

# «Использование метода SSA в машинном обучении для прогноза временных рядов»

Ежов Федор Валерьевич, группа 20.M03-мм

Санкт-Петербургский государственный университет  
Кафедра статистического моделирования

Научный руководитель: к.ф.-м.н., доцент Голяндина Н.Э.  
Рецензент: Программист, Майкрософт Израиль, Шлемов А.Ю.



Санкт-Петербург  
2022г.

# Постановка задачи

Рассмотрим временной ряд  $X_N = (x_1, \dots, x_N)$ .

**Задача:** прогнозирование временного ряда.

Нейронные сети (NN) подходят для решения этой задачи.

**Предобработка:** Singular Spectrum Analysis (SSA).

Методы, использующие предобработанные данные, называем гибридными (SSA-NN).

Метод SSA также может прогнозировать временные ряды самостоятельно.

**Что лучше?** Метод SSA или NN или SSA-NN.

# Singular Spectrum Analysis (SSA)

Считаем, что  $X_N = S_N + \xi_N$ , где  $S_N$  — сигнал,  $\xi_N$  — шум, случайный процесс с нулевым мат. ожиданием.

Входные данные:  $X_N = (x_1, \dots, x_N)$ .

Результат:  $\hat{S}_N = (\hat{s}_1, \dots, \hat{s}_N)$  — оценка сигнала ряда  $X_N$ .

Параметры:  $1 < L < N$  — длина окна.  $1 < r < L$  — количество компонент.

## Алгоритм выделения сигнала:

- 1 Траекторная матрица.

$$\mathbf{X} = [X_1 : \dots : X_K], K = N - L + 1; X_i = (x_i, \dots, x_{i+L-1}),$$

- 2 SVD.  $\mathbf{X} = \sum_{i=1}^L \sqrt{\lambda_i} U_i V_i^T$ ,

- 3 Оценка  $\hat{S}_N$  получается путем диагонального усреднения матрицы  $\hat{\mathbf{X}} = \sum_{i=1}^r \sqrt{\lambda_i} U_i V_i^T$ .

Результат:  $\hat{S}_N = (\hat{s}_1, \dots, \hat{s}_N)$ .

Подробнее в книге: Golyandina, Nekrutkin, Zhigljavsky, Analysis of Time Series Structure - SSA and Related Techniques, 2001.

# Singular Spectrum Analysis (SSA)

Модель сигнала конечного ранга:

$$s_n = \sum_{k=1}^P \sin(2\pi\omega_k n + \phi_k) \cdot e^{\alpha_k n} \cdot (\beta_{m_k} n^{m_k} + \dots + \beta_1 n + \beta_0)$$

Число компонент  $r$  для частных случаев:

- ❶  $r = 1$  для константы или экспоненциального ряда.
- ❷  $r = 2$  для синуса или косинуса.
- ❸  $r = m + 1$  для полинома степени  $m$ .

Рекомендации для выделения сигнала:

- ❶  $L = \lfloor \frac{N+1}{2} \rfloor$ , если сигнал конечного ранга.
- ❷  $L$  поменьше, если сигнал локально (приблизленно) конечного ранга.

# Обычные и гибридные методы

- MultiLayer Perceptron (MLP) (далее именуемая как ANN).
- Recurrent neural network (RNN).
- Gated recurrent unit (GRU).
- Long short-term memory (LSTM).

Гиперпараметры нейронных сетей:  $T$  – размер входного вектора,  $h$  – размер скрытого слоя.

Параметры (веса) нейронных сетей инициализируются случайно.  
Оптимизация параметров происходит с помощью «обратного распространения ошибки».

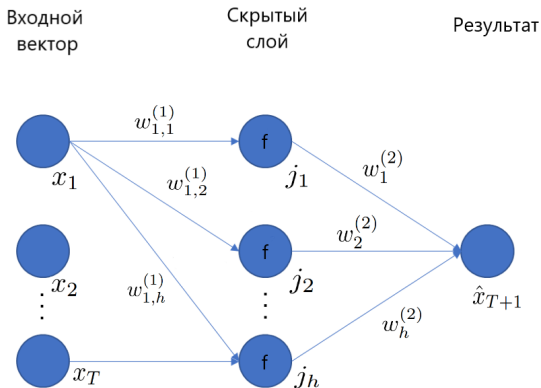
Обычные (негибридные) методы: ANN, RNN, GRU, LSTM.

Гибридные методы: SSA-ANN, SSA-RNN, SSA-GRU, SSA-LSTM.

## Отличие

На вход гибридным методам подается оценка сигнала ряда.

$$\hat{x}_{T+1} = \sum_{\ell=1}^h w_{\ell}^{(2)} f\left(\sum_{i=1}^T w_{i\ell}^{(1)} x_i + \theta^{(1)}\right) + \theta^{(2)}.$$



# Обычные и гибридные методы

- MultiLayer Perceptron (MLP) (далее именуемая как ANN).
- Recurrent neural network (RNN).
- Gated recurrent unit (GRU).
- Long short-term memory (LSTM).

Гиперпараметры нейронных сетей:  $T$  – размер входного вектора,  $h$  – размер скрытого слоя.

Параметры (веса) нейронных сетей инициализируются случайно.  
Оптимизация параметров происходит с помощью «обратного распространения ошибки».

Обычные (негибридные) методы: ANN, RNN, GRU, LSTM.

Гибридные методы: SSA-ANN, SSA-RNN, SSA-GRU, SSA-LSTM.

## Отличие

На вход гибридным методам подается оценка сигнала ряда.

Пусть  $\tau$ ,  $v$  и  $t = N - T$  номера последних строчек в каждой соответствующей части.

$$\mathbf{Z} = \left( \begin{array}{cccc|c} z_1 & z_2 & \cdots & z_T & z_{T+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ z_\tau & z_{\tau+1} & \cdots & z_{\tau+T} & z_{\tau+T+1} \\ z_{\tau+T+2} & z_{\tau+T+3} & \cdots & z_{\tau+2T+1} & z_{\tau+2T+2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ z_v & z_{v+1} & \cdots & z_{v+T} & z_{v+T+1} \\ z_{v+T+2} & z_{v+3} & \cdots & z_{v+2T+1} & z_{v+2T+2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ z_{N-T} & z_{N-T+1} & \cdots & z_{N-1} & z_N \end{array} \right).$$

Каждая выборка выделена цветом:  $\mathbf{Z}_{\text{train}}$ ,  $\mathbf{Z}_{\text{val}}$ ,  $\mathbf{Z}_{\text{test}}$ .



## Тренировочная выборка:

Для обучения SSA применяется к  $Z_{\text{train}} = (z_1, \dots, z_{\tau+T+1})$ .

## Валидационная/тестовая выборки:

Для прогноза элемента  $z_m$  применяем SSA к  $(z_{m-1-(\tau+T)}, \dots, z_{m-1})$ .

## Основные гиперпараметры гибридных методов:

- $L$  (размер окна SSA).
- $r$  (количество компонент в SSA для оценки сигнала).
- $T$  (размер входного вектора в нейронной сети).
- $h$  (размер скрытого слоя нейронной сети).

# Как сравнивать ?

**Задача:** Сравнить метод SSA или NN или SSA–NN.

Нужна методика сравнения.

Методика должна позволять:

- Сравнить различные группы методов корректно.
- Получить рекомендации по выбору гиперпараметров.
- Оценить устойчивость относительно параметров методов.
- Оценить устойчивость относительно случайных возмущений.

Прогноз методом SSA.

Выбор параметров SSA для гибридных методов.

**Сравнение методов:** После выбора лучших пар  $L$  и  $r$  для гибридных методов для гиперпараметров  $T$  и  $h$  задается сетка.  
В каждой ячейке  $T, h$  находится соответствующая ошибка прогноза.

	$T_1$	$T_2$	$\dots$	$T_n$
$h_1$	$\text{err}(T_1, h_1)$	$\text{err}(T_2, h_1)$	$\dots$	$\text{err}(T_n, h_1)$
$h_2$	$\text{err}(T_1, h_2)$	$\text{err}(T_2, h_2)$	$\dots$	$\text{err}(T_n, h_2)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$h_m$	$\text{err}(T_1, h_m)$	$\text{err}(T_2, h_m)$	$\dots$	$\text{err}(T_n, h_m)$

Усреднение по  $T$  или по  $h$ .

Прогноз методом SSA.

Выбор гиперпараметров SSA для гибридных методов.

Сравнение методов.

**Проверка устойчивости методов:** Дополнительная проверка устойчивости осуществляется с помощью случайной инициализации весов.

Также в методику включена различная визуализация в виде графиков или чисел.

Методика применялась к синтетическим и реальным данным.

Реальные данные:

- Indian Rain.
- Earth Orientation Parameters (EOP).
- Погода в Санкт-Петербурге.

На синтетических данных было исследовано:

- Влияние выбора параметра  $r$ .
- Влияние шума в ряде.
- Влияние длины ряда.
- Влияние красного шума.

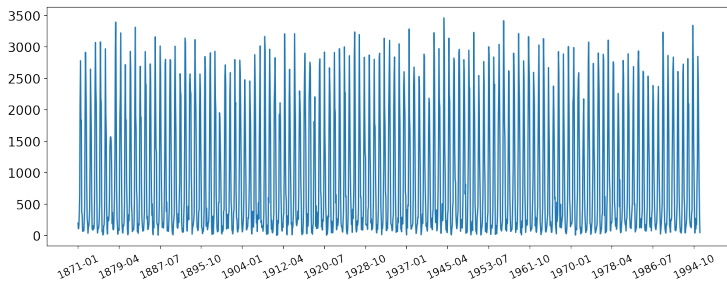
# Эксперименты. Indian Rain

Критика в Du, Zhao, Lei, 2017: гибридные методы не улучшают прогноз.

Рассмотрим временной ряд «Indian Rain» длины 1500 точек. Данные показывают средние месячные осадки в Индии.

Разбиение: 750 (тренировочная), 500 (валидационная), 250 (тестовая) точек.

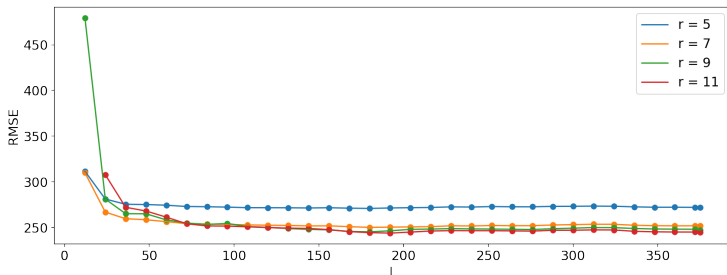
Гиперпараметры  $r = 7$ ,  $L = 375$ .



# Эксперименты. Indian Rain. Прогноз SSA

**Цель:** посмотреть для каких  $L$  и  $r$  прогноз SSA имеет оптимальную ошибку.

На графике: зависимость RMSE прогноза метода SSA на валидационной выборке относительно  $L$ .



# Эксперименты. Indian Rain. Сравнение методов

Зададим  $T = \{12, 48, \dots, 408\}$ ,  $h = \{10, 25, \dots, 100\}$ .

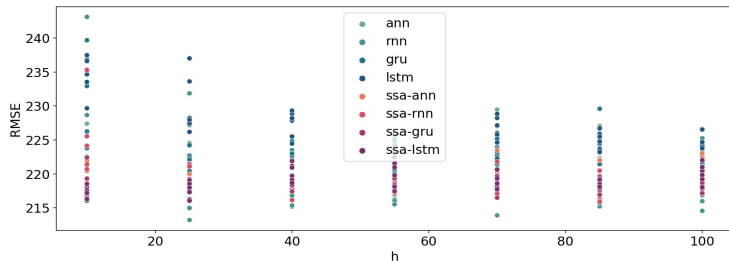
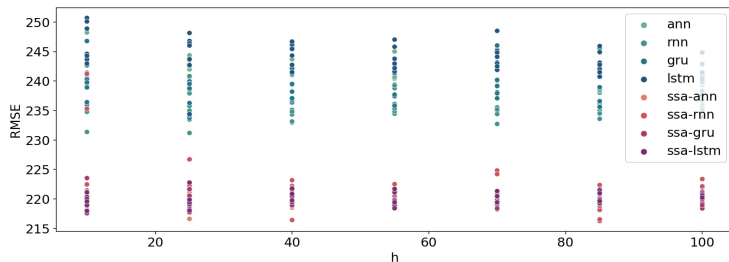
Усреднение ошибок по  $h$ . Результат:  $SSA-NN < SSA < NN$ .





# Эксперименты. Indian Rain. Проверка устойчивости

Зададим  $T = \{12, 156\}$ ,  $h = \{10, 25, \dots, 100\}$ .



- $SSA-NN < SSA < NN$
- SSA позволяет выбрать методы с меньшим размером  $T$  и  $h$ , что дает больше количество пар «признаки-предсказываемые значения», а также снижает сложность модели.

Объяснение:

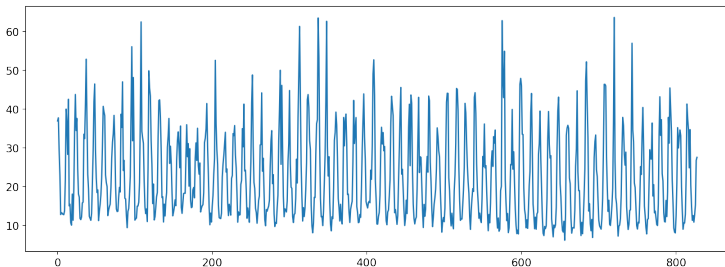
- Ряд имеет достаточную длину.
- Сигнал ряда простой и конечного ранга.
- Достаточный уровень шума.

# Эксперименты. Погода

Характеристика погоды в Санкт-Петербурге. Размер ряда 828 точек.  
Ряд похож на «Indian Rain».

Разбиение: 528, 150, 150 точек.

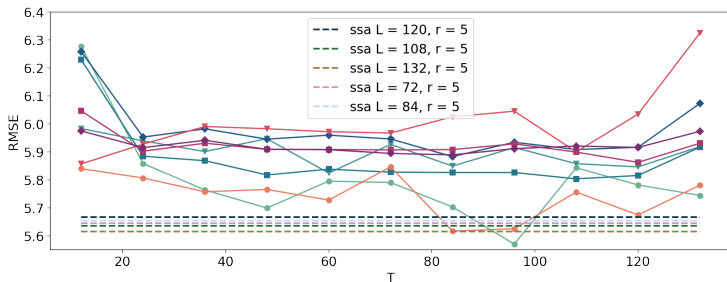
$L = 264$ ,  $r = 5$ .



Результаты:

- Результаты для NN и SSA–NN смешиваются.
- Средняя ошибка метода SSA наименьшая.

Вопрос таких результатов остается открытым.



# Эксперименты. Earth Orientation Parameters (EOP)

Метод SSA дал хорошие результаты ранее (Okhotnikov, Golyandina, 2019).

Рассмотрим ряд «x pole» из данных EOP.

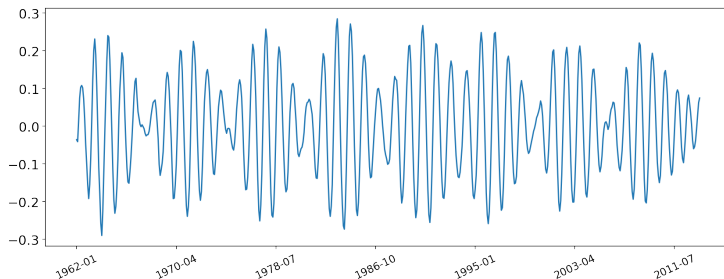
Перейдем от дневных к месячным значениям.

Вычтем тренд. Размер ряда 620.

Разбиение: 320, 150, 150 точек.

Сигнал неконечного ранга. Маленький шум.

$L = 78$ ,  $r = 18$ .



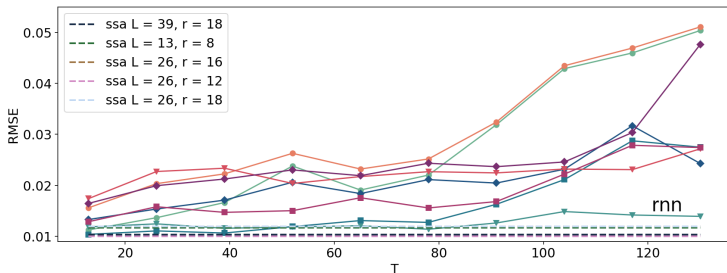
# Эксперименты. Earth Orientation Parameters (EOP)

Результаты:

- $SSA \leq NN < SSA-NN$ .
- Наилучший результат показал метод SSA ( $L = 26, r = 12$ ).
- Наилучшие параметры приводят к сильной аппроксимации методом SSA.

Возможные объяснения:

- Сигнал неконечного ранга.
- Маленький шум.



# Эксперименты. Таблица с результатами

Data	ssa-params		ann		rnn		gru		lstm		ssa
Rain	-		225.75		<b>223.07</b>		227.73		230.17		222.39
	$L = 375, r = 7$		220.80		220.58		<b>219.81</b>		220.46		222.39
EOP	-		0.027		<b>0.013</b>		0.016		0.020		0.011
	$L = 78, r = 18$		0.030		0.023		<b>0.018</b>		0.025		0.011
Погода	-		<b>5.803</b>		5.911		5.894		5.969		5.638
	$L = 264, r = 5$		<b>5.753</b>		5.868		5.912		5.895		5.638

# Эксперименты. Синтетические данные

Рассмотрим ряд  $\{Z_N : z_i = (\sin(2\pi \frac{i}{6}) + 2 \sin(2\pi \frac{i}{12}))\}$ .

Ранг ряда равен 4. Будем добавлять к ряду белый или красный шум.

Эксперименты:

- ❶  $Z_{650} + \text{белый шум } 1.5\varepsilon_i; \varepsilon_i \sim N(0, 1).$
- ❷  $Z_{650} + \text{белый шум } 0.3\varepsilon_i.$
- ❸  $Z_{650} + \text{красный шум } \xi_i, \text{ где } \xi_i = \xi_{i-1} + \sigma\varepsilon_i, \sigma = 1.2, \varepsilon_i \sim N(0, 1).$

experiment	ssa-params		ann	rnn	gru	lstm		ssa
1	-		1.657	1.635	<b>1.623</b>	1.625		1.586
	$L = 175, r = 4$		1.567	1.587	<b>1.566</b>	1.581		1.586
2	-		0.326	0.325	<b>0.321</b>	0.330		0.316
	$L = 175, r = 4$		<b>0.312</b>	0.316	0.313	0.317		0.316
3. Ряд	-		1.819	<b>1.396</b>	1.415	1.461		1.803
	$L = 84, r = 14$		2.031	1.968	<b>1.914</b>	2.043		1.803
3. Сигнал	-		<b>1.404</b>	1.679	1.697	1.577		1.398
	$L = 175, r = 4$		0.919	<b>0.856</b>	0.876	0.943		1.398



# Эксперименты. Синтетические данные

## Результаты эксперимента с красным шумом:

- $SSA-NN < SSA < NN$ , ошибка прогноза сигнала.
- $SSA-NN > SSA > NN$ , ошибка прогноза всего ряда.

Разные постановки задачи приводят к разным результатам.

experiment	ssa-params		ann	rnn	gru	lstm		ssa
1	-		1.657	1.635	<b>1.623</b>	1.625		1.586
	$L = 175, r = 4$		1.567	1.587	<b>1.566</b>	1.581		1.586
2	-		0.326	0.325	<b>0.321</b>	0.330		0.316
	$L = 175, r = 4$		<b>0.312</b>	0.316	0.313	0.317		0.316
3. Ряд	-		1.819	<b>1.396</b>	1.415	1.461		1.803
	$L = 84, r = 14$		2.031	1.968	<b>1.914</b>	2.043		1.803
3. Сигнал	-		<b>1.404</b>	1.679	1.697	1.577		1.398
	$L = 175, r = 4$		0.919	<b>0.856</b>	0.876	0.943		1.398

- Разработана методика, позволяющая сравнивать методы и помогать в выборе гиперпараметров.
- Продемонстрирован положительный результат гибридных методов на реальных данных.
- Гибридные методы не всегда работают точнее обычных. Осталось много вопросов.
- При применении гибридных методов важно помнить, что они решают задачу прогнозирования сигнала.