

Оценки, минимизирующие возможность потерь, и минимаксные оценки: сравнительный анализ*

А. И. Чуличков, Б. Юань

achulichkov@gmail.com

Физический Факультет МГУ им. М.В.Ломоносова, Москва, Ленинские горы, Дом 1, строение 2

Поставлена и решена задача оценивания значений функции в заданных точках области ее определения по результатам измерений конечного набора ее функционалов, выполненных с погрешностью. Показано, что с конечной погрешностью может быть оценена только конечномерная составляющая искомой функции, предложена точная конечномерная модель, позволяющая построить искомые оценки. Обсуждаются два метода оценивания. Первый метод минимизирует максимально возможную погрешность оценивания каждого значения функции в заданной точке. Считается, что погрешность измерения каждого линейного функционала с одной и той же возможностью принимают любое значение внутри заданного интервала. Для каждого оцениваемого значения функции построен интервал, которому может принадлежать это значение. Минимаксной оценкой является середина этого интервала, а погрешностью оценки – половина его длины. Концы каждого интервала определяются решениями задач линейного программирования. Второй метод оценивания основан на теоретико-возможностной модели измерений, в которой считается, что большие значения погрешности измерения каждого функционала менее возможны, чем малые. Критерием оценивания является возможность потерь. Метод оценивания минимизирует этот критерий и сводится к решению задачи линейного программирования. Оценки минимальной возможности потерь сравниваются с оценками, минимизирующими максимальную погрешность каждого значения функции. Обсуждаются различия минимаксных оценок и оценок минимальной возможности потерь. Приведен пример оценивания параметров реального спектрометрического эксперимента.

Ключевые слова: минимаксные оценки; теория возможностей; оценки минимальной возможности потерь; линейное программирование; анализ экспериментальных данных.

Estimation of minimum possibility of losses and minimax estimation: a comparative analysis*

A. Chulichkov, B. Yuan

Faculty of Physics M.V. Lomonosov Moscow State University, Moscow, Russia

The estimation of function values at specified points in its domain of definition based on the measurement results of the finite set of functionals is posed and solved. The measurements are distorted by a finite error. It is shown that with the finite error can be estimated only finite-dimensional component of unknown function. Exact finite-dimensional model, underlying the construction of required assessments is proposed. Two methods for estimation are discussed. The first method minimizes the maximal error of the estimation of each value of the function at a given point. It is believed that the measurement error of each linear functional may take any value within a given interval. For each of the estimated value of a function the interval that contains this value was constructed. The minimax estimate is the midpoint of this interval, and the error is the half of its length. The ends of each interval are determined

*Работа выполнена при финансовой поддержке РФФИ, проект № 14-07-00409

as solutions of linear programming problems. The second method of estimation is based on theoretical-possibility measurement models. It is believed that large values of the measurement error of each functional less possible than small. The estimation criterion is the possibility of losses. Estimation method minimizes this criterion and is reduced to the solution of a linear programming problem. The estimates of the minimum possible losses and the estimates that minimize the maximum error of each value of function is compared. Differences between the minimax estimates and estimates of the minimum possible losses are discussed. An example of estimation of specter based on the data of real spectrometric experiment is given.

Keywords: minimax estimation; the theory of possibilities; estimates of the minimum possible losses; linear programming; analysis of experimental data.

Введение

В современных экспериментальных исследованиях часто возникает необходимость восстанавливать характеристики исследуемых объектов по данным измерений, причем изучаемые характеристики лишь косвенно связаны с результатом измерительных экспериментов. Для таких исследований типичной является схема

$$\xi = Ag + \nu, \quad (1)$$

интерпретируемая следующим образом: результат измерения ξ есть искаженный шумом ν выходной сигнал Ag измерительного прибора A , на вход которого подан сигнал g от измеряемого объекта. Задача интерпретации измерения (1) состоит в том, чтобы извлечь из ξ наиболее точную версию входного сигнала g или вычислить значение заданной функции Ug . Исходными данными для решения такой задачи является математическая модель схемы измерений (1) и данные измерений ξ .

Примером таких экспериментов является измерение спектра электромагнитного излучения с помощью спектрометра [1]. В этом случае входным сигналом g является спектр электромагнитного излучения $g(\cdot)$, выходной сигнал $q = Ag$ спектрометра формируется согласно соотношению

$$q(\lambda) = \int_0^{\infty} a(\lambda, \lambda')g(\lambda')d\lambda', \quad \lambda \in [0, \infty). \quad (2)$$

Здесь $a(\cdot, \cdot)$ — аппаратная функция спектрометра, ее смысл состоит в том, что при подаче на вход спектрометра монохроматического спектра единичной интенсивности с длиной волны λ' на выходе спектрометра получим спектр $a(\lambda, \lambda')$, $\lambda \in [0, \infty)$.

Одним из широко распространенных подходов к решению задачи интерпретации измерений (1) состоит в решении интегрального уравнения Фредгольма 1 рода (2) на основании известной аппаратной функции $a(\cdot, \cdot)$ и данных измерений функции $q(\cdot)$, выполненных с погрешностью ν ; погрешность ν при этом считается либо ограниченной по норме [2], либо обладающей известными стохастическими свойствами [3, 4]. Однако, как показано в работах [2, 5, 6], такая задача может оказаться некорректно поставленной по Адамару, в частности, может оказаться неразрешимой, либо иметь неединственное решение, либо ее решение (псевдорешение [5], если уравнение $Ag = \xi$ неразрешимо) обладает неустойчивостью по отношению к возмущению функции $q(\cdot)$ или математической модели измерений. Для решения некорректно поставленных задач были предложены методы регуляризации,

основная идея которых состоит в наложении дополнительных ограничений на класс решений, в результате решение регуляризованной задачи оказывается единственным, и стремится к точному при стремлении погрешности измерения ν к нулю [7, 8, 9, 10]. Несмотря на значительные успехи в создании методов решения таких задач, интерес к ним не угасает и в настоящее время [11, 12].

Поиск регуляризованного решения часто производится минимизацией суммы функционала невязки $\|Ag - \xi\|$ и сглаживающего функционала $\Omega(g)$:

$$g_\alpha = \arg \min \{ \|Ag - \xi\| + \alpha \Omega(g) \}, \quad \alpha \geq 0, \quad (3)$$

так, чтобы при согласованном стремлении $\alpha \rightarrow 0$ и $\nu \rightarrow 0$ решение g_α стремилось к точному решению (или к псевдорешению).

Такой подход к решению задачи интерпретации измерений встречается ряд принципиальных трудностей. Во-первых, как правило, экспериментатора интересует наиболее *точная оценка* сигнала g (или сигнала $u = Ug$, где U — заданное преобразование). В задаче (3) минимизируемый функционал не имеет отношения к точности решения. Во-вторых, в реальных ситуациях погрешность не стремится к нулю, и возникают вопросы с выбором параметра $\alpha > 0$.

Другой подход к решению задач интерпретации измерений дает теория измерительно-вычислительных систем (ИВС), развиваемая в школе проф. Ю.П.Пытьева [13]. В теории ИВС эти задачи ставятся как задачи поиска такого преобразования R измерения ξ , результат которого наиболее близок к $u = Ug$. При этом сигнал $u = Ug$ интерпретируется как выходной сигнал Ug «идеального» прибора U , на вход которого подан сигнал g . Конкретная постановка задач оценивания сигнала u зависит от математической модели измерения (1) и критериев точности оценки. Например, если задать модель измерения (1), указав множества, которым принадлежат сигналы g и ν , $g \in G$, $\nu \in N$, то оценкой $R\xi = \hat{u}$ сигнала $u = Ug$ является

$$\hat{u} = \arg \inf_u \sup_g \{ \|u - Ug\| \mid \xi = Ag + \nu, \nu \in N, g \in G \}. \quad (4)$$

Оценка \hat{u} называется минимаксной, она минимизирует максимально возможную погрешность оценки u [13, 14].

Однако минимаксные оценки в силу постановки задачи (4) минимизируют ошибку в самых неблагоприятных ситуациях, которые, на взгляд исследователя, могут показаться не вполне реальными. Поясним это на простейшем примере.

Пусть в (1) результатом измерения является вектор ξ линейного пространства размерности, равной двум, с координатами (0.5, 3), линейный оператор A задан матрицей $A = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$, на координаты векторов g и ν наложены ограничения: $|\nu_i| \leq 1$, $g \in G = \{(g_1, g_2) : 0 \leq g_i \leq 1, i = 1, 2\}$. На рисунке 1, слева, показана область возможных значений вектора g в виде единичного квадрата. Образ этой области при отображении A показан на рисунке 1, справа, в виде параллелограмма. Точкой обозначен вектор ξ с координатами (0.5, 3). Координаты вектора Ag отличаются от координат ξ не более, чем на единицу, с другой стороны, принадлежат образу множества G при отображении A — т.е. области на плоскости в виде треугольника ABC . Полный прообраз этой области показан на рисунке 1, слева, в виде треугольника $A'B'C'$. Таким образом, результат измерения $\xi = (0.5, 3)$ в рамках принятой математической модели измерения, если и только если первая координата вектора g изменяется от 0 до 1, а вторая — от 0.5 до 1. Минимаксная

оценка первой координаты вектора g равна $\hat{g}_1 = 0.5$ с погрешностью 0.5, а вторая $\hat{g}_2 = 0.75$ с погрешностью 0.25. Однако образ $A\hat{g} = (0.5, 2)$ отличается от результата измерения $\xi = (0.5, 3)$ на вектор, одна из координат которого равна 1. Это означает, что получить результат ξ в измерении (1) при $g = \hat{g}$ можно только при максимальной измерительной погрешности второй координаты вектора Ag .

Реализация максимальной измерительной погрешности может показаться исследователю весьма неправдоподобной, поскольку измерительные эксперименты стараются организовать так, чтобы погрешности измерения были как можно меньше.

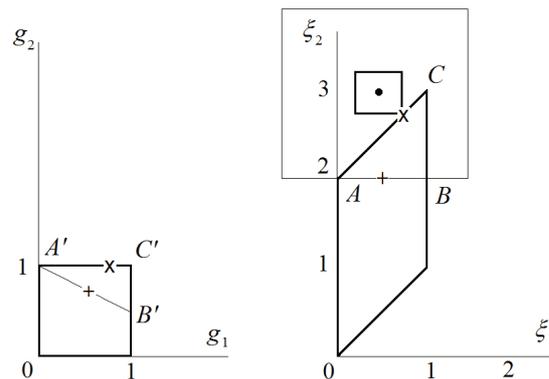


Рис. 1. Минимаксное оценивание вектора g в (1) и оценка, учитывающая, что малые значения погрешности измерений более возможны, чем большие

Учесть последнее замечание можно, отказавшись от предположения о том, что нет никаких предпочтений в значениях погрешности измерений, и указав, какие из их значений более возможны (правдоподобны), а какие — менее. В частности, если считать, что малые значения координат погрешности более возможны, чем большие, то точка Ag из треугольника ABC , координаты которой наиболее близки к координатам ξ , соответствует измерению с минимальными координатами измерительной погрешности. В этом смысле эта точка Ag наиболее правдоподобна из всех точек треугольника ABC . На рисунке 1, справа, это точка касания квадрата с центром в точке ξ с треугольником ABC , ее прообраз показан на рисунке 1, слева, знаком "x".

Математический формализм, позволяющий при оценивании параметров учесть предпочтения в значениях параметров математической модели измерения, дается в теории возможностей [15].

Впервые теория возможностей, основанная на теории нечетких множеств [16], была предложена Л.Заде в работе [17]. В отличие от «четкого» множества A , задаваемого индикатором $\chi_A(\cdot)$: $\chi_A(x) = 1$ для $x \in A$ и $\chi_A(x) = 0$ для $x \notin A$, нечеткое множество задается функцией принадлежности $\mu_A(\cdot)$, принимающей значение на отрезке $[0, 1]$, при этом если $\mu_A(x) = 0$, то элемент x не может принадлежать нечеткому множеству A , и чем больше $\mu_A(x)$, тем более возможным является утверждение, что $x \in A$; равенство $\mu_A(x) = 1$, означает, что включение $x \in A$ вполне возможно. Функция принадлежности интерпретируется как функция распределения возможностей включения x в A . Далее математическая теория возможностей развивалась в работах Д. Дюбуа, Г.Прадта [18, 19], Р. Ягера [20], Ю.П.Пытьева [15] и др.

В работах Ю.П.Пытьева предлагается трактовать возможность события как относительную оценку истинности данного события, его предпочтительности в сравнении с любым другим. Значение возможности события используется лишь для того, чтобы сравнить ее с возможностью любого другого события, указав, какое из двух более возможно или заключив, что они равновозможны. Это существенно расширяет возможности моделирования нечеткости, так как все распределения, значения которых связаны строго монотонно возрастающим преобразованием, оказываются эквивалентными. В отличие от вероятности, возможность не имеет событийно-частотной интерпретации, которая связывает её с экспериментом. Тем не менее теория возможностей позволяет математически моделировать реальность на основе опытных фактов, знаний, гипотез, суждений исследователей.

Математической моделью нечеткости в [15] является пространство с возможностью $(\Omega, P(\Omega), P)$, где Ω есть пространство элементарных событий, $P(\Omega)$ есть алгебра всех подмножеств Ω , и $P(\cdot) : P(\Omega) \rightarrow [0, 1]$ — возможностьная мера. Возможность $P(\cdot)$ определена как мера на $P(\Omega)$, и принимает значение на множестве $[0, 1]$, на котором определены бинарные операции «сложения» $+$ и «умножения» \bullet согласно равенствам $a+b = \max\{a, b\}$, $a \bullet b = \min\{a, b\}$, $(a, b) \in [0, 1] \times [0, 1]$. В работе [15] показано, что возможность любого события $A \in P(\Omega)$ может быть представлена в виде

$$P(A) = \sup_x \min\{p(x), \chi_A(x)\}, \quad (5)$$

где $\chi(\cdot)$ — индикатор A , $p(\cdot)$ — распределение возможностей, определенное как возможность одноточечного события $p(x) = P(\{x\})$, $x \in \Omega$. Формула (5), как и многие другие факты теории возможностей, получены из аналогичных формул теории вероятностей заменой операции интегрирования на \sup , суммы на \max , умножения — на \min . Так, например, события A и B независимы в теоретико-возможностном смысле, если $P(AB) = \min\{P(A), P(B)\}$; заметим, что это отличает рассматриваемый вариант теории возможностей от теории Л.Заде, в которой последнее равенство постулируется для всех событий A, B .

В теории возможностей аналогом случайных элементов некоторого линейного нормированного пространства \mathcal{R} являются нечеткие элементы, определенные как любая функция $\varphi(\cdot)$, заданная на Ω и принимающая значение в \mathcal{R} . Нечеткий элемент φ характеризуется распределением возможностей своих значений, заданным функцией $\pi^\varphi(\cdot)$, определенной на \mathcal{R} и принимающей значения в отрезке $[0, 1]$, значение $\pi^\varphi(y) = P(\{\varphi = y\})$ по определению есть возможность того, что нечеткий элемент φ примет значение y , $y \in \mathcal{R}$. На функцию $\pi^\varphi(\cdot)$ накладывается условие нормировки: $\max_{y \in \mathcal{R}} \{\pi^\varphi(y)\} = 1$.

Для того, чтобы формализовать утверждение «малые погрешности измерений более возможны, чем большие», достаточно считать погрешность $\nu \in \mathcal{R}_1$ в (1) нечетким элементом линейного нормированного пространства \mathcal{R}_1 и задать распределение возможностей его значений $\nu = z$ как монотонно убывающую функцию нормы $z \in \mathcal{R}_1$: $\pi^\nu(z) = \pi_0(\|z\|)$, $z \in \mathcal{R}_1$. Априорное знание $g \in G \subset \mathcal{R}_2$ в (4) о возможных значениях элемента $g \in \mathcal{R}_2$ может быть задано распределением возможности $\pi^g(y) = \chi_G(y)$, $y \in \mathcal{R}_2$. Если g и ν независимы, то совместное распределение возможностей пары $(g, \nu) \in \mathcal{R}_2 \times \mathcal{R}_1$ дается равенством $\pi^{g, \nu}(y, z) = \min\{\pi^g(y), \pi^\nu(z)\} = \min\{\chi_G(y), \pi_0(\|z\|)\}$. Если выполнено (1), то совместное

распределение тройки $(\xi, g, \nu) \in \mathcal{R}_2 \times \mathcal{R}_2 \times \mathcal{R}_1$ дается равенством

$$\begin{aligned} \pi^{\xi, g, \nu}(x, y, z) &= \begin{cases} \min\{\pi^g(y), \pi^\nu(z)\}, & \text{если } x = Ay + z, \\ 0, & \text{если } x \neq Ay + z, \end{cases} = \\ &= \begin{cases} \min\{\chi_G(y), \pi_0(\|x - Ay\|)\}, & \text{если } x = Ay + z, \\ 0, & \text{если } x \neq Ay + z. \end{cases} \end{aligned}$$

Распределение возможности пары нечетких элементов ξ, g дается "интегрированием":

$$\pi^{\xi, g}(x, y) = \sup_z \pi^{\xi, g, \nu}(x, y, z) = \min\{\chi_G(y), \pi_0(\|x - Ay\|)\}. \quad (6)$$

В общем случае в [15] решение о значении оцениваемой величины g определяется как нечеткий элемент δ с распределением переходной возможности $\pi^{\delta|\xi}(\cdot|\cdot)$, так, что при заданном значении $\xi = x$ значение $\pi^{\delta|\xi}(d|x)$ есть возможность считать d оценкой для g . Однако на практике в ряде случаев можно использовать "четкое" правило оценивания, задав функцию $d(\cdot)$, которая каждому значению $\xi = x$ ставит в соответствие оценку $d(x)$ нечеткого элемента g .

В [15] с каждым значением d оценки g предлагается связать значение возможности потерь $l(d, g)$, возникающих при использовании оценки d вместо истинного значения g . Тогда возможность потерь при решающем правиле $d(\cdot)$ равна $PL(d(\cdot)) = \sup_{x, g} \min\{l(d(x), g), \pi^{\xi, g}(x, y)\}$. Оптимальная оценка, минимизирующая возможность потерь, есть решение задачи на минимум

$$PL(d_*(\cdot)) = \arg \min_{d(\cdot)} PL(d(\cdot)). \quad (7)$$

В [15] показано, что оптимальную оценку можно получить, решая задачу (7) при каждом значении $\xi = x$.

В настоящей статье на примере решения задачи интерпретации данных спектрометрического эксперимента сравниваются два подхода, один из них основан на минимаксном оценивании параметров входного спектра, другой — на минимизации возможности потерь. В работе решение задачи интерпретации данных эксперимента (1) сводится к конечномерным задачам линейного программирования, при этом вместо замены операции интегрирования приближенной суммой, широко используемой при решении такого типа задач, предлагается точный метод дискретизации разложением по системе функций, полной в подпространстве, ортогональном нуль-пространству линейного интегрального оператора в (2). Показано, что в задачах интерпретации данных спектрометрии минимаксный подход не дает удовлетворительной оценки входного спектра, однако предположение о том, что большие значения измерительной погрешности менее возможны, чем малые, существенно улучшают оценку.

Математическая модель измерений

Пусть в (2) функции $g(\cdot)$, $a(\lambda, \cdot)$ являются элементами евклидова пространства $\mathcal{L}^2([0, \infty))$ при каждом $\lambda \in [0, \infty)$. Однако на практике в экспериментальных исследованиях измеряется не функция $q(\lambda)$, $\lambda \in [0, \infty)$, а конечный набор чисел. Если функция $a(\cdot, \cdot)$ задана своим представителем, являющимся непрерывной функцией на $[0, \infty) \times [0, \infty)$, то эти числа рассматриваются как значения функции $q(\cdot)$ в точках $\lambda_1, \dots, \lambda_n$. Результаты измерения значений $q(\lambda_i)$ сопровождаются погрешностью (шумом) ν_i , $i = 1, \dots, n$. В

частности, в рассматриваемом примере результатом измерения спектра являются числа ξ_1, \dots, ξ_n , интерпретируемые как искаженные шумом значения спектра на выходе спектрометра на длинах волн $\lambda_1, \dots, \lambda_n$:

$$\xi_i = q(\lambda_i) + \nu_i = \int_0^{\infty} a(\lambda_i, \lambda') g(\lambda') d\lambda' + \nu_i, \quad \lambda \in [0, \infty). \quad (8)$$

Заметив, что $\int_0^{\infty} a(\lambda_i, \lambda') g(\lambda') d\lambda' = (a_i, g)$, где $a_i \in \mathcal{L}^2([0, \infty))$ представлен непрерывной функцией $a(\lambda_i, \cdot)$, $i = 1, \dots, n$, а (\cdot, \cdot) — скалярное произведение в $\mathcal{L}^2([0, \infty))$, запишем (8) в виде

$$\xi_i = (a_i, g) + \nu_i, \quad i = 1, \dots, n, \quad (9)$$

элементы a_1, \dots, a_n считаем линейно независимыми. Погрешности измерений ν_i принадлежат заданным интервалам:

$$|\nu_i| \leq \varepsilon_i, \quad i = 1, \dots, n. \quad (10)$$

Как видно из (9), оценке с конечной погрешностью поддается лишь проекция $g \in \mathcal{L}^2([0, \infty))$ на линейную оболочку S_A элементов $a_1, \dots, a_n \in \mathcal{L}^2([0, \infty))$, совпадающую с ортогональным дополнением к нуль-пространству интегрального оператора в (8). Поскольку $a_i = a(\lambda_i, \cdot)$, $i = 1, \dots, n$, непрерывны, то элемент $Pg \in S_A$ можно задать его представителем $Pg(\cdot) = \sum_{i=1}^n \varphi_i a(\lambda_i, \cdot)$, для которого определены его значения в любой точке $\lambda' \in [0, \infty)$. Здесь $\varphi_1, \dots, \varphi_n$ — коэффициенты разложения проекции $Pg \in \mathcal{L}^2([0, \infty))$ по системе линейно независимых элементов $a(\lambda_i, \cdot) \in \mathcal{L}^2([0, \infty))$, $i = 1, \dots, n$.

Представив в (9) g в виде $g = Pg + (I - P)g$ и учтя, что $(I - P)g$ ортогонально всем a_i , $i = 1, \dots, n$, получим $\xi_i = \sum_{j=1}^n (a_i, a_j) \varphi_j + \nu_i$. Итак, результат измерения определяется значением конечного числа параметров $\varphi_1, \dots, \varphi_n$. Теперь ξ_i в (9) удобно представить как искаженный шумом ν_i результат измерения скалярного произведения двух n -мерных векторов $\vec{\varphi}$ и \vec{a}_i , заданных своими координатами: $\vec{\varphi} = (\varphi_1, \dots, \varphi_n)$, $\vec{a}_i = ((a_i, a_1), \dots, (a_i, a_n))$, $i = 1, \dots, n$:

$$\xi_i = (\vec{a}_i, \vec{\varphi})_n + \nu_i, \quad (11)$$

здесь $(\cdot, \cdot)_n$ — скалярное произведение в R^n .

Пусть требуется оценить значение проекции $Pg(\cdot)$ в точках $\lambda'_1, \dots, \lambda'_N \in [0, \infty)$. Эти значения будем рассматривать как координаты u_j вектора $\vec{u} \in \mathcal{R}^N$. Тогда

$$u_j = Pg(\lambda_j) = \sum_{i=1}^n \varphi_i a(\lambda_i, \lambda_j), \quad j = 1, \dots, N, \quad (12)$$

здесь φ_i , $i = 1, \dots, n$ — коэффициенты разложения Pg по линейно независимым элементам a_i , $i = 1, \dots, n$. Введя векторы $\vec{w}_j \in \mathcal{R}^n$, $i = j, \dots, N$, заданные своими координатами $\vec{w}_j = (a(\lambda_j, \lambda_1), \dots, a(\lambda_j, \lambda_n))$, $j = 1, \dots, N$, запишем координаты вектора $\vec{u} \in \mathcal{R}^N$ как скалярные произведения

$$u_j = (\vec{w}_j, \vec{\varphi})_n, \quad j = 1, \dots, N. \quad (13)$$

Оценка значений $Pg(\lambda_j)$, $j = 1, \dots, N$, проекции элемента $g \in \mathcal{L}^2([0, \infty))$ на S_A в точках $\lambda'_1, \dots, \lambda'_N \in [0, \infty)$ теперь получается оцениванием конечного числа координат (13) вектора $\vec{u} \in \mathcal{R}^N$ по результату (11) измерений n линейных функционалов конечномерного вектора $\vec{\varphi} \in \mathcal{R}^n$. Априорные ограничения на искомую оценку спектра (например, неотрицательность координат вектора \vec{u}) могут быть заданы в виде линейных неравенств

$$l_j \leq u_j \leq r_j, \quad -\infty \leq l_j < r_j \leq \infty, \quad j = 1, \dots, N. \quad (14)$$

В частности, если известно, что все координаты u_1, \dots, u_N неотрицательны, то $l_j = 0$, $r_j = +\infty$, $j = 1, \dots, N$.

Минимаксная оценка значения координат вектора \vec{u}

Рассмотрим задачу, в которой оценка каждой координаты u_j вектора \vec{u} определяется из принципа минимизации максимальной погрешности оценки.

Согласно (13) $u_j = (\vec{w}_j, \vec{\varphi})_n$, а $\vec{\varphi} \in \mathcal{R}^n$ в силу (9) и (10) удовлетворяет системе линейных неравенств

$$|(\vec{a}_i, \vec{\varphi})_n - \xi_i| \leq \varepsilon_i, \quad i = 1, \dots, n, \quad (15)$$

и системе линейных неравенств априорных ограничений (14)

$$l_k \leq (\vec{w}_k, \vec{\varphi})_n \leq r_k, \quad k = 1, \dots, N, \quad (16)$$

Решение системы неравенств (15) и (16) есть выпуклое замкнутое множество, и множество значений проекций всех элементов этого множества на любой вектор $\vec{w}_j \in \mathcal{R}^n$ есть отрезок. Поэтому искомая координата $u_j = (\vec{w}_j, \vec{\varphi})_n$ принадлежит интервалу $u_j \in [u_{j,\min}, u_{j,\max}]$, левый и правый края которого определяются как минимальное и максимальное значение скалярного произведения $(\vec{w}_j, \vec{\varphi})_n$ при ограничениях, заданных системой линейных неравенств (15)-(16), т.е. решениями соответствующих задач линейного программирования.

Определив границы соответствующих интервалов, минимаксную оценку \hat{u}_j каждой координаты u_j , $j = 1, \dots, N$, вектора \vec{u} найдем из задачи на минимакс

$$\hat{u}_j = \arg \min_{u'_j} \max_{u_{j,\min} \leq u_j \leq u_{j,\max}} |u_j - u'_j|.$$

Ее решение, очевидно, есть середина интервала $[u_{j,\min}, u_{j,\max}]$, а погрешность — половина его длины: $\hat{u}_j = (u_{j,\min} + u_{j,\max})/2$, $h_j = (u_{j,\max} - u_{j,\min})/2$, $j = 1, \dots, N$.

Оценки максимальной возможности

Как показано в предыдущих разделах настоящей статьи, задача интерпретации измерений (8) сводится к задаче оценивания значений проекции элемента $g \in \mathcal{L}^2([0, \infty))$ на линейную оболочку S_A линейно независимых элементов a_1, \dots, a_n , заданных непрерывными функциями; ищется оценка значений проекции в точках $\lambda_1, \dots, \lambda_N \in [0, \infty)$. Эти значения, рассматриваемые как координаты вектора $\vec{u} \in \mathcal{R}^N$, связаны линейным преобразованием $U \in (\mathcal{R}^n \rightarrow \mathcal{R}^N)$ с вектором $\vec{\varphi} \in \mathcal{R}^n$. Матрица этого линейного оператора задана в (13): $U_{j,i} = a(\lambda_j, \lambda_i)$, $i = 1, \dots, n$, $j = 1, \dots, N$. Вектор $\vec{\varphi}$ связан с вектором

$\vec{\xi} \in \mathcal{R}^n$ (с результатом измерения) соотношением (11). Таким образом, если задать векторы $\vec{\xi}, \vec{\nu} \in \mathcal{R}^n$ координатами $\vec{\xi} = (\xi_1, \dots, \xi_n)$, $\vec{\nu} = (\nu_1, \dots, \nu_n)$ и определить линейный оператор $A \in (\mathcal{R} \rightarrow \mathcal{R}^n)$ его матрицей $A_{ik} = (a_i, a_k)$, $i, k = 1, \dots, n$, то получим соотношения

$$\vec{\xi} = A\vec{\varphi} + \vec{\nu}, \quad \vec{u} = U\vec{\varphi}. \quad (17)$$

Задача состоит в том, чтобы, зная математическую модель измерения (17) и его результат $\vec{\xi}$, найти оценку вектора \vec{u} .

В соответствии с описанной во введении методикой оценивания, минимизирующей возможность потерь, зададим, во-первых, нечеткую модель погрешности измерения $\vec{\nu}$. Считая его координаты ν_i , $i = 1, \dots, n$, независимыми нечеткими величинами, зададим распределение возможностей $\pi^{\nu_i}(\cdot)$ для каждой из них соотношением

$$\pi^{\nu_i}(z) = \begin{cases} \pi_0\left(\frac{|z|}{\varepsilon_i}\right), & |z| \leq \varepsilon_i, \\ 0, & |z| > \varepsilon_i, \end{cases} \quad (18)$$

где $\pi_0(\cdot)$ монотонно убывает на интервале $[0, 1]$. Это есть формальное выражение априорной уверенности экспериментатора в том, что большие погрешности менее возможны, чем малые, а также в том, что ошибки ν_i , по модулю превосходящие ε_i , невозможны. Тогда совместное распределение возможностей нечеткого вектора погрешности $\vec{\nu} = (\nu_1, \dots, \nu_n)$ равно $\pi^\nu(\vec{z}) = \min\{\pi^{\nu_1}(z_1), \dots, \pi^{\nu_n}(z_n)\} = \pi_0\left(\max_{i=1, \dots, n} \frac{|z_i|}{\varepsilon_i}\right)$. Заметим, что $\max_{i=1, \dots, n} \frac{|z_i|}{\varepsilon_i}$ как функция вектора $\vec{z} = (z_1, \dots, z_n)$ обладает всеми свойствами нормы \vec{z} , и будем в дальнейшем использовать для этой функции обозначение $\|\vec{z}\|_\varepsilon$. Таким образом, $\pi^\nu(\vec{z}) = \pi_0(\|\vec{z}\|_\varepsilon)$.

Во-вторых, зададим априорное распределение возможностей нечеткого вектора $\vec{\varphi}$ как индикатор подмножества векторов \mathcal{R}^n , координаты которых удовлетворяют неравенствам (16):

$$\pi^\varphi(\vec{y}) = \begin{cases} 1, & \text{если } l_k \leq (\vec{w}_k, \vec{y})_n \leq r_k, \quad k = 1, \dots, N, \\ 0, & \text{в противном случае.} \end{cases}$$

Это есть формальное выражение априорной уверенности исследователя в том, что значения искомой проекции элемента g на S_A принадлежат "четкому" множеству, заданному системой линейных неравенств (16).

Эти две модели с учетом (17) и в предположении независимости нечеткой погрешности измерений $\vec{\nu}$ и $\vec{\varphi}$ позволяют записать совместное распределение возможности пары нечетких векторов $\vec{\xi}$ и $\vec{\varphi}$ в виде $\pi^{\xi, \varphi}(\vec{x}, \vec{y}) = \min\{\pi^\nu(\vec{x} - A\vec{y}), \pi^\varphi(\vec{y})\}$, действуя так же, как описано во введении при выводе формулы (6).

Далее, так как $\vec{u} = U\vec{\varphi}$ с возможностью единица, и нарушение этого равенства невозможно, то $\pi^{\xi, \varphi, u}(\vec{x}, \vec{y}, \vec{z}) = \pi^{\xi, \varphi}(\vec{x}, \vec{y})$, если $\vec{u} = U\vec{\varphi}$, и $\pi^{\xi, \varphi, u}(\vec{x}, \vec{y}, \vec{z}) = 0$, если $\vec{u} \neq U\vec{\varphi}$, то совместное распределение нечетких векторов \vec{u} и $\vec{\xi}$ есть

$$\pi^{\xi, u}(\vec{x}, \vec{z}) = \sup_{\vec{y}} \begin{cases} \min\{\pi_0(\|\vec{x} - A\vec{y}\|_\varepsilon), \pi^\varphi(\vec{y})\}, & \text{если } \vec{z} = U\vec{y}, \\ 0, & \text{в противном случае.} \end{cases}$$

Заметим, что максимальное по значению функции $\pi^{\xi,u}(\vec{x}, \cdot)$, т.е. решение задачи $\vec{z}_*(\vec{x}) = \arg \max_{\vec{z} \in \mathcal{R}^N} \{\pi^{\xi,u}(\vec{x}, \vec{z})\}$, достигается при $\vec{z}_*(\vec{x}) = U\vec{y}_*(\vec{x})$, где $\vec{y}_*(\vec{x})$ — решение задачи

$$\vec{y}_*(\vec{x}) = \arg \max_{\vec{y} \in \mathcal{R}^n} \{\min\{\pi_0(\|\vec{x} - A\vec{y}\|_\varepsilon), \pi^\varphi(\vec{y})\}\}. \quad (19)$$

Обозначим $\vec{d}(\cdot) : \mathcal{R}_n \rightarrow \mathcal{R}^N$ функцию, для каждого значения $\vec{\xi} = \vec{x}$ определяющую оценку $\vec{u} = \vec{d}(\vec{x})$ вектора \vec{u} , и $l(\cdot, \cdot) : \mathcal{R}^n \times \mathcal{R}^n \rightarrow [0, 1]$ — распределение возможностей потерь: значение $l(\vec{d}, \vec{u})$ есть возможность потерь в ситуации, когда вместо значения $\vec{u} = \vec{z}$ используется его значение \vec{d} . Тогда, как описано во введении, возможность потерь при оценке $\vec{d}(\cdot)$ и при заданном значении измерения $\vec{\xi} = \vec{x}$ равна $PL(\vec{d}(\vec{x})) = \sup_{\vec{y}} \min\{l(\vec{d}(\vec{x}), \vec{y}), \pi^{\xi,u}(\vec{x}, \vec{y})\}$.

Оценка $\vec{d}_*(\cdot)$, минимизирующая максимальную возможность потерь, для каждого $\vec{x} \in \mathcal{R}^n$ определяется решением задачи на минимум $PL(\vec{d}_*(\vec{x})) = \min_{\vec{d}(\cdot)} PL(\vec{d}(\vec{x}))$.

Как показано в [15], если функция возможности потерь задана так, что нулевая возможность потерь соответствуют верному решению, а любое неверное решение влечет потери с возможностью единица, то оценку максимальной возможности потерь определяет значение $\vec{d}_*(\vec{x})$, которое доставляет максимум возможности $\pi^{\xi,u}(\vec{x}, \vec{y})$ для каждого $\vec{\xi} = \vec{x}$. Такие оценки называются оценками максимальной апостериорной возможности. Таким образом, задача построения оценки, минимизирующей возможность потерь, свелась к задаче на максимум (19), и для ее решения следует найти максимум $\vec{y}_*(\vec{x})$ функции

$$\min\{\pi_0(\|\vec{x} - A\vec{y}\|_\varepsilon), \pi^\varphi(\vec{y})\}. \quad (20)$$

и вычислить оценку $\vec{d}_*(\vec{x}) = U\vec{y}_*(\vec{x})$.

В силу монотонности функции $\pi_0(\cdot)$ задача на максимум (20) эквивалентна задаче на минимум

$$\min_{\vec{y} : (\vec{w}_m, \vec{y})_n \in [l_m, r_m], m=1, \dots, n} \left\{ \max_{i=1, \dots, n} \frac{\left| x_i - \sum_{k=1}^n A_{ik} y_k \right|}{\varepsilon_i} \right\}. \quad (21)$$

Задача (21) сводится к задаче линейного программирования [13, 21] и по сути является модифицированной задачей минимизации невязки при решении системы линейных алгебраических уравнений. Заметим, что если значение минимума функционала в (21) больше единицы, то возможность такой оценки в силу (18) равна нулю, что свидетельствует о неадекватности используемой математической модели.

Описание и свойства оценки максимальной возможности суммируем в следующем утверждении.

Теорема. Пусть в схеме измерений (9) $a_i(\cdot)$, $i = 1, \dots, n$ — непрерывные функции, квадрат которых интегрируем на $[0, \infty)$, ν_i , $i = 1, \dots, n$, — независимые нечеткие элементы с распределением возможностей (18). Тогда оценкой максимальной возможности значений ортогональной проекции $Pg(\cdot)$ функции $g(\cdot)$ на линейную оболочку S_A функций

$a_i(\cdot)$, $i = 1, \dots, n$, в точках $\{\lambda_1, \dots, \lambda_N\} \in [0, \infty)$, удовлетворяющих условиям (12), (14) является $P\hat{g}(x_i) = \sum_{k=1}^n y_{*j} a_k(x_j)$, $i = 1, \dots, N$, где y_{*i} , $i = 1, \dots, N$, — решение задачи (21). Если значение минимума функционала в (21) больше единицы, то математическая модель измерения (9) не согласуется с его результатом.

Оценки максимальной возможности параметров спектрометрического эксперимента

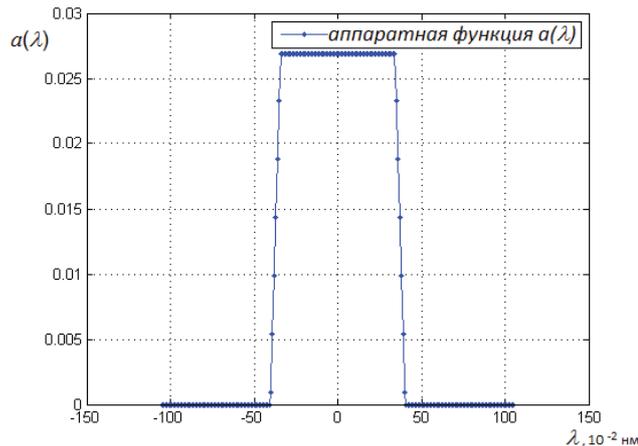


Рис. 2. График аппаратной функции двухщелевого спектрометра $a(\cdot)$, усл.ед.

Эффективность методов, предложенных в предыдущих разделах статьи, применялись для интерпретации результатов измерений интенсивности излучения газоразрядной лампы на длинах волн в окрестности эмиссионного спектра D -линии натрия (дублета с положением максимумов $\lambda_1 = 588.9950$ нм и $\lambda_2 = 589.5924$ нм). Измерения проводились на двулучевом спектрометре, ширины щелей спектрометра были выбраны так, чтобы аппаратная функция имела трапецевидную форму [1]. График аппаратной функции на интервале от -1 нм до 1 нм при фиксированных d_1 и d_2 приведен на рисунке 2, один отсчет по горизонтальной оси соответствует 0.01 нм. По вертикальной оси — условные единицы.

Спектр измерялся согласно схеме измерения $\xi(\lambda_i) = \int_{-\infty}^{\infty} a(\lambda - \lambda_i) f(\lambda) d\lambda + \nu_i$, $i = 1, \dots, 211$, в которой длины волн $\lambda_1, \dots, \lambda_{211}$ изменялись от 588.20 нм до 590.30 нм с шагом 0.01 нм. Результат измерения спектра — вектор с координатами $(\xi(\lambda_1), \dots, \xi(\lambda_{211}))$ — изображен на рисунке 3. Здесь по горизонтальной оси отложены значения двухсот одиннадцати длин волн, в которых измерялся спектр, по вертикальной оси отложена интенсивность измеренного спектра, равная числу зарегистрированных фотонов. Хотя экспериментальные данные были получены в результате реальных измерений, целью эксперимента была демонстрация возможностей математических методов интерпретации измерений, поэтому ширина аппаратной функции была специально выбрана существенно большей, чем расстояние между линиями дублета, в измеренном спектре они не разрешены.

Значение длин волн, в которых оценивалось значение входного спектра, были выбраны на том же отрезке от 588.20 до 590.30 нм с шагом 0.01 нм, таким образом, оцениваемый вектор имеет размерность 211. Считалось, что значение искомой проекции для всех 211 длин волн неотрицательны, то есть в (16) $u_{i,\min} = 0$, $u_{i,\max} = +\infty$, $i = 1, \dots, 211$.

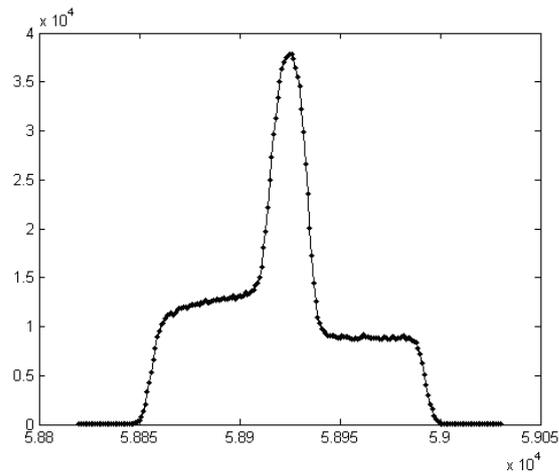


Рис. 3. Результат измерения спектра излучения Na : число зарегистрированных фотонов на каждой длине волны $\lambda_1, \dots, \lambda_{211}$

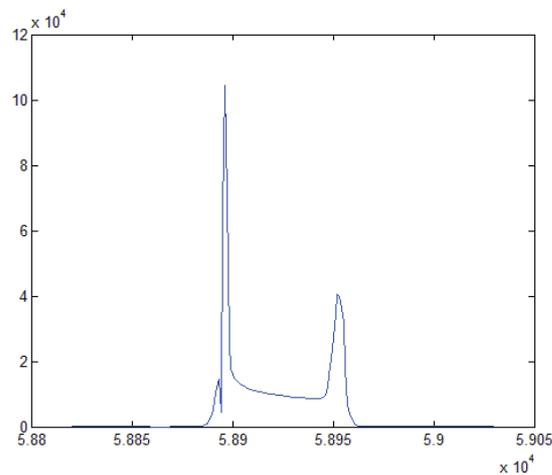


Рис. 4. Оценка входного спектра методом максимальной возможности, усл.ед.

Величина максимальной погрешности измерений ε_i для каждой длины волны λ_i считалась пропорциональной квадратному корню из ξ_i , $i = 1, \dots, 211$. Априорные ограничения на значения оцениваемого спектра состояли только в требовании неотрицательности.

Оценка значений проекции спектра на S_A в выбранных значениях длин волн, минимизирующая максимальную возможность потерь, приведена на рисунке 4. Видно, что вблизи максимумов дублета при $\lambda_1 = 588.9950$ нм и $\lambda_2 = 589.5924$ нм в оценке спектра имеются существенные максимумы.

Минимаксная оценка значений проекции спектра при тех же значениях длин волн приведена на рисунке 5. Верхняя кривая есть график $u_{j,\max}$ правой границы интервалов возможных значений оцениваемой координаты u_j , $i = 1, \dots, 211$. Нижний график (он практически совпадает с осью абсцисс) есть график левой границы $u_{j,\min}$ интервалов возможных значений оцениваемой координаты u_j , $j = 1, \dots, 211$. График свидетельствует, что практически для каждой координаты, рассматриваемой независимо от остальных координат, возможно ее значение, равное нулю. Минимаксная оценка координат u_j дается графиком, расположенным посередине между значениями $u_{j,\min}$ и $u_{j,\max}$. Из рисунка

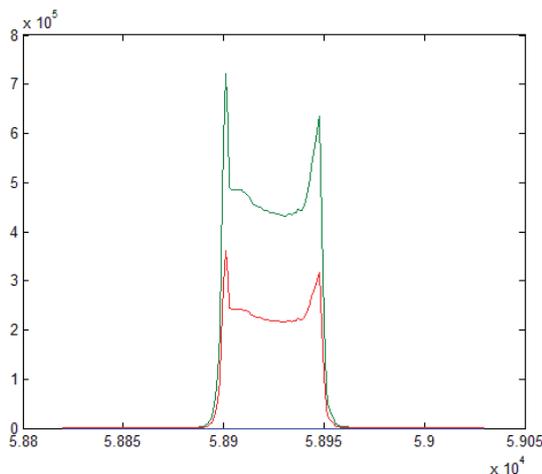


Рис. 5. Минимаксная оценка входного спектра, усл.ед.

видно, что в данных спектрометрического эксперимента содержится недостаточно информации для того, чтобы уверенно обнаружить дублет излучения Na .

Учет же дополнительной информации о том, что большие ошибки измерений менее возможны, чем малые, приводит к существенно лучшему результату, см. рисунок 4.

Заключение

В работе изучена задача оценивания входного сигнала измерительного прибора на основе математической модели, описывающей эксперимент, в котором измеряется конечное число линейных функционалов от входного сигнала прибора. Измерения сопровождаются погрешностью. Измеряемые в эксперименте линейные функционалы определяются как значение свертки входного сигнала с аппаратной функцией измерительного прибора; аппаратная функция считается непрерывной на своей области определения. Показано, что с конечной погрешностью можно оценить лишь конечномерную составляющую входного сигнала. Построены оценки, минимизирующие максимальную погрешность оценивания искомых значений, при этом используется априорная информация об оцениваемых величинах, заданная системой линейных неравенств, и предполагается, что измерительная погрешность может быть любой в пределах заданного интервала. Показано, что предположение о том, что большие погрешности измерений менее возможны, чем малые, могут существенно улучшить результат оценивания. Для формализации априорных предположений о том, что большие погрешности менее возможны, чем малые, построена возможностная модель измерения, поставлена и решена задача оценивания путем минимизации максимальной возможности потерь.

Эффективность методов иллюстрируются оценкой значений оптического спектра по данным, полученным с двухлучевого спектрометра с широкими ширинами щелей. Показано, что предположение о большей возможности малых значениях погрешности измерений позволяет получить более адекватное представление об измеряемом спектре, чем при оценивании, минимизирующем максимальную погрешность оценки.

Литература

- [1] Лебедева В. В. Экспериментальная оптика. 4-е изд. М.: Физический факультет, МГУ им. М.В.Ломоносова, 2005. 282 с.

- [2] Тихонов А. Н., Арсенин В. Я. Методы решения некорректных задач. М.: Наука, 1979.
- [3] Федотов А. М. Некорректные задачи со случайными ошибками в данных. Новосибирск: Наука, 1990.
- [4] Турчин В. А., Козлов В. П., Малкевич М. С. Использование методов математической статистики для решения некорректных задач // УФН, 1970. Т. 102, № 3. С. 345–385.
- [5] Иванов В. К., Васин В. В., Танана В. П. Теория линейных некорректных задач и ее приложения. М.: Наука, 1978.
- [6] Лаврентьев М. М. О некоторых некорректных задачах математической физики. Новосибирск: Изд-во СО АН СССР, 1962.
- [7] Тихонов А. Н., Гончарский А. В., Степанов В. В., Ягола А. Г. Численные методы решения некорректных задач. М.: Наука, 1990. 232 с.
- [8] Ягола А. Г., Ван Янфей, Степанова И. Э., Тутаренко В. Н. Обратные задачи и методы их решения. Приложения к геофизике. М.: Бином. Лаборатория знаний, 2014.
- [9] Самарский А. А., Вабищевич П. Н. Численные методы решения обратных задач математической физики. М.: Издательство ЛКИ, 2009. 480 с.
- [10] Ванник В. Н. Восстановление зависимостей по эмпирическим данным. М.: Наука, 1979.
- [11] Eichstadt S., Schmahling F., Wubbelier G., Kruger U., Elster C. A new approach to bandpass correction in spectrometer measurements using the Richardson-Lucy method // 16-th International Congress of Metrology, 2013. 14005.
- [12] Silva Neto A. J., Cella N. A regularized solution with weighted Bregman distances for the inverse problem of photoacoustic spectroscopy // Comput. Appl. Math, 2006. Vol.25. P. 139–165.
- [13] Пытьев Ю. П. Методы математического моделирования измерительно-вычислительных систем. М.: ФИЗМАТЛИТ, 2012. 428 с.
- [14] Лудов М. Л. Минимаксные методы оценивания. М.: Препринты ИПМ им. М.В.Келдыша. № 71, 2010. 87 с.
- [15] Пытьев Ю. П. Возможность как альтернатива вероятности. Математические и эмпирические основы, применение. М.: ФИЗМАТЛИТ, 2007. 464 с.
- [16] Zadeh L. Fuzzy Sets // Information and Control, 1965. Vol. 8. P. 235–350.
- [17] Zadeh L. Fuzzy sets as a basis for a theory of possibility // Fuzzy Sets and Systems, 1999. No. 100. P. 9–24.
- [18] Dubois D., Prade H. Possibility Theory, Probability Theory and Multiple-valued Logics: A Clarification // Annals of Mathematics and Artificial Intelligence, 2001. Vol. 32. P. 35–66.
- [19] Dubois D., Prade H. Formal representation of uncertainty // Decision-Making Process / Ed. by D. Bouyssou, D. Dubois, V. Pirlot, H. Prade. ISTE, London, 2009. P. 85–156.
- [20] Yager R. Conditional Approach to Possibility-Probability Fusion // IEEE Transactions on Fuzzy Systems, 2012. Vol. 20, No. 3. P. 526–535.
- [21] Кириллов К. В., Чуличков А. И. Редукция измерений в нечеткой модели эксперимента как решение задачи линейного программирования // Вестн. Моск. ун-та. Физ. Астрон, 1999. № 2. С. 62–64.

References

- [1] Lebedeva V. V. 2005. Experimental optics. 4th ed. M.: Faculty of Physics, Moscow State University. 282 p. (in Russ.)
- [2] Tikhonov A. N., Arsenin V. Y. 1979. Methods for solving ill-posed problems. M.: Nauka. (in Russ.)

- [3] Fedotov A. M. 1990. Ill-posed problems with random errors in the data. Novosibirsk: Nauka. (in Russ.)
- [4] Turchin V. A., Kozlov V. P., Malkevich M. C. 1970. The use of mathematical statistics methods for solving ill-posed problems. *Physics-Uspekhi (Advances in Physical Sciences)* 102(3):345–385. (in Russ.)
- [5] Ivanov V. K., Vasin V. V., Tanana V. P. 1978. The theory of linear ill-posed problems and its applications. M.: Nauka. (in Russ.)
- [6] Lavrent'ev M. M. 1962. On some ill-posed problems of mathematical physics. Novosibirsk: SD AS USSR Publ. (in Russ.)
- [7] Tikhonov A. N., Goncharsky A. V., Ctepanov V. V., Yagola A. G. 1990. Numerical methods for solving ill-posed problems. M.: Nauka. 232 p. (in Russ.)
- [8] Yagola A. G., Van Yunfey, Stepanova I. E., Titarenko V. N. 2014. Inverse problems and methods of their solution. Applications to Geophysics. M.: Binom. Laboratoria znanij. (in Russ.)
- [9] Samarsky A. A., Vabishevich P. N. 2009. Numerical methods for solving inverse problems of mathematical physics. M.: LKI Publisher. 480 p. (in Russ.)
- [10] Vapnik V. N. 1979. Restore dependency on empirical data. M.: Hayka. (in Russ.)
- [11] Eichstadt S., Schmahling F., Wubbelier G., Kruger U., Elster C. 2013. A new approach to bandpass correction in spectrometer measurements using the Richardson-Lucy method. *16-th International Congress of Metrology*. 14005.
- [12] Silva Neto A. J., Cella N. 2006. A regularized solution with weighted Bregman distances for the inverse problem of photoacoustic spectroscopy. *Comput. Appl. Math* 25:139–165.
- [13] Pyt'ev Yu. P. 2012. Mathematical modeling methods of measuring computer-aided systems. M.: FIZMATLIT. 428 p. (in Russ.)
- [14] Lidov M. L. 2010. Minimax estimation methods. M.: Preprint IPM. No. 71. 87 p. (in Russ.)
- [15] Pyt'ev Yu. P. 2007. The possibility as an alternative to probability. Mathematical and empirical basis, application. M.: FIZMATLIT. 464 p. (in Russ.)
- [16] Zadeh L. 1965. Fuzzy Sets. *Information and Control* 8:235–350.
- [17] Zadeh L. 1999. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems* 100:9–24.
- [18] Dubois D., Prade H. 2001. Possibility Theory, Probability Theory and Multiple-valued Logics: A Clarification. *Annals of Mathematics and Artificial Intelligence* 32:35–66.
- [19] Dubois D., Prade H. 2009. Formal representation of uncertainty. In: Bouyssou D., Dubois D., Pirlot M., Prade H. (eds) Decision-making process, ISTE, London. 85–156.
- [20] Yager R. 2012. Conditional Approach to Possibility-Probability Fusion. *IEEE Transactions on Fuzzy Systems* 20(3):526–535.
- [21] Kirillov K. V., Chulichkov A. I. 1999. The reduction of measurements in fuzzy model of the experiment as the solution of a linear programming problem. *Moscow University Physics Bulletin* 2:62–64. (in Russ.)