

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

На правах рукописи

Звонарев Никита Константинович

Структурные аппроксимации временных рядов

01.01.07 – Вычислительная математика

ДИССЕРТАЦИЯ

на соискание ученой степени

кандидата физико-математических наук

Научный руководитель — кандидат
физико-математических наук, доцент

Голяндина Нина Эдуардовна

Санкт-Петербург – 2018

Оглавление

Введение	4
Используемые обозначения	15
Глава 1. Сведения из теории временных рядов конечного ранга и теории оптимизации	18
1.1. Линейные рекуррентные формулы	18
1.2. Независимость ранга временного ряда от длины окна L	21
1.3. Методы оптимизации для нелинейной задачи наименьших квадратов	22
1.4. Непараметрический метод Кэдзоу	25
1.5. Метод квадратичного программирования «Active Set»	26
Глава 2. Свойства задачи аппроксимации временных рядов рядами конечного ранга	28
2.1. Необходимые свойства $\mathcal{D}_r$ и $\overline{\mathcal{D}_r}$	28
2.2. Параметризация множества $\mathcal{D}_r$	30
2.3. Дополнительные свойства $\overline{\mathcal{D}_r}$	37
2.4. Условия минимума в задаче аппроксимации рядами конечного ранга	38
2.5. Задача аппроксимации как задача оценивания сигнала	41
Глава 3. Численные методы решения задачи аппроксимации временных рядов рядами конечного ранга	47
3.1. Методы локальной оптимизации	47
3.2. Вычисление базисов пространств $\mathcal{Z}(A)$ и $\mathcal{Z}(A^2)$	55
3.3. Алгоритмы решения задачи	66

Глава 4. Матричный подход к задаче аппроксимации временного ряда рядами конечного ранга (метод Кэдзоу)	73
4.1. Сходимость метода Кэдзоу по подпоследовательностям	74
4.2. Постановка задачи поиска весов	78
4.3. Применение квадратичного программирования к поиску весов . .	86
4.4. Поиск весов с помощью минимизации гладкой функции	96
4.5. Быстрая реализация алгоритма Кэдзоу	98
4.6. Применение метода Кэдзоу к решению задачи аппроксимации временных рядов рядами конечного ранга	103
4.7. Применение алгоритма Кэдзоу к задаче оценивания сигнала . . .	106
 Глава 5. Результаты численных экспериментов по устойчивости и применимости модифицированного метода Гаусса-Ньютона и метода Кэдзоу	 116
5.1. Исследование устойчивости алгоритмов	116
5.2. Исследование свойств оценок сигнала с помощью статистического моделирования	123
5.3. Применение модифицированного метода Гаусса-Ньютона к данным экспрессии генов	131
5.4. Аппроксимация временными рядами конечного ранга при неизвестных параметрах авторегрессионного шума	137
 Заключение	 144
 Список литературы	 146

Введение

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 $\mathbf{X} = (x_1, \dots, x_N)^T \in \mathbb{R}^N$.

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

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

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

Рассмотрим следующий параметрический вид сигнала $\mathbf{S}$ в виде конечной суммы:

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

где $p_{m_k}(n)$ — многочлены от n степени m_k . Такой вид сигнала используется во многих приложениях, например, в теории обработки сигнала [1], в задачах идентификации линейных систем [2], задачах распознавания речи [3] и многих других. В приложениях к теории обработки сигнала часто считается, что сигнал

$\mathbf{S}$ в представлении (1) является суммой синусоид [1] или экспоненциально-модулированных синусоид [2].

Важной задачей обработки временных сигналов является оценка сигнала $\mathbf{S}$ по ряду наблюдений $\mathbf{X}$. Полученную оценку сигнала можно использовать для оценивания параметров сигнала [4, 5, 6, 7], построения оценки прогноза сигнала [8, 9, 10] и разложения сигнала на аддитивные составляющие [8, 5, 11].

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* $\mathbf{S}$, определённый следующим способом: для заданного натурального числа L , $1 < L < N$, называемого *длиной окна*, определим *оператор вложения* $\mathcal{T}_L : \mathbb{R}^N \rightarrow \mathbb{R}^{L \times (N-L+1)}$ как

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

Скажем, что ряд $\mathbf{S}$ имеет ранг r , если $\text{rank } \mathcal{T}_L(\mathbf{S}) = r < N/2$ 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 $\mathcal{T}_L(\mathbf{S})$ называется *траекторной матрицей ряда* $\mathbf{S}$.

Матрица $\mathcal{T}_L(\mathbf{S})$ является ганкелевой, то есть элементы на её побочных диагоналях равны. Перестановкой строк или столбцов матрицы $\mathcal{T}_L(\mathbf{S})$ в обратном порядке можно добиться равенства элементов на её диагоналях, то есть привести её к эквивалентному тёплицевому виду. Модель данных, где соответствующая сигналу $\mathbf{S}$ ганкелева матрица $\mathcal{T}_L(\mathbf{S})$ имеет конечный ранг, встречается в теории обработки сигналов [1, 12], задачах распознавания речи [3], теории управления и теории стохастических систем [2, 13]. Выбор значения ранга сигнала — отдельная задача и в данной работе не рассматривается. В дальнейшем

считаем ранг ряда r известным.

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

$$\mathbf{Y}^* = \arg \min_{\mathbf{Y} \in \overline{\mathcal{D}_r}} \|\mathbf{X} - \mathbf{Y}\|_{\mathbf{W}}^2, \quad (3)$$

где $\|\mathbf{Z}\|_{\mathbf{W}}^2 = \mathbf{Z}^T \mathbf{W} \mathbf{Z}$ — 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) [13, 14]. Для определённости, будем называть (3) задачей HSLRA в векторном виде.

Если ϵ — гауссовский шум с нулевым средним и ковариационной матрицей Σ , то решение задачи (3) с матрицей весов $\mathbf{W} = \Sigma^{-1}$ является оценкой максимального правдоподобия (ОМП) сигнала $\mathbf{S}$. Заметим, что для того чтобы оценка была ОМП, достаточно знать матрицу Σ 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 [15], следовательно, может содержать множество локальных минимумов. Для решения такой задачи можно либо использовать методы глобального поиска, либо локальный поиск с выбором достаточно близкого к оптимальному начального приближения. В работе разрабатываются методы локального поиска; также рассматривается и вопрос выбора начального приближения.

Для решения задачи (3) применяются численные методы. Для их построения используются два основных подхода. Первый — так называемый принцип *Variable projection*, рассмотренный в работах [16, 17] для более общего, чем в (3), случая произвольной аффинной структуры. В [17] после применения принципа

Variable projection используются методы Гаусса-Ньютона и Левенберга-Марквардта (регуляризованная версия метода Гаусса-Ньютона), которые являются методами локального поиска, но основаны на параметризации, отличной 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. Вторым недостатком — методы проявляют неустойчивость в ряде случаев, например, когда $\mathbf{S}$ является полиномиальным сигналом.

Второй подход — непараметрический метод итераций Кэдзоу, входящий в класс алгоритмов попеременных проекций (Alternating Projections) [1]. Метод решает задачу HSLRA в матричном виде:

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

$\|\mathbf{X}\|_{\mathbf{L}, \mathbf{R}}$ — порождённая скалярным произведением

$$\langle \mathbf{X}, \mathbf{Y} \rangle_{\mathbf{L}, \mathbf{R}} = \text{tr}(\mathbf{L} \mathbf{X} \mathbf{R} \mathbf{Y}^T) \quad (5)$$

матричная норма [18], $\mathcal{H} = \mathcal{H}_{L, K}$ — пространство ганкелевых матриц размера $L \times K$, $\mathcal{M}_r \subset \mathbb{R}^{L \times K}$ — множество матриц ранга, не превосходящего r , $\mathbf{L} \in \mathbb{R}^{L \times L}$, $\mathbf{R} \in \mathbb{R}^{K \times K}$ — положительно определённые матрицы весов, а $\mathbf{X}$, $\mathbf{Y}$ — матрицы, связанные с задачей (3) соотношениями $\mathbf{X} = \mathcal{T}_L(\mathbf{X})$ и $\mathbf{Y} = \mathcal{T}_L(\mathbf{Y})$.

Теория метода Кэдзоу тесно связана с теорией так называемых subspace-based методов и метода SSA (Singular Spectrum Analysis, анализ сингулярного спектра) [19, 20, 21, 8]. Метод можно распространить на случай косоугольной нормы, отличной от евклидовой [22]. 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 [23]). Вопрос выбора подходящих матриц $\mathbf{L}$, $\mathbf{R}$, дающих матрицу весов $\mathbf{W}$, близкую к Σ^{-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 [12, 24, 25]. Однако полученные в этих работах утверждения либо не учитывают вид матрицы $\mathbf{W}$, либо сильно ограничивают их множество. В работах [24, 25] исследована матричная постановка задачи (4), но в работе [24] матрицы $\mathbf{L}$, $\mathbf{R}$ фиксированы и равны единичным, в работе [25] лишь одна из двух матриц $\mathbf{L}$ или $\mathbf{R}$ произвольна. Для алгоритма Кэдзоу показана неоптимальность полученных им оценок [12, 26], но нет результата о том, насколько отсутствие оптимальности ухудшает качество полученной оценки, и можно ли снизить уровень «неоптимальности» путём выбора весов $\mathbf{L}$, $\mathbf{R}$.

В этой работе 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 [22] методе стохастической оптимизации.

Цели 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 [17], 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 недиагональных матриц весов $\mathbf{L}$ и $\mathbf{R}$.
4. Исследование асимптотических по соотношению сигнал/шум ошибок первого порядка для полученных методами оценок сигнала.

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 найдена гладкая параметризация и вид касательного подпространства, необходимый для построения методов локальной оптимизации.
2. Разработан метод вычисления базисов рядов конечного ранга, 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 [17] по скорости в случае ленточной матрицы весов $\mathbf{W}$ и по точности на полиномиальных сигналах.
4. Сформулирована задача поиска весов $\mathbf{L}$, $\mathbf{R}$ для метода Кэдзоу, 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 $\mathbf{L}$ и $\mathbf{R}$.
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.

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 $\mathbf{W}$. 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.

Апробация работы

Основные результаты обсуждались на семинарах кафедры статистического моделирования математико-механического факультета СПбГУ, семинаре кафедры статистики в School of Mathematics, Cardiff University (Великобритания, июнь 2017) и докладывались 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 [27] в научном издании, включенном в Перечень рецензируемых научных изданий, рекомендуемых ВАК. Работа [28] опубликована в научном издании, входящем в базы цитирования Web of Science и Scopus. Работа [27], 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 [29, 28, 30] постановка задачи, структура работы и введение принадлежат научному руководителю, основной текст написан совместно, а основные результаты получены соискателем. В частности, в работе [28] соискателю принадлежат основные теоретические результаты, в том числе теорема 1 о сходимости метода Кэдзоу по подпоследовательностям, а также численное моделирование оценки сигнала с помощью метода Кэдзоу. В [30] 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.

Основное содержание

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.

В **первой главе** приведён обзор существующих методов и свойств решения задачи аппроксимации (3) временного ряда рядом конечного ранга. Выписаны известные свойства множества $\mathcal{D}_r$ рядов ранга r . Приведён известный метод локальной оптимизации Гаусса-Ньютона, принцип Variable projection, а также применение принципа Variable projection к решению задачи (3). Описан непараметрический метод Кэдзоу для решения задачи (4) и его свойства. Выписан метод Active set решения задачи квадратичного программирования, использующийся в задаче поиска весов для метода Кэдзоу.

Вторая глава посвящена свойствам задачи (3). Построена параметризация множества $\mathcal{D}_r$, необходимая для построения методов локальной оптими-

зации. Получен параметрический вид касательного подпространства в точке, лежащей в $\mathcal{D}_r$. 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). Найден вид ошибок первого порядка по соотношению сигнал/шум при оценивании сигнала проекцией на множество $\mathcal{D}_r$, указана связь с границей Рао-Крамера.

Третья глава посвящена алгоритмам решения задачи (3). Приведён вид шага метода Гаусса-Ньютона для решения задачи (3) с помощью известного метода Variable projection (VPGN) в обозначениях работы в форме, удобной для реализации и сравнения. С помощью построенной в главе 2 параметризации введён модифицированный шаг Гаусса-Ньютона (MGN). Найдено различие между шагом метода VPGN и шагом метода MGN. Построены алгоритмы вычисления базисов рядов, использующихся в методе MGN, указан порядок обусловленности при их вычислении. Приведено сравнение методов VPGN и MGN по временной сложности при различных видах матрицы $\mathbf{W}$.

В **четвёртой главе** рассматривается матричный подход к решению задачи (3), заключающийся в применении метода Кэдзоу к задаче (4). Построено доказательство сходимости метода Кэдзоу по подпоследовательностям. Доказано соотношение между весами $\mathbf{W}$ в пространстве рядов, и матрицами $\mathbf{L}$ и $\mathbf{R}$, задающими косоугольное скалярное произведение в пространстве матриц, для эквивалентности задач (3) и (4). Сформулирована задача поиска положительно определённых матриц весов $\mathbf{L}$ и $\mathbf{R}$, дающих близкие к $\mathbf{W}_0 = \mathbf{\Sigma}^{-1}$ веса, доказана эквивалентность сформулированной задачи более простой задаче поиска весов по евклидовой метрике при некоторых условиях. Построена теория эффективного решения задачи поиска весов с использованием метода квадратичного программирования и с помощью оптимизации гладкой функции в параллелепипеде. Создана быстрая реализация метода Кэдзоу, указана её временная сложность. Найден вид ошибок первого порядка по соотношению сигнал/шум

при использовании оценки сигнала $\mathbf{S}$ методом Кэдзоу.

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

Используемые обозначения

$\mathbb{R}$ Множество вещественных чисел

$\mathbb{C}$ Множество комплексных чисел

i Мнимая единица

S, X, Y Временные ряды

$\mathcal{C}, \mathcal{Z}, \mathcal{D}$ Множества

$\emptyset$ Пустое множество

$\overline{\mathcal{D}}$ Замыкание множества $\mathcal{D}$

X, Y Матрицы

X, Y Векторы

$\mathcal{T}_L : \mathbb{R}^N \rightarrow \mathbb{R}^{L \times (N-L+1)}$ Оператор вложения

$\mathcal{V} : \mathbb{R}^{L \times K} \rightarrow \mathbb{R}^{LK}$ Оператор векторизации

ξ, η Случайные векторы

Ξ, H Распределения случайных векторов

$\xi(t) \Rightarrow_{t \rightarrow 0} \eta$ Слабая сходимость (распределений) случайных векторов

$\mathbb{E}\xi$ Математическое ожидание случайного вектора ξ

$\mathbb{P}(\|\xi\| < 1)$ Вероятность события

$\text{cov}(\xi)$ Ковариационная матрица случайного вектора ξ

$E_i \in \mathbb{R}^N$ i -й орт

$\mathbf{1}_N \in \mathbb{R}^N$ Вектор из N единиц

$\mathbf{0}_{L,K} \in \mathbb{R}^{L \times K}$ Нулевая матрица

$\mathbf{0}_N \in \mathbb{R}^N$ Нулевой вектор

$\mathbf{o}_{Z \rightarrow 0}(f(\mathbf{Z}))$ «о» малое от функции f при $\mathbf{Z} \rightarrow 0$

$\mathbf{e}_{i,j} \in \mathbb{R}^{L \times K}$ (i, j) -й матричный орт с 1 на позиции (i, j)

$\mathbf{F}^T$ Транспонирование матрицы $\mathbf{F}$

$\mathbf{I}_N \in \mathbb{R}^{N \times N}$ Единичная матрица размера $N \times N$

$\text{rank } \mathbf{F}$ Ранг матрицы $\mathbf{F}$

$\text{colspace } \mathbf{F}$ Пространство столбцов матрицы $\mathbf{F}$

$\text{rowspace } \mathbf{F}$ Пространство строк матрицы $\mathbf{F}$

$A_{\mathcal{C}}$ Вектор с индексами из множества $\mathcal{C}$

$\text{span}(V_1, \dots, V_n)$ Линейная оболочка векторов $V_1, \dots, V_n$

$\text{orthonorm}(\mathbf{F})$ Ортонормализация матрицы $\mathbf{F}$ по столбцам

$\mathbf{G}_{\mathcal{C},:}$ Матрица из строк матрицы $\mathbf{G}$ с индексами из множества $\mathcal{C}$

$\mathbf{G}_{:,\mathcal{C}}$ Матрица из столбцов матрицы $\mathbf{G}$ с индексами из множества $\mathcal{C}$

$(\mathbf{F})_{\mathbf{W}}^{\dagger} = (\mathbf{F}^T \mathbf{W} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{W}$ Взвешенная псевдообратная матрица

$\mathbf{F}^{\dagger} = (\mathbf{F})_{\mathbf{I}}^{\dagger}$ Псевдообратная матрица

$\Pi_{\mathbf{F}, \mathbf{W}} = \mathbf{F} (\mathbf{F})_{\mathbf{W}}^{\dagger}$ Взвешенный проектор на пространство $\text{colspace } \mathbf{F}$

$\Pi_{\mathcal{Z}, \mathbf{W}} = \Pi_{\mathbf{Z}, \mathbf{W}}$ Взвешенный проектор на линейное подпространство $\mathcal{Z}$, где матрица $\mathbf{Z}$ содержит произвольный базис $\mathcal{Z}$

$\Pi_{\mathcal{D}}$ Оператор проектирования на множество $\mathcal{D}$

$\mathcal{K}_i$ i -й элемент упорядоченного множества $\mathcal{K}$

$\mathcal{F}_N, \mathcal{F}_N^{-1}$ Прямое и обратное дискретное преобразование Фурье

$X * Y$ Ациклическая свёртка двух векторов

$A^2 = A * A$ Ациклическая свёртка вектора с самим собой

$\mathbf{X} * \mathbf{Y}$ Ациклическая свёртка двух матриц

$\text{cond } \mathbf{A}$ Число обусловленности матрицы $\mathbf{A}$

$\text{tr } \mathbf{A}$ След матрицы $\mathbf{A}$

$\text{diag}(\mathbf{A})$ Главная диагональ матрицы $\mathbf{A}$

$\text{diag}_{\mathbf{A}}(i)$ i -я диагональ матрицы $\mathbf{A}$

$\text{diag}(R)$ Диагональная матрица с вектором R на главной диагонали

$\langle X, Y \rangle_{\mathbf{W}} = X^T \mathbf{W} Y$ Скалярное произведение векторов X, Y по косоугольной евклидовой норме с весами $\mathbf{W}$

$\|X\|_{\mathbf{W}} = \sqrt{\langle X, X \rangle_{\mathbf{W}}}$ Косоугольная евклидова норма

$\langle \mathbf{X}, \mathbf{Y} \rangle_{\mathbf{L}, \mathbf{R}} = \text{tr}(\mathbf{L} \mathbf{X} \mathbf{R} \mathbf{Y}^T)$ Скалярное произведение матриц $\mathbf{X}, \mathbf{Y}$ по взвешенной матричной норме с матрицами весов $\mathbf{L}, \mathbf{R}$

$\langle \mathbf{X}, \mathbf{Y} \rangle_{\text{F}} = \langle \mathbf{X}, \mathbf{Y} \rangle_{\mathbf{I}_L, \mathbf{I}_K}$ Фробениусово скалярное произведение матриц $\mathbf{X}, \mathbf{Y}$

$\|\mathbf{X}\|_{\mathbf{L}, \mathbf{R}} = \sqrt{\langle \mathbf{X}, \mathbf{X} \rangle_{\mathbf{L}, \mathbf{R}}}$ Взвешенная матричная норма

Глава 1

Сведения из теории временных рядов конечного ранга и теории оптимизации

1.1. Линейные рекуррентные формулы

Основным объектом работы является временной ряд $\mathbf{S} \in \mathbb{R}^N$ — вектор-столбец длины N . В дальнейшем, под термином «ряд» будем понимать именно временной ряд. Формально определим упомянутые во введении объекты.

Определение 1.1.1. Ряд $\mathbf{S}$ имеет ранг r , если $\text{rank } \mathcal{T}_L(\mathbf{S}) = r < N/2$ для любого L такого, что $\min(L, N - L + 1) > r$, где $\mathcal{T}_L$ — оператор вложения, определённый в (2).

Определение 1.1.2. Множество рядов ранга r :

$$\mathcal{D}_{r,N} = \mathcal{D}_r = \{\mathbf{X} \in \mathbb{R}^N : \text{rank}(\mathcal{T}_{r+1}(\mathbf{X})) = r\},$$

где $N > 2r$.

$\overline{\mathcal{D}_r}$ — замыкание множества $\mathcal{D}_r$ по евклидовой норме.

Хорошо известно [31], что любой временной ряд вида (1) при достаточно большой длине ряда управляется *линейной рекуррентной формулой* (ЛРФ) некоторого порядка m :

$$s_i = \sum_{k=1}^m b_k s_{i-k}, \quad b_m \neq 0, \quad i = 1, \dots, N - m. \quad (1.1)$$

Один и тот же ряд может управляться многими различными ЛРФ разного порядка. ЛРФ минимального порядка r (при том единственная) называется минимальной, при этом соответствующий ей временной ряд имеет ранг r . Минимальная ЛРФ вместе с начальными значениями ряда однозначно определяет вид и параметры модели (1).

Равенства (1.1) можно переписать в векторном виде как $A^T \mathcal{T}_{r+1}(\mathbf{S}) = 0_{N-r}^T$, где $A = (b_r, \dots, b_1, -1)^T \in \mathbb{R}^{r+1}$. Вектор A , соответствующий минимальной ЛРФ, и первые r значений ряда $\mathbf{S}$ однозначно задают полный временной ряд $\mathbf{S}$. Следовательно, r коэффициентов ЛРФ порядка r и первые r значений ряда (вместе $2r$ параметров) можно выбрать как параметры ряда ранга r .

Однако эта параметризация не покрывает всё множество $\mathcal{D}_r$ [8, Theorem 5.1].

Давайте обобщим понятие ЛРФ.

Определение 1.1.3. Ряд $\mathbf{S}$ управляется обобщённой ЛРФ (ОЛРФ) порядка r , если $A^T \mathcal{T}_{r+1}(\mathbf{S}) = 0_{N-r}^T$ для некоторой $A \in \mathbb{R}^{r+1}$, $A \neq 0_{r+1}$.

Также мы рассмотрим ОЛРФ минимального порядка. Ключевая разница между ОЛРФ и обычной ЛРФ состоит в том, что последний коэффициент ОЛРФ не обязательно ненулевой, соответственно, формула не является «рекуррентной» в прямом смысле, она не обязана выражать последний элемент через предыдущие. Тем не менее, по крайней мере один коэффициент ОЛРФ должен быть ненулевым. Идея использовать произвольный ненулевой вектор A для задания соотношения, которому удовлетворяет ряд, используется в работах [16, 17].

Определение 1.1.4. *Характеристический многочлен* ОЛРФ(A) порядка r — это комплексный многочлен $g_A(z) = \sum_{k=0}^r a_{k+1} z^k$, у которого ведущие коэффициенты могут быть нулевыми [8].

Сигнальные корни ряда $\mathbf{S}$ — корни многочлена $g_A(z)$, соответствующего минимальной ОЛРФ(A), которая управляет рядом $\mathbf{S}$.

Следующая теорема предъявляет эквивалентный вид ряда, управляемого ОЛРФ(A) с ненулевыми значениями на краях вектора A (что эквивалентно ЛРФ с $b_r \neq 0$).

Теорема 1.1.1. [8, Theorem 5.3] Вещественный ряд $S = (s_1, \dots, s_N)^T$ управляется ОЛРФ(A), $A = (a_1, \dots, a_{r+1})^T$ с $a_1 \neq 0$, $a_{r+1} \neq 0$ тогда и только тогда, когда S имеет следующий вид:

$$s_n = \sum_{i=1}^k p_i(n) z_i^n,$$

где $z_1, \dots, z_k$ — сигнальные корни ряда S с кратностями $t_1, \dots, t_k$, $\sum_{i=1}^k t_k = r$, $p_i(n)$ — комплексный многочлен степени t_i , при этом коэффициенты многочленов $p_1, \dots, p_k$ задаются первыми r значениями ряда $s_1, \dots, s_r$.

Введём следующее определение, необходимое для дальнейшей работы.

Определение 1.1.5. $\mathcal{Z}_N(A) = \mathcal{Z}(A) = \{S \in \mathbb{R}^N : A^T \mathcal{T}_{r+1}(S) = 0_{N-r}^T\}$ — подпространство временных рядов, управляемых ОЛРФ(A), $A \in \mathbb{R}^{r+1}$.

Рассмотрим эквивалентный вид $\mathcal{Z}(A)$. Под $\mathbf{Q}^{M,d}$ будем обозначать оператор $\mathbb{R}^{d+1} \rightarrow \mathbb{R}^{M \times (M-d)}$, определённый следующим образом:

$$\mathbf{Q}^{M,d}(B) = \begin{pmatrix} b_1 & 0 & \dots & 0 \\ b_2 & b_1 & 0 & \vdots \\ \vdots & b_2 & \ddots & 0 \\ b_{d+1} & \vdots & \ddots & b_1 \\ 0 & b_{d+1} & \ddots & b_2 \\ \vdots & 0 & \ddots & \vdots \\ 0 & \dots & 0 & b_{d+1} \end{pmatrix}, \quad (1.2)$$

где $B = (b_1, \dots, b_{d+1})^T \in \mathbb{R}^{d+1}$ [32]. Тогда $\mathcal{Z}_N(A) = \mathcal{Z}(A)$ можно представить в другом виде как

$$\mathcal{Z}(A) = \{S \in \mathbb{R}^N : \mathbf{Q}^T(A)S = 0_{N-r}\}, \quad (1.3)$$

где $\mathbf{Q} = \mathbf{Q}^{N,r}$.

Будем использовать следующие обозначения: $\mathcal{Q}(A) = \text{colspace}(\mathbf{Q}(A))$, и под $\mathbf{Z}(A)$ мы определяем некоторую матрицу, столбцы которой составляют базис $\mathcal{Z}(A)$.

1.2. Независимость ранга временного ряда от длины окна

L

Для построения параметризации множества $\mathcal{D}_r$ необходимо указать несколько лемм, устанавливающие свойства множества $\mathcal{D}_r$ рядов ранга r и его замыкания $\overline{\mathcal{D}_r}$.

Предложение 1.2.1. *Временной ряд $\mathbf{S}$ имеет ранг r тогда и только тогда, когда существует L , $\min(L, N - L + 1) > r$, такое, что $\text{rank } \mathcal{T}_L(\mathbf{S}) = r$.*

Доказательство. Это следствие из [33, Prop. 5.4]. □

Следствие 1.2.1. *Временной ряд $\mathbf{S}$ имеет ранг r тогда и только тогда, когда $\text{rank } \mathcal{T}_{r+1}(\mathbf{S}) = r$.*

Замечание 1.2.1. *Если временной ряд $\mathbf{S}$ имеет ранг r , то $\text{rank } \mathcal{T}_{r^*}(\mathbf{S}) = r^*$ для любого $r^* \leq r$.*

В дальнейшем, когда мы говорим о временном ряде $\mathbf{S} \in \mathcal{D}_r$ и управляющей ей ОЛРФ(A), мы имеем в виду минимальную ОЛРФ, то есть $A \in \mathbb{R}^{r+1}$.

Таким образом, указанные свойства устанавливают независимость понятия ранга ряда от длины окна L при достаточно больших значениях N и $N - L + 1$.

1.3. Методы оптимизации для нелинейной задачи наименьших квадратов

1.3.1. Взвешенный метод Гаусса-Ньютона

Задача (3) взвешенного метода наименьших квадратов (МНК) с матрицей весов $\mathbf{W}$ может быть переписана как общая задача вида

$$\arg \min_P \|X - S(P)\|_{\mathbf{W}}^2, \quad (1.4)$$

где $P \in \mathbb{R}^p$ — вектор параметров, $S : \mathbb{R}^p \rightarrow \mathbb{R}^N$ — некоторая параметризация подмножества в $\mathbb{R}^N$, в котором ищется ближайшая к $X \in \mathbb{R}^N$ точка, при этом $S(P)$ — дифференцируемая по P функция, $\mathbf{W} \in \mathbb{R}^{N \times N}$ — симметричная положительно определённая матрица.

Если задача (1.4) нелинейна, то часто используются итеративные методы с линеаризацией на каждом шаге.

Определим взвешенную псевдообратную матрицу к $\mathbf{F}$ [34]:

$$(\mathbf{F})_{\mathbf{W}}^{\dagger} = (\mathbf{F}^T \mathbf{W} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{W}, \quad (1.5)$$

которая возникает при решении линейной задачи взвешенных наименьших квадратов $\min_P \|\mathbf{Y} - \mathbf{F}P\|_{\mathbf{W}}^2$ с $\mathbf{F} \in \mathbb{R}^{N \times p}$, так как её решением является $P^* = (\mathbf{F})_{\mathbf{W}}^{\dagger} \mathbf{Y}$. В частном случае $\mathbf{W} = \mathbf{I}_N$ (то есть случае обычной псевдообратной матрицы) $(\mathbf{F})_{\mathbf{I}_N}^{\dagger}$ будем обозначать просто как $\mathbf{F}^{\dagger}$.

Заодно будем обозначать проектор на линейную оболочку столбцов матрицы $\mathbf{F}$ как $\mathbf{\Pi}_{\mathbf{F}, \mathbf{W}} = \mathbf{F} (\mathbf{F})_{\mathbf{W}}^{\dagger}$.

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

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

где $\mathbf{J}_S(P_k)$ — матрица Якоби вектор-функции $S(P)$ в точке P_k .

Выбор шага γ — отдельная задача. Например, можно начать с $\gamma = 1$, и уменьшать шаг, если следующее значение хуже текущего (то есть значение функционала увеличивается).

Подробно теория метода Гаусса-Ньютона дана в [35].

В нашем случае наибольший интерес в задаче взвешенного метода наименьших квадратов представляет нахождение $S(P^*)$, где P^* — решение (1.4). Рассмотрим метод как последовательность $S(P_k)$, и мы получаем:

$$S(P_{k+1}) = S(P_k) + \gamma (\mathbf{J}_S(P_k))_{\mathbf{W}}^{\dagger} (X - S(P_k)). \quad (1.6)$$

Следующее замечание объясняет подход, который будет использоваться в главе 3 в предложенном там модифицированном методе Гаусса-Ньютона.

Замечание 1.3.1. Мы можем рассмотреть модификацию метода (1.6) путём замены $S(P_{k+1})$ на $\tilde{S}(P_{k+1})$, где $\|X - \tilde{S}(P_{k+1})\|_{\mathbf{W}} \leq \|X - S(P_{k+1})\|_{\mathbf{W}}$. Это оправдано в том случае, если $\tilde{S}(P_{k+1})$ вычисляется быстрее и/или устойчивее, чем $S(P_{k+1})$.

1.3.2. Принцип Variable projection

Пусть $P = \begin{pmatrix} B \\ C \end{pmatrix} \in \mathbb{R}^p$, $B \in \mathbb{R}^{p_1}$, $C \in \mathbb{R}^{p_2}$, $p_1 + p_2 = p$. Тогда задачу (взвешенного) метода наименьших квадратов (1.4) с

$$S(P) = \mathbf{G}(B)C, \quad (1.7)$$

где $\mathbf{G} : \mathbb{R}^{p_1} \rightarrow \mathbb{R}^{N \times p_2}$ — гладкая функция, можно рассмотреть как задачу проектирования вектора данных X на заданное множество:

$$\min_{Y \in \mathcal{D}} \|X - Y\|_{\mathbf{W}}^2 \quad \text{с} \quad \mathcal{D} = \{\mathbf{G}(B)C \mid \begin{pmatrix} B \\ C \end{pmatrix} \in \mathbb{R}^p\}.$$

Под обозначением $\{\varphi(z) \mid z \in \mathcal{C}\}$ подразумевается множество значений $\varphi(z)$ для $z \in \mathcal{C}$. Воспользуемся тем, что нам известно решение подзадачи

$$C^*(B) = \arg \min_C \|X - \mathbf{G}(B)C\|_{\mathbf{W}}^2,$$

а именно $C^*(B) = (\mathbf{G}(B))_{\mathbf{W}}^\dagger X$.

Обозначим $S^*(B) = \mathbf{G}(B)C^*(B)$, $\mathcal{G}(B) = \{\mathbf{G}(B)C \mid C \in \mathbb{R}^{p_2}\}$. Отсюда следует, что

$$S^*(B) = \arg \min_{S \in \mathcal{G}(B)} \|X - S\|_{\mathbf{W}}^2. \quad (1.8)$$

Таким образом, мы можем свести задачу к проекции на подмножество $\mathcal{D}^* \subset \mathcal{D}$ и тем самым к задаче минимизации только по части параметров, входящих в $S(P)$ нелинейно:

$$\min_{Y \in \mathcal{D}^*} \|X - Y\|_{\mathbf{W}}^2 \quad \text{с} \quad \mathcal{D}^* = \{S^*(B) \mid B \in \mathbb{R}^{p_1}\}. \quad (1.9)$$

Этот подход называется принципом «Variable projection» (см. [36] для случая евклидовой нормы).

1.3.3. Принцип Variable projection в задаче аппроксимации рядом конечного ранга

В [16, 17] применением принципа Variable projection была получена следующая эквивалентная (3) задача проекции на подмножество $\mathcal{D}^* \subset \overline{\mathcal{D}}_r$:

$$Y^* = \arg \min_{Y \in \mathcal{D}_r^*} \|X - Y\|_{\mathbf{W}} \quad \text{где} \quad \mathcal{D}^* = \{\Pi_{\mathcal{Z}(A), \mathbf{W}} X \mid A \in \mathbb{R}^{r+1}\},$$

что было переписано в виде задачи поиска ОЛРФ, управляющей рядом $\mathbf{S}^*$:

$$A^* = \arg \min_{A \in \mathcal{C}} \|\mathbf{X} - \mathbf{S}^*(A)\|_{\mathbf{W}},$$

где $\mathbf{S}^*(A) = \Pi_{\mathcal{Z}(A), \mathbf{W}}(\mathbf{X})$, при этом рассматривались следующие варианты $\mathcal{C}$: единичная r -мерная сфера [16] и r -мерное аффинное многообразие [17]. Полученная минимизационная задача решалась различными методами, в частности,

методом Гаусса-Ньютона. В удобных для данной работы терминах этот подход рассмотрен в главе 3.

Ниже приведена таблица обозначений этой работы и обозначений работ [16, 17], которая может быть удобна для сопоставления методов.

Таблица 1.1. Соответствие обозначений

Эта работа	$\mathbf{X}$	N	$r + 1$	1	$N - r$	A	$\mathbf{W}$	$\Gamma(A)$
Usevich & Markovsky	p_D	n_p	m	d	n	R	W	Γ

1.4. Непараметрический метод Кэдзоу

Один из стандартных численных методов решения задачи (4) — метод итераций Кэдзоу [1], принадлежащий классу методов попеременных проекций. Определим следующую последовательность матриц, начинающуюся с заданной матрицы $\mathbf{X}$:

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

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

Изначально метод рассматривался только для случая фробениусовской нормы $\|\cdot\|_F = \|\cdot\|_{\mathbf{I}_L, \mathbf{I}_K}$. Первые обобщения на случай диагональных матриц $\mathbf{L}$, $\mathbf{R}$ вместе с методами их подбора для соответствия задач (3) и (4) при $\mathbf{W} = \mathbf{I}_N$ разобраны в [22]. Доказательство сходимости метода Кэдзоу по подпоследовательностям для случая фробениусовской нормы приведено в [1].

Решение задачи (4) не всегда можно использовать для (3). Известно, что даже в случае диагональной матрицы $\mathbf{W} = \mathbf{I}_N$, не существует положительно определённых диагональных матриц $\mathbf{L}$ и $\mathbf{R}$ таких, что $\|\mathcal{T}_L(\mathbf{Z})\|_{\mathbf{L}, \mathbf{R}} = \|\mathbf{Z}\|_{\mathbf{W}}$ для любого $\mathbf{Z}$ [37]. Таким образом, при использовании метода Кэдзоу для решения задачи (3) необходимо решать задачу подбора матриц $\mathbf{L}$ и $\mathbf{R}$, приближённо дающих требуемую норму.

1.5. Метод квадратичного программирования «Active Set»

Для решения задачи поиска весов, использующихся в алгоритме Кэдзоу, используется так называемый «Primary Active Set» метод решения задач квадратичного программирования, описанный в [35].

В общем случае, задача квадратичного программирования выглядит следующим образом:

Задача 1.5.1.

$$\min_{X \in \mathcal{X}} \frac{1}{2} X^T \mathbf{G} X - V^T X, \quad (1.10)$$

$$\mathcal{X} = \{X \mid A_i^T X = p_i, \quad i \in \mathcal{Y};$$

$$A_i^T X \geq p_i, \quad i \in \mathcal{G}\}, \quad (1.11)$$

где $\mathbf{G} \in \mathbb{R}^{K \times K}$ — произвольная положительно определённая матрица, $V \in \mathbb{R}^K$ — произвольный вектор, $\mathcal{Y}$ и $\mathcal{G}$ — множества индексов, вектора $A_i \in \mathbb{R}^K$ вместе с $p_i \in \mathbb{R}$ задают ограничения.

Суть любого «Active Set» метода состоит в последовательном переборе подмножества ограничений, которые выполняются как равенство для промежуточной точки-кандидата в решение задачи квадратичного программирования. Такое множество называется *рабочим множеством*, и обозначается, как $\mathcal{W} \subset \mathcal{Y} \cup \mathcal{G}$, при этом всегда $\mathcal{Y} \subset \mathcal{W}$, а ограничения, лежащие в рабочем множестве, называются *активными*. Ниже представлена схема этого метода для решения задачи 1.5.1.

Алгоритм 1.5.1 ([35], стр. 472). *Вход*: параметры задачи квадратичного программирования: матрица $\mathbf{G}$, вектор V , множества $\mathcal{Y}$, $\mathcal{G}$ и коэффициенты условий A_i , p_i .

- 1: Найти начальную точку X_0 , удовлетворяющую условиям задачи (1.10), (1.11), и положить $\mathcal{W}_0$ — множество активных ограничений в этой точке.

2: Положить $k = 0$

3: Положить $G_k = V - \mathbf{G}X_k$

4: Решить подзадачу квадратичного программирования и найти множители Лагранжа u_i , $i \in \mathcal{W}_k$:

$$P_k^* = \arg \min_{P_k \in \mathcal{P}_k} \frac{1}{2} P_k^T \mathbf{G} P_k - G_k^T P_k, \quad (1.12)$$

$$\mathcal{P}_k = \{P_k \mid A_i^T P_k = 0, \quad i \in \mathcal{W}_k\}, \quad (1.13)$$

после чего u_i находятся из системы уравнений $\sum_{i \in \mathcal{W}_k} u_i A_i = \mathbf{G}P_k^* - G_K$.

5: **if** $P_k^* = 0_K$, и все $u_i \geq 0$, $i \in \mathcal{W}_k \cap \mathcal{G}$ **then**

6: **return** $X^* = X_k$ — решение задачи 1.5.1.

7: **end if**

8: **if** $P_k^* = 0_K$, но существует i такое, что $u_i < 0$ **then**

9: Взять $j = \arg \min_{j \in \mathcal{W}_k \cap \mathcal{G}} u_j$, положить $\mathcal{W}_{k+1} = \mathcal{W}_k \setminus \{j\}$, увеличить k на единицу и перейти к п. 3.

10: **end if**

11: Положить $\alpha_k = \min \left(1, \min_{i \notin \mathcal{W}_k, A_i^T P_k^* < 0} \frac{p_i - A_i^T X_k}{A_i^T P_k^*} \right)$.

12: Положить $X_{k+1} = X_k + \alpha_k P_k^*$.

13: **if** $\alpha_k < 1$ **then**

14: Положить $j = \arg \min_{i \notin \mathcal{W}_k, A_i^T P_k^* < 0} \frac{p_i - A_i^T X_k}{A_i^T P_k^*}$ и $\mathcal{W}_{k+1} = \mathcal{W}_k \cup \{j\}$.

15: **else**

16: Положить $\mathcal{W}_{k+1} = \mathcal{W}_k$.

17: **end if**

18: Положить $k = k + 1$ и перейти к пункту 3.

Глава 2

Свойства задачи аппроксимации временных рядов рядами конечного ранга

2.1. Необходимые свойства $\mathcal{D}_r$ и $\overline{\mathcal{D}_r}$

Для построения параметризации необходимо доказать несколько лемм, устанавливающих свойства множества $\mathcal{D}_r$ рядов ранга r и его замыкания $\overline{\mathcal{D}_r}$.

Предложение 2.1.1. $S \in \overline{\mathcal{D}_r}$ тогда и только тогда, когда S управляется минимальной ОЛРФ(A), $A \in \mathbb{R}^{r^*+1}$, $r^* \leq r$.

Доказательство. Доказательство проведём в два шага. Первый — мы покажем, что из того, что $S \in \overline{\mathcal{D}_r}$, следует, что S управляется минимальной ОЛРФ порядка не больше r . Вторым шагом — мы покажем, что любой ряд, управляющийся ОЛРФ порядка не больше $r - 1$, можно сколь угодно близко приблизить рядом из $\mathcal{D}_r$, что означает, что он лежит в $\overline{\mathcal{D}_r}$.

Рассмотрим множество матриц $\mathcal{M}_r \subset \mathbb{R}^{L \times K}$ ранга, не превышающего r , и множество $\mathcal{M}_{=r}$ матриц ранга ровно r , где мы полагаем $L = r+1$, $K = N-L+1$. Известно, что замыкание $\overline{\mathcal{M}_{=r}} = \mathcal{M}_r$, см. [38]. Равенство $\mathcal{T}_L^{-1}(\mathcal{M}_{=r} \cap \mathcal{H}) = \mathcal{D}_r$ выполняется по определению множества $\mathcal{D}_r$ 1.1.2 и следствию 1.2.1. Следовательно, $\overline{\mathcal{D}_r} = \overline{\mathcal{T}_L^{-1}(\mathcal{M}_{=r} \cap \mathcal{H})} = \mathcal{T}_L^{-1}(\overline{\mathcal{M}_{=r}} \cap \mathcal{H}) \subset \mathcal{T}_L^{-1}(\overline{\mathcal{M}_{=r}} \cap \mathcal{H}) = \mathcal{T}_L^{-1}(\mathcal{M}_r \cap \mathcal{H}) \stackrel{\text{def}}{=} \widehat{\mathcal{D}_r}$. Получаем, что $\widehat{\mathcal{D}_r} = \{X : \text{rank } \mathcal{T}_{r+1}(X) \leq r\}$. По предложению 1.2.1 получаем, что $\overline{\mathcal{D}_r} \subset \widehat{\mathcal{D}_r} = \bigcup_{0 \leq i \leq r} \mathcal{D}_i$. Первый шаг завершён.

Перейдём ко второму шагу: докажем, что $\overline{\mathcal{D}_r} = \widehat{\mathcal{D}_r}$. Нам необходимо показать, что мы можем приблизить любой ряд $S \in \widehat{\mathcal{D}_r}$ временным рядом $X \in \mathcal{D}_r$ с произвольной точностью. Для этого достаточно показать, что мы можем аппроксимировать $S \in \mathcal{D}_{r^*}$ с помощью $X \in \mathcal{D}_{r^*+1}$ сколь угодно близко, $r^* + 1 \leq r$,

а затем продолжить процесс по индукции, если это необходимо.

Пусть $\mathbf{S}$ управляется ОЛРФ(A^*), $A^* \in \mathbb{R}^{r^*+1}$, $A^* = (a_1, \dots, a_{r^*+1})^T$. Мы строим N линейно независимых экспоненциальных рядов $\mathbf{S}_i$ длины N , $\mathbf{S}_i = (e^{\lambda_i}, e^{2\lambda_i}, \dots, e^{N\lambda_i})^T$, управляемых ОЛРФ(A_i), $A_i = (e^{\lambda_i}, -1)^T$, где все λ_i различны, и выбираем ряд $\mathbf{S}_{i^*}$, который не управляется ОЛРФ(A^*). Тогда для любого положительного сколь угодно малого числа α мы получаем $\mathbf{X}(\alpha) = \mathbf{S} + \alpha \mathbf{S}_{i^*} \in \widehat{\mathcal{D}}_{r^*+1}$, так как ряд $\mathbf{X}(\alpha)$ управляется ОЛРФ(B_{i^*}), $B_{i^*} = (e^{\lambda_{i^*}} a_1, e^{\lambda_{i^*}} a_2 - a_1, e^{\lambda_{i^*}} a_3 - a_2, \dots, e^{\lambda_{i^*}} a_{r^*+1} - a_{r^*}, -a_{r^*+1})^T \in \mathbb{R}^{r^*+2}$.

Давайте теперь покажем, что $\mathbf{X}(\alpha) \notin \widehat{\mathcal{D}}_{r^*}$, что означает, что $\mathbf{X}(\alpha) \in \mathcal{D}_{r^*+1}$. Согласно замечанию 1.2.1, достаточно показать, что $\text{rank } \mathcal{T}_{r^*+1}(\mathbf{X}(\alpha)) = r^* + 1$. Известно, что $\text{rank } \mathcal{T}_{r^*+1}(\alpha \mathbf{S}_{i^*}) = 1$. Более того, заметим, что вектор $V \in \text{colspace } \mathcal{T}_{r^*+1}(\mathbf{S})$ тогда и только тогда, когда $V^T A^* = 0$, и $U \in \text{rowspace } \mathcal{T}_{r^*+1}(\mathbf{S})$ тогда и только тогда, когда $U^T \mathbf{Q}^{N-r^*, r^*}(A^*) = 0_{N-2r}^T$.

Нам известно, что $(V^*)^T A^* \neq 0$, и $(U^*)^T \mathbf{Q}^{N-r^*, r^*}(A^*) \neq 0_{N-2r}^T$, где V^* — произвольный столбец матрицы $\mathcal{T}_{r^*+1}(\alpha \mathbf{S}_{i^*})$, U^* — произвольный столбец матрицы $\mathcal{T}_{r^*+1}^T(\alpha \mathbf{S}_{i^*})$. Следовательно,

$$\text{rowspan } \mathcal{T}_{r^*+1}(\mathbf{S}) \cap \text{rowspan } \mathcal{T}_{r^*+1}(\alpha \mathbf{S}_{i^*}) = \emptyset,$$

$$\text{colspace } \mathcal{T}_{r^*+1}(\mathbf{S}) \cap \text{colspace } \mathcal{T}_{r^*+1}(\alpha \mathbf{S}_{i^*}) = \emptyset.$$

По [39, Theorem 2] мы получаем $\text{rank } \mathcal{T}_{r^*+1}(\mathbf{X}(\alpha)) = r^* + 1$. Доказательство предложения завершено. $\square$

Замечание 2.1.1. Для комплексного случая данное предложение было доказано в [40, Remark 1.46].

Это предложение означает эквивалентность определения $\overline{\mathcal{D}}_r$ как замыкания $\mathcal{D}_r$ и как множества рядов ранга не более r в постановке задачи (3).

Нам потребуются следующие свойства:

- Согласно следствию 1.2.1 и замечанию 1.2.1, $Y \in \mathcal{D}_r$ тогда и только тогда, когда существует ОЛРФ(A) порядка r , которая управляет Y .
- $\overline{\mathcal{D}_r} = \{Y : \exists A \in \mathbb{R}^{r+1}, A \neq 0_{r+1} : A^T \mathcal{T}_{r+1}(S) = 0_{N-r}^T\}$, т.е. $Y \in \mathcal{D}_r$ эквивалентно существованию ОЛРФ(A) порядка r , которая управляет Y . Это свойство следует из предложения 1.2.1 и предложения 2.1.1. Более того, это означает, что $\overline{\mathcal{D}_r} = \bigcup_{A \neq 0_{r+1}} \mathcal{Z}(A)$, где $\mathcal{Z}(A)$ — подпространство рядов, управляемых ОЛРФ(A).

2.2. Параметризация множества $\mathcal{D}_r$

2.2.1. Конструкция параметризации

Рассмотрим временной ряд $S_0 \in \mathcal{D}_r$, управляемый минимальной ОЛРФ(A_0) порядка r , заданной ненулевым вектором $A_0 = (a_1^{(0)}, \dots, a_{r+1}^{(0)})^T$. 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, так как ОЛРФ(A_0) инвариантна относительно умножения на константу. Давайте рассмотрим параметризацию, которая будет зависеть от индекса τ .

Вместо r значений в начале ряда, то есть начальных данных, которые рассматриваются в случае обычной ЛРФ, мы рассмотрим $\tau - 1$ значений в начале, и $r + 1 - \tau$ значений в конце ряда. Определим $\mathcal{I}(\tau) = \{1, \dots, N\} \setminus \{\tau, \dots, N - r - 1 + \tau\}$ и $\mathcal{K}(\tau) = \{1, \dots, r + 1\} \setminus \{\tau\}$ — два множества индексов размера r . Множество $\mathcal{I}(\tau)$ содержит индексы элементов ряда (которые мы назовём 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 $\mathcal{K}(\tau)$).

Пусть $\mathbf{G}_{\mathcal{C},:}$ обозначает матрицу состоящую из строк $\mathbf{G}$ с индексами из $\mathcal{C}$, а $\mathbf{G}_{:, \mathcal{C}}$ — матрицу, состоящую из столбцов матрицы $\mathbf{G}$ с индексами из $\mathcal{C}$. В случае

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

Например, $\mathbf{Z}_{\mathcal{I}(\tau),:} \in \mathbb{R}^{r \times r}$ обозначает матрицу, состоящую из строк матрицы $\mathbf{Z} \in \mathbb{R}^{N \times r}$ с номерами из $\mathcal{I}(\tau)$, $A_{\mathcal{K}(\tau)} \in \mathbb{R}^r$ обозначает вектор состоящий из элементов вектора $A \in \mathbb{R}^{r+1}$ под номерами из $\mathcal{K}(\tau)$.

Следующая теорема определяет параметризацию, которая будет использоваться далее в работе. Начиная с текущего момента одно и тоже обозначение $\mathbf{S}$ будет использоваться как параметризующее отображение, так и как временной ряд.

Теорема 2.2.1. Пусть $\mathbf{S}_0 \in \mathcal{D}_r$ — ряд, управляемый ОЛРФ(A_0), где $A_0 \in \mathbb{R}^{r+1}$, $a_{\tau}^{(0)} = -1$. Между окрестностью точки $((\mathbf{S}_0)_{\mathcal{I}(\tau)}, (A_0)_{\mathcal{K}(\tau)})^T \in \mathbb{R}^{2r}$ и пересечением окрестности точки $\mathbf{S}_0$ со множеством $\mathcal{D}_r$ единственным образом определено взаимно-однозначное отображение $\mathbf{S} : \mathbb{R}^{2r} \rightarrow \mathcal{D}_r$, которое строит по вектору параметров $(S_{(\tau)}, A_{(\tau)})^T$ ряд $\mathbf{S}$, и удовлетворяет следующим условиям для $\mathbf{S} = \mathbf{S}(S_{(\tau)}, A_{(\tau)})$: $(\mathbf{S})_{\mathcal{I}(\tau)} = S_{(\tau)}$, $\mathbf{S} \in \mathcal{D}_r$ и управляется ОЛРФ(A) такой, что $A_{\mathcal{K}(\tau)} = A_{(\tau)}$ и $a_{\tau} = -1$.

Предъявим явно такое отображение $\mathbf{S}$, которое удовлетворяет упомянутым в теореме 2.2.1 свойствам, тем самым, докажем заодно и теорему 2.2.1.

Предложение 2.2.1. Пусть $A_0 \in \mathbb{R}^{r+1}$, $a_{\tau}^{(0)} = -1$, $\mathbf{Z}_0 \in \mathbb{R}^{N \times r}$ — матрица, столбцы которой составляют базис пространства рядов, управляемых ОЛРФ(A_0).

1. Рассмотрим $(S_{(\tau)}, A_{(\tau)})^T \in \mathbb{R}_{2r}$, и определим такое A , что $A_{\mathcal{K}(\tau)} = A_{(\tau)}$, $a_{\tau} = -1$. Пусть $\mathbf{\Pi}_{\mathcal{Z}(A)}$ — ортогональный проектор на $\mathcal{Z}(A)$, $\mathbf{Z} = \mathbf{\Pi}_{\mathcal{Z}(A)} \mathbf{Z}_0$, $\mathbf{G} = \left(\mathbf{Z} \mathbf{Z}_{\mathcal{I}(\tau),:}^{-1} \right)$.

Тогда отображение $\mathbf{S}$ из теоремы 2.2.1 имеет явный вид

$$\mathbf{S}(S_{(\tau)}, A_{(\tau)}) = \mathbf{G} S_{(\tau)}. \quad (2.1)$$

2. *Отображение, обратное к $\mathbf{S}$, задано следующей формулой: $S_{(\tau)}(\mathbf{S}) = (\mathbf{S})_{\mathcal{I}(\tau)}$*

и

$$A_{(\tau)}(\mathbf{S}) = (-\hat{A}/\hat{a}_\tau)_{\mathcal{K}(\tau)}, \quad (2.2)$$

где $\hat{A} = \hat{A}(\mathbf{S}) = (\hat{a}_1, \dots, \hat{a}_{r+1})^T = (\mathbf{I}_{r+1} - \mathbf{P}_{\mathcal{L}(\mathbf{S})})A_0$, $\mathcal{L}(\mathbf{S}) = \text{colspace}(\mathcal{T}_{r+1}(\mathbf{S}))$,

$\mathbf{P}_{\mathcal{L}(\mathbf{S})}$ — ортогональный проектор на $\mathcal{L}(\mathbf{S})$.

Доказательство. Предложение 2.2.1 даёт параметризующее отображение, если мы докажем корректность (2.1), (2.2), докажем, что отображение $\mathbf{S}$ удовлетворяет указанным в теореме 2.2.1 условиям, единственно, является инъекцией и что (2.2) задаёт обратное отображение к $\mathbf{S}$.

Докажем корректность (2.1). Для начала покажем, что $\mathbf{Z}_{\mathcal{I}(\tau),:}$ невырождена и, следовательно, обратима. Это будет следовать из невырожденности $(\mathbf{Z}_0)_{\mathcal{I}(\tau),:}$ для любого базиса $\mathcal{Z}(A_0)$.

Представим A_0 как $A_0 = (0, \dots, 0, b_{r_m+1}, \dots, b_1, 0, \dots, 0)^T$, с r_b нулями в начале и r_e нулями в конце, $r_e + r_b + r_m = r$. Выберем матрицу $\mathbf{Z}_0^*$ таким образом, что $\mathbf{Z}_0^* = [\mathbf{Z}_{\text{begin}}, \mathbf{Z}_{\text{middle}}, \mathbf{Z}_{\text{end}}]$ состоит из трёх блоков: $\mathbf{Z}_{\text{begin}} = \begin{pmatrix} \mathbf{I}_{r_b} \\ \mathbf{0}_{N-r_b, r_b} \end{pmatrix}$,

$$\mathbf{Z}_{\text{middle}} = \begin{pmatrix} \mathbf{0}_{r_b, r_m} \\ \hat{\mathbf{Z}}_{\text{middle}} \\ \mathbf{0}_{r_e, r_m} \end{pmatrix}, \quad \mathbf{Z}_{\text{end}} = \begin{pmatrix} \mathbf{0}_{N-r_e, r_e} \\ \mathbf{I}_{r_e} \end{pmatrix}, \quad \text{где столбцы } \hat{\mathbf{Z}}_{\text{middle}} \in \mathbb{R}^{(N-r_b-r_e), r_m} \text{ со-}$$

ставляют базис рядов длины $N - r_b - r_e$, управляемых обычной ЛРФ с коэффициентами $-b_2/b_1, \dots, -b_{r_m+1}/b_1$. Так как $\{1, \dots, r_b\} \cup \{N - r_e + 1, \dots, N\} \subset \mathcal{I}(\tau)$ и любая подматрица матрицы $\hat{\mathbf{Z}}_{\text{middle}}$ порядка $r_m \times r_m$ невырождена [32, Prop. 2.3], мы получаем обратимость $(\mathbf{Z}_0^*)_{\mathcal{I}(\tau),:}$. Любую другую матрицу, содержащую базис $\mathcal{Z}(A_0)$ в качестве столбцов, можно представить в виде $\mathbf{Z}_0^* \mathbf{P}$, где $\mathbf{P} \in \mathbb{R}^{r \times r}$ — некоторая невырожденная матрица. Следовательно, будет обратима и $(\mathbf{Z}_0^* \mathbf{P})_{\mathcal{I}(\tau),:}$.

Теперь докажем обратимость $\mathbf{Z}_{\mathcal{I}(\tau),:}$. $\mathcal{Z}(A)$ — подпространство, ортогональ-

ное $\mathcal{Q}(A)$, поэтому, $\mathbf{P}_{\mathcal{Z}(A)}$ можно представить в виде непрерывной по $A \neq 0$ функции $\mathbf{P}_{\mathcal{Z}(A)} = \mathbf{I}_N - \mathbf{P}_{\mathcal{Q}(A)}$, $\mathcal{Q}(A)$ определено в (1.2). Рассмотрим матрицу $\mathbf{Z}_{\mathcal{I}(\tau),:}$. Её определитель — непрерывная функция от $\mathbf{Z}$. В свою очередь, $\mathbf{Z}$ непрерывно зависит от $A_{(\tau)}$. Так как $\mathbf{Z}(A_0) = \mathbf{Z}_0$, и у $(\mathbf{Z}_0)_{\mathcal{I}(\tau),:}$ ненулевой определитель, можно выбрать такую окрестность $(A_0)_{\mathcal{K}(\tau)}$, что определитель $\mathbf{Z}_{\mathcal{I}(\tau),:}$ не обращается в ноль, следовательно, матрица $\mathbf{Z}_{\mathcal{I}(\tau),:}$ обратима.

Покажем, что (2.1) не зависит от выбора базиса $\mathbf{Z}_0$: для любой невырожденной матрицы $\mathbf{P} \in \mathbb{R}^{r \times r}$ выполнено $(\mathbf{P}_{\mathcal{Z}(A)} \mathbf{Z}_0 \mathbf{P})((\mathbf{P}_{\mathcal{Z}(A)} \mathbf{Z}_0 \mathbf{P})_{\mathcal{I}(\tau),:})^{-1} = \mathbf{Z} \mathbf{Z}_{\mathcal{I}(\tau),:}^{-1}$.

Теперь покажем что выполнены свойства $\mathbf{S}$, заданные в теореме 2.2.1, то есть покажем, что $\mathbf{S} \in \mathcal{D}_r$, $\mathbf{S}$ управляется ОЛРФ(A) и $(\mathbf{S})_{\mathcal{I}(\tau)} = S_{(\tau)}$. $\mathbf{S}$ управляется ОЛРФ(A), так как каждый столбец матрицы $\mathbf{Z}$ управляется ОЛРФ(A). Покажем принадлежность $\mathbf{S} \in \mathcal{D}_r$. Рассмотрим матрицу $\mathcal{T}_{r+1}(\mathbf{S}_0)$, и выберем у неё подматрицу размера $r \times r$ с ненулевым определителем. Теперь рассмотрим определитель подматрицы $\mathbf{B}$ матрицы $\mathcal{T}_{r+1}(\mathbf{S}(S_{(\tau)}, A_{(\tau)}))$ с тем же расположением. Этот определитель — непрерывная функция от $(S_{(\tau)}, A_{(\tau)})$, так как (2.1) — непрерывная функция. Следовательно, мы можем выбрать такую окрестность $(S_{(\tau)}, A_{(\tau)})$, в которой определитель $\mathbf{B}$ не обращается в ноль, из чего получаем $\mathbf{S} \in \mathcal{D}_r$. Условие $(\mathbf{S})_{\mathcal{I}(\tau)} = S_{(\tau)}$ выполнено, так как

$$(\mathbf{S})_{\mathcal{I}(\tau)} = \left(\mathbf{Z}_{\mathcal{I}(\tau),:} \mathbf{Z}_{\mathcal{I}(\tau),:}^{-1} \right) S_{(\tau)} = S_{(\tau)}.$$

Объясним единственность отображения (2.1), удовлетворяющего упомянутым в теореме 2.2.1 свойствам. Пусть $\mathbf{S}_2$ — другое отображение, удовлетворяющее условиям теоремы 2.2.1, $\hat{\mathbf{S}}_2 = \mathbf{S}_2(S_{(\tau)}, A_{(\tau)}) \in \mathcal{D}_r$. Знаем, что $\hat{\mathbf{S}}_2 \in \mathcal{Z}(A)$. Следовательно, столбцы $\mathbf{Z}$ задают базис $\mathcal{Z}(A)$. Пусть $\hat{\mathbf{S}}_2 = \mathbf{Z}V$, $V \in \mathbb{R}^r$ — разложение $\hat{\mathbf{S}}_2$ по этому базису. Тогда должно быть выполнено $(\mathbf{Z}V)_{\mathcal{I}(\tau)} = S_{(\tau)}$. Но $\mathbf{Z}_{\mathcal{I}(\tau)}V = S_{(\tau)}$, а мы знаем, что $\mathbf{Z}_{\mathcal{I}(\tau)}$ обратима, следовательно, $V = \mathbf{Z}_{\mathcal{I}(\tau)}^{-1}S_{(\tau)}$, откуда следует, что $\hat{\mathbf{S}}_2 = \left(\mathbf{Z} \mathbf{Z}_{\mathcal{I}(\tau)}^{-1} \right) S_{(\tau)} = \mathbf{S}$. Единственность показана.

Покажем инъективность отображения $\mathbf{S}$. Выберем два таких неравных на-

бора параметров $(S_{(\tau)}^{(1)}, A_{(\tau)}^{(1)})^T, (S_{(\tau)}^{(2)}, A_{(\tau)}^{(2)})^T$ в окрестности $((S_0)_{\mathcal{I}(\tau)}, (A_0)_{\mathcal{K}(\tau)})^T$, и рассмотрим $\mathbf{X}_1 = \mathbf{S}(S_{(\tau)}^{(1)}, A_{(\tau)}^{(1)})$, $\mathbf{X}_2 = \mathbf{S}(S_{(\tau)}^{(2)}, A_{(\tau)}^{(2)})$. Если $S_{(\tau)}^{(1)} \neq S_{(\tau)}^{(2)}$, то очевидно $\mathbf{X}_1 \neq \mathbf{X}_2$, так как $(\mathbf{X}_1)_{\mathcal{I}(\tau)} \neq (\mathbf{X}_2)_{\mathcal{I}(\tau)}$. Пусть $S_{(\tau)}^{(1)} = S_{(\tau)}^{(2)}$, но $A_{(\tau)}^{(1)} \neq A_{(\tau)}^{(2)}$. Тогда вектор V_1 ортогонален $\text{colspace } \mathcal{T}_{r+1}(\mathbf{X}_1)$ тогда и только тогда, когда $V_1 = c_1 A^{(1)}$, вектор V_2 ортогонален $\text{colspace } \mathcal{T}_{r+1}(\mathbf{X}_2)$ тогда и только тогда, когда $V_2 = c_2 A^{(2)}$, где $A^{(1)} \in \mathbb{R}^{r+1}$, $A_{\mathcal{K}(\tau)}^{(1)} = A_{(\tau)}^{(1)}$, $a_{\tau}^{(1)} = -1$, $A^{(2)} \in \mathbb{R}^{r+1}$, $A_{\mathcal{K}(\tau)}^{(2)} = A_{(\tau)}^{(2)}$, $a_{\tau}^{(2)} = -1$, $c_1 \neq 0$, $c_2 \neq 0$ — произвольные. Но не существует таких c_1, c_2 , что $c_1 A^{(1)} = c_2 A^{(2)}$, следовательно $\mathbf{X}_1 \neq \mathbf{X}_2$. Инъективность доказана.

Докажем корректность (2.2). Согласно утверждению предложения, $A_{(\tau)}(\mathbf{S})$ — перенормировка вектора $\hat{A}$ такая, чтобы τ -й элемент был равен -1 . Докажем корректность определения $A_{(\tau)}(\mathbf{S})$, то есть возможность перенормировки $\hat{A}$. Рассмотрим матрицу $\mathbf{S} = \mathcal{T}_{r+1}(\mathbf{S})$, $\mathbf{S} \in \mathbb{R}^{(r+1) \times (N-r)}$. Пусть $\mathcal{J}$ будет таким подмножеством индексов, что подматрица $(\mathbf{S}_0)_{:, \mathcal{J}} \in \mathbb{R}^{(r+1) \times r}$ имеет ранг r , где $\mathbf{S}_0 = \mathcal{T}_{r+1}(\mathbf{S}_0)$. Тогда $\mathbf{P}_{\mathcal{L}(\mathbf{S})}$ можно записать в виде непрерывной в окрестности $\mathbf{S}_0$ функции $\mathbf{P}_{\mathcal{L}(\mathbf{S})} = \mathbf{P}_{\mathbf{S}_{:, \mathcal{J}}(\mathbf{S})}$, следовательно, можно выбрать такую окрестность, в которой $\hat{a}_{\tau}$ не обратится в нуль.

Объясним, что (2.2) является обратным отображением к отображению $\mathbf{S}$. Пусть $\mathbf{S} = \mathbf{S}(S_{(\tau)}, A_{(\tau)})$. Значения $S_{(\tau)} = \mathbf{S}_{\mathcal{I}(\tau)}$ берутся непосредственно из временного ряда. Ряд $\mathbf{S}$ управляется ОЛРФ($\hat{A}$), так как вектор $\hat{A}$ ортогонален $\text{colspace}(\mathcal{T}_{r+1}(\mathbf{S}))$ по своему определению. При этом ряд $\mathbf{S}$ управляется ОЛРФ(A); следовательно, A совпадает с $\hat{A}$ с точностью до умножения на константу. Следовательно, перенормировка $\hat{A}$ даёт требуемый вектор $A_{(\tau)}$.

Это соображение заканчивает доказательство теоремы 2.2.1 и предложения 2.2.1. □

2.2.2. Гладкость параметризации и производные

Теорема 2.2.2. Пусть $S_0 \in \mathcal{D}_r$ — ряд, управляемый ОЛРФ(A_0), где $a_\tau^{(0)} = -1$. Тогда параметризация $S(S_{(\tau)}, A_{(\tau)})$, введённая в теореме 2.2.1 — гладкий диффеоморфизм между окрестностью вектора $((S_0)_{\mathcal{I}(\tau)}, (A_0)_{\mathcal{K}(\tau)})^T \in \mathbb{R}^{2r}$ и пересечением окрестности S_0 с $\mathcal{D}_r$.

Доказательство. Достаточно показать, что $\Pi_{\mathcal{L}(S)}$ и $\Pi_{\mathcal{Z}(A)}$ из предложения 2.2.1 — гладкие (то есть бесконечно дифференцируемые) проекторы в окрестностях S_0 и Z_0 соответственно.

Так как $(S_0)_{:, \mathcal{J}}$ невырождена, $\Pi_{\mathcal{L}(S)} = S_{:, \mathcal{J}}((S_{:, \mathcal{J}})^T S_{:, \mathcal{J}})^{-1} S_{:, \mathcal{J}}^T$ — гладкая функция в окрестности S_0 . Так как $Q(A)$ — матрица полного ранга согласно определению (1.2), $\Pi_{\mathcal{Z}(A)} = I_N - Q(A)(Q^T(A)Q(A))^{-1}Q^T(A)$ — гладкая функция, за исключением точки $A = 0_{r+1}$.

Ясно, что остальные операции, указанные в параметризации, гладкие в своих соответствующих окрестностях. $\square$

Замечание 2.2.1. Отображение, сопоставляющее временному ряду S пару соответствующих ему параметров $(S_{(\tau)}, A_{(\tau)})^T$, тривиальным образом (т.е. с использованием той же аналитической формулы) гладко продолжается с пересечения окрестности с $\mathcal{D}_r$ на всю окрестность S_0 .

Замечание 2.2.2. Теорема 2.2.2 предъявляет гладкую параметризацию $\mathcal{D}_r$ в окрестности S_0 , что означает, что $\mathcal{D}_r$ является гладким многообразием.

Рассмотрим производные, соответствующие этой параметризации. Под $J_S = J_S(S_{(\tau)}, A_{(\tau)}) \in \mathbb{R}^{N \times 2r}$ будем обозначать матрицу Якоби отображения $S(S_{(\tau)}, A_{(\tau)})$.

По определению, касательное подпространство к $\mathcal{D}_r$ в точке S равно $\text{colspace } J_S(S_{(\tau)}, A_{(\tau)})$, где $S = S(S_{(\tau)}, A_{(\tau)})$. Касательное пространство инвариантно относительно выбора конкретной параметризации $\mathcal{D}_r$ в окрестности S .

Под A^2 определим ациклическую свёртку A самим с собой:

$$A^2 = (a_i^{(2)}) \in \mathbb{R}^{2r+1}, \quad a_i^{(2)} = \sum_{j=\max(1, i-r)}^{\min(i, r+1)} a_j a_{i-j+1}. \quad (2.3)$$

Теорема 2.2.3. *Касательное подпространство к $\mathcal{D}_r$ в точке S , управляемой ОЛРФ(A), имеет размерность $2r$ и равно $\mathcal{Z}(A^2)$.*

Перед этим докажем несколько вспомогательных лемм.

Удобно разделить параметры на две части, $S_{(\tau)}$ и $A_{(\tau)}$. Тогда $\mathbf{J}_S = [\mathbf{F}_S : \mathbf{F}_A]$, где $\mathbf{F}_S = (\mathbf{J}_S)_{:, \{1, \dots, r\}}$, $\mathbf{F}_A = (\mathbf{J}_S)_{:, \{r+1, \dots, 2r\}}$.

Лемма 2.2.1. $\mathbf{Q}(A)^T \mathbf{F}_S = \mathbf{0}_{N-r, r}$; $\text{colspace } \mathbf{F}_S = \mathcal{Z}(A)$.

Доказательство. Обозначим $\mathbf{F}_S = [F_{S,1}, \dots, F_{S,r}]$. Рассмотрим равенство $\mathbf{Q}^T(A)S = \mathbf{0}_{N-r}$, и продифференцируем его по $(S_{(\tau)})_{(i)}$, $i = 1, \dots, r$. Мы получаем $\mathbf{Q}^T(A)F_{S,i} = \mathbf{0}_{N-r}$, что означает, что $\text{colspace } \mathbf{F}_S \subset \mathcal{Z}(A)$. Тот факт, что $(\mathbf{F}_S)_{\mathcal{I}(\tau), :} = \mathbf{I}_r$, завершает доказательство. $\square$

Лемма 2.2.2. $\mathbf{Q}^T(A)\mathbf{F}_A = -\mathbf{S}_{\mathcal{K}(\tau), :}^T$, где $\mathbf{S} = \mathcal{T}_{r+1}(S)$; $\text{colspace } \mathbf{F}_A \subset \mathcal{Z}(A^2)$.

Доказательство. Обозначим $\mathbf{F}_A = [F_{A,1}, \dots, F_{A,r}]$. Рассмотрим $\mathbf{Q}^T(A)S = \mathbf{0}_{N-r}$ и продифференцируем это равенство по $(a_{\mathcal{K}(\tau)})_i$, то есть i -му элементу $a_{\mathcal{K}(\tau)}$, $i = 1, \dots, r$. Таким образом, мы получаем $\mathbf{Q}^T(E_j)S + \mathbf{Q}(A)F_{A,i} = \mathbf{0}_{N-r}$, $j = (\mathcal{K}(\tau))_i$ — i -й элемент множества $\mathcal{K}(\tau)$. Заметим, что $\mathbf{Q}^T(E_j)S$ — это столбец транспонированной траекторной матрицы $\mathbf{S}^T$ с номером, равным i -му элементу множества $\mathcal{K}(\tau)$. Следовательно, равенство $\mathbf{Q}^T(A)\mathbf{F}_A = -\mathbf{S}_{\mathcal{K}(\tau), :}^T$ доказано.

Чтобы доказать вторую часть леммы, рассмотрим матрицу $\mathbf{Q}^{N-r, r}(A) \in \mathbb{R}^{N \times (N-r)}$. Согласно первому утверждению, верно следующее:

$$(\mathbf{Q}^{N-r, r}(A))^T (\mathbf{Q}^{N, r}(A))^T \mathbf{F}_A = \mathbf{0}_{N-2r, r}.$$

Тогда

$$(\mathbf{Q}^{N-r, r}(A))^T = \sum_{i=1}^{r+1} a_i (\mathbf{Q}^{N-r, r}(E_i))^T, \quad (\mathbf{Q}^{N, r}(A))^T = \sum_{j=1}^{r+1} a_j (\mathbf{Q}^{N, r}(E_j))^T.$$

Заметим, что $(\mathbf{Q}^{N-r,r}(E_i))^T(\mathbf{Q}^{N,r}(E_j))^T = (\mathbf{Q}^{N,2r}(E_{i+j-1}))^T$. Таким образом,

$$\left(\sum_{i=1}^{r+1} \sum_{j=1}^{r+1} a_i a_j (\mathbf{Q}^{N,2r}(E_{i+j-1}))^T\right) \mathbf{F}_A = (\mathbf{Q}^{N,2r} \left(\sum_{i=1}^{r+1} \sum_{j=1}^{r+1} a_i a_j E_{i+j-1}\right))^T \mathbf{F}_A = \mathbf{0}_{N-2r,r},$$

что эквивалентно $\mathbf{Q}^T(A^2) \mathbf{F}_A = \mathbf{0}_{N-2r,r}$. $\square$

Теперь перейдём к доказательству теоремы.

Доказательство теоремы 2.2.3. Из Леммы 2.2.1 следует, что

$$\mathbf{Q}^T(A^2) \mathbf{F}_S = (\mathbf{Q}^{N-r,r}(A))^T \mathbf{Q}^T(A) \mathbf{F}_S = \mathbf{0}_{N-2r,r}.$$

Следовательно, $\text{colspace } \mathbf{J}_S = \mathcal{Z}(A) \subset \mathcal{Z}(A^2)$. Так как в пересечении окрестности точки $\mathbf{S}$ с $\mathcal{D}_r$ есть диффеоморфизм $\mathbf{S}$, матрица Якоби $\mathbf{J}_S$ имеет полный ранг $2r$. Следовательно, $\text{colspace } \mathbf{J}_S = \mathcal{Z}(A^2)$. $\square$

2.3. Дополнительные свойства $\overline{\mathcal{D}_r}$

Следующая теорема показывает, что $\overline{\mathcal{D}_r}$ не является гладким многообразием размерности $2r$ в каждой точке. На самом деле, $\overline{\mathcal{D}_r}$ является алгебраическим многообразием, где каждая точка $\overline{\mathcal{D}_r} \setminus \mathcal{D}_r$ является сингулярной (т.е. точкой, в окрестности которой нет диффеоморфизма).

Теорема 2.3.1. *Возьмём $\mathbf{S}_0 \in \overline{\mathcal{D}_r} \setminus \mathcal{D}_r$. Не существует диффеоморфизма между любым открытым множеством $\mathcal{P} \in \mathbb{R}^{2r}$ и пересечением окрестности $\mathbf{S}_0$ с $\overline{\mathcal{D}_r}$.*

Доказательство. Предположим, что ряд $\mathbf{S}_0 \in \mathcal{D}_{r_0}$, $r_0 < r$ такой, что $\mathbf{S}_0$ управляется ОЛРФ(A_0), $A_0 \in \mathbb{R}^{r_0+1}$, $A_0 = (a_1, \dots, a_{r_0+1})^T$. Идея доказательства следующая: если ранг ряда $\mathbf{S}_0$ меньше r , значит, добавлением ряда ранга 1 к $\mathbf{S}_0$ мы получим ряд, лежащий в $\mathcal{D}_r$. Найдём набор рядов ранга 1, с помощью которого легко получить противоречие с существованием диффеоморфизма.

Построим N линейно независимых экспоненциальных рядов $\mathbf{S}^{(i)}$ длины N , $\mathbf{S}^{(i)} = (e^{\lambda_i}, e^{2\lambda_i}, \dots, e^{N\lambda_i})^T$, управляемых ОЛРФ(A_i), $A_i = (e^{\lambda_i}, -1)^T$, где все λ_i различны. Тогда для любой α мы имеем $\mathbf{S}_0 + \alpha\mathbf{S}^{(i)} \in \overline{\mathcal{D}_r}$, так как ряд $\mathbf{S}_0 + \alpha\mathbf{S}^{(i)}$ управляется ОЛРФ(B_i) с $B_i = (e^{\lambda_i}a_1, e^{\lambda_i}a_2 - a_1, e^{\lambda_i}a_3 - a_2, \dots, e^{\lambda_i}a_{r_0+1} - a_{r_0}, -a_{r_0+1})^T \in \mathbb{R}^{r_0+2}$, $r_0+2 \leq r+1$. Более того, все N прямых $\mathcal{L}_i = \{\mathbf{S}_0 + \alpha\mathbf{S}^{(i)} | \alpha \in \mathbb{R}\}$ принадлежат $\overline{\mathcal{D}_r}$.

Предположим противное: пусть существует диффеоморфизм в окрестности $\mathbf{S}_0$ в $\overline{\mathcal{D}_r}$. Пусть $\mathcal{F}$ — касательное подпространство к $\overline{\mathcal{D}_r}$ в точке $\mathbf{S}_0$ ($\mathcal{F}$ имеет размерность $2r$). $\mathcal{F}$ должно содержать каждую прямую $\mathcal{L}_i - \mathbf{S}_0 = \{\alpha\mathbf{S}^{(i)} | \alpha \in \mathbb{R}\}$, $i = 1, \dots, N$. Мы получаем, что $\mathcal{F}$ — подпространство $\mathbb{R}^N$ размерности $2r$, которое должно содержать N линейно независимых рядов $\mathbf{S}^{(1)}, \dots, \mathbf{S}^{(N)}$. Следовательно, получено противоречие. $\square$

Теорема 2.3.1 объясняет, почему в дальнейшем используется предположение о принадлежности ряда $\mathbf{S}_0$ множеству $\mathcal{D}_r$, а не $\overline{\mathcal{D}_r}$.

2.4. Условия минимума в задаче аппроксимации рядами конечного ранга

В этом разделе мы перейдём к свойствам задачи (3). Для начала, мы определим стандартным образом понятия локального и глобального минимума.

Определение 2.4.1. Ряд $\mathbf{S}_0$ является точкой локального минимума задачи (3), если $\|\mathbf{X} - \mathbf{S}\|_{\mathbf{W}} \geq \|\mathbf{X} - \mathbf{S}_0\|_{\mathbf{W}}$ для любого временного ряда $\mathbf{S}$, лежащего в пересечении окрестности ряда $\mathbf{S}_0$ с $\overline{\mathcal{D}_r}$.

Для того чтобы определить глобальный минимум, введём множество, зависящее от ряда $\mathbf{X}$ (то есть многозначное отображение):

$$\Pi_{\overline{\mathcal{D}_r}}(\mathbf{X}) = \{\mathbf{S} \in \overline{\mathcal{D}_r} : \|\mathbf{X} - \mathbf{S}\|_{\mathbf{W}} = \inf_{\mathbf{Y} \in \overline{\mathcal{D}_r}} \|\mathbf{X} - \mathbf{Y}\|_{\mathbf{W}}\}. \quad (2.4)$$

Множество $\Pi_{\overline{\mathcal{D}_r}}(\mathbf{X})$ всегда непусто, так как инфимум внутри выражения всегда достижим из-за замкнутости $\overline{\mathcal{D}_r}$. Более того, множество $\Pi_{\overline{\mathcal{D}_r}}(\mathbf{X})$ замкнуто, так как является пересечением замкнутого множества $\overline{\mathcal{D}_r}$ и замкнутого шара радиуса $\inf_{\mathbf{Y} \in \overline{\mathcal{D}_r}} \|\mathbf{X} - \mathbf{Y}\|_{\mathbf{W}}$ по норме $\|\cdot\|_{\mathbf{W}}$.

Определение 2.4.2. Ряд $\mathbf{S}_0$ является точкой глобального минимума задачи (3), если $\mathbf{S}_0 \in \Pi_{\overline{\mathcal{D}_r}}(\mathbf{X})$, где $\Pi_{\overline{\mathcal{D}_r}}(\mathbf{X})$ определено в (2.4).

Лемма 2.4.1. Пусть $\mathbf{X} \notin \overline{\mathcal{D}_r} \setminus \mathcal{D}_r$. Тогда любая точка глобального минимума задачи (3) лежит в $\mathcal{D}_r$.

Доказательство. Если $\mathbf{X} \in \mathcal{D}_r$, то утверждение теоремы тривиально, так как $\Pi_{\overline{\mathcal{D}_r}}(\mathbf{X}) = \{\mathbf{X}\}$. Поэтому пусть $\mathbf{X} \notin \overline{\mathcal{D}_r}$.

Предположим противное. Пусть $\mathbf{S}_0$ — точка глобального минимума функции (3), $\mathbf{S}_0 \in \mathcal{D}_{r_0}$, $r_0 < r$; поэтому $\mathbf{X} \neq \mathbf{S}_0$. Построение $A_0, \mathbf{S}^{(i)}, A_i, B_i$ проводится в точности так же, как это делалось в доказательстве теоремы 2.3.1.

Введём скалярное произведение $\langle \mathbf{Z}, \mathbf{Y} \rangle_{\mathbf{W}} = \mathbf{Z}^T \mathbf{W} \mathbf{Y}$, порождающее норму $\|\cdot\|_{\mathbf{W}}$. Рассмотрим скалярные произведения $\langle \mathbf{X} - \mathbf{S}_0, \mathbf{S}^{(i)} \rangle_{\mathbf{W}}$ для $i = 1, 2, \dots, N$. Так как линейная оболочка всех $\mathbf{S}^{(i)}$ образует базис $\mathbb{R}^N$, найдётся такой индекс j , что $\langle \mathbf{X} - \mathbf{S}_0, \mathbf{S}^{(j)} \rangle_{\mathbf{W}} \neq 0$. Тогда возьмём $\mathbf{S}_1 = \mathbf{S}_0 + \frac{\langle \mathbf{X} - \mathbf{S}_0, \mathbf{S}^{(j)} \rangle_{\mathbf{W}}}{\|\mathbf{S}^{(j)}\|_{\mathbf{W}}^2} \mathbf{S}^{(j)}$, управляемый ОЛРФ(B_i) (следовательно, лежащий в $\overline{\mathcal{D}_r}$), и покажем, что $\|\mathbf{X} - \mathbf{S}_1\|_{\mathbf{W}} < \|\mathbf{X} - \mathbf{S}_0\|_{\mathbf{W}}$:

$$\langle \mathbf{X} - \mathbf{S}_0, \mathbf{X} - \mathbf{S}_0 \rangle_{\mathbf{W}} - \langle \mathbf{X} - \mathbf{S}_1, \mathbf{X} - \mathbf{S}_1 \rangle_{\mathbf{W}} = \frac{(\langle \mathbf{X} - \mathbf{S}_0, \mathbf{S}^{(j)} \rangle_{\mathbf{W}})^2}{\|\mathbf{S}^{(j)}\|_{\mathbf{W}}^2} > 0.$$

Таким образом, получено противоречие с тем, что $\mathbf{S}_0 \in \Pi_{\overline{\mathcal{D}_r}}(\mathbf{X})$. □

Замечание 2.4.1. Смысл леммы в том, что за исключением тривиального случая $\mathbf{X} \in \overline{\mathcal{D}_r}$, при котором $\Pi_{\overline{\mathcal{D}_r}}(\mathbf{X}) = \{\mathbf{X}\}$, любая точка глобального минимума $\mathbf{S}_0 \in \Pi_{\overline{\mathcal{D}_r}}(\mathbf{X})$ всегда лежит в $\mathcal{D}_r$. Это означает, что в нетривиальном случае решение задачи (3) достаточно искать в гладком многообразии $\mathcal{D}_r$.

Замечание 2.4.2. Абсолютно аналогичным образом можно доказать, что при $\mathbf{X} \notin \overline{\mathcal{D}_r} \setminus \mathcal{D}_r$ любой локальный минимум задачи (3) $\mathbf{S}^* \in \mathcal{D}_r$. Доказательство аналогично доказательству леммы 2.4.1, за исключением взятия $\mathbf{S}_1 = \mathbf{S}_0 + \alpha \mathbf{S}^{(j)}$, где $0 < \alpha < \langle \mathbf{X} - \mathbf{S}_0, \mathbf{S}^{(j)} \rangle_{\mathbf{W}} / \|\mathbf{S}^{(j)}\|_{\mathbf{W}}^2$, α можно брать сколь угодно малым.

Сформулируем необходимые условия локального минимума.

Лемма 2.4.2 (Необходимые условия локального минимума). *Если точка $\mathbf{S}_0 \in \mathcal{D}_r$, управляемая ОЛРФ(A_0), является точкой локального минимума в задаче (3), то $\Pi_{\mathcal{Z}(A_0^2), \mathbf{W}}(\mathbf{X} - \mathbf{S}_0) = 0_N$.*

Доказательство. Рассмотрим параметризацию из раздела 2.2. Возьмём подходящий индекс τ и отображение $\mathbf{S}(S_{(\tau)}, A_{(\tau)})$ из (2.1). Тогда, согласно теореме 2.2.2, целевая функция $\|\mathbf{X} - \mathbf{S}(S_{(\tau)}, A_{(\tau)})\|_{\mathbf{W}}^2$ является гладкой функцией в окрестности вектора $((\mathbf{S}_0)_{\mathcal{I}(\tau)}, (A_0)_{\mathcal{K}(\tau)})^T \in \mathbb{R}^{2r}$.

Подстановка целевой функции в [35, Theorem 2.2] с учётом теоремы 2.2.3 даёт утверждение леммы. $\square$

Заметим, что эти условия являются необходимыми, а не достаточными, так как согласно [35, Theorem 2.3] дополнительно требуется положительная определённость матрицы Гёссе целевой функции при использовании параметризации.

В дальнейшем нам понадобится проекция на множество $\overline{\mathcal{D}_r}$. Нам нужно понять, в каких случаях множество глобальных минимумов $\Pi_{\overline{\mathcal{D}_r}}(\mathbf{X})$ содержит одну точку (т.е. проекция однозначна) и какими свойствами обладает соответствующий проектор.

Лемма 2.4.3. *Пусть $\mathbf{S}_0 \in \mathcal{D}_r$ и управляется ОЛРФ(A_0). Тогда существует открытая окрестность $\mathcal{C}(\mathbf{S}_0)$ ряда $\mathbf{S}_0$ такая, что в ней определён бесконечно дифференцируемый по $\mathbf{X}$ оператор $\Pi_{\mathcal{D}_r}^{(\mathbf{S}_0)}(\mathbf{X})$, при этом $\Pi_{\overline{\mathcal{D}_r}}(\mathbf{X}) = \{\Pi_{\mathcal{D}_r}^{(\mathbf{S}_0)}(\mathbf{X})\}$ для*

любого $X \in \mathcal{C}(S_0)$, то есть $\Pi_{\overline{\mathcal{D}_r}}(X)$ состоит из одной точки, и матрица Якоби оператора $\Pi_{\mathcal{D}_r}^{(S_0)}(X)$ равна $\Pi_{Z(A_0^2), \mathbf{W}}$.

Доказательство. В окрестности S_0 множество $\overline{\mathcal{D}_r}$ имеет структуру сколь угодно раз дифференцируемого гладкого многообразия, так как можно выбрать столь малый шар $B_{S_0, \delta} = \{X : \|X - S_0\|_{\mathbf{W}} < \delta\}$, что $\overline{\mathcal{D}_r}$ совпадает в нём с $\mathcal{D}_r$, т.е. $B_{S_0, \delta} \cap \overline{\mathcal{D}_r} = B_{S_0, \delta} \cap \mathcal{D}_r$.

Тогда, согласно [38, Lemma 2.1], вокруг точки S_0 существует открытое множество $\mathcal{C}$, в котором оператор проекции $\Pi_{\mathcal{D}_r}^{(S_0)}(X)$ однозначен, бесконечно дифференцируем, и его производная равна $\Pi_{Z(A_0^2), \mathbf{W}}$, что следует из вида касательного многообразия к $\mathcal{D}_r$ в точке S_0 по теореме 2.2.3. $\square$

Замечание 2.4.3. Назовём любое измеримое относительно борелевской сигма-алгебры отображение $\hat{\Pi}$ из $\mathbb{R}^N$ в $\mathbb{R}^N$ локальным проектором на $\mathcal{D}_r$ в окрестности точки S_0 , если $\hat{\Pi}(X)$ совпадает с $\Pi_{\mathcal{D}_r}^{(S_0)}(X)$, полученным в лемме 2.4.3, для любого $X \in \mathcal{C}(S_0)$. Например, такое отображение можно получить, доопределив $\Pi_{\mathcal{D}_r}^{(S_0)}(X)$ нулём вне $\mathcal{C}(S_0)$.

В дальнейшем мы будем использовать только обозначение $\Pi_{\mathcal{D}_r}^{(S_0)}(X)$, подразумевая измеримое отображение, определённое во всём $\mathbb{R}^N$.

2.5. Задача аппроксимации как задача оценивания сигнала

Рассмотрим модель со случайным рядом $X = S_0 + t\xi$, где сигнал $S_0 \in \mathcal{D}_r$ управляется ОЛРФ(A_0), ξ — случайный вектор, имеющий распределение Ξ , t — малое вещественное число. Нас интересует распределение оценки сигнала $\hat{S}(t) = \Pi_{\mathcal{D}_r}^{(S_0)}(S_0 + t\xi)$, где $\Pi_{\mathcal{D}_r}^{(S_0)}(X)$ — локальный проектор на $\mathcal{D}_r$ в окрестности точки S_0 (см. лемму 2.4.3 и замечание 2.4.3).

Кроме того, помимо оценки сигнала нас интересует распределение оценки

параметров сигнала. Пусть задан гладкий диффеоморфизм T между областью $\mathcal{C}_{2r} \subset \mathbb{R}^{2r}$ и пересечением множества $\mathcal{D}_r$ с открытой окрестностью $\mathcal{C}_N$ точки S_0 . Под T^{-1} обозначим обратное отображение к отображению T ($\mathsf{S}_0 = \mathsf{T}(P_0)$, $P_0 \in \mathcal{C}_{2r}$, $(\mathsf{T}^{-1} \circ \mathsf{T})(P) = P$ для любого $P \in \mathcal{C}_{2r}$), область определения которого расширена так, что во всей окрестности $\mathcal{C}_N$ отображение T^{-1} является гладким. Для корректности, доопределим T^{-1} во всех остальных точках $\mathsf{Z} \notin \mathcal{C}_N$ нулём или иным измеримым способом. Обозначим матрицу Якоби $\mathsf{T}(P)$ в точке P_0 как $\mathbf{J}_0$, матрица Якоби $\mathsf{T}^{-1}(\mathsf{Z})$ в точке S_0 как $\mathbf{J}_0^-$. По правилу вычисления производной композиции отображений выполняется равенство $\mathbf{J}_0^- \mathbf{J}_0 = \mathbf{I}_{2r}$. Например, в качестве такого диффеоморфизма T можно взять параметризацию $\mathsf{S}(S_{(\tau)}, A_{(\tau)})$ из теоремы 2.2.1, где гладкость отображений объяснена в теореме 2.2.2 и замечании 2.2.1.

Представим оценки в следующем виде:

$$\widehat{\mathsf{S}}(t) = \Pi_{\mathcal{D}_r}^{(\mathsf{S}_0)}(\mathsf{S}_0 + t\xi) = \mathsf{S}_0 + t\Pi_{\mathcal{Z}(A_0^2), \mathbf{w}}(\xi) + \zeta(t), \quad (2.5)$$

$$\widehat{P}(t) = \mathsf{T}^{-1}(\Pi_{\mathcal{D}_r}^{(\mathsf{S}_0)}(\mathsf{S}_0 + t\xi)) = P_0 + t\mathbf{J}_0^- \Pi_{\mathcal{Z}(A_0^2), \mathbf{w}}(\xi) + \eta(t), \quad (2.6)$$

где $\zeta(t)$, $\eta(t)$ — случайные векторы в $\mathbb{R}_N$ и $\mathbb{R}_{2r}$ соответственно, зависящие от t .

Верно следующее:

Теорема 2.5.1. 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:

$$\frac{\zeta(t)}{t} \Rightarrow_{t \rightarrow 0} 0_N,$$

где $\zeta(t)$ определена в (2.5).

Доказательство. Лемма 2.4.3 даёт следующее соотношение для $\Pi_{\mathcal{D}_r}^{(\mathsf{S}_0)}(\mathsf{X})$:

$$\Pi_{\mathcal{D}_r}^{(\mathsf{S}_0)}(\mathsf{S}_0 + \mathsf{Z}) = \mathsf{S}_0 + \Pi_{\mathcal{Z}(A_0^2), \mathbf{w}}(\mathsf{Z}) + \mathbf{o}_{\mathsf{Z} \rightarrow 0}(\|\mathsf{Z}\|)$$

для любых $S_0 + Z \in \mathcal{C}(S_0)$.

Далее достаточно показать, что для любой липшицевой с параметром l ($|f(X) - f(Y)| \leq l\|X - Y\|_{\mathbf{W}}$ для любых X, Y) и ограниченной m ($|f(X)| \leq m$ для любого X) измеримой функции f , заданной на $\mathbb{R}^N$, интегралы

$$I_1(f, t) = \mathbb{E}(|f(G_S(t, \xi)) - f(0_N)|) = \int_{\mathbb{R}^N} |f(G_S(t, Y)) - f(0_N)| d\Xi(Y) \quad (2.7)$$

равномерно по f ограничены последовательностью, стремящейся к 0 при $t \rightarrow 0$ [41, §7], где

$$G_S(t, Y) = \frac{1}{t}(\Pi_{\mathcal{D}_r}^{(S_0)}(S_0 + tY) - S_0 - \Pi_{\mathcal{Z}(A_0^2), \mathbf{W}}(tY)).$$

Зафиксируем $\varepsilon > 0$, и выберем столь большой шар радиуса r_B , что

$$\mathbb{P}(\|\xi\|_{\mathbf{W}} \geq r_B) = \int_{\|Y\|_{\mathbf{W}} \geq r_B} 1 d\Xi(Y) < \varepsilon.$$

Выберем такое малое δ , что

$$\|\Pi_{\mathcal{D}_r}^{(S_0)}(S_0 + Z) - S_0 - \Pi_{\mathcal{Z}(A_0^2), \mathbf{W}}(Z)\|_{\mathbf{W}} \leq \frac{\varepsilon}{r_B} \|Z\|_{\mathbf{W}}$$

для любого $S_0 + Z \in B_{S_0, \delta}$, $B_{S_0, \delta} \subset \mathcal{C}$, $B_{S_0, \delta} = \{Y : \|Y - S_0\|_{\mathbf{W}} < \delta\}$. Возьмём $t = \frac{\delta}{r_B}$. Тогда:

$$\int_{\|Y\|_{\mathbf{W}} < r_B} |f(G_S(t, Y)) - f(0_N)| d\Xi(Y) \leq l\varepsilon.$$

При этом,

$$\int_{\|Y\|_{\mathbf{W}} \geq r_B} |f(G_S(t, Y)) - f(0_N)| d\Xi(Y) < 2m\varepsilon.$$

Таким образом, в (2.7) получаем $I_1(f, t) < l\varepsilon + 2m\varepsilon$, где правую часть неравенства можно сделать сколь угодно малой выбором ε . $\square$

Теорема 2.5.2. Пусть $S_0 \in \mathcal{D}_r$ и управляется ОЛРФ(A_0), ξ — случайный вектор, имеющий распределение Ξ , $\mathsf{T}(P)$ — диффеоморфизм между областью

в $\mathbb{R}^{2r}$ и пересечением некоторой открытой окрестности S_0 с $\mathcal{D}_r$, T^{-1} — обратное отображение к T , $S_0 = T(P_0)$. Тогда имеет место следующая слабая сходимость случайных величин:

$$\frac{\eta(t)}{t} \Rightarrow_{t \rightarrow 0} 0_{2r},$$

где $\eta(t)$ определена в (2.6).

Доказательство. Заметим, что из гладкости T^{-1} следует:

$$T^{-1}(\Pi_{\mathcal{D}_r}^{(S_0)}(S_0 + Z)) = P_0 + J_0^- \Pi_{Z(A_0^2), \mathbf{W}}(Z) + o_{Z \rightarrow 0}(\|Z\|).$$

Далее доказательство теоремы полностью аналогично доказательству теоремы 2.5.1. □

Из Теорем 2.5.1 и 2.5.2 получаем следующую слабую сходимость распределений:

$$\frac{1}{t}(\widehat{S}(t) - S_0) \Rightarrow_{t \rightarrow 0} \Pi_{Z(A_0^2), \mathbf{W}} \xi, \quad (2.8)$$

$$\frac{1}{t}(\widehat{P}(t) - P_0) \Rightarrow_{t \rightarrow 0} J_0^- \Pi_{Z(A_0^2), \mathbf{W}} \xi. \quad (2.9)$$

Эти предельные случайные величины естественно назвать *ошибками первого порядка по t* для оценки сигнала и параметров соответственно.

Теперь рассмотрим некоторую несмещённую оценку сигнала $\widetilde{S}(t)$, построенную по $\mathbf{X} = S_0 + t\xi$, где ξ — случайный вектор с распределением Ξ , и Ξ представляет из себя многомерное нормальное распределение с нулевым средним и ковариационной матрицей $\Sigma = \mathbf{W}^{-1}$. Оценка $\widetilde{S}(t)$ должна удовлетворять условиям регулярности [42]. Тогда выполняется неравенство Рао-Крамера, которое с учётом нормировки запишем в следующем виде:

$$\text{cov} \left(\frac{1}{t} \widetilde{S}(t) \right) \geq J_0 (J_0^T \mathbf{W} J_0)^{-1} J_0^T, \quad (2.10)$$

где $\mathbf{J}_0$ — матрица Якоби параметризации $\mathbf{T}(P)$ в точке P_0 , неравенство выполняется в смысле квадратичных форм [42]. Аналогично, рассмотрим несмещенную оценку $\tilde{P}(t)$ параметра P в параметризации $T(P)$, для которой верно

$$\text{cov} \left(\frac{1}{t} \tilde{P}(t) \right) \geq (\mathbf{J}_0^T \mathbf{W} \mathbf{J}_0)^{-1}.$$

Предложение 2.5.1. Пусть Ξ — многомерный нормальный вектор с нулевым средним и ковариационной матрицей $\Sigma = \mathbf{W}^{-1}$. Тогда оценка $\hat{S}(t)$ является асимптотически эффективной при $t \rightarrow 0$, то есть математическое ожидание её ошибки первого порядка $\Pi_{\mathcal{Z}(A_0^2), \mathbf{W}} \xi$ из (2.8) является нулевым, а ковариационная матрица этой величины достигает границы Рао-Крамера:

$$\text{cov} \Pi_{\mathcal{Z}(A_0^2), \mathbf{W}} \xi = \mathbf{J}_0 (\mathbf{J}_0^T \mathbf{W} \mathbf{J}_0)^{-1} \mathbf{J}_0^T = \Pi_{\mathcal{Z}(A_0^2), \mathbf{W}} \Sigma \Pi_{\mathcal{Z}(A_0^2), \mathbf{W}}^T. \quad (2.11)$$

Аналогично, оценка $\hat{P}(t)$ вместе с её ошибкой первого порядка (2.9) обладает ровно теми же свойствами

$$\text{cov} \mathbf{J}_0^- \Pi_{\mathcal{Z}(A_0^2), \mathbf{W}} \xi = (\mathbf{J}_0^T \mathbf{W} \mathbf{J}_0)^{-1}.$$

Доказательство. Доказательство производится прямой подстановкой явного вида проектора $\Pi_{\mathcal{Z}(A_0^2), \mathbf{W}}$ в (2.11) с учётом того, что $\text{colspace} \mathbf{J}_0 = \mathcal{Z}(A_0^2)$. $\square$

В частном случае $\mathbf{T}(P) = \mathbf{S}(S_{(\tau)}, A_{(\tau)})$ по теореме 2.2.1 мы получаем асимптотическую эффективность оценки параметров ОЛРФ и краевых значений. Возникает вопрос: верно ли это для параметризации вида (1)? Не всегда. Например, если ряд $\mathbf{S}_0$ включает в себя полином p_{m_k} степени больше нулевой, то представление (1) содержит меньше $2r$ параметров, следовательно, не является диффеоморфизмом между открытой областью в $\mathbb{R}^{2r}$ и пересечением окрестности ряда $\mathbf{S}_0$ с $\mathcal{D}_r$.

При исследовании случайных рядов и полученных оценок сигнала возникает задача сравнить полученную ошибку с границей Рао-Крамера для проверки

оптимальности оценок. Обычная мера качества оценки в таком случае — среднеквадратическая ошибка (mean squared error, MSE) по всем элементам ряда.

Предложение 2.5.2. *Рассмотрим ряд $\mathbf{S} = \mathbf{S}_0 + t\xi$, где $\mathbf{S}_0 \in \mathcal{D}_r$ и управляется ОЛРФ(A_0), $A_0 \in \mathbb{R}^{r+1}$, ξ имеет распределение Ξ — многомерное нормальное распределение с нулевым средним и ковариационной матрицей Σ . Тогда граница Рао-Крамера для средней по точкам ряда среднеквадратической ошибки $\mathbb{E}(\frac{1}{N}\|\hat{\mathbf{S}} - \mathbf{S}_0\|^2)$ оценки $\hat{\mathbf{S}}$ ряда $\mathbf{S}_0$ равна*

$$\frac{t^2}{N} \text{tr} \Pi_{\mathcal{Z}(A_0^2), \mathbf{W}} \Sigma \Pi_{\mathcal{Z}(A_0^2), \mathbf{W}} = \frac{t^2}{N} \text{tr}(\mathbf{Z}(A_0^2)(\mathbf{Z}^T(A_0^2)\Sigma^{-1}\mathbf{Z}(A_0^2))^{-1}\mathbf{Z}^T(A_0^2)), \quad (2.12)$$

где $\mathbf{W} = \Sigma^{-1}$.

В частном случае $\Sigma = \sigma^2 \mathbf{I}_N$ эта величина равна

$$\frac{t^2}{N} \text{tr} \Pi_{\mathcal{Z}(A_0^2), \mathbf{W}} \Sigma \Pi_{\mathcal{Z}(A_0^2), \mathbf{W}}^T = \frac{2rt^2\sigma^2}{N}. \quad (2.13)$$

Доказательство. Формула (2.12) является следствием формулы (2.11).

Докажем формулу (2.13). Так как

$$\Pi_{\mathcal{Z}(A_0^2), \mathbf{I}_N} \mathbf{I}_N \Pi_{\mathcal{Z}(A_0^2), \mathbf{I}_N} = \Pi_{\mathcal{Z}(A_0^2), \mathbf{I}_N}$$

и след матрицы-проектора $\Pi_{\mathcal{Z}(A_0^2), \mathbf{I}_N}$ равен размерности $2r$ подпространства, на которое идёт проекция, получаем требуемое. $\square$

Глава 3

Численные методы решения задачи аппроксимации временных рядов рядами конечного ранга

В этой главе мы рассмотрим методы локального поиска для решения задачи (3), с использованием построенной в главе 2 (раздел 2.2) параметризации $\mathcal{D}_r$.

3.1. Методы локальной оптимизации

Вначале мы опишем известные методы локальной оптимизации в терминах, введённых в главе 2, а в разделе 3.1.3 предложим новый метод численного решения задачи (3). Разделы 3.1.4–3.1.6 носят частично технический характер, и часть доказательств перенесена в них.

3.1.1. Взвешенный метод Гаусса-Ньютона для задачи аппроксимации рядом конечного ранга

Рассмотрим модификацию стандартного метода Гаусса-Ньютона (1.6), где параметризация $S(P) = \mathbf{S}(P)$ с $P = (S_{(\tau)}, A_{(\tau)})$, определённая в теореме 2.2.1 и зависящая от индекса τ , меняется на каждой итерации следующим образом. На $(k + 1)$ -м шаге параметризация строится в окрестности $A_0 = A^{(k)}$. Индекс τ , который определяет параметризацию, выбирается соответствующим максимальному по модулю значению A_0 . Так как параметризация не зависит от умножения A_0 на константу, всегда можно считать, что $a_\tau^{(0)} = -1$ и $|a_i^{(0)}| \leq 1$ для любого i , $1 \leq i \leq r + 1$.

Получаем следующий шаг:

$$P_{k+1} = P_k + \gamma (\mathbf{J}_S(P_k))_{\mathbf{W}}^\dagger (\mathbf{X} - \mathbf{S}(P_k)). \quad (3.1)$$

Чтобы применить метод, необходимо вычислить матрицу Якоби $\mathbf{J}_S(P_k)$ и значение $\mathbf{S}(S_{(\tau)}, A_{(\tau)})$. Формально, мы можем произвести эти вычисления; с другой стороны, прямое вычисление неустойчиво и очень затратно по времени.

3.1.2. Метод Variable projection для задачи аппроксимации рядом конечного ранга (метод VPGN)

Так как $S(P)$ при решении (3) имеет вид (1.7), применим принцип Variable projection. А именно, $S(P) = \mathbf{S}(S_{(\tau)}, A_{(\tau)})$, и согласно (2.1), в $\mathbf{S}(S_{(\tau)}, A_{(\tau)})$ часть параметров $S_{(\tau)}$ входит линейным образом.

Так же, как и в (2.1), мы рассмотрим задачу (3) в окрестности ряда $\mathbf{S}_0 \in \mathcal{D}_r$, и предположим, что ряд управляется ОЛРФ(A_0) такой, что $a_\tau^{(0)} = -1$. Заметим, что в окрестности A_0 вектор A однозначно задаётся с помощью $A_{(\tau)}$, и, наоборот, $A_{(\tau)} \in \mathbb{R}^r$ расширяется до $A \in \mathbb{R}^{r+1}$ с помощью вставки -1 на τ -ю позицию.

Подставим в (1.9) $\mathcal{D} = \overline{\mathcal{D}}_r$, $\mathcal{D}^* = \mathcal{D}_r^* \subset \overline{\mathcal{D}}_r$, $B = A_{(\tau)}$, $\mathbf{G}(B) = \mathbf{G}$, где $\mathbf{G} = \left(\mathbf{Z} \mathbf{Z}_{I(\tau)}^{-1}, : \right)$ (см. (2.1)), $C^*(B) = (\mathbf{G})_{\mathbf{W}}^\dagger \mathbf{X}$, $\mathbf{G}(B)C^*(B) = \Pi_{\mathcal{Z}(A), \mathbf{W}} \mathbf{X} \stackrel{\text{def}}{=} \mathbf{S}^*(A_{(\tau)})$. Тогда мы получаем задачу, эквивалентную проекции элемента на подмножество $\mathcal{D}_r^*$ вместо $\overline{\mathcal{D}}_r$, в которой параметры $S_{(\tau)}$ исключены:

$$\mathbf{Y}^* = \arg \min_{\mathbf{Y} \in \mathcal{D}_r^*} \|\mathbf{X} - \mathbf{Y}\|_{\mathbf{W}} \quad \text{с} \quad \mathcal{D}_r^* = \{\Pi_{\mathcal{Z}(A), \mathbf{W}} \mathbf{X} | A_{(\tau)} \in \mathbb{R}^r\}. \quad (3.2)$$

Перепишем задачу (3.2) только в виде задачи поиска параметра $A_{(\tau)}$:

$$A_{(\tau)}^* = \arg \min_{A_{(\tau)} \in \mathbb{R}^r} \|\mathbf{X} - \mathbf{S}^*(A_{(\tau)})\|_{\mathbf{W}}. \quad (3.3)$$

Таким образом, для численного решения (3), достаточно рассмотреть итерации по нелинейной части параметров. В точности этот подход использован в [16, 17].

Обозначим $\mathbf{J}_{\mathbf{S}^*}(A_{(\tau)})$ матрицу Якоби отображения $\mathbf{S}^*(A_{(\tau)})$, тогда итерация Гаусса-Ньютона для (3.3) выглядит как:

$$A_{(\tau)}^{(k+1)} = A_{(\tau)}^{(k)} + \gamma \left(\mathbf{J}_{\mathbf{S}^*}(A_{(\tau)}^{(k)}) \right)_{\mathbf{W}}^{\dagger} (\mathbf{X} - \mathbf{S}_{\mathbf{X}}^*(A_{(\tau)}^{(k)})), \quad (3.4)$$

где $\mathbf{J}_{\mathbf{S}^*}(A_{(\tau)}^{(k)})$ имеет вид (3.8), приведённый в разделе 3.1.4.

3.1.3. Модификация метода Variable projection для задачи аппроксимации рядом конечного ранга (метод MGN)

Предложим новый метод решения задачи (3) с помощью модификации итерации Гаусса-Ньютона.

Вернёмся к задаче с полным набором параметров $(S_{(\tau)}, A_{(\tau)})$ и применим подход, предложенный в замечании 1.3.1 с $\tilde{S}(P) = \mathbf{S}^*(A_{(\tau)})$. Это можно сделать, так как (1.8) выполнено с $\mathcal{G}(B) = \mathcal{Z}(A)$. Таким образом, мы можем рассмотреть $\mathbf{S}^*(A_{(\tau)}^{(k+1)}) \in \mathcal{D}_r^*$ как результат $(k+1)$ -й итерации вместо $\mathbf{S}(S_{(\tau)}^{(k+1)}, A_{(\tau)}^{(k+1)}) \in \overline{\mathcal{D}}_r$. Оказывается (см. раздел 3.2.1), что тогда мы можем производить более устойчивые вычисления. Предложенная модификация схожа с методом VPGN, использующим метод Variable projection (сравните с (3.4)), так как мы можем отбросить часть параметров $S_{(\tau)}$ и получить следующую итерацию:

$$A_{(\tau)}^{(k+1)} = A_{(\tau)}^{(k)} + \gamma \left(\left(\mathbf{J}_{\mathbf{S}}(S_{(\tau)}^{(k)}, A_{(\tau)}^{(k)}) \right)_{\mathbf{W}}^{\dagger} (\mathbf{X} - \mathbf{S}^*(A_{(\tau)}^{(k)})) \right)_{:, \{r+1, \dots, 2r\}}, \quad (3.5)$$

где $S_{(\tau)}^{(k)}$ взяты как соответствующие краевые значения $\mathbf{S}^*(A_{(\tau)}^{(k)})$.

Как и в методе Variable projection, $\mathbf{S}^*(A_{(\tau)}^{(k+1)}) \in \mathcal{D}_r^*$ для каждого k .

Следующая теорема 3.1.1 даёт вид итерации (3.5), подходящей для устойчивой реализации.

Теорема 3.1.1. *Итерация (3.5) эквивалентна следующей:*

$$A_{(\tau)}^{(k+1)} = A_{(\tau)}^{(k)} + \gamma \mathbf{M}_k^{\dagger} \mathbf{Q}^T(A_{(\tau)}^{(k)}) \mathbf{\Pi}_{\mathcal{Z}((A_{(\tau)}^{(k)})^2), \mathbf{W}} (\mathbf{X} - \mathbf{\Pi}_{\mathcal{Z}(A_{(\tau)}^{(k)}), \mathbf{W}}(\mathbf{X})), \quad (3.6)$$

где $\mathbf{M}_k = -\mathbf{S}_{\mathcal{K}(\tau), :}^T$, $\mathbf{S} = \mathcal{T}_{r+1} \left(\mathbf{S}^*(A_{(\tau)}^{(k)}) \right)$.

Теорема доказана в техническом разделе 3.1.5.

Итак, мы построили модификацию (3.6) шага (3.5) таким образом, чтобы иметь возможность уменьшить сложность вычислений, сведя основную часть к вычислению проекций на $\mathcal{Z}(A)$ и $\mathcal{Z}(A^2)$. Алгоритмы для их вычисления обсуждаются в разделе 3.2.1. Полный алгоритм, соответствующий предложенному методу — алгоритм 3.3.5.

Раздел 3.4 завершают три подраздела с техническими деталями относительно вида итерационных шагов методов VPGN, MGN и их сравнения.

3.1.4. Вычислительная форма метода Variable Projection (VPGN)

Вид шага (3.4) в явном виде содержится в [17, Proposition 3]. Предложим иной вид матрицы Якоби $\mathbf{J}_{S^*}$, более удобный для реализации и сравнения методов MGN и VPGN.

Лемма 3.1.1. *Верно следующее:*

$$\mathbf{\Pi}_{\mathcal{Z}(A), \mathbf{W}} \mathbf{X} = \mathbf{W}^{-1} \mathbf{Q}(A) \mathbf{\Gamma}^{-1}(A) \mathbf{Q}^T(A) \mathbf{X}, \quad (3.7)$$

где $\mathbf{\Gamma}(A) = \mathbf{Q}^T(A) \mathbf{W}^{-1} \mathbf{Q}(A)$, при этом матрица Якоби $\mathbf{J}_{S^*}(A_{(\tau)})$ отображения $S^*(A_{(\tau)}) = \mathbf{\Pi}_{\mathcal{Z}(A), \mathbf{W}} \mathbf{X}$ задана соотношением

$$\begin{aligned} (\mathbf{J}_{S^*}(A_{(\tau)}))_{:,i} = & -\mathbf{W}^{-1} \mathbf{Q}(A) \mathbf{\Gamma}^{-1}(A) \mathbf{Q}^T(E_j) \mathbf{\Pi}_{\mathcal{Z}(A), \mathbf{W}} \mathbf{X} - \\ & - \mathbf{\Pi}_{\mathcal{Z}(A), \mathbf{W}} \mathbf{W}^{-1} \mathbf{Q}(E_j) \mathbf{\Gamma}^{-1}(A) \mathbf{Q}^T(A) \mathbf{X}, \end{aligned} \quad (3.8)$$

где $j = (\mathcal{K}(\tau))_i$ — i -й элемент множества $\mathcal{K}(\tau)$.

Доказательство. Приведём доказательство первой формулы. Равенство

$$S^*(A_{(\tau)}) = \mathbf{\Pi}_{\mathcal{Z}(A), \mathbf{W}}(\mathbf{X}) \in \mathcal{D}_r^*$$

соответствует решению следующей квадратичной задачи:

$$S^*(A_{(\tau)}) = \arg \min_{\substack{\mathbf{Y} \\ \mathbf{Q}^T(A) \mathbf{Y} = 0}} \left(\frac{1}{2} \mathbf{Y}^T \mathbf{W} \mathbf{Y} - \mathbf{Y} \mathbf{W} \mathbf{X} \right),$$

что, используя формулы [35, (16.4), (16.15)], даёт нам выражение (3.7).

Доказательство формулы (3.8) строится путём взятия производной выражения (3.7) по a_k и применением следующей цепочки равенств:

$$\begin{aligned}
(\Pi_{Z(A), \mathbf{W}} \mathbf{X})'_{a_k} &= -\mathbf{W}^{-1} \mathbf{Q}(E_k) (\mathbf{Q}^T(A) \mathbf{W}^{-1} \mathbf{Q}(A))^{-1} \mathbf{Q}^T(A) \mathbf{X} - \\
&- \mathbf{W}^{-1} \mathbf{Q}(A) (\mathbf{Q}^T(A) \mathbf{W}^{-1} \mathbf{Q}(A))^{-1} \mathbf{Q}^T(E_k) \mathbf{X} + \mathbf{W}^{-1} \mathbf{Q}(A) (\mathbf{Q}^T(A) \mathbf{W}^{-1} \mathbf{Q}(A))^{-1} \cdot \\
&\cdot (\mathbf{Q}^T(E_k) \mathbf{W}^{-1} \mathbf{Q}(A) + \mathbf{Q}^T(A) \mathbf{W}^{-1} \mathbf{Q}(E_k)) (\mathbf{Q}^T(A) \mathbf{W}^{-1} \mathbf{Q}(A))^{-1} \mathbf{Q}^T(A) \mathbf{X} = \\
&= -\mathbf{W}^{-1} \mathbf{Q}(A) (\mathbf{Q}^T(A) \mathbf{W}^{-1} \mathbf{Q}(A))^{-1} \mathbf{Q}^T(E_k) \Pi_{Z(A), \mathbf{W}} \mathbf{X} - \\
&- \Pi_{Z(A), \mathbf{W}} \mathbf{W}^{-1} \mathbf{Q}(E_k) (\mathbf{Q}^T(A) \mathbf{W}^{-1} \mathbf{Q}(A))^{-1} \mathbf{Q}^T(A) \mathbf{X}.
\end{aligned}$$

□

3.1.5. Эквивалентные формы итераций для метода MGN

Введём следующие обозначения. Пусть $\mathbf{F}_{S,k} = (\mathbf{J}_{S,k})_{:, \{1, \dots, r\}}$ обозначает первые r столбцов матрицы Якоби $\mathbf{J}_S(S_{(\tau)}^{(k)}, A_{(\tau)}^{(k)}) = \mathbf{J}_{S,k}$, $\mathbf{F}_{A,k} = (\mathbf{J}_{S,k})_{:, \{r+1, \dots, 2r\}}$ обозначает последние r столбцов матрицы Якоби $\mathbf{J}_{S,k}$.

Докажем теорему 3.1.1 об эквивалентном виде итераций метода MGN.

Доказательство теоремы 3.1.1. Необходимо доказать эквивалентность (3.5) и (3.6). Следующее равенство верно для любой $\mathbf{F}$ полного ранга и положительно определённой $\mathbf{V}$: $(\mathbf{F})_{\mathbf{V}}^{\dagger} \Pi_{\mathbf{F}, \mathbf{W}} = (\mathbf{F}^T \mathbf{V} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{V} \mathbf{F} (\mathbf{F}^T \mathbf{W} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{W} = (\mathbf{F})_{\mathbf{W}}^{\dagger}$. Согласно теореме 2.2.3, $\text{colspace } \mathbf{J}_{S,k} = \mathcal{Z}((A^{(k)})^2)$, и, таким образом, $\Pi_{\mathbf{J}_{S,k}, \mathbf{W}} = \Pi_{\mathcal{Z}((A^{(k)})^2), \mathbf{W}}$.

Выберем матрицу $\mathbf{V}$ нужным образом. Определим $\hat{\mathbf{Q}}_k = \hat{\mathbf{Q}} = \begin{pmatrix} \mathbf{Q}^T(A^{(k)}) \\ \mathbf{F}_{S,k}^T \end{pmatrix}$ — матрицу полного ранга (см лемму 2.2.1), и положим $\mathbf{V} = \hat{\mathbf{Q}}^T \hat{\mathbf{Q}}$. Докажем, что

$$\begin{aligned}
&((\mathbf{J}_{S,k})_{\mathbf{V}}^{\dagger} \Pi_{\mathcal{Z}((A^{(k)})^2), \mathbf{W}} (\mathbf{X} - \Pi_{\mathcal{Z}(A^{(k)}), \mathbf{W}} (\mathbf{X})))_{:, \{r+1, \dots, 2r\}} = \\
&= \mathbf{M}_k^{\dagger} \mathbf{Q}^T(A^{(k)}) \Pi_{\mathcal{Z}((A^{(k)})^2), \mathbf{W}} (\mathbf{X} - \Pi_{\mathcal{Z}(A^{(k)}), \mathbf{W}} (\mathbf{X})),
\end{aligned}$$

где $\mathbf{M}_k = \mathbf{Q}^T(A^{(k)})\mathbf{F}_{A,k}$ согласно лемме 2.2.2.

Мы доказываем это равенство, используя технику блочных матриц. Заметим, что вычисление нижней половины строк матрицы $(\mathbf{J}_{S,k})_{\mathbf{V}}^{\dagger}$ требует вычисление только нижней половины строк матрицы $(\mathbf{J}_{S,k}^T \mathbf{V} \mathbf{J}_{S,k})^{-1}$.

$$\text{Таким образом, мы имеем } \widehat{\mathbf{Q}}\mathbf{J}_{S,k} = \begin{pmatrix} \mathbf{0}_{N-r,r} & \mathbf{M}_k \\ \mathbf{F}_{S,k}^T \mathbf{F}_{S,k} & \mathbf{F}_{S,k}^T \mathbf{F}_{A,k} \end{pmatrix},$$

$$\mathbf{J}_{S,k}^T \mathbf{V} \mathbf{J}_{S,k} = \begin{pmatrix} (\mathbf{F}_{S,k}^T \mathbf{F}_{S,k})^T \mathbf{F}_{S,k}^T \mathbf{F}_{S,k} & (\mathbf{F}_{S,k}^T \mathbf{F}_{S,k})^T \mathbf{F}_{S,k}^T \mathbf{F}_{A,k} \\ (\mathbf{F}_{S,k}^T \mathbf{F}_{A,k})^T \mathbf{F}_{S,k}^T \mathbf{F}_{S,k} & \mathbf{M}_k^T \mathbf{M}_k + (\mathbf{F}_{A,k}^T \mathbf{F}_{A,k})^T \mathbf{F}_{A,k}^T \mathbf{F}_{A,k} \end{pmatrix},$$

$$\begin{aligned} ((\mathbf{J}_{S,k}^T \mathbf{V} \mathbf{J}_{S,k})^{-1})_{:, \{r+1, \dots, 2r\}} &= \\ &= \begin{pmatrix} -(\mathbf{M}_k^T \mathbf{M}_k)^{-1} (\mathbf{F}_{S,k}^T \mathbf{F}_{A,k}) ((\mathbf{F}_{S,k}^T \mathbf{F}_{S,k})^T)^{-1} & (\mathbf{M}_k^T \mathbf{M}_k)^{-1} \end{pmatrix}, \end{aligned}$$

$$\mathbf{V} = \mathbf{Q}(A^{(k)})\mathbf{Q}^T(A^{(k)}) + \mathbf{F}_{S,k}\mathbf{F}_{S,k}^T, \quad \mathbf{J}_{S,k}^T \mathbf{V} = \begin{pmatrix} (\mathbf{F}_{S,k}^T \mathbf{F}_{S,k})^T \mathbf{F}_{S,k}^T \\ \mathbf{M}_k \mathbf{Q}^T(A^{(k)}) + (\mathbf{F}_{S,k}^T \mathbf{F}_{A,k})^T \mathbf{F}_{S,k}^T \end{pmatrix},$$

$$\begin{aligned} ((\mathbf{J}_{S,k}^T \mathbf{V} \mathbf{J}_{S,k})^{-1})_{:, \{r+1, \dots, 2r\}} \mathbf{J}_{S,k}^T \mathbf{V} &= ((\mathbf{J}_{S,k}^T \mathbf{V} \mathbf{J}_{S,k})^{-1} \mathbf{J}_{S,k}^T \mathbf{V})_{:, \{r+1, \dots, 2r\}} = \\ &= (\mathbf{M}_k^T \mathbf{M}_k)^{-1} \mathbf{M}_k \mathbf{Q}^T(A^{(k)}) = \mathbf{M}_k^{\dagger} \mathbf{Q}^T(A^{(k)}). \end{aligned}$$

□

Для реализации метода MGN используется именно вид итерации (3.6). Нам также понадобится другой вид итерации (3.5) для сравнения с методом VPGN. Определим

$$\mathbf{H}_k = (\mathbf{I}_N - \mathbf{\Pi}_{\mathcal{Z}(A^{(k)}), \mathbf{W}}) \mathbf{F}_{A,k}, \quad (3.9)$$

и рассмотрим следующую итерацию:

$$A_{(\tau)}^{(k+1)} = A_{(\tau)}^{(k)} + \gamma (\mathbf{H}_k)_{\mathbf{W}}^{\dagger} (\mathbf{X} - \mathbf{\Pi}_{\mathcal{Z}(A^{(k)}), \mathbf{W}}(\mathbf{X})). \quad (3.10)$$

Предложение 3.1.1. *Итерации (3.5) и (3.10) эквивалентны.*

Доказательство. Покажем, что

$$((\mathbf{J}_{S,k})_{\mathbf{W}}^{\dagger})_{:, \{r+1, \dots, 2r\}} = ((\mathbf{I}_N - \mathbf{\Pi}_{\mathcal{Z}(A^{(k)})}, \mathbf{W}) \mathbf{F}_{A,k})_{\mathbf{W}}^{\dagger}. \quad (3.11)$$

Чтобы это сделать, нужно представить $(\mathbf{J}_{S,k})_{\mathbf{W}}^{\dagger}$ в следующем виде:

$$(\mathbf{J}_{S,k})_{\mathbf{W}}^{\dagger} = \begin{pmatrix} \mathbf{F}_{S,k}^T \mathbf{W} \mathbf{F}_{S,k} & \mathbf{F}_{S,k}^T \mathbf{W} \mathbf{F}_{A,k} \\ \mathbf{F}_{A,k}^T \mathbf{W} \mathbf{F}_{S,k} & \mathbf{F}_{A,k}^T \mathbf{W} \mathbf{F}_{A,k} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{F}_{S,k}^T \mathbf{W} \\ \mathbf{F}_{A,k}^T \mathbf{W} \end{pmatrix}. \quad (3.12)$$

Вычислим нижнюю половину (3.12):

$$\begin{aligned} ((\mathbf{J}_{S,k})_{\mathbf{W}}^{\dagger})_{:, \{r+1, \dots, 2r\}} &= \\ &= (\mathbf{F}_{A,k}^T \mathbf{W} \mathbf{F}_{A,k} - \mathbf{F}_{A,k}^T \mathbf{\Pi}_{\mathbf{F}_{S,k}, \mathbf{W}}^T \mathbf{W} \mathbf{F}_{A,k})^{-1} (\mathbf{F}_{A,k}^T - \mathbf{F}_{A,k}^T \mathbf{\Pi}_{\mathbf{F}_{S,k}, \mathbf{W}}^T) \mathbf{W} = \\ &= (((\mathbf{I}_N - \mathbf{\Pi}_{\mathbf{F}_{S,k}, \mathbf{W}}) \mathbf{F}_{A,k})^T \mathbf{W} ((\mathbf{I}_N - \mathbf{\Pi}_{\mathbf{F}_{S,k}, \mathbf{W}}) \mathbf{F}_{A,k}))^{-1} ((\mathbf{I}_N - \mathbf{\Pi}_{\mathbf{F}_{S,k}, \mathbf{W}}) \mathbf{F}_{A,k})^T \mathbf{W}, \end{aligned}$$

так как $\mathbf{\Pi}_{\mathbf{F}_{S,k}, \mathbf{W}}^T \mathbf{W} \mathbf{\Pi}_{\mathbf{F}_{S,k}, \mathbf{W}}^T = \mathbf{\Pi}_{\mathbf{F}_{S,k}, \mathbf{W}}^T \mathbf{W}$. С учётом $\mathbf{\Pi}_{\mathcal{Z}(A^{(k)})}, \mathbf{W} = \mathbf{\Pi}_{\mathbf{F}_{S,k}, \mathbf{W}}$ получаем правую часть (3.11). $\square$

Также определим матрицу $\tilde{\mathbf{Q}}_k = \begin{pmatrix} \mathbf{Q}^T(A^{(k)}) \\ \mathbf{F}_{S,k}^T \mathbf{W} \end{pmatrix}$. Эта матрица невырождена, так как $\tilde{\mathbf{Q}}_k^T = [\mathbf{Q}(A^{(k)}) : \mathbf{W} \mathbf{F}_{S,k}]$, где $\text{colspace } \mathbf{W} \mathbf{F}_{S,k} = \mathcal{Z}(A^{(k)})$ из-за положительной определённости $\mathbf{W}$.

Для сравнения итерации метода MGN с итерацией метода VPGN нам понадобится следующее свойство.

Лемма 3.1.2. Для $\mathbf{H}_k$, заданной в (3.9), выполнено следующее равенство:

$$\mathbf{H}_k = \tilde{\mathbf{Q}}_k^{-1} \begin{pmatrix} -\mathbf{S}_{\mathcal{K},:}^T \\ \mathbf{0}_{r,r} \end{pmatrix}, \text{ где } \mathcal{K} = \mathcal{K}(\tau).$$

Доказательство. Покажем, что $\tilde{\mathbf{Q}}_k \mathbf{H}_k = \begin{pmatrix} -\mathbf{S}_{\mathcal{K},:}^T \\ \mathbf{0}_{r,r} \end{pmatrix}$, вычислив отдельно произведение верхних и нижних строк. Так как $\mathbf{Q}^T(A^{(k)}) \mathbf{\Pi}_{\mathcal{Z}(A^{(k)})}, \mathbf{W} = \mathbf{0}_{N-r, N}$, для

верхних $N - r$ строк имеем $\mathbf{Q}^T(A^{(k)})(\mathbf{I}_N - \mathbf{\Pi}_{\mathcal{Z}(A^{(k)})}, \mathbf{W})\mathbf{F}_{A,k} = \mathbf{Q}^T(A^{(k)})\mathbf{F}_{A,k} = -\mathbf{S}_{\mathcal{K},:}^T$.

Для нижних r строк получаем $\mathbf{F}_{S,k}^T \mathbf{W}(\mathbf{I}_N - \mathbf{\Pi}_{\mathcal{Z}(A^{(k)})}, \mathbf{W})\mathbf{F}_{A,k} = (\mathbf{F}_{S,k}^T \mathbf{W} - \mathbf{F}_{S,k}^T \mathbf{W})\mathbf{F}_{A,k} = \mathbf{0}_{r,r}$. $\square$

3.1.6. Сравнение вида итераций методов VPGN и MGN

Представим матрицу Якоби $\mathbf{J}_{S^*}(A_{(\tau)}^{(k)})$, заданную в (3.8), в виде суммы двух частей: $\mathbf{J}_{S^*}(A_{(\tau)}^{(k)}) = \tilde{\mathbf{H}}(A_{(\tau)}^{(k)}) + \hat{\mathbf{H}}(A_{(\tau)}^{(k)}) = \tilde{\mathbf{H}}_k + \hat{\mathbf{H}}_k$, где

$$(\tilde{\mathbf{H}}(A_{(\tau)}^{(k)}))_{:,i} = -\mathbf{W}^{-1}\mathbf{Q}(A^{(k)})\mathbf{\Gamma}^{-1}(A^{(k)})\mathbf{Q}^T(E_j)\mathbf{\Pi}_{\mathcal{Z}(A^{(k)})}, \mathbf{W}\mathbf{X},$$

$$(\hat{\mathbf{H}}(A_{(\tau)}^{(k)}))_{:,i} = -\mathbf{\Pi}_{\mathcal{Z}(A^{(k)})}, \mathbf{W}\mathbf{W}^{-1}\mathbf{Q}(E_j)\mathbf{\Gamma}^{-1}(A^{(k)})\mathbf{Q}^T(A^{(k)})\mathbf{X},$$

$j = \mathcal{K}(\tau)_i = \mathcal{K}_i$ — i -й элемент множества $\mathcal{K}$.

Лемма 3.1.3. *Выполняется следующее равенство: $\mathbf{H}_k = \tilde{\mathbf{H}}_k$.*

Доказательство. Докажем, что $\tilde{\mathbf{Q}}_k \tilde{\mathbf{H}}_k = \begin{pmatrix} -\mathbf{S}_{\mathcal{K},:}^T \\ \mathbf{0}_{r,r} \end{pmatrix}$, а затем используем результат леммы 3.1.2. Мы имеем $(\mathbf{Q}(A^{(k)})^T \tilde{\mathbf{H}}_k)_{:,i} = -\mathbf{Q}^T(E_j)\mathbf{\Pi}_{\mathcal{Z}(A^{(k)})}, \mathbf{W}\mathbf{X}$, следовательно, $\mathbf{Q}(A^{(k)})^T \tilde{\mathbf{H}}_k = -\mathbf{S}_{\mathcal{K},:}^T$, а также $\mathbf{F}_{S,k}^T \mathbf{W} \tilde{\mathbf{H}}_k = \mathbf{0}_{r,r}$, что доказывает утверждение. $\square$

Применим лемму 3.1.3 к шагу (3.4).

Предложение 3.1.2. *Следующая итерация эквивалентна (3.4):*

$$A_{(\tau)}^{(k+1)} = A_{(\tau)}^{(k)} + \gamma \left(\mathbf{H}(A_{(\tau)}^{(k)}) + \hat{\mathbf{H}}(A_{(\tau)}^{(k)}) \right)_{\mathbf{W}}^{\dagger} (\mathbf{X} - \mathbf{S}^*(A_{(\tau)}^{(k)})). \quad (3.13)$$

Замечание 3.1.1. *Предложение 3.1.1 и следствие 3.1.2 показывают, что итерация метода VPGN в виде (3.13) отличается от итерации метода MGN в виде (3.10) псевдообращением матрицы $\mathbf{H}_k + \hat{\mathbf{H}}_k$ вместо $\mathbf{H}_k$, при этом $\text{colspace } \mathbf{H}_k \subset \mathcal{Z}(A^{(k)})$.*

3.2. Вычисление базисов пространств $\mathcal{Z}(A)$ и $\mathcal{Z}(A^2)$

3.2.1. Общая теория

Для реализации шага итерации (3.6) предложенного алгоритма оптимизации MGN, нам необходимы эффективные алгоритмы вычисления ортонормированных базисов $\mathcal{Z}(A)$ и $\mathcal{Z}(A^2)$, где $\mathcal{Z}(B)$ обозначает пространство рядов, управляемых ОЛРФ(B). Заодно, мы можем использовать полученные алгоритмы для улучшения устойчивости шага (3.4) алгоритма VPGN. Как и ранее, под $\mathbf{Z}(B)$ мы имеем в виду матрицу, столбцы которой составляют некоторый базис $\mathcal{Z}(B)$. Нас интересует построение ортонормированных базисов для использования их при численном вычислении проекции на $\mathcal{Z}(B)$, чтобы получать более точные значения проекций.

Начнём с построения $\mathbf{Z}(A)$. Несмотря на то, что рассматриваемые ряды вещественнозначные, мы конструируем комплекснозначный базис, так как такая замена не влияет на проекцию $\mathbf{P}_{\mathcal{Z}(A), \mathbf{W}} X$ для любого вещественного вектора V и вещественной матрицы $\mathbf{W}$. Таким образом, мы хотим найти невырожденную матрицу $\mathbf{Z}(A) = \mathbf{Z} \in \mathbb{C}^{N \times r}$, которая бы удовлетворяла равенству $\mathbf{Q}^T(A)\mathbf{Z} = 0_{N-r}$.

Введём следующую матрицу-циркулянт $\mathbf{C}(A)$, зависящую от $A \in \mathbb{R}^{r+1}$:

$$\mathbf{C}(A) = \begin{pmatrix} a_1 & a_2 & \dots & \dots & a_{r+1} & 0 & \dots & 0 \\ 0 & a_1 & a_2 & \dots & \dots & a_{r+1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & a_1 & a_2 & \ddots & \ddots & a_{r+1} \\ a_{r+1} & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ a_3 & \dots & a_{r+1} & 0 & \dots & 0 & a_1 & a_2 \\ a_2 & \dots & \dots & a_{r+1} & 0 & \dots & 0 & a_1 \end{pmatrix}, \quad (3.14)$$

которую можно получить расширением матрицы $\mathbf{Q}^T(A)$. Получим $V \in \mathcal{Z}(A)$ тогда и только тогда, когда $\mathbf{C}(A)V \in \text{span}(E_{N-r+1}, \dots, E_N)$. Если $\mathbf{C}(A)$ невырождена, то мы можем найти базисные векторы V_k путём решения систем линейных уравнений $\mathbf{C}(A)V_k = E_{N-k+1}$, $k = 1, \dots, r$, с помощью дискретного преобразования Фурье [43], и затем применить ортонормализацию к $\mathbf{V}_r = [V_1 : \dots : V_r]$.

Обозначим $\mathcal{F}_N$ и $\mathcal{F}_N^{-1}$ прямое и обратное дискретное преобразование Фурье длины N . Для $\mathcal{F}_N(X) = Y$, $X, Y \in \mathbb{C}^N$, $X = (x_0, \dots, x_{N-1})^T$, $Y = (y_0, \dots, y_{N-1})^T$ выполняется $y_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j \exp(-\frac{i2\pi kj}{N})$. При этом определим преобразование Фурье от матрицы как $\mathcal{F}_N(\mathbf{X}) = [\mathcal{F}_N(X_1) : \dots : \mathcal{F}_N(X_r)]$, где $\mathbf{X} = [X_1 : \dots : X_r]$, и $\mathcal{F}_N^{-1}(\mathbf{Y})$ аналогично.

Следующая лемма — прямое применение теоремы о решении системы линейных уравнений с матрицей-циркулянтотом [43]. Обозначим g_A — характеристический многочлен ОЛРФ(A), см. определение 1.1.4.

Лемма 3.2.1. *Обозначим $\mathbf{V}_r = \mathcal{F}_N^{-1}(\mathbf{A}_g^{-1}\mathbf{R}_r)$, где $\mathbf{R}_r = \mathcal{F}_N([E_{N-r+1} : \dots : E_N])$, $\mathbf{A}_g = \text{diag}((g_A(\omega_0), \dots, g_A(\omega_{N-1})))^T$, $\omega_j = \exp(\frac{i2\pi j}{N})$. Тогда $\mathbf{Q}^T(A)\mathbf{V}_r = \mathbf{0}_{N-r,r}$, при этом диагональ матрицы $\mathbf{A}_g$ содержит собственные числа матрицы $\mathbf{C}(A)$.*

Замечание 3.2.1. 1. В качестве матрицы, содержащей базис $\mathcal{Z}(A)$, можно взять матрицу $\mathbf{Z}$, столбцы которой построены с помощью ортонормализации столбцов матрицы $\mathbf{V}_r$, заданной в лемме 3.2.1. Будем обозначать это преобразование как $\mathbf{Z} = \text{orthonorm}(\mathbf{V}_r)$. Так как $\mathbf{Q}^T(A)\mathbf{V}_r = \mathbf{0}_{N-r,r}$, то и $\mathbf{Q}^T(A)\mathbf{Z} = \mathbf{0}_{N-r,r}$.

2. Так как $\mathcal{F}_N^{-1}$ — преобразование, сохраняющее ортонормированность, матрица $\mathbf{Z} = \mathcal{F}_N^{-1}(\text{orthonorm}(\mathbf{A}_g^{-1}\mathbf{R}_r))$ состоит из ортонормированных столбцов.

Тем не менее, $\mathbf{C}(A)$ может оказаться вырожденной, например в случае линейных временных рядов с $s_n = an + b$, управляемых ОЛРФ(A), $A = (1, -2, 1)^T$.

Лемма 3.2.1 показывает, что собственные числа $\mathbf{C}(A)$ — значения многочлена $g_A(z)$ в узлах равномерной решётки на комплексной единичной окружности $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$: $\mathcal{W} = \{\exp(\frac{i2\pi j}{N}), j = 0, \dots, N-1\}$, $\mathcal{W} \subset \mathbb{T}$. Следовательно, невырожденность $\mathbf{C}(A)$ эквивалентна отсутствию корней полиномов $g_A(z)$ на $\mathcal{W}$. Следующая лемма помогает избежать проблем с нулевыми собственными числами. Определим унитарную матрицу $\mathbf{T}_M(\alpha) = \text{diag}((1, e^{i\alpha}, \dots, e^{i(M-1)\alpha}))^T$, где α — вещественное число.

Лемма 3.2.2. Пусть α — вещественное. Если ряд $\mathbf{X}$ управляется ОЛРФ(A), то ряд $\tilde{\mathbf{X}} = \mathbf{T}_N(\alpha)\mathbf{X}$ управляется $\tilde{A}(\alpha)$, где $\tilde{A}(\alpha) = (\mathbf{T}_{r+1}(-\alpha))A$. Более того, если ряд $\mathbf{Y}$ управляется ОЛРФ(A^2), то ряд $\tilde{\mathbf{Y}} = \mathbf{T}_N(\alpha)\mathbf{Y}$ управляется $(\tilde{A}(\alpha))^2$. К тому же, собственные значения $\mathbf{C}(\tilde{A})$ равны $g_{\tilde{A}}(\omega_j) = g_A(\tilde{\omega}_j)$, где $\tilde{\omega}_j = \omega_j e^{-i\alpha}$.

Доказательство. Лемма доказывается с помощью вычисления столбцов матриц $\tilde{A}^T(\alpha)\mathcal{T}_{r+1}(\tilde{\mathbf{X}}) = 0_{N-r}^T$ и $((\tilde{A}(\alpha))^2)^T\mathcal{T}_{r+1}(\tilde{\mathbf{Y}}) = 0_{N-r}^T$ по определению 1.1.3 и замечания о том, что $g_{A^2}(z) = (g_A(z))^2$. $\square$

Замечание 3.2.2. Возьмём $\alpha \in \mathbb{R}$, $-\pi/N < \alpha \leq \pi/N$ (так как α и $\alpha + 2\pi/N$ ведут к одному и тому же результату) такую, что $\mathbf{C}(\tilde{A})$ — невырождена, и вычислим ортонормированный базис $\tilde{\mathbf{Z}}$, используя лемму 3.2.1, такой, что $\mathbf{Q}^T(\tilde{A})\tilde{\mathbf{Z}} = \mathbf{0}_{N-r,r}$. Тогда $\mathbf{Z} = (\mathbf{T}_N(-\alpha))\tilde{\mathbf{Z}}$ — ортонормированная матрица такая, что $\mathbf{Q}^T(A)\mathbf{Z} = \mathbf{0}_{N-r,r}$. Подобным образом, для $\tilde{\mathbf{Z}}_2$ такой, что $\mathbf{Q}^T((\tilde{A})^2)\tilde{\mathbf{Z}}_2 = \mathbf{0}_{N-r,r}$, $\mathbf{Z}_2 = (\mathbf{T}_N(-\alpha))\tilde{\mathbf{Z}}_2$, мы получаем $\mathbf{Q}^T(A^2)\mathbf{Z}_2 = \mathbf{0}_{N-r,r}$.

Равенство $g_{\tilde{A}}(\omega_j) = g_A(\tilde{\omega}_j)$ означает, что собственные числа $\mathbf{C}(\tilde{A}(\alpha))$ можно вычислить как значения многочлена $g_A(\omega)$ в точках $\omega \in \mathcal{W}(\alpha)$, где

$$\mathcal{W}(\alpha) = \{\omega_j^{(\alpha)}, j = 0, \dots, N-1\}, \quad \omega_j^{(\alpha)} = \exp\left(i\left(\frac{2\pi j}{N} - \alpha\right)\right). \quad (3.15)$$

$\mathcal{W}(\alpha)$ — повернутая равномерная решётка на $\mathbb{T}$. Следовательно, можно добиться невырожденности $\mathbf{C}(\tilde{A}(\alpha))$ выбором подходящего α .

Цель выбора α состоит в том, чтоб уменьшить число обусловленности матрицы $\mathbf{C}(\tilde{A}(\alpha))$ как можно сильнее. Приближённо, это можно свести к задаче максимизации наименьшего собственного числа $|\lambda_{\min}(\alpha)| = \min_{z \in \mathcal{W}(\alpha)} |g_A(z)|$, так как модуль максимального собственного числа не превосходит $\max_{z \in \mathbb{T}} |g_A(z)|$. В точной арифметике любое ненулевое наименьшее собственное число соответствует невырожденности. Однако на практике нам требуется как можно меньшее число обусловленности для устойчивости и точности вычисления.

Объединяя леммы 3.2.1, 3.2.2 с замечаниями 3.2.1, 3.2.2, мы получаем следующий алгоритм вычисления $\mathbf{Z}$.

Алгоритм 3.2.1 (Вычисление базиса $\mathcal{Z}(A)$). *Вход:* число N , вектор коэффициентов $A \in \mathbb{R}^{r+1}$

- 1: Найти $\alpha_0 = \arg \max_{-\pi/N \leq \alpha < \pi/N} \min_{z \in \mathcal{W}(\alpha)} |g_A(z)|$ с помощью одномерного численного метода оптимизации.
- 2: Вычислить вектор $A_g = (a_{g,0}, \dots, a_{g,N-1})^T$ как $a_{g,j} = g_A(\exp(i(\frac{2\pi j}{N} - \alpha_0)))$, $j = 0, \dots, N-1$.
- 3: Вычислить матрицы $\mathbf{R}_r = \mathcal{F}_N([E_{N-r+1} : \dots : E_N])$, $\mathbf{L}_r = \mathbf{A}_g^{-1} \mathbf{R}_r$, где $\mathbf{A}_g = \text{diag } A_g$.
- 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 $\tilde{\mathbf{Z}} = \mathcal{F}_N^{-1}(\mathbf{U}_r)$.
- 6: **return** Матрица $\mathbf{Z} = (\mathbf{T}_N(-\alpha_0))\tilde{\mathbf{Z}}$, столбцы которой образуют ортонормированный базис $\mathcal{Z}(A)$.

Заметим, что вектор A_g состоит из собственных чисел матрицы $\mathbf{C}(\tilde{A})$.

Обсудим вычислительные свойства построенного алгоритма. Следующая теорема показывает порядок $|\lambda_{\min}(\alpha)|$ (и, следовательно, порядок числа обусловленности матрицы, которую надо обратить) в зависимости от длины ряда

N и зависящего от неё α .

Теорема 3.2.1. Пусть t обозначает максимальную кратность корня многочлена $g_A(z)$, лежащего на комплексной единичной окружности $\mathbb{T}$, $\lambda_{\min}(\alpha)$ — наименьшие по модулю собственные числа матриц $\mathbf{C}(\tilde{A}(\alpha))$.

1. Для любой вещественной последовательности $\alpha(N)$ выполняется $|\lambda_{\min}(\alpha(N))| = O(N^{-t})$ при $N \rightarrow +\infty$.
2. Существуют такая вещественная последовательность $\alpha(N)$, что $|\lambda_{\min}(\alpha(N))| = \Theta(N^{-t})$, то есть порядок N^{-t} достигается.

Доказательство. Обозначим $\angle(x, y)$ угол между двумя точками на $\mathbb{T}$, $0 \leq \angle(x, y) \leq \pi$. Докажем первый пункт. Рассмотрим корень $z_1 \in \mathbb{T}$ кратности t , тогда для любого α : $\min_{w \in \mathcal{W}(\alpha)} \angle(w, z_1) \leq \frac{\pi}{N}$ по принципу Дирихле. Зафиксируем любое $0 \leq \alpha_0 < 2\pi$, и выберем $w_0 = \arg \min_{w \in \mathcal{W}(\alpha_0)} \angle(w, z_1)$. Вычисление $g_A(z)$ в точке w_0 завершает доказательство первого пункта, так как $|z_1 - w_0| = O(1/N)$.

Чтобы доказать второй пункт, построим кусочную оценку $g_A(z)$ для $z \in \mathbb{T}$. Рассмотрим разложение $g_A(z) = p_A(z)q_A(z)$, где корни многочлена $p_A(z)$ лежат на $\mathbb{T}$, в то же время как корни $q_A(z)$ не лежат на $\mathbb{T}$. По построению, $\inf_{z \in \mathbb{T}} |q_A(z)| > 0$.

Обозначим $z_1, \dots, z_k$ корни $p_A(z)$ с кратностями $t_1, \dots, t_k$. Мы разбиваем окружность $\mathbb{T}$ на k полуоткрытых непересекающихся дуг $\mathcal{S}_1, \dots, \mathcal{S}_k$, $\mathbb{T} = \bigcup_{1 \leq i \leq k} \mathcal{S}_i$ таких, что $z_i \in \mathcal{S}_i$ для любого i , и $z_j \notin \overline{\mathcal{S}_i}$ для любого $j \neq i$ ($\overline{\mathcal{S}_i}$ обозначает замыкание $\mathcal{S}_i$), из чего следует $\inf_{z \in \mathcal{S}_i} \left| \frac{p_A(z)}{(z - z_i)^{t_i}} \right| > 0$.

Чтобы закончить доказательство, мы покажем, что существует такое $0 \leq \alpha = \alpha(N) < 2\pi$ что $\min_{w \in \mathcal{W}(\alpha(N)), 1 \leq i \leq k} \angle(w, z_i) = \Theta(1/N)$. Введём следующее множество, зависящее от $0 \leq \mu < \pi/N$, $z \in \mathbb{T}$:

$$\mathcal{B}_{z, \mu} = \{0 \leq \alpha < 2\pi : \min_{w \in \mathcal{W}(\alpha)} \angle(w, z) \leq \mu\}.$$

Это множество $\mathcal{B}_{z,\mu}$ имеет следующий явный вид:

$$\mathcal{B}_{z,\mu} = \bigcup_{0 \leq j \leq N-1} \left\{ \operatorname{Arg} \left(\exp \left(i \left(\frac{2\pi j}{N} + y \right) \right) / z \right) \mid -\mu \leq y \leq \mu \right\}.$$

Давайте объясним это выражение. Рассмотрим $\omega_j^{(\alpha)}$ из (3.15), и выберем все такие α_j , что $\angle(\omega_j^{(\alpha_j)}, z) \leq \mu$. Это означает, что полярный угол их отношения $\omega_j^{(\alpha_j)}/z$ лежит в промежутке $[-\mu, \mu]$, то есть $\omega_j^{(\alpha_j)}/z \in \{\exp(ix) \mid -\mu \leq x \leq \mu\}$. Сделаем эквивалентные преобразования: $\exp(i(\frac{2\pi j}{N} - \alpha_j)) \in \{z \exp(ix) \mid -\mu \leq x \leq \mu\}$,

$$\exp \left(i \left(\alpha_j - \frac{2\pi j}{N} \right) \right) \in \{\exp(iy)/z \mid -\mu \leq y \leq \mu\},$$

где $y = -x$, и в итоге

$$\alpha_j \in \{\operatorname{Arg}(\exp(i(\frac{2\pi j}{N} + y))/z) \mid -\mu \leq y \leq \mu\}. \quad (3.16)$$

Чтобы было выполнено $\min_{w \in \mathcal{W}(\alpha)} \angle(w, z) \leq \mu$, нужно, чтобы α равнялось одному из $\alpha_0, \dots, \alpha_{N-1}$. Объединение всех множеств вида (3.16) по $j = 0, \dots, N-1$ даёт $\mathcal{B}_{z,\mu}$.

Мера Лебега множества $\mathcal{B}_{z,\mu}$ на окружности $\mathbb{T}$ равна $\operatorname{mes} \mathcal{B}_{z,\mu} = 2\mu N$ для $\mu < \pi/N$. Возьмём $\mu = \frac{\pi}{2Nk}$ и рассмотрим $\mathcal{B} = \bigcup_{1 \leq i \leq k} \mathcal{B}_{\mu, z_i}$. Так как $\operatorname{mes} \mathcal{B} \leq \pi$, мы получаем, что $\operatorname{mes} \widehat{\mathcal{B}} \geq \pi$, где $\widehat{\mathcal{B}} = [0; 2\pi) \setminus \mathcal{B}$, что означает, что $\widehat{\mathcal{B}}$ не пустое множество. Таким образом, мы доказали, что для любой $\alpha \in \widehat{\mathcal{B}}$

$$\min_{w \in \mathcal{W}(\alpha), 1 \leq i \leq k} \angle(w, z_i) > \frac{\pi}{2Nk}.$$

Зафиксируем произвольное $\alpha_0 \in \widehat{\mathcal{B}}$ и рассмотрим любое $w \in \mathcal{W}(\alpha_0)$. Пусть i — такое, что $w \in \mathcal{S}_i$. Для любого i , $|w - z_i| = \Theta(1/N)$. Следовательно $|g_A(w)| = |q_A(w)| \left| \frac{p_A(w)}{(w - z_i)^{t_i}} \right| |(w - z_i)^{t_i}| \geq C \Theta(N^{-t_i})$, где $C > 0$ — некоторая константа. $\square$

Теперь рассмотрим алгоритм вычисления $\mathbf{Z}(A^2)$. Заметим, что $g_{A^2}(z) = (g_A(z))^2$, благодаря чему можно получить выигрыш по точности при вычислении $\mathcal{Z}(A^2)$ по сравнению с прямым вычислением A^2 и применением алгоритма

3.2.1. Более того, для увеличения точности можно использовать следующее соображение.

Замечание 3.2.3. Рассмотрим лемму 3.2.1 для $ОЛРФ(A^2)$: введём $\mathbf{R}_{2r} = \mathcal{F}_N([E_{N-2r+1} : \dots : E_N]) \in \mathbb{R}^{N \times 2r}$, $\mathbf{L}_{2r} = \mathbf{A}_g^{-2} \mathbf{R}_{2r}$, $\mathbf{V}_{2r} = \mathcal{F}_N^{-1} \mathbf{L}_{2r}$. Применим принцип замечания 3.2.1, но уже к $ОЛРФ(A^2)$. Из $\mathcal{Z}(A) \subset \mathcal{Z}(A^2)$ следует, что $\text{colspace } \mathbf{V}_r \subset \text{colspace } \mathbf{V}_{2r}$, а из этого следует, что $\text{colspace } \mathbf{L}_r \subset \text{colspace } \mathbf{L}_{2r}$. Поэтому ортонормализацию столбцов плохо обусловленной матрицы $\mathbf{L}_{2r}$ можно заменить на поиск ортонормированного базиса пространства столбцов матрицы $(\mathbf{I}_N - \Pi_{\mathbf{U}_r, \mathbf{I}_N}) \mathbf{L}_{2r}$ ранга r , что за счёт знания $\mathbf{U}_r$, полученной ортонормализацией $\mathbf{L}_r$, позволяет сделать процесс ортонормализации $\mathbf{L}_{2r}$ менее обусловленным.

Используя приведённые выше соображения, получим алгоритм 3.2.2.

Алгоритм 3.2.2 (Вычисление базиса $\mathcal{Z}(A^2)$). **Вход:** число N , вектор коэффициентов $A \in \mathbb{R}^{r+1}$.

- 1: Вычислить α_0 , $\mathbf{A}_g$, $\mathbf{U}_r$ таким же способом, как в алгоритме 3.2.1.
- 2: Вычислить $\mathbf{R}_{2r} = \mathcal{F}_N([E_{N-2r+1} : \dots : E_N])$, $\mathbf{L}_{2r} = \mathbf{A}_g^{-2} \mathbf{R}_{2r}$.
- 3: Вычислить $\hat{\mathbf{L}}_{2r} = (\mathbf{I}_N - \Pi_{\mathbf{U}_r, \mathbf{I}_N}) \mathbf{L}_{2r}$.
- 4: Найти матрицу $\hat{\mathbf{U}}_r \in \mathbb{C}^{N \times r}$, столбцы которой составляют ортонормированный базис $\text{colspace } \hat{\mathbf{L}}_{2r}$ (например, $\hat{\mathbf{U}}_r$ может состоять из первых r левых сингулярных векторов матрицы $\hat{\mathbf{L}}_{2r}$).
- 5: Вычислить $\tilde{\mathbf{Z}}_2 = \mathcal{F}_N^{-1}[\mathbf{U}_r : \hat{\mathbf{U}}_r]$.
- 6: **return** Матрица $\mathbf{Z} = (\mathbf{T}_N(-\alpha_0)) \tilde{\mathbf{Z}}_2$, столбцы которой образуют ортонормированный базис $\mathcal{Z}(A^2)$.

Замечание 3.2.4. Несмотря на то, что в алгоритме 3.2.2 используется матрица $\mathbf{A}_g^2$, благодаря её диагональному виду сохраняется порядок относительной ошибки элементов решения СЛАУ. Относительная ошибка вычисления величины $1/\lambda_{\min}$ может быть приближённо оценена как: $|\lambda_{\min}| \left| \frac{1}{\lambda_{\min}} - \right.$

$\frac{1}{\lambda_{\min} + \epsilon} \approx \left| \frac{\epsilon}{\lambda_{\min}} \right|$, где ϵ — ошибочный член, $|\epsilon|$ значительно меньше $|\lambda_{\min}|$, имеет порядок $O(N^t)$, а t определено в теореме 3.2.1. Более того, относительная ошибка величины $1/\lambda_{\min}^2$, рассчитанная как $|\lambda_{\min}^2| \left| \frac{1}{\lambda_{\min}^2} - \frac{1}{(\lambda_{\min} + \epsilon)^2} \right|$, имеет приближенное значение того же порядка.

3.2.2. Использование компенсированной схемы Горнера при вычислении базиса в алгоритмах 3.2.1 и 3.2.2

Вычисление базисов $\mathcal{Z}(A)$ и $\mathcal{Z}(A^2)$ — плохо обусловленная задача при наличии корней у полинома $g_A(z)$ на окружности $\mathbb{T}$, что видно из теоремы 3.2.1. Неустойчивость может проявляться и в пограничной ситуации, когда корни лежат возле окружности $\mathbb{T}$, так как $\lambda_{\min}(\alpha)$ непрерывно зависит от коэффициентов ОЛРФ(A). Поэтому применение алгоритмов 3.2.1 и 3.2.2 напрямую может привести к неустойчивости на практике. Мы предлагаем использовать так называемую «error-free» арифметику и компенсированную схему Горнера [44] (алгоритм `CompHorner`).

Объясним, как компенсированная схема Горнера может использоваться при вычислении базисов $\mathbf{Z}(A)$ и $\mathbf{Z}(A^2)$ с повышенной точностью. Схема Горнера может напрямую применяться в алгоритмах 3.2.1 и 3.2.2 для вычисления полиномов g_A .

Более того, с помощью схемы можно повысить точность вычисления $\mathbf{U}_r$ на шаге 4 алгоритма 3.2.1, что важно, если $\mathbf{L}_r$ плохо обусловлена. Заметим, что такой же подход может быть применён при вычислении $\hat{\mathbf{U}}_r$ на шаге 4 алгоритма 3.2.2. Рассмотрим матрицу $\mathbf{R}_r$, вычисляемую на шаге 3 алгоритма 3.2.1. Так как $(\mathbf{R}_r)_{k,:} = (\exp(\frac{i2\pi rk}{N}), \exp(\frac{i2\pi(r-1)k}{N}), \dots, \exp(\frac{i2\pi k}{N}))$, можно свести умножение $\mathbf{R}_r$ на вектор к вычислению полинома в точке $\exp(\frac{i2\pi k}{N})$. Следовательно, мы можем точнее посчитать умножение $\mathbf{R}_r$ на вектор с помощью схемы Горнера.

Чтобы использовать это свойство, рассмотрим новый способ вычисления

$\mathbf{U}_r$. Пусть $\mathbf{O}_r$ — такая, что $\mathbf{L}_r \mathbf{O}_r$ состоит из ортонормированных столбцов. Матрицу $\mathbf{O}_r$ можно найти с помощью SVD или QR разложения. Тогда выполнено $\mathbf{U}_r = \mathbf{A}_g^{-1}(\mathbf{R}_r \mathbf{O}_r)$. Такое представление $\mathbf{U}_r$ позволяет применить компенсированную схему Горнера к вычислению $\mathbf{R}_r \mathbf{O}_r$.

Алгоритм 3.2.3 (Вычисление базиса $\mathcal{Z}(A)$ с использованием компенсированной схемы Горнера). **Вход:** число N , вектор коэффициентов $A \in \mathbb{R}^{r+1}$.

- 1: Провести шаги 1 и 2 так же, как в алгоритме 3.2.3, но используя алгоритм *CompHorner* при вычислении полиномов.
- 2: Провести пункт 3 так же, как в алгоритме 3.2.3.
- 3: Вместо пункта 4: найти матрицу $\mathbf{O}_r$ такую, что $\mathbf{L}_r \mathbf{O}_r$ — ортонормированная, после чего вычислить $\mathbf{B}_r = \mathbf{R}_r \mathbf{O}_r$, используя алгоритм *CompHorner*, и $\mathbf{U}_r = \mathbf{A}_g^{-1}(\mathbf{B}_r)$, используя обычное матричное умножение.
- 4: Прodelать пункты 5, 6 так же, как в алгоритме 3.2.1.

Алгоритм 3.2.4 (Вычисление базиса $\mathcal{Z}(A^2)$ с использованием компенсированной схемы Горнера). **Вход:** число N , вектор коэффициентов $A \in \mathbb{R}^{r+1}$.

- 1: Вычислить α_0 , $\mathbf{A}_g$, $\mathbf{U}_r$ таким же способом, как в алгоритме 3.2.3.
- 2: Провести пункты 2 и 3 так же, как в алгоритме 3.2.2.
- 3: Вычислить $\hat{\mathbf{U}}_r$ иным способом: найти матрицу $\hat{\mathbf{O}}_r$ такую, что $\hat{\mathbf{L}}_{2r} \hat{\mathbf{O}}_r \in \mathbb{R}^{N \times r}$ — ортонормированная, после чего вычислить $\hat{\mathbf{B}}_r = \mathbf{R}_{2r} \hat{\mathbf{O}}_r$, используя алгоритм *CompHorner*, и $\hat{\mathbf{U}}_r = (\mathbf{I}_N - \Pi_{\mathbf{U}_r, \mathbf{I}_N}) \mathbf{A}_g^{-2}(\hat{\mathbf{B}}_r)$, используя матричное умножение.
- 4: Прodelать пункты 5, 6 так же, как в алгоритме 3.2.2.

3.2.3. Быстрое вычисление оптимального поворота

В пункте 1 алгоритма 3.2.1 требуется найти численное решение задачи

$$\alpha_0 = \arg \max_{-\pi/N \leq \alpha < \pi/N} \min_{z \in \mathcal{W}(\alpha)} |g_A(z)|. \quad (3.17)$$

Для решения этой задачи можно использовать любой метод поиска максимума на отрезке, например, метод Brent [45] или метод золотого сечения с разбиением на промежутки [46].

Проблема состоит в том, что вычисление значения $\min_{z \in \mathcal{W}(\alpha)} |g_A(z)|$ требует $O(Nr)$ времени, при том, что число точек в $\mathcal{W}(\alpha)$, лежащих ближе к корням $g_A(z)$, чем их соседние точки, не так и много — порядка числа корней полинома $g_A(z)$. Прямая оптимизация (3.17) требует много вычислений, поэтому применим эвристику, основывающуюся на том, что можно не перебирать все точки множества $\mathcal{W}(\alpha)$. Обоснованием эвристики послужит следующая лемма.

Лемма 3.2.3. Пусть f_1, f_2 — две ограниченные вещественные функции, определённые на компакте $\mathcal{T}$, и x_0 такое, что $f_1(x_0) = \max_{x \in \mathcal{T}} f_1(x)$ и $f_1(x_0) \leq f_2(x_0)$. Определим $f(x) = \min(f_1(x), f_2(x))$. Тогда $x_0 = \arg \max_{x \in \mathcal{T}} f(x)$.

Доказательство. Предположим противное. Пусть существует такое $\hat{x} \in \mathcal{T}$, что $f(\hat{x}) > f(x_0)$. Тогда $f_1(\hat{x}) \geq f(\hat{x}) > f(x_0) = f_1(x_0) = \max_{x \in \mathcal{T}} f_1(x)$. Получили противоречие. $\square$

Эвристика состоит в следующем: вычислим каждое $|g_A(w_j^{(0)})|$, т.е. $|g_A|$ в точках $\mathcal{W}(0)$, и выберем $j_0 = \arg \min_j |g_A(w_j^{(0)})|$. Будем поддерживать множество $\mathcal{U}$: если индекс j входит в $\mathcal{U}$, то в точке вида $w_j^{(\alpha)}$ будет вычисляться полином g_A . Инициализируем множество $\mathcal{U} = \{j_0 - 1, j_0, j_0 + 1\}$, где сложение и вычитание индексов подразумевается по модулю N . Дополнительно поясним, что $\mathcal{W}_{\mathcal{U}}(\alpha) = \{w_j^{(\alpha)} | j \in \mathcal{U}\}$. Добавление соседних точек связано с $\frac{2\pi}{N}$ -периодичностью функции $\min_j |g_A(w_j^{(\alpha)})|$ по α . Дальше найдём решение подзадачи

$$\hat{\alpha} = \arg \max_{-\pi/N \leq \alpha < \pi/N} \min_{z \in \mathcal{W}_{\mathcal{U}}(\alpha)} |g_A(z)|. \quad (3.18)$$

Если выполнено следующее условие:

$$\min_{z \in \mathcal{W}_{\mathcal{U}}(\hat{\alpha})} |g_A(z)| = \min_{z \in \mathcal{W}(\hat{\alpha})} |g_A(z)|, \quad (3.19)$$

то при $f_1(\alpha) = \min_{z \in \mathcal{W}_U(\alpha)} |g_A(z)|$, $f_2(\alpha) = f(\alpha) = \min_{z \in \mathcal{W}(\alpha)} |g_A(z)|$ с учётом леммы 3.2.3, мы получаем, что решение задачи (3.17) найдено, иначе нужно вновь найти индекс j_1 , в котором значение $|g_A(\omega_{j_1}^{(\hat{\alpha})})|$ достигает минимума, добавить j_1 с соседними точками $j_1 - 1$ и $j_1 + 1$ в множество $\mathcal{U}$ и повторить процедуру ещё раз. Равенство в условии (3.19) связано с тем, что $f_2(\alpha) \leq f_1(\alpha)$ при любом α .

Ниже выписан формальный алгоритм решения задачи (3.17) с помощью эвристики, которая может приводить к значительному уменьшению трудоёмкости.

Алгоритм 3.2.5 (Поиск оптимального поворота). *Вход:* число N , вектор коэффициентов $A \in \mathbb{R}^{r+1}$.

- 1: Вычислить $j_0 = \arg \min_{0 \leq j \leq N-1} |g_A(w_j^{(0)})|$.
- 2: Положить $\mathcal{U} = \{j_0 - 1, j_0, j_0 + 1\}$, $i = 0$.
- 3: **while** *TRUE* **do**
- 4: Численно решить оптимизационную задачу (3.18).
- 5: **if** выполнено условие (3.19) **then**
- 6: **return** $\hat{\alpha}$ — решение оптимизационной задачи (3.17).
- 7: **end if**
- 8: Положить $i = i + 1$.
- 9: Вычислить $j_{i+1} = \arg \min_j |g_A(w_j^{(\hat{\alpha})})|$.
- 10: Положить $\mathcal{U} = \mathcal{U} \cup \{j_{i+1} - 1, j_{i+1}, j_{i+1} + 1\}$.
- 11: **end while**

В наихудшем случае, алгоритм 3.2.5 добавит во множество $\mathcal{U}$ все индексы, и гарантированно завершит свою работу, так как условие (3.19) будет всегда верным. В наилучшем случае, алгоритм потребует всего 2 вычисления функции $\min_{z \in \mathcal{W}(\alpha)} |g_A(z)|$, что является большой оптимизацией при больших N . На практике не встречался случай, когда алгоритм 3.2.5 сделал бы больше 3 итераций.

Тем не менее, реальное число итераций может варьироваться в зависимости от количества корней у полинома $g_A(z)$.

3.3. Алгоритмы решения задачи

3.3.1. Вычисление взвешенных проекторов с заданным базисом

Как в модифицированном методе Гаусса-Ньютона (MGN) (3.6), так и в методе Variable projection Гаусса-Ньютона (VPGN) (3.4) вычисляются проекции вида $\Pi_{\mathbf{Z}, \mathbf{W}} \mathbf{X}$, матрица $\mathbf{Z}$ лежит в $\mathbb{R}^{N \times r}$ или $\mathbb{R}^{N \times 2r}$. Мы будем считать, что если матрица $\mathbf{W}$ — $(2p + 1)$ -диагональная, то она представлена своим разложением Холецкого $\mathbf{W} = \mathbf{C}^T \mathbf{C}$, $\mathbf{C}$ — верхнетреугольная с $(p + 1)$ диагональю; если же $\mathbf{W}^{-1}$ — $(2p + 1)$ -диагональная, то мы представляем $\mathbf{W} = \hat{\mathbf{C}}^{-1}(\hat{\mathbf{C}}^{-1})^T$, где $\mathbf{W}^{-1} = \hat{\mathbf{C}}^T \hat{\mathbf{C}}$ представлена аналогично своим разложением Холецкого.

Замечание 3.3.1. Как показано в разделе 1.3.1, вычисление выражения вида $\mathbf{B}^\dagger \mathbf{D}$ сводится к решению задачи обычного МНК, для чего можно использовать QR или SVD разложение матрицы $\mathbf{B}$.

Алгоритм 3.3.1 (Вычисление $\Pi_{\mathbf{Z}, \mathbf{W}} \mathbf{X}$). *Вход:* матрица $\mathbf{Z}$, временной ряд $\mathbf{X}$, разложение $\mathbf{W} = \mathbf{C}^T \mathbf{C}$ или $\mathbf{W}^{-1} = \hat{\mathbf{C}}^T \hat{\mathbf{C}}$.

- 1: **if** $\mathbf{W}$ — $(2p + 1)$ -диагональная **then**
- 2: Вычислить вектор $\mathbf{D} = \mathbf{C}\mathbf{X}$, матрицу $\mathbf{B} = \mathbf{C}\mathbf{Z}$.
- 3: **end if**
- 4: **if** $\mathbf{W}^{-1}$ — $(2p + 1)$ -диагональная **then**
- 5: Вычислить вектор $\mathbf{D} = (\hat{\mathbf{C}}^{-1})^T \mathbf{X}$, матрицу $\mathbf{B} = (\hat{\mathbf{C}}^{-1})^T \mathbf{Z}$.
- 6: **end if**
- 7: Вычислить $\mathbf{Q} = \mathbf{B}^\dagger \mathbf{D}$, см. замечание 3.3.1.
- 8: **return** Вычислить $\mathbf{Z}\mathbf{Q}$.

3.3.2. Вычисление $\Pi_{\mathcal{Z}(A), \mathbf{W}}\mathbf{X}$

Вычисление $\Pi_{\mathcal{Z}(A), \mathbf{W}}\mathbf{X}$ можно проводить с использованием особенностей подпространства $\mathcal{Z}(A)$ (см. раздел 3.2.1) и без них, как было предложено в [17].

Начнём с алгоритма, предложенного в статье [17]. Вычисление $\Pi_{\mathcal{Z}(A), \mathbf{W}}\mathbf{X}$ использует формулу (3.7), для которой нужно вычислять $\Gamma(A)$ и $\Gamma^{-1}(A)$. Мы приводим следующий алгоритм в том виде, как он использован в статье [17], с быстрым вычислением $\Gamma(A)$ и обратной к ней:

Алгоритм 3.3.2 (Вычисление решения $\Pi_{\mathcal{Z}(A), \mathbf{W}}\mathbf{X}$ методом из [17]). *Вход:* вектор A , временной ряд $\mathbf{X}$, матрица $\hat{\mathbf{C}}$.

- 1: Вычислить $(r + p + 1)$ -диагональную матрицу $\tilde{\mathbf{C}} = \hat{\mathbf{C}}\mathbf{Q}(A)$.
- 2: Вычислить $(2r + 2p + 1)$ -диагональную матрицу $\Gamma(A) = \tilde{\mathbf{C}}^T \tilde{\mathbf{C}}$.
- 3: Вычислить $(p + r + 1)$ -диагональное разложение Холецкого матрицы $\Gamma(A)$ для умножения $\Gamma^{-1}(A)$ справа на вектор.
- 4: **return** Вычислить $\Pi_{\mathcal{Z}(A), \mathbf{W}}\mathbf{X}$ по формуле (3.7).

Разработанная в разделе 3.2.1 теория позволяет построить другой способ вычисления проекции.

Алгоритм 3.3.3 (Вычисление решения $\Pi_{\mathcal{Z}(A), \mathbf{W}}\mathbf{X}$ с помощью свойств $\mathcal{Z}(A)$).

Вход: вектор A , временной ряд $\mathbf{X}$, разложение $\mathbf{W} = \mathbf{C}^T \mathbf{C}$ или $\mathbf{W}^{-1} = \hat{\mathbf{C}}^T \hat{\mathbf{C}}$.

- 1: Вычислить базис $\mathbf{Z}(A)$, используя алгоритм 3.2.1 или 3.2.3 в зависимости от требуемой точности.
- 2: **return** Вычислить $\Pi_{\mathbf{Z}(A), \mathbf{W}}\mathbf{X}$, используя алгоритм 3.3.1.

Ключевая разница между алгоритмами 3.3.2 и 3.3.3 состоит в замене вычисления образа $\mathcal{Q}(A)$ оператора $\mathbf{Q}(A)$ на вычисление его ядра $\mathcal{Z}(A)$.

3.3.3. Алгоритм VPGN

Ниже мы приводим алгоритм Гаусса-Ньютона для метода Variable Projection с итерациями (3.4), который был получен в [17] с использованием принципа Variable projection. Заметим, что быстрая реализация предлагаемого в [17] алгоритма вычисления $\Pi_{\mathbf{Z}(A), \mathbf{W}} \mathbf{X}$ доступна только когда $\mathbf{W}^{-1}$ — $(2p + 1)$ -диагональная.

Алгоритм 3.3.4 (Алгоритм Гаусса-Ньютона с помощью принципа Variable Projection и вычисления $\mathbf{S}^*$ из [17] (VPGN)). ***Вход:** временной ряд $\mathbf{X}$, вектор коэффициентов ОЛРФ A_0 , критерий остановки STOP.*

- 1: Вычислить $\tau = \arg \max_i |a_i^{(0)}|$.
- 2: Вычислить $\mathbf{S}_0 = \mathbf{S}^*(A_{(\tau)}^{(0)})$.
- 3: Положить $k = 0$.
- 4: **while** не STOP **do**
- 5: Вычислить $\mathbf{J}_{\mathbf{S}^*}(A_{(\tau)}^{(k)})$ согласно формуле (3.8), используя п. 1–3 алгоритма 3.3.2.
- 6: Вычислить $\Delta_k = \left(\mathbf{J}_{\mathbf{S}^*}(A_{(\tau)}^{(k)}) \right)_{\mathbf{W}}^{\dagger} (\mathbf{X} - \mathbf{S}_k)$, используя алгоритм 3.3.1.
- 7: Найти $\gamma_k = \arg \min_{0 \leq \gamma \leq 1} \|\mathbf{X} - \mathbf{S}^*(A_{(\tau)}^{(k)} + \gamma \Delta_k)\|_{\mathbf{W}}$, используя численный метод минимизации на отрезке.
- 8: Положить $A_{(\tau)}^{(k+1)} = A_{(\tau)}^{(k)} + \gamma_k \Delta_k$, $a_{\tau}^{k+1} = a_{\tau}^k$.
- 9: Вычислить $\mathbf{S}_{k+1} = \mathbf{S}^*(A_{(\tau)}^{(k+1)})$.
- 10: Положить $k = k + 1$.
- 11: **end while**
- 12: **return** Временной ряд $\mathbf{S}_k$.

Заметим, что вычисление $\mathbf{S}^*(A_{(\tau)}) = \Pi_{\mathbf{Z}(A_{(\tau)}), \mathbf{W}}$ можно производить одним из алгоритмов 3.3.2 или 3.3.3. В [17] был использован алгоритм 3.3.2.

3.3.4. Модифицированный алгоритм Гаусса-Ньютона MGN

Ниже мы приводим схему предлагаемого модифицированного метода Гаусса-Ньютона с итерациями в форме (3.6).

Алгоритм 3.3.5 (Модифицированный алгоритм Гаусса-Ньютона (MGN)).

Вход: временной ряд $\mathbf{X}$, вектор коэффициентов ОЛРФ A_0 , критерий остановки $STOP$.

- 1: Вычислить $\tau = \arg \max_i |a_i^{(0)}|$.
- 2: Вычислить $S_0 = S^*(A_{(\tau)}^{(0)})$.
- 3: Положить $k = 0$.
- 4: **while** не $STOP$ **do**
- 5: Вычислить $\mathbf{Z}((A^{(k)})^2)$, используя алгоритм 3.2.2 или 3.2.4 в зависимости от требуемой точности для вычисления проекции.
- 6: Вычислить $\Delta_k = \mathbf{M}_k^\dagger \mathbf{Q}^T(A^{(k)}) \Pi_{\mathcal{Z}((A^{(k)})^2), \mathbf{W}}(\mathbf{X} - S_k)$, используя алгоритм 3.3.1.
- 7: Найти $\gamma_k = \arg \min_{0 \leq \gamma \leq 1} \|\mathbf{X} - S^*(A_{(\tau)}^{(k)} + \gamma \Delta_k)\|_{\mathbf{W}}$, используя численный метод минимизации на отрезке.
- 8: Положить $A_{(\tau)}^{(k+1)} = A_{(\tau)}^{(k)} + \gamma_k \Delta_k$, $a_{\tau}^{k+1} = a_{\tau}^k$.
- 9: Вычислить $S_{k+1} = S^*(A_{(\tau)}^{(k+1)})$.
- 10: Положить $k = k + 1$.
- 11: **end while**
- 12: **return** Временной ряд S_k .

Подробно возможный критерий остановки $STOP$ рассматривается в разделе 5.1.2.

3.3.5. Сравнение вычислительной трудности алгоритмов

Сравним варианты алгоритмов 3.3.4 и 3.3.5.

Сравниваемые алгоритмы

Мы рассматриваем три фиксированных алгоритма:

1. Метод Variable Projection Гаусса-Ньютона (VPGN) с вычислением проектора через алгоритм 3.3.2 и шага через алгоритм 3.3.4.
2. Полуустойчивый метод Variable Projection Гаусса-Ньютона (S-VPGN) с вычислением проектора через алгоритм 3.3.3 и шага через алгоритм 3.3.4.
3. Устойчивый модифицированный метод Гаусса-Ньютона (MGN) с использованием алгоритма 3.3.3 и вычислением шага через алгоритм 3.3.5.

Замечание 3.3.2. *Алгоритм MGN допускает более устойчивую реализацию в случае сигналов, управляемых ОЛРФ с кратными корнями у характеристического многочлена (в частном случае, для полиномиальных сигналов). Это объясняется тем, что предложенный метод MGN работает с матрицами с числом обусловленности $O(N^t)$ в методе (см. теорему 3.2.1), где t — наибольшая степень корня характеристического полинома, в то время, как метод VPGN оперирует с матрицами с числом обусловленности $O(N^{2t})$ [17, Section 6.2].*

Численное сравнение упомянутых алгоритмов приведено в разделе (5.1.2).

Сравнение трудоёмкости алгоритмов 3.3.2 и 3.3.3

Найдём вычислительную сложность алгоритмов при $N \rightarrow \infty$. Заметим, что трудоёмкость поиска оптимального поворота (оптимальное α_0) в шаге 1 алгоритма 3.2.1 в асимптотику алгоритма 3.3.3 не включена.

- $\mathbf{W}^{-1}$ — $(2p + 1)$ -диагональная.

Асимптотическая сложность алгоритма 3.3.3 — $O(Nr^2 + Np^2 + rN \log N)$: алгоритмы 3.2.1, 3.2.3 — $O(rN \log N + Nr^2)$, вычисление проекции через алгоритм 3.3.1 — $O(Nr^2 + Np^2)$, МНК через QR-разложение требует $O(Nr^2)$ времени.

Асимптотическая сложность алгоритма 3.3.2 — $O(Nr^2 + Np^2)$.

- $\mathbf{W} — (2p + 1)$ -диагональная, $p > 0$.

Асимптотическая сложность алгоритма 3.3.3 — $O(Nr^2 + Np^2 + rN \log N)$, оценка времени работы такая же, как и в предыдущем случае.

Нет реализации алгоритма 3.3.2 метода с линейной по N асимптотикой из-за того, что $\mathbf{\Gamma}(\mathbf{A})$ — не ленточная.

Сравнение трудоемкости алгоритмов 3.3.5 и 3.3.4

Ниже мы приведём асимптотики трудоёмкости одной итерации алгоритма 3.3.5 и метода 3.3.4.

- $\mathbf{W}^{-1} — (2p + 1)$ -диагональная

Асимптотическая сложность итерации алгоритма 3.3.5 — $O(Nr^2 + Np^2 + rN \log N)$: алгоритмы 3.2.2, 3.2.4 — $O(rN \log N + Nr^2)$, вычисление проекции через Алгоритм 3.3.1 — $O(Nr^2 + Np^2)$; МНК через QR-разложение требует $O(Nr^2)$ времени.

Асимптотическая сложность итерации алгоритма 3.3.4 — $O(Nr^2 + Np^2)$: учитывая, что разложение Холецкого матрицы $\mathbf{\Gamma}(A_k)$ требуется считать один раз за итерацию за $O(Nr^2 + Np^2)$, умножение $\mathbf{\Gamma}^{-1}(A_k)$ на вектор стоит $O(N(r + p))$.

- $\mathbf{W} — (2p + 1)$ -диагональная, $p > 0$.

Асимптотическая сложность итерации алгоритма 3.3.5 — $O(Nr^2 + Np^2 + rN \log N)$, оценка времени работы такая же, как и в предыдущем случае.

Нет реализации алгоритма 3.3.4 с линейной по N асимптотикой из-за того, что $\mathbf{\Gamma}(\mathbf{A})$ — не ленточная.

Таким образом, если $\mathbf{W}^{-1}$ — $(2p + 1)$ -диагональная, то трудоёмкость предложенного метода MGN выше на $O(rN \log N)$, чем у метода VPGN. Однако, если $\mathbf{W}$ — $(2p + 1)$ -диагональная (а это случай шума-авторегрессии, что является вполне естественным предположением), то трудоёмкость метода MGN на порядок ниже.

Глава 4

Матричный подход к задаче аппроксимации временного ряда рядами конечного ранга (метод Кэдзоу)

В этой главе мы рассмотрим алгоритм численного решения задачи Hankel SLRA (4) — метод итераций Кэдзоу [1]. Определим следующую последовательность матриц, начинающуюся с заданной матрицы $\mathbf{X}$:

$$\mathbf{X}_0 = \mathbf{X}, \quad \mathbf{X}_{k+1} = \Pi_{\mathcal{H}} \Pi_{\mathcal{M}_r} \mathbf{X}_k, \quad k \geq 0, \quad (4.1)$$

где $\mathcal{H} = \mathcal{H}_{L,K}$ — пространство ганкелевых матриц размера $L \times K$, $\mathcal{M}_r \subset \mathbb{R}^{L \times K}$ — множество матриц ранга, не превосходящего r , $\mathbf{L} \in \mathbb{R}^{L \times L}$, $\mathbf{R} \in \mathbb{R}^{K \times K}$ — положительно определённые матрицы весов $\Pi_{\mathcal{H}}$ и $\Pi_{\mathcal{M}_r}$ — проекторы на соответствующие множества по матричной норме $\|\cdot\|_{\mathbf{L},\mathbf{R}}$, порождённой скалярным произведением (5).

В этой главе мы исследуем несколько вопросов, связанных с методом Кэдзоу. Мы покажем, что последовательность (4.1) содержит подпоследовательность, сходящуюся к $\mathcal{M}_r \cap \mathcal{H}$, изучим задачи поиска матриц весов $\mathbf{L}$ и $\mathbf{R}$ таких, что $\|\mathcal{T}_L(\mathbf{Z})\|_{\mathbf{L},\mathbf{R}} \approx \|\mathbf{Z}\|_{\mathbf{W}_0}$ для любого $\mathbf{Z}$, где матрица $\mathbf{W}$ из задачи (3) равна $\mathbf{W}_0 = \mathbf{\Sigma}^{-1}$ в этой главе, приведём алгоритмы решения этих задач и быструю реализацию метода Кэдзоу. В конце главы мы рассмотрим вопрос о применении метода Кэдзоу для задания начального приближения в методах, описанных в главе 3, в частности, в предлагаемом новом методе MGN, и найдём ошибку первого порядка при оценивании зашумлённого сигнала методом Кэдзоу.

4.1. Сходимость метода Кэдзоу по подпоследовательностям

4.1.1. Общие результаты о сходимости в гильбертовом пространстве

Рассмотрим задачу проектирования точки x на множество $\mathcal{H} \cap \mathcal{M}$ в гильбертовом пространстве $\mathbb{X}$ со скалярным произведением $\langle \cdot, \cdot \rangle$, где $\mathcal{H}$ и $\mathcal{M}$ замкнуты в порожденной скалярным произведением топологии, $\mathcal{H}$ — линейное подпространство, а также $\mathcal{M}$ замкнуто относительно операции умножения на скаляр, то есть, если $z \in \mathcal{M}$, то $\alpha z \in \mathcal{M}$ для любого вещественного α . Заметим, что $\mathcal{M}$ не обязательно выпуклое множество или линейное подпространство.

Задача имеет вид

$$y^* = \arg \min_{y \in \mathcal{H} \cap \mathcal{M}} \|x - y\|, \quad (4.2)$$

где $\|\cdot\|$ — это норма, порождённая скалярным произведением.

Чтобы привести схему алгоритма для решения данной задачи, введём проекторы на подмножество $\mathcal{M}$ и подпространство $\mathcal{H}$ по соответствующей норме $\|\cdot\|$: $P_{\mathcal{M}}$ — проектор на $\mathcal{M}$, $P_{\mathcal{H}}$ — проектор на $\mathcal{H}$. Заметим, что если проекция на $\mathcal{M}$ не определена однозначно, то будем предполагать, что в случае неоднозначности берётся любая ближайшая точка. Проектор на $\mathcal{H}$, очевидно, ортогонален, в то время как $P_{\mathcal{M}}$ ортогонален согласно следующему предложению.

Предложение 4.1.1. Пусть $\mathbb{X}$ — гильбертово пространство, $\mathcal{M} \subset \mathbb{X}$ — подмножество, замкнутое относительно операции умножения на скаляр, $P_{\mathcal{M}}$ — оператор проектирования на $\mathcal{M}$. Тогда для любого $x \in \mathbb{X}$ выполнено следующее равенство («теорема Пифагора»):

$$\|x\|^2 = \|x - P_{\mathcal{M}}x\|^2 + \|P_{\mathcal{M}}x\|^2.$$

Доказательство. Обозначим $y = P_{\mathcal{M}}x$. Так как

$$\|x\|^2 = \|x - y\|^2 + \|y\|^2 + 2\langle x - y, y \rangle,$$

нужно доказать, что $\langle x - y, y \rangle = 0$. Предположим противное: $\langle x - y, y \rangle \neq 0$. Тогда для

$$\gamma = \frac{\langle x, y \rangle}{\langle y, y \rangle}$$

$\langle x - \gamma y, \gamma y \rangle = \gamma \langle x - \gamma y, y \rangle = 0$ и, таким образом, $\|x - y\|^2 > \|x - \gamma y\|^2$:

$$\|x - y\|^2 - \|x - \gamma y\|^2 = \langle y, y \rangle - 2\langle x, y \rangle + \frac{\langle x, y \rangle^2}{\langle y, y \rangle} = \frac{\langle x - y, y \rangle^2}{\langle y, y \rangle} > 0.$$

Так как γy лежит в $\mathcal{M}$ согласно свойству $\mathcal{M}$, получено противоречие с тем фактом, что $y = \Pi_{\mathcal{M}}x$ — ближайшая точка к x . $\square$

Рассмотрим *итеративный метод попеременных проекций* для задачи (4.2), который задан следующим шагом итерации:

$$y_{k+1} = \Pi_{\mathcal{H}}\Pi_{\mathcal{M}}y_k, \text{ где } y_0 = x. \quad (4.3)$$

В следующей теореме рассмотрим сходимость последовательности (4.3).

Теорема 4.1.1. *Пусть выполняются условия Предложения 4.1.1, а также множество $\mathcal{M}$ и подпространство $\mathcal{H}$ топологически замкнуты. Тогда*

1. $\|y_k - \Pi_{\mathcal{M}}y_k\| \rightarrow 0$ при $k \rightarrow +\infty$, $\|\Pi_{\mathcal{M}}y_k - y_{k+1}\| \rightarrow 0$ при $k \rightarrow +\infty$.
2. Пусть $\mathcal{M} \cap B_1$ является компактом, где $B_1 = \{z : \|z\| \leq 1\}$ — замкнутый единичный шар. Тогда существует сходящаяся подпоследовательность точек $y_{i_1}, y_{i_2}, \dots$ такая, что её предел y^* лежит в $\mathcal{M} \cap \mathcal{H}$.

Доказательство. Воспользуемся следующими неравенствами:

$$\|y_k - \Pi_{\mathcal{M}}y_k\| \geq \|\Pi_{\mathcal{M}}y_k - y_{k+1}\| \geq \|y_{k+1} - \Pi_{\mathcal{M}}y_{k+1}\|. \quad (4.4)$$

Действительно, так как проекция $\Pi_{\mathcal{M}}z$ находится не дальше от z , чем любая другая точка из $\mathcal{M}$, и подобное рассуждение верно и для $\Pi_{\mathcal{H}}$, получаем, что $\|\Pi_{\mathcal{M}}y_k - z\| \geq \|z - \Pi_{\mathcal{M}}z\|$, где $z = y_{k+1}$, и $\|y_k - z\| \geq \|z - \Pi_{\mathcal{H}}z\|$, где $z = \Pi_{\mathcal{M}}y_k$.

1. Согласно неравенствам (4.4), последовательности $\|y_k - \Pi_{\mathcal{M}}y_k\|$, $k = 1, 2, \dots$, и $\|\Pi_{\mathcal{M}}y_k - y_{k+1}\|$, $k = 1, 2, \dots$, являются невозрастающими. Очевидно, что снизу они ограничены нулём. Таким образом, опять же согласно (4.4), они имеют одинаковый предел c .

Докажем, что $c = 0$, предполагая противное $c > 0$. Тогда существует $d > 0$ такое, что $\|y_k - \Pi_{\mathcal{M}}y_k\| > d$ и $\|\Pi_{\mathcal{M}}y_k - y_{k+1}\| > d$ для любых $k = 1, 2, \dots$. Согласно предложению 4.1.1, верно следующее равенство: $\|y_k\|^2 = \|y_k - \Pi_{\mathcal{M}}y_k\|^2 + \|\Pi_{\mathcal{M}}y_k\|^2$. Так как подпространство $\mathcal{H}$ линейно, следующее равенство тоже верно: $\|\Pi_{\mathcal{M}}y_k\|^2 = \|\Pi_{\mathcal{M}}y_k - \Pi_{\mathcal{H}}\Pi_{\mathcal{M}}y_k\|^2 + \|\Pi_{\mathcal{H}}\Pi_{\mathcal{M}}y_k\|^2 = \|\Pi_{\mathcal{M}}y_k - y_{k+1}\|^2 + \|y_{k+1}\|^2$. Следовательно,

$$\|y_k\|^2 = \|y_k - \Pi_{\mathcal{M}}y_k\|^2 + \|\Pi_{\mathcal{M}}y_k - y_{k+1}\|^2 + \|y_{k+1}\|^2.$$

Таким образом, $\|y_{k+1}\|^2 < \|y_k\|^2 - 2d^2$. Распиряя это неравенство подобным способом, мы получаем, что $\|y_{k+j}\|^2 < \|y_k\|^2 - 2jd^2$ для любых $j = 1, 2, \dots$. Выберем $k = 1$, и $j = \lceil \|y_k\|^2 / (2d^2) \rceil + 1$. Тогда $\|y_{k+j}\|^2 < 0$, чего не может быть. Таким образом, $c = 0$.

2. Рассмотрим последовательность $(\Pi_{\mathcal{M}}y_k, k = 1, 2, \dots)$. Она ограничена, так как $\|\Pi_{\mathcal{M}}z\| \leq \|z\|$ (согласно Предложению 4.1.1) и $\|\Pi_{\mathcal{H}}z\| \leq \|z\|$ для любого $z \in \mathbb{X}$. Последовательность лежит в компактном множестве, так как $\mathcal{M}$ замкнуто относительно операции умножения на скаляр, и мы можем растянуть единичный шар так, чтобы он покрывал последовательность. Поэтому можно выбрать сходящуюся подпоследовательность $(\Pi_{\mathcal{M}}y_{i_k})$; обозначим её предел как $y^* \in \mathcal{M}$, при этом $\|\Pi_{\mathcal{M}}y_{i_k} - y_{i_k+1}\| = \|\Pi_{\mathcal{M}}y_{i_k} - \Pi_{\mathcal{H}}\Pi_{\mathcal{M}}y_{i_k}\| \rightarrow 0$ при $k \rightarrow +\infty$. Так как $\mathcal{H}$ замкнуто, а пространство $\mathbb{X}$ банахово, проектор $\Pi_{\mathcal{H}}$ является непрерывным отображением. Зная, что $\|z - \Pi_{\mathcal{H}}z\|$ это композиция непрерывных отображений, получаем, что $\|y^* - \Pi_{\mathcal{H}}y^*\| = 0$, $y^* \in \mathcal{M} \cap \mathcal{H}$. Наконец, $\Pi_{\mathcal{H}}$ — непрерывное

отображение, и, таким образом, последовательность $(\Pi_{\mathcal{H}}\Pi_{\mathcal{M}}y_{i_k})$ сходится к y^* . Таким образом, y_{i_k+1} — требуемая подпоследовательность.

□

Замечание 4.1.1. Предложение 4.1.1 было доказано в [47] для частного случая, в то время как неравенства (4.4) являются обобщениями неравенств [48, (4.1)].

4.1.2. Применение результатов к задаче (4)

Применим теорему 4.1.1 к случаю матричной аппроксимации ганкелевыми матрицами неполного ранга (4). Пусть $\mathbb{X} = \mathbb{R}^{L \times K}$, то есть $\mathbb{X}$ — пространство матриц размера $L \times K$, снабжённое скалярным произведением (5), $\mathcal{H} \subset \mathbb{R}^{L \times K}$ — пространство ганкелевых матриц, $\mathcal{M} = \mathcal{M}_r \subset \mathbb{R}^{L \times K}$ — множество матриц ранга, не превосходящего r . Тогда шаг итерации (4.3) для метода попеременных проекций приводит к итерации Кэдзоу (4.1).

Известно, что множество $\mathcal{M}_r$ замкнуто по обычной фробениусовской норме [38], и, таким образом, замкнуто по любой норме, так как в пространстве матриц все нормы эквивалентны. Замкнутый единичный шар, очевидно, является компактом в конечномерном евклидовом пространстве. Таким образом, теорема 4.1.1 выполняется, и в последовательности метода Кэдзоу (4.1) существует сходящаяся подпоследовательность. Заметим, что существование сходящейся подпоследовательности может быть выведено из [1]. Однако, доказательство этого факта в данной работе основано на иных предположениях; в частности, упор сделан на теорему Пифагора для проекторов на множества, которые замкнуты относительно умножения на скаляр.

Следующее предложение показывает, что из себя представляет пересечение множеств $\mathcal{M}_r \cap \mathcal{H}$, в котором лежит предельная точка сходящейся подпоследовательности в последовательности (4.1), полученной методом Кэдзоу.

Предложение 4.1.2. Пусть $\min(L, N - L + 1) > r$. Тогда $\mathbf{X} \in \mathcal{M}_r \cap \mathcal{H}$ тогда и только тогда, когда $\mathbf{X} \in \overline{\mathcal{D}}_r$, где $\mathbf{X} = \mathcal{T}_L(\mathbf{X})$.

Доказательство. Это прямое следствие из предложения 1.2.1 и предложения 2.1.1. $\square$

4.2. Постановка задачи поиска весов

Рассмотрим оценку сигнала $\mathbf{S}_0$ из ряда $\mathbf{X} = \mathbf{S}_0 + \boldsymbol{\epsilon}$ с помощью решения задачи (4), где $\mathbf{S} \in \mathcal{D}_r$, $\boldsymbol{\epsilon}$ — случайный вектор с нулевым средним. Пусть $\boldsymbol{\Sigma}$ обозначает ковариационную матрицу вектора $\boldsymbol{\epsilon}$. Для того чтобы получить ОМП сигнала, необходимо решить задачу (3) с матрицей весов, равной $\boldsymbol{\Sigma}^{-1}$ или пропорциональной ей. Например, если $\boldsymbol{\epsilon}$ соответствует стационарному процессу, то вместо обратной к автоковариационной матрице можно брать матрицу весов в (3), обратную к автокорреляционной. Возникает вопрос: можно ли подобрать соответствующие матрицы $\mathbf{L}$, $\mathbf{R}$ из скалярного произведения в пространстве матриц (5) таким образом, чтобы достичь требуемых весов?

В этом разделе мы изучаем случай, когда ковариационная матрица $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_N \in \mathbb{R}^{N \times N}$ случайного вектора $\boldsymbol{\epsilon}$ является автоковариационной матрицей процесса авторегрессии порядка p (AR(p)).

4.2.1. Соотношение между скалярными произведениями

Рассмотрим соотношение между (полу)нормами $\|\mathcal{T}_L(\mathbf{Z})\|_{\mathbf{L}, \mathbf{R}}$ и $\|\mathbf{Z}\|_{\mathbf{W}}$, где $\|\mathbf{Z}\|_{\mathbf{W}}^2 = \mathbf{Z}^T \mathbf{W} \mathbf{Z}$, $\|\cdot\|_{\mathbf{W}}$ порождена скалярным произведением (5), $\mathbf{L} \in \mathbb{R}^{L \times L}$, $\mathbf{R} \in \mathbb{R}^{K \times K}$, $\mathbf{W} \in \mathbb{R}^{N \times N}$ — положительно (неотрицательно) определённые, $N = L + K - 1$. Обозначим элементы $\mathbf{L} = (l_{i,j})$, $\mathbf{R} = (r_{i,j})$, $\mathbf{W} = (w_{i,j})$.

Также мы определим ациклическую свёртку двух векторов $C = A * B$, $A = (a_1, \dots, a_{N_a})^T$, $B = (b_1, \dots, b_{N_b})^T$, $C = (c_1, \dots, c_{N_c})^T$: $c_i = \sum_{k,j:k+j-1=i} a_k b_j$,

$N_c = N_a + N_b - 1$. Обозначим i -ю диагональ матрицы $\mathbf{A} \in \mathbb{R}^M$ длины $M - |i|$, состоящую из элементов с индексами (j, k) , удовлетворяющих равенству $k - j = i$, как $\text{diag}_{\mathbf{A}}(i)$, $i = -M, \dots, M$. Определим свёртку двух матриц $\mathbf{C} = \mathbf{A} * \mathbf{B}$ для $\mathbf{A} \in \mathbb{R}^{L \times L}$, $\mathbf{B} \in \mathbb{R}^{K \times K}$, $\mathbf{C} \in \mathbb{R}^{N \times N}$: $c_{i,j} = \sum_{i_1+i_2-1=i} \sum_{j_1+j_2-1=j} a_{i_1,j_1} b_{i_2,j_2}$, где $\mathbf{A} = (a_{i_1,j_1})$, $\mathbf{B} = (b_{i_2,j_2})$, $\mathbf{C} = (c_{i,j})$.

Теорема 4.2.1. $\|\mathcal{T}_L(Z)\|_{\mathbf{L},\mathbf{R}} = \|Z\|_{\mathbf{W}}$ для любого ряда $Z \in \mathbb{R}^{N \times N}$ тогда и только тогда, когда $\mathbf{W} = \mathbf{W}(\mathbf{L}, \mathbf{R}) = \mathbf{L} * \mathbf{R}$. При этом i -я диагональ $\mathbf{L} * \mathbf{R}$ равна

$$\text{diag}_{\mathbf{L}*\mathbf{R}}(i) = \sum_{j+k=i} \begin{pmatrix} 0_{(|j|+|k|-|j+k|)/2} \\ \text{diag}_{\mathbf{L}}(j) * \text{diag}_{\mathbf{R}}(k) \\ 0_{(|j|+|k|-|j+k|)/2} \end{pmatrix}, \quad (4.5)$$

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

Доказательство. Чтобы показать равенство норм, необходимо показать равенство полускалярных произведений ортов, то есть вычислить:

$$\langle \mathcal{T}_L(E_i), \mathcal{T}_L(E_j) \rangle_{\mathbf{L},\mathbf{R}} = \text{tr}(\mathbf{L} \mathcal{T}_L(E_i) \mathbf{R} \mathcal{T}_L^T(E_j)),$$

так как эти элементы составляют матрицу $\mathbf{W}$.

Для того чтобы это сделать, необходимо представить скалярное произведение в виде линейной комбинации K^2 «полу-скалярных произведений»:

$$\begin{aligned} \langle \mathcal{T}_L(E_i), \mathcal{T}_L(E_j) \rangle_{\mathbf{L},\mathbf{R}} &= \sum_{k=1}^K \sum_{l=1}^K r_{m,k} \langle \mathcal{T}_L(E_i), \mathcal{T}_L(E_j) \rangle_{\mathbf{L}, \mathbf{e}_{m,k}} \\ &= \sum_{k=1}^K \sum_{m=1}^K r_{m,k} \text{tr}(\mathbf{L} \mathcal{T}_L(E_i) \mathbf{e}_{m,k} \mathcal{T}_L^T(E_j)). \end{aligned}$$

Таким образом, мы получили соотношение $\mathbf{W} = \sum_{k=1}^K \sum_{m=1}^K r_{m,k} \mathbf{W}_{m,k}$, где $\mathbf{W}_{m,k} = (w_{i,j}^{(m,k)})_{i,j=1}^N$ — матрицы следующего вида:

$$w_{i,j}^{(m,k)} = \text{tr}(\mathbf{L} \mathcal{T}_L(E_i) \mathbf{e}_{m,k} \mathcal{T}_L^T(E_j)) = \langle \mathbf{L} \mathcal{T}_L(E_i), \mathcal{T}_L(E_j) \mathbf{e}_{k,m} \rangle_{\mathbf{F}}.$$

Для удобства, мы обозначили $\langle \cdot, \cdot \rangle_F = \langle \cdot, \cdot \rangle_{\mathbf{I}_L, \mathbf{I}_K}$ фробениусово скалярное произведение.

Так как $\mathcal{T}_L(E_j)\mathbf{e}_{k,m}$ содержит только один ненулевой m -ый столбец, получаем

$$(\mathcal{T}_L(E_j)\mathbf{e}_{k,m})_{:,m} = \begin{cases} E_{j-k+1}, & 0 \leq j-k \leq L-1, \\ 0_L, & \text{в противном случае.} \end{cases}$$

Более того,

$$(\mathbf{L}\mathcal{T}_L(E_i))_{:,m} = \begin{cases} \mathbf{L}_{:,i-m+1}, & 0 \leq i-m \leq L-1, \\ 0_L, & \text{в противном случае.} \end{cases}$$

Поэтому $\mathbf{W}_{m,k}$ имеет следующий вид:

$$w_{i,j}^{(m,k)} = \begin{cases} l_{j-k+1,i-m+1}, & 0 \leq j-k \leq L-1, 0 \leq i-m \leq L-1, \\ 0, & \text{в противном случае.} \end{cases}$$

Мы можем переписать полученный результат в виде следующей свёртки:

$$w_{i,j} = \sum_{i_1+i_2-1=i} \sum_{j_1+j_2-1=j} l_{i_1,j_1} r_{i_2,j_2},$$

что означает $\mathbf{W} = \mathbf{L} * \mathbf{R}$ по определению свёртки двух матриц.

(i, j) -й элемент матрицы $\mathbf{W}$ лежит на $(j-i)$ -й диагонали $\mathbf{W}$, где $j-i = (j_1-i_1) + (j_2-i_2)$, что означает, что мы можем переписать полученную свёртку в следующем виде:

$$w_{i,j} = \sum_p \sum_{j_1-i_1=p} l_{i_1,j_1} r_{i+1-i_1,j+1-j_1},$$

изменение порядка суммирования которой завершает доказательство. $\square$

Частный случай теоремы 4.2.1 был рассмотрен в [22] и в [28].

Пусть Σ является автоковариационной матрицей процесса авторегрессии (AR) порядка p . Известно, что обратная к Σ матрица $\mathbf{W}_0 = \Sigma^{-1} - (2p+1)$ -диагональная [49]. Поэтому нам интересен случай $(2p+1)$ -диагональной $\mathbf{W}$.

Приведём очевидные следствия из теоремы 4.2.1.

Следствие 4.2.1. *Рассмотрим матрицы $\mathbf{L} \in \mathbb{R}^{L \times L}$, $\mathbf{R} \in \mathbb{R}^{K \times K}$ и $\mathbf{W}$ такие, что $\|\mathcal{T}_L(\mathbf{Z})\|_{\mathbf{L}, \mathbf{R}} = \|\mathbf{Z}\|_{\mathbf{W}}$ для любого $\mathbf{Z}$. Пусть $\mathbf{L}$ будет $(2p_1 + 1)$ -диагональной, и $\mathbf{R}$ — $(2p_2 + 1)$ -диагональной. Тогда матрица весов $\mathbf{W} = \mathbf{L} * \mathbf{R}$ — $(2p + 1)$ -диагональная, где $p = 2p_1 + 2p_2 + 1$ и i -я диагональ $\text{diag}_{\mathbf{W}}(i)$, $i = -p, \dots, p$, задана в (4.5).*

Важный частный случай — $(2p + 1)$ -диагональная $\mathbf{L}$ и диагональная $\mathbf{R}$, т.е. $p_1 = p$, $p_2 = 0$.

Следствие 4.2.2. *Рассмотрим $(2p + 1)$ -диагональную матрицу $\mathbf{L}$ и диагональную $\mathbf{R} = \text{diag}(R)$. Тогда матрица весов $\mathbf{W}$ является $(2p + 1)$ -диагональной, где $\text{diag}_{\mathbf{W}}(i) = \text{diag}_{\mathbf{L}}(i) * R$, $i = -p, \dots, p$.*

4.2.2. Эквивалентные формулировки задачи поиска весов

Нас интересует следующий вопрос: можно ли подобрать подходящие $\mathbf{L}$ и $\mathbf{R}$ для заданной обратной автоковариационной матрицы $\mathbf{W}_0 = \Sigma_N^{-1}$ процесса авторегрессии порядка p с коэффициентами $b_1, \dots, b_p$ так, что $\mathbf{W}_0 = \mathbf{L} * \mathbf{R}$. Тогда было бы возможно решать задачу (3) с весами, соответствующими авторегрессионному шуму, методами решения задачи (4).

К сожалению, уже для диагонального случая $p = 0$ не существует таких невырожденных $\mathbf{L}$, $\mathbf{R}$, что $\mathbf{W}(\mathbf{L}, \mathbf{R}) = \mathbf{L} * \mathbf{R} = \mathbf{I}_N$ [23]. Поэтому необходимо поставить задачу поиска $\mathbf{L}$ и $\mathbf{R}$, которые бы дали $\mathbf{W}(\mathbf{L}, \mathbf{R})$, близкую к Σ_N^{-1} .

Рассмотрим разумную меру между обратной ковариационной матрицей $\mathbf{W}_0 = \Sigma_N^{-1}$ и матрицей весов $\mathbf{W}$, которая не зависит от вида сигнала или его ранга. В качестве такой меры мы выбрали дивергенцию Кульбака-Лейблера [50] между многомерными нормальными распределениями с нулевыми средними и ковариационными матрицами Σ и $\mathbf{W}^{-1}$. По определению, дивергенция

выглядит следующим образом [51]:

$$D_{\text{KL}}(\mathcal{N}(0_N, \mathbf{W}_0^{-1}) \parallel \mathcal{N}(0_N, \mathbf{W}^{-1})) = \frac{1}{2} (\text{tr}(\mathbf{W}\mathbf{W}_0^{-1}) - N - \log(\det(\mathbf{W}\mathbf{W}_0^{-1}))).$$

Тем не менее, данную формулу трудно использовать на практике из-за вычислительной сложности задачи её минимизации. Поэтому мы перепишем

$$\log(\det(\mathbf{W}\mathbf{W}_0^{-1})) = \text{tr}(\log(\mathbf{W}\mathbf{W}_0^{-1})),$$

и разложим $\log(\mathbf{W}\mathbf{\Sigma})$ в ряд Тейлора в окрестности точки $\mathbf{I}_N$ до второго члена. Получим разложение дивергенции Кульбака-Лейблера второго порядка [52]:

$$D_{\text{KL2}}(\mathcal{N}(0_N, \mathbf{W}_0^{-1}) \parallel \mathcal{N}(0_N, \mathbf{W}^{-1})) = \frac{1}{2} \text{tr}(\mathbf{W}\mathbf{W}_0^{-1}\mathbf{W}\mathbf{W}_0^{-1}) - \text{tr}(\mathbf{W}\mathbf{W}_0^{-1}) + \frac{N}{2}. \quad (4.6)$$

Для краткости, будем обозначать $D_{\text{KL2}}(\mathcal{N}(0_N, \mathbf{W}_0^{-1}) \parallel \mathcal{N}(0_N, \mathbf{W}^{-1}))$ как $D_{\text{KL2}}(\mathbf{W}_0^{-1}, \mathbf{W}^{-1})$.

Так как задача минимизации $D_{\text{KL2}}(\mathbf{W}_0^{-1}, \mathbf{W}^{-1})$, где $\mathbf{W} = \mathbf{W}(\mathbf{L}, \mathbf{R})$, одновременно по $\mathbf{L}$ и $\mathbf{R}$ достаточно сложна, мы ограничимся случаем фиксированной матрицы $\mathbf{L}$ и поставим условие на невырожденность $\mathbf{R}$. Мы ограничим её число обусловленности $\text{cond } \mathbf{R} \leq 1/\alpha$, $0 < \alpha \leq 1$.

В качестве матрицы $\mathbf{L}$ выберем $\mathbf{L} = \mathbf{\Sigma}_L^{-1}$ — обратную автоковариационную матрицу процесса авторегрессии длины L с теми же коэффициентами $b_1, \dots, b_p$, что и у $\mathbf{L} = \mathbf{\Sigma}_N^{-1}$. Соответственно, по следствию 4.2.2, $\mathbf{R}$ ищется среди диагональных матриц.

Замечание 4.2.1. Выбор $\mathbf{L}$ имеет следующее объяснение. В случае $p = 0$ матрица $\mathbf{L}$ пропорциональна $\mathbf{I}_L$, то есть эта матрица является наименее обусловленной среди матриц размера $L \times L$. Более того, известно, что если N кратно L , то без ограничения на число обусловленности матрицы $\mathbf{R} = \text{diag}(R)$ можно добиться равномерных весов $\mathbf{W}_0 = \mathbf{I}_N$ [23]. При этом, согласно следствию 4.2.2, задачу подбора весов можно представить как подбор такого вектора R , что $\mathbf{1}_L * R \approx \mathbf{1}_N$.

Если же рассмотреть случай $p > 0$, то согласно [49] мы получим похожее соотношение для каждой из $2p + 1$ диагоналей, лишь за исключением $p - i$ значений на крае. Например, при $p = 1$ с точностью до умножения на константу выполнено: $\text{diag}_{\Sigma_N^{-1}}(0) = (1, 1 + b_1^2, \dots, 1 + b_1^2, 1)^T \in \mathbb{R}^N$, $\text{diag}_{\Sigma_L^{-1}}(0) = (1, 1 + b_1^2, \dots, 1 + b_1^2, 1)^T \in \mathbb{R}^L$, $\text{diag}_{\Sigma_N^{-1}}(1) = (-b_1, \dots, -b_1)^T \in \mathbb{R}^{N-1}$, $\text{diag}_{\Sigma_N^{-1}}(1) = (-b_1, \dots, -b_1)^T \in \mathbb{R}^{L-1}$, при этом мы хотим получить $\text{diag}_{\Sigma_N^{-1}}(0) \approx \text{diag}_{\Sigma_L^{-1}}(0) * R$, $\text{diag}_{\Sigma_N^{-1}}(1) \approx \text{diag}_{\Sigma_L^{-1}}(1) * R$.

Таким образом, мы ставим задачу поиска весов для $\mathbf{W}_0 = \Sigma_N^{-1}$, $\mathbf{L} = \Sigma_L^{-1}$ в следующем виде:

$$\mathbf{R}_{\text{KL2}}^* = \arg \min_{\substack{\text{cond } \mathbf{R} \leq 1/\alpha \\ \mathbf{W} = \Sigma_L^{-1} * \mathbf{R} \\ \mathbf{R} = \text{diag}(R)}} D_{\text{KL2}}(\Sigma_N, \mathbf{W}^{-1}). \quad (4.7)$$

Заметим, что если шум белый, то автоковариационная матрица Σ пропорциональна единичной, и задача сводится к задаче минимизации евклидового расстояния между диагоналями матриц $\mathbf{W}_0$ и $\mathbf{W}$, что приводит к следующей постановке:

$$\mathbf{R}_{\text{dist}}^* = \arg \min_{\substack{\text{cond } \mathbf{R} \leq 1/\alpha \\ \mathbf{R} = \text{diag}(R)}} \frac{1}{2} \|\mathbb{1}_L * R - \mathbb{1}_N\|^2. \quad (4.8)$$

Оказывается, задачи (4.7) и (4.8) эквивалентны во многих случаях, достаточно лишь, чтобы L было немногим больше половины длины ряда N . Будем обозначать под $\mathbf{0}_{L,K}$ нулевую матрицу размера $L \times K$.

Теорема 4.2.2. *Если или $p = 0$, или $p > 0$ и $L \geq K + p - 1$, то задачи (4.7) и (4.8) эквивалентны, то есть минимум функционалов достигается на одной и той же $\mathbf{R}$.*

Доказательство. Так как условия на $\mathbf{R}$ в (4.7) и (4.8) совпадают, достаточно доказать, что минимизируемые функционалы в (4.7) и (4.8) — одна и та же

квадратичная функция. Перепишем $\mathbb{1}_L * R = \mathbf{Q}^{N,L}(\mathbb{1}_L)R$, где $\mathbf{Q}^{N,L}$ определено в (1.2), и раскроем (4.8). Получим

$$\frac{1}{2}\|\mathbb{1}_L * R - \mathbb{1}_N\|^2 = \frac{1}{2}R^T \mathbf{S} R - L\mathbb{1}_K^T R + \frac{N}{2}, \quad (4.9)$$

где $\mathbf{S} = (s_{i,j})$, $s_{i,j} = L - |i - j|$, если $|i - j| \leq L$, и $s_{i,j} = 0$ в противном случае.

Для того чтобы раскрыть (4.7), введём разложение $\mathbf{W} = \sum_{k=1}^K r_k \mathbf{W}_k$ таким же способом, которым это было сделано в предложении 4.2.1, с $r_{k,k} = r_k$ и $\mathbf{W}_{k,k} = \mathbf{W}_k$:

$$\begin{aligned} D_{\text{KL2}}(\mathbf{W}_0^{-1}, \mathbf{W}^{-1}) &= \\ &= \frac{1}{2} \sum_{i=1}^K \sum_{j=1}^K r_i r_j \text{tr}(\mathbf{W}_i \mathbf{W}_0^{-1} \mathbf{W}_j \mathbf{W}_0^{-1}) - \sum_{i=1}^K r_i \text{tr}(\mathbf{W}_i \mathbf{W}_0^{-1}) + \frac{N}{2}. \end{aligned} \quad (4.10)$$

Представим $\mathbf{W}_k$ и $\mathbf{W}_0^{-1} = \mathbf{\Sigma} = \mathbf{\Sigma}_N$ в виде блочных матриц (заметим, что размер блока зависит от k):

$$\begin{aligned} \mathbf{W}_k &= \begin{pmatrix} \mathbf{0}_{k-1,k-1} & \mathbf{0}_{k-1,L} & \mathbf{0}_{k-1,K-k} \\ \mathbf{0}_{L,k-1} & \mathbf{\Sigma}_L^{-1} & \mathbf{0}_{L,K-k} \\ \mathbf{0}_{K-k,k-1} & \mathbf{0}_{K-k,L} & \mathbf{0}_{K-k,K-k} \end{pmatrix}, \\ \mathbf{\Sigma} &= \begin{pmatrix} \mathbf{\Sigma}_{k-1} & \mathbf{\Sigma}_{\text{top}} & \mathbf{\Sigma}_{\text{topright}} \\ \mathbf{\Sigma}_{\text{left}} & \mathbf{\Sigma}_L & \mathbf{\Sigma}_{\text{right}} \\ \mathbf{\Sigma}_{\text{bottomleft}} & \mathbf{\Sigma}_{\text{bottom}} & \mathbf{\Sigma}_{K-k} \end{pmatrix}. \end{aligned}$$

Обозначим γ_i значение автоковариационной функции процесса AR(p) с коэффициентами $b_1, \dots, b_p$ с лагом i ; $\gamma_{-i} = \gamma_i$. Тогда $\mathbf{\Sigma} = (\gamma_{i-j})$. Также обозначим $S_i = (\gamma_{i+L-1}, \dots, \gamma_i)^T \in \mathbb{R}^L$. Заметим, что $\mathbf{\Sigma}_L = [S_{-L+1} : \dots : S_0]$, $\mathbf{\Sigma}_{\text{right}} = [S_1 : \dots : S_{K-k}]$. Из уравнений Юла-Уолкера [53] следует, что $\gamma_k = \sum_{i=1}^p b_i \gamma_{k-i}$, $k \geq 1$. Таким образом, $S_k = \sum_{i=1}^p b_i S_{k-i}$, $k = 1, \dots, K - k$. В итоге, мы получили, что $\text{colspace } \mathbf{\Sigma}_{\text{right}} = \text{span}(S_{1-p}, \dots, S_0)$. По аналогии и симметрии, имеем $\text{colspace } \mathbf{\Sigma}_{\text{left}} = \text{span}(S_{-L+1}, \dots, S_{-L+p})$.

Произведение $\mathbf{W}_k$ и $\mathbf{W}_0^{-1} = \Sigma$ имеет следующий вид:

$$\mathbf{W}_k \mathbf{W}_0^{-1} = \begin{pmatrix} \mathbf{0}_{k-1,k-1} & \mathbf{0}_{k-1,L} & \mathbf{0}_{k-1,K-k} \\ \Sigma_L^{-1} \Sigma_{\text{left}} & \mathbf{I}_L & \Sigma_L^{-1} \Sigma_{\text{right}} \\ \mathbf{0}_{K-k,k-1} & \mathbf{0}_{K-k,L} & \mathbf{0}_{K-k,K-k} \end{pmatrix},$$

при этом выполняются соотношения

$$\text{colspace } \Sigma_L^{-1} \Sigma_{\text{left}} = \text{span}(E_1, \dots, E_p),$$

$$\text{colspace } \Sigma_L^{-1} \Sigma_{\text{right}} = \text{span}(E_{L-p+1}, \dots, E_L).$$

Это означает, что только первые p строк $\Sigma_{L,p}^{-1} \Sigma_{\text{left}}$ и только последние p строк матрицы $\Sigma_{L,p}^{-1} \Sigma_{\text{right}}$ ненулевые.

Вернёмся обратно к разложению (4.10). Очевидно, что $\text{tr}(\mathbf{W}_i \mathbf{W}_0^{-1}) = L$ для любого $i = 1, \dots, K$. Равенство

$$\text{tr}(\mathbf{W}_i \mathbf{W}_0^{-1} \mathbf{W}_j \mathbf{W}_0^{-1}) = \langle \mathbf{W}_i \mathbf{W}_0^{-1}, (\mathbf{W}_j \mathbf{W}_0^{-1})^T \rangle_{\text{F}}$$

означает, что результат зависит от того, как перекрываются ненулевые части $\mathbf{W}_i \mathbf{W}_0^{-1}$ и $(\mathbf{W}_j \mathbf{W}_0^{-1})^T$, см. рисунок 4.1. Когда $L \geq K + p - 1$, ненулевые части $\Sigma_L^{-1} \Sigma_{\text{left}}$ и $\Sigma_L^{-1} \Sigma_{\text{right}}$ внутри $\mathbf{W}_i \mathbf{W}_0^{-1}$ и $(\mathbf{W}_j \mathbf{W}_0^{-1})^T$ не перекрываются для любых $1 \leq i, j \leq K$. Следовательно, $\text{tr}(\mathbf{W}_i \mathbf{W}_0^{-1} \mathbf{W}_j \mathbf{W}_0^{-1}) = s_{i,j}$.

В случае $p = 0$, $\Sigma_{\text{left}} = \mathbf{0}_{L,k-1}$ и $\Sigma_{\text{right}} = \mathbf{0}_{L,K-k}$, мы получаем $\text{tr}(\mathbf{W}_i \mathbf{W}_0^{-1} \mathbf{W}_j \mathbf{W}_0^{-1}) = s_{i,j}$ для любого L .

Следовательно, мы доказали, что правая часть выражения (4.9) совпадает с правой частью (4.10). $\square$

Так как задачу (4.7) трудно решать, мы используем задачу (4.8) для нахождения весов для любого p . В теореме 4.2.2 показано, что функционал (4.8) квадратичный. Следовательно, с добавлением вспомогательной переменной, задача (4.8) может быть решена, используя методы квадратичного программирования, что показано в разделе 4.3 и [27].

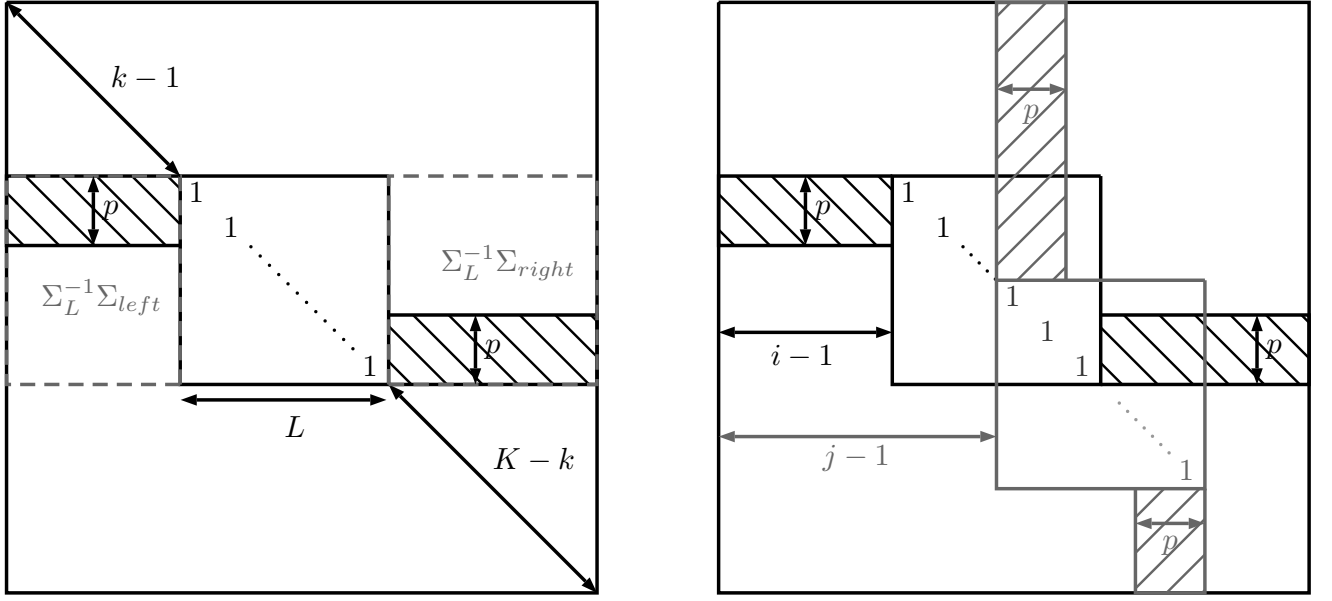


Рис. 4.1. Схематичные изображения матрицы $\mathbf{W}_k \mathbf{W}_0^{-1}$ (слева) и перекрывающихся матриц $\mathbf{W}_i \mathbf{W}_0^{-1}$ и $(\mathbf{W}_j \mathbf{W}_0^{-1})^T$ (справа).

4.3. Применение квадратичного программирования к поиску весов

В этом разделе мы рассмотрим алгоритм решения задачи поиска весов (4.8). Будем считать, что $\alpha < 1$ в (4.8), так как случай $\alpha = 1$ тривиален (вектор R состоит из одинаковых значений).

4.3.1. Эквивалентные формулировки

Преобразуем задачу (4.8): множество $\mathcal{R} = \{R : \text{cond}(\text{diag } \mathbf{R}) \leq 1/\alpha\}$ из (4.8) запишем в виде набора линейных ограничений, а в эквивалентной целевой функции (4.9) опустим константный член. Получим следующую эквивалентную формулировку:

Задача 4.3.1.

$$f^* = \min_{R \in \mathcal{R}} f(R), \quad \text{где} \quad f(R) = \frac{1}{2} R^T \mathbf{S} R - L \mathbf{1}_K^T R, \quad (4.11)$$

$$\mathcal{R} = \{R \mid r_i - \alpha r_j \geq 0, \quad i \neq j, \quad 1 \leq i, j \leq K\}, \quad (4.12)$$

где $\mathbf{S} = (s_{i,j})$, $s_{i,j} = L - |i - j|$ если $|i - j| \leq L$, и 0 в противном случае, $\mathbf{1}_K = (1, \dots, 1)^T$, $\mathbf{1}_K \in \mathbb{R}^K$.

Теперь задача 4.3.1 представлена в том виде, для которого разработана теория квадратичного программирования (КП) [54, р. 111]. Справедливо следующее:

Предложение 4.3.1. 1. Задача 4.3.1 имеет единственное решение R^* .

2. Её решение $R^* = (r_1^*, \dots, r_K^*)^T$ является палиндромом, то есть для любого индекса $1 \leq i \leq K$: $r_i^* = r_{K-i+1}^*$.

Доказательство. 1. Задача 4.3.1 — задача КП с набором из K^2 линейных ограничений (4.12) и целевой функцией (4.11), квадратичная форма в которой положительно определена, поэтому решение такой задачи единственно [35, р. 452].

2. Достаточно рассмотреть вектор $R^{**} = (r_K^*, \dots, r_1^*)$ и заметить, что $f(R^*) = f(R^{**})$ и $R^{**} \in \mathcal{R}$ в (4.12); значит, $R^* = R^{**}$, что и требовалось доказать.

□

Для решения задачи 4.3.1 можно использовать методы квадратичного программирования, но, к сожалению, K^2 линейных ограничений являются серьёзной проблемой при решении задачи на практике. Рассмотрим ещё две эквивалентных формулировки, свойства которых используются при решении задачи.

Сделаем дополнительное предположение: пусть в точке j веса достигают своего максимума (тем самым, снизим число линейных ограничений до линейного по K размера), а само j будем перебирать в цикле от 1 до $\lceil K/2 \rceil$. При этом воспользуемся тем фактом, что решение является палиндромом. Формально, задача выглядит следующим образом:

Задача 4.3.2.

$$f^{\star\star} = \min_{j=1, \dots, \lceil K/2 \rceil} f_j^{\star}, \quad (4.13)$$

$$f_j^{\star} = \min_{R \in \mathcal{R}_j} f(R), \quad (4.14)$$

$$\mathcal{R}_j = \{R \mid r_i = r_{K-i+1}, i = 1, \dots, \lceil K/2 \rceil; \quad (4.15)$$

$$r_i - \alpha r_j \geq 0, i = 1, \dots, j-1, j+1, \dots, \lceil K/2 \rceil; \quad (4.16)$$

$$r_j - r_i \geq 0, i = 1, \dots, j-1, j+1, \dots, \lceil K/2 \rceil\} \quad (4.17)$$

Чтобы не перебирать все возможные j , удобно уметь проверять, даёт ли полученное на очередной итерации решение глобальный минимум. Для такой проверки рассмотрим ещё одну эквивалентную формулировку задачи 4.3.1. Введём $r_{\max}$ — дополнительную переменную, хранящую максимальный вес, и перейдём к новым переменным $\hat{r}_i = r_{\max} - r_i$, $i = 1, \dots, K$ — разница между максимальным среди всех и текущим весом, при этом $r_{\max} \geq 0$, $\hat{r}_i \geq 0$. Между векторами $R = (r_1, \dots, r_K)^T$ и $\hat{R} = (\hat{r}_1, \dots, \hat{r}_K, r_{\max})^T$ существует простое линейное соответствие: $R = \mathbf{H}\hat{R}$, где $\mathbf{H} \in \mathbb{R}^{K \times (K+1)}$ — матрица следующего вида:

$$\mathbf{H} = \begin{pmatrix} -1 & 0 & 1 \\ & \ddots & \vdots \\ 0 & -1 & 1 \end{pmatrix}. \quad (4.18)$$

Условие $\text{cond}(\text{diag } \mathbf{R}) \leq 1/\alpha$, устанавливающее границу снизу для весов, в новых обозначениях записывается как $(1 - \alpha)r_{\max} - \hat{r}_i \geq 0$, $i = 1, \dots, K$.

Получаем следующую задачу квадратичного программирования с линейными ограничениями, но с нестрогой выпуклой целевой функцией:

Задача 4.3.3.

$$\hat{f}^{\star} = \min_{\hat{R} \in \hat{\mathcal{R}}} \hat{f}(\hat{R}), \quad \text{где} \quad \hat{f}(\hat{R}) = f(\mathbf{H}\hat{R}) = \frac{1}{2} \hat{R}^T \mathbf{H}^T \mathbf{S} \mathbf{H} \hat{R} - L \mathbf{1}_K^T \mathbf{H} \hat{R}, \quad (4.19)$$

$$\hat{\mathcal{R}} = \left\{ \hat{R} \mid \hat{r}_j \geq 0, \quad (1 - \alpha)r_{\max} - \hat{r}_j \geq 0, \quad j = 1, \dots, K \right\}. \quad (4.20)$$

Теорема 4.3.1. *Задачи 4.3.1, 4.3.2 и 4.3.3 эквивалентны.*

Доказательство. Требуется доказать, что $f^* \geq f^{**} \geq \hat{f}^* \geq f^*$, и что по решению одной из задач можно построить решение любой другой.

1. ($f^* \geq f^{**}$). Пусть R^* — решение задачи 4.3.1, то есть $R^* = \arg \min_{R \in \mathcal{R}} f(R)$. Положим $j^* = \arg \max_i r_i$. $R^* \in \mathcal{R}$ и R^* — палиндром, следовательно, $R^* \in \mathcal{R}_{j^*}$. Цепочка неравенств $f^* = f(R^*) \geq f_{j^*}^* = \min_{R \in \mathcal{R}_{j^*}} f(R) \geq \min_{j=1, \dots, \lceil K/2 \rceil} \min_{R \in \mathcal{R}_j} f(R) = f^{**}$ завершает доказательство пункта.
2. ($f^{**} \geq \hat{f}^*$). Пусть $j^* = \arg \min_{j=1, \dots, \lceil K/2 \rceil} f_j^*$, $R_{j^*}^* = \arg \min_{R \in \mathcal{R}_{j^*}} f(R)$ — решение задачи 4.3.2. Положим $r_{\max} = r_{j^*, j^*}$, $\hat{r}_i = r_{\max} - r_{j^*, i}$, $i = 1, \dots, K$, где $R_{j^*}^* = (r_{j^*, 1}, \dots, r_{j^*, K})^T$. Нетрудно заметить, что $\hat{R} \in \hat{\mathcal{R}}$, из чего получим, что $\hat{f}^* \leq \hat{f}(\hat{R}) = f(R_{j^*}^*) = f^{**}$.
3. ($\hat{f}^* \geq f^*$). Пусть $\hat{R}^* = \arg \min_{\hat{R} \in \hat{\mathcal{R}}} \hat{f}(\hat{R})$ — решение задачи 4.3.3. Положим $r_i = r_{\max} - \hat{r}_i$ и заметим, что $R \in \mathcal{R}$, из чего получим, что $f^* \leq f(R) = \hat{f}(\hat{R}) = \hat{f}^*$.

□

4.3.2. Общий алгоритм решения

Используя эквивалентность, доказанную в теореме 4.3.1, можем сформулировать следующий алгоритм решения Задачи 4.3.1:

Алгоритм 4.3.1. Вход: Параметры L , K , α .

- 1: Положить $j = 1$.
- 2: Решить подзадачу КП (4.14) $R_j^* = (r_{j,1}, \dots, r_{j,K})^T = \arg \min_{R \in \mathcal{R}_j} f(R)$.
- 3: Задать эквивалентный вектор $\hat{R}$ из (4.19), взяв $r_{\max} = r_{j,j}$, $\hat{r}_i = r_{\max} - r_{j,i}$.

4: **if** Вектор $\hat{R}$ — решение задачи 4.3.3 **then**
 5: **return** $R^* = R_j^*$ — вектор оптимальных весов
 6: **else**
 7: Положить $j = j + 1$ и перейти к пункту 2.
 8: **end if**

Таким образом, если встретился нужный индекс j , то алгоритму не понадобится перебирать оставшиеся индексы и решать подзадачу (4.14) лишний раз. Практические эксперименты показывают, что максимальный вес всегда находится на краях: алгоритм 4.3.1 останавливается уже при $j = 1$, то есть фактически сводится к задаче КП с положительно определённой квадратичной формой в целевой функции.

Для реализации алгоритма 4.3.1 необходимо разработать алгоритмы решения задач из пунктов 2 и 4, что и будет сделано в следующих разделах. Сначала в разделе 4.3.3 рассмотрим пункт 4, а в разделе 4.3.4 — пункт 2.

4.3.3. Проверка вектора на решение задачи 4.3.3

Предложим быстрый алгоритм, проверяющий, является ли заданный вектор $\hat{R}$ точкой, в которой достигается глобальный минимум в задаче 4.3.3, то есть реализующий пункт 4 из алгоритма 4.3.1. Для этого применим теорему о необходимом и достаточном условии минимума в задаче квадратичного программирования для задачи 4.3.3.

Теорема 4.3.2. 1. Рассмотрим вектор $C = (c_1, \dots, c_{K+1})^T = \mathbf{H}^T \mathbf{S} \mathbf{H} \hat{R} - L \mathbf{H}^T \mathbf{1}_K$. Тогда $\hat{R}$ является решением задачи 4.3.3 в том и только в том случае, если:

a. $C^T \hat{R} = 0$,

б. Существует вектор $U = (u_1, \dots, u_K)^T \in \mathbb{R}^K$ такой, что имеют

место неравенства $u_i \geq 0$, $i = 1, \dots, K$, $u_i \geq -c_i$, $i = 1, \dots, K$ и $(1 - \alpha) \sum_{i=1}^K u_i \leq c_{K+1}$.

2. Положим $u_i = \max(0, -c_i)$, $i = 1, \dots, K$. Тогда $\hat{R}$ является решением задачи 4.3.3 в том и только в том случае, если:

a. $C^T \hat{R} = 0$,

б. $(1 - \alpha) \sum_{i=1}^K u_i \leq c_{K+1}$.

Доказательство. Первый пункт данной теоремы является переформулировкой теоремы [54, теорема 9.2] для задачи 4.3.3, поэтому перейдём сразу к доказательству второго пункта.

Условие (а) в обоих пунктах одинаковое. Докажем эквивалентность условия (б).

Достаточность. Коэффициенты u_i выбраны так, чтобы условия $u_i \geq 0$ и $u_i \geq -c_i$ выполнялись. Если удовлетворено и условие (б), то выполняются все требования первого пункта.

Необходимость. Очевидно, что u_i выбраны наименьшими из всех тех, которые удовлетворяют условиям $u_i \geq 0$ и $u_i \geq -c_i$. Пусть $(1 - \alpha) \sum_{i=1}^K u_i > c_{K+1}$. Рассмотрим любой вектор $\tilde{U} = (\tilde{u}_1, \dots, \tilde{u}_N)^T$, удовлетворяющий условию $(1 - \alpha) \sum_{i=1}^K \tilde{u}_i \leq c_{K+1}$. Тогда существует индекс i такой, что $\tilde{u}_i < u_i$. Следовательно, либо $\tilde{u}_i < 0$, либо $\tilde{u}_i < -c_i$, из чего следует, что не существует такого вектора $\tilde{U}$, который бы удовлетворял условиям первого пункта. $\square$

Рассмотрим следующий алгоритм, базирующийся на теореме 4.3.2.

Алгоритм 4.3.2. Вход: предполагаемое решение $\hat{R}$, параметр α .

Результат: Булево значение: является ли $\hat{R}$ решением задачи 4.3.3.

1: Вычислить $C = \mathbf{H}^T \mathbf{S} \mathbf{H} \hat{R} - L \mathbf{H}^T \mathbf{1}_K$.

2: *if* $C^T \hat{R} \neq 0$ *then*

```

3:  return FALSE
4:  else
5:    Перейти к пункту 7.
6:  end if
7:  Выбрать в качестве вектора  $U = (u_1, \dots, u_K)^T$  следующие значения:  $u_i =$ 
     $\max(0, -c_i)$ ,  $i = 1, \dots, K$ .
8:  if  $(1 - \alpha) \sum_{i=1}^K u_i \leq c_{K+1}$  then
9:    return TRUE
10: else
11:  return FALSE
12: end if

```

4.3.4. Задача квадратичного программирования специального вида

Теперь переходим к пункту 2 алгоритма 4.3.1. Для решения подзадачи (4.14) при зафиксированном индексе j воспользуемся алгоритмом 1.5.1. Специфика задачи позволяет эффективно реализовать алгоритм.

Подзадача (4.14) при фиксированном j переписывается в терминах задачи 1.5.1 следующим образом: $X^* = R^*$, $X = R$, $\mathbf{G} = \mathbf{S}$, $V = L\mathbf{1}_K$, (1.10) состоит из (4.15), (1.11) состоит из (4.16) и (4.17).

Таким образом, необходимо объяснить, как находить начальную точку (п. 1 алгоритма 1.5.1), решение подзадачи квадратичного программирования и множители Лагранжа (п. 4 алгоритма 1.5.1) применительно к частному случаю задачи (4.14).

Обозначим вкратце те особенности, которые помогают получить быстрое решение:

- Для выполнения пункта 4 алгоритма 1.5.1 требуется уметь решать задачу (1.12) с ограничениями (1.13). Положив $\mathbf{A} = [A_i : i \in \mathcal{W}_k]$, $\mathbf{A} \in \mathbb{R}^{K \times m}$,

ограничения (1.13) можно записать как $\mathbf{A}^T P_k = 0_m$, где m — количество активных ограничений.

Для решения поставленной вспомогательной задачи (1.12) есть явная формула обобщённого метода наименьших квадратов:

$$P_k^* = \mathbf{A}^\perp ((\mathbf{A}^\perp)^T \mathbf{G} \mathbf{A}^\perp)^{-1} (\mathbf{A}^\perp)^T G_k,$$

где матрица $\mathbf{A}^\perp \in \mathbb{R}^{K \times (K-m)}$ состоит из столбцов, составляющих базис ортогонального дополнения к базису столбцов матрицы $\mathbf{A}$. Обычно m велико, поэтому $K - m$ мало, что позволяет быстро искать решение подзадачи.

В случае задачи 4.3.2 матрица $\mathbf{G}$ имеет простой вид, и её можно умножать на вектор за время $O(K)$. Матрица $\mathbf{A}$ — разрежённая, в ней содержится максимум 2 ненулевых коэффициента в каждом столбце. За счёт этого матрицу $\mathbf{A}^\perp$ можно также найти быстро и хранить, используя $O(K)$ памяти. Из-за быстрого умножения матрицы $\mathbf{A}$ для вычисления формулы обобщённого МНК удобно и эффективно использовать метод сопряжённых градиентов, вычислительная сложность которого $O(K(K - m))$.

Аналогично, за счёт разрежённости матрицы $\mathbf{A}$ можно быстро находить множители Лагранжа за время $O(K)$. Таким образом, вычислительная сложность одного шага алгоритма 1.5.1 при решении подзадачи (4.14) равна $O(K(K - m))$.

- Есть две стратегии выбора начальной точки: первая применяется при малом размере задачи и заключается в том, чтобы назначить первым t точкам, где t перебирается по целочисленной сетке до $\lfloor K/2 \rfloor$, максимальный вес (то есть (4.17)), а оставшимся — минимальный (то есть (4.16)), после чего выбрать то t , где достигается наименьшее значение целевой функции $f(R)$.

Вторая стратегия используется при $j = 1$ и основана на следующем наблюдении: решения задач при примерно одинаковом отношении N к L и одинаковом α схожи. Например, на рисунке 4.2 изображены два отнормированных (умноженных на N) решения задачи 4.3.1 при $N = 200$, $L = 60$, $\alpha = 0.1$ и $N = 1000$, $L = 300$, $\alpha = 0.1$ соответственно.

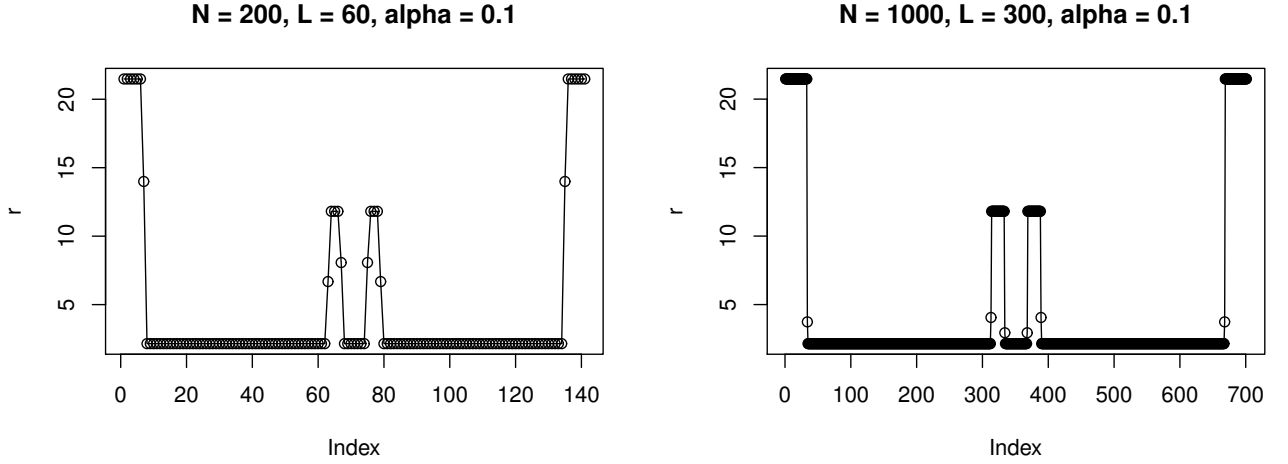


Рис. 4.2. Два решения задачи 4.3.1 при одинаковом отношении N к L .

Схема эвристики следующая: зафиксируем параметр масштаба $0 < \gamma < 1$ (на практике хорошими значениями γ являются $0.5 - 0.7$), найдём решение задачи при $N_\gamma \approx \gamma N$, $L_\gamma \approx \gamma L$, $\alpha_\gamma = \alpha$, после чего используем решение задачи меньшего порядка для выбора начального рабочего множества и начальной точки. Формально, алгоритм выглядит так:

Алгоритм 4.3.3. Вход: параметр γ .

Результат: Начальное рабочее множество ограничений $\mathcal{W}_0$ и стартовая точка X_0 .

- 1: Положить $K_\gamma = \lceil \gamma K \rceil$, $L_\gamma = \lceil \gamma L \rceil$, $N_\gamma = L_\gamma + K_\gamma - 1$.
- 2: Определить функцию $s(i) = \lceil \frac{(i-1)(K_\gamma-1)}{K_\gamma-1} + 1 \rceil$ перехода к уменьшенной размерности.

- 3: Найти масштабированное решение $R_\gamma = (r_{\gamma,1}, \dots, r_{\gamma,K_\gamma})^T$ задачи (4.14) при $K = K_\gamma$, $L = L_\gamma$, $\alpha = \alpha$, $j = j_\gamma = 1$.
- 4: Добавить к рабочему множеству $\mathcal{W}$ индексы всех ограничений вида (4.15).
- 5: **for all** $i = 2, 3, \dots, \lceil K/2 \rceil$ **do**
- 6: Вычислить $i_\gamma = s(i)$.
- 7: **if** $r_{\gamma,i_\gamma} - \alpha r_{\gamma,j_\gamma} = 0$ **then**
- 8: Добавить в $\mathcal{W}$ индекс ограничения вида $\{r_i - \alpha r_j \geq 0\}$ из (4.16), и перейти к следующей итерации.
- 9: **end if**
- 10: **if** $r_{\gamma,i_\gamma} = r_{\gamma,j_\gamma}$ **then**
- 11: Добавить в $\mathcal{W}$ индекс ограничения вида $\{r_j - r_i \geq 0\}$ из (4.17), и перейти к следующей итерации.
- 12: **end if**
- 13: **end for**
- 14: **for** $k = 1, 2, \dots$ **do**
- 15: Решить следующую задачу квадратичного программирования, вычислить значение точки минимума R^* :
- $$R^* = \arg \min_{R \in \mathbb{R}_k} f(R), \quad \mathbb{R}_k = \{R \mid A_i^T R = 0, \quad i \in \mathcal{W}\},$$
- где $f(R)$ определено в (4.11).
- 16: **if** для точки R^* выполняются все условия вида (4.16) и (4.17) **then**
- 17: **return** $X_0 = R^*$, $\mathcal{W}_0 = \mathcal{W}$.
- 18: **else**
- 19: Добавить к рабочему множеству $\mathcal{W}$ индекс любого ограничения, которое не выполняется.
- 20: **end if**
- 21: **end for**

Заметим, что шаг 5 алгоритма 4.3.3, необходимый для того, чтобы получить корректную начальную точку, совершит не более $\lceil K/2 \rceil$ итераций вследствие того, что каждая итерация на единицу уменьшает размерность подпространства $\mathbb{R}_k$, в которой лежит начальная точка R^* .

Таким образом, при использовании данной эвристики получаем рекурсивный алгоритм, который уменьшает размерность задачи в γ раз до тех пор, пока K не станет достаточно малым для использования первой стратегии.

Для решения задачи поиска весов было применено несколько нетривиальных переходов. Сформулированная в терминах минимизации квадратичного функционала задача 4.3.1 содержит K^2 линейных ограничений, что не позволяет эффективно искать решение на практике. Поэтому была выведена эквивалентная задача 4.3.2, решение которой – минимум из решений $K/2$ задач квадратичного программирования с количеством ограничений порядка K . Преимущество формулировки 4.3.2 состоит в том, что перебор задач КП можно остановить заранее, используя разработанный метод остановки, реализованный в алгоритме 4.3.2. На практике перебор ограничивается уже на первой итерации, вследствие чего решение находится эффективно.

4.4. Поиск весов с помощью минимизации гладкой функции

В этом разделе рассмотрим поиск весов при иных предположениях, чем в разделе 4.2. Рассмотрим задачу (4.7), но в случае белого нестационарного шума, т.е. с $\mathbf{W}_0^{-1} = \mathbf{\Sigma} = \text{diag}(\sigma_1^2, \dots, \sigma_N^2)^T$. Как и в разделе 4.2, для упрощения зафиксируем матрицу $\mathbf{\Sigma}_L^{-1} = \mathbf{I}_L$, соответствующую стационарному белому шуму.

Для указанного частного случая получим следующую эквивалентную за-

дачу:

$$\mathbf{R}_{\text{nonstat}}^* = \arg \min_{\substack{\text{cond } \mathbf{R} \leq 1/\alpha \\ \mathbf{R} = \text{diag}(R)}} \frac{1}{2} \|\Sigma \mathbf{Q}^{N,L}(\mathbb{1}_L)R - \mathbb{1}_N\|^2. \quad (4.21)$$

Оказывается, данную задачу можно свести к следующей задаче оптимизации в параллелепипеде:

$$\mathbf{R}_{\text{nonstateq}}^* = \arg \min_{\substack{\max \tilde{R} \leq 1 \\ \min \tilde{R} \geq \alpha}} \frac{1}{2} \left\| \frac{\langle \mathbb{1}_L, \Sigma \mathbf{Q}^{N,L}(\mathbb{1}_L) \tilde{R} \rangle}{\|\Sigma \mathbf{Q}^{N,L}(\mathbb{1}_L) \tilde{R}\|^2} \Sigma \mathbf{Q}^{N,L}(\mathbb{1}_L) \tilde{R} - \mathbb{1}_N \right\|^2, \quad (4.22)$$

где $\langle \cdot, \cdot \rangle$ обозначает обычное скалярное произведение.

Теорема 4.4.1. *Задачи (4.21) и (4.22) эквивалентны, т.е. $\mathbf{R}_{\text{nonstat}}^* = \mathbf{R}_{\text{nonstateq}}^*$.*

Доказательство. Сделаем замену переменных, введя дополнительный параметр c : пусть $R = c\tilde{R}$, при этом каждый из элементов $\tilde{R}$ ограничен снизу α , а сверху 1, таким образом, $\text{cond } \mathbf{R} \leq 1/\alpha$. Тогда задача (4.21) принимает вид

$$\arg \min_{\substack{\max \tilde{R} \leq 1 \\ \min \tilde{R} \geq \alpha}} \left(\min_c \frac{1}{2} \|\Sigma c \mathbf{Q}^{N,L}(\mathbb{1}_L) \tilde{R} - \mathbb{1}_N\|^2 \right). \quad (4.23)$$

Применим к полученной задаче принцип Variable projection (см. раздел 1.3.2) по новой переменной c , по которой задача минимизации является квадратичной. Мы получим:

$$c^* = \frac{\langle \mathbb{1}_L, \Sigma \mathbf{Q}^{N,L}(\mathbb{1}_L) \tilde{R} \rangle}{\|\Sigma \mathbf{Q}^{N,L}(\mathbb{1}_L) \tilde{R}\|^2},$$

и подстановка c^* в (4.23) даёт нам вид (4.22). □

Замечание 4.4.1. *В формулировке участвует функция (4.22), которая не является квадратичной, но является гладкой в параллелепипеде. Это означает, что для численной минимизации этой функции можно использовать достаточно эффективный метод L -BFGS-B [55].*

Замечание 4.4.2. В случае равных весов $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_N^2$ можно использовать решение задачи (4.21) вместо решения задачи (4.8) с использованием квадратичного программирования. Плюс данного подхода — простота, так как не требуется реализовывать сложные алгоритмы квадратичного программирования, если есть готовая реализация алгоритма *L-BFGS-B*. Минус подхода — не гарантируется сходимость оптимизационного метода к глобальному минимуму, в отличие от метода квадратичного программирования.

4.5. Быстрая реализация алгоритма Кэдзоу

4.5.1. Общая идея быстрой реализации

В этом разделе мы обсудим быструю реализацию метода итераций Кэдзоу, заданных в (4.1). Оригинальная идея была предложена в работах [56, 5], и представляет из себя следующее: рассмотрим одну итерацию метода Кэдзоу: $\mathbf{X}_{j+1} = \mathcal{T}_L^{-1}(\Pi_{\mathcal{H}}\Pi_{\mathcal{M}_r}\mathcal{T}_L(\mathbf{X}_j))$, и перепишем результат $\Pi_{\mathcal{M}_r}\mathcal{T}_L(\mathbf{X}_j)$ в виде суммы матриц единичного ранга: $\Pi_{\mathcal{M}_r}\mathcal{T}_L(\mathbf{X}_j) = \sum_{i=1}^r \hat{\mathbf{X}}_i$, $\text{rank } \hat{\mathbf{X}}_i = 1$. Затем мы можем вычислить $\mathbf{X}_{j+1}$ как $\mathbf{X}_{j+1} = \sum_{i=1}^r \mathcal{T}_L^{-1}(\Pi_{\mathcal{H}}(\hat{\mathbf{X}}_i))$. Проекторы $\Pi_{\mathcal{M}_r}$ и $\Pi_{\mathcal{H}}$ (в применении к матрицам ранга 1) могут быть реализованы эффективно с точки зрения трудоёмкости и потребляемой памяти благодаря использованию быстрого преобразования Фурье. Предлагается расширение оригинальной идеи со случая фробениусовской нормы на случай матричной нормы, порождённой скалярным произведением (5), в определении которой матрицы $\mathbf{L}$ и $\mathbf{R}$ являются ленточными.

4.5.2. Вычисление проектора $\Pi_{\mathcal{M}_r}$

Перейдём к первому шагу итерации метода Кэдзоу — вычислению проекции на множество $\mathcal{M}_r$. Приведём следующую теорему из [18], которая даёт решение задачи проектирования $\Pi_{\mathcal{M}_r} \mathcal{T}_L(\mathbf{X}_i)$.

Теорема 4.5.1. *[18, Theorem 1] Рассмотрим следующую задачу: пусть задана матрица $\mathbf{A} \in \mathbb{R}^{L \times K}$, требуется вычислить $\mathbf{B} = \Pi_{\mathcal{M}_r} \mathbf{A}$ по норме, порождённой скалярным произведением (5). Пусть $\mathbf{L} \in \mathbb{R}^{L \times L}$ и $\mathbf{R} \in \mathbb{R}^{K \times K}$ из (5) представлены в виде $\mathbf{L} = \mathbf{O}_L \mathbf{O}_L^T$, $\mathbf{R} = \mathbf{O}_R \mathbf{O}_R^T$, где $\mathbf{O}_L \in \mathbb{R}^{L \times L}$, $\mathbf{O}_R \in \mathbb{R}^{K \times K}$. Введём матрицу $\tilde{\mathbf{A}} = \mathbf{O}_L^T \mathbf{A} \mathbf{O}_R$, и вычислим её сингулярное разложение: обозначим первые r сингулярных чисел как $\sigma_1, \dots, \sigma_r$, сингулярные векторы как $U'_1, \dots, U'_r \in \mathbb{R}^L$, $V'_1, \dots, V'_r \in \mathbb{R}^K$. Тогда выполняется $\mathbf{B} = \sum_{i=1}^r \sigma_i U_i V_i^T$, где $U_i = (\mathbf{O}_L^T)^{-1} U'_i$, $V_i = (\mathbf{O}_R^T)^{-1} V'_i$.*

Рассмотрим $\mathbf{L}$ и $\mathbf{R}$ — $(2p_1 + 1)$ - и $(2p_2 + 1)$ -диагональные положительно определённые матрицы вместе с их разложениями Холецкого $\mathbf{O}_L \mathbf{O}_L^T = \mathbf{L}$ и $\mathbf{O}_R \mathbf{O}_R^T = \mathbf{R}$, где $\mathbf{O}_L$ и $\mathbf{O}_R$ являются $(p_1 + 1)$ - и $(p_2 + 1)$ -диагональными нижнетреугольными матрицами соответственно. Далее мы делаем подстановку в теорему с $\mathbf{A} = \mathbf{A}_j = \mathcal{T}_L(\mathbf{X}_j)$, $\mathbf{B} = \Pi_{\mathcal{M}_r} \mathcal{T}_L(\mathbf{X}_j)$, $\tilde{\mathbf{A}} = \tilde{\mathbf{A}} = \mathbf{O}_L^T \mathbf{A}_j \mathbf{O}_R$ при этом мы получаем желаемое представление $\Pi_{\mathcal{M}_r} \mathcal{T}_L(\mathbf{X}_j) = \sum_{i=1}^r \hat{\mathbf{X}}_i$ с $\hat{\mathbf{X}}_i = \sigma_i U_i V_i^T$, т.е. в виде суммы матриц ранга 1.

Для быстрого вычисления первых r сингулярных чисел и векторов матрицы $\tilde{\mathbf{A}}_j = \mathbf{O}_L^T \mathcal{T}_L(\mathbf{X}_j) \mathbf{O}_R$ можно использовать методы сингулярного разложения с использованием bidiagonalизации Ланцоша, подробности см. в [56, 5]. Чтобы использовать данный подход, мы должны реализовать быстрое умножение матрицы $\tilde{\mathbf{A}}_j$ на произвольный вектор $C \in \mathbb{R}^K$ справа или $D \in \mathbb{R}^L$ слева. Ключевая идея предложена в тех же работах [56, 5] и состоит в том, что мы не вычисляем $\tilde{\mathbf{A}}_j$ явно, а проводим умножение на вектор шаг за шагом, где шаг умножения

на $\mathcal{T}_L(\mathbf{X}_j)$ происходит с помощью быстрого преобразования Фурье. Приведём алгоритм.

Алгоритм 4.5.1 (Вычисление $\mathbf{O}_L^T \mathcal{T}_L(\mathbf{X}_j) \mathbf{O}_R C$). *Вход:* временной ряд $\mathbf{X}_j$, матрицы $\mathbf{O}_L$, $\mathbf{O}_R$, вектор C .

- 1: Вычислить вектор $\hat{C} = \mathbf{O}_R C$.
- 2: Вычислить вектор $B = \mathcal{T}_L(\mathbf{X}_j) \hat{C}$, используя алгоритм [5, Algorithm 1].
- 3: **return** Вектор $\mathbf{O}_L^T B$.

Мы применяем тот же алгоритм и к вычислению $\tilde{\mathbf{A}}_j^T D$, так как $\tilde{\mathbf{A}}_j^T D = \mathbf{O}_R^T \mathcal{T}_K(\mathbf{X}_j) \mathbf{O}_L D$.

4.5.3. Вычисление проектора $\Pi_{\mathcal{H}}$ от суммы матриц ранга 1

Теперь перейдём ко второму шагу итерации метода Кэдзоу (4.1) — вычислению $\Pi_{\mathcal{H}}$.

Предложение 4.5.1. Для произвольной матрицы $\mathbf{X} \in \mathbb{R}^{L \times K}$ выполнено следующее: $\Pi_{\mathcal{H}} \mathbf{X} = \sum_{i=1}^N c_i \mathcal{T}_L(E_i)$, где $C = (c_1, \dots, c_N)^T$, $C = \mathbf{W}^{-1} F$, $\mathbf{W} = \mathbf{L} * \mathbf{R}$, $F = (\langle \mathbf{X}, \mathcal{T}_L(E_1) \rangle_{\mathbf{L}, \mathbf{R}}, \dots, \langle \mathbf{X}, \mathcal{T}_L(E_N) \rangle_{\mathbf{L}, \mathbf{R}})^T$.

Доказательство. Рассмотрим постолбчатую векторизацию $\mathcal{V} : \mathbb{R}^{L \times K} \rightarrow \mathbb{R}^{LK}$, переводящую матрицу в вектор. В пространстве $\mathbb{R}^{LK}$ введём взвешенное скалярное произведение с положительно определённой матрицей $\mathbf{M} \in \mathbb{R}^{LK \times LK}$ такой, что $\langle X, Y \rangle_{\mathbf{M}} = \langle \mathcal{V}^{-1} X, \mathcal{V}^{-1} Y \rangle_{\mathbf{L}, \mathbf{R}}$. Таким образом, $\mathcal{V}$ — изоморфизм.

С помощью $\mathcal{V}$ оператор $\Pi_{\mathcal{H}}$ сводится к вычислению взвешенного МНК. Мы можем записать вектор $C = ([\mathcal{V} \mathcal{T}_L(E_1) : \dots : \mathcal{V} \mathcal{T}_L(E_N)])_{\mathbf{M}}^{\dagger} (\mathcal{V} \mathbf{X})$. Доказательство завершается раскрытием выражения по формуле (1.5), заменой скалярного произведения в $\mathbb{R}^{LK}$ на $\mathbb{R}^{L \times K}$ и применением теоремы 4.2.1. $\square$

Применим предложение 4.5.1 для матрицы $\mathbf{X} = \sum_{i=1}^r \hat{\mathbf{X}}_i$ — результата

применения проектора $\Pi_{\mathcal{M}_r}$, где матрица $\widehat{\mathbf{X}}_i$ имеет ранг 1. Получим:

$$\mathbf{X}_{j+1} = \mathcal{T}_L^{-1} \left(\Pi_{\mathcal{H}} \left(\sum_{i=1}^r \widehat{\mathbf{X}}_i \right) \right) = \mathbf{W}^{-1} \sum_{i=1}^r F_i, \quad (4.24)$$

где $\mathbf{W} = \mathbf{L} * \mathbf{R}$, см. следствие 4.2.1, $F_i = (\langle \widehat{\mathbf{X}}_i, \mathcal{T}_L(E_1) \rangle_{\mathbf{L}, \mathbf{R}}, \dots, \langle \widehat{\mathbf{X}}_i, \mathcal{T}_L(E_N) \rangle_{\mathbf{L}, \mathbf{R}})^T$.

Перепишем j -й элемент $F_i = (f_1^{(i)}, \dots, f_N^{(i)})^T$ в следующем виде:

$$\begin{aligned} f_j^{(i)} &= \langle \widehat{\mathbf{X}}_i, \mathcal{T}_L(E_j) \rangle_{\mathbf{L}, \mathbf{R}} = \text{tr}(\mathbf{L} \widehat{\mathbf{X}}_i \mathbf{R} \mathcal{T}_L^T(E_j)) = \\ &= \sigma_i \text{tr}(\mathbf{O}_{\mathbf{L}} U_i' V_i'^T \mathbf{O}_{\mathbf{R}}^T \mathcal{T}_L^T(E_j)) = \sigma_i \langle (\mathbf{O}_{\mathbf{L}} U_i') (\mathbf{O}_{\mathbf{R}} V_i')^T, \mathcal{T}_L(E_j) \rangle_{\mathbf{F}}. \end{aligned} \quad (4.25)$$

Таким образом, мы свели вычисление $\Pi_{\mathcal{H}}$ к задаче быстрого вычисления вектора $(\langle CD^T, \mathcal{T}_L(E_1) \rangle_{\mathbf{F}}, \dots, \langle CD^T, \mathcal{T}_L(E_N) \rangle_{\mathbf{F}})^T$ для произвольных $C \in \mathbb{R}^K$, $D \in \mathbb{R}^L$, решение которой указано в [5, Algorithm 2] (последний шаг 6 нам не нужен). Так как $\mathbf{W}$ — $(2p_1 + 2p_2 + 1)$ -диагональная, нам надо вычислить $(p_1 + p_2 + 1)$ -диагональное разложение Холецкого $\mathbf{O}_{\mathbf{W}}$ матрицы $\mathbf{W} = \mathbf{O}_{\mathbf{W}} \mathbf{O}_{\mathbf{W}}^T$ и использовать его для быстрого вычисления формулы (4.24). Заметим, что явно векторы U_i и V_i , как и матрицу $\widehat{\mathbf{X}}_i = \sigma_i U_i V_i^T$, вычислять не нужно, потому что они не используются в правой части формулы (4.25).

4.5.4. Алгоритм Кэдзоу

Ниже приведем схему быстро реализации алгоритма Кэдзоу.

Алгоритм 4.5.2 (Быстрый алгоритм Кэдзоу). ***Вход:** временной ряд $\mathbf{X}$, длина окна L , ленточные матрицы $\mathbf{L}$, $\mathbf{R}$, критерий остановки $STOP$.*

- 1: Вычислить разложения Холецкого матриц $\mathbf{L} = \mathbf{O}_{\mathbf{L}} \mathbf{O}_{\mathbf{L}}^T$ и $\mathbf{R} = \mathbf{O}_{\mathbf{R}} \mathbf{O}_{\mathbf{R}}^T$.
- 2: Вычислить ленточную матрицу $\mathbf{W} = \mathbf{L} * \mathbf{R}$ по формуле (4.5), её разложение Холецкого $\mathbf{W} = \mathbf{O}_{\mathbf{W}} \mathbf{O}_{\mathbf{W}}^T$.
- 3: Положить $j = 0$.
- 4: Положить $\mathbf{X}_0 = \mathbf{X}$.

5: **while** не **STOP do**

6: Вычислить $\sigma_i, U'_i, V'_i, i = 1, \dots, r$ — сингулярные числа и векторы матрицы $\mathbf{O}_L^T \mathcal{T}_L(\mathbf{X}_j) \mathbf{O}_R$, используя быстрый алгоритм сингулярного разложения (например, с использованием bidiagonalизации Ланцоша [56, 5]) с алгоритмом умножения вектора 4.5.1.

7: Вычислить векторы $F_i, i = 1, \dots, r$, из формулы (4.25) с помощью [5, Algorithm 2].

8: Вычислить $\mathbf{X}_{j+1} = (\mathbf{O}_W)^{-1}(\mathbf{O}_W^T)^{-1}(\sum_{i=1}^r F_i)$.

9: Положить $j = j + 1$.

10: **end while**

11: **return** $\mathbf{Y} = \mathbf{X}_j$.

4.5.5. Трудоёмкость быстрой реализации

Ниже мы рассмотрим трудоёмкость отдельных шагов алгоритма 4.5.2.

Матрицы $\mathbf{O}_L$ и $\mathbf{O}_R$ вычисляются за время $O(Lp_1^2)$ и $O(Kp_2^2)$ с использованием $O(Lp_1)$ и $O(Kp_2)$ места в памяти, $\mathbf{W}$ вычисляется за время $O(p_1p_2N \log N)$ и с использованием $O((p_1 + p_2)N)$ памяти с помощью предложения 4.2.1 и быстрого преобразования Фурье для вычисления свёрток, плюс требуется $O((p_1^2 + p_2^2)N)$ времени и $O((p_1 + p_2)N)$ места для вычисления $\mathbf{O}_W$. Одна итерация алгоритма SVD Ланцоша матрицы $\tilde{\mathbf{A}}_j$ требует $O(r(N \log N + p_1L + p_2K))$ времени с $O(rN)$ дополнительной памяти [56], а вычисление $\Pi_{\mathcal{H}}$ требует $O(r(N \log N + Lp_1 + Kp_2) + (p_1 + p_2)N)$ времени и $O(rN)$ памяти.

4.6. Применение метода Кэдзоу к решению задачи аппроксимации временных рядов рядами конечного ранга

4.6.1. Выбор метода для поиска весов

Обсудим применение алгоритма Кэдзоу (4.1) для решения задачи (3) вместе с методами поиска весов $\mathbf{L}$, $\mathbf{R}$ по заданной обратной автоковариационной матрице $\Sigma^{-1} = \mathbf{W}_0$, где $\mathbf{W}$ из (3) равна $\mathbf{W}_0$ в этом разделе.

Мы введём две комбинации методов: назовём *Cadzow(quadratic)* — метод Кэдзоу с вычислением весов с помощью алгоритма квадратичного программирования 4.3.1, и *Cadzow(box)* — метод Кэдзоу с вычислением весов с помощью минимизации функции (4.22) методом оптимизации в параллелепипеде, например, L-BFGS-B [55]. Проблема в том, что теоретические результаты и алгоритмы получены только для частных случаев задачи (4.7), поэтому рекомендации их применения в более общих случаях требуют объяснения.

Рассмотрим следующие сценарии применения:

1. $\mathbf{W}_0 = w_0 \mathbf{I}_N$, $w_0 > 0$ — случай белого шума

В этом случае решение задачи (4.7) (точнее, её частного случая (4.8)) $\mathbf{R}^*$ может быть вычислено точно с использованием алгоритма 4.3.1, либо вычислено приближённо с помощью минимизации функции (4.22). Соответственно, стоит использовать метод *Cadzow(quadratic)*, при этом $\mathbf{L} = \mathbf{I}_L$. Метод *Cadzow(box)* можно рассматривать как более простую в реализации альтернативу при наличии готовой реализации метода оптимизации в параллелепипеде, однако, оптимальность $\mathbf{R}^*$ не гарантирована.

2. $\mathbf{W}_0 = \text{diag}((w_1, \dots, w_N)^T)$ — нестационарный белый шум

В этом случае разумно применять метод *Cadzow(box)*, который создан как

раз для этого случая, $\mathbf{L} = \mathbf{I}_L$. Метод *Cadzow(quadratic)* работает только с равными весами.

3. $\mathbf{W}_0 = \Sigma_N^{-1} - (2p + 1)$ -диагональная обратная автоковариационная матрица процесса $\text{AR}(p)$

Для решения задачи (4.7) нет простого алгоритма, кроме случая, когда L достаточно большое, см. теорему 4.2.2, и задача становится эквивалентной (4.8). В качестве эвристики можно рассмотреть решение задачи (4.8), и взять $\mathbf{L} = \Sigma_L^{-1} - (2p + 1)$ -диагональная обратная автоковариационная матрица процесса $\text{AR}(p)$ с теми же коэффициентами, но меньшей длины L . В остальном данный пункт аналогичен первому.

4. Рассмотрим комбинацию предыдущих двух пунктов:

$$\mathbf{W}_0 = \mathbf{D}^{-1} \Sigma_N^{-1} \mathbf{D}^{-1}, \quad (4.26)$$

где $\Sigma_N^{-1} - (2p + 1)$ -диагональная обратная автоковариационная матрица процесса $\text{AR}(p)$, $\mathbf{D}$ — диагональная положительно определенная матрица. Шум с такой ковариационной матрицей можно называть *нестационарным красным шумом*, где $\text{diag}(\mathbf{D})$ можно назвать *огibaющей шума*.

Опять же, для решения задачи (4.7) нет простого алгоритма. В этом случае в качестве эвристики можно использовать метод *Cadzow(box)*, веса которого заданы матрицей $\mathbf{D}^{-2}$, $\mathbf{L} = \mathbf{L} = \Sigma_L^{-1}$.

В разделе 5.2.1 проведены численные эксперименты с алгоритмами *Cadzow(quadratic)* и *Cadzow(box)*.

4.6.2. Использование метода Кэдзоу для поиска начального значения для методов локального поиска

Дополнительная мотивация для изучения метода Кэдзоу следующая. Рассмотренные в главе 3 методы оптимизации сводились к итерации по вектору A из коэффициентов управляющей рядом ОЛРФ(A). Задача (3) может иметь много локальных минимумов [15], из-за чего при выборе неверного начального вектора A_0 алгоритм может найти лишь локальный, а не глобальный минимум. С другой стороны, метод Кэдзоу является непараметрическим, что позволяет свободно использовать его сразу же, без задания начального параметра.

Пусть мы хотим решить задачу (3) для заданного ряда X . Сделаем следующее с помощью алгоритма Кэдзоу: зафиксируем длину окна L , вычислим $X = \mathcal{T}_L(X)$, после чего применим итерацию (4.1) заданное число раз, обозначим его как n . Полученную матрицу переведем в ряд $X_n = \mathcal{T}_L^{-1}(X_n)$, где $X_n = (\Pi_{\mathcal{H}}\Pi_{\mathcal{M}_r})^n X$.

Результатом алгоритма Кэдзоу, вообще говоря, не является ряд $X_n \in \overline{\mathcal{D}}_r$. Согласно теореме 4.1.1 и предложению 4.1.2, у последовательности $X_1, X_2, \dots$ лишь существует сходящаяся к $\overline{\mathcal{D}}_r$ подпоследовательность.

Для того чтобы использовать методы локального поиска, необходимо найти начальную ОЛРФ(A_0). Нам понадобится следующее утверждение.

Предложение 4.6.1. *Пусть временной ряд $S \in \overline{\mathcal{D}}_r$, $\mathcal{T}_{r+1}(S)$ — траекторная матрица данного ряда, а $U \in \mathbb{R}^{r+1}$ — её $(r+1)$ -й левый сингулярный вектор. Тогда временной ряд S управляется ОЛРФ(U).*

Доказательство. Матрица $\mathcal{T}_{r+1}(S)$ неполного ранга, следовательно, её $(r+1)$ -й левый сингулярный вектор соответствует нулевому сингулярному числу. Это означает, что $U^T \mathcal{T}_{r+1}(S) = 0_{N-r}^T$, что совпадает с определением ряда, управляемого ОЛРФ(U). □

Способ выбора вектора A_0 в алгоритмах 3.3.4 и 3.3.5 состоит в следующем: для ряда $\mathbf{X}_n$, полученного методом Кэдзоу, есть основания предполагать его близость к $\overline{\mathcal{D}}_r$. Поэтому следует посчитать траекторную матрицу $\mathcal{T}_{r+1}(\mathbf{X}_n)$, вычислить у полученной матрицы $(r+1)$ -й левый сингулярный вектор U с помощью сингулярного разложения, и положить $A_0 = U$.

4.7. Применение алгоритма Кэдзоу к задаче оценивания сигнала

4.7.1. Линеаризованный алгоритм Кэдзоу

Предположим, что мы хотим применить алгоритм Кэдзоу (4.1) к задаче оценки сигнала ранга ровно r . Допустим, мы наблюдаем случайный ряд $\mathbf{X} = \mathbf{S}_0 + \xi$, где сигнал $\mathbf{S}_0$ такой, что $\text{rank } \mathcal{T}_L(\mathbf{S}_0) = r$ (т.е. $\mathbf{S}_0 \in \mathcal{D}_r$), ξ — случайный вектор с распределением Ξ . Нас интересует приближенное распределение оценки сигнала $\hat{\mathbf{S}}_n = \mathbf{X}_n$, полученной алгоритмом Кэдзоу

$$\hat{\mathbf{S}}_n = \mathcal{T}_L^{-1}(\Pi_{\mathcal{H}}\Pi_{\mathcal{M}_r})^n \mathcal{T}_L(\mathbf{X}) = (\mathcal{T}_L^{-1}\Pi_{\mathcal{H}}\Pi_{\mathcal{M}_r}\mathcal{T}_L)^n(\mathbf{X}) \quad (4.27)$$

при большом числе итераций n и малом шуме ξ , имеющем распределение Ξ ; длина окна L и длина ряда N удовлетворяют условию $\min(L, N - L + 1) > r$.

Подход следующий: в алгоритме Кэдзоу участвуют два оператора — $\Pi_{\mathcal{H}}$ и $\Pi_{\mathcal{M}_r}$, первый из которых является линейным, а второй нет. Заменяем нелинейный оператор на его линейное приближение, и изучим оценку получившегося линеаризованного алгоритма.

Известно [38], что в окрестности точки $\mathbf{S}_0 = \mathcal{T}_L(\mathbf{S}_0)$ множество $\mathcal{M}_r$ представляет из себя гладкое многообразие размерности $r(L+K) - r^2$, а касательное подпространство $\mathcal{N} = \mathcal{N}(\mathbf{S}_0)$ в точке $\mathbf{S}_0$ выражается следующим образом:

$$\mathcal{N}(\mathbf{S}_0) = \{\mathbf{X} \in \mathbb{R}^{L \times K} : \mathbf{U}^T \mathbf{X} \mathbf{V} = \mathbf{0}_{L-r, K-r}\}, \quad (4.28)$$

где $\mathbf{U} \in \mathbb{R}^{L \times (L-r)}$ и $\mathbf{V} \in \mathbb{R}^{K \times (K-r)}$ — матрицы полного ранга такие, что $\mathbf{U}^T \mathbf{S}_0 = \mathbf{0}_{L-r, K}$ и $\mathbf{S}_0 \mathbf{V} = \mathbf{0}_{L, K-r}$. Заметим, что определение корректно, так как $\mathcal{N}(\mathbf{S}_0)$ зависит только от $\text{colspace } \mathbf{S}_0$ и $\text{rowspace } \mathbf{S}_0$.

В дальнейшем, для удобства изложения, мы применим постолбчатую векторизацию $\mathcal{V} : \mathbb{R}^{L \times K} \rightarrow \mathbb{R}^{LK}$, переводящую матрицу в вектор. Пространство $\mathbb{R}^{LK}$ мы снабдим порождённой векторизацией $\mathcal{V}$ скалярным произведением $\langle X, Y \rangle = \langle \mathcal{V}^{-1} X, \mathcal{V}^{-1} Y \rangle_{\mathbf{L}, \mathbf{R}}$; таким образом, $\mathcal{V}$ — изоморфизм. Соответствующие множества обозначим так: $\hat{\mathcal{N}}(S) = \mathcal{V}(\mathcal{N}(\mathcal{V}^{-1}(S)))$, $\hat{\mathcal{H}} = \mathcal{V}(\mathcal{H})$, $\hat{\mathcal{M}}_r = \mathcal{V}(\mathcal{M}_r)$.

Замечание 4.7.1. *Существует малая окрестность точки $S_0 = \mathcal{V}(\mathbf{S}_0) = (\mathcal{V}\mathcal{T}_L)(\mathbf{S}_0)$, в которой проекцию на $\hat{\mathcal{M}}_r$ по порождённой скалярным произведением норме можно записать как $\Pi_{\hat{\mathcal{M}}_r}(S_0 + Z) = S_0 + \mathbf{\Pi}_{\hat{\mathcal{N}}(S_0)} Z + \mathbf{o}_{Z \rightarrow 0}(\|Z\|)$, где $\Pi_{\hat{\mathcal{M}}_r}(S_0 + Z)$ однозначно определён и бесконечно дифференцируем, $\mathbf{\Pi}_{\hat{\mathcal{N}}(S_0)}$ — проектор на линейное подпространство $\hat{\mathcal{N}}(S_0)$ по введённой норме. Это частный случай [38, Лемма 2.1].*

Так как, согласно замечанию 4.7.1, проектор $\Pi_{\hat{\mathcal{M}}_r}$ можно приблизить линейным оператором $\mathbf{\Pi}_{\hat{\mathcal{N}}(S_0)}$ в окрестности точки S_0 , можно рассмотреть *линеаризованный метод Кэдзоу* в $\mathbb{R}^{LK}$:

$$X_0 = X, \quad X_{k+1} = \mathbf{\Pi}_{\hat{\mathcal{H}}} \mathbf{\Pi}_{\hat{\mathcal{N}}(S_0)} X_k, \quad k \geq 0, \quad (4.29)$$

где $\mathbf{\Pi}_{\hat{\mathcal{H}}}$ и $\mathbf{\Pi}_{\hat{\mathcal{N}}(S_0)}$ — матрицы проектирования размера $(LK) \times (LK)$ на соответствующие линейные подпространства по введённой норме.

Так как $\hat{\mathcal{H}}$ и $\hat{\mathcal{N}}(S_0)$ — линейные подпространства, то свойства линеаризованного метода Кэдзоу гораздо проще исследовать по сравнению с исходным методом. Для начала, установим вид пересечения касательного подпространства $\mathcal{N}(\mathbf{S}_0)$ с множеством ганкелевых матриц.

Лемма 4.7.1. *Пусть $\mathbf{S}_0 \in \mathcal{D}_r$ управляется ОЛРФ(A_0), $A_0 \in \mathbb{R}^{r+1}$. Тогда $\hat{\mathcal{H}} \cap \hat{\mathcal{N}}((\mathcal{V}\mathcal{T}_L)(\mathbf{S}_0)) = (\mathcal{V}\mathcal{T}_L)(\mathcal{Z}(A_0^2))$.*

Доказательство. Заметим, что в определении (4.28) можно взять $\mathbf{U} = \mathbf{Q}^{L,r}(A_0)$, $\mathbf{V} = \mathbf{Q}^{K,r}(A_0)$, так как $\text{colspace}(\mathbf{S}_0) = \mathcal{Z}_L(A_0)$, $\text{colspace}(\mathbf{S}_0^T) = \mathcal{Z}_K(A_0)$, см. (1.3).

Для простоты изложения приведем доказательство на языке матриц, не используя отображение $\mathcal{V}$. Рассмотрим матрицу $\mathbf{X}$, лежащую в пересечении $\mathcal{H}$ и $\mathcal{N}(\mathbf{S}_0)$. Последнее означает, что существует $\mathbf{X} = (x_1, \dots, x_N)^T$: $\mathbf{X} = \mathcal{T}_L(\mathbf{X})$. Возьмём i -й столбец матрицы $\mathbf{Q}^{L,r}(A_0)$, j -й столбец матрицы $\mathbf{Q}^{K,r}(A_0)$, и вычислим (i, j) -й элемент матрицы $\mathbf{U}^T \mathbf{X} \mathbf{V}$ из (4.28). Получим:

$$\begin{aligned} \sum_{l=1}^{r+1} \sum_{p=1}^{r+1} a_l a_p x_{i+l-1+p+j-1-1} &= \sum_{l=1}^{r+1} \sum_{p=1}^{r+1} a_l a_p x_{(i+j-1)+p+l-2} = \\ &= \sum_{k=1}^{2r+1} a_k^{(2)} x_{(i+j-1)+k-1} = 0, \end{aligned}$$

где $a^{(2)}$ — коэффициенты ОЛРФ(A_0^2), см. определение (2.3). Перебор i от 1 до $L - r$ и j от 1 до $K - r$ покрывает весь промежуток $i + j - 1$ от 1 до $N - 2r$, что эквивалентно $\mathbf{Q}^T(A_0^2)\mathbf{X} = 0_{N-2r}$. Таким образом, согласно (1.3), все такие $\mathbf{X}$ составляют $\mathcal{Z}(A_0^2)$. $\square$

Замечание 4.7.2. Лемма 4.7.1 показывает, что пересечение касательного к $\mathcal{M}_r$ подпространства $\mathcal{N}(\mathbf{S}_0)$ с подпространством ганкелевых матриц $\mathcal{H}$ с точностью до изоморфизма совпадает с касательным пространством к $\mathcal{D}_r$ в точке $\mathbf{S}_0$, равным $\mathcal{Z}(A_0^2)$, см. теорему 2.2.3.

Следующая лемма устанавливает связь между проекцией на пересечение $\hat{\mathcal{H}} \cap \hat{\mathcal{N}}$ и пределом итераций линеаризованного метода Кэдзоу.

Предложение 4.7.1. Выполняется следующая сходимость матриц:

$$(\Pi_{\hat{\mathcal{H}}} \Pi_{\hat{\mathcal{N}}(S_0)})^n \rightarrow_{n \rightarrow +\infty} \Pi_{\hat{\mathcal{H}} \cap \hat{\mathcal{N}}(S_0)}.$$

Доказательство. Данное утверждение — частный случай [57, Theorem 13.7] для матриц $\Pi_{\hat{\mathcal{H}}}$ и $\Pi_{\hat{\mathcal{N}}(S_0)}$. $\square$

Рассмотрим ряд $\mathbf{X}$ и линеаризованный алгоритм Кэдзоу (4.29), который путём применения оператора $\mathcal{V}^{-1}$ к обеим частям принимает следующий эквивалентный вид:

$$\mathbf{X}_0 = \mathbf{X}, \quad \mathbf{X}_{k+1} = \Pi_{\mathcal{H}} \Pi_{\mathcal{N}(\mathbf{S}_0)} \mathbf{X}_k, \quad k \geq 0.$$

Применяя оператор $\mathcal{T}_L^{-1}$, получаем ещё один эквивалентный вид на языке рядов: $\mathbf{X}_n = (\mathcal{T}_L^{-1} \Pi_{\mathcal{H}} \Pi_{\mathcal{N}(\mathcal{T}_L(\mathbf{S}_0))} \mathcal{T}_L)^n \mathbf{X}$, или, иначе, $\mathbf{X}_n = \mathbf{P}_{\mathbf{S}_0}^n \mathbf{X}$, где

$$\mathbf{P}_{\mathbf{S}_0} = \mathcal{T}_L^{-1} \Pi_{\mathcal{H}} \Pi_{\mathcal{N}(\mathcal{T}_L(\mathbf{S}_0))} \mathcal{T}_L. \quad (4.30)$$

Оператор $\mathbf{P}_{\mathbf{S}_0} : \mathbb{R}^N \rightarrow \mathbb{R}^N$ может быть задан матрицей размера $N \times N$, так как все участвующие внутри (4.30) отображения линейны.

Следствие 4.7.1. Пусть $\mathbf{S}_0 \in \mathcal{D}_r$ управляется ОЛРФ(A_0), $A_0 \in \mathbb{R}^{r+1}$. Тогда $\mathbf{P}_{\mathbf{S}_0}^n \xrightarrow{n \rightarrow +\infty} \Pi_{\mathcal{Z}(A_0^2), \mathbf{W}(\mathbf{L}, \mathbf{R})}$, где $\mathbf{W}(\mathbf{L}, \mathbf{R}) = \mathbf{L} * \mathbf{R}$.

Доказательство. Доказательство проводится прямой подстановкой оператора $\mathcal{T}_L^{-1} \mathcal{V}^{-1}$ в результат леммы 4.7.1 и предложения 4.7.1 с учётом введённого в $\mathbb{R}^{LK}$ скалярного произведения $\langle \mathcal{V} \mathcal{T}_L(\mathbf{Y}), \mathcal{V} \mathcal{T}_L(\mathbf{Z}) \rangle = \langle \mathcal{T}_L(\mathbf{Y}), \mathcal{T}_L(\mathbf{Z}) \rangle_{\mathbf{L}, \mathbf{R}}$ и теоремы 4.2.1. $\square$

Предложение 4.7.2. Выполняется следующее:

$$\|(\mathbf{P}_{\mathbf{S}_0}^n - \Pi_{\mathcal{Z}(A_0^2), \mathbf{W}(\mathbf{L}, \mathbf{R})}) \mathbf{X}\|_{\mathbf{W}} \leq C \|\mathbf{X}\|_{\mathbf{W}} (|\lambda_{2r+1}| + \delta)^n, \quad (4.31)$$

где $C > 0$ — некоторая положительная константа, $\delta > 0$ — сколь угодно малое вещественное число, λ_{2r+1} — $(2r+1)$ -е по абсолютной величине собственное число матрицы $\mathbf{P}_{\mathbf{S}_0}$, при этом $|\lambda_{2r+1}| < 1$.

Доказательство. Из теоремы [58, стр. 630] с учётом следствия 4.7.1 следует, что модуль всех собственных чисел $\mathbf{P}_{\mathbf{S}_0}$ не превосходит единицы, при этом только единица является собственным числом с модулем, равным единице. Покажем, что матрица $\mathbf{P}_{\mathbf{S}_0}$ содержит только $2r$ собственных чисел, равных единице.

Мы знаем, что $\mathbf{P}_{S_0}Y = Y$ для любого $Y \in \mathcal{Z}(A_0^2)$. Предположим, что существует $Z \notin \mathcal{Z}(A_0^2)$, $\mathbf{P}_{S_0}Z = Z$. Тогда $\mathbf{P}_{S_0}^k Z \rightarrow_{k \rightarrow +\infty} Z$, но $Z \neq \mathbf{\Pi}_{\mathcal{Z}(A_0^2), \mathbf{W}(\mathbf{L}, \mathbf{R})} Z$. Получили противоречие со следствием 4.7.1. Таким образом, $|\lambda_{2r+1}| < 1$.

С учётом полученного факта о собственном числе λ_{2r+1} , утверждение предложения — следствие из [59, Theorem 2.12]. $\square$

Замечание 4.7.3 (Скорость сходимости). *Применим результат предложения 4.7.2 к определению скорости сходимости.*

Допустим, мы выполняем итерации линеаризованного алгоритма Кэдзоу $\mathbf{P}_{S_0}^k X$ по $k = 1, 2, \dots$, а в качестве критерия останова используем, например, $\|X^{(k)} - X^{(k-1)}\|_{\mathbf{W}} < \varepsilon$ или $\frac{\|X^{(k)} - X^{(k-1)}\|_{\mathbf{W}}}{\|X\|_{\mathbf{W}}} < \varepsilon$, где ε — небольшое положительное число. Путём обращения неравенств с учётом (4.31) с $\delta = 0$ получаем, что и в том, и в другом случае число шагов до останова приближённо равно

$$I \approx -c / \log(|\lambda_{2r+1}|), \quad (4.32)$$

где c — некоторая константа.

В свою очередь, на статистическом языке следствие 4.7.1 означает следующее:

Лемма 4.7.2. *Пусть $S_0 \in \mathcal{D}_r$ и управляется ОЛРФ(A_0), ξ — случайный вектор с распределением Ξ . Рассмотрим оценку сигнала $\tilde{S}_n = \mathbf{P}_{S_0}^n(S_0 + \xi)$, полученную линеаризованным алгоритмом Кэдзоу в модели $X = S_0 + \xi$, $\mathbf{P}_{S_0}^n$ определена в (4.30). Тогда имеет место следующая слабая сходимость распределений:*

$$\tilde{S}_n \Rightarrow_{n \rightarrow +\infty} S_0 + \mathbf{\Pi}_{\mathcal{Z}(A_0^2), \mathbf{W}(\mathbf{L}, \mathbf{R})} \xi. \quad (4.33)$$

Доказательство. Достаточно показать, что для любой липшицевой с параметром l ($|f(X) - f(Y)| \leq l\|X - Y\|_{\mathbf{W}}$ для любых X, Y) и ограниченной m ($|f(X)| \leq m$ для любого X) измеримой функции f , заданной на $\mathbb{R}^N$, интегралы

$$I_1(f, t) = \mathbb{E}(|f(\mathbf{P}_{S_0}^n(S_0 + \xi)) - f(S_0 + \mathbf{\Pi}_{\mathcal{Z}(A_0^2), \mathbf{W}(\mathbf{L}, \mathbf{R})} \xi)|) \quad (4.34)$$

равномерно по f ограничены последовательностью, стремящейся к 0 при $t \rightarrow 0$ [41, §7].

Зафиксируем $\varepsilon > 0$, и выберем столь большой шар радиуса r_B , что

$$\mathbb{P}(\|\xi\|_{\mathbf{W}} \geq r_B) < \varepsilon.$$

Согласно следствию 4.7.1, выберем такое большое n , что

$$\|(\mathbf{P}_{\mathbf{S}_0}^n - \mathbf{\Pi}_{\mathcal{Z}(A_0^2), \mathbf{W}(\mathbf{L}, \mathbf{R})})\mathbf{Z}\|_{\mathbf{W}} \leq \frac{\varepsilon}{r_B} \|\mathbf{Z}\|_{\mathbf{W}}$$

для любого $\mathbf{Z} \in \mathbb{R}^N$. Тогда:

$$\int_{\|\mathbf{Y}\|_{\mathbf{W}} < r_B} |f(\mathbf{P}_{\mathbf{S}_0}^n(\mathbf{S}_0 + \mathbf{Y})) - f(\mathbf{S}_0 + \mathbf{\Pi}_{\mathcal{Z}(A_0^2), \mathbf{W}(\mathbf{L}, \mathbf{R})}\mathbf{Y})| d\Xi(\mathbf{Y}) \leq l\varepsilon.$$

При этом,

$$\int_{\|\mathbf{Y}\|_{\mathbf{W}} \geq r_B} |f(\mathbf{P}_{\mathbf{S}_0}^n(\mathbf{S}_0 + \mathbf{Y})) - f(\mathbf{S}_0 + \mathbf{\Pi}_{\mathcal{Z}(A_0^2), \mathbf{W}(\mathbf{L}, \mathbf{R})}\mathbf{Y})| d\Xi(\mathbf{Y}) < 2m\varepsilon.$$

Таким образом, в (4.34) получаем $I_1(f, t) < l\varepsilon + 2m\varepsilon$, где правую часть неравенства можно сделать сколь угодно малой выбором ε . $\square$

Замечание 4.7.4. Если Ξ — распределение с нулевым математическим ожиданием и ковариационной матрицей Σ , то распределение случайного вектора $\mathbf{S}_0 + \mathbf{\Pi}_{\mathcal{Z}(A_0^2), \mathbf{W}(\mathbf{L}, \mathbf{R})}\xi$ имеет математическое ожидание $\mathbf{S}_0$ и ковариационную матрицу $\mathbf{\Pi}_{\mathcal{Z}(A_0^2), \mathbf{W}(\mathbf{L}, \mathbf{R})}\Sigma\mathbf{\Pi}_{\mathcal{Z}(A_0^2), \mathbf{W}(\mathbf{L}, \mathbf{R})}^T$.

Лемма 4.7.2 позволяет приближённо оценивать распределение оценки, полученной алгоритмом Кэдзоу при малом шуме с распределением Ξ .

4.7.2. Вид ошибки первого порядка оценки n -й итерации алгоритма Кэдзоу

Предположим, что мы наблюдаем случайный ряд $\mathbf{X} = \mathbf{S} + t\xi$ длины N , где $\mathbf{S} \in \mathcal{D}_r$, ξ имеет распределение Ξ . Вместе с оценкой сигнала линеаризованным алгоритмом Кэдзоу $\tilde{\mathbf{S}}_n(t) = \mathbf{P}_{\mathbf{S}_0}^n(\mathbf{S}_0 + t\xi)$ рассмотрим оценку $\hat{\mathbf{S}}_n(t)$ сигнала,

полученную с помощью n -й итерации алгоритма Кэдзоу, где $\widehat{S}_n$ определена в (4.27). Нас интересует следующий вопрос: есть ли сходимость разложений первого порядка по t оценки $\widetilde{S}_n(t)$ к оценке $\widehat{S}_n(t)$ при $t \rightarrow 0$? Иначе говоря, есть ли сходимость распределений оценки по алгоритму Кэдзоу к оценке по линеаризованному алгоритму Кэдзоу при $t \rightarrow 0$ (то есть при соотношении сигнал/шум, стремящемся к бесконечности)?

Для начала докажем следующее обобщение замечания 4.7.1. Под обозначением $\Pi_{\widehat{\mathcal{H}}}$ мы имеем в виду оператор проектирования в $\mathbb{R}^{LK}$, под $\mathbf{\Pi}_{\widehat{\mathcal{H}}}$ — соответствующую этому оператору матрицу.

Лемма 4.7.3. *Пусть $S_0 \in \mathcal{D}_r$ — ряд длины N , зафиксировано целое $k \geq 1$. Тогда вокруг точки $S_0 = \mathcal{V}(\mathbf{S}_0) = (\mathcal{V}\mathcal{T}_L)(S_0)$ существует открытое множество $\mathcal{U}_k \subset \mathbb{R}^{LK}$ такое, что сужение оператора $(\Pi_{\widehat{\mathcal{H}}}\Pi_{\widehat{\mathcal{M}}_r})^k$ на нём — гладкое (т.е. бесконечно дифференцируемое), и можно для любого сколь угодно малого $\varepsilon > 0$ выбрать шар $\mathcal{B}_r = \{Y : \|Y\| < \delta_k\}$ такого радиуса δ_k , что для любого $Z \in \mathcal{B}_r$:*

$$\|(\Pi_{\widehat{\mathcal{H}}}\Pi_{\widehat{\mathcal{M}}_r})^k(S_0 + Z) - (\mathbf{\Pi}_{\widehat{\mathcal{H}}}\mathbf{\Pi}_{\widehat{\mathcal{N}}(S_0)})^k(S_0 + Z)\| \leq \varepsilon\|Z\|. \quad (4.35)$$

Доказательство. Согласно замечанию 4.7.1, существует такое открытое множество $\mathcal{U}$, содержащее S_0 , что сужение $\widetilde{\Pi}_{\widehat{\mathcal{M}}_r} : \mathcal{U} \rightarrow \widehat{\mathcal{M}}_r$ оператора $\Pi_{\widehat{\mathcal{M}}_r}$ однозначно определено, бесконечно дифференцируемо, при этом в точке S_0 матрица-градиент $\widetilde{\Pi}_{\widehat{\mathcal{M}}_r}$ равна $\mathbf{\Pi}_{\widehat{\mathcal{N}}(S_0)}$.

Оператор $\Pi_{\widehat{\mathcal{H}}}$, очевидно, непрерывный и бесконечно дифференцируемый. Тогда композиция операторов $\Pi_{\widehat{\mathcal{H}}}\widetilde{\Pi}_{\widehat{\mathcal{M}}_r} : \mathcal{U} \rightarrow \mathbb{R}^{LK}$ — непрерывна и бесконечно дифференцируема. Для того чтобы корректно определить сужение оператора $(\Pi_{\widehat{\mathcal{H}}}\Pi_{\widehat{\mathcal{M}}_r})^k$ в открытой окрестности точки S_0 , сделаем следующее: положим $\mathcal{U}_1 = \mathcal{U}$, а для $i \geq 2$: $\mathcal{U}_i \subset \mathcal{U}$, $\mathcal{U}_i$ — следующий полный открытый прообраз множества: $\mathcal{U}_i = (\widetilde{\Pi}_{\widehat{\mathcal{H}}}\widetilde{\Pi}_{\widehat{\mathcal{M}}_r})^{-1}(\mathcal{U}_{i-1})$. Заметим, что все $\mathcal{U}_i$ содержат S_0 , так как $\Pi_{\widehat{\mathcal{H}}}\widetilde{\Pi}_{\widehat{\mathcal{M}}_r}(S_0) = S_0$.

В итоге получаем, что сужение $(\Pi_{\widehat{\mathcal{H}}} \Pi_{\widehat{\mathcal{M}}_r})^k : \mathcal{U}_k \rightarrow \mathbb{R}^{LK}$ — бесконечно дифференцируемый оператор в точке S_0 , при этом, пользуясь правилом вычисления матрицы-градиента композиции операторов, получаем, что градиент $(\Pi_{\widehat{\mathcal{H}}} \Pi_{\widehat{\mathcal{M}}_r})^k$ в точке S_0 равен $(\Pi_{\widehat{\mathcal{H}}} \Pi_{\widehat{\mathcal{N}}(S_0)})^k$; следовательно, в окрестности точки S имеется следующее представление:

$$(\Pi_{\widehat{\mathcal{H}}} \Pi_{\widehat{\mathcal{M}}_r})^k(S_0 + Z) = S_0 + (\Pi_{\widehat{\mathcal{H}}} \Pi_{\widehat{\mathcal{N}}(S_0)})^k(Z) + \mathbf{o}_{Z \rightarrow 0}(\|Z\|),$$

из чего следует (4.35). □

Теперь рассмотрим случайный вектор $\eta = (\mathcal{V}\mathcal{T}_L)(\xi)$ с распределением H . Введём случайную матрицу порядка $LK \times (n+1)$, построенную на основе оценок оценок сигнала, полученных методом Кэдзоу с $0, 1, \dots, n$ итерациями:

$$\widehat{\mathbf{E}}(t, \eta) = [\widehat{S}_0(t) : \dots : \widehat{S}_n(t)],$$

где $\widehat{S}_k(t) = \mathcal{V}\mathcal{T}_L(\widehat{\mathbf{S}}_k(t)) = (\Pi_{\widehat{\mathcal{H}}} \Pi_{\widehat{\mathcal{M}}_r})^k(S_0 + t\eta)$, $k = 0, \dots, n$.

Для корректности будем считать, что если проекция $\Pi_{\widehat{\mathcal{M}}_r}$ неоднозначна, то выбирается наименьшая в лексикографическом порядке точка.

Также мы введём матрицу следующего вида:

$$\widehat{\mathbf{M}}(\eta) = [(\Pi_{\widehat{\mathcal{H}}} \Pi_{\widehat{\mathcal{N}}(S_0)})^0 \eta : \dots : (\Pi_{\widehat{\mathcal{H}}} \Pi_{\widehat{\mathcal{N}}(S_0)})^n \eta],$$

через которую можем записать матрицу оценок $\widetilde{\mathbf{E}}(t, \eta)$, полученных линеаризованным методом Кэдзоу:

$$\widetilde{\mathbf{E}}(t, \eta) = [\widetilde{S}_0(t) : \dots : \widetilde{S}_n(t)] = [S_0 : \dots : S_0] + t\widehat{\mathbf{M}}(\eta),$$

где $\widetilde{S}_k(t) = \mathcal{V}\mathcal{T}_L(\widetilde{\mathbf{S}}_k(t))$, $k = 0, \dots, n$. Рассмотрим следующее представление:

$$\widehat{\mathbf{E}}(t, \eta) = [S_0 : \dots : S_0] + t\widehat{\mathbf{M}}(\eta) + \widehat{\mathbf{K}}(t, \eta), \quad (4.36)$$

где $\widehat{\mathbf{K}}(t, \eta)$ — случайная матрица размера $LK \times (n+1)$.

Справедлива следующая теорема:

Теорема 4.7.1. Пусть $S_0 \in \mathcal{D}_r$ и управляется ОЛРФ(A_0), ξ — случайный вектор. Определим случайный вектор $\eta = (\mathcal{V}\mathcal{T}_L)(\xi)$, имеющий распределение H . Тогда имеет место следующая слабая сходимость распределений:

$$\frac{1}{t}\widehat{\mathbf{K}}(t, \eta) \Rightarrow_{t \rightarrow 0} \mathbf{0}_{LK, n+1},$$

где $\widehat{\mathbf{K}}(t, \eta)$ определено в (4.36).

Доказательство. Пусть произвольная функция $f: \mathbb{R}^{LK \times (r+1)} \rightarrow \mathbb{R}$ измерима, липшицева с параметром l (для любых $\mathbf{X}, \mathbf{Y}$: $|f(\mathbf{X}) - f(\mathbf{Y})| \leq l \max_{0 \leq i \leq n} \|X_i - Y_i\|$, где $\mathbf{X} = [X_0 : \dots : X_n]$, $\mathbf{Y} = [Y_0 : \dots : Y_n]$) и ограничена m (для любой $\mathbf{X}$: $f(\mathbf{X}) \leq m$).

Достаточно показать [41, §7], что для любой f интегралы

$$I_1(t, f) = \mathbb{E}(|f(\widehat{\mathbf{K}}(t, \eta)/t) - f(\mathbf{0}_{LK, n+1})|) = \int_{\mathbb{R}^{LK}} |f(\widehat{\mathbf{K}}(t, Y)) - f(\mathbf{0}_{LK, n+1})| dH(Y) \quad (4.37)$$

равномерно по f ограничены последовательностью, стремящейся к 0 при $t \rightarrow 0$, где

$$\widehat{\mathbf{K}}(t, Y) = \frac{\widehat{\mathbf{E}}(t, Y) - [S_0 : \dots : S_0] - t\widehat{\mathbf{M}}(\eta)}{t}.$$

Зафиксируем $\varepsilon_2 > 0$, и выберем столь большой шар радиуса r_B , что

$$\mathbb{P}(\|\eta\| \geq r_B) = \int_{\|Y\| \geq r_B} 1 dH(Y) < \varepsilon_2.$$

По лемме 4.7.3 можно выбрать такие $\delta_1, \dots, \delta_n$, что выполняется (4.35) для $k = 1, \dots, n$ и $\varepsilon = \varepsilon_2/r_B$. Возьмём $\delta = \min(\delta_1, \dots, \delta_n)$, $t = \frac{\delta}{r_B}$. Тогда:

$$\int_{\|Y\| < r_B} |f(\widehat{\mathbf{K}}(t, Y)) - f(\mathbf{0}_{LK, n+1})| dH(Y) \leq l\varepsilon_2.$$

При этом,

$$\int_{\|Y\| \geq r_B} |f(\widehat{\mathbf{K}}(t, Y)) - f(\mathbf{0}_{LK, n+1})| dH(Y) < 2m\varepsilon_2.$$

Таким образом, в (4.37) получаем $I_1(f) < l\varepsilon_2 + 2m\varepsilon_2$, где правую часть неравенства можно сделать сколь угодно малой выбором ε . $\square$

Применим к результату теоремы 4.7.1 операторы $\mathcal{V}^{-1}$ и $\mathcal{T}_L^{-1}$. Введём следующие случайные матрицы размера $N \times (n+1)$. Пусть

$$\widehat{\mathbf{S}}(t, \xi) = [\widehat{\mathbf{S}}_0(t) : \dots : \widehat{\mathbf{S}}_n(t)]$$

состоит из оценок сигнала методом Кэдзоу определённых в (4.27), для $\mathbf{X} = \mathbf{S}_0 + t\xi$, с помощью $0, 1, \dots, n$ итераций. Затем введём матрицу, состоящую из оценок сигнала с помощью линеаризованного метода Кэдзоу следующего вида:

$$\widetilde{\mathbf{S}}(t, \xi) = [\widetilde{\mathbf{S}}_0(t) : \dots : \widetilde{\mathbf{S}}_n(t)],$$

где $\widetilde{\mathbf{S}}_k(t) = \mathbf{P}_{\mathbf{S}_0}^k(\mathbf{S}_0 + t\xi)$, $\mathbf{P}_{\mathbf{S}_0}^n$ определена в (4.30). Заметим, что выполняется равенство $\widehat{\mathbf{S}}(t, \xi) = [\mathbf{S}_0 : \dots : \mathbf{S}_0] + t\mathbf{M}(\xi)$, где $\mathbf{M}(\xi) = [\mathbf{P}_{(\mathbf{S}_0)}^0\xi : \dots : \mathbf{P}_{(\mathbf{S}_0)}^n\xi]$.

Рассмотрим представление:

$$\widetilde{\mathbf{S}}(t, \xi) = [\mathbf{S}_0 : \dots : \mathbf{S}_0] + t\mathbf{M}(\xi) + \mathbf{K}(t, \xi), \quad (4.38)$$

где $\mathbf{K}(t, \xi)$ — случайная матрица размера $N \times (n+1)$.

Мы получим следующее следствие.

Следствие 4.7.2. Пусть $\mathbf{S}_0 \in \mathcal{D}_r$ — ряд длины N , управляемый ОЛРФ(A_0), ξ — случайный вектор с распределением Ξ .

Тогда имеет место слабая сходимость распределений:

$$\frac{1}{t}\mathbf{K}(t, \xi) = \frac{1}{t}(\widetilde{\mathbf{S}}(t, \xi) - \widehat{\mathbf{S}}(t, \xi)) \Rightarrow_{t \rightarrow 0} \mathbf{0}_{N, n+1},$$

где $\mathbf{K}(t, \xi)$ определено в (4.38).

Следствие 4.7.2 позволяет приближённо оценивать распределение оценок сигнала, полученных с помощью алгоритма Кэдзоу, при малом шуме (то есть при малом t) на $1, 2, \dots, n$ -й итерации при помощи распределения оценки $\mathbf{M}(\xi)$ за счёт сходимости оценок $\widehat{\mathbf{S}}_n(t)$ и $\widetilde{\mathbf{S}}_n(t)$ при малых t .

Результаты численных экспериментов по устойчивости и применимости модифицированного метода Гаусса-Ньютона и метода Кэдзоу

5.1. Исследование устойчивости алгоритмов

5.1.1. Точность вычисления базисов

В разделе 3.2.1 упоминалось, что в алгоритмах 3.2.1 и 3.2.2 есть два этапа, наиболее чувствительных к точности вычислений: получение матриц $\mathbf{A}_g$ ($\mathbf{A}_g^2$) с помощью вычисления полиномов и матриц $\mathbf{U}_r$ и $\hat{\mathbf{U}}_r$ из $\mathbf{L}_r$ и $\hat{\mathbf{L}}_{2r}$ соответственно с помощью их ортогонализации. Теоретический анализ матриц $\mathbf{L}_r$, $\mathbf{L}_{2r}$ и тем более $\hat{\mathbf{L}}_{2r}$ представляется крайне трудным. Поэтому был проведён следующий численный эксперимент, показывающий порядок вырожденности матриц и полученную в результате ошибку при проектировании на пространства $\mathcal{Z}(A)$ и $\mathcal{Z}(A^2)$ при различных N и использовании алгоритма 3.2.1 для вычисления $\mathbf{Z}(A)$ и алгоритма 3.2.2 для вычисления $\mathbf{Z}(A^2)$.

В качестве ОЛРФ(A) была взята ОЛРФ с $A = (1, -3, 3, -1)^T$, соответствующий ей характеристический многочлен имеет вид $g_A(z) = (z - 1)^3$. Таким образом, эта ОЛРФ управляет последовательностями, представляющими из себя полином степени не более 2 согласно теореме 1.1.1, степень сигнального корня $t = 3$ (см. теорему 3.2.1). ОЛРФ(A^2) соответствует характеристический многочлен $g_{A^2}(z) = (z - 1)^6$, поэтому подпространство рядов, управляемых ОЛРФ(A^2), состоит из полиномиальных последовательностей степени не

более 5 согласно теореме 1.1.1. Были взяты N на логарифмической сетке от $N = 30$ до $N = 50000$. Для оценки точности вычисления базиса использовалась норма разницы между рядом, управляемым ОЛРФ, и его проекцией $\Delta_A = \|\mathbf{X}_N - \Pi_{\tilde{\mathbf{Z}}(A), \mathbf{I}_N} \mathbf{X}_N\|$, $\tilde{\mathbf{Z}}(A)$ — базис, вычисленный численным методом на компьютере (процессор *Intel Core i5-5250U*, архитектура *x86_64*). В качестве рядов брались $\mathbf{X}_N = (c_1 x_1^2, \dots, c_1 x_N^2)^T$ для алгоритма 3.2.1 и $\mathbf{X}_N = (c_1 x_1^5, \dots, c_1 x_N^5)^T$ для алгоритма 3.2.2, где x_i образуют равномерную решётку на отрезке $[-1; 1]$; положительные множители c_1 и c_2 выбраны так, чтобы $\|\mathbf{X}_N\| = 1$. Шаг 1 алгоритма 3.2.1 сводится к использованию алгоритма 3.2.5. В свою очередь, алгоритм оптимизации, используемый в шаге 5 алгоритма 3.2.5 — метод золотого сечения на пяти равномерных промежутках в $[-\pi/N; \pi/N]$. Более того, были измерены число обусловленности $\kappa(\mathbf{A}_g)$ матрицы $\mathbf{A}_g$, отношение первого к r -му сингулярному числу для $\mathbf{L}_r$ и $\widehat{\mathbf{L}}_{2r}$, обозначенные $\kappa(\mathbf{L}_r)$ и $\kappa(\widehat{\mathbf{L}}_{2r})$ соответственно. Результаты эксперимента представлены на рисунке 5.1. Для наглядности числа обусловленности домножены на 10^{-18} , чтобы привести к масштабу других графиков.

В легенде на рисунке 5.1 вдобавок к измеренным величинам указана оценка порядка полиномиального роста, построенная с помощью линейной регрессии. Например, порядок 3.00 означает кубический рост. Видим, что порядок Δ_A (3.01), Δ_{A^2} (2.88) и $\kappa(\mathbf{A}_g)$ (3.04) примерно равен $t = 3$, что согласуется с теоремой 3.2.1. Из рисунка 5.1 также видно, что, предположительно, порядок обусловленности $\kappa(\mathbf{L}_r)$ и $\kappa(\widehat{\mathbf{L}}_{2r})$ равен $O(N^{t-1})$. Таким образом, порядок ошибки вычисления проекции для алгоритмов 3.2.1 и 3.2.2 определяется обусловленностью $\mathbf{A}_g$.

Аналогично, ошибки вычисления проекции были сделаны и для алгоритмов 3.2.3 и 3.2.4, результаты представлены на рисунке 5.2. Видно, что при рассмотренных N компенсированная схема Горнера даёт настолько большое улучшение по точности, что порядок полиномиального роста $\kappa(\mathbf{A}_g)$ при указанных

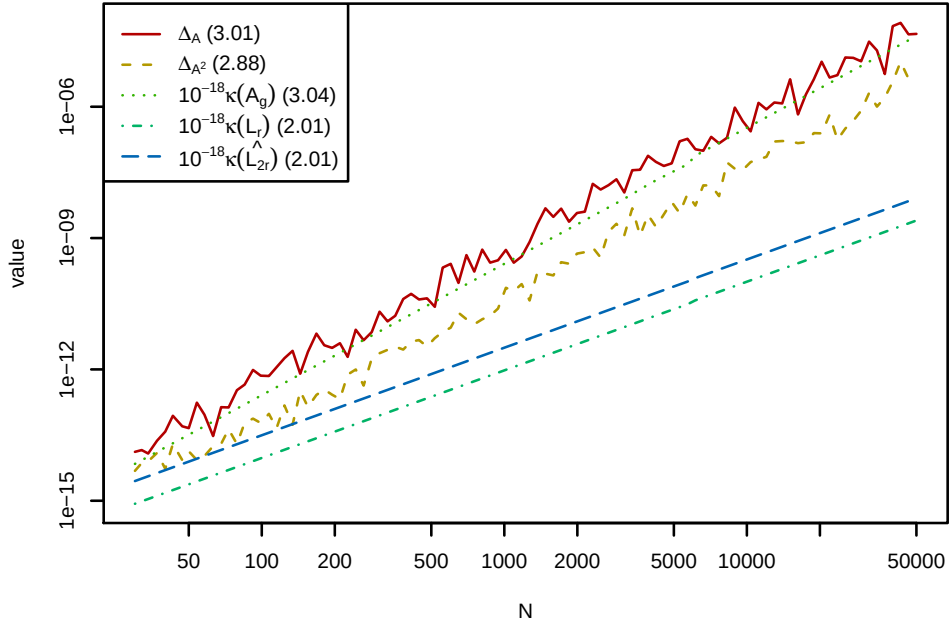


Рис. 5.1. Сравнение ошибок вычисления проекции для различных N и чисел обусловленности.

N был оценён как линейный.

5.1.2. Алгоритмы локального поиска

Для того чтобы сравнить локальные методы решения задачи (3) по точности оценки решения, необходимо построить такую пару из исходного временного ряда X и ряда $Y^* = S_0$, для которой было бы известно, что S_0 является точкой локального минимума задачи (3). Потом необходимо запустить алгоритм локального поиска из окрестности точки локального минимума и сравнить полученный результат с известным минимумом.

В качестве $Y_N^* = (c_1 x_1^2, \dots, c_1 x_N^2)^T$ взята квадратичная функция, управляемая ОЛРФ(A_0), $A_0 = (1, -3, 3, -1)^T$, где x_i — равномерная решётка на отрезке $[-1; 1]$, положительное c_1 взято так, чтобы $\|Y_N^*\| = 1$. Далее, был взят ряд $\hat{N}_N = (c_3 |x_1|, \dots, c_3 |x_N|)^T$ такой, что $\|\hat{N}_N\| = 1$, вычислен $N_N = \hat{N}_N - \Pi_{Z(A_0^2), w} \hat{N}_N$, и взят $X_N = Y_N^* + N_N$. Таким образом, пара $Y_N^* = S_0$ и $X = X_N$ удовлетворяет

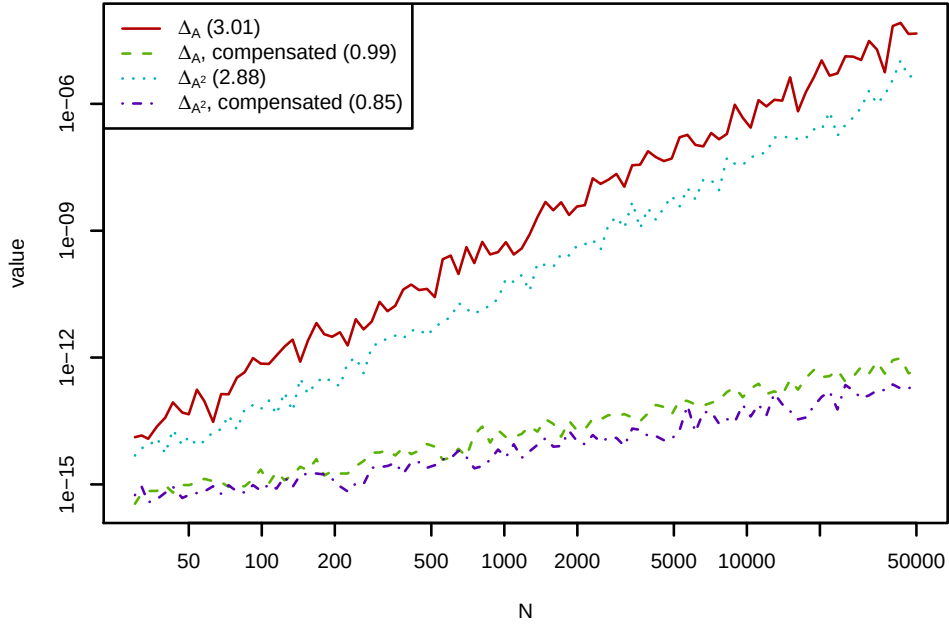


Рис. 5.2. Сравнение ошибки вычисления проекции при различных N с компенсированной схемой и без неё.

условиям леммы 2.4.2. Однако, лемма 2.4.2 даёт только необходимые условия локального минимума. Достаточное условие локального минимума (положительная определённость матрицы Гессе целевой функции $\|\mathbf{X} - \mathbf{S}(S_{(\tau)}, A_{(\tau)})\|_{\mathbf{W}}^2$ [35, Theorem 2.3]) было проверено численно при $N < 1000$.

Для численного построения примера необходимо строить проекцию на $\mathcal{Z}(A_0^2)$. Для этого нужен алгоритм, который очень точно строит базис $\mathcal{Z}(A_0^2)$. Подпространство $\mathcal{Z}(A_0)$ состоит из полиномиальных последовательностей степени не более 2, $\mathcal{Z}(A_0^2)$ из степени не более 5, что объяснено в подразделе 5.1.1. Для вычисления $\mathbf{Z}(A_0^2)$ с наибольшей точностью были вычислены полиномы Лежандра [60] от 0 до 5 порядка в точках x_i . Затем полученный базис был подвергнут дополнительной ортогонализации. Таким образом, пример построен. Теперь перейдём к сравнению методов.

Для простоты ограничимся случаем, когда $\mathbf{W}$ — единичная матрица. Мы провели сравнение методов VPGN, S-VPGN и MGN, описанных в разделе 3.3.5,

при различных N от 20 до 50000 с использованием компенсированной схемы Горнера для вычисления базисов подпространств $\mathcal{Z}(A)$ и $\mathcal{Z}(A^2)$. Детали применения алгоритмов для вычисления базисов описаны в разделе 5.1.1. В качестве меры ошибки было взято евклидово расстояние между результатом работы алгоритма $\tilde{Y}^*$ и точкой локального минимума Y_N^* . Дополнительно, была использована мера для расчёта того, насколько полученное решение $\tilde{Y}^*$ управляется последней вычисленной ОЛРФ(A^*). Для этого была вычислена невязка $\frac{\|Q^T(A^*)\tilde{Y}^*\|}{\|A^*\|}$.

В качестве алгоритма локального поиска по направлению Δ_k на шаге 8 алгоритмов 3.3.4 и 3.3.5 возьмём простой перебор $\gamma = 1, 1/2, 1/4, \dots, 2^{-16}, 0$ до тех пор, пока не выполнится улучшение решения относительно предыдущей итерации: $\|X - S^*(A_{(\tau)}^{(k)} + \gamma\Delta_k)\|_{\mathbf{W}} \leq \|X - S^*(A_{(\tau)}^{(k)})\|_{\mathbf{W}}$. В этом случае, в качестве критерия остановки можно использовать условие $\gamma = 0$, т.е. невозможность получить более близкое решение, чем на предыдущей итерации. Проблема состоит в том, что если мы вначале возьмём $\gamma = 1$ (соответствующее полному шагу метода Гаусса-Ньютона), а относительное изменение будет мало, скажем

$$\left\| \frac{S^*(A_{(\tau)}^{(k)} + \Delta_k) - S^*(A_{(\tau)}^{(k)})}{S^*(A_{(\tau)}^{(k)})} \right\| < \zeta, \quad (5.1)$$

где ζ имеет порядок корня из машинного эпсилон (мы используем $\zeta = 5 \cdot 10^{-8}$), то указанный локальный поиск нельзя применять вследствие недостаточной точности вычисления целевой функции $\|X - S^*(A_{(\tau)}^{(k)} + \gamma\Delta_k)\|_{\mathbf{W}}^2$, сводящейся к вычислению плохо обусловленных скалярных произведений.

Давайте модифицируем сам метод локального поиска по направлению для случая, когда выполнено условие (5.1). Как метод MGN, так и метод VPGN можно рассматривать как последовательность параметров $A_{(\tau)}^{(k)}$, так и последовательность рядов $S^*(A_{(\tau)}^{(k)})$. Если критерий по последовательности рядов не работает, можно использовать критерий по набору параметров. Положим $\gamma = 1$,

если $k \geq 0$ и $\|\Delta_{k-1}\| > \|\Delta_k\|$, и $\gamma = 0$ в противном случае, т.е. завершим алгоритм тогда, когда очередной шаг по параметрам $\|\Delta_k\| = \|A_{(\tau)}^{(k+1)} - A_{(\tau)}^{(k)}\|$ стал больше предыдущего из-за достижения локального минимума с максимальной возможной точностью. Таким образом, мы получим комбинацию из метода локального поиска по направлению и критерия останова, которая работает даже тогда, когда численное вычисление целевой функции может не показывать улучшения из-за недостаточной точности его вычисления.

В качестве начальных коэффициентов ОЛРФ был использован вектор $A_0 + 10^{-6}(1, 1, 1, 1)^T$. Результаты для метрики-расстояния до Y_N^* представлены на рисунке 5.3, для невязки — на рисунке 5.4.

Все три метода сошлись в одну и ту же окрестность Y_N^* . Заметим, что алгоритм S-VPGN не даёт существенного улучшения по расстоянию до решения по сравнению с методом VPGN, в то время как решения метода MGN оказываются существенно ближе решений методов S-VPGN и VPGN из-за более устойчивого вычисления шага. К тому же заметим, что используемая в локальном поиске и изображённая на рисунке 5.3 граница ζ из условия (5.1) повлияла на вид ошибки для алгоритма MGN — благодаря перебору длин шагов γ , который, условно говоря, работает над этой границей, удалось сделать решение близким, несмотря на падение точности при вычислении шага, но эта граница не повлияла на результат методов VPGN и S-VPGN из-за худшей точности вычисления шага. С точки зрения невязки решения алгоритмы S-VPGN и MGN идентичны и имеют значительное преимущество по устойчивости над методом VPGN.

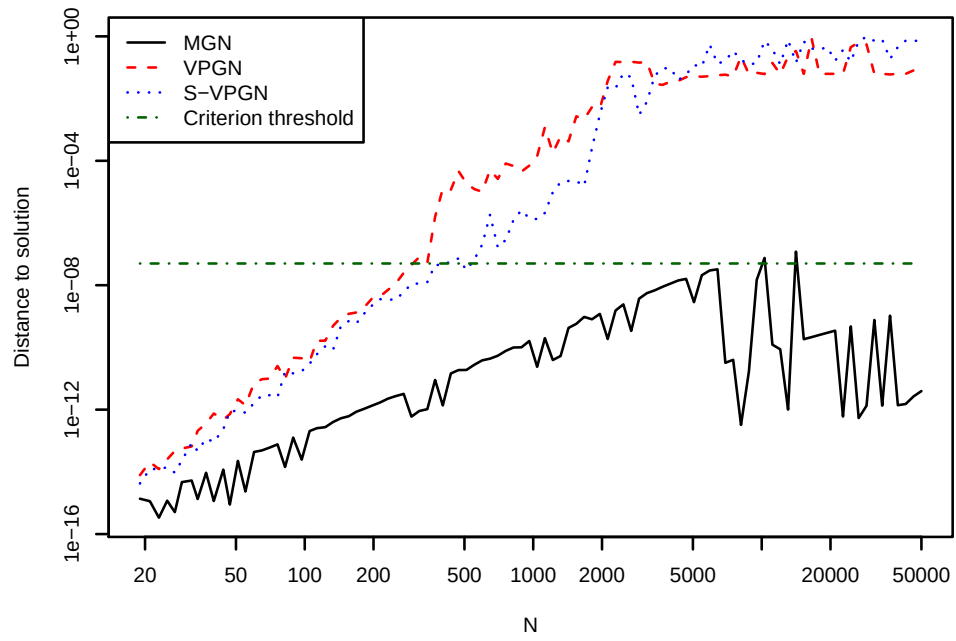


Рис. 5.3. Сравнение алгоритмов по расстоянию до ответа при различных N .

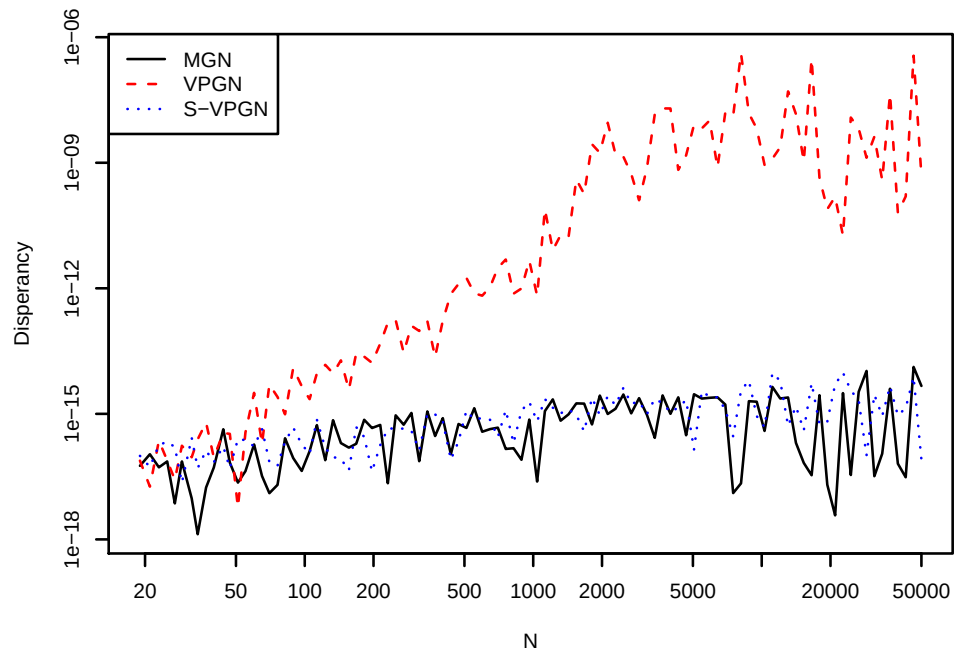


Рис. 5.4. Сравнение алгоритмов по невязке при различных N .

5.2. Исследование свойств оценок сигнала с помощью статистического моделирования

5.2.1. Исследование оценок, полученных методом Кэдзоу

Для исследования работы предложенных в главе 4 алгоритмов были проведено численное моделирование. Сравнение происходило на двух рядах длины $N = 40$ и ранга $r = 2$:

1. Синусоидальный сигнал $\mathbf{S}^{(1)} = (s_1^{(1)}, \dots, s_N^{(1)})^T$, где $s_k^{(1)} = 5 \sin \frac{2\pi k}{6}$,
2. Линейный сигнал $\mathbf{S}^{(2)} = (s_1^{(2)}, \dots, s_N^{(2)})^T$, где $s_k^{(2)} = 100 \frac{k}{N}$.

Были промоделированы ряды $\mathbf{X}_k^{(j)} = \mathbf{S}^{(j)} + \boldsymbol{\epsilon}_k(b_1)$, где $\boldsymbol{\epsilon}_k$ — независимые реализации длины N процесса авторегрессии порядка 1 с параметром b_1 : $\epsilon_i = b_1 \epsilon_{i-1} + \varepsilon_i$, где ε_i — гауссовский белый шум с нулевым средним и дисперсией σ^2 , k — номер реализации, $k = 1, \dots, M$. Мы рассматривали $b_1 = 0$ — случай белого шума и $b_1 = 0.9$ — случай красного шума. Затем к полученным промоделированным рядам был применен алгоритм *Cadzow(quadratic)* из раздела 4.6.1, и для полученных оценок $\widehat{\mathbf{S}}_k^{(j)}$ была оценена их точность с помощью MSE (mean squared error) — среднеквадратической ошибки от соответствующего сигнала $\mathbf{S}^{(j)}$, где среднее взято и по длине ряда, и по числу реализаций:

$$\text{MSE}_j = \frac{1}{M} \frac{1}{N} \sum_{k=1}^M \|\widehat{\mathbf{S}}_k^{(j)} - \mathbf{S}^{(j)}\|^2.$$

В качестве критерия остановки STOP в алгоритме *Cadzow(quadratic)* было использовано условие $\frac{\|\mathbf{X}_i - \mathbf{X}_{i-1}\|_{\mathbf{W}}}{\|\mathbf{X}_0\|_{\mathbf{W}}} < 10^{-6}$. Заметим, что в этом случае результаты алгоритмов *Cadzow(quadratic)* и *Cadzow(box)* из раздела 4.6.1 совпали (так как алгоритмы поиска весов дали одинаковый ответ). В качестве длины окна была взята половина длины ряда $L = 20$, следуя рекомендации в [61].

Для начала, мы промоделировали $M = 10000$ реализаций каждого ряда, и к одними и тем же реализациям мы применили метод *Cadzow(quadratic)* при различных α , использующихся при решении задач поиска весов (4.7) и (4.8). Напомним, что чем ближе α к 0, тем ближе матрица весов $\mathbf{R}$ в (4) к вырожденной, но тем лучше может быть задаваемая ими аппроксимация весов $\mathbf{W}$ в (3). Для MSE были построены 95-процентные выборочные доверительные интервалы («Simulation»), а так же вычислена теоретическая оценка MSE линейризованного алгоритма Кэдзоу («Theory») с помощью замечания 4.7.4 и граница Рао-Крамера с помощью предложения 2.5.2 («CRB»). Более того, было вычислено среднее число шагов, затраченное алгоритмом *Cadzow(quadratic)* («Number of iterations»), и на том же графике изображена величина (4.32) («Converted eigenvalue»), объясняющая число шагов.

Для начала, мы рассмотрим сигнал $\mathbf{S}^{(1)}$ и шум с $b_1 = 0$, то есть случай белого шума, и с двумя разными уровнями шума $\sigma = 1.5$ и $\sigma = 0.5$. Результаты представлены на рисунках 5.5 и 5.6.

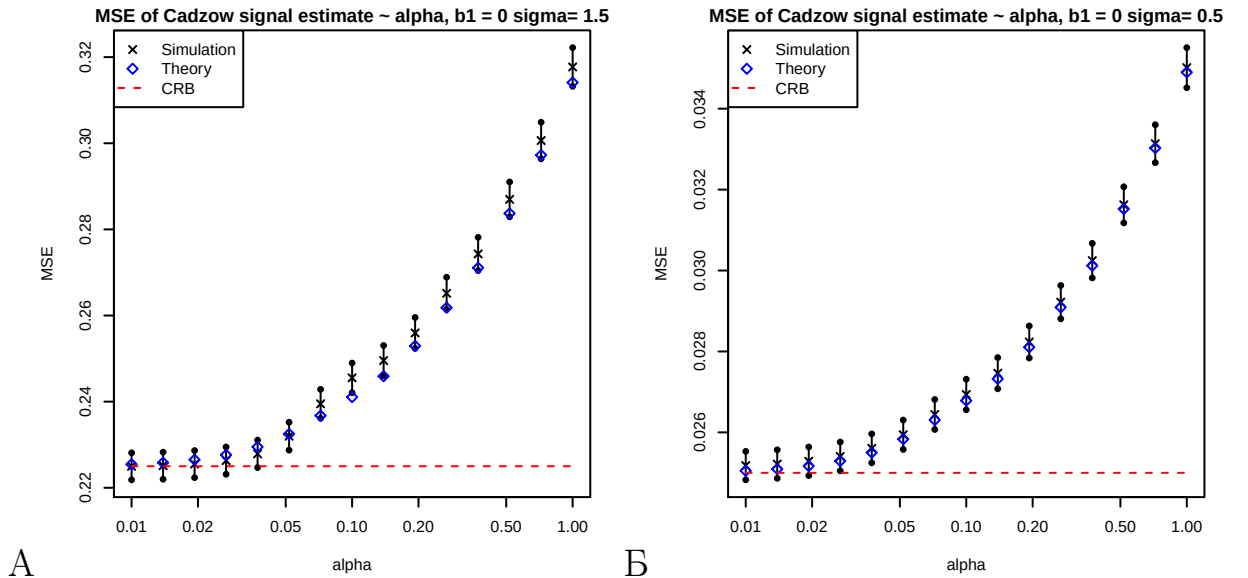


Рис. 5.5. Сравнение MSE для оценки сигнала $\mathbf{S}^{(1)}$ с помощью алгоритма *Cadzow(quadratic)* для белого шума при различных α и (А) $\sigma = 1.5$ (Б) $\sigma = 0.5$.

Прокомментируем результаты на рисунках 5.5 и 5.6. Заметим, что под-

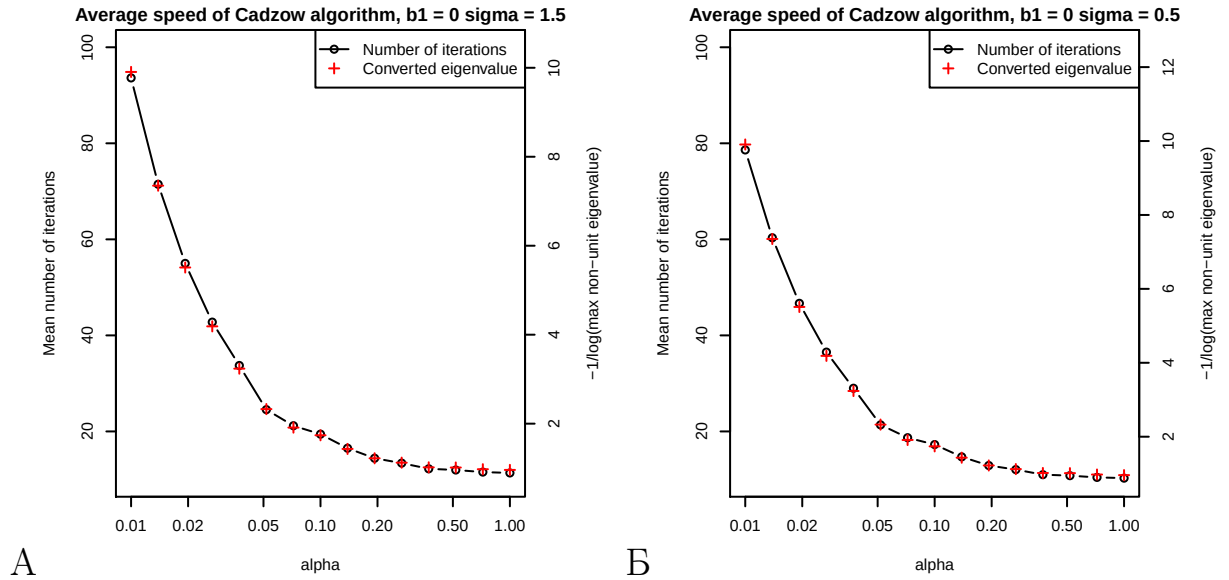


Рис. 5.6. Сравнение числа итераций алгоритма *Cadzow(quadratic)* для сигнала $S^{(1)}$ для белого шума при различных α и (А) $\sigma = 1.5$ (Б) $\sigma = 0.5$.

твердился теоретический результат леммы 4.7.2 для оценок ошибок алгоритма Кэдзоу, однако заметим, что в случае (Б) на рисунке 5.5, в котором соотношение сигнал/шум больше, чем в случае (А) ($\sigma = 0.5$ против $\sigma = 1.5$, что должно дать увеличение MSE примерно в 9 раз), практический эксперимент оказался ближе к теории (теоретические значение ближе к середине доверительных интервалов), что естественно, так как сходимость ошибок первого порядка возникает при соотношении сигнал/шум, стремящемся к бесконечности. Подтвердился и результат замечания 4.7.3: реальное сделанное алгоритмом число шагов достаточно точно описывается свойствами собственных чисел матрицы $\mathbf{P}_{(S_0)}$ из (4.30) (линии на рисунке 5.6 практически совпадают).

Уменьшение параметра α даёт сходимость и теоретической, и практической оценки к границе Рао-Крамера. Известно, что если N кратно L , то при отсутствии ограничения на число обусловленности матрицы $\mathbf{R}$ можно получить $\mathbf{L} * \mathbf{R} = \mathbf{I}_N$ [23]. Уменьшением α условие на вырожденность $\mathbf{R}$ в задаче (4.8) становится всё менее сильным, а в пределе при $\alpha = 0$ пропадает. Однако, платой за более высокую точность оценивания является увеличение времени работы,

что объясняется свойствами собственных чисел матрицы (4.30), см. замечание (4.7.3). Случай $\alpha = 1$ соответствует стандартному методу Кэдзоу с $\mathbf{L} = \mathbf{I}_L$, $\mathbf{R} = \mathbf{I}_K$.

Далее был рассмотрен сигнал $\mathbf{S}^{(1)}$ и шум с $b_1 = 0.9$, то есть случай красного шума, $\sigma = 0.1$, но рассмотрены два варианта выбора матрицы $\mathbf{L}$: один соответствовал модели шума ($\mathbf{L} = \Sigma_L^{-1}$), а второй нет ($\mathbf{L} = \mathbf{I}_L$, как надо было бы сделать в случае белого шума). Был применен метод *Cadzow(quadratic)*. Результаты представлены на рисунках 5.7 и 5.8.

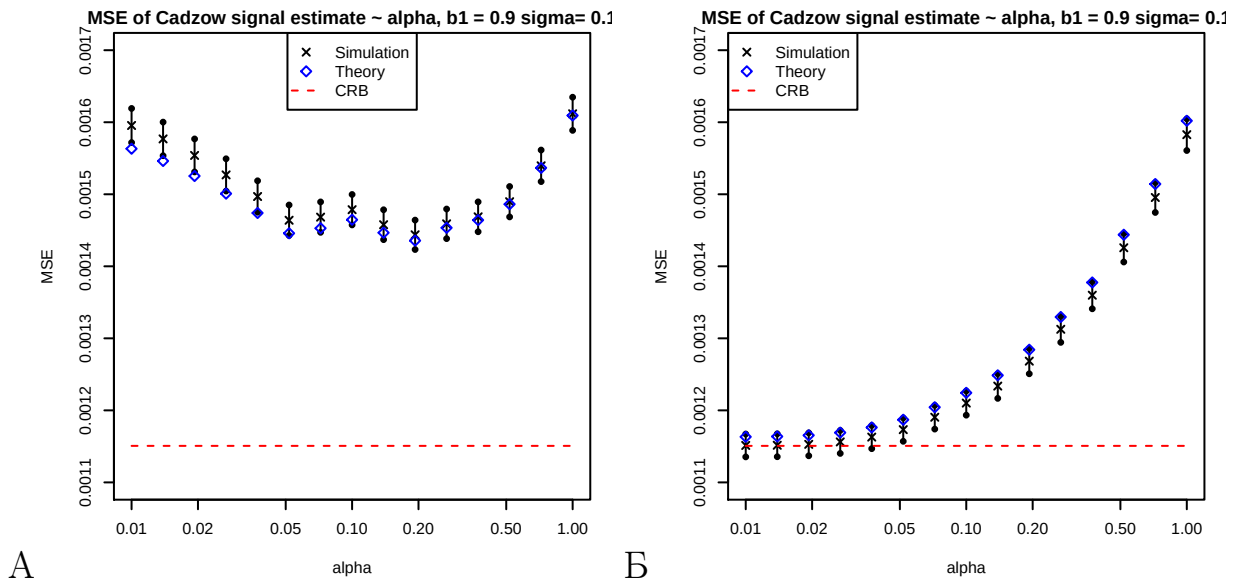


Рис. 5.7. Сравнение MSE для оценки сигнала $\mathbf{S}^{(1)}$ с помощью алгоритма *Cadzow(quadratic)* для красного шума при различных α и (А) диагональной $\mathbf{L}$ (Б) трехдиагональной $\mathbf{L}$.

Прокомментируем результаты на рисунке 5.7. Заметим, что выбор неправильной матрицы $\mathbf{L}$ ведет к тому, что уменьшение α не приводит к улучшению оценки. Выбором матрицы $\mathbf{L} = \Sigma_L^{-1}$ мы добились того, что уменьшение α ведёт ко всё более близкой к оценке Рао-Крамера ошибке оценивания сигнала, однако, в отличие от предыдущего случая белого шума, при таком выборе $\mathbf{L}$ невозможно выбрать матрицу $\mathbf{R}$ так, чтобы получить нужную $\mathbf{W}_0 = \mathbf{L} * \mathbf{R}$, $\mathbf{W}_0 = \Sigma_N^{-1}$ даже при отсутствии ограничения на её вырожденность! Тем не менее, полученная при малом α ошибка близка к границе Рао-Крамера, что

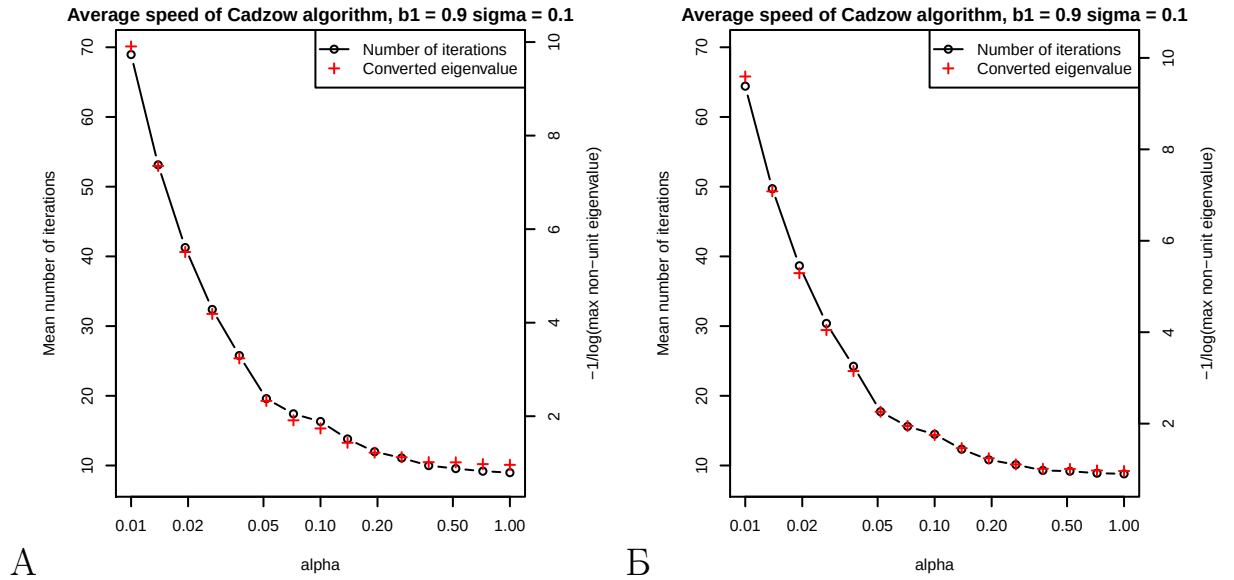


Рис. 5.8. Сравнение числа итераций алгоритма *Cadzow(quadratic)* для сигнала S_1 для красного шума при различных α и (А) диагональной $\mathbf{L}$ (Б) трёхдиагональной $\mathbf{L}$.

оправдывает сделанный нами выбор матрицы $\mathbf{L}$ в постановке задачи (4.7).

Полученные теоретические результаты были проверены и для полиномиального сигнала $S^{(2)}$ при белом шуме $b_1 = 0$ и $\sigma = 1$. Этот сигнал интересен тем, что содержит кратные сигнальные корни, см. определение 1.1.4, и поэтому прежде оценивание таких сигналов было затруднено. Результаты представлены на рисунке 5.9. Они аналогичны результатам для сигнала $S^{(1)}$, представленным на рисунках 5.5 и 5.6.

В следующем численном эксперименте рассмотрим зависимость ошибки оценки сигнала с помощью алгоритма *Cadzow(quadratic)* от числа шагов для разных значений α . Мы промоделировали $M = 1000$ реализаций, и изобразили MSE оценки вместе с 95-процентными доверительными интервалами, полученные на первых 15 итерациях *Cadzow(quadratic)*, при фиксированных α . В качестве теоретической оценки («Theory») мы использовали результаты следствия 4.7.2, в качестве теоретической оценки при большом числе итераций («Theory limit») использовали результат замечания 4.7.4, а также изобразили границу Рао-Крамера с помощью предложения 2.5.2 («CRB»).

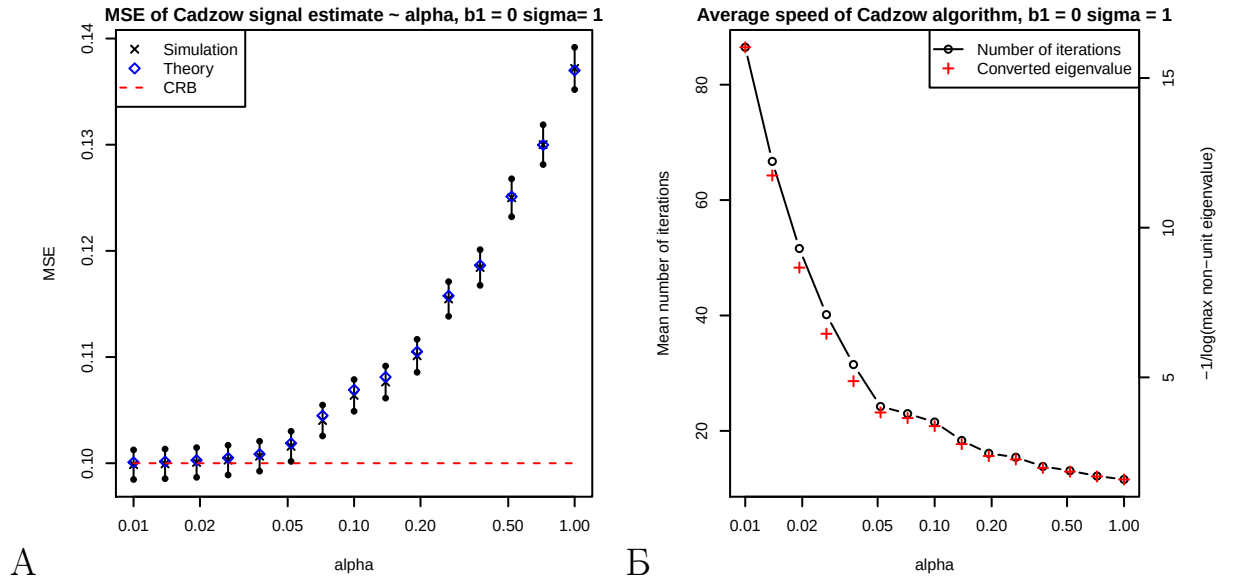


Рис. 5.9. (А) Сравнение MSE для оценки сигнала S_2 с помощью алгоритма *Cadzow(quadratic)* для белого шума при различных α , $\sigma = 1$ (Б) сравнение числа итераций алгоритма.

Результат для сигнала $S^{(1)}$, $b_1 = 0$, $\sigma = 1$ при $\alpha = 1$ и $\alpha = 0.1$ представлен на рисунке 5.10.

Пример на рисунке 5.10 (А) примечателен тем, что первая итерация с $\alpha = 1$ соответствует оценке, полученной методом SSA [8]. Мы видим, что при увеличении числа итераций полученная оценка по MSE не улучшается, что подтверждается как моделированием, так и следствием 4.7.2. Отсюда, в частности, можно сделать вывод, что нет особого смысла улучшать результат метода SSA с помощью последующих итераций. В то же время, использование $\alpha = 0.1$ позволяет уменьшить ошибку в пределе ценой повышения трудоёмкости, при этом ошибка на первой итерации существенно больше, чем при использовании SSA.

5.2.2. Исследование оценок, полученных модифицированным методом Гаусса-Ньютона (MGN)

В этом разделе на примере ряда $S^{(1)}$ из раздела 5.2.1 моделированием была исследована точность оценок сигнала $S^{(1)}$ с помощью модифицированного метода локального поиска Гаусса-Ньютона MGN, предложенного в разделе 3.3.

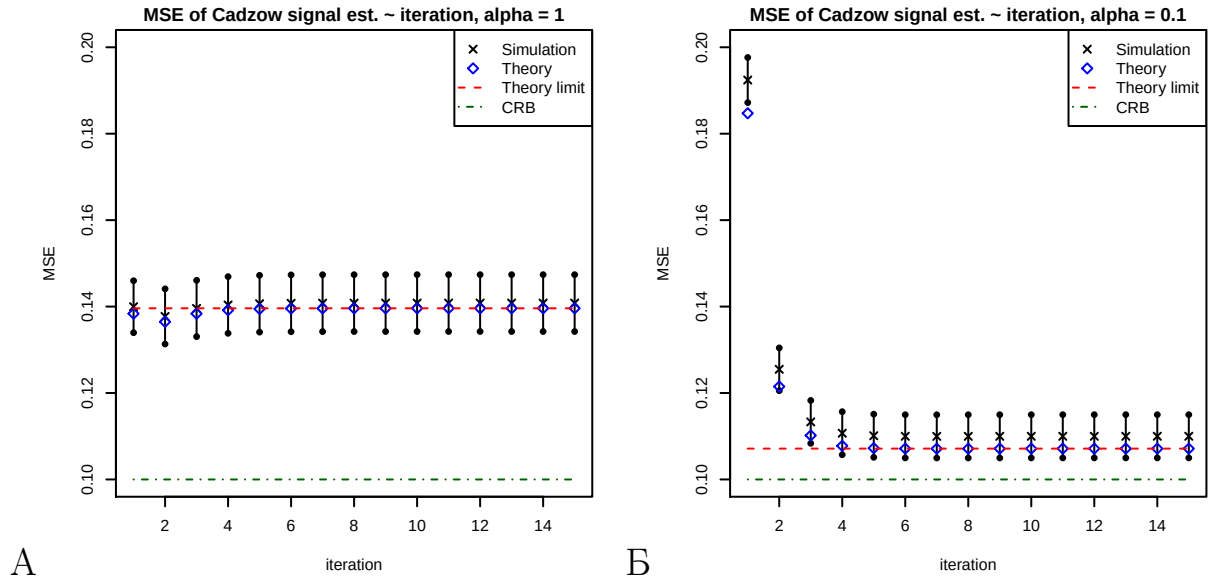


Рис. 5.10. Сравнение MSE для оценки сигнала S_1 с помощью алгоритма $Cadzow(quadratic)$ для белого шума и различном числе итераций (А) $\alpha = 1$ (Б) $\alpha = 0.1$.

Для выбора начального значения использовался метод $Cadzow(quadratic)$, описанный в разделе 4.6.1. Детали применения метода Кэдзоу для поиска начального значения описаны в подразделе 4.6.2. Был использован параметр $\alpha = 0.1$, остальные параметры и критерий останковки такие же, какие были использованы в разделе 5.2.1. Далее был применён метод MGN (см. алгоритм 3.3.5), критерий останковки тот же, что использовался в разделе 5.1.2.

Было сделано $M = 1000$ моделирований ряда $\mathbf{X} = \mathbf{S}^{(1)} + \boldsymbol{\epsilon}(b_1)$, к каждой реализации был применён алгоритм MGN и получена оценка сигнала $\tilde{\mathbf{S}}^{(1,k)} = (\tilde{s}_1^{(1,k)}, \dots, \tilde{s}_N^{(1,k)})^T$, $k = 1, \dots, M$. В качестве меры точности оценки взята поточечная (по элементам ряда) среднеквадратичная ошибка:

$$MSE_j = \frac{1}{M} \sum_{k=1}^M (s_j^{(1)} - \tilde{s}_j^{(1,k)})^2.$$

В качестве теоретической оценки MSE взят результат предложения 2.5.1, который даёт границу Рао-Крамера («Theory (CRB)») как асимптотическую оценку MSE. Для оценки, полученной моделированием, изображён 95-процентный выборочный доверительный интервал.

Результат для сигнала $S^{(1)}$ и шума с $b_1 = 0$ и двух значений $\sigma = 4$ и $\sigma = 1$ представлены на рисунке 5.11, а для случая $b_1 = 0.9$, $\sigma = 3.3$ и $\sigma = 0.1$ — на рисунке 5.12. Значения параметра σ выбраны для наглядности получившихся результатов.

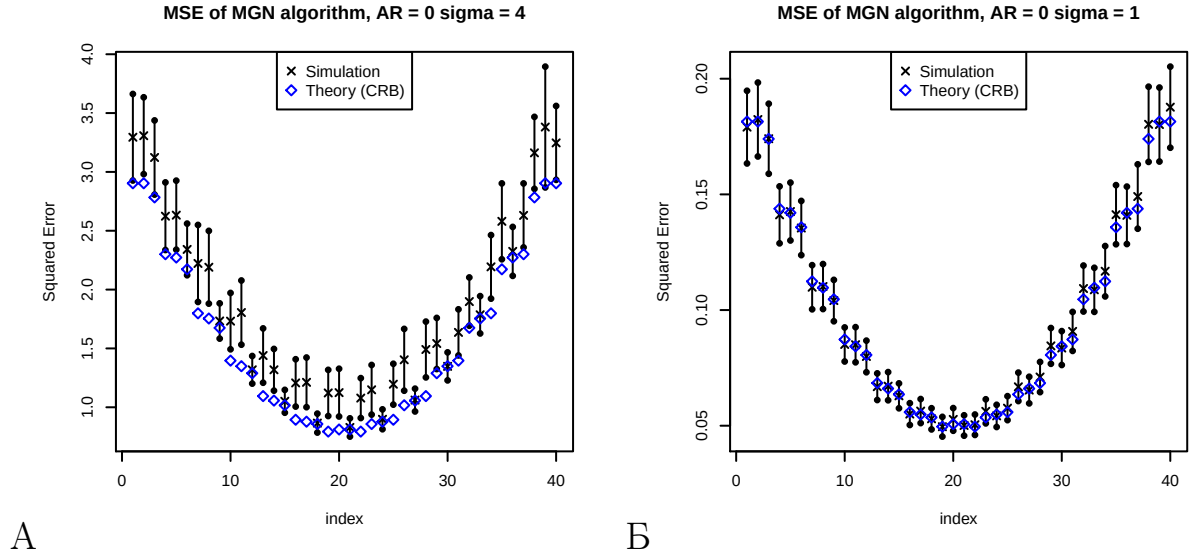


Рис. 5.11. Сравнение MSE для оценки сигнала $S^{(1)}$ методом MGN при белом шуме с (А) $\sigma = 4$ (Б) $\sigma = 1$.

В обоих случаях рисунки 5.11 и 5.12 показывают, что при увеличении соотношения сигнал/шум, то есть уменьшении σ , наблюдается сходимость результатов, полученных моделированием, к оценке, полученной в предложении 2.5.1, что согласуется с результатом теоремы 2.5.1. Более того, данные примеры демонстрируют, что метод Кэдзоу даёт хорошее начальное приближение для метода MGN при большом соотношении сигнал/шум: значение границы Рао-Крамера попадает в доверительный интервал; при малом же соотношении сигнал/шум асимптотический результат не работает.

Таким образом, при большом соотношении сигнал/шум ошибки полученных с помощью метода MGN оценок сигнала достигают границы Рао-Крамера, что является улучшением относительно оценивания сигнала с помощью метода Кэдзоу.

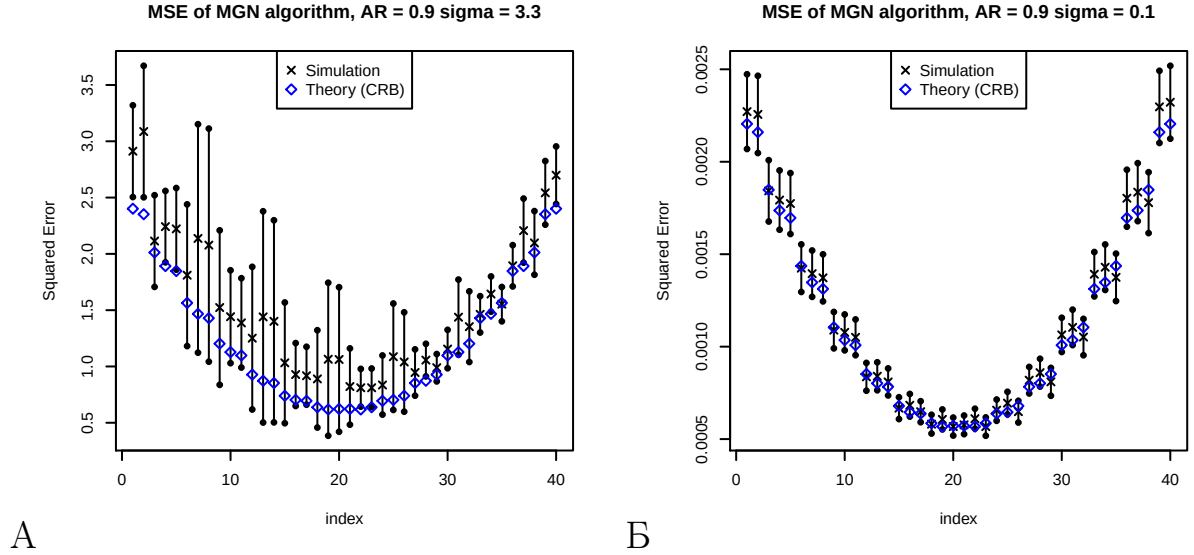


Рис. 5.12. Сравнение MSE для оценки сигнала $S^{(1)}$ методом MGN при красном шуме, $b_1 = 0.9$ с (А) $\sigma = 3.3$ (Б) $\sigma = 0.1$.

5.3. Применение модифицированного метода

Гаусса-Ньютона к данным экспрессии генов

В [62] построена модель экспрессии ряда генов, в которой можно выделить параметры, не зависящие от настроек микроскопа. Оценивание этих параметров позволяет описать динамику развития эмбриона в терминах активности рассматриваемых генов. Используемая модель сигнала соответствует ряду конечного ранга. Поэтому в этом разделе мы применим методы из раздела 3.3 к данным экспрессии генов, чтобы оценить сигнал и параметры. Полученные в разделе 2.5 теоретические результаты позволяют предположить, что полученные оценки будут более точными.

Исходные данные представляют из себя трёхмерные сканы эмбрионов дрозофилы (*Drosophila*), в каждой точке которых интенсивность свечения пропорциональна концентрации мРНК гена *bicoid* (*bcd*). Измерения проводились в течение трёх стадий развития: Cleavage, nc10-13 и nc14, соответственно, данные разбиты по этим трём группам. Трёхмерные координаты следующие: с головы

к хвосту эмбриона (AP), сверху вниз (DV) и изнутри к наружи (basal-apical, BA). Для каждого скана из исходных трёхмерных данных с помощью проекции на равномерную решётку по оси AP и оставления точек от 10 % до 90 % от длины эмбриона были получены два одномерных профиля — наружный (apical) и внутренний (basal). Подробное описание процесса сбора данных и построения профилей не является целью этого раздела, см. [62], где данный вопрос подробно разобран.

5.3.1. Модель данных

В [62] рассматривается следующая модель зависимости концентрации мРНК *bcd* $s(x)$ в зависимости от AP координаты x :

$$s(x) = C_{\text{anterior}} \lambda_{\text{anterior}}^x + C_{\text{shallow}} \lambda_{\text{shallow}}^x, \quad (5.2)$$

таким образом, модель из себя представляет сумму двух экспонент. Две компоненты в модели — это *anterior* (головная компонента, соответствующая резко затухающему паттерну в начале профиля), для которой $\lambda_{\text{anterior}} < 1$, и *shallow* (пологая компонента), для которой может быть как $\lambda_{\text{shallow}} < 1$ (затухание), так и $\lambda_{\text{shallow}} > 1$ (рост). В частном случае $\lambda_{\text{shallow}} = 1$ модель принимает вид «экспонента плюс константа». C_{anterior} и C_{shallow} — коэффициенты перед соответствующим экспонентам.

В качестве модели исходных данных использована модель $s(x) + \varepsilon(x)$, где $\varepsilon(x)$ представляет из себя шум.

5.3.2. Оценивание параметров сигнала

В [62] для оценивания параметров сигнала использован метод ESPRIT (точнее говоря, LS-ESPRIT [4, 5]). ESPRIT даёт оценки для $\lambda_{\text{anterior}}$ и λ_{shallow} , после чего C_{anterior} и C_{shallow} оцениваются по стандартному методу наименьших

квадратов, так как они линейно входят в модель (5.2). Так как считается, что головная компонента убывает быстро ($\lambda_{\text{anterior}} < 1$), а пологая компонента близка к константе (λ_{shallow} находится возле единицы), то оценки параметров упорядочены соответствующим образом.

В дополнение к оценкам ESPRIT для тех же самых профилей построена оценка $\lambda_{\text{anterior}}$ и λ_{shallow} с помощью Модифицированного метода Гаусса-Ньютона (MGN) с использованием алгоритма Cadzow в качестве метода поиска начального значения. Это возможно сделать, так как значения модели (5.2) на равномерной решётке представляют из себя ряд ранга 2. Полученная в результате ОЛРФ(A) имела ненулевую последнюю координату, соответственно, $\lambda_{\text{anterior}}$ и λ_{shallow} были вычислены как корни соответствующего характеристического ЛРФ полинома, см. определение 1.1.4. Шум $\varepsilon(x)$ имеет сложную структуру, так как состоит из шума от препроцессинга и вариабельной активности генов. Поэтому шум не удовлетворяет модели стационарного процесса, хотя не слишком сильно отличается от белого шума. В качестве более простого варианта была взята матрица $\mathbf{W} = \mathbf{I}_N$, соответствующая случаю белого шума.

Пример сравнения оценки сигнала для одного и того же наружного профиля методами ESPRIT и MGN представлен на рисунке 5.13, соответствующие оценкам головной и пологой части для каждого из методов представлены на рисунке 5.14. Заметим, что полученная с помощью метода MGN головная компонента более резко убывает, чем с помощью оценки по ESPRIT. Это связано с тем, что полученная MGN оценка лучше аппроксимировала резкий спад на левом краю ряда (среднеквадратичная ошибка аппроксимации по ESPRIT равна 13.62, для оценки по MGN равна 12.95).

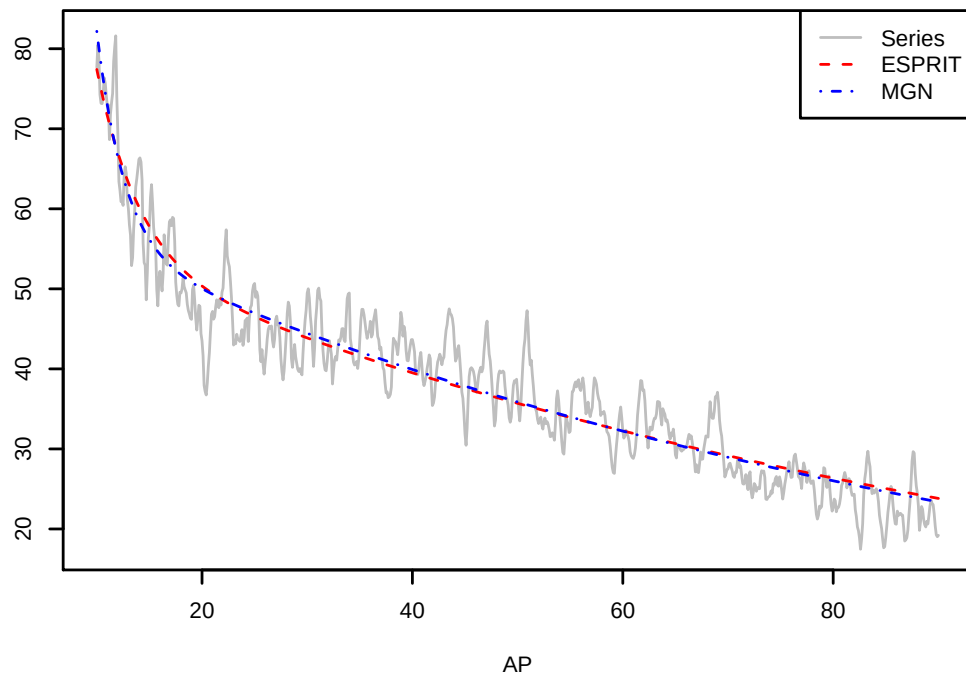


Рис. 5.13. Различные способы оценки сигнала в arisal профиле.

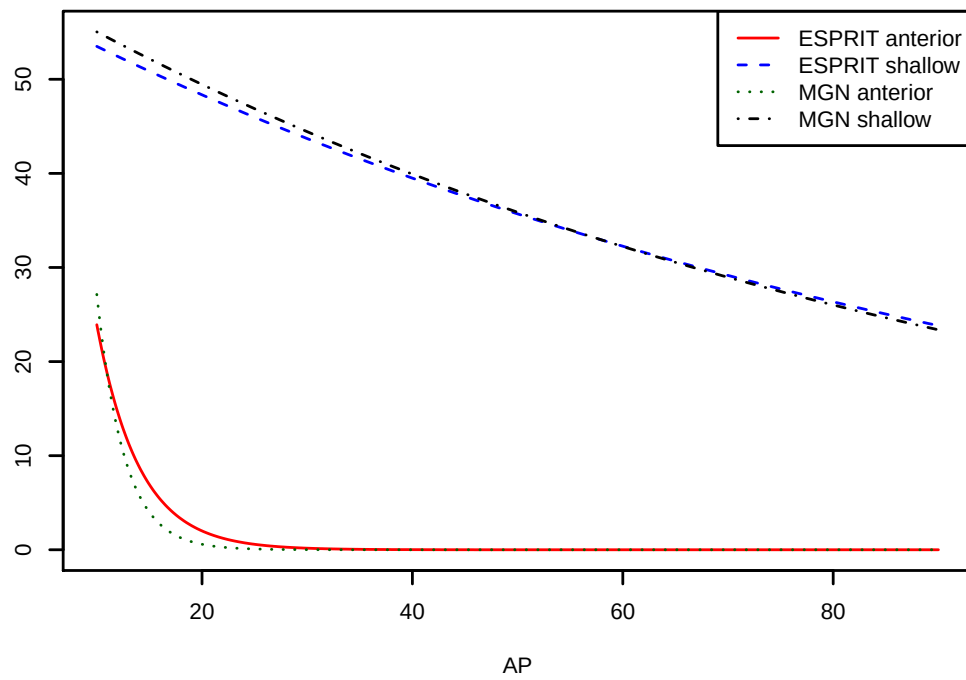


Рис. 5.14. Различные способы оценки частей сигнала в arisal профиле.

5.3.3. Сравнение характеристик полученных оценок

После очистки данных, описанной в [62], были построены доверительные интервалы по каждой из трёх групп эмбрионов: Cleavage, nc10-13 и nc14. В качестве основных рассматриваемых характеристик в [62] используется $\lambda_{\text{anterior}}^{(\text{apical})}$ — оценка головной экспоненты наружного профиля и $C^{ab} = \log(C_{\text{anterior}}^{(\text{apical})}/C_{\text{anterior}}^{(\text{basal})})$ — независимое от линейных преобразований отношение коэффициентов перед экспонентой наружного и внутреннего профиля. Для каждой из характеристик были построены 90-процентные доверительные интервалы для оценки по ESPRIT и для оценки по MGN. Результаты представлены в таблицах 5.1 и 5.2 (для краткости $\lambda_{\text{anterior}}^{(\text{apical})}$ записано как λ).

Таблица 5.1. Средние значения и границы 90-процентных доверительных интервалов характеристик $\lambda_{\text{anterior}}^{(\text{apical})}$ и C^{ab} , оценённых по методу ESPRIT.

	Среднее λ	λ 5%	λ 95%	Среднее C^{ab}	C^{ab} 5%	C^{ab} 95%
10-13c	0.879	0.862	0.895	-0.170	-0.300	-0.041
14c	0.871	0.862	0.880	0.652	0.508	0.797
Cleavage	0.815	0.785	0.846	0.283	0.124	0.443

Таблица 5.2. Средние значения и границы 90-процентных доверительных интервалов характеристик $\lambda_{\text{anterior}}^{(\text{apical})}$ и C^{ab} , оценённых по методу MGN.

	Среднее λ	λ 5%	λ 95%	Среднее C^{ab}	C^{ab} 5%	C^{ab} 95%
10-13c	0.845	0.827	0.863	-0.151	-0.249	-0.053
14c	0.849	0.836	0.862	0.589	0.467	0.710
Cleavage	0.764	0.729	0.799	0.167	0.053	0.282

Те же самые результаты в графическом виде с построенными для каждой из групп 80-процентными доверительными эллипсоидами представлены на рисунке 5.15.

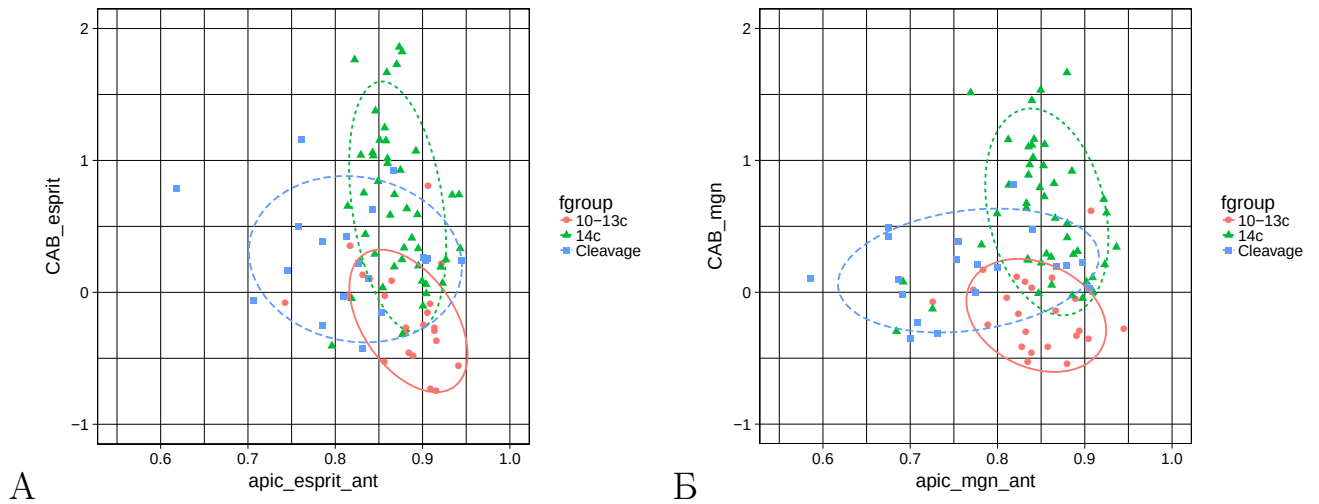


Рис. 5.15. $\lambda_{\text{anterior}}^{(\text{apical})}$ и C^{ab} для оценок по методу (А) ESPRIT (Б) MGN.

Более того, были сравнены ошибки аппроксимации сигнала, построенных различными методами. Результаты представлены в таблице 5.3. Видно, что оценённый с помощью метода MGN сигнал более точно аппроксимирует данные (среднеквадратичная ошибка аппроксимации по каждой паре группа/профиль уменьшилось). Сравнение средних оценок параметров в таблицах 5.1 и 5.2 показывает, что среднее значение характеристики $\lambda_{\text{anterior}}^{(\text{apical})}$ уменьшилось, и, следовательно, привело к более резкому поведению оценённой головной компоненты.

Таблица 5.3. Средние значения ошибок аппроксимации для оценок сигнала, полученных методами ESPRIT и MGN.

	apical, ESPRIT	apical, MGN	basal, ESPRIT	basal, MGN
10-13c	18.16	17.67	8.97	8.66
14c	17.02	16.79	4.10	3.95
Cleavage	22.64	21.64	24.26	22.70

5.4. Аппроксимация временными рядами конечного ранга при неизвестных параметрах авторегрессионного шума

Рассмотрим задачу оценки сигнала по наблюдаемому ряду $\mathbf{X} = \mathbf{S} + \boldsymbol{\epsilon}$ длины N , где $\mathbf{S}$ — сигнал конечного ранга, $\boldsymbol{\epsilon}$ — шум, например, представляющий из себя процесс авторегрессии порядка p : $\epsilon_i = \sum_{j=1}^p b_j \epsilon_{i-j} + \varepsilon_i$, где ε_i — гауссовский белый шум с нулевым средним и дисперсией σ^2 . На практике, коэффициенты процесса авторегрессии неизвестны, то есть неизвестен точный вид ковариационной матрицы $\boldsymbol{\Sigma}$, что не позволяет сразу решать задачу (3). Поэтому, задав порядок p процесса авторегрессии и ранг сигнала r , оценку сигнала можно построить с помощью максимизации функции правдоподобия:

$$\max_{\mathbf{Y} \in \overline{\mathcal{D}_r}, B=(b_1, \dots, b_p)^T, \sigma > 0} \log f_{\mathbf{X}, \mathbf{D}\boldsymbol{\Sigma}(B, \sigma)\mathbf{D}}(\mathbf{Y}), \quad (5.3)$$

где $f_{\mathbf{X}, \mathbf{D}\boldsymbol{\Sigma}(B, \sigma)\mathbf{D}}$ — функция правдоподобия многомерного нормального распределения со средним $\mathbf{X}$ и ковариационной матрицей $\mathbf{D}\boldsymbol{\Sigma}(B, \sigma)\mathbf{D}$, где $\boldsymbol{\Sigma}(B, \sigma) \in \mathbb{R}^{N \times N}$ — автоковариационная матрица процесса $\text{AR}(p)$ с параметрами $B = (b_1, \dots, b_p)^T$ и σ (в явном виде обратная автоковариационная матрица указана в [49]), $\mathbf{D}$ — заранее известная диагональная положительно определённая матрица, задающая огибающую шума (вид обратной матрицы приведён в (4.26)). Для решения этой задачи воспользуемся покоординатным спуском, решая поочерёдно задачу аппроксимацию рядом конечного ранга и поиска коэффициентов процесса авторегрессии. Ниже данный подход записан в виде алгоритма.

Алгоритм 5.4.1 (Покоординатный спуск в случае неизвестных коэффициентов авторегрессии). ***Вход:*** B_0 — начальный вектор параметров авторегрессии, T — число итераций.

1: Положить $i = 0$.

- 2: **while** $i < T$ или не выполнен критерий остановки **STOP do**
- 3: Положить $i = i + 1$.
- 4: Вычислить $Y_i = \arg \max_{Y \in \overline{\mathcal{D}_r}} \log f_{X, \mathbf{D}\Sigma(B_{i-1}, \sigma_{i-1})\mathbf{D}}(Y)$, что сводится к решению задачи (3) для ряда X и $\mathbf{W} = \mathbf{D}^{-1}\Sigma^{-1}(B_{i-1}, \sigma_{i-1})\mathbf{D}^{-1}$.
- 5: Вычислить $(B_i, \sigma_i) = \arg \max_{B=(b_1, \dots, b_p)^T, \sigma} \log f_{X, \mathbf{D}\Sigma(B, \sigma)\mathbf{D}}(Y_i)$, что сводится к задаче поиска коэффициентов процесса $AR(p)$ для ряда $\mathbf{D}^{-1}(X - Y_i)$ (например, с помощью алгоритма [63]).
- 6: **end while**
- 7: **return** Y_i .

Пример использования подхода

В качестве примера был взят ряд *US Unemployment, male* — ежемесячные данные по безработице в тысячах человек с 1948 по 1991 год, длина ряда равна $N = 408$, данные взяты из [64]. Тот же самый ряд рассматривался в качестве примера в статье [65], где для его исследования применялись техники улучшения разделимости метода SSA [8].

К этому ряду был применён алгоритм 5.4.1 в предположении, что ряд из себя представляет сумму сигнала конечного ряда и гетероскедастичного красного шума. Сделано предположение о том, что шум является «мультипликативным», то есть огибающая шума (диагональ матрицы $\mathbf{D}$) примерно пропорциональна тренду. Поэтому была сделана «грубая» оценка тренда с помощью применения метода SSA [8, 21] с небольшой длиной окна $L = 84$ и взятия первой собственной тройки, таким образом, метод SSA использовался для сглаживания исходного ряда [8]. Ряд вместе с оценкой тренда представлен на рисунке 5.16.

Так как задача (5.3) может иметь много локальных минимумов, алгоритм 5.4.1 может сходиться к различным решениям при различных начальных коэффициентах авторегрессии. Было взято $B_0 = (0.9, 0, \dots, 0)^T$ и сделано 10 итера-

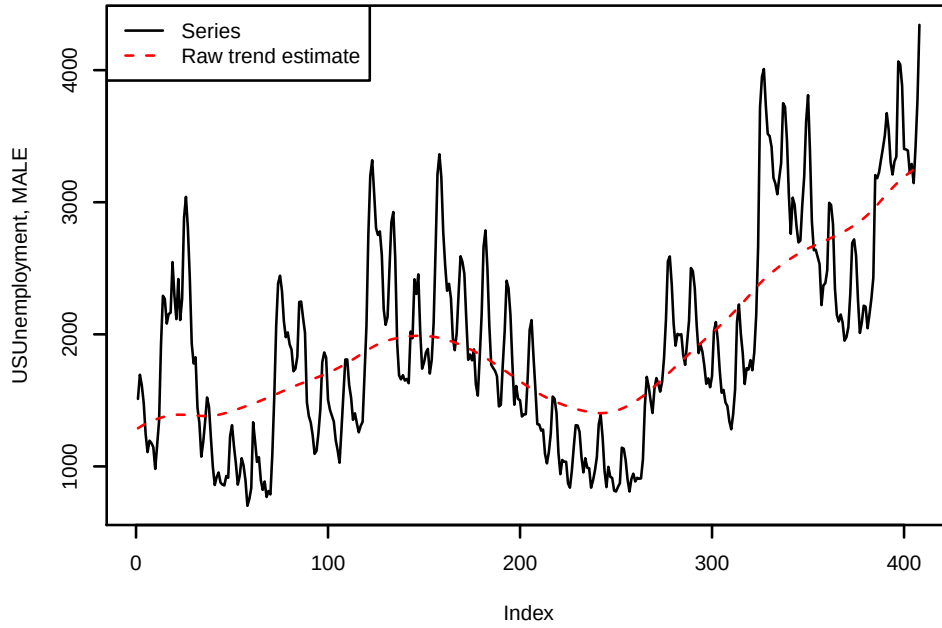


Рис. 5.16. Ряд *US Unemployment, male* и оценка тренда по методу SSA.

ций алгоритма 5.4.1, то есть изначально предполагалось, что шум красный, чтобы получить как можно более простую модель при подборе p . Выбор параметров был произведён по байесовскому информационному критерию (BIC) [66], который в данном случае выглядит как $-2 \log(f_{\mathbf{X}, \mathbf{D}\Sigma(B^*, \sigma^*)\mathbf{D}}(\mathbf{Y}^*)) + (2r + p) \log N$, где B^* , $\mathbf{Y}^*$, σ^* — решения задачи (5.3), $2r + p$ — общее число степеней свободы. Сравнением BIC с соседними значениями параметров был получен ранг ряда $r = 16$ и порядок процесса авторегрессии $p = 3$. Подзадача (3) решалась при помощи комбинации метода *Cadzow(box)* с $L = N/2$ и $\alpha = 0.3$, см. раздел 4.6.1 для выбора начальных значений для параметров сигнала, и предложенного модифицированного метода Гаусса-Ньютона MGN, см. раздел 3.3.5. Результат представлен на рисунке 5.17. Выбор параметров алгоритма Кэдзоу объясняется тем, что при меньших α или L итерации алгоритма Кэдзоу (4.1) испытывали неустойчивость.

Итоговые коэффициенты параметра авторегрессии: $B^* = (1.06264930,$

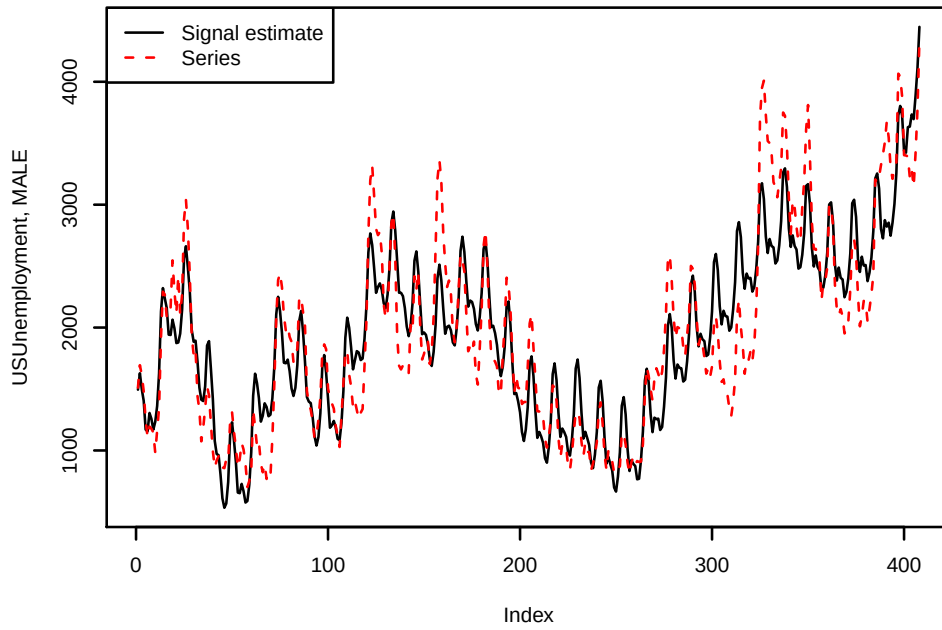


Рис. 5.17. Ряд *US Unemployment, male* и оценка его сигнала алгоритмом 5.4.1.

$0.02823641, -0.200730)^T$; управляющая сигналом ОЛРФ(A^*): $A^* = (-0.009937, 0.058891, -0.155361, 0.25028, -0.297177, 0.307467, -0.308517, 0.306784, -0.301952, 0.29789, -0.297655, 0.301633, -0.296959, 0.250715, -0.154484, 0.0578269, -0.009613)^T$.

Далее, у оценки сигнала были найдены сигнальные корни, было найдено представление ряда вида (1) с точностью до коэффициентов полиномов $p_{m_k}(n)$, а затем их коэффициенты найдены с помощью метода наименьших квадратов [8]. Сигнал представим в виде суммы сезонной компоненты (периодической части, состоящей из слабо модулированных синусоид с периодами $3.996 \approx 4$, $2.396 \approx 2.4$, $5.995 \approx 6$, $11.982 \approx 12$, $2.997 \approx 3$) и медленно меняющегося тренда (сумма двух растущих экспоненциальных сигналов и синусоид с большими периодами 51.695 и 182.413). Соответственно, исходный ряд мы представим как сумму сезонной компоненты, тренда и шума (то есть разности ряда и оценки сигнала). Результат такого разложения представлен на рисунке 5.18.

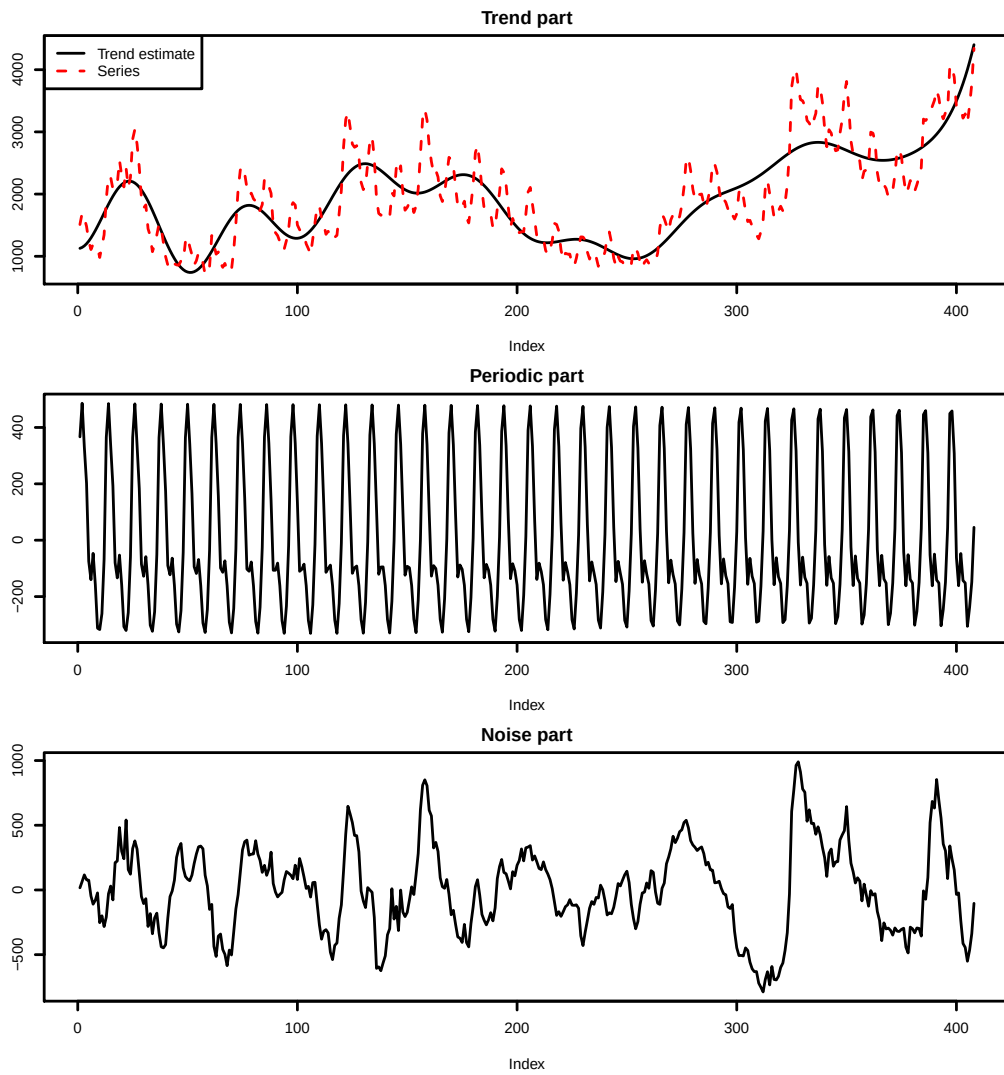


Рис. 5.18. Представление исходного ряда *US Unemployment, male* в виде суммы тренда, сезонной компоненты и шума.

Для того чтобы проверить корректность разложения, нужно удостовериться, что шум приближённо удовлетворяет модели гетероскедастичного процесса авторегрессии. Для этого был вычислен вектор остатков (то есть $\mathbf{C}(\mathbf{X} - \mathbf{Y}^*)$, где $\mathbf{C}^T \mathbf{C} = \mathbf{D}^{-1} \mathbf{\Sigma}^{-1}(B^*, 1) \mathbf{D}^{-1} = \mathbf{W}$ — разложение Холецкого), который должен приближённо представлять из себя вектор белого шума. Для него была вычислена периодограмма и оценка автокорреляционной функции, результаты представлены на рисунке 5.19. Отсутствие острых пиков и сильной закоррелированности подтверждают адекватность представленного разложения. Более того, вектор остатков был проверен с помощью теста на белый шум, предложенный в [67] (метод `whitenoise.test` в пакете `normwhn.test` языка R). Полученное p-value равно 0.3947.

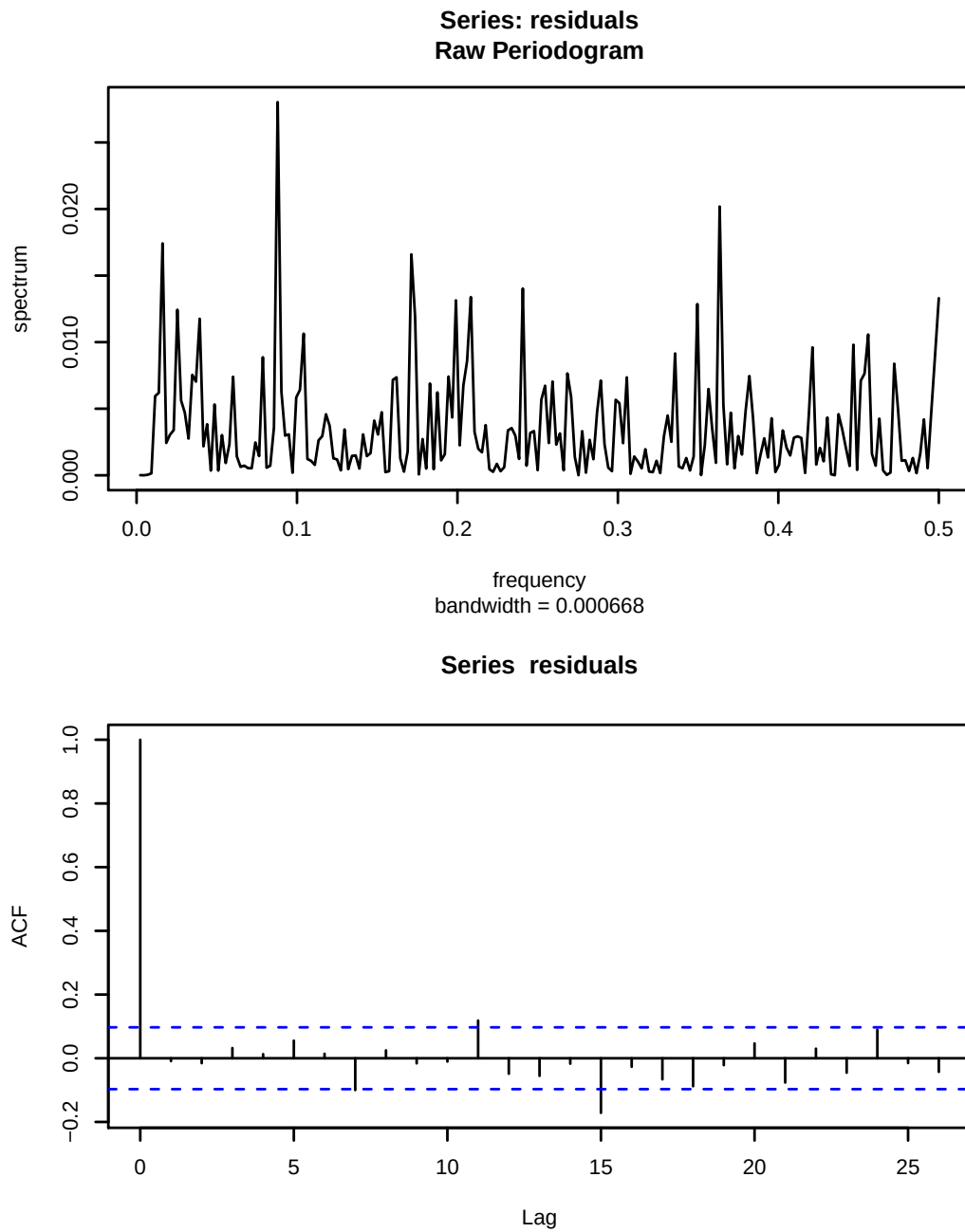


Рис. 5.19. Периодограмма и оценка автокорреляционной функции остатка.

Заключение

В данной работе рассматривались два подхода к решению задачи Hankel structured low-rank approximation (HSLRA) (3) — параметрический и непараметрический. В рамках параметрического подхода была предложена модификация метода Гаусса-Ньютона (MGN, алгоритм 3.3.5) для численного решения задачи (3). Было произведено сравнение с одним из лучших на данный момент алгоритмов локального поиска, основанном на принципе Variable projection (VPGN, алгоритм 3.3.4). Показано, что предложенный алгоритм MGN допускает более устойчивую реализацию в случае сигналов, управляемых ОЛРФ с кратными корнями у характеристического многочлена (в частном случае, для полиномиальных сигналов), см. раздел 5.1.2 и замечание 3.3.2. Сравнение по вычислительной сложности в разделе 3.3.5 показывает, что для случая матрицы весов $\mathbf{W} = \Sigma^{-1}$, соответствующей процессу авторегрессии, предложенный алгоритм MGN имеет существенно более низкую вычислительную сложность, чем VPGN, а в случае $(2p + 1)$ -диагональной матрицы весов $\mathbf{W}^{-1}$ предложенный метод обладает лишь немного большей вычислительной сложностью.

В рамках непараметрического подхода для решению задачи (4) был рассмотрен метод Кэдзоу, входящий в класс алгоритмов попеременных проекций (Alternating Projections). Был рассмотрен вопрос соотношения задач (3) и (4) для $(2p + 1)$ -диагональной матрицы весов $\mathbf{W}$, соответствующей авторегрессионному шуму порядка p . Были сформулированы задачи поиска необходимых матриц весов $\mathbf{L}$, $\mathbf{R}$ (4.7) и (4.8) для приближенного соответствия задач (3) и (4), доказана их эквивалентность в широком классе случаев (теорема 4.2.2) и построены численные методы их решения в разделах 4.3 и 4.4. В разделе 4.5 известная быстрая реализация метода Кэдзоу расширена на случай ленточных матриц весов $\mathbf{L}$ и $\mathbf{R}$. Реализация алгоритма Кэдзоу и корректность формулировки задачи поиска весов (4.7) были проверены с помощью численного мо-

делирования в разделе 5.2.1. Было показано, что при помощи рассмотренных методов можно уменьшить среднеквадратичную ошибку оценивания сигнала по сравнению со стандартным методом Кэдзоу. Более того, было показано, что метод Кэдзоу представляет хорошее начальное приближение для метода MGN.

Полученные асимптотические по соотношению сигнал/шум ошибки первого порядка для оценок сигнала с помощью проекции на множество $\mathcal{D}_r$ рядов ранга r (раздел 2.5) и с помощью линеаризованного алгоритма Кэдзоу (раздел 4.7) подтверждены с помощью статистического моделирования, представленного в разделе 5.2. Было показано, что при увеличении соотношения сигнал/шум полученные ошибки первого порядка становятся ближе к реально наблюдаемым ошибкам оценок.

Полученные алгоритмы были успешно применены к данным экспрессии генов (раздел 5.3), где с помощью предложенного метода MGN была повышена точность оценивания, и к временному ряду с неизвестной ковариационной матрицей шума (раздел 5.4), где удалось построить модель вида «сигнал плюс шум» для сигнала большого ранга и сильного шума сложной формы.

Предложенная теория основывается на исследованных в работе свойствах множества $\mathcal{D}_r$ рядов ранга r , в частности, на виде его касательного многообразия (теорема 2.2.3). Благодаря этому стало возможным построение алгоритмов вычисления базисов $\mathbf{Z}(A)$ и $\mathbf{Z}(A^2)$ пространств рядов, управляемых ОЛРФ(A) (алгоритм 3.2.1) и ОЛРФ(A^2) (алгоритм 3.2.2), используемых в модифицированном методе Гаусса-Ньютона MGN.

Таким образом, в диссертации были предложены новые методы для решения широко используемых на практике задач оценки сигнала и его параметров в классе сигналов конечного ранга. Получено теоретическое обоснование методов и их свойств, построены эффективные (быстрые и устойчивые) реализации. Это позволяет существенно расширить круг решаемых проблем и улучшить точность их решения.

Список литературы

1. Cadzow J. A. Signal enhancement: a composite property mapping algorithm // IEEE Trans. Acoust. 1988. Vol. 36, no. 1. P. 49–62.
2. Markovsky I. Structured low-rank approximation and its applications // Automatica. 2008. apr. Vol. 44, no. 4. P. 891–909.
3. Dendrinos M., Bakamidis S., Carayannis G. Speech enhancement from noise: A regenerative approach // Speech Communication. 1991. feb. Vol. 10, no. 1. P. 45–57.
4. Roy R., Kailath T. ESPRIT-estimation of signal parameters via rotational invariance techniques // IEEE Transactions on acoustics, speech, and signal processing. 1989. Vol. 37, no. 7. P. 984–995.
5. Golyandina N., Zhigljavsky A. Singular Spectrum Analysis for time series. Springer Science & Business Media, 2013.
6. Rouquette S., Najim M. Estimation of frequencies and damping factors by two-dimensional ESPRIT type methods // IEEE Transactions on signal processing. 2001. Vol. 49, no. 1. P. 237–245.
7. Wang Y., Chen J.-W., Liu Z. Comments on “estimation of frequencies and damping factors by two-dimensional ESPRIT type methods” // IEEE transactions on signal processing. 2005. Vol. 53, no. 8. P. 3348–3349.
8. Golyandina N., Nekrutkin V., Zhigljavsky A. Analysis of Time Series Structure: SSA and Related Techniques. Chapman&Hall/CRC, 2001.
9. Alexandrov T., Golyandina N. Automatic extraction and forecast of time series cyclic components within the framework of SSA // Proc. of the 5th St. Petersburg Workshop on Simulation. 2005. P. 45–50.
10. Некруткин В. Аппроксимирующие пространства и продолжения временных рядов // Статистические модели с приложениями в эконометрике и смежных областях/Под ред. П. ред. СМ Ермакова и ЮН Каштанова. СПб.: изд-

во НИИХ СПбГУ. 1999. С. 2–32.

11. Некруткин В. Разложения временных рядов // Главные компоненты временных рядов: метод «Гусеница»/под ред. ДЛ Данилова, АА Жиглявского. СПб.: Пресском. 1997. С. 194–227.
12. 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.
13. Markovsky I. Low Rank Approximation: Algorithms, Implementation, Applications (Communications and Control Engineering). Springer, 2011.
14. Exact and approximate modeling of linear systems: A behavioral approach / Ivan Markovsky, Jan C Willems, Sabine Van Huffel, Bart De Moor. SIAM, 2006. Vol. 11.
15. 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.
16. Usevich K., Markovsky I. Structured low-rank approximation as a rational function minimization // IFAC Proceedings Volumes. 2012. Vol. 45, no. 16. P. 722–727.
17. 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.
18. Allen G. I., Grotenick L., Taylor J. A generalized least-square matrix decomposition // Journal of the American Statistical Association. 2014. Vol. 109, no. 505. P. 145–159.
19. Broomhead D., King G. Extracting qualitative dynamics from experimental data // Physica D. 1986. Vol. 20. P. 217–236.
20. Vautard M., Ghil M. Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series // Physica D. 1989. Vol. 35. P. 395–424.

21. Elsner J. B., Tsonis A. A. Singular Spectrum Analysis: A New Tool in Time Series Analysis. Plenum, 1996.
22. 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.
23. Zhigljavsky A., Golyandina N., Gryaznov S. Deconvolution of a discrete uniform distribution // Stat Probabil Lett. 2016. Vol. 118. P. 37–44.
24. 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.
25. 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.
26. 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.
27. 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.
28. 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.
29. Звонарев Н., Голяндина Н. Итеративные алгоритмы взвешенной аппроксимации рядами конечного ранга // System Identification And Control Problems. SICPRO'15. 2015. P. 1371–1394.
30. Zvonarev N., Golyandina N. Modified Gauss-Newthon method in low-rank signal estimation. arXiv:1803.01419.
31. Hall M. Combinatorial Theory. Wiley-Interscience, 1998.

32. Usevich K. On signal and extraneous roots in singular spectrum analysis // Statistics and Its Interface. 2010. Vol. 3, no. 3. P. 281–295.
33. Heinig G., Rost. Algebraic Methods for Toeplitz-like Matrices and Operators (Operator Theory: Advances and Applications). Birkhäuser Verlag, 1985.
34. Stewart G. On scaled projections and pseudoinverses // Linear Algebra and its Applications. 1989. jan. Vol. 112. P. 189–193.
35. Nocedal J., Wright S. Numerical optimization. Springer Science & Business Media, 2006.
36. Golub G., Pereyra V. Separable nonlinear least squares: the variable projection method and its applications // Inverse problems. 2003. Vol. 19, no. 2. P. R1.
37. Gillard J., Zhigljavsky A. Weighted norms in subspace-based methods for time series analysis // Numerical Linear Algebra with Applications. 2016. Vol. 23, no. 5. P. 947–967.
38. Lewis A. S., Malick J. Alternating projections on manifolds // Mathematics of Operations Research. 2008. Vol. 33, no. 1. P. 216–234.
39. Bounds for the rank of the sum of two matrices : Rep. / Boeing Scientific Research Labs Seattle WA ; Executor: George Marsaglia : 1964.
40. Power Sums, Gorenstein Algebras, and Determinantal Loci / A. Iarrobino, A. Iarrobino, V. Kanev, S.L. Kleiman. Lecture Notes in Mathematics. Springer Berlin Heidelberg, 1999.
41. Shiryaev A. N. Convergence of probability measures. central limit theorem // Probability-1. Springer, 2016. P. 373–460.
42. Kay S. M. Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory (v. 1). Prentice Hall, 1993.
43. Davis P. J. Circulant matrices. American Mathematical Soc., 2012.
44. Graillat S., Ménissier-Morain V. Compensated Horner scheme in complex floating point arithmetic // Proceedings of the 8th Conference on Real Numbers and Computers, Santiago de Compostela, Spain. 2008. P. 133–146.

45. Brent R. P. Algorithms for Minimisation without Derivatives (Automatic Computation). Prentice Hall, 1972.
46. Kiefer J. Sequential minimax search for a maximum // Proceedings of the American Mathematical Society. 1953. mar. Vol. 4, no. 3. P. 502–502.
47. Gillard J., Zhigljavsky A. Optimization challenges in the structured low rank approximation problem // Journal of Global Optimization. 2013. Vol. 57, no. 3. P. 733–751.
48. Chu M. T., Funderlic R. E., Plemmons R. J. Structured low rank approximation // Linear Algebra and its Applications. 2003. Vol. 366, no. 0. P. 157 – 172. Special issue on Structured Matrices: Analysis, Algorithms and Applications.
49. Verbyla A. A note on the inverse covariance matrix of the autoregressive process // Australian & New Zealand Journal of Statistics. 1985. Vol. 27, no. 2. P. 221–224.
50. Kullback S., Leibler R. A. On information and sufficiency // The annals of mathematical statistics. 1951. Vol. 22, no. 1. P. 79–86.
51. Duchi J. Derivations for linear algebra and optimization // Berkeley, California. 2007.
52. Amari S.-i., Nagaoka H. Methods of information geometry, volume 191 of translations of mathematical monographs // American Mathematical Society. 2000. Vol. 13.
53. Theodoridis S. Machine learning: a Bayesian and optimization perspective. Academic Press, 2015.
54. Гавурин М., Малоземов В. Экстремальные задачи с линейными ограничениями. Учебное пособие. 1984.
55. A limited memory algorithm for bound constrained optimization / Richard H. Byrd, Peihuang Lu, Jorge Nocedal, Ciyou Zhu // SIAM Journal on Scientific Computing. 1995. sep. Vol. 16, no. 5. P. 1190–1208.
56. Korobeynikov A. Computation- and space-efficient implementation of SSA //

- Statistics and Its Interface. 2010. Vol. 3, no. 3. P. 357–368.
57. Von Neumann J. Functional Operators: The Geometry of Orthogonal Spaces. Annals of Mathematics Studies. Princeton University Press, 1950.
 58. Meyer C. D. Matrix analysis and applied linear algebra. SIAM: Society for Industrial and Applied Mathematics, 2001.
 59. Optimal rates of convergence of matrices with applications / Heinz H. Bauschke, J. Y. Bello Cruz, Tran T. A. Nghia et al. arXiv : math.OC/1407.0671v1.
 60. Belousov S. L. Tables of Normalized Associated Legendre Polynomials: Mathematical Tables Series. Pergamon, 2014.
 61. Golyandina N. On the choice of parameters in singular spectrum analysis and related subspace-based methods // Stat. Interface. 2010. Vol. 3, no. 3. P. 259–279.
 62. Two-exponential models of gene expression patterns for noisy experimental data / Theodore Alexandrov, Nina Golyandina, David Holloway et al. arXiv:1704.00351v1.
 63. Gardner G., Harvey A. C., Phillips G. D. A. Algorithm AS 154: An algorithm for exact maximum likelihood estimation of autoregressive-moving average models by means of kalman filtering // Applied Statistics. 1980. Vol. 29, no. 3. P. 311.
 64. Andrews D., Herzberg A. Data: a collection of problems from many fields for the student and research worker. Springer series in statistics. Springer-Verlag, 1985.
 65. Golyandina N., Shlemov A. Variations of singular spectrum analysis for separability improvement: non-orthogonal decompositions of time series. arXiv:1308.4022v2.
 66. Schwarz G. Estimating the dimension of a model // The Annals of Statistics. 1978. mar. Vol. 6, no. 2. P. 461–464.
 67. Velasco C., Lobato I. N. et al. A simple and general test for white noise // Econometric Society 2004 Latin American Meetings. No. 112. 2004.