На правах рукописи

Книжнерман Леонид Аронович

Оценки погрешности методов Ланцоша и Арнольди в точной и машинной арифметике

01.01.07 — вычислительная математика

АВТОРЕФЕРАТ
диссертации на соискание учёной степени
доктора физико-математических наук

Москва — 2012

Работа выполнена
в Федеральном государственном бюджетном учреждении науки
Институте вычислительной математики Российской академии наук

Официальные оппоненты:
доктор физико-математических наук, профессор Валерий Павлович Ильин (Федеральное государственное бюджетное учреждение науки Институт вычислительной математики и математической геофизики Сибирского отделения Российской академии наук, лаборатория вычислительной физики, г. н. с.)
доктор физико-математических наук Игорь Евгеньевич Капорин (Федеральное государственное бюджетное учреждение науки Вычислительный центр им. А. А. Дородницына Российской академии наук, с. н. с.)
доктор физико-математических наук Сергей Павлович Суетин (Федеральное государственное бюджетное учреждение науки Математический институт им. В. А. Стеклова Российской академии наук, отдел комплексного анализа, в. н. с.)

Ведущая организация:
Федеральное государственное бюджетное учреждение науки Институт проблем управления им. В. А. Трапезникова Российской академии наук

Защита состоится «___» ________ 2012 г. в ___ часов на заседании диссертационного совета Д 002.045.01 в Федеральном государственном бюджетном учреждении науки Институте вычислительной математики Российской академии наук по адресу: 119333, г. Москва, ул. Губкина, д. 8.

С диссертацией можно ознакомиться в библиотеке Федерального государственного бюджетного учреждения науки Института вычислительной математики Российской академии наук.

Автореферат разослан «___» ________ 2012 г.

Учёный секретарь
диссертационного совета Д 002.045.01
доктор физико-математических наук Г. А. Бочаров

Общая характеристика работы

Исторический обзор

Численное решение дифференциальных и интегральных уравнений после дискретизации часто сводится к выбору приближённого решения из подпространства Крылова[1]

$$\mathcal{K}^m(A,\varphi) \equiv \operatorname{span}\{A^0\varphi, A^1\varphi, \ldots, A^{m-1}\varphi\}, \tag{1}$$

где A — матрица размера $n \times n$ и φ — вектор (далее предполагаемый нормированным). При непосредственной работе со степенным базисом

$$(A^0\varphi, A^1\varphi, \ldots, A^{m-1}\varphi) \tag{2}$$

подпространства (1) вероятны вычислительные трудности, связанные с плохой обусловленностью этого базиса, поэтому обычно предпочитают работу с базисом, полученным в результате ортогонализации (2). Имеются два варианта ортогонализации, и выбор определяет возможность получения *априорных*[2] *оценок погрешности метода.*

1. Осуществляется в неявном виде ортогонализация по Граму–Шмидту базиса (2). Соответствующие методы в случаях эрмитовой и неэрмитовой матрицы A называются *методами Ланцоша и Арнольди.* Проведение классической (эрмитовой) ортогонализации позволяет при исследовании этих методов использовать технику ортогональных многочленов и многочленов Фабера. Если A — не матрица, а ограниченный оператор в гильбертовом пространстве, то для получения оценок можно работать со спектрами и использовать теорию логарифмического потенциала. Вклад в обоснование именно этих методов призвана внести представленная диссертация.

2. Все другие случаи: строится квазиортогональный базис или пара базисов. Методы этого подсемейства семейства крыловских методов (неэрмитовы обобщения метода Ланцоша и др.) также используются на практике, но в них случаются неприятности, когда в знаменателях расчётных формул появляются маленькие числа или ноль. Обычно эти методы можно использовать, когда имеется хороший предобуславливатель, но и это не даёт гарантии сходимости. Путей получения общих априорных оценок для методов этого подсемейства не видно.

---

[1] А. Н. Крылов, Изв. АН СССР, VII сер., Отд. матем. и ест. наук, № 4 (1931), 491–539.

[2] Т. е. не требующих знания результатов вычислений, в отличие от *апостериорных*; см. А. А. Амосов, Ю. А. Дубинский, Н. В. Копченова, *Вычислительные методы для инженеров*, М., Высшая школа, 1994: с. 61–62. Под погрешностью метода мы понимаем отклонение искомой величины от выдаваемого методом приближения (а не оценки промежуточных величин). См. с. 23–24 там же.

Метод Арнольди[3] вычисления собственных пар неэрмитовых матриц и несамосопряжённых операторов стал востребованным в 1980-х годах благодаря[4] работам Ю. Саада. Метод Арнольди применяется во многих задачах: собственно для вычисления спектра;[5] [6] как средство решения систем линейных алгебраических уравнений;[7] [8] для локализации спектра в итерационных методах решения линейных систем;[9] [10] при решении жёстких систем обыкновенных дифференциальных уравнений;[11] [12] для вычисления матричной экспоненты.[13] [14] [15] [16]

Пусть A — ограниченный оператор в гильбертовом пространстве $\mathcal{H}$, $\varphi \in \mathcal{H}$, $\|\varphi\| = 1$. Процесс Арнольди в $\mathcal{H}$ с A и φ основан на ортогонализации по Граму–Шмидту степенного базиса подпространства Крылова (1). Первые m шагов процесса Арнольди можно выразить соотношением

$$AQ = QH + h_{m+1,m}q_{m+1}e_m^T, \tag{3}$$

где $Q = (q_1, \ldots, q_m)$ — набор первых m ортонормированных векторов Арнольди, верхняя хессенбергова $m \times m$-матрица $H = (h_{ij})$ содержит коэффициенты рекурсии, а e_j есть j-й единичный вектор размерности m).

Числа Ритца (собственные значения H), которые считают приближениями к собственным значениям A, не обязаны лежать в спектре S оператора A; они находятся в «поле значений» (множестве значений отношения Рэлея) $\{\langle A\psi, \psi\rangle \mid \psi \in \mathcal{H}, \|\psi\| = 1\}$. Замыкание поля значений будем обозначать K.

Предпринятые до наших работ попытки оценить скорость сходимости метода Арнольди для матриц не привели, на наш взгляд, к получению законченных естественных результатов. Расстояние от собственного вектора до крыловского подпространства (1) оценено Ю. Саадом[17] через коэффициенты разложения φ по собственным векторам A и значе-

[3] W. E. Arnoldi, *Quart. Appl. Math.*, **9** (1951), 17–29.

[4] Х. Д. Икрамов, *Несимметричная проблема собственных значений*, М., Наука, 1991: § 16.

[5] A. Ruhe, *Lect. Notes in Math.*, **973** (1982), 104–120.

[6] Y. Saad, *Linear Algebra and its Applic.*, **34** (1980), 269–295.

[7] Y. Saad, *Math. Comp.*, **37** (1981), 105–126.

[8] Y. Saad, M. H. Schultz, *SIAM J. Sci. Stat. Comp.*, **7** (1986), 856–869.

[9] H. C. Elman, Y. Saad, P. E. Saylor, *SIAM J. Sci. Stat. Comp.*, **7** (1986), 840–855.

[10] D. Ho, F. Chatelin, M. Bennami, *Math. Mod. and Numer. Anal.*, **24** (1990), 53–65.

[11] P. N. Brown, Y. Saad, *SIAM J. Sci. Stat. Comp.*, **11** (1990), 450–481.

[12] C. W. Gear, Y. Saad, *SIAM J. Sci. Stat. Comp.*, **4** (1983), 583–601.

[13] E. Gallopulos, Y. Saad, *SIAM J. Sci. Stat. Comp.*, **13** (1992), 1236–1264.

[14] M. Hochbruck, C. Lubich, *J. Numer. Anal.*, **34** (1997), 1911–1925.

[15] M. Hochbruck, C. Lubich, H. Selhofer, *SIAM J. Sci. Comp.*, **19** (1998), 1552–1574.

[16] Y. Saad, *SIAM J. Numer. Anal.*, **29** (1992), 209–228.

[17] Y. Saad, *Numerical methods for large eigenvalue problems*, Manchester, Manchester Univ. Press, 1992: глава VI, § 7.

ния многочленов степени не выше $m-1$ в точках спектра:

$$\begin{aligned} &\operatorname{dist}\left(z_j, \mathcal{K}^m(A,\varphi)\right) \\ &\le \sum_{1\le k\le n,\ k\ne j} \frac{|\alpha_k|}{|\alpha_j|} \cdot \min_{p\in\mathbb{C}[\lambda],\ \deg p\le m-1,\ p(\lambda_j)=1}\ \max_{\lambda\in\operatorname{Sp}(A),\ \lambda\ne\lambda_j} |p(\lambda)|, \end{aligned} \tag{4}$$

где α_k — коэффициенты разложения $\varphi = \sum_{k=1}^n \alpha_k z_k$ входного вектора φ по набору нормированных собственных векторов z_k матрицы A. Коэффициенты α_k могут быть велики при больших перекосах где-то в спектре A. Таким образом, плохая обусловленность «неинтересной» моды может послужить причиной «развала» оценки для искомой собственной пары. Кроме того, оценить расстояние от собственного вектора до крыловского подпространства недостаточно: из малости этого расстояния не следует малость расстояния от искомого собственного вектора до какого-либо вектора Ритца. Стьюарт[18] распространяет результат Саада на недиагонализируемые матрицы, но оценивает то же расстояние, и его оценка также может содержать необоснованно большие величины. Стьюартом и Джа[19] погрешность аппроксимации собственного значения λ_j на шаге m оценена по порядку корнем m-ой степени из расстояния от соответствующего собственного вектора z_j до подпространства (1):

$$\min_{1\le l\le m} |\lambda_j - \theta_l| \le 4\left(\frac{\delta\|A\|}{\sqrt{1-\delta^2}}\right)^{1/m} \left(2\|A\| + \frac{\delta\|A\|}{\sqrt{1-\delta^2}}\right)^{1-1/m}, \tag{5}$$

где $\delta = \sin\angle\left(z_j, \mathcal{K}^m(A,\varphi)\right)$. Глава 9 диссертации показывает, что в одной из типичных ситуаций указанное расстояние убывает со скоростью геометрической прогрессии при росте m, а тогда из оценки (5) вряд ли можно извлечь что-либо полезное.

Для сравнения скажем, что одна из наших простейших оценок погрешности решения спектральной задачи имеет следующий вид: если есть конечное количество s собственных значений $\lambda_1, \ldots, \lambda_s$, не принадлежащих полю значений $\widehat{K}$ сужения оператора A на подпространство, дополнительное к интересующим нас s модам (это условие даёт возможность вывести из оценок расстояний от нужных собственных векторов до крыловского подпространства оценки ошибок соответствующих векторов Ритца), то оценка погрешности вычисления λ_j $(1 \le j \le s)$ равна $O(\rho_j^m m^c)$, где числа $0 < \rho_j < 1$ определяются взаимным расположением λ_j и $\widehat{K}$, c — константа, зависящая от формы границы $\widehat{K}$, и под знаком O не скрыты никакие перекосы (перекосы влияют на $\widehat{K}$). В общем

[18] G. W. Stewart, *Matrix algorithms. Volume II: eigensystems*, Philadelphia, SIAM, 2001: гл. 4, теорема 3.10.

[19] Zh. Jia, G. W. Stewart, *Math. Comp.*, 70 (2001), 637–647: следствие 4.2.

случае оценка не может быть существенно улучшена. Из этого следует, что оценки (4) и (5), игнорирующие поле значений сужения оператора, в принципе не могут даже гарантировать факт сходимости при $m \to \infty$ для оператора или семейства матриц, удовлетворяющих условиям нашей оценки.

В тот же период Ю. Саадом[20] была дана оценка погрешности решения системы линейных уравнений $Au = \varphi$ с помощью обобщённого метода минимальных невязок GMRES, использующего базис, построенный процессом Арнольди.[21] В этой оценке также (как и в (4)) фигурируют коэффициенты разложения правой части φ по системе собственных векторов A.

В [4], [5, § 4], [18, § 5] автором получены оценки погрешности *метода спектрального разложения Арнольди* (МСРА), т. е. варианта метода Арнольди, предназначенного для вычисления произведения матричной (операторной) функции на вектор. Если функция f аналитична в окрестности S и окрестности $\mathrm{Sp}(H)$ — спектра H, то МСРА в качестве приближения к вектору

$$u = f(A)\varphi \tag{6}$$

выдаёт вектор

$$u_m = Qf(H)e_1. \tag{7}$$

Из рекурсии (3) следует, что МСРА точен для многочленов степени $\leq m-1$, то есть

$$f(A)\varphi = Qf(H)e_1, \qquad f \in \mathbb{C}[t], \quad \deg f \leq m-1. \tag{8}$$

Если f аналитична на всём компакте K, то её можно разложить там в ряд Фабера и благодаря (8) свести оценку погрешности МСРА к оценке остатка ряда Фабера. Использование поля значений A в статье [4] благодаря оценке нормы резольвенты вне K освободило нас от необходимости явным образом отражать в оценке погрешности метода перекосы в спектре. Случай функций, которые могут иметь особенности между несколькими «внешними» собственными значениями A, рассмотрен нами в работах [5] и [18]; там с аналогичной целью мы использовали поле значений сужения A на подходящее инвариантное подпространство. А в работе [19] мы получили результаты, учитывающие спектр оператора (здесь — не матрицы) A; в частности, для FOM (метода полной ортогонализации вычисления $A^{-1}\varphi$) показана сходимость на некоторой последовательности шагов, если 0 лежит в неограниченной связной компоненте $\mathbb{C} \setminus \mathrm{Sp}(A)$, о чём не было речи в старой матричной теории.

[20] Y. Saad, *Iterative methods for sparse linear systems*, Philadelphia, SIAM, 2003: предложение 6.32.

[21] Г. И. Марчук, Ю. А. Кузнецов, Докл. АН СССР, **181** (1968), № 6, 1331–1334.

В [5, § 2], [18, § 3] мы получили оценки для точности вычисления собственных пар. Первое, что хочется сделать для оценки качества выделения собственной пары (λ, z), — воспользоваться формулой (8) и взять многочлен f степени не выше $m-1$, равный 1 в точке λ и принимающий как можно меньшие значения на остальной части спектра или на поле значений ограничения A на подходящее инвариантное подпространство («дополнительное» к $\mathbb{C}\mathrm{z}$). Так и делается, но ситуация осложняется ненормальностью оператора A и ненормальностью H. Часть оценок представляет собой обычные неравенства, но некоторые оценки получаются коши-адамаровского типа: оценивается верхний или нижний предел m-го корня ошибки в терминах функции Грина (с особенностью в бесконечности) поля значений или спектра. Нижним пределом приходится довольствоваться, когда поле значений ограничения A на инвариантное подпространство содержит искомое собственное значение. Наши оценки для спектральной задачи также сформулированы в терминах упомянутых функций Грина для подходящего сужения оператора; соответственно, в доказательствах используются элементы теории потенциала. То, что было сказано выше об отсутствии необходимости явно учитывать перекосы в спектре и об оценках, верных для подпоследовательностей, справедливо и здесь. Использование в формулировках функции Грина позволяет рассматривать произвольные, а не частные (отрезки, эллипсы) поля значений.

В статье [21] мы для частного случая $f(A) = (zI - A)^{-1}$ и $g = 1$ установили в терминах операторов жёсткие ограничения на A, при выполнении которых есть толк от квадратуры Гаусса–Арнольди,[22] [23] под которой понимается приближённое равенство $\langle f(A)\varphi, g(A)\varphi\rangle \approx \langle f(H)e_1, g(H)e_1\rangle$, где функции f и g аналитичны на спектрах A и H; до нас оставался открытым вопрос о том, быстрее ли сходится квадратура по сравнению с приближённым вычислением векторов $f(A)\varphi$ и $g(A)\varphi$ с помощью метода спектрального разложения Арнольди.

Метод Ланцоша[24] теоретически представляет собой метод Арнольди, применённый к самосопряжённому оператору: $A^* = A$. В этом случае матрица H коэффициентов рекурсии является вещественной симметричной трёхдиагональной, а многочлены Арнольди Q_i $(Q_i(A)\varphi = q_{i+1})$ ортогональны на вещественной оси и удовлетворяют трёхчленному рекуррентному соотношению. Компакт K превращается в отрезок вещественной оси; ряд Фабера функции f на K становится сдвинутым рядом Фурье

[22] D. Calvetti, S.-M. Kim, L. Reichel, *SIAM J. Matrix Anal. Appl.*, **26** (2005), 765–781.

[23] R. W. Freund, M. Hochbruck, *In: Linear Algebra for Large Scale and Real-Time Applications*; M. S. Moonen et al. (eds), Kluwer Academic Press, Amsterdam, 1993, 377–380.

[24] C. Lanczos, *J. Res. Nat. Bur. Standards: Sect. B*, **45** (1950), 225–280.

по многочленам Чебышёва. Эрмитовость матрицы H сильно облегчает построение теории в точной арифметике. Общая оценка погрешности *метода спектрального разложения Ланцоша* (МСРЛ; эрмитов аналог МСРА) была получена (для точной арифметики) в [2]; она позволила оценить ошибку вычисления произвольного собственного значения методом Ланцоша. До наших работ были известны лишь теоремы Каниэля[25] и Саада,[26] уточнённые Пэйджем,[27] которые касались аппроксимации со скоростью геометрической прогрессии нескольких собственных значений на одном краю спектра. Результаты Каниэля и Саада оценивают качество приближения j-го точного собственного значения λ_j j-ым же числом Ритца θ_j. Оценки эффективны, если число этих собственных значений существенно меньше, чем размерность подпространства Крылова m.

Используя равенство (8), В. Л. Друскин и автор получили априорные оценки ошибки вычисления не обязательно близкого к краю спектра собственного значения методом Ланцоша.

Исследование метода Ланцоша существенно усложняется при переходе к машинной арифметике (естественно, в случае матриц).[28] [29] [30] Было выяснено, что векторы Ланцоша q_j теряют ортогональность и даже становятся почти линейно зависимыми после того, как появится хорошее приближение к какому-либо собственному значению. Процесс Ланцоша с одной и той же парой (A, φ) может по-разному идти на компьютерах разных типов или даже на одном компьютере при разных способах компиляции программы. Причина в том, что в процессе Ланцоша при стандартной реализации вследствие теоретической трёхдиагональности H новый вектор Ланцоша ортогонализуется только к двум предыдущим, а не ко всем, в отличие от метода Арнольди. Поведению процесса Ланцоша в условиях машинной арифметики посвящены основополагающие работы К. Пэйджа,[31] [32] [33] в которых виртуозно вскрыт механизм появления неустойчивости в процессе Ланцоша и влияние неустойчивости на базовые матричные соотношения процесса. Оценки погрешности[34] можно извлечь

[25] S. Kaniel, *Math. Comp.*, **20** (1966), 369–378.

[26] Y. Saad, *SIAM J. Numer. Anal.*, **17** (1980), 687–706.

[27] Б. Парлетт, *Симметричная проблема собственных значений*, М., Мир, 1983: § 12.4, (12.4.17).

[28] С. К. Годунов, Г. П. Прокопов, *Ж. вычисл. матем. и матем. физ.*, **10** (1970), 1180–1190.

[29] Б. Парлетт, Op. cit.: гл. 13.

[30] G. H. Golub, D. P. O'Leary, *SIAM Review*, **31** (1989), 50–102.

[31] C. C. Paige, *The computation of eigenvalues and eigenvectors of very large sparse matrices*, Ph. D. thesis, London, Univ. of London, 1971.

[32] C. C. Paige, *J. Inst. Math. Applic.*, **18** (1976), 341–349.

[33] C. C. Paige, *Linear Algebra and its Applic.*, **34** (1980), 235–258.

[34] Говоря о погрешности крыловских методов в машинной арифметике, мы имеем в виду суммарный эффект от остановки на конкретном шаге m и от ошибок округления.

из эстетичной работы Гринбаум,[35] которая, используя работы Пэйджа, результат реального процесса Ланцоша на шаге m представила как результат идеального процесса Ланцоша с матрицей размерности $n+m$, но в упомянутой работе наложены очень жёсткие ограничения на параметры задачи и нельзя в общем случае получить оценку погрешности лучше корня четвёртой степени $\varepsilon^{1/4}$ из элементарной ошибки округления ε.

Оценка погрешности МСРЛ в машинной арифметике с правильным порядком по ε была дана в [3]. Для анализа влияния ошибок округления на процесс Ланцоша было построено чебышёвское рекуррентное соотношение с правой частью порядка ошибки округления; эта правая часть отвечает за неточность процесса. Получение оценки ошибки МСРЛ, в правой части которой слагаемое, отвечающее за ошибки округления, линейно по ε (что выведено из результатов Пэйджа), свелось к доказательству устойчивости чебышёвской рекурсии.

В [3], [11] был также в принципе объяснён *феномен Ланцоша* — свойство процесса Ланцоша рано или поздно вырабатывать приближения к собственным значениям с почти машинной точностью. Попытка объяснить феномен Ланцоша была предпринята Каллэм и Уиллафби,[36] однако в их работе используется ряд недоказанных предположений.[37]

Априорных оценок погрешности МСРЛ и МСРА до наших работ не было, если не считать оценки для решения систем линейных уравнений ($f(A)=A^{-1}$).

Итак, до нас:

• спектральные «оценки» в неэрмитовом случае или содержали в левой части промежуточные величины (невязки, расстояния от собственного вектора до крыловского подпространства), а не истинные отклонения чисел и векторов Ритца от собственных значений и векторов соответственно, или в правой части необоснованно содержали потенциально большие величины (результаты Саада, Стьюарта и др.). Непосредственно применимые (а не содержащие оценки лишь промежуточных величин) априорные результаты о сходимости чисел Ритца к собственным значениям относились лишь к краю спектрального интервала в эрмитовом случае при точной арифметике (оценки Каниэля и Саада);

• из матричных функций исследовалась только функция $f(A) = A^{-1}$, соответствующая решению системы линейных уравнений. Не было средств теоретического рассмотрения даже такой важной и популярной функции, как матричная экспонента $f(A)=\exp(-tA)$, $t\geq 0$, $A\geq 0$;

[35] A. Greenbaum, *Linear Algebra and its Applic.*, **113** (1989), 7–63.
[36] J. Cullum, R. A. Willoughby, *Linear Algebra and its Applic.*, **29** (1980), 63–90.
[37] Х. Д. Икрамов, *Итоги науки и техн.: Матем. анализ*, **20** (1982), 179–260: § 9.

• априорные оценки для простого процесса Ланцоша и квадратуры Гаусса–Ланцоша в машинной арифметике имели в лучшем случае правую часть порядка $n^3 m\varepsilon^{1/4}$ (результат Гринбаум и следствия из него);

• Были лишь частные попытки объяснить явление адаптации методов к спектру. Оценки для квадратуры Гаусса–Арнольди ничего не говорили о её эффективности: не было ясно, имеет ли (и если имеет, то при каких условиях на спектр) применение квадратуры преимущество перед вычислением соответствующих матричных функций.

Цель работы:

построение общей теории сходимости методов Ланцоша и Арнольди.

Актуальность темы

Методы Ланцоша и Арнольди — классические методы вычислительной математики. Они широко используются,[38] им уделяется много внимания. Однако имевшиеся до нас априорные результаты об их сходимости носили частный или незаконченный характер. Это и обусловило актуальность темы диссертации.

Научная новизна

В диссертации систематически изложена общая теория сходимости методов Ланцоша и Арнольди (в том числе в форме МСРЛ и МСРА), ликвидирующая имевшиеся качественные пробелы и содержащая прямые априорные оценки погрешности.

Теоретическая и практическая ценность

Результаты диссертации дают пользователям (математикам и программистам):

• принципиальную уверенность в том, что обсуждаемые методы при установленных условиях работают. В частности, прояснены классы матричных функций, к вычислению которых можно применять методы спектрального разложения Ланцоша и Арнольди;

• возможность априори оценивать объём вычислительной работы, достаточный для решения конкретной задачи, а также, зная вид аналитической зависимости границы ошибки от номера шага m, апостериорно прогнозировать фактическую ошибку путём подбора параметров (curve fitting);

[38] На момент написания данного данного предложения поисковая система Google выдаёт 79 тысяч ссылок на запрос “Lanczos method” и 16 тысяч — на запрос “Arnoldi method”.

• инструмент для исследования методов в условиях машинной арифметики: возмущённые чебышёвские и фаберовские рекуррентные соотношения.

Ссылки на описания некоторых приложений, к которым имеет отношение автор, даны в диссертации. Приложений, возможно, и не было бы, если б не была развита теория.

Методы исследования

Из линейной алгебры применены теоремы типа Гершгорина и теоремы о невязке с учётом и без учёта отделённости, из вычислительной математики — теорема о сходимости ньютоновского процесса решения системы нелинейных уравнений, из математического анализа — ортогональные многочлены, многочлены Фабера, специальные функции, конформные отображения, функция Грина и экстремальные многочлены.[39]

Положения, выносимые на защиту:

• Доказано, что при вычислении методом Арнольди «крайних» собственных пар сходимость имеет место со скоростью геометрической прогрессии, показатель которой выражается через значение функции Грина дополнения к полю значений сужения оператора на инвариантное подпространство, порождённое «неинтересной» частью спектра. Предложен метод спектрального разложения Арнольди для вычисления умноженной на вектор операторной функции; дана оценка погрешности этого метода в терминах коэффициентов ряда Фабера вычисляемой функции, построенного на поле значений.

• Исследован метод спектрального разложения Ланцоша для вычисления умноженной на вектор функции от эрмитовой матрицы. Для точной арифметики дана оценка погрешности метода в терминах коэффициентов ряда вычисляемой функции по многочленам Чебышёва, смещённым на спектральный интервал. Оценена погрешность вычисления произвольно расположенного в спектральном интервале собственного значения методом Ланцоша; в частности, показано, что числа Ритца сходятся к заданному собственному значению со скоростью геометрической прогрессии, показатель которой определяется отделённостью собственного значения в спектре.

[39]Автор полагает, что он был среди тех, кто в начале 1990-х годов стал активно внедрять использование поля значений, функции Грина и многочленов Фабера в исследование сходимости крыловских методов. Из «соратников» отметим М. Айерманна, О. Неванлинну и Л. Н. Трефевена с соавторами (список не претендует на полноту). Сейчас этими понятиями удивишь не очень многих вычислительных линейных алгебраистов: они (понятия) начали входить в учебники, а тогда их использование встречало заметное сопротивление.

• С помощью техники возмущённых чебышёвских рекурсий оценка погрешности метода спектрального разложения Ланцоша обобщена на случай машинной арифметики. Показано, что неустойчивость метода спектрального разложения Ланцоша в машинной арифметике не препятствует его практическому применению. Аналогичный результат получен для квадратуры Гаусса–Ланцоша. Доказаны конкретные утверждения, касающиеся феномена Ланцоша; в частности, установлено, что при выполнении ограничений, проистекающих из теории Пэйджа, минимальная ошибка вычисления хорошо отделённого собственного значения имеет примерно порядок машинной ошибки округления.

• В терминах линейных ограниченных операторов показано, почему на практике обсуждаемые методы работают лучше, чем гарантировано предыдущими оценками (адаптация к спектру). Адаптивные оценки погрешности сформулированы в терминах функции Грина неограниченной компоненты дополнения к операторному спектру. Получены результаты о сходимости на подпоследовательностях шагов процесса Ланцоша или Арнольди в ситуациях, когда процесс сходиться в обычном смысле не обязан. Проанализирована эффективность квадратуры Гаусса–Арнольди.

Обоснованность и достоверность результатов

Представленные в диссертации утверждения снабжены математическими доказательствами. Результаты численных экспериментов иллюстрируют многие из доказанных утверждений. Кроме того, справедливость положений представленной теории подтверждается многолетней практикой моделирования геофизических и иных процессов с помощью методов Ланцоша и Арнольди.

Апробация работы

Результаты диссертации излагались на юбилейной ланцошевской конференции (Рэйли, США, 1993), конференции по вычислительной линейной алгебре (Миловы, Чехия, 1997), конференции SIAM-GAMM-2006 (Дюссельдорф, Германия), на кафедре вычислительной математики Технического горного университета Фрайберга (Германия), в ВЦ РАН им. А. А. Дородницына, в ИПУ РАН им. В. А. Трапезникова, по нескольку раз обсуждались на семинарах в Институте вычислительной математики РАН и Институте прикладной математики РАН им. М. В. Келдыша (в том числе на семинаре им. К. И. Бабенко).

Публикации

Основные результаты диссертации изложены в статьях [2, 3, 4, 5, 6, 10, 11, 13, 17, 18, 21], опубликованных в рецензируемых научных журналах из списка, рекомендованного ВАК. Другие результаты диссертации опубликованы в работах [7] и [19] (последняя реферирована в Zentralblatt). Всего по теме диссертации, включая аннотации конференционных докладов и препринты, автором опубликована 21 работа, из них 12 — в рецензируемых научных изданиях.

В статьях, написанных совместно с В. Л. Друскиным, лично диссертант: использовал разложение в ряды Фурье–Чебышёва и специальные функции (В. Л. Друскин оценивал коэффициенты нужных рядов Фурье–Чебышёва с помощью сведения к оценке погрешности разностных схем для обыкновенных дифференциальных уравнений[40]); провёл выкладки, связанные с многочленами Чебышёва; разработал и использовал технику возмущённых чебышёвских рекуррентных соотношений. Включённые в диссертацию результаты, полученные совместно с В. Л. Друскиным и Э. Гринбаум, принадлежат их авторам в равной степени.

Разумеется, результаты из совместных статей, полученные без участия автора, в диссертацию не включены.

Ссылки на статьи автора по теме диссертации можно найти в нескольких книгах[41] [42] [43] [44] [45] [46] [47] и в статьях разных авторов.

Благодарности

Автор благодарит: В. Л. Друскина — за многолетнюю совместную деятельность по исследованию и практическому применению метода Ланцоша, а также за полезные обсуждения личных статей автора; Э. Гринбаум — за совместное написание нескольких статей по теории метода Ланцоша, а также за полезные обсуждения личных статей автора; за полезные обсуждения диссертации в целом или части отражённых в ней статей и за библиографическую поддержку — М. Айерманна, С. К. Асвадурова, Б. Бекерманна, А. Б. Богатырёва, Дж. Голуба, С. А. Горейнова,

[40] V. Druskin, L. Knizhnerman, *Radio Science*, **29** (1994), 937–953: § 5.

[41] С. К. Годунов, *Лекции по современным аспектам линейной алгебры*, Новосибирск, Научная книга (ИДМИ), 2002: предисловие.

[42] G. H. Golub, G. Meurant, *Matrices, moments and quadrature with applications*, Princeton and Oxford, Princeton Univ. Press, 2009: § 7.5, 8.0.

[43] A. Greenbaum, *Iterative methods for solving linear systems*, Philadelphia, SIAM, 1997: гл. 4.

[44] N. J. Higham, *Functions of matrices. Theory and computation*, Philadelphia, SIAM, 2008: § 13.5–13.6.

[45] G. Meurant, *The Lanczos and Congugate Gradient algorithms: From theory to finite precision computations*, Philadelphia, SIAM, 2006: § 3.11.

[46] B. N. Parlett, *The symmetric eigenvalue problem*, Philadelphia, SIAM, 1998: § 13.11.

[47] L. N. Trefethen, M. Embree, *Spectra and pseudospectra: the behavior of nonnormal matrices and operators*, Princeton, Princeton Univ. Press, 2005: гл. VI, § 28.

С. Н. Давыдычеву, Н. Л. Замарашкина, Х. Д. Икрамова, В. П. Ильина, Дж. Каллэм, А. В. Князева, $\boxed{\text{В. И. Лебедева}}$, $\boxed{\text{О. В. Локуциевского}}$, К. Любиха, Ю. М. Нечепуренко, Б. Парлетта, Л. Райхеля, А. Руэ, В. С. Рябенького, А. Скороходова, А. В. Собянина, В. З. Соколинского, С. П. Суетина, К. Торреса-Вердина, Е. Е. Тыртышникова, М. Хохбрук, О. Эрнста, участников перечисленных в пункте «Апробация» конференций и семинаров; Дэвида Бейли — за публикацию фортранного пакета повышенной точности[48] в Интернете; Центральную геофизическую экспедицию и Schlumberger–Doll Research — за организационную помощь.

Структура и объём работы

Диссертация состоит из двенадцати глав, десять из которых содержат математические результаты и распределены между четырьмя частями. Объём диссертации — 255 с. Список литературы содержит 187 наименований.

Краткое содержание работы

Глава 1 «Введение» описывает задачи, даёт исторический обзор и содержит информацию, обязательную для диссертаций.

Часть 1 «Методы Ланцоша и спектрального разложения Ланцоша в точной арифметике: неадаптивные оценки» содержит главы 2–3. Термин «неадаптивные» означает «не адаптированные к спектру», то есть зависящие лишь от спектрального интервала оператора в случае метода спектрального разложения Ланцоша и не зависящие от лакун в спектре, не связанных с отделённостью нужного собственного значения, в случае метода Ланцоша. Адаптивные оценки рассматриваются позже.

Глава 2 «Метод спектрального разложения Ланцоша в точной арифметике» содержит начальную информацию про МСРЛ. Пусть A — симметричная вещественная матрица размера $n \times n$, f — функция, определённая на спектральном интервале A, $\varphi \in \mathbb{R}^n$. Обозначим через λ_i, $i = 1, 2, \ldots, n,\ \ \lambda_{i+1} \geq \lambda_i$, собственные значения A $(\lambda_n > \lambda_1)$, а через z_i — соответствующую ортонормированную систему собственных векторов. Пусть разложение φ по собственным векторам имеет вид

$$\varphi = \sum_{i=1}^{n} \varphi_i z_i. \tag{9}$$

Напомним содержание m шагов процесса Ланцоша приближённого вычисления спектра оператора A. В подпространстве Крылова (1)

[48] D. N. Bailey, *ACM Trans. Math. Software*, **21** (1995), 379–387.

строится базис $q_1, \dots, q_m$, получаемый в результате ортогонализации по Граму–Шмидту последовательности $\varphi, A\varphi, \dots, A^{m-1}\varphi$. Ортогонализацию можно провести с помощью трёхчленной рекурсии

$$Aq_i = \beta_{i-1}q_{i-1} + \alpha_i q_i + \beta_i q_{i+1}, \qquad i = 1, 2, \dots, m-1,$$

где $\beta_0 q_0 \equiv 0$, $q_1 = \varphi$ и $\beta_i \geq 0$. Обозначим через H трёхдиагональную матрицу

$$H = \begin{pmatrix} \alpha_1 & \beta_1 & & \\ \beta_1 & \alpha_2 & \beta_2 & \\ & \ddots & \ddots & \ddots \\ & & \beta_{m-1} & \alpha_m \end{pmatrix},$$

через θ_i, $s_i = (s_{1i}, \dots, s_{mi})^{\scriptscriptstyle T}$, $i = 1, 2, \dots, m$, — собственные значения и соответствующие нормированные собственные векторы H. Положим $Q = (q_1 \dots q_m)$, $y_i = Qs_i$. Тогда (θ_i, y_i) — приближённые собственные пары оператора A, получаемые методом Ритца на $\mathcal{K}^m(A, \varphi)$. Положим

$$B = \frac{\lambda_n + \lambda_1}{\lambda_n - \lambda_1} I_n - \frac{2}{\lambda_n - \lambda_1} A, \qquad g(x) = f\left[\frac{\lambda_n + \lambda_1 - (\lambda_n - \lambda_1)x}{2}\right],$$

где I_n — единичная матрица порядка n. Рассмотрим ряд Фурье–Чебышёва функции g:

$$g(x) = \sum_{k=0}^{\infty} g_k T_k(x). \tag{10}$$

Мы оцениваем скорость сходимости МСРЛ через скорость сходимости ряда (10).

Теорема 1. *Пусть ряд (10) абсолютно сходится на отрезке* $[-1, 1]$. *Тогда для вектора (7) справедлива оценка погрешности*

$$\|u - u_m\| \leq 2 \sum_{k=m}^{\infty} |g_k| < +\infty. \tag{11}$$

В общем случае оценка (11) неулучшаема по порядку. Для многих функций, возникающих при решении дифференциальных уравнений, g_k выражаются через элементарные или специальные функции, что позволяет получить конкретизированные оценки погрешности: см. четыре следующих утверждения.

Предложение 1. *Пусть* $\lambda_1 \geq 0$, $f(A) = \exp(-tA)$, $t \geq 0$, $a = \frac{t\lambda_n}{2}$. *При* $m \leq a$ *справедлива оценка погрешности*

$$\|u - u_m\| \leq 2\left[\sqrt{\frac{2}{\pi}} + O\left(\frac{m}{a}\right)\right] \frac{\sqrt{a}}{m} \exp\left[-\frac{m^2}{2a} + O\left(\frac{m^4}{a^3}\right)\right].$$

Предложение 2. *Пусть* $\lambda_1 \geq 0$, $f(A) = \cos\left(t\sqrt{A}\right)$, $\tau = \frac{2}{\sqrt{\lambda_n}}$, $x = \frac{2t}{\tau}$, $\xi = \frac{2m}{x} - 1$. *Если* $0 < \xi \leq 1$, *то имеет место оценка погрешности*

$$\|u - u_m\| \leq \frac{1 + O\left(\frac{1}{x} + \sqrt{\xi}\right)}{\sqrt{\pi}\sqrt[4]{4m^2 - x^2}} \frac{\exp\left\{-x\left[\frac{(2\xi)^{1.5}}{3} + O(\xi^2)\right]\right\}}{\sqrt{\xi}}.$$

Введём в рассмотрение обратную функцию Жуковского $\Phi(w) = w + \sqrt{w^2 - 1}$.

Предложение 3. *Пусть* $\lambda_1 > 0$, $f(A) = \exp\left(-z\sqrt{A}\right)$, $z \geq 0$. *Справедлива оценка погрешности*

$$\|u - u_m\| \leq 4\Phi(c)^{-m}\left[1 - \Phi(c)^{-1}\right]^{-1}, \qquad c = \frac{\lambda_n + \lambda_1}{\lambda_n - \lambda_1}.$$

Предложение 4. *Пусть* $\lambda_1 \geq 0$, $\lambda_n \leq 1$, $f(A) = A^s$, $s \in \mathbb{N}\backslash\{0\}$. *Справедлива оценка погрешности*

$$\|u - u_m\| \leq \frac{2}{m}\sqrt{\frac{2s}{\pi}}\left[1 + O\left(\frac{m}{s}\right) + O\left(\frac{1}{s - m}\right)\right]\exp\left[-\frac{m^2}{s} + O\left(\frac{m^3}{s^2}\right)\right].$$

Глава 3 «Метод Ланцоша в точной арифметике» содержит доказательство аппроксимационных оценок для собственного значения, не зависящих от его номера в спектре. Первая из оценок не учитывает отделённость желаемого собственного значения в спектре, а вторая учитывает.

Предложение 5. *Предположим, что* $\|A\| \leq 1$, $\lambda_r = 0$, $\varphi_r \neq 0$ *и* $m \geq 3$. *Тогда*

$$\min_{1 \leq i \leq m} |\theta_i| \leq \frac{\left(\frac{4}{\varphi_r^2} - 2\right)^{\frac{1}{m-1}} - 1}{2}. \tag{12}$$

Когда $m \gg \log\left(\frac{4}{\varphi_r^2} - 2\right)$, то правая часть (12) близка к $\frac{\log\left(\frac{4}{\varphi_r^2} - 2\right)}{2m}$.

Определим величину

$$\delta = \min\left\{\max_{\lambda \in [\lambda_1, \lambda_{r-1}] \cup [\lambda_{r+1}, \lambda_n]} |f(\lambda)| \,\middle|\, f(X) \in \mathbb{R}[X],\ \deg f \leq m - 1,\ f(\lambda_r) = 1\right\},$$

где $1 < r < n$, $\varphi_r > 0$ и $m \geq 3$.

Теорема 2. *Если* $\|A\| \leq 1$ *и*

$$\delta < \frac{\varphi_r}{1 + \sqrt{1 - \varphi_r^2}},$$

то выполняется оценка

$$\min_{1\le j\le m} |\lambda_r - \theta_j| \le \frac{\sqrt{2(1-\varphi_r^2)}\,\delta}{\varphi_r\left(1-\delta^{\frac{1+\sqrt{1-\varphi_r^2}}{\varphi_r}}\right)}. \tag{13}$$

Правая часть оценки (13) убывает со скоростью геометрической прогрессии благодаря оптимизации многочлена f,[49] [50], например, по Золотарёву.

Часть 2 «Методы Ланцоша и спектрального разложения Ланцоша в машинной арифметике» состоит из глав 4–6 и также содержит неадаптивные оценки. Оценки из глав 4–5 являются аналогами оценок из глав 2–3 соответственно для машинной арифметики. Оценки главы 6 для *квадратуры Гаусса–Ланцоша* являются аналогами хорошо известных свойств квадратур типа Гаусса.

Глава 4 «Метод спектрального разложения Ланцоша в машинной арифметике» напоминает начальную информацию о машинной арифметике и содержит машинно-арифметические оценки погрешности МСРЛ. Буквой ε обозначается элементарная машинная ошибка округления. Будем придерживаться относительно машинной арифметики предположений из книги Парлетта.[51] Обозначения Q, H и т. п. будут относиться к соответствующим *реально вычисленным* объектам. Введём также следующие обозначения: c_1 — максимальное количество ненулевых элементов в строках A, $\varepsilon_0 \equiv 2(n+4)\varepsilon$, $\varepsilon_1 = \left(7 + c_1\frac{\|\,|A|\,\|}{\|A\|}\right)\varepsilon$, $\varepsilon_2 = \sqrt{2}\max(6\varepsilon_0, \varepsilon_1)$, $\eta = m^{2.5}\|A\|\varepsilon_2$. Пэйджем показано, что при условиях

$$m[6(n+4)\varepsilon + \varepsilon_1] \le 1, \qquad (n+4)\varepsilon < \frac{1}{24} \tag{14}$$

справедливо неравенство

$$\lambda_1 - \eta \le \theta_i \le \lambda_n + \eta, \quad i = 1, 2, \ldots, m. \tag{15}$$

Как принято в теории метода Ланцоша, мы считаем, что все действия с трёхдиагональной матрицей H, включая решение спектральной задачи, производятся точно.

Сделаем спектральное линейное преобразование, переводящее отрезок $[\lambda_1 - \eta, \lambda_1 + \eta]$, заведомо содержащий, в силу (15), собственные значения H, в отрезок $[-1, 1]$. А именно, положим

$$g(x) = f\left[\frac{(\lambda_n + \lambda_1) - (\lambda_n - \lambda_1 + 2\eta)x}{2}\right].$$

[49] В. И. Лебедев, *Ж. вычисл. матем. и матем. физ.*, **9** (1969), 1247–1252.

[50] Е. М. Никишин, В. Н. Сорокин, *Рациональные аппроксимации и ортогональность*, М., Наука, 1988: гл. 2, доказательство утверждения 8.3].

[51] Б. Парлетт, Op. cit.: § 2.4.

Введём в рассмотрение ряд Фурье–Чебышёва

$$g(x) = \sum_{k=0}^{\infty} g_k T_k(x), \qquad -1 \leq x \leq 1. \tag{16}$$

Теорема 3. *Пусть функция f определена на интервале $[\lambda_1 - \eta, \lambda_n + \eta]$ и ряд Фурье–Чебышёва (16) абсолютно сходится на отрезке $[-1, 1]$. Тогда вектор u_m определён корректно и при выполнении (14) справедлива оценка погрешности*

$$\|u - u_m\| \leq 4\sqrt{\frac{2}{\pi}} \|g\|_{L_2\left(]-1,1[;1/\sqrt{1-x^2}\right)} \frac{\|A\|}{\lambda_n - \lambda_1} m^3 \varepsilon_1 + \frac{10}{\sqrt{24}} \sqrt{m} \sum_{k=m}^{\infty} |g_k|. \tag{17}$$

Член оценки (17), отвечающий за машинную арифметику, имеет порядок машинной ошибки округления.

Пример запуска МСРЛ. В Горном техническом университете Фрайберга (Германия) одна из научных программ, написанная автором совместно с В. Л. Друскиным, вычисляющая матричную функцию $f(A) = A^{-1}\exp(-tA), \;\; t \geq 0$, была запущена с одними и теми же входными данными на персональных компьютерах, работавших под управлением операционных систем Windows и Linux. Трансляция была осуществлена фортранными компиляторами фирмы Intel: автором для Windows и в ИВМ РАН — для Linux. Таблица 1 показывает, что два процесса Ланцоша в конце концов выдали приближённое решение, совпадающее в пяти десятичных знаках. При этом коэффициенты ланцошевской рекурсии и значения приближённого решения по ходу процесса (от момента потери ортогональности векторами Ланцоша и до достижения сходимости заданного уровня) различались весьма значительно.

Глава 5 «Феномен Ланцоша и расположение чисел Ритца» содержит машинно-арифметические оценки погрешности процесса Ланцоша. Рассмотрим процесс Ланцоша с начальным вектором (9) и будем интересоваться качеством приближений к фиксированному собственному значению λ_r матрицы $A, \;\; 1 \leq r \leq n$. Без ограничения общности можно считать, что $\|A\| \leq 0.9$ и $\lambda_r = 0$.

Предложение 6. *Пусть $\|A\| \leq 0.9, \;\; \lambda_r = 0, \;\; \varphi_r \neq 0$ в (9) и $m \geq 3$. Предположим, что выполнены условия Пэйджа (14), а также условия*

$$9m^{2.5}\varepsilon_2 \leq 1, \qquad m^2 \varepsilon_1 \leq \frac{1}{4}\sqrt{\frac{0.19}{2}} |\varphi_r|.$$

Тогда справедлива оценка

$$\min_{1 \leq i \leq m} |\theta_i| = \min_{1 \leq i \leq m} |\theta_i - \lambda_r| \leq \frac{1}{2}\left[\left(\frac{10\sqrt{m}}{\sqrt{6}|\varphi_r|}\right)^{\frac{2}{m-2}} - 1\right]. \tag{18}$$

Таблица 1: Результаты вычисления матричной функции на двух компьютерах, работающих под управлением операционных систем Windows и Linux. j — номер шага процессов Ланцоша, α_j и β_j — коэффициенты ланцошевских рекурсий, U_j — значения приближённого решения в отдельном (контрольном) узле пространственной разностной сетки.

	Linux			Windows		
j	β_j	α_j	U_j	β_j	α_j	U_j
50			-0.15702E-14			-0.15702E-14
100	998.23	20434.	-0.16543E-09	998.23	20434.	-0.16543E-09
150			-0.15208E-07			-0.15208E-07
200	1987.9	2817.1	-0.79739E-07	1985.8	2813.8	-0.79744E-07
250			-0.62672E-06			0.63888E-06
300	5083.4	2987.7	-0.31798E-05	9944.3	9050.3	-0.32142E-05
350			-0.77681E-05			-0.76203E-05
400	10448.	12052.	-0.11058E-04	8183.5	5388.9	-0.11047E-04
450			-0.11686E-04			-0.11686E-04
500	11734.	11345.	-0.11666E-04	9273.5	5455.4	-0.11666E-04

Если

$$\frac{2}{m-1}\log\frac{10\sqrt{m}}{\sqrt{6}|\varphi_r|} \ll 1,$$

то правая часть оценки (18) примерно равна

$$\frac{1}{m}\log\frac{10\sqrt{m}}{\sqrt{6}|\varphi_r|}.$$

Таким образом, оценка имеет по m порядок $\frac{\log m}{m}$.

Если известна отделённость собственного значения, то ошибку аппроксимации можно оценить членом геометрической прогрессии с известными параметрами.

Теорема 4. *Пусть* $1 < r < n$, $\varphi_r > 0$, $\gamma \equiv \min_{1\le i\le n,\ i\ne r} |\lambda_r - \lambda_i| > 0$, $m \ge 3$, $f(X) \in \mathbb{R}[X]$, $\deg f \le m-3$, $f(\lambda_r) = 1$,

$$|f(x)| \le \begin{cases} 1 & \textit{при} \quad x \in [\lambda_1 - \eta, \lambda_n + \eta], \\ \mu \le 1 & \textit{при} \quad x \in [\lambda_1 - \eta, \lambda_n + \eta] \setminus]\lambda_{r-1} + \delta, \lambda_{r+1} - \delta[\end{cases}$$

и (помимо (14)) выполнены условия $D \le c_2\varphi_r$ *и* $m\varepsilon_2 \le c_3\gamma\|A\|^{-1}$, *где* $D \equiv \sqrt{m}\,\mu + \Delta$, $\delta \equiv c_4 m^3\varepsilon_2\|A\|$, $\Delta \equiv m^3\varepsilon_2\|A\|(\lambda_n - \lambda_1)^{-1}$, *и где* c_2 *и* c_3 *— достаточно малые, а* c_4 *— достаточно большая положительная* *константа. Тогда справедлива оценка*

$$\|A\|^{-1} \min_{1\le i\le m} |\theta_i - \lambda_r|$$
$$\lesssim \max\left[\sqrt{m}\varphi_r^{-2}\left(m^{1.5}\varepsilon_2\|A\|\gamma^{-1} + D\right), \varphi_r^{-1}(m^{1.5}\varepsilon_2 + D)\|A\|\gamma^{-1}\right].$$

При достаточно большом допустимом числе шагов оценка теоремы 4 имеет порядок машинной ошибки округления.

Когда матрица A является невырожденной, но не положительно определённой, то матрица $H = H_m$, порождённая m шагами процесса Ланцоша, может быть вырожденной. В этом случае аппроксиманта МСРЛ к решению системы линейных уравнений $Au = \varphi$ не определена на m-ом шаге процесса. Мы показываем, однако, что по меньшей мере одна из двух последовательных трёхдиагональных матриц, H_m или H_{m-1}, имеет спектр, чьё относительное расстояние до нуля больше, чем примерно нормированный квадрат расстояния от $\mathrm{Sp}(A)$ до нуля.

Предложение 7. *Положим*

$$d(z) \equiv \min_{i=1,2,\ldots,n} |z - \lambda_i| = \mathrm{dist}(z, \mathrm{Sp}(A)).$$

Пусть θ — собственное значение H_m, μ — собственное значение H_{m-1}, $\|A\| = 1$ и ни θ, ни μ не слишком близко к спектру A, именно,

$$\min\{d(\theta), d(\mu)\} > \left[(m+1)^3 + \sqrt{3}\,m^2\right]\varepsilon_2.$$

Тогда

$$\max\{|\theta|, |\mu|\} \ge \frac{d(0)^2 - a\varepsilon_1}{d(0) + 6.25\beta_m + \sqrt{(6.25\beta_m)^2 + 12.5\beta_m d(0) + c\varepsilon_1}}, \qquad (19)$$

где $a \le 12.5\ \sqrt{m} + O(\varepsilon)$. При условии

$$90.625\,[(2n+6)\varepsilon + m(3\varepsilon_0 + \varepsilon_1)] + 12.5\sqrt{m}\,\varepsilon_1 \le 8.5$$

неравенство (19) может быть заменено на

$$\max\{|\theta|, |\mu|\} \ge \frac{d(0)^2}{16} - \frac{12.5}{16}\sqrt{m}\,\varepsilon_1 + O(\varepsilon^2).$$

Попутно в этой главе получено обобщение теоремы Стилтьеса о разделении,[52] [53] улучшающее аналогичный результат из работы Гринбаум.[54]

[52] T. J. Stieltjes, *Annales scientifiques de É.N.S. 3e série*, **1** (1884), 409–426.
[53] Е. М. Никишин, В. Н. Сорокин, Op. cit.: гл. 2, § 8, упражнение 6.
[54] A. Greenbaum, Op. cit.

Глава 6 «Гауссова квадратурная формула, порождаемая простым процессом Ланцоша, и её приложения» посвящена квадратурной формуле типа Гаусса (Гаусса–Ланцоша, КФГЛ), порождаемой m первыми шагами *простого* (т. е. без реортогонализации) *процесса Ланцоша*. Под этим понимается квадратурная формула (для матриц это, точнее, формула приближённого суммирования)

$$\langle f(A)q_1, q_1\rangle \approx \langle f(H)e_1, e_1\rangle = \sum_{i=1}^{m} s_{1i}^2 f(\theta_i),$$

где f — функция, определённая и гладкая на отрезке $[\lambda_1 - \eta, \lambda_n + \eta]$ и где первое скалярное произведение берётся в $\mathbb{R}^n$, а второе — в $\mathbb{R}^m$. Нарушение ортогональности векторов Ланцоша в простом процессе Ланцоша не позволяет просто обосновать законность использования КФГЛ, порождаемой m первыми шагами простого процесса Ланцоша, и побуждает использовать вычислительно дорогую реортогонализацию, что нежелательно. Это порождает задачу обоснования КФГЛ в условиях машинной арифметики.

Теорема 5. *Пусть $\sigma \equiv \|A\|/(\lambda_n - \lambda_1)$. Пусть функции f и g определены на $[\lambda_1 - \eta, \lambda_n + \eta]$ и разлагаются в ряды*

$$f(\lambda) = \sum_{k=0}^{\infty} f_k T_k\left(\frac{\lambda_n + \lambda_1 - 2\lambda}{\lambda_n - \lambda_1 + 2\eta}\right), \qquad g(\lambda) = \sum_{k=0}^{\infty} g_k T_k\left(\frac{\lambda_n + \lambda_1 - 2\lambda}{\lambda_n - \lambda_1 + 2\eta}\right),$$

сходящиеся абсолютно на указанном отрезке. Тогда справедливы неравенства

$$\begin{aligned} &|\langle f(A)q_1, g(A)q_1\rangle - \langle f(H)e_1, g(H)e_1\rangle| \\ &\le O\left(m^3\sigma\varepsilon_2\right) \sum_{k+l\le 2m-2} |f_k g_l| + 2 \sum_{k+l\ge 2m-1} |f_k g_l| \end{aligned}$$

и

$$\left|\|f(A)q_1\|^2 - \|f(H)e_1\|^2\right| \le O\left(m^3\sigma\varepsilon_2\right) \sum_{k+l\le 2m-2} |f_k f_l| + 2 \sum_{k+l\ge 2m-1} |f_k f_l|.$$

Следствие 1. *В условиях теоремы 5 верна оценка*

$$|\langle f(A)q_1, q_1\rangle - \langle f(H)e_1, e_1\rangle| \le O\left(m^3\sigma\varepsilon_2\right) \sum_{k=0}^{2m-2} |f_k| + 2 \sum_{k=2m-1}^{\infty} |f_k|. \quad (20)$$

Бесконечный ряд в оценке (20) для КФГЛ начинается примерно с вдвое большего номера, чем в оценке (17) теоремы 3 для МСРЛ.

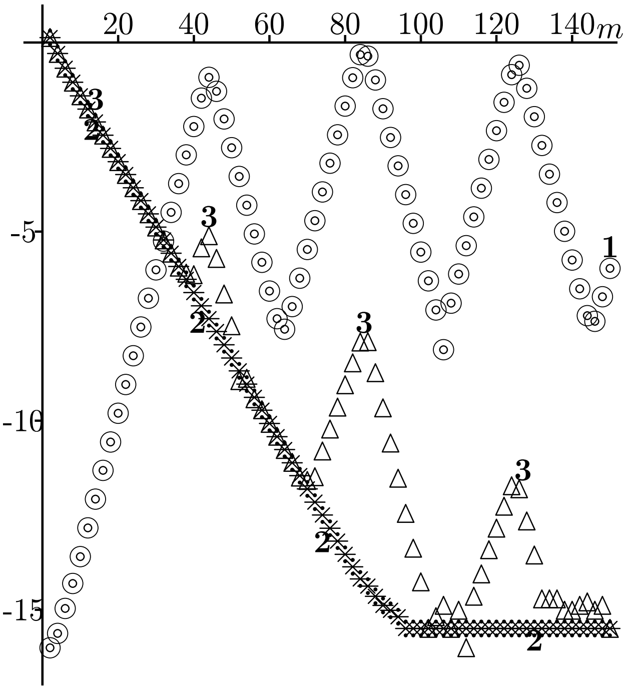

Рис. 1: Пример вычисления величины $\langle A^{-1}q_1, q_1\rangle$. Кривая 1 — график $\log_{10}|\langle q_1, q_m\rangle|$, кривая 2 — $\log_{10}|\langle A^{-1}q_1, q_1\rangle - \langle H^{-1}e_1, e_1\rangle|$, кривая 3 — $\log_{10}|\langle A^{-1}q_1, q_1\rangle - \langle QH^{-1}e_1, q_1\rangle|$.

В конце главы описываются приложения доказанных оценок: конструктивные — к методу спектрального разложения Ланцоша и решению некорректных задач с помощью вариационной регуляризации и теоретические — к феномену Ланцоша. В иллюстративных целях рассматривается пример со сравнением вычисления $\langle A^{-1}\varphi, \varphi\rangle$ с помощью КФГЛ и непосредственно с помощью МСРЛ. Матрица A положительно определена и имеет хорошо отделённое собственное значение. Приведены результаты счёта для чётных m (для нечётных m картинка несколько сдвинута, но качественно не отличается). Рисунок 1 показывает, что величина $\left\langle H^{-1}e_1, e_1\right\rangle$ стабильно сходится к решению, в то время как $\left\langle QH^{-1}e_1, q_1\right\rangle$ делает «подскоки», связанные, очевидно, с моментами потери ортогональности, происходящей при сходимости очередного дубля к хорошо отделённому собственному значению.

Часть 3 «Методы Арнольди и спектрального разложения Арнольди: неадаптивные оценки» содержит главы 7–9. «Неадаптивность» здесь понимается в том смысле, что оценки сформулированы в терминах поля значений оператора или сужения оператора на подходящее инвариантное подпространство, а не в терминах спектра. Здесь мы получаем аналоги результатов части 1 для несамосопряжённого случая; вместо многочленов Чебышёва и рядов Фурье–Чебышёва появляются многочлены и ряды Фабера; появляются также функции Грина, в терминах которых можно оценить поведение нужных экстремальных многочленов.

В главе 7 «Метод спектрального разложения Арнольди: общие оценки» оценивается погрешность вычисления вектора (6) с помощью метода Арнольди в несимметричном случае в условиях точной и машинной арифметики. Вместо разложения f в ряд Фурье–Чебышёва, применявшегося в симметричном случае, здесь используется разложение в ряд Фабера[55] на поле значений.

Процесс Арнольди в компьютерной арифметике с реортогонализацией может быть выражен формулами $AQ - QH = h_{m+1,m}q_{m+1}e_m^{\scriptscriptstyle T} + F$ и $Q^{\scriptscriptstyle T}Q - I = G$, где под Q, H и др. понимаются реально вычисленные на ЭВМ объекты и нормы $n \times n$-матрицы F и $m \times m$-матрицы G — умеренные кратные величины ε. Мы даём оценки погрешности в терминах величин $\|F\|$, $\|G\|$ и $\|Q^{\scriptscriptstyle T}q_{m+1}\|$, не вдаваясь в стандартную для вычислительной линейной алгебры задачу оценивания упомянутых норм. (Величины $\|F\|$ и $\max(\|G\|, \|Q^{\scriptscriptstyle T}q_{m+1}\|)$ не превосходят произведений, соответственно, $\varepsilon\|A\|$ и ε на одночлены от m, n с небольшими показателями степеней и коэффициентами.) Положим $\eta \equiv \|F\| + \max(\|G\|, \|Q^{\scriptscriptstyle T}q_{m+1}\|)\|A\|$. Предположим, что выполняются условия $\|G\| \le 0.5$, $\|F\| \le \|A\|$, $\|Q^{\scriptscriptstyle T}q_{m+1}\| \le 1$, $\|q_{m+1}\| \le 2$, $|h_{m+1,m}| \le 2\|A\|$. Они естественны и следуют из оценок точности вычисления скалярных произведений, линейных комбинаций векторов, умножения матрицы на вектор и т. п.[56] [57] при наложении на m, n и ε ограничений типа полиномиальных неравенств.

Свяжем с матрицей A множество K её значений отношения Рэлея. Будем рассматривать задачу вычисления (6) в предположении аналитичности f на K. Обозначим через ψ функцию, конформно и биективно отображающую внешность единичного круга $\{w \in \overline{\mathbb{C}} \mid |w| > 1\}$ на дополнение к K в расширенной комплексной плоскости $\overline{\mathbb{C}}$ при условиях $\psi(\infty) = \infty$ и $\psi'(\infty) > 0$, через Γ — границу K, а через Γ_ρ, $\rho > 1$, — линию уровня $\{\psi(w) \mid |w| = \rho\}$. Через Φ_k, $k \in \mathbb{N}$, обозначаем многочлены Фабера для K — целые части k-ой степени ψ^{-1}. Пусть

$$f(\lambda) = \sum_{k=0}^{\infty} f_k \Phi_k(\lambda) \tag{21}$$

— ряд Фабера функции f на K.

Теорема 6. *Если функция f аналитична в канонической области G_R, ограниченной кривой Γ_R с параметром $R > 1+\eta^{1/\alpha}$, то справедлива*

[55]П. К. Суетин, *Ряды по многочленам Фабера*, М., Наука, 1984.

[56]Б. Парлетт, Op. cit.: гл. 2.

[57]G. H. Wilkinson, *Rounding errors in algebraic processes*, N. Y., Prentice-Hall, 1964.

оценка погрешности

$$\|u-u_m\| \lesssim \max_{z\in\Gamma}|f(z)|\,\|F\|m^{2\alpha+2.5}\max_{0\le l\le m-1}\left(1+\frac{1}{\max(l,1)}+\eta^{1/\alpha}\right)^l \qquad (22)$$
$$+\sum_{k=m}^{\infty}|f_k|\left(1+\frac{1}{k}+\eta^{1/\alpha}\right)^k k^\alpha,$$

где константа $\alpha \ge 1$ и константа под знаком $\lesssim$ зависят от K.

Следствие 2. *Если $\varepsilon = 0$ (точная арифметика) и функция f аналитична на K, то справедлива оценка погрешности*

$$\|u-u_m\| \lesssim \sum_{k=m}^{\infty}|f_k|k^\alpha, \qquad (23)$$

где константа α и константа под знаком $\lesssim$ зависят от K.

С помощью недавнего результата Бекерманна[58] можно убрать степенной множитель из правой части (23).

Далее в главе с помощью примера с жордановыми клетками обосновано использование именно поля значений, а не матричного спектра, а также рассмотрено приложение к решению систем линейных алгебраических уравнений, в том числе с рестартами.

В методе Арнольди ортогональность векторов Арнольди устойчиво поддерживается,[59]. поэтому вопросы машинной арифметики гораздо менее актуальны для метода Арнольди, чем для метода Ланцоша, частным случаем которого метод Арнольди является лишь в точной арифметике. Эта глава — последняя, где фигурирует машинная арифметика.

Глава 8 «Оценки погрешности метода Арнольди и метода спектрального разложения Арнольди: случай нормальной матрицы и крайних изолированных собственных значений» начинает изучение влияния дискретности края спектра на поведение процесса Арнольди. В предыдущей главе метод Арнольди, применённый к матрице $A \in M_n(\mathbb{R})$ и вектору $\varphi \in \mathbb{R}^n$, исследован как средство приближённого вычисления векторов вида (6), где f — функция, аналитическая на множестве отношений Рэлея матрицы A. В настоящей главе изучается сходимость в методе Арнольди «крайних» собственных значений, не принадлежащих выпуклой оболочке K остальных собственных значений, и

[58]B. Beckermann, *C. R. Acad. Sci. Paris: Ser. I*, **340** (2005), 855–860.

[59]Хотя устойчивость вычисления конкретного базиса не гарантируется. См., например, A. Malyshev, M. Sadkane, *Linear Algebra and its Applic.*, **371** (2003), 315–331; J.-F. Carpraux, S. K. Godunov, S. V. Kuznetsov, *Linear Algebra and its Applic.*, **248** (1996), 137–160.

соответствующих собственных векторов. Даны аналоги теорем Каниэля и Саада из теории метода Ланцоша, сформулированные в терминах функции Грина (с особенностью в бесконечности) дополнения к K в расширенной комплексной плоскости. Затем исследуется метод Арнольди как метод вычисления (6) при ослабленном по сравнению с предыдущей главой ограничении на функцию f: ей разрешается иметь особенности между крайними собственными значениями A (f должна быть аналитична в области, содержащей K и искомые собственные значения). Показано, что при больших m оценённая скорость сходимости МСРА не зависит от выделенных собственных значений.

Глава 9 «Оценки погрешности метода Арнольди и метода спектрального разложения Арнольди: случай крайних изолированных собственных значений» содержит результаты, аналогичные результатам предыдущей главы, но не использующие предположения о нормальности. В терминах функции Грина дополнения к полю значений K сужения оператора на подпространство, дополнительное к искомым собственным векторам, даны более слабые (что обусловлено отсутствием теоремы о невязке с учётом отделённости) аналоги теорем Каниэля и Саада и последнего результата предыдущей главы.

Пусть $\mathcal{H}$ — гильбертово пространство, $A\colon \mathcal{H} \to \mathcal{H}$ — ограниченный линейный оператор, φ — элемент $\mathcal{H}$ с $\|\varphi\| = 1$. Мы рассмотрим процесс Арнольди с A и φ. Пусть $\lambda_1, \ldots, \lambda_s \quad (s \geq 1)$ — различные собственные значения A, $\xi = \left[\prod_{i=1}^{s}(A - \lambda_i I)\right] \varphi \in \mathcal{H}$, $\mathcal{G}$ — замыкание линейного подпространства $\operatorname{span}\left\{\xi, A\xi, A^2\xi, \ldots\right\}$. Будучи замыканием инвариантного подпространства, $\mathcal{G}$ само инвариантно, и мы можем определить ограничение $B = A|_{\mathcal{G}}\colon \mathcal{G} \to \mathcal{G}$ и замыкание K поля значений B.

Мы исследуем сходимость собственных значений $\lambda_1, \ldots, \lambda_s$ в предположениях, что $\lambda_1 \notin K, \ldots, \lambda_s \notin K$ и что $\varphi_1 \neq 0, \ldots, \varphi_s \neq 0$. Пусть Ψ — функция, конформно и биективно отображающая дополнение к единичному кругу $\left\{w \in \overline{\mathbb{C}} \mid |w| > 1\right\}$ на дополнение к K в $\overline{\mathbb{C}}$ при условиях $\Psi(\infty) = \infty,\ \Psi'(\infty) > 0$, Φ — функция, обратная к Ψ, $t_i = \Phi(\lambda_i)$. Определим величины $\delta_i = \mid t_i \mid^{-m} m^{c_5}/\varphi_i$, где константа $c_5 \geq 1$ зависит от формы K.

Теорема 7. *Метод Арнольди, применённый к оператору A и вектору φ, за $m > s$ шагов выработает такие приближённые собственные пары $(\theta_1, y_1), \ldots, (\theta_m, y_m)$, что при подходящей нумерации и нормализации верны оценки*

$$\lambda_i - \theta_i \lesssim \delta_i, \qquad \|z_i - y_i\| \lesssim \delta_i \qquad (i = 1, \ldots, s).$$

Обозначим через Φ_k $(k \in \mathbb{N})$ многочлены Фабера, построенные на

K, через Γ_ρ — изолинии функции Грина $\Gamma_\rho = \{\, \Psi(w) \mid w \in \mathbb{C}, |w| = \rho \,\}$, $\rho > 1$. По функции f, аналитической на K, можно построить ряд Фабера $f = \sum_{k=0}^{\infty} f_k \Phi_k$, сходящийся на K.

Теорема 8. *Пусть функция f аналитична в области, содержащей внутренность контура Γ_R $(R > 1)$ и окрестности точек $\lambda_1, \ldots, \lambda_s$. Тогда метод Арнольди, применённый к A и φ, за достаточно большое число шагов m выработает корректно определённую аппроксиманту u_m к u, такую что*

$$\varlimsup_{m \to \infty} \|u - u_m\|^{1/m} \le R^{-1}.$$

Часть 4 «Методы Ланцоша, Арнольди, спектрального разложения Ланцоша и спектрального разложения Арнольди: адаптивные оценки» содержит главы 10–11. Гл. 10 объясняет, почему результаты применения методов Ланцоша и Арнольди обычно лучше, чем обещано оценками из предыдущих глав. Вводится понятие адаптации обоих методов к спектру оператора (который в данном случае нельзя считать матрицей), то есть зависимости качества результатов от спектра. Приводятся количественные выражения утверждений о том, что лакуны в спектре ускоряют сходимость и что сходимость в значительной степени обусловлена спектром S, а не полем значений K. Гл. 11 в терминах операторов обсуждает качество *квадратуры Гаусса–Арнольди* для функции $\langle (zI - A)^{-1}\varphi, \varphi \rangle$ и связь между методом Арнольди и *рациональной аппроксимацией типа Паде функций марковского типа.*

Глава 10 «Об адаптации методов Ланцоша и Арнольди к спектру, или почему два этих метода так мощны» трактует адаптацию изучаемых методов к операторному спектру. Простое вычисление начального куска ряда Фурье–Чебышёва (10) и ряда Фабера (21) с подставленной туда матрицей даёт несколько лучшие оценки, чем (11) и (23). Причина успешного использования МСРЛ/МСРА заключается в так называемой адаптации обоих методов к спектру, то есть в зависимости их эффективности от спектра A, теоретический учёт лакун в котором способен приводить к лучшим оценкам, чем (11) и (23). Данная глава посвящена как раз этому аспекту поведения МСРЛ/МСРА и собственно методов Ланцоша и Арнольди как методов вычисления спектра. По причине нерелевантности в ненормальном случае матричного спектра работа ведётся с операторами.

Пусть A — ограниченный оператор в гильбертовом пространстве $\mathcal{H}$, $\varphi \in \mathcal{H}$, $S = \operatorname{Sp}(A)$, D — связная компонента дополнения $\overline{\mathbb{C}} \backslash S$, содержащая бесконечность, $E = \overline{\mathbb{C}} \backslash D \supseteq S$, $\operatorname{capl} E$ — логарифмическая ёмкость. Если $\operatorname{capl} E > 0$, то существует обобщённая функция Грина g для D;

в противном случае считаем, что $g \equiv +\infty$. Обозначим через r_k^A невязку FOM на k-ом шаге процесса Арнольди с A и φ (если приближённое решение МСРА не существует, мы полагаем $\left\|r_k^A\right\| = +\infty$).

Предложение 8. *Если $0 \notin E$, то для FOM справедлива оценка*

$$\varliminf_{m\to\infty} \left\|r_m^A\right\|^{1/m} \le e^{-g(0)}. \tag{24}$$

Адаптивный характер (24), так же, как и оценок ошибки из последующих утверждений, очевиден благодаря монотонности функции Грина[60] по области: чем меньше множество E, тем больше (нестрого) показатель $g(0)$ и, значит, тем лучше сходимость. Обсуждаемая монотонность является строгой для регулярных областей. Так как K выпукло и содержит S, справедливо включение $E \subseteq K$. Если g_K обозначает функцию Грина для K и $E \ne K$, то $g(0) > g_K(0)$, что и выражает преимущество (24) над неадаптивной оценкой (23) ввиду свойств сходимости рядов Фабера.

Предложение 9. *Пусть $A = A^*$ и $0 \notin S$. Тогда*

$$\varlimsup_{m\to\infty} \min\left[\left\|r_{m-1}^A\right\|, \left\|r_m^A\right\|\right]^{1/m} \le e^{-g(0)}, \tag{25}$$

где g — обобщённая функция Грина спектра S.

Заметим, что в отличие от (24) оценка (25) гарантирует хорошее поведение FOM для последовательности значений m, имеющей положительную нижнюю натуральную плотность (≥ 0.5).

Отметим, что на практике метод Ланцоша используется для решения знакоопределённых и знаконеопределённых ССЛАУ в течение длительного времени.[61] [62] Однако при обосновании использования метода в случае знаконеопределённых операторов были трудности, связанные с проблемой устойчивости.[63] Утверждая, что по крайней мере один из каждых двух последовательных ланцошевских шагов «хороший», предложение 9 как раз обосновывает применение МСРЛ в этом случае (несмотря на возможные «выскоки»).

Определим «критический уровень функции Грина» $R_0 = \inf\Big\{ R > 1 \,\Big|\, K \subset G_R \Big\}$, где G_R — каноническая область для компакта E.

[60] J. L. Walsh, *Interpolation and approximation by rational functions in the complex domain*, AMS Colloquium Publications, **20**, Providence, RI, AMS, 1969: § 4.1.

[61] B. N. Parlett, *Linear Algebra and its Applic.*, **29** (1980), 323–346.

[62] H. D. Simon, *Math. Comp.*, **42** (1984), 115–142.

[63] C. C. Paige, B. N. Parlett, H. A. van der Vorst, *Numer. Linear Algebra with Applic.*, **2** (1995), 115–133.

Теорема 9. *Пусть $R > R_0$, а функция f аналитична в G_R и непрерывна на $\overline{G_R}$. Тогда ошибка МСРА удовлетворяет неравенству*

$$\varlimsup_{m\to\infty} \|u - u_m\|^{1/m} \le R^{-1}. \tag{26}$$

Если $E \ne K$, то (26) экспоненциально сильнее, чем (11) и (23).

Пусть 0 — изолированное простое собственное значение A, z — соответствующий собственный вектор, $\xi = A\varphi$, $\mathcal{G}$ — замыкание линейного подпространства $\operatorname{span}\{\xi, A\xi, A^2\xi, \ldots\}$, $B = A|_{\mathcal{G}} \colon \mathcal{G} \to \mathcal{G}$. В следующей теореме спектр S, множество E и функция g относятся к оператору B (не A).

Теорема 10. *Метод Арнольди, применённый к оператору A и вектору φ, на m-ом шаге выработает приближённую собственную пару (θ_m, y_m), такую что при подходящей нормализации y_m верны оценки*

$$\varliminf_{m\to\infty} |\theta_m|^{1/m} \le e^{-g(0)}, \qquad \varliminf_{m\to\infty} \|z - y_m\|^{1/m} \le e^{-g(0)}. \tag{27}$$

Заметим, что добавление конечного числа точек к спектру не меняет обобщённых изолиний[64] функции Грина. Поэтому здесь нет необходимости совместно рассматривать несколько собственных значений, в отличие от теоремы 7.

Одна и та же последовательность значений m (которая может быть редкой) даёт два нижних предела в заключениях (27) теоремы 10.

Теорема 11. *Пусть $A = A^*$ и 0 — изолированный элемент S (и, следовательно, собственное значение оператора A). Метод Ланцоша, применённый к паре (A, φ), за $m \ge 2$ шагов выработает такие пары Ритца (θ_m, y_m) с подходящим образом нормализованными векторами Ритца y_m, что справедливы оценки*

$$\varlimsup_{m\to\infty} |\theta_m|^{1/m} \le e^{-g(0)}, \qquad \varlimsup_{m\to\infty} \left[\min\left(|\theta_{m-1}|, |\theta_m|\right)\right]^{1/m} \le e^{-2g(0)},$$

$$\varlimsup_{m\to\infty} \left[\min\left(\|z - y_{m-1}\|, \|z - y_m\|\right)\right]^{1/m} \le e^{-g(0)},$$

где z — нормированный собственный вектор A, соответствующий собственному значению 0, g — обобщённая функция Грина для S и $g(0)$ следует понимать как существующий предел $g(0) \equiv \lim_{\lambda\to 0,\ \lambda\ne 0} g(\lambda) >$ 0.

[64] J. L. Walsh, Op. cit.

Если ограничиться главной, экспоненциальной частью оценок, то можно сказать, что теорема 11 обобщает теоремы Каниэля и Саада в следующих двух направлениях. Во-первых, интересные собственные значения не обязаны лежать на краю спектра или быть соседними. Во-вторых, дополнительные (т. е. не связанные с отделённостью желаемых собственных значений) зазоры в спектре ускоряют сходимость, что как раз и выражает адаптацию. Отметим, что рассуждение на тему адаптации оценок Каниэля и Саада к спектру имеется в уже упомянутой книге Парлетта (§ 12.4).

Глава 11 «Квадратура Гаусса–Арнольди для функции $\langle (zI - A)^{-1}\varphi, \varphi\rangle$ и Паде-подобная рациональная аппроксимация функций марковского типа» посвящена квадратуре Гаусса–Арнольди и связанных вопросам. Пусть A — ограниченный оператор в гильбертовом пространстве $\mathcal{H}$ и φ — нормированный вектор из $\mathcal{H}$. Функция $f_m(z) = \big\langle (zI - H)^{-1}e_1, e_1\big\rangle$ ($z \notin \mathrm{Sp}(H)$) является $[m-1/m]$-аппроксимантой типа Паде в бесконечности с полюсами в точках θ_i ($1 \le i \le m$) к функции $f(z) = \big\langle (zI - A)^{-1}\varphi, \varphi\big\rangle$ ($z \notin S = \mathrm{Sp}(A)$). Аппроксимация $f(z) \approx f_m(z)$ обсуждается в нескольких работах.[65] [66] Она является частным случаем квадратуры Гаусса–Арнольди. Пусть Ω — неограниченная связная компонента $\mathbb{C}\backslash S$, $E = \mathbb{C}\backslash\Omega \supseteq S$.

Если оператор A нормален, то он обладает спектральным разложением,[67] а пара (A, φ) — спектральной мерой μ с носителем S; имеем

$$f(z) = \int\limits_S \frac{d\mu(t)}{z - t}. \tag{28}$$

Поскольку в терминах обобщённой функции Грина g области Ω мы имеем

$$\varlimsup_{m\to\infty} \operatorname{dist}\big[(zI - A)^{-1}\varphi, \mathcal{K}^m(A, \varphi)\big]^{1/m} \le e^{-g(z)}, \qquad z \in \Omega,$$

то два следующих утверждения являются критериями нетривиальности квадратуры Гаусса–Арнольди (мы считаем квадратуру нетривиальной, если с ростом m её ошибка уменьшается существенно быстрее, чем величина $\|(zI - A)^{-1}\varphi - Q(zI - H)^{-1}e_1\|$).

Предложение 10. *Пусть оператор A нормален и спектральная мера μ пары (A, φ) регулярна.[68] Если существует такая точка $z \in \Omega$,*

[65] D. Calvetti, S.-M. Kim, L. Reichel, *SIAM J. Matrix Anal. Appl.*, **26** (2005), 765–781.

[66] R. W. Freund, M. Hochbruck, In: *Linear Algebra for Large Scale and Real-Time Applications*, M. S. Moonen et al. (eds), Kluwer Academic Press, Amsterdam, 1993, 377–380.

[67] У. Рудин, *Функциональный анализ*, М., Мир, 1975: гл. 12.

[68] H. Stahl, V. Totik, *General Orthogonal Polynomials*, Encyclopedia of Math. Appl., **43**, Cambridge, Cambridge Univ. Press, 1992: гл. 3.

что $\overline{\lim}_{m\to\infty} |f(z) - f_m(z)|^{1/m} < e^{-g(z)}$, *то у множества* E *и, следовательно, у спектра* S *нет внутренних точек.*

Теорема 12. *Если спектр* S *нормального оператора* A *является частью аналитической кривой без самопересечений и не разделяет комплексную плоскость, то справедливы оценки*

$$\overline{\lim_{m\to\infty}} |f(z) - f_m(z)|^{1/m} < e^{-g(z)}, \quad z \in \mathbb{C}\setminus \operatorname{Co}(S),$$
$$\underline{\lim_{m\to\infty}} |f(z) - f_m(z)|^{1/m} < e^{-g(z)}, \quad z \in \Omega = \mathbb{C}\setminus E,$$

где $\operatorname{Co}(S)$ *— выпуклая оболочка* S.

Продемонстрировано, что замена нормального оператора с нетривиальной квадратурой на подобный ему в алгебраическом смысле с помощью ограниченного и непрерывно обратимого оператора может привести к потере нетривиальности (т. е. в ненормальном случае оценку в терминах спектра дать нельзя).

Для марковской функции $\hat{\mu}(z) = \int \frac{d\mu(t)}{z-t}$, где μ — положительная мера с компактным носителем бесконечной мощности $S = \operatorname{Supp}\mu$, лежащим в $\mathbb{R}$, сходимость диагональных аппроксимант Паде гарантируется теоремой Маркова и её уточнениями;[69] [70] [71] см. также препринт автора[72] по поводу сходимости в лакунах спектра S по меньшей мере на каждом втором шаге ассоциированного процесса Ланцоша.

В работе Гончара[73] (см. также работы Рахманова[74] [75]) исследована сходимость поддиагональных аппроксимант Паде к функции $\hat{\mu} + r$, где $r \in \mathbb{C}(t)$, $r(\infty) = 0$, — «рациональное возмущение» $\hat{\mu}$, при условии $\mu'(t) > 0$ почти всюду на выпуклом замыкании S (из чего следует, что S — отрезок); даны оценки сходимости полюсов аппроксимант к полюсам r. В статье Рахманова[76] рассмотрен случай носителя меры — объединения нескольких отрезков вещественной прямой и вещественного рационального возмущения $r \in \mathbb{R}(t)$. В статье Гончара и С. П. Суетина[77] получены аналогичные результаты для комплекснозначных мер некоторых типов на вещественной прямой.

---

[69] A. Markoff, *Acta Math.*, **19** (1895), 93–104.
[70] Е. М. Никишин, В. Н. Сорокин, Op. cit.: гл. 2, § 6.
[71] H. Stahl, V. Totik, Op. cit.: гл. 6.
[72] L. Knizhnerman, Schlumberger-Doll Research, Res. Note EMG-001-95-12 (1995), 18 p.
[73] А. А. Гончар, *Матем. сб.*, **97 (139)** (1975), 607–629.
[74] Е. А. Рахманов, *Матем. сб.*, **103 (145)** (1977), 237–252.
[75] Е. А. Рахманов, *Матем. сб.*, **118 (160)** (1982), 104–117.
[76] Е. А. Рахманов, *Матем. сб.*, **104 (146)** (1977), 271–291.
[77] А. А. Гончар, С. П. Суетин, *Современные проблемы математики*, **5** (2004), 3–67.

В статье С. П. Суетина[78] отмечено, что если S не лежит на прямой, то предельные точки полюсов диагональных аппроксимант Паде могут быть всюду плотны в $\overline{\mathbb{C}}$. Эта и родственные трудности в теории аппроксимации Паде возникают из-за квази-, а не эрмитовой ортогональности знаменателей аппроксимант. Тот факт, что соответствующие квазиортогональные многочлены «хорошо себя ведут», приходится выводить из алгебраической специфики случая.[79] [80]

Мы представляем рациональное возмущение с простыми полюсами в виде, удобном для проведения процесса Арнольди. На рациональное возмущение налагается условие вещественности суммарного вычета и его положительности.

Предложение 11. *Пусть $s \geq 1$, $\lambda_1, \ldots, \lambda_s$ — попарно различные комплексные числа, $c_1, \ldots, c_s$ — ненулевые комплексные числа, такие что $0 < c_1 + \cdots + c_s \in \mathbb{R}$. Существует такая трёхдиагональная матрица $T \in \mathbb{C}^{s \times s}$, что $\left\langle (zI - T)^{-1} e_1, e_1 \right\rangle = \sum_{j=1}^{s} \frac{c_j}{z - \lambda_j}$, $z \in \mathbb{C} \backslash \{\lambda_1, \ldots, \lambda_s\}$.*

Оставшиеся утверждения этой главы показывают, с какой скоростью метод Арнольди локализует полюсы рационального возмущения $d^2 \sum_{j=1}^{s} \frac{c_j}{z - \lambda_j}$, $\sum_{j=1}^{s} c_j = 1$, $d > 0$, функции марковского типа (28). Мы приведём здесь одно из них.

Матрицу A_1 из условия предложения 12 можно найти с помощью предложения 11.

Предложение 12. *Пусть μ — положительная мера с компактным носителем $S \subset \mathbb{C}$, A_2 — оператор умножения на независимую переменную в $L_{2,\mu}(S)$, $K = \mathrm{Co}(S)$ — замыкание поля значений A_2, $s \geq 1$, числа $\lambda_1, \ldots, \lambda_s \in \Omega$ попарно различны, $c_1, \ldots, c_s$ — ненулевые комплексные числа, $c_1 + \cdots + c_s = 1$, $d > 0$, матрица $A_1 \in \mathbb{C}^{s \times s}$ такова, что*

$$\left\langle (zI - A_1)^{-1} e_1, e_1 \right\rangle = \sum_{j=1}^{s} \frac{c_j}{z - \lambda_j}, \qquad z \in \mathbb{C} \backslash \{\lambda_1, \ldots, \lambda_s\},$$

z_j — нормированный собственный вектор A_1, отвечающий собственному значению λ_j, $\mathcal{H} = \mathbb{C}^s \bigoplus L_{2,\mu}(S)$, $A = A_1 \bigoplus A_2$ — ортогональная прямая сумма пространств и соответствующая прямая сумма операторов. Процесс Арнольди с оператором A в $\mathcal{H}$ и начальным вектором

$$\varphi = \frac{1}{\sqrt{d^2 + \mu(S)}} \, (de_1, \mathbf{1})^{\tau} \in \mathcal{H},$$

[78] С. П. Суетин, *Успехи матем. наук*, **57** (2002), 45–142: пункт 3 введения.

[79] С. П. Суетин, *Матем. сб.* **191** (2000), 81–114.

[80] С. П. Суетин, *Матем. сб.* **194** (2003), 63–92.

где **1** *— постоянная функция «единица» из* $L_{2,\mu}(S)$, *за* $m > s$ *шагов выработает такие пары Ритца* (θ_j, y_j), $j = 1, \ldots, m$, *что при подходящей их нумерации верны следующие оценки. Если при данном* j $(1 \le j \le s)$ $\lambda_j \notin K$, *то*

$$\overline{\lim_{m\to\infty}} |\lambda_j - \theta_j|^{1/m} \le e^{-g(\lambda_j) - \min_{1\le k\le s} g(\lambda_k)}, \qquad \overline{\lim_{m\to\infty}} |z_j - y_j|^{1/m} \le e^{-g(\lambda_j)}.$$

В общем случае $(\lambda_j \notin E)$

$$\underline{\lim}_{m\to\infty} |\lambda_j - \theta_j|^{1/m} \le e^{-g(\lambda_j) - \min_{1\le k\le s} g(\lambda_k)}, \qquad \underline{\lim}_{m\to\infty} |z_j - y_j|^{1/m} \le e^{-g(\lambda_j)}. \quad (29)$$

Здесь g *— обобщённая функция Грина области* Ω *с особенностью в бесконечности.*

В гл. 12 (заключительной) сформулированы краткие выводы, перечислены нерешённые задачи, упомянуты подходы других авторов и высказаны благодарности.

Машинной арифметике посвящены главы 4–6 и частично 7. Результаты глав 10 и 11 применимы только к ограниченным операторам в бесконечномерных гильбертовых пространствах. Результаты, касающиеся машинной арифметики, применимы только к матрицам. Остальные результаты применимы как к операторам, так и к матрицам или классам матриц неограниченной размерности, причём пространство обычно можно считать комплексным, даже если приведённая формулировка вещественна.

Публикации автора по теме диссертации

[1] *Друскин В. Л., Книжнерман Л. А.* Использование операторных рядов по ортогональным многочленам при вычислении функций от самосопряжённых операторов и обоснование феномена Ланцоша. Деп. в ВИНИТИ. 1987. № 1535-В87. 47 с.

[2] *Друскин В. Л., Книжнерман Л. А.* Два полиномиальных метода вычисления функций от симметричных матриц // Ж. вычисл. матем. и матем. физ. 1989. Т. 29. № 12. С. 1763–1775.

[3] *Друскин В. Л., Книжнерман Л. А.* Оценки ошибок в простом процессе Ланцоша при вычислении функций от симметричных матриц и собственных значений // Ж. вычисл. матем. и матем. физ. 1991. Т. 31. № 7. С. 970–983.

[4] *Книжнерман Л. А.* Вычисление функций от несимметричных матриц с помощью метода Арнольди // Ж. вычисл. матем. и матем. физ. 1991. Т. 31. № 1. С. 5–16.

[5] *Книжнерман Л. А.* Оценки погрешности метода Арнольди: случай нормальной матрицы // Ж. вычисл. матем. и матем. физ. 1992. Т. 32. № 9. С. 1347–1360.

[6] *Druskin V., Knizhnerman L.* The Lanczos optimization of a splitting-up method to solve homogeneous evolutionary equations // J. Comput. Appl. Math. 1992. V. 42. № 2. P. 221–231.

[7] *Druskin V., Knizhnerman L.* Evaluation for Krylov subspace approximation to internal eigenvalues of large symmetric matrices and bounded self-adjoint operators with continuous spectrum // Schlumberger-Doll Research. 1992. Res. note EMG-001. 20 p.

[8] *Druskin V., Knizhnerman L.* Error bounds for Lanczos method to compute matrix-vector functions // Cornelius Lanczos International Centenary Conference. Conference Program. December 12–17, 1993. North Carolina State University, Raleigh, North Carolina. P. 85.

[9] *Knizhnerman L.* The quality of approximations to a separated eigenvalue and location of values" // Cornelius Lanczos International Centenary Conference. Conference Program. December 12–17, 1993. North Carolina State University, Raleigh, North Carolina. P. 90.

[10] *Druskin V., Knizhnerman L.* Krylov subspace approximation of eigenpairs and matrix functions in exact and computer arithmetic // Numer. Linear Algebra with Appl. 1995. V. 2. № 3. P. 205–217.

[11] *Книжнерман Л. А.* Качество аппроксимаций к хорошо отделённому собственному значению и расположение «чисел Ритца» в простом процессе Ланцоша // Ж. вычисл. матем. и матем. физ. 1995. Т. 35. № 10. С. 1459–1475.

[12] *Knizhnerman L.* On adaptation of the Lanczos method to the spectrum // Schlumberger-Doll Research. 1995. Res. Note EMG-001-95-12. 18 p.

[13] *Книжнерман Л. А.* Простой процесс Ланцоша: оценки погрешности гауссовой квадратурной формулы и их приложения // Ж. вычисл. матем. и матем. физ. 1996. Т. 36. № 11. С. 5–19.

[14] *Knizhnerman L.* On adaptation of the Arnoldi method to the spectrum // Schlumberger-Doll Research. 1996. Report EMG-001-96-03. 13 p.

[15] *Knizhnerman L.* On adaptation of the Lanczos and Arnoldi methods to the spectrum // Workshop on Iterative Methods and Parallel Computing, June 16-21, 1997, Milovy, Czech Republic. Abstracts.

[16] *Druskin V., Greenbaum A., Knizhnerman L.* Using nonorthogonal Lanczos vectors in the computation of matrix functions // SIAM J. Sci. Comp. 1998. V. 19. № 1. P. 38–54.

[17] *Greenbaum A., Druskin V. L., Knizhnerman L. A.* On solving indefinite symmetric linear systems by means of the Lanczos method // Ж. вычисл. матем. и матем. физ. 1999. Т. 39. № 3. С. 371–377.

[18] *Knizhnerman L.* Error bounds for the Arnoldi method: a set of extreme eigenpairs // Linear Algebra and its Applic. 1999. V. 296. P. 191–211.

[19] *Knizhnerman L.* Adaptation of the Lanczos and Arnoldi methods to the spectrum, or why the two Krylov subspace methods are powerful // Чебышёвский сборник. 2002. V. 3. № 2. P. 141–164.

[20] *Knizhnerman L.* The quality of the Gauss–Arnoldi quadrature for $\langle (zI - A)^{-1}\varphi, \varphi\rangle$ and application to Padé-like approximation, In: SIAM-GAMM ALA 2006, Dusseldorf. Abstracts of the Applied Linear Algebra Conference SIAM-GAMM, 2006.

[21] *Книжнерман Л. А.* Квадратура Гаусса–Арнольди для функции $\langle (zI - A)^{-1}\varphi, \varphi\rangle$ и Паде-подобная рациональная аппроксимация функций марковского типа // Матем. сб. 2008. Т. 199. № 2. С. 27–48.