Заводская Лаборатория. – 2002. – Т. 68, № 1. – С. 118-126.

Интервальный анализ данных: развитие и перспективы

А.П. Вощинин

Введение

Для автора, как и для многих его коллег, чье профессиональное становление пришлось на середину 60-х годов, журнал «Заводская Лаборатория» (ЗЛ) сыграл исключительную роль в формировании и выборе научного направления. В то время секцию журнала «Математические методы эксперимента» возглавлял В.В. Налимов – пытливый и яркий ученый, интересы которого простирались от философии и математической статистики до химии и психологии. Благодаря его энергии и энтузиазму математическая статистика и ее сравнительно новый раздел «планирование эксперимента» получили в СССР мощный импульс и привлекли много приверженцев. Процесс становления молодых ученых и их приобщения к новому направлению проходил в творческой атмосфере открытых и доброжелательных дискуссий, царившей на многочисленных семинарах, региональных и Всесоюзных конференциях. Результатом деятельности сформировавшегося вокруг ЗЛ «незримого коллектива» ученых и практиков стали статьи и монографии по статистическим методам обработки данных и планированию эксперимента, методические рекомендации и пособия по применению этих методов в различных отраслях промышленности [1-6], кандидатские и докторские диссертации. Во многих технических Вузах стали читаться курсы по автоматизации и планированию эксперимента, включающие раздел статистических методов.

Направление «планирование эксперимента» стало модным и привлекло не только добросовестных исследователей, но и временных «попутчиков», чья цель состояла в скорейшей защите диссертации. Они увидели в планировании эксперимента удобный и простой инструмент, который позволял придать наукообразие несложным экспериментальным исследованиям. Типичные диссертации этого сорта, как правило, имели названия – «Применение планирования эксперимента при исследовании … (далее следовало название специфического объекта или промышленной установки)». Структура диссертации строилась по «типовой» схеме: описание объекта, математические основы планирования эксперимента (часто просто переписанные из книг В.В. Налимова и В.Г. Горского), описание плана и хода проведения эксперимента, регрессионный анализ модели с прилагаемым актом экономической эффекта.

В 70-е годы, возможно как реакция на профанацию идей математической статистики, появляется ряд статей, содержащих критику случаев некорректного применения статистических методов вообще и регрессионного анализа в частности, публикуются полемические материалы, в которых утверждаются, что исходные предпосылки регрессионного анализа далеки от реальности и редко выполняются на практике. В наиболее заостренной форме такие утверждения были суммированы в статьях [7, 8]. Автор статьи [8] утверждал, что математическую статистику нельзя считать наукой в строгом смысле этого слова, т.к. невозможно проверить на практике достоверность полученных с ее помощью результатов. В качестве примера «неверифицируемых» характеристик автор приводил понятие генеральной совокупности, доверительного интервала на неизвестное среднее случайной величины, ошибки первого и второго рода при проверке гипотез и др., которые невозможно проверить на реальных экспериментальных установках. Им ставилась под сомнение сама возможность воспроизвести статистический эксперимент в различных лабораториях. Эти тезисы, не всегда обоснованные и высказанные в эмоциональной форме, встретили в целом отрицательную реакцию.

Вместе с тем, практические приложения регрессионного анализа выявили ограничения слишком жесткой системы его исходных допущений. Замечания вкратце сводились к следующим тезисам:

- Гипотеза о нормальном распределении шума на практике выполняется далеко не всегда, ее проверка или не проводится вообще или выполняется на незначительных выборках. Неопределенность данных может иметь нестатистическую природу и включать систематическую составляющую, ошибки округления и группирования данных, методические погрешности. Даже если ошибки эксперимента случайные, они, как правило, действуют неаддитивно и меняются во времени, т.е. образуют нестационарный случайный процесс, что делает понятия генеральной совокупности и воспроизводимости несостоятельными.

- Разделение переменных на измеряемые точно «независимые» переменные и измеряемые с ошибками «зависимые» переменные для большого числа приложений выглядит искусственным и слишком жестким. В частности, эта гипотеза нарушается при идентификации динамических объектов, обработке фотоизображений, построении градуировочных кривых, сглаживании неявных функций и функций, заданных в полярных координатах.

Наиболее широкий класс практических задач, в которых нарушались классические постулаты регрессионного анализа, был вызван появлением в эти годы нового направления – «автоматизация технологических процессов». В его рамках экспериментальная модель (ЭМ) объекта была нужна не только и не столько для предсказания выходных значений переменных в режиме «совета» оператору, сколько для проверки гипотез, нахождения критических точек и областей, а также для решения задач оптимизации.
Общим моментом для перечисленных задач было или невыполнение предпосылок регрессионного анализа или отсутствие обоснованных алгоритмов их решения в рамках статистических методов.

Указанные обстоятельства привели к тому, что в 70-80-е годы в незримом коллективе, связанном с ЗЛ, начинает внедряться более широкая парадигма понимания природы неопределенности, которая формировалась в это время в различных областях прикладной математики. Она базировалась на трактовке неопределенности в более широком контексте, включающем не только случайность, но и незнание, неединственность возможных исходов, вариабельность переменных (не только случайную, но и систематическую), семантическую неопределенность целей, инструкций и оценок эксперта, многокритериальность при решении задач оптимизации. Новые подходы к описанию неопределенности вызвали появление концепции трехзначной логики («истина», «ложь», «не определено»), теорию нечетких множеств и нечетких чисел, метод программирования в ограничениях, концепцию недоопределённых моделей, теорию принятия решений в многокритериальных задачах [9, 10, 11] и, наконец, интервальный анализ, описанию которого и посвящена настоящая статья.

Развитие парадигмы интервального анализа

Интервал (отрезок), как объект действия, был впервые использован древними греками. В греческом языке слово «дихотомия» означает процесс последовательного деления отрезка на две равные части. Понятие совершенной пропорции (золотого сечения) в архитектуре также связано с делением отрезка длиной (a+b) на две неравные части в отношении: $0.5\cdot(\sqrt{5}-1)$. Совершенство и красоту золотого сечения, реализованного русскими зодчими, мы до сих пор можем видеть в пропорциях сохранившихся храмов, например храма Покрова на Нерли.

В контексте данной статьи отрезок [x] трактуется как множество возможных значений неизвестной истинной величины *x*, т.е. как ограниченный *интервал* ее *неопределённости*

$$[x] = \{ x \mid x^- \le x \le x^+ \},$$

задаваемый нижней и верхней границами x^-, x^+. Предполагается, что неизвестное истинное значение переменной достоверно лежит внутри интервала [x] и все значения внутри интервала считаются «равновозможными», т.е. на интервале не определяется никакой вероятностной меры. Понятие «равной возможности» не следует трактовать как равномерное распределение случайной величины на интервале, т.к. операции с равномерно распределенными величинами приводят к изменению распределения результата (сумма равномерных распределений стремится к нормальному распределению). Результатом же операций с интервалами всегда является интервал.

Основы интервального анализа были изначально заложены теорией измерений в *метрологии,* где интервал неопределенности вводится естественным образом. Предполагается, что имеется измеренное неточным прибором величины значение *x* неизвестной измеряемой величины x_0, и известна абсолютная ошибка измерения Δ (или относительная ошибка измерения δ). Тогда границы интервала неопределенности измеряемой величины x_0 определяются из условия

$$[x_0] = [x^-; x^+] = [x-\Delta; x+\Delta] = [x\cdot(1-\delta); x\cdot(1+\delta)] \qquad (1)$$

В метрологии используется понятие косвенного измерения $y = f(x_1, x_2, \ldots x_k)$, когда изучаемая величина *y* определена как известная функция от нескольких измеренных величин. Ошибку косвенного измерения Δy при известных ошибках измеренных величин Δx_i. обычно находят путем разложения функции в ряд Тейлора. Хотя этот подход до сих пор остается в арсенале метрологов, однако, он связан только с одной задачей, а именно, – определением интервального значения известной функции при интервальных значениях аргументов. Необходимость решения других прикладных задач с интервальными параметрами требовала новых подходов и методов. По этой причине появление в 70-80-е годы многочисленных работ по интервальной арифметике и интервальному анализу вызвало большой интерес специалистов. Развитие методов «интервального исчисления» проходило по двум направлениям.

На Западе (прежде всего в Германии) интервальный анализ развивался как средство автоматического учета ошибок округления при проведении численного решения задач на компьютерах. При этом результат выдается не в виде числа, а в виде интервала. Это направление на Западе получило название *интервальные вычисления* (reliable/validated/scientific computing). В его рамках решались две основные задачи: проверка точности существующих алгоритмов, анализ их интервальной сходимости и устойчивости и разработка новых алгоритмов для решения типовых задач, обеспечивающих минимальную ошибку интервального результата. Теоретической базой интервальных вычислений является *интервальная арифметика*. В «классической» интервальной арифметике [13] операции с положительными (не включающими нуля) интервалами определяются следующим образом:

$$[a]+[b] = [a^-+b^-;\ a^++b^+];$$
$$[a]-[b] = [a^--b^+;\ a^+-b^-]; \qquad (2)$$
$$[a][b] = [a^-b^-;\ a^+b^+];$$
$$[a]/[b] = [a^-/b^+;\ a^+/b^-].$$

Введение правил интервальной арифметики было связано с необходимостью автоматизации интервальных вычислений. Однако, при формализации операций с интервалами, которые являются более сложными объектами, чем числа, была потеряна техника символьных преобразований. В частности, в интервальной арифметике не выполняется дистрибутивный закон, т.е. $[a]+[b])\cdot[c] \neq [a]\cdot[c]+[b]\cdot[c]$, отсутствуют обратные элементы ($[a]/[a] \neq 1$, $[a]-[a] \neq 0$), интервальная арифметика оперирует только со значениями и «не видит» имен переменных, т.е. в ее рамках нельзя привести подобные члены. Эти ограничения не были существенными для рассматриваемого класса задач интервальных вычислений, где источником ошибок выступает ошибка округления, а ошибка результата, вызванная накоплением ошибок округления в ходе арифметических операций, принципиально зависит от выбранного вычислительного алгоритма.

Ученые СССР развивали *интервальный анализ* или интервальную математику как теоретическую основу для решения практических задач с неопределённостью в исходных данных и параметрах моделей [16-24]. При этом главная цель состояла не в автоматизации вычислений, а в нахождении области возможных значений результата с учетом структуры данных и функций, заданных в символьном виде. В этих условиях ограничения интервальной арифметики делали невозможным ее формальное использование. Направление интервального анализа активно развивалось (и развивается) в Сибирском отделении АН, в научных центрах и ВУЗах Москвы, Ленинграда и других городов. Были созданы научные школы, проводились семинары и конференции по данному направлению. В 1992 г. журнал ЗЛ инициировал дискуссию по применению интервального анализа данных, как альтернативу классическому регрессионному анализу при нарушении его исходных гипотез. Издательство Сибирского отделения АН СССР начало выпускать журнал «Вычислительные технологии». Все это способствовало публикации в СССР многочисленных статей и ряда книг по интервальному анализу.

Базовый *принцип интервального анализа* формулируется следующим образом: интервал неопределенности результата есть множество всех его возможных значений, получаемых при варьировании переменных и параметров задачи в границах известных интервалов. В частности, если дана функция (или функционал) $[y] = f([\mathbf{x}])$, интервального векторного аргумента $[\mathbf{x}] = ([x_1],\ldots[x_i],\ldots[x_k])$, то границы y^-, y^+ интервала неопределенности значения функции определяются как решение двух задач на экстремум:

$$y^- = \min_{x\in[X]} f(\mathbf{x}), \qquad y^+ = \max_{x\in[X]} f(\mathbf{x}). \qquad (3)$$

Таким образом, для определения границ необходимо найти наибольшее и наименьшее значение обычной функции векторного аргумента $y=f(\mathbf{x})$, представленной в символьном виде, когда ее аргументы меняются в заданных интервалах [X]. Результат арифметической операции, проведенной по формуле (3), зависит от того, имеем мы дело с одним интервалом или двумя. В частности, результат арифметических операций, проводимых с двумя интервалами [a] и [b] (даже имеющими одинаковые границы) в интервальном анализе совпадает с правилами интервальной арифметики (2). Однако, если рассматривается один интервал, то в интервальном анализе справедливо: $[a]-[a]=0$, $[a]\cdot[a]=[(a^-)^2;\ (a^+)^2]$, $[a]/[a]=1$. Следовательно, в интервальном анализе существуют обратные элементы и выполняется дистрибутивный закон.

Сравнение двух направлений показывает, что хотя они и базируются на интервальной форме описания неопределенности, однако решают разные задачи, базируются на разных гипотезах и используют различные подходы к определению результирующего интервала неопределенности. Однако, формальная близость «интервальной терминологии» обеих направлений, простота и привлекательность формальных операций интервальной арифметики привели к тому, что в ряде отечественных исследований методологию интервальных вычислений стали неправомерно применять к решению задач интервального анализа. При этом авторы (см. например [25]) искусственно вводили алгоритм вычислений, что приводило к путанице в постановке задачи и к дезориентации читателей. Возможно по этой причине, последователи интервальных вычислений, желая более четко позиционироваться, заменили название своего журнала *Interval Computations* на *Reliable computations*, в котором уже отсутствовало прилагательное «интервальный».

В дальнейшем в статье обсуждаются задачи только интервального анализа. Объяснения и анализ возможностей интервального анализа ведется на простых иллюстративных примерах, допускающих наряду с аналитическим и графическое решение.

Интервал неопределенности эмпирических позиномов

В технических, экологических и экономических приложениях наряду с объективными законами широко используются так называемые «эмпирические модели» или формулы, которые включают экспериментальные, неточно задаваемые переменные. Эмпирические модели являют типичный пример интервальных функций $[y] = f([x_1], [x_2], \dots[x_k])$. Для одного и того же явления часто существует несколько альтернативных эмпирических формул, отличающихся как структурой, так и числом переменных. При этом включение переменной в модель с одной стороны увеличивает ее полноту, т.е. снижает неопределенность, а с другой увеличивает общую ошибку модели, связанную с неточностью измерения переменных. Это дает основания говорить об оптимальной структуре модели, при которой достигается наименьшая степень общей неопределённости. Рассмотрим проблему выбора оптимальной структуры на примере так называемых *эмпирических позиномов.*

Простейший позином определяется как произведение степенных функций

$$y = x_1^{\beta_1} \cdot x_2^{\beta_2} \dots \cdot x_k^{\beta_k}, \qquad (4)$$

где переменные x_i и показатели степени β_i могут любые принимать неотрицательные значения. Предположим, что показатели степени определены точно, а значения переменных при фиксированных условиях заданы в интервальной форме $[x_i]$. Тогда, учитывая структуру позинома, легко получить, что границы интервала неопределенности $[y] = [y^-; y^+]$ определяются как

$$[y] = [y^-; y^+] = [\, y^- = (x_1^-)^{\beta_1} \cdot (x_2^-)^{\beta_2} \dots \cdot (x_л^-)^{\beta_k};\ y^+ = (x_1^+)^{\beta_1} \cdot (x_2^+)^{\beta_2} \dots \cdot (x_л^+)^{\beta_k}\,]$$

Если заданы точечные измеренные значения x_i и их относительные ошибки измерений δ_i, то границы результата *y* целесообразно определить через точечное значение выходной переменной $y = x_1^{\beta_1} \dots x_k^{\beta_k}$ и ее относительную ошибку δ_y. В общем случае значение δ_y выражается через относительные ошибки переменных следующим образом: $(1 \pm \delta_y) = (1 \pm \delta_1)^{\beta_1} \cdot \dots \cdot (1 \pm \delta_k)^{\beta_k}$. При небольших (меньше 0.3-0.5) значениях δ_i справедлива более простая приближенная формула $\delta_y \cong \sum_{i=1}^{k} \beta_i \delta_i$, которая позволяет записать границы интервала неопределенности в виде:

$$y^- = y \cdot (1 - \delta_y); \quad y^+ = y \cdot (1 + \delta_y). \qquad (5)$$

Допустим, что формула (4) включает все переменные x_i, влияющие на величину *y*. Тогда величину δ_y можно трактовать как относительную ширину интервала неопределенности и рассматривать ее как удобную численную характеристику неопределенности. Если неопределенность отсутствует, т.е. значение *y* определено точно, то $\delta_y = 0$. Если же $\delta_y \geq 1$, то имеется полная неопределенность, т.к. не определен даже знак результата

Рассмотрим более общий случай, когда в эмпирическую модель (4) включено только $m < k$ переменных, что приводит к появлению неточности из-за её упрощения. В этом случае (при игнорировании систематической составляющей) результирующая неопределенность результата определяется двумя слагаемыми: первое из которых связано с неполнотой модели, второе – с ошибками измерения переменных. Если предположить, что переменные модели проранжированы по убыванию влияния на переменную *y*, т.е. вклад переменной уменьшается с увеличением ее номера, то результирующую неопределенность можно представить в виде зависимости от числа m переменных, включенных в модель

$$\delta_y(m) = \delta_y(\text{неполноты}) + \delta_y(\text{измерений}) = e^{-\gamma\cdot m} + \sum_{i=1}^{m} \beta_i \delta_i \qquad (6)$$

Указанная функция имеет единственный минимум (см. рис.1).

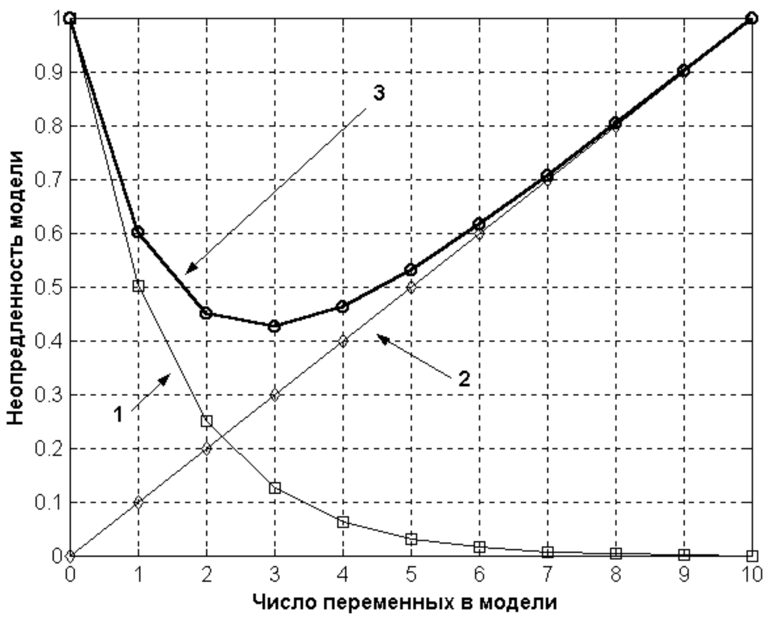

Рис. 1. Неопределенность эмпирической модели

На рис.1 представлены три кривых. Кривая 1 отражает изменение первого слагаемого в (6), связанного с неполнотой модели. Естественно, что при отсутствии переменных (m = 0) неопределенность модели максимальна, т.е. равна единице. При увеличении числа переменных структурная неопределенность снижается. В данном примере коэффициент $\gamma = 0.69$ в уравнении (6) выбран из предположения, что первая переменная (m=1) снижает неопределенность модели на 50%. Вклад остальных переменных убывает по экспоненте. Кривая 2 отражает изменение второго слагаемого в (6), связанного с увеличение суммарной ошибки измерения переменных. Она построена в предположении, что все степени позинома равны единице и все переменные заданы с одинаковой точностью, а именно: $\delta_i = \delta = 0.1$ (10%).

Кривая 3, равная сумме двух предыдущих кривых, для рассматриваемого случая достигает минимума в точке m = 3. Увеличение числа переменных сверх этого значения не снижает неопределенность модели, т.к. при этом начинает превалировать вклад ошибок измерения. Приведенный подход сравнительно легко распространяется и на более сложные эмпирические модели.

Определение корней интервального уравнения

Задача нахождения корней уравнений при интервально заданных коэффициентах часто возникает в многочисленных приложениях, в частности в задачах на экстремум, при нахождении критических точек и точек пересечения эмпирических моделей. Графическое решение двух простейших задач такого рода приведено на рис.2.

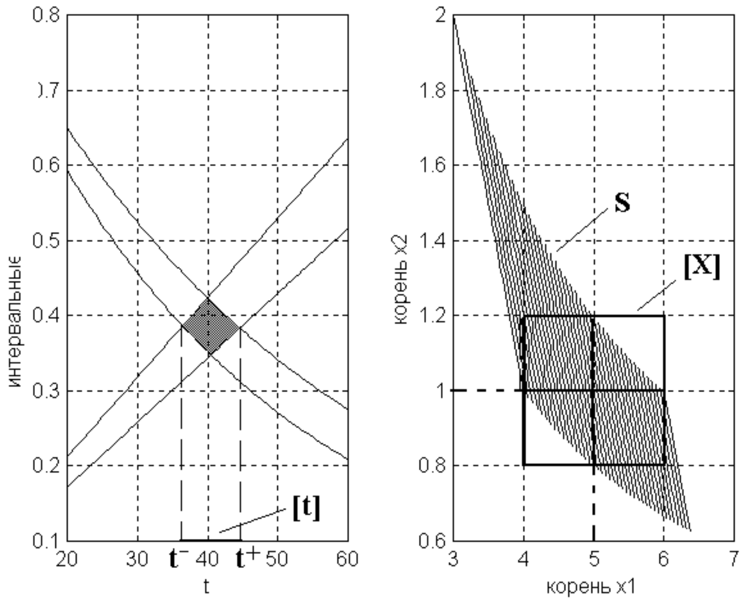

Рис.2

Рис 2а. Интервал неопределенности единственного корня;
Рис.2б. Область изменения двух корней.

На рис. 2а относится к случаю интервального уравнения $\exp^{-[k]\cdot t} = [\lambda]\cdot t$, которое используется при расчете длины опасной зоны загрязнения поверхности токсическим веществом. На рисунке изображены коридора левой и правой части уравнения, когда его параметры меняются в пределах: [k] = [0.0215, 0.0261], [λ] = [0.0086, 0.0106]. Интервал неопределенности корня уравнения $[t] = [t^-; t^+] = [36.44, 44.58]$ определен как проекция на ось t точек пересечения двух граничных кривых: $\exp^{-(k+)\cdot t} = \lambda^+ t$ и $\exp^{-(k-)\cdot t} = \lambda^- t$.

На рис. 2б показан результат решения квадратного уравнения $x^2 - [5; 7]\cdot x + [4; 6] = 0$. Множество получаемых пар корней уравнения при изменении коэффициентов уравнения внутри заданных пределов образует область S, вытянутая форма которой определяется действием теоремы Виета. Для замены нелинейной области S двумя интервальными корнями можно использовать внешний прямоугольник $[X]_{out} = \{[3; 6.37]; [0.63; 2]\}$, описанный вокруг области S. Он является довольно грубой аппроксимацией, т.к. содержит более 50% комбинаций корней, не принадлежащих области S.

Для нахождения более приемлемого интервального приближения можно использовать *метод фиксированного аргумента* [22], идея которого состоит в следующем. Центр искомого прямоугольника [X] помещается в точку (x_1 = 5, x_2 = 1), найденную по средним интервальным значениям коэффициентов уравнения. По теореме Виета корни уравнения связаны отношениями, которые должны выполняться при любых значения коэффициентов. Это позволяет записать интервальную систему двух уравнений:

$$x_1+x_2 = -[b_1] = [5; 7], \quad x_1 \cdot x_2 = [b_0] = [4; 6].$$

Подставляя в эту систему среднее значение корня x_1 = 5, получаем два интервальных уравнения относительно корня x_2: 5+ [x_2] = [5; 7] и 5·[x_2] = [4; 6], которые легко решаются относительно неизвестного интервала [x_2]. Беря в качестве решения системы наиболее «узкое» решение, получаем [x_2] = [0.8; 1.2]. Проделав аналогичные действия для фиксированного корня x_2=1, имеем интервал [x_1] = [4; 6]. Результирующая прямоугольная область [X] == { [4; 6]; [0.8; 1.2] } показана на рис.2б. Она дает лучшее приближение области S, чем внешняя аппроксимация.

Метод может быть распространен на случай определения интервального значения корней полинома n-ой степени, если использовать многомерную теорему Виета. Эта задача является важной при анализе линейных динамических систем с неточно заданной моделью.

Расчет градуировочной кривой по интервальным данным

В теории измерений важное место занимает задача построения градуировочной характеристики измерительного прибора по неточным данным эксперимента. Рассмотрим формулировку задачу для простейшего линейного случая. Пусть функция преобразования датчика (*прямая модель*) имеет вид линейной функции

$$y = b_1 + b_2 \cdot x$$

где x- измеряемая величина, y- измеренное с помощью датчика ее значение. Цель состоит в том, чтобы по данным эксперимента определить *обратную модель* $x = a_1+a_2 \cdot y$, называемую градуировочной характеристикой и оценить коридор ее ошибок. В рамках статистической парадигмы эта задача относится к классу наиболее сложных и слабо формализуемых. Ее решению посвящен отдельный раздел статистики – теория калибровки [25]. Трудности вызываются следующими причинами. Если по данным эксперимента, найдены несмещенные статистические оценки коэффициентов b_i прямой модели, то при переходе к обратной модели это свойство теряется, кроме того, возникают теоретические сложности при построении доверительного интервала. Если же рассчитываются оценки коэффициентов обратной модели, то зависимой становится переменная y, измеряемая с ошибками, что приводит к нарушению предпосылок регрессионного анализа.

В рамках интервального анализа задача решается достаточно просто даже при наличии ошибок в обеих переменных x и y. Покажем это на простом примере реальных данных построения градуировочной характеристики датчика для измерения вязкости [23]. Был проведен активный эксперимент на трех уровнях значения вязкости (0; 1; 2) с повторением каждого опыта 12 раз. Диапазон изменения переменной y был определен по результатам параллельных опытов, ошибка установки уровней вязкости x была принята равной Δ=0.05. Исходные экспериментальные данные и результаты их обработки показаны на рис. 3.

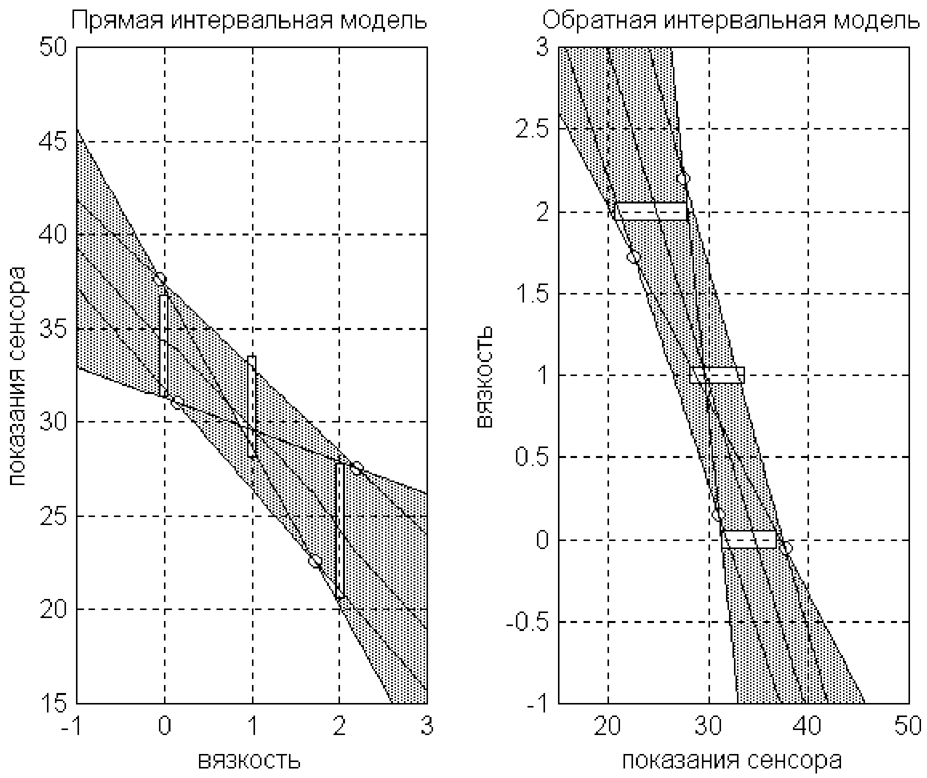

Рис.3 Коридоры неопределенности прямой (3а) и обратной моделей (3б)

Опытам в эксперименте на рис. 3 соответствуют прямоугольные области неопределенности, Границы коридора ошибок прямой модели (рис. 3а) имеют вид кусочно-линейных функций, образованных четырьмя экстремальными прямыми. Прямые участки, образующие коридор, построены по конечному числу экстремальных наборов точек. В частности, в диапазоне изменения вязкости $0 \le x \le 2$, границы интервала неопределенности прямой модели построены по верхним и нижним вершинам прямоугольников. Центральная прямая построена как полусумма границ коридора ошибок. Границы средней части коридора ошибок на рис. 3а образуют прямую интервальную модель

$$[y] = [31.81; 37.39] + [-5.375; -4.475]\cdot x \quad \text{при } 0 \le x \le 2. \tag{7}$$

Учитывая взаимно однозначное соответствие между переменными, на основе модели (7) легко получить обратную интервальную модель

$$[x] = -[b_1]/[b_2] + -1/[b_2]\cdot y = [5.917; 8.357] + [-0.223; -0.186]\cdot y, \tag{8}$$

которая и является искомой градуировочной характеристикой. Из рис. 3б видно, что модель (8) имеет приемлемую точность в сравнительно узком диапазоне изменения y, а именно $29 \le y \le 30.6$. За его пределами, где коридор ошибок измерительного прибора задается сочетанием других границ, неточность модели резко возрастает. Этот факт нельзя считать методической слабостью интервального анализа, т.к. он отражает объективную особенность коридора ошибок обратной градуировочной характеристики.

При статистическом подходе к калибровке сначала рассчитывают точечные оценки прямой модели, а затем получают обратную модель и ее доверительный коридор ошибок. Это приводит к градуировочной характеристике, несимметричной относительно границ, т.е к смещенным оценкам обратной модели. В интервальном анализе этого не происходит, т.к. вначале строятся границы коридора и находятся обратные от них, а затем определяется градуировочная характеристика как полусумма найденных границ. Это обеспечивает ее симметричность относительно границ интервала неопределенности. Преимущество интервального анализа также в том, что в отличие от статистического подхода, оба варианта расчета - а) построение прямой модели и ее обращение; б) непосредственное построение обратной модели, приводят к одним и тем же результатам.

Проверка гипотез

В приложениях часто необходимо проверять различного рода гипотезы относительно параметров эмпирических моделей. В частности, линейная динамическая система является устойчивой, если корни ее характеристического полинома имеют отрицательные действительные части. Если корни заданы в интервальной форме, то необходимо иметь процедуру проверки гипотез и определить понятие достоверности гипотезы. Рассмотрим два случая проверки гипотез.

1. Пусть задан *интервал* неопределенности [x] неизвестной величины *x и известная константа d*. Необходимо проверить гипотезу H_0: $x < d$, против альтернативной гипотезы H_1: $x > d$, где x неизвестное истинное значение. Результат проверки и его достоверность будет зависеть от положения константы *d* относительно границ интервала [x]. Очевидно, что если $d > x^+$, то гипотеза H_0 принимается безусловно, т.е. с достоверностью единица. Аналогично, если $d < x^-$, то с достоверностью единица принимается альтернативная гипотеза H_1.

Когда константа *d* попадает внутрь интервала, то возможны обе ситуации: $x<d$, $x>d$. (Возможность строгого равенства $x=d$, как и теории вероятностей, равна нулю). Определим *уровень достоверности заданного отношения* α как отношение длины отрезка, на котором отношение выполняется, к длине всего интервала (x^+-x^-). (При генерации конечного числа точек интервале [x] это соответствует отношению числу случаев выполнения отношения к общему числу точек). Таким образом, общее правило вычисления достоверности гипотез записывается в виде:

- Если $d \geq x^+$, то гипотеза H_0: $x<d$ имеет достоверность $\alpha_{x<d} = 1$,
  а гипотеза H_1: $x > d$ достоверность $\alpha_{x>d} = 0$.

- Если константа *d* лежит внутри интервала [x], то уровни
  достоверности гипотез H_0: $x < d$, H_1: $x > d$ связаны отношениями

$$\alpha_{x<d} = \frac{d - x^-}{x^+ - x^-}, \quad \alpha_{x>d} = 1\text{-}\alpha_{x<d}. \qquad (9)$$

Гипотезу целесообразно принимать, только если уровень ее достоверности превосходит заранее заданную, достаточно большую величину α_0, например α_0=0.95. Введенное определение достоверности близко к понятию уровня значимости при проверке статистических гипотез, хотя отличается от него как по способу вычисления, так и по смыслу.

Пример. Определить достоверность гипотезы H_0:x<0 при интервально заданном коэффициенте [x] = [-1; 1]. По формуле (3), полагая d = 0, получаем $\alpha_{x<0} = \alpha_{x>0} = 0.5$, т.е. в этом случае как нулевая, так и альтернативная гипотеза имеют одинаковую достоверность. Как трактовать этот результат и какое решение принять? Возможны два варианта: 1) отвергнуть нулевую гипотезу, как имеющую малую достоверность и «закрыть вопрос»; 2) Учитывая, что достоверность нулевой и альтернативной гипотез совпадает, трактовать этот факт как полную неопределенность относительно знака коэффициента, и принять гипотезу x=0. Подчеркнем, что такое решение, хотя и является достаточно разумным, является «волевым актом» эксперта, т.к. формально достоверность выполнения условия $x=d$ равна нулю.

2. Пусть даны *два интервала* [x] и [y]. Необходимо проверить гипотезу H_0:x<y против H_1:x>y. Если интервалы не пересекаются и интервал [x] лежит левее интервала [y], то очевидно, что гипотеза H_0:x<y выполняется с уровнем $\alpha_{x<y} = 1$. Если интервалы пересекаются, (но не вложены друг в друга), то отношение между ними можно задать на плоскости (рис. 4).

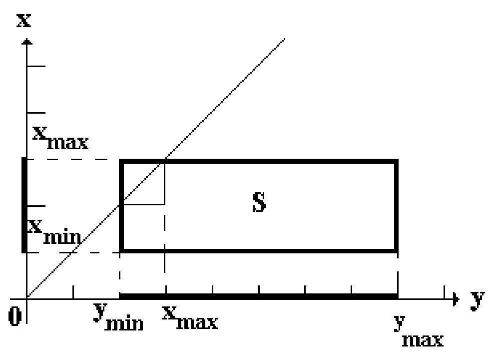

Рис.4. Отношение между интервалами

На рис. 4 наклонная линия, проведенная под 45 градусов, задает множество точек, где выполняется отношение x = y. Отмеченный жирными линиями прямоугольник определяет множество всех возможных комбинаций пар $\{x_i\ y_i\}$ внутри интервалов [x] и [y]. Для всех точек слева от наклонной линии выполняется отношение x > y, для точек справа от линии справедливо x < y. Уровень достоверности соответствующего отношения определяется как отношение площади «отсеченной» наклонной линией к площади всего прямоугольника. В частности, если $x^- < y^-$ и $x^+ < y^+$ для достоверность гипотез определяется соотношениями:

$$\alpha_{x<y} == 1 - \frac{0.5\cdot(x^+ - y^-)^2}{(y^+ - y^-)\cdot(x^+ - x^-)}, \quad \alpha_{x>y} == \frac{0.5\cdot(x^+ - y^-)^2}{(y^+ - y^-)\cdot(x^+ - x^-)}. \tag{10}$$

Пример. Для уровня достоверности $\alpha_0 = 0.95$ проверить гипотезу H_0:x<y, если заданы интервалы неопределенности x = [1; 3], y = [2; 8]. Используя формулу (10), получаем $\alpha_{x<y}$ = 1-0.042 = 0.958. Так как $\alpha_{x<y}$ > 0.95, то гипотеза принимается.
Возникает вопрос, нельзя ли при наличии двух интервалов [x] и [y] вместо гипотезы H_0: $x < y$ проверять гипотезу H_0: x-y < 0, используя правило (9), а не (10)? Ответ – отрицательный, т.к. при вычислении достоверности гипотезы случай двух неопределенных величин x и y нельзя свести к более простой ситуации одной неопределенной величины (x-y). Этот же факт имеет место и при проверке статистических гипотез.

Идентификация конических сечений по интервальным данным

Приложение интервального анализа к задачам сглаживания данных снимает многие проблемы, возникающие при использовании статистических методов. В частности, становится излишней концепция «черного ящика», основанная на разделении переменных на зависимые и независимые. Модель может быть задана в неявной форме или в полярных координатах и зависеть от многих переменных, каждая из которых определена с погрешностью.

Проиллюстрируем это на примере определение вида конического сечения построенного по интервальным данным. Известно, что общее уравнение конических сечений задается в виде

$$a_{11}x^2 + a_{22}y^2 + 2a_{12}yx + 2a_{13}x + 2a_{23}y + a_{33} = 0. \tag{11}$$

При известных коэффициентах a_{ij} уравнения (11) существуют три инварианты, определяющие свойства кривой второго порядка и не зависящие от ее положения на плоскости

$$I = a_{11}+a_{22}, \quad D = a_{11}a_{22}-(a_{12})^2, \quad A = \det C, \tag{12}$$

где С - (3×3) симметричная матрица с элементами a_{ij}. Если $D > 0$ и $(A/I) < 0$, то уравнение описывает эллипс или круг, при $D < 0$ гиперболу, а при $D = 0$ параболу.

Пусть имеются экспериментальные «интервальные точки» ($[x_i]$, $[y_i]$), принадлежащие кривой конического сечения, т.е. исходные данные эксперимента представляются в виде интервальных векторов [X] и [Y], которые на плоскости формируют прямоугольные области неопределенности (рис.5).

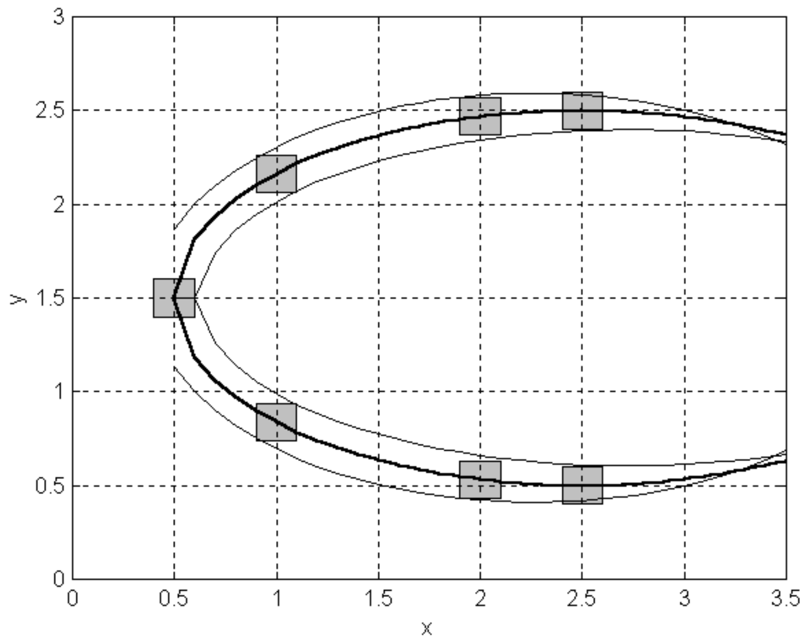

Рис.5 Исходные данные и интервальная модель конического сечения

Так как в рамках интервального анализа разделение переменных на входные и выходные является условным, в качестве «выходной» переменной в коническом сечении может быть выбрана любая базисная функция в уравнении (11).

Для рассматриваемого ниже примера было выбрано следующее представление:

$$y^2 = b_1 + b_2 \cdot y + b_3 \cdot x + b_4 \cdot x^2 + b_5 \cdot x\, y \tag{13}$$

Для нахождения неизвестных коэффициентов модели (13) необходимо рассчитать интервальный вектор $[\mathbf{Z}] = \{[y_i]^2\}$ и интервальную матрицу базисных функций правой части уравнения $[\mathbf{F}] = \{1, [y_i], [x_i], [x_i]^2, ([yi]\cdot[xi])\}$. При этом задача интервального анализа данных сводится к решению прямоугольной интервальной системы уравнений

$$[\mathbf{F}]\cdot\mathbf{B} = [\mathbf{Z}], \tag{14}$$

относительно вектора коэффициентов **B.** Решение системы имеет вид выпуклого многогранника S, который содержит все возможные значения коэффициентов. Для перехода к интервальным коэффициентам, как и в случае на рис. 2б, необходимо область S аппроксимировать прямоугольным «брусом», который задает искомый интервальный вектор **[B]**. Для его построения может быть использован метод фиксированного параметра [22]. Однако в данном случае проще построить границы коридора ошибок модели, определив два экстремальных конических сечения, проходящих по соответствующим вершинам прямоугольников. Указанные границы, показанные на рис.5, образуют следующую интервальную модель:

$$y^2 = [-0.344; -0.18]\, x^2 + [0.99; 1.54]\cdot x + [2.99; 3.002]\cdot y + [-2.81; -2.79] \qquad (15)$$

Средняя линия на рис.5 проведена методом наименьших квадратов через средние точки интервальных наблюдений.

Для определения типа конического сечения необходимо привести интервальную модель (15) к каноническому виду (11) и вычислить интервальные значения инвариант (12). При этом, применяя принцип (3), можно получить следующие результаты:

$$[I] = [-1.344; -0.18], \quad [D] = [0.18; 0.344], \quad [A/I] = [-0.2416; -0.0416].$$

Используя приведенные выше правила проверки гипотез, получаем, что обе гипотезы $H{:}D > 0$ и $H{:}(A/I) < 0$ выполняются с уровнем достоверности $\alpha = 1$. Это означает, что в рассмотренном случае коническое сечение является эллипсом.

Заключение

Приведенные иллюстративные примеры показывают, что применение интервального анализа позволяет снять многие проблемы и методические сложности, возникающие при решении прикладных задач статистическими методами. В рамках интервального анализа неопределенность исходных данных может иметь разные источники и природу. Интервал неопределенности позволяет описать широкий класс неопределенных, неоднозначных, вариабельных и неточных исходных данных. Значения ошибок в исходных данных могут колебаться в широких пределах. Результаты, полученные с помощью в рамках парадигмы интервального анализа, имеют ясную и четкую интерпретацию в терминах интервалов и областей неопределенности. Основной проблемой интервального анализа является корректное определение интервалов неопределенности на основе различной исходных данных при наличии различных источников неопределенности переменной. В этом плане представляются перспективными работы [20,27], в которых внутри интервала неопределенности выделяют как систематическую, так и случайную составляющие. Приведенные результаты дают основание сделать заключить, что применение интервального анализа к широкому спектру прикладных задач с ограниченными ошибками и неопределенностью в данных имеет больших перспективы.

Литература

1. Налимов В.В. Чернова Н.А., Статистические методы планирования экстремальных экспериментов, М.: Наука, 1965, 340 с.
2. Налимов В.В. Чернова Н.А., Теория эксперимента, М.: Наука, 1971.
3. Горский В.Г., Адлер Ю.П., Талалай А.М., Планирование промышленного эксперимента . Модели статики, М.: Изд-во Металлургия, 1974
4. Демиденко Е.З. Линейная и нелинейная регрессия, м.: Финансы и статистика, 1981..
5. Горский В.Г., Планирование кинетических экспериментов, М.: Наука, 1984.
6. Горский В.Г., Современные статистические методы обработки и планирования экспериментов в химической технологии, Труды Международной школы «Инженерно-химическая наука для передовых технологий», Казань, 1997.
7. Тутубалин В.Н. Границы применимости (Вероятностно-статистические методы и их возможности), М.:, Знание, 1977.
8. Алимов Ю.И., Альтернатива методу математической статистики, М.:, Знание, 1980.

9. Заде Л.А. Понятие лингвистической переменной и его применение к понятию приближенного решения, М.: Мир, 1976
10. Нариньяни А.С., Недоопредленность в системе представления и обработки знаний, Изв. АН СССР, “Техническая Кибернетика”, №5, 1986.
11. Hyvonen E. Constraint reasoning based on interval Arithmetic: the tolerance propagation approach, Artificial Intelligence, v.58, 1992.
12. Алефельд Г., Херцберг Ю. Введения в интервальные вычисления.- М.: Мир, 1987, 370 с.
13. Moore R. E. , Interval Analysis, Englewood, Cliffs, N.Y., Prentice-Hall, 1966.
14. Шокин Ю.И. Интервальный анализ, Новосибирск, Наука, 1981, 281 с.
15. Калмыков С.А., Шокин Ю.И., Юлдашев З.Х. Методы интервального анализа, Новосибирск: Наука, 1986.
16. Канторович Л.В. О некоторых новых подходах к вычислительным методам и обработке наблюдений, Сибирский математический журнал, т.3, №5, 1962.
17. Вощинин А.П., Сотиров Г.Р. Оптимизация в условиях неопределенности, Изд. МЭИ-СССР, Техника, НРБ, 1989.
18. Вощинин А.П., Бочков А.Ф., Сотиров Г.Р. Интервальный анализ данных как альтернатива регрессионному анализу, ЗЛ, 1990, No 7, 76-81 с.
19. Куржанский А.Б. Задача идентификации – теория гарантированных оценок, Автоматика и Телемеханика, №4, 1991.
20. Design of Experiments and Data Analysis: New Trends and Results, ANTAL, Moscow, 1993.
21. Gorsky V., Shvetzova-Shilovskaya T., Voschnin A. Risk assessment of accident involving environmental high-toxicity substances, Journal of Hazardous Materials, No78, 2000.
22. Вощинин А.П. Метод анализа данных с интервальными ошибками в задачах проверки гипотез и оценивания параметров неявных линейно параметризованных функций, Заводская лаборатория, том 66, №3, 2000.
23. Вощинин А.П., Скибицкий Н.В. Интервальный метод калибровки, Журнал "Системы и датчики, №7, 2000.
24. Шарый С.П. Интервальные алгебраические задачи и их численное решение, Диссертация на соискание ученой степени доктора физико-математических наук, 2000, 322 с.
25. Проблемы обеспечения безопасности и эксплуатационной надежности химических производств, Итоги науки и техники, серия "Процессы и аппараты химической технологии", том 19, под ред. Кафарова В.В., ВИНИТИ, Москва., 1973 186 с.
26. Семенов Л.А., Сирая Т.Н. Методы построения градуировочных характеристик средств измерения, М.: Издательство стандартов, 1986, 127 с.
27. Орлов А.И., Асимптотический линейный регрессионный анализ для интервальных данных, Заводская лаборатория, в печати.