

Пример 2. Для примера 2 § 3 получены следующие значения норм невязок:

$$\begin{aligned} c_1 &= 2; & c_2 &= 1.732; & c_3 &= 1.414; \\ c_4 &= 0.974; & c_5 &= 0.974; & c_6 &= 0.458; \\ c_7 &= 0.294_{10} - 2; & c_8 &= 0.155_{10} - 3; & c_9 &= 0.679_{10} - 5. \end{aligned}$$

Ясно, что невязка постоянно убывает.

ЛИТЕРАТУРА

1. Булгаков А. Я., Годунов С. К. Численное определение одного из критериев качества устойчивости систем линейных дифференциальных уравнений с постоянными коэффициентами. — Новосибирск, 1981. — 58 с. (Препринт/АН СССР, Сиб. отделение, ИМ).
2. Годунов С. К., Прокопов Г. П. Вариационный подход к решению больших систем линейных уравнений, возникающих в сильно эллиптических задачах. — М., 1968. — 40 с. (Препринт/АН СССР, ИИМ).
3. Воеводин В. В. О методах сопряженных градиентов. — Журн. вычисл. математики и мат. физики, 1979, т. 19, № 5, с. 1313—1317.
4. Hestenes M. R., Stiefel E. Methods of conjugate gradients for solving linear systems. — Nat. Bur. Standards. J. Res., 1952, v. 49, p. 409—436.
5. Федоренко Р. П. Приближенное решение задач оптимального управления. — М.: Наука, 1978. — 488 с.
6. Фаддеев Д. К., Фаддеева В. Н. Вычислительные методы линейной алгебры. — М.—Л.: Физматгиз, 1963. — 734 с.
7. Годунов С. К. Решение систем линейных уравнений. — Новосибирск: Наука, Сиб. отделение, 1980. — 177 с.
8. Форсайт Дж., Моулер К. Численное решение систем линейных алгебраических уравнений. — М.: Мир, 1969. — 167 с.
9. Булгаков А. Я. Вычисление экспоненты от асимптотически устойчивой матрицы. — В кн.: Вычислительные методы линейной алгебры. Новосибирск: Наука, 1985, с. 4—17.

О СХОДИМОСТИ ОРТОГОНАЛЬНО-СТЕПЕННОГО МЕТОДА РАСЧЕТА СПЕКТРА

В. И. КОСТИН, Ш. И. РАЗЗАКОВ

ВВЕДЕНИЕ

В работе обсуждается вопрос об устойчивости и эффективных оценках скорости сходимости ортогонально-степенного метода Воеводина, применяемого для решения полной проблемы собственных значений.

Под полной проблемой собственных значений понимается задача нахождения всех собственных значений матрицы A . Часто при этом пытаются найти и соответствующие им собственные и присоединенные векторы. Достаточно обширная библиография по решению полной проблемы содержится в монографиях и статьях В. В. Воеводина [1, 2], В. Н. Кублановской [3], В. Н. Фаддеевой [4, 5]. Напомним, что собственными значениями матрицы A называются корни ее характеристического полинома, т. е. корни уравнения

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} = \\ &= (-1)^n (\lambda^n - p_1 \lambda^{n-1} - p_2 \lambda^{n-2} - \dots - p_n) = 0. \end{aligned}$$

Все методы численного решения полной проблемы можно разделить на две группы: прямые и итерационные. Большинство прямых методов

включает в качестве обязательного этапа вычисление коэффициентов характеристического полинома $P(\lambda)$. Как известно, соответствующие p_i являются, с точностью до знака, суммами всех главных миноров порядка i матрицы A . Непосредственное вычисление коэффициентов p_i чрезвычайно громоздко и требует огромного числа операций для матриц, размерность которых превосходит 10—15. Поэтому совершенно естественно появление специальных вычислительных приемов, упрощающих определение таких коэффициентов. Наиболее употребительные из прямых методов основаны на том, что с помощью конечного числа подобных преобразований исходная матрица приводится к матрице более простого вида, для которой легко находятся коэффициенты характеристического многочлена. Далее для определения собственных значений используются какие-либо методы приближенного вычисления корней полинома.

Основное достоинство прямых методов — большая скорость работы, единообразная схема вычислений для матрицы произвольного вида. Слабость прямых методов — в их численной неустойчивости, что почти целиком является следствием того, что корни многих полиномов чрезвычайно чувствительны к малым относительным возмущениям коэффициентов. Это верно не только для полиномов, имеющих кратные корни, но и для полиномов, расположение корней которых, на первый взгляд, кажется очень хорошим (см., например, [6]).

Практическое применение прямых методов, ввиду их вычислительной неустойчивости, ограничено лишь некоторыми классами специальных матриц. Итерационные методы в отличие от прямых более устойчивы к возмущениям элементов матрицы. В итерационных методах собственные значения вычисляются как пределы некоторых числовых последовательностей без предварительного определения коэффициентов характеристического многочлена. При этом в большинстве случаев одновременно вычисляются собственные векторы или некоторые другие векторы, связанные с собственными простыми соотношениями.

В настоящее время созданы эффективные итерационные методы решения полной проблемы собственных значений эрмитовой матрицы (см. [7]) и нормальной матрицы (см. [8]). Ряд работ посвящен разработке и теоретическому обоснованию методов решения полной проблемы собственных значений произвольных матриц, т. е. таких матриц, у которых допускалось бы как произвольное распределение собственных значений, так и наличие жордановых клеток. Так, в работах Дж. Дж. Френсиса (см. [9, 10]) и В. Н. Кублановской [11] предложен QR -алгоритм. В. В. Воеводины разработан метод (см. [12]), который обладает теми же характеристиками сходимости, что и QR -алгоритм. Однако в отличие от QR -алгоритма метод Воеводина является самоисправляющимся, т. е. накапливающаяся за шаг погрешность учитывается только один раз, ибо каждое приближение можно считать начальным. Это обстоятельство и привело к тому, что задача настоящей работы состоит в получении квалифицированных оценок для ортогонально-степенного метода Воеводина. Вместе с тем надо отметить, что приведенный метод получения оценок пригоден также и для исследования QR -алгоритма (см. § 7).

Итерационные методы требуют для своей реализации выполнения очень большого объема вычислительной работы. Даже для матриц небольшого порядка этот объем настолько велик, что без помощи ЭВМ выполнить его невозможно.

Применяя ЭВМ для решения задач линейной алгебры, следует учитывать структуру разрядной сетки, используемой для представления чисел в машине. Эта структура диктует «принцип неопределенности», в силу которого каждое число, каждый вектор и каждая матрица, хранящиеся в машине или участвующие в промежуточных вычислениях, не отличимы от достаточно близких к ним. Детальное рассмотрение вопросов, возникающих при решении на ЭВМ задач линейной алгебры с учетом разрядной сетки ЭВМ, можно найти в [13—15]. Для характеристики разрядной сетки ЭВМ, как правило, используются две машинные

постоянные. Обозначим их, следуя [13], через ε_1 и ε_2 . ε_2 — машинная постоянная такая, что модуль любого машинного числа заключен в пределах $\rho_{\min}$, $\rho_{\max}$ таких, что $\rho_{\min} \leq \varepsilon_2$, $\rho_{\max} \geq 1/\varepsilon_2$. ε_1 — машинная постоянная такая, что наименьшее машинное число, большее единицы, превосходит $1 + \varepsilon_1$. Машина не различает (см., например, [13]) числа a и b , если

$$|a - b| \leq \frac{1}{2} \varepsilon_1 |a| + \varepsilon_2.$$

Следовательно, возникает вопрос об устойчивости проблемы собственных значений. То, что в отдельных случаях проблема собственных значений может быть неустойчивой, ясно из следующих соображений. Известны примеры матриц, имеющих лишь простые собственные значения, однако при некотором определенном изменении матричных элементов в пределах точности машинного их задания получается матрица, имеющая кратное собственное значение с нелинейным элементарным делителем (см., например, [14], с. 93). В этом случае каноническая форма матрицы при изменении ее элементов в пределах точности задания претерпевает качественное изменение, переходя от чисто диагональной формы к общей канонической.

Более того, при работе на ЭВМ возникают ситуации, когда конечномерный оператор-матрица оказывается эквивалентен по своим спектральным свойствам бесконечномерным операторам. У таких матриц вместо дискретного спектра приходится рассматривать непрерывный спектр (точнее, «почти спектр»), который заполняет двумерные области на плоскости. Например, в работах [16, 17] рассматривается однопараметрическое семейство матриц размерности 25×25

$$A(w) = \begin{bmatrix} -1 & 10 & & & & \\ & -1 & 10 & & & \\ & & -1 & 10 & & 0 \\ \vdots & & & 0 & \ddots & \vdots \\ & & & & & -1 & 10 \\ w & & & & \dots & & -1 \end{bmatrix}.$$

Здесь по главной диагонали стоят -1 , а по наддиагонали числа 10. Все остальные элементы равны 0, за исключением элемента в позиции $(25, 1)$, равного малому числу w . Характеристическое уравнение матрицы $A(w)$ есть

$$\Psi(w, \lambda) = \det[A - \lambda I] = (-1 - \lambda)^{25} + 10^{24}w = 0.$$

Из него непосредственно следует, что если задать $w = 10^{-24}(\alpha e^{i\varphi})^{25}$, где $0 \leq \alpha \leq r < 1$, $0 \leq \varphi < 2\pi$, то собственные числа матрицы $A(w)$ заполняют круг радиуса r , в то время как машина БЭСМ-6, например, не различает матрицы $A(w)$ и $A(0)$.

Указанные примеры показывают, что понятия собственных значений и соответствующих им собственных векторов для некоторых несимметрических матриц не являются устойчивыми по отношению к малым возмущениям элементов матрицы. В связи с этим интересен вопрос о том, как QR-алгоритм и метод Воеводина будут вести себя при их применении к матрицам, обладающим «сплошным» или, точнее, «почти сплошным» спектром. В работе дается частичный ответ на этот вопрос. Для определенности мы провели исследование до конца в несколько более простом случае (метод Воеводина). Оказывается, что если спектр может быть разделен на две части окружностью, не содержащей точек «почти спектра», то метод Воеводина сходится к клеточно-треугольной матрице, диагональные клетки которой имеют спектры по разные стороны от этой окружности. Правда, нужно еще предположить, что ин-

вариантные подпространства рассматриваемой матрицы находятся в «общем положении» по отношению к координатным осям.

Очевидно, что проводя изучение поведения процесса для матриц с патологическим сплошным «почти спектром», не слишком удобно пользоваться жордановой нормальной формой. Отсюда несколько необычная формулировка предположений, на основе которых ведутся доказательства. Такие формулировки приводят к довольно сложным рассуждениям, гораздо более длинным, чем в первоначальном доказательстве (см. [12]). Однако именно благодаря им удалось разобраться в теории метода при его применении к матрицам с патологическими «почти собственными значениями» (см. [18]).

Напомним вычислительную схему метода Воеводина: для матрицы A строится последовательность унитарных матриц Q_j ($Q_j^* Q_j = I$) и правых треугольных матриц R_j ($r_{il}^{(j)} = 0$ при $i > l$) так, чтобы выполнялись формулы

$$\begin{aligned} A &= Q_1 R_1, \\ A Q_1 &= Q_2 R_2, \\ A Q_2 &= Q_3 R_3, \\ &\dots \\ A Q_{j-1} &= Q_j R_j \end{aligned} \quad (1)$$

и определяются матрицы

$$A_j = Q_j^* A Q_j = \begin{pmatrix} B^{(j)} & D^{(j)} \\ S^{(j)} & C^{(j)} \end{pmatrix}.$$

В работе В. В. Воеводина [12] доказано, что если выполнены следующие два предположения: а) собственные значения матрицы A удовлетворяют неравенствам

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_l| > |\lambda_{l+1}| \geq \dots \geq |\lambda_N| \geq 0;$$

б) прямая сумма корневых подпространств, отвечающих собственным значениям $\lambda_{l+1}, \lambda_{l+2}, \dots, \lambda_N$, имеет нулевое пересечение с подпространством, натянутым на координатные векторы $f_1, f_2, \dots, f_l$, то элементы клетки $S^{(j)}$ матрицы A_j сходятся к нулю как величины

$$S_{ik}^{(j)} = O \left(j^{2N(t-1)} \left(\frac{|\lambda_l|}{|\lambda_h|} \right)^j \right),$$

где t — размерность наибольшей жордановой клетки.

Основной вопрос работы — выбор критериев, эквивалентных а) и б) на некотором классе матриц, который допускает наличие пятен сплошного «почти спектра».

Комплексное число λ называется почти собственным значением оператора A , если для любого положительного ε существует вектор x , удовлетворяющий неравенству

$$\|Ax - \lambda x\| \leq \varepsilon \|x\|,$$

а соответствующий x называется почти собственным вектором.

При машинной реализации алгоритмов вычисления почти собственного значения естественно считать λ почти собственным значением матрицы A , если существует вектор x такой, что $\|Ax - \lambda x\| \leq \varepsilon_0 \|x\|$, где ε_0 — фиксированная малая величина, определяемая разрядной сеткой машины.

Гипотезу о разделении спектра окружностью сформулируем как утверждение об отсутствии на этой окружности точек почти спектра.

1. Гипотеза о разделении спектра. Существуют числа ρ, δ ($0 < \rho < 1, 0 < \delta < \rho$) такие, что

$$\min_{\|x\| \neq 0} \frac{\|(A - \rho \|A\| e^{i\varphi} I) x\|}{\|x\|} > \delta \|A\|$$

для любого вектора x и любого аргумента φ . Это условие означает, что на кольце $\rho\|A\|(1 - \delta/2\rho) < |\lambda| < \rho\|A\|(1 + \delta/2\rho)$ нет почти собственных значений матрицы A (предполагаем, что $\delta \gg \varepsilon_0$).

Пусть $P(\lambda) = \det(A - \lambda I)$. По гипотезе I $\det(A - \lambda I)$ не обращается в нуль при $|\lambda| = \rho\|A\|$. Поэтому определено число k , равное изменению аргумента $P(\lambda)$ при положительном обходе параметром λ окружности $|\lambda| = \rho\|A\|$. Хорошо известно, что k равно числу характеристических корней A , лежащих в круге $|\lambda| < \rho\|A\|$.

Заметим, что при малых изменениях матрицы A , вызванных, например, ошибками округления, число k не будет меняться (см. подробнее в § 6).

Обозначим через $\mathcal{L}_1^1$ линейную оболочку первых $N - k$ базисных векторов N -мерного пространства, в котором действует линейное преобразование, определяемое матрицей A , а через $\mathcal{L}_1$ и $\mathcal{L}_2$ инвариантные подпространства матрицы A , отвечающие собственным значениям, лежащим соответственно вне и внутри круга $|\lambda| = \rho\|A\|$. Напомним, что подпространство $\mathcal{L}_1$ пространства $\mathcal{L}$ называется инвариантным относительно A , если для каждого вектора x из $\mathcal{L}_1$ образ Ax лежит в $\mathcal{L}_1$.

Дадим теперь формулировку того, что $\mathcal{L}_1^1$ пересекается с $\mathcal{L}_2$ лишь в начале координат. Это требование и есть требование «общего положения» по отношению к координатным осям. Формулировку приведем опять в виде неравенств для сингулярных чисел некоторой прямоугольной матрицы, зависящей от параметров.

Пусть $\hat{A}(\lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k)})$ — матрица, составленная из первых $N - k$ столбцов матрицы произведения $(A - \lambda^{(1)}I)(A - \lambda^{(2)}I) \dots (A - \lambda^{(k)}I)$ и $\sigma_1(\hat{A}(\lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k)}))$ — ее наименьшее сингулярное число.

II. Гипотеза «общего положения базиса» формулируется в виде неравенства

$$\sigma_1(\hat{A}(\lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k)})) \geq (\delta_1 \|A\|)^k$$

для любых $\lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k)}$ таких, что $|\lambda^{(j)}| \leq \rho\|A\|$. Из этой гипотезы следует, что $\mathcal{L}_1^1$ пересекается с $\mathcal{L}_2$ лишь в начале координат. Опять-таки можно показать, что из выполнения этой гипотезы для некоторой A следует ее справедливость для всех близких матриц (см. подробнее в § 6). Правда, при этом δ_1 придется заменить на несколько меньшее число, например, на $\delta_1/2$.

К сожалению, указанные гипотезы — как одна, так и другая — не допускают эффективной проверки. Несмотря на это их введение позволило понять поведение алгоритма в довольно сложной нестандартной ситуации.

Приведем интересный, на наш взгляд, пример матрицы, для которой заведомо не выполнена I гипотеза.

При описании примера будем для краткости называть ε -спектром матрицы A и обозначать $\Sigma_\varepsilon(A)$ множество ее ε -собственных значений.

Пусть матрица A размера 100×100 имеет блочно-диагональный вид

$$A = \begin{bmatrix} B^{(1)} & 0 \\ 0 & B^{(2)} \end{bmatrix},$$

в котором $B^{(1)}$ и $B^{(2)}$ — двухдиагональные матрицы размера 50×50

$$B^{(j)} = \begin{bmatrix} d_1^{(j)} & b_2^{(j)} & & 0 \\ & d_2^{(j)} & b_3^{(j)} & \\ & & \ddots & \\ 0 & & & d_{50}^{(j)} \end{bmatrix} \quad (j = 1, 2).$$

В нашем примере диагональные элементы матрицы $B^{(1)}$ — это комплексные числа, расположенные равномерно на объединении отрезков

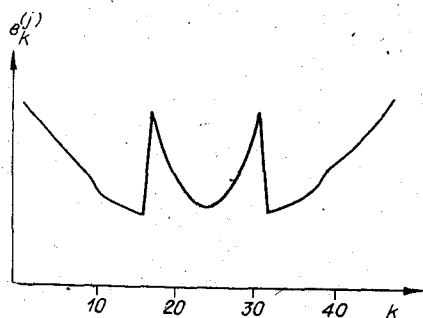


Рис. 1.

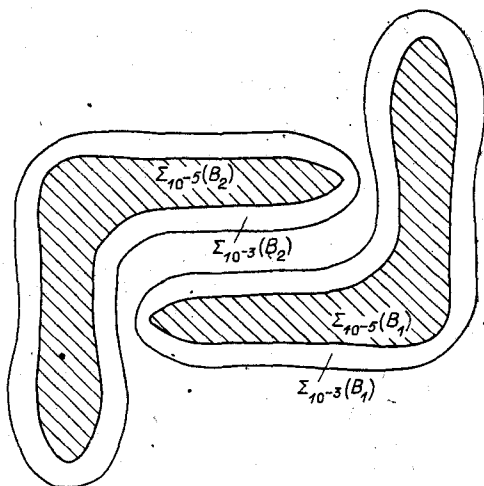


Рис. 2.

$[-1, 0] \cup [0, i]$ на комплексной плоскости,

$$d_k^{(1)} = \begin{cases} -1 + (k-1) \cdot 0,04, & 1 \leq k \leq 25; \\ i \cdot (k-25) \cdot 0,04, & 26 \leq k \leq 50. \end{cases}$$

Диагональные элементы матрицы $B^{(2)}$ получаются по правилу

$$d_k^{(2)} = -1,3 + 0,5i - d_k^{(1)}.$$

Внедиагональные элементы $b_k^{(j)}$ ($j=1, 2; k=2, 3, \dots, 50$) в обеих матрицах одинаковы и строятся по следующим формулам:

$$b_k^{(j)} = \begin{cases} (0,1 + (k-25)^2 \cdot 0,001) \cdot (0,5 + (k-5)^2 \cdot 0,04), & 2 \leq k \leq 10; \\ 0,1 + (k-25)^2 \cdot 0,001, & 11 \leq k \leq 17; \\ 0,2 + (k-25)^2 \cdot 0,011, & 18 \leq k \leq 32; \\ 0,1 + (k-25)^2 \cdot 0,001, & 33 \leq k \leq 40; \\ (0,1 + (k-25)^2 \cdot 0,001) \cdot (0,5 + (k-5)^2 \cdot 0,04), & 41 \leq k \leq 50. \end{cases}$$

Поведение $b_k^{(j)}$ в зависимости от k изображено на рис. 1.

На рис. 2 показаны полученные «пятна» ϵ -спектров матриц $B^{(1)}$ и $B^{(2)}$, объединение которых составляет ϵ -спектр матрицы A . Заштрихованная область соответствует $\epsilon = 10^{-5}$, а объемлющая ее область $\epsilon = 10^{-3}$. Расположение областей $\Sigma_{10^{-5}}(B^{(j)})$ и $\Sigma_{10^{-3}}(B^{(j)})$ позволяет сделать вывод, что точки, принадлежащие $\Sigma_{10^{-5}}(B^{(1)})$ с точностью, не меньшей 10^{-3} , не являются почти собственными значениями блока $B^{(2)}$, и наоборот. Иными словами, почти собственные значения матрицы A , вычисленные с точностью 10^{-5} , оказываются с точностью 10^{-3} разбитыми на две группы или, лучше сказать, на два «пятна». Причем геометрически очевидно, что разбиение ϵ -спектра матрицы A на два пятна не может быть произведено никакой окружностью.

Из определения ϵ -спектра матрицы очевидно, что ϵ -спектры унитарно подобных матриц совпадают, чего, однако, нельзя сказать о просто подобных матрицах. Таким образом, на рис. 2 изображен 10^{-5} -спектр всех матриц вида U^*AU , где A — построенная выше матрица 100-го порядка, а $U^*U = I$. Среди таких матриц, очевидно, имеются матрицы, не разбивающиеся на блоки.

Таким образом, матрицы, унитарно подобные матрице A , представляют собой примеры матриц, для которых не выполнена гипотеза I, хотя, с другой стороны, ϵ -спектры этих матриц при $\epsilon = 10^{-5}$ разбиваются на две группы.

Перейдем к более детальному изложению основных результатов работы, состоящей из восьми параграфов. Первые три параграфа носят вспомогательный характер, хотя представляют и самостоятельный интерес. В § 2, основываясь на гипотезе I, строится преобразующая матрица T , которая приводит исходную матрицу A к клеточно-диагональному виду:

$$TAT^{-1} = \hat{A} = \begin{bmatrix} \hat{B} & 0 \\ 0 & \hat{C} \end{bmatrix}.$$

В этом представлении матрица $\hat{B}$ имеет размерность $(N-k) \times (N-k)$, и для ее наименьшего сингулярного числа выполнена оценка

$$\sigma_1(\hat{B}) \geq \rho \|A\| \sqrt{1 + 2\delta_2^3}, \quad \delta_2 = \delta/2\rho.$$

Матрица $\hat{C}$ имеет размерность $k \times k$, и для ее сингулярного числа выполнена оценка

$$\sigma_N(\hat{C}) \leq \rho \|A\| \sqrt{1 - \delta_2^3}.$$

Следовательно, матрица T обладает таким свойством, что если пятна почти спектра исходной матрицы разделены окружностью, то она приводит матрицу A к блочно-диагональному виду. При этом сингулярные числа матриц $\hat{B}$ и $\hat{C}$ будут лежать по разные стороны этой окружности.

Кроме того, будет оценено число обусловленности преобразующей матрицы:

$$\mu(T) = \|T\| \cdot \|T^{-1}\| \leq 78\delta^{-7.5}.$$

При построении матрицы T используются вспомогательные леммы § 1.

Основываясь на гипотезе II, в § 3 показано, что $\mathcal{L}_1^1$ пересекается с $\mathcal{L}_2$ только в начале координат. Последнее утверждение является следствием теоремы, в котором доказано, что пространство $\mathcal{L}_1^1$ векторов вида $f_i = (f_1, f_2, \dots, f_{N-k}, 0, \dots, 0)'$ (штрих означает транспонирование) преобразуется преобразованием T в пространство $\mathcal{L}_1^1 = T\mathcal{L}_1^1$ такое, что для любого $l \in \mathcal{L}_1^1$, представленного как $l = (l_1, \dots, l_{N-k}, l_{N-k+1}, \dots, l_N)'$, выполняется оценка

$$\sqrt{|l_1|^2 + |l_2|^2 + \dots + |l_{N-k}|^2} \geq \frac{1}{\mu(T)} \left(\frac{\delta_1 \delta^{1.5}}{2} \right)^k \sqrt{|l_{N-k+1}|^2 + \dots + |l_N|^2}.$$

Если ввести понятие угла между подпространствами, то последнее неравенство означает, что

$$\text{tg} \angle (\mathcal{L}_1^1, \mathcal{L}_2) \geq \frac{(\delta_1 \delta^{1.5})^k}{2^k \mu(T)}.$$

На понятии угла между подпространствами более подробно остановимся в § 3, а пока напомним, что углом между подпространствами S и R назовем угол φ ($0 \leq \varphi \leq \pi/2$), определенный из равенства

$$\cos \varphi = \max_{r \in R, s \in S} \frac{|(r, s)|}{\|r\| \cdot \|s\|}.$$

Сформулируем исследуемый алгоритм с учетом вычислительных погрешностей. Как ранее отмечалось, метод Воеводина основан на разложении невырожденной матрицы в произведение унитарной матрицы и правой треугольной. В линейной алгебре имеются весьма эффективные алгоритмы, осуществляющие данное разложение. Наиболее популярное изложение алгоритмов с учетом вычислительных погрешностей можно найти в монографиях В. В. Воеводина [15], С. К. Годунова [13], Дж. Х. Уилкинсона [14].

Если учесть погрешности E_j , допускаемые при вычислении данного разложения и при перемножении матриц AQ_{j-1} , то вместо алгоритма (1) следует рассмотреть

$$\begin{aligned} AQ_0 &= Q_1 R_1 + E_1; \\ AQ_1 &= Q_2 R_2 + E_2; \\ &\dots \\ AQ_{j-1} &= Q_j R_j + E_j; \end{aligned} \quad (2)$$

$$A_j = Q_j^* A Q_j = Q_j^{-1} (A + H_j A) Q_j = (I + Q_j^{-1} H_j Q_j) \begin{bmatrix} B_j & D_j \\ S_j & C_j \end{bmatrix},$$

где $\|E_j\| \leq \varepsilon$, $\|Q_j^* Q_j - I\| = \|H_j\| < \varepsilon_j$; $\varepsilon, \varepsilon_j > 0$.

Для оценки скорости сходимости этого алгоритма поступим следующим образом. Для того чтобы матрица A_j была клеточно-треугольной (или почти клеточно-треугольной), достаточно, чтобы первые $N-k$ вектор-столбцов матрицы Q_j лежали (или почти лежали) на инвариантном подпространстве $\mathcal{L}_1$. Чтобы установить описанное свойство Q_j , достаточно показать малость углов φ_j , определенных из равенства

$$\cos \varphi_j = \max_{x \in \mathcal{L}_1} \frac{|(z, x)|}{\|z\| \cdot \|x\|}$$

для всех $z \in \mathcal{L}_j^{(N-k)} (\mathcal{L}_j^{(N-k)})$ — линейная оболочка первых $N-k$ вектор-столбцов матрицы Q_j .

Покажем, что $\sin \varphi_j \leq cL^j + \eta$, где $L < 1$, η — величина порядка $\varepsilon \cdot \delta^{-4,5}$ и c — константа.

Из (2) легко получается представление

$$(A + E_j)(A + E_{j-1}) \dots (A + E_1) Q_0 = Q_j R_j R_{j-1} \dots R_1,$$

откуда следует, что матрицы Q_j являются результатом разложения матрицы $(A + E_j)(A + E_{j-1}) \dots (A + E_2)(A + E_1)$ в произведение ортогональной матрицы Q_j на правую треугольную матрицу $R_j R_{j-1} \dots R_1$.

Следовательно, для того чтобы исследовать поведение первых $N-k$ столбцов матрицы Q_j , достаточно оценить углы φ_j между векторами y_j и подпространством $\mathcal{L}_1$.

Здесь

$$y_j = (\hat{A} + \hat{E}_j)(\hat{A} + \hat{E}_{j-1}) \dots (\hat{A} + \hat{E}_1) y_0,$$

$$\hat{A} = TAT^{-1}, \quad \|\hat{E}_j\| = \|TE_j T^{-1}\| \leq \mu(T) \cdot \varepsilon, \quad y_0 \in \mathcal{L}_1^1.$$

Такая оценка получена в § 4, где неравенство $\sin \varphi_j \leq cL^j + \eta$ доказывается в следующей конкретной форме:

$$\sin \varphi_j \leq \mu(T) \left[\left(\frac{2}{\delta_1 \delta^{1,5}} \right)^k L^j + \delta^6 \eta(\varepsilon) \right],$$

где

$$L = \frac{\rho \|A\| \sqrt{1 - \delta_2^2} + \mu(T) \cdot \varepsilon (\delta^3 + \delta^6 \operatorname{ctg} \varphi_0)}{\rho \|A\| \sqrt{1 + 2\delta_2^2} - \mu(T) \cdot \varepsilon (\delta^{1,5} + \operatorname{tg} \varphi_0)},$$

$$\eta(\varepsilon) = \frac{\varepsilon}{2\delta_2^3 (\rho \|A\| - \varepsilon \mu(T) + \sqrt{(\rho \|A\| - \varepsilon \mu(T))^2 - (\mu(T) \varepsilon \delta^2)})}.$$

Исследованию самого алгоритма посвящен § 5, в котором, основываясь на результатах § 4, оцениваются нормы матрицы S_j . Показано, что

$$\|S_j\| \leq 2\mu(T)(1 + \varepsilon_s) \|A\| \left[\left(\frac{2}{\delta_1 \delta^{1,5}} \right)^k L^j + \delta^6 \eta(\varepsilon) \right]$$

и $L < 1$, если $\mu(T) \delta^6 \eta(\varepsilon) < \operatorname{tg} \varphi_0 < (\mu(T) \eta(\varepsilon))^{-1}$.

Неравенство $\mu(T)\delta^e\eta(\epsilon) < \operatorname{tg} \varphi_0$ можно считать всегда выполненным; если это не так, то мы сразу получим $\sin \varphi_0 < \mu(T)\delta^e\eta(\epsilon)$. Для того чтобы удовлетворять второму неравенству, достаточно положить $\delta_1 = 2\delta^{-1.5} \cdot \sqrt{\mu^2(T)\eta(\epsilon)}$.

В § 6 доказывается непрерывная зависимость параметров δ и δ_1 от матрицы A .

Изучение полученных оценок в § 5—7 позволяет утверждать, что класс матриц, свойства которого даются основной теоремой (§ 5), является устойчивым по отношению к малым возмущениям элементов.

В § 7 исследован QR -алгоритм, в котором для обычного QR -алгоритма (без сдвигов) получены оценки скорости сходимости. В § 8 приведены результаты численных экспериментов.

§ 1. ВСПОМОГАТЕЛЬНЫЕ ЛЕММЫ

Пусть A — произвольная комплексная невырожденная матрица порядка N , и для нее выполнена гипотеза I, а именно: существуют числа ρ, δ ($0 < \delta < \rho < 1$) такие, что

$$\min_{\|x\| \neq 0} \frac{\|(A - \rho \|A\| e^{i\varphi} I)x\|}{\|x\|} > \delta \|A\|$$

для любого аргумента φ .

Пусть $\lambda_1, \lambda_2, \dots, \lambda_N$ — собственные значения матрицы A . По предположению матрица A не имеет собственных значений на кольце $\rho \|A\| (1 - \delta_2) \leq |\lambda| \leq \rho \|A\| (1 + \delta_2)$, $\delta_2 = \delta/2\rho$. Пусть y — внутри круга $|\lambda| < \rho \|A\|$ лежат k собственных значений $\lambda_1, \lambda_2, \dots, \lambda_k$, а вне этого круга остальные $\lambda_{k+1}, \lambda_{k+2}, \dots, \lambda_N$. Любая такая матрица может быть с помощью некоторой унитарной U ($U^*U = I$) записана в виде

$$A = U \begin{bmatrix} B & D \\ 0 & C \end{bmatrix} U^* = U \tilde{A} U^* \quad (1.1)$$

так, что B и C — квадратные матрицы размерностей $(N-k) \times (N-k)$ и $k \times k$, имеющие все собственные значения соответственно вне и внутри круга $|\lambda| = \rho \|A\|$. Разумеется, что при этом собственные значения матрицы B суть $\lambda_{k+1}, \lambda_{k+2}, \dots, \lambda_N$, а собственные значения матрицы C суть $\lambda_1, \lambda_2, \dots, \lambda_k$. Прямоугольная матрица D имеет $N-k$ строк и k столбцов.

Сформулированное утверждение следует из известной леммы Шура (см. [19], с. 242) о возможности приведения любой квадратной матрицы к треугольному виду с помощью подобного унитарного преобразования. Очевидно, что

$$|\lambda_i(B)| \geq \rho \|A\| (1 + \delta_2), \quad |\lambda_j(C)| \leq \rho \|A\| (1 - \delta_2). \quad (1.2)$$

Для матриц B и C верна следующая лемма.

Лемма 4. Если выполнена гипотеза I, то имеют место оценки

$$\|B^{-l}\| \leq 1/[\delta_2(\rho \|A\| (1 + \delta_2))^l], \quad (1.3)$$

$$\|C^l\| \leq (1/\delta_2)(\rho \|A\| (1 - \delta_2))^l, \quad (1.4)$$

где $l = 1, 2, 3, \dots$

Доказательство. Каждой исследуемой матрице A сопоставим разностное уравнение

$$G_{l+1} - (\rho \|A\|)^{-1} A G_l = \delta_0^l I, \quad (1.5)$$

где

$$\delta_0^l = \begin{cases} 1 & \text{при } l = 0; \\ 0 & \text{при } l \neq 0. \end{cases}$$

Ограниченное решение системы (1.5) представимо в виде

$$G_l = \begin{cases} \frac{1}{2\pi i} \int_{|\lambda|=1-\delta_2} \lambda^l [(\rho \|A\|)^{-1} A - \lambda I]^{-1} d\lambda & \text{при } l \geq 0; \\ \frac{1}{2\pi i} \int_{|\lambda|=1+\delta_2} \lambda^l [(\rho \|A\|)^{-1} A - \lambda I]^{-1} d\lambda & \text{при } l < 0. \end{cases} \quad (1.6)$$

Из (1.6) и гипотезы I, в силу которой $\|[(\rho \|A\|)^{-1} A - \lambda I]^{-1}\| < 1/\delta_2$ для всех λ из кольца $1 - \delta_2 < |\lambda| < 1 + \delta_2$, получим оценку

$$\|G_l\| \leq \begin{cases} (1/\delta_2)(1 - \delta_2)^l & \text{при } l > 0; \\ (1/\delta_2)(1 + \delta_2)^l & \text{при } l < 0. \end{cases} \quad (1.7)$$

Пусть матрица $\tilde{G}_l$ связана с матрицей G_l равенством

$$\tilde{G}_l = U G_l U^*. \quad (1.8)$$

Легко видеть, что уравнение (1.5) эквивалентно уравнению

$$\tilde{G}_{l+1} - (\rho \|A\|)^{-1} \tilde{A} \tilde{G}_l = \delta_0^l I. \quad (1.9)$$

Напомним, что матрица $\tilde{A}$ имеет вид

$$\tilde{A} = \begin{bmatrix} B & D \\ 0 & C \end{bmatrix}. \quad (1.10)$$

Решение уравнения (1.9) будем искать в виде

$$\tilde{G}_{l+1} = \begin{bmatrix} G_{l+1}^{(B)} & H_{l+1} \\ 0 & G_{l+1}^{(C)} \end{bmatrix}.$$

В силу (1.10) система уравнений (1.9) запишется как

$$\begin{bmatrix} G_{l+1}^{(B)} & H_{l+1} \\ 0 & G_{l+1}^{(C)} \end{bmatrix} = (\rho \|A\|)^{-1} \begin{bmatrix} B & D \\ 0 & C \end{bmatrix} \begin{bmatrix} G_l^{(B)} & H_l \\ 0 & G_l^{(C)} \end{bmatrix} + \delta_0^l \begin{bmatrix} I_{N-k} & 0 \\ 0 & I_k \end{bmatrix}.$$

Последняя система распадается на три системы:

$$\begin{aligned} G_{l+1}^{(B)} &= (\rho \|A\|)^{-1} B G_l^{(B)} + \delta_0^l I_{N-k}; \\ G_{l+1}^{(C)} &= (\rho \|A\|)^{-1} C G_l^{(C)} + \delta_0^l I_k; \\ H_{l+1} &= (\rho \|A\|)^{-1} [B H_l + D G_l^{(C)}]. \end{aligned} \quad (1.11)$$

Ограниченными решениями этих систем являются

$$\begin{aligned} G_l^{(B)} &= \begin{cases} -[(\rho \|A\|)^{-1} B]^{l-1} & \text{при } l \leq 0; \\ 0 & \text{при } l > 0; \end{cases} \\ G_l^{(C)} &= \begin{cases} [(\rho \|A\|)^{-1} C]^{l-1} & \text{при } l > 0; \\ 0 & \text{при } l \leq 0; \end{cases} \\ H_l &= \begin{cases} 0 & \text{при } l > 1; \\ D [(\rho \|A\|)^{-1} B]^{l-1} & \text{при } l \leq 1. \end{cases} \end{aligned}$$

Далее, в силу очевидных неравенств $\|G_l^{(B)}\| \leq \|G_l\|$, $\|G_l^{(C)}\| \leq \|G_l\|$ и неравенства (1.7), имеем

$$\begin{aligned} \|[(\rho \|A\|)^{-1} B]^l\| &\leq \delta_2^{-1} (1 + \delta_2)^l \quad \text{для } l < 0; \\ \|[(\rho \|A\|)^{-1} C]^l\| &\leq \delta_2^{-1} (1 - \delta_2)^l \quad \text{для } l > 0. \end{aligned}$$

Отсюда вытекают необходимые неравенства, доказывающие утверждение данной леммы:

$$\begin{aligned} \|B^{-l}\| &\leq (1/\delta_2) [\rho \|A\| (1 + \delta_2)]^{-l}; \\ \|C^l\| &\leq (1/\delta_2) [\rho \|A\| (1 - \delta_2)]^l, \end{aligned} \quad (1.12)$$

где $l = 1, 2, 3, \dots$

Следующая лемма посвящена исследованию матричного уравнения

$$XB - CX = D. \quad (1.13)$$

Как известно (см. § 18 главы 10 книги [20]), решение уравнения (1.13) единственно и может быть вычислено при помощи интеграла

$$X = \int_0^{\infty} e^{-\xi C} D e^{\xi B} d\xi,$$

когда спектры матриц B и C расположены в разных полуплоскостях комплексной плоскости.

Построим общее решение уравнения (1.13), когда спектры матриц B и C разделены кругом радиуса $\rho \|A\|$.

Лемма 2. Если выполнены оценки (1.2) и (1.12), то матричное уравнение относительно

$$LB - CL = D \quad (1.14)$$

имеет единственное решение, которое определяется формулой

$$L = \sum_{l=0}^{\infty} C^l D B^{-(l+1)}. \quad (1.15)$$

Ряд (1.15) сходится при любом D и $\|L\| \leq 1/2\delta^3$. (Матрицы L и D прямоугольные.)

Доказательство. Приступая к доказательству, заметим, что в силу (1.2) матрицы B и C не имеют общих собственных значений, что является необходимым и достаточным условием разрешимости уравнения (1.15) (см. [19], с. 207). Обозначим через L_n частную сумму ряда (1.15)

$$L_n = \sum_{l=0}^n C^l D B^{-(l+1)}$$

и оценим норму остаточного члена

$$\|L - L_n\| = \left\| \sum_{l=n+1}^{\infty} C^l D B^{-(l+1)} \right\| \leq \sum_{l=n+1}^{\infty} \|C^l\| \cdot \|D\| \cdot \|B^{-(l+1)}\|.$$

Огрубим последнее неравенство с помощью (1.12):

$$\begin{aligned} \|L - L_n\| &\leq \|D\| \sum_{l=n+1}^{\infty} (1/\delta_2)^2 [\rho \|A\| (1 - \delta_2)]^l [\rho \|A\| (1 + \delta_2)]^{-(l+1)} = \\ &= \frac{\|D\|}{\rho \|A\|} \cdot \frac{(1 - \delta_2)^{n+1}}{\delta_2^2 (1 + \delta_2)^{n+2}} \sum_{l=0}^{\infty} \frac{(1 - \delta_2)^l}{(1 + \delta_2)^l} = \\ &= \frac{1}{\rho \delta_2^2} \cdot \frac{1}{1 + \delta_2} \cdot \left(\frac{1 - \delta_2}{1 + \delta_2} \right)^{n+1} \cdot \frac{1}{1 - (1 - \delta_2)/(1 + \delta_2)} = \\ &= \frac{(1 - \delta_2)^{n+1}}{\rho \delta_2^2 (1 + \delta_2)^{n+1}} < \frac{8(1 - \delta_2)^{n+1}}{\delta^3 (1 + \delta_2)^{n+1}}. \end{aligned}$$

Из этих оценок, в свою очередь, с учетом $\delta_2 > 0$ без труда выводится, что $\lim_{n \rightarrow \infty} \|L - L_n\| = 0$. Непосредственно из (1.15) получим, что

$$\|L\| = \left\| \sum_{l=0}^{\infty} C^l D B^{-(l+1)} \right\| \leq \sum_{l=0}^{\infty} \|C^l\| \cdot \|D\| \cdot \|B^{-(l+1)}\|.$$

Откуда, имея в виду (1.12), запишем

$$\|L\| \leq \frac{\|D\|}{\rho \|A\| \delta_2^2 (1 + \delta_2)} \sum_{l=0}^{\infty} \left(\frac{1 - \delta_2}{1 + \delta_2} \right)^l \leq \frac{8\|D\|}{2\rho \|A\| \delta_2^2 (1 + \delta_2) \delta} \leq \frac{1}{2\delta^3}.$$

Непосредственной подстановкой (1.15) в (1.14) убеждаемся, что (1.15) удовлетворяет матричному уравнению $LB - CL = D$. Лемма доказана.

§ 2. ТЕОРЕМА О ПРЕДСТАВЛЕНИИ

Теорема. Если для матрицы A выполнена гипотеза I, то можно построить невырожденные матрицы P, R, U, T и прямоугольную матрицу S такие, что

$$TAT^{-1} = \begin{bmatrix} \hat{B} & 0 \\ 0 & \hat{C} \end{bmatrix}, \quad \text{где } T = \begin{bmatrix} P & S \\ 0 & R \end{bmatrix} U.$$

Матрица U унитарна, а T удовлетворяет оценкам

$$\|T\| \leq 6\delta^{-3}, \quad \|T^{-1}\| \leq 13\delta^{-4.5}.$$

В этом представлении матрица $\hat{B}$ имеет размерность $(N-k) \times (N-k)$ и для ее наименьшего сингулярного числа выполнена оценка

$$\sigma_1(\hat{B}) \geq \rho \|A\| \sqrt{1 + 2\delta_2^3} \quad (\delta_2 = \delta/2\rho).$$

Матрица C имеет размерность $k \times k$, и для ее наибольшего сингулярного числа выполнена оценка $\sigma_N(\hat{C}) \leq \rho \|A\| \sqrt{1 - \delta_2^3}$.

Доказательство. Как отмечалось ранее (см. § 1), любая квадратная матрица, для которой выполнена гипотеза I, может быть с помощью некоторой унитарной U ($U^*U = I$) записана в блочно-треугольном виде (ср. (1.1)):

$$A = U \begin{bmatrix} B & D \\ 0 & C \end{bmatrix} U^* = U \tilde{A} U^*. \quad (2.1)$$

Будем считать, что матрица A уже представлена в виде (2.1), и построим матрицы P, R и S .

Рассмотрим матричное уравнение относительно K :

$$B_1^* K B_1 - K = I \quad (2.2)$$

или $K B_1 - (B_1^{-1})^* K = (B_1^{-1})^*$, где $B_1 = (1/\rho \|A\|) B$. Матрица B невырожденная, что следует из (1.12). В силу леммы 2 (§ 1) решение уравнения (2.2) задается формулой

$$K = \sum_{i=1}^{\infty} (\rho \|A\|)^{2i} (B^* B)^{-i}, \quad (2.3)$$

в чем легко убедиться, заменяя в (1.14) L на K, B на B_1 и C на $(B_1^{-1})^*$. По той же лемме ряд (2.3) сходится. Из (2.3) видно, что K положительно определенная и симметричная, поэтому ее можно представить в виде $K = (P^{-1})^* P^{-1}$. (Например, можно положить $P = K^{-1/2}$.) Сингулярные числа матрицы обладают следующим свойством (см. [13], с. 12):

$$\sigma_1(A^{-1}) = \sigma_N^{-1}(A), \quad \sigma_2(A^{-1}) = \sigma_{N-1}^{-1}(A), \quad \dots, \quad \sigma_N(A^{-1}) = \sigma_1^{-1}(A).$$

Далее, по определению спектральной нормы

$$\|A\| = \sqrt{\max_{\|x\| \neq 0} \frac{(A^* A x, x)}{\|x\|^2}} = \sigma_N(A).$$

Пользуясь этим свойством сингулярных чисел, получим цепочку равенств

$$\begin{aligned} \frac{1}{\sigma_1^2(\hat{B})} &= \sigma_N^2(\hat{B}^{-1}) = \|\hat{B}^{-1}\|^2 = \max_{\|x\| \neq 0} \frac{(x, x)}{(\hat{B}x, \hat{B}x)} = \\ &= \max_{\|x\| \neq 0} \frac{(P^{-1} P x, P^{-1} P x)}{(P^{-1} B P x, P^{-1} B P x)} = \max_{\|x\| \neq 0} \frac{(Kz, z)}{(KBz, Bz)}, \end{aligned}$$

где $z = Px, K = (P^{-1})^* P^{-1}$.

Заменяя K его выражением (2.3), находим

$$\begin{aligned} \frac{1}{\sigma_1^2(\hat{B})} &= \max_{\|z\| \neq 0} \frac{\left(\sum_{l=1}^{\infty} (\rho \|A\|)^{2l} (B^*B)^{-l} z, z \right)}{\left(\sum_{l=1}^{\infty} (\rho \|A\|)^{2l} (B^*B)^{-l+1} z, z \right)} = \\ &= \max_{\|z\| \neq 0} \frac{\left(\sum_{l=1}^{\infty} (\rho \|A\|)^{2l} (B^*B)^{-l} z, z \right)}{\left(\left[(\rho \|A\|)^2 I + (\rho \|A\|)^2 \sum_{l=1}^{\infty} (\rho \|A\|)^{2l} (B^*B)^{-l} \right] z, z \right)} = \\ &= \frac{1}{(\rho \|A\|)^2} \max_{\|z\| \neq 0} \frac{(Kz, z)}{(z, z) + (Kz, z)}. \end{aligned}$$

Отсюда следует, что

$$\sigma_1^2(\hat{B}) \geq (\rho \|A\|)^2 (1 + \sigma_N^{-1}(K)). \quad (2.4)$$

Здесь мы использовали неравенство $(Kz, z) \leq \sigma_N(K) (z, z)$. Из (2.3), используя результат леммы 1, оценим $\sigma_N(K)$:

$$\begin{aligned} \sigma_N(K) = \|K\| &= \left\| \sum_{l=1}^{\infty} (\rho \|A\|)^{2l} (B^*B)^{-l} \right\| \leq \\ &\leq \sum_{l=1}^{\infty} (\rho \|A\|)^{2l} \|B^{-l}\| \|(B^{-l})^*\| = \sum_{l=1}^{\infty} (\rho \|A\|)^{2l} \|B^{-l}\|^2 \leq \\ &\leq \sum_{l=1}^{\infty} (\rho \|A\|)^{2l} \cdot \frac{1}{(\rho \|A\|)^{2l}} \cdot \frac{1}{\delta_2^2 (1 + \delta_2)^{2l}} = \frac{1}{\delta_2^2} \sum_{l=1}^{\infty} \frac{1}{(1 + \delta_2)^{2l}} = \\ &= 1/(2 + \delta_2 \delta_2^3) \leq 1/2\delta_2^3. \end{aligned}$$

Таким образом, $\sigma_N(K) \leq 1/2\delta_2^3$.

Учитывая эту оценку, огрубим неравенство (2.4): $\sigma_1^2(\hat{B}) \geq (\rho \|A\|)^2 (1 + 2\delta_2^3)$. Итак,

$$\sigma_1(\hat{B}) \geq \rho \|A\| \sqrt{1 + 2\delta_2^3}. \quad (2.5)$$

Отметим еще неравенство

$$\sigma_N(P^{-1}) = \sqrt{\sigma_N(K)} \leq \sqrt{1/2\delta_2^3}. \quad (2.6)$$

Из уравнения (2.2), в силу положительной определенности K , имеем $\sigma_1(B_1^*KB_1) = \sigma_1(K) + 1 > 1$. Кроме того, очевидно, что $\sigma_1(B_1^*KB_1) \leq \sigma_1(K) \|B_1\|^2$. Следовательно, $\sigma_1(K) \geq (\rho \|A\|)^2 / \|B\|^2 \geq 1$ (напомним, что $B_1 = (\rho \|A\|)^{-1}B$), поэтому

$$\|P\| = \|K^{-1}\|^{1/2} = 1/\sqrt{\sigma_1(K)} \leq 1. \quad (2.7)$$

Таким образом, при условии $|\lambda_j(B)| \geq \rho \|A\| (1 + \delta_2)$ построим матрицу P такую, что $\sigma_1(P^{-1}BP) \geq \rho \|A\| \sqrt{1 + 2\delta_2^3}$, и для нее получим оценки $\|P\| \leq 1$, $\|P^{-1}\| \leq (2\delta_2^3)^{-0.5}$. Отметим также, что

$$\|\hat{B}\| = \|P^{-1}BP\| \leq \mu(P) \|B\| \leq \|A\| / (2\delta_2^3)^{1/2}. \quad (2.8)$$

Решение аналогичного (2.2) уравнения

$$C_1^*MC_1 - M = -I, \quad (2.9)$$

где $C_1 = (\rho \|A\|)^{-1}C$, имеет вид

$$M = \sum_{l=0}^{\infty} (C_1^*C_1)^l = \sum_{l=0}^{\infty} (\rho \|A\|)^{-2l} (C^*C)^l. \quad (2.10)$$

Это утверждение обосновывается совершенно так же, как и формула (2.3) — ссылкой на лемму 2.

Из (2.10) видно, что M — положительно определенная и симметричная. Поэтому ее можно представить в виде $M = (R^{-1})^* R^{-1}$ (например, положить $R = M^{-1/2}$). Оценим $\sigma_N(\hat{C})$, где $\hat{C} = R^{-1}CR$:

$$\begin{aligned}\sigma_N^2(\hat{C}) &= \|\hat{C}\|^2 = \max_{\|x\| \neq 0} \frac{(\hat{C}x, \hat{C}x)}{(x, x)} = \\ &= \max_{\|x\| \neq 0} \frac{(R^{-1}CRx, R^{-1}CRx)}{(R^{-1}Rx, R^{-1}Rx)} = \max_{\|x\| \neq 0} \frac{((R^{-1})^* R^{-1}CRx, CRx)}{((R^*)^{-1} R^{-1}Rx, Rx)} = \\ &= \max_{\|z\| \neq 0} \frac{(MCz, Cz)}{(Mz, z)}.\end{aligned}$$

Таким образом,

$$\sigma_N^2(\hat{C}) = \max_{\|z\| \neq 0} \frac{(MCz, Cz)}{(Mz, z)}. \quad (2.11)$$

Подставляя в (2.11) вместо M его выражение из (2.10), имеем

$$\begin{aligned}\sigma_N^2(\hat{C}) &= \max_{\|z\| \neq 0} \frac{\left(\sum_{l=0}^{\infty} (\rho \|A\|)^{-2l} (C^*C)^l Cz, Cz \right)}{\left(\sum_{l=0}^{\infty} (\rho \|A\|)^{-2l} (C^*C)^l z, z \right)} = \\ &= \max_{\|z\| \neq 0} \frac{\left(\left[(\rho \|A\|)^2 \sum_{l=0}^{\infty} (\rho \|A\|)^{-2l} (C^*C)^l - I \right] z, z \right)}{\left(\sum_{l=0}^{\infty} (\rho \|A\|)^{-2l} (C^*C)^l z, z \right)} = \\ &= (\rho \|A\|)^2 \max_{\|z\| \neq 0} \frac{((M - I)z, z)}{(Mz, z)} = \\ &= (\rho \|A\|)^2 \max_{\|z\| \neq 0} \left[1 - \frac{(z, z)}{(Mz, z)} \right] \leq \\ &\leq (\rho \|A\|)^2 \left(1 - \max_{\|z\| \neq 0} \frac{(z, z)}{\sigma_N(M)(z, z)} \right).\end{aligned}$$

Окончательно

$$\sigma_N^2(\hat{C}) \leq (\rho \|A\|)^2 (1 - (\sigma_N(M))^{-1}). \quad (2.12)$$

Здесь использовано неравенство $(Mz, z) \leq \sigma_N(M)(z, z)$. Из (2.10), учитывая результаты леммы 1 (§ 1), оценим $\sigma_N(M)$:

$$\begin{aligned}\sigma_N(M) &= \left\| \sum_{l=0}^{\infty} (\rho \|A\|)^{-2l} (C^*C)^l \right\| \leq \\ &\leq \sum_{l=0}^{\infty} (\rho \|A\|)^{-2l} \|C^*\| \cdot \|C^l\| = \sum_{l=0}^{\infty} (\rho \|A\|)^{-2l} \|C^l\|^2,\end{aligned}$$

откуда, имея в виду (1.12), получим

$$\begin{aligned}\sigma_N(M) &\leq \sum_{l=0}^{\infty} (\rho \|A\|)^{-2l} (\rho \|A\|)^{2l} \frac{(1 - \delta_2)^{2l}}{\delta_2^2} = \\ &= \frac{1}{\delta_2^2} \sum_{l=0}^{\infty} (1 - \delta_2)^{2l} \leq \frac{1}{\delta_2^2 [1 - (1 - \delta_2^2)]} \leq \frac{1}{\delta_2^3},\end{aligned}$$

т. е. $\sigma_N(M) \leq \delta_2^{-3}$. С помощью последнего неравенства можно огрубить (2.12). При этом приходим к неравенству $\sigma_N^2(\hat{C}) \leq (\rho \|A\|)^2 (1 - \delta_2^3)$, или

$$\sigma_N(\hat{C}) \leq \rho \|A\| \sqrt{1 - \delta_2^3}. \quad (2.13)$$

Из уравнения (2.9), в силу положительной определенности M , имеем неравенство $\sigma_1(M) = \sigma_1(C_1^* M C_1) + 1 > 1$, из которого следует, что $\|R\| = \|M^{-1}\|^{0.5} < 1$. Таким образом, при условии $|\lambda_j(C)| \leq \rho \|A\| (1 - \delta_2)$ построена матрица R , обладающая свойством $\sigma_N(R^{-1}CR) \leq \rho \|A\| \sqrt{1 - \delta_2^2}$ и для нее получены оценки $\|R^{-1}\| \leq \delta_2^{-1.5}$, $\|R\| \leq 1$. Далее нетрудно видеть, что

$$\begin{bmatrix} P & S \\ 0 & R \end{bmatrix}^{-1} = \begin{bmatrix} P^{-1} & -P^{-1}SR^{-1} \\ 0 & R^{-1} \end{bmatrix}.$$

Матрицу S выберем так, чтобы выполнялось равенство

$$\begin{bmatrix} P^{-1}BP & 0 \\ 0 & R^{-1}CR \end{bmatrix} = \begin{bmatrix} P^{-1} & -P^{-1}SR^{-1} \\ 0 & R^{-1} \end{bmatrix} \begin{bmatrix} B & D \\ 0 & C \end{bmatrix} \begin{bmatrix} P & S \\ 0 & R \end{bmatrix}.$$

Перемножив матрицы в правой части последнего равенства, получим

$$\begin{bmatrix} P^{-1}BP & 0 \\ 0 & R^{-1}CR \end{bmatrix} = \begin{bmatrix} P^{-1}BP & P^{-1}BS + P^{-1}DR - P^{-1}SR^{-1}CR \\ 0 & R^{-1}CR \end{bmatrix}.$$

Для того чтобы удовлетворить последнему равенству, достаточно выбрать S так, чтобы $P^{-1}BS + P^{-1}DR - P^{-1}SR^{-1}CR = 0$, т. е. S должна удовлетворять матричному уравнению

$$BSR^{-1} - SR^{-1}C = -D.$$

В силу леммы 2, решение последнего уравнения задается формулой $S = LR$, где $L = \sum_{l=0}^{\infty} C^l DB^{-(l+1)}$. Отметим также, что $P^{-1}SR^{-1} = P^{-1}L$.

Завершим доказательство теоремы оценкой нормы преобразующей матрицы. В силу очевидных неравенств

$$\|T\| = \left\| \begin{bmatrix} P & S \\ 0 & R \end{bmatrix} \cdot U \right\| \leq \left\| \begin{bmatrix} P & S \\ 0 & R \end{bmatrix} \right\| \leq \|P\| + \|S\| + \|R\|;$$

$$\|T^{-1}\| = \left\| U^* \begin{bmatrix} P & S \\ 0 & R \end{bmatrix}^{-1} \right\| \leq \left\| \begin{bmatrix} P & S \\ 0 & R \end{bmatrix}^{-1} \right\| \leq \|P^{-1}\| + \|P^{-1}SR^{-1}\| + \|R^{-1}\|,$$

откуда с учетом полученных ранее оценок

$$\|T\| \leq 6\delta^{-3}, \quad \|T^{-1}\| \leq 13\delta^{-9/2}.$$

§ 3. «ОБЩНОСТЬ ПОЛОЖЕНИЯ» БАЗИСА ОТНОСИТЕЛЬНО ИНВАРИАНТНЫХ ПОДПРОСТРАНСТВ $\mathcal{L}_1, \mathcal{L}_2$

Напомним, что в унитарном пространстве, так же как и в евклидовом, угол φ между векторами x и y определяется из соотношения

$$\cos^2 \varphi = \frac{|(x, y)|^2}{(x, x)(y, y)}.$$

Угол между вектором y и некоторым подпространством S по определению есть наименьший угол между данным вектором и векторами из S :

$$\cos \varphi_1 = \max_{x \in S} \frac{|(x, y)|}{\|x\| \cdot \|y\|}.$$

Нетрудно видеть, что из всех векторов подпространства S наименьший угол с заданным вектором y образует ортогональная проекция вектора y на подпространство S .

Угол между подпространствами введем следующим образом. Пусть заданы два подпространства: S_l и S_r ($l \leq r$) (индекс внизу указывает

размерность подпространства). Найдём стационарные значения косинуса угла между вектором, лежащим в S_i , и вектором, принадлежащим к S_r . Эти значения можно искать, например, приравнявая нулю производные от вспомогательной функции Лагранжа

$$U(x, y) = \frac{|(x, y)|}{\|x\| \cdot \|y\|} + \frac{\lambda}{2} \|x\| + \frac{\mu}{2} \|y\|.$$

Те векторы, которые дают стационарное значение косинуса угла между ними, назовём стационарными векторами.

Оказывается (см. [21], с. 393, 405), в подпространствах S_i и S_r можно выделить соответственно l и r взаимно перпендикулярных стационарных векторов. Углы между подпространствами вводятся следующим образом. Выделим в S_i l взаимно перпендикулярных стационарных направлений. Пусть x — число не перпендикулярных к подпространству S_r стационарных векторов. Обозначим эти векторы через $x_1, x_2, \dots, x_x$, а их проекции на S_r — через $y_1, y_2, \dots, y_x$. Углы между x_1 и y_1, x_2 и $y_2, \dots, x_x$ и y_x обозначим через $\varphi_1, \varphi_2, \dots, \varphi_x$ и положим $\varphi_{x+1} = \varphi_{x+2} = \dots = \varphi_1 = \pi/2$. Углы $\varphi_1, \varphi_2, \dots, \varphi_l$ называются углами между подпространствами S_i и S_r . В дальнейшем вместо термина «углы между подпространствами» будем употреблять термин «угол между подпространствами», имея в виду при этом наибольшее из всех $\varphi_1, \varphi_2, \dots, \varphi_l$.

Оценим теперь угол между подпространствами $\mathcal{L}_1^1$ и $\mathcal{L}_2$.

Теорема 2. Если для матрицы A выполнена гипотеза II, то для любого вектора $l \in \tilde{\mathcal{L}}_1^1$, представленного в виде $l = (l_1, l_2, \dots, l_{N-k}, \dots, l_N)'$ выполняется оценка

$$\sqrt{|l_1|^2 + |l_2|^2 + \dots + |l_{N-k}|^2} \geq \frac{1}{\mu(T)} \left(\frac{\delta^{3/2} \delta_1}{2} \right)^k \sqrt{|l_{N-k+1}|^2 + \dots + |l_N|^2}, \quad (3.1)$$

где $\mu(T) = \|T\| \cdot \|T^{-1}\|$. (3.1) можно переписать в виде

$$\operatorname{tg} \varphi_0 \leq \mu(T) \cdot \left(\frac{2}{\delta^{3/2} \delta_1} \right)^k,$$

где $\varphi_0 \in (\tilde{\mathcal{L}}_1^1, \tilde{\mathcal{L}}_1)$.

Доказательство. Гипотеза II означает, что

$$\|(A - \lambda^{(1)}I)(A - \lambda^{(2)}I) \dots (A - \lambda^{(k)}I)f\| \geq (\delta_1 \|A\|)^k \|f\|. \quad (3.2)$$

В силу определения матрицы T перепишем неравенство (3.2):

$$\begin{aligned} & \|T^{-1}T(A - \lambda^{(1)}I)T^{-1}T(A - \lambda^{(2)}I)T^{-1} \dots T(A - \lambda^{(k)}I)T^{-1}Tf\| = \\ & = \left\| T^{-1} \begin{bmatrix} (\hat{B} - \lambda^{(1)}I) \dots (\hat{B} - \lambda^{(k)}I) & 0 \\ 0 & (\hat{C} - \lambda^{(1)}I) \dots (\hat{C} - \lambda^{(k)}I) \end{bmatrix} Tf \right\| \geq \\ & \geq (\delta_1 \|A\|)^k \|f\|. \end{aligned}$$

Обозначим $Tf = (g, h)'$. Перепишем последнее неравенство, округив его с помощью очевидных неравенств:

$$\begin{aligned} & \|Tz\| \geq \sigma_1(T) \|z\|, \quad \|Tz\| \leq \sigma_N(T) \|z\|; \\ & \left\| \begin{bmatrix} (\hat{B} - \lambda^{(1)}I) \dots (\hat{B} - \lambda^{(k)}I) g' \\ (\hat{C} - \lambda^{(1)}I) \dots (\hat{C} - \lambda^{(k)}I) h' \end{bmatrix} \right\| \geq \frac{1}{\mu(T)} (\delta_1 \|A\|)^k \|g'\|, \end{aligned} \quad (3.3)$$

где $g' \in \tilde{\mathcal{L}}_1$, $h' \in \tilde{\mathcal{L}}_2$ ($\tilde{\mathcal{L}}_1, \tilde{\mathcal{L}}_2$ — инвариантные подпространства матрицы TAT^{-1} , отвечающие собственным значениям, лежащим соответственно вне и внутри круга $|\lambda| = \rho \|A\|$). Пусть $y_1, y_2, \dots, y_k$ — линейно независимые столбцы присоединённых матриц к матрицам $T(A - \lambda_i I)T^{-1}$, соответствующие собственным значениям $\lambda_1, \lambda_2, \dots, \lambda_k$. Любой вектор $\tilde{h}$ из подпространства $\tilde{\mathcal{L}}_2$ представим в виде

$$\alpha_1 y_1 + \alpha_2 y_2 + \dots + \alpha_k y_k = \tilde{h}, \quad \tilde{h} \in \tilde{\mathcal{L}}_2.$$

Ясно, что матрица $(\hat{C} - \lambda_1 I)(\hat{C} - \lambda_2 I) \dots (\hat{C} - \lambda_k I)$ аннулирует любой вектор из подпространства $\tilde{\mathcal{L}}_2$, в том числе и вектор h' , т. е. $(\hat{C} - \lambda_1 I)(\hat{C} - \lambda_2 I) \dots (\hat{C} - \lambda_k I) h' = 0$. Следовательно,

$$\begin{aligned} \frac{1}{\mu(T)} (\delta_1 \|A\|)^k &\leq \frac{\|(\hat{B} - \lambda_1 I) \dots (\hat{B} - \lambda_k I) g'\|}{\sqrt{\|g'\|^2 + \|h'\|^2}} = \\ &= \frac{\|(\hat{B} - \lambda_1 I) \dots (\hat{B} - \lambda_k I) g'\|}{\sqrt{\|g'\|^2 + \|h'\|^2}} \leq \frac{\|(\hat{B} - \lambda_1 I)\| \cdot \|(\hat{B} - \lambda_2 I)\| \dots \|(\hat{B} - \lambda_k I)\| \cdot \|g'\|}{\|g'\| \cdot \sqrt{1 + (\|h'\|/\|g'\|)^2}}. \end{aligned}$$

Откуда

$$\frac{(\delta_1 \|A\|)^k}{\mu(T)} \leq \frac{\|\hat{B} - \lambda_1 I\| \dots \|\hat{B} - \lambda_k I\|}{\sqrt{1 + (\|h'\|/\|g'\|)^2}}. \quad (3.4)$$

В силу (2.8) и очевидного неравенства $|\lambda(A)| \leq \|A\|$ имеем

$$\|\hat{B} - \lambda_i I\| \leq 2\|A\|\delta^{-1.5}, \quad i = 1, 2, 3, \dots$$

Огрубив (3.4), получим

$$(\delta_1 \|A\|)^k / \mu(T) \leq (\|g'\|/\|h'\|) (\delta^{-1.5} \cdot 2\|A\|)^k.$$

Из последнего неравенства следует неравенство

$$\|g'\|/\|h'\| \geq (\delta_1 \delta^{1.5})^k / 2^k \mu(T), \quad (3.5)$$

которое доказывает утверждение теоремы.

§ 4. ОЦЕНКА УГЛОВ МЕЖДУ ПОДПРОСТРАНСТВАМИ $\mathcal{L}_1$ и $\mathcal{L}_j^{(N-k)}$

В § 2, 3 доказывалось, что если для произвольной комплексной матрицы A выполнены гипотезы I, II, то для матриц $\hat{B}$ и $\hat{C}$, определенных из равенства

$$\begin{bmatrix} \hat{B} & 0 \\ 0 & \hat{C} \end{bmatrix} = TAT^{-1} = \hat{A}, \quad (4.1)$$

имеют место оценки

$$\sigma_1(\hat{B}) \geq \rho \|A\| \sqrt{1 + 2\delta_2^3}, \quad \sigma_N(\hat{C}) \leq \rho \|A\| \sqrt{1 - \delta_2^3}. \quad (4.2)$$

Кроме того, для угла между подпространствами $\tilde{\mathcal{L}}_1^1$ и $\tilde{\mathcal{L}}_1$ получена оценка

$$\operatorname{tg} \varphi_0 \leq 2^k \mu(T) / (\delta^{1.5} \delta_1)^k. \quad (4.3)$$

По вектору $z_0 \in \mathcal{L}_1^1$ построим последовательность векторов z_j по рекуррентной формуле

$$z_j = (\hat{A} - \hat{E}_j) z_{j-1}, \quad j = 1, 2, \dots, \quad (4.4)$$

где $\|\hat{E}_j\| \leq \mu(T) \varepsilon$, $\mu(T) = \|T\| \cdot \|T^{-1}\|$, $\varepsilon > 0$. Пусть φ_j — угол между вектором z_j и подпространством $\tilde{\mathcal{L}}_1$.

Теорема 3. Если для матрицы A выполнены гипотезы I, II и, кроме того, $(\rho \|A\| [\sqrt{1 + 2\delta_2^3} - \sqrt{1 - \delta_2^3}] - (\varepsilon_1 + \varepsilon_4))^2 - 4\varepsilon_2\varepsilon_3 > 0$, то

$$\operatorname{tg} \varphi_j \leq \mu(T) \left[\left(\frac{2}{\delta^{3/2} \delta_1} \right)^k L^j + \delta^6 \eta(\varepsilon) \right],$$

$$L = \frac{\rho \|A\| \sqrt{1 - \delta_2^3} + \mu(T) \cdot \varepsilon (\delta^3 + \delta^5 \operatorname{ctg} \varphi_0)}{\rho \|A\| \sqrt{1 + 2\delta_2^3} - \mu(T) \cdot \varepsilon (\delta^{3/2} + \operatorname{tg} \varphi_0)},$$

$$\eta(\varepsilon) = \frac{\varepsilon}{2\delta_2^3 [\rho \|A\| - \varepsilon \mu(T) + \sqrt{[\rho \|A\| - \varepsilon \mu(T)]^2 - \mu^2(T) \varepsilon^2 \delta^2}]},$$

$$\varepsilon_1 = 2\mu(T) \delta^{3/2} \varepsilon, \quad \varepsilon_2 = \mu(T) \varepsilon, \quad \varepsilon_3 = \mu(T) \delta^6 \varepsilon, \quad \varepsilon_4 = \mu(T) \delta^3 \varepsilon.$$

При этом, если $\mu(T) \delta^6 \eta(\varepsilon) < \operatorname{tg} \varphi_0 < (\mu(T) \eta(\varepsilon))^{-1}$, то $L < 1$.

Доказательство. Обозначим $z_j = (g_j, h_j)'$, $g_j \in \tilde{\mathcal{L}}_1$, $h_j \in \tilde{\mathcal{L}}_2$. Учитывая (4.1), перепишем (4.4) следующим образом:

$$\begin{bmatrix} g_j \\ h_j \end{bmatrix} = \begin{bmatrix} \hat{B} + \hat{E}_{1j} \hat{E}_{2j} \\ \hat{E}_{3j} \hat{C} + \hat{E}_{4j} \end{bmatrix} \begin{bmatrix} g_{j-1} \\ h_{j-1} \end{bmatrix} = \begin{bmatrix} (\hat{B} + \hat{E}_{1j}) g_{j-1} + \hat{E}_{2j} h_{j-1} \\ \hat{E}_{3j} g_{j-1} + (\hat{C} + \hat{E}_{4j}) h_{j-1} \end{bmatrix}. \quad (4.5)$$

Оценим соотношение $\|h_j\|/\|g_j\|$, для которого, учитывая (4.5), получим

$$\frac{\|h_j\|}{\|g_j\|} = \frac{\|(\hat{C} + \hat{E}_{4j}) h_{j-1} + \hat{E}_{3j} g_{j-1}\|}{\|(\hat{B} + \hat{E}_{1j}) g_{j-1} + \hat{E}_{2j} h_{j-1}\|} \leq \frac{\|(\hat{C} + \hat{E}_{4j}) h_{j-1}\| + \|\hat{E}_{3j} g_{j-1}\|}{\|(\hat{B} + \hat{E}_{1j}) g_{j-1}\| - \|\hat{E}_{2j} h_{j-1}\|}. \quad (4.6)$$

Напомним, что

$$\begin{aligned} \hat{E}_j &= T E_j T^{-1} = \begin{bmatrix} P & S \\ O & R \end{bmatrix} U \begin{bmatrix} E_{1j} & E_{2j} \\ E_{3j} & E_{4j} \end{bmatrix} U^* \begin{bmatrix} P^{-1} & -P^{-1} S R^{-1} \\ O & R^{-1} \end{bmatrix} = \\ &= \begin{bmatrix} P & S \\ O & R \end{bmatrix} \begin{bmatrix} \tilde{E}_{1j} & \tilde{E}_{2j} \\ \tilde{E}_{3j} & \tilde{E}_{4j} \end{bmatrix} \begin{bmatrix} P^{-1} & -P^{-1} S R^{-1} \\ 0 & R^{-1} \end{bmatrix}. \end{aligned}$$

Перемножив матрицы в правой части, находим

$$\begin{aligned} \hat{E}_{1j} &= P \tilde{E}_{1j} P^{-1} + S \tilde{E}_{3j} P^{-1}, \quad \hat{E}_{2j} = -P \tilde{E}_{1j} P^{-1} S R^{-1} - S \tilde{E}_{3j} P^{-1} S R^{-1} + \\ &+ P \tilde{E}_{2j} R^{-1} + S \tilde{E}_{4j} R^{-1}, \quad \hat{E}_{3j} = R \tilde{E}_{3j} P^{-1}, \quad \hat{E}_{4j} = -R \tilde{E}_{3j} P^{-1} S R^{-1} + R \tilde{E}_{4j} R^{-1}. \end{aligned}$$

Отсюда на основании оценок, полученных в § 2, и в силу того, что $\|E_j\| = \|U E_j U^*\| \leq \|E_j\| < \varepsilon$, имеем

$$\begin{aligned} \|\hat{E}_{1j}\| &\leq 2\mu(T) \delta^{3/2} \varepsilon = \varepsilon_1, \quad \|\hat{E}_{2j}\| \leq \mu(T) \varepsilon = \varepsilon_2, \\ \|\hat{E}_{3j}\| &\leq \mu(T) \delta^6 \varepsilon = \varepsilon_3, \quad \|\hat{E}_{4j}\| \leq \mu(T) \delta^3 \varepsilon = \varepsilon_4. \end{aligned}$$

Из (4.6), принимая во внимание (4.2), с учетом только что полученных оценок нетрудно заключить, что

$$\frac{\|h_j\|}{\|g_j\|} \leq \frac{(\rho \|A\| \sqrt{1 - \delta_2^3} + \varepsilon_4) \|h_{j-1}\| / \|g_{j-1}\| + \varepsilon_3}{\rho \|A\| \sqrt{1 + 2\delta_2^3} - \varepsilon_1 - \varepsilon_2 \|h_{j-1}\| / \|g_{j-1}\|}.$$

Если ввести сокращенные обозначения

$$\eta_j = \|h_j\|/\|g_j\|, \quad M = \rho \|A\| \sqrt{1 + 2\delta_2^3} - \varepsilon_1, \quad m = \rho \|A\| \sqrt{1 - \delta_2^3} + \varepsilon_4, \quad (4.7)$$

то доказанное нами неравенство примет вид

$$\eta_j \leq \frac{m \eta_{j-1} + \varepsilon_3}{M - \varepsilon_2 \eta_{j-1}}. \quad (4.8)$$

Для того чтобы последовательность $\eta_0, \eta_1, \eta_2, \dots, \eta_j, \dots$ была убывающей, надо потребовать выполнения неравенств

$$\frac{m \eta_{j-1} + \varepsilon_3}{M - \varepsilon_2 \eta_{j-1}} \leq \eta_{j-1}. \quad (4.9)$$

Перепишем (4.9) следующим образом:

$$\begin{cases} \varepsilon_2 \eta_{j-1}^2 - (M - m) \eta_{j-1} + \varepsilon_3 \leq 0, \\ M - \varepsilon_2 \eta_{j-1} > 0. \end{cases} \quad (4.10)$$

Решая квадратное уравнение $\varepsilon_2 \eta_{j-1}^2 - (M - m) \eta_{j-1} + \varepsilon_3 = 0$, находим

$$\begin{aligned} \eta_{\min} &= \frac{M - m - \sqrt{(M - m)^2 - 4\varepsilon_2 \varepsilon_3}}{2\varepsilon_2} = \\ &= \frac{2\varepsilon_3}{M - m + \sqrt{(M - m)^2 - 4\varepsilon_2 \varepsilon_3}}, \\ \eta_{\max} &= \frac{M - m + \sqrt{(M - m)^2 - 4\varepsilon_2 \varepsilon_3}}{2\varepsilon_2} = \frac{\delta^6}{\eta_{\min}}. \end{aligned}$$

Тогда (4.10) можно переписать в виде

$$\begin{cases} (\eta_{j-1} - \eta_{\min})(\eta_{j-1} - \eta_{\max}) \leq 0, \\ \eta_{j-1} \leq \frac{M}{\varepsilon_2}. \end{cases} \quad (4.11)$$

Первое из неравенств (4.11) выполнено либо при а) $\eta_{j-1} \geq \eta_{\min}$, $\eta_{j-1} \leq \eta_{\max}$, либо при б) $\eta_{j-1} \leq \eta_{\min}$, $\eta_{j-1} \geq \eta_{\max}$.

Так как $\eta_{\max} > \eta_{\min}$, то условие б) не выполняется. Легко видеть, что $\eta_{\max} \leq M/\varepsilon_2$. Следовательно, решениями неравенства (4.9) являются точки отрезка $[\eta_{\min}, \eta_{\max}]$. Для всех $\eta_0 \in (\eta_{\min}, \eta_{\max})$ последовательность $\eta_0, \eta_1, \dots, \eta_j, \dots$ будет убывающей. Для оценки скорости сходимости последовательности $\eta_0, \eta_1, \dots$ к $\eta_{\min}$ обозначим

$$\xi_j = \eta_j - \eta_{\min}, \quad f(\xi_j) = \frac{\varepsilon_3 + m\eta_{\min} + m\xi_j}{M - \varepsilon_2\eta_{\min} - \varepsilon_2\xi_j} - \eta_{\min}. \quad (4.12)$$

Тогда неравенства (4.8) и (4.9) переписываются так:

$$\xi_j \leq f(\xi_{j-1}) \leq \xi_{j-1}, \quad (4.13)$$

причем $f(0) = 0$.

Функция $f(\xi)$ на интервале $[0, \eta_{\max} - \eta_{\min})$ монотонно убывающая, выпуклая, что следует из неравенств

$$\begin{aligned} f'(\xi) &= \frac{mM + \varepsilon_2 \varepsilon_3}{(M - \varepsilon_2 \eta_{\min} - \varepsilon_2 \xi)^2} \geq 0; \\ f''(\xi) &= \frac{2\varepsilon_2 (mM + \varepsilon_2 \varepsilon_3)}{(M - \varepsilon_2 \eta_{\min} - \varepsilon_2 \xi)^3} \geq 0. \end{aligned}$$

Для любой выпуклой функции f имеет место неравенство

$$f(q_1 x_1 + q_2 x_2) \leq q_1 f(x_1) + q_2 f(x_2) \quad (4.14)$$

для всех $q_1, q_2 > 0$, $q_1 + q_2 = 1$. Выбирая $q_1 = (\xi_{j-1} - \xi_j)/(\xi_{j-1} - \xi^*)$, $q_2 = 1 - \xi_j$, $x_1 = \xi^* = 0$, $x_2 = \xi_{j-1}$ и подставляя их в (4.14), получим $f(\xi_j) \leq (\xi_j/\xi_{j-1})f(\xi_{j-1})$, или, что то же самое,

$$f(\xi_j)/\xi_j \leq f(\xi_{j-1})/\xi_{j-1}. \quad (4.15)$$

Применяя несколько раз неравенство (4.13), имеем

$$\begin{aligned} \xi_j &\leq f(\xi_{j-1}) = \frac{f(\xi_{j-1})}{\xi_{j-1}} \xi_{j-1} \leq \frac{f(\xi_{j-1})}{\xi_{j-1}} f(\xi_{j-2}) = \\ &= \frac{f(\xi_{j-1})}{\xi_{j-1}} \cdot \frac{f(\xi_{j-2})}{\xi_{j-2}} \xi_{j-2} \leq \dots \leq \frac{f(\xi_{j-1})}{\xi_{j-1}} \dots \frac{f(\xi_0)}{\xi_0} \xi_0. \end{aligned}$$

Отсюда, в силу (4.15), $\xi_j \leq (f(\xi_0)/\xi_0)^j \xi_0$, или

$$\eta_j \leq [f(\xi_0)/\xi_0]^j \eta_0 + \eta_{\min}. \quad (4.16)$$

Из (4.14) видно, что если $\eta_0 \in (\eta_{\min}, \eta_{\max})$, то $f(\xi_0)/\xi_0 < 1$ или $[(\varepsilon_3 + \eta_0 m)/(M - \varepsilon_2 \eta_0) - \eta_{\min}]/(\eta_0 - \eta_{\min}) < 1$, с учетом последнего $L = (m + \varepsilon_3/\eta_0)/(M - \varepsilon_2 \eta_0) < 1$. Заменяя в (4.16) η_j , $f(\xi_0)$, ξ_0 , $\eta_{\min}$ их значениями из (4.7) и (4.12), получим

$$\frac{\|h_j\|}{\|g_j\|} \leq \left[\frac{\rho \|A\| \sqrt{1 + \delta_2^2} + \mu(T) \varepsilon (\delta^3 + \delta^6 \|g_0\|/h_0)}{\rho \|A\| \sqrt{1 + 2\delta_2^2} - \mu(T) \varepsilon (\delta^{3/2} + \|h_0\|/g_0)} \right]^j \times \\ \times \frac{\|h_0\|}{\|g_0\|} + \frac{\mu(T) \delta^3 \varepsilon}{\rho \|A\| - \mu(T) \varepsilon + \sqrt{[\rho \|A\| - \mu(T) \varepsilon]^2 - [\mu(T) \delta \varepsilon]^2}},$$

или короче:

$$\|h_j\|/\|g_j\| \leq L^j \|h_0\|/\|g_0\| + \mu(T) \delta^3 \eta(\varepsilon). \quad (4.17)$$

Неравенство (4.17) завершит доказательство теоремы 3, если заметить, что $\|h_j\|/\|z_j\| = \sin \varphi_j$, $\|g_j\|/\|z_j\| = \cos \varphi_j$ и огрубить его с помощью (4.3). Тогда запишем

$$\operatorname{tg} \varphi_j \leq \mu(T) \left[\left(\frac{2}{\delta^{3/2} \delta_1} \right)^k L^j + \delta^6 \eta(\varepsilon) \right]. \quad (4.18)$$

§ 5. ИССЛЕДОВАНИЕ СХОДИМОСТИ ОРТОГОНАЛЬНО-СТЕПЕННОГО МЕТОДА ВОЕВОДИНА С УЧЕТОМ ВЫЧИСЛИТЕЛЬНЫХ ПОГРЕШНОСТЕЙ

Построим последовательность матриц Q_i ($i = 1, 2, \dots$) по правилу:

$$\begin{aligned} (A + E_1)Q_0 &= Q_1 R_1; \\ (A + E_2)Q_1 &= Q_2 R_2; \\ (A + E_3)Q_2 &= Q_3 R_3; \\ &\dots \dots \dots \\ (A + E_j)Q_{j-1} &= Q_j R_j, \end{aligned} \quad (5.1)$$

где R_j — правые треугольные матрицы, Q_j — почти унитарные матрицы, т. е. $Q_j^* Q_j = I + H_j$, $\|H_j\| \leq \varepsilon_j$, $\varepsilon_j > 0$, $\|E_j\| \leq \varepsilon$, $\varepsilon > 0$. Из (5.1) легко получается представление

$$(A + E_j)(A + E_{j-1}) \dots (A + E_1)Q_0 = Q_j R_j R_{j-1} \dots R_1. \quad (5.2)$$

Обозначив $M_j = (A + E_j) \dots (A + E_1)$, $S_j = R_j R_{j-1} \dots R_1$, перепишем (5.2) в виде

$$M_j Q_0 = Q_j S_j. \quad (5.3)$$

Из представления (5.3) видно, что образ любого из первых $N - k$ столбцов матрицы Q_0 является линейной комбинацией первых $N - k$ столбцов матрицы Q_j .

В самом деле, если обозначить через y_i образ вектора $q_i^{(0)}$, т. е. $y_i = M_j q_i^{(0)}$, то исходя из того, что S — верхняя треугольная, имеем

$$y_i = S_{11}^{(j)} q_1^{(j)} + S_{21}^{(j)} q_2^{(j)} + \dots + S_{ii}^{(j)} q_i^{(j)}.$$

Следовательно, образы векторов $q_i^{(0)}$ при преобразовании M_j лежат на $N - k$ -плоскости, натянутой на первые $N - k$ столбцов матрицы Q_j . И наоборот, из представления $Q_j = M_j Q_0 S_j^{-1}$, в силу того, что матрица S_j^{-1} — верхняя треугольная, получим, что любой вектор, лежащий на $N - k$ -плоскости, натянутой на первые $N - k$ столбцов матрицы Q_j , представим в виде линейной комбинации образов первых $N - k$ столбцов матрицы Q_0 при преобразовании M_j . Исследуем линейную оболочку образов первых $N - k$ столбцов матрицы Q_0 с увеличением j .

Здесь $\mathcal{L}_1^{(0)}$ совпадает с линейной оболочкой векторов $q_1^{(0)}, q_2^{(0)}, \dots, q_{N-k}^{(0)}$ и $\mathcal{L}_j^{(N-k)}$ совпадает с линейной оболочкой $q_1^{(j)}, q_2^{(j)}, \dots, q_{N-k}^{(j)}$. Пусть

$$q_1^{(0)} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad q_2^{(0)} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad q_N^{(0)} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

— исходные базисные векторы пространства $\mathcal{L}$. В качестве нового базиса выберем столбцы матрицы Q_j : $q_1^{(j)}, q_2^{(j)}, \dots, q_N^{(j)}$. Как отмечалось ранее, векторы $q_1^{(j)}, q_2^{(j)}, \dots, q_{N-k}^{(j)}$ принадлежат подпространству $\mathcal{L}_j^{(N-k)}$. Ясно, что матрица Q_j будет матрицей перехода к новому базису и оператору A в этом базисе соответствует матрица

$$\tilde{A}_j = Q_j^{-1} A Q_j. \quad (5.4)$$

Пусть векторы $q_1^{(j)}, q_2^{(j)}, \dots, q_N^{(j)}$ нормированы. Предполагая для определенности, что $k > N/2$ (случай $k \leq N/2$ рассматривается аналогично), построим линейно независимую систему векторов $u_1, u_2, \dots, u_N$ по следующему правилу:

$$\begin{aligned} q_1^{(j)} &= \cos \varphi_1^{(j)} u_1 + \sin \varphi_1^{(j)} u_{N-k+1}; \\ q_2^{(j)} &= \cos \varphi_2^{(j)} u_2 + \sin \varphi_2^{(j)} u_{N-k+2}; \\ &\dots \dots \dots \\ q_{N-k}^{(j)} &= \cos \varphi_{N-k}^{(j)} u_{N-k} + \sin \varphi_{N-k}^{(j)} u_{2(N-k)}; \\ q_{N-k+1}^{(j)} &= \cos \varphi_1^{(j)} u_{N-k+1} - \sin \varphi_1^{(j)} u_1; \\ &\dots \dots \dots \\ q_{2(N-k)}^{(j)} &= \cos \varphi_{N-k}^{(j)} u_{2(N-k)} - \sin \varphi_{N-k}^{(j)} u_{N-k}; \\ q_{2(N-k)+1}^{(j)} &= u_{2(N-k)+1}; \\ &\dots \dots \dots \\ q_N^{(j)} &= u_N. \end{aligned}$$

Здесь $\varphi_i^{(j)}$ — угол между вектором $q_i^{(j)}$ и его ортогональной проекцией на подпространство $\mathcal{L}_1$. Обозначив через U матрицу, столбцами которой являются векторы $u_1, u_2, u_3, \dots, u_N$, получим

$$Q_j = U \cdot \Omega = U \times$$

$$\times \begin{bmatrix} \cos \varphi_1^{(j)} & -\sin \varphi_1^{(j)} & & & 0 \\ & \cos \varphi_{N-k}^{(j)} & & -\sin \varphi_{N-k}^{(j)} & \\ \sin \varphi_1^{(j)} & & \cos \varphi_1^{(j)} & & 0 \\ & \sin \varphi_{N-k}^{(j)} & & \cos \varphi_{N-k}^{(j)} & \\ 0 & & 0 & & 1 & \dots & 1 \end{bmatrix}$$

Подставляя в (5.4) вместо Q_j полученное выражение, имеем

$$\tilde{A}_j = Q_j^{-1} A Q_j = \Omega^{-1} U^{-1} A U \Omega = \Omega^* \bar{A} \Omega, \quad (5.5)$$

где $\bar{A} = U^{-1} A U$ — матрица оператора A в базисе $u_1, u_2, \dots, u_N$. Вследствие того, что линейная оболочка векторов $u_1, u_2, \dots, u_{N-k}$ — инвариантное подпространство, матрица $\bar{A}$ будет иметь вид

$$\bar{A} = U^{-1} A U = \begin{bmatrix} B & D \\ 0 & C \end{bmatrix}. \quad (5.6)$$

Клетки B, D, C — матрицы размеров $(N-k) \times (N-k)$, $(N-k) \times k$, $k \times k$ соответственно.

Подставляя (5.6) в (5.5) и обозначив

$$S = \begin{bmatrix} \sin \varphi_1^{(j)} & & & & \\ & \sin \varphi_2^{(j)} & & & \\ & & \ddots & & \\ 0 & & & & \\ & & & & \sin \varphi_{N-k}^{(j)} \end{bmatrix}, \quad C_1 = \begin{bmatrix} \cos \varphi_1^{(j)} & & & & \\ & \cos \varphi_2^{(j)} & & & \\ & & \ddots & & \\ 0 & & & & \\ & & & & \cos \varphi_{N-k}^{(j)} \end{bmatrix}, \quad C_2 = \begin{bmatrix} C_1 & 0 \\ 0 & I_{2k-N} \end{bmatrix} \quad (5.7)$$

(S — прямоугольная матрица), получим

$$\begin{aligned} \tilde{A}_j &= \begin{bmatrix} \tilde{A}_{1j} & \tilde{A}_{2j} \\ \tilde{A}_{3j} & \tilde{A}_{4j} \end{bmatrix} = \begin{bmatrix} C_1 & -S \\ S^* & C_2 \end{bmatrix}^{-1} \begin{bmatrix} B & D \\ 0 & C \end{bmatrix} \begin{bmatrix} C_1 & -S \\ S^* & C_2 \end{bmatrix} = \\ &= \begin{bmatrix} C_1 & S \\ -S^* & C_2 \end{bmatrix} \begin{bmatrix} B & D \\ 0 & C \end{bmatrix} \begin{bmatrix} C_1 & -S \\ S^* & C_2 \end{bmatrix}. \end{aligned}$$

Отсюда, перемножая матрицы в правой части, находим $\tilde{A}_{3j} = -S B C_1 - S D S + C_2 C S$. Тогда

$$\|\tilde{A}_{3j}\| \leq \|S\| (\|B\| \cdot \|C_1\| + \|S\| \cdot \|D\| + \|C_2\| \cdot \|C\|). \quad (5.8)$$

Из (5.7) оценим $\|S\|, \|C_1\|, \|C_2\|$:

$$\begin{aligned} \|S\| &= \sigma_N(S) = \max_{1 \leq i \leq N-k} \sin \varphi_i^{(j)}, \\ \|C_1\| &= \max_{1 \leq i \leq N-k} \cos \varphi_i^{(j)} \leq 1, \quad \|C_2\| \leq 1. \end{aligned}$$

В силу (5.6) и с учетом того, что $U^* U = I + H$, $\|H\| < \varepsilon_5$, имеем $\|B\| + \|C\| + \|D\| \leq 2(1 + \varepsilon_5) \|A\|$. Используя полученные неравенства, из (5.8) выводим оценку

$$\|\tilde{A}_{3j}\| \leq 2(1 + \varepsilon_5) \|A\| \max_{1 \leq i \leq N-k} \sin \varphi_i^{(j)}. \quad (5.9)$$

Углы $\varphi_i^{(j)}$ оценены в § 4. Подставляя оценку (4.18) в (5.9), получим окончательную оценку для $\|\tilde{A}_{3j}\|$:

$$\|\tilde{A}_{3j}\| \leq 2(1 + \varepsilon_5) \|A\| \mu(T) \left[\left(\frac{2}{\delta^{3/2} \delta_1} \right)^k L^j + \delta^6 \cdot \eta(\varepsilon) \right].$$

где

$$L = \frac{\rho \|A\| \sqrt{1 - \delta_2^3} + \mu(T) \cdot \varepsilon (\delta^3 + \delta^6 \operatorname{ctg} \varphi_0)}{\rho \|A\| \sqrt{1 + 2\delta_2^3} - \mu(T) \cdot \varepsilon (\delta^{3/2} + \operatorname{tg} \varphi_0)},$$

$$\eta(\varepsilon) = \frac{\varepsilon}{2\delta_2^3 [\rho \|A\| - \varepsilon \mu(T) + \sqrt{[\rho \|A\| - \varepsilon \mu(T)]^2 - \varepsilon^2 \delta_2^3 \mu^2(T)}}.$$

Подведя итог нашим рассуждениям, установим теорему.

Теорема. Если для матрицы A выполнены гипотезы I и II и, кроме того, $[\rho \|A\| (\sqrt{1 + 2\delta_2^3} - \sqrt{1 - \delta_2^3}) - \varepsilon_1 - \varepsilon_4]^2 - 4\varepsilon_2\varepsilon_3 > 0$, то последовательность матриц A_j , построенных по правилу (5.1), сходится к квазитреугольной матрице, причем для нижней внедиагональной клетки S_j матрицы A_j имеет место оценка

$$\|S_j\| \leq 2(1 + \varepsilon_5) \|A\| \mu(T) \left[\left(\frac{2}{\delta^{3/2} \delta_1} \right)^k L^j + \delta^6 \eta(\varepsilon) \right],$$

где

$$L = \frac{\rho \|A\| \sqrt{1 - \delta_2^3} + \mu(T) \cdot \varepsilon (\delta^3 + \delta^6 \operatorname{ctg} \varphi_0)}{\rho \|A\| \sqrt{1 + 2\delta_2^3} - \mu(T) \cdot \varepsilon (\delta^{3/2} + \operatorname{tg} \varphi_0)},$$

$$\eta(\varepsilon) = \frac{\varepsilon}{2\delta_2^3 [\rho \|A\| - \varepsilon \mu(T) + \sqrt{[\rho \|A\| - \varepsilon \mu(T)]^2 - [\mu(T) \varepsilon \delta]^2}},$$

$$\varepsilon_1 = 2\mu(T) \delta^{3/2} \varepsilon, \quad \varepsilon_2 = \mu(T) \varepsilon, \quad \varepsilon_3 = \mu(T) \delta^6 \varepsilon, \quad \varepsilon_4 = \mu(T) \delta^3 \varepsilon,$$

и $L < 1$, если $\mu(T) \delta^6 \eta(\varepsilon) < \operatorname{tg} \varphi_0 < [\mu(T) \eta(\varepsilon)]^{-1}$.

Неравенство $\mu(T) \delta^6 \eta(\varepsilon) < \operatorname{tg} \varphi_0$ можно считать всегда выполненным, поскольку если это не так, то мы сразу получим $\sin \varphi_0 \leq \mu(T) \delta^6 \eta(\varepsilon)$. Если последнее неравенство выполнено для всех $f \in \mathcal{L}_1^1$, то из (5.9) следует, что $\|S_j\| \leq 2(1 + \varepsilon_5) \|A\| \cdot \mu(T) \delta^6 \eta(\varepsilon)$. Для того чтобы удовлетворять второму неравенству, достаточно положить $\delta_1 = 2\delta^{-3/2} \cdot \sqrt{\mu^2(T) \eta(\varepsilon)}$.

§ 6. НЕПРЕРЫВНАЯ ЗАВИСИМОСТЬ ПАРАМЕТРОВ δ И δ_1 ОТ МАТРИЦЫ A

Для получения квалифицированных оценок скорости сходимости ортогонально-степенного метода Воеводина использовались два предположения об исходной матрице A . Они сформулированы в виде неравенств для наименьших сингулярных чисел некоторых матриц, зависящих от параметров. Напомним эти предположения.

1. Гипотеза о разделении спектра: существуют числа ρ, δ ($0 < \delta < \rho < 1$) такие, что выполняется неравенство

$$\min_{\|x\| \neq 0} \frac{\|(A - \rho \|A\| e^{i\varphi} I)x\|}{\|x\|} > \delta \|A\| \quad (6.1)$$

для любого вектора x и для всех $0 \leq \varphi \leq 2\pi$.

2. Вторая гипотеза выражена в виде неравенства

$$\sigma_1[\hat{A}(\lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k)})] \geq (\delta_1 \|A\|)^k \quad (6.2)$$

для всех $\lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k)}$ таких, что $|\lambda^{(j)}| \leq \rho \|A\|$, где $\hat{A}(\lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k)})$ — матрица, составленная из первых $N - k$ столбцов матрицы произведения $(A - \lambda^{(1)} I)(A - \lambda^{(2)} I) \dots (A - \lambda^{(k)} I)$. Исследуем оценки (6.1) и (6.2) для возмущенной матрицы $A + B$.

Покажем, что

$$\min_{\|x\| \neq 0} \frac{\|(A+B-\rho\|A\|e^{i\varphi}I)x\|}{\|x\|} > \left(\delta - \frac{\|B\|}{\|A\|}\right)\|A\|, \quad (6.3)$$

$$\sigma_1[(A+B)(\lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k)})] \geq (\delta_1\|A\|)^k - 2^{2k}\|A\|^k \frac{\|B\|/\|A\|}{1 - \|B\|/\|A\|}, \quad (6.4)$$

где через $(A+B)(\lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k)})$ обозначена матрица, составленная из первых $N-k$ столбцов матрицы произведения $(A+B-\lambda^{(1)}I)(A+B-\lambda^{(2)}I) \dots (A+B-\lambda^{(k)}I)$.

Приступая к доказательству, отметим очевидное неравенство

$$\sigma_1(A+B) \geq \sigma_1(A) - \sigma_N(B). \quad (6.5)$$

В силу (6.5) и (6.1) имеем

$$\begin{aligned} \min_{\|x\| \neq 0} \frac{\|(A+B-\rho\|A\|e^{i\varphi}I)x\|}{\|x\|} &= \sigma_1(A+B+\rho\|A\|e^{i\varphi}I) \geq \\ &\geq \sigma_1(A-\rho\|A\|e^{i\varphi}I) - \sigma_N(B) = \min_{\|x\| \neq 0} \frac{\|(A-\rho\|A\|e^{i\varphi}I)x\|}{\|x\|} - \\ &- \sigma_N(B) > \delta\|A\| - \|B\| = (\delta - \|B\|/\|A\|)\|A\|, \end{aligned}$$

из которого следует неравенство (6.3). Для получения оценки (6.4)

$\sigma_1[(A+B)(\lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k)})]$ перепишем следующим образом:

$$\begin{aligned} \sigma_1[(A+B)(\lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k)})] &= \min_{f \in \mathcal{L}_1^1, \|f\| \neq 0} \times \\ &\times \frac{\|(A+B-\lambda^{(1)}I)(A+B-\lambda^{(2)}I) \dots (A+B-\lambda^{(k)}I)f\|}{\|f\|}. \end{aligned}$$

Матрицу $(A+B-\lambda^{(1)}I)(A+B-\lambda^{(2)}I) \dots (A+B-\lambda^{(k)}I)$ представим в виде

$$\begin{aligned} &(A+B-\lambda^{(1)}I)(A+B-\lambda^{(2)}I) \dots (A+B-\lambda^{(k)}I) = \\ &= (A-\lambda^{(1)}I)(A-\lambda^{(2)}I) \dots (A-\lambda^{(k)}I) + \Phi(A, B, \lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k)}), \end{aligned}$$

где через $\Phi(A, B, \lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k)})$ обозначена та часть произведения, в которой в качестве множителя участвует матрица B . Тогда

$$\begin{aligned} \sigma_1[(A+B)(\lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k)})] &= \sigma_1[\hat{A}(\lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k)}) + \\ &+ \hat{\Phi}(A, B, \lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k)})], \text{ где } \hat{\Phi}(A, B, \lambda^{(1)}, \dots, \lambda^{(k)}) \end{aligned}$$

— прямоугольная матрица, составленная из первых $N-k$ столбцов матрицы $\hat{\Phi}(A, B, \lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k)})$. Согласно (6.2) и (6.5)

$$\begin{aligned} \sigma_1[\hat{A}(\lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k)}) + \hat{\Phi}(A, B, \lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k)})] &\geq \\ \geq \sigma_1[\hat{A}(\lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k)})] - \sigma_N[\hat{\Phi}(A, B, \lambda^{(1)}, \dots, \lambda^{(k)})] &\geq \\ \geq (\delta_1\|A\|)^k - \sigma_N[\hat{\Phi}(A, B, \lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k)})]. \end{aligned}$$

Таким образом,

$$\sigma_1[(A+B)(\lambda^{(1)}, \dots, \lambda^{(k)})] \geq (\delta_1\|A\|)^k - \sigma_N[\hat{\Phi}(A, B, \lambda^{(1)}, \dots, \lambda^{(k)})]. \quad (6.6)$$

Оценим $\sigma_N[\hat{\Phi}(A, B, \lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k)})]$. Легко проверить, что для

$\sigma_N[\widehat{\Phi}(A, B, \lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k)})]$ выполнена следующая цепочка неравенств:

$$\begin{aligned} \sigma_N[\widehat{\Phi}(A, B, \lambda^{(1)}, \dots, \lambda^{(k)})] &\leq C_k^1 (2\|A\|)^{k-1} \|B\| + \\ &+ C_k^2 (2\|A\|)^{k-2} \|B\|^2 + \dots + C_k^k (2\|A\|)^0 \|B\|^k \leq \\ &\leq (2\|A\|)^{k-1} \|B\| \left(C_k^1 + C_k^2 \frac{\|B\|}{2\|A\|} + \dots + C_k^k \left(\frac{\|B\|}{2\|A\|} \right)^{k-1} \right). \end{aligned}$$

Откуда, используя, что $\max_{m \leq k} C_k^m < 2^k$, запишем

$$\sigma_N[\widehat{\Phi}(A, B, \lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k)})] \leq 2^{2k} (\|A\|)^k \frac{\|B\|}{\|A\|} \left(1 + \frac{\|B\|}{2\|A\|} + \dots + \left(\frac{\|B\|}{2\|A\|} \right)^k \right).$$

Предполагая, что $\|B\|/\|A\| < 1$, находим

$$\sigma_N[\widehat{\Phi}(A, B, \lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k)})] \leq 2^{2k} \|A\|^k \frac{\|B\|/\|A\|}{1 - \|B\|/2\|A\|} < 4^k \|A\|^k \frac{\|B\|/\|A\|}{1 - \|B\|/\|A\|}.$$

Полученная оценка с учетом (6.6) доказывает неравенство (6.4).

§ 7. ОЦЕНКА СКОРОСТИ СХОДИМОСТИ QR-АЛГОРИТМА ФРЕНСИСА — КУБЛАНОВСКОЙ

Будем исследовать QR-алгоритм Френсиса — Кублановской, для которого установим теорему, аналогичную теореме из § 5.

Пусть A — произвольная матрица порядка N . Построим последовательность ортогональных матриц Q_i и правых треугольных матриц R_i так, чтобы выполнялись формулы

$$\begin{aligned} A &= Q_1 R_1, \quad A_1 = R_1 Q_1; \\ A_1 &= Q_2 R_2, \quad A_2 = R_2 Q_2; \\ &\dots \dots \dots \\ A_{l-1} &= Q_l R_l, \quad A_l = R_l Q_l. \end{aligned} \tag{7.1}$$

Легко показать, что для всех l матрицы A_l из (7.1) ортогонально подобны исходной матрице A . В самом деле

$$\begin{aligned} A_l &= R_l Q_l = Q_l^* (Q_l R_l) Q_l = Q_l^* (R_{l-1} Q_{l-1}) Q_l = Q_l^* Q_{l-1}^* (Q_{l-1} R_{l-1}) Q_{l-1} Q_l = \\ &= \dots = (Q_1 Q_2 \dots Q_l) A Q_1 Q_2 \dots Q_l. \end{aligned}$$

Обозначив $P_l = Q_1 Q_2 \dots Q_l$, получим

$$A_l = P_l^* A P_l. \tag{7.2}$$

Так как матрицы Q_i ортогональные, то будут ортогональными и матрицы P_i . Поэтому A_l ортогонально подобны A .

В [11] доказано, что если собственные значения матрицы удовлетворяют неравенствам

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_{N-k}| > |\lambda_{N-k+1}| \geq \dots \geq |\lambda_N| > 0$$

и прямая сумма корневых подпространств, отвечающих собственным значениям $\lambda_1, \lambda_2, \dots, \lambda_{N-k}$, имеет нулевое пересечение с подпространством, натянутым на координатные векторы $f_{N-k+1}, f_{N-k+2}, \dots, f_N$, то при достаточно большом l матрица A_l , построенная по правилу (7.1), станет сколь угодно близкой к правой квазиреугольной матрице.

Каждый шаг QR-алгоритма состоит из разложения невырожденной квадратной матрицы в виде произведения ортогональной и правой треугольной матрицы с последующим перемножением полученных Q_i и R_i в обратном порядке. Исследуем алгоритм (7.1), учитывая погрешности

$$\begin{aligned} A &= Q_1 R_1 - E_1^{(1)}, \quad A_1 = R_1 Q_1 + E_1^{(2)}; \\ A_1 &= Q_2 R_2 - E_2^{(1)}, \quad A_2 = R_2 Q_2 + E_2^{(2)}; \\ &\dots\dots\dots \\ A_{l-1} &= Q_l R_l - E_l^{(1)}, \quad A_l = R_l Q_l + E_l^{(2)}. \end{aligned} \tag{7.3}$$
$$A_l = P_l^* A P_l + E_l, \quad (7.4)$$
$$\begin{aligned} P_i &= Q_1 Q_2 \dots Q_i; \\ E_i &= Q_i^* Q_{i-1}^* \dots Q_2^* \widehat{E}_1 Q_2 \dots Q_i + Q_i^* Q_{i-1}^* \dots Q_3^* \widehat{E}_2 Q_3 \dots Q_i + \\ &\quad + \dots + Q_i^* \widehat{E}_{i-1} Q_i + \widehat{E}_i; \\ \widehat{E}_i &= Q_i^* E_i^{(1)} Q_i + E_i^{(2)}. \end{aligned}$$

Для оценки скорости сходимости алгоритма (7.3) поступим следующим образом. Для того чтобы матрица A_1 была клеточно-треугольной (или почти клеточно-треугольной), достаточно, чтобы первые $N-k$ вектор-столбцов матрицы P_i лежали (или почти лежали) на инвариантном подпространстве $\mathcal{L}_1$. Чтобы установить описанное свойство P_i , покажем малость углов φ_i , определенных из равенства

$$\cos \varphi_j = \max_{x \in \mathcal{L}} \frac{|(z, x)|}{\|z\| \cdot \|x\|}$$

$$\begin{aligned} P_l U_l &= Q_1 Q_2 \dots Q_{l-1} (Q_l R_l) R_{l-1} \dots R_2 R_1 = \\ &= P_{l-1} (A_l + E_l^{(1)}) U_{l-1} = P_{l-1} A_l U_{l-1} + P_{l-1} E_l^{(1)} U_{l-1}. \end{aligned} \quad (7.5)$$
$$P_{l-1}A_l = AP_{l-1} + P_{l-1}E_{l-1}.$$
$$\begin{aligned} P_l U_l &= A P_{l-1} U_{l-1} + P_{l-1} (E_{l-1} + E_l^{(1)}) U_{l-1} = \\ &= (A + P_{l-1} (E_{l-1} + E_l^{(1)}) P_{l-1}^*) P_{l-1} U_{l-1} = \\ &= (A + \overline{E}_l) P_{l-1} U_{l-1} = \dots = (A + \overline{E}_l) (A + \overline{E}_{l-1}) \dots (A + \overline{E}_1). \end{aligned}$$
$$P_l U_l = (A + \bar{E}_l) (A + \bar{E}_{l-1}) \dots (A + \bar{E}_1).$$

Следовательно, чтобы исследовать поведение первых $N - k$ столбцов матрицы P_i , достаточно оценить углы φ между векторами y_i и подпространством $\mathcal{L}_i$.

Здесь

$$y_l = (\hat{A} + \hat{E}_l)(\hat{A} + \hat{E}_{l-1}) \dots (\hat{A} + \hat{E}_1)y_0;$$

$$\hat{A} = TAT^{-1}, \|\hat{E}_l\| \leq 2\mu(T) \cdot l \cdot \varepsilon, y_0 \in \mathcal{L}_1.$$

Далее, дословным повторением рассуждений § 4 и 5 можно установить следующую теорему.

Теорема. Если для матрицы A выполнены гипотезы I и II и, кроме того, $[\rho \|A\| (\sqrt{1 + 2\delta_2^3} - \sqrt{1 - \delta_2^3}) - 2l(\varepsilon_1 + \varepsilon_4)]^3 - 16l^2\varepsilon_2\varepsilon_3 > 0$, то последовательность матриц A_l , построенных по правилу (7.3), сходится к квазирегулярной матрице, причем для нижней внедиагональной клетки матрицы A_l имеет место оценка

$$\|S_l\| \leq 2\mu(T) \|A\| \left[\left(\frac{2}{\delta_2^{3/2}\delta_1} \right)^l L^l(l) + \delta^6 \eta(l, \varepsilon) \right],$$

где

$$L(l) = \frac{\rho \|A\| \sqrt{1 - \delta_2^3} + 2\mu(T) l \varepsilon (\delta^3 + \delta^6 \operatorname{ctg} \varphi_0)}{\rho \|A\| \sqrt{1 + 2\delta_2^3} - 2\mu(T) l \varepsilon (\delta^{3/2} + \operatorname{tg} \varphi_0)};$$

$$\eta(l, \varepsilon) = \frac{2l\varepsilon}{2\delta_2^3 [\rho \|A\| - 2l\varepsilon\mu(T) + \sqrt{(\rho \|A\| - 2l\varepsilon\mu(T))^2 - (2l\varepsilon\delta\mu(T))^2}}.$$

§ 8. РЕЗУЛЬТАТЫ ЧИСЛЕННЫХ ЭКСПЕРИМЕНТОВ

Представлены результаты численных экспериментов, проведенных для ортогонально-степенного метода. Не будем приводить расчетные формулы метода — их можно найти в [22]. Метод отражений, который используется в [22] для вычисления матриц Q_{i+1} по матрицам AQ_i , довольно подробно описан в монографиях В. В. Воеводина [15], С. К. Годунова [13], Дж. Х. Уилкинсона [14]. Отметим только, что все расчеты осуществлялись на машине ЕС-1050 с двойной точностью.

Серия экспериментов проведена для семейства четырехпараметрических матриц $A(r_1, r_2, t_1, t_2)$ размерности $N \times N$:

$$A(r_1, r_2, t_1, t_2) = C + B + \Lambda + D,$$

где $C = (c_{ij})$ такова, что

$$c_{ij} = \begin{cases} r_1, & \text{если } i = j \leq N_1; \\ r_2, & \text{если } N_1 \leq i = j \leq N; \\ t_1, & \text{если } i = j - 1, j \neq 1, i, j \leq N_1; \\ t_2, & \text{если } i = j - 1, N_1 + 1 \leq i, j \leq N; \end{cases}$$

$c_{ij} = 0$ в противном случае.

Оставшиеся матрицы $B = (b_{ij})$, $\Lambda = (\lambda_{ij})$, $D = (d_{ij})$ определяются при помощи вектора p с координатами

$$p_i = \sin(i/7) \sqrt{2} / \sqrt{\sum_{i=1}^N [\sin(i/7)]^2}.$$

по следующим формулам:

$$d_{ij} = - \sum_{m=1}^2 t_m p_{N_1} p_i p_{N_1+1} p_j;$$

$$\lambda_{ij} = \begin{cases} 2p_i p_j r_1 + (p_{i+1} p_j + p_i p_{j+1}) t_1 & \text{при } j \neq N_1, i \neq N_1; \\ 2p_i p_j r_1 + p_j p_{i+1} t_1 & \text{при } j = N_1, i \neq N_1; \\ 2p_i p_j r_1 + p_i p_{j+1} t_1 & \text{при } j \neq N_1, i = N_1; \\ 2p_i p_j r_1 & \text{при } j = N_1, i = N_1; \\ 2p_i p_j r_2 + (p_{i+1} p_j + p_i p_{j+1}) t_2 & \text{при } N_1 + 1 \leq j \neq N, \\ & N_1 + 1 \leq i < N; \\ 2p_i p_j r_2 + p_j p_{j+1} t_2 & \text{при } N_1 + 1 \leq j = N, \\ & N_1 + 1 \leq i < N; \\ 2p_i p_j r_2 + p_i p_{j+1} t_2 & \text{при } N_1 + 1 \leq j \neq N, \\ & N_1 + 1 \leq i \leq N; \\ 2p_i p_j & \text{при } N_1 + 1 \leq i; j = N, \end{cases}$$

$$b_{ij} = \begin{cases} p_i p_j \sum_{m=1}^{N_1-1} p_m p_{m+1} t_1 + 2r_1 & \text{при } 1 \leq i, j \leq N_1; \\ p_i p_j \sum_{m=N_1+1}^{N-1} p_m p_{m+1} t_2 + 2r_2 & \text{при } N_1 + 1 \leq i, j \leq N. \end{cases}$$

Рассматриваемый класс матриц примечателен тем, что он ортогонально подобен некоторой двухдиагональной матрице вида

$$Z = \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix},$$

где Z_1 — квадратная двухдиагональная матрица порядка $N_1 \times N_1$, по главной диагонали которой стоят числа r_1 , а на наддиагонали числа t_1 . Z_2 — матрица того же вида, что и Z_1 , порядка $(N - N_1) \times (N - N_1)$, по главной диагонали которой стоят числа r_2 , а на наддиагонали числа t_2 . Как указывалось во введении, матрицы такого класса при определенных значениях параметров t_1, t_2 имеют непрерывный «почти спектр», заполняющий некоторые двумерные области комплексной плоскости. В следующих примерах изучалось поведение метода Воеводина на семействе матриц $A(r_1, r_2, t_1, t_2)$ при различных значениях параметров r_1, r_2, t_1, t_2 .

Отметим, что для матриц $A(r_1, r_2, t_1, t_2)$ гипотеза II выполнена.

Пример 1. Рассмотрим матрицу $A(1, 9, t_1, t_2)$ размера 20×20 . Для нее при разных значениях параметров t_1 и t_2 применялся ортогонально-степенной метод и исследовалось поведение максимального элемента κ поддиагональной прямоугольной матрицы S :

$$\kappa = \max_{\substack{N_1 < i < N \\ 1 < j < N - N_1}} |s_{ij}|.$$

Результаты расчета отражены в табл. 1, где M — число итераций.

Здесь приведены расчеты для двух матриц: $A(1, 9, 10, 10)$ и $A(1, 9, 10, 20)$. У матрицы $A(1, 9, 10, 10)$ почти собственные значения сосредоточены у единицы и у точки 9 достаточно хорошо разделены, следствием чего и является высокая точность выделения клеток. В случае матрицы $A(1, 9, 10, 20)$ при тех же значениях параметров r_1 и r_2 увеличение параметра t_2 приводит к тому, что около точки 9 увеличива-

Таблица 1

Значение	$t_1=t_2=10, \quad N=20, \quad N_1=10$			$t_1=10, \quad t_2=20, \quad N=20, \quad N_1=10$		
M	100	200	300	100	200	300
κ	10^{-10}	10^{-11}	10^{-11}	10^{-6}	10^{-7}	10^{-7}

Таблица 2

Значение	$r_1=5, r_2=4, N_1=15, t_i=10$			$r_1=1, r_2=2, N_1=10, t_1=t_2=10$			$r_1=1, r_2=2, N_1=10, t_1=t_2=30$		
	M	κ		M	κ		M	κ	
	100	200	300	100	200	300	100	200	300
	0,3	-0,5	10^{-1}	10^{-2}	1	10^{-1}	3,5	11	13

ется двумерная область, точки которой являются «почти» собственными значениями. Следовательно, кольцо, разделяющее пятна спектра, сужается (т. е. уменьшается величина δ), что, по-видимому, приводит к потере точности выделения клеток.

Пример 2. Рассмотрим то же свойство матриц $A(r_1, r_2, t_1, t_2)$ порядка 30×30 . В табл. 2 приведены результаты расчетов для трех матриц из этого семейства.

Для этих трех матриц гипотеза I не выполнена, следствием чего и является совершенно произвольное поведение κ в первом и втором случае. Для матрицы $A(1, 2, 30, 30)$ значения κ даже растут. В примерах 1 и 2 исследовался случай, когда гипотеза II выполнена. Следующий пример посвящен изучению случая, когда не выполняется гипотеза II.

Пример 3. Рассмотрим однопараметрическое семейство симметрических трехдиагональных матриц размерности 5×5 :

$$A(\tau) = \begin{bmatrix} 7 & 1 & & & 0 \\ 1 & 6+\tau & \tau & & \\ & \tau & 5+2\tau & \tau & \\ 0 & & \tau & 6+\tau & 1 \\ & & & 1 & 7 \end{bmatrix}.$$

Собственные значения матрицы $A(0)$ следующие: $\lambda_1 = \lambda_2 \approx 7,628$; $\lambda_3 = \lambda_4 \approx 5,382$; $\lambda_5 = 5$. Легко показать, что для достаточно маленьких τ собственные значения матрицы $A(\tau)$ будут лежать в интервалах

$$(7,628 - 3\tau, 7,628 + 3\tau), (5,382 - 3\tau, 5,382 + 3\tau), (5 - 3\tau, 5 + 3\tau).$$

Следовательно, $A(\tau)$ имеет изолированное собственное значение, т. е. для него выполнена гипотеза I. Вторая гипотеза для этих матриц не выполнена. В самом деле собственный вектор матрицы $A(0)$, соответствующий собственному значению $\lambda = 5$, имеет вид $f = (0, 0, 1, 0, 0)'$. Это значит, что третий вектор исходного базиса лежит в инвариантном подпространстве, отвечающем изолированному собственному значению. Легко проверить, что собственный вектор матрицы $A(\tau)$ для изолированного собственного значения из интервала $(5 - 3\tau, 5 + 3\tau)$ будет иметь вид $f = (\tau, 2\tau, 1, 2\tau, \tau)'$ и ясно, что при малых τ угол между третьим базисным вектором и вектором f будет пренебрежимо малым.

В табл. 3 приведены результаты расчетов для $A(\tau)$ при различных значениях параметра τ .

В табл. 4 показаны результаты экспериментов для того же семейства $A(\tau)$, но только после перенумерации базиса по следующему правилу:

$$e_1^{(1)} = e_1, \quad e_2^{(1)} = e_2, \quad e_3^{(1)} = e_5, \quad e_4^{(1)} = e_4, \quad e_5^{(1)} = e_3.$$

Таблица 3

Значение	$\tau=10^{-6}$			$\tau=10^{-7}$			$\tau=10^{-8}$			$\tau=10^{-10}$		
	M	κ		M	κ		M	κ		M	κ	
	100	200	300	100	200	300	100	200	300	100	200	300
	10^{-3}	10^{-2}	10^{-4}	10^{-4}	0,15	10^{-4}	10^{-6}	10^{-3}	10^{-2}	10^{-7}	10^{-4}	0,18

Таблица 4

Значе- ние	$\tau=10^{-6}$			$\tau=10^{-7}$			$\tau=10^{-8}$			$\tau=10^{-10}$		
	100	200	300	100	200	300	100	200	300	100	200	300
M	100	200	300	100	200	300	100	200	300	100	200	300
$\%$	10^{-9}	10^{-12}	10^{-16}	10^{-10}	10^{-13}	10^{-17}	10^{-12}	10^{-16}	10^{-19}	10^{-13}	10^{-17}	10^{-20}

В новом базисе для семейства $A(\tau)$ выполняются обе гипотезы и видно (см. табл. 4), что клетки выделяются с достаточно высокой точностью.

ЛИТЕРАТУРА

1. Воеводин В. В. Численные методы алгебры (теория и алгоритмы).— М.: Наука, 1966.— 248 с.
2. Воеводин В. В. Ошибки округления и устойчивость в прямых методах линейной алгебры.— М.: ВЦ МГУ, 1969.— 153 с.
3. 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.
4. Фаддеева В. Н., Кузнецов Ю. А. и др. Вычислительные методы линейной алгебры. Библиографический указатель, 1928—1974 гг.— Новосибирск, 1976.— 418 с.
5. Фаддеева В. Н. и др. Вычислительные методы линейной алгебры. Библиографический указатель, 1975—1980 гг.— Л., 1982.— 344 с.
6. Wilkinson J. H. Rounding errors in algebraic processes.— Notes Appl. Sci., 1963, N. 32.
7. Фаддеев Д. К., Фаддеева В. Н. Вычислительные методы линейной алгебры.— М.— Л.: Физматгиз, 1963.— 734 с.
8. Воеводин В. В. Некоторые методы решения полной проблемы собственных значений.— Журн. вычисл. математики и мат. физики, 1962, т. 2, № 1, с. 15—24.
9. Francis J. G. F. The QR-transformation. A unitary analogue to the LR-transformation. I.— Comput. J., 1961, v. 4, N 3, p. 265—271.
10. Francis J. G. F. The QR-transformation. A unitary analogue to the LR-transformation. II.— Comput. J., 1962, v. 4, N 4, p. 332—345.
11. Кублановская В. Н. Решение проблемы собственных значений для произвольной матрицы.— Тр. мат. ин-та АН СССР, 1962, т. 66, с. 113—135.
12. Воеводин В. В. Решение полной проблемы собственных значений степенными методами.— В кн.: Вычислительные методы и программирование. М.: МГУ, 1965, с. 7—55.
13. Годунов С. К. Решение систем линейных уравнений.— Новосибирск: Наука, Сиб. отд-ние, 1980.— 177 с.
14. Уилкинсон Дж. Х. Алгебраическая проблема собственных значений.— М.: Наука, 1970.— 564 с.
15. Воеводин В. В. Вычислительные основы линейной алгебры.— М.: Наука, 1977.— 304 с.
16. Годунов С. К., Булгаков А. Я. Устойчивость устойчивых матриц.— В кн.: Кубатурные формулы и вычислительная математика. Новосибирск: Наука, Сиб. отд-ние, 1980, с. 18—28.
17. Булгаков А. Я. Эффективно вычисляемый параметр качества устойчивости систем линейных дифференциальных уравнений с постоянными коэффициентами.— Сиб. мат. журн., 1980, т. 21, № 3, с. 32—41.
18. Раззаков Ш. И. Квалифицированные оценки скорости сходимости ортогонально-степенного метода Воеводина для произвольных матриц.— Новосибирск, 1983.— 36 с. (Препринт СО АН СССР, Сиб. отд-ние, ВЦ).
19. Гантмахер Р. Ф. Теория матриц. 3-е изд.— М.: Наука, 1967.— 576 с.
20. Беллман Р. Введение в теорию матриц.— М.: Наука, 1969.— 367 с.
21. Широков П. А. Тензорное исчисление. Казань: Изд-во Казан. ун-та, 1961.— 447 с.
22. Андреева Л. П. О практическом применении ортогонально-степенного метода.— В кн.: Вычислительные методы и программирование. М.: МГУ, 1965, с. 61—89.