Proposal for 'Visualization of Autoregressive(AR) Models' Project

Data Waves - 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,
- $\phi_1, \phi_2, \ldots, \phi_p$ 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.

Data Generation

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 will use it to generate data for AR models with varying parameters.

Data Analysis Plan

We will consider the following variations in AutoRegressive models, We will generate data for these models and create visualizations from the data.

1. AR(1) - Random Walk AR(1) - Having moderate positive correlation, decaying toward the mean over time.
2. AR(1) - Data will oscillate, alternating signs and showing a negative correlation
3. AR(2) - Data with oscillatory behavior

4. AR(2) - Data without oscillatory behavior
5. AR(3) - Data with Complex behavior

We will create a line graph and associated Autocorrelation graph for the above models to visualize the statistical information and how the data behaves. We will also create a shiny app that enables the user to change parameters and plot the line graph.

Plan for Poster

We will make a tree structure showing how the different visualizations we have explain the different features of AR model. We will include a small real world example to show how the AR model can be used to predict the future.

Weekly Plan of Attack

TaskName	Status	Assignee	Due.Date	Priority	Summary
Research and Background Setup	Completed	Libin, Saharsh, Ajeet, Navyasree	Nov 6	High	Research AR model theory, and types of processes, and decide on parameter settings for AR(1), AR(2), AR(3) models
Create proposal	Completed	Libin, Saharsh, Ajeet, Navyasree	Nov 13	High	Create a proposal for an AR model with a detailed attack plan.
Update the Proposal after peer review	TODO	Libin, Saharsh, Ajeet, Navyasree	Nov 18	High	Update the Proposal after incorporating the feedback and suggestions from Peer review and submission for instructor review.
Static Visualization in ggplot2 AR(1) - Random Walk	TODO	Saharsh	Nov 25	High	Generate time series data with $\phi_1 = 1$ for AR(1) model and Create static plots in ggplot2 for AR model for Random Walk.

Static Visualization in ggplot2 AR(1) - Positive	TODO	Ajeet	Nov 25	High	Generate time series data with $\phi_1 > 0$ for AR(1) model and Create static plots in ggplot2 for AR model for Positive Correlation.
Static Visualization in ggplot2 AR(1) - Negative	TODO	Navyasree	Nov 25	High	Generate time series data with $\phi_1 < 0$ for AR(1) model and Create static plots in ggplot2 for AR model for Negative Correlation.
Static Visualization in ggplot2 AR(2) - Positive	TODO	Libin	Nov 25	High	Generate time series data with $\phi_2 > 0$ for AR(2) model and Create static plots in ggplot2 for AR model without Oscillatory Behavior.
Static Visualization in ggplot2 AR(2) - Negative	TODO	Navyasree, Ajeet	Nov 25	High	Generate time series data with $\phi_2 < 0$ for AR(2) model and Create static plots in ggplot2 for AR model with Oscillatory Behavior.
Static Visualization in ggplot2 AR(3) - Complex Model	TODO	Saharsh, Libin	Nov 25	High	Generate time series data for AR(3) model and Create static plots in ggplot2 for AR model with Complex Behavior.
Shiny Interactive Plot Development	TODO	Libin, Saharsh, Ajeet, Navyasree	Dec 2	High	Develop Shiny app to allow users to adjust AR model parameters and visualize results in real-time, including dynamic plot updates.
Creating templates for Posters	TODO	Libin, Saharsh, Ajeet, Navyasree	Dec 2	High	Create templates for posters for peer review and feedback in class.

Final Analysis and Report	TODO	Libin, Saharsh, Ajeet, Navyasree	Dec 9	High	Summarize findings, discuss insights, and compile visualizations and interpretations for the final report.
Repository Organization and Documentation	TODO	Libin, Saharsh, Ajeet, Navyasree	Dec 9	High	Ensure clear structure and complete documentation, including README files for each folder, usage guidelines for plots, and Shiny app.
Project Write-Up	TODO	Libin, Saharsh, Ajeet, Navyasree	Dec 9	High	Create the project write-up.
Project Presentation/Poster for iShowcase	TODO	Libin, Saharsh, Ajeet, Navyasree	Dec 9	High	Create Project Presentation.

Final Repository Organization

- **data/**: Contains generated datasets and relevant metadata. It includes README, which details dataset parameters and generation processes.
- **plots/**: Includes static ggplot2 visualizations for AR models (AR(1), AR(2), and AR(3)) with various parameter combinations. README explains each plot and parameter choices.
- **shiny_app/**: Holds the Shiny app files for interactive plotting of AR models. README provides instructions for using the app and explains the interactive features.
- **docs/**: Project report, analysis write-ups, and weekly task updates. README outlines project goals, methods, and findings.
- **proposal.qmd:** Project proposal file in qmd format.
- **proposal.pdf**: Project proposal auto generated from proposal.qmd file.

Each folder will contain a README.md file explaining its contents and purpose in the project workflow.