать, что для норм Гельдера при $q \rightarrow \infty$ выполняется соотношение

$$\|x\|_q = \left(\sum_{k=1}^n |\xi_k|^q \right)^{1/q} \rightarrow \|x\|_\infty.$$

Задача 2. Пусть B неособенная матрица (т.е. $\det B \neq 0$) и $\|\cdot\|$ — векторная норма. Положим $\|x\|' = \|Bx\|$. Показать, что $\|\cdot\|'$ также норма, и найти выражение для порожденной этой нормой операторной матричной.

Задача 3. Для p и q , принимающих независимо друг от друга значения 1, 2, ∞ , найти

$$\sup_{x \in \mathbb{C}^n} \frac{\|x\|_p}{\|x\|_q}, \quad \sup_{A \in \mathbb{M}_n} \frac{\|A\|_p}{\|A\|_q}.$$

Обратить внимание на характер зависимости этих величин от n .

Задача 4. Показать, что норма Фробениуса не является операторной.

§2 Матричная геометрическая прогрессия и некоторые оценки

В этом параграфе, если не оговорено противное, мы всегда считаем, что задана произвольная векторная норма, а норма матрицы — всегда порожденная этой векторной операторная норма.

Пусть $A \in \mathbb{M}_n$ — некоторая матрица. Матричная геометрическая прогрессия — это последовательность матриц $\{A^s\}$. Поставим вопрос: при каких условиях $A^s \rightarrow 0$?

Лемма 1. Если $|\lambda| < 1$, ν — натуральное, то $C_s^\nu \lambda^{s-\nu} \rightarrow 0$ при $s \rightarrow \infty$.

Доказательство следует из того, что $C_s^\nu \leq s^\nu$. ■

Лемма 2. Пусть

$$D_n = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ 0 & 0 & \lambda & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix} \in \mathbb{M}_n. \quad (1)$$

Если $|\lambda| < 1$, то при $s \rightarrow \infty$ $D_n^s \rightarrow 0$.

Доказательство. Достаточно показать, что при всех k $D_n^s e_k \rightarrow 0$. Введем обозначение: $e_k = 0$ при $k \leq 0$. Тогда при всех $k \leq n$ $D_n e_k = \lambda e_k + e_{k-1}$ и методом индукции легко показывается, что

$$D_n^s e_k = \sum_{\nu=0}^s C_s^\nu \lambda^{s-\nu} e_{k-\nu}.$$

Действительно, при $s = 1$ это равенство очевидно. Покажем возможность индуктивного перехода от s к $s + 1$.

$$D_n^{s+1}e_k = \sum_{\nu=0}^s C_s^\nu \lambda^{s+1-\nu} e_{k-\nu} + \sum_{\nu=0}^s C_s^\nu \lambda^{s-\nu} e_{k-\nu-1}.$$

Совершая во второй сумме замену индекса $\nu = \nu' - 1$ и обозначая новый индекс ν' снова через ν , имеем

$$D_n^{s+1}e_k = \lambda^{s+1}e_k + \sum_{\nu=1}^s (C_s^\nu + C_s^{\nu-1}) \lambda^{s+1-\nu} e_{k-\nu} + e_{k-(s+1)} = \sum_{\nu=0}^{s+1} C_{s+1}^\nu \lambda^{(s+1)-\nu} e_{k-\nu},$$

и этим наше равенство доказано. При $s > k - 1$ слагаемые с номерами $\nu \geq k$ обращаются в ноль, так как тогда $e_{k-\nu} = 0$, и потому при таких s

$$D_n^s e_k = \sum_{\nu=0}^{k-1} C_s^\nu \lambda^{s-\nu} e_{k-\nu},$$

и используя лемму 1, получаем требуемое. ■

Определение. *Спектральным радиусом* матрицы A (обозначение $\rho(A)$) называется максимальный из модулей ее собственных чисел.

Теорема 1. *Для того чтобы $A^s \rightarrow 0$, необходимо и достаточно, чтобы спектральный радиус матрицы A был меньше единицы: $\rho(A) < 1$.*

Доказательство. 1) Необходимость докажем от противного. Пусть матрица A имеет собственное число λ , такое что $|\lambda| \geq 1$, и пусть z — соответствующий собственный вектор. Тогда $\|A^s z\| = |\lambda|^s \|z\| \geq \|z\|$ и $A^s z$ не стремится к нулевому вектору, а тем смым A^s — к нулевой матрице.

2) Достаточность легко показывается, если A имеет полную систему собственных векторов, но предполагать это мы не будем и воспользуемся для A канонической формой Жордана: $A = BDB^{-1}$, где

$$D = \begin{pmatrix} D_{n_1} & 0 & \dots & 0 \\ 0 & D_{n_2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & D_{n_l} \end{pmatrix},$$

D_{n_j} — матрицы вида (1) с собственными числами λ_j матрицы A на главной диагонали. Тогда $A^s = BD^s B^{-1}$, причем

$$D^s = \begin{pmatrix} D_{n_1}^s & 0 & \dots & 0 \\ 0 & D_{n_2}^s & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & D_{n_l}^s \end{pmatrix},$$

и в случае $\rho(A) < 1$ $D^s \rightarrow 0$. ■

Рассмотрим теперь сумму членов матричной геометрической прогрессии (ряд)

$$E + A + A^2 + A^3 + \dots \quad (2)$$

и рассмотрим вопрос о сходимости этого ряда. Сходимость матричного ряда понимается, естественно, как сходимость его частных сумм.

Теорема 2. Для того чтобы ряд (2) сходиллся, необходимо и достаточно условие $\rho(A) < 1$. В случае сходимости сумма ряда есть $S = (E - A)^{-1}$.

Доказательство. Необходимость следует из того, что общий член сходящегося ряда обязан стремиться к нулю, и теоремы 1.

Достаточность. Пусть S_m — частная сумма ряда. Тогда $(E - A)S_m = E - A^{m+1}$. Т.к. $\lambda = 1$ не есть собственное число матрицы A ($\rho(A) < 1$), то существует обратная матрица $(E - A)^{-1}$ и $S_m = (E - A)^{-1}(E - A^{m+1})$. Остается перейти к пределу при $m \rightarrow \infty$. ■

Следствие 1. Если $\|A\| < 1^*$, то ряд (2) сходится, матрица $E - A$ обратима и

$$\|(E - A)^{-1}\| \leq \frac{1}{1 - \|A\|}. \quad (3)$$

Доказательство ввиду свойства 8) операторной нормы требуется лишь для оценки (3) и следует из оценки частных сумм ряда (2):

$$\|S_m\| \leq 1 + \|A\| + \|A\|^2 + \dots + \|A\|^m \leq \frac{1}{1 - \|A\|}. \quad \blacksquare$$

Замечание 1. В доказанном следствии матрицу A можно заменить на $-A$, так что если $\|A\| \leq 1$, то существует обратная матрица $(E + A)^{-1}$ и

$$\|(E + A)^{-1}\| \leq \frac{1}{1 - \|A\|}.$$

[* Замечание 2. Доказанное следствие есть частный случай хорошо известной в функциональном анализе теоремы Банаха.*]

Определение. Говорят, что A — это матрица с диагональным преобладанием, если в каждой ее строке модуль диагонального элемента больше суммы модулей остальных элементов этой строки:

$$|a_{kk}| > \sum_{j \neq k} |a_{kj}|, \quad k = 1, 2, \dots, n.$$

Теорема 3. Матрица с диагональным преобладанием обратима (невырожденная).

Доказательство. Диагональные элементы матрицы A отличны от нуля, и она допускает представление $A = D(E + R)$, где

$$D = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}, \quad R = \begin{pmatrix} 0 & a_{12}/a_{11} & \dots & a_{1n}/a_{11} \\ a_{21}/a_{22} & 0 & \dots & a_{2n}/a_{22} \\ \dots & \dots & \dots & \dots \\ a_{n1}/a_{nn} & a_{n2}/a_{nn} & \dots & 0 \end{pmatrix}.$$

*Напомним, что $\|A\|$ — операторная норма.

Здесь $\|R\|_\infty < 1$, поэтому матрица $E + R$ обратима и A представлена в виде произведения двух обратимых матриц. ■

Замечание 3. Так как обратимость матрицы влечет и обратимость транспонированной, то для обратимости матрицы достаточно и “диагонального преобладания в столбцах”.

Определение. *Кругами Гершгорина* матрицы A называются круги на комплексной плоскости

$$\Lambda_k = \{ \lambda \in \mathbb{C} \mid |\lambda - a_{kk}| \leq \sum_{j \neq k} |a_{kj}| \}.$$

Теорема 4. Все собственные числа матрицы A содержатся в объединении ее кругов Гершгорина.

Доказательство. Если $\lambda \notin \Lambda_k$, то $|a_{kk} - \lambda| > \sum_{j \neq k} |a_{kj}|$, так что если $\lambda \notin \cup \Lambda_k$, то $A - \lambda E$ матрица с диагональным преобладанием и потому невырожденная, так что собственных чисел вне объединения кругов Гершгорина нет. ■

Замечание 4. Точно так же все собственные числа содержатся в кругах Гершгорина транспонированной матрицы.

Пусть исходные данные в какой-то задаче вычисления известны нам неточно. Ошибка в решении, вызванная неточностью исходных данных, называется *неустраняемой*. Займемся оценкой неустраняемой ошибки в задачах обращения матриц и решения систем линейных уравнений.

Теорема 5. Пусть матрица A обратима и матрица ΔA удовлетворяет неравенству $\|A^{-1}\| \|\Delta A\| < 1$. Тогда матрица $A + \Delta A$ также обратима и выполняются оценки

$$\|(A + \Delta A)^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \|\Delta A\|}, \quad \|(A + \Delta A)^{-1} - A^{-1}\| \leq \frac{\|A^{-1}\|^2 \|\Delta A\|}{1 - \|A^{-1}\| \|\Delta A\|}.$$

Доказательство. По следствию 1 матрица $E + A^{-1}\Delta A$ обратима и

$$\|(E + A^{-1}\Delta A)^{-1}\| \leq \frac{1}{1 - \|A^{-1}\| \|\Delta A\|}.$$

Т.к. $A + \Delta A = A(E + A^{-1}\Delta A)$, то существует $(A + \Delta A)^{-1} = (E + A^{-1}\Delta A)^{-1}A^{-1}$, откуда сразу же следует первая из доказываемых оценок. Докажем вторую:

$$(A + \Delta A)^{-1} - A^{-1} = [E - A^{-1}(A + \Delta A)](A + \Delta A)^{-1} = -A^{-1}\Delta A(A + \Delta A)^{-1}$$

и остается применить первую, уже доказанную, оценку. ■

Обратимся к задаче решения систем линейных алгебраических уравнений. Пусть x^* — решение системы линейных уравнений

$$Ax = y, \tag{4}$$

а $x^* + \Delta x$ — системы уравнений

$$(A + \Delta A)x = y + \Delta y.$$

Теорема 6. Пусть матрица A обратима и $\|A^{-1}\| \|\Delta A\| < 1^*$. Тогда

$$\|\Delta x\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \|\Delta A\|} [\|\Delta A\| \|x^*\| + \|\Delta y\|].$$

Доказательство. Введем еще вектор $x^* + \Delta_1 x$ — решение системы уравнений $Ax = y + \Delta y$, так что $\Delta_1 x = A^{-1} \Delta y$, $\|\Delta_1 x\| \leq \|A^{-1}\| \|\Delta y\|$. Тогда мы имеем два равенства:

$$\begin{aligned} (A + \Delta A)(x^* + \Delta_1 x) &= y + \Delta y + \Delta A(x^* + \Delta_1 x), \\ (A + \Delta A)(x^* + \Delta x) &= y + \Delta y \end{aligned}$$

и, вычитая из второго первое,

$$(A + \Delta A)(\Delta x - \Delta_1 x) = -\Delta A(x^* + \Delta_1 x),$$

откуда, используя теорему 5:

$$\|\Delta x - \Delta_1 x\| \leq \|(A + \Delta A)^{-1}\| \|\Delta A\| (\|x^*\| + \|\Delta_1 x\|) \leq \frac{\|A^{-1}\| \|\Delta A\|}{1 - \|A^{-1}\| \|\Delta A\|} (\|x^*\| + \|\Delta_1 x\|)$$

и далее

$$\|\Delta x\| \leq \|\Delta x - \Delta_1 x\| + \|\Delta_1 x\| \leq \frac{\|A^{-1}\| \|\Delta A\|}{1 - \|A^{-1}\| \|\Delta A\|} \|x^*\| + \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \|\Delta A\|} \|\Delta y\|,$$

что совпадает с доказываемым неравенством. ■

Замечание 5. Для того чтобы избежать в правой части оценки неизвестной нам величины $\|x^*\|$, можно воспользоваться неравенством $\|x^*\| \leq \|A^{-1}\| \|y\|$.

В теоремах 5 и 6 речь шла об *абсолютных* погрешностях. Введем *относительные* погрешности:

$$\delta A = \frac{\|\Delta A\|}{\|A\|}, \quad \delta y = \frac{\|\Delta y\|}{\|y\|}, \quad \delta A^{-1} = \frac{\|A + \Delta A\|^{-1} - \|A^{-1}\|}{\|A^{-1}\|}, \quad \delta x = \frac{\|\Delta x\|}{\|x^*\|}.$$

Определение. Пусть A невырожденная матрица. Числом обусловленности матрицы A называется

$$\mu(A) = \|A\| \|A^{-1}\|.$$

Здесь $\|A\|$ — операторная норма матрицы, так что число обусловленности, как и операторная матричная норма, связано с введенной в \mathbb{C}^n векторной нормой. Если последняя помечена каким-нибудь значком, то тем же значком помечается и $\mu(A)$. Например, евклидовой векторной норме соответствует $\mu_2(A)$.

Теорема 7. Пусть A обратимая матрица. Если $\mu(A) \delta A < 1$, то

$$\delta A^{-1} \leq \frac{\mu(A) \delta A}{1 - \mu(A) \delta A}, \quad \delta x \leq \frac{\mu(A)}{1 - \mu(A) \delta A} (\delta A + \delta y).$$

*Заметим, что в силу теоремы 5 тогда и $A + \Delta A$ обратима.

Доказательство. Поскольку $\mu(A) \delta A = \|A^{-1}\| \|\Delta A\|$, то в силу теоремы 5 матрица $A + \Delta A$ обратима. По той же теореме

$$\delta A^{-1} \leq \frac{\|A^{-1}\| \|\Delta A\|}{1 - \|A^{-1}\| \|\Delta A\|} = \frac{\mu(A) \delta A}{1 - \mu(A) \delta A}.$$

Перейдем к доказательству второй оценки. По теореме 6 (опять используя равенство $\|A^{-1}\| \|\Delta A\| = \mu(A) \delta A$)

$$\delta x \leq \frac{1}{1 - \mu(A) \delta A} \left(\mu(A) \delta A + \|A^{-1}\| \frac{\|\Delta y\|}{\|x^*\|} \right).$$

Остается оценить $\|x^*\|$ снизу из неравенства $\|y\| = \|Ax^*\| \leq \|A\| \|x^*\|$. ■

Замечание 6. Обратим внимание, что если произведение $\mu(A) \delta A$ настолько мало, что им можно пренебречь по сравнению с единицей, то в доказанных оценках $\mu(A)$ есть коэффициент при оценке δA^{-1} через сумму $\delta A + \delta y$.

Замечание 7. Поскольку $1 = \|E\| = \|A^{-1}A\| \leq \mu(A)$, то число обусловленности матрицы не может быть меньше единицы. В связи с доказанной теоремой 7 принято называть матрицу A и систему уравнений (4) хорошо обусловленными, если $\mu(A)$ невелико, и плохо обусловленными в противном случае. Сказанное, разумеется, не может служить математическим определением и связано с тем, с какой точностью ведутся вычисления и с какой точностью мы хотим получить результат.

Задача 1. Показать, что в случае самосопряженной матрицы A

$$\mu_2(A) = \frac{\max |\lambda_k|}{\min |\lambda_k|},$$

где λ_k — собственные числа матрицы A .

Задача 2. Показать, что для любой матрицы A и любой матричной нормы выполняется равенство

$$\lim_{s \rightarrow \infty} \sqrt[s]{\|A^s\|} = \rho(A).$$

Задача 3. Пусть все матрицы последовательности $\{A_s\}$ таковы, что $\rho(A_s) \leq q < 1$. Можно ли утверждать, что для матриц $B_s = A_s A_{s-1} \dots A_2 A_1$ выполняется соотношение $B_s \rightarrow 0$?

§3 Вопросы устойчивости в задаче на собственные значения

Рассмотрим вопрос, насколько могут изменяться собственные числа и векторы матрицы при малых ее возмущениях. Начнем с примеров, которые иллюстрируют трудности, которые могут здесь возникнуть.

Пример 1. Рассмотрим матрицу порядка n , зависящую от малого параметра ε :

$$A_\varepsilon = \begin{pmatrix} 1 & 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 1 \\ \varepsilon & 0 & 0 & \dots & 0 & 1 \end{pmatrix},$$

так что $A_\varepsilon \in \mathbb{M}_n$ отличается от “канонического ящика Жордана” с единицами на главной диагонали лишь одним элементом: “в левом нижнем углу” стоит $\varepsilon \geq 0$. У

матрицы A_0 $\lambda = 1$ является единственным собственным числом и имеет кратность n . При $\varepsilon > 0$ все собственные числа матрицы A_ε различны: $\lambda_k = 1 - \sqrt[n]{\varepsilon} e^{i2\pi k/n}$, где $k = 0, \dots, n-1$. Отсюда два вывода. Во-первых, кратность собственного числа неустойчива по отношению к малым возмущениям матрицы. Во-вторых, в случае кратного собственного числа его возмущение может иметь меньший порядок малости, чем возмущение самой матрицы. Так, в нашем примере возмущение лишь одного элемента матрицы A_0 на величину ε приводит к возмущению собственных чисел на величину порядка $\varepsilon^{1/n}$. При $n = 10$ и $\varepsilon = 10^{-10}$ для матрицы A_ε окажется $\lambda_0 = 1.1$.

Пример 2. Рассмотрим матрицы второго порядка

$$A_1 = \begin{pmatrix} 1 + \varepsilon & 0 \\ 0 & 1 - \varepsilon \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & \varepsilon \\ \varepsilon & 1 \end{pmatrix};$$

у обеих этих близких при малом ε матриц одни и те же собственные числа $1 + \varepsilon$ и $1 - \varepsilon$, но собственные векторы, нормированные условием $\|z\|_2 = 1$, у первой из них $(1, 0)$ и $(0, 1)$, а у второй $\frac{1}{\sqrt{2}}(1, 1)$ и $\frac{1}{\sqrt{2}}(1, -1)$.

Таким образом, в случае наличия близких собственных чисел малые возмущения матрицы могут кардинально менять систему ее собственных векторов.

Дальше будем рассматривать случай простого собственного числа.

Пусть дана матрица $A_0 \in \mathbb{M}_n$, λ_0 — простое собственное число этой матрицы и z_0 — соответствующий собственный вектор. Устойчивость λ_0 и z_0 означает, что малые изменения матрицы A_0 вызывают малые изменения λ_0 и z_0 , т.е. при малой матрице ΔA матрица $A_0 + \Delta A$ будет иметь собственное число $\lambda_0 + \Delta\lambda$ и соответствующий собственный вектор $z_0 + \Delta z$, где $\Delta\lambda$ и Δz малы. Поскольку собственный вектор определен с точностью до множителя, естественно считать его нормированным каким-либо условием и тем же условием нормированным вектор $z_0 + \Delta z$. Для определенности будем предполагать, что для некоторого вектора y_0 отлично от нуля скалярное произведение (y_0, z_0) , и в качестве нормирующего условия для собственных векторов близких к A матриц выберем

$$(y_0, z) = 1. \quad (1)$$

Теорема 1. Если λ_0 есть простое собственное число матрицы A_0 , то в некоторой окрестности этой матрицы собственные число и вектор (нормированный условием (1)) суть непрерывно дифференцируемые функции ее элементов*.

Доказательство начнем с λ . Пусть $P(A, \lambda)$ — характеристический полином матрицы A . Его коэффициенты суть непрерывно дифференцируемые (и даже полиномиальные) функции элементов матрицы A . Поскольку λ_0 — простое собственное число, то в точке (A_0, λ_0) будет $\frac{\partial}{\partial \lambda} P(A, \lambda) \neq 0$, и по теореме о неявных функциях в некоторой окрестности A_0 корень $\lambda(A)$ характеристического полинома (т.е. собственное число матрицы A) есть непрерывно дифференцируемая функция элементов матрицы A . Дальше мы будем рассматривать матрицы A лишь из этой окрестности.

*Это означает, что в некоторой области $\|A - A_0\| < \varepsilon$ ($\varepsilon > 0$) существуют непрерывно дифференцируемые функции $\lambda(A)$ и $z(A)$ элементов матрицы A , такие что $\lambda(A_0) = \lambda_0$, $z(A_0) = z_0$ и при всех A из этой области $\lambda(A)$ есть собственное число матрицы A , а $z(A)$ — соответствующий ему собственный вектор, нормированный условием (1). Как будет видно из доказательства, можно утверждать, что $\lambda(A)$ и $z(A)$ являются бесконечно дифференцируемыми функциями.

Матрица $A_0 - \lambda_0 E$ имеет нулевой определитель, и не умаляя общности будем считать, что ее первая строка либо нулевая, либо есть линейная комбинация остальных строк. Каждой близкой к A_0 матрице A и числу λ , близкому к λ_0 , сопоставим матрицу $B(A, \lambda)$, которая получается из матрицы $A - \lambda E$ заменой ее первой строки на вектор y_0 . Тогда матрица $B(A_0, \lambda_0)$ невырожденная, поскольку вектор z_0 есть *единственное* решение системы уравнений $B(A_0, \lambda_0)z = e_1$. Если возмущения матрицы A_0 и числа λ_0 достаточно малы, то матрица $B(A, \lambda)$ также будет невырожденной (см. теорему 5 предыдущего параграфа), причем элементы обратной матрицы $B^{-1}(A, \lambda)$ суть непрерывно дифференцируемые функции числа λ и элементов матрицы A . Именно только такие A и λ мы будем теперь рассматривать.

Система уравнений $B(A, \lambda)z = e_1$ определяет z как неявную непрерывно дифференцируемую функцию аргументов A и λ : $z(A, \lambda)$. Тогда $z(A) = z(A, \lambda(A))$ есть непрерывно дифференцируемая функция элементов матрицы A . Остается показать, что $z(A)$ есть собственный вектор матрицы A , соответствующий собственному числу $\lambda(A)$. Действительно, строки матрицы $A - \lambda(A)E$, начиная со второй по последнюю, линейно независимы, т.к. они же — строки неособенной матрицы $B(A, \lambda(A))$. Но сама матрица $A - \lambda(A)E$ особенная, и поэтому ее первая строка либо нулевая, либо есть линейная комбинация остальных. Поэтому, будучи ортогональным строкам со второй по n -ю этой матрицы, вектор $z(A)$ ортогонален и первой, т.е. есть собственный вектор матрицы A . ■

Степень устойчивости простого собственного числа и его собственного вектора зависит от величины их производных по элементам матрицы A . Получать соответствующие оценки мы будем в предположении, что *все* собственные числа матрицы A_0 простые.

Итак, пусть $\lambda_1, \dots, \lambda_n$ — все и притом различные собственные числа матрицы A_0 , z_1, \dots, z_n — соответствующие им собственные векторы. Комплексно сопряженные $\bar{\lambda}_1, \dots, \bar{\lambda}_n$ — собственные числа сопряженной матрицы A_0^* , и пусть y_1, \dots, y_n — соответствующие собственные векторы. Эти последние мы считаем фиксированными и в качестве нормирующего условия для вектора z_k и соответствующего вектора возмущенной матрицы выберем $(z, y_k) = 1$, так что $(z_k, y_j) = \delta_{kj}$.

Пусть ΔA — малое возмущение матрицы A_0 , а $\Delta \lambda_k$ и Δz_k — вызванные им возмущения собственных чисел и векторов (в силу нормирующего условия $(\Delta z_k, y_k) = 0$). Тогда

$$(A_0 + \Delta A)(z_k + \Delta z_k) = (\lambda_k + \Delta \lambda_k)(z_k + \Delta z_k).$$

Переходя к дифференциалам (т.е. пренебрегая слагаемыми порядка малости выше первого), имеем:

$$(dA)z_k + A_0 dz_k = \lambda_k dz_k + d\lambda_k z_k. \quad (2)$$

Умножим это равенство скалярно на y_j :

$$((dA)z_k, y_j) + (A_0(dz_k), y_j) = \lambda_k(dz_k, y_j) + d\lambda_k(z_k, y_j). \quad (3)$$

Здесь $(A_0(dz_k), y_j) = (dz_k, A_0^* y_j) = \lambda_j(dz_k, y_j)$, так что (3) может быть переписано в виде

$$((dA)z_k, y_j) = (\lambda_k - \lambda_j)(dz_k, y_j) + d\lambda_k(z_k, y_j). \quad (4)$$

В силу нормирующего условия $(z_k, y_k) = 1$, поэтому при $j = k$ равенство (4) дает

$$|d\lambda_k| = |((dA)z_k, y_k)| \leq \|dA\|_2 \|z_k\|_2 \|y_k\|_2 = c_k \|dA\|_2.$$

Здесь

$$c_k = \|z_k\|_2 \|y_k\|_2 = \frac{\|z_k\|_2 \|y_k\|_2}{|(z_k, y_k)|}.$$

Последнее выражение в этой формуле удобнее предыдущего в том отношении, что оно не связано с нормирующим условием — написанное отношение инвариантно относительно *любой* нормировки векторов z_k и y_k . Число c_k называется *коэффициентом перекоса* матрицы A_0 , соответствующим собственному числу λ_k . Очевидно, что всегда $c_k \geq 1$, а для эрмитовых матриц $c_k = 1$. Итак, задача нахождения простого собственного числа матрицы хорошо обусловлена, если только соответствующий этому числу коэффициент перекоса невелик.

Обратимся к собственным векторам. При $j \neq k$ $(z_k, y_j) = 0$ и из формулы (4) следует

$$(dz_k, y_j) = ((dA)z_k, y_j) / (\lambda_k - \lambda_j).$$

Разложим вектор dz_k по векторам z_j : $dz_k = \sum_{j=1}^n \alpha_{kj} z_j$, $\alpha_{kj} = (dz_k, y_j)$. Тогда в силу нормирующего условия $\alpha_{kk} = 0$ и при $j \neq k$

$$|\alpha_{kj}| = \frac{|((dA)z_k, y_j)|}{|\lambda_k - \lambda_j|} \leq \frac{\|dA\|_2 \|z_k\|_2 \|y_j\|_2}{|\lambda_k - \lambda_j|},$$

откуда

$$\frac{\|dz_k\|_2}{\|z_k\|_2} \leq \|dA\|_2 \sum_{j \neq k} \frac{\|z_j\|_2 \|y_j\|_2}{|\lambda_k - \lambda_j|} = \|dA\|_2 \sum_{j \neq k} \frac{c_j}{|\lambda_k - \lambda_j|}.$$

Здесь учтено, что $\alpha_{kk} = 0$ ввиду нормирующего условия. Итак, задача отыскания собственного вектора, соответствующего простому собственному числу, хорошо обусловлена, если это собственное число достаточно далеко отстоит от остальных собственных чисел и все коэффициенты перекоса невелики. Впрочем, существенную роль здесь играет еще порядок n матрицы.

В случае эрмитовой матрицы A можно указать оценку возмущения собственных чисел, вызванного возмущением ΔA , если ΔA также самосопряженная матрица. Эта оценка связана с экстремальными свойствами собственных чисел эрмитовых матриц.

Пусть матрица $A \in \mathbb{M}_n$ эрмитова, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ — ее собственные числа, z_1, \dots, z_n — соответствующие собственные векторы, выбранные так, что $(z_k, z_j) = \delta_{kj}$.

Теорема 2. *Справедливы равенства:*

$$\lambda_k = \max\{ (Ax, x) \mid \|x\|_2 = 1, (x, z_j) = 0 \text{ при } j = 1, \dots, k-1 \}.$$

Доказательство. Вектор x , удовлетворяющий выписанным условиям, разложим по векторам z_ν . Коэффициенты при z_ν для $\nu = 1, \dots, k-1$ окажутся равными нулю, и $x = \sum_{\nu=k}^n \alpha_\nu z_\nu$, причем $\sum_{\nu=k}^n |\alpha_\nu|^2 = \|x\|_2^2 = 1$. Поэтому

$$(Ax, x) = \sum_{\nu=k}^n \lambda_\nu |\alpha_\nu|^2 \leq \lambda_k \sum_{\nu=k}^n |\alpha_\nu|^2 = \lambda_k.$$

В то же время вектор z_k также удовлетворяет этим условиям и $(Az_k, z_k) = \lambda_k$. ■

Введем функцию аргументов $h_j \in \mathbb{C}^n$:

$$\varphi_A(h_1, \dots, h_{k-1}) = \max\{ (Ax, x) \mid \|x\|_2 = 1, (x, h_j) = 0, j = 1, \dots, k-1 \}.$$

Теорема 3 (минимально-максимальный принцип Куранта). *Справедливо равенство*

$$\min_{h_j} \varphi_A(h_1, \dots, h_{k-1}) = \lambda_k = \varphi_A(z_1, \dots, z_{k-1}).$$

Доказательство. Второе из доказываемых равенств следует из теоремы 2. Поэтому $\min \varphi_A \leq \lambda_k$. Докажем обратное неравенство. Рассмотрим для произвольных h_j систему однородных линейных уравнений

$$\sum_{\nu=1}^k \alpha_\nu(z_\nu, h_j) = 0, \quad j = 1, \dots, k-1.$$

Число уравнений здесь меньше числа неизвестных, и поэтому система заведомо имеет ненулевое решение. Пусть $\{\alpha_\nu\}$ — ненулевое решение, нормированное условием $\sum |\alpha_\nu|^2 = 1$. Положим $x = \sum_{\nu=1}^k \alpha_\nu z_\nu$. Тогда $\|x\|_2 = 1$, при $j = 1, \dots, k-1$ $(x, h_j) = 0$, т.к. числа α_ν удовлетворяют нашей системе однородных уравнений, и $(Ax, x) = \sum_{\nu=1}^k |\alpha_\nu|^2 \lambda_\nu \geq \lambda_k$. Итак, при любых h_j $\varphi_A(h_1, \dots, h_{k-1}) \geq \lambda_k$. ■

[* З а м е ч а н и е. Аналогичный минимально-максимальный принцип верен также для самосопряженных компактных операторов в гильбертовом пространстве.*]

Лемма. Пусть $A, B \in \mathbb{M}_n$ эрмитовы матрицы, λ_k^A и λ_k^B их собственные числа, расположенные в порядке убывания, $\mu_1 \geq \dots \geq \mu_n$ — собственные числа матрицы $A + B$. Тогда $\mu_k \leq \lambda_k^A + \lambda_1^B$.

Доказательство. Пусть $\{z_k\}$ собственные векторы матрицы A . Рассмотрим множество

$$\mathcal{M} = \{x \in \mathbb{C}^n \mid \|x\|_2 = 1, (x, z_j) = 0 \text{ при } j = 1, \dots, k-1\}.$$

По теореме 2 для $x \in \mathcal{M}$ $(Ax, x) \leq \lambda_k^A$ и $(Bx, x) \leq \lambda_1^B$ и потому $((A+B)x, x) \leq \lambda_k^A + \lambda_1^B$. Теперь имеем

$$\begin{aligned} \mu_k &= \min \varphi_{A+B}(h_1, \dots, h_{k-1}) \leq \varphi_{A+B}(z_1, \dots, z_{k-1}) = \\ &= \max\{ ((A+B)x, x) \mid x \in \mathcal{M} \} \leq \lambda_k^A + \lambda_1^B. \quad \blacksquare \end{aligned}$$

Следствие. В условиях леммы $\mu_k \leq \lambda_k^A + \|B\|$ ($\|B\|$ — любая операторная норма матрицы B).

Теорема 4. Пусть матрицы A и ΔA эрмитовы и λ_k и $\lambda_k + \Delta \lambda_k$ собственные числа матриц A и $A + \Delta A$, расположенные в порядке убывания. Тогда для любой операторной нормы матриц $|\Delta \lambda_k| \leq \|\Delta A\|$.

Доказательство. Согласно следствию $\lambda_k + \Delta \lambda_k \leq \lambda_k + \|\Delta A\|$, так что $\Delta \lambda_k \leq \|\Delta A\|$. Применяя то же следствие к матрице $A = (A + \Delta A) - \Delta A$: $\lambda_k \leq \lambda_k + \Delta \lambda_k + \|\Delta A\|$ и $-\|\Delta A\| \leq \Delta \lambda_k$. ■

Задача. Показать, что в условиях леммы для матрицы A при любом k найдется такая эрмитова матрица B с различными собственными числами, что $\mu_k = \lambda_k^A + \lambda_1^B$.

§4 Метод исключений Гаусса

Методы решения систем линейных уравнений делятся на *точные* и *итеративные*. Точные методы позволяют получить решение системы за конечное число арифметических операций (при этом предполагается, что эти арифметические операции осуществляются точно, без округлений). Итеративные методы позволяют построить последовательность векторов, сходящуюся к решению. В этом параграфе рассматривается наиболее употребительный точный метод.

Идея метода исключения для решения системы линейных уравнений заключается в следующем. Из первого уравнения первая неизвестная выражается через остальные, это выражение подставляется в следующие уравнения, и мы получаем систему на единицу меньшего порядка относительно оставшихся неизвестных, с которой поступаем так же. Эти вычисления продолжаются до тех пор, пока мы не приходим к одному линейному уравнению с одним неизвестным. Опишем этот процесс подробнее. Пусть решаемая система n уравнений есть $Ax = y$. Ради некоторого единообразия компоненты матрицы A будем обозначать через a_{kj}^1 , а вектора y через η_k^1 , и пусть искомое решение $x = (\xi_1, \dots, \xi_n)$.

Первый шаг вычислений заключается в том, что в предположении, что $a_{11}^1 \neq 0$ мы при $k = 2, \dots, n$ из уравнения с номером k будем вычитать первое, умноженное на $l_{k1} = a_{k1}^1/a_{11}^1$. Тогда в полученных уравнениях коэффициенты при ξ_1 окажутся равными нулю, и мы получим относительно ξ_2, \dots, ξ_n систему уравнений

$$a_{k2}^2 \xi_2 + \dots + a_{kn}^2 \xi_n = \eta_k^2, \quad k = 2, \dots, n.$$

где

$$a_{kj}^2 = a_{kj}^1 - l_{k1} a_{1j}^1, \quad \eta_k^2 = \eta_k^1 - l_{k1} \eta_1^1.$$

Если $a_{22}^2 \neq 0$, то с этой системой $(n-1)$ -го уравнения мы поступим так же, как с исходной. Вообще если мы уже получили систему уравнений

$$a_{km}^m \xi_m + \dots + a_{kn}^m \xi_n = \eta_k^m, \quad k = m, \dots, n$$

относительно неизвестных ξ_m, \dots, ξ_n и $a_{mm}^m \neq 0$, то переход к системе уравнений относительно ξ_{m+1}, \dots, ξ_n осуществляется по формулам

$$l_{km} = a_{km}^m/a_{mm}^m, \quad a_{kj}^{m+1} = a_{kj}^m - l_{km} a_{mj}^m, \quad \eta_k^{m+1} = \eta_k^m - l_{km} \eta_m^m, \\ k = m+1, \dots, n.$$

Вычисление коэффициентов этих систем называется прямым ходом. Для его осуществимости необходимо, чтобы все коэффициенты a_{mm}^m были отличны от нуля, и мы будем делать такое предположение. Получив, наконец, одно уравнение с одним неизвестным $a_{nn}^n \xi_n = \eta_n^n$, мы найдем $\xi_n = \eta_n^n/a_{nn}^n$. Остальные неизвестные находятся по формулам:

$$\xi_m = (\eta_m^m - a_{m,m+1}^m \xi_{m+1} - \dots - a_{mn}^m \xi_n)/a_{mm}^m, \quad m = n-1, n-2, \dots, 1.$$

Это — обратный ход.

Числа, вычисленные в процессе прямого хода, сосредоточим в двух матрицах

$$U = \begin{pmatrix} a_{11}^1 & a_{12}^1 & a_{13}^1 & \dots & a_{1n}^1 \\ 0 & a_{22}^2 & a_{23}^2 & \dots & a_{2n}^2 \\ 0 & 0 & a_{33}^3 & \dots & a_{3n}^3 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{nn}^n \end{pmatrix}, \quad L = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ l_{21} & 1 & 0 & \dots & 0 \\ l_{31} & l_{32} & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ l_{n1} & l_{n2} & l_{n3} & \dots & 1 \end{pmatrix}$$

и векторе $z = (\zeta_1, \zeta_2, \dots, \zeta_n) = (\eta_1^1, \eta_2^2, \dots, \eta_n^n)$. Как видно из приведенных выше формул, при обратном ходе вектор x получается как решение системы уравнений $Ux = z$ с верхней треугольной матрицей U .

Обратимся к вычислению правых частей уравнений. Рекуррентную формулу для правых частей перепишем, заменив m на $m - 1$, в виде

$$\eta_k^m = \eta_k^{m-1} - l_{km-1}\eta_{m-1}^{m-1} = \eta_k^{m-1} - l_{km-1}\zeta_{m-1}.$$

Отсюда

$$\begin{aligned} \zeta_m = \eta_m^m &= \eta_m^{m-1} - l_{mm-1}\zeta_{m-1} = \eta_m^{m-2} - l_{mm-2}\zeta_{m-2} - l_{mm-1}\zeta_{m-1} = \dots = \\ &= \eta_m^1 - l_{m1}\zeta_1 - l_{m2}\zeta_2 - \dots - l_{mm-1}\zeta_{m-1} \end{aligned}$$

или

$$l_{m1}\zeta_1 + l_{m2}\zeta_2 + \dots + l_{mm-1}\zeta_{m-1} + \zeta_m = \eta_m^1.$$

В левой части этого равенства стоит скалярное произведение m -ой строки матрицы L на вектор z , а в правой — компонента вектора y . Поэтому $Lz = y$, $LUx = y = Ax$. Последнее равенство выполняется для всех векторов x , и потому $A = LU$. Это представление матрицы A в виде произведения нижней треугольной матрицы L и верхней треугольной U называется ее *LU-разложением* или *LU-факторизацией*. Напомним, что условием получения такого разложения является отличие от нуля чисел a_{mm}^m , называемых *ведущими элементами*.

Итак, метод исключений может рассматриваться как процесс построения *LU*-разложения матрицы A , причем построение решения x сводится к решению двух систем уравнений с треугольными матрицами: $Lz = y$ и $Ux = z$. Матрицы L и U не зависят от вектора правых частей, и знание *LU*-разложения матрицы A позволяет достаточно просто решить систему уравнений $Ax = y$ при любом векторе y путем решения двух систем уравнений с треугольными матрицами — это требует примерно того же числа арифметических операций, как если бы мы знали обратную матрицу A^{-1} и использовали формулу $x = A^{-1}y$, причем построение обратной матрицы требует существенно больших вычислений, чем вычисление *LU*-разложения, особенно для матриц высокого порядка.

Метод в изложенной форме неприменим, если на некотором шаге прямого хода окажется $a_{mm}^m = 0$. Но плохо и в том случае, если ведущий элемент a_{mm}^m мал по абсолютной величине — он получен путем вычитания, и его малость означает обычно, что велика его относительная погрешность (мы ведем счет с округлением результатов арифметических действий), а тогда и все последующие результаты будут иметь

малую точность. Чтобы избежать этого, часто применяют перенумерацию уравнений или неизвестных, или того и другого. В связи с этим различают схемы исключения неизвестных с выбором ведущего элемента по строке, по столбцу или по всей матрице. На первом шаге в первом случае это означает такую перестановку уравнений, чтобы в полученной после перестановки системе оказалось $|a_{11}| \geq |a_{k1}|$ ($k = 2, \dots, n$), во втором — такую перенумерацию неизвестных, чтобы оказалось $|a_{11}| \geq |a_{1j}|$ ($j = 2, \dots, n$), а в последнем — такую перестановку уравнений и перенумерацию неизвестных, чтобы было $|a_{11}| \geq |a_{kj}|$ ($k, j = 1, \dots, n$). Подобные перестановки и (или) перенумерации осуществляются на каждом шаге прямого хода. Заметим еще, что если на некотором шаге все элементы некоторой строки матрицы U оказались равными нулю, то это означает линейную зависимость строк матрицы A , т.е. $\det A = 0$.

Задача. Каково LU -разложение трехдиагональной матрицы

$$A = \begin{pmatrix} b_1 & c_1 & 0 & \dots & 0 & 0 \\ a_2 & b_2 & c_2 & \dots & 0 & 0 \\ 0 & a_3 & b_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_n & b_n \end{pmatrix} ?$$

§5 Итеративные методы решения систем линейных уравнений

Метод простой итерации.

Пусть система линейных алгебраических уравнений записана в виде

$$x = Ax + y. \quad (1)$$

Метод простой итерации заключается в том, что, взяв произвольный вектор (начальное приближение) x_0 , строят последовательность векторов

$$x_s = Ax_{s-1} + y.$$

Очевидно, что если $x_s \rightarrow x^*$, то x^* — решение системы (1). Условимся решение системы всегда обозначать через x^* .

Теорема 1. Для того чтобы при любом начальном приближении x_0 метод простой итерации сходиллся: $x_s \rightarrow x^*$, необходимо и достаточно условие $\rho(A) < 1$.

Доказательство. Из равенств $x^* = Ax^* + y$, $x_s = Ax_{s-1} + y$ сразу же следует, что $x_s - x^* = A(x_{s-1} - x^*)$. Отсюда, по индукции, $x^s - x^* = A^s(x_0 - x^*)$. Поэтому для сходимости метода при любом начальном приближении необходимо и достаточно соотношение $A^s \rightarrow 0$, и остается воспользоваться теоремой 1 из §2 и добавить, что при $\rho(A) < 1$ решение x^* существует и единственно. ■

Замечание 1. Отметим связь метода итерации с рассмотренным в §2 рядом — суммой членов матричной геометрической прогрессии. Методом индукции легко показать, что

$$x_s = A^s x_0 + (E + A + A^2 + \dots + A^{s-1})y.$$

При условии $\rho(A) < 1$ первое слагаемое в правой части этого равенства стремится к нулевому вектору, а второе — к $(E - A)^{-1}y = x^*$.

З а м е ч а н и е 2. В случае $\rho(A) \geq 1$ метод итерации расходится почти при любом выборе начального приближения x_0 . В частности, если матрица A имеет полную систему собственных векторов, то, как легко видеть, метод расходится, если только при разложении вектора $x_0 - x^*$ по собственным векторам матрицы A хоть один из коэффициентов при собственных векторах, отвечающих собственным числам λ , таким что $|\lambda| \geq 1$, отличен от нуля.

Пусть теперь в \mathbb{C}^n задана некоторая норма $\|\cdot\|$, и нормы матриц ниже — это операторные нормы, порожденные введенной векторной.

Теорема 2. Если $\|A\| < 1$, то при любом начальном приближении x_0 метод итерации сходится: $x_s \rightarrow x^*$, и выполняются оценки:

$$\|x_s - x^*\| \leq \frac{\|A\|^s}{1 - \|A\|} \|x_1 - x_0\|, \quad (2)$$

$$\|x_s - x^*\| \leq \frac{\|A\|}{1 - \|A\|} \|x_s - x_{s-1}\|. \quad (3)$$

Доказательство. Сходимость следует из теоремы 1 и неравенства $\rho(A) \leq \|A\|$. Докажем оценку (2). Учитывая равенства $x_0 - x_1 = (E - A)x_0 - y$ и $x^* = (E - A)^{-1}y$, а также оценку $\|(E - A)^{-1}\| \leq 1/(1 - \|A\|)$, имеем

$$\|x_0 - x^*\| \leq \|(E - A)^{-1}\| \|(E - A)x_0 - y\| \leq \frac{1}{1 - \|A\|} \|x_1 - x_0\|,$$

и остается воспользоваться тем, что $x_s - x^* = A^s(x_0 - x^*)$.

Неравенство (3) немедленно следует из (2), если принять за новое начальное приближение $\tilde{x}_0 = x_{s-1}$, и тогда $\tilde{x}_1 = x_s$. ■

З а м е ч а н и е 3. Оценка (2) является *априорной*. Это значит, что ее правая часть может быть вычислена *до того*, как построено само приближение x_s , и она годится для того, чтобы заранее оценить, сколько шагов метода следует сделать, чтобы добиться необходимой точности. Оценка (3) *апостериорная* — ее правая часть вычисляется, когда x_s уже известно, и может служить для принятия решения о прекращении итераций, когда необходимая точность уже достигнута. Заметим одно преимущество оценки (3) — она остается верной, какие бы ошибки ни были допущены при вычислении x_j при $j = 1, \dots, s - 1$.

[* З а м е ч а н и е 4. Теорема 2 легко переносится на случай линейных уравнений в банаховом пространстве, и это связано с известной теоремой Банаха об обратимости оператора $I - A$, где A — линейный оператор в банаховом пространстве, удовлетворяющий условию $\|A\| < 1$. *]

Займемся вопросом о приведении системы уравнений к виду (1), если первоначально она задана в виде $Bx = z$. Остановимся на двух случаях.

1) Пусть $B = \{b_{kj}\}$ — матрица с диагональным преобладанием. Тогда систему $Bx = z$ можно переписать в виде $x = Ax + y$, где

$$A = \begin{pmatrix} 0 & -b_{12}/b_{11} & \dots & -b_{1n}/b_{11} \\ \dots & \dots & \dots & \dots \\ -b_{n1}/b_{nn} & -b_{n2}/b_{nn} & \dots & 0 \end{pmatrix}, \quad y = \begin{pmatrix} \zeta_1/b_{11} \\ \dots \\ \zeta_n/b_{nn} \end{pmatrix}. \quad (4)$$

Очевидно, что $\|A\|_\infty < 1$, и потому метод итерации будет сходиться.

2) Пусть B самосопряженная матрица. Взяв $\alpha \neq 0$, приведем систему к эквивалентной: $x = x - \alpha(Bx - z)$, так что $A = E - \alpha B$, и займемся вопросом о выборе α . Пусть μ_k — собственные числа матрицы B . Тогда собственные числа матрицы A суть $\lambda_k = 1 - \alpha\mu_k$. Если собственные числа матрицы B имеют разные знаки, то среди λ_k найдется хоть одно, большее единицы. Поэтому выбор числа α , гарантирующий сходимость метода итерации, возможен лишь при условии, что все μ_k имеют один знак. Для определенности будем считать, что они положительны (тогда самосопряженная матрица B называется положительной), так что следует выбрать $\alpha > 0$. Мы заинтересованы в том, чтобы норма $\|A\|_2 = \rho(A) = \max\{\alpha M - 1, 1 - \alpha m\}$, где $m = \min \mu_k$, $M = \max \mu_k$, была минимальной. Этому соответствует выбор $\alpha = \frac{2}{M+m}$, тогда $\rho(A) = \frac{M-m}{M+m} < 1$. Заметим, что если собственные числа B имеют сильный разброс (M во много раз больше, чем m), то $\rho(A)$ близко к 1, и сходимость метода итерации может оказаться довольно медленной. Собственные числа матрицы B обычно неизвестны. Легко видеть, что если выбрать число α так, что $0 < \alpha < \frac{2}{M'}$, где M' — любая верхняя граница для μ_k (например, $M' = \|B\|$), то окажется $\rho(A) < 1$, и метод итерации будет сходиться.

Метод Зайделя.

Это — другой итеративный способ решения системы уравнений (1). Предполагается заданным некоторое начальное приближение к решению x_0 . Введем обозначения для компонент последующих приближений: $x_s = (\xi_1^s, \dots, \xi_n^s)$. Вычислительные формулы метода:

$$\left. \begin{aligned} \xi_1^{s+1} &= a_{11}\xi_1^s + a_{12}\xi_2^s + \dots + a_{1n}\xi_n^s + \eta_1 \\ \xi_2^{s+1} &= a_{21}\xi_1^{s+1} + a_{22}\xi_2^s + \dots + a_{2n}\xi_n^s + \eta_2 \\ &\quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \xi_n^{s+1} &= a_{n1}\xi_1^{s+1} + \dots + a_{nn-1}\xi_{n-1}^{s+1} + a_{nn}\xi_n^s + \eta_n \end{aligned} \right\}.$$

Отличие метода Зайделя от метода простой итерации состоит в том, что как только получено значение некоторой компоненты очередного приближения, оно сразу же начинает использоваться при вычислении следующих компонент того же приближения. В векторной форме формулы метода можно записать так: $x_{s+1} = Rx_s + Lx_{s+1} + y$, где

$$R = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}, \quad L = \begin{pmatrix} 0 & 0 & \dots & 0 \\ a_{21} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn-1} & 0 \end{pmatrix}$$

(так что $A = R + L$), или $x_{s+1} = (E - L)^{-1}Rx_s + (E - L)^{-1}y$. Таким образом применение метода Зайделя к системе уравнений (1) — это то же самое, что применение метода простой итерации к равносильной (1) системе уравнений $x = (E - L)^{-1}Rx + (E - L)^{-1}y$, и потому из теоремы 1 сразу же вытекает, что для сходимости метода Зайделя при любом начальном приближении x_0 необходимо и достаточно неравенство $\rho((E - L)^{-1}R) < 1$. Это условие трудно проверять, и мы докажем некоторый достаточный признак сходимости, но сначала отметим, что известны примеры, когда метод простой итерации для системы уравнений (1) сходится, а метод Зайделя нет, и наоборот, когда сходится метод Зайделя, но не сходится метод простой итерации.

Теорема 3. Если $\|A\|_\infty < 1$, то метод Зайделя сходится при любом начальном приближении.

Доказательство. При $i = 1, 2, \dots, n$, считая $\beta_1 = 0$, положим

$$\beta_i = \sum_{j=1}^{i-1} |a_{ij}|, \quad \gamma_i = \sum_{j=i}^n |a_{ij}|.$$

Тогда $\max_i (\beta_i + \gamma_i) = \|A\|_\infty < 1$ и кроме того выполняется неравенство

$$\frac{\gamma_i}{1 - \beta_i} \leq \beta_i + \gamma_i \leq \|A\|_\infty, \quad \text{так как} \quad \beta_i + \gamma_i - \frac{\gamma_i}{1 - \beta_i} = \frac{\beta_i(1 - \beta_i - \gamma_i)}{1 - \beta_i} \geq 0.$$

Из формулы метода следует, что $x^* - x_{s+1} = L(x^* - x_{s+1}) + R(x^* - x_s)$. Пусть теперь номер ν таков, что $\|x^* - x_{s+1}\|_\infty = |\xi_\nu^* - \xi_\nu^{s+1}|$. При этом

$$\xi_\nu^* - \xi_\nu^{s+1} = \sum_{j=1}^{\nu-1} a_{\nu j}(\xi_j^* - \xi_j^{s+1}) + \sum_{j=\nu}^n a_{\nu j}(\xi_j^* - \xi_j^s),$$

так что

$$\begin{aligned} \|x^* - x_{s+1}\|_\infty &\leq \beta_\nu \|x^* - x_{s+1}\|_\infty + \gamma_\nu \|x^* - x_s\|_\infty, \\ \|x^* - x_{s+1}\|_\infty &\leq \frac{\gamma_\nu}{1 - \beta_\nu} \|x^* - x_s\|_\infty \leq \|A\|_\infty \|x^* - x_s\|_\infty. \end{aligned}$$

Так как последнее неравенство выполняется при всех s , то дальнейшее очевидно. ■

Метод Некрасова

Если первоначально система уравнений задана в виде $Bx = y$, причем все диагональные элементы матрицы B отличны от нуля, то она приводится к виду (1), если для A и y воспользоваться формулами, аналогичными формулам (4), которые применялись по отношению к методу простой итерации. Метод Зайделя, примененный к полученной таким путем системе (1), называется *методом Некрасова* для исходной системы. Таким образом, компоненты следующего приближения x^{s+1} в методе Некрасова находятся по компонентам предыдущего из уравнений

$$\left. \begin{aligned} b_{11}\xi_1^{s+1} &+ b_{12}\xi_2^s &+ b_{13}\xi_3^s &+ \dots &+ b_{1n}\xi_n^s &= \eta_1 \\ b_{21}\xi_1^{s+1} &+ b_{22}\xi_2^{s+1} &+ b_{23}\xi_3^s &+ \dots &+ b_{2n}\xi_n^s &= \eta_2 \\ \dots &\dots &\dots &\dots &\dots &\dots \\ b_{n1}\xi_1^{s+1} &+ b_{n2}\xi_2^{s+1} &+ b_{n3}\xi_3^{s+1} &+ \dots &+ b_{nn}\xi_n^{s+1} &= \eta_n \end{aligned} \right\}.$$

Теорема 4. Если выполнено хотя бы одно из условий

- а) матрица B имеет диагональное преобладание,
- б) матрица B положительно определена,

то метод Некрасова сходится при любом начальном приближении.

Доказательство. В случае а) утверждение теоремы непосредственно следует из теоремы 3.

Случай б). Положительно определенную матрицу B представим в виде суммы $B = R + D + R^*$ (R^* — матрица, сопряженная R), где

$$D = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}, \quad R = \begin{pmatrix} 0 & a_{12} & \dots & a_{1n} \\ 0 & 0 & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{pmatrix}.$$

Тогда формулы метода Некрасова приводятся к виду

$$(D + R^*)x_{s+1} + Rx_s = y, \quad x_{s+1} = -(D + R^*)^{-1}Rx_s + (D + R^*)^{-1}y,$$

и для доказательства сходимости метода достаточно показать, что $\rho((D + R^*)^{-1}R) < 1$. Пусть теперь λ и z собственное число и соответствующий собственный вектор этой матрицы, так что

$$Rz = \lambda Dz + \lambda R^*z = \lambda Bz - \lambda Rz.$$

Введем обозначения:

$$(Bz, z) = p > 0, \quad (Dz, z) = d > 0, \quad (Rz, z) = a + ib, \quad \text{так что } (R^*z, z) = a - ib.$$

Тогда $p = d + 2a$, и т.к. $(Rz, z) = \lambda(Bz, z) - \lambda(Rz, z)$, то $a + ib = \lambda(p - a - ib)$. Далее

$$\lambda = \frac{a + ib}{p - a - ib}, \quad |\lambda|^2 = \frac{a^2 + b^2}{(p - a)^2 + b^2} = \frac{a^2 + b^2}{a^2 + b^2 + p(p - 2a)} = \frac{a^2 + b^2}{a^2 + b^2 + pd} < 1.$$

Поскольку λ любое собственное число, нужное неравенство для спектрального радиуса $\rho((D + R^*)^{-1}R)$ доказано. ■

Итеративный метод с чебышевскими параметрами

Изложению метода предположим простые сведения о полиномах от матриц. Пусть дан полином $P_q \in \mathbb{P}_q$: $P_q(\lambda) = a_0\lambda^q + a_1\lambda^{q-1} + \dots + a_q$. Введем обозначение: если $A \in \mathbb{M}_n$ — матрица, то $P_q(A) = a_0A^q + a_1A^{q-1} + \dots + a_{q-1}A + a_qE \in \mathbb{M}_n$. Матрица $P_q(A)$ обладает следующими очевидными свойствами. Если матрица A самосопряженная, то такова же и $P_q(A)$. Если λ — собственное число матрицы A , а z — соответствующий собственный вектор: $Az = \lambda z$, то

$$P_q(A)z = a_0\lambda^q z + a_1\lambda^{q-1}z + \dots + a_q z = P_q(\lambda)z,$$

так что z является собственным вектором матрицы $P_q(A)$, соответствующим собственному числу $P_q(\lambda)$.

Теперь перейдем к самому методу. Как было показано в разделе, посвященном методу простой итерации, если B — положительная матрица, собственные числа которой μ_k удовлетворяют неравенствам $0 < m \leq \mu_k \leq M$, то при $0 < \alpha < 2/M$ сходится итеративный процесс решения системы уравнений $Bx = z$: $x_{s+1} = x_s - \alpha(Bx_s - z)$. Возникает вопрос, нельзя ли добиться лучшей сходимости, если использовать переменный шаг α_s , а именно положить $x_{s+1} = x_s - \alpha_s(Bx_s - z)$. Тогда

$$x_{s+1} - x^* = x_s - x^* - \alpha_s B(x_s - x^*) = (E - \alpha_s B)(x_s - x^*),$$

и если мы проделаем p таких шагов, то

$$x_p - x^* = \prod_{s=0}^{p-1} (E - \alpha_s B) (x_0 - x^*) = Q_p(B)(x_0 - x^*)$$

(при этой записи мы воспользовались тем, что матрицы $(E - \alpha_s B)$ коммутируют). Здесь $Q_p(\lambda) = \prod_{s=0}^{p-1} (1 - \alpha_s \lambda)$ есть полином степени p , такой что $Q_p(0) = 1$. Легко убедиться, что если $p = n$ и в качестве параметров мы выбираем $\alpha_s = 1/\mu_s$, то $x_n = x^*$: для любого собственного вектора z_s матрицы B $Q_n(B)z_s = Q_n(\mu_s)z_s = 0$, а любой вектор x можно представить как линейную комбинацию векторов z_s . Но собственные числа μ_s нам неизвестны, мы будем считать известными лишь их нижнюю и верхнюю границы: $0 < m \leq \mu_s \leq M$. Поэтому представляется естественным (независимо от соотношения между p и n) выбрать параметры α_s так, чтобы полином Q_p был равномерно мал на всем промежутке $[m, M]$. Но среди всех полиномов $Q_p(\lambda) \in \mathbb{P}_p$, удовлетворяющих условию $Q_p(0) = 1$, наименее уклоняется от нуля на $[m, M]$ полином Чебышева, "приведенный" к этому промежутку и нормированный указанным условием (см. следствие 3 в § 1 первой главы):

$$Q_p(\lambda) = (-1)^p \frac{T_p((2\lambda - M - m)/(M - m))}{T_p((M + m)/(M - m))} = \prod_{s=1}^p (1 - \alpha_s \lambda).$$

Здесь α_s — числа, обратные корням многочлена Q_p :

$$\alpha_s = \left(\frac{M - m}{2} t_s + \frac{M + m}{2} \right)^{-1},$$

где $t_s = \cos \frac{2s+1}{2p} \pi$ — корни полинома Чебышева. Матрица $Q_p(B)$ самосопряженная. Ее собственные числа равны $Q_p(\mu_s)$, причем $\mu_s \in [m, M]$. Поэтому

$$\|Q_p(B)\|_2 = \max |Q_p(\mu_s)| \leq \|Q_p\|_{C[m, M]} = \left[T_p \left(\frac{M + m}{M - m} \right) \right]^{-1}.$$

Это дает оценку погрешности:

$$\|x_p - x^*\|_2 \leq \frac{\|x_0 - x^*\|_2}{T_p \left(\frac{M + m}{M - m} \right)}.$$

Напомним, что $T_p \left(\frac{M + m}{M - m} \right)$ не меньше модуля значения в этой точке любого полинома степени p , который в точке $\lambda = 0$ равен единице.

Если при выбранном p точность приближения x_p нас не удовлетворяет (что обычно бывает при не слишком большом p), то процесс продолжают с циклическим повторением параметров α_s . Отметим еще, что порядок нумерации этих параметров не безразличен с точки зрения влияния допускаемых в процессе счета ошибок округления. Известен вполне благоприятный способ их нумерации (см. *В.И.Лебедев, С.А.Финогенов*).

Задача 1. Доказать, что правая часть оценки (3) всегда не больше правой части (2).

Задача 2. Показать, что в случае метода простой итерации при $\rho(A) < 1$ по любым $\varepsilon > 0$ и начальному приближению x_0 найдется такая постоянная C_ε , что

$$\|x_s - x^*\| \leq C_\varepsilon [\rho(A) + \varepsilon]^s.$$

§6 Обращение матриц

Построить матрицу $D = A^{-1}$, обратную данной, можно используя метод исключений Гаусса, поскольку j -й столбец матрицы D есть решение системы уравнений $Ax = e_j$. Так как все эти n систем уравнений имеют одну и ту же матрицу коэффициентов и отличаются лишь правыми частями, то естественно использовать LU-разложение матрицы A .

При наличии не слишком плохого приближения D_0 к A^{-1} можно для построения обратной матрицы использовать итеративный процесс

$$D_{s+1} = D_s(2E - AD_s). \quad (1)$$

Стоит обратить внимание, что эта вычислительная формула вполне аналогична формуле метода Ньютона (касательных) для решения числового уравнения $\frac{1}{t} - a = 0$ (об этом методе пойдет речь в следующей главе).

Формулу (1) удобно переписать в виде

$$R_s = E - AD_s, \quad D_{s+1} = D_s(E + R_s), \quad s = 0, 1, \dots \quad (2)$$

Здесь R_s — это невязка приближения D_s к решению D уравнения $AD = E$, и если D_0 достаточно близкое к D начальное приближение, то матрица R_0 будет малой.

Теорема. Для того чтобы при $s \rightarrow \infty$ матрицы D_s сходились к A^{-1} , необходимо и достаточно выполнение неравенства $\rho(R_0) < 1$. Если при этом* $\|R_0\| = q < 1$, то выполняется оценка

$$\|D_s - A^{-1}\| \leq \frac{q^{2^s}}{1 - q} \|D_0\|. \quad (3)$$

Доказательство. Заметим прежде всего, что если $\rho(R_0) < 1$, то тогда матрица $E - R_0 = AD_0$ неособенная, а значит, обратима и A . Преобразуем формулу для R_{s+1} :

$$R_{s+1} = E - AD_{s+1} = E - AD_s(E + R_s) = E - AD_s - AD_s R_s = R_s - AD_s R_s = R_s^2.$$

Отсюда легко следует, что $R_s = R_0^{2^s}$.

Докажем необходимость. Если $D_s \rightarrow A^{-1}$, то $R_s \rightarrow 0$. Поскольку $R_s = R_0^{2^s}$, то неравенство $\rho(R_0) < 1$ является для этого необходимым условием.

Докажем достаточность и оценку (3). Если $\rho(R_0) < 1$, то $R_s = R_0^{2^s} \rightarrow 0^+$, т.е. $AD_s \rightarrow E$ и $D_s \rightarrow A^{-1}$. Если $\|R_0\| = q < 1$, то $\|R_s\| = \|R_0^{2^s}\| \leq \|R_0\|^{2^s} = q^{2^s}$. Далее

$$A^{-1} - D_s = A^{-1} R_s, \quad \|A^{-1} - D_s\| \leq \|A^{-1}\| q^{2^s}$$

*Как обычно, считаем заданной некоторую векторную норму, и матричная — операторная, порожденная этой векторной.

†Здесь мы ссылаемся на теорему 1 из § 2, в то время как при доказательстве необходимости нам приходится ссылаться на *доказательство* этой теоремы, поскольку мы имеем дело с *подпоследовательностью* последовательности R_0^s .

и остается заметить, что

$$AD_0 = E - R_0, \quad A = (E - R_0)D_0^{-1}, \quad A^{-1} = D_0(E - R_0)^{-1},$$

$$\|A^{-1}\| \leq \|D_0\| \|(E - R_0)^{-1}\|,$$

причем $\|(E - R_0)^{-1}\| \leq \frac{1}{1-q}$ ■

Замечание. Следует подчеркнуть, что оценка (3) означает очень быструю сходимость процесса (2).

Следствие. Если A самосопряженная невырожденная матрица, то при $D_0 = \alpha A$, где $0 < \alpha < \frac{2}{(\rho(A))^2}$, процесс (2) сходится.

Доказательство. Если λ_k — собственные числа матрицы A (они вещественны), то при указанном выборе D_0 собственными числами матрицы R_0 будут $1 - \alpha\lambda_k^2$. Все они меньше 1, а неравенство $\alpha < \frac{2}{(\rho(A))^2}$ гарантирует, что они больше -1. ■

Следствие гарантирует сходимость процесса для самосопряженной матрицы, если взять, например, $D_0 = \frac{1}{\|A\|^2}A$ при любой операторной норме матрицы.

§7 Степенной метод

В задаче на отыскание собственных чисел и векторов матрицы A различают полную и частичную проблему собственных чисел. В первом случае требуется найти все собственные числа и векторы, а во втором — одно собственное число (чаще всего максимальное по модулю) и соответствующий собственный вектор. Степенной метод — это метод решения частичной проблемы собственных чисел.

Идея метода основана на анализе поведения последовательности векторов

$$y_k = A^k y_0 = Ay_{k-1}, \quad (1)$$

где y_0 произвольно взятый ненулевой вектор. Будем считать, что матрица A имеет полную систему собственных векторов, λ_j — ее собственные числа, а z_j — соответствующие собственные векторы, причем $|\lambda_1| \geq \dots \geq |\lambda_n|$. Пусть $y_0 = \sum_{j=1}^n \alpha_j z_j$ — разложение вектора y_0 по собственным векторам. Тогда

$$y_k = \sum_{j=1}^n \alpha_j \lambda_j^k z_j = \lambda_1^k \left[\alpha_1 z_1 + \sum_{j=2}^n \left(\frac{\lambda_j}{\lambda_1} \right)^k \alpha_j z_j \right]. \quad (2)$$

Основной случай

Будем сейчас считать, что $|\lambda_1| > |\lambda_2|$ (основной случай) и что $\alpha_1 \neq 0$ (это условие обычно выполнено для произвольно взятого y_0). Тогда суммой, стоящей в правой части (2), при больших k можно пренебречь, и мы имеем приближенное равенство $y_k \approx \lambda_1^k \alpha_1 z_1$. Пусть u — некоторый вектор, такой что $(z_1, u) \neq 0$. Положим

$$\varphi_k = (y_k, u) = \lambda_1^k \left[\alpha_1 (z_1, u) + \sum_{j=2}^n \left(\frac{\lambda_j}{\lambda_1} \right)^k \alpha_j (z_j, u) \right].$$

Тогда, как видно из этой формулы, при $k \rightarrow \infty$ $\varphi_k / \varphi_{k-1} \rightarrow \lambda_1$ и, более того,

$$\frac{\varphi_k}{\varphi_{k-1}} = \lambda_1 + \mathcal{O} \left(\left(\frac{\lambda_2}{\lambda_1} \right)^k \right).$$

Если $|\lambda_1| < 1$, то векторы y_k стремятся к нулевому, а при $|\lambda_1| > 1$ — к бесконечности (по норме). Это обычно неудобно, и поэтому на каждом шаге процесса часто осуществляют некоторую нормировку этих векторов, так что процесс выглядит так:

$$\tilde{x}_0 = y_0, \quad \text{и при } k = 0, 1, 2, \dots \quad x_k = \frac{\tilde{x}_k}{(\tilde{x}_k, u)}, \quad \tilde{x}_{k+1} = Ax_k. \quad (3)$$

Векторы \tilde{x}_k и x_k лишь числовыми множителями отличаются от вектора y_k : $x_k = \beta_k y_k$, и для введенных выше чисел φ_k имеем

$$\frac{\varphi_k}{\varphi_{k-1}} = \frac{(y_k, u)}{(y_{k-1}, u)} = \frac{(Ay_{k-1}, u)}{(y_{k-1}, u)} = \frac{(Ax_{k-1}, u)}{(x_{k-1}, u)} = (\tilde{x}_k, u),$$

поскольку $(x_{k-1}, u) = 1$, так что приближения к λ_1 и z_1 даются формулами

$$\lambda_1^{(k)} = (\tilde{x}_k, u), \quad z_1^k = x_k.$$

Вектор x_k , будучи нормирован условием $(x_k, u) = 1$, с точностью до вектора порядка малости $\mathcal{O}((\lambda_2/\lambda_1)^k)$ совпадает с собственным вектором z_1 , если последний нормирован тем же условием.

В качестве нормирующего вектора u удобно выбирать один из ортов e_j . После нескольких шагов процесса, когда наметилась некоторая стабилизация векторов x_k , целесообразно в качестве j взять номер максимальной по модулю компоненты x_k .

Фактически доказана теорема о сходимости метода, при формулировке которой будут использоваться введенные выше обозначения; кроме того, мы будем предполагать, что при всех k $(y_k, u) \neq 0$.

Теорема. Пусть выполнены условия:

- 1) $|\lambda_1| > |\lambda_2|$;
- 2) $(z_1, u) = 1$ (нормирующее условие для вектора z_1);
- 3) $\alpha_1 \neq 0$.

Тогда описанный формулами (3) процесс сходится: $\lambda_1^{(k)} \rightarrow \lambda_1$, $z_1^k \rightarrow z_1$, причем

$$\lambda_1^{(k)} = \lambda_1 + \mathcal{O}\left(\left(\frac{\lambda_2}{\lambda_1}\right)^k\right), \quad z_1^k = z_1 + \Delta z^k, \quad \|\Delta z^k\| = \mathcal{O}\left(\left(\frac{\lambda_2}{\lambda_1}\right)^k\right).$$

Замечание . Если в условиях теоремы $|\lambda_2| > |\lambda_3|$, $(u, z_2) \neq 0$ и $\alpha_2 \neq 0$, то характер сходимости $\lambda_1^{(k)}$ к λ_1 можно охарактеризовать более точно:

$$\lambda_1^{(k)} = \lambda_1 + c \left(\frac{\lambda_2}{\lambda_1}\right)^k + \mathcal{O}\left(\left(\frac{\lambda_2}{\lambda_1}\right)^{2k} + \left(\frac{\lambda_3}{\lambda_1}\right)^k\right),$$

где

$$c = \frac{\lambda_1}{\lambda_2} \frac{\alpha_2(z_2, u)}{\alpha_1(z_1, u)} (\lambda_2 - \lambda_1).$$

При сделанных предположениях такое же уточнение можно получить и по отношению к сходимости векторов $z_1^k = (\zeta_1^k, \dots, \zeta_n^k)$ к собственному вектору $z_1 = (\zeta_1, \dots, \zeta_n)$. Действительно,

$$y_k = \lambda_1^k \left[\alpha_1 z_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k z_2 + w_k \right], \quad \|w_k\| = \mathcal{O}\left(\left(\frac{\lambda_3}{\lambda_1}\right)^k\right),$$

и так как $z_1^k = x_k = y_k / (y_k, u)$, то для любого вектора v

$$(z_1^k, v) = \frac{\lambda_1^k [\alpha_1(z_1, v) + \alpha_2(\lambda_2/\lambda_1)^k(z_2, v) + \varepsilon_{k1}]}{\lambda_1^k [\alpha_1 + \alpha_2(\lambda_2/\lambda_1)^k(z_2, u) + \varepsilon_{k2}]}, \quad \varepsilon_{kj} = \mathcal{O} \left(\left(\frac{\lambda_1}{\lambda_3} \right)^k \right).$$

Отсюда нетрудно усмотреть, что

$$(z_1^k, v) = (z_1, v) + c(v) \left(\frac{\lambda_2}{\lambda_1} \right)^k + \mathcal{O} \left(\left(\frac{\lambda_2}{\lambda_1} \right)^{2k} + \left(\frac{\lambda_3}{\lambda_1} \right)^k \right),$$

Взяв в качестве v орты: $v = e_j$, имеем

$$\zeta_j^k = \zeta_j + c_j \left(\frac{\lambda_2}{\lambda_1} \right)^k + \mathcal{O} \left(\left(\frac{\lambda_2}{\lambda_1} \right)^{2k} + \left(\frac{\lambda_3}{\lambda_1} \right)^k \right),$$

Метод скалярных произведений.

Если матрица A самосопряженная, то располагая векторами x_k , построенными согласно (3), можно получить существенно лучшее приближение к λ_1 . Действительно, поскольку собственные векторы z_j попарно ортогональны, то

$$(y_k, y_k) = \alpha_1^2 \lambda_1^{2k} (z_1, z_1) + \sum_{j=2}^n \alpha_j^2 \lambda_j^{2k} (z_j, z_j),$$

аналогичный вид имеет и (y_k, y_{k-1}) , так что

$$\frac{(y_k, y_k)}{(y_k, y_{k-1})} = \lambda_1 \frac{1 + \sum_{j=2}^n \gamma_j \left(\frac{\lambda_j}{\lambda_1} \right)^{2k}}{1 + \sum_{j=2}^n \gamma_j \left(\frac{\lambda_j}{\lambda_1} \right)^{2k-1}}, \quad \gamma_j = \left(\frac{\alpha_j}{\alpha_1} \right)^2 \frac{(z_j, z_j)}{(z_1, z_1)}.$$

Таким образом

$$\tilde{\lambda}_1^{(k)} = \frac{(y_k, y_k)}{(y_k, y_{k-1})} = \frac{(\tilde{x}_k, \tilde{x}_k)}{(\tilde{x}_k, \tilde{x}_{k-1})} = \lambda_1 + \mathcal{O} \left(\left(\frac{\lambda_2}{\lambda_1} \right)^{2k} \right)$$

есть, вообще говоря, существенно лучшее приближение к λ_1 , чем $\lambda_1^{(k)}$. В этом и состоит метод скалярных произведений. Улучшить этим методом приближение к собственному вектору самосопряженной матрицы не удастся.

Случай $|\lambda_2| = |\lambda_1|$.

Остановимся кратко на случае $|\lambda_2| = |\lambda_1|$, считая для простоты матрицу A вещественной; тогда и векторы y_0 и u естественно считать вещественными. Из всех возможностей разберем три наиболее характерные.

1) λ_1 — кратное собственное число: $\lambda_1 = \dots = \lambda_m$, причем $|\lambda_1| > |\lambda_{m+1}|$. Тогда

$$y_k = \lambda_1^k \left[\sum_{j=1}^m \alpha_j z_j + \sum_{j=m+1}^n \alpha_j \left(\frac{\lambda_j}{\lambda_1} \right)^k z_j \right],$$

откуда видно, что $\lambda_1^{(k)}$ и z_1^k , построенные по формулам (3), сходятся соответственно к λ_1 и некоторому из собственных векторов матрицы A , соответствующих λ_1 , с той же быстротой, что указана в теореме: $\mathcal{O}((\lambda_{m+1}/\lambda_1)^k)$. То обстоятельство, что λ_1 кратное собственное число, в процессе счета замечено не будет.

2) $\lambda_2 = -\lambda_1$, $|\lambda_1| > |\lambda_3|$. В этом случае

$$y_k = \lambda_1^k \left[\alpha_1 z_1 + (-1)^k \alpha_2 z_2 + \sum_{j=3}^n \alpha_j \left(\frac{\lambda_j}{\lambda_1} \right)^k z_j \right].$$

Поэтому характерным признаком, по которому мы можем судить, что имеем дело с этим случаем, будет близость при больших k векторов x_k и x_{k-2} при существенном отличии x_{k-1} от x_k . Приближения к собственным числам и векторам можно вычислить по формулам

$$(\lambda_1^{(k)})^2 = (\lambda_2^{(k)})^2 = \frac{(y_k, u)}{(y_{k-2}, u)}, \quad z_m^k = c_m(y_k + \lambda_m^{(k)} y_{k-1}), \quad m = 1, 2.$$

Заниматься представлением правых частей этих формул через векторы \tilde{x}_k и x_k не будем.

3) $\lambda_1 = \rho e^{i\theta}$ и $\lambda_2 = \rho e^{-i\theta}$ — комплексно сопряженные собственные числа, $\rho > |\lambda_3|$. Векторы z_1 и z_2 имеют комплексно сопряженные компоненты. Поскольку векторы y_0 и u вещественные, то числа $\alpha_{1-2}(z_{1-2}, u) = r e^{\pm i\varphi}$ также комплексно сопряжены. Поэтому числа

$$(y_k, u) = \rho^k [2r \cos(k\theta + \varphi) + \mathcal{O}((\lambda_3/\rho)^k)]$$

ведут себя “неправильно”, что и служит признаком комплексно сопряженных наибольших по модулю собственных чисел. Модуль первых двух собственных чисел может быть приближенно найден по формуле

$$\rho^2 \approx \frac{(y_k, u)(y_{k-2}, u) - (y_{k-1}, u)^2}{(y_{k-1}, u)(y_{k-3}, u) - (y_{k-2}, u)^2},$$

так как

$$\begin{aligned} (y_k, u)(y_{k-2}, u) - (y_{k-1}, u)^2 &\approx \\ &\approx \rho^{2k-2} 4r^2 [\cos(k\theta + \varphi) \cos((k-2)\theta + \varphi) - \cos^2((k-1)\theta + \varphi)] = \\ &= -4r^2 \sin^2 \theta \rho^{2k-2}. \end{aligned}$$

Приводить формулы для вычисления θ не будем.

Метод обратных итераций.

Пусть нам требуется найти *наименьшее по модулю* собственное число λ_n матрицы A . Поскольку $\mu = \frac{1}{\lambda_n}$ — *наибольшее по модулю* собственное число матрицы $B = A^{-1}$, то можно найти λ_n , применяя степенной метод к матрице B . Обращать матрицу A нерационально, более выгодно вычислять на каждом шаге вектор y_k , решая систему уравнений $Ay_k = y_{k-1}$. При решении этих систем методом исключений следует учитывать, что все они имеют одну и ту же матрицу коэффициентов, так что естественно воспользоваться LU -разложением матрицы A .

В этом и состоит метод обратных итераций.

Метод обратных итераций можно применять для уточнения произвольного собственного числа λ матрицы A , если для него известно достаточно хорошее приближение $\lambda^{(0)}$. Следует воспользоваться тем, что $\lambda - \lambda^{(0)}$ есть собственное число матрицы $B = A - \lambda^{(0)}E$, минимальное по модулю, если $\lambda^{(0)}$ — хорошее приближение. Более того, сходимость метода обратных итераций для нахождения $\lambda - \lambda^{(0)}$ будет в этом случае очень быстрой.

Задача. Пусть для матрицы A $\lambda_1 = \rho e^{i\theta}$, $\lambda_2 = \rho e^{-i\theta}$, $\rho > |\lambda_3|$. Рассмотрев последовательность

$$p_k = \frac{(y_{k-1}, u)(y_{k+2}, u) - (y_k, u)(y_{k+1}, u)}{(y_{k-1}, u)(y_{k+1}, u) - (y_k, u)^2},$$

указать способ нахождения $\operatorname{Re} \lambda_1$ степенным методом в этом случае.

§8. Метод Крылова

Метод А.Н.Крылова — это метод решения полной проблемы собственных чисел, основанный на вычислении коэффициентов характеристического полинома матрицы A . С самого начала будем предполагать, что все собственные числа этой матрицы различны — в противном случае метод имеет некоторые осложнения, и останавливаться на этом случае мы не будем.

Метод основан на равенстве Кели–Гамильтона: пусть

$$P_n(\lambda) = \det(A - \lambda E) = (-1)^n \lambda^n + p_1 \lambda^{n-1} + \dots + p_n \quad -$$

характеристический полином матрицы A , тогда*

$$P_n(A) = (-1)^n A^n + p_1 A^{n-1} + \dots + p_n E = 0.$$

В случае, когда матрица A имеет полную систему собственных векторов, это равенство доказывается совсем просто. Действительно, пусть z — собственный вектор: $Az = \lambda z$. Тогда $P_n(A)z = P_n(\lambda)z = 0$ (т.к. λ — собственное число). Ввиду полноты системы собственных векторов и для любого вектора x будет $P_n(A)x = 0$, и значит $P_n(A)$ нулевая матрица.

Обратимся к методу. Возьмем произвольный ненулевой вектор y_0 и вычислим векторы

$$y_k = (\eta_1^{(k)}, \dots, \eta_n^{(k)}), \quad y_k = A^k y_0 = A y_{k-1}, \quad k = 1, \dots, n.$$

Для вычислений, разумеется, используется формула $y_k = A y_{k-1}$. Векторное равенство $P_n(A)y_0 = 0$, которое следует из равенства Кели–Гамильтона, перепишем покомпонентно, перенеся первое слагаемое в правую часть:

$$\left. \begin{aligned} \eta_1^{(n-1)} p_1 + \dots + \eta_1^{(0)} p_n &= (-1)^{n+1} \eta_1^{(n)} \\ \dots &\dots \\ \eta_n^{(n-1)} p_1 + \dots + \eta_n^{(0)} p_n &= (-1)^{n+1} \eta_n^{(n)} \end{aligned} \right\}. \quad (1)$$

Будем рассматривать (1) как систему линейных алгебраических уравнений относительно коэффициентов характеристического полинома и кратко запишем ее в виде

*Заметим, что полиномы с матричным аргументом уже рассматривались в § 5.

$Yp = (-1)^{n-1}y_n$ (матрица Y и вектор y_n известны, ищется вектор $p = (p_1, \dots, p_n)$). Пусть z_k собственные векторы матрицы A : $Az_k = \lambda_k z_k$ (напомним, что все λ_k предполагаются различными), и пусть $y_0 = \sum_{j=1}^n \alpha_j z_j$.

Теорема. Для того чтобы матрица Y системы уравнений (1) была невырожденной, необходимо и достаточно чтобы при всех j было $\alpha_j \neq 0^*$.

Доказательство. Введем обозначения для компонент собственных векторов $z_j = (\zeta_1^{(j)}, \dots, \zeta_n^{(j)})$ и определим матрицы Z , Λ и диагональную матрицу D :

$$Z = \begin{pmatrix} \zeta_1^{(1)} & \dots & \zeta_1^{(n)} \\ \dots & \dots & \dots \\ \zeta_n^{(1)} & \dots & \zeta_n^{(n)} \end{pmatrix}, \quad \Lambda = \begin{pmatrix} \lambda_1^{n-1} & \dots & 1 \\ \dots & \dots & \dots \\ \lambda_n^{n-1} & \dots & 1 \end{pmatrix}, \quad D = \begin{pmatrix} \alpha_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \alpha_n \end{pmatrix}.$$

Поскольку элемент k -й строки и j -го столбца матрицы Y есть

$$\eta_k^{(n-j)} = \sum_{\nu=1}^n \alpha_\nu \lambda_\nu^{n-j} \zeta_k^{(\nu)},$$

то $Y = ZD\Lambda$. Матрица Z невырождена, поскольку собственные векторы z_k линейно независимы, матрица Λ невырождена как матрица Вандермонда, а для того чтобы D была невырожденной, необходимо и достаточно, чтобы все коэффициенты α_j были отличны от нуля. ■

Замечание. Если среди собственных чисел матрицы A имеются равные, то $\det \Lambda = 0$, так что и $\det Y = 0$. В этом и состоит упоминавшееся в начале параграфа затруднение.

Пусть коэффициенты полинома P_n уже вычислены, и его корни λ_k — собственные числа матрицы A — найдены. Обратимся к нахождению собственных векторов. Построим полиномы

$$P_{n-1}^{(k)}(\lambda) = \frac{P_n(\lambda)}{(\lambda - \lambda_k)} = (-1)^n \lambda^{n-1} + p_1^{(k)} \lambda^{n-2} + \dots + p_{n-1}^{(k)}.$$

Очевидно, что $P_{n-1}^{(k)}(\lambda_j) = 0$ при $j \neq k$ и $P_{n-1}^{(k)}(\lambda_k) = P'_n(\lambda_k) \neq 0$ (так как λ_k — простые корни полинома P_n). Поэтому $P_{n-1}^{(k)}(A)z_j = P_{n-1}^{(k)}(\lambda_j)z_j = 0$ ($j \neq k$) и по этой причине $P_{n-1}^{(k)}(A)z_k = P'_n(\lambda_k)z_k$. Таким образом

$$P_{n-1}^{(k)}(A)y_0 = \sum_{j=1}^n \alpha_j P_{n-1}^{(k)}(A)z_j = \alpha_k P'_n(\lambda_k)z_k = \tilde{z}_k$$

есть собственный вектор матрицы A , соответствующий собственному числу λ_k . Остается еще заметить, что

$$\tilde{z}_k = (-1)^n y_{n-1} + p_1^{(k)} y_{n-2} + \dots + p_{n-1}^{(k)} y_0. \quad (2)$$

*Про такой вектор y_0 иногда говорят, что он “находится в общем положении относительно матрицы A ”.

Итак, алгоритм метода Крылова состоит в следующем:

- 1) выбирается произвольный ненулевой вектор y_0 и строятся векторы $y_k = Ay_{k-1}$, $k = 1, \dots, n$;
- 2) путем решения системы уравнений (1) находятся коэффициенты характеристического полинома $P_n(\lambda)$ матрицы A ;
- 3) собственные числа λ_k матрицы A находятся как корни полинома $P_n(\lambda)$;
- 4) вычисляются коэффициенты полиномов $P_{n-1}^{(k)}(\lambda) = P_n(\lambda)/(\lambda - \lambda_k)$;
- 5) собственные векторы \tilde{z}_k вычисляются по формулам (2).

Если определитель системы уравнений (1) оказался нулевым, то это означает, что либо у матрицы A имеется кратное собственное число, либо вектор y_0 выбран неудачно, и тогда его следует заменить.

§9 Метод Якоби

В этом параграфе матрицы будем считать вещественными.

Метод Якоби — это метод решения полной проблемы собственных чисел для симметричных матриц.

Напомним некоторые сведения из линейной алгебры. Матрица T называется ортогональной, если ее столбцы попарно ортогональны и сумма квадратов элементов каждого столбца равна единице, т.е. если* $T'T = E$. Произведение ортогональных матриц есть ортогональная матрица. Если A — симметричная матрица, а T — ортогональная, то $T'AT$ также симметричная, и собственные числа этих матриц совпадают. Симметричная матрица A может быть ортогональным преобразованием приведена к диагональному виду, т.е. существует такая ортогональная матрица T , что $T'AT = \Lambda$, где Λ — диагональная матрица. При этом диагональные элементы матрицы Λ — собственные числа матрицы A , а столбцы матрицы T — соответствующие собственные векторы.

Введем обозначения для квадратных корней из сумм квадратов всех элементов матрицы A (нормы Фробениуса) и ее диагональных элементов:

$$t^2(A) = \sum_{i=1}^n \sum_{j=1}^n a_{ij}^2, \quad d^2(A) = \sum_{i=1}^n a_{ii}^2.$$

Тогда $r^2(A) = t^2(A) - d^2(A)$ — сумма квадратов недиагональных элементов.

Лемма 1. Если A — симметричная матрица, а T — ортогональная и $B = T'AT$, то $t^2(B) = t^2(A)$.

Доказательство. Очевидно равенство† $t^2(A) = \text{tr}(A^2)$. Поэтому

$$t^2(B) = \text{tr}(B^2) = \text{tr}(T'AT T' A' T) = \text{tr}(T' A^2 T) = \text{tr}(A^2) = t^2(A) \quad \blacksquare$$

Элементарной ортогональной матрицей назовем ортогональную матрицу $T_{ij}(\varphi)$ ($i < j$), которая лишь четырьмя элементами отличается от единичной, а именно $t_{ii} = t_{jj} = \cos \varphi$, $-t_{ij} = t_{ji} = \sin \varphi$.

*Штрих означает транспонирование матрицы.

† $\text{tr}(A)$ — след матрицы A , сумма ее диагональных элементов, равная сумме собственных чисел.

Лемма 2. Пусть $B = T'_{ij}(\varphi)AT_{ij}(\varphi)$. Тогда матрица B лишь двумя столбцами и двумя строками с номерами i и j отличается от матрицы A (т.е. при $\mu \neq i, j$ и $\nu \neq i, j$ будет $b_{\mu\nu} = a_{\mu\nu}$). Кроме того $\hat{B} = T'(\varphi)\hat{A}T(\varphi)$. Здесь

$$\hat{A} = \begin{pmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{pmatrix}, \quad \hat{B} = \begin{pmatrix} b_{ii} & b_{ij} \\ b_{ji} & b_{jj} \end{pmatrix}, \quad T(\varphi) = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \quad -$$

матрицы второго порядка.

Доказательство очевидно. ■

Лемма 3. Для симметричной матрицы A и индексов $i < j$ найдется такой угол φ , что для матрицы $B = T'_{ij}(\varphi)AT_{ij}(\varphi)$ окажется $b_{ij} = b_{ji} = 0$. При таком выборе φ

$$b_{ii}^2 + b_{jj}^2 = a_{ii}^2 + a_{jj}^2 + 2a_{ij}^2. \quad (1)$$

Доказательство. Используя лемму 2, нетрудно получить, что при произвольном φ

$$b_{ij} = \frac{1}{2}(a_{jj} - a_{ii}) \sin 2\varphi + a_{ij} \cos 2\varphi.$$

Если $a_{ii} = a_{jj}$, то достаточно положить $\varphi = \pi/4$, а при $a_{ii} \neq a_{jj}$

$$\varphi = \frac{1}{2} \arctan \frac{2a_{ij}}{a_{ii} - a_{jj}}.$$

Формула (1) следует из леммы 1, примененной к матрицам \hat{A} и \hat{B} . ■

Метод Якоби состоит в построении последовательности матриц A_s ($A_0 = A$) с таким расчетом, чтобы при больших s матрицы A_s были близки к диагональным, так что их диагональные элементы близки к собственным числам матрицы A . Матрица A_{s+1} строится в виде $A_{s+1} = T'_s A_s T_s$, где $T_s = T_{ij}(\varphi)$, индексы $i = i_s$ и $j = j_s$ находятся из условия $|a_{ij}^{(s)}| = \max_{\mu \neq \nu} |a_{\mu\nu}^{(s)}|$, а $\varphi = \varphi_s$ вычисляется (согласно лемме 2) так, чтобы оказалось $a_{ij}^{(s+1)} = 0$. Напомним, что каждая следующая матрица отличается от предыдущей лишь двумя строками и столбцами.

Теорема. При $s \rightarrow \infty$ все недиагональные элементы матриц A_s стремятся к нулю.

Доказательство. Достаточно доказать, что $r^2(A_s) \rightarrow 0$. Согласно лемме 3 $t^2(A_{s+1}) = t^2(A_s)$. Все диагональные элементы матриц A_{s+1} и A_s , кроме двух, совпадают, и ввиду равенства (1) $d^2(A_{s+1}) = d^2(A_s) + 2 \left(a_{ij}^{(s)}\right)^2$, причем согласно выбору индексов i_s и j_s $\left(a_{ij}^{(s)}\right)^2 \geq r^2(A_s)/[n(n-1)]$ (максимальный из квадратов недиагональных элементов не меньше их среднего арифметического), и потому

$$d^2(A_{s+1}) \geq d^2(A_s) + 2r^2(A_s)/[n(n-1)],$$

$$r^2(A_{s+1}) = t^2(A_{s+1}) - d^2(A_{s+1}) \leq t^2(A_s) - d^2(A_s) - 2r^2(A_s)/[n(n-1)] = qr^2(A_s),$$

где $q = 1 - 2/[n(n-1)] < 1$. Методом индукции теперь легко показывается, что $r^2(A_s) \leq q^s r^2(A)$, откуда и следует, что $r^2(A_s) \rightarrow 0$. ■

Покажем, что диагональные элементы матриц A_s сходятся к собственным числам матрицы A . Пусть $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ — собственные числа матрицы A , а значит, и

A_s , и пусть $\mu_1^{(s)} \geq \mu_2^{(s)} \geq \dots \geq \mu_n^{(s)}$ — диагональные элементы матрицы A_s , расположенные в порядке убывания.

Следствие. При $s \rightarrow \infty$ выполняется соотношение $\mu_k^{(s)} \rightarrow \lambda_k$.

Доказательство. Представим матрицу A_s в виде $A_s = \Lambda_s + R_s$, где Λ_s — диагональная матрица, а у R_s на главной диагонали нули. По доказанной теореме $R_s \rightarrow 0$. Так как $\mu_k^{(s)}$ — собственные числа матрицы Λ_s , то по теореме 4 из §3 при всех k $|\lambda_k - \mu_k^{(s)}| \leq \|R_s\|_2 \leq t(R_s) = r(A_s)$. Мы воспользовались отмеченной в первом параграфе оценкой второй матричной нормы через норму Фробениуса. Остается воспользоваться доказанной теоремой, согласно которой $r(A_s) \rightarrow 0$ ■

Замечание. Из теоремы следует сходимость метода с быстротой геометрической прогрессии со знаменателем q , близким при большом n к единице. В действительности сходимость метода существенно быстрее.

Итак, за приближения к собственным числам матрицы A принимаются при большом s диагональные элементы матрицы A_s . За приближения к собственным векторам можно принять столбцы матрицы $T_1 T_2 \dots T_s$.

§10 Об ускорении сходимости

Пусть имеется сходящаяся числовая последовательность $a_s \rightarrow a^*$. Ускорением сходимости называется построение другой последовательности $\{b_s\}$, которая сходится к тому же пределу a^* , но быстрее, чем исходная. Для того, чтобы это было возможно, мы должны располагать некоторой дополнительной информацией о последовательности $\{a_s\}$. В этом параграфе будем предполагать известным, что $\{a_s\}$ сходится к пределу с быстротой общего члена геометрической прогрессии, т.е. что

$$a_s = a^* + q^s(c + \varepsilon_s), \quad (1)$$

где $c \neq 0$, $0 < |q| < 1$ и $\varepsilon_s \rightarrow 0$. Числа a^* , c и последовательность ε_s нам неизвестны (получение лучших, чем a_s , приближений к a^* и составляет цель ускорения сходимости), известен только факт существования такого представления. Что касается q , то следует различать два случая, когда q известно хотя бы приближенно, и неизвестно. Первый случай связан с методом Л.А. Люстерника ускорения сходимости метода итерации решения систем линейных уравнений.

Метод Люстерника.

Итак, нам известно, что a_s имеют представление (1), причем число q нам также известно. Из (1) следует, что $a_s - qa_{s-1} = (1-q)a^* + q^s(\varepsilon_s - \varepsilon_{s-1})$, и если мы положим $b_s = (a_s - qa_{s-1})/(1-q)$, то окажется $b_s = a^* + \frac{1}{1-q} q^s(\varepsilon_s - \varepsilon_{s-1})$, и таким образом $(b_s - a^*)/(a_s - a^*) = \frac{1}{1-q}(\varepsilon_s - \varepsilon_{s-1})/(c + \varepsilon_s) \rightarrow 0$. Это и означает, что b_s сходится к a^* быстрее, чем a_s .

Положение не изменится, если под a^* , a_s , c , ε_s мы будем понимать векторы. Тогда (1) — векторное равенство, и если b_s строить по формуле $b_s = (a_s - qa_{s-1})/(1-q)$, то т.к. $\|b_s - a^*\| \leq \frac{1}{1-q} |q|^s (\|\varepsilon_s\| + \|\varepsilon_{s-1}\|)$ и $\|a_s - a^*\| \geq |q|^s (\|c\| - \|\varepsilon_s\|)$, окажется $\|b_s - a^*\|/\|a_s - a^*\| \rightarrow 0$.

Обратимся теперь к методу итерации для решения системы линейных уравнений $x = Ax + y$: $x_{s+1} = Ax_s + y$. Сделаем предположения: 1) максимальное по моду-

лю собственное число матрицы A , имеющей полную систему собственных векторов, единственно, так что $1 > |\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|$; 2) при разложении вектора $x_0 - x^* = \sum_{j=1}^n \alpha_j z_j$ по собственным векторам матрицы A $\alpha_1 \neq 0$. Тогда

$$\begin{aligned} x_s - x^* &= A^s(x_0 - x^*) = \sum_{j=1}^n \alpha_j \lambda_j^s z_j = \lambda_1^s \left(\alpha_1 z_1 + \sum_{j=2}^n \alpha_j \left(\frac{\lambda_j}{\lambda_1} \right)^s z_j \right) = \\ &= \lambda_1^s (\alpha_1 z_1 + \varepsilon_s), \quad \|\varepsilon_s\| \rightarrow 0, \end{aligned}$$

т.е. последовательность $\{x_s\}$ имеет представление (1) при $q = \lambda_1$. В то же время

$$x_s - x_{s-1} = x_s - x^* - (x_{s-1} - x^*) = \lambda_1^{s-1} ((\lambda_1 - 1)\alpha_1 z_1 + \lambda_1 \varepsilon_s - \varepsilon_{s-1}),$$

откуда видно, что если для некоторого вектора u $(z_1, u) \neq 0$, то

$$\lambda_1^{(s)} = \frac{(x_s - x_{s-1}, u)}{(x_{s-1} - x_{s-2}, u)} \rightarrow \lambda_1,$$

и можно при больших s заменить $q = \lambda_1$ на $\lambda_1^{(s)}$ (в сущности мы применили для нахождения приближения к λ_1 степенной метод). В этом и состоит метод Люстерника, имеющий вычислительные формулы:

$$\lambda_1^{(s)} = \frac{(x_s - x_{s-1}, u)}{(x_{s-1} - x_{s-2}, u)}, \quad \tilde{x}_s = \frac{x_s - \lambda_1^{(s)} x_{s-1}}{1 - \lambda_1^{(s)}} = \frac{1}{1 - \lambda_1^{(s)}} x_s - \frac{\lambda_1^{(s)}}{1 - \lambda_1^{(s)}} x_{s-1}.$$

Вектор \tilde{x}_s , вообще говоря, ближе к x^* , чем x_s . Метод особенно эффективен, если $|\lambda_1|$ существенно больше, чем $|\lambda_2|$.

δ^2 -процесс Эйткена.

Рассмотрим второй случай, когда в представлении (1) числовой последовательности коэффициент q нам неизвестен. Пренебрегая малыми слагаемыми ε_j , будем находить b_s из системы уравнений относительно b_s , c и q

$$\left. \begin{aligned} a_{s-1} &= b_s + cq^{s-1} \\ a_s &= b_s + cq^s \\ a_{s+1} &= b_s + cq^{s+1} \end{aligned} \right\}. \quad (2)$$

Система легко решается:

$$q = \frac{a_{s+1} - a_s}{a_s - a_{s-1}}, \quad b_s = \frac{a_s - qa_{s-1}}{1 - q} = \frac{\begin{vmatrix} a_{s+1} & a_s \\ a_s & a_{s-1} \end{vmatrix}}{a_{s+1} - 2a_s + a_{s-1}}.$$

Последняя формула и применяется для вычисления b_s .

Лемма. При любом a выполняется равенство

$$b_s = a + \frac{\begin{vmatrix} a_{s+1} - a & a_s - a \\ a_s - a & a_{s-1} - a \end{vmatrix}}{a_{s+1} - 2a_s + a_{s-1}}. \quad (3)$$

Доказательство можно провести непосредственным вычислением, но оно следует также из того, что при тех же q и c число $b_s - a$ удовлетворяет равенствам (2), если из их левых частей вычесть a . ■

При вычислениях используют именно формулу (3), выбирая a близким к a_s , чтобы избежать существенного влияния ошибок округления в произведениях при вычислении определителя. В частности, если положить $a = a_s$, то

$$b_s = a_s + \frac{(a_{s+1} - a_s)(a_{s-1} - a_s)}{a_{s+1} - 2a_s + a_{s-1}}.$$

Для последовательностей рассматриваемого вида последовательность $\{b_s\}$ сходится к a^* быстрее, чем $\{a_s\}$:

Теорема. Если $0 < |q| < 1$ и $c \neq 0$, то

$$\frac{b_s - a^*}{a_s - a^*} \rightarrow 0.$$

Доказательство. Используя для b_s формулу (3) при $a = a^*$, имеем

$$\begin{aligned} b_s - a^* &= \frac{\begin{vmatrix} q^{s+1}(c + \varepsilon_{s+1}) & q^s(c + \varepsilon_s) \\ q^s(c + \varepsilon_s) & q^{s-1}(c + \varepsilon_{s-1}) \end{vmatrix}}{q^{s+1}(c + \varepsilon_{s+1}) - 2q^s(c + \varepsilon_s) + q^{s-1}(c + \varepsilon_{s-1})} = \\ &= q^s \frac{c(\varepsilon_{s+1} + \varepsilon_{s-1} - 2\varepsilon_s) + \varepsilon_{s+1}\varepsilon_{s-1} - \varepsilon_s^2}{q(c + \varepsilon_{s+1}) - 2(c + \varepsilon_s) + \frac{1}{q}(c + \varepsilon_{s-1})} = \frac{e_s}{Q_s} q^s. \end{aligned}$$

Здесь $e_s \rightarrow 0$ и $Q_s \rightarrow c \left(q - 2 + \frac{1}{q} \right) = \frac{c}{q}(1 - q)^2 \neq 0$. Поэтому

$$\frac{b_s - a^*}{a_s - a^*} = \frac{e_s}{(c + \varepsilon_s)Q_s} \rightarrow 0. \quad \blacksquare$$

Замечание. Как видно из доказательства, порядок сходимости b_s к a^* определяется формулой $b_s - a^* = \mathcal{O}(q^s(|\varepsilon_s| + |\varepsilon_{s+1}| + |\varepsilon_{s-1}|))$.

Вспомним степенной метод нахождения максимального по модулю собственного числа матрицы A . Как было показано в § 7, при соответствующих предположениях

$$\begin{aligned} \lambda_1^{(s)} &= \lambda_1 + c \left(\frac{\lambda_2}{\lambda_1} \right)^s + \mathcal{O} \left(\left(\frac{\lambda_2}{\lambda_1} \right)^{2s} + \left(\frac{\lambda_3}{\lambda_1} \right)^s \right) = \\ &= \lambda_1 + \left(\frac{\lambda_2}{\lambda_1} \right)^s \left(c + \mathcal{O} \left(\left(\frac{\lambda_2}{\lambda_1} \right)^s + \left(\frac{\lambda_3}{\lambda_2} \right)^s \right) \right). \end{aligned}$$

Итак, при тех условиях, которые указаны в § 7, последовательность $\{\lambda_1^{(s)}\}$ принадлежит к тому классу последовательностей, применение к которым δ^2 -процесса позволяет получить ускорение их сходимости. Последовательность $\{\mu_1^{(s)}\}$, полученная из $\{\lambda_1^{(s)}\}$ этим методом, сходится к λ_1 быстрее, причем, согласно замечанию,

$$\mu_1^{(s)} - \lambda_1 = \mathcal{O} \left(\left(\frac{\lambda_2}{\lambda_1} \right)^{2s} + \left(\frac{\lambda_3}{\lambda_1} \right)^s \right).$$

δ^2 -процесс применим и к методу скалярных произведений в случае самосопряженной матрицы A , а также к уточнению собственных векторов x_s , полученных степенным методом; в последнем случае речь идет о покомпонентном применении этого метода. Останавливаться на подробностях не будем.

Задача. Определить быстроту сходимости полученной с помощью δ^2 -процесса последовательности приближений к λ_1 в случае метода скалярных произведений.