УДК 519.614 doi 10.26089/NumMet.v17r105

ОРТОГОНАЛЬНО-СТЕПЕННОЙ МЕТОД РЕШЕНИЯ ЧАСТИЧНОЙ ПРОБЛЕМЫ СОБСТВЕННЫХ ЗНАЧЕНИЙ И ВЕКТОРОВ ДЛЯ СИММЕТРИЧНОЙ НЕОТРИЦАТЕЛЬНО ОПРЕДЕЛЕННОЙ МАТРИЦЫ

И. В. Киреев[1]

Предложена и обоснована экономичная версия метода сопряженных направлений для построения нетривиального решения однородной системы линейных алгебраических уравнений с вырожденной симметричной неотрицательно определенной квадратной матрицей. Предложено однопараметрическое семейство одношаговых нелинейных итерационных процессов вычисления собственного вектора, отвечающего наибольшему собственному значению симметричной неотрицательно определенной квадратной матрицы. Это семейство включает в себя степенной метод как частный случай. Доказана сходимость возникающих последовательностей векторов к собственному вектору, ассоциированному с наибольшим характеристическим числом матрицы. Предложена двухшаговая процедура ускорения сходимости итераций этих процессов, в основе которой лежит ортогонализация в подпространстве Крылова. Приведены результаты численных экспериментов.

Ключевые слова: собственный вектор, собственное значение, метод сопряженных направлений, подпространства Крылова.

1. Введение. Частичная проблема собственных значений [1] заключается в вычислении некоторого подмножества собственных значений матрицы и, если требуется, соответствующих им собственных векторов, когда интересны не все собственные значения, а только часть спектра матрицы, например граничная. Обзор методов решения этой задачи, разработанных до 2000 г., можно найти в [2], а в монографии [3] собрана библиография по исследованию этой проблемы с помощью подпространств Крылова [4].

Необходимо отметить, что на результат работы в компьютере любого из рассматриваемых алгоритмов оказывают значительное влияние погрешности округления [5–7] при выполнении арифметических операций. Однако степень этого влияния, как правило, является предметом самостоятельного исследования, поэтому ниже будем предполагать, что погрешности вычислений отсутствуют.

Обозначим через m порядок исследуемой симметричной неотрицательно определенной матрицы A, которую можно рассматривать как линейное преобразование евклидового пространства $\mathbb{R}^m$ вектор-столбцов со скалярным произведением

$$\langle \boldsymbol{a},\boldsymbol{b}\rangle = \sum_{1\leqslant k\leqslant m} \alpha_k\beta_k, \quad \boldsymbol{a}=(\alpha_1,\alpha_1,\dots,\alpha_m)', \quad \boldsymbol{b}=(\beta_1,\beta_2,\dots\beta_m)'\in\mathbb{R}^m;$$

здесь штрих означает транспонирование. В этих обозначениях $\|\boldsymbol{a}\| = \sqrt{\langle \boldsymbol{a},\boldsymbol{a}\rangle}$ есть длина вектора $\boldsymbol{a}$.

Вначале рассмотрим элементарный случай симметричной частичной проблемы, когда требуется построить ненулевой вектор, отвечающий нулевому собственному значению исследуемой матрицы. Символами $\mathbb{K}_A$ и $\mathbb{I}_A$ обозначаем ядро и образ матрицы A: $\mathbb{K}_A = \{\boldsymbol{k}\in\mathbb{R}^m \,|\, A\boldsymbol{k}=\boldsymbol{0}\}$ и $\mathbb{I}_A = \{\boldsymbol{i}\in\mathbb{R}^m \,|\, \exists\, \boldsymbol{j}\in\mathbb{R}^m : A\boldsymbol{j}=\boldsymbol{i}\}$.

В силу симметричности матрицы A пространство $\mathbb{R}^m = \mathbb{K}_A\oplus\mathbb{I}_A$ — прямая сумма и $\mathbb{K}_A\perp\mathbb{I}_A$, т.е. любой вектор $\boldsymbol{a}\in\mathbb{R}^m$ единственным образом представи́м в виде $\boldsymbol{a}=\boldsymbol{a}_{\mathrm{K}}+\boldsymbol{a}_{\mathrm{I}}$, где $\boldsymbol{a}_{\mathrm{K}}\in\mathbb{K}_A$, $\boldsymbol{a}_{\mathrm{I}}\in\mathbb{I}_A$ и $\boldsymbol{a}_{\mathrm{K}}\perp\boldsymbol{a}_{\mathrm{I}}$. Далее через $\boldsymbol{a}_1,\dots,\boldsymbol{a}_m$ обозначаем ортонормированный базис из собственных векторов матрицы A, упорядоченный по возрастанию соответствующих им характеристических чисел:

$$A\boldsymbol{a}_i=\lambda_i\boldsymbol{a}_i, \quad \langle\boldsymbol{a}_i,\boldsymbol{a}_j\rangle=\delta_{ij}, \quad i,j=1,2,\dots,m; \quad 0\leqslant\lambda_1\leqslant\lambda_2\ldots\leqslant\lambda_m; \tag{1}$$

здесь δ_{ij} — символ Кронекера. Поскольку ограничение преобразования A на $\mathbb{I}_A$ невырождено, то существует минимальное $\lambda_{\mathrm{I}}>0$ собственное значение матрицы A на подпространстве $\mathbb{I}_A$:

$$0<\lambda_{\mathrm{I}}\|\boldsymbol{a}_{\mathrm{I}}\|^2\leqslant\langle\boldsymbol{a}_{\mathrm{I}},A\boldsymbol{a}_{\mathrm{I}}\rangle\leqslant\lambda_m\|\boldsymbol{a}_{\mathrm{I}}\|^2 \quad \forall\boldsymbol{a}_{\mathrm{I}}\in\mathbb{I}_A, \quad \boldsymbol{a}_{\mathrm{I}}\neq\boldsymbol{0}. \tag{2}$$

[1] Институт вычислительного моделирования СО РАН, Академгородок, 50/44, 660036, Красноярск; доцент, e-mail: kiv@icm.krasn.ru

Числа λ_I и λ_m во многом определяют скорость сходимости итерационных процессов, связанных с матрицей A. Целью настоящей статьи является построение эффективных алгоритмов вычисления собственных векторов $\boldsymbol{a}_\mathrm{I}$ и $\boldsymbol{a}_m$.

2. Вычисление ненулевого вектора из ядра матрицы методом сопряженных направлений. Отметим следующий очевидный факт: если $\boldsymbol{a}$ — некоторое приближение к элементу из ядра $\mathbb{K}_A$ и в представлении $\boldsymbol{a} = \boldsymbol{a}_\mathrm{K} + \boldsymbol{a}_\mathrm{I}$ слагаемые отличны от нуля, то аппроксимирующий вектор $\boldsymbol{a}$ можно уточнить ортогонализацией его по отношению к $A\boldsymbol{a} \in \mathbb{I}_A$. Это соображение приводит к следующему алгоритму вычисления ненулевого вектора из $\mathbb{K}_A$.

Теорема 1. *Итерационный процесс*

$$\boldsymbol{a}^{n+1} = \boldsymbol{a}^n - \frac{\langle \boldsymbol{a}^n, A\boldsymbol{a}^n\rangle}{\|A\boldsymbol{a}^n\|^2} A\boldsymbol{a}^n, \tag{3}$$

где начальное приближение $\boldsymbol{a}^0 \notin \mathbb{K}_A$ задано и $n \geqslant 0$, сходится для вырожденной симметричной неотрицательно определенной матрицы A: $\lim\limits_{n\to\infty} \boldsymbol{a}^n \in \mathbb{K}_A$. Кроме того, для любого натурального n имеет место оценка $\|\boldsymbol{a}_\mathrm{I}^n\| \leqslant \left(\dfrac{\lambda_m - \lambda_\mathrm{I}}{\lambda_m + \lambda_\mathrm{I}}\right)^n \|\boldsymbol{a}_\mathrm{I}^0\|$, $\boldsymbol{a}_\mathrm{K}^n = \boldsymbol{a}_\mathrm{K}^0$, где $\boldsymbol{a}_\mathrm{I}^n \in \mathbb{I}_A$ и $\boldsymbol{a}_\mathrm{K}^n \in \mathbb{K}_A$ — ортогональные составляющая и проекция вектора $\boldsymbol{a}^n$ относительно подпространства $\mathbb{K}_A$, а λ_I определено в (2).

Доказательство. Поскольку вектор $\boldsymbol{a}^n$ единственным образом представи́м в виде $\boldsymbol{a}^n = \boldsymbol{a}_\mathrm{K}^n + \boldsymbol{a}_\mathrm{I}^n$, $\boldsymbol{a}_\mathrm{K}^n \in \mathbb{K}_A$, $\boldsymbol{a}_\mathrm{I}^n \in \mathbb{I}_A$ и $\boldsymbol{a}_\mathrm{K}^n \perp \boldsymbol{a}_\mathrm{I}^n$, то из (3) следует, что $\boldsymbol{a}_\mathrm{K}^{n+1} = \boldsymbol{a}_\mathrm{K}^n$ и $\boldsymbol{a}_\mathrm{I}^{n+1} = \boldsymbol{a}_\mathrm{I}^n - \dfrac{\langle \boldsymbol{a}_\mathrm{I}^n, A\boldsymbol{a}_\mathrm{I}^n\rangle}{\|A\boldsymbol{a}_\mathrm{I}^n\|^2} A\boldsymbol{a}_\mathrm{I}^n$. Поэтому,

$$\|\boldsymbol{a}_\mathrm{I}^{n+1}\|^2 = \left\|\boldsymbol{a}_\mathrm{I}^n - \frac{\langle \boldsymbol{a}_\mathrm{I}^n, A\boldsymbol{a}_\mathrm{I}^n\rangle}{\|A\boldsymbol{a}_\mathrm{I}^n\|^2} A\boldsymbol{a}_\mathrm{I}^n\right\|^2 = \|\boldsymbol{a}_\mathrm{I}^n\|^2 - \frac{\langle \boldsymbol{a}_\mathrm{I}^n, A\boldsymbol{a}_\mathrm{I}^n\rangle^2}{\|A\boldsymbol{a}_\mathrm{I}^n\|^2} = \|\boldsymbol{a}_\mathrm{I}^n\|^2 \left(1 - \frac{\langle \boldsymbol{a}_\mathrm{I}^n, A\boldsymbol{a}_\mathrm{I}^n\rangle^2}{\|\boldsymbol{a}_\mathrm{I}^n\|^2 \|A\boldsymbol{a}_\mathrm{I}^n\|^2}\right).$$

Однако тогда $\dfrac{\|\boldsymbol{a}_\mathrm{I}^{n+1}\|}{\|\boldsymbol{a}_\mathrm{K}^{n+1}\|} = \alpha_n \dfrac{\|\boldsymbol{a}_\mathrm{I}^n\|}{\|\boldsymbol{a}_\mathrm{K}^n\|}$, где $\alpha_n = \sqrt{1 - \dfrac{\langle \boldsymbol{a}_\mathrm{I}^n, A\boldsymbol{a}_\mathrm{I}^n\rangle^2}{\|\boldsymbol{a}_\mathrm{I}^n\|^2 \|A\boldsymbol{a}_\mathrm{I}^n\|^2}}$, так как предполагается, что $\boldsymbol{a}_\mathrm{I}^n \neq \mathbf{0}$. Оценим сверху величину α_n. В [1] показано, что при введенных ограничениях имеет место неравенство

$$\frac{\langle \boldsymbol{a}_\mathrm{I}^n, A\boldsymbol{a}_\mathrm{I}^n\rangle^2}{\|\boldsymbol{a}_\mathrm{I}^n\|^2 \|A\boldsymbol{a}_\mathrm{I}^n\|^2} \geqslant 4\left(\sqrt{\frac{\lambda_m}{\lambda_\mathrm{I}}} + \sqrt{\frac{\lambda_\mathrm{I}}{\lambda_m}}\right)^{-2}.$$

Следовательно, $\alpha_n^2 \leqslant 1 - 4\left(\sqrt{\dfrac{\lambda_m}{\lambda_\mathrm{I}}} + \sqrt{\dfrac{\lambda_\mathrm{I}}{\lambda_m}}\right)^{-2} = \left(\sqrt{\dfrac{\lambda_m}{\lambda_\mathrm{I}}} + \sqrt{\dfrac{\lambda_\mathrm{I}}{\lambda_m}}\right)^{-2}\left(\sqrt{\dfrac{\lambda_m}{\lambda_\mathrm{I}}} - \sqrt{\dfrac{\lambda_\mathrm{I}}{\lambda_m}}\right)^2 = \left(\dfrac{\lambda_m - \lambda_\mathrm{I}}{\lambda_m + \lambda_\mathrm{I}}\right)^2$.

Таким образом, $\dfrac{\|\boldsymbol{a}_\mathrm{I}^{n+1}\|}{\|\boldsymbol{a}_\mathrm{K}^{n+1}\|} = \alpha_n \dfrac{\|\boldsymbol{a}_\mathrm{I}^n\|}{\|\boldsymbol{a}_\mathrm{K}^n\|} \leqslant \dfrac{\lambda_\mathrm{A} - \lambda_\mathrm{I}}{\lambda_\mathrm{A} + \lambda_\mathrm{I}} \dfrac{\|\boldsymbol{a}_\mathrm{I}^n\|}{\|\boldsymbol{a}_\mathrm{K}^n\|} \leqslant \ldots \leqslant \left(\dfrac{\lambda_\mathrm{A} - \lambda_\mathrm{I}}{\lambda_\mathrm{A} + \lambda_\mathrm{I}}\right)^{n+1} \dfrac{\|\boldsymbol{a}_\mathrm{I}^0\|}{\|\boldsymbol{a}_\mathrm{K}^0\|}$.

Поскольку $\boldsymbol{a}_\mathrm{K}^{n+1} = \boldsymbol{a}_\mathrm{K}^0$, то последнее неравенство дает $\|\boldsymbol{a}_\mathrm{I}^n\| \leqslant \left(\dfrac{\lambda_m - \lambda_\mathrm{I}}{\lambda_m + \lambda_\mathrm{I}}\right)^n \|\boldsymbol{a}_\mathrm{I}^0\|$, что и завершает доказательство.

Следует отметить, что знакоопределенность матрицы A существенна для предложенного метода, поскольку в противном случае процесс, как правило, приводит к изотропному [8] вектору $\boldsymbol{a}$ матрицы A, для которого $\langle A\boldsymbol{a}, \boldsymbol{a}\rangle = 0$, и процесс (3) приходится останавливать. Кроме того, алгоритм (3) может быть осуществлен в любой метрике, в которой матрица A самосопряжена и, как показывает следующее утверждение, при незначительном увеличении вычислительных затрат может быть существенно улучшен.

Теорема 2. *Пусть A — симметричная неотрицательно определенная матрица порядка m, имеющая ненулевое ядро. Тогда итерационный процесс*

$$\nu_n = \frac{\langle \boldsymbol{c}^n, \boldsymbol{a}^{n-1}\rangle}{\langle \boldsymbol{c}^n, \boldsymbol{c}^n\rangle}, \quad \boldsymbol{a}^n = \boldsymbol{a}^{n-1} - \nu_n \boldsymbol{c}^n, \quad \sigma_n = \frac{\langle A\boldsymbol{a}^n, \boldsymbol{c}^n\rangle}{\langle \boldsymbol{c}^n, \boldsymbol{c}^n\rangle}, \quad \boldsymbol{c}^{n+1} = A\boldsymbol{a}^n - \sigma_n \boldsymbol{c}^n, \tag{4}$$

где начальное приближение $\boldsymbol{a}^0 \notin \mathbb{K}_A$ задано, $\boldsymbol{c}^1 = A\boldsymbol{a}^0$ и $n \geqslant 1$, не более чем за m шагов приведет к построению отличного от нуля вектора $\boldsymbol{a}^k$ из ядра матрицы A.

Доказательство. По существу, процесс (4) есть не что иное, как нестандартная форма процедуры ортогонализации степенной последовательности [8], порожденной симметричной матрицей A и вектором $\boldsymbol{a}^0$. Аналогично тому, как в монографии [9] доказана “конечность” метода сопряженных градиентов, докажем индукцией по n, что совокупность векторов $\boldsymbol{c}^1, \boldsymbol{c}^2, \dots, \boldsymbol{c}^n \in \mathbb{I}_A$ образует ортогональную систему.

Действительно, для $n = 1$ утверждение верно. Пусть $\langle \boldsymbol{c}^i, \boldsymbol{c}^j \rangle = 0$ при $1 \leqslant i < j \leqslant n$. Заметим, что если в (4) $\nu_n = 0$, то процесс завершается. Поэтому далее предполагаем, что $\nu_j \neq 0$ для $j \leqslant n$.

Из (4) следует, что $\langle \boldsymbol{c}^{n+1}, \boldsymbol{c}^n \rangle = 0$ и $\boldsymbol{c}^{n+1} = \sigma_{n-1}\boldsymbol{c}^{n-1} - \nu_n A\boldsymbol{c}^n + (1-\sigma_n)\boldsymbol{c}^n$, так как

$$\begin{aligned}\boldsymbol{c}^{n+1} &= A\boldsymbol{a}^n - \sigma_n \boldsymbol{c}^n = A\big(\boldsymbol{a}^{n-1} - \nu_n \boldsymbol{c}^n\big) - \sigma_n \boldsymbol{c}^n = A\boldsymbol{a}^{n-1} - \nu_n A\boldsymbol{c}^n - \sigma_n \boldsymbol{c}^n = \\ &= \boldsymbol{c}^n + \sigma_{n-1}\boldsymbol{c}^{n-1} - \nu_n A\boldsymbol{c}^n - \sigma_n\boldsymbol{c}^n = \sigma_{n-1}\boldsymbol{c}^{n-1} - \nu_n A\boldsymbol{c}^n + (1-\sigma_n)\boldsymbol{c}^n.\end{aligned}$$

Следовательно, $A\boldsymbol{c}^n = \dfrac{\sigma_{n-1}}{\nu_n}\,\boldsymbol{c}^{n-1} + \dfrac{1-\sigma_n}{\nu_n}\,\boldsymbol{c}^n - \dfrac{1}{\nu_n}\,\boldsymbol{c}^{n+1}$ и

$$\begin{aligned}\big\langle \boldsymbol{c}^{n+1}, \boldsymbol{c}^j \big\rangle &= \sigma_{n-1}\big\langle \boldsymbol{c}^{n-1}, \boldsymbol{c}^j \big\rangle - \nu_n \big\langle \boldsymbol{a}\boldsymbol{c}^n, \boldsymbol{c}^j \big\rangle + (1-\sigma_n)\big\langle \boldsymbol{c}^n, \boldsymbol{c}^j \big\rangle = \\ &= \sigma_{n-1}\big\langle \boldsymbol{c}^{n-1}, \boldsymbol{c}^j \big\rangle - \nu_n \big\langle \boldsymbol{c}^n, \boldsymbol{a}\boldsymbol{c}^j \big\rangle + (1-\sigma_n)\big\langle \boldsymbol{c}^n, \boldsymbol{c}^j \big\rangle = \\ &= \sigma_{n-1}\big\langle \boldsymbol{c}^{n-1}, \boldsymbol{c}^j \big\rangle - \frac{\nu_n}{\nu_j}\Big\langle \boldsymbol{c}^n, \sigma_{j-1}\boldsymbol{c}^{j-1} + (1-\sigma_j)\boldsymbol{c}^j - \boldsymbol{c}^{j+1} \Big\rangle + (1-\sigma_n)\big\langle \boldsymbol{c}^n, \boldsymbol{c}^j \big\rangle = \\ &= \frac{\nu_n}{\nu_j}\,\big\langle \boldsymbol{c}^n, \boldsymbol{c}^{j+1} \big\rangle + \sigma_{n-1}\big\langle \boldsymbol{c}^{n-1}, \boldsymbol{c}^j \big\rangle + \left[(1-\sigma_n) - \frac{\nu_n}{\nu_j}\,(1-\sigma_j)\right]\big\langle \boldsymbol{c}^n, \boldsymbol{c}^j \big\rangle - \frac{\nu_n}{\nu_j}\,\sigma_{j-1}\big\langle \boldsymbol{c}^n, \boldsymbol{c}^{j-1} \big\rangle,\end{aligned}$$

откуда в силу предположения индукции вытекает, что $\langle \boldsymbol{c}^{n+1}, \boldsymbol{c}^j \rangle = 0$ для $j < n-1$. Осталось доказать ортогональность векторов $\boldsymbol{c}^{n+1}$ и $\boldsymbol{c}^{n-1}$. Действительно,

$$\begin{aligned}\big\langle \boldsymbol{c}^{n+1}, \boldsymbol{c}^{n-1} \big\rangle &= \big\langle A\boldsymbol{a}^n - \sigma_n\boldsymbol{c}^n, \boldsymbol{c}^{n-1} \big\rangle = \big\langle A\boldsymbol{a}^n, \boldsymbol{c}^{n-1} \big\rangle = \big\langle \boldsymbol{a}^n, A\boldsymbol{c}^{n-1} \big\rangle = \\ &= \frac{1}{\nu_{n-1}}\,\big\langle \boldsymbol{a}^n, \sigma_{n-2}\boldsymbol{c}^{n-2} + (1-\sigma_{n-1})\boldsymbol{c}^{n-1} - \boldsymbol{c}^n \big\rangle = \frac{1}{\nu_{n-1}}\,\big\langle \boldsymbol{a}^n, \sigma_{n-2}\boldsymbol{c}^{n-2} + (1-\sigma_{n-1})\boldsymbol{c}^{n-1} \big\rangle = \\ &= \frac{1}{\nu_{n-1}}\,\big\langle \boldsymbol{a}^{n-1} - \nu_n\boldsymbol{c}^n, \sigma_{n-2}\boldsymbol{c}^{n-2} + (1-\sigma_{n-1})\boldsymbol{c}^{n-1} \big\rangle = \\ &= \frac{1}{\nu_{n-1}}\,\big\langle \boldsymbol{a}^{n-1}, \sigma_{n-2}\boldsymbol{c}^{n-2} + (1-\sigma_{n-1})\boldsymbol{c}^{n-1} \big\rangle = \frac{1}{\nu_{n-1}}\,\big\langle \boldsymbol{a}^{n-1}, \sigma_{n-2}\boldsymbol{c}^{n-2} \big\rangle = \\ &= \frac{\sigma_{n-2}}{\nu_{n-1}}\,\big\langle \boldsymbol{a}^{n-1}, \boldsymbol{c}^{n-2} \big\rangle = \frac{\sigma_{n-2}}{\nu_{n-1}}\,\big\langle \boldsymbol{a}^{n-2} - \nu_{n-1}\boldsymbol{c}^{n-1}, \boldsymbol{c}^{n-2} \big\rangle = \frac{\sigma_{n-2}}{\nu_{n-1}}\,\big\langle \boldsymbol{a}^{n-2}, \boldsymbol{c}^{n-2} \big\rangle = 0.\end{aligned}$$

Итак, система векторов $\boldsymbol{c}^1, \boldsymbol{c}^2, \dots, \boldsymbol{c}^n \in \mathbb{I}_A \subset \mathbb{R}^m$ взаимно ортогональна и потому обязательно $n < m$. Обозначим через $\boldsymbol{a}_{\mathrm{I}}$ нормированную ортогональную проекцию вектора $\boldsymbol{a}^0$ на подпространство $\mathbb{K}_A$. Поскольку $\langle \boldsymbol{c}^k, \boldsymbol{a}_{\mathrm{I}} \rangle = 0$ при $k \geqslant 1$, то $\langle \boldsymbol{a}^k, \boldsymbol{a}_{\mathrm{I}} \rangle = \langle \boldsymbol{a}^0, \boldsymbol{a}_{\mathrm{I}} \rangle \neq 0$, и по окончании процесса (4) последний вычисленный вектор $\boldsymbol{a}^n$ будет принадлежать ядру $\mathbb{K}_A$. Что и завершает доказательство.

Если наибольшее собственное значение λ_m известно, то применение алгоритмов (3) и (4) к симметричной неотрицательно определенной матрице $\lambda_m E - \boldsymbol{a}$, где E — единичная матрица порядка m, позволяет вычислить соответствующий собственный вектор.

3. Степенной метод вычисления собственного вектора, соответствующего наибольшему собственному значению. Как правило, само λ_m неизвестно, но в процессе вычисления собственного вектора можно получить и достаточно хорошее приближение к λ_m. Не ограничивая общности, далее будем считать, что характеристические числа матрицы A: $0 < \lambda_1 \leqslant \ldots \leqslant \lambda_{m-1} < \lambda_m$ — упорядочены по возрастанию и последнее собственное значение является простым.

Лемма. *Пусть* $\|\boldsymbol{a}\| = 1$ *и* $\langle \boldsymbol{a}, \boldsymbol{a}_m \rangle \neq 0$, *т.е.* $\varepsilon_a = \sqrt{1 - \langle \boldsymbol{a}, \boldsymbol{a}_m \rangle}^{\,2} < 1$. *Тогда* $\lambda_m \sqrt{1-\varepsilon_a^2} \leqslant \|A\boldsymbol{a}\| < \lambda_m$.

Доказательство. Обозначим через $\alpha_k = \langle \boldsymbol{a}, \boldsymbol{a}_k \rangle$, $k = 1, \dots, m$, координаты вектора A в ортонормированном базисе (1). Следовательно, $\boldsymbol{a} = \sum\limits_{1 \leqslant k \leqslant m} \alpha_k \boldsymbol{a}_k$, $\varepsilon_a^2 = \sum\limits_{1 \leqslant k < m} \alpha_k^2 < 1$, а в силу нормировки вектора $\boldsymbol{a}$ и упорядочивания собственных чисел по возрастанию имеем

$$\lambda_m^2 - \|A\boldsymbol{a}\|^2 = \sum_{1 \leqslant k < m} \big(\lambda_m^2 - \lambda_k^2\big)\alpha_k^2 \leqslant \sum_{1 \leqslant k < m} \big(\lambda_m^2 - \lambda_1^2\big)\alpha_k^2 = \big(\lambda_m^2 - \lambda_1^2\big)\varepsilon_a^2 \leqslant \lambda_m^2 \varepsilon_a^2,$$

т.е. $\|A\boldsymbol{a}\| \geqslant \sqrt{\lambda_m^2 - \lambda_m^2\varepsilon_a^2} = \lambda_m\sqrt{1-\varepsilon_a^2}$. Неравенство $\|A\boldsymbol{a}\| < \lambda_m$ — очевидно. Что и требовалось доказать.

Из этой леммы следует, что

$$0 < \frac{\lambda_m - \|A\boldsymbol{a}\|}{\lambda_m} = 1 - \frac{\|A\boldsymbol{a}\|}{\lambda_m} \leqslant \varepsilon_a^2,$$

т.е. если единичный вектор $\boldsymbol{a}$ отличается по норме на ε_a от собственного вектора, отвечающего наибольшему собственному значению λ_m матрицы A, то число $\|A\boldsymbol{a}\|$ приближает λ_m с относительной погрешностью, не превосходящей ε_a^2.

Для вычисления наибольшего собственного значения λ_m и соответствующего ему собственного вектора $\boldsymbol{a}_m \in \mathbb{R}^m$ часто используют степенной метод [10–12], сходимость одной из версий которого обсуждается в следующем утверждении.

Теорема 3. *Итерационный процесс степенного метода*

$$\lambda_m^n = \|A\boldsymbol{a}^n\|, \quad \boldsymbol{a}^{n+1} = \frac{1}{\lambda_m^n} A\boldsymbol{a}^n, \tag{5}$$

где начальное приближение $\boldsymbol{a}^0$, $\|\boldsymbol{a}^0\| = 1$, $\langle \boldsymbol{a}^0, \boldsymbol{a}_m\rangle \neq 0$ *задано и* $n \geqslant 0$, *сходится при* $n \to \infty : \lambda_m^n \to \lambda_m$ *и* $\boldsymbol{a}^n \to \operatorname{sgn}\langle \boldsymbol{a}^0, \boldsymbol{a}_m\rangle \cdot \boldsymbol{a}_m$. *Кроме того,*

$$\frac{\alpha_k^{n+1}}{\alpha_m^{n+1}} = \frac{\lambda_k}{\lambda_m} \cdot \frac{\alpha_k^n}{\alpha_m^n} \to 0; \quad 0 < \frac{\lambda_m - \lambda_m^n}{\lambda_m} < \varepsilon_n^2 < \left(\frac{\lambda_{m-1}}{\lambda_m}\right)^{2n} \frac{\varepsilon_0^2}{1-\varepsilon_0^2} \to 0.$$

Здесь α_k^n *обозначает* k*-ю координату вектора* $\boldsymbol{a}^n$ *в базисе* (1) *из собственных векторов матрицы* A*:*

$$\boldsymbol{a}^n = \sum_{1\leqslant k\leqslant m} \alpha_k^n \boldsymbol{a}_k; \quad \|\boldsymbol{a}^n\|^2 = \varepsilon_n^2 + (\alpha_m^n)^2 = 1, \quad \varepsilon_n > 0. \tag{6}$$

Доказательство. В ортонормированном базисе $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$ из (1) соотношение (5) эквивалентно равенствам $\lambda_m^n \alpha_k^n = (\lambda_k)^n \alpha_k^0$, $k = 1, \ldots, m$, причем $\alpha_m^n \neq 0$, так как $\alpha_m^0 = \langle \boldsymbol{a}^0, \boldsymbol{a}_m\rangle \neq 0$. Поэтому для $k \in \{1, \ldots, m-1\}$ имеем

$$\frac{\alpha_k^n}{\alpha_m^n} = \frac{\lambda_m^n \cdot \alpha_k^n}{\lambda_m^n \cdot \alpha_m^n} = \frac{(\lambda_k)^n \cdot \alpha_k^0}{(\lambda_m)^n \cdot \alpha_m^0} = \left(\frac{\lambda_k}{\lambda_m}\right)^n \frac{\alpha_k^0}{\alpha_m^0} \underset{n\to\infty}{\to} 0.$$

Поскольку при этом $\|\boldsymbol{a}^n\| = 1$, то процесс (5) сходится, а так как $\dfrac{\alpha_m^n}{\alpha_m^0} = \dfrac{\langle \boldsymbol{a}^n, \boldsymbol{a}_m\rangle}{\langle \boldsymbol{a}^0, \boldsymbol{a}_m\rangle} \geqslant 0$ для любого n, то $\operatorname{sgn}\langle \boldsymbol{a}^n, \boldsymbol{a}_m\rangle = \operatorname{sgn}\langle \boldsymbol{a}^0, \boldsymbol{a}_m\rangle$ и $\boldsymbol{a}^n \to \operatorname{sgn}\langle \boldsymbol{a}^0, \boldsymbol{a}_m\rangle \cdot \boldsymbol{a}_m$.

Оценим относительную погрешность аппроксимации λ_m числом λ_m^n. Так как $0 < \varepsilon_n < 1$, то

$$\frac{\varepsilon_n^2}{1-\varepsilon_n^2} = \frac{\varepsilon_n^2}{(\alpha_m^n)^2} = \sum_{1\leqslant k<m} \left(\frac{\alpha_k^n}{\alpha_m^n}\right)^2 = \sum_{1\leqslant k<m} \left[\left(\frac{\lambda_k}{\lambda_m}\right)^n \frac{\alpha_k^0}{\alpha_m^0}\right]^2 \leqslant \left(\frac{\lambda_{m-1}}{\lambda_m}\right)^{2n} \frac{\varepsilon_0^2}{1-\varepsilon_0^2} < \frac{\varepsilon_0^2}{1-\varepsilon_0^2}.$$

Следовательно, $\varepsilon_n < \varepsilon_0$ и $\varepsilon_n^2 \leqslant \varepsilon_0^2 \left(\dfrac{\lambda_{m-1}}{\lambda_m}\right)^{2n} \dfrac{1-\varepsilon_n^2}{1-\varepsilon_0^2} < \left(\dfrac{\lambda_{m-1}}{\lambda_m}\right)^{2n} \dfrac{\varepsilon_0^2}{1-\varepsilon_n^2}$. Что и требовалось доказать.

Несомненными преимуществами степенного метода (5) являются его уникальная устойчивость к погрешностям вычислений [4] и независимость от априорной информации о распределении собственных значений исследуемой матрицы.

4. Ортогонально-степенной метод построения собственного вектора, соответствующего наибольшему собственному значению. Алгоритм, являющийся комбинацией итерационных процессов из теорем 1 и 3, в какой-то степени наследует положительные качества последних методов и приводит к следующей процедуре вычисления наибольшего собственного значения и соответствующего собственного вектора.

Теорема 4. *Итерационный процесс ортогонально-степенного метода*

$$\lambda_m^n = \mu_n \frac{\langle A\boldsymbol{a}^n, A\boldsymbol{a}^n\rangle}{\langle A\boldsymbol{a}^n, \boldsymbol{a}^n\rangle} + (1-\mu_n)\langle A\boldsymbol{a}^n, \boldsymbol{a}^n\rangle,\ \boldsymbol{b}^n = \lambda_m^n \boldsymbol{a}^n - A\boldsymbol{a}^n;\ \nu_n = \frac{\langle \boldsymbol{b}^n, \boldsymbol{a}^n\rangle}{\|\boldsymbol{b}^n\|},\ \boldsymbol{a}^{n+1} = \frac{\boldsymbol{a}^n - \nu_n \boldsymbol{b}^n}{\|\boldsymbol{a}^n - \nu_n \boldsymbol{b}^n\|}, \tag{7}$$

где $\boldsymbol{a}^0$ *— заданное начальное приближение, такое, что* $\|\boldsymbol{a}^0\| = 1$, $\langle \boldsymbol{a}^0, \boldsymbol{a}_m \rangle \neq 0$, *и* $n \geqslant 0$, *порождает сходящиеся при* $n \to \infty$ *и* $\forall \mu_n \in [\mu_l, \mu_r] \subset \left(0, \dfrac{2\lambda_m}{\lambda_m - \lambda_1}\right)$ *последовательности* $\{\boldsymbol{a}^n\}$ *и* $\{\lambda_m^n\}$*:*

$$\frac{\alpha_k^{n+1}}{\alpha_m^{n+1}} = \frac{\alpha_k^n}{\alpha_m^n} \cdot \frac{\mu_n \lambda_k + (1-\mu_n)\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle}{\mu_n \lambda_m + (1-\mu_n)\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle} \to 0, \quad \lambda_m^n \to \lambda_m; \quad \boldsymbol{a}^n \to \operatorname{sgn}\langle \boldsymbol{a}^0, \boldsymbol{a}_m \rangle \cdot \boldsymbol{a}_m.$$

Здесь α_k^n *—* k*-я координата вектора* $\boldsymbol{a}^n$ *в ортонормированном базисе из собственных векторов матрицы* A*:*

$$\boldsymbol{a}^n = \sum_{1 \leqslant i \leqslant m} \alpha_i^n \boldsymbol{a}_i; \quad \|\boldsymbol{a}^n\|^2 = \varepsilon_n^2 + (\alpha_m^n)^2 = 1; \quad \varepsilon_n > 0, \quad k = 1, \ldots, m-1.$$

При $\mu_l \leqslant \mu_n < 1$ *имеют место неравенства*

$$0 < \frac{\mu_n \lambda_k + (1-\mu_n)\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle}{\mu_n \lambda_m + (1-\mu_n)\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle} \leqslant 1 - \mu_n \frac{\lambda_m - \lambda_k}{\lambda_m} \leqslant 1 - \mu_l \frac{\lambda_m - \lambda_{m-1}}{\lambda_m} < 1.$$

Если $\mu_n = 1$, *то шаг процесса* (7) *совпадает с аналогичным шагом степенного метода и*

$$\left.\frac{\mu_n \lambda_k + (1-\mu_n)\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle}{\mu_n \lambda_m + (1-\mu_n)\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle}\right|_{\mu_n = 1} = \frac{\lambda_k}{\lambda_m} \leqslant \frac{\lambda_{m-1}}{\lambda_m} < 1.$$

Если $1 < \mu_n \leqslant \mu_r$, *то*

$$\left|\frac{\mu_n \lambda_k + (1-\mu_n)\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle}{\mu_n \lambda_m + (1-\mu_n)\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle}\right| \leqslant \max\left\{\frac{\lambda_{m-1}}{\lambda_m}, \left|1 - \mu_r \frac{\lambda_m - \lambda_1}{\lambda_m}\right|, \left|1 - \mu_r \frac{\lambda_m - \lambda_{m-1}}{\lambda_m}\right|\right\} < 1.$$

Доказательство. Если в ортонормированном базисе $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$ из (1) $\boldsymbol{b}^n = \sum\limits_{1 \leqslant k \leqslant m} \beta_k^n a_k$, то

$$\begin{aligned}
\beta_k^n &= (\lambda_m^n - \lambda_k)\alpha_k^n; \quad \|\boldsymbol{b}^n\|^2 = \sum_{1 \leqslant k \leqslant m} (\alpha_k^n)^2 (\lambda_m^n - \lambda_k)^2; \\
\langle \boldsymbol{b}^n, \boldsymbol{a}^n \rangle &= \sum_{1 \leqslant k \leqslant m} \alpha_k^n \beta_k^n = \sum_{1 \leqslant k \leqslant m} (\alpha_k^n)^2 (\lambda_m^n - \lambda_k) \geqslant 0; \\
\nu_n &= \|\boldsymbol{b}^n\|^{-2} \sum_{1 \leqslant k \leqslant m} (\alpha_k^n)^2 (\lambda_m^n - \lambda_k); \quad \alpha_i^{n+1} = \|\boldsymbol{a}^{n+1}\|^{-2} (\alpha_i^n - \gamma_n \beta_i^n).
\end{aligned}$$

Для любого $k = 1, \ldots, m-1$ имеем

$$\frac{\alpha_k^{n+1}}{\alpha_m^{n+1}} = \frac{\alpha_k^n}{\alpha_m^n} \cdot \frac{\mu_n \lambda_k - (\mu_n - 1)\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle}{\mu_n \lambda_m - (\mu_n - 1)\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle}. \tag{8}$$

Поэтому, если $\mu_n \in [\mu_0, 1)$, где $0 < \mu_0 < 1$, то справедливы неравенства

$$0 < \frac{\mu_n \lambda_k + (1-\mu_n)\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle}{\mu_n \lambda_m + (1-\mu_n)\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle} \leqslant 1 - \mu_n \frac{\lambda_m - \lambda_k}{\lambda_m} \leqslant 1 - \mu_l \frac{\lambda_m - \lambda_{m-1}}{\lambda_m},$$

поскольку

$$\frac{\mu_n \lambda_k + (1-\mu_n)\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle}{\mu_n \lambda_m + (1-\mu_n)\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle} = 1 - \frac{\mu_n(\lambda_m - \lambda_k)}{\mu_n \lambda_m + (1-\mu_n)\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle} \leqslant 1 - \frac{\mu_n(\lambda_m - \lambda_k)}{\mu_n \lambda_m + (1-\mu_n)\lambda_m} = 1 - \mu_n \frac{\lambda_m - \lambda_k}{\lambda_m}.$$

В случае $\mu_n = 1$ равенство (8) принимает вид $\dfrac{\alpha_k^{n+1}}{\alpha_m^{n+1}} = \dfrac{\alpha_k^n}{\alpha_m^n} \cdot \dfrac{\lambda_k}{\lambda_m}$, т.е. совпадает с аналогичным соотношением из степенного метода (5).

При $\mu_n \in (1, \mu_r]$ в силу монотонности дробно-линейной функции имеем

$$\left|\frac{\mu_n\lambda_k + (1-\mu_n)\big\langle A\boldsymbol{a}^n, \boldsymbol{a}^n\big\rangle}{\mu_n\lambda_m + (1-\mu_n)\big\langle A\boldsymbol{a}^n, \boldsymbol{a}^n\big\rangle}\right| = \left|\frac{\lambda_k - (1-\mu_n^{-1})\big\langle A\boldsymbol{a}^n, \boldsymbol{a}^n\big\rangle}{\lambda_m - (1-\mu_n^{-1})\big\langle A\boldsymbol{a}^n, \boldsymbol{a}^n\big\rangle}\right| \leqslant \max_{0<x\leqslant(1-\mu_r^{-1})\lambda_m}\left|\frac{\lambda_k - x}{\lambda_m - x}\right| =$$
$$= \max\left\{\frac{\lambda_k}{\lambda_m}, \left|\frac{\lambda_k - \big(1-\mu_r^{-1}\big)\lambda_m}{\lambda_m - \big(1-\mu_r^{-1}\big)\lambda_m}\right|\right\} = \max\left\{\frac{\lambda_k}{\lambda_m}, \left|\frac{\mu_r\lambda_k - (\mu_r - 1)\lambda_m}{\lambda_m}\right|\right\}.$$

Следовательно,

$$\left|\frac{\mu_n\lambda_k + (1-\mu_n)\big\langle A\boldsymbol{a}^n, \boldsymbol{a}^n\big\rangle}{\mu_n\lambda_m + (1-\mu_n)\big\langle A\boldsymbol{a}^n, \boldsymbol{a}^n\big\rangle}\right| \leqslant \max_{0<k<m}\left\{\frac{\lambda_k}{\lambda_m}, \left|1 - \mu_r\,\frac{\lambda_m - \lambda_k}{\lambda_m}\right|\right\} \leqslant$$
$$\leqslant \max\left\{\frac{\lambda_{m-1}}{\lambda_m}, \left|1 - \mu_r\,\frac{\lambda_m - \lambda_1}{\lambda_m}\right|, \left|1 - \mu_r\,\frac{\lambda_m - \lambda_{m-1}}{\lambda_m}\right|\right\} < 1,$$

поскольку в силу условия теоремы выполнены неравенства $0 < \mu_r\,\dfrac{\lambda_m - \lambda_{m-1}}{\lambda_m} \leqslant \mu_r\,\dfrac{\lambda_m - \lambda_1}{\lambda_m} < 2$. Что и завершает доказательство.

Следует отметить, что левая граница интервала $\left(0, \dfrac{2\lambda_m}{\lambda_m - \lambda_1}\right)$ изменения параметра μ_n однозначно определена условием невырождения процесса (7):

$$0 < \big\langle \boldsymbol{b}^n, \boldsymbol{a}^n\big\rangle = \lambda_m^n - \big\langle A\boldsymbol{a}^n, \boldsymbol{a}^n\big\rangle = \mu_n\,\frac{\big\langle A\boldsymbol{a}^n, A\boldsymbol{a}^n\big\rangle}{\big\langle A\boldsymbol{a}^n, \boldsymbol{a}^n\big\rangle} + (1-\mu_n)\big\langle A\boldsymbol{a}^n, \boldsymbol{a}^n\big\rangle - \big\langle A\boldsymbol{a}^n, \boldsymbol{a}^n\big\rangle =$$
$$= \mu_n\,\frac{\big\langle A\boldsymbol{a}^n, A\boldsymbol{a}^n\big\rangle - \big\langle A\boldsymbol{a}^n, \boldsymbol{a}^n\big\rangle^2}{\big\langle A\boldsymbol{a}^n, \boldsymbol{a}^n\big\rangle}.$$

В силу неравенства Коши–Буняковского [8] числитель последней дроби неотрицателен, и мы приходим к условию $\mu_n > 0$.

Правая граница интервала изменения параметра μ_n определяется условием сходимости последовательностей $\left\{\dfrac{\alpha_k^n}{\alpha_m^n}\right\}$, которое будет выполнено, если удовлетворено неравенство

$$\max_{0<k<m,\ \|\boldsymbol{a}^n\|=1}\left|\frac{\mu_n\lambda_k + (1-\mu_n)\big\langle A\boldsymbol{a}^n, \boldsymbol{a}^n\big\rangle}{\mu_n\lambda_m + (1-\mu_n)\big\langle A\boldsymbol{a}^n, \boldsymbol{a}^n\big\rangle}\right| < 1. \tag{9}$$

Поскольку $0 < \lambda_k < \lambda_m$, то при $0 < \mu_n \leqslant 1$ последнее соотношение выполняется автоматически. При $\mu_n > 1$ в силу монотонности дробно-линейной функции неравенство (9) будет выполнено, если

$$\max\left\{\left|1 - \mu_n\,\frac{\lambda_m - \lambda_1}{\lambda_m}\right|, \left|1 - \mu_n\,\frac{\lambda_m - \lambda_{m-1}}{\lambda_m}\right|\right\} < 1. \tag{10}$$

Анализируя три случая взаимного расположения чисел $\dfrac{\lambda_1}{\lambda_m}$, $\dfrac{\lambda_{m-1}}{\lambda_m}$ и $1 - \dfrac{1}{\mu_m}$, приходим к выводу, что соотношение (10), а вместе с ним и неравенство (9), имеет место и при $1 < \mu_n < 2\lambda_m(\lambda_m - \lambda_1)^{-1}$, что и определяет правую границу интервала изменения μ_n.

Из доказательства теоремы 4 следует, что если $\mu_n \underset{n}{\equiv} 1$, то при точных вычислениях и одних и тех же начальных данных итерационный процесс (7) порождает последовательность векторов $\big\{\boldsymbol{a}^n\big\}$, полностью совпадающую с аналогичной последовательностью векторов степенного метода (5). Однако процесс (7) будет сходиться быстрее процесса (5), если положить, например,

$$\mu_n = \frac{\big\langle A\boldsymbol{a}^n, \boldsymbol{a}^n\big\rangle}{\big\langle A\boldsymbol{a}^n, \boldsymbol{a}^n\big\rangle - \lambda_1}\,. \tag{11}$$

В этом случае $\mu_n \geqslant \dfrac{\lambda_m}{\lambda_m - \lambda_1} > 1$ и $\dfrac{\alpha_k^{n+1}}{\alpha_m^{n+1}} = \dfrac{\alpha_k^n}{\alpha_m^n}\cdot\dfrac{\lambda_k - \lambda_1}{\lambda_m - \lambda_1}$, т.е. ортогонально-степенной метод сходится быстрее степенного, для которого $\dfrac{\alpha_k^{n+1}}{\alpha_m^{n+1}} = \dfrac{\alpha_k^n}{\alpha_m^n}\cdot\dfrac{\lambda_k}{\lambda_m}$.

Иной способ ускорения сходимости процесса (7) при μ_n, удовлетворяющим условию теоремы 4, возможен на базе метода сопряженных направлений и аналогичен процессу ортогонализации (4):

1) задано $\boldsymbol{a}^0$ — начальное приближение: $\|\boldsymbol{a}^0\| = 1$, $\langle \boldsymbol{a}^0, \boldsymbol{a}_m \rangle \neq 0$;

2) полагаем $\lambda_m^0 = \mu_0 \dfrac{\langle A\boldsymbol{a}^0, A\boldsymbol{a}^0 \rangle}{\langle A\boldsymbol{a}^0, \boldsymbol{a}^0 \rangle} + (1-\mu_0)\langle A\boldsymbol{a}^0, \boldsymbol{a}^0 \rangle$, $\boldsymbol{b}^0 = \lambda_m^0 \boldsymbol{a}^0 - A\boldsymbol{a}^0$, $\boldsymbol{c}^1 = \boldsymbol{b}^0$;

3) для $n \geqslant 1$ вычисляем:

$$\begin{aligned} \nu_n &= \frac{\langle \boldsymbol{c}^n, \boldsymbol{a}^{n-1} \rangle}{\|\boldsymbol{c}^n\|^2}; \quad \boldsymbol{a}^n = \frac{\boldsymbol{a}^{n-1} - \nu_n \boldsymbol{c}^n}{\|\boldsymbol{a}^{n-1} - \nu_n \boldsymbol{c}^n\|}; \\ \lambda_m^n &= \mu_n \frac{\langle A\boldsymbol{a}^n, A\boldsymbol{a}^n \rangle}{\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle} + (1-\mu_n)\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle; \\ \boldsymbol{b}^n &= \lambda_m^n \boldsymbol{a}^n - A\boldsymbol{a}^n; \quad \boldsymbol{c}^{n+1} = \boldsymbol{b}^n - \frac{\langle \boldsymbol{b}^n, \boldsymbol{c}^n \rangle}{\|\boldsymbol{c}^n\|^2} \boldsymbol{c}^n. \end{aligned} \tag{12}$$

Необходимо отметить, что алгоритм (12) построения собственного вектора, отвечающего наибольшему собственному значению, базируется на степенном методе (5), чем принципиально отличается от методов сопряженных градиентов, максимизирующих отношения Релея [13].

5. Результаты численных экспериментов. Была проведена серия численных экспериментов для матриц с собственными значениями

$$\begin{aligned} \lambda_1 &= \frac{a}{(m+1)^2} + 4\left[\sin\frac{\pi}{2(m+1)}\right]^2 = \frac{a+\pi^2}{(m+1)^2} + O\left(\frac{1}{(m+1)^4}\right), \\ \lambda_k &= \frac{a}{(m+1)^2} + 4\left[\sin\frac{k\pi}{2(m+1)}\right]^2, \quad k = 2, \ldots, m-1; \\ \lambda_m &= \frac{a}{(m+1)^2} + 4\left[\sin\frac{m\pi}{2(m+1)}\right]^2 = 4 + \frac{a-\pi^2}{(m+1)^2} + O\left(\frac{1}{(m+1)^4}\right). \end{aligned}$$

Подобные матрицы возникают при разностной аппроксимации первой краевой задачи для одномерного эллиптического уравнения с постоянным коэффициентом a на равномерной сетке. Тестовые расчеты проводились в системе Matlab с двойной точностью без режима накопления [4]. В этой связи, с целью уменьшения влияния ошибок округления на результаты работы алгоритма, матрица A выбиралась диагональной. Такой выбор гарантирует, что кососимметрическая составляющая погрешности вычисления произведения матрицы на вектор будет пренебрежимо мала.

Результаты вычислений для $m = 100$ и $a = 10.0$ в графической форме приводятся для одного и того же начального приближения $\boldsymbol{a}^0$ (полученного с помощью датчика случайных чисел) к собственному вектору $\boldsymbol{a}_m$: $\langle \boldsymbol{a}^0, \boldsymbol{a}_m \rangle \simeq 0.0557$. На всех графиках по оси абсцисс откладывается относительный номер итерации n/m, а по оси ординат — десятичный логарифм $\log_{10}\|\boldsymbol{a}^n - \boldsymbol{a}_m\|$ евклидовой нормы абсолютной погрешности с линейной интерполяцией значений между узлами.

Работа алгоритмов (3) и (4) для матрицы $\lambda_m E - A$ показана на рис. 1. График с номером 1 соответствует алгоритму (4), а с номером 2 — алгоритму (3). График 3 на рис. 1а получен при решении однородной системы линейных алгебраических уравнений с той же матрицей $\lambda_m E - A$ с помощью “наиболее экономичной” версией метода сопряженных градиентов [7, 12, 14] при том же начальном приближении $\boldsymbol{a}^0$.

Рисунки 2а–4а иллюстрируют работу алгоритмов (7), (12) и классического степенного метода (5) в трех случаях выбора параметра μ_n. На них графики с номером 1 соответствуют результатам вычислений по формулам (12), с номером 2 — ортогонально-степенному методу (7), а с номером 3 — алгоритму (5).

Рисунок 2 иллюстрирует ситуацию (11), когда

$$\mu_n = \frac{\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle}{\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle - \lambda_1}; \quad \lambda_m^n(\mu_n) = \frac{\langle A\boldsymbol{a}^n, A\boldsymbol{a}^n \rangle - \lambda_1 \langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle}{\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle - \lambda_1}.$$

Рисунок 2а — зумминг рис. 2б для первых итераций. Аналогичное взаимное расположение графиков имеет место и на рис. 2б, но скрыто из-за большего масштаба.

Рисунок 3а соответствует выбору

$$\mu_n \underset{n}{\equiv} 1; \quad \lambda_m^n = \frac{\langle A\boldsymbol{a}^n, A\boldsymbol{a}^n \rangle}{\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle}. \tag{13}$$

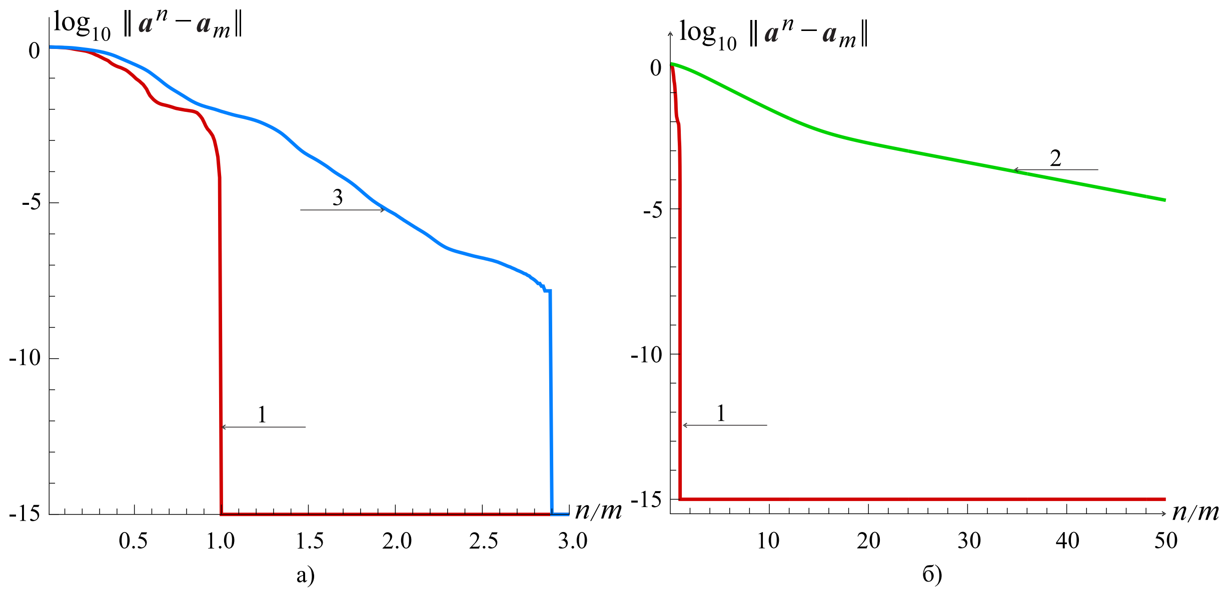


Рис. 1. Абсолютная погрешность вычисления собственного вектора $\boldsymbol{a}_m$ как нормированного решения однородной системы с матрицей $\lambda_m E - A$ методами сопряженных градиентов (а) и ортогонально-степенным алгоритмом (б)

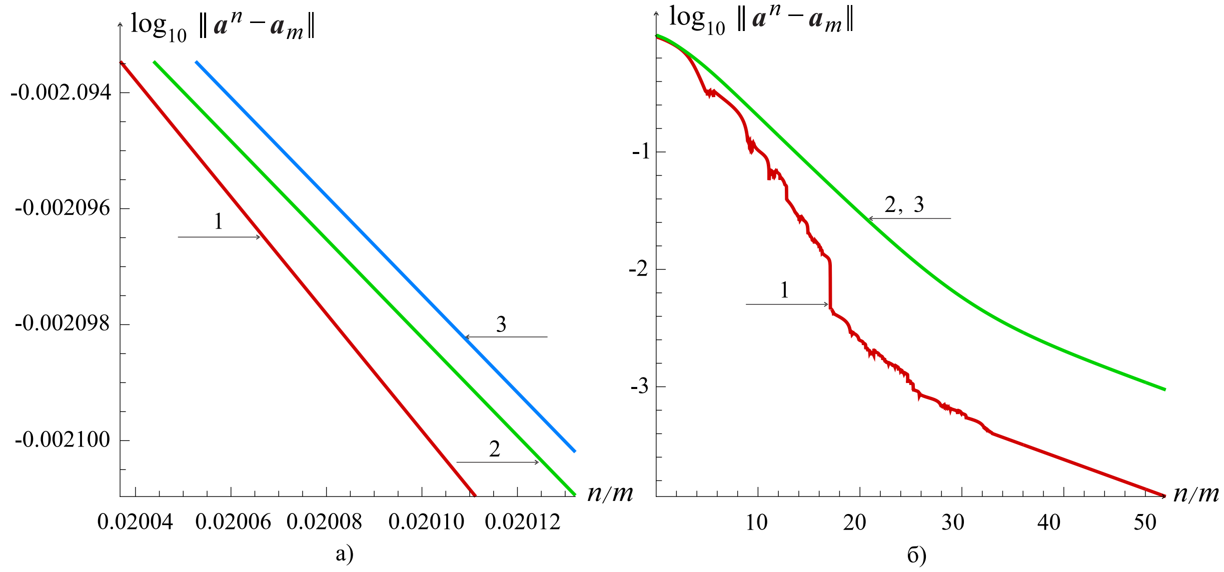


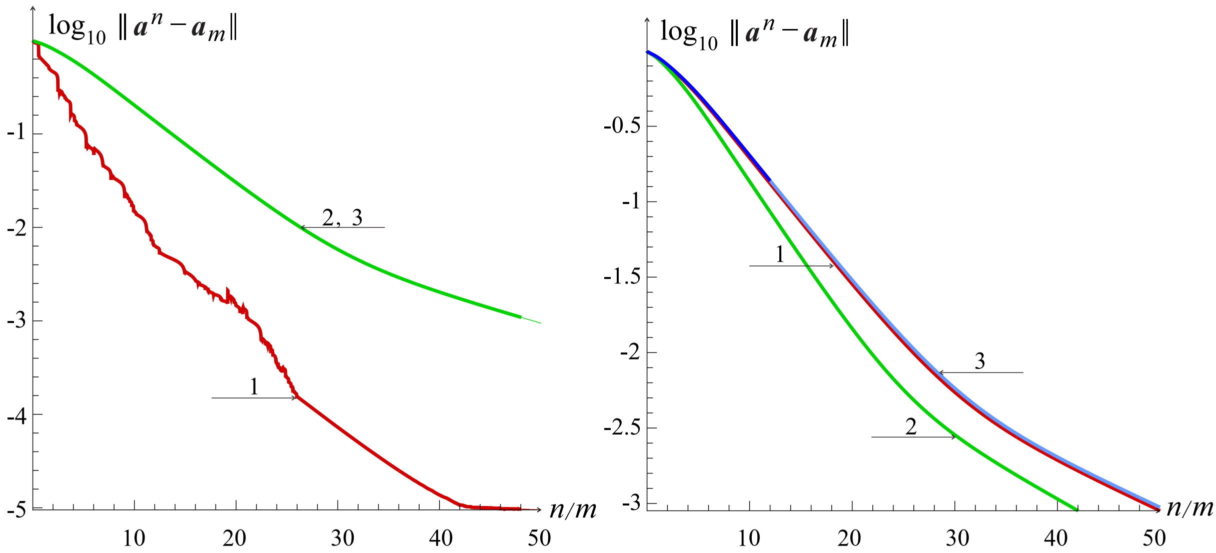
Рис. 2. Абсолютная погрешность вычисления собственного вектора $\boldsymbol{a}_m$ в случае $\mu_n = \dfrac{\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle}{\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle - \lambda_1}$

В этом случае, как отмечено выше, последовательность $\boldsymbol{a}^n$ из (7) совпадает с аналогичной последовательностью степенного метода (5). Заметим, что поскольку $\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle \leqslant \dfrac{\langle A\boldsymbol{a}^n, A\boldsymbol{a}^n \rangle}{\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle} \leqslant \lambda_m$, то относительная погрешность аппроксимации λ_m величиной λ_m^n из (13) не больше, чем аналогичная относительная погрешность степенного метода.

Ситуация на рис. 3б соответствует оптимальному в некотором смысле выбору параметра μ_n. Из соотношений (7) и (12) для относительной погрешности приближения максимального собственного значения λ_m величиной λ_m^n несложно получить следующее асимптотическое при $\varepsilon_n \to 0$ представление:

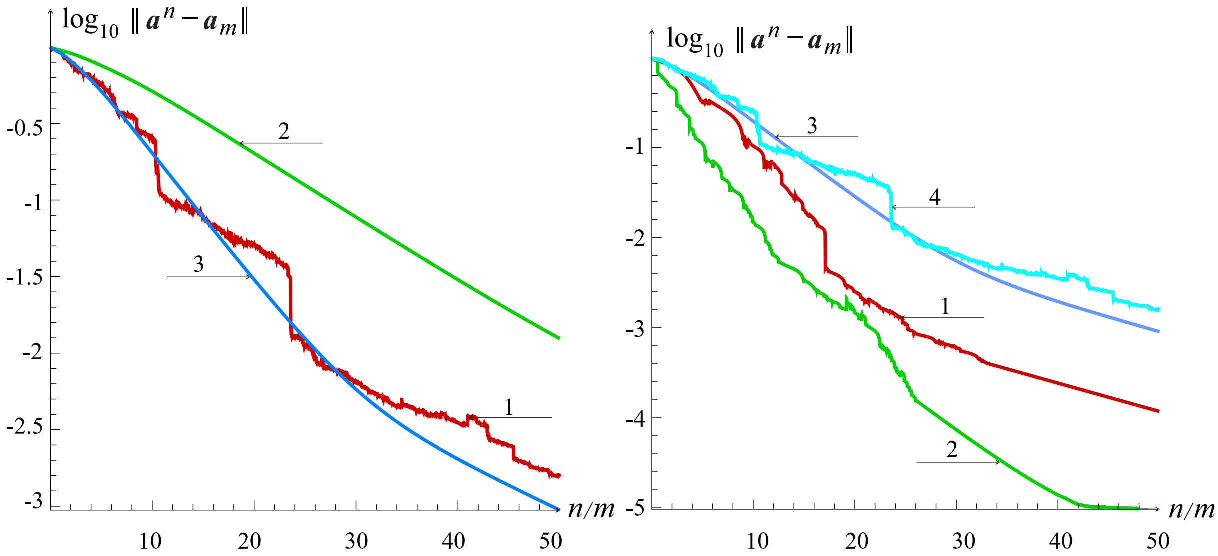
$$\frac{\lambda_m^n(\mu_n) - \lambda_m}{\lambda_m} = \sum_{1 \leqslant i < m} \left(\alpha_i^n\right)^2 f\left(1 - \frac{\lambda_i}{\lambda_m},\, \mu_n\right) + O\left(\varepsilon_n^4\right), \tag{14}$$

где ε_n определено в (6), а $f(t, \mu) = t(\mu t - 1)$ — квадратичная по t функция, зависящая от параметра μ,



а) б)

Рис. 3. Абсолютная погрешность вычисления собственного вектора $\boldsymbol{a}_m$:
а) при $\mu_n \equiv 1$, б) при $\mu_n \equiv 0.5 + \sqrt{0.5}$



а) б)

Рис. 4. Абсолютная погрешность вычисления собственного вектора $\boldsymbol{a}_m$;
а) при $\mu_n = \dfrac{\langle A\boldsymbol{a}^n, \boldsymbol{a}^n\rangle}{\sqrt{\langle A\boldsymbol{a}^n, A\boldsymbol{a}^n\rangle} + \langle A\boldsymbol{a}^n, \boldsymbol{a}^n\rangle}$; б) для четырех рассмотренных последовательностей μ_n: 1) 2б, 2) 3а, 3) 3б, 4) 4а

причем $0 < \mu \leqslant 2$ и $0 < t < 1$. Тогда решением задачи $\max\limits_{0\leqslant t\leqslant 1} \bigl|f(t,\mu)\bigr| \underset{\mu}{\to} \inf$ будет $\mu_* = 0.5 + \sqrt{0.5}$ и $\max\limits_{0\leqslant t\leqslant 1} \bigl|f(t,\mu_*)\bigr| = \mu_* - 1 = \sqrt{0.5} - 0.5 < 0.207$. Отметим, что в варианте (13) для относительной погрешности (14) справедливо равенство $\max\limits_{0\leqslant t\leqslant 1} \bigl|f(t,1)\bigr| = 0.25$, а в рассматриваемом ниже случае $\max\limits_{0\leqslant t\leqslant 1} \bigl|f(t,0.5)\bigr| = 0.5$.

Наконец, рис. 4а соответствует ситуации, когда выражение для λ_m^n совпадает с аналогичным из процесса (5), т.е.

$$\mu_n = \frac{\langle A\boldsymbol{a}^n, \boldsymbol{a}^n\rangle}{\sqrt{\langle A\boldsymbol{a}^n, A\boldsymbol{a}^n\rangle} + \langle A\boldsymbol{a}^n, \boldsymbol{a}^n\rangle}; \quad \lambda_m^n = \sqrt{\langle A\boldsymbol{a}^n, A\boldsymbol{a}^n\rangle} = \bigl\|A\boldsymbol{a}^n\bigr\|.$$

На рис. 4б приведены результаты работы всех четырех рассмотренных выше вариантов процесса (12). Графики показывают, что при $\mu_n \underset{n}{\equiv} 1$ собственный вектор $\boldsymbol{a}_m$ тестовой задачи вычисляется наиболее быстро следующим методом:

1) задано $\boldsymbol{a}^0$ — начальное приближение: $\|\boldsymbol{a}^0\| = 1$, $\langle \boldsymbol{a}^0, \boldsymbol{a}_m \rangle \neq 0$;

2) полагаем $\lambda_m^0 = \dfrac{\langle A\boldsymbol{a}^0, A\boldsymbol{a}^0 \rangle}{\langle A\boldsymbol{a}^0, \boldsymbol{a}^0 \rangle}$, $\boldsymbol{b}^0 = \lambda_m^0 \boldsymbol{a}^0 - A\boldsymbol{a}^0$, $\boldsymbol{c}^1 = \boldsymbol{b}^0$;

3) для $n \geqslant 1$ вычисляем

$$\begin{aligned} &\nu_n = \frac{\langle \boldsymbol{c}^n, \boldsymbol{a}^{n-1} \rangle}{\|\boldsymbol{c}^n\|^2}; \quad \boldsymbol{a}^n = \frac{\boldsymbol{a}^{n-1} - \nu_n \boldsymbol{c}^n}{\|\boldsymbol{a}^{n-1} - \nu_n \boldsymbol{c}^n\|}; \\ &\lambda_m^n = \frac{\langle A\boldsymbol{a}^n, A\boldsymbol{a}^n \rangle}{\langle A\boldsymbol{a}^n, \boldsymbol{a}^n \rangle}; \quad \boldsymbol{b}^n = \lambda_m^n \boldsymbol{a}^n - A\boldsymbol{a}^n; \quad \boldsymbol{c}^{n+1} = \boldsymbol{b}^n - \frac{\langle \boldsymbol{b}^n, \boldsymbol{c}^n \rangle}{\|\boldsymbol{c}^n\|^2} \boldsymbol{c}^n. \end{aligned} \tag{15}$$

Как было отмечено выше, этот итерационный процесс по существу является ускоренной (по технологии сопряженных направлений) версией степенного метода (5), что делает алгоритм (15) еще более привлекательным.

6. Заключение. В настоящей статье предложена и обоснована экономичная версия метода сопряженных направлений для вычисления ненулевого вектора из ядра вырожденной симметричной неотрицательно определенной квадратной матрицы.

Для симметричных неотрицательно определенных квадратных матриц построено однопараметрическое семейство одношаговых нелинейных итерационных процессов, включающее в себя и степенной метод. В каждом из этих процессов последующее приближение к собственному вектору, отвечающему наибольшему собственному значению, получается из предыдущего ортогонализацией к некоторому элементу подпространства Крылова текущей итерации. Доказана сходимость возникающей последовательности векторов к собственному вектору, ассоциированному с наибольшим характеристическим числом матрицы.

Предложена двухшаговая процедура ускорения сходимости итераций этих процессов, в основе которой лежит ортогонализация степенных последовательностей. Приведены результаты численных экспериментов, по результатам которых выбрана наиболее эффективная версия алгоритма, базирующегося на степенном методе.

Работа выполнена при финансовой поддержке РФФИ (проект 14–01–00130).

СПИСОК ЛИТЕРАТУРЫ

1. *Фаддеев Д.К., Фаддеева В.Н.* Вычислительные методы линейной алгебры. СПб: Лань, 2002.
2. *Gene G.H., van der Vorst H.A.* Eigenvalue computation in the 20th century // Journal of Computational and Applied Mathematics. 2000. **123**, N 1–2. 35–65.
3. *Watkins D.S.* The matrix eigenvalue problem: GR and Krylov subspace methods. Philadelphia: SIAM, 2007.
4. *Воеводин В.В., Кузнецов Ю.А.* Матрицы и вычисления. М.: Наука, 1984.
5. *Годунов С.К.* Современные аспекты линейной алгебры. Новосибирск: Научная книга, 1997.
6. *Meurant G., Strakoš Z.* The Lanczos and conjugate gradient algorithms in finite precision arithmetic // Acta Numerica. 2006. **15**. 471–542.
7. *Киреев И.В.* Экономичные критерии останова итераций в методе сопряженных градиентов // Вычислительные технологии. 2015. **20**, № 2. 44–55.
8. *Воеводин В.В.* Линейная алгебра. М.: Наука, 1980.
9. *Воеводин В.В.* Численные методы алгебры. Теория и алгорифмы. М.: Наука, 1966.
10. *Уилкинсон Дж.Х.* Алгебраическая проблема собственных значений. М.: Наука, 1970.
11. *Парлетт Б.* Симметричная проблема собственных значений. М.: Наука, 1983.
12. *Leslie H.L.* Discrete mathematics and its applications. Handbook of linear algebra, Boca Raton: CRC Press, 2013.
13. *Ovtchinnikov E.E.* Computing several eigenpairs of Hermitian problems by conjugate gradient iterations // J. Comput. Phys. 2008. **227**, N 22. 9477–9497.
14. *Горбаченко В.И.* Вычислительная линейная алгебра с примерами на MATLAB. СПб.: БХВ, 2011.

Поступила в редакцию
12.01.2016

An Orthogonal Power Method of Solving the Partial Eigenproblem for a Symmetric Nonnegative Definite Matrix

I. V. Kireev[1]

[1] *Institute of Computational Modeling, Siberian Branch of Russian Academy of Sciences; Akademgorodok 50/44, Krasnoyarsk, 660036, Russia; Ph.D., Associate Professor, e-mail: kiv@icm.krasn.ru*

Received January 12, 2016

Abstract: An efficient version of the conjugate direction method to find a nontrivial solution of a homogeneous system of linear algebraic equations with a singular symmetric nonnegative definite square matrix is proposed and substantiated. A one-parameter family of one-step nonlinear iterative processes to determine the eigenvector corresponding to the largest eigenvalue of a symmetric nonnegative definite square matrix is also proposed. This family includes the power method as a special case. The convergence of corresponding vector sequences to the eigenvector associated with the largest eigenvalue of the matrix is proved. A two-step procedure is formulated to accelerate the convergence of iterations for these processes. This procedure is based on the orthogonalization in Krylov subspaces. A number of numerical results are discussed.

Keywords: eigenvector, eigenvalue, conjugate direction method, Krylov subspaces.

References

1. D. K. Faddeev and V. N. Faddeeva, *Computational Methods of Linear Algebra* (Lan', St. Petersburg, 2002; Freeman, San Francisco, 1963).
2. G. H. Golub and H. A. van der Vorst, "Eigenvalue Computation in the 20th Century," J. Comput. Appl. Math. **123** (1–2), 35–65 (2000).
3. D. S. Watkins, *The Matrix Eigenvalue Problem: GR and Krylov Subspace Methods* (SIAM, Philadelphia, 2007).
4. V. V. Voevodin and Yu. A. Kuznetsov, *Matrices and Computations* (Nauka, Moscow, 1984) [in Russian].
5. S. K. Godunov, *Modern Aspects of Linear Algebra* (Nauchnaya Kniga, Novosibirsk, 1997; Amer. Math. Soc., Providence, 1998).
6. G. Meurant and Z. Strakoš, "The Lanczos and Conjugate Gradient Algorithms in Finite Precision Arithmetic," Acta Numer. **15**, 471–542 (2006).
7. I. V. Kireev, "Inexpensive Stopping Criteria in the Conjugate Gradient Method," Vychisl. Tekhnol. **20** (2), 44–55 (2015).
8. V. V. Voevodin, *Linear Algebra* (Nauka, Moscow, 1980) [in Russian].
9. V. V. Voevodin, *Numerical Methods of Algebra: Theory and Algorithms* (Nauka, Moscow, 1966) [in Russian].
10. J. H. Wilkinson, *The Algebraic Eigenvalue Problem* (Clarendon, Oxford, 1965; Nauka, Moscow, 1970).
11. B. N. Parlett, *The Symmetric Eigenvalue Problem* (Prentice-Hall, Englewood Cliffs, 1980; Nauka, Moscow, 1983).
12. H. L. Leslie (Ed.), *Discrete Mathematics and Its Applications. Handbook of Linear Algebra* (CRC Press, Boca Raton, 2014).
13. E. E. Ovtchinnikov, "Computing Several Eigenpairs of Hermitian Problems by Conjugate Gradient Iterations," J. Comput. Phys. **227** (22), 9477–9497 (2008).
14. V. I. Gorbachenko, *Computational Linear Algebra with Examples in MATLAB* (BKhV Press, St. Petersburg, 2011) [in Russian].