методов, для достижения поставленных требований по точности измерения. Описанный подход, позволяет произвести моделирование и расчет параметров измерительного канала, использующего в качестве чувствительного элемента NTC терморезистор, при этом каждая из дискретных моделей имеет собственные параметры и характеристики, может быть скорректирована независимо от других.

На данном этапе разработаны дискретные параметрические модели NTC терморезистора, и измерительной цепи построенной на основе терморезисторного делителя напряжения, а так же методы расчета параметров численных методов для аппроксимации температурной зависимости, что уже позволило автоматизировать процесс определения параметров элементов измерительной цепи и расчет коэффициентов аппроксимирующего полинома.

Результаты исследований изложенные в данной статье получены при финансовой поддержке Минобрнауки РФ в рамках реализации проекта "Создание высокотехнологичного производства по изготовлению информационно-телекоммуникационных комплексов спутниковой навигации ГЛОНАСС/GPS/ Galileo" по постановлению правительства №218 от 09.04.2010.

СПИСОК ЛИТЕРАТУРЫ

1. Беляев, А.О. Интегральные методы анализа элементов измерительного канала температуры на базе NTC терморезисторов// А.О. Беляев, Известия ЮФУ. Технические науки. Тематический выпуск "Компьютерные технологии в науке, инженерии и управлении" - Таганрог: Изд-во ТТИ ЮФУ, 2011. № 5. - С. 106-109

Беляев А.О.*, мл. науч. сотрудник, тел.: (8634) 311-143, e-mail: alexys@pisem.net, Научно-технический центр "Техноцентр" Южного федерального университета*

УДК 531.7

РЕГРЕССИОННАЯ МОДЕЛЬ НА ДАННЫХ С ИНТЕРВАЛЬНОЙ НЕОПРЕДЕЛЕННОСТЬЮ

Н.П.Ординарцева

Рассмотрены вопросы моделирования в предметной области, связанной с измерениями. Отмечена специфика результатов измерений как данных нечисловой природы с интервальной неопределённостью без вероятностной меры. Предложенный метод регрессионного моделирования на данных с интервальной неопределённостью актуален в практических задачах нахождения функциональных зависимостей, в частности, при построении градуировочных кривых средств измерений

Ключевые слова: интервальный анализ, регрессионная модель, данные с интервальной неопределенностью, градуировочные кривые

Актуальность постановки вопроса и его состояние

Во многих практических задачах возникает потребность построения функциональных зависимостей $Y = f(X)$ по экспериментальным данным $\{X_i\}$. Результат измерения физической величины (ФВ) представляет собой отображение реального физического свойства *X* на числовую ось *Y*.«*Измерение – гомоморфное отображение некоторой эмпирической системы с отношениями* ξ *на числовую систему с отношениями N, т.е.*

$$\xi = \left[\xi, R_\xi\right] \rightarrow N = \left[N, R_N\right],$$

где $\xi = \left[\xi, R_\xi\right]$ *- эмпирическая система с отношениями (*ξ *- множество эмпирических объектов;* R_ξ *- множество эмпирических отношений);* $N = \left[N, R_N\right]$ *- числовая система отношениями (N-множество числовых объектов,* R_N *- множество отношений)*» [1]. Отображение $\left[\xi, R_\xi\right] \rightarrow \left[N, R_N\right]$ является обязательной моделью любой измерительной процедуры. Любое измерение представляет собой отображение наблюдаемого фрагмента действительности в модельном метрологическом пространстве.

В большинстве работ, затрагивающих проблемы построения аксиоматических основ теории измерений то обстоятельство, что результаты измерений представляются действительными числами, служит основанием для приписывания этим результатам свойств действительных чисел. Однако результат измерений представляет собой величину как результат преобразования формы представления информации из аналоговой в числовую. Из последнего следует, что результатам измерений присущи те же свойства, что и объектам измерений (величинам), но не свойства действительных чисел как таковых, - результаты измерений являются данными ***нечисловой природы.***

При измерении мы имеем $\{X\}\xrightarrow{h}\{Y\}$, где *h* - функционал гомоморфного отображения $\{X\}$ на $\{Y\}$, причём существует обратное отображение h^{-1} множества $\{Y\}$ на множество $\{X\}$. Так как отсутствует требование равной мощности множеств *X* и *Y* (мощность множества – число элементов множества), то могут существовать элементы множества *Y*, в которые не отображается ни один из элементов множества *X*; отображение *Y* на *X* $\left\{Y \xrightarrow{h^{-1}} X\right\}$ изоморфно. Изоморфизм определяет характеристики качества измерений: «погрешность измерений» и «неопределённость измерений». Оба подхода в описании результата измерения не противоречат друг другу и совпадают по целям измерений, исходной информации, используемой для их определения, методам определения и, в конечном счёте, по количественным результатам определения характеристик погрешности и неопределенности; эти подходы не имеют различия в понимании вероятностного характера измерения, их нельзя противопоставлять: они имеют много общих элементов и дополняют друг друга. Различие между традиционным подходом, использующим понятие «погрешность измерений», и подходом на основе «неопределённости измерения», сводится к различию систем координат, относительно которых рассматривают значение измеряемой величины и результат измерений [2].

При рассмотрении погрешности измерений систему координат отсчёта совмещают со значением измеряемой величины *X*, наблюдая рассеяние результата измерений [3]

$$\Delta X = X_{изм} - \underset{\substack{\uparrow \\ X_{дейс}}}{X_0},$$

где X_0 и $X_{дейс}$ - соответственно истинное и действительное (условно принятое истинное) значение измеряемой ФВ. При подходе к оценке качества измерения на основе погрешности последняя является апостериорной оценкой достоверности приписанного значения, обуславливая неопределённость как присущую составляющую результата измерения; при рассмотрении неопределённости измерений – с результатом измерений *Y*, что и создаёт эффект рассеяния («интервал охвата») единственного значения измеряемой величины *X*.

В этой связи принципиально важно отметить то обстоятельство, что в ходе регрессионного эксперимента экспериментатор, снимая показания с измерительного прибора, может иметь только статистику $f(Y)$, но не $f(X)$. Очевидно, что в корректном решении поставленной регрессионной задачи необходимо отразить различие между $f(Y)$ и $f(X)$.

В результате измерения имеем измеренное значение ФВ X_j (перейдём к привычным буквенным обозначениям) и погрешность измерения неидеальным прибором Δ или δ (в абсолютной или относительной форме выражения). Последнее обстоятельство трактует результат измерения ФВ как ограниченный ***интервал её неопределённости***, задаваемый нижней и верхней границами X^-, X^+

$$[X] = \lfloor X : X^- \le X \le X^+ \rfloor = \\ = [X - \Delta; X + \Delta] = [X(1-\delta); X(1+\delta)],$$

при этом предполагается, что неизвестное истинное значение ФВ достоверно лежит внутри интервала $[X]$ и все значения внутри интервала считаются равновозможными, т.е. на интервале не определяется никакой вероятностной меры: нельзя сделать никакого предположения о вероятностном законе распределения случайной величины *X* внутри этого интервала равновозможных значений. Понятие «равной возможности» не следует трактовать как равномерное распределение случайной величины на интервале, так как операции с равномерно распределёнными величинами приводят к изменению распределения результата (сумма равномерных распределений стремится к нормальному распределению).

Таким образом, специфика результатов измерений состоит в том, что результаты измерений представляют ***данные нечисловой природы с интервальной неопределённостью без вероятностной меры***.

Задача систематизированного представления об операциях с данными нечисловой природы с интервальной неопределённостью без вероятностной меры в настоящее время является нерешённой, несмотря на широкое проявление этого вопроса во многих научно-исследовательских и производственных задачах. С учётом модели результата любого измерения как $\xi = \lfloor \xi, R_\xi \rfloor \to N = [N, R_N]$ отображения эмпирической системы с отношениями ξ на числовую систему с отношениями *N*, порождающими то обстоятельство, что результаты измерительного эксперимента представляют собой $f(Y)$, а не $f(X)$, исследование вопроса решения регрессионных задач для указанного типа данных нечисловой природы является ***актуальным***.

Критика некорректного применения исходных предпосылок регрессионного анализа в измерительном эксперименте

В практических вопросах восстановления данных по данным измерительного эксперимента имеется широкий класс задач, в которых нарушаются классические постулаты регрессионного анализа, основанные на гипотезе нормального распределения числовых данных. В числе «неверифицируемых» понятий исходных допущений классической теории можно привести такие, как [4]:

- генеральная совокупность;
- доверительный интервал на неизвестное среднее случайной величины;
- ошибки первого и второго рода при проверке гипотез.

Гипотеза о нормальном распределении шума на практике выполняется далеко не всегда, её проверка или не проводится вообще, или выполняется на незначительных выборках. Неопределённость данных может 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 жёстким. В частности, эта гипотеза нарушается при построении градировочных кривых средств измерений (СИ).

Для опознаваемых систематических погрешностей вводят поправки, а границы случайных погрешностей, выраженные через СКО и принятую доверительную вероятность, характеризуют рассеивание возможных приписанных значений величины. Анализ погрешностей содержит как статистические, так и нестатистические процедуры, что приводит к противоречию и математической некорректности в этом анализе. И пока не найден удовлетворительный способ идентификации, а тем более определения $\Delta_{\lim}$ и введения поправок для всей погрешности измерения ФВ. Не существует общепринятых средств комбинирования систематических и случайных погрешностей в одну «полную» погрешность, которая дала бы некоторое общее представление о том, насколько хорошо результат измерения соответствует истинному значению измеряемой ФВ [5].

Подход на основе интервального анализа

Существующие метрологические стандарты и руководства по представлению измерительных данных детально разработаны только для прямых измерений одной величины. Ошибку косвенного измерения Δy при известных ошибках измеренных величин Δx_i обычно находят путём разложения функции y в ряд Тейлора. Этот подход связан только с одной задачей – определением интервального значения известной функции y при интервальных значениях аргументов x_i. В интервальной математике подобная задача решаема с помощью производной Фреше [6].

Базовый принцип интервального анализа формулируется следующим образом: интервал неопределённости результата есть множество его возможных значений, получаемых при варьировании переменных и параметров задачи в границах известных интервалов. Результатом операций с интервалами всегда является интервал. Необходимость решения многих прикладных задач с интервальными параметрами требует новых подходов и методов, и в задачах измерений интервал неопределённости вводится самым что ни наесть естественным образом.

Конечно, предлагаемый подход на основе интервальной модели описания неопределённости измерительных данных нечисловой природы требует и предполагает много научной и практической полемики, и только затем нормативного закрепления в метрологии найденного оптимального решения, что полностью соответствует принципу стандартизации о постоянстве действия существующих нормативных документов и необходимости их периодической обновляемости [3].

Гибридная регрессионная модель, построенная на интервале $\left[X^{теор}, X^{эксп}\right]$ [7]

Задача восстановления функциональной зависимости $Y = f\left(X_1, X_2, ..., X_n\right)$, связывающей ожидаемую величину отклика *Y* с объясняющими переменными $X_1, X_2, ..., X_n$, является регрессионной задачей. С математической точки зрения любая регрессионная задача является ***некорректной, обратной и плохо обусловленной задачей***. Некорректность регрессионной задачи проявляется в том, что восстановление искомой зависимости порождения данных не имеет единственного решения.

То обстоятельство, что в поставленной задаче экспериментальные точки определяют саму эмпирическую зависимость, позволяет говорить о регрессионной задаче только как ***об обратно поставленной***. Кроме того, задачи регрессионного анализа, предполагающие восстановление зависимостей порождения данных - в силу неудачной традиции о нормальности распределения лежащих в её основе случайных величин (не встречающихся в реальной практике – а это не вполне адекватность постановки задачи), - являются ***плохо обусловленными задачами***. Известно, что важнейшее из требований к функции регрессии – ***непрерывность и достаточная «гладкость»***, предполагается в задачах метрологической и измерительной практики.

Изучение любого процесса начинается с построения его модели. Первый этап этого построения, как правило, не позволяет учесть всё многообразие влияющих факторов. Предлагаемый метод гибридного регрессионного анализа позволяет улучшить искомую модель Y и сделать последнюю более обоснованной и адекватной. Модель строится путём использования гибридных данных априорной теоретической модели и результатов планируемого физического эксперимента. Суть объединённого анализа теоретических и эмпирических данных в том, что в силу их некоторой избыточности имеется возможность уточнения исходных предпосылок. Например, теоретическая модель

$$Y_{теор} = f\left(X^{теор}{}_1, X^{теор}{}_2, ..., X^{теор}{}_m\right) \quad (1)$$

не учитывает случайную погрешность измерения, причиной которой могут быть стохастические явления, обусловленные люфтом, трением, гистерезисом, совместным влиянием влияющих величин и т.д. По понятной причине действие указанных факторов не может быть учтено, например, в практической метрологической задаче построения градуировочной характеристики СИ.

Результаты планируемого физического эксперимента представляют собой результаты совместных измерений входной и выходной величин искомой модели (входного и выходного сигнала СИ)

$$Y_{эксп} = \varphi\left(X^{эксп}{}_1, X_2, {}^{эксп}..., X^{эксп}{}_k\right), \quad (2)$$

выполненных с какой-то погрешностью. Гибридная модель представляет собой функцию, принадлежащую области, ограниченной зависимостями (1) и (2), т.е. построена на интервале $\left[X^{теор}, X^{эксп}\right]$ как иллюстрирует рисунок .

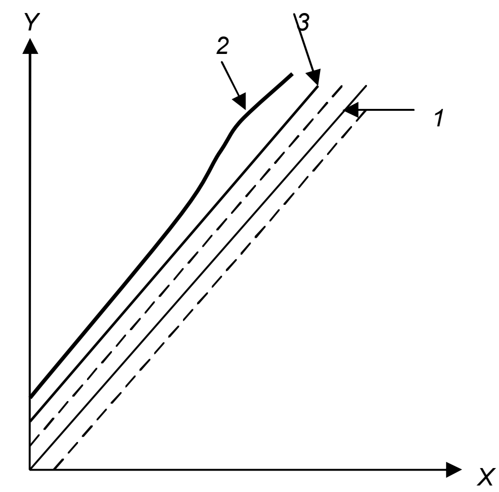

Рисунок - Область нахождения гибридной модели: 1 – теоретическая модель (с обозначенным пунктиром интервалом неопределённости результата измерения); 2 – регрессионная зави-

симость, построенная по результатам эксперимента; 3 – гибридная регрессионная модель

Для нахождения искомой гибридной модели воспользуемся декомпозиционными функциями, наиболее широко из которых в практических задачах нашли класс сумм квадратов отклонений, возникающих при применении метода наименьших квадратов (МНК).

Искомая регрессионная модель находится из условия

$$\begin{bmatrix} f\left(X^{теор}{}_1, X^{теор}{}_2, ..., X^{теор}{}_m\right) - \\ -\varphi\left(X^{эксп}{}_1, X^{эксп}{}_2, ..., X^{эксп}{}_k\right) \end{bmatrix}^2 \Rightarrow min$$

Неизвестными в искомой регрессионной гибридной модели будут коэффициенты ν_i и μ_j при аргументах $X^{теор}{}_{i/i=1-m}$ и $X^{эксп}{}_{j/j=1-k}$ соответственно, которые определяют, решая систему из $(m+k)$уравнений, что означает равенство нулю частных производных функций

$$\begin{cases} \dfrac{F(\nu_0,...,\nu_m,\mu_0,...,\mu_k)}{\partial \nu_{0-m}} = 0 \\ \dfrac{F(\nu_0,...,\nu_m,\mu_0,...,\mu_k)}{\partial \mu_{0-k}} = 0 \end{cases}$$

Заключение

Предлагаемый метод построения гибридной регрессионной модели на интервале $\left[X^{теор}, X^{эксп}\right]$ был использован в задаче измерения напряжений высоковольтными вакуумными делителями. На основе известного порядка элементов электрической схемы делителя была рассчитана первая граница интервала модели $X^{теор}$. Однако построенная теоретическая модель не могла отразить влияние на точность результата измерения напряжения таких факторов, как температура (не учтён ТКЕ – температурный коэффициент ёмкости), величина приложенного напряжения (нелинейность ВАХ), присутствие паразитных проводников (пылевые частицы, утечка по изоляционной оболочке конденсатора) и т.д. Дальнейшее уточнение априорной теоретической модели за счёт информационной избыточности и определение второй границы интервала $X^{эксп}$ нахождения модели было получено по результатам планируемого измерительного эксперимента, что позволило минимизировать стохастическую неопределённость индивидуальной характеристики высоковольтного вакуумного делителя. (Статья подготовлена в рамках реализации проекта «Создание высококачественных вакуумных конденсаторов» ФЦП «Научные и научно-педагогические кадры инновационной России (2009-20013 гг.)», Гос. контракт № П 489 от 13 мая 2010 г.)

СПИСОК ЛИТЕРАТУРЫ

1. Пфанцагль И. Теория измерений. [Текст] / И. Пфанцагль – М.: Мир, 1976. 248 с. // Johann Pfanzagl. Theory of measurement.- Physica-Verlag. Würzburg-Wien, 1971.
2. Кузнецов В.П. Сопоставительный анализ погрешности и неопределённости измерений [Текст] / В.П. Кузнецов // Измерительная техника. - 2003. - №8. С. 21-27.
3. Ординарцева Н.П. МЕТРОЛОГИЯ + СТАНДАРТИЗАЦИЯ + СЕРТИФИКАЦИЯ: учебное пособие [Текст] / Н. П. Ординарцева. – Пенза: Изд-во ПГУ, 2010. – 134 с. // Свидетельство о публикации в электронной библиотеке федеральной системы информационных образовательных ресурсов. Рег. № 73241/05-2011, Москва, 2011 г. [Электронный ресурс] / Н.П. Ординарцева. Режим доступа: http://window.edu.ru/window/library
4. Вощинин, А.П. Интервальный анализ данных: развитие и перспективы [Текст] / А.П. Вощинин // Заводская лаборатория. Диагностика материалов. – 2002. - №1. Том 68. С. 118-126.
5. Радев, Христо К. О подходах к измерению и его точности. Математическая, статистическая и компьютерная поддержка качества измерений: материалы Международного научно-практического семинара. – С-Петербург, ГНЦ РФ «ВНИИМ им. Д.И. Менделеева», 2009. [Электронный ресурс] / Христо К. Радев. Режим доступа: http://mscsmq.vniim.ru/files/2009/rus/1-radev-09-ru.pdf
6. Шокин, Ю.И. Интервальный анализ. [Текст] / Ю.И. Шокин. – Новосибирск: Наука, 1981. 112 с.
7. Ординарцева Н.П. Метод гибридного моделирования в регрессионном анализе [Текст] / Н.П. Ординарцева // Вопросы электроники. – 2012. - №1. С. 136-143.

***Ординарцева Н.П.**, Пензенский государственный университет, nat@rclink.ru, 440026, г. Пенза, ул. Красная 40, тел. (8-412)-368233, кандидат технических наук, доцент, докторант каф. МСК Пензенского госуниверситета //*

***Ordinartseva N. P.**, Penza State University, nat@rclink.ru, Russia, Penza, Krasnaya Street 40, (8-412)-368233, the Assistant of Professor MQS of PSU*