A family of adaptive Bayesian methods for statistical ill-posed inverse problems

Xavier Loizeau

Contents

Contents

Statistical ill-posed inverse problems

Notations:

- $\lambda u_j = \lambda_j u_j$;
- $\theta = \sum_{j \in \mathscr{J}} \theta_j u_j$;
- $\overline{\mathbb{P}}^n := \delta_{Y^n}$;
- $Y_j^n := \mathbb{E}_{\overline{\mathbb{P}}^n}[\langle u_j | Y^n \rangle]$.

Consider a class of inverse problems:

- Statistical model $\left(\mathcal{S}, \left(\mathbb{P}^n_{|\theta}\right)_{\theta \in \Theta, n \in \mathbb{N}}\right)$;
- observation $Y^n \sim \mathbb{P}^n_{|\lambda\theta^\circ}$;
- an ONB $\mathscr{U} = (u_j)_{j \in \mathscr{J}}$ of Θ that diagonalises λ.

Likelihood:

$$\frac{L^n_{Y^n|\boldsymbol{\theta}}(y^n, \theta)}{L^n_{Y^n|\boldsymbol{\theta}}(y^n, 0)} = \exp\left[-\frac{n}{2}\left(\sum_{j \in \mathscr{J}} \lambda_j^2 \theta_j^2 + 2 \sum_{j \in \mathscr{J}} \lambda_j \theta_j y_j + \Delta_1^n(\theta, y^n)\right)\right]$$

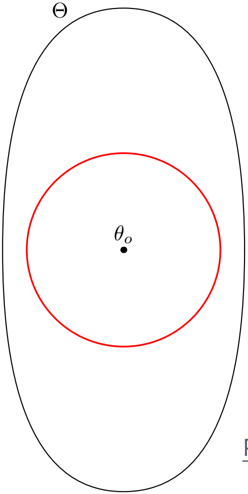
References: Engl et al. [2000], Cavalier [2011], Antoni [2012]

The frequentist Bayesian point of view

Consider:

- a true parameter $\theta^\circ \in \Theta$;
- $(\Theta, \mathscr{A})$ to be a measurable space;
- a sequence of priors about θ° denoted $\left(\mathbb{P}^n_{\boldsymbol{\theta}}\right)_{n\in\mathbb{N}} : \mathscr{A}^{\otimes\mathbb{N}} \to [0,1]^{\times\mathbb{N}}$;
- we have interest in the posterior distributions $\left(\mathbb{P}^n_{\boldsymbol{\theta}|Y^n}\right)_{n\in\mathbb{N}}$.

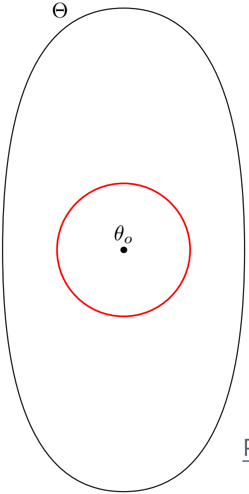
References: Schwartz [1965], Schwartz [1964], Ghosal et al. [2000]



- A sequence $(\phi_n)_{n\in\mathbb{N}}$ is a contraction rate at θ° if

$$\forall (c_n), c_n \to \infty, \quad \lim_{n\to\infty} \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^n_{\boldsymbol{\theta}|Y^n}\left(d(\boldsymbol{\theta},\theta^\circ)^2 \geq c_n\,\phi_n\right)\right] = 0;$$

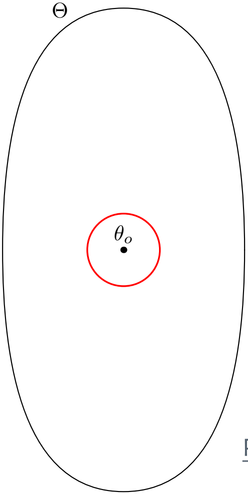
Reference: CASTILLO ET AL. [2008]



- A sequence $(\phi_n)_{n\in\mathbb{N}}$ is a contraction rate at θ° if

$$\forall (c_n), c_n \to \infty, \quad \lim_{n\to\infty} \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^n_{\boldsymbol{\theta}|Y^n}\left(d(\boldsymbol{\theta},\theta^\circ)^2 \geq c_n\,\phi_n\right)\right] = 0;$$

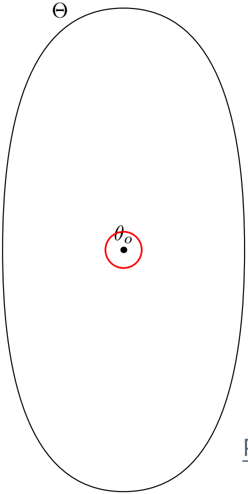
Reference: Castillo et al. [2008]



- A sequence $(\phi_n)_{n\in\mathbb{N}}$ is a contraction rate at θ° if

$$\forall\,(c_n)\,,c_n \to \infty, \qquad \lim_{n\to\infty} \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^n_{\boldsymbol{\theta}|Y^n}\left(d\left(\boldsymbol{\theta},\theta^\circ\right)^2 \geq c_n\,\phi_n\right)\right] = 0;$$

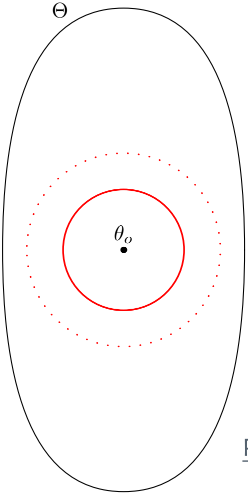
Reference: Castillo et al. [2008]



- A sequence $(\phi_n)_{n\in\mathbb{N}}$ is a contraction rate at θ° if

$$\forall (c_n), c_n \to \infty, \quad \lim_{n\to\infty} \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^n_{\boldsymbol{\theta}|Y^n}\left(d\left(\boldsymbol{\theta},\theta^\circ\right)^2 \geq c_n\,\phi_n\right)\right] = 0;$$

Reference: Castillo et al. [2008]



- A sequence $(\phi_n)_{n\in\mathbb{N}}$ is a contraction rate at θ° if

$$\forall (c_n), c_n \to \infty, \quad \lim_{n\to\infty} \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^n_{\boldsymbol{\theta}|Y^n}\left(d(\boldsymbol{\theta},\theta^\circ)^2 \geq c_n \phi_n\right)\right] = 0;$$

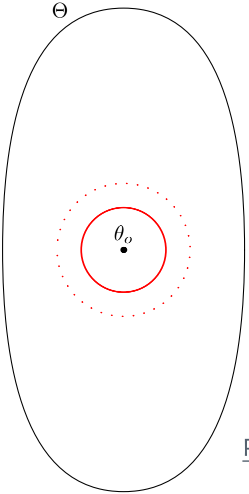
- and an exact contraction rate if in addition

$$\lim_{n\to\infty} \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^n_{\boldsymbol{\theta}|Y^n}\left(d(\boldsymbol{\theta},\theta^\circ)^2 \leq c_n^{-1} \phi_n\right)\right] = 0.$$

Reference: Castillo et al. [2008]

The frequentist Bayesian point of view

Contraction rate



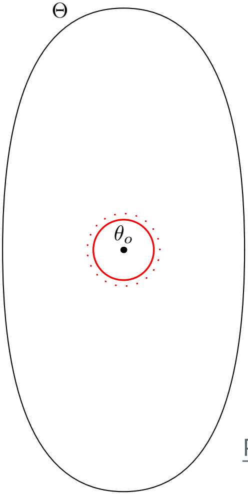
- A sequence $(\phi_n)_{n\in\mathbb{N}}$ is a contraction rate at θ° if

$$\forall\,(c_n)\,,c_n\to\infty,\qquad \lim_{n\to\infty}\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^n_{\boldsymbol{\theta}|Y^n}\left(d\left(\boldsymbol{\theta},\theta^\circ\right)^2\geq c_n\,\phi_n\right)\right]=0;$$

- and an exact contraction rate if in addition

$$\lim_{n\to\infty}\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^n_{\boldsymbol{\theta}|Y^n}\left(d\left(\boldsymbol{\theta},\theta^\circ\right)^2\leq c_n^{-1}\,\phi_n\right)\right]=0.$$

<u>Reference:</u> Castillo et al. [2008]



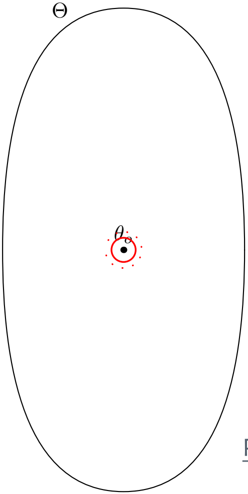
- A sequence $(\phi_n)_{n\in\mathbb{N}}$ is a contraction rate at θ° if

$$\forall (c_n), c_n \to \infty, \quad \lim_{n\to\infty} \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^n_{\boldsymbol{\theta}|Y^n}\left(d\left(\boldsymbol{\theta},\theta^\circ\right)^2 \geq c_n\,\phi_n\right)\right] = 0;$$

- and an exact contraction rate if in addition

$$\lim_{n\to\infty} \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^n_{\boldsymbol{\theta}|Y^n}\left(d\left(\boldsymbol{\theta},\theta^\circ\right)^2 \leq c_n^{-1}\,\phi_n\right)\right] = 0.$$

Reference: Castillo et al. [2008]



- A sequence $(\phi_n)_{n\in\mathbb{N}}$ is a contraction rate at θ° if

$$\forall\,(c_n),\,c_n\to\infty,\qquad \lim_{n\to\infty}\mathbb{E}^n_{\theta^{\circ}}\left[\mathbb{P}^n_{\boldsymbol{\theta}|Y^n}\left(d\left(\boldsymbol{\theta},\theta^{\circ}\right)^2\geq c_n\,\phi_n\right)\right]=0;$$

- and an exact contraction rate if in addition

$$\lim_{n\to\infty}\mathbb{E}^n_{\theta^{\circ}}\left[\mathbb{P}^n_{\boldsymbol{\theta}|Y^n}\left(d\left(\boldsymbol{\theta},\theta^{\circ}\right)^2\leq c_n^{-1}\,\phi_n\right)\right]=0.$$

Reference: Castillo et al. [2008]

- $(\Psi_n^\circ(\theta^\circ))_{n\in\mathbb{N}}$ is a lower bound at θ° for the contraction rate over the family of prior sequences $(\mathscr{G}_n)_{n\in\mathbb{N}}$ if:

$$\lim_{n\to\infty} \sup_{\mathbb{Q}_{\boldsymbol{\theta}}^n \in \mathscr{G}_n} \mathbb{E}_{\theta^\circ}^n \left[\mathbb{Q}_{\boldsymbol{\theta}|Y^n}^n \left(d^2\left(\theta^\circ, \boldsymbol{\theta}\right) \leq c_n^{-1} \Psi_n^\circ(\theta^\circ) \right) \right] = 0;$$

- it is called oracle rate over the family of priors if in addition there exists $\mathbb{P}_{\boldsymbol{\theta}}^n$ in $\mathscr{G}_n$ such that

$$\lim_{n\to\infty} \mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{\boldsymbol{\theta}|Y^n}^n \left(d^2\left(\theta^\circ, \boldsymbol{\theta}\right) \geq c_n \Psi_n^\circ(\theta^\circ) \right) \right] = 0.$$

Remark: these definitions generalize using the maximal risk, defining the uniform contraction rate and minimax optimal contraction rate.

Goal 1

Notations:

- $\Lambda_j := |\lambda_j|^{-2}$
- $\overline{\Lambda}_m := \frac{1}{m} \sum_{|j|=1}^{m} \Lambda_j$
- $b_m^2 := \sum_{|j|>m} |\theta_j^\circ|^2$

Likelihood:

$$\frac{L^n_{Y^n|\boldsymbol{\theta}}(y^n,\theta)}{L^n_{Y^n|\boldsymbol{\theta}}(y^n,0)} = \exp\left[-\frac{n}{2}\left(\sum_{j\in\mathcal{J}} \lambda_j^2\theta_j^2 + 2\sum_{j\in\mathcal{J}} \lambda_j\theta_j y_j + \Delta_1^n(\theta,y^n)\right)\right]$$

Goal:

- define a family of prior sequences indexed by a dimension parameter m_n;
- find conditions on Δ_1^n such that $\phi_n^{m_n}(\theta^\circ) := \left(\frac{m_n\overline{\Lambda}_{m_n}}{n} \vee b^2_{m_n}(\theta^\circ)\right)$ is a lower bound for the contraction rate of the posterior indexed by m_n;
- show that by selecting $m_n^\circ = \arg\min \phi_n^{m_n}(\theta^\circ)$ as the dimension parameter, $\Phi_n^\circ(\theta^\circ) := \min_{m_n} \phi_n^{m_n}(\theta^\circ)$ is a contraction rate
- construct an adaptive prior from this family

Iterated posteriors and self information carriers

In the spirit of Bunke and Johannes [2005], we then generate a posterior family by introducing an iteration parameter η:

- for $\eta = 1$, the prior distribution is $\mathbb{P}^{n,(1)}_{\boldsymbol{\theta}} = \mathbb{P}^{n}_{\boldsymbol{\theta}}$, the likelihood $\mathbb{P}^{n,(1)}_{Y^n|\boldsymbol{\theta}} = \mathbb{P}^{n}_{Y^n|\boldsymbol{\theta}}$ and the posterior distribution is $\mathbb{P}^{n,(1)}_{\boldsymbol{\theta}|Y^n} = \mathbb{P}^{n}_{\boldsymbol{\theta}|Y^n}$;

Iterated posteriors and self information carriers

In the spirit of BUNKE AND JOHANNES [2005], we then generate a posterior family by introducing an iteration parameter η:

- for $\eta = 1$, the prior distribution is $\mathbb{P}^{n,(1)}_{\boldsymbol{\theta}} = \mathbb{P}^{n}_{\boldsymbol{\theta}}$, the likelihood $\mathbb{P}^{n,(1)}_{Y^n|\boldsymbol{\theta}} = \mathbb{P}^{n}_{Y^n|\boldsymbol{\theta}}$ and the posterior distribution is $\mathbb{P}^{n,(1)}_{\boldsymbol{\theta}|Y^n} = \mathbb{P}^{n}_{\boldsymbol{\theta}|Y^n}$;
- for $\eta = 2$, we take the posterior for $\eta = 1$ as prior, hence, the prior is $\mathbb{P}^{n,(2)}_{\boldsymbol{\theta}} = \mathbb{P}^{n,(1)}_{\boldsymbol{\theta}|Y^n}$, the likelihood does not change $\mathbb{P}^{n,(2)}_{Y^n|\boldsymbol{\theta}} = \mathbb{P}^{n}_{Y^n|\boldsymbol{\theta}}$ and we compute the posterior with the same observations Y^n, which we note $\mathbb{P}^{n,(2)}_{\boldsymbol{\theta}|Y^n}$;

Iterated posteriors and self information carriers

In the spirit of Bunke and Johannes [2005], we then generate a posterior family by introducing an iteration parameter η:

- for $\eta = 1$, the prior distribution is $\mathbb{P}^{n,(1)}_{\boldsymbol{\theta}} = \mathbb{P}^{n}_{\boldsymbol{\theta}}$, the likelihood $\mathbb{P}^{n,(1)}_{Y^n|\boldsymbol{\theta}} = \mathbb{P}^{n}_{Y^n|\boldsymbol{\theta}}$ and the posterior distribution is $\mathbb{P}^{n,(1)}_{\boldsymbol{\theta}|Y^n} = \mathbb{P}^{n}_{\boldsymbol{\theta}|Y^n}$;
- for $\eta = 2$, we take the posterior for $\eta = 1$ as prior, hence, the prior is $\mathbb{P}^{n,(2)}_{\boldsymbol{\theta}} = \mathbb{P}^{n,(1)}_{\boldsymbol{\theta}|Y^n}$, the likelihood does not change $\mathbb{P}^{n,(2)}_{Y^n|\boldsymbol{\theta}} = \mathbb{P}^{n}_{Y^n|\boldsymbol{\theta}}$ and we compute the posterior with the same observations Y^n, which we note $\mathbb{P}^{n,(2)}_{\boldsymbol{\theta}|Y^n}$;
- ...

In the spirit of Bunke and Johannes [2005], we then generate a posterior family by introducing an iteration parameter η:

- for $\eta = 1$, the prior distribution is $\mathbb{P}^{n,(1)}_{\boldsymbol{\theta}} = \mathbb{P}^{n}_{\boldsymbol{\theta}}$, the likelihood $\mathbb{P}^{n,(1)}_{Y^n|\boldsymbol{\theta}} = \mathbb{P}^{n}_{Y^n|\boldsymbol{\theta}}$ and the posterior distribution is $\mathbb{P}^{n,(1)}_{\boldsymbol{\theta}|Y^n} = \mathbb{P}^{n}_{\boldsymbol{\theta}|Y^n}$;
- for $\eta = 2$, we take the posterior for $\eta = 1$ as prior, hence, the prior is $\mathbb{P}^{n,(2)}_{\boldsymbol{\theta}} = \mathbb{P}^{n,(1)}_{\boldsymbol{\theta}|Y^n}$, the likelihood does not change $\mathbb{P}^{n,(2)}_{Y^n|\boldsymbol{\theta}} = \mathbb{P}^{n}_{Y^n|\boldsymbol{\theta}}$ and we compute the posterior with the same observations Y^n, which we note $\mathbb{P}^{n,(2)}_{\boldsymbol{\theta}|Y^n}$;
- ...
- for any value of $\eta > 1$, the prior is $\mathbb{P}^{n,(\eta)}_{\boldsymbol{\theta}} = \mathbb{P}^{n,(\eta-1)}_{\boldsymbol{\theta}|Y^n}$ and we compute the posterior with the same likelihood $\mathbb{P}^{n,(\eta)}_{Y^n|\boldsymbol{\theta}} = \mathbb{P}^{n}_{Y^n|\boldsymbol{\theta}}$ and same observations Y^n which gives $\mathbb{P}^{n,(\eta)}_{\boldsymbol{\theta}|Y^n}$.

Goal 2

For our family of prior sequences, study the asymptotic with η:

- show that the posterior degenerates to a point mass

- give interpretation for the point mass

Remark: Bunke and Johannes [2005] show in some cases that the posterior concentrates around the generalized MLE.

Contents

Define:

- a sequence $(m_n)_{n\in\mathbb{N}}$, $\quad m_n \le n$;

Notations:

$\theta^\times$ prior mean;

s prior variance.

- the sequence of priors $\left(\mathbb{P}^n_{\boldsymbol{\theta}^{m_n}}\right)_{n\in\mathbb{N}}$ with density

$$p^n_{\boldsymbol{\theta}}(\theta) \propto \exp\left[-\frac{1}{2}\sum_{|j|\le m_n}\frac{\left|\theta_j^\times - \theta_j\right|^2}{s_j} + \Delta_2^n(\theta)\right]\prod_{|j|>m_n}\delta_{\theta_j^\times}(\theta_j);$$

- $\mathcal{G}_{m_n}$ is the family of priors generated by taking any sequence m_n.

References: Castillo et al. [2008], Rasmussen and Williams [2006], Johannes et al. [2016].

The posterior distribution is given by:

$$p^n_{\boldsymbol{\theta}^{m_n}|Y^n}(\theta, y^n) \propto \exp\left[-\frac{1}{2}\sum_{|j|\leq m_n}\frac{|\widehat{\theta}_{\eta,j}-\theta_j|^2}{\sigma^n_{\eta,j}} + \eta\Delta^n_2(\theta,y^n) + \Delta^n_1(\theta)\right].$$

Notations:

$$\widehat{\theta}_{\eta,j} := \frac{\frac{\theta^\times_j}{\eta n} + s_j y^n_j \lambda_j}{\frac{1}{\eta n} + s_j \lambda^2_j};$$

$$\sigma^n_{\eta,j} := \frac{s_j}{1 + \eta n s_j \lambda^2_j}.$$

Theorem Johannes & L. 2018: oracle contraction rate

Let Δ^n_2, Δ^n_1 be such that for any $0 \leq k_1, k_2 \leq 2$ and $j_1, j_2 \leq m_n$:

$$\mathbb{E}^n_{\boldsymbol{\theta}|Y^n}\left[\boldsymbol{\theta}^{k_1}_{j_1}\boldsymbol{\theta}^{k_2}_{j_2}\right] = \mathcal{O}_{\mathbb{P}^n_{\theta^\circ}}\left(\mathbb{E}_{\widetilde{\mathbb{P}}^n}\left[\boldsymbol{\theta}^{k_1}_{j_1}\boldsymbol{\theta}^{k_2}_{j_2}\right]\right)$$

Where $\widetilde{\mathbb{P}}^n = \mathcal{N}(\widehat{\theta}_\eta, \sigma^n_\eta)$.
Then $\mathbb{P}^m_{\boldsymbol{\theta}^{m_n}|Y^n}$ admits $\phi^{m_n}_n(\theta^\circ)$ as a contraction rate.

Consider the dimension parameter m as a random variable:

$$\mathbb{P}_M(\{M=m\}) = \frac{\exp[-\eta \cdot pen(m)]}{\sum_{k=1}^{n} \exp[-\eta \cdot pen(k)]};$$

$$\mathbb{P}^{n,(\eta)}_{M|Y^n} = \frac{\mathbb{P}_M}{\mathbb{P}_{Y^n}} \left(\int_\Theta \mathbb{P}_{Y^n|\boldsymbol{\theta}} \cdot \mathbb{P}_{\boldsymbol{\theta}|M} \, d\theta \right)^\eta;$$

$$\mathbb{P}^{n,(\eta)}_{\boldsymbol{\theta}^M|Y^n} = \sum_{m=1}^{n} \mathbb{P}^{n,(\eta)}_{\boldsymbol{\theta}|M=m,Y^n} \cdot \mathbb{P}^{n,(\eta)}_{M|Y^n}(\{M=m\});$$

$$\mathbb{E}^{n,(\eta)}_{\boldsymbol{\theta}^M|Y^n}[\boldsymbol{\theta}] = \sum_{m=1}^{n} \mathbb{E}^{n,(\eta)}_{\boldsymbol{\theta}|M=m,Y^n}[\boldsymbol{\theta}] \cdot \mathbb{P}^{n,(\eta)}_{M|Y^n}(\{M=m\}).$$

Theorem Johannes & L. 2016: Self informative Bayes carrier

As the number of iterations η tends to infinite, the distribution of $\boldsymbol{\theta}^M$ contracts around the projection estimator for which the dimension is given by the minimizer of the penalized contrast:

$$-\mu_m^{n,(1)}(y^n) + pen(m)$$

Where $\mu_m^{n,(1)}(y^n)$ is the log of the marginal of Y^n:

$$\exp\left[\mu_{m_n}^{n,(1)}(y^n)\right] \propto \int_{\theta\in\Theta} p_{\boldsymbol{\theta}|Y^n}(\theta, y^n)\, d\theta$$

Contents

The model and its likelihood

<u>Data:</u> $Y^n(x) = \int_0^x \lambda(\theta^\circ)(s)ds + \int_0^x \frac{1}{\sqrt{n}} dW(s), \qquad Y_j^n = \lambda_j \theta_j^\circ + \frac{1}{\sqrt{n}} \xi_j$

<u>Likelihood:</u> $\frac{L^n_{Y^n|\boldsymbol{\theta}}(y^n, \theta)}{L^n_{Y^n|\boldsymbol{\theta}}(y^n, 0)} = \exp\left[-\frac{n}{2} \sum_{j\in\mathbb{N}} \lambda_j^2 \theta_j^2 + n \sum_{j\in\mathbb{N}} \lambda_j \theta_j Y_j \right]$

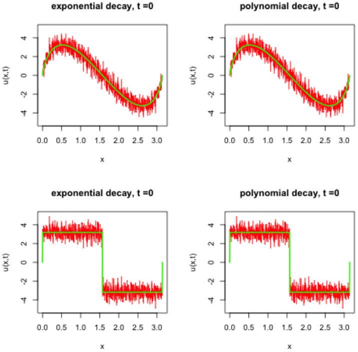
Notations:

$$\sum_{j\geq 1} \theta_j^\circ \lambda_j^t \cdot u_j(x);$$

$$\sum_{j\geq 1} Y_j \cdot u_j(x);$$

$$\lambda_j^t = \exp\left[-(j+1)^2 \cdot t \right]$$

$$\lambda_j^t = \exp\left[-2 \cdot \log(j+1) \cdot t \right]$$



Optimality results: oracle optimality

For any increasing and unbounded sequence $(c_n)_{n\in\mathbb{N}}$

Theorem Johannes & L. 2016: Oracle optimality

$$\lim_{n\to\infty} \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^n_{\boldsymbol{\theta}^{m_n^\circ}|Y^n}\left(d^2\left(\theta^\circ,\boldsymbol{\theta}^{m_n^\circ}\right) \geq c_n^{-1}\Phi_n^\circ(\theta^\circ)\right)\right] = 1;$$

$$\lim_{n\to\infty} \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^n_{\boldsymbol{\theta}^{m_n^\circ}|Y^n}\left(d^2\left(\theta^\circ,\boldsymbol{\theta}^{m_n^\circ}\right) \leq c_n\Phi_n^\circ(\theta^\circ)\right)\right] = 1;$$

$$\lim_{n\to\infty} \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^n_{\boldsymbol{\theta}^{M}|Y^n}\left(d^2\left(\theta^\circ,\boldsymbol{\theta}^{M}\right) \leq c_n\Phi_n^\circ(\theta^\circ)\right)\right] = 1;$$

Johannes et al. [2016] exhibit cases where $(c_n)_{n\in\mathbb{N}}$ can be replaced by a constant with $\eta = 1$. We generalized this property for any $\eta \in [2,\infty]$.

Contents

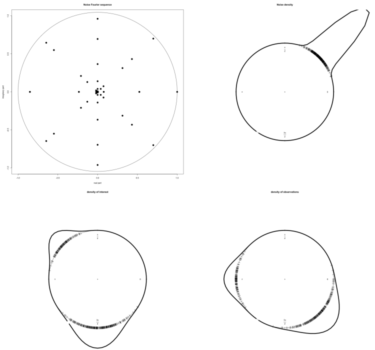
The model and its likelihood

Data:

$$Y^n = (Y_k)_{k\in[\![1,n]\!]} = (X_k + \epsilon_k)_k,$$
$$f^X(x) = \sum_{j\in\mathbb{Z}} \theta_j^\circ \exp[-2i\pi j\cdot]$$
$$f^\epsilon(x) = \sum_{j\in\mathbb{Z}} \lambda_j \exp[-2i\pi j\cdot]$$
$$f^Y(x) = \sum_{j\in\mathbb{Z}} \theta_j^\circ \lambda_j \exp[-2i\pi j\cdot]$$

Likelihood:

$$L^n_{Y^n|\boldsymbol{\theta}}(y^n,\theta) = \prod_{k=1}^{n}\left(\sum_{j\in\mathbb{Z}} \lambda_j \cdot \theta_j \cdot e_j(y_k)\right)$$



Optimality results

Assume: $\|f^Y\|_{l^1} \leq C < 2$,
we hence obtain

$$\Delta_1^n(\theta, y^n) = -2\left(1 + \sum_{k=1}^{n} \frac{\log(f^Y(y_k)) - f^Y(y_k)}{n}\right).$$

Goal: apply previous result by showing convergence of moments.

Reasons to hope for a positive result:

- existing results around equivalence of these experiments,
- proven optimality of the posterior mean estimator,
- promising simulations.

Further developments

- Consider $\mathbb{L}^p$-norms, $p \neq 2$;
- contemplate empirical Bayes for partially unknown operator;
- selecting prior variance sequence instead of dimension parameter.

Summary

- Family of Bayesian methods indexed by an iteration parameter;
- frequentist "model selection" method is a limit case;
- optimal contraction rate of the posterior distributions in Gaussian case;
- hints of good behavior for more general models;
- optimality of the estimators given by the posterior means for a larger class of models.