Chapter 12

Generators of Markov Processes

> This lecture is concerned with the infinitessimal generator of a Markov process, and the sense in which we are able to write the evolution operators of a homogeneous Markov process as exponentials of their generator.

Take our favorite continuous-time homogeneous Markov process, and consider its semi-group of time-evolution operators K_t. They obey the relationship $K_{t+s} = K_t K_s$. That is, composition of the operators corresponds to addition of their parameters, and vice versa. This is reminiscent of the exponential functions on the reals, where, for any $k \in \mathbb{R}$, $k^{(t+s)} = k^t k^s$. In the discrete-parameter case, in fact, $K_t = (K_1)^t$, where integer powers of operators are defined in the obvious way, through iterated composition, i.e., $K^2 f = K \circ (Kf)$. It would be nice if we could extend this analogy to continuous-parameter Markov processes. One approach which suggests itself is to notice that, for any k, there's another real number g such that $k^t = e^{tg}$, and that e^{tg} has a nice representation involving integer powers of g:

$$e^{tg} = \sum_{i=0}^{\infty} \frac{(tg)^i}{i!}$$

The strategy this suggests is to look for some other operator G such that

$$K_t = e^{tG} \equiv \sum_{i=0}^{\infty} \frac{t^i G^i}{i!}$$

Such an operator G is called the *generator* of the process, and the purpose of this chapter is to work out the conditions under which this analogy can be carried through.

In the exponential function case, we notice that g can be extracted by taking the derivative at zero: $\frac{d}{dt} e^{tg}\big|_{t=0} = g$. This suggests the following definition.

Definition 150 (Infinitessimal Generator) *Let K_t be a continuous-parameter semi-group of linear operators on L, where L is a normed linear vector space, e.g., L_p for some $1 \leq p \leq \infty$. Say that a function $f \in L$ belongs to* $\mathrm{Dom}(G)$ *if the limit*

$$\lim_{h \downarrow 0} \frac{K_h f - K_0 f}{h} \equiv Gf \tag{12.1}$$

exists in an L-norm sense (Definition 151). The operator G defined through Eq. 12.1 is called the infinitessimal generator of the semi-group K_t.

Definition 151 (Limit in the L-norm sense) *Let L be a normed vector space. We say that a sequence of elements $f_n \in L$ has a limit f in L when*

$$\lim_{n \to \infty} \|f_n - f\| = 0 \tag{12.2}$$

This definition extends in the natural way to continuously-indexed collections of elements.

Lemma 152 (Generators are Linear) *For every semi-group of homogeneous transition operators K_t, the generator G is a linear operator.*

PROOF: Exercise 12.1. □

Lemma 153 (Invariant Distributions of a Semi-group Belong to the Null Space of Its Generator) *If ν is an invariant distribution of a semi-group of Markov operators M_t with generator G, then $G\nu = 0$.*

PROOF: Since ν is invariant, $M_t \nu = \nu$ for all t (Theorem 136). Hence $M_t \nu - \nu = M_t \nu - M_0 \nu = 0$, so, applying the definition of the generator (Eq. 12.1), $G\nu = 0$. □

Remark: The converse assertion, that $G\nu = 0$ implies ν is invariant under M_t, requires extra conditions.

There is a conjugate version of this lemma.

Lemma 154 (Invariant Distributions and the Generator of the Time-Evolution Semigroup) *If ν is an invariant distribution of a Markov process, and the time-evolution semi-group K_t is generated by G, then, $\forall f \in$* $\mathrm{Dom}(G)$, *$\nu G f = 0$.*

PROOF: Since ν is invariant, $\nu K_t = \nu$ for all t, hence $\nu K_t f = \nu f$ for all $t \geq 0$ and all f. Since taking expectations with respect to a measure is a linear operator, $\nu(K_t f - f) = 0$, and obviously then (Eq. 12.1) $\nu G f = 0$. □

Remark: Once again, $\nu G f = 0$ for all f is not enough, in itself, to show that ν is an invariant measure.

You will usually see the definition of the generator written with f instead of $K_0 f$, but I chose this way of doing it to emphasize that G is, basically, the derivative at zero, that $G = dK/dt|_{t=0}$. Recall, from calculus, that the exponential function can k^t be defined by the fact that $\frac{d}{dt}k^t \propto k^t$ (and e can be defined as the k such that the constant of proportionality is 1). As part of our program, we will want to extend this differential point of view. The next lemma builds towards it, by showing that if $f \in \mathrm{Dom}(G)$, then $K_t f$ is too.

Lemma 155 (Operators in a Semi-group Commute with Its Generator) *If G is the generator of the semi-group K_t, and f is in the domain of G, then K_t and G commute, for all t:*

$$K_t G f = \lim_{t' \to t} \frac{K_{t'} f - K_t f}{t' - t} \tag{12.3}$$

$$= G K_t f \tag{12.4}$$

PROOF: Exercise 12.2. □

Definition 156 (Time Derivative in Function Space) *For every $t \in T$, let $u(t,x)$ be a function in L. When the limit*

$$u'(t_0, x) = \lim_{t \to t_0} \frac{u(t,x) - u(t_0,x)}{t - t_0} \tag{12.5}$$

exists in the L sense, then we say that $u'(t_0)$ is the time derivative *or* strong derivative *of $u(t)$ at t_0.*

Lemma 157 (Generators and Derivatives at Zero) *Let K_t be a homogeneous semi-group of operators with generator G. Let $u(t) = K_t f$ for some $f \in \mathrm{Dom}(G)$. Then $u(t)$ is differentiable at $t = 0$, and its derivative there is Gf.*

PROOF: Obvious from the definitions. □

Theorem 158 (The Derivative of a Function Evolved by a Semi-Group) *Let K_t be a homogeneous semi-group of operators with generator G, and let $u(t,x) = (K_t f)(x)$, for fixed $f \in \mathrm{Dom}(G)$. Then $u'(t)$ exists for all t, and is equal to $Gu(t)$.*

PROOF: Since $f \in \mathrm{Dom}(G)$, $K_t G f$ exists, but then, by Lemma 155, $K_t G f = G K_t f = Gu(t)$, so $u(t) \in \mathrm{Dom}(G)$ for all t. Now let's consider the time derivative of $u(t)$ at some arbitrary t_0, working from above:

$$\frac{(u(t) - u(t_0))}{t - t_0} = \frac{K_{t-t_0} u(t_0) - u(t_0)}{t - t_0} \tag{12.6}$$

$$= \frac{K_h u(t_0) - u(t_0)}{h} \tag{12.7}$$

Taking the limit as $h \downarrow 0$, we get that $u'(t_0) = Gu(t_0)$, which exists, because $u(t_0) \in \mathrm{Dom}(G)$. □

Corollary 159 (Initial Value Problems in Function Space) *$u(t) = K_t f$, $f \in \text{Dom}(G)$, solves the initial value problem $u(0) = f$, $u'(t) = Gu(t)$.*

PROOF: Immediate from the theorem. □

Remark: Such initial value problems are sometimes called *Cauchy problems*, especially when G takes the form of a differential operator.

Corollary 160 (Derivative of Conditional Expectations of a Markov Process) *Let X be a homogeneous Markov process whose time-evolution operators are K_t, with generator G. If $f\text{Dom}(G)$, then its condition expectation $\mathbf{E}\left[f(X_s)|X_t\right]$ has strong derivative $Gu(t)$.*

PROOF: An immediate application of the theorem.

We are now almost ready to state the sense in which K_t is the result of exponentiating G. This is given by the remarkable Hille-Yosida theorem, which in turn involves a family of operators related to the time-evolution operators, the "resolvents", again built by analogy to the exponential functions, and to Laplace transforms.

Recall that the *Laplace transform* of a function $f : \mathbb{R} \mapsto \mathbb{R}$ is another function, $\tilde{f}$, defined by

$$\tilde{f}(\lambda) \equiv \int_0^\infty e^{-\lambda t} f(t) dt$$

for positive λ. Laplace transforms arise in many contexts (linear systems theory, integral equations, etc.), one of which is *moment-generating functions* in basic probability theory. If Y is a real-valued random variable with probability law P, then the moment-generating function is

$$M_Y(\lambda) \equiv \mathbf{E}\left[e^{\lambda Y}\right] = \int e^{\lambda y} dP = \int e^{\lambda y} p(y) dy$$

when the density in the last expression exists. You may recall, from this context, that the distributions of well-behaved random variables are completely specified by their moment-generating functions; this is actually a special case of a more general result about when functions are uniquely described by their Laplace transforms, i.e., when f can be expressed uniquely in terms of $\tilde{f}$. This is important to us, because it turns out that the Laplace transform, so to speak, of a semi-group of operators is better-behaved than the semi-group itself, and we'll want to say when we can use the Laplace transform to recover the semi-group.

The analogy with exponential functions, again, is a key. Notice that, for any positive constant λ,

$$\int_{t=0}^\infty e^{-\lambda t} e^{tg} dt = \frac{1}{\lambda - g} \tag{12.8}$$

from which we could recover g, as that value of λ for which the Laplace transform is singular. In our analogy, we will want to take the Laplace transform of the semi-group. Just as $\tilde{f}(\lambda)$ is another real number, the Laplace transform of a semi-group of operators is going to be another operator. We will want that to be the inverse operator to $\lambda - G$.

Definition 161 (Resolvents) *Given a continuous-parameter homogeneous semi-group K_t, for each $\lambda > 0$, the* resolvent operator *or* resolvent R_λ *is the Laplace transform of K_t: for every $f \in L$,*

$$(R_\lambda f)(x) \equiv \int_{t=0}^{\infty} e^{-\lambda t}(K_t f)(x) dt \tag{12.9}$$

Remark 1: Think of K_t as a function from the real numbers to the linear operators on L. Symbolically, its Laplace transform would be $\int_0^\infty e^{-\lambda t} K_t dt$. The equation above just fills in the content of that symbolic expression.

Remark 2: The name "resolvent", like some of the other ideas an terminology of operator semi-groups, comes from the theory of integral equations; invariant densities (when they exist) are solutions of homogeneous linear Fredholm integral equations of the second kind. Rather than pursue this connection, or even explain what that phrase means, I will refer you to the classic treatment of integral equations by Courant and Hilbert (1953, ch. 3), which everyone else seems to follow *very closely*.

Remark 3: When the function f is a value (loss, benefit, utility, ...) function, $(K_t f)(x)$ is the expected value at time t when starting the process in state x. $(R_\lambda f)(x)$ can be thought of as the *net present expected value* when starting at x and applying a discount rate λ.

Definition 162 (Yosida Approximation of Operators) *The* Yosida approximation *to a semi-group K_t with generator G is given by*

$$K_t^{(\lambda)} \equiv e^{tG^{(\lambda)}} \tag{12.10}$$

$$G^{(\lambda)} \equiv \lambda(\lambda R_\lambda - I) = \lambda G R_\lambda \tag{12.11}$$

The domain of $G^{(\lambda)}$ is all of L, not just $\mathrm{Dom}(G)$.

Theorem 163 (Hille-Yosida Theorem) *Let G be a linear operator on some linear subspace $\mathcal{D}$ of L. G is the generator of a continuous semi-group of contractions K_t if and only if*

1. *$\mathcal{D}$ is dense in L;*
2. *For every $f \in L$ and $\lambda > 0$, there exists a unique $g \in \mathcal{D}$ such that $\lambda g - Gg = f$;*
3. *For every $g \in \mathcal{D}$ and positive λ, $\|\lambda g - Gg\| \geq \lambda \|g\|$.*

Under these conditions, the resolvents of K_t are given by $R_\lambda = (\lambda I - G)^{-1}$, and K_t is the limit of the Yosida approximations as $\lambda \to \infty$:

$$K_t f = \lim_{\lambda \to \infty} K_t^\lambda f, \ \forall f \in L \tag{12.12}$$

PROOF: See Kallenberg, Theorem 19.11. □

Remark 1: Other good sources on the Hille-Yosida theorem include Ethier and Kurtz (1986, sec. 1.2), and of course Hille's own book (Hille, 1948, ch. XII). I have not read Yosida's original work.

Remark 2: The point of condition (1) in the theorem is that for any $f \in L$, we can chose $f_n \in \mathcal{D}$ such that $f_n \to f$. Then, even if Gf is not directly defined, Gf_n exists for each n. Because G is linear and therefore continuous, Gf_n goes to a limit, which we can chose to write as Gf. Similarly for G^2, G^3, etc. Thus we can *define* $e^{tG}f$ as

$$\lim_{n\to\infty} \sum_{j=0}^{\infty} \frac{t^j G^j f_n}{j!}$$

assuming all of the sums inside the limit converge, which is yet to be shown.

Remark 3: The point of condition (2) in the theorem is that, when it holds, $(\lambda I - G)^{-1}$ is well-defined, i.e., there *is* an inverse to the operator $\lambda I - G$. This is, recall, what we would like the resolvent to be, if the analogy with exponential functions is to hold good.

Remark 4: If we start from the semi-group K_t and obtain its generator G, the theorem tells us that it satisfies properties (1)–(3). If we start from an operator G and see that it satsifies (1)–(3), the theorem tells us that it generates some semi-group. It might seem, however, that it doesn't tell us how to construct that semi-group, since the Yosida approximation involves the resolvent R_λ, which is defined in terms of the semi-group, creating an air of circularity. In fact, when we start from G, we *decree* R_λ to be $(\lambda I - G)^{-1}$. The Yosida approximations then is defined in terms of G and λ alone.

Corollary 164 (Stochastic Approximation of Initial Value Problems) *Let $u(0) = f$, $u'(t) = Gu(t)$ be an initial value problem in L. Then a stochastic approximation to $u(t)$ can be founded by taking*

$$\hat{u}(t) = \frac{1}{n}\sum_{i=1}^{n} f(X_i(t)) \tag{12.13}$$

where the X_i are independent copies of the Markov process corresponding to the semi-group K_t generated by G, with initial condition $X_i(0) = x$.

PROOF: Combine corollaries. □

12.1 Exercises

Exercise 12.1 (Generators are Linear) *Prove Lemma 152.*

Exercise 12.2 (Semi-Groups Commute with Their Generators) *Prove Lemma 155.*

1. *Prove Equation 12.3, restricted to* $t' \downarrow t$ *instead of* $t' \to t$. Hint*: Write* T_t *in terms of an integral over the corresponding transition kernel, and find a reason to exchange integration and limits.*

2. *Show that the limit as* $t' \uparrow t$ *also exists, and is equal to the limit from above.* Hint*: Re-write the quotient inside the limit so it only involves positive time-differences.*

3. *Prove Equation 12.4.*

Exercise 12.3 (Generator of the Poisson Counting Process) *Find the generator of the time-evolution operators of the Poisson counting process (Example 140).*