

Комбинированный метод идентификации параметров распределенной динамической модели

М.Г. Матвеев¹, А.В. Копытин¹, Е.А. Сирота¹

¹Воронежский государственный университет, Университетская площадь 1, Воронеж, Россия, 394018

Аннотация. Предложен комбинированный метод идентификации уравнений математической физики, описывающих динамику пространственно-распределенных процессов, на основе экспериментальных многомерных временных рядов. Первым этапом метода является получение МНК-оценок параметров многомерной авторегрессии. Однако, эти оценки оказываются смещенными в силу наличия ошибок в регрессорах. С целью уменьшения указанного смещения в качестве второй составляющей метода применяется расширенный фильтр Калмана. Приводится вычислительный эксперимент, подтверждающий эффективность предложенной методики.

1. Введение и постановка задачи

Продолжая исследования, начатые в [1-2], 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 начальными и граничными условиями имеет следующий общий вид:

$$\begin{aligned} \frac{\partial x}{\partial t} + v \frac{\partial x}{\partial l} &= D \frac{\partial^2 x}{\partial l^2}, \\ x(0, l) &= \varphi(l), \\ x(t, l^{\min}) &= f_1(t), \quad x(t, l^{\max}) = f_2(t), \end{aligned} \tag{1}$$

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

Источником информации о поведении системы являются данные натурных измерений переменной x_i^t с погрешностью ε_i^t в виде «белого шума» – $y_i^t = x_i^t + \varepsilon_i^t$ в последовательные моменты времени $t=0,1,\dots,n$ в узлах одномерной пространственной регулярной сетки $i=0,1,\dots,m$, 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 y_i^t .

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

$$\frac{x_i^{t+1} - x_i^t}{\Delta t} + \nu \frac{x_{i+1}^t - x_{i-1}^t}{2\Delta l} = D \frac{x_{i+1}^t - 2x_i^t + x_{i-1}^t}{\Delta l^2},$$

$$x_i^{t+1} = (b_1 + b_2)x_{i-1}^t + (1 - 2b_2)x_i^t + (b_2 - b_1)x_{i+1}^t; \quad b_1 = \frac{\nu\Delta t}{2\Delta l}; \quad b_2 = \frac{D\Delta t}{\Delta l^2}. \tag{2}$$

Уравнения (2) можно записать в виде

$$x_i^{t+1} = a_1 x_{i-1}^t + a_2 x_i^t + a_3 x_{i+1}^t, \quad t = 0, 1, \dots, n, \tag{3}$$

где a_i – оцениваемые регрессионные параметры, связанные с параметрами b_1 и b_2 следующими соотношениями

$$\begin{cases} b_1 + b_2 = a_1, \\ 1 - 2b_2 = a_2, \\ b_2 - b_1 = a_3. \end{cases} \tag{4}$$

В матричной форме выражение (3) принимает вид

$$\mathbf{x} = \mathbf{X}\mathbf{a}, \tag{5}$$

где

$$\mathbf{x} = \begin{pmatrix} x_i^1 \\ x_i^2 \\ \vdots \\ x_i^n \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} x_{i-1}^0 & x_i^0 & x_{i+1}^0 \\ x_{i-1}^1 & x_i^1 & x_{i+1}^1 \\ \vdots & \vdots & \vdots \\ x_{i-1}^{n-1} & x_i^{n-1} & x_{i+1}^{n-1} \end{pmatrix}, \quad \mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}.$$

Обозначим $\mathbf{y} = \mathbf{x} + \boldsymbol{\delta}$, $\mathbf{Y} = \mathbf{X} + \boldsymbol{\varepsilon}$, где $\boldsymbol{\delta}$ и $\boldsymbol{\varepsilon}$ – соответственно вектор и матрица ошибок измерений, т.е.

$$\boldsymbol{\delta} = \begin{pmatrix} \varepsilon_i^1 \\ \varepsilon_i^2 \\ \vdots \\ \varepsilon_i^n \end{pmatrix}, \quad \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_{i-1}^0 & \varepsilon_i^0 & \varepsilon_{i+1}^0 \\ \varepsilon_{i-1}^1 & \varepsilon_i^1 & \varepsilon_{i+1}^1 \\ \vdots & \vdots & \vdots \\ \varepsilon_{i-1}^{n-1} & \varepsilon_i^{n-1} & \varepsilon_{i+1}^{n-1} \end{pmatrix}.$$

В новых обозначениях уравнение (5) имеет вид

$$\mathbf{y} = \mathbf{Y}\mathbf{a} + (\boldsymbol{\delta} - \boldsymbol{\varepsilon}\mathbf{a}). \tag{6}$$

Исходя из представления (6), можно найти обычную МНК-оценку $\hat{\mathbf{a}}$ вектора параметров \mathbf{a}

$$\hat{\mathbf{a}} = (\mathbf{Y}^T \mathbf{Y})^{-1} \mathbf{Y}^T \mathbf{y} = \mathbf{a} + (\mathbf{Y}^T \mathbf{Y})^{-1} \mathbf{Y}^T (\boldsymbol{\delta} - \boldsymbol{\varepsilon}\mathbf{a}). \tag{7}$$

Однако, как показано, например, в [3], оценка (7) не является ни состоятельной, ни несмещённой. Следовательно, несостоятельной и смещённой является оценка $\hat{\mathbf{b}} = (\hat{a}_1 - \hat{a}_3, \hat{a}_1 + \hat{a}_3)^T / 2$ вектора параметров $\mathbf{b} = (b_1, b_2)^T$ уравнений (2).

Повышения точности решения задачи верификации процессов уравнения (1) с использованием МНК-оценок параметров авторегрессии (3) можно достичь на основе комбинирования методов идентификации. Такой подход получил широкое распространение в различных предметных областях (см., например, [4-6]). Комбинирование дает возможность компенсировать недостатки одних методов при помощи других и направлено на повышение качества параметрической идентификации, как одного из главных критериев эффективности модели. Мы предлагаем для решения задачи верификации использовать комбинацию МНК-оценок и фильтра Калмана. МНК-оценки дают начальную точку в пространстве поиска для реализации рекурсивного алгоритма Калмана, что очевидно лучше, чем произвольный выбор такой точки. Исследованию эффективности комбинации МНК-оценок и фильтра Калмана посвящена предлагаемая работа.

2. Фильтр Калмана

Как известно фильтр Калмана представляет собой рекурсивный алгоритм оптимального оценивания неизвестного состояния линейной динамической системы по зашумленным измерениям в дискретные моменты времени. Если система оказывается нелинейной, обычно используется процедура линеаризации. Полученный таким образом фильтр называется расширенным фильтром Калмана.

Рассматриваемая нами линейная динамическая модель может быть сформулирована следующим образом

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}(\mathbf{b})\mathbf{x}_k + \mathbf{u}_k + \boldsymbol{\xi}_k, \\ \mathbf{y}_k = \mathbf{H}\mathbf{x}_k + \boldsymbol{\eta}_k, \end{cases} \tag{8}$$

где $\mathbf{x}_k = (x_{i-2}^k, x_{i-1}^k, x_i^k, x_{i+1}^k, x_{i+2}^k)^T$ – вектор состояния, $\mathbf{A}(\mathbf{b})$ – матрица эволюции процесса, зависящая от постоянного вектора $\mathbf{b} = (b_1, b_2)^T$, $\mathbf{u}_k = (y_{i-2}^{k+1}, 0, 0, 0, y_{i+2}^{k+1})^T$ интерпретируется в терминах фильтра Калмана как вектор управляющих воздействий, $\boldsymbol{\xi}_k = (\varepsilon_{i-2}^{k+1}, 0, 0, 0, \varepsilon_{i+2}^{k+1})^T$ – шум процесса, $\mathbf{y}_k = (y_{i-1}^k, y_i^k, y_{i+1}^k)^T$ – вектор измерений, \mathbf{H} – матрица измерений, $\boldsymbol{\eta}_k = (\varepsilon_{i-1}^k, \varepsilon_i^k, \varepsilon_{i+1}^k)^T$ – измерительный шум.

$$\mathbf{A}(\mathbf{b}) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ b_1 + b_2 & 1 - 2b_2 & b_2 - b_1 & 0 & 0 \\ 0 & b_1 + b_2 & 1 - 2b_2 & b_2 - b_1 & 0 \\ 0 & 0 & b_1 + b_2 & 1 - 2b_2 & b_2 - b_1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

Поскольку \mathbf{b} – постоянный вектор, вполне естественно положить

$$\mathbf{b}_{k+1} = \mathbf{b}_k. \tag{9}$$

Тогда система (8) вместе с предположением (9) может быть переформулирована как нелинейная модель

$$\begin{cases} \begin{pmatrix} \mathbf{x}_{k+1} \\ \mathbf{b}_{k+1} \end{pmatrix} = \begin{pmatrix} \mathbf{A}(\mathbf{b}_k)\mathbf{x}_k \\ \mathbf{b}_k \end{pmatrix} + \begin{pmatrix} \mathbf{u}_k \\ 0 \end{pmatrix} + \begin{pmatrix} \boldsymbol{\xi}_k \\ 0 \end{pmatrix}, \\ \mathbf{y}_k = (\mathbf{H} \quad 0) \begin{pmatrix} \mathbf{x}_k \\ \mathbf{b}_k \end{pmatrix} + \boldsymbol{\eta}_k, \end{cases} \tag{10}$$

и расширенный фильтр Калмана может быть применен для оценки вектора состояния, содержащего \mathbf{b}_k в качестве своих компонент. Однако, для этого нам необходимо иметь оценки

матриц ковариации $\mathbf{Q} = \text{cov}(\boldsymbol{\xi}_k)$ и $\mathbf{R} = \text{cov}(\boldsymbol{\eta}_k)$, начальную оценку вектора состояния $\begin{pmatrix} \hat{\mathbf{x}}_{0|0} \\ \hat{\mathbf{b}}_{0|0} \end{pmatrix}$ и оценку ковариационной матрицы $\mathbf{P}_{0|0}$ начальной оценки вектора состояния.

Пусть $\mathbf{e} = \mathbf{y} - \mathbf{Y}\hat{\mathbf{a}}$ – вектор остатков регрессии (6). В качестве оценки дисперсии измерительной погрешности ε_i^t возьмем величину $v_0 = \mathbf{e}^T \mathbf{e} / (n - 3) / (1 + \hat{\mathbf{a}}^T \hat{\mathbf{a}})$. Тогда

$$\mathbf{Q} = \text{diag}\{v_0, 0, 0, 0, v_0, 0, 0\}, \quad \mathbf{R} = \text{diag}\{v_0, v_0, v_0\},$$

где diag обозначает диагональную матрицу.

В качестве оценки $\hat{\mathbf{x}}_{0|0}$ естественно взять вектор $(y_{i-2}^0, y_{i-1}^0, y_i^0, y_{i+1}^0, y_{i+2}^0)^T$, а в качестве оценки $\hat{\mathbf{b}}_{0|0}$ МНК-оценку $\hat{\mathbf{b}} = (\hat{a}_1 - \hat{a}_3, \hat{a}_1 + \hat{a}_3)^T / 2$. Оценка ковариационной матрицы вектора МНК-оценок $\hat{\mathbf{a}}$ имеет вид $\mathbf{V}_a = (v_{ij}) = \mathbf{e}^T \mathbf{e} / (n - 3)(\mathbf{Y}^T \mathbf{Y})^{-1}$. Тогда в качестве оценки ковариационной матрицы вектора $\hat{\mathbf{b}}$ возьмем следующую матрицу \mathbf{V}_b

$$V_b = \begin{pmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} v_{11} & v_{13} \\ v_{31} & v_{33} \end{pmatrix} \begin{pmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{pmatrix},$$

и матрица $P_{0|0}$ примет вид

$$P_{0|0} = \begin{pmatrix} \text{diag}\{v_0, v_0, v_0, v_0, v_0\} & 0 \\ 0 & V_b \end{pmatrix}.$$

Алгоритм расширенного фильтра Калмана в нашем случае выглядит следующим образом (см. [7])

$$\left\{ \begin{aligned} \begin{pmatrix} \hat{x}_{0|0} \\ \hat{b}_{0|0} \end{pmatrix} &= \begin{pmatrix} (y_{i-2}^0, y_{i-1}^0, y_i^0, y_{i+1}^0, y_{i+2}^0)^T \\ \hat{b} \end{pmatrix}, \quad P_{0|0} = \begin{pmatrix} \text{diag}\{v_0, v_0, v_0, v_0, v_0\} & 0 \\ 0 & V_b \end{pmatrix}, \\ \text{Для } k &= 1, 2, \dots, n, \\ \begin{pmatrix} \hat{x}_{k|k-1} \\ \hat{b}_{k|k-1} \end{pmatrix} &= \begin{pmatrix} A(\hat{b}_{k-1|k-1})\hat{x}_{k-1|k-1} + u_k \\ \hat{b}_{k-1|k-1} \end{pmatrix}, \\ P_{k|k-1} &= \begin{pmatrix} A(\hat{b}_{k-1|k-1}) & \frac{\partial}{\partial b}(A(\hat{b}_{k-1|k-1})\hat{x}_{k-1|k-1}) \\ 0 & I \end{pmatrix} P_{k-1|k-1} \begin{pmatrix} A(\hat{b}_{k-1|k-1}) & \frac{\partial}{\partial b}(A(\hat{b}_{k-1|k-1})\hat{x}_{k-1|k-1}) \\ 0 & I \end{pmatrix}^T \\ &+ \begin{pmatrix} Q & 0 \\ 0 & 0 \end{pmatrix}, \\ K_k &= P_{k|k-1} (H \ 0)^T \left((H \ 0) P_{k|k-1} (H \ 0)^T + R \right)^{-1}, \\ P_{k|k} &= (I - K_k (H \ 0)) P_{k|k-1}, \\ \begin{pmatrix} \hat{x}_{k|k} \\ \hat{b}_{k|k} \end{pmatrix} &= \begin{pmatrix} \hat{x}_{k|k-1} \\ \hat{b}_{k|k-1} \end{pmatrix} + K_k (y_k - H\hat{x}_{k|k-1}). \end{aligned} \right. \tag{11}$$

В качестве итоговой оценки вектора параметров b будем использовать последнее значение оценки $\hat{b}_{n|n}$, полученное при $k = n$.

3. Результаты вычислительного эксперимента

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

$$x(t, l) = \exp\left(\frac{v}{2D}\left(l - \frac{vt}{2}\right)\right) (\exp(-Dt)\sin(l) + \exp(-4Dt)\sin(2l) + \exp(-9Dt)\sin(3l)). \tag{12}$$

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; $n = 1000$.

Далее к значениям x_i^t решения (12) в узлах пространственно-временной сетки добавим смоделированную с помощью генератора случайных чисел погрешность в виде «белого шума» с дисперсией σ и по полученным значениям y_i^t найдем сначала МНК-оценки параметров v и D , затем оценки этих параметров с помощью фильтра Калмана, используя произвольный выбор начальной оценки вектора b , а затем с помощью комбинированного метода. Результаты идентификации для трех различных значений σ представлены в таблице 1.

Поскольку вычисления проводились в рамках модельного эксперимента с известными значениями параметров $v = 2$ и $D = 3$, возникает естественное желание сравнить средние

значения оценок с истинными значениями параметров. Результаты сравнения показаны в таблице 2.

Таблица 1. Сравнение качества идентификации по статистическому критерию.

| σ | МАРЕ ^a | | | | | |
|--------------|-------------------|-------|----------------|------|-----------------------|------|
| | МНК | | Фильтр Калмана | | Комбинированный метод | |
| | ν | D | ν | D | ν | D |
| 0,001 | 7,52 | 1,02 | 3,17 | 2,74 | 0,95 | 0,3 |
| 0,005 | 96,3 | 16,3 | 11,11 | 4,58 | 4,21 | 0,67 |
| 0,01 | 158,93 | 26,34 | 32,2 | 8,36 | 6,82 | 1,19 |

^a Средняя абсолютная ошибка в процентах.

Таблица 2. Средние значения оценок параметров.

| σ | МНК | | Фильтр Калмана | | Комбинированный метод | |
|--------------|--------------|------|----------------|------|-----------------------|------|
| | ν | D | ν | D | ν | D |
| | 0,001 | 2,14 | 3,03 | 1,94 | 2,92 | 1,98 |
| 0,005 | 4 | 3,49 | 1,78 | 2,86 | 1,97 | 2,98 |
| 0,01 | 5,03 | 3,79 | 1,36 | 2,75 | 1,94 | 2,97 |

4. Заключение

Результаты проведенных экспериментов подтверждают справедливость предположения об эффективности комбинирования положительных качеств различных методов идентификации. Действительно, как показывают результаты таблиц 1 и 2, МНК-оценки имеют смещение и при больших погрешностях значительное. С другой стороны, применение фильтра Калмана с произвольным выбором начальной оценки вектора \mathbf{b} (в нашем случае $\hat{\mathbf{b}}_{00} = (0;0)^T$) также не обеспечивает получение удовлетворительных оценок. Более того, при другом выборе начальной оценки, например, $\hat{\mathbf{b}}_{00} = (1;1)^T$, качество оценивания становится еще хуже даже при невысоком уровне помех. В то же время если использовать для фильтра Калмана смещенные МНК-оценки в качестве начальных оценок вектора параметров, то качество оценивания возрастает на порядок по сравнению с обычным МНК и на несколько процентов по сравнению с прямым использованием фильтра Калмана даже в случае удачно подобранной точки $(0;0)^T$.

5. Литература

- [1] Matveev, M.G. Modeling of nonstationary distributed processes on the basis of multidimensional time series / M.G. Matveev, A.V. Kopytin, E.A. Sirota, E.A. Kopytina // Procedia Engineering. – 2017. – Vol. 201. – P. 511-516. DOI: 10.1016/j.proeng.2017.09.643.
- [2] Матвеев, М.Г. Верификация процесса конвективной диффузии на основе анализа многомерных временных рядов / М.Г. Матвеев, Е.А. Сирота, М.Е. Семенов, А.В. Копытин // Аналитика и управление данными в областях с интенсивным использованием данных: сборник научных трудов XIX Международной конференции DAMDID / RCDL'2017 (10-13 октября 2017 года, Москва, МГУ, Россия). – Москва: ФИЦ ИУ РАН, 2017. – С. 433-437.
- [3] Айвазян, С.А. Основы эконометрики, т.2 / С.А. Айвазян. – М: Юнити, 2001. – 432 с.
- [4] Fogler, H.R. A pattern recognition model for forecasting / H.R. Fogler // Management science. – 1974. – Vol. 20(8). – P. 1178-1189. DOI: 10.1287/mnsc.20.8.1178.
- [5] Conejo, A.J. Day-ahead electricity price forecasting using the wavelet transform and ARIMA models / A.J. Conejo, M.A. Plazas, R. Espinola, A.B. Molina // IEEE transaction on power systems. – 2005. – Vol. 20(2). – P. 1035-1042. DOI: 10.1109/TPWRS.2005.846054.

- [6] Alfares, H.K. Electric load forecasting: literature survey and classification of methods / H.K. Alfares, M. Nazeeruddin // *International Journal of Systems Science*. – 2002. – Vol. 33(1). – P. 23-34. DOI: 10.1080/00207720110067421.
- [7] Chui, C.K. Kalman filtering with real-time applications / C.K. Chui, G. Chen. – Berlin: Springer-Verlag Berlin Heidelberg, 2009. – 230 p.

Combined method for identifying the parameters of a distributed dynamic model

M.G. Matveev¹, A.V. Kopytin¹, E.A. Sirota¹

¹Voronezh State University, Universitetskaya sq. 1, Voronezh, Russia, 394018

Abstract. A combined method for identifying the equations of mathematical physics describing the dynamics of spatially-distributed processes on the basis of experimental multidimensional time series is proposed. The first stage of the method is the derivation of OLS estimators of multidimensional autoregression parameters. However, these estimates are biased due to the presence of errors in the regressors. To reduce this bias, the extended Kalman filter as the second component of the method is used. A computational experiment confirming the effectiveness of the proposed technique is presented.

Keywords: partial differential equations, parametric identification, multidimensional autoregression, combined method, Kalman filter.