



ЭФФЕКТ РЕДКОЙ ВЫБОРКИ ПРИ ОЦЕНКЕ НАПРАВЛЕННЫХ СВЯЗЕЙ ПО ВРЕМЕННЫМ РЯДАМ

Д. А. Смирнов, Б. П. Безручко

В различных областях исследований возникает задача обнаружения и количественной оценки направленных связей (взаимных воздействий) между системами по дискретным записям их колебаний – временным рядам. В данной работе показано, что при использовании для ее решения традиционных характеристик «причинности по Грейнджеру» результаты существенно зависят от интервала выборки (шага дискретизации по времени). Выявлены причины и характер влияния интервала выборки на численные значения оценок связи и объяснено, почему при большом интервале (редкой выборке) могут быть получены ошибочные выводы о двунаправленной связи в случае однонаправленно связанных систем. Такой эффект редкой выборки продемонстрирован как для линейных, так и для нелинейных систем в различных режимах динамики.

Ключевые слова: Связанные нелинейные системы, временные ряды, однонаправленная связь, двунаправленная связь, причинность по Грейнджеру.

Введение

Исследования связанных колебательных систем – традиционная проблематика теории колебаний и нелинейной динамики. Многочисленные работы посвящены изучению условий синхронизации различных систем (см., например, [1–4]) и соответствующей обратной задаче – оценке степени синхронности колебаний по временным рядам, то есть дискретным последовательностям значений наблюдаемых величин (см., например, [5–9]). В последние годы среди обратных задач нелинейной динамики все большее внимание уделяется задаче выявления «направленных связей» (взаимных воздействий) по наблюдаемым данным, что актуально в механике [10], радиофизике [11,12], геофизике [13–15] и биофизике [16–21]. В ряде постановок фундаментально важен вопрос о том, является ли связь одно- или двунаправленной [14,16–21]. Для решения таких задач оказалось полезным понятие «причинности по Грейнджеру», предложенное в эконометрике [22,23] и становящееся все более популярным в физических науках [14–16].

Говорят, что система Y влияет на систему X «по Грейнджеру», если прогноз будущего поведения X улучшается при учете данных об Y по сравнению с прогнозом, основанным только на данных об X . Ненулевое улучшение прогноза (УП) ассоциируют с наличием физического воздействия Y на X , а ненулевые УП «в обе стороны» обычно интерпретируются как признак двунаправленной связи. Прогноз осуществляется на один временной шаг Δt – интервал выборки, который определяется процедурой измерений, например, частотой дискретизации аналого-цифрового преобразователя. При этом в физических исследованиях обычно не принимают во внимание возможность сложной зависимости УП от Δt . В математической литературе отмечалось, что при достаточно редкой выборке (большом Δt) даже в случае *однонаправленно* связанных систем могут наблюдаться ненулевые УП *в обе стороны*, то есть «ложные связи» [24,25]. Это принципиальное обстоятельство остается недооцененным в работах по анализу временных рядов, возможно, потому что не проводился анализ количественных значений «ложных связей» и не было показано, насколько значителен может быть эффект ложных связей из-за редкой выборки на практике.

В данной работе проводится такой количественный анализ, и на примере эталонных колебательных систем с однонаправленной связью демонстрируется типичность больших значений УП в «ложном» направлении. Показаны условия, когда этот эффект наиболее силен, так что УП в «ложном» направлении могут даже превосходить УП в «правильном» направлении. Ниже, в п. 1 вводится формальное определение причинности по Грейнджеру. Затем демонстрируется эффект редкой выборки на линейных осцилляторах (п. 2) и нелинейных отображениях (п. 3). Выводы представлены в Заключение.

1. Причинность по Грейнджеру

Здесь мы рассматриваем лишь такую постановку задачи, в которой есть две исследуемых системы и отсутствует любая третья система, которая влияла бы на них. Наличие «скрытой» третьей системы сильно осложняло бы выводы о характере взаимосвязей [23]. Но и в этой элементарной постановке двух систем, с которой необходимо уметь справляться для решения более сложных задач, возникают непростые вопросы. Итак, пусть $(X(t), Y(t))$ – двухмерный случайный процесс. Наблюдаются значения в дискретные моменты времени с интервалом выборки Δt : $x_n = X(n\Delta t)$, $y_n = Y(n\Delta t)$, где n – целое число, которое обычно называют дискретным временем. Обозначим $x_n^- = \{x_{n-k}\}_{k=1}^\infty$ и $y_n^- = \{y_{n-k}\}_{k=1}^\infty$ – последовательности значений x_n и y_n , предшествующие n -му моменту времени. Из всех способов индивидуального (без учета Y) прогноза x_n наименьший средний квадрат ошибки имеет место в случае $x_n^{\text{ind}} = M[x_n | x_n^-]$, где $M[x_n | x_n^-]$ – математическое ожидание x_n при условии x_n^- . Обозначим средний квадрат ошибки $\sigma_{x,\text{ind}}^2 = M[(x_n - x_n^{\text{ind}})^2]$. Наилучший совместный (с учетом Y) прогноз $x_n^{\text{jo}} = M[x_n | x_n^-, y_n^-]$ дает ошибку $\sigma_{x,\text{jo}}^2 = M[(x_n - x_n^{\text{jo}})^2]$. Нормированная величина УП $G_{y \rightarrow x} = (\sigma_{x,\text{ind}}^2 - \sigma_{x,\text{jo}}^2) / \sigma_{x,\text{ind}}^2$ характеризует причинность (влияние) по Грейнджеру в направлении $Y \rightarrow X$ [23]. Аналогично определяется влияние $X \rightarrow Y$.

Эта идея была впервые реализована [22] для стационарных гауссовских процессов (x_n, y_n) . Было использовано то обстоятельство, что такой процесс единственным образом описывается двухмерным линейным уравнением авторегрессии (АР) вида

$$\begin{aligned} x_n &= \sum_{k=1}^{\infty} a_{x,k} x_{n-k} + \sum_{k=1}^{\infty} b_{x,k} y_{n-k} + \xi_n, \\ y_n &= \sum_{k=1}^{\infty} a_{y,k} y_{n-k} + \sum_{k=1}^{\infty} b_{y,k} x_{n-k} + \psi_n, \end{aligned} \quad (1)$$

где (ξ_n, ψ_n) – двухмерный гауссовский белый шум с нулевым средним, дисперсиями $(\sigma_{\xi}^2, \sigma_{\psi}^2)$ и ковариацией $M[\xi_n \psi_n] = \gamma$. Условие, что шум «белый», эквивалентно минимуму ошибки прогноза [26], так что имеет место $\sigma_{\xi}^2 = \sigma_{x,\text{joi}}^2$ и $\sigma_{\psi}^2 = \sigma_{y,\text{joi}}^2$. Далее, процесс x_n подчиняется и одномерному уравнению АР, то есть первому уравнению (1) с нулевыми $b_{x,k}$ и белым шумом ξ'_n , дисперсия которого $\sigma_{\xi'}^2 = \sigma_{x,\text{ind}}^2$. Теперь по дисперсиям шумов $\sigma_{\xi}^2, \sigma_{\xi'}^2$ рассчитывается величина $G_{y \rightarrow x}$. Аналогично рассчитывается $G_{x \rightarrow y}$.

Чтобы оценить теоретические величины $G_{y \rightarrow x}$ и $G_{x \rightarrow y}$ по конечному временному ряду $\{x_n, y_n\}_{n=1}^N$, все суммы в уравнениях (1) ограничивают p -м членом (вместо ∞) и оценивают коэффициенты и дисперсии шумов в АР-моделях порядка p с помощью стандартного метода наименьших квадратов. Ниже в численных экспериментах длина ряда N велика, так что величина p выбирается просто настолько большой, чтобы результаты оценивания практически перестали меняться с ее дальнейшим увеличением (а именно, $p = 10$ оказалось достаточным во всех примерах). Статистическая значимость отличия оценок $G_{y \rightarrow x}$ и $G_{x \rightarrow y}$ от нуля проверяется с помощью F -теста Фишера [27].

Уравнения (1) при любом Δt – это вполне допустимое представление исходной системы (X, Y) , верно отражающее свойства ее временных реализаций с шагом выборки Δt . Но при этом важен вопрос: как величины $G_{y \rightarrow x}$ и $G_{x \rightarrow y}$ зависят от Δt ? Если нет реального воздействия $Y \rightarrow X$, то естественно было бы ожидать $G_{y \rightarrow x} = 0$ при любом Δt или, по меньшей мере, $G_{y \rightarrow x} \ll 1$ и $G_{y \rightarrow x} < G_{x \rightarrow y}$. Однако ниже показано, что эти ожидания не оправдываются и, более того, относительная мера «ложной связи» $r = G_{y \rightarrow x} / G_{x \rightarrow y}$ может даже существенно превышать единицу.

2. Эффект редкой выборки в линейных системах

В качестве наглядного примера рассмотрим достаточно «физичную» модельную систему – стохастические диссипативные осцилляторы с дискретным временем

$$\begin{aligned} X(t) &= A_{X,1} X(t-1) + A_{X,2} X(t-2) + \Xi(t), \\ Y(t) &= A_{Y,1} Y(t-1) + A_{Y,2} Y(t-2) + kX(t-1) + \Psi(t), \end{aligned} \quad (2)$$

где Ξ, Ψ – независимые гауссовские белые шумы с нулевым средним и дисперсиями σ_{Ξ}^2 и σ_{Ψ}^2 , k – коэффициент связи, а собственный период колебаний X (положение пика в спектре мощности) и его время релаксации (определяет ширину пика) даются выражениями $A_{X,1} = 2 \cos(2\pi/T_X) \exp(-1/\tau_X)$ и $A_{X,2} = -\exp(-2/\tau_X)$ [28]. Все

аналогично для осциллятора Y . Система (2) может быть получена как приближенная разностная схема или результат интегрирования некоторых дифференциальных уравнений для колебательных систем. Она также может быть и в точности реализована на практике с помощью электронных [29,30] или оптических [31,32] систем с запаздывающими обратными связями, где параметр k представлял бы собой интенсивность физического воздействия.

В качестве стартовых значений для анализа были приняты небольшие значения периодов колебаний $T_X = T_Y = 5.0$ и времен релаксации $\tau_X = \tau_Y = 4.0$, дисперсии шумов $\sigma_\Xi^2 = \sigma_\Psi^2 = 1$. Сила однонаправленной связи $X \rightarrow Y$ определялась значением $k = 0.3$. Величины $G_{y \rightarrow x}$ и $G_{x \rightarrow y}$ оценивались по временным рядам достаточно большой длины $N = 10^5$, чтобы сделать пренебрежимо малыми статистические флуктуации. Рис. 1, *a-в* показывают, что однонаправленная связь адекватно отражается величинами $G_{y \rightarrow x}$ и $G_{x \rightarrow y}$ при $\Delta t = 1$: $G_{y \rightarrow x} = 0$ и $G_{x \rightarrow y} > 0$. Однако с ростом Δt положительными становятся и значения $G_{y \rightarrow x}$, то есть обнаруживается «ложная связь», что мы назовем эффектом редкой выборки. Так, $G_{y \rightarrow x}$ максимально при $\Delta t = 3$, где $r = 0.08$ достаточно велико – «нефизичное» улучшение прогноза составляет почти 10% от «правильного», то есть ложные связи весьма заметны. При дальнейшем росте Δt обе величины $G_{y \rightarrow x}$ и $G_{x \rightarrow y}$ уменьшаются, а в пределе $\Delta t \rightarrow \infty$ имеем $G_{y \rightarrow x} = G_{x \rightarrow y} = 0$, так как зависимость между значениями x, y в сильно удаленные последовательные моменты n и $n - 1$ пропадает, то есть процесс (x_n, y_n) становится белым шумом (см. $\Delta t = 15$ на рис. 1, *a, б*).

Наглядное пояснение эффекта ненулевых $G_{y \rightarrow x}$ при $\Delta t > 1$ дано на рис. 2. Заметим сначала, что в данном примере двумерный вектор $X(t), X(t - 1)$ содержит *полную* информацию о распределении X во все будущие моменты времени, в частности, об $X(t + 1)$ (то есть X – марковский процесс второго порядка [26]). Учет каких-либо более давних значений X и любых прошлых значений Y бесполезен для улучшения прогноза будущих значений X . Сплошные линии и кружки иллюстрируют это на рис. 2, *a* для $\Delta t = 1$. В терминах моделей с пространствами состояний вектор $X(t), X(t - 1)$ определяет *состояние* процесса X .

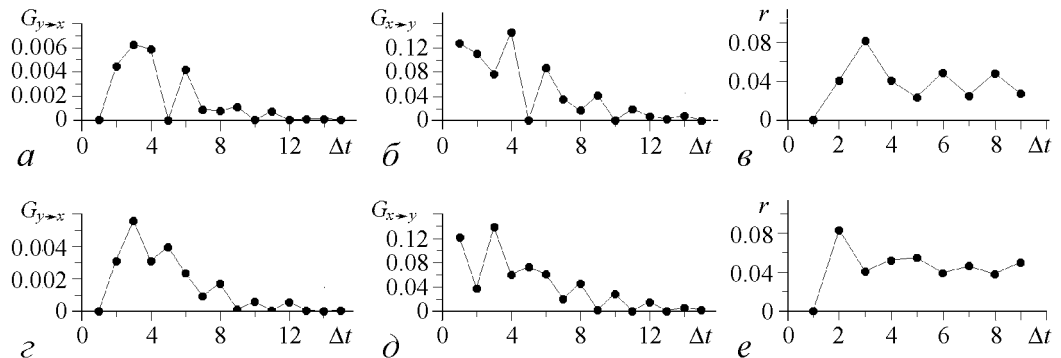


Рис. 1. Характеристики причинности по Грейнджеру для однонаправленно связанных осцилляторов (2) при $T_X = T_Y = 5.0$ (*a-в*) и $T_X = T_Y = 4.4$ (*г-е*)

Напротив, при $\Delta t = 2$ *неполная* информация о векторе $X(t), X(t - 1)$ содержится в наблюдаемом *прореженном* ряде $X(t), X(t - 2), X(t - 4), \dots$ (неполноту и означают пунктирные стрелки на рис. 2, б), так что прогноз X на основе этих данных не является наилучшим из возможных. При учете данных из ряда Y может быть получена дополнительная информация о «скрытом» значении $X(t - 1)$ (крестик на рис. 2, б) за счет корреляции между $X(t - 1)$ и $Y(t)$. Таким образом, необходимым условием «ложной связи» (УП «в обратную сторону») является то, что при редкой выборке состояние ведущего процесса *не полностью* определяется по наблюдаемым данным.

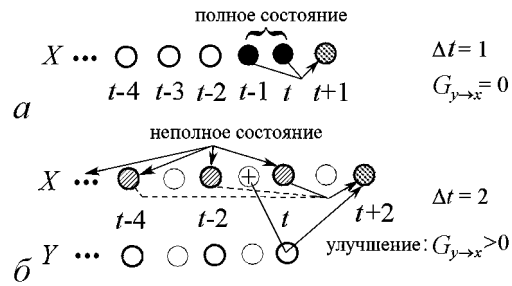


Рис. 2. Иллюстрации к оценке причинности по Грейнджеру для однонаправленно связанных осцилляторов (2): а – нет улучшения прогноза в «неправильном» направлении при $\Delta t = 1$; б – эффект редкой выборки при $\Delta t = 2$

Иллюстрацией этого утверждения является и «стробоскопический эффект» на рис. 1, а, б, где оба УП малы при $\Delta t = 5$: $G_{y \rightarrow x} \approx 2 \cdot 10^{-5}$, $G_{x \rightarrow y} = 1 \cdot 10^{-3}$. Причина этого состоит в том, что при $\Delta t = T_X = T_Y$ оба осциллятора представляют собой практически процессы авторегрессии *первого* порядка. Конкретно, представление (1) при $\Delta t = 5$ дает $a_{x,1} = 0.29$ и остальные $|a_{x,k}| < 0.002$ для осциллятора X . Если же ведущая система – АР-процесс первого порядка, то ненулевое УП в его направлении невозможно, так как его состояние – скаляр, то есть *полностью* определяется наблюдаемыми данными. Действительно, при нарушении равенства $\Delta t = T_X = T_Y$, а именно, при $T_X = T_Y = 4.4$ и $\Delta t = 5$ значения УП становятся значительно большими (рис. 1, з–е).

Положительные $G_{y \rightarrow x}$ и r наблюдаются в широких диапазонах значений параметров, как показано на рис. 3 при $\Delta t = 2$. Так, при малых шумах в ведомой системе σ_{Ψ}^2 лучше восстанавливаются состояния ведущей системы по данным об Y , а поэтому оказываются большими значения $G_{y \rightarrow x} > 0.06$, и $r > 0.1$ (рис. 3, а, б). Отношение r превышает единицу при $T_X = 3.0$ (рис. 3, з) или при $T_Y = 3.5$ (рис. 3, е). Отметим, что r столь велико только для неидентичных осцилляторов, тогда как для идентичных оно остается в пределах $r < 1$. Рис. 3, ж–к показывают, что r поразному зависит от двух времен релаксации. Это отношение растет с ростом τ_X и падает с ростом τ_Y (последнее обусловлено тем, что информация о состоянии X хуже восстанавливается по данным от Y). Следовательно, большие значения $G_{y \rightarrow x}$ должны наблюдаться, если время релаксации ведущей системы сравнительно велико, а ведомой – мало. Действительно, при $\tau_X = 10$, $\tau_Y = 1$, $\sigma_{\Psi}^2 = 0.01$ и $T_X = T_Y = 5.0$ было получено значение $G_{y \rightarrow x} = 0.14$, что превышает все значения, показанные на рис. 3.

Осцилляторы с непрерывным временем, задаваемые дифференциальными уравнениями, демонстрируют тот же эффект редкой выборки с вполне аналогичными зависимостями УП от параметров. Такие примеры не представлены здесь для краткости. Для уточнения терминологии заметим, что наблюдающиеся «нефизичные» положительные значения $G_{y \rightarrow x}$ не являются ложными, ведь улучшение прогноза дей-

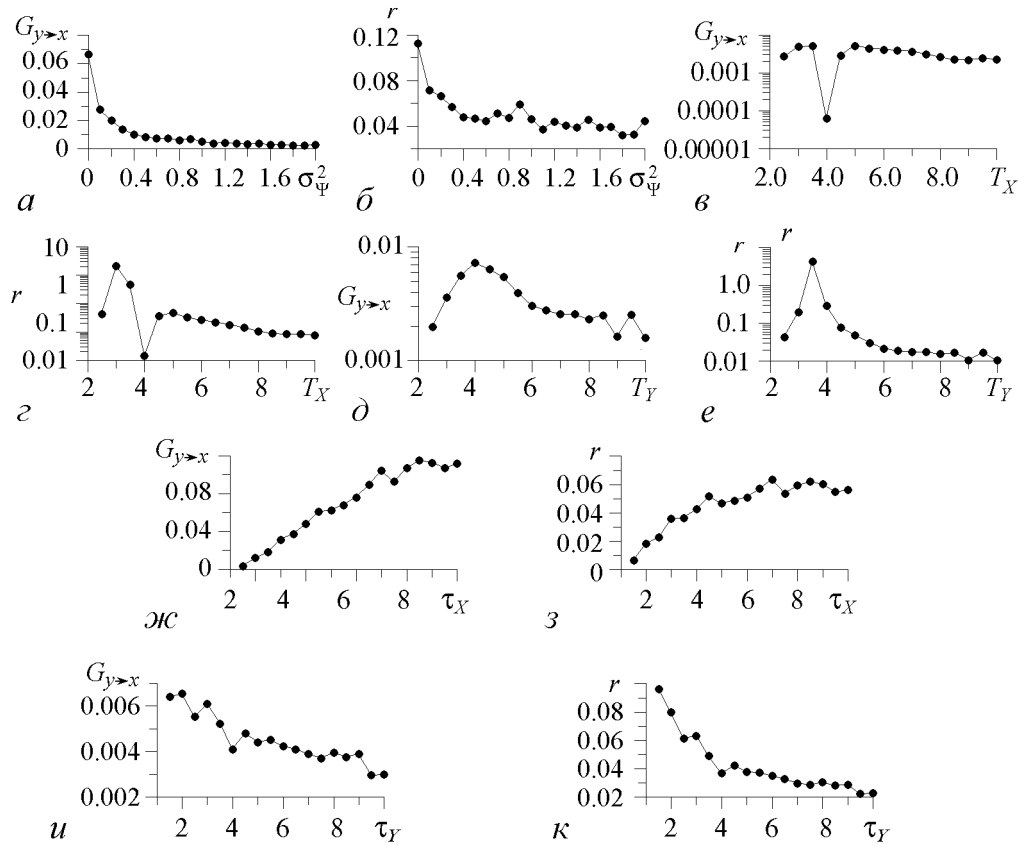


Рис. 3. Характеристики причинности по Грейнджеру для осцилляторов (2) при $\Delta t = 2$ в зависимости от параметров, указанных на осях абсцисс, при $T_X = T_Y = 5.0$, $\tau_X = \tau_Y = 4$, $\sigma_{\Xi}^2 = \sigma_{\Psi}^2 = 1$, $k = 0.3$

ствительно имеет место. Ошибочны лишь выводы о наличии воздействия $Y \rightarrow X$, сделанные из неравенства $G_{y \rightarrow x} > 0$. Поэтому мы не используем термин «ложные УП», а говорим о «ложных связях».

3. Эффект редкой выборки в нелинейных системах

При наличии нелинейности в динамике исследуемых систем эффект редкой выборки полностью сохраняется, так как его принципиальной причиной является не конкретная форма уравнений, а неполнота информации о состояниях систем в наблюдаемых данных. Изучение нелинейного случая лишь более сложно технически. Рассмотрим стохастически возмущенное отображение Спротта [33] – двухмерное квадратичное отображение вида

$$X(t) = 1 + a_X X(t-1) - b_X X^2(t-2) + \Xi(t), \quad (3)$$

где $\Xi(t)$ – последовательность независимых случайных величин, равномерно распределенных на отрезке $[-\varepsilon_{\Xi}/2, \varepsilon_{\Xi}/2]$. При $a_X = 0.4$, $b_X = 1$, $\varepsilon_{\Xi} = 0$ оно демонстрирует хаотический режим с ляпуновским показателем $\lambda = 0.08$ (рис. 4, а). Будем рассматривать ниже только интервал выборки $\Delta t = 2$. Запишем первое уравнение (3) для

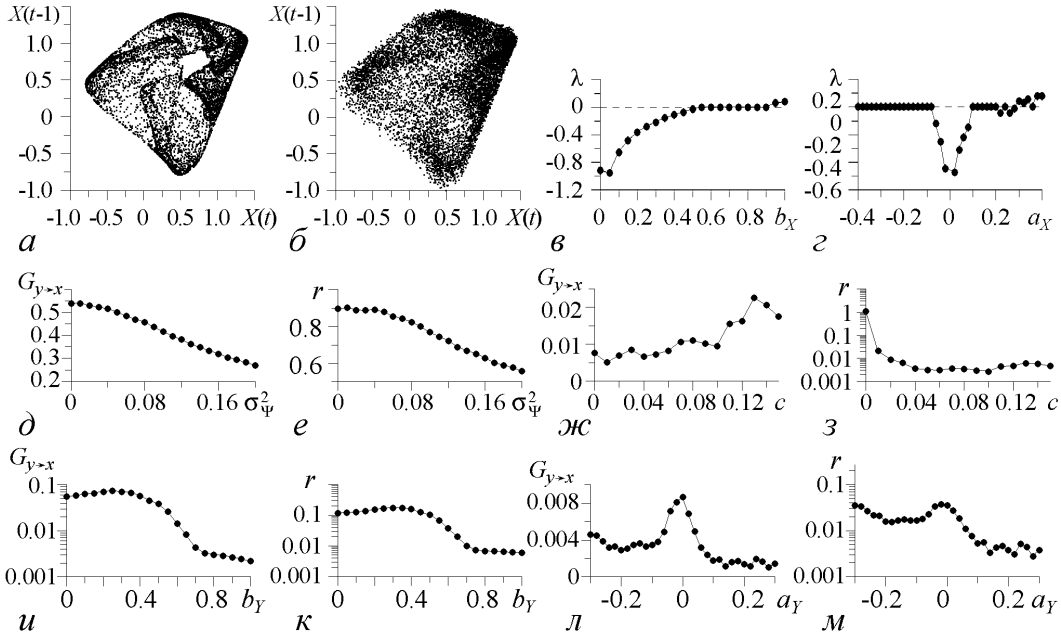


Рис. 4. Результаты для отображений Спротта (4): a – фазовая траектория X при нулевом шуме, b – при $\varepsilon_{\Xi} = 0.1$; v, z – показатель Ляпунова при $\varepsilon_{\Xi} = 0$; d – m – характеристики причинности по Грейнджеру в зависимости от параметров, указанных по осям абсцисс, при прочих фиксированных; d, e – системы (3) и (4) ($a_X = 0.4, b_X = 1, \varepsilon_{\Xi} = 0.1$); $ж$ – $м$ – системы (3) и (5) ($a_X = 0.4, a_Y = 0.3, b_X = b_Y = 1, \varepsilon_{\Xi} = \varepsilon_{\Psi} = 0.1, c = 0.13$)

моментов времени $t, t - 1$ и $t - 2$ и выразим $X(t)$ через $X(t - 2)$ и $X(t - 4)$. Получим при достаточно малом ε_{Ξ} (например, $\varepsilon_{\Xi} < 0.1$) приближенное описание X в виде нелинейного АР-процесса второго порядка с многочленом четвертой степени и дисперсией ошибки прогноза $\sigma_{x, \text{ind}}^2 \approx \sigma_{\Xi}^2(1 + a_X^2 + 4b_X^2(M[X^2])^2)$.

Пусть ведомая система описывается простым соотношением

$$Y(t) = X(t - 1) + \Psi(t), \quad (4)$$

где $\Psi(t)$ – последовательность независимых случайных величин, равномерно распределенных на отрезке $[-\varepsilon_{\Psi}/2, \varepsilon_{\Psi}/2]$. При $\varepsilon_{\Psi} = 0$ АР-модель с двумя переменными X, Y дает прогноз $X(t)$ на основе $X(t - 2), Y(t - 2), X(t - 4), Y(t - 4), \dots$ с дисперсией ошибки $\sigma_{x, \text{joit}}^2 = \sigma_{\Xi}^2(1 + a_X^2)$. Отсюда $G_{y \rightarrow x} = \frac{4b_X^2(M[X^2])^2}{1 + a_X^2 + 4b_X^2(M[X^2])^2} > 0$ и, в частности, при $a_X = 0.4, b_X = 1$ и $\varepsilon_{\Xi} = 0.1$ (рис. 4, b) получим весьма большую величину $G_{y \rightarrow x} = 0.6$. Близки и численные оценки (рис. 4, e), которые получены по временным рядам системы (3) и (4) с помощью подгонки АР-моделей четвертого порядка с многочленом четвертой степени (нелинейные оценки причинности по Грейнджеру [34–36]). С помощью таких моделей получены и оценки УП во всех примерах ниже, так как при указанных значениях порядка и степени АР-моделей стабилизировались результаты оценивания.

Если система Y задается аналогичным отображением

$$Y(t) = 1 + a_Y Y(t - 1) - b_Y Y^2(t - 2) + \Psi(t) + cX(t - 1), \quad (5)$$

то нелинейные оценки УП демонстрируют меньшие, но тоже значимо отличные от нуля $G_{y \rightarrow x}$ в широком диапазоне значений параметров (рис. 4, ж-м). При этом эффект ложных связей из-за редкой выборки наблюдается как для возмущенных хаотических, так и для возмущенных регулярных режимов ведомой системы (показатели Ляпунова представлены на рис. 4, в, з), например, для неподвижной точки при $b_Y = 0$ и хаотического режима при $b_Y = 1$ (рис. 4, и, к). Судя по величине r , эффект наиболее силен при умеренных силах связи, а по $G_{y \rightarrow x}$ – при более сильных связях (рис. 4, ж, з), но в любом случае – при таких a_Y, b_Y , которые соответствуют регулярным режимам (рис. 4, и-м). Последнее имеет место по той причине, что регулярные режимы ведомой системы обеспечивают наиболее точное следование ее состояния за состояниями ведущей системы и, значит, лучшее восстановление информации о скрытых состояниях ведущей системы. Собственная же хаотическая динамика ведомой системы снижает степень взаимозависимости состояний двух систем и возможности восстановления информации об X по Y .

Рассмотрим также более известную нелинейную систему – отображение Эно –

$$X(t) = 1 - a_X X^2(t-1) + b_X X(t-2) + \Xi(t), \quad (6)$$

где $\Xi(t)$ – то же, что в предыдущем примере. При $a_X = 1, b_X = 0.4, \varepsilon_\Xi = 0$ отображение демонстрирует хаотический режим с показателем Ляпунова $\lambda = 0.1$ (рис. 5, а, в), а шум приводит к некоторому его «размыванию» (рис. 5, б). Пусть Y задается уравнением (4) при $\varepsilon_\Psi = 0$. Тогда, аналогично предыдущему примеру, работая с более сложными алгебраическими выражениями, при условии малых шумов получим $G_{y \rightarrow x} \approx b_X^2 / (1 + 4a_X^2 M[X^2] + b_X^2) > 0$. Эта величина вновь положительна, но ее численные значения гораздо меньше, чем для отображения Спротта, даже для качественно схожих режимов динамики. Это подтверждается и численными оценками на рис. 5, з, д. Разница между двумя примерами имеет место за счет разных нелинейных функций, связывающих последовательные значения переменных, и, следовательно, разных потерь информации при прореживании и разных возможностей

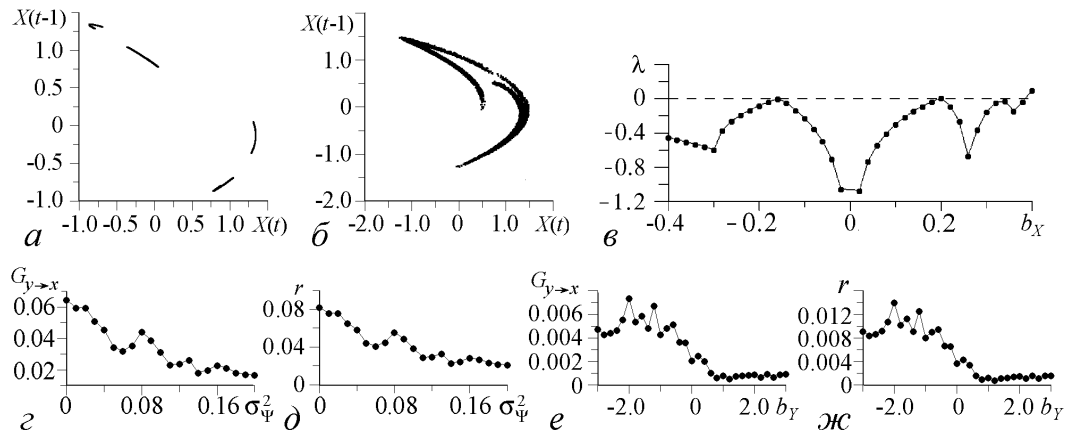


Рис. 5. Результаты для отображений Эно (6): а – фазовая траектория X при нулевом шуме, б – при $\varepsilon_\Xi = 0.1$; в – показатель Ляпунова при $\varepsilon_\Xi = 0$; г-ж – характеристики причинности по Грейнджеру в зависимости от параметров, указанных на осях абсцисс при прочих фиксированных; з, д – системы (6) и (4) ($a_X = 1, b_X = 0.4, \varepsilon_\Xi = 0.1$); е, ж – системы (6) и (7) ($b_X = 0.4, a_X = a_Y = 1, \varepsilon_\Xi = \varepsilon_\Psi = 0.1, c = 0.1$)

ее восстановления при учете Y . В основном, различаются, как показывает более детальный анализ, эти последние возможности.

Для системы Y , заданной отображением Эно

$$Y(t) = 1 - a_Y Y^2(t-1) + b_Y Y(t-2) + \Psi(t) + cX(t-1), \quad (7)$$

оценки УП показаны на рис. 5, *е, ж*. Эффект ложной связи вновь наблюдается как для возмущенных хаотических, так и для возмущенных регулярных режимов ведомой системы (показатели Ляпунова даны на рис. 5, *в*). Однако, в отличие от отображений Спротта, количественные значения «ложных связей» заметно меньше.

Таким образом, эффект ложных связей из-за редкой выборки в случае нелинейных систем также имеет место, как и для линейных примеров. При этом важно отметить, что численные значения УП могут сильно различаться даже для систем, схожих по динамике и «формулам» в уравнениях динамики. Эти значения зависят от детального характера нелинейных функций, связывающих последовательные значения переменных.

Наконец, отметим, что непременным условием эффекта редкой выборки является наличие стохастического возмущения (шума) в исследуемых системах: в строго детерминированных системах улучшения прогноза в обе стороны теоретически равны нулю из-за того, что согласно теореме Такенса [37] индивидуальный прогноз абсолютно точен при использовании вектора временных задержек (индивидуальной предыстории процесса) достаточно большой размерности. Таким образом, роль шумов в данной задаче существенная, как и во многих других задачах анализа временных рядов, см. например, работы, посвященные собственно вопросам моделирования и прогноза [38,39]. С практической точки зрения приведенное утверждение о случае детерминированных систем справедливо для систем невысокой размерности. При их высокой размерности и недостаточно богатой статистике индивидуальный прогноз может быть далеко не точным, что может вести к аналогичным ложным выводам о наличии связи, то есть различие между стохастическими и высокоразмерными детерминированными системами должно быть практически не заметно.

Заключение

При анализе направленных связей по временным рядам с помощью характеристик причинности по Грейнджеру следует учитывать, что интервал выборки, обусловленный процедурой измерений, может существенно влиять на результаты оценивания. Игнорирование этого фактора может приводить к ошибочным выводам о характере взаимодействия исследуемых систем, если они имеют меньший «собственный временной шаг». Как показано в данной работе, при достаточно большом интервале выборки характеристики связи (улучшения прогноза) могут оказаться положительными «в обе стороны» для систем с однонаправленной связью, то есть ошибочно свидетельствовать о двунаправленной связи. Причем эти значения могут быть достаточно большими, а в случае неидентичных систем УП в «ложном» направлении может даже существенно превосходить УП в «правильном» направлении. Показано также, что этот эффект редкой выборки проявляется как для линейных, так и для нелинейных систем, а для последних – в режимах как регулярной, так и

хаотической динамики. Причина такой универсальности состоит в том, что эффект определяется неполнотой информации о векторе состояния ведущей системы в наблюдаемых данных, а не видом оператора эволюции и даже не режимом динамики.

Отсюда следует, в частности, что активно развиваемая идея выбора оптимальных (не малых) временных лагов при оценке нелинейной причинности по Грейнджеру [40,41] также может вести к ложным выводам о двунаправленной связи, так что ее применение требует осторожности и учета эффекта редкой выборки. Таким образом, надежные выводы о двунаправленной связи на практике должны делаться на основе специального теста с учетом конечной величины интервала выборки. Вариант такого теста представлен в параллельной работе [42].

Работа выполнена при поддержке РФФИ (гранты 11-02-00599, 11-05-01139, 12-02-00377), программ РАН и ФЦП «Научные и научно-педагогические кадры инновационной России на 2009-2013 годы» (контракт № 14.В37.21.0247).

Библиографический список

1. *Блехман И.И.* Синхронизация динамических систем. М.: Наука, 1971.
2. *Пиковский А.С., Розенблюм М.Г., Куртс Ю.* Синхронизация. Фундаментальное нелинейное явление. М.: Техносфера, 2003.
3. *Анищенко В.С., Астахов В.В., Вадивасова Т.Е., Нейман А.Б., Стрелкова Г.И., Шиманский-Гайер Л.* Нелинейные эффекты в хаотических и стохастических системах. Москва-Ижевск: Институт компьютерных исследований, 2003.
4. *Кузнецов А.П., Емельянова Ю.П., Сатаев И.Р., Тюрюкина Л.В.* Синхронизация в задачах. Саратов: ООО издательский центр «Наука», 2010.
5. *Rosenblum M.G., Pikovsky A.S., Kurths J., Schafer C., Tass P.A.* Phase synchronization: from theory to data analysis // in Moss F., Gielen S (eds). Neuro-informatics. Handbook of Biological Physics (v. 4). New York, Elsevier Science, 2000. P. 279.
6. *Вадивасова Т.Е., Анищенко В.С.* Взаимосвязь частотных и фазовых характеристик хаоса. Два критерия синхронизации // Радиотехника и электроника. 2004. Т. 49, № 1. С.1.
7. *Hramov A.Ye., Koronovskii A.A.* An approach to chaotic synchronization // Chaos. 2004. Vol. 14, № 3. P. 603.
8. *Hramov A.E., Koronovskii A.A., Ponomarenko V.I., Prokhorov M.D.* Detection of synchronization from univariate data using wavelet transform // Phys. Rev. E. 2007. Vol. 75, № 5. 056207.
9. *Pavlov A.N., Sosnovtseva O.V., Pavlova O.N., Mosekilde E., Holstein-Rathlou N.-H.* Characterizing multimode interaction in renal autoregulation // Physiological Measurement. 2008. Vol. 29. P. 945.
10. *Kralemann B., Cimponeriu L., Rosenblum M., Pikovsky A., Mrowka R.* Uncovering interaction of coupled oscillators from data // Phys. Rev. E. 2007. Vol. 76, № 5. 055201(R).
11. *Bezruchko B., Ponomarenko V., Rosenblum M.G., Pikovsky A.S.* Characterizing direction of coupling from experimental observations // Chaos. 2003. Vol.13, № 1. P. 179.

12. *Hung Y.-C., Hu C.-K.* Chaotic communication via temporal transfer entropy // *Phys. Rev. Lett.* 2008. Vol. 101. 244102.
13. *Palus M., Novotna D.* Quasi-biennial oscillations extracted from the monthly NAO index and temperature records are phase-synchronized // *Nonlinear Processes in Geophysics.* 2006. Vol.13, № 3. P.287.
14. *Мохов И.И., Смирнов Д.А.* Исследование взаимного влияния процессов Эль-Ниньо – Южное колебание и Северо-Атлантического и Арктического колебаний нелинейными методами // *Изв. РАН. Физика атмосферы и океана.* 2006. Т. 42, № 5. С.650.
15. *Мохов И.И., Смирнов Д.А., Карпенко А.А.* Оценки связи изменений глобальной приповерхностной температуры с разными естественными и антропогенными факторами на основе данных наблюдений // *Доклады академии наук.* 2012. Т. 443, № 2. С. 225.
16. *Pereda E., Quian Quiroga R., Bhattacharya J.* Nonlinear multivariate analysis of neurophysiological signals // *Progr. in Neurobiology.* 2005. Vol.77. P. 1.
17. *Brea J., Russell D.F., Neiman A.B.* Measuring direction in the coupling of biological oscillators: A case study for electroreceptors of paddlefish // *Chaos.* 2006. Vol. 16, № 2. 026111.
18. *Безручко Б.П., Пономаренко В.И., Прохоров М.Д., Смирнов Д.А., Тасс П.А.* Моделирование и диагностика взаимодействия нелинейных колебательных систем по хаотическим временным рядам (приложения в нейрофизиологии) // *Успехи физических наук.* 2008. Т. 178. № 3. С. 323.
19. *Smirnov D., Barnikol T., Barnikol U., Bezruchko B.P., Hauptmann C., Buehrle C., Maarouf M., Sturm V., Freund H.-J., Tass P.A.* The generation of parkinsonian tremor as revealed by directional coupling analysis // *Europhysics Letters.* 2008. Vol. 83. 20003.
20. *Сысоева М.В., Сысоев И.В.* Математическое моделирование динамики энцефалограммы во время эпилептического припадка // *Письма в ЖТФ.* 2012. Т. 38. Вып. 3. С. 103.
21. *Филина Е.В.* Динамика локальных потенциалов мозга при абсанс-эпилепсии: эмпирическое моделирование // *Известия вузов. Прикладная нелинейная динамика.* 2011. Т. 19, № 4. С.109.
22. *Granger C.W.J.* Investigating causal relations by econometric models and cross-spectral methods // *Econometrica.* 1969. Vol. 37, № 3. P. 424.
23. *Granger C.W.J.* Testing for causality. A personal viewpoint // *J. Economic Dynamics and Control.* 1980. Vol. 2. P. 329.
24. *Sims C.A.* Discrete approximations to continuous time distributed lags in econometrics // *Econometrica.* 1971. Vol. 39. P. 545.
25. *Marcellino M.* Some consequences of temporal aggregation in empirical analysis // *J. Business and Economic Statistics.* 1999. Vol. 17. P. 129.
26. *Бокс Дж., Дженкинс Г.* Анализ временных рядов. Прогноз и управление. Ч. 1, 2. М.: Мир, 1974.

27. Себер Дж. Линейный регрессионный анализ. М.: Мир, 1974.
28. Timmer J., Lauk M., Pfleger W., Deuschl G. Cross-spectral analysis of physiological tremor and muscle activity: I. Theory and application to unsynchronized electromyogram // Biol. Cybern. 1998. Vol. 78. P. 349.
29. Кузнецов С.П. Бифуркации удвоения в простой модели распределенной системы // Изв. вузов. Радиофизика. 1982. Т. 25, № 11. С. 1364.
30. Кузнецов С.П. Универсальность и подобие в поведении связанных систем Фейгенбаума // Изв. вузов. Радиофизика. 1985. Т. 28, № 8. С. 991.
31. Ikeda K. Multiple-valued stationary state and its instability of the transmitted light by a ring cavity system // Opt. Commun. 1979. Vol. 30. P. 257.
32. Lang R., Kobayashi K. External optical feedback effects on semiconductor injection laser properties // IEEE J. Quantum Electron. 1980. Vol. 16, № 1. P. 347.
33. Elhadj Z., Sprott J.C. A minimal 2-D quadratic map with quasi-periodic route to chaos // Int. J. Bifurcation Chaos. 2008. Vol. 18, № 5. P. 1567.
34. Ancona N., Marinazzo D., Stramaglia S. Radial basis function approach to nonlinear Granger causality of time series // Phys. Rev. E. 2004. Vol. 70, № 5. 056221.
35. Marinazzo D., Pellicoro M., Stramaglia S. Kernel method for nonlinear Granger causality // Phys. Rev. Lett. 2008. Vol. 100. 144103.
36. Hlavackova-Schindler K., Palus M., Vejmelka M., Bhattacharya J. Causality detection based on information-theoretic approaches in time series analysis // Physics Reports. 2007. Vol. 441, № 1. P. 1.
37. Takens F. Detecting strange attractors in turbulence // Lec. Notes in Math. 1981. Vol. 898. P. 366.
38. Molkov Ya.A., Mukhin D.N., Loskutov E.M., Feigin A.M., Fidelin G.A. Using the minimum description length principle for global reconstruction of dynamic systems from noisy time series // Phys. Rev. E. 2009. Vol. 80, № 4. 046207
39. Яхно Ю.В., Мольков Я.И., Мухин Д.Н., Лоскутов Е.М., Фейгин А.М. Реконструкция оператора эволюции как способ анализа электрической активности мозга при эпилепсии // Известия вузов. Прикладная нелинейная динамика. 2011. Т. 19, № 6. С. 156.
40. Vlachos I., Kugiumtzis D. Nonuniform state-space reconstruction and coupling detection // Phys. Rev. E. 2010. Vol. 82. № 1. 016207.
41. Faes L., Nollo G., Porta A. Information-based detection of nonlinear Granger causality in multivariate processes via a nonuniform embedding technique // Phys. Rev. E. 2011. V. 83, № 5. 051112.
42. Smirnov D.A., Bezruchko B.P. Spurious causalities due to low temporal resolution: Towards detection of bidirectional coupling from time series // Europhys. Lett. 2012. Vol. 100. P. 10005.

Саратовский государственный
университет им. Н.Г. Чернышевского
Саратовский филиал ИРЭ
им. В.А. Котельникова РАН

Поступила в редакцию 3.10.2012
После доработки 29.04.2013

EFFECT OF RARE SAMPLING ON ESTIMATION OF DIRECTIONAL COUPLINGS FROM TIME SERIES

D. A. Smirnov, B. P. Bezruchko

The problem of detection and quantitative estimation of directional couplings (mutual influences) between systems from discrete records of their oscillations (time series) arises in different fields of research. This work shows that results of the traditional «Granger causality» approach depend essentially on a sampling interval (a time step). We have revealed the causes and character of the influence of a sampling interval on numerical values of coupling estimates. As well, we have explained why one can get erroneous conclusions about bidirectional coupling for unidirectionally coupled systems in the case of a large sampling interval (rare sampling). The rare sampling effect is demonstrated both for linear and nonlinear systems in different dynamical regimes.

Keywords: Coupled nonlinear systems, time series, unidirectional coupling, bidirectional coupling, Granger causality.

Смирнов Дмитрий Алексеевич – родился в Саратовской обл. (1977), окончил факультет нелинейных процессов Саратовского государственного университета (1999). Работает в Саратовском филиале Института радиотехники и электроники им. В.А. Котельникова РАН ведущим научным сотрудником. Защитил диссертации на соискание ученой степени кандидата физико-математических наук (2001) и доктора физико-математических наук (2010) в СГУ в области теории колебаний, нелинейной динамики и их приложений для анализа временных рядов в биофизике и геофизике. Автор монографий (в соавторстве с Б.П. Безручко) «Математическое моделирование и хаотические временные ряды» (Саратов, изд-во ГосУНЦ «Колледж» 2005) и «Extracting knowledge from time series: An introduction to nonlinear empirical modeling» (Berlin, Heigelberg, Springer-Verlag 2010). Опубликовал более 50 научных статей в рецензируемых отечественных и международных журналах.



410019 Саратов, ул. Зеленая, 38
Саратовский филиал Института радиотехники и электроники
имени В.А. Котельникова РАН
E-mail: smirnovda@yandex.ru

Безручко Борис Петрович – родился в Саратовской обл. (1946), окончил физический факультет Саратовского государственного университета (1969). Работает заведующим кафедрой динамического моделирования и биомедицинской инженерии СГУ и заведующим лабораторией моделирования в нелинейной динамике в Саратовском филиале Института радиотехники и электроники им. В.А. Котельникова РАН. Защитил диссертации на соискание ученой степени кандидата физико-математических наук в (1980) и доктора физико-математических наук (1994) в СГУ в области теории колебаний и радиофизики. Автор монографий (в соавторстве с Д.А. Смирновым) «Математическое моделирование и хаотические временные ряды» (Саратов, изд-во ГосУНЦ «Колледж» 2005) и «Extracting knowledge from time series: An introduction to nonlinear empirical modeling» (Berlin, Heigelberg, Springer-Verlag 2010). Опубликовал более 100 научных статей в рецензируемых отечественных и международных журналах.



410012 Саратов, ул. Астраханская, 83
Саратовский государственный университет им. Н.Г. Чернышевского
E-mail: bezruchkobp@gmail.com