Project Write-up of 'Visualization of Autoregressive(AR) Models' Project

DataWaves - Libin N George, Ajeet Singh, Saharsh Bhave, Navyasree Madhu

Objective

Presenting and visualizing technical topic in statistics, focusing on Autoregressive (AR) models, to demonstrate how parameters affect patterns in time series data.

Introduction

This project aims to explore and visually demonstrate how autoregressive (AR) models operate, specifically how different parameters and AR orders impact the structure and behavior of a time series. AR models are essential tools in time series analysis, with broad applications across economics, finance, weather forecasting, and more. Understanding the visual and statistical characteristics of AR(1), AR(2), and AR(3) models, as well as how different parameter values influence the model's output, provides insight into data pattern recognition and prediction capabilities.

We will use a generated dataset to simulate various AR model scenarios. This approach allows us to precisely control parameters and observe how each setting impacts the time series and how to enable precise, tailored demonstrations of AR model behavior. Our visualizations will feature multiple plots for AR(1), AR(2), and AR(3) models, showcasing stationary and non-stationary processes, cyclical patterns, and convergence to the mean. We will also develop an interactive Shiny app that allows users to work on different AR model parameters and visualize the resulting time series. This approach helps to understand through direct engagement.

Real-World Applications of AR Models

Autoregressive(AR) models are foundational tools in time series analysis and have significant real-world applications across various domains. In finance, they are used to forecast stock prices, assess risks, and optimize portfolios. In meteorology, AR models aid in weather prediction and climate analysis, while in healthcare, they help track disease outbreaks and forecast hospital resource needs. Industries like energy and telecommunications rely on AR models for demand forecasting and network traffic prediction. These applications highlight the power of AR models in analyzing historical patterns to make informed predictions, enabling better decision-making and resource optimization.

Generic AR Model Equation

The general formula for an auto regressive model of order p , or AR(p), is:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \epsilon_t$$

where:

- X_t is the value of the time series at time t,
- $\phi_1, \phi_2, \dots, \phi_p$ are the auto regressive coefficients,
- p is the order of the AR model, and
- ϵ_t is a white noise error term at time t, with mean zero and constant variance.

Justification of approach

To understand the behavior of AR (AutoRegressive) models under different configurations, we can simulate time series data using the arima.sim() function in R. This function is well-suited for generating synthetic data that follows ARIMA processes, allowing us to study how different AR and MA configurations affect the patterns in a time series. Specifically, we use it to generate data for AR models with varying parameters. In the data analysis plan, we did the following variations in AutoRegressive models and generate data for these models and created visualizations from the data.

1. AR(1) - Random Walk
2. AR(1) - Having moderate positive correlation, decaying toward the mean over time.
3. AR(1) - Data will oscillate, alternating signs and showing a negative correlation
4. AR(2) - Data with oscillatory behavior
5. AR(2) - Data without oscillatory behavior
6. AR(3) - Data with Complex behavior We created a line graph and associated Autocorrelation graph for the above models to visualize the statistical information and how the data behaves. We have included a shiny app that enables the user to change parameters and plot the line graph.

Code

1. AR(1) with positive and negative coefficients