2
Processos de Markov

Para preparar o leitor, discutiremos aqui a construção de cadeias e processos de Markov em espaços de estados discretos, primeiro a tempo discreto e depois a tempo contínuo. Assim, poderemos finalmente fazer uma definição geral de processo de Markov como uma família de probabilidades.

2.1
Cadeias de Markov a tempo discreto

Quando falamos de um processo estocástico no sentido mais geral, temos uma família de variáveis aleatórias (v.a.'s) $\{X_j\}_{j\in J}$ (onde J é um conjunto de índices), definidas num espaço de probabilidade $(\Omega, \mathcal{F}, P)$. Se o processo evolui no tempo, dizemos que o processo é discreto no caso em que $J \subseteq \mathbb{N}$ ou contínuo se $J \subseteq [0, +\infty)$, embora, podem ser considerados conjuntos mais gerais. Agora, seja S um conjunto finito ou enumerável, que chamamos espaço de estados do processo e que definimos como os possíveis valores que podem tomar as v.a.'s $\{X_j\}_{j\in J}$. Também, definimos uma matriz estocástica $(p(x,y))_{(x,y)\in S^2}$ como uma matriz de números não negativos, tais que,

$$\sum_{y\in S} p(x,y) = 1, \quad \forall x \in S.$$

Claramente, se S é um conjunto finito então a matriz é finita. Suponhamos que $\{X_n\}_{n\geq 0}$ são v.a.'s definidas no espaço de probabilidade $(\Omega, \mathcal{F}, P)$. Então, $\{X_n\}_{n\geq 0}$ é uma cadeia de Markov com probabilidade de transição $p_n(x,y)$ se, para cada $n \geq 0$ e qualquer escolha de $x_0, x_1, \cdots, x_{n-1}, x, y \in S$ temos

$$P\left(\omega \in \Omega : X_{n+1}(\omega) = y | X_n(\omega) = x, \cdots, X_0(\omega) = x_0\right) = p_n(x,y). \tag{2.1}$$

A condição (2.1) diz que, dado o atual estado x, a evolução futura é totalmente independente da evolução passada, isto é, não depende da escolha de $x_0, x_1, \cdots, x_{n-1}$, esta condição é chamada de propriedade de Markov. Além disso, se o lado direito de (2.1) não depende do tempo n, então o processo diz-se homogêneo no tempo.
Assim, dada uma cadeia de Markov $\{X_n\}_{n\geq 0}$ homogênea no tempo e tomando o conjunto $[X_0 = x_0] := \{\omega \in \Omega : X_0(\omega) = x_0\}$[1], chamamos distribuição inicial do processo, a medida denotada por μ, definida por

$$\mu(x_0) := P\left(X_0 = x_0\right), \quad \forall x_0 \in S.$$

[1]Daqui em diante, às vezes, vamos usar esta definição sem indicação para o leitor.

Pela equação (2.1), e dada uma cadeia de Markov $\{X_n\}_{n\geq 0}$ com matriz de transição $(p(x,y))_{(x,y)\in S^2}$ e distribuição inicial μ temos que

$$\begin{aligned}
&P\left(X_0 = x_0, \cdots, X_n = x_n\right) \\
&= P\left(X_0 = x_0, \cdots, X_{n-1} = x_{n-1}\right) P\left(X_n = x_n | X_0 = x_0, \cdots, X_{n-1} = x_{n-1}\right),
\end{aligned}$$

pela propriedade de Markov temos que a equação acima é igual a

$$\begin{aligned}
&P\left(X_0 = x_0, \cdots, X_{n-1} = x_{n-1}\right) \underbrace{P\left(X_n = x_n | \; X_{n-1} = x_{n-1}\right)}_{p(x_{n-1}, x_n)} \\
&\vdots \\
&= P\left(X_0 = x_0\right) p(x_0, x_1) p(x_1, x_2) \cdots p(x_{n-1}, x_n) \\
&= \mu(x_0) p(x_0, x_1) p(x_1, x_2) \cdots p(x_{n-1}, x_n).
\end{aligned} \tag{2.2}$$

Uma pergunta natural, mas não trivial de responder é: "Será que dada uma matriz estocástica p e uma probabilidade $\mu \in S$, existe uma cadeia de Markov com matriz de transição p e distribuição μ?"

Sem perda de generalidade, seja μ a medida de Dirac no ponto x_0. Suponhamos que queremos construir um número finito m de v.a.'s $(X_0, X_1, \cdots, X_m)$ com a propriedade de Markov para $0 \leq n \leq m-1$ e com distribuição inicial $\mu = \delta_{x_0}$. Seja $\Omega = S^{m+1}$ o espaço vetorial de dimensão $m+1$ com valores em S, e seja $\mathcal{F}$ o conjunto de todos os subconjuntos de Ω, isto é $\mathcal{F} = \mathfrak{P}(\Omega)$. Para cada $\omega = (s_0, \cdots, s_m) \in \Omega$ definimos

$$P^{x_0}(w) := \delta_{x_0}(s_0) p(s_0, s_1) p(s_1, s_2) \cdots p(s_{m-1}, s_m). \tag{2.3}$$

Assim, podemos definir as v.a.'s por $X_j(\omega) := s_j$ com $0 \leq j \leq m$. Então conseguimos definir P^{x_0} sobre o espaço S^{m+1} tal que as v.a.'s X_j (projeções) têm a propriedade markoviana (que resulta da construção) com matriz de transição p e estado inicial x_0.

Agora, pretendemos estender a construção ao processo infinito $\{X_n\}_{n\geq 0}$ no espaço $\Omega = S^\infty$, cujos elementos são sequências infinitas $\omega = (s_0, s_1, s_2, \cdots)$ com $s_j \in S$, tomando $X_n(\omega) = s_n$, isto é, considerando

$$\begin{aligned}
X_n : \Omega &\longrightarrow S \\
\omega = (s_0, s_1 \cdots) &\longmapsto s_n.
\end{aligned}$$

A σ-álgebra produto $\mathcal{F}$ em Ω é gerada pelos cilindros (eventos com um número finito de coordenadas). Com um espaço de estado S enumerável é suficiente considerar conjuntos da forma

$$\{\omega = (s_0, s_1, s_2, \cdots) : X_0(\omega) = s_0, X_1(\omega) = s_1, \cdots, X_m(\omega) = s_m\}.$$

Seja $\mathcal{C}_0$ a classe de tais conjuntos obtidos por deixar os vetores $(s_0, \cdots, s_m)$ variarem sobre S^{m+1}. Definimos a função P^x em $\mathcal{C}_0$ por

$$P^x (X_0 = s_0, X_1 = s_1, \cdots, X_m = s_m) = \delta_x(s_0)p(s_0, s_1) \cdots p(s_{m-1}, s_m).$$

Este é um primeiro passo para dar solução à nossa construção, o segundo passo é fornecido pelo Teorema de extensão de Kolgomorov, que garante que para cada x, a medida de probabilidade P^x, existe no espaço produto infinito, de tal forma que a probabilidade de cada cilindro é igual à probabilidade dada por (2.3). Se queremos começar com um estado inicial X_0 aleatório com distribuição μ, tomamos em Ω a medida P^μ definida em $A \in \mathcal{F}$ por

$$P^\mu(A) = \sum_{x \in S} \mu(x) P^x(A).$$

Portanto a resposta à pergunta é afirmativa, e dizemos que uma cadeia de Markov está bem determinada pela sua matriz de transição e pela sua distribuição inicial.

2.2
Cadeias de Markov a tempo contínuo

Vamos construir uma cadeia de Markov $\{X_t\}_{t \geq 0}$ com $0 \leq t < \infty$, mas num espaço de estados enumerável S. Como S é enumerável nossa cadeia tem que se mover em saltos, assim a evolução da cadeia é descrita da seguinte forma. No estado x esperamos um tempo aleatório, e saltamos aleatoriamente a um novo estado y. Depois esperamos um outro tempo aleatório e saltamos aleatoriamente a um estado z, e assim por diante. Então, temos que determinar duas coisas:

1. A distribuição de probabilidade dos tempos aleatórios de espera em cada estado $x \in S$.
2. O mecanismo para escolher o próximo estado quando ocorre um salto.

Notemos que a propriedade de Markov nos diz que a distribuição do tempo até ao salto seguinte, só pode depender da localização atual x e não pode depender do tempo passado.

Teorema 2.2.1. *A única distribuição contínua com a propriedade de não ter memória é a distribuição exponencial.*

Demonstração. Afirmamos que a v.a. X com distribuição exponencial de parâmetro $\lambda > 0$ tem perda de memória. De fato, note que, $\forall s, t > 0$

$$\begin{aligned} P(X > s + t | X > t) &= \frac{P(X > s + t, X > t)}{P(X > t)} \\ &= \frac{P(X > s + t)}{P(X > t)} \\ &= e^{-\lambda(s+t)} e^{\lambda t} \\ &= e^{-\lambda s} \\ &= P(X > s). \end{aligned}$$

A igualdade anterior significa que a probabilidade de X ser maior a $t + s$ sabendo que X é maior que t, só depende amplitude do intervalo, ou seja, essa probabilidade coincide com a probabilidade de X ser maior que s. Agora vejamos que é única. Seja X v.a. contínua com perda de memória e seja $F(t) = P(X > t)$. Então, $F(t + s) = F(t)F(s)$, $\forall t, s > 0$ e isto implica que $F(2) = F^2(1)$ e $F\left(\frac{1}{2}\right) = F^{\frac{1}{2}}(1)$. Em geral, se $a \in \mathbb{Q}$ temos que $F(a) = F^a(1)$. Assim a única função contínua que satisfaz a relação acima para $a \geq 0$ com $a \in \mathbb{Q}$, é $F(a) = F^a(1) = e^{a \ln F(1)}$. Agora basta tomar $\lambda = -\ln(F(1))$. □

Voltando, a perda de memória força que o tempo de espera em x seja distribuído exponencialmente. Seja $(c(x))^{-1}$ a sua média. Então, $c(x)$ é a taxa de salto do estado x. Sempre que a cadeia está em x consideramos que o tempo de espera T_x tem distribuição exponencial, ou seja, $\forall t > 0$ $P(T_x > t) = e^{-c(x)t}$. Quando a cadeia "salta", a propriedade de Markov diz que a escolha do seguinte estado depende só do atual estado x. Assim os saltos são descritos pela matriz estocástica $p(x, y)$, onde $p(x, y)$ é a probabilidade de estar em x e saltar para y. Como p não depende do tempo a probabilidade de saltar de x a y não depende do tempo em que o salto ocorre.

Então, isto sugere que para construir uma cadeia de Markov $\{X_t\}_{t\geq 0}$ a tempo contínuo com saltos exponencialmente distribuídos com parâmetros $c(x)$ e com $p(x, y)$ como matriz de transição, tomamos uma cadeia de Markov em tempo discreto $\{Y_n\}_{n\geq 0}$ com matriz de transição $p(x, y)$ e ajustamos os tempos de espera para produzir os tempos corretos, isto é, exponencialmente distribuídos com média $(c(x))^{-1}$.

Seja $x \in S$ o estado inicial e seja $(\Omega_1, \mathcal{H}_1, P_1^x)$ o espaço de probabilidade no qual está definida a cadeia discreta $\{Y_n\}_{n\geq 0}$, com matriz de transição $p(x, y)$ e estado inicial x. Além disso, independentemente de $\{Y_n\}_{n\geq 0}$, seja $\{\tau_j\}_{j\geq 0}$ uma sequência de v.a.'s independentes e identicamente distribuídas (i.i.d), com distribuição exponencial de média 1, no espaço $(\Omega_2, \mathcal{H}_2, P_2)$. Tomamos

$(\Omega, \mathcal{H}, \mathbf{P}^x) = (\Omega_1 \otimes \Omega_2, \mathcal{H}_1 \otimes \mathcal{H}_2, P_1^x \otimes P_2)$. Note que os estados que a cadeia toma em tempo contínuo $\{X_t\}_{t\geq 0}$, são $x = Y_0, Y_1, \cdots$, isto é a cadeia toma os mesmos valores de $\{Y_n\}_{n\geq 0}$. Definimos os tempos de espera por $\sigma_n = (c(Y_n))^{-1}\tau_n$. Claramente, $\{\sigma_n\}_{n\geq 0}$ não é independente da cadeia $\{Y_n\}_{n\geq 0}$, mas condicionada à variável discreta Y_n, temos que a variável σ_n é independente de $\{\sigma_k, Y_k\}_{0\leq k<n}$, e além disso, tem distribuição exponencial com média $(c(Y_n))^{-1}$.
Agora, defina $T_0 = 0, T_n = \sigma_0 + \sigma_1 + \cdots + \sigma_n$, para $n \geq 1$. Então

$$X_t = \sum_{n=0}^{\infty} Y_n \mathbf{1}_{\{[T_n, T_{n+1})\}}(t).$$

Note que X_t está definida para todo tempo $0 \leq t < \infty$, se $T_n \nearrow \infty$ quando $n \nearrow \infty$. De fato, isto acontece quase certamente no caso em que exista uma constante $C_0 < \infty$, tal que $c(x) \leq C_0$, $\forall x \in S$, pois, teríamos que $(c(Y_n))^{-1}$ não se anula quando $n \nearrow \infty$. Portanto, daqui para a frente vamos assumir este fato. A nossa construção pode ser repetida para cada estado inicial $x \in S$. Defina a probabilidade de transição por $p_t(x, y) = \mathbf{P}^x(X_t = y)$. Assim, temos a seguinte propriedade para todos os tempos $0 \leq t_0 < t_1 < t_2 < \cdots < t_n$ e estados iniciais $x_0, x_1, x_2, \cdots, x_n$:

$$\mathbf{P}^x\left(X_{t_0} = x_0, \cdots, X_{t_n} = x_n\right) = p_{t_0}(x, x_0) p_{t_1 - t_0}(x_0, x_1) \cdots p_{t_n - t_{n-1}}(x_{n-1}, x_n). \tag{2.4}$$

A igualdade anterior, implica que $\{X_t\}_{t\geq 0}$ tem a propriedade markoviana,

$$\mathbf{P}^x\left(X_{t_n} = x_n | X_{t_{n-1}} = x_{n-1}, \cdots, X_{t_0} = x_0\right) = p_{t_n - t_{n-1}}(x_{n-1}, x_n).$$

Seja D_S o espaço de funções $\xi : [0, +\infty) \to S$ tal que $t \mapsto \xi(t)$, com a propriedade que $\xi(t) = \lim_{s\downarrow t} \xi(s)$ (contínua à direita) e $\xi(t^-) = \lim_{s\uparrow t} \xi(s)$ existe (o limite à esquerda existe). Então D_S é o espaço de funções definidas em $[0, +\infty)$ contínuas à direita e com limite à esquerda. Daqui em diante, conjuntos de funções com esta propriedade vamos chamar càdlàg (do francês "continue à droite, limite à gauche").
Seja $\mathcal{F}$ a σ-álgebra em D_S gerada pela projeção $\pi_t : D_S \to S$ tal que para todo $t \geq 0$, $\pi_t(\xi) = \xi(t)$. Assim, podemos pensar em $X_. = \{X_t\}_{0\leq t<\infty}$ como uma v.a. tal que

$$\begin{array}{rcl} X_. : (\Omega, \mathcal{H}, \mathbf{P}^x) & \longrightarrow & (D_S, \mathcal{F}, P^x) \\ \omega & \longmapsto & X_.(\omega), \end{array}$$

com $P^x(A) = \mathbf{P}^x(X_. \in A)$, para todo $A \in \mathcal{F}$. Isto define uma família de probabilidades $\{P^x\}_{x\in S}$ em D_S. Então, para uma v.a. Z e para $x \in S$, podemos definir a esperança $E^x[Z] = \int_{D_S} Z dP^x$. A probabilidade de transição pode se

expressar por $p_t(x, y) = \mathbf{P}^x\left(X_t = y\right) = P^x\left(\xi(t) = y\right)$.
Queremos expressar a propriedade de Markov numa forma mais poderosa. Para tal, seja $\{\theta_t\}_{t\geq 0}$ a aplicação de translações no espaço D_S, definida por $\theta_t\xi(s) = \xi(s+t)$. O efeito da aplicação de translação θ_t é reiniciar o processo no tempo t. Para um evento $A \in \mathcal{F}$ temos que a imagem inversa

$$\theta_t^{-1}(A) = \{\xi \in D_S : \theta_t(\xi) \in A\}$$

é o evento onde "A acontece do tempo t em diante".
Seja $\mathcal{F}_t = \sigma\left(\xi(s) : 0 \leq s \leq t\right)$ a σ-álgebra natural em D_S. Então, para todos os eventos $A \in \mathcal{F}$ e $x \in S$ temos que

$$P^x\left(\theta_t^{-1}(A)|\mathcal{F}_t\right)(\xi) = P^{\xi(t)}(A), \tag{2.5}$$

para P^x-quase todo ξ. O lado esquerdo da equação (2.5) é a probabilidade condicional de um evento que ocorre do tempo t em diante, condicionada ao passado até ao tempo t. Para deduzir (2.5) de (2.4), basta verificar que

$$E^x[\mathbf{1}_B \cdot \mathbf{1}_A \circ \theta_t] = E^x[\mathbf{1}_B \cdot P^{\xi(t)}(A)],$$

primeiro para cilindros $A \in \mathcal{F}$ e $B \in \mathcal{F}_t$, e depois estender a todos os eventos pelo teorema $\pi - \lambda$ (ver Apêndice 1 de [9]). Acima E^x é a esperança com respeito a P^x.

2.3 Processo de Markov geral

Seja $\mathcal{S}$ um espaço métrico e considere em S a σ-álgebra de Borel que denotamos por $\mathcal{B}_{\mathcal{S}}$. Considere para cada $t \in [0, \infty)$ considere as projeções $\pi_t(\omega) = \omega(t)$. Seja $\mathcal{F}$ a menor σ-álgebra em $D_{\mathcal{S}}$ na qual todas as projeções π_t são mensuráveis. Para $t \in [0, \infty)$, seja $\mathcal{F}_t$ a menor σ-álgebra em $D_{\mathcal{S}}$ na qual todas as projeções π_s são mensuráveis, para $s \leq t$.

Definição 2.3.1. *Um processo de Markov é uma coleção $\{P^x\}_{x\in\mathcal{S}}$ de probabilidades em $D_{\mathcal{S}}$ tal que*

1. $P^x\left(\omega \in D_{\mathcal{S}} : \omega(0) = x\right) = 1$.
2. *$\forall A \in \mathcal{F}$ a função $x \mapsto P^x(A)$ é mensurável em $\mathcal{S}$.*
3. *$P^x(\theta_t^{-1}(A)|\mathcal{F}_t)(\omega) = P^{\omega(t)}(A)$ para P^x-quase certamente ω, para todo $x \in \mathcal{S}$ e para todo $A \in \mathcal{F}$.*

Na definição anterior, *1.* diz que x é o valor inicial para a cadeia partindo da medida P^x, e o ponto *3.* é a propriedade de Markov.

Para começar o processo com uma distribuição μ diferente de δ_x, definimos P^μ em $D_\mathcal{S}$ por

$$P^\mu(A) = \int_\mathcal{S} P^x(A)\mu(dx),\ \forall A \in \mathcal{F}.$$

A probabilidade de transição $p_t(x, dy)$ é definida para todo $t \geq 0$, $x \in \mathcal{S}$ e $B \in \mathcal{B}_\mathcal{S}$ por

$$p_t(x, B) = P^x\left(\pi_t \in B\right). \tag{2.6}$$

As equações de Chapman-Kolgomorov

$$p_{t+s}(x, B) = \int_\mathcal{S} p_s(y, B) p_t(x, dy), \tag{2.7}$$

são consequência da propriedade de Markov.

Agora, para funções $\mathcal{F}$- mensuráveis e limitadas em $\mathcal{S}$ e $t \geq 0$ definimos a função $S_t f$ em $\mathcal{S}$ por

$$S_t f(x) = E^x[f(\pi_t)] = \int_\mathcal{S} f(x) p_t(x, dy). \tag{2.8}$$

A mensurabilidade de $S_t f$ é consequência do ponto 2. da Definição 2.3.1. Denotamos a norma do supremo para funções mesuráveis e limitadas por

$$\|f\|_\infty := \sup_{x \in \mathcal{S}} |f(x)|, \tag{2.9}$$

e assim, é fácil ver que:

$$\|S_t f\|_\infty \leq \|f\|_\infty, \tag{2.10}$$

portanto S_t leva funções mensuráveis e limitadas em $\mathcal{S}$, em funções mensuráveis e limitadas em $\mathcal{S}$. Pela linearidade da integral também temos a linearidade de S_t. Portanto S_t é um operador linear definido no espaço das funções mensuráveis e limitadas. Logo, pela propriedade de Markov temos que

$$\begin{aligned} S_{s+t} f(x) &= E^x[f(\pi_{s+t})] \\ &= E^x[E^x[f(\pi_{s+t})|\mathcal{F}_s]] \\ &= E^x[E^{\pi_s}[f(\pi_t)]] \\ &= E^x[S_t f(\pi_s)] \\ &= S_s S_t f(x). \end{aligned}$$

Daqui a pouco, vamos ver que $\{S_t\}_{t\geq 0}$ forma um semigrupo. Além disso, pela propriedade (2.10) dizemos que $\{S_t\}_{t\geq 0}$ é um semigrupo de contração. Afirmamos, sem demostração, que $\{P^x\}_{x\in\mathcal{S}}$ são unicamente determinadas pelo semigrupo $\{S_t\}_{t\geq 0}$ (ver [11], página 9).

Definição 2.3.2. *Seja* $C_b(\mathcal{S}) = \{f : \mathcal{S} \to \mathbb{R},$ *contínua e limitada em* $\mathcal{S}\}$.

Vamos considerar em $C_b(\mathcal{S})$ a norma definida em (2.9).

Definição 2.3.3. *Um processo de Markov* $\{P^x\}_{x\in\mathcal{S}}$ *é um processo de Feller se,* $C_b(\mathcal{S})$ *é fechado sobre a ação do semigrupo, isto é, se* $f \in C_b(\mathcal{S})$ *então* $S_t f \in C_b(\mathcal{S})$, $\forall t \geq 0$.

2.4 Semigrupos e geradores

Esta seção não pretende abarcar toda a teoria de semigrupos, só aquela teoria que precisamos nesta dissertação.

Definição 2.4.1. *Uma família* $\{S_t\}_{t\geq 0}$ *de operadores limitados em* $C_b(\mathcal{S})$ *diz-se um semigrupo se:*

1. $S_0 = I$.
2. $S_{t+s} = S_t S_s$, $\forall s, t \geq 0$.

Definição 2.4.2. *Uma família* $\{S_t\}_{t\geq 0}$ *é um semigrupo fortemente contínuo se* $\lim_{t\downarrow 0} \|S_t f - f\|_\infty = 0$, $\forall f \in C_b(\mathcal{S})$. *Se cada* S_t *é uma contração, isto é, satisfaz* (2.10), *então* $\{S_t\}_{t\geq 0}$ *diz-se um semigrupo de contração.*

Lema 2.4.3. *Suponha que* $\{S_t\}_{t\geq 0}$ *é um semigrupo de contração fortemente contínuo em* $C_b(\mathcal{S})$. *Então, para cada* $f \in C_b(\mathcal{S})$ *a função* $t \to S_t f$ *é uma função uniformemente contínua.*

Demonstração. Para $t, h \geq 0$ temos

$$\|S_{t+h} f - S_t f\|_\infty = \|S_t(S_h - I)f\|_\infty \leq \|S_h f - f\|_\infty,$$

e para $0 \leq h \leq t$ temos,

$$\|S_{t-h} f - S_t f\|_\infty = \|S_{t-h}(S_h - I)f\|_\infty \leq \|S_h f - f\|_\infty.$$

Nos dois casos anteriores usamos o fato que S_t é uma contração. Assim, pela Definição 2.4.2 temos que o lado direito das expressões anteriores se anulam quando $h \to 0$. Logo a função é uniformemente contínua. □

Definição 2.4.4. *O gerador infinitesimal ou, simplesmente, gerador de um semigrupo* $\{S_t\}_{t\geq 0}$, *é o operador* L *definido por*

$$Lf = \lim_{t\to 0} \frac{S_t f - f}{t}, \tag{2.11}$$

com domínio $D(L)$ *que é o conjunto das funções* $f \in C_b(\mathcal{S})$ *para as quais o limite existe. A convergencia e no sentido da norma de* $C_b(\mathcal{S})$, *isto é, na norma do supremo.*

Lema 2.4.5. *Suponha que $\{S_t\}_{t\geq 0}$ é um semigrupo de contração fortemente contínuo em $C_b(\mathcal{S})$ com gerador L. Então valem as seguintes afirmações:*

1. *Se $f \in C_b(\mathcal{S})$ e $t \geq 0$. Então, $\int_0^t S_s f ds \in D(L)$ e*

$$S_t f - f = L \int_0^t S_s f ds.$$

2. *Se $f \in D(L)$ e $t \geq 0$. Então, $S(s)f \in D(L)$ e*

$$\frac{d}{dt} S_t f = L S_t f = S_t L f.$$

3. *Se $f \in D(L)$ e $t \geq 0$. Então,*

$$S_t f - f = \int_0^t L S_s f ds = \int_0^t S_s L f ds.$$

Corolário 2.4.6. *Suponha que $\{S_t\}_{t\geq 0}$ é um semigrupo de contração fortemente contínuo em $C_b(\mathcal{S})$ com gerador L. Então, $D(L)$ é denso em $C_b(\mathcal{S})$.*

Para a demonstração do Lema (2.4.5) e do Corolário (2.4.6), veja [4],página 9 e 10.

2.5 Medida de Probabilidade Invariante

Seja $\{P^x\}_{x\in\mathcal{S}}$ um processo de Markov a tempo contínuo que é um processo de Feller. Assuma que S_t é um semigrupo de contração no espaço de Banach $C_b(\mathcal{S})$. Seja $M_1(\mathcal{S})$ o espaço das medidas de probabilidade em $\mathcal{S}$, que também é um espaço métrico quando munido da seguinte convergência. A convergência em $M_1(\mathcal{S})$ é dada pela convergência fraca definida por

$$\mu_n \xrightarrow[n\uparrow\infty]{w} \mu \Leftrightarrow \int f(x)\mu_n(dx) \xrightarrow[n\uparrow\infty]{} \int f(x)\mu(dx), \, \forall f \in C_b(\mathcal{S}).$$

Para $\mu \in M_1(\mathcal{S})$ e $t \geq 0$, definimos $\mu S_t \in M_1(\mathcal{S})$ por

$$\int f(x)\mu S_t(dx) = \int S_t f(x)\mu(x), \, \forall f \in C_b(\mathcal{S}).$$

Assim, se $B \in \mathcal{B}_{\mathcal{S}}$ temos

$$\mu S_t(B) = \int P^x(X_t \in B)\mu(dx).$$

A medida de probabilidade μS_t é interpretada como a distribuição no tempo t de um processo com gerador L e semigrupo S_t, quando a distribuição inicial desse processo é μ.

Agora, dizemos que μ é invariante para o processo $\{X_t\}_{t\geq 0}$ se

$$\mu S_t = \mu, \ \forall t \geq 0.$$

Seja $\mathcal{I}$ o conjunto das medidas invariantes para $\{X_t\}_{t\geq 0}$. Agora, invariância de μ implica que se o estado inicial X_0 tem distribuição μ, então X_t também tem distribuição μ para todo $t \geq 0$. Além disso, o processo $\{X_t\}_{t\geq 0}$ é estacionário o que significa que a distribuição do processo transladado no tempo é a mesma que a distribuição do processo original.

Definição 2.5.1. *Para cada operador linear L, um subespaço linear Y de $D(L)$ é um cerne de L se o gráfico de L é o fecho do gráfico de $L\mid_Y$.*

Teorema 2.5.2. *Seja L o gerador de um semigrupo de contração fortemente contínuo S_t em $C_b(\mathcal{S})$ definido por um processo de Markov. Seja μ uma medida de probabilidade em $\mathcal{S}$. Seja Y um cerne para L. Então μ é invariante para X_t se e somente se*

$$\int Lf(x)\mu(dx) = 0, \ \forall f \in Y.$$

Demonstração. $\Rightarrow$) Suponhamos que μ é invariante e $f \in D(L)$. Como $\dfrac{S_t f - f}{t} \underset{n\uparrow\infty}{\longrightarrow} Lf$, uniformemente, então

$$\int Lf d\mu = \lim_{t\to 0}\int \frac{S_t f - f}{t} d\mu = \lim_{t\to 0}\frac{1}{t}\left(\int S_t f d\mu - \int f d\mu\right)$$
$$= \lim_{t\to 0}\frac{1}{t}\left(\int f d\mu - \int f d\mu\right) = 0.$$

$\Leftarrow$) Agora, assumamos que $\int Lf d\mu = 0, \ \forall f \in Y$. Pela definição de cerne temos que, $\forall f \in D(L)$, $\exists \{g_n\}_{n\geq 0} \in Y$ tal que $Lg_n \underset{n\uparrow\infty}{\longrightarrow} Lf$, uniformemente, quando $n \to \infty$. Logo $\int Lf d\mu = 0, \ \forall f \in D(L)$.

Fixada $f \in D(L)$, pelo Lema 2.4.3, temos que $S_t f \in D(L), \forall t \geq 0$. Assim, usando o Teorema de Fubini temos

$$\int S_t f d\mu - \int f d\mu = \int_0^t \left(\int L[S_s f] d\mu\right) ds = 0,$$

isto implica que

$$\int f d(\mu S_t) = \int f d\mu, \ \forall f \in D(L).$$

Além disso, pelo Corolário 2.4.6, temos que $D(L)$ é denso em $C_b(\mathcal{S})$. Assim, temos que $\mu S_t = \mu$. □