# Understanding Autocorrelation and Partial Autocorrelation Functions (ACF and PACF)

Introduction

András Kis · Follow

4 min read · Jul 26, 2024



In time series analysis, understanding the relationship between observations at different points in time is crucial. Two important tools for this are the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF). This article will guide you through the concepts of ACF and PACF, how to interpret their plots, and provide real-life examples and code snippets to enhance your understanding.

## What is Autocorrelation?

Autocorrelation, also known as serial correlation, measures the relationship between a time series and a lagged version of itself over successive time intervals. Simply put, it tells you how similar the data points are to each other at different time lags.

**Example:** Imagine you are tracking the temperature of a city every day. If today's temperature is similar to yesterday's, and yesterday's temperature is similar to the day before, we say that the temperature data is autocorrelated.

### The Autocorrelation Function (ACF)

The ACF plots the correlation of the time series with itself at different lags. This helps in identifying patterns such as seasonality, trends, and the persistence of values over time.

- **Lag 1:** Correlation between observations at time $t$ and $t-1$
- **Lag 2:** Correlation between observations at time $t$ and $t-2$
- And so on…

### Interpreting ACF Plots

When you look at an ACF plot, you'll see bars at each lag. The height of the bar represents the correlation coefficient at that lag.

- **Significant Lag:** If a bar extends beyond the significance bounds, it indicates significant autocorrelation at that lag.

- **Gradual Decline:** A gradual decline in bar heights suggests a long-term dependency in the data.
- **Seasonal Patterns:** Regular spikes at certain lags suggest seasonality in the data.

### Partial Autocorrelation

Partial autocorrelation measures the correlation between observations at two time points, accounting for the values of the observations at all shorter lags. This helps isolate the direct relationship between observations at different lags, removing the influence of intermediary observations.

### The Partial Autocorrelation Function (PACF)

The PACF plot shows the partial correlation of the time series with itself at different lags.

### Interpreting PACF Plots

The PACF plot helps determine the order of an autoregressive model (AR model).

- **Significant Lag:** A significant spike at a particular lag suggests the inclusion of that lag in the AR model.

- **Cut-off Point:** The lag at which the PACF plot cuts off helps determine the maximum lag to include in the AR model.

### Real-Life Example: Stock Prices

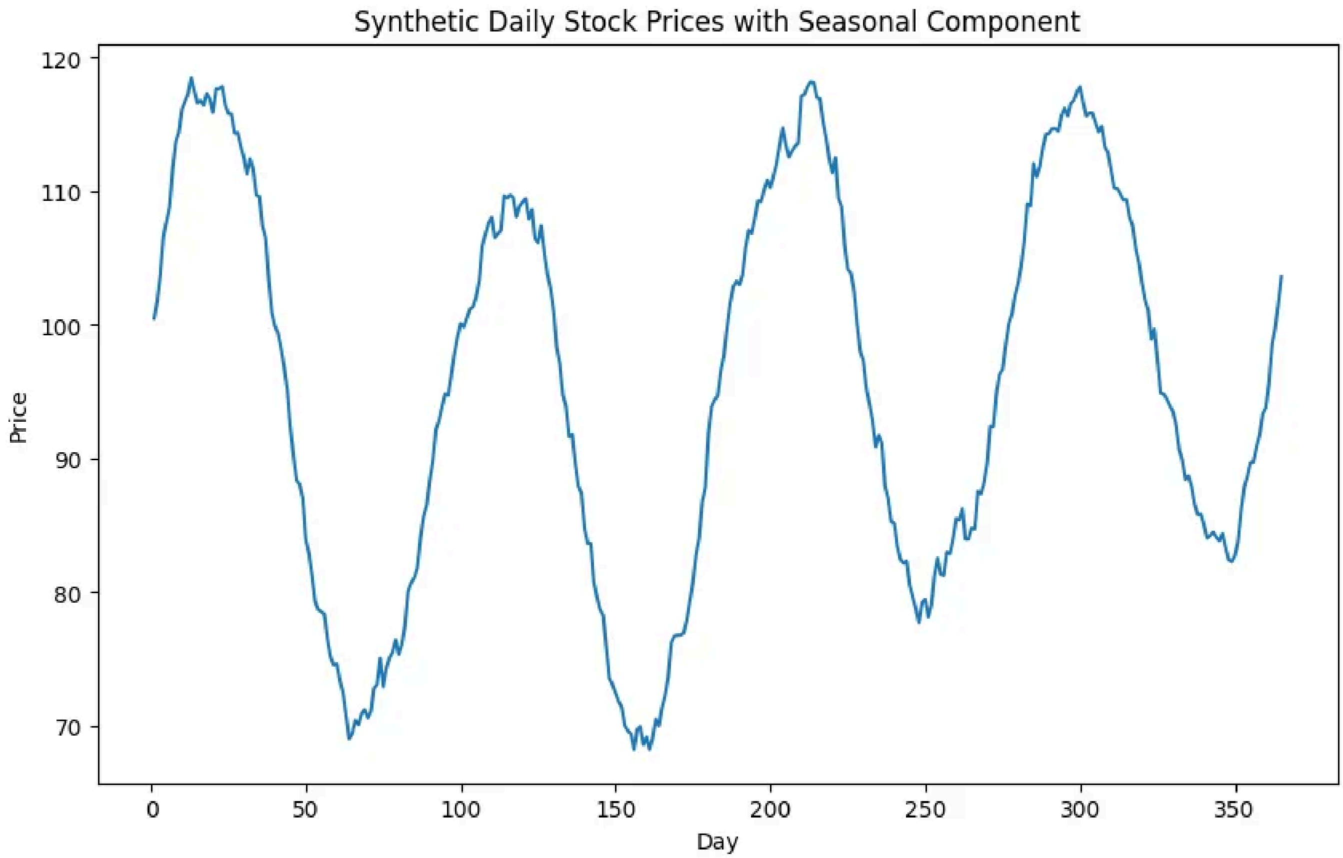
Consider the daily closing prices of a stock. If today's price is influenced by the previous day's price, and yesterday's price is influenced by the day before, the prices show autocorrelation. The ACF and PACF plots of stock prices can help identify trends and the influence of past prices on the current price.

Code Example:

```
import pandas as pd
import matplotlib.pyplot as plt
from statsmodels.graphics.tsaplots import plot_acf, plot_pacf
```

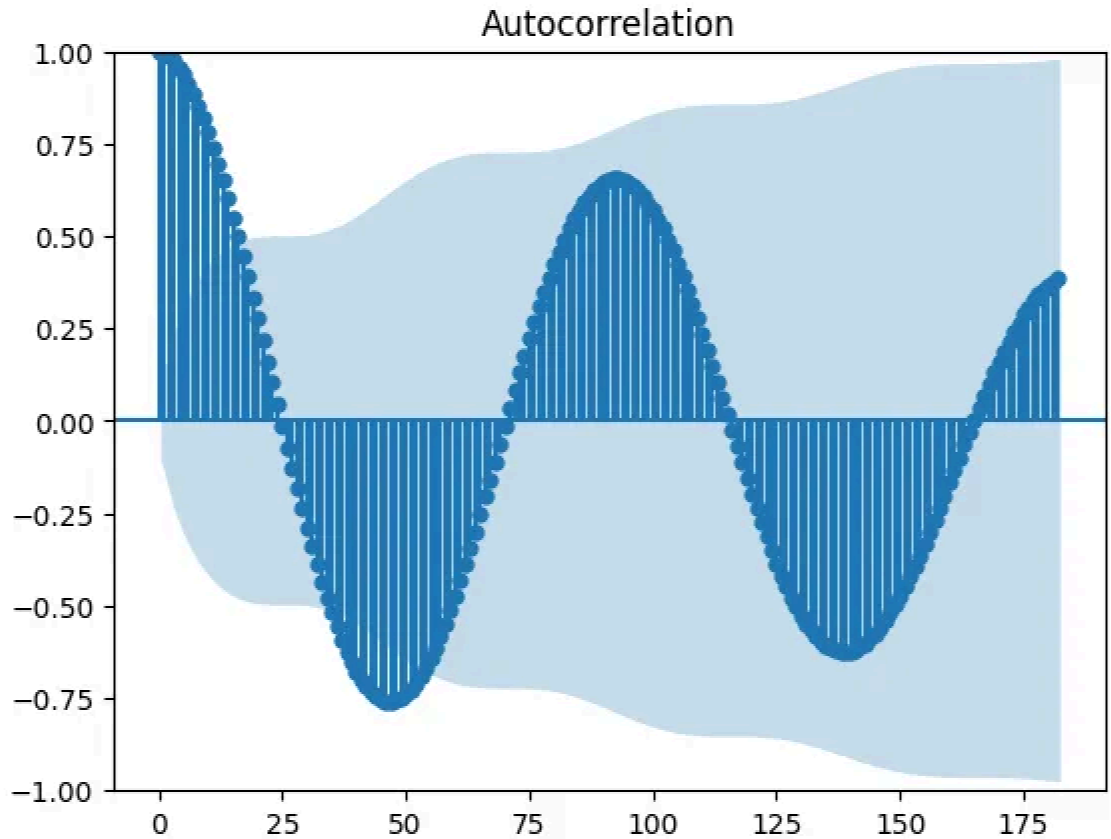
```
# Example data: daily stock prices
data = pd.read_csv('synthetic_stock_prices.csv')

# Plot the time series data
plt.figure(figsize=(10, 6))
plt.plot(data['Day'], data['Price'])
plt.title('Synthetic Daily Stock Prices with Seasonal Component')
plt.xlabel('Day')
plt.ylabel('Price')
plt.show()
```



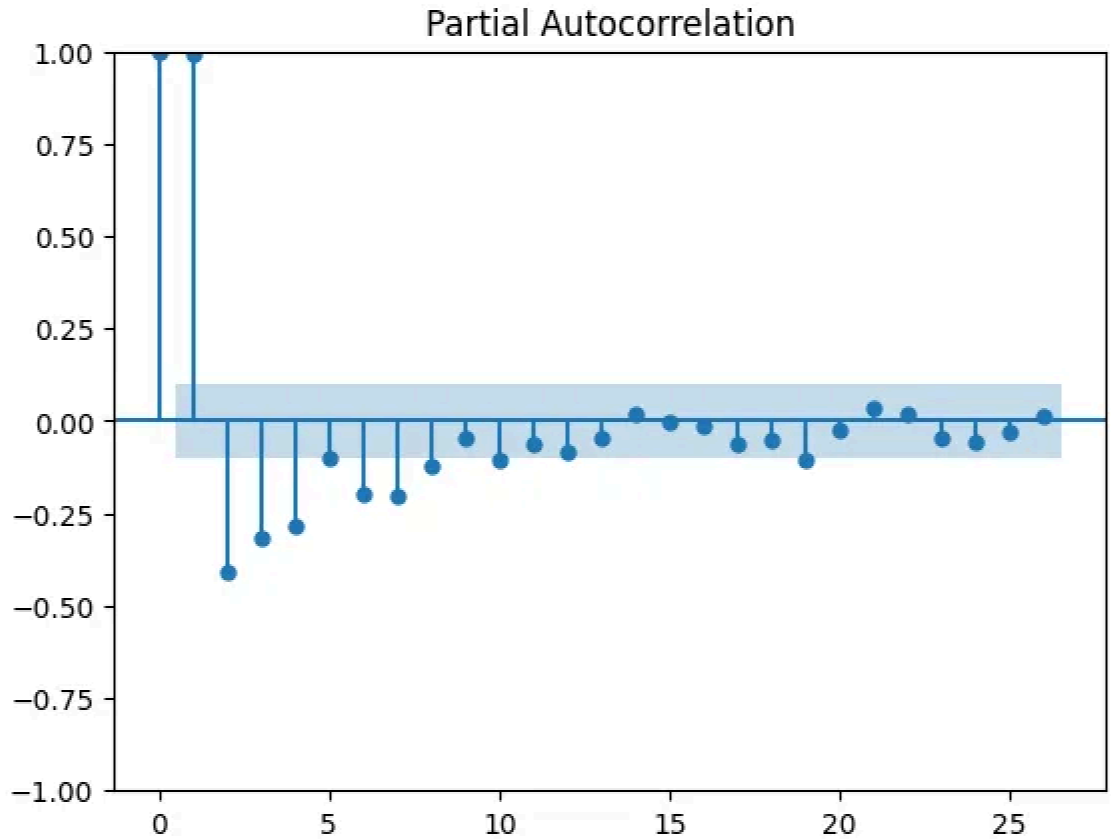

Output

```
# ACF plot
plot_acf(data['Price'], lags=len(data)//2)
plt.show()
```




Output

```
# PACF plot
plot_pacf(data['Price'])
plt.show()
```




Output

- **ACF Plot:** The ACF plot shows a significant sinusoidal pattern with a periodicity, indicating strong seasonal behavior. The slow decay in the autocorrelations suggests a non-stationary series with periodic fluctuations.
- **PACF Plot:** The PACF plot shows significant spikes at lag 1, 2, 3, and 4, indicating that these lags are important in modeling the underlying data.

Given these observations, the data should be transformed to achieve stationarity before applying an AR model. After you achieve stationarity, you can then fit an AR model, potentially including terms up to lag 4, as indicated by the PACF plot.

## Conclusion

Understanding ACF and PACF is vital for analyzing time series data. These functions help in identifying patterns, trends, and the order of autoregressive models, making them essential tools in fields such as finance, meteorology, and economics. By interpreting ACF and PACF plots, you can gain insights into the structure of your time series data and make informed decisions based on historical patterns.

Time Series Analysis | Time Series Forecasting | Autoregressive | Statistics | Data Pre Processing

## Written by András Kis

5 Followers · 19 Following

Follow

Hi! I'm Andras Nandor Kis, passionate about data analysis, ML, AI and MLOps. I share my continuous learning journey. Let's connect and grow together!