Maxiset point of view for signal detection in inverse problems

F. Autin, M. Clausel, J.-M. Freyermuth et C. Marteau

Aix-Marseille Université

Rencontres de Statistique Avignon-Marseille 2019.

Physical domain :

$$Y = \mathbf{A}f + \varepsilon\xi.$$

$\mathbf{A} \ : \ H \to K$ compact operator, ξ Gaussian White Noise.

Sequential domain :

$$y_k = \mathbf{b_k}\theta_k + \varepsilon\xi_k, \quad k \in \mathbb{N}^*,$$

where $\mathbf{b}^2 := (\mathbf{b_k}^2)_k$ is the sequence of eigenvalues values of the operator $\mathbf{A}^\star\mathbf{A}$ and ξ_k are i.i.d $\mathcal{N}(0,1)$. Here the noise level $\varepsilon > 0$ is known.

Goal : we aim at detecting whether the signal f is the null function (hypothesis $\mathbf{H_0}$) or not (alternative hypothesis), using the observations $y = (y_k)_k$.

Definition (Hypotheses testing problems)

Indirect testing problem

$$\mathbf{H_0} : \theta = 0_{l_2(\mathbb{N}^*)} \quad \text{against} \quad \mathbf{H_1} : \theta \in \Theta, \ \|\theta\|^2 \geq r_\varepsilon^2.$$

Direct testing problem

$$\mathbf{H_0} : \theta = 0_{l_2(\mathbb{N}^*)} \quad \text{against} \quad \tilde{\mathbf{H}}_1 : \mathbf{b}.\theta \in \tilde{\Theta}, \ \|\mathbf{b}.\theta\|^2 \geq \mu_\varepsilon^2.$$

Above :

- $r = (r_\varepsilon)_{\varepsilon>0}$ and $\mu = (\mu_\varepsilon)_{\varepsilon>0}$ are decreasing and non negative sequences, called the chosen rates of detection,
- Θ and $\tilde{\Theta}$ are subset of $l_2(\mathbb{N}^*)$.

Signal detection in inverse problem: testing procedures

Definition (α-testing procedure)

Fix $0 < \alpha < 1$. An α-testing procedure $\Delta_{\alpha,\varepsilon} = \Delta_{\alpha,\varepsilon}(y)$ is a measurable function of $y = (y_k)_{k\in\mathbb{N}^*}$, such that $\Delta_{\alpha,\varepsilon} \in \{0,1\}$ and

$$\mathbb{P}_{0_{l_2(\mathbb{N}^*)}}(\Delta_{\alpha,\varepsilon} = 1) \leq \alpha.$$

Convention :

- $\Delta_{\alpha,\varepsilon} = 1$ means H_0 is rejected
- $\Delta_{\alpha,\varepsilon} = 0$ means H_0 is not rejected

For a given truncation that is an increasing sequence of non negative integers $(D_\varepsilon)_{\varepsilon>0}$.

Definition

Indirect testing Procedure : $\mathbf{b} = (\mathbf{b_k})_{\mathbf{k}\in\mathbb{N}^*}$ is known.

$$\Delta^{IP}_{\alpha,\varepsilon} = \mathbf{1}_{\{T_{D_\varepsilon} > t_{\alpha,\varepsilon}\}}, \text{ where } T_{D_\varepsilon} = \sum_{k=1}^{D_\varepsilon} \mathbf{b_k}^{-2} y_k^2$$

Direct testing Procedure : $\mathbf{b} = (\mathbf{b_k})_{\mathbf{k}\in\mathbb{N}^*}$ is unknown.

$$\Delta^{DP}_{\alpha,\varepsilon} = \mathbf{1}_{\{S_{D_\varepsilon} > s_{\alpha,\varepsilon}\}} \quad \text{where} \quad S_{D_\varepsilon} = \sum_{k=1}^{D_\varepsilon} y_k^2,$$

where $t_{\alpha,\varepsilon}$ and $s_{\alpha,\varepsilon}$ denote appropriate thresholds.

Let $(\alpha, \beta) \in (0,1)^2$ be respectively the Type-I and the Type-II errors. Fix a set $\Theta \subset l_2(\mathbb{N}^*)$. Let $\bullet \in \{IP, DP\}$.

Definition (β-separation rate for $\Delta^{\bullet}_{\alpha,\varepsilon}$)

The β-separation rates over Θ for $\Delta^{\bullet}_{\alpha,\varepsilon}$ is defined as :

$$R_\varepsilon(\Delta^{IP}_{\alpha,\varepsilon}, \Theta, \beta) = \inf\left\{ r_\varepsilon > 0, \sup_{\theta\in\Theta,\, \|\theta\|^2 \geq r_\varepsilon^2} \mathbb{P}_\theta(\Delta^{IP}_{\alpha,\varepsilon} = 0) \leq \beta \right\}.$$

$$R_\varepsilon(\Delta^{DP}_{\alpha,\varepsilon}, \tilde{\Theta}, \beta) = \inf\left\{ \mu_\varepsilon > 0, \sup_{\theta\in\tilde{\Theta},\, \|\mathbf{b}.\theta\|^2 \geq \mu_\varepsilon^2} \mathbb{P}_\theta(\Delta^{DP}_{\alpha,\varepsilon} = 0) \leq \beta \right\}.$$

The (α, β)-minimax separation rates associated to the testing problems are then defined as $r^\star = (r^\star_\varepsilon)_{\varepsilon>0}$ and $\mu^\star = (\mu^\star_\varepsilon)_{\varepsilon>0}$ where :

$$r^\star_\varepsilon := \inf_{\Delta_{\alpha,\varepsilon}} R_\varepsilon(\Delta_{\alpha,\varepsilon}, \Theta, \beta) \ \text{ and } \ \mu^\star_\varepsilon := \inf_{\Delta_{\alpha,\varepsilon}} R_\varepsilon(\Delta_{\alpha,\varepsilon}, \tilde{\Theta}, \beta)$$

Theorem (B. Laurent, J.-M. Loubes and C. Marteau (2010))

Consider mildly or severely ill-posed model. Then, with a good choice of $(D_\varepsilon)_\varepsilon$:

1. *in the Indirect testing problem (IP), $\Delta^{IP}_{\alpha,\varepsilon}$ is (α,β)-minimax optimal over of Besov balls,*
2. *in the Direct testing problem (DP), $\Delta^{DP}_{\alpha,\varepsilon}$ is (α,β)-minimax optimal over Besov ellipsoids.*
3. *Any (α,β)-minimax optimal testing procedure over ellipsoids balls for (DP) is (α,β)-minimax optimal procedure over ellipsoids balls for (IP). The converse are wrong.*

Corollary (B. Laurent, J.-M. Loubes and C. Marteau (2010))

According to the minimax sense, $\Delta^{DP}_{\alpha,\varepsilon}$ outperforms $\Delta^{IP}_{\alpha,\varepsilon}$ for the problem of signal detection.

Questions on the minimax point of view :

- Can other signals be detected by our testing procedures at the chosen (α, β)-minimax separation rates ?
- What about the performance of our testing procedures for other rates of detection ?

$\Longrightarrow$ Could maxiset approach tackle this problem ?

Definition (Maxiset of $\Delta_\alpha^\bullet$ for a chosen rate of detection)

Indirect testing problem

$$MS(\Delta_\alpha^{IP}, r, \beta) = \left\{\theta : \forall \varepsilon \in (0,1), \left[\|\theta\|^2 \geq r_\varepsilon^2 \Rightarrow \mathbb{P}_\theta\left[\Delta_{\alpha,\varepsilon}^{IP} = 0\right] \leq \beta\right]\right\}.$$

Direct testing problem

$$MS(\Delta_\alpha^{DP}, \mu, \beta) = \left\{\theta : \forall \varepsilon \in (0,1), \left[\|\mathbf{b}.\theta\|^2 \geq \mu_\varepsilon^2 \Rightarrow \mathbb{P}_\theta\left[\Delta_{\alpha,\varepsilon}^{DP} = 0\right] \leq \beta\right]\right\}.$$

Remark : The faster the chosen rates of detection $r = (r_\varepsilon)_\varepsilon$ and $\mu = (\mu_\varepsilon)_\varepsilon$ the thinner the maxisets.

Lemma (Indirect testing problem)

There exist $C_{\max}(\alpha,\beta)$ *and* $C_{\min}(\alpha,\beta)$ *such that, for all* $\varepsilon \in (0,1)$ *:*

$$(i) \qquad \sum_{k=1}^{D_\varepsilon} \theta_k^2 > C_{\max}(\alpha,\beta)\varepsilon^2 \sqrt{\sum_{k=1}^{D_\varepsilon} b_k^{-4}} \quad \Rightarrow \mathbb{P}_\theta(\Delta_{\alpha,\varepsilon}^{IP} = 0) \leq \beta,$$

$$(ii) \qquad \sum_{k=1}^{D_\varepsilon} \theta_k^2 \leq C_{\min}(\alpha,\beta)\varepsilon^2 \sqrt{\sum_{k=1}^{D_\varepsilon} b_k^{-4}} \quad \Rightarrow \mathbb{P}_\theta(\Delta_{\alpha,\varepsilon}^{IP} = 0) > \beta.$$

Remark : A consequence of (ii) is :

$$\|\theta\|^2 = C_{\min}(\alpha,\beta)\varepsilon_o^2 \sqrt{\sum_{k=1}^{D_{\varepsilon_o}} b_k^{-4}} \text{ for some } 0 < \varepsilon_o < 1 \Longrightarrow \mathbb{P}_\theta(\Delta_{\alpha,\varepsilon_o}^{IP} = 0) > \beta.$$

Lemma (Direct testing problem)

There exist $C'_{\min}(\alpha,\beta)$ *and* $C'_{\max}(\alpha,\beta)$ *such that, for all* $\varepsilon \in (0,1)$ *:*

$$(i) \qquad \sum_{k=1}^{D_\varepsilon} b_k^2\theta_k^2 > C'_{\max}(\alpha,\beta)\varepsilon^2\sqrt{D_\varepsilon} \quad \Rightarrow \mathbb{P}_\theta(\Delta^{DP}_{\alpha,\varepsilon}=0) \le \beta,$$

$$(ii) \qquad \sum_{k=1}^{D_\varepsilon} b_k^2\theta_k^2 \le C'_{\min}(\alpha,\beta)\varepsilon^2\sqrt{D_\varepsilon} \quad \Rightarrow \mathbb{P}_\theta(\Delta^{DP}_{\alpha,\varepsilon}=0) > \beta.$$

Remark : A consequence of (ii) is :

$$\|\mathbf{b}.\theta\|^2 = C'_{\min}(\alpha,\beta)\varepsilon_o^2\sqrt{D_{\varepsilon_o}} \text{ for some } 0<\varepsilon_o<1 \Longrightarrow \mathbb{P}_\theta(\Delta^{DP}_{\alpha,\varepsilon_o}=0) > \beta.$$

Comment :

- Poorly informative setting $\Rightarrow$ Robust version of Maxisets.

Definition (Robustness)

A set $\mathcal{H} \subset l_2(\mathbb{N}^*)$ is said to be robust if for any filter $h \in l_\infty(\mathbb{N}^*)$ with l_∞-norm smaller than 1 :

$$\theta \in \mathcal{H} \quad \Rightarrow \quad h.\theta \in \mathcal{H}.$$

Heuristic idea : focusing on detected signal θ or $\mathbf{b}.\theta$ with the smallest information.

Definition

For any $C > 0$ we introduce the following subset of $l_2(\mathbb{N}^*)$:

$$\mathcal{F}_{r,D}(C) = \left\{ \theta \in l_2(\mathbb{N}^*),\ \forall \varepsilon \in (0,1);\ \sum_{k>D_\varepsilon} \theta_k^2 < r_\varepsilon^2 - C\varepsilon^2 \sqrt{\sum_{k=1}^{D_\varepsilon} b_k^{-4}} \right\},$$

$$\mathcal{G}_{\mu,D}(C) = \left\{ \theta \in l_2(\mathbb{N}^*),\ \forall \varepsilon \in (0,1);\ \sum_{k>D_\varepsilon} b_k^2 \theta_k^2 < \mu_\varepsilon^2 - C\varepsilon^2 \sqrt{D_\varepsilon} \right\}.$$

- Maxisets are empty for fast rates of detection.

Corollary (Indirect testing problem)

$$MS(\Delta_\alpha^{IP}, r, \beta) \neq \emptyset \iff r_\varepsilon^2 \succeq \varepsilon^2 \sqrt{\sum_{k=1}^{D_\varepsilon} b_k^{-4}}, \quad \forall \varepsilon \in (0,1).$$

Corollary (Direct testing problem)

$$MS(\Delta_\alpha^{DP}, \mu, \beta) \neq \emptyset \iff \mu_\varepsilon^2 \succeq \varepsilon^2 \sqrt{D_\varepsilon}, \quad \forall \varepsilon \in (0,1).$$

Remarks :

- Direct testing Procedure can detect signals with small information while Indirect procedure can not.
- For the Indirect testing Procedure, the worse the operator, the further the horizon of detection.

Theorem

Consider the two testing procedures Δ_α^{IP} *and* Δ_α^{DP} *as defined previously.*
Denote as $MS^{filt}(\Delta_\alpha^{IP}, r, \beta)$ *and* $MS^{filt}(\Delta_\alpha^{DP}, r, \beta)$ *the respective* **robust maxisets** *associated with chosen rates* $r = (r_\varepsilon)_{\varepsilon>0}$ *and* $\mu = (\mu_\varepsilon)_{\varepsilon>0}$ *that are beyond the horizon of detection. Then :*

1. $MS^{filt}(\Delta_\alpha^{IP}, r, \beta) = \mathcal{F}_{r,D}$. *More precisely,*

$$\mathcal{F}_{r,D}(C_{\max}(\alpha,\beta)) \subset MS^{filt}(\Delta_\alpha^{IP}, r, \beta) \subset \mathcal{F}_{\sqrt{2}r,D}(C_{\min}(\alpha,\beta)),$$

2. $MS^{filt}(\Delta_\alpha^{DP}, \mu, \beta) = \mathcal{G}_{\mu,D}$. *More precisely,*

$$\mathcal{G}_{\mu,D}(C_{\max}(\alpha,\beta)) \subset MS^{filt}(\Delta_\alpha^{DP}, \mu, \beta) \subset \mathcal{G}_{\sqrt{2}\mu,D}(C_{\min}(\alpha,\beta)).$$

Corollary

Choose $r = (r_\varepsilon)_{\varepsilon>0}$, *and* $\mu = (\mu_\varepsilon)_{\varepsilon>0}$ *such that, for any* $\varepsilon \in (0,1)$, $\mu_\varepsilon \succeq \mathbf{b}_{\mathbf{D}_\varepsilon} r_\varepsilon$. *Then,*

$$MS(\Delta_\alpha^{IP}, r) \subset MS(\Delta_\alpha^{DP}, \mu).$$

According to the maxiset sense, $\Delta_{\alpha,\varepsilon}^{DP}$ *outperforms* $\Delta_{\alpha,\varepsilon}^{IP}$ *for the problem of signal detection.*

Concluding remarks :

We presented many reasons to prefer the Direct testing Procedure (DP) for signal detection in inverse problem :

1. looking at minimax performance (optimality),
2. looking at maxiset performance (horizon of detection, embedding).

Bibliography :

Maxiset point of view for signal detection in inverse problems.
F. Autin, M. Clausel, J.-M. Freyermuth and C. Marteau.
To appear in Math. Meth. of Statist.

Merci de votre attention