Несколько задач, связанных с анализом временных рядов

Пимахов Кирилл Юрьевич, гр. 16.М03-мм

Санкт-Петербургский государственный университет
Прикладная математика и информатика
Статистическое моделирование

Научный руководитель: к.ф.-м.н., доц. Некруткин В.В.
Рецензент: к.ф.-м.н., Пепелышев А.Н.

Санкт-Петербург
2018г.

Basic SSA: задача восстановления сигнала

- Ряд (сигнал): $F_N = (f_0, f_1, \ldots, f_{N-1})$, задаваемый минимальной рекуррентной формулой

$$f_n = \sum_{k=1}^{d} b_k f_{n-k}, \quad d \leqslant n \leqslant N.$$

- Помеха: $\mathrm{E}_N = (e_0, e_1, \ldots, e_{N-1})$.
- Наблюдаемый ряд: $\mathrm{F}_N(\delta) = \mathrm{F}_N + \delta \mathrm{E}_N$.

Цель — оценить сигнал F_N.

Basic SSA: траекторная матрица

- Траекторная матрица $L \times K$ ряда $F_N = (f_0, f_1, \ldots, f_{N-1})$, $L + K = N + 1$:

$$\mathbf{H} = \begin{pmatrix} f_0 & f_1 & \cdots & f_{K-1} \\ f_1 & \ddots & & \vdots \\ \vdots & & \ddots & f_{L+K-2} \\ f_{L-1} & \cdots & f_{L+K-2} & f_{L+K-1} \end{pmatrix}.$$

- d — порядок рекуррентной формулы, задающей F_N. Если $d \leq \max(L, K)$, то $\operatorname{rank} \mathbf{H} = d$.
- $\mathbf{E}$ — аналогичная траекторная матрица помехи.

- $\widehat{\mathbf{H}}$ — сумма d главных элементарных матриц сингулярного разложения $\mathbf{H}(\delta) = \mathbf{H} + \delta\mathbf{E}$.
- $\widehat{\mathrm{F}}_N(\delta) = \mathcal{S}\widehat{\mathbf{H}}$ — восстановленный сигнал (диагональное усреднение).
- Ошибка восстановления траекторной матрицы:

$$\Delta_\delta(\mathbf{H}) = \widehat{\mathbf{H}} - \mathbf{H}.$$

- Максимальная ошибка восстановления исходного сигнала:

$$\|\widehat{\mathrm{F}}_N(\delta) - \mathrm{F}_N\|_{\max} = \max_{0\le i\le N-1} |\widehat{f}_i(\delta) - f_i| = \\ = \max_{0\le i\le N-1} \left|\mathcal{S}\big(\Delta_\delta(\mathbf{H})\big)_{[i]}\right|.$$

Формализация задачи: обозначения

- $\mathbf{H}$ — траекторная матрица сигнала F_N, $\operatorname{rank}\mathbf{H} = d$.
- $\mathbf{H}(\delta)$ — траекторная матрица возмущенного сигнала $\mathrm{F}_N(\delta)$.
- $\mathbf{A} = \mathbf{H}\mathbf{H}^{\mathrm{T}}, \quad \mathbf{A}(\delta) = \mathbf{H}(\delta)\mathbf{H}^{\mathrm{T}}(\delta)$.
- $\mathbb{U}_0^{\perp}$ — подпространство, порожденное собственными векторами $\mathbf{A}$, соответствующими ненулевым собственным числам (размерность d).
- $\mathbb{U}_0^{\perp}(\delta)$ — подпространство, образованное d главными собственными векторами матрицы $\mathbf{A}(\delta)$.
- $\mathbf{P}_0^{\perp}$ и $\mathbf{P}_0^{\perp}(\delta)$ — соответственно операторы ортогонального проектирования на $\mathbb{U}_0^{\perp}$ и $\mathbb{U}_0^{\perp}(\delta)$.

Формализация задачи

В. В. Некруткин (SII, 2010).

Ошибка восстановления траекторной матрицы:

$$\Delta_\delta(\mathbf{H}) = \widehat{\mathbf{H}} - \mathbf{H} = \left(\mathbf{P}_0^\perp(\delta) - \mathbf{P}_0^\perp\right)\mathbf{H}(\delta) + \delta\mathbf{P}_0^\perp\mathbf{E}.$$

Нас интересует близость исходного и восстановленного сигналов при увеличении длины ряда $N \to \infty$.

Точность аппроксимации сигнала F_N зависит от $\|\mathbf{P}_0^\perp(\delta) - \mathbf{P}_0^\perp\|$.

Цель работы и известные результаты

- Сигнал:
$$f_n = a^{nT(N)/N}, \quad a > 1$$
и помеха:
$$e_n = b^{nT(N)/N} \cos(\xi n + \phi), \quad b > 1, \quad n = 0, 1, \ldots, N.$$
Будут рассмотрены случаи $T(N) = N$ и $T(N) = T_0$.
- Задача: исследовать максимальную ошибку восстановления исходного ряда F_N при $N \to \infty$ для различных вариантов $T(N)$.
- Ранее (Ivanova, Nekrutkin, 2017): при $b = 1$, $L/N \to \alpha \in (0, 1)$ и $T(N) = N$ максимальная ошибка восстановления не стремится к нулю, а при $T(N) = T_0$ имеет порядок $O(N^{-1})$.

Примеры рядов при $T(N) = 10$ и различных N

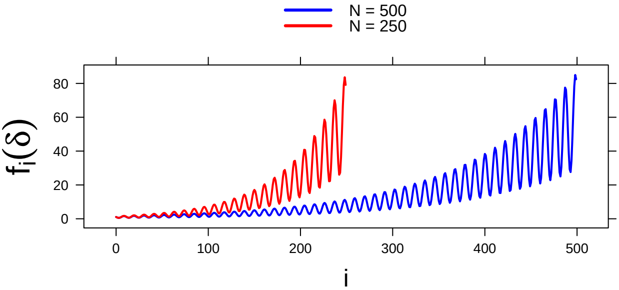

Рис.: Примеры рядов $F_N(\delta)$ при $T(N) = 10$, $a = 1.5$, $b = 1.5$, $\xi = \sqrt{3}/3$, $\phi = \sqrt{2}$, $\delta = 0.5$ и $L = \lfloor (N+1)/2 \rfloor$.

Основные результаты

Доказано, что в случае $L/N \to \alpha$, $\alpha \in (0,1)$

- при $T(N) = N$ и $1 < b^2 < a$ максимальная ошибка восстановления $r_{\max}(N)$, вообще говоря, не стремится к нулю, сделано предположение о порядке асимптотики ошибки: $r_{\max}(N) = O(b^N)$;
- при $T(N) = T_0$ для любого соотношения $a, b > 1$ $r_{\max}(N) = O(N^{-1})$;
- при $T(N) = N$ и $b = a > 1$ норма разности проекторов $\|\mathbf{P}_0^{\perp}(\delta) - \mathbf{P}_0^{\perp}\|$ не стремится к нулю.

  Численные эксперименты: $r_{\max}(N) \nrightarrow 0$.

Иллюстрации: максимальная ошибка восстановления при $1 < b^2 < a$ и $T(N) = N$

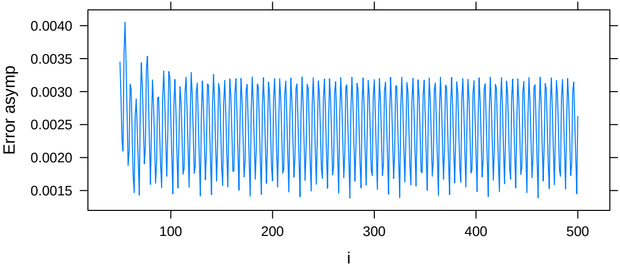

Рис.: Поведение $b^{-N} \max_i |\hat{f}_i - f_i|$ при $T(N) = N$, $a = 1.05$, $b = 1.01$, $\xi = \sqrt{3}/3$, $\phi = \sqrt{2}$, $\delta = 0.01$ и $L = \lfloor (N+1)/2 \rfloor$.

Иллюстрации: максимальная ошибка восстановления при $b = a$ и $T(N) = N$

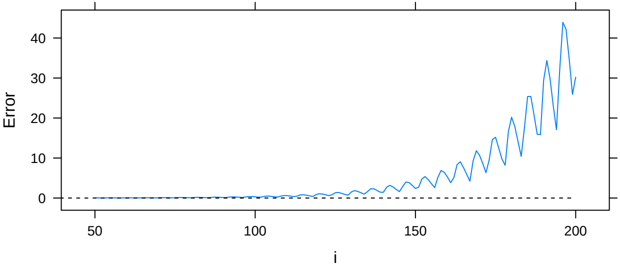

Рис.: Поведение максимальной ошибки восстановления при $T(N) = N$, $a = b = 1.05$, $\xi = \sqrt{3}/3$, $\phi = \sqrt{2}$, $\delta = 0.01$ и $L = \lfloor (N+1)/2 \rfloor$.

Использованные результаты.

В. В. Некруткин (SII, 2010).

$\mathbf{B}(\delta) = \mathbf{A}(\delta) - \mathbf{A} = \mathbf{H}(\delta)\mathbf{H}(\delta)^{\mathrm{T}} - \mathbf{H}\mathbf{H}^{\mathrm{T}}$.

$\mu_{\min}$ — минимальное положительное собственное число $\mathbf{H}\mathbf{H}^{\mathrm{T}}$.

Если $\exists \delta_0$, что при $|\delta| < \delta_0$ $\|\mathbf{B}(\delta)\|/\mu_{\min} < 1/4$, то

$$\|\mathbf{P}_0^{\perp}(\delta) - \mathbf{P}_0^{\perp} - \mathbf{L}(\delta)\| \le 16\, C\, \frac{\|\mathbf{S}_0\mathbf{B}(\delta)\|^2}{1 - 4\|\mathbf{B}(\delta)\|/\mu_{\min}},$$

$$\|\mathbf{P}_0^{\perp}(\delta) - \mathbf{P}_0^{\perp} - \mathbf{W}_1(\delta)\| \le 16C \left(\frac{\|\mathbf{B}(\delta)\|}{\mu_{\min}}\right)^2 \frac{1}{1 - 4\|\mathbf{B}(\delta)\|/\mu_{\min}},$$

где $\mathbf{L}(\delta)$, $\mathbf{W}_1(\delta)$ — некоторые матрицы, выражающиеся через сигнал и помеху, а $\mathbf{S}_0$ — псевдообратная к $\mathbf{H}\mathbf{H}^{\mathrm{T}}$.

Техника исследования максимальной ошибки восстановления: $T(N) = N$

Доказательство отсутствия сходимости к нулю в случае $T(N) = N$ и $1 < b^2 < a$:

- Используется представление

$$\Delta_\delta(\mathbf{H}) = \left(\mathbf{P}_0^\perp(\delta) - \mathbf{P}_0^\perp - \delta\,\mathbf{V}_0^{(1)}\right)\mathbf{H}(\delta) + \delta\,\mathbf{V}_0^{(1)}\mathbf{H}(\delta) + \delta\,\mathbf{P}_0^\perp\mathbf{E}.$$

- Доказывается, что $\left\|\left(\mathbf{P}_0^\perp(\delta) - \mathbf{P}_0^\perp - \delta\,\mathbf{V}_0^{(1)}\right)\mathbf{H}(\delta)\right\| \to 0$;
- Исследуется нижний правый элемент r_{LK} матрицы $\mathbf{V}_0^{(1)}\mathbf{H}(\delta) + \mathbf{P}_0^\perp\mathbf{E}$;
- $|r_{LK}|$ имеет порядок b^N в том смысле, что $\limsup_N b^{-N}|r_{LK}| = c_{r_\infty}$ и, вообще говоря, $c_{r_\infty} \neq 0$.

Техника исследования максимальной ошибки восстановления: $T = T_0$

Доказательство сходимости к нулю в случае $T(N) = T_0$:

- Используется представление

$$\Delta_\delta(\mathbf{H}) = \left(\mathbf{P}_0^\perp(\delta) - \mathbf{P}_0^\perp - \mathbf{L}(\delta)\right)\mathbf{H}(\delta) + \mathbf{L}(\delta)\mathbf{H}(\delta) + \delta\,\mathbf{P}_\mathbf{0}^\perp\mathbf{E}.$$

- Доказывается, что $\left\|\left(\mathbf{P}_0^\perp(\delta) - \mathbf{P}_0^\perp - \mathbf{L}(\delta)\right)\mathbf{H}(\delta)\right\| = O(1/N)$;
- Доказывается, что максимальный по модулю элемент матрицы $\mathbf{L}(\delta)\mathbf{H}(\delta) + \delta\,\mathbf{P}_\mathbf{0}^\perp\mathbf{E}$ имеет вид $O(1/N)$.

  Ход доказательства — такой же, как в Ivanova, Nekrutkin, 2017.

Заключение

В исследуемой модели с сигналом $f_n = a^{nT(N)/N}$ с $a > 1$ и помехой $e_n = b^{nT(N)/N} \cos(\xi n + \phi)$ с $b > 1$:

- Показано, что при $T(N) = N$ и $1 < b^2 < a$ максимальная ошибка восстановления, вообще говоря, не стремится к нулю.
- Доказано, что при $T(N) = N$ и при $a = b > 1$ норма разности проекторов не стремится к нулю, а принимая во внимание

  $$\Delta_\delta(\mathbf{H}) = \widehat{\mathbf{H}} - \mathbf{H} = \left(\mathbf{P}_0^\perp(\delta) - \mathbf{P}_0^\perp\right)\mathbf{H}(\delta) + \delta\mathbf{P}_0^\perp\mathbf{E}$$

  и численные эксперименты, можно предположить, что и максимальная ошибка восстановления не стремится к нулю.
- Доказано, что при $T(N) = T_0$ и любых $a, b > 1$ максимальная ошибка восстановления стремится к нулю и имеет порядок $O(N^{-1})$.