401-3382-25L: Log-Sobolev Inequalities and Markov Semigroups

Lecture 1 – Introduction to Markov semigroups

Lecturer: Yuansi Chen *Week 1-2* *Spring 2025*

High level motivation of this course. Analytic, probabilistic and geometric aspects of Markov semigroup, and their interplay. More discussed in class.

Key concepts:

- Markov processes
- Markov semigroup
- Poincaré inequality and Log-Sobolev inequality
- Objects in Markov semigroups
  - Kernels, Chapman-Kolmogorov equations, Infinitesimal generator, carré du champ operator, Fokker-Planck equation, etc.
- Example: heat semigroup on Euclidean space

1.1 Introduction

We start with some *probability space* $(\Omega, \Sigma, \mathbb{P})$, where

- Ω is sample space, which is the set of all possible outcomes
- Σ is event space, which is the set of events, where each event is a subset of outcomes in the sample space
- $\mathbb{P}$ is a probability function, which assigns, to each event in Σ, a probability, which is a number between 0 and 1.

A *random variable* is a measurable function $X : \Omega \to E$ from the sample space Ω to a measurable space $(E, \mathcal{F})$.

Example 1. *Take the standard Euclidean space* $\mathbb{R}^n$ *equipped with its Borel* σ*-field.*

A *process* $(X_t)_{t\geq 0}$ is a family of random variables constructed on $(\Omega, \Sigma, \mathbb{P})$, with values in E. Let $\mathcal{F}_t := \sigma(X_u : u \leq t)$, $t \geq 0$, be the *natural filtration* of $(X_t)_{t\geq 0}$.

Markov property The Markov property indicates that for $t > s$, the law of X_t given $\mathcal{F}_s$ is the law of X_t given X_s. That is,

$$\mathbb{P}(X_t \in A \mid \mathcal{F}_s) = \mathbb{P}(X_t \in A \mid X_s).$$

We also assume the Markov process is *time homogeneous* (which is the only case we consider), that is, the law of X_t given X_s is the law of X_{t-s} given X_0.

Markov process The process $(X_t)_{t\geq 0}$ satisfying the Markov property is said to be a Markov process.

Markov semigroup associated with a Markov process Given a Markov process $(X_t)_{t\geq 0}$, we associate a semigroup to the process as follows

$$P_t f(x) = \mathbb{E}\left[f(X_t) \mid X_0 = x\right], t \geq 0$$

for every bounded and measurable function $f : E \to \mathbb{R}$, for every $x \in E$.

Key properties of Markov semigroup

(i) For every $t \geq 0$, P_t is a linear operator sending bounded measurable functions on $(E, \mathcal{F})$ to bounded measurable functions.

(ii) $P_0 = \mathrm{Id}$, (initial condition)

(iii) $P_t \mathbf{1} = \mathbf{1}$, (mass conservation)

(iv) If $f \geq 0$, then $P_t f \geq 0$, (positivity preserving)

(v) For every $t, s \geq 0$, $P_{t+s} = P_t \circ P_s$, (semigroup property)

Proof. as exercise □

Invariant measure (or stationary measure) A measure μ on $(E, \mathcal{F})$ is said to be *invariant for P (or stationary)* if

$$\int_E P_t f d\mu = \int_E f d\mu, \forall t \geq 0,$$

for every bounded and measurable f.

The semigroup (P_t) can be seen as acting on measures as well through the following duality

$$\int_E f d(\mu P_t) = \int_E P_t f d\mu.$$

If μ is a probability measure, then μP_t is the law of X_t when μ is the law of X_0.

The last condition required to deal with a Markov semigroup is the following *continuity property*

(vi) For every $f \in \mathbb{L}^2(\mu)$, $P_t f$ converges to f in $\mathbb{L}^2(\mu)$ as $t \to 0$ (continuity property).

General construction of Markov semigroups (without relying on X) (Def 1.2.2) A family of operators $(P_t)_{t\geq 0}$ defined on the bounded measurable functions on a state space $(E, \mathcal{F})$ with invariant σ-finite measure μ satisfying the properties (i)-(vi) is called a Markov semigroup of operators.

1.2 Quick introduction of log-Sobolev inequalities

Variance If f is a square integral function with respect to μ, its *variance* is

$$\mathrm{Var}_\mu(f) = \int f^2 d\mu - \left(\int f d\mu\right)^2 = \mathrm{Var}_{X\sim\mu}(f(X))$$

Entropy If f is a nonnegative function, integrable with respect to μ, then its *entropy* is

$$\mathrm{Ent}_\mu(f) = \int f \log f d\mu - \left(\int f d\mu\right) \log \left(\int f d\mu\right).$$

Poincaré inequality We say μ satisfies the Poincaré inequality with constant C if

$$\mathrm{Var}_\mu(f) \leq C \int |\nabla f|^2 \, d\mu,$$

for every function f whose gradient belongs to $\mathbb{L}^2(\mu)$.

Log-Sobolev inequality We say μ satisfies the logarithmic-Sobolev inequality with constant C if

$$\mathrm{Ent}_\mu(f^2) \leq 2C \int |\nabla f|^2 \, d\mu.$$

for every function f whose gradient belongs to $\mathbb{L}^2(\mu)$.

We just introduce them for the purpose explaining the title of the course. We will discuss them in more details later.

1.3 Objects of interest in Markov semigroups

- Kernels
- Chapman-Kolmogorov equations
- Infinitesimal generator
- carré du champ operator
- Fokker-Planck equation

Kernel representation (Prop 1.2.3) Given a good measure space $(E, \mathcal{F}, \mu)$, if $P\mathbf{1} = \mathbf{1}$, and P is bounded on $\mathbb{L}^1(\mu)$, P can be represented as a probability kernel $p(x, A)$,

$$Pf(x) = \int_E f(y)p(x, dy)$$

for every bounded or positive measurable function f on E and (μ-almost) every $x \in E$.

We know how to define P_t starting from X_t (we call it Markov semigroup associated with a Markov process). How to construct a Markov process from P_t defined using (i)-(vi)?

Chapman-Kolmogorov equation Let $(P_t)_{t\geq 0}$ on $\mathbb{L}^2(\mu)$. The semigroup property $P_t \circ P_s = P_{t+s}$ translates to the kernels $p_t(x, dy)$ via the composition property, for all $t, s \geq 0, x \in E$,

$$p_{t+s}(x, dy) = \int_{z\in E} p_t(z, dy)p_s(x, dz).$$

Using the above equation, one may construct, starting from any point $x \in E$, a Markov process $(X_t)_{t\geq 0}$ on E by specifying the distribution of $(X_{t_1}, \dots, X_{t_k}), 0 \leq t_1 \leq \cdots \leq t_k$, as

$$\mathbb{E}\left[f(X_{t_1}, \dots, X_{t_k})\right] = \int f(y_1, \dots, y_k)p_{t_k - t_{k-1}}(y_{k-1}, dy_k) \cdots p_{t_1}(x, dy_1).$$

Given a Markov semigroup $(P_t)_{t\geq 0}$ with its kernel $p_t(\cdot, \cdot)$, we therefore have candidates (by specifying its law on a finite-dimensional space) of a Markov processes starting from a point $x \in E$. However, in general, the finite dimensional description is not enough to characterize the full law of the Markov process. It needs some extra work such as assuming regular paths of X_t, but we don't deal with it in this course. See e.g. Kolmogorov extension theorem for more details.

YC — Lecture 01 stopped at Chapman-Kolmogorov equation, after deriving it for heat semigroup

It is pretty clear that from the semigroup property, that (P_t) is completely determined by its behavior when t tends to 0. It is natural to consider differentiating $P_t f$ at $t = 0$.

Infinitesimal generator The *infinitesimal generator* L associated with a Markov semigroup $(P_t)_{t\geq 0}$ is the operator defined by

$$Lf := \lim_{t\to 0^+} \frac{P_t f - f}{t},$$

for all functions f for which the above limit exists.

Remark 1. *Due to technical difficulties which is hard to resolve in a short time, we don't specify the set of functions for which the limit exists. The domain of L is usually worth thinking twice when working with Markov semigroups. In general, according to the Hille-Yosida theory for Markov semigroups on Banach space $\mathfrak{B}$, there exists a dense linear subspace of $\mathfrak{B}$, on which the derivative at $t = 0$ exists in $\mathfrak{B}$. See [BGL13] Chap 1.4.1 and Appendix A.1.*

Carré du champ operator *Carré du champ operator* is the bilinear map

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

defined for f and g on a vector subspace $\mathcal{A}$ (also an algebra) of the domain of L. To lighten the notation, we set $\Gamma(f) = \Gamma(f, f)$.

Remark 2. *One may ask why it is called carré du champ? In French, carré = square, champ = field. In the simple example of $L = \Delta$ the Laplacian on $\mathbb{R}^n$, we have*

$$\Gamma(f, g) = \nabla f \cdot \nabla g.$$

So $\Gamma(f, f) = |\nabla f|^2$.

Proposition 1.3.1. *The carré du champ operator is positive on $\mathcal{A}$ in the sense that*

$$\Gamma(f) \geq 0, f \in \mathcal{A}.$$

Proof. We have

$$P_t(f^2)(x) = \mathbb{E}\left[f(X_t)^2 \mid X_0 = x\right] \geq \mathbb{E}\left[f(X_t) \mid X_0 = x\right]^2 = (P_t(f)(x))^2$$

where the inequality follows from Cauchy-Schwarz inequality (or Jensen's inequality). Taking the limit $t \to 0$, we have

$$L(f^2) \geq 2fLf.$$

□

Proposition 1.3.2. *If L is a second-order differential operator of the form*

$$Lf(x) = \sum_{ij} a_{ij}(x)\partial_{ij} f(x) + \sum_i b_i(x)\partial_i f(x)$$

then the carré du champ operator recovers a weighted squared gradient

$$\Gamma(f)(x) = \sum_{ij} a_{ij}(x)\partial_i f(x)\partial_j f(x)$$

Proof. exercise □

Fokker-Planck equations provide a dual point of view on the Chapman-Kolmogorov equation.

Fokker-Planck equations Given a Markov semigroup $(P_t)_{t\geq 0}$ with kernel $p_t(x,y)dy$ and generator L. The function $p_t(x,y), t > 0, x \in E$ is a solution to the following partial differential equation

$$\partial_t p_t(x,y) = L_x p_t(x,y), p_0(x,y)dy = \delta_x \tag{1.1}$$

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

$$\partial_t P_t f = L P_t f. \tag{1.2}$$

The dual equation is called the *Fokker-Planck equation or Kolmogorov forward equation*, for $t > 0$,

$$\partial_t p_t(x,y) = L_y^* p_t(x,y), p_0(x,y)dy = \delta_x \tag{1.3}$$

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

$$\int_E f L^* g dy = \int_E g L f dy,$$

for suitable functions f, g. Eq. (1.1) describes how the probability density function evolves over time (forward in time). In contrast, the Kolmogorov backward equation Eq. (1.2) describes the evolution of expected functionals of the process.

1.3.1 Other properties of a Markov semigroup

Reversibility (or symmetry). A Markov semigroup (P_t) is said to be *symmetric* with respect to the invariant measure μ, or μ is *reversible* for (P_t), if for all functions $f, g \in \mathbb{L}^2(\mu)$ and $t \geq 0$,

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

If in addition, P_t admits a kernel $p_t(x,y)$ with respect to μ, then p_t is a symmetric function on $E \times E$. Additionally, its infinitesimal generator L is a symmetric operator on $\mathbb{L}^2(\mu)$, that is, $\int_E fLgd\mu = \int_E gLfd\mu$.

Proposition 1.3.3. *If μ is reversible for (P_t), then μ is invariant measure for (P_t).*

Proof. Take g to be $\mathbf{1}$, the all one function. □

Contractivity If μ is stationary, then P_t extends to a continuous operator on $\mathbb{L}^p(\mu)$ for any $q \in [0, \infty]$. Moreover, P_t is a contraction:

$$\|P_t f\|_q \le \|f\|_q, \forall q \in [1, \infty], \forall f \in \mathbb{L}^q(\mu).$$

Proof. Let f be bounded and in $\mathbb{L}^q(\mu)$. By Jensen's inequality we have $|P_t f(x)|^q \le P_t(|f|^q)(x)$, pointwise. Integrating and using stationarity, we get

$$\int_E |P_t f|^q \, d\mu \le \int_E P_t(|f|^q) d\mu = \int_E |f|^q \, d\mu.$$

Since bounded functions are dense in $\mathbb{L}^q(\mu)$, we conclude.

□

For the following properties, see [BGL13] Chap 1, we cover them when needed.

- Abstract martingale problem
- Dirichlet form and spectral decomposition.
- Ergodicity
- Convergence to equilibrum

1.4 First example: heat semigroup

Gaussian measure on $\mathbb{R}^n$

$$\gamma^n(A) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_A \exp\left(-\frac{|x|^2}{2}\right) dx$$

Its density with respect to Lebesgue measure is $\frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{|x|^2}{2}\right)$.

Brownian motion Brownian motion is a process $(B_t)_{t\geq 0}$ in $\mathbb{R}^n$ such that $B_0 = 0$ and satisfying

1. (independent increments) For all $0 < t_1 < \cdots < t_k$, the random variables $(B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_k} - B_{t_{k-1}})$ are mutually independent.
2. (law of the increments) For all $0 \leq s < t < \infty$,
$$B_t - B_s \sim \mathcal{N}(0, (t-s)\mathbb{I}_n)$$
3. (continuity of the paths) Almost surely, $t \mapsto B_t$ is continuous.

In this section, we focus on the process

$$\tilde{B}_t := B_{2t}, t \geq 0.$$

Derive its associated Markov semigroup

$$\begin{aligned} P_t f(x) &= \mathbb{E}\left[f(\tilde{B}_t) \mid \tilde{B}_0 = x\right] \\ &\overset{(i)}{=} \mathbb{E}_{G\sim\mathcal{N}(0,\mathbb{I}_n)}\left[f(x+\sqrt{2t}G)\right] \\ &= \int f(x+\sqrt{2t}z)\frac{1}{(2\pi)^{n/2}}\exp\left(-\frac{|z|^2}{2}\right)dz \end{aligned}$$

(i) use the Gaussian law of the increments.

Find an invariant measure $d\mu = dx$ the Lebesgue measure is invariant. Note that it is not a bounded measure on $\mathbb{R}^n$.

Derive its kernel Using change of variable formula for $y = x + \sqrt{2t}z$, we have

$$\begin{aligned} P_t f(x) &= \int f(x+\sqrt{2t}z)\frac{1}{(2\pi)^{n/2}}\exp\left(-\frac{|z|^2}{2}\right)dz \\ &= \int f(y)\frac{1}{(4\pi t)^{n/2}}\exp\left(-\frac{|y-x|^2}{4t}\right)dy \end{aligned} \tag{1.4}$$

We can identify the kernel

$$p_t(x, dy) = \frac{1}{(4\pi t)^{n/2}}\exp\left(-\frac{|y-x|^2}{4t}\right)dy$$

Derive its Chapman-Kolmogorov equation It appears as a consequence of the fact that the sum of two independent Gaussian vectors in $\mathbb{R}^n$ with respective covariance matrices $2t\mathbb{I}_n$ and $2s\mathbb{I}_n$ is a Gaussian vector with covariance $2(t+s)\mathbb{I}_n$.

Derive its infinitesimal generator We would like to calculate

$$Lf := \lim_{t\to 0^+} \frac{P_t f - f}{t},$$

so take derivative of P_t with respect to t at the neighborhood of 0. For semigroups on $\mathbb{R}^n$, the trick is to observe in Eq. (1.4) that we can move t into f and then take derivative. For $t > 0$,

$$\begin{aligned}
\frac{d}{dt}P_t f &= \int \left\langle \frac{\sqrt{2}}{2} t^{-\frac{1}{2}} z, \nabla f(x+\sqrt{2t}z) \right\rangle \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{|z|^2}{2}\right) dz \\
&\overset{(i)}{=} 0 + \sum_{i=1}^n \int \partial_i^2 f(x+\sqrt{2t}z) \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{|z|^2}{2}\right) dz
\end{aligned}$$

(i) applies integration by parts on each coordinate ($u' = z_i \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{|z|^2}{2}\right)$), the boundary terms are zero, because of the Gaussian tail and conditions on f. For example, we could consider the infinitesimal generator on the class of smooth ($\mathcal{C}^\infty$) compacted supported functions.

Finally, take the limit $t \to 0^+$, we obtain

$$Lf = \Delta f$$

where $\Delta f = \sum_{i=1}^n \partial_i^2 f$ is the Laplacian on $\mathbb{R}^n$ defined on smooth functions.

Derive its carré du champ

$$\begin{aligned}
\Gamma(f,g) &= \frac{1}{2}\left[L(fg) - fLg - gLf\right] \\
&= \frac{1}{2}\left[\Delta(fg) - f\Delta g - g\Delta f\right] \\
&= \sum_{i=1}^n (\partial_i f)(\partial_i g) \\
&= \nabla f \cdot \nabla g.
\end{aligned}$$

Note that

$$\begin{aligned}
&\partial_i(fg) = (\partial_i f)g + (\partial_i g)f \\
&\partial_i^2(fg) = (\partial_i^2 f)g + (\partial_i f)(\partial_i g) + (\partial_i^2 g)f + (\partial_i f)(\partial_i g).
\end{aligned}$$

Derive its Fokker-Planck equation Since μ is also reversible to P_t, then

$$L^* = L = \Delta.$$

The Fokker-Planck equation becomes

$$\partial_t p_t = \Delta p_t, \quad p_0(y)dy = \delta_x.$$

This is exactly the heat equation, which admits a solution in the form

$$y \mapsto p_t(x, y) = \frac{1}{(4\pi t)^{n/2}} \exp\left(-\frac{|y-x|^2}{4t}\right)$$

1.5 Exercises

Exercise 1 (Generator of Langevin semigroup). *Consider the solution to the following Langevin stochastic differential equation*

$$dX_t = -\nabla U(X_t)dt + \sqrt{2}dB_t.$$

Derive its associated Markov semigroup and its infinitesimal generator L. First, without justification on regularity conditions. Second, think about the domain of L

Exercise 2. *Prove Proposition 1.3.2.*

Exercise 3 (Uniqueness of Kolmogorov forward equation solution). *Let $(P_t)_{t\geq 0}$ be a Markov semigroup, with μ is its stationary measure and L is its infinitesimal generator. Suppose $u_t = P_t f$ solves Kolmogorov forward equation*

$$\partial_t u_t = Lu_t, u_0 = f.$$

Prove that if a solution u_t exists and satisfies the semigroup property, then it is unique.

Exercise 4 (Heat semigroup). *Let $(P_t)_{t\geq 0}$ be the heat semigroup on $\mathbb{R}^n$.*

1. *Prove that if f is bounded on $\mathbb{R}^n$, then*

$$\sup_x P_t f(x) \leq \sup_x f(x).$$

2. *Suppose $f > 0$. Prove that for any $t > 0, x, y \in \mathbb{R}^n$*

$$P_t f(x) \leq P_t f(y) \exp\left(\frac{|x-y|^2}{4t}\right).$$

Exercise 5. *Let $(P_t)_{t\geq 0}$ be a Markov semigroup. Prove that for any differentiable function f with bounded gradient on $\mathbb{R}^n$:*

$$|\nabla P_t f| \leq P_t |\nabla f|.$$

Bibliography

[BGL13] Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348. Springer Science & Business Media, 2013.