Adaptive Bayes Test for monotonicity

Journées MAS - Clermont Ferrand

J-B. Salomond

CREST - Dauphine

August 30, 2012

Contents

Introduction

Consider the regression problem

$$Y_i = f(i/n) + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(0, \sigma^2) \tag{1}$$

where $f : [0, 1] \to \mathbb{R}$ and $\sigma \in \mathbb{R}^+$ are unknown

Introduction

Consider the regression problem

$$Y_i = f(i/n) + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(0, \sigma^2) \tag{1}$$

where $f : [0, 1] \to \mathbb{R}$ and $\sigma \in \mathbb{R}^+$ are unknown

Aim

We want to test, for $0 < \alpha \leq 1$

$$H_0 : f \searrow \text{ versus } H_1 : f \text{ is not } \searrow \text{ and } f \in H_\alpha(L)$$

Introduction cont'd

More precisely we want a test that

- require no regularity assumptions under the null
- has good asymptotic properties (consistency)
- does not depend on the regularity α under the alternative
- achieves the "optimal" separation rate for a wide variety of α
- is easy to implement

Why ?

Shape constrained estimation appears in a variety of models (Drug Response, Global Warming, Survival Analysis, . . .)

Why ?

Shape constrained estimation appears in a variety of models (Drug Response, Global Warming, Survival Analysis, . . .)

Many tests in the frequentist literature

Why ?

Shape constrained estimation appears in a variety of models (Drug Response, Global Warming, Survival Analysis, . . .)
Many tests in the frequentist literature but no Bayesian results . . .

Our Model

We build a prior on f by considering a piecewise constant approximation

$$f_{\omega,k} = \sum_{i=1}^{k} \omega_i \mathbb{1}_{](i-1)/k, i/k]}$$

and a prior

$$\pi : \begin{cases} k & \sim \pi_k \\ \omega_i & \overset{iid}{\sim} g \\ \sigma & \sim h \end{cases}$$

Test in a nutshell

Idea for a test

We will test the monotonicity of the sequence $(\omega_i)_i$ and thus need

- A criteria for monotonicity
- Conditions on the prior such that $f_{\omega,k}$ concentrates around f

Bayes Test

We define

$$H(\omega, k) = \max_{1<i<j\leq k} (\omega_j - \omega_i)$$

Bayes Test

We define

$$H(\omega, k) = \max_{1<i<j\leq k} (\omega_j - \omega_i)$$

Test

We consider the test

$$\delta_n^\pi = \mathbb{1}\left\{\pi\left(H(\omega, k) > M_n^k | Y_n\right) > 1/2\right\}$$

Bayes Test

We define

$$H(\omega, k) = \max_{1<i<j\leq k} (\omega_j - \omega_i)$$

Test

We consider the test

$$\delta_n^\pi = \mathbb{1}\left\{\pi\left(H(\omega, k) > M_n^k | Y_n\right) > 1/2\right\}$$

Remarks

- This is a modified version of the Bayes Factor
- We reject monotonicity if we have sufficiently strong evidence that f is not $\searrow$

Some definitions

Consistency

We will say that δ_n^π is consistent if

$$\begin{aligned} \sup_{f \searrow} \mathrm{E}_0^n(\delta_n^\pi) &= o(1) \\ \sup_{d(f,\searrow)>\rho, f\in H_\alpha(L)} \mathrm{E}_0^n(1-\delta_n^\pi) &= o(1) \end{aligned} \tag{2}$$

Separation rate

We define the separation rate ρ_n of our test as the minimal sequence r_n such that

$$\begin{aligned} \sup_{f \in \searrow} \mathrm{E}_0^n(\delta_n^\pi) &= o(1) \\ \sup_{d(f,\searrow)>r_n, f\in H_\alpha(L)} \mathrm{E}_0^n(1-\delta_n^\pi) &= o(1) \end{aligned} \tag{3}$$

Main Theorem

Theorem

Let $M_n^k = M_0\sqrt{k\log(n)/n}$ *and let* π *be a prior on* $f_{\omega,k}$ *such that* $\omega_i \overset{iid}{\sim} g$ *and* $k \sim \pi_k$, $\sigma \sim h$. *Assume that g and h put mass on* $\mathbb{R}$ *and* $\mathbb{R}^{+*}$ *respectively and that* π_k *is such that there exist positive constants* C_d *and* C_u *such that*

$$e^{-C_d kL(k)} \leq \pi_k(k) \leq e^{-C_u kL(k)}$$

Where $L(k)$ *is either* $\log(k)$ *or* 1. *Consider the test*

$$H_0 : f \searrow \text{ versus } H_1 : f \text{ not } \searrow, f \in H_\alpha(L)$$

let $\rho_n = M(n/\log(n))^{-\alpha/(2\alpha+1)}$, *then the test* δ_n^π *is consistent and achieve the separation rate* ρ_n.

Sketch of the proof

For a fixed k, let ω_0 be such that

$$\omega_0 = \operatorname{argmin}\{\omega, KL(f_{\omega,k}, f)\}$$

We first prove

Result

$$\pi\left(\max_i |\omega_i - \omega_i^0| \geq C\xi_n^k | Y^n\right) \leq 1/2 + o_{P_0^n}(1). \tag{4}$$

where ξ_n^k is such that

$$\xi_n^k := \begin{cases} (n/\log(n))^{-1/3} & \text{if } f \searrow \\ \rho_n & \text{if } f \text{ not } \searrow,\ f \in H_\alpha(L) \end{cases}$$

Sketch of the proof cnt'd

- For $f \searrow$

$$H(\omega, k) \leq 2 \max_i |\omega_i - \omega_i^0|$$

Sketch of the proof cnt'd

- For $f \searrow$

$$H(\omega, k) \leq 2 \max_i |\omega_i - \omega_i^0|$$

- For f not $\searrow$, and $f \in H_\alpha(L)$

$$H(\omega, k) \geq d_n(f_{\omega^0,k}, \searrow) - 2 \max_i |\omega_i - \omega_i^0|$$

$$\geq \rho_n - ||f_{\omega^0,k} - f||_n - 2 \max_i |\omega_i - \omega_i^0|$$

Sketch of the proof cnt'd

Results

- $||f_{\omega^0,k} - f||_n \leq \rho_n/4$ given some consistency results
- $M_n^k \leq \xi_n^k/4$ with posterior probability $\to 1$

Sketch of the proof cnt'd

Results

- $||f_{\omega^0,k} - f||_n \leq \rho_n/4$ given some consistency results
- $M_n^k \leq \xi_n^k/4$ with posterior probability $\rightarrow 1$

Under H_0

$$\pi\left(H(\omega, k) > M_n^k | Y^n\right) \leq \pi\left(\max|\omega_i - \omega_i^0| > \xi_n^k/8\right)$$

Sketch of the proof cnt'd

Results

- $||f_{\omega^0,k} - f||_n \leq \rho_n/4$ given some consistency results
- $M_n^k \leq \xi_n^k/4$ with posterior probability $\to 1$

Under H_0

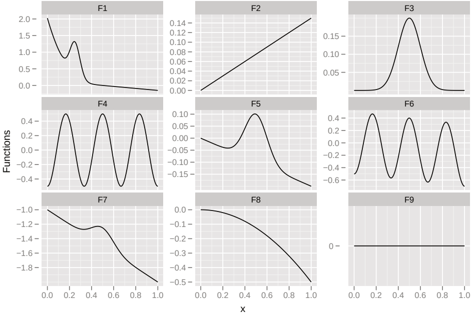
$$\pi\left(H(\omega,k) > M_n^k | Y^n\right) \leq \pi\left(\max|\omega_i - \omega_i^0| > \xi_n^k/8\right)$$

Under H_1

$$\pi\left(H(\omega,k) \leq M_n^k | Y_n\right) \leq \pi\left(\max_i |\omega_i - \omega_i^0| \geq \frac{\rho_n - ||f_{\omega^0,k} - f||_n - M_n^k}{4} | Y^n\right)$$

Simulated Examples

We run our test for nine functions adapted from the frequentist literature



Simulated Examples cnt'd

- We propose a generic choice for the prior
- We study the behaviour of our test for various values of n
- Compare our results with those obtained in the literature

Some specific choice of prior

We choose

- $k \sim \mathcal{P}(\lambda)$
- $\sigma | k \sim IG(\alpha, \beta)$
- $\omega_i | k, \sigma^2 \sim \mathcal{N}(m, \sigma^2/\mu)$

With specific choices for the hyper-parameters

Some specific choice of prior

We choose

- $k \sim \mathcal{P}(\lambda)$
- $\sigma | k \sim IG(\alpha, \beta)$
- $\omega_i | k, \sigma^2 \sim \mathcal{N}(m, \sigma^2/\mu)$

With specific choices for the hyper-parameters

- We get a closed form formulation for the posterior distribution.
- All the computations are straightforward

Results of the Simulation study

f	σ^2	Barraud et al. $n = 100$	Akakpo et al. $n = 100$	Bayes Test, n : 100	250	500	1000
f_1	0.01	99	99	100.0	100.0	100.0	100.0
f_2	0.01	99	100	99.1	100.0	100.0	100.0
f_3	0.01	99	98	99.6	100.0	100.0	100.0
f_4	0.01	100	99	100.0	100.0	100.0	100.0
f_5	0.004	99	99	99.6	100.0	100.0	100.0
f_6	0.006	98	99	100.0	100.0	100.0	100.0
f_7	0.01	76	68	20.1	40.5	61.6	85.4
f_8	0.01	-	-	0.5	1.0	0.6	0.4
f_9	0.01	-	-	6.2	6.2	4.4	3.4

Skip Real Data

Real Datas

We apply our procedure to Global Warming dataset

Real Datas

We apply our procedure to Global Warming dataset

Data

We model the annual temperature anomalies since the 1850's.

Real Datas

We apply our procedure to Global Warming dataset

Data

We model the annual temperature anomalies since the 1850's. Used in the frequentist literature to perform isotonic regression

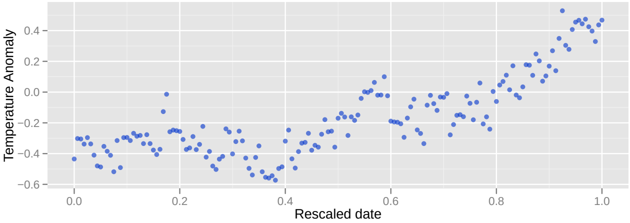
Real Datas

We apply our procedure to Global Warming dataset

Data

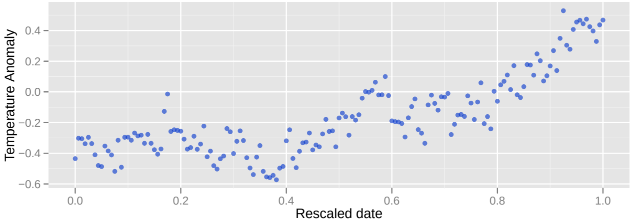
We model the annual temperature anomalies since the 1850's. Used in the frequentist literature to perform isotonic regression
$\rightarrow$ Is there a monotone trend ?

Real Data cnt'd



We run the MCMC sampler using $K = 10^5$ iterations

Real Data cnt'd



We run the MCMC sampler using $K = 10^5$ iterations $\rightarrow$ non monotonic trend (!)

Summary

We have

- A consistent test for monotonicity
- that achieve an “optimal” separation rate
- that is easy to implement

Summary

We have

- A consistent test for monotonicity
- that achieve an "optimal" separation rate
- that is easy to implement

Problems

- We don't control for the constant
- No idea how the BF behave