Математические методы исследования
Mathematical methods of investigation

DOI: https://doi.org/10.26896/1028-6861-2020-86-1-62-74

ЗАДАЧА ВОССТАНОВЛЕНИЯ ЗАВИСИМОСТЕЙ ПО ДАННЫМ С ИНТЕРВАЛЬНОЙ НЕОПРЕДЕЛЕННОСТЬЮ

© **Сергей Петрович Шарый**

Институт вычислительных технологий СО РАН, Россия, 630090, г. Новосибирск, пр. Академика Лаврентьева, 6;
e-mail: shary@ict.nsc.ru

Статья поступила 16 апреля 2019 г. Поступила после доработки 30 апреля 2019 г. Принята к публикации 17 мая 2019 г.

Рассмотрена задача восстановления зависимостей по данным с неопределенностью, которая не описывается теоретико-вероятностными законами, но ограничена по величине и имеет интервальный характер, т.е. выражается интервалами возможных значений данных. Исследован наиболее общий случай, когда интервалы являются результатами измерений как в независимых (предикторных) переменных, так и в зависимой (критериальной) переменной. Введены понятия слабой и сильной согласованности данных и параметров функциональной зависимости. Формулировки задач сведены к исследованию и оцениванию различных множеств решений для интервальной системы уравнений, построенной по обрабатываемым данным. Подробно рассмотрено сильное согласование параметров и данных как более практичное, более адекватное реальности и обладающее лучшими теоретическими свойствами. Оценки параметров зависимости, получаемые с учетом сильного согласования, имеют полиномиальную вычислительную сложность, робастны, почти всегда имеют конечную вариабельность, а также лишь частично подвержены так называемому парадоксу Е. З. Демиденко. Предложена вычислительная технология решения задачи восстановления линейной зависимости в условиях интервальной неопределенности данных и с учетом требования сильного согласования. Ее основой служит техника, основанная на применении так называемого распознающего функционала множества решений задачи — специального отображения, которое знаком своих значений распознает принадлежность точки множеству решений и одновременно дает количественную меру этой принадлежности. Обсуждаются свойства распознающего функционала. Оценкой параметров восстанавливаемой зависимости принимается точка максимума этого функционала, которая обеспечивает наилучшее согласование параметров и данных (или их наименьшее рассогласование). Соответственно, практическая реализация этого подхода, названного «методом максимума согласования», сводится к численному нахождению безусловного максимума распознающего функционала — вогнутой негладкой функции. В заключение работы приведен конкретный пример решения задачи восстановления линейной функции по данным измерений с интервальной неопределенностью.

Ключевые слова: задача восстановления зависимостей; интервальная неопределенность данных; согласование параметров и данных; сильное согласование; интервальная система уравнений; объединенное множество решений; допусковое множество решений; распознающий функционал.

DATA FITTING PROBLEM UNDER INTERVAL UNCERTAINTY IN DATA

© **Sergey P. Shary**

Institute of Computational Technologies SB RAS, Lavrentieva prosp., 6, Novosibirsk, 630090, Russia;
e-mail: shary@ict.nsc.ru

Received April 16, 2019. Revised April 30, 2019. Accepted May 17, 2019.

We consider the data fitting problem under uncertainty, which is not described by probabilistic laws, but is limited in magnitude and has an interval character, i.e., is expressed by the intervals of possible data values. The most general case is considered when the intervals represent the measurement results both in independent (predictor) variables and in the dependent (criterial) variables. The concepts of weak and strong compatibility of data and parameters of functional dependence are introduced. It is shown that the resulting formulations of problems are reduced to the study and estimation of various solution sets for an interval system of equations constructed from the processed data. We discuss in detail the strong compatibility of the parameters and data, as more practical, more adequate to the reality and possessing better theoretical properties. The estimates of the function parameters, obtained in view of the strong compatibility, have a polynomial computational complexity, are robust, almost always have finite variability, and are also only partially affected by the so-called Demidenko paradox. We also propose a computational technology for solving the problem of constructing a linear functional dependence under interval data uncertainty and take into account the requirement of strong compatibility. It is based on the application of the so-called recognizing functional of the problem solution set — a special mapping, which recognizes, by the sign of the values, whether a point belongs to the solution set and simultaneously provides a quantitative measure of this membership. The properties of the recognizing functional are discussed. The maximum point of this functional is taken as an estimate of the parameters of the functional dependency under construction, which ensures the best compatibility between the parameters and data (or their least discrepancy). Accordingly, the practical implementation of this approach, named "maximum compatibility method," is reduced to the computation of the unconditional maximum of the recognizing functional — a concave non-smooth function. A specific example of solving the data fitting problem for a linear function from measurement data with interval uncertainty is presented.

Keywords: data fitting problem; interval data uncertainty; compatibility of the parameters and data; weak compatibility; strong compatibility; interval system of equations; united solution set; tolerable solution set; recognizing functional.

Введение

Предмет работы — развитие методов анализа данных, которые неточны и имеют интервальную неопределенность. Мы рассматриваем задачу восстановления линейной зависимости вида

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n, \qquad (1)$$

в которой $x_1, x_2, \ldots, x_n$ — независимые переменные (называемые также входными или предикторными переменными); y — зависимая переменная (называемая также выходной или критериальной переменной), а $\beta_0, \beta_1, \ldots, \beta_n$ — некоторые коэффициенты. Эти неизвестные коэффициенты должны быть определены на основе ряда измерений значений $x_1, x_2, \ldots, x_n$ и y. Будем считать, что всего имеется m измерений, результатами которых являются

$$\begin{matrix} x_{11}, & x_{12}, & \ldots, & x_{1n}, & y_1 \\ x_{21}, & x_{22}, & \ldots, & x_{2n}, & y_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ x_{m1}, & x_{m2}, & \ldots, & x_{mn}, & y_m \end{matrix} \qquad (2)$$

(первый нижний индекс обозначает номер измерения).

Рассматриваемая задача восстановления зависимостей — одна из классических задач обработки данных, которую наглядно иллюстрирует рис. 1: требуется найти прямую, которая «наилучшим образом приближает» множество точек, полученных в результате измерений или наблюдений.

В практических задачах восстановления зависимостей данные почти всегда неточны, поскольку на результаты измерений влияют внешние неконтролируемые факторы, измерительные приборы не являются абсолютно точными и т.п. Далее будем работать с неопределенностями и неточностями в данных с помощью методов интервального анализа (см., например, [1 – 4]). При этом для результатов измерений заданными считаются интервальные оценки, которым принадлежат истинные значения измеряемых величин. В частности, в рассматриваемой нами задаче оценивания параметров линейной зависимости мы считаем, что даны интервалы для x_{ij} и y_i:

$$x_{ij} \in \boldsymbol{x}_{ij} = [\inf \boldsymbol{x}_{ij}, \sup \boldsymbol{x}_{ij}]$$

$$\text{и } \; y_i \in \boldsymbol{y}_i = [\inf \boldsymbol{y}_i, \sup \boldsymbol{y}_i],$$

где жирный шрифт обозначает (в соответствии с неформальным международным стандартом [5]) интервальность соответствующих величин; inf и sup — операции взятия нижней и верхней границ интервала.

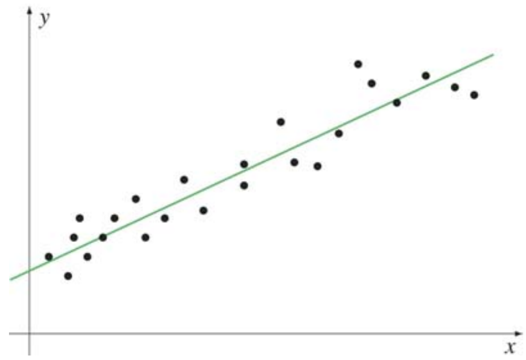

Рис. 1. Иллюстрация задачи восстановления линейной зависимости

Fig. 1. Illustration of the data fitting problem

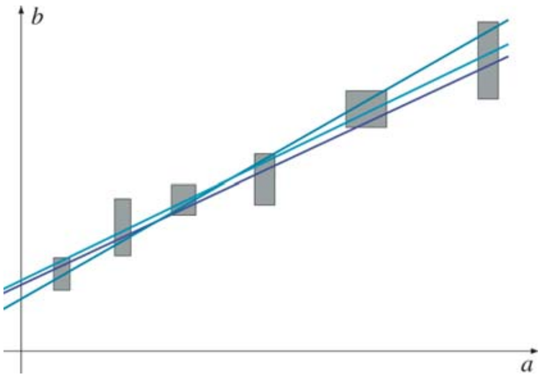

Рис. 2. Восстановление линейной зависимости по интервальным данным измерений

Fig. 2. Construction of a linear function dependence from the interval measurement data

В подобной постановке задача впервые была рассмотрена в 1962 году Л. В. Канторовичем в статье [6], где также были кратко намечены доступные методы ее решения. В последующие годы значительные результаты в теории и практическом решении задачи восстановления зависимостей по интервальным данным были получены в работах А. П. Вощинина и его учеников [7 – 9], В. А. Суханова [10], Н. М. Оскорбина с коллегами [11], С. И. Жилина [12], С. И. Спивака и учеников (см. в частности, [13]), Б. Т. Поляка и С. А. Назина [14], С. И. Кумкова [15], А. Л. Померанцева и О. Е. Родионовой [16], А. А. Подружко и А. С. Подружко [17] и других исследователей. Этому же вопросу посвящены статьи [18 – 22], развивающие так называемый метод максимума согласования. При этом в подавляющем большинстве работ рассматривалась упрощенная (хотя и очень важная с практической точки зрения) версия общей задачи, когда входные (предикторные) переменные задаются точно, а ограниченные неопределенности присущи только выходным (критериальным) переменным восстанавливаемой зависимости. В данной работе дается решение наиболее общей постановки, в которой невырожденные интервалы возможных значений заданы как для входных, так и для выходных переменных.

Первая зарубежная публикация на рассматриваемую тему принадлежит Ф. Швеппе [23], дальнейшие результаты и их обзор можно найти, например, в [1, 24, 25]. Следует отметить, что исследования по анализу данных с интервальной неопределенностью часто находятся в более широком контексте так называемых «ограниченных неопределенностей» (англоязычные термины — bounded uncertainty, set membership estimation, и др.), когда множествами, описывающими неопределенности интересующих нас величин, выступают не только обычные интервалы или их обобщения, но и многогранники, эллипсоиды и другие более общие параметризованные классы ограниченных множеств.

Существуют исследования, в которых рассматриваются интервальные неопределенности измерений (наблюдений), но в рамках теоретико-вероятностных моделей реальности [26 – 28]. Здесь же мы рассматриваем ситуацию, когда аппарат теории вероятностей не применим к описанию погрешностей измерений.

Итак, результаты измерений неточны, и мы предполагаем, что они являются некоторыми интервалами, которые дают двусторонние границы точных значений измеренных величин. Будем считать, что в результате i-го измерения получаются такие интервалы $\boldsymbol{x}_{i1}$, $\boldsymbol{x}_{i2}$, ..., $\boldsymbol{x}_{in}$, $\boldsymbol{y}_i$, что истинное значение x_1 лежит в $\boldsymbol{x}_{i1}$, истинное значение x_2 — в $\boldsymbol{x}_{i2}$ и т.д. вплоть до y, истинное значение которого лежит в $\boldsymbol{y}_i$. Иными словами, вместо (2) рассматриваем интервальные данные

$$\begin{array}{ccccc} \boldsymbol{x}_{11}, & \boldsymbol{x}_{12}, & \ldots, & \boldsymbol{x}_{1n}, & \boldsymbol{y}_1, \\ \boldsymbol{x}_{21}, & \boldsymbol{x}_{22}, & \ldots, & \boldsymbol{x}_{2n}, & \boldsymbol{y}_2, \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \boldsymbol{x}_{m1}, & \boldsymbol{x}_{m2}, & \ldots, & \boldsymbol{x}_{mn}, & \boldsymbol{y}_m. \end{array} \qquad (3)$$

Необходимо найти или оценить коэффициенты β_j, $j = 0, 1, ..., n$, для которых линейная функция (1) «наилучшим образом» приближает данные (3) (рис. 2). При этом идеальным является случай, когда график восстанавливаемой зависимости «проходит через все точки измерений», совершенно так же, как, например, в задаче интерполирования.

Но в условиях неточности данных, когда результат каждого измерения-наблюдения вместо точки представляет собой целое множество возможных значений рассматриваемой величины, само понятие «прохождения через точки наблюдений» имеет неоднозначный смысл. Теперь множества неопределенности измерений приобретают структуру, что вызывает необходимость различать те или иные случаи прохождения графика функции через эти множества. Это объясняется, в частности, тем, что входы и выходы системы (соответствующие независимым аргументам функции и зависимым переменным) отличаются друг от друга по функциональному назначению, а их измерения могут выполняться отличными друг от друга способами или даже в разное время.

Необходимость различать эти случаи порождает понятия слабого согласования и сильного согласования данных и параметров восстанавливаемой зависимости. В данной работе в продолжение статей [20, 21] обсуждаются практические методы нахождения оценок параметров зависимостей, удовлетворяющих условию сильной согласованности с данными. Показано, что они имеют лучшие теоретические свойства и практически более удобны. Построение зависимостей в смысле обычного (слабого) согласования было подробно рассмотрено ранее [18, 19, 22].

Слабое и сильное согласование данных и параметров зависимости

Подставляя формально данные (3) в выражение (1) для восстанавливаемой функции, получаем интервальную $m \times (n + 1)$-систему линейных алгебраических уравнений

$$\begin{array}{ccccccccccc} \beta_0 & + & \boldsymbol{x}_{11}\beta_1 & + & \boldsymbol{x}_{12}\beta_2 & + & \ldots & + & \boldsymbol{x}_{1n}\beta_n & = & \boldsymbol{y}_1, \\ \beta_0 & + & \boldsymbol{x}_{21}\beta_1 & + & \boldsymbol{x}_{22}\beta_2 & + & \ldots & + & \boldsymbol{x}_{2n}\beta_n & = & \boldsymbol{y}_2, \\ \vdots & & \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ \beta_0 & + & \boldsymbol{x}_{m1}\beta_1 & + & \boldsymbol{x}_{m2}\beta_2 & + & \ldots & + & \boldsymbol{x}_{mn}\beta_n & = & \boldsymbol{y}_m, \end{array} \tag{4}$$

или (кратко)

$$\boldsymbol{X}\beta = \boldsymbol{y}, \tag{5}$$

где $\boldsymbol{X} = (\boldsymbol{x}_{ij})$ — интервальная $m \times (n + 1)$-матрица с первым столбцом из единиц; $\boldsymbol{y} = (\boldsymbol{y}_i)$ — интервальный m-вектор и $\beta = (\beta_0, \beta_1, ..., \beta_n)^{\mathrm{T}}$ — вектор $n + 1$ неизвестных. Аналогично обычному неинтервальному случаю оценивание параметров восстанавливаемой зависимости можно считать «решением» этой интервальной системы уравнений.

В идеале график восстанавливаемой зависимости должен проходить через все «точки данных», которые теперь являются брусами в пространстве R^{n+1}. Обычно их называют «брусами неопределенности замеров». В этом контексте естественным представляется следующее.

Определение 1. Набор параметров $\beta_0, \beta_1, ..., \beta_n$ рассматриваемой линейной зависимости слабо согласуется (или просто согласуется) с интервальными экспериментальными данными $(\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, ..., \boldsymbol{x}_{in}, \boldsymbol{y}_i)$, $i = 1, 2, ..., m$, если для каждого наблюдения i в пределах измеренных интервалов найдутся такие значения $x_{i1} \in \boldsymbol{x}_{i1}$, $x_{i2} \in \boldsymbol{x}_{i2}$, ..., $x_{in} \in \boldsymbol{x}_{in}$ и $y_i \in \boldsymbol{y}_i$, что

$$\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_n x_{in} = y_i.$$

В соответствии с этим определением прохождение графика конструируемой зависимости через «точку» данных, ставшую брусом, понимается просто как ее непустое пересечение с этим брусом (см. рис. 2).

С использованием формального языка логики предикатов определение множества наборов параметров $\beta = (\beta_0, \beta_1, ..., \beta_n)^{\mathrm{T}}$, которые со-

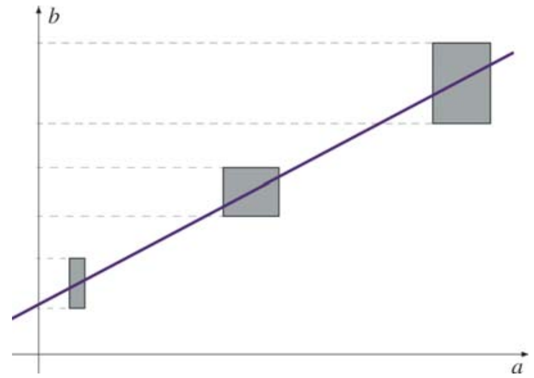

Рис. 3. Иллюстрация сильного согласования параметров линейной модели и интервальных данных измерений

Fig. 3. Illustration of strong compatibility between the linear model parameters and interval measurement data

гласуются с данными (3), выглядит следующим образом:

$$\{\beta \in \mathrm{R}^{n+1} \mid (\exists X \in \boldsymbol{X})(\exists y \in \boldsymbol{y})(X\beta = y)\}.$$

В интервальном анализе оно называется объединенным множеством решений $\Xi_{uni}(\boldsymbol{X}, \boldsymbol{y})$ интервальной системы линейных алгебраических уравнений (4), (5) и неформально описывается как

$$\Xi_{uni}(\boldsymbol{X}, \boldsymbol{y}) = \{\beta \in \mathrm{R}^{n+1} \mid X\beta = y$$
$$\text{для некоторых } X \in \boldsymbol{X} \text{ и } y \in \boldsymbol{y}\}.$$

Но «раздувшиеся» точки-брусы данных приобретают уже дополнительную структуру, которой не было у исходных бесконечно малых точек. Они становятся прямыми декартовыми произведениями интервалов, имеющих разный содержательный смысл, которые отвечают входным (независимым) переменным и выходной (зависимой) переменной. Как следствие, различные грани бруса неопределенности замера имеют разное значение (на рис. 2 это вертикальные и горизонтальные стороны прямоугольников), а задача восстановления зависимостей по неточным данным приобретает новый смысл. Теперь важно знать, как именно график восстанавливаемой зависимости проходит через брус неопределенности замера.

Если процесс измерения значений входа и выхода разорван во времени и разделен на этапы, когда выходы измеряются после фиксации значений входов, то более адекватно понимание «согласования», при котором ограничение на выходе должно выполняться равномерно при любых значениях входов. Эта ситуация описывается уже другим определением.

Определение 2. Набор параметров $\beta_0, \beta_1, ..., \beta_n$ рассматриваемой линейной зависимости сильно согласуется с интервальными экспериментальными данными $(\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, ..., \boldsymbol{x}_{in}, \boldsymbol{y}_i)$, $i = 1, 2, ..., m$, если для каждого наблюдения i для любых значений $x_{i1} \in \boldsymbol{x}_{i1}, x_{i2} \in \boldsymbol{x}_{i2}, ..., x_{in} \in \boldsymbol{x}_{in}$ в пределах измеренных интервалов найдется такое $y_i \in \boldsymbol{y}_i$, что

$$\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_n x_{in} = y_i.$$

Это определение наглядно иллюстрирует рис. 3. Множество параметров зависимости (1), согласующихся с данными задачи в смысле определения 2, на формальном языке описывается следующим образом:

$$\{\beta \in \mathrm{R}^{n+1} \,|\, (\forall X \in \boldsymbol{X})(\forall y \in \boldsymbol{y})(X\beta = y)\}.$$

В интервальном анализе это множество называется допусковым множеством решений $\Xi_{tol}(\boldsymbol{X}, \boldsymbol{y})$ интервальной линейной системы уравнений (4), (5), так как исторически оно возникло из решения практических задач, в которых фигурируют некоторые «допуски» на параметры объекта [4, 29]. Неформально

$$\Xi_{tol}(\boldsymbol{X}, \boldsymbol{y}) = \{\beta \in \mathrm{R}^{n+1} \mid \text{для любой } X \in \boldsymbol{X} \text{ выполнено } X\beta \in \boldsymbol{y}\}.$$

Задачу восстановления зависимостей по неточным данным будем решать по следующей общей схеме:

1) вводим количественную «меру согласования» (слабого или сильного) параметров зависимости и интервальных данных;

2) находим точку максимума этой меры и берем ее в качестве оценки параметров.

Ясно, что при достаточно разумном выборе «меры согласования» оценка параметров по данному способу всегда будет получена. Но совершенно необязательно, что реальное согласование полученных параметров и данных в самом деле будет иметь место. Иными словами, как и в традиционном неинтервальном случае, иногда может не существовать набора параметров, согласующихся с данными, т.е. линии, проходящей через все брусы неопределенности замеров в нужном нам смысле, обычном или сильном.

Основной вопрос, возникающий в связи с намеченным планом, состоит в том, какой взять количественную меру сильного согласования/несогласования параметров и данных?

Существуют естественные требования, которым эта мера должна удовлетворять. В процессе решения могут возникнуть две качественно отличные друг от друга ситуации, когда множество решений пусто и когда оно непусто. Выявление этого отличия можно возложить на нашу меру согласования/несогласования. При непустом множестве решений она должна быть положительной или по крайней мере неотрицательной для точек из этого множества, на которых согласование в самом деле достигается. Для точек вне множества решений, на которых согласования нет, она может быть отрицательной. Таким образом, знак величины этой меры будет служить признаком пустоты или непустоты множества решений. Кроме того, в случае непустого множества решений для точек его границы мера согласования должна быть не больше, чем для точек из его внутренности. Далее подробно рассмотрим сильное согласование параметров линейной зависимости и интервальных данных.

Интервальные системы линейных уравнений

При решении задачи восстановления линейной зависимости по данным с интервальной неопределенностью возникает интервальная система (4), (5) линейных алгебраических уравнений

$$\begin{aligned} \beta_0 + \boldsymbol{x}_{11}\beta_1 + \boldsymbol{x}_{12}\beta_2 + \ldots + \boldsymbol{x}_{1n}\beta_n &= \boldsymbol{y}_1, \\ \beta_0 + \boldsymbol{x}_{21}\beta_1 + \boldsymbol{x}_{22}\beta_2 + \ldots + \boldsymbol{x}_{2n}\beta_n &= \boldsymbol{y}_2, \\ &\ldots \\ \beta_0 + \boldsymbol{x}_{m1}\beta_1 + \boldsymbol{x}_{m2}\beta_2 + \ldots + \boldsymbol{x}_{mn}\beta_n &= \boldsymbol{y}_m, \end{aligned}$$

или (кратко)

$$\boldsymbol{X}\beta = \boldsymbol{y},$$

где $X = (\boldsymbol{x}_{ij})$ — интервальная $m \times (n + 1)$-матрица и $\boldsymbol{y} = (\boldsymbol{y}_i)$ — интервальный m-вектор. Это формальная запись, обозначающая семейство точечных линейных систем $X\beta = y$ той же структуры, что и интервальная система, с $X \in \boldsymbol{X}$ и $y \in \boldsymbol{y}$. Каждая система линейных алгебраических уравнений $X\beta = y$, матрица которой X взята из интервальной матрицы $\boldsymbol{X}$, а правая часть y из $\boldsymbol{y}$, может иметь решения, которые во многих ситуациях имеет смысл рассматривать совместно, единой совокупностью, т.е. объединив их. На этом пути мы получаем так называемое объединенное множество решений

$$\Xi_{uni}(\boldsymbol{X}, \boldsymbol{y}) = \{\beta \in \mathrm{R}^{n+1} \mid \text{существуют такие } X \in \boldsymbol{X} \text{ и } y \in \boldsymbol{y}, \text{ что } X\beta = y\}$$

(англоязычный термин — united solution set). Оно формализует, по-видимому, наиболее простое и естественное понимание «решения» интервальной системы уравнений. Этому множеству и различным способам его нахождения и оценивания посвящено огромное количество работ (см., в частности, [1, 3, 4, 14, 18, 19, 22, 25]).

Сильное согласование параметров и данных диктует другое понимание решений интервальной системы уравнений. Ему соответствует так называемое допусковое множество решений интервальной линейной системы уравнений — множество

$$\Xi_{tol}(\boldsymbol{X}, \boldsymbol{y}) = \{\beta \in \mathrm{R}^{n+1} \mid \text{для любой}$$

$$X \in \boldsymbol{X} \text{ справедливо } X\beta \in \boldsymbol{y}\}$$

(англоязычный термин — tolerable solution set). Это множество всевозможных векторов β, для которых произведение $X\beta$ попадает в интервалы правых частей $\boldsymbol{y}$ при любых $X \in \boldsymbol{X}$. Допусковое множество решений может оказаться пустым, если интервалы правой части «слишком узки» в сравнении с интервалами элементов матрицы. Тогда произведение $\boldsymbol{X}\beta$ получает «большой размах», который может не уместиться в «коридорах» правых частей системы.

Нетрудно понять, что всегда $\Xi_{tol}(\boldsymbol{X}, \boldsymbol{y}) \in \Xi_{uni}(\boldsymbol{X}, \boldsymbol{y})$, т.е. допусковое множество решений является подмножеством объединенного множества решений. В терминах задачи восстановления зависимостей это означает, что если имеет место сильное согласование параметров и данных, то тем более справедливо обычное согласование.

Существует несколько результатов, дающих аналитические описания допускового множества решений для интервальных линейных систем. В частности, теорема И. Рона [4, 30] представляет допусковое множество решений в виде решения системы линейных алгебраических неравенств. Поскольку задача решения таких систем имеет полиномиальную сложность (см., например [31]), то из теоремы И. Рона следует, что в общем случае распознавание пустоты/непустоты допускового множества решений и отыскание точки из него также могут выполняться за полиномиальное от размеров задачи время. Для этого особенно удобны развитые методы линейного программирования и реализующие их готовые пакеты программ и процедуры.

Чрезвычайно важным для понимания свойств сильного согласования параметров и интервальных данных в задаче восстановления линейной зависимости является следующий результат И. А. Шарой [32, 33][1].

Критерий неограниченности допускового множества решений. Непустое допусковое множество решений интервальной системы линейных алгебраических уравнений неограниченно тогда и только тогда, когда в матрице системы есть линейно зависимые неинтервальные столбцы.

[1] В работе [32] для наименования допускового множества решений интервальных систем уравнений используется устаревший термин «допустимое множество решений».

Напомним, что множество векторов линейного пространства называется линейно зависимым, если существуют такие скаляры, не все равные нулю, что линейная комбинация векторов с этими скалярами равна нулевому вектору. Приведенный критерий ограниченности показывает, что допусковое множество решений неограниченно лишь в исключительных случаях, которые заведомо не выполняются, если входные переменные имеют существенные интервальные неопределенности. Это следует из того, что при сложении интервалов их ширина не уменьшается (см. следующий раздел), и потому нетривиальная линейная комбинация невырожденных интервалов никогда не сможет занулиться.

Метод распознающего функционала

Кратко изложим известные результаты о допусковом множестве решений, опубликованные ранее, в частности, в [4, 29]. Далее понадобится классическая интервальная арифметика IR — алгебраическая система, образованная интервалами $\boldsymbol{x} = [\inf \boldsymbol{x}, \sup \boldsymbol{x}]$ вещественной оси R так, что для любой арифметической операции «$*$» из множества $\{+, -, \cdot, /\}$ результат операции между интервалами определяется «по представителям», т.е. как

$$\boldsymbol{x} * \boldsymbol{y} = \{x * y \mid x \in \boldsymbol{x}, y \in \boldsymbol{y}\}.$$

Развернутые конструктивные формулы для арифметических операций выглядят следующим образом:

$$\boldsymbol{x} + \boldsymbol{y} = [\inf \boldsymbol{x} + \inf \boldsymbol{y}, \sup \boldsymbol{x} + \sup \boldsymbol{y}],$$

$$\boldsymbol{x} - \boldsymbol{y} = [\inf \boldsymbol{x} - \sup \boldsymbol{y}, \sup \boldsymbol{x} - \inf \boldsymbol{y}],$$

$$\boldsymbol{x} * \boldsymbol{y} = [\min S, \max S],$$

где $S = \{\inf \boldsymbol{x} * \inf \boldsymbol{y}, \ \inf \boldsymbol{x} * \sup \boldsymbol{y}, \ \sup \boldsymbol{x} * \inf \boldsymbol{y}, \sup \boldsymbol{x} * \sup \boldsymbol{y}\}$,

$$\boldsymbol{x}/\boldsymbol{y} = \boldsymbol{x} * [1/\sup \boldsymbol{y}, 1/\inf \boldsymbol{y}] \text{ при } 0 \in \boldsymbol{y}.$$

Отправной точкой дальнейших построений является следующий результат: для интервальной $m \times (n + 1)$-системы линейных алгебраических уравнений $X\beta = y$ точка $\beta \in \mathrm{R}^{n+1}$ принадлежит допусковому множеству решений $\Xi_{tol}(\boldsymbol{X}, \boldsymbol{y})$ тогда и только тогда, когда

$$\boldsymbol{X} \cdot \beta \in \boldsymbol{y}, \qquad (6)$$

где «$\cdot$» — интервальное матричное умножение. Справедливость этого описания следует из свойств интервального матрично-векторного умножения и определения допускового множества решений (см. [4, 29]). Преобразуем включение (6) в другую форму. Прежде всего перепишем (6) в виде равносильной системы покомпонентных интервальных включений. По определению

$$(\boldsymbol{X}\cdot\beta)_i = \sum_{j=0}^{n}\boldsymbol{x}_{ij}\beta_j,\ i = 1, 2, ..., m,$$

где удобно считать, что самый первый столбец матрицы $\boldsymbol{X}$, составленный из единиц, имеет нулевой номер. Тогда вместо (6) можно написать

$$\sum_{j=0}^{n}\boldsymbol{x}_{ij}\beta_j \in \boldsymbol{y}_i,\ i = 1, 2, ..., m. \qquad (7)$$

Обозначим

$$\operatorname{mid}\boldsymbol{y}_i = \frac{1}{2}(\sup\boldsymbol{y}_i - \inf\boldsymbol{y}_i) \text{ — середина интервала } \boldsymbol{y}_i,$$

$$\operatorname{rad}\boldsymbol{y}_i = \frac{1}{2}(\inf\boldsymbol{y}_i + \sup\boldsymbol{y}_i) \text{ — радиус интервала } \boldsymbol{y}_i.$$

Представим правые части включений (7) в виде сумм средних точек и уравновешенных интервалов $[-\operatorname{rad}\boldsymbol{y}_i, \operatorname{rad}\boldsymbol{y}_i]$:

$$\sum_{j=0}^{n}\boldsymbol{x}_{ij}\beta_j \in \operatorname{mid}\boldsymbol{y}_i + [-\operatorname{rad}\boldsymbol{y}_i, \operatorname{rad}\boldsymbol{y}_i],\ i = 1, 2, ..., m.$$

Добавляя теперь к обеим частям включений по $(-\operatorname{mid}\boldsymbol{y}_i)$, получим

$$\sum_{j=0}^{n}\boldsymbol{x}_{ij}\beta_j - \operatorname{mid}\boldsymbol{y}_i \in [-\operatorname{rad}\boldsymbol{y}_i, \operatorname{rad}\boldsymbol{y}_i],\ i = 1, 2, ..., m.$$

Но включение интервала в уравновешенный интервал $[-\operatorname{rad}\boldsymbol{y}_i, \operatorname{rad}\boldsymbol{y}_i]$ можно эквивалентным образом записать как неравенство

$$\left|\operatorname{mid}\boldsymbol{y}_i - \sum_{j=0}^{n}\boldsymbol{x}_{ij}\beta_j\right| \le \operatorname{rad}\boldsymbol{y}_i,\ i = 1, 2, ..., m,$$

что равносильно

$$\operatorname{rad}\boldsymbol{y}_i - \left|\operatorname{mid}\boldsymbol{y}_i - \sum_{j=0}^{n}\boldsymbol{x}_{ij}\beta_j\right| \ge 0,\ i = 1, 2, ..., m.$$

Поэтому в целом

$$X\cdot\beta \in \boldsymbol{y} \Leftrightarrow \operatorname{rad}\boldsymbol{y}_i - \left|\operatorname{mid}\boldsymbol{y}_i - \sum_{j=0}^{n}\boldsymbol{x}_{ij}\beta_j\right| \ge 0,$$

$i = 1, 2, ..., m$.

Наконец, с помощью операции минимума можно свернуть по i конъюнкцию неравенств в правой части полученной логической эквивалентности:

$$\boldsymbol{X}\cdot\beta \in \boldsymbol{y} \Leftrightarrow \min_{1\le i\le m}\left\{\operatorname{rad}\boldsymbol{y}_i - \left|\operatorname{mid}\boldsymbol{y}_i - \sum_{j=0}^{n}\boldsymbol{x}_{ij}\beta_j\right|\right\} \ge 0.$$

Мы пришли к следующему результату:

Теорема. Пусть $\boldsymbol{X}$ — интервальная $m\times(n + 1)$-матрица, $\boldsymbol{y}$ — интервальный m-вектор. Тогда выражением

$$\operatorname{Tol}(\beta, \boldsymbol{X}, \boldsymbol{y}) = \min_{1\le i\le m}\left\{\operatorname{rad}\boldsymbol{y}_i - \left|\operatorname{mid}\boldsymbol{y}_i - \sum_{j=0}^{n}\boldsymbol{x}_{ij}\beta_j\right|\right\}$$

задается отображение Tol: $\mathbb{R}^{n+1}\times\mathbb{IR}^{m\times(n+1)}\times\mathbb{IR}^m \to \mathbb{R}$ такое, что принадлежность точки $\beta\in\mathbb{R}^{n+1}$ допусковому множеству решений $\Xi_{tol}(\boldsymbol{X}, \boldsymbol{y})$ интервальной линейной системы уравнений $\boldsymbol{X}\beta = \boldsymbol{y}$ равносильна неотрицательности Tol в точке β, т.е.

$$\beta \in \Xi_{tol}(\boldsymbol{X}, \boldsymbol{y}) \Leftrightarrow \operatorname{Tol}(\beta, \boldsymbol{X}, \boldsymbol{y}) \ge 0.$$

Таким образом, допусковое множество решений $\Xi_{tol}(\boldsymbol{X}, \boldsymbol{y})$ интервальной линейной системы является множеством уровня

$$\{\beta\in\mathbb{R}^{n+1} \mid \operatorname{Tol}(\beta, \boldsymbol{X}, \boldsymbol{y}) \ge 0\}$$

отображения Tol. Будем называть это отображение «распознающим функционалом» допускового множества решений, так как его областью значений является числовое множество $\mathbb{R}$, т.е. вещественная ось, а посредством знака своих значений Tol «распознает» принадлежность точки множеству $\Xi_{tol}(\boldsymbol{X}, \boldsymbol{y})$.

Конспективно изложим свойства распознающего функционала. Их подробные доказательства можно найти в [4, 29].

Прежде всего, функционал Tol непрерывен по всем своим переменным. Функционал Tol — вогнутый по β всюду в $\mathbb{R}^{n+1}$ и, как следствие, унимодальный. Функционал $\operatorname{Tol}(\beta, \boldsymbol{X}, \boldsymbol{y})$ — полиэдральный, т.е. его подграфик — полиэдральное множество, а график составлен из кусков гиперплоскостей (рис. 4). Наконец, $\operatorname{Tol}(\beta, \boldsymbol{X}, \boldsymbol{y})$ достигает конечного максимума на всем пространстве $\mathbb{R}^{n+1}$.

Если $\operatorname{Tol}(\beta, \boldsymbol{X}, \boldsymbol{y}) > 0$, то β — точка топологической внутренности $\operatorname{int}\Xi_{tol}(\boldsymbol{X}, \boldsymbol{y})$ допускового множества решений. Поясним, что точка топологической внутренности — это точка множества, принадлежащая ему вместе с некоторым шаром (относительно какой-то нормы), имеющим центр в этой точке. Следовательно, точки из внутренности множества остаются принадлежащими этому множеству даже при их малых «шевелениях», что

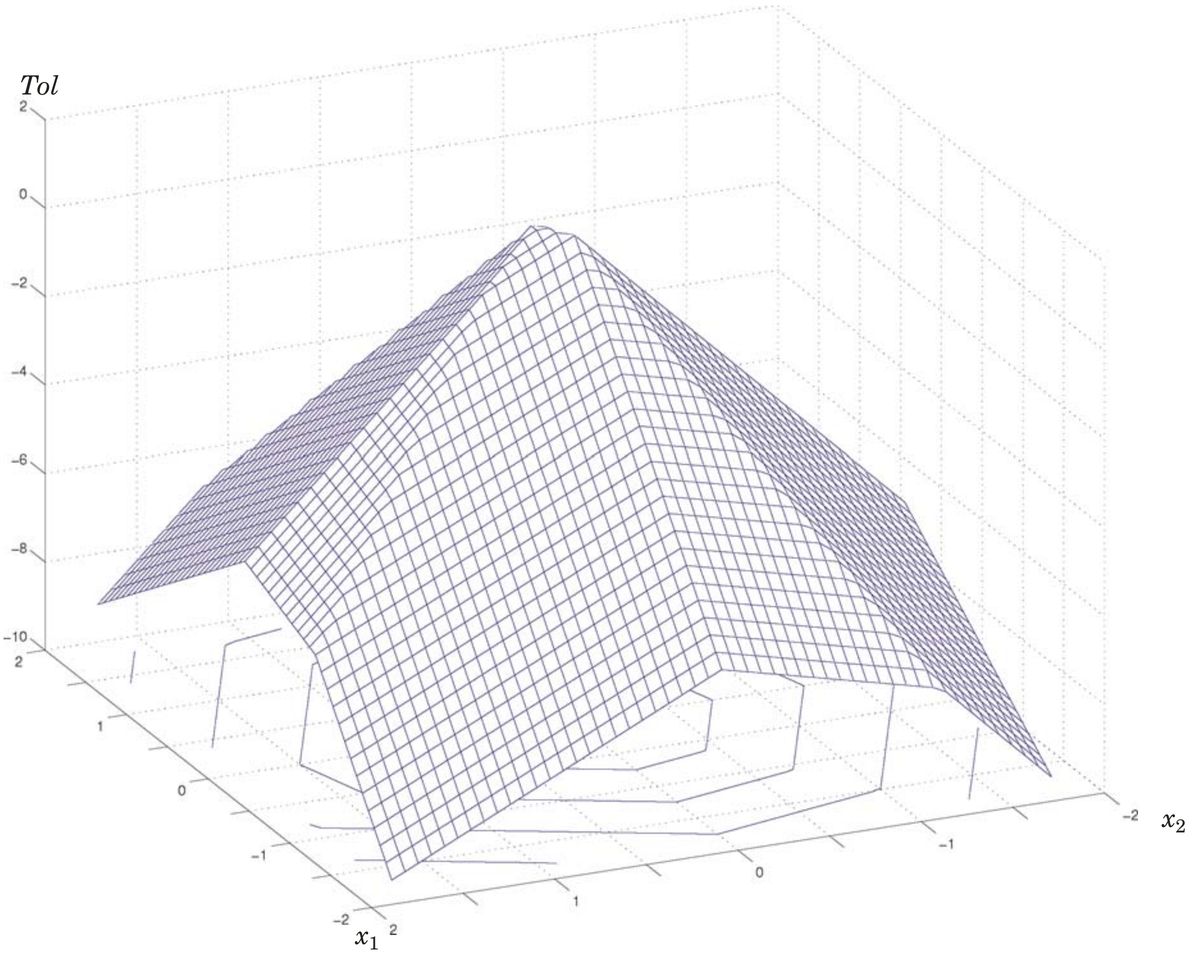

Рис. 4. Типичный график распознающего функционала допускового множества решений

Fig. 4. A typical graph of the recognizing functional for a tolerable solution set

нередко важно для практики. При определенных необременительных условиях на интервальную систему $\boldsymbol{X}\beta = \boldsymbol{y}$ верно и обратное, т.е. из принадлежности $\beta \in \operatorname{int} \Xi_{tol}(\boldsymbol{X}, \boldsymbol{y})$ следует строгое неравенство $\mathrm{Tol}(\beta, \boldsymbol{X}, \boldsymbol{y}) > 0$ [4, 29].

Как следствие сформулированных результатов, с помощью распознающего функционала можно выполнить исследование непустоты/пустоты допускового множества решений интервальной системы линейных алгебраических уравнений $\boldsymbol{X}\beta = \boldsymbol{y}$ по следующей схеме. Решаем задачу безусловной максимизации распознающего функционала $\mathrm{Tol}(\beta, \boldsymbol{X}, \boldsymbol{y})$. Пусть найденный максимум достигается в точке $\beta \in \mathrm{R}^{n+1}$ и равен

$$U = \max_{\beta \in \mathrm{R}^{n+1}} \mathrm{Tol}(\beta, \boldsymbol{X}, \boldsymbol{y}).$$

Тогда:

если $U \geq 0$, то $\tau \in \Xi_{tol}(\boldsymbol{X}, \boldsymbol{y}) \in \varnothing$, т.е. допусковое множество решений системы $\boldsymbol{X}\beta = \boldsymbol{y}$ непусто и τ лежит в нем;

если $U > 0$, то $\tau \in \operatorname{int} \Xi_{tol}(\boldsymbol{X}, \boldsymbol{y}) \in \varnothing$, и принадлежность точки τ допусковому множеству решений устойчива к малым возмущениям $\boldsymbol{X}$ и $\boldsymbol{y}$;

если $U < 0$, то $\Xi_{tol}(\boldsymbol{X}, \boldsymbol{y}) = \varnothing$, т.е. допусковое множество решений интервальной системы линейных уравнений $\boldsymbol{X}\beta = \boldsymbol{y}$ пусто.

Сама величина $U = \max \mathrm{Tol}$ может служить количественной мерой «запаса разрешимости» задачи (при $U \geq 0$) или же ее «дефицита разрешимости» (при $U < 0$).

Метод максимума согласования: «сильная версия»

Результаты предшествующего раздела можно положить в основу подхода к решению задачи восстановления линейной зависимости по неточным данным, которое удовлетворяет требованию сильной согласованности данных и параметров.

В соответствии с планом, намеченным в конце первого раздела, необходимо ввести «меру сильного согласования / несогласования» параметров и данных. Напомним, что при непустом допусковом множестве решений она должна быть положительной для точек из этого множества, на которых «сильное согласование» в самом деле достигается. Для точек вне допускового множества решений, на которых «сильного согласования» нет, она может быть отрицательной. Таким образом, на роль меры согласования очень подходит распознающий функционал Tol. В частности, Tol различает точки границы и внутренности допускового множества решений.

Мы приходим к методу оценивания параметров линейной зависимости по неточным данным, который будем называть методом максимума согласования.

Оценкой параметров берем точку, в которой достигается наибольшее значение распознающего функционала Tol. Следствие теории предшествующего раздела:

если $\max \mathrm{Tol} \geq 0$, то найденная точка лежит во множестве параметров, сильно согласующихся с данными;

если $\max \mathrm{Tol} < 0$, то не существует параметров, сильно согласующихся с данными, но найденная точка минимизирует «несогласованность» параметров и данных.

Содержательная интерпретация метода максимума согласования в случае непустого множества решений $\Xi_{tol}(\boldsymbol{X}, \boldsymbol{y})$ может быть, например, такой: оценка параметров, т.е. аргумент, на котором достигается $\max \mathrm{Tol}$, — это последняя точка, которая остается в непустом допусковом множестве решений при равномерном сужении вектора правой части относительно его середины [4, 29].

Отметим также, что для случая точечных данных сильная версия метода максимума согласования (так же, как и обычная «слабая») совпадает с так называемым чебышевским сглаживанием данных, которое давно и успешно применяется при обработке данных (см., например, [34]). В самом деле, если матрица входных данных $\boldsymbol{X}$ и вектор выходных данных $\boldsymbol{y}$ — точечные (неинтервальные), т.е. $\boldsymbol{X} = X = (x_{ij})$ и $\boldsymbol{y} = y = (y_i)$, то для всех индексов i, j

$$\mathrm{rad}\, \boldsymbol{y}_i = 0,\ \mathrm{mid}\, \boldsymbol{y}_i = y_i,\ \boldsymbol{x}_{ij} = x_{ij}.$$

Распознающий функционал допускового множества решений принимает при этом вид

$$\mathrm{Tol}(\beta, \boldsymbol{X}, \boldsymbol{y}) = \min_{1 \leq i \leq m} \left\{ -\left| y_i - \sum_{j=0}^{n} x_{ij} \beta_j \right| \right\} =$$

$$-\max_{1 \leq i \leq m} \left| y_i - \sum_{j=0}^{n} x_{ij} \beta_j \right| =$$

$$= -\max_{1 \leq i \leq m} |(X\beta)_i - y_i| = ||X\beta - y||_\infty.$$

Здесь посредством $||\cdot||\infty$ обозначается чебышевская норма (∞-норма) вектора в конечномерном пространстве R^m, которая определяется как

$$||z||_\infty = \max_{1 \leq i \leq m} |z_i|.$$

Тогда

$$\max_{\beta \in \mathrm{R}^{n+1}} \mathrm{Tol}(\beta) = \max_{\beta \in \mathrm{R}^{n+1}} (-||X\beta - y||_\infty) = -\min_{\beta \in \mathrm{R}^{n+1}} ||X\beta - y||_\infty,$$

поскольку $\max(-f(\beta)) = -\min f(\beta)$. Таким образом, максимизация распознающего функционала равносильна в этом случае минимизации чебышевской нормы невязки решения.

Обсуждение

Каковы преимущества метода максимума согласования в представленной нами сильной версии? Прежде всего, это полиномиальная сложность вычислительных алгоритмов, которые обеспечивают нахождение оценки. Напомним, что в слабой версии метода максимума согласования распознавание пустоты/непустоты и оценивание объединенного множества решений являются NP-трудными задачами, которые требуют для своего решения в общем случае экспоненциально сложных алгоритмов [4].

Другим важным преимуществом сильной версии метода максимума согласования является робастность оценок, понимаемая как их устойчивость к возмущениям в данных. Она вытекает из того факта, что допусковое множество решений интервальных систем линейных алгебраических уравнений является наиболее устойчивым из всех множеств решений и переход к нему имеет регуляризующий эффект [35].

Еще одно преимущество «сильного» метода максимума согласования, тесно связанное с предыдущим, — конечная вариабельность (изменчивость) оценок, получаемых с его помощью. Критерий И. А. Шарой неограниченности допускового множества решений [32, 33] имеет своим следствием то обстоятельство, что множество параметров, удовлетворяющих условию сильного согласования, почти всегда оказывается ограниченным. Если для независимых (предикторных) переменных данные существенно интервальны, то эта ограниченность имеет место даже тогда, когда составленная из них матрица имеет неполный ранг и когда измерений меньше, чем неизвестных параметров. Отмеченное свойство иллюстрирует пример, приведенный в следующем разделе.

Наконец, восстановление зависимостей по интервальным данным с помощью сильной версии максимума согласования лишь частично подвержено так называемому «парадоксу Е. З. Демиденко», сформулированному в [36]. Его суть может быть кратко и емко выражена фразой «чем лучше, тем хуже». Дело в том, что присутствие неопределенности в данных является нежелательным явлением, которое искажает истинную картину реальности, а потому уменьшение этой неопределенности, т.е. сужение интервалов данных, является позитивным фактом, который, казалось бы, должен приветствоваться и приводить к лучшему решению задачи восстановления зависимостей. Но при более узких интервалах исходных данных объединенное множество решений интервальной системы уравнений также сужается и может вообще сделаться пустым. Выбирать из него параметры становится труднее, чем для широких исходных данных задачи. Иными словами, чем выше точность исходных данных и меньше их интервальная неопределенность, тем хуже условия для оценивания параметров. И наоборот, чем шире интервальные неопределенности, чем меньше знаем о точных значениях измеряемых величин, тем лучше для процесса оценивания параметров и тем более богатый набор результатов можно получить. Именно так обстоит дело для слабой версии метода максимума согласования и многих других методов восстановления зависимостей по интервальным данным, которые явно или неявно основываются на слабом согласовании параметров и данных и, как следствие, на использовании объединенного множества решений.

Анализ «парадокса Е. З. Демиденко» и способов его преодоления можно найти в работах [18, 19]. Но для сильного согласования параметров и данных ситуация меняется. Чем шире интервалы данных по входным (предикторным) переменным восстанавливаемой зависимости, тем меньше допусковое множество решений и тем труднее выбирать из него подходящие параметры! Для интервалов выходных (критериальных) переменных все остается по-прежнему. Но в целом поведение оценок сильного согласования является парадоксальным «по Демиденко» лишь частично.

Реализация

Развитая в предшествующих разделах теория будет практичной и реально полезной лишь в том случае, когда в нашем распоряжении имеются эффективные методы для нахождения величины max Tol, т.е. максимума распознающего функционала допускового множества решений. Свойства распознающего функционала рассмотрены выше, и они являются достаточно благоприятными для применения эффективных численных методов оптимизации.

В общем случае задача вычисления max Tol — это задача безусловной максимизации вогнутой негладкой целевой функции. Ее решение может опираться на методы негладкой выпуклой оптимизации, интенсивно развиваемые уже в течение нескольких десятков лет различными научными школами у нас в стране и за рубежом. Автор, в частности, давно и успешно использует результаты работ Н. З. Шора и его сотрудников из Института кибернетики НАН Украины (см., в частности, [37, 38]).

Более десятилетия автором свободно распространяется программа tolsolvty, которую можно загрузить с веб-сайта «Интервальный анализ и его приложения» — http: //www.nsc.ru/interval (раздел «Программное обеспечение» и далее «Некоторые интервальные программы для Scilab» или «Некоторые интервальные программы для MATLAB»). Программа предназначена для численного нахождения безусловного максимума распознающего функционала Tol и использует в качестве основы алгоритм ralgb5, созданный П. И. Стецюком (ему посвящена статья [38]). Фактически, tolsolvty — очень хорошая и проверенная временем реализация метода максимума согласования в сильном смысле, которую можно рекомендовать для решения практических задач.

Сравнительно недавно появилась возможность использовать для нахождения максимума распознающего функционала Tol методы отделяющих плоскостей, предложенные Е. А. Нурминским [39] и развитые далее Е. А. Воронцовой [40]. На веб-сайте «Интервальный анализ и его приложения» выложена свободная программа tolspaclip, реализующая метод отделяющих плоскостей с дополнительным отсечением, предназначенная для тех же целей, что и tolsolvty. Методы отсекающих плоскостей хорошо работают при размерностях пространства параметров до нескольких тысяч.

Пример

Рассмотрим в качестве примера восстановление линейной зависимости

$$y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3, \tag{8}$$

по интервальным данным двух измерений, приведенным ниже:

x_1	x_2	x_3	y
[98, 100]	[97, 99]	[96, 98]	[190, 210]
[99, 101]	[98, 100]	[97, 99]	[200, 220]

Отметим, что в этих данных брусы неопределенности двух замеров существенно «налегают» друг на друга: их пересечением является брус с непустой внутренностью, размеры которого сравнимы с размерами исходных брусов данных.

Для определения параметров зависимости (8) приходим к интервальной системе линейных алгебраических уравнений

$$\begin{aligned} &[98, 100]\beta_1 + [97, 99]\beta_2 + [96, 98]\beta_3 = [190, 210] \\ &[99, 101]\beta_1 + [98, 100]\beta_2 + [97, 99]\beta_3 = [200, 220], \end{aligned} \quad (9)$$

которая является недоопределеной. Более того, интервальная матрица системы содержит матрицу неполного ранга 1

$$\begin{pmatrix} 98 & 98 & 98 \\ 99 & 99 & 99 \end{pmatrix}$$

Объединенное множество решений системы (9) является неограниченным, так что восстановление зависимости на основе обычного понятия согласования (Определение 1) представляется затруднительным. Тем не менее и в этих неблагоприятных условиях допусковое множество решений интервальной системы уравнений непусто и ограниченно (см. критерий И. А. Шарой).

Нахождение максимума распознающего функционала этой системы с помощью программы tolsolvty дает значение $\max \mathrm{Tol} = 3{,}9698$, которое достигается в точке

$$\arg\max \mathrm{Tol} = (2{,}0603,\ 3 \cdot 10^{-6},\ 2{,}1 \cdot 10^{-6}).$$

Ее можно взять в качестве оценки коэффициентов.

Полученный ответ приводит к гипотезе о том, что значимым в восстанавливаемой зависимости является только первый коэффициент β_1, тогда как β_2 и β_3, будучи почти нулевыми, на исследуемую зависимость никак не влияют. Для исследования этого вопроса нужно предпринять более детальное оценивание допускового множества решений с помощью методов, описанных, к примеру, в [4, 29].

Внутренние оценки допускового множества решений системы (7) оказываются малыми по размерам интервалами, тогда как его внешние оценки — интервалы примерно [–2, 4] по каждой из трех координат. Это свидетельствует о значительной протяженности множества решений, которое в целом является тонкой «пластинкой» со значительными размерами. Получается, что гипотеза о «незначимости» коэффициентов β_2 и β_3 не может быть решена чисто математическими средствами и для своего подтверждения или опровержения требует дополнительных содержательных соображений.

Итоги и выводы

В задачах восстановления зависимостей по данным с интервальной неопределенностью следует различать возможные способы согласования интервальных данных с параметрами конструируемой зависимости. В частности, необходимо ввести понятия сильного и слабого согласования данных и параметров, которые соответствуют различной роли входных (предикторных) переменных и выходных (критериальных) переменных в процессе измерения.

Рассмотренный в работе метод максимума согласования — перспективный метод восстановления зависимостей для данных с интервальными неопределенностями и неточностями, основанный на максимизации распознающего функционала множества решений задачи. Он является обобщением метода чебышевского сглаживания данных и может служить хорошей альтернативой традиционным методам регрессионного анализа, использующим теоретико-вероятностные модели ошибок (в частности, популярному методу наименьших квадратов). В работе рассмотрена модификация метода максимума согласования для случая, когда параметры зависимости и данные должны удовлетворять условию сильного согласования. Показаны ее преимущества перед «слабой версией» и некоторыми другими методами восстановления зависимостей по интервальным данным. Наконец, даны рекомендации по практической реализации предложенных подходов.

Благодарности

Автор благодарен сотрудникам ВНИИМ им. Д. И. Менделеева докт. техн. наук А. Г. Чуновкиной и А. А. Королевой за полезные обсуждения и помощь в работе.

ЛИТЕРАТУРА

1. **Жолен Л., Кифер М., Дидри О., Вальтер Э.** Прикладной интервальный анализ. — Москва – Ижевск: Издательство «РХД», 2007. — 468 с.
2. Интервальный анализ и его приложения. — Тематический веб-сайт, http://www.nsc.ru/interval.
3. **Moore R. E., Kearfott R. B., Cloud M. J.** Introduction to Interval Analysis. — Philadelphia: SIAM, 2009. — 223 p.
4. **Шарый С. П.** Конечномерный интервальный анализ. — Новосибирск: Институт вычислительных технологий СО РАН, 2019. — 631 с. Электронная книга, доступная на http://www.nsc.ru/interval/InteBooks.
5. **Kearfott R. B., Nakao M., Neumaier A., Rump S., Shary S. P., van Hentenryck P.** Standardized notation in interval analysis / Вычислительные Технологии. 2010. Т. 15. № 1. С. 7 – 13.
6. **Канторович Л. В.** О некоторых новых подходах к вычислительным методам и обработке наблюдений / Сибирский математический журнал. 1962. Т. 3. № 5. С. 701 – 709.
7. **Вощинин А. П., Бочков А. Ф., Сотиров Г. Р.** Метод анализа данных при интервальной нестатистической ошибке / Заводская лаборатория. 1990. Т. 56. № 7. С. 76 – 81.
8. **Вощинин А. П.** Интервальный анализ данных: развитие и перспективы / Заводская лаборатория. Диагностика материалов. 2002. Т. 68. № 1. С. 118 – 126.

9. **Скибицкий Н. В.** Построение прямых и обратных статических характеристик объектов по интервальным данным / Заводская лаборатория. Диагностика материалов. 2017. Т. 83. № 1. Ч. I. С. 87 – 93.
10. **Суханов В. А.** Исследование эмпирических зависимостей: нестатистический подход. — Барнаул: Издательство Алтайского университета, 2007. — 290 с.
11. **Оскорбин Н. М., Максимов А. В., Жилин С. И.** Построение и анализ эмпирических зависимостей методом центра неопределенностей / Известия Алтайского государственного университета. 1998. № 1. С. 37 – 40.
12. **Zhilin S. I.** On fitting empirical data under interval error / Reliable Computing. 2005. Vol. 11. P. 433 – 442. DOI: 10.1007/s11155-005-0050-3.
13. **Спивак С. И., Кантор О. Г., Юнусова Д. С.** Идентификация и информативность моделей количественного анализа многокомпонентных смесей / Журнал Средневолжского математического общества. 2016. Т. 18. № 3. С. 153 – 163.
14. **Поляк Б. Т., Назин С. А.** Оценивание параметров в линейных многомерных системах с интервальной неопределенностью / Проблемы управления и информатики. 2006. № 1, 2. С. 103 – 115.
15. **Кумков С. И.** Обработка экспериментальных данных ионной проводимости расправленного электролита методами интервального анализа / Расплавы. 2010. № 3. С. 79 – 89.
16. **Померанцев А. Л., Родионова О. Е.** Построение многомерной градуировки методом простого интервального оценивания / Журнал аналитической химии. 2006. Т. 61. № 10. С. 1032 – 1047.
17. **Подружко А. А., Подружко А. С.** Интервальное представление полиномиальных регрессий. — М.: Эдиториал УРСС, 2003. — 47 с.
18. **Шарый С. П.** Разрешимость интервальных линейных уравнений и анализ данных с неопределенностями / Автоматика и Телемеханика. 2012. № 2. С. 111 – 125.
19. **Шарый С. П., Шарая И. А.** Распознавание разрешимости интервальных уравнений и его приложения к анализу данных / Вычислительные технологии. 2013. Т. 18. № 3. С. 80 – 109.
20. **Шарый С. П.** Сильная согласованность в задачах восстановления зависимостей по интервальным данным / Вестник ЮУрГУ. Серия «Математика. Механика. Физика». 2017. Т. 9. № 1. С. 39 – 48.
21. **Шарый С. П.** Сильная согласованность в задаче восстановления зависимостей при интервальной неопределенности данных / Вычислительные технологии. 2017. Т. 22. № 2. С. 150 – 172.
22. **Шарый С. П.** Метод максимума согласования для восстановления зависимостей по данным с интервальной неопределенностью / Известия РАН. Теория и системы управления. 2017. № 6. С. 3 – 19.
23. **Schweppe F. C.** Recursive state estimation: unknown but bounded errors and system inputs / IEEE Trans. on Automatic Control. 1968. AC-13. P. 22 – 28.
24. **Combettes P. L.** Foundations of set-theoretic estimation / Proc. IEEE. 1993. Vol. 81. N 2. P. 182 – 208.
25. **Milanese M., Norton J., Piet-Lahanier H., Walter E., eds.** Bounding Approaches to System Identification. — New York: Plenum Press, 1996. — 567 p. DOI: 10.1007/978-1-4757-9545-5.
26. **Лемешко Б. Ю., Постовалов С. Н.** О решении задач статистического анализа интервальных наблюдений / Вычислительные технологии. 1997. Т. 2. № 1. С. 28 – 36.
27. **Орлов А. И., Луценко Е. В.** Системная нечеткая интервальная математика. — Краснодар: Издательство КубГАУ, 2014. — 600 с.
28. **Орлов А. И.** Статистика интервальных данных (обобщающая статья) / Заводская лаборатория. Диагностика материалов. 2015. Т. 81. № 3. С. 61 – 69.
29. **Шарый С. П.** Решение интервальной линейной задачи о допусках / Автоматика и телемеханика. 2004. № 7. С. 147 – 162.
30. **Rohn J.** Inner solutions of linear interval systems / Interval Mathematics 1985 / K. Nickel, ed. Lecture Notes in Computer Science 212. — Berlin: Springer, 1986. P. 157 – 158.
31. **Схрейвер А.** Теория линейного и целочисленного программирования. Т. 1. — М.: Мир, 1991. — 360 с.
32. **Шарая И. А.** Ограничено ли допустимое множество решений интервальной системы? / Вычислительные технологии. 2004. Т. 9. № 3. С. 108 – 112.
33. **Sharaya I. A.** On unbounded tolerable solution sets / Reliable Computing. 2005. Vol. 11. N 5. P. 425 – 432. DOI: 10.1007/s11155-005-0049-9.
34. **Ремез Е. Я.** Основы численных методов чебышевского приближения. — Киев: Наукова думка, 1969. — 624 с.
35. **Shary S. P.** Interval regularization for imprecise linear algebraic equations. Статья, депонированная в репозитории arXiv.org 27 сентября 2018 года, номер arXiv: 1810.01481. — 21 с.
36. **Демиденко Е. З.** Комментарий II к статье А. П. Вощинина, А. Ф. Бочкова и Г. Р. Сотирова «Метод анализа данных при интервальной нестатистической ошибке» / Заводская лаборатория. 1990. Т. 56. № 7. С. 83 – 84.
37. **Шор Н. З., Журбенко Н. Г.** Метод минимизации, использующий операцию растяжения пространства в направлении разности двух последовательных градиентов / Кибернетика. 1971. № 3. С. 51 – 59.
38. **Стецюк П. И.** Субградиентные методы ralgb5 и ralgb4 для минимизации овражных выпуклых функций / Вычислительные технологии. 2017. Т. 22. № 2. С. 127 – 149.
39. **Nurminski E. A.** Separating plane algorithms for convex optimization / Mathematical Programming. 1997. Vol. 76. P. 373 – 391. DOI: 10.1007/BF02614389.
40. **Воронцова Е. А.** Линейная задача о допусках для интервальной модели межотраслевого баланса / Вычислительные технологии. 2017. Т. 22. № 2. С. 67 – 84.

REFERENCES

1. **Jaulin L., Kieffer M., Didrit O., Walter E.** Applied Interval Analysis. — London, Springer, 2001. — 379 p. DOI: 10.1007/978-1-4471-0249-6.
2. Interval analysis and its applications. — A thematic web site, URL: http://www.nsc.ru/interval [in Russian].
3. **Moore R. E., Kearfott R. B., Cloud M. J.** Introduction to Interval Analysis. — Philadelphia: SIAM, 2009. — 223 p.
4. **Shary S. P.** Finite-dimensional Interval Analysis. — Novosibirsk: Institute of Computational Technologies SB RAS, 2019. — 631 p. An electronic book, accessible at http://www.nsc.ru/interval/InteBooks [in Russian].
5. **Kearfott R. B., Nakao M., Neumaier A., Rump S., Shary S. P., van Hentenryck P.** Standardized notation in interval analysis / Vychislit. Tekhnol. 2010. Vol. 15. B 1. P. 7 – 13.
6. **Kantorovich L. V.** On some new approaches to computational methods and observations processing / Sib. Matem. Zh. 1962. Vol. 3. N 5. P. 701 – 709 [in Russian].
7. **Voshchinin A. P., Bochkov A. F., Sotirov G. R.** A method of data analysis under interval non-statistical error / Zavod. Lab. 1990. Vol. 56. N 7. P. 76 – 81 [in Russian].
8. **Voshchinin A. P.** Interval data analysis: development and perspectives / Zavod. Lab. Diagn. Mater. 2002. Vol. 68. N 1. P. 118 – 126 [in Russian].
9. **Skibitskiy N. V.** Construction of direct and inverse static characteristics of the objects by interval data / Zavod. Lab. Diagn. Mater. 2017. Vol. 83. N 1. Part I. P. 87 – 93 [in Russian].
10. **Sukhanov V. A.** Study of Empirical Dependencies: Nonstatistical Approach. — Barnaul: Izd. Altai. Gos. Univ., 2007. — 290 p. [in Russian].
11. **Oskorbin N. M., Maksimov A. V., Zhilin S. I.** Construction and analysis of empirical dependencies by the method of uncertainty center / Izv. Altai. Gos. Univ. 1998. N 1. P. 37 – 40 [in Russian].
12. **Zhilin S. I.** On fitting empirical data under interval error / Reliable Computing. 2005. Vol. 11. P. 433 – 442. DOI: 10.1007/s11155-005-0050-3.
13. **Spivak S. I., Kantor O. G., Yunusova D. S.** Identification and inf ormativity of models for quantitative analysis of multi-

component mixtures / Zh. Srednevolzh. Matem. Obsh. 2016. Vol. 18. N 3. P. 153 – 163 [in Russian].

14. **Polyak B. T., Nazin S. A.** Estimation of parameters in linear multidimensional systems under interval uncertainty / Journal of Automation and Information Sciences. 2006. Vol. 38. N 2. P. 1 – 5.
15. **Kumkov S. I.** Processing experimental data of ionic conduction of molten electrolyte by methods of interval analysis / Rasplavy. 2010. N 3. P. 79 – 89 [in Russian].
16. **Pomerantsev A. L., Rodionova O. Ye.** Construction of a multivariate calibration by the simple interval calculation method / Journal of Analytical Chemistry. 2006. Vol. 61. N 10. P. 952 – 966. DOI: 10.1134/S1061934806100030.
17. **Podruzhko A. A., Podruzhko A. S.** Interval Representation of Polynomial Regressions. — Moscow: Editorial URSS, 2003. – 47 p. [in Russian].
18. **Shary S. P.** Solvability of interval linear equations and data analysis under uncertainty / Automation and Remote Control. 2012. Vol. 73. P. 310 – 322. DOI: 10.1134/S0005117912020099.
19. **Shary S. P., Sharaya I. A.** Recognition of solvability of interval equations and its applications to data analysis / Vychisl. Tekhnol. 2013. Vol. 18. N 3. P. 80 – 109 [in Russian].
20. **Shary S. P.** Strong compatability in data fitting problems based on interval data / Vestn. Yuzhno-Ural. Gos. Univ. Ser. Matem. Mekh. Fiz. 2017. Vol. 9. N 1. P. 39 – 48 [in Russian].
21. **Shary S. P.** Strong compatibility in data fitting problem under interval uncertainty / Vychisl. Tekhnol. 2017. Vol. 22. N 2. P. 150 – 172 [in Russian].
22. **Shary S. P.** Maximum compatibility method for data fitting under interval uncertainty / Journal of Computer and Systems Sciences International. 2017. Vol. 56. Issue 6. P. 897 – 913. DOI: 10.1134/S1064230717050100.
23. **Schweppe F. C.** Recursive state estimation: unknown but bounded errors and system inputs / IEEE Trans. on Automatic Control. 1968. AC-13. P. 22 – 28.
24. **Combettes P. L.** Foundations of set-theoretic estimation / Proc. IEEE. 1993. Vol. 81. N 2. P. 182 – 208.
25. **Milanese M., Norton J., Piet-Lahanier H., Walter E., eds.** Bounding Approaches to System Identification. — New York: Plenum Press, 1996. — 567 p. DOI: 10.1007/978-1-4757-9545-5.
26. **Lemeshko B. Yu., Postovalov S. P.** On solving problems of statistical analysis of interval observations / Vychisl. Tekhnol. 1997. Vol. 2. N 1. P. 28 – 36 [in Russian].
27. **Orlov A. I., Lutsenko E. V.** System Fuzzy Interval Mathematics. — Krasnodar: Izd. Kuban Gos. Agrar. Univ., 2014. — 600 p. [in Russian].
28. **Orlov A. I.** Statistics of interval data (generalizing paper) / Zavod. Lab. Diagn. Mater. 2015. Vol. 81. N 3. P. 61 – 69 [in Russian].
29. **Shary S. P.** An interval linear tolerance problem / Automation and Remote Control. 2004. Vol. 65. P. 1653 – 1666. DOI: 10.1023/B: AURC. 0000044274. 25098. da.
30. **Rohn J.** Inner solutions of linear interval systems / Interval Mathematics 1985 / K. Nickel, ed. Lecture Notes in Computer Science 212. — Berlin: Springer, 1986. P. 157 – 158.
31. **Schrijver A.** Theory of Linear and Integer Programming. — Chichester-New York: Wiley, 1998. — 471 p.
32. **Sharaya I. A.** Is the tolerable solution set bounded? / Vychisl. Tekhnol. 2004. Vol. 9. N 3. P. 108 – 112 [in Russian].
33. **Sharaya I. A.** On unbounded tolerable solution sets / Reliable Computing. 2005. Vol. 11. N 5. P. 425 – 432. DOI: 10.1007/s11155-005-0049-9.
34. **Remez E. Ya.** Principles of Numerical Methods of Chebyshev Approximation. — Kiev: Naukova Dumka, 1969. — 624 p. [in Russian].
35. **Shary S. P.** Interval regularization for imprecise linear algebraic equations. Deposited in arXiv.org Sept. 27, 2018, arXiv N 1810.01481. — 21 p. [in Russian].
36. **Demidenko E. Z.** Comment II on the article by A. P. Voshchinin, A. F. Bochkov, G. R. Sotirov “A method of data analysis under interval non-statistical error” / Zavod. Lab. 1990. Vol. 56. N 7. P. 83 – 84 [in Russian].
37. **Shor N. Z., Zhurbenko N. G.** A minimization method using the operation of extension of the space in the direction of the difference of two successive gradients / Cybernetics. 1971. Vol. 7. N 3. P. 450 – 459. DOI: 10.1007/BF01070454.
38. **Stetsyuk P. I.** Subgradient methods ralgb5 and ralgb4 for minimization of ravine-like convex functions / Vychisl. Tekhnol. 2017. Vol. 22. N 2. P. 127 – 149 [in Russian].
39. **Nurminski E. A.** Separating plane algorithms for convex optimization / Mathematical Programming. 1997. Vol. 76. P. 373 – 391. DOI: 10.1007/BF02614389.
40. **Vorontsova E. A.** Linear tolerance problem for input-output models with interval data / Vychisl. Tekhnol. 2017. Vol. 22. N 2. P. 67 – 84 [in Russian].