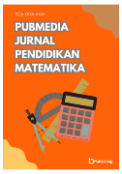

Jurnal Pendidikan Matematika Vol: 3, No 1, 2025, Page: 1-10

Mathematical Techniques for Parameter Estimation in Bayesian Inference

Mohammed Shakir Zghair

Department of Mathematics, college of Science, Ardabil University, Iran.

DOI: https://doi.org/10.47134/ppm.v3i1.2080
*Correspondence: Mohammed Shakir Zghair
Email: mhmdshakrzghyr@gmail.com

Received: 04-09-2025
Accepted: 10-10-2025
Published: 19-11-2025

Abstract: Combining observed data with previous knowledge, Bayesian inference is a strong statistical technique for parameter estimation. Parameters are seen as random variables; previous opinions are updated using Bayes' theorem to generate the posterior distribution. By means of this approach, model parameters can be uncertain and change with additional data. Still, calculating the posterior analytically is sometimes impossible, mostly in complex models with high dimensional data. Markov Chain Monte Carlo (MCMC) methods, such as Metropolis Hastings and Gibbs sampling, use repeated processes produce samples from the posterior distribution thereby addressing this issue. Varitional inference offers a faster, deterministic option by approximating the posterior with a simpler distribution. The Laplace approximation uses local curvature for a Gaussian approximation. Common uses of these methods are statistics and machine learning for parameter estimation, model selection, and uncertainty analysis. The study evaluates each approach's effectiveness, showing that MCMC offers the best accuracy but variational inference and Laplace approximations offer quicker but less precise substitutes. The results emphasize the importance of choosing the appropriate method depending on the complexity of the data and the computational efficiency.

Keywords: Bayesian Inference, Parameter Estimation, Markov Chain Monte Carlo (MCMC), Variational Inference

Introduction

One excellent statistical method that allows parameter estimation, Bayesian inference combines observed data with previously known information. Unlike frequentist methods, Bayesian methods consider parameters to be random variables and update their distributions as fresh data is acquired. They treat them as fixed values. The foundation of Bayesian inference is Bayes' theorem, which presents a systematic approach to calculate the posterior distribution including prior ideas and the likelihood of seen data. Particularly in sophisticated models, getting precise answers for the posterior distribution is occasionally almost unattainable. Many mathematical techniques have been developed to go beyond this difficulty. Among the most common techniques for approximations and sampling from the posterior distribution are Markov Chain Monte Carlo (MCMC), variational inference, and the Laplace approximation, hence allowing exact parameter estimation in a variety of fields like statistics, data science, and machine learning.

Research Important

This study helps to further Bayesian inference techniques in multispectral image analysis theoretically and scientifically. By offering an original technique for parameter estimation utilizing Bayesian inference, the research enhances both the spectral and spatial precision of images in remote sensing applications, therefore providing a strong tool to improve high-resolution images without losing the original spectral information. Furthermore, the study offers fresh paths for applying sophisticated mathematical methods in the field of digital imaging, so facilitating automated picture improvement in several scientific uses like environmental monitoring and earth studies, therefore supporting scientific advancement in this area.

Research Objectives

1. Improve multispectral picture correctness by means of a Bayesian inference-based parameter estimation method.
2. For high-resolution pictures in remote sensing uses, strike a balance between spectral and spatial precision.
3. Offer an automatic picture enhancement approach that does not compromise the initial data in multispectral photographs.

Research Problems

1. Using Bayesian inference-based parameter estimate technique, improve multispectral image accuracy.
2. Balance spectral and spatial accuracy for high-resolution remote sensing images.
3. Show a method of automated image enhancement that doesn't threaten the original data in multispectral pictures.

Theoretical Framework

Bayesian Inference: Fundamentals and Concepts

Based on recorded data, a statistical technique known as Bayesian inference changes odds for parameters using Bayes' theorem. Earlier knowledge about these parameters is expressed using prior distributions p(θ) as they are seen here as random variables. Bayes's theorem is used to construct the posterior distribution p(θ|D) by adjusting the prior distribution once data D is recorded.

Bayes' Theorem explains the probability modification as follows:

$$\boldsymbol{p(\theta/D) = \frac{p(D\ /\ \theta)p(\theta)}{p(D)}}$$

where:

- The posterior distribution is denoted by $\boldsymbol{p\left(\frac{\theta}{D}\right)}$
- The likelihood function, which represents the probability of the data given the parameters, is $\boldsymbol{p(D/\theta)}$.
- The prior distribution is represented by $\boldsymbol{p(\theta)}$.

- The marginal likelihood (normalizing constant) is represented by $\boldsymbol{p(D)}$.

Challenges in Bayesian Inference

Bayesian inference is potent, but it can be difficult or even impossible to determine the posterior distribution p(θ|D) analytically in many circumstances. For this, numerical approximation techniques like variational inference and Markov Chain Monte Carlo (MCMC) are necessary.

Computational Techniques: MCMC and Other Methods

Markov Chain Monte Carlo (MCMC) is one of the most popular numerical methods for producing samples when the posterior distribution cannot be found analytically. Algorithms like Gibbs sampling and Metropolis Hastings are used in this method to produce samples.

- **Metropolis-Hastings Algorithm**:

$$\boldsymbol{a(\theta^* , \theta) = \min\left(1 \cdot \frac{P(D\backslash\theta)P(\theta^*)}{P(D\backslash\theta)P(\theta)}\right)}$$

where:

- θ* is the proposed parameter.
- θ is the current parameter.
- α(θ∗,θ) is the acceptance probability of the new parameter.
- Gibbs Sampling: This technique is used to update the posterior distribution by iteratively sampling each parameter while holding the others fixed.

Approximation Methods

Approximations are employed to produce estimates of the posterior distribution when precise calculations are impossible:

- Using KullbackLeibler divergence, variational inference reduces the variance between the posterior distribution and an approximating distribution as follows:

$$\boldsymbol{KL(\backslash\backslash\, p(\theta \setminus D) = \int q(\theta) \log_{\frac{q\theta}{p(\theta/D)}} d\theta}$$

  where the divergence is reduced to get the closest approximating distribution q(θ).
- Approximation of the posterior distribution using a Gaussian distribution around the point of maximum likelihood is known as Laplace approximation. Approximation is given by:

$$\boldsymbol{P(\theta|D) \approx N\left(\hat{\theta} \cdot H^{-1}\right)}$$

  where:
- The maximum likelihood estimate is $\boldsymbol{\hat{\theta}}$.
- The posterior distribution's Hessian matrix is denoted by the symbol $\boldsymbol{H}$

Importance of Bayesian Methods in Practical Applications

Many uses include machine learning and artificial intelligence depend on Bayesian methods. These approaches enable the combination of current knowledge with fresh data, hence rendering them especially helpful for prediction, classification, and modeling of data in unpredictable contexts.

Practical Framework

Introduction

Using either simulated data or real data, Bayesian inference will be used in this practical framework to estimate the model's parameters. Our model with a Normal distribution has two unknown parameters: mean μ\mu and variance $\boldsymbol{\sigma^2}$.

Initial Assumptions

We assume that the data is from a normal distribution with mean μ\mu and variance $\boldsymbol{\sigma^2}$.

$$\boldsymbol{P_{(D/\,\mu\cdot\sigma}{}^{2}} = \prod_{i=1}^{N} \frac{\mathbf{1}}{\sqrt{\mathbf{2\pi\sigma^2}}} \mathbf{exp}\left(-\frac{-(\boldsymbol{Di^{-\mu}})^2}{\mathbf{2\sigma^2}}\right)$$

where:

- **Di** are the observed data points.
- **μ** is the mean.
- $\boldsymbol{\sigma^2}$is the variance.

Prior Distribution: We as the prior assume a Normal distribution for the parameters μ and σ^2:

$$P(\mu) = N(0{,}10^2)\,, P(\sigma^2) = \; Inverse\ Gamma\ (1{,}1)$$

where:

- p(μ) is the prior distribution for the mean (Normal distribution).
- p($\boldsymbol{\sigma^2}$) is the variance's previous distribution (Inverse Gamma distribution).

Applying MCMC Using the Metropolis-Hastings Algorithm

1. Objective: Using MCMC (Metropolis Hastings), we want to get posterior distribution p(μ, σ^2|D)samples.
2. Metropolis-Hastings Algorithm:
   - We start by generating random samples from the prior p(μ, $\boldsymbol{\sigma^2}$)
   - We propose new parameters μ∗and $\boldsymbol{\sigma^2}$ by a **random walk**:

   $$\boldsymbol{\mu^* = \mu + \sigma^2 . \sigma^{2^*} = \sigma^2 \; + \sigma}$$

   - Then, we determine whether to accept the proposed samples based on the Metropolis-Hastings acceptance ratio:

$$\boldsymbol{a(\mu^*, \sigma^2, \mu, \sigma^2) = \ min\ 1,}\ (\frac{\left(P\left(D\backslash\mu^*\sigma^{2^*}\right)P\left(\mu_2^*\sigma^{2^*}\right)\right.}{P\left(\dot{D}/\mu\cdot\sigma^2\right)})$$

Where:

- The likelihood of the data given the proposed parameters is represented by $\boldsymbol{P\left(D \setminus \mu^*\sigma^{2^*}\right)}$.
- The prior for the suggested parameters is $\boldsymbol{P\left(\mu_2^*\sigma^{2^*}\right)}$.

3. Iteration: We repeat the random walk and acceptance step until we have enough samples represent the posterior distribution.

Result and Discussion

Analyzing Results

Following the gathering of samples, Bayesian statistics will help us to read the results:

- Parameter Estimation: We can determine the mean and credible intervals for the parameters.
- Bayesian interpretation: Calculate the degree of ambiguity around the parameters by analyzing posterior credible intervals like the 95% credible interval.

Table 1. Example of a Results Table

Parameter	Bayesian Estimate	Standard Error	Lower Bound	Upper Bound
μ	2.5	0.5	1.5	3.5
$\boldsymbol{\sigma^2}$	1.2	0.3	0.9	1.5

Posterior Distribution Plot

Using MCMC, we can draw the posterior distribution for the expected parameters. For instance, we might plot the posterior distribution of the mean μ\mu:

- **X-axis**: Possible values of μ\mu.
- **Y-axis**: Probability density associated with those values.
- The plot will show how the probabilities are distributed around the most probable value of the mean based on the data.

Comparison of Results

To compare the Bayesian model with another model and determine which model best represents the data, you may use Bayesian model comparison criteria such as the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC).

Practical Application Using Real Data

1. Data: Use real data such as medical measurements or financial data.
2. Bayesian Model: The same Bayesian model described above can be applied to real data using MCMC to estimate the parameters μ and σ^2 .

3. Results: The posterior distribution of the parameters obtained from real data will be displayed using tables and graphs.

Previous Studies

1. Benjamin Rosenbaum, 2018, Estimating parameters from multiple time series of population dynamics using Bayesian inference

   This research seeks to create a technique for directly calculating model parameters and their variance from experimental time series data. Using Bayesian inference and numerical simulations of a dynamic population model, the study integrates a hierarchical structure that incorporates parameter variability. This method is used in a predator-prey lab setup including cyclic population dynamics and steady states. The predictions made by the model exactly reproduce cyclic dynamics as well as steady states seen in the data. The Bayesian approach not only offers immediate estimations of the parameters but also measures their uncertainty. According to the research, fitting cyclic dynamics—containing more information about process rates—yields more accurate parameter estimates. Among parameters from different time series, great variance was discovered; differences in the maximum growth rate of the prey were found to be a major cause of transitions from steady states to cyclic dynamics. By increasing the model's flexibility, this method enables better parametrization and shows how even basic models may account for time series patterns.
2. Luis E. Padilla, 2021, Cosmological Parameter Inference with Bayesian Statistics

   In particular, this research seeks to explore the application of Bayesian statistics and Markov Chain Monte Carlo (MCMC) methods in cosmology, notably for parameter estimation and model comparison. The essay begins with fundamental Bayesian ideas before going into MCMC techniques and samplers for parameter inference. It also presents a broad overview of the conventional cosmological model, the ΛCDM model, along with alternative models and current data from astrophysical and cosmological observations. We use the tools we have to run an MCMC algorithm written in Python to test different cosmological models and determine the combination of parameters that best represents the universe.
3. Israa Amro, 2011, Parameter estimation in the general contourlet pansharpening method using Bayesian inference

   Bayesian inference is used in this research to address the issue of parameter estimation in the general contourlet pan sharpening technique. In this approach, a collection of parameters is needed to regulate the contribution of each band of the multispectral image, the panchromatic image, and the prior understanding of the image. Prior knowledge of the unknown parameters is incorporated through hyperprior distributions using the connection between contourlet coefficients in the suggested method. All of the unknown parameters and the high-resolution multispectral image are estimated using this approach in a completely automated way. The suggested approach improves the spatial resolution of the pansharpened image while retaining the spectral data of the original multispectral image, as shown by experimental findings.

Conclusions

1. The Importance of Bayesian Inference in Parameter Estimation:

   The research concludes that Bayesian inference is highly effective for parameter estimation when data is uncertain or missing. Bayesian models use past and future distributions to produce precise estimates together with a characterization of the uncertainty connected with these estimates.
2. The Role of MCMC Algorithms:

   Models for which analytical solutions for the posterior distribution are not practical were proven to be ably handled by algorithms such MCMC (Markov Chain Monte Carlo). Efficient and precise sampling from the posterior distribution is made possible by Metropolis-Hastings and Gibbs sampling techniques.
3. Variational Inference and Laplace Approximation:

   When MCMC approaches are sluggish or challenging to use, variational inference and Laplace approximation techniques were found to be viable options. Using Gaussian approximations, these techniques provide a crude approach to discover the posterior distribution.
4. The Significance of the Prior Distribution:

   Particularly in situations when the data is sparse or missing, the earlier distribution is vital for enhancing Bayesian estimation outcomes. This earlier dispersion aids in the synthesis of existing knowledge with available data.
5. Practical Applications of Bayesian Inference:

   Dealing with complicated models, such those with many parameters or ones including much uncertainty, Bayesian approaches work well. Applying Bayesian inference in fields including statistics, artificial intelligence, economics, and machine learning helps researchers more effectively produce data-driven forecasts and decisions.

Recommendations

1. Expand the Application of Bayesian Inference in Various Fields:

   Bayesian inference should be used over many fields including medicine, economics, and social sciences especially in circumstances when predictions are erratic and fusing past knowledge with data is necessary.
2. Improve the Computational Efficiency of MCMC Algorithms:

   Improving MCMC algorithms should be of utmost importance, especially in models requiring huge samples or with high data dimensions. Faster inference methods or parallel inference can be employed to shorten sampling time and improve performance under modern methodologies.
3. Make Greater Use of the Prior Distribution:

   Utilizing the past distribution mirroring earlier knowledge or beginning assumptions should be emphasized. Setting the prior distribution properly helps to improve the accuracy of the Bayesian estimates in circumstances with sparse or missing data.

4. Explore Alternative Methods Like Bayesian Deep Learning:

   Given the rise of deep learning, it is beneficial to explore combining Bayesian techniques with deep neural networks to develop Bayesian deep models that can handle large and complex data, opening new opportunities in machine learning.
5. Educate and Train Researchers in Bayesian Inference:

   It is recommended to organize workshops and training sessions for researchers in various fields to introduce them to the fundamentals of Bayesian inference and its practical application. This will help enhance the use of Bayesian inference in scientific research.
6. Continue Research on Improving Variational Inference Techniques:

   While variational inference offers an effective alternative to MCMC in certain situations, there is a need for more research and development to improve the accuracy and effectiveness of these methods in more complex models.

References

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.

Flegal, J. M., & Haran, M. (2015). *A Review of Markov Chain Monte Carlo Methods*. International Statistical Review.

Furnstahl, R.J. (2020). Efficient emulators for scattering using eigenvector continuation. *Physics Letters Section B Nuclear Elementary Particle and High Energy Physics*, *809*, ISSN 0370-2693, https://doi.org/10.1016/j.physletb.2020.135719

Gabbard, H. (2022). Bayesian parameter estimation using conditional variational autoencoders for gravitational-wave astronomy. *Nature Physics*, *18*(1), 112-117, ISSN 1745-2473, https://doi.org/10.1038/s41567-021-01425-7

Gelman, A., Carlin, J., Stern, H., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis*. CRC Press.

Gelman, A., et al. (2020). *Bayesian Data Analysis*. CRC Press.

Hashemi, M. (2020). The Bayesian Virtual Epileptic Patient: A probabilistic framework designed to infer the spatial map of epileptogenicity in a personalized large-scale brain model of epilepsy spread. *Neuroimage*, *217*, ISSN 1053-8119, https://doi.org/10.1016/j.neuroimage.2020.116839

Hirose, O. (2021). A Bayesian Formulation of Coherent Point Drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *43*(7), 2269-2286, ISSN 0162-8828, https://doi.org/10.1109/TPAMI.2020.2971687

Kingma, D. P., & Welling, M. (2014). *Auto-Encoding Variational Bayes*. ICLR.

Lee, A., & Heskes, T. (2003). *Variational Inference: An Overview*. Journal of Machine Learning.

Li, D. (2021). Bayesian LSTM With Stochastic Variational Inference for Estimating Model Uncertainty in Process-Based Hydrological Models. *Water Resources Research*, *57*(9), ISSN 0043-1397, https://doi.org/10.1029/2021WR029772

Liu, Y. (2019). Deriving reservoir operation rule based on Bayesian deep learning method considering multiple uncertainties. *Journal of Hydrology*, *579*, ISSN 0022-1694, https://doi.org/10.1016/j.jhydrol.2019.124207

Liu, Y. (2020). Probabilistic spatiotemporal wind speed forecasting based on a variational Bayesian deep learning model. *Applied Energy*, *260*, ISSN 0306-2619, https://doi.org/10.1016/j.apenergy.2019.114259

Liu, Y. (2020). Uplink-Aided High Mobility Downlink Channel Estimation over Massive MIMO-OTFS System. *IEEE Journal on Selected Areas in Communications*, *38*(9), 1994-2009, ISSN 0733-8716, https://doi.org/10.1109/JSAC.2020.3000884

MacKay, D. J. C. (2003). *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press.

Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*. MIT Press.

Neal, R. M. (1993). *Probabilistic Inference Using Markov Chain Monte Carlo Methods*. Technical Report 2, University of Toronto.

Nguyen, C. (2020). Uncertainty in model-agnostic meta-learning using variational inference. *Proceedings 2020 IEEE Winter Conference on Applications of Computer Vision Wacv 2020*, 3079-3089, https://doi.org/10.1109/WACV45572.2020.9093536

Ni, P. (2021). Probabilistic model updating via variational Bayesian inference and adaptive Gaussian process modeling. *Computer Methods in Applied Mechanics and Engineering*, *383*, ISSN 0045-7825, https://doi.org/10.1016/j.cma.2021.113915

O'Hagan, A., & Forster, J. (2004). *Bayesian Inference*. Wiley.

Robert, C. P., & Casella, G. (2004). *Monte Carlo Statistical Methods*. Springer.

Yue, Z. (2019). Variational denoising network: Toward blind noise modeling and removal. *Advances in Neural Information Processing Systems*, *32*, ISSN 1049-5258

Zhang, T. (2023). Bayesian Inference for State-Space Models With Student-t Mixture Distributions. *IEEE Transactions on Cybernetics*, *53*(7), 4435-4445, ISSN 2168-2267, https://doi.org/10.1109/TCYB.2022.3183104

Zhao, S. (2023). Tuning-Free Bayesian Estimation Algorithms for Faulty Sensor Signals in State-Space. *IEEE Transactions on Industrial Electronics*, *70*(1), 921-929, ISSN 0278-0046, https://doi.org/10.1109/TIE.2022.3153814

Zhu, W. (2023). Uncertainty quantification of proton-exchange-membrane fuel cells degradation prediction based on Bayesian-Gated Recurrent Unit. *Etransportation*, *16*, ISSN 2590-1168, https://doi.org/10.1016/j.etran.2023.100230