

ИПМ им.М.В.Келдыша РАН • Электронная библиотека

Препринты ИПМ • Препринт № 37 за 2019 г.



ISSN 2071-2898 (Print) ISSN 2071-2901 (Online)

Ильин И.Э., Орлов Ю.Н.

Статистическое распознавание динамических систем с хаосом

Рекомендуемая форма библиографической ссылки: Ильин И.Э., Орлов Ю.Н. Статистическое распознавание динамических систем с хаосом // Препринты ИПМ им. М.В.Келдыша. 2019. № 37. 24 с. doi:10.20948/prepr-2019-37

URL: http://library.keldysh.ru/preprint.asp?id=2019-37

Ордена Ленина ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ имени М.В.Келдыша Российской академии наук

И.Э. Ильин, Ю.Н. Орлов

Статистическое распознавание динамических систем с хаосом

Ильин И.Э., Орлов Ю.Н.

Статистическое распознавание динамических систем с хаосом

В работе описывается метод анализа временных рядов, порожденных нелинейными динамическими системами с шумовой составляющей в виде Фильтрация шума осуществляется оцениванием носителя возмущения. плотности совместного распределения значений ряда и его приращений в зависимости от мелкости разбиения многомерной гистограммы. Приводятся примеры разных типов возмущений: аддитивная добавка, возмущение случайное размерности, ШУМ В реализации динамической системы, переключение, вложение нескольких процессов.

Ключевые слова: динамические системы, оптимальное разбиение гистограммы, согласованный уровень значимости, фильтрация шума **Ilin I.E., Orlov Yu.N.**

Statistical recognition of the dynamical systems with chaotic perturbation

In this paper we consider a method of analyzing time series, generated by nonlinear dynamical systems with noise component in the form of perturbation. Noise filtering is carried out by evaluating the domain of density of the joint distribution function of values and its increments, depending on the fineness of the partitioning of the multidimensional histogram. Examples of different types of perturbations are given: additive noise, dimension perturbation, noise in the implementation of a dynamic system, random switching, injection of several processes.

Keywords: dynamical systems, optimal histogram clusterization, self-consistent significance level, noise filtration

Работа выполнена при поддержке гранта РФФИ, проект № 19-01-00602.

Содержание

1. Методы фильтрации шума	3		
 Распознавание нелинейной динамической системы	9		
		5. Переключение процессов	16
		6. Вложение процессов	18
7. Анализ нестационарности	21		
Заключение	23		
Список источников	24		

1. Методы фильтрации шума

Во многих задачах статистического распознавания образов надо очистить сигнал (искомый неизвестный образ) от случайных возмущений, связанных с неточностью измерений и неточностью функционирования самой изучаемой системы. Задачи фильтрации в линейной постановке достаточно подробно исследованы, тогда как нелинейная фильтрация изучена менее детально. Целью данной работы является демонстрация возможностей метода анализа носителя эмпирической плотности совместного распределения случайной величины и ее приращений для распознавания нелинейно коррелированных временных рядов. Сложность такого анализа состоит в том, что сама мелкость разбиения гистограммы для кластеризации наблюдаемых значений случайной величины является определенным фильтром и требует оптимизации.

Кратко изложим основные подходы к линейной фильтрации шумов при анализе стационарных случайных процессов. Напомним [1], что случайный процесс x(t) называется стационарным в широком смысле, если его математическое ожидание a = Mx(t) не зависит от t, а автокорреляционная функция $(AK\Phi)$ A(t,s) = M(x(t)-a)(x(s)-a) = B(t-s) зависит только от разности моментов времени. Спектральным представлением действительного скалярного случайного процесса с $AK\Phi$ B(t) называется представление

$$B(t) = \int_{-\infty}^{+\infty} e^{i\lambda t} dQ(\lambda) . \tag{1}$$

Функция $Q(\lambda)$ называется спектральной функцией, а порожденная ей мера – спектральной мерой.

Пусть x(t) – стационарный в широком смысле процесс, B(t) и $Q(\lambda)$ – его автокорреляционная и спектральная функции. Преобразование $x(t) \to y(t)$ процесса x(t) называется линейным фильтром, если процесс y(t) представляется в виде

$$y(t) = \int_{-\infty}^{\infty} h(t - s)x(s)ds .$$
 (2)

Функция h(t) называется импульсной переходной функцией фильтра, а ее Фурье-образ $H(\omega)$ называется частотной характеристикой фильтра.

Известно (теорема Вольда о разложении), что всякий стационарный процесс x(t) может быть представлен единственным образом в виде суммы некоррелированных между собой сингулярного (детерминированного) процесса и регулярного (полностью случайного).

Сингулярная составляющая стационарного случайного процесса определяется как некоторая динамическая система, а регулярная составляющая ищется на основе следующего утверждения.

Для того чтобы стационарный процесс y(t) был регулярным, необходимо и достаточно, чтобы он представлял собой преобразование фильтрации (2) некоторого случайного процесса x(t) с независимыми приращениями (т.е. белого шума).

В этой связи линейные преобразования фильтрации приобрели большую практическую значимость в анализе и моделировании случайных процессов. Однако линейная теория имеет естественные ограничения по эффективности своей применимости.

Рассмотрим в качестве типичного примера задачу фильтрации шума при измерении некоторого полезного сигнала x(t). Процесс измерения состоит в том, что прибор известным образом искажает сигнал, умножая его на некоторую функцию C(t) (пусть она для простоты постоянна), а также добавляет в показания некоторый шум V(t) в виде стационарного случайного процесса с нулевым средним и заданной дисперсией σ_V^2 . Итак, наблюдается величина

$$z(t) = Cx(t) + V(t). \tag{3}$$

По ней требуется построить оптимальную в среднеквадратичном оценку измеряемой величины x(t), если эта величина и шум не коррелированны.

Линейная по z оценка \hat{x} при нулевых средних имеет вид

$$\hat{x}(t) = az(t). \tag{4}$$

Параметр a выбирается из условия

$$\frac{\partial}{\partial a} \sum_{k} (x_k - \hat{x}_k)^2 = -2 \sum_{k} (x_k - \hat{x}_k) z_k = 0.$$

Подставляя в это уравнение связь $z_k = Cx_k + V_k$ и учитывая, что $\sum_k x_k V_k = 0$, получаем

$$\sum_{k} (Cx_k + V_k)(x_k - aCx_k - aV_k) = (C - aC^2)\sigma_x^2 - a\sigma_V^2 = 0,$$

откуда следует

$$a = \frac{C\sigma_x^2}{C^2\sigma_x^2 + \sigma_V^2} \,. \tag{5}$$

Если шума нет, то $\sigma_V^2 = 0$ и, естественно, a = 1/C. Но если шум есть, то наряду с гипотезой об отсутствии корреляции сигнала и шума надо еще знать дисперсию σ_V^2 шума, что на практике, как правило, не реализуется. Следовательно, полученное решение (5) полезно главным образом теоретически как пример существования метода фильтрации.

Пусть далее наблюдаемый сигнал (3) дискретный, а C = 1.

Если известна не только дисперсия шума и фильтруемого сигнала, но и более общие характеристики – их автокорреляционные функции $B_X(n)$, $B_V(n)$

и ковариационная функция $A_{\chi\chi}(k,k+n)=B_{\chi\chi}(n)$, то можно построить соответствующий фильтр (фильтр Винера, см., напр. [2]). Пусть h(n) – искомая импульсная характеристика фильтра. Тогда на выходе из него вместо z_n наблюдается сигнал

$$y_n = \sum_{k=0}^{n} z_k h(n-k)$$
. (6)

Будем строить фильтр, исходя из минимизации среднего квадрата ошибки фильтрации $(y_n-x_n)^2$. При этом сама ошибка оказывается ортогональной в смысле среднего значения фильтруемому сигналу: $M(z_n(y_n-x_n))=0$. Это условие, записанное через АКФ, называется уравнением Винера-Хопфа:

$$B_{ZX}(j) = \sum_{k=0}^{n} B_{Z}(j-k)h(k).$$
 (7)

Если время наблюдения бесконечно, то (7) представляет собой свертку функций, так что их фурье-образы (спектральные функции, согласно теореме Винера-Хинчина) входят в виде произведения, что позволяет определить частотную характеристику фильтра:

$$H_{opt}(\omega) = \frac{S_{zx}(\omega)}{S_{z}(\omega)}, \quad B(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S(\omega) e^{i\omega t} d\omega.$$
 (8)

Если сигнал и шум не коррелированы, то $S_{zx}(\omega) = S_z(\omega) = S_x(\omega) + S_V(\omega)$ и частотная характеристика фильтра имеет вид

$$H_{opt}(\omega) = \frac{S_{\chi}(\omega)}{S_{\chi}(\omega) + S_{V}(\omega)}.$$
 (9)

Однако спектральная функция шума $S_V(\omega)$ не определяется непосредственно по измеряемому сигналу на входе фильтра. Следовательно, для успешной фильтрации надо иметь «библиотеку» подходящих корреляционных функций, после чего методом проб и ошибок подбирать ту, которая действительно будет фильтровать сигнал.

Если время наблюдения в (7) конечно, то вместо (8) рассматривается выборочная ковариационная матрица суммарного сигнала, построенная по выборке некоторой длины N:

$$B_{z}(n) = M(Z_{n}Z_{n}^{T}), \ Z_{n}^{T} = (z_{n}, z_{n-1}, ..., z_{n-N+1}).$$
 (10)

В итоге решением уравнения Винера-Хопфа является оптимальный фильтр с передаточной функцией $\boldsymbol{H}_n^T = (h_0, h_1, ..., h_{N-1})$, такой, что

$$H_{n,opt} = B_z^{-1}(n)r_{zx}(n), \ r_{zx}^T(n) = M(x_n Z_n^T).$$
 (11)

Однако следует признать, что величина $r_{\chi\chi}(n)$ в правой части (11) неизвестна. Аналогичный вывод приходится сделать и относительно более сложных задач. Возникает вопрос: можно ли реализовать на практике

некоторый способ распознавания случайного процесса, если его свойства априори не известны?

Предположим, что между элементами x и y двух временных рядов имеется прямая функциональная связь, т.е. $y = \varphi(x)$, $x, y \in [0;1]$. Тогда, построив совместное распределение вероятностей $f_N(x,y)$ по выборке длины N, мы обнаружим, что его носитель находится в квадратах $y_i = \varphi(x_j)$, в соответствии с разбиением гистограммы на n равных кластеров. Точность, с которой мы можем говорить о такой функциональной связи, равна точности позиционирования случайных величин в классовых интервалах, т.е. 1/n. Уровень достоверности полученной оценки равен интегралу от плотности распределения по выбранной доле носителя. Поскольку в данном случае вне отмеченных квадратов нет точек носителя совместного распределения, уровень достоверности равен единице.

Если функциональной связи нет, то при фиксированном номере ј интервала по первому аргументу мы обнаружим отличные от нуля значения функции $f_N(j,i)$ для нескольких номеров i интервалов по второму аргументу. При этом с увеличением длины выборки носитель совместного распределения занимает все большую долю области разбиения гистограммы. Это означает, что путем потери точности можно получить достоверную оценку функциональной связи даже в отсутствие таковой, но будет ли это удовлетворять исследователя? Насколько точно нужно позиционировать искомое значение, чтобы вероятность его принадлежности определенному интервалу была не исчезающе малой, но и сам интервал существенно отличался бы от всего множества значений случайной величины? Для этого вместо априори задаваемого значимости следует ввести согласованный критерий совместной оценки точности и уровня значимости [3].

Пусть

$$\delta = \int_{\Omega} dx dy \tag{12}$$

есть мера множества $\Omega(x,y)$, принадлежащего носителю совместного распределения, на котором можно говорить о функциональной связи между x и y. Величина δ будет точностью, с которой установлена эта связь, а величина

$$\alpha = \int_{\Omega} f(x, y) dx dy \tag{13}$$

будет давать уровень достоверности найденной связи. Как величина δ , так и α зависят от множества $\Omega(x,y)$. Выберем такую часть этого множества, для которого $\alpha + \delta = 1$. Величину $\delta = 1 - \alpha$ будем называть согласованным (с множеством Ω) уровнем значимости распознавания.

Тогда фильтрацию можно осуществить следующим образом. Отметим ту ячейку гистограммы совместного распределения $f_N(j,i)$, в которой для j-го классового интервала значение функции $f_N(j,i)$ максимально. Таких ячеек при

данной мелкости разбиения будет n, т.е. потенциальная точность распознавания будет равна $\delta = 1/n$. Если при этом сумма отмеченных вероятностей будет приблизительно равна величине $1-\delta$, то задачу распознавания считаем законченной. Если же уровень доверия оказался существенно меньше, то увеличиваем меру носителя, участвующего в распознавании, включая в него в каждом j-ом столбце одну из двух соседних ячеек дополнительно к отмеченной ранее, именно ту из них, в которой вероятность больше. Тем самым повысится величина интеграла в (13), но ухудшится точность (12), она станет равной $\delta = 2/n$. Постепенно ухудшая точность, добиваемся согласования значимости с уровнем доверия, так что $\alpha + \delta = 1$.

После того как определена полоса носителя Ω , в которой непараметрически была распознана зависимость y(x) в виде собственно самого этого множества Ω , переходим к фильтрации. Рассматриваем пару значений (x_1, y_1) . Если эта пара принадлежит множеству Ω , то она выбрасывается из анализируемой совокупности, а если она лежит в ячейке, не входящей в это множество, то остается. В итоге остаются только пары, не связанные функциональной зависимостью. Их статистику можно изучить отдельно от основной зависимости, отфильтрованной на предыдущем этапе.

2. Распознавание нелинейной динамической системы

Рассмотрим задачу распознавания хорошо известной динамической системы с квадратичной нелинейностью — так называемой логистической системы, задаваемой рекуррентно формулой

$$x_{n+1} = 4x_n(1 - x_n). (14)$$

Совместная выборочная плотность распределения $f(x_{n+1},x_n)$ двух последовательных координат этой системы при мелкости разбиения r=1/100 показана на рис. 1 для выборки длины N=20 тыс.

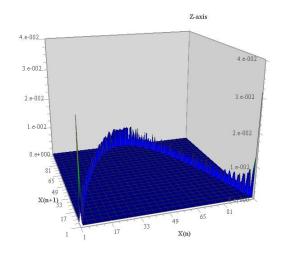


Рис. 1 — Совместная выборочная плотность распределения $f(x_{n+1},x_n)$ для логистической системы при мелкости разбиения r=1/100

Несмотря на то что визуально носитель распределения отчетливо представляет собой параболу из (14), сумма вероятностей по ячейкам с максимальными значениями совместной плотности составляет всего лишь 0,450. Однако сумма площадей всех ячеек в единичном квадрате с ненулевым значением совместной вероятности равна 0,028, так что данную систему можно распознать на уровне доверия, близком к 1-0,028=0,972.

Согласованным уровнем значимости при распознавании носителя совместного распределения назовем площадь В единичном квадрате, покрываемую наименьшей частью носителя, сумма вероятностей по которой равна дополнению этого уровня до единицы. Для рассматриваемой на рис. 1 системы этот уровень равен 0,025 и достигается при суммировании по трем ячейкам с наибольшими значениями вероятностей в каждом столбце x(n)(таких ячеек в данном примере 216 из 10 тыс.), а также включая еще 34 из оставшихся 64 ячеек. Суммарная вероятность по указанным ячейкам, т.е. уровень доверия распознавания данной динамической системы, равна 0,975.

При укрупнении разбиения уровень значимости меняется примерно пропорционально величине ячейки разбиения (рис. 2).

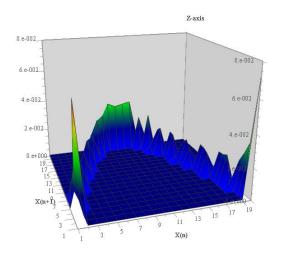


Рис. 2 — Совместная выборочная плотность распределения $f(x_{n+1},x_n)$ для логистической системы при мелкости разбиения r=1/20

Сумма по максимальным ячейкам приблизительно такая же (0,449), как и при мелкости разбиения r = 1/100, но ненулевых ячеек в этом случае 56 из 400, так что согласованный уровень значимости будет не лучше, чем 0,14, что примерно в 5 раз хуже, чем в предыдущем случае.

Рассмотрим теперь некоторую модификацию системы (14) с тем, чтобы введение дополнительных возмущений не нарушило бы равномерной ограниченности порождаемого временного ряда. Если незначительно изменить коэффициент при нелинейности, то совместные распределения будут похожи. Например, пусть

$$x_{n+1} = 3.9x_n(1 - x_n). (15)$$

Совместная выборочная плотность распределения $f(x_{n+1},x_n)$ двух последовательных значений координаты системы (15) при мелкости разбиения r=1/100 показана на рис. З для выборки длины N=20 тыс.

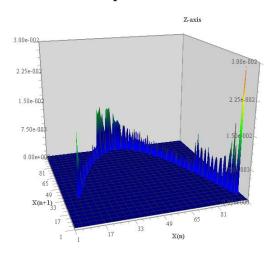


Рис. 3 — Совместная выборочная плотность распределения $f(x_{n+1},x_n)$ для системы (15) при мелкости разбиения r=1/100

Согласованный уровень значимости, на котором распознается эта система, равен 0,022. Возникает вопрос: насколько эффективен описанный метод при распознавании зашумленной динамической системы?

Далее рассматриваются различные типы возмущений.

3. Возмущение размерности

Рассмотрим вместо динамической системы (15) систему $x_{n+1} = 3.9x_n (1-x_n) + bx_{n-1}, \tag{16}$

где b есть некоторое малое число, например, b = 0.013. Не зная того, что реальная система двумерна, будем искать носитель совместного распределения, считая систему одномерной. Соответствующая выборочная плотность совместного распределения показана на рис. 4.

Несмотря на то что распределения на рис. 3 и 4 визуально похожи, их носители не совпадают по 151 ячейке. В 112 ячейках из отмеченных различающихся элементов носителей система (16) присутствует, а система (15) – нет, в 39 ячейках — наоборот. Носитель распределения для системы (15) содержится в 239 ячейках, носитель распределения для системы (16) содержится в 312 ячейках, общих ячеек 200. Интеграл от модуля разности совместных распределений по несовпадающей части их носителей равен 0,374.

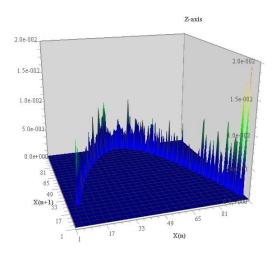


Рис. 4 — Совместная выборочная плотность распределения $f(x_{n+1},x_n)$ для системы (16) для b=0.013 при мелкости разбиения r=1/100

С увеличением параметра b отклонения в распределениях становятся более заметны (рис. 5).

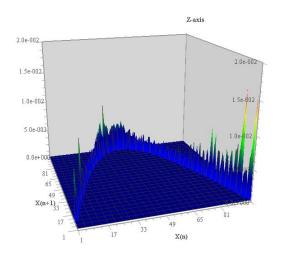


Рис. 5 — Совместная выборочная плотность распределения $f(x_{n+1},x_n)$ для системы (16) для b=0.029 при мелкости разбиения r=1/100

Видно, что некоторая динамическая система распознается, но не совсем та, которой соответствует распределение на рис. 3. Носители распределений не совпадают уже по 288 ячейкам. В 242 из них система (16) присутствует, а система (15) — нет, в 46 ячейках — наоборот, общих ячеек 193. Интеграл от модуля разности совместных распределений по несовпадающей части их носителей равен 0,661.

В примере на рис. 5 носитель распределения содержится в 435 ячейках, согласованный уровень значимости распознавания равен 0,04, т.е. эта система распознается примерно в 1,5 раза хуже, чем невозмущенная. Тем не менее эта точность вполне приемлема с практической точки зрения.

Итак, мы достаточно точно распознали динамическую систему, содержащуюся в носителе распределения, представленного на рис. 5. Однако насколько точно мы распознали при этом невозмущенную систему? Ведь именно это есть предмет нашего анализа. Для этого надо рассмотреть расстояние между носителями (рис. 3 и 5) максимальных совместных вероятностей в норме *С*. В данном случае оно также оказалось равным 0,04. Таким образом, малое возмущение размерности системы не явилось препятствием для ее достаточно точной аппроксимации.

Рассмотрим теперь ситуацию, когда реальная динамическая система многомерна. Близкой по структуре к рассматриваемой логистической модели является система Лоренца, в дифференциальной форме имеющая вид [4]

$$\begin{cases} \dot{x} = \sigma(y - x), \\ \dot{y} = \lambda x - y - xz, \\ \dot{z} = -\mu z + xy. \end{cases}$$
(17)

Здесь σ , λ , μ – некоторые числовые параметры.

После дискретизации трехмерного вектора (для простоты с одним и тем же шагом h) получим систему

$$\begin{cases} x_{n+1} = x_n + \sigma h(y_n - x_n), \\ y_{n+1} = y_n(1-h) + \lambda h x_n - h x_n z_n, \\ z_{n+1} = (1-\mu h) z_n + h x_n y_n. \end{cases}$$
(18)

Предположим, что после ее численного решения анализируется одномерный временной ряд, например, значений z_n . Можно ли по совместному распределению значений этого ряда понять, что система не одномерна? На рис. 6 представлен фрагмент численной реализации координаты z_n из системы (18).

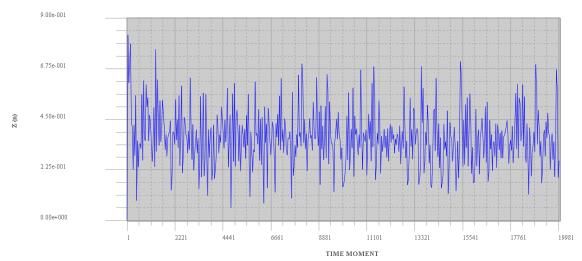


Рис. 6 – Пример временного ряда, порожденного одной из координат системы Лоренца

На рис. 7 показана выборочная плотность совместного распределения $f(z_{n+1},z_n)$.

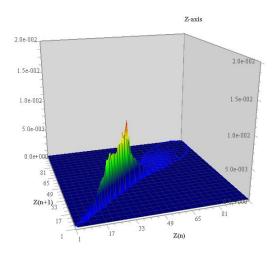


Рис. 7 — Совместная выборочная плотность распределения $f(z_{n+1}, z_n)$ для системы (18) при мелкости разбиения r = 1/100

Видно, что при малых значениях координаты z добавка от высших размерностей мала, так как при малом расчетном шаге вычислений главной частью является $z_{n+1} \approx (1 - \mu h)z_n$, и потому практически наблюдается линейная авторегрессионная модель, но при больших значениях координат влияние размерности оказывается существенным и система демонстрирует хаотическое позволяющее провести идентификацию функциональной зависимости. Следовательно, в общем случае неверная оценка размерности задачи будет препятствием для точного распознавания динамической системы. Одним ИЗ индикаторов наличия скрытой размерности является неоднозначность функциональной связи, демонстрируемой носителем на рис. 7 в области больших значений z.

4. Аддитивный шум

Рассмотрим теперь другой тип возмущения системы (15). Будем считать, что в самой динамической системе (15) содержится некоторое случайное искажение функциональной зависимости. Положим

$$x_{n+1} = 3.9x_n(1 - x_n) + \alpha_{n+1}, \tag{19}$$

где α_n есть некоторая малая случайная добавка. Дальнейшее зависит от распределения случайной величины α . Если оно имеет малую дисперсию, то система (15) по данным (19) распознается хорошо, а если большую, то распознавание несколько хуже.

Пусть α имеет равномерное распределение на промежутке, который много меньше размаха анализируемой выборки, например (0; 0,02). В этом случае система идентифицируется достаточно точно. Совместная выборочная

плотность распределения $f(x_{n+1},x_n)$ двух последовательных координат системы (17) при мелкости разбиения r=1/100 показана на рис. 8 для выборки длины N=20 тыс.

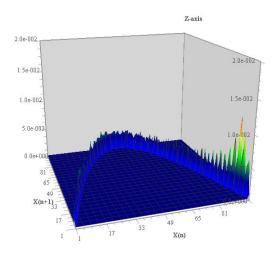


Рис. 8 – Совместная выборочная плотность распределения $f(x_{n+1},x_n)$ для системы (17) при мелкости разбиения r=1/100

Согласованная точность распознавания этой системы оказалась равной 0,052. В силу нелинейности исходной системы (15) возмущение привело к несколько большей погрешности в распознавании системы, чем сумма исходной погрешности (0,022) и шума с амплитудой 0,020. Для сравнения на рис. 9 приведено совместное распределение для системы, полученной из (19) при постоянном значении $\alpha = 0,01$, что приближенно отвечает ситуации, когда дисперсия распределения случайной величины α мала.

$$x_{n+1} = 3.9x_n(1 - x_n) + 0.01. (20)$$

Эта система распознается на том же уровне значимости, что и исходная система (15), т.е. с ошибкой, равной 0,022.

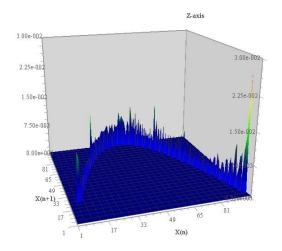


Рис. 9 — Совместная выборочная плотность распределения $f(x_{n+1},x_n)$ для системы (20) при мелкости разбиения r=1/100

Рассмотрим теперь другой тип зашумления, когда наблюдаемый процесс представляет собой результат реализации динамической системы (15) с шумовой составляющей, т.е. временной ряд имеет вид

$$y_n = x_n + \alpha_n,$$

 $x_{n+1} = 3.9x_n(1 - x_n).$ (21)

В этом примере α имеет равномерное распределение на [0;1]. Соответствующая выборочная плотность представлена на рис. 10.

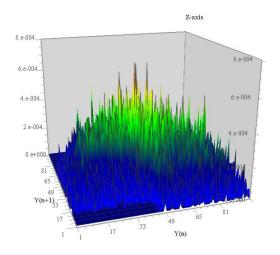


Рис. 10 — Совместная выборочная плотность распределения $f(y_{n+1},y_n)$ для системы (21) при мелкости разбиения r=1/100

Хотя распределение на рис. 10 представляется хаотичным, расположение кривой локальных максимумов по методу, описанному в п.1, оказывается все же идентифицируемым (не особенно точно) как парабола вида (15):

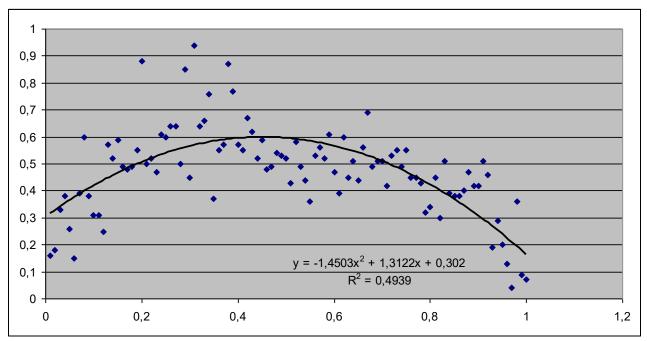


Рис. 11 – Распределение максимумов совместного распределения на рис. 10

Интеграл (13) по отмеченным ячейкам, занимающим 1/100 от площади единичного квадрата, равен всего лишь 0,038, что, конечно, слишком мало для уровня доверия. Расширив меру носителя, содержащего указанную полосу максимумов, до половины площади квадрата, мы получим величину интеграла, также равного приблизительно 0,5. На этом согласованном уровне значимости распознается система (15) по данным (21) при выбранной мелкости разбиения и зашумленности, амплитуда которой равна размаху самой системы.

Заметим теперь, что точность, с которой определяется распределение случайного процесса, зависит от мелкости разбиения, которое само является своеобразным фильтром. Согласно методике оптимального разбиения гистограммы, описанной в [5], число n ячеек одномерного разбиения на равные классовые интервалы определяется при заданной длине N выборки по формуле

$$n\sqrt{-\frac{\pi}{2}\ln(1-(1-1/n)^2)} = \frac{\sqrt{N}}{\Sigma_N(n)}, \quad \Sigma_N(n) = \sum_{i=1}^n \sqrt{f_N(i)(1-f_N(i))}.$$
 (22)

Здесь $f_N(i)$ есть эмпирическая частота попадания значений выборки в указанную ячейку с номером i. Для выборки длины N=20 тыс. оптимальным в этом смысле оказывается разбиение на n=20 классовых интервалов. Соответствующая плотность совместного распределения приведена на рис. 12. Она гораздо более точно функционально идентифицируема, чем плотность той же самой системы на рис. 10 при другой мелкости разбиения.

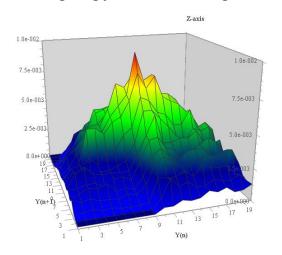


Рис. 12 — Совместная выборочная плотность распределения $f(y_{n+1}, y_n)$ для системы (21) при оптимальной мелкости разбиения r = 1/20

Ячейки максимумов совместного распределения образуют полосу мерой 0,05, в которой содержится 0,10 этого распределения (см. рис. 13). Расширяя эту полосу до 0,35 площади квадрата, находим, что величина интеграла (13) по ней приблизительно равна 0,65, т.е. 0,35 — это и есть согласованный уровень значимости в идентификации динамической системы (15).

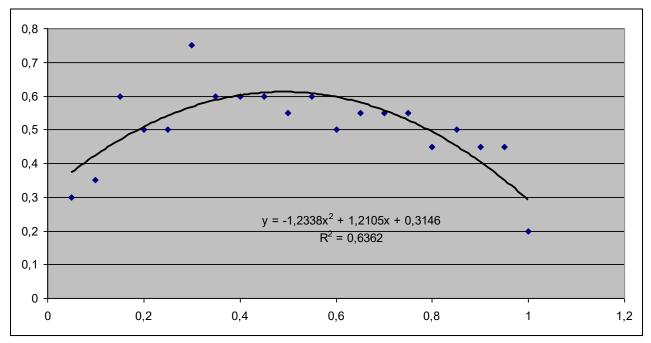


Рис. 13 – Распределение максимумов совместного распределения на рис. 12

Отметим, что, хотя коэффициент при $-x^2$ в аппроксимации на рис. 13 примерно в 3 раза отличается от правильного коэффициента системы (15), все же эта система содержится в полосе интегрирования указанной доли носителя. Этот результат представляется достаточно хорошим, поскольку «на глаз» динамическая система весьма слабо просматривается на рис. 12 и вообще не видна на рис. 10.

5. Переключение процессов

Если случайный процесс реализуется в зависимости от того, какому множеству принадлежит значение некоторого функционала от фрагмента траектории (т.е. фактически от скользящего выборочного распределения реализации процесса), то задачей является определение не только самой системы, но и условий переключения. В простейшем случае переключение обусловлено тем, какому множеству принадлежит определенное (допустим, текущее) значение ряда. Такой модификацией исследуемой нами системы (15) является, например, следующая:

$$x_{n+1} = \begin{cases} 3.9x_n(1-x_n), & x_n \notin (0.5-\beta; 0.5+\beta); \\ random\alpha_{n+1}, & x_n \in (0.5-\beta; 0.5+\beta). \end{cases}$$
 (23)

В зависимости от полуширины промежутка переключения β совместное выборочное распределение меняется более или менее заметно. Статистически обнаружить переключение можно только в случаях, когда либо вероятность попадания значений ряда в зону переключения не меньше согласованного уровня значимости, на котором распознается динамическая система, либо ширина зоны переключения не меньше мелкости разбиения.

Например, для $\beta = 0.01$ в выборке длиной N = 20 тыс. количество переключений в системе (23) равно 295, а для $\beta = 0.2$ оно равно 9024, что составляет почти половину выборки. Совместные выборочные распределения для этих вариантов представлены на рис. 14 и 15.

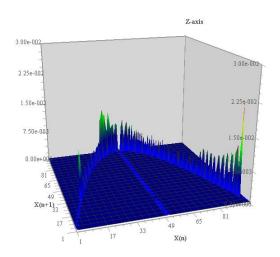


Рис. 14 — Совместная выборочная плотность распределения $f(x_{n+1},x_n)$ для системы (23) при мелкости разбиения r=1/100 и переключении $\beta=0.01$

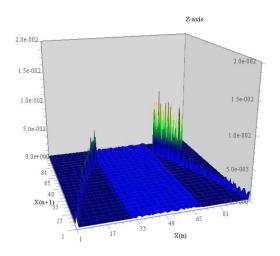


Рис. 15 — Совместная выборочная плотность распределения $f(x_{n+1},x_n)$ для системы (23) при мелкости разбиения r=1/100 и переключении $\beta=0,2$

В приведенных примерах области динамического поведения и хаотизации могут быть распознаны на уровне значимости 0,05. Если же полуширина переключения $2\beta < r$ (т.е. $\beta < 0,005$), то для системы (23) этот эффект не обнаруживается данным методом.

Усложним условие переключения, дополнив его, например, условием положительного приращения значения ряда на предыдущем шаге. Пусть

$$x_{n+1} = \begin{cases} 3.9x_n(1-x_n), & x_n \notin (0.5-\beta; 0.5+\beta); \\ random\alpha_{n+1}, & x_n \in (0.5-\beta; 0.5+\beta) \bigcap x_n > x_{n-1}; . \\ 3.9x_n(1-x_n), & x_n \in (0.5-\beta; 0.5+\beta) \bigcap x_n \le x_{n-1}. \end{cases}$$
(24)

Тогда совместная плотность имеет вид (рис. 16):

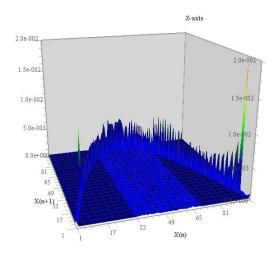


Рис. 16 — Совместная выборочная плотность распределения $f(x_{n+1}, x_n)$ для системы (24) при мелкости разбиения r = 1/100 и переключении $\beta = 0.2$

Переключение в этом примере происходит в 38 % случаев, но поскольку максимум совместного распределения в области переключения по-прежнему отвечает динамической системе (15), режим переключения не распознается. Уровень доверия при моделировании ряда (24) системой, лежащей в максимуме полосы носителя совместного распределения на рис. 16, примерно равен 0,55. Следовательно, в этом примере статистически выявить точное условие переключения не удается. Метод, с помощью которого можно было бы статистически выявить условие переключения, пока неизвестен.

6. Вложение процессов

Вложением двух процессов с шагом вложения p будем называть временной ряд, задаваемый условиями:

$$x(pn) = random \alpha(pn);$$

 $x(pn+k) = g(x(pn-p+k)), k = 1,2,..., p-1.$ (25)

В частности, для p=2 все четные элементы ряда имеют равномерное распределение на [0;1], а все нечетные — отвечают некоторой динамической системе. Фигурирующая в (25) функция g(x) задается, например, правой частью выражения (15). На рис. 17 и 18 показаны плотности совместных распределений при вложении различных порядков.

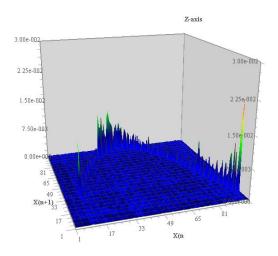


Рис. 17 — Совместная выборочная плотность распределения $f(x_{n+1},x_n)$ для системы (25) при мелкости разбиения r=1/100 и вложении с p=7

Динамическая система на рис. 17 легко опознается по максимуму распределения и отфильтровывается: пары последовательных значений, принадлежащие ее носителю, исключаются из выборки. Оставшиеся данные анализируются вновь, и для них выясняется, что их распределение равномерно.

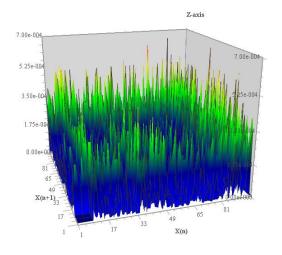


Рис. 18 — Совместная выборочная плотность распределения $f(x_{n+1},x_n)$ для системы (25) при мелкости разбиения r=1/100 и вложении с p=2

При высокой частоте вложения, как на рис. 18, динамическая система не распознается независимо от мелкости разбиения. Причина здесь не в том, что четное вложение хаотическое, а в том, что фильтрующее окно лаговой переменной в построении совместного распределения выбрано неправильно. В этом случае даже детерминированное вложение двух динамических систем распознаваться не будет. Рассмотрим, например, четно-нечетное вложение логистического отображения (15) и отображения Бернулли (см. [6]), выделяющего дробную часть числа:

$$x_{2n+1} = 3.9x_{2n-1}(1 - x_{2n-1});$$

$$x_{2n} = \begin{cases} 2x_{2n-2}, & 0 \le x_{2n-2} < 1/2; \\ 2x_{2n-2} - 1, & 1/2 \le x_{2n-2} \le 1. \end{cases}$$
(26)

Если не знать применяемого лага, то получается «игольчатое» совместное распределение, представленное на рис. 19. При правильном лаге распознавание таких систем проводится точно (рис. 20).

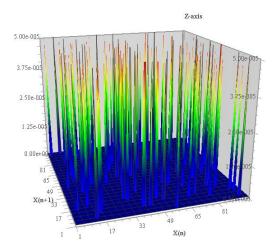


Рис. 19 — Совместная выборочная плотность распределения $f(x_{n+1},x_n)$ для системы (26) при мелкости разбиения r=1/100

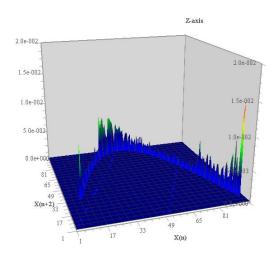


Рис. 20 — Совместная выборочная плотность распределения $f(x_{n+2},x_n)$ для системы (26) при мелкости разбиения r=1/100

Правильным для распознавания является такой лаг, при котором получается наилучшая точность распознавания функциональной зависимости. Следовательно, для фильтрации вложения, как и для сложной стратегии переключения между процессами, необходимо иметь некоторую априорную информацию о системе, чтобы искомый процесс можно было обнаружить статистическим методом из анализа совместного распределения.

Подчеркнем, что упорядоченность вложения (каждый второй или каждый третий и т.п. элемент ряда должен быть взят из заданного распределения) не играет существенной роли. Важен фильтрующий лаг, с которым надо строить плотность совместного распределения. Например, если принять, что вложение отображения Бернулли из (26) осуществляется в логистическое отображение в случайный момент времени в соответствии с тем, например, является ли величина $\inf(1/rand(n))$ четной или нечетной, то эти динамические системы распознаются точно (рис. 21), хотя само правило переключения остается при этом неизвестным.

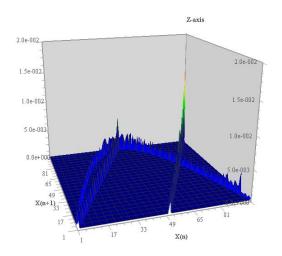


Рис. 21 — Совместная выборочная плотность распределения $f(x_{n+1}, x_n)$ для системы (26) при случайном переключении, мелкость разбиения r = 1/100

В примере на рис. 21 вложение фактически является переключением и имеет свойства, аналогичные рассмотренным в п. 5.

7. Анализ нестационарности

В заключение рассмотрим вопрос об однородности выборок, полученных при моделировании динамических систем с возмущением. Для анализа используем индекс нестационарности, введенный в [3]. Он определяется следующим образом.

Сначала строятся выборочные функции распределения по выборкам длины N, меняющимся в доступных исследователю пределах. Затем вычисляются расстояния между независимыми выборками в норме C. В частности, для анализа эволюции целесообразно рассматривать так называемые встык-выборки. Обозначим через $G_N(\rho)$ эмпирическую функцию распределения расстояний $\rho(N)$ между двумя встык-выборками длины N:

$$\rho(N) = \|F_{1,N}(x) - F_{2,N}(x)\|_{C}.$$
(27)

Численное решение относительно ρ уравнения

$$G_N(\rho) = 1 - \rho \tag{28}$$

дает значение согласованного уровня стационарности $\rho^*(N)$ как вероятности того, что расстояние между выборками длины N больше ρ^* . Для стационарных распределений соответствующая стационарная точка $\varepsilon = \varepsilon_0(N)$ уровня значимости определяется как решение уравнения

$$1 - K \left(\sqrt{\frac{N}{2}} \varepsilon \right) = \varepsilon, \tag{29}$$

где K(z) есть функция Колмогорова. Для стационарных распределений положение стационарной точки не зависит от вида распределения F.

Индексом нестационарности называется отношение

$$J(N) = \frac{\rho^*(N)}{\varepsilon_0(N)}. (30)$$

Ряд стационарный, если для его выборок выполняется условие $J(N) \le 1$, и нестационарный в противном случае.

Визуально рассмотренные модификации динамической системы (15) весьма сложно различимы. На рис. 22 представлен составной ряд из фрагментов длиной 4 тыс. данных для моделей (15), (16), (19), (23) и (25). Выяснилось, что собственно логистическая система (14), как и ее модификации (15) и (16), стационарны на всех длинах выборки. Индекс нестационарности для них колеблется около значения J=0.85. Те же свойства и у рядов (19) и (23), образованных добавлением шума и переключениями. Нестационарным на относительно малых выборках оказался только процесс вложения (25), который становится стационарным на выборках длин более 3 тыс.



Рис. 22 – Составные фрагменты изучаемых рядов длиной 4 тыс. данных

На рис. 23 показана зависимость индекса нестационарности построенного составного ряда от длины выборки.

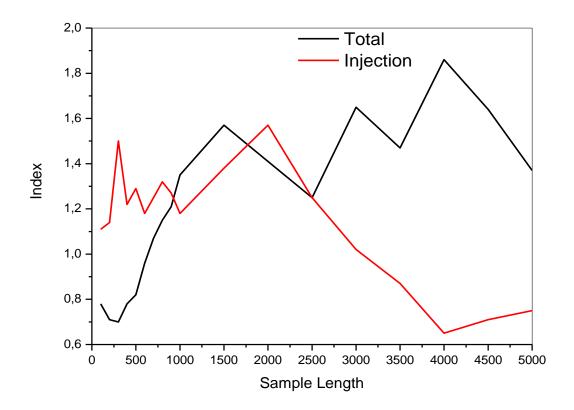


Рис. 23 – Индексы нестационарности исследуемых рядов

Анализ индекса нестационарности показал, что для составного ряда на рис. 22, составленного из фрагментов различных рядов длиной 4 тыс., наибольшая нестационарность наблюдается именно на длине 4 тыс. данных, отвечающих «кванту сборки». На малых выборках этот ряд стационарный, но на длинах, превышающих 800 данных, фрагменты составного ряда уже становятся статистически различимы.

Заключение

В данной работе были приведены примеры статистического распознавания динамических систем с шумом. При наличии шумового возмущения точность распознавания зависит от оптимальности разбиения области многомерной гистограммы. Выяснилось, что многие типы возмущений, оставаясь малыми по сравнению с динамикой основной системы, могут быть отфильтрованы с практически приемлемой точностью. Существенно, что для такой фильтрации не требуется априорных знаний о характеристиках входного сигнала.

Показано также, что распознавание нетривиального управления даже при относительно небольших лагах запаздывания невозможно без дополнительных указаний на способ управления. Можно лишь распознать эффект наличия нелокального управления. Это, в частности, означает, что чисто статистический анализ данных имеет ограничения по точности извлечения информации.

Вопрос об отыскании непараметрического метода для дешифровки интегрального по времени и, возможно, нелинейного процесса стохастического управления остается открытым.

Список источников

- 1. Королюк В.С., Портенко Н.И., Скороход А.В., Турбин А.Ф. Справочник по теории вероятностей и математической статистике. М.: Наука, 1985. 640 с.
- 2. Шахтарин Б.И. Случайные процессы в радиотехнике. Линейные преобразования. М.: «Гелиос APB», 2006.
- 3. Орлов Ю.Н. Кинетические методы исследования нестационарных временных рядов. М.: МФТИ, 2014. 276 с.
- 4. Кузнецов С.П. Динамический хаос. М.: Физматлит, 2001. 296 с.
- 5. Орлов Ю.Н. Оптимальное разбиение гистограммы для оценивания выборочной плотности распределения нестационарного временного ряда // Препринты ИПМ им. М.В. Келдыша. 2013. № 14. 26 с.

URL: http://library.keldysh.ru/preprint.asp?id=2013-14

6. Аникин В.М., Голубенцев А.Ф. Аналитические модели детерминированного хаоса. — М.: Физматлит, 2007. - 326 с.