УДК 004.942
ГРНТИ 27.35.63

ИССЛЕДОВАНИЕ МНК-ОЦЕНОК ПРИ ИДЕНТИФИКАЦИИ ТЕРМОДИНАМИЧЕСКИХ ПРОЦЕССОВ В АТМОСФЕРЕ

М.Г. МАТВЕЕВ, доктор технических наук, профессор
ВУНЦ ВВС «ВВА имени профессора Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)
А.В. ИВАНОВ, кандидат технических наук, доцент
ВУНЦ ВВС «ВВА имени профессора Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)
В.В. МАКЕЕВА
ВУНЦ ВВС «ВВА имени профессора Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)

В настоящей статье представлен анализ разностного уравнения и авторегрессии. Проведено исследование смещения методом наименьших квадратов оценок параметров авторегрессии. Разработана модельная зависимость стандартной ошибки и величины смещения оценки параметров от погрешности наблюдения. С помощью программы Excel были получены МНК-оценки параметров при различной интенсивности помех и оценки величины смещения при различных интенсивностях помех.

Ключевые слова: разностная схема; авторегрессия; смещение; метод наименьших квадратов; оценки параметров; МНК-оценки.

THE LSM-ASSESSMENTS INVESTIGATION IN THE IDENTIFICATION OF THERMODYNAMIC PROCESSES IN THE ATMOSPHERE

M.G. MATVEEV, Doctor of Technical Sciences, Professor
MESC AF "N.E. Zhukovsky and Y.A. Gagarin Air Force Academy" (Voronezh)
A.V. IVANOV, Candidate of Technical Sciences, Associate Professor
MESC AF "N.E. Zhukovsky and Y.A. Gagarin Air Force Academy" (Voronezh)
V.V. MAKEEVA
MESC AF "N.E. Zhukovsky and Y.A. Gagarin Air Force Academy" (Voronezh)

This article presents an analysis of the difference equation and autoregression. A study of bias least-squares estimates of the parameters of the autoregression. The model dependence of the standard error and the displacement of the parameter estimation on the observation error is developed. Using the program Excel, the resulting estimation of parameters at different intensity of the disturbance and estimate the magnitude of displacement at different intensities of the interference.

Keywords: difference scheme; autoregression; bias; method of least squares; parameters estimation; LSM-assessments.

Введение. Важной задачей метеорологической службы Военно-воздушных космических сил (ВКС) является анализ динамических процессов изменения атмосферной температуры в зоне действия авиации, основанный на краткосрочном (на 1 день) прогнозировании различных температурных профилей над заданной территорией. В частности, на различных геопотенциальных уровнях, соответствующих слоям атмосферы, где действует авиация, мы будем рассматривать скалярные поля атмосферной температуры, представленные регулярной сеткой, в узлах i которой производится ежедневное измерение атмосферной температуры – $y_i(t)$. Временное изменение температуры в каждом узле сетки представляет собой наблюдаемый дискретный нестационарный случайный процесс. Задача построения математических моделей динамических распределенных процессов по экспериментальным данным была и остается актуальной задачей различных практических приложений. Для ее решения предлагаются разнообразные подходы, обзоры которых можно найти, например, в работах [1–3]. Возможен подход, основывающийся на представлении динамической модели в форме разностного уравнения, параметры которого определяются как среднеквадратичные оценки на основе наблюдений пространственно распределенной переменной этого уравнения [4, 5]. При этом разностное уравнение удобно рассматривать как авторегрессионную модель временного ряда наблюдений, нахождение оценок параметров которой базируется на методе наименьших квадратов (МНК).

Известна проблема смещенности МНК-оценок при параметрической идентификации авторегрессионной модели нестационарных временных рядов [6, 7]. Если временной ряд относится к классу TS-рядов [7], то возможности параметрической идентификации ограничиваются методами «остационирования» ряда: либо последовательным применением оператора конечных разностей, либо методами коинтегрирования [7]. Процессы «остационаривания» могут обеспечить получение несмещенных оценок, но они достаточно сложны, громоздки и требуют подтверждения ряда статистических гипотез, проще говоря, не обеспечивают эффективности оценок. Поэтому часто в практических задачах пытаются использовать простые и эффективные МНК-оценки. Если смещенность таких оценок мала или удается уменьшить смещение до приемлемых значений [8], то смещением можно пренебречь, например, при условии, когда его величина соизмерима с величиной стандартной ошибки статистически значимой оценки.

Целью предлагаемой работы является исследование условий, при которых можно ожидать получения приемлемых МНК-оценок параметров разностного уравнения по результатам наблюдений нестационарных временных рядов его переменных.

Анализ разностного уравнения и авторегрессии. Пусть модель динамического распределенного процесса описывается системой приведенных разностных схем [9]:

$$y_i^{k+1} = (y^k)^T a; \quad \forall i \tag{1}$$

с заданными начальными и краевыми условиями:

$$y_i^0 = c, \quad y_{i-1}^k = b_1^k, \quad y_{i+1}^k = b_2^k, \forall k; \tag{2}$$

где i – дискретные значения пространственной координаты, k – дискретное время; $y^k = (y_{i-1}^k; y_i^k; y_{i+1}^k)^T$; $y_i = (y_i^0; y_i^1; ...; y_i^k)^T$ – векторы переменных процесса, c – начальное значение в нулевой момент времени; b_1, b_2 – наблюдаемые векторы значений краевых условий; $a = (a_1; a_2; a_3)^T$ – вектор параметров разностной схемы; T – знак транспонирования. В соответствии с условием консервативности $\sum_{i=1}^{3} a_i = 1$.

К схеме (1) приводится, например, разностный вид уравнения конвективной диффузии.

Переменные процесса наблюдаются в каждом i-ом узле с некоторой погрешностью e:

$$x_i^k = y_i^k + e_i^k \tag{3}$$

Будем считать, что случайный процесс $e_i^k, k = 0;1;\ldots$ имеет характер нормального «белого шума» во всех пространственных узлах i. Таким образом, в каждом i-ом узле формируется временной ряд наблюдений x_i^k. Запишем выражение (1) с учетом наблюдаемых значений (3):

$$x_i^{k+1} = (x^k)^T a - (e^k)^T a + e_i^{k+1}, \qquad (4)$$

где $(e^k)^T = (e_{i-1}^k; e_i^k; e_{i+1}^k)$.

Вычислим скалярное произведение $(e^k)^T a$:

$$(e^k)^T a = \tilde{e}_i^{k+1}. \qquad (5)$$

Перепишем выражение (4) в виде:

$$x_i^{k+1} = (x^k)^T a + \delta_i^{k+1}, \qquad (6)$$

где $\delta_i^{k+1} = e_i^{k+1} - \tilde{e}_i^{k+1}$.

Поскольку $\sum_{i=1}^{3} a_i = 1$, выражение (5) можно рассматривать как аффинную комбинацию случайных величин – компонент вектора e^k. При этом случайная величина $\tilde{e}_i^{k+1}$ сохраняет свойства нормального «белого шума» [10].

Для перехода к уравнению авторегрессии найдем условное математическое ожидание (6):

$$\hat{x}_i^{k+1} = M(x_i^{k+1} \mid X) = (x^k)^T a + M(\delta_i^{k+1} \mid X) \qquad (7)$$

или с учетом $M(e_i^{k+1} \mid X) = M(e_i^{k+1}) = 0$

$$\tilde{x}_i^{k+1} = (x^k)^T a - M(\tilde{e}_i^{k+1} \mid X), \qquad (8)$$

где $X = X_{k \times 3}$ – матрица результатов наблюдений в узлах $i-1; i; i+1$, формируемая в моменты времени от 0 до k.

Выражение (8) отличается от классической авторегрессии на величину $M(\tilde{e}_i^{k+1} \mid X)$, что является следствием зависимости помехи от объясняющих регрессоров, обусловленной выражением (3). Это порождает смещение МНК-оценки параметров a.

Исследование смещения МНК-оценки параметров авторегрессии. Необходимые условия минимума критерия наименьших квадратов позволяют получить МНК-оценку параметров авторегрессии в виде [11]:

$$\tilde{a} = (X^T X)^{-1} X^T x^{k+1}. \qquad (9)$$

СОДЕРЖАНИЕ

Вектор наблюдаемых значений $(x^{k+1})^T = (x_i^1;...;x_i^{k+1})$ перепишем с учетом (6) следующим образом:

$$x^{k+1} = Xa + \delta^{k+1}, \tag{10}$$

где $\delta^{k+1} = (\delta_i^1;...;\delta_i^{k+1})^T$.

Тогда выражение для вычисления оценки принимает вид:

$$\widetilde{a} = (X^T X)^{-1} X^T (Xa + \delta^{k+1}) = a - (X^T X)^{-1} X^T \delta^{k+1}. \tag{11}$$

Найдем математическое ожидание левой и правой части (11). Так как вектор e^{k+1} не зависит от X, получим $M(\widetilde{a}) = a - M[(X^T X)^{-1} X^T (\widetilde{e}^{k+1} \mid X)]$.

Тогда вектор абсолютных величин смещения оценок параметров будет определяться выражением:

$$\Delta = a - M(\widetilde{a}) = M[(X^T X)^{-1} X^T (\widetilde{e}^{k+1} \mid X)] \neq 0. \tag{12}$$

Математическое ожидание $M[(X^T X)^{-1} X^T (\widetilde{e}^{k+1} \mid X)]$ следует рассматривать как условное, т. е. как вектор со случайными компонентами. Очевидно, что в этом случае и смещение Δ является случайным вектором.

Наличием смещения можно пренебречь в случаях, когда величина стандартной ошибки оценки параметров существенно больше величины смещения. Величина стандартной ошибки оценки параметров при фиксированном объеме выборки зависит от погрешности наблюдения e. Величина смещения также зависит от этой погрешности. Стандартная ошибка рассчитывается при МНК-оценке параметров, а величину смещения можно рассчитать с помощью выражения (11). Построим эти зависимости на основе модельного эксперимента.

Модельная оценка зависимости стандартной ошибки и величины смещения оценки параметров от погрешности наблюдения. Процесс (12) был воспроизведен с заданными значениями параметров: $a_1 = 0{,}5;\ a_2 = 0{,}3;\ a_3 = 0{,}2$. В результате была получена выборка из 100 строк значений переменных $y^k = (y_{i-1}^k; y_i^k; y_{i+1}^k)$, $k = 1; 2;...; 100$, фрагмент которой показан в таблице 1.

Таблица 1 – Фрагмент модельной статистической выборки

K	y_{i-1}	y_i	y_{i+1}
1	5,1	3,9	6,9
2	5,2	5,1	6,8
3	2,3	5,5	6,7
4	5,4	5,6	6,6
5	5,5	5,7	6,5
6	5,6	5,8	6,4
7	5,7	5,8	6,3
8	5,8	5,8	6,2
9	5,9	5,9	6,1

СОДЕРЖАНИЕ

К полученным значениям переменных $y^k = (y_{i-1}^k; y_i^k; y_{i+1}^k)$ были добавлены случайные гауссовские переменные с нулевым математическим ожиданием и переменной дисперсией – $k\sigma^2$, где $\sigma^2 = 1;\quad k \in \{0{,}01;\ 0{,}1;\ 0{,}5\}$.

Окончательно сформированная выборка использовалась для получения МНК-оценок параметров моделируемого процесса.

При различной интенсивности помех были получены МНК-оценки параметров $\hat{a}_1; \hat{a}_2; \hat{a}_3$, приведенные в таблице 2.

Таблица 2 – МНК-оценки параметров при различной интенсивности помех

	a_1	Стандартная ошибка оценки	a_2	Стандартная ошибка оценки	a_3	Стандартная ошибка оценки
Исходные значения параметров	0,5		0,3		0,2	
k=0,01	0,499	0,001	0,301	0,001	0,201	0,001
k=0,1	0,497	0,011	0,302	0,011	0,201	0,011
k=0,5	0,486	0,051	0,307	0,050	0,206	0,052

По полученным выборкам для различных интенсивностей помех были рассчитаны оценки величины смещения по формуле:

$$\hat{\Delta} = (X^T X)^{-1} X^T \delta^{k+1}. \tag{13}$$

Результаты расчетов помещены в таблицу 3.

Таблица 3 – Оценки величины смещения при различных интенсивностях помех

	$\hat{\Delta} a_1$	$\hat{\Delta} a_2$	$\hat{\Delta} a_3$
k=0,01	0,0004	-0,0007	0,0003
k=0,1	-0,003	0,006	-0,003
k=0,5	-0,04	0,08	-0,04

Сравнительные зависимости изменения стандартных ошибок оценок параметров и абсолютной величины их смещения от погрешности наблюдений показаны на рисунке 1.

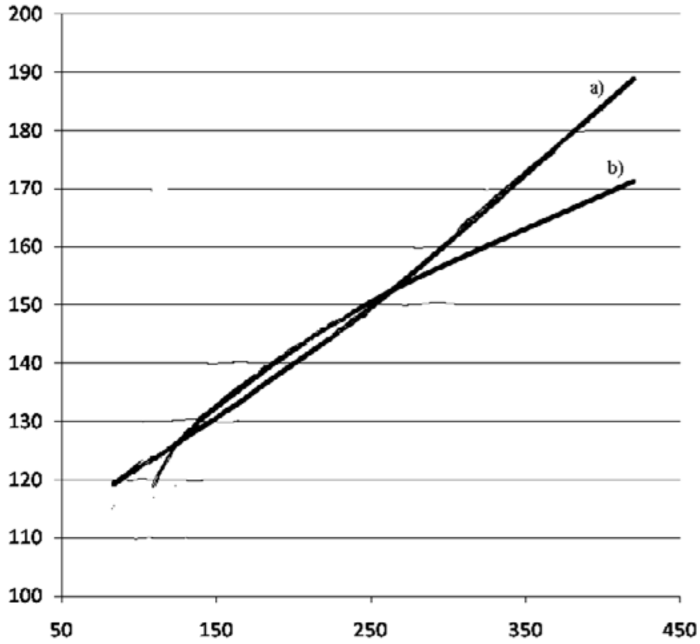

Рисунок 1 – Стандартные ошибки и абсолютные погрешности в логарифмических координатах

а) стандартные ошибки оценок параметров; б) абсолютные погрешности оценок смещения

МНК-оценки обладают свойством эффективности при выполнении условий [12]: $\sigma_e^2 = const$; $\operatorname{cov}(e_i; e_j) = 0$, которые выполняются в рассматриваемом случае. Тогда применение методов обеспечивающих несмещенные, но не обязательно эффективные оценки может привести к ситуации, показанной на рисунке 2. В такой ситуации, смещенные эффективные МНК-оценки будут более предпочтительны, чем несмещенные оценки со значительной дисперсией.

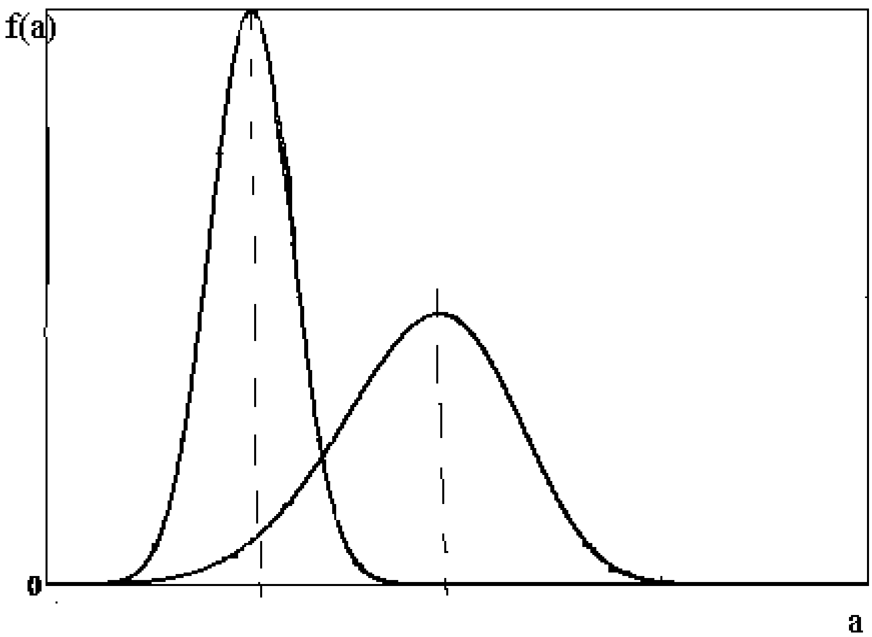

Рисунок 2 – Смещенные эффективные МНК-оценки

Выводы. Полученные результаты показывают, что при небольших помехах величина смещения МНК-оценок параметров приведенной разностной схемы (1) существенно меньше стандартной ошибки и таким смещением можно пренебречь. В этом случае, вместо громоздких алгоритмов стационирования временного ряда наблюдений или использования метода инструментальной переменной [12] можно применять прямые МНК-оценки без существенной потери их адекватности. С увеличением помех величина смещения остается того же порядка, что и стандартная ошибка. Это означает, что смещение не выходит, например, за границы доверительного интервала с 5% уровнем значимости.

Если учесть, что измерения атмосферной температуры, производимые современными приборами, достаточно точны и, кроме того, подвергаются процедурам усреднения, что также уменьшает разброс значений, то применение МНК-оценок при обработке временных рядов атмосферной температуры можно считать оправданным для определенного класса задач моделирования. Такой вывод подтверждается многочисленными положительными результатами моделирования температурных полей, полученными в исследованиях Военного учебно-научного центра Военно-воздушных сил «Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина.

СПИСОК ЛИТЕРАТУРЫ

1. Дубенко Т.И. Идентификация и оценивание параметров в стохастических системах, описываемых уравнениями с частными производными // Автоматика и телемеханика. 1983. Вып. 12. С. 5-19.

2. Muller T.G, Timmer G. Parameter Identification techniques differential equations // International Journal of Bifurcation and Chaos. 2004. Vol. 14, No. 6. P. 2053-2060.

3. Бойко И.В., Кривулин Н.П. Методы идентификации динамических систем // Программные системы: теория и приложения. 2014. №5(23). С. 79-96.

4. Матвеев М.Г. Сирота Е.А. Разработка и исследование статистических моделей нестационарного многомерного временного ряда атмосферных температур в условиях неоднородности // Информационные технологии. 2014. № 12. С. 20-24.

5. Матвеев М.Г., Михайлов В.В., Сирота Е.А. Комбинированная прогностическая модель нестационарного многомерного временного ряда для построения пространственного профиля атмосферной температуры // Информационные технологии. 2016. Т. 22, №2. С. 89-94.

6. Patterson, K. An Introduction to Applied Econometrics: A Time Series Approach / K. Patterson. New York: Palgrave, 2000. 832 p

7. Гребенюк Е. А. Методы анализа нестационарных временных рядов с неявными изменениями свойств // Автоматика и телемеханика. 2005. №12. С. 3–29.

8. Заусаева М.А., Зотеев В.Е. Определение параметров двумерных динамических процессов на основе разностных схем // Вестн. Сам. гос. техн. ун-та. Сер. Физ.-мат. науки. 2010. Т. 20, №1. С. 154–161.

9. Моделирование динамики атмосферных процессов на основе анализа многомерных временных рядов. / М.Г. Матвеев и др. // Воздушно-космические силы. Теория и практика. 2017. №1 (1). С. 98-110.

10. Бахтин В.И. Введение в прикладную статистику. Минск: БГУЭ, 2012. 99 с.

11. Демиденко Е.З. Линейная и нелинейная регрессия. М.: Финансы и статистика, 1981. 302 с.

12. Носко В.Н. Эконометрика. Кн.2. М.: Дело, 2011. 259 с.

REFERENCES

1. Dubenko T.I. Identifikatsiya i otsenivanie parametrov v stokhasticheskikh sistemakh, opisyvaemykh uravneniyami s chastnymi proizvodnymi // Avtomatika i telemekhanika. 1983. Vyp. 12. P. 5-19.

2. Muller T.G, Timmer G. Parameter Identification techniques differential equations // International Journal of Bifurcation and Chaos. 2004. Vol. 14, No. 6. P. 2053-2060.

3. Bojko I.V., Krivulin N.P. Metody identifikatsii dinamicheskikh sistem // Programmnye sistemy: teoriya i prilozheniya. 2014. №5(23). P. 79-96.

4. Matveev M.G. Sirota E.A. Razrabotka i issledovanie statisticheskikh modelej nestatsionapnogo mnogomepnogo vpemennogo pyada atmosfepnykh tempepatup v usloviyakh neodnopodnosti // Informatsionnye tekhnologii. 2014. № 12. S. 20-24.

5. Matveev M.G., Mikhajlov V.V., Sirota E.A. Kombinirovannaya prognosticheskaya model' nestatsionarnogo mnogomernogo vremennogo ryada dlya postroeniya prostranstvennogo profilya atmosfernoj temperatury // Informatsionnye tekhnologii. 2016. T. 22, №2. P. 89-94.

6. Patterson, K. An Introduction to Applied Econometrics: A Time Series Approach / K. Patterson. New York: Palgrave, 2000. 832 p.

7. Grebenyuk E. A. Metody analiza nestatsionarnykh vremennykh ryadov s neyavnymi izmeneniyami svojstv // Avtomatika i telemekhanika. 2005. №12. C. 3–29.

8. Zausaeva M.A., Zoteev V.E. Opredelenie parametrov dvumernykh dinamicheskikh protsessov na osnove raznostnykh skhem // Vestn. Sam. gos. tekhn. un-ta. Ser. Fiz.-mat. nauki. 2010. T. 20, №1. S. 154–161.

9. Modelirovanie dinamiki atmosfernykh protsessov na osnove analiza mnogomernykh vremennykh ryadov. / M.G. Matveev i dr. // Vozdushno-kosmicheskie sily. Teoriya i praktika. 2017. №1 (1). S. 98-110.

10. Bakhtin V.I. Vvedenie v prikladnuyu statistiku. Minsk: BGUEH, 2012. 99 p.

11. Demidenko E.Z. Linejnaya i nelinejnaya regressiya. M.: Finansy i statistika, 1981. 302 p.

12. Nosko V.N. EHkonometrika. Kn.2. M.: Delo, 2011. 259 p.

Матвеев Михаил Григорьевич, доктор технических наук, профессор, старший научный сотрудник 21 отдела научно-исследовательского 2 управления научно-исследовательского научно-исследовательского центра (проблем применения, обеспечения и управления авиацией Военно-воздушных сил), Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж), Россия, 394064, г. Воронеж, ул. Старых Большевиков, 54А.

Иванов Алексей Владимирович, кандидат технических наук, доцент, начальник 2 управления научно-исследовательского научно-исследовательского центра (проблем применения, обеспечения и управления авиацией Военно-воздушных сил), Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж), Россия, 394064, г. Воронеж, ул. Старых Большевиков, 54А.

Макеева Виктория Владимировна, младший научный сотрудник 21 отдела научно-исследовательского 2 управления научно-исследовательского научно-исследовательского центра (проблем применения, обеспечения и управления авиацией Военно-воздушных сил), Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж), Россия, 394064, г. Воронеж, ул. Старых Большевиков, 54А, Toryka_M@mail.ru

СОДЕРЖАНИЕ