Изв. вузов «ПНД», т.10, № 1-2, 2002 УДК 681.3, 519.67

УСТОЙЧИВОСТЬ ЧИСЛЕННЫХ ОЦЕНОК ХАРАКТЕРИСТИК ВРЕМЕННЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

К. Е. Бобров, А. М. Искольдский

Обсуждаются численные методы анализа данных (конечных упорядоченных последовательностей натуральных двоичных кодов), отвечающих фрагментам траекторий, полученных численным решением конечного числа нелинейных обыкновенных дифференциальных уравнений. Эти уравнения представляют детерминированные диссипативные хаотические динамические системы.

Измеряемые свойства таких последовательностей характеризуются оценкой, которой считается код, полученный в результате переработки входной последовательности данных по заданному алгоритму, реализованному на ЭВМ и не предполагающего участия эксперта. Формализуется понятие устойчивости оценки. Исследуется устойчивость получаемых оценок по отношению к моделируемым численно малым вариациям параметров схемы измерения, а также - к параметрам алгоритма обработки. Рассматриваются примеры последовательностей данных, параметры которых (разрядность, шаг по времени, длина последовательности) являются типичными для многих реальных экспериментов.

Показано, что оценка, получаемая по алгоритму, основанному на анализе свойств, существенно зависящих от поведения траектории в каждой точке, является неустойчивой. В то же время оценка, получаемая по алгоритму, основанному на анализе свойств, определяемых некоторым «осредненным» для различных точек поведением траектории, является устойчивой.

Введение

Процедуры, описанные в данной работе, имеют отношение к анализу маломодовых моделей быстропротекающих процессов в электрофизике (см., например, [1, 2]). Представляет интерес рассмотреть последовательности данных, параметры которых (разрядность, шаг по времени, длина последовательности) являются типичными для этих (и многих других) реальных экспериментов.

При вычислении оценок величин, характеризующих временные последовательности, важно формализовать и определить их устойчивость по отношению к малым (в определенном ниже смысле) изменениям режимов регистрации. Мы будем моделировать численно следующие изменения режимов регистрации: шага квантования по времени, квантования по амплитуде, длины последовательности. Проверка устойчивости оценок, получаемых в результате работы вычислительных процедур, позволяет сделать вывод о том, является ли эта величина характеристикой данной последовательности. Оценка не должна зависеть от параметров процедуры обработки, а сами процедуры не должны предполагать участия эксперта.

В работе исследуются конечные упорядоченные последовательности кодов, полученных в результате численного решения нелинейных дифференциальных уравнений, представляющих динамическую систему, которая находится в режиме детерминированного хаоса. Результаты численного или натурного экспериментов представляют собой упорядоченные последовательности кодов (отсчетов), которые интерпретируются как целые (конструктивные) числа. В случае натурного эксперимента аналого-цифровой преобразователь (АЦП) выдает на каждом такте код, интерпретируемый как одно из целых чисел в диапазоне, определяемом разрядностью АЦП. В случае численного эксперимента формат чисел, получаемых на каждом такте работы расчетного алгоритма, может быть и форматом с плавающей точкой. Множество таких чисел конечно. Для удобства будем производить нормировку и округлять до ближайшего целого числа данные, выдаваемые программой решения системы дифференциальных уравнений.

Предполагаем, что оценка, получаемая по алгоритму, основанному на анализе свойств, существенно зависящих от поведения траектории в каждой точке, будет неустойчивой. В то же время оценка, получаемая по алгоритму, основанному на анализе свойств, определяемых некоторым «осредненным» для различных точек поведением траектории, будет устойчивой.

1. Реализации и процедуры вычислений

В результате преобразования, осуществляемого АЦП, получаем упорядоченные двоичные коды разрядностью k (k - разрядность АЦП). В численном эксперименте используем разностные схемы решения дифференциальных уравнений. Далее применяется стандартная процедура округления числа до ближайшего целого в сочетании с процедурой нормировки на диапазон кодов, определяемых заданной разрядностью. Для моделирования перехода от k-разрядного АЦП к АЦП с меньшей разрядностью используем процедуру сдвига вправо исходных кодов. Для получения последовательностей с более редким шагом применяется процедура прореживания.

Для изучения свойств аттрактора динамической системы по результатам регистрации конечного фрагмента траектории одной переменной, используется процедура Паккарда и Такенса [3, 4], основанная на теореме Такенса [5, 6]. Когда имеется только одна последовательность кодов конечной длины N, строится массив размерностью $(N - n \times m) \times n$, в котором каждая следующая последовательность получается сдвигом относительно исходной на некоторое фиксированное и одинаковое для всех сдвигов количество кодов m. Обычно n называется размерностью соответствующего псевдофазового портрета.

Будем называть строку длиной n массива, построенного по процедуре Паккарда и Такенса, вектором размерностью n.

При описанных ниже вычислениях обычно применяется понятие евклидовой метрики. В используемых нами процедурах реализуется иной критерий близости: векторы в n-мерном пространстве считаются близкими, если равны все коды, составляющие векторы. Возможность использования такого критерия близости демонстрируется, например, в [7].

Для вычисления оценки минимальной размерности вложения аттрактора динамической системы в [8, 9] предлагается нижеприведенный порядок действий, основанный на следующем положении. Назовем вектор, который используется для сравнения с другими векторами той же размерности, опорным. Если следующий за опорным вектором отсчет функционально связан с предыдущими, то, встретив близкий (по евклидовой метрике) к опорному вектор, мы вправе ожидать близости отсчетов, следующих непосредственно за опорным и близким к нему.

Для оценок используются величины: ρ - евклидово расстояние между опорным вектором, составленным из первых n отсчетов вектора размерностью $(n+1)$, и вектором, составленным из первых n отсчетов вектора размерностью $(n+1)$, сравниваемого с опорным; r - евклидово расстояние между $(n+1)$-м отсчетом опорного вектора и $(n+1)$-м отсчетом сравниваемого с ним вектора.

Строится график функциональной зависимости ρ от r. Затем оценивается степень достижения искомого числа (интерпретируемого как размерность вложения) по «монотонности огибающей сверху к этому графику в области малых ρ при повышении n», либо - по отсутствию пар векторов, у которых первые n отсчетов близки, а $(n+1)$-е - нет.

Для расчетов была предложена следующая процедура, основанная на положениях работ [8, 9].

Процедура 1. Производится сравнение с опорным вектором векторов размерностью n, составленных из первых n отсчетов векторов размерностью $(n+1)$. Если вектора размерностью n близки, проверяется близость $(n+1)$-х отсчетов векторов размерностью $(n+1)$. Данное n считается искомой оценкой, если сочетания векторов, построенных из данной последовательности, у которых первые n отсчетов близки, а $(n+1)$-е - нет, отсутствуют.

Пары, для которых первые n отсчетов близки, а $(n+1)$-е - нет, получили название «ложных ближайших соседей» (false nearest neighbours, FNN [10 - 12]) . В этих же работах для получения искомой оценки предлагается оценивать количество FNN, получающихся на каждом шаге при увеличении n. Когда это число становится малым и меняется слабо при увеличении n, данное n считается искомой оценкой. В наших расчетах на основании положений этих работ была предложена следующая более «гибкая» процедура.

Процедура 2. Также, как и в предыдущем случае, сравниваются векторы размерностью n. Если при увеличении n доля FNN не превышает некоторое, наперед заданное число (верхний порог) и меняется при каждом последующем увеличении n не более, чем в некоторое другое наперед заданное число раз (нижний порог), данное n считается искомой оценкой.

Если результат работы процедуры является устойчивым, можно заключить, что он характеризует обрабатываемую последовательность данных. В этом случае результаты работы двух описанных выше процедур можно рассматривать как оценку минимальной размерности вложения аттрактора динамической системы, от которой получена данная последовательность кодов.

Для получения оценки размерности вложения аттрактора динамической системы в работах [13-16] предлагается такая схема действия.

Исследуется зависимость энтропийной или корреляционной размерности при увеличении размерности псевдофазового портрета, который строится из n последовательностей, полученных сдвигом. Строится график зависимости логарифма числа разных векторов размерности n или корреляционной суммы для векторов размерности n от логарифма максимального расстояния, при котором вектора в n-мерном пространстве еще считаются близкими в евклидовой метрике. Такое построение делается для некотрого набора этих максимальных расстояний. Поделив логарифм числа разных векторов размерности n или логарифм отношения корреляционной суммы к квадрату числа векторов N для векторов размерности n на логарифм максимального расстояния ε, можно получить числа, интерпретируемые соответственно как оценка энтропийной или корреляционной размерности. Эти величины определяют характерный наклон графиков. Затем предлагается определить число n, начиная с которого графики этих зависимостей идут параллельно и близко друг от друга, и это число называется в качестве оценки для минимальной размерности вложения.

Используя положения работ [11-14, 6], мы разработали и применили в расчетах следующую процедуру.

Процедура 3. При фиксированной разрядности n подсчитывается число разных векторов $\tilde{N}$ размерностью n. Затем логарифм по основанию 2 от этого числа делится на разрядность, то есть на логарифм по основанию 2 от $1/\varepsilon$

$$d_E=[\ln(\tilde{N})/\ln(1/\varepsilon)] = [\log_2(\tilde{N})/\log_2(e)]/[\log_2(1/\varepsilon)\ \log_2(e)]= (1)/k)\log(\tilde{N}).$$

Здесь $1/\varepsilon$ - число кодов разрядности k. Как и в предыдущем случае находится такая размерность массива n, при которой d_E изменяется не больше, чем на некоторую наперед заданную величину порога l. Получившееся число n - «первый результат» *процедуры 3.* Конструктивное действительное число d_E, соответствующее той размерности массива n, при которой d_E меняется не больше, чем на 1 - «второй результат» *процедуры 3.*

Если оценка, получаемая в результате применения *процедуры 3* к данной последовательности кодов, устойчива, она может рассматриваться как характеристика этой последовательности. Результаты ее работы можно рассматривать как оценки минимальной размерности вложения аттрактора соответствующей динамической системы и энтропийной размерности аттрактора динамической системы, от которой получена данная последовательность кодов.

При эксперименте необходимо выбрать оптимальные для работы процедур обработки параметры регистрации - длину последовательности N, шаг квантования сигнала по времени τ, разрядность АЦП k, а также параметры процедур обработки последовательностей - величины сдвига m и порога l. Получаемые оценки должны быть устойчивы в определенном ниже смысле. Если для данной модели оцениваемая величина известна априори, то представляет интерес ее сравнение с полученной оценкой.

Под устойчивостью результатов *процедур 1* и *2*, а также первого результата *процедуры 3* будем понимать ее неизменность при независимых, малых в определенном ниже смысле вариациях N, τ, k, m, l. Под устойчивостью второго результата *процедуры 3* будем понимать его относительное изменение не более, чем на $1/k$ при таких же вариациях этих параметров.

На основании анализа результатов численных экспериментов и априорных предположений об оптимальных параметрах регистрации и процедур обработки сформулирован ряд требований к последовательностям.

- Мгновенное значение амплитуды сигнала не должно выходить за допустимые границы.
- Число отсчетов N в последовательности для *процедуры 1* равно 2^{n+k}, а для *процедур 2* и *3* - 2^{nk}.
- Шаг по времени τ выбирается таким образом, чтобы при заданной длине N в последовательности было не менее двух экстремумов.
- Последовательности с разрядностью по амплитуде ниже 4 не рассматриваются.
- После получения последовательности m выбирается как число отсчетов между экстремумами ближайшими к началу. Для *процедуры 2* значение верхнего порога выбирается равным $1/\log_2(N)$, а значение нижнего порога - равным 2. Для *процедуры 3* значение порога l выбирается как $2/k$.

Установленные требования обеспечивают соответствие псевдофазового портрета «случаю общего положения». В частности, портрет должен быть достаточно «представительным» (требования к длине последовательности) и «невырожденным», например, в диагональ (требования к величине сдвига и шагу по времени).

Если эти требования не удовлетворяются, данная последовательность считается непригодной для работы дальнейших процедур обработки, так как при работе с такой последовательностью не следует ожидать устойчивости результатов и их близости априорно известным характеристикам соответствующей гладкой динамической системы.

2. Рассматриваемые модели

2.1. Система Лоренца [17]:

$$\dot{X} = \sigma(Y - X),$$
$$\dot{Y} = -XZ + rX - Y,$$
$$\dot{Z} = XY - bZ.$$

При σ=10, b=8/3, r=26 система находится в режиме странного аттрактора. Рассматривается реализация X-компоненты системы. Энтропийная размерность странного аттрактора динамической системы в данном режиме равна 2.08 [15, 16], минимальная размерность вложения - 3.

2.2. Система Рассела - Хансона [18]:

$$\dot{X} = Y(Z - 1 + Y^2) + \gamma X,$$
$$\dot{Y} = X(3Z + 1 - X^2) + \gamma Y,$$
$$\dot{Z} = -2Z(\nu + XY).$$

При ν=1.1, γ=0.87 система находится в режиме странного аттрактора. Рассматривается реализация X-компоненты системы. Энтропийная размерность странного аттрактора динамической системы в данном режиме равна 2.38 [18], минимальная размерность вложения - 3.

2.3. Система Ресслера [19]:

$$\dot{X} = -Z - Y,$$
$$\dot{Y} = X + aY,$$
$$\dot{Z} = b + Z(X - c).$$

При a=0.15, b=0.2, c=10 система находится в режиме странного аттрактора. Рассматривается реализация X-компоненты системы. Энтропийная размерность странного аттрактора динамической системы в данном режиме равна 2.03 [20], минимальная размерность вложения - 3.

3. Результаты вычислений

3.1. Процедура 1.

Система уравнений	Разрядность	Длина последовательности		
	k	$N/2$	N	$N \times 2$
Лоренца	6		4	
	5	3	6	8
	4		7	
Рассела - Хансона	6		5	
	5	4	6	9
	4		8	
Ресслера	6		5	
	5	4	6	9
	4		8	

Оценки не являются устойчивыми по отношению к выбранным изменениям разрядности данных и длине последовательности ни в одном из трех рассматриваемых примеров. На основании этого полагаем, что данные оценки не могут характеризовать выбранные последовательности.

3.2. Процедура 2.

Система уравнений	Разрядность	Длина последовательности		
	k	$N/2$	N	$N \times 2$
Лоренца	6		3	
	5	3	3	5
	4		4	
Рассела - Хансона	6		3	
	5	3	3	4
	4		5	
Ресслера	6		3	
	5	3	3	3
	4		3	
Ресслера	Шаг по времени	Сдвиг для построения последовательности		
	τ	$m/2$	m	$m \times 2$
	$\tau\times2$		3	
	τ	3	3	3
	$\tau/2$		3	

Только в последнем из трех рассматриваемых примеров оценка является устойчивой по отношению к выбранным изменениям разрядности данных, длины последовательности, шага по времени и величины сдвига. Дальнейшие численные эксперименты показали также устойчивость данной оценки (в последнем примере) при l=3 и l=5. На основании этого полагаем, что лишь в этом примере данная оценка может характеризовать выбранную последовательность.

3.3. Процедура 3. Первый результат.

Система уравнений	Разрядность	Длина последовательности		
	k	$N/2$	N	$N \times 2$
Лоренца	6	3		
	5	3	3	3
	4		3	
Рассела - Хансона	6		3	
	5	3	3	3
	4		3	
Ресслера	6		3	
	5	3	3	3
	4		3	
	Шаг по времени	Сдвиг для построения последовательности		
	τ	$m/2$	m	$m \times 2$
	$\tau\times2$		3	
Ресслера	τ	3	3	3
	$\tau/2$		3	

Дальнейшие численные эксперименты во всех трех рассматриваемых примерах также показали устойчивость данной оценки при l=3 и l=5. На основании этого полагаем, что полученные для этих примеров оценки могут характеризовать выбранные последовательности.

Процедура 3. Второй результат.

Система уравнений	Разрядность	Длина последовательности		
	k	$N/2$	N	$N \times 2$
Лоренца	6		2.61	
	5	2.57	2.59	2.59
	4		2.54	
	Шаг по времени	Сдвиг для построения последовательности		
	τ	$m/2$	m	$m\times2$
	$\tau\times2$		2.58	
	τ	2.59	2.59	2.62
	$\tau/2$		2.60	
Рассела - Хансона	6		2.39	
	5	2.40	2.42	2.44
	4		2.47	
	Шаг по времени	Сдвиг для построения последовательности		
	τ	$m/2$	m	$m\times2$
	$\tau\times2$		2.43	
	τ	2.41	2.42	2.42
	$\tau/2$		2.44	
Ресслера	6		2.49	
	5	2.50	2.51	2.54
	4		2.58	
	Шаг по времени	Сдвиг для построения последовательности		
	τ	$m/2$	m	$m\times2$
	$\tau\times2$		2.53	
	τ	2.51	2.51	2.52
	$\tau/2$		2.54	

Во всех рассматриваемых примерах изменения оценок не превышают $^1/_5$. На основании этого полагаем, что данные оценки могут характеризовать выбранные последовательности.

Выводы

Получены оценки путем переработки входных последовательностей данных по заданному алгоритму, реализованному на ЭВМ и не предполагающему участие эксперта. Исследована устойчивость оценок по отношению к численно моделируемым малым вариациям параметров схемы измерения, а также - к параметрам процедур обработки. Рассмотрены примеры последовательностей данных, параметры которых (разрядность, шаг по времени, длина последовательности) являются типичными для многих реальных экспериментов.

Показано, что оценка, получаемая по алгоритму, основанному на анализе свойств, существенно зависящих от поведения траектории в каждой точке таких последовательностей данных (то есть локальных свойств), является неустойчивой. В то же время оценка, получаемая по алгоритму, основанному на анализе свойств, определяемых некоторым «осредненным» для различных точек поведением траектории, является устойчивой. *Процедуры 1* и *2* обработки данных можно отнести к первому типу, *процедуру 3* - ко второму. *Процедура 2* демонстрирует, тем не менее, лучшую устойчивость результатов, чем *процедура 1*. Кроме того, анализ результатов численных экспериментов, выполненных непосредственно по другим известным методам, в частности, упомянутым в тексте, также подтвердил справедливость этих выводов. Оценки минимальной размерности вложения, полученные по *процедурам 3* и *2* для численной модели системы уравненийРесслера, совпадают с априорно известными. Отклонения оценок величины d_E, получаемых по *процедуре 3*, от априорно известных связаны, по нашему мнению, с весьма низкой разрядностью рассматриваемых последовательностей данных. Это обстоятельство необходимо учитывать при анализе данных реальных экспериментов, для которых получение таких последовательностей является типичным.

В предлагаемых процедурах операции с плавающей точкой либо отсутствуют, либо их количество минимально. Это позволяет легко реализовать соответствующие вычисления с использованием алгоритмов и аппаратных структур распараллеливания при обработке больших объемов данных.

Библиографический список

1. *Волков Н. Б., Искольдский А. М*. Об аналогии между начальными стадиями зарождения турбулентности и электрического взрыва проводников // Письма в ЖЭТФ. 1990. Т. 51, вып. 11. С. 560.

2. *Волков Н. Б., Зубарев Н. М., Зубарева О. В., Шкатов В. Т.* Динамическое прерывание тока и вихревые структуры в токонесущей плазмоподобной среде // Письма в ЖТФ. 1996. Т. 22, вып. 13. С. 43.

3. *Packard N. H., Crutchfield J. P., Farmer J. D., Shaw R. S.* Geometry from a Time Series // Phys. Rev. Lett. 1980. Vol. 45. P. 712.

4. *Broomhead D. S. and King G. P.* Extracting qualitative dynamics from experimental data // Physica D. 1986. Vol. 20, №. 2. P. 217.

5. *Takens F.* Detecting strange attractors in turbulence / Lect. Notes in Math. Vol. 898. N. Y.: Springer, 1981. P. 366.

6. *Sauer T., Yorke J., Casdagli M.* Embedology // Journal of Statistical Physics. 1991. Vol. 65. P. 579.

7. *Малинецкий Г. Г, Потапов А. Б.* Современные проблемы нелинейной динамики. М.: Эдиториал УРСС, 2000. 336 с.

8. *Лукащук С. Н., Предтеченский А. А., Фалькович Г. Е., Черных А. И.* О

вычислении размерности аттракторов по экспериментальным данным. Препринт № 280. Институт автоматики и электрометрии. Новосибирск, 1985.

9. *Лукащук С. Н., Фалькович Г. Е., Черных А. И.* О вычислении размерности аттракторов по экспериментальным данным // ПМТФ. 1989. № 1(173). С. 99. СО АН СССР.

10. *Kennel M. B., Brown R., Abarbanel H. D. I.* Determining embedding dimension for phase-space reconstruction using a geometrical construction // Phys. Rev. A. 1992. Vol. 45, № 6. P. 3403.

11. *Abarbanel H. D. I., Kennel M. B.* Local false nearest neighbors and dynamical dimensions from observed chaotic data // Phys. Rev E. 1993. Vol. 47. P. 3057.

12. *Schreiber T.* Efficient neighbor searching in nonlinear time series analysis // Int. J. Bifurcations and Chaos. 1995. Vol. 5. P. 349.

13. *Grassberger P., Procaccia I.* Estimation of the Kolmogorov Entropy from a Chaotic Signal // Phys. Rev. A. 1983. Vol. 28. P. 2591.

14. *Grassberger P., Procaccia I.* Characterization of Strange Attractors // Phys. Rev. Lett. 1983. Vol. 50. P. 346.

15. *Grassberger P., Procaccia I.* Measuring the Strangeness of Strange Attractors // Physica D. 1983. Vol. 9. P. 189.

16. *Grassberger P., Procaccia I.* Dimensions and Entropies of Strange Attractors from a Fluctuating Dynamics Approach // Physica D. 1983. Vol. 13. P. 34.

17. *Лоренц Э.* Странные аттракторы. М.: Мир. 1981. С. 88.

18. *Russel D. A., Hanson J. D., Ott E. //* Phys. Rev. Lett. 1980. Vol. 45. P. 1175.

19. *Rössler O. E.* // Phys. Lett. A. 1976. Vol. 57. P. 397.

20. *Wolf A., Swift J. B., Swinney H. L. and Vastano J. A.//* Physica D. 1985. Vol. 16. P. 285.

Институт электрофизики
УрО РАН

Поступила в редакцию 4.07.01
после переработки 17.05.02

STABILITY OF NUMERICAL ESTIMATIONS OF TIME SERIES CHARACTERISTICS

K. E. Bobrov, A. M. Iskoldsky

The numerical methods of a data analysis (finite ordered sequences of natural binary codes), corresponding to fragments of trajectories, obtained by the numerical solving of a finite number of nonlinear ordinary differential equations are discussed. These equations represent the determined dissipative chaotic dynamic systems.

Measuring properties of such sequences are characterized by the estimation which assumes the code, obtained as a result of processing of source sequence of data by given algorithm, realized on the computer and not supposing the participation of the expert. The concept of stability of the estimation is formalized. The stability of obtained estimations in relation to numerically simulated small variations of registration scheme parameters and to parameters of the algorithm of processing is investigated. Examples of data sequences with typical parameters (digit capacity, time step, length of sequence) for many real experiments are considered.

It is shown that estimation, obtained by the algorithm based on the analysis of properties essentially depended on behavior of trajectory in every point, is unstable. At the same time, the estimation, obtained by the algorithm based on the analysis of properties, essentially depended on some «average» for different points behavior of trajectory, is stable.

Бобров Константин Евгеньевич- родился в 1966 году. Окончил физико-технический факультет Уральского политехнического института по специальности «электроника и автоматика физических установок» (1992). С 1992 года работает в лаборатории моделирования электрофизических процессов Института электрофизики Уральского Отделения РАН в должности программиста. E-mail: cn@ami.uran.ru

Искольдский Александр Михайлович - родился в 1939 году. Окончил Новосибирский университет по специальности «физика плазмы» (1963). С 1963 по 1968 год учился в аспирантуре и работал в Новосибирском институте ядерной физики, с 1968 по 1980 в Институте автоматики и элетрометрии СОАН. Был зам. директора этого института, зам. главного редактора журнала «Автометрия». В 1985 году защитил докторскую диссертацию по специальности «электрофизика». Долгое время был председателем Совета по автоматизации научных исследований. В начале Томского филиала СО АН, затем Уральского филиала АН. В настоящее время - заведующий лабораторией моделирования электрофизических процессов Института электрофизики УрО РАН. Имеет открытие и несколько изобретений. E-mail: ami@ami.uran.ru