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

Рассматривается задача прогнозирования финансовых временных рядов. Основными особенностями таких временных рядов являются высокая волатильность и высокая попарная ковариация. Классическим подходом к решению задачи является выполнение прогноза в исходном пространстве. Новый метод заключается в переходе в пространство попарных расстояний между временными рядами, осуществлении прогноза в нем и переходе обратно в исходное пространство. Для его реализации необходимо ввести функцию расстояния между временными рядами (метрику), которая должна удовлетворять определенным свойствам. В данной статье изучаются эти свойства и проводятся сравнения различных метрик на основе численных экспериментов.

Keywords Временные ряды · Метрика · Ковариация

1 Introduction

В текущей статье исследуется задача погружения временных рядов в метрическое пространство. Таким образом, набору временных рядов ставится в соответствие матрица попарных расстояний и появляется возможность перейти от прогнозирования набора временных рядов к прогнозированию матрицы попарных расстояний. При этом выбор метрики осуществляется так, чтобы по полученной матрице расстояний можно было восстановить прогноз для набора временных рядов.

В статистике, обработке сигналов и многих других областях под временным рядом понимаются последовательно измеренные через некоторые (зачастую равные) промежутки времени данные. 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. Типичный пример — предсказание цены открытия биржи основываясь на предыдущей её деятельности.

Одними из хорошо известных, классических методов прогнозирования временных рядов являются экспоненциальное сглаживание (англ. Exponential Smoothing) [1], LSTM (англ. Long Short-Term Memory) [2], ARIMA (англ. autoregressive integrated moving average) [3]. Главным отличием исследуемого метода от вышеперечисленных является то, что временные ряды прогнозируются при помощи прогнозирования матрицы попарных расстояний.

В качестве простейшей метрики рассматривается ковариация между временными рядами. [4] Таким образом, для набора временных рядов получаем матрицу ковариации. Стоит заметить, что матрица ковариации (матрица попарных расстояний) вычисляется в каждый момент времени. Альтернативные варианты метрики выбираются из класса ядер [5].

Численные эксперименты проводятся на трех видах данных: синтетические, сигналы коры головного мозга, финансовые временные ряды. Эксперимент состоит из выполнения прогноза временного ряда при помощи прогнозирования матрицы попарных расстояний. В качестве прогностической модели выбирается линейная регрессия с использованием SSA (англ. Singular Spectrum Analysis) [6]. По результатам экспериментов проводится анализ точности прогноза и его устойчивости в зависимости от выбранной метрики и вида данных. Цель эксперимента состоит в оптимальном выборе функции попарных расстояний для выполнения прогноза.

2 Problem statement

Пусть $\mathrm{X} = \{\mathbf{x} = [x_1, \dots, x_\mathrm{N}] | x_i \in R\}$ — множество временных рядов, заданных своей реализацией. Обозначим через

$$\chi = [x^{(1)}, \dots, x^{(n)}]^T \tag{1}$$

заданный набор из n временных рядов. Через $\chi_t = [x_t^{(1)}, \dots, x_t^{(n)}]^T$ обозначим значение ряда χ в момент времени t. Определим функцию расстояния между временными рядами: $d : \mathrm{X} \times \mathrm{X} \to R$, удовлетворяющую условиям Мерсера [7]. Обозначим расстояние между временными рядами $x^{(i)} = [x_1^{(i)}, \dots, x_t^{(i)}], x^{(j)} = [x_1^{(j)}, \dots, x_t^{(j)}]$ следующим образом:

$$d(x^{(i)}, x^{(j)}) = d_t(i, j) \tag{2}$$

Таким образом, в каждый момент времени t ряду X поставлена в соответствие матрица попарных расстояний $\Sigma_t \in \mathcal{S}_n^+$ (симметричная, положительно полуопределенная матрица):

$$\chi_t \Rightarrow \Sigma_t = \begin{pmatrix} d_t(x^{(1)}, x^{(1)}) & d_t(x^{(1)}, x^{(2)}) & \dots & d(x_t^{(1)}, x^{(n)}) \\ d_t(x^{(2)}, x^{(1)}) & d_t(x^{(2)}, x^{(2)}) & \dots & d_t(x^{(2)}, x^{(n)}) \\ \vdots & \vdots & \ddots & \vdots \\ d_t(x^{(n)}, x^{(1)}) & d_t(x^{(n)}, x^{(2)}) & \dots & d_t(x^{(n)}, x^{(n)}) \end{pmatrix} \tag{3}$$

Дадим описание двум используемым прогностическим моделям:

2.1 Прогнозирование

2.1.1 Авторегрессия

Пусть имеем реализацию матрицы попарных расстояний из N компонент: $\Sigma = [\Sigma_1, \Sigma_2, \dots, \Sigma_\mathrm{N}]$. Пусть $\Sigma_{\mathrm{N}+1}$ - прогнозируемая матрица. Авторегрессионная модель:

$$\hat{\Sigma}_{\mathrm{N}+1} = \sum_{i=1}^{N} a_i \Sigma_i + \epsilon_{\mathrm{N}+1} \tag{4}$$

где $a_1, \dots, a_\mathrm{N}$ — параметры модели (коэффициенты авторегрессии), $\epsilon_{\mathrm{N}+1} \in R^{\mathrm{N}\times\mathrm{N}}$ — белый шум. Вектор параметров определяется с помощью решения задачи оптимизации:

$$a = \arg\min_a \|\Sigma_{\mathrm{N}+1} - \hat{\Sigma}_{\mathrm{N}+1}\|_F^2 \tag{5}$$

2.1.2 Многомерная гусеница (MSSA)

Пусть имеем реализацию матрицы попарных расстояний из N компонент: $\Sigma = [\Sigma_1, \Sigma_2, \dots, \Sigma_\mathrm{N}]$. Определим L - ширина окна, $\mathrm{K} = \mathrm{N} - \mathrm{L} + 1$. В силу того, что матрица симметричная, то для каждой компоненты матрицы попарных расстояний, где $i \leqslant j$, построим матрицу траекторий (матрица Генкеля):

$$\mathrm{H}_{ij}=\begin{pmatrix} d_1(x^{(i)},x^{(j)}) & d_2(x^{(i)},x^{(j)}) & \dots & d_{\mathrm{K}}(x^{(i)},x^{(j)}) \\ d_2(x^{(i)},x^{(j)}) & d_3(x^{(i)},x^{(j)}) & \dots & d_{\mathrm{K}+1}(x^{(i)},x^{(j)}) \\ \vdots & \vdots & \ddots & \vdots \\ d_{\mathrm{L}}(x^{(i)},x^{(j)}) & d_{\mathrm{L}+1}(x^{(i)},x^{(j)}) & \dots & d_{\mathrm{N}}(x^{(i)},x^{(j)}) \end{pmatrix}. \tag{6}$$

Строим блочную матрицу траекторий:

$$\mathrm{H}=[\mathrm{H}_{11}:\mathrm{H}_{12}:\mathrm{H}_{13}:\mathrm{H}_{1s}:\mathrm{H}_{22}:\ldots:\mathrm{H}_{ss}] \tag{7}$$

Сингулярное разложение симметричной, положительно полуопределенной матрицы $\mathrm{HH}^T=\mathrm{U}\Lambda\mathrm{U}^T$, где Λ - диагональная матрица $L\times L$ собственных значений $\lambda_1\geq\ldots\geq\lambda_L$, $\mathrm{U}=[\mathrm{u}_1,\ldots,\mathrm{u}_L]$ - ортогональная матрица собственных векторов матрицы HH^T. Пусть $d=\max\{i|\lambda_i>0, i\in\{1,\ldots,\mathrm{L}\}\}$. Тогда исходную матрицу траекторий можно представить в виде:

$$\mathrm{H}=\sum_{i=1}^{d}\sqrt{\lambda_i}u_iv_i^T=\mathrm{H}^{(1)}+\ldots+\mathrm{H}^{(d)}=\sum_{i=1}^{r}\mathrm{H}^{(i)}+\sum_{i=r+1}^{d}\mathrm{H}^{(i)}, \tag{8}$$

где $v_i=H^Tu_i/\sqrt{\lambda_i}$ $\mathrm{H}_r=\sum\limits_{i=1}^{r}\mathrm{H}^{(i)}$ — сигнальные компоненты, а $\mathrm{H}-\mathrm{H}_r$ — шумовые. Полученную матрицу можно представить в виде:

$$\mathrm{H}_r=[\widetilde{\mathrm{H}}_{11}:\widetilde{\mathrm{H}}_{12}:\ldots:\widetilde{\mathrm{H}}_{ss}] \tag{9}$$

где каждая из матриц $\widetilde{\mathrm{H}}_{ij}\in R^{\mathrm{N}-\mathrm{L}+1\times\mathrm{L}}$. Восстановим исходные расстояния с помощью антидиагонального усреднения элементов соответствующих матриц:

$$\widetilde{d}_t(i,j)=\frac{1}{2t-1}\sum_{k,l:k+l-1=t}\widetilde{h}_{kl}^{(ij)}, t\in\{1,\ldots,N\} \tag{10}$$

где $\widetilde{h}_{kl}^{(ij)}$ — элемент на пересечении k-ой строки и l-го столбца матрицы H_{ij}. Обозначим через $u_j^{\triangledown}$ L - 1 первые компоненты собственного вектора u_j, а π_j — последнюю компоненту вектора u_j, $(j=1,\ldots,r)$. Определим $v^2=\sum\limits_{j=1}^{r}\pi_j^2$ и:

$$\mathcal{R}=\frac{1}{1-v^2}\sum_{j=1}^{r}\pi_ju_j^{\triangledown}. \tag{11}$$

При $v^2<1$ возможен прогроз:

$$[d_{\mathrm{N}+1}(1,1),d_{\mathrm{N}+1}(1,2),\ldots,d_{\mathrm{N}+1}(s,s)]=\mathcal{R}^TZ \tag{12}$$

где $Z=[Z^{(11)},\ldots,Z^{(ss)}]^T$, $Z^{(ij)}=[\widetilde{d}_{\mathrm{N}-\mathrm{L}+2}(i,j),\ldots,\widetilde{d}_{\mathrm{N}}(i,j)]$.

2.2 Выбор функции попарных расстояний

Пусть $\mathcal{F}=\{d^{(1)},\ldots,d^{(s)}\}$ — 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:

$$d_{\mathrm{acc}}=\arg\min_{d\in\mathcal{F}}\sigma^2(d)=\arg\min_{d\in\mathcal{F}}\|\hat{\Sigma}(d)-\Sigma(d)\|_{\mathrm{F}}^2. \tag{13}$$

Пусть $\widetilde{x}^{(i)}=x^{(i)}+\varepsilon_i$, $i=1,\ldots,n$ — зашумленный исходный набор временных рядов, где $\varepsilon_i\in R^{\mathrm{N}}$. Пусть $\widetilde{\sigma}^2(d)$ — точность прогноза матрицы попарных расстояний зашумленного набора временных рядов при использовании $d\in\mathcal{F}$. Тогда наболее устойчивый выбор функции попарных расстояний:

$$d_{\text{stable}} = \arg\min_{d \in \mathrm{F}} |\widetilde{\sigma}^2(d) - \sigma^2(d)| \tag{14}$$

Список литературы

[1] Everette S. Gardner Jr. Exponential smoothing: The state of the art. Journal of Forecasting, 1985.

[2] J. Schmidhuber S. Hochreiter. Long short-term memory. Neural Computation, 1997.

[3] Jenkins Box. Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day, 1970.

[4] Steven Diamond Stephen Boyd, Enzo Busseti and Ronald N. Kahn. Multi-period trading via convex optimization. Foundations and Trends in Optimization, 2017.

[5] John Shawe-Taylor and Nello Cristianini. Kernel methods for pattern analysis. Cambridge university press, 2004.

[6] Robert Vautard, Pascal Yiou, and Michael Ghil. Singular-spectrum analysis: A toolkit for short, noisy chaotic signals. Physica D: Nonlinear Phenomena, 58(1-4):95–126, 1992.

[7] Benyamin Ghojogh, Ali Ghodsi, Fakhri Karray, and Mark Crowley. Reproducing kernel hilbert space, mercer's theorem, eigenfunctions, nystr\"om method, and use of kernels in machine learning: Tutorial and survey. arXiv preprint arXiv:2106.08443, 2021.