Временные ряды

Введение в основы и тд

Виктор Горячко
13 ноября 2017 г.
НИУ ВШЭ ФКН

Содержание

Временной ряд

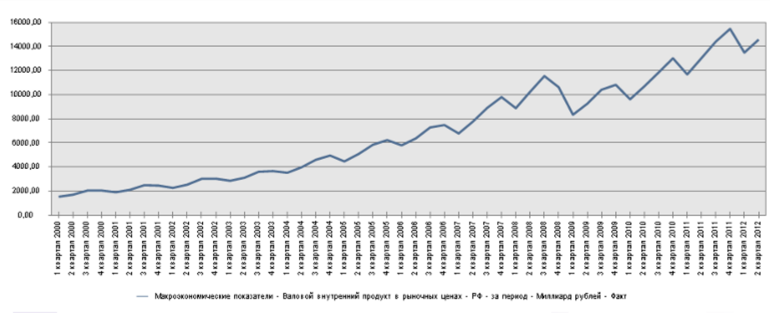


Рис. 1: Пример временного ряда

$y_0, y_1, \dots, y_t, \dots$— временной ряд, где $y_i \in R$
$\hat{y}_{i+d}(w) = f_t(y_1, \dots, y_t; w)$— модель временного ряда, w - вектор параметров модели

Примеры задач

Временные ряды встречаются в различных задачах, например:

- рыночные цены
- объем продаж в торговых сетях
- объемы потребления и цены чего-либо
- дорожный трафик
- уровень сахара крови человека

Временной ряд

Выделяют 4 явления во временных рядах:

- тренды – плавно изменяющаяся компонента, описывающая чистое влияние долговременных факторов;
- сезонность – состоит из последовательности почти повторяющихся циклов;
- разладки – смена модели ряда.
- случайность - содержание случайной компоненты

Временной ряд

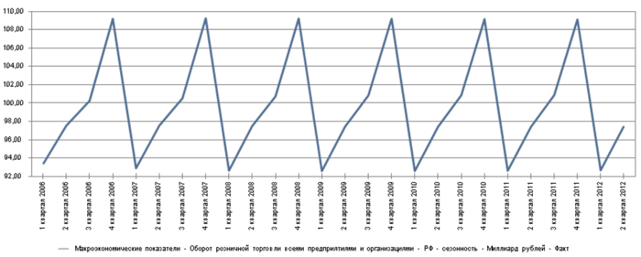
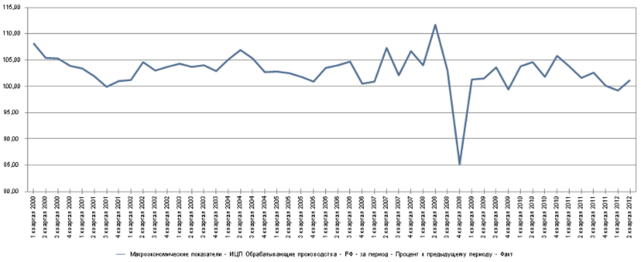


Рис. 2: Пример сезонного временного ряда



Временной ряд

Стадии анализа ВР

- графическое представление и описание поведения ВР;
- выделение и удаление закономерных составляющих ВР, зависящих от времени: тренда, сезонных и циклических составляющих;
- выделение и удаление низко- или высокочастотных составляющих процесса (фильтрация);
- исследование случайной составляющей ВР, оставшейся после удаления перечисленных выше составляющих;
- построение математической модели для описания случайной составляющей и проверка ее адекватности;
- прогнозирование будущего развития процесса, представленного ВР;
- исследование взаимодействий между различными ВР.

Методы анализа временных рядов

- Авторегрессионные модели
- ARMA, ARIMA, GARCH,...
- Нейросетевые модели
- Адаптивные методы краткосрочного прогнозирования
- Прогнозирование плотности
- Квантильная регрессия
- Спектральный анализ
- Скрытые марковские модели
- Фильтр Калмана

Экспоненциальное скользящее среднее

Простая регрессионная модель $\hat{y}_{t+1} = c$
$\sum_{i=0}^{t} \beta^{t-i} \cdot (y_i - c)^2 \to min_c$
$\beta \in (0, 1)$
Найдем решение.
$c = \hat{y}_{t+1} = \frac{\sum_{i=0}^{t} \beta^i y_{t-i}}{\sum_{i=0}^{t} \beta^i}$
Аналогично можно найти для $\hat{y}_t$ и записать рекуррентную формулу для ЭСС
$\hat{y}_{t+1} = \hat{y}_t + \alpha(y_t - \hat{y}_t) = \alpha y_t + (1 - \alpha)\hat{y}_t$, где $\alpha \in (0, 1)$ параметр сглаживания.
При $\alpha \to 1$ получим $\hat{y}_{t+1} = y_t$, а при $\alpha \to 0$ получаем $\hat{y}_{t+1} = \bar{y}$

Модель Хольта

Линейны тренд без сезонных эффектов

$\hat{y}_{t+d} = a_t + b_t d,$

где a_t, b_t - коэффициенты линейного тренда

Рекуррентная формула

$a_t = \alpha_1 y_t + (1 - \alpha_1)(a_{t-1} + b_{t-1})$

$b_t = \alpha_2(a_t - a_{t-1}) + (1 - \alpha_2) b_{t-1}$

Модель Тейла-Вейджа

Линейный тренд с аддитивной сезонностью периода s:

$$\hat{y}_{t+d} = (a_t + b_t d) + \theta_{t+(d mod s)-s}$$

$a_t + b_t d$ - тренд
$\theta_1, ...\theta_{s-1}$ - сезонный профиль периода s
Рекуррентная формула:
$a_t = \alpha_1(y_t - \theta_{t-s}) + (1 - \alpha_1)(a_{t-1} + b_{t-1})$
$b_t = \alpha_2(a_t - a_{t-1}) + (1 - \alpha_2)b_{t-1}$ $\theta_t = \alpha_3(y_t - a_t) + (1 - \alpha_3)\theta_{t-s}$

Модель Уинтерса

Мультипликативная сезонность периода s:
$\hat{y}_{t+d} = a_t \cdot \theta_{t+(d mod s)-s}$
Рекуррентная формула
$a_t = \alpha_1(\frac{y_t}{\theta_{t-s}}) + (1-\alpha_1)a_{t-1}$
$b_t = \alpha_2(\frac{y_t}{a_t}) + (1-\alpha_2)\theta_{t-s}$

Модели с трендом и сезонностью

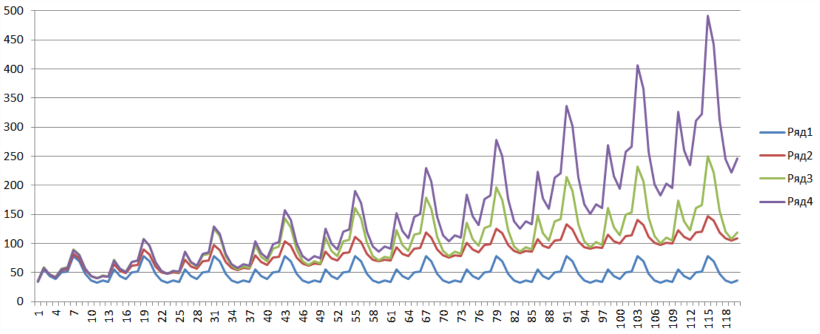


Рис. 4: Модели с трендом и сезонностью

1 - сезонный 2 - аддитивная сезонность (Модель Тейла-Вейджа) 3 - мультипликативная сезонность, линейный тренд (Модель Уинтерса) 4 - мультипликативная сезонность, экспоненциальный тренд

Прогнозирование плотности распределения

В данном методе требуется спрогнозировать числовую величину, оценив её эмпирическое распределение. Это существенно отличается от стандартных задач прогнозирования, в которых вычисляется либо точечная, либо интервальная оценка прогнозируемой величины. Можно использовать следующие методы для нахождения оценки плотности распределения:

- вероятностном интеграле преобразования (Probability integral transform, PIT).
- уравнения Фоккера–Планка
- Алгоритм CHS (Construction of Horizon Se- ries)
- Аппроксимация условной плотности на основе вейвлет-преобразования

Квантильная регрессия

Пусть функция потерь имеет следующий вид:

$$F(y_t, \hat{y}_t) = \begin{cases} a(y_t - \hat{y}_t) & y_t \geq \hat{y}_t \\ b(y_t - \hat{y}_t) & y_t \leq \hat{y}_t \end{cases}$$

Тогда чтобы найти $\hat{y}_{t+1}$ нужно минимизировать $\sum_{i:y_i \geq \hat{y}_{t+1}} w^{t-i}\theta|y_i - \hat{y}_{t+1}| + \sum_{i:y_i \leq \hat{y}_{t+1}} w^{t-i}(1-\theta)|y_i - \hat{y}_{t+1}| \to min_{\hat{y}_{t+1}}$ где $\theta = \frac{a}{a+b}$

Спектральный анализ

Решить задачу линейной множественной регрессии, где зависимая переменная -наблюдаемый временной ряд, а независимые переменные или регрессоры: функции синусов всех возможных (дискретных) частот. Такая модель линейной множественной регрессии может быть записана как: $y_t = a_0 + \sum_{k=1}^{q}(a_k \cos(\lambda_k t) + b_k \sin(\lambda_k t))$ где $\lambda = 2\pi\nu_k$, коэффициенты a_k при косинусах и коэффициенты b_k при синусах - это коэффициенты регрессии, показывающие степень, с которой соответствующие функции коррелируют с данными.

Скрытые марковские модели

$X = x_1, \ldots, x_n$ множество состояний
$V = v_1, \ldots v_m$ алфавит, из которого выбираем наблюдаемые y
q_t состояние во время t
y_t наблюдаемая во время t
$a_{ij} = p(q_{t+1} = x_j | q_t = x_i)$ вероятность перехода из i в j
$b_j(k) = p(v_k | x_j)$ вероятность получить v_k в состоянии j
Начальное распределение $\pi = \pi_j$, где $= p(q_1 = x_j)$

Источники

- https://www.youtube.com/watch?v=RdTxLXmbvjY
- https://logic.pdmi.ras.ru/ sergey/teaching/mlbayes/06-hmm.pdf
- http://statsoft.ru/home/textbook/modules/sttimser.htmlspectrum
- http://www.machinelearning.ru/wiki/images/b/b5/Koval2009ms.pdf
- https://mipt.ru/upload/1d5/f$_{g}ieq - arphcxl1tgs.pdf$
- http://www.isa.ru/jitcs/images/stories/2009/01/3$_{1}3.pdf$
- http://elib.bsu.by/bitstream/123456789/160135/1/Лобач.pdf
- http://www.machinelearning.ru/wiki/index.php?title=Временной
- http://www.machinelearning.ru/wiki/index.php?title=Прогнозировани
- https://university.prognoz.ru/biu/ru/Временные%2C
- http://www.machinelearning.ru/wiki/index.php?title=Модель –
- http://www.machinelearning.ru/wiki/index.php?title=Модель –