Предсказание временных рядов

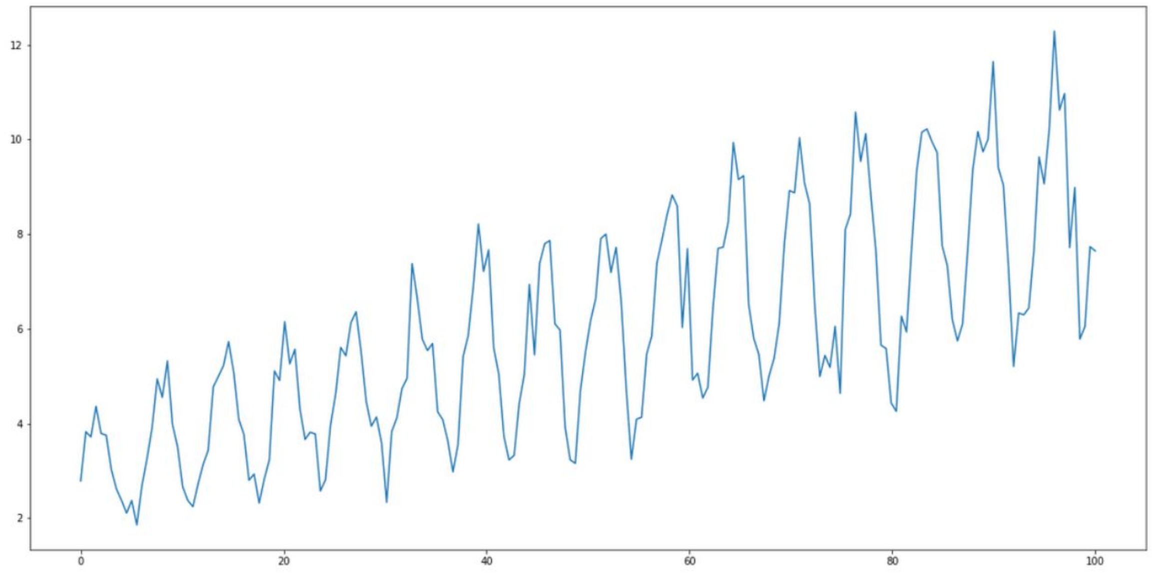
Лекция 2

Глава 1

Moving Average & ARMA

Начнем с примера

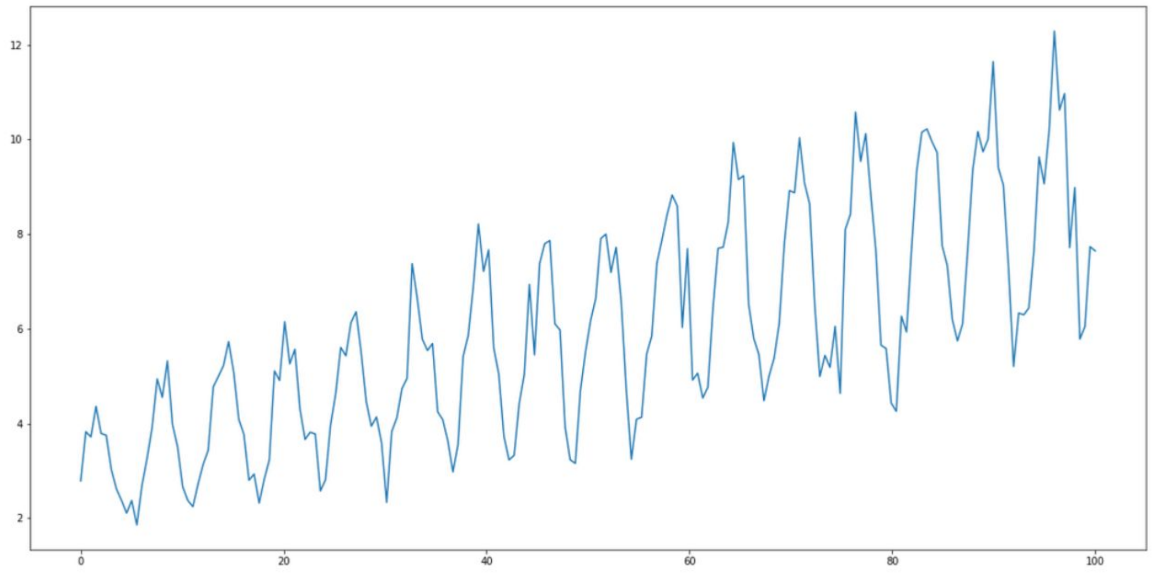
Что мы можем сказать про этот временной ряд?



Начнем с примера

Что мы можем сказать про этот временной ряд?

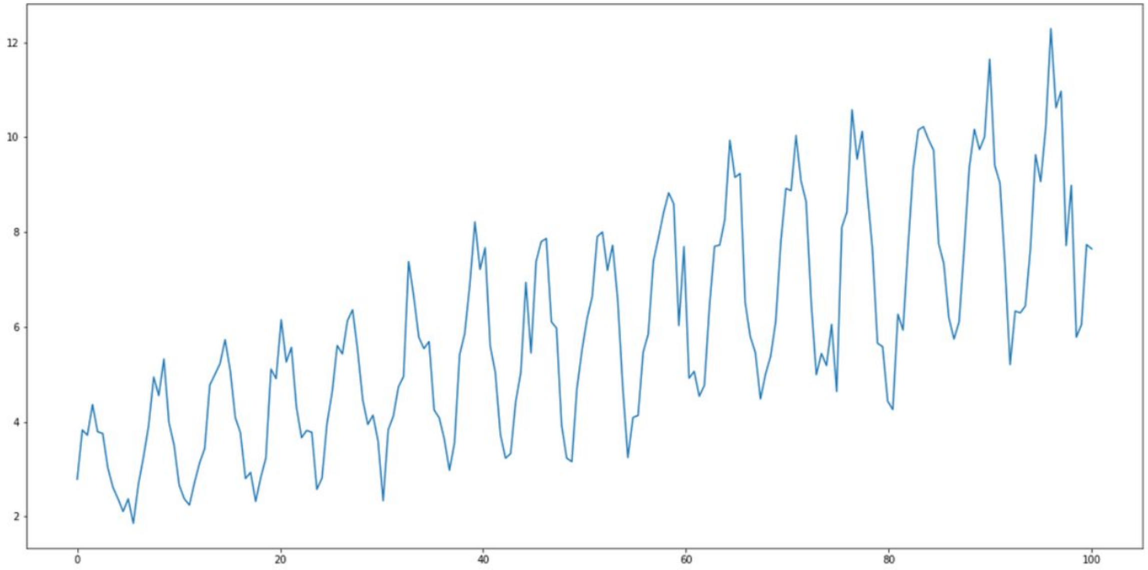
- Сезонность
- Тренд



Начнем с примера

Что мы можем сказать про этот временной ряд?

- Сезонность
- Тренд
- Не стационарный



Напоминание

Стационарность - свойство, характеризующее отсутствие тренда (постоянное матожидание), и постоянную дисперсию (ковариация зависит только от лага ***l***)

Определение

Гомоскедастичность - свойство, определяющее наличие постоянной дисперсии в каждом из сезонов.

Три вопроса

1. Можно ли ряд привести к стационарному виду?

Три вопроса

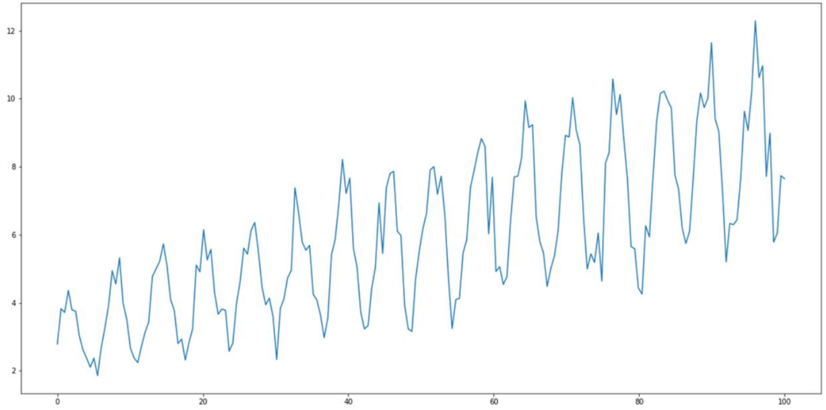
1. Можно ли ряд привести к стационарному виду?
2. А как?

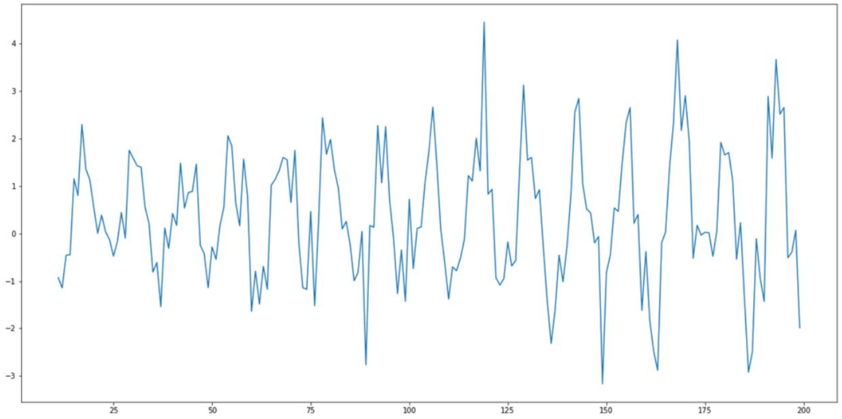
Три вопроса

1. Можно ли ряд привести к стационарному виду?
2. А как?
3. А обратно?

Пример. Продолжение

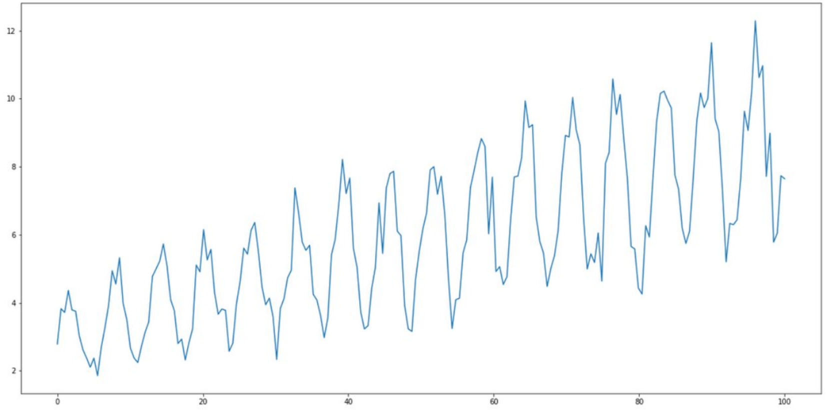
Приведем ряд к стационарному виду:

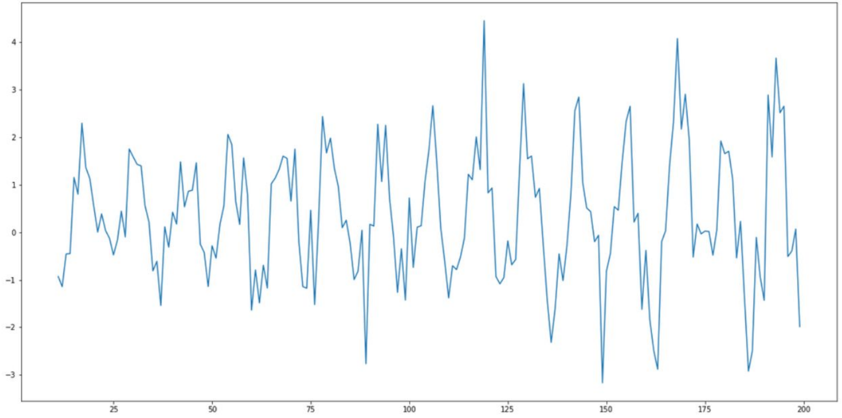




Пример. Продолжение

Приведем ряд к стационарному виду:





Ничего не напоминает?

Вопрос

Для того, чтобы привести ряд к стационарному виду, всегда ли хватит один раз сделать такую процедуру?

Теорема Волда

Каждый ***слабо стационарный*** временной ряд можно представить в виде ***скользящего среднего*** бесконечного порядка $MA(\infty)$

Такое представление будет называться ***представлением скользящим средним*** для временных рядов.

$$Y_t = \sum_{i=1}^{\infty} b_i \epsilon_{t-i} + \nu_t$$

- Y_t - рассматриваемый временной ряд
- ϵ_{t-i} - белый шум
- b_i - коэффициенты скользящего среднего
- ν_t - детерминированная компонента (характеризует тренд)

А коэффициенты b_i такие, что:

- нулевой коэффициент равен 1
- ряд сходится абсолютно
- отсутствуют члены с $j < 0$
- не зависят от t

Пример

На самом деле все несложно - посмотрим на примере лагового оператора ***L(i)*** - он возвращает значение, которое было ***i*** позиций назад.

В таком случается, утверждается, что теорема Волда может быть записана так:

$$\gamma(L) = \sum_{k=0}^{\infty} \gamma_k L_k$$

p.s. не пугайтесь, тут просто эпсилон на гамму заменили :)

Замечание 1

Понятное дело, что построить $MA(\infty)$ мы не можем - поэтому построим ***MA(d)***

Замечание 2

Понятное дело, что построить $MA(\infty)$ мы не можем - поэтому построим ***MA(d)***

Более того, мы даже тренд найти можем (как например?)

Замечание 3

Понятное дело, что построить $MA(\infty)$ мы не можем - поэтому построим ***MA(d)***

Более того, мы даже тренд найти можем (как например?)

Но как найти ϵ_{t-i} ?..

Замечание 3

Понятное дело, что построить $MA(\infty)$ мы не можем - поэтому построим ***MA(d)***

Более того, мы даже тренд найти можем (как например?)

Но как найти ϵ_{t-i} ?..

Надо что-то ***обучить***! Но что?

AR(p)

Для этого обратимся к еще одной модельке - ***авторегрессионной модели***

$$Y_t = \sum_{j=1}^{p} Y_{t-j}\alpha_j + \epsilon_t$$

Эта модель полагается только на предыдущие значения, а значит ее мы можем обучить. Как пример, получив на вход **2,3,4** - она скажет **5**.

AR(p)

Для этого обратимся к еще одной модельке - ***авторегрессионной модели***

$$Y_t = \sum_{j=1}^{p} Y_{t-j}\alpha_j + \epsilon_t$$

она нам и нужна!

Эта модель полагается только на предыдущие значения, а значит ее мы можем обучить. Как пример, получив на вход **2,3,4** - она скажет **5**.

И?

Что делать дальше?

ARMA(p,d)

Все просто! Объединим их:

$$Y_t = \sum_{i=1}^{p} Y_{t-i}\alpha_i + \sum_{j=1}^{d} b_j \epsilon_{t-j} + \mu_t + \epsilon_t$$

ARMA(p,d)

Все просто! Объединим их:

Это от AR

$$Y_t = \sum_{i=1}^{p} Y_{t-i}\alpha_i + \sum_{j=1}^{d} b_j \epsilon_{t-j} + \mu_t + \epsilon_t$$

Это от MA

ARMA(p,d)

Все просто! Объединим их:

Помним что это такое?

$$Y_t = \sum_{i=1}^{p} Y_{t-i}\alpha_i + \sum_{j=1}^{d} b_j \epsilon_{t-j} + \mu_t + \epsilon_t$$

ARMA(p,d)

Оптимизация: все тот же **MSE**

ARMA(p,d)

Оптимизация: все тот же **MSE**

Окей, допустим, параметры она найдет - что делать с **p** и **d**?

ARMA(p,d)

Оптимизация: все тот же **MSE**

Окей, допустим, параметры она найдет - что делать с **p** и **d**?

- Самый простой способ - угадать

ARMA(p,d)

Оптимизация: все тот же **MSE**

Окей, допустим, параметры она найдет - что делать с **p** и **d**?

- Самый простой способ - угадать
- Есть еще один, но он вам не понравится…

ARMA(p,d)

Подбор **p**:

- Идем от 0 элемента до конца
- Для каждого элемента считаем корреляцию с набором всех предыдущих - это называется **PACF**
- Параллельно строим график значимости - выбираем предел значимости (обычно 0.05 или 0.01) и с помощью (например) таблицы критических значения распределения Стьюдента, мы находим критический порог для текущего уровня значимости.

ARMA(p,d)

Подбор **q**:

- Делаем аналогично, но теперь считаем корреляцию для **лагов**. Это называется **ACF**.

ARMA(p,d)

Круто, мы теперь примерно понимаем как работает ARMA.

И по теореме Волда даже понимаем, что она реально работает.

Но она работает не со всеми рядами…

Как фиксить?

ARMA(p,d)

Круто, мы теперь примерно понимаем как работает ARMA.

И по теореме Волда даже понимаем, что она реально работает.

Но она работает не со всеми рядами…

Как фиксить?

Мы ведь знаем как приводить ряд к стационарному виду!

Глава 2

ARIMA

ARIMA(p,q,d)

Давайте просто сделаем ряд стационарным и добавим что-то про тренд!

$$\Delta^q Y_t = c + \sum_{i=1}^{p} \Delta^q Y_{t-i}\alpha_i + \sum_{j=0}^{d} b_j \epsilon_{t-j}$$

Δ^q - оператор взятия разности **q** раз - сначала в самом ряду, потом в полученном и т.д.

Как подобрать q?

Как это можно сделать?

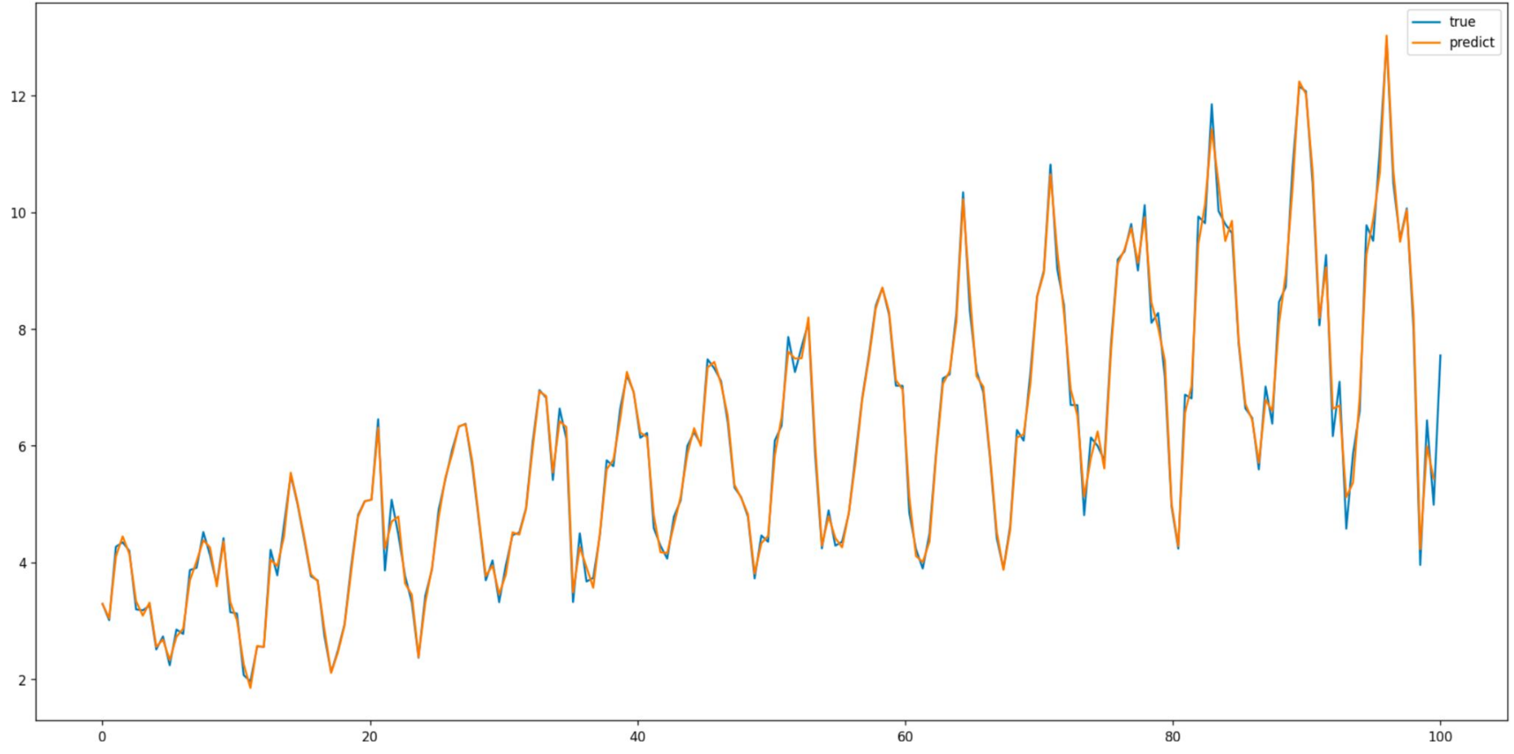
Как подобрать q?

Как это можно сделать?

На самом деле несложно:

- Берем 1 раз разность.
- Проверяем на стационарность (например тестом Дики-Фуллера)
- Если все ок - то заканчиваем. Иначе берем разность еще раз.

Как итог, ARIMA(p,q,d)



Вопросы?


Я вам
ЗАПРЕЩАЮ
использовать ARMA на не стационарном ряду
вы нормальные вообще?
3847Ь