## Секция «Математика и механика»

### Трендовый, декомпозиционный и комбинированный подходы прогнозирования временных рядов на основе метода «Гусеница»-SSA

***Щелкалин Виталий Николаевич***

*Аспирант*

*Харьковский национальный университет радиоэлектроники, факультет прикладной математики и менеджмента, Харьков, Украина*

*E-mail: schelkalin.vitaly@gmail.com*

К настоящему моменту времени создано большое количество мат. моделей и методов анализа и прогнозирования временных рядов (в. р.) и наметилась тенденция комбинирования их идей с целью получения лучших характеристик конечной, в смысле формирования, модели.

Трендовый подход в работе реализовывается путём совместного использования методов «Гусеница»-SSA (SSA) и Бокса-Дженкинса (BJ), т.е. путём добавления к линейной рекуррентной формуле (ЛРФ) SSA модели АРПСС (1), (2). Впоследствии модель приняла вид (3), а все её параметры уже подстраивались методом оптимизации целевой функции интегрированного критерия точности и адекватности. Модель (3) названа автором моделью авторегрессии – спектрально проинтегрированного скользящего среднего (АРСПСС). Идентифицируемый методом SSA сигнал достаточно хорошо аппроксимируется моделями АРПСС (см. модель (4)). Для случая нескольких сезонных составляющих и экзогенных переменных модели обобщаются подобным образом как, например, модель (4) обращается в (6).

$$y_t = g(B)s_t + \frac{c(B)}{\nabla a(B)}e_t; \Rightarrow \tag{1}$$

$$\Rightarrow y_t = g(B)y_t + \frac{c'(B)}{\nabla a'(B)}e_t; \Rightarrow \tag{2}$$

$$\Rightarrow g'(B)y_t = \frac{c''(B)}{\omega(B)a''(B)}e_t; \tag{3}$$

$$\begin{cases} \hat{s}_t = \dfrac{f(B)}{\nabla d(B)}\varepsilon_t; \\ y_t = \hat{s}_t + \dfrac{c(B)}{\nabla a(B)}e_t. \end{cases} \Rightarrow y_t = \frac{f(B)}{\nabla d(B)}\varepsilon_t + \frac{c(B)}{\nabla a(B)}e_t; \tag{4}$$

$$y_t = g(B)\frac{f(B)}{\nabla d(B)}\varepsilon_t + \frac{c'''(B)}{\nabla a'''(B)}e_t; \tag{5}$$

$$y_t = \left(\sum_{i=1}^{N}\frac{\omega_i'(B)}{\delta_i'(B)} + \sum_{i=1}^{N}\frac{\omega_i(B)}{\delta_i(B)}\right)B^{m_i'+m_i}x_t + \frac{\prod_{i=1}^{S_n'} f^{S_i'}(B)}{\prod_{i=1}^{S_n'}\nabla_{S_i'}^{D_i'}\prod_{i=1}^{S_n'} d^{S_i'}(B)}\varepsilon_t +$$

$$+\frac{\prod_{i=1}^{S_n} c^{S_i}(B)}{\prod_{i=1}^{S_n}\nabla_{S_i}^{D_i}\prod_{i=1}^{S_n'} a^{S_i}(B)}e_t. \tag{6}$$

где $s_t$ — идентифицированный методом SSA сигнал; $g(B) = \sum_{i=1}^{L-1} g_{L-i}B^i$; $L$ — длина окна метода SSA или $g(B) = \sum_{i=1}^{R-1} g''_{R-i}B^i$, где $R$ — порядок ЛРФ минимальной длины; $B$ — оператор задержки; $g_i$, $i = \overline{1, L-1}$ — коэффициенты миннорм ЛРФ [1]; $g'(B) = 1 - g(B)$; $\nabla = 1 - B$. Оператор $g(B)$ может применяться непосредственно к процессу $y_t$, тогда (1) запишется как (2), а в (3) — $\omega(B) \equiv \nabla$. Остальные обозначения, описаны в [1].

Суть же предложенных декомпозиционного и комбинированного методов прогнозирования состоит в декомпозиции в. р. (прогнозируемого и экзогенных) методом SSA на составляющие, которые в свою очередь могут быть разложены на компоненты с более простой для идентификации структурой (рис. 1); в отборе каким-либо методом из этих составляющих конструктивных и отбрасывании деструктивных; в идентификации для тех из конструктивных составляющих, которые имеют упреждающий характер на прогнозируемый в. р., или наоборот интервал запаздывания которых меньше требуемого интервала упреждения прогнозирования, моделей BJ (при декомпозиционном подходе (формула 7)) или других видов мат. моделей с наиболее подходящими структурами для каждой конкретной составляющей в. р. (при комбинированном подходе) и расчёт их прогнозов с требуемым упреждением; в использовании полученных моделей или как гребёнки фильтров (в случае моделирования сигналов), или как ансамбля моделей, подавая на вход MISO модели или использовать как составляющие комбинированной модели, параметры которой далее совместно подстраиваются оптимизационным методом [1].

$$\hat{s}_t^{(j)} = \frac{\omega^{(j)}(B)}{\delta^{(j)}(B)} y_{t-m^{y^{s(j)}}} + \sum_{i=1}^{N} \frac{\omega_x^{s^{(j)}i}(B)}{\delta_x^{s^{(j)}i}(B)} x^i_{t-m^{x^{i w^{(j)}}}} +$$

$$+ \sum_{i=1, i\neq j}^{M} \frac{\omega_s^{s^{(j)}i}(B)}{\delta_s^{s^{(j)}i}(B)} \hat{s}^{(i)}_{t-m^{s^{(i)}}} + \frac{\prod_{i=1}^{S_n^{s^{(j)}}} c^{S_i^{s^{(j)}}}(B)}{\prod_{i=1}^{S_n^{s^{(j)}}} \nabla^{D_i^{s^{(j)}}}_{S_i^{s^{(j)}}} \prod_{i=1}^{S_n^{s^{(j)}}} a^{S_i^{s^{(j)}}}(B)} e_t^{s^{(j)}};$$

$$\vdots$$

$$\hat{s}_t^{(k)} = \frac{\omega^{(k)}(B)}{\delta^{(k)}(B)} y_{t-m^{y^{s(k)}}} + \sum_{i=1}^{N} \frac{\omega_x^{s^{(k)}i}(B)}{\delta_x^{s^{(k)}i}(B)} x^i_{t-m^{x^{i s^{(k)}}}} +$$

$$+ \sum_{i=1, i\neq k}^{M} \frac{\omega_s^{s^{(k)}i}(B)}{\delta_s^{s^{(k)}i}(B)} \hat{s}^{(i)}_{t-m^{s^{(i)}}} + \frac{\prod_{i=1}^{S_n^{s^{(k)}}} c^{S_i^{s^{(k)}}}(B)}{\prod_{i=1}^{S_n^{s^{(k)}}} \nabla^{D_i^{s^{(k)}}}_{S_i^{s^{(k)}}} \prod_{i=1}^{S_n^{s^{(k)}}} a^{S_i^{s^{(k)}}}(B)} e_t^{s^{(k)}};$$

$$y_t = \sum_{i=1}^{M} \frac{\omega^i(B)}{\delta^i(B)} \hat{s}^{(i)}_{t-m_i^s} + \sum_{i=1}^{N} \frac{\omega_x^{yi}(B)}{\delta_x^{yi}(B)} x^i_{t-m_i} + \frac{\prod_{i=1}^{S_n} c^{S_i}(B)}{\prod_{i=1}^{S_n} \nabla_{S_i}^{D_i} \prod_{i=1}^{S_n} a^{S_i}(B)} e_t. \tag{7}$$

Предложенные в работе методы являются некоторым промежуточным подходом между классическими регрессионными и современными нейросетевыми и более формализованы по выбору структуры и экономны по временным затратам, при этом их модели являются квазиоптимальными по детализации.

## Литература

1. Щелкалин В. Н. Математические модели и методы, основанные на совместном использовании идей методов «Гусеница»-SSA и Бокса-Дженкинса // Труды IX Международной конференции «Идентификация систем и задачи управления», SICPRO'12, Москва, 2012. С. 1035 – 1048.

**Иллюстрации**

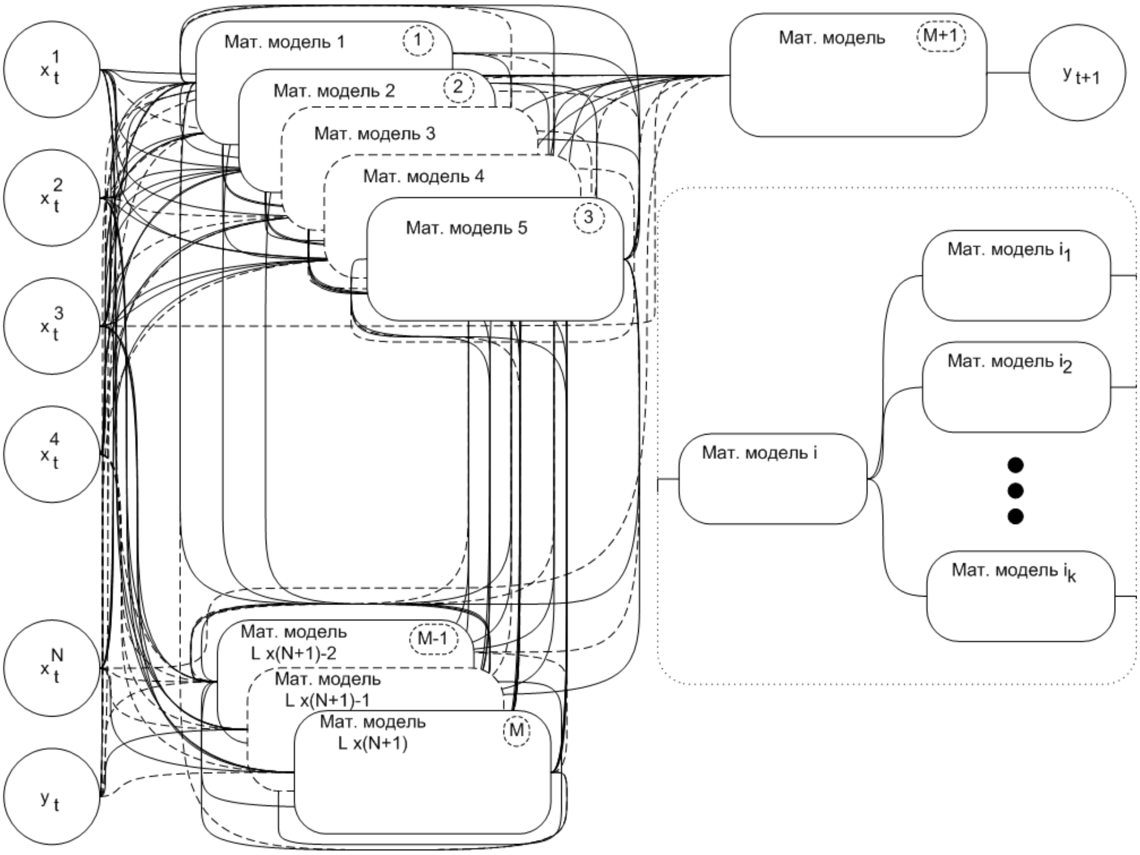


Рис. 1: Структурная схема модели комбинированного метода прогнозирования