УДК 681.511.4

О МНОГОЭКСТРЕМАЛЬНОСТИ В ЗАДАЧАХ ОЦЕНИВАНИЯ СОСТОЯНИЯ СИСТЕМ ДЕТЕРМИНИРОВАННОГО ХАОСА

С.М. Елсаков, В.И. Ширяев

ON MULTIEXTREMALITY IN THE ESTIMATION PROBLEMS OF DETERMINISTIC CHAOS SYSTEMS

S.M. Elsakov, V.I. Shiryaev

Рассматриваются нелинейные динамические системы, привлекаемые либо для описания протекания некоторых процессов, либо для описания остатков при использовании линейных моделей. Для этих систем формулируется задача оценивания состояния (а также параметрической идентификации) на основе метода максимального правдоподобия. Для решения поставленной задачи предлагается использовать однородные алгоритмы глобальной оптимизации. Приводятся основные теоремы о линейном виде критерия для однородных алгоритмов глобальной оптимизации, о достаточном условии сходимости однородного алгоритма глобальной оптимизации. Приводятся результаты оценивания состояния для одномерной динамической нелинейной системы.

Ключевые слова: детерминированный хаос, нелинейная система, глобальная оптимизация.

The nonlinear dynamic systems used for the modeling of some processes or remainders when using linear models are considered. The problem of state estimation (as well as parametric identification) on the basis of maximum likelihood method is formulated. To solve the problem it is proposed to use homogeneous algorithms of global optimization. The main theorems about the linear type of criterion for homogeneous algorithms of global optimization and about the sufficient condition of convergence of homogeneous algorithm of global optimization are given. The results of state estimation for one-dimensional dynamic nonlinear system are given.

Keywords: deterministic chaos, nonlinear system, global optimization.

Задачи прогнозирования различных процессов, как в технике, так и в экономике, и в социальной сфере могут решаться с применением методов теории автоматического управления. Однако существенное увеличение скорости протекания всех процессов, а также высокие требования к точности прогнозов вынуждают отказываться от использования линейных моделей и переходить к нелинейным моделям, поскольку многие процессы (рис. 1,2) имеют хаотический характер [1,2, 13,16-18, 19].

1. Системы детерминированного хаоса

Сейчас все более широкое применение для описания процессов в сложных системах находит теория детерминированного хаоса. Эта теория применяется и для описания процессов на финансовом рынке [1,2] и на рынке акций [16] и в теории связи [17, 18] и т.д. Системы детерминированного хаоса задаются в виде нелинейной динамической системы:

$$\begin{cases} x_{k+1} = F(x_k) + \xi_k; \\ y_{k+1} = G(x_{k+1}) + \eta_{k+1}, \end{cases} \qquad (1)$$

где $x_k \in R^N$ – вектор состояния системы, $y_k \in R^M$ – вектор наблюдения, $\xi_k \in R^N$ – шумы при движении системы, $\eta_k \in R^M$ – помехи в канале наблюдения. Эти системы могут являться описаниями процессов протекающих в самой системе, т.е. описывать хаотичные системы. А также могут применяться для описания систем, на которые действуют хаотичные возмущения, хотя сам объект описывается линейными уравнениями:

$$\begin{cases} x'_{k+1} = Ax'_k + \xi'_k; \\ y'_{k+1} = Cx'_{k+1} + \eta'_{k+1}, \end{cases} \qquad (2)$$

где $x'_k \in R^N$ – вектор состояния системы, $y'_k \in R^M$ – вектор наблюдения, $\xi'_k \in R^N$ – шумы

Елсаков Сергей Михайлович - аспирант кафедры прикладной математики ЮУрГУ; elsakov@prima.susu.ac.ru.

Ширяев Владимир Иванович - д.т.н., профессор, заведующий кафедрой прикладной математики ЮУрГУ; vis@prima.susu.ac.ru.

Elsakov Sergey Mikhailovich - post-graduate student of applied mathematics department of SUSU; elsakov@prima.susu.ac.ru.

Shiryaev Vladimir Ivanovich - PhD, professor, head of applied mathematics department of SUSU; vis@prima.susu.ac.ru.

при движении системы, $\eta'_k \in R^M$ – помехи в канале наблюдения.

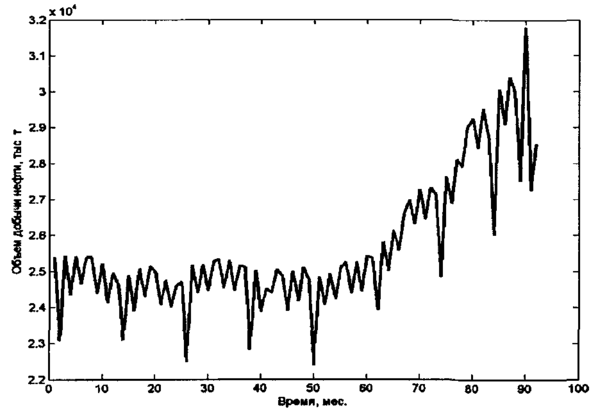

Рис. 1. Изменение объемов добычи нефти

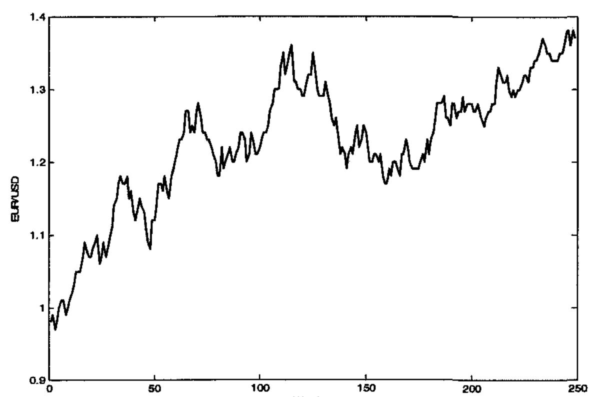

Рис. 2. Изменение курса евро к доллару

В этом случае вектор состояния можно расширить, включив в него хаотические переменные, тем самым, снизив уровни ошибок. Использование уравнений (1), (2) для описания процессов протекающих в некоторой системе возможно, только если ошибки ξ_k, η_k - являются белым шумом. Имея экспериментальные данные y_k, есть возможность воспользоваться арсеналом [5] методов идентификации, оценивания состояния линейных систем. Однако нередко на практике оказывается, что ошибки ξ_k, η_k представляют собой не белый шум. В этом случае применение вышеназванных методов является некорректным, а результаты недостоверными [9, 11, 12-14, 15].

При линеаризации нелинейной модели (1) в остатки кроме ошибок η_k вносится ошибка модели ζ_k, которая имеет детерминированную составляющую. В результате шумы η_k также перестают быть белыми. Кроме того, для линейных систем характерен принцип суперпозиции, в результате чего, двукратное применение линейной модели к остаткам линейной модели лишено смысла, для нелинейных моделей принцип суперпозиции не выполняется, что позволяет применять нелинейную модель несколько раз. Примеры удачного применения такого подхода можно найти в [4].

Применение подхода, основанного на позиции крайнего пессимизма [11, 12, 14] - минимаксного подхода, когда шумы представляются ограниченными, приводит к очень затрубленным результатам.

2. Постановка задачи оценивания состояния

Рассматривается задача идентификации (оценивания состояния) нелинейной динамической системы описываемой уравнениями:

$$\begin{cases} x_{k+1} = F(x_k) + \xi_k; \\ y_{k+1} = G(x_{k+1}) + \eta_{k+1}, \end{cases} \qquad (3)$$

где $x_k \in R^N$ – вектор состояния системы, $y_k \in R^M$ – вектор наблюдения, $\xi_k \in R^N$ – нормально распределенные с матрицей ковариаций Q шумы при движении системы, $\eta_k \in R^M$ – нормально распределенные с матрицей ковариаций R помехи в канале наблюдения.

В общем случае, неизвестные параметры этой системы могут быть включены в вектор состояния.

2.1. Рекуррентный алгоритм оценивания состояния с поиском глобального минимума

Рассмотрим применение метода максимального правдоподобия к системе (3). Пусть известны оценка xp_k вектора состояния x_k на k шаге, и выполнено измерения y_k. Пусть v_k реализовавшиеся значение вектора состояния x_k, а w_k – реализовавшиеся значение вектора ошибки ξ_k. Можно записать логарифмическую функцию правдоподобия

$$\begin{aligned} L(v_k, w_k) &= (v_k - x_k)^T D_k^{-1}(v_k - x_k) + \\ &+ (G(F(v_k) + w_k) - y_{k+1})^T R^{-1} \times \\ &\times (G(F(x_1) + x_2) - y_{k+1}) + w_k^T Q^{-1} w_k, \end{aligned} \qquad (4)$$

где D_k – матрица ковариаций вектора xp_k. Оценкой вектора x_{k+1} будет вектор $xp_{k+1} = F(v_k^0) + w_k^0$,

$$L(v_k^0, w_k^0) = \operatorname{abs}\min L(v_k, w_k), \qquad (5)$$

с матрицей ковариации $D_{k+1} = -\left[\partial^2 L(x) / \partial x_i \partial x_j\right]^{-1}$.

В случае, если функции $F(\cdot)$ и $G(\cdot)$ линейные, то задача (5) приводит к соотношениям фильтра Калмана для линейной системы. В случае нелинейности функции $F(\cdot)$ или $G(\cdot)$ задача (5) переходит в класс многоэкстремальных задач оптимизации. В этом случае распределение оценки xp_{k+1} будет отличным от нормального распределения, и матрица ковариации будет носить оценочный характер. Применение этого алгоритма будет на каждой итерации связано с потерей части информации о распределении оценки xp_{k+1} из-за аппроксимации ее нормальным распределением.

2.2. Рекуррентный алгоритм оценивания состояния с поиском всех локальных минимумов

Здесь при минимизации (4) с целью увеличения хранимой о функции распределения xp_{k+1} информации предлагается сохранять не только значение глобального минимума, но и значения всех локальных минимумов, т.е. фактически решается нелинейное уравнение

$$\operatorname{grad} L(v_k, w_k) = \bar{0}. \qquad (6)$$

Каждому найденному локальному минимуму ставится в соответствие своя оценка $xp^j_{k+1} = F(v^j_{k+1}) + w^j_{k+1}$, которая соответствует точке, в которой достигается локальный минимум функции (4). Матрица ковариаций в каждой точке выбирается, как и в предыдущем алгоритме. Кроме того каждой оценке ставится в соответствие вероятность, показывающая, что истинное значение оценивается именно этой оценкой. Вероятности выбираются на основе значения локальных минимумов, таким образом, чтобы в сумме они давали единицу. А минимальное значение локального минимума соответствовало максимальной вероятности.

2.3. Алгоритм оценивания состояния на основе нескольких последовательных измерений

В двух рассмотренных рекуррентных алгоритмах при переходе от шага к шагу происходит потеря информации о плотности распределения оценки. Эта потеря связана с тем, что плотность распределения не получается хранить в параметризованном виде, а требуется сохранять в виде функции. Чтобы избежать передачи плотности распределения с шага на шаг можно одновременно обрабатывать данные нескольких измерений. Для этого рассмотрим систему (3) на k шаге и будем совместно обрабатывать ρ измерений. Запишем логарифмическую функцию правдоподобия

$$L(xp_{k+1}, xp_k, \dots, xp_{k-p+2}) =$$
$$= \sum_{i=k-p+2}^{k+1} (G(x_i) - y_i)^T R^{-1} (G(x_i) - y_i) +$$
$$+ \sum_{i=k-p+2}^{k} (F(x_{i+1}) - x_i)^T Q^{-1} (F(x_{i+1}) - x_i). \qquad (7)$$

Запишем необходимое условие максимума в виде (8).

Предполагая, что $\partial F(x_i)/\partial x \neq 0$, из первого уравнения можно выразить переменную x_{k-p+3}, из следующих уравнений аналогично можно выразить остальные переменные, кроме переменной x_{k-p+2}. Затем можно подставить выраженные значения в функцию (7) - получится функция с количеством переменных равным количеству переменных состояния. Решая задачу глобальной оптимизации для этой функции можно определить наилучшую оценку по нескольким последним измерениям. Альтернативным путем решения указанной задачи может быть решения последнего уравнения из системы (8) с учетом выполненных подстановок. Матрица ковариаций для полученной оценки может быть получена описанным ранее способом.

$$\begin{cases} R^{-1}(G(x_{k-p+2}) - y_{k-p+2}) \dfrac{\partial G(x_{k-p+2})}{\partial x} + \\ +Q^{-1}(F(x_{k-p+2}) - x_{k-p+3}) \dfrac{\partial F(x_{k-p+2})}{\partial x} = 0; \\ \dots \\ R^{-1}(G(x_j) - y_j) \dfrac{\partial G(x_j)}{\partial x} + \\ +Q^{-1}(F(x_j) - x_{j+1}) \dfrac{\partial F(x_j)}{\partial x} + \\ +Q^{-1}(x_j - F(x_{j-1})) = 0, \ j = k-p+3 \dots k; \\ \dots \\ R^{-1}(G(x_{k+1}) - y_{k+1}) \dfrac{\partial G(x_{k+1})}{\partial x} + \\ +Q^{-1}(x_{k+1} - F(x_k)) = 0. \end{cases} \qquad (8)$$

3. Алгоритмы решения задач глобальной оптимизации

Рассматривая различные постановки задачи оценивания состояния системы детерминированного хаоса, в плане численных методов постоянно возникает задача многоэкстремальной (глобальной) оптимизации. Кроме того, среди трех постановок задач все задачи и рекуррентные и не рекуррентные свелись к задаче глобальной оптимизации одинаковой размерности. Важным этапом является выбор адекватного алгоритма для решения задачи многоэкстремальной оптимизации. Задачи глобальной оптимизации широко встречаются на практике [3, 6-8]. Среди алгоритмов глобальной оптимизации традиционно выделяют [8] детерминированные и стохастические алгоритмы. И те и другие алгоритмы организуют процесс порождения точек допустимого множества, в которых необходимого провести вычисления целевой функции в зависимости от проведенных вычислений. Одно такое вычисление обычно называют *испытанием*. Детерминированные алгоритмы организуют процесс порождения этих точек детерминированным образом, а стохастические используют датчик псевдослучайных чисел.

Однородным называют такой алгоритм глобальной оптимизации, который для двух функций различающихся только на константу порождает одинаковые последовательности точек испытаний [7].

Для того, чтобы алгоритм учитывал особенности целевой функции, можно использовать модели целевой функций, которые могут быть представлены в функциональном виде, а именно при проведении k испытаний определены две функции: $m_k(x)$ и $s_k(x)$ к которым будем предъявлять следующие требования:

- $m_k(x_i) = f(x_i), i = \overline{1,k}$;
- $s_k(x_i) = 0, i = \overline{1,k}$;
- $s_k(x) > 0, x \neq x_i, i = \overline{1,k}$;
- $m_k(x)$ и $s_k(x)$ – липшицевы с константой Липшица L_m, $k = 1,\ldots$ и L_s, $k = 1,\ldots$ соответственно.

Сам алгоритм глобальной оптимизации задается в виде

$$x_{k+1} = \operatorname{abs}\max_{x\in X} P\left(s_k(x), m_k(x)\right).$$

Для однородных алгоритмов доказаны теоремы:

Теорема 1. Если критерий $P\left(s_k(x), m_k(x)\right)$ дважды непрерывно дифференцируемый и для функций $m_k(x)$ и $s_k(x)$ выполняются дополнительные условия:

- $m_k^f(x) = m_k^g(x) + C$ ($f(x) = g(x) + C$);
- $s_k^f(x) = s_k^g(x)$ ($f(x) = g(x) + C$),

то для однородного алгоритма критерий $P\left(s_k(x), m_k(x)\right)$ может быть представлен в виде

$$P\left(s_k(x), m_k(x)\right) = Cm_k(x) + p(s_k(x)),$$

где $C = \text{const}$ и $p(s_k(x))$ – некоторая функция.

Теорема 2. Для того чтобы множество предельных точек порождаемых однородным алгоритмом с критерием $P(s_k(x), m_k(x))$ совпадало с множеством глобальных минимумов липшицевой функции $f(x)$ с константой Липшица L на компакте X достаточно, чтобы

$$-\max_{x\in X}\left(P(s_k(x), m_k(x))\right) \le \min_{x\in X}\max_{i=\overline{1,k}}\left(f(x_i) - K\|x - x_i\|\right),$$

где $K > L$.

Теорема 3. Если для функции $s_k(x)$ выполняется условие

$$s_k(x) \ge \min_{i=\overline{1,k}}\|x - x_i\|,$$

то множество предельных точек порождаемых линейным алгоритмом с критерием $Ks_k(x) - m_k(x)$, где $K > L + L_m$, будет совпадать с множеством глобальных минимумов липшицевой функции $f(x)$ с константой Липшица L.

В рамках однородных алгоритмов могут быть представлены некоторые существующие алгоритмы, а также были построены новые алгоритмы на основе триангуляции Делоне, реберно-симплексных разверток и т.д. [7].

Для решения поставленных задач можно привлекать однородные алгоритмы в зависимости от размерности решаемой задачи.

4. Пример

Рассмотрим систему [13], задаваемую одномерным отображением

$$\begin{cases} x_{k+1} = 1 - ax_k^2 + \xi_k; \\ y_{k+1} = x_{k+1} + \eta_k, \end{cases} \quad (9)$$

где ξ_k – случайная величина с дисперсией R, а η_k – с дисперсией Q.

Для сравнения работы рассматривалось два алгоритма оценивания состояния: фильтр Калмана примененный к линеаризованной модели и не рекуррентный алгоритм с обработкой трех последних измерений. Параметр a был выбран $a = 1{,}85$. Интенсивности шумов изменялись. Для каждой интенсивности шумов выбиралось случайное начальное значение x_0. Начальное приближение для фильтра Калмана также выбиралось случайно. На рис. 3 изображен пример ряда получаемого с помощью уравнения (9), $R = Q = 0{,}1$. На рис. 4, представлена зависимость ошибки оценивания состояния при различных интенсивностях шумов.

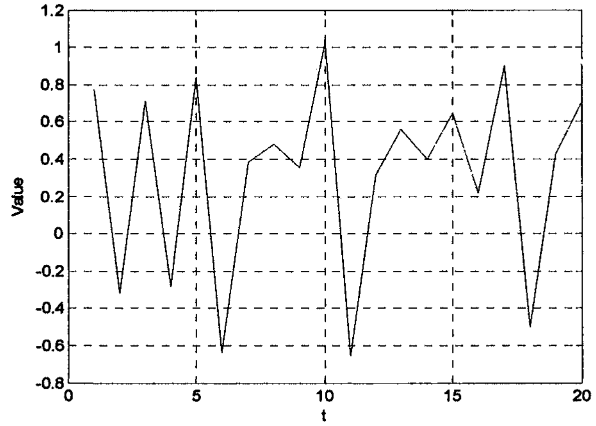

Рис. 3. Пример хаотического ряда

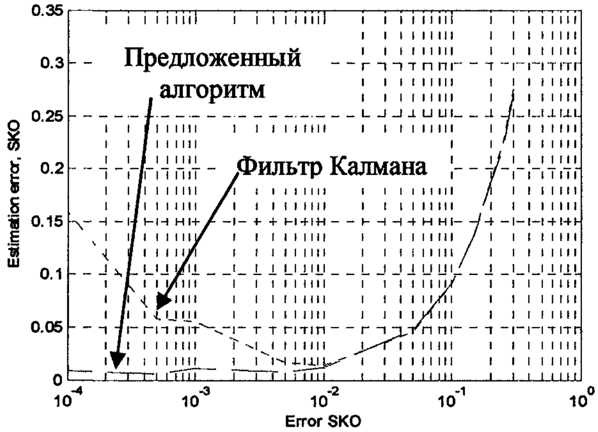

Рис. 4. Зависимость точности оценивания от уровня помех

Совместная обработка показала лучший результат при использовании малых шумов, когда ошибки при линеаризации модели для применения фильтра Калмана велики. В этом случае с одной стороны шумы подаваемые на фильтр Калмана являются не белыми, а с другой присутствует неточность модели.

При больших шумах эффективность алгоритмов практически одинакова. Однако с увеличением размерности и, как следствие, усложнением нелинейной модели, увеличивается порог шумов до которых алгоритм совместного распределения является лучшим.

Выводы

Использование моделей детерминированного хаоса является необходимостью при моделировании сложных объектов, а также при моделировании остатков линейных объектов. Задача оценивания состояния системы детерминированного хаоса приводит к многоэкстремальной задаче оптимизации, которая может быть решена с использованием однородных алгоритмов глобальной оптимизации. При численном решении задачи оценивания состояния системы (9), рассмотренный алгоритм показывает лучшие результаты до некоторого (небольшого) уровня шума, а затем демонстрирует результаты схожие с результатами применения фильтра Калмана.

Литература

1. Mandelbrot, B. A Multifractal Walk Down Wall Street/B. Mandelbrot/IScientific American. -1999.

2. Peters, E. Fractal Market Analysis. Applying Chaos Theory to Investment & Economics / E. Peters. -New York: John Wiley & Sons, 1994.

3.Антонов, М. О. Нахождение оптимального расположения радиомаяков в разностно-дальномерной системе посадки летательного аппарата / М. О. Антонов, С. М. Елсаков, В. И. Ширяев // Авиакосмическое приборостроение. — 2005. — №11. -С. 41-45.

4. Афанасьева, К. Е. Идентификация состояния и прогнозирование регионального рынка / К. Е. Афанасьева, В. И. Ширяев // Проблемы управления. -2007. -№3.

5. Бек, В. В. Интегрированные системы терминального управления / В. В. Бек, Ю. С. Вишняков, А. Р. Махлин. -М: Наука, 1989. -224 с.

б.Гергель, В. П. Современные методы принятия оптимальных решений / В. П. Гергель и др. - Н. Новгород: Изд-во ННГУ, 2002. - 189 с.

7. Елсаков, С. М. Об однородных алгоритмах глобальной оптимизации / СМ Елсаков, В. И. Ширяев // Информационный бюллетень ассоциации математического программирования: материалы 13-й Всероссийской конференции «Математическое программирование и приложения», Екатеринбург, 26 февраля - 2 марта 2007 г. -2007.-№11.-С.37-38.

8.Жиглявский, А. А. Методы поиска глобального экстремума / А. А. Жигляевский, А. Г. Жилинскас. - М.: Наука, 1991. - 248 с.

9.Калман Р.Е. Идентификация систем с шумами //Успехи математических наук, 1985. - Т.40, вып. 4(244).-С. 27-41.

10. Красовский, Н. Н. Управление и стабилизация при недостатке информации / Н. Н. Красовский // Изв. РАН. Техническая кибернетика. — 1993. -Ns 1. -С. 148-151.

П.Кунцевич, В. М. Гарантированное оценивание фазового состояния и параметров линейных динамических систем / В. М Кунцевич // Проблемы управления и информатики. - 2005. - № 5. - С 18-25.

12. Куржанский, А. Б. Идентификация нелинейных процессов - гарантированные оценки / А. Б. Куржанский, В. Д. Фурасов // Автоматика и телемеханика. -1999. -№6.-С. 70-87.

13. Смирнов, Д. А. Метод оценки параметров одномерных отображений по хаотическим рядам / Д. А. Смирнов, В. С. Вläскин, В. И. Пономаренко //ПисьмаβЖТФ. -2005. -Т. 31, №3.

14. Черноусько, Ф. Л. Об оптимальном эллипсе и дальнем оценивании для динамических систем, подверженных неопределённым возмущениям / Ф. Л. Черноусько //Кибернетика и системный анализ. - 2002. -№2.- С. 85-94.

15. Ширяев, В. И. Синтез управления линейными системами при неполной информации / В. И. Ширяев //Изв. РАН. Техническая кибернетика. -1994. -№3.- С. 229-237.

16. Никуличев, Е. В. Модели хаоса для процессов изменения курса акций / Е. В. Никуличьев // Exponenta Pro. Математика в приложениях. - 2003.-№1(1).-С49-52.

17.Leland, W.E. On the self-similar nature of ethernet traffic (extended version) / W. E. Lelandet al. //IEEE/ACM Transactions of Networking. - 1993. - №2(1).-P. 1-15.

18. Осин, А. В. Влияние самоподобности речевого трафика на качество обслуживания в телекоммуникационных сетях / А. В. Осин // Автореф. дис. канд. техн. наук. - 2005.

19. Красовский, А. А. Аттракторы и синтез управления в критических режимах /А. А. Красовский // Изв. РАН. Теория и системы управления. - 1996.-№3.-С. 5-14.

Поступила в редакцию 6 октября 2007 г.