

Задачи анализа временных рядов, теория метода «Анализ Сингулярного Спектра» SSA (Семестр 6)

Яковлев Денис Михайлович, гр.21.Б04-мм

Санкт-Петербургский государственный университет
Прикладная математика и информатика
Вычислительная стохастика и статистические модели

3 курс (бак.) «Производственная практика
(научно-исследовательская работа)»

Санкт-Петербург, 2024

Задачи анализа временных рядов

Задачи анализа временных рядов, теория
метода «Анализ Сингулярного Спектра» SSA
(Семестр 6)

Яковлев Денис Михайлович, гр.21.Б04-мм

Санкт-Петербургский государственный университет
Прикладная математика и информатика
Вычислительная стохастика и статистические модели

3 курс (бак.) «Производственная практика
(научно-исследовательская работа)»

Санкт-Петербург, 2024

Научный руководитель к.ф.-м.н., доцент Некруткин В.В.,
кафедра статистического моделирования

Моя отметка за работу Дениса Яковлева – 5 с минусом (очень хорошо). 5 за результат, минус – за стиль работы.

Задачи анализа временных рядов, теория метода «Анализ Сингулярного Спектра» SSA (Семестр 6)

Яковлев Денис Михайлович, гр.21.Б04-мм

Санкт-Петербургский государственный университет
Прикладная математика и информатика
Вычислительная стохастика и статистические модели

3 курс (бак.) «Производственная практика
(научно-исследовательская работа)»

Санкт-Петербург, 2024

Задачи анализа временных рядов

Задачи анализа временных рядов, теория
метода «Анализ Сингулярного Спектра» SSA
(Семестр 6)

Яковлев Денис Михайлович, гр.21.Б04-мм

Санкт-Петербургский государственный университет
Прикладная математика и информатика
Вычислительная стохастика и статистические модели

3 курс (бак.) «Производственная практика
(научно-исследовательская работа)»

Санкт-Петербург, 2024

Научный руководитель к.ф.-м.н., доцент Некруткин В.В.,
кафедра статистического моделирования

- Цель работы — решение теоретических и прикладных задач анализа временных рядов с применением метода SSA.
- Используемый метод — SSA (Singular Spectrum Analysis), или «Анализ Сингулярного Спектра» (сокращенно, АСС).
- Задача — исследование поведения ошибок восстановления ряда в асимптотическом случае.

Задачи анализа временных рядов

└ Введение

Введение

- Цель работы — решение теоретических и прикладных задач анализа временных рядов с применением метода SSA.
- Используемый метод — SSA (Singular Spectrum Analysis), или «Анализ Сингулярного Спектра» (сокращенно, АСС).
- Задача — исследование поведения ошибок восстановления ряда в асимптотическом случае.

Целью этой работы является решение теоретических и прикладных задач по анализу временных рядов с применением знаний о методе SSA (Singular Spectrum Analysis), или «Анализ Сингулярного Спектра» (сокращенно, АСС). Ознакомиться с методом АСС можно в [1], а в [2] описывается теоретическая часть метода АСС. Общая задача состоит в том, чтобы оценить ошибки восстановления ряда.

Определение 1 (Сигнал)

Вещественный сигнал $H = (h_0, h_1, \dots, h_n, \dots)$, управляется линейной рекуррентной формулой (ЛРФ) порядка d

$$h_n = \sum_{k=1}^d a_k h_{n-k}, \quad n \geq d, \quad a_d > 0.$$

$$H_N = (h_0, h_1, \dots, h_N).$$

Определение 2 (Помеха)

Помеха — $E_N = (e_0, \dots, e_N)$.

Задачи анализа временных рядов

└ Обозначения

Обозначения

Определение 1 (Сигнал)

Вещественный сигнал $H = (h_0, h_1, \dots, h_n, \dots)$, управляется линейной рекуррентной формулой (ЛРФ) порядка d

$$h_n = \sum_{k=1}^d a_k h_{n-k}, \quad n \geq d, \quad a_d > 0.$$

$$H_N = (h_0, h_1, \dots, h_N).$$

Определение 2 (Помеха)

Помеха — $E_N = (e_0, \dots, e_N)$.

Введём необходимые для постановки задачи определения и известные результаты. Начнём с определений. Рассматривается вещественный сигнал H , причем предполагается, что ряд управляется линейной рекуррентной формулой (ЛРФ) порядка d с $a_d > 0$, которая является минимальной в том смысле, что не существует ЛРФ меньшего порядка, управляющей рядом H . Так как нас интересует использование метода SSA, будем рассматривать отрезки длины N . Кроме того, вводится помеха E , некоторый ряд из вещественных значений.

Определение 3 (Траекторная матрица)

Пусть $1 \leq L < N$ — длина окна, $K = N - L + 1$. Тогда траекторной матрицей вещественного сигнала $H_N = (h_0, h_1, \dots, h_N)$ называется матрица \mathbf{H} размера $L \times K$ вида:

$$\begin{pmatrix} h_0 & h_1 & h_2 & \dots & h_{K-1} \\ h_1 & h_2 & h_3 & \dots & h_K \\ h_2 & h_3 & h_4 & \dots & h_{K+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ h_{L-1} & h_L & h_{L+1} & \dots & h_{N-1} \end{pmatrix}$$

Замечание 1 (Размеры матриц)

Считаем, что все рассматриваемые матрицы размера $L \times K$, если не оговорено иначе.

Задачи анализа временных рядов

└ Обозначения

Обозначения

Определение 3 (Траекторная матрица)

Пусть $1 \leq L < N$ — длина окна, $K = N - L + 1$. Тогда траекторной матрицей вещественного сигнала $H_N = (h_0, h_1, \dots, h_N)$ называется матрица \mathbf{H} размера $L \times K$ вида:

$$\begin{pmatrix} h_0 & h_1 & h_2 & \dots & h_{K-1} \\ h_1 & h_2 & h_3 & \dots & h_K \\ h_2 & h_3 & h_4 & \dots & h_{K+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ h_{L-1} & h_L & h_{L+1} & \dots & h_{N-1} \end{pmatrix}$$

Замечание 1 (Размеры матриц)

Считаем, что все рассматриваемые матрицы размера $L \times K$, если не оговорено иначе.

Пользуясь методом SSA, введём траекторные матрицы для сигнала и помехи. Хотя траекторные матрицы не используются в явном виде при решении поставленных задач, само определение должно помочь в понимании последующих определений и результатов.

Замечание 2 (Ранг матрицы)

$d = \text{rank } \mathbf{H} < \min(L, K)$ - ранг траекторной матрицы \mathbf{H} , образованной от сигнала \mathbf{H} , управляемого ЛРФ порядка d .

Определение 4 (Матрица возмущённого сигнала)

Пусть $\delta \in \mathbb{R}$ — формальный параметр возмущения, \mathbf{E} — траекторная матрица \mathbf{E}_N , $\mathbf{H} = \mathbf{H}_N$ соответственно. Тогда $\mathbf{H}(\delta) = \mathbf{H} + \delta \mathbf{E}$ — матрица возмущённого сигнала.

Замечание 3

$\mathbf{H}(0) = \mathbf{H}$.

Задачи анализа временных рядов

└─Обозначения

Обозначения

Замечание 2 (Ранг матрицы)

$d = \text{rank } \mathbf{H} < \min(L, K)$ - ранг траекторной матрицы \mathbf{H} , образованной от сигнала \mathbf{H} , управляемого ЛРФ порядка d .

Определение 4 (Матрица возмущённого сигнала)

Пусть $\delta \in \mathbb{R}$ — формальный параметр возмущения, \mathbf{E} — траекторная матрица \mathbf{E}_N , $\mathbf{H} = \mathbf{H}_N$ соответственно. Тогда $\mathbf{H}(\delta) = \mathbf{H} + \delta \mathbf{E}$ — матрица возмущённого сигнала.

Замечание 3

$\mathbf{H}(0) = \mathbf{H}$.

В отношении введённой траекторной матрицы также необходимо уточнить её ранг, который напрямую связан с порядком линейной рекуррентной формулы d . Теперь рассмотрим формальное определение сигнала с помехой. Пусть формальный параметр возмущения δ вещественный, \mathbf{E} — траекторная матрица помехи, а \mathbf{H} — траекторная матрица сигнала. Тогда $\mathbf{H}(\delta) = \mathbf{H} + \delta \mathbf{E}$ — сигнал с параметризуемой помехой.

Определение 5 (Возмущённая матрица для сингулярного разложения)

$$\mathbf{A}(\delta) = \mathbf{H}(\delta)\mathbf{H}(\delta)^T = \mathbf{H}\mathbf{H}^T + \delta(\mathbf{H}\mathbf{E}^T + \mathbf{E}\mathbf{H}^T) + \delta^2\mathbf{E}\mathbf{E}^T,$$

$\mathbf{A}(\delta)$ — возмущённая матрица размера $L \times L$.

Замечание 4

При $\delta = 0$, $\mathbf{A}(0) = \mathbf{A} = \mathbf{H}\mathbf{H}^T$.

Определение 6 (Наименьшее и наибольшее ненулевые собственные числа)

Пусть $\{\mu_n\}_{n=1}^L$ — собственные числа \mathbf{A} . Тогда μ_{\min}, μ_{\max} — наименьшее и наибольшее ненулевые собственные числа.

Задачи анализа временных рядов

└ Обозначения

Обозначения

Определение 5 (Возмущённая матрица для сингулярного разложения)

$$\mathbf{A}(\delta) = \mathbf{H}(\delta)\mathbf{H}(\delta)^T = \mathbf{H}\mathbf{H}^T + \delta(\mathbf{H}\mathbf{E}^T + \mathbf{E}\mathbf{H}^T) + \delta^2\mathbf{E}\mathbf{E}^T,$$

$\mathbf{A}(\delta)$ — возмущённая матрица размера $L \times L$.

Замечание 4

При $\delta = 0$, $\mathbf{A}(0) = \mathbf{A} = \mathbf{H}\mathbf{H}^T$.

Определение 6 (Наименьшее и наибольшее ненулевые собственные числа)

Пусть $\{\mu_n\}_{n=1}^L$ — собственные числа \mathbf{A} . Тогда μ_{\min}, μ_{\max} — наименьшее и наибольшее ненулевые собственные числа.

Как известно из метода SSA, для отделения сигнала от помехи рассматриваем указанную матрицу. Кроме того, из свойств матрицы \mathbf{A} следует, что все её собственные числа неотрицательны и вещественны.

Определение 7 (Собственные подпространства матрицы A)

Пусть $\{U_n\}_{n=1}^L$ — собственные вектора размера $L \times 1$ матрицы A . Тогда \mathbb{U}_0 — собственное подпространство, соответствующее нулевым собственным числам, а \mathbb{U}_0^\perp — соответствует ненулевым собственным числам.

Определение 8 (Ортогональные проекторы собственных подпространств)

Пусть I — единичная матрица размера $L \times L$. Тогда P_0 — ортогональный проектор на \mathbb{U}_0 , а $P_0^\perp = I - P_0$ — ортогональный проектор на \mathbb{U}_0^\perp . P_0, P_0^\perp — матрицы размера $L \times L$.

Замечание 5

$$P_0^\perp H = H.$$

Задачи анализа временных рядов

— Обозначения

Обозначения

Определение 7 (Собственные подпространства матрицы A)

Пусть $\{U_n\}_{n=1}^L$ — собственные вектора размера $L \times 1$ матрицы A . Тогда \mathbb{U}_0 — собственное подпространство, соответствующее нулевым собственным числам, а \mathbb{U}_0^\perp — соответствует ненулевым собственным числам.

Определение 8 (Ортогональные проекторы собственных подпространств)

Пусть I — единичная матрица размера $L \times L$. Тогда P_0 — ортогональный проектор на \mathbb{U}_0 , а $P_0^\perp = I - P_0$ — ортогональный проектор на \mathbb{U}_0^\perp . P_0, P_0^\perp — матрицы размера $L \times L$.

Замечание 5

$$P_0^\perp H = H.$$

Продолжая разговор про матрицу A , заметим, что её собственные вектора (левые сингулярные векторы) можно разбить на два подпространства: собственное подпространство *ненулевых* и *нулевых* собственных чисел. Более того, для введения и объяснения результатов нам понадобится понятие ортогональных проекторов.

Заметим, что, в связи с таким разбиением проекторов, матрица $H = IH = (P_0^\perp + P_0)H = P_0^\perp H$, что следует из сингулярного разложения матрицы A .

Замечание 6 (Про матричные нормы)

Полагаем $\|\cdot\|$ — спектральная норма. В иных случаях явно указываем норму.

Например, максимальная норма $\|\cdot\|_{\max}$.

Определение 9 (Псевдообратная матрица)

Пусть \mathbf{S}_0 — матрица размера $L \times L$, псевдообратная к $\mathbf{H}\mathbf{H}^T$.

Положим $\mathbf{S}_0^{(0)} = -\mathbf{P}_0$ и $\mathbf{S}_0^{(k)} = \mathbf{S}_0^k$ для $k \geq 1$, $\|\mathbf{S}_0^{(k)}\| = 1/\mu_{\min}^k$.

Определение 10 (Возмущение матрицы $\mathbf{A}(\delta)$)

Зависящую от δ часть матрицы $\mathbf{A}(\delta)$ будем обозначать $\mathbf{B}(\delta)$.

Тогда $\mathbf{B}(\delta) = \delta(\mathbf{H}\mathbf{E}^T + \mathbf{E}\mathbf{H}^T) + \delta^2\mathbf{E}\mathbf{E}^T$.

Задачи анализа временных рядов

└ Обозначения

Обозначения

Замечание 6 (Про матричные нормы)

Полагаем $\|\cdot\|$ — спектральная норма. В иных случаях явно указываем норму.

Например, максимальная норма $\|\cdot\|_{\max}$.

Определение 9 (Псевдообратная матрица)

Пусть \mathbf{S}_0 — матрица размера $L \times L$, псевдообратная к $\mathbf{H}\mathbf{H}^T$.

Положим $\mathbf{S}_0^{(0)} = -\mathbf{P}_0$ и $\mathbf{S}_0^{(k)} = \mathbf{S}_0^k$ для $k \geq 1$, $\|\mathbf{S}_0^{(k)}\| = 1/\mu_{\min}^k$.

Определение 10 (Возмущение матрицы $\mathbf{A}(\delta)$)

Зависящую от δ часть матрицы $\mathbf{A}(\delta)$ будем обозначать $\mathbf{B}(\delta)$.

Тогда $\mathbf{B}(\delta) = \delta(\mathbf{H}\mathbf{E}^T + \mathbf{E}\mathbf{H}^T) + \delta^2\mathbf{E}\mathbf{E}^T$.

В случае, если не указано иное, воспринимаем норму как спектральную. Определим \mathbf{S}_0 — псевдообратную матрицу к $\mathbf{H}\mathbf{H}^T$. Положим $\mathbf{S}_0^{(0)} = -\mathbf{P}_0$ и $\mathbf{S}_0^{(k)} = \mathbf{S}_0^k$ для $k \geq 1$, $\|\mathbf{S}_0^{(k)}\| = 1/\mu_{\min}^k$. Также пригодится для представления результатов возмущение матрицы $\mathbf{A}(\delta)$.

Определение 11 (Ошибка восстановления)

Пусть $\tilde{H}_N(\delta) = (\tilde{h}_0(\delta), \tilde{h}_1(\delta), \dots, \tilde{h}_N(\delta))$ — восстановленный сигнал из $H_N(\delta) = (h_0 + \delta e_0, h_1 + \delta e_1, \dots, h_N + \delta e_N)$, полученный методом SSA.

Тогда $r_i(N) = \tilde{h}_i(\delta) - h_i$, $0 \leq i < N$ — ошибка восстановления.

Определение 12 (Оператор ганкелевизации)

\mathcal{S} — оператор диагонального усреднения (ганкелевизации).

Замечание 7 (Оператор ганкелевизации для траекторной матрицы сигнала)

$$\mathcal{S}H = H.$$

Задачи анализа временных рядов

└ Обозначения

Обозначения

Определение 11 (Ошибка восстановления)

Пусть $\tilde{H}_N(\delta) = (\tilde{h}_0(\delta), \tilde{h}_1(\delta), \dots, \tilde{h}_N(\delta))$ — восстановленный сигнал из $H_N(\delta) = (h_0 + \delta e_0, h_1 + \delta e_1, \dots, h_N + \delta e_N)$, полученный методом SSA.

Тогда $r_i(N) = \tilde{h}_i(\delta) - h_i$, $0 \leq i < N$ — ошибка восстановления.

Определение 12 (Оператор ганкелевизации)

\mathcal{S} — оператор диагонального усреднения (ганкелевизации).

Замечание 7 (Оператор ганкелевизации для траекторной матрицы сигнала)

$\mathcal{S}H = H$.

Итак, перейдём к постановке общей задачи. Для этого введём ошибку восстановления, как разность между восстановленным сигналом и самим сигналом. Восстановленный сигнал получаем из сигнала с параметризуемой помехой методом SSA. Заметим, что по построению траекторной матрицы, она является ганкелевой. Поэтому использование оператора ганкелевизации её не изменит.

Определение 13 (Разность матриц)

$\Delta_\delta(\mathbf{H}) = \mathbf{P}_0^\perp(\delta)\mathbf{H}(\delta) - \mathbf{P}_0^\perp\mathbf{H}$ — разность матриц возмущённого и невозмущённого сигналов.

Постановка 1

Пусть \mathbf{N} — некоторая матрица размера $L \times L$. Так как $\|\mathcal{S}\mathbf{A}\|_{\max} \leq \|\mathbf{A}\|_{\max}$ и $\|\mathbf{A}\|_{\max} \leq \|\mathbf{A}\|$ для любой конечномерной матрицы \mathbf{A} :

$$\begin{aligned} \max_{0 \leq n < N} |r_i(N)| &= \|\mathcal{S}\Delta_\delta(\mathbf{H})\|_{\max} \leq \|\Delta_\delta(\mathbf{H})\|_{\max} \\ &\leq \left\| (\mathbf{P}_0^\perp(\delta) - \mathbf{P}_0^\perp - \mathbf{N})\mathbf{H}(\delta) \right\| + \left\| \mathbf{N}\mathbf{H}(\delta) + \delta\mathbf{P}_0^\perp\mathbf{E} \right\|_{\max}. \end{aligned} \quad (1)$$

Общая задача: подобрать такую \mathbf{N} , чтобы правая часть (1) стремилась к нулю.

Задачи анализа временных рядов

Обозначения и постановка общей задачи

Обозначения и постановка общей задачи

Определение 13 (Разность матриц)

$\Delta_\delta(\mathbf{H}) = \mathbf{P}_0^\perp(\delta)\mathbf{H}(\delta) - \mathbf{P}_0^\perp\mathbf{H}$ — разность матриц возмущённого и невозмущённого сигналов.

Постановка 1

Пусть \mathbf{N} — некоторая матрица размера $L \times L$. Так как $\|\mathcal{S}\mathbf{A}\|_{\max} \leq \|\mathbf{A}\|_{\max}$ и $\|\mathbf{A}\|_{\max} \leq \|\mathbf{A}\|$ для любой конечномерной матрицы \mathbf{A} :

$$\begin{aligned} \max_{0 \leq n < N} |r_i(N)| &= \|\mathcal{S}\Delta_\delta(\mathbf{H})\|_{\max} \leq \|\Delta_\delta(\mathbf{H})\|_{\max} \\ &\leq \left\| (\mathbf{P}_0^\perp(\delta) - \mathbf{P}_0^\perp - \mathbf{N})\mathbf{H}(\delta) \right\| + \left\| \mathbf{N}\mathbf{H}(\delta) + \delta\mathbf{P}_0^\perp\mathbf{E} \right\|_{\max}. \end{aligned} \quad (1)$$

Общая задача: подобрать такую \mathbf{N} , чтобы правая часть (1) стремилась к нулю.

Введём дополнительное обозначение для разности матриц возмущённого и невозмущённого сигналов перед постановкой общей задачи. Общая задача состоит в следующем: подобрать такую матрицу \mathbf{N} , чтобы правая часть стремилась к нулю.

Формулировка 1 (Оценка нормы)

Оценить выражение сверху

$\forall n \in \mathbb{N} : \left\| \mathbf{P}_0^\perp(\delta) - \mathbf{P}_0^\perp - \sum_{p=1}^n \mathbf{W}_p(\delta) \right\|$, $\mathbf{W}_p(\delta)$ — матрицы размера $L \times L$ такие, что

$$\mathbf{W}_p(\delta) = (-1)^p \sum_{l_1 + \dots + l_{p+1} = p, l_j \geq 0} \mathbf{W}_p(l_1, \dots, l_{p+1}; \delta),$$

а

$$\mathbf{W}_p(l_1, \dots, l_{p+1}; \delta) = \mathbf{S}_0^{(l_1)} \mathbf{B}(\delta) \mathbf{S}_0^{(l_2)} \dots \mathbf{S}_0^{(l_p)} \mathbf{B}(\delta) \mathbf{S}_0^{(l_{p+1})}.$$

Задачи анализа временных рядов

└ Формулировка

Формулировка

Формулировка 1 (Оценка нормы)

Оценить выражение сверху

$\forall n \in \mathbb{N} : \left\| \mathbf{P}_0^\perp(\delta) - \mathbf{P}_0^\perp - \sum_{p=1}^n \mathbf{W}_p(\delta) \right\|$, $\mathbf{W}_p(\delta)$ — матрицы размера $L \times L$ такие, что

$$\mathbf{W}_p(\delta) = (-1)^p \sum_{l_1 + \dots + l_{p+1} = p, l_j \geq 0} \mathbf{W}_p(l_1, \dots, l_{p+1}; \delta),$$

а

$$\mathbf{W}_p(l_1, \dots, l_{p+1}; \delta) = \mathbf{S}_0^{(l_1)} \mathbf{B}(\delta) \mathbf{S}_0^{(l_2)} \dots \mathbf{S}_0^{(l_p)} \mathbf{B}(\delta) \mathbf{S}_0^{(l_{p+1})}.$$

Сформулируем первую теоретическую задачу, которую можно увидеть на слайде. В качестве матрицы \mathbf{N} будем рассматривать суммы матриц из $\mathbf{W}_p(\delta)$. Для чего это нужно? Дело в том, что первое слагаемое оценки не обязательно стремится к нулю, поэтому мы хотим подобрать такие матрицы \mathbf{N} , чтобы правая часть стремилась к нулю. Также хотим, чтобы структура \mathbf{N} была простой и понятной в использовании. Следующий известный результат позволит объяснить, почему был выбран именно такая матрица \mathbf{N} .

Воспользуемся вспомогательными теоремами и леммами из [2].

Теорема 1 (Теорема 2.1)

Пусть $\delta_0 > 0$ и

$$\|\mathbf{B}(\delta)\| < \mu_{\min}/2$$

для всех $\delta \in (-\delta_0, \delta_0)$. Тогда для возмущённого проектора $\mathbf{P}_0^\perp(\delta)$ верно представление:

$$\mathbf{P}_0^\perp(\delta) = \mathbf{P}_0^\perp + \sum_{p=1}^{\infty} \mathbf{W}_p(\delta).$$

Задачи анализа временных рядов

└ Известные результаты

Известные результаты

Воспользуемся вспомогательными теоремами и леммами из [2].

Теорема 1 (Теорема 2.1)

Пусть $\delta_0 > 0$ и

$$\|\mathbf{B}(\delta)\| < \mu_{\min}/2$$

для всех $\delta \in (-\delta_0, \delta_0)$. Тогда для возмущённого проектора $\mathbf{P}_0^\perp(\delta)$ верно представление:

$$\mathbf{P}_0^\perp(\delta) = \mathbf{P}_0^\perp + \sum_{p=1}^{\infty} \mathbf{W}_p(\delta).$$

Объясним матрицы $\mathbf{W}_p(\delta)$ на следующем результате. Теорема 1 показывает, что матрицы $\mathbf{W}_p(\delta)$ представляют собой члены разложения бесконечного ряда возмущённого проектора. Поэтому можем сколько угодно близко приближать норму к 0 при ограничениях теоремы 1.

Теорема 2 (Теорема 2.3)

Если $\delta_0 > 0$ и $\frac{\|\mathbf{B}(\delta)\|}{\mu_{\min}} < \frac{1}{4}$ для всех $\delta \in (-\delta_0, \delta_0)$, то проектор $\mathbf{P}_0^\perp(\delta)$ существует и

$$\left\| \mathbf{P}_0^\perp(\delta) - \mathbf{P}_0^\perp \right\| \leq 4C \frac{\|\mathbf{S}_0 \mathbf{B}(\delta) \mathbf{P}_0\|}{1 - 4 \|\mathbf{B}(\delta)\| / \mu_{\min}}$$

где $C = e^{1/6} / \sqrt{\pi} \approx 0.667$.

Лемма 1 (Лемма 6.1)

Если $0 < \beta < 1/4$, $k \geq 0$, то

$$\sum_{p=k}^{\infty} \binom{2p}{p} \beta^p \leq C \frac{(4\beta)^k}{1 - 4\beta}, \quad C = e^{1/6} / \sqrt{\pi}.$$

Задачи анализа временных рядов

└ Известные результаты

Известные результаты

Теорема 2 (Теорема 2.3)

Если $\delta_0 > 0$ и $\frac{\|\mathbf{B}(\delta)\|}{\mu_{\min}} < \frac{1}{4}$ для всех $\delta \in (-\delta_0, \delta_0)$, то проектор $\mathbf{P}_0^\perp(\delta)$ существует и

$$\left\| \mathbf{P}_0^\perp(\delta) - \mathbf{P}_0^\perp \right\| \leq 4C \frac{\|\mathbf{S}_0 \mathbf{B}(\delta) \mathbf{P}_0\|}{1 - 4 \|\mathbf{B}(\delta)\| / \mu_{\min}}$$

где $C = e^{1/6} / \sqrt{\pi} \approx 0.667$.

Лемма 1 (Лемма 6.1)

Если $0 < \beta < 1/4$, $k \geq 0$, то

$$\sum_{p=k}^{\infty} \binom{2p}{p} \beta^p \leq C \frac{(4\beta)^k}{1 - 4\beta}, \quad C = e^{1/6} / \sqrt{\pi}.$$

Опишем результаты: теорема 2 ставит ограничения для возможности оценки нормы разности возмущённого и невозмущённого проекторов, в то время как лемма 1 используется для доказательства как этой теоремы, так и полученного результата.

Следствие 1

В условиях Теоремы 1, $\forall n \in \mathbb{N}$:

$$\left\| \mathbf{P}_0^\perp(\delta) - \mathbf{P}_0^\perp - \sum_{p=1}^n \mathbf{W}_p(\delta) \right\| \leq 4^{n+1} C \left(\frac{\|\mathbf{B}(\delta)\|}{\mu_{\min}} \right)^{n+1} \frac{1}{1 - 4 \|\mathbf{B}(\delta)\| / \mu_{\min}},$$

где $C = e^{1/6} / \sqrt{\pi}$.

Задачи анализа временных рядов

└ Полученные результаты

Пользуясь известными результатами, получаем следующую оценку для произвольного натурального n .

Полученные результаты

Следствие 1

В условиях Теоремы 1, $\forall n \in \mathbb{N}$:

$$\left\| \mathbf{P}_0^\perp(\delta) - \mathbf{P}_0^\perp - \sum_{p=1}^n \mathbf{W}_p(\delta) \right\| \leq 4^{n+1} C \left(\frac{\|\mathbf{B}(\delta)\|}{\mu_{\min}} \right)^{n+1} \frac{1}{1 - 4 \|\mathbf{B}(\delta)\| / \mu_{\min}},$$

где $C = e^{1/6} / \sqrt{\pi}$.

Постановка 2 (Линейный сигнал с гармониками)

Рассмотрим при $n = 0, 1, \dots, N - 1$ линейный сигнал $h_n = \theta_1 n + \theta_0$, где $\theta_1 \neq 0$, и помеху, которая является линейной комбинацией гармоник

$$e_n = \sum_{l=1}^r \tau_l \cos(2\pi n \omega_l + \varphi_l),$$

где $\tau_l \neq 0, \omega_l \neq \omega_p$ при $l \neq p$ и $0 < \omega_l < 1/2$.

Пусть $\mathbf{N} = \mathbf{W}_1$. Тогда при N — нечётном, $L = (N + 1)/2 = K$ верна теорема из [3] о том, что

$$\max_{0 \leq n < N} |r_n(N)| = O(N^{-1}).$$

Задачи анализа временных рядов

Постановка

Постановка

Постановка 2 (Линейный сигнал с гармониками)

Рассмотрим при $n = 0, 1, \dots, N - 1$ линейный сигнал $h_n = \theta_1 n + \theta_0$, где $\theta_1 \neq 0$, и помеху, которая является линейной комбинацией гармоник

$$e_n = \sum_{l=1}^r \tau_l \cos(2\pi n \omega_l + \varphi_l),$$

где $\tau_l \neq 0, \omega_l \neq \omega_p$ при $l \neq p$ и $0 < \omega_l < 1/2$.

Пусть $\mathbf{N} = \mathbf{W}_1$. Тогда при N — нечётном, $L = (N + 1)/2 = K$ верна теорема из [3] о том, что

$$\max_{0 \leq n < N} |r_n(N)| = O(N^{-1}).$$

Теперь рассмотрим постановку теоретической задачи №2, для которой понадобится результат теоретической задачи №1. Рассматривается линейный сигнал с линейной комбинацией гармоник. Верна теорема о том, что когда матрица $\mathbf{N} = \mathbf{W}_1$, длина ряда N — нечётное и длина окна $L = K$, то есть, для квадратных траекторных матриц, максимальная ошибка восстановления ряда может быть оценена сверху некоторой константой, умноженной на N^{-1} . Возникает вопрос: можно ли расширить множество используемых траекторных матриц, не нарушая результат теоремы? Для этого подберём другую матрицу \mathbf{N} .

Формулировка и известные результаты

Формулировка 2

Обобщить результат [3] с $L = K$ до $L/N \rightarrow \alpha \in (0, 1)$ с помощью выбора $\mathbf{N} = \mathbf{W}_1 + \mathbf{W}_2$.

Воспользуемся вспомогательными теоремами и леммами из [3]

Лемма 2 (Лемма 1)

При $N \rightarrow \infty$ имеет место соотношение $\|\mathbf{H}\mathbf{E}^T\|_{\max} = O(N)$.

Лемма 3 (Лемма 2)

При $N \rightarrow \infty$ имеет место соотношение $\|\mathbf{P}_0^\perp \mathbf{E}\|_{\max} = O(N^{-1})$.

Лемма 4 (Лемма 3)

При $N \rightarrow \infty$ имеет место соотношение $\|\mathbf{S}_0 \mathbf{E}\| = O(N^{-4})$.

Задачи анализа временных рядов

└ Формулировка и известные результаты

Формулировка и известные результаты

Формулировка 2

Обобщить результат [3] с $L = K$ до $L/N \rightarrow \alpha \in (0, 1)$ с помощью выбора $\mathbf{N} = \mathbf{W}_1 + \mathbf{W}_2$.

Воспользуемся вспомогательными теоремами и леммами из [3]

Лемма 2 (Лемма 1)

При $N \rightarrow \infty$ имеет место соотношение $\|\mathbf{H}\mathbf{E}^T\|_{\max} = O(N)$.

Лемма 3 (Лемма 2)

При $N \rightarrow \infty$ имеет место соотношение $\|\mathbf{P}_0^\perp \mathbf{E}\|_{\max} = O(N^{-1})$.

Лемма 4 (Лемма 3)

При $N \rightarrow \infty$ имеет место соотношение $\|\mathbf{S}_0 \mathbf{E}\| = O(N^{-4})$.

Теперь сформулируем задачу и приведём известные результаты из [3], которые понадобятся при оценке максимального модуля ошибки восстановления ряда

Замечание 8

При $\mathbf{N} = \mathbf{W}_1 + \mathbf{W}_2$ рассматривается неравенство

$$\max_{0 \leq i < N} |r_i(N)| \leq \left\| (\mathbf{P}_0^\perp(\delta) - \mathbf{P}_0^\perp - (\mathbf{W}_1 + \mathbf{W}_2)) \mathbf{H}(\delta) \right\| + \left\| (\mathbf{W}_1 + \mathbf{W}_2) \mathbf{H}(\delta) + \delta \mathbf{P}_0^\perp \mathbf{E} \right\|_{\max}.$$

Предложение 1

Пусть $L/N \rightarrow \alpha \in (0, 1)$. Тогда для любого δ

$$\left\| \left(\mathbf{P}_0^\perp(\delta) - \mathbf{P}_0^\perp - \mathbf{W}_1(\delta) - \mathbf{W}_2(\delta) \right) \mathbf{H}(\delta) \right\| = O(N^{-1}).$$

Предложение 2

В условиях Предложения 1

$$\left\| (\mathbf{W}_1(\delta) + \mathbf{W}_2(\delta)) \mathbf{H}(\delta) \right\|_{\max} = O(N^{-1})$$

Задачи анализа временных рядов

└ Полученные результаты

Полученные результаты

Замечание 8

При $\mathbf{N} = \mathbf{W}_1 + \mathbf{W}_2$ рассматривается неравенство

$$\max_{0 \leq i < N} |r_i(N)| \leq \left\| (\mathbf{P}_0^\perp(\delta) - \mathbf{P}_0^\perp - (\mathbf{W}_1 + \mathbf{W}_2)) \mathbf{H}(\delta) \right\| + \left\| (\mathbf{W}_1 + \mathbf{W}_2) \mathbf{H}(\delta) + \delta \mathbf{P}_0^\perp \mathbf{E} \right\|_{\max}.$$

Предложение 1

Пусть $L/N \rightarrow \alpha \in (0, 1)$. Тогда для любого δ

$$\left\| (\mathbf{P}_0^\perp(\delta) - \mathbf{P}_0^\perp - \mathbf{W}_1(\delta) - \mathbf{W}_2(\delta)) \mathbf{H}(\delta) \right\| = O(N^{-1}).$$

Предложение 2

В условиях Предложения 1

$$\left\| (\mathbf{W}_1(\delta) + \mathbf{W}_2(\delta)) \mathbf{H}(\delta) \right\|_{\max} = O(N^{-1})$$

В качестве полученных результатов выделим Предложение 1 и Предложение 2, которые являются переформулировками предложений из [3] для рассматриваемого случая обобщения. В результате получим обобщение результата с $L = K$ до $L/N \rightarrow \alpha \in (0, 1)$.

Формулировка 3

На основе теоретического результата задачи №2 проиллюстрировать результат для ряда

$$\tilde{h}_n = n + 3 \cos(2\pi n\omega + \varphi),$$

где $\omega = 1/4$, $\varphi = \pi/8$, $n = 0, 1, \dots, N-1$, $N = 9 \dots 200$, длина окна $L = \lfloor N/3 \rfloor$.

Задачи анализа временных рядов

└ Приложение. Результаты вычислительных экспериментов

Приложение. Результаты вычислительных экспериментов

Формулировка 3
На основе теоретического результата задачи №2 проиллюстрировать результат для ряда
$$\tilde{h}_n = n + 3 \cos(2\pi n\omega + \varphi),$$

где $\omega = 1/4$, $\varphi = \pi/8$, $n = 0, 1, \dots, N-1$, $N = 9 \dots 200$, длина окна $L = \lfloor N/3 \rfloor$.

Для подтверждения результатов теоретической задачи №2 проведём вычислительный эксперимент. Рассмотрим следующий ряд при длине окна L приблизительно равной длине ряда N , делённой на 3. Проиллюстрируем результаты.



Рис. 1: Максимальные ошибки восстановления ряда в зависимости от длины ряда при $\tilde{h}_n = n + 3 \cos(\pi n/2 + \pi/8)$.

Задачи анализа временных рядов

└ Приложение. Результаты вычислительных экспериментов

Приложение. Результаты вычислительных экспериментов

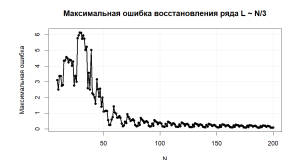


Рис. 1: Максимальные ошибки восстановления ряда в зависимости от длины ряда при $\tilde{h}_n = n + 3 \cos(\pi n/2 + \pi/8)$.

Прокомментируем результаты вычислительного эксперимента: по рисунку 1 убедимся, что максимальные по модулю ошибки восстановления ряда стремятся к нулю с ростом N .



Рис. 2: Максимальные ошибки восстановления ряда, умноженные на N , в зависимости от длины ряда N для $x_n = n + 3 \cos(\pi n/2 + \pi/8)$.

Задачи анализа временных рядов

└ Приложение. Результаты вычислительных экспериментов

Приложение. Результаты вычислительных экспериментов

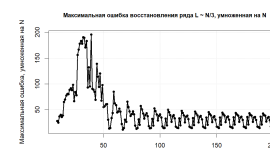


Рис. 2: Максимальные ошибки восстановления ряда, умноженные на N , в зависимости от длины ряда N для $x_n = n + 3 \cos(\pi n/2 + \pi/8)$.

Рисунок 2 демонстрирует, что после умножения ряда рисунка 1 на N максимальные по модулю ошибки восстановления ряда становятся ограниченными, что подтверждает результат обобщения Теоремы из [3].

- 1 Поставлена общая теоретическая задача;
- 2 Дана оценка $\|\mathbf{P}_0^\perp(\delta) - \mathbf{P}_0^\perp - \sum_{p=1}^n \mathbf{W}_p(\delta)\| \forall n \in \mathbb{N}$;
- 3 Обобщён случай асимптотической разделимости линейного сигнала с линейной комбинацией гармоник с $L = K$ до $L/N \rightarrow \alpha \in (0, 1)$;
- 4 Проведан вычислительный эксперимент, подтверждающий результаты обобщения теоремы из [3].




Задачи анализа временных рядов

└ Заключение

Заключение

- Поставлена общая теоретическая задача;
- Дана оценка $\|\mathbf{P}_0^\perp(\delta) - \mathbf{P}_0^\perp - \sum_{p=1}^n \mathbf{W}_p(\delta)\| \forall n \in \mathbb{N}$;
- Обобщён случай асимптотической разделимости линейного сигнала с линейной комбинацией гармоник с $L = K$ до $L/N \rightarrow \alpha \in (0, 1)$;
- Проведан вычислительный эксперимент, подтверждающий результаты обобщения теоремы из [3].




В ходе проделанных работ были изучены теоретические свойства метода SSA, поставлена общая теоретическая задача, дана оценка **нормы разностей ортогональных проекций с дубль-вэ для произвольного n** , обобщён случай асимптотической разделимости линейного сигнала с линейной комбинацией гармоник с $L = K$ до $L/N \rightarrow \alpha \in (0, 1)$, а также проделан вычислительный эксперимент, подтверждающий результаты обобщения теоремы из [3].

-  GOLYANDINA, N., NEKRUTKIN, V. AND ZHIGLJAVSKY, A. (2001). *Analysis of Time Series Structure. SSA and Related Techniques*. Champan & Hall/CRC, Boca Raton-London-New York-Washington D.C.
-  NEKRUTKIN, V. (2010). *Perturbation expansions of signal subspaces for long signals*. *Statistics and Its Interface*. **3**, 297–319.
-  Н.Зенкова, В.Некруткин. Об асимптотической разделимости линейных сигналов с гармониками методом анализа сингулярного спектра, Вестник СПбГУ. Математика. Механика. Астрономия. 2022. Т. 9 (67). Вып. 2.

Задачи анализа временных рядов

└ Список литературы

Список литературы

-  GOLYANDINA, N., NEKRUTKIN, V. AND ZHIGLJAVSKY, A. (2001). *Analysis of Time Series Structure. SSA and Related Techniques*. Champan & Hall/CRC, Boca Raton-London-New York-Washington D.C.
-  NEKRUTKIN, V. (2010). *Perturbation expansions of signal subspaces for long signals*. *Statistics and Its Interface*. **3**, 297–319.
-  Н.Зенкова, В.Некруткин. Об асимптотической разделимости линейных сигналов с гармониками методом анализа сингулярного спектра, Вестник СПбГУ. Математика. Механика. Астрономия. 2022. Т. 9 (67). Вып. 2.

На данном слайде представлен список основных источников, используемых в моей работе. Спасибо за внимание. Я готов ответить на Ваши вопросы.