



РЕКОНСТРУКЦИЯ ПРИ НАЛИЧИИ СКРЫТЫХ ПЕРЕМЕННЫХ: МОДИФИЦИРОВАННЫЙ АЛГОРИТМ БОКА *

Б.П. Безручко, Д.А. Смирнов, И.В. Сысоев

Рассматривается задача оценки параметров модели по временным рядам в условиях, когда часть переменных не наблюдается и/или данные сильно зашумлены. С помощью введенных количественных критериев сравниваются различные алгоритмы решения задачи на основе метода множественной стрельбы – алгоритм Бока и его модифицированный вариант, допускающий разрывы фазовой траектории модели. Показано, что модифицированный метод существенно эффективнее для хаотических рядов большой длины и предъявляет более мягкие требования к стартовым догадкам, что существенно расширяет возможности его практического применения.

Введение

Задача построения математической модели¹, описывающей сложную динамику исследуемого объекта, по временным реализациям экспериментально наблюдаемых величин $\eta(t)$, возникает во многих областях науки и техники и имеет различные постановки [1]. Мы рассматриваем постановку, в которой структура модельных уравнений полностью известна из «первых принципов» или других соображений:

$$dy/dt = f(y, c), \quad (1)$$

где y – модельная переменная (D -мерный вектор состояния), а c – P -мерный вектор параметров, который необходимо найти. Задача сводится к оценке неизвестных параметров c_1, \dots, c_P по наблюдаемым данным, которые имеют вид дискретной последовательности отсчетов – временного ряда $\{\eta_i\} = \{\eta_1, \dots, \eta_N\}$, где i – номер

*Статья написана по материалам доклада на VII Международной школе «Хаотические автоколебания и образование структур», 1-6 октября 2004, Саратов, Россия.

¹В нелинейной динамике ее называют реконструкцией динамических систем, в статистике – идентификацией систем.

отсчета, N – длина ряда². Такая постановка встречается, например, в аэродинамике (оценка коэффициентов аэродинамических сил при летном эксперименте с самолетом [2]), лазерной физике (оценка скоростей перехода между уровнями рабочего вещества газового лазера [3]), радиотехнике (моделирование электрических цепей с сегнетоэлектрическими или полупроводниковыми нелинейностями [4, 5]), клеточной биологии (оценка параметров модели нейрона [6], модели сигнального пути клеток определенного вида [7]).

Стандартная процедура реконструкции модели (1) предполагает знание векторного временного ряда значений y , то есть требуется иметь ряды всех D компонент, формируемых по данным эксперимента из ряда наблюдаемой $\{\eta_i\}$. Например, одна из переменных может просто совпадать с наблюдаемой, то есть измеряется непосредственно, а другие – получаются из нее дифференцированием, интегрированием, использованием временной задержки, весовым суммированием и другими способами. Однако нередко из-за невозможности измерить нужную величину вследствие зашумленности или иных причин удается сформировать ряды лишь для $l < D$ компонент y . ($D - l$) переменных, ряды которых не могут быть получены из данных наблюдения, называют *скрытыми*. Наличие скрытых переменных сильно усложняет задачу, требуя использования специальных подходов.

В этой ситуации основная тяжесть задачи построения модели переносится на этап оценки параметров, причем в число искомых величин кроме параметров вводят начальные условия. Обычно для оценки параметров используется метод наименьших квадратов³, который формулируется в данной ситуации следующим образом. Нужно подобрать такие начальные условия модели y^0 и параметры c , чтобы обеспечить минимальное отклонение решения модельных уравнений y от данных экспериментального наблюдения y_{η}^l . Сравнение проводится только по l нескрытым компонентам y . То есть требуется минимизировать функцию:

$$S(y^0, c) = \sum_{i=1}^N (y_i^l(y^0, c) - y_{\eta i}^l)^2 = \min, \quad (2)$$

где $y_i^l(y^0, c)$ – l -мерный вектор «нескрытой» части решения уравнений (1), $y_{\eta i}^l$ – аналогичные значения, полученные из ряда наблюдаемой. Минимизацию (2) проводят с помощью численных итерационных методов, отправляясь от некоторых *стартовых догадок* для искомых величин y^0, c .

В случае хаотических рядов траектория модели очень чувствительна к начальным условиям, поэтому «рельеф» целевой функции (2) при значительном N весьма изрезан и имеет множество локальных минимумов, а область сходимости в глобальный минимум, соответствующий наиболее точным оценкам, очень узка, то есть требуются исключительно удачные стартовые догадки, чтобы туда попасть. Обойти трудности обещает изящный технический трюк – алгоритм множественной

² Будем называть величину η просто наблюдаемой и ограничимся случаем, когда она скалярна, – наиболее сложным, но очень распространенным.

³ Следует сказать, что более статистически предпочтительными были бы оценки, полученные методом максимального правдоподобия, но его использование на практике затруднительно из-за сложности функции максимального правдоподобия и в данной работе отвлекло бы от основной цели исследования. Отметим, что, если шум является аддитивным, нормальным и дельта-коррелированным, оба метода совпадают.

стрельбы [8, 9]. Но, как показывает опыт, он требует дополнительных усилий для его успешной реализации. Эта проблема уже привлекала к себе внимание [3, 10], но систематический ее анализ не проводился. В данной работе мы ввели количественные критерии эффективности процедуры оценки параметров и проводим анализ условий работоспособности наиболее продуктивных алгоритмов реализации метода множественной стрельбы (раздел 1). В численном эксперименте (раздел 2) на примере эталонной хаотической системы, в том числе при добавлении измерительных шумов, мы демонстрируем, что для длинных рядов наиболее эффективен модифицированный алгоритм Бока, ослабляющий требование непрерывности фазовой траектории модели на интервале наблюдения.

Актуальность поставленной задачи объясняется тем, что хаотическое поведение изучаемых систем и дефицит априорной информации о значениях искомых величин типичны в исследовательской практике. Кроме параметров, рассматриваемые методы позволяют получить значения самих скрытых переменных. Учитывая, что эти величины могут быть в принципе недоступны прямому измерению, рассматриваемая процедура выступает в роли универсального «измерительного прибора».

1. Методы оценки параметров при наличии скрытых переменных

1.1. Метод начального условия

Так называют метод оценки параметров, состоящий в минимизации непосредственно функции (2), где N – длина всего наблюдаемого временного ряда [11]. На практике при требуемых достаточно больших⁴ N он, как правило, неприменим, особенно в случае хаоса, поскольку из-за экспоненциальной чувствительности траектории модели к начальным условиям y^0 область сходимости в глобальный минимум целевой функции (2) столь мала, что в нее почти невероятно попасть. Недостатки этого подхода в сравнении с другими продемонстрированы, в частности, в [9] и [5], и поэтому мы не будем его подробно рассматривать.

1.2. Метод множественной стрельбы⁵

1.2.1. Алгоритм Бока. Увеличить допустимые значения N и погрешностей в стартовых догадках позволяет модификация метода начального условия. Она направлена на то, чтобы хотя бы на промежуточных этапах минимизации целевой функции снизить чувствительность траектории модели к начальным условиям y^0 . Это достигается путем разбиения исходного ряда на L более коротких фрагментов (n – длина фрагмента, $N = Ln$) и использования начальных условий на этих фрагментах $y_1^0, y_{n+1}^0, \dots, y_{(L-1)n+1}^0$ в качестве дополнительных искомых величин – аргументов

⁴Без достаточного количества точек нельзя выудить из сигнала информацию при наличии шумов, кроме того, важно, чтобы используемый ряд отражал все временные масштабы объекта.

⁵Название «метод множественной стрельбы» происходит от аналогии с методом решения краевых задач для обыкновенных дифференциальных уравнений. Поскольку множественная стрельба лежит в основе нескольких методов, первоначальный алгоритм, предложенный Боком, будем именовать «алгоритм Бока».

целевой функции S :

$$S(\mathbf{y}_1^0, \dots, \mathbf{y}_{(L-1)n+1}^0, \mathbf{c}) =$$

$$= \sum_{i=1}^L \sum_{j=1}^n \left(\mathbf{y}_{(i-1)n+j}^l \left(\mathbf{y}_{(i-1)n+1}^0, \mathbf{c} \right) - \mathbf{y}_{\eta(i-1)n+j}^l \right)^2 = \min, \quad (3)$$

где нижний индекс $\mathbf{y}_{(i-1)n+1}^0$ соответствует первой точке i -го фрагмента. Чтобы избежать большого числа независимых неизвестных, увеличивающего дисперсию оценок, вводится условие непрерывности траектории модели – «сшивания» фрагментов:

$$\mathbf{y}_{in+1}^0 = \mathbf{y}_{in+1} \left(\mathbf{y}_{(i-1)n+1}^0, \mathbf{c} \right), \quad i = 1, \dots, L - 1. \quad (4)$$

Минимизация (3) при условии (4) – это задача условной минимизации. При произвольном выборе стартовых догадок для всех искомым величин траектория модели, как правило, состоит из L нестыкующихся фрагментов. Но в процессе работы итерационного метода условие (4) удовлетворяется все лучше и лучше, так что в итоге траектория модели оказывается непрерывной. Заметим, что при $L = 1, n = N$ алгоритм Бока превращается в метод начального условия.

Хотя в [9] утверждается, что для эффективности алгоритма Бока не требуется стартовых догадок, близких к истинным значениям, *практика показывает, что это чаще всего не так* (см. ниже, раздел 2). Алгоритм только до некоторой степени расширяет возможности оценки параметров по хаотическому ряду по сравнению с «наивным» методом начального условия. Все же итоговое условие (4) часто оказывается для хаотических систем слишком жестким и при достаточно большой длине N удается попасть только в локальные минимумы целевой функции (3).

1.2.2. Кусочные методы. Чтобы точнее оценить параметры, проще всего разбить ряд на короткие сегменты (l штук) и проводить реконструкцию по каждому из сегментов отдельно. Итоговая оценка получается как среднее арифметическое $\bar{\mathbf{c}} = \frac{1}{l} \sum_{i=1}^l \mathbf{c}_i$. Такой подход называют «кусочным» [10]; при использовании на отдельных сегментах алгоритма Бока будем говорить о *кусочном методе множественной стрельбы*. Проблема здесь состоит в том, что для короткого сегмента оценка может быть существенно смещенной (даже оценки максимального правдоподобия лишь асимптотически не смещены) и это смещение может не устраняться при усреднении по l кускам. Поэтому погрешности оценок для кусочного метода больше, чем для исходного алгоритма Бока, если при использовании последнего удастся найти глобальный минимум.

1.2.3. Модифицированный метод. Опираясь на статистические соображения [10], можно утверждать, что использование сразу всего объема данных позволяет получать оценки с меньшим смещением, чем при кусочных подходах. Поэтому мы предлагаем обратить внимание на модификацию алгоритма Бока (далее – *модифицированный метод*), которая уже использовалась в работе [3] для нехаотических рядов вынужденно, поскольку эти ряды составлялись из нескольких измерений. Возможность построения такой модификации также упоминалась в [10]. Модифицированный метод состоит в отказе от непрерывности траектории модели в некоторые моменты времени внутри интервала наблюдения, то есть условие (4) не требуется для $(\nu - 1)$ моментов времени. Это означает, что начальные условия модели в эти

ν моментов, включая начальный, становятся независимыми искомыми величинами. В отличие от (3), где разрывность траекторий была следствием способа получения данных, мы предлагаем искусственно ввести возможность разрывов. Мы выберем их распределенными равномерно внутри интервала наблюдения. По сравнению с кусочным методом параметры с всегда удерживаются одинаковыми на всех сегментах. Такой подход сочетает ослабление требований к непрерывности траекторий модели с использованием всего длинного ряда в целом для оценки параметров. Он имеет два свободных параметра: ν – число сегментов с независимыми начальными условиями и L – число подсегментов, на которые делится каждый сегмент для реализации алгоритма Бока, где $N = \nu Ln$. При $\nu = 1$ модифицированный метод сводится к алгоритму Бока. Этот подход, на наш взгляд незаслуженно, не получил до сих пор широкого распространения. По всей видимости, это связано с тем, что нарушение гладкости траектории на первый взгляд не сулит хорошей модели и настораживает.

2. Сравнительный анализ в численном эксперименте

2.1. Методика сравнения

Чтобы охарактеризовать и сопоставить эффективность алгоритмов, нами использовались следующие способы представления информации и количественные меры. На плоскостях параметров c_{i_1}, c_{i_2} (рис. 1, *a-в*), в шкале серого отмечены результаты выбора соответствующих стартовых догадок. Белым показаны стартовые точки, откуда достигается глобальный минимум, то есть получаются очень точные оценки параметров. Оттенками серого – точки, позволившие при оптимизации достичь лишь локального минимума, из которых получается оценка, далекая от истинных значений параметров – чем хуже, тем темнее. Для удобства на осях откладываются значения, нормированные так, чтобы идеальной догадке (истинным значениям параметров c_i^0) соответствовала центральная часть рисунка: $b_i = (c_i - c_i^0) / c_i^0$. Точка $\mathbf{b} = \mathbf{0}$ соответствует глобальному минимуму, если пренебречь шумами, включая погрешности вычислений. Размер светлой области на рис. 1 характеризует эффективность мето-

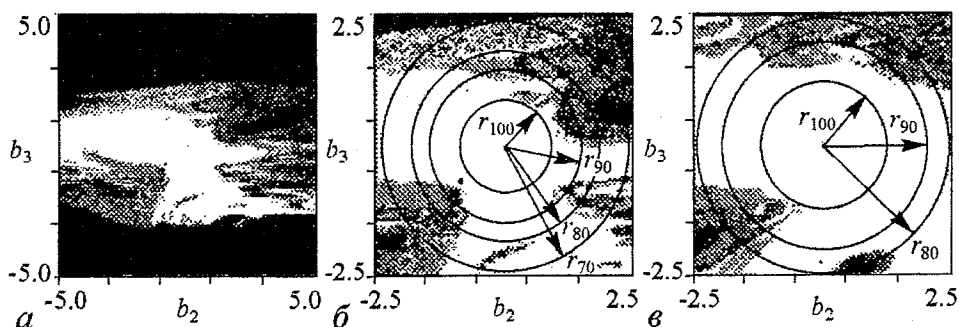


Рис. 1. Плоскость нормированных стартовых догадок для параметров системы Лоренца (сечение пространства параметров плоскостью $b_1 = 0$). Белым обозначены точки, откуда достигается глобальный минимум, оттенками серого – локальные минимумы (интенсивность зависит от их удалённости от глобального), черным – алгоритм минимизации расходится: *a, б* – алгоритм Бока при $L = 30, n = 35$; *в* – модифицированный метод с $L = 15, n = 35, \nu = 2$

да – чем она больше, тем работоспособнее алгоритм. Области хороших догадок в типичном случае сильно изрезаны, поэтому в качестве интегральной характеристики эффективности метода использовалась мера, показывающая долю белых точек в круге радиуса r с центром в начале координат⁶. Обозначим через μ относительную площадь светлой области внутри круга с центром в точке $\mathbf{b} = \mathbf{0}$, это – оценка вероятности попадания в глобальный минимум, если ошибка в стартовых догадках не больше r . Чем больше μ при данном r , тем эффективнее метод. Терпимость метода к ошибкам в стартовых догадках усредненно характеризуется величиной радиуса r_μ , при которой вероятность попадания в глобальный минимум не меньше заданного μ (см. рис. 1, б, в). Для определенности далее будем, как правило, использовать r_{100} – радиус области 100-процентной сходимости.

2.2. Оценка параметров системы Лоренца

Методику исследования эффективности методов (работоспособности при длинных рядах и «терпимости» к стартовым догадкам) проиллюстрируем на конкретном примере системы Лоренца

$$\begin{aligned}\dot{y}_1 &= c_1 (y_2 - y_1) \\ \dot{y}_2 &= -y_2 + y_1 (c_3 - y_3) \\ \dot{y}_3 &= -c_2 y_3 + y_1 y_2\end{aligned}\tag{5}$$

с параметрами $c_1^0 = 10$, $c_2^0 = 8/3$, $c_3^0 = 46$ по решению, полученному при начальном условии $y_1 = -7.60$, $y_2 = -12.37$, $y_3 = 38.66$, произвольно выбранном на хаотическом аттракторе (старший ляпуновский показатель $\lambda_1 = 1.23$ [9]). Уравнения интегрировались методом Рунге – Кутты 4-го порядка с шагом 0.001, интервал выборки составлял 0.002 (около 280 точек на характерном периоде T). В качестве ряда скалярной наблюдаемой использовалась последовательность значений переменной y_1 . Для большей реалистичности к переменной добавлялся нормальный шум наблюдения: $\eta = y_1 + \xi$. Две остальные переменные y_2 и y_3 считались скрытыми, их временные ряды – неизвестными.

Угадать истинные (идеальные, наилучшие) стартовые догадки маловероятно, поэтому, следуя работам [5, 9], при оценке параметров в качестве стартовых догадок для всех переменных модели $\mathbf{y}_1^0, \mathbf{y}_{n+1}^0, \dots, \mathbf{y}_{(L-1)n+1}^0, \dots, \mathbf{y}_{(v-1)Ln+(L-1)n+1}^0$ использовались значения наблюдаемой в соответствующие моменты времени. Такой подход неплохо зарекомендовал себя, хотя он и не идеален⁷. Для минимизации целевой функции использовался обобщенный метод Гаусса – Ньютона [9].

На рис. 1, а, б в разных масштабах представлены результаты исследования сходимости в глобальный минимум для алгоритма Бока (п. 2.2) при различных старто-

⁶Наиболее информативная картинка получится, если строить шар в пространстве всех стартовых догадок, но для удобства иллюстрации мы ограничимся двумерной характеристикой.

⁷Варианты использования идеальных стартовых догадок для скрытых переменных, а также идеальных догадок с добавленным к ним шумом также были рассмотрены нами достаточно подробно, но результаты оказались качественно сходны с приведенными ниже.

вых догадках для параметров. Картины получены в случае длины ряда, при которой алгоритм Бока дает максимальный радиус стопроцентной сходимости. На рис. 1 приведены сечения пространства параметров плоскостью $b_1 = 0$, поскольку в рассматриваемом случае ошибки в параметре b_1 наименее критичны. Как видно из рис. 1, *а, б*, область сходимости алгоритма Бока в случае незашумленного ряда достаточно широка – диаметр окружности r_{100} на плоскости нормированных параметров больше 1.0, то есть допускаются ошибки в стартовых догадках больше 100%. Существуют и очень удаленные от истинных значений области, откуда также достигается сходимость в глобальный минимум (см. рис. 1, *а*). Однако модифицированный метод еще менее требователен к стартовым догадкам для параметров: сравните рис. 1, *б* и рис. 1, *в*, где значения r_{100} , r_{90} , r_{80} существенно больше, а соответствующая r_{100} белая область – шире.

Рис. 2 позволяет оценить зависимость результатов моделирования от уровня измерительного шума, добавлявшегося к ряду аддитивно и имевшего нормальное распределение. Из рис. 2, *а*, где построены графики меры $\mu(r)$ при различных уровнях шума, видно, что работоспособность подхода сохраняется и при наличии шума умеренной интенсивности: горизонтальная площадка при $\mu = 1$, соответствующая радиусу 100-процентной сходимости, несколько укорачивается, но остается весьма значительной – r_{100} уменьшается при использовании исходного алгоритма с 1.1 до 0.7, а модифицированного (при $\nu = 2$) – с 1.2 до 0.9. Аналогичные выводы можно сделать из рис. 2, *б*, где показаны зависимости $r_{100}(N)$ для незашумленного ряда (при $\nu = 1$ и $\nu = 4$) и для ряда, к которому был добавлен 1-процентный шум.

О преимуществах модифицированного метода свидетельствует более светлый тон рис. 3, где на плоскостях параметров длина сегмента n – количество сегментов L для алгоритма Бока и модифицированного подхода (с $\nu = 2$) интенсивностью се-

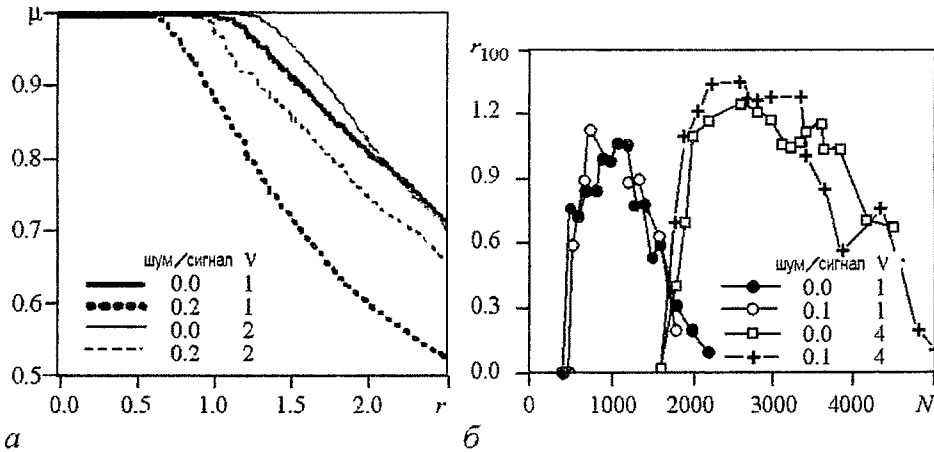


Рис. 2. Иллюстрация зависимости результатов моделирования от уровня шума: *а* – мера $\mu(r)$ для алгоритма Бока (жирные кривые) и для модифицированного метода (тонкие кривые) при различных уровнях шума, числами показано отношение шум/сигнал в единицах стандартных отклонений; *б* – зависимость радиуса стопроцентной сходимости r_{100} от общей длины ряда N при разных уровнях шума

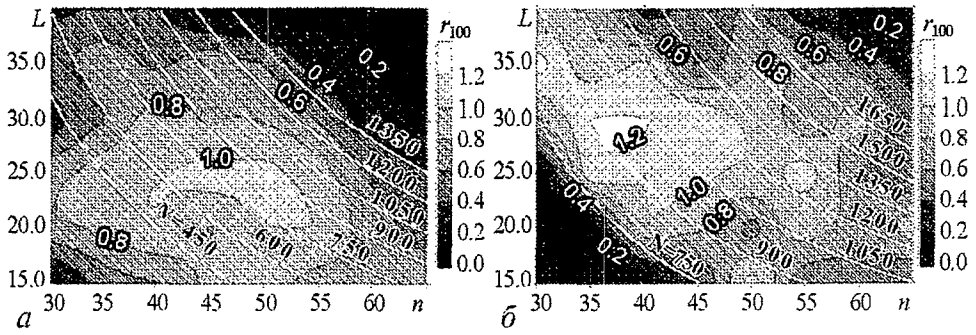


Рис. 3. Зависимость $r_{100}(L, n)$ для алгоритма Бока (а) и модифицированного подхода при $\nu = 2$ (б). Чем темнее область, тем меньше в ней r_{100} . На ограничивающих области линиях уровня указана точная величина r_{100} , нанесенные на рисунок белые гиперболы соответствуют значениям постоянной длины ряда $N = \text{const}$

рого показана зависимость $r_{100}(L, n)$: чем темнее область, тем меньшему значению r_{100} она соответствует. Точные значения r_{100} указаны на ограничивающих разные области линиях уровня. Максимально достижимое значение r_{100} для модифицированного метода больше, чем для алгоритма Бока: не только область с $r_{100} \geq 1$ больше, но и внутри нее выделяется область с $r_{100} > 1$. Кроме того, рассматривая белые гиперболы (линии $N = \text{const}$) на рис. 3, можно сделать вывод, что преимущества модифицированного метода сильнее проявляются для более длинных рядов.

Последние выводы подтверждаются и уточняются на рис. 4, а, где показана величина r_{100} в зависимости от N для разного числа сегментов ν (параметры L и n подбирались так, чтобы получить максимальный r_{100})⁸. Холмообразная форма зависимостей $r_{100}(N)$ определяется двумя факторами: при очень малых N не хватает точек для подгонки параметров, а при больших сказывается экспоненциальная чувствительность к начальным условиям – за время $\tau_\lambda = 1/\lambda_1$ малые возмущения достигают макромасштабов. Кривые, соответствующие большим значениям ν , достигают больших значений r_{100} и сдвинуты вправо, в сторону более длинных рядов. При этом шире становятся и интервалы по оси N с ненулевыми значениями r_{100} , то есть с увеличением числа разрывов диапазон длин ряда, в котором можно получить хорошую модель, возрастает.

Проведенное исследование (см. рис. 4, б) показывает, что оптимальное значение длины сегмента Ln связано с ляпуновским временем τ_λ и числом разрывов ν . Для случаев $\nu = 1 \div 4$ оптимальные длины соответствуют 1–3 ляпуновским временам (см. верхнюю шкалу по оси абсцисс), смещаясь в сторону меньших значений с ростом ν . Это можно объяснить со следующих позиций. Шансы на успех реконструкции зависят от допустимой длины сегмента Ln , на которой малые возмущения начальных условий еще не слишком возрастают, а также, от числа $P + \nu D$ свободных параметров, которые нужно подогнать по ряду. Увеличивая число разрывов, мы ослабляем чувствительность к начальным условиям, но одновременно увеличиваем количество подгоняемых величин. При $\nu = 1$ их всего шесть, при $\nu = 2$ – девять,

⁸Для этого удобно использовать белые гиперболы на рис. 3 – линии $N = \text{const}$, выбирая L и n , соответствующие более светлым областям.

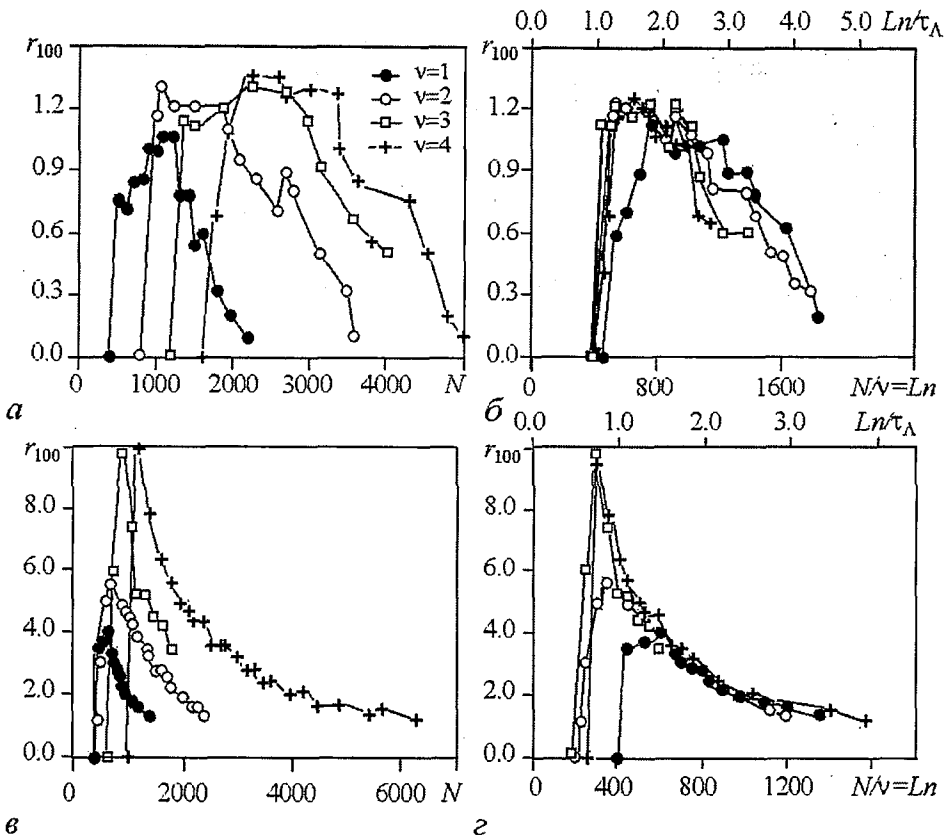


Рис. 4. Зависимость радиуса стопроцентной сходимости r_{100} : а – от общей длины используемого ряда N при различном числе разрывов ($v - 1$); б – от длины сегмента (Ln); в, г – аналогичные зависимости, построенные при использовании в качестве стартовых догадок для скрытых переменных их истинных значений (идеальных догадок), с добавлением к ряду 1-процентного нормального шума наблюдения

а при $v = 4$ – уже пятнадцать. Поэтому с ростом общего количества подгоняемых величин результат реконструкции более критичен к ошибкам в начальных условиях и интервал значений Ln с ненулевым r_{100} сужается.

Рис. 4, в, г аналогичны рис. 4, а, б, но при их построении использовались не реалистичные, то есть взятые только из наблюдаемой, а идеальные догадки для скрытых переменных, что заметно облегчает процесс моделирования. Это объясняет основное отличие этих графиков – кривые в целом лежат намного выше. На рис. 4, г наилучшие результаты соответствуют длине сегмента, немного меньшей одного ляпуновского времени, что, на наш взгляд, объясняется введением в ряд наблюдаемой аддитивного шума.

Сходные результаты были получены и при других значениях параметров: $c_1^0 = 10$, $c_2^0 = 8/3$, $c_3^0 = 28$ – см. рис. 5, где приведена зависимость, аналогичная рис. 3. Видно, что в целом фон на рис. 5, б светлее, чем на рис. 5, а, то есть модифицированный метод дает в среднем больший радиус сходимости. Кроме того, на рис. 5, б выделяется область с $r_{100} \geq 1.0$, которой нет на рис. 5, а. При этом преимущества модифицированного метода проявляются при больших длинах ряда.

Аналогично исследовался случай моделирования по временным рядам, записанным при использовании в качестве наблюдаемой величины y_2 .

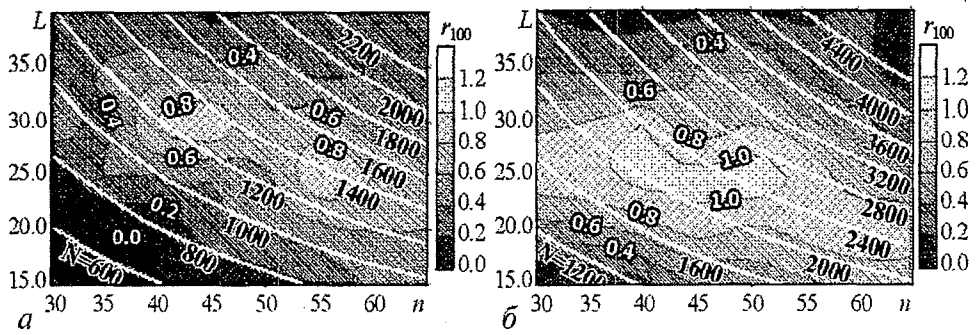


Рис. 5. Зависимость $r_{100}(L, n)$ для алгоритма Бока (а) и модифицированного подхода при $\nu = 2$ (б). Значения параметров: $c_1^0 = 10$, $c_2^0 = 8/3$, $c_3^0 = 28$

Заключение

С помощью введенного количественного критерия проведено сравнение работоспособности различных реализаций метода множественной стрельбы для оценки параметров при наличии скрытых переменных. На примере реконструкции параметров системы Лоренца показано, что исходный алгоритм Бока уступает его модификации, заключающейся в допуске разрывов траектории модели при сохранении единых значений параметров.

Продемонстрирована степень влияния на результат оценивания параметров длины используемого ряда, количества его сегментов, выбора наблюдаемой и способа задания стартовых догадок для скрытых переменных. Показано, что шансы на успех оценки значений параметров возрастают с увеличением длины ряда, если одновременно увеличивается число допустимых разрывов траектории модели. При этом длина сегмента, на котором соблюдается требование непрерывности траектории, должна быть тем меньше и ближе к ляпуновскому времени, чем больше количество разрывов.

Модифицированный метод имеет ряд преимуществ по сравнению с алгоритмом Бока, поскольку он свободен от ограничений, связанных с высокой чувствительностью к начальным условиям: накапливающиеся вдоль фазовой траектории невязки отбрасываются в разрешенных этим подходом разрывах. Он позволяет использовать ряды большой длины, что дает возможность уточнить оценки параметров за счет введения в рассмотрение дополнительного объема данных, причем такие оценки статистически предпочтительнее, чем получаемые с помощью кусочного метода множественной стрельбы (п. 2.2.2). При использовании рядов фиксированной длины он в большинстве случаев предъявляет меньшие требования к стартовым догадкам для искомых параметров, а иногда позволяет получить достаточно точные оценки, когда с помощью алгоритма Бока это невозможно сделать ни при каких стартовых догадках, даже равных истинным значениям.

Показано, что влияние шума (вплоть до 0.2 от уровня сигнала) на процесс моделирования при достаточных длинах ряда не существенно для обеих методик. Следует сказать, что точность оценки параметров при условии попадания в глобальный минимум целевой функции и для выбранной длины ряда наибольшая для

алгоритма Бока, предъявляющего самые жесткие требования при подгонке, меньше для модифицированного метода, и еще меньше для кусочного метода множественной стрельбы. Если же учесть, что попасть в глобальный минимум гораздо проще при использовании модифицированного метода, то он в итоге имеет практические преимущества даже в смысле точности оценок. Но следует учитывать и то, что, так как модифицированный метод предъявляет меньшие требования к адекватности используемой модели, существует опасность при слишком малом размере сегментов успешно подогнать под наблюдаемый ряд «чужую» модель.

Работа выполнена при поддержке грантов РФФИ (05-02-16305), CRDF (REC-006) и гранта Президента РФ для молодых ученых (МК-1067.2004.2).

Библиографический список

1. *Льонг.* Идентификация систем. Теория для пользователей. - М.: Наука, 1991. 432 с.
2. *Макаренко Н.Г.* Эмбедология и нейропрогноз // Лекции по нейроинформатике, М.: МИФИ, 2003. С. 2-67.
3. *Horbelt W., Timmer J., Bunner M.J., Meucci R. and Ciofini M.* Identifying physical properties of a CO₂ laser by dynamical modeling of measured time series // Phys. Rev. E, 2001. Vol.64. P. 016222.
4. *Hegger R., Kantz H., Schmuser F., Diestelhorst M., Kapsch R.-P., and Beige H.* Dynamical properties of a ferroelectric capacitors observed through nonlinear time series analysis // Chaos, 1998. Vol. 8, № 3. P. 727-754.
5. *Timmer J., Rust H., Horbelt W., and Voss H.* Parametric, nonparametric and parametric modelling of a chaotic circuit time series// Phys. Lett. A. 2000. Vol. 274. P. 123-134.
6. *Tokuda I., Parlitz U., Illing L., Kennel M. and Abarbanel H.D.I.* Parameter estimation for neurons // Experimental Chaos, Proceedings of the 7th Experimental Chaos Conference, San Diego, USA, 2002.
7. *Swameye I., Muller T.G., Timmer J., Sandra O., and Klingmuller U.* Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by databased modeling // Proc. Natl. Acad. Sci. USA. 2003. Vol. 100. P. 1028-1033.
8. *Bock H.* in: K. Ebert, P. Deuffhard W. Jager (Eds.), Modelling of Chemical Reaction Systems, Springer Berlin, 1981. Vol. 18. Chap. 8. P. 102-125.
9. *Baake E., Baake M., Bock H.G., and Briggs K.M.* Fitting ordinary differential equations to chaotic data // Phys. Rev. A. 1992. Vol. 45. No. 8. P. 5524-5529.
10. *Pisarenko V.F., Sornette D.* Statistical methods of parameter estimation for deterministically chaotic time series. Phys. Rev. E. 2004. Vol. 69. P. 036122.
11. *Horbelt W.* Maximum likelihood estimation in dynamical systems: PhD thesis, University of Freiburg, 2001. <http://webber.physik.uni-freiburg.de/horbelt/diss/>.

*Саратовский государственный
университет
Саратовское отделение ИРЭ РАН*

Поступила в редакцию 22.11.2004

RECONSTRUCTION WITH HIDDEN VARIABLES: MODIFIED BOCK'S APPROACH

B.P. Bezruchko, D.A. Smirnov, I.V. Sysoev

The task of model parameter estimation from experimental time series under the assumption that some variables cannot be measured or their series are very noisy is considered. Based on proposed quantitative criteria two algorithms – original Bock's approach and its modification that allows discontinuity of model trajectory are compared. The modified method is shown to be significantly more efficient for long series. Also with the modified approach the worse starting guesses for parameters to be estimated are acceptable.



Сысоев Илья Вячеславович – выпускник лицея Колледжа прикладных наук СГУ. Соросовский студент факультета нелинейных процессов СГУ, стипендиат Президента и Правительства России, а также Фонда некоммерческих программ «Династия». С 2004 года аспирант и по совместительству инженер НИИ ФНС СГУ, а также соисполнитель ряда грантов РФФИ, CRDF, Министерства образования и науки РФ. Область научных интересов: математическое моделирование по временным рядам радиофизических и биологических систем в условиях дефицита информации об объекте. В том числе разработка подходов в случае, когда часть переменных не может быть измерена. Автор 16 публикаций в отечественных и зарубежных изданиях.