Back to Normal!
Gaussianizing posterior distributions for cosmological probes

Robert L. Schuhmann[1], Benjamin Joachimi[1] and Hiranya V. Peiris[1]

[1]Dept. of Physics and Astronomy
University College London
Gower Place
London, UK WCE1 6BT
email: robert.schuhmann.13@ucl.ac.uk

Abstract. We present a method to map multivariate non-Gaussian posterior probability densities into Gaussian ones via nonlinear Box-Cox transformations, and generalizations thereof. This is analogous to the search for normal parameters in the CMB, but can in principle be applied to any probability density that is continuous and unimodal. The search for the optimally Gaussianizing transformation amongst the Box-Cox family is performed via a maximum likelihood formalism. We can judge the quality of the found transformation *a posteriori*: qualitatively via statistical tests of Gaussianity, and more illustratively by how well it reproduces the credible regions. The method permits an analytical reconstruction of the posterior from a sample, *e.g.* a Markov chain, and simplifies the subsequent joint analysis with other experiments. Furthermore, it permits the characterization of a non-Gaussian posterior in a compact and efficient way. The expression for the non-Gaussian posterior can be employed to find analytic formulae for the Bayesian evidence, and consequently be used for model comparison.

Keywords. methods: data analysis, methods: numerical, cosmology: observations

---

1. The Idea

The quest for normal parameters, where the cosmological likelihood is of approximately multivariate Gaussian form, has proven to be highly advantageous for rapid computation of CMB likelihoods, *e.g.* CMBfit, CMBwarp, PICO (Kosowsky *et al.* 2002, Chu *et al.* 2003, Jimenez *et al.* 2004, Sandvik *et al.* 2004, Fendt & Wandelt 2007). To achieve the same for systems in which this is not feasible analytically, we computationally search for a parameter transformation which renders the transformed probability density approximately Gaussian. This allows for a reliable reconstruction of the posterior distribution, *e.g.* from an MCMC sample, and enormously simplifies the reporting and further usage of the multivariate posterior constraints, via data compression.

Our proposed method can be used to characterize and to reproduce high-dimensional probability densities (see Fig. 1). Instead of publishing an MCMC sample ($\gtrsim 10^5$ numbers), the posterior can be reconstructed simply from the first and second moments of the Gaussianized distribution, and the optimal transformation parameters. This provides substantial data compression, while simultaneously providing versatility and accuracy. Further, there is no need for smoothing of contours, in contrast with density estimates based on MCMC samples.

2. The Tools: Box-Cox Transformations and their Kin

We employ families of one-dimensional transformations mapping cosmological parameters to normal parameters. The transformations are defined by a set of parameters for
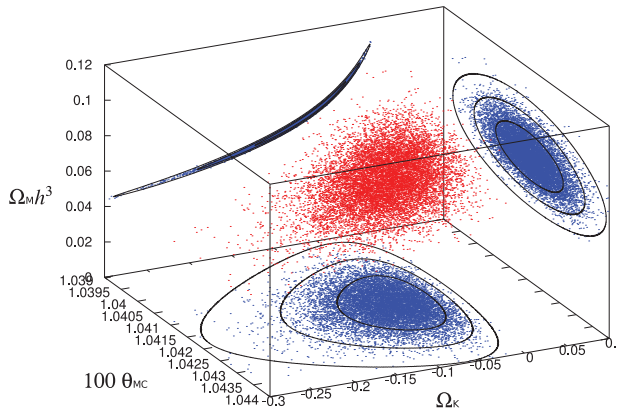
Figure 1. *3D marginalized MCMC sample from Planck XVI (2013) (red), the marginal 2D sample (blue), and 2D contours of the full 3D analytic posterior (black), as reconstructed via Box-Cox transformations.*
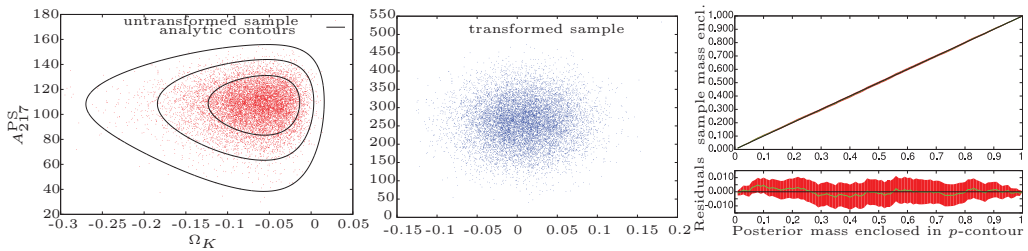

Figure 2. *left to right: 1) 2D MCMC sample from Planck XVI (2013) (red) and reconstructed analytic posterior (black). 2) Sample, as Gaussianized with ABC transformation (blue). 3) CC plot showing the validity of the reconstructed black contours. "Residuals" shows the deviation of the point sample probability mass in a given contour from the probability mass in the analytic posterior.*

each dimension. One particularly useful family was introduced by Box & Cox (1964):

$$x \mapsto B_{a,\lambda}(x) = \begin{cases} [\lambda^{-1}(x+a)^{\lambda} - 1] & (\lambda \neq 0) \\ \log(x+a) & (\lambda = 0). \end{cases} \qquad (2.1)$$

Notice that this family is continuous in λ. As a useful generalization for removing residual kurtosis we propose the ABC transformation,

$$x \mapsto F_{a,\lambda,t}(x) = \begin{cases} t^{-1} \sinh[t\, B_{a,\lambda}(x)] & (t > 0) \\ B_{a,\lambda}(x) & (t = 0) \\ t^{-1} \operatorname{arsinh}[t\, B_{a,\lambda}(x)] & (t < 0). \end{cases} \qquad (2.2)$$

Our algorithm can be applied to arbitrary parameterized transformations, suitable for various forms of non-Gaussianity.

3. The Method: Optimization and Verification

Given a sample from a non-Gaussian posterior distribution, we find the optimal transformation parameters by maximizing the log-likelihood of the transformed sample after a given transformation (Joachimi & Taylor 2011). Typically, degeneracies appear in the transformation parameters. To avoid these, we penalize transformations too far from the identity.

The quality of the reproduced contours can be evaluated via cross-contour (CC) plots: given a certain posterior contour region of enclosed probability mass m, compute the
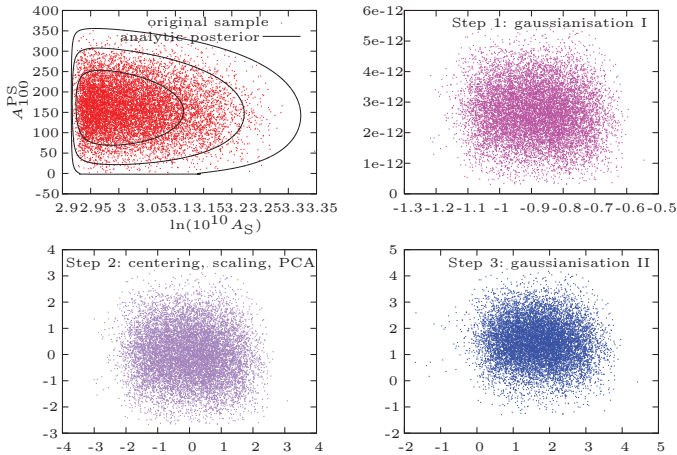
Figure 3. *From top left to bottom right: 1) Two-pass optimization of 2D sample from Planck XVI (2013), showing A_{100}^{PS} vs. $\ln(10^{10}A_s)$, and reconstructed posterior (black contours). 2) After first Gaussianization transformation (all y values plus 0.91585740528). 3) After linear reshaping. 4) After second Gaussianization transformation.*

fraction f of points inside. Plotting f over m, we can determine whether every credible region of the analytic posterior encloses the correct probability mass of the point sample, *i.e.* whether $m \simeq f$ within the margin of noise. To quantify the sampling noise, we perform a non-parametric bootstrap to find the 95% confidence interval for every f (see Fig. 2).

4. Two-pass transformations

If the transformed posterior is not sufficiently Gaussian (as, *e.g.*, determined via a CC plot), the process can be repeated. We have implemented the following protocol for the two-pass transformation (see Fig. 3):

Step 1: Optimize the parameters of the first transformation.

Step 2: Linear reshaping: Center the sample, scale every direction to unit variance, then rotate into the PCA eigenbasis.

Step 3: Optimize the parameters of the second transformation.

Note that this two-pass protocol is more effective for Gaussianization than one single transformation, because the transformation families employed (Box-Cox, ABC) do not form groups – in particular, this means that two subsequent transformations do not result in a transformation of the same family.

5. Outlook

Transforming the unnormalized posterior distribution to a Gaussian form can allow for a simple and fast computation of the marginal likelihood, and hence used for model comparison (Schuhmann *et al.*, in preparation).

References

Box, G. E. P. & Cox, D. R. 1964, *J. Roy. Stat. Soc.* B 26/2, 211-252

Chu, M., Kaplinghat, M., & Knox, L. 2003, *ApJ*, 596, 725

Fendt, W. A. & Wandelt, B. D. 2007, *ApJ*, 654, 2 (astro-ph/0606709)

Jimenez, R., Verde, L., & Peiris, H. 2004, *Phys. Rev. D*, 70, 023005 (astro-ph/0404237)

Joachimi, B. & Taylor, A. N. 2011, *MNRAS*, 416(2), 1010-1022, (arXiv:1103.3370)

Kosowsky, A., Milosavljevic, M., & Jimenez, R. 2002, *Phys. Rev. D*, 66, 063007

Planck Collaboration (Ade *et al.*) 2013, *A&A*, in press (arXiv:1311.1657)

Sandvik, H. B., Tegmark, M., Wang, X., & Zaldarriaga, M. 2004, *Phys. Rev. D*, 69, 063005 (astro-ph/0311544v2)