

УДК 519.246.8+519.254+519.237.7

Варианты метода «Гусеница»-SSA для анализа многомерных временных рядов

Н.Э. Голяндина

*Санкт-Петербургский государственный университет,
Математико-механический факультет*
Россия, 198504, Санкт-Петербург, Петродворец, Ульяновский пр., 28
E-mail: nina@ng1174.spb.edu

В.В. Некруткин

*Санкт-Петербургский государственный университет,
Математико-механический факультет*
Россия, 198504, Санкт-Петербург, Петродворец, Ульяновский пр., 28
E-mail: nekr@vn1068.spb.edu

Д.В. Степанов

*Санкт-Петербургский государственный университет,
Математико-механический факультет*
Россия, 198504, Санкт-Петербург, Петродворец, Ульяновский пр., 28
E-mail: dmitrys_2000@mail.ru

Ключевые слова: разложение временных рядов на аддитивные компоненты, многомерные временные ряды, делимость временных рядов, анализ сингулярного спектра

Key words: time series decomposition into additive components, multidimensional time series, time series separability, Singular Spectrum Analysis

В данной работе исследуются вопросы делимости аддитивных компонент двумерных рядов с помощью двух вариантов метода SSA (Singular Spectrum Analysis): многомерного MSSA и комплексного CSSA. На модельных и реальных примерах проведено сравнение этих методов между собой, а также с результатами применения SSA к одномерным рядам по отдельности. Дополнительную информацию о методе и соответствующем программном обеспечении можно найти на сайте по адресу <http://www.gistatgroup.com/gus/>

VARIANTS OF THE “CATERPILLAR”-SSA METHOD FOR ANALYSIS OF MULTIDIMENSIONAL TIME SERIES / N.E. Golyandina (Mathematical Department, St.Petersburg State University, 28 Ulyanovskij pr., St.Petersburg Petrodvorets 198504, Russia, E-mail: nina@ng1174.spb.edu), V.V.Nekrutkin (Mathematical Department, St.Petersburg State University, 28 Ulyanovskij pr., St.Petersburg Petrodvorets 198504, Russia, E-mail: nekr@vn1068.spb.edu), D.V. Stepanov (Mathematical Department, St.Petersburg State University, 28 Ulyanovskij pr., St.Petersburg Petrodvorets 198504, Russia, E-mail: dmitrys_2000@mail.ru). Two variants of Singular Spectrum Analysis (briefly, SSA) method, multidimensional MSSA and complex CSSA, are compared with each other for two-dimensional time series. Analogously to the standard SSA, the main object of the investigation is the separability of additive components of a series. Several model and real-life examples serve as an illustration of theoretical results. Additional information on the SSA method, its abilities and the corresponding software can be found on the site <http://www.gistatgroup.com/cat/>

1. Введение

Одним из мощных методов анализа временных рядов является метод «Гусеница» (называемый SSA – Singular Spectrum Analysis – в зарубежной литературе). Хотя первыми публикациями, относящимися к этому методу, считаются статьи [1] и [2], его идеи неоднократно и независимо появлялись (и появляются до сих пор) в различных областях знаний, связанных с обработкой временных рядов. Один из обзоров публикаций на эту тему можно найти в сборнике [3].

Наиболее исследованным и апробированным является одномерный метод SSA, предназначенный для анализа одного вещественнозначного временного ряда. Именно этой тематике (в основном) посвящены обе имеющиеся в настоящее время монографии, затрагивающие SSA: элементарное введение в метод [4] и содержащая его теоретическое обоснование книга [5].

Для одномерного ряда базовый метод SSA состоит в преобразовании исходного ряда в многомерный, сингулярном разложении получившейся траекторной матрицы, группировке членов разложения и последующем восстановлении (см. раздел 2 настоящей работы). При этом часто оказывается возможным выделить отдельные аддитивные составляющие исходного ряда, такие как тренд (гладкая и медленно меняющаяся часть ряда), различные колебательные и периодические компоненты, а также шумовую компоненту.

Решающими для объяснения успеха (или неуспеха) применения метода к рядам различной природы, а также для выбора его параметров, являются понятия слабой и сильной делимости рядов (см. основные понятия и определения в разделе 3, подробное описание, проиллюстрированное на множестве реальных и модельных примеров можно найти в [5]). В этом смысле теорию одномерного метода SSA можно считать в основном законченной.

Имеются, однако, статьи (см., например, [6] или [7]), где идеи метода одномерного SSA распространяются на случай нескольких временных рядов. В этом случае ожидаемым результатом является одновременное разложение нескольких рядов на согласованные компоненты. Но достаточно полной теории для многомерного SSA (MSSA – Multi-channel SSA) не существует. Настоящая работа частично восполняет этот пробел.

При этом мы ограничиваемся исследованием следующих вопросов. Во-первых, в основном рассматриваются только двумерные временные ряды. Это позволяет избежать излишней громоздкости и дает возможность получать результаты в замкнутой форме. Кроме того, здесь мы затрагиваем одну из проблем, связанных с MSSA – проблему построения траекторной матрицы многомерного ряда. Дело в том, что структура траекторных матриц, рассматриваемых в [6] и [7], выглядит естественно, но не ясно, насколько она является канонической. Более того, сложно описать класс альтернативных способов построения траекторных матриц для многомерного ряда. Исключение составляют двумерные ряды, где альтернативой является комплексный SSA (CSSA – Complex SSA), обладающий многими свойствами одномерного (и, тем самым, очевидной «каноничностью»). Коротко такое расширение упоминается в [8]. Таким образом, сравниваются три варианта SSA для анализа двумерных рядов: последовательное применение одномерного SSA к обоим компонентам векторного, двумерный SSA и комплексный SSA.

Второе ограничение связано с делимостью. Как будет видно, понятие делимости без труда обобщается как на MSSA, так и на CSSA, хотя и

приводит к несколько различным результатам (собственно говоря, исследованию этих различий и посвящена настоящая работа). В то же время при работе с реальными рядами делимость проявляет себя двумя способами: как приближенная делимость и как асимптотическая делимость (следует отметить, что для «длинных» рядов часто удается совместить эти проявления). Проиллюстрируем эти различия для одномерного метода SSA.

На практике асимптотическая делимость «работает» для достаточно длинных рядов, и, соответствующим образом, основной параметр L («длина гусеницы» или «длина окна»), определяющий размеры траекторной матрицы, должен быть примерно равным половине длины ряда. Например, можно показать (см. [5]), что при этом выборе L любой сигнал асимптотически отделяется от гауссовского белого шума. Другой теоретический результат, также подтверждаемый практикой, говорит, что любые две гармоники с различными частотами (и амплитудами) также асимптотически делимы.

Приближенная делимость (для относительно коротких рядов) как правило, связана с разделением составляющих слабо зашумленных периодик. В этом случае нужно стараться, чтобы обе размерности траекторной матрицы (то есть длина окна L и $K = N - L + 1$, где N – длина ряда) были пропорциональны периоду T исследуемой последовательности. Это является сильным, но с практической точки зрения несколько формальным ограничением, поскольку, как уже говорилось, приближенная делимость часто сочетается с асимптотической.

В настоящей работе мы рассматриваем только точную (неасимптотическую) делимость, которая на практике проявляется как приближенная. Тем самым, определяется набор тестовых рядов для сравнения MSSA с CSSA: мы рассматриваем условия точной делимости либо (двумерных) гармоник, либо постоянных рядов (имитирующих тренды, близкие к константе) и гармоник. Другие интересные примеры, связанные, скажем, с делимостью существенно непостоянных трендов от годовых (может быть, модулированных) осцилляций, решаются в рамках SSA только асимптотически.

Наконец, мы не рассматриваем варианты SSA, отличные от базового (такие как SSA с однократным и двукратным центрированием или Теплицев SSA), а также методы прогноза и нахождения моментов разладки многомерных рядов, основанные на идеях SSA. Для одномерного случая этот материал можно найти в [5].

Настоящая работа состоит из четырех разделов, включая введение. Во втором из них излагается алгоритмическая структура метода SSA в одномерном, двумерном и комплексном варианте. Третий раздел посвящен понятию делимости и его особенностям. Там же обсуждаются сходство и различие вариантов SSA, в том числе и на модельных примерах. Оказывается, что ни один из вариантов SSA не является абсолютно лучшим по сравнению с другими. В то же время, например, комплексный SSA обладает рядом привлекательных особенностей, которые позволяют считать его альтернативным к двумерному.

Последний раздел содержит обработки нескольких реальных примеров и имеет иллюстративный характер.

Доказательства всех утверждений опущены. Они вполне элементарны, но, как правило, достаточно громоздки.

2. Описание базового алгоритма

В этом параграфе описаны схемы одномерного метода SSA, его комплексной формы CSSA, а также многомерной модификации MSSA. Схема MSSA будет подробно описана для двумерных временных рядов с комментариями о способе обобщения на их произвольное число.

Пусть имеются два ряда $F = (f_0, \dots, f_{N-1})$ и $G = (g_0, \dots, g_{N-1})$ длины $N > 2$. Их можно рассматривать как двумерный временной ряд (F, G) или как комплексный ряд $C = (c_0, \dots, c_{N-1})$, где $c_k = f_k + ig_k$, $0 \leq k < N$. Для анализа временных рядов выберем целый параметр L , $1 < L < N$. Его называют «длина гусеницы» или «длина окна». Схему одномерного метода рассмотрим для ряда F . Далее будем предполагать, что ряды ненулевые.

Базовый алгоритм состоит из четырех шагов. Первые два шага образуют этап разложения, а последние два – этап восстановления.

2.1. Шаг 1: Вложение

Одномерный метод.

Процедура вложения есть преобразование исходного одномерного ряда в последовательность L -мерных векторов, число которых равно $K = N - L + 1$:

$$X_i = (f_{i-1}, \dots, f_{i+L-2})^T, 1 \leq i \leq K.$$

Эти вектора образуют *траекторную матрицу* $\mathbf{X} = [X_1 : \dots : X_K]$ ряда F . Записывая матрицу более подробно, заметим, что $x_{ij} = f_{i+j-2}$, т. е. матрица \mathbf{X} имеет одинаковые элементы на диагонали $i + j = \text{const}$:

$$\mathbf{X} = \begin{pmatrix} f_0 & f_1 & f_2 & \cdots & f_{K-1} \\ f_1 & f_2 & f_3 & \cdots & f_K \\ f_2 & f_3 & f_4 & \cdots & f_{K+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_{L-1} & f_L & f_{L+1} & \cdots & f_{N-1} \end{pmatrix}.$$

Комплексный метод.

Траекторная матрица строится абсолютно аналогично, поэтому метод можно считать естественным обобщением одномерного:

$$\mathbf{X} = \begin{pmatrix} c_0 & c_1 & c_2 & \cdots & c_{K-1} \\ c_1 & c_2 & c_3 & \cdots & c_K \\ c_2 & c_3 & c_4 & \cdots & c_{K+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{L-1} & c_L & c_{L+1} & \cdots & c_{N-1} \end{pmatrix} = \mathbf{F} + i\mathbf{G},$$

где \mathbf{F} и \mathbf{G} – траекторные матрицы вещественных рядов F и G соответственно. Размерность траекторной матрицы \mathbf{X} равна $L \times K$.

Двумерный метод.

Построим последовательности L -мерных векторов для рядов F и G :

$$F_i = (f_{i-1}, \dots, f_{i+L-2})^T, G_i = (g_{i-1}, \dots, g_{i+L-2})^T, 1 \leq i \leq K.$$

В каждой из последовательностей число векторов равно $K = N - L + 1$. *Траекторной матрицей* двумерного ряда, порожденной длиной окна L , назовем матрицу

$$\mathbf{X} = [F_1 : \dots : F_K : G_1 : \dots : G_K] = [\mathbf{F}, \mathbf{G}].$$

Другими словами, траекторная матрица двумерного ряда представляет собой расположенные последовательно траекторные матрицы \mathbf{F} и \mathbf{G} одномерных рядов F и G , полученные при одинаковой длине окна L . Таким образом, размерность получившейся матрицы равна $L \times 2K$.

Замечание 2.1. Ясно, как изменится первый шаг метода при произвольном числе рассматриваемых рядов произвольной (не обязательно одинаковой) длины. Будем считать, что наблюдается система из S временных рядов $F^{(k)} = (f_i^{(k)})_{i=0}^{N_k-1}$, где $k=1, \dots, S$. Параметр N_k , таким образом, есть длина k -го ряда. Выбираем L такое, что $1 < L < N_k$ для любого k . Для каждого k вычислим $K_k = N_k - L + 1$ векторов

$$X_i^{(k)} = (f_{i-1}^{(k)}, \dots, f_{i+L-2}^{(k)})^T, \quad 1 \leq i \leq K_k.$$

Тогда траекторная матрица многомерного ряда $(F^{(1)}, \dots, F^{(S)})$ будет иметь вид

$$\mathbf{X} = [X_1^{(1)} : \dots : X_{K_1}^{(1)} : \dots : X_1^{(S)} : \dots : X_{K_S}^{(S)}] = [\mathbf{X}^{(1)} : \dots : \mathbf{X}^{(S)}],$$

где $\mathbf{X}^{(i)}$ – траекторная матрица ряда $F^{(i)}$, соответствующая длине окна L .

Размерность матрицы \mathbf{X} равна $L \times \sum_{k=1}^S K_k$.

Замечание 2.2. При фиксированных длинах рядов и длине окна L траекторная матрица однозначно определяет ряды, по которым она построена.

2.2. Шаг 2: Сингулярное разложение

Результатом этого шага является сингулярное разложение траекторной матрицы ряда.

Одномерный и двумерный методы.

Рассмотрим матрицу $\mathbf{S} = \mathbf{X}\mathbf{X}^T$, где \mathbf{X} – траекторная матрица одномерного или двумерного вещественного временного ряда, соответствующая длине окна L .

Поскольку \mathbf{S} положительно полуопределена, ее собственные числа неотрицательны. Обозначим через $\lambda_1, \dots, \lambda_L$ собственные числа матрицы \mathbf{S} , взятые в порядке убывания ($\lambda_1 \geq \dots \geq \lambda_L \geq 0$) и через U_1, \dots, U_L ортонормированную систему собственных векторов матрицы \mathbf{S} , соответствующих этим собственным числам.

Пусть $d = \max\{i, \text{таких, что } \lambda_i > 0\}$. Обозначив $V_i = \frac{1}{\sqrt{\lambda_i}} \mathbf{X}^T U_i$

($i=1, \dots, d$), получим разложение траекторной матрицы:

$$(1) \quad \mathbf{X} = \mathbf{X}_1 + \dots + \mathbf{X}_d, \text{ где } \mathbf{X}_i = \sqrt{\lambda_i} U_i V_i^T.$$

Отметим, что ортонормированные вектора V_i являются собственными векторами матрицы $\mathbf{X}^T \mathbf{X}$, соответствующими тем же собственным числам λ_i .

В стандартной терминологии $\sqrt{\lambda_i}$ называются *сингулярными числами*, U_i и V_i – *левыми* и *правыми сингулярными векторами* матрицы \mathbf{X} соответственно. Каждая из матриц \mathbf{X}_i имеет ранг 1. Поэтому их можно назвать *элементарными матрицами*. Набор $(\sqrt{\lambda_i}, U_i, V_i)$ мы будем называть *i -й собственной тройкой* сингулярного разложения.

На геометрическом языке система собственных векторов U_i задает ортонормированный базис в линейном пространстве, порождаемом столбцами исходной матрицы \mathbf{X} . Аналогично, вектора V_i задают ортонормированный базис линейного пространства, порождаемого строками матрицы \mathbf{X} .

В другой (статистической) терминологии вектора U_i называются *собственными*, вектора V_i – *факторными*, направление, задаваемое i -м собственным вектором U_i – i -м *главным направлением*, вектор $Z_i = \sqrt{\lambda_i} V_i$, составленный из проекций векторов X_i на i -е главное направление – *вектором i -х главных компонент*.

Прокомментируем результат этапа разложения для двумерного случая. При таком подходе $\mathbf{S} = \mathbf{X}\mathbf{X}^T = \mathbf{F}\mathbf{F}^T + \mathbf{G}\mathbf{G}^T$,

$$V_k = \frac{1}{\sqrt{\lambda_k}} \begin{pmatrix} \mathbf{F}^T \\ \mathbf{G}^T \end{pmatrix} U_k = \frac{1}{\sqrt{\lambda_k}} \begin{pmatrix} \mathbf{F}^T U_k \\ \mathbf{G}^T U_k \end{pmatrix}$$

и

$$\mathbf{X} = \sum_{k=1}^d U_k (U_k^T \mathbf{F}, U_k^T \mathbf{G}) = \sum_{k=1}^d [U_k U_k^T \mathbf{F} : U_k U_k^T \mathbf{G}].$$

Таким образом, для каждого из рядов получено разложение столбцов их траекторных матриц по общему базису (U_1, \dots, U_d) :

$$(2) \quad \mathbf{F} = \sum_{k=1}^d U_k U_k^T \mathbf{F}, \quad \mathbf{G} = \sum_{k=1}^d U_k U_k^T \mathbf{G}.$$

Эти разложения, однако, не обязаны быть сингулярными разложениями матриц \mathbf{F} и \mathbf{G} .

Комплексный метод.

В этом случае $\mathbf{S} = \mathbf{X}\mathbf{X}^*$, где \mathbf{X}^* – эрмитово-сопряженная к \mathbf{X} матрица,

$$\mathbf{S} = (\mathbf{F} + i\mathbf{G})(\mathbf{F}^T - i\mathbf{G}^T) = (\mathbf{F}\mathbf{F}^T + \mathbf{G}\mathbf{G}^T) + i(\mathbf{G}\mathbf{F}^T - \mathbf{F}\mathbf{G}^T).$$

Заметим, что \mathbf{S} является эрмитовой и имеет неотрицательные собственные числа $\lambda_1 \geq \dots \geq \lambda_d \geq 0$. Снова обозначим d число ненулевых собственных чисел матрицы \mathbf{S} . Пусть U_1, \dots, U_d – (комплекснозначные) собственные вектора, соответствующие собственным числам $\lambda_1, \dots, \lambda_d$, причем $U_j U_k^* = \delta_{jk}$, где δ_{jk} – символ Кронекера.

Тогда, если при $k=1, \dots, d$ положить $V_k = \mathbf{X}^* U_k / \sqrt{\lambda_k}$, то сингулярным разложением матрицы \mathbf{X} является разложение

$$(3) \quad \mathbf{X} = \mathbf{X}_1 + \dots + \mathbf{X}_d, \text{ где } \mathbf{X}_k = \sqrt{\lambda_k} U_k V_k^*.$$

Как и в одномерном случае, V_k является собственным вектором матрицы $\mathbf{X}^* \mathbf{X}$, соответствующим собственному числу λ_k , причем $V_j V_k^* = \delta_{jk}$.

Обозначим $U_k = P_k + iQ_k$, тогда

$$V_k = \frac{1}{\sqrt{\lambda_k}} (\mathbf{F}^T - i\mathbf{G}^T) U_k = \frac{1}{\sqrt{\lambda_k}} \left((\mathbf{F}^T P_k + \mathbf{G}^T Q_k) + i(\mathbf{F}^T Q_k - \mathbf{G}^T P_k) \right)$$

и

$$(4) \quad \mathbf{X} = \sum_{k=1}^d (P_k + iQ_k) \left((P_k^T \mathbf{F} + Q_k^T \mathbf{G}) - i(Q_k^T \mathbf{F} - P_k^T \mathbf{G}) \right).$$

Разделяя вещественные и мнимые слагаемые в (4), получаем представления

$$(5) \quad \mathbf{F} = \sum_{k=1}^d (P_k P_k^T + Q_k Q_k^T) \mathbf{F} + (P_k Q_k^T - Q_k P_k^T) \mathbf{G},$$

$$(6) \quad \mathbf{G} = \sum_{k=1}^d (P_k P_k^T + Q_k Q_k^T) \mathbf{G} + (Q_k P_k^T - P_k Q_k^T) \mathbf{F}.$$

Замечание 2.3. Сингулярные разложения (1) и (3) единственны тогда и только тогда, когда все собственные числа λ_k различны. Это очевидное утверждение является, тем не менее, существенным для SSA.

Замечание 2.4. Сравнение (2) с (5) и (6) показывает, что для двумерного варианта разложение каждого из рядов зависит только от траекторной матрицы этого конкретного ряда и собственных векторов, полученных на основе траекторной матрицы двумерного ряда. Если же используется комплексный подход, то на разложение каждого ряда траекторная матрица другого влияет непосредственно. Конечно, если мы производим сингулярные разложения траекторных матриц рядов F и G по отдельности, то они являются «независимыми».

Замечание 2.5. Если рассматриваемые ряды совпадают ($F=G$), то в комплексном варианте метода $\mathbf{S} = \mathbf{F}\mathbf{F}^T + \mathbf{G}\mathbf{G}^T = 2\mathbf{F}\mathbf{F}^T$, что совпадает с матрицей \mathbf{S} двумерного варианта. В этом случае у каждого из собственных векторов $U_k = P_k + iQ_k$ либо $P_k = 0$, либо $Q_k = 0$. А это означает, что разложения (2) и (5), (6), полученные этими двумя методами, будут идентичными. Поскольку при удвоении всех элементов матрицы \mathbf{S} ее ортонормированная система собственных векторов не изменяется, то это разложение будет совпадать также с последовательным одномерным сингулярным разложением каждого из рядов по отдельности. Конечно, здесь подразумевается, что длина окна L во всех разложениях одна и та же.

Замечание 2.6. Вид левых и правых сингулярных векторов, трактуемых в SSA как временные ряды, является очень важным для следующего шага метода – группировки (см., например, [3] или [5]). При этом для одномерного SSA левые и правые сингулярные вектора обладают определенной симметрией, так как в этих случаях сингулярные разложения траекторных матриц с длиной окна L и $K = N - L + 1$ эквивалентны. Комплексный SSA (с точностью до того, что вместо транспонирования здесь рассматривается эрмитово сопряжение) также обладает подобным свойством.

Для двумерного SSA такая симметрия отсутствует, так как столбцы и строки траекторной матрицы имеют здесь различный смысл: в то время как каждый столбец траекторной матрицы \mathbf{X} – это отрезок длины L одного из рядов F и G , строка матрицы \mathbf{X} состоит из «пристыкованных» отрезков длины K рядов F и G .

Сингулярные вектора наследуют эту структуру: если левый сингулярный вектор (собственный вектор) по-прежнему трактуется как (одномерный) временной ряд длины L , то правый сингулярный вектор (факторный вектор) скорее следует трактовать как два одномерных ряда длины K , первый из которых относится к ряду F , а второй – к ряду G .

Замечание 2.7. Даже если собственное число λ комплексной эрмитовой матрицы простое, то существует множество нормированных на единицу собственных векторов этой матрицы, соответствующих λ . Действительно, если X – собственный вектор такой матрицы, то вектор cX , где c – произвольное комплексное число, по модулю равно 1 , тоже является собственным и имеет ту же длину. Поэтому левые и правые сингулярные вектора U_k и V_k в

комплексном случае не имеют смысла по отдельности. Их свойства проявляются только в виде элементарной матрицы $\mathbf{X}_k = \sqrt{\lambda_k} U_k V_k^*$, которая определяется однозначно, если λ_k – простое собственное число матрицы \mathbf{S} .

Замечание 2.8. Нахождение собственных чисел и соответствующей системы ортогональных собственных векторов комплексной эрмитовой матрицы $\mathbf{S} = \mathbf{A} + i\mathbf{B}$ размера $L \times L$ сводится к такой же задаче для действительной матрицы

$$\mathbf{D} = \begin{pmatrix} \mathbf{A} & -\mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix},$$

которая имеет размер $2L \times 2L$. Эта матрица является симметрической, так как $(-\mathbf{B}) = \mathbf{B}^T$. Каждому действительному собственному значению λ и соответствующему собственному вектору $X = Y + iZ$ исходной матрицы \mathbf{D} соответствуют собственное число λ кратности два и пара собственных векторов $(Y, Z)^T, (-Z, Y)^T$ действительной матрицы \mathbf{D} . В самом деле, если

$$(\mathbf{A} + i\mathbf{B})(Y + iZ) = \mathbf{A}Y - \mathbf{B}Z + i(\mathbf{A}Z + \mathbf{B}Y) = \lambda(Y + iZ),$$

то

$$\begin{pmatrix} \mathbf{A} & -\mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix} \begin{pmatrix} Y \\ Z \end{pmatrix} = \begin{pmatrix} \mathbf{A}Y - \mathbf{B}Z \\ \mathbf{B}Y + \mathbf{A}Z \end{pmatrix} = \lambda \begin{pmatrix} Y \\ Z \end{pmatrix},$$

$$\begin{pmatrix} \mathbf{A} & -\mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix} \begin{pmatrix} -Z \\ Y \end{pmatrix} = \begin{pmatrix} -\mathbf{A}Z - \mathbf{B}Y \\ -\mathbf{B}Z + \mathbf{A}Y \end{pmatrix} = \lambda \begin{pmatrix} -Z \\ Y \end{pmatrix}.$$

Верно и обратное. При этом легко видеть, что из ортогональности вещественных собственных векторов следует ортогональность комплексных и наоборот. Неоднозначность определения собственных векторов Y и Z матрицы \mathbf{D} соответствует множественности нормированных собственных векторов, соответствующих собственному числу λ эрмитовой матрицы \mathbf{S} .

2.3. Шаг 3. Группировка

Процедура группировки формально одинакова для всех рассматриваемых разновидностей SSA.

На основе разложения (1) процедура группировки делит все множество индексов $\{1, \dots, d\}$ на m непересекающихся подмножеств I_1, \dots, I_m .

Пусть $I = \{i_1, \dots, i_p\}$. Тогда *результатирующая* матрица \mathbf{X}_I , соответствующая группе I , определяется как $\mathbf{X}_I = \mathbf{X}_{i_1} + \dots + \mathbf{X}_{i_p}$.

Такие матрицы вычисляются для $I = I_1, \dots, I_m$, тем самым разложение (1) может быть записано в сгруппированном виде:

$$(7) \quad \mathbf{X} = \mathbf{X}_{I_1} + \dots + \mathbf{X}_{I_m}.$$

Процедура выбора множеств I_1, \dots, I_m и называется *группировкой собственных троек*.

2.4. Шаг 4. Диагональное усреднение

Одномерный и комплексные методы.

На последнем шаге базового алгоритма каждая матрица сгруппированного разложения переводится в новый ряд длины N .

Пусть \mathbf{Y} – матрица размера $L \times K$ с элементами y_{ij} , $1 \leq i \leq L$, $1 \leq j \leq K$. Положим $L^* = \min(L, K)$, $K^* = \max(L, K)$ и $N = L + K - 1$. Пусть $z_{ij} = y_{ij}$, если $L < K$ и $z_{ij} = y_{ji}$ в остальных случаях.

Диагональное усреднение переводит матрицу \mathbf{Y} в ряд (g_0, \dots, g_{N-1}) по формуле

$$(8) \quad g_k = \begin{cases} \frac{1}{k+1} \sum_{j=1}^{k+1} z_{j,k-j+2} & \text{для } 0 \leq k < L^* - 1, \\ \frac{1}{L^*} \sum_{j=1}^{L^*} z_{j,k-j+2} & \text{для } L^* - 1 \leq k < K^*, \\ \frac{1}{N-k} \sum_{j=k-K^*+2}^{N-K^*+1} z_{j,k-j+2} & \text{для } K^* \leq k < N. \end{cases}$$

Это выражение соответствует усреднению элементов матрицы вдоль «диагоналей» $i+j=k+2$: выбор $k=0$ дает $g_0 = y_{11}$, для $k=1$ получаем $g_1 = (y_{12} + y_{21})/2$ и т.д. Применяв диагональное усреднение к матрицам \mathbf{X}_{I_k} , полученным на этапе группировки, приходим к разложению исходного ряда в сумму m рядов.

Такая процедура обладает свойством оптимальности: траекторная матрица $\tilde{\mathbf{Y}}$ ряда, полученного из некоторой матрицы \mathbf{Y} с помощью диагонального усреднения, является ближайшей к \mathbf{Y} по матричной норме среди всевозможных ганкелевых (в нашей терминологии, траекторных) матриц той же размерности. При этом, если исходная матрица \mathbf{Y} является траекторной для некоторого временного ряда F , то полученный в результате диагонального усреднения ряд G совпадает с F .

Отметим, что матричная норма вещественной матрицы равняется корню квадратному из суммы квадратов всех элементов матрицы. Для комплексной матрицы определение аналогично.

Двумерный метод.

Каждая матрица \mathbf{X}_{I_k} размера $L \times 2K$ в сгруппированном разложении (7) разбивается на последовательно расположенные матрицы размера $L \times K$: $\mathbf{X}_{I_k} = [\mathbf{X}_{I_k}^{(1)} : \mathbf{X}_{I_k}^{(2)}]$. После этого для каждой из матриц $\mathbf{X}_{I_k}^{(j)}$, $j=1,2$, производится диагональное усреднение по формуле (8). В результате каждое слагаемое в правой части (7) порождает двумерный ряд – аддитивную компоненту ряда (F, G) .

Таким образом, результатом применения базового алгоритма SSA к (одномерному, многомерному или комплексному) ряду является его представление в виде суммы m рядов. Параметрами метода являются длина окна L и способ группировки элементарных матриц.

2.5. Модельные примеры: сингулярные разложения траекторных матриц

В этом разделе мы опишем собственные числа и пространства $\mathbf{\Lambda}^{(L)}$, $\mathbf{\Lambda}^{(K)}$, натянутые на левые и правые сингулярные вектора U_1, \dots, U_d и V_1, \dots, V_d соответственно для двух простейших модельных рядов и разных вариантов метода SSA. Эти результаты понадобятся нам при изучении делимости конкретных временных рядов.

2.5.1. Константные ряды. Рассмотрим два постоянных вещественных ряда $P^{(1)} = (p_0^{(1)}, \dots, p_{N-1}^{(1)})$ и $P^{(2)} = (p_0^{(2)}, \dots, p_{N-1}^{(2)})$, где $p_k^{(1)} = p_1$, $p_k^{(2)} = p_2$ и $p_1 p_2 \neq 0$.

Пусть L – длина окна и $K = N - L + 1$. Нас интересуют сингулярные разложения траекторных матриц рядов $P^{(1)}$, $P^{(1)} + iP^{(2)}$ и $(P^{(1)}, P^{(2)})$.

Предложение 2.1. 1) Порядок d сингулярного разложения траекторной матрицы постоянного ряда равен 1 в одномерном, комплексном и двумерном случаях. Во всех трех случаях пространство $\Lambda^{(L)}$ натянуто на L -мерный вектор $(1, \dots, 1)^T$.

2) Для одномерного ряда $P^{(1)}$ собственное число λ равно $p_1^2 LK$, а пространство $\Lambda^{(K)}$ натянуто на K -мерный вектор $(1, \dots, 1)^T$.

3) Для комплексного ряда $P^{(1)} + iP^{(2)}$ собственное число λ равно $(p_1^2 + p_2^2)LK$, а пространство $\Lambda^{(K)}$ натянуто на K -мерный вектор $(1, \dots, 1)^T$.

4) Для двумерного ряда $(P^{(1)}, P^{(2)})$ собственное число λ также равно $(p_1^2 + p_2^2)LK$, а пространство $\Lambda^{(K)}$ натянуто на $2K$ -мерный вектор $(v_1, \dots, v_K, w_1, \dots, w_K)^T$, где $v_j = p_1$ и $w_j = p_2$.

2.5.2. Гармонические ряды. Рассмотрим ряды $S^{(1)} = (s_0^{(1)}, \dots, s_{N-1}^{(1)})$ и $S^{(2)} = (s_0^{(2)}, \dots, s_{N-1}^{(2)})$ с

$$s_k^{(1)} = A \cos(2\pi\omega k + \varphi_1), \quad s_k^{(2)} = B \cos(2\pi\omega k + \varphi_2),$$

где $0 < \omega < 1/2$, $0 \leq \varphi_1, \varphi_2 < 2\pi$ и $A, B \neq 0$. Как и раньше, мы изучаем сингулярные разложения траекторных матриц рядов $S^{(1)}$, $S^{(1)} + iS^{(2)}$ и $(S^{(1)}, S^{(2)})$. Для краткости такие ряды мы будем называть гармоническими.

Зафиксируем длину окна L , положим $K = N - L + 1$ и рассмотрим сингулярные разложения траекторных матриц гармонических рядов во всех трех вариантах. Опишем сначала пространства $\Lambda^{(L)}$ и $\Lambda^{(K)}$, натянутые на левые и правые сингулярные вектора соответственно.

Предложение 2.2. 1) Порядок d сингулярного разложения траекторных матриц одномерного и двумерного гармонических рядов равен 2. Для комплексных гармонических рядов $d = 1$, если $A = B$ и $|\varphi_1 - \varphi_2| = \pi/2 \pmod{\pi}$, в остальных случаях $d = 2$.

2) Если $d = 2$, то для всех трех вариантов гармонического ряда пространство $\Lambda^{(L)}$ натянуто на вектора

$$\begin{aligned} & (1, \cos(2\pi\omega), \dots, \cos(2\pi(L-1)\omega))^T, \\ & (0, \sin(2\pi\omega), \dots, \sin(2\pi(L-1)\omega))^T. \end{aligned}$$

3) Для одномерного гармонического ряда и для комплексного гармонического ряда в случае $d = 2$ пространство $\Lambda^{(K)}$ натянуто на вектора

$$\begin{aligned} & (1, \cos(2\pi\omega), \dots, \cos(2\pi(K-1)\omega))^T, \\ & (0, \sin(2\pi\omega), \dots, \sin(2\pi(K-1)\omega))^T. \end{aligned}$$

4) Для двумерного гармонического ряда пространство $\Lambda^{(K)}$ натянуто на $2K$ -мерные вектора

$$(c_1^{(1)}, \dots, c_K^{(1)}, c_1^{(2)}, \dots, c_K^{(2)})^T, (s_1^{(1)}, \dots, s_K^{(1)}, s_1^{(2)}, \dots, s_K^{(2)})^T,$$

где

$$\begin{aligned} c_j^{(1)} &= A \cos(2\pi(j-1)\omega + \varphi_1), \quad c_j^{(2)} = B \cos(2\pi(j-1)\omega + \varphi_2), \\ s_j^{(1)} &= A \sin(2\pi(j-1)\omega + \varphi_1) \quad \text{и} \quad s_j^{(2)} = B \sin(2\pi(j-1)\omega + \varphi_2). \end{aligned}$$

5) Для комплексного ряда в случае $A = B$ и $|\varphi - \varphi_2| = \pi/2 \pmod{\pi}$ одномерные пространства $\Lambda^{(L)}$ и $\Lambda^{(K)}$ натянуты соответственно на вектора $(1, a, \dots, a^{L-1})^T$ и $(1, a, \dots, a^{K-1})^T$, где $a = e^{i2\pi\omega}$.

Изучим более подробно случай, когда числа $L\omega$ и $K\omega$ – целые. Эта ситуация, как будет видно в дальнейшем, характерна для делимости с участием гармонических рядов. Нас будут в основном интересовать собственные числа сингулярных разложений траекторных матриц (пространства $\Lambda^{(L)}$ и $\Lambda^{(K)}$) уже описаны в предложении 2.2).

Предложение 2.3. Пусть числа $L\omega$ и $K\omega$ – целые. Тогда собственные числа сингулярных разложений траекторных матриц гармонических рядов имеют следующий вид.

1) Для одномерной гармонике $S^{(1)}$ собственные числа одинаковы: $\lambda_1 = \lambda_2 = A^2 L K / 4$.

2) Для двумерной гармонике $(S^{(1)}, S^{(2)})$ собственные числа одинаковы: $\lambda_1 = \lambda_2 = (A^2 + B^2) L K / 4$.

3) Для комплексной гармонике $S^{(1)} + iS^{(2)}$ в случае $d = 2$ собственные числа имеют вид $C^2 L K$ и $D^2 L K$, где

$$(9) \quad C^2 = (A^2 + B^2 - 2AB \sin \varphi) / 4, \quad D^2 = (A^2 + B^2 + 2AB \sin \varphi) / 4$$

и $\varphi = \varphi_2 - \varphi_1$.

4) Для комплексной гармонике $S^{(1)} + iS^{(2)}$, если $A = B$ и $|\varphi - \varphi_2| = \pi/2 \pmod{\pi}$, то $\lambda = A^2 L K$.

Замечание 2.9. Результаты предложений 2.2 и 2.3 показывают существенные различия между вещественными и комплексным вариантом SSA для гармонических рядов. Во-первых, если ряды обладают одинаковыми амплитудами и сдвинуты на $\pi/2$ друг относительно друга, то в комплексном случае мы приходим к одномерному разложению, в то время как вещественные SSA всегда дают двумерные разложения. Во-вторых (это более важно для практики), собственные числа комплексного SSA, как правило, различны (при целых $L\omega$ и $K\omega$ они одинаковы только если ряды софазные или противофазные), в то время как при тех же условиях на L , K и ω вещественные варианты SSA дают одинаковые собственные числа. При этом максимальное отличие между собственными числами достигается при $|\varphi - \varphi_2| \approx \pi/2$ или при $|\varphi - \varphi_2| \approx 3\pi/2$. В то же время предложение 2.2 показывает, что во всех случаях гармонические ряды вполне распознаваемы по сингулярным векторам разложений своих траекторных матриц.

Как указывалось в разделе 2, итогом работы алгоритма SSA является разложение временного ряда на аддитивные составляющие. Покажем, как выглядят эти составляющие для комплексного SSA и гармонического ряда $F + iG$.

Предложение 2.4. Пусть $L\omega$ и $K\omega$ – целые. Тогда SSA-разложение комплексного ряда $F + iG$ на аддитивные слагаемые имеет вид

$$s_k^{(1)} + i s_k^{(2)} = C e^{i(2\pi\omega k + \phi_1)} + D e^{-i(2\pi\omega k + \phi_2)}, \quad k = 0, \dots, N-1,$$

где C и D – вещественные постоянные, удовлетворяющие (9). При этом фазы ϕ_1 и ϕ_2 могут быть найдены из уравнений

$$C \cos \phi_1 = (A \cos \varphi_1 - B \sin \varphi_2) / 2,$$

$$C \sin \phi_1 = (A \sin \varphi_1 + B \cos \varphi_2) / 2,$$

и

$$\begin{aligned} D\cos\phi_2 &= (A\cos\phi_1 + B\sin\phi_2)/2, \\ D\sin\phi_2 &= (-A\sin\phi_1 + B\cos\phi_2)/2. \end{aligned}$$

3. Разделимость

Пусть ряд F (одномерный, двумерный или комплексный) является суммой m рядов и задачей является нахождение этих слагаемых. Возникает вопрос, в каких случаях можно восстановить аддитивные составляющие ряда F с помощью сингулярного разложения его траекторной матрицы? Для ответа на этот вопрос вводится понятие разделимости одномерных рядов.

Для одномерных рядов понятие разделимости подробно обсуждается в [5]. Здесь мы распространим это понятие на комплексные и двумерные ряды.

3.1. Основные понятия и определения

Пусть $F^{(1)}$, $F^{(2)}$, $G^{(1)}$, $G^{(2)}$ – временные ряды длины N , $F = F^{(1)} + F^{(2)}$, $G = G^{(1)} + G^{(2)}$. Выберем некоторую длину окна L . Обозначим за \mathbf{X} , $\mathbf{X}^{(1)}$ и $\mathbf{X}^{(2)}$ траекторные матрицы рядов (F, G) , $(F^{(1)}, G^{(1)})$ и $(F^{(2)}, G^{(2)})$ в случае двумерного подхода, и рядов $F + iG$, $F^{(1)} + iG^{(1)}$ и $F^{(2)} + iG^{(2)}$, если используется комплексный метод. Если же нас интересует разделимость одномерных рядов $F^{(1)}$ и $F^{(2)}$, то это будут траекторные матрицы рядов F , $F^{(1)}$ и $F^{(2)}$.

Обозначим $\Lambda^{(L)}$, $\Lambda^{(L,1)}$ и $\Lambda^{(L,2)}$ линейные пространства, порожденные столбцами траекторных матриц \mathbf{X} , $\mathbf{X}^{(1)}$ и $\mathbf{X}^{(2)}$ соответственно. Аналогичные обозначения $\Lambda^{(K)}$, $\Lambda^{(K,1)}$ и $\Lambda^{(K,2)}$ будем использовать для пространств, порожденных столбцами транспонированных матриц \mathbf{X}^T , $(\mathbf{X}^{(1)})^T$ и $(\mathbf{X}^{(2)})^T$ (можно сказать, что они порождены строками траекторных матриц). Как и раньше, для комплексного случая вместо транспонирования рассматривается эрмитово сопряжение.

Очевидно, что во всех этих случаях $\mathbf{X} = \mathbf{X}^{(1)} + \mathbf{X}^{(2)}$. Предположим теперь, что матрицы $\mathbf{X}^{(1)}$ и $\mathbf{X}^{(2)}$ получаются из некоторого сингулярного разложения матрицы \mathbf{X} с помощью группировки элементарных матриц \mathbf{X}_k (см. этап «Группировка» с $m=2$). Тогда, поскольку матрицы $\mathbf{X}^{(1)}$ и $\mathbf{X}^{(2)}$ являются траекторными, базовый алгоритм SSA приводит к естественному представлению исходного ряда в виде суммы двух других.

В этом случае разложение $\mathbf{X} = \mathbf{X}^{(1)} + \mathbf{X}^{(2)}$ соответствует представлению пространства $\Lambda^{(L)}$ в виде прямой суммы двух ортогональных линейных подпространств (*собственных подпространств*) $\Lambda^{(L,1)}$ и $\Lambda^{(L,2)}$, а также $\Lambda^{(K)}$ как прямой суммы ортогональных подпространств $\Lambda^{(K,1)}$ и $\Lambda^{(K,2)}$.

Эти соображения приводят к естественному определению разделимости рядов. Мы приведем соответствующее определение только для двумерных рядов, поскольку для одномерных и комплексных оно полностью аналогично.

Определение 3.1. Будем говорить, что ряды $(F^{(1)}, G^{(1)})$ и $(F^{(2)}, G^{(2)})$ слабо L -разделимы если пространство $\Lambda^{(L,1)}$ ортогонально $\Lambda^{(L,2)}$ и пространство $\Lambda^{(K,1)}$ ортогонально $\Lambda^{(K,2)}$.

Несколько следующих утверждения и определений также приводится только для двумерных рядов. Перенос их на комплексный (или одномерный) случай тривиален. Например, для переноса предложения 3.1 с двумерных рядов на комплексные достаточно заменить в (10) и (11) знак транспонирования на знак эрмитова сопряжения.

Предложение 3.1. Пусть ряды $(F^{(1)}, G^{(1)})$ и $(F^{(2)}, G^{(2)})$ слабо L -разделимы. Рассмотрим некоторые сингулярные разложения траекторных матриц $\mathbf{X}^{(1)}$ и $\mathbf{X}^{(2)}$:

$$(10) \quad \mathbf{X}^{(1)} = \sum_k \sqrt{\lambda_{1k}} U_{1k} V_{1k}^T, \quad \mathbf{X}^{(2)} = \sum_m \sqrt{\lambda_{2m}} U_{2m} V_{2m}^T.$$

Тогда разложение

$$(11) \quad \mathbf{X} = \mathbf{X}^{(1)} + \mathbf{X}^{(2)} = \sum_k \sqrt{\lambda_{1k}} U_{1k} V_{1k}^T + \sum_m \sqrt{\lambda_{2m}} U_{2m} V_{2m}^T$$

является сингулярным.

Следствие 3.1. Пусть двумерные ряды $(F^{(1)}, G^{(1)})$ и $(F^{(2)}, G^{(2)})$ слабо L -разделимы. Тогда существует такое сингулярное разложение траекторной матрицы \mathbf{X} ряда (F, G) и такая группировка элементарных матриц этого разложения на две группы, что получившиеся результирующие матрицы являются траекторными матрицами рядов $(F^{(1)}, G^{(1)})$ и $(F^{(2)}, G^{(2)})$.

В формулировке следствия 3.1 утверждается существование подходящего сингулярного разложения траекторной матрицы \mathbf{X} . Конечно, это ограничение можно опустить, если все собственные числа этого разложения различны (тогда сингулярное разложение единственно). В то же время, если некоторые собственные числа совпадают, причем одни из них относятся к сингулярному разложению матрицы $\mathbf{X}^{(1)}$, а другие – к сингулярному разложению матрицы $\mathbf{X}^{(2)}$, то возникающая неопределенность может затруднить практическое разделение рядов $(F^{(1)}, G^{(1)})$ и $(F^{(2)}, G^{(2)})$. Для формального устранения таких ситуаций необходимо усилить понятие разделимости.

Определение 3.2. Если ряды $(F^{(1)}, G^{(1)})$ и $(F^{(2)}, G^{(2)})$ слабо L -разделимы и множество собственных чисел сингулярного разложения траекторной матрицы одного ряда не пересекается с множеством собственных чисел сингулярного разложения второго ряда (то есть если $\lambda_{1k} \neq \lambda_{2m}$ для всех k и m), то ряды $(F^{(1)}, G^{(1)})$ и $(F^{(2)}, G^{(2)})$ называются *сильно L -разделимыми*.

Предложение 3.2. Пусть двумерные ряды $(F^{(1)}, G^{(1)})$ и $(F^{(2)}, G^{(2)})$ сильно L -разделимы. Тогда для любого сингулярного разложения траекторной матрицы \mathbf{X} ряда (F, G) существует такая группировка элементарных матриц этого разложения на две группы, что получившиеся результирующие матрицы являются траекторными матрицами рядов $(F^{(1)}, G^{(1)})$ и $(F^{(2)}, G^{(2)})$.

Изучим условия слабой разделимости подробнее. Пусть имеются два ряда $F = F^{(1)} + F^{(2)}$ и $G = G^{(1)} + G^{(2)}$ длины N . Выбираем некоторую длину окна L . Выпишем для наглядности необходимые нам траекторные матрицы $\mathbf{X}^{(k)}$, $k=1,2$, в одномерном, комплексном и двумерном вариантах.

Одномерный метод (ряд $F^{(k)}$):

$$\mathbf{X}^{(k)} = \begin{pmatrix} f_0^{(k)} & \cdots & f_{K-1}^{(k)} \\ \vdots & \ddots & \vdots \\ f_{L-1}^{(k)} & \cdots & f_{N-1}^{(k)} \end{pmatrix}.$$

Комплексный метод (ряд $F^{(k)} + iG^{(k)}$):

$$\mathbf{X}^{(k)} = \begin{pmatrix} f_0^{(k)} + ig_0^{(k)} & \cdots & f_{K-1}^{(k)} + ig_{K-1}^{(k)} \\ \vdots & \ddots & \vdots \\ f_{L-1}^{(k)} + ig_{L-1}^{(k)} & \cdots & f_{N-1}^{(k)} + ig_{N-1}^{(k)} \end{pmatrix}.$$

Двумерный метод (ряд $(F^{(k)}, G^{(k)})$):

$$\mathbf{X}^{(k)} = \begin{pmatrix} f_0^{(k)} & \cdots & f_{K-1}^{(k)} & g_0^{(k)} & \cdots & g_{K-1}^{(k)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ f_{L-1}^{(k)} & \cdots & f_{N-1}^{(k)} & g_{L-1}^{(k)} & \cdots & g_{N-1}^{(k)} \end{pmatrix}.$$

В следующем утверждении слабая разделимость рядов выражается в терминах строк и столбцов их траекторных матриц $\mathbf{X}^{(k)}$.

Предложение 3.3. Пусть $K = N - L + 1$.

1) **Одномерный метод.** Ряды $F^{(1)}$ и $F^{(2)}$ слабо L -разделимы тогда и только тогда, когда выполнены следующие два условия:

a) для любых $0 \leq k, m < K$

$$f_k^{(1)} f_m^{(2)} + \dots + f_{k+L-1}^{(1)} f_{m+L-1}^{(2)} = 0;$$

b) для любых $0 \leq k, m < L$

$$f_k^{(1)} f_m^{(2)} + \dots + f_{k+K-1}^{(1)} f_{m+K-1}^{(2)} = 0.$$

2) **Комплексный метод.** Ряды $F^{(1)} + iG^{(1)}$ и $F^{(2)} + iG^{(2)}$ слабо L -разделимы тогда и только тогда, когда выполнены следующие два условия:

a) для любых $0 \leq k, m < K$

$$\begin{cases} f_k^{(1)} f_m^{(2)} + \dots + f_{k+L-1}^{(1)} f_{m+L-1}^{(2)} = -(g_k^{(1)} g_m^{(2)} + \dots + g_{k+L-1}^{(1)} g_{m+L-1}^{(2)}), \\ f_k^{(1)} g_m^{(2)} + \dots + f_{k+L-1}^{(1)} g_{m+L-1}^{(2)} = g_k^{(1)} f_m^{(2)} + \dots + g_{k+L-1}^{(1)} f_{m+L-1}^{(2)}; \end{cases}$$

b) для любых $0 \leq k, m < L$

$$\begin{cases} f_k^{(1)} f_m^{(2)} + \dots + f_{k+K-1}^{(1)} f_{m+K-1}^{(2)} = -(g_k^{(1)} g_m^{(2)} + \dots + g_{k+K-1}^{(1)} g_{m+K-1}^{(2)}), \\ f_k^{(1)} g_m^{(2)} + \dots + f_{k+K-1}^{(1)} g_{m+K-1}^{(2)} = g_k^{(1)} f_m^{(2)} + \dots + g_{k+K-1}^{(1)} f_{m+K-1}^{(2)}. \end{cases}$$

3) **Двумерный метод.** Ряды $(F^{(1)}, G^{(1)})$ и $(F^{(2)}, G^{(2)})$ слабо L -разделимы тогда и только тогда, когда выполнены следующие два условия:

a) для любых $0 \leq k, m < K$

$$\begin{cases} f_k^{(1)} f_m^{(2)} + \dots + f_{k+L-1}^{(1)} f_{m+L-1}^{(2)} = 0, \\ f_k^{(1)} g_m^{(2)} + \dots + f_{k+L-1}^{(1)} g_{m+L-1}^{(2)} = 0, \\ f_k^{(2)} g_m^{(1)} + \dots + f_{k+L-1}^{(2)} g_{m+L-1}^{(1)} = 0, \\ g_k^{(1)} g_m^{(2)} + \dots + g_{k+L-1}^{(1)} g_{m+L-1}^{(2)} = 0; \end{cases}$$

b) для любых $0 \leq k, m < L$

$$f_k^{(1)} f_m^{(2)} + \dots + f_{k+K-1}^{(1)} f_{m+K-1}^{(2)} + g_k^{(1)} g_m^{(2)} + \dots + g_{k+K-1}^{(1)} g_{m+K-1}^{(2)} = 0.$$

Из предложения 3.3 видно, что условия слабой разделимости для одномерного, комплексного и двумерного вариантов метода SSA, вообще говоря, различны: из выполнения условий для одного метода не следует выполнение условий для другого. Можно, однако, отметить следующий факт.

Следствие 3.2. Если ряды $F^{(1)}$ и $F^{(2)}$, $G^{(1)}$ и $G^{(2)}$, $F^{(1)}$ и $G^{(2)}$, а также $G^{(1)}$ и $F^{(2)}$ слабо L -разделимы, то двумерные ряды $(F^{(1)}, G^{(1)})$ и $(F^{(2)}, G^{(2)})$, а также комплексные ряды $F^{(1)} + iG^{(1)}$ и $F^{(2)} + iG^{(2)}$ слабо L -разделимы.

3.2. Модельный пример: разделимость константных рядов от гармоник

Рассмотрим два ряда $F^{(1)}$ и $F^{(2)}$ длины N . Пусть каждый из этих рядов имеет составляющую константного вида и гармоническую составляющую. Другими словами, пусть $F^{(1)} = P^{(1)} + S^{(1)}$ и $F^{(2)} = P^{(2)} + S^{(2)}$, где

$$P^{(1)} = (p_0^{(1)}, \dots, p_{N-1}^{(1)}), \quad P^{(2)} = (p_0^{(2)}, \dots, p_{N-1}^{(2)}),$$

$$S^{(1)} = (s_0^{(1)}, \dots, s_{N-1}^{(1)}), \quad S^{(2)} = (s_0^{(2)}, \dots, s_{N-1}^{(2)}),$$

$$p_k^{(1)} = p_1, \quad p_k^{(2)} = p_2, \quad s_k^{(1)} = A \cos(2\pi\omega_1 k + \varphi_1) \quad \text{и} \quad s_k^{(2)} = B \cos(2\pi\omega_2 k + \varphi_2).$$

Предполагается, что $A, B, p_1, p_2 \neq 0$, а $0 < \omega_1, \omega_2 < 1/2$.

В этом разделе мы рассмотрим вопросы разделимости (слабой и сильной) одномерных рядов $P^{(1)}$ и $S^{(1)}$, двумерных рядов $(P^{(1)}, P^{(2)})$ и $(S^{(1)}, S^{(2)})$, а также комплексных рядов $P^{(1)} + iP^{(2)}$ и $S^{(1)} + iS^{(2)}$. Несмотря на кажущуюся искусственность и примитивность такой постановки задачи, она важна для тех практических ситуаций, когда у ряда наблюдаются регулярные осцилляции вокруг медленно меняющегося тренда. Отметим, что в двумерном и комплексном случаях мы ставим задачу разделения согласованных компонент, т. е. пары константных рядов от пары гармоник.

3.2.1. Слабая разделимость. Начнем со слабой разделимости рядов.

Предложение 3.4. 1) Одномерные ряды $P^{(1)}$ и $S^{(1)}$ слабо разделимы тогда и только тогда, когда $L\omega$ и $K\omega$ – целые.

2) Если частоты ω и ω_2 различны, то двумерные ряды $(P^{(1)}, P^{(2)})$ и $(S^{(1)}, S^{(2)})$ слабо разделимы тогда и только тогда, когда $L\omega$, $K\omega$, $L\omega_2$ и $K\omega_2$ – целые.

3) Если частоты одинаковы ($\omega = \omega_1 = \omega_2$), то двумерные ряды $(P^{(1)}, P^{(2)})$ и $(S^{(1)}, S^{(2)})$ слабо разделимы в трех случаях: а) если $L\omega$ и $K\omega$ – целые, б) если $L\omega$ – целое, $Ap_1 = -Bp_2$, а $\varphi_1 = \varphi_2$ и с) если $L\omega$ – целое, $Ap_1 = Bp_2$, а $\varphi_1 = \varphi_2 + \pi$.

4) Комплексные ряды $P^{(1)} + iP^{(2)}$ и $S^{(1)} + iS^{(2)}$ слабо разделимы тогда и только тогда, когда числа $L\omega$, $K\omega$, $L\omega_2$ и $K\omega_2$ – целые.

Замечание 3.1. Поскольку здесь мы рассматриваем гармоники с (вообще говоря) разными частотами ω и ω_2 , то последовательный одномерный метод позволяет для рядов $F^{(1)}$ и $F^{(2)}$ выбирать свою длину окна. В двумерном и комплексном вариантах для анализа всей системы рядов используется один параметр. Значит, если частоты ω и ω_2 различны, то последовательный одномерный метод может привести к разделимости там, где двумерный и комплексный методы потерпят неудачу. В то же время (при различных частотах), если какой-то двумерный метод дает разделимость с длиной окна L , то и дважды примененный одномерный метод с таким же L даст разделимость.

Замечание 3.2. В случае гармонических рядов с одинаковыми частотами условия слабой разделимости для комплексного и одномерного методов оказываются одинаковыми и, вообще говоря, более жесткими, чем для двумерного метода.

Например, пусть $N = 40$, $\omega = 1/30$, $L = 30$, $p_1 = 3$ и $p_2 = -1.5$, $A = 1$ и $B = 2$, $\varphi_1 = \varphi_2 = 0$. Последовательное применение метода к каждому из рядов по

отдельности и комплексный вариант не приведут к разделимости. Двумерная же схема дает слабую разделимость.

В то же время такие отличия с практической точки зрения следует признать малозначимыми, так как для реализации преимущества двумерного метода требуется наличие жесткой связи между значениями постоянных рядов и амплитудами гармоник. В целом можно сказать, что для одинаковых частот все три метода дают схожие результаты в смысле слабой разделимости.

3.2.2. Сильная разделимость. Напомним, что слабо L -разделимые ряды называются сильно L -разделимыми, если множество собственных чисел разложения траекторной матрицы одного ряда не пересекается с множеством собственных чисел разложения второго ряда.

Рассмотрим условия сильной разделимости тех же одномерных рядов $P^{(1)}$ и $S^{(1)}$, двумерных рядов $(P^{(1)}, P^{(2)})$ и $(S^{(1)}, S^{(2)})$, а также комплексных рядов $P^{(1)} + iP^{(2)}$ и $S^{(1)} + iS^{(2)}$ в предположении, что $\omega_1 = \omega_2 = \omega$.

Прежде всего, нам нужно обеспечить условия слабой разделимости рядов. Согласно предложению 3.4, слабая разделимость имеет место для всех вариантов SSA, если числа $L\omega$ и $K\omega$ – целые. Будем предполагать эти условия выполненными.

Далее, если при сингулярном разложении траекторной матрицы суммы двух рядов эти ряды оказались слабо разделимыми, то для отыскания собственных чисел разложения достаточно рассматривать сингулярные разложения траекторных матриц каждого ряда по отдельности. Следовательно, нам осталось воспользоваться результатами раздела 2.5, касающимися собственных чисел. При этом будем предполагать, что для комплексного гармонического ряда $d=2$. Иначе говоря, мы не рассматриваем случай $A=B$, $|\varphi_1 - \varphi_2| = \pi/2 \pmod{\pi}$, который в некотором смысле является исключительным.

В следующем предложении собраны (с учетом сделанных выше замечаний) нужные нам результаты о собственных числах сингулярных разложений траекторных матриц одномерного ряда $P^{(1)} + S^{(1)}$, двумерного ряда $(P^{(1)} + S^{(1)}, P^{(2)} + S^{(2)})$, а также комплексного ряда $P^{(1)} + S^{(1)} + i(P^{(2)} + S^{(2)})$. Все разложения трехмерны ($d=3$), первое собственное число λ_1 всегда соотносится с константным рядом, два других (λ_2 и λ_3) – с гармоническим.

Предложение 3.5. Пусть выполнены следующие условия: а) $\omega = \omega_1 = \omega_2$; б) числа $L\omega$ и $K\omega$ – целые; с) либо $A \neq B$, либо $|\varphi_1 - \varphi_2| \neq \pi/2 \pmod{\pi}$. Тогда

1) Собственные числа, порожденные одномерным рядом $P^{(1)} + S^{(1)}$, имеют вид $\lambda_1 = p_1^2 LK$; $\lambda_2 = \lambda_3 = A^2 LK/4$.

2) Собственные числа, порожденные двумерным рядом $(P^{(1)} + S^{(1)}, P^{(2)} + S^{(2)})$, имеют вид $\lambda_1 = (p_1^2 + p_2^2) LK$; $\lambda_2 = \lambda_3 = (A^2 + B^2) LK/4$.

3) Собственные числа, порожденные комплексным рядом $P^{(1)} + S^{(1)} + i(P^{(2)} + S^{(2)})$, имеют вид

$$\lambda_1 = (p_1^2 + p_2^2) LK; \lambda_2 = (A^2 + B^2 + 2AB \sin \varphi) LK/4,$$

$$\lambda_3 = (A^2 + B^2 - 2AB \sin \varphi) LK/4,$$

где $\varphi = \varphi_1 - \varphi_2$.

Тем самым (в условиях предложения 3.5) мы приходим к следующему утверждению.

Следствие 3.3. 1) Если $A^2 \neq 4p_1^2$, то одномерные ряды $P^{(1)}$ и $S^{(1)}$ сильно разделимы. Аналогично, если $B^2 \neq 4p_2^2$, то ряды $P^{(2)}$ и $S^{(2)}$ сильно разделимы.

2) Если $A^2 + B^2 \neq 4(p_1^2 + p_2^2)$, то двумерные ряды $(P^{(1)}, P^{(2)})$ и $(S^{(1)}, S^{(2)})$ сильно разделимы.

3) Если $A^2 + B^2 \pm 2AB \sin(\varphi_1 - \varphi_2) \neq 4(p_1^2 + p_2^2)$, то комплексные ряды $P^{(1)} + iP^{(2)}$ и $S^{(1)} + iS^{(2)}$ сильно разделимы.

Замечание 3.3. Тем самым условия сильной разделимости (при наличии слабой) у рассматриваемых методов не являются эквивалентными. В частности, в условиях предложения 3.5,

а) если нет сильной разделимости только для одной из пар $(P^{(1)}, S^{(1)})$, $(P^{(2)}, S^{(2)})$, то двумерный метод даст сильную разделимость констант от гармоник, так как в этом случае $A^2 + B^2 \neq 4(p_1^2 + p_2^2)$;

б) наоборот, если $A^2 + B^2 = 4(p_1^2 + p_2^2)$, но $A^2 \neq 4p_1^2$ и $B^2 \neq 4p_2^2$, то константные ряды, не разделимые сильно с парой гармоник двумерным вариантом метода, сильно разделимы последовательно одномерным вариантом;

в) наконец, если $A^2 = 4p_1^2$, $B^2 = 4p_2^2$ и $\sin \varphi \neq 0$, то константные ряды, не разделимые сильно с парой гармоник ни последовательным одномерным, ни двумерным методами, будут сильно разделимы комплексным вариантом SSA.

Эти различия, однако, не имеют серьезного практического значения. Нежелательных соотношений $A^2 = 4p_1^2$ и $B^2 = 4p_2^2$ легко избежать, прибавив к рядам $P^{(1)} + S^{(1)}$ и $P^{(2)} + S^{(2)}$ любые ненулевые константы. Конечно, эти константы легко (как аналитически, так и экспериментально) подобрать так, чтобы сильная разделимость константного ряда от гармонического имела место и в комплексном варианте SSA.

3.3. Модельный пример: разделимость разночастотных гармоник

3.3.1. Слабая разделимость. Рассмотрим четыре гармонических вещественных временных ряда $R^{(1)}$, $R^{(2)}$, $S^{(1)}$ и $S^{(2)}$ длины N . Предположим, что частоты рядов $R^{(1)}$ и $R^{(2)}$ одинаковы (и равны ω_1), а также что частоты $S^{(1)}$ и $S^{(2)}$ одинаковы (и равны ω_2), причем $\omega_1 \neq \omega_2$. Другими словами, мы предполагаем, что (в естественных обозначениях) члены временных рядов $R^{(1)}$, $R^{(2)}$, $S^{(1)}$ и $S^{(2)}$ имеют вид

$$\begin{aligned} r_k^{(1)} &= A_1 \cos(2\pi\omega_1 k + \varphi_1), & r_k^{(2)} &= B_1 \cos(2\pi\omega_1 k + \varphi_2), \\ s_k^{(1)} &= A_2 \cos(2\pi\omega_2 k + \phi_1), & s_k^{(2)} &= B_2 \cos(2\pi\omega_2 k + \phi_2), \end{aligned}$$

$k = 0, \dots, N-1$. Как и раньше, $A_1, A_2, B_1, B_2 > 0$.

Наша задача – исследовать слабую разделимость согласованных гармоник, т. е. разделимость рядов $R^{(1)}$ и $S^{(1)}$, $(R^{(1)}, R^{(2)})$ и $(S^{(1)}, S^{(2)})$, а также $R^{(1)} + iR^{(2)}$ и $S^{(1)} + iS^{(2)}$.

Следующее предложение дает полный ответ на поставленный вопрос.

Предложение 3.6. Пусть $\omega_1 \neq \omega_2$.

1) Ряды $R^{(1)}$ и $S^{(1)}$ слабо разделимы тогда и только тогда, когда числа $L(\omega_2 - \omega_1)$, $L(\omega_2 + \omega_1)$, $K(\omega_2 - \omega_1)$ и $K(\omega_2 + \omega_1)$ – целые.

2) Двумерные ряды $(R^{(1)}, R^{(2)})$ и $(S^{(1)}, S^{(2)})$ слабо разделимы в пяти случаях:

$$\left\{ \begin{array}{l} L(\omega_2 - \omega_1) \\ L(\omega_2 + \omega_1) \\ K(\omega_2 - \omega_1) \\ K(\omega_2 + \omega_1) \end{array} \right\} \text{ – целые, } \left\{ \begin{array}{l} L(\omega_2 - \omega_1) \text{ – целое} \\ L(\omega_2 + \omega_1) \text{ – целое} \\ K(\omega_2 - \omega_1) \text{ – целое,} \\ A_1 A_2 = B_1 B_2 \\ |\varphi_2 - \varphi_1 + \phi_2 - \phi_1| = \pi \end{array} \right\}, \left\{ \begin{array}{l} L(\omega_2 - \omega_1) \text{ – целое} \\ L(\omega_2 + \omega_1) \text{ – целое} \\ K(\omega_2 + \omega_1) \text{ – целое,} \\ A_1 A_2 = B_1 B_2 \\ |\varphi_2 - \varphi_1 + \phi_1 - \phi_2| = \pi \end{array} \right\},$$

$$\left\{ \begin{array}{l} L(\omega_2 - \omega_1) \text{ – целое} \\ L(\omega_2 + \omega_1) \text{ – целое} \\ A_1 A_2 = B_1 B_2 \\ |\varphi_2 - \varphi_1| = \pi \\ \phi_2 - \phi_1 = 0 \end{array} \right\}, \left\{ \begin{array}{l} L(\omega_2 - \omega_1) \text{ – целое} \\ L(\omega_2 + \omega_1) \text{ – целое} \\ A_1 A_2 = B_1 B_2 \\ \varphi_2 - \varphi_1 = 0 \\ |\phi_2 - \phi_1| = \pi \end{array} \right\}.$$

3) Комплексные ряды $R^{(1)} + iR^{(2)}$ и $S^{(1)} + iS^{(2)}$ слабо разделимы в трех случаях:

$$\left\{ \begin{array}{l} L(\omega_2 - \omega_1) \\ L(\omega_2 + \omega_1) \\ K(\omega_2 - \omega_1) \\ K(\omega_2 + \omega_1) \end{array} \right\} \text{ – целые, } \left\{ \begin{array}{l} L(\omega_2 - \omega_1) \text{ – целое} \\ K(\omega_2 - \omega_1) \text{ – целое} \\ A_1 = A_2 = B_1 = B_2 \\ |\varphi_2 - \varphi_1| = |\varphi_2 - \varphi_1| = \pi/2 \end{array} \right\},$$

$$\left\{ \begin{array}{l} L(\omega_2 - \omega_1) \text{ – целое} \\ K(\omega_2 - \omega_1) \text{ – целое} \\ A_1 = A_2 = B_1 = B_2 \\ |\varphi_2 - \varphi_1| = |\varphi_1 - \varphi_2| = \pi/2 \end{array} \right\}.$$

Мы снова видим, что методы не эквивалентны в смысле слабой разделимости, причем как двумерный, так и комплексный SSA дают дополнительные возможности по сравнению с последовательным применением одномерного SSA. Однако эти различия с практической точки зрения не очень существенны, так как связаны со специфическими соотношениями как между амплитудами гармоник, так и между их фазами. Основным условием слабой разделимости следует считать то условие, при котором числа $L(\omega_2 - \omega_1)$, $L(\omega_2 + \omega_1)$, $K(\omega_2 - \omega_1)$, $K(\omega_2 + \omega_1)$ являются целыми.

3.3.2. Сильная разделимость. Теперь, в условиях раздела 3.3.1, мы займемся сильной разделимостью гармонических временных рядов: одномерных $R^{(1)}$ и $S^{(1)}$, двумерных $(R^{(1)}, R^{(2)})$ и $(S^{(1)}, S^{(2)})$, а также комплексных $R^{(1)} + iR^{(2)}$ и $S^{(1)} + iS^{(2)}$. Как и раньше, предполагается, что $\omega_1 \neq \omega_2$.

Кроме того, предполагается, что числа $L\omega_1$, $L\omega_2$, $K\omega_1$, $K\omega_2$, являются целыми. Это условие обеспечивает слабую разделимость упомянутых рядов. Кроме того, оно соответствует условиям предложения 2.3, дающего нам явный вид собственных чисел разложения траекторных матриц гармонических рядов.

Наконец, это требование (см. предложения 3.4 и 3.5) является основным и для делимости константных рядов от гармоник.

Будет считаться выполненным еще одно, менее существенное требование. А именно, мы предполагаем для простоты, что $|\varphi_1 - \varphi_2| \neq \pi/2 \pmod{\pi}$ и $|\phi_1 - \phi_2| \neq \pi/2 \pmod{\pi}$. Тогда (см. предложение 2.2) для комплексного SSA каждая из гармоник $R^{(1)} + iR^{(2)}$, $S^{(1)} + iS^{(2)}$ будет порождать двумерное сингулярное разложение своей траекторной матрицы независимо от соотношения амплитуд A_1 и B_1 или A_2 и B_2 .

Следующее утверждение аналогично предложению 3.5. Мы снова рассматриваем сингулярные разложения траекторных матриц рядов $R^{(1)} + S^{(1)}$, $(R^{(1)} + S^{(1)}, R^{(2)} + S^{(2)})$ и $R^{(1)} + S^{(1)} + i(R^{(2)} + S^{(2)})$ и интересуемся соответствующими собственными числами. Все разложения имеют порядок 4, первые два собственных числа (λ_1 и λ_2) относятся к гармоникам $R^{(1)}$ и $R^{(2)}$, в то время как собственные числа λ_3 и λ_4 – к гармоникам $S^{(1)}$ и $S^{(2)}$.

Предложение 3.7. Пусть выполнены следующие условия: а) $\omega_1 \neq \omega_2$; б) числа $L\omega_1$, $L\omega_2$, $K\omega_1$, $K\omega_2$ – целые; в) $|\varphi_1 - \varphi_2| \neq \pi/2 \pmod{\pi}$ и $|\phi_1 - \phi_2| \neq \pi/2 \pmod{\pi}$. Тогда

1) Собственные числа, порожденные одномерным рядом $R^{(1)} + S^{(1)}$, имеют вид $\lambda_1 = \lambda_2 = A_1^2 LK/4$; $\lambda_3 = \lambda_4 = A_2^2 LK/4$.

2) Собственные числа, порожденные двумерным рядом $(R^{(1)} + S^{(1)}, R^{(2)} + S^{(2)})$, имеют вид $\lambda_1 = \lambda_2 = (A_1^2 + B_1^2)LK/4$, $\lambda_3 = \lambda_4 = (A_2^2 + B_2^2)LK/4$.

3) Собственные числа, порожденные комплексным рядом $R^{(1)} + S^{(1)} + i(R^{(2)} + S^{(2)})$, имеют вид

$$\lambda_1 = (A_1^2 + B_1^2 + 2A_1B_1 \sin \varphi)LK/4, \quad \lambda_2 = (A_1^2 + B_1^2 - 2A_1B_1 \sin \varphi)LK/4,$$

$$\lambda_3 = (A_2^2 + B_2^2 + 2A_2B_2 \sin \phi)LK/4, \quad \lambda_4 = (A_2^2 + B_2^2 - 2A_2B_2 \sin \phi)LK/4$$

где $\varphi = \varphi_1 - \varphi_2$, $\phi = \phi_1 - \phi_2$.

Обсудим эти результаты с точки зрения идеологии SSA. Поскольку мы интересуемся делимостью разночастотных гармоник, нас не волнует совпадение собственных чисел, относящихся к одной из гармоник. Наоборот, совпадение λ_1 с λ_2 или λ_3 с λ_4 может быть удобно на практике для отыскания собственных троек, относящихся к одной компоненте ряда. Наоборот, совпадение, скажем, λ_2 и λ_3 означает вполне реальную возможность «перемешивания» компонент. Тем самым мы получаем, что в условиях предложения 3.7 имеют место следующие утверждения.

Следствие 3.4. 1) Одномерные гармонические ряды $R^{(1)}$ и $S^{(1)}$ сильно делимы тогда и только тогда, когда $A_1 \neq A_2$. Аналогично, одномерные ряды $R^{(2)}$ и $S^{(2)}$ сильно делимы тогда и только тогда, когда $B_1 \neq B_2$.

2) Двумерные гармонические ряды $(R^{(1)}, R^{(2)})$ и $(S^{(1)}, S^{(2)})$ сильно делимы тогда и только тогда, когда $A_1^2 + B_1^2 \neq A_2^2 + B_2^2$.

3) Комплексные гармонические ряды $R^{(1)} + iR^{(2)}$ и $S^{(1)} + iS^{(2)}$ сильно делимы тогда и только тогда, когда $A_1^2 + B_1^2 \pm 2A_1B_1 \sin \varphi \neq A_2^2 + B_2^2 \pm 2A_2B_2 \sin \phi$.

Замечание 3.4. Как и в предыдущих примерах, условия сильной делимости гармонических рядов зависят от того, какой из вариантов SSA

(последовательный одномерный, двумерный или комплексный) применяется. При этом, если для одномерного и двумерного методов условия сильной разделимости зависят только от амплитуд гармоник, для комплексного варианта существенным являются также и их фазы. В частности, в условиях предложения 3.7

- a) если нет сильной разделимости только для одной из гармоник, то двумерный вариант дает сильную разделимость пар гармоник одинаковой частоты;
- b) если $A_1 \neq A_2$ и $B_1 \neq B_2$, но $A_1^2 + B_1^2 = A_2^2 + B_2^2$, то пары гармоник, не разделимые сильно двумерным вариантом метода, будут сильно разделимы по отдельности;
- c) если $A_1 = A_2 = B_1 = B_2$ и $\sin \varphi \neq \sin \phi$, то только комплексный вариант метода разделяет разночастотные гармоники.

Последний случай представляет особый интерес, так как разделимость гармоник с одинаковой (на практике – с близкой) амплитудой представляет серьезную проблему в одномерном SSA. Как показано (по крайней мере, на уровне точной разделимости), двумерный метод не решает эту проблему, в то время как комплексный, вообще говоря, способствует ее решению.

В то же время, если ряды имеют более сложную структуру, чем просто сумма двух гармоник, то большая разница между собственными числами, описывающими одну гармоническую компоненту (то есть между λ_3 и λ_2 или между λ_5 и λ_4) комплексного ряда, может сослужить плохую службу. Действительно, тогда появится опасность смешения одной из составляющих гармоники с другими компонентами ряда, например, с одной из компонент тренда или с шумовой компонентой.

Представляется, что последняя ситуация является более распространенной. В этом смысле, возможно, MSSA имеет преимущество над CSSA.

4. Примеры обработки рядов методами SSA

В этом разделе мы (на примере двумерного ряда продаж вин в Австралии) кратко опишем технологию применения двумерного варианта MSSA. Методики применения одномерного SSA и комплексного CSSA аналогичны. В заключение на том же примере будут продемонстрированы различия между тремя модификациями SSA.

4.1. Методика применения MSSA

Рассмотрим два временных ряда продаж вин (FORT и DRY – крепленые и сухие вина) в Австралии, ежемесячно, с января 1980 года по ноябрь 1993 года ($N=167$), в тысячах литров [9]. Ряды изображены на рис.1 и рис.2. Видно, что ряд FORT имеет почти постоянный тренд, а ряд DRY – тренд, близкий к линейному и с небольшим углом наклона. Также видно, что в рядах присутствуют сезонные колебания, причем годовое поведение продаж вин имеет достаточно сложную форму, что соответствует разложению сезонных компонент рядов на (возможно, модулированные) гармоники с почти всеми частотами из набора $(1/12, 1/6=2/12, 1/4=3/12, 1/3=4/12, 1/2.4=5/12, 1/2=6/12)$.

Посмотрим, к какому разложению приведет применение к этим рядам двумерного варианта метода MSSA. Для улучшения разделимости возьмем длину окна L кратной периоду сезонной компоненты, т. е. 12-ти. Кроме того,

возьмем L примерно равным половине длины ряда: $L = 84$. Тогда K также равняется 84.

На рис. 3, 4 и 5 приведены логарифмы собственных чисел сингулярного разложения траекторной матрицы рядов для двумерного MSSA, а также соответствующие собственные и факторные вектора. Форма графика собственных чисел говорит о том, что, скорее всего, первые 12 собственных троек соответствуют сигналу, а остальные – шуму.

Глядя на собственные и факторные вектора (рис. 4 и 5), можно идентифицировать первую и шестую собственные тройки как соответствующие главным образом тренду. Напомним, что первая половина факторных векторов соответствует первому ряду, а вторая – соответственно, второму.

Видно, что тренд не до конца разделился с более высокочастотными составляющими ряда. Анализ показывает, что произошло небольшое перемешивание тренда с сезонными компонентами рядов.

Далее, предложение 2.2 говорит о том, что гармоническим составляющим ряда соответствуют собственные и факторные вектора, также являющиеся гармоническими (в случае приближенной делимости – приближенно гармоническими) рядами. Анализ собственных и факторных векторов (см. рис. 4 и 5), показывает, что 2-я и 3-я собственные тройки соответствуют годовой периодичности (период 12), 4-я и 5-я – четырехмесячному периоду, 7-я и 8-я – полугодовой (период 6) периодичности, а частоты $1/2.4$ и $1/3$ смешались и содержатся в собственных тройках 9–12. Компонента периода 2 в данном разложении не выделилась.

Для нахождения пар собственных векторов, соответствующих одной гармонике, очень удобным оказывается попарное изображение собственных векторов на двумерных скаттерграммах (рис.6). В силу ортогональности собственных векторов при таком изображении правильный (выпуклый или звездчатый) многоугольник, возможно, несколько искаженный, идентифицирует гармонику. Смещение гармоник может быть выявлено с помощью периодограммного анализа, а также на основе анализа собственных троек, имеющих близкие собственные числа.

На рис. 7 и 8 представлены первые шесть компонент разложения исходных рядов, полученные без группировки (т.е. операция диагонального осреднения применена непосредственно к каждой элементарной матрице). Рис. 7 соответствует разложению первого ряда (FORT), а рис. 8 – разложению второго ряда (DRY).

Такая «элементарная» группировка не приводит к разделимым, а, следовательно, интерпретируемым компонентам ряда. Для того чтобы согласованно разложить каждый из рядов на тренд, сезонную компоненту и шум, нужно сгруппировать собственные тройки согласно их порядковым номерам на 3 группы: $I_1 = \{1,6\}$, $I_2 = \{2-5,7-12\}$, I_3 – все остальные собственные тройки.

На рис. 9–11 приведено соответствующее разложение исходных рядов. Видно, что тренд (рис. 9) достаточно гладкий (т.е., смешение с сезонной компонентой незначительное; к тому же вклад шестой собственной тройки небольшой). С помощью рис. 10 можно сопоставить сезонное поведение продаж вин для рассматриваемых рядов.

Для более детального изучения сезонных компонент может оказаться полезным сопоставление их гармонических составляющих (рис. 12–14). На рис. 12 и 13 изображены годовые и полугодовые гармонические составляющие

рядов. Их фазы смещены друг относительно друга примерно на четверть периода. На рис. 14 изображены суммы остальных выделенных гармоник. Видно, что они почти совпадают, как по фазе, так и по амплитуде. Чтобы проинтерпретировать такое соответствие, нужна дополнительная информация о специфике употребления и продаж вин в Австралии, а также о технологии сбора данных.

4.2. Сравнение с другими вариантами

Проведем сравнение результатов рассмотренного выше MSSA с CSSA и с применением одномерного SSA к каждому из рядов по отдельности. В таблице 1 приведены результаты идентификации первых четырнадцати собственных троек разложений траекторных матриц с $L = 84$. Числа в клетках таблицы означают периоды гармоник (возможно, модулированных), присутствующих в соответствующей собственной тройке. Жирным шрифтом выделены периоды тех гармоник, наличие которых существенно.

Таблица 1. Идентификация собственных троек для разложения SSA, MSSA и CSSA

	FORT, SSA	DRY, SSA	FORT+DRY, MSSA	FORT+DRY, CSSA
1	тренд	тренд	тренд+12	тренд
2	12	12	12	12
3	12	12	12	4
4	4	4	4	4
5	4	4	4	6
6	6	3+6+2.4	тренд+6+4+3+2.4	3+12+2.4
7	6	3+6	6	12+2.4+3
8	2.4	2.4+3+6	6	2.4+3+12
9	2.4	3+6	3+2.4	2.4
10	3	2.4+3+6	3+2.4	3
11	3	2.4	2.4+3	6
12	шум	тренд	2.4+3	тренд
13	шум	тренд	шум	шум
14	шум	2	шум	шум

Опишем результаты сравнения:

- в смысле слабой делимости одномерный метод показал себя лучше: для ряда DRY выделился более подробный тренд и обнаружилась гармоника периода 2; однако еще раз обратим внимание, что при применении SSA к рядам по отдельности нужно проводить дополнительную работу по сопоставлению компонент рядов, имеющих одинаковый смысл.;
- в смысле сильной делимости результат применения одномерного метода сильно зависит от ряда: для ряда FORT вообще не произошло перемешивания компонент, в то время как для ряда DRY сильно перемешались три гармонические компоненты;
- метод MSSA частично убирает перемешивание для ряда DRY, однако добавляет его в ряд FORT;
- в одномерных SSA разложениях, а также в MSSA, компоненты, соответствующие одной гармонической составляющей ряда, идут подряд, что облегчает их идентификацию;

- в CSSA такой порядок компонент нарушен, что согласуется с теорией: из-за сдвига по фазе годовых (и полугодовых) гармонических составляющих исходных рядов соответствующие собственные тройки расположены не подряд; это привело, например, к смешению 12-месячной гармоник с более высокочастотными (заметим, что для комплексного случая «потерянная» (неучтенная при восстановлении) компонента приводит к искажению фазы восстановленной компоненты ряда, в то время как для MSSA – к искажению амплитуды).

Таким образом, ни один из методов не приводит к полному разделению всех интерпретируемых компонент для обоих рядов. Если сравнивать MSSA с CSSA, то можно говорить о некотором преимуществе двумерного SSA над комплексным: единственными компонентами, которые существенно смешиваются MSSA, являются слабые гармоники с периодами 3 и 2.4, в то время как CSSA не позволяет «чисто» выделить основную 12-месячную гармонику.

В то же время, если бы амплитуды двух основных (12-месячной и 4-месячной) гармоник были бы почти одинаковы, то, возможно, результат сравнения MSSA с CSSA был бы противоположным.

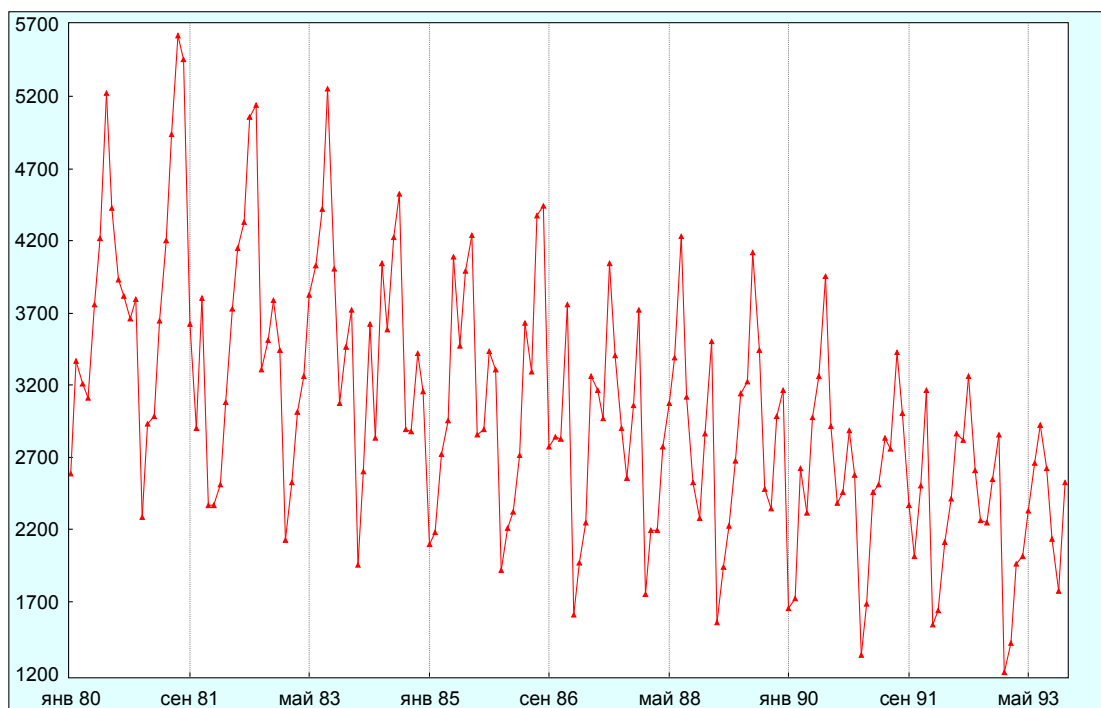


Рис. 1. Временной ряд FORT (объем продаж крепленых вин в Австралии)

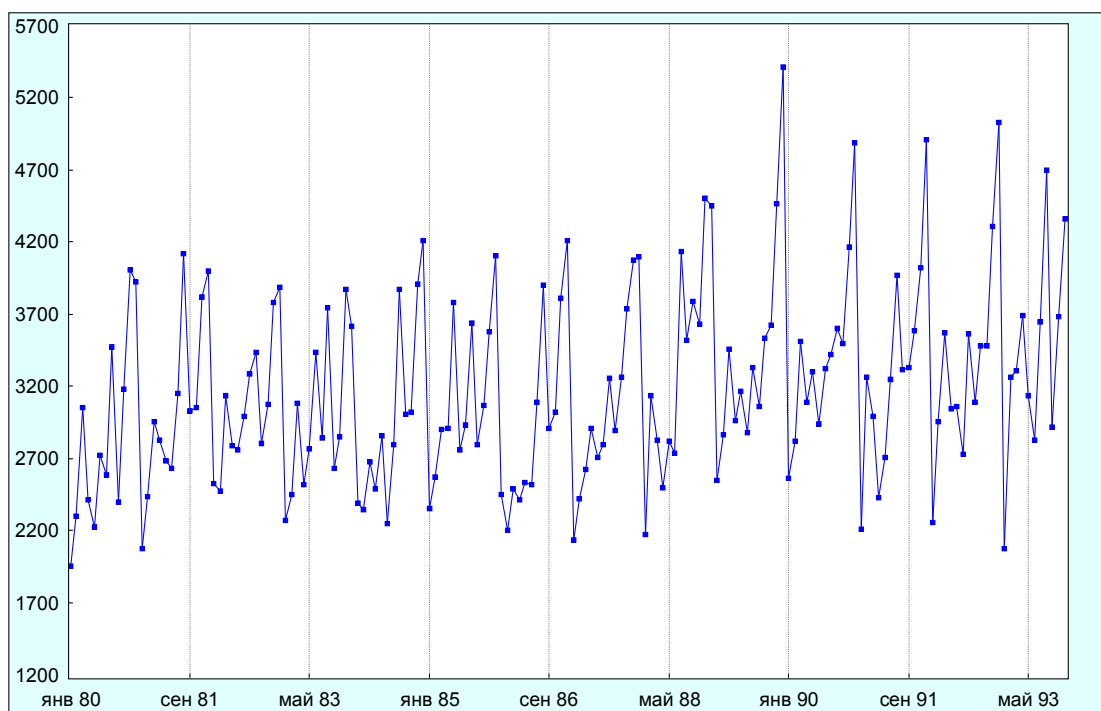


Рис. 2. Временной ряд DRY (объем продаж сухих вин в Австралии)

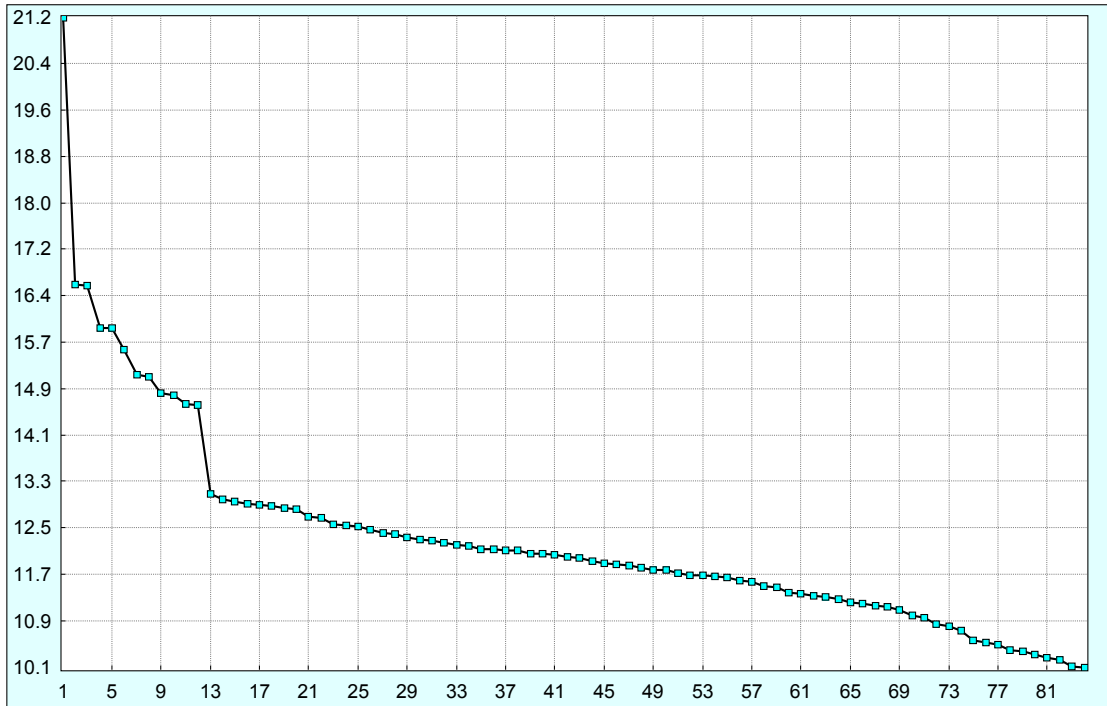


Рис. 3. MSA: логарифмы собственных чисел

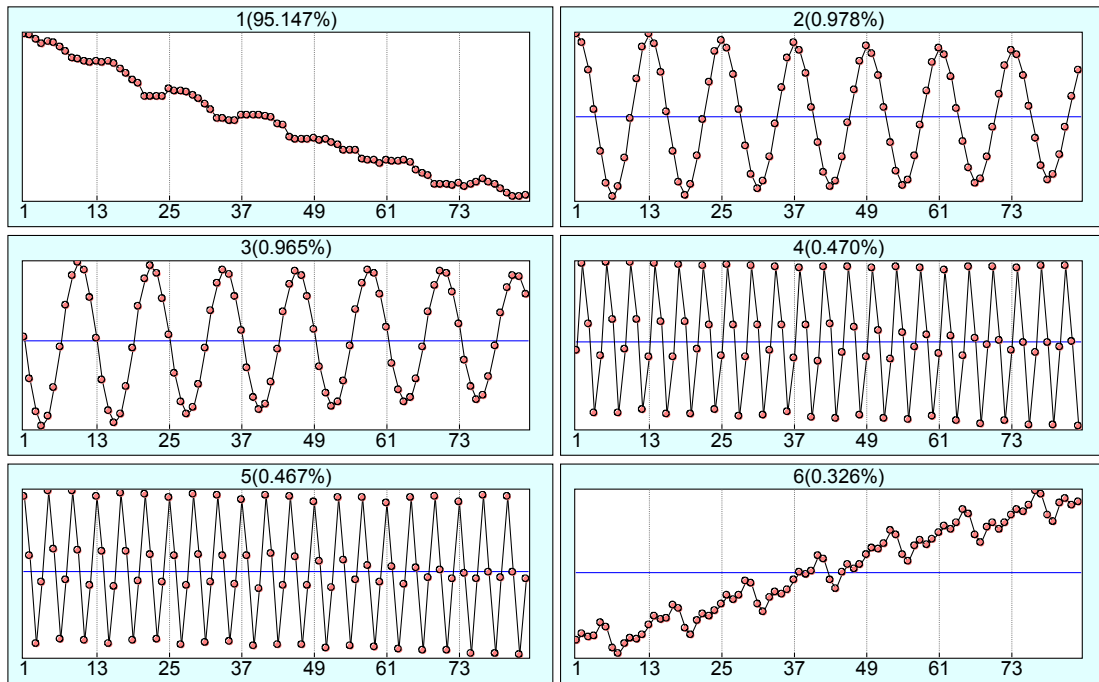


Рис. 4. MSA: собственные вектора

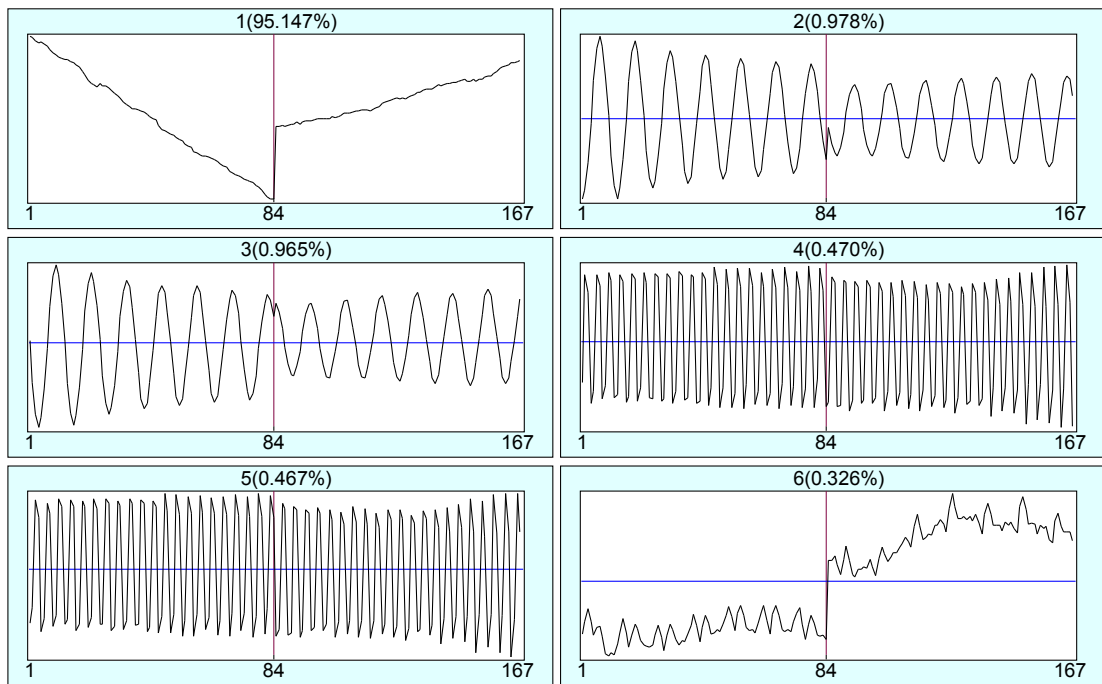


Рис. 5. MSSA: факторные вектора

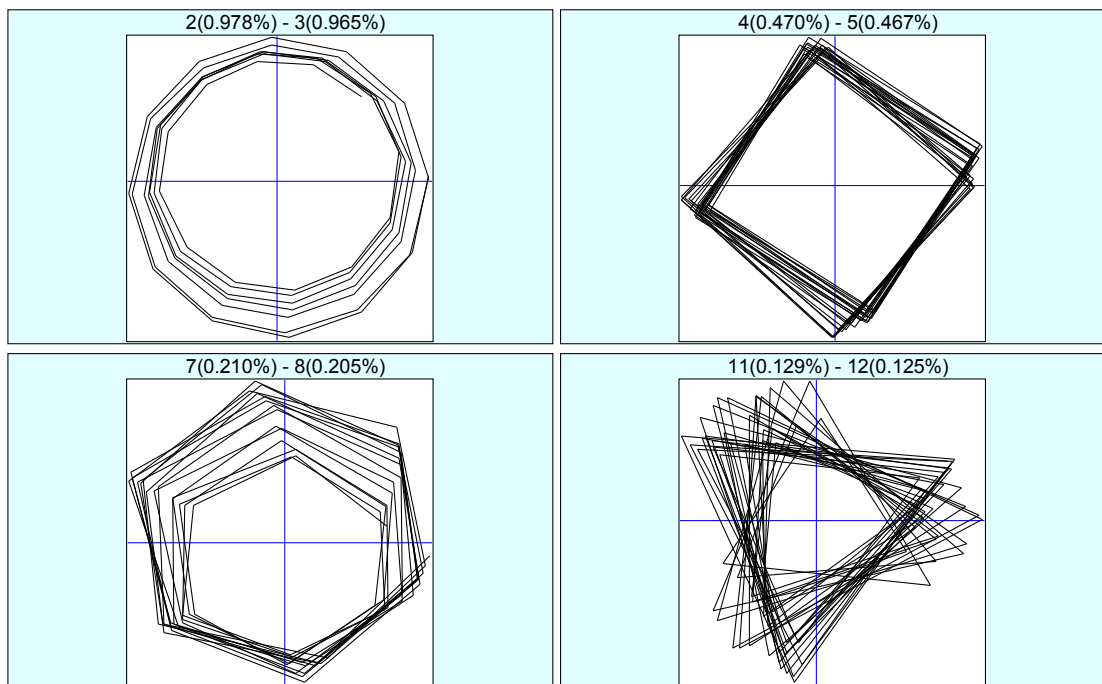


Рис. 6. MSSA: двумерные скаттерграммы собственных векторов

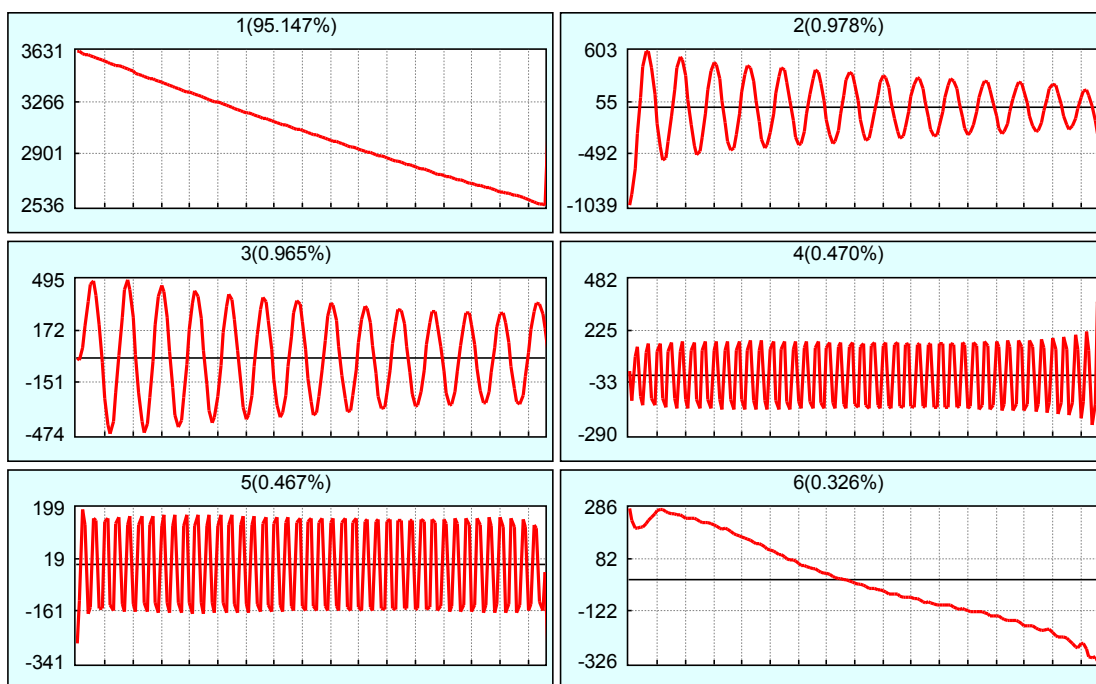


Рис. 7. MSSA-разложение ряда FORT

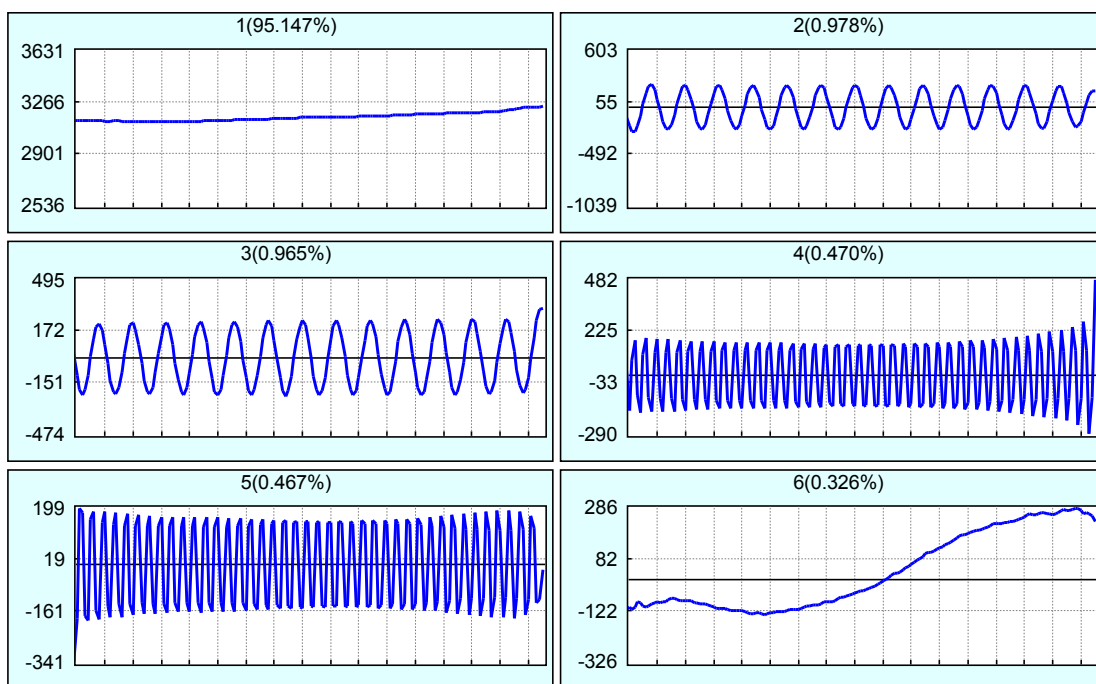


Рис. 8. MSSA-разложение ряда DRY

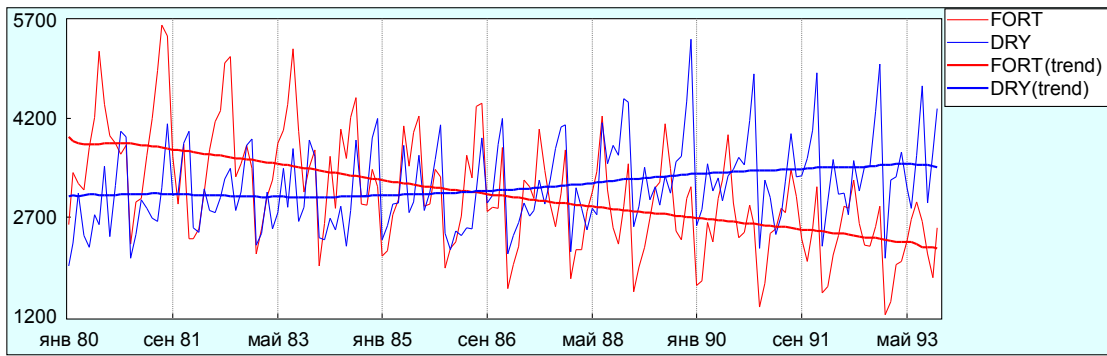


Рис. 9. MSSA: тренды

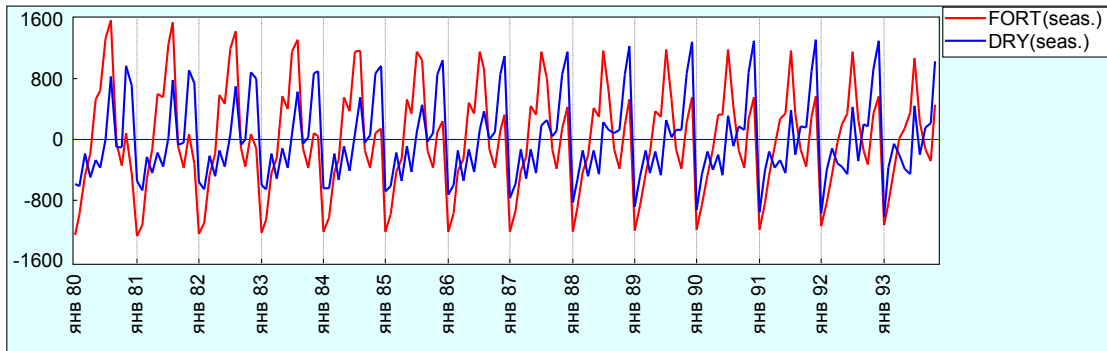


Рис. 10. MSSA: сезонные компоненты

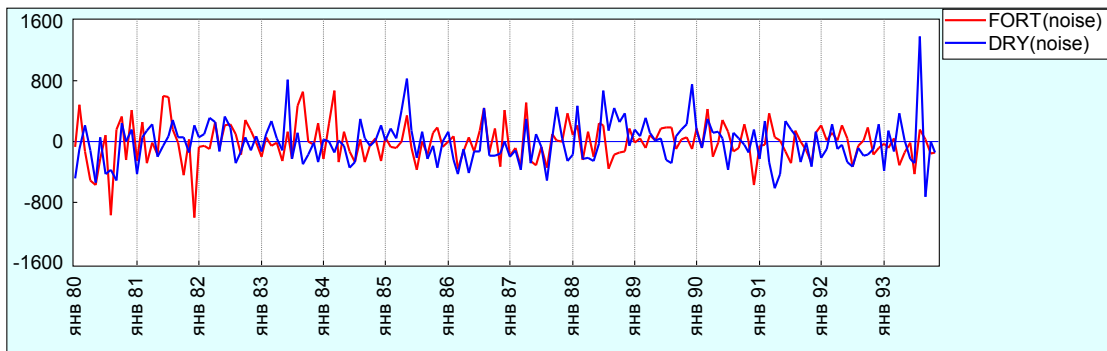


Рис. 11. MSSA: шумовые компоненты

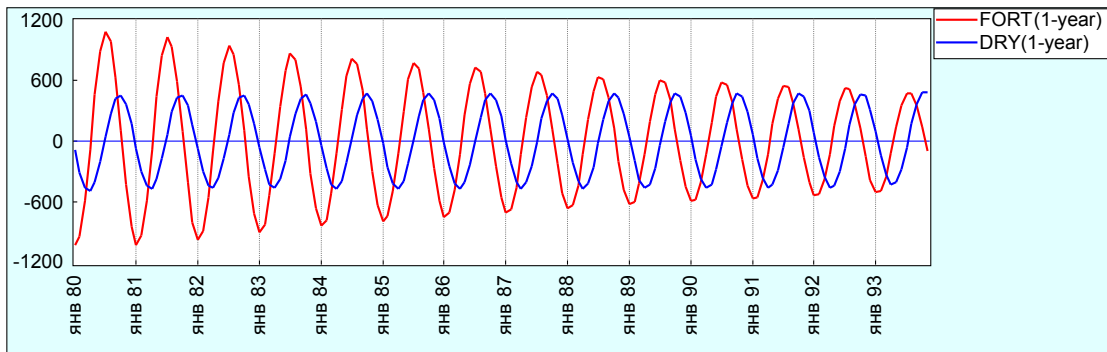


Рис. 12. MSSA: годовая периодичность, сдвиг на четверть периода

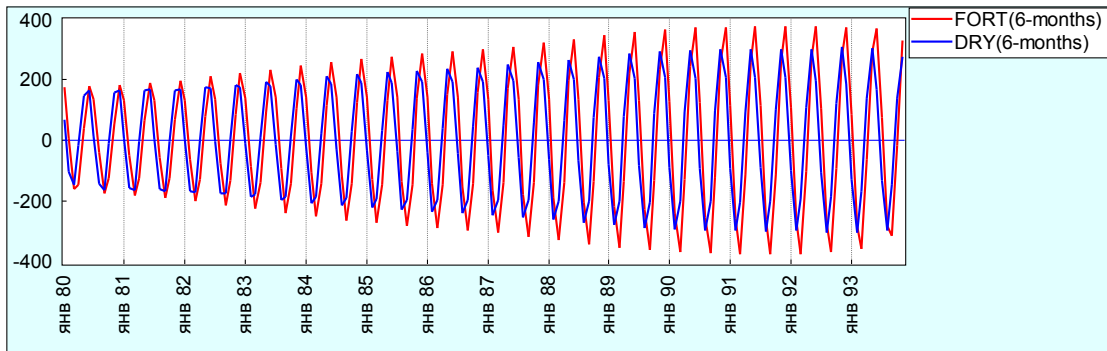


Рис. 13. MSSA: полугодовая периодичность, сдвиг на четверть периода

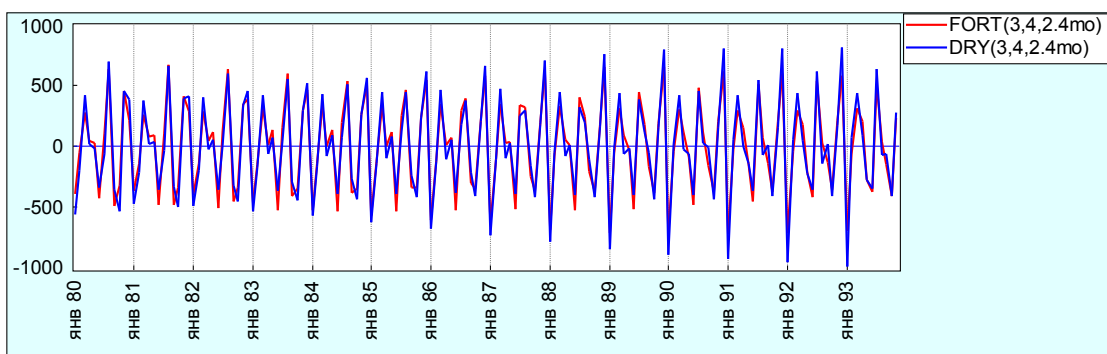


Рис. 14. MSSA: сумма 3, 4, 2.4 – месячных периодичностей, совпадение фаз

Список литературы

1. Broomhead D.S., King G.P. Extracting qualitative dynamics from experimental data // *Physica D*. 1986. Vol. 20. С. 217-236.
2. Broomhead D.S., King G.P. On the qualitative analysis of experimental dynamical systems // *Nonlinear Phenomena and Chaos* / Ed. by S. Sarkar. Bristol: Adam Hilger. 1986. P. 113-144.
3. Главные компоненты временных рядов: метод «Гусеница» / Под. ред. Д.Л. Данилова, А.А. Жиглявского. Санкт-Петербург: Изд-во СПбГУ. 1997. 307 с. <http://www.gistatgroup.com/gus/>
4. Elsner J., Tsonis A. Singular Spectrum Analysis. A New Tool in Time Series Analysis. New York: Plenum Press, 1996. 163 p.
5. Golyandina N., Nekrutkin V., Zhigljavsky A. Analysis of Time Series Structure: SSA and Related Techniques. Boca Raton: Chapman & Hall/CRC. 2001. 305 p.
6. Plaut G., Vautard R. Spells of low-frequency oscillations and weather regimes in the northern hemisphere // *Journal of the Atmospheric Sciences*. 1994. Vol. 51. P. 210-236.
7. Голяндина Н.Э., Данилов Д.Л. Обработка многомерных временных рядов с помощью метода «Гусеница» // Главные компоненты временных рядов: метод «Гусеница» / Под ред. Д.Л. Данилова, А.А. Жиглявского. 1997. Санкт-Петербург: Изд-во СПбГУ, С. 105-131.
8. Keppenne C., Lall U. Complex singular spectrum analysis and multivariate adaptive regression splines applied to forecasting the southern oscillation // *Exp. LongLead Forcst. Bull.* 1996. <http://www.cpc.ncep.noaa.gov/products/predictions/experimental/bulletin/Mar96/article13.html>
9. Time Series Data Library. <http://www-personal.buseco.monash.edu.au/~hyndman/TSDL/>