MAE326 - Aplicações de Processos Estocásticos

Revisão.

Cadeias de Markov em Tempo Contínuo.

É uma coleção de variáveis aleatórias (v.a.)
$\{X_t, t \in T\}$

indexadas por um conjunto T = $\mathbb{R}_+ \equiv \{ t \in \mathbb{R} : t \geq 0 \}$,
que podemos imaginar ser o tempo, satisfazendo uma condição a respeito da dependência entre essas variáveis em tempos distintos (condição de Markov abaixo).

As v.a. assumem valores num conjunto S finito ou infinito enumerável, que podemos chamar de "espaço de estados". Esse conjunto deve descrever o conjunto de todas as "situações possíveis" do sistema de interesse num determinado instante. Deve conter toda a informação necessária para que se possa identificar "as probabilidades de cada evolução a partir deste instante".

$$S = \{x_1, x_2, \dots, x_N\} \quad \text{ou} \quad S = \{x_1, x_2, \dots\}$$

Obs: Um conjunto é dito enumerável se posso associar um número natural a cada um de seus elementos, seu "ID" ou "número de identificação"; dessa forma posso ir "enumerando", fazer uma chamada a partir desta "lista de presença", tal que todo elemento seja eventualmente chamado.

Exercícios: O conjunto dos números naturais é claramente enumerável. O conjunto de números inteiros é enumerável? O conjunto dos números racionais é enumerável? O conjunto dos números reais é enumerável? Para cada caso, se a resposta for afirmativa, apresente uma possível "lista de chamada". Por outro lado, suponha que você suspeita que algum desses conjuntos não é enumerável, como "mostrar isso matematicamente"? Um conjunto "não-enumerável" seria "mais infinito" que o conjunto dos números naturais?

Condição de Markov: *Dizemos que $\{X_t, t\in T\}$ é uma cadeia de Markov se*

$$P(X_{t_{n+1}}=j \mid X_{t_0}=i_0, X_{t_1}=i_1, \dots, X_{t_{n-1}}=i_{n-1}, X_{t_n}=i) =$$

$$P(X_{t_{n+1}}=j \mid X_{t_n}=i)$$

para toda coleção de instantes $0\le t_0 \le t_1 \le \dots \le t_n \le t_{n+1} < \infty$

e estados (elementos de S) $i_0, i_1, \dots, i_{n-1}, i, j.$

Obs: considere os três eventos:

"Passado" = $\{X_{t_0}=i_0, X_{t_1}=i_1, \dots, X_{t_{n-1}}=i_{n-1}\}$

"Presente" = $\{X_{t_n}=i\}$

"Futuro" = $\{X_{t_{n+1}}=j\}$

a condição de Markov estabelece que

P("Futuro" | "Passado" $\cap$ "Presente") = P("Futuro" | "Presente")

Isso implica que "conhecendo o presente, o passado e futuro são independentes"

Ou seja:

P("Passado" $\cap$ "Futuro" | "Presente") =

P("Passado" | "Presente") . P("Futuro" | "Presente")

(exercício: demonstre isso a partir da condição de Markov)

Definição: *Dizemos que a cadeia é homogênea se*

$$P(X_{t+s}=j \mid X_s=i) = P(X_t=j \mid X_0=i) \qquad (1)$$

para todo i, j em S e instantes $s\ge 0$ e $t\ge 0$.

Obs: O lado esquerdo de (1) pede a prob. da cadeia estar em j após um tempo extra t, se sei que "agora", quando meu relógio indica o instante s, a cadeia está em i. A identidade indica que essa prob. não depende do "horário do relógio". Pense num exemplo que esta identidade vale e num exemplo na qual não vale.

O lado direito da identidade (1), para cada $t \geq 0$ é uma matriz na qual cada termo é identificado por dois estados, i (linha) e j (coluna).

Matriz $P(t)$ com

$$(P(t))_{ij} \equiv P(X(t)=j \mid X(0)=i) \equiv P_{ij}(t) \quad \text{para } i,j \in S \text{ e } t \geq 0$$

termo de na linha i e coluna j

indica que é uma definição

para simplificar a notação.

Obs: Note que $P(0) = \mathbb{1}$ = "Matriz identidade"

$$(\mathbb{1})_{ij} = \delta_{ij} = \begin{cases} 1 & \text{se } i=j \\ 0 & \text{caso contrário} \end{cases}$$

"função δ de Kronecker"

Para todo $t \geq 0$ essa matriz $\mathbb{P}(t)$ é uma "matriz estocástica", ou seja, satisfaz:

$$P_{ij}(t) \geq 0 \quad \text{e} \quad \sum_j P_{ij}(t) = 1$$

(Porquê?)

$$\left(\sum_i P_{ij}(t) = ? \right)$$

Obs: Em geral, achar uma expressão concreta para $\mathbb{P}(t)$ não é simples, mesmo no caso em que S é finito. Vamos ver alguns exemplos, principalmente nessa situação finita, na qual temos um problema de álgebra linear (no caso geral, surgem diversas questões e fenômenos novos; envolve, basicamente, uma área da matemática conhecida como análise funcional)

Obs: Para uma cadeia de Markov em tempo Discreto, a matriz $\mathbb{P}$ definida por:

$$(\mathbb{P})_{ij} = P(X_1 = j \mid X_0 = i)$$

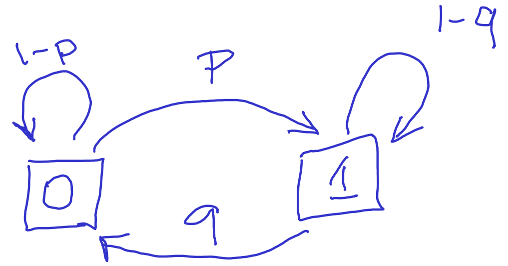
é chamada de Matriz de Transição da Cadeia de Markov (em tempo discreto).

Alguns exemplos:

Tempo discreto:

1) $S = \{0, 1\}$, $\{X_n\}_{n \geq 0}$ com

$$\mathbb{P} = \begin{array}{c} \\ 0 \\ 1 \end{array}\!\!\begin{array}{c} \begin{array}{cc} 0 & \;\;1 \end{array} \\ \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix} \end{array}$$



sendo p e q dois parâmetros, com $0 \leq p \leq 1$ e $0 \leq q \leq 1$.

2) *Problema do "duelo triplo" (ex 4, lista de revisão 1)*

S = { 000, 001, ..., 111 }; cada trio de valores 1 e 0, 1 = "vivo" e 0 = "morto" (lembre que é apenas uma brincadeira de paintball...), indica a situação de cada jogador, respectivamente, A, B e C. Você pode, claro, representar S de outras formas.

Por definição do problema (certo?) a cadeia é homogênea. Por exemplo, temos

$$P_{ij} = 0 \quad \text{se} \quad i = 111 \;\; e \;\; j = 000 \qquad \text{(Porquê?)}$$

Tempo contínuo:

1) $S = \{0, 1\}$, $\{X_t\}_{t \geq 0}$ **com**

$$P_{01}(t) = P(\text{" } N(t) \text{ é ímpar"}) \qquad P_{11}(t) = P(\text{" } N(t) \text{ é par"})$$

onde $N(t)$ ***, para*** $t \geq 0$ ***, são v. a. independentes e com distribuição Poisson com parâmetro*** λ .

Exercício: encontre a fórmula para todos os quatro elementos da matriz 2 por 2

$\mathbb{P}(t)$

Obs: Voltaremos a discutir este exemplo. Ainda não recordamos a noção da matrix de taxas de uma cadeia em tempo contínuo, mas você já pode ir pensando: quais são as taxas desta cadeia?

2) Exercício 4 da lista "exercíciosRevisão2.pdf" (três funcionários que alternam entre "trabalhar" e "tomar um cafesinho")

Para a cadeia $\{X_t\}_{t\geq 0}$ podemos tomar o mesmo espaço de estados S do problema do duelo triplo, com 1 = "trabalhando" e 0 = "tomando café".
Para $\{Y_t\}_{t\geq 0}$ temos uma cadeia de dois estados. É o mesmo $\mathbb{P}(t)$ que no exemplo acima?

Equação de Chapman-Kolmogorov.

$$P_{ij}(s+t) = \sum_{k\in S} P_{ik}(s)\, P_{kj}(t) \quad \text{para todo } i,j \text{ em } S$$

Demonstração: $(X_t \equiv X(t))$

$$P_{ij}(s+t) \equiv P(X(t+s)=j \mid X(0)=i) \overset{\text{porquê?}}{=} \sum_{k\in S} \underbrace{P(X(t+s)=j, X(s)=k \mid X(0)=i)}_{\circledast}$$

$$\circledast = \frac{P(X(t+s)=j, X(s)=k, X(0)=i)}{P(X(0)=i)}$$

$$= \underbrace{P(X(t+s)=j \mid X(s)=k, X(0)=i)}_{P(X(t+s)=j \mid X(s)=k)} \cdot \frac{P(X(s)=k, X(0)=i)}{P(X(0)=i)} \overset{\text{definição}}{=} \underbrace{P(X(s)=k \mid X(0)=i)}_{P_{ik}(s)}$$

Propr. Markov

$$\underbrace{P(X(t+s)=j \mid X(s)=k)}_{P(X(t)=j \mid X(0)=k)}$$

Homogeneidade

$$\underbrace{P(X(t)=j \mid X(0)=k)}_{P_{kj}(t)}$$

Notação

e segue a equação de Chapman-Kolmogorov. □

Para cada $t \geq 0$, $\mathbb{P}(t)$ é uma matrix com linhas e colunas indexadas por elementos de S, o espaço de estados. Então essa identidade de Chapman-Kolmogorov diz que

$$\mathbb{P}(s+t) = \mathbb{P}(s) \cdot \mathbb{P}(t)$$

produto de matrizes.

A informação disponível sobre uma cadeia de Markov num determinado instante é representada por uma distribuição de probabilidade em S.

Distribuição de probabilidades no espaço de estados S.

Dizemos que π *é uma distribuição de probabilidade em S se for uma função* $\pi : S \to \mathbb{R}$ *tal que*

$$\pi(i) \geq 0 \ , \forall i \in S \quad e \quad \sum_{i \in S} \pi(i) = 1$$

Notação: $\pi = \{\pi(i)\}_{i \in S}$ *que podemos imaginar como um "vetor linha", uma matriz com apenas uma linha e colunas indexadas por S.*

Ao estudar uma cadeia de Markov frequentemente começamos fixando a distribuição de probabilidade no instante zero, π_0 *a distribuição inicial da cadeia.*

E estamos interessados na distribuiçao de probabilidade no instante t ≥ 0 *, que é dado por*

$$\pi_t(j) = \sum_{i \in S} \pi_0(i) P_{ij}(t) \qquad \text{(verifique!)}$$

ou, usando a notação usual de matrizes $\pi_t = \pi_0 \mathbb{P}(t) \ , \ \forall t \geq 0$

Cada estado da cadeia de interesse representa "a presente situação" de algum sistema que pode ser bastante complicado e que foi iniciado numa situação bastante particular, digamos , "o sistema começa vazio e com todos os seus componentes em perfeito estado de funcionamento". Frequentement há alguma v.a. de interesse (por exemplo, o "lucro" ou "custo do sistema") e estamos preocupados com o 1) valor médio desta v.a "ao longo de um grande período de tempo de funcionamento do sistema" ou 2)"no valor esperado desta v.a. num "instante t grande".

São duas questões distintas, mas intimamente relacionadas. Basicamente elas *têm a ver com a "perda de memória" de uma cadeia de Markov, que acontece sob certas condições, e faz com que "o particular estado inicial" se torne irrelevante. Mais precisamente, temos que o limite*

$$\lim_{t \uparrow \infty} P_{ij}(t)$$

vai existir (sob condições) e não depende de i (estado inicial). Vocês devem ter visto estes resultados para cadeias de Markov em tempo discreto. Vamos recaptular e notar que cadeias em tempo contínuo envolvem basicamente as mesmas idéias.

Definição: dizemos que uma distribuição de probabilidade $\{\pi(i)\}_{i \in S}$ ***é invariante pela evolução da cadeia de Markov com*** $\{\mathbb{P}(t)\}_{t \geq 0}$

se $$\pi(j) = \sum_{i \in S} \pi(i) P_{ij}(t), \quad \forall\, t \in T$$ *(caso discreto ou contínuo)*

Exercícios: 1) Mostre que π, com $\pi(0) = \frac{q}{p+q}$ e $\pi(1) = \frac{p}{p+q}$

é invariante para o exemplo da cadeia discreta de dois estados acima.

2) Mostre que π, com $\pi(0) = \frac{1}{2}$ e $\pi(1) = \frac{1}{2}$

é invariante para o exemplo da cadeia contínua de dois estados acima.

3) Considere a generalização do exercício 6, listaRevisão1.pdf: N bolas numeradas de 1 até N são colocadas de uma forma qualquer entre duas urnas A e B; a cada instante escolho, com distribuição uniforme, um número entre 1 e N+k, sendo k um número inteiro não-negativo qualquer; se o número escolhido for menor ou igual a N, troco a bola correspondente de urna; caso contrário as bolas permanecem onde estão e passo ao instante seguinte, quando o procedimento se repete. $X_n, n \geq 0$ ***indica o número de bolas na urna A no instante n. Verifique que esta é uma cadeia de Markov (modelo de urnas de Ehrenfest). Qual é o espaço de estados? Considere a distribuição Binomial(N, p); ela é invariante?***

Equação de Balanceamento: Suponha que $\{X_t\}_{t\geq 0}$ *é uma cadeia de Markov em tempo contínuo em S com probabilidades de transição* $\{\mathbb{P}(t)\}_{t\geq 0}$ *tal que* $\{\pi(i)\}_{i\in S}$ *é uma distribuição de probabilidade estacionária, ou seja*

$$\pi(i) = \sum_{j\in S} \pi(j) P_{ij}(t) \quad , \quad \forall \; t \geq 0$$

Podemos reescrever essa equação como

$$\pi(i) = \sum_{j\in S} \pi(j) P_{ij}(t) = \sum_{j\in S\setminus\{i\}} \pi(j) P_{ij}(t) + \underbrace{\pi(i) P_{ii}(t)}_{\text{quando } j=i}$$

ou seja, $j\in S$ mas $j \neq i$

de onde segue que, para todo i em S e $t \geq 0$

"Há balanceamento no fluxo de prob."

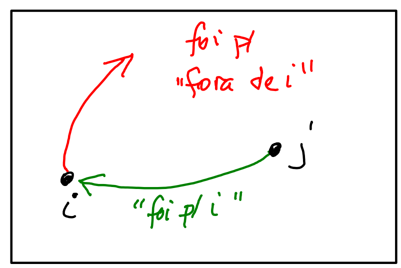
$$(*) \quad \pi(i)\big(1 - P_{ii}(t)\big) = \sum_{\substack{j\in S \\ j\neq i}} \pi(j) P_{ji}(t)$$

"Estar em i no instante 0 mas não estar em i no instante t"

"Saiu do estado i indo para outro estado qualquer"

"Estar em outro estado j, diferente de i, mas está em i no instante t"

"foi para i a partir de um outro estado qualquer"



S espaço de estados

"fluxo de prob. para i" = "fluxo de prob. para fora de i" $\forall i$

No caso discreto a equação de balanceamento pode ser escrita em termos da Matriz de transição em um passo:

$$\pi(i)\big(1 - P_{ii}\big) = \sum_{\substack{j\in S \\ j\neq i}} \pi(j) P_{ji} \quad \text{para todo } i\in S$$

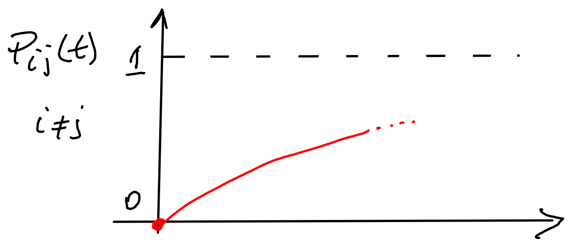
Nesse caso discreto a matriz de transição $\mathbb{P}$ *é conhecida e a solução deste sistema de equações fornece a distribuição de probabilidade estacionária da cadeia.*

Mas no caso de cadeias de Markov em tempo contínuo não temos $\mathbb{P}(t)$!
Note que (*) *estabelece uma identidade, para todo t, entre duas funções do tempo, uma no lado esquerdo e outra no lado direito.*

Como são os termos da matriz de transição em tempo contínuo?

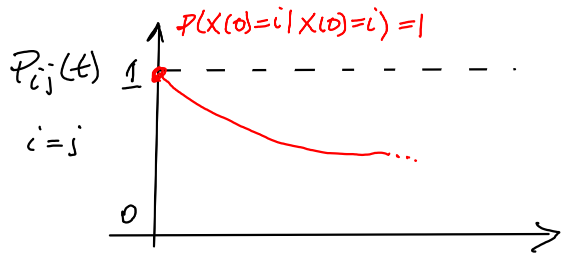
Duas situações para $P_{ij}(t) \equiv P(X(t)=j \mid X(0)=i)$

① $i \neq j$



$P(X(0)=j \mid X(0)=i)=0$
$i \neq j$

② $i = j$



Sem entrar em detalhes matemáticos, parece claro que

$\frac{d}{dt} P_{ij}(t) \geq 0$ para $i \neq j$ (a prob. começa em 0 e só pode aumentar...)

$\frac{d}{dt} P_{ij}(t) \leq 0$ para $i \neq j$ (a prob. começa em 1 e só pode diminuir ...)

Notação: Defina a Matriz de taxas Q, cujos componentes são estas derivadas no instante zero

$$\underbrace{(Q)_{ij}}_{\text{componente da matriz } Q \text{ na linha } i \text{ e coluna } j.} \overset{\text{notação}}{=} q_{ij} \overset{\text{definição}}{\equiv} \frac{d}{dt} P_{ij}(t)\Big|_{t=0} = P'_{ij}(0)$$

derivada calculada em $t=0$

Obs: esta matriz indica a velocidade ("taxas") com que as probabilidades $P_{ij}(t)$ ***variam. Como indicado acima, temos***

$$q_{ij} \geq 0 \quad i \neq j \quad e \quad q_{ii} \leq 0$$

ou seja, apenas os termos da diagonal de Q podem ser negativos.

Notação: defino q_i como

$$q_i = -q_{ii} \geq 0$$

Note que $\sum_{j\in S} P_{ij}(t) = 1$ para todo $i\in S$ e $t\geq 0$

e "portanto" (pergunte para o professor de cálculo...)

$$\frac{d}{dt}\left(\sum_{j\in S} P_{ij}(t)\right) \overset{?}{=} \sum_{j\in S} \frac{d}{dt} P_{ij}(t) = 0$$

ou seja, $\sum_{j\in S} q_{ij} = 0$, para todo $i\in S$.

Então, derivando os dois lados da equação de balanceamento $(*)$

acima, temos que uma distribuição $\{\pi(i)\}_{i\in S}$ *é estacionária se e somente se*

$$\pi(i) q_i = \sum_{\substack{j\in S \\ j\neq i}} \pi(j) q_{ji} \quad , \forall\, i\in S$$

Obs: no caso discreto, a maneira mais simples para definirmos uma cadeia de Markov é indicarmos sua matriz de transição (em um passo) $\mathbb{P}$.

No caso contínuo, a maneira mais natural é apresentar sua Matriz de Taxas Q.

Exemplo: Considere a cadeia de dois estados com

$$Q = \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix} \quad \text{para} \quad \lambda > 0 \text{ e } \mu > 0$$

Qual é a distribuição estacionária?

Note que a relação entre $\mathbb{P}(t)$ *e* Q *é dada pela equação*

$$\left.\frac{d}{dt}\mathbb{P}(t)\right|_{t=0} = Q \quad \text{com} \quad \mathbb{P}(0) = \mathbb{1} \ (\text{matriz identidade})$$

Condição de Markov e v.a. exponenciais.

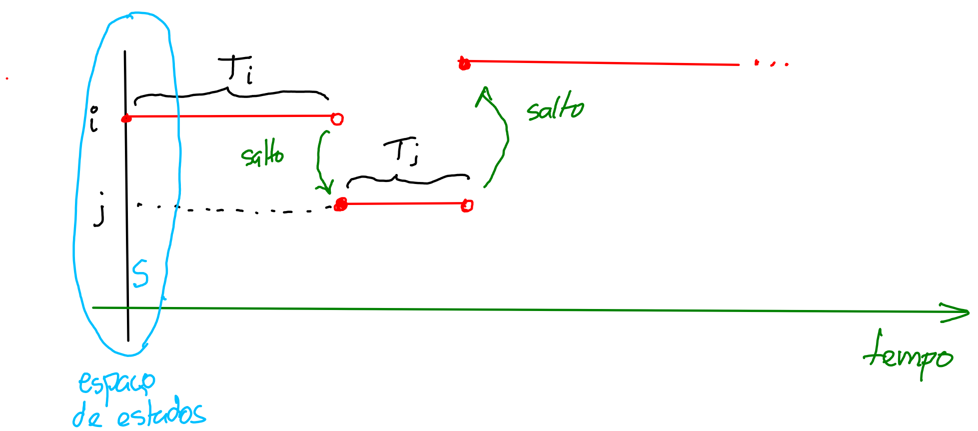
Seja $\{X_t\}_{t\geq 0}$ *uma cadeia de Markov em tempo contínuo em S*

Suponha que esta cadeia começa num certo estado i em S, ou seja, suponha que a distribuição de probabilidade no instante 0, $\{\pi_0(k)\}_{k\in S}$

é dada por

$$\pi_0(k) = P(X(0)=k) = \delta_{ik} \underset{\text{definição}}{\equiv} \begin{cases} 1 & \text{se } k=i \\ 0 & \text{caso contrário} \end{cases}$$

O que vai acontecer ao longo do tempo? O "natural" seria imaginar que a cadeia permanece certo tempo aleatório T_i *neste estado inicial i, ao cabo do qual "salta" para outro estado j; uma vez que alcança este estado j, permanece um tempo aleatório* T_j *quando "salta" novamente e assim sucessivamente.*



A condição de Markov garante uma enorme simplicidade.
Tanto com respeito aos sucessivos tempos de permanência como na maneira como são escolhidos os sucessivos saltos visitados.

Basicamente temos que estes sucessivos tempos de permanência são independentes e tem distribuição exponencial. As sucessivas transições ocorrem conforme uma cadeia de Markov em tempo discreto.

Obs: a variável aleatória exponencial aparece basicamente por ser uma v.a. contínua que tem a propriedade de "não ter memória", ou seja, se T é uma v.a. com distribuição exponencial, então

$$P(T > s+t \mid T > s) = P(T > t)$$ Verifique!

Em palavras: suponha que T seja o tempo que certo aparelho demora para falhar. Se sei que o aparelho estava funcionando no instante s ($\{T > s\}$), a probabilidade dele funcionar por mais um tempo t (lado esquerdo da identidade) é a mesma que se tivesse acabado de ligar um "aparelho novo" (lado direito da identidade). "O aparelho não lembra que já funcionou pelo tempo s".

Pode parecer surpreendente, mas há diversas situações "reais" de "tempos de falha" que tem esta propriedade de "perda de memória". Em física quântica, por exemplo, isto é exatamente o que acontece para tempos de fissão de átomos instáveis.

Obs: Existe uma v.a. discreta (que vocês certamente conhecem) que também tem esta propriedade de "perda de memória". Qual é?

Por conta deste resultado, temos a construção de cadeias de Markov em tempo contínuo que já mencionei (e aparece no exercício 7 da listaRevisão1.pdf).

Uma cadeia de Markov em tempo contínuo, com matriz de taxas Q, pode ser construida em duas etapas independentes.

I) Parâmetros dos sucessivos tempos exponenciais de permanência em cada estado.

Os sucessivos tempos de permanência têm distribuição exponencial e são independentes. Os parâmetros dos sucessivos tempos em cada estado são iguais.

Defino uma função, que denoto por q, de S nos números reais não negativos

$$q: S \to \mathbb{R}_+$$

de tal forma que q(i) $= q_i$ é o parâmetro associado ao estado i. Mais precisamente, se i é um estado, denote por

$\{T_i^k\}_{k\geq 1}$ *a coleção de sucessívos tempos de permanência neste estado i.*

(ou seja, na primeira visita, permanece pelo tempo T_i^1 *, quando salta para outro estado; se ele retornar eventualmente, vai permanecer por um tempo* T_i^2 *; e assim sucessivamente).*

Temos que $\{T_i^k\}_{k\geq 1}$, $i \in S$

são v.a. exponenciais independentes, com $T_i^k \sim exp(q_i)$

II) Matriz de transição da cadeia (discreta) "imersa": $\mathbb{P}$
Ao final de cada tempo de permanência num dado estado i em S a cadeia salta para um estado j com probabilidade $(\mathbb{P})_{ij} = P_{ij}$

A relação entre estre a função q e a matriz P com a matriz de taxas Q é dada por: $\mathbb{P}$

$$q_{ij} = \overset{q(i)}{q_i}\, \mathbb{P}_{ij} \quad \text{se} \quad i \neq j$$

e $$q_{ii} = -q_i$$

Demonstração: É um pouco longa, pois vou detalhar para tentar ser claro.
(Exercício, complete os argumentos)

Considere $i \neq j$

$$q_{ij} = \frac{d}{dt} P_{ij}(t)\Big|_{t=0} = \lim_{\Delta t \downarrow 0} \frac{P_{ij}(\Delta t) - \underbrace{P_{ij}(0)}_{=0}}{\Delta t} \quad \text{(Porquê?)}$$

Podemos decompor $P_{ij}(\Delta t) = P(X(\Delta t) = j \mid X(0) = i)$ em 2 casos

$$P_{ij}(\Delta t) = P(\overbrace{T_i < \Delta t, \text{"No instante } T_i \text{ salta para } j\text{"}, \text{"}X_{\Delta t} = j\text{"}}^{A} \mid X(0) = i)$$

$$+ P(\overbrace{T_i < \Delta t, \text{"No instante } T_i \text{ salta de } i \text{ para algum } k \neq j\text{"}, \text{"}X_{\Delta t} = j\text{"}}^{B} \mid X(0) = i)$$

O evento (A) exige 3 condições:

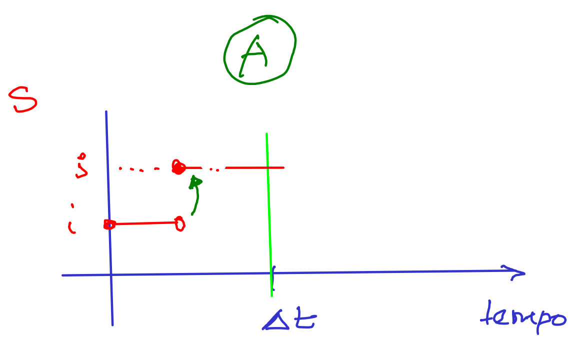
$$A = A_1 \cap A_2 \cap A_3$$

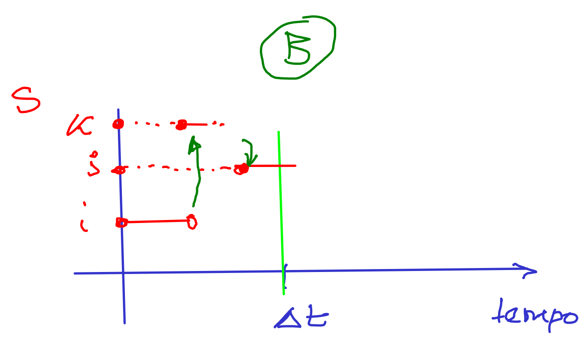
"o tempo de permanência em i é menor que Δt, A_1
No instante em que deixa i, salta diretamente para j, A_2
no instante Δt (A_3) está em j"

O evento (B) também exige 3 condições; considera a possibilidade de que houve mais que um salto na transição entre o estado i e o estado j durante o intervalo de tempo Δt.

$$B = B_1 \cap B_2 \cap B_3$$

"o tempo de permanência em i é menor que Δt, B_1
No instante em que deixa i, salta para $k \neq j$, B_2
no instante Δt (B_3) está em j"





Note que os dois casos são "pouco prováveis", pois exigem, entre outras condições, que

"$T_i < \Delta t$" com $T_i \sim \exp(q_i)$

Exercício: Verifique que $P(T_i < \Delta t) = 1 - e^{-q_i \Delta t} = q_i \Delta t + o(\Delta t)$

Notação: Dizer que $f(x)=o(x)$, para uma função f

com $f(0)=0$ significa que $\lim\limits_{x\downarrow 0}\frac{f(x)}{x}=0$ ou seja

"f vai a zero mais rápido que linearmente"

e Dizer que $f(x)=o(x^2)$ significa que $\lim\limits_{x\downarrow 0}\frac{f(x)}{x^2}=0$

"f vai a zero mais rápido que quadraticamente".

Exemplos: Se $f(x)=x^2$ então $f(x)=o(x)$

mas $f(x) \cancel{=} o(x^2)$

Se $f(x)=1-e^{-\lambda x}$, para $\lambda\in\mathbb{R}$, verifique que

$$f(x)=\lambda x+o(x)$$

Usando esta notação, vejamos os dois termos da decomposição de $P_{ij}(\Delta t)$ ***acima***

$$P(A\,|\,X(\Delta t)=i) \underset{\uparrow}{=} P(A_1\,|\,X(\Delta t)=i)\;P(A_2\cap A_3\,|\,X(\Delta t)=i)$$

pela independência entre T_i e "o que acontece depois"

mas

$$P(A_1\,|\,X(\Delta t)=i) = P(T_i<\Delta t\,|\,X(\Delta t)=i) = q_i\Delta t + o(\Delta t)$$

e

$$\underbrace{P(A_2\cap A_3}\,|\,X(\Delta t)=i) = P_{ij} + o(\Delta t)$$

como $T_j\sim exp(q_j)$ $(q_j>0)$
essa prob. vai a 1 quando $\Delta t\downarrow 0$

{"saltou de i para j"} $\cap$ ({"permaneceu em j até Δt"} $\cup$ {"Saiu de j mas retornou até o instante Δt"})

"cadeia de Markov discreta imersa saltar de i para $j = P_{ij}$"

essa prob. é $o(\Delta t)$ pois envolve acontecer de pelo menos duas v.a. exponenciais serem muito pequenas.

Portanto

$$P(A \mid X(\Delta t)=i) = \left(q_i \Delta t + o(\Delta t)\right) \cdot \left(P_{ij} + o(\Delta t)\right)$$

$$= q_i P_{ij} \Delta t + o(\Delta t)$$

Agora, vejamos o termo com $\boxed{B}$

$$P(B \mid X(\Delta t)=i) = P(B_1 \mid X(\Delta t)=i)\, P(B_2 \cap B_2 \mid X(\Delta t)=i) = o(\Delta t)$$

envolve "$T_i < \Delta t$" cuja prob. é $o(\Delta t)$

envolve pelo menos dois saltos mais rápidos que Δt: o de i para k e o salto final para j; estes saltos são indep., cada um deles $o(\Delta t)$

Portanto: prob. $= o(\Delta t^2)$

Então, no caso $i \neq j$ que estamos considerando

$$q_{ij} = \frac{d}{dt} P_{ij}(t)\Big|_{t=0} = \lim_{\Delta t \downarrow 0} \frac{P_{ij}(\Delta t)}{\Delta t}$$

$$= \lim_{\Delta t \downarrow 0} \frac{q_i P_{ij} \Delta t + o(\Delta t)}{\Delta t} = q_i P_{ij}$$

O caso $i=j$ é mais simples. Usando argumentos similares, verifique que

$$q_{ii} = \frac{d}{dt} P_{ii}(t)\Big|_{t=0} = \lim_{\Delta t \downarrow 0} \frac{\overbrace{P(X(\Delta t)=i \mid X(0)=i)}^{1-q_i\Delta t + o(\Delta t)} - \overbrace{P(X(0)=i \mid X(0)=i)}^{1}}{\Delta t}$$

$$= q_i$$

$\square$

A dinâmica de uma cadeia de Markov em tempo contínuo.

Equações de Kolmogorov:

"Adiantada" ("Forward"):

$$\dot{P}_{ij}(t) = \sum_{k \in S} P_{ik}(t)\, q_{kj} \quad \textbf{ou} \quad \dot{\mathbb{P}}(t) = \mathbb{P}(t).Q$$

notação matricial

"Retrógrada" ("Backward"):

$$\dot{P}_{ij}(t) = \sum_{k \in S} q_{ik}\, P_{kj}(t) \quad \textbf{ou} \quad \dot{\mathbb{P}}(t) = Q.\mathbb{P}(t)$$

onde $\dot{P}_{ij}(t) \overset{\text{notação}}{=} \frac{d}{dt} P_{ij}(t)$ e $\left(\dot{\mathbb{P}}(t)\right)_{ij} = \dot{P}_{ij}(t)$

Demonstração:

As duas equações são bastante parecidas mas não são equivalentes. A equação retógada é mais "fundamental", no sentido que é satisfeita sob hipóteses mais fracas. No nível deste curso podemos supor que essas duas equações são sempre válidas. Vou apresentar apenas uma "demonstração intuitiva" da equação adiantada, que desconsidera algumas questões matemáticas (que você pode tentar identificar).

A equação diferencial considera

$$\dot{P}_{ij}(t) = \frac{d}{dt} P_{ij}(t) \underset{\text{definição}}{=} \lim_{\Delta t \downarrow 0} \frac{P_{ij}(t+\Delta t) - P_{ij}(t)}{\Delta t}$$

Começamos com $P_{ij}(t+\Delta t) = P(X(t+\Delta t) = j \mid X(0) = i)$

$$= \sum_{k \in S} P(X(t+\Delta t) = j, X(t) = k \mid X(0) = i)$$

$$= \sum_{k \in S} \frac{P(X(t+\Delta t)=j,\ X(t)=k,\ X(0)=i)}{P(X(0)=i)}$$

$$= \sum_{k \in S} P(X(t+\Delta t)=j \mid X(t)=k,\ X(0)=i) \cdot P(X(t)=k \mid X(0)=i)$$

propr. Markov

$$= \sum_{k \in S} P_{ik}(t)\, P_{kj}(\Delta t) = \sum_{\substack{k \in S \\ k \neq j}} P_{ik}(t)\, P_{kj}(\Delta t) + P_{ij}(t)\, P_{jj}(\Delta t)$$

Usando argumentos que já apresentamos, verifique que

$$P_{kj}(\Delta t) = q_k P_{kj} \Delta t + o(\Delta t) \qquad \underline{\text{se}} \qquad k \neq j$$

matriz de transição da cadeia imersa.

<u>Obs:</u> Note que isso é verdade pois $P_{kj}(\Delta t) = P(X(\Delta t)=j \mid X(0)=k)$ envolve basicamente $\underbrace{\{T_k < \Delta t\}}_{\text{prob.} \,=\, q_k \Delta t + o(\Delta t^2)}$ e, em seguida, $\underbrace{\text{"saltar de } k \text{ para } j\text{"}}_{\text{prob.} \,=\, P_{kj}}$

De forma análoga, verifique que

$$P_{jj}(\Delta t) = 1 - q_j \Delta t + o(\Delta t)$$

Então, temos

$$P_{ij}(t+\Delta t) = \sum_{\substack{k \in S \\ k \neq j}} P_{ik}(t)\, P_{kj}(\Delta t) + P_{ij}(t)\, P_{jj}(\Delta t)$$

$$= \sum_{\substack{k\in S \\ k\neq j}} P_{ik}(t)\left(q_k P_{kj}\Delta t + o(\Delta t)\right)$$

$$+ P_{ij}(t)\left(1 - q_j\Delta t + o(\Delta t)\right)$$

todos os termos $o(\Delta t^2)$ são coletados aqui

$$= \sum_{\substack{k\in S \\ k\neq j}} P_{ik}(t)\, q_k P_{kj}\Delta t + P_{ij}(t) - P_{ij}(t)\, q_j \Delta t + o(\Delta t)$$

Rearranjando os termos, e dividindo por Δt

$$\frac{P_{ij}(t+\Delta t) - P_{ij}(t)}{\Delta t} = \sum_{\substack{k\in S \\ k\neq j}} \frac{P_{ik}(t)\, q_k P_{kj}\cancel{\Delta t}}{\cancel{\Delta t}} - \frac{P_{ij}(t)\, q_j \cancel{\Delta t}}{\cancel{\Delta t}} + \frac{o(\Delta t)}{\Delta t}$$

$\Delta t \downarrow 0$, $\Delta t \downarrow 0$, 0

$$\frac{d}{dt}P_{ij}(t) = \sum_{\substack{k\in S \\ k\neq j}} P_{ik}(t)\, q_k P_{kj} - P_{ij}(t)\, q_j$$

$$= \sum_{k\in S} P_{ik}(t)\, q_{kj}$$

que é a equação de Kolmogorov adiantada. $\square$

Note que $\left.\frac{d}{dt}\mathbb{P}(t)\right|_{t=0} = \mathbb{Q}$ **matriz de taxas**

As equações de Kolmogorov estabelecem o valor da derivada para todo t não-negativo

Adiantada $\quad \frac{d}{dt}\mathbb{P}(t) = \mathbb{P}(t) \cdot Q$

Retrógrada $\quad \frac{d}{dt}\mathbb{P}(t) = Q\,\mathbb{P}(t)$

A solução da equação $\frac{d}{dt} f(t) = a f(t)$ com a = constante e $f(0) = 1$ é $f(t) = e^{at}$

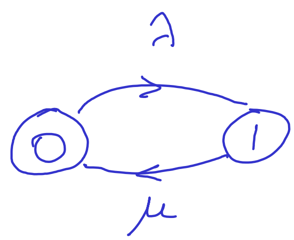
Quando S é finito, todas essas matrizes são finitas. A solução da equação de Kolmogorov é

$$\mathbb{P}(t) = e^{tQ} = \sum_{k=0}^{\infty} \frac{t^k}{k!} Q^n$$

Exemplo. Considere a cadeia de Markov em tempo contínuo mais simples: Uma cadeia com dois estados, S = { 0 , 1}

A matriz de taxas mais geral é

$$Q = \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix}$$



A solução para $\mathbb{P}(t)$ *envolve um sistema de equações diferenciais bastante simples e sua solução pode ser encontrada em diversos livros de processos estocásticos (por exemplo, item 6.11, página 381, do livro Intr. Prob. Models, S. Ross)*

Vou ilustrar a solução atraves da exponencial de matrizes. O cálculo de potencias de Q não parece simples, mas podemos usar álgebra linear elementar e começar encontrando uma transformação A que diagonaliza Q.

Queremos encontrar A, inversível, tal que

$$A^{-1} Q A = D = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}, \quad A = \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{00} \end{bmatrix}$$

ou seja $QA = AD$

$$\begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix} \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{00} \end{bmatrix} = \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{00} \end{bmatrix} \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}$$

que implica

$$\begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix} \underbrace{\begin{bmatrix} a_{00} \\ a_{10} \end{bmatrix}}_{v_1} = d_1 \begin{bmatrix} a_{00} \\ a_{10} \end{bmatrix}$$

$$\begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix} \underbrace{\begin{bmatrix} a_{01} \\ a_{00} \end{bmatrix}}_{v_2} = d_2 \begin{bmatrix} a_{01} \\ a_{00} \end{bmatrix}$$

ou seja, v_1 e v_2 são autovetores de Q com autovalores $\overset{0}{\overset{\shortparallel}{d_1}}$ e $\overset{-(\lambda+\mu)}{\overset{\shortparallel}{d_2}}$ (resp.)

Resolvendo, temos $A = \begin{bmatrix} 1 & -\lambda \\ 1 & \mu \end{bmatrix}$ e $D = \begin{bmatrix} 0 & 0 \\ 0 & -(\lambda+\mu) \end{bmatrix}$

e $$\mathbb{P}(t) = A \cdot e^{tD} \cdot A^{-1} = A \begin{bmatrix} 1 & 0 \\ 0 & e^{-(\lambda+\mu)t} \end{bmatrix} A^{-1}$$

$$= \frac{1}{\mu+\lambda} \begin{bmatrix} \mu + \lambda e^{-(\lambda+\mu)t} & \lambda\left(e^{-(\lambda+\mu)t} - 1\right) \\ \mu\left(e^{-(\lambda+\mu)t} - 1\right) & \lambda + \mu e^{-(\lambda+\mu)t} \end{bmatrix}$$