Прогнозирование временных рядов

1. Временные ряды.

Временным рядом называется последовательность значений признака y, измеряемого через постоянные временные интервалы:

$$y_1, \dots, y_T, \dots, \quad y_t \in \mathbb{R}.$$

Таким образом, данные оказываются упорядочены относительно неслучайных моментов времени, и, значит, **в отличие от случайных выборок, могут содержать в себе дополнительную информацию**, которую можно извлечь.

Примеры временных рядов — это *ряды среднедневных цен на акции определённой компании, среднемесячного уровня безработицы*, измеренного в течение нескольких лет, *среднегодового уровня производства товара*.

Прогнозирование временного ряда.

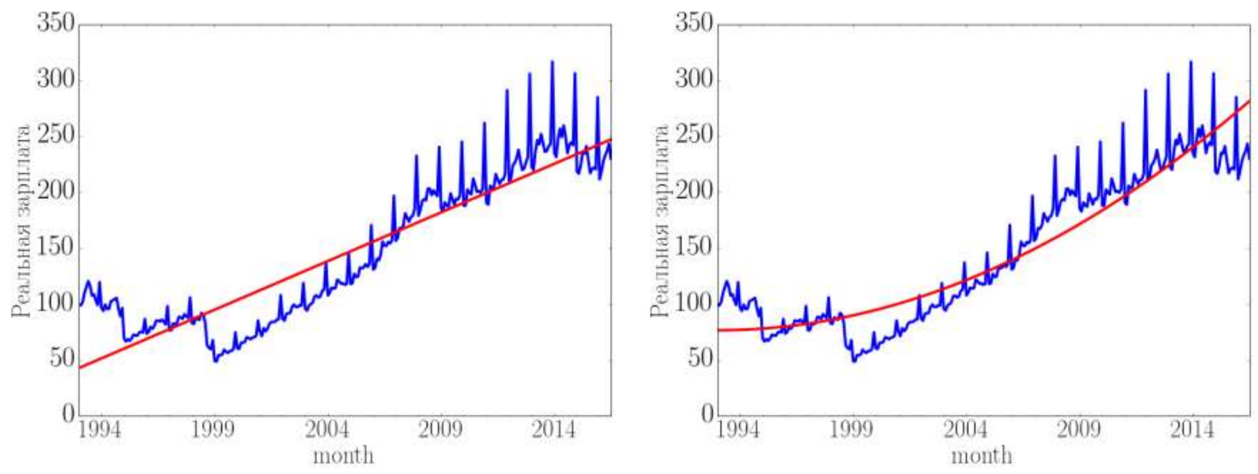
Интерес представляет задача прогнозирования временных рядов. Подразумевается, что, зная значение признака в прошлом, можно предсказать его в будущем. Формально задача ставится как поиск функции f_T :

$$y_{T+d} \approx f_T(y_T, \dots, y_1, d) \equiv \hat{y}_{T+d|T},$$

где d ∈ {1, . . . , D} — отсрочка прогноза, D — горизонт прогнозирования.

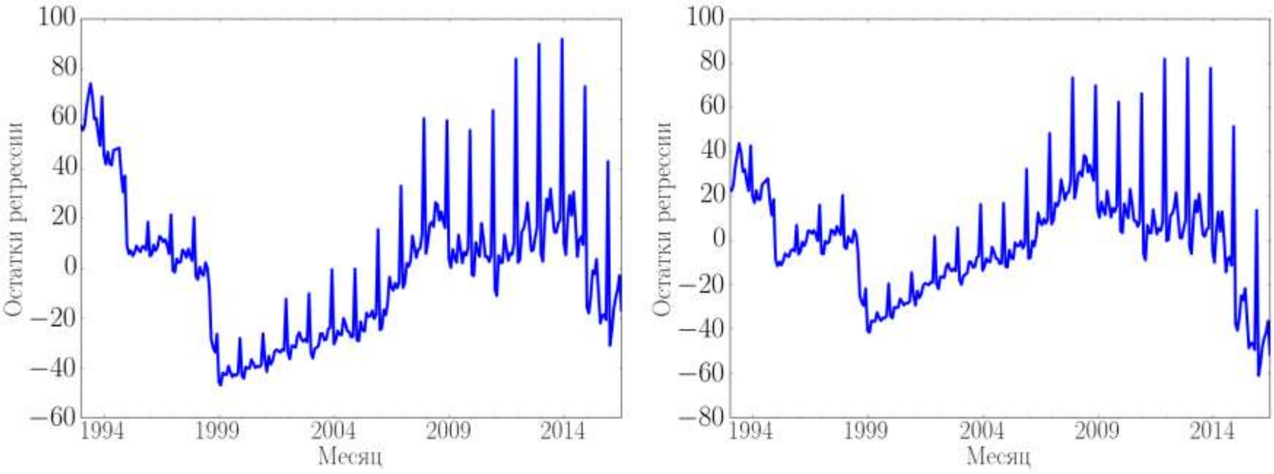
В задаче анализа временных рядов предполагается, что данные в прошлом каким-то образом связаны с данными в будущем. Чем сильнее они связаны, тем больше имеется информации о поведении временного ряда в будущем и тем точнее можно сделать прогноз.

Можно попробовать свести задачу прогнозирования временного ряда к задаче обучения с учителем. Процесс разворачивается во времени, поэтому кажется логичным задать признаки, связанные со временем и попробовать решить задачу, применяя модель регрессии. Регрессия может быть линейной или, например, квадратичной:



Применение модели линейной (слева) и квадратичной (справа) регрессии к задаче прогнозирования временного ряда.

Однако это решение слишком простое, чтобы быть хорошим. Отклонения (остатки) такой регрессии далеко не похожи на случайный шум, в них остаётся большая часть структуры, которая не была учтена в регрессионной модели. Чем больше структуры временного ряда учитывается в модели, тем лучшее предсказание она даёт.



Остатки модели линейной (слева) и квадратичной (справа) регрессии в задаче прогнозирования временного ряда.

Компоненты временных рядов

Рассмотрим несколько понятий, которыми можно описать поведение временных рядов:

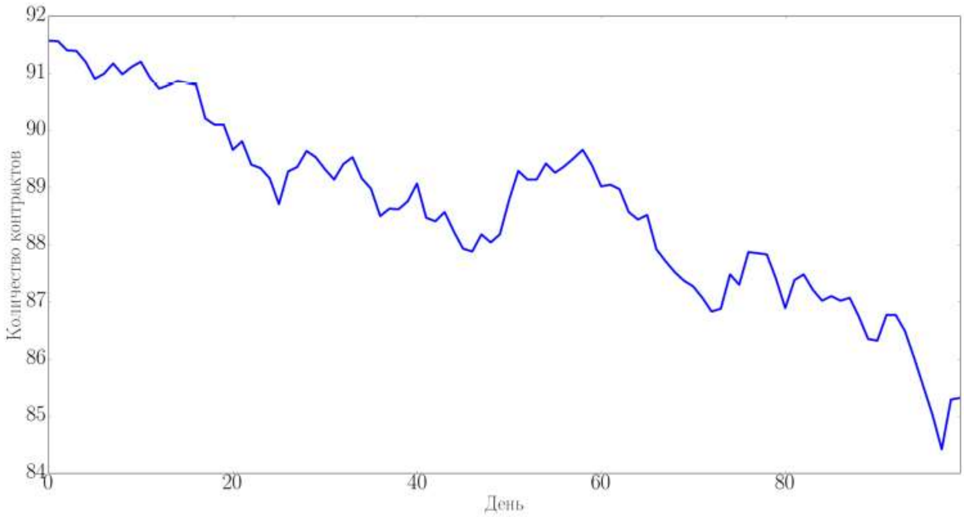
✓ **Тренд** — плавное долгосрочное изменение уровня ряда. Эту характеристику можно получить, наблюдая ряд в течение достаточно долгого времени.

✓ **Сезонность** — циклические изменения уровня ряда *с постоянным периодом*. Например, потребление электроэнергии является высоким в течение дня и низким ночью, или онлайн-продажи увеличиваются во время Рождества, прежде чем снова замедлиться.

✓ **Цикл** — изменение уровня ряда с *переменным периодом*. Такое поведение часто встречается в рядах, связанных с продажами, и объясняется циклическими изменениями экономической активности. В экономике выделяют циклы длиной 4 − 5 лет, 7 − 11 лет, 45 − 50 лет и т. д. Другой пример ряда с такой характеристикой — это солнечная активность, которая соответствует, например, количеству солнечных пятен за день. Она плавно меняется с периодом, который составляет несколько лет, причём сам период также меняется во времени.

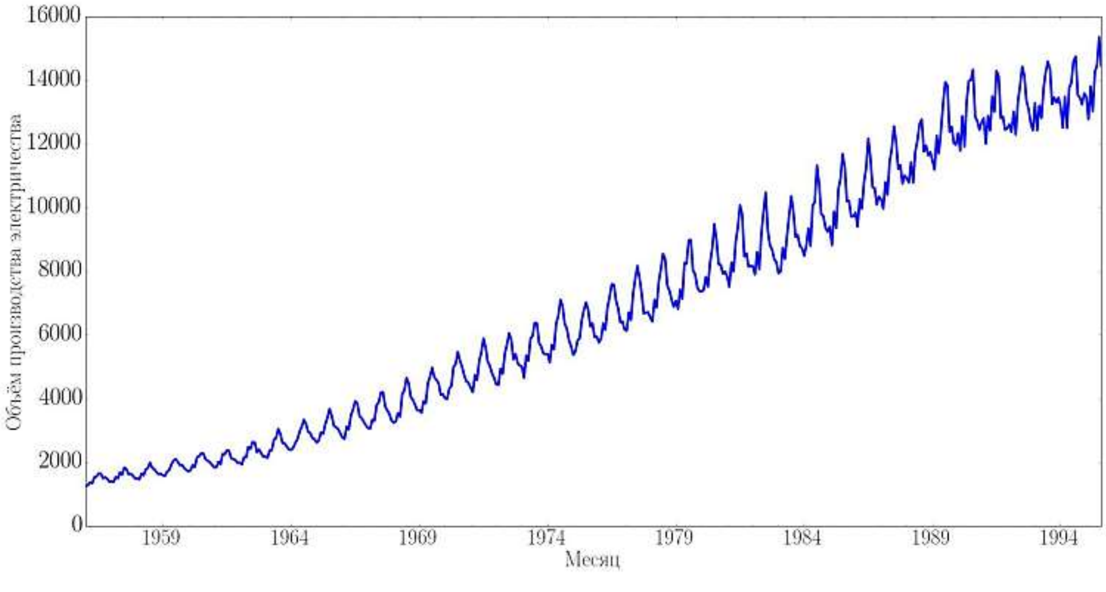
✓ **Ошибка** — непрогнозируемая случайная компонента ряда. Сюда включены все те характеристики временного ряда, которые сложно измерить (например, слишком слабые)

В качестве примера временного ряда можно рассмотреть количество контрактов за день в сокровищнице США. На графике виден **хорошо выраженный понижающийся тренд**, который можно описать линейной функцией. На этом участке в данных **не наблюдается ни циклов, ни сезонности**. По-видимому, всё, что не удаётся описать трендом, является ошибкой.



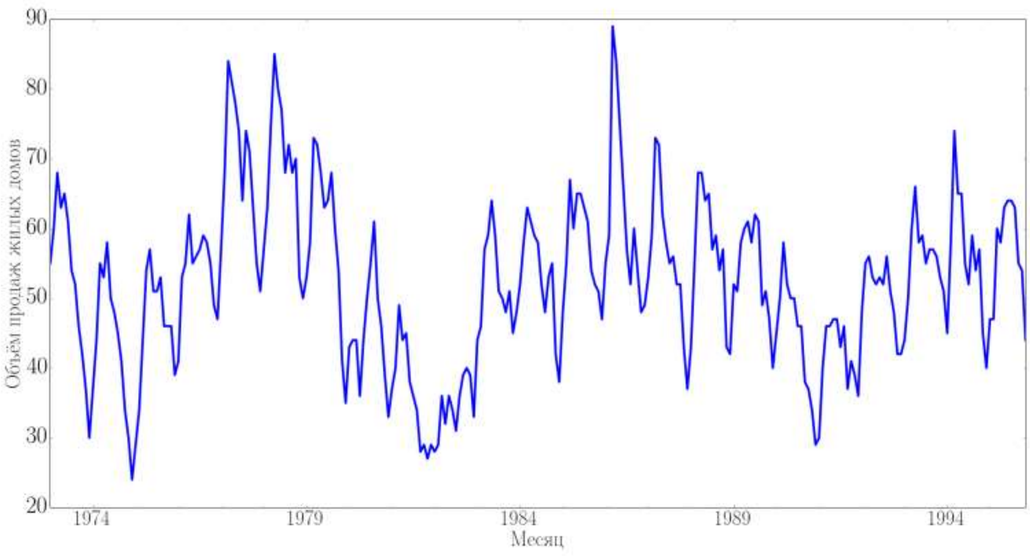
Количество контрактов за день в сокровищнице США

На следующем графике показаны суммарные объёмы электричества, произведённого за месяц в Австралии. **На графике**, как и в предыдущем случае, **виден тренд**, на этот раз повышающийся. Кроме того, **наблюдается годовая сезонность**: значение признака совершает колебания, минимум которых всегда приходится на зиму, а максимум — на середину лета. Это легко объяснить тем, что зимой электричества необходимо меньше всего, это самый тёплый сезон в Австралии



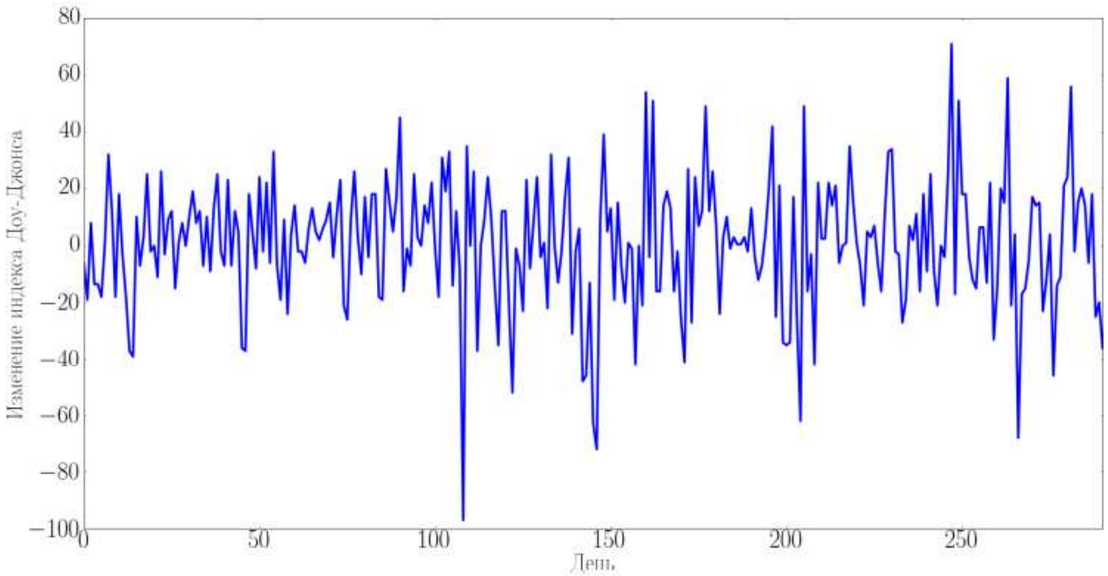
Суммарный объём электричества, произведённого за месяц в Австралии

Следующий пример — суммарный объём проданной жилой недвижимости в Америке за месяц, данные так же собраны за несколько лет. **На графике наблюдается сочетание двух основных компонент. Первая компонента — это годовая сезонность** (минимум всегда приходится на зиму, а максимум — на середину лета), а **вторая — это циклы**, связанные с изменением среднего уровня экономической активности (период в данном случае составляет 7-9 лет).



Суммарный объём проданной жилой недвижимости (в млн кв. м.) в Америке за месяц

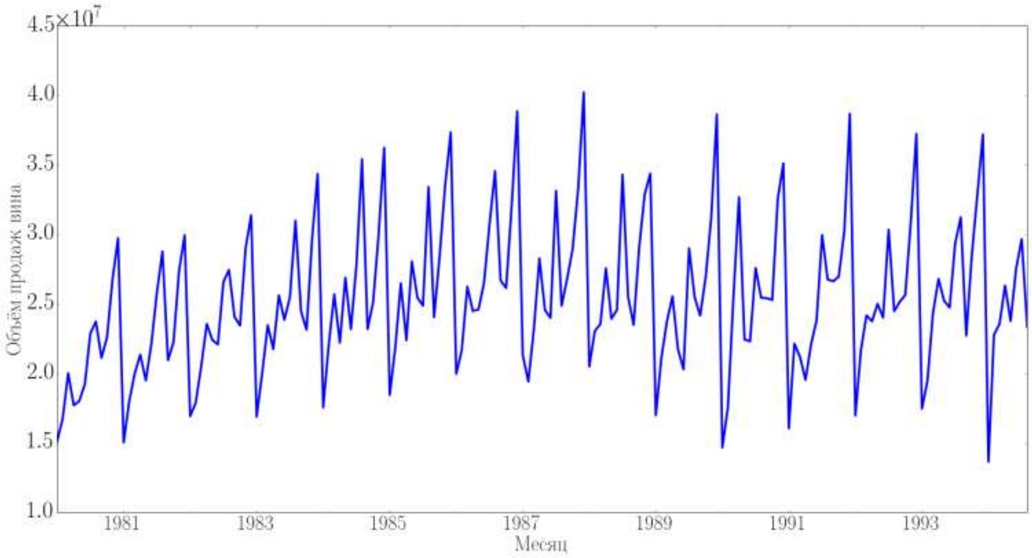
На следующем рисунке показаны ежедневные изменения индекса Доу-Джонса. Глядя на этот график, сложно сказать, присутствует ли в данных какая-то систематическая компонента: **явно нет ни тренда, ни сезонности, ни цикла**. По всей видимости, ряд представляет собой что-то похожее на случайную ошибку. Однако даже такие ряды можно прогнозировать.



Ежедневное изменение индекса Доу-Джонса

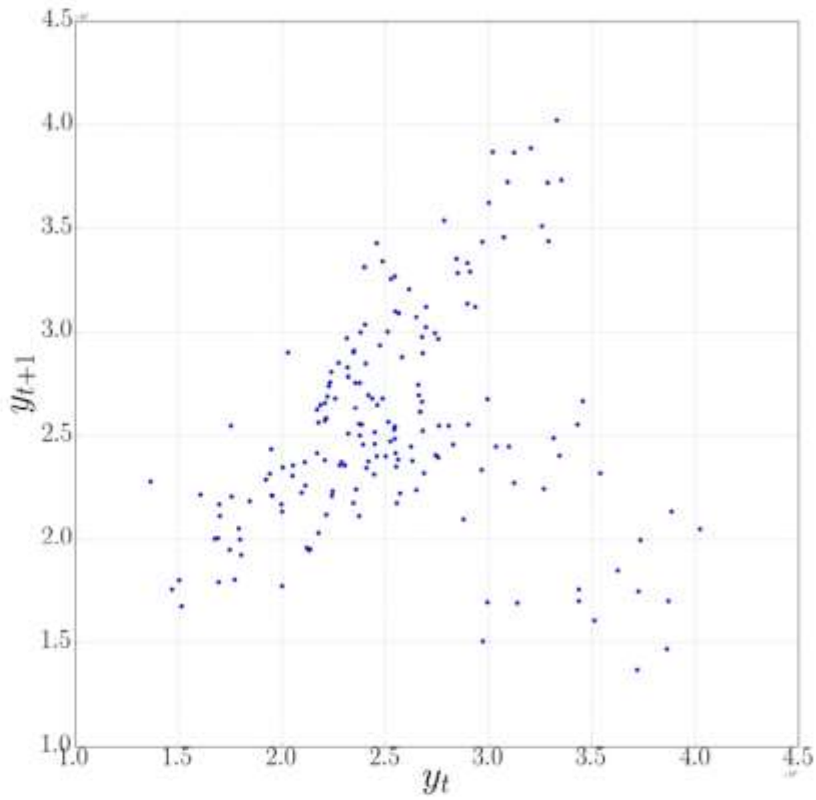
2. Автокорреляция

Одной из важнейших характеристик временного ряда является автокорреляция. Рассмотрим суть этой характеристики на примере данных о суммарном объёме продаж вина в Австралии за месяц на протяжении почти 15 лет.



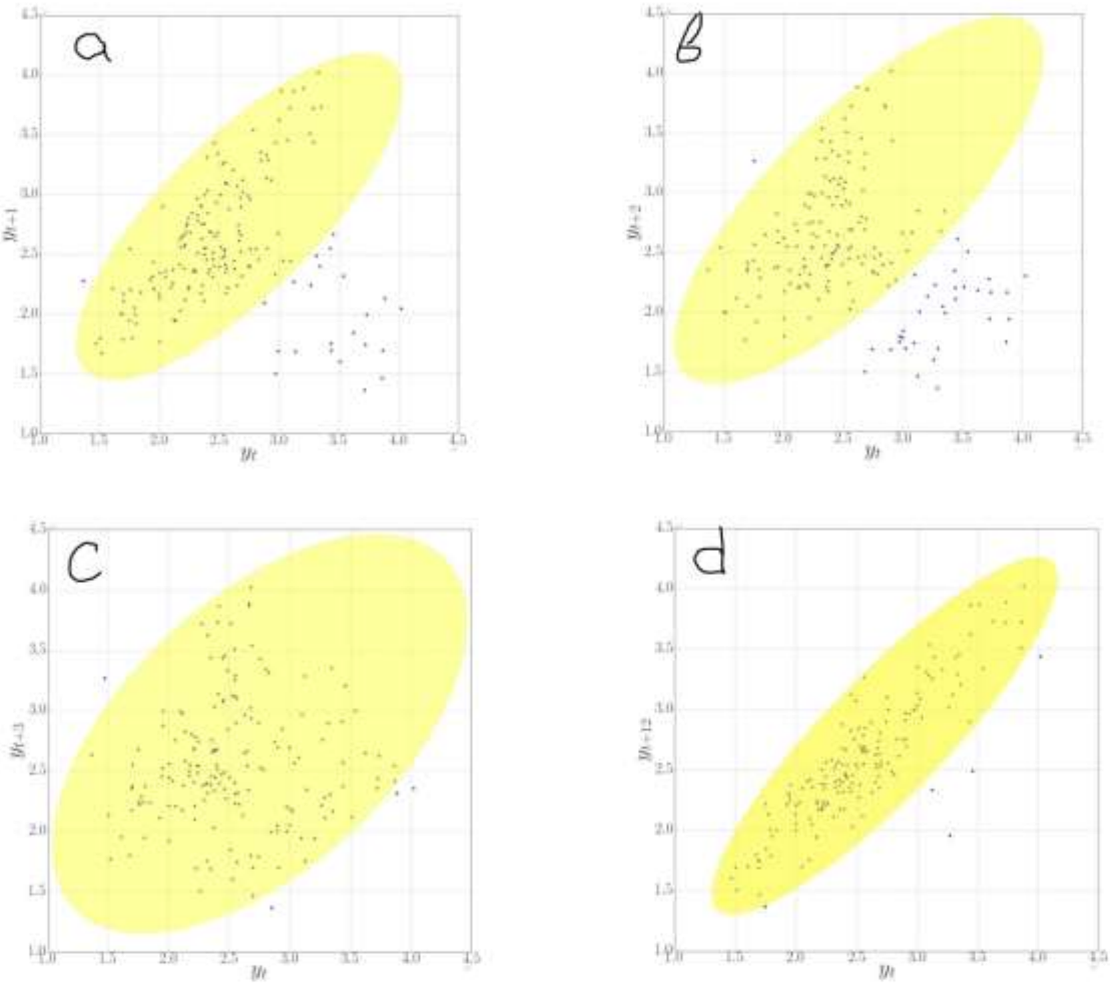
Месячный объём продаж вина в Австралии, в бутылках

Этот ряд обладает ярко выраженной годовой сезонностью: максимум продаж за год приходится на декабрь, а затем, в январе, происходит существенное падение.



Связь между значениями объёма продаж вина в соседние месяцы, по горизонтали отложен объём продаж в месяц t, по вертикали — в следующий месяц, t + 1, каждая точка задаёт продажи в 2 соседних месяца

Видно, что большая часть точек на графике группируется вокруг главной диагонали. Это говорит о том, что в основном значения продаж в соседние месяцы похожи. Ещё одно подмножество точек выделяется в правом нижнем углу, оно связано с падением продаж от декабря к январю, которое было видно на предыдущем графике.



Связь между продажами в соседние месяцы (a), через месяц (b), через два месяца(c) и через год (d).

Если построить аналогичный график, но по вертикальной оси отложить yt+2 (рис.b), то видно, что точки в основном облаке начинают «расплываться» вокруг главной диагонали, то есть сходство между продажами через месяц уменьшается по сравнению с соседними месяцами. Если посмотреть связь между продажами через два месяца (рис. c), то облако станет ещё шире, а сходство — ещё меньше. Однако если рассмотреть продажи в одни и те же месяцы соседних лет (рис. d), то видно, что точки на графике снова стягиваются к главной диагонали. Это значит, что значения продаж в одни и те же месяцы соседних лет очень сильно похожи.

2.1. Вычисление автокорреляции

Количественной характеристикой сходства между значениями ряда в соседних точках является автокорреляционная функция (или просто автокорреляция), которая задаётся следующим соотношением:

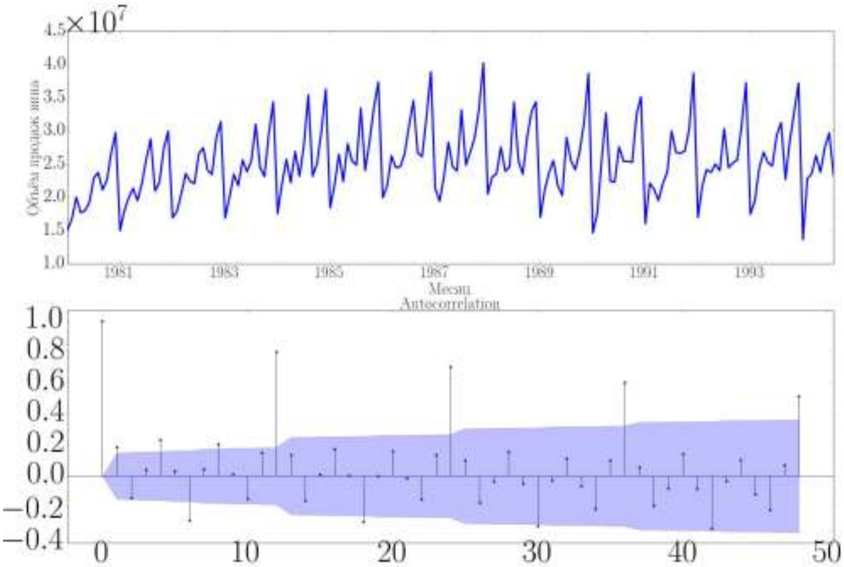
$$r_\tau = \frac{\mathbb{E}((y_t - \mathbb{E}y)(y_{t+\tau} - \mathbb{E}y))}{\mathbb{D}y}.$$

Автокорреляция — это уже встречавшаяся ранее корреляция Пирсона между исходным рядом и его версией, сдвинутой на несколько отсчётов. Количество отсчётов, на которое сдвинут ряд, называется лагом автокорреляции (τ). Значения, принимаемые автокорреляцией такие же, как и у коэффициента Пирсона: $r_\tau \in [-1, 1]$. Вычислить автокорреляцию по выборке можно, заменив в формуле математическое ожидание на выборочное среднее, а дисперсию — на выборочную дисперсию:

$$r_\tau = \frac{\sum_{t=1}^{T-\tau}(y_t - \bar{y})(y_{t+\tau} - \mathbb{E}y))}{\sum_{t=1}^{T-\tau}((y_t - \bar{y}))^2}.$$

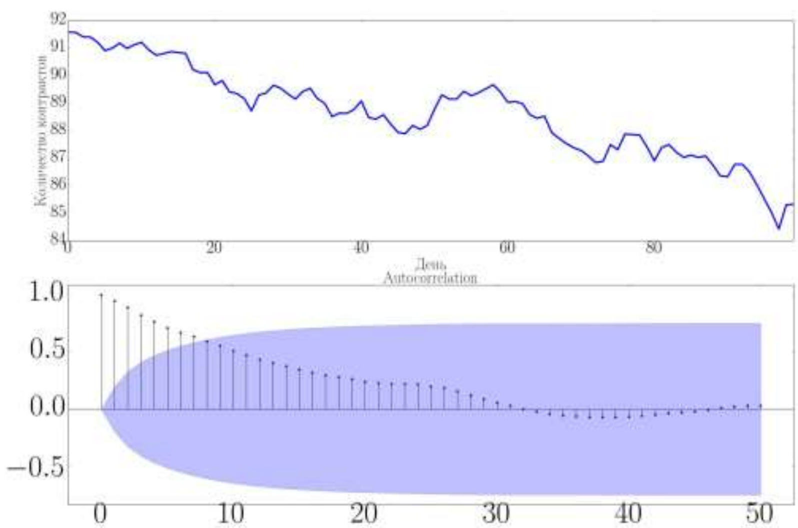
2.2. Коррелограммы

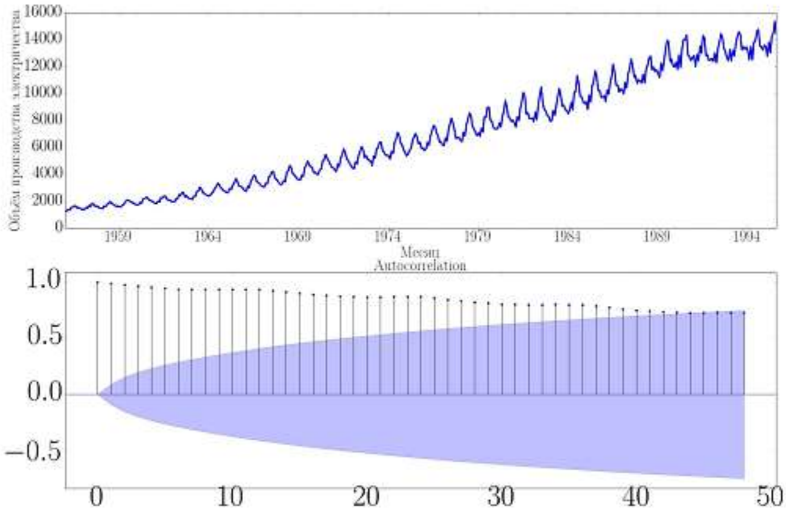
Анализировать величину автокорреляции при разных значениях лагов удобно с помощью графика, который называется **коррелограммой**. По оси ординат на нём откладывается автокорреляция, а по оси абсцисс — размер лага τ.



На рисунке показан пример коррелограммы для исследуемых ранее данных о месячных продажах вина в Австралии. На графике видно, что автокорреляция принимает большие значения в лагах, кратных сезонному периоду. **Такой вид коррелограммы типичен для данных с выраженной сезонностью.**

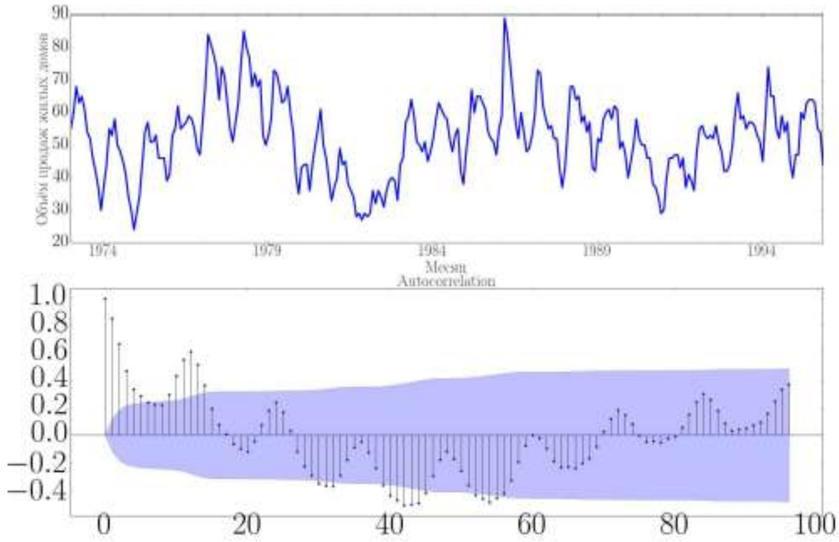
Так выглядит коррелограмма для данных с **ярко выраженным трендом**. Автокорреляция тем больше, чем меньше величина лага τ, и с ростом τ она начинает постепенно убывать, при этом автокорреляция может начать колебаться вокруг горизонтальной оси, соответствующей её нулевому значению.



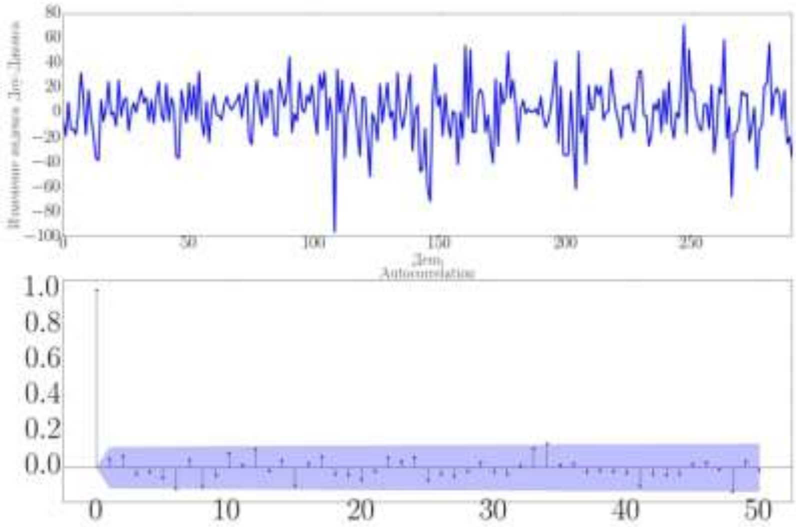


Коррелограмма построена для временного ряда, в котором присутствуют **и тренд, и сезонность**. На ней можно наблюдать оба описанных ранее эффекта, *однако тренд настолько сильный, что практически нейтрализует влияние сезонности* (следствие которой — наличие пиков в лагах, кратных периоду сезона).

Коррелограмма для ряда, в котором есть и **сезонность, и цикл**. Для самого первого лага, кратного сезонному периоду, виден пик, однако далее положение этого пика смещается: следующий пик не приходится на 2, 3 или 4 года. Это происходит, потому что в ряде есть циклы, период которых плавно меняется.



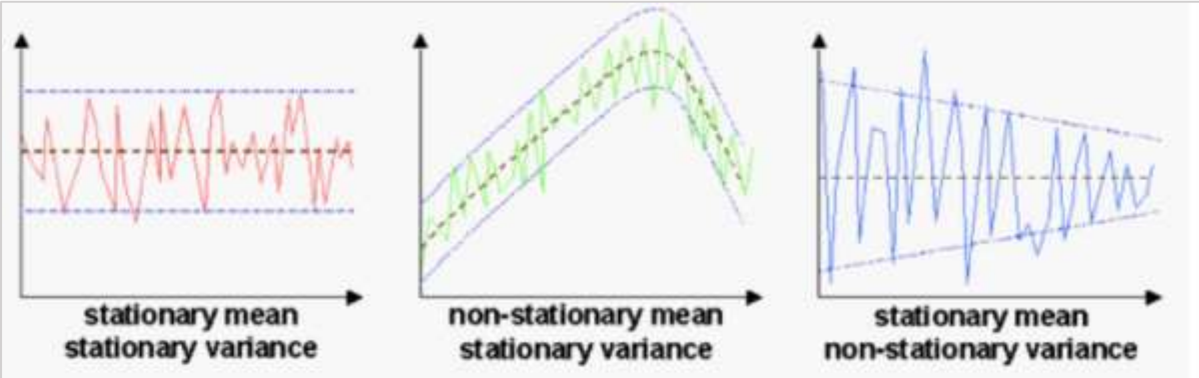
Значимость автокорреляции. На всех показанных коррелограммах изображён синий коридор вокруг горизонтальной оси. Это коридор значимости отличия корреляции от нуля. Автокорреляции, которые изображены вне этого коридора, значимо отличаются от нуля.

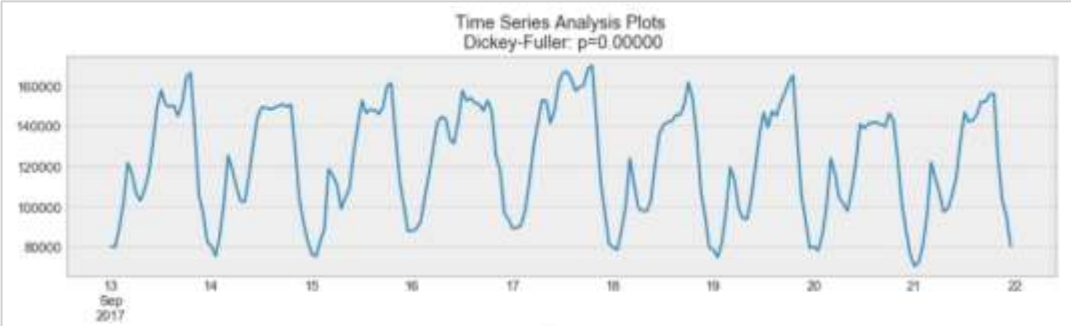


На коррелограмме по данным о ежедневном изменении индекса Доу-Джонса, ни одна из корреляций не выходит за пределы коридора значимости, а значит ни одна из них не является значимо отличающейся от нуля.

3. Стационарность

*Временной ряд называется **стационарным**, если его статистические свойства* (постоянное среднее и дисперсия) *не изменяются со временем*.



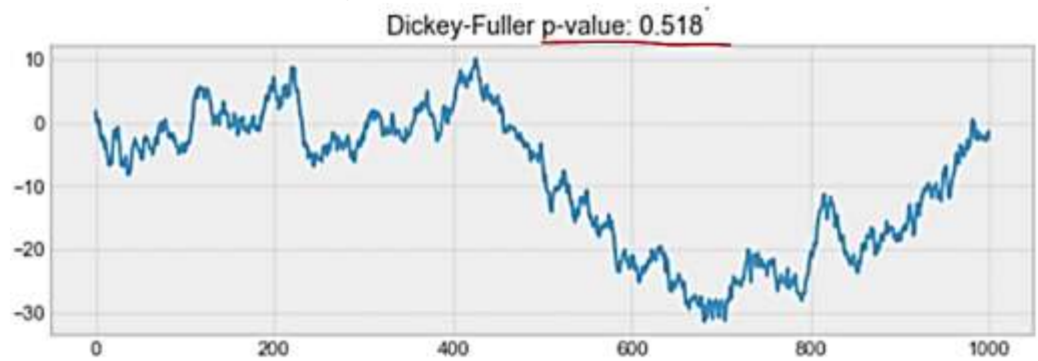


Пример стационарного процесса

В **нестационарных временных рядах** статистические свойства меняются со временем. Они показывают сезонные эффекты, тренды и другие структуры, которые зависят от временного показателя. *Пример — международные перелеты авиакомпаний. Количество пассажиров на тех или иных направлениях меняется в зависимости от сезонности.*

Критерий Дики-Фуллера

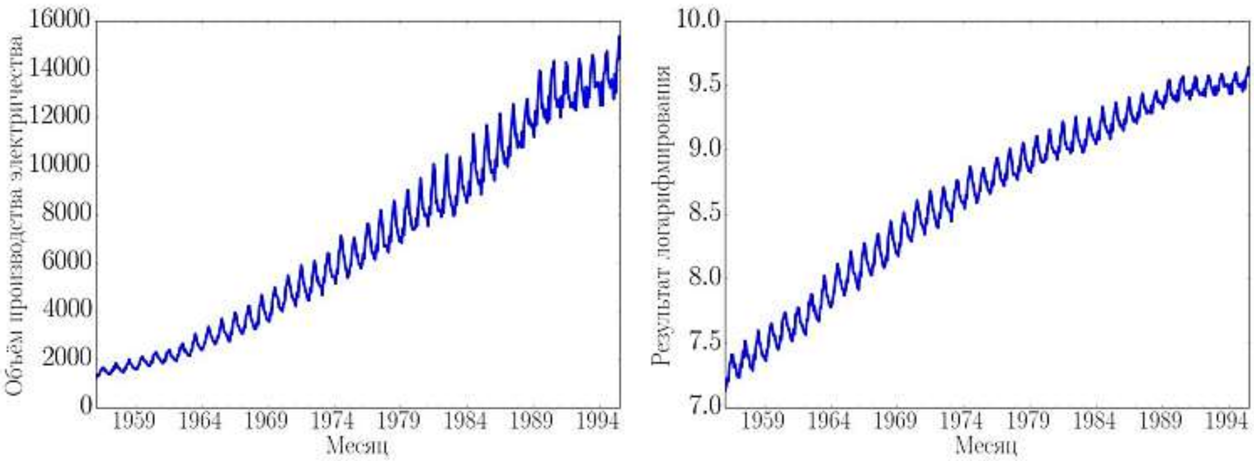
Гипотезу о стационарности можно проверить с помощью критерия Дики-Фуллера.
Не вдаваясь в технические детали теста Дики-Фуллера, он проверяет нулевую гипотезу о наличии единичного корня.
Если да, то ***p*> 0** и процесс **не стационарный**.
В противном случае ***p* = 0**, нулевая гипотеза отклоняется, и процесс считается **стационарным**.
В качестве примера приведенный ниже процесс не является стационарным. Обратите внимание на то, что среднее значение не является постоянным во времени.



Пример нестационарного процесса

Как сделать ряд стационарным?

При работе с нестационарными временными рядами используется ряд стандартных трюков, чтобы сделать их стационарными. В случае, **если во временном ряде монотонно по времени изменяется дисперсия, применяется специальное преобразование, стабилизирующее дисперсию**. Очень часто в качестве такого преобразования выступает **логарифмирование**.

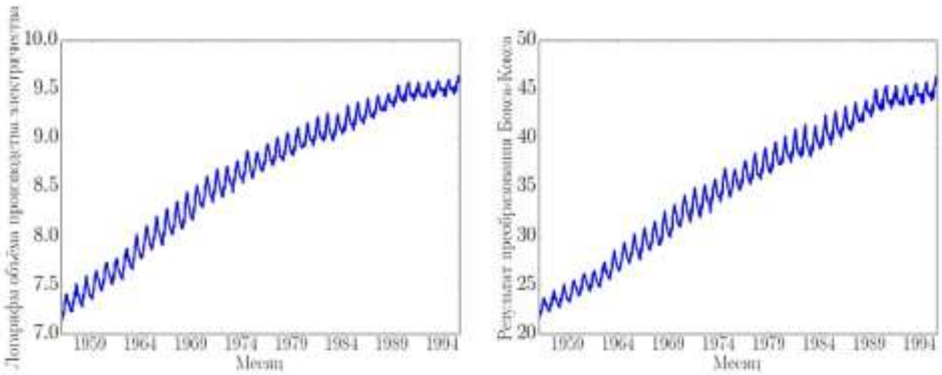


Временной ряд до и после логарифмирования. Видно, что после логарифмирования размах колебаний в начале и конце ряда становится очень похожим, и дисперсия примерно стабилизируется.

Логарифмирование принадлежит к семейству преобразований Бокса-Кокса.

$$y'_t = \begin{cases} \ln y_t, & \lambda = 0, \\ \left(y_t^\lambda - 1\right)/\lambda, & \lambda \neq 0. \end{cases}$$

Параметр λ определяет, как именно будет преобразован ряд: λ = 0 — это логарифмирование, λ = 1 — тождественное преобразование ряда, а при других значениях λ — степенное преобразование. Значение параметра можно подбирать так, чтобы дисперсия была как можно более стабильной во времени.



Слева — результат логарифмирования,

справа — преобразование Бокса-Кокса с параметром λ = 0.27, при котором дисперсия выглядит более стабильной, чем при логарифмировании.

Ещё один важный способ, который позволяет сделать ряд стационарным, — это **дифференцирование**, переход к попарным разностям соседних значений:

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

Для нестационарного ряда часто оказывается, что получаемый после дифференцирования ряд является стационарным.

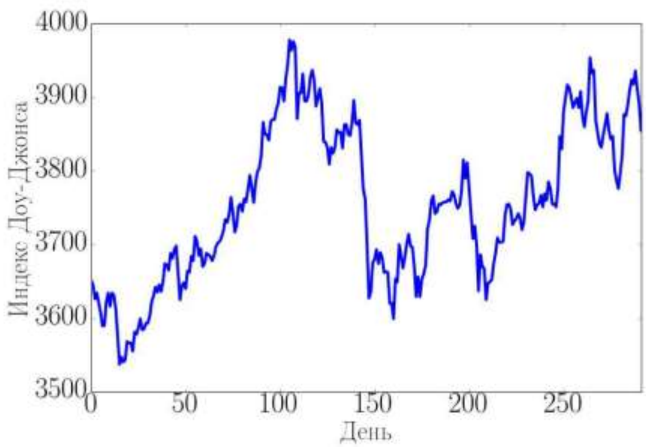
Также может применяться **сезонное дифференцирование ряда**, переход к попарным разностям значений в соседних сезонах. Если длина периода сезона составляет s, то новый ряд задаётся разностями

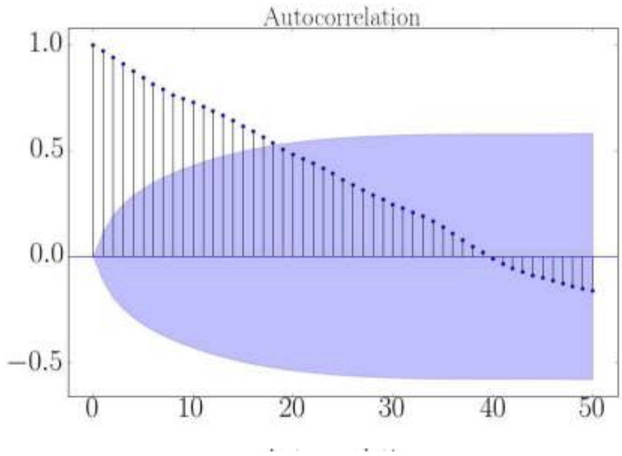
$$y'_t = y_t - y_{t-s}.$$

Сезонное и обычное дифференцирование могут применяться к ряду в любом порядке. Однако *если у ряда есть ярко выраженный сезонный профиль, то рекомендуется начинать с сезонного дифференцирования, уже после такого преобразования может оказаться, что ряд стационарен*.

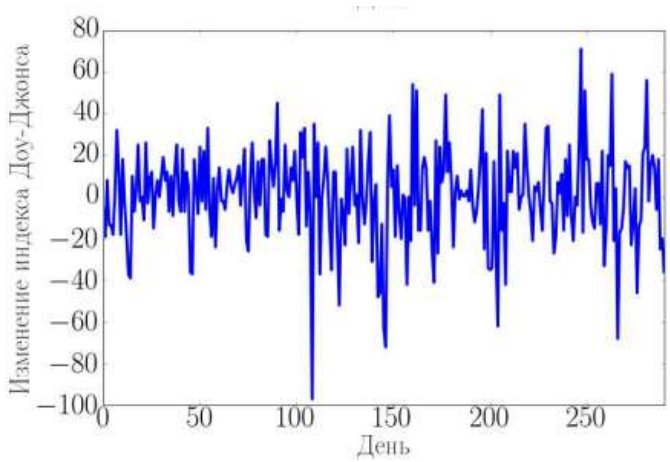
Пример

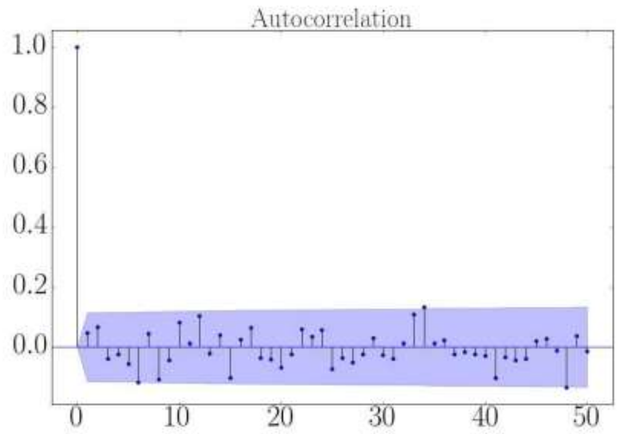
Ниже на рисунке ряд значений индекса Доу-Джонса и его автокорреляционная функция. Видно, что этот ряд достаточно сильно нестационарен — имеется ярко выраженный тренд.





От этого тренда удаётся полностью избавиться, продифференцировав ряд.





4. Модели прогнозирования класса ARMA

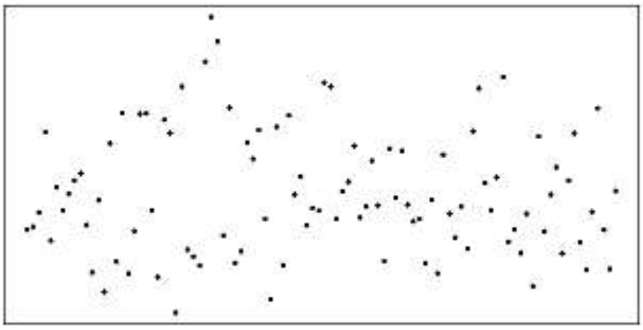
4.1. Авторегрессионная модель (AR). В ней значения в будущем определяются как значения из прошлого, умноженные на коэффициенты:

$$y_t = \alpha + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t.$$

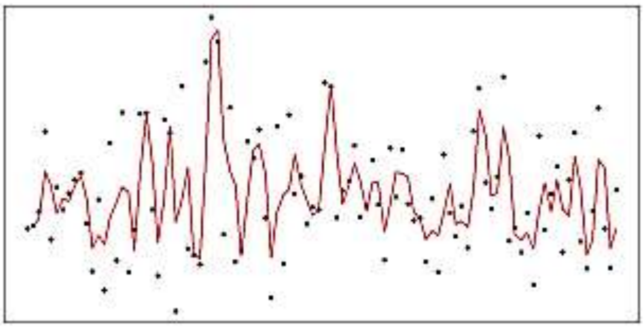
В этом регрессионном уравнении y_t — это отклик, $y_{t-1}, y_{t-2}, \ldots, y_{t-p}$ — признаки, $\alpha, \phi_1, \phi_2, \ldots, \phi_p$ — параметры модели, которые необходимо оценить, ε_t — шумовая компонента, описывает отклонения значений ряда от данного уравнения. Такая модель называется моделью авторегрессии порядка p (AR(p)). В этой модели yt представляет собой линейную комбинацию p предыдущих значений ряда и шумовой компоненты.

4.2. Скользящее среднее (MA).

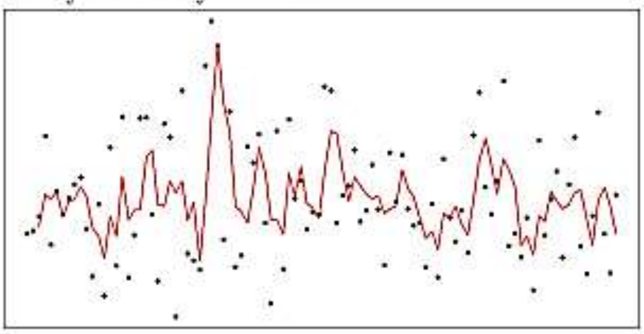
Следующий класс моделей — это скользящее среднее. Чтобы лучше понимать, как они устроены, можно рассмотреть независимый, одинаково распределённый во времени шум ε_t:



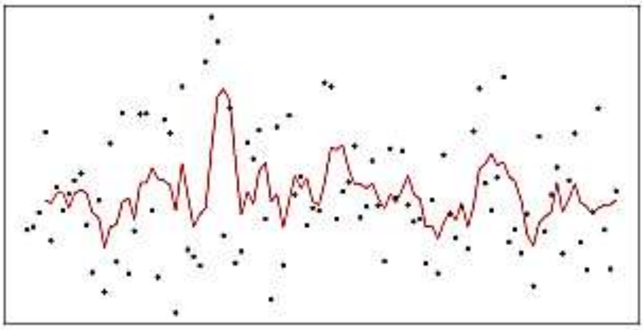
(a) Независимый, одинаково распределённый во времени случайный шум



(b) Среднее по двум соседним точкам



(c) Среднее по трём соседним точкам



(d) Среднее по четырём соседним точкам

То, что получается в результате такого усреднения, — это уже не простая выборка с независимыми, одинаково распределёнными элементами. Соседние значения на красной линии очень похожи друг на друга, потому что в их вычислении используются одни и те же шумовые компоненты. Данную идею можно обобщить и записать следующую модель ряда:

$$y_t = \alpha + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q},$$

где $\boldsymbol{\varepsilon_t, \varepsilon_{t-1}, \ldots, \varepsilon_{t-q}}$ – значения шума в q предыдущих моментах времени,

α, θ₁, θ₂, . . . , θq — это параметры модели, которые необходимо оценить.

Такая модель называется моделью скользящего среднего порядка q (**MA(q)**). В ней предполагается, что значение ряда yt — это линейная комбинация q последних значений шумовой компоненты.

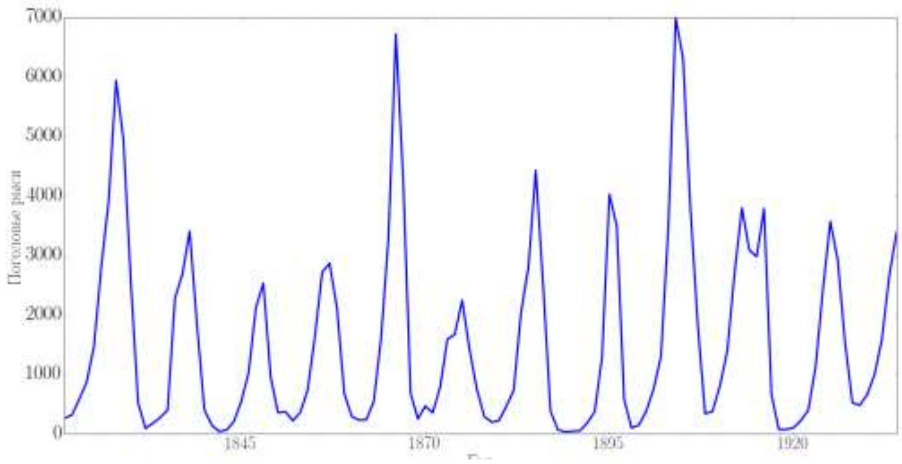
4.3. Авторегрессионная модель скользящей средней (ARMA)

Если взять авторегрессионную модель порядка p (AR(p)) и модель скользящего среднего порядка q (MA(q)) и сложить то, что находится у них в правых частях, то в результате получим это модель ARMA(p, q), она выглядит следующим образом:

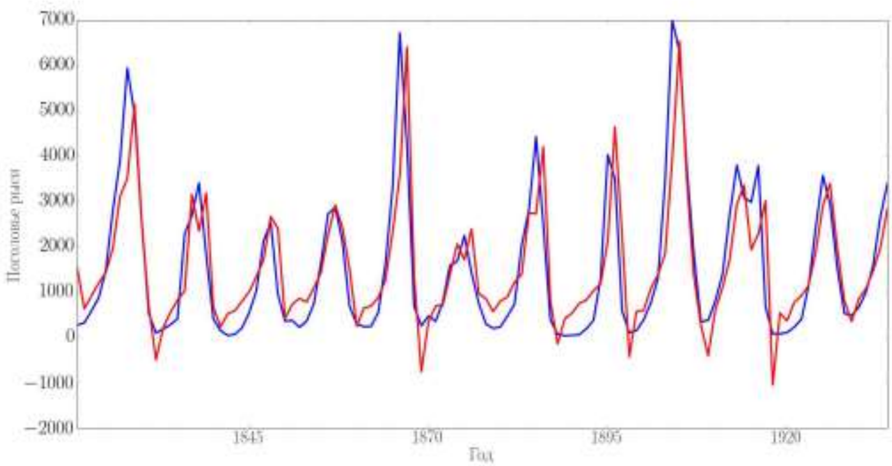
$$y_t = \alpha + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}.$$

Теорема Вольда утверждает, что любой стационарный временной ряд может быть описать моделью ARMA(p, q) с правильным подбором значений параметров p, q.

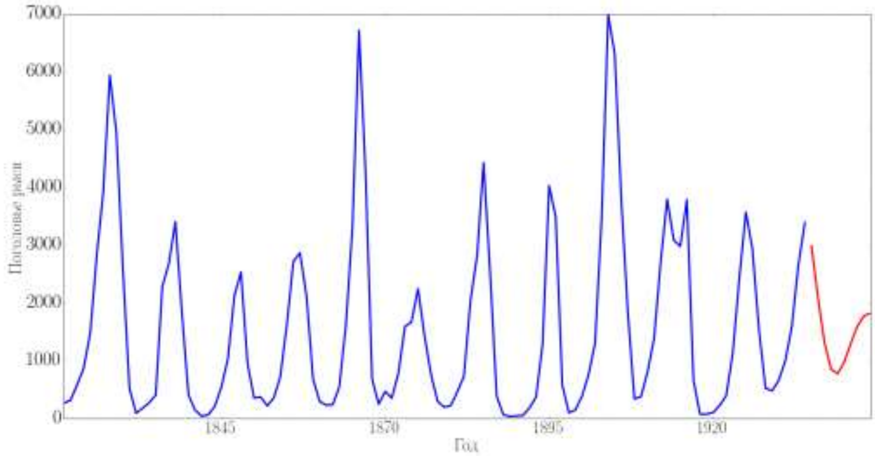
Пример. Для демонстрации работы модели ARMA(p, q) можно рассмотреть данные о поголовье рыси. Ряд стационарен, а значит в классе ARMA(p, q) для него можно найти достаточно хорошее описание.



Модель **ARMA(2, 2)** даёт результат, который достаточно сильно похож на исходный ряд. Модель не во всех точках близка к истинному значению ряда, однако результат всё равно намного лучше, чем если бы для приближения использовалась регрессия на линейный или квадратичный временной тренд.



Модель ARMA(2, 2) можно использовать и для построения прогноза, то есть решения той задачи, которая была изначально поставлена.



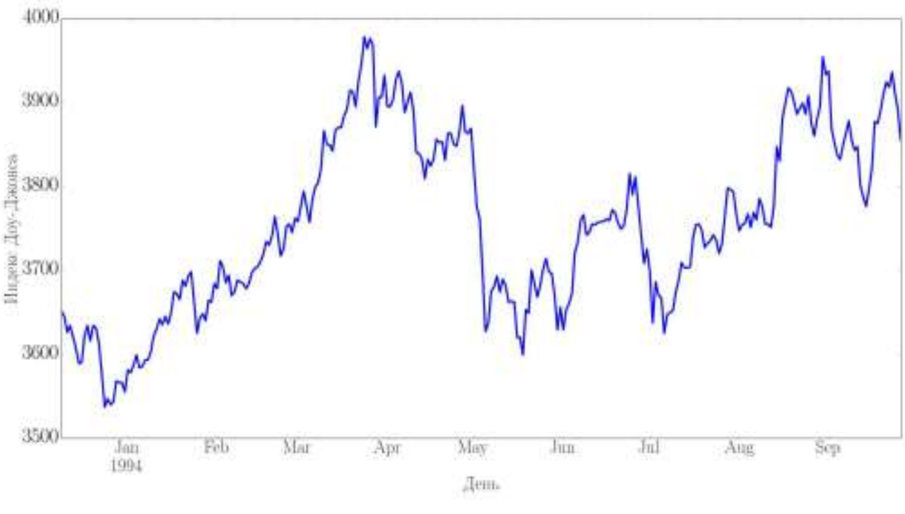
5. Модели класса ARIMA

Модели типа ARIMA — это обобщение модели класса ARMA.

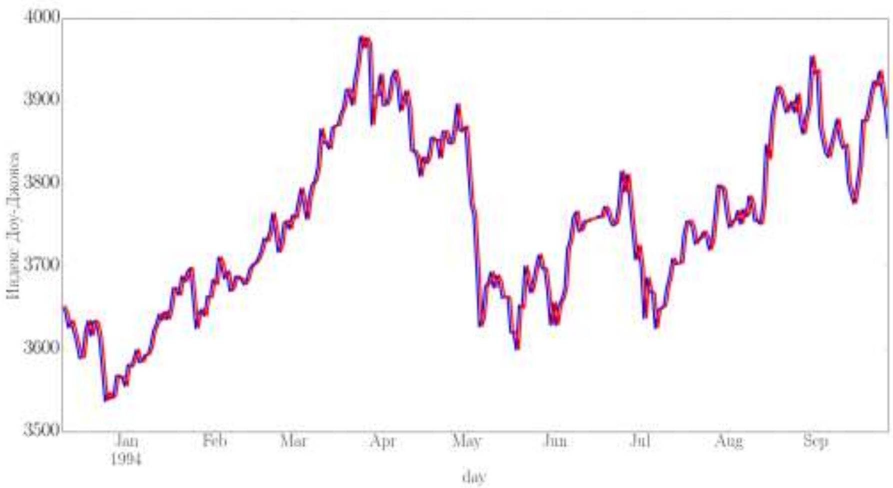
При помощи дифференцирования нестационарный ряд можно сделать стационарным, а любой стационарный ряд может быть описан моделью ARMA(p, q).

Эти две идеи и лежат в основе моделей класса ARIMA. Модель ARIMA(p, d, q) — это модель ARMA(p, q) для d раз продифференцированного ряда.

Рассмотрим на примере: Даны 300 значений индекса Доу-Джонса. Этот ряд не стационарен, но ранее было показано, что стационарен ряд его первых разностей.



Из этого следует, что для ряда разностей можно подобрать достаточно хорошую модель в классе ARMA. Если сделать это, а затем произвести операцию, обратную дифференцированию, то в результате будет получена модель ARIMA для исходного ряда.



На данном рисунке показана модель ARIMA(0, 1, 0). В этой модели происходит одно дифференцирование и не используется ни одной компоненты авторегрессии и скользящего среднего, и это немного странно, но результат в любом случае лучше, чем то, что можно было бы получить с помощью регрессии ряда на временные признаки.

6. Модели для временных рядов с сезонными компонентами

Для работы с сезонными временными рядами используются модели **SARMA** и **SARIMA**. Это расширения моделей ARMA и ARIMA соответственно, добавляющие в них сезонные условия.

SARMA. Пусть ряд имеет сезонный период длины S. Тогда можно взять модель ARMA(p, q):

$$y_t = \alpha + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q},$$

добавить к этой модели P авторегрессионных компонент, но не предыдущих, а с шагом, равным периодом сезонности:

$$+\phi_S y_{t-S} + \phi_{2S} y_{t-2S} + \cdots + \phi_{PS} y_{t-PS}$$

и Q компонент скользящего среднего, также с шагом, равным периодом сезонности:

$$+\theta_S \varepsilon_{t-S} + \theta_{2S} \varepsilon_{t-2S} + \cdots + \theta_{PS} \varepsilon_{t-QS}.$$

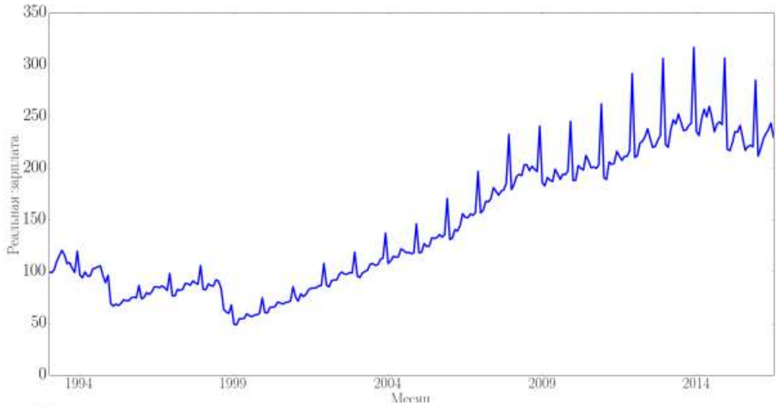
Результат — это модель SARMA(p, q) × (P, Q).

Модель **SARIMA**(p, d, q) × (P, D, Q) — модель SARMA(p, q) × (P, Q) для ряда, к которому d раз было применено обычное дифференцирование и D раз — сезонное. Такую модель часто называют просто ARIMA: первая буква не пишется, но подразумевается, что сезонная компонента тоже может быть.

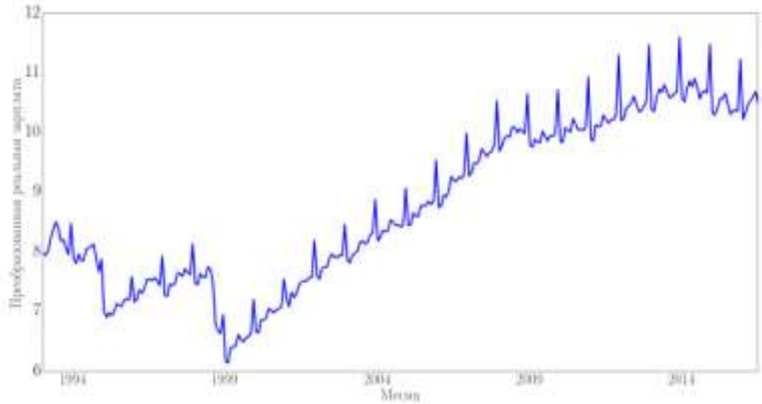
Пример

Для демонстрации рассмотренных моделей будет использоваться временной ряд реальной заработной платы в России.

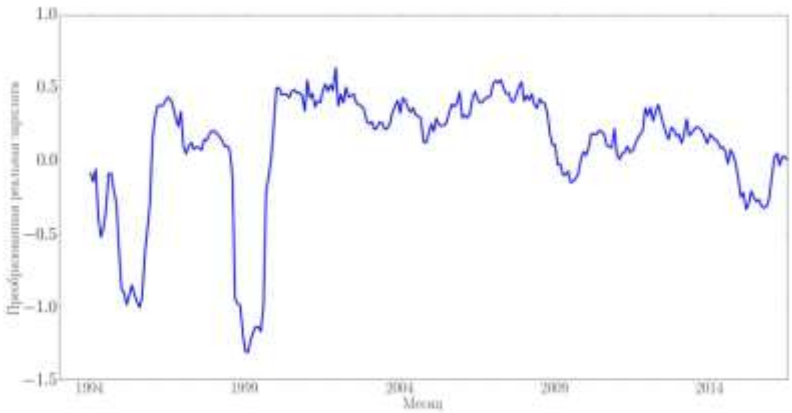
1. Ряд нестационарный, видно, что меняется дисперсия: разброс скачков в начале совсем не такой, как ближе к концу. Критерий Дики-Фуллера не отвергает гипотезу о том, что этот ряд нестационарный (p = 0.2265).



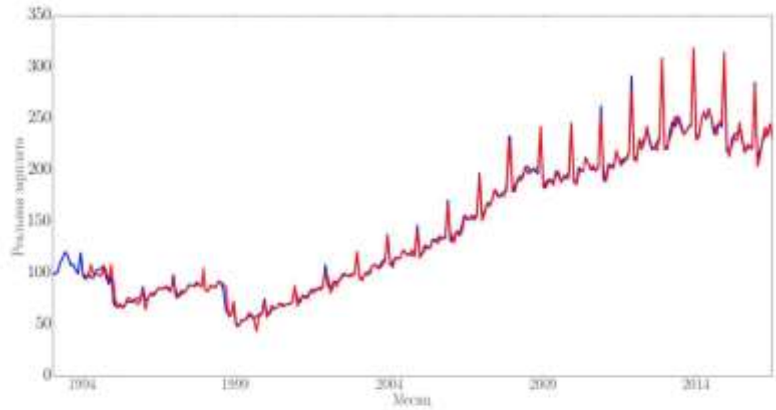
2. Ряд после применения преобразования Бокса-Кокса с параметром λ = 0.22. Критерий Дики-Фуллера всё ещё не отвергает для этого ряда гипотезу о нестационарности (p = 0.1661). Это можно объяснить наличием в ряду сезонности и тренда.



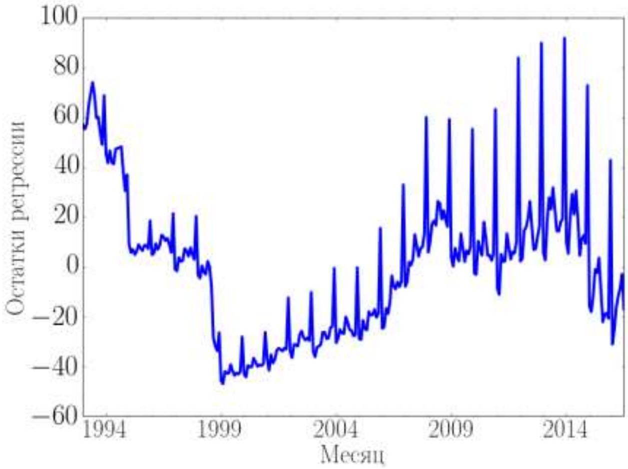
3. После применения к ряду сезонного дифференцирования критерий Дики-Фуллера отвергает гипотезу о нестационарности (p = 0.01). Относительно этого ряда можно говорить, что он стационарный, а значит, можно попытаться подобрать для него модель в классе ARMA или даже сезонную модель.

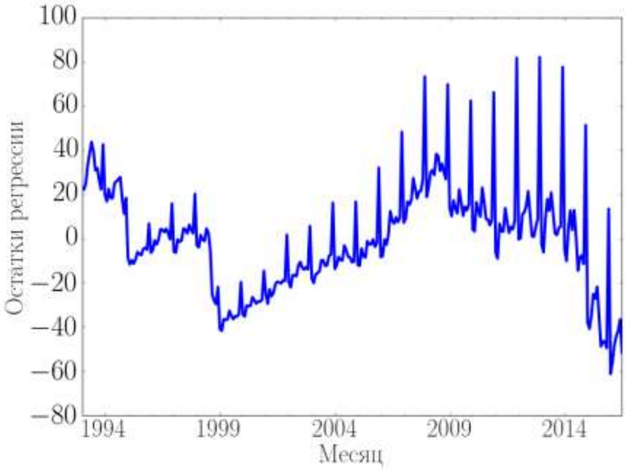


4. После обратных преобразований к преобразованию Бокса-Кокса и сезонному дифференцированию. Красная линия на графике — это предсказание модели, видно, что она достаточно хорошо описывает исходные данные, а значит, можно надеяться, что и прогнозы она будет давать хорошие.

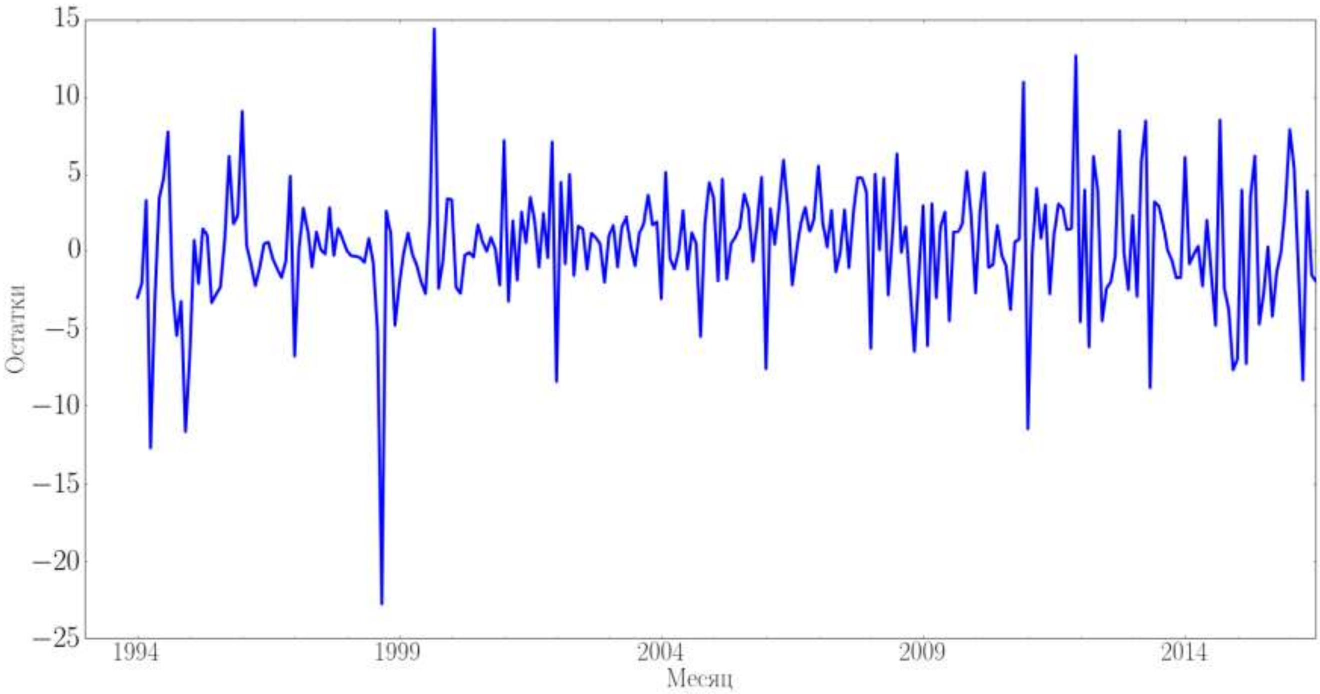


При применении регрессии с линейным или квадратичным трендом по времени в остатках этой модели было видно достаточно много структуры, а значит, что в данных оставалось много информации, которую не учитывает модель.





А вот остатки для построенной модели SARIMA уже гораздо больше похожи на белый шум. Выброс в остатках — это кризис 1998 года, который плохо описывается построенной моделью. Тем не менее, в этих остатках уже практически не имеется структуры, а значит, полученный результат лучше, чем при использовании линейной регрессии.



7. Подбор параметров

У моделей класса ARIMA есть несколько групп параметров. Параметры d, D, q, Q, p, P можно считать гиперпараметрами, поскольку они определяют структуру и количество коэффициентов в самой модели ARIMA.

Параметры d, D

Параметры d, D, которые задают порядки дифференцирования, необходимо подбирать так, чтобы ряд стал стационарным. Ранее уже упоминалось, что всегда рекомендуется начинать с сезонного дифференцирования, потому что уже после него ряд может оказаться стационарным. Дело в том, что выгодно дифференцировать ряд как можно меньше раз, потому что с **увеличением количества дифференцирований растёт дисперсия итогового прогноза**.

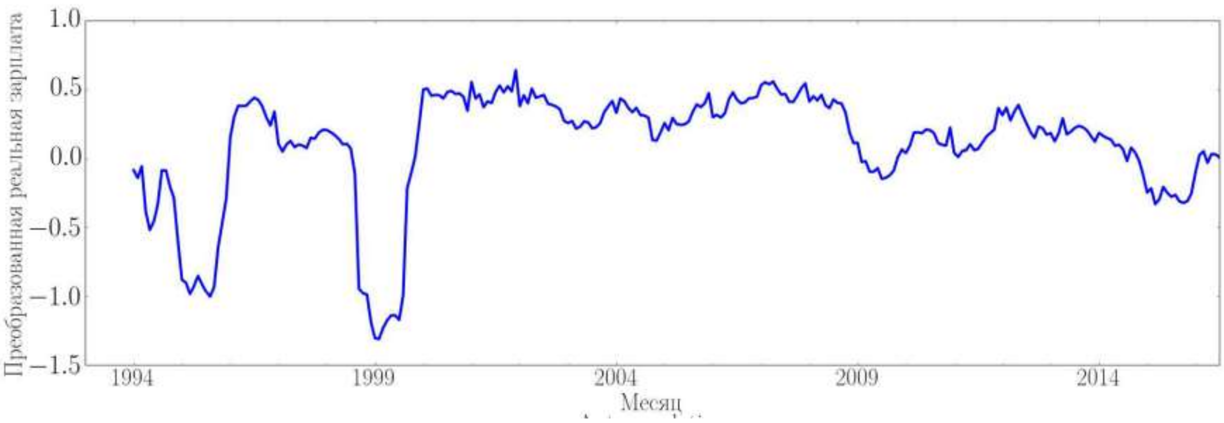
Параметры q, Q, p, P

К сожалению, гиперпараметры q, Q, p, P нельзя выбирать из принципа максимума правдоподобия. Например, чем больше значение параметра p, тем больше параметров φ и тем лучше это уравнение описывает данные. Чем больше значения гиперпараметров, тем больше параметров в модели и тем она сложнее. Таким образом, с увеличением значения этих гиперпараметров значение правдоподобия может только увеличиваться. Поэтому для сравнения моделей с разным количеством параметров необходим другой критерий. В качестве искомого критерия можно использовать, например, критерий AIC:

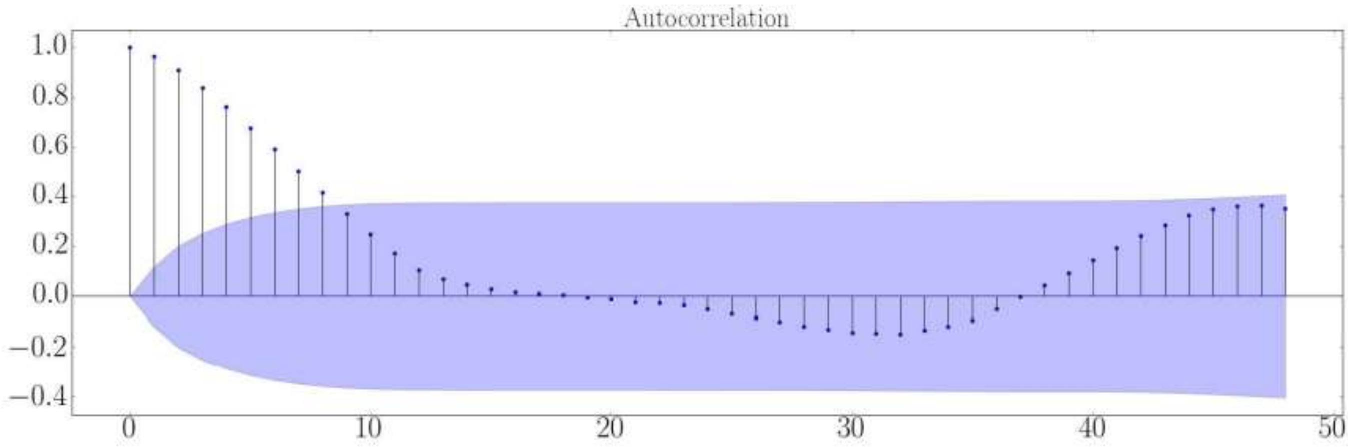
$$AIC = -2 \ln L + 2k,$$

где L — правдоподобие, k = P + Q + p + q + 1 — число параметров в модели. Оптимальной по критерию AIC будет модель с наименьшим значением этого критерия. Такая модель, с одной стороны, будет достаточно хорошо описывать данные, а с другой — содержать не слишком большое количество параметров.

В конечном итоге значения параметров q, Q, p, P определяются перебором: из разных значений гиперпараметров выбираются те, у которых значение критерия AIC будет минимальным.



На рисунке ряд реальной заработной платы в России после преобразования Бокса-Кокса и сезонного дифференцирования, а далее автокорреляционная функция этого ряда.



Определение параметров q и p по коррелограмме

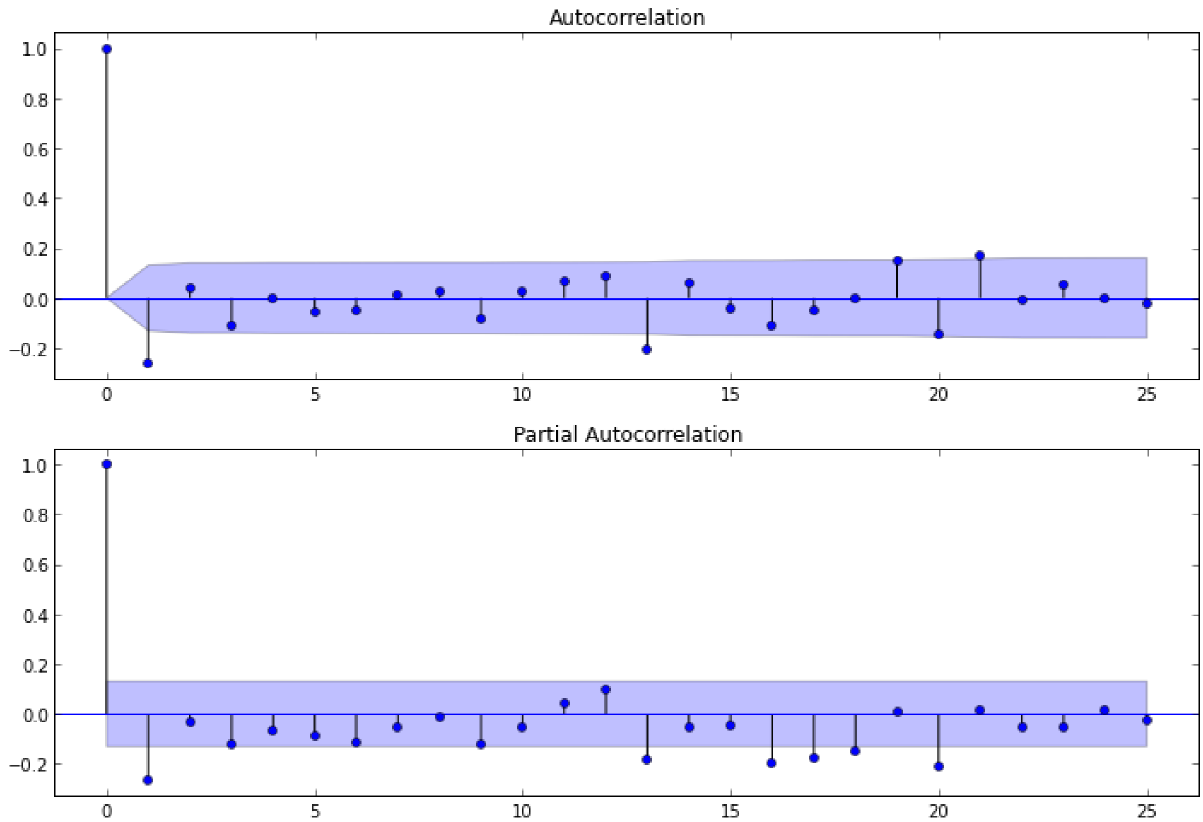
https://habr.com/ru/post/207160/

Для их определения нам надо изучить **автокорреляционную** (**ACF**) и частично **автокорреляционную**(**PACF**) функции для ряда первых разностей.

ACF поможет нам определить **q**, т. к. по ее коррелограмме можно определить количество автокорреляционных коэффициентов сильно отличных от 0 в модели MA

PACF поможет нам определить **p**, т. к. по ее коррелограмме можно определить максимальный номер коэффициента сильно отличный от 0 в модели AR.

В графиках ACF и PACF по оси X откладываются номера лагов, а по оси Y значения соответствующих функций. Нужно отметить, что количество лагов в функциях и определяет число значимых коэффициентов.



После изучения коррелограммы **PACF** можно сделать вывод, что p = 1, т.к. на ней только 1 лаг сильно отличен от нуля. По коррелограмме **ACF** можно увидеть, что q = 1, т.к. после лага 1 значении функций резко падают.

8. Анализ остатков

Анализ остатков — это техника, которая помогает понять, есть ли у прогнозирующей модели небольшие недостатки, которые можно устранить доработкой, или же фундаментальные проблемы.

Остатки — это разность между фактом и прогнозом:

$$\hat{\varepsilon}_t = y_t - \hat{y}_t.$$

Остатки оценивают ошибку, то есть шумовую компоненту, которую наблюдать невозможно. При построении модели делаются предположения об этой шумовой компоненте, и логично, что свойства остатков должны согласовываться с выдвинутыми предположениями:

- ✓ **Несмещённость**, то есть в среднем остатки должны быть равны нулю.
- ✓ **Стационарность**, то есть отсутствие зависимости от времени. Таким образом, остатки во времени должны быть распределены примерно одинаково.
- ✓ **Неавтокоррелированность**, то есть отсутствие зависимости от предыдущих наблюдений.

Хороший практический пример анализа временного ряда:

https://www.kaggle.com/code/andreshg/timeseries-analysis-a-complete-guide