УДК 517.9

ПРИМЕНЕНИЕ МЕТОДА ИНСТРУМЕНТАЛЬНЫХ ПЕРЕМЕННЫХ ДЛЯ ПАРАМЕТРИЧЕСКОЙ ИДЕНТИФИКАЦИИ РАСПРЕДЕЛЕННОЙ ДИНАМИЧЕСКОЙ СИСТЕМЫ

А. В. Копытин, Е. А. Копытина

Воронежский государственный университет

Поступила в редакцию 22.11.2018 г.

Аннотация. Предложен вариант метода инструментальных переменных для параметрической идентификации уравнений математической физики, описывающих динамику пространственно-распределенных процессов, на основе экспериментальных многомерных временных рядов. Проведенный вычислительный эксперимент показывает значительное улучшение качества оценок параметров уравнения по сравнению с методом наименьших квадратов.

Ключевые слова: уравнения в частных производных; параметрическая идентификация; многомерная авторегрессия; МНК; метод инструментальных переменных.

Annotation. A variant of the method of instrumental variables for the parametric identification of equations of mathematical physics describing the dynamics of spatially-distributed processes on the basis of experimental multidimensional time series is proposed. The computational experiment performed shows a significant improvement in the quality of estimates of the parameters of the equation in comparison with the ordinary least squares method.

Keywords: partial differential equations; parametric identification; multidimensional autoregression; OLS; instrumental variables method.

ВВЕДЕНИЕ

Хорошо известна проблема параметрической идентификации моделей динамических систем. Исследованию этой проблемы посвящено много работ (см., например., [1–10]). При этом использование известных подходов ограничено рядом условий: часть подходов основана на обработке статистики, полученной в результате активного эксперимента, который невозможно провести для некоторых объектов исследования. Многие работы ограничены рамками моделей, описываемых обыкновенными дифференциальными уравнениями.

Продолжая исследования, начатые в [11–13], будем рассматривать широкий класс пространственно-распределенных динамических систем, для которых характерны диффузионные процессы, процессы адвекции или их сочетание. Соответствующее дифференциальное уравнение в частных производных с начальным и граничными условиями имеет следующий общий вид:

$$\frac{\partial y}{\partial t} + v\frac{\partial y}{\partial l} = D\frac{\partial^2 y}{\partial l^2}, \tag{1}$$
$$y(0,l) = \phi(l),$$
$$y(t,l^{\min}) = f_1(t), \quad y(t,l^{\max}) = f_2(t),$$

где v – скорость адвекции, D – коэффициент диффузии, l – пространственная координата.

Источником информации о поведении системы являются данные натурных измерений переменной y_i^k с погрешностью ε_i^k в виде «белого шума» – $x_i^k = y_i^k + \varepsilon_i^k$ в последовательные моменты времени $\{t_k\}_{k=0}^n$ в узлах одномерной пространственной регулярной сетки $\{l_i\}_{i=0}^m$, т. е. многомерный временной ряд. Рассмотрение одномерной сетки ничем не ограничивает дальнейшие исследования, зато позволяет избежать громоздких построений, характерных для плоских и объемных пространств.

Задача заключается в верификации процессов конвективной диффузии на основе анализа многомерных временных рядов и разработке алгоритмов параметрической идентификации механистической модели с постоянными коэффициентами по наблюдаемым значениям x_i^k.

МЕТОДЫ И МАТЕРИАЛЫ

Для решения задачи составим явную четырехточечную разностную схему для уравнения (1):

$$\frac{y_i^{k+1}-y_i^k}{\Delta t}+v\frac{y_{i+1}^k-y_{i-1}^k}{2\Delta l}=D\frac{y_{i+1}^k-2y_i^k+y_{i-1}^k}{\Delta l^2},$$

$$y_i^{k+1}=(b_1+b_2)y_{i-1}^k+(1-2b_2)y_i^k+(b_2-b_1)y_{i-1}^k, \quad (2)$$

где $b_1=\dfrac{v\Delta t}{2\Delta l}$, $b_2=\dfrac{D\Delta t}{\Delta l^2}$.

Уравнения (2) можно записать в виде

$$y_i^{k+1}=a_1y_{i-1}^k+a_2y_i^k+a_3y_{i-1}^k,\ k=0,1,\dots,n-1, \quad (3)$$

где a_1, a_2, a_3 – оцениваемые регрессионные параметры, связанные с параметрами b_1 и b_2 следующими соотношениями:

$$\begin{cases} b_1+b_2=a_1, \\ 1-2b_2=a_{2,} \\ b_2-b_1=a_{3.} \end{cases} \quad (4)$$

Запишем уравнения (3) при $k=1,\dots,n-1$ в матричной форме

$$\mathbf{y}=\mathbf{Ya}, \quad (5)$$

где

$$\mathbf{y}=\begin{pmatrix} y_i^2 \\ y_i^3 \\ \vdots \\ y_i^n \end{pmatrix},\quad \mathbf{Y}=\begin{pmatrix} y_{i-1}^1 & y_i^1 & y_{i+1}^1 \\ y_{i-1}^2 & y_i^2 & y_{i+1}^2 \\ \vdots & \vdots & \vdots \\ y_{i-1}^{n-1} & y_i^{n-1} & y_{i+1}^{n-1} \end{pmatrix},\quad \mathbf{a}=\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}.$$

Обозначим $\mathbf{x}=\mathbf{y}+\boldsymbol{\varepsilon}$, $\mathbf{X}=\mathbf{Y}+\mathbf{E}$, где $\boldsymbol{\varepsilon}$ и $\mathbf{E}$ – соответственно вектор-столбец и матрица ошибок измерений:

$$\boldsymbol{\varepsilon}=\begin{pmatrix} \varepsilon_i^2 \\ \varepsilon_i^3 \\ \vdots \\ \varepsilon_i^n \end{pmatrix},\quad \mathbf{E}=\begin{pmatrix} \varepsilon_{i-1}^1 & \varepsilon_i^1 & \varepsilon_{i+1}^1 \\ \varepsilon_{i-1}^2 & \varepsilon_i^2 & \varepsilon_{i+1}^2 \\ \vdots & \vdots & \vdots \\ \varepsilon_{i-1}^{n-1} & \varepsilon_i^{n-1} & \varepsilon_{i+1}^{n-1} \end{pmatrix}.$$

В новых обозначениях уравнение (5) примет вид

$$\mathbf{x}=\mathbf{Xa}+(\boldsymbol{\varepsilon}-\mathbf{Ea}). \quad (6)$$

Уравнение (6) позволяет найти обычную МНК-оценку $\hat{\mathbf{a}}$ вектора параметров $\mathbf{a}$:

$$\hat{\mathbf{a}}=(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{x}=\mathbf{a}+(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T(\boldsymbol{\varepsilon}-\mathbf{Ea}). \quad (7)$$

Найдем математическое ожидание левой и правой частей выражения (7):

$$M(\hat{\mathbf{a}})=\mathbf{a}+M[(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T(\boldsymbol{\varepsilon}-\mathbf{Ea})]. \quad (8)$$

Как видно из полученного равенства (8), математическое ожидание оценки $\hat{\mathbf{a}}$ отличается от истинного вектора $\mathbf{a}$ на величину $M[(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T(\boldsymbol{\varepsilon}-\mathbf{Ea})]\neq 0$, которая интерпретируется как смещение компонент вектора оценки.

Наличие смещения может существенно повлиять на оценку параметров b_1 и b_2, определяемых на основе системы уравнений (4). Эту проблему можно уменьшить за счет использования инструментальных переменных [14, 15].

Метод инструментальных переменных предполагает наличие набора переменных $\mathbf{Z}$, называемых инструментами. Инструменты должны быть некоррелированны с ошибкой $\boldsymbol{\varepsilon}-\mathbf{Ea}$ и, напротив, как можно сильнее коррелированны с регрессорами $\mathbf{X}$. Количество инструментов должно быть не меньше количества регрессоров.

Как только инструменты выбраны, можно применить двухшаговый МНК для оценивания вектора $\mathbf{a}$. На первом шаге находятся обычные МНК-оценки $\hat{\mathbf{B}}=(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{X}$ матрицы параметров $\mathbf{B}$ уравнения регрессии $\mathbf{X}=\mathbf{ZB}+\mathbf{V}$. В результате получаем следующие оценки исходных переменных $\mathbf{X}$:

$$\hat{\mathbf{X}}=\mathbf{Z}\hat{\mathbf{B}}=\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{X}=\mathbf{P_Z X}, \quad (9)$$

где $\mathbf{P_Z}=\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T$.

На втором шаге также обычным МНК оценивается исходная модель (6) с заменой регрессоров $\mathbf{X}$ на их оценки (9), полученные на первом шаге:

$$\hat{\mathbf{a}}_{IV}=(\hat{\mathbf{X}}^T\hat{\mathbf{X}})^{-1}\hat{\mathbf{X}}^T\mathbf{x}=(\mathbf{X}^T\mathbf{P_Z}^T\mathbf{P_Z}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{P_Z}^T\mathbf{x}.$$

Учитывая, что $\mathbf{P_Z}^T=\mathbf{P_Z}$, $\mathbf{P_Z}^T\mathbf{P_Z}=\mathbf{P_Z}$, окончательно получаем формулу оценок метода инструментальных переменных:

$$\hat{\mathbf{a}}_{IV}=(\mathbf{X}^T\mathbf{P_Z X})^{-1}\mathbf{X}^T\mathbf{P_Z x},\quad \mathbf{P_Z}=\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T. \quad (10)$$

В нашем случае в качестве инструментов предлагается выбрать набор из пяти лагированных переменных:

$$\mathbf{Z} = \begin{pmatrix} x_{i-2}^0 & x_{i-1}^0 & x_i^0 & x_{i+1}^0 & x_{i+2}^0 \\ x_{i-2}^1 & x_{i-1}^1 & x_i^1 & x_{i+1}^1 & x_{i+2}^1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{i-2}^{n-2} & x_{i-1}^{n-2} & x_i^{n-2} & x_{i+1}^{n-2} & x_{i+2}^{n-2} \end{pmatrix}.$$

Выбранные таким образом инструменты будут с одной стороны некоррелированы с ошибкой $\boldsymbol{\varepsilon} - \mathbf{Ea}$, поскольку погрешности измерений ε_i^k в разные моменты времени независимы, а с другой – линейно выражать регрессоры $\mathbf{X}$.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Для проведения исследований удобно воспользоваться данными модельного эксперимента. Для этого необходимо найти решение исходного дифференциального уравнения (1) с заданными значениями параметров v и D, которые легко пересчитываются в параметры b_1 и b_2 разностной схемы. Затем выполнить регулярную дискретизацию полученного решения и добавить погрешность ε в виде «белого шума» с различной интенсивностью. Полученные статистические данные будут использованы для получения оценок $\hat{b}_1$ и $\hat{b}_2$ параметров разностной схемы. Таким образом, модельный эксперимент позволяет провести наглядное сравнение исходных значений параметров и их оценок при различных методах получения оценок и различных интенсивностях помехи.

Рассмотрим задачу (1) на отрезке [1; 3] с такими функциями ϕ, f_1 и f_2 что ее решение имеет вид:

$$y(t,l) = \\ = e^{\frac{v}{2D}\left(l - \frac{vt}{2}\right)} \left(e^{-Dt} \sin l + e^{-4Dt} \sin 2l + e^{-9Dt} \sin 3l\right). \tag{11}$$

Пусть значения параметров v и D равны соответственно 2 и 3; шаг по пространственной координате $\Delta l = 0{,}1$; шаг по времени $\Delta t = \Delta l^2 / (4D)$, что соответствует условиям Куранта для обеспечения устойчивости аппроксимирующей разностной схемы; $n = 1001$.

Далее к значениям y_i^k решения (11) в узлах пространственно-временной сетки добавим смоделированную с помощью генератора случайных чисел погрешность в виде «белого шума» со стандартным отклонением σ и по полученным значениям x_i^k найдем сначала МНК-оценки $\hat{\mathbf{a}}$ параметров уравнения (6), оценки этих параметров с помощью метода инструментальных переменных и оценки $\hat{b}_1$ и $\hat{b}_2$ с помощью системы (4) и предложенной методики. Очевидно, что результаты параметрической идентификации могут существенно зависеть от интенсивности «белого шума», задаваемой стандартным отклонением случайной погрешности наблюдений σ. Для проведения эксперимента были выбраны три уровня погрешности: $\sigma = 0{,}001$; $\sigma = 0{,}005$; $\sigma = 0{,}01$. Результаты идентификации представлены в табл. 1.

Данные табл. 1 показывают, что при всех трех уровнях погрешности средняя абсолютная ошибка в процентах метода инструментальных переменных значительно меньше, чем ошибка МНК, хотя и является достаточно большой.

Таблица 1

Средняя абсолютная ошибка в процентах (MAPE)

		σ		
		0,001	0,005	0,01
МНК				
	v	7,934	97,825	150,964
	D	1,018	16,222	26,048
Метод инструментальных переменных				
	v	4,864	25,468	56,656
	D	0,536	2,697	6,286

Таблица 2

Средние значения оценок параметров

		σ		
		0,001	0,005	0,01
МНК				
	v	2,144	3,956	5,013
	D	3,029	3,487	3,781
Метод инструментальных переменных				
	v	1,999	2,047	2,160
	D	2,993	2,991	3,022

Модельный эксперимент по каждому методу получения оценок и при каждом уровне интенсивности помехи повторялся 500 раз, что позволило получить среднеарифметические значения оценок параметров v и D. Средние значения оценок можно рассматривать как хорошее приближение к их математическому ожиданию. В этом случае разницу между средним значением оценки и ее истинным значением можно принять за величину смещения. Результаты сравнения средних значений исследуемых оценок при различных интенсивностях помех получаются на основании данных табл. 2.

Сравнение средних значений с истинными значениями параметров показывает, что величина смещения оценок МНК существенна даже при небольших погрешностях и возрастает с ростом погрешности. При этом смещение оценок метода инструментальных переменных на порядки меньше и остается приемлемым даже при самом высоком уровне погрешности, что подтверждает целесообразность применения данного метода.

ЗАКЛЮЧЕНИЕ

Результаты проведенных экспериментов показывают, что использование для оценки параметров дифференциального уравнения (1) метода наименьших квадратов может привести к существенным искажениям истинных значений параметров в условиях высокого уровня погрешностей наблюдений за многомерными временными рядами в узлах разностной схемы.

Предложенный вариант метода инструментальных переменных, как показывают результаты табл. 1 и 2, существенно улучшает качество оценки.

СПИСОК ЛИТЕРАТУРЫ

1. A Bayesian approach to parameter estimation in HIV dynamical models / H. Putter [et al.] // Statistics in Medicine. – 2002. – Vol. 21. – P. 2199–2214.

2. *Huang, Y.* Hierarchical Bayesian methods for estimation of parameters in a longitudinal HIV dynamic system / Y. Huang, D. Liu, H. Wu // Biometrics. – 2006. – Vol. 62. – P. 413–423.

3. *Huang, Y.* A Bayesian approach for estimating antiviral efficacy in HIV dynamic models / Y. Huang, H. Wu // Journal of Applied Statistics. – 2006. – Vol. 33. – P. 155–174.

4. Parameter estimation for differential equations: a generalized smoothing approach (with discussion) / J. O. Ramsay [и др.] // Journal of the Royal Statistical Society. Series B. – 2007. – Vol. 69. – P. 741–796.

5. *Liang, H.* Parameter estimation for differential equation models using a framework of measurement error in regression models / H. Liang, H. Wu // Journal of the American Statistical Association. – 2008. – Vol. 103. – P. 1570–1583.

6. *Chen, J.* Efficient local estimation for time-varying coefficients in deterministic dynamic models with applications to HIV-1 dynamics /

J. Chen, H. Wu // Journal of the American Statistical Association. – 2008. – Vol. 103. – P. 369–384.

7. *Cao, J.* Penalized nonlinear least squares estimation of time-varying parameters in ordinary differential equations / J. Cao, J. Z. Huang, H. Wu // Journal of Computational and Graphical Statistics. – 2012. – Vol. 21. – P. 42–56.

8. *Muller, T.* Fitting parameters in partial differential equations from partially observed noisy data / T. Muller, J. Timmer // Physical Review, D. – 2002. – Vol. 171. P. 1–7.

9. *Muller, T.* Parameter identification techniques for partial differential equations / T. Muller, J. Timmer // International Journal of Bifurcation and Chaos. – 2004. – Vol. 14. – P. 2053–2060.

10. Parameter estimation of partial differential equation models / X. Xun [et al.] // Journal of the American Statistical Association. – 2013. – Vol. 108. – P. 1009–1020.

11. Modeling of nonstationary distributed processes on the basis of multidimensional time series / M. G. Matveev [et al.] // Procedia Engineering. – 2017. – Vol. 201. – P. 511–516.

12. Verification of the convective diffusion process based on the analysis of multidimensional time series / M. G. Matveev [et al.] // CEUR Workshop Proceedings. – 2017. – Vol. 2022. – P. 354–358.

13. *Копытин, А. В.* Применение расширенного фильтра Калмана для идентификации параметров распределенной динамической системы / А. В. Копытин, Е. А. Копытина, М. Г. Матвеев // Вестник Воронеж. гос. ун-та. Сер. Системный анализ и информационные технологии. – Воронеж, 2018. – № 3. – С. 44–50.

14. *Bowden, R. J.* Instrumental variables / R. J. Bowden, D. A. Turkington. – New York : Cambridge University Press, 1984. – 227 p.

15. *White, H.* Asymptotic theory for econometricians / H. White. – New York : Academic Press, 2001. – 264 p.

Копытин А. В. – к. ф.-м. н., доцент кафедры информационных технологий управления, факультет компьютерных наук, Воронежский государственный университет.
E-mail: alexkopytin@gmail.com

Копытина Е. А. – ассистент кафедры информационных технологий управления, факультет компьютерных наук, Воронежский государственный университет.
E-mail: zhemkaterina@yandex.ru

Kopytin A. V. – Ph. D. of Physical and Mathematical Sciences, Associate Professor, Department of Information Technologies in Management, Computer Sciences Faculty, Voronezh State University.
E-mail: alexkopytin@gmail.com

Kopytina E. A. – Postgraduate Student, Department of Information Technologies in Management, Computer Sciences Faculty, Voronezh State University.
E-mail: zhemkaterina@yandex.ru