

ИЗМЕРИТЕЛЬ-ЭКСТРАПОЛЯТОР ВРЕМЕННЫХ ПРОЦЕССОВ

В. И. Долгов, Г. П. Азбукин

Харьков

Задача измерения процесса с предсказанием будущего его значения характерна для многих практических ситуаций, возникающих в автоматических измерителях, как автономных, так и входящих в системы управления и контроля различного назначения. Суть ее заключается в том, что значение процесса λ_n в момент времени t_n необходимо предсказать по выборке $\bar{\lambda} = \{\lambda_1 \dots \lambda_{n-1}\}$, представляющей совокупность известных значений процесса в предыдущие моменты времени t_1, \dots, t_{n-1} , либо по выборке $\bar{u} = \{u_1, \dots, u_{n-1}\}$ совокупности известных значений другого процесса, статистически и функционально связанного с экстраполируемым. Как правило, эту задачу в указанных системах приходится решать в реальном масштабе времени, т. е. в ритме поступления наблюдаемых данных, а это значит, что для решения задачи экстраполяции отводится ограниченное время, определяемое интервалом дискретности системы. Отмеченное обстоятельство вынуждает к поиску и исследованию алгоритмов экстраполяции, которые бы обеспечивали высокую точность экстраполяции и вместе с тем допускали бы простую техническую реализацию, позволяющую провести все необходимые вычисления к очередному такту работы системы. Следует отметить, что теория экстраполяции развивается успешно. Хорошо разработана теория линейной экстраполяции стационарных [1, 2] и нестационарных случайных процессов [3, 4]. Определенные результаты получены и в теории нелинейной экстраполяции [5].

Современные вычислительные машины позволяют реализовать многие из алгоритмов, полученные в указанных выше работах, однако подход, развитый в них, оказывается неудовлетворительным: решение задач отыскивается в классе наперед введенных операторов (линейных или нелинейных), а наиболее употребительной мерой точности экстраполяции является среднеквадратическая ошибка.

Самая последовательная в постановке задачи и близкая по смыслу к рассматриваемой проблеме в настоящее время теория статистических решений, которая позволяет получить общее решение задачи оптимальной экстраполяции для произвольного критерия оптимальности при наличии выборок любого объема. На основании этой теории получены эффективные алгоритмы оптимальной фильтрации процессов [6, 7].

Главная трудность использования оптимальных методов в практике экстраполяции — сложность большинства оптимальных процедур, связанная с необходимостью запоминания и обработки достаточно большого (в пределе бесконечного) объема накопленных к текущему моменту времени ранее произведенных решений или ранее наблюдаемых данных. Ее эффективно устраняют путем поиска и исследования рекуррентных алгоритмов экстраполяции, позволяющих реализовать учет ранее приобретенных сведений о прогнозируемом процессе и требующих проведения конечного числа вычислительных операций при ограниченном объеме запоминаемых данных независимо от длительности интервала наблюдения.

Настоящая статья посвящена отысканию рекуррентных алгоритмов оптимальной экстраполяции нормальных процессов на основании теории статистических решений.

Постановка и общее решение задачи

Обработке подлежит последовательность выборочных значений процесса $\bar{u} = \{u_1, \dots, u_{n-1}\}$ в моменты времени t_1, \dots, t_{n-1} из интервала наблюдения $[(0, T), t_{n-1} = (n-1)T_0]$, статистические свойства которого описываются заданием условной плотности вероятностей $P(\bar{u} / \bar{\lambda}_{n-1}, \bar{\alpha}_{n-1})$ совокупности существенных с точки зрения измерений параметров $\bar{\lambda}_{n-1} = \{\lambda_1, \dots, \lambda_{n-1}\}$ из этого же интервала и в общем случае совокупности дополнительных мешающих (несущественных) параметров $\bar{\alpha}_{n-1} = \{\alpha_1, \dots, \alpha_{n-1}\}$, не подлежащих измерению или оценке в устройстве обработки результатов наблюдений. Для упрощения записи будем считать все переменные, фигурирующие в формулах, скалярными. Задана априорная статистика несущественных параметров $\bar{\alpha}$ — функция распределения $P(\bar{\alpha})$, которая позволяет определить функцию

правдоподобия $P(\bar{u}/\lambda_1, \dots, \lambda_{n-1})$ совокупности существенных параметров $\lambda_1, \dots, \lambda_{n-1}$ на всех периодах повторения сигнала из рассматриваемого интервала наблюдения $(0, T)$. Известна априорная статистика отслеживаемого процесса, т. е. задана функция распределения $P(\lambda_1, \dots, \lambda_n)$ на рассматриваемом интервале $(0, t_n)$.

Требуется по совокупности выборочных значений процесса $\bar{u} = \{u_1, \dots, u_{n-1}\}$ из интервала наблюдения $(0, T)$, статистически связанного с процессом $\bar{\lambda}_{n-1} = \{\lambda_1, \dots, \lambda_{n-1}\}$, предсказать наилучшим образом значение процесса λ_n в момент времени t_n и рассчитать теоретически предельную дисперсию полученной оценки.

Первым этапом решения задачи, как следует из теории статистических решений, является нахождение апостериорной плотности распределения $P(\lambda_n/\bar{u}_{n-1})$ для интересующего нас значения параметра в момент времени t_n по выборочным значениям процесса из интервала наблюдения $(0, T)$.

На основании формулы Байеса легко получить

$$P(\lambda_n/\bar{u}_{n-1}) = \frac{\int P(\lambda_1 \dots \lambda_n) P(\bar{u}_{n-1}/\bar{\lambda}_{n-1}, \bar{a}_{n-1}) P(\bar{a}_{n-1}) \times}{\int P(\bar{\lambda}_{n-1}, \lambda_n) P(\bar{a}_{n-1}) P(\bar{u}_{n-1}/\bar{\lambda}_{n-1}, \bar{a}_{n-1}) \times} \times \frac{d\bar{a}_{n-1} d\bar{\lambda}_{n-1}}{d\bar{a}_{n-1} d\bar{\lambda}_{n-1}}. \quad (1)$$

Эта плотность вероятности содержит наиболее полную информацию о параметре λ_n , учитывающую априорные сведения о процессе $\lambda(t)$, и сведения, накопленные на всем интервале наблюдений $(0, T)$.

На втором этапе в теорию статистических решений вводится критерий качества предсказания, который выражается через функцию потерь $W(\lambda_n, \lambda_{on})$, характеризующую плату за ошибки в решениях, когда в качестве предсказанного значения принимается λ_{on} , в то время как истинное значение процесса в момент t_n есть λ_n .

Наилучшее правило выбора предсказанного значения λ из всего множества значений определяется из условия минимума среднего риска, т. е. значения функции потерь, усредненного по апостериорному распределению (1).

В большинстве практических ситуаций апостериорное распределение (1) оказывается симметричной функцией, а в этом случае для широкого класса симметричных функций потерь минимизация среднего риска приводит к универсальной оптимальной оценке прогнозируемого значения процесса λ_{on} в виде максимального (среднего) значения апостериорного распределения $P(\lambda_n/\bar{u}_{n-1})$.

Дисперсия оценки λ_{on} определяется (n, n) элементом матрицы ковариаций результирующего апостериорного распределения $P(\lambda_1, \dots, \lambda_{n-1}, \lambda_n / \bar{u}_{n-1})$.

Экстраполяция нормальных процессов

Для упрощения будем считать, что несущественные параметры отсутствуют.

При нормальном процессе априорное распределение совокупности значений процесса в моменты времени t_1, \dots, t_n выразим в виде

$$P(\lambda_1, \dots, \lambda_n) = k \exp \left\{ -\frac{1}{2} (\bar{\lambda}^{(n)} - \bar{\lambda}_{cp})^T V (\bar{\lambda}^{(n)} - \bar{\lambda}_{cp}) \right\}. \quad (2)$$

Здесь $\bar{\lambda}^{(n)} = \{\lambda_1 \dots \lambda_n\}$ — вектор-столбец значений процесса в моменты времени t_1, \dots, t_n ; $\bar{\lambda}_{cp} = \{\lambda_{cp1}, \dots, \lambda_{cpn}\}$ — соответствующий вектор-столбец средних значений; $V = R^{-1}$ — матрица, обратная корреляционной матрице R процесса; T — знак транспонирования.

Будем считать, что функция правдоподобия $P(\bar{u}_{n-1} / \bar{\lambda}_{n-1})$ совокупности значений параметра $\lambda_1, \dots, \lambda_{n-1}$ из интервала $(0, T)$ допускает параболическую аппроксимацию, т. е. пусть

$$P(\bar{u}_{n-1} / \bar{\lambda}_{n-1}) = k \exp \left\{ -\frac{1}{2} (\bar{\lambda}^{(n-1)} - \bar{\lambda}^*)^T A (\bar{\lambda}^{(n-1)} - \bar{\lambda}^*) \right\}. \quad (3)$$

Здесь

$$\bar{\lambda}^{(n-1)} = \{\lambda_1, \dots, \lambda_{n-1}\}; \quad \bar{\lambda}^* = \{\lambda_1^*, \dots, \lambda_{n-1}^*\}$$

— вектор-столбец, составленный из оценок максимального правдоподобия; A — матрица с элементами

$$A_{ij} = -\frac{\partial^2}{\partial \lambda_i \partial \lambda_j} \ln P(\bar{u}_{n-1} / \bar{\lambda}_{n-1}^*).$$

Будем искать выражение для оптимальной оценки прогнозируемого значения λ_{on} через оптимальные оценки отфильтрованных значений процесса в предыдущие моменты времени. Заметим, что решение задачи фильтрации процессов на основе теории статистических решений приводит к оптимальным оценкам процесса в виде средних (максимальных) значений соответствующих апостериорных распределений. Учитывая последнее, запишем интересующее нас апостериорное распределение совокупности оцениваемых параметров $\lambda_1, \dots, \lambda_n$ в виде

$$P(\lambda_1, \dots, \lambda_n / \bar{u}_{n-1}) = P(\lambda_1, \dots, \lambda_{n-1} / \bar{u}_{n-1}) P(\lambda_n / \bar{\lambda}_{n-1}). \quad (4)$$

Явные выражения для сомножителей в (4) легко найти, если учесть, что для элементов матрицы, обратной корреляционной матрице нормального процесса, справедливо соотношение

$$V_{ij}^{(n)} = V_{ij}^{(n-1)} + \frac{V_{in}^{(n)}V_{nj}^{(n)}}{V_{nn}^{(n)}}. \quad (5)$$

Здесь индекс сверху обозначает порядок матрицы.

Таким образом, выражение (5) дает связь между элементами матриц, обратных корреляционным, различных порядков.

Используя (5), легко получить вместо (2)

$$P(\lambda_1, \dots, \lambda_n) = P(\lambda_1, \dots, \lambda_{n-1}) P(\lambda_n/\lambda_{n-1}); \quad (6)$$

сомножители выражения (6) можно представить в виде

$$P(\lambda_1, \dots, \lambda_{n-1}) = k \exp \left\{ -\frac{1}{2} (\bar{\lambda}^{(n-1)} - \bar{\lambda}_{cp})^T V^{(n-1)} (\bar{\lambda}^{(n-1)} - \bar{\lambda}_{cp}) \right\}; \quad (7)$$

$$P(\lambda_n/\lambda_{n-1}) = k \exp \left\{ -\frac{1}{2} (\bar{\lambda}^{(n)} - \lambda_{cp}) \frac{\bar{V}\bar{V}^T}{V_{nn}} (\bar{\lambda}^{(n)} - \lambda_{cp}) \right\}, \quad (8)$$

где $\bar{V} = \{V_{1n}^{(n)} \dots V_{n-1n}^{(n)}\}$ — вектор-столбец, составленный из соответствующих элементов матрицы $V^{(n)}$. Но тогда в (4)

$$P(\lambda_1, \dots, \lambda_{n-1}/\bar{u}_{n-1}) = \frac{P(\bar{u}_{n-1}/\bar{\lambda}_{n-1}) P(\lambda_1, \dots, \lambda_{n-1})}{P(\bar{u}_{n-1})} = \quad (9)$$

$$= k \exp \left\{ -\frac{1}{2} (\bar{\lambda}^{(n-1)} - \bar{\lambda}^{**})^T C (\bar{\lambda}^{(n-1)} - \bar{\lambda}^{**}) \right\}.$$

Здесь

$$C^{-1} = A^{(n-1)} + V^{(n-1)} \quad (10)$$

и

$$\bar{\lambda}^{**(n-1)} = C^{(n-1)} A^{(n-1)} (\bar{\lambda}^* - \bar{\lambda}_{cp}) + \bar{\lambda}_{cp}. \quad (11)$$

Выражение (11) характеризует значение оптимальных оценок процесса $\lambda_1, \dots, \lambda_{n-1}$, сформированных в текущий момент времени t_{n-1} — эффект оптимальной фильтрации, а выражение (10) служит для определения элементов матриц ковариаций результирующего апостериорного распределения, т. е. характеризует точность фильтрации. Интересующее нас прогнозируемое распределение (4) также будет нормальным (как произведение нормальных распределений), и его мы представим в виде

$$P(\lambda_1, \dots, \lambda_n/\bar{u}_{n-1}) = k \exp \left\{ -\frac{1}{2} (\bar{\lambda}^{(n)} - \bar{\lambda}_0)^T C^{-1} (\bar{\lambda}^{(n)} - \bar{\lambda}_0) \right\}. \quad (12)$$

Для упрощения расчетов в дальнейшем рассмотрим случай диагональной матрицы A , что обычно соответствует независимости функций правдоподобия на отдельных интервалах наблюдения.

Подставляя (9) и (8) в (4) и сравнивая результат с (12), можно найти интересующие нас достаточные статистики прогнозируемого распределения

$$\bar{\lambda}_o = \bar{\lambda}_{cp} - C^{-1(n-1)} C^{(n)} (\bar{\lambda}_{cp} - \bar{\lambda}^{**}); \quad (13)$$

$$C^{(n)} = \left\| \begin{array}{cc} B + B\bar{V}^T \beta \bar{V} B & -\beta B\bar{V} \\ -\beta \bar{V}^T B & \beta \end{array} \right\|. \quad (14)$$

Здесь

$$\beta = (V_{nn} + \bar{V}^T C \bar{V}) \cdot \frac{1}{V_{nn}^2}, \quad (15)$$

$$B = C^{(n-1)} - \frac{C^{(n-1)} \bar{V}^T \bar{V} C^{(n-1)}}{V_{nn} + \bar{V}^T C^{(n-1)} \bar{V}}. \quad (16)$$

Из (14) и (16) находим

$$-\beta \bar{V}^T B = -\frac{\bar{V}^T C^{(n-1)}}{V_{nn}}. \quad (17)$$

Следовательно, в развернутом виде

$$C_{ni}^{(n)} = -\frac{1}{V_{nn}} \sum_{k=1}^{n-1} V_{nk} C_{ki}^{(n-1)}. \quad (18)$$

Учитывая (18), из (13) можно найти выражение для прогнозируемого значения процесса в виде

$$\lambda_{on} = \sum_{k=1}^{n-1} \frac{V_{nk}}{V_{nn}} (\lambda_{cpk} - \lambda_k^{**})^{(n-1)} + \lambda_{cpn}. \quad (19)$$

Таким образом, нам удалось выразить оценку процесса для момента времени t_n через значения оптимальных оценок, сформированных в момент времени t_{n-1} . Следует отметить, что в выражении (19) фактически имеем дело с оценками для предыдущих интервалов наблюдения, уточняемых в результате расширения этого интервала. В то же время для получения алгоритма следающего типа хотелось бы обрабатывать оценки, поступающие в ритме поступления наблюдаемых данных и не требующие уточнения.

Это удастся сделать, если выразить элементы матриц ковариаций, сформированных на текущем цикле измерения, через элементы матриц ковариаций, полученных в предыдущие моменты, т. е. если суметь разделить данные, приобретаемые на теку-

шем цикле измерения, и данные, полученные ранее. Можно показать, что для элементов матриц ковариаций различных порядков при диагональности матрицы A справедливо соотношение

$$C_{ik}^{(l)} = C_{ik}^{(l-1)} - \frac{A_{il}^{(l)}}{1 - A_{il}^{(l)}C_{il}^{(l)}} C_{il}^{(l)}C_{ik}^{(l)}, \quad (20)$$

где $i, k = 1, 2, \dots, l-1$.

Тогда для оценок внутри интервала можно получить выражение

$$(\lambda_i^{**} - \lambda_{cpi})^{(n-1)} = (\lambda_i^{**} - \lambda_{cpi})^{(l)} + \sum_{k=l+1}^{n-1} \frac{A_{kk}^{(k)}C_{ki}^{(k)}}{1 - A_{kk}^{(k)}C_{kk}^{(k)}} (\lambda_k^* - \lambda_k^{**}). \quad (21)$$

Подставляя (21) в (19), окончательно получаем

$$\lambda_{on} = - \sum_{k=1}^{n-1} \frac{V_{nk}^{(n)}}{V_{nn}} (\lambda_k^{**} - \lambda_{cpi})^{(k)} - \frac{1}{V_{nn}} \sum_{i=1}^{n-2} \sum_{k=i+1}^{n-1} \frac{V_{ni}^{(n)}C_{ik}^{(k)}}{1 - A_{kk}^{(k)}C_{kk}^{(k)}} \times \\ \times A_{kk}^{(k)} (\lambda_k^* - \lambda_k^{**})^{(k)} + \lambda_{cpi}. \quad (22)$$

В этом выражении

$$C_{ik}^{(k)} = - \frac{(1 - A_{kk}C_{kk})}{V_{kk}} \sum_{j=1}^{k-1} V_{kj}^{(k)}C_{ji}^{(k-1)}, \quad (i = 1, \dots, k-1); \quad (23)$$

$$C_{kk}^{(k)} = \left(A_{kk} + V_{kk} - \sum_{l, m=1}^{k-1} V_{kl}B_{lm}V_{mk} \right)^{-1}, \quad i = k. \quad (24)$$

Уравнение (15) будем рассматривать как уравнение для определения дисперсии прогнозируемого значения оценки, а уравнение (22) — трактовать как уравнение следящего измерителя.

Анализируя выражение (22), отметим, что объем памяти, необходимый для вычисления оценок, определяется видом априорного закона распределения экстраполируемых параметров. Элементы матрицы $V_{ni}^{(n)}$ являются весовыми коэффициентами, определяющими вклад каждой из уточненных оценок параметров, относящихся к предыдущим интервалам наблюдения, в результирующее значение параметра на следующий такт работы системы.

Естественно поэтому наиболее просто реализуемые рекуррентные алгоритмы обработки информации получить для процессов типа марковских, обладающих ограниченным последствием, т. е. $V_{ni} = 0$ начиная с некоторого i .

В качестве иллюстрации рассмотрим процесс с независимыми стационарными вторыми приращениями.

В соответствии с [8] для такого процесса имеем

$$\lambda_n - 2\lambda_{n-1} + \lambda_{n-2} = \delta. \quad (25)$$

Для независимых и нормально распределенных δ априорное совместное распределение вероятностей $\lambda_1, \dots, \lambda_n$

$$P(\lambda_1, \dots, \lambda_n) = kP(\lambda_1)P(\lambda_2/\lambda_1) \exp \left\{ -\frac{1}{2\sigma_\delta^2} \sum_{i=3}^n (\lambda_i - 2\lambda_{i-1} + \lambda_{i-2})^2 \right\}. \quad (26)$$

Для интересующих нас элементов матрицы, обратной к корреляционной, получаем

$$V_{nn} = \frac{1}{\sigma_\delta^2}, \quad V_{nn-2} = -\frac{1}{\sigma_\delta^2}, \quad (27)$$

$$V_{nn-1} = -\frac{2}{\sigma_\delta^2}, \quad V_{ni} = 0, \quad i = n-3, n-4, \dots$$

Подставляя (27) в (22), найдем уравнение для оптимального значения оценки λ_{oa} :

$$\lambda_{oa} = -(\lambda_{n-2}^{**} - \lambda_{cpn-2})^{(n-2)} + 2(\lambda_{n-1}^{**} - \lambda_{cpn-1})^{(n-1)} - \frac{C_{n-2n-1}^{(n-1)}}{\sigma_{n-1}^2 - \hat{\sigma}_{n-1}^2} (\lambda_{n-1}^* - \lambda_{n-1}^{**})^{(n-1)} + \lambda_{cpn}, \quad \gamma = \frac{C_{n-2n-1}^{(n-1)}}{\sigma_{n-1}^2 - \hat{\sigma}_{n-1}^2}; \quad (28)$$

$$\lambda_{n-1}^{**} = \lambda_{on-1} + \frac{\hat{\sigma}_{n-1}^2}{\sigma_{n-1}^2} (\lambda_{n-1}^* - \lambda_{on-1}), \quad x = \frac{\hat{\sigma}_{n-1}^2}{\sigma_{n-1}^2}.$$

Для дисперсии предельных ошибок экстраполяции параметра λ_{on} в соответствии с (15) и (20) имеем

$$\frac{1}{\sigma_n^2} = \frac{1}{\sigma_\delta^2 + 4\hat{\sigma}_{n-1}^2 - 4C_{n-1n-2}^{(n-1)} + C_{n-2n-2}^{(n-1)}}, \quad (29)$$

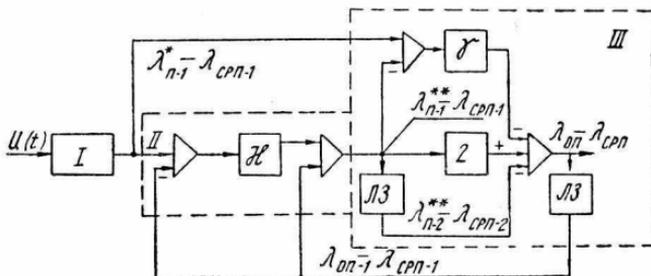
где $C_{n-1n-2}^{(n-1)}$ и $C_{n-2n-2}^{(n-1)}$ определяются рекуррентно в виде

$$C_{n-1n-2}^{(n-1)} = \left(1 - \frac{\hat{\sigma}_{n-1}^2}{\sigma_{n-1}^2} \right) (2\hat{\sigma}_{n-2}^2 - C_{n-3n-2}^{(n-2)}); \quad (30)$$

$$C_{n-2n-2}^{(n-1)} = C_{n-2n-2}^{(n-2)} - \frac{1}{\sigma_{n-1}^2} \left(1 - \frac{\hat{\sigma}_{n-1}^2}{\sigma_{n-1}^2} \right)^{-1} [C_{n-2n-1}^{(n-1)}]^2. \quad (31)$$

Согласно (28), на рисунке представлена схема измерителя-экстраполятора, которая работает следующим образом. На вход блока I поступает колебание $u(t)$, содержащее параметр $\lambda(t)$, который необходимо измерить и предсказать его будущее значение; на выходе блока I формируется оценка по максимуму функции правдоподобия λ_{n-1}^* и подается на вход блока II, осуществляющего фильтрацию оценок максимального правдоподобия и

формирующего в результате этого преобразования оценку по максимуму функции апостериорной вероятности λ_{n-1}^{**} , которая подается на вход блока III, предсказывающего будущее значение параметра λ_{0n} ; в данной схеме преобразования выполняются над смещенными оценками, что не является принципиальным. Таким образом, приведенная схема в виде следящего фильтра с обратными связями, состоящая из сумматоров, линий задержек и умножителей, позволяет измерять оптимальным образом текущее значение параметра и предсказывать его будущее значение.



ВЫВОДЫ

1. Алгоритмы, полученные в результате общего решения задачи оптимальной экстраполяции, весьма сложны.
2. Предложен метод преобразования алгоритмов оптимальной экстраполяции в рекуррентную форму.
3. Сравнительно простую техническую реализацию в виде следящих фильтров при сохранении высокой точности экстраполяции обеспечивают рекуррентные алгоритмы для процессов, обладающих марковскими свойствами.

ЛИТЕРАТУРА

1. А. Н. Колмогоров. Интерполирование и экстраполирование стационарных случайных последовательностей. Изв. АН СССР, сер. Матем., 5, № 1, 1941.
2. N. Wiener. The extrapolation, interpolation and smothing of stationary time series, n. y. John wiley, 1949.
3. R. S. Davis. On the theory of prediction of nonstationary stochastic processes. J. Appl. Phys., 23, № 9, 1952.
4. G. W. Preston. On the theory of prediction of nonstationary stochastic processes. J. Appl. Phys., 24, № 2, 1953.
5. А. М. Яглом. Примеры оптимального нелинейного экстраполирования стационарных случайных процессов. Труды VI Всесоюзного совещания по теории вероятностей и математической статистике, Вильнюс. Изд-во полит. и науч. лит-ры, 1962.
6. Г. П. Тартаковский (под редакцией). Вопросы статистической теории радиолокации, т. II. Изд-во «Советское радио», 1964.
7. С. Е. Фалькович. Оценка параметров сигнала. Изд-во «Советское радио», 1970.
8. А. М. Яглом. Корреляционная теория процесса со случайными стационарными n -ми приращениями. «Математический сборник», т. 37 (79), в. 1, 1955.