

Санкт-Петербургский государственный университет

Н. Э. Голяндина

**Метод «Гусеница»-SSA:
анализ временных рядов**

У ч е б н о е п о с о б и е

Санкт-Петербург
2004

УДК 519.2+681.3
ББК 22.172
Г60

Рецензенты: доц. В.В. Некруткин (С.-Петербург. гос. ун-т),
ст. научн. сотр. В.Н. Солнцев (С.-Петербург. гос. ун-т)

*Печатается по постановлению
Редакционно-издательского совета
С.-Петербургского государственного университета*

Голяндина Н.Э.

Г60 Метод «Гусеница»-SSA: анализ временных рядов: Учеб.
пособие. — СПб., 2004. — 76 с.

В основу учебного пособия положен курс лекций «Главные компоненты временных рядов», читаемый на кафедре статистического моделирования математико-механического факультете СПбГУ. Пособие содержит материалы части курса, посвященной анализу временных рядов с помощью метода «Гусеница»-SSA. Кроме описания теории метода, приведены указания по его практическому применению, включая выбор параметров, а также подробно разобранный пример применения метода для анализа реальных данных.

Предназначено для студентов и аспирантов математико-механического факультета.

ББК 22.172

© Н. Э. Голяндина, 2004
© С.-Петербургский
государственный
университет, 2004

Содержание

Введение	3
1. Базовый алгоритм метода «Гусеница»-SSA	5
1.1. Первый этап: разложение	5
1.2. Второй этап: восстановление	6
1.3. Комментарии к алгоритму	7
1.3.1. Вложение	7
1.3.2. Сингулярное разложение	9
1.3.3. Группировка	10
1.3.4. Диагональное усреднение	12
2. Разделимость	14
2.1. Основные определения и свойства	14
2.2. Примеры разделимости	16
2.3. Приближенная и асимптотическая разделимость	17
2.4. Разделимость сигнала и шума	20
2.5. Разделимость и преобразование Фурье	21
2.6. Меры разделимости. Взвешенная корреляция.	25
2.7. Сильная разделимость	26
3. Пространства, порождаемые рядами	28
3.1. Ряды конечного ранга	28
3.2. Примеры рядов конечного ранга и характеристики их сингулярного разложения	29
3.2.1. Экспоненциально-модулированный гармонический ряд	30
3.2.2. Полином	31
4. Выбор параметров	32
4.1. Группировка	32
4.1.1. Общие правила группировки	32
4.1.2. Группировка для выделения тренда и сглаживания	33
4.1.3. Группировка для выделения колебаний	34
4.1.4. Группировка для нахождения полного разложения ряда	39
4.1.5. Дополнительные характеристики, используемые при группировке	40

4.2.	Выбор длины окна	42
4.2.1.	Общие принципы	42
4.2.2.	Выбор длины окна для выделения тренда и сглаживания	44
4.2.3.	Длина окна для выделения периодик	45
4.2.4.	Длина окна и полное разложение ряда	47
4.2.5.	Дополнительная информация для выбора дли- ны окна	47
5.	Модификации базового метода SSA	49
5.1.	Центрирование в SSA	49
5.2.	Стационарные временные ряды и теплицев вариант SSA	53
5.3.	Близкие собственные числа и последовательный ме- тод SSA	54
6.	Приложение А: сингулярное разложение матриц	56
6.1.	Существование и единственность	56
6.2.	Матричная форма сингулярного разложения	58
6.3.	Ортогональные матричные разложения и SVD	60
6.4.	Оптимальные свойства сингулярного разложения	61
7.	Приложение В: пример анализа временного ряда	64
	Литература	74

Введение

Учебное пособие посвящено мощному и быстро развивающемуся методу анализа временных рядов «Гусеница»-SSA. Метод разрабатывался независимо в России и в Великобритании и США под разными именами, «Гусеница» и SSA (Singular Spectrum Analysis), отсюда и его полное название. Описание метода, его применений и ссылки на литературу можно найти в книгах [1, 2, 3]. Отечественная история развития метода вместе со ссылками на соответствующую литературу кратко изложена в [2]. Ссылки на зарубежную литературу можно найти в [1, 3]. При изложении мы будем придерживаться обозначений и понятий книги [3], а также будем приводить без ссылок и доказательств содержащиеся там результаты.

Анализ временных рядов является очень широкой областью. Поэтому прокомментируем, в чем именно заключается специфика рассматриваемого метода.

Опишем кратко, как работает метод (строгое описание алгоритма содержится в разделе 1). Для анализа временного ряда выбирается целый параметр L ; назовем его «длина окна». Параметр L может выбираться достаточно произвольно. При достаточно большой длине ряда и достаточно большом L результаты не будут зависеть от длины окна. Затем на основе ряда строится траекторная матрица, столбцами которой являются скользящие отрезки ряда длины L : с первой точки по L -ю, со второй по $(L + 1)$ -ю и т. д. Следующий шаг — это сингулярное разложение траекторной матрицы в сумму элементарных матриц. Каждая элементарная матрица задается набором из собственного числа и двух сингулярных векторов — собственного и факторного.

Предположим, что исходный временной ряд является суммой нескольких рядов. Теоретические результаты позволяют при некоторых условиях определить по виду собственных чисел, собственных и факторных векторов, что это за слагаемые и какой набор элементарных матриц соответствует каждому из них. Суммируя элементарные матрицы внутри каждого набора и затем переходя от результирующих матриц к ряду, мы получаем разложение ряда на аддитивные слагаемые, например, на сумму тренда, периодики и шума или на сумму низкочастотной и высокочастотной составляющих. Возможность разбить совокупность элементарных матриц на группы, соответствующие интерпретируемым аддитивным со-

ставляющим ряда, тесно связана с понятием *разделимости* рядов, которое будет рассмотрено в разделе 2.

Таким образом, целью метода является разложение временного ряда на интерпретируемые аддитивные составляющие. При этом метод не требует стационарности ряда, знания модели тренда, а также сведений о наличии в ряде периодических составляющих и их периодах. При таких слабых предположениях метод «Гусеница»-SSA может решать различные задачи, такие как, например, выделение тренда, обнаружение периодик, сглаживание ряда, построение полного разложения ряда в сумму тренда, периодик и шума.

Платой за такой широкий спектр возможностей при достаточно слабых предположениях является, во-первых, существенно неавтоматическая группировка компонент сингулярного разложения траекторной матрицы ряда для получения составляющих исходного ряда. Во-вторых, отсутствие модели не позволяет проверить гипотезы о наличии в ряде той или иной составляющей (этот недостаток объективно присущ непараметрическим методам). Для проверки подобных гипотез требуется построение модели, которое, в свою очередь, может быть проведено на основе информации, получаемой с помощью метода «Гусеница»-SSA. Отметим также, что рассматриваемый непараметрический метод позволяет получить результаты, часто лишь незначительно менее точные, чем многие параметрические методы при анализе ряда с известной моделью.

Остановимся кратко на содержании учебного пособия. Первый раздел содержит описание алгоритма метода и комментарии к нему. В разделах 2 и 3 приведены теоретические сведения о методе. В разделе 4 находится подробное описание выбора параметров метода, которое основано на приведенных перед этим теоретических результатах. Материалы этого раздела важны для практического применения метода. Следующий, пятый, раздел является дополнительным и содержит модификации базового алгоритма, которые, в зависимости от вида анализируемых временных рядов, могут позволить получить лучшие результаты. Приложение А, в которое помещены общие сведения о сингулярном разложении матриц, добавлено в пособие, так как этап сингулярного разложения является ключевым в методе «Гусеница»-SSA. Чтобы продемонстрировать, как описанный метод может быть применен к реальным временным рядам, а также для того, чтобы проиллюстрировать теоретические результаты, в приложение В включен подробно разобранный пример анализа временного ряда методом «Гусеница»-SSA.

1. Базовый алгоритм метода «Гусеница»-SSA

Пусть $N > 2$. Рассмотрим вещественнозначный временной ряд $F = (f_0, \dots, f_{N-1})$ длины N . Будем предполагать, что ряд F — ненулевой, т. е. существует, по крайней мере, одно i , такое что $f_i \neq 0$. Обычно считается, что $f_i = f(i\Delta)$ для некоторой функции $f(t)$, где t — время, а Δ — некоторый временной интервал, однако это не будет играть особой роли в дальнейшем. Более того, числа $0, \dots, N-1$ могут быть интерпретированы не только как дискретные моменты времени, но и как некоторые метки, имеющие линейно-упорядоченную структуру.

Нумерация значений временного ряда начинается с $i = 0$, а не стандартно с $i = 1$ только из-за удобства обозначений.

Базовый алгоритм состоит из двух дополняющих друг друга этапов, разложения и восстановления.

1.1. Первый этап: разложение

Шаг 1. Вложение

Процедура вложения переводит исходный временной ряд в последовательность многомерных векторов.

Пусть L — некоторое целое число (*длина окна*), $1 < L < N$. Процедура вложения образует $K = N - L + 1$ векторов вложения

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

имеющих размерность L . Если нам нужно будет подчеркнуть размерность X_i , то мы будем называть их векторами L -вложения.

L -Траекторная матрица (или просто *траекторная матрица*) ряда F

$$\mathbf{X} = [X_1 : \dots : X_K]$$

состоит из векторов вложения в качестве столбцов.

Другими словами, траекторная матрица — это матрица

$$\mathbf{X} = (x_{ij})_{i,j=1}^{L,K} = \begin{pmatrix} f_0 & f_1 & f_2 & \dots & f_{K-1} \\ f_1 & f_2 & f_3 & \dots & f_K \\ f_2 & f_3 & f_4 & \dots & f_{K+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_{L-1} & f_L & f_{L+1} & \dots & f_{N-1} \end{pmatrix}. \quad (1)$$

Очевидно, что $x_{ij} = f_{i+j-2}$ и матрица \mathbf{X} имеет одинаковые элементы на «диагоналях» $i + j = \text{const}$. Таким образом, траекторная матрица является *ганкелевой*. Существует взаимно-однозначное соответствие между ганкелевыми матрицами размерности $L \times K$ и рядами длины $N = L + K - 1$.

Шаг 2. Сингулярное разложение

Результатом этого шага является сингулярное разложение (SVD = Singular Value Decomposition) траекторной матрицы ряда.

Пусть $\mathbf{S} = \mathbf{X}\mathbf{X}^T$. Обозначим $\lambda_1, \dots, \lambda_L$ *собственные числа* матрицы \mathbf{S} , взятые в неубывающем порядке ($\lambda_1 \geq \dots \geq \lambda_L \geq 0$) и U_1, \dots, U_L — ортонормированную систему *собственных векторов* матрицы \mathbf{S} , соответствующих собственным числам.

Пусть $d = \max\{i : \lambda_i > 0\}$. Если обозначить $V_i = \mathbf{X}^T U_i / \sqrt{\lambda_i}$, $i = 1, \dots, d$, то сингулярное разложение матрицы \mathbf{X} может быть записано как

$$\mathbf{X} = \mathbf{X}_1 + \dots + \mathbf{X}_d, \quad (2)$$

где $\mathbf{X}_i = \sqrt{\lambda_i} U_i V_i^T$. 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 (2).

1.2. Второй этап: восстановление

Шаг 3. Группировка

На основе разложения (2) 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 (2) может быть записано в сгруппированном виде

$$\mathbf{X} = \mathbf{X}_{I_1} + \dots + \mathbf{X}_{I_m}. \quad (3)$$

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

Шаг 4. Диагональное усреднение

На последнем шаге базового алгоритма каждая матрица сгруппированного разложения (3) переводится в новый ряд длины N .

Пусть \mathbf{Y} — некоторая $L \times K$ матрица с элементами y_{ij} , где $1 \leq i \leq L$, $1 \leq j \leq K$. Положим $L^* = \min(L, K)$, $K^* = \max(L, K)$ и $N = L + K - 1$. Пусть $y_{ij}^* = y_{ij}$, если $L < K$, и $y_{ij}^* = y_{ji}$ иначе. Диагональное усреднение переводит матрицу \mathbf{Y} в ряд g_0, \dots, g_{N-1} по формуле

$$g_k = \begin{cases} \frac{1}{k+1} \sum_{m=1}^{k+1} y_{m, k-m+2}^* & \text{для } 0 \leq k < L^* - 1, \\ \frac{1}{L^*} \sum_{m=1}^{L^*} y_{m, k-m+2}^* & \text{для } L^* - 1 \leq k < K^*, \\ \frac{1}{N-k} \sum_{m=k-K^*+2}^{N-K^*+1} y_{m, k-m+2}^* & \text{для } K^* \leq k < N. \end{cases} \quad (4)$$

Выражение (4) 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. Заметим, что если матрица \mathbf{Y} является траекторной матрицей некоторого ряда (h_0, \dots, h_{N-1}) (другими словами, если матрица \mathbf{Y} является ганкелевой), то $g_i = h_i$ для всех i .

Применяя диагональное усреднение (4) к результирующим матрицам \mathbf{X}_{I_k} , мы получаем ряды $\tilde{F}^{(k)} = (\tilde{f}_0^{(k)}, \dots, \tilde{f}_{N-1}^{(k)})$, и, следовательно, исходный ряд (f_0, \dots, f_{N-1}) раскладывается в сумму m рядов:

$$f_n = \sum_{k=1}^m \tilde{f}_n^{(k)}. \quad (5)$$

1.3. Комментарии к алгоритму

Формальное описание базового алгоритма SSA требует некоторых пояснений. Коротко прокомментируем шаги алгоритма.

1.3.1. Вложение

Процедура вложения может рассматриваться как отображение, переводящее одномерный временной ряд $F = (f_0, \dots, f_{N-1})$

в многомерный ряд X_1, \dots, X_K , состоящий из векторов вложения $X_i = (f_{i-1}, \dots, f_{i+L-2})^T \in \mathbb{R}^L$, где $K = N - L + 1$.

Единственным параметром этого шага алгоритма является длина окна L , целое число, такое что $2 \leq L \leq N - 1$.

Вложение является стандартной процедурой анализа временных рядов. После вложения одномерного ряда в многомерное пространство дальнейший анализ зависит от цели исследования.

Для специалистов в области динамических систем стандартной является техника получения эмпирического распределения попарных расстояний между векторами вложения X_i и X_j и последующее вычисление так называемой корреляционной размерности ряда. Эта размерность связана с фрактальной размерностью динамической системы, порождаемой рядом. Заметим, что в этом подходе L должно быть относительно небольшим, а K — очень большим (формально, $K \rightarrow \infty$).

Если L достаточно большое, то можно рассмотреть каждый вектор X_i L -вложения как отдельный ряд и исследовать динамику различных характеристик исходного ряда. Самый простой пример — это хорошо известный метод «скользящего среднего».

Другой пример можно предложить, если исходный ряд является (локально) стационарным. Тогда мы можем разложить каждый вектор вложения X_i по некоторому фиксированному базису (например, базису разложения Фурье или вейвлет-базису) и изучать динамику такого разложения.

Метод SSA ближе к динамическим методам и параметр L должен выбираться, вообще говоря, достаточно большим, чтобы каждый из векторов L -вложения содержал в себе существенную часть поведения исходного ряда $F = (f_0, \dots, f_{N-1})$.

Результатом этого шага является именно матрица \mathbf{X} , а не набор ее столбцов (векторов вложения), так как упорядоченность элементов внутри каждого вектора, как и упорядоченность самих векторов, существенна в методе SSA.

Заметим, что траекторная матрица (1) обладает очевидным свойством симметрии: транспонированная матрица \mathbf{X}^T также является траекторной матрицей того же ряда (f_0, \dots, f_{N-1}) с длиной окна, равной $K = N - L + 1$ вместо L .

1.3.2. Сингулярное разложение

Сингулярное разложение является основой математической части метода. Большинство свойств сингулярного разложения справедливо для произвольной матрицы $L \times K$, но ганкелева структура траекторной матрицы добавляет ряд особых свойств.

Перечислим основные определения и свойства, более подробно описанные в приложении А.

Определим сингулярное разложение произвольной ненулевой $L \times K$ -матрицы $\mathbf{X} = [X_1 : \dots : X_K]$ как разложение \mathbf{X} в виде

$$\mathbf{X} = \sum_{i=1}^d \sqrt{\lambda_i} U_i V_i^T, \quad (6)$$

где λ_i ($i = 1, \dots, L$) — собственные числа матрицы $\mathbf{S} = \mathbf{X}\mathbf{X}^T$, упорядоченные по убыванию,

$$d = \max\{i, \text{такие что } \lambda_i > 0\} = \text{rank } \mathbf{X},$$

$\{U_1, \dots, U_d\}$ — соответствующая система ортонормированных собственных векторов матрицы \mathbf{S} и $V_i = \mathbf{X}^T U_i / \sqrt{\lambda_i}$.

В стандартной терминологии $\sqrt{\lambda_i}$ называются *сингулярными числами*; U_i и V_i — *левыми* и *правыми сингулярными векторами* матрицы \mathbf{X} соответственно. Набор $(\sqrt{\lambda_i}, U_i, V_i)$ называется *i -й собственной тройкой* матрицы \mathbf{X} . Заметим, что строки и столбцы траекторной матрицы являются отрезками исходного ряда. Поэтому левый и правый сингулярные векторы также имеют временную структуру и могут рассматриваться как временные ряды.

Если определить $\mathbf{X}_i = \sqrt{\lambda_i} U_i V_i^T$, то представление (6) может быть переписано в виде (2), т. е. как представление \mathbf{X} в виде суммы элементарных матриц. При этом собственное число λ_i является характеристикой вклада матрицы \mathbf{X}_i в разложение.

Так как сингулярное разложение имеет дело с матрицей \mathbf{X} как единым целым, оно не инвариантно относительно перестановки ее столбцов X_1, \dots, X_K . Более того, равенство (6) показывает, что сингулярное разложение обладает свойством симметричности: V_1, \dots, V_d образуют ортонормированную систему собственных векторов матрицы $\mathbf{X}^T \mathbf{X}$, соответствующих тем же самым собственным числам λ_i .

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

Сингулярное разложение является единственным разложением по биортогональному базису (с точностью до одновременной замены знака у сингулярных векторов и некоторой свободы в выборе собственных векторов при наличии кратных собственных чисел).

Отметим также, что сингулярное разложение — минимальное разложение в сумму элементарных матриц (ранга 1) и обладает рядом оптимальных свойств, описанных в приложении.

Свойства оптимальности и статистические аналогии (столбцы матрицы \mathbf{X} рассматриваются как L -мерная выборка объема K) приводят к следующей терминологии. Векторы U_i будем называть *собственными*, векторы 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*.

1.3.3. Группировка

Шаг группировки очень тесно связан с понятием делимости двух рядов, подробно описанным в следующем разделе. Поэтому коротко остановимся на главных моментах.

Для простоты положим $m = 2$ и возьмем $I_1 = I = \{i_1, \dots, i_r\}$ и $I_2 = \{1, \dots, d\} \setminus I$, где $1 \leq i_1 < \dots < i_r \leq d$.

Целью этапа группировки является разделение аддитивных компонент ряда. Предположим, что временной ряд F равен сумме двух рядов, $F^{(1)}$ и $F^{(2)}$, т. е. $f_i = f_i^{(1)} + f_i^{(2)}$ для $i = 0, \dots, N-1$. Зафиксируем длину окна L и обозначим \mathbf{X} , $\mathbf{X}^{(1)}$ и $\mathbf{X}^{(2)}$ соответствующие L -траекторные матрицы рядов F , $F^{(1)}$ и $F^{(2)}$.

Рассмотрим сингулярное разложение (2) траекторной матрицы \mathbf{X} . Мы будем говорить, что ряды $F^{(1)}$ и $F^{(2)}$ (слабо) *разделимы разложением* (2), если существует такой набор индексов $I \subset \{1, \dots, d\}$, что $\mathbf{X}^{(1)} = \sum_{i \in I} \mathbf{X}_i$ и, следовательно, $\mathbf{X}^{(2)} = \sum_{i \notin I} \mathbf{X}_i$.

В случае делимости вклад $\mathbf{X}^{(1)}$, первой компоненты в разложении $\mathbf{X} = \mathbf{X}^{(1)} + \mathbf{X}^{(2)}$, естественно выражается как доля соответствующих собственных чисел: $\sum_{i \in I} \lambda_i / \sum_{i=1}^L \lambda_i$.

Посмотрим на понятие делимости с разных сторон. Зафиксируем набор индексов $I = I_1$ и рассмотрим соответствующую результирующую матрицу \mathbf{X}_{I_1} . Если эта матрица и, следовательно, матрица $\mathbf{X}_{I_2} = \mathbf{X} - \mathbf{X}_{I_1}$ ганкелевы, то они обязательно будут являться траекторными матрицами некоторых временных рядов, которые делимы данным разложением (2). Более того, если матрицы \mathbf{X}_{I_1} и \mathbf{X}_{I_2} близки по структуре к ганкелевым, то найдутся такие ряды $F^{(1)}$ и $F^{(2)}$, что $F = F^{(1)} + F^{(2)}$ и траекторные матрицы этих рядов будут близки к \mathbf{X}_{I_1} и \mathbf{X}_{I_2} соответственно (проблема нахождения этих рядов обсуждается ниже). В этом случае мы будем говорить, что ряды *приближенно делимы*. Таким образом, целью шага группировки (т. е. процедуры объединения индексов в группы) является идентификация нескольких групп I_1, \dots, I_m , таких что матрицы $\mathbf{X}_{I_1}, \dots, \mathbf{X}_{I_m}$ удовлетворяют (3) и близки по структуре к ганкелевым.

Прокомментируем шаг группировки с точки зрения многомерной геометрии. Пусть $\mathbf{X} = [X_1 : \dots : X_K]$ — траекторная матрица временного ряда F , $F = F^{(1)} + F^{(2)}$, и ряды $F^{(1)}$ и $F^{(2)}$ делимы разложением (2), что соответствует разделению множества индексов $\{1, \dots, d\}$ на I и $\{1, \dots, d\} \setminus I$. Разложение (3) с $m = 2$ означает, что U_1, \dots, U_d , базис траекторного пространства $\mathfrak{L}^{(L)}$, разбивается на две группы базисных векторов. Это соответствует представлению $\mathfrak{L}^{(L)}$ как произведения двух ортогональных линейных подпространств (*собственных подпространств*) $\mathfrak{L}^{(L,1)} = \mathfrak{L}(U_i, i \in I)$ и $\mathfrak{L}^{(L,2)} = \mathfrak{L}(U_i, i \notin I)$, порожденных $U_i, i \in I$, и $U_i, i \notin I$, соответственно. Делимость двух рядов $F^{(1)}$ и $F^{(2)}$ означает, что матрица \mathbf{X}_I , чьи столбцы являются проекциями векторов вложения X_1, \dots, X_K на собственное пространство $\mathfrak{L}^{(L,1)}$, — это в точности траекторная матрица ряда $F^{(1)}$.

Несмотря на то что несколько формальных критериев делимости будут представлены в следующем разделе, процедуру разделения членов сингулярного разложения на группы (шаг группировки) трудно полностью формализовать. Эта процедура основана на анализе сингулярных векторов U_i, V_i и собственных чисел λ_i в сингулярном разложении (2) и (6). Принципы и методы идентификации компонент разложения подробно описаны в разделе 4.

Так как каждая матричная компонента сингулярного разложения полностью определяется соответствующей сингулярной тройкой, мы будем говорить о группировке собственных троек, а не элементарных матриц \mathbf{X}_i .

Заметим, что в случае суммы двух рядов ($m = 2$), рассмотренной выше, более естественно оказывается говорить не о разделении компонент, а о выделении одной компоненты (например, в задаче выделения сигнала). В этом случае нас интересует только одна группа индексов, а именно I .

1.3.4. Диагональное усреднение

Если компоненты ряда разделимы и собственные тройки сгруппированы правильно, то все матрицы в разложении (3) являются ганкелевыми и мы немедленно получаем разложение (5) исходного ряда: для каждого k и n , $\tilde{f}_n^{(k)}$ равняется любому из элементов $x_{ij}^{(k)}$ на побочной диагонали

$$\{(i, j), \text{ такие что } i + j = n + 2\}$$

матрицы \mathbf{X}_{I_k} .

На практике, однако, такая ситуация не реалистична. В общем случае, никакая из побочных диагоналей не будет состоять из одинаковых элементов. Поэтому возникает необходимость в формальной процедуре преобразования произвольной матрицы в ганкелеву и, следовательно, во временной ряд. В качестве такой процедуры рассматривается процедура *диагонального усреднения*, которая задает значения ряда $\tilde{F}^{(k)}$ как среднее значение элементов матрицы \mathbf{X}_{I_k} вдоль соответствующей побочной диагонали.

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 (матричная норма равняется корню из суммы квадратов всех элементов матрицы). Будем говорить, что матрица $\tilde{\mathbf{Y}}$ получена из матрицы \mathbf{Y} с помощью *ганкелизации* и обозначать соответствующий оператор ганкелизации как \mathcal{H} .

Применив процедуру ганкелизации ко всем матричным компонентам разложения (3), мы получим другое разложение:

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

где $\tilde{\mathbf{X}}_{I_i} = \mathcal{H}\mathbf{X}_{I_i}$.

Правильные длина окна и группировка приводят к разложению (3), в котором результирующие матрицы \mathbf{X}_{I_k} являются почти ганкелевыми, что соответствует приближенной разделимости.

Так как все матрицы в правой части разложения (7) являются ганкелевыми, то каждая матрица однозначно определяет временной ряд $\tilde{F}^{(k)}$ и мы, таким образом, получаем (5), разложение исходного ряда.

Процедуру вычисления временных рядов $\tilde{F}^{(k)}$ (т. е. построение группы I_k плюс диагональное усреднение матрицы \mathbf{X}_{I_k}) будем называть *восстановлением компоненты ряда $\tilde{F}^{(k)}$ по собственным тройкам* с индексами из I_k .

2. Разделимость

Пусть временной ряд F является суммой m рядов и задачей является нахождение этих слагаемых. В результате базового алгоритма SSA мы также получаем m рядов. Возникает вопрос, в каких случаях мы можем так выбрать параметр алгоритма L и так сгруппировать собственные тройки, чтобы в SSA-разложении получить m исходных рядов; т. е. можно ли восстановить аддитивные составляющие ряда F с помощью сингулярного разложения его траекторной матрицы? Для ответа на этот вопрос нам необходимо ввести понятие разделимости рядов.

2.1. Основные определения и свойства

Пусть $F_N^{(1)}$ и $F_N^{(2)}$ — ряды длины N и $F_N = F_N^{(1)} + F_N^{(2)}$. При выборе длины окна, равной L , каждый из рядов $F_N^{(1)}$, $F_N^{(2)}$ и F_N порождает L -траекторную матрицу: $\mathbf{X}^{(1)}$, $\mathbf{X}^{(2)}$ и \mathbf{X} .

Обозначим $\mathfrak{L}^{(L,1)}$ и $\mathfrak{L}^{(L,2)}$ линейные пространства, порожденные столбцами траекторных матриц $\mathbf{X}^{(1)}$ и $\mathbf{X}^{(2)}$ соответственно. Аналогичные обозначения $\mathfrak{L}^{(K,1)}$ и $\mathfrak{L}^{(K,2)}$ будем использовать для пространств, порожденных столбцами транспонированных матриц $(\mathbf{X}^{(1)})^T$ и $(\mathbf{X}^{(2)})^T$, $K = N - L + 1$ (можно сказать, что они порождены строками траекторных матриц).

Определение 2.1. Будем говорить, что ряды $F_N^{(1)}$ и $F_N^{(2)}$ *слабо L -разделимы* (или, для краткости, просто слабо разделимы), если $\mathfrak{L}^{(L,1)} \perp \mathfrak{L}^{(L,2)}$ и $\mathfrak{L}^{(K,1)} \perp \mathfrak{L}^{(K,2)}$.

Приведем утверждения, иллюстрирующие данное определение.

Положим $F_N^{(1)} = (f_0^{(1)}, \dots, f_{N-1}^{(1)})$ и $F_N^{(2)} = (f_0^{(2)}, \dots, f_{N-1}^{(2)})$.

Предложение 2.1. Пусть $K = N - L + 1$. Ряды $F_N^{(1)}$ и $F_N^{(2)}$ слабо разделимы тогда и только тогда, когда

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

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

Таким образом, необходимым и достаточным условием слабой разделимости является ортогональность отрезков длины L ряда $F^{(1)}$ и отрезков длины L ряда $F^{(2)}$, а также отрезков длины K ряда $F^{(1)}$ и отрезков длины K ряда $F^{(2)}$.

Предложение 2.2. Пусть ряды $F_N^{(1)}$ и $F_N^{(2)}$ слабо L -разделимы. Рассмотрим некоторые сингулярные разложения траекторных матриц $\mathbf{X}^{(1)}$ и $\mathbf{X}^{(2)}$:

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

Тогда разложение

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

также является сингулярным.

Это утверждение показывает, что в случае разделимости представление ряда $F_N = F_N^{(1)} + F_N^{(2)}$ является естественным с точки зрения сингулярного разложения его траекторной матрицы \mathbf{X} .

Следствие 2.1. Пусть ряды $F_N^{(1)}$ и $F_N^{(2)}$ слабо L -разделимы. Тогда существует такое сингулярное разложение траекторной матрицы \mathbf{X} ряда F_N , что его можно разбить на две части, являющиеся сингулярными разложениями траекторных матриц рядов $F_N^{(1)}$ и $F_N^{(2)}$.

В результате выполнения этапа разложения базового алгоритма SSA мы получаем некоторое разложение траекторной матрицы ряда, которое вовсе не обязано соответствовать разделимости двух рядов. Поэтому необходимо усилить понятие разделимости.

Определение 2.2. Если ряды $F_N^{(1)}$ и $F_N^{(2)}$ слабо L -разделимы и множество собственных чисел разложения траекторной матрицы одного ряда не пересекается с множеством собственных чисел сингулярного разложения второго ряда ($\lambda_{1k} \neq \lambda_{2m}$ для всех k и m), то будем говорить, что ряды $F_N^{(1)}$ и $F_N^{(2)}$ *сильно L -разделимы*.

Предложение 2.3. Пусть ряды $F_N^{(1)}$ и $F_N^{(2)}$ сильно L -разделимы. Тогда любое сингулярное разложение траекторной матрицы X ряда F_N можно разбить на две части, являющиеся сингулярными разложениями траекторных матриц рядов $F_N^{(1)}$ и $F_N^{(2)}$.

В дальнейшем будем рассматривать в основном слабую разделимость и опускать слово «слабая», если не нужно это подчеркнуть.

2.2. Примеры разделимости

Рассмотрим простые тестовые примеры. В качестве набора рядов возьмем константный ряд, экспоненту, синус/косинус, экспоненту, умноженную на косинус, и полином (линейную функцию). Полученные результаты очевидным образом переносятся на разделимость линейных комбинаций из этих рядов.

Сначала приведем таблицу, в которой знаком + отмечены пары рядов, для которых существуют параметры функций и параметры метода L и $K = N - L + 1$, при которых они разделимы. Табл. 1 показывает, что условия разделимости являются достаточно жесткими.

Таблица 1. Точная разделимость

	const	cos	exp	exp cos	$ak + b$
const	–	+	–	–	–
cos	+	+	–	–	–
exp	–	–	–	+	–
exp cos	–	–	+	+	–
$ak + b$	–	–	–	–	–

Приведем примеры условий и параметров, при которых пары рядов, помеченные в табл. 1 плюсами, будут разделимы.

Пример 2.1. Отделимость от константного ряда

Рассмотрим константный временной ряд $F_N^{(1)}$ с $f_n^{(1)} \equiv c \neq 0$, $0 \leq n \leq N-1$. Тогда необходимые и достаточные условия отделимости от него ряда $F_N^{(2)}$ следующие:

- 1) Ряд $F_N^{(2)}$ имеет целый период T ; L/T и K/T — целые;
- 2) $f_0^{(2)} + \dots + f_{T-1}^{(2)} = 0$.

Например, для $f_n^{(2)} = \cos(2\pi n/T + \phi)$ и длины ряда N , такой что $(N+1)/T$ — целое, ряд $F_N^{(2)}$ будет отделим от константы при выборе длины окна $L < N$, делящейся нацело на T .

Пример 2.2. *Отделимость от экспоненциального ряда*

Пусть $f_n^{(1)} = e^{\alpha n}$. Тогда необходимые и достаточные условия отделимости от него ряда $F_N^{(2)}$ следующие:

- 1) Ряд $e^{\alpha n} F_N^{(2)}$ имеет целый период T ; L/T и K/T — целые;
- 2) $\sum_{m=0}^{T-1} e^{\alpha m} f_m^{(2)} = 0$.

Например, для $f_n^{(2)} = e^{-\alpha n} \cos(2\pi n/T + \phi)$ и длины ряда N , такой что $(N+1)/T$ — целое, ряд $F_N^{(2)}$ будет отделим от константы при выборе длины окна $L < N$, кратной T .

Пример 2.3. *Отделимость от гармонического ряда*

Рассмотрим ряд с элементами $f_n^{(1)} = \cos(2\pi\omega n + \phi)$, предположив, что $0 < \omega < 1/2$ и $L, K > 2$. Положим $T = 1/\omega$.

Приведем пример отделимости от гармонического ряда. Ряд $f_n^{(2)} = \cos(2\pi\tilde{\omega}n + \tilde{\phi})$, $\tilde{T} = 1/\tilde{\omega}$, отделим от ряда $F^{(1)}$, если $\omega \neq \tilde{\omega}$ и L и K (а следовательно, и $N+1$) делятся нацело как на T , так и на \tilde{T} . Естественно, в качестве второго ряда можно рассмотреть и константу.

Пример 2.4. *Отделимость от гармонического ряда с экспоненциальной амплитудой*

Предположим, что $f_n^{(1)} = e^{\alpha n} \cos(2\pi\omega n + \phi)$. Тогда ряд с общим членом $f_n^{(2)} = e^{-\alpha n} \cos(2\pi\tilde{\omega}n + \tilde{\phi})$, $\tilde{T} = 1/\tilde{\omega}$, отделим от ряда $F^{(1)}$, если $\omega \neq \tilde{\omega}$ и L и K (а следовательно, и $N+1$) делятся нацело как на T , так и на \tilde{T} .

Пример 2.5. *Отделимость от линейного ряда*

Пусть $f_n^{(1)} = an + b$ с $a \neq 0$. Как следует из табл. 1, не существует ряда, отделимого от линейного. Такой же результат справедлив и для полиномов старшей степени.

2.3. Приближенная и асимптотическая разделимость

Как уже отмечалось выше, условия разделимости являются слишком жесткими и вряд ли выполнимы в реальных задачах. Поэтому необходимо ввести понятие приближенной разделимости. Так

как определение точной разделимости сводится к требованию ортогональности отрезков рядов, то естественно вводить определение приближенной разделимости через меру ортогональности этих отрезков.

Введем обозначения. Для любого ряда $F_N = (f_0, \dots, f_{N-1})$ положим

$$F_{i,j} = (f_{i-1}, \dots, f_{j-1}), \quad 1 \leq i \leq j < N.$$

Пусть $F_N^{(1)} = (f_0^{(1)}, \dots, f_{N-1}^{(1)})$, $F_N^{(2)} = (f_0^{(2)}, \dots, f_{N-1}^{(2)})$.

Для $i, j \geq 1$ и $M \leq N - 1 - \max(i, j)$ определим

$$\rho_{i,j}^{(M)} = \frac{\left(F_{i,i+M-1}^{(1)}, F_{j,j+M-1}^{(2)} \right)}{\|F_{i,i+M-1}^{(1)}\| \|F_{j,j+M-1}^{(2)}\|} \quad (8)$$

в предположении, что знаменатель положителен. Здесь (\cdot, \cdot) — это обычное скалярное произведение векторов, а $\|\cdot\|$ — евклидова норма. Если знаменатель в (8) равен нулю, то будем предполагать, что $\rho_{i,j}^{(M)} = 0$.

Значение $\rho_{i,j}^{(M)}$ имеет смысл косинуса угла между векторами $F_{i,i+M-1}^{(1)}$ и $F_{j,j+M-1}^{(2)}$. В статистической терминологии величину $\rho_{i,j}^{(M)}$ называют *коэффициентом корреляции*.

Определение 2.3. Ряды $F_N^{(1)}, F_N^{(2)}$ называются (слабо) ε -разделимыми при длине окна L , если

$$\rho^{(L,K)} \stackrel{\text{def}}{=} \max \left(\max_{1 \leq i,j \leq K} |\rho_{i,j}^{(L)}|, \max_{1 \leq i,j \leq L} |\rho_{i,j}^{(K)}| \right) < \varepsilon, \quad (9)$$

$K = N - L + 1$. Если число ε мало, то ряды называются *приближенно разделимыми*.

Замечание 2.1. Приближенная разделимость может возникнуть, например, в следующих случаях:

1) Ряды $F_N^{(1)}$ и $F_N^{(2)}$ возникли в результате небольшого искажения точно разделимых рядов.

2) Ряды $F_N^{(1)}$ и $F_N^{(2)}$ разделимы, но параметры, при которых эта разделимость имеет место, искажены.

Введем теперь понятие асимптотической делимости. Рассмотрим два бесконечных временных ряда, $F^{(1)} = (f_0^{(1)}, \dots, f_n^{(1)}, \dots)$ и $F^{(2)} = (f_0^{(2)}, \dots, f_n^{(2)}, \dots)$. Пусть для каждого $N > 2$ ряды $F_N^{(1)}$ и $F_N^{(2)}$ состоят из первых N членов рядов $F^{(1)}$ и $F^{(2)}$ соответственно. Выбирая последовательность длин окон $1 < L = L(N) < N$, мы получаем последовательность *максимальных коэффициентов корреляции* $\rho_N = \rho^{(L(N), K(N))}$, определенных равенством (9).

Определение 2.4. Если $\rho^{(L(N), K(N))} \rightarrow 0$ при некоторой последовательности $L = L(N)$, $N \rightarrow \infty$, то временные ряды $F^{(1)}$ и $F^{(2)}$ называются *асимптотически $L(N)$ -разделимыми*. Если $F^{(1)}$ и $F^{(2)}$ асимптотически $L(N)$ -разделимы при любом выборе последовательности $L(N)$, таком что $L(N) \rightarrow \infty$ и $K(N) \rightarrow \infty$, то они называются *регулярно асимптотически разделимыми* или просто *асимптотически разделимыми*.

Приведем примеры асимптотической делимости и одновременно исследуем скорость сходимости в (9). Начнем с таблицы, аналогичной табл. 1 (табл. 2).

Таблица 2. Асимптотическая делимость

	const	cos	exp	exp cos	$ak + b$
const	–	+	+	+	–
cos	+	+	+	+	+
exp	+	+	+	+	+
exp cos	+	+	+	+	+
$ak + b$	–	+	+	+	–

Табл. 2 показывает, что асимптотическая делимость имеет место для гораздо более широкого класса рядов, чем точная делимость. Заметим, что если каждое слагаемое в одной сумме рядов асимптотически делимо со всеми слагаемыми другой суммы, то суммы также будут асимптотически делимы.

В табл. 3 приведены условия и скорость асимптотической делимости.

В заголовках строк и столбцов таблицы приведен вид k -го члена ряда. Предполагается, что $\alpha \neq 0$, $\beta \neq 0$, $\omega_1 \in (0, 0.5]$ и $\omega_2 \in (0, 0.5]$. В ячейках таблицы приведены порядки скорости сходимости к нулю максимальных корреляций, определенных в (9),

Таблица 3. Асимптотическая разделимость

	const	$\cos(2\pi\omega_2 k)$	$\exp(\beta k)$
const	—	$1/\min(L, K)$	$1/\sqrt{\min(L, K)}$
$\cos(2\pi\omega_1 k)$	$1/\min(L, K)$	$1/\min(L, K)$ ($\omega_1 \neq \omega_2$)	$1/\sqrt{\min(L, K)}$
$\exp(\alpha k)$	$1/\sqrt{\min(L, K)}$	$1/\sqrt{\min(L, K)}$	$e^{-C \min(L, K)}$ ($\alpha\beta < 0$)

при L и K , стремящихся к бесконечности. В двух ячейках в скобках приведены условия, необходимые и достаточные для асимптотической разделимости. Асимптотические результаты для полинома совпадают с результатами для константы, также одинаковые результаты получаются для экспоненты и косинуса с амплитудой, меняющейся по экспоненте.

Кроме порядка сходимости, для небольших L и K важную роль могут играть параметры рядов.

Приведем два примера:

1. $f_n^{(1)} = \text{const}$, $f_n^{(2)} = e^{\alpha n}$. Если величина $|\alpha|$ близка к нулю, то истинный порядок (9) имеет вид $(|\alpha| \min(L, K))^{-1/2}$. Следовательно, чем меньше α , тем больше должны быть L и K .

2. $f_n^{(1)} = \cos(2\pi\omega_1 n)$, $f_n^{(2)} = \cos(2\pi\omega_2 n)$. Если частоты ω_1 и ω_2 близки, то порядок сходимости имеет вид $C|\omega_1 - \omega_2|/\min(L, K)$. Таким образом, чем ближе частоты гармонических рядов, тем хуже они разделяются.

2.4. Разделимость сигнала и шума

Ранее мы рассматривали шум скорее как детерминированную составляющую ряда (например, сумму гармоник с близкими амплитудами).

Приведем результат, демонстрирующий как можно на вероятностном языке ввести понятие (асимптотической) разделимости.

Пусть $(\Omega, \mathcal{F}, \mathbf{P})$ — вероятностное пространство и пусть $F^{(1)}$, $F^{(2)}$ — две бесконечные последовательности.

Определение 2.5. Если существует такая последовательность $L = L(N)$, что последовательность случайных величин $\rho^{(L,K)}$, определенная в (9), (8), стремится к нулю по вероятности при $N \rightarrow \infty$, то мы будем называть случайные последовательности $F^{(1)}$ и $F^{(2)}$ *стохастически разделимыми*.

Наибольший интерес представляет случай, когда $F^{(1)}$ — неслучайная последовательность («сигнал»), а $F^{(2)}$ — чисто случайная последовательность («шум»).

Приведем результат для случая, когда $F^{(2)} = (f_1^{(2)}, f_2^{(2)}, \dots)$ — гауссовский белый шум, т. е. $f_i^{(2)}$ — независимые одинаково распределенные величины, имеющие нормальное распределение с нулевым средним.

Предложение 2.4. *Любой бесконечный детерминированный ряд $F^{(1)}$ стохастически отделим от гауссовского белого шума $F^{(2)}$, если $L, K \rightarrow \infty$ и $L/K \rightarrow a > 0$.*

2.5. Разделимость и преобразование Фурье

Для описания конечных, но достаточно длинных рядов удобным (особенно для стационарных рядов) является язык разложения Фурье. Пусть $F_N = (f_0, \dots, f_{N-1})$ — ряд длины N .

Определение 2.6. Разложение

$$f_n = c_0 + \sum_{k=1}^{[N/2]} \left(c_k \cos(2\pi n k/N) + s_k \sin(2\pi n k/N) \right), \quad (10)$$

где $0 \leq n < N$ и $s_{N/2} = 0$ для четного N , называется *разложением Фурье* ряда F_N .

Так как нулевой член c_0 равняется среднему значению ряда, то для центрированного ряда он равен нулю.

Разложение (10) — это разложение по ортогональному базису. В другой форме (пусть N — четное) оно может быть записано как

$$f_n = c_0 + \sum_{k=1}^{[(N-1)/2]} \sqrt{c_k^2 + s_k^2} \cos(2\pi n k/N + \varphi_k) + c_{N/2} (-1)^k.$$

Можно считать, что коэффициент c_0 умножается на косинус с частотой 0, а $c_{N/2}$ — на косинус с частотой 0.5.

Введем понятие периодограммы как объекта, характеризующего вклад частот в разложение Фурье ряда.

Определение 2.7. *Периодограммой* ряда $F_N = (f_0, \dots, f_{N-1})$ будем называть функцию Π_f^N , которая задана на множестве частот $\{k/N, k = 0, \dots, [N/2]\}$ формулой

$$\Pi_f^N(k/N) = \frac{N}{2} \begin{cases} 2c_0^2 & \text{для } k = 0, \\ c_k^2 + s_k^2 & \text{для } 0 < k < N/2, \\ 2c_{N/2}^2 & \text{для } k = N/2. \end{cases} \quad (11)$$

Последний случай имеет место только при четном N .

Рассмотрим теперь разложение Фурье (10) двух рядов $F^{(1)}$ и $F^{(2)}$ длины N и обозначим соответствующие коэффициенты $c_k^{(j)}$ и $s_k^{(j)}$, $j = 1, 2$. Обозначим

$$d_k = \begin{cases} c_k^{(1)}c_k^{(2)} + s_k^{(1)}s_k^{(2)} & \text{для } k \neq 0 \text{ и } N/2, \\ 2c_k^{(1)}c_k^{(2)} & \text{для } k = 0 \text{ или } N/2, \end{cases}$$

Предложение 2.5. 1. *Скалярное произведение двух рядов имеет вид*

$$(F^{(1)}, F^{(2)}) \stackrel{\text{def}}{=} \sum_{k=0}^{N-1} f_n^{(1)} f_n^{(2)} = \frac{N}{2} \sum_{k=0}^{[N/2]} d_k.$$

2. *Норма $\|F\| = \sqrt{(F, F)}$ ряда (10) имеет вид*

$$\|F\|^2 = \sum_{k=0}^{[N/2]} \Pi_f^N(k/N). \quad (12)$$

Равенство (12) показывает, что значение (11) периодограммы в точке k/N описывает вклад гармонической компоненты с частотой $\omega = k/N$ в сумму (10). Более того, (12) объясняет коэффициент N^{-1} в определении периодограммы.

Определение 2.8. Значение $\Pi_f^N(k/N)$ называется *мощностью* частоты k/N . Набор частот $\omega_k = k/N$ с положительными мощностями будем называть *носителем* периодограммы. Если носитель периодограммы принадлежит некоторому интервалу $[a, b]$, то этот интервал будем называть диапазоном частот периодограммы.

Так как разделимость рядов описывается в терминах ортогональности их отрезков, а скалярное произведение двух рядов может быть выражено через коэффициенты разложения Фурье, то можно посмотреть на понятие разделимости с точки зрения разложения Фурье и периодограмм.

Сформулируем достаточное условие для разделимости двух рядов. Пусть, как обычно, $F_N = F_N^{(1)} + F_N^{(2)}$ — временной ряд длины N , L — длина окна, $K = N - L + 1$.

Предложение 2.6. Если носители периодограмм отрезков ряда $F_N^{(1)}$ длины L не пересекаются с носителями периодограмм отрезков ряда $F_N^{(2)}$ длины L и аналогичное утверждение справедливо для отрезков рядов длины K , то ряды $F_N^{(1)}$ и $F_N^{(2)}$ слабо разделимы.

Посмотрим с этой точки зрения на (достаточное) условие разделимости двух гармонических рядов с различными периодами T_1 и T_2 : если

$$L = k_1 T_1 = k_2 T_2, \quad K = m_1 T_1 = m_2 T_2$$

с целыми k_1, k_2, m_1, m_2 , то для любого отрезка i -го ряда ($i = 1, 2$) длины L его периодограмма содержит только одну частоту $\omega_i^{(L)} = 1/k_i$ с ненулевой мощностью. Для периодограмм отрезков рядов длины K соответствующая единственная частота с положительной мощностью имеет вид $\omega_i^{(K)} = 1/m_i$. Очевидно, что условия предложения 2.6 в данном случае выполнены.

Приведем пример, показывающий, что периодограммы дают только достаточное условие разделимости.

Пример 2.6. Пусть $N = 399$, $a = 1.005$, $T = 200$ и члены рядов $F^{(1)}$ и $F^{(2)}$ имеют вид

$$f_n^{(1)} = a^{-n}, \quad f_n^{(2)} = a^n \cos(2\pi n/T).$$

Тогда выбор длины окна $L = 200$ (следовательно, $K = 200$) приводит к точной разделимости рядов, однако носители периодограмм отрезков первого и второго ряда пересекаются.

Рассмотрим асимптотическую разделимость рядов с точки зрения периодограмм.

Введем меру спектральной ортогональности рядов $F_N^{(1)}$ и $F_N^{(2)}$.

Определение 2.9. Величину

$$\rho_{12}^{(\Pi)} \stackrel{\text{def}}{=} \frac{\sqrt{\Pi_{f_1}^N(k/N) \Pi_{f_2}^N(k/N)}}{\sqrt{\sum_{k=0}^{[N/2]} \Pi_{f_1}^N(k/N)} \sqrt{\sum_{k=0}^{[N/2]} \Pi_{f_2}^N(k/N)}} \quad (13)$$

будем называть *спектральным коэффициентом корреляции* рядов $F_N^{(1)}$ и $F_N^{(2)}$.

Спектральный коэффициент корреляции (13) имеет смысл меры спектральной ортогональности двух рядов. По модулю он всегда не меньше обычного коэффициента корреляции. Следовательно, если он равен нулю, то ряды ортогональны.

Предложение 2.7. *Если спектральные коэффициенты корреляции отрезков ряда $F^{(1)}$ длины $L = L(N)$ и отрезков ряда $F^{(2)}$ длины L равномерно стремятся к нулю при $N \rightarrow \infty$ и аналогичное утверждение справедливо для отрезков рядов длины $K = K(N)$, то ряды $F^{(1)}$ и $F^{(2)}$ асимптотически разделимы.*

Так как для стационарных бесконечных рядов F периодограммы их начальных отрезков F_N в некотором смысле сходятся при $N \rightarrow \infty$ к спектральной мере ряда F , то справедлив следующий результат.

Предложение 2.8. *Если носители спектральных мер двух стационарных рядов не пересекаются, то эти ряды будут асимптотически разделимы при $L \rightarrow \infty$ и $K \rightarrow \infty$.*

Примером применения предложения 2.8 является асимптотическая разделимость двух гармоник с разными частотами.

2.6. Меры разделимости. Взвешенная корреляция.

Выше мы построили две меры разделимости, набор обычных коэффициентов корреляции и набор спектральных коэффициентов корреляции между отрезками рядов. Равенство нулю всех обычных коэффициентов корреляции между соответственно столбцами и строками траекторных матриц одного и другого рядов является необходимым и достаточным условием слабой разделимости. Равенство нулю всех спектральных коэффициентов корреляции между столбцами и строками траекторных матриц рядов является только достаточным условием слабой разделимости.

Введем еще одно условие разделимости, которое является лишь необходимым, но оно очень простое и легко проверяется.

Пусть $L^* = \min(L, K)$ и $K^* = \max(L, K)$. Введем веса

$$w_i = \begin{cases} i + 1 & \text{для } 0 \leq i \leq L^* - 1, \\ L^* & \text{для } L^* \leq i < K^*, \\ N - i & \text{для } K^* \leq i \leq N - 1. \end{cases} \quad (14)$$

Определим взвешенное скалярное произведение рядов $F_N^{(1)}$ и $F_N^{(2)}$ длины N как

$$(F_N^{(1)}, F_N^{(2)})_w \stackrel{\text{def}}{=} \sum_{i=0}^{N-1} w_i f_i^{(1)} f_i^{(2)}. \quad (15)$$

Определение 2.10. Ряды $F_N^{(1)}$ и $F_N^{(2)}$ называются *w-ортогональными* при длине окна, равной L , если $(F_N^{(1)}, F_N^{(2)})_w = 0$.

Предложение 2.9. Если два ряда разделимы при длине окна, равной L , то они *w-ортогональны*.

Веса (14) появляются в силу того, что разделимость выражается в терминах траекторных матриц рядов, но элементы рядов входят в траекторную матрицу в разном количестве — i -й элемент ряда повторяется в траекторной матрице w_i раз. Веса w_i в скалярном произведении (15) имеют форму трапеции. Если L мало по отношению к N , то почти все веса равны. Если $L \approx N/2$, то влияние крайних элементов ряда гораздо меньше, чем вклад центральных точек ряда.

2.7. Сильная разделимость

До сих пор мы рассматривали условия слабой разделимости рядов. К сожалению, отсутствие сильной разделимости при наличии слабой делает практически невозможным разделение этих двух рядов. Из-за совпадающих собственных чисел тот вариант сингулярного разложения, который имеет место на практике, вообще говоря, не будет разделять ряды.

Приведем простой пример, показывающий эффект отсутствия сильной разделимости.

Пример 2.7. *Слабая и сильная разделимость*

Пусть $N = 3$, $L = K = 2$ и рассмотрим ряды

$$F^{(1)} = (1, -a, a^2), \quad F^{(2)} = (1, a^{-1}, a^{-2}), \quad a \neq 0.$$

Тогда траекторные матрицы $\mathbf{X}^{(1)}$ и $\mathbf{X}^{(2)}$ будут иметь вид

$$\mathbf{X}^{(1)} = \begin{pmatrix} 1 & -a \\ -a & a^2 \end{pmatrix}, \quad \mathbf{X}^{(2)} = \begin{pmatrix} 1 & a^{-1} \\ a^{-1} & a^{-2} \end{pmatrix}.$$

Слабая разделимость проверяется легко. В то же время, матрицы $\mathbf{X}^{(1)}(\mathbf{X}^{(1)})^T$ и $\mathbf{X}^{(2)}(\mathbf{X}^{(2)})^T$ имеют только одно положительное собственное число и сингулярные значения матриц $\mathbf{X}^{(1)}$ и $\mathbf{X}^{(2)}$ записываются в виде

$$\sqrt{\lambda^{(1)}} = 1 + a^2 \quad \text{и} \quad \sqrt{\lambda^{(2)}} = 1 + a^{-2}$$

соответственно. Таким образом, для любого $a \neq 1$ ряды $F^{(1)}$ и $F^{(2)}$ сильно разделимы, но при $a = 1$ сильная разделимость теряется и остается только слабая.

Действительно, если $a \neq 1$, то сингулярное разложение траекторной матрицы \mathbf{X} ряда $F^{(1)} + F^{(2)}$ однозначно определено:

$$\begin{aligned} \mathbf{X} &= \begin{pmatrix} 2 & -a + a^{-1} \\ -a + a^{-1} & a^2 + a^{-2} \end{pmatrix} = \\ &= (1 + a^2)U_1V_1^T + (1 + a^{-2})U_2V_2^T = \mathbf{X}^{(1)} + \mathbf{X}^{(2)}, \end{aligned}$$

где $U_1 = V_1 = (1, -a)^T / \sqrt{(1 + a^2)}$ и $U_2 = V_2 = (1, a^{-1})^T / \sqrt{(1 + a^{-2})}$.

В то же время, когда $a = 1$, существует бесконечно много сингулярных разложений матрицы \mathbf{X} :

$$\mathbf{X} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = 2U_1V_1^T + 2U_2V_2^T,$$

где $\{U_1, U_2\}$ — произвольный ортонормированный базис в \mathbf{R}^2 , а векторы V_1 и V_2 определяются как $V_1 = \mathbf{X}^T U_1 / 2$ и $V_2 = \mathbf{X}^T U_2 / 2$. Из всех таких разложений только одно является подходящим. Оно соответствует выбору $U_1 = (1, 1)^T / \sqrt{2}$ и $U_2 = (1, -1)^T / \sqrt{2}$. Только такое разложение дает требуемое разделение $\mathbf{X} = \mathbf{X}^{(1)} + \mathbf{X}^{(2)}$.

На практике недостаток сильной делимости (при наличии слабой, может быть асимптотической, делимости) проявляется, когда матрица $\mathbf{X}\mathbf{X}^T$ имеет близкие собственные числа, что приводит к неустойчивости вычисления конкретного сингулярного разложения. Например, можно рассмотреть случай суммы двух гармоник с разными частотами, но одинаковыми амплитудами. В этом случае все 4 ненулевых собственных числа будут асимптотически равны при L и K , стремящихся к бесконечности. Другим примером является сумма тренда сложной формы и гармоники. Тогда тренд раскладывается на большое число составляющих и вклад одной из этих составляющих может совпасть с вкладом гармоники.

3. Пространства, порождаемые рядами

Предположим, что $F_N = F_N^{(1)} + F_N^{(2)}$ и известно, что ряды $F_N^{(1)}$ и $F_N^{(2)}$ разделимы. Это означает, что в сингулярном разложении ряда F_N часть слагаемых относится к сингулярному разложению ряда $F_N^{(1)}$, а остальные — к сингулярному разложению ряда $F_N^{(2)}$.

Важный вопрос, который возникает при практическом применении метода, состоит в том, сколько слагаемых относится к первому ряду и как их идентифицировать. А этот вопрос, в свою очередь, связан с понятием размерности и структуры пространств, порождаемых временными рядами (или их компонентами).

3.1. Ряды конечного ранга

Рассмотрим временной ряд $F_N = (f_0, \dots, f_{N-1})$ с $N \geq 3$ и зафиксируем длину окна L ($1 < L < N$).

Как и раньше, в результате процедуры вложения мы получаем последовательность векторов вложения:

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

$\mathfrak{L}^{(L)} = \mathfrak{L}^{(L)}(F_N) \stackrel{\text{def}}{=} \text{span}(X_1, \dots, X_K)$ — траекторное пространство ряда F_N .

Определение 3.1. Пусть $0 \leq d \leq L$. Если $\dim \mathfrak{L}^{(L)} = d$, то мы будем говорить, что *ряд F_N имеет L -ранг d* и записывать это как $\text{rank}_L(F_N) = d$. Для нулевого ряда положим $\dim \mathfrak{L}^{(L)} = 0$.

Очевидно, что равенство $\text{rank}_L(F_N) = d$ может быть справедливым, только если

$$d \leq \min(L, K). \quad (16)$$

При фиксированном d будем называть L , для которого выполнено условие (16), допустимым. Заметим, что если длина окна, равная L , является допустимой, то и длина окна, равная $N - L + 1$, также допустима.

Определение 3.2. Если равенство $\text{rank}_L(F_N) = d < N/2$ имеет место для любого допустимого L , то будем говорить, что *ряд F_N имеет ранг d* ($\text{rank}(F_N) = d$). Если такое d существует, то ряд F_N будем называть *рядом конечного ранга*.

Предложение 3.1. Если ненулевой ряд $F_N^{(1)}$ отделим от ненулевого ряда $F_N^{(2)}$ при длине окна L , то $\text{rank}_L(F_N^{(i)}) < \min(L, K)$, $i = 1, 2$.

Предложение 3.2. 1.

$$\text{rank}_L(F_N) = \text{rank } \mathbf{X} = \text{rank } \mathbf{X}\mathbf{X}^T = \text{rank } \mathbf{X}^T\mathbf{X}.$$

2. Ортонормированная система собственных векторов $\{U_i\}_{i=1}^d$, соответствующих положительным собственным числам $\lambda_1 \geq \dots \geq \lambda_d$ матрицы $\mathbf{X}\mathbf{X}^T$, составляет ортонормированный базис пространства $\mathfrak{L}^{(L)}$.

3. Ранг $\text{rank}_L(F_N)$ равняется порядку (числу ненулевых собственных чисел) сингулярного разложения траекторной матрицы \mathbf{X} .

Следующее утверждение описывает множество рядов конечного ранга.

Предложение 3.3. Любой ряд, являющийся линейной комбинацией произведений полиномов, экспонент и косинусов, является рядом конечного ранга.

Рассмотрим бесконечный в обе стороны временной ряд $F = (\dots, f_{-1}, f_0, f_1, \dots)$. Для него можно рассмотреть аналогичное определение ранга как ранга траекторной матрицы с бесконечным числом столбцов.

Определение 3.3. Будем говорить, что бесконечный ряд F управляется линейной рекуррентной формулой размерности d , если найдутся такие коэффициенты a_1, \dots, a_d , что для любого i

$$f_{i+d} = \sum_{k=1}^d a_k f_{i+d-k}, \quad a_d \neq 0. \quad (17)$$

Предложение 3.4. Любой бесконечный ряд является рядом конечного ранга d тогда и только тогда, когда он управляется линейной рекуррентной формулой размерности d .

3.2. Примеры рядов конечного ранга и характеристики их сингулярного разложения

Приведем примеры рядов конечного ранга, их ранг, а также вид соответствующих сингулярных векторов.

3.2.1. Экспоненциально-модулированный гармонический ряд

Рассмотрим ряд F_N с

$$f_n = Ae^{\alpha n} \cos(2\pi\omega n + \phi), \quad (18)$$

$\omega \in [0, 1/2]$, $\phi \in [0, 2\pi)$ и обозначим $T = 1/\omega$. В зависимости от параметров такой ряд имеет ранг либо 1 либо 2.

1. *Экспоненциально-модулированный гармонический ряд с частотой $\omega \in (0, 1/2)$*

Если $\omega \in (0, 1/2)$, то для любых $L \geq 2$ и $N \geq L+1$ сингулярное разложение тракторной матрицы имеет два члена (ряд ранга 2). Оба собственных (и факторных) вектора имеют такую же форму (18) с той же самой частотой ω и показателем экспоненты α .

Если $\alpha \leq 0$, то при $L, K \rightarrow \infty$ оба ненулевых собственных числа будут асимптотически равны.

2. *Экспоненциально-модулированный пилообразный ряд с частотой $\omega = 1/2$*

Если $\omega = 1/2$ и $\sin \phi \neq 0$, то f_n пропорционально $(-e^\alpha)^n$. В этом случае при любом L сингулярное разложение будет иметь один член. Соответствующие собственный и факторный векторы будут иметь такую же пилообразную форму, как и исходный ряд.

3. *Экспоненциальный ряд ($\omega = 0$)*

Если $\omega = 0$ и $\cos \phi \neq 0$, то f_n пропорционально $e^{\alpha n}$ и мы имеем дело с экспоненциальным рядом. Для любых N и L ранг ряда равен 1. Соответствующие собственный и факторный векторы будут иметь такую же экспоненциальную форму, как и исходный ряд с тем же показателем α .

4. *Гармонический ряд ($\alpha = 0$)*

Это частный случай экспоненциально-модулированного гармонического ряда. Поэтому ранг ряда равен 2 при $\omega < 1/2$ и равен 1 при $\omega = 1/2$. Сингулярные векторы являются гармониками с той же частотой ω . Если $\omega < 1/2$, а K и L делятся нацело на $T = 1/\omega$, то оба собственных числа совпадают. Если $L\omega$ ($K\omega$) — целое, то фазы пары собственных (факторных) векторов в силу ортогональности будут отличаться на $\pi/2$. Поэтому их элементы в этом случае имеют вид $C \cos(2\pi\omega n + \tilde{\phi})$ и $C \sin(2\pi\omega n + \tilde{\phi})$.

3.2.2. Полином

Рассмотрим полиномиальный ряд F_N с

$$f_n = \sum_{k=0}^m a_k n^k, \quad a_m \neq 0.$$

Если F_N является полиномом степени m , то порядок соответствующего сингулярного разложения (и ранг ряда) при $L \geq m + 1$ и $N \geq L + m$ равняется $m + 1$ и все сингулярные векторы также являются полиномами степени, не превосходящей m . В частности, линейный ряд имеет ранг, равный 2, и два линейных собственных (и факторных) векторы.

4. Выбор параметров

В этом разделе мы обсудим роль параметров базового метода SSA и принципы их выбора. Будем предполагать, что рассматриваемый ряд состоит из медленно меняющегося тренда, колебательных компонент и шума.

Конечно, выбор параметров зависит от данных и от того, какую задачу анализа временного ряда нужно решить. Обсудим задачу выбора параметров для некоторых основных задач.

В базовом методе SSA есть два параметра. Первый — это целое L , длина окна, а второй параметр является структурным — это способ группировки собственных троек (компонент сингулярного разложения траекторной матрицы ряда).

4.1. Группировка

Пусть длина окна L фиксирована и мы уже имеем сингулярное разложение траекторной матрицы исходного ряда. Тогда следующим шагом является группировка членов сингулярного разложения, соответствующая (приближенно) разделимым при данном L составляющим ряда (предполагаем, что приближенная разделимость имеет место).

4.1.1. Общие правила группировки

1. *Если мы восстановили компоненту ряда только с помощью одной собственной тройки и оба сингулярных вектора имеют похожую форму, то восстановленная компонента будет иметь примерно такую же форму.* Это правило означает, что, имея дело с единственной собственной тройкой, часто можно предсказать поведение соответствующей компоненты временного ряда. Например, если оба сингулярных вектора собственной тройки похожи на линейные ряды, то соответствующая составляющая ряда также будет близкой к линейной. Если сингулярные векторы имеют экспоненциальную форму, то и компонента ряда будет такой же. Монотонные сингулярные векторы соответствуют монотонной компоненте ряда. Синусоидальные векторы порождают гармоническую составляющую ряда.
2. *Если $L \ll K$, то факторный вектор в собственной тройке более похож на восстановленную компоненту ряда, чем соб-*

ственный вектор. Следовательно, можно приближенно предсказать результат восстановления, основываясь на виде факторного вектора.

3. *Если восстановить ряд с помощью нескольких собственных троек, с одним и тем же диапазоном частот $[a, b]$ соответствующих сингулярных векторов, то носитель периодограммы восстановленного ряда будет также в основном принадлежать отрезку $[a, b]$.*
4. *Чем больше сингулярное значение в собственной тройке, тем больше вклад соответствующей восстановленной компоненты ряда.* Собственное число можно рассматривать в качестве характеристики этого вклада.

Перейдем теперь к решению конкретных задач.

4.1.2. Группировка для выделения тренда и сглаживания

1. Тренды

Согласно нашему определению, трендом является медленно меняющаяся компонента ряда, не содержащая колебательных компонент. Предположим, что ряд состоит только из тренда. Тогда несколько первых сингулярных векторов будут также медленно меняться. Примером тому служат экспоненциальный или полиномиальный тренды.

Рассмотрим некоторый ряд F , для которого мы будем предполагать, что его тренд (приближенно) сильно отделим от других компонент. Это означает, что среди собственных троек ряда F можно выделить те, которые приближенно соответствуют компонентам сингулярного разложения тренда.

Таким образом, *чтобы извлечь тренд из ряда, нужно собрать все собственные тройки с медленно меняющимися сингулярными векторами.*

Порядковые номера таких собственных троек зависят не только от самого тренда $F^{(1)}$, но также и от остатка $F^{(2)} = F - F^{(1)}$. Рассмотрим два противоположных случая. В первом случае пусть ряд F имеет сильную ярко-выраженную тенденцию $F^{(1)}$ с относительно небольшой колебательно-шумовой компонентой $F^{(2)}$. Тогда большинство собственных троек, соответствующих тренду, будет иметь первые позиции в сингулярном разложении траекторной матрицы

всего ряда. Конечно, некоторые из собственных троек будут иметь меньшие собственные числа, особенно если тренд имеет сложную структуру.

В противоположной ситуации рассмотрим ряд, являющийся суммой больших колебаний на уровне слабой общей тенденции. В этом случае первые собственные тройки будут соответствовать колебаниям.

2. Сглаживание

Задача сглаживания подобна задаче извлечения тренда. Ее особенностью является то, что мы можем сглаживать любой ряд, даже тот, который не имеет очевидной трендовой составляющей. Для сглаживания ряда достаточно собрать все собственные тройки с медленно меняющимися сингулярными векторами. В случае стационарного ряда с убывающей спектральной плотностью (например, процесса авторегрессии первого порядка с положительным коэффициентом) собственные тройки, соответствующие низкочастотным составляющим ряда, будут находиться главным образом среди первых.

4.1.3. Группировка для выделения колебаний

1. Гармонический ряд

Начнем с рассмотрения чистой гармоника с частотой ω и некоторыми фазой и амплитудой. Так как мы предполагаем, что компонента $F^{(1)}$ исходного ряда приближенно сильно отделима от $F^{(2)} = F - F^{(1)}$ при длине окна L , мы можем рассчитывать, что две (или одна, если $\omega = 1/2$) собственные тройки сингулярного разложения траекторной матрицы ряда F соответствуют $F^{(1)}$. Проблема заключается только в том, как идентифицировать эти собственные тройки среди всех остальных.

Пусть $\omega < 1/2$. Как было сформулировано выше, гармоника, заданная в (18) с $\alpha = 0$, имеет размерность 2 и порождает сингулярные векторы с той же частотой ω .

Рассмотрим идеальную ситуацию, когда $L\omega$ и $K\omega$ — целые, $K = N - L + 1$.

Возьмем для определенности левые сингулярные векторы (т. е. собственные векторы). В идеальной ситуации собственные векторы — это косинус и синус с периодом T и с одной и той же амплитудой и фазой. То же самое верно и для факторных векторов. Двумерное изображение двух таких собственных (и факторных) векторов может упростить поиск таких пар, так как косинус и синус с одинаковыми частотой, фазой и амплитудой при двумерном изображении (например, косинус — по оси X , синус — по оси Y) создают точки, лежащие на окружности.

Если $T = 1/\omega$ — целое (T является периодом гармоник), то получаем вершины правильного T -угольника. Для рациональных частот $\omega = q/p < 1/2$ с взаимно простыми целыми p и q , точки являются вершинами правильного p -угольника. На рис. 1 приведены двумерные изображения четырех пар синус/косинус с нулевой фазой, одинаковыми амплитудами и частотами $1/12$, $10/53$, $2/5$ и $5/12$.

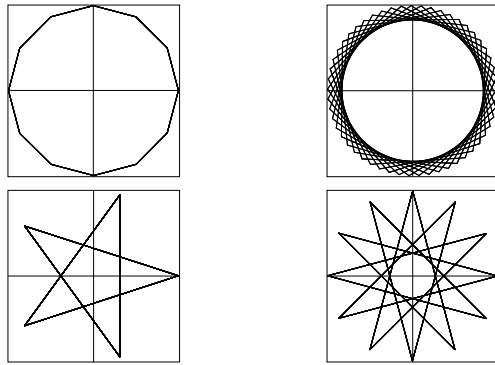


Рис. 1. Двумерные диаграммы синуса/косинуса.

Так как собственные числа, порожденные гармоникой, асимптотически при $L, K \rightarrow \infty$ равны, то на практике они оказываются близки друг к другу, по крайней мере, когда L и K в несколько раз больше $1/\omega$. Это помогает находить пару собственных троек, порожденную гармоникой, рассматривая только последовательные собственные тройки. Заметим, что если период гармоник сравним с длиной ряда, то соответствующие собственные числа могут довольно сильно отличаться.

Если ряд F содержит несколько гармоник, то каждая из них может быть идентифицирована на основе двумерных графиков последовательных сингулярных векторов. В более сложных случаях, когда, например, гармоника выделилась плохо, можно использовать периодограммный анализ сингулярных векторов.

Частоты гармоник могут быть оценены как с помощью периодограмм, так и на основе попарных двумерных графиков сингулярных векторов.

Рассмотрим две собственные тройки, которые соответствуют гармонике с частотой ω_0 . Тогда двумерная диаграмма собственных векторов может быть представлена как двумерная кривая с евклидовыми координатами

$$\begin{aligned} x(n) &= r(n) \cos(2\pi\omega(n)n + \phi(n)), \\ y(n) &= r(n) \sin(2\pi\omega(n)n + \phi(n)), \end{aligned}$$

где функции r , ω и ϕ близки к константам и $\omega(n) \approx \omega_0$. Полярные координаты узлов кривой имеют вид $(r(n), \delta(n))$, а $\delta(n)$ выражается как $\delta(n) = 2\pi\omega(n)n + \phi(n)$.

Так как $\Delta_n \stackrel{\text{def}}{=} \delta(n+1) - \delta(n) \approx 2\pi\omega_0$, можно оценить ω_0 усреднением приращений полярных углов Δ_n ($n = 0, \dots, L-1$). Точно такая же процедура может быть применена к паре факторных векторов.

Если период гармоник равен 2, т. е. $\omega = 1/2$, то ситуация проще, так как в этом случае сингулярное разложение состоит только из одной собственной тройки, а соответствующие собственный и факторный векторы имеют пилообразную форму.

2. Периодические компоненты общего вида

Рассмотрим более подробно случай выделения периодической компоненты $F^{(1)}$ из ряда F . Если целое T является периодом этой компоненты, то разложение Фурье будет иметь вид

$$f_n^{(1)} = \sum_{k=1}^{[T/2]} \sqrt{a_k^2 + b_k^2} \cos(2\pi kn/T + \phi_k). \quad (19)$$

Следовательно, размерность такого ряда не превосходит $T-1$, так как каждая из гармонических компонент порождает либо две (при $k \neq T/2$), либо одну ($k = T/2$) собственную тройку. При больших L

все компоненты в (19) будут приближенно сильно разделимы, если все их мощности $a_k^2 + b_k^2$ различны.

Таким образом, при условиях:

- (а) ряд $F^{(1)}$ — (приближенно) строго отделим от $F^{(2)}$ в сумме $F = F^{(1)} + F^{(2)}$ при длине окна L ,
- (б) все ненулевые мощности $a_k^2 + b_k^2$ в разложении (19) различны,
- (в) L — достаточно большое,

можно выделить все собственные тройки, соответствующие периодическому ряду $F^{(1)}$, а также его гармоническим составляющим, в сингулярном разложении траекторной матрицы всего ряда F .

Для того чтобы выполнить это разделение, достаточно идентифицировать те собственные тройки, которые соответствуют всем гармоникам с частотами k/T , и сгруппировать их.

Например, если известно, что в ряде F присутствует сезонная компонента и данные помесечные, то необходимо искать периодики с частотами $1/12$ (годовая), $1/6$ (полугодовая), $1/4$, $1/3$ (квартальная), $1/2.4$, и $1/2$. Естественно, некоторые из них могут отсутствовать.

Если некоторые из ненулевых мощностей в (19) совпадают, то проблема идентификации собственных троек имеет свою специфику. Предположим, например, что $a_1^2 + b_1^2 = a_2^2 + b_2^2$. Если L — большое, то четыре сингулярных числа в сингулярном разложении траекторной матрицы ряда F будут (приближенно) одинаковыми и соответствующие сингулярные векторы будут равны линейным комбинациям гармоник с двумя частотами: $\omega_1 = 1/T$ и $\omega_2 = 2/T$.

Следовательно, компоненты с этими частотами не будут сильно разделимы. В этом случае периодограммный анализ сингулярных векторов может быть очень полезен. Если периодограмма имеет острые пики в окрестности частот ω_1 и ω_2 , то соответствующие собственные тройки должны быть сопоставлены ряду $F^{(1)}$.

3. Модулированные периодики

Случай амплитудно-модулированных периодик более сложен, так как форма модуляции амплитуды неизвестна. Однако пример экспоненциально-модулированной гармоники (18) с $\alpha \neq 0$ показывает, что иногда идентификация такого рода компонент может быть выполнена.

Начнем с рассмотрения такого ряда. Заметим, что несмотря на простоту, экспоненциально-модулированная гармоника может быть рассмотрена, например, как аддитивная компонента некоторого эконометрического ряда, описывающего экспоненциальный рост с экспоненциально-модулированной сезонной компонентой.

В целом, ситуация аналогична случаю чистой гармоники. Если частота ω меньше $1/2$ и длина окна L достаточно большая, то в сингулярном разложении мы получаем две собственные тройки с примерно равными сингулярными значениями и сингулярными векторами, имеющими такую же, как у исходного ряда, форму. Поэтому двумерная диаграмма этой пары собственных (факторных) векторов будет иметь спиралеобразную форму, легко различимую на графике.

Если период гармоники равен 2, то компонента имеет вид $(-a)^n$, где $a = e^\alpha$. Сингулярные векторы единственной собственной тройки, порожденные такой компонентой, имеют точно такую же модулированную пилообразную форму.

Если ряд, модулирующий гармонику, не является экспоненциальным, то выделение такой компоненты гораздо сложнее. Рассмотрим случай, когда это все же можно сделать.

Если амплитуда модулированной гармоники $F^{(1)}$ меняется медленно, то ее диапазон частот сосредоточен вокруг главной частоты, которая может быть ясно увидена на периодограмме модулированной гармонике. Если длина окна L и величина $K = N - L + 1$ большие, то все сингулярные векторы этого ряда будут иметь одни и те же свойства. Следовательно, если ряд $F^{(1)}$ (приближенно) сильно отделим от ряда $F^{(2)}$ в сумме $F = F^{(1)} + F^{(2)}$, то можно ожидать, что диапазон частот ряда $F^{(2)}$ имеет маленькое с точки зрения мощности пересечение с диапазоном частот модулированной гармонике $F^{(1)}$. Таким образом, анализируя периодограммы сингулярных векторов собственных троек, можно идентифицировать большинство из тех, которые (приближенно) описывают временной ряд $F^{(1)}$.

В случае произвольной периодики (не обязательно гармонике) ситуация аналогична. В этом случае каждый член в сумме (19) умножается на некоторую функцию, модулирующую амплитуду, и каждая частота k/T порождает группу соседних частот. Поэтому при условии (приближенной) сильной делимости рядов $F^{(1)}$ и $F^{(2)}$, нужно искать сингулярные векторы сингулярного разложения

траекторной матрицы ряда F , такие что частоты их периодограмм сосредоточены вокруг частот k/T .

4.1.4. Группировка для нахождения полного разложения ряда

Задача нахождения полного разложения ряда эквивалентна идентификации собственных троек сингулярного разложения траекторной матрицы ряда, соответствующих тренду, различным колебательным компонентам и шуму. Принципы группировки для идентификации тренда и колебательных компонент были описаны выше.

Что касается шума, мы всегда должны иметь в виду присущую этому понятию неопределенность при недостатке строгой математической модели. С практической точки зрения, естественный путь выделения шума — это отнесение к нему тех собственных троек, которые не содержат ни тренда, ни колебаний. Однако, поступая таким образом, нужно обратить внимание на следующие моменты.

1. Если диапазон частот шума включает в себя частоту гармонической компоненты сигнала, то гармоническая компонента, восстановленная по соответствующим собственным тройкам сингулярного разложения, будет содержать также часть шума, соответствующую этой частоте. Аналогично, выделенный тренд захватывает низкочастотную часть шума. Если частотные диапазоны сигнала и шума не пересекаются, то такой эффект не возникает.
2. Если амплитуда гармонической компоненты сигнала мала, а шум большой, то собственные числа, соответствующие этой гармонике и шуму, могут быть близки друг к другу. Поэтому может оказаться невозможным отделить гармонику от шума на основе анализа собственных троек для всего ряда. Говоря более строго, гармоническая и шумовая компоненты не будут сильно разделены. Этот эффект исчезает асимптотически при $N \rightarrow \infty$.
3. Сингулярное разложение траекторной матрицы чистого шума при больших N , L и K скорее всего будет содержать, по крайней мере, несколько первых собственных троек, похожих на порожденные гармониками. Компоненты исходного ряда,

восстановленные на их основе, также выглядят похожими на гармоники (обычно сильно модулированные). Поэтому нужны тщательный контроль и осторожность при интерпретации восстановленных компонент.

Приведенное выше обсуждение имеет отношение и к задаче снижения уровня шума, которая является частным случаем рассмотренной проблемы.

4.1.5. Дополнительные характеристики, используемые при группировке

Есть ряд дополнительных характеристик собственных троек сингулярного разложения траекторной матрицы исходного ряда, которые могут помочь правильно их сгруппировать для выделения разных компонент ряда. Обсудим две такие характеристики.

1. Сингулярные числа

Как уже упоминалось выше, если N , L и K достаточно большие, то каждая гармоническая, отличная от пилообразной, компонента порождает две собственных тройки с близкими сингулярными числами. Более того, подобная ситуация происходит, если мы имеем сумму нескольких гармоник с различными частотами и близкими амплитудами. Хотя соответствующие сингулярные векторы не обязательно будут соответствовать чистым гармоникам из-за возможного смешивания, но они все же могут формировать пары с близкими сингулярными числами и похожими формами.

Таким образом, ярко выраженные плато на графике собственных чисел могут служить подсказкой к выбору парных собственных троек.

Также может быть полезна форма графика собственных чисел. Как правило, чисто шумовой ряд порождает медленно убывающую последовательность собственных чисел. Если шум добавлен к сигналу, описываемому несколькими собственными тройками с большими сингулярными числами, то скачок в собственных числах может отделять собственные тройки, относящиеся к сигналу, от остальных.

Заметим, что, вообще говоря, нет формальной процедуры, позволяющей находить такой скачок. Более того, для сложного сигнала и большого шума собственные тройки могут перемешиваться по отношению к порядку, задаваемому убыванием сингулярных чисел.

В любом случае, сингулярные числа дают важную, но только вспомогательную информацию для группировки. Структура сингулярных векторов является более существенной.

2. w -Корреляции

Как уже отмечалось раньше, необходимым условием для (приближенной) делимости двух рядов является (приближенно) нулевые w -корреляции восстановленных компонент. С другой стороны, собственные тройки, входящие в одну и ту же группу, могут соответствовать сильно коррелированным составляющим ряда.

Таким образом, естественной вспомогательной информацией является матрица абсолютных значений w -корреляций, соответствующих полному разложению (в таком разложении каждая группа соответствует одной матричной компоненте сингулярного разложения траекторной матрицы ряда). Матрица модулей w -корреляций для ряда F с

$$f_n = e^{n/400} + \sin(2\pi n/17) + 0.5 \sin(2\pi n/10) + \varepsilon_n, \quad (20)$$

$n = 0, \dots, 339$, стандартным гауссовским белым шумом ε_n и для $L = 85$ представлена на рис. 2 (w -корреляции для первых 30 восстановленных компонент изображены в 20-цветовой шкале от белого до черного, соответствующей абсолютным значениям от 0 до 1).

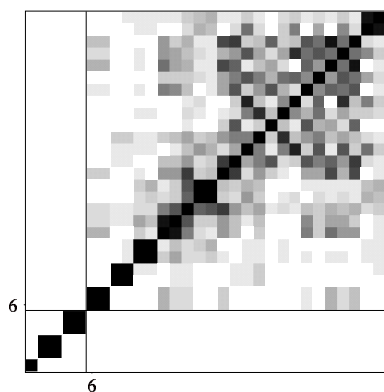


Рис. 2. Ряд (20): матрица w -корреляций.

Вид матрицы является индикатором для правильной группировки: первая собственная тройка описывает экспоненциальный

тренд, две пары последующих соответствуют двум гармоникам, большой пестрый квадрат соответствует шумовым компонентам. Результат разложения соответствует теоретическим результатам о асимптотической делимости компонент ряда.

4.2. Выбор длины окна

Длина окна является основным параметром базового алгоритма SSA: неправильный выбор этого параметра может привести к тому, что никакая группировка не приведет к хорошему разложению ряда на составляющие.

Выбор правильной длины окна зависит от решаемой задачи и предварительной информации, известной о ряде. В общем случае нет универсальных правил и безусловных рекомендаций для выбора длины окна. Основная проблема состоит во влиянии длины окна как на слабую, так и на сильную делимость, т. е. на ортогональность соответствующих временных рядов и на близость собственных чисел.

Однако есть несколько основных принципов для выбора длины окна L , которые имеют как теоретическое, так и практическое обоснование. Обсудим их.

4.2.1. Общие принципы

1. Длина окна и длина ряда

- Сингулярные разложения траекторных матриц одного и того же ряда длины N , соответствующие выбору длины окна L и $N - L + 1$ эквивалентны (с точностью до симметрии: левые сингулярные векторы \leftrightarrow правые сингулярные векторы). Следовательно, для анализа структуры временного ряда базовым методом SSA не имеет смысла брать длину окна, большую чем половина длины ряда.
- С учетом предыдущего замечания, чем больше длина окна, тем более детальным получается разложение исходного ряда. Таким образом, наиболее детальное разложение достигается при выборе длины окна, приблизительно равной половине длины ряда ($L \sim N/2$). Исключением являются ряды конечного ранга, для которых при любом L , большем чем ранг ряда d , и $N > 2d - 1$ число ненулевых компонент в сингулярном разложении равно d и не зависит от длины окна.

2. Длина окна и слабая разделимость

- Так как результаты относительно слабой разделимости аддитивных составляющих временных рядов в большинстве своем асимптотические (при $L, K \rightarrow \infty$), как правило, чтобы достичь лучшей разделимости, нужно выбирать большую длину окна. Другими словами, *маленькая длина окна может привести к смешиванию интерпретируемых компонент ряда*.
- Если длина окна L достаточно большая (скажем, несколько дюжин), то *результаты (слабой) разделимости устойчивы по отношению к небольшим изменениям L* .
- С другой стороны, *для решения особых задач могут быть конкретные рекомендации по выбору длины окна*, соответствующие относительно небольшим N и L (см., например, раздел, посвященный выбору длины окна для выделения периодик).

3. Длина окна и близость собственных чисел

Отрицательное влияние близких собственных чисел, относящихся к разным компонентам исходного ряда (т. е. отсутствие сильной разделимости, когда (приближенная) слабая разделимость имеет место), трудно формализуется в терминах длины окна. Часто отсутствие сильной разделимости (смешивание компонент ряда) не может быть преодолено только за счет выбора длины окна.

Отметим два момента, относящиеся к проблеме близких собственных чисел.

- *Для рядов со слишком сложной структурой слишком большая длина окна L может привести к нежелательному разложению интересующих нас компонент ряда*, в частности, к смешиванию их с другими компонентами ряда. Это не очень приятное свойство, особенно потому, что сильное уменьшение L может привести к значительному ухудшению качества (слабой) разделимости.
- В свою очередь, иногда *даже маленькое изменение значения L может уменьшить смешивание и привести к лучшему разделению компонент ряда*, т. е. обеспечить переход от слабой к сильной разделимости. В любом случае, всегда стоит попробовать несколько вариантов выбора длины окна.

4.2.2. Выбор длины окна для выделения тренда и сглаживания

1. Задача выделения тренда

В задаче выделения тренда противоречие между слабой и сильной разделимостью проявляется наиболее часто. Так как тренд — это относительно гладкая кривая, его отделимость от шума и колебаний требует достаточно больших значений длины окна L .

С другой стороны, если тренд имеет сложную структуру, то при больших L он может описываться большим количеством собственных троек с относительно небольшими сингулярными числами. Некоторые из этих сингулярных чисел могут иметь значения, близкие к сингулярным числам, соответствующим колебаниям и/или шуму.

Задача выделения тренда сложна и существует большое разнообразие возможных ситуаций. Опишем коротко, на качественном уровне, два противоположных случая: когда тренд может быть выделен довольно легко и когда выбор длины окна для выделения тренда труден или даже невозможен.

2. Тренды: устойчивая разделимость

Пусть $F = F^{(1)} + F^{(2)}$, где $F^{(1)}$ — тренд и $F^{(2)}$ — остаток. Будем предполагать следующее.

(а) Ряд $F^{(1)}$ — «простой». Понятие «простоты» мы понимаем так:

- С теоретической точки зрения ряд $F^{(1)}$ хорошо аппроксимируется рядом конечного и небольшого ранга d (например, он похож на экспоненциальный, $d = 1$, линейный $d = 2$, квадратичный, $d = 3$, и т. д.).
- Нас интересует выделение общей тенденции ряда.
- В терминах частотного анализа, периодограмма ряда $F^{(1)}$ сосредоточена в области низких частот.
- В терминах разложения ряда, первые несколько собственных троек сингулярного разложения траекторной матрицы ряда даже для больших L дают достаточно хорошую его аппроксимацию.

(б) Предполагается, что ряд $F^{(1)}$ намного больше по значениям, чем ряд $F^{(2)}$ (например, если $\|F^{(1)}\| \gg \|F^{(2)}\|$).

Пусть эти предположения верны и длина окна L обеспечивает разделение (слабое, приближенное) временных рядов $F^{(1)}$ и $F^{(2)}$. Тогда можно ожидать, что в сингулярном разложении траекторной матрицы ряда F первые несколько собственных троек будут соответствовать $F^{(1)}$, т. е. они будут иметь сингулярные числа, большие чем собственные тройки, соответствующие ряду $F^{(2)}$. Другими словами, будет достигнута сильная разделимость. Более того, длина окна L , достаточная для разделимости, в этом случае не обязана быть очень большой в силу «простоты» тренда.

3. Тренды: сложный случай

Более сложная ситуация возникает, если мы хотим выделить тренд более точно (поэтому он более «сложный»), и остаток $F^{(2)}$ также имеет сложную структуру (например, это сильная шумовая компонента) с большой нормой $\|F^{(2)}\|$. Тогда большое L может вызвать не только смещение порядковых номеров собственных троек, соответствующих $F^{(1)}$ и $F^{(2)}$, но и привести к отсутствию сильной разделимости. Конечно, существует большое количество промежуточных ситуаций.

4. Сглаживание

Вообще говоря, рекомендации относительно выбора длины окна для сглаживания ряда подобны тем, которые были даны выше для решения задачи выделения тренда, так как эти две задачи тесно связаны друг с другом. Прокомментируем выбор длины окна для сглаживания на языке частотного анализа.

Трактуя сглаживание как удаление высокочастотной части исходного ряда, мы должны выбирать длину окна L достаточно большой, чтобы обеспечить разделение низкочастотной и высокочастотной частей спектра. Если мощности низкочастотных составляющих существенно больше, чем мощности высокочастотных составляющих, то задача сглаживания довольно проста и сводится к группировке нескольких первых троек. Как правило, чем больше L , тем более узкий диапазон низких частот можно извлечь. Более сложная задача возникает при сглаживании ряда с близкими мощностям высоких и низких частот.

4.2.3. Длина окна для выделения периодик

Выбор длины окна для выделения периодической компоненты $F^{(1)}$ из суммы $F = F^{(1)} + F^{(2)}$ имеет некоторые особенности, связанные с соотношением между длиной окна и величиной периода.

Вообще говоря, эти особенности не отличаются в случаях выделения чистых гармоник, сложных периодик и даже модулированных периодик. Поэтому рассмотрим все эти случаи вместе.

1. При решении задачи выделения периодической компоненты с периодом T естественно измерять длину ряда N и длину окна L в числе периодов. В частности, если ряд $F^{(1)}$ асимптотически отделим от ряда $F^{(2)}$, то, чтобы достичь разделения, длина ряда должна быть, как правило, порядка нескольких периодов.
2. Для относительно коротких рядов, предпочтительно учитывать условия для точной (не асимптотической) разделимости. Если известно, что временной ряд имеет периодическую компоненту с известным периодом T (например, сезонную компоненту), то предпочтительно взять длину окна L пропорциональной величине этого периода. Заметим, что, с теоретической точки зрения, $N-1$ (или, что то же самое, $K = N-L+1$) также должно быть пропорционально T . Но практически более важна пропорциональность меньшего из L и $K = N-L+1$.
3. В случае длинных рядов требование, чтобы L/T и $(N-1)/T$ были целыми, не столь важно. В этом случае можно рекомендовать брать длину окна как можно больше (близкой к $N/2$, если позволяют возможности компьютера). Тем не менее, даже для случая длинных рядов лучше выбрать L , делящееся нацело на T .
4. Если ряд $F^{(2)}$ содержит периодическую компоненту с периодом $T_1 \approx T$, то, чтобы извлечь $F^{(1)}$, нужна большая длина окна, чем для случая, когда такой компоненты нет.
5. Так как две гармонических компоненты с одинаковыми амплитудами порождают одинаковые сингулярные числа (асимптотически при L и K , стремящихся к бесконечности), то большая длина окна может привести к потере сильной разделимости и, как следствие, смещению компонент ряда. Если, к тому же, частоты двух гармоник (почти) равны, то проявляется противоречие между необходимостью как слабой, так и сильной разделимости: близкие частоты требуют для разде-

лимости большой длины окна, но большая длина окна приводит к близким сингулярным значениям.

4.2.4. Длина окна и полное разложение ряда

При одновременном выделении разных компонент ряда все аспекты выбора длины окна должны приниматься во внимание. Имеет смысл начинать всегда с выбора длины окна, близкой к половине ряда (по крайней мере, достаточно большой) и кратной периоду периодической компоненты, если такая присутствует в ряде и ее период известен.

Для коротких рядов и/или рядов со сложной структурой эти простые рекомендации могут быть невыполнимыми или недостаточными. Тогда выбор длины окна становится более сложной задачей.

В более сложных случаях лучшие результаты можно получить, используя последовательный метод SSA, в котором после выделения одной из компонент с некоторым L анализируется остаток (снова методом SSA) с возможно уже другой длиной окна L .

4.2.5. Дополнительная информация для выбора длины окна

Если мы не имеем какой-либо дополнительной информации о ряде из той области, данные из которой он содержит, то единственным источником информации для выбора длины окна является временной ряд сам по себе. Часто форма графика ряда может стать эффективным индикатором его свойств: например, показывать наличие тренда или явной периодичности. В то же самое время, невозможно описать все возможные способы предварительного анализа ряда, которые могут помочь при выборе L , особенно из-за того, что соответствующий алгоритм должен быть быстрым (в сравнении с сингулярным разложением большой матрицы) и соответствовать специфике решаемой задачи.

Мы рассмотрим только рекомендации для выбора длины окна, основанные на периодограммном анализе исходного ряда или его части:

1. Если разрешение периодограммы ряда F достаточно хорошее (т. е. если ряд достаточно длинный), то периодограмма может помочь в определении периодов гармонических компонент ряда. Более того, присутствие различных диапазонов частот с

большими мощностями может быть индикатором возможных естественных компонент исходного ряда в соответствии с этими диапазонами.

2. Одним из достаточных условий приближенной слабой разделимости двух рядов является малость спектральных корреляций для всех отрезков первого ряда длины L (и также $K = N - L + 1$) и отрезков второго ряда такой же длины. Предположим, что:

- периодограммы всех отрезков ряда F длины L и K имеют одинаковую структуру,
- эта структура характеризуется наличием различных и далеко отстоящих друг от друга диапазонов частот с большими мощностями.

В этом случае выбор длины окна, равной L , скорее всего приведет к разделению ряда F на компоненты, соответствующие этим диапазонам частот.

Критерием правильности выбора длины окна является этап группировки. Успешная группировка собственных троек означает, что длина окна была выбрана правильно.

5. Модификации базового метода SSA

С помощью модификаций SSA часто удается улучшить базовый метод для некоторых специфических классов рядов и для рядов со сложной структурой. В этом разделе мы рассмотрим несколько таких классов рядов и опишем соответствующие модификации. А именно мы будем иметь дело со следующими рядами и задачами:

1. Временной ряд колеблется вокруг линейного тренда, и мы хотим извлечь этот линейный тренд.
2. Временной ряд имеет стационарное поведение, и мы хотим извлечь несколько гармонических компонент из этого ряда.
3. Временной ряд имеет сложную структуру (например, тренд сложной формы или гармонические компоненты с близкими амплитудами), и, следовательно, для любой длины окна происходит смешивание компонент ряда.

Техника, используемая для решения первых двух задач, в некотором смысле аналогична; имея дополнительную информацию о временном ряде, мы получаем некоторое разложение траекторной матрицы ряда, отличное от сингулярного и настроенное на структуру ряда.

Отсутствие сильной разделимости (когда слабая разделимость имеет место) является одной из основных проблем базового метода SSA. Одним из возможных путей для преодоления проблемы совпадающих собственных чисел является увеличение собственного числа, соответствующего некоторой компоненте ряда, с помощью добавления к ряду подобной компоненты. Другим способом является последовательное применение SSA. В этом случае мы можем извлечь сначала одну из компонент ряда, а затем снова применить SSA к остатку после выделения этой компоненты. Рассмотрим эти подходы более подробно.

5.1. Центрирование в SSA

Рассмотрим следующее расширение базового варианта SSA. Предположим, что длина окна L уже выбрана. Для $K = N - L + 1$ возьмем некоторую матрицу \mathbf{A} размерности $L \times K$ и перейдем от

траекторной матрицы \mathbf{X} ряда F к матрице $\mathbf{X}^* = \mathbf{X} - \mathbf{A}$. Пусть $\mathbf{S}^* = \mathbf{X}^*(\mathbf{X}^*)^T$. Обозначим λ_i и U_i ($i = 1, \dots, d$) ненулевые собственные числа и соответствующие ортонормированные собственные векторы матрицы \mathbf{S}^* . Положив $V_i = (\mathbf{X}^*)^T U_i / \sqrt{\lambda_i}$, мы вместо стандартного сингулярного разложения (2) получим разложение

$$\mathbf{X} = \mathbf{A} + \sum_{i=1}^d \mathbf{X}_i^* \quad (21)$$

с $\mathbf{X}_i^* = \sqrt{\lambda_i} U_i V_i^T$. На этапе группировки матрица \mathbf{A} войдет в одну из результирующих матриц как слагаемое. В частности, сама матрица может породить отдельную компоненту временного ряда с помощью процедуры ее диагонального усреднения.

Если матрица \mathbf{A} ортогональна всем \mathbf{X}_i^* , то матричное разложение (21) приводит к соотношению

$$\|\mathbf{X}\|_{\mathcal{M}}^2 = \|\mathbf{A}\|_{\mathcal{M}}^2 + \sum_{i=1}^d \|\mathbf{X}_i^*\|_{\mathcal{M}}^2$$

квадратов матричных норм Фробениуса соответствующих матриц.

Здесь мы рассмотрим два таких способа выбора матрицы \mathbf{A} .

Однократное центрирование — это центрирование строк траекторной матрицы. Здесь

$$\mathbf{A} = \mathcal{A}(\mathbf{X}) = [\mathcal{E}_1(\mathbf{X}) : \dots : \mathcal{E}_1(\mathbf{X})],$$

где каждая i -я компонента вектора $\mathcal{E}_1(\mathbf{X})$ ($i = 1, \dots, L$) равняется среднему значению i -х компонент векторов вложения X_1, \dots, X_K .

Таким образом, при однократном центрировании мы вместо $\text{span}(X_1, \dots, X_K)$ рассматриваем пространство $\text{span}(X_1^{(c)}, \dots, X_K^{(c)})$ с $X_i^{(c)} = X_i - \mathcal{E}_1(\mathbf{X})$. Другими словами, мы сдвигаем центр тяжести векторов вложения и затем используем сингулярное разложение полученной матрицы. Однократное центрирование является стандартной процедурой в анализе главных компонент, примененном к траекторной матрице как матрице многомерных статистических данных.

При *двойном центрировании* сингулярное разложение применяется к матрице, полученной из траекторной матрицы ряда с помощью последовательного вычитания из каждого ее элемента сначала среднего по строкам, а потом, в применении к получившейся

матрице, среднего по столбцам. Другими словами,

$$\mathbf{A} = \mathcal{A}(\mathbf{X}) + \mathcal{B}(\mathbf{X})$$

с $\mathcal{B}(\mathbf{X}) = [\mathcal{E}_{12}(\mathbf{X}) : \dots : \mathcal{E}_{12}(\mathbf{X})]^T$, где j -я компонента вектора $\mathcal{E}_{12}(\mathbf{X})$ ($j = 1, \dots, K$) равняется среднему всех компонент вектора $X_j^{(c)}$.

Заметим, что при однократном центрировании слагаемое \mathbf{A} имеет ту же форму, что и слагаемые разложения (21) при условии, что мы имеем дело с нормализованным вектором средних $U_{0(1)} = \mathcal{E}_1(\mathbf{X}) / \|\mathcal{E}_1(\mathbf{X})\|$ среди набора собственных векторов U_i . Действительно, $\mathbf{A} = U_{0(1)} Z_{0(1)}^T$ с $Z_{0(1)} = \|\mathcal{E}_1(\mathbf{X})\| \mathbf{1}_K$ (каждая из компонент вектора $\mathbf{1}_K \in \mathbb{R}^K$ равняется единице).

При двойном центрировании мы к собственным векторам добавляем кроме $U_{0(1)}$ еще один вектор $U_{0(2)} = \mathbf{1}_L / \sqrt{L}$. Тогда

$$\mathbf{A} = U_{0(1)} Z_{0(1)}^T + U_{0(2)} Z_{0(2)}^T$$

с $Z_{0(2)} = \sqrt{L} \mathcal{E}_{12}(\mathbf{X})$. Определим

$$\lambda_{0(1)} = \|Z_{0(1)}\| = \|\mathcal{E}_1(\mathbf{X})\| \sqrt{K} \text{ и } \lambda_{0(2)} = \|Z_{0(2)}\| = \|\mathcal{E}_{12}(\mathbf{X})\| \sqrt{L}.$$

Более того, пусть $V_{0(1)} = Z_{0(1)} / \sqrt{\lambda_{0(1)}}$ и $V_{0(2)} = Z_{0(2)} / \sqrt{\lambda_{0(2)}}$. Назовем $(U_{0(i)}, V_{0(i)}, \lambda_{0(i)})$ ($i = 1, 2$) *тройками первого и второго средних*.

Так как $\mathcal{A}(\mathbf{X})$ и $\mathcal{B}(\mathbf{X})$ ортогональны друг другу и всем остальным составляющим разложения, то для двойного центрирования справедливо

$$\|\mathbf{X}\|_{\mathcal{M}}^2 = \lambda_{0(1)} + \lambda_{0(2)} + \sum_{i=1}^d \lambda_i$$

(для однократного центрирования слагаемое $\lambda_{0(2)}$ опускается). Следовательно, доли троек средних среди всех троек равны

$$\lambda_{0(1)} / \|\mathbf{X}\|_{\mathcal{M}}^2, \quad \lambda_{0(2)} / \|\mathbf{X}\|_{\mathcal{M}}^2 \text{ и } \lambda_i / \|\mathbf{X}\|_{\mathcal{M}}^2.$$

Заметим, что базовый вариант метода SSA не использует никакого центрирования. Тем не менее есть ситуации, когда центрирование может дать некоторое преимущество.

Для вариантов SSA с центрированием можно естественным образом обобщить понятие делимости компонент ряда.

Тогда можно показать, что условия делимости гармоники от константного тренда с помощью SSA с однократным центрированием слабее, чем с помощью базового варианта SSA. При этом константному тренду соответствует только одна тройка, а именно тройка первого среднего. Аналогично можно показать, что условия делимости гармоники от линейного тренда с помощью SSA с двойным центрированием слабее, чем с помощью базового варианта SSA (с помощью базового варианта линейный тренд вообще точно не отделим ни от какого ряда). При этом линейному тренду соответствует обе тройки средних (и только они). Отсюда если ряд F может быть представлен в форме $F = F^{(1)} + F^{(2)}$, где $F^{(1)}$ — константный или линейный ряд, а ряд $F^{(2)}$ колеблется вокруг нуля, то применение SSA с соответственно однократным или двойным центрированием может дать улучшение делимости.

Разница между различными вариантами метода наиболее ярко проявляется при небольших длинах ряда; для длинных рядов как базовый вариант, так и вариант с центрированием дадут приблизительно одинаковые результаты.

Представляет интерес сравнение варианта SSA с двойным центрированием как метода выделения линейного тренда с линейным регрессионным анализом. Заметим, что эти два метода имеют различное происхождение и могут приводить к различным результатам. Что касается линейной регрессии, то это формальная процедура, служащая для линейной аппроксимации ряда по методу наименьших квадратов. Результатом ее является линейная функция, даже если ряд вообще не имеет линейной тенденции. Совершенно другой подход с точки зрения SSA с двойным центрированием, при котором линейная компонента выделяется (как правило, приближенно), только если достаточно сильная линейная тенденция действительно существует.

Рассмотрим разницу между методами на примере ряда F_N с $f_n = an + b + 5 \sin(2\pi n/T)$, $a = 1$, $b = 0$, периодом $T = 10$, длиной ряда $N = 19$. Возьмем длину окна L равной 10. При группировке троек первого и второго среднего мы получим в точности компоненту ряда, описываемую равенством $f_n^{(1)} = n$. При применении линейной регрессии в качестве оценок a и b получатся искаженные результаты: $\tilde{a} = 0.77$ и $\tilde{b} = 2.22$. Это расхождение является след-

ствием того, что при линейной регрессии ставится специфическая задача получения наилучшего предсказания, а не выделения тренда как такового.

5.2. Стационарные временные ряды и теплицев вариант SSA

Если длина N ряда F не очень велика и предполагается, что ряд F стационарен, то обычно рекомендуется заменить матрицу $\mathbf{S} = \mathbf{X}\mathbf{X}^T$ некоторой другой матрицей, учитывая стационарность исходного ряда. Заметим сперва, что мы можем рассматривать L -ковариационную матрицу $\mathbf{C} = \mathbf{S}/K$ вместо \mathbf{S} для получения сингулярного разложения траекторной матрицы \mathbf{X} . Действительно, разница между сингулярными разложениями на основе собственных чисел и собственных векторов матриц \mathbf{S} и \mathbf{C} лежит только в величине собственных чисел (у матрицы \mathbf{S} они в K раз больше); собственные векторы у них одинаковы. Следовательно, использование матриц \mathbf{S} и \mathbf{C} приводит к одному и тому же результату.

Обозначим $c_{ij} = c_{ij}(N)$ элементы L -ковариационной матрицы \mathbf{C} . Если временной ряд является стационарным и $K \rightarrow \infty$, то $\lim c_{ij} = R_f(|i-j|)$ при $N \rightarrow \infty$, где $R_f(k)$ обозначает соответствующее сдвигу k значение ковариационной функции ряда.

Следовательно, главной идеей является рассмотрение теплицевой модификации L -ковариационной матрицы с одинаковыми значениями \tilde{c}_{ij} на каждой диагонали $|i-j| = k$. Естественно, сходимость $\tilde{c}_{ij} \rightarrow R_f(|i-j|)$ должна сохраняться.

Возьмем одну из стандартных оценок ковариационной функции. Тогда для временного ряда $F = (f_0, \dots, f_{N-1})$ и длины окна L рассмотрим матрицу $\tilde{\mathbf{C}}$ с элементами

$$\tilde{c}_{ij} = \frac{1}{N - |i-j|} \sum_{m=0}^{N-|i-j|-1} f_m f_{m+|i-j|}, \quad 1 \leq i, j \leq L,$$

вместо L -ковариационной матрицы базового варианта метода SSA $\mathbf{C} = \mathbf{S}/K$ с элементами

$$c_{ij} = \frac{1}{K} \sum_{m=0}^{K-1} f_{m+i-1} f_{m+j-1}, \quad 1 \leq i, j \leq L.$$

Получив *теплицеву* L -ковариационную матрицу $\tilde{\mathbf{C}}$, мы вычислим ее ортонормированные собственные векторы H_1, \dots, H_L и затем рассмотрим разложение траекторной матрицы

$$\mathbf{X} = \sum_{i=1}^L H_i Z_i^T,$$

где $Z_i = \mathbf{X}^T H_i$. Таким образом, мы получим теплицево ортогональное матричное разложение. Положив $\lambda_i = \|Z_i\|^2$ и $Q_i = Z_i/\sqrt{\lambda_i}$ (здесь мы формально предполагаем, что $\tilde{\mathbf{C}}$ имеет полный ранг), мы придем к разложению траекторной матрицы \mathbf{X} в сумму элементарных матриц, подобно тому как это делалось при сингулярном разложении. Группировка и диагональное усреднение могут быть произведены стандартным образом. Заметим, что числа λ_i (их можно назвать квадратами *теплицевых сингулярных чисел*), вообще говоря, не совпадают с собственными числами матрицы $\tilde{\mathbf{C}}$.

Если исходный ряд является суммой константного ряда и некоторого стационарного ряда, то одним из способов сделать его стационарным — это произвести центрирование всего ряда целиком с помощью вычитания среднего значения, вычисленного по всему ряду. Другой способ — это применить операцию однократного центрирования к траекторной матрице ряда.

Теплицев вариант SSA может иметь преимущество при применении его к стационарному ряду, особенно это справедливо для коротких рядов. Однако надо иметь в виду, что для нестационарных рядов базовый вариант метода будет давать лучший результат. Также заметим, что теплицево разложение не является оптимальным (минимальным), что может привести к разложению компонент ряда на большее число слагаемых по сравнению с сингулярным разложением и тем самым затруднить идентификацию и усилить проблему близких собственных чисел (проблему отсутствия сильной делимости).

5.3. Близкие собственные числа и последовательный метод SSA

Как уже обсуждалось, близкие собственные числа в сингулярном разложении создают трудности, преодолеть которые сложно с помощью только варьирования параметра L . Тем не менее, можно

предложить несколько технологий, помогающих решить эту проблему. Одна из них — это добавление к компоненте ряда ей подобной, что усиливает (в общем случае, изменяет) ее вклад. Например, если в ряде присутствует константный тренд, то можно к ряду добавить некоторую константу. Если в ряде присутствует гармоническая составляющая и период ее известен, то к ряду можно добавить гармонику с таким же периодом. В этом случае вклад гармоники для модифицированного ряда изменится (будет больше или меньше в зависимости от соотношения фаз исходной и добавленной гармоник). Такие преобразования ряда в некоторых случаях позволяют изменить соотношение собственных чисел, соответствующих разным компонентам, и, следовательно, преодолеть проблему отсутствия сильной разделимости.

Другим способом решения проблемы смешения компонент временного ряда (т. е. проблемы совпадающих собственных чисел для слабо разделимых компонент) является последовательное применение SSA к ряду (так называемый *последовательный метод SSA*).

Опишем двухшаговый последовательный SSA. Сначала мы извлекаем несколько компонент ряда с помощью некоторой длины окна L_1 . Затем снова применяем базовый метод SSA к остаткам и извлекаем оставшиеся компоненты ряда. Длина окна L_2 на втором шаге, вообще говоря, отличается от L_1 .

Имея два набора компонент ряда, мы можем затем группировать их вместе. Например, если на первом шаге мы извлекли основную часть тренда, а на втором — его уточняющие компоненты, то затем мы можем их сложить и получить уточненный тренд.

Стандартной ситуацией, при которой последовательный SSA хорошо работает, является следующая. Предположим, что тренд ряда имеет сложную форму. Если мы выберем большую длину окна L , то некоторые компоненты тренда ряда (а сложный тренд раскладывается на много компонент) могут смешаться с другими компонентами ряда. При выборе маленького L мы сможем выделить тренд, но это может привести к смешиванию остальных интересующих нас компонент ряда (так как при маленьких длинах окна приближенная разделимость может отсутствовать).

Одним из способов решения этой проблемы является последовательное применение SSA, сначала с маленькой длиной окна L_1 для выделения тренда, а затем применение SSA к оставшейся части ряда с большим L_2 .

6. Приложение А: сингулярное разложение матриц

Этот раздел посвящен сингулярному разложению вещественных матриц, которое составляет основной математический аппарат базового метода SSA.

6.1. Существование и единственность

Пусть \mathbf{X} — ненулевая матрица с $L > 1$ строками и $K > 1$ столбцами. Тогда $\mathbf{S} \stackrel{\text{def}}{=} \mathbf{X}\mathbf{X}^T$ — симметричная матрица $L \times L$. Как и любая симметричная матрица, матрица \mathbf{S} имеет L линейно-независимых *собственных векторов*, т. е. линейно независимых векторов U_1, \dots, U_L , таких что

$$\mathbf{S}U_i = \lambda_i U_i,$$

где λ_i — вещественные числа, называемые *собственными числами* матрицы \mathbf{S} . Линейная оболочка собственных векторов U_i , $i = 1, \dots, L$, называется *собственным подпространством*.

Мы можем выбрать собственные векторы U_i *ортонормальными*, т. е. $(U_i, U_j) = 0$ для $i \neq j$ (свойство ортогональности) и $\|U_i\| = 1$ (нормировка), где (X, Y) — скалярное произведение векторов, а $\|X\| = \sqrt{(X, X)}$ — *норма* вектора X .

Более того, матрица \mathbf{S} является *неотрицательно определенной*, т. е. $\lambda_i \geq 0$ для всех $i = 1, \dots, L$.

Будем предполагать, что собственные числа λ_i расположены в *убывающем порядке*: $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L \geq 0$.

Обозначим d — число ненулевых собственных чисел матрицы \mathbf{S} . Если $d < L$, $\lambda_d > 0$ и $\lambda_{d+1} = 0$, то все остальные собственные числа с номерами, большими d , являются нулевыми. Если $\lambda_L > 0$, то $d = L$. Так как d равняется рангу матрицы \mathbf{X} , мы получаем неравенство $d \leq \min(L, K)$.

Для $1 \leq i \leq d$ положим

$$V_i = \frac{1}{\sqrt{\lambda_i}} \mathbf{X}^T U_i. \quad (22)$$

Предложение 6.1. *Векторы U_i и V_i обладают следующими свойствами.*

1. Пусть $1 \leq i, j \leq d$. Тогда $(V_i, V_j) = 0$ для $i \neq j$ и $\|V_i\| = 1$. Если $i > d$, то $\mathbf{X}^T U_i = \mathbf{0}_K \in \mathbf{R}^K$, где $\mathbf{0}_K$ — нулевой вектор.
2. V_i является собственным вектором матрицы $\mathbf{X}^T \mathbf{X}$, соответствующим собственному числу λ_i .
3. Если $1 \leq i \leq d$, то

$$U_i = \frac{1}{\sqrt{\lambda_i}} \mathbf{X} V_i.$$

4. Если $K > d$, то все остальные $K - d$ собственных векторов матрицы $\mathbf{X}^T \mathbf{X}$ соответствуют нулевым собственным числам.
5. Имеет место следующее равенство:

$$\mathbf{X} = \sum_{i=1}^d \sqrt{\lambda_i} U_i V_i^T. \quad (23)$$

Определение 6.1. Равенство (23) называется *сингулярным разложением (SVD)* матрицы \mathbf{X} , числа $\sqrt{\lambda_i}$ — *сингулярными числами* матрицы \mathbf{X} , а векторы U_i и V_i называются *левым* и *правым сингулярными векторами* матрицы \mathbf{X} . Набор $(\sqrt{\lambda_i}, U_i, V_i)$ называется *i -й собственной тройкой* матрицы \mathbf{X} .

Обсудим единственность сингулярного разложения (23). Конечно, единственность в данном случае не может пониматься буквально. Во-первых, так как $(-U_i)$ также является собственным вектором матрицы \mathbf{S} , соответствующим собственному числу λ_i , то в (23) мы можем заменить одну или более пар (U_i, V_i) на $(-U_i, -V_i)$, не меняя слагаемых в правой части (23).

Более того, совпадающие собственные числа порождают бесконечное число разных сингулярных разложений. Если, например, $\lambda \stackrel{\text{def}}{=} \lambda_1 = \lambda_2 > \lambda_3$, то определено только двумерное собственное подпространство, соответствующее собственному числу λ , и в качестве собственных векторов могут быть выбраны любые векторы U_1 и U_2 , составляющие ортонормированный базис этого подпространства. Таким образом, векторы U_1 и U_2 (также, как и векторы V_1, V_2) не однозначно определены. Это означает, что если $\lambda_1 = \lambda_2 > \lambda_3$, то обе матрицы $\sqrt{\lambda_1} U_1 V_1^T$ и $\sqrt{\lambda_2} U_2 V_2^T$ имеют смысл только в сумме (сумма не зависит от выбора U_1 и U_2), но не каждая по отдельности.

Учитывая приведенные рассуждения, единственность сингулярного разложения может быть сформулирована в виде следующего утверждения.

Предложение 6.2. Пусть P_1, \dots, P_L и Q_1, \dots, Q_L — некоторые ортонормированные системы в \mathbf{R}^L и \mathbf{R}^K , соответственно. Предположим, что существуют такие неотрицательные константы $c_1 \geq \dots \geq c_L \geq 0$, что

$$\mathbf{X} = \sum_{i=1}^L c_i P_i Q_i^T. \quad (24)$$

Рассмотрим сингулярное разложение (23) матрицы \mathbf{X} . Тогда:

1. $c_d > 0$ и $c_{d+1} = \dots = c_L = 0$.
2. $c_i^2 = \lambda_i$ для $1 \leq i \leq d$.
3. Для каждого $i = 1, \dots, d$ вектор P_i является собственным вектором матрицы $\mathbf{X}\mathbf{X}^T$, соответствующим собственному числу λ_i .
4. $Q_i = \mathbf{X}^T P_i / \sqrt{\lambda_i}$ ($i = 1, \dots, d$).
5. Если все числа c_i различны, то (24) совпадает с (23) с точностью до знака U_i и V_i .

Следствие 6.1. Пусть $I \subset \{1, \dots, d\}$, $J = \{1, \dots, d\} \setminus I$. Положим

$$\mathbf{X}_I = \sum_{i \in I} \sqrt{\lambda_i} U_i V_i^T$$

и $\mathbf{X}_J \stackrel{\text{def}}{=} \mathbf{X} - \mathbf{X}_I$. Тогда разложение

$$\mathbf{X}_J = \sum_{i \in J} \sqrt{\lambda_i} U_i V_i^T$$

является сингулярным разложением матрицы \mathbf{X}_J .

6.2. Матричная форма сингулярного разложения

Сингулярное разложение (23) может быть переписано в матричной форме следующим образом. Положим $\mathbf{U}_d = [U_1 : \dots : U_d]$, $\mathbf{V}_d = [V_1 : \dots : V_d]$, и пусть Λ_d — диагональная $d \times d$ матрица с собственными числами λ_i в качестве i -х диагональных элементов. Тогда (23) может быть записано в виде

$$\mathbf{X} = \mathbf{U}_d \Lambda_d^{1/2} \mathbf{V}_d^T, \quad (25)$$

который является стандартной *матричной формой сингулярного разложения*.

Равенство (25) может быть переписано в форме, известной как квази-диагональное представление матрицы \mathbf{X} . Хорошо известно, что при соответствующем выборе ортонормированного базиса в \mathbf{R}^L любая симметричная $L \times L$ матрица имеет диагональное представление. Как следует из (25), можно выбрать подходящие базисы в \mathbf{R}^L и \mathbf{R}^K так, чтобы получить аналогичное представление для прямоугольной матрицы \mathbf{X} .

Пусть $\mathbf{U} = [U_1 : \dots : U_L]$ и $\mathbf{V} = [V_1 : \dots : V_K]$. (Заметим, что в случае $d < K$ мы выбираем V_{d+1}, \dots, V_K как ортонормированную систему собственных векторов, соответствующих нулевым собственным числам матрицы.)

Матрицы \mathbf{U} и \mathbf{V} — $L \times L$ и $K \times K$ *унитарные матрицы* (или *матрицы вращения*). Для матрицы \mathbf{U} это означает, что для любых векторов $X, Y \in \mathbf{R}^L$ равенство $(\mathbf{U}X, \mathbf{U}Y) = (X, Y)$ справедливо и, следовательно, матрица \mathbf{U} , рассматриваемая как линейное отображение $\mathbf{R}^L \mapsto \mathbf{R}^L$, сохраняет векторные нормы и углы между векторами. Другой характеристикой свойства вращения является тождество $\mathbf{U}^{-1} = \mathbf{U}^T$.

Обозначим Λ матрицу такой же размерности, как и исходная матрица \mathbf{X} с диагональными элементами $\lambda_{ii} = \lambda_i$ при $1 \leq i \leq d$ и всеми остальными элементами, равными нулю. Тогда (25) может быть переписано в виде

$$\mathbf{X} = \mathbf{U}\Lambda^{1/2}\mathbf{V}^T \quad \text{or} \quad \Lambda^{1/2} = \mathbf{U}^T\mathbf{X}\mathbf{V}. \quad (26)$$

Равенство (26) имеет смысл *квази-диагонального представления* матрицы \mathbf{X} . При подходящих базисах U_1, \dots, U_L в \mathbf{R}^L и V_1, \dots, V_K в \mathbf{R}^K (другими словами, с помощью двух правильно подобранных вращений) любая прямоугольная $L \times K$ матрица имеет квази-диагональное представление $\Lambda^{1/2}$. Термин «вращение» здесь использован, так как переход от одного ортонормированного базиса к другому базису линейного пространства осуществляется с помощью матрицы вращения, такой как \mathbf{U} и \mathbf{V} .

6.3. Ортогональные матричные разложения и SVD

В сингулярном разложении (23) положим $\mathbf{X}_i = \sqrt{\lambda_i} U_i V_i^T$. Тогда (23) может быть переписано в виде

$$\mathbf{X} = \mathbf{X}_1 + \dots + \mathbf{X}_d. \quad (27)$$

Матрицы \mathbf{X}_i имеют ранг, равный 1 (будем называть такие матрицы *элементарными*). Более того, матрицы \mathbf{X}_i являются *биортогональными* в том смысле, что $\mathbf{X}_i \mathbf{X}_j^T = \mathbf{0}_{LL}$ и $\mathbf{X}_i^T \mathbf{X}_j = \mathbf{0}_{KK}$ для $i \neq j$.

Предложение 6.3. *Сингулярное разложение (27) является единственным (в смысле предложения 6.2) разложением матрицы \mathbf{X} на биортогональные элементарные матрицы. Среди всех разложений на элементарные матрицы оно является минимальным (состоит из минимального количества элементарных матриц).*

Рассмотрим линейное пространство $\mathcal{M}_{L,K}$ вещественных $L \times K$ матриц со стандартными операциями сложения и умножения на константу. Очевидно, что размерность такого пространства равна LK . Определим *скалярное произведение матриц* следующим образом. Пусть $\mathbf{X} = (x_{ij})_{i,j=1}^{L,K}$ и $\mathbf{Y} = (y_{ij})_{i,j=1}^{L,K}$ — матрицы из $\mathcal{M}_{L,K}$. Тогда

$$\langle \mathbf{X}, \mathbf{Y} \rangle_{\mathcal{M}} = \sum_{i=1}^L \sum_{j=1}^K x_{ij} y_{ij}. \quad (28)$$

Стандартным образом равенство

$$\|\mathbf{X}\|_{\mathcal{M}}^2 = \langle \mathbf{X}, \mathbf{X} \rangle_{\mathcal{M}} = \sum_{i=1}^L \sum_{j=1}^K x_{ij}^2$$

задает квадрат матричной нормы (обычно называемой *матричной нормой Фробениуса*), а тождество $\text{dist}_{\mathcal{M}}(\mathbf{X}, \mathbf{Y}) = \|\mathbf{X} - \mathbf{Y}\|_{\mathcal{M}}$ имеет смысл расстояния между матрицами \mathbf{X} и \mathbf{Y} .

Скалярное произведение (28) — это обычное скалярное произведение векторов в \mathbb{R}^{LK} (с элементами x_{ij} и y_{ij}) и не зависит от прямоугольной структуры матриц. В частности, (28) не зависит от перестановки матричных элементов и, следовательно, не учитывает

многие важные матричные характеристики, такие как, например, ранг матрицы.

Сингулярное разложение (27) является разложением на \mathcal{M} -ортогональные матрицы, $\|\mathbf{X}_i\|^2 = \lambda_i$. Поэтому

$$\|\mathbf{X}\|_{\mathcal{M}}^2 = \lambda_1 + \dots + \lambda_d$$

и собственное число λ_i можно рассматривать как вклад матрицы $\mathbf{X}_i = \sqrt{\lambda_i}U_iV_i^T$ (или вклад собственной тройки $(\sqrt{\lambda_i}, U_i, V_i)$) в сингулярное разложение (27).

6.4. Оптимальные свойства сингулярного разложения

Оптимальные свойства сингулярного разложения основаны на экстремальных свойствах собственных чисел/векторов. Эти свойства хорошо известны.

Обозначим \mathcal{M}_k множество всех матриц, ранг которых не превышает k .

Пусть $(\sqrt{\lambda_i}, U_i, V_i)$ — собственные тройки сингулярного разложения (23). Оптимальные свойства SVD могут быть сформулированы следующим образом.

Предложение 6.4.

1. $\min_{\mathbf{Y} \in \mathcal{M}_k} \|\mathbf{X} - \mathbf{Y}\|_{\mathcal{M}}^2 = \sum_{i=k+1}^d \lambda_i$.
2. Если взять

$$\mathbf{Y}_0 = \sum_{i=1}^k \sqrt{\lambda_i}U_iV_i^T \in \mathcal{M}_k,$$

то имеет место равенство

$$\|\mathbf{X} - \mathbf{Y}_0\|_{\mathcal{M}}^2 = \min_{\mathbf{Y} \in \mathcal{M}_k} \|\mathbf{X} - \mathbf{Y}\|_{\mathcal{M}}^2.$$

Предложение 6.4 говорит о том, что *сумма первых k членов сингулярного разложения дает наилучшую аппроксимацию исходной матрицы матрицей ранга, не превосходящего k* . Первый пункт предложения дает вид ошибки такой аппроксимации.

Сформулируем данное свойство оптимальности несколько по-другому. Пусть $\mathbf{X} = [X_1 : \dots : X_K]$.

Предложение 6.5. *Минимум*

$$\min_{\mathcal{L}} \sum_{i=1}^K \text{dist}^2(X_i, \mathcal{L})$$

по всем подпространствам размерности, не превосходящей k , достигается на $\mathcal{L} = \text{span}(U_1, \dots, U_k)$.

Естественной характеристикой этих двух (эквивалентных) оптимальных аппроксимаций является

$$\frac{\|\mathbf{X} - \mathbf{Y}_0\|_{\mathcal{M}}^2}{\|\mathbf{X}\|_{\mathcal{M}}^2} = \frac{\lambda_{k+1} + \dots + \lambda_d}{\lambda_1 + \dots + \lambda_d}.$$

Если иметь дело не с оптимальной аппроксимацией и положить

$$\mathbf{X}_I = \sum_{i \in I} \sqrt{\lambda_i} U_i V_i^T$$

с $I = \{j_1, \dots, j_k\} \subset \{1, \dots, d\}$, $j_1 > \dots > j_k$, и $k < d$, то

$$1 - \frac{\|\mathbf{X} - \mathbf{X}_I\|_{\mathcal{M}}^2}{\|\mathbf{X}\|_{\mathcal{M}}^2} = \frac{\lambda_{j_1} + \dots + \lambda_{j_k}}{\lambda_1 + \dots + \lambda_d}. \quad (29)$$

Определение 6.2. Величина, задаваемая формулой (29), называется *долей собственных троек с номерами j_1, \dots, j_k* .

Другое описание оптимальных свойств сингулярного разложения связано с так называемыми главными векторами набора векторов $X_1, \dots, X_K \in \mathbf{R}^L$. Пусть $X, P \in \mathbf{R}^L$, $X \neq \mathbf{0}_L$, $\|P\| = 1$. Тогда $(X, P)P$ является проекцией X на одномерное пространство $\mathfrak{L}_P = \text{span}(P)$, и $c = |(X, P)|$ — норма этой проекции. Число $c = c(P)$ можно рассматривать как меру качества аппроксимации вектора X пространством \mathfrak{L}_P : чем больше $c = c(P)$, тем лучше X аппроксимируется $\text{span}(P)$.

Если мы хотим найти P , такое что \mathfrak{L}_P аппроксимирует набор векторов X_1, \dots, X_K наилучшим образом, то мы приходим к следующей оптимизационной задаче: найти такой вектор P_0 , что $\|P_0\| = 1$ и

$$\nu_1 \stackrel{\text{def}}{=} \sum_{i=1}^K (X_i, P_0)^2 = \max_P \sum_{i=1}^K (X_i, P)^2, \quad (30)$$

где максимум в правой части (30) берется по всем $P \in \mathbf{R}^L$ с $\|P\| = 1$. Решение этой задачи может быть описано в терминах сингулярного разложения.

Предложение 6.6. *Рассмотрим матрицу $\mathbf{X} = [X_1 : \dots : X_K]$ и ее сингулярное разложение (23). Тогда:*

1. Вектор $P_0 = U_1$ дает решение задачи (30) с $\nu_1 = \lambda_1$.
2. Пусть P_0 – решение следующей оптимизационной задачи:

$$\nu_k \stackrel{\text{def}}{=} \sum_{i=1}^K (X_i, P_0)^2 = \max_P^{(k)} \sum_{i=1}^K (X_i, P)^2, \quad (31)$$

где максимум в правой части (31) берется по всем $P \in \mathbf{R}^L$, таким что $\|P\| = 1$ и $(P, U_i) = 0$ при $1 \leq i < k$. Если $k \leq d$, то вектор $P_0 = U_k$ является решением задачи (31) и $\nu_k = \lambda_k$. Если $k > d$, то $\nu_k = 0$.

Предложение 6.6 позволяет нам назвать вектор U_i *i -м главным вектором* набора векторов X_1, \dots, X_K .

Положим $c_j(U_i) = (X_j, U_i)$. Так как

$$X_j = \sum_{i=1}^d c_j(U_i) U_i,$$

коэффициент $c_j(U_i)$ называется *i -й главной компонентой* вектора X_j . А вектор

$$Z_i = (c_1(U_i), \dots, c_K(U_i))^T = \mathbf{X}^T U_i$$

естественно назвать *вектором i -х главных компонент*.

Заметим, что, в силу (22), $Z_i = \sqrt{\lambda_i} V_i$ и сингулярное разложение (23) можно интерпретировать как одновременное разложение всех столбцов матрицы \mathbf{X} по базису из их главных векторов. Такая интерпретация является стандартной в анализе главных компонент в статистике, когда столбцы матрицы \mathbf{X} образуют L -мерную выборку объема K . Заметим, что термин *факторные векторы*, применяемый для обозначения векторов V_i , также имеет статистическое происхождение.

Точно таким же образом векторы V_i являются главными векторами для строк матрицы \mathbf{X} , $\sqrt{\lambda_i} U_i$ – векторы из главных компонент. Таким образом, сингулярное разложение (23) порождает две системы главных векторов и два разложения по этим системам.

7. Приложение В: пример анализа временного ряда

Продемонстрируем методику применения метода «Гусеница»-SSA на примере анализа временного ряда FORT (объемы месячных продаж крепленых вин в Австралии с января 1984 года по июнь 1994 года (рис. 3*)). Так как мы будем рассматривать как сам ряд длины $N = 174$, так и его подряд, состоящий из первых 120 точек, то там, где это необходимо, будем использовать обозначения FORT174 для всего ряда и FORT120 для укороченного ряда. Приведенные графики будут одновременно являться иллюстрацией теоретических аспектов метода.

Визуальный анализ ряда, изображенного на рис. 3, говорит о том, что ряд имеет тренд, который должен хорошо описываться линейной функцией либо убывающей экспонентой, а также сезонное поведение довольно сложной меняющейся формы. Периодограмма центрированного ряда (для увеличения резкости изображения она построена по первым 168 точкам, так как 168 кратно двенадцати, числу месяцев в году) подтверждает такое предположение (рис. 4).

Поставим задачу поиска разложения данного ряда на три компоненты — тренд, сезонную компоненту и шум.

Выбор длины окна. Посчитаем, какая может быть размерность у сигнала (состоящего из тренда и периодической компоненты). Тренд ряда FORT скорее всего описывается 1–2 собственными тройками, сезонная компонента раскладывается на гармонические составляющие с частотами $1/12$ (годовая периодичность), $1/6=2/12$ (полугодовая), $1/4=3/12$, $1/3=4/12$, $1/2.4=5/12$, $1/2=6/12$. Все гармоники с частотами, меньшими, чем $1/2$, имеют размерность 2, гармоника с периодом два имеет размерность 1 (мы предполагаем, что амплитуды каждой их гармоник меняются по экспоненте, возможно с разными показателями для разных частот). Таким образом, предполагаемая размерность ряда не должна превосходить 13. Если бы ряд не содержал шума и компоненты ряда были строго разделимы, то для получения разложения ряда на тренд и сезонную компоненту достаточно было бы взять длину окна, равной 13. Однако для реальных данных точная разделимость практически не встречается, поэтому необходимо воспользоваться результатами о

*Рис. 3–13 см. на с. 69–74

приближенной (асимптотической) разделимости медленно меняющегося тренда и гармоник. Для достижения лучшей точности мы должны выбрать длину окна большой, близкой к половине длины ряда (так как скорость сходимости ошибки разделимости к нулю имеет порядок $1/\min(L, K)$, где $K = N - L + 1$, N — длина ряда). Так как кратность L и K периоду должна улучшать отделимость периодической компоненты, причем более важна кратность меньшего из этих чисел, то выберем длину окна, равной $L = 84$ ($N = 174$, следовательно, $K = 91$).

Подход к идентификации собственных троек. Рассмотрим результат сингулярного разложения траекторной матрицы ряда при таком выборе длины окна. На рис. 5 приведено изображение собственных векторов из первых шести собственных троек сингулярного разложения. Заметим, что вид факторных векторов, или правых сингулярных векторов, практически не отличается от вида левых, так как длина окна близка к половине длины ряда. Если бы мы взяли длину окна меньше, то собственные векторы имели бы более регулярный вид по сравнению с факторными, которые содержали бы в себе, например, изменение амплитуд для гармонических компонент. Для идентификации собственных троек воспользуемся результатами о виде собственных векторов, соответствующих тренду и гармоникам при условии их приближенной разделимости.

Идентификация тренда. Начнем с идентификации тренда. Мы знаем, что сингулярные (в частности, собственные) векторы имеют в целом такой же вид, как и компонента исходного ряда, которой они соответствуют. Поэтому на одномерных диаграммах собственных векторов нужно найти медленно меняющиеся собственные векторы. В данном случае только один, а именно первый собственный вектор, имеет требуемый вид. Так как получилось, что в данном случае тренд описывается единственной собственной тройкой, то это означает, что тренд аппроксимируется экспонентой. Чем сложнее форма тренда, тем больше его (приближенная) размерность и тем большее число собственных троек ему соответствует.

Идентификация гармоник. Займемся теперь идентификацией гармонических (возможно, с меняющейся амплитудой) компонент, порожденных сезонной компонентой исходного ряда. На рис. 5 видно, что собственные тройки с номерами 2–6, возможно, соответствуют каким-либо гармоникам, так как имеют регу-

лярное периодическое поведение. Однако поскольку гармоника с периодом большим, чем 2, порождает две собственные тройки (косинус с периодом 2 порождает только одну собственную тройку, сингулярные векторы которой имеют пилообразный вид, а таких на одномерных диаграммах в данном случае не обнаружено), то искать относящиеся к гармоникам пары собственных троек удобнее на двумерных диаграммах. Так как при достаточно большой длине ряда соответствующая пара собственных чисел имеет близкие значения, то достаточно рассматривать двумерные диаграммы собственных векторов из соседних, упорядоченных по собственным значениям, собственных троек. На рис. 6 можно различить регулярные двумерные изображения, образующие двумерные траектории с вершинами, лежащими на кривой, имеющей спиралеобразную форму. Это означает, что соответствующая пара собственных векторов порождена модулированной гармонической компонентой исходного ряда. Таким образом, получаем, что собственные тройки (используем аббревиатуру ET — eigentriples) ET2,3 соответствуют периоду 12, ET4,5 — периоду 4, ET6,7 — периоду 6, ET8,9 — периоду 2.4 (говоря о дробном периоде, мы имеем в виду гармонику с частотой, обратной этому периоду, в данном случае — $5/12$), ET10,11 — периоду 3.

Вспомогательные характеристики. Рассмотрим, какая дополнительная информация может помочь для идентификации собственных троек (или подтвердить то, что получено). Рис. 7, на котором изображены логарифмы собственных значений, подтверждает найденные пары собственных троек (каждой паре соответствует «ступенька»).

Другим полезным средством идентификации является матрица w -корреляций между восстановленными компонентами ряда. На рис. 8 изображены первые 6 формально восстановленных, по каждой собственной тройке отдельно, компонент исходного ряда (будем называть такие восстановленные компоненты *элементарными*). Напомним, что w -корреляция — это взвешенная корреляция между восстановленными рядами и равенство ее нулю является необходимым условием разделимости компонент ряда. Рис. 9 подтверждает проведенную идентификацию тем, что внутри пары, порождаемой гармоникой, w -корреляция высока, а между парами и трендом — близка к нулю (это показывает белый цвет соответствующих ячеек матрицы корреляций).

Отделение сигнала от шума. Отдельно остановимся на вопросе отделения компонент, относящихся к сигналу, от шумовых компонент. Во-первых, нерегулярное поведение сингулярных векторов может говорить о принадлежности их к набору, порожденному шумовой компонентой (эти случаи нужно не путать с перемешиванием компонент, порожденным отсутствием сильной разделимости рядов). Во-вторых, медленное, почти без скачков, убывание собственных значений с некоторого номера также говорит об этом. В-третьих, большой набор собственных троек, порождающих коррелирующие друг с другом восстановленные компоненты, скорее всего относится к шуму. Рис. 9 показывает, что собственные тройки с номерами 14–84 как раз образуют такой блок. Вопрос вызывает пара ЕТ12,13. С одной стороны, она довольно хорошо отделяется от остатка. С другой стороны, периодограмма ряда показывает, что период у восстановленной на основе пары ЕТ12,13 компоненты равен примерно 2.33. Такой период, необъяснимый наличием сезонности, может быть порожден шумом, однако, в принципе, причиной может быть также сильная модуляция гармоник с периодом 2.4. На данном конкретном ряде невозможно сделать достоверный вывод, поэтому мы все же отнесем ЕТ12,13 к шуму.

И последним подтверждением правильности разделения (естественно, приближенного) сигнала и шума является проверка на принадлежность к шуму компоненты ряда, полученной с помощью восстановления по, так сказать, «шумовым» собственным тройкам, стандартными статистическими методами. На рис. 10 изображено разложение исходного ряда на три компоненты — тренд (ЕТ1; на фоне исходного ряда), периодику (ЕТ2–11) и шум (ЕТ12–84). Статистические критерии подтверждают (не отвергают) то, что третья компонента является реализацией белого шума (вероятностные уровни трех разных критериев независимости наблюдений больше, чем 0.4).

Разложение исходного ряда на компоненты. Таким образом, рис. 10, полученный на основе группировки элементов сингулярного разложения и последующего диагонального усреднения, дает решение поставленной в начале раздела задачи построения разложения исходного ряда на интерпретируемые аддитивные компоненты. Табл. 4 содержит значения w -корреляций между изображенными на рис. 10 компонентами.

Таблица 4. w -Корреляции

	ET1	ET2-11	ET12-84
ET1	1	0	0
ET2-11	0	1	0.016
ET12-84	0	0.016	1

Кроме общего разложения может представлять интерес более подробное исследование ряда. Для нашего ряда представляет интерес более детальное исследование нерегулярного на вид поведения сезонной компоненты. На рис. 11 представлены первые три компоненты разложения сезонной компоненты на гармоники. Обращает на себя внимание уменьшение амплитуды годовой гармоники, более-менее постоянное поведение полугодовой и увеличение амплитуды 4-месячной гармоники. Заметим, что стандартные статистические методы обычно предполагают либо неизменность амплитуды (аддитивные модели), либо ее изменение, одинаковое для всех гармоник, пропорциональное тренду (мультипликативные модели). Видно, что в данном случае эти оба варианта не годятся.

Проблема отсутствия сильной разделимости. При анализе исходного временного ряда нам не пришлось столкнуться с проблемой перемешивания компонент ряда, т. е. с проблемой отсутствия сильной разделимости, вызванной близкими собственными числами (весами), соответствующими разным компонентам. Чтобы продемонстрировать эту проблему, укоротим исходный ряд до первых 120 точек. Соответственно выберем длину окна, равной 60. На рис. 12 изображена матрица w -корреляций. Сравнивая ее с матрицей, изображенной на рис. 9, мы видим, что по-прежнему, как и при анализе полного ряда, собственные тройки, начиная с двенадцатой, можно отнести к шуму. Однако темный блок, который образуют ET8-11, показывает, что, скорее всего, две гармонические компоненты смешались. Периодограммный анализ собственных векторов подтверждает это — смешались гармоники с частотами $1/3$ и $1/2.4$. Для выделения всей сезонной компоненты целиком такое смешение не столь существенно (оно лишь вызывает неудобства при идентификации). Если же мы хотели бы выделить, например, квартальную (трехмесячную) компоненту отдельно, то отсутствие сильной разделимости не позволило бы нам это сделать. Сравнение периодограмм (почти) полного исходного ряда (рис. 4) и его первых

120 точек (рис. 13) объясняет, почему проблемы отсутствия сильной разделимости не было для исходного ряда: вклады частот $1/3$ и $1/2.4$ в первом случае немного различаются, в то время как во втором случае они практически совпадают.

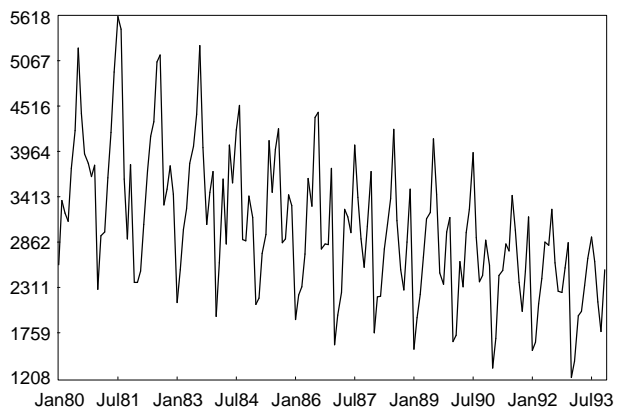


Рис. 3. FORT174: график исходного ряда.

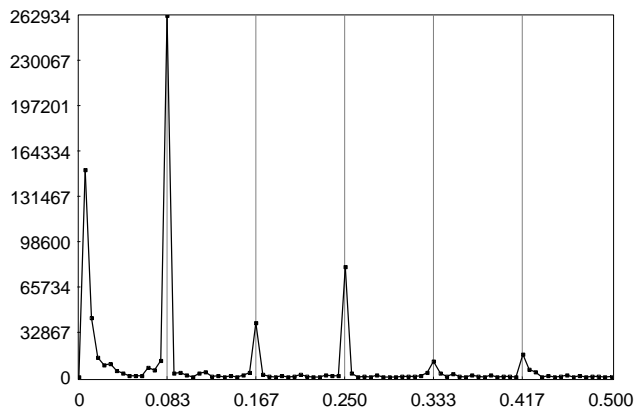


Рис. 4. FORT174: периодограмма центрированного ряда (первые 168 точек).

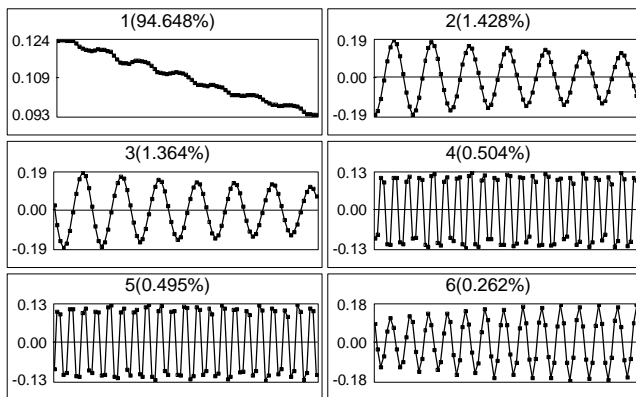


Рис. 5. FORT174: одномерные диаграммы собственных векторов.

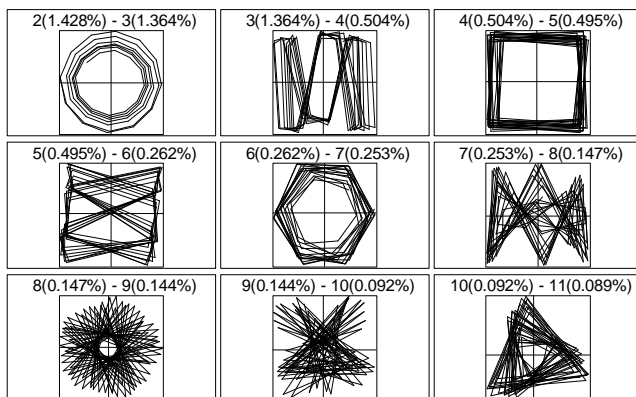


Рис. 6. FORT174: двумерные диаграммы собственных векторов.

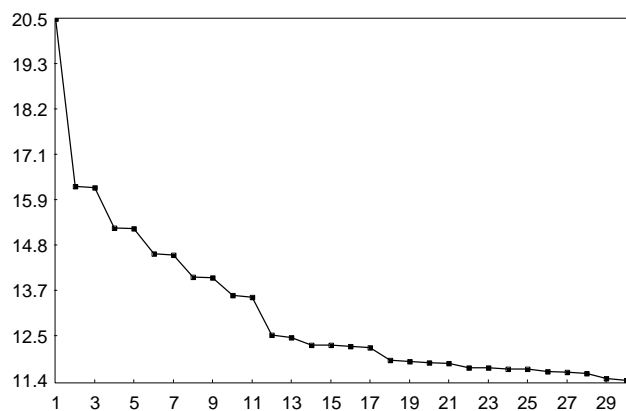


Рис. 7. FORT174: логарифмы первых 30 собственных чисел.

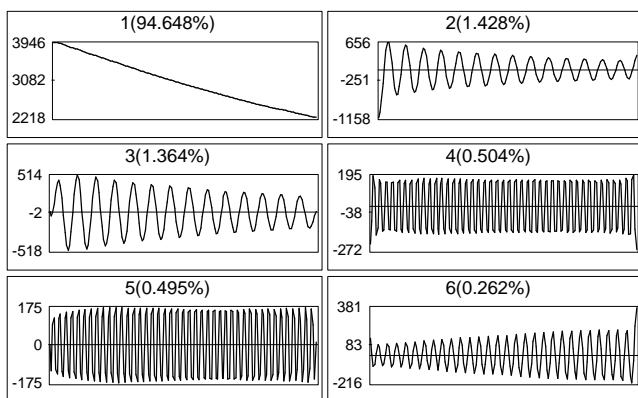


Рис. 8. FORT174: элементарные восстановленные компоненты ряда.

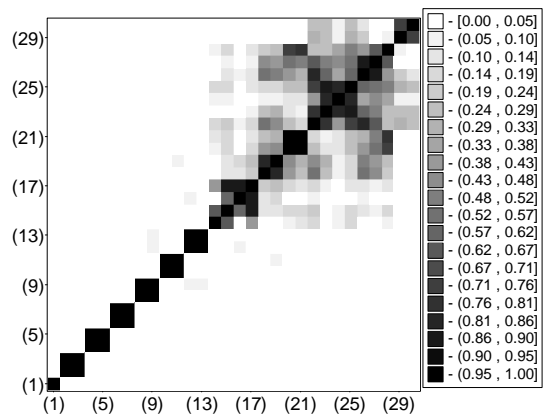


Рис. 9. FORT174: матрица w -корреляций элементарных восстановленных компонент ряда.

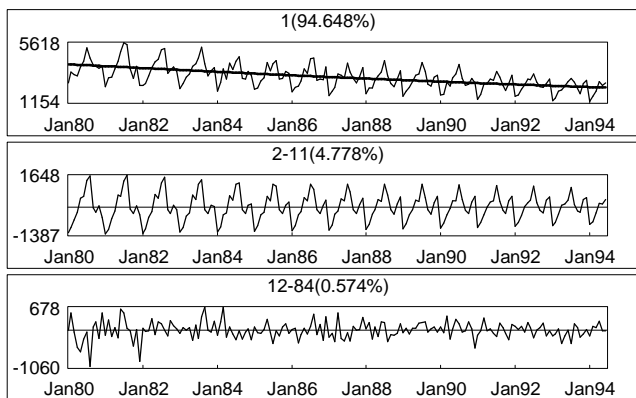


Рис. 10. FORT174: разложение ряда на тренд (вверху, на фоне исходного ряда), сезонную компоненту (в середине) и шум (внизу).

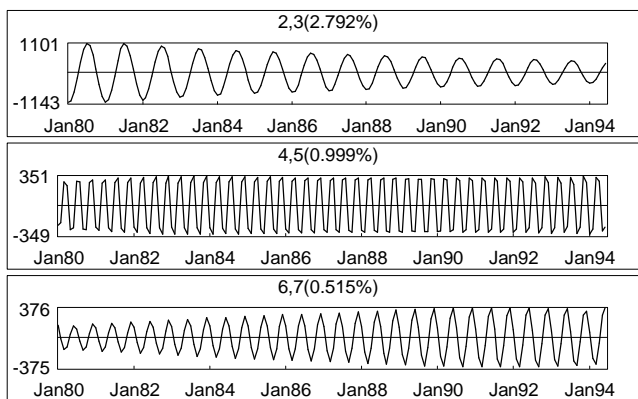


Рис. 11. FORT174: разложение сезонной компоненты ряда на гармоники.

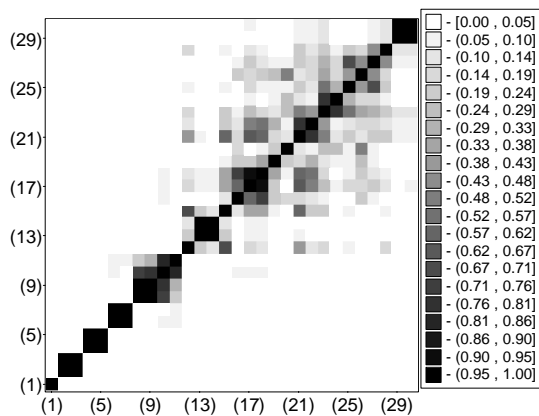


Рис. 12. FORT120: матрица w -корреляций элементарных восстановленных компонент ряда.

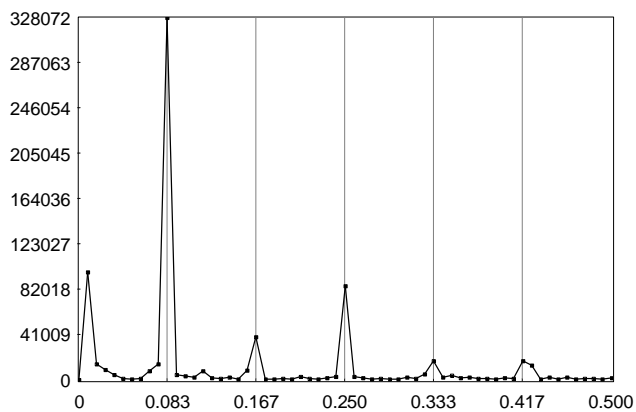


Рис. 13. FORT120: периодограмма центрированного ряда.

Литература

1. *Elsner J.B. and Tsonis A.A.* Singular Spectrum Analysis. A New Tool in Time Series Analysis. New York and London: Plenum Press, 1996. 164 p.
2. *Главные компоненты временных рядов: метод «Гусеница»* / Под ред. Д.Л. Данилова, А.А. Жиглявского. СПб: Пресском, 1997. 308 с.
3. *Golyandina N., Nekrutkin V., and Zhigljavsky A.* Analysis of Time Series Structure: SSA and Related Techniques. London: Chapman & Hall/CRC, 2001. 305 p.