УДК681.516
DOI: 10.22213/2413-1172-2020-4-77-84

Анализ и прогнозирование параметров авторегрессионного процесса *p*-го порядка

А. А. Шерстнева, кандидат технических наук, Сибирский государственный университет телекоммуникаций и информатики, Новосибирск, Россия

При решении задач расчета показателей производительности и надежности инфокоммуникационных систем классическими методами делают ряд допущений, например, об экспоненциальном распределении исходных переменных, что не всегда соответствует реальной сетевой ситуации. Кроме того, исходные переменные являются случайными величинами, сбор и обработка которых осуществляются системой мониторинга. Для получения как можно более точных результатов расчета показателей необходимо проводить большое число измерений случайных величин. В этом смысле известные формулы теории телетрафика обладают некоторой неточностью. Один из эффективных путей решения данной проблемы – применение методов регрессионного анализа. Целью статьи является решение задач прогнозирования тренда данных для расчета параметров инфокоммуникационных систем.

Статья направлена на идентификацию авторегрессионного процесса для прогнозирования параметров инфокоммуникационных систем. Применение классического метода регрессионного анализа, такого как МНК-оценка параметров модели ряда, имеет некоторые ограничения. В статье предлагается альтернативный подход к решению обозначенной проблемы через теорию фильтрации сигналов.

Наряду с классическим оцениванием методом наименьших квадратов одной из возможностей является решение эмпирических уравнений Юла – Уокера. В статье приводится методика решения по алгоритму Левинсона – Дурбина. Помимо теоретических выкладок разработана программа, позволяющая автоматизировать процесс вычислений.

В зависимости от количества наблюдений выбирается критерий, минимизация которого приводит к выбору истинного порядка модели. Оценка осуществляется с использованием программных средств, а именно программы математического моделирования Matlab. В связи с этим в статье рассматривается не только возможность выбора критерия с теоретической точки зрения, но и его практическая реализация. Также приведены условия использования и показана наиболее эффективная методика выбора критерия в зависимости от количества наблюдений.

Ключевые слова: регрессия, анализ, метод наименьших квадратов, уравнения Юла – Уокера, алгоритм Левинсона – Дурбина.

Введение

При рассмотрении статистических методов оценивания параметров моделей временных рядов и их прогнозировании предполагают, что последовательность наблюдений ряда достаточно велика, чтобы проявились асимптотические свойства оценок. Получаемые оценки при этом близки к их истинным теоретическим значениям. Необходимым условием является достаточность периода основания прогноза для выявления тенденции временного ряда по предыдущим наблюдениям и возможности прогнозирования поведения процесса на будущее.

При рассмотрении короткого ряда невозможно делать долгосрочное прогнозирование, имея в виду только прошлые наблюдения. В таком случае привлекается дополнительная информация о прогнозируемом процессе. Эта проблема решается с помощью экспертных систем для сбора информации о предполагаемых свойствах ряда.

Однако существуют и другие проблемы, которые требуют детальной проработки. При анализе коротких рядов сложные модели не рассматриваются для описания трендов, так как их оценивание требует больших выборок. В этом случае оказывается невозможным применение традиционных статистических методов оценки и определение доверительных интервалов. Если период основания прогноза мал, то возникают проблемы, связанные с недостатком априорной информации. МНК-оценка параметров модели ряда, применяющаяся в области регрессионного анализа, ненадежна. В этом смысле цифровая фильтрация находит широкое применение в разнообразных задачах оценивания. Популярность обусловлена возможностью эффективного решения технических вопросов и реализацией через программы математического моделирования. На базе теории фильтрации сигналов созданы системы навигации и управления космическими аппаратами и другими подвижными объектами. Предлагаемая теория также находит широкое применение при расчете надежности разных инфокоммуникационных систем.

Область применения цифровой фильтрации относится к беспроводным сетям и системам связи, беспроводным системам безопасности, радиолокационным системам, дистанционному зондированию окружающей среды и медицинским системам. Фильтры в составе всевозможных систем радиочастотной и СВЧ-связи широко распространены и особенно актуальны сегодня, когда беспроводная связь обещает предоставить голосовой доступ и доступ к данным кому угодно, где угодно и в любое время.

Постановка задачи

Технические процессы, характеризующие функционирование сложных технических систем, как правило, представляются математическими моделями. При исследовании математических моделей получаем системы уравнений с несколькими показателями. Предопределенные переменные используются в моделях с запаздыванием, в которых каждое новое состояние зависит не только от времени, но и от предыдущего состояния.

В качестве примера на рисунке 1 в виде графа состояний приведена математическая модель контроля работоспособности технического состояния функционального блока.

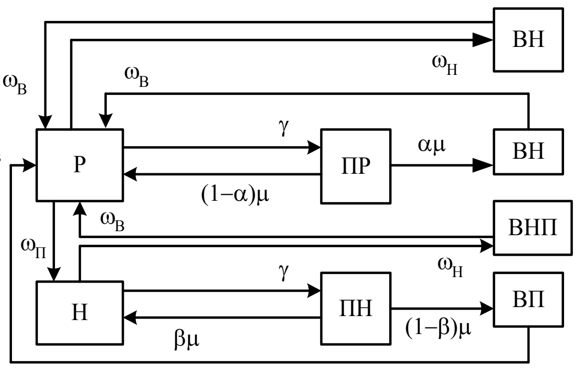

Рис. 1. Граф состояний функционального блока

Fig. 1. The state graph of the functional block

Работоспособность функционального блока контролируется двумя видами контроля – непрерывным и периодическим. Выделены соответствующие состояния. Функциональный блок включен в рабочую конфигурацию и находится в работоспособном состоянии (Р) и неработоспособном из-за отказа, который возможно обнаружить только периодическим контролем (Н). Введены состояния проверки ПР и ПН соответственно. После обнаружения отказа функциональный блок восстанавливается соответственно виду контроля, который обнаружил отказ, т. е. ВН, ВНП, ВП. Особенность этой модели в том, что в ней учитываются ошибки контроля 1-го (α) и 2-го (β) рода. Между состояниями функциональный блок переходит из-за событий, описываемых интенсивностью отказов $\left(\omega_{0}=\omega_{Н}+\omega_{П}\right)$, обнаруживаемых непрерывным ($\omega_Н$) и периодическим ($\omega_П$) контролем, а также интенсивностью проведения контроля (γ), его окончания (μ) и интенсивностью восстановления ($\omega_В$). Используя матричный метод расчета, получаем систему уравнений (1), включающую ряд случайных переменных параметров:

$$\left.\begin{aligned} K_П &= \frac{\omega_П\omega_В + \left(\omega_Н + \gamma(1-\beta)\right)\left(\omega_0 + \alpha\gamma\right)}{\omega_В\left(\omega_0 + \gamma(1-\beta)\right) + \left(\omega_Н + \gamma(1-\beta)\right)\left(\omega_0 + \alpha\gamma\right)} \\ t_Н &= \frac{\omega_П}{\left(\omega_Н + \gamma(1-\beta)\right)\left(\omega_0 + \alpha\gamma\right)} \\ K_Г &= 1 - K_П. \end{aligned}\right\} \qquad (1)$$

В (1) приведены формулы для расчета коэффициента простоя ($K_П$) и коэффициента готовности ($K_Г$) функционального блока, среднее время необнаруженного отказа ($t_Н$). В состояние необнаруженного отказа функциональный блок попадает в том случае, если отказ обнаруживается только периодическим контролем [1].

Вычисление значений каждого показателя, входящего в расчетные формулы, рассматривается как модель многомерной регрессии, фундаментальные исследования которой приводят к оценке параметров авторегрессионного процесса [2, 3].

Цель исследования – идентификация авторегрессионного процесса для прогнозирования параметров инфокоммуникационных систем.

Актуальной является задача оценивания параметров авторегрессионной модели и дисперсии. Для решения задачи определения порядка модели пользуются, как правило, упрощенным способом. Метод заключается в допущении того, что временной ряд является авторегрессионным процессом, и тогда порядок авторегрессии увеличивают до тех пор, пока сумма квадратов ошибок в схеме оценивания параметров авторегрессии при МНК-подходе не перестанет ощутимо уменьшаться. Указанный метод имеет очевидные недостатки, например, сложность вычислений параметров авторегрессионного процесса, обусловленного порядком модели. МНК-подход эффективен для порядка модели не более двух [4–6]. Задача заключается в рассмотрении моделей более высокого порядка. Как правило, при построении моделей временных рядов критерии качества подгонки моделей применяются для сравнения моделей между собой. Поскольку оценки коэффициентов проводятся путем оптимизации, фактически речь идет о выборе порядка модели, т. е. о сравнении моделей с различным числом параметров. Абсолютные критерии типа стандартного коэффициента множественной детерминации не применяются. В статье предлагается альтернативный подход к решению обозначенной проблемы. В зависимости от количества наблюдений выбирается критерий, минимизация которого приводит к выбору истинного порядка модели. Оценка осуществляется с использованием программных средств, а именно программы математического моделирования Matlab. В связи с этим в статье рассматривается не только возможность выбора критерия с теоретической точки зрения, но и его практическая реализация. Также приведены условия использования и показана наиболее эффективная методика выбора критерия в зависимости от количества наблюдений.

Теоретическая часть

Авторегрессионная модель создается с помощью программы математического моделирования Matlab. Функция filter с входным вектором (b, a, z) фильтрует последовательность данных $Z[m]$ при использовании цифрового фильтра с коэффициентами $a_n \ (n = 0, 1, 2, \ldots, p)$ и $b_n \ (n = 0, 1, 2, \ldots, p)$.

Выходная последовательность X_m определяется выражением

$$x[m] = \frac{1}{a_0}\left(\sum_{n=0}^{q} b_1 z[m-1] - \sum_{n=1}^{p} a_1 x[m-n]\right),$$

где p – порядок фильтра обратной связи; q – порядок фильтра прямой связи.

Порядок фильтра прямой связи должен быть $q = 0$, чтобы создать авторегрессионую модель

$$x[m] = \frac{1}{a_0}\left(b_0 z[m] - \sum_{n=1}^{p} a_n x[m-n]\right).$$

Это выражение показывает, что коэффициенты фильтра b_0 и a_0 должны быть равны единице для создания авторегрессионной модели порядка p:

$$x[m] = z[m] - \sum_{n=1}^{p} a_n x[m-n],$$

где $b = b_0 = 1$, $a = (1, a_1, \ldots, a_p)$; последовательность $Z[m]$ – белый шум.

На рисунке 2 показана блок-схема функции filter с входной последовательностью $x[m]$ и выходной последовательностью $y[m]$.

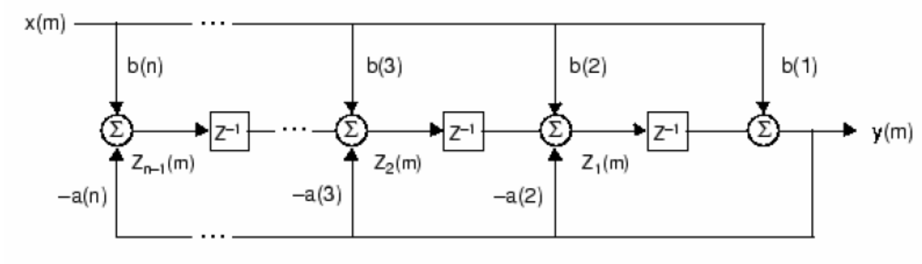

Рис. 2. Блок-схема функции filter Matlab

Fig. 2. Block diagram of filter function

Выражение авторегрессионной модели может быть представлено в виде [7–9]

$$\sum_t \left(X_t - \left(\underbrace{-a_1}_{c_1} X_{t-1} \underbrace{-a_2}_{c_2} X_{t-2} - \ldots \underbrace{-a_k}_{c_k} X_{t-k} \right) \right)^2 = \\ = \sum_t Z_t^2 .$$

С учетом того, что $-a_i = c_i$, выражение, приведенное выше, можно интерпретировать как сумму ошибок в зависимости от параметра c_i :

$$q(c) = \sum_t \left(X_t - \left(c_1 X_{t-1} c_2 X_{t-2} \ldots c_k X_{t-k} \right) \right)^2 = \sum_t Z_t^2 .$$

Это выражение приводит к методу МНК с параметрами $c_1 \ldots c_k$ и матричной записи, состоящей из $(N - k)$ уравнений и k неизвестных:

$$\underbrace{\begin{bmatrix} X_{k+1} \\ X_{k+2} \\ \vdots \\ X_N \end{bmatrix}}_{X} - \underbrace{\begin{bmatrix} X_{k1} & X_{k-1} & \cdots & X_1 \\ X_{k+1} & X_k & \cdots & X_2 \\ \vdots & & \ddots & \\ X_{N-1} & X_{N-2} & \cdots & X_{N-k} \end{bmatrix}}_{H} \underbrace{\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{bmatrix}}_{c} = \underbrace{\begin{bmatrix} z_{k+1} \\ z_{k+2} \\ \vdots \\ z_N \end{bmatrix}}_{Z},$$

$$N \geq 2k. \qquad (2)$$

Минимальное количество наблюдений $n = 2k$ для решения системы уравнений, которая привела бы к интерполяции экспериментальных данных. Поэтому для оценки методом наименьших квадратов требуется $n > 2k$ наблюдений. Параметр c_i, который минимизирует сумму квадратов ошибок, необходимо вычислить $2k$ раз.

МНК-оценка фильтра с коэффициентами $\hat{a} = (\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_k)$ и дисперсией σ_Z^2 определяется с помощью матрицы проекции, также известной как матрица влияния, и отображает вектор значений отклика (зависимых переменных) к вектору прогнозируемых значений.

Другой возможностью оценки параметров фильтра и дисперсии является расчет эмпирических уравнений Юла – Уокера [10, 11].

При соответствующей оценке ковариационной функции авторегрессионного процесса эмпирические уравнения Юла – Уокера могут быть рассчитаны в матричной форме:

$$\begin{pmatrix} \hat{c}_{XX}(0) & \hat{c}_{XX}(1) & \cdots & \hat{c}_{XX}(k-1) \\ \hat{c}_{XX}(1) & \hat{c}_{XX}(0) & \cdots & \hat{c}_{XX}(k-2) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{c}_{XX}(k-1) & \hat{c}_{XX}(k-2) & \cdots & \hat{c}_{XX}(0) \end{pmatrix} \begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \vdots \\ \hat{a}_k \end{pmatrix} =$$

$$= - \begin{pmatrix} \hat{c}_{XX}(1) \\ \hat{c}_{XX}(2) \\ \vdots \\ \hat{c}_{XX}(k) \end{pmatrix},$$

$$N \geq k + 1. \qquad (3)$$

Необходимо только $k + 1$ количество наблюдений, чтобы получить

$$\hat{c}_{XX}(0), \hat{c}_{XX}(1), \hat{c}_{XX}(k).$$

Дисперсию можно оценить, используя эмпирическое уравнение Юла – Уокера,

$$\hat{\sigma}_Z^2 = \hat{c}_{XX}(0) + \sum_{n=1}^{k} \hat{a}_n \hat{c}_{XX}(n).$$

С помощью системы уравнений Юла – Уокера получается матрица, в которой совпадают коэффициенты на главной диагонали и все коэффициенты на диагоналях, равноотстоящих от главной. Матрица такого вида является симметричной теплицевой матрицей. Если рекурсивный фильтр авторегрессионного процесса устойчив, матрица коэффициентов также положительно определена. Однако предположив, что параметры $\hat{c}_{XX}(0), \hat{c}_{XX}(1), \hat{c}_{XX}(k)$ известны заранее, система уравнений может быть решена однозначно для $\hat{a}_1, \hat{a}_2, \hat{a}_k$ по алгоритму Левинсона – Дурбина.

Необходимо отметить, что матрица коэффициентов (2) не может быть решена с помощью алгоритма Левинсона – Дурбина, потому что это несимметричная теплицева матрица и в большинстве случаев – переопределенная система уравнений, которая приводит к МНК-оценке.

Преимуществом алгоритма Левинсона – Дурбина является и рекурсивная структура расчета, которая рассчитывает коэффициенты для порядков < k, что позволяет быстро оценить порядок модели.

В связи с тем, что оценочная дисперсия $\hat{\sigma}_{Z,k}^2$ практически постоянна при $k > p$, существуют методы для оценки истинного порядка авторегрессионного процесса. Одним из них является критерий Акаике (АК) [12]:

$$\text{АК}(k) = \ln \hat{\sigma}_Z^2 + \frac{k}{2N}, \qquad (4)$$

где k, которое приводит к минимуму АК(k), является оценкой истинного порядка модели p. Критерий Акаике является смещенной оценкой и верным выбором, если число наблюдений N мало. Информационный критерий Акаике базируется на обобщении принципа максимального правдоподобия. Случайное возмущение при этом является гауссовым.

Другим методом оценки порядка модели является критерий Риссанена (РК) [13]:

$$\text{РК}(k) = \ln \hat{\sigma}_Z^2 + \frac{\ln(N)}{2N}. \qquad (5)$$

Оценка истинного порядка модели (k) достигается также за счет минимизации критерия $\text{РК}(k)$. В отличие от критерия Акаике критерий Риссанена является несмещенной оценкой и подходит в том случае, когда количество наблюдений N велико [14–17].

Экспериментальная часть

Первоначально авторегрессионный процесс генерируется через фильтрацию белого шума с помощью рекурсивного цифрового фильтра с коэффициентами $a_1 = 0{,}5$; $a_2 = 0{,}3$; $a_3 = 0{,}1$; $a_4 = 0{,}7$; $a_5 = 0{,}3$. Возможности программы математического моделирования Matlab позволяют выполнить эту процедуру.

Задача состоит в идентификации авторегрессионного процесса по алгоритму Левинсона – Дурбина. Необходимо найти параметры рекурсивного фильтра $a_i, i = 0, 1, 2, 3, 4, 5,$ и дисперсию $\hat{\sigma}_Z^2$.

Решением поставленной задачи стало написание функции covfct для нахождения выборочной ковариационной функции и levindur в среде Matlab. С помощью этих двух функций возможно идентифицировать авторегрессионный процесс.

Быстрый алгоритм Левинсона – Дурбина вычисляет решение для системы уравнений согласно эмпирическому уравнению Юла – Уокера с $O(k^2)$ плавающей точкой. В таблице 1 приведены оценочные коэффициенты $AR(p)$-процесса, решающего эмпирическое уравнение Юла – Уокера по алгоритму Левинсона – Дурбина.

Таблица 1. **Оценочные коэффициенты *AR*(*p*)-процесса**

Table 1. **Estimated coefficients of the *AR*(*p*) process**

Коэффициент	Расчетное значение	Ошибка
$\hat{a}_1$	0,4766	0,0234
$\hat{a}_2$	0,2881	0,0119
$\hat{a}_3$	0,0913	0,0087
$\hat{a}_4$	0,6695	0,0305
$\hat{a}_5$	0,2895	0,0105
	Σ	0,0850

Дисперсию также можно оценить, решив эмпирическое уравнение Юла – Уокера:

$$\sigma_Z^2 = 0{,}8918.$$

Следующей задачей для рассмотрения стало определение порядка модели. Первоначально оценка параметров $a_i, i = 0, 1, 2, 3, 4, 5,$ и дисперсии осуществляется для неверного порядка $p = 4$ и $p = 6$.

По теоретическим выкладкам (4), (5) проводится оценка порядка модели и дисперсии для порядка k ($k = 1, \ldots, 10$) согласно алгоритму Левинсона – Дурбина.

Повторная оценка порядка модели показывает лучшее приближение к реальным значениям. В таблице 2 приведены расчетные коэффициенты с использованием неверного порядка моделей $k = 4$ и $k = 6$.

На рисунке 3 приведены графики оценки дисперсии (*a*, *b*) и порядка модели: с помощью критерия Акаике (*c*, *d*), критерия Риссанена (*e*, *f*).

Оценка порядка модели с помощью критерия Акаике и Риссанена – это k, которое минимизирует функцию. Оба минимума находятся при $k = 5 = p$, что является истинным порядком модели. Поэтому в данном случае оба критерия оценивают порядок авторегрессионного процесса.

Заключение

В статье рассматривается интеллектуальный анализ данных и приведена методология идентификации авторегрессионного процесса p-го порядка через оценку коэффициентов рекурсивного цифрового фильтра и дисперсии.

Дисперсия находится через решение системы уравнений Юла – Уокера. Выборочная ковариационная функция определяется программно в среде Matlab. Параметры модели оценены с помощью алгоритма Левинсона – Дурбина, полученные коэффициенты авторегресионного процесса максимально приближены к реальным значениям модели.

Оценка порядка модели достигается за счет минимизации критерия Аккаике либо Риссанена. Выбор критерия целесообразно проводить исходя из количества наблюдений. Для нахождения порядка модели также была написана программа по заданным параметрам. Повторные оценки и заранее неверно заданный критерий подтверждают достоверность результатов исследования.

Таблица 2. **Расчетные коэффициенты**

Table 2. **Calculated coefficients**

Порядок k	$\hat{a}_1$	$\hat{a}_2$	$\hat{a}_3$	$\hat{a}_4$	$\hat{a}_5$	$\hat{a}_6$	σ_Z^2
4	0,3087	0,2855	0,0087	0,5801	0	0	0,9733
5	0,4766	0,2881	0,0913	0,6695	0,2895	0	0,8918
6	0,4874	0,3130	0,0947	0,6802	0,3072	0,0372	0,8905

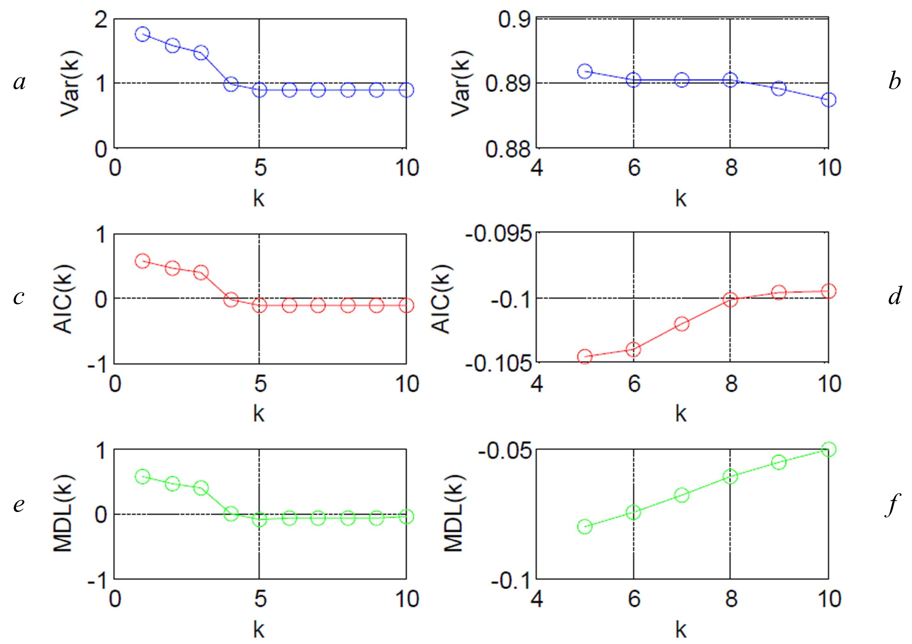

Рис. 3. Оценка дисперсии и порядка модели с использованием критериев Акаике, Риссанена

Fig. 3. Estimation of variance and order of the model using the Akaike, Rissanen criteria

Построение сложносоставной модели с множеством параметров и созависимостями внутри самой модели является комплексной задачей, требующей тщательного анализа и проработки каждого показателя. Одним из способов улучшения качества прогнозов можно рассматривать составление «частных» моделей, описывающих конкретный показатель и его созависимости. В статье рассматривается построение модели типа авторегрессии для описания поведения отдельного параметра. Это необходимо сделать, поскольку между переменными системы существуют сложные взаимосвязи, часто нелинейные, для описания которых необходимо подобрать адекватные функциональные зависимости. Проблема оценки параметров модели, описываемой системами взаимосвязанных уравнений, существенно сложнее, чем оценивание одномерных моделей. Кроме того, необходимо принимать во внимание, что с течением времени меняются не только параметры системы, но и структура самих уравнений, иными словами, устаревает вид функциональных зависимостей, описывающих ее состояние. Ответом может стать создание модели, работающей в режиме реального времени, способной к обучению и переобучению исходя из действительной ситуации инфокоммуникационной системы.

Библиографические ссылки

1. *Шерстнева А. А.*, *Шерстнева О. Г.* Анализ сети связи с учетом показателей надежности // Вестник Рязанского гос. радиотехнического ун-та. 2020. № 73 (2). С. 52–58.

2. *Домбровский В. В.*, *Объедко Т. Ю.* Оптимальные стратегии прогнозирующего управления системами со случайными параметрами, описываемыми многомерной регрессионной моделью с марковским переключением режимов // Вестник Томского гос. ун-та. 2019. № 48. С. 4–12.

3. *Домбровский В. В.*, *Пашинская Т. Ю.* Прогнозирующее управление системами с марковскими скачками и авторегрессионным мультипликативным шумом с марковским переключением режимов // Вестник Томского гос. ун-та. 2018. № 44. С. 4–9.

4. *Афанасьев А. А.*, *Власов Р. С.* Особенности выделения сегментов анализа речевого сигнала при его обработке // Перспективные технологии в средствах передачи информации : сборник тр. конф. / Владимирский гос. ун-т, 2019. С. 156–160.

5. *Моттль В. В.*, *Красоткина О. В.*, *Ежова Е. О.* Непрерывное обобщение информационного критерия Акаике для оценивания нестационарной регрессионной модели временного ряда с неизвестной степенью изменчивости коэффициентов // Математические методы распознавания образов. 2009. № 1. С. 52–55.

6. *Вилков А. П.*, *Родионова Т. Е.* Использование систем одновременных уравнений для получения моделей описания технических объектов // Современные проблемы проектирования, производства и эксплуатации радиотехнических систем. 2016. № 10. С. 175–177.

7. Св. о регистрации электронного ресурса № 17760 / О. Г. Шерстнева, А. А. Шерстнева. Программа имитации функционирования телекоммуникационной сети с учетом реальных показателей надежности. 29.12.2011 г.

8. *Wickham H.* Elegant graphics for data analysis. 2nd ed. Springer, 2016, 213 p.

9. *Athanasopoulos G.*, *Hyndman R.J.*, *Kourentzes N.*, *Petropoulos F.* Forecasting with temporal hierarchies. European J. of Operational Research, 2017, no. 262, pp. 60-74.

10. *Bergmeir C.*, *Hyndman R.J.*, *Benítez J.M.* Bagging exponential smoothing methods using STL decomposition and Box-Cox transformation. International J. of Forecasting, 2016, no. *3*2, pp. 303-312.

11. *Sherstneva A.*, *Sherstneva O.* Analysis statistical information for data trend forecasting. International Ural conference on Electrical power Engineering. IEEE, 2020, pp. 153-158.

12. *Кантарович Г. Г.* Экономическая статистика. Эконометрика. М. : ГУ-ВШЭ, 2000.

13. *Judge G.G.*, *Griffits W.E.*, *Hill R.C.* The theory and practice of econometrics. NY, John Willey and Sons, 1985.

14. *Bergmeir C.*, *Hyndman R.J.*, *Koo B.* A note on the validity of cross-validation for evaluating autoregressive time series prediction. Computational Statistics and Data Analysis, 2018, no. 120, pp.70-83.

15. *Wickramasuriya S.L.*, *Athanasopoulos G.* Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization. J. American Statistical Association, 2019, no. 114, pp. 804-819.

16. *Harrell F.E.* Regression modeling strategies: With applications to linear models, logistic and ordinal regression, and survival analysis. 2nd ed. NY, Springer, 2015, 568 p.

17. *Madsen K.*, *Nielsen H.B.*, *Tingleff O.* Methods for Non-linear Least Squares Problem Cobenhavn. Technical University of Denmark, 2004, 30 p.

References

1. Sherstneva A.A., Sherstneva O.G. [Analysis of the communication network taking into account reliability indicators]. *Vestnik Ryazanskogo gos. radiotekhnicheskogo un-ta*, 2020, no. 73, pp. 52-58 (in Russ.).

2. Dombrovskii V.V., Ob"edko T.Yu. [Optimal strategies for predictive control of systems with random parameters described by a multidimensional regression model with Markov switching of modes]. *Vestnik Tomskogo gos. un-ta*, 2019, no. 48, pp. 4-12 (in Russ.).

3. Dombrovskij V.V., Pashinskaya T.YU. [Predictive control of systems with Markov jumps and autoregressive multiplicative noise with Markov mode switching]. *Vestnik Tomskogo gos. un-ta*, 2018, no. 44, pp. 4-9 (in Russ.).

4. Afanas'ev A.A., Vlasov R.S. *Osobennosti vydeleniya segmentov analiza rechevogo signala pri ego obrabotke* [Features of the selection of segments of the analysis of a speech signal during its processing]. *Perspektivnye tekhnologii v sredstvakh peredachi informatsii : cbornik tr. konf.* [Perspective technologies in the means of information transmission]. Proc. of conf., 2019, pp. 156-160.

5. Mottl' V.V., Krasotkina O.V., Ezhova E.O. [Continuous generalization of the Akaike information criterion for estimating a non-stationary regression model of a time series with an unknown degree of variability of the coefficients]. *Matematicheskie metody raspoznaniya obrazov*, 2009, no. 1, pp. 52-55 (in Russ.).

6. Vilkov A.P., Rodionova T.E. [Using systems of simultaneous equations to obtain models for describing technical objects]. *Sovremennye problem proektirovaniya, proizvodstva i ekspluatacii radiotekhnicheskih sistem*, 2016, no. 10, pp. 175-177 (in Russ.).

7. Sherstneva O.G., Sherstneva A.A. *Programma imitatsii funktsionirovaniya telekommunikatsionnoi seti s uchetom real'nykh pokazatelei nadezhnosti* [A program for simulating the functioning of a telecommunication network taking into account real reliability indicators]. Sv. o registratsii elektronnogo resursa № 17760, 29.12.2011 g.

8. Wickham H. Elegant graphics for data analysis. 2td ed. Springer, 2016, 213 p.

9. Athanasopoulos G., Hyndman R.J., Kourentzes N., Petropoulos F. Forecasting with temporal hierarchies. *European J. of Operational Research*, 2017, no. 262, pp. 60-74.

10. Bergmeir C., Hyndman R.J., Benítez J.M. Bagging exponential smoothing methods using STL decomposition and Box-Cox transformation. *International J. of Forecasting*, 2016, no. *3*2, pp. 303-312.

11. Sherstneva A., Sherstneva O. Analysis statistical information for data trend forecasting. *International Ural conference on Electrical power Engineering*. IEEE, 2020, pp. 153-158.

12. Kantarovich G.G. *Ekonomicheskaya statistika. Ekonometrika* [Economic statistics. Econometrics]. Moscow, GU-VShE Publ., 2000.

13. Judge G.G., Griffits W.E., Hill R.C. The theory and practice of econometrics. NY, John Willey and Sons, 1985.

14. Bergmeir C., Hyndman R.J., Koo B. A note on the validity of cross-validation for evaluating autoregressive time series prediction. *Computational Statistics and Data Analysis*, 2018, no. 120, pp. 70-83.

15. Wickramasuriya S.L., Athanasopoulos G. Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization. *J. American Statistical Association*, 2019, no. 114, pp. 804-819.

16. Harrell F.E. Regression modeling strategies: With applications to linear models, logistic and ordinal regression, and survival analysis. 2nd ed. NY, Springer, 2015, 568 p.

17. Madsen K., Nielsen H.B., Tingleff O. Methods for Non-linear Least Squares Problem Cobenhavn. Technical University of Denmark, 2004, 30 p.

Analysis and Forecasting of Parameters for p-Order Autoregressive Process

A.A. Sherstneva, PhD in Engineering, Siberian State University of Telecommunications and Information Sciences, Novosibirsk, Russia

Several assumptions are made while solving the problems of calculating the performance and reliability metrics of info-communication systems using classical methods – for instance, the exponential distribution of source variables does not always correspond to real data. Moreover, source variables are random variables collected and processed by the monitoring system. To obtain the most accurate results of the indicator calculation, it is necessary to conduct a large number of measurements of random variables. In this sense, the well-known formulas of teletraffic theory have some inaccuracy. One of the effective ways to solve this problem is the use of regression analysis. The paper purposes for estimating the metrics in info-communication systems for data trend forecasting.

The paper aims to identify the autoregressive process for parameters forecasting for info-communication systems. The use of a classical regression analysis method, such as least squares estimation, has several limitations. The paper proposes an alternative approach to solving the indicated problem through the filtration theory.

Along with the classical least squares estimation, an alternative opportunity is to solve the empirical Yule-Walker equations. The paper presents a solution technique using the Levinson-Durbin algorithm. In addition to theoretical calculations, a program has been developed to automate the computation process.

Depending on the number of observations, a criterion is selected, its minimization leads to the model's real order. The estimation is carried out using the mathematical modeling program Matlab. In this regard, the paper considers the possibility of choosing a criterion from a theoretical point of view and its practical implementation. The conditions of use are also given; the most effective method for criterion choosing is shown depending on the number of observations.

Keywords: regression, analysis, least squares estimation, Yule-Walker equations, Levinson-Durbin algorithm.

Получено 06.11.2020

Образец цитирования

Шерстнева А. А. Анализ и прогнозирование параметров авторегрессионного процесса *p*-го порядка // Вестник ИжГТУ имени М. Т. Калашникова. 2020. Т. 3, № 4. С. 77–84. DOI: 10.22213/2413-1172-2020-4-77-84.

For Citation

Sherstneva A.A. [Analysis and Forecasting of Parameters for *p*-Order Autoregressive Process]. *Vestnik IzhGTU imeni M.T. Kalashnikova*, 2020, vol. 23, no. 4, pp. 77-84 (in Russ.). DOI: 10.22213/2413-1172-2020-4-77-84.