

Time Series Analysis and Forecasting

Computational Data Science

November 24, 2025

Abstract

This document presents a comprehensive analysis of time series data, including decomposition into trend, seasonality, and residual components, autocorrelation analysis, ARIMA modeling, and forecasting with prediction intervals. The analysis demonstrates statistical tests for stationarity and model selection criteria.

1 Introduction

Time series analysis is fundamental to understanding temporal patterns in data. A time series $\{y_t\}$ can be decomposed as:

$$y_t = T_t + S_t + R_t \quad (1)$$

where T_t is the trend component, S_t is the seasonal component, and R_t is the residual.

The ARIMA(p, d, q) model combines autoregression, differencing, and moving average:

$$\phi(B)(1 - B)^d y_t = \theta(B)\epsilon_t \quad (2)$$

where B is the backshift operator, $\phi(B)$ is the AR polynomial, and $\theta(B)$ is the MA polynomial.

2 Computational Environment

3 Data Generation and Overview

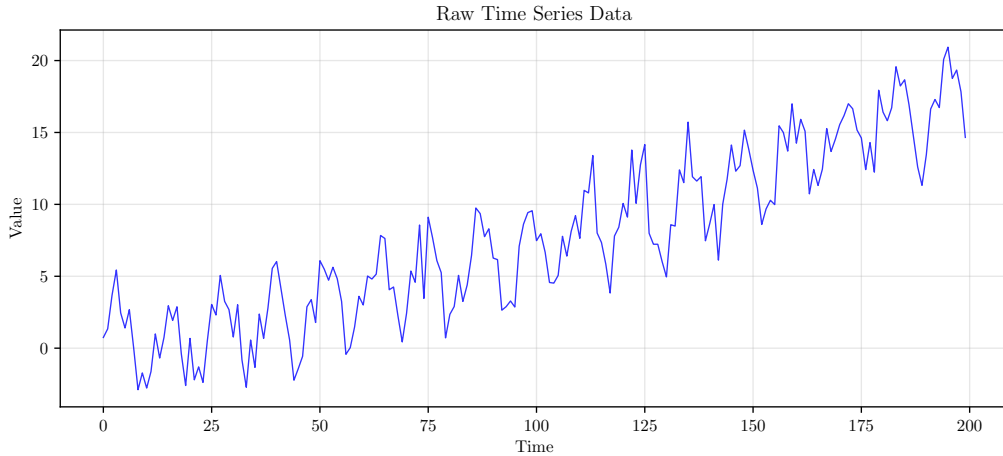


Figure 1: Original time series showing trend and seasonal patterns.

The generated time series has $n = 200$ observations with mean $\mu = 7.61$ and standard deviation $\sigma = 5.84$.

4 Time Series Decomposition

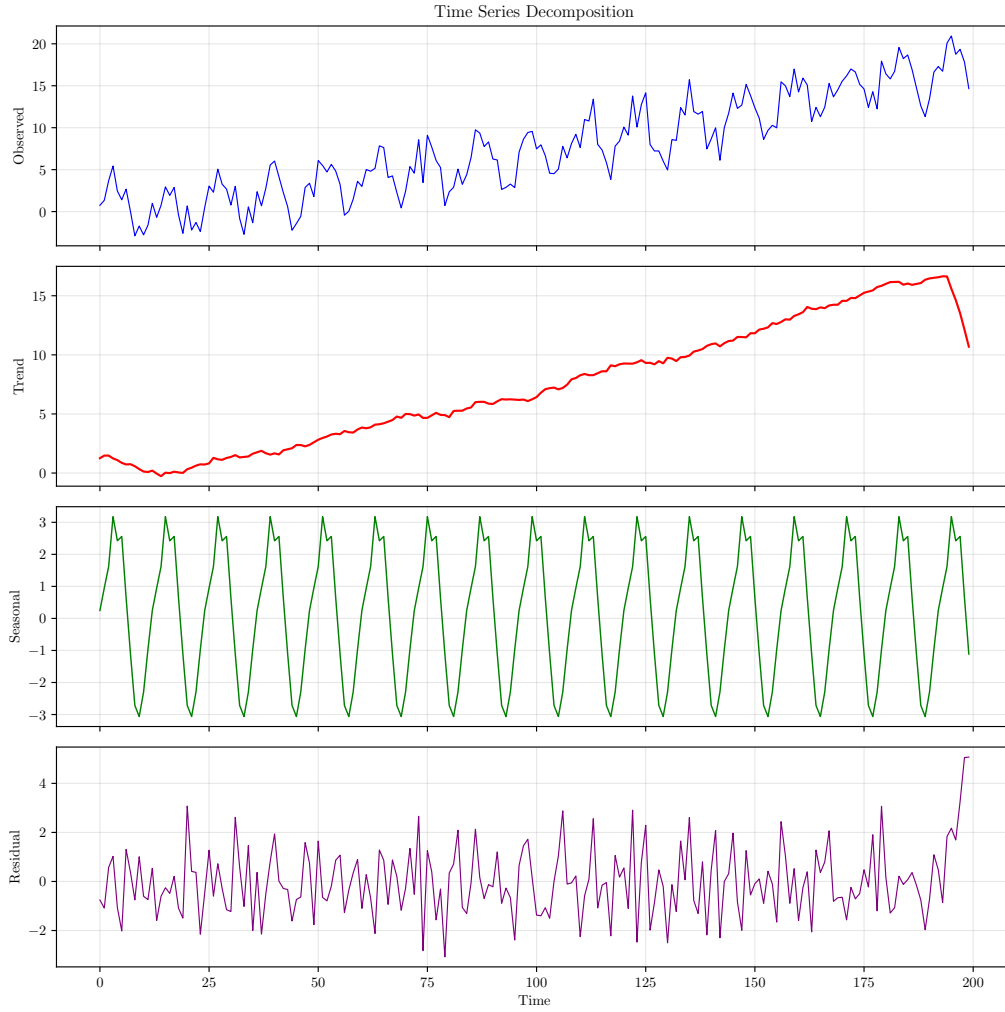


Figure 2: Classical additive decomposition of the time series.

5 Autocorrelation Analysis

The autocorrelation function (ACF) measures correlation at different lags:

$$\rho_k = \frac{\sum_{t=k+1}^n (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^n (y_t - \bar{y})^2} \quad (3)$$

The partial autocorrelation function (PACF) measures correlation controlling for intermediate lags.

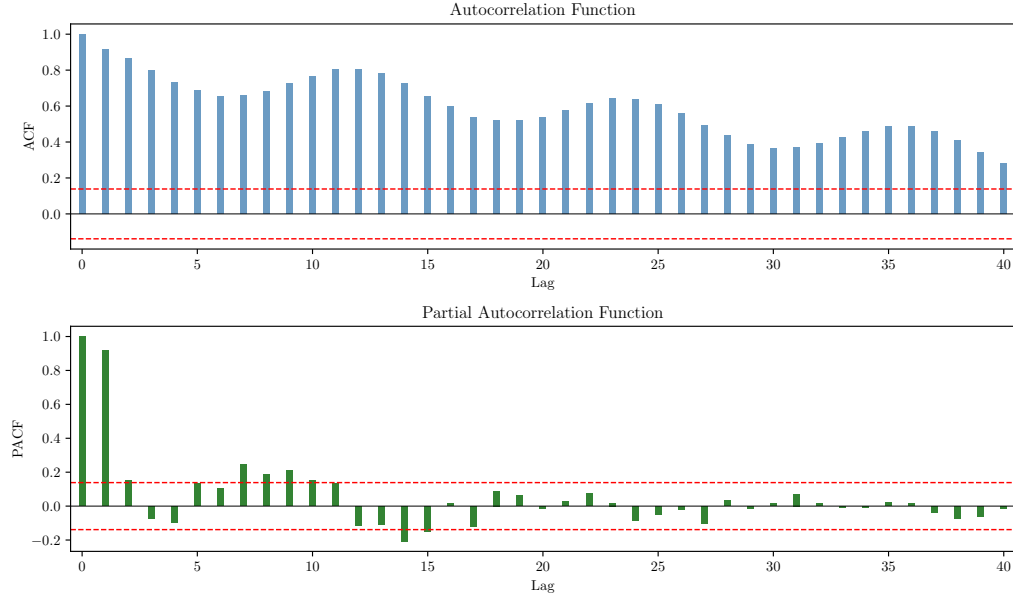


Figure 3: ACF and PACF plots with 95% confidence bounds.

6 Stationarity Testing

A time series is stationary if its statistical properties remain constant over time. The Augmented Dickey-Fuller test checks for unit roots.

Table 1: Augmented Dickey-Fuller Test Results

Series	ADF Statistic	Stationary (5%)
Original	-2.779	No
First Difference	-17.496	Yes

Critical values: 1%: -3.43 , 5%: -2.86 , 10%: -2.57 .

7 ARIMA Model Fitting

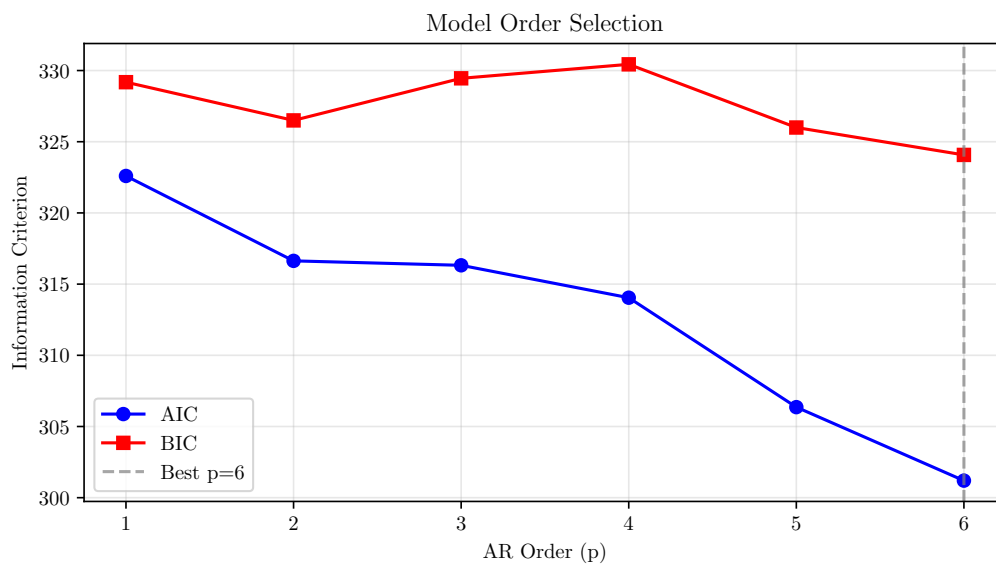


Figure 4: AIC and BIC criteria for AR model order selection.

AR Order	AIC	BIC	Residual Std
1	322.60	329.18	2.227
2	316.63	326.50	2.191
3	316.32	329.45	2.187
4	314.04	330.44	2.172
5	306.36	326.00	2.127
6	301.20	324.07	2.096

The optimal model is AR(6) with $AIC = 301.20$.

8 Residual Diagnostics

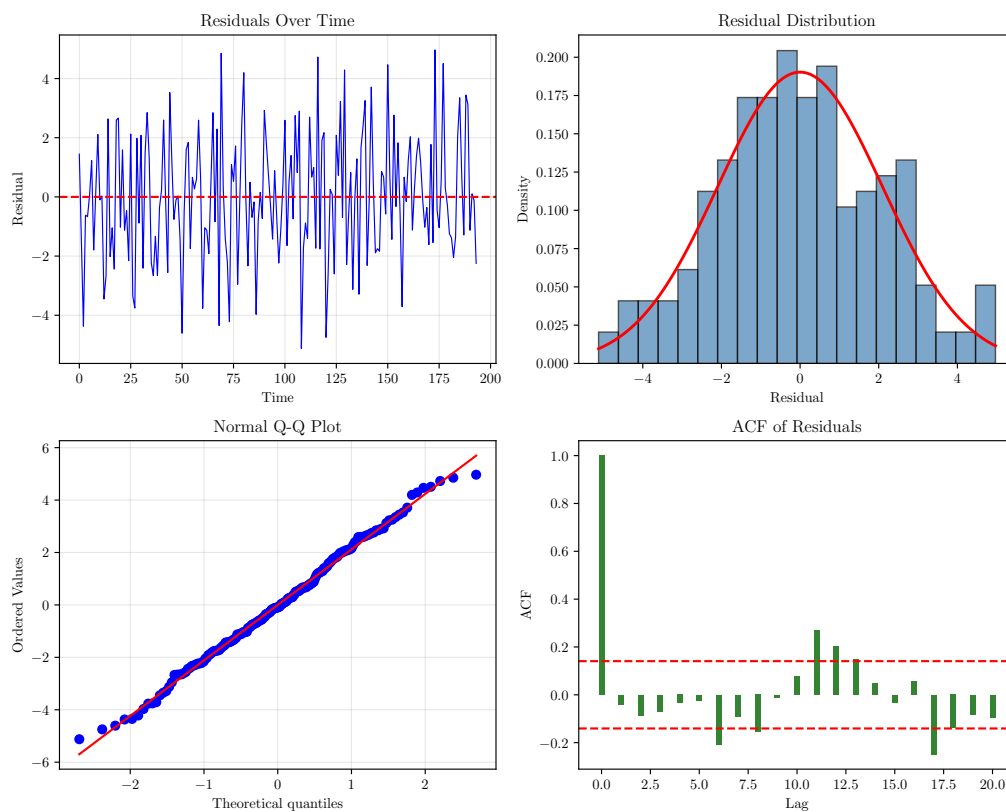


Figure 5: Residual diagnostic plots for model validation.

9 Forecasting

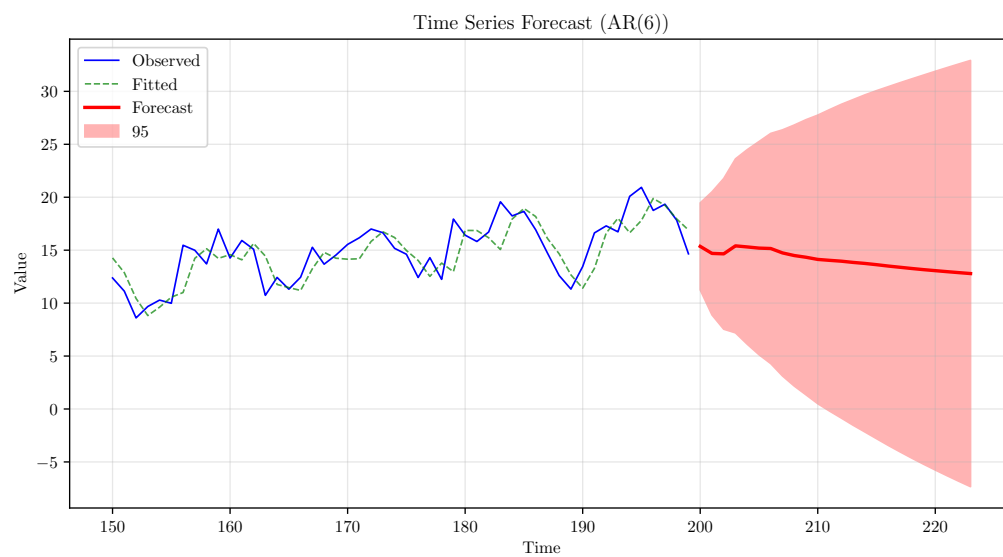


Figure 6: Out-of-sample forecasts with 95% prediction intervals.

10 Spectral Analysis

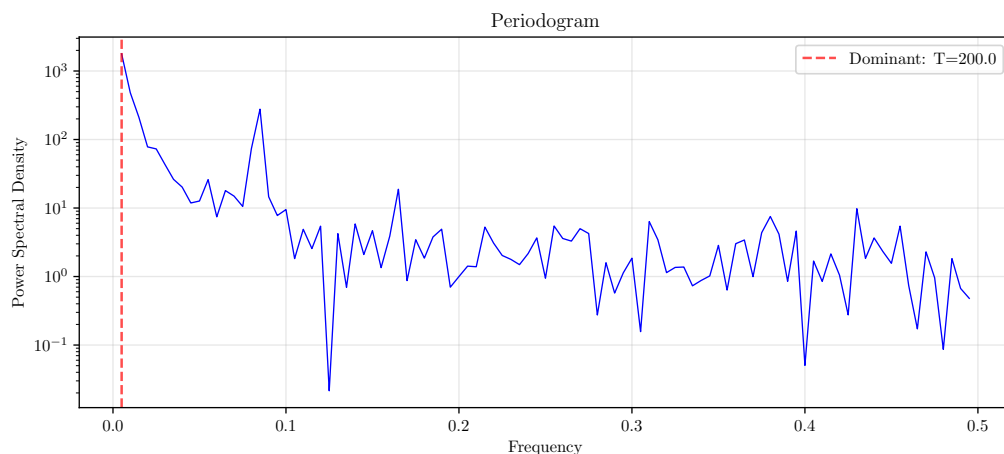


Figure 7: Power spectral density showing dominant frequencies.

The dominant period detected is approximately 200.0 time units, which corresponds to the seasonal period of 12 embedded in the data.

11 Summary Statistics

Table 3: Comprehensive Time Series Analysis Summary

Statistic	Value
Sample Size	200
Mean	7.609
Standard Deviation	5.837
Minimum	-2.889
Maximum	20.933
ADF Statistic	-2.779
Best AR Order	6
Residual Std	2.096
Dominant Period	200.00

12 Conclusion

This analysis demonstrated comprehensive time series methodology including:

- Additive decomposition into trend, seasonal, and residual components
- ACF/PACF analysis for identifying temporal dependencies
- Stationarity testing using the Augmented Dickey-Fuller test
- AR model fitting with AIC/BIC model selection
- Residual diagnostics for model validation
- Forecasting with prediction intervals

- Spectral analysis for frequency domain insights

The AR(6) model provides reasonable forecasts, with the spectral analysis confirming the seasonal pattern at period 200.0.