

УДК 333.555.777.99

# ИТЕРАТИВНЫЕ АЛГОРИТМЫ ВЗВЕШЕННОЙ АППРОКСИМАЦИИ РЯДАМИ КОНЕЧНОГО РАНГА

Н.Э. Голяндина

*Санкт-Петербургский государственный университет,  
Математико-механический факультет*

Россия, 198504, Санкт-Петербург, Петродворец, Университетский пр., 28

E-mail: [nina@gistatgroup.com](mailto:nina@gistatgroup.com)

Н.К. Звонарев

*Санкт-Петербургский государственный университет,  
Математико-механический факультет*

Россия, 198504, Санкт-Петербург, Петродворец, Университетский пр., 28

E-mail: [mb\\_93@mail.ru](mailto:mb_93@mail.ru)

**Ключевые слова:** Временные ряды, алгоритм Cadzow, косоугольное SVD-разложение, SSA

В работе рассматриваются вопросы аппроксимации временного ряда рядами конечного ранга. Эта задача возникает в задачах обработки сигналов, в частности, при анализе зашумленных сигналов. В работе рассматривается постановка задачи аппроксимации, т.е. нахождения ряда конечного ранга, ближайшего к исходному. Возникает оптимизационная задача, в которой целевая функция, судя по всему, имеет много локальных минимумов. Один из методов локального поиска (итерации Cadzow) уже хорошо известен, и описан в данной работе. Целевая функция имеет вид взвешенного евклидова расстояния, однако, итерации Cadzow могут работать только с весами специфичного вида, в то время как зачастую требуются именно единичные веса, порождающие обычную евклидову метрику, поэтому были построены и рассмотрены несколько новых методов, подход которых состоит в том, чтобы добиться единичных или близких к ним весов. Для всех методов было произведено сравнение на модельных примерах, из которого видно преимущество новых алгоритмов в задаче оценки зашумленного сигнала, по сравнению с обычными итерациями Cadzow.

## 1. Введение

Рассмотрим задачу выделения сигнала  $\mathbb{Y} = (y_1, \dots, y_N)$  из наблюдаемого зашумлённого сигнала  $\mathbb{X} = \mathbb{Y} + \mathbb{N}$ , где  $\mathbb{Y}$  обладает некоторой заданной структурой, а именно,  $\mathbb{Y}$  управляется некоторой *линейной рекуррентной формулой* (ЛРФ) порядка  $r$ :

$$y_n = \sum_{i=1}^r a_i y_{n-i}, \quad n = r+1, \dots, N.$$

Вообще говоря, ряды, управляемые ЛРФ, могут быть записаны в параметрической форме в виде  $y_n = \sum_i P_i(n) \exp(\alpha_i n) \cos(2\pi\omega_i n + \psi_i)$ . Однако, параметрический подход к задаче не приводит к хорошим оценкам, так как параметров много, и их оценки неустойчивы.

Хорошо зарекомендовали себя так называемые *subspace-based* методы. Идея таких методов следующая: зафиксируем длину окна  $L$ ,  $1 \leq L \leq N$ ,  $K = N - L + 1$ , и построим по ряду  $\mathbf{Y}$  траекторную матрицу

$$\mathbf{Y} = \begin{pmatrix} y_1 & y_2 & \dots & y_K \\ y_2 & y_3 & \dots & y_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & \dots & y_N \end{pmatrix}.$$

Если ряд управляется минимальной ЛРФ порядка  $r$ ,  $r < \min(L, K)$ , то  $\text{rank } \mathbf{Y} = r < L$ . Таким образом,  $\mathbf{Y}$  — ганкелева матрица ранга  $r$ .

Пусть  $\mathbf{X}$  — траекторная матрица ряда  $\mathbf{X}$ . Тогда задачу оценивания  $\mathbf{Y}$  можно рассматривать как задачу аппроксимации матрицы  $\mathbf{X}$  ганкелевой матрицей ранга  $\leq r$ :

$$(1) \quad \|\mathbf{Y} - \mathbf{X}\|_F^2 \rightarrow \min_{\substack{\text{rank } \mathbf{Y} \leq r \\ \mathbf{Y} \in \mathcal{H}}}$$

Этой задаче посвящено много работ, например, [1], [2]. Методы решения — итеративные, например, метод Cadzow состоит из альтернативных проекций на множество ганкелевых матриц и матриц ранга  $\leq r$ . Хотя целевая функция не унимодальная, и сходимость к глобальному минимуму не гарантируется, тем не менее, задача (1) считается достаточно хорошо исследованной.

Заметим, что задача (1) эквивалентна задаче

$$(2) \quad \sum_{i=1}^n w_i (x_i - y_i)^2 \rightarrow \min_{\substack{\mathbf{Y}: \text{rank } \mathbf{Y} \leq r \\ \mathbf{Y} \in \mathcal{H}}},$$

$$\text{где } w_i = \begin{cases} i & \text{для } i = 1, \dots, L-1, \\ L & \text{для } i = L, \dots, K, \\ N - i + 1 & \text{для } i = K+1, \dots, N. \end{cases}.$$

На краях ряда вес меньше, чем в середине, то есть (2) не является задачей стандартного (не взвешенного) МНК для ряда (заметим, что чем меньше  $L$ , тем ближе веса к равномерным).

Целью данной работы было рассмотреть методы, решающие задачу (2) с равными весами  $w_i$ , и сравнить результаты с точки зрения точности оценивания сигнала  $\mathbf{Y}$ . Все рассматриваемые методы являются итерационными. Если интерес представляет оценка сигнала, которая не обязательно управляется ЛРФ, то в качестве оценки сигнала можно брать первую итерацию с целью уменьшения трудоёмкости. Таким образом, рассматриваемые методы сравнивались по точности оценки сигнала на первой итерации и в пределе. Заметим, что известный метод SSA можно рассматривать как одну итерацию метода Cadzow.

Рассматриваются следующие методы:

1. Cadzow — альтернативные проекции.
2. Weighted Cadzow — метод со взвешенным SVD с весами  $1/w_i$  на побочных диагоналях. На каждой итерации для взвешенного SVD приходится использовать итерационный алгоритм, что сильно увеличивает его трудоёмкость.
3. Extended Cadzow — метод, в котором ряд расширяется так, чтобы веса крайних точек формально стали такими же, как и у центральных. Добавленные точки считаются пропусками и для их заполнения на каждой итерации используется итерационный ЕМ-алгоритм, то есть трудоёмкость также большая.
4. Метод сведения к обобщенному SVD когда  $N$  кратно  $L$ . В пространстве строк траекторной матрицы вводится  $\mathbf{S}$ -скалярное произведение с

$$\mathbf{S} = \text{diag}(1, \alpha, \alpha, \dots, \alpha, 1, \alpha, \dots, 1),$$

где единицы стоят на  $1, L+1, 2L+1, \dots, K$  месте и аппроксимация в (1) идет по соответствующей норме Фробениуса. Для  $\alpha = 0$  задача (1) сводится к (2) с единичными весами, однако минимум берётся по произвольным (не только ганкелевым) матрицам ранга  $\leq r$ , то есть решается не задача аппроксимации рядами, управляемыми ЛРФ. Для  $\alpha \approx 0$  ( $\alpha > 0$ ) веса получаются близкими к одинаковыми. Заметим, что метод 1 (метод Cadzow) является частным случаем с  $\alpha = 1$ . Будем называть этот метод Cadzow ( $\alpha$ ).

5. В методе Nat-Cadzow идея метода 4 модифицируется: вводится иное скалярное произведение в пространстве строк, веса которого получены аппроксимацией весов, используемых в методе 2. В остальном методы совпадают.

## 2. Постановка задачи

### 2.1. Задача о приближении временного ряда рядом конечного $L$ -ранга

Рассмотрим временной ряд  $\mathbb{X} = (x_1, \dots, x_N)$  длины  $N \geq 3$ . Зафиксируем длину окна  $L$  ( $1 < L < N$ ),  $K = N - L + 1$ . Также рассмотрим последовательность векторов:

$$(3) \quad X_i = (x_i, \dots, x_{i+L-1})^T, \quad i = 1, \dots, K.$$

Траекторным пространством ряда  $\mathbb{X}$  назовём

$$\mathcal{X}^{(L)}(\mathbb{X}) = \mathcal{X}^{(L)} = \text{span}(X_1, \dots, X_{N-L+1}).$$

**Определение 1.** Пусть  $0 \leq r \leq L$ . Будем говорить, что ряд  $\mathbb{X}$  имеет  $L$ -ранг  $r$ , если  $\dim \mathcal{X}^{(L)} = r$ .

Заметим, что ряд  $\mathbb{X}$  может иметь  $L$ -ранг  $r$  только тогда, когда

$$(4) \quad r \leq \min(L, N - L + 1).$$

Скажем, что при фиксированном  $r$  длина окна  $L$  является *допустимой*, если для неё выполнено условие (4).

В дальнейшем будет предполагаться, что  $L$  не превосходит  $K$ , так как транспонирование не изменит ситуацию, а строчный ранг матрицы равен её столбцовому рангу.

Пусть  $\mathbf{X}_N$  — множество всех временных рядов длины  $N$ ,  $\mathbf{X}_N^r$  — множество всех временных рядов длины  $N$   $L$ -ранга  $r$ . Для заданных  $\mathbb{X} \in \mathbf{X}_N$  — исходный временной ряд,  $1 < L < N$  — длина окна и  $r$ , удовлетворяющего условию (4), рассмотрим задачу:

$$(5) \quad f_q(\mathbb{Y}) \rightarrow \min_{\mathbb{Y} \in \mathbf{X}_N^r}, \quad f_q(\mathbb{Y}) = \sum_{i=1}^N q_i (x_i - y_i)^2,$$

где  $f_q(\mathbb{Y}) = \sum_{i=1}^N q_i (x_i - y_i)^2$ ,  $y_i$  —  $i$ -е измерение ряда  $\mathbb{Y}$ , а  $q_1, \dots, q_N$ ,  $q_i \geq 0$ ,  $i = 1, \dots, N$  — некоторые неотрицательные веса. В частном случае, рассматривается целевая функция  $f(\mathbb{Y}) = \rho^2(\mathbb{X}, \mathbb{Y})$  — квадрат евклидова расстояния в  $\mathbf{R}^N$ . Она совпадает с  $f_q(\mathbb{Y})$  при  $q_i = 1$ ,  $i = 1, \dots, N$ .

## 2.2. Отображение на множество ганкелевых матриц

Пусть  $\mathbb{X}$  — временной ряд длины  $N$ , а  $\mathbf{X} \in \mathcal{H}$  — матрица, где  $\mathcal{H} = \mathcal{H}^{L \times K}$  — множество всех ганкелевых матриц размера  $L \times K$ . Тогда между  $\mathbf{X}_N$  — множеством всех временных рядов длины  $N$  и  $\mathcal{H}$  можно построить отображение  $\mathcal{T}$ , действующее по правилу

$$\mathcal{T}(\mathbb{X}) = \mathbf{X} : \hat{x}_{l,k} = x_{l+k-1}, \quad \mathbf{X} = (\hat{x}_{l,k}), \quad \mathbb{X} = (x_1, \dots, x_N).$$

Нетрудно заметить, что это отображение является биективным.

## 2.3. Эквивалентные целевые функции задачи (5)

Так как есть взаимно-однозначное соответствие между пространством рядов и ганкелевыми матрицами, задачу (5) можно записать на матричном языке.

В пространстве рядов целевая функция явным образом задаётся через скалярное произведение

$$(6) \quad \langle \mathbb{X}, \mathbb{Y} \rangle_q = \sum_{i=1}^N q_i x_i y_i,$$

где  $q_i$  — положительные веса.

Рассмотрим два скалярных произведения в пространстве матриц, являющихся расширениями обычного фробениусова скалярного произведения.

Введём

$$(7) \quad \langle \mathbf{X}, \mathbf{Y} \rangle_{1, \mathbf{M}} = \sum_{i=1}^L \sum_{j=1}^K m_{i,j} x_{i,j} y_{i,j}.$$

для матрицы  $\mathbf{M}$  с положительными элементами и

$$(8) \quad \langle \mathbf{X}, \mathbf{Y} \rangle_{2, \mathbf{S}} = \text{tr}(\mathbf{X} \mathbf{S} \mathbf{Y}^T)$$

для положительно определённой (или неотрицательно определённой для полунормы) матрицы  $\mathbf{S}$ .

Заметим, что если матрица  $\mathbf{M}$  состоит из всех единиц, т.е.  $m_{i,j} = 1$ , и если  $\mathbf{S}$  — единичная матрица, то оба скалярных произведения совпадают с обычным фробениусовым.

**Утверждение 1.** 1. Пусть  $\mathbf{X} = \mathcal{T}(\mathbb{X})$ ,  $\mathbf{Y} = \mathcal{T}(\mathbb{Y})$ . Тогда  $\langle \mathbb{X}, \mathbb{Y} \rangle_q = \langle \mathbf{X}, \mathbf{Y} \rangle_{1, \mathbf{M}}$  тогда и только тогда, когда

$$(9) \quad q_i = \sum_{\substack{1 \leq l \leq L \\ 1 \leq k \leq K \\ l+k-1=i}} m_{l,k}.$$

2. Для диагональной матрицы  $\mathbf{S}$ ,  $\langle \mathbf{X}, \mathbf{Y} \rangle_{1, \mathbf{M}} = \langle \mathbf{X}, \mathbf{Y} \rangle_{2, \mathbf{S}}$  тогда и только тогда, когда

$$(10) \quad m_{l,k} = s_{k,k}.$$

*Доказательство.* Для доказательства первой части утверждения заметим, что

$$\langle \mathbf{X}, \mathbf{Y} \rangle_{1, \mathbf{M}} = \sum_{i=1}^L \sum_{j=1}^K m_{i,j} x_{i+j-1} y_{i+j-1},$$

Доказательство второй части следует из того, что для диагональной матрицы  $\mathbf{S}$

$$\langle \mathbf{X}, \mathbf{Y} \rangle_{2, \mathbf{S}} = \sum_{l=1}^L \sum_{k=1}^K s_{k,k} x_{l,k} y_{l,k}.$$

□

Заметим, что вторая матричная норма с диагональной матрицей  $\mathbf{S}$  является частным случаем первой. Однако, ценность записи первой нормы в виде второй состоит в том, что аппроксимация матрицами меньшего ранга по первой норме — это сложная задача при неединичных весах  $m_{i,j}$ , а аппроксимация по второй норме — естественная задача, решаемая с помощью косоугольного сингулярного разложения.

### 3. Основной алгоритм Cadzow

#### 3.1. Постановка задачи для основного алгоритма

Рассмотрим задачу (5) со следующим весами  $w_1, \dots, w_N$ ,  $q_i = w_i$ :

$$(11) \quad w_i = \begin{cases} i & \text{для } i = 1, \dots, L-1, \\ L & \text{для } i = L, \dots, K, \\ N-i+1 & \text{для } i = K+1, \dots, N. \end{cases}$$

Видно, что такие  $w_i$  получаются из (9) при  $m_{l,k} = 1$ ,  $1 \leq l \leq L$ ,  $1 \leq k \leq K$ .  $w_i$  — количество элементов на  $i$ -й побочной диагонали матрицы порядка  $L \times K$ .

#### 3.2. Проектор на множество ганкелевых матриц

Следующим шагом, необходимым для постройки алгоритма, является проекция из  $\mathbb{R}^{L \times K}$  на множество  $\mathcal{H}$ .

Заданной матрице  $\mathbf{Y} \in \mathbb{R}^{L \times K}$  сопоставим  $\mathbf{X} = \Pi_{\mathcal{H}}(\mathbf{Y})$ ,  $\mathbf{X} \in \mathbb{R}^{L \times K}$  следующим способом:

$$x_{l,k} = \sum_{i+j=l+k} y_{ij} / w_{l+k-1}, \quad \mathbf{X} = (x_{l,k}), \quad \mathbf{Y} = (y_{l,k}).$$

Ещё один смысл  $w_i$  — это количество различных пар целых чисел  $(l, k)$  таких, что  $1 \leq l \leq L$ ,  $1 \leq k \leq K$ ,  $l+k-1 = i$ .

По сути, данный алгоритм усредняет матрицу на её побочных диагоналях.

#### 3.3. Проектор на множество матриц меньшего ранга

Рассмотрим множество  $\mathcal{M}_r$  — подмножество  $\mathbb{R}^{L \times K}$ , состоящих из матриц ранга, не превышающих  $r$ . Введем еще один оператор  $\Pi_{\mathcal{M}_r}$  на пространстве  $\mathbb{R}^{L \times K}$  с фробениусовской нормой, который отображает матрицу в  $\mathcal{M}_r$ . Представим матрицу  $\mathbf{P}$  в виде её сингулярного разложения:

$$\mathbf{P} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T,$$

где  $\mathbf{U}$  — ортогональная матрица порядка  $L \times L$ ,  $\mathbf{\Sigma}$  — квазидиагональная матрица порядка  $L \times K$  с неотрицательными диагональными элементами, расположенными в невозрастающем порядке,  $\mathbf{V}$  — ортогональная матрица порядка  $K \times K$ . Пусть  $\mathbf{\Sigma} = (\sigma_1, \dots, \sigma_L)$  — вектор, состоящий из диагональных элементов матрицы  $\mathbf{\Sigma}$ . Тогда  $\mathbf{\Sigma}_r = (\sigma_{lk}^r)$  — следующая матрица:

$$\sigma_{ij}^r = \begin{cases} \sigma_i & \text{при } i = j, i \leq r, \\ 0 & \text{в противном случае.} \end{cases}$$

Тогда оператор  $\Pi_{\mathcal{M}_r}$  по свойствам сингулярного разложения можно представить следующим образом:

$$\Pi_{\mathcal{M}_r}(\mathbf{P}) = \mathbf{U} \mathbf{\Sigma}_r \mathbf{V}^T.$$

В результате, полученная матрица будет иметь ранг, не превышающий  $r$ . Стоит заметить, что данный оператор не является линейным, потому что множество матриц ранга, не превышающего  $r$ , не является линейным подпространством. Более того, оно не является выпуклым множеством. К тому же, результат проекции не всегда однозначен. Это можно заметить, когда получившиеся  $\sigma_i$  одинаковы.

### 3.4. Описание алгоритма

Пусть  $\mathbb{X}$  — исходный временной ряд. Сопоставим ему ганкелеву матрицу  $\mathbf{X}_0 = \mathcal{T}(\mathbb{X})$ , а затем рассмотрим последовательность матриц:

$$\mathbf{X}_k = \Pi_H(\Pi_{\mathcal{M}_r}(\mathbf{X}_{k-1})), \quad k = 1, 2, \dots$$

Данная последовательность имеет подпоследовательность, которая сходится к некоторой  $\tilde{\mathbf{X}}$  такой, что  $\tilde{\mathbb{X}} = \mathcal{T}^{-1}(\tilde{\mathbf{X}})$  является рядом конечного  $L$ -ранга  $r$  [1].

## 4. Cadzow-подобный алгоритм задачи приближения ряда рядом конечного ранга (Extended Cadzow)

### 4.1. Введение

Как уже было показано в разделе 3.1., базовый алгоритм Cadzow решает задачу с весами  $q_i$ ,  $i = 1, \dots, N$ , не являющимися единичными. Метрика, которую порождают данные веса, является естественной для матриц  $\mathbf{R}^{L \times K}$ , однако в практических задачах куда естественнее рассматривать обычную евклидову метрику на множестве временных рядов длины  $N$ . Поэтому в следующих двух секциях мы предполагаем, что все  $q_i$  равны 1, и будем строить приближенные решения такой задачи. В этом и состоит основная цель данной работы.

### 4.2. Отображение на множество траекторных псевдоматриц

Введём несколько определений. *Траекторной матрицей* ряда  $\mathbb{X}$  назовем матрицу  $\mathbf{X} = [X_1 : \dots : X_K]$ . Эта матрица имеет размер  $L \times K$ , и её ранг равен размерности траекторного пространства  $\mathcal{X}^{(L)}$ .

Рассмотрим отображение  $\mathcal{X}$  из множества пар индексов  $(i, j)$  во множество вещественных чисел  $\mathbf{R}$ . Множество, на котором определено это отображение, назовем носителем отображения  $\mathcal{X}$ , и будем обозначать как  $\text{supp}\mathcal{X}$ . Тогда отображение  $\mathcal{X}$ , которое удовлетворяет следующим свойствам

$$(12) \quad \begin{aligned} \mathcal{X}(i, j) &= x_{i+j-L}, \quad (i, j) \in \text{supp}\mathcal{X}, \\ \text{supp}\mathcal{X} &= \{(i, j) | 1 \leq i \leq L, 1 \leq i + j - L \leq N\}, \end{aligned}$$

$$\mathcal{X} = \begin{pmatrix} & x_1 & x_2 & \cdots & x_K & x_{K+1} & \cdots & x_N \\ & \ddots & x_2 & x_3 & \ddots & x_{K+1} & \ddots & \ddots \\ x_1 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & x_N \\ x_1 & x_2 & \cdots & x_L & x_{L+1} & \cdots & x_N & \end{pmatrix},$$

назовём *траекторной псевдоматрицей* ряда  $\mathbb{X}$ . Обозначим множество всех возможных траекторных псевдоматриц как  $\mathfrak{X}$ . Теперь удобно записать определение (12) в виде отображения  $\tilde{T} : \mathcal{X}_N \rightarrow \mathfrak{X}$ , такого, что  $\tilde{T}(\mathbb{X}) = \mathcal{X}$ . Это отображение является взаимно-однозначным, поэтому справедливо следующее равенство:  $\tilde{T}^{-1}(\mathcal{X}) = \mathbb{X}$ .

Ещё одно простое, но важное свойство — сужение  $\mathcal{X}|_{L \leq j \leq N}$  можно рассматривать как траекторную матрицу  $\mathbf{X}$ .

### 4.3. Проектор на множество матриц меньшего ранга

Обозначим множество матриц размера  $L \times N + L - 1$  как  $\mathbf{R}^{L \times N + L - 1}$ . Подмножество  $\mathbf{R}^{L \times N + L - 1}$ , состоящее из матриц ранга, не превосходящего  $r$ , обозначим как  $\mathcal{M}_r^{L \times N + L - 1} = \widetilde{\mathcal{M}}_r$ . Для заданной псевдоматрицы  $\mathcal{X}$  рассмотрим задачу:

$$g(\mathbf{Y}) \rightarrow \min_{\mathbf{Y} \in \widetilde{\mathcal{M}}_r},$$

где  $g(\mathbf{Y}) = \sum_{(i,j) \in \text{supp} \mathcal{X}} (x_{i,j} - y_{i,j})^2$ ,  $\mathcal{X}(i,j) = (x_{i,j})$ ,  $\mathbf{Y} = (y_{i,j})$ . Её решение соответствует оператору проектирования

$$(13) \quad \Pi_{\widetilde{\mathcal{M}}_r} : \mathfrak{X} \rightarrow \widetilde{\mathcal{M}}_r.$$

Заметим, что у нас уже есть алгоритм для решения этой задачи — это ЕМ-метод для заполнения матриц с пропусками, описанный ранее. Метод строит последовательность матриц, определённых следующим соотношением:

$$\begin{aligned} \mathbf{Y}_0 &= (y_{i,j}^0), \quad y_{i,j}^0 = \begin{cases} x_{i,j}, & (i,j) \in \text{supp} \mathcal{X} \\ \tilde{x}_{i,j}, & (i,j) \notin \text{supp} \mathcal{X} \end{cases}, \\ \mathbf{Y}_n &= \Pi_{\mathcal{M}_r}(\mathbf{Y}_0 \odot \mathbf{M} + \mathbf{Y}_{n-1} \odot (\mathbf{U} - \mathbf{M})), \\ \mathbf{U} &\in \mathbf{R}^{L \times N + L - 1}, \quad \mathbf{U} = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}, \\ \mathbf{M} &\in \mathbf{R}^{L \times N + L - 1}, \quad \mathbf{M} = (m_{i,j}), \quad m_{i,j} = \begin{cases} 1, & (i,j) \in \text{supp} \mathcal{X} \\ 0, & (i,j) \notin \text{supp} \mathcal{X} \end{cases}, \end{aligned}$$

где  $\Pi_{\mathcal{M}_r}$  — проектор на множество  $\mathcal{M}_r$  во фробениусовской норме, ранее описанный в пункте 3.3., а  $\tilde{x}_{i,j}$  — начальная оценка скрытых параметров ЕМ-алгоритма.

Он сходится, однако только к локальному минимуму, поэтому нельзя говорить, что он находит точное решение задачи.



#### 4.4. Проектор на множество траекторных псевдоматриц

Теперь построим оператор  $\Pi_{\mathfrak{X}}$ , действующий из множества  $\widetilde{\mathcal{M}}_r$  во множество  $\mathfrak{X}$ . Заданной матрице  $\mathbf{Y} \in \mathcal{M}_r$  сопоставим траекторную псевдоматрицу  $\mathcal{X} = \Pi_{\mathfrak{X}}(\mathbf{Y})$  следующим способом:

$$\mathcal{X}(i, j) = \frac{1}{L} \sum_{\substack{k+l=i+j \\ (k, l) \in \text{supp } \mathcal{X}}} y_{k, l}, \quad \mathbf{Y} = (y_{i, j}),$$

$$\text{supp } \mathcal{X} = \{(i, j) | 1 \leq i \leq L, 1 \leq i + j - L \leq N\}.$$

Получаем, что оператор выполняет усреднение на побочных диагоналях матрицы, которые имеют смысл при сужении до траекторной псевдоматрицы.

#### 4.5. Описание алгоритма

Пусть  $\mathbb{X}$  — исходный временной ряд. Сопоставим ему траекторную псевдоматрицу  $\mathcal{X}_0 = \tilde{T}(\mathbb{X})$ , а затем рассмотрим последовательность матриц:

$$\mathcal{X}_k = \Pi_{\mathfrak{X}}(\Pi_{\widetilde{\mathcal{M}}_r}(\mathcal{X}_{k-1})), \quad k = 1, 2, \dots,$$

где  $\Pi_{\widetilde{\mathcal{M}}_r}$  определён в (13). Данная последовательность имеет подпоследовательность, которая сходится к некоторой  $\tilde{\mathcal{X}}$  такой, что  $\tilde{\mathbb{X}} = \tilde{T}^{-1}(\tilde{\mathcal{X}})$  является рядом конечного  $L$ -ранга  $r$ . Это следует из свойства о том, что траекторная матрица является сужением траекторной псевдоматрицы.

Главное отличие от базового алгоритма Cadzow состоит в том, что используются несколько иные объекты, и они введены ради сохранения евклидовой метрики на пространстве рядов при решении задачи.

#### 4.6. Практические рекомендации

Один из оставшихся вопросов — начальная оценка неизвестных параметров  $(\tilde{x}_{i, j})$  при проектировании ЕМ-алгоритмом. Различные варианты рассмотрены в [5]. Хорошим вариантом начальной оценки при вычислении  $\mathcal{X}_1$  служит выборочное среднее. Использование векторного прогнозирования, описанного в [3], на  $L - 1$  шагов вперед и столько же шагов назад существенно улучшают работу алгоритма в случае применения для нестационарного ряда.

Важное улучшение, которое подсказывает сам вид алгоритма — при вычислении  $\mathcal{X}_2, \mathcal{X}_3, \dots$  в качестве начальной оценки скрытых параметров использовать их последние значения, полученные при предыдущем проектировании на матрицы меньшего ранга перед усреднением значений на диагоналях.

## 5. Ещё один Cadzow-подобный алгоритм (Weigthed Cadzow)

### 5.1. Взвешенный проектор на множество матриц неполного ранга

Пусть нам дана матрица  $\mathbf{X} \in \mathbb{R}^{L \times K}$ ,  $\mathbf{X} = (x_{l,k})$ , и  $\mathbf{M} \in \mathbb{R}^{L \times K}$ ,  $\mathbf{M} = (m_{l,k})$ , причём для каждого  $1 \leq l \leq L$ ,  $1 \leq k \leq K$ :  $0 < m_{l,k} \leq 1$ . Тогда рассмотрим следующую задачу:

$$g(\mathbf{Y}) \rightarrow \min_{\mathbf{Y} \in \mathcal{M}_r},$$

где  $g(\mathbf{Y}) = \sum_{l=1}^L \sum_{k=1}^K m_{l,k} (x_{l,k} - y_{l,k})^2$ ,  $\mathbf{Y} = (y_{l,k})$ . Её решение соответствует оператору проектирования

$$(14) \quad \hat{\Pi}_{\tilde{\mathcal{M}}_r} : \mathbb{R}^{L \times K} \rightarrow \tilde{\mathcal{M}}_r.$$

Для данной задачи не найдено точное решение, однако известны методы, которые дают хороший результат. Один из них описан в [5]. Идея метода состоит в построении последовательности матриц, определённой следующим соотношением:

$$\begin{aligned} \mathbf{Y}_0 &= \mathbf{X}, \quad \mathbf{Y}_n = \Pi_{\mathcal{M}_r}(\mathbf{X} \odot \mathbf{M} + \mathbf{Y}_{n-1} \odot (\mathbf{U} - \mathbf{M})), \\ \mathbf{U} &\in \mathbb{R}^{L \times K}, \quad \mathbf{U} = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}, \end{aligned}$$

где  $\Pi_{\mathcal{M}_r}$  — проектор на множество  $\mathcal{M}_r$  во фробениусовской норме, ранее описанный в пункте 3.3..

### 5.2. Описание алгоритма

Мы продолжаем строить приближенные решения задачи (5). Теперь мы будем использовать алгоритм проектирования на множество матриц неполного ранга, описанный в разделе 5.1.. Для него требуется определить матрицу  $\mathbf{M}$ .

$$(15) \quad \mathbf{M} : m_{l,k} = \frac{1}{w_{l+k-1}},$$

где  $w_i$  определены в (11). Таким образом, подставив  $m_i$  в (9), мы получим  $q_i = 1$ ,  $i = 1, \dots, N$ , следовательно,  $\mathcal{H}$  во взвешенной норме с весами  $m_{l,k}$  изоморфно  $\mathbf{X}_N$  с единичными весами.

Теперь пусть  $\mathbb{X}$  — исходный временной ряд. Сопоставим ему ганкелеву матрицу  $\mathbf{X}_0 = \mathcal{T}(\mathbb{X})$ , а затем рассмотрим последовательность матриц:

$$\mathbf{X}_k = \Pi_H(\hat{\Pi}_{\mathcal{M}_r}(\mathbf{X}_{k-1})), \quad k = 1, 2, \dots,$$

где  $\Pi_H$  определён в разделе 3.2.,  $\hat{\Pi}_{\mathcal{M}_r}$  определён в (14).

Данная последовательность имеет подпоследовательность, которая сходится к некоторой  $\tilde{\mathbf{X}}$  такой, что  $\tilde{\mathbf{X}} = \mathcal{T}^{-1}(\tilde{\mathbf{X}})$  является рядом конечного  $L$ -ранга  $r$ .

## 6. Применение подхода, связанного с использованием косоугольного SVD—разложения

В предыдущих частях работы рассматривались алгоритмы решения задачи (5), использующие ЕМ-алгоритм для проектирования на множество матриц меньшего ранга. Один из недостатков такого подхода — большая трудоемкость, чем у основного алгоритма Cadzow. В данной части будет приведен подход, рассмотренный в статье [2], указаны его недостатки и приведён другой алгоритм, показавший лучшие результаты в некоторых численных экспериментах. К тому же, в оригинальной работе [2] отсутствует взвешенное проектирование на пространство ганкелевых матриц, что при использовании алгоритма при  $\alpha = 0$  даёт совсем некорректные результаты. Мы вводим взвешенное проектирование и рассматриваю алгоритм уже с его использованием. В дальнейшем, в работе будет показано, как использование такого подхода влияет на слабую разделимость рядов и как параметр  $\alpha$  влияет на скорость сходимости алгоритма.

### 6.1. Взвешенный проектор на множество матриц неполного ранга, использующий косоугольное SVD—разложение

Пусть  $\mathbf{S} \in \mathbb{R}^{K \times K}$  — симметричная неотрицательно определённая матрица. Тогда определим полунорму  $\|\cdot\|_{\mathbf{S}}$  на пространстве  $\mathbb{R}^{K \times K}$  следующим образом:

$$(16) \quad \|\mathbf{X}\|_{\mathbf{S}}^2 = \text{tr}(\mathbf{X}\mathbf{S}\mathbf{X}^T).$$

Теперь пусть нам дана матрица  $\mathbf{X} \in \mathbb{R}^{L \times K}$ . Тогда рассмотрим следующую задачу:

$$\|\mathbf{X} - \mathbf{Y}\|_{\mathbf{S}} \rightarrow \min_{\mathbf{Y} \in \mathcal{M}_r}.$$

Для её решения проделаем следующее: сначала представим матрицу  $\mathbf{S}$  в виде  $\mathbf{S} = \mathbf{O}_{\mathbf{S}}^T \mathbf{O}_{\mathbf{S}}$ . Такое представление может быть получено с помощью вычисления корня из матрицы или разложения Холецкого. Затем вычислим матрицу  $\mathbf{B} = \mathbf{X} \mathbf{O}_{\mathbf{S}}^T$ . К полученной матрице применим оператор  $\Pi_{\mathcal{M}_r}$ , после чего в качестве ответа следует взять  $\mathbf{Y} = \Pi_{\mathcal{M}_r}(\mathbf{B})(\mathbf{O}_{\mathbf{S}}^T)^{\dagger}$ ,  $(\mathbf{O}_{\mathbf{S}}^T)^{\dagger}$  обозначает псевдообратную матрицу Мура-Пенроуза к матрице  $\mathbf{O}_{\mathbf{S}}^T$ . Полученный ответ

будет правильным, если пространство столбцов матрицы  $\mathbf{X}$  содержится в пространстве столбцов матрицы  $\mathbf{S}$  [4]. В качестве доказательства приводится следующая цепочка равенств:

$$\begin{aligned} \|\mathbf{X} - \mathbf{Y}\|_{\mathbf{S}}^2 &= \text{tr}((\mathbf{X} - \mathbf{Y})\mathbf{S}(\mathbf{X} - \mathbf{Y})^T) = \\ &= \text{tr}((\mathbf{X}\mathbf{O}_{\mathbf{S}}^T - \mathbf{Y}\mathbf{O}_{\mathbf{S}}^T)(\mathbf{X}\mathbf{O}_{\mathbf{S}}^T - \mathbf{Y}\mathbf{O}_{\mathbf{S}}^T)^T) = \\ &= \text{tr}((\mathbf{B} - \Pi_{\mathcal{M}_r}(\mathbf{B}))(\mathbf{B} - \Pi_{\mathcal{M}_r}(\mathbf{B}))^T) = \|\mathbf{B} - \Pi_{\mathcal{M}_r}(\mathbf{B})\|^2, \end{aligned}$$

где  $\|\cdot\|$  обозначает норму Фробениуса в пространстве  $\mathbb{R}^{L \times K}$ .  $\Pi_{\mathcal{M}_r}$  является проектором на  $\mathcal{M}_r$  в пространстве  $\mathbb{R}^{L \times K}$  с фробениусовской нормой, что доказывает минимальность полученного расстояния. Осталось заметить, что ранг матрицы  $\mathbf{Y}$  не превосходит  $r$ , так как ее ранг не превосходит ранг ее сомножителя  $\Pi_{\mathcal{M}_r}(\mathbf{B})$ . Таким образом, получим проектор

$$(17) \quad \Pi_{\mathcal{M}_r}^{\mathbf{S}} : \quad \mathbf{Y} = \Pi_{\mathcal{M}_r}^{\mathbf{S}}(\mathbf{X}).$$

Применённый здесь подход является частным случаем косоугольного SVD-разложения, описанного в [4].

## 6.2. Подход к решению задачи, описанный в [2]

Следующий важный шаг — найти такую матрицу  $\mathbf{S}$ , чтобы полунорма  $\|\cdot\|_{\mathbf{S}}$  соотносилась с расстоянием  $f_q(\mathbb{Y}) = f(\mathbb{Y})$ , встречающимся в (5), то есть выполнялось равенство (9) при всех  $q_i = 1$ . Воспользуемся следующей леммой:

**Лемма 1.** [2] Пусть  $\mathbb{X} \in \mathbb{X}_N$ ,  $\mathbf{X} = \mathcal{T}(\mathbb{X}) \in \mathbb{R}^{L \times K}$ . Если  $h = N/L$  — целое, тогда  $\text{tr}(\mathbf{X}\mathbf{S}\mathbf{X}^T) = \mathbb{X}^T \mathbb{X}$ , где  $\mathbf{S}$  — диагональная матрица со следующими диагональными элементами:

$$s_{k,k} = \begin{cases} 1, & \text{если } k = jL + 1 \text{ для некоторых } j = 0, \dots, h-1, \\ 0, & \text{в противном случае} \end{cases}.$$

*Доказательство.* В предположении, что  $h = N/L$  — целое, получим  $N = hL$  и  $K = (h-1)L + 1$ . По определению, элементы матрицы  $\mathbf{X}$   $x_{l,k} = \hat{x}_{l+k-1}$ , где  $(\hat{x}_k)$  — элементы ряда  $\mathbb{X}$ . Получим следующее:

$$\begin{aligned} \text{tr}(\mathbf{X}\mathbf{S}\mathbf{X}^T) &= \sum_{l=1}^L \sum_{k=1}^K \sum_{k'=1}^K x_{l,k} s_{k,k'} x_{l,k'} = \sum_{l=1}^L \sum_{k=1}^K s_{k,k} x_{l,k}^2 = \\ &= \sum_{l=1}^L \sum_{j=0}^{h-1} x_{l,jL+1}^2 = \sum_{l=1}^L \sum_{j=0}^{h-1} \hat{x}_{l+jL}^2 = \sum_{n=1}^N \hat{x}_n^2. \end{aligned}$$

□

Первая проблема, которая встречается при этом подходе — нулевые элементы на диагоналях матрицы  $\mathbf{S}$ . Таким образом, ранг матрицы  $\mathbf{S}$  заведомо меньше  $K$ . В [2] предложено заменить нули на диагоналях на некоторое малое  $\alpha$ , чтобы исправить проблему. Это исправляет проблему ранга, но, как будет видно дальше, оставляет свойства метода, решающего задачу (5) плохими.

**Следствие 1.** В предположениях леммы 1: пусть  $h = N/L$  — целое, и матрица  $\mathbf{S}$  — диагональная со следующими диагональными элементами:

$$(18) \quad s_{k,k} = \begin{cases} 1, & \text{если } k = jL + 1 \text{ для некоторых } j = 0, \dots, h-1, \\ \alpha, & \text{в противном случае,} \end{cases}$$

где  $0 \leq \alpha \leq 1$ . Тогда веса  $q_i$ , определённые в (9), выглядят следующим образом

$$q_i = \begin{cases} 1 + (i-1)\alpha & \text{для } i = 1, \dots, L-1, \\ 1 + (L-1)\alpha & \text{для } i = L, \dots, K-1, \\ 1 + (N-i)\alpha & \text{для } i = K, \dots, N. \end{cases}$$

*Доказательство.* Достаточно просуммировать  $s_{k,k}$  правильное число раз, равное размеру  $i$ -й побочной диагонали. Заметим, что при  $\alpha = 1$  полученные веса совпадают с весами (11) из постановки задачи для основного алгоритма.  $\square$

Используя соотношение (10) и взяв  $\mathbf{S}$ , описанную в лемме 1, мы получим следующую  $\mathbf{M}$ :

$$\mathbf{M} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 1 & 0 & \dots & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 & 1 & 0 & \dots & \dots & 1 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 & 1 & 0 & \dots & \dots & 1 \end{pmatrix}.$$

Таким образом, расстояние берётся только по  $h$  столбцам вместо  $K$ , а домножение на  $\mathbf{O}_S^T$  обнуляет у матрицы  $\mathbf{X}$   $K - h$  столбцов. Поэтому требование на соответствие пространств столбцов отсутствует, и полученный результат будет соответствовать решению другой задачи, в которой у траекторной матрицы убираются столбцы с повторяющимися индексами исходного ряда.

### 6.3. Нахождение новой весовой матрицы

Подойдём к задаче со стороны матриц: найдём матрицу  $\mathbf{S}$  такую, что полученная норма  $\|\cdot\|_{\mathbf{S}}$  будет похожа на расстояние в  $\mathbb{R}^{L \times K}$  с матрицей  $\mathbf{M}$ , полученной в (15). Как можно заметить в доказательстве леммы 1, нас устроит только диагональная  $\mathbf{S}$ . Рассмотрим множество  $\mathbf{Z}^{L \times K} \subset \mathbb{R}^{L \times K}$  — матрицы, у которых элементы в столбцах равны. Разумным выбором станет матрица  $\mathbf{Z} \in \mathbb{R}^{L \times K}$ ,  $\mathbf{Z} = (z_{l,k})$ ,  $z_{l,k} = s_{k,k}$  такая, что

$$\|\mathbf{M} - \mathbf{Z}\| \rightarrow \min_{\mathbf{Z} \in \mathbf{Z}^{L \times K}}.$$

Решение сводится к усреднению элементов матрицы  $\mathbf{M}$  по столбцам. В итоге, полученная матрица  $\hat{\mathbf{S}}$  будет иметь следующие диагональные элементы:

$$(19) \quad \hat{s}_{k,k} = \frac{1}{L} \sum_{l=1}^L m_{l,k}.$$

В дальнейшем, будем решать общую задачу и с весовой матрицей  $\mathbf{Z}$ , а полученный ответ будем рассматривать для решения задачи (5) с весами  $q_i = 1$ .

## 6.4. Взвешенный проектор на множество ганкелевых матриц

Главное отличие этого подхода от предыдущих — в пределах одной побочной диагонали элементы имеют разные веса. Формально, задачу можно поставить так: найти проекцию из  $\mathbb{R}^{L \times K}$  на множество  $\mathcal{H}$  в норме  $\|\cdot\|_{\mathbf{S}}$ , где  $\mathbf{S}$  определена в (19) или же в (18). Это эквивалентно следующему:

$$\sum_{l=1}^L \sum_{k=1}^K s_{k,k} (x_{l,k} - \hat{y}_{l+k-1})^2 \rightarrow \min_{\mathbb{Y} \in \mathbf{X}_N},$$

где  $\mathbf{X} = x_{l,k}$ ,  $\mathbb{Y} = (\hat{y}_n)$ , а  $\mathbf{Y} = \mathcal{T}(\mathbb{Y})$  — результат проекции. Перепишем задачу следующим образом:

$$\sum_{p=1}^N \sum_{l+k-1=p} s_{k,k} (x_{l,k} - \hat{y}_p)^2 \rightarrow \min_{\mathbb{Y} \in \mathbf{X}_N}.$$

Видно, что для каждого  $p \in 1, \dots, N$  задачу можно решить отдельно. Зафиксируем  $p$ , получим

$$\sum_{l+k-1=p} s_{k,k} (x_{l,k} - \hat{y}_p)^2 \rightarrow \min_{\hat{y}_p \in \mathbb{R}},$$

после чего для нахождения минимума возьмём производную левой части по  $\hat{y}_p$  и приравняем ее к 0. Получим:

$$\sum_{l+k-1=p} -2s_{k,k} (x_{l,k} - \hat{y}_p) = 0,$$

что эквивалентно

$$\hat{y}_p = \frac{\sum_{l+k-1=p} s_{k,k} x_{l,k}}{\sum_{l+k-1=p} s_{k,k}}.$$

Оператор проектирования обозначим оператором  $\Pi_{\mathcal{H}}^{\mathbf{S}}: \mathbf{Y} = \Pi_{\mathcal{H}}^{\mathbf{S}}(\mathbf{X})$ .

## 6.5. Описание алгоритма

Предъявим ещё один алгоритм решения задачи (5). Пусть  $\mathbb{X}$  — исходный временной ряд. Сопоставим ему ганкелеву матрицу  $\mathbf{X}_0 = \mathcal{T}(\mathbb{X})$ , а затем рассмотрим последовательность матриц:

$$\mathbf{X}_k = \Pi_{\mathcal{H}}^{\mathbf{S}}(\Pi_{\mathcal{M}_r}^{\mathbf{S}}(\mathbf{X}_{k-1})), \quad k = 1, 2, \dots,$$

где  $\Pi_{\mathcal{M}_r}^{\mathbf{S}}$  определён в (17). Данная последовательность имеет подпоследовательность, которая сходится к некоторой  $\tilde{\mathbf{X}}$  такой, что  $\tilde{\mathbf{X}} = \mathcal{T}^{-1}(\tilde{\mathbf{X}})$  является рядом конечного  $L$ -ранга  $r$ . Уже упомянутое главное преимущество данного алгоритма — его скорость работы практически равна скорости базового алгоритма Cadzow за счет быстрой проекции на матрицы неполного ранга.

## 6.6. Нахождение весов ряда

Один из оставшихся вопросов — какие итоговые веса  $\hat{q}_i$  мы получим, и насколько сильно они будут отличаться от единичных весов. Ответ дан в следующих утверждениях:

**Утверждение 1.** Матричные веса  $\hat{s}_k$ , определённые в (19), равны

$$\hat{s}_{k,k} = \begin{cases} \frac{1}{L} \left( \frac{k}{L} + \sum_{j=k}^{L-1} \frac{1}{j} \right), & k = 1, \dots, L-1, \\ \hat{s}_{N-k+1, N-k+1}, & k = K-L+2, \dots, K, \\ 1/L, & \text{в противном случае.} \end{cases}$$

*Доказательство.* Достаточно подставить в (19)  $m_{l,k}$ , определённые в (15).  $\square$

**Утверждение 2.** Пусть  $N \geq 4(L-1)$ . Тогда веса  $\hat{q}_i$ , определённые в (9), выглядят следующим образом

$$\hat{q}_i = \begin{cases} \frac{i(i+1)}{2L^2} + \frac{i}{L}(1 + H_{L-1} - H_i), & 1 \leq i \leq L-1, \\ 1 + \frac{2iL-i-i^2}{2L^2} + \frac{L-i}{L}(H_{L-1} - H_{i-L}), & L \leq i \leq 2L-1, \\ \hat{q}_{N-i+1}, & N-2L+2 \leq i \leq N, \\ 1, & \text{в противном случае,} \end{cases}$$

где  $H_0 = 0$ , а  $H_i = \sum_{j=1}^i 1/j$  —  $i$ -е гармоническое число.

*Доказательство.* Для  $1 \leq i \leq L-1$ , имеем

$$\begin{aligned} \hat{q}_i &= \sum_{j=1}^i s_j = \sum_{j=1}^i \frac{1}{L} \left( \frac{j}{L} + \sum_{k=j}^{L-1} 1/k \right) = \frac{\sum_{j=1}^i j}{L^2} + \frac{1}{L} \sum_{j=1}^i \sum_{k=j}^{L-1} 1/k = \\ &= \frac{i(i+1)}{2L^2} + \frac{1}{L} \sum_{k=1}^{L-1} \sum_{j=1}^{\min(k,i)} 1/k = \frac{i(i+1)}{2L^2} + \frac{1}{L} \sum_{k=1}^{L-1} \frac{\min(k,i)}{k} = \\ &= \frac{i(i+1)}{2L^2} + \frac{1}{L} \sum_{k=1}^i \frac{k}{k} + \frac{1}{L} \sum_{k=i+1}^{L-1} \frac{i}{k} = \frac{i(i+1)}{2L^2} + \frac{i}{L}(1 + H_{L-1} - H_i). \end{aligned}$$

Для  $L \leq i \leq 2L-1$ , получаем

$$\begin{aligned} \hat{q}_i &= \sum_{j=1}^L S_{i-L+j} = \sum_{j=i-L+1}^{L-1} s_j + \frac{i-L+1}{L} = \\ &= \frac{i-L+1}{L} + \frac{1}{L^2} \sum_{j=i-L+1}^{L-1} j + \frac{1}{L} \sum_{j=i-L+1}^{L-1} \sum_{k=j}^{L-1} 1/k = \\ &= \frac{i-L+1}{L} + \frac{2iL-i-i^2}{2L^2} + \frac{1}{L} \sum_{k=i-L+1}^{L-1} \sum_{j=i-L+1}^k 1/k = \\ &= \frac{i-L+1}{L} + \frac{2iL-i-i^2}{2L^2} + \frac{1}{L} (2L-i-1 + (L-i) \sum_{k=i-L+1}^{L-1} 1/k) = \\ &= \frac{i-L+1+2L-i-1}{L} + \frac{2iL-i-i^2}{2L^2} + \frac{L-i}{L} (H_{L-1} - H_{i-L}) = \\ &= 1 + \frac{2iL-i-i^2}{2L^2} + \frac{L-i}{L} (H_{L-1} - H_{i-L}). \end{aligned}$$

Для правой части ряда веса будут симметричными, а для центральной необходимо  $L$  раз просуммировать  $s_k = 1/L$ .  $\square$

Отнормированные веса  $q_i$  при  $\alpha = 1$  (стандартная постановка),  $\alpha = 0$  (единичные веса),  $\alpha = 0.1$  и  $\hat{q}_i$  при  $N = 40$ ,  $L = 8$  представлены на рисунке 1.

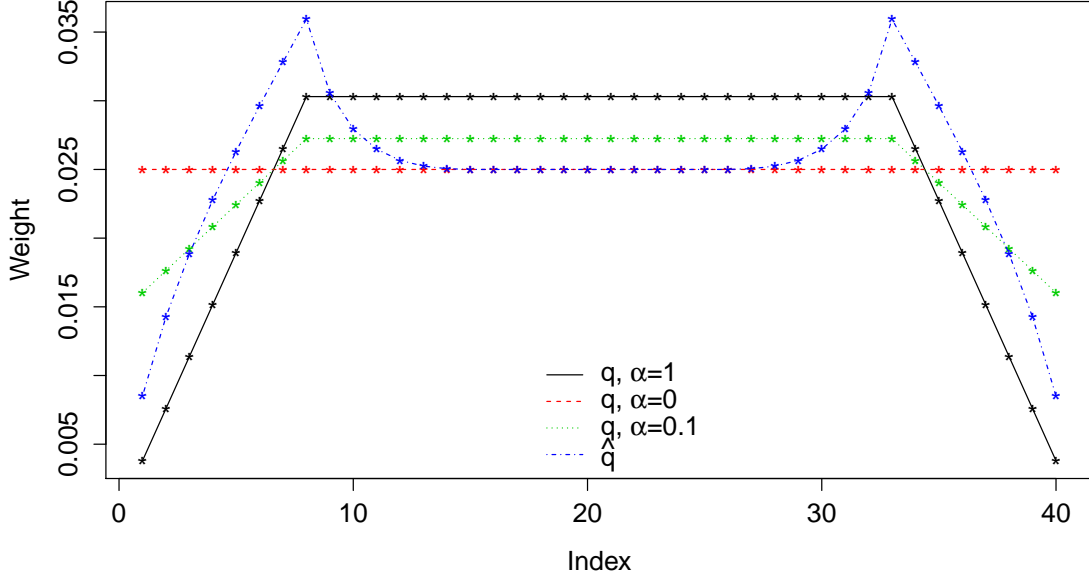


Рис. 1. Веса ряда

## 6.7. Слабая разделимость

Введём ещё одну характеристику алгоритмов, показывающую, насколько хорошо они раскладывают временной ряд на его аддитивные компоненты. Основным применением описанных алгоритмов является задача оценки сигнала, поэтому данное качество нам важно для получения как можно более точной оценки.

Пусть  $\mathbf{S} \in \mathbb{R}^{K \times K}$  — симметричная неотрицательно определённая матрица,  $\mathbb{X}_1$  и  $\mathbb{X}_2$  — два разных временных ряда длины  $N$ ,  $\mathbf{X}^1$ ,  $\mathbf{X}^2$  — их траекторные матрицы. Тогда коэффициентом корреляции  $i$ -го и  $j$ -го столбца назовём следующую величину:

$$(20) \quad \rho_{i,j}^c = \frac{(X_i^1, X_j^2)}{\|X_i^1\| \|X_j^2\|},$$

где  $X_i^k$  —  $i$ -й столбец матрицы  $\mathbf{X}^k$ ,  $k = 1, 2$ ,  $(\cdot, \cdot)$  — скалярное произведение,  $\|\cdot\|$  — евклидова норма. Коэффициентом корреляции  $i$ -й и  $j$ -й строки назовём следующую величину:

$$(21) \quad \rho_{i,j}^r = \frac{(X^{1,i}, X^{2,j})_{\mathbf{S}}}{\|X^{1,i}\|_{\mathbf{S}} \|X^{2,j}\|_{\mathbf{S}}},$$



где  $X^{k,i}$  —  $i$ -я строчка матрицы  $\mathbf{X}^k$ ,  $k = 1, 2$ , а  $(\cdot, \cdot)_{\mathbf{S}}$  — скалярное произведение с матрицей  $\mathbf{S}$  в  $\mathbb{R}^K$ , определённая следующим образом:  $(X, Y)_{\mathbf{S}} = ((X\mathbf{S})^T, Y^T)$ , так как  $X, Y$  — вектор-строчки,  $\|\cdot\|_{\mathbf{S}}$  — норма, порождённая этим скалярным произведением. Скажем, что ряды  $\mathbb{X}_1$  и  $\mathbb{X}_2$  *слабо  $\varepsilon$ -разделимы*, если

$$(22) \quad \rho = \max(\max_{1 \leq i, j \leq K} |\rho_{i,j}^c|, \max_{1 \leq i, j \leq L} |\rho_{i,j}^r|) < \varepsilon.$$

Нас будет интересовать порядок  $\varepsilon$  при различных матрицах  $\mathbf{S}$  и рядах  $\mathbb{X}_1 = (c, c, \dots)$  — некоторая константа и  $\mathbb{X}_2 = (\cos(2\pi\omega k)), k = 1, 2, \dots$ , а так же при различных  $L$  и  $K$  при условии, что мы будем брать только  $N = L + K - 1$  компонент ряда. Когда  $\mathbf{S}$  — единичная матрица, ответ известен:  $\varepsilon$  имеет порядок  $1/\min(L, K)$ . Этот результат может быть найден в [3].

**Лемма 2.** Пусть  $\mathbf{S}$  определена аналогично Следствию 1: Если  $h = N/L$  — целое, тогда  $\mathbf{S}$  — диагональная матрица со следующими диагональными элементами:

$$s_{k,k} = \begin{cases} 1, & \text{если } k = jL + 1 \text{ для некоторых } j = 0, \dots, h-1, \\ \alpha, & \text{в противном случае} \end{cases},$$

где  $0 \leq \alpha \leq 1$ . Тогда  $\rho$  имеет порядок  $1/\min(L, N/L + \alpha(K - N/L))$  при  $L, K \rightarrow \infty$ .

*Доказательство.* Необходимо оценить порядки следующих величин:

$$\rho_{i,j}^c = \frac{\sum_{k=j}^{j+L-1} \cos(2\pi\omega k)}{\sqrt{L(\sum_{k=j}^{j+L-1} \cos^2(2\pi\omega k))}},$$

$$\rho_{i,j}^r = \frac{\sum_{k=1}^K s_{k,k} \cos(2\pi\omega(j+k-1))}{\sqrt{(\sum_{k=1}^K s_{k,k})(\sum_{k=1}^K s_{k,k} \cos^2(2\pi\omega(j+k-1)))}}.$$

Для доказательства используем следующие факты:

$$\sum_{k=1}^n \cos(ak + b) = \csc(a/2) \sin(an/2) \cos\left(\frac{an + a + 2b}{2}\right),$$

$$\sum_{k=1}^n \cos^2(ak + b) = \frac{1}{4}(\csc(a) \sin(2an + a + 2b) - \csc(a) \sin(a + 2b) + 2n),$$

для любых вещественных  $a, b$  и положительного целого  $n$ . Таким образом, когда ряд  $\mathbb{X}_2$  не представляет из себя константу, числитель в  $\rho_{i,j}^c$  имеет порядок  $O(1)$ , а знаменатель —  $O(L)$ . Первая часть доказана, и её доказательство целиком аналогично случаю, когда  $\mathbf{S}$  — единичная матрица.

Для доказательства второй части необходимо разбить суммы и в числи-

теле, и в знаменателе согласно соответствующим значениям  $s_{k,k}$ :

$$\begin{aligned}
\sum_{k=1}^K s_{k,k} \cos(2\pi\omega(j+k-1)) &= \sum_{\substack{1 \leq k \leq K \\ s_{k,k}=1}} \cos(2\pi\omega(j+k-1)) + \\
\sum_{\substack{1 \leq k \leq K \\ s_{k,k}=\alpha}} \alpha \cos(2\pi\omega(j+k-1)) &= O(1) + \alpha O(1) = O(1), \\
\sum_{k=1}^K s_{k,k} &= N/L + \alpha(K - N/L), \\
\sum_{k=1}^K s_{k,k} \cos^2(2\pi\omega(j+k-1)) &= \sum_{\substack{1 \leq k \leq K \\ s_{k,k}=1}} \cos^2(2\pi\omega(j+k-1)) + \\
\sum_{\substack{1 \leq k \leq K \\ s_{k,k}=\alpha}} \alpha \cos^2(2\pi\omega(j+k-1)) &= O(N/L) + \alpha O(K - N/L).
\end{aligned}$$

□

*Замечание:* для выполнения леммы дополнительно требуется, чтобы ряды, составленные из элементов  $(j, j+L, \dots, (h-1)L+j)$ , где  $1 \leq j \leq L$ , также не представляли из себя константу, что эквивалентно тому, что  $\omega \neq k/L$  для любого целого  $k$ .

Таким образом, разделимость константы и синуса становится хуже, чем при обычном варианте: при  $\alpha$ , близких к нулю, оптимальным выбором  $L$  будет  $L \approx \sqrt{N}$ , и, таким образом, получаем порядок разделимости  $1/\sqrt{N}$ .

**Лемма 3.** Пусть  $\mathbf{S}$  определена аналогично Утверждению 1. Тогда  $\rho$  имеет порядок  $\max\left(1/L, \frac{H_L}{\sqrt{NK}}\right)$  при  $L, K \rightarrow \infty$ , где  $H_L$  —  $L$ -е гармоническое число.

*Доказательство.* Необходимо оценить порядки следующих величин:

$$\begin{aligned}
\rho_{i,j}^c &= \frac{\sum_{k=j}^{j+L-1} \cos(2\pi\omega k)}{\sqrt{L(\sum_{k=j}^{j+L-1} \cos^2(2\pi\omega k))}}, \\
\rho_{i,j}^r &= \frac{\sum_{k=1}^K \hat{s}_{k,k} \cos(2\pi\omega(j+k-1))}{\sqrt{(\sum_{k=1}^K \hat{s}_{k,k})(\sum_{k=1}^K \hat{s}_{k,k} \cos^2(2\pi\omega(j+k-1)))}}.
\end{aligned}$$

Доказательство порядка для  $\rho_{i,j}^c$  целиком аналогично предыдущему пункту, поэтому сразу перейдем ко второму. Докажем корреляцию только первых строчек — для остальных доказательство будет целиком аналогичным. Рассмотрим числитель  $\rho_{1,1}^r$ :

$$\begin{aligned}
\sum_{k=1}^K \hat{s}_{k,k} \cos(2\pi\omega k) &= \sum_{k=1}^{L-1} \hat{s}_{k,k} \cos(2\pi\omega k) + \sum_{k=L}^{K-L+1} \frac{\cos(2\pi\omega k)}{L} + \\
\sum_{k=K-L+2}^K \hat{s}_{k,k} \cos(2\pi\omega k) &= I_1 + I_2 + I_3,
\end{aligned}$$

который разбился на три части. Для центральной справедлива оценка  $O(1/L)$ , а для левой части доказательство аналогично правой.

$$\begin{aligned} \left| \sum_{k=1}^{L-1} \frac{1}{L} \left( \frac{k}{L} + \sum_{j=k}^{L-1} \frac{1}{j} \right) \cos(2\pi\omega k) \right| &= \left| \sum_{k=1}^{L-1} \frac{k \cos(2\pi\omega k)}{L^2} + \right. \\ &+ \left. \frac{1}{L} \sum_{k=1}^{L-1} \sum_{j=k}^{L-1} \frac{\cos(2\pi\omega k)}{j} \right| \leq \left| \sum_{k=1}^{L-1} \frac{k \cos(2\pi\omega k)}{L^2} \right| + \left| \frac{1}{L} \sum_{k=1}^{L-1} \sum_{j=k}^{L-1} \frac{\cos(2\pi\omega k)}{j} \right|. \end{aligned}$$

Используя тот факт, что

$$\begin{aligned} \sum_{k=1}^n k \cos(ak + b) &= -\frac{1}{4} \csc^2(a/2) (-(n+1) \cos(an+b) + \\ &+ n \cos(an+a+b) + \cos b), \end{aligned}$$

получаем:

$$\begin{aligned} \left| \sum_{k=1}^{L-1} \frac{k \cos(2\pi\omega k)}{L^2} \right| &= O(1/L), \quad \left| \frac{1}{L} \sum_{k=1}^{L-1} \sum_{j=k}^{L-1} \frac{\cos(2\pi\omega k)}{j} \right| = \\ \left| \frac{1}{L} \sum_{j=1}^{L-1} \sum_{k=1}^j \frac{\cos(2\pi\omega k)}{j} \right| &\leq \left| \frac{1}{L} \sum_{j=1}^{L-1} \frac{\hat{c}}{j} \right| = O\left(\frac{H_L}{L}\right), \end{aligned}$$

где  $\hat{c}$  — константа. Для знаменателя нужно рассмотреть следующие суммы:

$$\sum_{k=1}^K \hat{s}_{k,k} = N/L$$

по определению, а следующую составляющую просто оценить снизу:

$$\sum_{k=1}^K \hat{s}_{k,k} \cos^2(2\pi\omega(j+k-1)) \geq \sum_{k=1}^K \frac{1}{L} \cos^2(2\pi\omega(j+k-1)) = O\left(\frac{K}{L}\right).$$

□

Для рядов  $\mathbb{X} = (\sin(\frac{2k\pi}{6}))$ ,  $k = 1, \dots, 120$ , и  $\mathbb{Y} = (1, 1, \dots, 1)$  были посчитаны порядки слабой  $\varepsilon$ -разделимости при  $L = 10, 20, 24, 30, 60$  и различных матрицах  $\mathbf{S}$ . Результат представлен на рисунке 2. Стоит заметить, что при небольших  $L$  порядки слабой разделимости совпадают, потому что одинаковые для всех  $\mathbf{S}$  корреляции столбцов по модулю больше, чем корреляции строк во всех случаях.

## 6.8. Небольшое исследование скорости алгоритма в применении к задаче очистки ряда от шума

Для начала, следует ответить на некоторые оставшиеся вопросы. Первый из них — в каком пространстве будет лежать результат работы алгоритма, если мы будем использовать  $\mathbf{S}$  из леммы 1 при  $\alpha = 0$ .

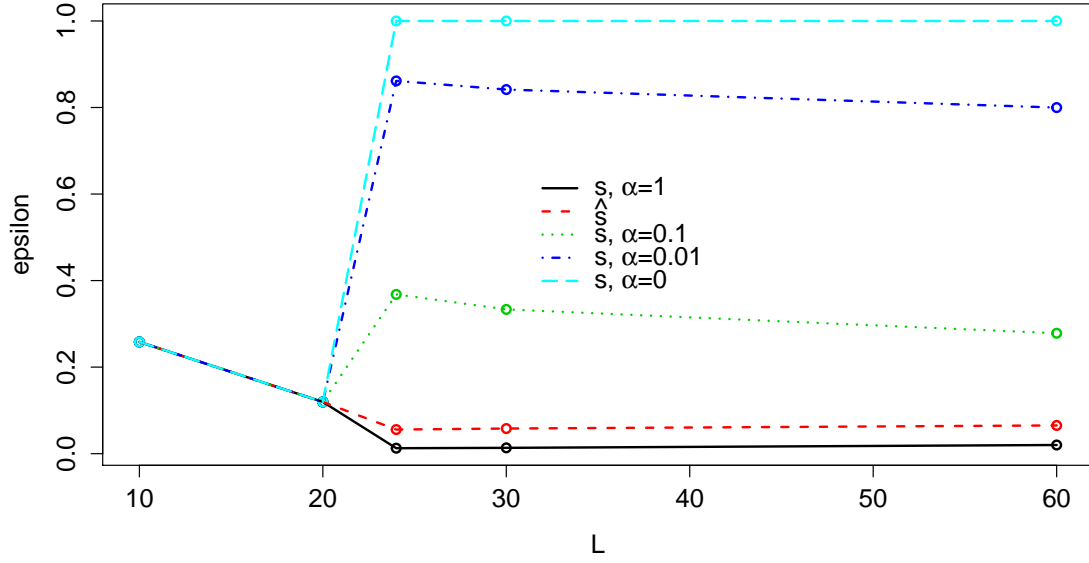


Рис. 2. Порядок слабой разделимости

Рассмотрим временной ряд  $\mathbb{X} = (x_1, \dots, x_N)$  длины  $N \geq 3$ . Зафиксируем длину окна  $L$  ( $1 < L < N$ ) такую, что  $N/L = h$  — целое. Также рассмотрим последовательность векторов (непересекающиеся вектора вложения):

$$\hat{X}_i = (x_{1+(i-1)L}, \dots, x_{iL})^T, \quad i = 1, \dots, h.$$

Непересекающимся траекторным пространством ряда  $\mathbb{X}$  назовем

$$\hat{\mathcal{X}}^{(L)}(\mathbb{X}) = \hat{\mathcal{X}}^{(L)} = \text{span}(\hat{X}_1, \dots, \hat{X}_h).$$

**Определение 2.** Пусть  $0 \leq r \leq L$ . Будем говорить, что ряд  $\mathbb{X}$  имеет непересекающийся  $L$ -ранг  $r$ , если  $\dim \hat{\mathcal{X}}^{(L)} = r$ .

Заметим, что ряд  $\mathbb{X}$  может иметь непересекающийся  $L$ -ранг  $r$  только тогда, когда

$$(23) \quad r \leq \min(L, h).$$

Скажем, что при фиксированном  $r$  длина окна  $L$  является *допустимой*, если для неё выполнено условие (23).

Тогда  $\hat{\mathcal{X}}_N^r$  — множество всех временных рядов длины  $N$  непересекающегося  $L$ -ранга  $r$  — и есть множество, в котором будет лежать результат алгоритма. Важные замечания: во-первых,  $\mathcal{X}_N^r \subset \hat{\mathcal{X}}_N^r$ , так как соответствующие траекторные подпространства вкладываются друг в друга из-за того, что непересекающиеся вектора вложения являются подмножеством векторов вложения. Во-вторых, диагональное усреднение является биективным отображением — усреднение происходит только по одной компоненте матрицы. Поэтому, алгоритм Cadzow при  $\alpha = 0$  вырождается в одну операцию проектирования на матрицы меньшего ранга.

Следующий вопрос: достаточно хорошо ли устроено пространство  $\hat{X}_N^r$ , чтобы проекцию на него использовать в качестве ответа на задачу об очистке зашумленного сигнала шума. Был взят следующий сигнал:

$$\mathbb{X} = (x_1, \dots, x_N), \quad x_i = 5 \sin \frac{2k\pi}{6}, \quad k = 1, \dots, N, \quad N = 40.$$

Алгоритмы запускались на следующем наборе рядов:

$$\mathbb{Z}^j = \mathbb{X} + R_j^T, \quad j = 1, \dots, M, \quad M = 1000.$$

где каждый  $R_j$  —  $N$  гауссовский белый шум с математическим ожиданием 0 и дисперсией, равной 1. Были взяты различные матрицы  $\mathbf{S}$  и построена оценка отклонения от  $\mathbb{X}$  при различном числе итераций. Параметры:  $L = 8$  (соответственно,  $h = 5$ ),  $r = 2$ . Результаты представлены на рисунке 3.

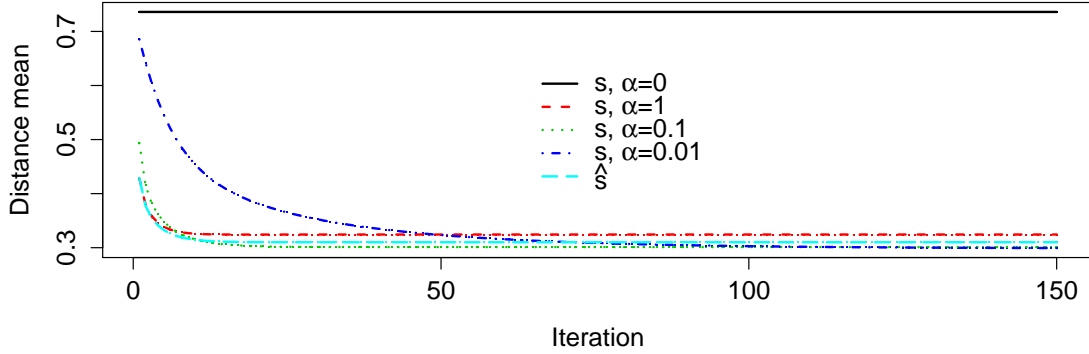


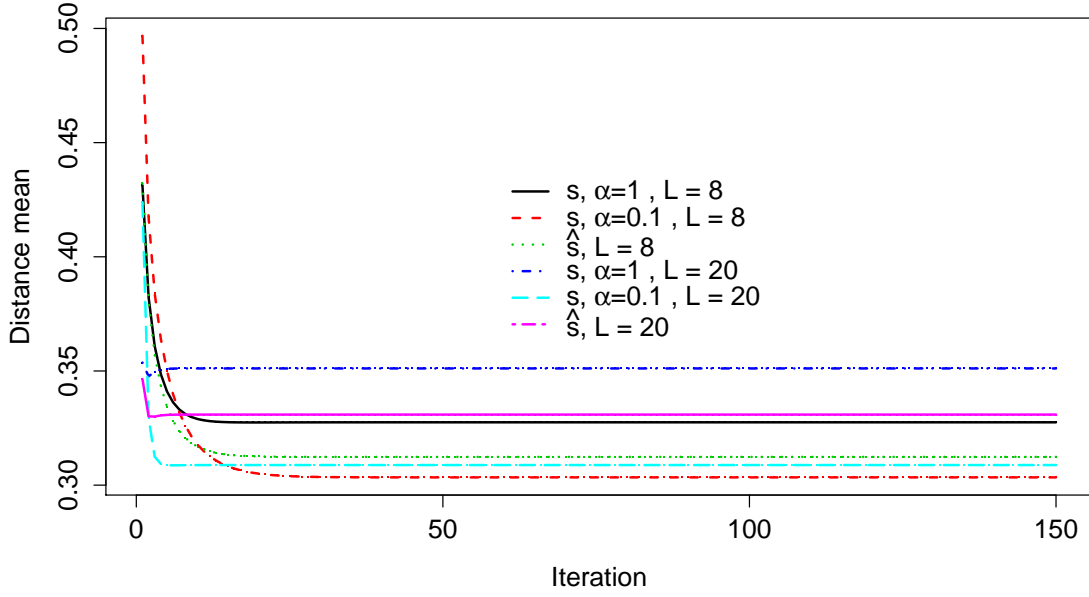
Рис. 3. Оценка отклонения от  $\mathbb{X}$  при различном числе итераций

Видно, что при  $\alpha = 0$  оценка, построенная проекцией на  $\hat{X}_N^r$ , является достаточно плохой. При  $\alpha$ , близких к нулю, алгоритм на первой итерации дает ряд, достаточно близкий к  $\hat{X}_N^r$ , но в дальнейшем алгоритм сходится к хорошей оценке, но с низкой скоростью: чем ближе  $\alpha$  к нулю, тем медленнее сходимость. Дополнительно, был проведён эксперимент с  $L_1 = 8$  и  $L_2 = 20$ , где результаты размещены на одном рисунке 4. Видно, что чем меньше  $L$ , тем ближе веса к единичным, но, с другой стороны, падает скорость сходимости.

## 7. Сходимость алгоритмов

Для новых введенных алгоритмов еще не показаны свойства, касающиеся их сходимости. Известный результат для базового алгоритма может быть найден в [1]. Рассмотрим общий случай. Для этого нам понадобятся следующие утверждения:

**Утверждение 2.** Пусть  $X$  — гильбертово пространство,  $\mathcal{M}$  — замкнутое относительно умножения на скалярную величину подмножество ( $x \in \mathcal{M} \Rightarrow$



**Рис. 4.** Оценка отклонения от  $\mathcal{X}$  при различном числе итераций и различных  $L$

$\alpha x \in \mathcal{M} \forall \alpha \in \mathbb{R}$ ),  $\Pi_{\mathcal{M}}$  — оператор проектирования на  $\mathcal{M}$ <sup>1</sup>. Тогда для любого  $x \in X$  выполняется теорема Пифагора:  $\|x\|^2 = \|x - \Pi_{\mathcal{M}}x\|^2 + \|\Pi_{\mathcal{M}}x\|^2$ .

*Доказательство.* Представим множество  $\mathcal{M}$  в виде  $\mathcal{M} = \bigcup_{l \in L} l$ , где  $L$  — множество всех прямых, лежащих в  $\mathcal{M}$  и проходящих через 0, и  $l \cap m = \{0\}$  для любых  $l, m \in L, l \neq m$ . Тогда операция проектирования может быть записана так: вначале мы выбираем прямую  $l$  такую, что  $\text{dist}(x, l) \rightarrow \min_{l \in L}$ , после чего  $y = \Pi_{\mathcal{M}}x$  — это проекция  $x$  на  $l$ . Получаем нужное свойство.  $\square$

Теперь рассмотрим  $\mathbb{R}^{L \times K}$ , множество матриц ранга, не превосходящего ранга  $r$   $\mathcal{M}_r$ , множество ганкелевых матриц  $\mathcal{H}$ , при этом нас устроит любая порожденная скалярным произведением метрика,  $\Pi_{\mathcal{M}_r}$  и  $\Pi_{\mathcal{H}}$  — проекторы на соответствующее пространства в зафиксированной метрике. Более того, все утверждения справедливы и для траекторных псевдоматриц в алгоритме Extended Cadzow. Пусть  $\mathbf{X}_0 \in \mathcal{H}$ ,  $\mathbf{X}_k = \Pi_{\mathcal{H}}\Pi_{\mathcal{M}_r}\mathbf{X}_{k-1}$ ,  $k = 1, 2, \dots$

**Утверждение 3.** [?] Справедливы следующие неравенства:

$$\|\mathbf{X}_k - \Pi_{\mathcal{M}_r}\mathbf{X}_k\| \geq \|\Pi_{\mathcal{M}_r}\mathbf{X}_k - \mathbf{X}_{k+1}\| \geq \|\mathbf{X}_{k+1} - \Pi_{\mathcal{M}_r}\mathbf{X}_{k+1}\|.$$

*Доказательство.* Для проектора  $\Pi_{\mathcal{H}}$  справедливо:

$$\|\Pi_{\mathcal{M}_r}\mathbf{X}_k - \mathbf{X}_{k+1}\| = \|\Pi_{\mathcal{M}_r}\mathbf{X}_k - \Pi_{\mathcal{H}}\Pi_{\mathcal{M}_r}\mathbf{X}_k\| \leq \|\Pi_{\mathcal{M}_r}\mathbf{X}_k - \mathbf{Y}\|$$

<sup>1</sup>На самом деле, проекция может быть определена неоднозначно, но дальше будем предполагать, что оператор однозначен. В данном случае, можно выбирать любой вариант для проекции, это никак не влияет на истинность утверждений.

для любого  $\mathbf{Y} \in \mathcal{H}$ , в частности, и для  $\mathbf{Y} = \mathbf{X}_k$ . Оставшееся неравенство доказывается аналогично.  $\square$

**Утверждение 4.** В топологии, порожденной фиксированной метрикой (эквивалентной порожденной фробениусовской нормой) множество матриц  $\mathcal{M}_r$  ранга, не превосходящего  $r$ , является замкнутым.<sup>2</sup>

*Доказательство.* Опустим случай, когда  $r = \min(L, K)$  — тогда  $\mathcal{M}_r = \mathbb{R}^{L \times K}$ , и оно очевидно замкнуто.

Возьмем матрицу  $\mathbf{X} \in \mathbb{R}^{L \times K}$ . Если её ранг больше  $r$ , тогда существует квадратная подматрица порядка  $r + 1$ , определитель которой не равен нулю. Если же ранг меньше, либо равен  $r$ , то определители всех таких подматриц равны нулю.

Построим следующее отображение:

$$f(\mathbf{X}) = \sum_{\mathbf{Y} \in \mathcal{Y}} |\det \mathbf{Y}|,$$

где  $\mathcal{Y}$  — множество всех подматриц матрицы  $\mathbf{X}$  порядка  $r + 1$ . Заметим, что по построению  $f$  — непрерывное отображение.

Теперь рассмотрим любую сходящуюся последовательность из  $\mathcal{M}_r$ :  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}^*$  — предельная точка. Знаем, что для любого  $k = 1, 2, \dots$ :  $f(\mathbf{X}_k) = 0$ . Пользуясь непрерывностью отображения  $f$  получаем, что  $f(\mathbf{X}^*) = 0$ . А это означает то, что  $\mathbf{X}^* \in \mathcal{M}_r$ .  $\square$

**Теорема 1.** 1.  $\|\mathbf{X}_k - \Pi_{\mathcal{M}_r} \mathbf{X}_k\| \rightarrow 0$  при  $k \rightarrow +\infty$ ,  $\|\Pi_{\mathcal{M}_r} \mathbf{X}_k - \mathbf{X}_{k+1}\| \rightarrow 0$  при  $k \rightarrow +\infty$ .

2. Существует сходящаяся подпоследовательность  $\mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \dots$  такая, что ее предел  $\mathbf{X}^*$  лежит в  $\mathcal{M}_r \cap \mathcal{H}$ .

*Доказательство.* 1. Согласно утверждению 3, последовательности  $(\|\mathbf{X}_k - \Pi_{\mathcal{M}_r} \mathbf{X}_k\|)$  и  $(\|\Pi_{\mathcal{M}_r} \mathbf{X}_k - \mathbf{X}_{k+1}\|)$ ,  $k = 1, 2, \dots$  являются невозрастающими. Очевидно, они ограничены снизу нулем. Поэтому они имеют одинаковый предел  $c$ , опять же согласно 3.

Докажем, что  $c = 0$ . Предположим противное: существует  $d > 0$  такое, что для любого  $k = 1, 2, \dots$ :  $\|\mathbf{X}_k - \Pi_{\mathcal{M}_r} \mathbf{X}_k\| > d$ ,  $\|\Pi_{\mathcal{M}_r} \mathbf{X}_k - \mathbf{X}_{k+1}\| > d$ . Согласно утверждению 2, справедливо следующее:

$$\begin{aligned} \|\mathbf{X}_k\|^2 &= \|\Pi_{\mathcal{M}_r} \mathbf{X}_k\|^2 + \|\mathbf{X}_k - \Pi_{\mathcal{M}_r} \mathbf{X}_k\|^2 = \\ &= \|\Pi_{\mathcal{M}_r} \mathbf{X}_k - \mathbf{X}_{k+1}\|^2 + \|\Pi_{\mathcal{M}_r} \mathbf{X}_k - \mathbf{X}_{k+1}\|^2 + \|\mathbf{X}_{k+1}\|^2. \end{aligned}$$

Таким образом,  $\|\mathbf{X}_{k+1}\|^2 < \|\mathbf{X}_k\|^2 - 2d^2$ . Расписывая неравенство аналогично дальше, получим, что для любого  $j = 1, 2, \dots$ :  $\|\mathbf{X}_{k+j}\|^2 < \|\mathbf{X}_k\|^2 - 2jd^2$ . Возьмем любое  $k$ , например  $k = 1$ , и  $j = \lceil \|\mathbf{X}_k\|^2 / (2d^2) \rceil + 1$ . Тогда  $\|\mathbf{X}_{k+j}\|^2 < 0$ , чего не может быть.

<sup>2</sup>Для алгоритма Extended Cadzow  $\mathcal{M}_r$ , вообще говоря, не является даже множеством матриц, но оно является непрерывной проекцией замкнутого множества матриц ранга, не превосходящего  $r$ , порядка  $L \times N + L - 1$ .

2. Рассмотрим последовательность  $(\Pi_{\mathcal{M}_r} \mathbf{X}_k)$ ,  $k = 1, 2, \dots$ . Она ограничена, так как  $\|\Pi_{\mathcal{M}_r} \mathbf{Y}\| \leq \|\mathbf{Y}\|$  и  $\|\Pi_{\mathcal{H}} \mathbf{Y}\| \leq \|\mathbf{Y}\|$  для любого  $\mathbf{Y} \in \mathbb{R}^{L \times K}$  (это справедливо, например, и по утверждению 2). Тогда мы можем выбрать из нее сходящуюся подпоследовательность  $(\Pi_{\mathcal{M}_r} \mathbf{X}_{i_k})$ ,  $\mathbf{X}^*$  — её предел, при этом  $\|\Pi_{\mathcal{M}_r} \mathbf{X}_{i_k} - \mathbf{X}_{i_k+1}\| = \|\Pi_{\mathcal{M}_r} \mathbf{X}_{i_k} - \Pi_{\mathcal{H}} \Pi_{\mathcal{M}_r} \mathbf{X}_{i_k}\| \rightarrow 0$  при  $k \rightarrow +\infty$ . Учитывая, что  $\|\mathbf{Y} - \Pi_{\mathcal{H}} \mathbf{Y}\|$  — композиция непрерывных отображений, получаем, что  $\|\mathbf{X}^* - \Pi_{\mathcal{H}} \mathbf{X}^*\| = 0$  и зная, что  $\mathcal{M}_r$  — замкнутое множество, получаем, что  $\mathbf{X}^* \in \mathcal{M}_r \cap \mathcal{H}$ . Осталось увидеть, что последовательность  $(\Pi_{\mathcal{H}} \Pi_{\mathcal{M}_r} \mathbf{X}_{i_k})$  — сходящаяся, так как  $\Pi_{\mathcal{H}}$  — непрерывное отображение, и ее предел равен  $\mathbf{X}^*$ . Получаем, что  $\mathbf{X}_{i_k+1}$  — требуемая подпоследовательность.

□

## 8. Исследование алгоритмов в задаче восстановления рядов и подведение итогов

Была произведена проверка набора известных алгоритмов для приближения временных рядов в задаче восстановлении рядов. Важно понимать, мы наблюдаем зашумленные ряды, и к зашумленным рядам применяем алгоритмы, пытаясь оценить именно исходный ряд. Исследуемые ряды имеет вид:

$$\mathbb{X} = (x_1, \dots, x_N), \quad x_k = 5 \sin \frac{2k\pi}{6}, \quad k = 1, \dots, N, \quad N = 40.$$

$$\mathbb{Y} = (y_1, \dots, y_N), \quad y_k = 5 \sin \frac{2(k + 2/3)\pi}{6} e^{k/20}, \\ k = 1, \dots, N, \quad N = 40.$$

Алгоритмы запускались на следующем наборе рядов:

$$\mathbb{Z}_j^1 = \mathbb{X} + R_j^T, \quad j = 1, \dots, M, \quad M = 1000, \\ \mathbb{Z}_j^2 = \mathbb{Y} + R_j^T, \quad j = 1, \dots, M, \quad M = 1000.$$

где каждый  $R_j$  — гауссовский белый шум длины  $N$  с математическим ожиданием 0 и дисперсией, равной 1. При этом везде выбрано  $r = 2$ , а оптимальной шириной окна во всех случаях оказалось  $L = 20$ .

Для каждого алгоритма было рассмотрено:

- Используемые веса ряда  $q_i$
- $P$  — оценка среднеквадратического отклонения до  $\mathbb{X}$  или  $\mathbb{Y}$  (оценивание сигнала) на одной итерации алгоритма
- $P$  — оценка среднеквадратического отклонения до  $\mathbb{X}$  или  $\mathbb{Y}$  (оценивание сигнала) на большом ( $T = 100$ ) числе итераций.
- Трудоёмкость



### 8.0.1. Алгоритм Cadzow при $\alpha = 1$ .

- Имеет веса  $q_i = w_i$ , описанные в (11).
- $P = 0.3749$  для ряда  $\mathbb{X}$  и  $P = 0.3762$  для ряда  $\mathbb{Y}$  на одной итерации алгоритма. Алгоритм хорошо работает на одной итерации благодаря хорошим свойствам слабой разделимости рядов.
- $P = 0.3714$  для ряда  $\mathbb{X}$  и  $P = 0.3661$  для  $\mathbb{Y}$  на ста итерациях.
- Алгоритм быстро сходится и не требует дополнительных итераций для проектирования на множество матриц неполного ранга.

### 8.0.2. Алгоритм Cadzow при $\alpha = 0.1$ .

- Имеет веса  $q_i$ , описанные в Следствии 1.
- $P = 0.4312$  для ряда  $\mathbb{X}$  и  $P = 0.4434$  для ряда  $\mathbb{Y}$  на одной итерации алгоритма. Не самые лучшие результаты объясняются плохой слабой разделимостью рядов.
- $P = 0.3254$  для ряда  $\mathbb{X}$  и  $P = 0.3220$  для  $\mathbb{Y}$  на ста итерациях. Хорошие результаты объясняются близкими к единичным весами.
- Алгоритм не требует дополнительных итераций для проектирования на множество матриц неполного ранга, однако чем меньше  $\alpha$ , тем медленнее он сходится.

### 8.0.3. Алгоритм Cadzow с матрицей $\hat{\mathbf{S}}$ .

- Имеет веса  $\hat{q}_i$ , описанные в Утверждении 2.
- $P = 0.3643$  для ряда  $\mathbb{X}$  и  $P = 0.3718$  для ряда  $\mathbb{Y}$  на одной итерации алгоритма.
- $P = 0.3494$  для ряда  $\mathbb{X}$  и  $P = 0.3439$  для  $\mathbb{Y}$  на ста итерациях.
- Алгоритм быстрый и не требует дополнительных итераций для проектирования на множество матриц неполного ранга. Практически не уступает основному алгоритму Cadzow.

### 8.0.4. Алгоритм Weighted Cadzow.

- Имеет веса  $q_i = 1$  для любого  $i = 1, \dots, N$ .
- $P = 0.3640$  для ряда  $\mathbb{X}$  и  $P = 0.3734$  для ряда  $\mathbb{Y}$  на одной итерации алгоритма, но серьёзно “притягивается” на краях шумом.
- $P = 0.3393$  для ряда  $\mathbb{X}$  и  $P = 0.3343$  для  $\mathbb{Y}$  на ста итерациях.
- Алгоритм требует дополнительные итерации для проектирования на множество матриц неполного ранга, поэтому достаточно медленный.

### 8.0.5. Алгоритм Extended Cadzow.

- Имеет веса  $q_i = 1$  для любого  $i = 1, \dots, N$  (но целевая функция не в точности соответствует исходной, так как рассматривается расширенный ряд).
- $P = 0.3330$  для ряда  $\mathbb{X}$  и  $P = 0.3474$  для ряда  $\mathbb{Y}$  на одной итерации алгоритма. Может “притягиваться” на краях шумом.
- $P = 0.3171$  для ряда  $\mathbb{X}$  и  $P = 0.3191$  для  $\mathbb{Y}$  на ста итерациях. Алгоритм очень нестабилен для нестационарных рядов, если в качестве оценки скрытых переменных брать среднее по исходному ряду, а не использовать векторное прогнозирование.
- Алгоритм требует дополнительные итерации для проектирования на множество матриц неполного ранга, и вследствие своей специфики, очень медленный.

Изображения, на которых показана ошибка восстановления в точке для различных рядов и методов, представлены ниже. Оценки среднеквадратичного отклонения собраны в таблицу [1](#).

**Таблица 1.** Сравнение алгоритмов

$P$ :	$\mathbb{X}, T = 1$	$\mathbb{Y}, T = 1$	$\mathbb{X}, T = 100$	$\mathbb{Y}, T = 100$
Cadzow, $\alpha = 1$	0.3749	0.3762	0.3714	0.3661
Cadzow, $\alpha = 0.1$	0.4312	0.4434	0.3254	0.3220
Cadzow $\hat{\mathbf{S}}$	0.3643	0.3718	0.3494	0.3439
Weighted Cadzow	0.3640	0.3734	0.3393	0.3343
Extended Cadzow	0.3330	0.3474	0.3171	0.3191

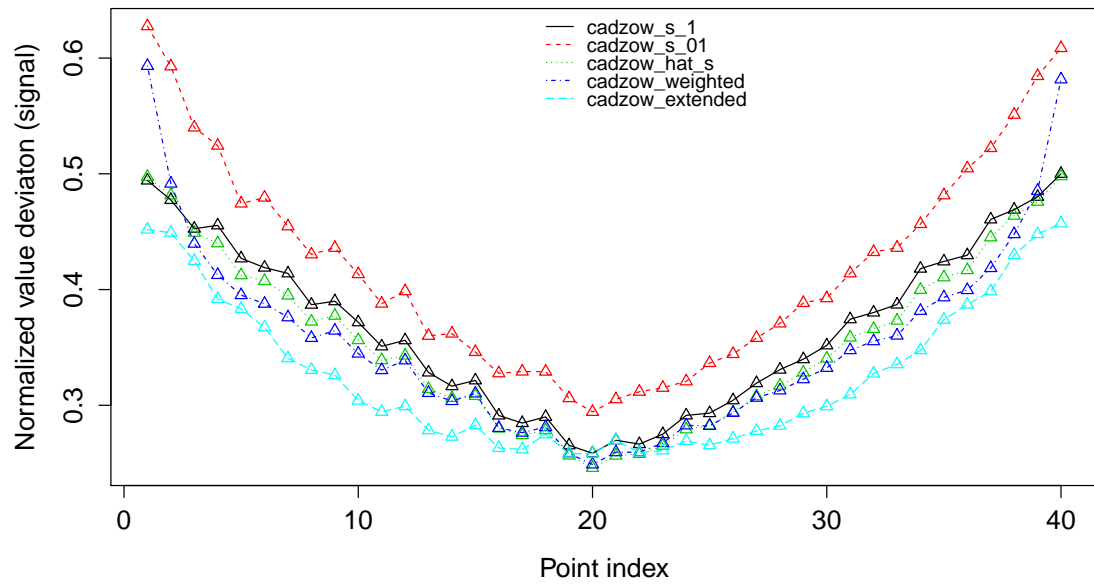


Рис. 5. Ошибка восстановления в точке на одной итерации для ряда X

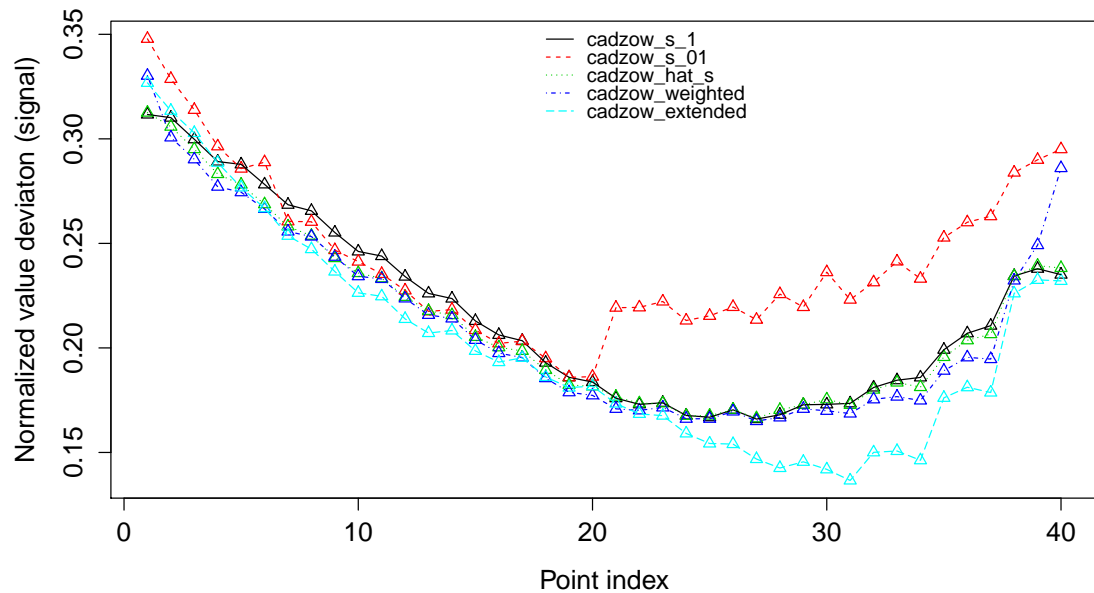


Рис. 6. Нормированная ошибка восстановления в точке на одной итерации для ряда Y

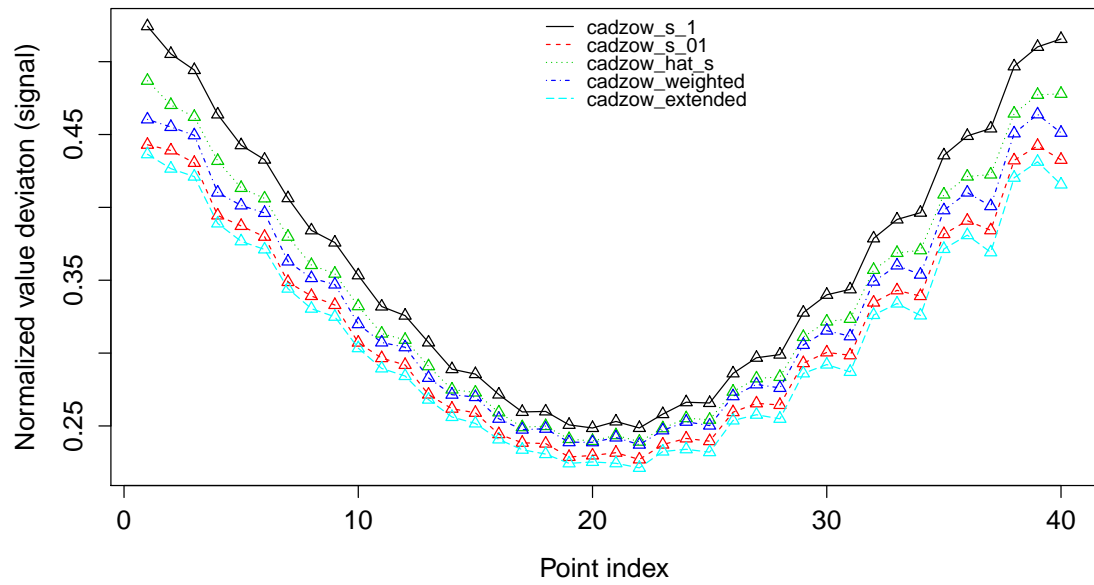


Рис. 7. Ошибка восстановления в точке на ста итерациях для ряда X

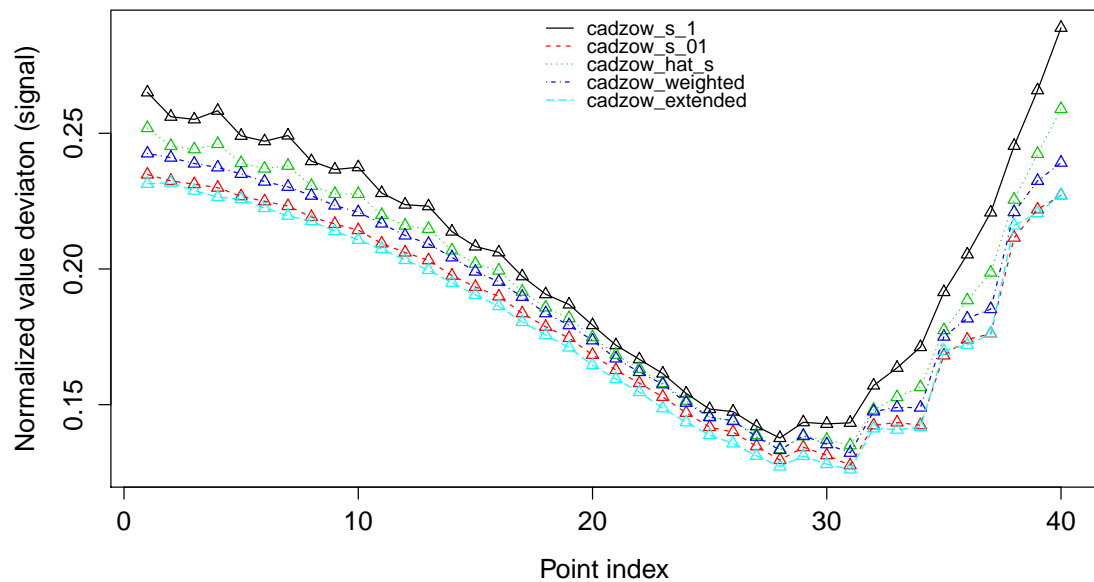


Рис. 8. Нормированная ошибка восстановления в точке на ста итерациях для ряда Y

## Заключение

Таким образом, в работе были рассмотрены и исследованы разные методы решения задачи выделения сигнала, управляемого ЛРФ. Также, было проведено численное сравнение методов.

Перечислим вкратце полученные в результате исследований итоги:

- Weghted Cadzow и Extended Cadzow — очень трудоёмкие алгоритмы, в то время как основной алгоритм и его модификации, использующие косоугольное SVD-разложение, примерно одинаково быстры.
- Чем меньше  $L$ , тем веса  $q_i$  ближе к единичным, и алгоритмы сходятся к лучшему ответу.
- Чем меньше  $L$ , тем скорость сходимости меньше; чем больше  $L$ , тем скорость сходимости выше.
- Для алгоритма Cadzow при различных  $\alpha$ : при  $\alpha = 0$  имеем сходимость не к тому множеству рядов, при  $\alpha > 0$  — сходимость к рядам конечного  $L$ -ранга есть, но чем меньше  $\alpha$ , тем она медленнее.
- В рассмотренных примерах, на одной итерации лучшим оказался трудоёмкий благодаря внутренним итерациям метод Extended Cadzow для длины окна порядка половины длины ряда. В пределе лучшим оказались самые медленно-сходящиеся методы Cadzow( $\alpha$ ) при маленьком  $\alpha$  и маленькой длине окна.

## Список литературы

1. J. A. Cadzow. Signal enhancement: a composite property mapping algorithm. *IEEE Trans. Acoust.*, 36(1):49–62, 1988.
2. J. Gillard and A. A. Zhigljavsky. Stochastic algorithms for solving structured low-rank matrix approximation problems. *Communication in Nonlinear Science and Numerical Simulation*, 2014.
3. N. Golyandina, V. Nekrutkin, and A. Zhigljavsky. *Analysis of Time Series Structure: SSA and Related Techniques*. Chapman&Hall/CRC, 2001.
4. Nina Golyandina and Alex Shlemov. Variations of singular spectrum analysis for separability improvement: non-orthogonal decompositions of time series. 2013.
5. Nathan Srebro, Tommi Jaakkola, et al. Weighted low-rank approximations. In *ICML*, volume 3, pages 720–727, 2003.