# Итеративные алгоритмы взвешенной аппроксимации рядами конечного ранга

Н. К. Звонарев и Н. Э. Голяндина

Санкт-Петербургский государственный университет

29 января 2015 г.

# Постановка задачи

- Дано: $\mathbb{X}$ — временной ряд, $\mathbb{X} \in \mathcal{X}_N$ — множество всех временных рядов длины $N$.
- $\mathcal{X}_N^r$ — подмножество $\mathcal{X}_N$ рядов некоторой структуры.
- Найти: $\mathbb{Y} = (y_1, \ldots, y_N)$: $f_q(\mathbb{Y}) \to \min, \mathbb{Y} \in \mathcal{X}_N^r$,
  $f_q(\mathbb{Y}) = \sum\limits_{i=1}^{N} q_i(x_i - y_i)^2$, $q_i \geq 0$ — неотрицательные веса.
- Приложение в статистике: наблюдаем $\mathbb{X} = \mathbb{Y} + \varepsilon$, где $\mathbb{Y} \in \mathcal{X}_N^r$ — сигнал, $\varepsilon$ — белый гауссовский шум с нулевым матожиданием. Требуется найти оценку для $\mathbb{Y}$. В качестве оценки можно рассматривать решение задачи.

# Ряды, управляемые ЛРФ

## Определение

*Ряд, управляемый линейной рекуррентной формулой порядка $r$:*
$\mathbb{X} = (x_1, \ldots, x_N)$, $\quad x_n = \sum_{i=1}^{r} a_i x_{n-i}$, $\quad n = r+1, \ldots, N$.

Параметрический вид:

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

Можно показать эквивалентность поиска ряда, управляемого ЛРФ, и поиска ряда конечного $L$-ранга при наложении дополнительных условий.

# Постановка на матричном языке

1. Временной ряд $\mathbb{X} = (x_1, \ldots, x_N)$.
2. $\mathbf{X} = \mathcal{T}(\mathbb{X})$, где $\mathcal{T}$ — биективное отображение на множество траекторных матриц:

$$\mathbf{X} = \begin{pmatrix} x_1 & x_2 & \cdots & x_K \\ x_2 & \iddots & \iddots & x_{K+1} \\ \vdots & \iddots & \vdots & \vdots \\ x_L & x_{L+1} & \cdots & x_N \end{pmatrix}$$

3. Найти $\mathbf{Y}$: $||\mathbf{X} - \mathbf{Y}||_? \to \min$, $\mathbf{Y} \in \mathcal{H} \cap \mathcal{M}_r$, где $\mathcal{H}$ — множество ганкелевых матриц, $\mathcal{M}_r$ — матрицы ранга $\le r$.
4. $\mathbb{Y} = \mathcal{T}^{-1}(\mathbf{Y})$

# Структура сигнала

Временной ряд $\mathbb{X} = (x_1, \ldots, x_N)$.
Параметры: $1 < L < N$ — *длина окна*, $K = N - L + 1$, $0 \leq r \leq L$ — ранг.
Вектора вложения: $X_i = (x_i, \ldots, x_{i+L-1})^{\mathrm{T}}, \qquad i = 1, \ldots, K$.

## Определение

*Траекторное пространство:* $\mathfrak{X}^{(L)}(\mathbb{X}) = \mathfrak{X}^{(L)} = \mathrm{span}(X_1, \ldots, X_K)$.

## Определение

*$L$-ранг ряда $\mathbb{X}$ равен $r$* $\Leftrightarrow \dim \mathfrak{X}^{(L)} = r$.

$\mathcal{X}_N$ — множество всех временных рядов длины $N$. $\mathcal{X}_N^r$ — множество всех временных рядов длины $N$ $L$-ранга, не превосходящего $r$.

# Стандартное решение

$\mathbf{Y}$: $||\mathbf{X}-\mathbf{Y}||_? \to \min$, $\mathbf{Y} \in \mathcal{H} \cap \mathcal{M}_r$, где $\mathcal{H}$ — множество ганкелевых матриц, $\mathcal{M}_r$ — матрицы ранга $\leq r$.

## Решение

*Переменные проекции:*

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

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

$\Pi_{\mathcal{H}}$ и $\Pi_{\mathcal{M}_r}$ зависят от $||\cdot||_?$. Варианты:

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

# Эквивалентность скалярных произведений и норм

1. $< \mathbb{X}, \mathbb{Y} >_q = \sum_{i=1}^{N} q_i x_i y_i$
2. $< \mathbf{X}, \mathbf{Y} >_M = \sum_{l=1}^{L} \sum_{k=1}^{K} m_{l,k} x_{l,k} y_{l,k}$
3. $< \mathbf{X}, \mathbf{Y} >_{\mathbf{C}} = \mathrm{tr}(\mathbf{X}\mathbf{C}\mathbf{Y}^{\mathrm{T}})$.

## Утверждение

*1. Пусть* $\mathbf{X} = \mathcal{T}(\mathbb{X})$, $\mathbf{Y} = \mathcal{T}(\mathbb{Y})$. *Тогда* $< \mathbb{X}, \mathbb{Y} >_q = < \mathbf{X}, \mathbf{Y} >_M$ *тогда и только тогда, когда*

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

*2. Для диагональной матрицы* $\mathbf{C}$, $< \mathbf{X}, \mathbf{Y} >_M = < \mathbf{X}, \mathbf{Y} >_{\mathbf{C}}$ *тогда и только тогда, когда*

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

# Проектор $\Pi_{\mathcal{M}_r}$

Варианты:

1. $||\mathbf{X}||_M$, 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. $||\mathbf{X}||_M$, общий случай: EM-подобный алгоритм.
3. $||\mathbf{X}||_{\mathbf{C}}$: косоугольное SVD-разложение. $\mathbf{C} = \mathbf{O}_{\mathbf{C}}^T\mathbf{O}_{\mathbf{C}}$, $\mathbf{B} = \mathbf{X}\mathbf{O}_{\mathbf{C}}^T$, $\Pi_{\mathcal{M}_r}^{\mathbf{C}}(\mathbf{X}) = \Pi_{\mathcal{M}_r}(\mathbf{B})(\mathbf{O}_{\mathbf{C}}^T)^\dagger$.

# Проектор $\Pi_{\mathcal{H}}$

Варианты:

1. $||\mathbf{X}||_M$, $m_{l,k}$ 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. $||\mathbf{X}||_{\mathbf{S}}$, $\mathbf{S}$ — диагональная: взвешенное диагональное усреднение с весами $s_{i,i}$, где $\mathbf{S} = (s_{i,j})$.

# Варианты алгоритма Cadzow

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

1. Все $m_{l,k} = 1$ — базовый алгоритм Cadzow ($\Pi_{\mathcal{M}_r}$ — обычный SVD, $\Pi_{\mathcal{H}}$ — диаг. усреднение).

## Замечание

*Эквивалентные веса базового алгоритма:*

$$q_i = w_i = \begin{cases} i & \textit{для } i = 1, \ldots, L-1, \\ L & \textit{для } i = L, \ldots, K, \\ N - i + 1 & \textit{для } i = K+1, \ldots, N, \end{cases} .$$

*что не всегда естественно.*

2. $||\mathbf{X}||_M$, $m_{l,k} = 1/w_{l+k-1}$ — Weighted Cadzow ($\Pi_{\mathcal{M}_r}$ — EM, $\Pi_{\mathcal{H}}$ — диаг. усреднение). Для него все $q_i = 1$.

# Варианты Cadzow с косоугольным SVD

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

1. $\mathbf{C} = \mathrm{diag}(1, \alpha, \alpha, \ldots, \alpha, 1, \alpha, \ldots, 1)$, где единицы стоят на 1, $L+1$, $2L+1$, ..., $K$ месте, $0 \le \alpha \le 1$ (Cadzow($\alpha$))

**Теорема**

*Эквивалентные веса алгоритма Cadzow($\alpha$):*

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

2. $\hat{\mathbf{C}} = \mathrm{diag}(C)$, где $C$ — усреднение матрицы $\mathbf{M}$, примененной в алгоритме Weighted Cadzow, по столбцам (Hat-Cadzow).

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

# Различные проблемы

- Базовый алгоритм Cadzow — имеет $q_i$, далекие от оптимальных
- Weighted Cadzow — работает долго, так как использует итерационный EM-алгоритм для проектирования на множество матриц ранга $\leq r$
- Cadzow($\alpha$) с косоугольным SVD: при $\alpha = 0$ — вообще нет сходимости к нужному множеству, $\alpha \approx 0$, $\alpha > 0$: медленно сходится, проблемы со слабой разделимостью, требуется, чтобы $N$ было кратно $L$

$$\alpha = 0 \implies \mathbf{C} = \mathrm{diag}(1, 0, 0, \ldots, 0, 1, 0, \ldots, 1) \quad —$$

вырожденная матрица

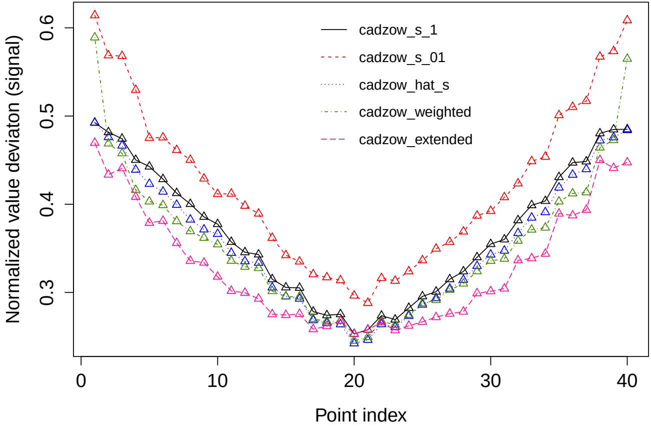
# Алгоритм Extended Cadzow

$$\mathcal{X} = \begin{pmatrix} & & & x_1 & x_2 & \cdots & x_K & x_{K+1} & \cdots & x_N \\ & & \iddots & x_2 & x_3 & \iddots & x_{K+1} & \iddots & \iddots & \\ & x_1 & \iddots & \iddots & \iddots & \iddots & \iddots & x_N & & \\ x_1 & x_2 & \cdots & x_L & x_{L+1} & \cdots & x_N & & & \end{pmatrix}.$$

1. $q_i = 1$, $i = 1, \ldots, N$, но задача несколько иная
2. $\tilde{\mathcal{T}} : \tilde{\mathcal{T}}(\mathbb{X}) = \mathcal{X}$ — биекция между рядами и матрицами с пропусками.
3. Проектор на множество матриц неполного ранга $\Pi_{\widetilde{\mathcal{M}}_r}$ (EM-алгоритм).
4. Проектор на множество траекторных псевдоматриц $\Pi_{\mathfrak{X}}$ (диагональное усреднение).
5. Сам алгоритм: $\mathbb{Y} = \tilde{\mathcal{T}}^{-1}((\Pi_{\mathfrak{X}} \circ \Pi_{\widetilde{\mathcal{M}}_r})^I(\tilde{\mathcal{T}}(\mathbb{X})))$.

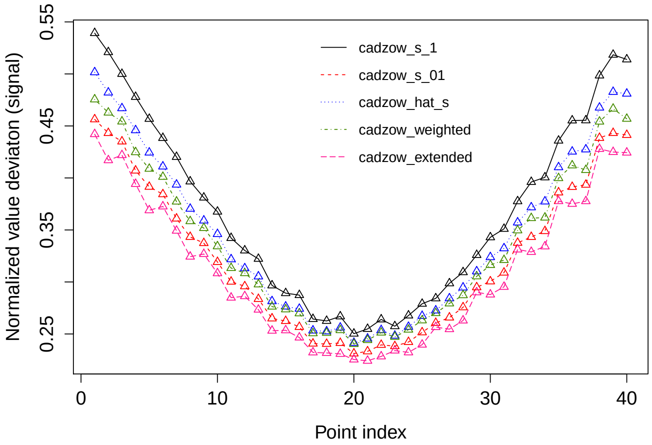
# Пример №1 (одна итерация)

Задача оценки сигнала: $N = 40$, $L = 20$, $r = 2$, $\mathbb{X} = (x_1, \ldots, x_N)$, $x_k = 5 \sin \frac{2k\pi}{6}$.



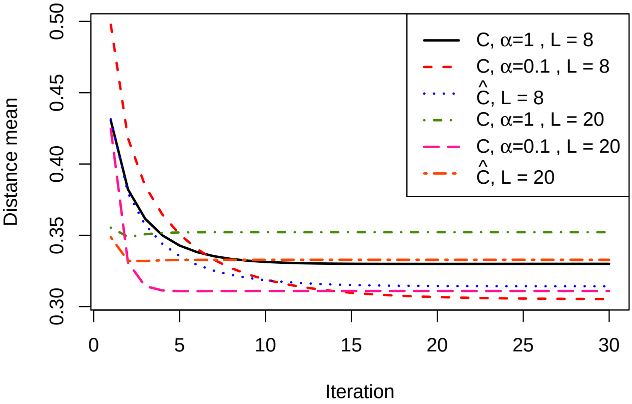
# Пример №2 (много итераций)

Задача оценки сигнала: $N = 40$, $L = 20$, $r = 2$, $\mathbb{X} = (x_1, \ldots, x_N)$, $x_k = 5 \sin \frac{2k\pi}{6}$.



# Пример №3

Сходимость алгоритма: $N = 40$, $L_1 = 8$, $L_2 = 20$, $r = 2$,
$\mathbb{X} = (x_1, \ldots, x_N)$, $x_k = 5 \sin \frac{2k\pi}{6}$.



# Выводы

Рассмотрен широкий класс алгоритмов, проведено сравнение на модельном примере:

- Доказана сходимость по подпоследовательностям для всех алгоритмов
- Единичные веса достигаются только на алгоритмах с вложенными итерациями (Extended, Weighted Cadzow), которые являются медленными.
- Алгоритмы с более близкими к единичным весами дают лучшую аппроксимацию в пределе, но медленее сходятся и имеют проблемы с разделимостью:
  Cadzow($\alpha$): малое $L$, $\alpha \Rightarrow$ плохая разделимость
- Лучший метод среди всех – Extended Cadzow. Среди методов без внутренних итераций на одной лучше всего работает Cadzow(1) и Hat-Cadzow, в пределе — Cadzow(0.1)