Итерационные методы решения проблемы собственных значений

Метод вращений (метод Якоби)

Для симметричной матрицы A при отыскании собственных значений и собственных векторов в настоящее время наиболее употребительным является метод вращений (метод Якоби). При его обосновании исходят из того, что определение собственных значений и собственных векторов симметричной матрицы A равносильно построению диагональной матрицы Λ и ортогональной матрицы T, связанных соотношением

$$\Lambda = T^{-1} A T = T^{\mathrm{T}} A T . \tag{1}$$

Диагональные элементы матрицы Λ будут искомыми собственными значениями, а столбцы матрицы T — столбцами координат собственных векторов, соответствующих этим собственным значениям. При приближенном вычислении матриц Λ и T строят последовательности матриц

$$A_0 = A, \quad A_1, \quad A_2, \quad \ldots \quad A_k, \quad \ldots \quad \to \Lambda ,$$

$$T_0 = E, \quad T_1, \quad T_2, \quad \ldots \quad T_k, \quad \ldots \quad \to T ,$$

по формулам

$$A_{k+1} = T_{ij}^{\mathrm{T}} A_k T_{ij}, \qquad T_{k+1} = T_k T_{ij}, \tag{2}$$

где T_{ij} — матрица простых вращений. Матрица T_k равна произведению всех матриц T_{ij}, примененных при построении матриц A_0, A_1, A_2, … , A_k, причем матрицы T_{ij} в этом произведении перемножаются слева направо в том порядке, в каком они применялись. На k-м шаге принимают $\Lambda \approx A_k$, $T \approx T_k$.

Матрицу T_{ij} в формуле (2) строят следующим образом. В матрице A_k выбирают наибольший по модулю недиагональный элемент $a_{ij}^{(k)}$ и строят матрицу простого вращения

$$T_{ij} = \begin{array}{c} \begin{pmatrix} 1 & & & \cdot & & & & \cdot & & & \\ & \cdot & & \cdot & & & & \cdot & & & \\ & & 1 & \cdot & & & & \cdot & & & \\ \dots & \dots & \dots & \cos\alpha & \dots & \dots & \dots & -\sin\alpha & \dots & \dots & \dots \\ & & & \cdot & 1 & & & \cdot & & & \\ & & & \cdot & & \cdot & & \cdot & & & \\ & & & \cdot & & & 1 & \cdot & & & \\ \dots & \dots & \dots & \sin\alpha & \dots & \dots & \dots & \cos\alpha & \dots & \dots & \dots \\ & & & \cdot & & & & \cdot & 1 & & \\ & & & \cdot & & & & \cdot & & \cdot & \\ & & & \cdot & & & & \cdot & & & 1 \end{pmatrix} \end{array} \begin{array}{l} \\ \\ \\ i\text{-я} \\ \text{строка} \\ \\ \\ \\ j\text{-я} \\ \text{строка} \\ \\ \\ \end{array}, \quad (3)$$

выбирая угол α так, чтобы в матрице A_{k+1} обратился в нуль элемент $a_{ij}^{(k+1)}$. Заметим, что если в правой части первого из соотношений (2) провести умножение матриц, то элемент $a_{ij}^{(k+1)}$ матрицы A_{k+1} будет иметь вид

$$a_{ij}^{(k+1)} = (\cos^2\alpha - \sin^2\alpha)\, a_{ij}^{(k)} + \cos\alpha \cdot \sin\alpha\, (a_{jj}^{(k)} - a_{ii}^{(k)}).$$

Из равенства нулю этого выражения получаем

$$\operatorname{tg} 2\alpha = \frac{2a_{ij}^{(k)}}{(a_{ii}^{(k)} - a_{jj}^{(k)})}, \quad |\alpha| \le \frac{\pi}{4}. \qquad (4)$$

Чтобы записать матрицу T_{ij}, нужно знать $\cos\alpha$ и $\sin\alpha$. Их легко найти по формулам

$$\cos 2\alpha = \frac{1}{\sqrt{1 + \operatorname{tg}^2 2\alpha}},$$

$$\cos\alpha = \sqrt{\frac{1+\cos 2\alpha}{2}}, \quad \sin\alpha = \pm\sqrt{\frac{1-\cos 2\alpha}{2}}. \qquad (5)$$

Чтобы выполнялось условие $|\alpha| \le \frac{\pi}{4}$, знак у корня в формуле для $\sin\alpha$ выбирают тот же, что и у выражения

$$a_{ij}^{(k)}(a_{ii}^{(k)} - a_{jj}^{(k)}). \qquad (6)$$

При вычислении элементов матрицы удобно пользоваться формулами

$$a_{ml}^{(k+1)} = \begin{cases} a_{ml}^{(k)}, & m,l \neq i,j; \\ a_{mi}^{(k)} \cos\alpha + a_{mj}^{(k)} \sin\alpha, & m = i,\ l \neq i,j \text{ или } l = i,\ m \neq i,j; \\ -a_{mi}^{(k)} \sin\alpha + a_{mj}^{(k)} \cos\alpha, & m = j,\ l \neq i,j \text{ или } l = j,\ m \neq i,j; \\ a_{ii}^{(k)} \cos^2\alpha + 2a_{ij}^{(k)} \cos\alpha \sin\alpha + \\ \qquad + a_{jj}^{(k)} \sin^2\alpha, & m = l = i; \\ a_{ii}^{(k)} \sin^2\alpha - 2a_{ij}^{(k)} \cos\alpha \sin\alpha + \\ \qquad + a_{jj}^{(k)} \cos^2\alpha, & m = l = j; \\ (\cos^2\alpha - \sin^2\alpha)\, a_{ij}^{(k)} + \\ \quad + \cos\alpha \sin\alpha\, (a_{jj}^{(k)} - a_{ii}^{(k)}), & m = i,\ l = j \text{ или } m = j,\ l = i, \end{cases}$$

которые получаются при сравнении элементов матриц левой и правой частей первого равенства из (2) (элементы $a_{ij}^{(k+1)}$, $a_{ji}^{(k+1)}$ при выбранном α по формулам (4) и (5) равны нулю).

Метод вращений обладает высокой скоростью сходимости.

Для доказательства этого утверждения сумму квадратов всех недиагональных элементов матрицы A_k обозначим через $t(A_k)$. Учитывая, что матрица A_{k+1} отличается от матрицы A_k лишь строками и столбцами с номерами i, j, из равенства $A_{k+1} = T_{ij}^{-1} A_k T_{ij}$ непосредственным подсчетом элементов в правой части при выбранном угле α в соответствии с формулой (4) получим соотношение

$$t(A_{k+1}) = t(A_k) - 2(a_{ij}^{(k)})^2. \qquad (7)$$

В то же время, поскольку $a_{ij}^{(k)}$ — максимальный по абсолютной величине недиагональный элемент матрицы A_k и всего недиагональных элементов в этой матрице $n(n-1)$, то выполняется неравенство

$$t(A_k) \leq n(n-1)(a_{ij}^{(k)})^2.$$

Следовательно,

$$(a_{ij}^{(k)})^2 \geq \frac{1}{n(n-1)} \cdot t(A_k).$$

Тогда из равенства (7) получаем

$$t(A_{k+1}) \leq t(A_k) - \frac{2}{n(n-1)} \cdot t(A_k) = q \cdot t(A_k),$$

где $q = 1 - \dfrac{2}{n(n-1)}$, причем $0 \leq q < 1$, так как $n \geq 2$.

Отсюда вытекает цепочка неравенств

$$t(A_k) \leq q \cdot t(A_{k-1}) \leq q^2 \cdot t(A_{k-2}) \leq \ldots \leq q^k \cdot t(A_0),$$

и поскольку $A_0 = A$, то

$$t(A_k) \leq q^k \cdot t(A).$$

Таким образом, $t(A_k) \to 0$ при $k \to \infty$ со скоростью, не меньшей скорости сходимости геометрической прогрессии со знаменателем $q = 1 - \dfrac{2}{n(n-1)} < 1$.

Оценим предельное количество арифметических операций, требуемых для выполнения одного шага алгоритма метода вращений, т. е. для вычисления вектора $X^{(k+1)}$ и $\lambda_i^{(k)}$:

$$C(n) = 2n^2 + O(n) \ (n \to \infty)$$

Пример Методом вращений найти собственные значения и собственные векторы матрицы

$$A = \begin{pmatrix} 1 & -3 & -1 \\ -3 & 1 & 1 \\ -1 & 1 & 5 \end{pmatrix}.$$

Решение. Положим $A_0 = A$ и по формуле (2) при $k = 0$ будем строить матрицы A_1 и T_1. Так как в матрице $A_0 = A$ $\max \left| a_{ij} \right| = \left| a_{12} \right| = 3$, то $i = 1$, $j = 2$ и

$$T_{ij} = T_{12} = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Угол 2α определяется по формуле

$$\operatorname{tg} 2\alpha = \frac{2a_{ij}}{a_{ii} - a_{jj}} = \frac{-6}{1-1} \to -\infty .$$

Такое значение $\operatorname{tg} 2\alpha$ указывает на то, что угол $2\alpha = -\frac{\pi}{2}$. Следовательно, $\alpha = -\frac{\pi}{4}$, $\cos\alpha = \frac{1}{\sqrt{2}}$, $\sin\alpha = -\frac{1}{\sqrt{2}}$ и

$$T_{12} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

По формулам (2) вычислим матрицы A_1 и T_1:

$$A_1 = T_{12}^{-1} A_0 T_{12} = \begin{pmatrix} 4 & 0 & -\sqrt{2} \\ 0 & -2 & 0 \\ -\sqrt{2} & 0 & 5 \end{pmatrix},$$

$$T_1 = T_0 T_{12} = E T_{12} = T_{12}.$$

По тем же формулам (2) построим матрицы A_2 и T_2. Так как в матрице A_1 $\max \left| a_{ij}^{(1)} \right| = \left| a_{13} \right| = \sqrt{2}$, то $i = 1$, $j = 3$ и

$$T_{ij} = T_{13} = \begin{pmatrix} \cos\alpha & 0 & -\sin\alpha \\ 0 & 1 & 0 \\ \sin\alpha & 0 & \cos\alpha \end{pmatrix}.$$

Угол 2α определяется по формуле

$$\operatorname{tg} 2\alpha = \frac{2a_{ij}^{(1)}}{a_{ii}^{(1)} - a_{jj}} = \frac{-\dfrac{4}{\sqrt{2}}}{4-5} = \frac{4}{\sqrt{2}} = 2\sqrt{2}.$$

Поэтому

$$\cos 2\alpha = \frac{1}{\sqrt{1+\operatorname{tg}^2 2\alpha}} = \frac{1}{\sqrt{1+8}} = \frac{1}{3},$$

$$\cos\alpha = \sqrt{\frac{1+\cos 2\alpha}{2}} = \sqrt{\frac{1+\frac{1}{3}}{2}} = \sqrt{\frac{2}{3}},$$

$$\sin\alpha = \sqrt{\frac{1-\cos 2\alpha}{2}} = \sqrt{\frac{1-\frac{1}{3}}{2}} = \sqrt{\frac{1}{3}}.$$

Поскольку $a_{13}^{(1)} \cdot (a_{11}^{(1)} - a_{33}^{(1)}) = -\sqrt{2}\,(4-5) > 0$, то по формуле (5) для $\sin\alpha$ выбран знак “$+$” и матрица

$$T_{13} = \begin{pmatrix} \sqrt{\frac{2}{3}} & 0 & -\frac{1}{\sqrt{3}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{3}} & 0 & \sqrt{\frac{2}{3}} \end{pmatrix}.$$

Теперь по первой формуле из (2) вычислим матрицу

$$A_2 = T_{13}^{-1} A_1 T_{13} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 6 \end{pmatrix} = \Lambda$$

и по второй формуле из (2) — матрицу

$$T = T_2 = T_{12} T_{13} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \sqrt{\frac{2}{3}} \end{pmatrix}.$$

По найденным матрицам Λ и T выписываем собственные значения $\lambda_1 = 3$, $\lambda_2 = -2$, $\lambda_3 = 6$ и соответствующие им собственные векторы матрицы A

$$e_1' = \frac{1}{\sqrt{3}}(1,-1,1)^{\mathrm{T}}, \quad e_2' = \frac{1}{\sqrt{2}}(1,1,0)^{\mathrm{T}}, \quad e_3' = \frac{1}{\sqrt{6}}(-1,1,2)^{\mathrm{T}}.$$

В рассмотренных примерах процесс закончился соответственно на втором и первом шагах. В общем же случае такой процесс бесконечен, так как на каждом шаге вместо нулевых недиагональных элементов могут возникать ненулевые. Однако при выполнении каждого шага уменьшается с большой скоростью сумма квадратов недиагональных элементов. Процесс прекращают, когда все недиагональные элементы становятся пренебрежимо малыми по абсолютной величине.

Метод вращений хорошо реализуется средствами компьютерных технологий, причем существует большое количество вычислительных схем для его реализации. При любой из них метод вращений представляет собой исключительно компактную процедуру, которая обеспечивает высокую точность вычислений. Точность нахождения собственных значений и собственных векторов этим методом оказывается сравнимой с длиной машинного слова. Стандартные программы к методу вращений содержатся во многих пакетах математического обеспечения.

QR - алгоритм

В настоящее время наиболее надежным и употребительным методом решения полной проблемы собственных значений, т. е. проблемы отыскания всех характеристических чисел произвольной действительной или комплексной квадратной матрицы A n-го порядка, является QR-алгоритм, предложенный в начале 1960-х гг. независимо и почти одновременно В.Н. Кублановской и Дж. Френсисом.

При выполнении QR-алгоритма строят последовательность подобных матриц

$$A_0 = A,\ A_1,\ A_2, \ldots, A_k, \ldots \tag{1}$$

по следующему правилу:

Матрицу A_{k-1}, $k = 1, 2, \ldots,$ представляют ее QR-разложением

$$A_{k-1} = Q_k R_k \tag{2}$$

и полагают

$$A_k = Q_k^{-1} A_{k-1} Q_k = Q_k^{-1} (Q_k R_k) Q_k = R_k Q_k .$$

Фактически матрицы A_k получаются из матрицы A_{k-1}, представленной ее QR-разложением (2), путем перестановки множителей в этом разложении.

Таким образом, построение последовательности (1) подобных матриц реализуется по следующим рекуррентным формулам:

$$\begin{aligned}
&A_0 = Q_1 R_1, \quad A_1 = Q_1^{-1} A_0 Q_1 = Q_1^{-1} (Q_1 R_1) Q_1 = R_1 Q_1, \\
&A_1 = Q_2 R_2, \quad A_2 = Q_2^{-1} A_1 Q_2 = Q_2^{-1} (Q_2 R_2) Q_2 = R_2 Q_2, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\
&A_{k-1} = Q_k R_k,\ A_k = Q_k^{-1} A_{k-1} Q_k = Q_k^{-1} (Q_k R_k) Q_k = R_k Q_k, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.
\end{aligned} \tag{3}$$

При этом матрица A_k при любом k ортогонально (унитарно) подобна матрице A, так как

$$A_k = Q_k^{-1} A_{k-1} Q_k = Q_k^{-1} Q_{k-1}^{-1} A_{k-2} Q_{k-1} Q_k = \ldots =$$

$$= Q_k^{-1} Q_{k-1}^{-1} \ldots Q_1^{-1} A Q_1 \ldots Q_{k-1} Q_k = \qquad (4)$$

$$= (Q_1 \ldots Q_{k-1} Q_k)^{-1} A (Q_1 \ldots Q_{k-1} Q_k) = P_k^{-1} A P_k,$$

где $P_k = Q_1 Q_2 \ldots Q_{k-1} Q_k$ — ортогональная (унитарная) матрица.

При достаточно общих предположениях последовательность подобных матриц (1) сходится к треугольной матрице. Диагональные элементы такой треугольной матрицы являются ее характеристическими числами и характеристическими числами всех подобных матриц A_k, $k = 0, 1, 2, \ldots$, в частности матрицы $A_0 = A$.

На практике процесс итераций по формулам (3) останавливают на некотором k-м шаге, когда все поддиагональные элементы матрицы A_k становятся по абсолютной величине меньше заданного ε, и за приближенные значения характеристических чисел матрицы A принимают диагональные элементы матрицы A_k.

Важным приемом ускорения QR-алгоритма является использование сдвигов ν_k, $k = 1, 2, \ldots$, для повышения скорости убывания по абсолютной величине поддиагональных элементов матрицы A_{k-1}. С этой целью на k-м шаге строят QR-разложение не для матрицы A_{k-1}, а для матрицы $A_{k-1} - \nu_k E$. Если построено QR-разложение

$$A_{k-1} - \nu_k E = Q_k R_k$$

матрицы $A_{k-1} - \nu_k E$, то полагают

$$A_k = Q_k^{-1} A_{k-1} Q_k = Q_k^{-1} (Q_k R_k + \nu_k E) Q_k = R_k Q_k + \nu_k E.$$

При этом снова матрица A_k ортогонально (унитарно) подобна матрице A_{k-1}, так как

$$Q_k^{-1} A_{k-1} Q_k = Q_k^{-1} (Q_k R_k + \nu_k E) Q_k = R_k Q_k + \nu_k E = A_k.$$

В силу этого для матрицы A_k снова выполняется цепочка соотношений (4), из которой следует ортогональное (унитарное) подобие матрицы A_k матрице A.

На практике обычно в качестве сдвига ν_k принимают элемент матрицы A_{k-1}, стоящий в ее нижнем правом углу, или собственное значение подматрицы второго порядка

$$\begin{pmatrix} a_{n-1,n-1}^{(k-1)} & a_{n-1,n}^{(k-1)} \\ a_{n,n-1}^{(k-1)} & a_{n,n}^{(k-1)} \end{pmatrix},$$

которое ближе по величине к элементу $a_{n,n}^{(k-1)}$.

Наличие комплексно сопряженных пар собственных чисел у данной вещественной матрицы A не является, вообще говоря, препятствием для применения QR-алгоритма; просто в этом случае предельной матрицей для последовательности $(\mathbf{A}k)$ будет матрица квазитреугольного, иначе, блочно-треугольного, вида. Каждой комплексной паре собственных чисел в такой матрице будет соответствовать диагональный 2×2-блок, причем сходимость здесь наблюдается по форме матрицы, а не поэлементно (т.е. элементы внутри этих блоков могут изменяться без видимой зависимости от k при сохранении неизменными их собственных чисел).

Теоретическим фундаментом для QR-алгоритма служит следующая теорема о вещественном *разложении Шура*.

Теорема *Для любой вещественной $n\times n$-матрицы* **A** *найдется такая вещественная ортогональная $n\times n$-матрица* **Q**, *что*

$$\mathbf{Q}^T\mathbf{A}\mathbf{Q} = \begin{pmatrix} R_{11} & R_{12} & \dots & R_{1m} \\ 0 & R_{22} & \dots & R_{2m} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & R_{mm} \end{pmatrix},$$

где R_{ii} — либо вещественное число, либо 2×2-матрица с комплексно-сопряженными собственными значениями.

Рассмотренный выше **QR**-алгоритм в разных вариациях настроен на нахождение всех собственных чисел произвольных вещественных $n \times n$-матриц и, в отличие от метода вращений Якоби, ориентированного на симметричные матрицы, мало приспособлен для вычисления собственных векторов. Хотя в рамках этого метода, разумеется, можно путем значительных вычислительных затрат, связанных с преобразованиями самих матриц ортогональных преобразований, находить и собственные векторы, но чаще здесь прибегают к использованию обратного степенного метода, тем более что задача нахождения *всех* собственных векторов — сравнительно редкая.

Пример. С помощью QR-алгоритма найти характеристические числа матрицы

$$A = \begin{pmatrix} 7 & 1 & 1 \\ 0 & 3 & -1 \\ 0 & -3 & 5 \end{pmatrix}.$$

Решение. Чтобы показать, как проводятся шаги QR-алгоритма без сдвига и со сдвигом, начнем с того, что первый шаг проведем без сдвига. Поэтому будем строить QR-разложение матрицы $A_0 = A$. Сначала эту матрицу вращениями приведем к треугольному виду. Для этого следует обнулить лишь элемент $a_{32} = -3$. С этой целью составим матрицу вращения

$$T_{32} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}$$

и найдем $\cos\alpha$ и $\sin\alpha$ из равенства нулю элемента матрицы $T_{32} A_0$ в позиции $(3, 2)$, т. е. из равенства

$$3\sin\alpha - 3\cos\alpha = 0.$$

Отсюда получаем $\operatorname{tg}\alpha = 1$. Следовательно,

$$\cos\alpha = \frac{1}{\sqrt{1+\operatorname{tg}^2\alpha}} = \frac{1}{\sqrt{2}}, \qquad \sin\alpha = \frac{\operatorname{tg}\alpha}{\sqrt{1+\operatorname{tg}^2\alpha}} = \frac{1}{\sqrt{2}}$$

и

$$T_{32} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Теперь вычислим произведение

$$T_{32}A_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 7 & 1 & 1 \\ 0 & 3 & -1 \\ 0 & -3 & 5 \end{pmatrix} = \begin{pmatrix} 7 & 1 & 1 \\ 0 & \frac{6}{\sqrt{2}} & -\frac{6}{\sqrt{2}} \\ 0 & 0 & \frac{4}{\sqrt{2}} \end{pmatrix} = R_0 .$$

Из этого равенства получаем QR-разложение

$$A_0 = T_{32}^{-1} R_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \cdot \begin{pmatrix} 7 & 1 & 1 \\ 0 & \frac{6}{\sqrt{2}} & -\frac{6}{\sqrt{2}} \\ 0 & 0 & \frac{4}{\sqrt{2}} \end{pmatrix} = Q_0 R_0 .$$

В заключение цикла построим матрицу

$$A_1 = R_0 Q_0 = \begin{pmatrix} 7 & 0 & \sqrt{2} \\ 0 & 6 & 0 \\ 0 & -2 & 2 \end{pmatrix} .$$

Следующий шаг проведем со сдвигом $\nu = a_{32}^{(1)} = 2$. Поэтому будем строить QR-разложение матрицы

$$A_1 - \nu E = A_1 - 2E = \begin{pmatrix} 5 & 0 & \sqrt{2} \\ 0 & 4 & 0 \\ 0 & -2 & 0 \end{pmatrix} .$$

Как и в предыдущем случае, сначала эту матрицу приведем к треугольному виду. С этой целью также следует обнулить лишь элемент $a_{32}^{(1)} = -2$ матрицы $A_1 - 2E$. Для этого составим матрицу вращения

$$T_{32} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}$$

и найдем $\cos\alpha$ и $\sin\alpha$ из равенства нулю элемента матрицы $T_{32}(A_1 - 2E)$ в позиции $(3, 2)$, т. е. из равенства

$$4\sin\alpha - 2\cos\alpha = 0.$$

Отсюда получаем $\operatorname{tg}\alpha = \frac{1}{2}$. Поэтому

$$\cos\alpha = \frac{1}{\sqrt{1+\operatorname{tg}^2\alpha}} = \frac{2}{\sqrt{5}}, \qquad \sin\alpha = \frac{\operatorname{tg}\alpha}{\sqrt{1+\operatorname{tg}^2\alpha}} = \frac{1}{\sqrt{5}}$$

и

$$T_{32} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ 0 & \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix}.$$

Теперь вычислим произведение

$$T_{32}(A_1 - 2E) = \begin{pmatrix} 5 & 0 & \sqrt{2} \\ 0 & 2\sqrt{5} & 0 \\ 0 & 0 & 0 \end{pmatrix} = R_1.$$

Из этого равенства получаем QR-разложение

$$A_1 - 2E = T_{32}^{-1} R_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ 0 & -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} 5 & 0 & \sqrt{2} \\ 0 & 2\sqrt{5} & 0 \\ 0 & 0 & 0 \end{pmatrix} = Q_1 R_1 .$$

В заключение цикла построим матрицу

$$A_2 = R_1 Q_1 + 2E = \begin{pmatrix} 5 & -\frac{\sqrt{2}}{\sqrt{5}} & \frac{2\sqrt{2}}{\sqrt{5}} \\ 0 & 4 & 2 \\ 0 & 0 & 0 \end{pmatrix} + 2E = \begin{pmatrix} 7 & -\frac{\sqrt{2}}{\sqrt{5}} & \frac{2\sqrt{2}}{\sqrt{5}} \\ 0 & 6 & 2 \\ 0 & 0 & 2 \end{pmatrix}.$$

Матрица A_2 уже треугольная с характеристическими числами $\lambda_1 = 7$, $\lambda_2 = 6$, $\lambda_3 = 2$. Эти же значения являются характеристическими числами матрицы A.

Степенной метод

При вычислении наибольшего по абсолютной величине собственного значения λ_1 матрицы A и соответствующего ему собственного вектора $X_1 = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$ можно пользоваться итерационным методом, называемым степенным. Он состоит в следующем. В качестве нулевого приближения к искомому вектору X_1 берут произвольный вектор

$$X_1^{(0)} = (x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)})^{\mathrm{T}}$$

и строят последовательность приближений

$$X_1^{(1)} = A X_1^{(0)},\ X_1^{(2)} = A^2 X_1^{(0)}, \ldots,\ X_1^{(k)} = A^k X_1^{(0)}, \ldots \qquad (1)$$

Если эта последовательность векторов сходится при $k \to \infty$ к вектору X_1, то он будет искомым собственным вектором, а для соответствующих координат векторов последовательности (1) будут выполняться соотношения

$$\lim_{k \to \infty} \frac{x_i^{(k)}}{x_i^{(k-1)}} = \lambda_1, \quad i = 1, 2, \ldots, n. \qquad (2)$$

В качестве вектора X_1 приближенно можно принять вектор $X_1^{(k)} = A^k X_1^{(0)}$ при достаточно большом значении k, а за собственное значение λ_1 — любое из соотношений (3),

$$\frac{x_i^{(k)}}{x_i^{(k-1)}}, \quad i = 1, 2, \ldots, n$$

или среднее арифметическое таких отношений, т. е. положить

$$\lambda_1 \approx \lambda_1^{(k)} = \frac{1}{n} \sum_{i=1}^{n} \frac{x_i^{(k)}}{x_i^{(k-1)}}. \qquad (4)$$

Процесс приостанавливают, когда $(k-1)$-е и k-е приближения к λ_1 достаточно мало отличаются одно от другого.

Если процесс окажется несходящимся, последнее легко заметить по сильно “прыгающим” значениям отношений (3) при изменении k, то следует изменить начальный вектор $X_1^{(0)}$ и процесс повторить.

Чтобы избежать чрезмерного роста по абсолютной величине координат векторов $A^k X_1^{(0)}$, $k = 1, 2, \ldots$, иногда целесообразно все их координаты умножать на какие-либо числа α_k, меньшие единицы. Например, удобным оказывается деление всех координат вектора $A^k X_1^{(0)}$ на модуль его первой координаты или нормирование этих векторов. При этом вместо последовательности векторов (1) получают последовательность векторов

$$Y_1^{(1)} = \alpha_1 A X_1^{(0)},\ Y_1^{(2)} = \alpha_2 A^2 X_1^{(0)}, \ldots,\ Y_1^{(k)} = \alpha_k A^k X_1^{(0)}, \ldots .$$

Для получения приближений к собственному значению λ_1 в этом случае следует брать отношения соответствующих координат векторов $A Y_1^{(k)}$ и $Y_1^{(k)}$, а за вектор X_1 — вектор $Y_1^{(k)}$ или $A Y_1^{(k)}$. Более подробные рекомендации об этом можно найти в [3] и [34]. Там же можно познакомиться с различными модификациями степенного метода и с обратным степенным методом.

Приведем обоснование рассматриваемого метода.

Пусть действительная квадратная матрица A n-го порядка имеет полную систему линейно независимых собственных векторов $e'_1, e'_2, \ldots, e'_n$, соответствующих собственным значениям $\lambda_1, \lambda_2, \ldots, \lambda_n$. Таким образом, выполняются соотношения

$$A e'_i = \lambda_i e'_i, \quad i = 1, 2, \ldots, n. \tag{5}$$

Будем предполагать, что выполняются также неравенства

$$\left|\lambda_1\right| > \left|\lambda_2\right| \geq \left|\lambda_3\right| \geq \ldots \geq \left|\lambda_n\right|. \tag{6}$$

В пространстве E_n с базисом $e_1, e_2, \ldots, e_n$ возьмем произвольный вектор $X_1^{(0)}$ и разложим его по базису $e'_1, e'_2, \ldots, e'_n$ собственных векторов матрицы A. Тогда получим

$$X_1^{(0)} = c_1 e'_1 + c_2 e'_2 + \ldots + c_n e'_n = \sum_{j=1}^{n} c_j e'_j . \tag{7}$$

Построим последовательность векторов $X_1^{(1)}$, $X_1^{(2)}, \ldots,$ $X_1^{(k)}$, ... с помощью итерационной процедуры

$$X_1^{(k)} = A X_1^{(k-1)} = A^k X_1^{(0)}, \quad k = 1, 2, \ldots$$

Учитывая соотношения (1), (5) и (7), получим

$$X_1^{(k)} = A^k X_1^{(0)} = \sum_{j=1}^{n} c_j \lambda_j^k e'_j, \quad k = 1, 2, \ldots \tag{8}$$

Пусть в базисе $e_1, e_2, \ldots, e_n$ пространства E_n вектор $X_1^{(k)}$ имеет столбец координат $(x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)})^{\mathrm{T}}$, а собственный вектор e'_j — столбец координат $(\alpha_{1j}, \alpha_{2j}, \ldots, \alpha_{nj})^{\mathrm{T}}$, т. е. вектор

$$e'_j = \sum_{i=1}^{n} \alpha_{ij} e_i .$$

Подставляя это выражение для e'_j в равенство (8), получим

$$X_1^{(k)} = \sum_{j=1}^{n} c_j \lambda_j^k \sum_{i=1}^{n} \alpha_{ij} e_i = \sum_{i=1}^{n} e_i \sum_{j=1}^{n} c_j \alpha_{ij} \lambda_j^k, \quad k = 1, 2, \ldots .$$

Отсюда видно, что i-й координатой вектора $X_1^{(k)}$ является

$$x_i^{(k)} = \sum_{j=1}^{n} c_j \alpha_{ij} \lambda_j^k . \tag{9}$$

Аналогично получим

$$x_i^{(k+1)} = \sum_{j=1}^{n} c_j \alpha_{ij} \lambda_j^{k+1} . \tag{10}$$

Разделив равенство (10) на равенство (9), придем к соотношению

$$\frac{x_i^{(k+1)}}{x_i^{(k)}} = \frac{c_1 \alpha_{i1} \lambda_1^{k+1} + \ldots + c_n \alpha_{in} \lambda_n^{k+1}}{c_1 \alpha_{i1} \lambda_1^k + \ldots + c_n \alpha_{in} \lambda_n^k} . \tag{11}$$

Предположим, что $c_1 \neq 0$ и $\alpha_{i1} \neq 0$. Этого можно добиться, выбирая надлежащим образом исходный вектор $X_1^{(0)}$ и базис $e_1, e_2, \ldots, e_n$ пространства E_n. Равенство (11) преобразуем к виду

$$\frac{x_i^{(k+1)}}{x_i^{(k)}} = \lambda_1 \cdot \frac{1 + \dfrac{c_2 \alpha_{i2}}{c_1 \alpha_{i1}} \left(\dfrac{\lambda_2}{\lambda_1} \right)^{k+1} + \ldots + \dfrac{c_n \alpha_{in}}{c_1 \alpha_{i1}} \left(\dfrac{\lambda_n}{\lambda_1} \right)^{k+1}}{1 + \dfrac{c_2 \alpha_{i2}}{c_1 \alpha_{i1}} \left(\dfrac{\lambda_2}{\lambda_1} \right)^{k} + \ldots + \dfrac{c_n \alpha_{in}}{c_1 \alpha_{i1}} \left(\dfrac{\lambda_n}{\lambda_1} \right)^{k}} . \quad (12)$$

Отсюда следует, что при $|\lambda_1| > |\lambda_2|$

$$\lim_{k \to \infty} \frac{x_i^{(k+1)}}{x_i^{(k)}} = \lambda_1 , \quad i = 1, 2, \ldots, n ,$$

или приближенно

$$\lambda_1 \approx \frac{x_i^{(k+1)}}{x_i^{(k)}} , \quad i = 1, 2, \ldots, n .$$

За приближенное значение λ_1 можно брать найденное по этой формуле значение при любом $i = 1, 2, \ldots, n$ или среднее арифметическое этих значений.

Вектор $X_1^{(k)} = A^k X_1^{(0)}$ приближенно представляет собой собственный вектор e'_1 матрицы A, соответствующий собственному значению λ_1. Действительно, из равенства (8) получаем

$$X_1^{(k)} = A^k X_1^{(0)} = c_1 \lambda_1^k e'_1 + \sum_{j=2}^{n} c_j \lambda_j^k e'_j =$$

$$= c_1 \lambda_1^k \left[e'_1 + \sum_{j=2}^{n} \frac{c_j}{c_1} \left(\frac{\lambda_j}{\lambda_1} \right)^k e'_j \right]. \quad (13)$$

Поскольку $\left(\frac{\lambda_j}{\lambda_1}\right)^k \to 0$ при $k \to \infty$, то при достаточно большом значении k

$$X_1^{(k)} = A^k X_1^{(0)} \approx c_1 \lambda_1^k e'_1,$$

т. е. вектор $X_1^{(k)} = A^k X_1^{(0)}$ пропорционален собственному вектору e'_1, принадлежащему собственному значению λ_1, и, следовательно, также является собственным вектором, принадлежащим собственному значению λ_1.

В случае, когда наибольшее по модулю собственное значение матрицы A является s-кратным, обоснование рассматриваемого метода проводится аналогично, с незначительными уточнениями в преобразованиях правой части равенства (11) при переходе к равенству, соответствующему равенству (12). При этом результат сохраняется прежний. То же относится и к равенствам (8) и (13).

Пример 1. Степенным методом найти наибольшее по абсолютной величине собственное значение λ_1 и соответствующий ему собственный вектор $X_1 = (x_1, x_2, x_3)^{\mathrm{T}}$ матрицы

$$A = \begin{pmatrix} 13 & 16 & 16 \\ -5 & -7 & -6 \\ -6 & -8 & -7 \end{pmatrix}.$$

Решение. За нулевое приближение вектора X_1 примем вектор $X_1^{(0)} = (1, 1, 1)^{\mathrm{T}}$ и построим последовательность векторов $X_1^{(k)} = A^k X_1^{(0)}$ при $k = 1, 2, \ldots, 10$. Причем на каждом шаге вычислим $\lambda_1^{(k)}$ по формуле (4). Результаты приведены в табл. 1.

Как видно, значения $\lambda_1^{(k)}$ колеблются около числа -3, причем $\lambda_1^{(8)}$, $\lambda_1^{(9)}$ и $\lambda_1^{(10)}$ мало отличаются одно от другого. Поэтому за λ_1 можно принять $\lambda_1 \approx \lambda_1^{(9)} \approx -3$, а за собственный вектор X_1 — вектор

$$X_1 \approx X_1^{(9)} = A^9 X_1^{(0)} = (157581;\ -78762;\ -78789)^{\mathrm{T}},$$

или вектор

$$X_1 \approx \frac{1}{157581} A^9 X_1^{(0)} \approx (1; -0,49982; -0,49999)^{\mathrm{T}},$$

или единичный вектор

$$X_1 \approx \frac{A^9 X_1^{(0)}}{\left| A^9 X_1^{(0)} \right|} \approx (0,8165; -0,4081; -0,4083)^{\mathrm{T}}.$$

Таблица 1

$X_1^{(0)}$	$A X_1^{(0)}$	$A^2 X_1^{(0)}$	$A^3 X_1^{(0)}$	$A^4 X_1^{(0)}$	$A^5 X_1^{(0)}$
1 1 1	45 – 18 – 21	– 39 27 21	261 – 120 – 129	– 591 309 297	2013 – 990 – 1005
$\lambda_1^{(k)}$	2	– 1,1222	– 5,7599	– 2,3806	– 3,3313

Продолжение табл. 1

$A^6X_1^{(0)}$	$A^7X_1^{(0)}$	$A^8X_1^{(0)}$	$A^9X_1^{(0)}$	$A^{10}X_1^{(0)}$
– 5751 2895 2877	17589 – 8772 – 8793	– 52383 26217 26193	157581 – 78762 – 78789	– 472263 236193 236133
– 2,8813	– 3,0483	– 2,9819	– 3,0068	– 2,9975

При переходе, например, от вектора $A^8X_1^{(0)}$ к вектору $A^9X_1^{(0)}$, чтобы избежать больших по абсолютной величине координат, вместо вектора $A^8X_1^{(0)}$ можно было взять вектор

$$Y_1^{(8)} = \frac{1}{52383}A^8X_1^{(0)} \approx (-1; 0,50049; 0,50003)^{\mathrm{T}}.$$

Тогда вместо вектора $A^9X_1^{(0)}$ мы имели бы вектор

$$AY_1^{(8)} = (3,0082; -1,50358; -1,50409)^{\mathrm{T}}$$

и по формуле (4) получили бы

$$\lambda_1^{(9)} = \frac{1}{3}\left(\frac{3,00824}{-1} + \frac{-1,50358}{0,50049} + \frac{-1,50409}{0,50003}\right) \approx -3.$$

Степенной метод в комбинации с ***методом исчерпывания*** можно применять к вычислению других собственных значений и собственных векторов матрицы. Для обоснования такого подхода нам потребуются следующие утверждения.

Утверждение 1. *Характеристические многочлены матрицы* A *и ее транспонированной матрицы* A^{T} *совпадают.*

Следствие. *Транспонированная матрица* A^{T} *имеет, с учетом кратностей, те же собственные значения, что и матрица* A.

Утверждение 2. *Если вектор* x' *является собственным вектором матрицы* A^{T}, *принадлежащим собственному значению* λ, *т. е. если* $A^{\mathrm{T}}x' = \lambda x'$, *то выполняется соотношение*

$$(x')^{\mathrm{T}} A = \lambda (x')^{\mathrm{T}}.$$

Утверждение 3. *Собственный вектор* x *матрицы* A, *принадлежащий собственному значению* λ, *ортогонален каждому собственному вектору* x' *матрицы* A^{T}, *принадлежащему собственному значению* $\mu \neq \lambda$.

Теперь перейдем к изложению **метода исчерпывания**. Пусть уже найдено собственное значение λ_1 и собственный вектор x_1 матрицы A, принадлежащий этому собственному значению. Для вычисления второго по абсолютной величине собственного значения λ_2 и соответствующего ему собственного вектора x_2 матрицы A найдем степенным методом или из системы $(A^{\mathrm{T}} - \lambda_1 E)\, u_1 = 0$ собственный вектор u_1 матрицы A^{T}, принадлежащий собственному значению λ_1, и положим

$$x'_1 = \frac{u_1}{(x_1, u_1)}.$$

Так, построенный собственный вектор x_1' матрицы A^{T} удовлетворяет условию $(x_1, x_1') = x_1^{\mathrm{T}} x_1' = 1$, поскольку

$$x_1^{\mathrm{T}} x_1' = \frac{x_1^{\mathrm{T}} u_1}{(x_1, u_1)} = \frac{(x_1, u_1)}{(x_1, u_1)} = 1.$$

Далее построим матрицу

$$A_1 = A - \lambda_1 x_1 (x_1')^{\mathrm{T}}.$$

Все собственные векторы $x_1, x_2, \ldots, x_n$ матрицы A являются собственными векторами матрицы A_1, причем соответствующие собственные значения также сохраняются, за исключением λ_1, которому в матрице A_1 соответствует нулевое собственное значение.

Наибольшее по абсолютной величине собственное значение μ_1 и ему принадлежащий собственный вектор v_1 матрицы A_1 (их опять можно найти степенным методом) являются соответственно вторым по абсолютной величине собственным значением λ_2 и ему принадлежащим собственным вектором x_2 матрицы A. С матрицей A_1 можно поступить аналогично и т. д. В итоге можно найти все собственные значения и собственные векторы матрицы A. Поясним такой подход на примере.

Пример 2. Степенным методом в сочетании с методом исчерпывания найти собственные значения и собственные векторы матрицы

$$A = \begin{pmatrix} 2 & -5 & -3 \\ -1 & -2 & -3 \\ 3 & 15 & 12 \end{pmatrix}.$$

Решение. Сначала степенным методом для матрицы A найдем собственное значение λ_1 и принадлежащий ему собственный вектор $X_1 = (x_1, x_2, x_3)^{\mathrm{T}}$. Для этого за нулевое приближение вектора X_1 примем, например, вектор $X_1^{(0)} = (1, 1, 1)^{\mathrm{T}}$ и построим векторы $X_1^{(k)} = A^k X_1^{(0)}$ при $k = 1, 2, \ldots, 10$. Результаты вычислений приведены в табл. 2.

Таблица 2

$X_1^{(0)}$	$A X_1^{(0)}$	$A^2 X_1^{(0)}$	$A^3 X_1^{(0)}$	$A^4 X_1^{(0)}$	$A^5 X_1^{(0)}$
1	– 6	– 72	– 540	– 3564	– 22356
1	– 6	– 72	– 540	– 3564	– 22356
1	30	252	1728	11016	68040

$A^6 X_1^{(0)}$	$A^7 X_1^{(0)}$	$A^8 X_1^{(0)}$	$A^9 X_1^{(0)}$	$A^{10} X_1^{(0)}$
–137052	–831060	–5012604	–30154356	–181162330
–137052	–831060	–5012604	–30154356	–181162330
414072	2501928	15064056	90541800	543723200

Теперь по формуле (4) при $k = 10$ находим

$$\lambda_1 \approx \frac{1}{3}\left(\frac{-181162330}{-30154356} + \frac{-181162330}{-30154356} + \frac{543723200}{90541800}\right) \approx 6.$$

За вектор X_1 примем вектор

$$X_1 \approx \frac{A^{10} X_1^{(0)}}{181162330} \approx (-1; -1; 3)^{\mathrm{T}}.$$

Чтобы найти собственное значение λ_2 и собственный вектор X_2 матрицы A, сначала из системы $(A^{\mathrm{T}} - \lambda_1 E)\, U = 0$ или

степенным методом найдем собственный вектор $U_1 = (1; 5; 3)^{\mathrm{T}}$ матрицы A^{T}, принадлежащий собственному значению $\lambda_1 = 6$, и положим

$$X_1' = \frac{U_1}{(X_1, U_1)} = \frac{1}{3}(1; 5; 3)^{\mathrm{T}}.$$

Затем сконструируем матрицу

$$A_1 = A - \lambda_1 X_1 (X_1')^{\mathrm{T}} = A - 6 \cdot \begin{pmatrix} -1 \\ -1 \\ 3 \end{pmatrix} \cdot \frac{1}{3}(1; 5; 3) = \begin{pmatrix} 5 & 5 & 3 \\ 1 & 8 & 3 \\ -3 & -15 & -6 \end{pmatrix}.$$

Для матрицы A_1 степенным методом найдем собственное значение μ_1 и ему принадлежащий собственный вектор V_1. Для этого за нулевое приближение собственного вектора V_1 примем, например, вектор $V_1^{(0)} = (1, 1, 1)^{\mathrm{T}}$ и построим векторы $V_1^{(k)} = A^k V_1^{(0)}$ при $k = 1, 2, 3$. Результаты вычислений приведены в табл. 3.

Таблица 3

$V_1^{(0)}$	$A_1 V_1^{(0)}$	$A_1^2 V_1^{(0)}$	$A_1^3 V_1^{(0)}$
1	12	36	108
1	12	36	108
1	– 24	– 72	– 216

Теперь по формуле (4) при $k = 3$ находим

$$\mu_1 \approx \frac{1}{3}\left(\frac{108}{36} + \frac{108}{36} + \frac{-216}{-72}\right) = 3$$

и за вектор V_1 примем вектор

$$V_1 = \frac{A^3 V_1^{(0)}}{108} = (1; 1; -2)^{\mathrm{T}}.$$

Поэтому для матрицы A собственное значение $\lambda_2 = \mu_1 = 3$ и собственный вектор $X_2 = V_1 = (1;1;-2)^{\mathrm{T}}$. Собственное значение $\lambda_3 = 3$ матрицы A найдем из соотношения

$$\lambda_1 + \lambda_2 + \lambda_3 = a_{11} + a_{22} + a_{33},$$

а собственный вектор $X_3 = (2;-1;1)^{\mathrm{T}}$ — из системы

$$(A - \lambda_3 E)X = 0.$$