На правах рукописи

Звонарев Никита Константинович

Структурные аппроксимации временных рядов

Специальнoть 01.01.07 –
вычислительная математика

Автореферат
диссертации на соискание учёной степени
кандидата физико-математических наук

Санкт-Петербург – 2018

Работа выполнена в федеральном государственном бюджетном образовательном учреждении высшего образования «Санкт-Петербургский государственный университет».

Научный руководитель: **Голяндина Нина Эдуардовна**, кандидат физико-математических наук, доцент

Официальные оппоненты: **Шевляков Георгий Леонидович**, доктор физико-математических наук, профессор, Санкт-Петербургский политехнический университет Петра Великого, профессор кафедры прикладной математики

Антонов Антон Александрович, кандидат физико-математических наук, финансовый математик в ООО «Эксперт-Система»

Ведущая организация: Федеральное государственное бюджетное образовательное учреждение высшего образования «Вологодский государственный университет»

Защита состоится «14» июня 2018 г. в 11 часов 00 минут на заседании диссертационного совета Д 212.232.49 на базе Санкт-Петербургского государственного университета по адресу: 198504, Санкт-Петербург, Старый Петергоф, Университетский пр., д. 28, ауд. 405.

С диссертацией можно ознакомиться в Научной библиотеке им. М. Горького Санкт-Петербургского государственного университета по адресу: 199034, Санкт-Петербург, Университетская наб., 7-9 и на сайте:
`https://disser.spbu.ru/files/disser2/disser/bprpwWD9w4.pdf`.

Автореферат разослан «____» ____________ 2018 г.

Ученый секретарь диссертационного совета,
доктор физико-математических наук Чурин Ю.В.

Общая характеристика работы

Актуальность темы. *Временные ряды* встречаются во многих областях науки. Чаще всего, временной ряд — последовательно измеренный через равноотстоящие промежутки времени вещественный показатель некоторого процесса. Кроме того, название временной ряд сохраняется и в том случае, когда измерения сделаны не во временных точках, а равномерно вдоль пространственной координаты. Будем представлять временной ряд длины $N \geq 2$ как вектор-столбец $\mathsf{X} = (x_1, \ldots, x_N)^{\mathrm{T}} \in \mathbb{R}^N$.

Для того чтобы дальнейшее исследование временных рядов стало возможным, строится модель их представления. Предположим, что временной ряд имеет следующий вид:

$$x_n = s_n + \epsilon_n\,, \quad n = 1, 2, \ldots, N\,,$$

где X — *ряд наблюдений* , $\mathsf{S} = (s_1, \ldots, s_N)^{\mathrm{T}}$ — *сигнальная составляющая* (или просто *сигнал*) и $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_N)^{\mathrm{T}}$ — *шумовая составляющая* (или просто *шум*). Предполагается, что сигнальная составляющая имеет некоторую структуру, а шумовая составляющая является реализацией некоего случайного процесса с нулевым математическим ожиданием.

Рассмотрим следующий параметрический вид сигнала S в виде конечной суммы:

$$s_n = \sum_k p_{m_k}(n) \exp(\alpha_k n) \sin(2\pi\omega_k n + \phi_k), \tag{1}$$

где $p_{m_k}(n)$ — многочлены от n степени m_k. Такой вид сигнала используется во многих приложениях, например, в теории обработки сигнала [5], в задачах идентификации линейных систем [6], задачах распознавания речи [7] и многих других. В приложениях к теории обработки сигнала часто считается, что сигнал S в представлении (1) является суммой синусоид [5] или экспоненциально-модулированных синусоид [6].

Важной задачей анализа временных рядов является оценка сигнала S по ряду наблюдений X. Полученную оценку сигнала можно использовать для оценивания параметров сигнала, построения оценки прогноза сигнала и разложения сигнала на аддитивные составляющие [8].

Часто явный вид сигнала (1) неизвестен, то есть неизвестно количество ненулевых слагаемых k, количество ненулевых частот ω_k и так далее. Вместо этого фиксируют так называемый *ранг ряда* S, определённый следующим способом. Для заданного натурального числа L, $1 < L < N$, называемого *длиной окна*, определим *оператор вложения* $\mathcal{T}_L:\ \mathbb{R}^N \to \mathbb{R}^{L\times(N-L+1)}$ как

$$\mathcal{T}_L(\mathsf{S}) = \begin{pmatrix} s_1 & s_2 & \dots & s_{N-L+1} \\ s_2 & s_3 & \dots & \vdots \\ \vdots & \vdots & \dots & s_{N-1} \\ s_L & s_{L+1} & \dots & s_N \end{pmatrix}. \tag{2}$$

Скажем, что временной ряд S имеет ранг r если $\operatorname{rank}\mathcal{T}_L(\mathsf{S}) = r < N/2$ для любого L такого, что $\min(L, N-L+1) > r$ [8]. Например, сумма двух экспонент, как и синусоидальный сигнал или линейная функция, имеет ранг 2.

Матрица $\mathcal{T}_L(\mathsf{S})$ является ганкелевой, то есть элементы на её побочных диагоналях равны. Перестановкой строк или столбцов матрицы $\mathcal{T}_L(\mathsf{S})$ в обратном порядке можно добиться равенства элементов на её диагоналях, то есть привести её к эквивалентному тёплицевому виду. Модель данных, где соответствующая сигналу S ганкелева матрица $\mathcal{T}_L(\mathsf{S})$ имеет конечный ранг, встречается в теории обработки сигналов [5, 9], задачах распознавания голоса [7], теории управления и теории стохастических систем [6, 10]. Выбор значения ранга сигнала — отдельная задача и в данной работе не рассматривается. В дальнейшем считаем ранг временного ряда r известным.

Множество всевозможных временных рядов ранга r обозначим как $\mathcal{D}_r$. Тогда для решения задачи оценивания сигнала можно использовать решение следующей нелинейной задачи взвешенных наименьших квадратов:

$$\mathsf{Y}^\star = \operatorname*{arg\,min}_{\mathsf{Y}\in\overline{\mathcal{D}_r}} \|\mathsf{X}-\mathsf{Y}\|^2_{\mathbf{W}}, \tag{3}$$

где $\|\mathsf{Z}\|^2_{\mathbf{W}} = \mathsf{Z}^{\mathrm{T}}\mathbf{W}\mathsf{Z}$ — квадрат косоугольной евклидовой нормы, $\mathbf{W}\in\mathbb{R}^{N\times N}$ — положительно определённая матрица весов, $\overline{\mathcal{D}_r}$ — замыкание множества $\mathcal{D}_r$. Общепринятое название данной задачи — *Hankel structured low-rank approximation* (HSLRA) [10]. Для определённости, будем называть (3) задачей HSLRA в векторном виде.

Если $\boldsymbol{\epsilon}$ — гауссовский шум с нулевым средним и ковариационной матрицей $\mathbf{\Sigma}$, то решение задачи (3) с матрицей весов $\mathbf{W} = \mathbf{\Sigma^{-1}}$ является оценкой максимального правдоподобия (ОМП) сигнала S. Заметим, что для того чтобы оценка была ОМП, достаточно знать матрицу $\mathbf{\Sigma}$ с точностью до константы, так как решение задачи (3) не зависит от умножения $\mathbf{W}$ на положительную константу.

Задача (3) сформулирована как задача поиска глобального минимума, но известно, что целевая функция является невыпуклой [11], следовательно, может содержать множество локальных минимумов. Для решения такой задачи можно либо использовать методы глобального поиска, либо локальный поиск с выбором начального приближения, достаточно близкого к глобальному минимуму. В работе разрабатываются методы локального поиска; также рассматривается и вопрос выбора начального приближения.

Для решения задачи (3) используются численные методы. Для их построения используются два основных подхода. Первый — так называемый принцип Variable projection, рассмотренный в работах [12, 13] для более общего, чем в (3), случая произвольной аффинной структуры. В [13] после применения принципа Variable projection используются методы Гаусса-Ньютона и Левенберга-Марквардта (регуляризованная версия метода Гаусса-Ньютона), которые являются методами локального поиска, но используют параметризацию, отличную от (1). Несмотря на свои достоинства (самые главные из которых — сходимость к локальному минимуму и большая эффективность по времени работы в случае диагональной $\mathbf{W}$), методы обладают рядом недостатков. Первый — методы чувствительны к форме матрицы $\mathbf{W}$, то есть, например, если $\mathbf{W}$ является хотя бы трёхдиагональной, то быстрая реализация метода невозможна. Второй недостаток — методы проявляют неустойчивость в ряде случаев, например, когда S является полиномиальным сигналом.

Второй подход — непараметрический метод итераций Кэдзоу, входящий в класс алгоритмов попеременных проекций (Alternating Projections) [5]. Метод решает задачу HSLRA в матричном виде:

$$\mathbf{Y}^\star = \operatorname*{arg\,min}_{\mathbf{Y}\in\mathcal{M}_r\cap\mathcal{H}} f_{\mathbf{L},\mathbf{R}}(\mathbf{Y}), \quad f_{\mathbf{L},\mathbf{R}}(\mathbf{Y}) = \|\mathbf{X}-\mathbf{Y}\|^2_{\mathbf{L},\mathbf{R}}, \tag{4}$$

$\|\mathbf{X}\|_{\mathbf{L},\mathbf{R}}$ — порождённая скалярным произведением $\langle\mathbf{X},\mathbf{Y}\rangle_{\mathbf{L},\mathbf{R}} = \operatorname{tr}(\mathbf{LXRY}^{\mathrm{T}})$ матричная норма, $\mathcal{H} = \mathcal{H}_{L,K}$ — пространство ганкелевых матриц размера $L\times K$, $\mathcal{M}_r \subset \mathbb{R}^{L\times K}$ — множество матриц ранга, не превосходящего r, а $\mathbf{X}$, $\mathbf{Y}$ — матрицы, связаные с задачей (3) соотношениями $\mathbf{X} = \mathcal{T}_L(\mathsf{X})$ и $\mathbf{Y} = \mathcal{T}_L(\mathsf{Y})$.

Теория метода Кэдзоу тесно связана с теорией так называемых subspace-based методов и метода SSA (Singular Spectrum Analysis, анализ сингулярного спектра) [8]. Метод можно распространить на случай косоугольной нормы, отличной от евклидовой [14]. Основные проблемы метода — локальные свойства предельной точки, полученной методом Кэдзоу, неизвестны; к тому же, в большинстве случаев метод решает задачу (4) с матрицами весов $\mathbf{L}$, $\mathbf{R}$, не дающими $\mathbf{W}$ в постановке (3) (что ведёт к тому, что метод не может обеспечить ОМП даже в случае белого гауссовского шума [15]). Вопрос выбора подходящих матриц $\mathbf{L}$, $\mathbf{R}$, дающих матрицу весов $\mathbf{W}$, близкую к $\mathbf{\Sigma}^{-1}$, остаётся открытым. Также, быстрая реализация метода Кэдзоу была создана только для диагональной $\mathbf{W}$.

Асимптотические свойства ошибок оценки сигнала с помощью решения задачи (3) рассматривались в работах [9, 16, 17]. Однако полученные в этих работах утверждения либо не учитывают вид матрицы $\mathbf{W}$ [16], либо множество таких матриц $\mathbf{W}$ сильно ограничено [17]. Для алгоритма Кэдзоу известна неоптимальность полученных им оценок: в смысле недостижения алгоритмом локального минимума [18] и в смысле качества оценки сигнала [9]. Однако нет результата о том, насколько критично для оценивания сигнала то, что локальный минимум в задаче (4) точно не находится, и можно ли улучшить качество оценки путём выбора весов $\mathbf{L}$, $\mathbf{R}$.

В диссертации рассматривается и развивается как параметрический, так и непараметрический подход. Предлагаются новый метод локального поиска, базирующийся на одном из стандартных для нелинейной задачи наименьших квадратов методе Гаусса-Ньютона, и расширение непараметрического метода Кэдзоу. Ключевыми факторами при выборе методов для исследования являлись: возможность построить эффективную по времени реализацию каждого из методов, возможность использовать метод для недиагональных матриц весов и находить асимптотические свойства оценок, полученных методами при оценивании сигнала в широком классе $\mathcal{D}_r$ сигналов ранга r. Более того, необходимость изучения как непараметрического, так и параметрического подхода объясняется тем, что для решения задачи (3) с использованием методов локального поиска требуется начальное приближение, близкое или лежащее в окрестности глобального минимума. Такое приближение может быть получено с помощью непараметрического метода.

Все исследуемые в данной работе методы оптимизации являются детерминированными, однако, развитие детерминированных методов способно улучшить и качество методов случайного поиска. В частности, метод Кэдзоу используется на одном из шагов в предложенном в статье [14] методе стохастической оптимизации.

Цели и задачи диссертационной работы: Основными целями являются:

1. Разработка и эффективная реализация модифицированного численного метода Гаусса-Ньютона решения задачи HSLRA (3), обладающего большей устойчивостью, чем метод Variable projection [13], сравнение методов по виду итерации, временной сложности и устойчивости.
2. Постановка задачи поиска матриц весов $\mathbf{L}$ и $\mathbf{R}$ для метода Кэдзоу для соответствия задач HSLRA в векторном (3) и матричном (4) виде, теоретическое обоснование и разработка эффективных численных методов её решения.
3. Построение быстрой реализации метода Кэдзоу для случая недиаго-

нальных матриц весов **L** и **R**.

4. Исследование асимптотических по соотношению сигнал/шум ошибок первого порядка для полученных методами оценок сигнала.

Научная новизна. Все результаты, выносимые на защиту, являются новыми.

Теоретическая и практическая значимость. Результаты, полученные в данной работе, позволяют улучшить точность решения задачи HSLRA и расширить область применения методов к случаю недиагональной матрицы весов **W**. Полученные теоретические результаты могут послужить основой для дальнейшего исследования в области структурной аппроксимации. С помощью численных экспериментов было показано, что реализованные алгоритмы могут успешно применяться для решения задачи HSLRA.

Методы исследования. В работе применяются методы линейной алгебры, теория гладких многообразий, теория численных методов оптимизации и решения систем линейных алгебраических уравнений (СЛАУ), теория вероятностей, математическая статистика и функциональный анализ. Для реализации алгоритмов использовались языки программирования R и C.

Положения, выносимые на защиту:

1. Для множества временных рядов $\mathcal{D}_r$ ранга r найдены гладкая параметризация и вид касательного подпространства, необходимые для построения методов локальной оптимизации.
2. Разработан метод вычисления базисов подпространств временных рядов ранга r, теоретически обоснована его корректность и устойчивость, создана устойчивая реализация.
3. На основе предложенной параметризации и алгоритма вычисления базисов разработан и эффективно реализован модифицированный метод Гаусса-Ньютона. Доказано, что алгоритм превосходит метод Variable projection [13] по скорости в случае ленточной матрицы весов **W** и по точности на полиномиальных сигналах.
4. Сформулирована задача поиска весов **L**, **R** для соответствия задач (3) и (4), теоретически обоснована её постановка, построен и реализован алгоритм решения с помощью метода квадратичного программирования.
5. Построена быстрая реализация метода Кэдзоу в случае недиагональных матриц весов **L** и **R**.
6. Найдены виды асимптотических ошибок первого порядка для оценок сигнала с помощью проекции на множество $\mathcal{D}_r$ и с помощью линеаризованного алгоритма Кэдзоу, получен результат про соотношение с границей Рао-Крамера.

Апробация результатов. Основные результаты обсуждались на семинарах кафедры статистического моделирования СПбГУ, семинаре кафедры статистики в School of Mathematics, Cardiff University (Великобритания, июнь 2017) и на международной конференции «Идентификация систем и задачи управления» SICPRO'15 (Москва, 26–29 января 2015 г.). Часть результатов диссертации была получена в ходе работ по гранту РФФИ (проект РФФИ 16-04-00821).

Публикации. По теме диссертационной работы опубликована работа [1] в научном издании, включенном в Перечень рецензируемых научных изданий, рекомендуемых ВАК. Работа [2] опубликована в научном издании, входящем в базы цитирования Web of Science и Scopus. Работа [1], в которой построена формулировка задачи поиска весов для задачи HSLRA, доказаны теоремы 1 и 2 об эквивалентных формулировках задач квадратичного программирования, построен алгоритм быстрого поиска весов, полностью выполнена соискателем. В работах [2–4] постановка задачи, структура работы и введение принадлежат научному руководителю, основной текст написан совместно, а основные результаты получены соискателем. В частности, в работе [2] соискателю принадлежат основные теоретические результаты, в том числе теорема 1 о сходимости метода Кэдзоу по подпоследовательностям, а также проведено численное моделирование оценки сигнала с помощью метода Кэдзоу. В [3] теоремы 2.1, 2.3 и 2.4 о параметризации множества рядов конечного ранга и вида его касательного подпространства, а также алгоритм 5.5 модифицированного метода Гаусса-Ньютона принадлежат соискателю.

Личный вклад автора. Все представленные в диссертации результаты получены лично автором.

Структура и объем диссертации. Диссертация состоит из введения, пяти глав, заключения и библиографии. Общий объём диссертации составляет 151 страницу. В тексте содержится 4 таблицы и 21 рисунок. Библиография работы состоит из 67 наименований.

Содержание работы

В **первой главе** приведён обзор существующих методов и свойств решения задачи (3) аппроксимации временного ряда рядом конечного ранга. Описаны следующие понятия: принцип Variable projection [13], метод Active set решения задачи квадратичного программирования, известные свойства множества $\mathcal{D}_r$ временных рядов ранга r, устанавливающие независимость определения временного ряда ранга r от длины окна.

Приведены понятия, использующиеся в дальнейшем. Будем говорить, что временной ряд управляется *обобщённой линейной рекуррентной формулой* (ОЛРФ) порядка r, если $A^{\mathrm{T}}\mathcal{T}_{r+1}(\mathsf{S}) = 0$ для некоторого $A \in \mathbb{R}^{r+1}$, $A \neq 0$. Ключевая разница между ОЛРФ и обычной линейной рекуррентной формулой вида

$$s_i = \sum_{k=1}^{r} b_k s_{i-k}, \quad b_r \neq 0, \quad i = 1, \dots N-m$$

состоит в том, что последний коэффициент ОЛРФ не обязательно ненулевой. Идея использовать произвольный ненулевой вектор A для задания линейного соотношения, которому удовлетворяет временной ряд, используется в работах [12, 13].

Для пространства временных рядов, управляемых ОЛРФ(A), вводится обозначение $\mathcal{Z}_N(A) = \mathcal{Z}(A)$. Введён оператор $\mathbf{Q}^{M,d}$: $\mathbb{R}^{d+1} \to \mathbb{R}^{M\times(M-d)}$, определённый как $\mathbf{Q}^{M,d}(B) = [Q_1 : \dots : Q_M]$, где $B = (b_1, \dots, b_{d+1})^{\mathrm{T}} \in \mathbb{R}^{d+1}$, $Q_1 = (b_1, \dots, b_{d+1}, 0, \dots, 0)^{\mathrm{T}}$, $Q_2 = (0, b_1, \dots, b_{d+1}, 0, \dots, 0)^{\mathrm{T}}$, …, $Q_M = (0, \dots, 0, b_1, \dots, b_{d+1})^{\mathrm{T}}$. Тогда $\mathcal{Z}(A)$ можно представить как $\mathcal{Z}(A) = \{\mathsf{S} : \mathbf{Q}^{\mathrm{T}}(A)\mathsf{S} = 0\}$, где $\mathbf{Q} = \mathbf{Q}^{N,r}$. Под $\mathbf{Z}(A)$ будем понимать некоторую матрицу, столбцы которой составляют базис $\mathcal{Z}(A)$.

Указан вид итерации взвешенного метода Гаусса-Ньютона для задачи взвешенного метода наименьших квадратов вида $\arg\min_P \|X - S(P)\|^2_{\mathbf{W}}$, где $P \in \mathbb{R}^p$ — вектор параметров, $S : \mathbb{R}^p \to \mathbb{R}^N$ — дифференцируемая по P функция, $\mathbf{W} \in \mathbb{R}^{N\times N}$ — симметричная положительно определённая матрица весов.

Одна итерация алгоритма Гаусса-Ньютона с шагом γ выглядит следующим образом:

$$P_{k+1} = P_k + \gamma \left(\mathbf{J}_S(P_k)\right)^{+}_{\mathbf{W}} (X - S(P_k)), \tag{5}$$

где $\mathbf{J}_S(P_k)$ — матрица Якоби вектор-функции $S(P)$, $(\mathbf{F})^{+}_{\mathbf{W}} = (\mathbf{F}^{\mathrm{T}}\mathbf{W}\mathbf{F})^{-1}\mathbf{F}^{\mathrm{T}}\mathbf{W}$ — взвешенная псевдообратная матрица к $\mathbf{F}$. Указана возможность модифицировать алгоритм Гаусса-Ньютона путём замены точки $S(P_k)$ на более близкую к X.

Приведен метод итераций Кэдзоу [5] — численный метод решения задачи HSLRA (3) в матричном виде (4):

$$\mathbf{X}_0 = \mathbf{X}, \quad \mathbf{X}_{k+1} = \Pi_{\mathcal{H}}\Pi_{\mathcal{M}_r}\mathbf{X}_k, \quad k \geq 0,$$

где $\Pi_{\mathcal{H}}$ и $\Pi_{\mathcal{M}_r}$ — проекторы на соответствующие множества по норме $\|\cdot\|_{\mathbf{L},\mathbf{R}}$. Указаны известные свойства метода Кэдзоу.

Вторая глава посвящена свойствам задачи (3). Построена параметризация множества $\mathcal{D}_r$, необходимая для использования методов локальной оптимизации. Рассмотрим временной ряд $\mathsf{S}_0 \in \mathcal{D}_r$, управляемый минимальной ОЛРФ(A_0) порядка r, заданной ненулевым вектором $A_0 = (a_1^{(0)}, \ldots, a_{r+1}^{(0)})^{\mathrm{T}}$. Рассмотрим индекс τ такой, что $a_\tau^{(0)} = -1$. Если ОЛРФ(A_0) ненулевая, то перенормировкой вектора A_0 такой индекс τ всегда можно найти.

Вместо r значений в начале временного ряда, которые рассматриваются в случае обычной ЛРФ, мы рассмотрим $\tau - 1$ значений в начале, и $r + 1 - \tau$ значений в конце ряда. Определим $\mathcal{I}(\tau) = \{1, \ldots, N\} \setminus \{\tau, \ldots, N - r - 1 + \tau\}$ и $\mathcal{K}(\tau) = \{1, \ldots, r + 1\} \setminus \{\tau\}$ — два множества индексов размера r. Множество $\mathcal{I}(\tau)$ содержит индексы элементов временного ряда (назовём их краевыми значениями), которых достаточно для того, чтобы найти все элементы ряда при помощи ОЛРФ(A) (или, если точнее, с помощью элементов A с индексами из $\mathcal{K}(\tau)$, $A = (a_1, \ldots, a_{r+1})^{\mathrm{T}}$).

Обозначим $B_{\mathcal{C}}$ вектор, состоящий из элементов вектора B с номерами из $\mathcal{C}$.

Теорема 1. *Пусть $\mathsf{S}_0 \in \mathcal{D}_r$ — временной ряд, управляемый ОЛРФ(A_0), где $A_0 \in \mathbb{R}^{r+1}$, $a_\tau^{(0)} = -1$. Между окрестностью точки $((\mathsf{S}_0)_{\mathcal{I}(\tau)}, (A_0)_{\mathcal{K}(\tau)})^{\mathrm{T}} \in \mathbb{R}^{2r}$ и пересечением окрестности точки S_0 со множеством $\mathcal{D}_r$ единственным образом определено взаимно-однозначное отображение $\mathsf{S} : \mathbb{R}^{2r} \to \mathcal{D}_r$, которое строит по вектору параметров $(S_{(\tau)}, A_{(\tau)})^{\mathrm{T}}$ ряд S и удовлетворяет следующим условиям для $\mathsf{S} = \mathsf{S}(S_{(\tau)}, A_{(\tau)})$: $(\mathsf{S})_{\mathcal{I}(\tau)} = S_{(\tau)}$, $\mathsf{S} \in \mathcal{D}_r$ и управляется ОЛРФ(A) такой, что $A_{\mathcal{K}(\tau)} = A_{(\tau)}$ и $a_\tau = -1$.*

Теорема 2. *Пусть $\mathsf{S}_0 \in \mathcal{D}_r$ — временной ряд, управляемый ОЛРФ(A_0), где $a_\tau^{(0)} = -1$. Тогда параметризация $\mathsf{S}(S_{(\tau)}, A_{(\tau)})$, введённая в теореме 1 — гладкий диффеоморфизм между окрестностью вектора $((\mathsf{S}_0)_{\mathcal{I}(\tau)}, (A_0)_{\mathcal{K}(\tau)})^{\mathrm{T}} \in \mathbb{R}^{2r}$ и пересечением окрестности S_0 с $\mathcal{D}_r$.*

Таким образом, теорема 2 предъявляет гладкую параметризацию $\mathcal{D}_r$ в окрестности S_0, что означает, что множество $\mathcal{D}_r$ является гладким многообразием.

Далее рассмотрены производные, соответствующие этой параметризации. $\mathbf{J}_{\mathsf{S}} = \mathbf{J}_{\mathsf{S}}(S_{(\tau)}, A_{(\tau)}) \in \mathbb{R}^{N \times 2r}$ обозначает матрицу Якоби отображения $\mathsf{S}(S_{(\tau)}, A_{(\tau)})$. По определению, *касательное подпространство* к $\mathcal{D}_r$ в точке S — это colspace $\mathbf{J}_{\mathsf{S}}(S_{(\tau)}, A_{(\tau)})$, где $\mathsf{S} = \mathsf{S}(S_{(\tau)}, A_{(\tau)})$. Через A^2 обозначим ациклическую свёртку A самим с собой: $A^2 = (a_i^{(2)}) \in \mathbb{R}^{2r+1}$, $a_i^{(2)} = \sum_{j=\max(1,i-r)}^{\min(i,r+1)} a_j a_{i-j+1}$.

Теорема 3. *Касательное подпространство к $\mathcal{D}_r$ в точке S, управляемой ОЛРФ(A), имеет размерность $2r$ и равно $\mathcal{Z}(A^2)$.*

Доказано, что всё множество $\overline{\mathcal{D}_r}$ не является гладким многообразием. Найдены условия локального минимума решения задачи аппроксимации временного ряда рядом конечного ранга (3).

В конце главы найден вид ошибок первого порядка по соотношению сигнал/шум при оценивании сигнала проекцией на множество $\mathcal{D}_r$. Рассмотрим случайный временной ряд $\mathsf{S}_0+t\xi$, где сигнал $\mathsf{S}_0 \in \mathcal{D}_r$ управляется ОЛРФ(A_0), ξ — случайный вектор, имеющий распределение Ξ, t — малое вещественное число. Нас интересует распределение оценки сигнала $\widehat{\mathsf{S}}(t)$:

$$\widehat{\mathsf{S}}(t) = \Pi_{\mathcal{D}_r}^{(\mathsf{S}_0)}(\mathsf{S}_0 + t\xi) = \mathsf{S}_0 + t\mathbf{\Pi}_{\mathcal{Z}(A_0^2),\mathbf{W}}(\xi) + \zeta(t), \tag{6}$$

где проектор на линейную оболочку столбцов некоторой матрицы $\mathbf{F}$ обозначен как $\mathbf{\Pi}_{\mathbf{F},\mathbf{W}} = \mathbf{F}\left(\mathbf{F}\right)^{+}_{\mathbf{W}}$, $\Pi_{\mathcal{D}_r}^{(\mathsf{S}_0)}(\mathsf{X})$ — проектор на $\mathcal{D}_r$ в окрестности точки S_0.

Теорема 4. *Пусть $\mathsf{S}_0 \in \mathcal{D}_r$ и управляется ОЛРФ(A_0), ξ — случайный вектор, имеющий распределение Ξ. Тогда имеет место следующая слабая сходимость случайных величин: $\frac{\zeta(t)}{t} \Rightarrow_{t\to 0} 0$, где $\zeta(t)$ определена в (6).*

Таким образом, $\zeta(t)$ представляет из себя малую относительно стремящегося к нулю t случайную величину, где t задаёт соотношение сигнал/шум.

Результаты второй главы опубликованы в работе [3].

Третья глава посвящена алгоритмам решения задачи (3). Приведён вид шага метода Гаусса-Ньютона для решения задачи (3) с помощью прямой подстановки введённой во второй главе параметризации $\mathsf{S}(S_{(\tau)}, A_{(\tau)})$ в метод (5), $P = (S_{(\tau)}, A_{(\tau)})$.

Приведён вид итерации уже известного принципа Variable Projection [13] в обозначениях работы в форме, удобной для реализации и сравнения. Так как итерация Variable Projection Гаусса-Ньютона (VPGN) является неустойчивой и затратной по времени в ряде случаев, нас интересует построение иной итерации. Шаг (5) модернизируется путём замены точки $\mathsf{S}(P_k)$ на более близкую к X, что приводит к итерации следующего вида (Модифицированная итерация Гаусса-Ньютона, MGN):

$$A_{(\tau)}^{(k+1)} = A_{(\tau)}^{(k)} + \gamma \mathbf{M}_k^{+} \mathbf{Q}^{\mathrm{T}}(A^{(k)}) \mathbf{\Pi}_{\mathcal{Z}((A^{(k)})^2),\mathbf{W}}(\mathsf{X} - \mathbf{\Pi}_{\mathcal{Z}(A^{(k)}),\mathbf{W}}(\mathsf{X})), \tag{7}$$

где $\mathbf{M}_k = -\mathbf{S}^{\mathrm{T}}_{\mathcal{K}(\tau),:}$, $\mathbf{S} = \mathcal{T}_{r+1}\left(\mathbf{\Pi}_{\mathcal{Z}(A^{(k)}),\mathbf{W}}\mathsf{X}\right)$, $\mathbf{G}_{:,\mathcal{C}}$ обозначает матрицу, состоящую из столбцов произвольной матрицы $\mathbf{G}$ с индексами из $\mathcal{C}$. Ключевая особенность состоит в том, что итерация MGN сводится к вычислению проекторов на $\mathcal{Z}(A)$ и $\mathcal{Z}(A^2)$, в отличие от итерации VPGN.

С помощью использования преобразования Фурье построены алгоритмы вычисления базисов пространств временных рядов $\mathcal{Z}(A)$ и $\mathcal{Z}(A^2)$, необходимых в шаге (7), за счёт чего был достигнут выигрыш по трудоёмкости и устойчивости относительно метода VPGN. Обозначим $\mathcal{F}_N$ и $\mathcal{F}_N^{-1}$ прямое и обратное дискретное преобразование Фурье длины N. Для $\mathcal{F}_N(X) = Y$, $X, Y \in \mathbb{C}^N$, $X = (x_0, \ldots, x_{N-1})^{\mathrm{T}}$, $Y = (y_0, \ldots, y_{N-1})^{\mathrm{T}}$ выполняется $y_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j \exp(-\frac{\mathrm{i}2\pi kj}{N})$. При этом доопределим преобразование Фурье от матрицы как $\mathcal{F}_N(\mathbf{X}) = [\mathcal{F}_N(X_1) : \ldots : \mathcal{F}_N(X_r)]$, где $\mathbf{X} = [X_1 : \ldots : X_r]$, и $\mathcal{F}_N^{-1}(\mathbf{Y})$ аналогично. Также определим комплексный многочлен $g_A(z) = \sum_{k=0}^{r} a_{k+1} z^k$, у которого ведущие коэффициенты могут быть нулевыми, множество $\mathcal{W}(\alpha) = \{\omega_j^{(\alpha)},\ j = 0, \ldots, N-1\}$, где $\omega_j^{(\alpha)} = \exp\left(\mathfrak{i}\left(\frac{2\pi j}{N} - \alpha\right)\right)$, матрицу $\mathbf{T}_M(\alpha) = \mathrm{diag}((1, e^{\mathrm{i}\alpha}, \ldots, e^{\mathrm{i}(M-1)\alpha}))^{\mathrm{T}}$, E_i — единичный орт. Ниже приведён алгоритм вычисления матрицы $\mathbf{Z}(A)$, составленной из базисных векторов $\mathcal{Z}(A)$.

Алгоритм 1. ***Вход:*** *число N, вектор коэффициентов $A \in \mathbb{R}^{r+1}$.*

1: *Найти $\alpha_0 = \arg\max_{-\pi/N \le \alpha < \pi/N} \min_{z \in \mathcal{W}(\alpha)} |g_A(z)|$ с помощью одномерного численного метода оптимизации.*

2: *Вычислить вектор $A_g = (a_{g,0}, \ldots, a_{g,N-1})^{\mathrm{T}}$ как $a_{g,j} = g_A(\exp(\mathfrak{i}(\frac{2\pi j}{N} - \alpha_0))$, $j = 0, \ldots, N-1$.*

3: *Вычислить матрицы $\mathbf{R}_r = \mathcal{F}_N([E_{N-r+1} : \ldots : E_N])$, $\mathbf{L}_r = \mathbf{A}_g^{-1}\mathbf{R}_r$, где $\mathbf{A}_g = \mathrm{diag}\, A_g$.*

4: *Найти матрицу $\mathbf{U}_r \in \mathbb{C}^{N \times r}$, состоящую из ортонормированных столбцов матрицы $\mathbf{L}_r$ (например, $\mathbf{U}_r$ может состоять из r левых сингулярных векторов $\mathbf{L}_r$).*

5: *Вычислить $\widetilde{\mathbf{Z}} = \mathcal{F}_N^{-1}(\mathbf{U}_r)$.*

6: ***return*** *Матрица $\mathbf{Z} = (\mathbf{T}_N(-\alpha_0))\widetilde{\mathbf{Z}}$, столбцы которой образуют ортонормированный базис $\mathcal{Z}(A)$.*

Завершается глава сравнением методов VPGN и MGN (7) по временной сложности. Показано, что для $(2p+1)$-диагональной матрицы $\mathbf{W}$, соответствующей процессу авторегрессии порядка p (AR(p)), предложенный метод MGN требует $O(N(r+p)^2 + rN\log N)$ времени для итерации, в то время как у VPGN нет быстрой реализации. Результаты третьей главы опубликованы в работе [3].

В **четвёртой главе** рассматривается матричный подход к решению задачи (3), заключающийся в применении метода Кэдзоу к задаче (4). Доказана сходимость метода Кэдзоу по подпоследовательностям.

Получено соотношение между весами $\mathbf{W}$ в пространстве временных рядов и матрицами $\mathbf{L}$ и $\mathbf{R}$, задающими косоугольное скалярное произведение в пространстве матриц, для эквивалентности задач (3) и (4). Определим ациклическую свёртку двух векторов $C = A * B$, $A = (a_1, \ldots, a_{N_a})^{\mathrm{T}}$, $B = (b_1, \ldots, b_{N_b})^{\mathrm{T}}$, $C = (c_1, \ldots, c_{N_a+N_b-1})^{\mathrm{T}}$: $c_i = \sum\limits_{k,j:k+j-1=i} a_k b_j$. Обозначим i-ю диагональ матрицы $\mathbf{A} \in \mathbb{R}^{N\times N}$ длины $N-|i|$, состоящую из элементов с индексами (j,k), удовлетворяющих равенству $k-j=i$, как $\mathrm{diag}_{\mathbf{A}}(i)$, нулевой вектор произвольной длины M как $0_M \in \mathbb{R}^M$.

Теорема 5. $\|\mathcal{T}_L(\mathsf{Z})\|_{\mathbf{L},\mathbf{R}} = \|\mathsf{Z}\|_{\mathbf{W}}$ *для любого временного ряда* Z *тогда и только тогда, когда* $\mathbf{W} = \mathbf{L} * \mathbf{R}$, *где* i*-я диагональ* $\mathbf{L} * \mathbf{R}$ *определена как*

$$\mathrm{diag}_{\mathbf{L}*\mathbf{R}}(i) = \sum_{j+k=i} \begin{pmatrix} 0_{(|j|+|k|-|j+k|)/2} \\ \mathrm{diag}_{\mathbf{L}}(j) * \mathrm{diag}_{\mathbf{R}}(k) \\ 0_{(|j|+|k|-|j+k|)/2} \end{pmatrix},$$

$i = -N, \ldots, N$.

Далее изучен случай, когда ковариационная матрица шума $\mathbf{\Sigma} = \mathbf{\Sigma}_N$ является автоковариационной матрицей процесса $\mathrm{AR}(p)$. К сожалению, уже для диагонального случая $p = 0$ не существует таких невырожденных $\mathbf{L}$, $\mathbf{R}$, что $\mathbf{W}(\mathbf{L},\mathbf{R}) = \mathbf{L} * \mathbf{R}$ была бы равна единичной матрице [15]. Поэтому необходимо поставить задачу поиска $\mathbf{L}$ и $\mathbf{R}$, которые бы дали $\mathbf{W}(\mathbf{L},\mathbf{R})$, близкую к $\mathbf{W}_0 = \mathbf{\Sigma}_N^{-1}$.

Для упрощения поиска $\mathbf{L}$, $\mathbf{R}$ таких, что $\mathbf{W}(\mathbf{L},\mathbf{R}) \approx \mathbf{W}_0$, мы зафиксируем левую матрицу $\mathbf{L} = \mathbf{\Sigma}_L^{-1}$ равной обратной автоковариационной матрице процесса $\mathrm{AR}(p)$, с теми же коэффициентами, что и у $\mathbf{\Sigma}_N$. Задача сформулирована следующим образом:

$$\mathbf{R}^{\star}_{\mathrm{KL2}} = \mathop{\arg\min}_{\substack{\mathrm{cond}\,\mathbf{R} \le 1/\alpha \\ \mathbf{W} = \mathbf{\Sigma}_L^{-1} * \mathbf{R} \\ \mathbf{R} = \mathrm{diag}(R)}} D_{\mathrm{KL2}}(\mathbf{\Sigma}_N, \mathbf{W}^{-1}), \tag{8}$$

где $0 < \alpha \le 1$ задаёт ограничение на число обусловленности матрицы $\mathbf{R}$, D_{KL2} — разложение дивергенции Кульбака-Лейблера между многомерными нормальными распределениями с ковариационными матрицами $\mathbf{\Sigma}_N$ и $\mathbf{W}^{-1}$ в ряд Тейлора до второго порядка:

$$D_{\mathrm{KL2}}(\mathrm{N}(0, \mathbf{W}_0^{-1}) \| \mathrm{N}(0, \mathbf{W}^{-1})) = \frac{1}{2}\mathrm{tr}(\mathbf{W}\mathbf{W}_0^{-1}\mathbf{W}\mathbf{W}_0^{-1}) - \mathrm{tr}(\mathbf{W}\mathbf{W}_0^{-1}) + \frac{N}{2}.$$

Кроме того, введена более простая задача поиска весов по эвклидовой норме:

$$\mathbf{R}^{\star}_{\mathrm{dist}} = \mathop{\arg\min}_{\substack{\mathrm{cond}\,\mathbf{R} \le 1/\alpha \\ \mathbf{R} = \mathrm{diag}(R)}} \frac{1}{2}\|\mathbb{1}_L * R - \mathbb{1}_N\|^2, \tag{9}$$

где $\mathbb{1}_M$ обозначает вектор из M единиц.

Теорема 6. *Если или $p = 0$, или $p > 0$ и $L \geq K + p - 1$, то задачи* (8) *и* (9) *эквивалентны, то есть минимум функционалов достигается на одной и той же* $\mathbf{R}$.

Построена теория эффективного решения задачи (9) с использованием метода квадратичного программирования и с помощью оптимизации гладкой функции в параллелепипеде. Предложена быстрая реализация метода Кэдзоу с временной сложностью итерации $O(r(N \log N + Lp_1 + Kp_2) + (p_1 + p_2)N)$, где матрица $\mathbf{L}$ — $(2p_1 + 1)$-диагональная, $\mathbf{R}$ — $(2p_2 + 1)$-диагональная, $N = L + K - 1$.

В конце главы найден вид ошибок первого порядка при использовании оценки сигнала линеаризацией метода Кэдзоу. Рассмотрим случайный временной ряд $\mathsf{X} = \mathsf{S}_0 + \xi$, где сигнал $\mathsf{S}_0 \in \mathcal{D}_r$, ξ — случайный вектор с распределением Ξ. Определим матрицу оператора, задающего один шаг линеаризованного метода Кэдзоу, как

$$\mathbf{P}_{\mathsf{S}_0} = \mathcal{T}_L^{-1} \Pi_{\mathcal{H}} \Pi_{\mathcal{N}(\mathcal{T}_L(\mathsf{S}_0))} \mathcal{T}_L, \tag{10}$$

где $\mathcal{N}(\mathcal{T}_L(\mathsf{S}_0))$ обозначает касательное подпространство к пространству $\mathcal{M}_r$ матриц ранга r в точке $\mathcal{T}_L(\mathsf{S}_0)$.

Лемма 1. *Пусть $\mathsf{S}_0 \in \mathcal{D}_r$ и управляется ОЛРФ(A_0), ξ — случайный вектор с распределением Ξ. Рассмотрим оценку сигнала $\widetilde{\mathsf{S}}_n = \mathbf{P}_{\mathsf{S}_0}^n(\mathsf{S}_0 + \xi)$, полученную линеаризованным алгоритмом Кэдзоу в модели $\mathsf{X} = \mathsf{S}_0 + \xi$, $\mathbf{P}_{\mathsf{S}_0}^n$ определена в* (10)*. Тогда имеет место следующая слабая сходимость распределений:*

$$\widetilde{\mathsf{S}}_n \Rightarrow_{n \to +\infty} \mathsf{S}_0 + \mathbf{\Pi}_{\mathcal{Z}(A_0^2), \mathbf{W}(\mathbf{L}, \mathbf{R})} \xi. \tag{11}$$

Результаты четвёртой главы опубликованы в работах [1, 2, 4].

В **пятой главе** проведены численные эксперименты. Показано, что построенная итерация метода Гаусса-Ньютона MGN (7) обладает большей устойчивостью, чем шаг VPGN при вычислении точки локального минимума. С помощью статистического моделирования подтверждены результаты о величине ошибок первого порядка для проекции на множество временных рядов ранга r и линеаризованного метода Кэдзоу. Приведены примеры применения метода модифицированного метода Гаусса-Ньютона MGN к данным экспрессии генов и к временному ряду с неизвестной ковариационной матрицей шума.

В **Заключении** подведены основные итоги диссертационной работы. Отмечено, что в диссертации были предложены новые методы для решения широко используемых на практике задач оценки сигнала и его параметров в классе сигналов конечного ранга. Перечислены ключевые результаты. А именно, указано, что получено теоретическое обоснование модифицированного метода Гаусса-Ньютона и метода Кэдзоу для решения задач (3) и (4), а также их свойств, построены эффективные (быстрые и устойчивые) реализации. Это позволяет существенно расширить круг решаемых проблем на случай авторегрессионного шума и улучшить точность их решения.

Список публикаций

1. Звонарев Н. К. Поиск весов в задаче взвешенной аппроксимации временным рядом конечного ранга // Вестник Санкт-Петербургского университета. Серия 1. Математика. Механика. Астрономия. 2016. Т. 3, № 4.
2. Zvonarev N., Golyandina N. Iterative algorithms for weighted and unweighted finite-rank time-series approximations // Statistics and Its Interface. 2017. Vol. 10, no. 1. P. 5–18.
3. Zvonarev N., Golyandina N. Modified Gauss-Newthon method in low-rank signal estimation. arXiv:1803.01419.
4. Звонарев Н., Голяндина Н. Итеративные алгоритмы взвешенной аппроксимации рядами конечного ранга // System Identification And Control Problems. SICPRO'15. 2015. P. 1371–1394.

Цитированная литература

5. Cadzow J. A. Signal enhancement: a composite property mapping algorithm // IEEE Trans. Acoust. 1988. Vol. 36, no. 1. P. 49–62.
6. Markovsky I. Structured low-rank approximation and its applications // Automatica. 2008. apr. Vol. 44, no. 4. P. 891–909.
7. Dendrinos M., Bakamidis S., Carayannis G. Speech enhancement from noise: A regenerative approach // Speech Communication. 1991. feb. Vol. 10, no. 1. P. 45–57.
8. Golyandina N., Nekrutkin V., Zhigljavsky A. Analysis of Time Series Structure: SSA and Related Techniques. Chapman&Hall/CRC, 2001.
9. Tufts D., Shah A. Estimation of a signal waveform from noisy data using low-rank approximation to a data matrix // IEEE Transactions on Signal Processing. 1993. apr. Vol. 41, no. 4. P. 1716–1721.

10. Markovsky I. Low Rank Approximation: Algorithms, Implementation, Applications (Communications and Control Engineering). Springer, 2011.
11. Ottaviani G., Spaenlehauer P.-J., Sturmfels B. Exact solutions in structured low-rank approximation // SIAM Journal on Matrix Analysis and Applications. 2014. Vol. 35, no. 4. P. 1521–1542.
12. Usevich K., Markovsky I. Structured low-rank approximation as a rational function minimization // IFAC Proceedings Volumes. 2012. Vol. 45, no. 16. P. 722–727.
13. Usevich K., Markovsky I. Variable projection for affinely structured low-rank approximation in weighted 2-norms // Journal of Computational and Applied Mathematics. 2014. Vol. 272. P. 430–448.
14. Gillard J., Zhigljavsky A. Stochastic methods for Hankel structured low rank approximation // Proceedings of 21th International Symposium on Mathematical Theory of Networks and Systems. 2014. P. 961–964.
15. Zhigljavsky A., Golyandina N., Gryaznov S. Deconvolution of a discrete uniform distribution // Stat Probabil Lett. 2016. Vol. 118. P. 37–44.
16. Kukush A., Markovsky I., Van Huffel S. Consistency of the structured total least squares estimator in a multivariate errors-in-variables model // Journal of Statistical Planning and Inference. 2005. Vol. 133, no. 2. P. 315–358.
17. The element-wise weighted total least-squares problem / Ivan Markovsky, Maria Luisa Rastello, Amedeo Premoli et al. // Computational Statistics & Data Analysis. 2006. Vol. 50, no. 1. P. 181–209.
18. De Moor B. Total least squares for affinely structured matrices and the noisy realization problem // IEEE Transactions on Signal Processing. 1994. Vol. 42, no. 11. P. 3104–3113.