Adaptive Bayes test for monotonicity

Jean-Bernard Salomond[1]

Bayesian Young Statisticians Meeting (BAYSM), Milan June, 5-6, 2013
Paper no. 37

[1] CREST and Université Paris Dauphine, Paris, France
jean.bernard.salomond@ensae.fr

Abstract

We propose a Bayesian procedure to test monotonicity of a regression function in a Bayesian setting where the Bayes Factor approach yield poor results. We propose a Bayesian procedure that is consistent and achieves the optimal separation rate. Furthermore, our testing procedure is straightforward to implement which is a great advantages compare to the frequentist tests.

Keywords: Bayesian nonparametric; Asymptotic properties of tests; Nonparametric testing

1 Introduction

Shape constrained models are of growing interest in the non parametric field. Among them monotone constrains are very popular and have been widely studied in the literature as such hypotheses arise naturally in many applications. [Barlow et al., 1972] and [Mukerjee, 1988] among others proposed a shape constrain estimator of monotonic regression functions. These methods are widely applied in practice. For instance [Bornkamp and Ickstadt, 2009] consider monotone function when modeling the response to a drug as a function of the dose and [Neittaanmäki et al., 2008] use a monotone representation for environmental data.

In this paper we propose a Bayesian procedure to test for monotonicity constrains. We consider the Gaussian regression setting

$$Y_i = f(i/n) + \sigma\epsilon_i,\ \epsilon_i \overset{iid}{\sim} \mathcal{N}\left(0,1\right), \sigma > 0,\ i = 1, \dots, n, \tag{1}$$

and want to test

$$H_0 : f \searrow \text{ versus } H_1 : f \text{ is not } \searrow . \tag{2}$$

Note that in this setting both the null and the alternative hypotheses are non parametric. For this test, it appears that the Bayes Factor, which is the standard Bayesian approach to testing, lead to poor results. To tackle this issue we consider an alternative approach and test

$$H_0^a : \tilde{d}(f, \mathcal{F}) \leq M_n \quad \text{versus} \quad H_1^a : \tilde{d}(f, \mathcal{F}) > \tau \tag{3}$$

where $\tilde{d}(f, \mathcal{F})$ is a distance between f and the set of monotone non increasing function $\mathcal{F}$ and τ a threshold, which can be calibrated a priori given some knowledge on how far from monotonicity is still acceptable as approximately monotone. However, such a knowledge may not be available, we thus propose an automatic calibration of τ such that our test has good asymptotic properties. This ideas is similar to the one proposed in [Rousseau, 2007] and to the approximation of a point null hypothesis by a interval hypothesis testing, see also [Verdinelli and Wasserman, 1998]. To perform such a test we consider the $\gamma_0 - \gamma_1$ loss with fixed $\gamma_0, \gamma_1 > 0$ and thus our procedure can be define as

$$\delta_n^\pi := \begin{cases} 0 \text{ if } \pi\left(\tilde{d}(f, \mathcal{F}) \leq \tau_n | X_n\right) \geq \frac{\gamma_0}{\gamma_0 + \gamma 1} \\ 1 \text{ otherwise} \end{cases} \tag{4}$$

This test is straightforward to implement and will only require sampling under the posterior.

2 Main results

Similarly to the frequentist test, we consider α-Hölderian alternatives with $\alpha \leq 1$. Under some mild conditions on the prior that do not depend on the regularity α under the alternative, we get an explicit threshold τ such that our testing procedure defined in (4) satisfies

$$\begin{aligned} \sup_{f \in \mathcal{F}} E_f^n(\delta_n^\pi) = o(1) \\ \sup_{f, d_n(f,\mathcal{F}) > \rho, f \in \mathcal{H}(\alpha, L)} E_f^n(1 - \delta_n^\pi) = o(1) \end{aligned} \tag{5}$$

furthermore, if $\rho_n(\alpha) = M\left(n/\log(n)\right)^{-\alpha/(2\alpha+1)}$ then (5) is still valid with $\rho = \rho_n(\alpha)$. Thus our procedure is consistent and has the optimal adaptive separation rate.

References

[Barlow et al., 1972] Barlow, R. E., Bartholomew, D. J., Bremner, J. M., and Brunk, H. D. (1972). *Statistical inference under order restrictions. The theory and application of isotonic regression.* John Wiley & Sons, London-New York-Sydney. Wiley Series in Probability and Mathematical Statistics.

[Bornkamp and Ickstadt, 2009] Bornkamp, B. and Ickstadt, K. (2009). Bayesian nonparametric estimation of continuous monotone functions with applications to dose-response analysis. *Biometrics*, 65(1):198–205.

[Groeneboom, 1985] Groeneboom, P. (1985). Estimating a monotone density. In *Proceedings of the Berkeley conference in honor of Jerzy Neyman and Jack Kiefer, Vol. II (Berkeley, Calif., 1983)*, Wadsworth Statist./Probab. Ser., pages 539–555, Belmont, CA. Wadsworth.

[Khazaei et al., 2012] Khazaei, S., Rousseau, J., and Balabdaoui, F. (2012). Nonparametric bayesian estimation of densities under monotonicity constraint.

[Mukerjee, 1988] Mukerjee, H. (1988). Monotone nonparameteric regression. *Ann. Statist.*, 16(2):741–750.

[Neittaanmäki et al., 2008] Neittaanmäki, P., Rossi, T., Majava, K., and Pironneau, O. (2008). Monotonic regression for assessement of trends in environmental quality data.

[Prakasa Rao, 1970] Prakasa Rao, B. L. S. (1970). Estimation for distributions with monotone failure rate. *Ann. Math. Statist.*, 41:507–519.

[Robertson et al., 1988] Robertson, T., Wright, F. T., and Dykstra, R. L. (1988). *Order restricted statistical inference.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Ltd., Chichester.

[Rousseau, 2007] Rousseau, J. (2007). Approximating interval hypothesis: p-values and Bayes factors. In *Bayesian statistics 8*, Oxford Sci. Publ., pp. 417–452. Oxford: Oxford Univ. Press.

[Salomond, 2013] Salomond, J-B. (2013). Convergence rate and consistency at 0 for monotone non increasing densities. *ArXiv e-prints*

[Verdinelli and Wasserman, 1998] Verdinelli, I. and L. Wasserman (1998). Bayesian goodness-of-fit testing using infinite-dimensional exponential families. *Ann. Statist. 26*(4), 1215–1241.