Journées de Statistique
Marne La Vallée, January 2005

Contraction properties of Feynman-Kac semigroups

Pierre DEL MORAL, Laurent MICLO
Lab. J. Dieudonné, Nice Univ., LATP Univ. Provence, Marseille

Notations $(E, \mathcal{E})$ measurable space

$$\begin{array}{rcl} \mathcal{P}(E) & = & \{\mu \text{ probability on E}\} \\ \mathcal{B}_b(E) & = & \{f : E \to \mathbb{R} \text{ bounded } + \mathcal{E}\text{-measurable}\} \end{array}$$

- $(\mu, f) \in \mathcal{P}(E) \times \mathcal{B}_b(E) \quad \longrightarrow \quad \mu(f) = \int f(x)\mu(dx)$

- $M(x, dx')$ Markov kernel on E

$$\mu M(dx') \;=\; \int \mu(dx) M(x, dx') \quad \text{and} \quad M(f)(x) = \int M(x, dx') f(x')$$

and, with X_n Markov chain $\sim\ M(x, dy)$

$$M^n(x, dz) = \int M^{n-1}(x, dy) M(y, dz) = \mathbb{P}_x(X_n \in dz)$$

Feynman-Kac models

- X_n Markov chain $M_n(x, dy)$ on $(E, \mathcal{E})$
- Potential function $G_n : E \to [0, \infty)$

$\rightsquigarrow$ *Feynman-Kac measures* ($\forall$ test funct. $f : E \to \mathbb{R}$)

$$\eta_n(f) = \gamma_n(f)/\gamma_n(1) \quad \text{with} \quad \gamma_n(f) = \mathbb{E}_{\eta_0}(f(X_n) \prod_{0 \le p < n} G_p(X_p))$$

Ex.:

- Updated models $\widehat{\gamma}_n(f) = \gamma_n(fG_n) \longrightarrow \widehat{\eta}_n(f) = \widehat{\gamma}_n(f)/\widehat{\gamma}_n(1) = \eta_n(fG_n)/\eta_n(G_n)$.
- $G_n = e^{-\beta V_n} \Rightarrow \eta_n = n$-th marginal path. measure $\mathbb{Q}_n = \frac{1}{\mathcal{Z}_n} \exp\{-\beta \sum_{0 \le p < n} V_p(X_p)\} \, d\mathbb{P}_n$
- $G_n = 1_A \Rightarrow \eta_n =Law(X_n \mid \forall 0 \le p < n \quad X_p \in A)$

Note: $\mathcal{Z}_n = \gamma_n(1)$ and $\gamma_n(f) = \eta_n(f) \prod_{0 \le p < n} \eta_p(G_p) \Rightarrow \frac{1}{n} \log \mathcal{Z}_n = \frac{1}{n} \sum_{0 \le p < n} \log \eta_p(G_p)$

Feynman-Kac semigroups

- Unnormalized models

$$\gamma_n = \gamma_p Q_{p,n} \longrightarrow \textbf{linear semigroup} \quad Q_{p,n}(f)(x) = \mathbb{E}_{p,x}(f(X_n) \prod_{p \leq q < n} G_q(X_q))$$

- Normalized models

$$\eta_n = \Phi_{p,n}(\eta_p) \longrightarrow \textbf{nonlinear semigroup } \Phi_{p,n} = \Phi_{n-1,n} \circ \cdots \circ \Phi_{p,p+1} \text{ with}$$

$$\Phi_{p,p+1}(\eta) = \Psi_p(\eta) M_{p+1} \quad \text{and} \quad \Psi_p(\eta)(dx) = \frac{1}{\eta(G_p)} \, G_p(x) \, \eta(dx)$$

Pb.: Asymptotic stability, contraction properties, functional entropy inequalities, decays estimates...

Applications/Motivations

- **<u>Particle physics:</u> Markov evolution** $X_n \in$ **Absorbing medium** $G(x) = e^{-V(x)} \in [0,1]$

$$X_n^c \in E^c = E \cup \{c\} \xrightarrow{absorption} \widehat{X}_n^c \xrightarrow{exploration} X_{n+1}^c$$

Absorption/killing: $\longrightarrow \widehat{X}_n^c = X_n^c$, with proba $G(X_n^c)$; otherwise the particle is killed and $\widehat{X}_n^c = c$.

$$\Downarrow$$

$$\begin{array}{rcl} A & = & \{x \;:\; G(x) = 0\} \longrightarrow \text{Hard obstacles} \\ T & = & \inf\{n \geq 0 \;;\; \widehat{X}_n^c = c\} \longrightarrow \text{Absorption time} \quad X_{T+n}^c = \widehat{X}_{T+n}^c = c \end{array}$$

$$\Downarrow$$

Feynman-Kac models (G, X_n) **:** $\quad \gamma_n = \text{Law}(X_n^c \;;\; T \geq n) \quad$ and $\quad \eta_n = \text{Law}(X_n^c \mid T \geq n)$

- **Genetics/Stoch. algo** ($\rightarrow$ particle methods = local perturbation/stoch. linearization)

$$\xi_n \in E^N \xrightarrow{\text{selection}} \widehat{\xi}_n \in E^N \xrightarrow{\text{mutation}} \xi_{n+1} \in E^N \quad \text{"iid"} \sim \Phi_{n-1,n}\left(\tfrac{1}{N}\textstyle\sum_{1\leq i\leq N} \delta_{\xi^i_{n-1}}\right)$$

  - *Selection transition*

$$\widehat{\xi}_n = (\widehat{\xi}^i_n)_{1\leq i\leq N} \text{"iid"} \quad \Psi_n\Big(\frac{1}{N}\sum_{1\leq j\leq N} \delta_{\xi^j_{n-1}}\Big) = \sum_{1\leq j\leq N} \frac{G_n(\xi^j_n)}{\sum_{1\leq k\leq N} G_n(\xi^k_n)} \delta_{\xi^j_n}$$

  - *Mutation transition* $\widehat{\xi}^i_n \rightsquigarrow \xi^i_{n+1} \sim M_{n+1}(\widehat{\xi}^i_n, \centerdot)$

Particle occupation measures:

$$\eta^N_n = \frac{1}{N}\sum_{i=1}^N \delta_{\xi^i_n} \longrightarrow_{N\to\infty} \eta_n \quad \text{and} \quad \gamma^N_n(\centerdot) = \eta^N_n(\centerdot) \times \prod_{0\leq p<n} \eta^N_p(G_p) \longrightarrow_{N\to\infty} \gamma_n(\centerdot)$$

- **Advanced signal processing** ($\rightarrow$ filtering/hidden Markov chains/Bayesian methodology)
  - *Signal process:* X_n = Markov chain $\in E$, $\eta_0 = \text{Law}(X_0)$
  - *Observation/Sensor eq.:* $Y_n = H_n(X_n, V_n) \in F$ with $\mathbb{P}(H_n(x_n, V_n) \in dy_n) = g_n(x_n, y_n)\ \lambda_n(dy_n)$

Example:

$$Y_n = h_n(X_n) + V_n \in F = \mathbb{R}, \text{ with Gaussian noise } V_n = \mathcal{N}(0,1)$$

$$\Downarrow$$

$$\mathbb{P}(h_n(x_n) + V_n \in dy_n) = (2\pi)^{-1/2} e^{-\frac{1}{2}(y_n - h_n(x_n))^2} \quad dy_n = \underbrace{\exp\left[h_n(x_n)y_n - h_n^2(x_n)/2\right]}_{g_n(x_n,y_n)} \quad \underbrace{\mathcal{N}(0,1)(dy_n)}_{\lambda_n(dy_n)}$$

Prediction/filtering $\rightarrow$ Feynman-Kac representation $G_n(x_n) = g_n(x_n, y_n)$

$$\eta_n = \text{Law}(X_n \mid \forall 0 \leq p < n \quad Y_p = y_p) \quad \text{and} \quad \widehat{\eta}_n = \text{Law}(X_n \mid \forall 0 \leq p \leq n \quad Y_p = y_p)$$

- **<u>Statistics:</u>** ($\rightarrow$ Sequential MCMC and Feynman-Kac-Metropolis models)

Metropolis potential [π target measure]+[(K,L) pair Markov transitions]

$$G(y_1,y_2)=\frac{\pi(dy_2)L(y_2,dy_1)}{\pi(dy_1)K(y_1,dy_2)}$$

Th.: (**Time reversal formula**), [*A. Doucet, P.DM; (Séminaire Probab. 2003)*]

$$\mathbb{E}^L_\pi(f_n(Y_n,Y_{n-1}\ldots,Y_0)|Y_n=y)=\frac{\mathbb{E}^K_y(f_n(Y_0,Y_1,\ldots,Y_n)\ \{\prod_{0\leq p<n}G(Y_p,Y_{p+1})\})}{\mathbb{E}^K_y(\{\prod_{0\leq p<n}G(Y_p,Y_{p+1})\})}$$

In addition :

$\oplus$ *FK-Metropolis n-marginal:* $\lim_{n\rightarrow\infty}\eta_n=\pi$ (cv. decays $\perp\pi$)

$\oplus$ *Nonhomogeneous models:* (π_n,L_n,K_n)

$\pi_n(dy)\propto e^{-\beta_n V(y)}\ \lambda(dy)$, cooling schedule $\beta_n\uparrow\infty$, mutation s.t. $\pi_n=\pi_n K_n$, and Law$(X_0)=\pi_0$

$$\Downarrow$$

$$G_n(y_1,y_2)=\exp\left[-(\beta_{n+1}-\beta_n)V(y_1)\right]\Longrightarrow\eta_n=\pi_n$$

Spectral analysis (*time homogeneous models*)

- $G\ M(h) = \lambda\ h\ (> 0)$, $\Rightarrow G = \lambda h / M(h) \Longrightarrow \widehat{\gamma}_n(fM(h)) \propto \Psi_h(\eta_0) M_h^n(f)$

  with

  $$\Psi_h(\eta_0)(dx) = \frac{1}{\eta_0(h)}\ h(x)\ \eta_0(dx) \quad \text{and} \quad M_h(x,dy) = \frac{1}{M(h)(x)}\ M(x,dy)h(y)$$

  Then

  $$\Psi_{M(h)}(\widehat{\eta}_n) = \Psi_h(\eta_0) M_h^n(f) = \text{Law}(Y_n) \quad \text{with} \quad Y_n \ \text{ Markov chain } \sim M_h$$

- $\Phi_{0,n} = \Phi^n$ contractive $\Rightarrow \exists\ !\ \eta_\infty = \Phi(\eta_\infty)$

  If M μ-reversible then: $M(Gh) = \lambda h\ (> 0) \iff \eta_\infty = \Psi_h(\mu)$ and $\eta_\infty(G) = \lambda$

Note:

(H)$_m$**:** $M^m(x, \centerdot) \ \geq\ \epsilon\ M^m(y, \centerdot)$ and $G_n(x)\ \geq\ r\ G_n(y) \Rightarrow h = d\eta_\infty/d\mu \in [\epsilon/r^m, r^m/\epsilon]$

h-**Relative entropy**, h convex, $h(ax, ay) = ah(x, y) \in \mathbb{R} \cup \{\infty\}$, $h(1, 1) = 0$

$$H(\eta, \mu) = \int h(d\eta, d\mu) = \int g(d\eta/d\mu)\ d\mu \qquad \text{with} \quad g(x) = h(x, 1)$$

Ex.:

- *Total variation and* $\mathbb{L}_p$*-norms:* $g(t) \propto |t - 1|^p$
- *Boltzmann or Shannon-Kullback entropy:* $g(t) = t \log t$
- *Havrda-Charvat entropy order* $p > 1$*:* $g(t) = \frac{1}{p-1}(t^p - 1)$
- *Kakutani-Hellinger integrals order* $\alpha \in (0, 1)$*:* $g(t) = t - t^\alpha$

Dobushin's contraction coefficient

$$\beta(M) =_{def.} \sup_{x,y} \|M(x, \centerdot\,) - M(y, \centerdot\,)\|_{tv} = \sup_{\eta,\mu} \frac{\|\eta M - \mu M\|_{tv}}{\|\eta - \mu\|_{tv}}$$

Th. [L.Miclo, M. Ledoux, P.DM (PTRF 2003)]

$$H(\mu M, \eta M) \leq \beta(M) H(\mu, \eta)$$

Ex.: $M(x, \centerdot\,) \geq \epsilon M(y, \centerdot\,) \Longrightarrow \beta(M) \leq (1-\epsilon) \Longrightarrow H(\mu M, \eta M) \leq (1-\epsilon) H(\mu, \eta)$

$$\Downarrow$$

Corollary *(Filtering with a wrong initial condition* $\eta'_0 \leadsto \eta'_n \longrightarrow \widehat{\eta'_n}$*)*

$$\mathbb{E}(\text{Ent}(\widehat{\eta_n} \mid \widehat{\eta'_n})) \leq \mathbb{E}(\text{Ent}(\eta_n \mid \eta'_n)) \leq \left[\prod_{p=1}^{n} \beta(M_p)\right] \text{Ent}(\eta_0 \mid \eta'_0)$$

Feynman-Kac semigroups

$$\Phi_{p,n}(\mu)(f) = \frac{\mu Q_{p,n} f}{\mu Q_{p,n} 1} = \frac{\mu(G_{p,n} P_{p,n}(f))}{\mu(G_{p,n})} \qquad \text{with} \quad G_{p,n} =_{def.} Q_{p,n}(1) \quad \text{and} \quad P_{p,n}(f) =_{def.} \frac{Q_{p,n}(f)}{Q_{p,n}(1)}$$

Note:

$$\beta(P_{p,n}) = \sup_{\mu,\eta} \|\Phi_{p,n}(\eta) - \Phi_{p,n}(\mu)\|_{tv} \quad \text{and let} \quad g_{p,n} =_{def.} \sup_{x,y} G_{p,n}(x)/G_{p,n}(y)$$

$$\downarrow$$

Th. $h(x,y)$ suff. regular. $\Rightarrow$ $\exists\ \alpha_h(t)\ \uparrow$ s.t.

$$H(\Phi_{p,n}(\mu), \Phi_{p,n}(\eta)) \ \leq\ \alpha_h(g_{p,n})\ \beta(P_{p,n})\ H(\mu,\eta)$$

Ex.:

$\alpha_h(t) = t$ (total variation norm and Boltzmann entropy), $\alpha_h(t) = t^{1+p}$ (Havrda-Charvat and Kakutani-Hellinger integrals of order p, $\alpha_h(t) = t^3$ ($\mathbb{L}_2$-norms),...

Contraction estimates (H)$_m$: $(M^m(x,\centerdot)\ \geq\ \epsilon\ M^m(y,\centerdot)$ and $G_n(x)\ \geq\ r\ G_n(y)\ )$

Lemma

$$(H)_m \Longrightarrow \forall p\geq 0 \qquad \sup_{n\geq p} g_{p,n}\leq r^m/\epsilon \quad \text{and} \quad \beta(P_{p,p+nm})\leq (1-\epsilon^2/r^{m-1})^n$$

$$\Downarrow$$

Th.

$$\|\Phi_{p,p+nm}(\eta)-\Phi_{p,p+nm}(\mu)\|_{tv}\ \leq\ (1-\epsilon^2/r^{m-1})^n \qquad \Big(\ =(1-\epsilon^2)^n \quad \Longleftarrow\ (H)_1\Big)$$

and

$$H(\Phi_{p,p+nm}(\mu),\Phi_{p,p+nm}(\eta))\ \leq\ \alpha_h(r^m/\epsilon)\ (1-\epsilon^2/r^{m-1})^n\ H(\mu,\eta)$$

Extensions: *nonhomogeneous models, continuous time FK semigroups,...*

Some application model areas

- Asymptotic stability of optimal filters.
- Uniform estimates for particle approximation models (w.r.t time parameter).
- Spectral analysis of FK-Schrödinger s.g., new probab./particle interpretations.
- Long time behavior of Feynman-Kac-Metropolis models.
- Stability properties of infinite population genetic models.