Дальневосточный математический журнал. 2018. Т. 18. № 2. С. 219–232

УДК 519.65
MSC2010 97M50

© В. Г. Назаров[1]

О повышении точности вычислений в задаче нахождения химического состава среды

В работе рассматривается задача нахождения химического состава среды методом многократного просвечивания этой среды коллимированным рентгеновским излучением. При этом изучается вопрос о возможности повышения точности решения путем проведения нескольких серий повторных измерений прошедшего через среду излучения. Показано, что при некоторых предположениях о поведении ошибок измерений излучения, ошибки решения стремятся к нулю с ростом числа проведенных измерений. В качестве иллюстрации приводятся результаты расчетов, выполненных для конкретного вещества.

Ключевые слова: *радиография сплошной среды, нахождение химического состава среды, сингулярное разложение матрицы, точность вычислений.*

Введение

Радиографические методы зондирования среды и, в частности, определение ее химического состава являются эффективным способом исследования в тех случаях, когда требуется выполнить неразрушающий контроль изделия или когда непосредственный доступ к объекту исследования затруднен или нежелателен. Такая ситуация может возникать в научных исследованиях, в таможенном деле, в медицине, в химической промышленности. В ходе радиографического зондирования требуемые результаты можно получить намного быстрее, чем при лабораторных исследованиях. Количество научных публикаций на эту тему как российских, так и зарубежных исследователей остается стабильно высоким. При этом авторы используют различные подходы к решению рассматриваемой задачи, и эти подходы могут заметно отличаться в зависимости от конкретной задачи. Среди публикиций отметим [1–3].

В данной работе рассматривается задача нахождения химического состава среды методом многократного просвечивания этой среды коллимированным рентгеновским излучением. Математическая постановка задачи приводит к системам линейных алгебраических уравнений с плохо обусловленными матрицами [4, 5]. При большом числе переменных решения систем могут иметь значительные ошибки. По

[1] Институт прикладной математики ДВО РАН, 690041, г. Владивосток, ул. Радио, 7. Электронная почта: naz@iam.dvo.ru

этой причине изучается вопрос повышения точности решения путем проведения нескольких серий повторных измерений. Показано, что при некоторых предположениях о поведении ошибок правой части системы, ошибки решения стремятся к нулю с ростом числа проведенных измерений. Построен (абстрактный) пример переопределенной системы уравнений, для которой нормальное обобщенное решение системы заметно хуже приближенного решения, предложенного автором.

Для рассматриваемой задачи приводятся результаты расчетов, выполненных для конкретного вещества. В целом представленные в работе результаты могут быть весьма полезны в практическом приложении, в частности, при таможенном досмотре грузов и багажа.

1. Предварительные замечания и постановка задачи

Рассматриваемая задача обладает определенной спецификой, которая станет ясна из дальнейшего, поэтому, перед тем как приводить её точную формулировку, введем необходимые обозначения и сделаем ряд пояснений. Сначала кратко остановимся на задаче химии, которая более подробно обсуждалась в [4, 5].

Пусть исследуемый образец G_0 является однородным по химическому составу веществом X_0, причем все химические элементы (или простые химические соединения, которые мы дальше также будем называть элементами), входящие в состав X_0, присутствуют в некотором заранее заданном перечне элементов $X_1,...,X_N$, который нам известен. Образец G_0 имеет толщину l и подвергается облучению потоком фотонов, коллимированным как по направлению, так и по энергии и идущим вдоль некоторой фиксированной прямой (см. рис. 1).

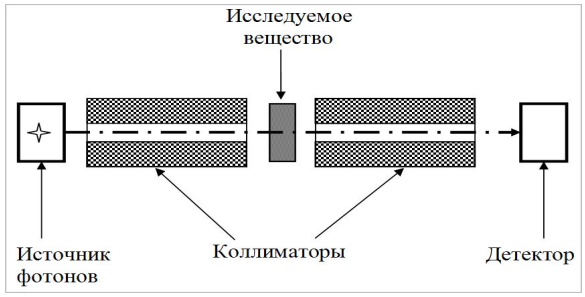

Рис. 1. Схема проведения эксперимента

В ходе каждого измерительного эксперимента все фотоны имеют некоторую энергию E_k из фиксированного (дискретного) набора энергий излучения

$$0.1 \text{ МэВ} = E_1 < E_2 < \ldots < E_{\overline{N}} = 20 \text{ МэВ}; \quad N \leqslant \overline{N}. \tag{1}$$

Далее при проведении численных экспериментов, мы будем пользоваться числовыми данными для конкретных веществ и энерий, взятых из таблиц [8], где приводятся данные для 20 значений энергии, поэтому мы ограничимся случаем $\overline{N} = 20$.

В таблице 1 приведен полный перечень значений этих энергий E_j с указанием порядкового номера энергии j.

Таблица 1. Соответствие между номером энергии излучения j и значением энергии E_j(МэВ)

j	1	2	3	4	5	6	7	8	9	10
E_j(МэВ)	0.1	0.15	0.2	0.3	0.4	0.5	0.6	0.8	1	1.25
j	11	12	13	14	15	16	17	18	19	20
E_j(МэВ)	1.5	2	3	4	5	6	8	10	15	20

Пусть $h_k = h(E_k)$ — плотность потока излучения, входящего в G_0, $H_k = H(E_k)$ — плотность потока излучения, выходящего из G_0, для энергии E_k, $k = 1, \ldots, N$, $\mu_{0k} = \mu_0(E_k)$ — коэффициент ослабления излучения для вещества X_0, $\mu_{xik} = \mu_{xi}(E_k)$ — коэффициенты ослабления излучения для X_i, $i = 1, \ldots, N$, ρ_0 — плотность вещества X_0, ρ_{xi} — плотность X_i, w_i — массовая доля элемента X_i, входящего в состав вещества X_0. Химический состав вещества X_0 образца G_0 нам не известен и подлежит определению по результатам измерения входящих и выходящих из G_0 потоков излучения $h(E_k)$ и $H(E_k)$ для известных E_k.

Расположение коллиматоров перед исследуемым веществом и после него позволяет выделить из начального потока излучения преимущественно только те фотоны, которые не вступали во взаимодействие с веществом и своей начальной энергии не потеряли. В таком случае и при введенных обозначениях уравнение переноса излучения [6, 7] принимает простой вид и его следствием является экспоненциальное затухание плотности потока излучения в веществе. Поэтому для каждого значения энергии E_k мы можем записать равенство $H_k = h_k \exp(-l\mu_{0k})$; $\quad k = 1, \ldots, N$, откуда $\ln(H_k/h_k) = -l\mu_{0k}$, или

$$-l\mu_{0k} = \ln(H_k/h_k).$$

Согласно [8, 9], коэффициент ослабления μ_{0k} для вещества X_0 связан с коэффициентами ослабления $\mu_{xik} = \mu_{xi}(E_k)$ входящих в его состав элементов $X_1, \ldots, X_N$ равенством

$$\mu_{0k} = \rho_0 \sum_{i=1}^{N} w_i \frac{\mu_{xik}}{\rho_{xi}}.$$

Подставляя это выражение в предыдущую формулу, получаем следующую систему уравнений

$$\sum_{i=1}^{N} \frac{\mu_{xik}}{\rho_{xi}} \cdot (l\rho_0 w_i) = \ln \frac{h_k}{H_k}; \quad k = 1, \ldots, N. \tag{2}$$

В этой системе известная величина l намеренно помещена не в правую часть, а в левую, для того чтобы все уравнения были безразмерными.

Массовые доли w_i элементов X_i, входящих в состав вещества X_0, по своему определению удовлетворяют соотношению

$$\sum_{i=1}^{N} w_i = 1 \tag{3}$$

и условиям

$$w_i \geqslant 0; \quad i = 1, \ldots, N. \tag{4}$$

В итоге задача нахождения химического состава формулируется так.

Задача 1. *Найти величины* $\rho_0, w_i, i = 1, \ldots, N$, *удовлетворяющие уравнениям* (2), (3) *и неравенствам* (4) *при условии, что все остальные величины, входящие в* (2), *нам известны.*

Обоснованность и целесообразность такой постановки задачи химии достаточно подробно обсуждалась в [4, 5]. Перепишем (2) в виде $Ax = b$, или

$$\sum_{i=1}^{N} A_{ki} x_i = b_k; \qquad k = 1, \ldots, N, \tag{5}$$

где $A_{ki} = \mu_{xik}/\rho_{xi} = \mu_{xi}(E_k)/\rho_{xi}$, $x_i = l\rho_0 w_i$, $b_k = \ln(h_k/H_k)$. Отметим, что в такой записи коэффициенты матрицы A есть массовые коэффициенты ослабления излучения для химических элементов, и матрица не зависит от геометрии задачи.

Далее будем рассматривать (5) как систему линейных алгебраических уравнений, в которой A и b известны, а x — неизвестный вектор, $x^T = (l\rho_0 w_1, \ldots, l\rho_0 w_N)$.

Исследования, проведенные ранее в [4] для различных групп элементов $X_1, \ldots, X_N$ для значений N от 2 до 10, показали, что матрица A остается невырожденной практически при любом выборе значений энергий просвечивания образца $E_1, \ldots, E_N$, входящих в набор (1). Далее будем считать, что во всех рассматриваемых случаях матрица A не вырождена, тогда задача химии имеет решение и оно единственно [5].

Вместе с этим в [4, 5] было показано, что даже небольшие ошибки правой части системы (5), вызванные неточностью измерений входящего и выходящего потоков излучения h_k и H_k, могут привести к значительным ошибкам решения. При этом, с одной стороны, величина максимально возможной ошибки решения системы (5) быстро растет с ростом числа N, а с другой, при заданном N и фиксированном наборе $X_1, \ldots, X_N$ существенно зависит от выбора набора энергий $E_1, \ldots, E_N$ и может при этом меняться на несколько порядков. Для изучения вопроса о возможной ошибке решения задачи химии вернемся к рассмотрению системы уравнений (5) и введем следующие обозначения.

Пусть $b = b_T + \delta b$, где b_T — вектор "точных" значений правой части системы (5), то есть таких значений $b_k = \ln(h_k/H_k)$, $k = 1, \ldots, N$, которые мы имели бы в случае отсутствия ошибок измерения, δb — вектор возмущения, вызванный измерительными ошибками, $x = x_T + \delta x$, где x_T — точное, а x — возмущенное решение (5), так что выполняются равенства $A(x_T + \delta x) = b_T + \delta b$, $Ax_T = b_T$ и $A(\delta x) = \delta b$. Тогда

$$\delta x = A^{-1} \delta b. \tag{6}$$

Напомним известные факты о сингулярном разложении матрицы [10, 11]. Пусть $\lambda_1, \ldots, \lambda_N$ — собственные числа матрицы $G = A^T A$, а $v_1, \ldots, v_N$ — собственные векторы G, соответствующие $\lambda_1, \ldots, \lambda_N$. Поскольку A не вырождена, то все $\lambda_i > 0$. Пусть V

и U есть $N\times N$ матрицы такие, что столбцы V образованы векторами $v_1,...,v_N$, а столбцы U — векторами $u_i = Av_i/\sqrt{\lambda_i}$. Тогда справедливо разложение

$$A = USV^T, \tag{7}$$

где S есть диагональная матрица, $S = diag\{\sigma_1,...,\sigma_N\}$; $\sigma_i = \sqrt{\lambda_i}$; $i = 1,...,N$. При этом обе матрицы V и U ортогональные и справедлива формула

$$A^{-1} = VS^{-1}U^T, \tag{8}$$

причем $S^{-1} = diag\left\{\sigma_1^{-1},...,\sigma_N^{-1}\right\}$. Числа σ_i называются сингулярными числами матрицы A, а векторы $v_1,...,v_N$ — сингулярными векторами матрицы A.

Далее для удобства будем считать, что собственные числа матрицы G всегда занумерованы так, что $\lambda_1 \leqslant \lambda_2 \leqslant ... \leqslant \lambda_N$, тогда $\sigma_1^{-1} \geqslant \sigma_2^{-1} \geqslant ... \geqslant \sigma_N^{-1}$.

Спектральную норму матрицы A обозначим $||A||_S$, она имеет вид [10]

$$||A||_S = \max_{||x||\,=\,1} ||Ax|| = \sqrt{\lambda_{\max}} = \sigma_{\max}, \tag{9}$$

а число обусловленности C_S матрицы A определяется формулой $C_S = C_S(A) = ||A||_S \cdot ||A^{-1}||_S$ и равно

$$C_S = \sqrt{\lambda_{\max}/\lambda_{\min}} = \sigma_{\max}/\sigma_{\min}. \tag{10}$$

Из равенства $\delta x = A^{-1}\delta b$ и (8) при $\delta b = u_i$ получаем

$$\delta x^{(i)} = A^{-1}u_i = VS^{-1}U^Tu_i = \sigma_i^{-1}v_i. \tag{11}$$

Отсюда несложно увидеть, что если в качестве множества возмущений правой части системы (5) взять единичную сферу $\Omega = \left\{\delta b \,\middle|\, \delta b \in \mathbb{R}^N, ||\delta b|| = 1\right\}$, то множество $M = A^{-1}(\Omega) \subset \mathbb{R}^N$, состоящее из возмущений решения x_T, будет эллипсоидом, главные полуоси которого направлены вдоль сингулярных векторов $v_1,...,v_N$ и по длине равны $\sigma_1^{-1},...,\sigma_N^{-1}$.

Пусть также $\widehat{\Omega} = \left\{\delta b \,\middle|\, \delta b \in \mathbb{R}^N,\ ||\delta b|| \leqslant 1\right\}$ и $\widehat{M} = A^{-1}(\widehat{\Omega})$.

2. Построение множества, содержащего решение задачи

Далее мы будем использовать следующие обозначения.

Через $E^{(p)} = \left(E_1^{(p)}, E_2^{(p)}, ..., E_N^{(p)}\right)$ будет обозначаться какой-нибудь вектор, сформированный из поднабора энергий $\left\{E_1^{(p)}, E_2^{(p)}, ..., E_N^{(p)}\right\}$, входящих в набор (1) и удовлетворяющих условию $E_1^{(p)} < E_2^{(p)} < ... < E_N^{(p)}$. Нетрудно увидеть, что всего существует $C_N^{20} = \frac{20!}{N!\cdot(20-N)!}$ таких векторов и они естественным образом (лексикографически) упорядочены. Таким образом, верхний индекс p в записи $E^{(p)}$ будет указывать на порядковый номер такого вектора при данном упорядочивании. Подробнее это поясняется в описании после таблицы 2.

Каждому вектору $E^{(p)}$ соответствует (невырожденная) матрица $A = A\left(E^{(p)}\right)$ системы (5), элементы которой $A_{ki} = \mu_{xik}/\rho_{xi} = \mu_{xi}\left(E_k^{(p)}\right)/\rho_{xi}$, $k,i = 1,...,N$.

Будем проводить серии измерительных экспериментов по просвечиванию образца. Каждая серия состоит из N экспериментов, в которых образец просвечивается на каждой энергии $E_1^{(p)}, E_2^{(p)}, \ldots, E_N^{(p)}$ и находятся (с ошибками) величины $b_k = \ln(h_k/H_k)$; $k=1,\ldots,N$. По завершении n-й серии мы получаем вектор $b^{(p;n)} = b_T^{(p)} + \delta b^{(p;n)}$. Здесь и далее в верхних индексах первое число (p) указывает на то, что данный вектор был получен для набора энергий $E^{(p)}$, а второе число (n) указывает на то, что данный вектор был получен в n-й серии экспериментов. Вектор точного результата серии экспериментов $b_T^{(p)}$ не зависит от номера серии, но зависит от набора энергий $E^{(p)}$, на которых происходило просвечивание. Векторы $b_T^{(p)} = \left(b_{T,1}^{(p)}, \ldots, b_{T,N}^{(p)}\right)$ и $\delta b^{(p;n)} = \left(\delta b_1^{(p;n)}, \ldots, \delta b_N^{(p;n)}\right)$ по отдельности нам не известны, но мы будем предполагать следующее:

1) для каждой энергии $E_k^{(p)}$, $k=1,\ldots,N$ нам известно число $r_k > 0$ такое, что при любом n

$$\left|\delta b_k^{(p;n)}\right| = \left|b_{T,k}^{(p)} - b_k^{(p;n)}\right| \leqslant r_k; \tag{12}$$

2) ошибки измерений $\delta b_k^{(p;n)}$ для каждого k и n могут принимать любое значение на промежутке $[-\,r_k,\ \ r_k]$, причем для любых ε и k таких, что $0 < \varepsilon \leqslant r_k$, $1 \leqslant k \leqslant N$, существуют целые $m = m(\varepsilon, k)$ и $n = n(\varepsilon, k)$ такие, что $\delta b_k^{(p;m)} \in [-r_k, -r_k + \varepsilon]$ и $\delta b_k^{(p;n)} \in [r_k - \varepsilon, r_k]$;

3) измерительный прибор не имеет систематических ошибок.

Таким образом, согласно условию 2), при желании, выполнив достаточно измерений, мы можем сколь угодно близко подойти к любой грани параллелепипеда, в котором находятся все возможные ошибки $\delta b^{(p;i)}$ правой части системы (5).

Обозначим также

$$\Pi^{(p)} = \{z|\ z = (z_1, \ldots, z_N) \in \mathbb{R}^N,\ |z_k| = r_k, \quad k = 1, \ldots, N\},$$
$$\widehat{\Pi}^{(p)} = \{z|\ z = (z_1, \ldots, z_N) \in \mathbb{R}^N,\ |z_k| \leqslant r_k, \quad k = 1, \ldots, N\},$$
$$T^{(p)} = A^{-1}(\Pi^{(p)}), \quad \widehat{T}^{(p)} = A^{-1}(\widehat{\Pi}^{(p)}), \quad x_T = A^{-1}(b_T^{(p)}), \quad \delta x^{(p;n)} = A^{-1}(\delta b^{(p;n)}).$$

Проведем n-ю серию экспериментов, просвечивая образец на энергиях $E_1^{(p)}$, $E_2^{(p)}, \ldots, E_N^{(p)}$, получим вектор $b^{(p;n)} = b_T^{(p)} + \delta b^{(p;n)}$. Отсюда справедлива формула

$$b_T^{(p)} \pm \delta b^{(p;n)} \in b_T^{(p)} + \widehat{\Pi}^{(p)}, \tag{13}$$

откуда, в частности,

$$b_T^{(p)} \in b_T^{(p)} + \delta b^{(p;n)} + \widehat{\Pi}^{(p)}, \tag{14}$$

значит $A^{-1}\left(b_T^{(p)}\right) \in A^{-1}\left(b_T^{(p)}\right) + A^{-1}\left(\delta b^{(p;n)}\right) + A^{-1}\left(\widehat{\Pi}^{(p)}\right)$, то есть

$$x_T \in x_T + \delta x^{(p;n)} + \widehat{T}^{(p)}. \tag{15}$$

Векторы x_T и $\delta x^{(p;n)}$ по отдельности нам не известны, но известен вектор $x_T + \delta x^{(p;n)}$ и множество $\widehat{T}^{(p)}$, поэтому для любого $E^{(p)} = \left(E_1^{(p)}, E_2^{(p)}, \ldots, E_N^{(p)}\right)$ и любой

n-й серии экспериментов мы можем указать множество $x_T+\delta x^{(p;n)}+\widehat{T}^{(p)}$, содержащее искомый вектор x_T. Пусть

$$P_n = \bigcap_{i=1}^{n} (b_T^{(p)} + \delta b^{(p;i)} + \widehat{\Pi}^{(p)}), \tag{16}$$

$$Q_n = \bigcap_{i=1}^{n} (x_T + \delta x^{(p;i)} + \widehat{T}^{(p)}), \tag{17}$$

тогда $b_T^{(p)} \in P_n \subset P_{n-1} \subset ... \subset P_2 \subset P_1$, $x_T \in Q_n \subset Q_{n-1} \subset ... \subset Q_2 \subset Q_1$ и

$$x_T = A^{-1}(b_T^{(p)}) \in A^{-1}(P_n) = Q_n. \tag{18}$$

Если среди набора полученных данных $b^{(p;1)},...,b^{(p;n)}$ будет много таких, для которых ошибки $\delta b^{(p;i)}$ находятся близко к границе $\partial\widehat{\Pi}^{(p)} = \Pi^{(p)}$, и эти ошибки достаточно равномерно распределены по углам, то множество Q_n, содержащее x_T, будет "маленьким". В этом случае мы можем сказать, что искомое решение задачи x_T хорошо локализовано. Несложно увидеть, что в "идеальном" случае, если произойдет так, что точки $b^{(p;1)} = b_T^{(p)} + \delta b^{(p;1)}$ и $b^{(p;2)} = b_T^{(p)} + \delta b^{(p;2)}$ окажутся в противоположных вершинах параллелепипеда $\Pi^{(p)}$ (а значит $\delta b^{(p;1)} = -\delta b^{(p;2)}$), мы получим

$$Q_2 = \bigcap_{i=1}^{2} (x_T + \delta x^{(p;i)} + \widehat{T}^{(p)}) = x_T,$$

то есть для нахождения точного решения задачи достаточно двух "удачных" измерений. Довольно курьезно, что наиболее "ценными" являются именно те измерения, которые содержат наибольшие ошибки. И напротив, измерения с маленькими ошибками $\delta b^{(p;i)}$ вносят небольшой вклад в локализацию решения.

Проиллюстрируем сказанное на простом примере, в котором в качестве неизвестного вещества X_0 был взят гидрид бария BaH_2, так что в данном случае $N=2$. Выбор этого вещества (для иллюстрации) вызван главным образом тем, что для него можно сделать достаточно хороший рисунок. В таблице 2 приводятся основные характеристики матрицы $A = A(E^{(k)})$ для BaH_2 для ряда значений вектора $E^{(k)}$.

Таблица 2. Некоторые основные характеристики матрицы $A = A(E^{(k)})$ для BaH_2 для некоторых значений вектора $E^{(k)}$.

Набор энергий $E^{(k)}$	Длины полуосей σ_1^{-1}, σ_2^{-1} и угол наклона φ	Число C_S
$E^{(169)} = (13, 20)$	$\sigma_1^{-1} = 34.85$, $\sigma_2^{-1} = 11$, $\varphi = -50.8°$	$C_S = C_{Smin} = 3.16$
$E^{(136)} = (10, 11)$	$\sigma_1^{-1} = 16397$, $\sigma_2^{-1} = 5.98$, $\varphi = -65.9°$	$C_S = C_{Smax} = 2741$
$E^{(2)} = (1, 3)$	$\sigma_1^{-1} = 5.45$, $\sigma_2^{-1} = 0.443$, $\varphi = -8.6°$	$C_S = 12.3$

В первой строке таблицы запись $E^{(169)}=(13,20)$ означает, что просвечивание происходило на энергиях E_{13} и E_{20} из перечня (1) — эта пара образует 169-й набор при

естественном упорядочивании всевозможных выборок двух энергий из (1). Пользуясь таблицей 1, можно увидеть, что $E_{13}=3$ МэВ, $E_{20}=20$ МэВ. Далее идут значения главных полуосей эллипса возмущения решений σ_1^{-1} и σ_2^{-1}, затем угол φ между осью x_1 и направлением главной полуоси эллипса и, наконец, число обусловленности C_S для матрицы $A(E^{(169)})$.

Во второй строке таблицы приводятся те же данные для набора энергий $E^{(136)}=$ $=(10{,}11)$, при котором число обусловленности матрицы $A=A(E^{(136)})$ имеет наибольшее значение среди всех возможных наборов (всего же их для данного случая 190).

В третьей строке приводятся данные для набора энергий $E^{(2)}=(1{,}3)$, при котором наибольшая полуось $\sigma_1^{-1}=5.45$ эллипса $A^{-1}(\Omega)$ имеет наименьшую длину. Именно такой набор энергий целесообразно использовать в процессе нахождения множеств P_n и Q_n. Далее на рисунке 2 представлены одновременно несколько графиков. Они строились для случая $E^{(169)}=(13{,}20)$, поскольку эксцентриситет эллипса при этом минимален и рисунок воспринимается лучше, чем в случае набора энергий $E^{(2)}$ $=$ $=$ $(1{,}3)$.

В центре находится квадрат $\Pi^{(169)}$, вершины которого составляют множество $\bigl\{z|\ z=(z_1,z_2)\in\mathbb{R}^2,\ |z_k|=1/\sqrt{2},\ k=1,2\bigr\}$. Нетрудно увидеть, что этот квадрат вписывается в окружность $\Omega=\bigl\{\delta b|\ \delta b\in\mathbb{R}^2,||\delta b||=1\bigr\}$, которая на рисунке не показана. В квадрате $\Pi^{(169)}$ находятся двадцать векторов — возмущений правой части $\bigl\{\delta b^{(169;1)},\ldots,\delta b^{(169;20)}\bigr\}$, которые были получены с помощью генератора случайных чисел.

Эллипс есть множество $M=A^{-1}(\Omega)$, внутри него находится косоугольный параллелепипед $T^{(169)}=A^{-1}(\Pi^{(169)})$, внутри которого можно заметить двадцать векторов — возмущений решения $\bigl\{\delta x^{(169;1)},\ldots,\delta x^{(169;20)}\bigr\}$. О множествах P_{20} и Q_{20} будет сказано позже.

Обратимся снова к процессу нахождения множеств P_n и Q_n. Пусть

$$\Bigl\{b^{(p;1)},\ldots,b^{(p;n)}\Bigr\} \tag{19}$$

есть полный набор данных, полученных в результате n серий экспериментов, $n\geqslant 2$. Выберем из него N пар векторов, которые обозначим (f_1,g_1), $(f_2,\mathrm{g}_2),\ldots,(f_N,\mathrm{g}_N)$, таких, что f_1 есть тот вектор $b^{(p;i)}$ из набора (19), у которого первая координата $b_1^{(p;i)}$ максимальна среди всех первых координат векторов из (19), а g_1 есть тот вектор $b^{(p;j)}$ из (19), у которого первая координата $b_1^{(p;i)}$ минимальна среди всех первых координат векторов из (19). У вектора f_2 максимальна вторая координата, а у g_2 минимальна вторая координата и так далее. Ясно, что такие точки лежат на выпуклой оболочке множества (19) и вносят наибольший вклад в локализацию решения.

Для пары (f_1,g_1) построим пару гиперплоскостей, проходящих через точки f_1- $-(r_1,0,\ldots,0)$ и $\mathrm{g}_1+(r_1,0,\ldots,0)$ и ортогональных первой координатной оси. Для пары (f_2,g_2) построим пару гиперплоскостей, проходящих через точки $f_2-(0,r_2,0,\ldots,0)$ и $\mathrm{g}_2+(0,r_2,0,\ldots,0)$ и ортогональных второй координатной оси и так далее.

Таким образом, для векторов $f_k=(f_{k1},\ldots,f_{kN})$ и $\mathrm{g}_k=(\mathrm{g}_{k1},\ldots,\mathrm{g}_{kN})$ справедливы

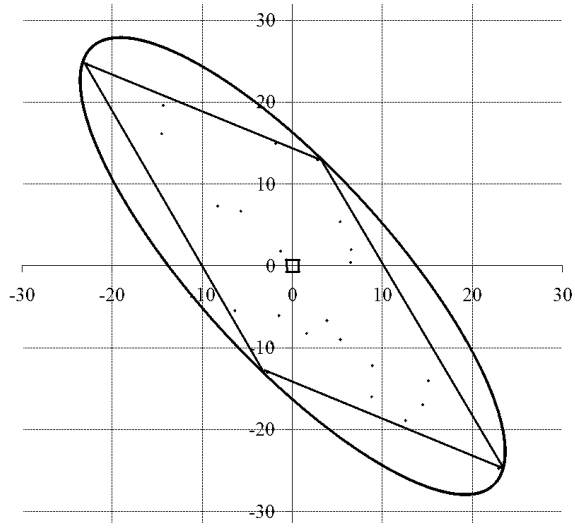

Рис. 2. Множество $A^{-1}(\Omega)$ — эллипс, $A^{-1}(\Pi)$ — параллелепипед, Π — квадрат со стороной $\sqrt{2}$ в центре рисунка.

равенства

$$f_{kk} = \max_{i\,=\,1,\ldots,n} b_k^{(p;i)}, \qquad \mathrm{g}_{kk} = \min_{i\,=\,1,\ldots,n} b_k^{(p;i)}, \quad k = 1, \ldots, N. \tag{20}$$

Пусть

$$P = \left\{ z \;\middle|\; z \in \mathbb{R}^N,\ f_{kk} - r_k \leqslant z_k \leqslant \mathrm{g}_{kk} + r_k, \quad k = 1, \ldots, N \right\}, \tag{21}$$

тогда справедливо следующее почти очевидное утверждение.

Утверждение 1. *Множество P совпадает с множеством P_n, определенным формулой* (16)

Доказательство. Из (16) несложно увидеть, что

$$P_n = \left\{ z \;\middle|\; z \in \mathbb{R}^N,\ b_k^{(p;i)} - r_k \leqslant z_k \leqslant b_k^{(p;i)} + r_k,\ k = 1, \ldots, N,\ i = 1, \ldots, n \right\}. \tag{22}$$

Зафиксируем произвольную точку $z \in P_n$, тогда $b_k^{(p;i)} - r_k \leqslant z_k \leqslant b_k^{(p;i)} + r_k$. Взяв в этих неравенствах сначала $\max_{i\,=\,1,\ldots,n}$ от левой части, а затем $\min_{i\,=\,1,\ldots,n}$ от правой части из равенств (20) получим неравенства $f_{kk} - r_k \leqslant z_k \leqslant \mathrm{g}_{kk} + r_k$, значит $P_n \subset P$.

Наоборот, если для произвольной точки $z \in P$ выполняются все неравенства

$$f_{kk} - r_k \leqslant z_k \leqslant \mathrm{g}_{kk} + r_k, \quad k = 1, \ldots, N,$$

то из равенств (20) тем более будут выполняться все неравенства

$$b_k^{(p;i)} - r_k \leqslant z_k \leqslant b_k^{(p;i)} + r_k, \quad k = 1, \ldots, N,\ i = 1, \ldots, n.$$

Таким образом, $P \subset P_n$ и утверждение доказано. □

В итоге множество $P = P_n$ всегда содержит точку $b_T^{(p)}$ и оказывается прямоугольным параллелепипедом, а множество $Q_n = A^{-1}(P_n) = A^{-1}(P)$, содержащее точку $x_T = A^{-1}(b_T^{(p)})$, также оказывается параллелепипедом, но, как правило, не прямоугольным. После нахождения всех вершин P_n несложно найти вершины Q_n и, при необходимости, диаметр Q_n и проекции $pr_k(Q_n)$ на k-ю координатную ось пространства решений.

Пусть

$$x_{L,k} = \min_{x \in Q_n} x_k, \quad x_{R,k} = \max_{x \in Q_n} x_k,$$

$x_a = (x_{a,1}, \ldots, x_{a,N})$, $x_{a,k} = 0.5(x_{L,k} + x_{R,k})$, $k = 1, \ldots, N$, тогда, поскольку Q_n выпукло, $x_a \in Q_n$. Вектор x_a можно назвать приближенным решением задачи, построенным по результатам n серий измерений. Ясно, что для любой точки $x \in Q_n$ (и, в частности, для $x = x_T$)

$$||x_a - x|| \leqslant \frac{1}{2} \sum_{k=1}^{N} \left[(x_{R,k} - x_{L,k})^2 \right]^{1/2}.$$

Теперь вернемся ненадолго к рисунку 1, а точнее, к множеству P_{20}, о котором там говорилось. По результатам проведенных расчетов из формул (20) были получены следующие числовые значения: $f_{11} = 0.5650$, $g_{11} = -0.5075$, $f_{22} = 0.6928$, $g_{22} = -0.6539$. Поскольку $r_1 = r_2 = 1/\sqrt{2} = 0.7071$, то из (26) получаем

$$P = \left\{ z \;\middle|\; z \in \mathbb{R}^2,\ -0.1421 \leqslant z_k \leqslant 0.1996,\ -0.0143 \leqslant z_k \leqslant 0.0532 \right\}.$$

В итоге первоначальный квадрат $\widehat{\Pi}^{(p)}$ уменьшает свои размеры по первой и второй осям примерно в 4 и 20 раз соответственно. Во столько же раз должны сократиться длины сторон параллелепипеда $A^{-1}(\widehat{\Pi}^{(p)})$.

Отметим кратко следующее. Несложно увидеть, что при выполнении предположений 1)–3) из утверждения 1 и формул (20), (21) следует, что с ростом числа серий измерений n последовательность множеств Q_n стягивается к точному решению x_T системы (5). Однако на практике предположения 1)–3) не выполняются, в частности, все числа $r_1, \ldots, r_N$ экспериментатору известны лишь приблизительно. Поэтому, увеличивая количество серий измерений n, мы не сможем локализовать искомый вектор $x_T \in Q_n$ со сколь угодно высокой точностью.

3. Замечание о нормальном обобщенном решении задачи

По результатам данных, полученных в n сериях измерений, можно построить нормальное обобщенное решение задачи x_{opt} [10]. Для этого рассмотрим переопределенную систему уравнений

$$Ax = b^{(p;1)}$$
$$Ax = b^{(p;2)}$$
$$\ldots$$
$$Ax = b^{(p;n)},$$

которую запишем в виде

$$\widehat{A}x = \widehat{b}, \tag{23}$$

где $\widehat{A}$ — прямоугольная ("вертикальная") матрица размером $nN \times N$, а $\widehat{b}$ — вектор-столбец размером nN, составленный из векторов $b^{(p;1)}, \ldots, b^{(p;n)}$. Тогда

$$x_{opt} = \widehat{A}^{+} \cdot \widehat{b}, \tag{24}$$

где $\widehat{A}^{+}$ — матрица размером $N \times nN$, псевдообратная к матрице $\widehat{A}$. Естественно задать вопрос: всегда ли это решение x_{opt} принадлежит множеству Q_n? В общем случае ответ отрицательный. Более того, такое решение (оно существует всегда) может оказаться намного хуже вышеупомянутого приближенного решения x_a в том смысле, что $||x_{opt} - x_T||$ может оказаться намного больше, чем $||x_a - x_T||$.

Ниже мы построим (абстрактный) пример того, когда это происходит, а предварительно напомним определение псевдообратной матрицы Мура – Пенроуза [12] и докажем одно несложное, но полезное утверждение общего характера, которое нам потребуется при построении примера.

Определение 1. Пусть $C = C(n, m)$ — комплексная матрица размером $n \times m$ и $C^* = C^*(m, n)$ — сопряженная с ней матрица. Матрица $C^+ = C^+(m, n)$ называется псевдообратной к матрице C, если существуют такие квадратные матрицы $F = F(m, m)$ и $G = G(n, n)$, что выполняются равенства

$$CC^+C = C, \qquad (a)$$
$$C^+ = FC^*, \qquad (b)$$
$$C^+ = C^*G. \qquad (c)$$

Справедливо следующее утверждение.

Утверждение 2. *Если $C = C(N, N)$ — (комплексная) невырожденная матрица, а $\widehat{C} = \widehat{C}(nN, N)$ — "вертикальная" матрица, составленная из n экземпляров матрицы C, тогда матрица $\widehat{C}^{+} = \widehat{C}^{+}(N, nN)$, псевдообратная к $\widehat{C}$, имеет вид*

$$\widehat{C}^{+} = \frac{1}{n}\left(C^{-1}, \ldots, C^{-1}\right)$$

Доказательство. состоит в несложной проверке равенств $(a), (b), (c)$, при этом

$$F(N, N) = \frac{1}{n} C^{-1} \cdot C^{*-1},$$

а матрица $G(nN, nN)$ состоит из $n \times n$ блоков вида

$$\frac{1}{n^2} C^{*-1} \cdot C^{-1}.$$

□

Пример. *Рассмотрим систему* (23), *в которой $N = 2$, $n = 10$ и $A = I$ — единичная матрица размером 2×2, тогда из утверждения 2 следует равенство $\widehat{A}^{+} = 0.1 \cdot (I, \ldots, I)$. Пусть точное решение x_T системы* (5) *(нам не известное) есть:*

$x_T = (1,1)$, максимальные измерительные ошибки $r_1 = r_2 = 1$ и в результате десяти серий измерений ($n = 10$) были получены следующие значения для правых частей системы (23)*: $b^{(1)} = (1.99, 0.01)$; $b^{(2)} = (0.01, 1.99)$; $b^{(k)} = (b_1^{(k)}, b_2^{(k)})$, где*

$$b_1^{(k)} = 1.9 + 0.08 \cos \frac{\pi(k-1)}{4}; \quad b_2^{(k)} = 1.9 + 0.08 \sin \frac{\pi(k-1)}{4}; \quad k = 3, \dots, 10.$$

Тогда для компонент $x_{opt} = (x_{opt,1}, x_{opt,2})$ из (24) *и утверждения 2 получаем*

$$x_{opt,1} = \frac{1}{10}[b_1^{(1)} + b_1^{(2)} + \sum_{k=1}^{8} b_1^{(k)}] = 1.72; \quad x_{opt,2} = 1.72.$$

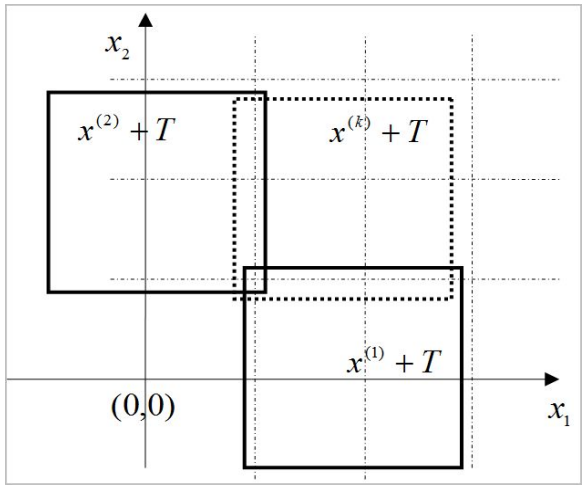

Рис. 3. Построение множества Q_2 в примере.

Теперь обратимся к рис. 3. На нем схематично показан процесс нахождения приближенного решения x_a. Согласно представленным данным в правой нижней части рисунка находится множество $x^{(1)} + T$, в левой верхней части — множество $x^{(2)} + T$. Пересечение этих двух квадратов есть множество Q_2. Это квадрат, центр которого – точка (1,1); $Q_2 = \{x|\ x \in \mathbb{R}^2,\ 0.99 \leqslant x_1 \leqslant 1.01,\ 0.99 \leqslant x_2 \leqslant 1.01\}$. На рис. 3 пунктиром приближенно показано место, в котором располагаются все остальные множества $x^{(k)} + T$. Несложно увидеть, что $Q_2 \subset x^{(k)} + T$; $k = 3, \dots, 10$, поэтому цепочка включений $Q_1 \supset Q_2 \supset \dots \supset Q_{10}$ стабилизируется на множестве Q_2 и приближенное решение $x_a = (1,1)$. Таким образом, в данном примере $x_a = x_T$, $||x_T - x_{opt}|| \approx 1.018$, что значительно больше, чем диаметр $d(Q_2) \approx 2.83 \cdot 10^{-2}$ множества Q_2, содержащего точное решение x_T. Отметим также, что случай системы (23) *сводится к только что рассмотренному путем n-кратного поблочного умножения* (23) *слева на матрицу A^{-1}.*

Заключение

В заключение отметим, что предложенный метод повышения точности определения химического состава исследуемого образца G_0 может быть использован тогда, когда у экспериментатора есть возможность многократно проводить серии просвечиваний образца, когда ранее сформулированные условия 1)–3) достаточно хорошо выполняются и все числа $r_1, \ldots, r_N$ известны. В этом случае можно надеяться на заметное повышение точности вычислений.

Список литературы

[1] Sergei Osipov, Sergei Chakhlov, Andrey Batranin, Oleg Osipov, Van Bak Trinh, Juriy Kytmanov, "Theoretical study of a simplified implementation model of a dual-energy technique for computed tomography", *NDT and E International*, **98**, (2018), 63–69.

[2] S. P. Osipov, V. A. Udod, Yanzhao Wang, "Identification of Materials in X-Ray Inspections of Objects by the Dual-Energy Method", *Russian Journal of Nondestructive Testing*, **53 (8)**, (2017), 568–587.

[3] В. А. Клименов, С. П. Осипов, А. К. Темник, "Идентификация вещества объекта контроля методом дуальных энергий", *Дефектоскопия*, **11**, (2013), 40–50.

[4] В. Г. Назаров, "Метод сингулярного разложения матрицы в задаче нахождения химического состава среды", *Сибирские электронные математические известия*, **14**, (2017), 821–837.

[5] В. Г. Назаров, "Выбор оптимальных значений энергии излучения в задаче нахождения химического состава среды", *Математическое моделирование*, **30**, (2018), 91–102.

[6] Д. С. Аниконов, А. Е. Ковтанюк, И. В. Прохоров, *Использование уравнения переноса в томографии*, Логос, М., 2000.

[7] D. S. Anikonov, A. E. Kovtanyuk, I. V. Prokhorov, *Transport Equation and Tomography*, VSP, Utrecht-Boston, 2002.

[8] J. H. Hubbell, S. M. Seltzer, *Tables of X – Ray Mass Attenuation Coefficients and Mass Energy Absorption Coefficients 1 Kev to 20 Mev for Elements Z = 1 to 92 and 48 Additional Substances of Dosimetric Interest*, Preprint NISTIR-5632, Nat. Inst. of Standard and Technology, Gaithersburg, 1995.

[9] M. J. Berger, J. H. Hubbell, S. M. Seltzer, J. Chang, J. S. Coursey, R. Sukumar, D. S. Zucker, "XCOM: Photon Cross Section Database. National Institute of Standards and Technology. Gaithersburg. MD.", 2005, http://www.physics.nist.gov/xcom.

[10] С. К. Годунов, А. Г. Антонов, О. П. Кирилюк, В. И. Костин, *Гарантированная точность решения систем линейных уравнений в евклидовых пространствах*, Наука. Сиб. отд-ние, Новосибирск, 1988.

[11] Дж. Форсайт, М. Малькольм, К. Моулер, *Машинные методы математических вычислений*, Мир, М., 1980.

[12] Ф. Р. Гантмахер, *Теория матриц*, Наука, М., 1988.

Поступила в редакцию
5 апреля 2018 г.

Nazarov V. G. On increase of calculation accuracy at the problem of determining the chemical composition of a medium. *Far Eastern Mathematical Journal*. 2018. V. 18. No 2. P. 219–232.

ABSTRACT

In work the problem of a finding of a chemical composition of a homogeneous medium by a method of collimated multiple X-ray irradiation is considered. The question on possibility of increase of accuracy of the solution is thus studied by carrying out of several series of repeated measurements of passing radiation. It is shown that under some natural assumptions on measurement errors of passing radiation, the solution errors tend to zero with growth of number of the measurements fulfilled. Results of the calculations executed for concrete substance are shown by way of illustration.

Key words: *radiography of a continuous medium, finding the chemical composition of a medium, singular value decomposition, calculation accuracy.*