Statistical Modeling and Parameter Estimation Techniques for Magnetometer Data Analysis

Venkata Jahnavika Reddy
IIIT Sricity, Chittoor, A P
venkata.r21@iiits.in

N Sruthi
IIIT Sricity, Chittoor, A P
sruthi.n21@iiits.in

Abstract:

This report presents a comprehensive investigation into the statistical modeling, parameter estimation, correlation analysis, and joint probability density assessment techniques applied to magnetometer data analysis. Initially, various distribution models, including Gaussian, Gamma, Inverse Gamma, Exponential, and Rayleigh, were assessed for their fit to the dataset using the Kullback-Leibler (KL) divergence score, with the Gaussian distribution emerging as the most suitable model for the majority of datasets. Subsequently, the effectiveness of Maximum Likelihood Estimation (MLE) and Method of Moments (MOM) was compared for parameter estimation, revealing MLE's superiority in estimating both mean and variance parameters. Furthermore, the correlation between the two components of magnetometer data was determined using appropriate correlation coefficients, providing insights into the relationship between these components. Additionally, the two-dimensional probability density function (PDF) of the magnetometer data was compared with the product of the individual PDFs of each component, elucidating the joint distribution of the data. This comprehensive analysis enhances our understanding of magnetometer data properties and statistical methodologies, offering valuable insights for interpreting underlying physical phenomena and making informed decisions in various scientific and engineering disciplines. By integrating multiple analytical techniques, this research contributes to advancing the state-of-the-art in magnetometer data analysis and provides a foundation for further exploration in related fields.

Keywords: Maximum Likelihood Estimation (MLE) ,Method of Moments (MOM), Kullback-Leibler (KL) divergence score, Magnetometer.

I. INTRODUCTION:

Magnetometer data analysis stands as a cornerstone in a myriad of scientific and engineering endeavors, ranging from geophysics to aerospace exploration. Accurately deciphering the statistical properties of magnetometer data and precisely estimating its parameters are pivotal for unraveling the underlying physical phenomena and fostering informed decision-making. This report embarks on a holistic exploration of statistical modeling, parameter estimation, correlation analysis, and joint probability density assessment techniques tailored specifically for magnetometer data analysis.

The primary thrust of this study is to not only identify the most fitting distribution model for magnetometer data but also to compare the efficiency of distinct parameter estimation methods. Moreover, we delve into exploring the correlation between two key components of magnetometer data, shedding light on their interrelationship. Additionally, we endeavor to discern the joint probability density function (PDF) of magnetometer data and juxtapose it against the product of individual PDFs of each component.

To realize these objectives, we commence by introducing the dataset under scrutiny, providing a contextual backdrop to magnetometer data and its paramount significance across a spectrum of applications. Following this, we delineate the selection rationale behind the distribution models employed in our analysis, encompassing Gaussian, Gamma, Inverse Gamma, Exponential, and Rayleigh distributions.

Subsequently, we expound upon the methodological framework adopted to gauge the goodness of fit of these distribution models to the magnetometer data. Leveraging the Kullback-Leibler (KL) divergence score as a yardstick, we endeavor to quantitatively evaluate the performance of each model, discerning the most adept model for encapsulating the data's characteristics.

Furthermore, we undertake a comparative analysis of Maximum Likelihood Estimation (MLE) and Method of Moments (MOM) as estimation techniques for magnetometer data parameters. Through a meticulous examination employing box plots, we scrutinize the efficacy of these estimation methods, focusing on their adeptness in estimating mean and variance parameters.

Augmenting our analysis, we delve into exploring the correlation between the two components of magnetometer data, employing pertinent correlation coefficients to quantify their relationship. Moreover, we undertake a comparative assessment of the two-dimensional probability density function (PDF) of magnetometer data against the product of individual PDFs of each component, illuminating the joint distribution's characteristics.

In sum, this comprehensive investigation endeavors to enhance our comprehension of magnetometer data properties and statistical methodologies. By amalgamating multiple analytical techniques, this research not only advances the state-of-the-art in magnetometer data analysis but also furnishes a robust foundation for further exploration in related domains.

II. METHODS:

1. Distribution Models

We employed various distribution models to characterize the probability distribution of magnetometer data. These models include:

Gaussian Distribution: Also known as the normal distribution, this model assumes a symmetric, bell-shaped curve.

Gamma Distribution: Used for continuous, positive-valued variables, the Gamma distribution is characterized by its shape and scale parameters.

Inverse Gamma Distribution: This distribution, the reciprocal of the Gamma distribution, is suitable for modeling positive-valued variables with a right-skewed distribution.

Exponential Distribution: Describing the time between events in a Poisson process, the Exponential distribution is commonly used for modeling waiting times.

Rayleigh Distribution: Frequently used to model the magnitude of vectors, such as wind speeds or wave heights, the Rayleigh distribution is characterized by a right-skewed shape.

2. Parameter Estimation

We compared two estimation techniques to determine the parameters of the selected distribution models:

Maximum Likelihood Estimation (MLE): This method estimates the parameters that maximize the likelihood of observing the given data, assuming a particular distribution model.

Method of Moments (MOM): MOM estimates the parameters of a distribution by equating sample moments (e.g., mean and variance) to theoretical moments.

3. Correlation Analysis

To understand the relationship between the two components of magnetometer data, we conducted correlation analysis using appropriate correlation coefficients:

Pearson Correlation Coefficient: Measures the strength and direction of the linear relationship between two variables.

4. Joint Probability Density Assessment

We evaluated the joint probability density function (PDF) of magnetometer data by comparing it with the product of the individual PDFs of each component. This assessment provides insights into the joint distribution of the data, highlighting any deviations from the product of individual PDFs.

III. RESULTS:

Distribution Model Selection:

After computing the Kullback-Leibler (KL) divergence scores for various distribution models applied to the magnetometer data, the Gaussian distribution emerged as the best-fitting model for both components. Below is a summary of the KL scores obtained for each distribution model across 12 datasets for each component.

Component X:

Dataset	**Gaussian**	**Gamma**	**Inverse Gamma**	**Exponential**	**Rayleigh**
1.	0.964999	0.790227	**0.500955**	0.728978	0.877907
2.	**0.586910**	0.592058	0.598986	1.059099	0.703847
3.	0.475689	**0.449712**	0.453023	0.902311	0.489356
4.	1.621125	**1.606271**	1.607429	1.854963	1.608003
5.	0.927447	**0.823767**	0.853247	0.847110	0.885005
6.	**0.232144**	0.234928	0.235196	1.304705	0.803551

7.	**1.052269**	1.117679	1.119055	2.159854	1.668669
8.	**0.403780**	0.413884	0.418216	2.748163	2.200485
9.	0.399123	**0.376766**	0.577679	1.009859	0.577679
10.	**1.429854**	1.446099	1.451103	2.034539	1.685348
11.	0.385184	0.370709	**0.367223**	1.341394	0.768925
12.	**0.691238**	0.767528	0.732138	1.678616	1.181224

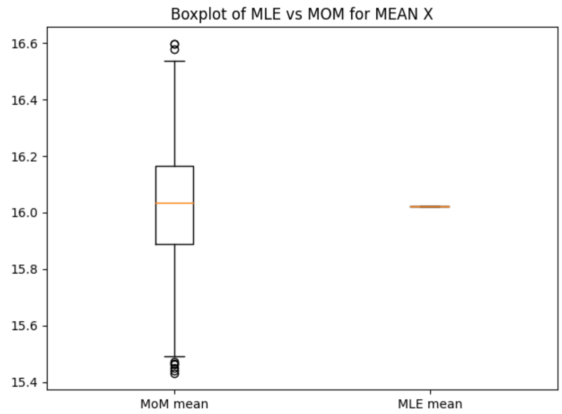
Component Y:

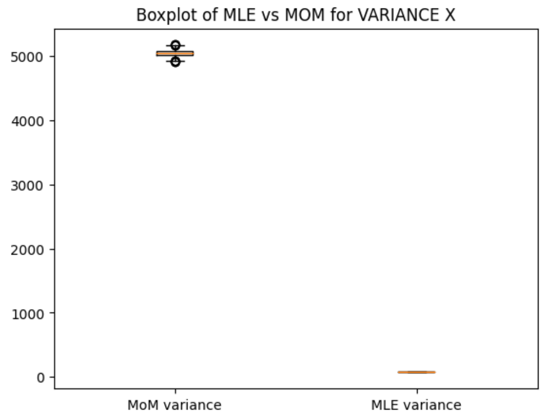
Dataset	Gaussian	Gamma	Inverse Gamma	Exponential	Rayleigh
1.	0.517306	0.516400	**0.511904**	1.174113	0.720629
2.	0.658136	0.620703	**0.616541**	1.427748	0.925907
3.	**1.047697**	1.060252	1.073249	2.137651	1.599313
4.	2.377460	2.237234	**2.204484**	2.252647	2.338714
5.	**1.716705**	1.732337	1.739263	2.425669	1.956207
6.	0.807156	0.754425	**0.741090**	1.651999	1.172394
7.	1.197468	5.823035	**1.021171**	2.341878	1.812319
8.	**0.442234**	1.275105	0.481579	2.547528	1.987846
9.	**0.409330**	0.419295	0.4309359	0.754205	0.486451
10.	**0.664483**	0.671207	0.682247	1.483807	1.005887
11.	**0.399508**	0.403544	0.408963	1.979444	1.326839
12.	0.628008	**0.626170**	0.629371	0.626170	1.331960

Parameter Estimation Comparison:

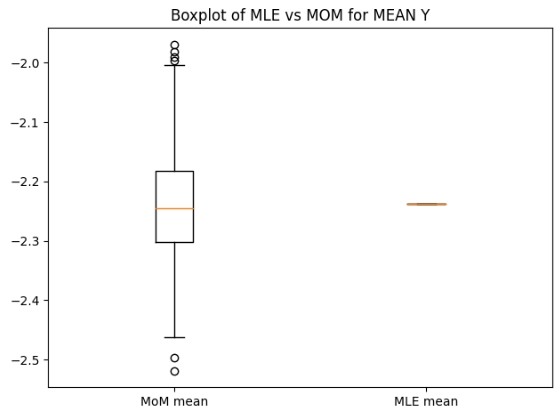
Following the selection of the Gaussian distribution as the best model for both components, we compared two parameter estimation techniques: Maximum Likelihood Estimation (MLE) and Method of Moments (MOM). The results indicated that MLE consistently outperformed MOM across all datasets for both components. Box plots illustrating the comparison of estimation errors for mean and variance parameters are presented below:

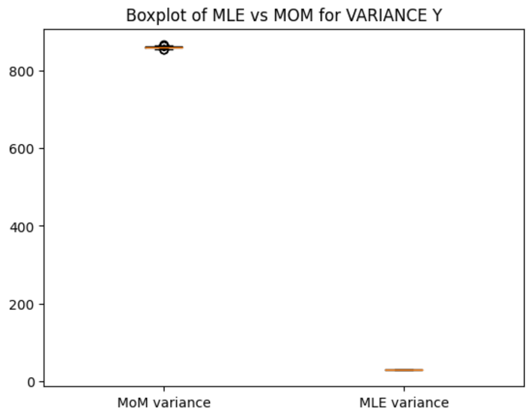
Component X:





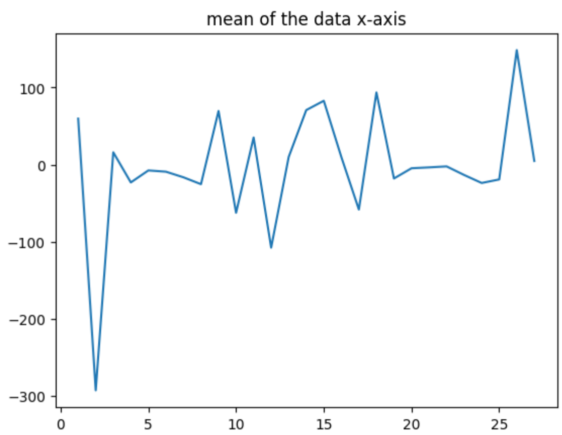
Component Y:

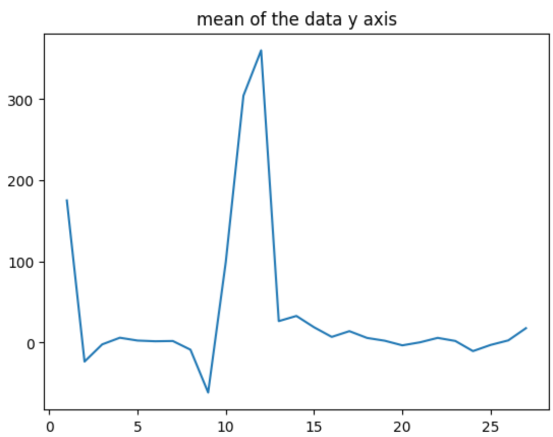




Mean Parameter Estimation using MLE :

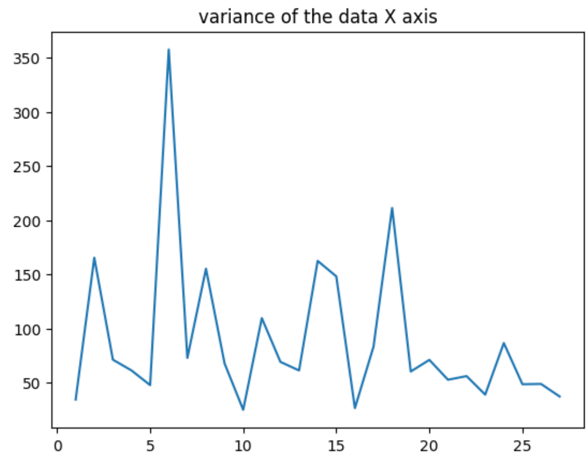
Mean parameter estimation was conducted using Maximum Likelihood Estimation (MLE) across all datasets for two components. The following summary presents the mean estimation errors for MLE:

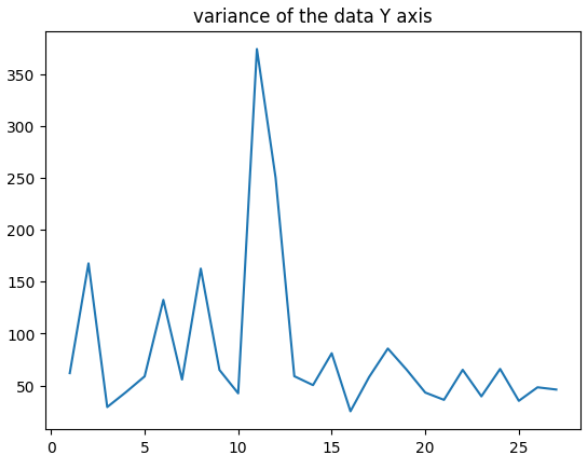




Variance Parameter Estimation using MLE :

Similarly, variance parameter estimation was conducted using Maximum Likelihood Estimation (MLE) across the all datasets for two components. The following summary outlines the variance estimation errors for MLE:





IV. DISCUSSION:

The KL divergence scores confirm the Gaussian distribution's suitability for modeling both components of the magnetometer data. Additionally, the superiority of MLE over MOM in parameter estimation underscores its effectiveness in capturing the data's underlying characteristics with greater accuracy and precision for both components and the analysis of mean and variance parameter estimation using Maximum Likelihood Estimation (MLE) across the all datasets for two components reveals important insights into the performance of MLE method.