

ОЦЕНКА ЧИСЛА ОБУСЛОВЛЕННОСТИ БОЛЬШОЙ ПЛОХО ОБУСЛОВЛЕННОЙ МАТРИЦЫ МЕТОДОМ МОНТЕ — КАРЛО

В. С. Антюфеев^{1,2}

¹ *Институт вычислительной математики и математической геофизики СО РАН, 630090, Новосибирск*

² *Новосибирский государственный университет, 630090, Новосибирск*

УДК 519.6

Рассматриваются плохо обусловленные детерминированные (неслучайные) матрицы систем линейных алгебраических уравнений со случайной ошибкой вектора правой части. Исследуется число ν обусловленности матрицы системы. Показано: если случайный вектор ошибки правой части системы и матрица системы удовлетворяют некоторым естественным условиям, величина числа ν может быть значительно уменьшена сравнительно со «стандартной» величиной.

Ключевые слова: плохо обусловленная матрица, число обусловленности, случайный вектор ошибки, функция распределения

Введение

0.1. Рассмотрим систему линейных алгебраических уравнений $A\mathbf{x} = \mathbf{b}$. Пусть $\Delta\mathbf{b}$ и $\Delta\mathbf{x}$ — ошибки правой части \mathbf{b} и решения \mathbf{x} , соответственно. Относительные ошибки решения и правой части связаны неравенством

$$\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq \nu \cdot \frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|}, \quad (1)$$

где число обусловленности $\nu := \|A^{-1}\| \cdot \|A\|$ определяет «качество» матрицы A . Его величина может быть использована при определении коэффициента регуляризации системы. Число ν в неравенстве (1) нельзя заменить на меньшее. Однако, в некотором смысле оценка (1) является грубой. Ниже мы уточним смысл этого утверждения и покажем, как можно значительно уменьшить величину ν в оценке.

0.2. Естественным считать, что вектор ошибки $\Delta\mathbf{b}$ является случайным. Если компоненты этого вектора взаимно-независимы, естественно также считать, что случайный радиус-вектор $\mathbf{s} := \Delta\mathbf{b}/\|\Delta\mathbf{b}\|$ его направления является изотропным, то есть случайная точка \mathbf{s} распределена равномерно на единичной сфере $S^{n-1} \subset R^n$. Будем считать, что векторы \mathbf{x}, \mathbf{b} фиксированы (неслучайные).

В работе предложен новый способ оценки числа обусловленности, соответствующего норме $\|\cdot\|_2$ с учетом случайности вектора ошибки. Ниже изложены соображения, которые приводят к такой оценке.

0.3. Неравенство (1) эквивалентно неравенству

$$\frac{|\Delta\mathbf{x}|}{|\Delta\mathbf{b}|} \cdot \frac{|\mathbf{b}|}{|\mathbf{x}|} \leq \|A^{-1}\| \cdot \|A\|,$$

которое является произведением неравенств

$$\frac{|\Delta\mathbf{x}|}{|\Delta\mathbf{b}|} \leq \|A^{-1}\| \quad (2)$$

и

$$\frac{|\mathbf{b}|}{|\mathbf{x}|} \leq \|A\|.$$

Неравенство (2) относится к случайным векторам $\Delta \mathbf{x}, \Delta \mathbf{b}$. Перепишем его в эквивалентной форме:

$$|A^{-1} \mathbf{s}| \leq \|A^{-1}\| \quad (3)$$

0.4. Если матрица A плохо обусловлена, длина радиус-вектора $\{\mathbf{e} = A^{-1} \mathbf{s}\}$ испытывает значительные вариации. Для иллюстрации рассмотрим двумерный случай. На Рис.1 точки \mathbf{s}_k расположены равномерно на единичной окружности S^1 , а соответствующие точки $\{\mathbf{e}_k = A^{-1} \mathbf{s}_k\}$ расположены на эллипсе E . Из рисунка видно, что с большой вероятностью выполняется соотношение

$$|A^{-1} \mathbf{s}| \ll \max_{\mathbf{s} \in S^1} |A^{-1} \mathbf{s}| \equiv \|A^{-1}\| \quad (4)$$

Поэтому от неравенства (4) для неслучайной величины следует перейти к соотношению для распределения случайной величины $|A^{-1} \mathbf{s}|$.

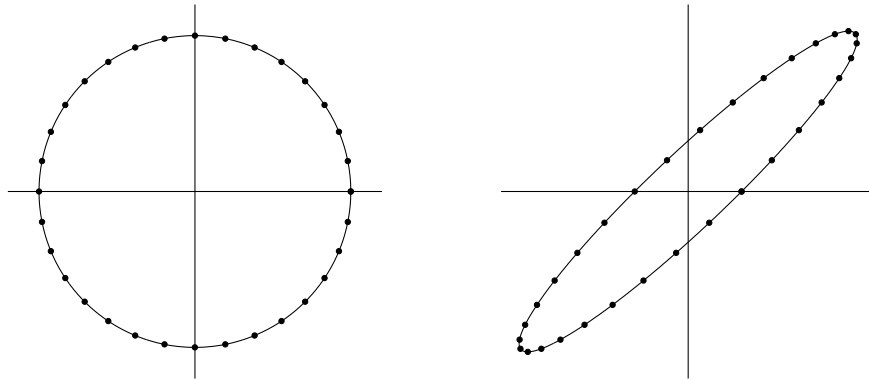


Рис. 1: Расположение точек $\mathbf{s}_k \in S^1$ и $\mathbf{e}_k \in E$

1 Случайная величина \mathbf{L}

1.1. Запишем выражение для длины \mathbf{L} радиус-вектора точки $\mathbf{e} = A^{-1} \mathbf{s} \in E$:

$$\mathbf{L} := |A^{-1} \mathbf{s}| = (A^{-1} \mathbf{s} \cdot A^{-1} \mathbf{s})^{1/2} = (\mathbf{s} \cdot (A^{-1})^* A^{-1} \mathbf{s})^{1/2} = ((A^* A)^{-1} \mathbf{s} \cdot \mathbf{s})^{1/2}$$

Здесь \mathbf{s} случайная точка, поэтому \mathbf{L} — случайная величина.

Пусть $\lambda_1^2 > \dots > \lambda_n^2 > 0$ — собственные числа матрицы $(A^* A)^{-1}$. Выполняются соотношения [3]:

$$1/\|A\| = \lambda_n \leq \mathbf{L} = \left(\sum_{k=1}^n \lambda_k^2 s_k^2 \right)^{1/2} \leq \lambda_1 = \|A^{-1}\| \quad (5)$$

1.2. \mathbf{s} — случайная точка, поэтому неравенство (3) — оценка случайной величины $|A^{-1} \mathbf{s}|$. Оно выполняется с вероятностью 1. Перепишем его в эквивалентной вероятностной форме:

$$P(\mathbf{L} < \lambda_1) = 1 \quad (6)$$

В свою очередь, (6) можно обобщить. Пусть F — функция распределения [1] случайной величины \mathbf{L} . Тогда соотношение

$$P(\mathbf{L} < t) = F(t)$$

является обобщением (6).

1.3. Рассмотрим вычисление значений функции F . Пусть μ — мера Лебега на сфере S^{n-1} . Точка \mathbf{s} распределена равномерно на S^{n-1} (см. (0.2), поэтому для любого (измеримого) множества $M \subset S^{n-1}$ выполняется соотношение

$$P(\mathbf{s} \in M) = \mu(M)/\mu(S^{n-1}). \quad (7)$$

Таким образом,

$$F(t) = \begin{cases} 0, & t \leq \lambda_n \\ \frac{\mu(\mathbf{L} < t)}{\mu(S^{n-1})}, & \lambda_n < t < \lambda_1 \\ 1, & t \geq \lambda_1 \end{cases} \quad (8)$$

Величину $\mu(\mathbf{L} < t)$ из (8) (и вместе с ней функцию F) не удается найти аналитически. Вместо этого мы вычисляем методом Монте — Карло выборочную функцию распределения величины \mathbf{L} . С этой целью моделируем последовательность случайных точек \mathbf{s}_k , $k = 1, 2, \dots$ равномерно на S^{n-1} , пользуясь известными формулами. Для каждой точки \mathbf{s}_k вычисляем $t_k = |A^{-1}\mathbf{s}_k|$. Числа t_k , $k = 1, 2, \dots$ образуют последовательность выборочных значений случайной величины \mathbf{L} . Если количество N_s точек \mathbf{s}_k достаточно велико, строим приближенную функцию распределения.

Все t_k , $k = 1, 2, \dots, N_s$ удовлетворяют неравенству $\lambda_N \leq t_k \leq \lambda_1$. Для каждого t_i подсчитаем количество N_i чисел t_k : $t_k < t_i$. Функцию F_{N_s} зададим в узлах t_k , $k = 1, 2, \dots, N_s$: $F_{N_s}(t_k) = N_i/N_s$. Кусочно-постоянная аппроксимация функции F_{N_s} между соседними узлами приводит к выборочной функции распределения. Ниже будем рассматривать кусочно-линейную аппроксимацию, чтобы получать удобные графики.

Случайная величина \mathbf{L} и ее функция распределения F определены выбором конечной последовательности $\Lambda(1:n) = \{\lambda_k, k = 1, \dots, n\}$, поэтому ниже будем использовать обозначения $\mathbf{L} \equiv \mathbf{L}_{\Lambda(1:n)}$, $F \equiv F_{\Lambda(1:n)}$.

Далее мы рассматриваем лишь числовые последовательности $\Lambda(1:n)$, где n велико, а λ_n/λ_1 мало, не упоминая матрицы, из которых получены.

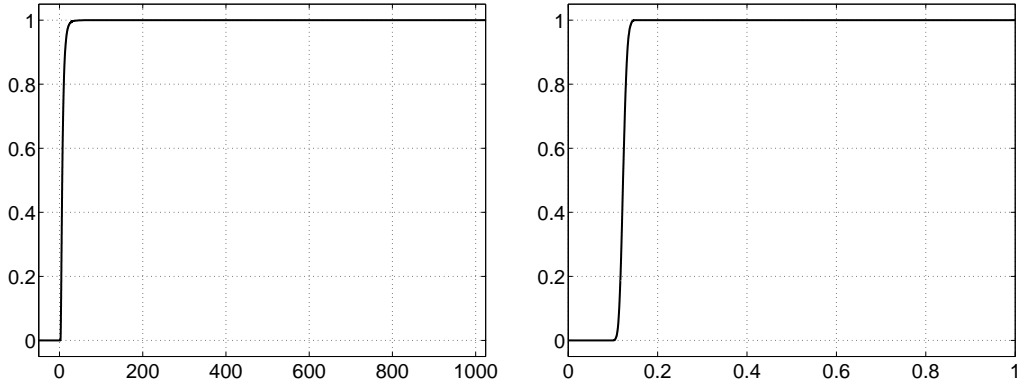


Рис. 2: Слева: $\Lambda^1 = \{\lambda_k = 2^{10}, 2^9, \dots, 1\}$; справа: $\Lambda^2 = \{\lambda_k = 1.01^0, 1.01^{-1}, \dots, 1.01^{-300}\}$.

Для двух таких последовательностей Λ^1, Λ^2 приведены графики соответствующих функций $F_{\Lambda^1}, F_{\Lambda^2}$. Они похожи на график функции Хевисайда χ . В следующем разделе доказана теорема, объясняющая это.

2 Теорема сходимости

Здесь показано, что для рассматриваемых конечных числовых последовательностей $\Lambda(1:n)$ (см. раздел 1) $F_{\Lambda(1:n)} \simeq \chi$. Чтобы доказать соответствующее предельное соотношение, образуем вспомогательную «нормированную» последовательность $\bar{\Lambda}(1:n) = \{\bar{\lambda}_k = \lambda_k/\lambda_1, k = 1, \dots, n\}$. Здесь $\bar{\lambda}_1 = 1$, $\bar{\lambda}_n \simeq 0$. Конечную «нормированную» последовательность $\bar{\Lambda}(1:n)$ можно рассматривать как часть бесконечной последовательности $\{\bar{\lambda}_k, k = 1, 2, \dots\}$, где $\bar{\lambda}_k \rightarrow 0$. Для таких последовательностей выполняется следующая **теорема**

Теорема. Пусть $\bar{\lambda}_k \rightarrow 0$; $\mathbf{s}^n = (s_1^n, \dots, s_n^n)$ — случайная точка, распределенная равномерно на единичной сфере S^{n-1} ; $\mathbf{L}_{\bar{\Lambda}(1:n)} = (\bar{\lambda}_1^2(s_1^n)^2 + \dots + \bar{\lambda}_n^2(s_n^n)^2)^{1/2}$. Тогда для всех $t \in R$

$$\lim_{n \rightarrow \infty} F_{\bar{\Lambda}(1:n)}(t) = \chi(t), \quad (9)$$

При этом сходимость $F_{\bar{\Lambda}(1:n)} \rightarrow \chi$ при любом $\varepsilon > 0$ является равномерной на множестве $R \setminus (0, \varepsilon)$.

Доказательство. Ниже все формулы, содержащие Γ -функцию, взяты из [2]. Сначала преобразуем соотношение (9).

2.1. Согласно (5), $0 < \bar{\lambda}_n \leq \mathbf{L}_{\bar{\lambda}(1:n)} \leq \bar{\lambda}_1 = 1$. Следовательно, $F_{\bar{\lambda}(1:n)}(t) \equiv 0$ при $t \in (-\infty, \bar{\lambda}_n]$, и $F_{\bar{\lambda}(1:n)}(t) \equiv 1$ при $t \in [1, +\infty)$ то есть $F_{\bar{\lambda}(1:n)}(t) \equiv \chi(t)$ при $t \in (-\infty, 0] \cup [1, +\infty)$. Теперь вместо (9) достаточно доказать, что при $t \in (0, 1)$

$$\lim_{n \rightarrow \infty} F_{\bar{\lambda}(1:n)} = 1, \quad (10)$$

причем сходимость равномерная на множестве $[\varepsilon, 1)$. при любом $\varepsilon \in (0, 1)$.

2.2. Поскольку $F_{\bar{\lambda}(1:n)}(t) = P(\mathbf{L}_{\bar{\lambda}(1:n)} < t) = 1 - P(\mathbf{L}_{\bar{\lambda}(1:n)} \geq t)$, то (10) эквивалентно соотношению

$$P(\mathbf{L}_{\bar{\lambda}(1:n)} \geq t) \stackrel{(7)}{=} \frac{\mu(\mathbf{L}_{\bar{\lambda}(1:n)} \geq t)}{\mu(S^{n-1})} = \frac{\mu\left(\sum_{k=1}^n \bar{\lambda}_k^2 (s_k^n)^2 \geq t^2\right)}{\mu(S^{n-1})} \rightarrow 0.$$

Итак, вместо (9) будем доказывать эквивалентное соотношение

$$\frac{\mu\left(\sum_{k=1}^n \bar{\lambda}_k^2 (s_k^n)^2 \geq t^2\right)}{\mu(S^{n-1})} \rightarrow 0 \quad (11)$$

2.3. Докажем вспомогательное неравенство. Пусть функция $f \geq 0$ определена на множестве S , $h > 0$. Тогда

$$\int_S f = \int_{f(\mathbf{x}) < h} + \int_{f(\mathbf{x}) \geq h} \geq h \cdot \mu\{f(\mathbf{x}) \geq h\}$$

Отсюда

$$\mu\{f(\mathbf{x}) \geq h\} \leq \frac{1}{h} \cdot \int_S f \quad (12)$$

2.4. Пусть в (12) $f(\mathbf{s}) = \sum \bar{\lambda}_k^2 (s_k^n)^2$, $h = t^2$. Оценим числитель дроби в (11):

$$\mu\left(\sum_{k=1}^n \bar{\lambda}_k^2 (s_k^n)^2 \geq t^2\right) \leq \frac{1}{t^2} \int_{S^{n-1}} \sum_{k=1}^n \bar{\lambda}_k^2 (s_k^n)^2 d\mathbf{s}^n = \left(\frac{1}{t^2} \sum_{k=1}^n \bar{\lambda}_k^2\right) \cdot \int_{S^{n-1}} (s_k^n)^2 d\mathbf{s}^n \quad (13)$$

Интеграл в правой части (13) не зависит от индекса k , поэтому его можно вынести за знак суммы, считая, что $k : 1 \leq k \leq n$ является произвольным.

2.5. Вычислим интеграл в правой части (13).

$$\int_{S^{n-1}} (s_k^n)^2 d\mathbf{s}^n = 2 \frac{\pi^{(n-1)/2}}{\Gamma((n-1)/2)} \cdot \int_0^\pi \cos^2 \varphi \sin^{n-2} \varphi d\varphi = 2 \frac{\pi^{(n-1)/2}}{\Gamma((n-1)/2)} \int_0^\pi (\sin^{n-2} \varphi - \sin^n \varphi) d\varphi$$

Здесь

$$\int_0^\pi \sin^n \varphi d\varphi = \sqrt{\pi} \cdot \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n+2}{2}\right)}$$

Ниже используем два простых утверждения [2]:

2.6.1. если $c_n \rightarrow 0$, то

$$\frac{1}{n} \sum_{k=1}^n c_k \rightarrow 0$$

2.6.2. Пусть $\lim_{k \rightarrow \infty} a_k = 0$ и $a_k \geq b_k \geq 0$ для всех k . Тогда $\lim_{k \rightarrow \infty} b_k = 0$.

2.7. Вернемся к доказательству (11):

$$0 \leq \frac{\mu\left(\sum_{k=1}^n \bar{\lambda}_k^2 (s_k^n)^2 \geq t^2\right)}{\mu(S^{n-1})} \leq \dots$$

учитывая 2.4, 2.5 и формулу $\mu(S^{n-1}) = 2 \frac{\pi^{n/2}}{\Gamma(n/2)}$,

$$\begin{aligned}
\ldots &\leq \left(\frac{1}{t^2} \sum_{k=1}^n \bar{\lambda}_k^2 \right) \cdot \frac{2 \frac{\pi^{(n-1)/2}}{\Gamma((n-1)/2)}}{2 \frac{\pi^{n/2}}{\Gamma(n/2)}} \cdot \sqrt{\pi} \cdot \left(\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} - \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n+2}{2}\right)} \right) = \\
&= \left(\frac{1}{t^2} \sum_{k=1}^n \bar{\lambda}_k^2 \right) \cdot \left(1 - \frac{\Gamma\left(\frac{n}{2}\right) \cdot \Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right) \cdot \Gamma\left(\frac{n+2}{2}\right)} \right) = \ldots
\end{aligned}$$

используя формулу $\Gamma(x+1) = x \cdot \Gamma(x)$ для $x = (n-1)/2$, $x = n/2$ и учитывая $t \geq \varepsilon$,

$$\ldots = \left(\frac{1}{t^2} \sum_{k=1}^n \bar{\lambda}_k^2 \right) \cdot \left(1 - \frac{\Gamma\left(\frac{n}{2}\right) \cdot \Gamma\left(\frac{n-1}{2}\right) \cdot \left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right) \cdot \Gamma\left(\frac{n}{2}\right) \cdot \left(\frac{n}{2}\right)} \right) = \frac{1}{t^2} \cdot \frac{1}{n} \sum_{k=1}^n \bar{\lambda}_k^2 \leq \frac{1}{\varepsilon^2} \left[\frac{1}{n} \sum_{k=1}^n \bar{\lambda}_k \right]$$

Возвращаясь к началу 2.7., имеем:

$$0 \leq \frac{\mu\left(\sum_{k=1}^n \bar{\lambda}_k^2 (s_k^n)^2 \geq t^2\right)}{\mu(S^{n-1})} \leq \frac{1}{\varepsilon^2} \left[\frac{1}{n} \sum_{k=1}^n \bar{\lambda}_k \right], \quad (14)$$

причем, согласно 2.6.1,

$$\frac{1}{\varepsilon^2} \left[\frac{1}{n} \sum_{k=1}^n \bar{\lambda}_k \right] \rightarrow 0,$$

а, согласно 2.6.2.,

$$\frac{\mu\left(\sum_{k=1}^n \bar{\lambda}_k^2 (s_k^n)^2 \geq t^2\right)}{\mu(S^{n-1})} \rightarrow 0,$$

причем здесь сходимость равномерная, так как мажоранта в (14) не зависит от t . Доказательство теоремы закончено. \square

Вернемся к конечной последовательности $\Lambda(1:n) = \{\lambda_k > 0, k = 1, \dots, n\}$, где n велико, а λ_n/λ_1 мало. Ей соответствует вспомогательная последовательность $\bar{\Lambda}(1:n)$. Поскольку $\mathbf{L}_{\Lambda(1:n)}$ и $\mathbf{L}_{\bar{\Lambda}(1:n)}$ связаны соотношением

$$\mathbf{L}_{\Lambda(1:n)} = (\lambda_1^2 (s_1^n)^2 + \dots + \lambda_n^2 (s_n^n)^2)^{1/2} = \lambda_1 \cdot (\bar{\lambda}_1^2 (s_1^n)^2 + \dots + \bar{\lambda}_n^2 (s_n^n)^2)^{1/2} = \lambda_1 \mathbf{L}_{\bar{\Lambda}(1:n)},$$

то функции $F_{\Lambda(1:n)}$ и $F_{\bar{\Lambda}(1:n)}$ связаны соотношением

$$F_{\Lambda(1:n)}(t) = F_{\bar{\Lambda}(1:n)}(t/\lambda_1)$$

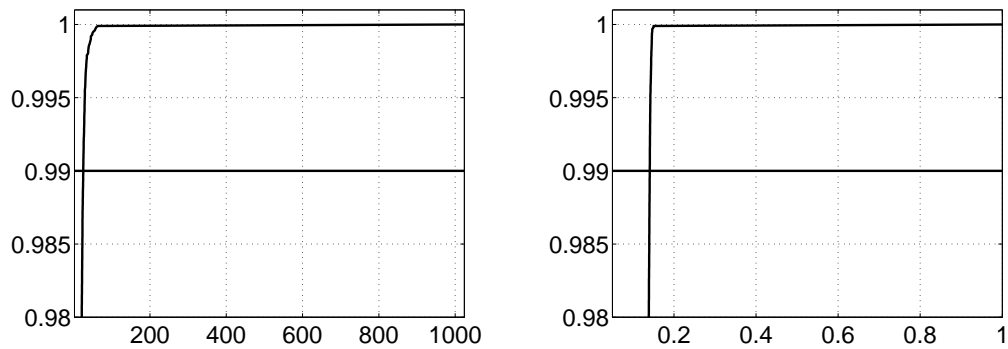
Согласно доказанной теореме, при больших n $F_{\bar{\Lambda}(1:n)} \simeq \chi$ за пределами интервала $(0, \varepsilon)$. Следовательно, $F_{\Lambda(1:n)} \simeq \chi$ за пределами интервала $(0, \lambda_1 \varepsilon)$.

В следующем разделе рассмотрим вопросы использования полученных результатов.

3 Замена значения коэффициента ν на ν_ε

Согласно определению и обозначениям, коэффициент обусловленности $\nu(A)$ — это наименьшее число, удовлетворяющее (с вероятностью 1) неравенству $\mathbf{L} \cdot \|A\| \leq \nu$.

Выберем такое малое $\varepsilon > 0$, что в расчетах можно пренебречь некоторым случайным событием, если его вероятность меньше ε . Найдем численное решение уравнения $F_{\bar{\Lambda}(1:n)}(t) = 1 - \varepsilon$ (см. Рис.3 — фрагмент Рис.2). Поскольку $P(\mathbf{L} < L_\varepsilon) \equiv F_{\bar{\Lambda}(1:n)}(L_\varepsilon) = 1 - \varepsilon$, то $P(\mathbf{L} \geq t_\varepsilon) = 1 - P(\mathbf{L} < L_\varepsilon) = \varepsilon$. Другими словами, событие $\{\mathbf{L} \geq t_\varepsilon\}$ является настолько маловероятным, что им можно пренебречь в расчетах.

Рис. 3: Графическое решение уравнений $F_{\Lambda^k}(t) = 1 - \varepsilon$ ($k = 1, 2$).

Пусть $\nu_\varepsilon = L_\varepsilon \cdot \|A\|$. Тогда с вероятностью $1 - \varepsilon$ выполняется неравенство

$$\frac{|\Delta x|}{|\Delta b|} \cdot \frac{|b|}{|x|} \equiv \mathbf{L} \cdot \|A\| \leq \nu_\varepsilon.$$

Таким образом, коэффициент ν можно заменить на ν_ε (хотя, см. ниже, $\nu_\varepsilon/\nu \ll 1$).

Отношение значений ν_ε и ν равно $\nu_\varepsilon/\nu = L_\varepsilon\|A\|/\lambda_1\|A\| = L_\varepsilon/\lambda_1$. Вычислим ν_ε/ν для последовательностей Λ^1, Λ^2 (см. раздел 1, заголовок к рисунку 2). Пусть $\varepsilon = 0.01$.

1) Последовательность Λ^1 : $L_\varepsilon = 0.2481$, $\lambda = 1024$, $\nu_\varepsilon/\nu = 0.00024$.

2) Последовательность Λ^2 : $L_\varepsilon = 0.1411$, $\lambda = 1$, $\nu_\varepsilon/\nu = 0.14$.

Значения дроби ν_ε/ν в 1), 2) существенно различаются. Видимо, причина заключена в том, что последовательность 1.01^{-k} (Λ^2) сходится к нулю гораздо медленнее, чем 2^{-k} (Λ^1).

Список литературы

- [1] Феллер В. Введение в теорию вероятностей и ее приложения. М.: Мир, 1984.
- [2] Фихтенгольц Г.М. Курс дифференциального и интегрального исчисления. М.: Наука, 1969.
- [3] Беллман Р. Введение в теорию матриц. М.: Наука, 1969.

Виктор Степанович Антюфеев — д.ф.-м.н., ст. науч.сотр. Института
вычислительной математики и математической геофизики СО РАН;
e-mail: ant@osmf.sscc.ru.

Дата поступления — 30 мая 2017 г.