Lectures on finite Markov processes

Plan :

1- Finite Markov generators

2 - Semi groups

3 - Spectral gap

4 - Relative entropy

5 - Simulated annealing

6 - A first look at diffusion processes

1- Finite Markov generators

Process: here means continuous time
(chain: discrete time)

In the 5 first lectures, the state space will be a finite set V.

Markov property:

A family $X := (X(t))_{t\geq 0}$ of random variables, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in V, is said to be Markovian when for any

$$0 \leq t_0 < t_1 < \cdots < t_{n+1}, \quad n \in \mathbb{Z}_+$$

we have for the conditional laws:

$$\mathcal{L}(X(t_{n+1}) \mid X(t_n), \ldots, X(t_0))$$

$$= \mathcal{L}(X(t_{n+1}) \mid X(t_n))$$

namely, for any $x_0, x_1, \ldots, x_{n+1} \in V$

such that $\mathbb{P}[X(t_0) = x_0, \ldots X(t_n) = x_n] > 0$,

$$\mathbb{P}[X(t_{n+1}) = x_{n+1} \mid X(t_n) = x_n, \dots X(t_0) = x_0]$$

$$= \mathbb{P}[X(t_{n+1}) = x_{n+1} \mid X(t_n) = x_n]$$

Equivalently, to come back to Markov chains, for any (strictly) increasing sequence $(t_n)_{n \in \mathbb{Z}_+}$, $(X(t_n))_{n \in \mathbb{Z}_+}$ has to be a (V-valued) Markov chain.

In English, X is a Markov process if at any time (say t_n), the future $(X(t_{n+1}))$ depends on the past $(X(t_n), X(t_{n-1}), \dots, X(t_0))$ only through the present $(X(t_n))$.

As the past can be richer than the past behavior of X, here is a natural extension:

Let $(\mathcal{F}(t))_{t \geq 0}$ be a filtration such that $(X(t))_{t \geq 0}$ is adapted.

The process X is said to be Markovian with respect to $(\mathcal{F}(t))_{t\geq 0}$ if $\forall\, t, s \geq 0$

$$\mathcal{L}(X(s+t) \mid \mathcal{F}(s))$$
$$= \mathcal{L}(X(s+t) \mid X(s)).$$

The notion of Markovianity given at the beginning was with respect to the filtration generated by X.

The Markov process X is said to be time-homogeneous if there exists a family of matrices $(P_t)_{t\geq 0}$ such that $\forall\, t, s \geq 0$, $\forall\, x, y \in V$,

$$\mathbb{P}[X(s+t) = y \mid X(s) = x] = P_t(x,y)$$

Except otherwise mentioned, from now on the Markov processes will be assumed to be time-homogeneous.

With the above definitions, a finite Markov process can be quite wild. Consider μ a probability measure on V which is not concentrated on a point.

Let $(X(t))_{t\geq 0}$ be a family of independent variables distributed according to μ.

It is a Markov process and we have

$\forall t>0, \quad \forall x, y \in V, \quad P_t(x, y) = \mu(y)$

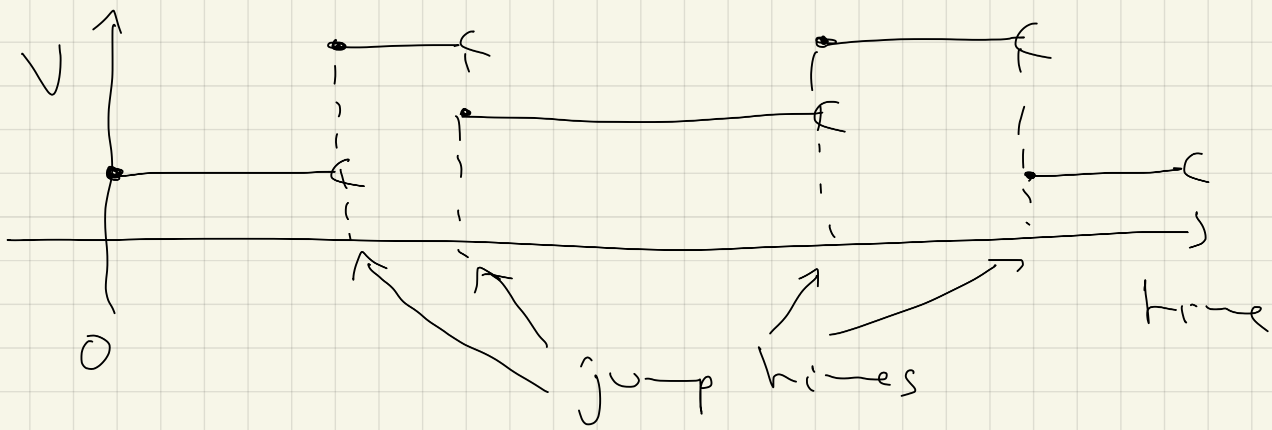
(P_0 is always the identity matrix).

It can be shown that the trajectory

$$\mathbb{R}_+ \ni t \longmapsto X(t) \qquad (1)$$

is not measurable (a.s.)

To avoid this situation, we impose some regularity on the trajectories: (1) is assumed to be <u>càdlàg</u>, ie. right continuous with left limits. Namely they look like (with $|V|=3$)



A process $(X(t))_{t\geq 0}$ whose trajectories are càdlàg is said to be <u>strong Markov</u> if for any finite stopping-time (with respect to $(\mathcal{F}(t))_{t\geq 0}$) τ, we have

$\forall\, t\geq 0,\ \forall\, y\in V$

$$\mathbb{P}[X(\tau+t)=y \mid \mathcal{F}(\tau)]$$

$$= P_r(X(\tau), y)$$

where $\mathcal{F}(\tau)$ is the σ-field of events before τ. Note that the random variables

$$\Omega \ni \omega \mapsto X(\tau(\omega))(\omega)$$

is indeed measurable due to the càdlàg assumption on the trajectories.

Finite càdlàg Markov processes are conveniently described through Markov generators.

A $V \times V$ matrix $L := (L(x,y))_{x,y \in V}$ is a <u>Markov generator</u> if

(i) $\forall x \neq y \in V, \quad L(x,y) \geq 0$

(ii) $\forall x \in V, \quad \sum_{y \in V} L(x,y) = 0$

In particular

$$L(x) := -L(x,x) \geq 0.$$

Denote $\mathcal{F}(V)$ and $\mathcal{M}(V)$ the sets of real functions and probability measures on V.

$f \in \mathcal{F}(V)$ and $\mu \in \mathcal{P}(V)$

are seen respectively as

column and row vectors

$(f(x))_{x \in V}$ and $(\mu(x))_{x \in V}$.

Consequently L acts on

$\mathcal{F}(V)$ and $\mathcal{P}(V)$ respectively on

the right and on the left.

E.g. $\forall f \in \mathcal{F}(V)$, $\forall x \in V$,

$$L[f](x) = \sum_{y \in V} L(x,y) f(y)$$

Martingale problem:

A càdlàg process $X := (X(t))_{t \geq 0}$

is a solution to the martingale

problem associated to L if

for any $f \in \mathcal{F}(V)$, $(M^f(t))_{t \geq 0}$

is a martingale, where

$\forall\ t \geq 0$,

$$M^f(t) = f(X(t)) - f(X(0)) - \int_0^t L[f](X(s))\,ds$$

The main goal of this lecture is to show:

<u>Theorem 1</u>:

A solution to the martingale problem associated to L is a strong Markov process

This may seem strange at first view, but martingale problems are the simplest way to introduce and to work with càdlàg Markov processes, and it is even truer when the state space is more complicated than a finite set.

The *initial law* of X is the distribution of $X(0)$.

If $\mu \in \mathcal{P}(V)$ is given, we say that X is a solution to the martingale problem associated to V and starting from μ, when the initial law of X is μ.

When $\mu = \delta_x$, the Dirac mass at $x \in V$, we also say that X starts from x.

In this situation we will write $X_x := (X_x(t))_{t \geq 0}$ instead of X.

The main step of the proof of Theorem 1 is

Proposition 2

Assume that X_x is a solution to the martingale problem associated to L and starting from x.

Define the first time jump as

$$\tau = \inf\{ t > 0 : X_x(t) \neq X_x(t-) \}$$

- If $L(x) = 0$, we have

  $\tau = +\infty$: the process stays at x.

- If $L(x) > 0$, then

$$X_x(\tau) \perp\!\!\!\perp \tau$$

$$\mathcal{L}(\tau) = \text{exponential of parameter } L(x)$$

$$\mathcal{L}(X_x(\tau)) = \frac{L(x, \cdot)}{L(x)}$$

proof

Fix $y \neq x$ and denote $f = \mathbb{1}_{\{y\}}$ the indicator function of y.

There exists a martingale $(M^f(t))_{t\geq 0}$ such that for any $t \geq 0$,

$$f(X(t)) = f(X(0)) + \int_0^t Lf(X(s))\,ds + M^f(t)$$

Let us replace t by $t\wedge\tau$:

$$f(X(t\wedge\tau)) = f(X(0)) + \int_0^{t\wedge\tau} Lf(X(s))\,ds + M^f(\tau\wedge t)$$

We have $f(X(0)) = \mathbb{1}_y(x) = 0$

and for $s < \tau$, we have

$X(s) = x$, so that

$$f(X(t\wedge\tau)) = Lf(x)(t\wedge\tau) + M^f(\tau\wedge t)$$

$$= L(x,y)(t\wedge\tau) + M^f(\tau\wedge t)$$

Note that by definition $M^f(0) = 0$ and that $(M^f(\tau\wedge t))_{t\geq 0}$ is a martingale. It follows that

$$\mathbb{E}[M^f(\tau\wedge t) \mid \mathcal{F}(0)] = M^f(\tau\wedge 0) = 0$$

and thus

$$\mathbb{E}[f(X(t\wedge\tau))] = L(x,y)\,\mathbb{E}[t\wedge\tau] + \underbrace{\mathbb{E}[M^f(\tau\wedge t)]}_{0}$$

$$= L(x,y)\,\mathbb{E}[t\wedge\tau]$$

$$\parallel \quad (2)$$

$$\mathbb{P}[X(t\wedge\tau)=y]$$

$$= \mathbb{P}[t\geq\tau,\ X(\tau)=y]$$

Summing (2) over $y \in V\setminus\{x\}$, we get

$$\mathbb{P}[t\geq\tau] = -L(x)\,\mathbb{E}[t\wedge\tau]$$

$$= -L(x)\int_0^t \mathbb{P}[s\leq\tau]\,ds \quad (3)$$

Indeed by Fubini,

$$\int_0^t \mathbb{E}[\mathbb{1}_{s\leq\tau}]\,ds$$

$$= \mathbb{E}\left[\int_0^t \mathbb{1}_{s\leq\tau}\,ds\right]$$

$$= \mathbb{E}[t\wedge\tau]$$

Define $F(t) = \mathbb{P}[\tau > t]$ for all $t \geq 0$, (3) writes

$$1 - F(t) = -L(x,x) \int_0^t F(s)\, ds$$

which is integrated into

$$F(t) = e^{-L(x)t} \qquad (4)$$

If $L(x) = 0$, we get

$$\mathbb{P}[\tau > t] = 1 \quad \forall\, t \geq 0$$

$$\Rightarrow \tau = +\infty \text{ a.s.}$$

If $L(x) > 0$, (4) shows that $\tau \sim \exp(L(x))$

Coming back to (2) and replacing $\mathbb{E}[t \wedge \tau] = L(x)(1 - \mathbb{P}[\tau > t])$

$$= L(x)\, \mathbb{P}[\tau \leq t]$$

as we have just see, we get

$$\mathbb{P}[X(\tau) = y,\ \tau \leq t] = \frac{L(x,y)}{L(x)} \mathbb{P}[\tau \leq t]$$

Letting $t \to +\infty$, we get

$$\mathbb{P}[X(\tau) = y] = \frac{L(x,y)}{L(x)}$$

and finally

$$\mathbb{P}[X(\tau) = y, \tau \leq t] = \mathbb{P}[X(\tau) = y]\, \mathbb{P}(\tau \leq t)$$

which shows $X(\tau) \perp\!\!\!\perp \tau$

$\square$

The same proof, replacing the expectation $\mathbb{E}$ by the conditional expectation $\mathbb{E}[\,\cdot\,| X(0)]$ enables us to get that if X is a solution to the martingale problem, then for any $t \geq 0$ and $y \in V$,

$$\mathbb{P}[\tau > t, X(\tau) = y \,|\, X(0)]$$

$$= e^{-L(X_0)t} \, \frac{L(X(0), y)}{L(X(0))} \, \mathbb{1}_{y \neq X(0)}$$

We can iterate this result

to give a probabilistic description of X. Write $\tau_1 = \tau$ as above and consider the second jump time:

$$\tau_2 = \inf\{t > \tau_1 : X(t) \neq X(t-)\}$$

(assuming $\tau_1 < +\infty$).

Then conditionally on $\mathcal{F}_{\tau_1}$, if $L(X(\tau_1)) > 0$, we have

$$\tau_2 - \tau_1 \sim \exp(L(X(\tau_1)))$$

$$X(\tau_2) \sim \frac{L(X(\tau_1), \cdot)}{L(X(\tau_1))}$$

$$\tau_2 - \tau_1 \perp\!\!\!\perp X(\tau_2)$$

According to the above arguments, to see these results, it is sufficient to see that $(\underbrace{X(\tau_1 + t)}_{=: (\tilde{X}(t))_{t \geq 0}})_{t \geq 0}$ is a solution to the martingale problem associated to L, under the conditional law $\mathbb{P}[\,\cdot \mid \mathcal{F}(\tau_1)]$, where the underlying filtration is now

$$\left(\tilde{\mathcal{F}}_{(r)}\right)_{r\geq 0} := \left(\mathcal{F}_{(\tau_1+r)}\right)_{r\geq 0}.$$

To check this, write $\forall r\geq 0$,

$$f(X(\tau_1+r)) = f(X(\tau_1)) + \int_{\tau_1}^{\tau_1+r} L[f](X(s))\,ds$$

$$+ \tilde{M}^f(r)$$

with

$$\tilde{M}^f(r) = M^f(\tau_1+r) - M^f(\tau_1)$$

And indeed the stopping theorem shows that $\forall\, r\geq s\geq 0$,

$$\mathbb{E}\left[\tilde{M}^f(r+s) - \tilde{M}^f(r) \,\middle|\, \tilde{\mathcal{F}}_{(r)}\right]$$

$$= \mathbb{E}\left[M^f(\tau_1+r+s) - M^f(\tau_1+r) \,\middle|\, \mathcal{F}_{(\tau_1+r)}\right]$$

$$= 0 \qquad (5)$$

Iterating this procedure to get the conditional distributions of all the couples (time, position) of the jumps, we see that

X, solution to the martingale problem associated to L and starting from μ,

has the same law (on the set of càdlàg trajectories) as the process $\hat{X}$ constructed as follows

- sample $\hat{X}(0) \sim \mu$
- sample $\hat{\tau}_1 \sim \exp(L(\hat{X}(0)))$
- take $\hat{X}(t) = \hat{X}(0)$ for $t \in (0, \tau_1)$
- sample $\hat{X}(\hat{\tau}_1) \sim \dfrac{L(\hat{X}(0), \cdot)}{L(\hat{X}(0))}$
- sample $\hat{\tau}_2 - \hat{\tau}_1 \sim \exp(L(\hat{X}(\hat{\tau}_1)))$
- take $\hat{X}(t) = X(\hat{\tau}_1)$ for $t \in (\tau_1, \tau_2)$
- sample $\hat{X}(\hat{\tau}_2) \sim \dfrac{L(\hat{X}(\hat{\tau}_1), \cdot)}{L(\hat{X}(\hat{\tau}_1))}$
- sample $\hat{\tau}_3 - \hat{\tau}_2 \sim \exp(L(\hat{X}(\hat{\tau}_2)))$

etc., at least as long as

we do not end up with a

$\hat{\tau}_n = +\infty$, in which case we take $\hat{X}(t) = \hat{X}(\hat{\tau}_{n-1})$ for $t \in (\hat{\tau}_{n-1}, \infty)$.

Conversely it can be checked that the above construction leads to a solution to the martingale problem associated to L and starting from μ.

Thus we have obtained:

Corollary 3

Let be given a Markov generator L and $\mu \in \mathcal{P}(V)$. There exists a solution to the martingale problem associated to L and starting from μ. Furthermore its law is unique.

The above uniqueness in law is the main ingredient to show

Theorem 1. Let X be a solution to the martingale problem associated to L. Consider τ a finite stopping time (if we take τ to be a constant time, it will show the Markov property of X).

As in (5) we show that conditionally on $\mathcal{F}(\tau)$, $(X(\tau+t))_{t\geq 0}$ is still a solution to the martingale problem associated to L.

Its law is thus determined by its initial distribution which is the Dirac mass at $X(\tau)$ (don't forget we work here conditionally on $\mathcal{F}(\tau)$).

This amounts to the wanted dependence on the past of the future only going through the present position $X(\tau)$.

It remains to check the time-homogeneity of X.

Fix $t \geq 0$ and take $\tau \equiv s$

The above arguments show that

$(X(s+t))_{t \geq 0}$ conditioned by

$\mathcal{F}(s)$ has the same law

as $(X_{X(s)}(t))_{t \geq 0}$

In particular, fixing $t \geq 0$

and $y \in V$,

$$\mathbb{P}[X(s+t) = y \mid \mathcal{F}(s)]$$

$$= \mathbb{P}[X_{X(s)}^{(t)} = y]$$

$$= P_t(X(s), y)$$

with

$$P_t(x, y) := \mathbb{P}[X_x(t) = y]$$