УДК 004.02

ПОСТАНОВКА ЗАДАЧ ОБРАБОТКИ ДАННЫХ НА ОСНОВЕ КРИТЕРИЯ МИНИМУМА ПРОТЯЖЕННОСТИ

Вовк С. М. – канд. физ.-мат. наук, доцент, доцент кафедры компьютерных наук и информационных технологий Днипровского национального университета имени Олеся Гончара, Днипро, Украина.

АННОТАЦИЯ

Актуальность. Для обработки данных, содержащих аномальные значения, и получения разреженных решений или решений малой протяженности может быть использовано требование минимизировать протяженность функции, применяемой для поиска решения. Объектом исследования в данной работе является процесс постановки задач обработки данных на основе указанного требования, которое далее упоминается как критерий минимума протяженности. Целью данной работы является разработка подхода к постановке задач обработки данных на основе данного критерия.

Метод. На основе критерия минимума протяженности предложен новый подход к постановке задач аппроксимации данных и к постановке обратных задач с прямым линейным оператором, решение которых имеет малую протяженность или является разреженным, в условиях, когда исходные данные содержат шум и аномальные значения. Постановка задачи аппроксимации получена путем задания параметрической модели данных и применения критерия минимума протяженности к невязке решения. Постановка обратной задачи получена путем применения критерия минимума протяженности к решению задачи и к его невязке. Представлены частные случаи этой постановки и отмечено, что она обобщает постановку задачи регуляризации Тихонова. Предложенные постановки задач сформулированы в виде задач минимизации соответствующих функционалов, построенных на основе «супермножества» стоимостных функций. В общем случае указанные функционалы не являются ни выпуклыми, ни унимодальными, и их минимизация может оказаться трудоемкой задачей.

Результаты. Предложенные постановки задач обобщают те постановки, которые выполнены на основе критериев наименьших квадратов и/или наименьших модулей. Численное моделирование задачи аппроксимации линейной функцией зашумленных данных в условиях наличия шума импульсного типа, а также в условиях наличия мешающего фрагмента экспоненциальной функции подтвердило целесообразность предложенной постановки и ее результативность. Численное моделирование обратной задачи, которой отвечала переопределенная система линейных алгебраических уравнений с грубыми ошибками в ее правой части и разреженным решением, также подтвердило целесообразность использования критерия минимума протяженности для ее постановки.

Выводы. Постановка задач обработки данных на основе критерия минимума протяженности является целесообразной в условиях, когда часть исходных данных является грубо искаженной и/или когда искомое решение имеет малую протяженность. Постановки, основанные на критерии минимума протяженности, позволяют расширить круг решаемых задач.

КЛЮЧЕВЫЕ СЛОВА: обработка, критерий, протяженность, аппроксимация, обратная задача.

НОМЕНКЛАТУРА

$\mathbf{A}$ – линейный оператор (матрица);
A – неизвестный параметр;
$\mathbf{b}$ – вектор;
B – неизвестный параметр;
b – параметр одной из стоимостных функций;
$\mathbf{D}$ – матрица;
E – функционал строгой протяженности;
$E^{(\alpha,\beta,q)}$ – функционал нестрогой протяженности;
$f(x)$ – произвольная функция;
$g(x)$ – функция, описывающая исходные данные;
$\mathbf{H}$ – линейный оператор (матрица);
$k_S^{(\alpha,\beta,q)}$ – нормирующий коэффициент;
N – количество элементов;
$\mathbf{n}$ – вектор, описывающий шум;
$p(\xi)$ – закон распределения ξ;
q – свободно настраиваемый параметр;
$R(\mathbf{x})$ – регуляризатор решения $\mathbf{x}$;
$s(x;\boldsymbol{\theta})$ – функция, описывающая модель данных;
$\mathbf{u}$ – вектор решения;
$\mathbf{x}$ – вектор, описывающий входной сигнал;
x_n – n-ое значение аргумента x;
$\mathbf{y}$ – вектор, описывающий выходной сигнал;
α – свободно настраиваемый параметр;
β – свободно настраиваемый параметр;
γ – параметр регуляризации;
$\boldsymbol{\theta}$ – вектор неизвестных параметров;
θ – неизвестный параметр;
χ – характеристическая функция множества;
ψ – произвольная стоимостная функция;
$\psi_S^{(\alpha,\beta,q)}$ – стоимостная функция супермножества;
σ – стандартное отклонение шума Гаусса или параметр масштаба шума Коши;
$\|\ldots\|^2$ – квадрат евклидовой нормы.

ВВЕДЕНИЕ

Успешность решения задач обработки данных определяется наличием априорных сведений о шумовой обстановке, в которой данные были получены, а также о модели формирования данных и модели искомого решения. В дальнейшем будем полагать, что модель формирования данных является линейной и полностью известной. Тогда если априорные сведения о шумовой обстановке являются статистически полными, а модель искомого решения известна с точностью

DOI 10.15588/1607-3274-2019-1-15

до нескольких неизвестных параметров, то для постановки задач обработки данных может быть применен классический подход, основанный на критерии максимального правдоподобия [1]. Неполное знание статистических характеристик шумовой обстановки приводит к необходимости формулировки задач обработки данных на основе робастного подхода, который позволяет обеспечить устойчивость процесса обработки данных к «загрязнениям» неизвестной статистической природы [2]. С другой стороны, неполное знание математических моделей, которые описывают искомое решение, приводит к необходимости формулировки таких критериев и/или ограничений, которые в достаточно общей форме описывают желаемые свойства решения, например, свойство ограниченности его энергии [3]. В данной работе представлены постановки задач обработки данных, сформулированные на основе критерия минимума протяженности, который выражает свойство «иметь минимальную протяженность». Это свойство в равной степени используется как для постановки задач обработки данных в условиях наличия в них аномальных значений, так и для постановки задач, в которых решение имеет малую протяженность или же является разреженным, а данные содержат аномальные значения.

Объектом исследования в данной работе является процесс постановки задач обработки данных на основе критерия минимума протяженности.

Предмет исследования составляют постановки задач аппроксимации данных, содержащих шум и аномальные значения, а также постановки обратных задач с прямым линейным оператором, решение которых априорно является разреженным или имеет малую протяженность, а исходные данные содержат шум и аномальные значения.

Целью данной работы является разработка и описание подхода к постановке задач обработки данных на основе критерия минимума протяженности.

1 ПОСТАНОВКА ЗАДАЧИ

Постановка задачи заключается в том, чтобы на основе критерия минимума протяженности получить математические формулировки задач обработки данных с такими характеристиками: 1) искомое решение задается параметрической моделью с небольшим числом неизвестных параметров, а исходные данные искажены аддитивным шумом и содержат аномальные значения; 2) искомое решение является разреженным или имеет малую протяженность, а исходные данные представляют его размытую копию, возникающую из-за воздействия линейной системы формирования данных и искаженную из-за шума и аномальных значений. Эти задачи представляют собой соответственно задачу аппроксимации (приближения) данных и обратную задачу с прямым линейным оператором, которая в дискретном случае сводится к решению системы линейных алгебраических уравнений с неточно заданной правой частью.

2 ОБЗОР ЛИТЕРАТУРЫ

Классический подход к постановке задач обработки данных основан на использовании критерия максимального правдоподобия [1]. В предположении гауссовского закона распределения шума этот критерий становится квадратичным критерием, приводя к задаче «наименьших квадратов» [4]. Несомненным достоинством квадратичного критерия является то, что для линейной модели данных задача «наименьших квадратов» имеет аналитическое решение. Однако этот критерий является неудовлетворительным для обработки тех данных, которые в своем составе имеют аномальные значения (выбросы, грубые ошибки, пропуски значений и т.п.) [5]. Аналогичным (но в меньшей степени) недостатком обладает также и модульный критерий, который возникает из критерия максимального правдоподобия в предположении лапласовского закона распределения шума и приводит к постановке задачи «наименьших модулей» [6]. В этих условиях целесообразным является робастный подход, который основан на использовании критерия обобщенного максимального правдоподобия и который предназначен для решения задач обработки данных при их неполном статистическом описании [2]. Робастный подход часто реализуется на основе стоимостных функций с горизонтальными асимптотами [7], которые обеспечивают конечную величину точки удаления (rejection point) [8]. При этом выбор стоимостной функции завершается установкой значений ее настроечных параметров.

Основу робастного подхода составляет идея метода М-оценивания [2]. Согласно [2], метод М-оценивания неизвестного параметра θ заключается в решении задачи:

$$\min_{\theta} \sum_{n=1}^{N} \psi[x_n;\theta], \qquad (1)$$

где x_n есть n-й элемент данных, а ψ есть произвольная функция. Функцию ψ (это обозначение взято из [9], а в [2] используется символ ρ) называют стоимостной функцией [10], функцией потерь, функцией контраста [7] или весовой функцией [9] в зависимости от придаваемого ей смысла, который может и не иметь вероятностной трактовки [8]. Соотношение (1) выражает достаточно общую идею постановки любой задачи в виде задачи оптимизации. Однако важным является факт использования в нем аппарата стоимостных функций, которые позволяют формировать нужные критерии обработки. В частности, использование в качестве функции ψ логарифма совместной плотности вероятности и предположения о полном статистическом описании данных приводит к критерию максимума правдоподобия [1]. Детализация этой функции на основе априорных сведений о природе шума позволяет получить постановки задач типа

DOI 10.15588/1607-3274-2019-1-15

«наименьших квадратов», «взвешенных наименьших квадратов», «наименьших модулей» и т.п.

В [11–12] предложено развитие идеи этого подхода. Оно заключается в построении «супермножества» стоимостных функций, которое управляется набором из трех свободных параметров и которое содержит в себе характеристическую функцию множества, необходимую для получения оценок типа моды (mode-type estimator). Стоимостная функция, которая задает это супермножество, имеет вид [11]:

$$\psi_S^{(\alpha,\beta,q)}(x) = k_S^{(\alpha,\beta,q)}[(1+|x/\alpha|^q)^{\beta/q} - 1] \qquad (2)$$

где α – параметр сглаживания ($0<\alpha<\infty$); q – параметр степени сглаживания ($0<q<\infty$); β – параметр формы стоимостной функции, причем $-\infty<\beta\le 1$ и $\beta<q$; $k_S^{(\alpha,\beta,q)} = 1/[(1+|x_1/\alpha|^q)^{\beta/q} - 1]$ – нормировочный коэффициент, в котором x_1 является точкой нормировки функции (2) на единицу, то есть $\psi_S^{(\alpha,\beta,q)}(x_1)=1$ (обычно $x_1=1$). Параметры α, β и q имеют смысл свободных параметров, которые позволяют менять поведение стоимостной функции (2). Элементами супермножества являются многие известные стоимостные функции, а разработанный в [12] механизм их преобразования позволяет выполнять настройку процесса обработки на текущую шумовую обстановку.

Применение априорных сведений о шумовой обстановке и свойствах искомого решения находит свое выражение в постановке задачи обработки данных как задачи минимизации некоторого составного функционала, где одна часть функционала отвечает за невязку решения, а вторая часть – за получаемое решение. Так, в [13] приведена традиционная постановка задачи обработки сигнала, представленного в виде: $\mathbf{y} = \mathbf{Hx} + \mathbf{n}$, где $\mathbf{H}$ есть известный линейный оператор, который задает процесс формирования данных, $\mathbf{x}$ есть искомое решение, а $\mathbf{n}$ описывает шум. В предположении гауссовской природы шума, эта постановка имеет вид задачи минимизации целевой функции:

$$J(\mathbf{x}) = \| \mathbf{y} - \mathbf{Hx} \|^2 + \gamma R(\mathbf{x}), \qquad (3)$$

которая фактически получается как результат использования байесовского подхода. В (3) первое слагаемое отвечает за достоверность данных, а второе слагаемое является «регуляризатором», который описывает априорную информацию о решении. Например, если известно, что решение $\mathbf{x}$ должно быть гладким (низкочастотным) сигналом, то регуляризатор $R(\mathbf{x})$ должен отвечать за энергию выхода высокочастотного фильтра. Если же решение является кусочно-постоянным, то регуляризатор может задавать «полную вариацию» (total variation) решения [13].

В [14–15] приведены различные постановки задачи удаления шума из кусочно-постоянного сигнала, формулируемые в виде задач кусочно-постоянной аппроксимации. Они представлены как функционалом, состоящим из одной части, так и составным «обобщенным» функционалом, состоящим из взвешенной суммы основной и регуляризирующей частей. При этом для построения функционалов или их частей используются различные функции потерь, которые построены в виде расстояний в пространствах L_0, L_1 и L_2, а также в виде некоторых функций этих расстояний, что позволяет формировать «жесткие» (hard) и «мягкие» (soft) версии функции потерь и соответствующих им функционалов [14].

В [16] приведена постановка задачи восстановления разреженных данных, искаженных импульсным шумом. В отличие от (3), эта задача сформулирована как задача минимизации составного функционала, в котором первый член задает невязку решения в метрике L_1, а второй член исполняет роль штрафной функции, обеспечивающей получение разреженного решения. При этом в качестве штрафной функции используется непрерывная кусочно-квадратичная функция $\psi(x) = \begin{cases} |x| - bx^2; & |x| \le 1/(2b) \\ 1/(4b); & |x| > 1/(2b) \end{cases}$, которую называют «слабо выпуклой функцией разреженности» (weakly convex sparseness function) или «минимаксно-вогнутой штрафной функцией» (minimax-concave penalty function) [16]. Важным является то, что для умеренных значений аргумента $|x| \le 1/(4b)$ эта функция близка к модульной функции, а для $|x| > 1/(2b)$ она выходит на константу.

Наиболее известной и широко используемой на практике является постановка задачи регуляризации Тихонова [3], которая применительно к решению системы линейных алгебраических уравнений $\mathbf{Au} = \mathbf{b}$ с неточно заданной правой частью $\mathbf{b}$ имеет вид:

$$\min_{\mathbf{u}} \{ \| \mathbf{Au} - \mathbf{b} \|^2 + \gamma \| \mathbf{u} \|^2 \}. \qquad (4)$$

В (4) оба члена регуляризирующего функционала построены на основе критерия наименьших квадратов. Это обеспечивает возможность получения решения задачи в аналитическом виде, где значение параметра регуляризации γ выбирается, например, на основе принципа обобщенной невязки [3]. Если в первом члене (4) матрицу $\mathbf{A}$ заменить на тождественную матрицу, а во втором члене вместо $\mathbf{u}$ использовать $\mathbf{Du}$, где $\mathbf{D}$ есть матрица формирования конечных разностей первого порядка, то получим задачу аппроксимации исходных данных $\mathbf{b}$ некоторым «гладким» решением $\mathbf{u}$, которое не требует задания параметрической модели.

DOI 10.15588/1607-3274-2019-1-15

3 МАТЕРИАЛЫ И МЕТОДЫ

Критерий минимума протяженности основан на математической формулировке понятия «протяженность функции». Применительно к абстрактной функции $f(x)$, строгая формулировка этого понятия может быть задана функционалом [17]:

$$E[f(x)] = \int_{-\infty}^{\infty} \chi[f(x)]dx, \quad (5)$$

где $\chi[f(x)] = \begin{cases} 1; & f(x) \neq 0 \\ 0; & f(x) = 0 \end{cases}$ есть характеристическая функция множества значений $f(x)$, которая играет роль идеальной стоимостной функции и которая задает жесткое разбиение всего множества значений функции $f(x)$ на ненулевые значения и нулевое значение. Нестрогая формулировка этого понятия, которая использует (2), может быть задана функционалом:

$$E^{(\alpha,\beta,q)}[f(x)] = \int_{-\infty}^{\infty} \psi_S^{(\alpha,\beta,q)}[f(x)]dx. \quad (6)$$

Так как стоимостная функция (2) дает возможность плавного перехода от ненулевого значения функции к ее нулевому значению через «почти нулевое» значение, то (6) можно использовать для случаев, когда $f(x)$ не является строго финитной функцией и/или когда ее значения искажены шумом во всей рассматриваемой области.

Критерий минимума протяженности заключается в требовании минимизировать протяженность функции, используемой для поиска решения. При этом такая функция может представлять собой как искомое решение, так и невязку искомого решения. На основе данного критерия возможны формулировки различных задач обработки данных. Далее представлены постановки задач аппроксимации данных заданной параметрической моделью и обратных задач с прямым линейным оператором, выполненных с использованием (6).

Задача аппроксимации наблюдаемой функции $g(x)$ заданной модельной функцией $s(x;\theta)$, где θ есть вектор неизвестных параметров, имеет вид:

$$\min_{\theta} E^{(\alpha,\beta,q)}[g(x) - s(x;\theta)]; \quad x \in X, \quad (7)$$

где X есть область наблюдения $g(x)$. Развернутая запись задачи (7) есть:

$$\min_{\theta}\left\{k_S^{(\alpha,\beta,q)} \cdot \int_X \left[\left(1 + \left|\frac{g(x) - s(x;\theta)}{\alpha}\right|^q\right)^{\beta/q} - 1\right]dx\right\}, \quad (8)$$

в которой коэффициент $k_S^{(\alpha,\beta,q)}$ удерживается из-за того, что $k_S^{(\alpha,\beta,q)} > 0$ для $0 < \beta < 1$ и $k_S^{(\alpha,\beta,q)} < 0$ для $-\infty < \beta < 0$. Однако предельный переход $\beta \to \pm 0$ приводит (8) к виду

$$\min_{\theta}\left\{k_S^{(\alpha,0,q)} \cdot \int_X \ln\left(1 + \left|\frac{g(x) - s(x;\theta)}{\alpha}\right|^q\right)dx\right\} = \\ = \min_{\theta}\left\{\int_X \ln\left(1 + \left|\frac{g(x) - s(x;\theta)}{\alpha}\right|^q\right)dx\right\}, \quad (9)$$

где $k_S^{(\alpha,0,q)} = 1/\ln[1 + |x_1/\alpha|^q] > 0$. В дискретном случае (8) и (9) принимают вид:

$$\min_{\theta}\left\{k_S^{(\alpha,\beta,q)} \sum_{n=1}^{N}\left[\left(1 + \left|\frac{g(x_n) - s(x_n;\theta)}{\alpha}\right|^q\right)^{\beta/q} - 1\right]\right\}, \quad (10)$$

и

$$\min_{\theta}\left\{\sum_{n=1}^{N} \ln\left(1 + \left|\frac{g(x_n) - s(x_n;\theta)}{\alpha}\right|^q\right)\right\}, \quad (11)$$

где N есть число дискретных отсчетов x_n функции $g(x)$ и ее модели $s(x;\theta)$, которые вовлекаются в обработку для получения оценки θ.

Из (7) можно получить постановки таких частных задач, как задача аппроксимации константой:

$$\min_{A} E^{(\alpha,\beta,q)}[g(x) - A], \quad (12)$$

задача аппроксимации линейной функцией:

$$\min_{A,B} E^{(\alpha,\beta,q)}[g(x) - (Ax + B)], \quad (13)$$

задача аппроксимации синусоидальной функцией:

$$\min_{A,B,\omega} E^{(\alpha,\beta,q)}\{g(x) - [A\sin(\omega x) + B\cos(\omega x)]\} \quad (14)$$

и многие другие. Отметим, что в общем случае функционал в (8) или в (9) не является ни выпуклой, ни унимодальной функцией неизвестных параметров. Поэтому аналитическое решение задач (12)–(14) невозможно, а их численное решение в большинстве случаев будет ограничено методами оптимизации нулевого порядка. Это естественно увеличивает вычислительную трудоемкость получения решения.

Необходимость применения критерия минимума протяженности для решения задач аппроксимации возникает тогда, когда часть значений функции $g(x)$

DOI 10.15588/1607-3274-2019-1-15

грубо искажена, например, случайными выбросами или другими аномалиями. В этих условиях функция $s(x;\theta)$ может правильно приближать $g(x)$ только на части X_1 области наблюдения X, причем из-за случайного характера грубых искажений область X_1 в общем случае является неизвестной. Тогда при отсутствии шума можно потребовать, чтобы протяженность области несовпадения функции $g(x)$ с ее моделью $s(x;\theta)$ была минимальной. Это означает целесообразность постановки задачи минимизации функционала (5), где $f(x) = g(x) - s(x;\theta)$. Если же значения функции $g(x)$ не только грубо искажены на части области наблюдения, но и зашумлены на всей области наблюдения, то в этом случае имеет смысл постановка (7), в которой свободные параметры должны быть настроены на текущую шумовую обстановку.

Важный класс задач обработки данных составляют обратные задачи с прямым линейным оператором. В дискретном случае на основе изоморфизма линейных операторов и матриц эти задачи сводятся к решению систем линейных алгебраических уравнений вида $\mathbf{Au} = \mathbf{b}$. В условиях, когда вектор решения $\mathbf{u}$ является разреженным, а вектор исходных данных $\mathbf{b}$ содержит шум и аномальные значения, имеет смысл постановка такой задачи минимизации:

$$\min_{\mathbf{u}}\{E^{(\alpha_1,\beta_1,q_1)}[\mathbf{Au}-\mathbf{b}]+\gamma E^{(\alpha_2,\beta_2,q_2)}[\mathbf{u}]\}, \qquad (15)$$

где γ есть параметр регуляризации, который устанавливает компромисс между протяженностью невязки и протяженностью получаемого решения. Частными случаями (15) являются задачи, в которых главная и/или регуляризирующая части функционала построены на основе квадратичного критерия. Так, для $\alpha_1 \to \infty$ и $q_1 = 2$ из (15) получаем задачу

$$\min_{\mathbf{u}}\{\| \mathbf{Au}-\mathbf{b} \|^2 + \gamma E^{(\alpha_2,\beta_2,q_2)}[\mathbf{u}]\}, \qquad (16)$$

в которой главный член этого составного функционала построен на основе критерия наименьших квадратов, а регуляризирующий член – на основе критерия минимума протяженности. Необходимость постановки задачи (16) возникает тогда, когда невязка решения содержит шум, но не содержит аномальных значений, а решение имеет малую протяженность или является разреженным. Если в (15) положить $\alpha_2 \to \infty$ и $q_2 = 2$, то получаем задачу:

$$\min_{\mathbf{u}}\{E^{(\alpha_1,\beta_1,q_1)}[\mathbf{Au}-\mathbf{b}]+\gamma \| \mathbf{u} \|^2\}, \qquad (17)$$

в которой главный член функционала построен на основе критерия минимума протяженности, а регуляризирующий член – на основе критерия наименьших квадратов. Наконец, если в (15) положить $\alpha_1 \to \infty$, $\alpha_2 \to \infty$ и $q_1 = q_2 = 2$, то имеем традиционную задачу регуляризации Тихонова (4). Таким образом, задача (15) является обобщением задач (4), (16) и (17).

На основе свойств стоимостной функции (2) можно получить, что для $\alpha_1 \to \infty$ и $\alpha_2 \to \infty$ выбор $q_1 = 1$ и/или $q_2 = 1$ приводит к постановкам задач, в которых используется модульный критерий. В частности, для значений $q_1 = 2$ и $q_2 = 1$ получим постановку задачи регрессии lasso (least absolute shrinkage and selection operator) Тибширани [18].

Указанные постановки задач могут быть модифицированы таким образом, чтобы в одном из членов составного функционала фигурировало неизвестное решение, а в другом члене – его производная. Так, если предположить существование первой производной $\mathbf{u}'$ решения $\mathbf{u}$, то на базе (17) можно сформулировать задачу

$$\min_{\mathbf{u}}\{E^{(\alpha_1,\beta_1,q_1)}[\mathbf{Au}-\mathbf{b}]+\gamma \| \mathbf{u}' \|^2\}, \qquad (18)$$

в которой главный член составного функционала построен на основе критерия минимума протяженности для невязки решения, а регуляризирующий член – на основе критерия наименьших квадратов для производной решения. Таким образом, главный член составного функционала в (18) «обрабатывает» шум и аномальные значения в правой части системы линейных алгебраических уравнений, а регуляризирующий член ограничивает энергию первой производной получаемого решения, делая его гладким [19].

4 ЭКСПЕРИМЕНТЫ

Для подтверждения целесообразности и демонстрации результативности предложенных постановок задач было выполнено численное моделирование задачи аппроксимации данных, представляющих собой линейный тренд, и обратной задачи с прямым линейным оператором как задачи решения комплекснозначной системы линейных алгебраических уравнений с действительным решением, но с шумом и аномальными значениями в ее правой части. Важная цель моделирования состояла в том, чтобы показать существование значений свободных параметров, которые приводят к желаемым результатам. При этом задаваемое значение α всегда пересчитывалось для заданных значений β и q посредством процедуры выравнивания поведения стоимостной функции (2) в нуле [12].

Моделирование задачи аппроксимации исходных данных, представляющих собой линейный тренд, выполнялось на основе функциональной зависимости: $s(x; A, B) = Ax + B$, где $A = 1$ и $B = 1$, в диапазоне значений аргумента $x \in [0; 1]$. Далее полученные данные подвергались искажениям двух следующих видов. Первый вид искажений заключался в добавлении

DOI 10.15588/1607-3274-2019-1-15

к исходным данным шума импульсного типа, распределенного по закону Коши с нулевым параметром сдвига: $p(\xi) = (\sigma/\pi)\cdot(\xi^2+\sigma^2)^{-1}$, где $\sigma = 0{,}1$ – параметр масштаба шума Коши. Отметим, что данный вид искажений является принципиально неприемлемым для методов, построенных на основе квадратичного критерия. При этом теоретически наилучшими значениями свободных параметров функционала $E^{(\alpha,\beta,q)}$ являются следующие [11]: $\alpha = \sigma$, $\beta = 0$ и $q = 2$.

Второй вид искажений имитировал наличие аномального участка зависимости и заключался в замене второй половины линейного тренда сегментом экспоненциального тренда, заданного формулой: $s_2(x;C,D) = 0{,}5\cdot\exp[-(x-0{,}5)/10]$, где $x \in [0{,}5;1]$. Таким образом, исходные данные представляли собой кусочно-непрерывную кривую, состоящую из двух одинаковых по протяженности сегментов, которые стыковались в центре окна наблюдения. Далее эти данные были искажены аддитивным шумом, имеющим распределение Гаусса с нулевым математическим ожиданием: $p(\xi) = (\sqrt{2\pi}\sigma)^{-1}\cdot\exp[-\xi^2/2\sigma^2]$, где $\sigma = 0{,}05$ – стандартное отклонение шума. Для получения рабочих значений параметров α, β и q была выполнена настройка функционала $E^{(\alpha,\beta,q)}$ на заданный вид шумовой обстановки с помощью таких множеств значений: для α это множество имело вид $\{10^{-5};10^{-4};10^{-3};10^{-2};10^{-1};1\}$, для β это было множество $\{0{,}5;\,0{,}25;\,0;\,-1;\,-2;\,-4;\,-16\}$ и для q это было множество $\{0{,}5;\,1;\,1{,}5;\,2;\,3;\,5;\,10\}$. В результате настройки были получены такие рабочие значения свободных параметров: $\alpha = 10^{-2}$, $\beta = -1$ и $q = 10$.

Моделирование обратной задачи с прямым линейным оператором выполнялось в виде задачи решения комплекснозначной системы линейных алгебраических уравнений $\mathbf{Au} = \mathbf{b}$ с действительным вектором решением $\mathbf{u}$, но с шумом и аномальными значениями в $\mathbf{b}$. Эта система отвечала математической модели формирования данных, которая описывала значения поперечной компоненты электрического поля дипольных источников излучения, расположенных вдоль прямой линии с параллельными и направленными перпендикулярно ей в одну и ту же сторону электрическими моментами. Она имела такой развернутый вид [20]:

$$\sum_{n=1}^{N} u_n\left[k^2 - \frac{ik}{r_{nj}} - \frac{1}{r_{nj}^2}\right]\cdot\frac{\exp(-ikr_{nj})}{r_{nj}} = b_j;\ j = 1,\ldots,J\,, \quad (19)$$

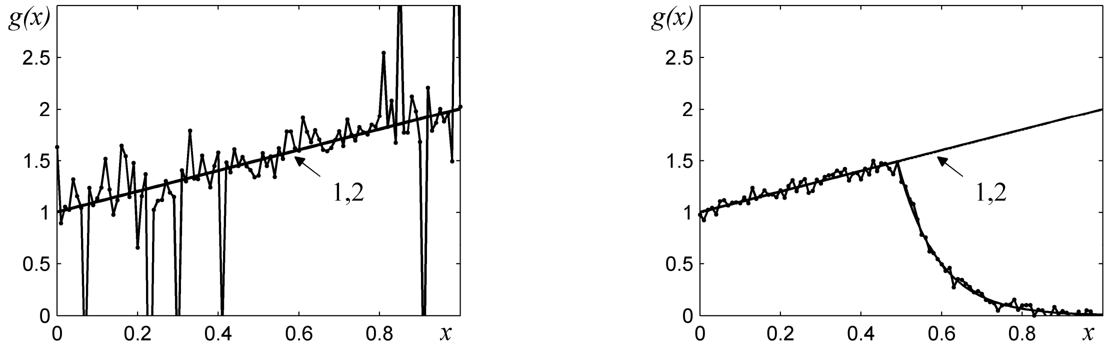
где u_n – значение амплитуды дипольного источника, b_j – значение поперечной компоненты электрического поля, $r_{jn} = \sqrt{(x_n^{(s)} - x_j^{(m)})^2 + z^2}$ – расстояние между точкой $x_n^{(s)}$ на линии дипольных источников и точкой $x_j^{(m)}$ на линии измерений электрического поля, z – расстояние между линией источников и линией измерений, $k = 2\pi/\lambda$ – волновое число и $i = \sqrt{-1}$. При моделировании прямой задачи полагалось, что $N = 100$, $J = 100$, $x_n^{(s)} = n\cdot\Delta x^{(s)}$, $x_j^{(m)} = j\cdot\Delta x^{(m)}$, $\Delta x^{(s)} = \Delta x^{(m)} = \Delta x = 1$, $\Delta x/\lambda = 1/10$, $z/\lambda = 1$, причем $u_n = 1$ для n=40, 50, 60, 80 и $u_n = 0$ для остальных n. Также считалось, что комплексные значения b_j искажены аддитивным гауссовским шумом с нулевым математическим ожиданием, стандартное отклонение которого составляло около 5% от максимального абсолютного значения элементов вещественной и мнимой частей вектора $\mathbf{b}$. Аномальные значения моделировались выбросами с равномерно распределенной амплитудой в интервале [–1,1], а вероятность их появления была равна 0,1 . Последнее означало, что около 10% общего числа элементов данных были полностью разрушены, и априорно было неизвестно, какие именно элементы данных являлись некорректными. Из-за наличия в данных выбросов и предположения о разреженности решения, вместо (19) решалась задача (15) при таких значениях свободных параметров и параметра регуляризации: $\alpha_1 = 5\cdot10^{-3}$, $\alpha_2 = 10^{-4}$, $\beta_1 = \beta_2 = 0$, $q_1 = q_2 = 2$ и $\gamma = 0{,}1$. Эти значения были получены в результате настройки процесса обработки.

5 РЕЗУЛЬТАТЫ

На рис. 1 приведены результаты моделирования задачи аппроксимации исходных данных, представляющих линейный тренд. Рис. 1а отображает первый вид искажений, когда данные искажены аддитивным шумом импульсного типа с законом распределения Коши. Видно, что аппроксимирующая прямая визуально совпала с искомой прямой. Рис. 1б отображает второй вид искажений, когда данные линейного тренда искажены заменой его фрагмента фрагментом экспоненциального тренда, а также аддитивным шумом Гаусса. Видно, что и в этом случае аппроксимирующая прямая визуально совпала с искомой прямой.

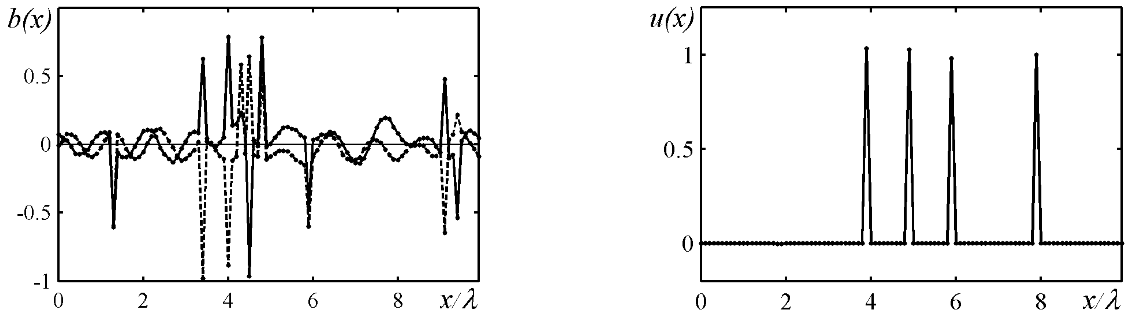
На рис. 2 приведены результаты моделирования обратной задачи, связанной с определением параметров дипольных источников излучения по данным измерений электрической компоненты ближнего поля. Рис. 2а отображает исходные данные задачи в виде действительной (сплошная кривая) и мнимой (пунктирная кривая) частей вектора $\mathbf{b}$, которые искажены шумом Гаусса и выбросами. Рис. 2б отображает результат решения обратной задачи, который визуально совпадает с заданным распределением дипольных источников. Здесь хорошо видно, что полученное решение состоит из четырех диполей, что и было задано при моделировании прямой задачи.

DOI 10.15588/1607-3274-2019-1-15



а б

Рисунок 1 – Аппроксимация линейной функцией:
а – исходные данные искажены шумом Коши; б – исходные данные состоят из сегментов линейного и экспоненциального трендов, искаженных шумом Гаусса; прямые 1 и 2 отвечают истинной и полученной линейным функциям



а б

Рисунок 2 – Решение переопределенной комплекснозначной системы линейных алгебраических уравнений с разреженным действительным вектором решения и грубыми ошибками в ее правой части:
а – исходные данные, б – результат решения

6 ОБСУЖДЕНИЕ

Как и следовало ожидать, качество решения задачи, сформулированной на основе критерия минимума протяженности, определяется результатом настройки значений свободных параметров α, β и q. Отсюда следут вывод, что если для заданного вида шумовой обстановки априорно известно, какие значения свободных параметров являются наилучшими, то их необходимо использовать. Это подтверждает, в частности, рис. 1а. Другими видами шумовой обстановки, для которых можно легко указать наилучшие значения свободных параметров, являются те, которые обусловлены аддитивным шумом с законами распределения Гаусса, Лапласа, обобщенным нормальным законом распределения и обобщенным законом распределения Коши [10]. Если же априорная информация о шумовой обстановке отсутствует или шумовая обстановка является сложной (например, на рис. 1б половина линейного тренда «случайно» искажена фрагментом другого тренда), то в таких случаях необходимо тем или иным способом настроить свободные параметры на их наилучшие значения. Это можно сделать путем численного моделирования аналогичных задач, для которых решение априорно известно.

Процесс настройки значений свободных параметров для примера на рис. 1б показал, что заданный вид искажений оказался неприемлемым не только для методов аппроксимации, построенных на основе квадратичного или модульного критериев, но и для тех, которые построены на основе логарифмических стоимостных функций с модульным или квадратичным внутренним сглаживанием (т.е. когда $\beta = 0$ и $q = 1$ или $q = 2$, соответственно). Последнее обусловлено априорным предположением о том, что решению должен отвечать глобальный минимум функционала. Однако в условиях наличия большого числа аномальных значений искомому решению может отвечать не глобальный, а локальный минимум. Эти вопросы требуют отдельных исследований.

Следует также отметить, что результаты моделирования обратной задачи показали возможность применения предложенного подхода для решения подобных задач на основе критерия минимума протяженности.

В целом, результаты моделирования подтверждают целесообразность и демонстрируют результативность предложенных постановок задач.

DOI 10.15588/1607-3274-2019-1-15

ВЫВОДЫ

Рассмотрена актуальная проблема постановки задач обработки данных, содержащих аномальные значения, и получения разреженных решений или решений малой протяженности. Предложенные постановки задач основаны на критерии минимума протяженности и являются обобщением известных постановок, основанных на использовании квадратичного и/или модульного критериев.

Научная новизна полученных результатов состоит в получении нового подхода к постановке задач обработки данных в условиях сложной шумовой обстановки, а также для получения разреженных решений или решений малой протяженности. В частности, новый подход к задаче аппроксимации данных, которые заданы параметрической моделью, заключается в ее формулировке в виде задачи минимизации функционала протяженности, который применяется к невязке решения. Для обратных задач с прямым линейным оператором предлагаемая постановка задачи сформулирована как задача минимизации составного функционала, в котором основная и регуляризирующая части построены на основе функционала протяженности.

Практическое значение полученных результатов заключается в том, что они могут быть использованы для постановки конкретных задач обработки данных в условиях, когда часть исходных данных является грубо искаженной и/или когда искомое решение является разреженным или имеет малую протяженность. Перспективы дальнейших исследований связаны с построением эффективных численных методов решения задач, основанных на критерии минимума протяженности.

Постановки задач, основанные на критерии минимума протяженности, позволяют расширить круг решаемых задач.

БЛАГОДАРНОСТИ

Работа выполнена в рамках госбюджетной научно-исследовательской темы Днипровского национального университета «Методы и информационные технологии цифровой обработки многоканальных данных» (номер государственной регистрации 0116U001297).

ЛІТЕРАТУРА / ЛИТЕРАТУРА

1. Millar R. B. Maximum Likelihood Estimation and Inference: With Examples in R, SAS and ADMB / R. B. Millar. – New York: Wiley, 2011. – 376 p.
2. Huber P. Robust statistics. 2nd ed. / P. Huber, E. M. Ronchetti. – Hoboken : Wiley, 2009. – 370 p.
3. Тихонов А. Н. Регуляризирующие алгоритмы и априорная информация / А. Н. Тихонов, А. В. Гончарский, В. В. Степанов, А. Г. Ягола. – М. : Наука, 1983. – 200 с.
4. Wolberg J. Data Analysis Using the Method of Least Squares: Extracting the Most Information from Experiments / J. Wolberg. – Berlin : SpringerVerlag, 2005. – 250 p.
5. Chandola V. Anomaly detection: A survey / V. Chandola, A. Banerjee, V. Kumar // ACM Computing Surveys. – 2009. – Vol. 41, № 3. – P. 15–58.
6. Elgmati E. A. Quartile Method Estimation of Two-Parameter Exponential Distribution Data with Outliers / E. A. Elgmati, N. B. Gredni // International Journal of Statistics and Probability. – 2016. – Vol. 5, № 5. – P. 12–15.
7. Shevlyakov G. L. Robustness in data analysis: criteria and methods. / G. L. Shevlyakov, N. O. Vilchevski. – Utrecht : VSP, 2002. – 310 p.
8. Hampel F. R. Robust statistics: the approach based on influence functions / F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, W. A. Stahel. – Hoboken, NJ : Wiley, 2011. – 502 p. DOI=10.1002/9781118186435.
9. Демиденко Е. З. Оптимизация и регрессия / Е. З. Демиденко. – М. : Наука, 1989. – 296 с.
10. Aysal T. C. Meridian filtering for robust signal processing / T. C. Aysal, K. E. Barner // IEEE Trans. on Signal Processing. – 2007. – V. 55, № 8. – P. 3949–3962.
11. Borulko V. F. Minimum-duration filtering / V. F. Borulko, S. M. Vovk // Радіоелектроніка, інформатика, управління. – 2016. – № 1. – С. 7–14. DOI: 10.15588/1607-3274-2016-1-1.
12. Vovk S. M. General approach to building the methods of filtering based on the minimum duration principle / S. M. Vovk // Radioelectronics and Communications Systems. – 2016. – V. 59, № 7. – P. 281–292. DOI: 10.3103/S0735272716070013.
13. Selesnick I. W. Enhanced Sparsity by Non-Separable Regularization / I. W. Selesnick, İ. Bayram // IEEE Transactions on Signal Processing. – 2016. – Vol. 64, № 9. – P. 2298 2313. DOI: 10.1109/TSP.2016.2518989.
14. Little M. A. Generalized methods and solvers for noise removal from piecewise constant signals. I. Background theory / M. A. Little, N. S. Jones // Proceedings of the Royal Society A. – 2011. – Vol. 467. – P. 3088–3114. DOI: 10.1098/rspa.2010.0671.
15. Little M. A. Generalized methods and solvers for noise removal from piecewise constant signals. II. New methods / M. A. Little, N. S. Jones // Proceedings of the Royal Society A. – 2011. – Vol. 467. – P. 3115–3140. DOI: 10.1098/rspa.2010.0674.
16. Liu Q. Robust Sparse Recovery via Weakly Convex Optimization in Impulsive Noise / Q. Liu, C. Yang, Y. Gu, H. C. So // Signal Processing. – 2018. – Vol. 152. – P. 84–89. DOI: =https://doi.org/10.1016/j.sigpro.2018.05.020.
17. Titchmarsh E. C. The theory of functions / E. C. Titchmarsh. – New York : Oxford University Press, 1939. – 454 p.
18. Tibshirani R. Regression shrinkage and selection via the lasso: a retrospective / R. Tibshirani // Journal of the Royal Statistical Society Series B. – 2011. – Vol. 73, № 3. – P. 273–282. DOI= https://doi.org/10.1111/j.1467-9868.2011.00771.x
19. Вовк С. М. Метод численного дифференцирования зашумленных данных с выбросами / С. М. Вовк // Радіоелектроніка, інформатика, управління. – 2017. – № 3 – С. 44–52. DOI: 10.15588/1607-3274-2017-3-5].
20. Вовк С. М. Двойственный метод минимума пространственной протяженности для робастного оценивания параметров дипольных источников излучения / С. М. Вовк, В. Ф. Борулько // Радіоелектроніка, інформатика, управління. – 2014. – № 2. – С. 8–17. DOI: 10.15588/1607-3274-2014-2-1.

Статья поступила в редакцию 26.12.2018.
После доработки 09.01.2019.

DOI 10.15588/1607-3274-2019-1-15

УДК 004.02

ПОСТАНОВКА ЗАДАЧ ОБРОБКИ ДАНИХ НА ОСНОВІ КРИТЕРІЮ МІНІМУМУ ПРОТЯЖНОСТІ

Вовк С. М. – канд. фіз.-мат. наук, доцент, доцент кафедри комп'ютерних наук та інформаційних технологій Дніпровського національного університету імені Олеся Гончара, Дніпро, Україна.

АНОТАЦІЯ

Актуальність. Для обробки даних, які вміщують аномальні значення, та отримання розріджених рішень або рішень малої протяжності може бути використана вимога мінімізувати протяжність функції, використовуваної для пошуку рішення. Об'єктом дослідження в даній роботі є процес постановки задач обробки даних на основі зазначеної вимоги, яка далі згадується як критерій мінімуму протяжності. Метою даної роботи є розробка підходу до постановки задач обробки даних на основі даного критерію.

Метод. На основі критерію мінімуму протяжності запропоновано новий підхід до постановки задач апроксимації даних та до постановки обернених задач з прямим лінійним оператором, рішення яких має малу протяжність або є розрідженим, в умовах, коли дані містять шум та аномальні значення. Постановка задачі апроксимації отримана шляхом завдання параметричної моделі даних і застосування критерію мінімуму протяжності до відхилу рішення. Постановка оберненої задачі отримана шляхом застосування критерію мінімуму протяжності до рішення задачі та до його відхилу. Представлені окремі випадки цієї постановки та відзначено, що вона узагальнює постановку задачі регуляризації Тихонова. Запропоновані постановки задач сформульовані у вигляді задач мінімізації відповідних функціоналів, побудованих на основі «супермножини» вартісних функцій. У загальному випадку зазначені функціонали не є ані опуклими, ані унімодальними, і їх мінімізація може виявитися трудомістким завданням.

Результати. Запропоновані постановки задач узагальнюють ті постановки, які виконані на основі критеріїв найменших квадратів та/або найменших модулів. Чисельне моделювання задачі апроксимації лінійною функцією зашумленних даних в умовах наявності шуму імпульсного типу, а також в умовах наявності фрагмента експоненційної функції, який заважає, підтвердило доцільність запропонованої постановки та її результативність. Чисельне моделювання оберненої задачі, якій відповідала перевизначена система лінійних алгебраїчних рівнянь з грубими помилками в її правій частині та розрідженим рішенням, також підтвердило доцільність використання критерію мінімуму протяжності для її постановки.

Висновки. Постановка задач обробки даних на основі критерію мінімуму протяжності є доцільною в умовах, коли частина даних є грубо спотвореною та/або коли шукане рішення має малу протяжність. Постановки, засновані на критерії мінімуму протяжності, дозволяють розширити коло вирішуваних задач.

КЛЮЧОВІ СЛОВА: обробка, критерій, протяжність, апроксимація, обернена задача.

UDC 004.02

PROBLEM STATEMENTS OF DATA PROCESSING BASED ON CRITERION OF MINIMUM EXTENT

Vovk S. M. – PhD, Associate Professor, Associate Professor of the Department of Computer Science and Information Technology, Oles Honchar Dnipro National University, Ukraine.

ABSTRACT

Context. In order to process the data containing anomalous values as well as to obtain the sparse solutions or solutions with small extent, the requirement to minimize the extent of the function used to find solution can be used. In this paper the object of the study is the process of setting the data processing problems on the basis of this requirement, which is further referred to as the criterion of minimum extent.

Objective. The goal of this work is the development of an approach to the formulation of the data processing problems based on criterion of minimum extent.

Method. On the basis of minimum extent criterion, a new approach is proposed. This approach allows to formulate the data approximation problem as well as the inverse problem with a direct linear operator and with a solution of small extent or with a sparse solution in conditions that the initial data contain noise and anomalous values. The statement of the approximation problem is obtained by setting a parametric data model and applying the criterion of minimum extent to the solution residual. The statement of the inverse problem is obtained by applying the criterion of minimum extent to the solution of the problem and to the solution residual. The special cases of this statement are presented and it is noted that this statement generalizes the statement of the Tikhonov regularization problem. The proposed problem statements are formulated as minimization problems for the corresponding functionals constructed on the basis of the "superset" of cost functions. In the general case, the indicated functionals are neither convex nor unimodal, and their minimization can be a laborious task.

Results. The proposed problem statements generalize those that are performed on the basis of the least squares criterion and/or least modules criterion. Numerical simulation of the problem of approximation by a linear function of noisy data in the presence of impulsive noise, as well as in the presence of an interfering fragment of exponential function, confirmed the feasibility of the proposed statement and its effectiveness. Numerical simulation of the inverse problem, corresponded to the overdetermined system of linear algebraic equations with gross errors in its right-hand side and sparse solution, also confirmed the feasibility of using the criterion of minimum extent for its formulation.

Conclusions. The problem statement of data processing which is based on the criterion of minimum extent is expedient under conditions when the part of the data is roughly distorted and/or when the desired solution has a small extent. The statements based on the criterion of minimum extent allow us to expand the range of the problems to be solved.

KEYWORDS: processing, criterion, extent, approximation, inverse problem.

DOI 10.15588/1607-3274-2019-1-15

REFERENCES

1. Millar R. B. Maximum Likelihood Estimation and Inference: With Examples in R, SAS and ADMB. New York, Wiley, 2011, 376 p.
2. Huber P., Ronchetti E. M. Robust statistics. 2nd ed. Hoboken, Wiley, 2009, 370 p.
3. Tihonov A. N., Goncharskij A. V., Stepanov V. V., Jagola A. G. Reguljarizirujushhie algoritmy i apriornaja informacija. Moscow, Nauka, 1983, 200 p.
4. Wolberg J. Data Analysis Using the Method of Least Squares: Extracting the Most Information from Experiments. Berlin, SpringerVerlag, 2005, 250 p.
5. Chandola V., Banerjee A., Kumar V. Anomaly detection: A survey, *ACM Computing Surveys*, 2009, Vol. 41, No. 3, pp. 15–58.
6. Elgmati E. A., Gredni N. B. Quartile Method Estimation of Two-Parameter Exponential Distribution Data with Outliers, *International Journal of Statistics and Probability*, 2016, Vol. 5, No. 5, pp. 12–15.
7. Shevlyakov G. L., Vilchevski N. O. Robustness in data analysis: criteria and methods. Utrecht, VSP, 2002, 310 p.
8. Hampel F. R. Ronchetti E. M., Rousseeuw P. J., Stahel W. A. Robust statistics: the approach based on influence functions. Hoboken, NJ, Wiley, 2011, 502 p. DOI=10.1002/9781118186435.
9. Demidenko E. Z. Optimizacija i regressija. Moscow, Nauka, 1989, 296 с.
10. Aysal T. C., Barner K. E. Meridian filtering for robust signal processing, *IEEE Trans. on Signal Processing*, 2007, V. 55, No. 8, pp. 3949–3962.
11. Borulko V. F., Vovk S. M. Minimum-duration filtering, *Radio Electronics, Computer Science, Control,* 2016, No. 1, pp. 7–14. DOI: 10.15588/1607-3274-2016-1-1.
12. Vovk S. M. General approach to building the methods of filtering based on the minimum duration principle, *Radioelectronics and Communications Systems*, 2016, V. 59, No. 7, pp. 281–292. DOI: 10.3103/S0735272716070013.
13. Selesnick I. W., Bayram İ. Enhanced Sparsity by Non-Separable Regularization, *IEEE Transactions on Signal Processing,* 2016, Vol. 64, No. 9, pp. 2298–2313. DOI: 10.1109/TSP.2016.2518989.
14. Little M. A., Jones N. S. Generalized methods and solvers for noise removal from piecewise constant signals. I. Background theory, *Proceedings of the Royal Society A,* 2011, Vol. 467, pp. 3088–3114. DOI:10.1098/rspa.2010.0671.
15. Little M. A., Jones N. S. Generalized methods and solvers for noise removal from piecewise constant signals. II. New methods, *Proceedings of the Royal Society A,* 2011, Vol. 467, pp. 3115–3140. DOI: 10.1098/rspa.2010.0674.
16. Liu Q., Yang C., Gu Y., So H. C. Robust Sparse Recovery via Weakly Convex Optimization in Impulsive Noise, *Signal Processing*, 2018, Vol. 152, pp. 84–89. DOI: =https://doi.org/10.1016/j.sigpro.2018.05.020.
17. Titchmarsh E. C. The theory of functions. New York, Oxford University Press, 1939, 454 p.
18. Tibshirani R. Regression shrinkage and selection via the lasso: a retrospective, *Journal of the Royal Statistical Society Series B,* 2011, Vol. 73, No. 3, pp. 273–282. DOI= https://doi.org/10.1111/j.1467-9868.2011.00771.x
19. Vovk S. M. Metod chislennogo differencirovanija zashumlennyh dannyh s vybrosami, *Radio Electronics, Computer Science, Control,* 2017, No. 3, pp. 44–52. DOI: 10.15588/1607-3274-2017-3-5].
20. Vovk S. M., Borul'ko V. F. Dvojstvennyj metod minimuma prostranstvennoj protjazhennosti dlja robastnogo ocenivanija parametrov dipol'nyh istochnikov izluchenija, *Radio Electronics, Computer Science, Control,* 2014, No. 2, pp. 8–17. DOI: 10.15588/1607-3274-2014-2-1.

DOI 10.15588/1607-3274-2019-1-15