Advanced Methods in Applied Statistics: Article Write-up of "Breaking the spell of Gaussianity: Forecasting with higher order Fisher matrices" (Sellentin et al., 2014)

Nikolai Plambech Nielsen (LPK331)

March 7, 2023

1 Introduction

A major part of experimental physics is data analysis of experiments and further refinement of experimental design. Robust methods for data analysis are introduced in this course, namely MCMC bootstrapping. This robustness however comes at the cost of computational time. A low cost alternative to MCMC bootstrapping for parameter estimation is the use of the Fisher Matrix, which approximates the likelihood as a Gaussian with mean at the MLE parameters. This is an easily calculated quantity, but just as MCMC bootstrapping sacrifices computational time for robustness, so does the Fisher Matrix sacrifice robustness for computational time.

In the above paper, authors Elena Sellentin, Miguel Quartin and Luca Amendola present the "DALI" method (Derivative Approximation for LIkelihoods), which serves as a middle ground between the two extremes of the Fisher Matrix and MCMC bootstrapping.

Note: Throughout the paper, Einstein summation convention is used: repeated indices in a formula imply summation over that index.

2 Review

In the paper, the authors assume a general (negative) log-likelihood (LLH) $\mathcal{L} = -\ln(P)$, dependent on a set of n parameters, labelled by p_α, with $\alpha = 1, \dots, n$. They then expand the LLH in a Taylor series to fourth order about the maximum likelihood estimator (MLE). This results in an approximate LLH of the form

$$\mathcal{L} = N + \frac{1}{2}F_{\alpha\beta}\Delta p_\alpha \Delta p_\beta + \frac{1}{3!}S_{\alpha\beta\gamma}\Delta p_\alpha \Delta p_\beta \Delta p_\gamma + \frac{1}{4!}Q_{\alpha\beta\gamma\delta}\Delta p_\alpha \Delta p_\beta \Delta p_\gamma \Delta p_\delta + \mathcal{O}(5), \tag{1}$$

where

$$F_{\alpha\beta} = \mathcal{L}_{,\alpha\beta}, \quad S_{\alpha\beta\gamma} = \mathcal{L}_{,\alpha\beta\gamma}, \quad Q_{\alpha\beta\gamma\delta} = \mathcal{L}_{,\alpha\beta\gamma\delta} \tag{2}$$

(with subscript "$,\alpha$" denoting a partial derivative w.r.t. the parameter α), $\Delta p_\alpha = p_\alpha - \hat{p}_\alpha$ is the deviation of parameter p_α from the MLE value of $\hat{p}_\alpha$, and lastly N is a normalization. In this expression, three tensors appear. These are rank two, three and four respectively: $F_{\alpha\beta}$ is the Fisher Matrix, $S_{\alpha\beta\gamma}$ is dubbed the "Flexion" and $Q_{\alpha\beta\gamma\delta}$ is dubbed the "Quarxion". Note that for linear models (and Gaussian likelihoods in general), the Flexion and Quarxion are identically zero. This method then also serves as an explicit check of Gaussianity In the frequentist approach these can be simplified to

$$S = 3\boldsymbol{\mu}_{,\alpha\beta}M\boldsymbol{\mu}_{,\gamma}\Delta p_\alpha \Delta p_\beta \Delta p_\gamma, \quad Q = (4\boldsymbol{\mu}_{,\alpha\gamma\delta}M\boldsymbol{\mu}_{,\beta} + 3\boldsymbol{\mu}_{,\delta\gamma}M\boldsymbol{\mu}_{,\beta\alpha})\Delta p_\alpha \Delta p_\beta \Delta p_\gamma \Delta p_\delta, \tag{3}$$

where M is the inverse of the covariance in the data-space, i.e., the sample covariance, and $\boldsymbol{\mu}$ is a vector of the model predictions for the data.

The problem with this expansion is that the Flexion term is never globally negative (which is required, if the probability density is to remain positive) due to it being qubic in Δ. Similarly, the Quarxion term is not guaranteed to be globally negative either. This shortcoming can be alleviated by expanding in order of derivative instead of parameter deviations. For example, up to second order in the derivatives of the model, the LLH is

$$\mathcal{L} = \frac{1}{2}\boldsymbol{\mu}_{,\alpha}M\boldsymbol{\mu}_{,\beta}\Delta p_\alpha \Delta p_\beta + \left[\frac{1}{2}\boldsymbol{\mu}_{,\alpha\beta}M\boldsymbol{\mu}_{,\gamma}\Delta p_\alpha \Delta p_\beta \Delta p_\gamma + \frac{1}{8}\boldsymbol{\mu}_{,\delta\gamma}M\boldsymbol{\mu}_{,\beta\alpha}\Delta p_\alpha \Delta p_\beta \Delta p_\gamma \Delta p_\delta\right] + \mathcal{O}(3). \tag{4}$$

This approximation is positive definite and normalizable, provided that the matrix M is positive definite. This approximation is dubbed the "doublet-DALI". To third order the approximation is dubbed the "triplet-DALI", and is

$$\begin{aligned}\mathcal{L} =& \frac{1}{2}\boldsymbol{\mu}_{,\alpha} M \boldsymbol{\mu}_{,\beta} \Delta p_\alpha \Delta p_\beta + \left[\frac{1}{2}\boldsymbol{\mu}_{,\alpha\beta} M \boldsymbol{\mu}_{,\gamma} \Delta p_\alpha \Delta p_\beta \Delta p_\gamma + \frac{1}{8}\boldsymbol{\mu}_{,\delta\gamma} M \boldsymbol{\mu}_{,\beta\alpha} \Delta p_\alpha \Delta p_\beta \Delta p_\gamma \Delta p_\delta\right] \\ &+ \left[\frac{1}{6}\boldsymbol{\mu}_{,\delta} M \boldsymbol{\mu}_{,\beta\alpha\gamma} \Delta p_\alpha \Delta p_\beta \Delta p_\gamma \Delta p_\delta + \frac{1}{12}\boldsymbol{\mu}_{,\alpha\beta\delta} M \boldsymbol{\mu}_{,\gamma\tau} \Delta p_\alpha \Delta p_\beta \Delta p_\gamma \Delta p_\delta \Delta p_\tau \right. \\ &\left. + \frac{1}{72}\boldsymbol{\mu}_{,\alpha\beta\delta} M \boldsymbol{\mu}_{,\delta\tau\sigma} \Delta p_\alpha \Delta p_\beta \Delta p_\gamma \Delta p_\delta \Delta p_\tau \Delta p_\sigma\right] + \mathcal{O}(4).\end{aligned} \tag{5}$$

These two equations are, in essence, the DALI method. Armed with these equations, the authors apply the method to type Ia supernova data ("SNeIa" data). In particular they apply it to the Union2.1 catalogue of SNeIa data (Amanulla et al. 2010). The results are reproduced below in Figure 1, where we see that the DALI method is capable of recovering the non-Gaussianities present in the data set.

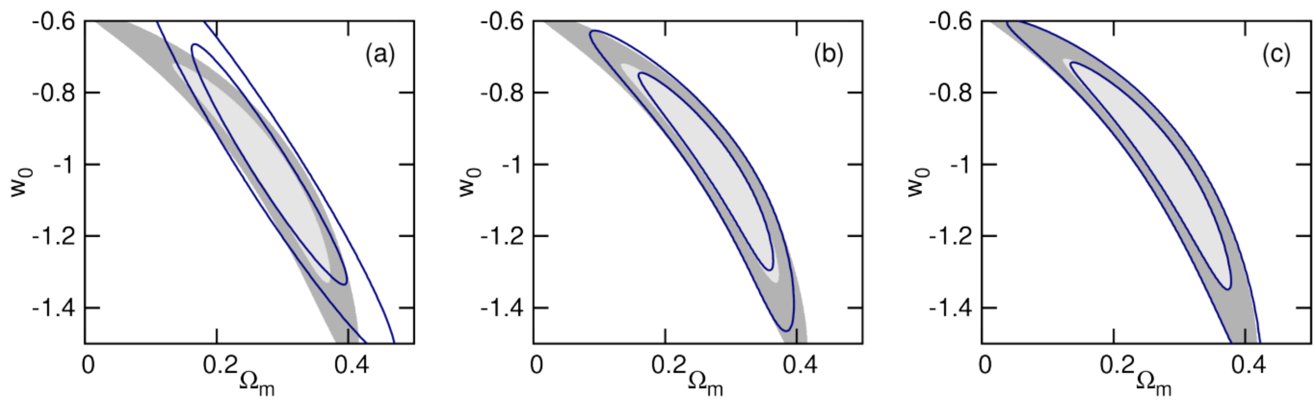

Figure 1. Comparison of the full, non-approximated posterior of the SNeIa Union2.1 catalogue (grey) with different approximations (dark-blue). In this plot only we fix $w_a = 0$ (i.e., assume what is often called "wCDM" model). The confidence contours are drawn at the 1 and 2σ confidence levels. Panel (a): The Fisher Matrix approximation; panel (b): Eq. (15), the doublet-DALI approximation of the posterior includes well the non-Gaussianities; panel (c): Eq. (16), the triplet-DALI approximation captures the non-Gaussianities even better.

Figure 1: The DALI method applied to to SNeIa data from the Union2.1 catalogue, from the article. The LLH is calculated as a function of the mass-density Ω_m and the present dark energy equation of state parameter w_0.

3 Conclusion

The authors set out to expand the Fisher Matrix formalism, to allow it to reproduce non-Gaussianities that are often present in physical models. In this case the expansion is in the form of expanding the likelihood in orders of derivative of the model. By rearranging the sum and using the frequentist formalism, this results in normalizable expressions for the likelihood at any order. In the paper, the authors present expressions up to third order in the derivatives of the model. They then apply these expressions on type Ia supernova data and are able to represent non-Gaussianities in the data.

4 References

Amanullah et al., 2010, Ap.J., 716, 712, 1004.1711, ADS