

Mathematical Forecasting Methods

Лекция 6

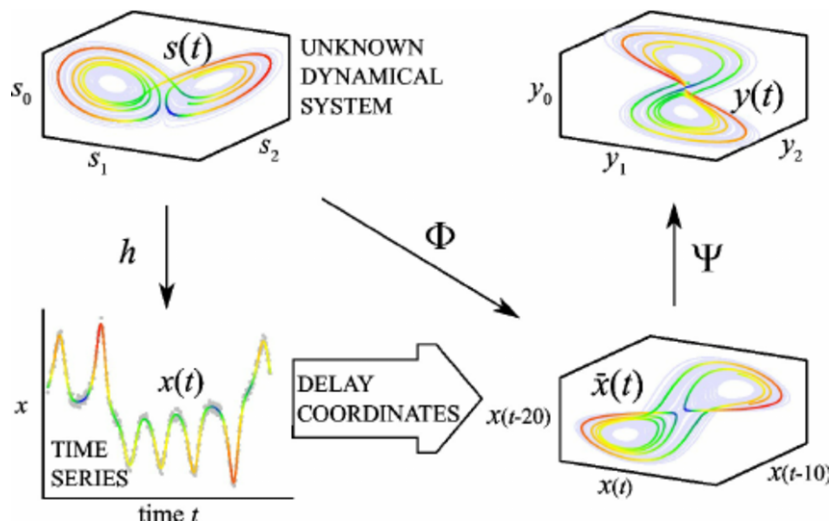
МФТИ

Осень, 2023

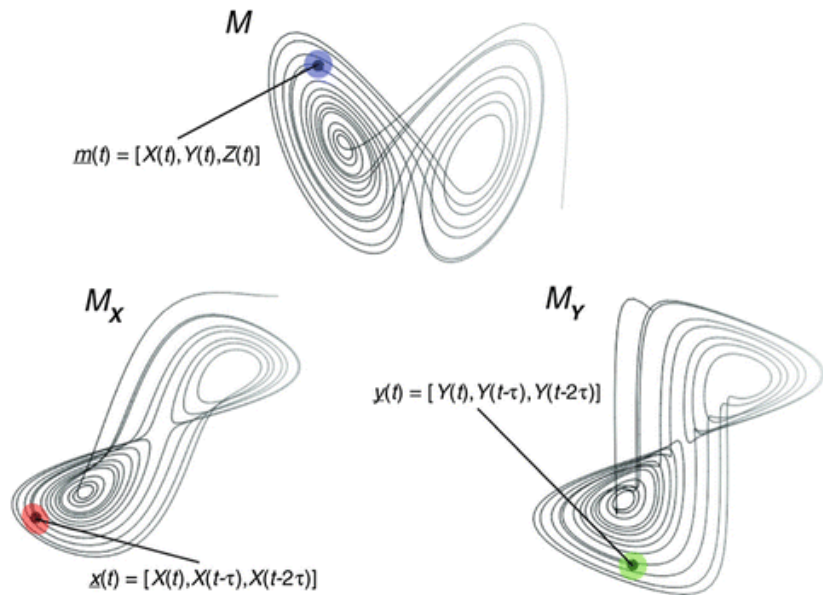
Краткое повторение

- ▶ Математические основы динамических систем позволяют осуществлять анализ временных рядов.
- ▶ Теорема Такенса позволяет использовать вектора задержек для восстановления внутренней структуры динамической системы.
- ▶ При выполнении условия $m \geq 2d + 1$, где d — размерность вложения, возможно реконструировать пространство состояний системы.
- ▶ В частности, выводы теоремы Такенса используются в ССМ. Алгоритм ССМ является аналогом статистического теста и оценивает причинно-следственную связь двух временных рядов.

Краткое повторение



Cross Convergent Mapping



Краткое повторение

1. Пусть имеются 2 временных ряда $\{x_1, x_2, \dots, x_T\}$ и $\{y_1, y_2, \dots, y_T\}$ длины T .
2. Для ряда $\{x_t\}_{t=1}^T$ формируем векторы предыстории размерности E с шагом по времени τ :

$$\mathbf{x}_t = (x_t, x_{t-\tau}, x_{t-2\tau}, \dots, x_{t-(E-1)\tau})$$

3. В пространстве \mathbb{R}^E (*фазовом пространстве*) такие векторы предыстории образуют *фазовую траекторию* системы – множество $M_x = \{\mathbf{x}_t \mid t = 1 + (E-1)\tau, \dots, T\}$.
4. Пусть требуется построить предсказание для значения y_t . Для этого сначала найдем $E+1$ векторов из M_x , ближайших к \mathbf{x}_t (в терминах, например, стандартной метрики d в \mathbb{R}^n). Пусть им соответствуют временные индексы t_1, \dots, t_{E+1} , отсортированные в порядке от ближайшей точки до наиболее удалённой.

$$d_i = d(\mathbf{x}_t, \mathbf{x}_{t_i}), \quad d_1 < d_2 < \dots < d_{E+1}$$

Алгоритм Convergent Cross Mapping

5. Тогда оценка значения y_t строится следующим образом в виде взвешенной суммы значений ряда в моменты времени t_1, \dots, t_{E+1} :

$$\hat{y}_t | M_x = \sum_{i=1}^{E+1} \omega_i y(t_i)$$

$$\omega_i = \frac{u_i}{\sum_{j=1}^{E+1} u_j}, \quad u_i = \exp\left(-\frac{d_i}{d_1}\right), \quad i = 1, \dots, E + 1$$

6. Для того, чтобы посудить о существовании зависимости между рядами $\{x_t\}_{t=1}^T$ и $\{y_t\}_{t=1}^T$, рассчитаем теперь коэффициент корреляции Пирсона:

$$C_{yx} = \left[\rho(y, \hat{y} | M_x) \right]^2$$

Применение рядов Фурье для анализа временных рядов

Исследуется функция $x(t)$, заданная на отрезке $[0, T]$.

Фурье-анализ позволяет представить функцию $x(t)$ в виде:

$$x(t) = a_0 + \sum_{k=1}^{+\infty} a_k \cos\left(2\pi \frac{k}{T} t + \theta_k\right)$$

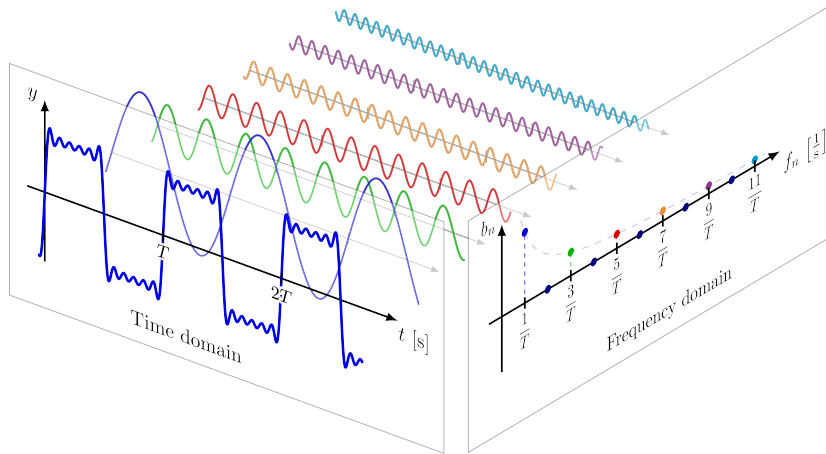
T – отрезок, где функция определена (длительность сигнала)

a_k – амплитуда k -ой гармонической составляющей,

θ_k – начальная фаза k -ой гармонической составляющей.

Разложение в ряд Фурье позволяет представить непрерывную функцию $x(t)$, определенную на отрезке $[0, T]$ в виде бесконечного ряда тригонометрических функций с определёнными амплитудами и фазами, также рассматриваемых на отрезке $[0, T]$.

Применение рядов Фурье для анализа временных рядов



Глобальные и локальные методы прогнозирования

В анализе и прогнозировании временных рядов можно выделить две группы методов:

- ▶ глобальные – параметры аппроксимирующей функции находятся посредством использования всех известных значений ряда; основное приложение – получение глобальных характеристик системы (ARMA, SSA)
- ▶ локальные методы основаны на принципе локальной аппроксимации (LA)

Глобальные и локальные методы прогнозирования

Преимущества локальных методов прогноза нерегулярных (хаотических, квазипериодических) рядов:

- ▶ применение не требует априорной информации о системе, породившей временной ряд
- ▶ нет необходимости в построении специфической модели динамики исследуемого ряда
- ▶ использования для прогнозирования наиболее близких к стартовой точке (стартовому вектору) значений ряда
- ▶ использование меньшего количества исходных данных
- ▶ меньше ограничений на горизонт прогнозирования

Матричное представление временного ряда

- ▶ Вектор задержек:

$$\mathbf{x}_t = (x_t, x_{t+1}, \dots, x_{t+\tau-1})^T \in \mathbb{R}^\tau$$

- ▶ Матрица задержек:

$$\begin{aligned} \mathbf{X} &= (\mathbf{x}_1 \quad \mathbf{x}_2 \quad \mathbf{x}_3 \quad \dots \quad \mathbf{x}_\tau \quad \dots \quad \mathbf{x}_n) = \\ &= \begin{pmatrix} x_1 & x_2 & x_3 & \dots & x_\tau & \dots & x_n \\ x_2 & x_3 & x_4 & \dots & x_{\tau+1} & \dots & x_{n+1} \\ x_3 & x_4 & x_5 & \dots & x_{\tau+2} & \dots & x_{n+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_\tau & x_{\tau+1} & x_{\tau+2} & \dots & x_{2\tau-1} & \dots & x_N \end{pmatrix} \in \mathbb{R}^{\tau \times n} \end{aligned}$$

Здесь предполагаем, что $n > \tau$.

SVD: напоминание

Для матрицы \mathbf{X} существует (и, вообще говоря, не единственно) сингулярное разложение:

$$\underset{\tau \times n}{\mathbf{X}} = \underset{\tau \times \tau}{U} \underset{\tau \times n}{\Sigma} \underset{n \times n}{V^T}$$

Здесь:

- ▶ $U \in \mathbb{R}^{\tau \times \tau}$ – ортогональная матрица собств. векторов $\mathbf{X}\mathbf{X}^T$ (столбцы U образуют ортонормированный базис в \mathbb{R}^{τ})
- ▶ $\Sigma \in \mathbb{R}^{\tau \times n}$ – прямоугольная диагональная матрица сингулярных чисел \mathbf{X} , причем $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{\tau}$:

$$\Sigma = \begin{pmatrix} \sigma_1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \dots & \vdots \\ 0 & \dots & \sigma_{\tau} & 0 & \dots & 0 \end{pmatrix}$$

- ▶ $V \in \mathbb{R}^{n \times n}$ – ортогональная матрица собств. векторов $\mathbf{X}^T\mathbf{X}$ (столбцы V образуют ортонормированный базис в \mathbb{R}^n)

Low Rank Approximation

Одним из важных свойств сингулярного разложения матрицы является возможность построения для неё *низкорангового приближения*, а именно:

- ▶ Пусть $A \in \mathbb{R}^{m \times n}$, $A = U\Sigma V^T$ – сингулярное разложение A .
- ▶ Рассмотрим матрицу Σ и её приближение ранга r – матрицу $\tilde{\Sigma}$, где подматрица Σ_1 имеет размер $r \times r$:

$$\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix}, \quad \tilde{\Sigma} = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix}$$

- ▶ **Теорема:** среди всех матриц B ранга не более r норма разности матриц A и $\tilde{A} = U\tilde{\Sigma}V^T$ минимальна:

$$\tilde{A} = \arg \min_{\text{rank}(B) \leq r} \|A - B\| \quad (\text{верно для } \|\cdot\|_2, \text{ для } \|\cdot\|_F)$$

- ▶ Матрица \tilde{A} также может быть получена, как $\tilde{A} = U_1 \Sigma_1 V_1^T$, где U_1 , V_1 – первые r столбцов матриц U и V .

Метод SSA и алгоритм SSA-сглаживания

Метод SSA (Singular Spectral Analysis) основан на использовании сингулярного разложения матрицы задержек для анализа временного ряда.

Алгоритм SSA-сглаживания:

1. Для матрицы задержек \mathbf{X} построим её низкоранговое приближение $\tilde{\mathbf{X}}$ с помощью SVD.
2. Из элементов матрицы $\tilde{\mathbf{X}}$ получим сглаженные оценки значений временного ряда x_1, \dots, x_N . Заметим, что в исходной матрице эти значения стоят на нескольких позициях одновременно:

$$\mathbf{X} = \begin{pmatrix} x_1 & x_2 & x_3 & \dots & x_\tau & \dots & x_n \\ x_2 & x_3 & x_4 & \dots & x_{\tau+1} & \dots & x_{n+1} \\ x_3 & x_4 & x_5 & \dots & x_{\tau+2} & \dots & x_{n+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_\tau & x_{\tau+1} & x_{\tau+2} & \dots & x_{2\tau-1} & \dots & x_N \end{pmatrix}$$

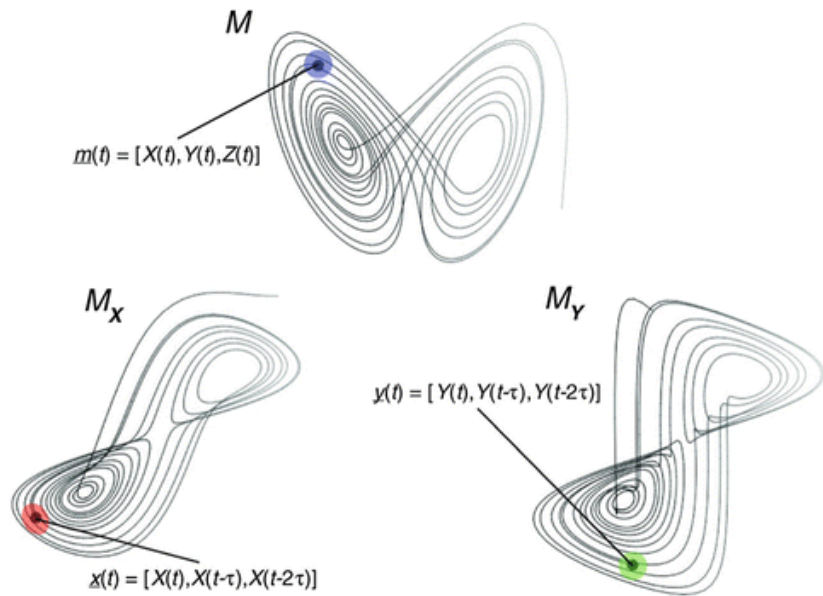
Метод SSA и алгоритм SSA-сглаживания

3. Отсюда получаем формулу для сглаженных оценок:

$$\hat{x}_s = \begin{cases} \frac{1}{s} \sum_{i=1}^s \tilde{x}_{i,s-i+1}, & 1 \leq s \leq \tau \\ \frac{1}{\tau} \sum_{i=1}^{\tau} \tilde{x}_{i,s-i+1}, & \tau \leq s \leq n \\ \frac{1}{N-s+1} \sum_{i=1}^{N-s+1} \tilde{x}_{i+s-n,n-i+1}, & n \leq s \leq N \end{cases}$$

Такое сглаживание помогает снизить шум.

Cross Convergent Mapping



Метод Local Approximation (LA)

1. Перейдём к векторному представлению временного ряда с помощью векторов задержек и матрицы задержек. В этом методе, однако, удобно изменить порядок следования элементов вектора по времени – имеем следующее обозначение:

$$\mathbf{x}_t = (x_t, x_{t-1}, \dots, x_{t-\tau+1})^T \in \mathbb{R}^\tau$$

2. Матрица задержек:

$$\mathbf{X} = (\mathbf{x}_\tau \quad \mathbf{x}_{\tau+1} \quad \dots \quad \mathbf{x}_{N-2} \quad \mathbf{x}_{N-1} \quad \mathbf{x}_N) \in \mathbb{R}^{\tau \times (N-\tau+1)}$$

3. Пусть требуется построить предсказание значения временного ряда x_{N+1} . Возьмём предшествующий данному моменту вектор задержек \mathbf{x}_N .
4. Зададимся параметром S и найдём в матрице \mathbf{X} S векторов – ближайших соседей для \mathbf{x}_N в терминах, например, евклидовой нормы. Обозначим их $\mathbf{x}_{i_1}, \dots, \mathbf{x}_{i_S}$.

Метод Local Approximation (LA)

5. Предсказание для значения x_{N+1} построим с помощью параметризованной функции f :

$$x_{N+1} = f(\mathbf{x}_N | \theta)$$

Например, линейная аппроксимация: $f(\mathbf{x} | \theta) = \theta_0 + \theta_1^T \mathbf{x}$.
Необходимо лишь определить оптимальный набор параметров θ .

6. Параметры θ определим по локальной окрестности вектора \mathbf{x}_N в пространстве векторов задержек:

$$\hat{\theta} = \arg \min \sum_{s=1}^S \left(x_{i_s+1} - f(\mathbf{x}_{i_s} | \theta) \right)^2$$

7. Окончательно, предсказание для $N + 1$ -ого значения:

$$x_{N+1} = f(\mathbf{x}_N | \hat{\theta})$$

- ▶ Временной ряд можно аппроксимировать отрезком его ряда Фурье.
- ▶ Вектор задержек и матрица задержек – это удобные способы многомерного представления временного ряда.
- ▶ Метод SSA основан на сингулярном разложении матрицы задержек ряда и, как правило, переходе к её низкоранговому приближению.
- ▶ Метод LA позволяет предсказывать значения временного ряда по окрестности соответствующего вектора задержек в своём пространстве (идея, схожая с алгоритмом CCM).