

Аппроксимация временных рядов рядами конечного ранга

Звонарев Никита Константинович

Санкт-Петербургский государственный университет
Кафедра статистического моделирования

Научный руководитель: к.ф.-м.н., доц. Голяндина Н. Э.

Рецензент: к.ф.-м.н., доц. Коробейников А. И.



Санкт-Петербург
2015г.

Постановка задачи

Дан $\mathbb{X} = \mathbb{S} + \mathbb{N}$, $\mathbb{X} = (x_1, \dots, x_N)$.

Требуется построить оценку сигнала.

\mathbb{N} — белый шум.

\mathbb{S} — сигнал, управляемый *линейной рекуррентной формулой*

Определение

Ряд, управляемый ЛРФ порядка r :

$$\mathbb{S} = (s_1, \dots, s_N), \quad s_n = \sum_{i=1}^r a_i s_{n-i}, \quad n = r+1, \dots, N, \quad a_r \neq 0.$$

Параметрический вид рядов, управляемых ЛРФ:

$$s_n = \sum_i P_i(n) \exp(\alpha_i n) \cos(2\pi\omega_i n + \psi_i).$$

Параметрический подход не работает из-за неустойчивости

Ряды конечного ранга

$\mathbb{X} = (x_1, \dots, x_N)$ — ряд.

$1 < L < N$ — длина окна, $K = N - L + 1$.

Траекторная матрица ряда \mathbb{X} :

$$\mathbf{X} = \mathcal{T}(\mathbb{X}) = \begin{pmatrix} x_1 & x_2 & \dots & x_K \\ x_2 & x_3 & \dots & x_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_L & x_{L+1} & \dots & x_N \end{pmatrix}.$$

Заметим, что \mathbf{X} — ганкелева ($\mathbf{X} \in \mathcal{H}$).

Определение

Ряд \mathbb{S} имеет L -ранг r если $\text{rank } \mathbf{S} = r$, где $\mathbf{S} = \mathcal{T}(\mathbb{S})$.

Ряд \mathbb{S} управляется минимальной ЛРФ порядка $r \Rightarrow \mathbb{S}$ имеет L -ранг r .

$\mathbf{X}_N^r = \{\mathbb{X} : \mathbb{X} = (x_1, \dots, x_N), \mathbb{X} \text{ имеет } L\text{-ранг } \leq r\}$

Эквивалентные задачи

$$\mathbb{X} = (x_1, \dots, x_N), \mathbf{X} = \mathcal{T}(\mathbb{X}).$$

Задача аппроксимации матрицы:

$$\|\mathbf{X} - \mathbf{Y}\|_F^2 \rightarrow \min_{\mathbf{Y} \in \mathcal{M}_r \cap \mathcal{H}}, \quad (1)$$

где \mathcal{M}_r — множество матриц ранга $\leq r$.

$\hat{\mathbb{S}} = \mathcal{T}^{-1}(\mathbf{Y})$ — оценка \mathbb{S} .

Задача аппроксимации рядов (эквивалентная (1)):

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

где

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

трапециевидные веса.

Задача: как сделать веса q_i равными?

План работы

I часть Теория

1. Метод решения задачи аппроксимации в общем случае
2. Применение теории к аппроксимации рядами конечного ранга
3. Выбор норм, соответствие между весами для аппроксимации рядов и матриц
4. Нахождение проекторов в выбранных нормах

II часть Алгоритмы и их особенности

III часть Численные эксперименты

1. Сравнение на модельном примере
2. Реальный пример

Общая схема итераций

X — гильбертово пространство

\mathcal{H} — замкнутое (топологически) линейное подпространство.

\mathcal{M} — замкнутое относительно умножения на константу подмножество, т.е. если $\mathbf{z} \in \mathcal{M}$, то $\alpha \mathbf{z} \in \mathcal{M}$ для любого $\alpha \in \mathbb{R}$.

Требуется спроектировать заданную точку \mathbf{x} на $\mathcal{H} \cap \mathcal{M}$:

$$\|\mathbf{x} - \mathbf{y}\| \rightarrow \min_{\mathbf{y}}, \text{ где } \mathbf{y} \in \mathcal{H} \cap \mathcal{M}$$

Метод переменных проекций (Alternating projections, AP):

$\mathbf{y}_{k+1} = \Pi_{\mathcal{H}} \Pi_{\mathcal{M}} \mathbf{y}_k$, где $\mathbf{y}_0 = \mathbf{x}$, $\Pi_{\mathcal{M}}$ — проектор на \mathcal{M} , $\Pi_{\mathcal{H}}$ — проектор на \mathcal{H} .

Основная теорема

Один шаг алгоритма AP: $\mathbf{y}_{k+1} = \Pi_{\mathcal{H}}\Pi_{\mathcal{M}}\mathbf{y}_k$, где $\mathbf{y}_0 = \mathbf{x}$.

Теорема

Пусть X — гильбертово пространство, $M \subset X$ — замкнутое относительно умножения на константу подпространство, множество M и подпространство \mathcal{H} топологически замкнуты, Π_M — проектор на M . Тогда

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

Обратно к задаче

Ряды: $\sum_{i=1}^N q_i (x_i - y_i)^2 \rightarrow \min, \mathbf{Y} \in \mathcal{X}_N^r$, заданы $q_i > 0$.

Матрицы: $\mathbf{Y}: \|\mathbf{X} - \mathbf{Y}\|_? \rightarrow \min, \mathbf{Y} \in \mathcal{H} \cap \mathcal{M}_r$. Так как q_i задаются, то требуется другая (не фробениусовская) норма.

Применяем алгоритм к нашему случаю, и получаем решение. Переменные проекции (метод [Cadzow](#)):

$$\mathbf{Y}_R = (\Pi_{\mathcal{H}} \circ \Pi_{\mathcal{M}_r})^R(\mathbf{X}),$$

$\Pi_{\mathcal{H}}$ — проектор на ганкелевы матрицы, $\Pi_{\mathcal{M}_r}$ — проектор на матрицы ранга $\leq r$, R — число итераций.

Теорема применима к задаче \Rightarrow существование сходящейся к нужному множеству подпоследовательности.

Варианты норм

Ряды: $\|\mathbb{X}\|_q^2 = \sum_{i=1}^N q_i x_i^2$, все $q_i \geq 0$

Матрицы:

- 1 $\|\mathbf{X}\|_{1,M}^2 = \sum_{l=1}^L \sum_{k=1}^K m_{l,k} (x_{l,k})^2$, все $m_{l,k} \geq 0$
- 2 Частный случай предыдущего пункта: $m_{l,k} = 1$.
- 3 $\|\mathbf{X}\|_{2,C}^2 = \text{tr}(\mathbf{X}\mathbf{C}\mathbf{X}^T)$, \mathbf{C} — симметричная, неотрицательно определенная матрица порядка $K \times K$ (+ диагональная в нашем случае)

Проблемы:

- 1 Как подобрать $m_{i,j}$, соответствующие q_i ?
- 2 Как вычислить $\Pi_{\mathcal{H}}$, $\Pi_{\mathcal{M}_r}$ по нужной норме?

Эквивалентность скалярных произведений и норм

- 1 $\langle \mathbb{X}, \mathbb{Y} \rangle_q = \sum_{i=1}^N q_i x_i y_i$
- 2 $\langle \mathbf{X}, \mathbf{Y} \rangle_{1,M} = \sum_{l=1}^L \sum_{k=1}^K m_{l,k} x_{l,k} y_{l,k}$
- 3 $\langle \mathbf{X}, \mathbf{Y} \rangle_{2,C} = \text{tr}(\mathbf{X} \mathbf{C} \mathbf{Y}^T).$

Утверждение

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

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

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

$$m_{l,k} = c_{k,k}.$$

Проектор $\Pi_{\mathcal{M}_r}$

Варианты:

- 1 $\|\mathbf{X}\|_{1,M}$, все $m_{l,k} = 1$: через стандартное SVD-разложение.
 $\Pi_{\mathcal{M}_r}(\mathbf{X}) = \mathbf{U}\Sigma_r\mathbf{V}^T$, где $\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^T$.
- 2 $\|\mathbf{X}\|_{2,C}$: косоугольное SVD-разложение. Сводится к п. 1
- 3 $\|\mathbf{X}\|_{1,M}$, общий случай: EM-подобный алгоритм.

Проектор $\Pi_{\mathcal{H}}$

Общий случай (взвешенное диагональное усреднение):

$$\mathbf{X} = \Pi_{\mathcal{H}}(\mathbf{Y}), \quad x_{l,k} = \frac{\sum_{i,j: i+j=l+k} m_{i,j} y_{i,j}}{\sum_{i,j: i+j=l+k} m_{i,j}}.$$

Варианты:

- 1 $\|\mathbf{X}\|_{1,M}$, $m_{l,k}$ на побочных диагоналях равны, в частности $m_{l,k} = 1$: диагональное усреднение.

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

- 2 $\|\mathbf{X}\|_{2,C}$, C — диагональная: взвешенное диагональное усреднение с весами $m_{i,j} = c_{i,i}$, где $C = (c_{i,j})$.

Варианты алгоритма Cadzow

$$\|\mathbf{X}\|_{2,C}^2 = \text{tr}(\mathbf{X}\mathbf{C}\mathbf{X}^T)$$

- $\mathbf{C} = \mathbf{I}_K$ — базовый алгоритм **Cadzow** (Cadzow, 1988) ($\Pi_{\mathcal{M}_r}$ — используя обычное SVD, $\Pi_{\mathcal{H}}$ — диаг. усреднение).
- $\mathbf{C} = \text{diag}(1, \alpha, \alpha, \dots, \alpha, 1, \alpha, \dots, 1)$, где единицы стоят на $1, L+1, 2L+1, \dots, K$ месте, $0 \leq \alpha \leq 1$.

Cadzow(α) (Zhigljavsky, 2015)

$\Pi_{\mathcal{M}_r}$ — косоугольное SVD, $\Pi_{\mathcal{H}}$ — взвешенное диагональное усреднение.

$$\|\mathbf{X}\|_{1,M}^2 = \sum_{l=1}^L \sum_{k=1}^K m_{l,k}$$

- $\|\mathbf{X}\|_{1,M}$, $m_{l,k}$ такие, что $q_i = 1$ — **Weighted Cadzow** ($\Pi_{\mathcal{M}_r}$ — EM, $\Pi_{\mathcal{H}}$ — диаг. усреднение). Для него все $q_i = 1$.
- **Extended Cadzow** (использует искусственное продолжение ряда в обе стороны)

Различные особенности

$$\mathbf{C} = \text{diag}(1, \alpha, \alpha, \dots, \alpha, 1, \alpha, \dots, 1) \text{ (Cadzow}(\alpha)\text{)}.$$

Теорема

Эквивалентные веса ряда в алгоритме Cadzow(α):

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

Достаточно взять $\alpha = 0$, чтобы получить равномерные веса?

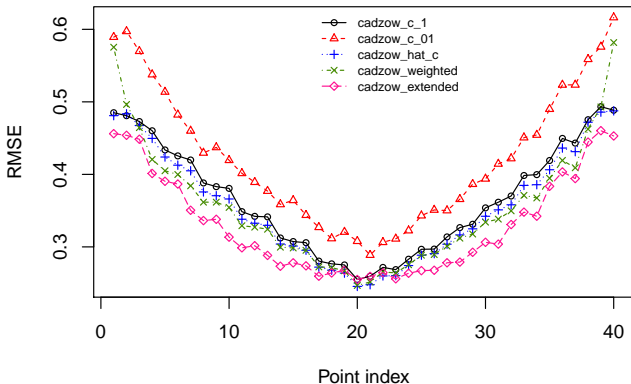
Нет!

- $\alpha = 0$ — другая задача
- $\alpha \approx 0$, $\alpha > 0$: медленно сходится, проблемы со слабой делимостью, требуется, чтобы $N \vdots L$.
- $\alpha = 1$ — базовый алгоритм Cadzow

Сравнение (одна итерация)

Задача оценки сигнала: $N = 40$, $L = 20$, $r = 2$,

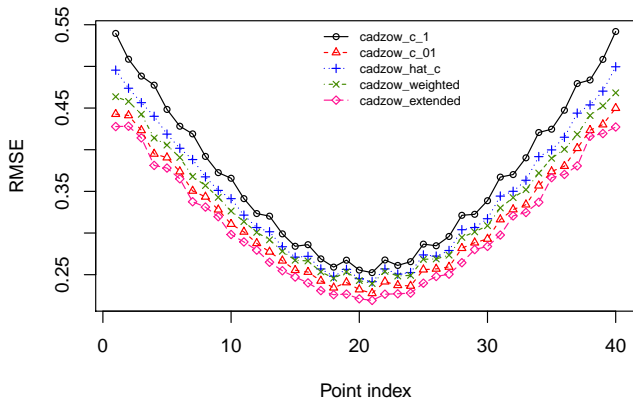
$\mathbb{X} = (x_1, \dots, x_N)$, $x_k = 5 \sin \frac{2k\pi}{6} + \varepsilon_k$; ε_k , $k = 1, \dots, N$, — гауссовский белый шум, $\mathbb{E}\varepsilon_k = 0$, $\mathbb{D}\varepsilon_k = 1$, 1000 реализаций.



Сравнение (100 итераций)

Задача оценки сигнала: $N = 40$, $L = 20$, $r = 2$,

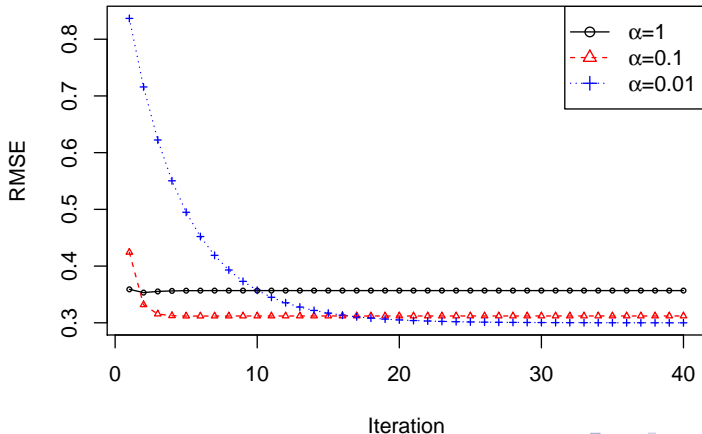
$\mathbb{X} = (x_1, \dots, x_N)$, $x_k = 5 \sin \frac{2k\pi}{6} + \varepsilon_k$; ε_k , $k = 1, \dots, N$, — гауссовский белый шум, $\mathbb{E}\varepsilon_k = 0$, $\mathbb{D}\varepsilon_k = 1$, 1000 реализаций.



Сравнение (зависимость RMSE от числа итераций)

Задача оценки сигнала: $N = 40$, $L = 20$, $r = 2$,

$\mathbb{X} = (x_1, \dots, x_N)$, $x_k = 5 \sin \frac{2k\pi}{6} + \varepsilon_k$; ε_k , $k = 1, \dots, N$, — гауссовский белый шум, $\mathbb{E}\varepsilon_k = 0$, $\mathbb{D}\varepsilon_k = 1$, 1000 реализаций.



Реальный пример

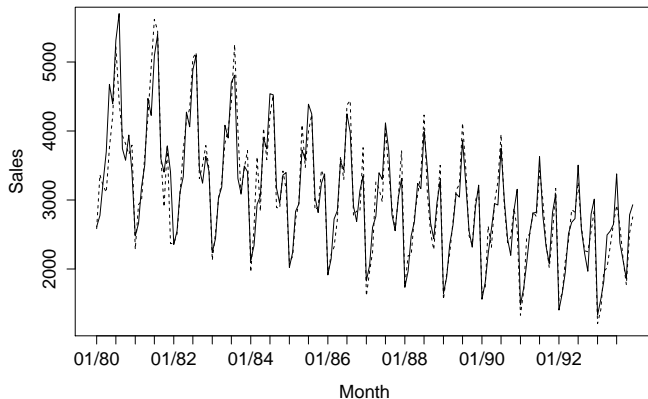
Ряд 'Fortified wine', первые 168 значений.

Схема выбора параметра α для применения метода Cadzow(α):

- Построена параметрическая модель “похожего” ряда с помощью SSA + ESPRIT
- С помощью моделирования метода Cadzow(α) получено качество оценки модельного сигнала и точность аппроксимации ряда
- Найден промежуток значений α , на котором улучшение точности аппроксимации даёт улучшение RMSE оценки сигнала
- Метод применён к исходному ряду с параметром α , соответствующим наибольшей точности

Реальный пример: наилучшее приближение

Получено алгоритмом Cadzow(0.2). Пунктирная линия — исходный ряд, сплошная линия — аппроксимация рядом конечного ранга.



Итоги

- Рассмотрен широкий класс итерационных алгоритмов решения задачи для оценивания сигнала \mathbb{S} в модели $\mathbb{X} = \mathbb{S} + \mathbb{N}$
- Доказано существование сходящейся подпоследовательности для всех алгоритмов
- Установлено соотношение между разделимостью, скоростью сходимости и точностью в пределе на численных примерах
- Проведено численное сравнение на модельном и реальном примере
- Получена скорость разделимости для одного примера в алгоритме $\text{Cadzow}(\alpha)$, рекомендации к быстрой реализации алгоритмов (в приложении)