
ПОГРУЖЕНИЕ ВРЕМЕННЫХ РЯДОВ С ВЫСОКОЙ ВОЛАТИЛЬНОСТЬЮ В МЕТРИЧЕСКОЕ ПРОСТРАНСТВО

A PREPRINT

Эйнуллаев Алтай

Кафедра интеллектуальных систем
Московский физико-технический институт
Долгопрудный
einullaev.ae@phystech.edu

Яковлев Константин

Кафедра интеллектуальных систем
Московский физико-технический институт
Долгопрудный
iakovlev.kd@phystech.edu

ABSTRACT

Решается задача прогнозирования финансовых временных рядов. Основными особенностями таких временных рядов являются высокая волатильность и высокая попарная ковариация. Базовым подходом к решению задачи является выполнение прогноза в исходном пространстве. Предлагаемый метод COF (correlation optimization forecasting) заключается в переходе в пространство попарных расстояний между временными рядами, прогнозировании в нем и переходе обратно в исходное пространство. Для его реализации вводится функцию расстояния между временными рядами, которая должна удовлетворять определенным свойствам. В данной статье изучаются эти свойства и проводятся сравнения различных функций попарного расстояния на основе численных экспериментов.

Keywords Набор временных рядов · LSTM · COF · Корреляция · ARIMA · Матрица попарных расстояний

1 Introduction

Решается задача прогнозирования набора временных рядов, используя функцию попарного расстояния. Прогнозируемые временные ряды обладают следующими свойствами: высокая попарная корреляция и высокая дисперсия. Набору временных рядов ставится в соответствие и прогнозируется матрица попарных расстояний. При этом функция расстояния выбирается так, чтобы по полученной матрице расстояний можно было восстановить прогноз для набора временных рядов.

Временной ряд — последовательно измеренные через равные промежутки времени данные. [1] Прогнозирование временных рядов заключается в построении модели для предсказания будущих событий основываясь на известных событиях прошлого, предсказания будущих данных до того как они будут измерены. В [2] для прогнозирования использовался алгоритм экспоненциального сглаживания (англ. Exponential Smoothing). В [3] — LSTM (англ. Long Short-Term Memory), в [4] — ARIMA (англ. autoregressive integrated moving average). В работах [5], [6] набор коррелированных временных рядов прогнозируется с помощью глубоких нейронных сетей и метода экспоненциального сглаживания, соответственно.

Предлагается использовать свойство высокой коррелированности набора временных рядов при прогнозировании. Вместо непосредственного прогнозирования временных рядов, прогнозируется матрица попарных расстояний. Прогноз набора временных рядов восстанавливается по матрице путем решения задачи оптимизации. В качестве функции попарных расстояний рассматривается ковариация между временными рядами [7]. Для набора временных рядов задана матрица ковариации, вычисляемая в каждый момент времени.

Численные эксперименты проводятся на наборе финансовых временных рядов. Эксперимент состоит в прогнозировании матрицы попарных расстояний, оценке качества прогноза, восстановлении прогноза в исходное пространство и оценке его точности. В качестве прогностической модели рассматривается ARIMA и LSTM. По результатам экспериментов проводится анализ точности прогноза в зависимости от выбранной функции попарных расстояний и вида данных. Цель эксперимента состоит в оптимальном выборе функции попарных расстояний для выполнения прогноза и сравнение работы метода с ARIMA и LSTM.

2 Постановка задачи

Обозначим через $\mathbf{X} \in R^{n \times m}$ заданный набор из n временных рядов:

$$\mathbf{X} = [\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}]^T. \quad (1)$$

Через $\mathbf{X}_t \in R^{n \times t}$ обозначим $t < m$ первых столбцов \mathbf{Y} :

$$\mathbf{X}_t = [\mathbf{x}_{1:t}^{(1)}, \dots, \mathbf{x}_{1:t}^{(n)}]^T. \quad (2)$$

Определим функцию расстояния между временными рядами удовлетворяющую условиям Мерсера [8]:

$$d(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) = d(\mathbf{x}^{(j)}, \mathbf{x}^{(i)}),$$

$$\Sigma \in S_n : \Sigma_{ij} = d(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) \text{ — неотрицательно определена.}$$

Обозначим расстояние между временными рядами $\mathbf{x}^{(i)} = [x_1^{(i)}, \dots, x_t^{(i)}]$, $\mathbf{x}^{(j)} = [x_1^{(j)}, \dots, x_t^{(j)}]$ следующим образом:

$$d_t(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) = d_t(i, j). \quad (3)$$

Таким образом, в каждый момент времени t набору временных рядов \mathbf{X}_t поставлена в соответствие матрица попарных расстояний $\Sigma_t \in S_n^+$ (симметричная, неотрицательно определенная матрица):

$$\mathbf{X}_t \Rightarrow \Sigma_t = \begin{pmatrix} d_t(\mathbf{x}^{(1)}, \mathbf{x}^{(1)}) & d_t(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) & \dots & d_t(\mathbf{x}^{(1)}, \mathbf{x}^{(n)}) \\ d_t(\mathbf{x}^{(2)}, \mathbf{x}^{(1)}) & d_t(\mathbf{x}^{(2)}, \mathbf{x}^{(2)}) & \dots & d_t(\mathbf{x}^{(2)}, \mathbf{x}^{(n)}) \\ \vdots & \vdots & \ddots & \vdots \\ d_t(\mathbf{x}^{(n)}, \mathbf{x}^{(1)}) & d_t(\mathbf{x}^{(n)}, \mathbf{x}^{(2)}) & \dots & d_t(\mathbf{x}^{(n)}, \mathbf{x}^{(n)}) \end{pmatrix}. \quad (4)$$

2.1 Выбор функции попарных расстояний

Считаем, что в каждый момент времени временной ряд задается своими прошлыми L значениями. Следовательно, набор временных рядов задается матрицей $\mathbf{X}_t \in R^{n \times L}$:

$$\mathbf{X}_t = \begin{pmatrix} x_{t-L+1}^{(1)} & x_{t-L+2}^{(1)} & \dots & x_t^{(1)} \\ x_{t-L+1}^{(2)} & x_{t-L+2}^{(2)} & \dots & x_t^{(2)} \\ \dots & \dots & \dots & \dots \\ x_{t-L+1}^{(n)} & x_{t-L+2}^{(n)} & \dots & x_t^{(n)} \end{pmatrix} \quad (5)$$

i -й столбец матрицы \mathbf{X}_t обозначим через вектор \mathbf{y}^i . А i -ю строку матрицы \mathbf{X}_t обозначим через вектор $\mathbf{x}^{(i)}$. Рассмотрим три способа вычисления матрицы Σ_t .

Тензорное произведение:

$$\Sigma_t^1 = \frac{1}{L} \sum_{i=1}^L \mathbf{y}^i \mathbf{y}^{iT} \quad (6)$$

Центрированное тензорное произведение:

Введем вектор $\mathbf{m} \in R^n$ следующим образом:

$$\mathbf{m}_j = \frac{1}{L} \sum_{i=1}^L y_j^i. \quad (7)$$

А теперь матрицу Σ_t^2 , при построении которой, в отличие от Σ_t^1 , процентрируем векторы \mathbf{y}^i с помощью \mathbf{m} :

$$\Sigma_t^2 = \frac{1}{L} \sum_{i=1}^L (\mathbf{y}^i - \mathbf{m})(\mathbf{y}^i - \mathbf{m})^T \quad (8)$$

Корреляция Пирсона:

Для каждой строки $\mathbf{x}^{(j)}$, $j \in 1, \dots, n$ матрицы \mathbf{X}_t введем величину:

$$\sigma_j = \sqrt{\frac{1}{L} \sum_{i=1}^L (\mathbf{x}_i^{(j)} - \mathbf{m}_i)^2}. \quad (9)$$

Теперь введем матрицу Σ_t^3 следующим образом:

$$(\Sigma_t^3)_{ij} = \frac{(\Sigma_t^2)_{ij}}{\sigma_i \sigma_j}. \quad (10)$$

2.2 Восстановление прогноза

После прогнозирования матрицы Σ_{t+1} стоит задача восстановления прогноза для набора временных рядов. Для этого в каждом случае нужно решить задачу оптимизации:

$$\hat{\mathbf{y}}^{t+1} = \arg \min_{\mathbf{y} \in R^n} \|\Sigma_{t+1} - \hat{\Sigma}_{t+1}\|_F^2. \quad (11)$$

В силу того, что данная задача оптимизации имеет либо два решения либо нулевое решение, то при условии положительных значений временного ряда, решаем задачу условной оптимизации с условием: $\mathbf{y} \succ 0$.

$$\hat{\mathbf{y}}^{t+1} = \arg \min_{\mathbf{y} \in R_+^n} \|\Sigma_{t+1} - \hat{\Sigma}_{t+1}\|_F^2. \quad (12)$$

2.3 Задача прогнозирования набора временных рядов

В качестве критериев качества прогноза временных рядов используются

$$\text{MAPE} = \frac{1}{m} \sum_{i=1}^m \frac{|x_t^{(i)} - \hat{x}_t^{(i)}|}{|x_t^{(i)}|}, \quad (13)$$

$$\text{MSE} = \frac{1}{m} \sum_{i=1}^m (x_t^{(i)} - \hat{x}_t^{(i)})^2. \quad (14)$$

Также выбираются $p < m$ временных рядов из набора так, чтобы отказ от прогнозирования соответствующих $m - p$ временных рядов существенно повышал качество MAPE, MSE.

2.4 Выбор функции попарных расстояний

Пусть $\mathcal{F} = \{d^{(1)}, \dots, d^{(s)}\}$ — множество функций попарных расстояний, из которого нужно выбрать оптимальный вариант. Пусть $\hat{\Sigma}(d)$ — спрогнозированная, а $\Sigma(d)$ — точная матрица попарных расстояний при использовании $d \in \mathcal{F}$. Тогда оптимальный по точности выбор функции попарных расстояний:

$$d_{\text{acc}} = \arg \min_{d \in \mathcal{F}} \sigma^2(d) = \arg \min_{d \in \mathcal{F}} \|\hat{\Sigma}(d) - \Sigma(d)\|_F^2. \quad (15)$$

3 Вычислительный эксперимент

Целью эксперимента является сравнение предложенного способа прогнозирования временных рядов с ARIMA и LSTM. Также сравниваются различные способы вычисления матрицы попарных расстояний, путем анализа ошибки прогноза. В качестве модели для прогноза матрицы попарных расстояний также рассмотрим LSTM, ARIMA. Проведем эксперименты как на синтетических наборах временных рядов, так и на реальных. Точность прогноза будем оценивать с помощью MAPE (15) и MSE (16).

3.1 Реальные данные.

Прогнозирование с помощью LSTM, ARIMA.

Рассмотрим набор цен на электричество в Германии [9]. Набор состоит из $n = 24$ высоко коррелированных временных рядов. В силу большого количество выбросов в 14 из 24 временных рядов, будем прогнозировать 10 из них. Подберем параметры ARIMA с помощью GridSearch. Также выберем оптимальный горизонт прогноза для прогноза ARIMA и LSTM с помощью метода сломанной трости. На Рис.3 показана зависимость MAPE от горизонта прогноза для LSTM и ARIMA.

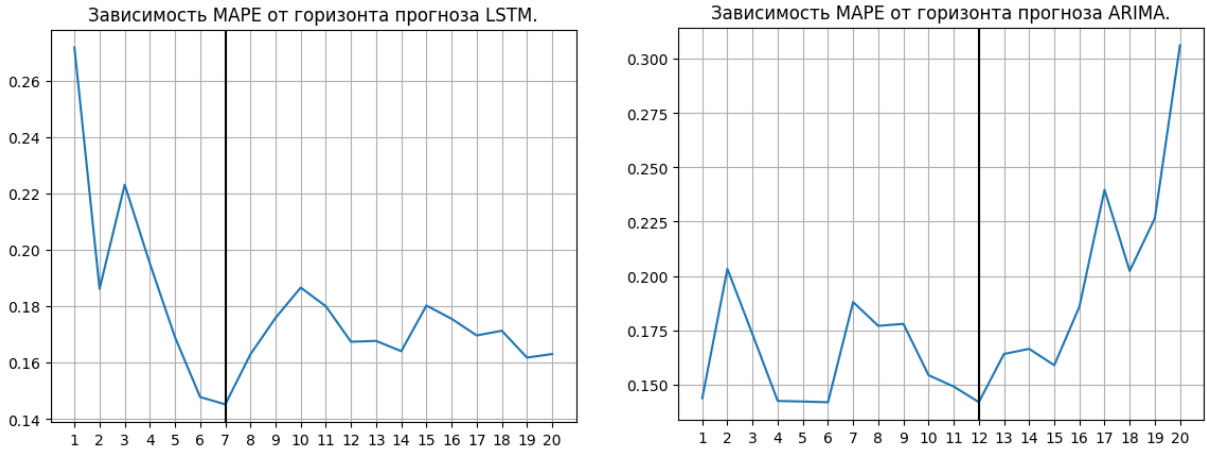


Рис. 1: Зависимость MAPE от горизонта прогнозирования LSTM, ARIMA.

Получаем, что оптимальный горизонт прогноза для LSTM — 7, для ARIMA — 12. В Таблице 2 представлены ошибки прогнозов при оптимальном горизонте прогноза.

$\text{MAPE}_{\text{LSTM}}$	MSE_{LSTM}	$\text{MAPE}_{\text{ARIMA}}$	$\text{MSE}_{\text{ARIMA}}$
0.15	59.14	0.14	76.62

Таблица 1: Ошибки прогноза LSTM и ARIMA на реальном наборе при оптимальном горизонте прогноза.

Можем заметить, что несмотря на больший горизонт прогноза, $\text{MAPE}_{\text{ARIMA}}$ практически совпадает с $\text{MAPE}_{\text{LSTM}}$. Однако точность MSE_{LSTM} лучше, чем $\text{MSE}_{\text{ARIMA}}$.

Прогнозирование с помощью COF.

Матрицу попарных расстояний будем считать по $L = 10$ предыдущим значениям. При прогнозировании матрицы попарных расстояний с помощью ARIMA снова используем GridSearch оптимальных параметров. Определим оптимальный горизонт прогноза матрицы попарных расстояний для LSTM и ARIMA.

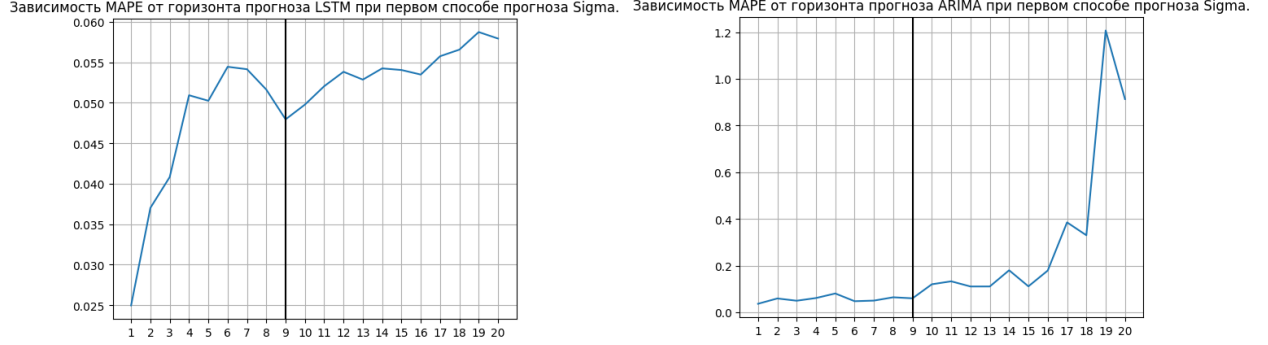


Рис. 2: Зависимость MAPE от горизонта прогнозирования матрицы с помощью LSTM, ARIMA.

В обоих случаях будем прогнозировать матрицы попарных расстояний с горизонтом прогнозирования 9. В Таблице 2 указаны ошибки LSTM, ARIMA прогнозов матрицы попарных расстояний при различных способах ее подсчета. Таким образом, LSTM лучше прогнозирует матрицу Σ . Теперь по каждой спрогнозированной матрице восстановим прогноз исходного набора временных рядов.

	$MAPE_{LSTM}$	MSE_{LSTM}	$MAPE_{ARIMA}$	MSE_{ARIMA}
Σ^1	0.052	30112	0.057	34091
Σ^2	0.990	466	1.680	937
Σ^3	1.720	0.056	2.440	0.157

Таблица 2: Ошибки прогноза матрицы при различных способах вычисления Σ_t .

3.2 Восстановление прогноза

Восстановим прогноз, решив условную задачу оптимизации (12). В Таблице 3 представлены ошибки восстановления прогноза.

	$MAPE_{LSTM}$	MSE_{LSTM}	$MAPE_{ARIMA}$	MSE_{ARIMA}
Σ^1	0.46	359.67	0.36	293.57
Σ^2	0.29	222.46	0.30	225.41
Σ^3	0.43	437.85	0.50	918.49

Таблица 3: Ошибки восстановленного прогноза при различных способах вычисления Σ_t .

Заметим, что и в случае LSTM и в случае ARIMA наименьшая ошибка прогноза в случае второго способа подсчета Σ .

4 Сравнение результатов

Сравним результаты прогноза набора временных рядов с помощью ARIMA, LSTM, COF-ARIMA, COF-LSTM. Приведем результаты в Таблицу 4.

Алгоритм	MAPE	MSE
LSTM	0.15	59.14
COF-LSTM-1	0.46	359.67
COF-LSTM-2	0.29	222.46
COF-LSTM-3	0.43	437.85
ARIMA	0.14	76.62
COF-ARIMA-1	0.36	293.57
COF-ARIMA-2	0.30	225.41
COF-ARIMA-3	0.50	918.49

Таблица 4: Ошибки прогноза набора временных рядов различных методов.

Таким образом, даже лучшие результаты COF алгоритма значительно уступают простым ARIMA, LSTM. Причиной этому служит недостаточно точное предсказание матрицы Σ .

5 Заключение

Предложили новый метод прогнозирования набора высокоррелированных временных рядов, основанного на прогнозировании матрицы попарных расстояний с помощью ARIMA и LSTM и восстановления прогноза с помощью решения задачи оптимизации. Сравнили три способа вычисления матрицы попарных расстояний. Сравнили результаты работы метода с ARIMA и LSTM. Предложенный метод уступает в точности обоим методам. Причиной этому является недостаточная точность прогнозирования матрицы. Одним из способов улучшения точности является использование других методов прогнозирования, например, генеративных моделей.

Список литературы

- [1] Robert H Shumway, David S Stoffer, and David S Stoffer. *Time series analysis and its applications*, volume 3. Springer, 2000.
- [2] Everette S. Gardner Jr. Exponential smoothing: The state of the art. *Journal of Forecasting*, 1985.
- [3] J. Schmidhuber S. Hochreiter. Long short-term memory. *Neural Computation*, 1997.
- [4] Jenkins Box. *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day, 1970.
- [5] Razvan-Gabriel Cirstea, Darius-Valer Micu, Gabriel-Marcel Muresan, Chenjuan Guo, and Bin Yang. Correlated time series forecasting using multi-task deep neural networks. In *Proceedings of the 27th acm international conference on information and knowledge management*, pages 1527–1530, 2018.
- [6] Ana Corberán-Vallet, José D Bermúdez, and Enriqueta Vercher. Forecasting correlated time series with exponential smoothing models. *International Journal of Forecasting*, 27(2):252–265, 2011.
- [7] Steven Diamond Stephen Boyd, Enzo Busseti and Ronald N. Kahn. Multi-period trading via convex optimization. *Foundations and Trends in Optimization*, 2017.
- [8] Benyamin Ghogh, Ali Ghodsi, Fakhri Karray, and Mark Crowley. Reproducing kernel hilbert space, mercer’s theorem, eigenfunctions, nyström method, and use of kernels in machine learning: Tutorial and survey. *arXiv preprint arXiv:2106.08443*, 2021.
- [9] Цена на электричество в Германии. URL: http://svn.code.sf.net/p/mlalgorithms/code/Group474/Uvarov2017SuperpositionForecasting/data/electricity_price_german/GermanSpotPrice.csv.