Прогнозирование временного ряда.

Рассмотрим временной ряд $F_N = (f_0, \ldots, f_{N-1})$ с $N \geq 3$ и зафиксируем длину окна L $(1 < L < N)$.

В результате процедуры вложения мы получаем последовательность векторов вложения:

$$X_i^{(L)} = X_i = (f_{i-1}, \ldots, f_{i+L-2})^{\mathrm{T}}, \qquad i = 1, \ldots, K,$$

Итерационный метод.

Линейная рекуррентная формула (ЛРФ):

$$f_{i+d} = \sum_{k=1}^{d} a_k f_{i+d-k}, \qquad 0 \leq i \leq N-d-1, \quad a_d \neq 0.$$

Перейдем к прогнозированию временных рядов методом гусеницы. Для начала определимся с тем, что мы будем понимать под продолжением ряда. Числовой ряд $\{x_i\}_{i=1}^{n+1}$ называется продолжением ряда $\{x_i\}_{i=1}^{n}$, если порождаемая им при гусеничной обработке выборка лежит в той же гиперплоскости, что и у исходного ряда.

Из собственных векторов $(U^{<1>}, U^{<2>}, \ldots, U^{<L>})$ матрицы $S = XX^T$ можем выбрать базис. Выбираются векторы, соответствующие максимальным собственным значениям матрицы S. Обозначим матрицу из базисных векторов через $M = [U^{<1>}, \ldots, U^{<r>}]$.

Любой вектор из ганкелевой матрицы $X^{<i>}$ можно разложить по выбранному базису

$$X^{<i>} = p_1^i U^{<1>} + p_2^i U^{<2>} + \cdots + p_r^i U^{<r>}$$

Матричная запись $\quad Mp^{<i>} = X^{<i>}$

Вектор параметров :

$$p^{<i>} = \begin{pmatrix} p_1^i \\ p_2^i \\ \vdots \\ p_r^i \end{pmatrix}$$

можно найти по формуле

$$p^{<i>} = (M^T M)^{-1} M^T X^{<i>}.$$

Траекторная матрица принимает

$$X = \begin{pmatrix} f_1 & f_2 & \cdots & f_{N-L+1} \\ f_2 & f_3 & \cdots & \boxed{f_{N-L+2}} \\ \vdots & \vdots & \ddots & \vdots \\ f_L & f_{L+1} & \cdots & f_N \end{pmatrix} \qquad \begin{pmatrix} f_{N-L+2} \\ \cdots \\ f_{N+1} \end{pmatrix}$$

Для предсказания значения f_{N+1} рассмотрим вектор столбец

$$X^{N-L+2} = (f_{N-L+1}, \cdots, f_{N+1})^T.$$

Разложение вектора X^{N-L+2} примет вид

$$[U^1 U^2 \ldots U^r] \begin{pmatrix} p_1^{N-L+2} \\ \cdots \\ p_r^{N-L+2} \end{pmatrix} = \begin{pmatrix} f_{N-L+2} \\ \cdots \\ f_{N+1} \end{pmatrix}$$

Так как число уравнений больше числа неизвестных , то мы можем найти вектор параметров $(p_1^i,\ p_2^i, \ldots, p_r^i)$ из системы уравнений

$$\begin{bmatrix} U_1^1 & U_1^2 & \cdots & U_1^r \\ \vdots & \vdots & \cdots & \vdots \\ U_{L-1}^1 & U_{L-1}^2 & \cdots & U_{L-1}^r \end{bmatrix} \begin{pmatrix} p_1^i \\ \cdots \\ p_r^i \end{pmatrix} = \begin{pmatrix} X_1^i \\ \cdots \\ X_{L-1}^i \end{pmatrix}$$

Систему назовем усеченной системой.

$$V_* = \begin{bmatrix} U_1^1 & U_1^2 & \cdots & U_1^r \\ \vdots & \vdots & \cdots & \vdots \\ U_{L-1}^1 & U_{L-1}^2 & \cdots & U_{L-1}^r \end{bmatrix}$$

Параметры ($p_1^{N-L+2} \ldots p_r^{N-L+2}$) найдем из усеченной системы

$$V_* \begin{pmatrix} p_1^{N-L+2} \\ \cdots \\ p_r^{N-L+2} \end{pmatrix} = \begin{pmatrix} f_{N-L+2} \\ \cdots \\ f_N \end{pmatrix}$$

$$p^{<i>} = (V_*^T V_*)^{-1} V_*^T X^{<i>}$$

Тогда предсказанное значение найдем из соотношения

$$f_{N+1} = \begin{bmatrix} U_{L-1}^1 & U_{L-1}^2 & \cdots & U_{L-1}^r \end{bmatrix} \begin{pmatrix} p_1^{N-L+2} \\ \cdots \\ p_r^{N-L+2} \end{pmatrix}$$

- обобщенное продолжение рассматриваемого ряда.

Повторить .

«Гусеница»-SSA – универсальный метод для решения задач общего назначения, таких как выделение тренда, обнаружение периодичностей, корректировка на сезонность, сглаживание, подавление шума.

Среди прочего можно отметить:

- Поиск трендов разного разрешения;
- Сглаживание;
- Извлечение сезонных составляющих;
- Одновременное извлечение циклов с малыми и большими периодами;
- Выделение периодичностей с разной амплитудой;
- Одновременное извлечение сложных трендов и периодичностей;
- Поиск структуры в коротких временных рядах.

Прогнозирование временных рядов.
Обработка пропущенных значений.