Тензорный метод SSA в задаче декомпозиции многомерных временных рядов

Сёмкин Кирилл[1] Вадим Стрижов[2]

Аннотация

Декомпозиция временного ряда часто применяется для получения его структуры: разложения на простые и/или интерпретируемые составляющие, выявление периодичности, избавление от шума и т.д. В случае же набора нескольких рядов игнорирование взаимосвязей между ними может приводить к некачественным и ложным разложениям. В данной работе предлагается метод декомпозиции, учитывающий фактор связанности и основанный на классическом методе SSA (Гусеница) и тензорных разложениях. Предложенный подход сравнивается с похожим mSSA, основанным на матричных разложениях, в математических свойствах и при обработке синтетических и реальных данных (потребление электроэнергии, акселерометрия).

Введение

Выявление структуры временного ряда — краеугольная задача математического прогнозирования. Любой метод, независимо от его назначения и цели: предсказание, восстановление пропусков в данных, декомпозиция и т.д., напрямую или косвенно имеет свои предположения насчёт этой структуры. Например, модели стационарных временных рядов ([18], [10]) предполагают рассматриваемый случайный процесс стационарным в широком смысле. Применяемые методы регрессионного анализа ([11], [8], [5]) рассматривают время в качестве независимых переменных (регрессов) и вводят параметрическую модель $f(\mathbf{\Theta}, t)$, отражающую структуру ряда. Также часто применяются методы на основе динамических систем ([15], [2], [17]) , предполагающие существование таковой и порождающей наблюдаемый временной ряд напрямую или косвенно. Один из подходов такого типа — **SSA**, широко применяется в задачах прогнозирования и разложения на составляющие сигналы.

Классическая постановка задачи декомпозиции временного ряда состоит в его разложении на сезонную, циклическую, тренд и шумовую составляющие. Циклическую и трендовую часть часто объединяют в одну компоненту. Также разложение может быть аддитивным (исходный ряд — сумма компонент) или мультипликативным (исходный ряд — произведение компонент). Существует большое количество методов разложения на основе техники скользящего среднего MA ([5], [4], [3]), подбирая сначала авторегрессионную трендовую компоненту, затем усреднением остатка ряда получается сезонная компонента. Данный класс приёмов активно используется в эконометрике и финансах. Также возможно разложение сигнала в ряд фурье по гармоникам или полиномам, и связанные с ним более продвинутые методы (вейвлет преобразование и др.). Данное разложение удобно для амплитудно-частотного анализа сигнала и его фильтрации, но оно имеет недостаток фиксированного базиса функций, по которой раскладывается ряд. Метод SSA же позволяет

[1]semkin.ki@phystech.edu

[2]vadim.swifton@gmail.com

получить адаптивный базис разложения, связанный с пространством скрытой динамической системы, порождающей наблюдаемый сигнал. Если полученные базисные векторы возможно разделить на группы, отвечающие своему порождаемому сигналу, то мы получаем желанную декомпозицию исходного ряда.

В случае же набора временных рядов, порождённых одной системой и часто одновременно измеренных (например, показания акселерометра и гироскопа; количество реагентов типа A и B при протекании их реакции и т.д.), их разложение порознь обычными методами скорей всего ни отразит их общей структуры, ни даст качественных, интерпретируемых компонент. Для решения данной проблемы существует простая модификация 'Гусеницы' - **mSSA**, описание которой будет приведено ниже. Она успешно применяется на практике: например, изучение фазовой синхронизации [9] или восстановление пропусков в временных рядов [1].

В данной работе вводится ещё одна модификация: **tSSA**, в которой предлагается использовать мощь тензорных разложений вместо матричных. В большинстве своём для них нет простых и понятных алгоритмов, но существуют достаточно быстрые приближённые, что в случае больших выборок будет вычислительно эквивалентно матричным. С другой стороны, представление временных рядов в виде тензора и его декомпозиция потенциально более выразительна и способна выявить более тонкие взаимодействия между рядами. Применение и исследование данного метода можно найти в [12], [13], где авторы сворачивают одномерный сигнал ЭКГ мозговой активности в трёхмерный тензор и получают его разложение по IMFs (intrinsic mode functions) определённых частот. В работе [19] авторы применяют ту же технику но для анализа сигнала отклика механических систем, а также предлагают сведение задачи тензорного разложения к задаче выпуклой оптимизации. В работах [6] и [7] исследователи работают с гиперспектральными изображениями, которые сами по своей природе представляются в виде многомерных массивов, и тензорное разложение позволяет учесть не только взаимодействие в цветовых каналах изображения, но и пространственно-спектральное. Метод успешно применяется для сегментации изображений земной поверхности.

Далее будет дано математическое описание алгоритмов mSSA и tSSA, их особенности и сравнение в подходах к декомпозиции многомерных рядов. После данные методы тестируются на разложении набора синтетических рядов (гармонические функции с разными частотами), а также на реальных данных: потребление электроэнергии частным домохозяйством в течении года и данными акселерометра движущегося пешехода. Анализируется ошибка, возникающая при разложении рядов, интерпретируемость разложения, качество предсказания на отложенных значениях выборки.

Теоретическая часть

Пусть имеем набор временных рядов $\{x_i(t_j)\}_{i=1}^m$, где сетка по времени $t_j \in \overrightarrow{1, N}$. Стоит задача разложения временных рядов на аддитивные компоненты:

$$\begin{gathered}
x_1(t) = f_1(t) + f_2(t) + \ldots + f_{n_1}(t) \\
x_2(t) = g_1(t) + g_2(t) + \ldots + g_{n_2}(t) \\
\ldots
\end{gathered}$$

Формулировка задачи слишком общая, ведь таких разложений для произвольной функции можно построить сколь угодно много и каким угодно способом: можно раскладывать по индикаторным функциям (любую измеримую функцию можно так разложить), можно в ряды Фурье, и т.д. Мы же будем полагаться на свойства вводимых далее методов с некоторыми предположениями для поиска этого разложения.

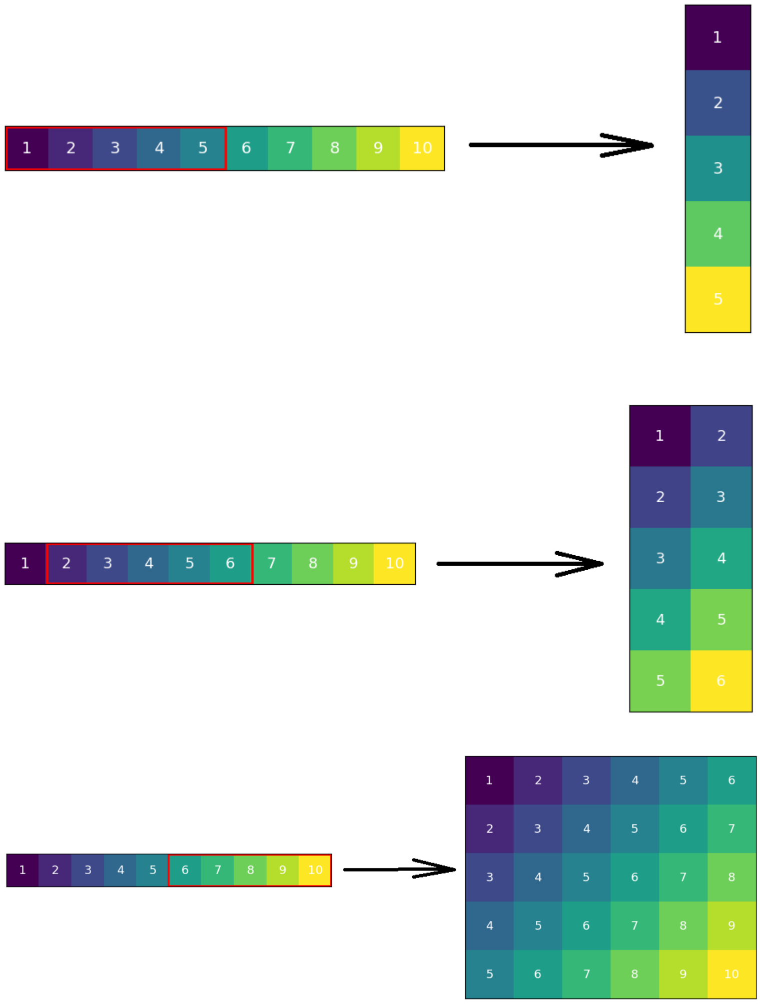

Рис. 1: Построение ганкелевой матрицы

Метод SSA

Пусть имеем пока один времнной ряд. Подход 'Гусеницы' предполагает наличие некоторой динамической системы, порождающей наблюдаемый ряд, и опирается на теорему Такенса об аппроксимации многообразия, в котором лежат траектории системы, *векторами временных задержек* ([16]). Для построения этих векторов выбирается длина окна L, далее это окно 'прикладывается' к разным частям временного ряда (см. рис.1). Формально k-ый вектор задержек есть $\mathbf{x}_k = (x(t_k)\ x(t_{k+1})\ \ldots\ x(t_{k+L-1}))$. Далее данные вектора собираются в матрицу T_{ij} по столбцам, у которой в результате все антидиагональные элементы $i+j = const$ равны. Такие матрицы называются *ганкелевой*, а в терминах временных рядов матрица T называется *траекторной* или матрицей задержек.

Далее к матрице применяется SVD-разложение $T = \sum^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i$ (знак транспонирования для $\mathbf{v}_i$ опускается), получая таким образом тот самый адаптивный ортонормированный базис $\{\mathbf{u}_i\}_{i=1}^r$ в пространстве векторов задержек. Компоненты с малыми σ_i интерпретируются как шумовые.

Для декомпозиции исходного ряда предлагается следующее: пусть $x(t) = f_1(t) + f_2(t)$ и пусть мы построили ганкелевы матрицы X_1, X_2 для $f_1(t)$ и $f_2(t)$, тогда $T = X_1 + X_2$. Как известно, SVD-разложение матриц единственно (разложение по ортонормированному базису), т.е. для матриц X_1 и X_2 это разложение по тем же базисам $\{\mathbf{u}_i\}_{i=1}^r$, $\{\mathbf{v}_i\}_{i=1}^r$. Главное предположение здесь — возможность разделить эти базисы на две непересекающиеся группы: $(\{\mathbf{u}_i^1\}_{i=1}^{r_1}, \{\mathbf{v}_i^1\}_{i=1}^{r_1})$ и $(\{\mathbf{u}_i^2\}_{i=1}^{r_2}, \{\mathbf{v}_i^2\}_{i=1}^{r_2})$, где $r_1 + r_2 = r$. Тогда возможно вычленить декомпозицию исходного ряда из его сингулярного разложения:

$$T = \sum_{i}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i = \sum_{i}^{r_1} \sigma_i^1 \mathbf{u}_i^1 \mathbf{v}_i^1 + \sum_{i}^{r_2} \sigma_i^2 \mathbf{u}_i^2 \mathbf{v}_i^2 = X_1 + X_2$$

Получив X_1, X_2 можно однозначно восстановить $\{f_1(t_j)\}, \{f_2(t_j)\}$, для этого достаточно взять первый столбец и последнюю строки в этих матрицах и соединить в один вектор друг за другом — это мгновенно следует из способа построения матриц векторов задержек и их ганкелевой структуры. Таким же образом можно разделять спектральное разложение на большее число групп и по той же процедуре извлекать составляющие сигнала, если гипотеза возможности такого разделения верна.

Вводимое предположение имеет следующий смысл: т.к. вектора задержек для $f_1(t)$ и $f_2(t)$ лежат в ортогональных друг другу пространствах, порождённых $\{\mathbf{u}_i^1\}_{i=1}^{r_1}$ и $\{\mathbf{u}_i^2\}_{i=1}^{r_2}$, то и исходные вектора задержек двух рядов ортогональны друг другу:

$$f_1(t_i)f_2(t_i) + \ldots + f_1(t_{i+L-1})f_2(t_{i+L-1}) =< f_{1,L}, f_{2,L} >= 0,\ \forall i \in \overrightarrow{1, N-L+1}$$

С точки зрения динамических систем, между векторами задержек и многообразием траекторий скрытой системы существует диффеоморфизм, таким образом существование предполагаемого разделение говорит о разложимости скрытого многообразия в прямую сумму подпространств.

Как показано в подробной книге о методе SSA [15], многие простые типы функций невозможно строго разделить таким образом. Тем не менее, в асимптотическом приближении $N \to \infty, L \to \infty$ они разделяются, т.е. их корреляция в этом случае $< f_{1,L}, f_{2,L} >\to 0$, см. табл.1. Также там показано, что если $\exists \tilde{L} : \forall L > \tilde{L} \hookrightarrow rank(T) = r = const$, то исходный ряд в точности описывается авторегрессионной моделью.

Таблица 1: Асимптотическая разделимость компонент (Golyandina)

	const	cos	exp	exp (cos)	ak + b
const	-	+	+	+	-
cos	+	+	+	+	+
exp	+	+	+	+	+
exp (cos)	+	+	+	+	+
ak + b	-	+	+	+	-

Т.к. любые данные содержат в себе ошибки (погрешности измерений, неточности вычислений и т.д.), даже идеально разделяемые ряды на практике не получится чётко разделить по спектральному разложению. Поэтому применяется простая процедура: полученные в ходе разделения главных компонент X_1, X_2 *ганкелизуются*, т.е. каждая антидиагональ матриц усредняется и заменяется на это среднее.

вставить оценку ошибки такого приближения

Таким образом оценивать полученную декомпозицию можно по получившейся невязке. Также часто для разделения главных компонент используют *относительную близость спектральных чисел*, т.е. по сути происходит их кластеризация. В той же [15] есть некоторые обоснования этого для нескольких классов функций.

Метод mSSA

Пусть теперь имеем m временных рядов $\{x_i(t_j)\}_{i=1}^m$, где $t_j \in \overrightarrow{1, N}$. Предполагая ту же гипотезу о природе порождения этих сигналов, будет действовать похожим на SSA способом. Для этого опять выбираем длину окна L и строим траекторные матрицы для каждого ряда $T_1, T_2, \ldots, T_m$. Теперь конкатенируем все матрицы в одну $T = [T_1\ T_2 \ldots T_m]$, которую далее раскладываем с помощью SVD:

$$T = \sum_i^r \sigma_i \mathbf{u}_i \mathbf{v}_i \Leftrightarrow \begin{cases} T_1 = \sum\limits_i^r \sigma_i \mathbf{u}_i \mathbf{v}_i^1 \\ T_2 = \sum\limits_i^r \sigma_i \mathbf{u}_i \mathbf{v}_i^2 \\ \dots \\ T_m = \sum\limits_i^r \sigma_i \mathbf{u}_i \mathbf{v}_i^m \end{cases}$$

Здесь каждая главная компонента в пространстве строк $\mathbf{v}_i$ представляется в виде конкатенации: $\mathbf{v}_i = (\mathbf{v}_i^1 \dots \mathbf{v}_i^m)$. Т.о. получаем некоторое разложение траекторных матриц, которое не является SVD-разложением в общем случае (т.к. компоненты $\mathbf{v}_i^k$ не обязаны быть ортогональны между собой). Тем не менее базис столбцов у всех сигналов одинаковый $\{\mathbf{u}_i\}_{i=1}^r$ и ортонормированный, 'сингулярные' числа тоже одинаковые. Таким образом общее столбцовое пространство связывает траекторные матрицы каждого ряда и учитывает их общую природу порождения.

Дальнейший ход действий для декомпозиции аналогичен методу SSA. Можно рассматривать разложение каждой траекторной матрицы отдельно, но т.к. все они имеют общий набор 'сингулярных' чисел, и группировка главных компонент часто делается на их основе, то разделение на группы происходит одинаковое для всех траекторных матриц. Получаем итоговое разложение каждого сигнала в виде:

$$\begin{gathered} x_1(t) = \hat{f}_1(t) + \hat{f}_2(t) + \ldots + \hat{f}_n(t) \\ x_2(t) = \hat{g}_1(t) + \hat{g}_2(t) + \ldots + \hat{g}_n(t) \\ \dots \\ x_m(t) = \hat{h}_1(t) + \hat{h}_2(t) + \ldots + \hat{h}_n(t) \end{gathered}$$

Метод tSSA

Данный подход основан на сборке того же набора рядов $\{x_i(t_j)\}_{i=1}^m$ в тензор и применения уже *тензорного* разложения вместо рассматриваемого до этого SVD. Сразу отметим, что способов упаковки и разложений для тензоров можно предложить огромное количество (для базового ознакомления можно, например, обратиться к [14]; также см. предлагаемые работы в Введение), являясь ещё скудно изученной областью в приложении к временным рядам.

В данной работе предлагается вдохновлённый классическим SSA способ получения *траекторного тензора* $\mathbf{T}$: как и в mSSA, строятся траекторные матрицы для каждого ряда $T_1, \dots, T_m$, после чего данные матрицы состыковываются друг с другом по третьему измерению (индексу) тензора. Т.о. измерениям $\mathbf{T}$ соответствуют столбцы и строки временных задержек, а также номер сигнала. Далее, к полученному тензору применяется *каноническое разложение* (Canonical polyadic decomposition, CPD), как наиболее схожее с SVD разложение (см. рис.2).

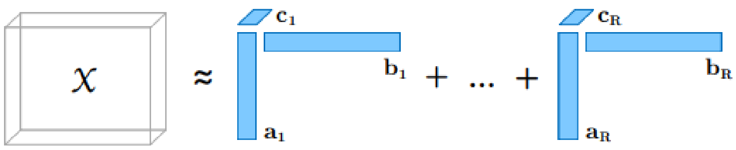

Рис. 2: Иллюстрация канонического разложения

С математической точки зрения это выглядит так:

$$\mathbf{T} = \sum_i^r \mathbf{a}_i \otimes \mathbf{b}_i \otimes \mathbf{c}_i \Leftrightarrow \begin{cases} T_1 = \sum_i^r \mathbf{c}_i[1] \cdot \mathbf{a}_i \mathbf{b}_i \\ T_2 = \sum_i^r \mathbf{c}_i[2] \cdot \mathbf{a}_i \mathbf{b}_i \\ \dots \\ T_m = \sum_i^r \mathbf{c}_i[m] \cdot \mathbf{a}_i \mathbf{b}_i \end{cases}$$

Здесь записано разложение исходного тензора в сумму тензорных произведений набора векторов $\mathbf{a}_i, \mathbf{b}_i, \mathbf{c}_i$, которые можно для удобства состыковать в матрицы A, B, C. Если зафиксировать третий индекс, то получим разложение каждой траекторной матрицы T_i, где $\mathbf{c}_i[k]$ — это k-ая компонента вектора $\mathbf{c}_i$. Таким образом разложение происходит по одному базису как для столбцов, так и для строк у каждой матрицы задержек! Роль же сингулярных чисел играют строки матрицы C. Дальнейший ход действий для получения декомпозиции аналогичен SSA.

Полученное разложение обладает *рядом преимуществ* по сравнению с mSSA: теперь имеем общий базис как у столбцов, так и у строк матриц T_i, что ещё сильнее связывает каждый сигнал; тем более строки у T_i - это те же вектора задержек, только другой размерности, поэтому странно раскладывать их для каждой T_i по своему базису, если столбцы раскладываются по одному. Также данное разложение более гибкое, т.к. для каждой матрицы задержки имеется свой набор $\mathbf{c}_i[k]$, $k \in \overrightarrow{1, r}$ 'сингулярных' чисел, что позволяет каждому сигналу подстроиться под базисные векторы индивидуально. Таким образом каждую траекторную матрицу можно раскладывать также индивидуально, что было затруднительно в mSSA. Ещё одним положительным моментом является свойство CPD-разложения: при затратно проверяемых, но весьма необременительных условиях оно единственно.

Тем не менее есть и некоторые недостатки по сравнению с матричным разложением: теперь базисы в пространстве столбцов не составляют ортогональную систему (чего не скажешь о mSSA), в пространстве строк аналогично (у mSSA та же проблема). Поэтому имеет смысл объединять слишком скоррелированые векторы в этих базисах. Также их необходимо нормировать для преемственности с предыдущими методами, нормы векторов скорректируют изначальные 'сингулярные' числа (далее по умолчанию предполагаем нормированность). Ещё эти числа могут быть отрицательными, что уменьшает их интерпретируемость.

О других проблемах рассмотренных методов ещё немного в следующем разделе.

Трудности рассмотренных методов

В данном разделе хочется априори проанализировать несколько 'белых пятен' и проблем предложенных алгоритмов.

Проблема выбора группировки

Как и в обычном SSA, как и в других рассматриваемых методах, необходимо группировать полученную сумму факторов в разложении матриц T_i. Для SSA и mSSA это возможно делать по близости сингулярных чисел, хотя нет абсолютно никакой гарантии в правильности данного подхода. В tSSA из-за того, что эти числа могут быть отрицательными и более разреженными, такой метод отбора затруднителен. Качество полученной группировки можно оценить по совокупной ошибке при ганкелизации полученных факторов. Но предложить быстрый метода хорошего выбора пока затруднительно. Возможно в tSSA разложении можно разбивать полученные базисные векторы на почти ортогональные подсистемы, что будет служить знаком существования возможности группирования. В других методах столбцовый базис уже составляет ортогональную систему.

Сложность вычислений

Предложенные методы работают с разложением матриц и тензоров, что при больших объёмах выборок и выбора длины окна L может привести к невозможности их использования, по крайней мере стандартных алгоритмов. Например, в табл.3 приведено время работы и требования по памяти для SVD-разложения из библиотеки *scipy*. Поэтому для размеров временных рядов $\gtrsim 10^4$ необходимо искать более продвинутые решения.

Рис. 3: Сложность SVD-разложения (scipy)

n	10^3	10^4	10^5
Время	0.1 s	0.9 min	16 h
Mem(A)	8 Mb	0.8 Gb	80 Gb

Для CPD-разложения всё сложней. Несмотря на его хорошие свойства, вычисление канонического ранга r является NP-трудной задачей. Поэтому при применении tSSA его необходимо подбирать. Из-за неточности этого подбора возможна большая ошибка аппроксимации полученным разложением исходного тензора $\mathbf{T}$, а слишком больший выбор r приводит к большому времени вычисления и потребляемой памяти. Достаточно быстрый итерационный приближённый алгоритм ALS (Alternating Least Squares) всё равно оперирует матрицами, их свёртками и нахождениями минимума матричных функций, что катастрофично для тензоров больших размеров.

Вычислительный эксперимент

Анализ ошибки

Список литературы

[1] Anish Agarwal, Abdullah Alomar и Devavrat Shah. *On Multivariate Singular Spectrum Analysis and its Variants*. 2020. arXiv: 2006.13448 [cs.LG].

[2] Ricky T. Q. Chen и др. *Neural Ordinary Differential Equations*. 2018. arXiv: 1806.07366 [cs.LG].

[3] Robert B. Cleveland и др. «STL: A Seasonal-Trend Decomposition Procedure Based on Loess (with Discussion)». В: *Journal of Official Statistics* 6 (1990), с. 3—73.

[4] Estela Dagum и Silvia Bianconcini. *Seasonal Adjustment Methods and Real Time Trend-Cycle Estimation*. Авг. 2016. ISBN: 978-3-319-31820-2.

[5] W. Enders и John Wiley & Sons. *Applied Econometric Time Series*. Wiley series in probability and statistics. Wiley, 2010. ISBN: 9788126543915. URL: https://books.google.ru/books?id=vzJOCgAAQBAJ.

[6] Hang Fu и др. «Tensor Singular Spectrum Analysis for 3-D Feature Extraction in Hyperspectral Images». В: *IEEE Transactions on Geoscience and Remote Sensing* 61 (2023), с. 1—14. DOI: 10.1109/TGRS.2023.3272669.

[7] Hang Fu и др. «Three-dimensional singular spectrum analysis for precise land cover classification from UAV-borne hyperspectral benchmark datasets». В: *ISPRS Journal of Photogrammetry and Remote Sensing* 203 (2023), с. 115—134. ISSN: 0924-2716. DOI: https://doi.org/10.1016/j.isprsjprs.2023.07.013. URL: https://www.sciencedirect.com/science/article/pii/S0924271623001946.

[8] William H. Greene. *Econometric Analysis*. Fifth. Pearson Education, 2003. ISBN: 0-13-066189-9. URL: http://pages.stern.nyu.edu/~wgreene/Text/econometricanalysis.htm.

[9] Andreas Groth и Michael Ghil. «Multivariate singular spectrum analysis and the road to phase synchronization». В: *Phys. Rev. E* 84 (3 сент. 2011), с. 036206. DOI: 10.1103/PhysRevE.84.036206. URL: https://link.aps.org/doi/10.1103/PhysRevE.84.036206.

[10] J.D. Hamilton. *Time Series Analysis*. Book collections on Project MUSE т. 10. Princeton University Press, 1994. ISBN: 9780691042893. URL: https://books.google.ru/books?id=B8_1UBmqVUoC.

[11] Robin John Hyndman и George Athanasopoulos. *Forecasting: Principles and Practice*. English. 2nd. Australia: OTexts, 2018.

[12] Samaneh Kouchaki и Saeid Sanei. «Tensor based singular spectrum analysis for nonstationary source separation». В: *2013 IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*. 2013, с. 1—5. DOI: 10.1109/MLSP.2013.6661921.

[13] Samaneh Kouchaki и др. «Tensor Based Singular Spectrum Analysis for Automatic Scoring of Sleep EEG». В: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 23.1 (2015), с. 1—9. DOI: 10.1109/TNSRE.2014.2329557.

[14] Stephan Rabanser, Oleksandr Shchur и Stephan Günnemann. *Introduction to Tensor Decompositions and their Applications in Machine Learning*. 2017. arXiv: 1711.10781 [stat.ML].

[15] D. Stepanov и N. Golyandina. «SSA-based approaches to analysis and forecast of multidimensional time series». не определен. В: *Proceedings of the 5th St.Petersburg Workshop on Simulation*. 2005, с. 293—298.

[16] Floris Takens. «Detecting Strange Attractors in Turbulence». В: *Dynamical Systems and Turbulence, Warwick 1980*. Под ред. David Rand и Lai-Sang Young. Т. 898. Lecture Notes in Mathematics. Berlin: Springer, 1981. Гл. 21, с. 366—381. ISBN: 978-3-540-11171-9. DOI: 10.1007/bfb0091924. URL: http://dx.doi.org/10.1007/bfb0091924.

[17] Anastasios A. Tsonis и др. «Convergent Cross Mapping: Theory and an Example». В: *Advances in Nonlinear Geosciences*. Под ред. Anastasios A. Tsonis. Cham: Springer International Publishing, 2018, с. 587—600. ISBN: 978-3-319-58895-7. DOI: 10.1007/978-3-319-58895-7_27. URL: https://doi.org/10.1007/978-3-319-58895-7_27.

[18] Granville Tunnicliffe Wilson. «Time Series Analysis: Forecasting and Control,5th Edition, by George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel and Greta M. Ljung, 2015. Published by John Wiley and Sons Inc., Hoboken, New Jersey, pp. 712. ISBN: 978-1-118-67502-1». В: *Journal of Time Series Analysis* 37 (март 2016), n/a—n/a. DOI: 10.1111/jtsa.12194.

[19] Dan Yang и др. «Improved Tensor-Based Singular Spectrum Analysis Based on Single Channel Blind Source Separation Algorithm and Its Application to Fault Diagnosis». В: *Applied Sciences* 7.4 (2017). ISSN: 2076-3417. DOI: 10.3390/app7040418. URL: https://www.mdpi.com/2076-3417/7/4/418.