Notes

March 20^{th}, 2012

1 The infinitesimal generator of a Markov Chain

In these notes we introduce informally the infinitesimal generator of a homogeneous continuous-time Markov process and we show its expression for the case of Markov chains.

Let us consider the following family of operators, $\{T_t\}_{t\geq 0}$, indexed by the variable $t \geq 0$ and acting on the bounded functions $f : E \to \mathbb{R}$. Since we deal with Markov chains, we will assume implicitly that the state space is numerable.

$$T_t f(i) = \mathbb{E}[f(X(t))|X(0) = i] = \mathbb{E}_i[f(X(t))] \ .$$

By time-homogeneity we have also that $T_t f(i) = \mathbb{E}[f(X(t+s))|X(s) = i]$, for any $s \geq 0$. The family $\{T_t\}_t$, satisfies the following relation

$$\begin{aligned} T_t(T_s f)(i) &= \mathbb{E}[T_s f(X(t))|X(0) = i] = \mathbb{E}[\mathbb{E}[f(X(s+t))|X(t),\, X(0) = i]|X(0) = i] \\ &= \mathbb{E}[f(X(s+t))|X(0) = i] = T_{t+s} f(i) \ , \end{aligned}$$

that is $T_t\, T_s = T_{s+t}$, that expresses the fact that this family of operators forms a semigroup under composition.

If we denote by $||f|| = \sup_{i\in E} |f(i)|$, we have that given a function f such that $||f|| \leq 1$, also $||T_t f|| \leq 1$ that generally is expressed by the fact that the family $\{T_t\}_t$ is a contraction semigroup.

In addition it is also strongly continuous as we have that

$$\lim_{t\downarrow 0} T_t f = f \ .$$

All the limits have to be interpreted in the sup norm, this converges naturally implies also the weaker pointwise convergence. For example the limit above means that

$$||T_t f - f|| \overset{t\downarrow 0}{\to} 0 \ .$$

The infinitesimal generator, $\mathcal{A}$, is informally the derivative at time 0 of the continuous semigroup, that is, given a function f we define its value at f as

$$\mathcal{A} f = \lim_{t\downarrow 0} \frac{T_t f - f}{t}$$

whenever such limit exists, and we define the domain of the generator as the set of functions, $\mathcal{D}(\mathcal{A})$, for which the limit above exists.

The limit again is in the sup norm, that is, if we take a function $f \in \mathcal{D}(\mathcal{A})$ we have that

$$||\frac{T_t f - f}{t} - \mathcal{A} f|| \overset{t\downarrow 0}{\to} 0 \ .$$

Once we know that $f \in \mathcal{D}(\mathcal{A})$ then we also have the pointwise convergence

$$\mathcal{A} f(i) = \lim_{t\downarrow 0} \frac{T_t f(i) - f(i)}{t}, \quad \forall i \in E \ .$$

1.1 Backwards and Forward Kolmogorov equations

In this section we informally compute the derivative of the semigroup $\{T_t\}_{t\geq 0}$ at a time $t > 0$, we than have

$$\begin{matrix} \text{KBE}: \\ \\ \text{KFE}: \end{matrix} \quad T_t' = \lim_{s\downarrow 0} \frac{T_{t+s} f - T_t f}{s} = \begin{matrix} \lim_{s\downarrow 0} \frac{T_s(T_t f) - T_0(T_t f)}{s} \\ \\ \lim_{s\downarrow 0} \frac{T_t(T_s f) - T_t(T_0 f)}{s} \end{matrix} = \begin{matrix} \lim_{s\downarrow 0} \frac{(T_s - T_0)}{s} T_t f = \mathcal{A}\, T_t f \\ \\ \lim_{s\downarrow 0} T_t \frac{T_s f - T_0 f}{s} = T_t\, \mathcal{A} f \end{matrix}$$

where $T_0 = I$, with I being the identity operator. Therefore we can write $\mathcal{A} = T_0'$ and in general using the KFE

$$T_t' = T_t\, \mathcal{A}$$

with boundary conditions $T_0 = I$. It follows that at least formally one expects that

$$T_t f = e^{t\mathcal{A}} f$$

where the exponential of the operator $\mathcal{A}$ is formally defined as

$$e^{t\mathcal{A}} = \sum_{n=0}^{\infty} \frac{t^n}{n!} \mathcal{A}^n \ .$$

1.2 Computing the infinitesimal generator for the HCMC

In this section we compute the infinitesimal generator for the case of a homogeneous continuous time Markov Chain with transition probability $\boldsymbol{P} = (p_{ij})$ and mean sojourn times $\mu_i = \rho_i^{-1}$. If we define the diagonal rate matrix $\boldsymbol{R} = \text{diag}(\rho_i)$ we have that the rate transition matrix, $\boldsymbol{Q} = (q_{ij}) = \boldsymbol{R}(\boldsymbol{P} - \mathbb{I})$, where $\mathbb{I}$ is the identity matrix.

Given the function f, we compute the value of the generator applied to this function evaluated at point i in the following way

$$\begin{aligned} \mathcal{A}f(i) &= \lim_{t\downarrow 0} \frac{\mathbb{E}_i[f(X(t))] - f(i)}{t} \\ &= \lim_{t\downarrow 0} \frac{\mathbb{E}_i[f(X(t)); N(t) = 0] - f(i)}{t} + \lim_{t\downarrow 0} \frac{\mathbb{E}_i[f(X(t)); N(t) = 1]}{t} + \lim_{t\downarrow 0} \frac{\mathbb{E}_i[f(X(t)); N(t) \geq 2]}{t} \ , \qquad (1) \end{aligned}$$

where we denoted by $N(t)$ the number of transitions in the interval $[0, t]$ and used the notation $\mathbb{E}[f(X); A] = \mathbb{E}[f(X)\, 1\{A\}]$, for a given event A.

For the first term we have that

$$\begin{aligned} \lim_{t\downarrow 0} \frac{\mathbb{E}_i[f(X(t)); N(t) = 0] - f(i)}{t} &= \lim_{t\downarrow 0} \frac{\mathbb{E}_i[f(i); N(t) = 0] - f(i)}{t} = \lim_{t\downarrow 0} f(i) \frac{\mathbb{P}_i\{N(t) = 0\} - 1}{t} \\ &= \lim_{t\downarrow 0} f(i) \frac{e^{-\rho_i t} - 1}{t} = -f(i)\, \rho_i = f(i)\, q_{ii} \end{aligned}$$

For the second term we have that

$$\lim_{t\downarrow 0} \frac{\mathbb{E}_i[f(X(t)); N(t) = 1]}{t} = \lim_{t\downarrow 0} \sum_j p_{ij} f(j) \frac{\mathbb{P}_i\{N(t) = 1, X(T_1) = j\}}{t} = \sum_j f(j)\, \rho_i\, p_{ij} = \sum_{j\neq i} f(j)\, q_{ij}$$

In the second equality we used the fact that

$$\begin{aligned} \mathbb{P}_i\{N(t) = 1, X(T_1) = j\} &= \mathbb{P}_i\{T_1 < t, T_2 > t, X(T_1) = j\} = \mathbb{P}_i\{Y_1 < t, Y_1 + Y_2 > t, X(T_1) = j\} \\ &= \rho_i e^{-t\rho_i} \frac{1 - e^{-t(\rho_j - \rho_i)}}{\rho_j - \rho_i} = t\rho_i e^{-t\rho_i} + o(t) \ , \end{aligned}$$

where T_1 and T_2 are the first two transition epochs and $Y_1 = T_1$, $Y_2 = T_2 - T_1$ are the first two inter-transition times. $Y_1|X(0) = i \sim \text{Exp}(\rho_i)$ and $Y_2|X(T_1) = j \sim \text{Exp}(\rho_j)$ are independent exponential random variables.

For the last term we have that

$$\lim_{t\downarrow 0} \frac{\mathbb{E}_i[f(X(t)); N(t) \geq 2]}{t} = \lim_{t\downarrow 0} \sum_j p_{ij} f(j) \frac{\mathbb{P}_i\{N(t) \geq 2, X(T_1) = j\}}{t} = 0$$

In the last equality we used the fact that

$$\begin{aligned} \mathbb{P}_i\{N(t) \geq 2, X(T_1) = j\} &= \mathbb{P}_i\{T_1 < t, T_2 < t, X(T_1) = j\} = \mathbb{P}_i\{Y_1 < t, Y_1 + Y_2 < t, X(T_1) = j\} \\ &= \frac{\rho_i}{\rho_i - \rho_j}(1 - e^{-t\rho_j}) - \frac{\rho_j}{\rho_i - \rho_j}(1 - e^{-t\rho_i}) = o(t) \ . \end{aligned}$$

Substituting the expressions above in (1) we finally get

$$\mathcal{A}f(i) = \sum_j q_{ij}\, f(j) = (\boldsymbol{Q}\ f)_i$$

where in the last equation we identified the function f with the column vector $(f(i))_{i\in E}$.

1.3 Computing the transition probabilities

In this section we show how to compute the transition probabilities $p_{ij}(t) = \mathbb{P}_i\{X(t) = j\}$ using the infinitesimal generator.

Choose the Kronecker function $f(i) = \delta_{ij}$. We have that

$$T_t f(i) = \mathbb{E}_i[f(X(t))] = \mathbb{E}_i[1\{X(t) = j\}] = \mathbb{P}_i\{X(t) = j\} = p_{ij}(t)$$

Using the KFE we have that

$$p'_{ij}(t) = T'_t f(i) = T_t\, A\, f(i) = \sum_{k,h} p_{ik}(t)\, q_{kh}\, \delta_{hj} = \sum_k p_{ik}(t)\, q_{kj}$$

and in matrix form, by denoting $\boldsymbol{P}(t) = (p_{ij}(t))$ the t-transition matrix, we have

$$\boldsymbol{P}'(t) = \boldsymbol{P}(t)\, \boldsymbol{Q}$$

that with boundary condition $\boldsymbol{P}(0) = \mathbb{I}$ gives the solution

$$\boldsymbol{P}(t) = e^{t\boldsymbol{Q}} \ ,$$

where $e^{t\boldsymbol{Q}} = \sum_{n\geq 0} t^n\, \boldsymbol{Q}^n / n!$.