Breaking the Spell of Gaussianity: Forecasting with higher order Fisher Matrices

Article by Elena Sellentin, Miguel Quartin and Luca Amendola. Presentation by Nikolai Plambech Nielsen

Published April 2, 2014
Presented March 8, 2023

Introduction

- MCMC: Robust parameter estimation, taking into account the non-Gaussianity of the system, but computationally expensive.
- Fisher Matrix: Quick parameter estimation, but approximates every likelihood as a Gaussian distribution. Includes no explicit check of non-Gaussianity.
- DALI (introduced in article): Relatively quick parameter estimation, takes into account non-Gaussianities.

General likelihood form

Authors assume a general Gaussian form of the likelihood:

$$L = N \exp\Big[-\frac{1}{2}[\boldsymbol{m} - \boldsymbol{\mu}(\boldsymbol{p})]^T M [\boldsymbol{m} - \boldsymbol{\mu}(\boldsymbol{p})]\Big] \tag{1}$$

where $\boldsymbol{m}$ is vector of measurements, $\boldsymbol{\mu}(\boldsymbol{p})$ is theoretical predictions corresponding to the measurements (evaluated at the parameters $\boldsymbol{p}$), and M is the inverse covariance matrix.

Note: Parametric dependence is only in the model μ. In appendix and future work the authors generalize to covariances with parametric dependence.

Fisher Matrix

Negative Hessian of LLH. In frequentist regime this reduces to

$$F_{\alpha\beta} = \langle \mathcal{L}_{,\alpha} \mathcal{L}_{,\beta} \rangle, \quad \mathcal{L} = \ln(L). \tag{2}$$

LLH is

$$\mathcal{L} \approx N - \frac{1}{2} \boldsymbol{\mu}_{,\alpha} M \boldsymbol{\mu}_{,\beta} \Delta p_\alpha \Delta p_\beta \equiv N + F, \tag{3}$$

where, $\Delta p_\alpha = p_\alpha - \hat{p}_\alpha$ is deviation from MLE parameters. Note parabolic in parameters, so always Gaussian.

[0]Subscript ”$,\alpha$” means a partial derivative w.r.t. α.

DALI: Taylor expansion of likelihood

To fourth order in parameters:

$$\begin{aligned}\mathcal{L} \approx N &- \frac{1}{2} F_{\alpha\beta} \Delta p_\alpha \Delta p_\beta \\ &- \frac{1}{3!} S_{\alpha\beta\gamma} \Delta p_\alpha \Delta p_\beta \Delta p_\gamma \\ &- \frac{1}{4!} Q_{\alpha\beta\gamma\delta} \Delta p_\alpha \Delta p_\beta \Delta p_\gamma \Delta p_\delta,\end{aligned} \tag{4}$$

with

$$F_{\alpha\beta} = \mathcal{L}_{,\alpha\beta}, \quad S_{\alpha\beta\gamma} = \mathcal{L}_{,\alpha\beta\gamma}, \quad Q_{\alpha\beta\gamma\delta} = \mathcal{L}_{,\alpha\beta\gamma\delta} \tag{5}$$

Problem: Q- and S-terms not globally negative, and thus non-normalizable.

DALI: Collecting terms in orders of derivatives - Doublet DALI

To second order in the derivatives of the model we have the "doublet DALI"

$$\mathcal{L} \approx N + F - \left[\frac{1}{2} \boldsymbol{\mu}_{,\alpha\beta} M \boldsymbol{\mu}_{,\gamma} \Delta p_\alpha \Delta p_\beta \Delta p_\gamma \right. \\ \left. + \frac{1}{8} \boldsymbol{\mu}_{,\delta\gamma} M \boldsymbol{\mu}_{,\beta\alpha} \Delta p_\alpha \Delta p_\beta \Delta p_\gamma \Delta p_\delta \right] \\ = N + F + S \tag{6}$$

Leading second order term at big p is now negative, if M is positive definite, so entire approximation is normalizable.

DALI: Collecting terms in orders of derivatives - Triplet DALI

To third order, the "triplet DALI":

$$\mathcal{L} \approx N + F + S - \left[\frac{1}{6}\boldsymbol{\mu}_{,\delta} M \boldsymbol{\mu}_{,\beta\alpha\gamma} \Delta p_\alpha \Delta p_\beta \Delta p_\gamma \Delta p_\delta \right.$$
$$+ \frac{1}{12}\boldsymbol{\mu}_{,\alpha\beta\delta} M \boldsymbol{\mu}_{,\gamma\tau} \Delta p_\alpha \Delta p_\beta \Delta p_\gamma \Delta p_\delta \Delta p_\tau$$
$$\left. + \frac{1}{72}\boldsymbol{\mu}_{,\alpha\beta\delta} M \boldsymbol{\mu}_{,\delta\tau\sigma} \Delta p_\alpha \Delta p_\beta \Delta p_\gamma \Delta p_\delta \Delta p_\tau \Delta p_\sigma \right]$$
$$= N + F + S + Q. \tag{7}$$

Applications

Applied to type Ia supernova data (Union2.1 catalogue)

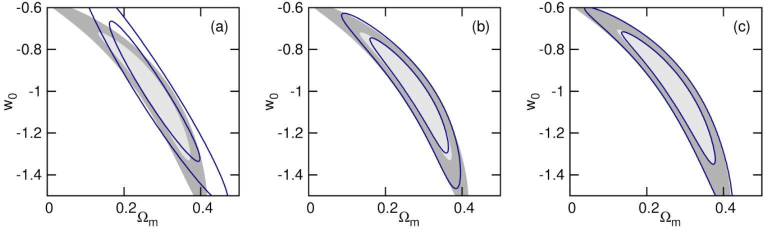

Figure 1. Comparison of the full, non-approximated posterior of the SNeIa Union2.1 catalogue (grey) with different approximations (dark-blue). In this plot only we fix $w_a = 0$ (i.e., assume what is often called "wCDM" model). The confidence contours are drawn at the 1 and 2σ confidence levels. Panel (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: The DALI method applied to mock SNeIa data, 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.

Applications

Applied to mock type Ia supernova data

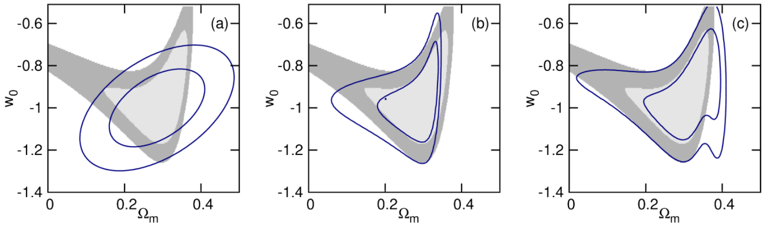

Figure 2. Same as Figure 1 but for the mock catalogue of 1000 SNeIa (see text) and marginalizing over w_a, which results in a heavily non-Gaussian grey posterior. Again the DALI methods capture the shape of the posterior much better than the Fisher Matrix. Note that the doublet-DALI is a very good compromise between speed and shape accuracy.

Figure: The DALI method applied 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. Here marginalized over $w_a \in (-\infty, \infty)$, from the "CPL" model.