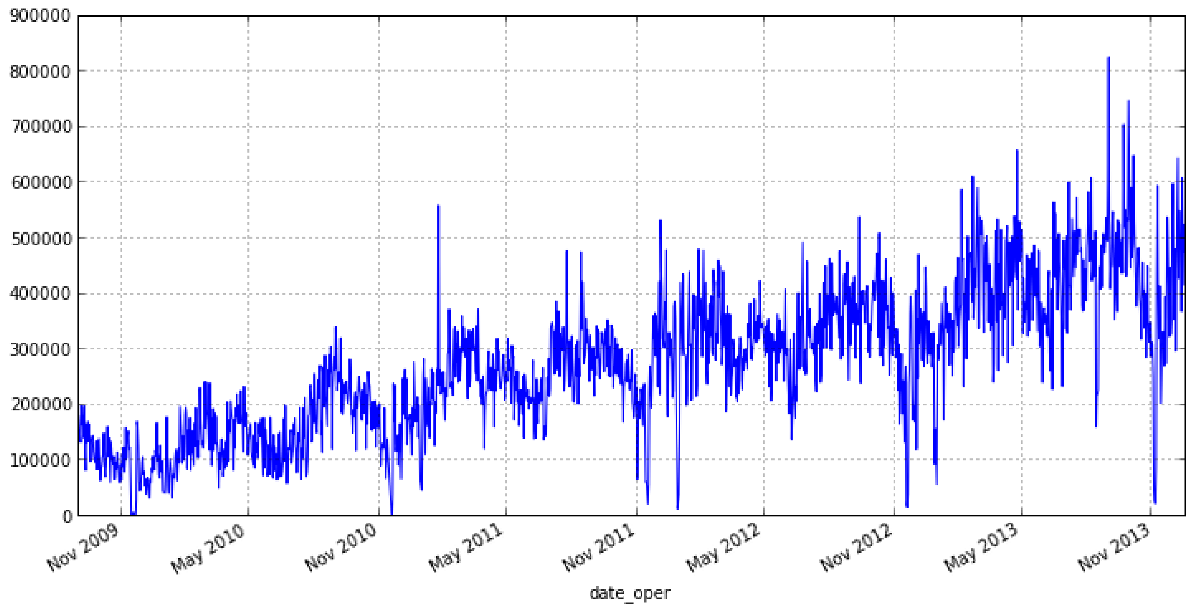
Занятие 38. Временные ряды

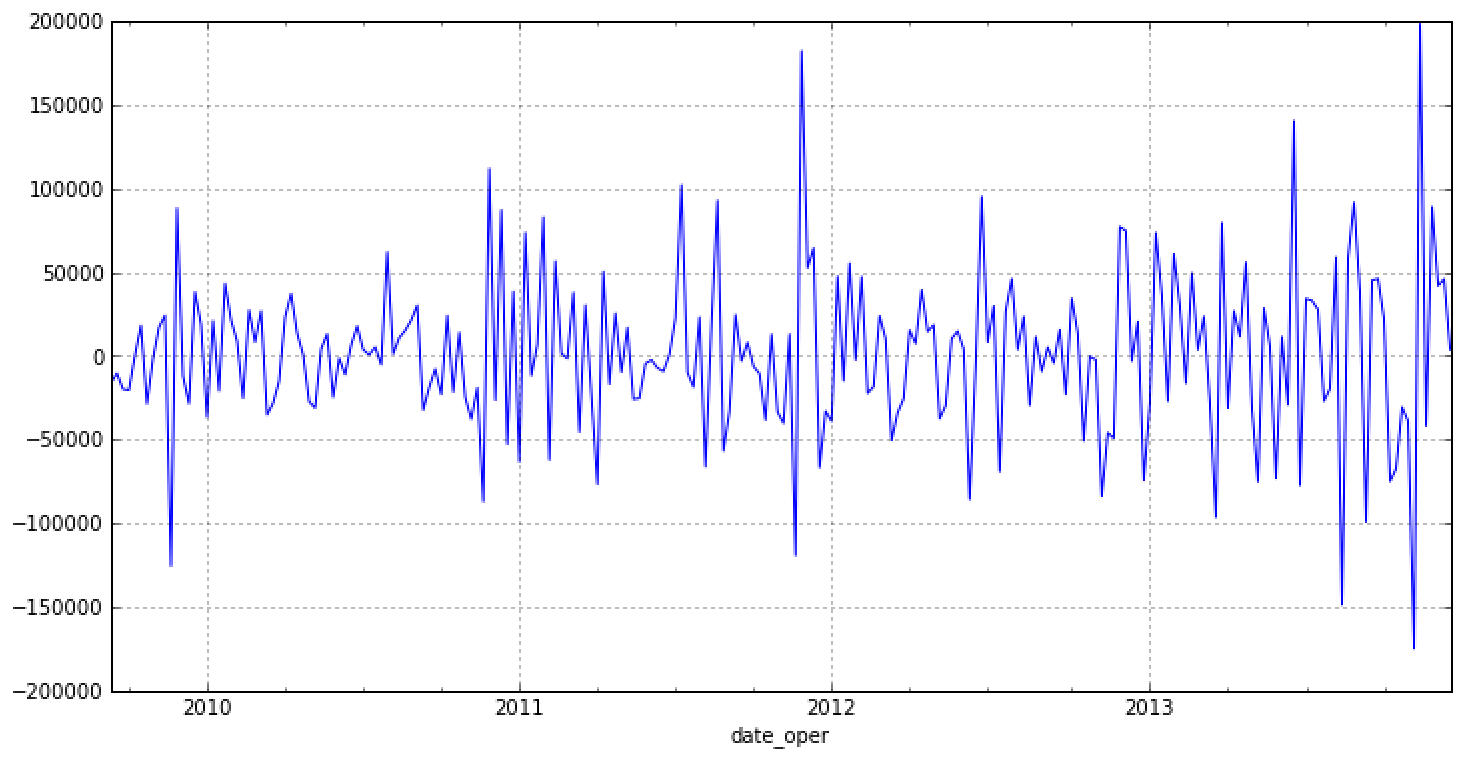
Talk about...

- Основные характеристики и компоненты временных рядов
- Стационарность и её значение
- Простое экспоненциальное сглаживание
- Метод Хольта
- Метод Хольта-Винтерса
- ARIMA



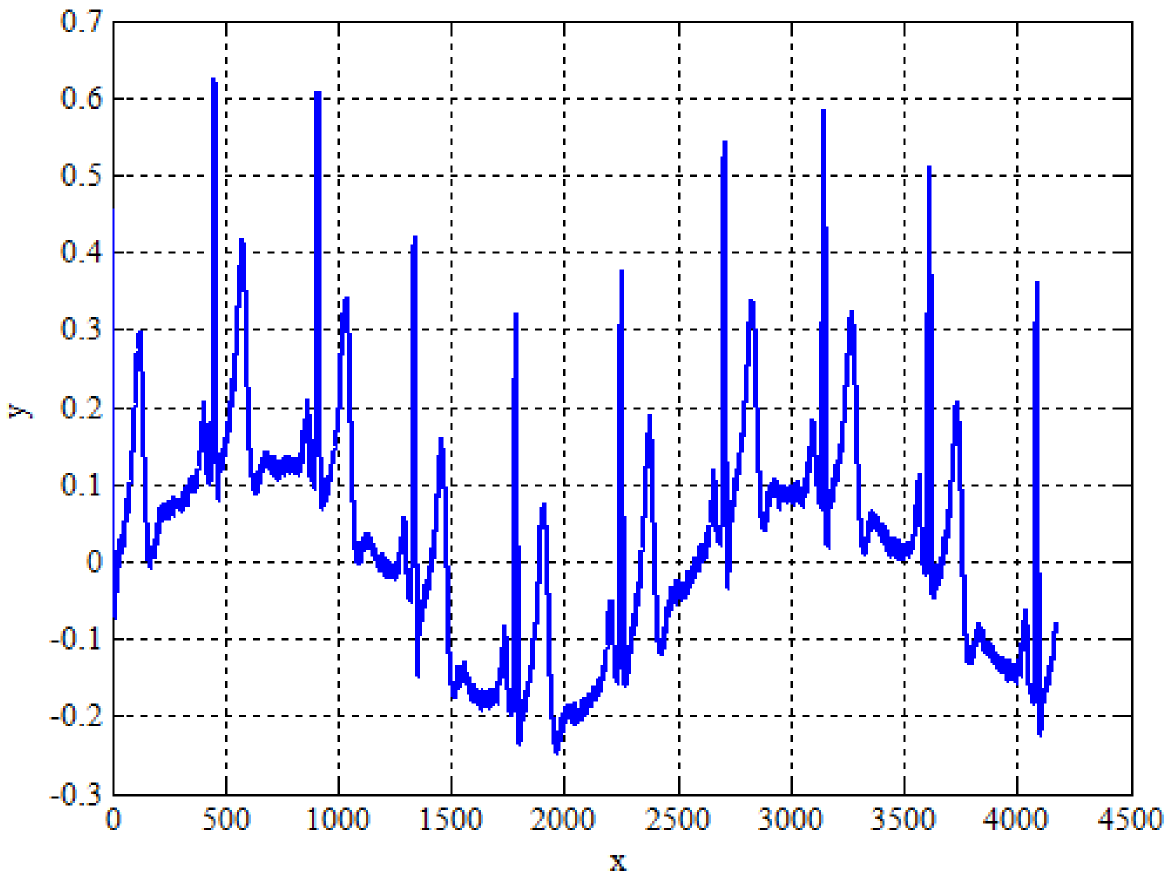
Teach Me Skills

1. Что такое временные ряды?



2. **Характеристики и компоненты временных рядов**

- Тренд
- Сезонность
- Цикличность
- Случайные колебания



3. Стационарность

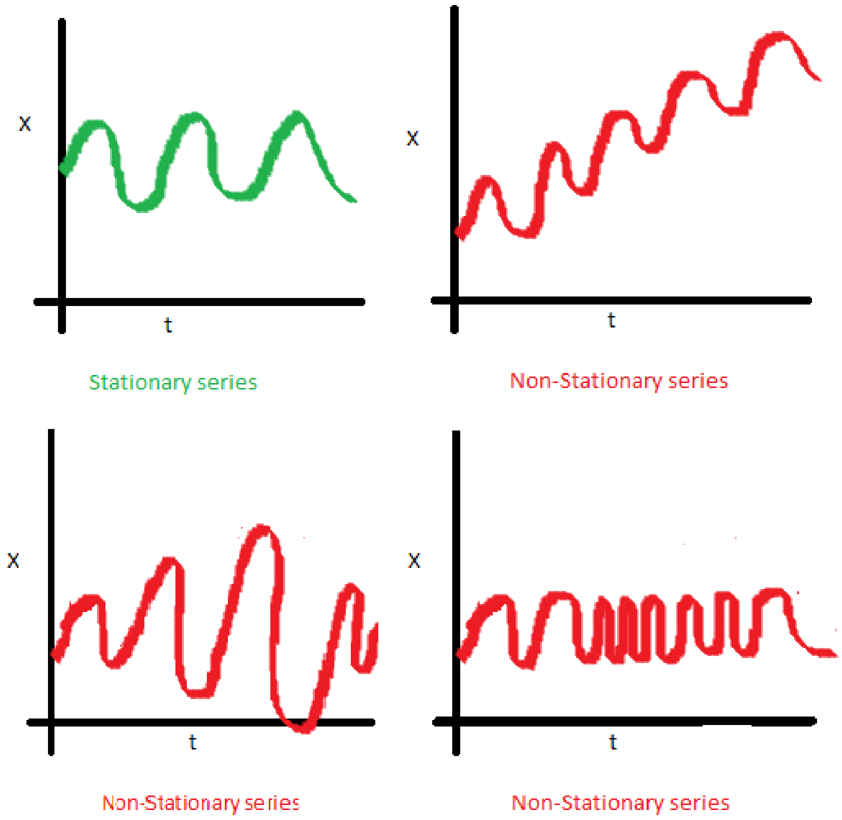
Teach Me Skills

Определение:

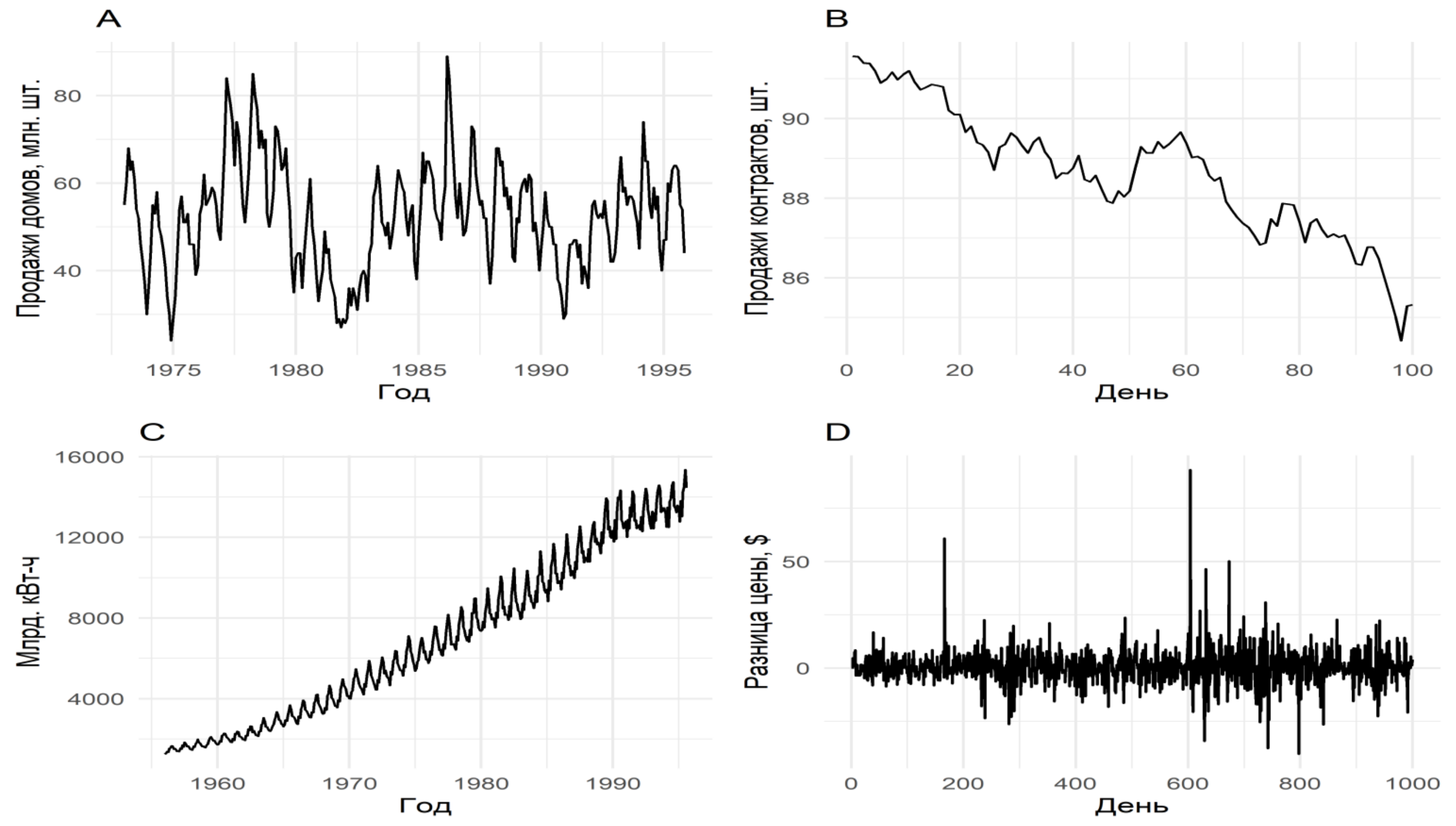
- Временной ряд считается стационарным, если его статистические свойства, такие как среднее, дисперсия и автокорреляция, не меняются со временем.

То есть, для стационарного ряда:

- Среднее значение остается постоянным.
- Дисперсия (разброс данных) остается постоянной.
- Корреляция между двумя точками данных зависит только от расстояния между этими точками, а не от абсолютного времени их наблюдения.



3. Стационарность



4. **Лаг во временных рядах**

Teach Me Skills

- Математическое определение лага

Пусть y_t — это значение временного ряда в момент времени t. Тогда значение ряда с лагом 1 в момент времени t будет:

$$y_{t-1}$$

Аналогично, значение ряда с лагом 2 в момент времени t будет:

$$y_{t-2}$$

и так далее.

Для лага k обозначение будет:

$$y_{t-k}$$

5. Тест Дики-Фуллера

Teach Me Skills

Рассмотрим следующую авторегрессионную модель первого порядка:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \cdots + \delta_p \Delta y_{t-p} + \varepsilon_t$$

где:

- Δ - оператор разностей первого порядка, $\Delta y_t = y_t - y_{t-1}$
- y_t - значение временного ряда в момент времени t
- α - константа
- βt - линейный тренд
- γ - коэффициент при значении ряда с предыдущим шагом времени
- $\delta_1, \ldots, \delta_p$ - коэффициенты лаговых значений разностей
- ε_t - ошибка в момент времени t

Тестирование гипотезы проводится для параметра γ:

$H_0 : \gamma = 0$ (наличие единичного корня, ряд нестационарен)

$H_1 : \gamma < 0$ (отсутствие единичного корня, ряд стационарен)

5. Тест Дики-Фуллера

Teach Me Skills

Статистика теста Дики-Фуллера рассчитывается следующим образом:

$$DF = \frac{\hat{\gamma}}{SE(\hat{\gamma})}$$

где:

- $\hat{\gamma}$ - оцененное значение параметра γ
- $SE(\hat{\gamma})$ - стандартная ошибка оценки γ

Если полученное значение статистики DF меньше критического значения (например, -2,89, -3,44 или -3,87 для уровней значимости 10%, 5% и 1% соответственно), то нулевая гипотеза отвергается, и ряд признается стационарным.

Точные критические значения зависят от конкретных условий модели (наличие константы, тренда и т. д.) и определяются на основе таблиц или с помощью специализированных программных продуктов.

5. Тест Дики-Фуллера

- Пример расчета

$y = [5, 6, 7, 8, 10, 11, 12, 14, 15, 17]$

1. **Рассчитываем разности первого порядка:**

$$\Delta y = [1, 1, 1, 2, 1, 1, 2, 1, 2]$$

1. **Подставляем значения в модель регрессии:**

В упрощенном примере рассмотрим модель без тренда и сезонности:

$$\Delta y_t = \alpha + \gamma y_{t-1} + \varepsilon_t$$

1. **Оценка параметров модели регрессии:**

С использованием метода наименьших квадратов (или другого метода регрессионного анализа) оцениваем параметры модели. Предположим, что результаты следующие:

$$\hat{\alpha} = 0.8, \hat{\gamma} = -0.5$$

1. **Расчет стандартной ошибки:**

Предположим, что стандартная ошибка для γ составляет $SE(\hat{\gamma}) = 0.2$.

1. **Вычисление статистики теста:**

$$DF = \frac{\hat{\gamma}}{SE(\hat{\gamma})} = \frac{-0.5}{0.2} = -2.5$$

1. **Сравнение с критическим значением:**

Если DF < -2.89 (критическое значение для 10% уровня значимости), то нулевая гипотеза о наличии единичного корня отвергается.

В нашем примере, -2.5 больше, чем -2.89, следовательно, нулевая гипотеза не может быть отвергнута на уровне значимости 10%.

Этот пример упрощен и показывает базовый процесс. В реальных условиях, для расчета такого теста требуются специализированные инструменты и более длинные временные ряды.

6. **Преобразование Бокса-Кокса**

Teach Me Skills

- Преобразование Бокса-Кокса используется для стабилизации дисперсии временных рядов. Если дисперсия временного ряда изменяется со временем, это может нарушить предположения многих статистических методов, и преобразование Бокса-Кокса может помочь сделать дисперсию более постоянной.

Дан временной ряд y_t и параметр преобразования λ. Преобразование Бокса-Кокса определяется следующим образом:

1. Если $\lambda \neq 0$:

$$y_t^{(\lambda)} = \frac{y_t^\lambda - 1}{\lambda}$$

1. Если $\lambda = 0$:

$$y_t^{(\lambda)} = \ln(y_t)$$

где:

- $y_t^{(\lambda)}$ — это преобразованное значение в момент времени t после применения преобразования Бокса-Кокса.
- λ — это параметр преобразования, который может быть выбран таким образом, чтобы минимизировать некоторую меру искажения в преобразованном ряду, например, используя метод максимального правдоподобия.

7. Простое экспоненциальное сглаживание

Teach Me Skills

- y_t — наблюдаемое значение временного ряда на момент времени t.
- $\hat{y}_{t+1}$ — прогнозное значение на момент времени $t+1$ на основе всей доступной информации до момента t (включительно).
- α — параметр сглаживания, который принимает значение между 0 и 1 ($0 \leq \alpha \leq 1$).

Формула простого экспоненциального сглаживания:

$$\hat{y}_{t+1} = \alpha y_t + (1 - \alpha)\hat{y}_t$$

где:

- αy_t — взвешенное значение наблюдения на момент t.
- $(1 - \alpha)\hat{y}_t$ — взвешенное прогнозное значение на момент t на основе информации до момента $t - 1$.

Выбор α влияет на степень сглаживания. Когда α близко к 1, ряд сглаживается слабо и больший вес придается недавним наблюдениям. Когда α близко к 0, ряд сглаживается сильнее, и больший вес придается более ранним значениям.

Простое экспоненциальное сглаживание особенно эффективно для рядов без тренда и сезонности.

8. Метод Хольта

Teach Me Skills

- y_t — наблюдаемое значение временного ряда на момент времени t.
- $\hat{y}_{t+m}$ — прогнозное значение на m шагов вперед от момента времени t.
- l_t — уровень (сглаженное значение) на момент времени t.
- b_t — тренд (наклон) на момент времени t.
- α — параметр сглаживания для уровня (0 ≤ α ≤ 1).
- β — параметр сглаживания для тренда (0 ≤ β ≤ 1).

Формулы метода Хольта:

1. **Уровень:**

   $$l_t = \alpha y_t + (1-\alpha)(l_{t-1} + b_{t-1})$$

2. **Тренд:**

   $$b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1}$$

3. **Прогноз:**

   $$\hat{y}_{t+m} = l_t + mb_t$$

$$\hat{y}_{t+m} = l_t + mb_t$$

где:

- αy_t — взвешенное наблюдаемое значение на момент t.
- $(1-\alpha)(l_{t-1} + b_{t-1})$ — взвешенное прогнозное значение на момент t, основанное на предыдущем уровне и тренде.
- $\beta(l_t - l_{t-1})$ — взвешенное изменение уровня между текущим и предыдущим моментами времени.
- $(1-\beta)b_{t-1}$ — взвешенный тренд на предыдущем моменте времени.

Выбор параметров α и β обычно основывается на оптимизации, например, с использованием метода наименьших квадратов или метода максимального правдоподобия, чтобы минимизировать ошибки прогноза на исторических данных.

8. Метод Хольта

Teach Me Skills

- Пример. Временной ряд представляет собой количество проданных товаров в магазине за 5 месяцев:

Месяц	Продажи
1	100
2	105
3	108
4	107
5	110

Предположим начальные значения:

- $l_1 = 100$ (первое наблюдение)
- $b_1 = 5$ (первоначальное предположение о тренде)

Параметры сглаживания:

- $\alpha = 0.5$
- $\beta = 0.4$

Расчёт:

1. **Месяц 2:**
   - $l_2 = \alpha \times y_2 + (1 - \alpha) \times (l_1 + b_1) = 0.5 \times 105 + 0.5 \times (100 + 5) = 105$
   - $b_2 = \beta \times (l_2 - l_1) + (1 - \beta) \times b_1 = 0.4 \times (105 - 100) + 0.6 \times 5 = 5$
2. **Месяц 3:**
   - $l_3 = \alpha \times y_3 + (1 - \alpha) \times (l_2 + b_2) = 0.5 \times 108 + 0.5 \times (105 + 5) = 109$
   - $b_3 = \beta \times (l_3 - l_2) + (1 - \beta) \times b_2 = 0.4 \times (109 - 105) + 0.6 \times 5 = 5.6$

Таким образом, можно продолжить расчёты для следующих месяцев.

9. Модель Хольта-Винтерса

Teach Me Skills

- y_t — наблюдаемое значение временного ряда на момент времени t.
- l_t — уровень (сглаженное значение) на момент времени t.
- b_t — тренд (наклон) на момент времени t.
- s_t — сезонный компонент на момент времени t.
- α — параметр сглаживания для уровня ($0 \leq \alpha \leq 1$).
- β — параметр сглаживания для тренда ($0 \leq \beta \leq 1$).
- γ — параметр сглаживания для сезонности ($0 \leq \gamma \leq 1$).
- m — длина сезонного периода (например, 12 для месячных данных с годовой сезонностью).

Формулы модели Хольта-Винтерса:

1. **Уровень**:
   $l_t = \alpha \frac{y_t}{s_{t-m}} + (1-\alpha)(l_{t-1} + b_{t-1})$
2. **Тренд**:
   $b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1}$
3. **Сезонность**:
   $s_t = \gamma \frac{y_t}{l_t + b_t} + (1-\gamma)s_{t-m}$
4. **Прогноз** на h периодов вперед:
   $\hat{y}_{t+h} = (l_t + hb_t)s_{t-m+h_{mod}m}$
   где $h_{mod}m$ - остаток от деления h на m.

9. **Модель Хольта-Винтерса**

Teach Me Skills

Пример: Временной ряд представляет собой количество проданных товаров в магазине за 8 месяцев:

Месяц	Продажи
1	100
2	120
3	110
4	130
5	105
6	125
7	115
8	135

Предположим, что сезонность имеет период в 4 месяца.

Предположим начальные значения:

- $l_1 = 100$
- $b_1 = 5$
- $s_1 = s_2 = s_3 = s_4 = 1$ (первоначальные множители сезонности)

Параметры сглаживания:

- $\alpha = 0.5$
- $\beta = 0.4$
- $\gamma = 0.6$

Расчет:

1. **Месяц 5:**
   - $l_5 = \alpha \frac{y_5}{s_1} + (1-\alpha)(l_4 + b_4) = 0.5 \times \frac{105}{1} + 0.5 \times (130 + 5) = 120$
   - $b_5 = \beta(l_5 - l_4) + (1-\beta)b_4 = 0.4 \times (120 - 130) + 0.6 \times 5 = 4$
   - $s_5 = \gamma \frac{y_5}{l_5 + b_5} + (1-\gamma)s_1 = 0.6 \times \frac{105}{124} + 0.4 = 1.01$
2. **Месяц 6:**
   - $l_6 = \alpha \frac{y_6}{s_2} + (1-\alpha)(l_5 + b_5)$
   - $b_6 = \beta(l_6 - l_5) + (1-\beta)b_5$
   - $s_6 = \gamma \frac{y_6}{l_6 + b_6} + (1-\gamma)s_2$

4. **ARIMA и её компоненты**

Teach Me Skills

ARIMA (AutoRegressive Integrated Moving Average) — это популярный класс моделей для прогнозирования временных рядов. ARIMA объединяет три основные компоненты: авторегрессию (AR), интегрирование (I) и скользящее среднее (MA).

Авторегрессионная модель (AR)

Авторегрессионная модель определяет значение временного ряда в данный момент времени на основе его предыдущих значений.

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \epsilon_t$$

где:

- y_t — значение временного ряда в момент времени t
- c — константа
- $\phi_1, \phi_2, \ldots, \phi_p$ — параметры авторегрессии
- ϵ_t — белый шум
- p — порядок авторегрессии

Интегрированная часть (I)

Интегрирование в ARIMA означает разности между последовательными значениями временного ряда. Это делается для того, чтобы сделать ряд стационарным. Обычно, если первые разности ряда стационарны, говорят, что ряд интегрирован первого порядка.

Для однократного интегрирования:

$$y'_t = y_t - y_{t-1}$$

где:

- y'_t — первая разность временного ряда

Параметр d в модели ARIMA указывает, сколько раз нужно интегрировать ряд, чтобы он стал стационарным.

4. **ARIMA и её компоненты**

Teach Me Skills

Подбор параметров p, d, q
Для выбора оптимальных параметров p,d,q для модели ARIMA часто используют методы, такие как:
ACF (автокорреляционная функция) и PACF (частичная автокорреляционная функция) для определения порядка p и q.
Тесты на стационарность, такие как тест Дики-Фуллера, для определения порядка d.
Критерии информативности, такие как AIC (акаике) и BIC (байесовский информационный критерий), для сравнения различных моделей.

Модель скользящего среднего (MA)

Модель скользящего среднего представляет значение временного ряда как комбинацию белого шума и его предыдущих значений.

$$y_t = c + \epsilon_t + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q}$$

где:

- $\theta_1, \ldots, \theta_q$ — параметры скользящего среднего
- q — порядок скользящего среднего

4. **Критерии информативности: AIC и BIC**

Teach Me Skills

Критерии информативности являются методами оценки и сравнения статистических моделей, основываясь на их качестве и сложности. Они помогают определить, какая модель лучше описывает данные, учитывая не только качество приближения, но и количество параметров, используемых в модели.

1. AIC (Критерий Акаике)

Критерий Акаике или AIC (Akaike Information Criterion) оценивает качество модели относительно других моделей. Меньшее значение AIC предпочтительнее.

Математическая формулировка:

$$AIC = 2k - 2\ln(L)$$

где:

- k — количество параметров в модели
- L — правдоподобие модели (likelihood)

2. BIC (Байесовский информационный критерий)

BIC (Bayesian Information Criterion) также измеряет качество модели, но он включает в себя штраф за увеличение числа параметров, что делает его более строгим по сравнению с AIC, особенно при больших объемах данных.

Математическая формулировка:

$$BIC = \ln(n) \cdot k - 2\ln(L)$$

где:

- n — количество наблюдений (размер выборки)
- k — количество параметров в модели
- L — правдоподобие модели

Практическое занятие:

Прогнозирование временных рядов

Домашнее задание:
Прогнозирование временного ряда