Poster title

Name
Supervisor

Background

Consider an indirect Gaussian sequence space model consisting of:

- an unknown parameter of interest $\left(\theta_j^\circ\right)_{j\in\mathbb{N}} = \theta^\circ$,
- a decreasing multiplicative sequence $\left(\lambda_j\right)_{j\in\mathbb{N}} = \lambda$ converging to 0,
- observations $\left(Y_j\right)_{j\in\mathbb{N}} = Y$, contaminated by an additive independent centered Gaussian noise with variance n^{-1},

$$Y = \left(\theta_j^\circ \cdot \lambda_j + \sqrt{n}^{-1} \cdot \xi_j\right)_{j\in\mathbb{N}}, \quad (\xi)_{j\in\mathbb{N}} \sim_{iid} \mathcal{N}(0,1).$$

The goal is to recover θ° and derive an upper bound.

The he model

For any index j, an unbiased estimator of θ_j° is Y_j/λ_j. Hence, an intuitive class of estimators are the projection estimators: $\tilde{\theta}^m = \left(Y_j/\lambda_j \mathbb{1}_{\{j\leq m\}}\right)_{j\in\mathbb{N}}$ with m in $\mathbb{N}$. The model selection method offers a data driven way to select m in this context:

$$G_n := \max\left\{1 \leq j \leq n : n^{-1}\lambda_j^{-2} \leq \lambda_1^{-2}\right\},$$

$$\widehat{m} := \underset{m\in[\![1,G_n]\!]}{\arg\min}\left\{3m - \sum_{j=1}^{m} Y_j^2\right\}, \qquad \widehat{\theta} := \left(\widehat{\theta}_j^{\widehat{m}}\right)_{j\in\mathbb{N}}.$$

It is shown in **?**, in the direct case, that this estimator is consistent, converges in probability and $\mathbb{L}^2$-norm, noted $\|\cdot\|$, with minimax optimal rate over some Sobolev ellipsoid:

$$\Theta^\circ := \Theta^\circ(\mathfrak{a}, L^\circ)\left\{\theta : \sum_{j=1}^{\infty}\frac{1}{\mathfrak{a}_j}\theta_j^2 < L^\circ\right\}.$$

Prior work

We adopt a Bayesian point of view:

- the parameter $\boldsymbol{\theta}$ is a random variable with prior $\mathbb{P}_{\boldsymbol{\theta}}$,
- given $\boldsymbol{\theta}$, the likelihood of Y is $\mathbb{P}^n_{Y|\boldsymbol{\theta}} = \mathcal{N}\left(\boldsymbol{\theta}\lambda, n^{-1}\mathbb{I}\right)$,
- we are interested in the posterior distribution $\mathbb{P}_{\theta^n|Y} \propto \mathbb{P}^n_{Y|\boldsymbol{\theta}} \cdot \mathbb{P}_{\boldsymbol{\theta}}$.

In the spirit of **?**, we then generate a posterior family by introducing an iteration parameter η:

- for $\eta = 1$, the prior distribution is $\mathbb{P}_{\boldsymbol{\theta}^1} = \mathbb{P}_{\boldsymbol{\theta}}$, the likelihood $\mathbb{P}^n_{Y^1|\boldsymbol{\theta}^1} = \mathbb{P}^n_{Y|\boldsymbol{\theta}}$ and the posterior distribution is $\mathbb{P}^n_{\boldsymbol{\theta}^1|Y^1} = \mathbb{P}^n_{\boldsymbol{\theta}|Y}$,
- for $\eta = 2$, we take the posterior for $\eta = 1$ as prior, hence, the prior distribution is $\mathbb{P}^n_{\boldsymbol{\theta}^2} = \mathbb{P}^n_{\boldsymbol{\theta}^1|Y^1}$, the likelihood is kept the same $\mathbb{P}^n_{Y^2|\boldsymbol{\theta}^2} = \mathbb{P}^n_{Y|\boldsymbol{\theta}}$ and we compute the posterior distribution with the same observations Y, which we note $\mathbb{P}^n_{\boldsymbol{\theta}^2|Y^2}$,
- ...
- for any value of $\eta > 1$, the prior is $\mathbb{P}^n_{\boldsymbol{\theta}^\eta} = \mathbb{P}^n_{\boldsymbol{\theta}^{\eta-1}|Y^{\eta-1}}$ and we compute the posterior with the same likelihood $\mathbb{P}_{Y^\eta|\boldsymbol{\theta}^\eta} = \mathbb{P}^n_{Y|\boldsymbol{\theta}}$ and same observation Y which gives $\mathbb{P}^n_{\boldsymbol{\theta}^\eta|Y^\eta}$.

This iteration procedure corresponds to giving more and more weight to the observations and make the prior knowledge vanish.

Within this framework we define the family of estimators:

$$\widehat{\theta}^{(\eta)} := \mathbb{E}^n_{\boldsymbol{\theta}^\eta|Y^\eta}[\boldsymbol{\theta}],$$

and call self-informative limit the limit of the estimate with $\eta \to \infty$.

We are interested in the behavior of the family $\left(\mathbb{P}^n_{\boldsymbol{\theta}^\eta|Y^\eta}\right)_{\eta\in\mathbb{N}^\star}$ as n and/or η tend to infinite.

In particular, the question of oracle and minimax concentration (resp. convergence) is answered for any element of the family of posterior distributions (resp. posterior means), including when η tends to infinite.

The problem

- Consider a random hyper-parameter M, with values in a subset of $\mathbb{N}$, acting like a threshold:

$$\forall j > m, \quad \mathbb{P}_{\boldsymbol{\theta}_j|M=m} = \delta_0,$$
$$\forall j \leq m, \quad \mathbb{P}_{\boldsymbol{\theta}_j|M=m} = \mathcal{N}(0,1).$$

- if we denote $\mathbb{P}_M$ the distribution of M (to be specified later), then

$$\mathbb{P}^n_{\boldsymbol{\theta}|Y} = \sum_{m\in\mathbb{N}} \mathbb{P}^n_{\boldsymbol{\theta}|M=m,Y} \cdot \mathbb{P}_{M=m|Y}{}^n.$$

- Hence, given M, the posterior is

$$\forall j > m, \quad \boldsymbol{\theta}_j|M = m, Y \sim \delta_0,$$
$$\forall j \leq m, \quad \boldsymbol{\theta}_j|M = m, Y \sim \mathcal{N}\left(\frac{Y_j \cdot n \cdot \lambda_j}{1 + n \cdot \lambda_j^2}, \frac{1}{1 + n \cdot \lambda_j^2}\right).$$

Remark: the family of hierarchical priors with deterministic threshold M is called family of sieve priors.

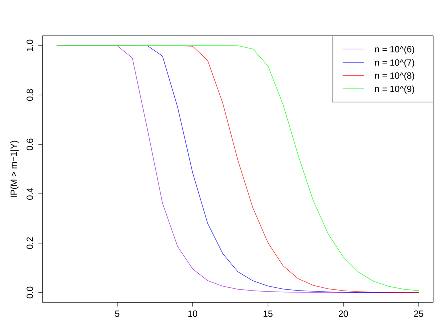


Figure: Survival function of M for different values of n

Main results

In **?**, under a pragmatic Bayesian point of view; that is, the existence of a true parameter θ° is accepted; it is shown that, by choosing $\mathbb{P}_M$ suitably:

- the estimator $\widehat{\theta}^{(1)}$ converges with,
  - oracle optimal rate for the quadratic risk which means, $\forall\theta^\circ \in \Theta^\circ, \exists C^\circ \in [1,\infty[: \forall n \in \mathbb{N}, \exists\Phi_n^\circ \in \mathbb{R}$:

$$\inf_{m\in\mathbb{N}} \mathbb{E}^n_{\theta^\circ}\left[\left\|\tilde{\theta}^m - \theta^\circ\right\|^2\right] \geq \Phi_n^\circ,$$
$$\mathbb{E}^n_{\theta^\circ}\left[\left\|\widehat{\theta}^{(1)} - \theta^\circ\right\|^2\right] \leq C^\circ\Phi_n^\circ;$$

  - minimax optimal rate for the maximal risk over Θ°, that is to say, $\exists C^\star \in [1,\infty[: \forall n \in \mathbb{N}, \exists\Phi_n^\star \in \mathbb{R}$:

$$\inf_{\tilde{\theta}} \sup_{\theta^\circ\in\Theta^\circ} \mathbb{E}^n_{\theta^\circ}\left[\left\|\tilde{\theta} - \theta^\circ\right\|^2\right] \geq \Phi_n^\star,$$
$$\sup_{\theta^\circ\in\Theta^\circ} \mathbb{E}^n_{\theta^\circ}\left[\left\|\widehat{\theta}^{(1)} - \theta^\circ\right\|^2\right] \leq C^\star\Phi_n^\star,$$

  where $\inf_{\tilde{\theta}}$ is taken over all possible estimators of θ°;

- the posterior distribution concentrates with,
  - oracle optimal rate for the quadratic loss which means, $\forall\theta^\circ \in \Theta^\circ, \exists K^\circ \in [1,\infty[$:

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

  - minimax optimal rate Θ°, that is to say, for any unbounded sequence $K_n \in \mathbb{R}^{\mathbb{N}}$:

$$\lim_{n\to\infty} \sup_{\theta^\circ\in\Theta^\circ} \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^n_{\boldsymbol{\theta}^1|Y^1}\left(\|\boldsymbol{\theta} - \theta^\circ\|^2 \leq K_n\Phi_n^\star\right)\right] = 1.$$

Iterated posterior distributions

Note that in the framework of our hierarchical prior, we have:

$$\mathbb{P}^n_{\boldsymbol{\theta}^\eta|Y^\eta} = \sum_{m\in\mathbb{N}} \mathbb{P}^n_{\boldsymbol{\theta}^\eta|M^\eta=m,Y^\eta} \cdot \mathbb{P}^n_{M^\eta=m|Y^\eta},$$
$$\widehat{\theta}^{(\eta)} = \left(\mathbb{E}^n_{\boldsymbol{\theta}^\eta|M^\eta\geq j,Y^\eta}[\boldsymbol{\theta}_j] \cdot \mathbb{P}^n_{M^\eta|Y^\eta}(M^\eta \geq j)\right)_{j\in\mathbb{N}}.$$

Hence, we first compute $\boldsymbol{\theta}_j^\eta|M^\eta, Y^\eta$:

$$\forall j \in \mathbb{N}, \quad \boldsymbol{\theta}_j^\eta|M^\eta \geq j, Y^\eta \sim \mathcal{N}\left(\frac{\eta \cdot Y_j \cdot n \cdot \lambda_j}{1 + \eta \cdot n \cdot \lambda_j^2}, \frac{1}{1 + n \cdot \eta \cdot \lambda_j^2}\right),$$
$$\boldsymbol{\theta}_j^\eta|M^\eta < j, Y^\eta \sim \delta_0;$$

and then fix the distribution of M^1: $\forall m \in [\![1, G_n]\!]$,

$$\mathbb{P}_{M^1}(M = m) \propto \exp\left(-3 \cdot \eta \cdot \frac{m}{2}\right) \cdot \prod_{j=1}^{m}\left(1 + n \cdot \eta \cdot \lambda_j^2\right)^2.$$

Which gives the family of posterior distributions:

$$\mathbb{P}^n_{M^\eta|Y^\eta}(m) \propto \exp\left[-\frac{\eta}{2}\left(3m - \sum_{j=1}^{m}\frac{\eta\left(Y_j \cdot n \cdot \lambda_j^2\right)^2}{1 + \eta \cdot n \cdot \lambda_j^2}\right)\right].$$

Publications and preprints

Consider the limit of the family of posteriors as η tends to infinite:

$$\lim_{\eta\to\infty} \mathbb{P}^n_{\boldsymbol{\theta}^\eta|M^\eta=m,Y^\eta} = \delta_{\tilde{\theta}^m},$$

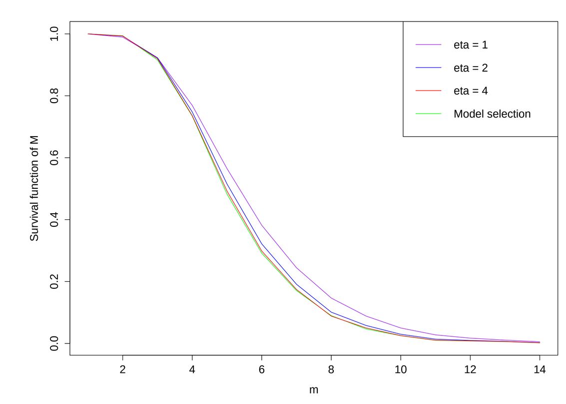
where $\tilde{\theta}^m$ is the projection estimator on the first m dimensions.
The distribution of M tends to a point mass:

$$\lim_{\eta\to\infty} \mathbb{P}^n_{M^\eta|Y^\eta} = \delta_{\widehat{m}},$$

where $\widehat{m}$ is the choice given by the frequentist model selection presented earlier.

The self-informative limit is equal to the model selection estimator, $\widehat{\theta}$, presented above.

Figure: Survival function of M for different values of η



Bibliography

Define the following quantities:

$$\mathfrak{b}_m := \sum_{j=m+1}^{\infty}(\theta^\circ)^2, \quad \Lambda_j := \lambda_j^{-2}, \quad m \cdot \overline{\Lambda}_m := \sum_{j=1}^{m}\Lambda_j,$$

$$m_n^\circ := \underset{m\in[\![1,G_n]\!]}{\arg\min}\left[\mathfrak{b}_m \vee n^{-1}m\overline{\Lambda}_m\right], \quad \Phi_n^\circ := \left[\mathfrak{b}_{m_n^\circ} \vee n^{-1}m_n^\circ\overline{\Lambda}_{m_n^\circ}\right],$$

$$m_n^\star := \underset{m\in[\![1,G_n]\!]}{\arg\min}\left[\mathfrak{a}_m \vee n^{-1}m\overline{\Lambda}_m\right], \quad \Phi_n^\star := \left[\mathfrak{a}_{m_n^\star} \vee n^{-1}m_n^\star\overline{\Lambda}_{m_n^\star}\right].$$

It is important to note that:

- $\Phi_n^\star$ is the minimax optimal rate over Θ°,
- Φ_n° is the oracle optimal rate over the projection estimators.

Personal Informations

Last degree before doctorate	MSc in Mathematical statistic, Rennes1 University, 1^{st} of July 2015
Member of the RTG	From the 5^{th} of October 2015 until the 5^{th} of October 2018
Date of doctoral degree	Anticipated, 5^{th} of October 2018
Occupation following doctorate	NA
Page of individual report in the report	?