Review of Maximum Likelihood Estimation: Methods, Applications, and Biases

José Martín Véliz Zambrano.
Tecnológico de Monterrey, Querétaro, México.

Abstract

Maximum Likelihood Estimation (MLE) is a key statistical method involving parameter estimation. The focus of this review is on papers discussing the theoretical and practical aspects of MLE including its principles, advantages and disadvantages, application in different settings such as logistic regression and forecasting, and bias reduction in its estimates. The paper concludes by highlighting MLE's diverse applications, potential bias issues, and suggestions for future research in the field.

1. Introduction

Maximizing likelihood estimation (MLE), a significant method in parameter estimation for statistics boasts a wide array of applications fundamental to the field. These applications range from serving as the backbone for hypothesis testing to playing a crucial role in shaping confidence intervals. At its core, MLE is an optimization algorithm that seeks to find the parameter values which make the observed data most likely. These estimated parameters can then be used for effective data modelling and prediction.

However, despite its prevalent use and significant contribution, MLE possesses inherent limitations that users must be wary of. From a theoretical standpoint, MLE does not always guarantee a unique or existing set of parameter values, a fact that can complicate the understanding and interpretation of results. This brings into focus the critical need for discernment and proper statistical rigour when utilizing MLE since blindly following the principles of MLE might result in misleading outcomes, hence the caution required when dealing with results gleaned from its application.

2. Elaboration of the Theory and Principles of MLE

In his seminal work, Myung[3] delves into the intricacies underpinning the Maximum Likelihood Estimation (MLE). A standout concept elucidated is that of the Probability Density Function (PDF), a mathematical function that provides the probabilities of occurrence of different possible outcomes. Alongside this, Myung[3] also explores the functions indexed by the model's parameters, focusing on how they can affect potential data outcomes. It is these parameters that MLE seeks to estimate by maximising the probability of the observed data, given the estimated parameters, a process also known as maximising the likelihood function.

However, a challenge arises with the use of MLE - it may not always yield a unique or even an existing set of parameter values. This inability signifies that the process of MLE doesn't always guarantee a convergent, optimised maximisation of the likelihood function. Furthermore, this possibility brings to light the complexity of the application of MLE, and advises caution for researchers, noting that a superior or optimum fit should not be misinterpreted as a definite sign of model correctness[3].

3. MLE's role in Logistic Regression and Forecasting

Czepiel's research[1] underlines the pre-eminence of MLE in estimating parameters in logistic regression models. This study reveals that MLE takes precedence over the least-squares estimation, strengthening its position as a preferred estimation method in logistic regression. A notable element in this process is the deployment of the Newton-Raphson method, a numerical technique that solves systems of nonlinear equations effectively, thereby aiding in parameter estimation.

Further supplementing this understanding, the work of Sura & Candès[4] provides critical insights into the manifestation of bias, and the behaviour of the likelihood-ratio test in high-dimensional logistic regression models when using MLE. Their findings flag the need for an improved understanding and handling of these nuances when implementing MLE in complex modelling scenarios.

Moreover, the usage of MLE extends into the realm of regression and forecasting problems, where it stands superior in performance. As Awasthi et al[5] affirm, MLE showcases its versatility by offering better excess risk bounds, aptly capturing prior knowledge and allowing optimization of various target metrics during the forecasting process.

4. Decoding Bias Reduction in MLE

A groundbreaking contribution to the field has been through the work of Firth[2], placing a spotlight on bias reduction strategies when using MLE in regular parametric problems. Firth[2] devises a new method that works to modify the score function to ameliorate bias, thereby proposing a way to use MLE without falling into bias-related pitfalls. This method offers a novel approach to reducing bias, which can greatly augment the utility of MLE.

Nonetheless, it's crucial to note that its application isn't universally viable, and it can demand a significant computational toll. As such, there's apparent room for future research in exploring methods for bias reduction that balance computational efficiency and broad usability. The interplay between bias reduction, computational demand, and ease of use highlights a fertile ground for potential developments in the field.

5. Conclusion

MLE is a powerful tool in statistics but also complex enough to warrant exercising caution in interpreting results. While it’s suitable for logistic regression models and superior in regression and forecasting tasks, its potential bias in specific settings requires further attention. Future research should focus on how to further reduce bias in MLE, ensure replicability, and enhance its applications using computation-friendly methods.

6. References

[1]Czepiel, S. A. (). Maximum Likelihood Estimation of Logistic Regression Models: Theory and Implementation.

[2]Firth, D. (). Bias reduction of maximum likelihood estimates.

[3]Myung, I.J. (2003). Tutorial on maximum likelihood estimation. Journal of Mathematical Psychology, 47(1), 90-100.

[4]Sura, P., & Candès, E. J. (2018). A modern maximum-likelihood theory for high-dimensional logistic regression.

[5]Awasthi, P., Das, A., Sen, R., & Suresh, A. T.(). On the benefits of maximum likelihood estimation for Regression and Forecasting. Google Research.