УДК 519.2:005

ОЦЕНКА ИНФОРМАЦИОННОЙ ПРОГНОЗИРУЕМОСТИ СТОХАСТИЧЕСКИХ ПРОЦЕССОВ

А.В. ОВСЯННИКОВ

Белорусский государственный университет
пр. Независимости, 4, Минск, 220030, Беларусь

Поступила в редакцию 10 октября 2014

Приведены необходимые теоретические сведения, обеспечивающие получения алгоритмов параметрической оценки информационной прогнозируемости стохастических процессов. Рассмотрены примеры алгоритмов оценивания информационной прогнозируемости для процессов, описываемых оптимальными разностными схемами, эквивалентными прогностическим моделям в виде стохастических дифференциальных уравнений.

Ключевые слова: информационная прогнозируемость, стохастический процесс, одношаговая плотность перехода, алгоритм оценки прогнозируемости.

Введение

Введенная в работах [1, 2] функция информационной прогнозируемости (ИП), в общем случае, нестационарного стохастического процесса и его параметров, определяется изменением во времени количества информации Фишера относительно этих параметров или самого процесса в целом. Удобство такого подхода к определению прогнозируемости, в отличии от классических спектрально-корреляционных и энтропийных методов, состоит в следующем. Во-первых, в его естественной связи с информацией Фишера и, соответственно, алгоритмами оценивания, качественными и количественными характеристиками этих алгоритмов. Во-вторых, функция ИП строго положительна и имеет конкретное численное выражение для любого будущего момента времени. В-третьих, этот подход применим и к ИП стохастических систем, поведение которых описывается соответствующими стохастическими процессами.

Особый интерес представляет взаимосвязь функции ИП с величиной горизонта прогностической модели процесса, заданной стохастическим дифференциальным уравнением (СДУ) или нестационарной плотностью. Если предположить, что такая модель известна с точностью до постоянных коэффициентов, то вычисленная функция ИП количественно точно, с заданной достоверностью будет определять информацию о процессе в любые будущие моменты времени, соответствующие заданному горизонту прогностической модели.

В том случае, когда априорные данные о прогностической модели ограничены ее видом, требуется решать задачу параметрического оценивания (фильтрации) неизвестных коэффициентов модели по результатам текущих наблюдений. В рамках этой задачи нашли применение методы, работающие на коротких временных участках, в частности, адаптивные методы [3,4] и методы с использованием ε-стационарности наблюдаемой части временного ряда [5]. Однако, в первом случае рассматриваются лишь линейные разностные схемы с адаптацией коэффициентов экспоненциального сглаживания прогностических моделей, а во втором требуется решение дополнительной задачи по определению границ участка ε-стационарности [6] для построения соответствующих эволюционных уравнений выборочной функции распределения. В связи с этим в данной работе для оценки неизвестных коэффициентов параметрической прогностической модели предлагается использовать выборочную многомерную плотность вероятности (ВМП), сформированную из одношаговых плотностей перехода (ОПП) марковского процесса.

Цель работы состоит в исследовании и получении теоретических инструментов, обеспечивающих определение информационной прогнозируемости стохастических процессов с использованием алгоритмов оценки параметров прогностической модели.

Теоретические инструменты параметрической оценки информационной прогнозируемости стохастического процесса

Будем рассматривать процесс, модель которого в непрерывном времени представляется СДУ: $\dot{y}(t)+a(t,y)=g(t)\zeta(t)$, где $a(t,\xi)$, $g(t)$ – известные детерминированные функции удовлетворяющие условию Липшица, $\zeta(t)$ – нормальный белый шум с нулевым средним $\langle\zeta(t)\rangle=0$ и дельтообразной корреляционной функцией $\langle\zeta(t)\zeta(t')\rangle=N\delta(t'-t)/2$, N – односторонняя спектральная плотность. Оптимальная разностная схема, при понимании СДУ в форме Ито, имеет вид [7]

$$y_{i+1}=y_i-\Delta a_i+\Delta g_i\zeta_i,\ \ i=\overline{1,k}, \tag{1}$$

где $\Delta=t_{i+1}-t_i$ малый интервал времени, на котором слева и справа интегрируется СДУ, $a_i=a(t_i,y_i)$, $\zeta_i=(1/\Delta)\int_{t_i}^{t_i+\Delta}\zeta(t)dt$, $b(t_i)=\lim\limits_{\Delta\to 0}(N/2\Delta)\int_{t_i}^{t_i+\Delta}g^2(t)dt$ – коэффициент диффузии (в случае $g=\mathrm{const}$ этот коэффициент равен $b=Ng^2/2$), $M_i=y_i-\Delta a_i$ – математическое ожидание, $D_i=b(t_i)\Delta$ – дисперсия. Глобальная среднеквадратическая погрешность разностной схемы (1) определяется величиной [7]: $\sigma\le\Delta\left(\int_0^T M\left[(a'_y g(t))^2\right]dt\right)^{1/2}$, $T=k\Delta$. В дальнейшем, там где это не вызывает сомнений, зависимость функций от аргументов опускается.

При наблюдении на интервале времени $[t_1\ ;t_k]$ имеется последовательность $Y=[y_1\ ,...,y_k]$ и соответствующая ей ВМП $P_k=P(Y\,|\,X)$, где $X=\{x_q\}$, $q=\overline{1,Q}$ – набор независимых параметров ВМП. В качестве такого набора параметров ВМП могут выступать, например, математическое ожидание процесса, его дисперсия, любой другой параметр, зависящий от времени.

Согласно работам [1,2] ИП одного из параметров x_q ВМП P_k, в общем случае, нестационарного стохастического процесса определяется выражением

$$IP_{x_q}(t_k)=\left\langle\left(\partial\ln P_k/\partial x_q\right)^2\right\rangle_Y, \tag{2}$$

а ИП процесса в целом есть неотрицательная функция

$$IP_Y(t_k)=\left\langle\left(\partial\ln P_k/\partial t_k\right)^2\right\rangle_Y=\sum_q\dot{x}_q^2 IP_{x_q}(t_k). \tag{3}$$

Таким образом, оценка ИП состоит в последовательном вычислении ВМП P_k и, далее, непосредственном использовании формул (2) и (3). Наличие априорной информации об элементах уравнения (1) или параметрах ВМП приводит соответственно к параметрическим или непараметрическим алгоритмам. В этой статье рассматриваются только параметрические алгоритмы. В этом случае ВМП удобно представить в виде

$$P_k=C_k^{-1}\exp\left(-\sum_{i=0}^{k-1}B_{i+1,i}\right), \tag{4}$$

где $B_{i+1,i}=B[y_{i+1},y_i,X_i]$ – семейство известных параметрических функций, $X_i=X(t_i,\Theta)$ – значение параметра ВМП в i-тый момент времени, Θ – набор неизвестных параметров плотности, требующих оценки, P_1 – плотность вероятности начальной координаты,

$B_{1,0} = B[y_1, X_1]$, C_k – нормирующий множитель. Если выбрать функции $B_{i+1,i} = -\ln(c_{i+1,i}\pi_{i+1,i})$, где $\pi_{i+1,i}$ – одношаговая плотность перехода, $c_{i+1,i}$ – нормирующие множители, связь между выражением (4) и выборочным эмпирическим функционалом (ВЭФ) определяется зависимостью $W_k = k^{-1}\left[\sum_{i=0}^{k-1} B_{i+1,i} - \ln C_k^{-1}\right]$. В этом случае функция $B_{i+1,i}$ представляет собой частную информационную функцию потерь. Положив $C_k = \left(\prod_{i=0}^{k-1} c_{i+1,i}\right)$, получим

$$P_k = \prod_{i=0}^{k-1} \pi_{i+1,i} \tag{5}$$

и, следовательно, ВЭФ принимает вид $W_k = -\ln(P_k)/k$.

Теперь, с учетом (5), входящие в формулу (2) компоненты можно представить в виде

$$\ln P_k = \sum_{i=0}^{k-1} \ln \pi_{i+1,i},$$

$$\left(\frac{\partial \ln P_k}{\partial x}\right)^2 = \left(\sum_{i=0}^{k-1} \frac{\partial \ln \pi_{i+1,i}}{\partial x_i}\right)^2 = \sum_{i=0}^{k-1}\left(\frac{\partial \ln \pi_{i+1,i}}{\partial x_i}\right)^2 + 2\sum_{i=0}^{k-2}\sum_{j=i+1}^{k-1} \frac{\partial \ln \pi_{i+1,i}}{\partial x_i}\frac{\partial \ln \pi_{j+1,j}}{\partial x_j}.$$

На основании последних формул, рекуррентная формула вычисления информационной прогнозируемости (2) следующая: $IP_{x_q}(t_k) = IP_{x_q}(t_{k-1}) + IP_{x_q}(\Delta t_k)$, где одношаговая ИП

$$IP_{x_q}(\Delta t_k) = \left\langle \left(\frac{\partial \ln \pi_{k,k-1}}{\partial x_{k-1}}\right)^2 \right\rangle_Y + \left\langle 2\sum_{i=0}^{k-2} \frac{\partial \ln \pi_{i+1,i}}{\partial x_i}\frac{\partial \ln \pi_{k,k-1}}{\partial x_{k-1}} \right\rangle_Y.$$

В качестве задаваемой (или считающейся известной) ОПП, используется плотность формальной конструкции – обобщенное нормальное распределение

$$\pi_{i+1,i} = B(u)\exp\left(-A(u)\left(\left|y_{i+1} - M_i\right| / \sqrt{D_i}\right)^u\right),\ u \ge 0{,}5, \tag{6}$$

где $A(u) = \left(\Gamma(3/u)\Gamma^{-1}(1/u)\right)^{u/2}$, $B(u) = A(u)^{1/u} / (2\sqrt{D_i}\Gamma(1+1/u))$, $\Gamma(z)$ – гамма-функция. Наиболее часто такие модели используются с параметром $u \in \{0{,}5;1;2\}$.

Далее, для иллюстрации излагаемого подхода, ограничимся рассмотрением задачи оценки прогнозируемости параметров гауссовского процесса (математического ожидания и дисперсии), являющейся, пожалуй, самой распространенной в рамках задачи прогнозируемости, т.е. $x = [M_i, D_i]$ и u = 2. Формула (5) примет следующий известный вид

$$P_k = \left[(2\pi)^k \det(D_k)\right]^{-1/2} \exp\left(-\frac{1}{2}(Y - M_k)^T D_k^{-1}(Y - M_k)\right), \tag{7}$$

где вектор $M_k = \{M_i\}$, матрица $D_k = \mathrm{diag}\{D_i\}$, $D_i = f(t_i, d)$ $i = \overline{1,k}$, $d = \{d_v\}$ – набор неизвестных параметров функции масштаба (дисперсии) $v = \overline{0,V}$. Тогда, для прогнозируемости математического ожидания, справедлива формула

$$IP_x(t_k) = \sum_{i=0}^{k-1} \int_{y_{i+1}} \int_{y_i} \left(\frac{\partial \ln \pi_{i+1,i}}{\partial x}\right)^2 P(y_i, y_{i+1}) dy_i dy_{i+1},\ x = [M_i, D_i].$$

Таким образом, для ОПП (6) с параметром u = 2, на основании последних выражений, получаем:

$$IP_M(t_{k+m}) = tr(D_{k+m}^{-1}) = \sum_{i=1}^{k} D_{i+m}^{-1}, \tag{8}$$

$$IP_D(t_{k+m}) = 0{,}5\sum_{i=1}^{k} D_{i+m}^{-2}, \quad m = 0,1,\ldots \tag{9}$$

Здесь индекс m означает прогнозируемость соответствующих параметров стохастического процесса по результатам его наблюдения на интервале $[t_1,...,t_k]$ на будущие моменты времени t_{k+m}. На интервале $[t_1,...,t_k]$ формируется оценка вектора неизвестных $\Theta_k^* = d_k^*$, входящих в функции D. Таким образом, задача оценки прогнозируемости параметров $[M, D]$ и самого стохастического процесса в целом y сводится к оценке выборочной дисперсии D^*, точнее коэффициентов d^*. Такую оценку можно получить, используя стандартные статистические процедуры, применяемые к формулам (6), (7) и ВЭФ W_k.

Далее необходимо сделать следующее важное замечание. Поскольку для выборочных моментов нестационарных рядов отсутствуют утверждения об их состоятельности как оценок моментов соответствующих распределений, то нет необходимости определять их так, чтобы они давали несмещенную оценку, как в случае стационарных рядов [5]. При этом использование таких выборочных моментов не предполагает их асимптотической несмещенности, т.к. асимптотическая сходимость не определена для таких процессов ни в слабом смысле, ни по вероятности. Однако в данной работе неизвестные функции D_i, входящие в формулы (8), (9), считаются параметрически заданными, оценке подлежат только неизвестные коэффициенты d, и, в этом случае, применение метода максимального правдоподобия приводит к оценкам, которые, как и все оценки этого вида, являются состоятельными, асимптотически эффективными и асимптотически нормальными [8, 9].

Метод максимального правдоподобия дает систему оценок для вектора d в точке $d = d^*$:

$$\frac{\partial W_k}{\partial d_l} = \frac{\partial W_k}{\partial D_k}\frac{\partial D_k}{\partial d_l} = 0, \quad \frac{\partial W_k}{\partial D_k} = \frac{\partial W_k}{\partial D_k} = \frac{1}{2}\det(D_k)^{-1} - \frac{1}{2}(Y - M_k)^T D_k^{-2}(Y - M_k)\,. \tag{10}$$

Рассмотрим пример оценки прогнозируемости для стохастического процесса следующего вида: $\dot{y}(t) + \mu y(t) = \sqrt{gt}\zeta(t)$, $t > 0$, $y(t_1) = y_1$ – начальное значение координаты, $\mu > 0$ – известный параметр, $g > 0$ – неизвестный параметр. Оптимальная разностная схема, эквивалентная приведенному СДУ: $y_{i+1} = (1-\Delta\mu)y_i + \Delta\sqrt{gt_i}\zeta_i$. Математическое ожидание и дисперсия ОПП соответственно равны $M_i = (1-\Delta\mu)y_i$, $D_i = dt_i$, $d = gN\Delta/2$. Тогда, с учетом соотношений (10), получаем в явном виде оценку

$$d_k^* = \frac{1}{k}\sum\nolimits_{i=1}^{k}\frac{(y_{i+1} - M_i)^2}{t_i}\,, \tag{11}$$

или в случае $\Delta << 1$

$$d_k^* = \frac{1}{k}\sum\nolimits_{i=1}^{k}\frac{(y_{i+1} - ry_i)^2}{t_i}\,, \quad r = \exp(-\Delta\mu)\,, \tag{12}$$

откуда неизвестное значение g равно:

$$g_k^* = \frac{2}{N\Delta k}\sum\nolimits_{i=1}^{k}\frac{(y_{i+1} - ry_i)^2}{t_i}\,.$$

Теперь, с учетом определенных величин, оценки (8), (9) примут окончательный вид

$$IP_M^*(t_{k+m}) = (d_k^*)^{-1} tr(t_{k+m}^{-1}) = (d_k^*)^{-1}\sum\nolimits_{i=1}^{k} t_{i+m}^{-1}\,, \tag{13}$$

$$IP_D^*(t_{k+m}) = 0,5(d_k^*)^{-2}\sum\nolimits_{i=1}^{k} t_{i+m}^{-2}, \quad m = 0,1,...\,. \tag{14}$$

Если положить $t_i = i\Delta$, то оценки (13) и (14) примут вид

$$IP_M^*(t_{k+m}) = (\Delta d_k^*)^{-1}\left[H(k+m) - H(m)\right], \tag{15}$$

$$IP_D^*(t_{k+m}) = 0,5(\Delta d_k^*)^{-2}\left[\Psi_1(1+m) - \Psi_1(1+k+m)\right], \tag{16}$$

где $\Psi_n(z) = \Psi^{(n)}(z)$, $\Psi(z) = \Gamma_z{}'(z)/\Gamma(z)$, $H(n)$ – n-ое гармоническое число.

Отметим, разность $H(k+m)-H(m)$ представляет собой скользящее по m временное окно, в котором производится суммирование частичного гармонического ряда длинной k элементов. При использовании известной приближенной формулы Эйлера, формула (15) может быть заменена на следующую $IP_M^*(t_{k+m}) \approx (\Delta d_k^*)^{-1}\ln(1+k/m)$.

Прогнозируемость стохастического процесса в целом, определяемая формулой (3), на основании (13) и (14) равна

$$IP_Y(t_{k+m}) = \sum_q \dot{x}_q^2 IP_{x_q}(t_k) = 0{,}5\sum_{i=1}^k t_{i+m}^{-2}, \quad m = 0,1,\dots ,$$

и, если $t_i = i\Delta$, то

$$IP_Y^*(t_{k+m}) = 0{,}5(\Delta)^{-2}\left[\Psi_1(1+m)-\Psi_1(1+k+m)\right], \quad m = 0,1,\dots . \tag{17}$$

Однако оценки типа (11) и (12) в явном виде, возможны лишь для ограниченного класса СДУ. В практически важных и более реалистичных случаях приходится прибегать к методам последовательного поиска решения. Так, например, рассмотрим СДУ более общего вида: $\dot{y}(t)+\mu y(t) = \sqrt{\sum_{v=0}^V g_v t^v}\zeta(t)$, $y(t_1) = y_1$ – начальное значение координаты, $\mu > 0$ – известные параметры, $g_0, g_1 > 0$ – неизвестный параметр. Такого рода СДУ могут быть использованы для описания процессов с медленными нестационарными изменениями, процессов установления, процессов ухода из контрольной зоны, процессов, описывающих метрологические характеристики аппаратно-технических средств в теории надежности. Оптимальная разностная схема, эквивалентная приведенному СДУ: $y_{i+1} = (1-\Delta\mu)y_i + \Delta\sqrt{\sum_{v=0}^V g_v t_i^v}\zeta_i$. Математическое ожидание и дисперсия ОПП равны $M_i = (1-\Delta\mu)y_i$, $D_i = \sum_{v=0}^V d_v t_i^v$, $d_v = g_v N\Delta/2$, $v = \overline{0,V}$. Тогда на основании соотношений (10) получаем систему уравнений для точки $d = d^*$:

$$\frac{\partial W_k}{\partial d_v} = -\frac{1}{k}\sum_{i=1}^k \frac{\partial \ln\pi_{i+1,i}}{\partial d_v}\Bigg|_{d=d^*} = -\frac{1}{k}\sum_{i=1}^k t_i^v D_i^{-1}\left[(y_{i+1}-ry_i)^2 D_i^{-1}-1\right]_{d=d^*} = 0,\ \Delta \ll 1,\ v = \overline{0,V}.$$

Для решения этого уравнения можно воспользоваться известными методами стохастической аппроксимации, структура которых, в общем случае, имеет вид

$$d_{i+1}^* = d_{i+1}^0 - \mathbf{K}_{i+1}\frac{\partial \ln\pi_{i+1,i}}{\partial d_v}\Bigg|_{d=d^*},\ i = \overline{1,k-1} \tag{18}$$

где d_{i+1}^0 – найденное одним из известных способов [8,10] предварительное значение оценки; $\mathbf{K}_{i+1}$ - матрица, определяющая конкретный вид стохастической аппроксимации [8]; начальные условия для уравнения (18) d_1^0, $\mathbf{K}_1$ выбираются исходя из имеющейся в наличии априорной информации. Рекуррентные оценки максимального правдоподобия (18) при общих условиях регулярности состоятельны, асимптотически эффективны и асимптотически нормальны [9]. Уравнения (18) образуют замкнутую систему оценки параметров d_v^*.

Моделирование алгоритмов оценки прогнозируемости стохастического процесса

Исследуем оценку прогнозируемости стохастического процесса, описанного выше: $\dot{y}(t)+\mu y(t) = \sqrt{gt}\zeta(t)$. Для оценки параметра масштаба d_k^* удобно воспользоваться последовательным алгоритмом экспоненциального сглаживания, эквивалентным (11), (12):

$$d_{i+1}^* = d_i^* + \frac{1}{i+1}\left(\frac{(y_{i+1}-ry_i)^2}{t_i} - d_i^*\right),\ i = \overline{1,k-1}. \tag{19}$$

На рис. 1 представлены результаты оценки функции прогнозируемости математического ожидания $IP_M^*(t_{k+m})$ (формула (15); $\Delta = 0{,}1$; $\mu = 1$) в сопоставлении с функцией $IP_M(t_{k+m})$ (при известном d = 1) для двух значений количества шагов прогноза m = 100 (кривые 1, 3) и m = 500 (кривые 2, 4) при изменении объема накопленных данных по которым производилась оценка d_k^* ($k=\overline{1,2000}$). Кривые 3 и 4 получены с использованием алгоритма оценки (19). На рисунке 2 приведены зависимости прогнозируемости стохастического процесса в целом, полученные по формуле (17). Кривая 1 соответствует величине накопления k = 50, а кривая 2 – k = 100. Изменение прогнозируемости показано на интервале прогнозных значений индекса $m=\overline{50,400}$.

Полученные в ходе моделирования результаты физически обоснованы. Во-первых, при одинаковых объемах накопления данных о стохастическом процессе (величина k) с увеличением индекса m ИП математического ожидания уменьшается (рис. 1). Во-вторых, оценка ИП, полученная на основе приведенных алгоритмов, адекватна теоретической (рис. 1). В-третьих, с увеличением интервала наблюдения (величина k) возрастает и величина ИП (рис. 2).

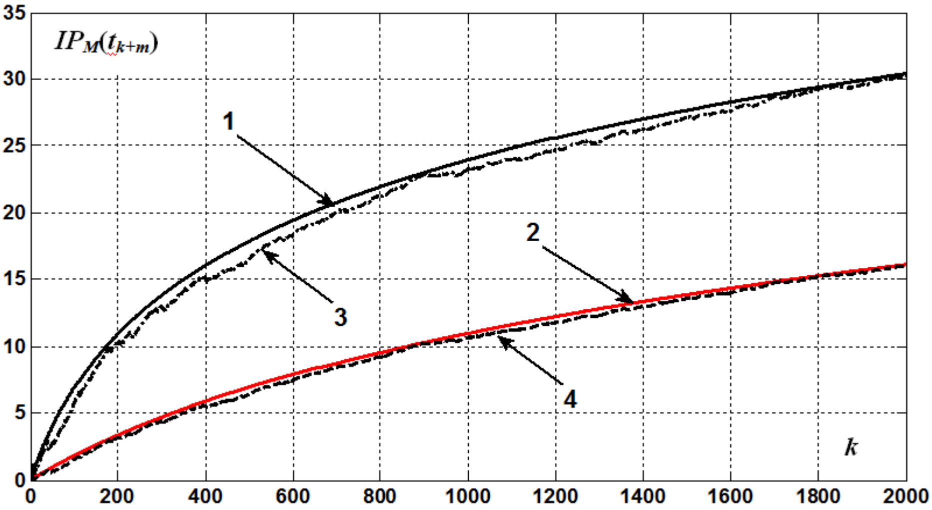

Рис.1. Информационная прогнозируемость математического ожидания: 1 – $IP_M(t_{k+100})$; 2 – $IP_M(t_{k+500})$; 3 – $IP_M^*(t_{k+100})$; 4 – $IP_M^*(t_{k+500})$.

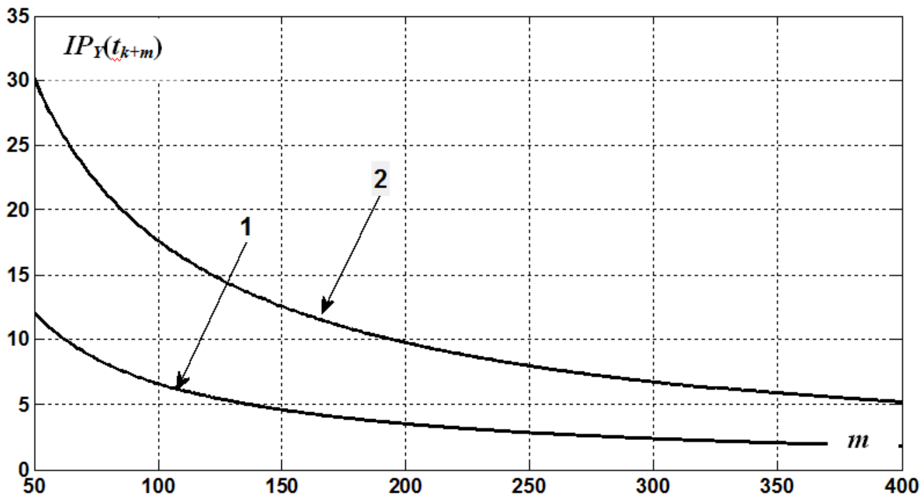

Рис.2. Информационная прогнозируемость процесса: 1 – $IP_Y(t_{50+m})$; 2 – $IP_Y(t_{100+m})$

Заключение

Рассмотрены вопросы параметрической оценки информационной прогнозируемости стохастических процессов. Показано, что эта оценка связана с оценкой параметров ВМП, формируемой по результатам наблюдений. Выбор вида ОПП, как компонента ВМП, может быть осуществлен из практических соображений простоты вычислений и адекватности представления реальной ситуации. ОПП являются удобными аналитическими моделями с физически понятными при их использовании результатами.

С теоретической точки зрения оценка информационной прогнозируемости интересна и как классификационный параметр стохастического процесса, изменяющийся во времени, и как параметр, функционально связанный с самим прогнозируемым процессом. Практическая значимость этой оценки, осуществляемой в реальном времени, состоит в том, что она может способствовать выбору наиболее адекватного метода прогнозирования или адаптивной его настройке.

ASSESSMENT OF INFORMATION PREDICTABILITY OF STOCHASTIC PROCESSES

A.V. AUSIANNIKAU

Abstract

The necessary theoretical information for parameter estimation algorithms informational predictability of stochastic processes are given. We consider examples of algorithms for estimating the predictability of information for the processes described by the optimal difference schemes, equivalent prognostic models in the form of stochastic differential equations are considered.

Список литературы

1. *Овсянников А.В.* // Докл. БГУИР. 2014, № 6 (84). С. 48–54.
2. *Овсянников А.В., Козел В.М.* // Докл. БГУИР. 2014. №. 8 (86). С. 48–53.
3. *Лукашин Ю.П.* Адаптивные методы краткосрочного прогнозирования временных рядов. М., 2003.
4. *Давнис В.В., Тинякова В.И.* Адаптивные модели: анализ и прогноз в экономических системах. Воронеж, 2006.
5. *Орлов Ю.Н., Осминин К.П.* Нестационарные временные ряды: Методы прогнозирования с примерами анализа финансовых и сырьевых рынков. М., 2011.
6. *Орлов Ю.Н., Осминин К.П.* // Информационные технологии и вычислительные системы. 2008. № 3. С. 3–13.
7. *Никитин Н.Н., Разевиг В.Д.* // Журнал вычислит. матем. и математич. физики. 1978. Т 18, № 1. С.106–117.
8. *Цыпкин Я.З.* Информационная теория идентификации. М., 1995.
9. *Ибрагимов И.А., Хасьминский Р.З.* Асимптотическая теория оценивания. М., 1979.
10. *Тартаковский Г.П.* Теория информационных систем. М., 2005.