Estimation et intervalles de crédibilité pour le taux de reproduction de la Covid19 par échantillonnage Monte Carlo Langevin proximal

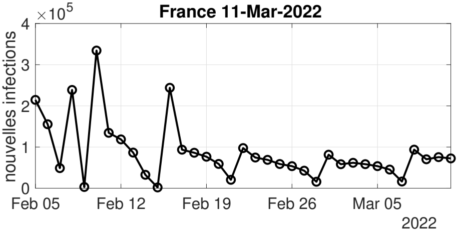
P. Abry[1,†], G. Fort[2,‡], B. Pascal[3], N. Pustelnik[1,4]

Colloque GRETSI 2022, Nancy

1. CNRS, ENS de Lyon, Laboratoire de Physique, France,
2. CNRS, Institut de Mathématique de Toulouse, France,
3. CNRS, Université de Lille, CRIStAL, France
4. UC Louvain, Belgium

† Soutenu partiellement par la bourse 80PRIME-2021 CNRS
‡ Financé en partie par la *Fondation Simone et Cino Del Duca, Institut de France*

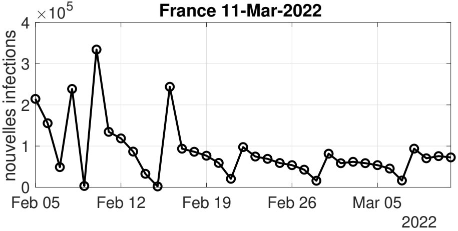
Décompte quotidien du nombre de nouveaux cas



données collectées
*par l'*Université Johns Hopkins
auprès des Agences de Santé Publique

Enjeux principaux de la surveillance épidémiologique

Décompte quotidien du nombre de nouveaux cas



données collectées
*par l'*Université Johns Hopkins
auprès des Agences de Santé Publique

Élaborer des mesures sanitaires adaptées et évaluer leur efficacité nécessite :

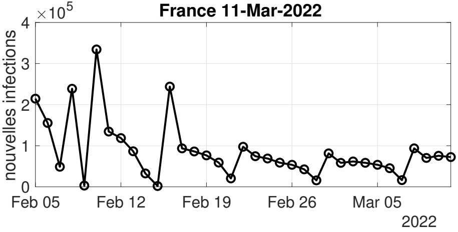
$\longrightarrow$ des outils de surveillance performants *modèle épidémiologique,*

$\longrightarrow$ robustes à la mauvaise qualité des données *gérer les valeurs aberrantes,*

$\longrightarrow$ accompagnés de niveaux de confiance fiables *intervalles de crédibilité.*

Enjeux principaux de la surveillance épidémiologique

Décompte quotidien du nombre de nouveaux cas



données collectées
*par l'*Université Johns Hopkins
auprès des Agences de Santé Publique

Élaborer des mesures sanitaires adaptées et évaluer leur efficacité nécessite :

- $\longrightarrow$ des outils de surveillance performants *modèle épidémiologique,*
- $\longrightarrow$ robustes à la mauvaise qualité des données *gérer les valeurs aberrantes,*
- $\longrightarrow$ accompagnés de niveaux de confiance fiables *intervalles de crédibilité.*

Indicateur clé : taux de reproduction R_0 (Liu et col, 2018, *PNAS*)

"nombre moyen de cas secondaires générés par une personne contagieuse typique"

$\Longrightarrow$ relaxé en un **taux de reproduction effectif** R_t au jour t

(Cori et col, 2013, *Am Journal of Epidemiology*)

Z_t : nouvelles infections au jour t

$$\mathbb{P}(Z_t|Z_1,\ldots,Z_{t-1}) = \text{Poisson}\left(p_t(\theta)\right), \quad p_t(\theta) = R_t \sum_{u=1}^{\tau_\phi} \Phi_u Z_{t-u} + O_t$$

$\Phi \Rightarrow$ délai aléatoire entre infections primaire et secondaire

Paramètres à estimer $\boldsymbol{\theta} = (\mathbf{R}, \mathbf{O})$

- $\mathbf{R} = (R_1, \cdots, R_T)$ les taux de reproduction ;
- $\mathbf{O} = (O_1, \cdots, O_T)$ valeurs aberrantes.

(Pascal et col, 2022, IEEE Trans Sig Process ; Artigas et col, 2022, EUSIPCO ; Fort et col, 2022, *preprint*)

Z_t : nouvelles infections au jour t

$$\mathbb{P}(\mathrm{Z}_t|\mathrm{Z}_1,\ldots,\mathrm{Z}_{t-1}) = \mathrm{Poisson}\left(p_t(\theta)\right), \quad p_t(\theta) = \mathrm{R}_t \sum_{u=1}^{\tau_\phi} \Phi_u \mathrm{Z}_{t-u} + \mathrm{O}_t$$

$\Phi \Rightarrow$ délai aléatoire entre infections primaire et secondaire

Paramètres à estimer $\boldsymbol{\theta} = (\mathbf{R}, \mathbf{O})$

- $\mathbf{R} = (\mathrm{R}_1, \cdots, \mathrm{R}_T)$ les taux de reproduction ;
- $\mathbf{O} = (\mathrm{O}_1, \cdots, \mathrm{O}_T)$ valeurs aberrantes.

(Pascal et col, 2022, IEEE Trans Sig Process ; Artigas et col, 2022, EUSIPCO ; Fort et col, 2022, *preprint*)

Log-vraisemblance du modèle de Poisson

$$-f(\boldsymbol{\theta}) := \begin{cases} \sum_{t=1}^{T} \left(\mathrm{Z}_t \ln p_t(\boldsymbol{\theta}) - p_t(\boldsymbol{\theta})\right) & \text{si } \boldsymbol{\theta} \in \mathcal{D} = \{\theta \,|\, \forall t,\ p_t(\theta) \geq 0\}, \\ -\infty & \text{sinon.} \end{cases}$$

Z_t : nouvelles infections au jour t

$$\mathbb{P}(\mathrm{Z}_t|\mathrm{Z}_1,\ldots,\mathrm{Z}_{t-1}) = \text{Poisson}\left(p_t(\theta)\right), \quad p_t(\theta) = \mathrm{R}_t \sum_{u=1}^{\tau_\phi} \Phi_u \mathrm{Z}_{t-u} + \mathrm{O}_t$$

$\Phi \Rightarrow$ délai aléatoire entre infections primaire et secondaire

Paramètres à estimer $\quad \boldsymbol{\theta} = (\mathbf{R}, \mathbf{O})$

– $\mathbf{R} = (\mathrm{R}_1, \cdots, \mathrm{R}_T)$ les taux de reproduction ;

– $\mathbf{O} = (\mathrm{O}_1, \cdots, \mathrm{O}_T)$ valeurs aberrantes.

(Pascal et col, 2022, IEEE Trans Sig Process ; Artigas et col, 2022, EUSIPCO ; Fort et col, 2022, *preprint*)

Log-vraisemblance du modèle de Poisson

$$-f(\boldsymbol{\theta}) := \begin{cases} \sum_{t=1}^{T} \left(\mathrm{Z}_t \ln p_t(\boldsymbol{\theta}) - p_t(\boldsymbol{\theta})\right) & \text{si } \boldsymbol{\theta} \in \mathcal{D} = \{\theta \,|\, \forall t,\ p_t(\theta) \geq 0\}, \\ -\infty & \text{sinon.} \end{cases}$$

Log-distribution *a priori*

- $\mathrm{R}_t - 2\mathrm{R}_{t-1} + \mathrm{R}_{t-2} \sim \text{Laplace}(\lambda_{\mathrm{R}})$
- $\mathrm{O}_t \sim \text{Laplace}(\lambda_{\mathrm{O}})$

$$\Longrightarrow -g(\mathrm{A}\boldsymbol{\theta}) = -\lambda_{\mathrm{R}} \|\mathrm{D}_2 \mathbf{R}\|_1 - \lambda_{\mathrm{O}} \|\mathbf{O}\|_1, \qquad \mathrm{D}_2 \in \mathbb{R}^{T-2\times T} \text{ : matrice du laplacien}$$

Z_t : nouvelles infections au jour t

$$\mathbb{P}(Z_t | Z_1, \ldots, Z_{t-1}) = \text{Poisson}\left(p_t(\theta)\right), \quad p_t(\theta) = R_t \sum_{u=1}^{\tau_\phi} \Phi_u Z_{t-u} + O_t$$

$\Phi \Rightarrow$ délai aléatoire entre infections primaire et secondaire

Paramètres à estimer $\quad \boldsymbol{\theta} = (\mathbf{R}, \mathbf{O})$

- $\mathbf{R} = (R_1, \cdots, R_T)$ les taux de reproduction ;
- $\mathbf{O} = (O_1, \cdots, O_T)$ valeurs aberrantes.

(Pascal et col, 2022, IEEE Trans Sig Process ; Artigas et col, 2022, EUSIPCO ; Fort et col, 2022, *preprint*)

Log-vraisemblance du modèle de Poisson

$$-f(\boldsymbol{\theta}) := \begin{cases} \sum_{t=1}^{T} \left(Z_t \ln p_t(\boldsymbol{\theta}) - p_t(\boldsymbol{\theta})\right) & \text{si } \boldsymbol{\theta} \in \mathcal{D} = \{\theta \,|\, \forall t,\ p_t(\theta) \geq 0\}, \\ -\infty & \text{sinon.} \end{cases}$$

Log-distribution *a priori*

- $R_t - 2R_{t-1} + R_{t-2} \sim \text{Laplace}(\lambda_{\mathrm{R}})$
- $O_t \sim \text{Laplace}(\lambda_{\mathrm{O}})$

$$\Longrightarrow -g(\mathrm{A}\boldsymbol{\theta}) = -\lambda_{\mathrm{R}} \|\mathrm{D}_2 \mathbf{R}\|_1 - \lambda_{\mathrm{O}} \|\mathbf{O}\|_1, \qquad \mathrm{D}_2 \in \mathbb{R}^{T-2 \times T} \ : \text{matrice du laplacien}$$

$$\mathrm{A} = \begin{pmatrix} \mathrm{D}_2 & 0 \\ 0 & \mathrm{I} \end{pmatrix} \in \mathbb{R}^{(2T-2) \times 2T} \Rightarrow \overline{\mathrm{A}} = \begin{pmatrix} \overline{\mathrm{D}}_2 & 0 \\ 0 & \mathrm{I} \end{pmatrix} \in \mathbb{R}^{2T \times 2T}, \ \textbf{inversible}$$

(Fort et col, 2022, *preprint*)

Formalisme bayésien pour la propagation de la Covid19

estimer $\{\mathrm{R}_t,\ t = 1, \ldots, T\} \equiv$ échantillonner sous une loi a posteriori[†] de la forme :

$$\pi(\boldsymbol{\theta}) \propto \exp\left(-f(\boldsymbol{\theta}) - g(\mathrm{A}\boldsymbol{\theta})\right) \mathbb{1}_{\mathcal{D}}(\boldsymbol{\theta})$$

(Artigas et col, 2022, EUSIPCO ; Fort et col, 2022, *preprint*)

[†] π est définie à une constante de normalisation près.

Formalisme bayésien pour la propagation de la Covid19

estimer $\{\mathrm{R}_t,\ t = 1, \ldots, T\} \equiv$ échantillonner sous une loi a posteriori[†] de la forme :

$$\pi(\boldsymbol{\theta}) \propto \exp\left(-f(\boldsymbol{\theta}) - g(\mathrm{A}\boldsymbol{\theta})\right) \mathbb{1}_{\mathcal{D}}(\boldsymbol{\theta})$$

- $\boldsymbol{\theta} \in \mathbb{R}^d$ vecteur des paramètres,
- f différentiable,
- $\mathcal{D} \subset \mathbb{R}^d$ domaine admissible,
- g convexe, non différentiable,
- $\mathrm{A} \in \mathbb{R}^{c \times d}$ opérateur linéaire,

  $c \leq d$, A de rang plein.

(Artigas et col, 2022, EUSIPCO ; Fort et col, 2022, *preprint*)

[†] π est définie à une constante de normalisation près.

Formalisme bayésien pour la propagation de la Covid19

estimer $\{\mathrm{R}_t,\ t = 1, \dots, T\} \equiv$ échantillonner sous une loi a posteriori† de la forme :

$$\pi(\boldsymbol{\theta}) \propto \exp\left(-f(\boldsymbol{\theta}) - g(\mathrm{A}\boldsymbol{\theta})\right) \mathbb{1}_{\mathcal{D}}(\boldsymbol{\theta})$$

- $\boldsymbol{\theta} \in \mathbb{R}^d$ vecteur des paramètres,
- f différentiable,
- $\mathcal{D} \subset \mathbb{R}^d$ domaine admissible,
- g convexe, non différentiable,
- $\mathrm{A} \in \mathbb{R}^{c \times d}$ opérateur linéaire,

  $c \leq d$, A de rang plein.

(Artigas et col, 2022, EUSIPCO ; Fort et col, 2022, *preprint*)

Méthode de Monte Carlo par chaîne de Markov

† π est définie à une constante de normalisation près.

Formalisme bayésien pour la propagation de la Covid19

estimer $\{\mathrm{R}_t,\ t = 1, \dots, T\} \equiv$ échantillonner sous une loi a posteriori[†] de la forme :

$$\pi(\boldsymbol{\theta}) \propto \exp\left(-f(\boldsymbol{\theta}) - g(\mathrm{A}\boldsymbol{\theta})\right) \mathbb{1}_{\mathcal{D}}(\boldsymbol{\theta})$$

- $\boldsymbol{\theta} \in \mathbb{R}^d$ vecteur des paramètres,
- f différentiable,
- $\mathcal{D} \subset \mathbb{R}^d$ domaine admissible,
- g convexe, non différentiable,
- $\mathrm{A} \in \mathbb{R}^{c \times d}$ opérateur linéaire, $c \leq d$, A de rang plein.

(Artigas et col, 2022, EUSIPCO ; Fort et col, 2022, *preprint*)

Méthode de Monte Carlo par chaîne de Markov

1) générer une suite aléatoire $\{\boldsymbol{\theta}^n,\ n \in \mathbb{N}\}$ telle que

- $\boldsymbol{\theta}^{n+1}$ ne dépend que de $\boldsymbol{\theta}^n$,
- à convergence, i.e., lorsque $n \to +\infty$, $\boldsymbol{\theta}^n \sim \pi$,

[†] π est définie à une constante de normalisation près.

Formalisme bayésien pour la propagation de la Covid19

estimer $\{\mathrm{R}_t,\ t = 1, \ldots, T\} \equiv$ échantillonner sous une loi a posteriori[†] de la forme :

$$\pi(\boldsymbol{\theta}) \propto \exp\left(-f(\boldsymbol{\theta}) - g(\mathrm{A}\boldsymbol{\theta})\right) \mathbb{1}_{\mathcal{D}}(\boldsymbol{\theta})$$

- $\boldsymbol{\theta} \in \mathbb{R}^d$ vecteur des paramètres,
- f différentiable,
- $\mathcal{D} \subset \mathbb{R}^d$ domaine admissible,
- g convexe, non différentiable,
- $\mathrm{A} \in \mathbb{R}^{c \times d}$ opérateur linéaire,
  $c \leq d$, A de rang plein.

(Artigas et col, 2022, EUSIPCO ; Fort et col, 2022, *preprint*)

Méthode de Monte Carlo par chaîne de Markov

1) générer une suite aléatoire $\{\boldsymbol{\theta}^n,\ n \in \mathbb{N}\}$ telle que
   - $\boldsymbol{\theta}^{n+1}$ ne dépend que de $\boldsymbol{\theta}^n$,
   - à convergence, i.e., lorsque $n \to +\infty$, $\boldsymbol{\theta}^n \sim \pi$,

2) calculer les estimateurs bayésiens, e.g., les intervalles de crédibilité,

à partir des échantillons $\{\boldsymbol{\theta}^n,\ n \geq N\}$ pour $N \gg 1$

[†] π est définie à une constante de normalisation près.

Algorithme de type Hastings-Metropolis

1) proposition gaussienne : $\boldsymbol{\theta}^{n+\frac{1}{2}} = \mu(\boldsymbol{\theta}^n) + \sqrt{2\gamma}\xi^{n+1}, \quad \xi^{n+1} \sim \mathcal{N}_d(0, \mathrm{C})$;

affranchie de la contrainte $\boldsymbol{\theta} \in \mathcal{D}$;

Algorithme de type Hastings-Metropolis

1) <u>proposition gaussienne :</u> $\boldsymbol{\theta}^{n+\frac{1}{2}} = \mu(\boldsymbol{\theta}^n) + \sqrt{2\gamma}\xi^{n+1}, \quad \xi^{n+1} \sim \mathcal{N}_d(0, \mathrm{C})$;

affranchie de la contrainte $\boldsymbol{\theta} \in \mathcal{D}$;

2) <u>acceptation-rejet :</u> $\boldsymbol{\theta}^{n+1} = \boldsymbol{\theta}^{n+\frac{1}{2}}$ aléatoirement, et seulement si $\boldsymbol{\theta}^{n+\frac{1}{2}} \in \mathcal{D}$;

$\boldsymbol{\theta}^{n+1} = \boldsymbol{\theta}^n$, sinon. (Kent, 1978, *Adv Appl Probab*)

Algorithme de type Hastings-Metropolis

1) proposition gaussienne : $\boldsymbol{\theta}^{n+\frac{1}{2}} = \mu(\boldsymbol{\theta}^n) + \sqrt{2\gamma}\xi^{n+1}, \quad \xi^{n+1} \sim \mathcal{N}_d(0, \mathrm{C})$;

affranchie de la contrainte $\boldsymbol{\theta} \in \mathcal{D}$;

2) acceptation-rejet : $\boldsymbol{\theta}^{n+1} = \boldsymbol{\theta}^{n+\frac{1}{2}}$ aléatoirement, et seulement si $\boldsymbol{\theta}^{n+\frac{1}{2}} \in \mathcal{D}$;

$\boldsymbol{\theta}^{n+1} = \boldsymbol{\theta}^n$, sinon. (Kent, 1978, *Adv Appl Probab*)

Cas π régulière : dynamique de Langevin (Roberts & Tweedie, 1996, *Bernoulli*)

$$\mu(\boldsymbol{\theta}) = \boldsymbol{\theta} + \gamma\nabla \ln \pi(\boldsymbol{\theta}), \ \gamma > 0$$

$\Longrightarrow$ déplacement vers les régions de forte probabilité

Algorithme de type Hastings-Metropolis

1) proposition gaussienne : $\boldsymbol{\theta}^{n+\frac{1}{2}} = \mu(\boldsymbol{\theta}^n) + \sqrt{2\gamma}\xi^{n+1}, \quad \xi^{n+1} \sim \mathcal{N}_d(0, \mathrm{C})$;

affranchie de la contrainte $\boldsymbol{\theta} \in \mathcal{D}$;

2) acceptation-rejet : $\boldsymbol{\theta}^{n+1} = \boldsymbol{\theta}^{n+\frac{1}{2}}$ aléatoirement, et seulement si $\boldsymbol{\theta}^{n+\frac{1}{2}} \in \mathcal{D}$;

$\boldsymbol{\theta}^{n+1} = \boldsymbol{\theta}^n$, sinon. (Kent, 1978, *Adv Appl Probab*)

Cas π régulière : dynamique de Langevin (Roberts & Tweedie, 1996, *Bernoulli*)

$$\mu(\boldsymbol{\theta}) = \boldsymbol{\theta} + \gamma\nabla \ln \pi(\boldsymbol{\theta}), \ \gamma > 0$$

$\Longrightarrow$ déplacement vers les régions de forte probabilité

Cas π non différentiable : Langevin proximal $\pi \propto \exp\left(-f - g(\mathrm{A}\cdot)\right) \mathbb{1}_{\mathcal{D}}$

- f différentiable de gradient ∇f,
- g non-différentiable, convexe,

  d'opérateur proximal $\mathrm{prox}_{\rho g} = (\mathrm{I} + \rho\partial g)^{-1}$, $\rho > 0$ connu explicitement.

But : comparaison de différentes approches proximales pour la construction de μ.

Objectif : terme de dérive $\mu(\boldsymbol{\theta})$ adapté à $\pi \propto \exp\left(-f - g(\mathsf{A}\cdot)\right) \mathbb{1}_{\mathcal{D}}$, g non régulière.

$$\operatorname{prox}_{\gamma g(\mathsf{A}\cdot)}(\boldsymbol{\theta}) = \underset{\boldsymbol{\varphi} \in \mathbb{R}^d}{\operatorname{argmin}} \left(\frac{1}{2} \|\boldsymbol{\theta} - \boldsymbol{\varphi}\|_2^2 + \gamma g(\mathsf{A}\boldsymbol{\varphi}) \right)$$

Objectif : terme de dérive $\mu(\boldsymbol{\theta})$ adapté à $\pi \propto \exp\left(-f - g(\mathsf{A}\cdot)\right) \mathbb{1}_{\mathcal{D}}$, g non régulière.

$$\mathrm{prox}_{\gamma g(\mathsf{A}\cdot)}(\boldsymbol{\theta}) = \underset{\boldsymbol{\varphi} \in \mathbb{R}^d}{\mathrm{argmin}} \left(\frac{1}{2} \|\boldsymbol{\theta} - \boldsymbol{\varphi}\|_2^2 + \gamma g(\mathsf{A}\boldsymbol{\varphi}) \right)$$

Méthodes primales

Objectif : terme de dérive $\mu(\boldsymbol{\theta})$ adapté à $\pi \propto \exp\left(-f - g(\mathsf{A}\cdot)\right)\mathbb{1}_{\mathcal{D}}$, g non régulière.

$$\operatorname{prox}_{\gamma g(\mathsf{A}\cdot)}(\boldsymbol{\theta}) = \underset{\boldsymbol{\varphi}\in\mathbb{R}^d}{\operatorname{argmin}} \left(\frac{1}{2}\|\boldsymbol{\theta} - \boldsymbol{\varphi}\|_2^2 + \gamma g(\mathsf{A}\boldsymbol{\varphi})\right)$$

Méthodes primales

- Dérive de Moreau : *approximation régulière de g par son enveloppe de Moreau*

$$\mu^{\mathtt{M}}(\boldsymbol{\theta}) = \boldsymbol{\theta} - \gamma\nabla f(\boldsymbol{\theta}) - \frac{\gamma}{\rho}\mathsf{A}^\top(\mathsf{I} - \operatorname{prox}_{\rho g})\mathsf{A}\boldsymbol{\theta}, \quad \rho = \gamma$$

(Durmus et col, 2018, *SIAM J Imaging Sci* ; Luu et col, 2020, *Methodol Comput Appl Probab*)

Objectif : terme de dérive $\mu(\boldsymbol{\theta})$ adapté à $\pi \propto \exp\left(-f - g(\mathsf{A}\cdot)\right) \mathbb{1}_{\mathcal{D}}$, g non régulière.

$$\mathrm{prox}_{\gamma g(\mathsf{A}\cdot)}(\boldsymbol{\theta}) = \operatorname*{argmin}_{\boldsymbol{\varphi} \in \mathbb{R}^d} \left(\frac{1}{2} \|\boldsymbol{\theta} - \boldsymbol{\varphi}\|_2^2 + \gamma g(\mathsf{A}\boldsymbol{\varphi}) \right)$$

Méthodes primales

- Dérive de Moreau : *approximation régulière de g par son enveloppe de Moreau*

$$\mu^{\mathtt{M}}(\boldsymbol{\theta}) = \boldsymbol{\theta} - \gamma \nabla f(\boldsymbol{\theta}) - \frac{\gamma}{\rho} \mathsf{A}^\top (\mathsf{I} - \mathrm{prox}_{\rho g}) \mathsf{A}\boldsymbol{\theta}, \quad \rho = \gamma$$

(Durmus et col, 2018, *SIAM J Imaging Sci* ; Luu et col, 2020, *Methodol Comput Appl Probab*)

- Dérive PGdec : si $\mathsf{A}\mathsf{A}^\top = \nu \mathsf{I}$, avec $\nu > 0 \Longrightarrow$ forme explicite de $\mathrm{prox}_{\gamma g(\mathsf{A}\cdot)}$

$$\mu^{\mathtt{PGdec}}(\boldsymbol{\theta}) = \mathrm{prox}_{\gamma\, g(\mathsf{A}\cdot)} \left(\boldsymbol{\theta} - \gamma\, \nabla f(\boldsymbol{\theta})\right)$$

étendu à $g(\mathsf{A}\cdot) = \sum_{i=1}^{I} g_i(\mathsf{A}_i\cdot)$, avec $\mathsf{A}_i\mathsf{A}_i^\top = \nu_i \mathsf{I}$, $\nu_i > 0$ (Fort et col, 2022, *preprint*)

Objectif : terme de dérive $\mu(\boldsymbol{\theta})$ adapté à $\pi \propto \exp\left(-f - g(\mathrm{A}\cdot)\right) \mathbb{1}_{\mathcal{D}}$, g non régulière.

$$\operatorname{prox}_{\gamma g(\mathrm{A}\cdot)}(\boldsymbol{\theta}) = \underset{\boldsymbol{\varphi} \in \mathbb{R}^d}{\operatorname{argmin}} \left(\frac{1}{2} \|\boldsymbol{\theta} - \boldsymbol{\varphi}\|_2^2 + \gamma g(\mathrm{A}\boldsymbol{\varphi}) \right)$$

Méthodes primales

- Dérive de Moreau : *approximation régulière de g par son enveloppe de Moreau*

$$\mu^{\mathrm{M}}(\boldsymbol{\theta}) = \boldsymbol{\theta} - \gamma \nabla f(\boldsymbol{\theta}) - \frac{\gamma}{\rho} \mathrm{A}^\top (\mathrm{I} - \operatorname{prox}_{\rho g}) \mathrm{A}\boldsymbol{\theta}, \quad \rho = \gamma$$

(Durmus et col, 2018, *SIAM J Imaging Sci* ; Luu et col, 2020, *Methodol Comput Appl Probab*)

- Dérive PGdec : si $\mathrm{A}\mathrm{A}^\top = \nu \mathrm{I}$, avec $\nu > 0 \Longrightarrow$ forme explicite de $\operatorname{prox}_{\gamma g(\mathrm{A}\cdot)}$

$$\mu^{\mathrm{PGdec}}(\boldsymbol{\theta}) = \operatorname{prox}_{\gamma\, g(\mathrm{A}\cdot)} \left(\boldsymbol{\theta} - \gamma\, \nabla f(\boldsymbol{\theta})\right)$$

étendu à $g(\mathrm{A}\cdot) = \sum_{i=1}^{I} g_i(\mathrm{A}_i\cdot)$, avec $\mathrm{A}_i \mathrm{A}_i^\top = \nu_i \mathrm{I}$, $\nu_i > 0$ (Fort et col, 2022, *preprint*)

- Dérive marche aléatoire : $\mu^{\mathrm{RM}}(\boldsymbol{\theta}) = \boldsymbol{\theta}$

Monte Carlo Langevin proximal

Objectif : terme de dérive $\mu(\boldsymbol{\theta})$ adapté à $\pi \propto \exp\left(-f - g(\mathsf{A}\cdot)\right) \mathbb{1}_{\mathcal{D}}$, g non régulière.

Objectif : terme de dérive $\mu(\boldsymbol{\theta})$ adapté à $\pi \propto \exp\left(-f - g(\mathrm{A}\cdot)\right) \mathbb{1}_{\mathcal{D}}$, g non régulière.

$\mathrm{A} \in \mathbb{R}^{c \times d}$, $c \leq d$ de **rang plein** $\Longrightarrow$ extension **inversible** $\overline{\mathrm{A}} \in \mathbb{R}^{d \times d}$

terme de dérive dual $\tilde{\mu}(\tilde{\boldsymbol{\theta}})$, $\tilde{\boldsymbol{\theta}} = \overline{\mathrm{A}}\boldsymbol{\theta}$, adapté à $\tilde{\pi} \propto \exp\left(-f(\overline{\mathrm{A}}^{-1}\cdot) - \overline{g}\right) \mathbb{1}_{\mathcal{D}}(\overline{\mathrm{A}}^{-1}\cdot)$

Monte Carlo Langevin proximal

Objectif : terme de dérive $\mu(\boldsymbol{\theta})$ adapté à $\pi \propto \exp\left(-f - g(\mathrm{A}\cdot)\right) \mathbb{1}_{\mathcal{D}}$, g non régulière.

$\mathrm{A} \in \mathbb{R}^{c \times d}$, $c \leq d$ de **rang plein** $\Longrightarrow$ extension **inversible** $\overline{\mathrm{A}} \in \mathbb{R}^{d \times d}$

terme de dérive dual $\tilde{\mu}(\tilde{\boldsymbol{\theta}})$, $\tilde{\boldsymbol{\theta}} = \overline{\mathrm{A}}\boldsymbol{\theta}$, adapté à $\tilde{\pi} \propto \exp\left(-f(\overline{\mathrm{A}}^{-1}\cdot) - \overline{g}\right) \mathbb{1}_{\mathcal{D}}(\overline{\mathrm{A}}^{-1}\cdot)$

Méthodes duales

Objectif : terme de dérive $\mu(\boldsymbol{\theta})$ adapté à $\pi \propto \exp\left(-f - g(\mathsf{A}\cdot)\right) \mathbb{1}_{\mathcal{D}}$, g non régulière.

$\mathsf{A} \in \mathbb{R}^{c \times d}$, $c \leq d$ de **rang plein** $\Longrightarrow$ extension **inversible** $\overline{\mathsf{A}} \in \mathbb{R}^{d \times d}$

terme de dérive dual $\tilde{\mu}(\tilde{\boldsymbol{\theta}})$, $\tilde{\boldsymbol{\theta}} = \overline{\mathsf{A}}\boldsymbol{\theta}$, adapté à $\tilde{\pi} \propto \exp\left(-f(\overline{\mathsf{A}}^{-1}\cdot) - \overline{g}\right) \mathbb{1}_{\mathcal{D}}(\overline{\mathsf{A}}^{-1}\cdot)$

Méthodes duales

- Dérive de Moreau duale :

$$\tilde{\mu}^{\mathtt{M}}(\tilde{\boldsymbol{\theta}}) = \tilde{\boldsymbol{\theta}} - \gamma \overline{\mathsf{A}}^{-\top} \nabla f(\overline{\mathsf{A}}^{-1}\tilde{\boldsymbol{\theta}}) - \frac{\gamma}{\rho}(\mathsf{I} - \operatorname{prox}_{\rho\overline{g}})\tilde{\boldsymbol{\theta}}, \quad \rho = \gamma$$

Monte Carlo Langevin proximal

Objectif : terme de dérive $\mu(\boldsymbol{\theta})$ adapté à $\pi \propto \exp\left(-f - g(\mathsf{A}\cdot)\right) \mathbb{1}_{\mathcal{D}}$, g non régulière.

$\mathsf{A} \in \mathbb{R}^{c \times d}$, $c \leq d$ de **rang plein** $\Longrightarrow$ extension **inversible** $\overline{\mathsf{A}} \in \mathbb{R}^{d \times d}$

terme de dérive dual $\tilde{\mu}(\tilde{\boldsymbol{\theta}})$, $\tilde{\boldsymbol{\theta}} = \overline{\mathsf{A}}\boldsymbol{\theta}$, adapté à $\tilde{\pi} \propto \exp\left(-f(\overline{\mathsf{A}}^{-1}\cdot) - \overline{g}\right) \mathbb{1}_{\mathcal{D}}(\overline{\mathsf{A}}^{-1}\cdot)$

Méthodes duales

- Dérive de Moreau duale :

$$\tilde{\mu}^{\mathtt{M}}(\tilde{\boldsymbol{\theta}}) = \tilde{\boldsymbol{\theta}} - \gamma \overline{\mathsf{A}}^{-\top} \nabla f(\overline{\mathsf{A}}^{-1}\tilde{\boldsymbol{\theta}}) - \frac{\gamma}{\rho}(\mathsf{I} - \mathrm{prox}_{\rho \overline{g}})\tilde{\boldsymbol{\theta}}, \quad \rho = \gamma$$

- Dérive PGdual :

$$\tilde{\mu}^{\mathtt{PG}}(\tilde{\boldsymbol{\theta}}) = \mathrm{prox}_{\gamma \overline{g}}\left(\tilde{\boldsymbol{\theta}} - \gamma \overline{\mathsf{A}}^{-\top} \nabla f(\overline{\mathsf{A}}^{-1}\tilde{\boldsymbol{\theta}})\right)$$

(Artigas, 2022, *EUSIPCO* ; Fort et col, 2022, *preprint*)

Monte Carlo Langevin proximal

Objectif : terme de dérive $\mu(\theta)$ adapté à $\pi \propto \exp(-f - g(\mathsf{A}\cdot))\,\mathbb{1}_{\mathcal{D}}$, g non régulière.

$\mathsf{A} \in \mathbb{R}^{c\times d}$, $c \leq d$ de **rang plein** $\Longrightarrow$ extension **inversible** $\overline{\mathsf{A}} \in \mathbb{R}^{d\times d}$

terme de dérive dual $\tilde{\mu}(\tilde{\theta})$, $\tilde{\boldsymbol{\theta}} = \overline{\mathsf{A}}\boldsymbol{\theta}$, adapté à $\tilde{\pi} \propto \exp\left(-f(\overline{\mathsf{A}}^{-1}\cdot) - \overline{g}\right)\mathbb{1}_{\mathcal{D}}(\overline{\mathsf{A}}^{-1}\cdot)$

Méthodes duales

- Dérive de Moreau duale :

$$\tilde{\mu}^{\mathrm{M}}(\tilde{\boldsymbol{\theta}}) = \tilde{\boldsymbol{\theta}} - \gamma\overline{\mathsf{A}}^{-\top}\nabla f(\overline{\mathsf{A}}^{-1}\tilde{\boldsymbol{\theta}}) - \frac{\gamma}{\rho}(\mathsf{I} - \mathrm{prox}_{\rho\overline{g}})\tilde{\boldsymbol{\theta}}, \quad \rho = \gamma$$

- Dérive PGdual :

$$\tilde{\mu}^{\mathrm{PG}}(\tilde{\boldsymbol{\theta}}) = \mathrm{prox}_{\gamma\,\overline{g}}\left(\tilde{\boldsymbol{\theta}} - \gamma\,\overline{\mathsf{A}}^{-\top}\nabla f(\overline{\mathsf{A}}^{-1}\tilde{\boldsymbol{\theta}})\right)$$

(Artigas, 2022, *EUSIPCO* ; Fort et col, 2022, *preprint*)

- Dérive marché aléatoire duale : $\tilde{\mu}^{\mathrm{RM}}(\tilde{\boldsymbol{\theta}}) = \tilde{\boldsymbol{\theta}}$

Exemple : MCMC pour la régression logistique pénalisée

Modèle

- $\mathsf{X} \in \mathbb{R}^{N \times d}$: matrice des covariables,
- $\boldsymbol{\theta}^* \in \mathbb{R}^d$: vecteur de régression constant par blocs,
- $\mathsf{Y} \in \{0, 1\}^N$: vecteur de réponses binaires

$$\mathsf{Y}_j \sim \text{Bernoulli}\left((1 + \exp(-(\mathsf{X}\boldsymbol{\theta}^*)_j))^{-1} \right), \quad \text{indépendantes.}$$

Exemple : MCMC pour la régression logistique pénalisée

Modèle

- $\mathsf{X} \in \mathbb{R}^{N \times d}$: matrice des covariables,
- $\boldsymbol{\theta}^* \in \mathbb{R}^d$: vecteur de régression constant par blocs,
- $\mathsf{Y} \in \{0,1\}^N$: vecteur de réponses binaires

$$\mathsf{Y}_j \sim \text{Bernoulli}\left((1+\exp(-(\mathsf{X}\boldsymbol{\theta}^*)_j))^{-1}\right), \quad \text{indépendantes.}$$

Log-distribution *a posteriori*

$\mathsf{D}_1 \in \mathbb{R}^{d-1 \times d}$: gradient discret

$$\ln \pi_{\mathrm{t}}(\boldsymbol{\theta}) = \mathsf{Y}^\top \mathsf{X}\boldsymbol{\theta} - \sum_{j=1}^{N} \ln\left(1+\exp((\mathsf{X}\boldsymbol{\theta})_j)\right) - \lambda \|\mathsf{D}_1\boldsymbol{\theta}\|_1$$

Exemple : MCMC pour la régression logistique pénalisée

Modèle
- $X \in \mathbb{R}^{N \times d}$: matrice des covariables,
- $\boldsymbol{\theta}^* \in \mathbb{R}^d$: vecteur de régression constant par blocs,
- $Y \in \{0, 1\}^N$: vecteur de réponses binaires

$$Y_j \sim \text{Bernoulli}\left((1 + \exp(-(X\boldsymbol{\theta}^*)_j))^{-1}\right), \quad \text{indépendantes.}$$

Log-distribution *a posteriori* $D_1 \in \mathbb{R}^{d-1 \times d}$: gradient discret

$$\ln \pi_t(\boldsymbol{\theta}) = Y^\top X\boldsymbol{\theta} - \sum_{j=1}^{N} \ln\left(1 + \exp((X\boldsymbol{\theta})_j)\right) - \lambda \|D_1\boldsymbol{\theta}\|_1$$

PGdec : $\|D_1\boldsymbol{\theta}\|_1 = \underbrace{\|D_{1,p}\boldsymbol{\theta}\|_1}_{\text{lignes paires}} + \underbrace{\|D_{1,i}\boldsymbol{\theta}\|_1}_{\text{lignes impaires}}$, $\quad D_{1,p}D_{1,p}^\top = D_{1,i}D_{1,i}^\top = \nu I$, $\nu = 1$;

Exemple : MCMC pour la régression logistique pénalisée

Modèle

- $\mathrm{X} \in \mathbb{R}^{N \times d}$: matrice des covariables,
- $\boldsymbol{\theta}^* \in \mathbb{R}^d$: vecteur de régression constant par blocs,
- $\mathrm{Y} \in \{0,1\}^N$: vecteur de réponses binaires

$$\mathrm{Y}_j \sim \text{Bernoulli}\left(\left(1 + \exp(-(\mathrm{X}\boldsymbol{\theta}^*)_j)\right)^{-1}\right), \quad \text{indépendantes.}$$

Log-distribution *a posteriori* $\mathrm{D}_1 \in \mathbb{R}^{d-1 \times d}$: gradient discret

$$\ln \pi_{\mathrm{t}}(\boldsymbol{\theta}) = \mathrm{Y}^\top \mathrm{X}\boldsymbol{\theta} - \sum_{j=1}^{N} \ln\left(1 + \exp((\mathrm{X}\boldsymbol{\theta})_j)\right) - \lambda \|\mathrm{D}_1\boldsymbol{\theta}\|_1$$

`PGdec` : $\|\mathrm{D}_1\boldsymbol{\theta}\|_1 = \underbrace{\|\mathrm{D}_{1,p}\boldsymbol{\theta}\|_1}_{\text{lignes paires}} + \underbrace{\|\mathrm{D}_{1,i}\boldsymbol{\theta}\|_1}_{\text{lignes impaires}}$, $\mathrm{D}_{1,p}\mathrm{D}_{1,p}^\top = \mathrm{D}_{1,i}\mathrm{D}_{1,i}^\top = \nu \mathrm{I}$, $\nu = 1$;

`*dual` : $\overline{\mathrm{D}_1} = \begin{pmatrix} -1 & 0 \dots 0 \\ 0 & \\ \vdots & \mathrm{D}_1 \\ 0 & \end{pmatrix} \Longrightarrow$ extension inversible de D_1.

Exemple : MCMC pour la régression logistique pénalisée

Modèle

- $\mathsf{X} \in \mathbb{R}^{N \times d}$: matrice des covariables,
- $\boldsymbol{\theta}^* \in \mathbb{R}^{d}$: vecteur de régression constant par blocs,
- $\mathsf{Y} \in \{0,1\}^N$: vecteur de réponses binaires

$$\mathsf{Y}_j \sim \text{Bernoulli}\left(\left(1 + \exp(-(\mathsf{X}\boldsymbol{\theta}^*)_j)\right)^{-1} \right), \quad \text{indépendantes.}$$

Log-distribution *a posteriori*

$\mathsf{D}_1 \in \mathbb{R}^{d-1 \times d}$: gradient discret

$$\ln \pi_{\mathrm{t}}(\boldsymbol{\theta}) = \mathsf{Y}^\top \mathsf{X}\boldsymbol{\theta} - \sum_{j=1}^{N} \ln\left(1 + \exp((\mathsf{X}\boldsymbol{\theta})_j)\right) - \lambda \|\mathsf{D}_1\boldsymbol{\theta}\|_1$$

`PGdec` : $\|\mathsf{D}_1\boldsymbol{\theta}\|_1 = \underbrace{\|\mathsf{D}_{1,p}\boldsymbol{\theta}\|_1}_{\text{lignes paires}} + \underbrace{\|\mathsf{D}_{1,i}\boldsymbol{\theta}\|_1}_{\text{lignes impaires}}$, $\quad \mathsf{D}_{1,p}\mathsf{D}_{1,p}^\top = \mathsf{D}_{1,i}\mathsf{D}_{1,i}^\top = \nu \mathsf{I}$, $\nu = 1$;

`*dual` : $\overline{\mathsf{D}}_1 = \begin{pmatrix} -1 & 0 \dots 0 \\ 0 & \\ \vdots & \mathsf{D}_1 \\ 0 & \end{pmatrix} \Longrightarrow$ extension inversible de D_1.

Données $N = 2.10^3$, $d = 20$

X : v.a. de Rademacher indépendantes, lignes normalisées à 1.

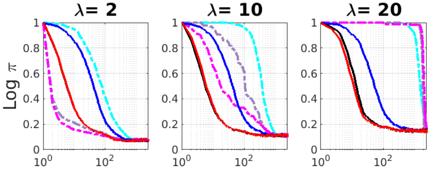
Indicateur de convergence :

$$\mathrm{Log}\,\pi = \frac{\ln \pi_t(\boldsymbol{\theta}^n) - \ln \pi_t^*}{\ln \pi_t(\boldsymbol{\theta}^1) - \ln \pi_t^*}, \qquad \ln \pi_t^* = \max_{\boldsymbol{\theta}\in\mathbb{R}} \ln \pi_t(\boldsymbol{\theta})$$

régions de grande probabilité

Comparaison des méthodes

primales *en traits pleins*	duales *en pointillés*
`RW`	`RWdual`
`M`	`Mdual`
`PGdec`	`PGdual`



– gain à utiliser l'ordre un vs. `RW` ;

– méthodes primales : les meilleures à λ petit ;

– méthodes duales : les plus rapides pour λ moyen à grand, correctes pour λ petit.

$\Longrightarrow$ `Mdual` et `PGdual` bonnes performances ; robustes au choix de λ

[Mdual] dérive de type Moreau sur le dual $\quad \gamma = \rho$

$$\mathbf{R}^{n+\frac{1}{2}} = \mathbf{R}^n - \gamma \overline{\mathsf{D}}_2^{-1} \overline{\mathsf{D}}_2^{-\top} \nabla_{\mathsf{R}} f(\boldsymbol{\theta}^n) - \overline{\mathsf{D}}_2^{-1} \left[0\,;0\,;\mathsf{D}_2 \boldsymbol{\theta}^n - \mathrm{prox}_{\gamma \lambda_{\mathrm{R}} \|\cdot\|_1}(\mathsf{D}_2 \mathbf{R}^n)\right] + \sqrt{2\gamma}\xi_{\mathsf{R}}^{n+1},$$

$$\mathbf{O}^{n+\frac{1}{2}} = -\gamma_{\mathsf{O}} \nabla_{\mathsf{O}} f(\boldsymbol{\theta}^n) + \mathrm{prox}_{\gamma_{\mathsf{O}} \lambda_{\mathrm{O}} \|\cdot\|_1}(\mathbf{O}^n) + \sqrt{2\gamma_{\mathsf{O}}}\xi_{\mathsf{O}}^{n+1}\ ;$$

[PGdual] dérive de type proximal-gradient sur le dual

$$\mathbf{R}^{n+\frac{1}{2}} = \overline{\mathsf{D}}_2^{-1} \mathrm{prox}_{\gamma \lambda_{\mathrm{R}} \|(\cdot)_{3:T}\|_1} \left(\overline{\mathsf{D}}_2 \mathbf{R}^n - \gamma \overline{\mathsf{D}}_2^{-\top} \nabla_{\mathsf{R}} f(\boldsymbol{\theta}^n) \right) + \sqrt{2\gamma}\xi_{\mathsf{R}}^{n+1},$$

$$\mathbf{O}^{n+\frac{1}{2}} = \mathrm{prox}_{\gamma_{\mathsf{O}} \lambda_{\mathrm{O}} \|\cdot\|_1} \left(\mathbf{O}^n - \gamma_{\mathsf{O}} \nabla_{\mathsf{O}} f(\boldsymbol{\theta}^n) \right) + \sqrt{2\gamma_{\mathsf{O}}}\xi_{\mathsf{O}}^{n+1}.$$

$\nabla_{\mathsf{R}/\mathsf{O}}$: gradient partiel par rapport à R/O ;

[`Mdual`] dérive de type Moreau sur le dual $\quad \gamma = \rho$

$$\mathbf{R}^{n+\frac{1}{2}} = \mathbf{R}^n - \gamma \overline{\mathsf{D}}_2^{-1} \overline{\mathsf{D}}_2^{-\top} \nabla_{\mathsf{R}} f(\boldsymbol{\theta}^n) - \overline{\mathsf{D}}_2^{-1} \left[0\,;0\,;\mathsf{D}_2\boldsymbol{\theta}^n - \mathrm{prox}_{\gamma\lambda_{\mathrm{R}}\|\cdot\|_1}(\mathsf{D}_2\mathbf{R}^n)\right] + \sqrt{2\gamma}\xi_{\mathsf{R}}^{n+1},$$

$$\mathbf{O}^{n+\frac{1}{2}} = -\gamma_{\mathsf{O}} \nabla_{\mathsf{O}} f(\boldsymbol{\theta}^n) + \mathrm{prox}_{\gamma_{\mathsf{O}}\lambda_{\mathrm{O}}\|\cdot\|_1}(\mathbf{O}^n) + \sqrt{2\gamma_{\mathsf{O}}}\xi_{\mathsf{O}}^{n+1}\ ;$$

[`PGdual`] dérive de type proximal-gradient sur le dual

$$\mathbf{R}^{n+\frac{1}{2}} = \overline{\mathsf{D}}_2^{-1} \mathrm{prox}_{\gamma\lambda_{\mathrm{R}}\|(\cdot)_{3:T}\|_1} \left(\overline{\mathsf{D}}_2\mathbf{R}^n - \gamma \overline{\mathsf{D}}_2^{-\top} \nabla_{\mathsf{R}} f(\boldsymbol{\theta}^n)\right) + \sqrt{2\gamma}\xi_{\mathsf{R}}^{n+1},$$

$$\mathbf{O}^{n+\frac{1}{2}} = \mathrm{prox}_{\gamma_{\mathsf{O}}\lambda_{\mathrm{O}}\|\cdot\|_1}\left(\mathbf{O}^n - \gamma_{\mathsf{O}} \nabla_{\mathsf{O}} f(\boldsymbol{\theta}^n)\right) + \sqrt{2\gamma_{\mathsf{O}}}\xi_{\mathsf{O}}^{n+1}.$$

$\nabla_{\mathsf{R}/\mathsf{O}}$: gradient partiel par rapport à R/O ;

Perturbations gaussiennes :

- $\xi_{\mathsf{R}}^{n+1} \sim \mathcal{N}(0, \overline{\mathsf{D}}_2^{-1} \overline{\mathsf{D}}_2^{-\top})$,
- $\xi_{\mathsf{O}}^{n+1} \sim \mathcal{N}(0, \mathsf{I})$;

Paramètres :

- $(\lambda_{\mathrm{R}}, \lambda_{\mathrm{O}}) = (3.5\,\sigma_{\mathbf{Z}}\sqrt{6}/4, 0.05)$,
- $\gamma_{\mathsf{O}} = \gamma(\lambda_{\mathrm{R}}/\lambda_{\mathrm{O}})^2$,
- γ ajusté pour accepter 25% des $\boldsymbol{\theta}^{n+\frac{1}{2}}$.

[Mdual] dérive de type Moreau sur le dual $\quad \gamma = \rho$

$$\mathbf{R}^{n+\frac{1}{2}} = \mathbf{R}^n - \gamma \overline{\mathrm{D}}_2^{-1}\overline{\mathrm{D}}_2^{-\top}\nabla_{\mathrm{R}} f(\boldsymbol{\theta}^n) - \overline{\mathrm{D}}_2^{-1}\left[0\,;0\,;\mathrm{D}_2\boldsymbol{\theta}^n - \mathrm{prox}_{\gamma\lambda_{\mathrm{R}}\|\cdot\|_1}(\mathrm{D}_2\mathbf{R}^n)\right] + \sqrt{2\gamma}\,\xi_{\mathrm{R}}^{n+1},$$

$$\mathbf{O}^{n+\frac{1}{2}} = -\gamma_{\mathrm{O}}\nabla_{\mathrm{O}} f(\boldsymbol{\theta}^n) + \mathrm{prox}_{\gamma_{\mathrm{O}}\lambda_{\mathrm{O}}\|\cdot\|_1}(\mathbf{O}^n) + \sqrt{2\gamma_{\mathrm{O}}}\,\xi_{\mathrm{O}}^{n+1}\ ;$$

[PGdual] dérive de type proximal-gradient sur le dual

$$\mathbf{R}^{n+\frac{1}{2}} = \overline{\mathrm{D}}_2^{-1}\,\mathrm{prox}_{\gamma\lambda_{\mathrm{R}}\|(\cdot)_{3:T}\|_1}\left(\overline{\mathrm{D}}_2\mathbf{R}^n - \gamma\overline{\mathrm{D}}_2^{-\top}\nabla_{\mathrm{R}} f(\boldsymbol{\theta}^n)\right) + \sqrt{2\gamma}\,\xi_{\mathrm{R}}^{n+1},$$

$$\mathbf{O}^{n+\frac{1}{2}} = \mathrm{prox}_{\gamma_{\mathrm{O}}\lambda_{\mathrm{O}}\|\cdot\|_1}\left(\mathbf{O}^n - \gamma_{\mathrm{O}}\nabla_{\mathrm{O}} f(\boldsymbol{\theta}^n)\right) + \sqrt{2\gamma_{\mathrm{O}}}\,\xi_{\mathrm{O}}^{n+1}.$$

$\nabla_{\mathrm{R}/\mathrm{O}}$: gradient partiel par rapport à R/O ;

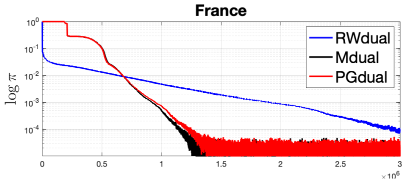
Perturbations gaussiennes :

- $\xi_{\mathrm{R}}^{n+1} \sim \mathcal{N}(0, \overline{\mathrm{D}}_2^{-1}\overline{\mathrm{D}}_2^{-\top})$,
- $\xi_{\mathrm{O}}^{n+1} \sim \mathcal{N}(0, \mathrm{I})$;

Paramètres :

- $(\lambda_{\mathrm{R}}, \lambda_{\mathrm{O}}) = (3.5\,\sigma_{\mathbf{Z}}\sqrt{6}/4, 0.05)$,
- $\gamma_{\mathrm{O}} = \gamma(\lambda_{\mathrm{R}}/\lambda_{\mathrm{O}})^2$,
- γ ajusté pour accepter 25% des $\boldsymbol{\theta}^{n+\frac{1}{2}}$.

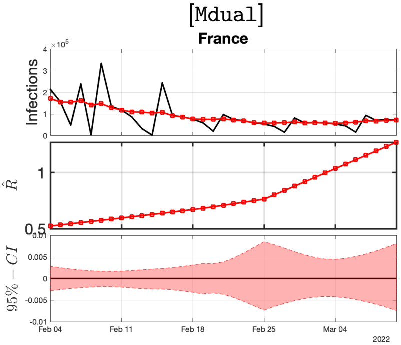
Vitesse de convergence des chaînes de Markov

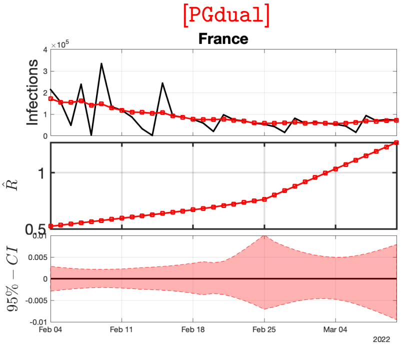


Estimation par intervalles de crédibilité

– Nombre de nouvelles infections débruités $\mathbf{Z}^{(D)} = \mathbf{Z} - \widehat{\mathbf{O}}$

– Coefficient de reproduction $\widehat{\mathbf{R}}$



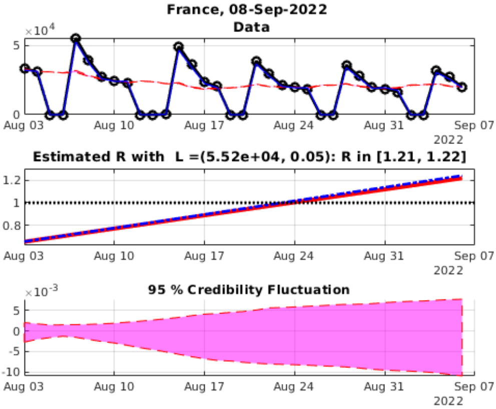


$\Longrightarrow$ crucial pour appuyer des mesures sanitaires sur l'estimée R_T

Nombre de cas débruité et intervalles de crédibilité de R_t postés quotidiennement

`https://perso.ens-lyon.fr/patrice.abry/`

`https://perso.math.univ-toulouse.fr/gfort/project/opsimore-2/`



Pistes de recherche et perspectives :

- Réglage automatique piloté par les données des hyperparamètres $\gamma_{R/O}$, $\lambda_{R/O}$;
- Gestion de données reportées de moins en moins régulièrement.