ИПМ им.М.В.Келдыша РАН • Электронная библиотека

Препринты ИПМ • Препринт № 164 за 2018 г.

ISSN 2071-2898 (Print)
ISSN 2071-2901 (Online)

Оселедец И.В., Бочев М.А., Катруца А.М., Овчинников Г.В.

Как оптимизировать предобусловливатели в методе сопряжённых градиентов: стохастический подход

Рекомендуемая форма библиографической ссылки: Как оптимизировать предобусловливатели в методе сопряжённых градиентов: стохастический подход / И.В.Оселедец [и др.] // Препринты ИПМ им. М.В.Келдыша. 2018. № 164. 26 с. doi:10.20948/prepr-2018-164
URL: http://library.keldysh.ru/preprint.asp?id=2018-164

ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
имени М.В.КЕЛДЫША
Р о с с и й с к о й а к а д е м и и н а у к

И.В.Оселедец, М.А.Бочев, А.М.Катруца, Г.В.Овчинников

Как оптимизировать предобусловливатели в методе сопряжённых градиентов: стохастический подход

Москва — 2018

Оселедец И.В., Бочев М.А., Катруца А.М., Овчинников Г.В.

Как оптимизировать предобусловливатели в методе сопряжённых градиентов: стохастический подход

Метод сопряжённых градиентов (conjugate gradient method, CG) обычно используют с предобусловливанием, позволяющим улучшить эффективность и надёжность метода. Многие предобусловливатели включают параметры, выбор которых зачастую является нетривиальной задачей. Существует немало оценок сходимости, на основе которых можно оптимизировать параметры предобусловливателя. Однако, эти оценки обычно выполняются для всех векторов начальных приближений, другими словами, они отражают наихудшую скорость сходимости. Чтобы отследить среднюю скорость сходимости, в этой работе предлагается простой стохастический подход. Он основан на выполнении серии пробных запусков метода сопряжённых градиентов со случайными векторами начальных приближений и даёт функционал, который можно использовать для оптимизации параметров предобусловливателя в методе сопряжённых градиентов. Представлены численные эксперименты, показывающие что оптимизация данного функционала обычно даёт лучшие значения параметров предобусловливателя, чем оптимизация на основе спектрального числа обусловленности.

Ключевые слова: метод сопряжённых градиентов, предобусловливание, число обусловленности, кластеры собственных чисел, неполное разложение Холесского

I.V. Oseledets, M.A. Botchev, A. Katrutsa and G.V. Ovchinnikov

How to optimize preconditioners for the conjugate gradient method: a stochastic approach

The conjugate gradient method (CG) is usually used with a preconditioner which improves efficiency and robustness of the method. Many preconditioners include parameters and a proper choice of a preconditioner and its parameters is often not a trivial task. Although many convergence estimates exist which can be used for optimizing preconditioners, they typically hold for all initial guess vectors, reflecting the worst convergence rate. To account for the mean convergence rate instead, in this paper, we follow a simple stochastic approach. It is based on trial runs with random initial guess vectors and leads to a functional which can be used to monitor convergence and to optimize preconditioner parameters in CG. Presented numerical experiments show that optimization of this new functional usually yields a better parameter value than optimization of the functional based on the spectral condition number.

Key words: conjugate gradient method, preconditioners, condition number, eigenvalue clustering, relaxed incomplete Cholesky preconditioner

Работа выполнена при поддержке гранта РФФИ 17-01-00854-а.

1. Введение

Предобусловливание — важный инструмент улучшения сходимости итерационных методов решения систем линейных алгебраических уравнений [1, 2, 3]. Эффективные предобусловливатели не столько улучшают число обусловленности матрицы системы, сколько, что ещё более важно, часто приводят к кластерному расположению её собственных значений. В нестационарных методах, таких как метод сопряжённых градиентов (conjugate gradient method, CG) [4, 5] или обобщённый метод минимальных невязок (generalized minimal residual method, GMRES) [2, 6], подобное повышение разреженности (кластерности) обычно проявляется в так называемой «суперлинейной» сходимости [7, 8].

Многие предобусловливатели включают в себя определённые параметры, и, хотя основные классы предобусловливателей хорошо изучены [1, 2, 3], практический выбор параметров предобусловливателей часто представляет собой нетривиальную задачу. Для оптимизации параметров предобусловливателей на практике могут быть использованы различные функционалы (целевые функции) [9], такие как спектральный радиус [10], числа Ритца [11, Chapter 8.4] предобусловленной матрицы, так называемое число K-обусловленности [12, 13], подходящая норма матрицы перехода [14, 15], близость предобусловленной матрицы к единичной матрице во фробениусовой норме [9] и след предобусловленной матрицы [16]. Все эти функционалы характеризуют сходимость предобусловленного итерационного метода и имеют одно общее свойство: они базируются на определённых оценках сходимости, которые справедливы для *всех* векторов начального приближения. В этом смысле они отражают наихудший сценарий сходимости, и естественно возникает вопрос, насколько адекватны эти функционалы для выбора параметров предобусловливателей на практике. Не будет ли лучше, например, отслеживать *среднюю* скорость сходимости, нежели *наихудшую* скорость сходимости?

В этой работе предлагается попытка ответа на этот вопрос. Несложный анализ сходимости метода сопряжённых градиентов показывает, что имеется непустое открытое множество векторов начального приближения, для которых наблюдается повышенная скорость сходимости метода. Это говорит о том, что средняя скорость сходимости действительно может являться более адекватной мерой сходимости, чем наихудшая скорость. Основываясь на этом наблюдении, мы предлагаем стохастический подход оптимизации параметров предобусловливателей. В данном подходе, основанном на тестовых прогонах метода со случайными векторами начального приближения, предлагается использовать новый функционал оптимизации. Функционал может быть использован для отслеживания сходимости и оптимизации параметров предобусловливателей в методе сопряжённых градиентов. Представленные численные эксперименты показывают, что оптимизация параметров предобусловливателей в методе сопря-

жённых градиентов по отношению к данному функционалу обычно даёт лучшие значения параметров, чем оптимизация по отношению к функционалу, основанному на спектральном числе обусловленности.

Следует подчеркнуть, что предложенный подход в настоящее время представляет ограниченный практический интерес, если решается только одна система уравнений. Дело в том, что предложенная процедура оптимизации основана на тестовых прогонах со значительным числом векторов начального приближения, а это означает существенные вычислительные затраты. Однако бывают ситуации, когда должно решаться большое количество линейных систем с одной и той же матрицей — например, в неявных схемах интегрирования по времени. В таких ситуациях оптимизация предобусловливателя может привести к значительному улучшению сходимости, и наш подход представляет практический интерес.

Заметим, что решение большого числа линейных систем с одной и той же матрицей является направлением активных исследований (из последних работ укажем [17, 18]). Вклад нашего стохастического оптимизационного подхода здесь состоит в том, что он органично может быть использован в сочетании с методами решения большого числа линейных систем с одной и той же матрицей и привести к дальнейшему снижению вычислительных затрат.

В данной работе предполагаем, что линейные системы вида

$$Ax = b \tag{1}$$

должны решаться для симметричной положительно определённой матрицы $A \in \mathbb{R}^{m\times m}$ и многих векторов правой части $b \in \mathbb{R}^m$. Данная статья организована следующим образом. В разделе 2 рассмотрена стохастическая мера сходимости (функционал сходимости) для стационарных линейных итерационных методов, т.е., для итераций вида

$$Mx_{k+1} = Nx_k + b, \tag{2}$$

где $M - N = A$ и M — невырожденная матрица. В этом же разделе также вводится подобный функционал сходимости для нестационарных нелинейных итерационных методов, таких как метод сопряжённых градиентов. Вопрос, представляет ли этот функционал новую меру сходимости, отличную от классической оценки сходимости, рассмотрен в разделе 3. Здесь мы показываем, что существует открытое множество векторов начального приближения, для которых метод сопряжённых градиентов сходится быстрее, чем показывает классическая оценка сходимости. Это показывает, что предложенный функционал сходимости отличен от классического. Численные эксперименты представлены в разделе 5. Здесь показано, как наша техника оптимизации может быть использована для нахождения наилучшего значения релаксационного параметра α в

широко известном предобусловливателе неполного разложения Холесского без заполнения [19, 3, 20]. Этот предобусловливатель обозначим $\mathrm{RIC}_\alpha(0)$. Выводы работы изложены в разделе 6.

2. Средняя оценка сходимости

Итерационные методы хорошо изучены [1, 4, 21, 22, 2, 3], и известны классические оценки их сходимости. Оценка сходимости обычно имеет вид

$$\|x_* - x_k\|_* \leq C q_A^k \|x_* - x_0\|_*, \tag{3}$$

где x_* — точное решение системы (1), x_k — приближение решения на k-ой итерации, $\|\cdot\|_*$ некоторая векторная норма, $C > 0$ некоторая константа и $q_A > 0$ константа, зависящая от матрицы A. Оценка (3) является наихудшей оценкой среди *всех* начальных приближений x_0, хотя естественно вместо наихудшей оценки рассмотреть *среднюю оценку сходимости* итерационного метода. В данной работе мы исследуем такой подход, который, насколько мы знаем, ранее не был изучен. Рассмотрим вектор начальной ошибки $x_* - x_0$, который является *случайным вектором* с независимыми и одинаково распределёнными элементами, среднее которых равно 0, а стандартное отклонение равно 1, сгенерированными, например, из стандартного нормального распределения $N(0, 1)$. Тогда вектор ошибки $x_* - x_k$ — также случайный вектор, и мы можем определить его математическое ожидание e_k как

$$e_k^2 = \mathbb{E}(\|x_* - x_k\|_*^2). \tag{4}$$

Возникает вопрос о возможности получения для e_k оценки вида

$$e_k \sim C\mu_A^k, \tag{5}$$

где $\mu_A > 0$ — константа, зависящая от A и определяющая скорость сходимости. Необходимо аккуратно определять *асимптотическое* поведение метода в (4), поскольку некоторые методы (например, метод сопряжённых градиентов) в точной арифметике сходятся за m итераций. Тем не менее для больших m оценки вида (4) интересны и дают полезную информацию о сходимости метода.

Сначала рассмотрим важный частный случай — стационарный линейный итерационный метод (2),

$$x_* - x_{k+1} = G(x_* - x_k), \qquad G = M^{-1}N,$$

где G — это матрица итераций и $M - N = A$. Используя классический результат Хатчинсона о стохастической оценке следа матрицы [23] и обозначив $d_k = x_* - x_k$, $k \geqslant 0$, мы получаем в евклидовой норме оценку

$$e_k^2 = \mathbb{E}\left(G^k d_0, G^k d_0\right) = \mathbb{E}\left((G^k)^\top G^k d_0, d_0\right) = \mathrm{Tr}((G^k)^\top G^k) = \|G^k\|_F^2,$$

где $\|\cdot\|_F$ — фробениусова норма матрицы, а Tr обозначает след матрицы. Используя формулу Гельфанда

$$\lim_{k\to\infty} \|G^k\|_*^{1/k} = \rho(G),$$

где $\|\cdot\|_*$ — любая норма* и $\rho(G)$ — спектральный радиус матрицы G, получаем следующую оценку

$$e_k \sim \rho(G)^k. \tag{6}$$

Таким образом, для рассматриваемого случая стационарного линейного итерационного метода наихудшая оценка сходимости одновременно является и оценкой сходимости в среднем [24].

Для нелинейных нестационарных итерационных методов, таких как метод сопряжённых градиентов (CG) или метод минимальных невязок (MINRES), аналогичный анализ является более сложным и выходит за рамки данной работы. Тем не менее, мы экспериментально, с помощью метода Монте-Карло и эмпирической оценки скорости сходимости, получили, что поведение таких методов существенно отличается от поведения линейных стационарных методов, то есть наихудшая оценка сходимости существенно хуже (завышена), чем приближённая средняя оценка сходимости. Поскольку аналитическое выражение для средней оценки сходимости отсутствует, мы получим вычислимую меру сходимости, аналогичную (6). Как k итераций стационарного линейного метода (2) эквивалентно выполнению k умножений матрицы итераций G на вектор, так и k итераций нестационарного итерационного метода можно представить как действие некоторого нелинейного преобразования $\mathcal{G}$, определяющего метод, k раз†. Таким образом, можно записать

$$x_* - x_k = \mathcal{G}[x_* - x_{k-1}] = \mathcal{G}^k[x_* - x_0], \tag{7}$$

где $\mathcal{G}$ — это отображение, эквивалентное применению одной итерации нелинейного метода, а $\mathcal{G}^k$ обозначает применение этого отображения k раз. Тогда практический способ наблюдения сходимости метода — это выполнение k итераций для некоторого числа случайных начальных приближений $x_0^{(i)}$. Действительно, определим стохастический функционал сходимости как

$$F_s \equiv \frac{1}{n}\sum_{i=1}^{n} \|x_* - x_k^{(i)}\|_*, \tag{8}$$

*Норма не обязательно должна быть операторной нормой (т.е. быть порождённой векторной нормой), но если норма операторная, то предел достигается сверху.

†Для метода сопряжённых градиентов, отображение $\mathcal{G}$ зависит от векторов приближённого решения на двух последних итерациях x_k и x_{k-1} и двух соответствующих A-сопряжённых направлений p_k и p_{k-1}.

где n — число случайных начальных приближений и $x_k^{(i)}$ — приближение решения на k-ой итерации с начальным приближением $x_0^{(i)}$. Возникает вопрос: как вычислять функционал F_s, если точное решение x_*, как правило, неизвестно на практике. Для метода сопряжённых градиентов эту трудность легко обойти, если вычислять F_s для A-нормы. Можно также воспользоваться следующим подходом, работающим для любых итерационных решателей: при вычислении F_s считать точное решение x_* равным нулю, что соответствует решению линейной системы с нулевой правой частью $b = 0$. Везде далее в данной работе F_s вычислялся для $x_* = 0$ и евклидовой нормы. Заметим, что по аналогии с (4) мы можем записать

$$F_s = \frac{1}{n}\sum_{i=1}^{n} \|x_* - x_k^{(i)}\|_* = \frac{1}{n}\sum_{i=1}^{n} \|\mathcal{G}^k[x_* - x_0^{(i)}]\|_* \approx \mathbb{E}_{x_0}\|\mathcal{G}^k[x_* - x_0]\|_*.$$

Возможное применение данного подхода, рассматриваемое в данной работе, заключается в оптимизации параметра предобусловливателя в методе сопряжённых градиентов для получения более быстрой сходимости. Очевидно, что такой оптимизационный процесс, основанный на многократных пробных запусках итерационного метода, является вычислительно затратным (один пробный запуск означает выполнение nk итераций метода). Такие затраты оправданы, только если полученный оптимальный предобусловливатель будет использован для проведения большого числа итераций, например, при решении различных систем с одинаковой матрицей и различными правыми частями. Именно поэтому мы сделали такое предположение при постановке задачи решения системы (1).

Поскольку для стационарного линейного итерационного метода стохастический функционал сходимости оказался эквивалентен классической мере сходимости (то есть, спектральному радиусу матрицы итераций), возникает вопрос совпадает ли стохастический функционал сходимости F_s с некоторой известной мерой сходимости для нестационарного итерационного метода. В следующем разделе мы рассмотрим этот вопрос для метода сопряжённых градиентов. Хорошо изученный классический функционал сходимости для метода сопряжённых градиентов (см., например, [1, 4, 22, 2, 3]) основан на числе обусловленности κ матрицы системы A:

$$\|x_* - x_k\|_A \leqslant 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^k \|x_* - x_0\|_A. \tag{9}$$

Эта оценка в общем случае может быть пессимистичной, поскольку она не отражает часто наблюдаемую сверхлинейную скорость сходимости метода сопряжённых градиентов [7, 19]. Тем не менее она может быть использована для

получения скорости сходимости метода сопряжённых градиентов. Таким образом, вместе с (8) мы рассматриваем соответствующий классический функционал сходимости

$$F_c \equiv \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k. \tag{10}$$

Заметим, что оценка (9) может быть улучшена в предположении о кластеризации спектра матрицы A [19, 2, 7]. В следующем разделе мы покажем, что даже в отсутствие предположений о кластеризации спектра существует открытое множество начальных векторов, для которых метод сопряжённых градиентов показывает более быструю сходимость, чем та, которую даёт классическая оценка (9). Это означает, что предлагаемый стохастический функционал (8) существенным образом отличается от классического функционала (10).

3. Вектор начального приближения и сходимость сопряжённых градиентов

Анализ, приведённый в этом разделе, основан на подходах и результатах работы [7]. Пусть A — симметричная положительно определённая $m \times m$ матрица, $z_1, \ldots, z_m$ — ортонормальные собственные вектора A, а $0 < \lambda_1 \leqslant \cdots \leqslant \lambda_m$ — соответствующие им собственные числа. Для простоты обозначений в этом разделе мы опускаем индекс $\cdot_*$ в обозначении точного решения x_*. Для итерационных приближений x_j, $j = 1,2,\ldots$, метода сопряжённых градиентов выполняется свойство оптимальности:

$$\|x - x_j\|_A = \min_{q \in \Pi_j^{\mathrm{ct1}}} \|q(A)(x - x_0)\|_A, \tag{11}$$

где Π_j^{ct1} — множество всех многочленов степени не выше j со свободным членом, равным единице 1. Пусть q — невязочный многочлен метода сопряжённых градиентов, т.е. многочлен, для которого достигается минимум в соотношении (11), и пусть

$$x - x_0 = \sum_{i=1}^{m} \gamma_i z_i.$$

Легко проверить, что свойство оптимальности (11) можно записать в виде

$$\|x - x_j\|_A^2 = \sum_{i=1}^{m} \lambda_i (\gamma_i q(\lambda_i))^2 \leqslant \sum_{i=1}^{m} \lambda_i (\gamma_i \tilde{q}(\lambda_i))^2, \qquad \forall \tilde{q} \in \Pi_j^{\mathrm{ct1}}. \tag{12}$$

Кроме того,

$$q(t) = \frac{(\theta_1 - t) \ldots (\theta_j - t)}{\theta_1 \ldots \theta_j}, \tag{13}$$

где корни $\theta_1, \ldots, \theta_j$ многочлена $q(t)$ — числа Ритца процесса сопряжённых градиентов.

Рассмотрим теперь такой вектор начального приближения $\bar{x}_0$, что

$$x - \bar{x}_0 = \sum_{i=2}^{m} \gamma_i z_i, \tag{14}$$

и обозначим $\bar{q}(t)$ невязочный многочлен процесса сопряжённых градиентов, начатого на начальном приближении $\bar{x}_0$. Аналогично соотношению (13) для $\bar{q}(t)$ выполняется

$$\bar{q}(t) = \frac{(\bar{\theta}_1 - t) \ldots (\bar{\theta}_j - t)}{\bar{\theta}_1 \ldots \bar{\theta}_j},$$

где $\bar{\theta}_1, \ldots, \bar{\theta}_j$ — числа Ритца процесса сопряжённых градиентов, начатого на начальном приближении $\bar{x}_0$. Заметим, что если в соотношении (12) выбрать в качестве $\tilde{q}(t)$ чебышёвский минимаксный полином на интервале $[\lambda_1,\lambda_m]$, то получим классическую оценку сходимости

$$\begin{aligned} \|x - x_j\|_A^2 &\leqslant \sum_{i=1}^{m} \lambda_i (\gamma_i \tilde{q}(\lambda_i))^2 \leqslant \\ &\leqslant \max_i \tilde{q}(\lambda_i)^2 \sum_{i=1}^{m} \lambda_i \gamma_i^2 \leqslant \max_{\lambda \in [\lambda_1, \lambda_m]} \tilde{q}(\lambda)^2 \sum_{i=1}^{m} \lambda_i \gamma_i^2 = \\ &= \max_{\lambda \in [\lambda_1, \lambda_m]} \tilde{q}(\lambda)^2 \cdot \|x - x_0\|_A^2 = 4C_1^{2j} \|x - x_0\|_A^2, \end{aligned} \tag{15}$$

где

$$C_1 = \frac{\sqrt{\kappa_1} - 1}{\sqrt{\kappa_1} + 1}, \qquad \kappa_1 = \frac{\lambda_m}{\lambda_1}.$$

Для процесса сопряжённых градиентов, начатого на начальном приближении $\bar{x}_0$, соответствующая оценка сходимости имеет вид

$$\|x - \bar{x}_j\|_A^2 \leqslant 4C_2^{2j} \|x - \bar{x}_0\|_A^2, \qquad C_2 = \frac{\sqrt{\kappa_2} - 1}{\sqrt{\kappa_2} + 1}, \qquad \kappa_2 = \frac{\lambda_m}{\lambda_2}. \tag{16}$$

Теорема 1. *Пусть вектор начального приближения x_0 в итерационном процессе сопряженных градиентов выбран так, что первая компонента γ_1 начальной ошибки $x - x_0$ мала по отношению к остальным компонентам, то есть пусть существует такая константа $\delta > 0$, что*

$$\frac{\lambda_1 \gamma_1^2}{\sum_{i=2}^{m} \lambda_i \gamma_i^2} = \frac{\lambda_1 \gamma_1^2}{\|x - \bar{x}_0\|_A^2} \leqslant 4C_2^{2j} \delta, \quad j = 1, \ldots, J, \tag{17}$$

где x_0 определён в уравнении (14)*. Тогда сходимость итерационного процесса сопряженных градиентов на первых J итерациях определяется константой C_2, а не константой C_1 (см. соотношения* (15),(16)*) в том смысле, что*

$$\|x - x_j\|_A^2 \leqslant 4(1+\delta)C_2^{2j}\|x - x_0\|_A^2, \quad \textit{for } j = 1, \ldots, J. \tag{18}$$

Доказательство. Выбирая в (12) в качестве полинома $\tilde{q}(t)$ невязочный полином $\bar{q}(t)$ итерационного процесса сопряжённых градиентов, начатого в $\bar{x}_0$, получаем

$$\begin{aligned}
\|x - x_j\|_A^2 &= \sum_{i=1}^{m} \lambda_i(\gamma_i q(\lambda_i))^2 \leqslant \sum_{i=1}^{m} \lambda_i(\gamma_i \bar{q}(\lambda_i))^2 = \\
&= \lambda_1(\gamma_1 \bar{q}(\lambda_1))^2 + \sum_{i=2}^{m} \lambda_i(\gamma_i \bar{q}(\lambda_i))^2 = \\
&= \lambda_1(\gamma_1 \bar{q}(\lambda_1))^2 + \|x - \bar{x}_j\|_A^2 \leqslant \lambda_1(\gamma_1 \bar{q}(\lambda_1))^2 + 4C_2^{2j}\|x - \bar{x}_0\|_A^2 \leqslant \\
&\leqslant \lambda_1\gamma_1^2 + 4C_2^{2j}\|x - \bar{x}_0\|_A^2,
\end{aligned}$$

где последняя оценка выполняется, поскольку

$$0 \leqslant \bar{q}(t) \leqslant 1.$$

Данное неравенство справедливо потому, что $\bar{q}(t)$ — монотонно невозрастающая функция на интервале $[0,\lambda_2]$ и $\bar{q}(0) = 1$. Используя предположение (17), получаем

$$\begin{aligned}
\|x - x_j\|_A^2 &\leqslant \lambda_1\gamma_1^2 + 4C_2^{2j}\|x - \bar{x}_0\|_A^2 \leqslant \\
&\leqslant \delta 4C_2^{2j}\|x - \bar{x}_0\|_A^2 + 4C_2^{2j}\|x - \bar{x}_0\|_A^2 \leqslant 4(1+\delta)C_2^{2j}\|x - x_0\|_A^2,
\end{aligned}$$

что завершает доказательство. □

Нетрудно заметить, что последнюю теорему можно обобщить на случай, когда несколько первых компонентов начальной ошибки малы по отношению к её остальным компонентам. Действительно, обозначим

$$C_s = \frac{\sqrt{\kappa_s} - 1}{\sqrt{\kappa_s} + 1}, \qquad \kappa_s = \frac{\lambda_m}{\lambda_s}. \tag{19}$$

Тогда выполняется следующий результат.

Теорема 2. *Пусть вектор начального приближения x_0 в итерационном процессе сопряженных градиентов выбран так, что первые $s-1$ компонент γ_1, ...,*

γ_{s-1} начальной ошибки $x - x_0$ малы по отношению к её остальным компонентам, то есть пусть существует такая константа $\delta > 0$, что

$$\frac{\sum_{i=1}^{s-1} \lambda_i \gamma_i^2}{\sum_{i=s}^{m} \lambda_i \gamma_i^2} = \frac{\sum_{i=1}^{s-1} \lambda_i \gamma_i^2}{\|x - \bar{x}_0\|_A^2} \leqslant 4C_s^{2j}\delta, \quad \textit{for } j = 1, \ldots, J, \tag{20}$$

где

$$x - \bar{x}_0 = \sum_{i=s}^{m} \gamma_i z_i, \tag{21}$$

Тогда сходимость итерационного процесса сопряженных градиентов на первых J итерациях определяется константой C_s, а не константой C_1 (см. соотношение (15)*) в том смысле, что*

$$\|x - x_j\|_A^2 \leqslant 4(1 + \delta)C_s^{2j}\|x - x_0\|_A^2, \quad j = 1, \ldots, J. \tag{22}$$

Доказательство. Доказательство повторяет доказательство Теоремы 1. В соотношении (12) выбираем в качестве $\tilde{q}(t)$ невязочный полином $\bar{q}(t)$ процесса сопряжённых градиентов, начатого на векторе $\bar{x}_0$, определённом в (21). Используя оценку сходимости

$$\|x - \bar{x}_j\|_A^2 \leqslant 4C_s^{2j}\|x - \bar{x}_0\|_A^2,$$

выполняющуюся для этого процесса, и предположение (20), получаем (22). □

Теоремы 1, 2 можно проиллюстрировать следующим численным примером. Пусть A — диагональная матрица размерности $m = 1000$, с диагональными элементами

$$1, 2, 3, \ldots, 1000.$$

Выберем, кроме того, вектор правой части b так, чтобы все компоненты вектора точного решения были равны единице, а вектор начального приближения x_0 — так, чтобы все компоненты $x - x_0$ кроме первой — независимые и одинаково распределённые величины класса $N(0, 1)$ (стандартного нормального распределения). Пусть первая компонента $x - x_0$ равняется γ_1.

На рис. 1 представлены ошибка процесса сопряжённых градиентов, оценки сходимости Чебышёва (15), (16) и величины $\gamma_1 q(\lambda_1)$, $\gamma_1 \bar{q}(\lambda_1)$. Как видим, на первых итерациях (примерно до итерации 75 для $\gamma_1 = 0.05$ и до итерации 25 для $\gamma_1 = 5$) величины $\gamma_1 q(\lambda_1)$ и $\gamma_1 \bar{q}(\lambda_1)$ практически одинаковы и не меняются. Это означает, что процесс сопряжённых градиентов сходится так, как если бы первая компонента ошибки равнялась нулю. Как ясно видно на первом графике

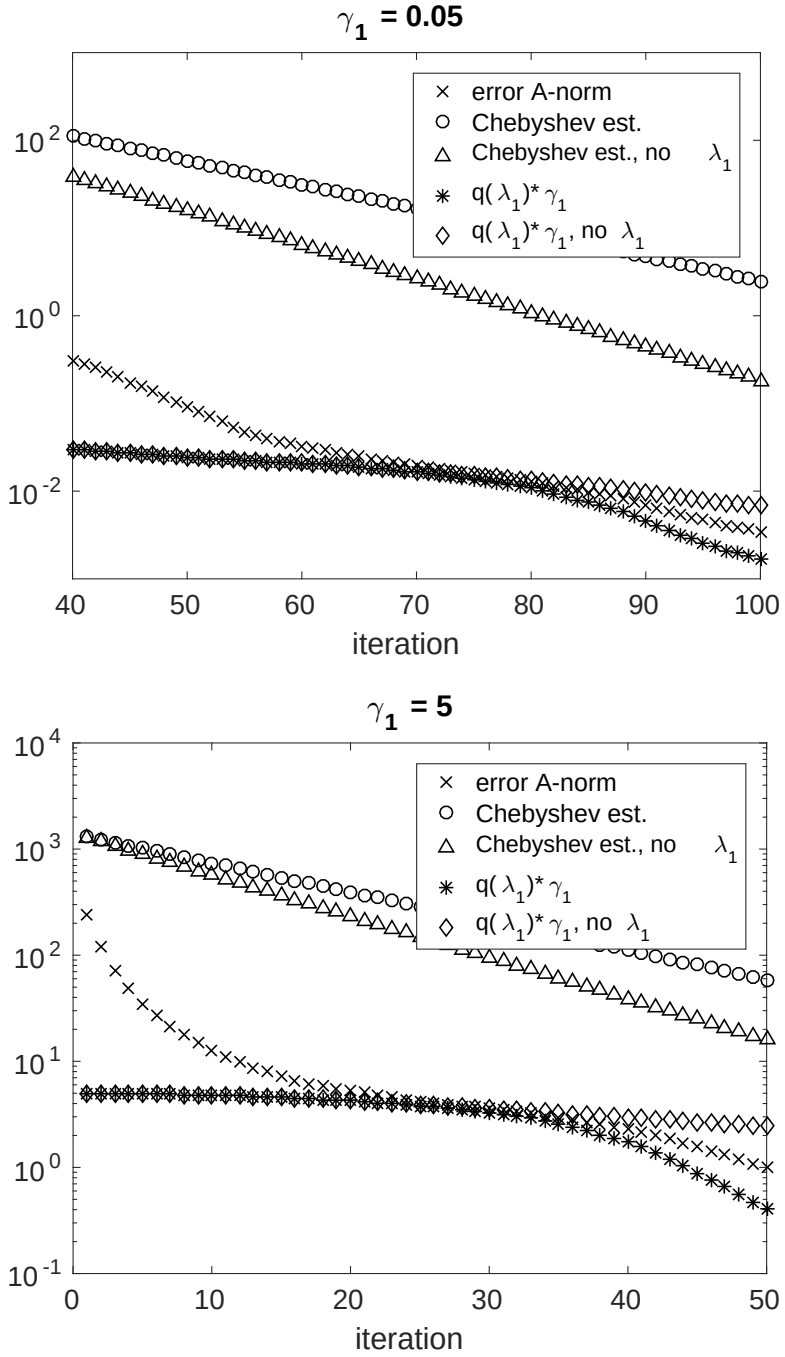


Рис. 1. Сходимость сопряжённых градиентов для векторов начального приближения с $\gamma_1 = 0.05$ (верхний график) и $\gamma_1 = 5$ (нижний график): норма ошибки $\|x - x_j\|_A$ (кривая $\times$), оценка $2C_1^j$ (кривая $\circ$), оценка $2C_2^j$ (кривая $\triangle$), величины $\gamma_1 q(\lambda_1)$ и $\gamma_1 \bar{q}(\lambda_1)$ (кривые $*$ и $\diamond$ соответственно).

рис. 1, примерно до итерации 75 скорость убывания A-нормы ошибки (кривая $\times$) соответствует улучшенной оценке Чебышёва (кривая $\triangle$), что подтверждает оценку (18). Первая компонента ошибки игнорируется итерационным процессом до тех пор, пока она не становится сравнимой по величине со всей нормой ошибки (т.е. пока кривая $\times$ не пересечёт кривую $*$). Начиная с этого момента,

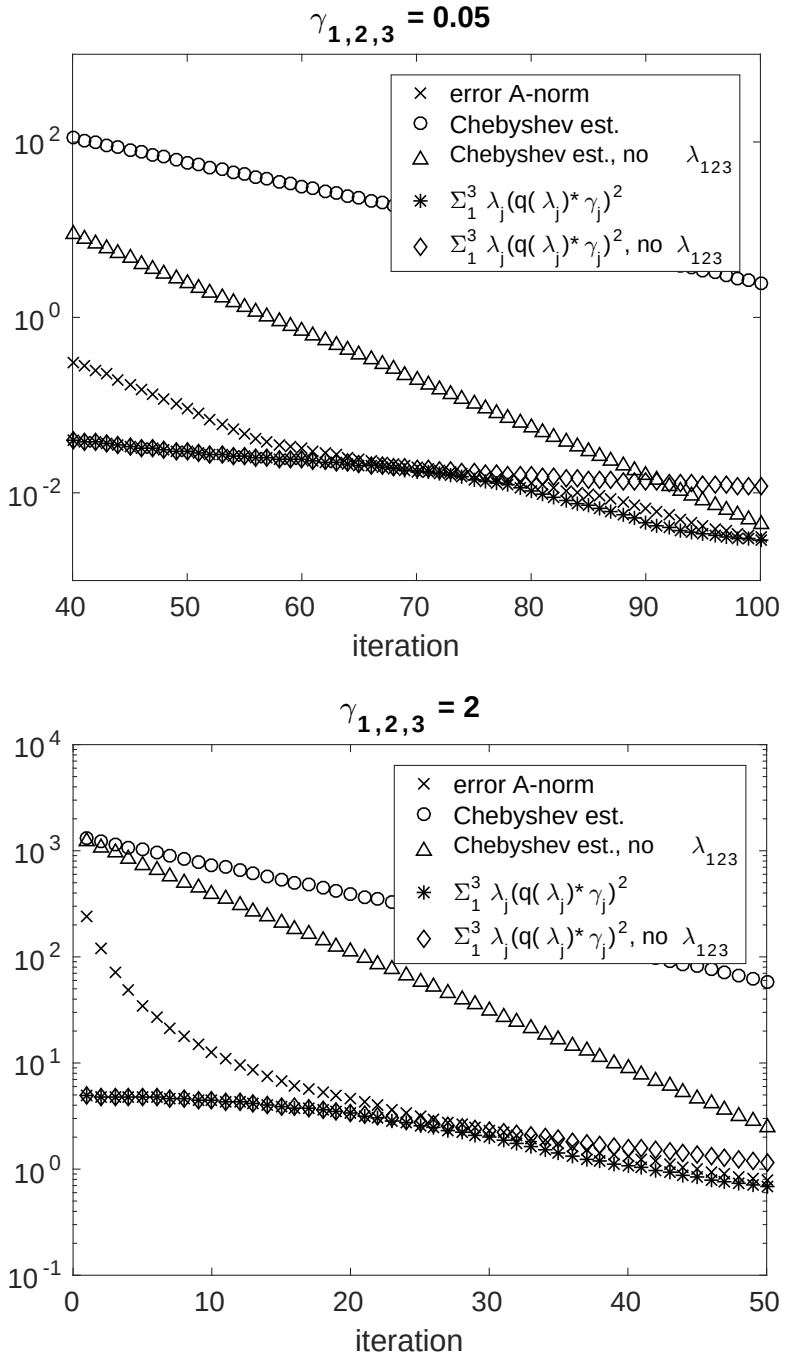


Рис. 2. Сходимость сопряжённых градиентов для векторов начального приближения с $\gamma_1 = \gamma_2 = \gamma_3 = 0.05$ (верхний график) и $\gamma_1 = \gamma_2 = \gamma_3 = 2$ (нижний график): норма ошибки $\|x - x_j\|_A$ (кривая $\times$), оценка $2C_1^j$ (кривая $\circ$), оценка $2C_2^j$ (кривая $\triangle$), величины $\gamma_1 q(\lambda_1)$ и $\gamma_1 \bar{q}(\lambda_1)$ (кривые $*$ и $\diamond$ соответственно).

$\gamma_1 q(\lambda_1)$ начинает убывать, гася первую компоненту ошибки. Величина $\gamma_1 \bar{q}(\lambda_1)$ остаётся практически неизменной и с определённого момента превосходит норму ошибки. Отметим, что в точках пересечения кривых $\times$ и $*$ (итерация 75 для $\gamma_1 = 0.05$ и итерация 25 для $\gamma_1 = 5$) оценка (17) выполняется в обоих случаях для примерно одного и того же значения δ, для $\delta < 10^{-3}$. Таким образом, эту величину можно считать «пороговой»: всё, что меньше этой величины, процесс сопряжённых градиентов «считает» достаточно малым.

Результаты аналогичного теста представлены на рис. 2. Для этого теста значения параметров берутся теми же, за исключением того, что теперь три первые компоненты ошибки γ_1, γ_2, γ_3 получают заданные значения (0.05 и 2). В соответствии с этим, сравнительный процесс сопряжённых градиентов с полиномом $\bar{q}(t)$ (с кривой сходимости $\diamond$) начат теперь в x_0, определённом, следуя соотношению (21) с $s = 4$.

Замечание 1. *Заметим, что наши результаты сходимости можно рассматривать как дополнение к классическим оценкам сходимости Ван дер Слюйса и Ван дер Ворста [7]. Действительно, наши результаты описывают возможное поведение сходимости сопряжённых градиентов в начальной фазе процесса, т.е., до того как определённые компоненты ошибки «задавлены» и сопряжённые градиенты начинают сходиться «суперлинейно». Эта суперлинейная фаза сходимости видна в кривых $\times$ на рис. 1 и 2 после того, как они пересекают кривые $\diamond$.*

4. Вычислительная сложность процедуры оптимизации

Для оптимизации параметра предобусловливателя в соответствии с функционалом сходимости F_s мы используем метод Брента [25]. Этот метод одномерной оптимизации, использующий метод золотого сечения в сочетании с обратно-квадратичной аппроксимацией целевой функции, требует одного вычисления целевой функции на каждой итерации. Таким образом, суммарную сложность выполнения оптимизации можно оценить как Kns умножений предобусловленной матрицы на вектор, где K — это количество итераций предобусловленного метода сопряжённых градиентов (выбор этого параметра обсуждается ниже в разделе 5.2), n — это число случайных начальных приближений, s — это количество итераций в методе Брента, необходимое для получения оптимального значения параметра с заданной точностью.

Предложенная процедура оптимизации параметра предобусловливателя эффективнее поиска по сетке, если количество итераций метода Брента s существенно меньше количества точек в сетке. В нашем случае, мы ограничили пространства поиска до отрезка $\alpha \in [0.9,1]$, в котором обычно [26],[3, раздел 13.2.1] лежит оптимальный параметр предобусловливателя $\mathrm{RIC}_\alpha(0)$. В численных экспериментах мы обычно наблюдали сходимость метода Брента не более чем за

25 итераций, в то время как поиск по сетке с такой же точностью требовал более 100 тестовых точек в сетке.

Для оптимизации классического функционала сходимости F_c в численных экспериментах (раздел 5) также использовался метод Брента. Для вычисления числа обусловленности в F_c мы использовали стандартную функцию поиска собственных значений разреженных матриц из библиотеки NumPy на языке Python (эта функция совпадает с командой `eigs` в Octave и MATLAB и также использует пакет ARPACK [27]). Заметим, что поиск собственных значений для вычисления F_c может быть слишком вычислительно затратным на практике и делается только для сравнения параметров, полученных при оптимизации F_s и F_c.

5. Численный эксперимент

В этом разделе представлено сравнение классического функционала сходимости и предложенного стохастического функционала для выбора оптимального параметра в предобусловливателе $\mathrm{RIC}_\alpha(0)$ [19]. Мы начинаем раздел с описания тестовой задачи.

5.1. Тестовая задача. В тестах использовались линейные системы, полученные стандартной конечно-разностной аппроксимацией второго порядка следующей краевой задачи для неизвестной функции $u(x,y)$:

$$\begin{aligned} &-(D_1 u_x)_x - (D_2 u_y)_y = g(x,y), \quad (x,y) \in \Omega = [0,1] \times [0,1], \\ &u(x,y)|_{\partial\Omega} = 0, \end{aligned} \tag{23}$$

где индексы $\cdot_x$, $\cdot_y$ обозначают частные производные по отношению к x и y, соответственно. Мы рассматриваем два случая: в первом случае коэффициенты $D_{1,2}$ берутся тождественно равными единице во всей области Ω, а во втором — разрывными:

$$D_1 = \begin{cases} 1000, & \text{если } (x,y) \in [\frac{1}{4},\frac{3}{4}] \times [\frac{1}{4},\frac{3}{4}], \\ 1, & \text{иначе,} \end{cases} \qquad D_2 = \frac{1}{2} D_1.$$

Функция правой части $g(x,y)$ выбирается так, чтобы значения функции

$$u(x,y) = \sin(\pi x)\sin(\pi y) \tag{24}$$

на конечно-разностной сетке являлись элементами вектора точного решения дискретизированной краевой задачи.

5.2. Сравнение двух функционалов. Для сравнения предложенного функционала (8) с классическим функционалом (10) рассмотрены четыре тестовые линейные системы. Эти линейные системы были получены из тестовой задачи (23) с правой частью (24) и следующими наборами параметров:

1. $m = 2500$ (сетка 52×52), постоянные коэффициенты $D_{1,2}$;
2. $m = 2500$ (сетка 52×52), разрывные коэффициенты $D_{1,2}$;
3. $m = 10000$ (сетка 102×102), постоянные коэффициенты $D_{1,2}$;
4. $m = 10000$ (сетка 102×102), разрывные коэффициенты $D_{1,2}$.

В экспериментах для вычисления функционала F_s использовалась евклидова норма. Кроме того, мы использовали $n = 50$ начальных приближений, которых было достаточно для сходимости (см. также раздел 5.3).

Следуя работам [26],[3, Section 13.2.1], мы искали оптимальное значение параметра α предобусловливателя $\mathrm{RIC}_\alpha(0)$ на интервале $[0.9, 1]$. Число итераций K, используемое в функционалах сходимости F_sи F_c, выбиралось так, чтобы заданная точность достигалась для разумного, но ещё не оптимизированного значения α. В численных экспериментах, представленных в этой работе, для уменьшения нормы невязки $\|r_k\|/\|r_0\|$ (где $r_k = b - Ax_k$) задавалась точность 10^{-7}. Значения K для такой заданной точности представлены в Таблице 1. Пусть α_c^* — оптимальное значение α, полученное оптимизацией функционала F_c, а α_s^* — оптимизацией F_s. Для поиска как α_c^*, так и α_s^* использовался оптимизационный процесс Брента, для которого была выбрана точность 10^{-5}, достаточная в данном случае. Вычисленные таким образом значения α_c^* и α_s^* представлены в Таблице 1. Здесь незначащие (учитывая точность оптимизационного процесса) цифры приводятся в скобках.

Таблица 1. Число итераций K, используемое при вычислении функционалов F_s и F_c (для получения решения с точностью 10^{-7}), и соответствующие оптимальные параметры α_s^* и α_c^*

Тестовый пример	K	α_s^*	α_c^*
$m = 2500$, постоянные $D_{1,2}$	20	$0.98257(07)$	$0.99618(02)$
$m = 2500$, разрывные $D_{1,2}$	30	$0.97671(44)$	$0.99999(47)$
$m = 10000$, постоянные $D_{1,2}$	35	$0.99245(52)$	$0.99900(93)$
$m = 10000$, разрывные $D_{1,2}$	45	$0.99451(65)$	$0.99999(33)$

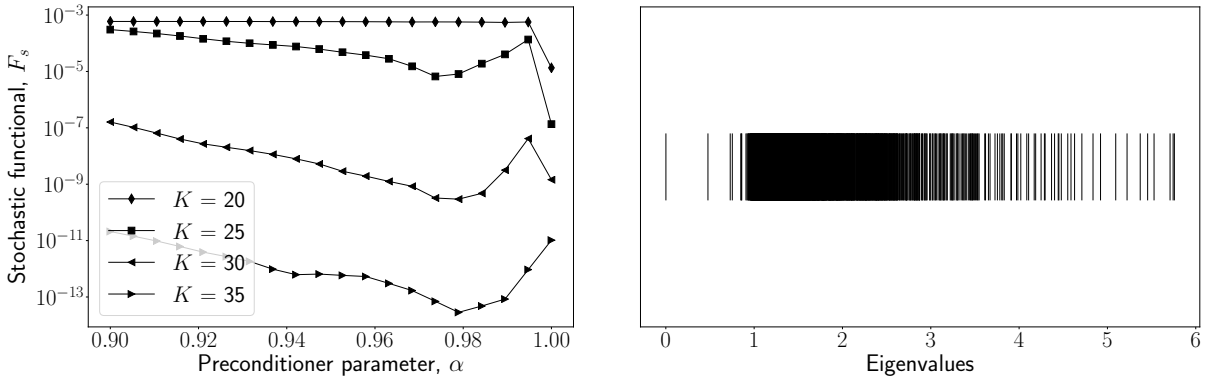
На рис. 3–6 представлены зависимости обоих функционалов от параметра предобусловливания α (где график (а) соответствует стохастическому функционалу, а график (в) — классическому). На графиках видим, что стохастический функционал F_s, как правило, имеет минимум близко к единице, в то время как F_c — чаще в точности в единице. Распределения собственных значений предобусловленной матрицы для значения α_s^* представлены на рис. 3б),4б),5б) и 6б). Для значения α_c^* аналогичные графики представлены на рис. 3г),4г),5г) и 6г).

На графиках распределения собственных значений видно, что значение α_s^* даёт бóльшую разреженность спектра на маленьких собственных числах, чем значение α_c^*. Следовательно, можно ожидать, что метод сопряжённых градиентов будет сходиться быстрее для α_s^*, чем для α_c^*. Это подтверждается графиками сходимости на рис. 3д),4д),5д) и 6д).

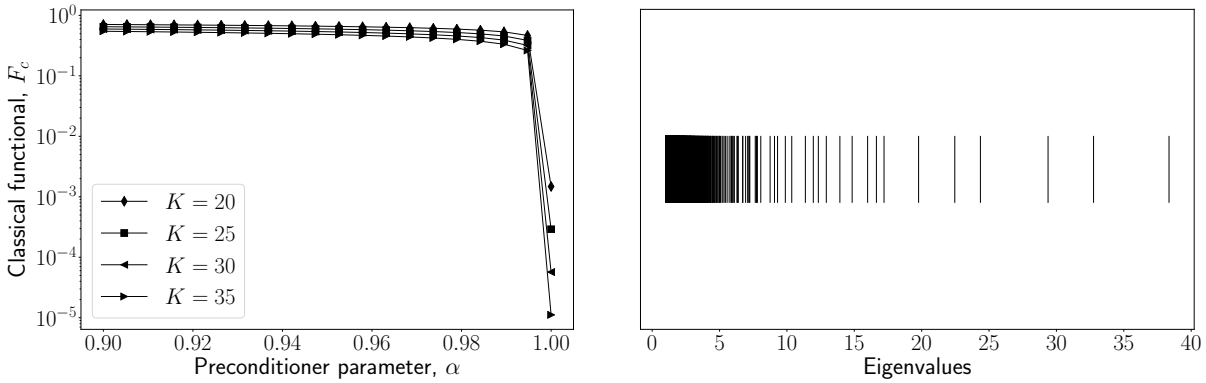
5.3. Количество случайных начальных приближений n. Вычислительные затраты на решение задачи оптимизации существенно зависят от выбора числа начальных приближений n (см. раздел 4). Как указано выше, во всех экспериментах использовалось значение $n = 50$. В этом разделе на тестовой задаче с $m = 2500$ и постоянными коэффициентами $D_{1,2}$ мы проверяем, насколько результаты экспериментов чувствительны к выбору n. Оказывается, что такие же результаты могут быть получены с использованием меньшего числа начальных приближений n. В соответствии с Таблицей 1, для этого тестового примера было взято $K = 20$ итераций. Зависимость предлагаемого стохастического функционала сходимости F_s от параметра предобусловливателя α для нескольких значений n показана на рис. 7. Как видно из графиков, чем больше n, тем более гладким оказывается график, но уже $n = 10$ достаточно для адекватного представления рассматриваемой зависимости. На рис. 8 показаны доверительные интервалы для F_s.

6. Заключение

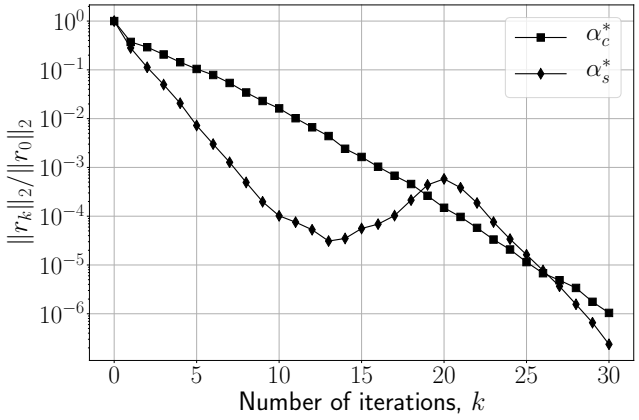
В данной работе предложен стохастический подход оценки скорости сходимости итерационных линейных решателей. Предлагаемая нами оценка (которую мы называем стохастическим функционалом сходимости) основан на отслеживании *средней* скорости сходимости для для определённого числа векторов начального приближения, выбираемых случайным образом. Для линейных стационарных итерационных методов показано, что предложенный стохастический функционал сходимости совпадает с классической оценкой сходимости, основанной на спектральном радиусе матрицы перехода итераций. Для метода сопряжённых градиентов, являющегося нелинейным нестационарным методом, проделанные анализ и численные эксперименты указывают, что стохастический функционал сходимости даёт более точную меру сходимости, чем классическая оценка, использующая спектральное число обусловленности предобусловленной матрицы. Кроме того, в данной работе продемонстрировано, как предложенный стохастический функционал сходимости может быть использован для оптимизации параметров предобусловливателей в методе сопряжённых градиентов. Представлены численные эксперименты для сопряжённых градиентов, предобусловленных неполным разложением Холесского $\mathrm{RIC}_\alpha(0)$. Численные эксперименты показывают, что предложенный стохастический функци-



а) **б)**

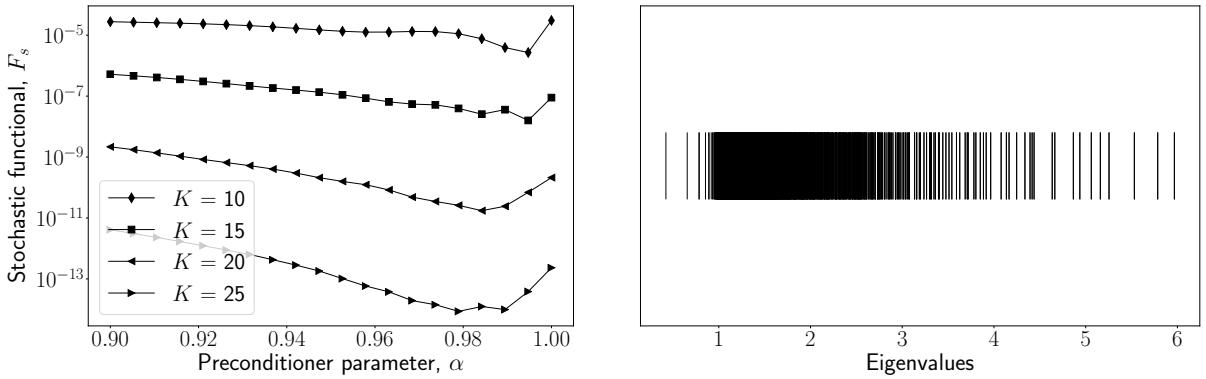


в) **г)**

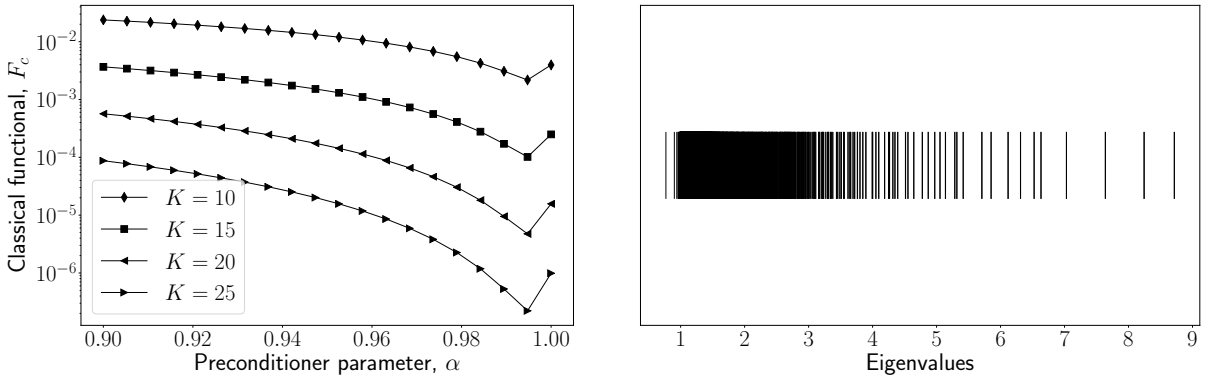


д)

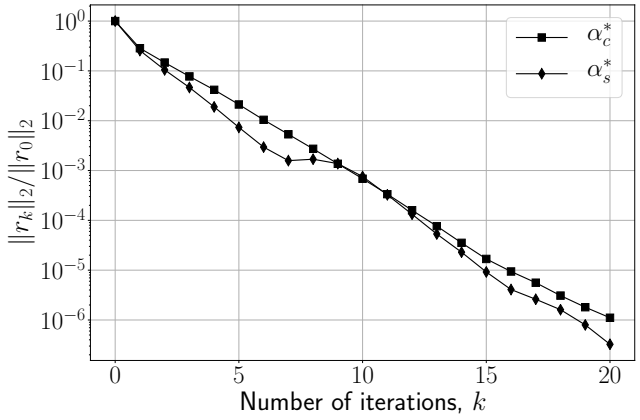
Рис. 3. Тестовый пример с параметром $m = 2500$ и постоянными коэффициентами $D_{1,2}$. Рис. 3а), 3в): зависимость функционалов F_s, F_c от параметра предобусловливателя α для различного числа итераций K. Рис. 3б), 3г): собственные значения предобусловленной матрицы A для оптимального значения параметра α_s^* (найденного для F_s, рис. 3б)) и α_c^* (найденного для F_c, рис. 3г)). Рис. 3д): сходимость относительной нормы невязки для метода сопряжённых градиентов с предобусловливателем $\mathrm{RIC}_\alpha(0)$ для α_s^* и α_c^*.



а) **б)**

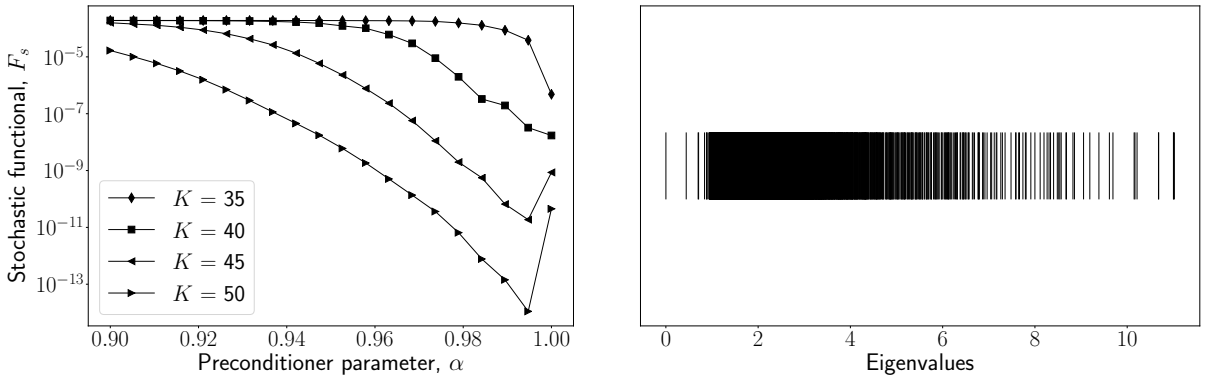


в) **г)**

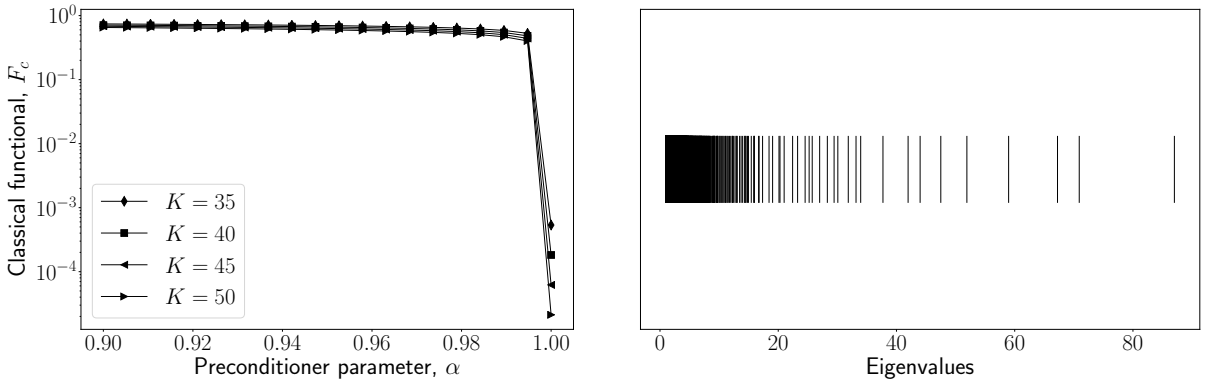


д)

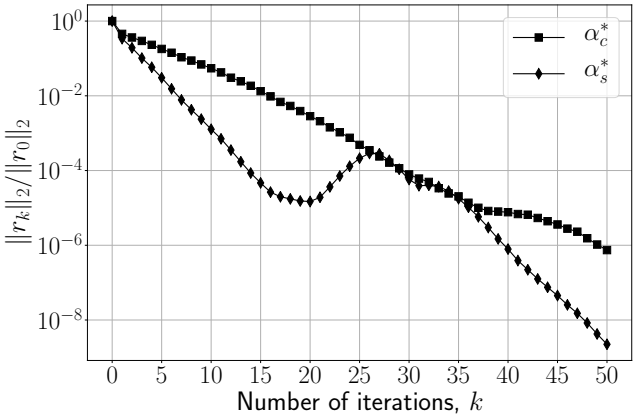
Рис. 4. Тестовый пример с параметром $m = 2500$ и постоянными коэффициентами $D_{1,2}$. Рис. 4а), 4в): зависимость функционалов F_s, F_c от параметра предобусловливателя α для различного числа итераций K. Рис. 4б), 4г): собственные значения предобусловленной матрицы A для оптимального значения параметра α_s^* (найденного для F_s, рис. 4б)) и α_c^* (найденного для F_c, рис. 4г)). Рис. 4д): сходимость относительной нормы невязки для метода сопряжённых градиентов с предобусловливателем $\mathrm{RIC}_\alpha(0)$ для α_s^* и α_c^*.



а) **б)**

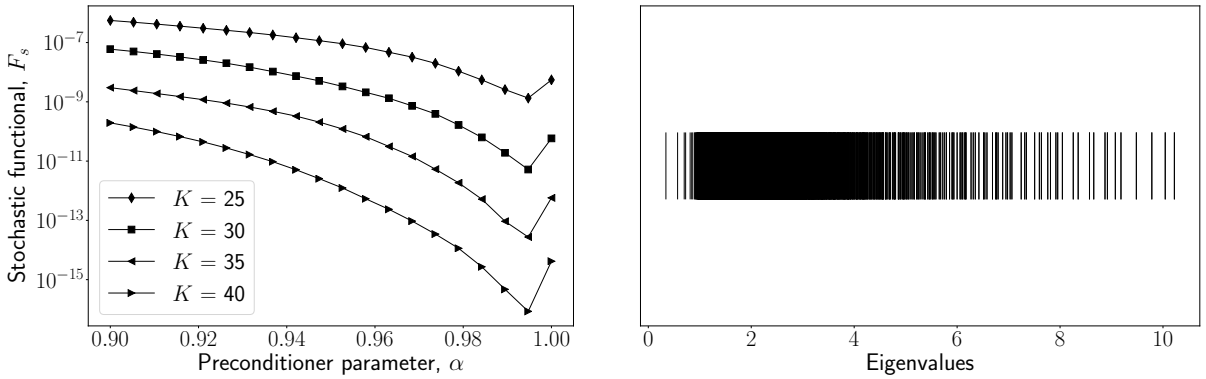


в) **г)**

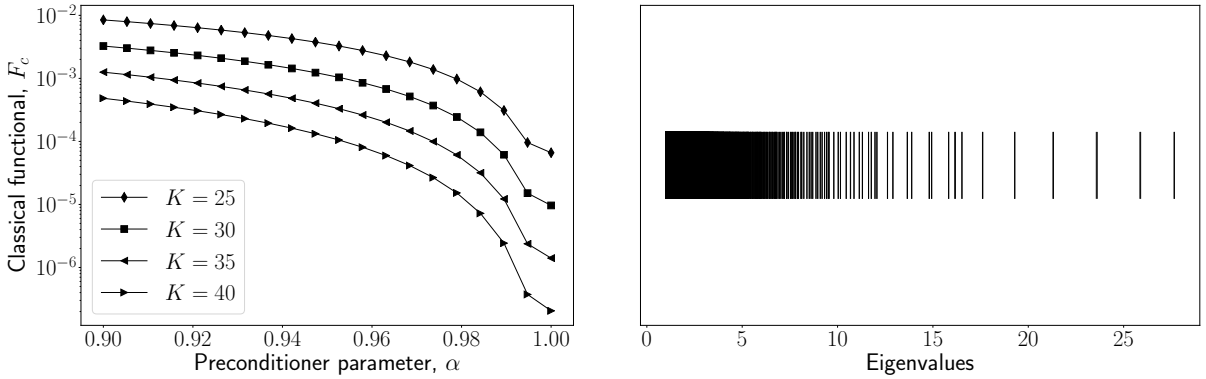


д)

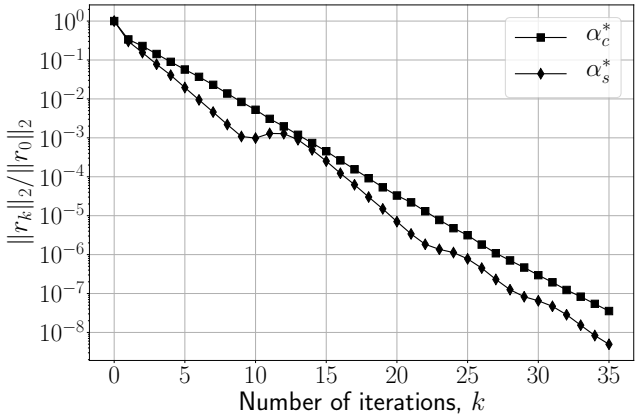
Рис. 5. Тестовый пример с параметром $m = 10000$, разрывные коэффициенты $D_{1,2}$. Рис. 5а), 5в): зависимость функционалов F_s и F_c от параметра предобусловливателя α для различного числа итераций K. Рис. 5б), 5г): собственные значения предобусловленной матрицы A для оптимального значения параметра α_s^* (найденного для F_s, рис. 5б)) и α_c^* (найденного для F_c, рис. 5г)). Рис. 5д): Сходимость относительной нормы невязки для метода сопряжённых градиентов с предобусловливателем $\mathrm{RIC}_\alpha(0)$ для α_s^* и α_c^*.



а) **б)**



в) **г)**



д)

Рис. 6. Тестовый пример с параметром $m = 10000$, постоянные коэффициенты $D_{1,2}$. Рис. 6а), 6в): зависимость функционалов F_s и F_c от параметра предобусловливателя α для различного числа итераций K. Рис. 6б), 6г): собственные значения предобусловленной матрицы A для оптимального значения параметра α_s^* (найденного для F_s, рис. 6б)) и α_c^* (найденного для F_c, рис. 6г)). Рис. 6д): Сходимость относительной нормы невязки для метода сопряжённых градиентов с предобусловливателем $\mathrm{RIC}_\alpha(0)$ для α_s^* и α_c^*.

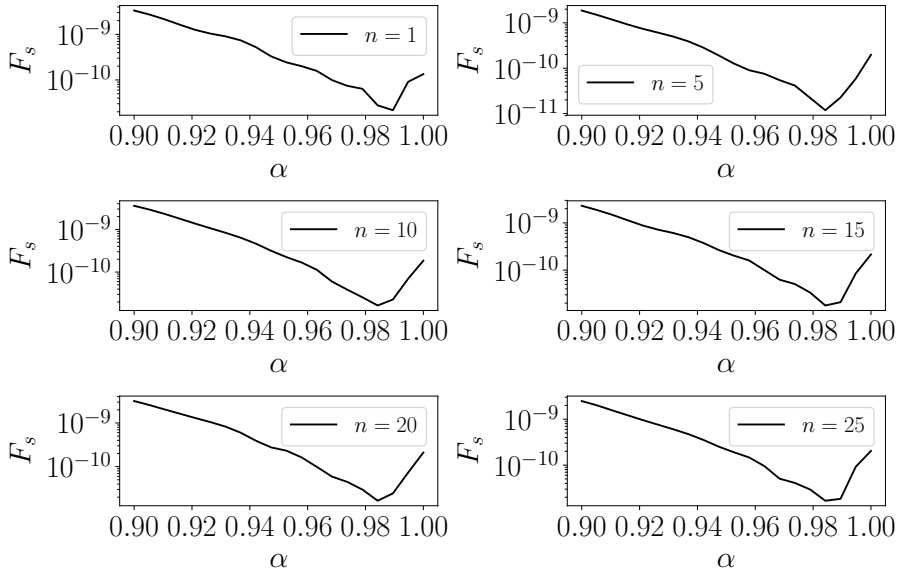


Рис. 7. Зависимость стохастического функционала сходимости F_s от параметра предобусловливателя α для $K = 20$ и различного числа начальных приближений n. Параметры тестового примера: $m = 2500$, постоянные коэффициенты $D_{1,2}$. График для $n = 50$, используемого в экспериментах, неотличим от графика для $n = 25$.

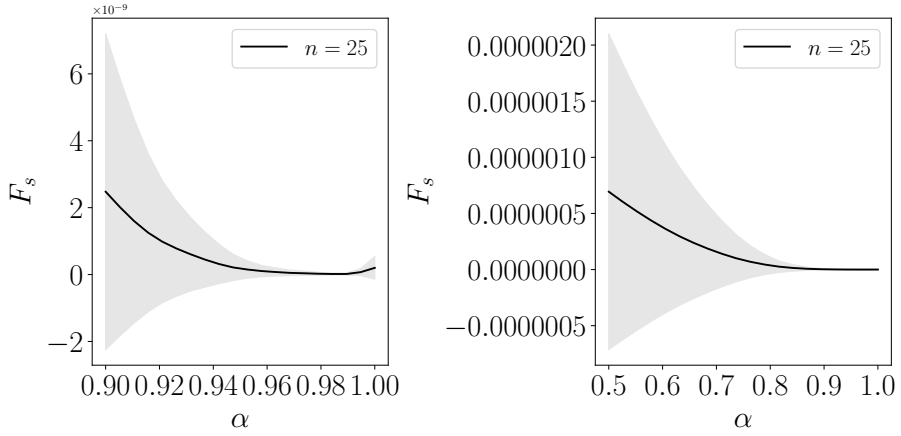


Рис. 8. Доверительный интервал (показан серым цветом) стохастического функционала F_s для $\alpha \in [0.9,1]$ (слева) и $\alpha \in [0.5,1]$ (справа). Тестовый пример с параметром $m = 2500$ и постоянным коэффициентом $D_{1,2}$.

онал сходимости даёт лучший способ оптимизации параметров в $\mathrm{RIC}_\alpha(0)$, чем минимизация спектрального числа обусловленности.

Простой анализ сходимости, представленный в работе, показывает, что классическая оценка сходимости, основанная на спектральном числе обусловленности, может быть улучшена для определённых векторов начального приближения. Интересной нерешённой задачей является вопрос о том, могут ли другие оценки сходимости, в частности, оценки, показывающие «суперлиней-

ную» сходимость, быть улучшены для определённых векторов начального приближения. Мы предполагаем, что ответ на этот вопрос утвердительный, и оставляем его рассмотрение на будущее.

Другим интересным продолжением этой работы может являться оптимизация предобусловливателей по отношению к нескольким параметрам, актуальная, например, для циркулянтных предобусловливателей [28, 29, 30]. В этом случае могут быть успешно использованы градиентные оптимизационные методы в сочетании с методами автоматического дифференцирования (см. нашу недавнюю работу [24]).

В заключение отметим, что наша стохастическая процедура оптимизации, предположительно, может быть применена в сочетании с решением многих «похожих» линейных систем (т.е., например, систем со многими векторами правых частей) методами так называемых «утилизированных» подпространств Крылова [17, 18, 31]. В частности, предложенный оптимизационный процесс можно выполнять на множестве заданных правых частей, начиная с неоптимизированного предобусловливателя и выполняя оптимизацию по ходу решения линейных систем. Мы надеемся проверить работоспособность этого подхода в будущем.

Список литературы

[1] Axelsson O. Iterative solution methods. — Cambridge : Cambridge University Press, 1994.

[2] Saad Y. Iterative Methods for Sparse Linear Systems. — 2d edition. — SIAM, 2003. — Available from `http://www-users.cs.umn.edu/~saad/books.html`.

[3] van der Vorst H. A. Iterative Krylov methods for large linear systems. — Cambridge University Press, 2003.

[4] Golub G. H., Van Loan C. F. Matrix computations. — 3 edition. — JHU Press, 2012.

[5] Hestenes M. R., Stiefel E. Methods of conjugate gradients for solving linear systems. — NBS Washington, DC, 1952. — Vol. 49.

[6] Saad Y., Schultz M. H. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems // SIAM Journal on scientific and statistical computing. — 1986. — Vol. 7, no. 3. — P. 856–869.

[7] van der Sluis A., van der Vorst H. A. The rate of convergence of conjugate gradients // Numer. Math. — 1986. — Vol. 48. — P. 543–560.

[8] van der Vorst H. A., Vuik C. The superlinear convergence of GMRES // J. Comput. Appl. Math. — 1993. — Vol. 48. — P. 327–341.

[9] Benzi M. Preconditioning techniques for large linear systems: A survey // Journal of Computational Physics. — 2002. — Vol. 182. — P. 418–477.

[10] Varga R. S. Matrix Iterative Analysis. — Prentice-Hall, 1962.

[11] Numerical Linear Algebra for High-Performance Computers / Jack J. Dongarra, Iain S. Duff, Danny C. Sorensen, Henk A. van der Vorst. — Philadelphia, PA : SIAM, 1998. — P. 342.

[12] Kaporin I. E. New convergence results and preconditioning strategies for the conjugate gradient method // Numerical linear algebra with applications. — 1994. — Vol. 1, no. 2. — P. 179–210.

[13] Kaporin I. Scaling, preconditioning, and superlinear convergence in GMRES-type iterations // Matrix Methods: Theory, Algorithms and Applications: Dedicated to the Memory of Gene Golub. — World Scientific, 2010. — P. 273–295.

[14] Bochev M. A., Krukier L. A. Iterative solution of strongly nonsymmetric systems of linear algebraic equations // Russian Comput. Mathematics and Math. Physics. — 1997. — Vol. 37, no. 11. — P. 1241–1251.

[15] Botchev M. A., Golub G. H. A class of nonsymmetric preconditioners for saddle point problems // SIAM Journal on Matrix Analysis and Applications. — 2006. — Vol. 27, no. 4. — P. 1125–1149.

[16] Parameter estimates for the relaxed dimensional factorization preconditioner and application to hemodynamics / Michele Benzi, Simone Deparis, Gwenol Grandperrin, Alfio Quarteroni // Computer Methods in Applied Mechanics and Engineering. — 2016. — Vol. 300. — P. 129–145.

[17] Recycling Krylov subspaces for CFD applications and a new hybrid recycling solver / Amit Amritkar, Eric de Sturler, Katarzyna Świrydowicz et al. // Journal of Computational Physics. — 2015. — Vol. 303. — P. 222–237.

[18] Benner P., Feng L. Recycling Krylov subspaces for solving linear systems with successively changing right-hand sides arising in model reduction // Model Reduction for Circuit Simulation. — Springer, 2011. — P. 125–140.

[19] Axelsson O., Lindskog G. On the eigenvalue distribution of a class of preconditioning methods // Numerische Mathematik. — 1986. — Vol. 48, no. 5. — P. 479–498.

[20] Meurant G. Computer solution of large linear systems. — Elsevier, 1999. — Vol. 28.

[21] Greenbaum A. Iterative methods for solving linear systems. — SIAM, 1997. — Vol. 17.

[22] Meurant G. The Lanczos and Conjugate Gradient Algorithms: from theory to finite precision computations. — SIAM, 2006.

[23] Hutchinson M. F. A stochastic estimator of the trace of the influence matrix for Laplacian smoothing splines // Communications in Statistics-Simulation and Computation. — 1990. — Vol. 19, no. 2. — P. 433–450.

[24] Katrutsa A., Daulbaev T., Oseledets I. Deep multigrid: learning prolongation and restriction matrices // arXiv preprint arXiv:1711.03825. — 2017.

[25] Brent R. P. Algorithms for minimization without derivatives. — Courier Corporation, 2013.

[26] van der Vorst H. A. ICCG and related methods for 3D problems on vector computers // Computer Physics Communications. — 1989. — Vol. 53, no. 1–3. — P. 223–235.

[27] ARPACK: a collection of subroutines designed to solve large scale eigenvalue problems. — `http://www.caam.rice.edu/software/ARPACK/`.

[28] Chan T. F. An optimal circulant preconditioner for Toeplitz systems // SIAM journal on scientific and statistical computing. — 1988. — Vol. 9, no. 4. — P. 766–771.

[29] Oseledets I., Tyrtyshnikov E. A unifying approach to the construction of circulant preconditioners // Linear algebra and its applications. — 2006. — Vol. 418, no. 2-3. — P. 435–449.

[30] Tyrtyshnikov E. E. Optimal and superoptimal circulant preconditioners // SIAM Journal on Matrix Analysis and Applications. — 1992. — Vol. 13, no. 2. — P. 459–473.

[31] Computing reduced order models via inner-outer Krylov recycling in diffuse optical tomography / Meghan O'Connell, Misha E Kilmer, Eric de Sturler, Serkan Gugercin // SIAM Journal on Scientific Computing. — 2017. — Vol. 39, no. 2. — P. B272–B297.

Оглавление