

СТРУКТУРНО-ПАРАМЕТРИЧЕСКАЯ ИДЕНТИФИКАЦИЯ СТОХАСТИЧЕСКОЙ МОДЕЛИ ИЗМЕРЕНИЙ НА ЖОРДАНОВОЙ ТРАЕКТОРИИ

Аннотация. Рассмотрены особенности структурно-параметрической идентификации стохастической модели измерений на жордановой траектории. Получены формулы для расчета оценок максимального правдоподобия. Соответствующие выводы подтверждены результатами вычислительных экспериментов.

Ключевые слова: модель, авторегрессия, Жорданова кривая, идентификация.

Постановка задачи

Проектирование информационно-измерительных технологий неразрушающего контроля качества и мониторинга состояния технических объектов требует знание моделей измерений информативных параметров [1]. Как правило, при неразрушающем контроле, измерение таких параметров выполняется с одинаковым шагом вдоль некоторой траектории по поверхности изделий. В этом случае измерения информативного параметра z можно представить в виде ряда

$$z = \{z_i\} = z_1, z_2, \dots, z_i, \dots, z_n; \quad z \in \mathbb{R}^n \quad (1)$$

где i - номер точки на траектории, $i \in [1, n]$; n - количество точек измерений; z - вектор измерений информативного параметра. Рассмотрим случай, когда в каждой точке i траектории, информативный параметр z случайным образом принимает то или иное значение, что связано с неоднородностью объекта, а дисперсия случайных ошибок измерений незначительна, и ей можно пренебречь. В этом случае z , как правило, автокоррелированная последовательность случайных величин и для её формального описания могут использоваться известные стохастические модели временных рядов, за исключением того, что значения ряда зависят не от времени t , а от точки измерения i .

Особый случай представляют последовательности, полученные путем измерения информативного параметра вдоль жордановой траектории, например, окружности, когда за последней точкой измерений, с номером $i = n$, следует первая $i = 1$. Это приводит к тому, между членами последовательности z , расположенными на концах, имеется корреляционная связь. Данное свойство замкнутости, которое не учитывают известные модели временных рядов, вносит свои особенности, как в алгоритмы формирования последовательностей измерений, так и в алгоритмы структурно-параметрической идентификации их моделей.

Модель замкнутых автокоррелированных последовательностей

Для формального описания последовательности z можно использовать модель случайного процесса, учитывающего свойства замыкания. Такая модель (2), полученная на основе авторегрессионной модели m -го порядка, подробно рассмотрена в [2] и [3].

$$z_i = \begin{cases} \sum_{k=1}^m \alpha_k z_{n-k+1} + \xi_1, & i = 1, \\ \sum_{k=1}^{i-1} \alpha_k z_{i-k} + \sum_{k=i}^m \alpha_k z_{n-k+i} + \xi_i, & 1 < i \leq m, \\ \sum_{k=1}^m \alpha_k z_{i-k} + \xi_i, & m < i \leq n; \end{cases} \quad (2)$$

где $\xi = \{\xi_i\}$; $\xi \in \mathbb{R}^n$ - стационарная последовательность независимых случайных величин; $\alpha = \{\alpha_1, \dots, \alpha_m\}$, $\alpha \in \mathbb{R}^m$ - регулируемые параметры модели, m - порядок.

Подстановкой $m=1$ в (2) получим модель, прообразом которой является случайный Марковский процесс для дискретного времени, а при $m=2$ получим Юловскую модель:

$$z_i = \begin{cases} \alpha_1 z_n + \xi_1, & i = 1, \\ \alpha_1 z_{i-1} + \xi_i, & i > 1; \end{cases} \quad m = 1.$$

$$z_i = \begin{cases} \alpha_1 z_n + \alpha_2 z_{n-1} + \xi_1, & i = 1, \\ \alpha_1 z_1 + \alpha_2 z_n + \xi_2, & i = 2, \\ z_i = \alpha_1 z_{i-1} + \alpha_2 z_{i-2} + \xi_i, & 2 < i \leq n; \end{cases} \quad m = 2.$$

Основные свойства замкнутых автокоррелированных последовательностей

Математическое ожидание всех членов последовательности равно

$$M[z_i] = m_\xi \left[1 - \sum_{k=1}^m \alpha_k \right]^{-1}, \quad (3)$$

где $M\{\xi_i\} = m_\xi$ - математическое ожидание последовательности $\{\xi_i\}$.

Если обозначить $\omega = \exp(2\pi i / n)$ - корень степени n из единицы, где $i = \sqrt{-1}$, то ковариационная функция может быть представлена в виде

$$R(k) = \sigma_\xi^2 \frac{1}{n} \sum_{i=1}^n \left[\omega^{ki} \left| 1 - \sum_{h=1}^m \alpha_h \omega^{hi} \right|^{-2} \right], \quad k \in [-n, n], k \in \mathbb{Z}. \quad (4)$$

где $D\{\xi_i\} = \sigma_\xi^2$ - дисперсия последовательности ξ . В [3] показано, что $R(k)$ - всегда функция действительных значений, а также что $R(k) = R(-k) = R(n - k)$, таким образом ковариационная функция симметричная относительно точки $k = n / 2$.

Дисперсия последовательности может быть определена как $D[z_i] = R(0)$, а корреляционная функция как $r(k) = R(k) / R(0)$.

Для модели первого порядка ($m=1$), выше приведенные формулы могут быть представлены в компактном виде

$$M[z_i] = \frac{m_\xi}{1 - \alpha_1}, \quad D[z_i] = \sigma_\xi^2 \frac{\alpha_1^n + 1}{(1 - \alpha_1^n)(1 - \alpha_1^2)}; \quad (5)$$

$$R(k) = \sigma_\xi^2 \frac{\alpha_1^{|k|} + \alpha_1^{n-|k|}}{(1 - \alpha_1^n)(1 - \alpha_1^2)}, \quad r(k) = \frac{\alpha_1^{|k|} + \alpha_1^{n-|k|}}{1 + \alpha_1^n}. \quad (6)$$

Последовательность z можно выразить как взвешенную сумму независимых случайных величин ξ только при условии, что ни один из корней характеристического уравнения $P_m(x) = 1 - \alpha_1 x - \alpha_2 x^2 - \dots - \alpha_m x^m$ не лежит на единичной окружности. Замкнутые последовательности одинаковой длины n отвечают требованию слабой стационарности независимо от того, лежат ли корни $P_m(x)$ внутри единичной окружности или за её пределами. Это приводит к тому, что единственная ковариационная структура последо-

вательности z , может быть представлена множеством линейных моделей вида (2), как это показано далее.

Положим, что вектор z имеет обратную ковариационную матрицу R_z^{-1} и может быть представлен моделью (2), которую, используя [3], запишем в виде

$$\prod_{k=1}^m (I - x_k A) z = \xi,$$

где A - ортогональная матрица оператора сдвига назад ($A^T A = I$); I - единичная матрица; $x_1^{-1}, \dots, x_k^{-1}, \dots, x_m^{-1}$ - корни $P_m(x)$, $|x_k^{-1}| \neq 1, \forall k$.

Из [3] также получим формулу для R_z^{-1} и представим ее следующим образом:

$$R_z^{-1} = \sigma_\xi^{-2} P_m(A^T) P_m(A) = \sigma_\xi^{-2} \prod_{k=1}^m (I - x_k A^T)(I - x_k A). \quad (7)$$

Для сомножителей в (7) всегда выполняется равенство

$$\begin{aligned} (I - x_k A^T)(I - x_k A) &= x_k^2 (I - x_k^{-1} A^T)(I - x_k^{-1} A), \forall x_k \in \mathbb{R}, \\ (I - x_k A^T)(I - \overline{x_k} A) &= x_k \overline{x_k} (I - x_k^{-1} A^T)(I - \overline{x_k}^{-1} A), \forall x_k \in \mathbb{Z}, \end{aligned} \quad (8)$$

где предполагается, что для комплексных корней взяты обратные значения каждого из членов сопряженной пары $\{x_k, \overline{x_k}\}$. При этом, если, например, действительный корень x_k^{-1} или комплексно-сопряженная пара $\{x_k^{-1}, \overline{x_k}^{-1}\}$ лежат вне единичного круга, то их обратные значения - внутри, и наоборот.

Из (8) следует, что любая из стохастических моделей

$$\prod_{k=1}^m (I - x_k^{\pm 1} A) z = c \xi,$$

будет иметь ту же ковариационную матрицу R_z при надлежащем выборе константы c . В этой связи возникает неоднозначность в выборе параметров модели.

Назначение стохастической модели (2) - адекватное формальное описание вероятностной структуры измерений объекта. Для фиксированного порядка модели такое представление должно быть единственным. Если ξ - вектор нормальных случайных величин, то единственность параметров модели фиксированного порядка m подразуме-

вает, что матрица R_z должна однозначно соответствовать единственно возможной последовательности коэффициентов $\{a_1, \dots, a_m\}$ характеристического уравнения и, следовательно, его корней $x_1^{-1}, \dots, x_m^{-1}$. Из равенства (8) выходит, что такая единственность может быть достигнута путем ввода ограничения касательно их расположения относительно единичного круга на комплексной плоскости. Предлагается выбрать вариант, когда все корни лежат вне единичного круга. Это, во-первых, упростит процедуру идентификации и оценивания модели, как будет показано далее, во-вторых, алгоритм формирования замкнутых последовательностей будет иметь меньшую вычислительную сложность [3]. При таком выборе, из (8) также доказывается, что значение дисперсии σ_ξ^2 будет минимально возможным, при котором существует матрица R_z .

Оценка параметров модели

Пусть случайный вектор $z = \{z_1, \dots, z_i, \dots, z_n\}$ принял определенное значение и по данной реализации необходимо оценить параметры модели (2), при условии, что компоненты вектора $\xi = \{\xi_1, \dots, \xi_i, \dots, \xi_n\}$ имеют нормальный закон распределения: $\xi_i \sim N(m_\xi, \sigma_\xi^2), \forall i$, а все корни характеристического уравнения лежат вне единичного круга.

Можно показать, что среднее значение $\bar{\xi}$ независимых случайных величин ξ связано со средним значением \bar{z} последовательности z , соотношением

$$\bar{\xi} = \bar{z} \left[1 - \sum_{k=1}^m \alpha_k \right]. \quad (9)$$

Из (9) следует, что α можно оценить по центрированному ряду $\{z_i - \bar{z}\}$, полагая, что $m_\xi \approx \bar{\xi} = 0$. Далее предположим, что последовательность z центрирована.

Если для последовательности z известен её вектор параметров α , то по формуле (2) можно последовательно вычислить всё множество значений $\xi_i(\alpha / z), \forall i$, поскольку, согласно условию, первые элементы последовательности зависят от последних. При этом, так

как компоненты $\{\xi_i\}$ распределены нормально, то плотность распределения вероятностей случайного вектора ξ равна

$$W(\xi) = (2\pi)^{-\frac{n}{2}} \sigma_\xi^{-n} \exp\left(-\frac{1}{2\sigma_\xi^2} \sum_{i=1}^n \xi_i^2\right). \quad (10)$$

Следовательно, условная логарифмическая функция правдоподобия для параметров (α, σ_ξ) , найденная из (10), примет вид

$$l(\alpha, \sigma_\xi / z) = -n \ln \sigma_\xi - \frac{n}{2} \ln 2\pi - \frac{1}{2\sigma_\xi^2} S(\alpha / z), \quad (11)$$

где $S(\alpha / z) = \sum_{i=1}^n \xi_i^2(\alpha / z)$. С учетом (2), получим, что

$$S(\alpha / z) = \left(z_1 - \sum_{k=1}^m \alpha_k z_{n-k+1}\right)^2 + \sum_{i=2}^m \left(z_i - \sum_{k=1}^{i-1} \alpha_k z_{i-k} - \sum_{k=i}^m \alpha_k z_{n-k+i}\right)^2 + \sum_{i=m+1}^n \left(z_i - \sum_{k=1}^m \alpha_k z_{i-k}\right)^2, \quad (12)$$

где предполагается, что при $m=1$ вторая сумма не участвует в расчете.

Путем максимизации функции (11) могут быть получены *оценки условного максимального правдоподобия (СМЛ-оценки)*. Как указано в [4], для их нахождения, в общем случае, нужно использовать численные методы, а не решать систему уравнений в частных производных, приравненных к нулю. Это связано с тем, что информация, содержащаяся в функции правдоподобия, выражается только такими оценками и вторыми её производными, когда в интересующей области функции применима квадратичная аппроксимация [4]. Необходимые для численных расчетов начальные приближения оценок можно получить прямым дифференцированием функции $l(\alpha_1, \sigma_\xi / z)$. Например, для модели первого порядка, решив систему уравнений в частных производных, относительно неизвестных α_1 и σ_ξ , получим выражения для таких приближений оценок:

$$\alpha_1^* = r_1^*, \quad \sigma_\xi^* = \sigma_z^* \sqrt{1 - (r_1^*)^2}, \quad (13)$$

где $r_1^* = \left(z_1 z_n + \sum_{i=2}^n z_i z_{i-1}\right) / \sum_{i=1}^n z_i^2$, $\sigma_z^{*2} = \frac{1}{n} \sum_{i=1}^n z_i^2$ - значение выборочной дисперсии z .

Для второго порядка ($m=2$), аналогичным образом, получены оценки

$$\alpha_1^* = \frac{r_1^* (1 - r_2^*)}{1 - r_1^{*2}}, \quad \alpha_2^* = \frac{r_2^* - r_1^{*2}}{1 - r_1^{*2}}, \quad \sigma_\xi^* = \sigma_z^* \sqrt{(1 - \alpha_1^* r_1^* - \alpha_2^* r_2^*)}, \quad (14)$$

где $r_2^* = \left(z_1 z_{n-1} + z_2 z_n + \sum_{i=3}^n z_i z_{i-2} \right) / \sum_{i=1}^n z_i^2$.

Вычисления оценок $\hat{\alpha} = \{\alpha_1^*, \dots, \alpha_m^*\}$ для моделей более высоких порядков показывают, что они также, как для первого и второго порядков, совпадают с оценками, найденными из уравнений Юла-Уокера, при условии, что выборочные коэффициенты автокорреляции вычисляются по формуле

$$r_k^* = \frac{\sum_{i=1}^{n-k} z_i z_{i+k} + \sum_{i=1}^k z_{n-k+i} z_i}{\sum_{i=1}^n z_i^2}, \quad k \in (1, n), \quad (15)$$

где k - расстояние между членами последовательности на жордановой траектории (задержка). В этом случае вектор $\hat{\alpha}$ может быть определен по формуле

$$\hat{\alpha} = \hat{P}_m^{-1} \hat{r}_m, \quad (16)$$

где

$$\hat{r}_m = \begin{pmatrix} r_1^* \\ r_2^* \\ \vdots \\ r_m^* \end{pmatrix}, \quad \hat{P}_m = \begin{pmatrix} 1 & r_1^* & r_2^* & \cdots & r_{m-1}^* \\ r_1^* & 1 & r_1^* & \cdots & r_{m-2}^* \\ \vdots & \vdots & \vdots & & \vdots \\ r_{m-1}^* & r_{m-2}^* & r_{m-3}^* & \cdots & 1 \end{pmatrix}. \quad (17)$$

Оценка σ_ξ^* может быть вычислена следующим образом:

$$\sigma_\xi^* = \sigma_z^* \sqrt{\left(1 - \sum_{k=1}^m \alpha_k^* r_k^* \right)}. \quad (18)$$

Полученные таким образом оценки Юла-Уокера могут быть выбраны в качестве начальных приближений для поиска более точных СМЛ-оценок, только если корни $P_m(x)$ лежат вне единичного круга, как было определено выше.

На рис.1. показаны гистограммы СМЛ-оценок и оценок Юла-Уокера полученных в результате оценивания 100 000 последовательностей первого порядка, размером $n=25$ и $n=90$ и с параметрами:

$\alpha_1 = 0.5$, $\sigma_\xi = 1$, $m_\xi = 0$. Максимум функции (11) вычислялся методом сопряженных градиентов.

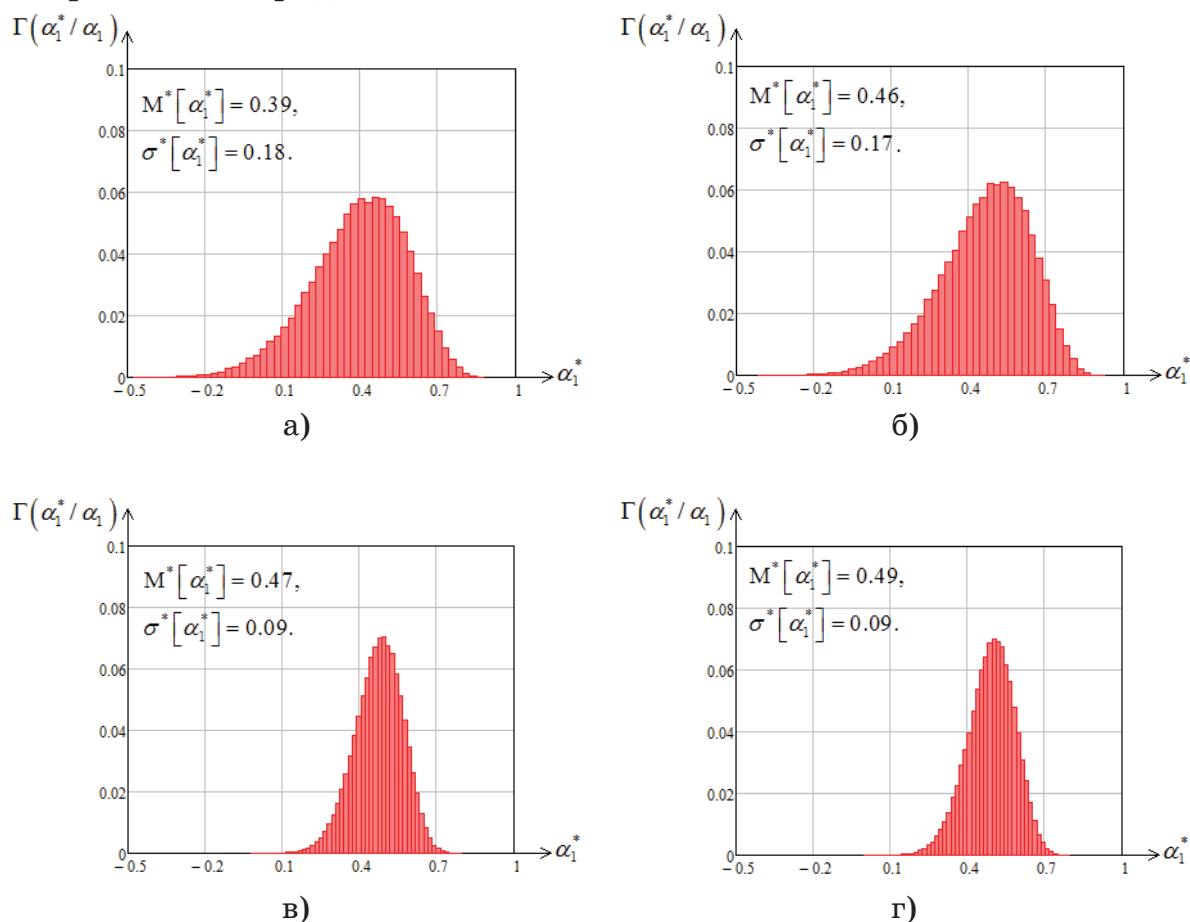


Рисунок 1 - Гистограммы оценок α_1^* :

а) Юла-Уокера, $n=25$; б) СМЛ, $n=25$; в) Юла-Уокера, $n=90$; г) СМЛ, $n=90$;

На рис.1 также указаны средние значения и выборочные значения среднеквадратического отклонения оценок. Как видно, более точными и эффективными являются оценки максимума условного правдоподобия (СМЛ-оценки), однако для всех полученных оценок замечено смещение в сторону занижения по модулю. Как было обнаружено по результатам вычислительных экспериментов, с ростом размера оцениваемой последовательности данный эффект уменьшается, а значения оценок Юла-Уокера и СМЛ-оценок, практически идентичны.

Следует заметить, что для поиска СМЛ-оценок численными методами, помимо начальных приближений, необходимо указать ограничения для допустимых значений оценок $\{\alpha_1^*, \dots, \alpha_m^*\}$ исходя из по-

ложений корней характеристического уравнения за пределами единичного круга на комплексной плоскости.

Из (15) следует, что $r^*(k) = r_k^* = r_{n-k}^*$, $0 < k < n$, таким образом, коррелограмма, также как и теоретическая функция, симметричная относительно точки $n/2$.

В условное правдоподобие $l(\alpha, \sigma_\xi / z)$ данные входят только через условную сумму квадратов (12). Отсюда следует, что оценки максимума условного правдоподобия (СМЛ-оценки) те же, что и оценки наименьших квадратов.

Также могут быть получены *оценки максимального правдоподобия (МЛ-оценки)*. Плотность вероятности случайного вектора z равна

$$W(z) = (2\pi)^{-\frac{n}{2}} (\det R_z)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} z^T R_z^{-1} z\right). \quad (19)$$

Из формулы (19) определим безусловную логарифмическую функцию правдоподобия, которая примет вид

$$L(\alpha, \sigma_\xi / z) = l(\alpha, \sigma_\xi / z) + \ln(|\det B|), \quad (20)$$

где $\det B = \prod_{j=1}^n P_m(\omega^j)$ - определитель матрицы линейного оператора перехода от вектора z к вектору ξ , а $|\det B|$ - его модуль.

Для вычисления МЛ-оценок также необходимо использовать численные методы.

Выразим $\det B$ через корни $x_1^{-1}, \dots, x_m^{-1}$ характеристического уравнения:

$$\det B = \prod_{j=1}^n P_m(\omega^j) = \prod_{k=1}^m \prod_{j=1}^n (1 - x_k \omega^j) = \prod_{k=1}^m \left(1 - x_k \sum_{j=1}^n \omega^j + x_k^2 \sum_{\substack{j_1=1, \\ j_2=1}}^n \omega^{j_1} \omega^{j_2} - \dots + (-x_k)^n \prod_{j=1}^n \omega^j \right), \quad (21)$$

где $w_j = \omega^j$. В каждой скобке выражения (21) все слагаемые, кроме первого и последнего, равны нулю, следовательно, определитель равен

$$\det B = \prod_{k=1}^m (1 - x_k^n). \quad (22)$$

Так как, согласно введенному выше ограничению, $|x_k^{-1}| > 1$, то из (22) следует, что $\det B \approx 1$ для некоторого небольшого n . В этом

случае СМЛ-оценки почти те же, что и МЛ-оценки максимального правдоподобия и членом $\ln(|\det B|)$ в (20) можно пренебречь. Вычислительные эксперименты на 100 000 последовательностях малых размеров ($n = 10$) показывают, что выборочная корреляция между МЛ- и СМЛ-оценками, составляет 0,99.

Ковариационная матрица МЛ-оценок $\hat{\alpha}$ приближенно равна [4]

$$R_{\hat{\alpha}} \approx n^{-1} (1 - r_m^T \alpha) P_m^{-1}, \quad (23)$$

где r_m и P_m - вектор и матрица теоретических корреляций, с формами, аналогичными (17). Оценки ковариаций можно получить, заменив в (23) неизвестные параметры их МЛ-оценками и выборочными корреляциями (17). Диагональные элементы матрицы $R_{\hat{\alpha}}$ - дисперсии соответствующих МЛ-оценок.

Структурная идентификация

Структура рассматриваемой модели определяется её порядком m . Для предварительной его оценки можно использовать частную автокорреляционную функцию, периодограмму, кумулятивную периодограмму и другие характеристики, применяемые для анализа временных рядов [4].

Значения частной автокорреляционной функции должны рассчитываться с учетом формул (15) и (16).

В отношении периодограммы была замечена особенность, что при определенных значениях параметров $\{\alpha_1, \dots, \alpha_m\}$, на ней могут стабильно присутствовать значительные низкочастотные пики, как, например, на рис.2б. Это приводит к тому, что по анализу графиков таких последовательностей, в полярной системе координат, как, например, на рис.2а, можно ошибочно предположить наличие в последовательности детерминированной составляющей, напоминающей эллипс или другую гладкую кривую, в зависимости от вклада соответствующих гармоник.

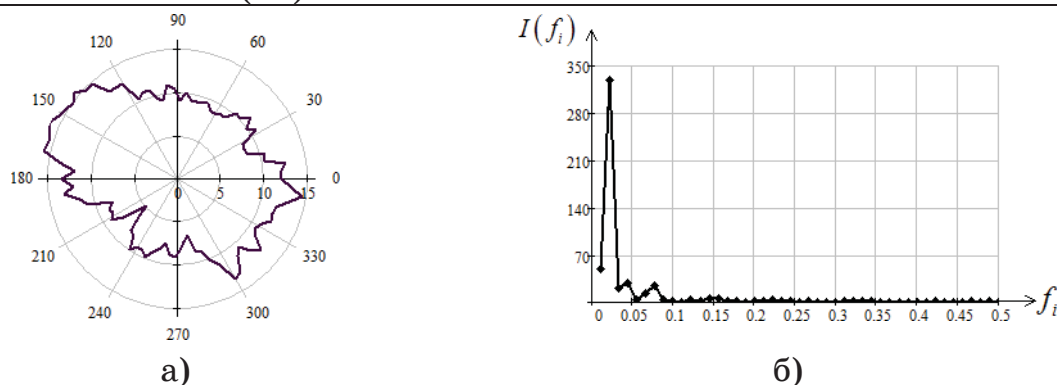


Рисунок 2 - График последовательности с параметрами $n = 90$, $m_\xi = 1$, $\sigma_\xi = 1$, $\alpha = 0.9$ (а), и её периодограмма (б), где $I(f_i)$ - интенсивность на частоте $f_i = i / n$

После предварительного анализа ряда, для предполагаемого порядка модели m , необходимо оценить вектор параметров $\hat{\alpha}$ и вычислить остаточные ошибки $\hat{\xi} = \{\xi_i^*\}$, выразив их из формулы (2), которую можно записать в матричном виде, как в [3]. Последние нужно исследовать на наличие автокорреляции, например, по Q-тесту Бокса-Пирса. Наличие в остатках периодического отклонения от случайности можно выявить по виду кумулятивной периодограммы [4].

Объективное решение о величине порядка модели можно принять по показателю внешнего информационного критерия, например, AIC:

$$AIC = 2(m + 1) - 2L(\hat{\alpha}, \sigma_\xi^* / z),$$

где $L(\hat{\alpha}, \sigma_\xi^* / z)$ - значение максимума логарифмической функции правдоподобия (20); $\hat{\alpha}, \sigma_\xi^*$ - найденные ML- или SML-оценки. Значение m , при котором показатель AIC минимальный, принимается за порядок модели. Аналогичным образом могут использоваться и другие информационные критерии, например BIC:

$$BIC = 2n \ln(\sigma_\xi^*) + (m + 1) \ln(n).$$

Проверить значимо ли отличаются оценки $\hat{\alpha}$ от нуля, можно вычислив их стандартные отклонения по формуле (23).

Практическую значимость модели можно оценить по показателю

$$R_m^2 = 1 - \frac{\sigma_\xi^{*2}}{\sigma_z^{*2}}.$$

Дифференциальная энтропия

Энтропия случайного вектора z , с учетом того, что $\det B \approx 1$, равна

$$H(z) = H(B^{-1}\xi) = H(\xi) - \ln(\det B) \approx H(\xi).$$

При условии нормального закона распределения вектора ξ , получим формулу

$$H(z) \approx n \ln \sigma_\xi + \frac{n}{2} (1 + \ln 2\pi).$$

Таким образом, оценка энтропии последовательности z пропорциональна дисперсии остатков и числу измерений.

Выводы

1. Рассмотрены особенности идентификации стохастической модели измерений на жордановой траектории, когда за последней точкой измерений, с номером $i = n$, следует первая $i = 1$. Данное свойство, в сравнении с моделями временных рядов, обеспечивает отсутствие "концевых эффектов" при оценивании модели.

2. Для параметров модели получены формулы для расчета оценок Юла-Уокера, максимального правдоподобия и оценок максимума функции условного правдоподобия. Даны рекомендации относительно их использования.

ЛИТЕРАТУРА

1. Малайчук В.П. Математическая дефектоскопия: Монография / В.П. Малайчук, А.В. Мозговой.-Д.:Системные технологии, 2005. -180 с.
2. Кошулян А.В. Моделирование автокоррелированных последовательностей измерений вдоль жордановой траектории / А.В. Кошулян, В.П. Малайчук -Д.: Системные технологии. Региональный межвузовский сборник научных трудов. - Выпуск 1(8), 2012, с. 9-17.
3. Кошулян А.В. Стохастическая модель автокоррелированных измерений на жордановых траекториях - Д.: Системные технологии. Региональный межвузовский сборник научных трудов. - Выпуск 2(79), 2012, с. 162-173.
4. Бокс Дж., Дженкинс Г. Анализ временных рядов, прогноз и управление: Пер. с англ.//Под ред. В.Ф. Писаренко. – М.: Мир, 1974, кн. 1. – 406 с.