Автоматизация и управление технологическими процессами и производствами (05.13.06)

УДК 621.398
DOI: 10.24160/1993-6982-2019-3-108-115

Анализ подходов к решению задачи построения математической модели объекта по неточным экспериментальным данным

Н.В. Скибицкий

Проанализированы источники неопределённости и особенности применения различных моделей описания неточных данных, а также подходы к решению задач построения прямой и обратной статических характеристик системы по неточным данным, базирующихся на разных моделях описания неопределённости.

Показано, что применение подхода, наиболее часто используемого на практике, основанного на вероятностной (статистической) модели описания неопределённости, целесообразно лишь в случае, когда неопределённость связана только со случайной вариабельностью. Описание других источников неопределённости в рамках данной модели затруднительно. В этом случае при описании помехи приоритетным является нормальное распределение, на постулировании которого основаны последующие выводы, а для описания механизма её действия чаще всего используют модель аддитивной помехи на выходе с нулевым значением ошибки на входе системы. В предположении выполнения указанных предпосылок для решения задачи построения прямой и обратной статических характеристик взят аппарат регрессионного анализа. Все оптимальные свойства получаемых в этом случае оценок параметров справедливы только для прямой модели, при выполнении следующей достаточно жёсткой системы предположений, которые часто нарушаются на практике в силу того, что

- реальные законы распределения ошибок разнообразны и часто далеки от нормального;
- для установления действительного вида функции распределения необходимо проведение испытаний, число которых должно быть тем больше, чем большим выбирается значение доверительной вероятности;
- ошибки измерений входных величин существенны, что приводит к смещённым оценкам параметров модели, а ошибки измерений выхода содержат как случайные, так и систематические составляющие, вследствие чего метод наименьших квадратов не является оптимальным в статистическом смысле;
- случайные ошибки могут быть зависимыми, а их корреляционная матрица неизвестной и часто, хотя и необоснованно, используемые в этом случае обычные оценки наименьших квадратов не будут обеспечивать наименьшие дисперсии;
- помеха измерения и/или установки факторов может описываться не только аддитивной моделью, используемой в большинстве случаев на практике, но и мультипликативной или аддитивно-мультипликативной.

В рамках регрессионного анализа нет теоретически обоснованного метода построения обратной характеристики и её доверительного интервала, поэтому применение статистического подхода к решению задачи построения обратных характеристик приводит к серьезным трудностям, а формальное применение регрессионного анализа даёт результаты, далекие от истинных:

- разделение переменных на точно измеряемые входные и измеряемые с ошибками выходные нарушается при построении обратных характеристик;
- аналитическое определение доверительного интервала для предсказанного значения входной переменной теоретически необосновано и представляется крайне затруднительным;
- существенным недостатком доверительного интервала значения ошибки является невозможность суммирования её составляющих, так как доверительный интервал суммы не равен сумме доверительных интервалов слагаемых;
- из-за существенной нелинейности обратного преобразования модель ошибки является очень сложной и не может быть описана в терминах абсолютной или относительной ошибок.

С учётом этого наиболее полной характеристикой точности решения задачи построения обратной функции представляется интервал её неопределённости, а перспективным является применение подхода, основанного на использовании интервальной модели, позволяющей описать широкий класс неопределённых, вариабельных и неточных исходных данных.

Ключевые слова: модель, неопределённость, прямая и обратная статическая характеристика, регрессионный анализ, интервальный анализ.

Для цитирования: Скибицкий Н.В. Анализ подходов к решению задачи построения математической модели объекта по неточным экспериментальным данным // Вестник МЭИ. 2019. № 3. С. 108—115. DOI: 10.24160/1993-6982-2019-3-108-115.

Analysis of Approaches to Constructing a Plant Mathematical Model Based on Inaccurate Experimental Data

N.V. Skibitskiy

The article analyzes the sources of uncertainty and peculiarities of using different models for describing inaccurate data, as well as approaches to constructing the forward and reverse static characteristics of a system from inaccurate data based on different uncertainty description models.
It is shown that application of the approach based on a probabilistic (statistical) uncertainty description model, which is most frequently used in practice, is expedient only in the case when the uncertainty is solely due to random variability. The description of other uncertainty sources within the framework of this model involves difficulties. In this case, interference is predominantly described by a normal distribution, the postulation of which serves as the basis for inferring the subsequent conclusions. As to the interference action mechanism, it is most frequently described using the model of additive interference at the output with a zero error value at the system input. Assuming that these prerequisites are fulfilled, the regression analysis techniques were taken to solve the problem of constructing the forward and reverse static characteristics. All optimal properties of the parameter estimates obtained in this case are only valid for the forward model, provided that the following sufficiently stringent system of assumptions is valid:
(i) The actual error distribution laws are various and often far from being normal.
(ii) To establish the actual form of the distribution function, it is necessary to carry out tests; the higher the confidence probability, the larger the number of tests should be.
(iii) The input quantities measurement errors are essential, which leads to biased estimates of the model parameters, and the output quantity measurement errors contain both random and systematic components, due to which the least-squares method is not optimal in the statistical sense.
(iv) Random errors can be dependent, and their correlation matrix is unknown. Therefore, the usual least-squares estimates, which are frequently, although unreasonably, used in this case, will not provide the smallest variance values.
(v) The factor measurement and/or setting interference can be described not only by the additive model that is used in the majority of practical cases, but also by a multiplicative or an additive-multiplicative model.
In the regression analysis framework, there is no theoretically grounded method for constructing the reverse characteristic and its confidence interval. Therefore, attempts to apply the statistical approach to constructing reverse characteristics involve serious difficulties, and formal application of regression analysis yields results that are far from being true:
● The separation of variables into accurately measured input ones and output ones measured with errors is upset in constructing reverse characteristics.
● Analytical determination of the confidence interval for the predicted input variable value is not theoretically grounded and seems to be extremely difficult.
● A significant drawback of the error value confidence interval is the impossibility of summing up its components, because the confidence interval of the sum is not equal to the sum of the confidence intervals of the terms.
● Owing to essential nonlinearity of the reverse transformation, the error model is very complex and cannot be described in terms of absolute or relative errors.
In view of the above, the reverse function uncertainty interval seems to be most complete parameter characterizing the accuracy of solving the reverse function construction problem. In this regard, application of the approach based on the interval model that makes it possible to describe a wide class of uncertain, variable and inaccurate initial data is most promising.

Key words: model, uncertainty, forward and reverse static characteristics, regression analysis, interval analysis.

For citation: Skibitskiy N.V. Analysis of Approaches to Constructing a Plant Mathematical Model Based on Inaccurate Experimental Data. Bulletin of MPEI. 2019;3:108—115. (in Russian). DOI: 10.24160/1993-6982-2019-3-108-115.

Введение

Решение задачи управления или прогноза поведения системы — завершающий этап длительного процесса, включающего построение математической модели. Исследователь часто имеет дело с неточно известными данными, а, значит, модель оказывается известной лишь приближенно, поэтому важно правильно оценить помехи, действующие на объект, что невозможно без выбора адекватной формы описания данных, полученных в условиях неопределённости.

Поставлена задача анализа подходов к построению прямых и обратных моделей объектов по неточным экспериментальным данным.

Модели описания неопределённости

Модели, позволяющие получить аналитическое описание процесса путем обработки неточных экспериментальных данных, применяются для решения широкого класса прикладных задач. На практике часто используется линейная по параметрам статическая характеристика объекта, позволяющая решать так называемые прямые задачи, связанные с оценкой выходного значения y при заданном входном значении x [1, 2]. Применительно к системе с одним входом и одним выходом она может быть представлена в виде:

$$y = f(x) = b_1\varphi_1(x) + \ldots + b_m\varphi_m(x), \tag{1}$$

где $\varphi_i(x)$ — известные базисные функции; b_i — неизвестные параметры, подлежащие определению.

При исследовании следящих систем, в метрологии и т. д. помимо прямой задачи необходимо решать и обратную, т. е. находить оценку x при заданном значении y [3, 4]:

$$x(y) = f^{-1}(b_1,\ldots,b_m,x) = \psi(a_1,\ldots,a_m,y), \tag{2}$$

где a_j — неизвестные параметры, выражаемые через параметры b_i.

Очевидно, что задача имеет смысл, только если функция $f(x)$ зависит от единственной входной переменной x и является однозначной и монотонной.

Для описания неточно известных данных применяют три модели: вероятностную, нечеткую и интервальную.

Вероятностная модель используется для описания неточных данных в предположении, что они являются случайными величинами [2, 5]. Исчерпывающим описанием непрерывной случайной величины x служит её функция плотности вероятности $w(x)$. На практике исследователь, не имея возможности на основе ограниченных наблюдений получить точное описание функции плотности вероятности, оперирует с её статистической оценкой.

Интервальная модель предполагает задание интервала неопределённости на значения неточно известных данных, описываемого своими нижней и верхней границами [6]. Для его определения может быть использован широкий спектр априорной информации: результаты экспериментов, сведения об абсолютных и относительных ошибках, ошибках округления, экспертные данные и т. п.

Нечёткая модель в своей основе содержит понятия нечёткого множества и функции принадлежности, задаваемой обычно экспертным путем и определяющей степень принадлежности возможного значения нечёткой переменной нечёткому множеству [7]. Модель пригодна для описания широкого спектра источников неопределённости, однако её применение встречает методологические трудности при сравнении и ранжировании нечётких чисел и сглаживании нечётких данных.

Вероятностный подход к построению прямых статических характеристик объекта

Данный подход получил наибольшее распространение. В его основе — данные эксперимента, представляющие собой множества $\{x_k, y_k, k = 1...N\}$, где x_k, y_k — значения входной и выходной переменных в k-м эксперименте. Последовательность действий в этом случае включает ряд этапов [1, 2, 8].

1. *Выбор модели описания неточности измерений.* Для описания механизма действия ошибок чаще всего используют модель аддитивной помехи

$$y = y_0 + e_y;\; x = x_0 + e_x, \tag{3}$$

где y_0, x_0 — точные значения переменных; e_x, e_y — значения ошибки измерения входной и выходной переменных.

Выделяют три случая, каждый из которых вносит специфику в процедуру обработки данных:

● аддитивную модель ошибки измерения y, когда считается, что значения величины x фиксируются с высокой точностью на заранее заданных уровнях:

— с постоянной дисперсией:

$$e_x \equiv 0;\; e_y \neq 0;\; \sigma^2_{e_y} = \mathrm{const}; \tag{4}$$

— в предположении наличия неоднородности дисперсии в опытах:

$$e_x \equiv 0;\; e_y \neq 0;\; \sigma^2_{e_y} = \mathrm{var}; \tag{5}$$

● аддитивную модель ошибки измерения x, когда считается, что

$$e_x \neq 0;\; e_y \equiv 0; \tag{6}$$

● аддитивную ошибку в обеих переменных x и y, когда ошибки присутствуют как во входной, так и в выходной переменных:

$$e_y \neq 0;\; ex \neq 0. \tag{7}$$

2. *Выбор вида функции.* При выборе вида функции стараются подобрать по возможности простую, желательно линейную по параметрам функцию, примеры которых приведены в [1].

3. *Проведение эксперимента.* Повысить качество полученных оценок позволяют методы теории планирования эксперимента [9, 10]. При планировании активного эксперимента считается, что значения входной величины x устанавливаются экспериментатором с высокой точностью, и ошибками в её значениях можно пренебречь. При проведении пассивного эксперимента считается, что исследователь вынужден ограничиться результатами доступных наблюдений, а ошибки присутствуют как во входной, так и в выходной переменных.

4. *Оценка коэффициентов статической характеристики.* Цель — поиск оценок $\hat{b}_i$ неизвестных коэффициентов b_i выбранной функции вида (1). Наиболее детально в статистике разработана процедура оценивания параметров при выполнении условия (4) и нормальном распределении аддитивной ошибки с нулевым математическим ожиданием и постоянной дисперсией σ^2. Данная задача получила название задачи построения регрессионной модели, а для её решения применяется аппарат регрессионного анализа [2, 5]. Методы оценивания в этом случае базируются на минимизации некоторой нормы отклонений предсказанных значений $\hat{y}_k$, определяемых как

$$\hat{y}_k = \hat{b}_1\varphi_1(x_k) + \hat{b}_2\varphi_2(x_k) + ... + \hat{b}_m\varphi_m(x_k), \tag{8}$$

от полученных в эксперименте y_k. Обычно в качестве нормы используется сумма квадратов отклонений, и вектор неизвестных оценок $\hat{\mathbf{b}} \in R^m$ определяется методом наименьших квадратов (МНК) по формуле

$$\hat{\mathbf{b}} = (F^T F)^{-1} F^T \mathbf{y}, \tag{9}$$

где F — $(N\times m)$-матрица значений базисных функций $\varphi_i(x_k)$; $\mathbf{y}\in R^N$ — вектор значений переменной y_k в N опытах.

При сделанных допущениях относительно ошибок МНК-оценки (9) имеют нормальное распределение, являются состоятельными, несмещёнными и эффективными.

При выполнении условия (5) для определения оценок параметров используют взвешенный МНК и оценки рассчитывают следующим образом:

$$\hat{\mathbf{b}} = (F^T W^{-1} F)^{-1} F^T W^{-1} \mathbf{y},$$

где матрица $W \in R^{N\times N}$ принимается диагональной с элементами σ_k^2, представляющими дисперсии ошибок переменной y_k и определяемыми в результате обработки параллельных опытов.

Доверительный интервал предсказания зависит от ковариационной матрицы оценок параметров D, которая при справедливости условий (4) или (5) выглядит как

$$D = s^2(F^TF)^{-1} \text{ или } D = (F^TW^{-1}F)^{-1},$$

где s^2 — выборочная оценка дисперсии ошибки e_y.

Знание ковариационной матрицы позволяет записать дисперсию $\sigma^2(\hat{y}(x))$ ошибки предсказания по модели (8):

$$\sigma^2(\hat{y}(x)) = \boldsymbol{\varphi}^T(x) D \boldsymbol{\varphi}(x).$$

Если ошибка имеет нормальное распределение, то коридор ошибок для предсказанного значения выходной переменной по (8) определяют в виде доверительного интервала:

$$P\left\{\hat{y}(x) - t_{\alpha,N-m}\sigma(\hat{y}(x)) \le y(x) \le \hat{y}(x) + t_{\alpha,N-m}\sigma(\hat{y}(x))\right\} \le \\ \le (1-\alpha),$$

где $t_{\alpha,N-m}$ — квантиль распределения Стьюдента для доверительной вероятности α и числа степеней свободы $N - m$.

Вместе с тем оптимальность оценок неизвестных параметров сохраняется только при выполнении определённых условий. Так, при справедливости соотношений (6) или (7) применение (9) для установления оценок $\hat{b}_i$ приводит, как правило, к смещённым и несостоятельным оценкам. Поэтому рекомендуется применение методов конфлюэнтного анализа [11], что требует дополнительной информации и редко позволяет получить обоснованные результаты для моделей общего вида.

Условия, обеспечивающие получение достоверных результатов в случае выполнения соотношения (7), рассмотрены в [8] при анализе влияния погрешностей регистрации данных на точность математического описания. Они определяют требования к точности регистрации отдельных переменных в предположении, что ошибки измерения по каждой переменной представляют собой случайные величины с нулевым математическим ожиданием, не коррелированны с ошибками по другим переменным и со значениями самих переменных.

При распределениях погрешностей, значительно отличающихся от гауссовских, а также наличии выбросов оценки МНК имеют большие погрешности. В таких случаях предлагается применять робастные методы построения зависимостей [12] в предположении, что распределения погрешностей не очень сильно отличаются от гауссовского распределения. Использование робастного подхода для линейной статической характеристики обычно приводит к системе нелинейных относительно оцениваемых параметров уравнений, решаемых численными методами. В данном случае особое значение приобретают итерационные схемы нахождения оценок. Важную роль играет также выбор начальных приближений для оцениваемых параметров. Робастные методы для построения нелинейных функций получаются аналогично, но с возрастанием вычислительной сложности. Аналитическое исследование робастных методов весьма затруднительно, поэтому в основном используются методы статистического моделирования.

Следует отметить, что на практике модели ошибок не исчерпываются случаями (4) — (7). Так, имеют место случаи мультипликативной помехи как на входе, так и на выходе, а так же комбинированной модели воздействия ошибок [13, 14].

Построение обратных статических характеристик объектов

Проанализированная статическая характеристика позволяет решать так называемые прямые задачи, связанные с оценкой значения y при заданном значении x. В ряде случаев необходимо решать и обратную задачу, т. е. искать оценку x при заданном значении y в соответствии с (2).

В рамках статистического подхода построение обратной функции по неточным данным, как правило, осуществляется за два шага. Сначала по экспериментальным данным находят оценки коэффициентов прямой функции, а затем алгебраически осуществляют переход к обратной зависимости. Тогда оценки коэффициентов обратной зависимости определяются неточными оценками коэффициентов прямой функции и предсказанное значение $\hat{x}(y)$ содержит ошибки. При этом даже для простейшего линейного случая переход от прямой модели к обратной является нетривиальной задачей.

Пусть при нормально распределенных аддитивных ошибках измерения y с нулевым математическим ожиданием найдены несмещенные оценки параметров и записана прямая линейная модель

$$\hat{y} = \hat{b}_1 + \hat{b}_2 x.$$

Допустим, что оценки $\hat{b}_2$, $\hat{b}_2$ — независимые случайные величины с нормальным распределением с известными дисперсиями, тогда оценка предсказанного значения $\hat{y}$ также нормально распределена, а обратная модель определяется линейной функцией

$$\hat{x} = \hat{a}_1 + \hat{a}_2 y;\; \hat{a}_1 = -\left(\hat{b}_1 / \hat{b}_2\right);\; \hat{a}_2 = \left(1 / \hat{b}_2\right).$$

Распределение предсказанного значения $\hat{x}$ зависит от распределения оценок $\hat{a}_1, \hat{a}_2$. Первая из оценок $\hat{a}_1$— отношение двух нормально распределенных случайных величин, описываемое распределением Коши, причём дисперсия в этом случае принципиально не может быть указана, так как определяющий её интеграл расходится, а определённое значение математического ожидания отсутствует, т.е. использование данной оценки неправомерно [5]. Вторая представляет собой величину, обратную к нормальной случайной величине, и теоретически имеет неограниченный диапазон, что связано с возможностью деления на ноль.

Определение доверительного интервала для предсказанного значения переменной $\hat{x}$ при статистическом подходе затруднительно. Действительно, доверительный интервал прямой модели определяется границами, рассчитываемыми по формуле

$$\hat{y}(x) \pm s t_{(\alpha, N-m)} \sqrt{1 + \varphi^T(x) D \varphi(x)}, \qquad (10)$$

что не позволяет записать аналитически границы для $\hat{x}$, поскольку не удаётся записать функцию, обратную к (10). Ситуация еще больше осложняется, если используется не линейная, а другие функции.

Ограниченность гипотез статистического подхода

Практические приложения регрессионного анализа показали, что его исходные предпосылки часто далеки от реальности. Так, широкое распространение нормального распределения погрешностей в практике измерений обычно объясняется центральной предельной теоремой теории вероятностей, утверждающей, что если результат наблюдения является суммой многих случайных слабо взаимозависимых величин, каждая из которых вносит малый вклад относительно общей суммы, то при увеличении числа слагаемых распределение центрированного и нормированного результата стремится к нормальному. Вместе с тем, основное воздействие на ошибку оказывают погрешности приборов и результатов измерений, особенность которых состоит в их большом разнообразии.

Результаты исследования распределения погрешности различных электроизмерительных приборов [14] показали, что закон распределения этих погрешностей часто существенно отличается от нормального, и в 1974 г. был введен в действие ГОСТ 8.011—72, устанавливающий, что при сообщении результата измерения целесообразно указывать вид распределения, были стандартизованы модели равномерного, трапецеидального, треугольного, нормального и двухмодального распределений. Кроме того, результирующая погрешность измерений складывается из ряда составляющих, и если их рассматривать как случайные величины, то суммирование погрешностей сведется к суммированию случайных величин. Вместе с тем при суммировании случайных величин форма закона распределения суммы резко отличается от формы распределения составляющих.

Нельзя забывать и о существенно различном вкладе отдельных составляющих в результат, и что возможность идентификации закона распределения экспериментальных данных ограничена малостью объема выборки. Кроме того, характеристика точности результатов измерений предполагает указания границ интервала, в котором данная погрешность находится с заданной вероятностью. Для этих целей используются квантильные оценки. Традиционно считают, что ошибки результатов измерения распределены нормально (хотя это часто не так) и дают рекомендации по пользованию таблицей квантилей стандартного нормального распределения, что приводит к ошибочным результатам.

Погрешность средств измерения и статическая характеристика преобразования

Задача построения прямой и обратной функций играет важную роль в метрологии, где к числу нормируемых метрологических характеристик средств измерений относится градуировочная характеристика (ГХ), устанавливающая зависимость (1) выходного сигнала y измерительного преобразователя от входного сигнала x [1]. Процедура построения ГХ детально проработана в рамках вероятностного подхода и предусматривает построение по экспериментальным данным прямой ГХ с последующим её использованием для нахождения входной величины x в соответствии с (2) [1, 4, 15]. Иногда предлагают сразу определять обратную функцию для непосредственного нахождения значения входной величины. Так, решение задачи получения обратной зависимости методом максимального правдоподобия при наличии ошибки установки входных переменных предложено в [16].

На практике выбор подхода основывают на соотношении между погрешностями измерений входной и выходной величин.

Если погрешности измерений входных величин пренебрежимо малы по сравнению с погрешностями выходных, то считается, что первоначально целесообразно строить прямую ГХ, а затем для определения x по y использовать обратную функцию. Методы построения ГХ в этом случае базируются на методах регрессионного анализа в предположении, что результаты измерений выходных величин описываются моделью

$$y_i = y_{i0} + e_{yi},$$

где e_{yi} — независимая нормально распределенная случайная погрешность с нулевым средним и известной дисперсией.

Если пренебрежимо малы погрешности измерений выходных величин, то целесообразно сразу строить обратную ГХ. Для этого применяют те же методы, что и в первом случае, и полученную ГХ непосредственно используют для оценивания значений x по y.

Если относительные погрешности измерений входных и выходных величин одного порядка, то статистические методы построения ГХ опираются на методы конфлюентного анализа.

При использовании статистических методов оценивание погрешности ГХ предусматривает построение доверительных интервалов, что позволяет получить доверительный коридор для зависимости $y = f(x)$.

Современный подход к выражению неопределенности измерений отражен в [17 — 19], где отмечается, что:

● понятие «неопределенности» как количественной характеристики является относительно новым в истории измерений и его появление связано с тем, что после того, как найдены оценки всех ожидаемых составляющих погрешности, а в результат измерения внесены соответствующие поправки, все ещё остаётся сомнение в том, насколько точно он соответствует значению измеряемой величины;

● термины «погрешность» и «неопределённость» не являются синонимами, поскольку погрешность результата измерения — идеализированное понятие и на практике её точное значение неизвестно, а неопределенность результата измерения отражает отсутствие точного знания значения измеряемой величины.

В [19] отмечено, что «неопределённость измерения — параметр, относящийся к результату измерения и характеризующий разброс значений, которые могли бы быть обоснованно приписаны измеряемой величине», а идеальный метод для выражения неопределённости измерения должен предоставлять возможность указать интервал, в пределах которого, можно предполагать, находится большая часть распределения значений, которые обоснованно могут быть приписаны подлежащей измерению величине. В соответствии с этим неопределённость в результате измерения состоит из нескольких составляющих, группирующихся в две категории в соответствии со способом их численного оценивания:

● составляющие, оцениваемые статистическими методами;

● составляющие, оцениваемые другими способами.

В этой связи интерес представляет интервальный подход к оценке погрешности измерения [20, 21]. Его актуальность связана, в том числе, с тем, что неточность измерительного прибора задается максимальной абсолютной погрешностью Δ или классом точности прибора, связанного с величиной относительной погрешности δ. При наличии единичного измерения x и погрешности Δ (или δ) интервал возможных значений неизвестного истинного значения x_0 выглядит как:

$$x_0 = x \pm \Delta \text{ (или } x_0 = x \pm \delta\%). \quad (11)$$

Тогда от представления (11) легко перейти к интервальному представлению результата, и при заданной абсолютной (относительной) погрешности измерения интервал неопределённости истинного значения равен:

$$[x]\colon x - \Delta \le x_0 \le x + \Delta = x_{\min} \le x_0 \le x_{\max}, \quad \text{или } [x]\colon x - \delta|x| \le x_0 \le x + \delta|x|). \quad (12)$$

Принципиальным моментом в этом случае является гипотеза о том, что истинное значение наверняка содержится внутри интервала (12).

Заключение

Проанализированы источники неопределенности и особенности применения различных моделей описания неточных данных в научных задачах.

Показано, что использование наиболее популярной вероятностной модели целесообразно в случае, когда неопределённость при описании данных связана только со случайной вариабельностью, описание других источников в рамках указанной модели затруднительно.

Достоинство регрессионного анализа моделей при нулевых ошибках во входной переменной заключается не только в глубокой теоретической проработке, но и в наличии пакетов программ статистического анализа. Все это обусловило его широкое применение при построении математических моделей.

Вместе с тем при вероятностном подходе все оптимальные свойства оценок параметров справедливы только для прямой модели, причём при выполнении достаточно жёсткой системы предположений, которые часто нарушаются на практике, а именно:

● реальные законы распределения ошибок разнообразны и часто далеки от не всегда оправданно постулируемого нормального, на применении которого основаны многие последующие выводы;

● для установления действительного вида функции распределения необходимо проведение испытаний, число которых должно быть тем больше, чем большим выбирается значение доверительной вероятности;

● случайные ошибки могут быть зависимыми, а их корреляционная матрица неизвестной и часто, хотя и необоснованно, используемые в этом случае оценки МНК не будут обеспечивать наименьшие дисперсии;

● ошибки измерений выхода содержат как случайные, так и систематические составляющие, вследствие чего МНК не является оптимальным в статистическом смысле;

● помеха измерения и/или установки факторов может описываться не только используемой на практике в большинстве случаев аддитивной моделью, но и мультипликативной или аддитивно-мультипликативной.

В рамках регрессионного анализа:

● разделение переменных на точно измеряемые входные и измеряемые с ошибками выходные для большого числа приложений является искусственным и нарушается при построении обратных характеристик;

● аналитическое определение доверительного интервала для предсказанного значения входной переменной теоретически не обосновано и представляется крайне затруднительным, существенным недостатком доверительного значения ошибки является невозможность суммирования её составляющих;

● из-за нелинейности обратного преобразования модель ошибки является очень сложной, и поэтому наиболее полной характеристикой точности решения задачи построения обратной функции представляется интервал её неопределённости, полученный по интервальной модели и позволяющий описать широкий класс неточных исходных данных.

Литература

1. **Семенов Л.А., Сирая Т.Н.** Методы построения градуировочных характеристик средств измерений. М.: Изд-во стандартов, 1986.
2. **Орлов А.И.** Прикладная статистика. М.: Экзамен, 2004.
3. **Сирая Т.Н.** Методы обработки данных при измерениях и метрологические модели // Измерительная техника. 2018. № 1. С. 9—14.
4. **Лячнев В.В., Сирая Т.Н., Довбета Л.И.** Метрологические основы теории измерительных процедур. СПб.: Элмор, 2011.
5. **Климов Г.П.** Теория вероятностей и математическая статистика. М.: Изд-во МГУ, 2011.
6. **Вощинин А.П.** Интервальный анализ данных: развитие и перспективы // Заводская лаборатория. Диагностика материалов. 2002. № 1. С. 118—126.
7. **Заде Л.А.** Понятие лингвистической переменной и его применение к принятию приближенных решений. М.: Мир, 1976.
8. **Бородюк В.П., Лецкий Э.К.** Статистическое описание промышленных объектов. М.: Энергия, 1971.
9. **Налимов В.В., Голикова Т.И.** Логические основания планирования эксперимента. М.: Металлургия, 1980.
10. **Авдеев Б.Я.** Планирование измерительного эксперимента. СПб.: Изд-во СПбГЭЕУ «ЛЭТИ», 2005.
11. **Грешилов А.А., Стакун В.А., Стакун А.А.** Статистические задачи принятия решений с элементами конфлюентного анализа. М.: Радио и связь, 1998.
12. **Вощинин А.П., Скибицкий Н.В.** Интервальный подход к выражению неопределенности измерений и калибровке цифровых измерительных систем // Заводская лаборатория. Диагностика материалов. 2007. № 11. С. 66—71.
13. **Вощинин А.П., Скибицкий Н.В.** Интервальный подход к анализу однофакторных мультисенсорных систем // Датчики и системы. 2006. № 4. С. 2—7.
14. **Алексеева И.У.** Теоретическое и экспериментальное исследование законов распределения погрешностей, их классификация и методы оценки их параметров: автореферат дис. … канд. техн. наук. Л.: Изд-во ЛПИ, 1975.
15. **Введение** к «Руководству по выражению неопределенности измерения и сопутствующим документам. Оценивание данных измерений» / под ред. В.А. Слаева, А.Г. Чуновкиной. СПб.: Профессионал, 2011.
16. **Горяинов С.В.** Разработка статистических методов построения градуировочных характеристик мультисенсорных систем: автореферат дис. … канд. техн. наук. М.: Изд-во МЭИ, 1997.
17. **Слаев В.А., Чуновкина А.Г.** Неопределенность измерений: пути использования международного документа // Приборы. 2004. № 9. С. 1—6.
18. **Оценивание** данных измерений — роль неопределенности измерений при оценке соответствия / под

References

1. **Semenov L.A., Siraya T.N.** Metody Postroeniya Graduirovochnykh Kharakteristik Sredstv Izmereniy. M.: Izd-vo Standartov, 1986. (in Russian).
2. **Orlov A.I.** Prikladnaya Statistika. M.: Ekzamen, 2004. (in Russian).
3. **Siraya T.N.** Metody Obrabotki Dannykh pri Izmereniyakh i Metrologicheskie Modeli. Izmeritel'naya Tekhnika. 2018;1:9—14. (in Russian).
4. **Lyachnev V.V., Siraya T.N., Dovbeta L.I.** Metrologicheskie Osnovy Teorii Izmeritel'nykh Protsedur. SPb.: Elmor, 2011. (in Russian).
5. **Klimov G.P.** Teoriya Veroyatnostey i Matematicheskaya Statistika. M.: Izd-vo MGU, 2011. (in Russian).
6. **Voshchinin A.P.** Interval'nyy Analiz Dannykh: Razvitie i Perspektivy. Zavodskaya Laboratoriya. Diagnostika Materialov. 2002;1:118—126. (in Russian).
7. **Zade L.A.** Ponyatie Lingvisticheskoy Peremennoy i Ego Primenenie k Prinyatiyu Priblizhennykh Resheniy. M.: Mir, 1976. (in Russian).
8. **Borodyuk V.P., Letskiy E.K.** Statisticheskoe Opisanie Promyshlennykh Ob′ektov. M.: Energiya, 1971. (in Russian).
9. **Nalimov V.V., Golikova T.I.** Logicheskie Osnovaniya Planirovaniya Eksperimenta. M.: Metallurgiya, 1980. (in Russian).
10. **Avdeev B.Ya.** Planirovanie Izmeritel'nogo Eksperimenta. SPb.: Izd-vo SPbGEEU «LETI», 2005. (in Russian).
11. **Greshilov A.A., Stakun V.A., Stakun A.A.** Statisticheskie Zadachi Prinyatiya Resheniy s Elementami Konflyuentnogo Analiza. M.: Radio i Svyaz', 1998. (in Russian).
12. **Voshchinin A.P., Skibitskiy N.V.** Interval'nyy Podkhod k Vyrazheniyu Neopredelennosti Izmereniy i Kalibrovke Tsifrovykh Izmeritel'nykh Sistem. Zavodskaya Laboratoriya. Diagnostika Materialov. 2007;11:66—71. (in Russian).
13. **Voshchinin A.P., Skibitskiy N.V.** Interval'nyy Podkhod k Analizu Odnofaktornykh Mul'tisensornykh Sistem. Datchiki i Sistemy. 2006;4:2—7. (in Russian).
14. **Alekseeva I.U.** Teoreticheskoe i Eksperimental'noe Issledovanie Zakonov Raspredeleniya Pogreshnostey, ikh Klassifikatsiya i Metody Otsenki ikh Parametrov: Avtoreferat Dis. … Kand. Tekhn. Nauk. L.: Izd-vo LPI, 1975. (in Russian).
15. **Vvedenie** k «Rukovodstvu po Vyrazheniyu Neopredelennosti Izmereniya i Soputstvuyushchim Dokumentam. Otsenivanie Dannykh Izmereniy». Pod Red. V.A. Slaeva, A.G. Chunovkinoy. SPb.: Professional, 2011. (in Russian).
16. **Goryainov S.V.** Razrabotka Statisticheskikh Metodov Postroeniya Graduirovochnykh Kharakteristik Mul'tisensornykh Sistem: Avtoreferat Dis. … Kand. Tekhn. Nauk. M.: Izd-vo MEI, 1997. (in Russian).
17. **Slaev V.A., Chunovkina A.G.** Neopredelennost' Izmereniy: Puti Ispol'zovaniya Mezhdunarodnogo Dokumenta. Pribory. 2004;9:1—6. (in Russian).
18. **Otsenivanie** Dannykh Izmereniy — Rol' Neopredelennosti Izmereniy Pri Otsenke Sootvetstviya. Pod

ред. В.А. Слаева, А.Г. Чуновкиной. СПб.: Профессионал, 2014.

19. **ГОСТ Р 54500.3—2011.** Неопределённость измерения. Ч. 3. Руководство по выражению неопределенности измерения.

20. **Вощинин А.П., Скибицкий Н.В.** Интервальный метод калибровки // Датчики и системы. 2000. № 9. С. 52—60.

21. **Скибицкий Н.В.** Построение прямых и обратных статических характеристик объектов по интервальным данным // Заводская лаборатория. Диагностика материалов. 2017. № 1. С. 87—93.

Red. V.A. Slaeva, A.G. Chunovkinoy. SPb.: Professional, 2014. (in Russian).

19. **GOST R 54500.3—2011.** Neopredelennost' Izmereniya. Ch. 3. Rukovodstvo po Vyrazheniyu Neopredelennosti Izmereniya.

20. **Voshchinin A.P., Skibitskiy N.V.** Interval'nyy Metod Kalibrovki. Datchiki i Sistemy. 2000;9:52—60. (in Russian).

21. **Skibitskiy N.V.** Postroenie Pryamykh i Obratnykh Staticheskikh Kharakteristik Ob′ektov po Interval'nym Dannym. Zavodskaya Laboratoriya. Diagnostika Materialov. 2017;1:87—93. (in Russian).

Сведения об авторе:

Скибицкий Никита Васильевич — доктор технических наук, профессор кафедры управления и информатики НИУ «МЭИ», e-mail: SkibitskyNV@mpei.ru

Information about author:

Skribitsky Nikita V. — Dr.Sci. (Techn.), Professor of Control and Informatics Dept., NRU MPEI, e-mail: SkibitskyNV@mpei.ru

Статья поступила в редакцию: *06.04.2018*
The article received to the editor: *06.04.2018*