Une Introduction aux méthodes particulaires en ingénierie stochastique

P. Del Moral

INRIA Centre Bordeaux-Sud Ouest

Journée de rencontre EDP/Proba, IHP Mars 2008

Outline

Le filtrage de processus stochastiques

- X_t**:=Signal=Processus stochastique**

ingénierie/physique/biologique/économique :

- Cibles non coopérative (militaire : missile, char, avion,...).
- Physiques (fluides : tornades, cyclones, modèles océano, pression/température/coeff de diffusivité,...).
- Financiers (actifs boursiers, portefeuilles, volatilités,...).
- Signaux (parole, codages/transmission informations,...)

Dynamiques et aléas :

- Equations d'évolution physiques (Ex. $\sum_i u_i \overrightarrow{F}_i = \overrightarrow{A}$)
- Perturbations et aléas :
  - Erreurs de modèles $\oplus$ Perturbations extérieures.
  - Contrôles et paramètres inconnus.

    $\leadsto$ lois a priori $\oplus$ (inconnues= réalisation de v.a.)

Modèles de Filtrage de processus stochastiques

- **Y_t=Observations partielles et bruitées de X_t :**

ingénierie/physique/biologique/économique :

- Ingénierie : Radar, Sonar, GPS, ...
- Physiques (capteurs de pression/température/...).
- Financiers (actifs boursiers, portefeuilles, volatilités,...).
- Statistique (données réelles : médecine, pharmaceutiques, politique, éco-finances,...).

Dynamiques et aléas :

- Observations partielles : mélanges, coordonnées partielles.
- Perturbations et aléas :
  - erreurs de mesures (bruits thermiques).
  - perturbations extérieures.
  - erreurs de modèle.

Objectifs

Calculer/Simuler/Estimer récursivement le flot de mesures

$$t \in \mathbb{R}_+ \quad \text{ou} \quad t = n \in \mathbb{N} \longrightarrow \eta_t = \mathrm{Loi}(X_t \mid Y_0, \ldots, Y_t)$$

Remarques

- Filtrage trajectoriel : $X_t = (X'_0, \ldots, X'_t) \in E_t$

$$\Updownarrow$$

$$\eta_t = \mathrm{Loi}((X'_0, \ldots, X'_t) \mid (Y_0, \ldots, Y_t)) = \mathrm{Loi}(X_t \mid Y_0, \ldots, Y_t)$$

Vocabulaire "équivalent" :

- Assimilation de données (météo).
- Chaines de Markov cachées (stat. fréquentiste).
- Lois a posteriori (a priori=$\mathrm{Loi}(X)$) (stat. baysiennes).

L'heuristique du filtrage particulaire

Dynamique de population de N "individus"/particules" t.q.

$$(\widehat{\xi}_t^1, \ldots, \widehat{\xi}_t^N) \in E_t^N \rightsquigarrow \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^N \delta_{\widehat{\xi}_t^i} = \mathrm{Loi}(X_t \mid (Y_0, \ldots, Y_t))$$

Méthode de simulation effective

- Prédiction $\rightsquigarrow$ simulation de N transitions locales du signal.
- Correction$\rightsquigarrow$ processus de naissance et mort (taille fixe N).
  - On tue les individus sur des sites peu vraisemblables.
  - On multiplie les individus sur des sites plus vraisemblables.

$\rightsquigarrow$ Mod. trajectoriels : $X_t = (X_0', \ldots, X_t')$ $\rightsquigarrow$ Arbre Généalogique

$$\lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^N \delta_{\text{ligne ancestrale}_t(i)} = \mathrm{Loi}((X_0', \ldots, X_t') \mid (Y_0, \ldots, Y_t))$$

Analyse d'événements rares

2 Ingrédients

- 1 Processus Physique/biologique/economique : files d'attentes, réseau télécom, portefeuille, volatilité d'actifs, température, mouvements de fluides,...
- 1 fonction potentiel (type énergie ou indicatrice/restriction): dépassements de niveaux critiques, saturation, niveaux de propagations d'épidémies, dispersion de radioactivité, ruines,...

Objectifs

- Calcul des probabilités d'événements rares.
- Calculer les lois des trajectoires du processus évoluant en régime $\rightsquigarrow$ prédiction $\oplus$ contrôle.

L'heuristique des estimations particulaires

Flot de mesures à complexité croissante

- l'événement rare = cascade d'événements intermédiaires (moins) rares (énergies ou niveaux $\uparrow$, passerelles physiques).
- Flot de lois conditionnelles

$$n \to \eta_n = \text{Loi}(\text{processus} | \text{une série de } n \text{ évts intermédiaires} \downarrow)$$

- Les probabilités d'événements rares= Cts de normalisation.

Méthode particulaire heuristique

(Simulation arbre de défauts de type généalogique $\oplus$ % réussites)

- Explorations/Propositions locales des espaces d'états.
- Branchements-Selection des individus $\in$ régimes critiques $\uparrow$.

4 Exemples de flots de mesures "cibles"

1. $\eta_n = \mathrm{Loi}((X_0, \ldots, X_n) \mid \forall 0 \leq p \leq n \quad X_p \in A_p)$
2. $\eta_n(dx) \propto e^{-\beta_n V(x)} \lambda(dx)$ avec $\beta_n \uparrow$
3. $\eta_n(dx) \propto 1_{A_n}(x) \lambda(dx)$ avec $A_n \downarrow$
4. $\eta_n = \mathrm{Loi}_\pi((X_0, \ldots, X_n) \mid X_n = x_n)$.

4 Heuristiques particulaires :

1. Transitions locales $\oplus$ Selection des individus $\in A_p$
2. Transitions MCMC mes. inv. $= \eta_{n-1}$
   $\oplus$ Selection des individus $\propto G_n(x) = e^{-(\beta_n - \beta_{n-1})V(x)}$
3. Transitions MCMC mes. inv. $= \eta_{n-1}$
   $\oplus$ Selection des individus $\propto G_n(x) = 1_{A_n}(x)$
4. Transitions locales $\oplus$ Selection $G(x_1, x_2) = \frac{\pi(dx_2)K(x_2, dx_1)}{\pi(dx_1)M(x_1, dx_2)}$

Flot (non linéaires) de mesures de probabilités [Temps discret]

- Transitions Markov $K_{n,\eta}$

$$\eta_n(dy) = \left(\eta_{n-1} K_{n,\eta_{n-1}}\right)(dy) := \int \eta_{n-1}(dx)\ K_{n,\eta_{n-1}}(x,dy)$$

Exemple : Prédicteurs opt. $\eta_n := \mathrm{Loi}(X_n \mid (Y_0,\ldots,Y_{n-1}))$

$$\eta_n \xrightarrow{\text{Correction}} \widehat{\eta}_n = \Psi_n(\eta_n) \xrightarrow{\text{Prédiction}} \eta_{n+1} = \widehat{\eta}_n M_{n+1}$$

Transf. de Boltzmann-Gibbs : $G_n(x)$=vraisemblance du site x

$$\Psi_n(\eta_n)(dx) := \frac{1}{\eta_n(G_n)}\ G_n(x)\ \eta_n(dx) = \eta_n S_{n,\eta_n}(dx)$$

avec le transport $\forall \epsilon_n G_n \leq 1$

$$S_{n,\eta_n}(x,dy) := \epsilon_n G_n(x)\ \delta_x(dy) + (1-\epsilon_n G_n(x))\ \Psi_n(\eta_n)(dy)$$

Chaines de Markov $\overline{X}_n$ non linéaires t.q. $\eta_n = \mathrm{Loi}(\overline{X}_n)$

- **Transitions locales :**

$$\mathbb{P}(\overline{X}_n \in dx_n \mid \overline{X}_{n-1}) = K_{n,\eta_{n-1}}(\overline{X}_{n-1}, dx_n)$$

- **Mesure de McKean (processus canonique) :**

$$\mathbb{P}_n(d(x_0,\ldots,x_n)) = \eta_0(dx_0)\ K_{1,\eta_0}(x_0,dx_1)\ldots K_{n,\eta_{n-1}}(x_{n-1},dx_n)$$

Interprétation particulaire de type champ moyen

- Chaine Markov $\xi_n = (\xi_n^1,\ldots,\xi_n^N) \in E_n^N$ t.q.

$$\eta_n^N := \frac{1}{N}\sum_{1\leq i\leq N} \delta_{\xi_n^i} \ \simeq_{N\uparrow\infty}\ \eta_n$$

- Transitions locales "approchées" $(\forall 1 \leq i \leq N)$

$$\xi_{n-1}^i \rightsquigarrow \xi_n^i \sim K_{n,\eta_{n-1}^N}(\xi_{n-1}^i, dx_n)$$

Modèle champ moyen temps discret

Figure schématique : $\xi_n \in E_n^N \rightsquigarrow \xi_{n+1} \in E_{n+1}^N$

$$\begin{matrix} \xi_n^1 & \xrightarrow{K_{n+1,\eta_n^N}} & \xi_{n+1}^1 \\ \vdots & & \vdots \\ \xi_n^i & \longrightarrow & \xi_{n+1}^i \\ \vdots & & \vdots \\ \xi_n^N & \longrightarrow & \xi_{n+1}^N \end{matrix}$$

Idée intuitive :

$$\begin{aligned} \eta_n^N \simeq_{N\uparrow\infty} \eta_n \quad & \Longrightarrow \quad K_{n+1,\eta_n^N} \simeq_{N\uparrow\infty} K_{n+1,\eta_n} \\ & \Longrightarrow \quad \xi_n^i \text{ copies i.i.d. de } \overline{X}_n \end{aligned}$$

QQ Avantages

- Modèle de champ moyen=linéarisation/perturbation stoch. :

$$\eta_n^N = \eta_{n-1}^N K_{n,\eta_{n-1}^N} + \frac{1}{\sqrt{N}}\ W_n^N$$

  avec $W_n^N \simeq W_n$ Champs gaussiens centrés et $\perp$.
- $\eta_n = \eta_{n-1} K_{n,\eta_{n-1}}$ stable $\Rightarrow$ non propagation des erreurs locales

$$\Longrightarrow \text{control uniforme des erreurs globales / temps}$$

- "Pas besoin" d'étudier la cv à l'équilibre de modèles MCMC.
- Grille/Maillage stochastique adaptatif.
- Non linéarité du syst. $\rightsquigarrow$ interactions bénéfiques.
- Algo. naturel et facile a simuler, etc.

L'exemple $\rightsquigarrow$ Modèles de Feynman-Kac

(Correction,Prédiction)=(G_n, M_n)=(**selection, explorations**)

$$\eta_n = \Phi_n(\eta_{n-1}) := \Psi_{G_{n-1}}(\eta_{n-1})M_n \simeq_{N\uparrow\infty} \eta_n^N$$

"Solution" : X_n Markov $\sim$ transitions M_n

$$\eta_n(f_n) = \frac{\gamma_n(f_n)}{\gamma_n(1)} \quad \text{with} \quad \gamma_n(f_n) = \mathbb{E}\left(f_n(X_n)\prod_{0\leq p<n} G_p(X_p)\right)$$

Formule multiplicative $\rightsquigarrow$ est. non biaisé des Cts. de normalisation

$$\mathbb{E}\left(\prod_{0\leq p<n} G_p(X_p)\right) = \prod_{0\leq p<n} \eta_p(G_p) \simeq_{N\uparrow\infty} \prod_{0\leq p<n} \eta_p^N(G_p)$$

$\supset$ Exemples $\oplus$ TOUTES les heuristiques précédentes :
G_n=vraisemblances, $G_n = 1_{A_n}$, $G_n = e^{-(\beta_n-\beta_{n-1})V}$, $G_n = e^{-\beta V}$,etc.

Algorithme stochastique de type génétique

$$\left.\begin{array}{c}\xi_n^1\\ \vdots\\ \xi_n^i\\ \vdots\\ \xi_n^N\end{array}\right] \xrightarrow{S_{n,\eta_n^N}} \left[\begin{array}{ccc}\widehat{\xi}_n^1 & \xrightarrow{M_{n+1}} & \xi_{n+1}^1\\ \vdots & & \vdots\\ \widehat{\xi}_n^i & \longrightarrow & \xi_{n+1}^i\\ \vdots & & \vdots\\ \widehat{\xi}_n^N & \longrightarrow & \xi_{n+1}^N\end{array}\right]$$

Acceptation/Rejet-Selection : [horloges géométriques]

$$S_{n,\eta_n^N}(\xi_n^i, dx)$$

$$:= \epsilon_n G_n(\xi_n^i)\ \delta_{\xi_n^i}(dx) + \left(1 - \epsilon_n G_n(\xi_n^i)\right)\ \sum_{j=1}^N \frac{G_n(\xi_n^j)}{\sum_{k=1}^N G_n(\xi_n^k)} \delta_{\xi_n^j}(dx)$$

Ex. : $G_n = 1_A \rightsquigarrow G_n(\xi_n^i) = 1_A(\xi_n^i)$

Algorithmes stochastiques "équivalents" :

- Algorithmes génétiques, processus de branchements.
- Méthodes de Monte Carlo séquentielles.
- Algorithmes de "population Monte Carlo", Diffusion Monte Carlo (DMC), Quantum Monte Carlo (QMC), ...
- Botanique des processus d'adaptation $\sim \neq$ domaines d'applications :
  bootsrapping, selection, pruning-enrichment, reconfiguration, cloning, go with the winner, spawning, condensation, grouping, rejuvenations, harmony searches, biomimetics, ...

Theorie "asymptotique" TCL,PGD, PDM,...(n,N). QQ exemples :

- Mesure McKean particulaire

$$\frac{1}{N}\sum_{i=1}^{N}\delta_{(\xi_0^i,\ldots,\xi_n^i)} \simeq_N \mathrm{Loi}(\overline{X}_0,\ldots,\overline{X}_n) \quad \text{et} \quad \eta_n^N = \frac{1}{N}\sum_{i=1}^{N}\delta_{\xi_n^i} \simeq_N \eta_n$$

- Proc. empiriques : $\sup_{n\geq 0}\sup_{N\geq 1}\sqrt{N}\,\mathbb{E}(\|\eta_n^N-\eta_n\|_{\mathcal{F}_n}^p)<\infty$
- Inégalités de concentration uniformes :

$$\sup_{n\geq 0}\mathbb{P}(|\eta_n^N(f_n)-\eta_n(f_n)|>\epsilon)\leq c\ \exp-(N\epsilon^2)/(2\ \sigma^2)$$

- Propagations du chaos : $\mathbb{P}_{n,q}^N := \mathrm{Loi}(\xi_n^1,\ldots,\xi_n^q)$

$$\mathbb{P}_{n,q}^N \simeq \eta_n^{\otimes q} + \frac{1}{N}\ \partial^1\mathbb{P}_{n,q} + \ldots + \frac{1}{N^k}\ \partial^k\mathbb{P}_{n,q} + \frac{1}{N^{k+1}}\ \partial^{k+1}\mathbb{P}_{n,q}^N$$

avec $\sup_{N\geq 1}\|\partial^{k+1}\mathbb{P}_{n,q}^N\|_{\mathrm{tv}}<\infty$ et $\sup_{n\geq 0}\|\partial^1\mathbb{P}_{n,q}\|_{\mathrm{tv}}\leq c\ q^2$.

Flot de mesures de probabilités (sg. non linéaires)

- **(sens faible)** : Générateurs infinitesimaux $L_{t,\eta}$

$$\frac{d}{dt}\eta_t(f) = \eta_t L_{t,\eta_t}(f) := \int_E \eta_t(dx)\ L_{t,\eta_t}(f)(x)$$

- **Exemple FKS :** $X_t \simeq \left(L \overset{ex.}{=} \frac{1}{2}\Delta\right) - \text{processus} \oplus$ potentiel V.

$$\eta_t(f) := \frac{\gamma_t(f)}{\gamma_t(1)} \quad \text{avec} \quad \gamma_t(f) = \mathbb{E}\left(f(X_t)\exp\left\{-\int_0^t V(X_s)ds\right\}\right)$$

$$\begin{aligned}
\frac{d}{dt}\gamma_t(f) &= \gamma_t(L^V(f)) \quad \text{Schrodinger op.} \quad L^V := L - V \\
\frac{d}{dt}\eta_t(f) &= \eta_t L_{\eta_t}(f) \\
&:= \int \eta_t(dx)\left\{L(f)(x) + V(x)\int (f(y) - f(x))\eta_t(dy)\right\}
\end{aligned}$$

Interprétation champ moyen

- **Processus de Markov** $\xi_t = (\xi_t^i)_{1\leq i\leq N}$ générateur infinitésimal

$$\mathcal{L}_t(F)(x^1,\ldots,x^N) := \sum_{i=1}^N L^{(i)}_{t,\frac{1}{N}\sum_{i=1}^N \delta_{x^i}} F(x^1,\ldots,x^i,\ldots,x^N)$$

- Eq. évolution des mesures d'occupation $\eta_t^N := \frac{1}{N}\sum_{i=1}^N \delta_{\xi_t^i}$

$$d\eta_t^N(f) = \eta_t^N L_{t,\eta_t^N}(f)dt + \frac{1}{\sqrt{N}}\, dM_t^N(f)$$

avec

$$\langle M^N(f)\rangle_t = \int_0^t \eta_s^N \Gamma_{L_{s,\eta_s^N}}(f,f)\, ds$$

Exemple : Modèle FKS $\rightsquigarrow$ Systèmes de particules de type Moran

- $(\xi_t^i)_{1\leq i\leq N}$ = L-explorations $\oplus$ sauts interactifs (V-intensité)

$$\mathcal{L}_t(F)(x^1,\ldots,x^N)$$

$$= \sum_{i=1}^N L^{(i)} F(x^1,\ldots,x^i,\ldots,x^N) + \sum_{i=1}^N V(x^i)$$

$$\times \int \left(F(x^1,\ldots,y^i,\ldots,x^N) - F(x^1,\ldots,x^i,\ldots,x^N)\right) m(x)(dy^i)$$

avec $m(x) = N^{-1}\sum_{i=1}^N \delta_{x^i}$.

- Théorie asymptotique $\sim$ modèles à temps discret.

Horloges Géométriques $\rightsquigarrow$ Horloges Exponentielles

Théorie asymptotique

Modèle FKS ⊕ Systèmes de particules de type Moran

- **Estimations particulaires**

$$\mathbb{E}\left(f(X_t)e^{\int_0^t V(X_s)ds}\right) = \eta_t(f)\ e^{-\int_0^t \eta_s(V)ds}$$
$$\simeq_N \eta_t^N(f)\ e^{-\int_0^t \eta_s^N(V)ds} \quad \text{(non biaisé)}$$

- Etats fondamentaux op. Schrodinger : ($\supset$ Méthodes DMC, QMC) (v.p. λ ⊕ état fondamental h (L μ-réversible))

$$\lim_{N,t\to\infty} \eta_t^N(dx) \propto h(x)\ \mu(dx) \quad \text{et} \quad e^{-\int_0^t \eta_s^N(V)ds} \simeq e^{-\lambda t}$$

- **Théorie asymptotique "∼" modèles à temps discret.**

QQ références

Modèles à temps discret

- Feynman-Kac formulae. Genealogical and interacting particle systems, Springer (2004) ⊕ Refs.
- joint work with Doucet A., Jasra A. Sequential Monte Carlo Samplers. JRSS B (2006).
- joint work with A. Doucet. Particle Motions in Absorbing Medium with Hard and Soft Obstacles. Stochastic Analysis and Applications, vol. 22 (2004).

QQ références

Modèles à temps continu

- joint work with L. Miclo. A Moran particle system approximation of Feynman-Kac formulae. *Stochastic Processes and their Applications,* Vol. 86, 193-216 (2000).
- joint work with L. Miclo. `Particle Approximations of Lyapunov Exponents` $\sim$ `FKS Sg.`. ESAIM PS, vol. 7, pp. 169-207 (2003).
- joint work with L. Miclo. On the strong propagation of chaos for interacting particle approximations of Feynman-Kac formulae. *Stochastic Analysis and Applications (2006)*.
- Travaux de Mathias Rousset (INRIA Lille, eq. SIMPAF)

  $\oplus$ B. Jourdain, T. Lelièvre, G. Stoltz (INRIA Rocq., eq. MICMAC).

Pub! : Nouvelle génération d'algorithmes MCMC en interaction

- **i-MCMC algorithms [temps discret] :**
  - joint work with A. Doucet. `Interacting Markov Chain Monte Carlo Methods For Solving Nonlinear Measure-Valued Eq.`, HAL-INRIA RR-6435, (Feb. 2008).
  - joint work with B. Bercu and A. Doucet. `Fluctuations of Interacting Markov Chain Monte Carlo Models.` HAL-INRIA RR-6438, (Feb. 2008).
  - joint work with C. Andrieu, A. Jasra, A. Doucet. *Non-Linear Markov chain Monte Carlo via self-interacting approximations.* Tech. report, Dept of Math., Bristol Univ. (2007).
  - joint work with A. Brockwell and A. Doucet. *Sequentially interacting Markov chain Monte Carlo.* Tech. report, Dept. of Statistics, Univ. of British Columbia (2007).