In The Name Of God

Shiraz University

Estimation Theory

Dr. Somayeh Afrasiabi

HW 3

Alireza Garmsiri

40260337

Introduction

In many fields of science and engineering, parameter estimation is essential for understanding and modeling complex systems based on observed data. One common technique used for parameter estimation is **Least Squares Estimation (LSE)**, where the objective is to minimize the sum of the squared differences between observed values and predicted values. In this project, we apply the Least Squares Estimation method to estimate a parameter A from noisy data generated by an autoregressive process, specifically using a simple linear model with added Gaussian noise.

This report explores how the parameter A is estimated from a set of noisy observations. The noise added to the data introduces uncertainty in the estimation process. We simulate this process for various sample sizes and observe the effects on the accuracy and variability of the estimated parameter. The code generates multiple trials of estimates for different sample sizes, calculates the mean and standard deviation of the estimates, and then visualizes the results.

Implementation

Problem Description

In this experiment, the true parameter A is observed indirectly through noisy data. The observations are generated using a model where the true value of A is added to random Gaussian noise. The objective is to estimate the value of A from these noisy observations using the Least Squares Estimation method.

The model for the noisy observations is described as:

The observed value is the true value A plus random Gaussian noise. This creates uncertainty in the observation, making it a typical scenario in many real-world applications where data is corrupted by random fluctuations (noise).

For each trial, we estimate the value of A based on the noisy observations and calculate the mean and standard deviation of the estimates across multiple trials to understand the reliability of the estimation process.

Methodology

Simulation of Observations:

In each trial, a set of N observations is generated, where each observation consists of the true parameter A added to Gaussian noise. The standard deviation of the noise is determined by A, making the noise scale proportional to the true parameter.

Estimate Calculation:

The parameter A is estimated from the noisy data using a formula derived from the Least Squares Estimation method. The method computes an estimate for A by minimizing the sum of squared errors between the observed values and predicted values. The estimate is then stored for further analysis.

Monte Carlo Simulations:

To understand the impact of random noise on the estimation, the process is repeated M times for each sample size NNN. This is done using **Monte Carlo simulations**, which provide a way to account for randomness in the estimation process. Each trial produces an estimate of A, and these estimates are collected for analysis.

Statistics Calculation:

For each sample size NNN, the mean and standard deviation of the estimates are calculated across all trials. These statistics provide insight into the accuracy (mean) and variability (standard deviation) of the estimates.

Result

Effect of Sample Size on Estimation:

The simulation was run for different values of N (the number of observations), ranging from 5 to 400, with increments of 20. For each N, 10,000 trials were performed to estimate A, and the results were stored.

The **Gaussian Distribution Plot** reveals that as N increases, the distribution of the estimates becomes more concentrated around the true value of A. This shows that with more data, the estimation process becomes more accurate.

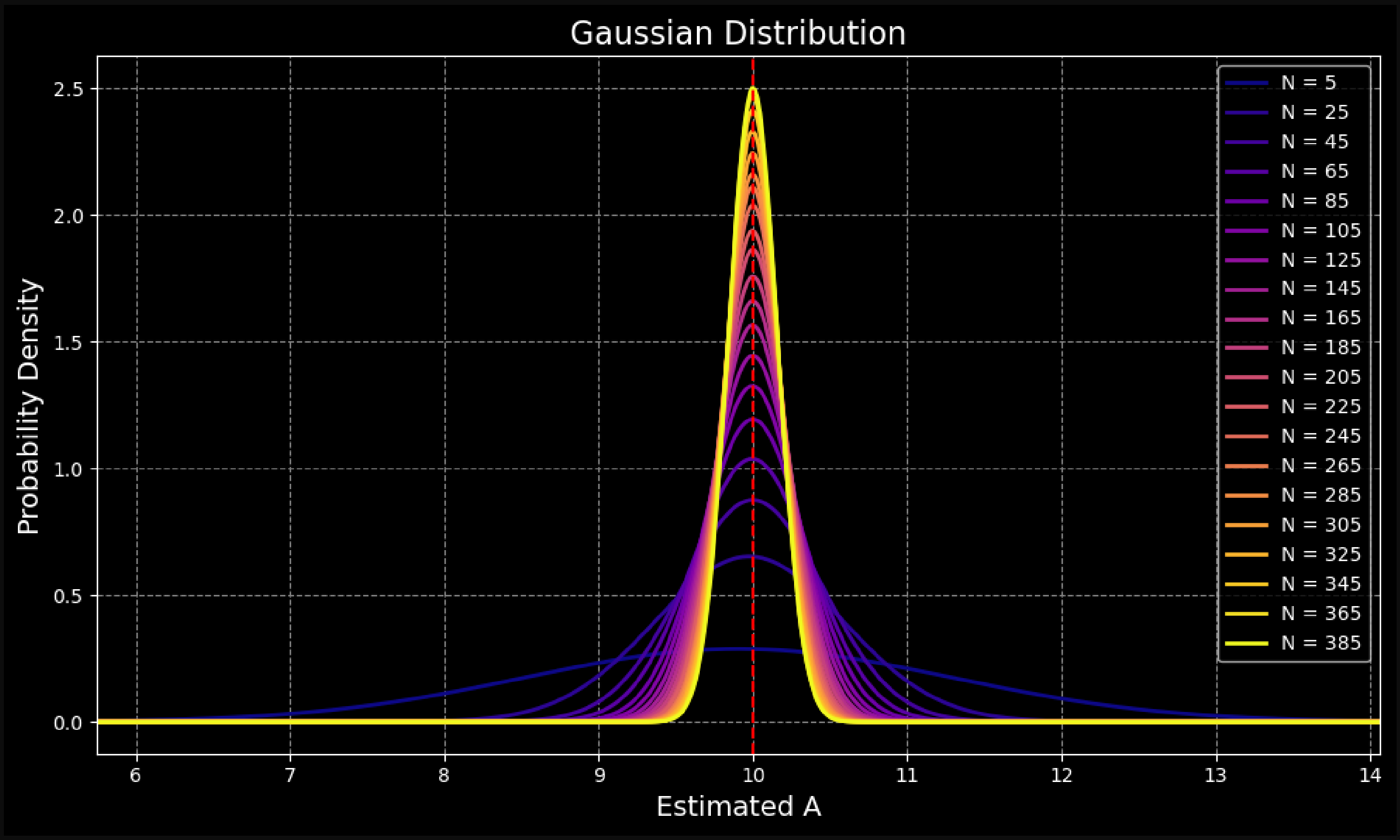
The **Mean Bias Plot** demonstrates that the bias in the estimates decreases as the sample size increases. This is expected because larger sample sizes reduce the uncertainty in the estimates.

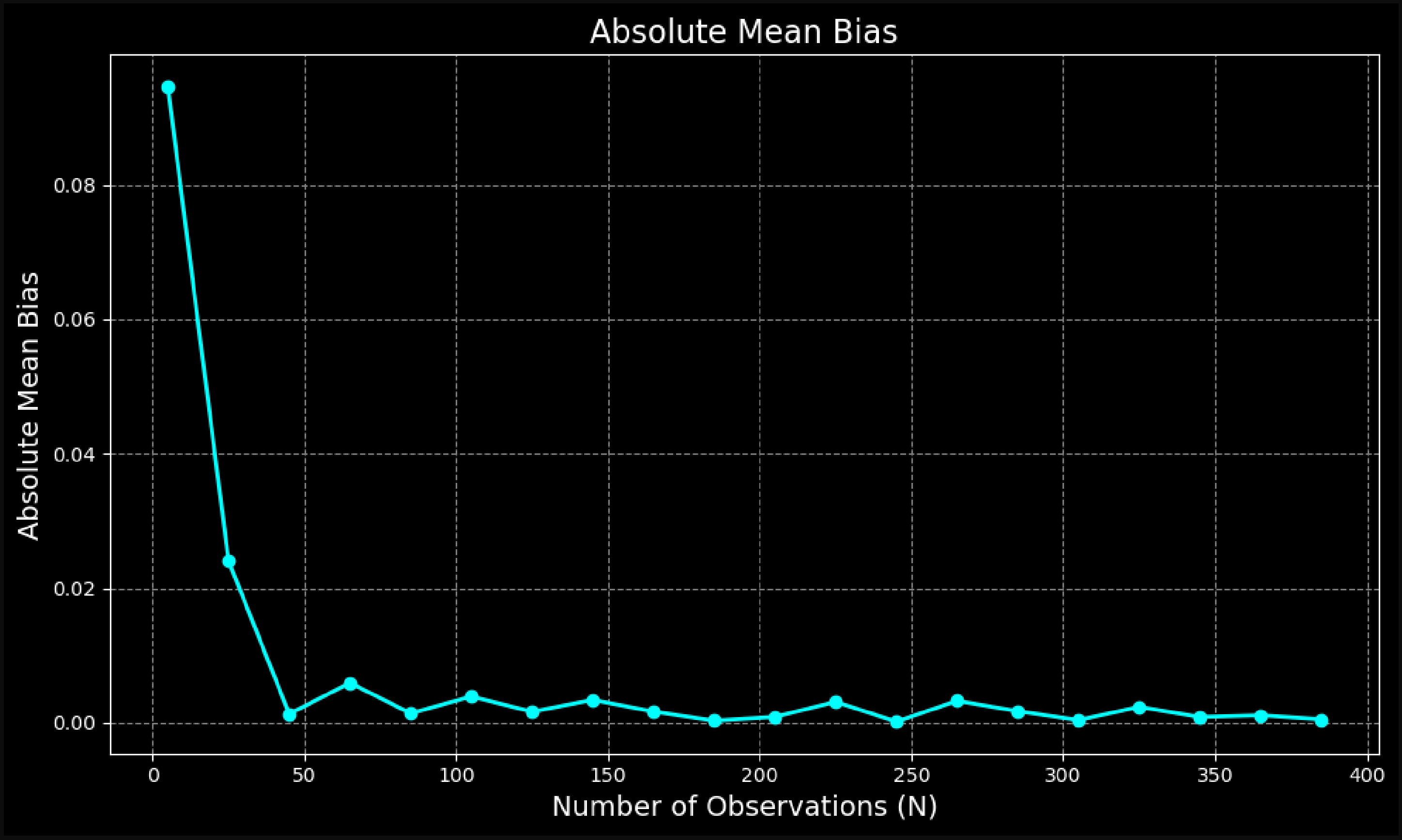
Effect of Noise on Estimation:

The noise standard deviation was varied by changing the value of A, since the noise is proportional to A. The results show that as the noise level increases, the spread of the estimated values also increases. The estimates become less reliable and more spread out around the true value, particularly for small N.

The higher noise levels lead to greater uncertainty in the estimation, which is evident in the broader Gaussian distributions.

Plots


Gaussian Distribution
Probability Density
Estimated A
2.5
2.0
1.5
1.0
0.5
0.0
6
7
8
9
10
11
12
13
14
N = 5
N = 25
N = 45
N = 65
N = 85
N = 105
N = 125
N = 145
N = 165
N = 185
N = 205
N = 225
N = 245
N = 265
N = 285
N = 305
N = 325
N = 345
N = 365
N = 385


Absolute Mean Bias
Absolute Mean Bias
Number of Observations (N)
0.08
0.06
0.04
0.02
0.00
0
50
100
150
200
250
300
350
400

Discussion

Impact of Sample Size:

As expected, the accuracy of the estimates improves with larger sample sizes. This is a well-known result in statistics, where increasing the number of observations leads to more precise estimates. The reduction in bias with increasing N reflects the central limit theorem, where the estimator becomes more concentrated around the true value as more data is collected.

The plots clearly show that for smaller values of NNN, the estimates have a larger spread and a higher bias. As N increases, the bias decreases, and the estimates become more stable and concentrated around the true parameter.

Impact of Noise:

The noise in the data significantly affects the accuracy of the estimates. For higher noise levels, the estimates are more spread out, reflecting the increased uncertainty. This highlights the importance of controlling noise in practical applications, as noise reduces the reliability of parameter estimation.

Precision and Variability:

The standard deviation of the estimates is an important measure of precision. The results demonstrate that with larger sample sizes, the precision of the estimates improves, as evidenced by the narrower Gaussian distributions. Smaller sample sizes lead to higher variability in the estimates, as shown by the broader distributions.

Conclusion

This experiment demonstrates the application of **Least Squares Estimation** for parameter estimation in a noisy system. By simulating the estimation process for different sample sizes and noise levels, we can see how the accuracy and reliability of the estimates are affected by these factors. The results show that larger sample sizes lead to more precise and accurate estimates, while higher noise levels increase the spread and bias of the estimates.

Key findings include:

Increasing sample size improves the accuracy and stability of the estimates.

Higher noise levels result in less accurate estimates, with increased spread and bias.

The **bias** in the estimates decreases with larger sample sizes, and the **standard deviation** becomes smaller as the number of observations increases.