Bootcamp Data Science
Zajęcia 4

Przemysław Spurek

Szeregi czasowe

Time-Series Data, Time-related data – dane zmieniające się wraz z upływem czasu; dane zawierające serie (szeregi) wartości/wielkości zmieniających się w czasie.

Szeregi czasowe

Time-Series Data, Time-related data – dane zmieniające się wraz z upływem czasu; dane zawierające serie (szeregi) wartości/wielkości zmieniających się w czasie.

Szereg czasowy

Szereg czasowy – ciąg obserwacji pewnego zjawiska w kolejnych jednostkach czasu [def. statystyczna].

Szeregi czasowe

Cel analizy szeregów czasowych:

- Zbudowanie modelu pewnego zjawiska/procesu w oparciu o obserwowane zmiany w czasie pewnych mierzalnych wielkości opisujących ten proces.

Szeregi czasowe

Cel analizy szeregów czasowych:

- Zbudowanie modelu pewnego zjawiska/procesu w oparciu o obserwowane zmiany w czasie pewnych mierzalnych wielkości opisujących ten proces.
- Ogólne założenie: obserwowany przebieg składa się z:
  - Części systematycznej (trend, składowa stała, wahania sezonowe i cykliczne) – w oparciu, o które buduje się model.
  - Części przypadkowej (szumu, wahań przypadkowych).

Szeregi czasowe

Cel analizy szeregów czasowych:

- Zbudowanie modelu pewnego zjawiska/procesu w oparciu o obserwowane zmiany w czasie pewnych mierzalnych wielkości opisujących ten proces.
- Ogólne założenie: obserwowany przebieg składa się z:
  - Części systematycznej (trend, składowa stała, wahania sezonowe i cykliczne) – w oparciu, o które buduje się model.
  - Części przypadkowej (szumu, wahań przypadkowych).
- Wymienione składniki – czynniki determinujące rozważane zjawisko. W analizie szeregów dąży się do ich wyodrębnienia i pomiaru – dekompozycja szeregu czasowego.

Szeregi czasowe

Cel analizy szeregów czasowych:

- Zbudowanie modelu pewnego zjawiska/procesu w oparciu o obserwowane zmiany w czasie pewnych mierzalnych wielkości opisujących ten proces.
- Ogólne założenie: obserwowany przebieg składa się z:
  - Części systematycznej (trend, składowa stała, wahania sezonowe i cykliczne) – w oparciu, o które buduje się model.
  - Części przypadkowej (szumu, wahań przypadkowych).
- Wymienione składniki – czynniki determinujące rozważane zjawisko. W analizie szeregów dąży się do ich wyodrębnienia i pomiaru – dekompozycja szeregu czasowego.
- Przy użyciu otrzymanego modelu można dokonywać predykcji (eksploracji) przebiegu szeregu lub jego składowych.

Szeregi czasowe

Podstawowa struktura szeregów czasowych:

- **Trend** (tendencja rozwojowa) – reprezentuje ogólny kierunek rozwoju zjawiska (systematyczne zmiany jakim podlega zjawisko); rozróżnia się na przykład trend liniowy lub nieliniowy.

Szeregi czasowe

Podstawowa struktura szeregów czasowych:

- **Trend** (tendencja rozwojowa) – reprezentuje ogólny kierunek rozwoju zjawiska (systematyczne zmiany jakim podlega zjawisko); rozróżnia się na przykład trend liniowy lub nieliniowy.
- **Składowa okresowa** (wahania okresowe / regularne odchylenia od tendencji rozwojowej) – składnik powtarzający się cyklicznie.

Szeregi czasowe

Podstawowa struktura szeregów czasowych:

- **Trend** (tendencja rozwojowa) – reprezentuje ogólny kierunek rozwoju zjawiska (systematyczne zmiany jakim podlega zjawisko); rozróżnia się na przykład trend liniowy lub nieliniowy.
- **Składowa okresowa** (wahania okresowe / regularne odchylenia od tendencji rozwojowej) – składnik powtarzający się cyklicznie.
- **Szum** (zakłócenia, wahania przypadkowe).

`https://github.com/przem85/bootcamp/blob/master/statistics/D15_Z01.ipynb`

lags (opóźnienia)

- Modelowanie szeregów czasowych zakłada związek między daną obserwacją, a poprzednią.

lags (opóźnienia)

- Modelowanie szeregów czasowych zakłada związek między daną obserwacją, a poprzednią.
- Poprzednie obserwacje w szeregu czasowym nazywane są lags (opóźnieniem). Obserwacja w poprzednim kroku to lag1, obserwacja sprzed dwóch kroków czasowych to lag2, i tak dalej.

lags (opóźnienia)

- Modelowanie szeregów czasowych zakłada związek między daną obserwacją, a poprzednią.
- Poprzednie obserwacje w szeregu czasowym nazywane są lags (opóźnieniem). Obserwacja w poprzednim kroku to lag1, obserwacja sprzed dwóch kroków czasowych to lag2, i tak dalej.
- Pandas posiada specjalny typ wykresu umożliwiający zbadanie zależności między obserwacją, a opóźnieniem - lag_plot.

lags (opóźnienia)

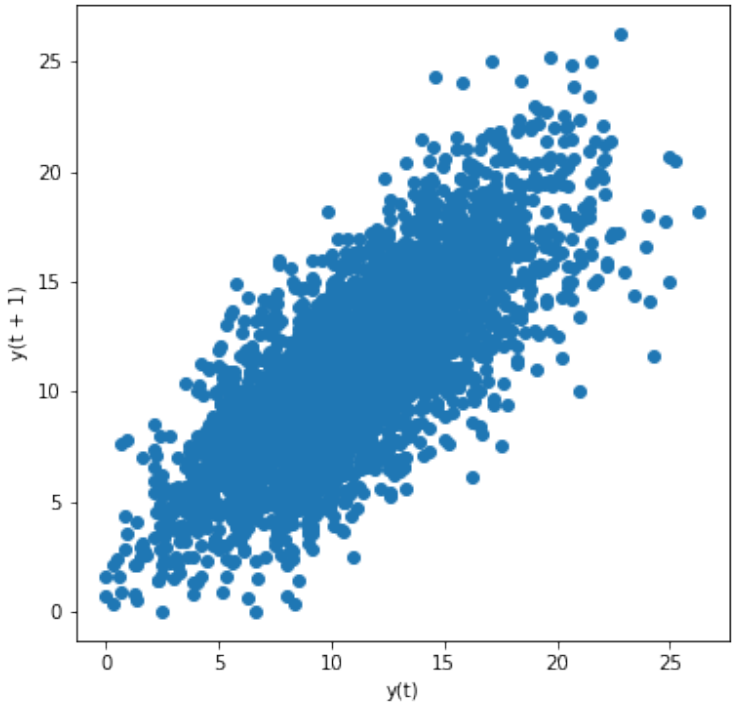
- Modelowanie szeregów czasowych zakłada związek między daną obserwacją, a poprzednią.
- Poprzednie obserwacje w szeregu czasowym nazywane są lags (opóźnieniem). Obserwacja w poprzednim kroku to lag1, obserwacja sprzed dwóch kroków czasowych to lag2, i tak dalej.
- Pandas posiada specjalny typ wykresu umożliwiający zbadanie zależności między obserwacją, a opóźnieniem - lag_plot.
- Sporządza ona wykres obserwacji w czasie t na osi x i obserwacji lag1 $(t-1)$ na osi y.

lags (opóźnienia)

- Modelowanie szeregów czasowych zakłada związek między daną obserwacją, a poprzednią.
- Poprzednie obserwacje w szeregu czasowym nazywane są lags (opóźnieniem). Obserwacja w poprzednim kroku to lag1, obserwacja sprzed dwóch kroków czasowych to lag2, i tak dalej.
- Pandas posiada specjalny typ wykresu umożliwiający zbadanie zależności między obserwacją, a opóźnieniem - lag_plot.
- Sporządza ona wykres obserwacji w czasie t na osi x i obserwacji lag1 $(t-1)$ na osi y.
  - Jeśli punkty skupiają się wzdłuż linii przekątnej od lewego dolnego rogu do górnego prawego rogu wykresu, sugeruje ona dodatnią korelację.
  - Jeśli punkty skupiają się wzdłuż przekątnej od górnego lewego do prawego dolnego to sugeruje to ujemną korelację. Im bardziej obserwacje przylegają do przekątnej tym silniejsza korelacja, a im bardziej są one rozproszone tym słabsza.

lags (opóźnienia)

https://github.com/przem85/bootcamp/blob/master/statistics/D15_Z02.ipynb



Autokorelacja

- Możemy ocenić siłę i rodzaj zależności pomiędzy obserwacjami, a ich lags (opóźnieniami).

Autokorelacja

- Możemy ocenić siłę i rodzaj zależności pomiędzy obserwacjami, a ich lags (opóźnieniami).
- W statystyce nazywa się to korelacją, a kiedy obliczane są wartości opóźnień w szeregach czasowych, nazywa się autokorelacją. Wartość korelacji obliczona między dwiema grupami liczb, takimi jak obserwacje i ich wartościami lag1, daje liczbę między -1 i 1.

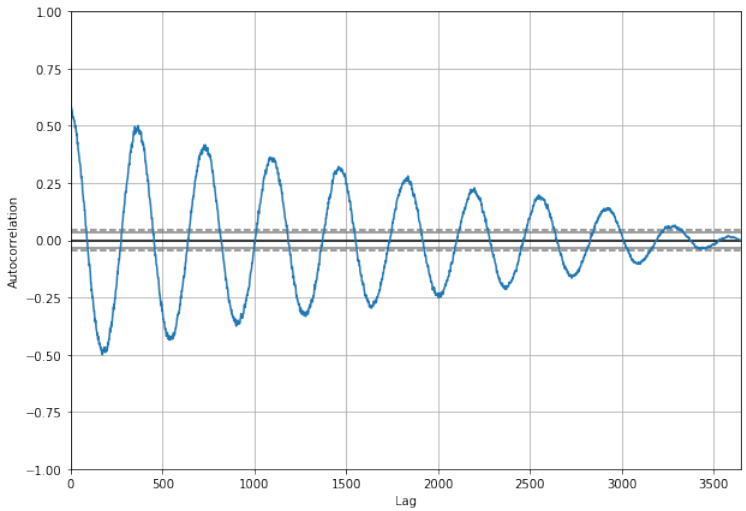
Autokorelacja

- Możemy ocenić siłę i rodzaj zależności pomiędzy obserwacjami, a ich lags (opóźnieniami).
- W statystyce nazywa się to korelacją, a kiedy obliczane są wartości opóźnień w szeregach czasowych, nazywa się autokorelacją. Wartość korelacji obliczona między dwiema grupami liczb, takimi jak obserwacje i ich wartościami lag1, daje liczbę między -1 i 1.
- Znak liczby wskazuje odpowiednio ujemną lub dodatnią korelację. Wartość bliska zeru sugeruje słabą korelację, podczas gdy wartość bliżej -1 lub 1 wskazuje na silną korelację.

Autokorelacja

- Możemy ocenić siłę i rodzaj zależności pomiędzy obserwacjami, a ich lags (opóźnieniami).
- W statystyce nazywa się to korelacją, a kiedy obliczane są wartości opóźnień w szeregach czasowych, nazywa się autokorelacją. Wartość korelacji obliczona między dwiema grupami liczb, takimi jak obserwacje i ich wartościami lag1, daje liczbę między -1 i 1.
- Znak liczby wskazuje odpowiednio ujemną lub dodatnią korelację. Wartość bliska zeru sugeruje słabą korelację, podczas gdy wartość bliżej -1 lub 1 wskazuje na silną korelację.
- Wartości korelacji (zwane współczynnikami korelacji) można obliczyć dla każdej obserwacji i różnych wartości opóźnienia. Po obliczeniu można utworzyć wykres, aby lepiej zrozumieć, jak relacja ta zmienia się wraz ze wzrostem opóźnienia.

Autokorelacja



Otrzymany wykres przedstawia opóźnienie wzdłuż osi x i korelację na osi y. Przerywane linie wskazują obszar krytyczny, powyżej tego obszaru korelacje są statystycznie znaczące.
Widzimy, że dla naszego zestawu danych mamy cykle silnej ujemnej i pozytywnej korelacji. Ujemne oznaczają związek obserwacji w przeciwnych porach roku. Fale sinusoidalne, jak widać w tym przykładzie, są mocnym znakiem sezonowości w zestawie danych.

Stacjonarność szeregu czasowego

Stwierdzono, że szereg czasowy jest stacjonarny, jeśli jego właściwości statystyczne, takie jak średnia, wariancja pozostają niezmienne w czasie.

Stacjonarność szeregu czasowego

Stwierdzono, że szereg czasowy jest stacjonarny, jeśli jego właściwości statystyczne, takie jak średnia, wariancja pozostają niezmienne w czasie.

Ale dlaczego jest to ważne?

Stacjonarność szeregu czasowego

Stwierdzono, że szereg czasowy jest stacjonarny, jeśli jego właściwości statystyczne, takie jak średnia, wariancja pozostają niezmienne w czasie.

Ale dlaczego jest to ważne?

- Większość modeli szeregów czasowych zakłada, że szereg czasowy jest stacjonarny.
- Intuicyjnie możemy to rozumieć, że jeśli szereg czasowy ma stałe zachowanie z upływem czasu, to istnieje bardzo wysokie prawdopodobieństwo, że będzie to następowało w przyszłości.

Stacjonarność szeregu czasowego

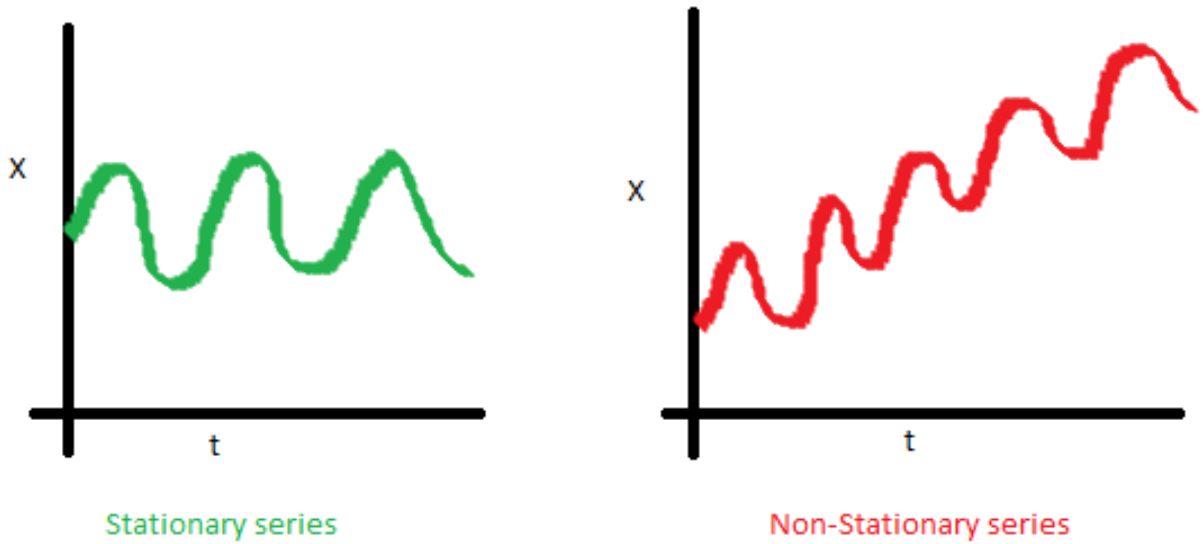
Stacjonarność jest określona przy użyciu bardzo skomplikowanych kryteriów. Jednak w celach praktycznych można założyć, że szereg czasowy jest stacjonarny, jeśli ma stałe właściwości statystyczne w czasie, tj. następujące:

- stała średnia
- stała wariancja
- autokowariancja nie zależy od czasu.

Stacjonarność szeregu czasowego

Istnieją trzy podstawowe kryteria dla szeregów czasowych, które pomagają zweryfikować stacjonarność:

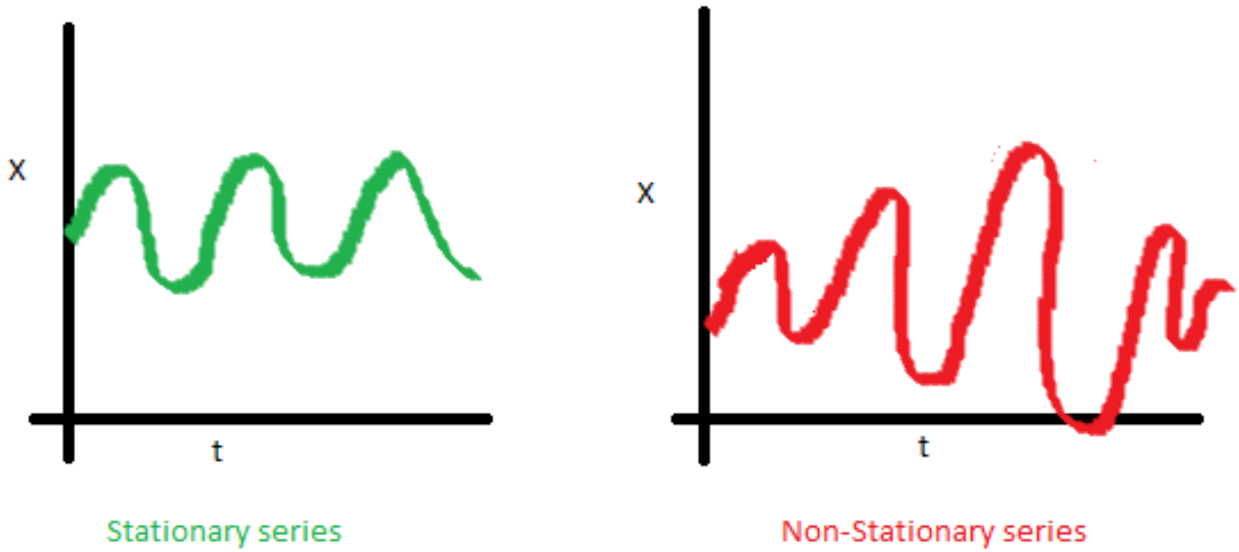
- Średnia w szeregu czasowym nie powinna być funkcją czasu, raczej powinna być stała.



Stacjonarność szeregu czasowego

Istnieją trzy podstawowe kryteria dla szeregów czasowych, które pomagają zweryfikować stacjonarność:

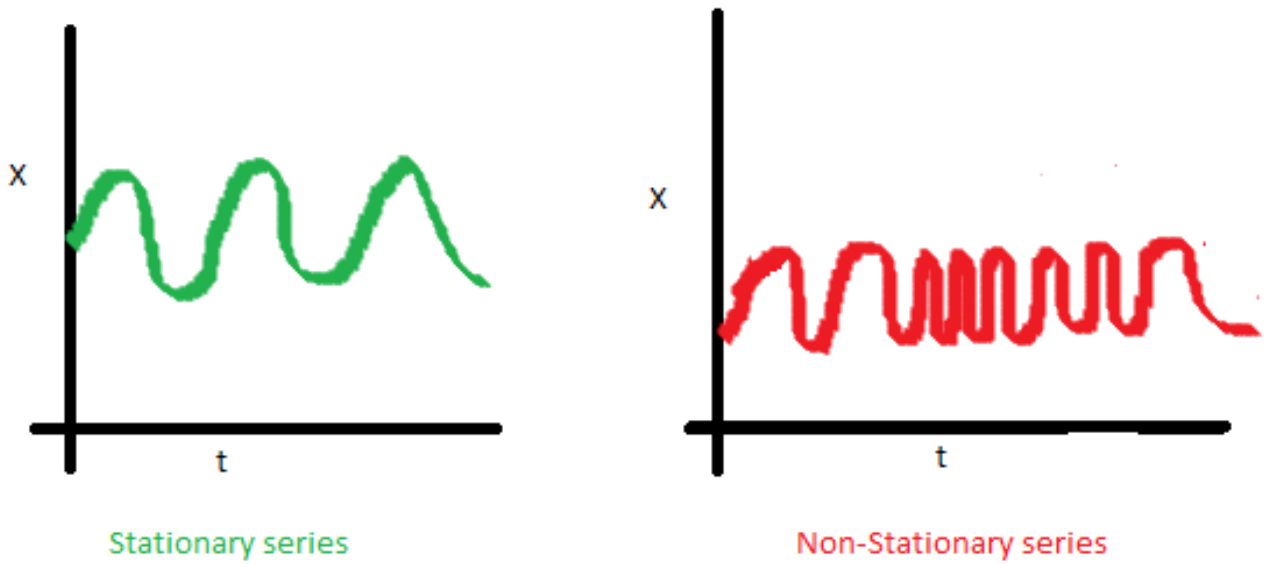
- Wariancja w szeregu czasowym nie powinna być funkcją czasu. Ta właściwość jest znana jako homoscedasticity.



Stacjonarność szeregu czasowego

Istnieją trzy podstawowe kryteria dla szeregów czasowych, które pomagają zweryfikować stacjonarność:

- Korelacja nie powinna być funkcją czasu. Na poniższym wykresie zauważysz, że rozrzut staje się mniejszy wraz ze wzrostem czasu. Stąd też korelacja nie jest stała dla czerwonego wykresu.



Random Walk

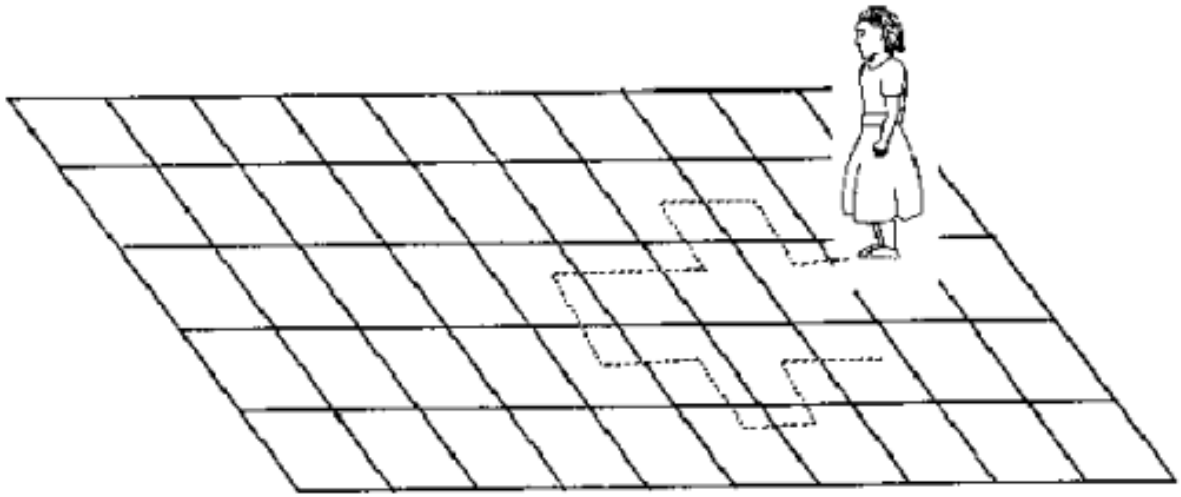
Jest to najbardziej podstawowa koncepcja szeregów czasowych. Możesz znać tę koncepcję dobrze. Ale niektórzy myślą, że błądzenie losowe jest procesem stacjonarnym.

Random Walk

Jest to najbardziej podstawowa koncepcja szeregów czasowych. Możesz znać tę koncepcję dobrze. Ale niektórzy myślą, że błądzenie losowe jest procesem stacjonarnym.

Przykład

Wyobraź sobie dziewczynę poruszającą się losowo na olbrzymiej szachownicy. W tym przypadku kolejna pozycja dziewczyny zależy tylko od ostatniej pozycji.



Random Walk

- Teraz wyobraź sobie, że siedzisz w innym pokoju i nie możesz zobaczyć dziewczyny.

Random Walk

- Teraz wyobraź sobie, że siedzisz w innym pokoju i nie możesz zobaczyć dziewczyny.
- Chcesz przewidzieć pozycję dziewczyny w kolejnych krokach.

Random Walk

- Teraz wyobraź sobie, że siedzisz w innym pokoju i nie możesz zobaczyć dziewczyny.
- Chcesz przewidzieć pozycję dziewczyny w kolejnych krokach.
- Jak dokładnie jesteś w stanie to zrobić?

Random Walk

- Teraz wyobraź sobie, że siedzisz w innym pokoju i nie możesz zobaczyć dziewczyny.
- Chcesz przewidzieć pozycję dziewczyny w kolejnych krokach.
- Jak dokładnie jesteś w stanie to zrobić?
- Oczywiście nasze przewidywania staną się coraz bardziej niedokładne, w każdym kolejnym kroku.

Random Walk

- Teraz wyobraź sobie, że siedzisz w innym pokoju i nie możesz zobaczyć dziewczyny.
- Chcesz przewidzieć pozycję dziewczyny w kolejnych krokach.
- Jak dokładnie jesteś w stanie to zrobić?
- Oczywiście nasze przewidywania staną się coraz bardziej niedokładne, w każdym kolejnym kroku.
- W $t = 0$ dokładnie wiemy, gdzie jest dziewczyna. W następnym kroku może tylko przejść do 8 kwadratów, a zatem prawdopodobieństwo spadnie z 1 do 1/8 itd.

Random Walk

Teraz spróbujmy sformułować tę sytuację w postaci szeregu czasowego:

$$X(t) = X(t-1) + Er(t),$$

gdzie $Er(t)$ jest błędem w punkcie czasowym t. Jest to przypadkowa decyzja dziewczyny o przejściu na następne pole w każdym punkcie czasu t.

Random Walk

Teraz spróbujmy sformułować tę sytuację w postaci szeregu czasowego:

$$X(t) = X(t-1) + Er(t),$$

gdzie $Er(t)$ jest błędem w punkcie czasowym t. Jest to przypadkowa decyzja dziewczyny o przejściu na następne pole w każdym punkcie czasu t.

Teraz, możemy zapisać sytuację w chwili t za pomocą wartości w chwili 0 i kolejnych kroków:

$$X(t) = X(0) + \mathrm{Sum}(Er(1), Er(2), Er(3), \ldots, E(t))$$

`https://github.com/przem85/bootcamp/blob/master/statistics/D15_Z03.ipynb`

Random Walk

Czy średnia jest stała?

Random Walk

Czy średnia jest stała?

$$E(X(t)) = E(X(0)) + \mathrm{Sum}(E(Er(1)), E(Er(2)), E(Er(3)), \ldots, E(E(t)))$$

Random Walk

Czy średnia jest stała?

$$E(X(t)) = E(X(0)) + \mathrm{Sum}(E(Er(1)), E(Er(2)), E(Er(3)), \ldots, E(E(t)))$$

Wiemy, że wartość oczekiwana błędu jest równa zero, ponieważ jest to błądzenie losowe.
Ponadto mamy:

$$E[X(t)] = E[X(0)] = Constant$$

Random Walk

Czy wariancja jest niezmienna?

$$Var(X(t)) = Var(X(0)) + \mathrm{Sum}(Var(Er(1)), Var(Er(2)), \ldots, Var(E(t)))$$

Random Walk

Czy wariancja jest niezmienna?

$$Var(X(t)) = Var(X(0)) + \mathrm{Sum}(Var(Er(1)), Var(Er(2)), \ldots, Var(E(t)))$$

Ponieważ $Var(X(0)) = 0$ oraz $Var(Er(i)) = Var(error)$

$$Var(X(t)) = t \cdot Var(error).$$

Random Walk

Czy wariancja jest niezmienna?

$$Var(X(t)) = Var(X(0)) + \mathrm{Sum}(Var(Er(1)), Var(Er(2)), \ldots, Var(E(t)))$$

Ponieważ $Var(X(0)) = 0$ oraz $Var(Er(i)) = Var(error)$

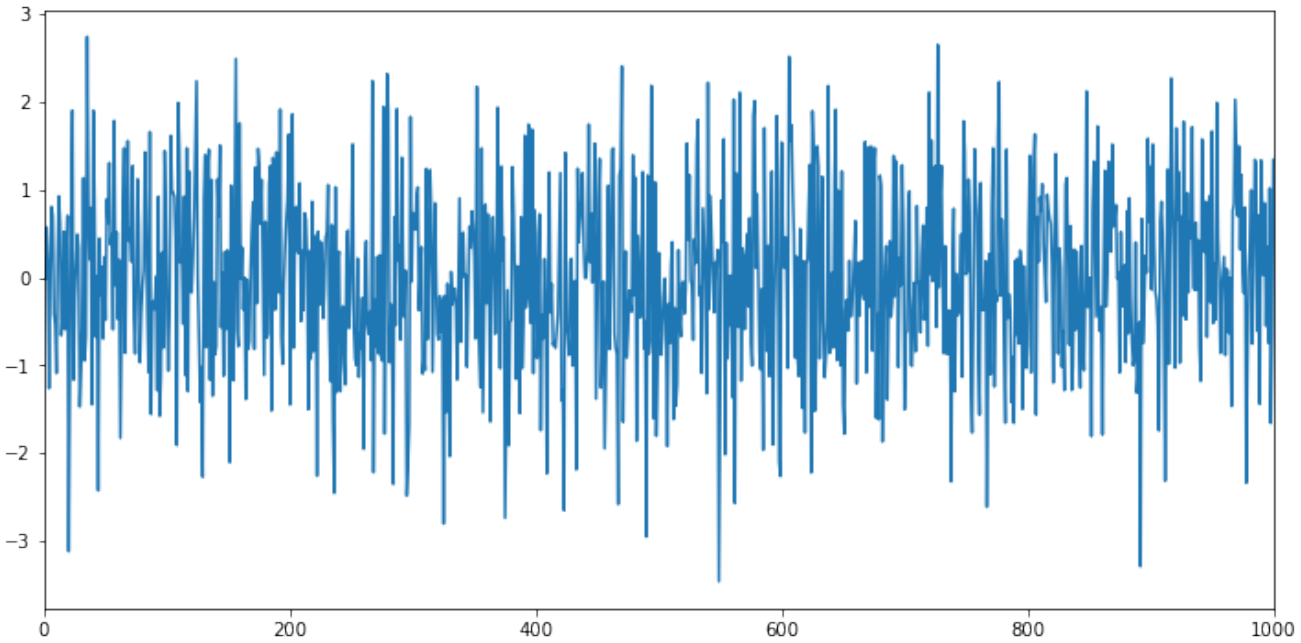
$$Var(X(t)) = t \cdot Var(error).$$

Stąd wniosek, że błądzenie losowe nie jest procesem stacjonarnym, ponieważ wariancja jest zmienna w czasie.

Współczynnik Rho

Teraz wypróbujmy różne wartości $\rho = 0$:

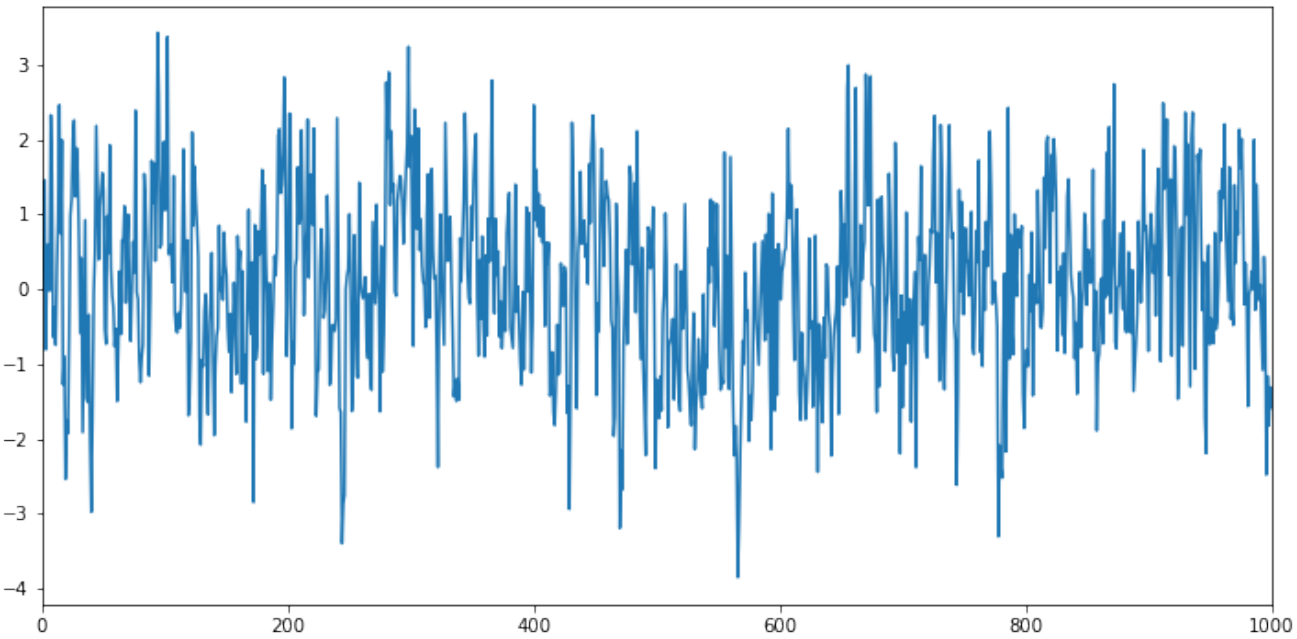
$$X(t) = \rho \cdot X(t-1) + Er(t)$$



Współczynnik Rho

Teraz wypróbujmy różne wartości $\rho = 0.5$:

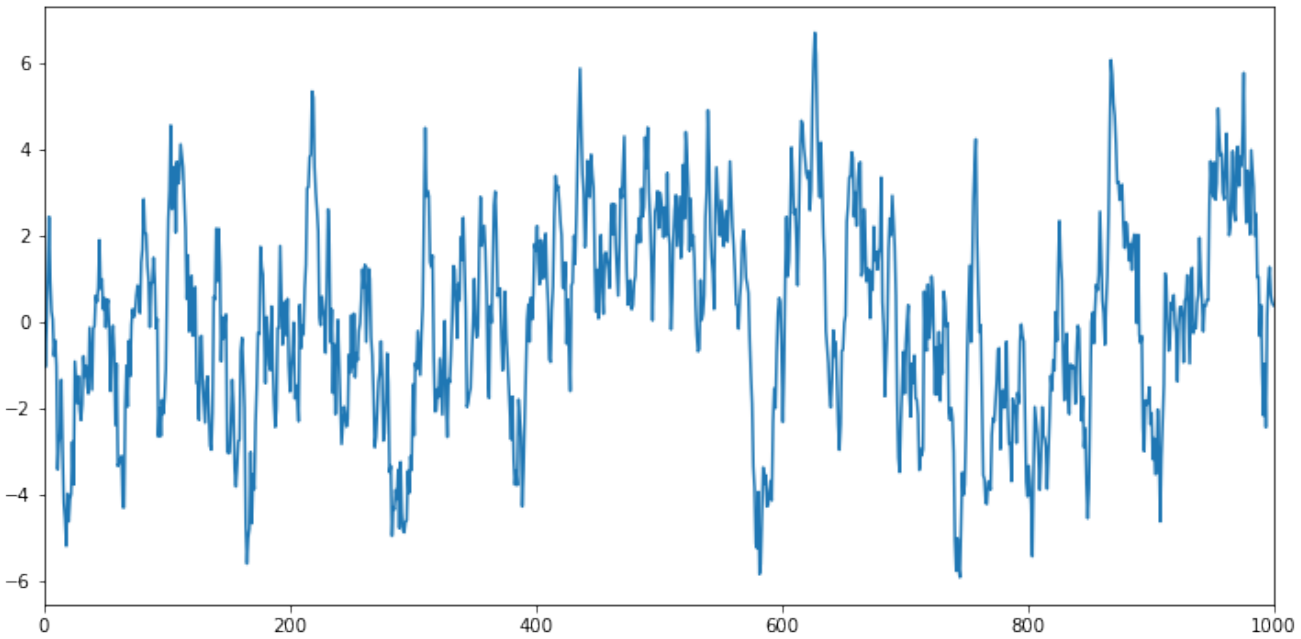
$$X(t) = \rho \cdot X(t-1) + Er(t)$$



Współczynnik Rho

Teraz wypróbujmy różne wartości $\rho = 0.9$:

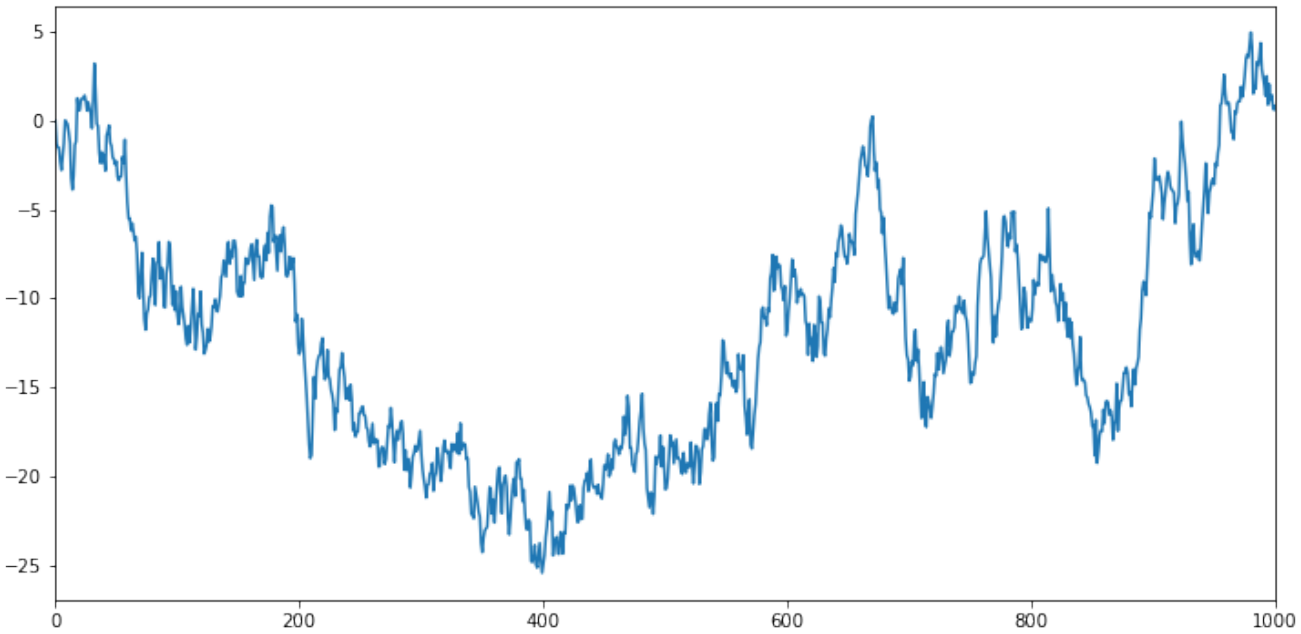
$$X(t) = \rho \cdot X(t-1) + Er(t)$$



Współczynnik Rho

Teraz wypróbujmy różne wartości $\rho = 1$:

$$X(t) = \rho \cdot X(t-1) + Er(t)$$



Random Walk

Jak widzimy w zależności od parametru ρ dostajemy albo szereg stacjonarny albo nie. Dla $\rho = 1$ mamy ewidentnie szereg niestacjonarny. W naszym przypadku mamy:

$$E[X(t)] = \rho * E[X(t-1)]$$

Jak widzimy, jeżeli ρ jest małe, to powoduje zmniejszanie się wartości oczekiwanej (dla $\rho = 0$ wartość oczekiwana jest zawsze zero). Gdy ρ jest równe 1, nie mamy efektu ściągania (w kolejnych krokach) $E[X(t)]$ do zera.

Dickey Fuller Test

Nasze rozumowanie prowadzi do Testu Dickey-Fullera. Aby otrzymać dokładną postać wystarczy przekształcić:

$$X(t) = \rho \cdot X(t-1) + Er(t)$$

do postaci:

$$X(t) - X(t-1) = (\rho - 1) \cdot X(t-1) + Er(t)$$

Dickey Fuller Test

Nasze rozumowanie prowadzi do Testu Dickey-Fullera. Aby otrzymać dokładną postać wystarczy przekształcić:

$$X(t) = \rho \cdot X(t-1) + Er(t)$$

do postaci:

$$X(t) - X(t-1) = (\rho - 1) \cdot X(t-1) + Er(t)$$

Aby wykonać test musimy sprawdzić, czy $\rho - 1$ jest znacznie różny od zera, czy nie.

Dickey Fuller Test

Nasze rozumowanie prowadzi do Testu Dickey-Fullera. Aby otrzymać dokładną postać wystarczy przekształcić:

$$X(t) = \rho \cdot X(t-1) + Er(t)$$

do postaci:

$$X(t) - X(t-1) = (\rho - 1) \cdot X(t-1) + Er(t)$$

Aby wykonać test musimy sprawdzić, czy $\rho - 1$ jest znacznie różny od zera, czy nie.
Jeśli hipoteza zerowa zostanie odrzucona, otrzymamy stacjonarny szereg czasowy.

Dickey Fuller Test

- Test Dickey-Fuller: jest to jeden ze statystycznych testów sprawdzających stacjonarność.
- Tutaj hipotezą zerową jest, że szereg czasowy jest niestacjonarny.
- Wyniki testu obejmują statystykę testu oraz niektóre wartości krytyczne dla różnych poziomów ufności.
- Jeśli statystyka testowa jest mniejsza niż wartość krytyczna, możemy odrzucić hipotezę zerową i powiedzieć, że szereg czasowy jest stacjonarny.

`https://github.com/przem85/bootcamp/blob/master/statistics/D15_Z04.ipynb`

Co zrobić gdy szereg jest niestacjonarny?

Stacjonarność

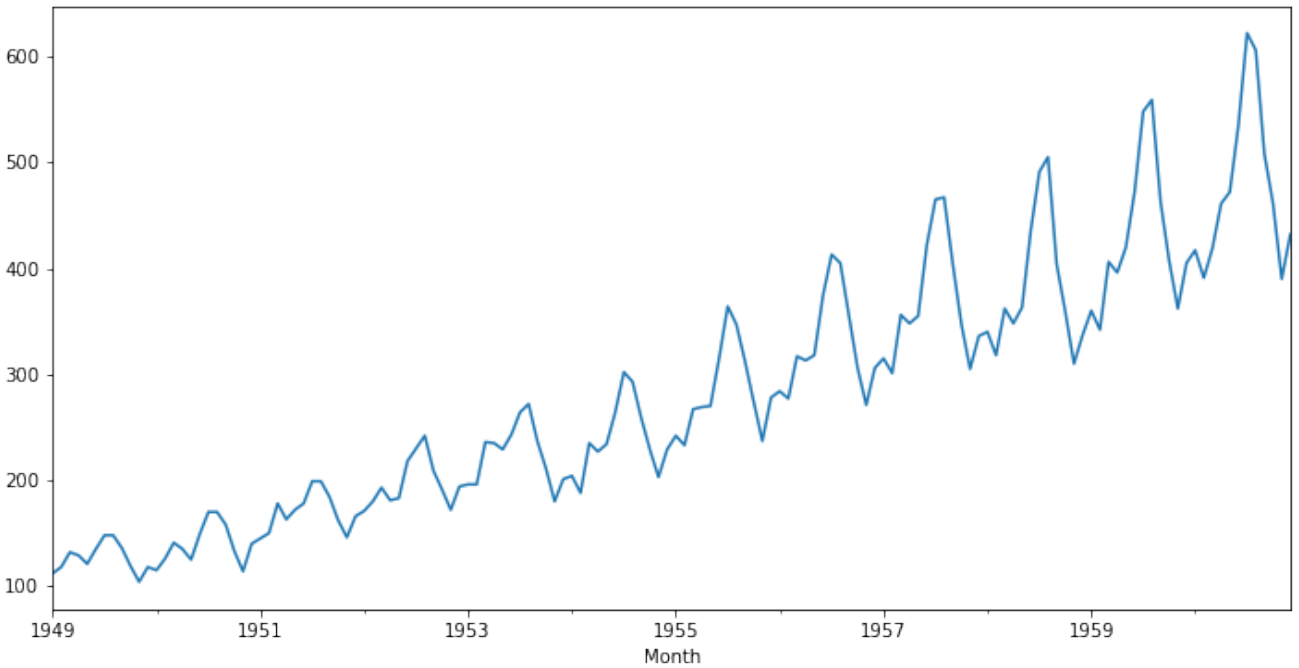
Stacjonarność jest określona przy użyciu bardzo skomplikowanych kryteriów. Jednak w celach praktycznych można założyć, że szereg czasowy jest stacjonarny, jeśli ma stałe właściwości statystyczne w czasie, tj. następujące:

- stała średnia,
- stała wariancja,
- autokowariancja nie zależy od czasu.

Stacjonarność

Rozważmy dane opisujące ilość pasażerów latających liniami lotniczymi:

`https://github.com/przem85/bootcamp/blob/master/statistics/D15_Z05.ipynb`

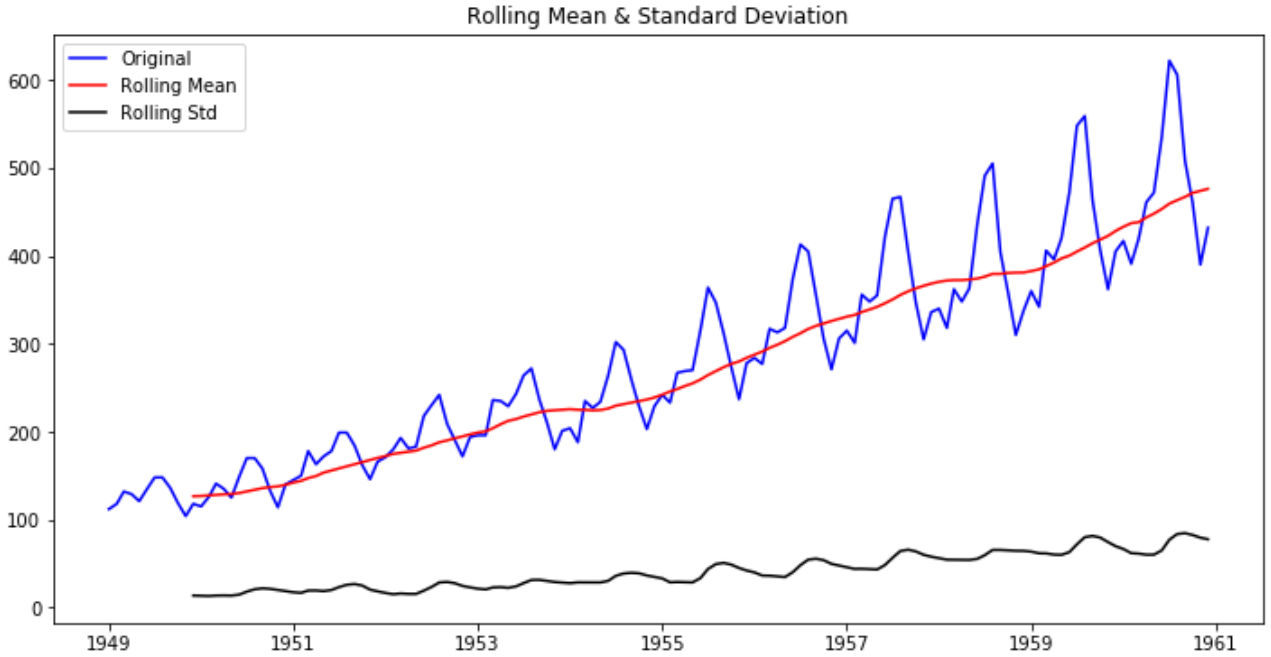


Stacjonarność

Jest oczywiste, że istnieje ogólna tendencja wzrostowa w danych wraz z sezonowymi wahaniami. Jednak nie zawsze możliwe jest takie wizualne wnioskowanie (zobaczymy takie przypadki później). Więc bardziej formalnie możemy sprawdzić stacjonarność używając:

- możemy wyznaczyć średnią ruchomą lub ruchomą wariancji i sprawdzić, czy zmienia się ona z czasem. Średnia ruchoma lub ruchoma wariancja oznacza, że w każdej chwili t przeanalizujemy średnią/wariancję z ostatniego roku, tzn. w ciągu ostatnich 12 miesięcy. Ale znowu jest to bardziej wizualna technika.
- Test Dickey-Fuller

Stacjonarność




```
Results of Dickey-Fuller Test:
Test Statistic                  0.815369
p-value                         0.991880
#Lags Used                     13.000000
Number of Observations Used    130.000000
Critical Value (1%)             -3.481682
Critical Value (5%)             -2.884042
Critical Value (10%)            -2.578770
dtype: float64
```

Stacjonarność

- Chociaż założenie stacjonarności jest przyjmowane w wielu modelach, w praktyce żaden szereg czasowy nie jest stacjonarny.
- Mamy sposoby, aby szereg stał się stacjonarny.

Uwaga

Należy pamiętać, że jest to prawie niemożliwe, ale staramy się doprowadzić do sytuacji, w której jest on jak najbliżej bycia stacjonarnym.

Stacjonarność

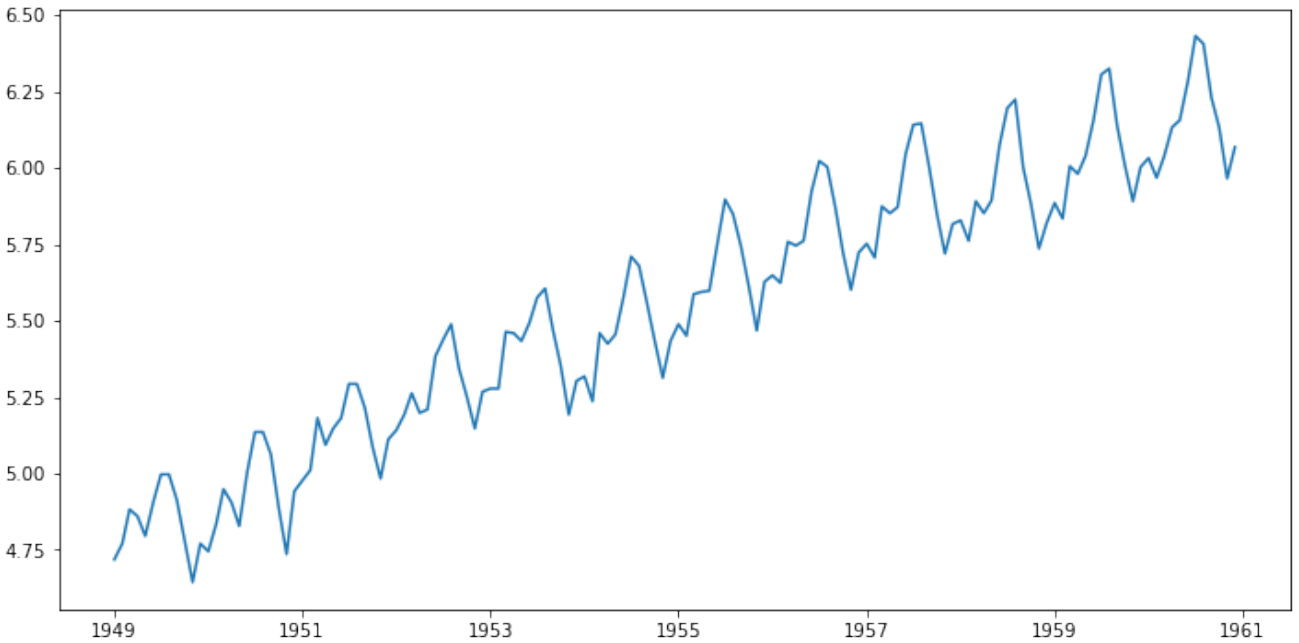
Istnieją dwa główne powody, które powodują niestacjonarność:

- Trend - zmienna średnia w czasie. Na przykład, w naszym przypadku zauważyliśmy, że przeciętnie liczba pasażerów rośnie z upływem czasu.
- Sezonowość - zmiany w określonych przedziałach czasowych. Np. ludzie mogą mieć skłonność do zakupu samochodów w danym okresie.

Podstawową zasadą jest modelowanie lub szacowanie trendu i sezonowości w szeregu czasowym oraz usunięcie ich, aby uzyskać stacjonarne szeregi czasowe. Następnie dla takiego szeregu czasowego można zastosować modele statystyczne. Ostatnim krokiem byłoby przekształcenie prognozowanych wartości do pierwotnej skali poprzez dodanie trendu i sezonowości.

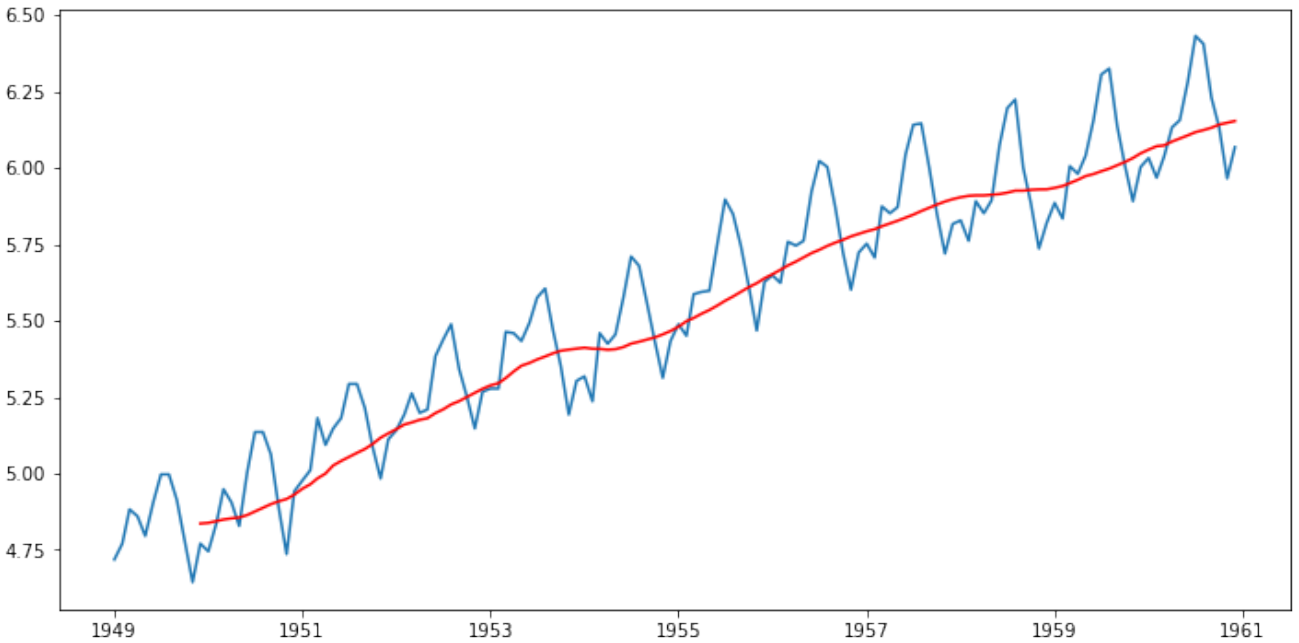
Stacjonarność

Czasami pomaga wzięcie logarytmu z danych.



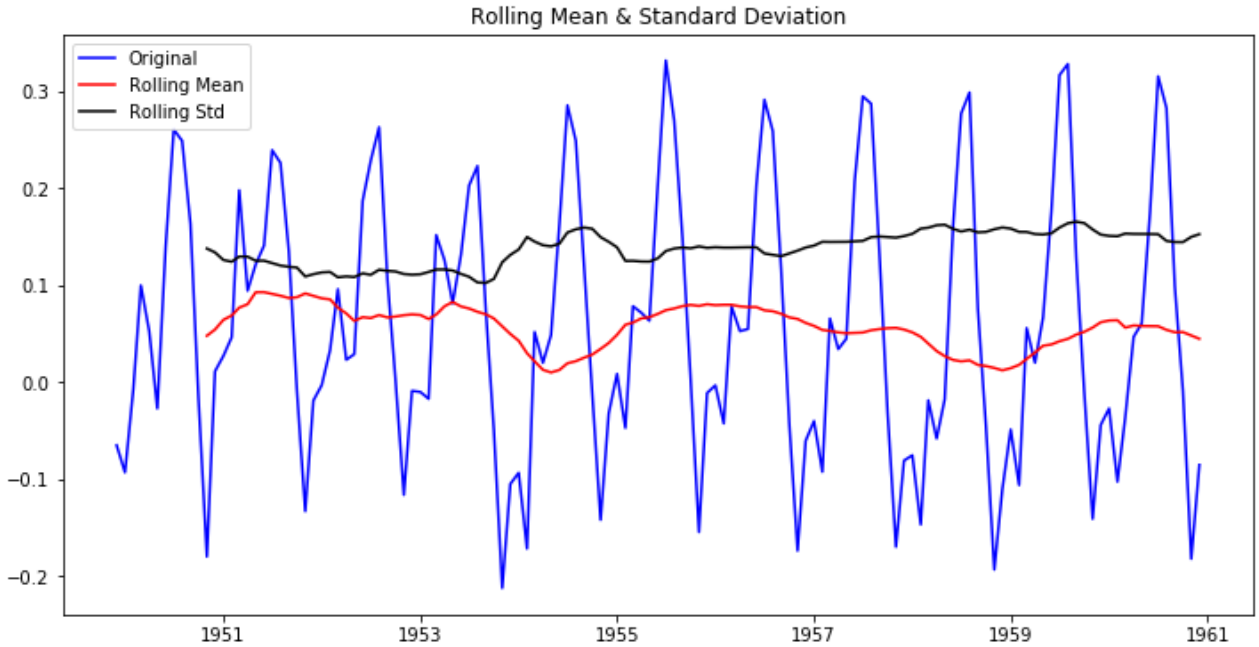
Stacjonarność

Odejmijmy średnią kroczącą.



Stacjonarność

Odejmijmy średnią kroczącą.



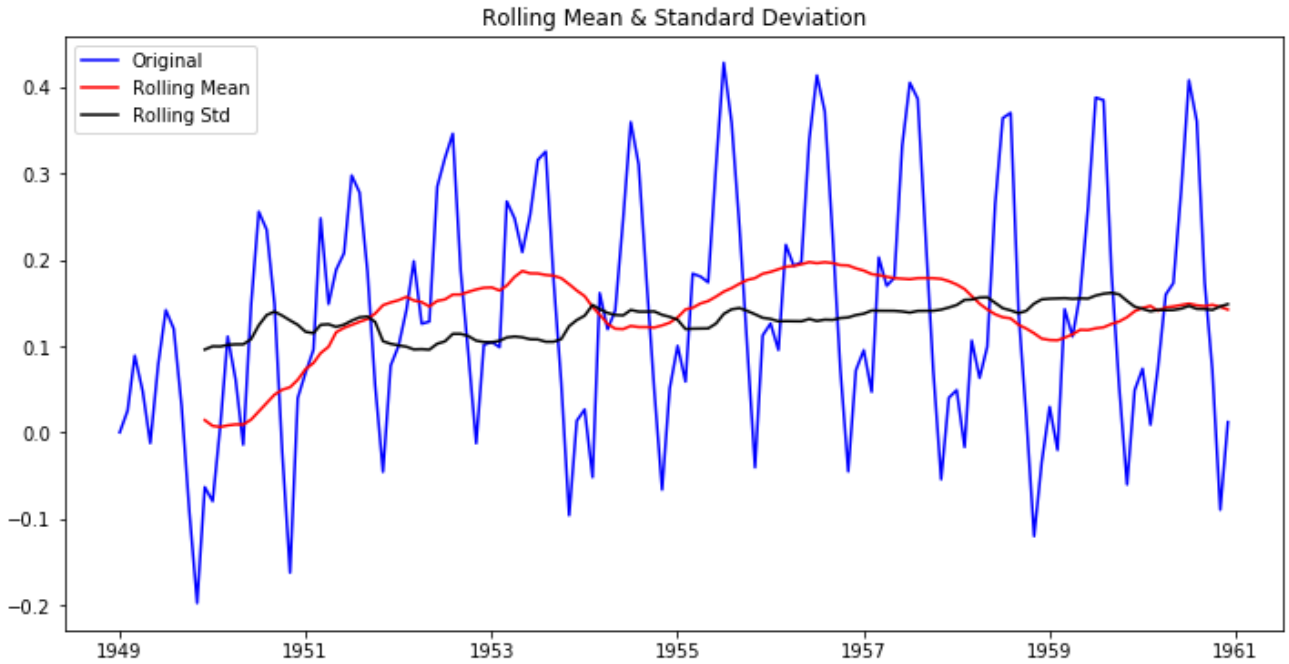

```
Results of Dickey-Fuller Test:
Test Statistic                  -3.162908
p-value                          0.022235
#Lags Used                      13.000000
Number of Observations Used    119.000000
Critical Value (1%)             -3.486535
Critical Value (5%)             -2.886151
Critical Value (10%)            -2.579896
dtype: float64
```

Stacjonarność

- Jednakże wadą w tym szczególnym podejściu jest to, że okres czasu musi być ściśle określony. W tym przypadku możemy przyjąć średnie roczne, ale w skomplikowanych sytuacjach, takich jak prognozowanie cen akcji, trudno jest dobrać tą stałą.
- Przyjmujemy zatem "ważoną średnią ruchliwą", gdzie bliższe wartości mają wyższe wagi. Istnieje wiele technik przypisywania wag. My użyjemy średniej harmonicznej.

`http://pandas.pydata.org/pandas-docs/stable/computation.html#exponentially-weighted-moment-functions`

Stacjonarność – modelowanie trendu i sezonowości




```
Results of Dickey-Fuller Test:
Test Statistic                  -3.601262
p-value                          0.005737
#Lags Used                      13.000000
Number of Observations Used    130.000000
Critical Value (1%)             -3.481682
Critical Value (5%)             -2.884042
Critical Value (10%)            -2.578770
dtype: float64
```

Stacjonarność – modelowanie trendu i sezonowości

Proste techniki redukcji trendu omówione wcześniej nie działają we wszystkich przypadkach, szczególnie tych o wysokiej sezonowości. Teraz omówimy dwa sposoby usuwania tendencji i sezonowości:

- Różnicowanie - za pomocą różnych różnic czasowych,
- Rozkład - modelowanie zarówno tendencji, jak i sezonowości, i usuwanie ich z modelu.

Stacjonarność - differencing

Jedną z najczęstszych metod radzenia sobie z trendem i sezonowością jest różnicowanie. W tej technice w każdej chwili czasowej rozważamy różnicę obserwacji z tą z poprzedniej chwili. Działa to głównie poprawiając stacjonarność.

Dlaczego różnicujemy szeregi czasowych? Różnicowanie to metoda przekształcania szeregów czasowych. Może być użyta do usunięcia zależności czasowych, takich jak trendy i sezonowość. Różnicowanie odbywa się przez odjęcie poprzedniej obserwacji od obecnej obserwacji.

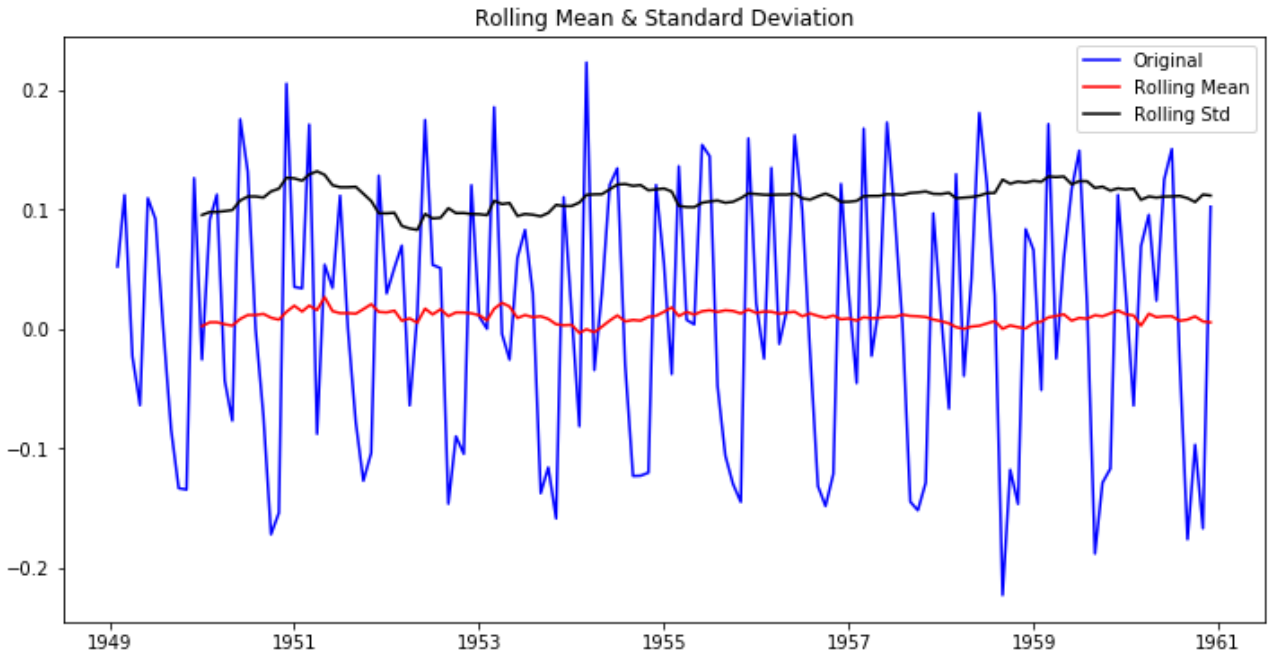
$$\text{Różnica (t)} = \text{obserwacja (t)} - \text{obserwacja (t-1)}$$

`https://github.com/przem85/bootcamp/blob/master/statistics/D15_Z06.ipynb`

Stacjonarność - differencing

Wróćmy do naszych danych opisujących ilość ludzi latających samolotami.

`https://github.com/przem85/bootcamp/blob/master/statistics/D15_Z07.ipynb`




```
Results of Dickey-Fuller Test:
Test Statistic                  -2.717131
p-value                          0.071121
#Lags Used                      14.000000
Number of Observations Used    128.000000
Critical Value (1%)             -3.482501
Critical Value (5%)             -2.884398
Critical Value (10%)            -2.578960
dtype: float64
```

Stacjonarność

Dekompozycja szeregów czasowych opiera się na myśleniu o szeregu czasowym jako kombinacji:

- trendu,
- sezonowości,
- szumu.

Szereg czasowy zazwyczaj składa się z części powtarzającej się i szumu:

- Level: Średnia wartość w szeregu czasowym.
- Trend: Wzrastająca lub malejąca wartość w szeregu czasowym.
- Seasonality: powtarzający się cykl krótkoterminowy w szeregu czasowym.
- Noise: losowe zaburzenia w szeregu czasowym.

Stacjonarność

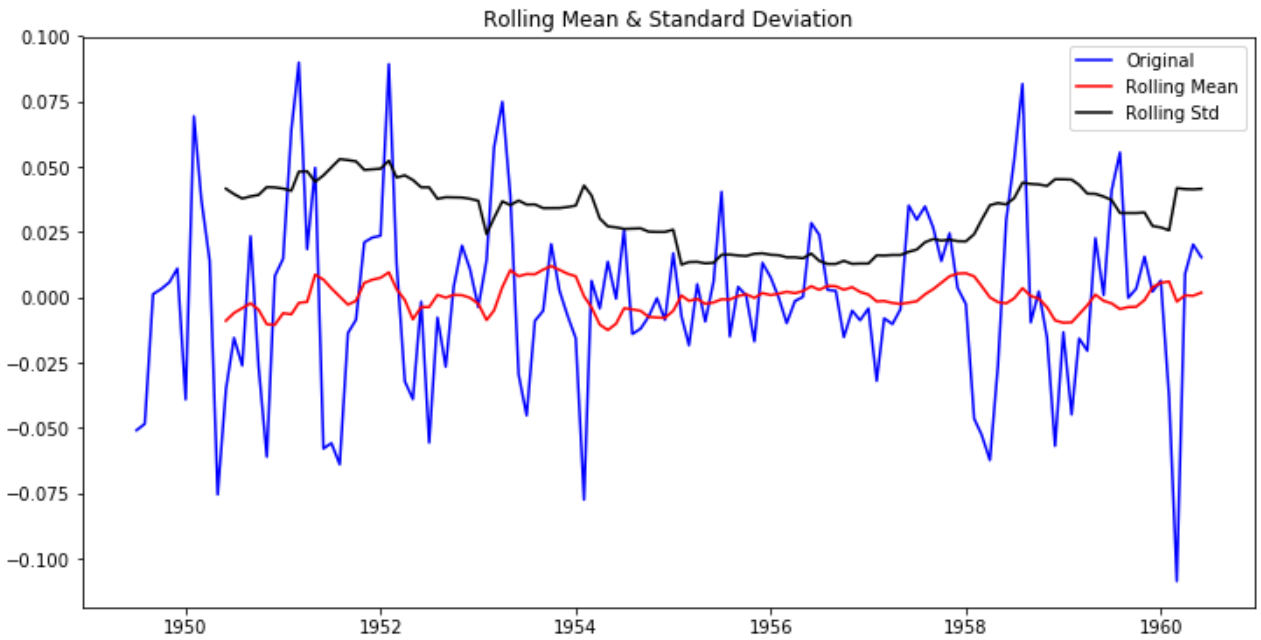
Uważa się, że szereg czasowy jest agregatem lub kombinacją tych czterech elementów. Wszystkie szeregi czasowe mają poziom (level) i szum. Elementy trendu i sezonowości są opcjonalne.

- Additive Model
  y(t) = Level + Trend + Seasonality + Noise
- Multiplicative Model
  y(t) = Level * Trend * Seasonality * Noise

`https://github.com/przem85/bootcamp/blob/master/statistics/D15_Z08.ipynb`

Stacjonarność

Wróćmy do naszych danych opisujących ilość ludzi latających samolotami.
`https://github.com/przem85/bootcamp/blob/master/statistics/D15_Z09.ipynb`




```
Results of Dickey-Fuller Test:
Test Statistic                -6.332387e+00
p-value                        2.885059e-08
#Lags Used                     9.000000e+00
Number of Observations Used    1.220000e+02
Critical Value (1%)           -3.485122e+00
Critical Value (5%)           -2.885538e+00
Critical Value (10%)          -2.579569e+00
dtype: float64
```

ARMA

Modele ARMA są powszechnie stosowane w modelowaniu szeregów czasowych. W modelu ARMA, AR oznacza auto-regression, a MA oznacza średnią ruchomą.

Pamiętaj, że modele AR i MA zakładają stacjonarności szeregów czasowych.

Przykład:

- Obecne PKB kraju $x(t)$ zależy od ubiegłorocznego PKB, tzn. $x(t-1)$.

Przykład:

- Obecne PKB kraju $x(t)$ zależy od ubiegłorocznego PKB, tzn. $x(t-1)$.
- Zakłada się, że całkowity koszt PKB w kraju w roku podatkowym jest uzależniony od wydajności zakładów produkcyjnych/usługowych w poprzednim roku oraz nowo zakładanych zakładów w bieżącym roku.

AR

Przykład:

- Obecne PKB kraju $x(t)$ zależy od ubiegłorocznego PKB, tzn. $x(t-1)$.
- Zakłada się, że całkowity koszt PKB w kraju w roku podatkowym jest uzależniony od wydajności zakładów produkcyjnych/usługowych w poprzednim roku oraz nowo zakładanych zakładów w bieżącym roku.
- Głównym składnikiem PKB jest jednak ta pierwsza część.

AR

Przykład:

- Obecne PKB kraju $x(t)$ zależy od ubiegłorocznego PKB, tzn. $x(t-1)$.
- Zakłada się, że całkowity koszt PKB w kraju w roku podatkowym jest uzależniony od wydajności zakładów produkcyjnych/usługowych w poprzednim roku oraz nowo zakładanych zakładów w bieżącym roku.
- Głównym składnikiem PKB jest jednak ta pierwsza część.

Stąd możemy napisać równanie PKB jako:

$$x(t) = \alpha * x(t-1) + error(t)$$

AR

$$x(t) = \alpha * x(t-1) + error(t)$$

- To równanie jest znane jako model AR(1).
- Cyfra (1) oznacza, że następne wystąpienie jest wyłącznie zależne od poprzedniej instancji.
- Skalar α jest współczynnikiem, którego szukamy, aby zminimalizować błąd.
- Zauważmy, że $x(t-1)$ jest rzeczywiście powiązane z $x(t-2)$ w ten sam sposób.
- Zatem wszelkie zaburzenia $x(t)$ stopniowo zanikają w przyszłości.

MA

Rozważmy inny przykład, aby zrozumieć model MA:

MA

Rozważmy inny przykład, aby zrozumieć model MA:

- Producent wytwarza określony typ torebek, który był dostępny na rynku od dłuższego czasu. Będąc konkurencyjnym rynkiem, sprzedaż torebek była zerowa przez wiele dni.

MA

Rozważmy inny przykład, aby zrozumieć model MA:

- Producent wytwarza określony typ torebek, który był dostępny na rynku od dłuższego czasu. Będąc konkurencyjnym rynkiem, sprzedaż torebek była zerowa przez wiele dni.
- Pewnego dnia producent przeprowadził eksperyment z projektem i wyprodukował inny rodzaj torebek.

MA

Rozważmy inny przykład, aby zrozumieć model MA:

- Producent wytwarza określony typ torebek, który był dostępny na rynku od dłuższego czasu. Będąc konkurencyjnym rynkiem, sprzedaż torebek była zerowa przez wiele dni.
- Pewnego dnia producent przeprowadził eksperyment z projektem i wyprodukował inny rodzaj torebek.
- Ten typ torebki nie był dostępny w żadnym miejscu na rynku. Tak więc producent był w stanie sprzedać cały zapas 1000 torebek (można nazwać to jako $x(t)$).

MA

Rozważmy inny przykład, aby zrozumieć model MA:

- Producent wytwarza określony typ torebek, który był dostępny na rynku od dłuższego czasu. Będąc konkurencyjnym rynkiem, sprzedaż torebek była zerowa przez wiele dni.
- Pewnego dnia producent przeprowadził eksperyment z projektem i wyprodukował inny rodzaj torebek.
- Ten typ torebki nie był dostępny w żadnym miejscu na rynku. Tak więc producent był w stanie sprzedać cały zapas 1000 torebek (można nazwać to jako $x(t)$).
- Zapotrzebowanie było tak wysokie, że zapasy torebek wyczerpały się. W rezultacie około 100 klientów nie mogło kupić tej torebki.

MA

Rozważmy inny przykład, aby zrozumieć model MA:

- Producent wytwarza określony typ torebek, który był dostępny na rynku od dłuższego czasu. Będąc konkurencyjnym rynkiem, sprzedaż torebek była zerowa przez wiele dni.
- Pewnego dnia producent przeprowadził eksperyment z projektem i wyprodukował inny rodzaj torebek.
- Ten typ torebki nie był dostępny w żadnym miejscu na rynku. Tak więc producent był w stanie sprzedać cały zapas 1000 torebek (można nazwać to jako $x(t)$).
- Zapotrzebowanie było tak wysokie, że zapasy torebek wyczerpały się. W rezultacie około 100 klientów nie mogło kupić tej torebki.
- Możemy tą lukę interpretować jako błąd w danym punkcie czasowym. Z czasem torba straciła swój efekt „wow". Ale pozostało kilku klientów, którzy nadal chcą kupić torbę.

MA

Taki proces można opisać za pomocą formuły:

$$x(t) = \beta * error(t-1) + error(t)$$

Uwaga

Zauważyłeś różnicę między modelem MA i AR? W modelu MA szum/szok szybko zanika wraz z upływem czasu. Model AR modeluje trwały efekt.

ARMA

- Główna różnica między modelem AR i MA oparta jest na korelacji pomiędzy obserwacjami w szeregu czasowym w różnych punktach czasowych.
- Korelacja między punktami $x(t)$, a $x(t-n)$ dla dużych n w modelu MA zawsze wynosi zero (intuicja z poprzedniego przykładu).
- Jednak korelacja $x(t)$ i $x(t-n)$ maleje znacznie wolniej w modelu AR.
- Różnice te są wykorzystywana do wyboru modelu AR lub modelu MA.

ARMA

Zanim zdecydujemy, który model ma być używany, musimy przyjrzeć się autokorelacji.

Sezonowość w szeregu czasowym można badać za pomocą correlogramów, które przedstawiają graficznie i numerycznie funkcję autokorelacji (ACF). Wykresy wbudowane w pandas oraz statsmodels standaryzują dane przed obliczaniem autokorelacji. Biblioteki te odejmują średnie i dzielą dane przez odchylenie standardowe.

Partial autocorrelations.

- Inną użyteczną metodą sprawdzania zależności w szeregach czasowych jest sprawdzenie częściowej funkcji autokorelacji (PACF).
- Jest to rozszerzenie autokorelacji, w którym usunięto zależność od elementów pośrednich.
- Biorąc pod uwagę szereg czasowy z_t, Partial autocorrelations (autokorelacja częściowa), jest autokorelacją między z_t, a z_{t+k} z usuniętą liniową zależnością między z_t, a $z_{t+1}, \ldots, z_{t+k-1}$.

ARMA

Gdy mamy stacjonarny szereg czasowy, musimy odpowiedzieć na dwa podstawowe pytania:

- Czy jest to proces AR lub MA?
- Jaki wybrać parametr dla naszego procesu AR lub MA?

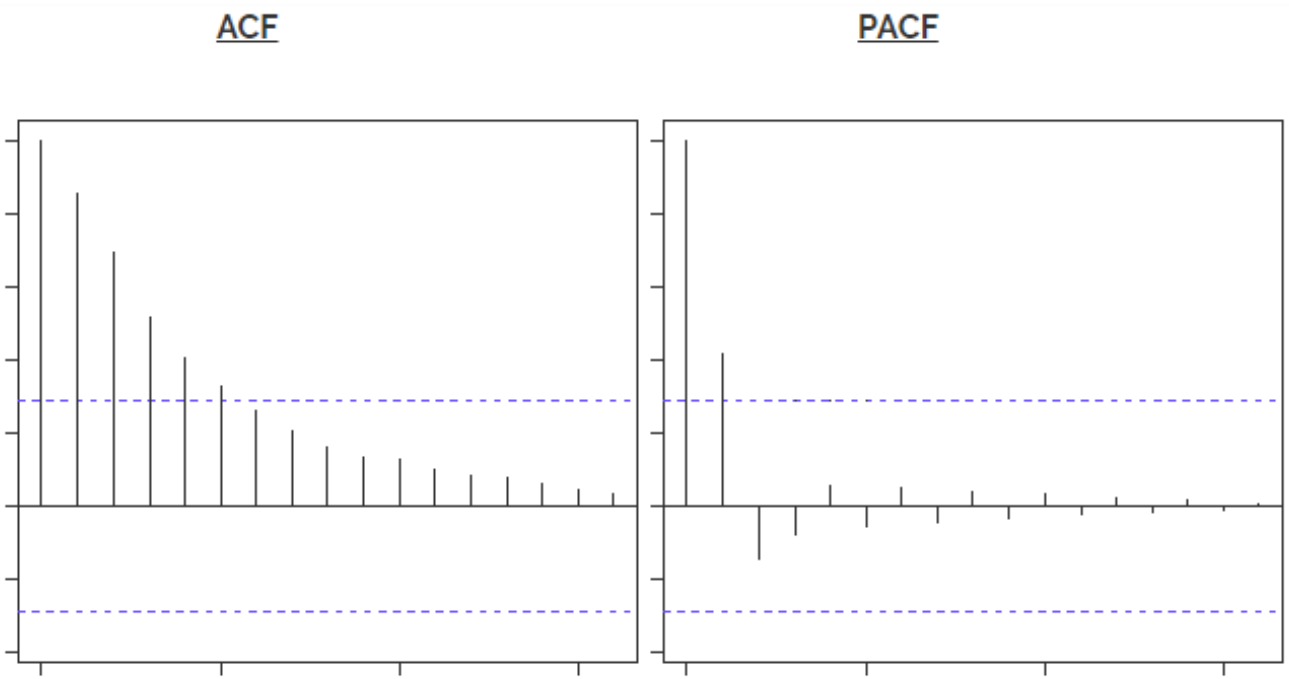
ARMA

Na pierwsze pytanie można odpowiedzieć przy użyciu Auto-correlation function (ACF).

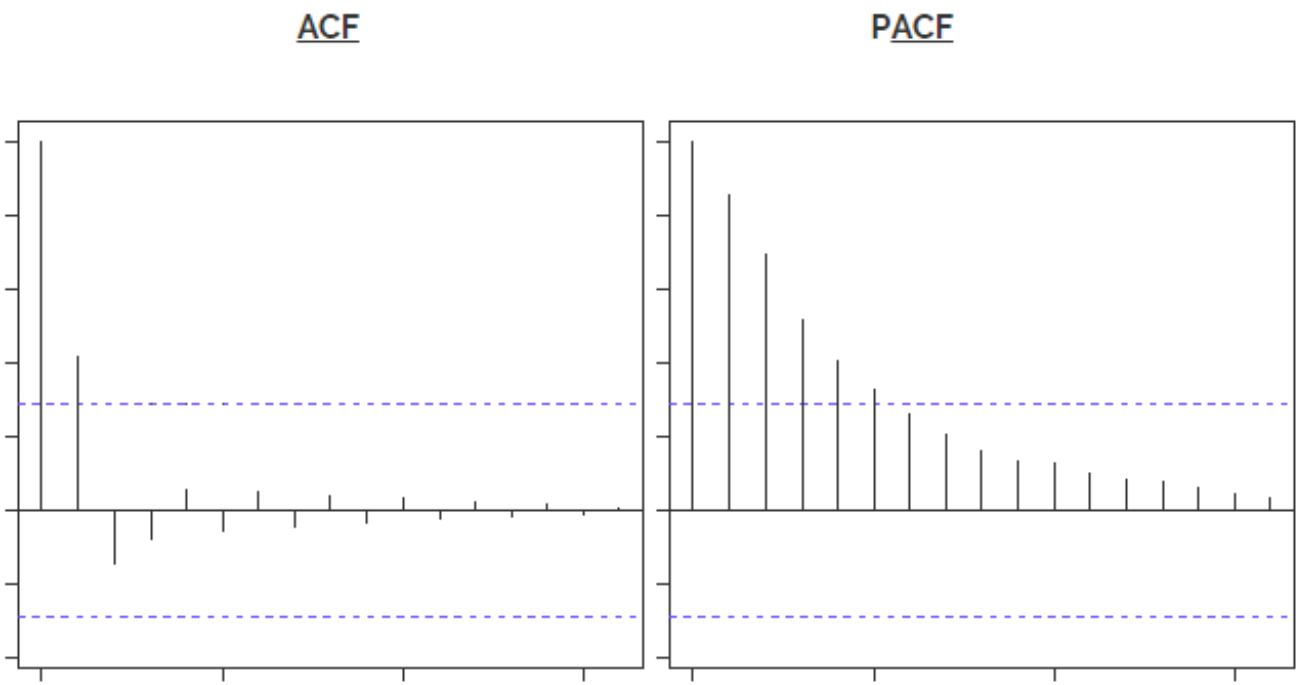
- ACF jest wykresem łącznej korelacji między różnymi opóźnieniami. Interesuje nas korelacja $x(t)$ z $x(t-1)$, $x(t-2)$ i tak dalej.
- W modelu AM dużych opóźnień nie mamy żadnej korelacji pomiędzy $x(t)$ a $x(t-n-1)$. Stąd ACF odcina się na n-tym opóźnieniu. Łatwiej jest znaleźć opóźnienie dla szeregu MA niż AR.

ARMA

- W przypadku szeregu AR ta korelacja stopniowo ustępuje bez jakiejkolwiek wartości odcięcia.
- Jeśli znajdziemy częściową korelację każdego opóźnienia, zostanie ona przerwana po stopniu serii AR.
- Na przykład: jeśli mamy serię AR (1) i jeśli wykluczamy efekt pierwszego opóźnienia ($x(t-1)$), nasze drugie opóźnienie ($x(t-2)$) jest niezależne od $x(t)$. Stąd funkcja częściowej korelacji (PACF) spadnie gwałtownie po pierwszym opóźnieniu.



Clearly, the graph above has a cut off on PACF curve after 2nd lag which means this is mostly an AR(2) process.



Clearly, the graph above has a cut off on ACF curve after 2nd lag which means this is mostly a MA(2) process.

Statystyka Durbin-Watsona

- Innym popularnym testem korelacji szeregowej jest statystyka Durbin-Watsona.
- Statystyczna wartość DW mieści się w przedziale 0 - 4.
- Dodatnia korelacja jest związana z wartościami DW poniżej 2, a ujemna korelacja z wartościami DW powyżej 2.
- Wartość statystyki Durbin-Watsona wynosi blisko 2, jeśli błędy nie są ze sobą powiązane.

Poleca się stosowanie następujących zasad:

Poleca się stosowanie następujących zasad:

- **Reguła 1:** Jeśli ACF wykazuje rozkłady wykładnicze, a PACF ma skok przy opóźnieniu 1 oraz nie ma korelacji dla innych opóźnień, to użyj jednego parametru autoregresji (p).

Poleca się stosowanie następujących zasad:

- **Reguła 1:** Jeśli ACF wykazuje rozkłady wykładnicze, a PACF ma skok przy opóźnieniu 1 oraz nie ma korelacji dla innych opóźnień, to użyj jednego parametru autoregresji (p).
- **Zasada 2:** Jeśli ACF ma kształt sinusoidy lub powtarzający się wzorzec w kształcie funkcji wykładniczej, a PACF ma skoki przy opóźnieniach 1 i 2 oraz nie ma korelacji dla innych opóźnień, to użyj parametrów autoregresji (p) równej dwa.

Poleca się stosowanie następujących zasad:

- **Reguła 1:** Jeśli ACF wykazuje rozkłady wykładnicze, a PACF ma skok przy opóźnieniu 1 oraz nie ma korelacji dla innych opóźnień, to użyj jednego parametru autoregresji (p).
- **Zasada 2:** Jeśli ACF ma kształt sinusoidy lub powtarzający się wzorzec w kształcie funkcji wykładniczej, a PACF ma skoki przy opóźnieniach 1 i 2 oraz nie ma korelacji dla innych opóźnień, to użyj parametrów autoregresji (p) równej dwa.
- **Zasada 3:** Jeśli ACF ma skok w punkcie 1 i nie ma korelacji z innymi opóźnieniami, a PACF maleje wykładniczo, to użyj parametru średniej ruchomej (q) równego jeden.

Poleca się stosowanie następujących zasad:

- **Reguła 1:** Jeśli ACF wykazuje rozkłady wykładnicze, a PACF ma skok przy opóźnieniu 1 oraz nie ma korelacji dla innych opóźnień, to użyj jednego parametru autoregresji (p).
- **Zasada 2:** Jeśli ACF ma kształt sinusoidy lub powtarzający się wzorzec w kształcie funkcji wykładniczej, a PACF ma skoki przy opóźnieniach 1 i 2 oraz nie ma korelacji dla innych opóźnień, to użyj parametrów autoregresji (p) równej dwa.
- **Zasada 3:** Jeśli ACF ma skok w punkcie 1 i nie ma korelacji z innymi opóźnieniami, a PACF maleje wykładniczo, to użyj parametru średniej ruchomej (q) równego jeden.
- **Zasada 4:** Jeśli ACF ma skoki przy opóźnieniach 1 i 2 i nie ma korelacji z innymi opóźnieniami, a PACF ma kształt sinusoidy lub zespół rozkładów wykładniczych, to użyj parametrów średniej ruchomej (q) równej dwa.

Poleca się stosowanie następujących zasad:

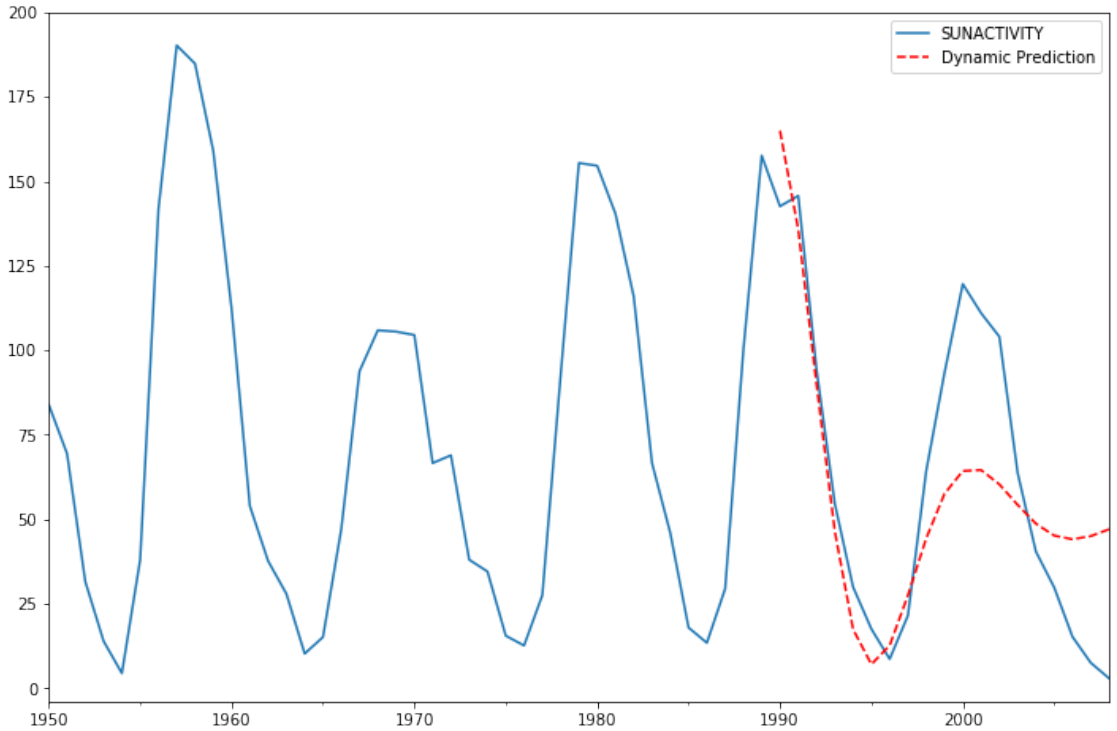
- **Reguła 1:** Jeśli ACF wykazuje rozkłady wykładnicze, a PACF ma skok przy opóźnieniu 1 oraz nie ma korelacji dla innych opóźnień, to użyj jednego parametru autoregresji (p).
- **Zasada 2:** Jeśli ACF ma kształt sinusoidy lub powtarzający się wzorzec w kształcie funkcji wykładniczej, a PACF ma skoki przy opóźnieniach 1 i 2 oraz nie ma korelacji dla innych opóźnień, to użyj parametrów autoregresji (p) równej dwa.
- **Zasada 3:** Jeśli ACF ma skok w punkcie 1 i nie ma korelacji z innymi opóźnieniami, a PACF maleje wykładniczo, to użyj parametru średniej ruchomej (q) równego jeden.
- **Zasada 4:** Jeśli ACF ma skoki przy opóźnieniach 1 i 2 i nie ma korelacji z innymi opóźnieniami, a PACF ma kształt sinusoidy lub zespół rozkładów wykładniczych, to użyj parametrów średniej ruchomej (q) równej dwa.
- **Zasada 5:** Jeśli ACF i PACF wygląda jak funkcja wykładnicza zaczynająca się od opóźnienia 1, to użyj parametru autoregresji (p =1) i parametru średniej ruchu (q=1).

ARMA

https://github.com/przem85/bootcamp/blob/master/statistics/D15_Z10.ipynb

Zadnie

Proszę dobrać optymalny parametr modelu ARMA dla danych.



ARMA

https://github.com/przem85/bootcamp/blob/master/statistics/D15_Z11.ipynb

https://github.com/przem85/bootcamp/blob/master/statistics/D15_Z12.ipynb

Zadanie

Proszę nauczyć model AR.

ARIMA

Wprowadzenie do ARIMA.
ARIMA oznacza **Auto-Regressive Integrated Moving Averages**.
Prognozowanie za pomocą ARIMA szeregów stacjonarnych jest niczym innym, jak równanie liniowe (jak regresja liniowa). Współczynniki predykcyjne zależą od parametrów (p, d, q) modelu ARIMA:

ARIMA

Wprowadzenie do ARIMA.
ARIMA oznacza **Auto-Regressive Integrated Moving Averages**.
Prognozowanie za pomocą ARIMA szeregów stacjonarnych jest niczym innym, jak równanie liniowe (jak regresja liniowa). Współczynniki predykcyjne zależą od parametrów (p, d, q) modelu ARIMA:

- Liczba AR (Auto-Regressive) (p): liczba lagów zmiennej zależnej. Na przykład jeśli p wynosi 5, predyktorami dla $x(t)$ będą $x(t-1), \ldots, x(t-5)$.

ARIMA

Wprowadzenie do ARIMA.
ARIMA oznacza **Auto-Regressive Integrated Moving Averages**.
Prognozowanie za pomocą ARIMA szeregów stacjonarnych jest niczym innym, jak równanie liniowe (jak regresja liniowa). Współczynniki predykcyjne zależą od parametrów (p, d, q) modelu ARIMA:

- Liczba AR (Auto-Regressive) (p): liczba lagów zmiennej zależnej. Na przykład jeśli p wynosi 5, predyktorami dla $x(t)$ będą $x(t-1), \ldots, x(t-5)$.
- Liczba MA (Moving Average) (q): liczba błędów w równaniu predykcyjnym. Na przykład jeśli q wynosi 5, predykcjami dla $x(t)$ będzie $e(t-1), \ldots, e(t-5)$, gdzie $e(i)$ jest różnicą między średnią ruchomą w chwili t, a wartością $X(t)$.

ARIMA

Wprowadzenie do ARIMA.
ARIMA oznacza **Auto-Regressive Integrated Moving Averages**.
Prognozowanie za pomocą ARIMA szeregów stacjonarnych jest niczym innym, jak równanie liniowe (jak regresja liniowa). Współczynniki predykcyjne zależą od parametrów (p, d, q) modelu ARIMA:

- Liczba AR (Auto-Regressive) (p): liczba lagów zmiennej zależnej. Na przykład jeśli p wynosi 5, predyktorami dla $x(t)$ będą $x(t-1), \ldots, x(t-5)$.
- Liczba MA (Moving Average) (q): liczba błędów w równaniu predykcyjnym. Na przykład jeśli q wynosi 5, predykcjami dla $x(t)$ będzie $e(t-1), \ldots, e(t-5)$, gdzie $e(i)$ jest różnicą między średnią ruchomą w chwili t, a wartością $X(t)$.
- Liczba różnicowań (Differences) (d): stopień wykorzystanego różnicowania. Albo możemy przekazać $X(t) - X(t-1)$ i użyć $d = 0$ lub przekazać pierwotną zmienną i użyć $d = 1$. Obie generują takie same wyniki.

ARIMA

https://github.com/przem85/bootcamp/blob/master/statistics/D15_Z13.ipynb

https://github.com/przem85/bootcamp/blob/master/statistics/D15_Z14.ipynb

Zadanie

Proszę nauczyć model ARIMA.

ARIMA

- Aby dobrze nauczyć model nie możemy uczyć na tych samych danych, co testujemy.
- W tym celu dzielimy nasze dane na zbiór treningowy i testowy.
- Możemy powtórzyć proces dzielenia szeregu czasowego na zestawy podciągów i testować wielokrotnie.
- Metoda ta używa coraz to większą i większą historię w procesie uczenia.

`https://github.com/przem85/bootcamp/blob/master/statistics/D15_Z15.ipynb`

Uczenie nadzorowane

- Problem prognozowania szeregów czasowych możemy sformułować w klasycznej formie nauczania nadzorowanego.
- Uczenie nadzorowane polega na analizie zmiennych wejściowych X i zmiennej wyjściowej y i użyciu algorytmów, które uczą się zależności między X, a y.

$$y = f(X)$$

- Celem jest przybliżenie rzeczywistej zależności tak dobrze, że gdy masz nowe dane wejściowe X, możesz przewidzieć zmienne wyjściowe y dla tych danych.

Uczenie nadzorowane

- Problem przewidywania wartości szeregu czasowego można sformułować jako problem uczenia maszynowego.
- Biorąc pod uwagę sekwencję liczb (szereg czasowy), możemy przetransformować dane, tak aby wyglądały jak dane używane w nauczaniu nadzorowanym.
- Możemy to zrobić używając poprzednich kroków czasowych jako zmiennych wejściowych i użyć następnego kroku jako zmiennej wyjściowej.

Uczenie nadzorowane

Przykład:
Wyobraźmy sobie, że mamy szereg czasowy dany w następujący sposób:

time, measure
1, 100
2, 110
3, 108
4, 115
5, 120

Uczenie nadzorowane

Możemy troszkę przerobić te dane używając wartości w poprzednim kroku czasowym, aby przewidzieć wartość w następnym kroku.

W naszym przypadku wyglądało by to tak:
X, y
?, 100
100, 110
110, 108
108, 115
115, 120
120, ?

Uczenie nadzorowane

Oto kilka obserwacji:

- Widzimy, że poprzednim krokiem czasowym jest wejście (X), a następnym krokiem jest wyjście (y).
- Widzimy, że porządek między obserwacjami jest zachowywany i musi być zachowany w przypadku korzystania z tego zestawu danych do nauczania maszynowego.
- Widzimy, że nie mamy poprzedniej wartości, którą możemy użyć do przewidywania pierwszej wartości w sekwencji. Usuniemy ten wiersz, ponieważ nie możemy go użyć.
- Możemy również zobaczyć, że nie mamy znanej kolejnej wartości dla ostatniej wartości w szeregu czasowym. Możemy też usunąć tę wartość podczas uczenia.

Uczenie nadzorowane

Uwaga

Użycie kroków poprzedzających czas, aby przewidzieć następny krok czasu, nazywa się **sliding window method**.

Liczba poprzednich kroków czasowych nazywana jest szerokością okna.

Uczenie nadzorowane

Liczba obserwacji zarejestrowanych w danym czasie może być różna:

- Univariate Time Series: są to zestawy danych, w których za każdym razem obserwuje się tylko jedną zmienną, taką jak temperatura mierzona co godzinę.
- Multivariate Time Series: są to zestawy danych, w których za każdym razem obserwuje się dwie lub więcej zmiennych.

Uczenie nadzorowane

Przykład:
Załóżmy, że w każdym kroku czasowym mamy dwa zestawy danych.
Przyjmijmy również, że chodzi tylko o przewidywanie: measure2

time, measure1, measure2
1, 0.2, 88
2, 0.5, 89
3, 0.7, 87
4, 0.4, 88
5, 1.0, 90

Uczenie nadzorowane

Możemy przekształcić ten szereg czasowy do problemu nauczania nadzorowanego z szerokością okna równą 1.
Oznacza to, że użyjemy poprzednich wartości dla measure1, measure2.
Następnie będziemy próbować przewidzieć kolejną wartość dla measure2.

X1, X2, X3, y
?, ?, 0.2 , 88
0.2, 88, 0.5, 89
0.5, 89, 0.7, 87
0.7, 87, 0.4, 88
0.4, 88, 1.0, 90
1.0, 90, ?, ?

Uczenie nadzorowane

Możemy zauważyć, że podobnie jak w przypadku poprzednim, musimy usunąć pierwsze i ostatnie rzędy w celu uruchomienia algorytmu nauczania maszynowego.
Gdy chcemy przewidzieć zarówno measure1 jak i measure2, to również zastosujemy to podejście:

X1, X2, y1, y2
?, ?, 0.2, 88
0.2, 88, 0.5, 89
0.5, 89, 0.7, 87
0.7, 87, 0.4, 88
0.4, 88, 1.0, 90
1.0, 90, ?, ?

Uczenie nadzorowane

Możemy zauważyć, że podobnie jak w przypadku poprzednim, musimy usunąć pierwsze i ostatnie rzędy w celu uruchomienia algorytmu nauczania maszynowego.
Gdy chcemy przewidzieć zarówno measure1 jak i measure2, to również zastosujemy to podejście:

X1, X2, y1, y2
?, ?, 0.2, 88
0.2, 88, 0.5, 89
0.5, 89, 0.7, 87
0.7, 87, 0.4, 88
0.4, 88, 1.0, 90
1.0, 90, ?, ?

Nieliczne metody nauczania nadzorowanego mogą poradzić sobie z przewidywaniem wielu wartości, ale niektóre sieci neuronowe i różne modyfikacje klasycznych algorytmów dają sobie radę.

Uczenie nadzorowane

Podstawową funkcją, która pomaga przekształcić szereg czasowy do postaci wymaganej przez algorytmy uczenia maszynowego, jest funkcja `shift()`.

`https://github.com/przem85/bootcamp/blob/master/statistics/D15_Z16.ipynb`

`https://github.com/przem85/bootcamp/blob/master/statistics/D15_Z17.ipynb`