

A NOTE ON MARKOV DIFFUSION

RUIHAN XU

1. MARKOV SEMIGROUP

A **semigroup** $\mathbf{P} = (P_t)_{t \geq 0}$ is a family of operators acting on some suitable function space with the semigroup property $P_t \circ P_s = P_{t+s}$, $P_0 = Id$. We can define a **Markov semigroup** on a measurable state space (E, \mathcal{F}) and Markov process $\{X_t\}_{t \geq 0}$ by

$$P_t f(x) = \mathbb{E}[f(X_t) | X_0 = x].$$

Given transition function $p_t(x, y)$, which can also be interpreted as the probability density at y at time t when $X_0 = x$, we have

$$P_t f(x) = \int_E f(y) p_t(x, y) dy.$$

By duality, the semigroup also act on the set of measure ν via

$$\int_E P_t f d\nu = \int_E f d(P_t^* \nu),$$

where $P_t^* \nu$ is the law of X_t if ν is the law of X_0 . μ is called an **invariant measure** if $P_t^* \mu = \mu$ for every $t \geq 0$.

By Jensen's inequality, for every convex function ϕ $t \geq 0$ and measurable function f on E , we have

$$P_t(\phi(f)) \geq \phi(P_t f).$$

2. INFINITESIMAL GENERATORS AND CARRÉ DU CHAMP OPERATORS

2.1. Infinitesimal Generators. Considering a family of bounded linear operators $(P_t)_{t \geq 0}$ on a Banach space \mathcal{B} with semigroup properties and continuity $P_t f \rightarrow f$ for every f and $t \rightarrow 0$, the Hille-Yosida theory indicates that there is a dense linear subspace of \mathcal{B} , called the **domain** \mathcal{D} of the semigroup $(P_t)_{t \geq 0}$, on which the derivative at $t = 0$ of P_t exists in \mathcal{B} . The operator that maps $f \in \mathcal{D}$ to this derivative $Lf \in \mathcal{B}$ of $P_t f$ at $t = 0$ is a linear (usually unbounded) operator, called the **infinitesimal generator** of the semigroup $(P_t)_{t \geq 0}$, denoted L . Applied to a Markov semigroup $\mathbf{P} = (P_t)_{t \geq 0}$ with state space (E, \mathcal{F}) and invariant measure μ , The infinitesimal generator L of \mathbf{P} in $\mathcal{B} = L^2(\mu)$ is called the Markov generator of the semigroup $\mathbf{P} = (P_t)_{t \geq 0}$ with $L^2(\mu)$ -domain $\mathcal{D}(L)$.

The linearity of the operators P_t , together with the semigroup property, shows that L is the derivative of P_t at any time $t > 0$:

$$\begin{aligned} \frac{1}{s} [P_{t+s} - P_t] &= P_t \left(\frac{1}{s} [P_s - Id] \right) = \left(\frac{1}{s} [P_s - Id] \right) P_t, \\ \partial_t P_t &= L P_t = P_t L \quad \text{by letting } s \rightarrow 0. \end{aligned}$$

Note that replacing L with cL for some $c > 0$ amount to the time change $t \rightarrow ct$.

Recall that $L = \lim_{t \rightarrow 0^+} \frac{P_t - \text{Id}}{t}$. For convex function ϕ , if $f \in \mathcal{D}(L)$ and $\phi(f) \in \mathcal{D}(L)$, we have

$$L\phi(f) \geq \phi'(P_0 f) Lf = \phi'(f) Lf.$$

Note that if μ is the invariant measure, then $\int_E P_t f d\mu = \int_E f d\mu$, thus $Lf = 0$.

Example 2.1 (jump Markov process). Let X_t be a jump Markov process on \mathcal{X} with rate kernel $\lambda(x, dy)$, then it's holding rate at state x is

$$\Lambda(x) = \int_{\mathcal{X}} \lambda(x, dy),$$

which means its holding time at state x is exponentially distributed with rate $\Lambda(x)$ and the density that it chooses y as next state is $\frac{\lambda(x, dy)}{\Lambda(x)}$.

Then infinitesimal generator L of jump Markov process X_t is

$$(Lf)(x) = \int_{\mathcal{X}} (f(y) - f(x)) \lambda(x, dy).$$

2.2. Carré du Champ Operators. Assume that we are given a vector subspace \mathcal{A} of the domain $\mathcal{D}(L)$ such that for every pair (f, g) of functions in \mathcal{A} , the product fg is in the domain $\mathcal{D}(L)$ (\mathcal{A} is an algebra). Then the bilinear map

$$\Gamma(f, g) = \frac{1}{2} [L(fg) - fLg - gLf]$$

defined for every $(f, g) \in \mathcal{A} \times \mathcal{A}$ is called the **carré du champ operator** of the Markov generator L . To lighten the notation, we set $\Gamma(f) = \Gamma(f, f)$.

Example 2.2. The Laplacian $L = \Delta$ on \mathbb{R}^n gives rise to the standard carré du champ operator $\Gamma(f, g) = \nabla f \cdot \nabla g$ (the usual scalar product of the gradients of f and g) for smooth functions f, g on \mathbb{R}^n .

Note that $\phi(x) = x^2$ is convex, thus in the limit as $t \rightarrow 0$ $L(f^2) \geq 2fLf$. It follows that the carré du champ operator is positive on \mathcal{A} in the sense that

$$\Gamma(f, f) \geq 0.$$

By Cauchy-Schwarz inequality it immediately yields

$$\Gamma(f, g)^2 \leq \Gamma(f) \Gamma(g), \quad (f, g) \in \mathcal{A} \times \mathcal{A}.$$

3. FOKKER-PLANCK EQUATIONS

Recall that given a Markov semigroup $\mathbf{P} = (P_t)_{t \geq 0}$ with infinitesimal generator L , there might exist a density kernel $p_t(x, y)$ with respect to some measure m , such that

$$P_t f(x) = \int_E f(y) p_t(x, y) dm(y).$$

Then $p_t(x, y)$ is the solution of the heat equation

$$\partial_t p_t(x, y) = L_x p_t(x, y), \quad p_0(x, y) dm(y) = \delta_x,$$

where L_x denotes the operator L acting on the x variables. This expresses that

$$\partial_t P_t f = L P_t f.$$

One may also consider the dual equation, called the **Fokker-Planck equation** (or Kolmogorov forward equation)

$$\partial_t p_t(x, y) = L_y^* p_t(x, y),$$

where L^* is the adjoint of L with respect to the reference measure m in the sense that

$$\int_E f L^* g dm = \int_E g L f dm.$$

Proposition 3.1. *Consider infinitesimal generator L on a diffusion process $\{X_t\}_t \geq 0 \in \mathbb{R}^n$*

$$dX_t = m(t, X_t)dt + \sigma(t, X_t)dW_t$$

where W_t is standard Brownian motion in \mathbb{R}^n . Define $\Sigma = \sigma\sigma^\top$ Then we have

$$\begin{aligned} Lf &= \sum_{i=1}^n m_i(x, t) \partial_{x_i} f + \frac{1}{2} \sum_{i,j} \Sigma_{ij}(x, t) \frac{\partial^2 f}{\partial x_i \partial x_j} \\ &= m(x, t) \cdot \nabla f + \frac{1}{2} \text{Tr}(\Sigma(x, t) \nabla^2 f), \\ L^* p &= - \sum_{i=1}^n \frac{\partial}{\partial x_i} (m_i(x, t) p(x)) + \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} (\Sigma_{ij}(x, t) p(x)) \\ &= -\nabla \cdot (m(x, t) p(x)) + \frac{1}{2} \nabla \cdot (\Sigma(x, t) \nabla p(x)). \end{aligned}$$

4. DIRICHLET FORMS AND SPECTRAL DECOMPOSITION

4.1. Dirichlet Forms. We call a Markov semigroup $\mathbf{P} = (P_t)_{t \geq 0}$ a **symmetric Markov semigroup** with respect to the invariant μ (or reversible measure) if for all function $f, g \in \mathbb{L}^2(\mu)$ and $t \geq 0$ we have

$$\int_E f P_t g d\mu = \int_E g P_t f d\mu.$$

If the semigroup $\mathbf{P} = (P_t)_{t \geq 0}$ admits density kernels $p_t(x, y)$, then $p_t(x, y)$ is symmetric. Differentiating the equation leads to

$$\int_E f L g d\mu = \int_E g L f d\mu.$$

For a symmetric Markov semigroup $\mathbf{P} = (P_t)_{t \geq 0}$ with infinitesimal generator L , reversible measure μ and carré du champ operator Γ on a class \mathcal{A} of functions on E , consider the (symmetric) bilinear operator

$$\mathcal{E}(f, g) = \int_E \Gamma(f, g) d\mu, \quad (f, g) \in \mathcal{A} \times \mathcal{A}.$$

Note that since $\int_E \Gamma(f, g) d\mu = \frac{1}{2} \int_E L(fg) - f Lg - g Lf d\mu = - \int_E f Lg d\mu$, we have

$$\mathcal{E}(f, g) = \int_E \Gamma(f, g) d\mu = - \int_E f Lg d\mu$$

A direct computation gives

$$\partial_t \int_E (P_t f)^2 d\mu = 2 \int_E P_t f L P_t f d\mu = -2\mathcal{E}(P_t f),$$

and thus

$$\int_E f^2 d\mu - \int_E (P_t f)^2 d\mu = 2 \int_0^t \mathcal{E}(P_s f) ds \geq 2t\mathcal{E}(P_t f).$$

The inequality is by noting that

$$\begin{aligned} \partial_t \mathcal{E}(P_t f) &= -\partial_t \int_E P_t f L P_t f d\mu \\ &= -\int_E (L P_t f)^2 d\mu - \int_E P_t f L^2 P_t f d\mu \\ &= -2 \int_E (L P_t f)^2 d\mu. \end{aligned}$$

Changing t into $\frac{t}{2}$ we have

$$\begin{aligned} \mathcal{E}(P_t f) &= \frac{1}{t} \left[\int_E f^2 d\mu - \int_E (P_{t/2} f)^2 d\mu \right] \\ &= \frac{1}{t} \left[\int_E f^2 d\mu - \int_E f P_{t/2} P_{t/2} f d\mu \right] \\ &= \frac{1}{t} \left[\int_E f^2 d\mu - \int_E f P_t f d\mu \right] \\ &= \frac{1}{t} \left[\int_E f(f - P_t f) d\mu \right]. \end{aligned}$$

4.2. Spectral Decomposition. For the generator of a symmetric semigroup, we can analyze it by conducting a spectral decomposition. Assume that $(e_k)_{k \in \mathbb{N}}$ is a Hilbertian basis of $\mathbb{L}_2(\mu)$ consisting of eigenfunctions of L with corresponding sequence of eigenvalues $(\lambda_k)_{k \in \mathbb{N}}$. Hence

$$-L e_k = \lambda_k e_k, \quad k \in \mathbb{N}.$$

The reason of negative sign can be justified by observing that

$$\mathcal{E}(e_k) = \int_E \Gamma(e_k, e_k) d\mu = - \int_E e_k L e_k d\mu = \lambda_k \int_E e_k^2 d\mu = \lambda_k,$$

where $\mathcal{E}(f)$ is sometimes interpreted as the energy of f .

Since $P_t = e^{tL}$, we can decompose a function $f = \sum_{k \in \mathbb{N}} f_k e_k$ and have

$$P_t f = \sum_{k \in \mathbb{N}} e^{-\lambda_k t} f_k e_k.$$

And the density kernel can be represented as

$$p_t(x, y) = \sum_{k \in \mathbb{N}} e^{-\lambda_k t} e_k(x) e_k(y)$$

5. POINCARÉ INEQUALITY

The spectral decomposition shows that as time evolves, $P_t f$ will converge to a constant, which corresponds to eigenvalue $\lambda_0 = 0$. The speed of convergence is actually governed by the inverse of spectral gap, which is defined as

$$\kappa := \sup\left\{\frac{\text{var}_\pi(f)}{\mathcal{E}(f)}; f \in \mathcal{D}(\mathcal{E})\right\},$$

where π is the invariant measure, and $\mathcal{D}(\mathcal{E})$ is the core domain of Dirichlet form \mathcal{E} . We can describe spectral gap in an equivalent inequality form, which is called a κ -Poincaré inequality (κ -PI or $P(\kappa)$) or spectral gap inequality:

$$\text{var}_\pi(f) = \int_E f^2 d\pi - \left(\int_E f d\pi\right)^2 \leq \kappa \mathcal{E}(f).$$

Applying the Poincaré inequality $P(\kappa)$ to an eigenfunction e_k of opposite infinitesimal generator $-L$ yields

$$\int_E e_k^2 d\pi = \text{var}_\pi(e_k) \leq \kappa \mathcal{E}(e_k) = -\kappa \int_E e_k L e_k d\pi = \kappa \lambda_k \int_E e_k^2 d\pi,$$

which means any nonzero eigenvalue λ_k is greater than $\frac{1}{\kappa}$. The Poincaré inequality immediately admits the following useful lemma

Lemma 5.1. *Under the Poincaré inequality $P(\kappa)$, for a function $f \in \mathcal{D}(\mathcal{E})$ such that $f = 0$ outside a set $A \in \mathcal{F}$ with $\pi(A) < 1$, it holds that*

$$\int_A f^2 d\pi \leq \frac{\kappa}{1 - \pi(A)} \mathcal{E}(f).$$

To see the reason that spectral gap controls convergence rate, we introduce the variance decay theorem

Theorem 5.2. *Given a Markov Triple (E, μ, Γ) with associated Markov semigroup $\mathbf{P} = (P_t)_{t \geq 0}$, the following assertions are equivalent:*

- (1) (E, μ, Γ) satisfies a Poincaré inequality $P(C)$.
- (2) For every function $f \in \mathbb{B}\mathbb{R}$ in $\mathbb{L}^2(\mu)$, and every $t \geq 0$,

$$\text{var}_\mu(P_t f) \leq e^{-2t/C} \text{var}_\mu(f).$$

- (3) For every function $f \in \mathbb{L}^2(\mu)$, there exists a constant $c(f) > 0$ (possibly depending on f) such that, for every $t \geq 0$,

$$\text{var}_\mu(P_t f) \leq c(f) e^{-2t/C}.$$

The proof is straight forward by noting that

$$\begin{aligned} \frac{d}{dt} \text{var}(P_t f) &= \int_E 2(P_t f) L(P_t f) d\pi - 2 \int_E P_t f d\pi \int_E L P_t f d\pi \\ &= -2 \mathcal{E}(P_t f) \leq \frac{2}{C} \text{var}(P_t f). \end{aligned}$$

5.1. Perturbation propositions.

Proposition 5.3 (Holley–Stroock perturbation principle). *Assume that the Markov Triple (E, μ, Γ) satisfies a Poincaré inequality $P(C)$. Moreover, suppose there exists a constant $b \in (0, \infty)$ such that*

$$\frac{1}{b} \leq \frac{d\mu'}{d\mu} \leq b$$

Then μ' satisfies a Poincaré inequality $P(b^2C)$.

Proposition 5.4 (Tensorization of Poincaré Inequalities). *If (E_1, μ_1, Γ_1) and (E_2, μ_2, Γ_2) satisfy Poincaré inequalities with respective constants C_1 and C_2 , then the product Markov Triple $(E_1 \times E_2, \mu_1 \otimes \mu_2, \Gamma_1 \oplus \Gamma_2)$ satisfies a Poincaré inequality with constant $C = \max(C_1, C_2)$.*

Theorem 5.5 (Muckenhoupt’s criterion on a real line \mathbb{R}). *Let μ be a probability measure on the Borel sets of \mathbb{R} with density p with respect to the Lebesgue measure, and let m be a median of μ (i.e. $\mu([m, +\infty)) \geq \frac{1}{2}$ and $\mu((-\infty, m]) \geq \frac{1}{2}$). Set*

$$B_+ = \sup_{x > m} \mu([x, +\infty)) \int_m^x \frac{1}{p(t)} dt$$

$$B_- = \sup_{x < m} \mu((-\infty, x]) \int_x^m \frac{1}{p(t)} dt$$

Then, a necessary and sufficient condition in order that μ satisfies a Poincaré inequality $P(C)$ is that $B = \max(B_+, B_-) < \infty$. Moreover, $B \leq 2C \leq 8B$ where C is the Poincaré constant of μ .