

Гауссовские процессы. Часть 2

Максим Кочуров

МГУ им. М.В. Ломоносова

Лекция 5



Содержание

① Введение

② Подход с гауссовскими процессами

- Введение
- Непериодическая часть
- Периодическая часть
- Модель
- Признаки Фурье



Временные ряды, классический подход

Если в данных есть сезонность, обычно используют **STL** декомпозицию (“**Seasonal and Trend decomposition using Loess**”). Однако:

- Параметры не интерпретируемы, есть только декомпозиция
- Нет оценок неопределённости
- Недостаточная гибкость

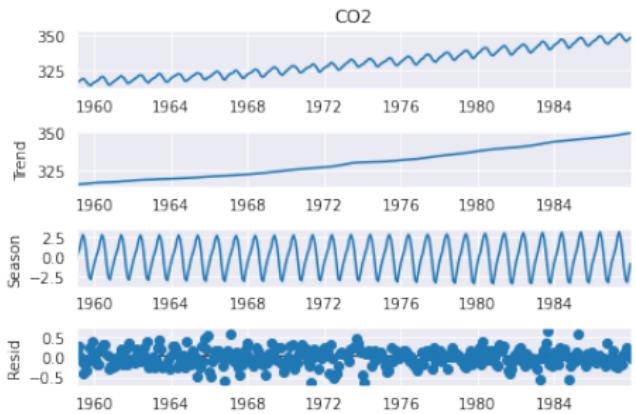


Figure: STL декомпозиция для данных об уровне CO₂, Statsmodels



Декомпозиция с гауссовскими процессами

Гауссовские процессы поддерживают дополнительные по сравнению с STL предположения

- Детализация сезонности (год + квартал + месяц + неделя)
- Модели с точками излома (changepoint).
- Гибкая функция правдоподобия
- Модели панельной регрессии
- Пропущенные значения



Типичная модель

Типичная модель аддитивная

$$x_t \sim \underbrace{g(t)}_{\text{тренд}} + \underbrace{s(t)}_{\text{периодичность}} + \underbrace{h(t)}_{\text{праздники}}$$

Источник

См. препринт о моделях Prophet [1]. Каждая модель временного ряда уникальна.



Напоминание

$x \in \mathbb{R}^n, y \in \mathbb{R}$

$$Y \sim \mathcal{GP}(m(x), k(x, x'))$$



Напоминание

$x \in \mathbb{R}^n$, $y \in \mathbb{R}$

$$Y \sim \mathcal{GP}(m(x), k(x, x'))$$

$$\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} m(x_1) \\ \vdots \\ m(x_N) \end{bmatrix}, \begin{bmatrix} k(x_1, x_1) & \dots & k(x_1, x_N) \\ \vdots & \ddots & \vdots \\ k(x_N, x_1) & \dots & k(x_N, x_N) \end{bmatrix} \right)$$

- ① \mathcal{GP} гауссовский процесс – многомерное нормальное распределение с матожиданием $m(t)$ и ковариацией $k(x, x')$



Напоминание

$x \in \mathbb{R}^n$, $y \in \mathbb{R}$

$$Y \sim \mathcal{GP}(m(x), k(x, x'))$$

$$\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} m(x_1) \\ \vdots \\ m(x_N) \end{bmatrix}, \begin{bmatrix} k(x_1, x_1) & \dots & k(x_1, x_N) \\ \vdots & \ddots & \vdots \\ k(x_N, x_1) & \dots & k(x_N, x_N) \end{bmatrix} \right)$$

- ① \mathcal{GP} гауссовский процесс – многомерное нормальное распределение с матожиданием $m(t)$ и ковариацией $k(x, x')$
- ② $m(x)$ – функция среднего
 - Линейная регрессия: $m(x) = x^\top \beta$
 - Константа (в т.ч. ноль): $m(x) = c$
 - Произвольная функция: $m(x) = \sin(x)$



Напоминание

$x \in \mathbb{R}^n$, $y \in \mathbb{R}$

$$Y \sim \mathcal{GP}(m(x), k(x, x'))$$

$$\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} m(x_1) \\ \vdots \\ m(x_N) \end{bmatrix}, \begin{bmatrix} k(x_1, x_1) & \dots & k(x_1, x_N) \\ \vdots & \ddots & \vdots \\ k(x_N, x_1) & \dots & k(x_N, x_N) \end{bmatrix} \right)$$

- ① \mathcal{GP} гауссовский процесс – многомерное нормальное распределение с матожиданием $m(t)$ и ковариацией $k(x, x')$
- ② $m(x)$ – функция среднего
 - Линейная регрессия: $m(x) = x^\top \beta$
 - Константа (в т.ч. ноль): $m(x) = c$
 - Произвольная функция: $m(x) = \sin(x)$
- ③ $k(x, x')$ – ядерная функция, мера сходства между x и x'
 - $[K]_{ij} = k(x_i, x_j)$ симметричная положительно определённая матрица



Непериодическая часть (среднее)

- Модели роста
- Модели линейного тренда
- Модели с точками излома
(changepoint)



Непериодическая часть (среднее)

- Модели роста
- Модели линейного тренда
- Модели с точками излома (changepoint)

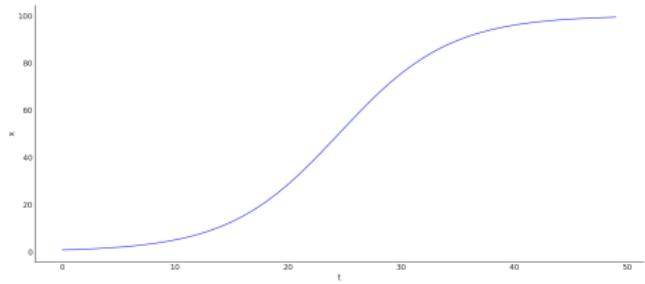


Figure: Модель роста

$$x = \frac{c}{1 + \exp(-k(t - m))}$$



Непериодическая часть (среднее)

- Модели роста
- **Модели линейного тренда**
- Модели с точками излома (changepoint)

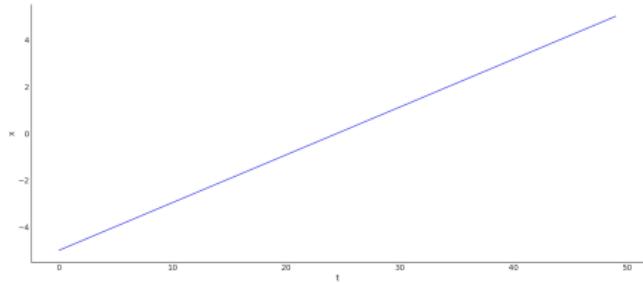


Figure: Модель линейного тренда

$$x = \frac{c}{1 + \exp(-k(t - m))}$$



Непериодическая часть (среднее)

- Модели роста
- Модели линейного тренда
- Модели с точками излома
(changepoint)

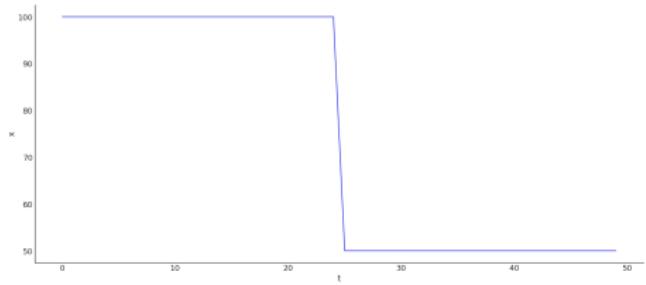


Figure: Модель с точкой излома

$$x = \begin{cases} c_1, & t < m \\ c_2, & t \geq m \end{cases}$$



Непериодическая часть (среднее)

- Модели роста
- Модели линейного тренда
- Модели с точками излома (changepoint)

Расширения

Возможны расширения, например, зависящее от времени насыщение в модели роста, см. [1].

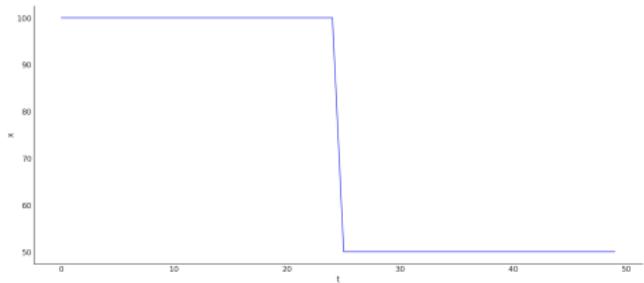


Figure: Модель с точкой излома

$$x = \begin{cases} c_1, & t < m \\ c_2, & t \geq m \end{cases}$$



Праздники

$$h(t) = \text{is-holiday}(t)$$

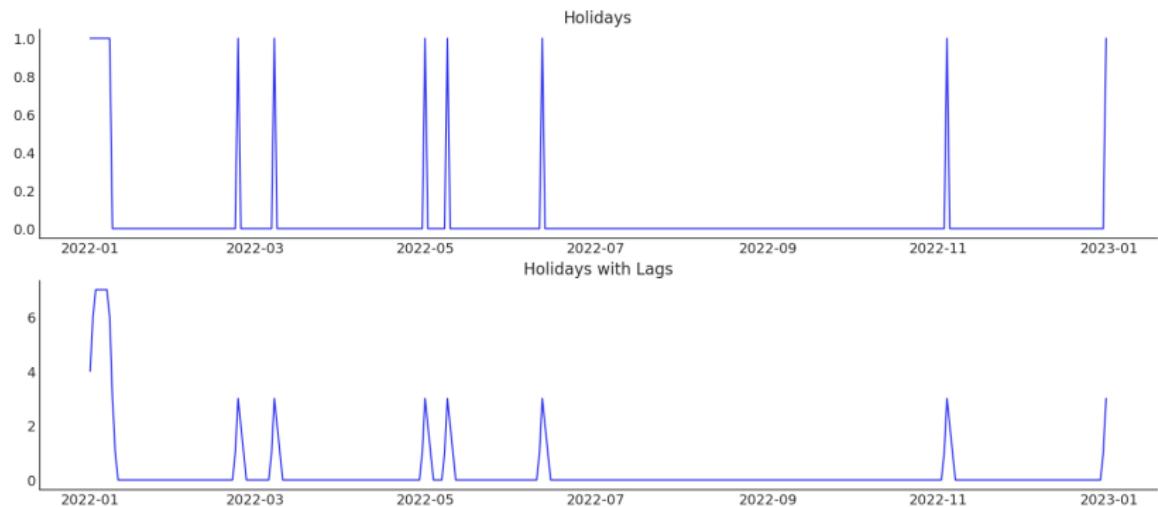


Figure: Признаки выходных дней



Периодическая часть (ковариация)

Здесь важна степень детализации.
Можно использовать несколько
периодических ядер.

- год
- квартал
- месяц
- неделя

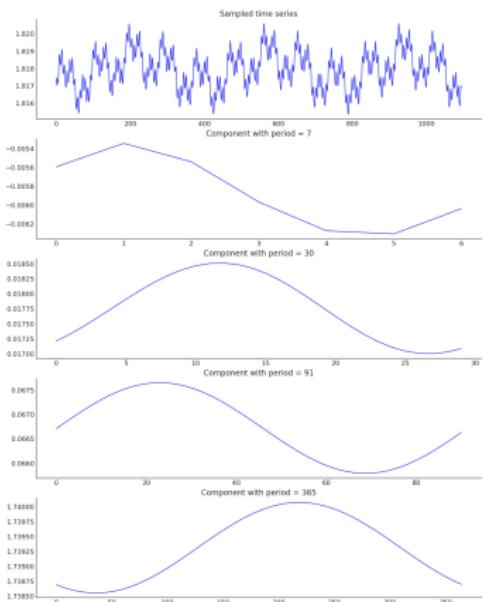


Figure: Сезонная декомпозиция



Масштабы длины для периодической части

Гиперпараметры

Выбор масштаба длины на основе здравого смысла

- Неделя – динамика может измениться за пару дней ($l_s \approx 3$)
- Месяц – динамика может измениться за неделю ($l_s \approx 7$)
- Квартал – динамика может измениться за месяц ($l_s \approx 30$)
- Год – динамика может измениться за квартал ($l_s \approx 90$)



Масштабы длины для периодической части

Гиперпараметры

Выбор масштаба длины на основе здравого смысла

- Неделя – динамика может измениться за пару дней ($l_s \approx 3$)
- Месяц – динамика может измениться за неделю ($l_s \approx 7$)
- Квартал – динамика может измениться за месяц ($l_s \approx 30$)
- Год – динамика может измениться за квартал ($l_s \approx 90$)

На практике

Вместо периодического ядра используют признаки Фурье (Fourier features).



Итоговая модель

$$m(t) = \underbrace{g(t)}_{\text{апериодичная часть}} + \underbrace{h(t)}_{\text{выходные}}$$

$$\begin{aligned} k(*, *) = & \quad \text{Periodic}(p = 365, l = 90) \\ & + \quad \text{Periodic}(p = 90, l = 30) \\ & + \quad \text{Periodic}(p = 30, l = 7) \\ & + \quad \text{Periodic}(p = 7, l = 3) \end{aligned}$$



Итоговая модель

$$m(t) = \underbrace{g(t)}_{\text{апериодичная часть}} + \underbrace{h(t)}_{\text{выходные}}$$

$$\begin{aligned} k(*, *) &= \alpha_{365} \text{Periodic}(p = 365, l = 90) \\ &\quad + \alpha_{90} \text{Periodic}(p = 90, l = 30) \\ &\quad + \alpha_{30} \text{Periodic}(p = 30, l = 7) \\ &\quad + \alpha_7 \text{Periodic}(p = 7, l = 3) \end{aligned}$$

Чего не хватает

- ❶ Веса при периодических компонентах



Итоговая модель

$$m(t) = \underbrace{g(t)}_{\text{апериодичная часть}} + \underbrace{h(t)}_{\text{выходные}}$$

$$\begin{aligned} k(*, *) &= \alpha_{365} \text{Periodic}(p = 365, l = 90) \\ &\quad + \alpha_{90} \text{Periodic}(p = 90, l = 30) \\ &\quad + \alpha_{30} \text{Periodic}(p = 30, l = 7) \\ &\quad + \alpha_7 \text{Periodic}(p = 7, l = 3) \\ &\quad + \beta \text{ExpQuad}(l = 90) + \text{WhiteNoise}(\gamma) \end{aligned}$$

Чего не хватает

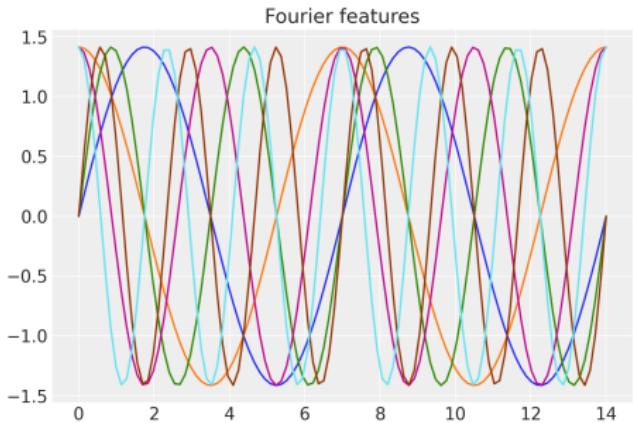
- ① Веса при периодических компонентах
- ② Отклонения от тренда $g(t)$



Более эффективная периодичность

Некоторые оптимизации недоступны, когда в модели есть хотя бы одно периодичное ядро.

- Можно добавить признаки Фурье как регрессоры
- Это позволяет добавить в модель разумную периодичность





Выбор признаков

У каждого периода есть
оптимальное число компонент
Фурье (опрядок)

- неделя: 3
- месяц: 10
- год: 5

```
from collections import namedtuple
from enum import Enum

Season = namedtuple("Season", "period,order")

class Daily(Season, Enum):
    Week = 7.0, 3
    Year = 365.25, 5
    Month = 365.25 / 12, 10

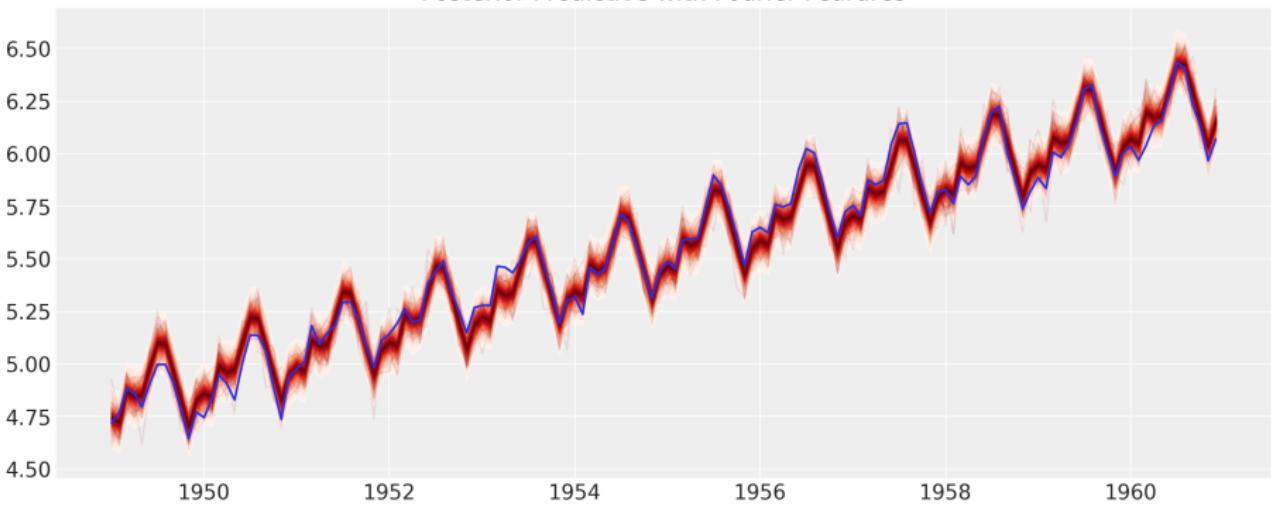
class Monthly(Season, Enum):
    Year = 12, 5
    Month = 1, 10
```



Применение признаков Фурье

- Модели типа Prophet
- Очень быстрая и доступная альтернатива гауссовским процессам
- Полезны для байесовских моделей

Posterior Predictive with Fourier Features





Интеграция в модель

```
alpha = pm.Normal("alpha", 0, sigma, shape=Monthly.Year.order * 2)
features = fourier_series(months, season=Monthly.Year)
seasonality = features @ alpha

expectation = bias + trend + seasonality
```



Библиография I



T. SJ and L. B.
Forecasting at scale.
2017.