

УДК 681.51:519.6

РОБАСТНЫЙ АЛГОРИТМ НЕПАРАМЕТРИЧЕСКОЙ ИДЕНТИФИКАЦИИ ТЕХНИЧЕСКИХ СИСТЕМ

^{1,2}Воскобойников Ю.Е., ¹Боева В.А.

¹ФГБОУ ВО «Новосибирский государственный архитектурно-строительный университет» (Сибстрин), Новосибирск, e-mail: voscob@mail.ru;

²ФГБОУ ВО «Новосибирский государственный технический университет», Новосибирск

Многие реальные технические системы описываются математическими моделями, аналитически представленными в виде интегрального вольтеровского уравнения I рода с разностным ядром. Задача непараметрической идентификации для подобных систем будет сводиться к построению оценки импульсной переходной функции идентифицируемой системы на основе измеренных (с шумами) значений входного и выходного сигналов. Такая постановка задачи традиционно определяется как некорректная, поскольку может не удовлетворять одному или нескольким условиям корректности, в частности условию устойчивости решения к погрешностям задания исходных данных. Для построения решения, удовлетворяющего условиям корректности, применяют как детерминированные, так и статистические регуляризирующие алгоритмы. В недавней работе авторов был предложен устойчивый алгоритм, суть которого заключается в переходе от решения вольтеровского уравнения I рода к решению уравнению II рода, что является уже корректной задачей. Вычисление первых производных входного и выходного сигналов, содержащихся в интегральном уравнении II рода, на основе аппарата сглаживающих кубических сплайнов позволило снизить негативное влияние шумов измерений на общую ошибку идентификации. К сожалению, сглаживающие кубические сплайны неудовлетворительно фильтруют аномальные измерения, и это приводит к значительным ошибкам идентификации. Поэтому в данной работе предлагается двухшаговый алгоритм идентификации. На первом шаге аномальные измерения эффективно удаляются с использованием пространственно-локального нелинейного фильтра, а на втором – по отфильтрованным значениям строятся сглаживающие сплайны, вычисляются производные и находится импульсная переходная функция. Выполненные исследования показали эффективность предлагаемого алгоритма идентификации.

Ключевые слова: задача идентификации, уравнение Вольтера I рода, уравнение Вольтера II рода, сглаживающие кубические сплайны, алгоритм решения уравнения Вольтера II рода, фильтрация аномальных измерений

ROBUST ALGORITHM OF NONPARAMETRIC IDENTIFICATION OF TECHNICAL SYSTEMS

^{1,2}Voskoboynikov Yu.E., ¹Boeva V.A.

¹Novosibirsk State University of Architecture and Civil Engineering (Sibstrin), Novosibirsk, e-mail: voscob@mail.ru;

²Novosibirsk State Technical University, Novosibirsk

Mathematical models that classify many of real technical systems have the theoretic representation in form of integral Volterra equation of the first kind with a difference kernel. For such systems, the non-parametric identification problem reduces to the estimation of pulse transition characteristics of the system from the registered noise-contaminated values of input and output signals. The problem posing this way is traditionally ill-posed due to the violation of one or several well-posedness conditions, partly, the solution stability condition concerning noise terms in initial data. To formulate solution for the problem complied with correctness conditions different regularization methods, both deterministic and statistical, exist. In their recent paper, authors propose the stable identification algorithm based on transition from original integral Volterra equation of the first kind solution to the well-posed second kind integral equation solution. Smoothing cubic splines calculation for first derivatives of the input and output signals in new integral equation allows to reduce the negative impact of measurement noises on total level of identification error. Unfortunately, smoothing cubic splines algorithm infiltrates badly abnormal data that follows to sufficient increase of identification errors. For all the reasons aforementioned, in this paper authors present the new two-step identification algorithm. Firstly, the effective local-spatial non-linear filtration procedure of abnormal data takes the place. Then, smoothing cubic splines calculating on filtered data bases, derivatives calculation and pulse transition characteristic estimation at once follow. All the studies investigated prove the confident efficiency of algorithm proposed.

Keywords: identification problem, Volterra integral equation of the first kind, Volterra integral equation of the second kind, Volterra integral equation of the second kind calculation algorithm, abnormal data infiltration

Интегральные модели достаточно часто используются для описания и моделирования технических систем [1, с. 12–18]. Если система является стационарной (т.е. ее параметры не меняются во времени), то для ее описания применяется интегральное уравнение Вольтера I рода [2, с. 25–29]:

$$\int_0^t k(t-\tau)\varphi(\tau)d\tau = f(t), t \in [0, T], \quad (1)$$

где $k(t)$ – импульсная переходная функция (ИПФ) системы; $\varphi(\tau), f(t)$ – ее входной

и выходной сигналы. Техническую реализуемость системы определяет условие

$$k(t) = 0 \text{ при } t < 0. \quad (2)$$

Применительно к модели (1), задача непараметрической идентификации [2, 3] системы сводится к вычислению оценки для $k(t)$ по измеренным значениям входного и выходного сигналов идентифицируемой системы.

Сформулированная задача идентификации некорректна, поскольку могут не удовлетворяться одно или несколько условий

корректности по Адамару [4, с. 15–18; 5, с. 9]. Наиболее часто нарушается условие устойчивости решения к погрешностям (ошибкам) исходных данных, когда малые погрешности могут вызвать сколь угодно большие ошибки решения. Для построения устойчивых решений некорректных задач используются специальные методы регуляризации, как детерминированные [4, с. 38–56], так и статистические [5, с. 124–129]. В работе авторов [6] был построен эффективный алгоритм, вычисляющий оценку ИПФ из решения интегрального уравнения Вольтера II рода, что является уже корректно поставленной задачей. Однако при этом необходимо было вычислить первые производные от входного и выходного сигналов, что в свою очередь также является неустойчивой задачей. Для устойчивого вычисления этих первых производных использовались сглаживающие кубические сплайны (СКС). Применяемый алгоритм оценивания оптимального параметра сглаживания СКС (из условия минимума среднеквадратической ошибки сглаживания [5, с. 65–67]) позволил вычислять первые производные с приемлемыми ошибками, что обусловило устойчивое решение задачи идентификации даже при больших значениях погрешностей измерения выходного и входного сигналов идентифицируемой системы.

В ряде случаев зарегистрированные значения входного и выходного сигналов могут содержать аномальные измерения, которые значительно (в разы) отличаются от близлежащих измерений, и их появление можно объяснить наличием шумовой составляющей (импульсного шума), дисперсия которой существенно (на два и более порядков) отличается от дисперсии шумов измерений других значений сигналов. Аномальные измерения могут быть обусловлены сбоями в измерительной системе, канале связи, регистрирующей аппаратуре и т.п. К сожалению, СКС очень слабо сглаживает такие импульсные шумы, и это обуславливает большие ошибки в вычислении первых производных, что в свою очередь приводит к существенным ошибкам идентификации ИПФ динамической системы.

В данной работе предлагается новый двухшаговый алгоритм идентификации, на первом шаге которого эффективно фильтруются аномальные измерения (т.е. удаляется импульсный шум) с использованием пространственно-локальных фильтров, а на втором шаге осуществляется построение СКС по фильтрованным значениям входного и выходного сигнала, вычисления первых производных и построение устойчивого решения с использованием алгоритма, изложенного в работе [6]. Особое внимание

уделяется оцениванию оптимальных параметров используемых алгоритмов обработки, что позволяет получать решения с наименьшей ошибкой идентификации.

Метод исследования

В работе [6] устойчивый алгоритм идентификации строится переходом от уравнения (1) к уравнению задачи идентификации, которое при выполнении условия (2) имеет вид

$$\int_0^t \varphi(t - \tau)k(\tau)d\tau = f(t), t \in [0, T]. \quad (3)$$

Путем дифференцирования этого уравнения по переменной t и выполнения несложных преобразований, приходим к интегральному уравнению Вольтера II рода вида

$$k(t) + \frac{1}{\varphi(0)} \int_0^t \varphi'(t - \tau)k(\tau)d\tau = \frac{f'(t)}{\varphi(0)}, \quad t \in [0, T]. \quad (4)$$

Для устойчивого дифференцирования зашумлённых входного и выходного сигналов идентифицируемой системы в работе [6] использовались сглаживающие кубические сплайны (СКС). Напомним, что функция $S_\varphi(x)$ называется кубическим сплайном на интервале $[0, T]$ с узлами

$$0 = t_1 < t_2 < \dots < t_N = T, \quad (5)$$

если: а) на каждом отрезке $[t_i, t_{i+1}), i = 1, \dots, N - 1$, функция $S_\varphi(t)$ представляет собой полином третьей степени вида

$$S_\varphi(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3; \quad (6)$$

б) функция $S_\varphi(t)$ является дважды непрерывно дифференцируемой на всем интервале $[0, T]$.

В работе [6] для устойчивого вычисления первых производных входного и выходного сигналов использовался естественный сглаживающий кубический сплайн – функция $S_{\varphi, \alpha}(t)$, которая одновременно удовлетворяет естественным краевым условиям относительно второй производной $S''_{\varphi, \alpha}(t): S''_{\varphi, \alpha}(t_1) = S''_{\varphi, \alpha}(t_N) = 0$ и доставляет минимальное значение функционалу [5, с. 42–44; 7, с. 124–132]:

$$F_\alpha(S) = \alpha \int_{t_1}^{t_N} |S''(t)|^2 dt + \sum_{i=1}^n p_i^{-1} (\varphi_i - S(t_i))^2, \quad (7)$$

где $p_i^{-1}, i = 1, \dots, N$ – весовые множители, $\varphi_i = \varphi(t_i), i = 1, \dots, N$ – значения функции

$\varphi(t)$ в узлах сетки (5), для которой строится СКС. В рассматриваемой задаче идентификации это зашумленные значения входного и выходного сигналов идентифицируемой системы. Меняя параметр сглаживания α в интервале $[0, \infty)$, можно менять ошибку сглаживания. Выбор параметра α из условия минимума среднеквадратической ошибки сглаживания рассмотрен в работах [5, 6]. Вычисление коэффициентов a_i, b_i, c_i, d_i при заданном параметре α подробно изложено в [5, с. 45–49].

Рассмотрим поведение сглаживающего сплайна при наличии аномальных измерений (АИ). В качестве иллюстрации на рис. 1 сплошной линией показаны точные значения входного сигнала, точечной кривой – зашумленные значения (относительный уровень шума 10%), штрих-точечной кривой – измеренные значения, содержащие АИ (вероятность появления АИ равна 0,02).

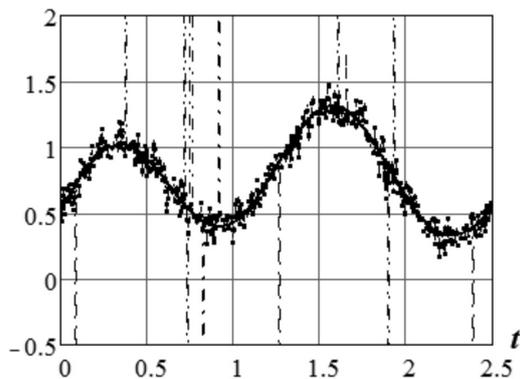


Рис. 1. Измеренные значения входного сигнала

На рис. 2 точечной кривой показаны значения СКС, построенного по измерениям, не содержащим АИ, штрих-точечная кривая – значения СКС, построенного по измерениям, содержащим АИ. Параметр сглаживания для этих двух сплайнов был одинаков и равен $\alpha = \alpha_{opt}$, где α_{opt} – оценка оптимального параметра сглаживания (подробнее см. [6]). Видно, что второй СКС содержит осцилляции, вызванные АИ, которые дадут существенные ошибки при вычислении первой производной и, соответственно, большую ошибку в решении уравнения (4). Неудовлетворительное сглаживание АИ обусловлено вторым слагаемым $\sum_{i=1}^n p_i^{-1} (\varphi_i - S(t_i))^2$ функционала (7) (слагаемое называется функционалом невязки). Наличие квадрата разности $(\varphi_i - S(t_i))^2$ «заставляет» сплайн

«приближаться» к аномальному измерению и тем самым не удается эффективно сгладить данное АИ.

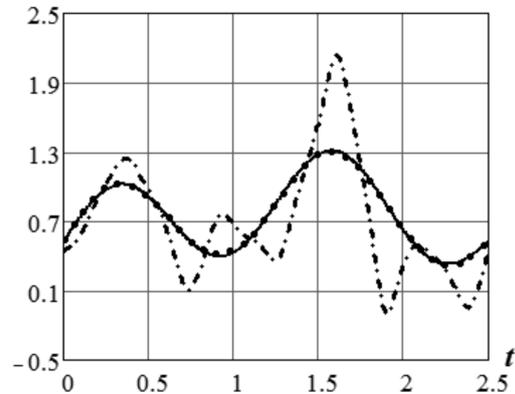


Рис. 2. Значения сглаживающих сплайнов

Для того чтобы сделать алгоритм идентификации нечувствительным (т.е. робастным) к АИ, возможны два подхода:

– заменить квадратичный функционал невязки в (7) на другой, аналогичный функционалу, используемому в методах робастного оценивания коэффициентов уравнения регрессии при наличии АИ (например, на сумму модулей);

– предварительно удалить (отфильтровать) АИ, а затем сгладить оставшийся шум сглаживающим кубическим сплайном.

Первый подход, когда строится робастный сглаживающий сплайн, достаточно сложен и трудоемок [8, 9]. Поэтому в данной работе предлагается робастный алгоритм идентификации, являющийся реализацией второго подхода.

Результат исследования. Робастный двухшаговый алгоритм идентификации

Предположим, что зашумленные значения входного и выходного сигналов допускают следующие представления:

$$\tilde{\varphi}_i = \varphi(t_i) + \xi_i; \tilde{f}_i = f(t_i) + \eta_i, i = 1, \dots, N, \quad (8)$$

где ξ_i, η_i – шумы измерения. Для статистического описания АИ введем следующую модель шумов ξ_i, η_i . Шум измерения входного сигнала ξ_i с вероятностью $1 - p_\xi$ имеет нулевое среднее и дисперсию σ_ξ^2 , а с вероятностью p_ξ – нулевое среднее и дисперсию $C_\xi \sigma_\xi^2$, где $C_\xi \gg 1$, т.е. с вероятностью p_ξ появляется импульсная составляющая шума измерения. Если эта вероятность равна нулю, то такой шум (без импульсной составляющей) будем называть однород-

ным. Шум измерения выходного сигнала η_i с вероятностью $1 - p_\eta$ имеет нулевое среднее и дисперсию σ_η^2 , а с вероятностью p_η – нулевое среднее и дисперсию $C_\eta \sigma_\eta^2$, где $C_\eta \gg 1$. На рис. 1 показаны АИ при $p_\xi = 0,02$ и $C_\xi = 400$.

$$\hat{\phi}_j^{KF} = \text{aver}_K(\hat{\phi}_i^{MF} : j - K \leq i \leq j + K, |\hat{\phi}_i^{MF} - \hat{\phi}_j^{MF}| \leq \Delta_\phi), \quad (9)$$

где $\hat{\phi}_i^{MF}$ – результат фильтрации медианным фильтром (МФ), определяемый выражением

$$\hat{\phi}_j^{MF} = \text{med}_L(\tilde{\phi}_{j-L}, \tilde{\phi}_{j-L+1}, \dots, \tilde{\phi}_j, \dots, \tilde{\phi}_{j+L}), \quad (10)$$

где med_L – функция, вычисляющая медиану из $2L + 1$ значений, попавших в апертуру фильтра (указаны в скобках). Заметим, что МФ хорошо фильтрует импульсные шумы и сохраняет контрастные составляющие в отфильтрованном сигнале. Усреднение в КФ (сглаживание оставшихся однородных шумов) происходит только для значений $\hat{\phi}_i^{MF}$ из интервала $[\hat{\phi}_i^{MF} - \Delta_\phi, \hat{\phi}_i^{MF} + \Delta_\phi]$, что предотвращает сглаживание контрастных составляющих точного сигнала. Таким образом, КФ устраняет импульсные шумы и успешно сглаживает однородные остаточные шумы, при этом размеры апертур удовлетворяют неравенству $K \geq L$. Если величина $K = 0$, то КФ превращается в МФ. Если величина $L = 0$, то КФ превращается в интервальный фильтр скользящего среднего. Величину Δ_ϕ можно определить соотношением: $\Delta_\phi = 2 \cdot \sigma_\xi^2$. Аналогичный КФ используется для фильтрации выходного сигнала.

На втором этапе по отфильтрованным значениям входного $\{\hat{\phi}_i^{KF}\}$ и выходного $\{f_i^{KF}\}$ сигналов строятся сглаживающие кубические сплайны $S_{\phi,\alpha}(t), S_{f,\alpha}(t)$ с производными $S'_{\phi,\alpha}(t), S'_{f,\alpha}(t)$ соответственно. Параметр сглаживания α выбирался на основе критерия оптимальности, приведенного в работе [6] и подробно рассмотренного в работе [5]. Этот критерий позволяет вычислить оценку α_w для оптимального параметра сглаживания α_{opt} , который минимизирует СКО сглаживания. После построения сплайнов и вычисления их первых производных формируется матрица Φ размером $(N-1) \times (N-1)$, элементы которой определяются как: при $i < j, \Phi_{i,j} = 0$; при $i \geq j, \Phi_{j,i} = \int_{t_k}^{t_{k+1}} S'_{\phi,\alpha}(\tau) d\tau = b_k + 2c_k(t_{k+1} - t_k) + 3d_k(t_{k+1} - t_k)^2$, где $k = j - i$. Сформированная матрица Φ позволяет аппроксимировать уравнение (4) системой ли-

нейных алгебраических уравнений (СЛАУ) вида (подробнее в работе [6]):

$$\left(I + \frac{1}{S_{\phi,\alpha}(t_1)} \Phi \right) \hat{k} = \frac{1}{S_{\phi,\alpha}(t_1)} \hat{f}', \quad (11)$$

где I – единичная матрица размером $(N-1) \times (N-1)$, а вектор $\hat{f}' = \begin{bmatrix} S'_{f,\alpha}(t_1) \\ \vdots \\ S'_{f,\alpha}(t_{N-1}) \end{bmatrix}$ составлен из значений первой производной выходного сигнала в узлах сетки. Решением этой системы будет вектор \hat{k} , проекции которого являются оценками для значений $k(t_i)$ ИПФ идентифицируемой системы.

Результаты численных исследований

Приведем некоторые результаты вычислительного эксперимента, выполненного для исследования точности идентификации предложенным алгоритмом. В качестве тестовой ИПФ была взята ИПФ, график которой показан сплошной кривой на рис. 3 (соответствует колебательному звену второго порядка). Значения выходного и входного сигналов искажались погрешностями, подчиняющимися нормальному распределению с относительными

уровнями $\delta_f = \frac{\|\tilde{f} - f\|}{\|f\|}, \delta_\phi = \frac{\|\tilde{\phi} - \phi\|}{\|\phi\|}$, где

$\|\cdot\|$ – евклидова норма вектора, f, ϕ – векторы, составленные из точных значений выходного и входного сигналов, $\tilde{f}, \tilde{\phi}$ – векторы зашумленных значений. В отсутствии АИ $\delta_\phi = 0,10, \delta_f = 0,10$, при наличии аномальных измерений ($p_\xi = 0,02, C_\xi = 400, p_\eta = 0,02, C_\eta = 400$) – $\delta_\phi = 0,358, \delta_f = 0,421$. На рис. 3 штрих-точечной кривой показаны значения ИПФ, вычисленные алгоритмом работы [6], когда сглаживающие сплайны строились по АИ и имели из-за этого большие ошибки сглаживания (см. штрих-точечную кривую рис. 2). После сглаживания СКС имеем относительные уровни ошибок сглаживания $\delta_\phi = 0,311, \delta_f = 0,591$.

В этом случае относительная ошибка идентификации $\delta_k = \frac{\|\hat{k} - k\|}{\|k\|}$ была равна 1,258,

которая говорит об очень плохой точности идентификации.

Приведем некоторые количественные характеристики предлагаемого двухступенчатого алгоритма идентификации. После выполнения на первом шаге фильтрации с использованием КФ относительные уровни остаточных шумов были равны: $\delta_\phi = 0,052$, $\delta_f = 0,061$. Сравнивая с уровнями шумов исходных сигналов ($\delta_\phi = 0,358$, $\delta_f = 0,421$), видим существенное уменьшение уровней шума и удаление импульсных составляющих. После сглаживания СКС остаточного шума имеем: $\delta_\phi = 0,038$, $\delta_f = 0,049$. На рис. 3 точечной кривой показаны значения ИПФ, вычисленные двухшаговым алгоритмом идентификации, при этом относительная ошибка идентификации составила $\delta_k = 0,138$. Заметим, что относительная ошибка идентификации алгоритмом работы [6] при отсутствии аномальных измерений была равна 0,131.

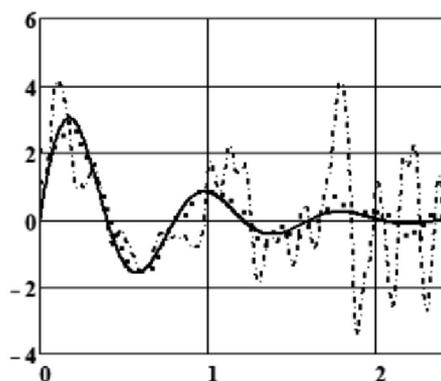


Рис. 3. Импульсная переходная функция и ее оценки

Заключение

Предложенный двухшаговый алгоритм идентификации позволяет решать задачу идентификации с приемлемой точностью

даже при наличии в измеряемых сигналах идентифицируемой системы импульсного шума. Отчасти это обусловлено тем, что матрица СЛАУ хорошо обусловлена (число обусловленности зависит от полосы спектра мощности входного сигнала и чаще всего не превышает 1000), и поэтому ошибки, возникающие при дифференцировании выходного сигнала, оказывают существенно меньшее влияние на точность построенного решения по сравнению с ошибками правой части уравнения (3), решаемого с использованием регуляризирующего алгоритма на основе дискретного преобразования Фурье [5]. Следует отметить, что общие затраты машинного времени на построение решения предложенным алгоритмом возрастают (из-за пространственно-локальной фильтрации) незначительно и это увеличение составляет примерно 25–30%.

Список литературы

1. Сидоров Д.Н. Методы анализа интегральных динамических моделей: теория и приложения. Иркутск: Изд-во ИГУ, 2013. 293 с.
2. Бойков И.В. Аналитические и численные методы идентификации динамических систем. Пенза: ПГУ, 2016. 396 с.
3. Первушин В.Ф. О непараметрической идентификации линейных динамических объектов // Вестник Томского государственного университета. Сер. Управление, вычислительная техника и информатика. 2013. № 4 (25). С. 95–104.
4. Тихонов А.Н. Методы решения некорректных задач. М.: Наука, 1986. 285 с.
5. Воскобойников Ю.Е. Математическая обработка эксперимента в молекулярной газодинамике. Новосибирск: Наука, 1984. 238 с.
6. Воскобойников Ю.Е., Боева В.А. Новый устойчивый алгоритм непараметрической идентификации технических систем // Современные наукоемкие технологии. 2019. № 5. С. 25–29.
7. Wang Y. Smoothing Splines Methods and Applications. Ser. Monographs on Statistics and Applied Probability v. 121. A Chapman & Hall book. 2011. 347 p.
8. Wang B., Shi W., Miao Z. Comparative Analysis for Robust Penalized Spline Smoothing Methods. Mathematical Problems in Engineering. 2014. № 4. P. 3–9.
9. Zhang H. A Robust Spline Filter Algorithm Based on M-Estimate Theory. Advanced Materials Research. 2013. V. 655–657. P. 909–912.
10. Воскобойников Ю.Е. Метод устранения артефактов вейвлет-фильтрации // Фундаментальные исследования. 2017. № 8. С. 246–250.