ОЦЕНКИ СХОДИМОСТИ ОРТОГОНАЛЬНО-СТЕПЕННОГО МЕТОДА*)

С. К. Годунов, С. В. Кузнецов

§ 1. Постановка задачи и используемые понятия

В предлагаемой работе исследуется влияние погрешностей на результат ортогонально-степенного алгоритма, с помощью которого осуществляется ортогонально-подобное приведение квадратной матрицы A к клеточно-треугольному виду. Каждая диагональная клетка полученной матрицы должна иметь равные по модулю собственные значения. Однако ввиду ограниченной точности вычислений спектр клетки состоит из собственных значений, которые расположены в кольцевом слое $a_j \leqslant |\lambda| \leqslant b_j$, где j — номер диагональной клетки.

Число разрядов в компьютерном представлении матричных элементов конечно, что приводит к необходимости рассматривать вместо точного спектра матрицы A ее ε-спектр $\Lambda_\varepsilon(A)$, т. е. множество чисел λ, при которых имеет место неравенство $\|(A-\lambda I)^{-1}\| \geqslant 1/(\varepsilon\|A\|)$. В [1–4] приведены разнообразные примеры, показывающие, что ε-спектр $\Lambda_\varepsilon(A)$ может распадаться на «пятна» совсем не малых размеров даже при малых ε. Каждая точка, принадлежащая «пятну», является собственным значением той или иной матрицы, отличающейся от матрицы A по норме не более чем на величину $\varepsilon\|A\|$. Ясно, что все точки, входящие в одно из таких «пятен», принадлежат одной и той же диагональной клетке. Пусть окружность $|\lambda| = \rho\|A\|$ разделяет спектр так, что норма $\|(A-\lambda I)^{-1}\|$ не слишком велика на этой окружности. Тогда собственные числа, отвечающие точкам, лежащим внутри и вне круга, ограниченного окружностью $|\lambda| = \rho\|A\|$, окажутся в разных диагональных клетках. В качестве критерия дихотомии спектра окружности (т. е. числовой характеристики удаленности окружности $|\lambda| = \rho\|A\|$ от точек спектра) мы выбираем величину $\omega(\rho) = \|H\|$ (см. [5, 6]), где

$$H = \frac{\rho^2\|A\|^2}{2\pi}\int_0^{2\pi}(A^* - \rho\|A\|e^{-i\theta}I)^{-1}(A-\rho\|A\|e^{i\theta}I)^{-1}\,d\theta.$$

Очевидно, что значение $\omega(\rho)$ очень велико, если окружность $|\lambda| = \rho\|A\|$ пересекается с каким-либо «пятном» ε-спектра.

Как показано в [7], точки спектра, разделенные окружностью $|\lambda| = \rho\|A\|$, где ρ такое, что $\omega(\rho)$ не слишком большое, окажутся после работы

---

*) Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (код проекта 93-011-1515).

алгоритма в различных диагональных клетках, если исследуемая матрица A записана в подходящем базисе. В [7] не учитывались погрешности, неизбежные при реализации алгоритма. Основная цель настоящей работы — оценить погрешность приведения матрицы A к клеточно-треугольному виду с учетом влияния погрешностей, возникающих при реализации алгоритма. Этому посвящены § 2–4. В § 6 доказываются неравенства, позволяющие использовать оценки сходимости ортогонально-степенного метода для анализа влияния погрешностей на результат QR-алгоритма (см. [8–11]).

Так же, как в [7], здесь используются критерий дихотомии спектра окружности $\omega(\rho)$ и параметр $l(\rho)$, характеризующий расположение координатного базиса относительно инвариантных для матрицы A подпространств. Пусть окружность $|\lambda| = \rho\|A\|$ с не слишком большим $\omega(\rho)$ делит спектр матрицы A так, что N_0 собственных значений расположены в круге, ограниченном этой окружностью, а остальные $N_\infty = N - N_0$ точек спектра лежат вне круга. Для того чтобы после применения ортогонально-степенного алгоритма собственные значения λ, удовлетворяющие условиям $|\lambda| < \rho\|A\|$ и $|\lambda| > \rho\|A\|$, оказались в разных клетках, необходимо выполнение следующего условия: линейная оболочка первых N_∞ векторов координатного базиса, используемого для записи матрицы A, не содержит векторов, коллинеарных (или «почти коллинеарных») векторам подпространства, инвариантного для A и отвечающего всем точкам спектра, заключенным внутри круга. Степень «неколлинеарности» мы характеризуем введенным в [7] параметром $l(\rho)$, который не должен быть слишком большим. Параметр $l(\rho)$ определяется на основе следующего канонического представления матрицы A:

$$A = U^* \begin{bmatrix} I_{N_\infty} & -L \\ 0 & I_{N_0} \end{bmatrix} \begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix} \begin{bmatrix} I_{N_\infty} & L \\ 0 & I_{N_0} \end{bmatrix} U, \tag{1.1}$$

где матрица U унитарная, спектр клетки B (размеров $N_\infty \times N_\infty$) лежит вне круга, т. е. $|\lambda| > \rho\|A\|$, а спектр клетки C (размеров $N_0 \times N_0$) — внутри круга, т. е. $|\lambda| < \rho\|A\|$. Определим матрицы X и Y размеров $N_\infty \times N_\infty$ и $N_0 \times N_\infty$ соответственно из равенства $U \begin{bmatrix} I_{N_\infty} \\ 0 \end{bmatrix} = \begin{bmatrix} X \\ Y \end{bmatrix}$ и положим $l(\rho) = \|Y(X + LY)^{-1}\|$ (если матрица A имеет собственное значение λ, лежащее на окружности $|\lambda| = \rho\|A\|$, то параметр $l(\rho)$ не определен). Заметим, что параметр $l(\rho)$ конечен, если и только если матрица $X + LY$ не вырождена.

Нетрудно показать, что $\|L\| = \operatorname{ctg}\varphi$, где φ — угол раствора между инвариантными подпространствами, отвечающими собственным значениям, лежащим вне и внутри круга $|\lambda| = \rho\|A\|$. Угол раствора, как обычно, определяется из следующего условия: число $2\sin(\varphi/2)$ равно наибольшему из расстояний между концами единичных векторов одного подпространства и ближайших к ним единичных векторов другого подпространства. Параметр $l(\rho)$ можно связать с некоторым угловым расстоянием между линейной оболочкой первых N_∞ векторов координатного базиса и инвариантным подпространством для матрицы A, отвечающим собственным значениям, бо́льшим $\rho\|A\|$ по модулю. Именно: $l(\rho) = \operatorname{tg}\alpha$, где α — угол раствора, измеренный в специальной евклидовой метрике, отличной от исходной. Эта метрика определяется с помощью скалярного произведения $[\xi, \eta]$, связанного с каноническим видом матрицы A и вычисляемого по формуле

$$[\xi, \eta] = (x_{N_\infty}, y_{N_\infty}) + (Lx_{N_0}, y_{N_\infty}) + (x_{N_\infty}, Ly_{N_0}) + (x_{N_0}, [I + L^* L] y_{N_0}),$$

в которой векторы x_{N_∞}, y_{N_∞} размерности n_∞ и векторы x_{N_0}, y_{N_0} размерности N_0 определяются в соответствии с равенствами

$$\xi = U \begin{bmatrix} x_{N_\infty} \\ y_{N_0} \end{bmatrix}, \qquad \eta = U \begin{bmatrix} x_{N_\infty} \\ y_{N_0} \end{bmatrix}.$$

В такой метрике инвариантные подпространства матрицы A, отвечающие точкам спектра внутри и вне круга с радиусом $\rho\|A\|$, будут ортогональны.

Нам удобно считать, что речь идет о разделении спектра единичной окружностью $|\lambda| = \rho\|A\| = 1$. Это предположение не ограничивает общности, как следует из замечаний, сделанных ниже. При выбранной нормировке имеем

$$\omega = \omega(1/\|A\|) = \|H\|,$$
$$H = \frac{1}{2\pi} \int_0^{2\pi} (A^* - e^{-i\theta} I)^{-1} (A - e^{i\theta} I)^{-1} \, d\theta.$$

В [6] при помощи канонического представления (1.1) изучалась последовательность матриц G_m $(-\infty < m < +\infty)$, однозначно определяемых условиями

$$\|G_m\| \to 0 \quad (m \to \pm\infty),$$
$$G_0 = G_{-1} + I, \quad G_m = AG_{m-1} \quad (m \neq 0).$$

Было получено представление

$$G_m = \begin{cases} -U^* \begin{bmatrix} B^m & B^m L \\ 0 & 0 \end{bmatrix} U & \text{при} \quad m \leqslant -1, \\[2ex] U^* \begin{bmatrix} 0 & -LC^m \\ 0 & C^m \end{bmatrix} U & \text{при} \quad m \geqslant 0 \end{cases}$$

и выведена оценка

$$\|G_m\| < \begin{cases} \sqrt{\omega} \left(\frac{\omega}{1+\omega} \right)^{\frac{m}{2}} & \text{при} \quad m \geqslant 0, \\[2ex] \sqrt{\omega} \left(\frac{\omega}{1+\omega} \right)^{\frac{|m|-1}{2}} & \text{при} \quad m < 0. \end{cases} \tag{1.2}$$

В нашем исследовании существенно используются вытекающие из оценок (1.2) неравенства

$$\begin{aligned} &\|B^{-p}\| \leqslant \sqrt{\omega} \left(\tfrac{\omega}{1+\omega} \right)^{\frac{p-1}{2}} && (p = 1, 2, 3, \ldots), \\ &\|C^p\| \leqslant \left\| \begin{bmatrix} -LC^p \\ C^p \end{bmatrix} \right\| \leqslant \sqrt{\omega} \left(\tfrac{\omega}{1+\omega} \right)^{\frac{p}{2}} && (p = 0, 1, 2, \ldots), \end{aligned} \tag{1.3}$$

из которых, в частности, следуют неравенства

$$|\lambda_j(B)| \geqslant \sqrt{\frac{1+\omega}{\omega}} > 1 \qquad (1 \leqslant j \leqslant N_\infty),$$

$$|\lambda_j(C)| \leqslant \sqrt{\frac{\omega}{1+\omega}} < 1 \qquad (1 \leqslant j \leqslant N_0).$$

Иными словами,

$$\begin{aligned} |\lambda_N(A)| &= |\lambda_{N_\infty}(B)| \geqslant |\lambda_{N-1}(A)| \geqslant \cdots \geqslant |\lambda_{N_0+1}(A)| \\ &= |\lambda_1(B)| > 1 > |\lambda_{N_0}(A)| \\ &= |\lambda_{N_0}(C)| \geqslant |\lambda_{N_0-1}(A)| \geqslant \cdots \geqslant |\lambda_1(A)| \geqslant 0. \end{aligned}$$

Мы будем также пользоваться вытекающими из (1.2) оценками

$$\sqrt{1+\|L\|^2} \leqslant \sqrt{\omega}, \qquad 1+\|L\| \leqslant \sqrt{2\omega}. \tag{1.4}$$

Напомним, что ортогонально-степенной алгоритм осуществляется по схеме

$$C_0 = I, \qquad AC_{j-1} = C_j R_j, \tag{1.5}$$

где $C_j^* C_j = I$, а R_j $(j = 1, 2 \ldots)$ — верхние треугольные матрицы.

Если матрицу A умножить на положительный скаляр, то это не изменит определяемых ортогональных сомножителей C_j, а все R_j умножатся на тот же множитель, что и матрица A. Ввиду того, что сходимость алгоритма обеспечивается поведением последовательности $\{C_j\}$, выбор нормировки не влияет на сходимость. Поэтому мы ограничимся рассмотрением окружности $|\lambda| = 1$, разрешая принимать норме $\|A\|$ любые положительные значения и полагая $\rho = 1/\|A\|$.

Для учета неизбежных при реальном осуществлении алгоритма вычислительных погрешностей, следует модифицировать соотношения (1.5) и заменить их равенствами

$$C_0 = I, \qquad (A + \Delta_j)C_{j-1} = C_j[\widetilde{R}_j + \widetilde{\nabla}_j], \qquad C_j^* C_j = I + \widetilde{E}_j, \tag{1.6}$$

в которых

$$\|\Delta_j\| \leqslant \varepsilon\|A\|, \quad \|\widetilde{\nabla}_j\| \leqslant \varepsilon\|A\|, \quad \|\widetilde{E}_j\| \leqslant \varepsilon \quad (\varepsilon \ll 1). \tag{1.7}$$

Очевидно, что $\widetilde{E}_j^* = \widetilde{E}_j$, поэтому

$$I + \widetilde{E}_j = (I + \widehat{E}_j^*)(I + \widehat{E}_j) \qquad (\,\|\widehat{E}_j\| \leqslant \varepsilon/[2(1-\varepsilon)]\,). \tag{1.8}$$

При этом можно предположить, что $\widehat{E}_j$ — верхняя треугольная матрица. Обозначим через $\widehat{C}_j$ ортогональную матрицу вида

$$\widehat{C}_j = C_j(I + \widehat{E}_j)^{-1}, \quad \widehat{C}_j^* \widehat{C}_j = I.$$

Заметим, что

$$\begin{aligned} \|C_j^* A C_j - \widehat{C}_j^* A \widehat{C}_j\| &= \|(I + \widehat{E}_j^*)\widehat{C}_j^* A \widehat{C}_j (I + \widehat{E}_j) - \widehat{C}_j^* A \widehat{C}_j\| \\ &\leqslant (2\|\widehat{E}_j\| + \|\widehat{E}_j\|^2)\,\|\widehat{C}_j^* A \widehat{C}_j\| \\ &\leqslant \left(\frac{\varepsilon}{1-\varepsilon} + \frac{\varepsilon^2}{4(1-\varepsilon)^2}\right)\|A\| \leqslant \frac{\varepsilon}{(1-\varepsilon)^2}\|A\|. \end{aligned} \tag{1.9}$$

Векторы, компоненты которых образуют первые N_∞ столбцов матрицы C_j, можно рассматривать как базисные векторы некоторого подпространства, которое в идеальном случае (при нулевых погрешностях) при $j \to \infty$ стремится к инвариантному для A подпространству, отвечающему всем собственным числам, бо́льшим единицы по модулю (бо́льшим $\rho\|A\|$). Мы будем рассматривать только эти N_∞ столбцов матриц C_j и исследовать влияние погрешностей на поведение при $j \to \infty$ подпространства, порожденного указанными столбцами. Исследование основано на глубокой аналогии QR-алгоритма с методом ортогональной прогонки (см. [12]) решения хорошо обусловленных краевых задач для дифференциальных или разностных уравнений. Эта аналогия привела нас к идее использовать вспомогательные дискретные краевые задачи при анализе накопления погрешностей (см. § 3).

Отметим также, что поскольку матрицы $I + \widehat{E}_j$, $(I + \widehat{E}_j)^{-1}$ верхние треугольные, линейные оболочки первых N_∞ столбцов матриц C_j, $\widehat{C}_j$ совпадают. Матрицу, образованную первыми N_∞ столбцами матрицы C_j, разобьем на клетки P_j, Q_j размеров $N_\infty \times N_\infty$ и $N_0 \times N_\infty$ так, что

$$C_j = \begin{bmatrix} P_j & H_j \\ Q_j & F_j \end{bmatrix}. \tag{1.10}$$

Аналогично представим матрицы $\widehat{C}_j$, $\widetilde{E}_j$, $\widetilde{\nabla}_j$ и $\widetilde{R}_j$ в виде

$$\widehat{C}_j = \begin{bmatrix} \widehat{P}_j & \dots \\ \widehat{Q}_j & \dots \end{bmatrix}, \qquad \widetilde{E}_j = \begin{bmatrix} E_{11}^{(j)} & E_{12}^{(j)} \\ E_{21}^{(j)} & E_{22}^{(j)} \end{bmatrix},$$
$$\widetilde{\nabla}_j = \begin{bmatrix} \nabla_{11}^{(j)} & \nabla_{12}^{(j)} \\ \nabla_{21}^{(j)} & \nabla_{22}^{(j)} \end{bmatrix}, \qquad \widetilde{R}_j = \begin{bmatrix} R_j^{(\infty)} & \dots \\ 0 & R_j^{(0)} \end{bmatrix}. \tag{1.11}$$

Из формул (1.8), (1.10) и (1.11) следует равенство

$$\begin{bmatrix} \widehat{P}_j \\ \widehat{Q}_j \end{bmatrix} = \begin{bmatrix} P_j \\ Q_j \end{bmatrix} \left[I_{N_\infty} + \widehat{E}_j^{(\infty)} \right]^{-1}, \tag{1.12}$$

где $\widehat{E}_j^{(\infty)}$ — верхний главный минор размеров $N_\infty \times N_\infty$ треугольной матрицы $\widehat{E}_j$. Поскольку столбцы матрицы $\widehat{C}_j$ ортонормированы, а столбцы матрицы C_j приближенно ортонормированы, имеем

$$\begin{gathered} \widehat{P}_j^* \widehat{P}_j + \widehat{Q}_j^* \widehat{Q}_j = I_{N_\infty}, \\ P_j^* P_j + Q_j^* Q_j = I_{N_\infty} + E_{11}^{(j)}, \\ P_j^* H_j + Q_j^* F_j = E_{12}^{(j)}. \end{gathered} \tag{1.13}$$

Используя соотношения (1.6) и разложения (1.11) матриц, получаем

$$[A + \Delta_j] \begin{bmatrix} P_{j-1} \\ Q_{j-1} \end{bmatrix} = \begin{bmatrix} P_j \\ Q_j \end{bmatrix} \left[R_j^{(\infty)} + \nabla_{11}^{(j)} \right] + \begin{bmatrix} H_j \\ F_j \end{bmatrix} \nabla_{21}^{(j)}, \tag{1.14}$$

где $P_0 = I_{N_\infty}$, $Q_0 = 0$. Отметим, что справедливо равенство

$$\begin{bmatrix} H_j \\ F_j \end{bmatrix} \nabla_{21}^{(j)} = \begin{bmatrix} P_j \\ Q_j \end{bmatrix} \nabla_j^{(\infty)}, \tag{1.15}$$

где $\nabla_j^{(\infty)}$ определяется по формуле

$$\nabla_j^{(\infty)} = \left(I_{N_\infty} + E_{11}^{(j)}\right)^{-1} E_{12}^{(j)} \nabla_{21}^{(j)}. \tag{1.16}$$

Действительно, умножая правую и левую части (1.15) на матрицу $[P_j^* Q_j^*]$ и учитывая (1.13), приходим к равенству (1.16). Матрица $I_{N_\infty} + E_{11}^{(j)}$ (см. выше) обратима вследствие неравенства (1.7), причем при $\varepsilon < 1$ справедливы оценки

$$\|(I_{N_\infty} + E_{11}^{(j)})^{-1}\| \leqslant 1/(1-\varepsilon), \qquad \|E_{11}^{(j)}\| \leqslant \|\widetilde{E}_j\| \leqslant \varepsilon. \tag{1.17}$$

Из (1.17) при $\varepsilon \leqslant 1/2$ легко получить оценку

$$\|\nabla_j^{(\infty)}\| \leqslant (\varepsilon^2/(1-\varepsilon))\|A\| \leqslant \varepsilon\, \|A\|.$$

Таким образом, в силу (1.15) равенство (1.14) можно переписать в виде

$$[A + \Delta_j] \begin{bmatrix} P_{j-1} \\ Q_{j-1} \end{bmatrix} = \begin{bmatrix} P_j \\ Q_j \end{bmatrix} \left[R_j^{(\infty)} + \nabla_{11}^{(j)} + \nabla_j^{(\infty)}\right].$$

Далее мы используем следующие обозначения:

$$R_j = R_j^{(\infty)}, \quad \Delta_j^{(1)} = \Delta_j, \quad \Delta_j^{(2)} = \nabla_{11}^{(j)} + \nabla_j^{(\infty)}, \quad \Delta_j^{(3)} = E_{11}^{(j)}.$$

Кроме того, переобозначим N_∞ через k. В принятых обозначениях ортогонально-степенной алгоритм можно записать в виде следующих соотношений:

$$\begin{gathered} \begin{bmatrix} P_0 \\ Q_0 \end{bmatrix} = \begin{bmatrix} I_k \\ 0 \end{bmatrix}, \quad \left(A + \Delta_j^{(1)}\right) \begin{bmatrix} P_{j-1} \\ Q_{j-1} \end{bmatrix} = \begin{bmatrix} P_j \\ Q_j \end{bmatrix} \left[R_j + \Delta_j^{(2)}\right], \\ P_j^* P_j + Q_j^* Q_j = I + \Delta_j^{(3)}. \end{gathered} \tag{1.18}$$

Несложно проверить, что

$$\begin{aligned} &\|\Delta_j^{(1)}\| \leqslant \varepsilon\|A\|, \\ &\|\Delta_j^{(2)}\| \leqslant \|\nabla_{11}^{(j)}\| + \|\nabla_j^{(\infty)}\| \leqslant 2\varepsilon\|A\|, \\ &\|\Delta_j^{(3)}\| = \|E_{11}^{(j)}\| \leqslant \varepsilon. \end{aligned} \tag{1.19}$$

Так как параметр $l(\rho)$ предполагается конечным, ранги матриц

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} \overline{P}_0 \\ \overline{Q}_0 \end{bmatrix} = U \begin{bmatrix} P_0 \\ Q_0 \end{bmatrix}, \qquad \overline{P}_0 + L\overline{Q}_0$$

равны $k = N_\infty$, т. е. матрица $\overline{P}_0 + L\overline{Q}_0$ не вырождена. Только при соблюдении этого условия формула

$$l(\rho) = \|\overline{Q}_0(\overline{P}_0 + L\overline{Q}_0)^{-1}\| = \|Y(X+LY)^{-1}\| \tag{1.20}$$

имеет смысл. Ниже, в § 2, полагая

$$\begin{bmatrix} \overline{P}_j \\ \overline{Q}_j \end{bmatrix} = U \begin{bmatrix} P_j \\ Q_j \end{bmatrix}, \tag{1.21}$$

мы оценим $l_j(\rho) = \|\overline{Q}_j(\overline{P}_j + L\overline{Q}_j)^{-1}\|$ и $l_0(\rho) = l(\rho)$. Если в (1.21) заменим P_j, Q_j на $\widehat{P}_j$, $\widehat{Q}_j$ и $\overline{P}_j$, $\overline{Q}_j$ на $\overline{\overline{P}}_j$, $\overline{\overline{Q}}_j$, т. е.

$$\begin{bmatrix} \overline{\overline{P}}_j \\ \overline{\overline{Q}}_j \end{bmatrix} = U \begin{bmatrix} \widehat{P}_j \\ \widehat{Q}_j \end{bmatrix}, \tag{1.22}$$

и вычислим $\widehat{l}_j(\rho) = \|\overline{\overline{Q}}_j(\overline{\overline{P}}_j + L\overline{\overline{Q}}_j)^{-1}\|$ вместо $l_j(\rho)$, то получим то же самое числовое значение $\widehat{l}_j(\rho) = l_j(\rho)$. Это утверждение справедливо в силу равенства $\overline{Q}_j(\overline{P}_j + L\overline{Q}_j)^{-1} = \overline{\overline{Q}}_j(\overline{\overline{P}}_j + L\overline{\overline{Q}}_j)^{-1}$, вытекающего из (1.12), (1.21) и (1.22).

§ 2. Ключевые оценки влияния вычислительных погрешностей

Приступим к детальному исследованию ортогонально-степенного метода. В § 1 было показано, что при учете вычислительных погрешностей определяющие соотношение (1.6), (1.7) имеют вид (ср. (1.18))

$$\begin{bmatrix} P_0 \\ Q_0 \end{bmatrix} = \begin{bmatrix} I_k \\ 0 \end{bmatrix}, \quad \left(A + \Delta_j^{(1)}\right) \begin{bmatrix} P_{j-1} \\ Q_{j-1} \end{bmatrix} = \begin{bmatrix} P_j \\ Q_j \end{bmatrix} \left[R_j + \Delta_j^{(2)}\right],$$
$$P_j^* P_j + Q_j^* Q_j = I + \Delta_j^{(3)},$$

где R_j $(j = 1, 2, \dots)$ — верхние треугольные матрицы. Мы предполагаем, что погрешность вычисления мала, т. е. имеют место неравенства (ср. (1.19))

$$\begin{aligned} &\|\Delta_j^{(1)}\| \leqslant \varepsilon\|A\|, \\ &\|\Delta_j^{(2)}\| \leqslant \|\nabla_{11}^{(j)}\| + \|\nabla_j^{(\infty)}\| \leqslant 2\varepsilon\|A\|, \\ &\|\Delta_j^{(3)}\| = \|E_{11}^{(j)}\| \leqslant \varepsilon. \end{aligned} \tag{2.1}$$

Используя каноническое представление (1.1) матрицы A, можно представить $A + \Delta_j^{(1)}$ в виде

$$A + \Delta_j^{(1)} = U^* \begin{bmatrix} I_k & -L \\ 0 & I_{N-k} \end{bmatrix} \begin{bmatrix} B + \Delta_{11}^{(j)} & \Delta_{12}^{(j)} \\ \Delta_{21}^{(j)} & C + \Delta_{22}^{(j)} \end{bmatrix} \begin{bmatrix} I_k & L \\ 0 & I_{N-k} \end{bmatrix} U,$$

где
$$\begin{bmatrix} \Delta_{11}^{(j)} & \Delta_{12}^{(j)} \\ \Delta_{21}^{(j)} & \Delta_{22}^{(j)} \end{bmatrix} = \begin{bmatrix} I_k & L \\ 0 & I_{N-k} \end{bmatrix} U\Delta_j^{(1)}U^* \begin{bmatrix} I_k & -L \\ 0 & I_{N-k} \end{bmatrix}.$$
Очевидны следующие оценки (см. (1.4), (2.1)):
$$\|\Delta_{pq}^{(j)}\| \leqslant \left\| \begin{bmatrix} I & L \\ 0 & I \end{bmatrix} \right\| \|\Delta_j^{(1)}\| \left\| \begin{bmatrix} I & -L \\ 0 & I \end{bmatrix} \right\| \leqslant (1+\|L\|)^2 \|\Delta_j^{(1)}\|$$
$$\leqslant \varepsilon(1+\|L\|)^2\|A\| \leqslant 2\varepsilon\omega\|A\| = [2\varepsilon\omega(\rho)]/\rho = \delta/2. \tag{2.2}$$
Здесь опущены индексы, указывающие размеры единичных клеток. Обозначим через $\overline{P}_j$, $\overline{Q}_j$ матрицы, определенные при помощи равенства
$$\begin{bmatrix} \overline{P}_j \\ \overline{Q}_j \end{bmatrix} = U \begin{bmatrix} P_j \\ Q_j \end{bmatrix}. \tag{2.3}$$
Ввиду ортогональности матрицы U ($U^*U = I$) имеем
$$\overline{P}_j^*\overline{P}_j + \overline{Q}_j^*\overline{Q}_j = P_j^*P_j + Q_j^*Q_j = I + \Delta_j^{(3)}. \tag{2.4}$$
Условие (2.4) существенно используется в §3. Для $\overline{P}_j$, $\overline{Q}_j$ справедливы соотношения
$$\begin{bmatrix} B+\Delta_{11}^{(j)} & \Delta_{12}^{(j)} \\ \Delta_{21}^{(j)} & C+\Delta_{22}^{(j)} \end{bmatrix} \begin{bmatrix} I & L \\ 0 & I \end{bmatrix} \begin{bmatrix} \overline{P}_{j-1} \\ \overline{Q}_{j-1} \end{bmatrix} = \begin{bmatrix} I & L \\ 0 & I \end{bmatrix} \begin{bmatrix} \overline{P}_j \\ \overline{Q}_j \end{bmatrix} \left[R_j + \Delta_j^{(2)}\right]. \tag{2.5}$$

Лемма 2.1. *Если матрица $\overline{P}_m + L\overline{Q}_m$ не вырождена (следовательно, определено конечное значение параметра $l_m = \|\overline{Q}_m(\overline{P}_m + L\overline{Q}_m)^{-1}\|$) и справедливы оценки*
$$l_m \leqslant 2(\omega l_0 + \sqrt{\omega}), \tag{2.6}$$
$$4\varepsilon\omega/\rho = \delta < 1/[\sqrt{\omega} + \omega + 2\omega l_0\sqrt{\omega}], \tag{2.7}$$
то матрицы $\overline{P}_{m+1} + L\overline{Q}_{m+1}$, R_{m+1}, $\widetilde{R}_{m+1} = R_{m+1} + \Delta_{m+1}^{(2)}$ также не вырождены.

Доказательство. Воспользуемся вытекающим из (2.5) равенством
$$(\overline{P}_{m+1} + L\overline{Q}_{m+1})\widetilde{R}_{m+1}(\overline{P}_m + L\overline{Q}_m)^{-1}$$
$$= B\left[I + B^{-1}\Delta_{11}^{(m+1)} + B^{-1}\Delta_{12}^{(m+1)}\overline{Q}_m(\overline{P}_m + L\overline{Q}_m)^{-1}\right], \tag{2.8}$$
невырожденностью матрицы B и оценкой
$$\|B^{-1}\Delta_{11}^{(m+1)} + B^{-1}\Delta_{12}^{(m+1)}\overline{Q}_m(\overline{P}_m + L\overline{Q}_m)^{-1}\|$$
$$\leqslant \tfrac{\delta}{2}(1+l_m)\|B^{-1}\| \leqslant \tfrac{\delta}{2}[1 + 2(\omega l_0 + \sqrt{\omega})]\sqrt{\omega} < \tfrac{1}{2}, \tag{2.9}$$
в которой использованы формулы (1.3), (2.2), (2.6), (2.7). Оценки (1.3), (2.9) показывают, что правая часть в (2.8) не вырождена и, следовательно, не вырождены все матричные сомножители в левой части, в том числе матрицы $\overline{P}_{m+1} + L\overline{Q}_{m+1}$, $\widetilde{R}_{m+1} = R_{m+1} + \Delta_{m+1}^{(2)}$. Лемма 2.1 доказана.

Лемма 2.2. *Если матрицы $\overline{P}_m + L\overline{Q}_m$, $\widetilde{R}_m$, $\widetilde{R}_{m-1}, \dots, \widetilde{R}_1$ не вырождены и верно соотношение*

$$4\varepsilon\omega/\rho = \delta < 1/[4(1+\omega)(\omega l_0 + \sqrt{\omega})], \tag{2.10}$$

то справедлива оценка

$$l_m = \|\overline{Q}_m(\overline{P}_m + L\overline{Q}_m)^{-1}\| \leqslant 2(\omega l_0 + \sqrt{\omega}).$$

Доказательство. Используя обозначения

$$\begin{gathered}\widetilde{R}_r = R_r + \Delta_r^{(2)}, \quad T_p = \widetilde{R}_{p+1}^{-1}\widetilde{R}_{p+2}^{-1}\cdots\widetilde{R}_m^{-1}(\overline{P}_m + L\overline{Q}_m)^{-1}, \quad Y_q = \overline{Q}_q T_q, \\ Z_q = (\overline{P}_q + L\overline{Q}_q)T_q, \quad \mathcal{L} = \overline{Q}_0(\overline{P}_0 + L\overline{Q}_0)^{-1} = Y_0 Z_0^{-1} \\ (0 \leqslant p \leqslant m-1,\ 0 \leqslant q \leqslant m,\ 0 \leqslant r \leqslant m),\end{gathered} \tag{2.11}$$

можно переписать соотношения (2.5) следующим образом:

$$\begin{gathered}\begin{bmatrix} B + \Delta_{11}^{(r)} & \Delta_{12}^{(r)} \\ \Delta_{21}^{(r)} & C + \Delta_{22}^{(r)} \end{bmatrix} \begin{bmatrix} \overline{P}_{r-1} + L\overline{Q}_{r-1} \\ \overline{Q}_{r-1} \end{bmatrix} \widetilde{R}_r^{-1} = \begin{bmatrix} \overline{P}_r + L\overline{Q}_r \\ \overline{Q}_r \end{bmatrix}, \\ \begin{bmatrix} B + \Delta_{11}^{(r)} & \Delta_{12}^{(r)} \\ \Delta_{21}^{(r)} & C + \Delta_{22}^{(r)} \end{bmatrix} \begin{bmatrix} \overline{P}_{r-1} + L\overline{Q}_{r-1} \\ \overline{Q}_{r-1} \end{bmatrix} T_{r-1} = \begin{bmatrix} \overline{P}_r + L\overline{Q}_r \\ \overline{Q}_r \end{bmatrix} T_r, \\ \begin{bmatrix} B + \Delta_{11}^{(r)} & \Delta_{12}^{(r)} \\ \Delta_{21}^{(r)} & C + \Delta_{22}^{(r)} \end{bmatrix} \begin{bmatrix} Z_{r-1} \\ Y_{r-1} \end{bmatrix} = \begin{bmatrix} Z_r \\ Y_r \end{bmatrix}.\end{gathered} \tag{2.12}$$

Очевидно (см. (2.11)), что

$$Z_m = I, \qquad Y_0 = \mathcal{L}Z_0, \qquad l_0 = \|\mathcal{L}\|. \tag{2.13}$$

Для упрощения записи в (2.12) положим

$$\mathcal{D}_{r-1} = \Delta_{11}^{(r)} Z_{r-1} + \Delta_{12}^{(r)} Y_{r-1}, \quad \mathcal{E}_{r-1} = \Delta_{21}^{(r)} Z_{r-1} + \Delta_{22}^{(r)} Y_{r-1}$$

и воспользуемся вытекающими из (2.2) соотношениями

$$\begin{aligned}\|\mathcal{D}_{r-1}\| &\leqslant [(4\varepsilon\omega(\rho))/\rho]\max\{\|Y_{r-1}\|, \|Z_{r-1}\|\} = \delta\max\{\|Y_{r-1}\|, \|Z_{r-1}\|\}, \\ \|\mathcal{E}_{r-1}\| &\leqslant [(4\varepsilon\omega(\rho))/\rho]\max\{\|Y_{r-1}\|, \|Z_{r-1}\|\} = \delta\max\{\|Y_{r-1}\|, \|Z_{r-1}\|\}.\end{aligned}$$

Очевидно, что матрицы $Y_0, Z_0, Y_1, Z_1, \dots, Y_m, Z_m$ могут рассматриваться как решение следующей дискретной краевой задачи:

$$\begin{aligned} &Y_0 = \mathcal{L}Z_0, \qquad Z_r = BZ_{r-1} + \mathcal{D}_r, \\ &Z_m = I, \qquad\ \ Y_r = CY_{r-1} + \mathcal{E}_r, \end{aligned} \qquad 0 \leqslant r \leqslant m. \tag{2.14}$$

В § 3 будет показано (лемма 3.3), что при выполнении условий (1.3), (2.10) решение этой краевой задачи удовлетворяет оценке (см. (3.9))

$$\|Y_m\| \leqslant \max_{0 \leqslant r \leqslant m}\{\|Y_r\|, \|Z_r\|\} \leqslant 2(\omega l_0 + \sqrt{\omega}).$$

Отметим, что (см. формулу (2.11) и лемму 2.1)

$$Y_m = \overline{Q}_m(\overline{P}_m + L\overline{Q}_m)^{-1}, \qquad \|Y_m\| = l_m.$$

Теперь воспользуемся методом индукции. Предположим, что δ одновременно удовлетворяет неравенствам (2.7) и (2.10). Поскольку $\omega \geqslant 1$, в силу невырожденности матрицы $\overline{P}_0 + L\overline{Q}_0$ (см. § 1) и тривиальности неравенства $l_0 \leqslant 2(\omega l_0 + \sqrt{\omega})$ на основе лемм 2.1 и 2.2 устанавливаем последовательно для $m = 0, 1, \ldots$ неравенство $l_m \leqslant 2(\omega l_0 + \sqrt{\omega})$. Одновременно проверяем невырожденность матриц $\overline{P}_m + LQ_m$, $\widetilde{R}_m$ для всех m.

Отметим, что при тех же предположениях, что и выше, для величины $l_m = \|\overline{Q}_m(\overline{P}_m + L\overline{Q}_m)^{-1}\|$ можно вывести оценку более сильную, чем (2.6). Действительно, воспользуемся равенством $l_m = \|Y_m\|$, где Y_m — матрица, входящая в решение дискретной краевой задачи (2.14). Используя установленную ниже оценку (3.10) этого решения, получаем

$$l_m \leqslant l_0\sqrt{1+\omega}[\omega/(1+\omega)]^m + 4\delta(1+\omega)(\omega l_0 + \sqrt{\omega})^2. \tag{2.15}$$

Таким образом, выразив δ через ω с помощью (2.2), сформулируем основной результат этого параграфа в виде следующей теоремы.

Теорема 2.1. *Пусть справедливы формулы* (1.1), (1.3), (1.4), (2.1). *Тогда при*

$$\varepsilon < \min\{\rho/[4\omega\sqrt{\omega}(1 + 2\sqrt{\omega} + \omega l_0)], \rho/[16\omega(1+\omega)(\omega l_0 + \sqrt{\omega})]\} \tag{2.16}$$

имеет место оценка

$$l_m = \|\overline{Q}_m(\overline{P}_m + L\overline{Q}_m)^{-1}\| \leqslant l_0\sqrt{1+\omega}[\omega/(1+\omega)]^m + \frac{16\varepsilon}{\rho}\omega(1+\omega)(\omega l_0 + \sqrt{\omega})^2,$$

где матрицы $\overline{P}_m$, $\overline{Q}_m$ определены равенством (2.3). *Кроме того, матрицы $R_1, R_2, \ldots, R_m$ не вырождены.*

Замечание 2.1. Поскольку $\omega \geqslant 1$, легко убедиться в справедливости неравенства

$$4\omega\sqrt{\omega}(1 + 2\sqrt{\omega} + \omega l_0) < 16\omega(1+\omega)(\omega l_0 + \sqrt{\omega}) < 32(1 + l_0)\omega^3.$$

Поэтому условие (2.16) можно заменить более жестким, но более простым неравенством

$$\frac{\varepsilon\omega^3}{\rho} < \frac{1}{32(1+l_0)}.$$

§ 3. Оценки решений вспомогательной дискретной краевой задачи

Пусть матрицы Y_r, Z_r $(r = 1, 2, \ldots, m)$ определены как решение краевой задачи (2.14). Будем предполагать, что коэффициенты и правые части удовлетворяют условиям

$$\begin{gathered} \|\mathcal{L}\| = l_0, \qquad \|\mathcal{D}_r\| \leqslant d, \qquad \|\mathcal{E}_r\| \leqslant d, \\ \|B^{-p}\| \leqslant \sqrt{\omega}\,[\omega/(1+\omega)]^{\frac{p-1}{2}} \qquad (p \geqslant 1), \\ \|C^p\| \leqslant \sqrt{\omega}\,[\omega/(1+\omega)]^{\frac{p}{2}} \qquad (p \geqslant 0). \end{gathered} \tag{3.1}$$

Лемма 3.1. *При условии* (3.1) *решение* Y_r, Z_r ($r = 0,1,2,\dots,m$) *однозначно определено и удовлетворяет оценкам*

$$\begin{aligned} \|Z_r\| &\leqslant \sqrt{\omega}\,[\omega/(1+\omega)]^{\frac{m-r-1}{2}} + 2(1+\omega)\sqrt{\omega}d, \\ \|Y_r\| &\leqslant l_0\omega\,[\omega/(1+\omega)]^{\frac{m+r-1}{2}} + 2(1+\omega)\sqrt{\omega}\left[l_0\omega[\omega/(1+\omega)]^{\frac{r}{2}} + 1\right] d. \end{aligned} \tag{3.2}$$

Доказательство. Легко проверить, что решение представимо в виде

$$\begin{aligned} Z_p &= B^{-(m-p)} + \sum_{i=p}^{m-1} B^{-(m-i)}\mathcal{D}_i \quad (0 \leqslant p \leqslant m-1), \\ Z_r &= C^r\mathcal{L}Z_0 + \sum_{i=0}^{r-1} C^i\mathcal{E}_{r-i} \qquad (0 \leqslant r \leqslant m) \end{aligned} \tag{3.3}$$

и потому единственно. Неравенства (3.2) непосредственно следуют из формул (3.3) и условий леммы. Начнем с оценки $\|Z_r\|$:

$$\begin{aligned} \|Z_r\| &\leqslant \sqrt{\omega}\left\{[\omega/(1+\omega)]^{\frac{m-r-1}{2}} + d\sum_{i=r}^{m-1}[\omega/(1+\omega)]^{\frac{m-r-1}{2}}\right\} \\ &\leqslant \sqrt{\omega}\left\{[\omega/(1+\omega)]^{\frac{m-r-1}{2}} + d\sum_{p=0}^{\infty}[\omega/(1+\omega)]^{\frac{p}{2}}\right\} \\ &= \sqrt{\omega}\left\{[\omega/(1+\omega)]^{\frac{m-r-1}{2}} + d\{(1+\omega) + \sqrt{\omega(1+\omega)}\}\right\} \\ &\leqslant \sqrt{\omega}\,[\omega/(1+\omega)]^{\frac{m-r-1}{2}} + 2(1+\omega)\sqrt{\omega}d. \end{aligned}$$

В частности,

$$\|Z_0\| \leqslant \sqrt{\omega}\,[\omega/(1+\omega)]^{\frac{m-1}{2}} + 2(1+\omega)\sqrt{\omega}d.$$

Используя (3.3), можно получить также следующую оценку $\|Y_r\|$:

$$\begin{aligned} \|Y_r\| &\leqslant l_0\sqrt{\omega}\,[\omega/(1+\omega)]^{\frac{r}{2}}\|Z_0\| + d\sum_{p=0}^{\infty}\sqrt{\omega}[\omega/(1+\omega)]^{\frac{p}{2}} \\ &\leqslant l_0\sqrt{\omega}\,[\omega/(1+\omega)]^{\frac{r}{2}}\|Z_0\| + 2(1+\omega)\sqrt{\omega}d \\ &\leqslant l_0\omega\,[\omega/(1+\omega)]^{\frac{m+r-1}{2}} + 2(1+\omega)\sqrt{\omega}d\left[l_0\sqrt{\omega}\,[\omega/(1+\omega)]^{\frac{r}{2}} + 1\right]. \end{aligned}$$

Лемма 3.1 доказана

Непосредственно из леммы 3.1 вытекает

Лемма 3.2. *Пусть выполнены условия* (3.1). *Тогда при любом* m *решение* Y_r, Z_r ($r = 1,2,\dots,m$) *задачи* (2.14) *удовлетворяет оценкам*

$$\begin{aligned} \|Z_r\| &\leqslant \sqrt{\omega} + 2(1+\omega)\sqrt{\omega}d, \\ \|Y_r\| &\leqslant \omega l_0 + 2(1+\sqrt{\omega}l_0)(1+\omega)\sqrt{\omega}d. \end{aligned} \tag{3.4}$$

Предположим, что можно выразить правые части $\mathcal{D}_r$, $\mathcal{E}_r$ системы (2.14) *через* Z_r, Y_r *так, что*

$$\|\mathcal{D}_r\| \leqslant \delta \max\{\|Y_r\|, \|Z_r\|\}, \quad \|\mathcal{E}_r\| \leqslant \delta \max\{\|Y_r\|, \|Z_r\|\}. \tag{3.5}$$

Тогда в условиях (3.1) *можно положить*

$$d = \max\{\max_r \|\mathcal{D}_r\|, \max_r \|\mathcal{E}_r\|\} \leqslant \delta \max\{\max_r \|Y_r\|, \max_r \|Z_r\|\} = \delta M. \tag{3.6}$$

Здесь введено обозначение

$$M = \max\{\max_r \|Y_r\|, \max_r \|Z_r\|\}. \tag{3.7}$$

Лемма 3.3. *Пусть выполнены условия* (3.1), (3.5), (3.6) *и неравенство*

$$\delta < 1/[4(1+\omega)(\omega l_0 + \sqrt{\omega})]. \tag{3.8}$$

Тогда число M, *определенное формулой* (3.7), *можно оценить следующим образом:*

$$M < 2(\omega l_0 + \sqrt{\omega}). \tag{3.9}$$

Доказательство. Используя (3.6) и заменяя d в правых частях неравенств (3.4) на δM, получаем

$$\begin{aligned} M &= \max_{i,j}\{\|Y_i\|, \|Z_j\|\} \\ &\leqslant \max\left\{\sqrt{\omega} + 2(1+\omega)\sqrt{\omega}\delta M,\ \omega l_0 + 2(\sqrt{\omega} + \omega l_0)(1+\omega)\delta M\right\} \\ &\leqslant (\sqrt{\omega} + \omega l_0)[1 + 2(1+\omega)\delta M], \end{aligned}$$

$$M[1 - 2(\sqrt{\omega} + \omega l_0)(1+\omega)\delta] \leqslant \sqrt{\omega} + \omega l_0.$$

Воспользовавшись (3.8), приходим к неравенству $M/2 < \sqrt{\omega} + l_0\omega$, которое эквивалентно утверждению леммы 3.3.

Лемма 3.4. *Пусть выполнены условия* (3.1), (3.5), (3.8). *Тогда решение краевой задачи* (2.14) *удовлетворяет следующей оценке:*

$$\|Y_m\| \leqslant l_0\sqrt{1+\omega}\,[\omega/(1+\omega)]^m + 4\delta(1+\omega)(\omega l_0 + \sqrt{\omega})^2. \tag{3.10}$$

Доказательство основано на применении неравенства (3.2), которое в несколько огрубленном виде вместе с (3.6) дает оценку

$$\|Y_m\| \leqslant l_0\sqrt{1+\omega}\,[\omega/(1+\omega)]^m + 2(1+\omega)(\omega l_0 + \sqrt{\omega})\delta M.$$

Остается воспользоваться оценкой (3.9) константы M.

§ 4. Оценка погрешности приведения матрицы к каноническому виду

Цель настоящего параграфа — изучить поведение матриц $C_j^* A C_j$ при $j \to \infty$. Воспользовавшись неравенством (см. (1.9))

$$\|C_j^* A C_j - \widehat{C}_j^* A \widehat{C}_j\| \leqslant \frac{\varepsilon}{(1-\varepsilon)^2}\|A\|, \tag{4.1}$$

мы ограничимся подробным изучением таких произведений $\widehat{C}_j^* A \widehat{C}_j$, у которых сомножители $\widehat{C}_j$ строго ортогональны ($\widehat{C}_j^* \widehat{C}_j = I$). Напомним (см. (1.11), (1.12), (2.15)) представления

$$\widehat{C}_j = \begin{bmatrix} \widehat{P}_j & \dots \\ \widehat{Q}_j & \dots \end{bmatrix} = \begin{bmatrix} \widehat{C}_{11}^{(j)} & \widehat{C}_{12}^{(j)} \\ \widehat{C}_{21}^{(j)} & \widehat{C}_{22}^{(j)} \end{bmatrix}, \qquad \widehat{P}_j^* \widehat{P}_j + \widehat{Q}_j^* \widehat{Q}_j = I_k,$$

и оценку

$$l_j = l_j(\rho) \leqslant l_0 \sqrt{1+\omega}[\omega/(1+\omega)]^j + 4\delta(1+\omega)(\omega l_0 + \sqrt{\omega})^2, \tag{4.2}$$

где параметр $l_j(\rho)$ выражается через матрицы $\overline{\overline{P}}_j$, $\overline{\overline{Q}}_j$, определенные равенствами

$$\begin{bmatrix} \overline{\overline{P}}_j \\ \overline{\overline{Q}}_j \end{bmatrix} = U \begin{bmatrix} \widehat{P}_j \\ \widehat{Q}_j \end{bmatrix} = U \begin{bmatrix} \widehat{C}_{11}^{(j)} \\ \widehat{C}_{21}^{(j)} \end{bmatrix},$$

с помощью формулы (см. (1.20)) $l_j = l_j(\rho) = \|\overline{\overline{Q}}_j(\overline{\overline{P}}_j + L\overline{\overline{Q}}_j)^{-1}\|$. Имеем

$$\begin{gathered} \|\overline{\overline{P}}_j\| \leqslant 1, \quad \|\overline{\overline{Q}}_j\| \leqslant 1, \quad \|\overline{\overline{P}}_j + L\overline{\overline{Q}}_j\| \leqslant 1 + \|L\|, \\ \|\overline{\overline{Q}}_j\| = \|\overline{\overline{Q}}_j(\overline{\overline{P}}_j + L\overline{\overline{Q}}_j)^{-1}(\overline{\overline{P}}_j + L\overline{\overline{Q}}_j)\| \leqslant (1 + \|L\|)\, l_j. \end{gathered} \tag{4.3}$$

Введем обозначение

$$V_j = U\widehat{C}_j = \begin{bmatrix} V_{11}^{(j)} & V_{12}^{(j)} \\ V_{21}^{(j)} & V_{22}^{(j)} \end{bmatrix} = \begin{bmatrix} \overline{\overline{P}}_j & \dots \\ \overline{\overline{Q}}_j & \dots \end{bmatrix}.$$

Используя оценки (4.2) и (2.15), перепишем последнее неравенство в (4.3) следующим образом:

$$\begin{aligned} \|V_{21}^{(j)}\| \leqslant (1 + \|L\|)\, l_j \leqslant l_0 \sqrt{2\omega(1+\omega)}&[\omega/(1+\omega)]^j \\ + 4\delta\sqrt{2\omega}(1+\omega)(\omega l_0 + \sqrt{\omega})^2 &= \varepsilon_j. \end{aligned} \tag{4.4}$$

Ввиду ортогональности матриц V_j справедливы равенства

$$[V_{11}^{(j)}]^* V_{11}^{(j)} + [V_{21}^{(j)}]^* V_{21}^{(j)} = I_k, \qquad V_{21}^{(j)}[V_{21}^{(j)}]^* + V_{22}^{(j)}[V_{22}^{(j)}]^* = I_{N-k},$$
$$V_{11}^{(j)}[V_{11}^{(j)}]^* + V_{12}^{(j)}[V_{12}^{(j)}]^* = I_k,$$

которые вместе с (4.4) приводят к оценкам

$$\begin{aligned} &\|\,[V_{11}^{(j)}]^* V_{11}^{(j)} - I_k\| = \|V_{11}^{(j)}[V_{11}^{(j)}]^* - I_k\| \leqslant \varepsilon_j^2, \\ &\|\,[V_{22}^{(j)}]^* V_{22}^{(j)} - I_{N-k}\| = \|V_{22}^{(j)}[V_{22}^{(j)}]^* - I_{N-k}\| \leqslant \varepsilon_j^2, \\ &\|V_{21}^{(j)}\| = \sqrt{\|I_k - V_{11}^{(j)}[V_{11}^{(j)}]^*\|} = \sqrt{\|I_k - [V_{11}^{(j)}]^* V_{11}^{(j)}\|} \\ &\qquad = \sqrt{\|I_k - \overline{\overline{P}}_j^* \overline{\overline{P}}_j\|} = \|\overline{\overline{Q}}_j\| \leqslant \varepsilon_j. \end{aligned}$$

Пусть $\widetilde{V}_j = \begin{bmatrix} \widetilde{V}_{11}^{(j)} & 0 \\ 0 & \widetilde{V}_{22}^{(j)} \end{bmatrix}$ — клеточно-диагональная ортогональная матрица с клетками

$$\widetilde{V}_{11}^{(j)} = V_{11}^{(j)}([V_{11}^{(j)}]^* V_{11}^{(j)})^{-\frac{1}{2}}, \qquad \widetilde{V}_{22}^{(j)} = V_{22}^{(j)}([V_{22}^{(j)}]^* V_{22}^{(j)})^{-\frac{1}{2}}.$$

Из неравенств

$$\begin{aligned} &\|\widetilde{V}_{11}^{(j)} - V_{11}^{(j)}\| \leqslant 1/[(1-\varepsilon_j^2)^{\frac{1}{2}}] - 1, \\ &\|\widetilde{V}_{22}^{(j)} - V_{22}^{(j)}\| \leqslant 1/[(1-\varepsilon_j^2)^{\frac{1}{2}}] - 1, \\ &\|V_{11}^{(j)}\| \leqslant \varepsilon_j, \end{aligned}$$

где $\varepsilon_j < 1$, следует оценка

$$\begin{aligned} \|\widetilde{V}_j - V_j\| &\leqslant \varepsilon_j + 1/[(1-\varepsilon_j^2)^{\frac{1}{2}}] - 1 \leqslant \varepsilon_j + \tfrac{1}{2}\varepsilon_j^2(1 + \varepsilon_j^2 + \varepsilon_j^4 + \ldots) \\ &= \varepsilon_j + \varepsilon_j^2/[2(1-\varepsilon_j^2)]. \end{aligned}$$

Во всяком случае при $\varepsilon_j < 1/2$ можно утверждать, что $\|\widetilde{V}_j - V_j\| \leqslant 2\varepsilon_j$. Определим ортогональную матрицу $\widetilde{C}_j = U^* \widetilde{V}_j$. Заметим, что верны равенства

$$\widetilde{C}_j^* A \widetilde{C}_j = \begin{bmatrix} [\widetilde{V}_{11}^{(j)}]^* & 0 \\ 0 & [\widetilde{V}_{22}^{(j)}]^* \end{bmatrix} \begin{bmatrix} B & D \\ 0 & C \end{bmatrix} \begin{bmatrix} \widetilde{V}_{11}^{(j)} & 0 \\ 0 & \widetilde{V}_{22}^{(j)} \end{bmatrix}$$

$$= \begin{bmatrix} [\widetilde{V}_{11}^{(j)}]^* B \widetilde{V}_{11}^{(j)} & [\widetilde{V}_{11}^{(j)}]^* D \widetilde{V}_{22}^{(j)} \\ 0 & [\widetilde{V}_{22}^{(j)}]^* C \widetilde{V}_{22}^{(j)} \end{bmatrix}.$$

Поэтому $\widetilde{C}_j^* A \widetilde{C}_j$ — верхняя клеточно-треугольная матрица, диагональные клетки которой ортогонально-подобны клеткам B, C в каноническом представлении (1.1) исследуемой матрицы A.

Очевидно, что $\|\widetilde{C}_j - \widehat{C}_j\| \leqslant 2\varepsilon_j$ при $\varepsilon_j < 1/2$. Так же, как доказывали (1.9), устанавливаем оценку

$$\|\widetilde{C}_j^* A \widetilde{C}_j - \widehat{C}_j^* A \widehat{C}_j\| \leqslant 4\varepsilon_j(1+\varepsilon_j)\|A\|. \tag{4.5}$$

Рассматривая (4.5) совместно с (4.1) при $\varepsilon < 1$, $\varepsilon_j < 1/2$, получаем неравенство

$$\|C_j^* A C_j - \widetilde{C}_j^* A \widetilde{C}_j\| \leqslant [\varepsilon/(1-\varepsilon)^2 + 4\varepsilon_j(1+\varepsilon_j)]\|A\|,$$

которое означает, что при достаточно малых ε, ε_j «почти-ортогональные» матрицы C_j (см. § 6), рассчитанные в процессе выполнения ортогонально-степенного алгоритма, осуществляют подобное приведение матрицы A к клеточно-треугольному виду с определенной точностью.

§ 5. Выбор стартового базиса перед началом алгоритма

В предшествующих параграфах было исследовано влияние вычислительных погрешностей на результаты ортогонально-степенного алгоритма, с помощью которого строится подобное ортогональное преобразование, приводящее матрицу A к клеточно-диагональному виду. В основу доказательств положены две (кроме предположения малости погрешностей, допускаемых на каждом шаге алгоритма) гипотезы, связанные с матрицей A. Первая заключается в возможности разделения спектра матрицы A окружностью с радиусом $\rho\|A\|$. Это позволяет представить матрицу A в виде

$$A = \frac{1}{\rho} U^* \begin{bmatrix} I_k & -L \\ 0 & I_{N-k} \end{bmatrix} \begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix} \begin{bmatrix} I_k & L \\ 0 & I_{N-k} \end{bmatrix} U, \tag{5.1}$$

где матрица U ортогональна ($U^*U = I$), а матрицы L, B, C при $p \geqslant 1$ удовлетворяют оценкам (1.3), (1.4). Согласно второй гипотезе составная клеточная матрица

$$\begin{bmatrix} X \\ Y \end{bmatrix} = U \begin{bmatrix} I_k \\ 0 \end{bmatrix}, \tag{5.2}$$

образованная с помощью матрицы U, такова, что матрица $X + LY$ не вырождена и величина

$$l = \|Y(X + LY)^{-1}\| \tag{5.3}$$

не слишком велика. Вторая гипотеза обеспечивает удовлетворительное функционирование алгоритма за счет ограниченности выбора возможных ортогональных матриц U в (5.1). Удобно заменить вторую гипотезу другой — более грубой, но наглядной.

Покажем сначала, что при условии

$$\|Y\| < 1/\sqrt{2\omega} \tag{5.4}$$

величина l будет ограниченной. Из (5.2) следует, что все сингулярные числа σ_j ($\sigma_k \geqslant \sigma_{k-1} \geqslant \ldots \geqslant \sigma_1 > 0$) матрицы, стоящей в левой части равенства (5.2), одинаковы и равны единице (см., например, [13]). В частности,

$$\sigma_{\min}\left(\begin{bmatrix} X \\ Y \end{bmatrix}\right) = \sigma_1\left(\begin{bmatrix} X \\ Y \end{bmatrix}\right) = 1, \quad \left\|\begin{bmatrix} X \\ Y \end{bmatrix}\right\| = \sigma_{\max}\left(\begin{bmatrix} X \\ Y \end{bmatrix}\right) = 1.$$

Поэтому

$$\sigma_{\min}([X]) = \sigma_{\min}\left(\begin{bmatrix} X \\ 0 \end{bmatrix}\right) = \sigma_{\min}\left(\begin{bmatrix} X \\ Y \end{bmatrix} - \begin{bmatrix} 0 \\ Y \end{bmatrix}\right) \geqslant 1 - \|Y\|,$$

$$\begin{aligned} \sigma_{\min}([X + LY]) &= \|(X + LY)^{-1}\|^{-1} \\ &\geqslant \sigma_{\min}([X]) - \|L\|\,\|Y\| \geqslant 1 - (1 + \|L\|)\,\|Y\|. \end{aligned}$$

В силу неравенства $1 + \|L\| \leqslant \sqrt{2\omega}$ (см. (1.3), (1.4)) получаем

$$\|Y(X + LY)^{-1}\| \leqslant \|Y\|/(1 - \sqrt{2\omega}\,\|Y\|). \tag{5.5}$$

Отметим, что линейная оболочка первых k столбцов ортогональной матрицы U^* из (5.1) совпадает с инвариантным относительно A подпространством, отвечающим точкам спектра A, модуль которых больше величины $\rho\|A\|$. Ортогональный проектор $\Pi_{\mathrm{inv}(k)} = \Pi^*_{\mathrm{inv}(k)} = \Pi^2_{\mathrm{inv}(k)}$ на это инвариантное подпространство определяется формулой

$$\Pi_{\mathrm{inv}(k)} = U^* \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} U.$$

Пусть Π_k — ортогональный проектор на линейную оболочку первых k ортов: $\Pi_k = \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix}$. Очевидно, что

$$\Pi_k U^* \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} I_k \\ 0 \end{bmatrix} = \begin{bmatrix} I_k \\ 0 \end{bmatrix} = U^* \begin{bmatrix} X \\ Y \end{bmatrix},$$

$$\Pi_{\mathrm{inv}(k)} U^* \begin{bmatrix} X \\ Y \end{bmatrix} = U^* \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} U U^* \begin{bmatrix} X \\ Y \end{bmatrix} = U^* \begin{bmatrix} X \\ 0 \end{bmatrix}.$$

Поэтому

$$[\Pi_k - \Pi_{\mathrm{inv}(k)}] \begin{bmatrix} I_k \\ 0 \end{bmatrix} = [\Pi_k - \Pi_{\mathrm{inv}(k)}] U^* \begin{bmatrix} X \\ Y \end{bmatrix} = U^* \begin{bmatrix} 0 \\ Y \end{bmatrix},$$

$$\|Y\| \leqslant \|\Pi_k - \Pi_{\mathrm{inv}(k)}\|.$$

Предположим, что

$$\|[\Pi_k - \Pi_{\mathrm{inv}(k)}]\| < 1/(\sqrt{2\omega}) < 1. \tag{5.6}$$

Тогда будет выполнено неравенство (5.4), что дает нам возможность вывести из (5.5) оценку

$$l(\rho) = \|Y(X + LY)^{-1}\| \leqslant \frac{\|\Pi_k - \Pi_{\mathrm{inv}(k)}\|}{1 - \sqrt{2\omega}\,\|\Pi_k - \Pi_{\mathrm{inv}(k)}\|}.$$

Таким образом, доказана следующая

Лемма 5.1. *Если матрица A представима в виде* (5.1), *выполнены условия* (1.3), (1.4) *и параметр $l(\rho)$ определен формулами* (5.2), (5.3), *то* (*в силу неравенства* (5.6) *для $\Pi_k - \Pi_{\mathrm{inv}(k)}$*) *параметр $l(\rho)$ конечен:*

$$l(\rho) \leqslant \frac{\|\Pi_k - \Pi_{\mathrm{inv}(k)}\|}{1 - \sqrt{2}\omega(\rho)\|\Pi_k - \Pi_{\mathrm{inv}(k)}\|}.$$

Напомним, что норма разности двух ортогональных проекторов может быть истолкована как величина синуса угла, образованного теми подпространствами, на которые эти проекторы осуществляют проектирование.

Поясним, как лемма 5.1 может быть использована. Прежде чем переходить к ортогонально-степенному алгоритму для выделения инвариантных подпространств матрицы A, предположим, что выбран некоторый ортонормированный базис, координаты векторов которого образуют столбцы ортогональной матрицы W. При этом ортогональный проектор Π'_k на линейную оболочку первых k векторов этого базиса удовлетворяет неравенству

$$\|\Pi_{\mathrm{inv}(k)} - \Pi'_k\| < 1/[2\sqrt{2}\omega(\rho)]. \tag{5.7}$$

С помощью W построим ортогонально-подобную матрице A матрицу $A_0 = W^* A W$. Очевидно, что

$$\Pi_k = W^* \Pi'_k W, \quad \Pi^{(0)}_{\mathrm{inv}(k)} = W^* \Pi_{\mathrm{inv}(k)} W, \tag{5.8}$$

где так же, как и выше, Π_k — ортогональный проектор на линейную оболочку k векторов координатного базиса, тогда как $\Pi^{(0)}_{\mathrm{inv}(k)}$ — ортогональный проектор на k-мерное инвариантное подпространство матрицы A_0, отвечающее точкам спектра λ таким, что $|\lambda| > \rho\|A_0\| = \rho\|A\|$. Нетрудно проверить, что для матриц A и A_0 соответствующие величины $\omega(\rho)$ одинаковы. Следствием (5.7), (5.8) являются соотношения

$$\|\Pi^{(0)}_{\mathrm{inv}(k)} - \Pi_k\| = \|\Pi_{\mathrm{inv}(k)} - \Pi'_k\|, \qquad \|\Pi^{(0)}_{\mathrm{inv}(k)} - \Pi_k\| < 1/[2\sqrt{2}\omega(\rho)].$$

Последнее неравенство позволяет утверждать, что параметр $l(\rho)$ для матрицы A_0 конечен и можно применять ортогонально-степенной алгоритм для выделения инвариантных подпространств, отвечающих точкам спектра λ таким, что $|\lambda| < \rho\|A_0\|$ и $|\lambda| > \rho\|A_0\|$.

Мы показали, что перед обращением к алгоритму целесообразно провести выбор стартового базиса, который задается с помощью матрицы W и удовлетворяет, если это возможно, условию (5.7). Оказывается, что этот базис можно составить из левых сингулярных векторов матрицы A^p — некоторой степени матрицы A. При этом сингулярные векторы упорядочены так, что с увеличением номера базисного вектора отвечающее ему сингулярное число не возрастает. Иными словами, базисные векторы упорядочены по убыванию сингулярных чисел.

Представим A^p в виде суммы $A^p = A_1 + A_2$, где

$$A_1 = \frac{1}{\rho^p} U^* \begin{bmatrix} B^p & B^p L \\ 0 & 0 \end{bmatrix} U, \qquad A_2 = \frac{1}{\rho^p} U^* \begin{bmatrix} 0 & -LC^p \\ 0 & C^p \end{bmatrix} U.$$

Следует отметить, что ввиду (1.3), (1.4)

$$\sigma^{(1)}_{N-k+1} \geqslant 1/[\rho^p \|B^{-p}\|] \geqslant [(1+\omega(\rho))^{\frac{p-1}{2}}]/[\rho^p (\omega(\rho))^{\frac{p}{2}}], \tag{5.9}$$

где $\sigma^{(1)}_j$ — сингулярные числа матрицы A_1, упорядоченные по убыванию:

$$\sigma^{(1)}_N \geqslant \sigma^{(1)}_{N-1} \geqslant \ldots \geqslant \sigma^{(1)}_{N-k+1} > \sigma^{(1)}_{N-k} = \sigma^{(1)}_{N-k-1} = \ldots = \sigma^{(1)}_1 = 0.$$

С другой стороны (см. (1.3), (1.4)), имеем $\|A_2\| \leqslant \omega^{\frac{p+1}{2}}/[\rho^p (1+\omega)^{\frac{p}{2}}]$. Поэтому справедлива формула

$$\|A^p - A_1\| = \|A_2\| \leqslant \frac{[\omega(\rho)]^{\frac{p+1}{2}}}{\rho^p [1+\omega(\rho)]^{\frac{p}{2}}}, \tag{5.10}$$

а базис, составленный из k левых сингулярных векторов матрицы A_1, соответствующих ее ненулевым сингулярным числам, будет иметь вид $U^* \begin{bmatrix} I_k \\ 0 \end{bmatrix}$, т. е. этот базис будет совпадать с одним из возможных базисов подпространства, инвариантного относительно матрицы A и отвечающего точкам спектра $\lambda(A)$ таким, что $|\lambda| > \rho\|A\|$. Ортогональный проектор на это подпространство мы условились обозначать через $\Pi_{\mathrm{inv}(k)}$. Ортогональный проектор на линейную оболочку левых сингулярных векторов матрицы A^p, отвечающих ее сингулярным числам $\sigma_N, \sigma_{N-1}, \dots, \sigma_{N-k+1}$, обозначим символом $\Pi_k^{(p)}$.

Дальнейшие рассуждения опираются на известную теорему, доказательство которой приведено, например, в [14]. Сформулируем эту теорему в удобной для нас форме, предварительно определив понятие «зазора»:

- Пусть M и M_1 — две близкие матрицы. Предположим, что сингулярные числа $\sigma_j(M_1)$ $(1 \leqslant j \leqslant N)$ матрицы M_1 разбиты на две группы так, что каждое сингулярное число $\sigma_j(M_1)$ $(N-k+1 \leqslant j \leqslant N)$ первой группы больше всех сингулярных чисел $\sigma_j(M_1)$ $(1 \leqslant j \leqslant N-k)$ второй группы. *Зазором* Δ между двумя этими группами назовем наименьшую разность между сингулярными числами из разных групп: $\sigma_{N-k+1}(M_1) - \sigma_{N-k}(M_1) = \Delta$.

Теорема 5.1. *Если*

$$\|M - M_1\| < \delta_0 \leqslant \Delta/2, \tag{5.11}$$

то ортогональные проекторы Π, Π_1 *на линейные оболочки сингулярных векторов, отвечающих наибольшим* k *сингулярным числам матриц* M *и* M_1 *соответственно, подчинены неравенству*

$$\|\Pi - \Pi_1\| < \frac{\delta_0/\Delta}{1 - \delta_0/\Delta} \leqslant \frac{2\delta_0}{\Delta}.$$

Применяя теорему 5.1 к матрицам $M = A^p$, $M_1 = A_1$, мы можем определить (см. (5.9), (5.10))

$$\Delta = \frac{[1+\omega(\rho)]^{\frac{p-1}{2}}}{\rho^p[\omega(\rho)]^{\frac{p}{2}}}, \quad \delta_0 = \frac{[\omega(\rho)]^{\frac{p+1}{2}}}{\rho^p[1+\omega(\rho)]^{\frac{p}{2}}}, \quad \frac{\delta_0}{\Delta} = \frac{[\omega(\rho)]^{\frac{2p+1}{2}}}{[1+\omega(\rho)]^{\frac{2p-1}{2}}}.$$

Отметим, что при достаточно большом p таком, что

$$\frac{4\sqrt{2\omega(\rho)}\,[\omega(\rho)]^{\frac{2p+1}{2}}}{[1+\omega(\rho)]^{\frac{2p-1}{2}}} = \frac{4\sqrt{2}[\omega(\rho)]^{\frac{2p+3}{2}}}{[1+\omega(\rho)]^{\frac{2p-1}{2}}} < 1, \tag{5.12}$$

выполнено условие (5.11) и справедлива оценка

$$\sqrt{2\omega(\rho)}\,\|\Pi_{\mathrm{inv}(k)} - \Pi_k^{(p)}\| < 1/2.$$

Выбирая стартовый базис из сингулярных векторов матрицы A^p, упорядоченных по убыванию отвечающих им сингулярных чисел, и тем самым полагая $\Pi' = \Pi_k^{(p)}$, мы обеспечиваем выполнение неравенства (5.7).

В реальных ситуациях при выборе степени p необходимо учитывать неизбежные вычислительные погрешности при вычислении матрицы A^p и построении базиса из ее сингулярных векторов. Мы ограничимся тем, что покажем, как учесть погрешность вычисления матрицы A^p. Обозначив результат вычисления через $[A^p]_{\rm comp}$, предположим, что после анализа процедуры расчета мы убедились в справедливости оценки

$$\|[A^p]_{\rm comp} - A^p\| \leqslant \varepsilon_p \|A\|^p = \varepsilon_p \frac{(\rho\|A\|)^p}{\rho^p} = \frac{\varepsilon_p}{\rho^p}.$$

Определив матрицы $M = [A^p]_{\rm comp}$, $M_1 = A_1$ и полагая

$$\Delta = \frac{[1+\omega(\rho)]^{\frac{p-1}{2}}}{\rho^p[\omega(\rho)]^{\frac{p}{2}}}, \qquad \delta_0 = \frac{1}{\rho^p}\left\{\frac{[\omega(\rho)]^{\frac{p+1}{2}}}{[1+\omega(\rho)]^{\frac{p}{2}}} + \varepsilon_p\right\},$$

$$\frac{\delta_0}{\Delta} = \frac{[\omega(\rho)]^{p+\frac{1}{2}}}{[1+\omega(\rho)]^{p-\frac{1}{2}}} + \varepsilon_p \frac{[\omega(\rho)]^{\frac{p}{2}}}{[1+\omega(\rho)]^{\frac{p-1}{2}}},$$

применяем теорему 5.1. Для того чтобы параметр $l(\rho)$ был конечным, при использовании базиса из упорядоченных сингулярных векторов матрицы $[A^p]_{\rm comp}$ можно рекомендовать выбирать степень p и характеристику ε_p точности вычисления матрицы A^p так, чтобы выполнялось неравенство

$$\frac{4\sqrt{2}\omega^{\frac{p}{2}}}{(1+\omega)^{\frac{p-1}{2}}}\left[\sqrt{\omega}\left(\frac{\omega}{1+\omega}\right)^{\frac{p}{2}} + \varepsilon_p\right] < 1. \tag{5.13}$$

Условие (5.13) более жесткое, чем условие (5.12): последнее есть частный случай условия (5.13) при $\varepsilon_p = 0$.

§ 6. Оценки для QR-алгоритма

Хорошо известно, что ортогонально-степенной метод эквивалентен с точностью до вычислительных погрешностей так называемому QR-алгоритму. В этом параграфе мы приводим для полноты изложения несложные неравенства, с помощью которых оценки из § 1–5 могут быть использованы при анализе влияния погрешностей на результат, полученный с помощью QR-алгоритма.

Процедура QR-алгоритма состоит в последовательном вычислении ортогональных матриц Q_j и верхних треугольных матриц R_j, связанных равенствами

$$\begin{aligned} &A + \Delta_0' = Q_1R_1, && A_1 = R_1Q_1 + \Delta_0'', \\ &A_{j-1} + \Delta_{j-1}' = Q_jR_j, && A_j = R_jQ_j + \Delta_j'' \qquad (1 \leqslant j \leqslant n+1). \end{aligned} \tag{6.1}$$

Учет вычислительных погрешностей в этих равенствах осуществляется с помощью матриц погрешностей Δ_j', Δ_j'', которые мы подчиняем условиям

$$\begin{aligned} &\|\Delta_0'\| \leqslant \varepsilon\|A\|, && \|\Delta_1''\| \leqslant \varepsilon\|A_1\|, \\ &\|\Delta_{j-1}'\| \leqslant \varepsilon\|A_{j-1}\|, && \|\Delta_j''\| \leqslant \varepsilon\|A_j\| \qquad (1 \leqslant j \leqslant n+1). \end{aligned} \tag{6.2}$$

Матрицы Q_j $(j = 1, \dots, n+1)$ будем считать приближенно ортогональными, что означает следующее:

$$Q_j^*Q_j = I + Z_j, \quad \text{где} \quad \|Z_j\| \leqslant \varepsilon. \tag{6.3}$$

Интересующие нас неравенства обеспечивают приводимые ниже леммы.

Лемма 6.1. *Пусть матрицы* A_j $(1 \leqslant j \leqslant N)$, *определяемые равенствами* (6.1), *удовлетворяют условиям* (6.2), (6.3). *Тогда справедливы следующие оценки:*

$$\|A_j\| < \frac{1}{(1-\varepsilon)^3}\,\|A_{j-1}\|, \qquad \|A_j\| < \frac{1}{(1-\varepsilon)^{3n}}\,\|A\|. \tag{6.4}$$

Прежде чем переходить к доказательству леммы, заметим, что из (6.2), (6.4) вытекают неравенства

$$\|\Delta'_{j-1}\| < \frac{\varepsilon}{(1-\varepsilon)^{3n}}\,\|A\|, \qquad \|\Delta''_j\| < \frac{\varepsilon}{(1-\varepsilon)^{3n}}\,\|A\|. \tag{6.5}$$

Доказательство леммы 6.1. В силу (6.1) справедливо равенство $A_j = Q_j^{-1}A_{j-1}Q_j + Q_j^{-1}\Delta'_{j-1}Q_j + \Delta''_j$, которое вместе с (6.2) служит основанием оценки $\|A_j\| \leqslant \|Q_j^{-1}\|\,\|Q_j\|\,\|A_{j-1}\| + \varepsilon\|Q_j^{-1}\|\,\|Q_j\|\,\|A_{j-1}\| + \varepsilon\|A_j\|$, эквивалентной (при $\varepsilon < 1$) неравенству

$$\|A_j\| \leqslant \frac{1+\varepsilon}{1-\varepsilon}\,\|Q_j^{-1}\|\,\|Q_j\|\,\|A_{j-1}\| \leqslant \frac{1}{(1-\varepsilon)^2}\|Q_j^{-1}\|\,\|Q_j\|\,\|A_{j-1}\|\,. \tag{6.6}$$

Из (6.3) находим

$$\begin{aligned} &\|Q_j^*\| = \|Q_j\| \leqslant \sqrt{\|I+Z_j\|} \leqslant \sqrt{1+\varepsilon} \leqslant \frac{1}{\sqrt{1-\varepsilon}},\\ &\|Q_j^{-1}\| \leqslant \sqrt{\|(I+Z_j)^{-1}\|} \leqslant \frac{1}{\sqrt{1-\varepsilon}}\,. \end{aligned} \tag{6.7}$$

Очевидно, что (6.6) и (6.7) приводят к требуемым неравенствам (6.4).

Лемма 6.2. *Справедливы следующие равенства:*

$$\begin{aligned} Q_1Q_2\ldots Q_{n+1}R_{n+1} &= AQ_1Q_2\ldots Q_n + \Delta'_0Q_1Q_2\ldots Q_n\\ &+\sum_{j=1}^{n-1} Q_1Q_2\ldots Q_j(\Delta'_j+\Delta''_j)Q_{j+1}Q_{j+2}\ldots Q_n + Q_1Q_2\ldots Q_n(\Delta'_n+\Delta''_n), \end{aligned} \tag{6.8}$$

$$\begin{aligned} A_n &= Q_n^{-1}Q_{n-1}^{-1}\ldots Q_1^{-1}AQ_1Q_2\ldots Q_n + Q_n^{-1}Q_{n-1}^{-1}\ldots Q_1^{-1}\Delta'_0Q_1Q_2\ldots Q_n\\ &+\sum_{j=1}^{n-1}[Q_n^{-1}Q_{n-1}^{-1}\ldots Q_{j+1}^{-1}(\Delta'_j+\Delta''_j)Q_{j+1}\ldots Q_n] + \Delta''_n. \end{aligned} \tag{6.9}$$

Доказательство проводится индукцией по n с использованием равенств (6.1).

Введем обозначения $C_j = Q_1Q_2\ldots Q_j$,

$$\begin{aligned} \widetilde{\Delta}_n^{(0)} &= Q_n^{-1}Q_{n-1}^{-1}\ldots Q_1^{-1}\Delta'_0Q_1Q_2\ldots Q_n\\ &\quad+\sum_{j=1}^{n-1}[Q_n^{-1}Q_{n-1}^{-1}\ldots Q_{j+1}^{-1}(\Delta'_j+\Delta''_j)Q_{j+1}\ldots Q_n] + \Delta''_n, \end{aligned} \tag{6.10}$$

$$\widehat{\Delta}_n = \Delta_0' Q_1 Q_2 \ldots Q_n + Q_1 Q_2 \ldots Q_n (\Delta_n' + \Delta_n'') + \sum_{j=1}^{n-1} Q_1 Q_2 \ldots Q_j (\Delta_j' + \Delta_j'') Q_{j+1} Q_{j+2} \ldots Q_n \qquad (6.11)$$

и с их помощью перепишем (6.8), (6.9) в виде

$$AC_n + \widehat{\Delta}_n = C_{n+1} R_{n+1}, \qquad A_n = C_n^{-1} A C_n + \widetilde{\Delta}_n^{(0)}. \qquad (6.12)$$

Из представлений (6.10), (6.11) и оценок (6.5), (6.7) при $7n\varepsilon/2 < 1$ следуют неравенства

$$\|\widetilde{\Delta}_n^{(0)}\| \leqslant \frac{\varepsilon\|A\|}{(1-\varepsilon)^n} + \sum_{j=1}^{n-1} \frac{1}{(1-\varepsilon)^{n-j}} \frac{2\varepsilon\|A\|}{(1-\varepsilon)^{3j}} + \frac{\varepsilon\|A\|}{(1-\varepsilon)^{3n}} < \frac{2n\varepsilon\|A\|}{(1-\varepsilon)^{3n}},$$

$$\begin{aligned}\|\widehat{\Delta}_n\| &\leqslant \frac{\varepsilon\|A\|}{(1-\varepsilon)^{n/2}} + \sum_{j=1}^{n-1} \frac{1}{(1-\varepsilon)^{n-j}} \frac{2\varepsilon\|A\|}{(1-\varepsilon)^{3j}} + \frac{\varepsilon\|A\|}{(1-\varepsilon)^{7n/2}} \\ &< \frac{(2n+1)\varepsilon\|A\|}{(1-\varepsilon)^{7n/2}} < \frac{(2n+1)\varepsilon\|A\|}{(1-\varepsilon)^{7n/2}}.\end{aligned}$$

Лемма 6.3. *Матрица C_n «почти ортогональна», т. е.*

$$\|C_n^* C_n - I\| \leqslant n\varepsilon/(1-n\varepsilon), \quad \text{где } n\varepsilon < 1. \qquad (6.13)$$

Доказательство. Из (6.3) вытекают равенства $Q_n^* - Q_n^{-1} = [(I+Z_n) - I]Q_n^{-1} = Z_n Q_n^{-1}$, в силу которых

$$\begin{aligned}&\|Q_n^* - Q_n^{-1}\| \leqslant \|Z_n\|\,\|Q_n^{-1}\| \leqslant \varepsilon/\sqrt{1-\varepsilon}, \\ &\|C_n\| = \|Q_1 Q_2 \ldots Q_n\| \leqslant 1/(1-\varepsilon)^{n/2}.\end{aligned} \qquad (6.14)$$

Преобразуем разность $C_n^* - C_n^{-1}$:

$$\begin{aligned}C_n^* - C_n^{-1} &= Q_n^* Q_{n-1}^* \ldots Q_1^* - Q_n^{-1} Q_{n-1}^{-1} \ldots Q_1^{-1} \\ &= (Q_n^* - Q_n^{-1}) Q_{n-1}^* Q_{n-2}^* \ldots Q_1^* + Q_n^{-1}(Q_{n-1}^* - Q_{n-1}^{-1}) Q_{n-2}^* \ldots Q_1^* \\ &\qquad + \ldots + Q_n^{-1} Q_{n-1}^{-1} \ldots Q_2^{-1}(Q_1^* - Q_1^{-1}) \\ &= Z_n Q_n^{-1} Q_{n-1}^* Q_{n-2}^* \ldots Q_1^* + Q_n^{-1} Z_{n-1} Q_{n-1}^{-1} Q_{n-2}^* Q_{n-3}^* \ldots Q_1^* \\ &\qquad + \ldots + Q_n^{-1} Q_{n-1}^{-1} \ldots Q_2^{-1} Z_1 Q_1^{-1}.\end{aligned}$$

С помощью (6.3), (6.7) получаем неравенство $\|C_n^* - C_n^{-1}\| \leqslant n\varepsilon/(1-\varepsilon)^{n/2}$, которое вместе с (6.14) приводит к требуемой оценке (6.13).

Обозначив $\widetilde{\Delta}_n = -\widetilde{\Delta}_n^{(0)} + (C_n^* - C_n^{-1}) A C_n$, перепишем второе равенство в (6.12) в виде $C_n^* A C_n = A_n + \widetilde{\Delta}_n$. Ясно, что при $3n\varepsilon < 1$

$$\|\widetilde{\Delta}_n\| < \left[\frac{2n}{(1-\varepsilon)^{3n}} + \frac{n}{(1-\varepsilon)^n}\right] \varepsilon\,\|A\| < \frac{3n\varepsilon\,\|A\|}{(1-\varepsilon)^{3n}}.$$

Таким образом, мы доказали, что матрица A_n, полученная в результате n шагов QR-алгоритма (см. (6.1)-(6.3)), при достаточно малом ε ($\varepsilon \ll 1/(4n)$) мало отличается от матрицы $C_n^* A C_n$, полученной в результате применения ортогонально-степенного метода, именно:

$$\|C_n^* A C_n - A_n\| < 3n\varepsilon \, \|A\|/(1-\varepsilon)^{3n}.$$

При этом матрицы C_n определяются по следующим формулам:

$$A + \widehat{\Delta}_0 = C_1 R_1,$$

$$AC_j + \widehat{\Delta}_j = C_{j+1} R_{j+1}, \qquad \|C_j^* C_j - I\| \leqslant n\varepsilon/(1-n\varepsilon),$$

$$\|\widehat{\Delta}_j\| \leqslant (2n+1)\varepsilon \, \|A\|/(1-4n\varepsilon) \qquad (1 \leqslant j \leqslant n-1).$$

Сформулированное утверждение позволяет воспользоваться результатами проведенного в § 1–5 анализа ортогонально-степенного метода для исследования матриц A_n, получаемых при выполнении QR-алгоритма.

ЛИТЕРАТУРА

1. **Годунов С. К., Кирилюк О. П., Костин В. И.** Спектральные портреты матриц. Новосибирск, 1990. (Препринт / АН СССР. Сиб. отд-ние. Ин-т математики; № 3).
2. **Kostin V. I.** On definition of matrices' spectra // Proc. of the Second Internat. Sympos. at Montpelier. Amsterdam: North-Holland, 1991. P. 407–414.
3. **Малышев А. Н.** Введение в вычислительную линейную алгебру. Новосибирск: Наука, 1991.
4. **Годунов С. К.** Гарантированная точность в несимметричных спектральных задачах // Дополнение к книге: Малышев А. Н. Введение в вычислительную линейную алгебру. Новосибирск: Наука, 1991. С. 204–223.
5. **Годунов С. К.** Задача о дихотомии спектра матрицы // Сиб. мат. журн. 1986. Т. 27, № 5. С. 24–37.
6. **Булгаков А. Я., Годунов С. К.** Круговая дихотомия матричного спектра // Сиб. мат. журн. 1988. Т. 29, № 5. С. 59–70.
7. **Годунов С. К.** Критерии сходимости ортогонально-степенных методов спектрального анализа матриц // Тр. Ин-та математики /АН СССР. Сиб. отд-ние. 1989. Т. 15. С. 4–12.
8. **Fransis J. D. F.** The QR-transformation. The unitary analogue to the LR-transformation. I // Comput. J. 1961. V. 4, N 3. P. 265–271.
9. **Fransis J. D. F.** The QR-transformation. The unitary analogue to the LR-transformation. II // Comput. J. 1962. V. 4, N 4. P. 332–345.
10. **Кублановская В. Н.** О некоторых алгоритмах для решения полной проблемы собственных значений // Журн. вычисл. математики и мат. физики. 1961. Т. 1, № 4. С. 555-570.
11. **Воеводин В. В.** Некоторые методы решения полной проблемы собственных значений // Журн. вычисл. математики и мат. физики. 1961. Т. 2, № 1. С. 15–24.
12. **Годунов С. К.** О численном решении краевых задач для систем линейных обыкновенных дифференциальных уравнений // Успехи мат. наук. 1961. Вып. 3. С. 171–173.
13. **Годунов С. К.** Решение систем линейных уравнений. Новосибирск: Наука, 1980.
14. **Годунов С. К., Антонов А. Г., Кирилюк О. П., Костин В. И.** Гарантированная точность решения систем линейных уравнений в евклидовых пространствах. Новосибирск: Наука, 1988.