



RAFAEL POLLI CARNEIRO

Uma Abordagem Geométrica no Estudo de Processos Markovianos

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro da
“CAPES”

Santo André, 2018



Universidade Federal do ABC

Centro de Matemática, Computação e Cognição

Rafael Polli Carneiro

Uma Abordagem Geométrica no Estudo de Processos Markovianos

Orientador: Prof. Dr. Cristian Favio Coletti

Dissertação de mestrado apresentada ao Centro de
Matemática, Computação e Cognição para
obtenção do título de Mestre em Matemática

ESTA É A VERSÃO ORIGINAL DA TESE, TAL COMO
SUBMETIDA À COMISSÃO JULGADORA.

Santo André, 2018

AGRADECIMENTOS

Agradeço aos meus pais, Jacira Polli Carneiro e Celso Carneiro, e a minha família por estarem sempre ao meu lado e por me ensinarem a importância dos estudos. Sem eles, nada disto seria possível.

Ao meu orientador Cristian Favio Coletti e a professora Sandra Maria Zapata Yepes por compartilharem seus conhecimentos comigo, por me ajudarem no desenvolvimento da dissertação e pela amizade.

Aos meus amigos, em especial ao Washigton Fernandes, pelas contínuas palavras de incentivo e apoio incondicional. Tê-los ao meu lado, dividindo risadas, frustrações e alegrias é algo que guardarei na memória com carinho.

À esta universidade e seu corpo docente que contribuíram com meu aperfeiçoamento profissional e pessoal. Também gostaria de agradecer ao matemático Elon Lages Lima que, graças aos seus livros, me despertou o interesse pela matemática.

Finalmente, agradeço ao auxílio financeiro da CAPES a minha pesquisa.

A utopia está lá no horizonte. Me aproximo
dois passos, ela se afasta dois passos.
Caminho dez passos e o horizonte corre dez passos.
Por mais que eu caminhe, jamais alcançarei.
Para que serve a utopia? Serve para isso:
para que eu não deixe de caminhar.
- Eduardo Galeano

RESUMO

Neste trabalho estamos interessados em descrever as cadeias de Markov a tempo discreto, cujas matrizes de transição tenham determinante positivo, em função de processos markovianos a tempo contínuo, e vice-versa. Tal procedimento é possível se estudarmos as propriedades algébricas do grupo de matrizes de Lie, dado pelas matrizes estocásticas, e as propriedades algébricas de seu espaço tangente, o qual será uma álgebra de Lie. Ainda mais, ao estudarmos o coproduto de um elemento do centro do espaço tangente, como o Casimir, obtemos uma forma de construir uma matriz geradora de um processo de Markov a tempo contínuo. E, em posse desta matriz geradora, somos capazes de estabelecer um sistema de partículas de dois sítios, o qual, sob certas transformações, podemos estender para sistemas de N -partículas. Para encerrar, obtemos operadores de autodualidade para sistemas de N -partículas ($N \geq 2$) e provamos a versão estocástica do Teorema de Noether.

Palavras-chave: Sistema de Partículas, Grupo de Matrizes de Lie, Coproduto

ABSTRACT

In this work we are interested in a way to describe discrete Markov chains, whose transition matrices have positive determinant, in function of continuous Markov chains and vice versa. Such procedure is possible if we study the algebraic properties of the group of Lie matrices given by the stochastic matrices and its tangent space, which satisfy the condition of being a Lie algebra. Also, if we took the coproduct of an element in the center of the tangent space, just like the Casimir element, we obtain a generator matrix of a continuous Markov process. And, given this generator matrix, we are capable to define a particle system of two sites, in which, under some algebraic transformations, we can extend to a particle system of N sites. Finally, we construct a self-duality operator to a particle system of N sites ($N \geq 2$) and we prove the stochastic version of Noether's Theorem.

Keywords: Particle System, Matrix Lie Group, Coproduct

CONTEÚDO

1	INTRODUÇÃO	1
2	UMA ABORDAGEM SUCINTA SOBRE PROCESSOS MARKOVIANOS . . .	5
2.1	Espaço de Probabilidades	5
2.2	Variáveis Aleatórias	7
2.3	Processo de Poisson	11
2.4	Cadeias de Markov	12
3	GRUPOS CLÁSSICOS	17
3.1	Conceitos Gerais	17
3.2	Grupos topológicos	20
3.3	Grupo de Matrizes de Lie	22
4	ÁLGEBRAS	27
4.1	Conceitos Básicos	27
4.2	Álgebras de Lie	30
4.2.1	Definições Gerais	31
4.2.2	Álgebras de Lie semi-simples e a Representação Adjunta . .	34
4.2.3	A Álgebra Linear Especial de Lie $sl_2(\mathbb{R})$	35
4.2.4	O Espaço Tangente e suas propriedades algébricas	36
4.3	Biálgebras	40
5	MATRIZES ESTOCÁSTICAS	43
5.1	O grupo das Matrizes Estocásticas e seu Espaço Tangente	44
5.2	Relação entre Matrizes Estocásticas e Matrizes Geradoras	47
5.2.1	Aplicação do Teorema 5.3	49
5.3	Construindo uma Matriz Estocástica a partir do Espaço Tangente .	51
6	SISTEMAS DE PARTÍCULAS	55
6.1	Uma breve Introdução	55

6.2	Dinâmica do Sistema de Partículas	58
6.3	Propriedades e o Gerador de um Sistema de Partículas	60
6.4	Estabelecendo um sistema de partículas a partir da matriz de taxas da seção 5.3	63
6.5	Autodualidade	65
6.6	Uma Versão do Teorema de Noether para o caso estocástico	70
7	CONCLUSÃO	75
	APÊNDICE	77
A.	DEMONSTRAÇÃO DA PROPOSIÇÃO 5.1	79
1	Notação	79
2	Demonstração	81
B.	PROGRAMA PARA CÁLCULO DA ÁLGEBRA SEMI-SIMPLES DA SEÇÃO 5.3	95
	Referência Bibliográfica	99

1

INTRODUÇÃO

Ao estudarmos cadeias de Markov a tempo discreto e a tempo contínuo nos interessamos na seguinte pergunta: "Existe alguma maneira de, através de uma matriz geradora de um processo de Markov a tempo contínuo, gerarmos uma cadeia de Markov a tempo discreto?". Além do mais, será que o caminho inverso também é possível? Em face destas perguntas, tentaremos estudar relações entre estes processos por meio das matrizes estocásticas e das matrizes geradoras, que caracterizam estas cadeias de Markov.

Jeremy G. Summer em seu artigo [\[Summer\]](#) observa, ao estudar as propriedades algébricas do grupo de matrizes de Lie \mathcal{G} , dado pelas matrizes 2×2 estocásticas, e de seu espaço tangente $T_1(\mathcal{G})$, que nada mais é que uma álgebra de Lie, que é possível decompor um subgrupo de \mathcal{G} em função de uma matriz geradora Q de um processo de Markov a tempo contínuo, pela seguinte forma

$$\mathcal{G}^0 = \bigcup_{t \in \mathbb{R}} e^{Q^t} \mathcal{H},$$

com \mathcal{G}^0 dado pelas matrizes estocásticas de determinante positivo e \mathcal{H} um subgrupo normal de \mathcal{G}^0 . Desta forma, para matrizes quadradas 2×2 , obtemos uma resposta positiva para a primeira pergunta acima.

Além do mais, Frank Redig no artigo [\[Redig\]](#) apresenta um procedimento de se construir cadeias de Markov a tempo contínuo através de uma álgebra de Lie \mathfrak{g} qualquer. Este procedimento consiste em calcular o coproduto do elemento do Casimir da álgebra \mathfrak{g} , cujo resultado será uma cadeia de Markov a tempo contínuo. Sendo assim, considerando que o espaço tangente do grupo das matrizes estocásticas é uma álgebra de Lie, obtemos uma forma de construir cadeias de Markov a tempo contínuo através de uma estrutura algébrica que representa os processos markovianos a tempo discreto.

Em posse destes resultados, este trabalho visa generalizar o resultado de [\[Summer\]](#) para grupo de matrizes estocásticas de dimensão qualquer e, através do espaço tangente deste grupo, desejamos obter uma matriz geradora, calculada através do coproduto do elemento do Casimir desta álgebra.

Para encerrar, iremos, a partir da matriz geradora L' , obtida pelo coproduto do elemento do Casimir discutido acima, construir um sistema de partículas de dois sítios e cujo gerador seja dado pela própria L' . Consequentemente, em posse do gerador do sistema de duas partículas, estabeleceremos um gerador para um sistema de N -partículas, tomando como base o trabalho de [Belitsky and Schütz]. Em seguida, estudamos operadores autoduais para o sistema de partículas de N sítios, tomando como base para cálculo o autodual do sistema mais simples, de duas partículas e de gerador dado pela matriz L' . Para a construção dos autoduais nos baseamos nos trabalhos de [Daniela] e [Frank Redig and Federico Sau]. Finalmente, concluímos nosso trabalho provando a versão estocástica do Teorema de Noether. Em suma, nosso objetivo de trabalho pode ser descrito pelo diagrama da Figura 1.

Os capítulos desta dissertação distribuem-se da seguinte forma:

capítulo 2: Breve discussão sobre cadeias de Markov a tempo discreto e contínuo.

capítulo 3: Aqui descrevemos as propriedades básicas de grupos, grupos topológicos e mostramos que a componente conexa da identidade de um grupo topológico é um subgrupo normal, fechado e conexo. Além do mais, mostramos que o grupo linear geral satisfaz a condição de ser um grupo topológico. Em seguida, definimos os grupos de matrizes de Lie.

capítulo 4: Introduzimos a noção de álgebras, álgebras de Lie e espaços tangentes. Mostramos também a importante propriedade do espaço tangente ser uma álgebra de Lie.

capítulo 5: Em posse dos resultados dos capítulos anteriores, mostramos que o grupo das matrizes estocásticas é um grupo de matrizes de Lie e calculamos seu espaço tangente. Em seguida, propriedades referentes a componente conexa do grupo das matrizes estocásticas são demonstradas. Tais propriedades nos possibilitam decompor a componente conexa por meio de uma matriz geradora, utilizando a exponencial de matrizes. Esta decomposição, como desejado no **Procedimento 1** (vide Figura 1), nos possibilita gerar matrizes estocásticas através de uma matriz geradora. Em seguida, damos alguns exemplos de como gerar matrizes de transição por meio da decomposição da componente conexa do grupo das matrizes estocásticas. Para finalizar o capítulo, obtemos uma matriz geradora, como descrito no **Procedimento 2** (Figura 1), a partir de uma subálgebra semi-simples do espaço tangente do grupo das matrizes 3×3 estocásticas.

capítulo 6: Inicialmente abordamos os conceitos básicos de um sistema de partículas e como este pode ser visto como um processo de Markov a tempo contínuo. Em seguida, de acordo com o desejado no **Procedimento 3** (vide Figura 1) construiremos um sistema de duas partículas através da matriz geradora obtida no final do capítulo 5. Com o sistema de duas partículas estabelecido, estendemos este sistema para um sistema de N-partículas, o qual nos baseamos no trabalho de [Belitsky and Schütz]. Dados que a partir desta metodologia, somos capazes de criar sistemas de N-partículas, iremos estudar operadores de autodualidade para estes sistemas. Tomando como referência os trabalhos de [Daniela] e [Frank Redig and Federico Sau], estabelecemos uma função autodual para nosso sistema de duas partículas e, a partir do gerador de um sistema de N-partículas, já estabelecido, estendemos o autodual do sistema de duas para o sistema de N-partículas. Para encerrar o capítulo, enunciamos e provamos o Teorema de Noether, versão estocástica.

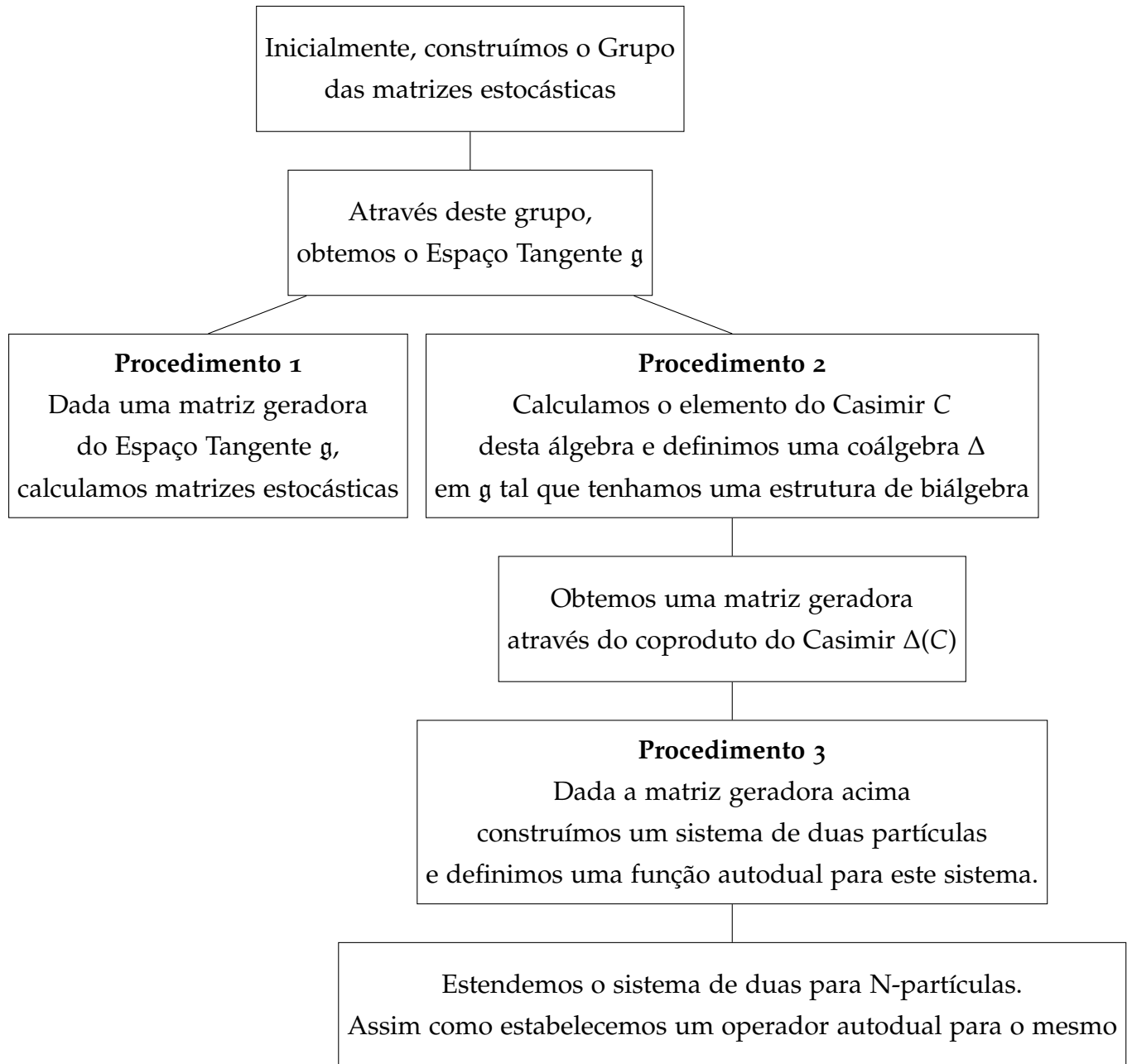


Figura 1: Objetivos do trabalho

2

UMA ABORDAGEM SUCINTA SOBRE PROCESSOS MARKOVIANOS

Neste capítulo introduziremos os alicerces básicos de probabilidade para enunciarmos o conceito de cadeias de Markov a tempo discreto e contínuo. As ideias e demonstrações apresentadas aqui podem ser vistas nos livros [Barry James], [Feller] e [Durrett1].

2.1 ESPAÇO DE PROBABILIDADES

Para estudarmos a noção da probabilidade de que certos eventos ocorram, como, por exemplo, a probabilidade de uma moeda lançada 10 vezes dar cara em todos lançamentos, devemos estabelecer certas condições para termos esta medida de probabilidade bem definida. Para isto, consideremos o conjunto Ω como o conjunto de todos os resultados possíveis de um certo experimento, em nosso exemplo teríamos $\Omega = \{\text{cara, coroa}\}$, ou, um outro exemplo, poderíamos ter $\Omega = \mathbb{R}$, com Ω representando o tempo de vida possível de uma população qualquer. (observe que o conjunto Ω pode englobar muito mais valores do que alguém esperaria na prática.)

Associado ao conjunto Ω temos a σ -álgebra \mathcal{A} (vide o texto de medida [Pedro Fernandez] ou [Carlos Isnard]) representando todos os eventos possíveis de um experimento. Utilizando ainda o exemplo do lançamento de uma moeda 10 vezes, podemos ilustrar como eventos possíveis deste experimento os eventos: “a quantidade de vezes que a moeda deu cara foi um número par” ou “o número de vezes que a moeda deu cara é maior que o número de vezes que a moeda deu coroa”... Uma σ -álgebra \mathcal{A} é um subconjunto do conjunto de partes de Ω , i.e., $\mathcal{A} \subseteq \mathcal{P}(\Omega)$, que satisfaz as condições:

(i) $\Omega, \emptyset \in \mathcal{A}$;

(ii) Dada uma sequência de eventos $(A_i)_{i \in \mathbb{N}}$, com $A_i \in \mathcal{A}$, para todo natural i . Então,

$$\bigcup_{i \in \mathbb{N}} A_i \in \mathcal{A};$$

(iii) Dado o evento $A \in \mathcal{A}$ então $\Omega \setminus A \in \mathcal{A}$.

Uma σ -álgebra importante em probabilidade é a σ -álgebra de Borel, definida da seguinte forma: seja Ω um espaço topológico e S o conjunto formado por todos os abertos deste espaço, então, a σ -álgebra gerada por S

$$\mathcal{B} = \bigcap_{\substack{\mathcal{A} \subseteq \mathcal{P}(\Omega); \\ S \subseteq \mathcal{A}}} \mathcal{A}, \quad \text{com } \mathcal{A} \text{ uma } \sigma\text{-álgebra},$$

é denotada por σ -álgebra de Borel, cujos elementos são chamados de borelianos. Segue abaixo uma caracterização, cuja demonstração pode ser vista em [Carlos Isnard], da σ -álgebra de Borel para o \mathbb{R}^n :

Proposição 2.1. *Todo intervalo da reta real é um boreliano, e o conjunto de todos intervalos limitados I em \mathbb{R}^n gera a σ -álgebra de Borel no \mathbb{R}^n .*

Seja Ω um conjunto não vazio, \mathcal{A} uma σ -álgebra de subconjuntos de Ω . Dizemos que $P : \mathcal{A} \rightarrow [0, 1]$ é uma medida de probabilidade se

(i) $P[\Omega] = 1$;

(ii) Para toda sequência enumerável de eventos dois a dois disjuntos, $(A_i)_{i \in \mathbb{N}}$, com $A_i \in \mathcal{A}$ para todo natural, tivermos

$$P[\cup_{i=1}^{\infty} A_i] = \sum_{i=1}^{\infty} P[A_i].$$

Em outras palavras, P deve ser uma medida σ -aditiva. Denotaremos a terna (Ω, \mathcal{A}, P) por espaço de probabilidade e, salvo quando não houver confusão, denotaremos nosso espaço de probabilidade por apenas Ω .

Para encerrar esta seção, definimos o conceito de eventos independentes e probabilidade condicional. Seja (Ω, \mathcal{A}, P) um espaço de probabilidade e $A, B \in \mathcal{A}$ eventos quaisquer. Dizemos que os eventos A e B são independentes quando a igualdade

$$P[A \cap B] = P[A]P[B]$$

é satisfeita. Além do mais, definimos a probabilidade condicional $P[A|B]$ como sendo a probabilidade de um evento A ocorrer dado que o evento B já ocorreu. Esta probabilidade é definida por

$$P[A|B] = \frac{P[A \cap B]}{P[B]},$$

com $P[B] \neq 0$. No caso em que tivermos $P[B] = 0$ definimos a probabilidade condicional como sendo $P[A|B] = P[A]$.

2.2 VARIÁVEIS ALEATÓRIAS

Em probabilidade é comum associar a espaços de probabilidade (Ω, \mathcal{A}, P) funções que avaliam, de alguma maneira, os resultados possíveis do experimento descrito por Ω . Por exemplo, suponhamos que certo pesquisador deseja estudar o tempo de vida de uma população de bactérias armazenadas em ensaios clínicos. Sabemos que o tempo de vida destas bactérias não será o mesmo, ou seja, existe uma aleatoriedade a ser considerada quando estudamos o tempo de vida destas bactérias. Portanto, considerando Ω como o conjunto que descreve todo tempo de vida possível das bactérias em estudo, podemos, então, associar uma função $X : \Omega \rightarrow \mathbb{R}$ que nos informa o tempo de vida das bactérias, e tal função será chamada de variável aleatória.

Mais formalmente, uma variável aleatória é uma função $X : (\Omega_1, \mathcal{A}_1) \rightarrow (\Omega_2, \mathcal{A}_2)$ entre σ -álgebras, $\mathcal{A}_1 \subseteq \mathcal{P}(\Omega_1)$ e $\mathcal{A}_2 \subseteq \mathcal{P}(\Omega_2)$, satisfazendo

$$\forall B \in \mathcal{A}_2 \quad \text{tem-se } X^{-1}(B) \in \mathcal{A}_1,$$

ou seja, são funções mensuráveis (vide [Carlos Isnard]). Comumente, as variáveis aleatórias são classificadas como discretas ou contínuas e, tais denominações, tem relação com a forma como avaliamos as probabilidades dos eventos ocorrerem para cada variável.

Uma variável aleatória X é classificada como discreta se o conjunto dos possíveis resultados de um experimento Ω for enumerável, denotado por $\{x_1, x_2, \dots\}$, e, para cada resultado possível, tivermos

$$\sum_{i=1}^{|\Omega|} P[X = x_i] = 1$$

enquanto que, uma variável aleatória é chamada de contínua quando o conjunto de resultados possíveis puder ser descrito pela reta real, caso unidimensional, ou pelo \mathbb{R}^n , caso multidimensional, e existir uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ (com $n \in \mathbb{N}$ um natural fixado), chamada de função densidade, tal que, para todo boreliano B , tenhamos

$$P[X \in B] = \int_B f dP,$$

com $[X \in B] = X^{-1}(B)$. Segue abaixo alguns exemplos das variáveis aleatórias mais comuns:

Exemplo 2.2.1. Uma variável aleatória X tem distribuição de probabilidade bernoulli de parâmetro θ , denotada por $X \sim \text{bernoulli}(\theta)$, quando os únicos valores possíveis de X são os valores 1 ou 0, com probabilidade

$$P[X = 1] = \theta \quad \text{e} \quad P[X = 0] = 1 - \theta.$$

Exemplo 2.2.2. Dizemos que a variável $X : \Omega \rightarrow \mathbb{N} \cup \{0\}$ tem distribuição Poisson de parâmetro $\theta \geq 0$, denotada por $X \sim \text{Poisson}(\theta)$, quando, para todo $i = 0, 1, 2, \dots$, tivermos

$$P[X = i] = e^{-\theta} \frac{\theta^i}{i!}.$$

Estes dois exemplos foram de variáveis aleatórias discretas. A seguir, fornecemos exemplos de variáveis contínuas:

Exemplo 2.2.3. Uma variável aleatória X tem distribuição uniforme no intervalo $[a, b]$, $X \sim U[a, b]$, com $a < b$, se X tiver função de densidade $f : \mathbb{R} \rightarrow \mathbb{R}$ valendo

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{se } x \in [a, b]; \\ 0 & \text{caso contrário.} \end{cases}$$

Exemplo 2.2.4. X tem distribuição exponencial com parâmetro $\theta > 0$, $X \sim \exp(\theta)$, se X tiver função de densidade $f : \mathbb{R} \rightarrow \mathbb{R}$ satisfazendo

$$f(x) = \begin{cases} \theta e^{-\theta x} & \text{se } x \geq 0; \\ 0 & \text{caso contrário.} \end{cases}$$

Segue da variável X ter distribuição exponencial com parâmetro θ a importante propriedade, chamada de perda de memória, que diz: a probabilidade do evento " $X > a + b$ ", com a, b constantes, dado que o evento " $X > a$ " já ocorreu é igual a probabilidade do evento " $X > b$ " ocorrer. Em outras palavras,

$$P[X > a + b | X > a] = P[X > b].$$

Esta propriedade ocorre pois

$$\begin{aligned} P[X > a + b | X > a] &= \frac{\int_{a+b}^{\infty} \theta e^{-\theta x} dx}{\int_a^{\infty} \theta e^{-\theta x} dx} \\ &= \frac{e^{-\theta(a+b)}}{e^{-\theta a}} \\ &= e^{-\theta b} \\ &= P[X > b]. \end{aligned}$$

Para encerrar nossos exemplos de variáveis contínuas, enunciamos a variável de uso mais rotineiro em aplicações:

Exemplo 2.2.5. Uma variável aleatória tem distribuição normal de parâmetros μ e σ , $X \sim N(\mu, \sigma^2)$, se sua função de densidade for

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

A função de densidade da variável normal é conhecida pelo formato de sino, como exemplificada na Figura 2.

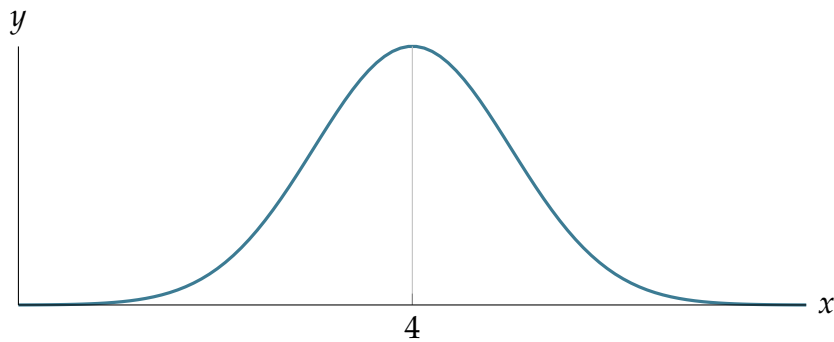


Figura 2: Função densidade de uma normal $X \sim N(4, 1)$.

Seja X uma variável aleatória discreta ou contínua, com f a função densidade no caso de variável contínua, e Ω o espaço de probabilidade. Definimos como média ou esperança de X , denotada por $E[X]$, e variância de X , denotada por $\text{Var}[X]$, as seguintes medidas resumo:

$$E[X] = \int_{\Omega} X dP = \int_{\Omega} x f(x) dx \quad \text{e} \quad \text{Var}[X] = E[(X - E[X])^2].$$

A esperança acima foi definida para uma variável aleatória contínua, no caso da variável ser discreta, a média será dada por:

$$E[X] = \sum_{i \in X(\Omega)} i P[X = i].$$

A esperança pode ser interpretada como uma medida que nos informa onde os valores em estudo se concentram, enquanto a variância calcula a dispersão dos dados ao redor da média. Outra medida muito usada é o desvio padrão, que nada mais é que a raiz quadrada da variância.

Exemplo 2.2.6. Uma variável aleatória $X \sim \text{Poisson}(\theta)$ tem média e variância dados por $E[X] = \theta$ e $\text{Var}[X] = \theta$. Isto é verdade, pois basta observar que para

$$E[X] = \sum_{i=0}^{\infty} i e^{-\theta} \frac{\theta^i}{i!}$$

temos

$$\begin{aligned} E[X] &= \sum_{i=1}^{\infty} e^{-\theta} \frac{\theta^i}{(i-1)!} \\ &= \theta \sum_{i=1}^{\infty} e^{-\theta} \frac{\theta^{(i-1)}}{(i-1)!} \\ &= \theta. \end{aligned}$$

Para calcular a variância, utilizamos a igualdade $\text{Var}[X] = E[X^2] - (E[X])^2$ (vide [Barry James]). Logo, para $E[X^2]$ teremos

$$\begin{aligned} E[X^2] &= \sum_{i=0}^{\infty} i^2 e^{-\theta} \frac{\theta^i}{i!} \\ &= \theta \sum_{i=0}^{\infty} (i+1) e^{-\theta} \frac{\theta^i}{i!} \\ &= \theta(\theta + 1). \end{aligned}$$

Desta forma,

$$\text{Var}[X] = \theta(\theta + 1) - \theta^2 = \theta.$$

Exemplo 2.2.7. Uma variável aleatória com distribuição exponencial de parâmetro θ , $X \sim \text{exp}(\theta)$, tem esperança $E[X] = 1/\theta$ e variância $\text{Var}[X] = 1/\theta^2$. Para a demonstração deste fato, vide [Barry James].

Finalmente, dizemos que duas variáveis aleatórias X_1 e X_2 , definidas no mesmo espaço de probabilidade (Ω, \mathcal{A}, P) , são independentes se, para quaisquer eventos $A_1, A_2 \in \mathcal{A}$, tivermos

$$P([X_1 \in A_1] \cap [X_2 \in A_2]) = P[X_1 \in A_1] \cdot P[X_2 \in A_2].$$

Da mesma forma, dizemos que uma sequência de variáveis aleatórias $(X_i)_{i \in \mathbb{N}}$, definidas no mesmo espaço de probabilidade, são independentes e identicamente distribuídas, denotadas por i.i.d, se as variáveis X_i forem conjuntamente independentes e possuírem mesma distribuição de probabilidade. Por exemplo, dizemos que a sequência $(X_i)_{i \in \mathbb{N}}$ é i.i.d com distribuição uniforme $U[0, 1]$ quando as variáveis forem independentes entre si e $X_i \sim U[0, 1]$, para $i \in \mathbb{N}$ um valor qualquer.

2.3 PROCESSO DE POISSON

Definiremos abaixo o conceito de um Processo de Poisson de taxa λ na reta real, e ilustraremos algumas de suas importâncias.

Definição 2.1. Dada uma sequência de variáveis aleatórias independente T_1, T_2, T_3, \dots com distribuição exponencial de parâmetro λ . Dizemos que a variável aleatória definida por

$$N(t) = \max\{n \in \mathbb{N}; T_1 + T_2 + \dots + T_n \leq t\},$$

com $t \geq 0$ e $N(0) = 0$, satisfaz o Princípio de Processo de Poisson de taxa λ , denotado por $\text{PPP}(\lambda)$.

Um Processo de Poisson pode ser visto como a contagem de incidências de um certo evento num dado intervalo de tempo. Um exemplo interessante, o qual pode ser visto com mais detalhes em [Barry James], é o processo de contagem de telefonemas recebidos numa central telefônica que, de fato, será um processo de Poisson se considerarmos as seguinte hipóteses:

- (i) A probabilidade da central telefônica receber telefonemas no intervalo $[s, s + t)$ não depender do valor s , para $t > 0$ qualquer;
- (ii) Os eventos de chamadas telefônicas ocorridas em intervalos de tempo disjuntos são independentes;
- (iii) Chamadas telefônicas não ocorrem ao mesmo tempo.

As demonstrações dos resultados abaixo podem ser vistas em [Durrett].

Proposição 2.2. Seja $N(t)$, com $t \geq 0$, um Processo de Poisson de taxa λ . Então, $N(t)$ tem como distribuição de probabilidade uma Poisson de taxa $t\lambda$, i.e.,

$$P[N(t) = i] = e^{-t\lambda} \frac{(t\lambda)^i}{i!}.$$

Em particular, a quantidade de ocorrências medidas pela variável N no intervalo $[a, b]$ é dada por uma Poisson de taxa $(b - a)\lambda$. Em outras palavras, $N[a, b] \sim \text{Poisson}((b - a)\lambda)$.

Exemplo 2.3.1. Suponhamos que um site de comércio eletrônico tenha uma quantidade de acessos por minuto modelada por um processo de Poisson de taxa 10^3 pessoas por

minuto. Dessa forma, podemos modelar a quantidade média de pessoas que acessam o site em intervalos de tempo quaisquer. Por exemplo, no período do meio-dia até às 15h temos que a quantidade média de acessos será da ordem de $3 \cdot 10^3$ pessoas.

Outro fato importante é o seguinte:

Proposição 2.3. *Seja $N(t)$, com $t \geq 0$, um Processo de Poisson de taxa λ . Então, as quantidades de ocorrências em intervalos disjuntos são dadas por variáveis independentes, ou seja, dados os valores positivos a, b , temos que a variável de contagem $N(a + b) - N(a)$ é independente da variável $N(c)$, para todo $c \leq a$.*

2.4 CADEIAS DE MARKOV

Seja Ω um conjunto enumerável, o qual chamaremos de espaço de estados, $(U_n)_{n \in \mathbb{N}}$ uma sequência i.i.d com $U_1 \sim U[0, 1]$, e $F : \Omega \times [0, 1] \rightarrow \Omega$ uma função. Dizemos que a sequência de variáveis aleatórias $(X_n)_{n \in \mathbb{N}}$, com espaço de estados Ω (i.e., a variável aleatória X_n , para qualquer natural n , toma valores no espaço Ω), é uma Cadeia de Markov a tempo discreto se a sequência $(X_n)_{n \geq 0}$ satisfazer

$$X_n = F(X_{n-1}, U_n), \quad \text{para } n \geq 1,$$

com ponto inicial X_0 . O ponto inicial do processo markoviano pode ser um valor fixado inicialmente ou pode ser dado por uma distribuição de probabilidade.

Decorre, da definição de cadeias de Markov, a seguinte caracterização dos processos markovianos a tempo discreto:

Proposição 2.4. *Dado Ω um conjunto enumerável e $(X_n)_{n \in \mathbb{N}}$ uma sequência de variáveis aleatórias de espaço de estados Ω . Então, $(X_n)_{n \in \mathbb{N}}$ é uma Cadeia de Markov a tempo discreto se, e somente se, para qualquer natural n e para quaisquer valores a_0, a_1, \dots, a_n pertencentes ao espaço de estados, tivermos*

$$P[X_n = a_n | X_{n-1} = a_{n-1}, X_{n-2} = a_{n-2}, \dots, X_0 = a_0] = P[X_n = a_n | X_{n-1} = a_{n-1}] = p(a_{n-1}, a_n),$$

com $p : \Omega \times \Omega \rightarrow [0, 1]$ a matriz que satisfaz $\sum_{b \in \Omega} p(a, b) = 1$, para todo $a \in \Omega$. Dizemos, também, que toda matriz cujas linhas somam um é uma matriz de transição.

Desta caracterização segue o seguinte exemplo

Exemplo 2.4.1. Um centro meteorológico, situado em São Paulo, avalia a probabilidade dos seguintes eventos ocorrerem: o dia iniciar chuvoso, ensolarado ou com forte neblina. Denotando por 0, 1, 2 os eventos acima, respectivamente, o centro meteorológico conclui que a condição climática no início de cada dia pode ser descrita por uma cadeia de Markov $(X_n)_{n \in \mathbb{N}}$ com espaço de estados $\Omega = \{0, 1, 2\}$ e matriz de transição dada por

$$p = \begin{matrix} & \begin{matrix} 0 & 1 & 2 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \end{matrix} & \begin{pmatrix} 0,6 & 0,3 & 0,1 \\ 0,05 & 0,7 & 0,25 \\ 0,3 & 0,3 & 0,4 \end{pmatrix} \end{matrix}.$$

Esta cadeia de Markov pode ser visualizada pelo grafo da Figura 3. Além do mais,

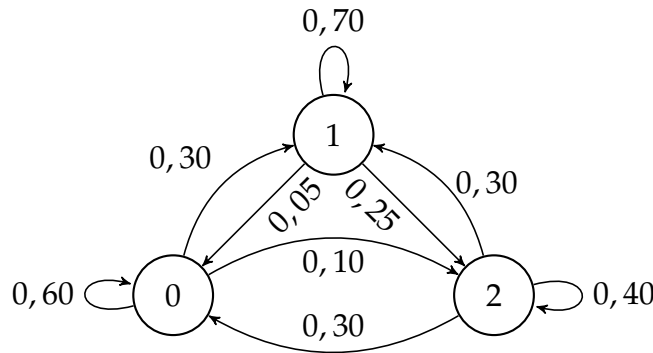


Figura 3: Grafo da matriz de transição p da Cadeia de Markov a tempo discreto com espaço de estados $\Omega = \{0, 1, 2\}$, descrita no Exemplo 2.4.1.

se considerarmos que os eventos “chuvoso”, “ensolarado” e “com forte neblina”, tem probabilidade de ocorrer, no dia seguinte, dada por

$$\pi(0) = 0,5, \quad \pi(1) = 0,3 \quad \text{e} \quad \pi(2) = 0,2,$$

temos que a probabilidade de, daqui a 10 dias, o dia amanhecer chuvoso é igual a

$$P[X_{10} = 0] = \pi(0)p^{10}(0,0) + \pi(1)p^{10}(1,0) + \pi(2)p^{10}(2,0) = 0,25005.$$

Dado uma cadeia de Markov $(X_n)_{n \in \mathbb{N}}$ associada ao espaço de estados Ω e com matriz de transição p . Dizemos que uma distribuição de probabilidade $\pi : \Omega \rightarrow [0, 1]$ é invariante se

$$\sum_{b \in \Omega} \pi(b)P[X_1 = a | X_0 = b] = \pi(a), \text{ para todo } a \in \Omega.$$

Em outras palavras, uma medida de probabilidade sobre o espaço de estados Ω é invariante se, para qualquer $n \in \mathbb{N}$, e para qualquer $a \in \Omega$, tivermos

$$P[X_n = a] = \pi(a),$$

ou seja, a probabilidade de estar num estado qualquer de seu grafo é um invariante em relação a qualquer instante de tempo discreto $n \in \mathbb{N}$.

Exemplo 2.4.2. Consideremos a cadeia de Markov do Exemplo 2.4.1. Então, uma medida invariante π para esta cadeia deve satisfazer

$$\sum_{b=0}^2 \pi(b)P[X_1 = a|X_0 = b] = \pi(a), \forall a \in \Omega = \{0, 1, 2\}.$$

Logo, a igualdade acima é equivalente a

$$[\pi(0) \ \pi(1) \ \pi(2)] \begin{pmatrix} 0,6 & 0,3 & 0,1 \\ 0,05 & 0,7 & 0,25 \\ 0,3 & 0,3 & 0,4 \end{pmatrix} = [\pi(0) \ \pi(1) \ \pi(2)].$$

Resulta, da igualdade acima, que a medida invariante para o processo markoviano em questão será dada por:

$$\pi(0) = 1/4 \quad \pi(1) = 1/2 \quad \pi(2) = 1/4.$$

Abaixo, estabelecemos um critério o qual auxilia no cálculo de medidas invariantes para cadeias de Markov a tempo discreto.

Definição 2.2. Dada uma cadeia de Markov $(X_n)_{n \in \mathbb{N}}$ com espaço de estados Ω . Dizemos que uma medida de probabilidade π sobre Ω é reversível se, para todo elemento $a, b \in \Omega$, tivermos

$$\pi(a)P[X_1 = b|X_0 = a] = \pi(b)P[X_1 = a|X_0 = b],$$

como ilustrado na Figura 4.

Decorre de uma probabilidade π ser reversível para a cadeia $(X_n)_n$ que a mesma, também, será uma medida invariante, pois, para todos estados $a, b \in \Omega$, temos

$$\pi(a)P[X_1 = b|X_0 = a] = \pi(b)P[X_1 = a|X_0 = b]$$

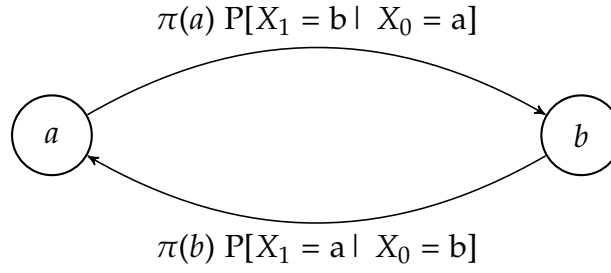


Figura 4: Exemplo de transição entre estados a e b tal que a probabilidade de se iniciar no ponto a e ir para o ponto b é igual ao caminho reverso, i.e., a probabilidade de começar no estado b e ir para o estado a .

onde, tomando a somatória em relação aos estados $b \in \Omega$ em ambos os lados da igualdade, obtemos

$$\pi(a) \sum_{b \in \Omega} P[X_1 = b | X_0 = a] = \sum_{b \in \Omega} \pi(b) P[X_1 = a | X_0 = b],$$

e, como $\sum_{b \in \Omega} P[X_1 = b | X_0 = a] = 1$, temos a igualdade desejada. Portanto, π é uma medida invariante.

Uma abordagem natural a cadeias de Markov a tempo discreto é estender o tempo em que as transições de estados ocorrem para algo contínuo. Sob esta perspectiva, definimos os processos markovianos a tempo contínuo:

Definição 2.3. Seja Ω o conjunto representando os espaços de estado e $(X_t, t \geq 0)$ uma sequência de variáveis aleatórias, indexadas pelos valores reais e positivos $t \geq 0$, assumindo valores em Ω . Dizemos que $(X_t, t \geq 0)$ é uma cadeia de Markov a tempo contínuo se, para toda sequência de tempos positivos $s_1 < s_2 < \dots < s_n$, tivermos

$$P[X_{t+s_n} = a | X_{s_1} = a_1, \dots, X_{s_{n-1}} = a_{n-1}, X_{s_n} = b] = P[X_t = a_t | X_0 = b],$$

para $a, b, a_1, a_2, \dots, a_{n-1} \in \Omega$, e $t \geq 0$.

Decorre desta definição o seguinte exemplo.

Exemplo 2.4.3. Consideremos o processo de Poisson de taxa 1, dado por $N(t)$ com $t \geq 0$, e a cadeia de Markov a tempo discreto X_n , dada no Exemplo 2.4.1. Logo, seja $t \geq 0$ um instante de tempo, definimos o processo

$$Y_t = X_{N(t)}.$$

Tal processo será uma cadeia de Markov a tempo contínuo o qual espera tempos aleatórios de transição, dados por exponenciais de taxa 1, para mudar de estados. Além do mais, as transições são dadas por processos de Markov a tempo discreto.

Dada uma cadeia de Markov a tempo contínuo $(X_t, t \geq 0)$ estaremos interessados em estudar as seguintes taxas

$$\lim_{s \rightarrow 0} \frac{P[X_{t+s} = j | X_t = i]}{s},$$

para i, j elementos do espaço de estados do processo. Caso o limite acima exista, denotaremos, então, este limite por taxa de transição do estado i para o estado j , denotado por $c(i, j)$, i.e.,

$$c(i, j) = \lim_{s \rightarrow 0} \frac{P[X_{t+s} = j | X_t = i]}{s}.$$

Em posse destas taxas de transição definimos a seguinte matriz

$$Q = (Q_{ij})_{i,j \in \Omega}, \text{ com } Q_{ij} = \begin{cases} c(i, j) & \text{se } i \neq j; \\ -\sum_{l \neq i} c(i, l) & \text{c.c.,} \end{cases}$$

com Ω o espaço de estados. A matriz Q , definida acima, será chamada de matriz geradora ou L-matriz.

Observação. A matriz geradora Q satisfaz a condição de que todas linhas somam zero.

Em posse de uma L-matriz qualquer somos capazes de criar um processo de Markov a tempo contínuo. Para a construção de processos markovianos a tempo contínuo, a partir de uma matriz geradora, consulte [Durrett1], capítulo 4.

Um resultado importante para cadeias de Markov a tempo contínuo são as equações de Kolmogorov. Considerando o processo de Markov $(X_t, t \geq 0)$, com espaço de estados Ω , definimos para $i, j \in \Omega$ a função:

$$p_t(i, j) = P[X_t = j | X_0 = i].$$

Por meio desta função $p_t(i, j)$ e da matriz geradora Q , associada ao processo markoviano, valem as seguintes igualdades:

$$p'(t) = Qp_t; \quad (\text{Equação Regressiva de Kolmogorov})$$

$$p'(t) = p_t Q. \quad (\text{Equação Progressiva de Kolmogorov})$$

Para as demonstrações destas igualdades veja [Durrett1].

3

GRUPOS CLÁSSICOS

Neste capítulo formularemos o conceito algébrico de grupos e algumas de suas propriedades. Em seguida, as estruturas de grupos topológicos e de grupo de matrizes de Lie serão discutidos. A importância de se estudar os grupos topológicos manifesta-se pelo fato do grupo linear geral ter posse dessa estrutura e, sendo assim, todo grupo de matrizes de Lie também será um grupo topológico. O principal resultado deste capítulo, dado pelo Teorema 3.3, nos fornecerá informação sobre a componente conexa do elemento identidade de um grupo topológico. Graças a este resultado, sabemos que a componente conexa da identidade será um subgrupo normal, fechado e conexo do grupo em questão. As demonstrações dos Teoremas, assim como os conceitos propostos neste capítulo, são baseados nas seguintes referências: [Stillwell] (para o estudo de grupos de matrizes de Lie), [John L. Kelley] (para o tratamento de grupos topológicos) e [Paulo A. Martin] (capítulo 1, noções básicas de grupos).

3.1 CONCEITOS GERAIS

Seja G um conjunto não nulo munido de um operador $\cdot : G \times G \rightarrow G$. Dizemos que a tupla (G, \cdot) é um grupo se as seguintes condições forem satisfeitas:

- (i) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ para todo $a, b, c \in G$ (*associatividade*);
- (ii) Existe um elemento $e \in G$ tal que, para todo $a \in G$, vale $a \cdot e = e \cdot a = a$ (*existência de identidade*);
- (iii) Para todo $a \in G$ existe um elemento $a^{-1} \in G$ tal que $a \cdot a^{-1} = a^{-1} \cdot a = e$ (*existência do elemento inverso*).

Daqui em diante, sempre quando não houver risco de confusão, denotaremos o grupo (G, \cdot) por apenas o grupo G .

Exemplo 3.1.1. (*O Grupo Linear Geral*) Um exemplo importante, o qual iremos trabalhar bastante no decorrer deste texto, é o do grupo linear geral, denotado por $GL(n, \mathbb{R})$. Este grupo é definido pelo conjunto

$$GL(n, \mathbb{R}) = \{A = (a_{ij})_{n \times n}; a_{ij} \in \mathbb{R} \text{ e } \det(A) \neq 0\}$$

munido da multiplicação usual entre matrizes.

Definição 3.1. Dado o grupo G e um subconjunto não nulo $H \subseteq G$. Dizemos que H é um subgrupo de G , denotado por $H \leq G$, se:

- (i) O elemento identidade do grupo G estiver contido em H ;
- (ii) Para todo $a \in H$ tivermos $a^{-1} \in H$;
- (iii) Dados dois elementos quaisquer de H , digamos $a, b \in H$, tivermos $a \cdot b \in H$.

Abaixo, alguns exemplos de subgrupos:

Exemplo 3.1.2. Seja G um grupo, dois exemplos imediatos decorrem da definição de subgrupo, são eles: $\{e\}$ e o próprio grupo G . Estes subgrupos são chamados de subgrupos triviais.

Exemplo 3.1.3. Seja o grupo $(\mathbb{Z}, +)$ dos inteiros munido da soma usual. O conjunto

$$H = \{\dots, -6, -4, -2, 0, 2, 4, 6, \dots\}$$

munido da soma usual dos inteiros é um subgrupo de \mathbb{Z} .

Exemplo 3.1.4. Seja G um grupo qualquer. Denotamos por centro do grupo G o conjunto

$$Z(G) = \{x \in G; \forall y \in G, xy = yx\}.$$

Decorre desta definição que $Z(G)$ é um subgrupo de G . Para que $Z(G)$ seja um subgrupo de G devemos checar apenas a propriedade (ii), pois as propriedades (i) e (iii) são imediatas. Portanto, dados os elementos $x \in Z(G)$ e $y \in G$ quaisquer, temos:

$$(xy)(xy)^{-1} = e \implies (yx)(y^{-1}x^{-1}) = e \implies xy^{-1} = y^{-1}x.$$

Logo, $Z(G)$ é um subgrupo de G .

Dado um subgrupo $H \leq G$ de um grupo G qualquer e os elementos $x, y \in G$. Denotaremos por xH , Hy e xHy como sendo os seguintes conjuntos:

$$\begin{aligned} xH &:= \{x \cdot a; a \in H\}; \\ Hy &:= \{a \cdot y; a \in H\}; \\ xHy &:= \{x \cdot a \cdot y; a \in H\}. \end{aligned}$$

Em função destas operações, temos a seguinte definição:

Definição 3.2. Seja G um grupo e H um subgrupo de G . Dizemos que H é um subgrupo normal de G , denotado por $H \triangleleft G$, se, para todo $x \in G$, tivermos a inclusão

$$xHx^{-1} \subseteq H.$$

Observe que a inclusão acima é equivalente a dizer que $xHx^{-1} = H$.

Exemplo 3.1.5. Os subgrupos triviais de um grupo qualquer são exemplos imediatos de subgrupos normais.

Exemplo 3.1.6. Dado um grupo G qualquer, o centro $Z(G)$ é um subgrupo normal de G , pois, dados $x \in G$ e $y \in Z(G)$ elementos quaisquer, temos

$$xyx^{-1} = xx^{-1}y \in Z(G)$$

e, portanto, $xZ(G)x^{-1} \subseteq Z(G)$.

Definição 3.3. Seja G um grupo qualquer e $H \leq G$ um subgrupo de G . Definimos, então, por o conjunto das coclasses de H em G como sendo o conjunto

$$G/H = \{Hx; x \in G\}.$$

Ao estudar as coclasses, ou classes laterais, dadas pelo subgrupo $H \leq G$, obtemos a seguinte decomposição do grupo G :

$$G = \bigcup_{x \in G} Hx,$$

que nada mais é que a união de todos elementos do conjunto G/H . Além do mais, se H for um subgrupo normal de G teremos que G/H terá estrutura de grupo se considerarmos o seguinte produto

$$\begin{aligned} \odot : G/H \times G/H &\rightarrow G/H \\ (Hx, Hy) &\mapsto H(xy). \end{aligned}$$

Para encerrar esta seção, enunciaremos o significado de uma função entre grupos ser um homomorfismo e, decorrente desta definição, enunciaremos o Teorema do Isomorfismo para grupos.

Definição 3.4. Seja $f : (G, \cdot) \rightarrow (G', \odot)$ uma função entre grupos. Então, dizemos que f é um homomorfismo de grupos se, para todo $x, y \in G$, tivermos

$$f(x \cdot y) = f(x) \odot f(y).$$

Além do mais, se f for bijetora, dizemos que f é um isomorfismo de grupos, denotados por $G \cong G'$.

Decorre desta definição que o núcleo de todo homomorfismo $f : G \rightarrow G'$ entre grupos, denotado por $\ker(f)$, é um subgrupo normal de G . Em posse destas definições, e do fato do núcleo de um homomorfismo ser um subgrupo normal, enunciamos o seguinte Teorema:

Teorema 3.1. Dado um homomorfismo entre grupos $f : G \rightarrow G'$. Então, o seguinte diagrama

$$\begin{array}{ccc} G & \xrightarrow{f} & \text{Im}(f) \\ \downarrow \pi & \nearrow \tilde{f} & \\ G/\ker(f) & & \end{array}$$

comuta, onde $\pi(x) = \ker(f)x$ e $\tilde{f}(\ker(f)x) = f(x)$, para todo $x \in G$. Caso f seja um isomorfismo teremos, então, que $G' \cong G/\ker f$.

Demonstração. Vide [Paulo A. Martin]. □

3.2 GRUPOS TOPOLÓGICOS

Seja G um conjunto não nulo e \mathcal{F} uma topologia de G . Dizemos que a terna (G, \mathcal{F}, \cdot) é um grupo topológico se:

- (i) As tuplas (G, \mathcal{F}) e (G, \cdot) forem um espaço topológico¹ e um grupo, respectivamente;

¹ Vide [Elon1] ou [Elon2] para o estudo mais aprofundado de temas relacionados a topologia ou espaços métricos.

- (ii) As funções $m : G \times G \mapsto G$ e $i : G \mapsto G$ dadas por $m(x, y) = x \cdot y$ e $i(x) = x^{-1}$ forem funções contínuas.

Observe que, fixado um ponto $a \in G$ qualquer, as funções de multiplicação à esquerda e à direita por a , denotadas por L_a e R_a , respectivamente,

$$\begin{array}{ll} L_a : G \rightarrow G & R_a : G \rightarrow G \\ x \mapsto a \cdot x & x \mapsto x \cdot a \end{array}$$

são homeomorfismo de espaços topológicos.

Exemplo 3.2.1. O grupo linear geral é um grupo topológico. Como já sabemos, vide [Elon1], o conjunto das matrizes $\mathcal{M}_n(\mathbb{R})$ tem estrutura de espaço topológico e, portanto, o grupo linear geral é um espaço topológico com topologia induzida por $\mathcal{M}_n(\mathbb{R})$. Finalmente, como as funções

$$\begin{array}{ll} m : GL(n, \mathbb{R}) \times GL(n, \mathbb{R}) \rightarrow GL(n, \mathbb{R}) & e \quad i : GL(n, \mathbb{R}) \rightarrow GL(n, \mathbb{R}) \\ (A, B) \mapsto A \cdot B & A \mapsto A^{-1} \end{array}$$

são contínuas, obtemos que $GL(n, \mathbb{R})$ é, de fato, um grupo topológico.

Abaixo segue o Teorema que nos permitirá avaliar as componentes conexas de grupos topológicos como subgrupos normais, fechados e conexos. Porém, antes de enunciarmos o Teorema, proporemos um Lema, cuja demonstração pode ser vista em [Elon1], que nos auxiliará na prova do Teorema 3.3.

Lema 3.2. *Dado uma função $f : (X, \mathcal{F}_1) \rightarrow (Y, \mathcal{F}_2)$ entre espaços topológicos e C_a a componente conexa do ponto $a \in X$. Se a função f for um homeomorfismo teremos, então, que a imagem $f(C_a)$ será a componente conexa do ponto $f(a)$ em Y .*

Teorema 3.3. *Dado um grupo topológico (G, \mathcal{F}, \cdot) e, denotemos por \mathbf{e} , o elemento identidade do grupo. Então, a componente conexa da identidade, denotada por G^0 , será um subgrupo normal, fechado e conexo de G .*

Demonstração. Sabemos que G^0 é um espaço conexo e fechado em G . Resta provarmos que G^0 é um subgrupo normal. Para provarmos que G^0 é um subgrupo, devemos checar que

P1) $\mathbf{e} \in G^0$;

P2) $x \in G^0$ implica $x^{-1} \in G^0$;

P3) $x, y \in G^0$ implica $x \cdot y \in G^0$.

A propriedade P1) é imediata. Para todo $x \in G^0$, tem-se que, como já observado, o operador $L_{x^{-1}} : G \rightarrow G$ com $L_{x^{-1}}(y) = x^{-1} \cdot y$ é uma função contínua. Logo, a imagem $L_{x^{-1}}(G^0)$ é um conexo contendo a identidade, pois $L_{x^{-1}}(x) = \mathbf{e}$. Portanto,

$$L_{x^{-1}}(G^0) \subseteq G^0$$

e, como $L_{x^{-1}}(\mathbf{e}) \in L_{x^{-1}}(G^0)$, obtemos que $x^{-1} \in G^0$, mostrando que P2) é satisfeita. Para a propriedade P3), observe que, para todo $x, y \in G^0$, temos $x^{-1} \in G^0$, pela propriedade P2). Então, concluímos que $\mathbf{e} \in L_x(G^0)$ e, por $L_x(G^0)$ ser um conexo contendo a identidade, obtemos

$$\mathbf{e} \in L_x(G^0) \subseteq G^0,$$

ou seja, $L_x(y) \in L_x(G^0)$. Sendo assim, $x \cdot y \in G^0$ e a propriedade P3) está provada.

Finalmente, provemos que G^0 é um subgrupo normal de G . Dado $x \in G$, um elemento qualquer, temos que os operadores de multiplicação à esquerda, L_x , e à direita, R_x , são funções homeomorfas entre G . Pelo Lema 3.2, obtemos, por G^0 ser a componente conexa de \mathbf{e} , que

$$xG^0x^{-1} = R_{x^{-1}} \circ L_x(G^0) = R_{x^{-1}}(C_{x \cdot \mathbf{e}}) = C_{x \cdot \mathbf{e} \cdot x^{-1}},$$

i.e., $xG^0x^{-1} = G^0$ e, a partir daí, concluímos que G^0 é um subgrupo normal de G . \square

3.3 GRUPO DE MATRIZES DE LIE

Iniciamos esta seção estabelecendo a definição de grupos de Lie:

Definição 3.5 (Grupos de Lie). Um grupo de Lie é uma variedade diferenciável que admite uma estrutura de grupo onde as operações de multiplicação e inversão são deriváveis.

Exemplo 3.3.1. O grupo $GL(n, \mathbb{R})$ é um grupo de Lie, visto que a multiplicação (resp. inversão) de matrizes é contínua.

Em posse da estrutura de grupos de Lie, estudaremos os grupos de matrizes de Lie. Dizemos que um subgrupo \mathcal{H} do grupo topológico $GL(n, \mathbb{R})$ é um grupo de matrizes de Lie se \mathcal{H} for um subconjunto fechado. Desta forma, pelo Teorema 3.3, concluímos que a componente conexa da identidade é um grupo de matrizes de Lie.

Abaixo segue alguns exemplos clássicos de grupo de matrizes de Lie.

Exemplo 3.3.2. Dado o grupo linear geral $GL(n, \mathbb{R})$. O subgrupo

$$SL(n, \mathbb{R}) := \{A \in GL(n, \mathbb{R}); \det(A) = 1\},$$

denominado por grupo linear especial, é um grupo de matrizes de Lie. Isto é verdade pois $SL(n, \mathbb{R})$ é a imagem inversa de um fechado, o conjunto $\{1\}$, por uma função contínua, o determinante.

Exemplo 3.3.3. Dado o grupo linear geral $GL(n, \mathbb{R})$, temos que o conjunto

$$O(n, \mathbb{R}) = \{A \in GL(n, \mathbb{R}); AA^t = A^t A = I_n\}$$

munido da multiplicação usual entre matrizes é um grupo de matrizes Lie, denotado por grupo ortogonal.

Exemplo 3.3.4. Seguindo o Exemplo acima, dos grupos ortogonais, temos que o grupo dado por

$$SO(n, \mathbb{R}) = \{A \in GL(n, \mathbb{R}); AA^t = A^t A = I_n \text{ e } \det(A) = 1\}$$

é um grupo de matrizes de Lie, chamado de grupo ortogonal especial ou de grupo das rotações.

Veremos, a seguir, que com o auxílio do operador exponencial de matrizes podemos estabelecer relações com os grupos de matrizes de Lie e seu espaço tangente (vide a subseção 4.2.4 para a definição de espaço tangente). Para isso, iniciamos esta discussão definindo a exponencial de matrizes:

Definição 3.6. A função exponencial de uma matriz quadrada $A \in \mathcal{M}_n(\mathbb{R})$ é dada por

$$\exp : \mathcal{M}_n(\mathbb{R}) \rightarrow \mathcal{M}_n(\mathbb{R})$$

$$A \mapsto \sum_{i=0}^{\infty} \frac{A^i}{i!},$$

com $A^0 = I_n$.

Observação. Note que a série acima é absolutamente convergente pois, como para todo $A \in \mathcal{M}_n(\mathbb{R})$ vale $\|A\| < \infty$, temos que a norma da série é dada pela seguinte desigualdade

$$\begin{aligned} \left\| \sum_{i=0}^{\infty} \frac{A^i}{i!} \right\| &\leq \sum_{i=0}^{\infty} \frac{\|A^i\|}{i!} \\ &= \sum_{i=0}^{\infty} \frac{\|A\|^i}{i!} \\ &= e^{\|A\|}. \end{aligned}$$

Decorre da série ser absolutamente convergente que a função exponencial está bem definida. Dada uma matriz $A \in \mathcal{M}_n(\mathbb{R})$ qualquer, denotaremos, para simplificar a notação, por e^A como sendo a matriz $\exp(A)$, i.e., $e^A := \exp(A)$.

Proposição 3.4. *Dadas as matrizes $A, B \in \mathcal{M}_n(\mathbb{R})$ quaisquer e λ um autovalor da matriz A , com $v \in \mathbb{R}^n$ seu autovetor (i.e., $Av = \lambda v$). Logo, as seguintes propriedades para a exponencial são válidas:*

$$P_1) \quad e^0 = I_n;$$

$$P_2) \quad \text{Se } AB = BA, \text{ então } e^{A+B} = e^A e^B;$$

$$P_3) \quad e^{BAB^{-1}} = B e^A B^{-1};$$

$$P_4) \quad \det(e^A) = e^{\text{tr}(A)};$$

$$P_5) \quad e^{tA}v = e^{t\lambda}v, \text{ para } t \text{ um real qualquer.}$$

Demonstração. vide [Stillwell] □

A partir da função exponencial podemos relacionar grupos de matrizes de Lie e seus espaços tangentes. Estas relações são dadas pelos seguintes teoremas (cuja demonstrações podem ser encontradas em [Stillwell, capítulo 7, p. 143 e 149]):

Teorema 3.5. *Seja G um grupo de matrizes de Lie e $T_1(G)$ seu espaço tangente. Então, a imagem do espaço tangente pela função exponencial é um subconjunto de G . Isto é:*

$$\exp(T_1(G)) \subseteq G.$$

Teorema 3.6. *Seja G um grupo de matrizes de Lie e $T_1(G)$ seu espaço tangente. Então, existe um número real, $r > 0$, suficientemente pequeno e um aberto U de $T_1(G)$ contendo a matriz $\mathbf{0} \in T_1(G)$ tal que*

$$\exp : U \subseteq T_1(G) \rightarrow B(I_n, r) \subseteq G$$

é uma bijeção.

4

ÁLGEBRAS

Introduziremos, neste capítulo, os conceitos necessários sobre álgebras para o nosso trabalho, tal qual, mostraremos como as propriedades de associatividade e existência de unidade em álgebras, descritas com o auxílio do produto tensorial, podem nos proporcionar uma forma de expressar estruturas algébricas denotadas por coálgebras, cuja importância, para nós, se dará pelo fato do coproduto do elemento do Casimir de uma certa álgebra de Lie (mais precisamente, do espaço tangente relativo ao grupo de matrizes de Lie dado pelas matrizes estocásticas) ser uma matriz que, por uma perspectiva de processos markovianos, será uma matriz geradora de uma cadeia de Markov a tempo contínuo. Abordaremos, também, a relação existente entre álgebras de Lie com os espaços tangentes de grupos de matrizes de Lie. Este capítulo segue de perto os textos de [San Martin], para o estudo das álgebras de Lie, as notas de aula de [Ken Brown], para a abordagem de coálgebras e biálgebras, assim como o texto de [Stillwell] para a demonstração do fato de que o espaço tangente do grupo de matrizes de Lie é uma álgebra de Lie.

4.1 CONCEITOS BÁSICOS

Um \mathbb{K} -espaço vetorial \mathcal{A} (\mathbb{K} denotando um corpo) munido de uma multiplicação $M : \mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A}$ é dito ser uma álgebra se, para todo $x, y, z \in \mathcal{A}$ e $\lambda, \beta \in \mathbb{K}$, as seguintes propriedades forem satisfeitas:

- (i) $M(x, y + z) = M(x, y) + M(x, z)$ (*distributiva à direita*);
- (ii) $M(x + y, z) = M(x, z) + M(y, z)$ (*distributiva à esquerda*);
- (iii) $M(\lambda x, \beta y) = \lambda \beta M(x, y)$.

Caso o operador de multiplicação satisfaça a propriedade

$$M(M(x, y), z) = M(x, M(y, z)), \forall x, y, z \in \mathcal{A} \quad (\text{associatividade})$$

dizemos que a álgebra em questão é associativa. Outra propriedade importante que algumas álgebras podem apresentar é a existência de unidade, i.e., existe um elemento $1_{\mathcal{A}} \in \mathcal{A}$ tal que, para todo $x \in \mathcal{A}$, vale a seguinte igualdade

$$M(x, 1_{\mathcal{A}}) = M(1_{\mathcal{A}}, x) = x.$$

Quando uma álgebra \mathcal{A} possuir identidade dizemos que \mathcal{A} é uma álgebra com unidade. Mais a frente, quando estivermos estudando coálgebras e biálgebras, estaremos apenas interessados em álgebras associativas com unidade.

Observe que o operador multiplicação nada mais é que um operador bilinear. Desta forma, pela propriedade universal do produto tensorial (vide [Mac Lane and Birkhoff]), existe uma aplicação linear $\tilde{M} : \mathcal{A} \otimes \mathcal{A} \rightarrow \mathcal{A}$ tal que o diagrama

$$\begin{array}{ccc} \mathcal{A} \times \mathcal{A} & \xrightarrow{\otimes} & \mathcal{A} \otimes \mathcal{A} \\ & \searrow M & \downarrow \tilde{M} \\ & & \mathcal{A} \end{array}$$

comuta. Além do mais, podemos definir uma aplicação linear $u : \mathbb{K} \rightarrow \mathcal{A}$ tal que $u(1_{\mathbb{K}}) = 1_{\mathcal{A}}$, ou seja, a função u irá satisfazer, para todo $x \in \mathcal{A}$ e $\lambda \in \mathbb{K}$, a igualdade

$$M(u(\lambda), x) = M(x, u(\lambda)).$$

Em outras palavras, uma álgebra associativa com unidade pode ser vista como uma terna (\mathcal{A}, M, u) o qual, \mathcal{A} é um \mathbb{K} -espaço vetorial e, $M : \mathcal{A} \otimes \mathcal{A} \rightarrow \mathcal{A}$ e $u : \mathbb{K} \rightarrow \mathcal{A}$ são aplicações lineares cujos diagramas

$$\begin{array}{ccc} \mathcal{A} \otimes \mathcal{A} \otimes \mathcal{A} & \xrightarrow{Id \otimes M} & \mathcal{A} \otimes \mathcal{A} \\ M \otimes Id \downarrow & & \downarrow M \\ \mathcal{A} \otimes \mathcal{A} & \xrightarrow{M} & \mathcal{A} \end{array}$$

Diagrama 1, associatividade.

$$\begin{array}{ccccc} & & \mathcal{A} \otimes \mathcal{A} & & \\ u \otimes Id \nearrow & & \downarrow M & & \nwarrow Id \otimes u \\ \mathbb{K} \otimes \mathcal{A} & & \mathcal{A} & & \mathcal{A} \otimes \mathbb{K} \\ s_1 \searrow & & & & \swarrow s_2 \end{array}$$

Diagrama 2, existência da unidade

comutam. Acima, temos que as funções Id, s_1 e s_2 representam

$$\begin{aligned} Id : \mathcal{A} &\rightarrow \mathcal{A} \\ x &\mapsto x \end{aligned}$$

$$\begin{aligned} s_1 : \mathbb{K} \otimes \mathcal{A} &\rightarrow \mathcal{A} \\ (\lambda, x) &\mapsto \lambda x \end{aligned}$$

$$\begin{aligned} s_2 : \mathcal{A} \otimes \mathbb{K} &\rightarrow \mathcal{A} \\ (x, \lambda) &\mapsto \lambda x. \end{aligned}$$

Observação. Com relação aos diagramas acima podemos notar que

- (i) O Diagrama 1 comutar significa que a multiplicação é associativa pois, dados os elementos $x, y, z \in \mathcal{A}$, temos

$$\begin{aligned} M \circ (Id \otimes M)(x \otimes y \otimes z) &= M(Id(x) \otimes M(y, z)) \\ &= M(x \otimes M(y, z)) \end{aligned}$$

e

$$\begin{aligned} M \circ (M \otimes Id)(x \otimes y \otimes z) &= M(M(x, y) \otimes Id(z)) \\ &= M(M(x, y) \otimes z) \end{aligned}$$

Logo, pela igualdade $M \circ (Id \otimes M) = M \circ (M \otimes Id)$, concluímos que $M(M(x, y) \otimes z) = M(x \otimes M(y \otimes z))$, ou seja, M é associativa.

- (ii) O Diagrama 2 representa a existência de identidade em \mathcal{A} . Esta afirmação é válida se observarmos que $\mathbb{K} \otimes \mathcal{A} \cong \mathcal{A} \otimes \mathbb{K}$ e, portanto, para todo $x \in \mathcal{A}$ vale

$$M \circ (u \otimes Id)(1 \otimes x) = M \circ (Id \otimes u)(x \otimes 1) = x$$

ou seja,

$$M(u(1_{\mathbb{K}}) \otimes x) = M(x \otimes u(1_{\mathbb{K}})) = x.$$

Desta forma, $u(1_{\mathbb{K}})$ é o elemento unitário da álgebra.

Exemplo 4.1.1. (*A álgebra tensorial*) Seja V um \mathbb{R} -espaço vetorial. Consideremos o seguinte conjunto:

$$T(V) := \bigoplus_{k \geq 0} V^{\otimes k},$$

com $V^{\otimes k} = V \otimes \cdots \otimes V$, k -vezes. O conjunto $T(V)$ será um \mathbb{R} -espaço vetorial se considerarmos como operadores de soma e multiplicação por escalar as seguintes aplicações

$$+ : T(V) \times T(V) \rightarrow T(V)$$

$$\left(\sum_k v_k, \sum_k w_k \right) \mapsto \sum_k (v_k + w_k),$$

com as somas acima finitas e $v_k, w_k \in V^{\otimes k}$.

$$\odot : \mathbb{R} \times T(V) \rightarrow T(V)$$

$$\left(\lambda, \sum_k w_k \right) \mapsto \sum_k (\lambda w_k),$$

Além do mais, podemos definir uma multiplicação neste espaço vetorial:

$$M : T(V) \otimes T(V) \rightarrow T(V)$$

$$\left(\sum_k v_1^k \otimes \cdots \otimes v_{i_k}^k, \sum_l w_1^l \otimes \cdots \otimes w_{i_l}^l \right) \mapsto \sum_{k,l} (v_1^k \otimes \cdots \otimes v_{i_k}^k \otimes w_1^l \otimes \cdots \otimes w_{i_l}^l),$$

com $v_1^k, \dots, v_{i_k}^k, w_1^l, \dots, w_{i_l}^l \in V$. Decorre do produto tensorial ser associativo e de $V^{\otimes k} \otimes \mathbb{R} \cong V^{\otimes k}$, para todo inteiro $k \geq 0$, que o espaço vetorial $T(V)$ será uma álgebra associativa com unidade, chamada de *álgebra tensorial* do espaço vetorial V . Aqui, a unidade da álgebra tensorial é dada pela unidade $1 \in \mathbb{R}$, do corpo dos reais.

Para a álgebra tensorial vale a seguinte propriedade universal:

Proposição 4.1 (Propriedade universal da álgebra tensorial). *Seja V um espaço vetorial e \mathcal{A} uma álgebra associativa com unidade. Então, para toda transformação linear $f : V \rightarrow \mathcal{A}$ existe um homeomorfismo de álgebras $\phi : T(V) \rightarrow \mathcal{A}$, entre a álgebra tensorial do espaço V e a álgebra \mathcal{A} , tal que o diagrama abaixo comuta.*

$$\begin{array}{ccc} V & \xrightarrow{i} & T(V) \\ & \searrow f & \downarrow \phi \\ & & \mathcal{A} \end{array}$$

Aqui, $i : V \rightarrow T(V)$ representa a aplicação inclusão do espaço vetorial na álgebra tensorial.

Demonstração. Basta considerar a função $\phi : T(V) \rightarrow \mathcal{A}$ tal que

$$\sum_k v_1^k \otimes \cdots \otimes v_{i_k}^k \mapsto \sum_k f(v_1^k) \otimes \cdots \otimes f(v_{i_k}^k),$$

com $v_1^k, \dots, v_{i_k}^k \in V$. □

4.2 ÁLGEBRAS DE LIE

Nesta seção introduziremos os conceitos de álgebras de Lie, ideais, álgebras semi-simples e espaços tangentes, e mostraremos como estes se relacionam.

4.2.1 Definições Gerais

Seja \mathfrak{g} um \mathbb{R} -espaço vetorial. Dizemos que \mathfrak{g} é uma álgebra de Lie se existir um operador $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$, chamado de colchete de Lie, satisfazendo

- (i) $[\cdot, \cdot]$ é um operador bilinear;
- (ii) $[X, X] = 0$, para todo $X \in \mathfrak{g}$;
- (iii) Para todo $X, Y, Z \in \mathfrak{g}$ vale a igualdade, chamada de identidade de Jacobi,

$$[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0.$$

Uma propriedade que decorre imediatamente do colchete de Lie é a propriedade antissimétrica, i.e., para todo $X, Y \in \mathfrak{g}$ vale $[X, Y] = -[Y, X]$. Para averiguarmos esta igualdade, observemos que

$$0 = [X + Y, X + Y] = [X, X] + [X, Y] + [Y, X] + [Y, Y],$$

ou seja, $[X, Y] = -[Y, X]$.

Exemplo 4.2.1. O \mathbb{R} -espaço vetorial das matrizes $n \times n$ com entradas nos números reais $\mathcal{M}_n(\mathbb{R})$ pode ser visto como uma álgebra de Lie se considerarmos o seguinte colchete (denotado por *comutador*):

$$\begin{aligned} [\cdot, \cdot] : \mathcal{M}_n(\mathbb{R}) \times \mathcal{M}_n(\mathbb{R}) &\rightarrow \mathcal{M}_n(\mathbb{R}) \\ (A, B) &\mapsto AB - BA. \end{aligned}$$

O operador $[\cdot, \cdot]$ é um colchete de Lie pois, para toda matriz $A, B, C \in \mathcal{M}_n(\mathbb{R})$ e para todo $\lambda \in \mathbb{R}$, temos

(i)

$$\begin{aligned} [\lambda \cdot A + B, C] &= (\lambda \cdot A + B)C - C(\lambda \cdot A + B) \\ &= \lambda \cdot AC + BC - \lambda \cdot CA - CB \\ &= \lambda \cdot (AC - CA) + (BC - CB), \end{aligned}$$

ou seja, $[\lambda \cdot A + B, C] = \lambda \cdot [A, C] + [B, C]$. A prova de que $[A, \lambda \cdot B + C] = \lambda \cdot [A, B] + [A, C]$ é similar ao que foi feito acima. Sendo assim, $[\cdot, \cdot]$ é um operador bilinear;

(ii) $[A, A] = 0$, imediato;

(iii) Somando as três igualdades abaixo

- $[A, [B, C]] = [A, BC - CB] = ABC - ACB - BCA + CBA$;
- $[C, [A, B]] = [C, AB - BA] = CAB - CBA - ABC + BAC$;
- $[B, [C, A]] = [B, CA - AC] = BCA - BAC - CAB + ACB$

obtemos a identidade de Jacobi, i.e.,

$$[A, [B, C]] + [C, [A, B]] + [B, [C, A]] = 0.$$

Desta forma, concluímos que $\mathcal{M}_n(\mathbb{R})$ munido do colchete $[A, B] = AB - BA$, para todo $A, B \in \mathcal{M}_n(\mathbb{R})$, é uma álgebra de Lie.

Exemplo 4.2.2. O espaço vetorial \mathbb{R}^3 munido do produto vetorial também é uma álgebra de Lie.

Definição 4.1. Seja \mathfrak{g} uma álgebra de Lie e $\mathfrak{h} \subseteq \mathfrak{g}$ um subespaço vetorial. Dizemos que \mathfrak{h} é uma subálgebra de \mathfrak{g} se, para todo $X, Y \in \mathfrak{h}$, tivermos $[X, Y] \in \mathfrak{h}$.

Um exemplo de uma subálgebra é o seguinte: seja \mathfrak{g} uma álgebra de Lie e \mathfrak{h} o subconjunto

$$\mathfrak{h} = \{X \in \mathfrak{g}; [X, Y] = 0, \text{ para todo } Y \in \mathfrak{g}\}.$$

Então \mathfrak{h} é uma subálgebra de \mathfrak{g} chamada de *centro da álgebra* de \mathfrak{g} . Para verificarmos que \mathfrak{h} é uma subálgebra, devemos checar que

- \mathfrak{h} é um espaço vetorial;
- para todo $X, Y \in \mathfrak{h}$ temos $[X, Y] \in \mathfrak{h}$.

Que \mathfrak{h} é um espaço vetorial é imediato, pois

(i) $[0, Y] = 0 \cdot [0, Y] = 0$, para todo $Y \in \mathfrak{g}$. Portanto $0 \in \mathfrak{h}$;

(ii) Dados $X, Y \in \mathfrak{h}$, temos que, para todo $Z \in \mathfrak{g}$, vale

$$[X + Y, Z] = [X, Z] + [Y, Z] = 0.$$

Logo, $X + Y \in \mathfrak{h}$;

(iii) Dados $X \in \mathfrak{h}$ e $\lambda \in \mathbb{R}$, vale, para todo $Y \in \mathfrak{g}$, que

$$[\lambda X, Y] = \lambda[X, Y] = 0.$$

Então, $\lambda X \in \mathfrak{h}$.

Finalmente, para provarmos que \mathfrak{h} é uma subálgebra de \mathfrak{g} , devemos mostrar que, para todo $X, Y \in \mathfrak{h}$, vale $[X, Y] \in \mathfrak{h}$. Para isso, observe que pela identidade de Jacobi temos, para todo $Z \in \mathfrak{g}$, a seguinte igualdade

$$[Z, [X, Y]] + [Y, [Z, X]] + [X, [Y, Z]] = 0.$$

Como $[Z, X] = [Y, Z] = 0$, obtemos $[Z, [X, Y]] = 0$ e, portanto, concluímos que $[X, Y] \in \mathfrak{h}$.

Outro exemplo de subálgebra é o seguinte:

Exemplo 4.2.3. O espaço

$$sl(n) = \{X \in \mathcal{M}_n(\mathbb{R}); tr(X) = 0\}$$

é uma subálgebra de $\mathcal{M}_n(\mathbb{R})$. Isto é verdade, visto que, para todo $X, Y \in sl(n)$ e para todo $\lambda \in \mathbb{R}$, temos

$$tr(\lambda \cdot X + Y) = \lambda \cdot tr(X) + tr(Y) = 0$$

e $0 \in sl(n)$ (aqui, 0 representa a matriz com entradas todas nulas). Portanto, $sl(n)$ é um subespaço vetorial e

$$\begin{aligned} tr([X, Y]) &= tr(XY - YX) \\ &= tr(XY) - tr(YX) \\ &= tr(XY) - tr(XY) = 0, \end{aligned}$$

ou seja, $[X, Y] \in sl(n)$.

Definição 4.2. Seja \mathfrak{g} uma álgebra de Lie e \mathfrak{h} um subespaço vetorial. Dizemos que \mathfrak{h} é um ideal de \mathfrak{g} se, para todo $X \in \mathfrak{g}$ e para todo $Y \in \mathfrak{h}$, tivermos $[X, Y] \in \mathfrak{h}$.

Da definição de ideal decorre imediatamente que todo ideal também é uma subálgebra, porém, a volta não é verdadeira. Pelo Exemplo 4.2.2 podemos construir um exemplo de subálgebra que não é um ideal. Considere o subespaço vetorial

$$\mathfrak{h} = \{\lambda \cdot (1, 0, 0); \lambda \in \mathbb{R}\}.$$

Temos que \mathfrak{h} é uma subálgebra de \mathbb{R}^3 pois, para todo $\lambda_1, \lambda_2 \in \mathbb{R}$, temos

$$\lambda_1 \cdot (1, 0, 0) \times \lambda_2 \cdot (1, 0, 0) = 0 \in \mathfrak{h}.$$

Agora, tome o vetor $(0, 1, 0)$. O produto vetorial entre $(0, 1, 0)$ e $(1, 0, 0)$ será igual a

$$(0, 1, 0) \times (1, 0, 0) = (0, 0, -1) \notin \mathfrak{h},$$

ou seja, \mathfrak{h} não é um ideal.

4.2.2 Álgebras de Lie semi-simples e a Representação Adjunta

Nesta subseção estabeleceremos o conceito de álgebras de Lie semi-simples e o conceito de representação adjunta. No estudo do elemento do Casimir de uma álgebra de Lie, i.e., no estudo de certos elementos do centro de uma álgebra, teremos que tais elementos possuem uma fórmula fechada quando sua álgebra for semi-simples. Devido a importância do elemento do Casimir em nosso estudo, nesta subseção iremos abordar estas propriedades.

Definição 4.3. (*Série derivada*) Seja \mathfrak{g} uma álgebra de Lie. Definimos por sua série derivada como sendo a sequência de ideais de $\mathfrak{g} : \mathfrak{g}^{(1)}, \dots, \mathfrak{g}^{(n)}, \mathfrak{g}^{(n+1)}, \dots$ satisfazendo

$$\mathfrak{g}^{(1)} = \mathfrak{g} \quad \text{e} \quad \mathfrak{g}^{(n+1)} = [\mathfrak{g}^{(n)}, \mathfrak{g}^{(n)}],$$

com $[\mathfrak{g}^{(n)}, \mathfrak{g}^{(n)}]$ sendo o seguinte conjunto:

$$[\mathfrak{g}^{(n)}, \mathfrak{g}^{(n)}] = \{[X, Y] \in \mathfrak{g}; X \in \mathfrak{g}^{(n)} \text{ e } Y \in \mathfrak{g}^{(n)}\}.$$

Observação. Dizemos que uma álgebra de Lie \mathfrak{g} é *solúvel* se existir um natural $n > 1$ tal que $\mathfrak{g}^{(n)} = \{0\}$.

Exemplo 4.2.4. Seja a álgebra de Lie dada por

$$\mathfrak{g} = \text{span} \left\{ X_1 = \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix}, \quad X_2 = \begin{pmatrix} 0 & 0 \\ 1 & -1 \end{pmatrix} \right\}$$

com colchete de Lie dado pelo comutador. Fazendo os cálculos, obtemos que a série derivada de \mathfrak{g} é dada por: $\mathfrak{g}^{(1)} = \mathfrak{g}$, $\mathfrak{g}^{(2)} = \text{span}\{X_2 - X_1\}$ e $\mathfrak{g}^{(3)} = \{0\}$. Pela igualdade, $\mathfrak{g}^{(3)} = \{0\}$, obtemos que \mathfrak{g} é uma álgebra de Lie solúvel.

Definição 4.4. Seja \mathfrak{g} uma álgebra de Lie. Dizemos que \mathfrak{g} é uma álgebra de Lie semi-simples se o único ideal solúvel \mathfrak{h} de \mathfrak{g} for o ideal dado por $\mathfrak{h} = \{0\}$.

A seguir, definimos o conceito de representação adjunta:

Definição 4.5. Seja \mathfrak{g} uma álgebra de Lie e $\mathfrak{gl}(\mathfrak{g})$ a álgebra de Lie dada pelas transformações lineares $h : \mathfrak{g} \rightarrow \mathfrak{g}$ com colchete de Lie dado pelo comutador, i.e., para todo $f, h \in \mathfrak{g}$ vale $[f, h] = f \circ h - h \circ f$. Então, chamamos de *representação adjunta* de \mathfrak{g} a seguinte aplicação:

$$\begin{aligned} \text{Ad} : \mathfrak{g} &\rightarrow \mathfrak{gl}(\mathfrak{g}) \\ X &\mapsto \text{Ad}(X) : \mathfrak{g} \rightarrow \mathfrak{g} \\ Y &\mapsto [X, Y]. \end{aligned}$$

Dados $X, Y \in \mathfrak{g}$ e λ um número real, temos, para todo $Z \in \mathfrak{g}$, que

$$\text{Ad}(X + \lambda Y)(Z) = [X + \lambda Y, Z] = [X, Z] + \lambda[Y, Z],$$

ou seja, $\text{Ad}(X + \lambda Y) = \text{Ad}(X) + \lambda \text{Ad}(Y)$. Além do mais,

$$\begin{aligned} \text{Ad}([X, Y])(Z) &= [[X, Y], Z] \\ &= -([Y, [Z, X]] + [X, [Y, Z]]) \quad (\text{igualdade obtida pela identidade de Jacobi}) \\ &= -(-[Y, [X, Z]] + [X, [Y, Z]]) \\ &= -(-\text{Ad}(Y) \circ \text{Ad}(X) + \text{Ad}(X) \circ \text{Ad}(Y))(Z) \\ &= [\text{Ad}(X), \text{Ad}(Y)](Z). \end{aligned}$$

Logo, obtemos que a representação adjunta é um homeomorfismo de álgebras de Lie.

4.2.3 A Álgebra Linear Especial de Lie $sl_2(\mathbb{R})$

Um caso específico de álgebra de Lie, que surgirá em nosso trabalho, será a álgebra linear especial $sl_2(\mathbb{R})$ sobre o corpo dos reais (“special linear Lie algebra”). Abaixo segue sua definição:

Definição 4.6. A álgebra $sl_2(\mathbb{R})$ é a álgebra de Lie das matrizes quadradas 2×2 com colchete de Lie dado pelo comutador e o qual as matrizes satisfazem a condição de terem seu traço igual a zero:

$$sl_2(\mathbb{R}) = \{A \in M_2(\mathbb{R}); \text{tr}(A) = 0\}.$$

A álgebra $sl(2, \mathbb{R})$ admite como base os vetores

$$E := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad F := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad H := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

com comutadores dados por

$$[E, F] = H, \quad [H, F] = -2F, \quad [H, E] = 2E.$$

Dentre as diversas propriedades de $sl_2(\mathbb{R})$, as quais podem ser vistas em [Lang], destacamos as três abaixo:

Proposição 4.2. *A álgebra $sl_2(\mathbb{R})$ é uma álgebra simples, i.e., esta álgebra contém apenas os ideais triviais, dados pela identidade e a própria álgebra.*

Proposição 4.3. *Um elemento de Casimir, denotado por C , da álgebra $sl_2(\mathbb{R})$ (i.e., um elemento do centro da álgebra), considerando a base acima, é dado por*

$$C = H^2 + 2(EF + FH).$$

Proposição 4.4. *A álgebra universal envelopante $U(sl_2(\mathbb{R}))$, vide Seção, é a álgebra livre gerada pelos elementos e, f, h que satisfazem as restrições*

$$ef - fh = h, \quad hf - fh = -2f, \quad he - eh = 2e.$$

4.2.4 O Espaço Tangente e suas propriedades algébricas

Seja G um subgrupo do grupo linear geral $GL(n, \mathbb{R})$. Dizemos que uma função $g : [0, 1] \rightarrow G$ é um caminho diferencial se g for diferenciável, ou seja, $g(t) = (a_{ij}(t))$ satisfaz a condição de que cada entrada $a_{ij} : [0, 1] \rightarrow \mathbb{R}$ seja uma função diferenciável. Com o conceito de caminhos diferenciáveis, definimos:

Definição 4.7. Seja G um subgrupo de $GL(n, \mathbb{R})$. Dizemos que o espaço tangente de G , denotado por $T_1(G)$, é o seguinte conjunto

$$T_1(G) = \{A \in \mathcal{M}_n(\mathbb{R}); \exists g : [0, 1] \rightarrow G \text{ um caminho diferenciável} \\ \text{tal que } g(0) = I_n \text{ e } g'(0) = A\},$$

com $\mathcal{M}_n(\mathbb{R})$ o conjunto das matrizes $n \times n$ e I_n a matriz identidade.

Ou seja, o espaço tangente $T_1(G)$ de um subgrupo G do grupo linear geral é o conjunto dos vetores tangentes a identidade em G . Além do mais, veremos a seguir que o espaço tangente satisfaz a propriedade de ser uma álgebra de Lie. Porém, antes de provarmos este fato, ilustremos o conceito de espaço tangente pelo seguinte exemplo:

Exemplo 4.2.5. Considere o grupo linear geral das matrizes 2×2 :

$$GL(2, \mathbb{R}) = \left\{ A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}; ad - bc \neq 0. \right\}$$

Sabemos que o conjunto $H = \{A \in GL(2, \mathbb{R}); \det(A) = 1\}$ é um subgrupo do grupo linear geral. Logo, todo caminho diferenciável $g : [0, 1] \rightarrow H$, satisfazendo $g(0) = I_2$ (I_2 a matriz identidade), deve satisfazer, se considerarmos

$$g(t) = \begin{pmatrix} a(t) & b(t) \\ c(t) & d(t) \end{pmatrix},$$

a restrição $(a(t)d(t) - b(t)c(t))' = 0$, ou seja, quando $t = 0$ devemos ter

$$a'(0)d(0) + a(0)d'(0) = b'(0)c(0) + b(0)c'(0).$$

Porém, dada a igualdade $g(0) = I_2$, temos que $a(0) = d(0) = 1$ e $c(0) = b(0) = 0$. Logo, nossa restrição para $g'(0)$ será dada por: $a'(0) + d'(0) = 0$. Sendo assim, obtemos que

$$g'(0) = \begin{pmatrix} a'(0) & b'(0) \\ c'(0) & -a'(0) \end{pmatrix}.$$

Ou seja, o espaço tangente de H é dado por:

$$T_1(H) = \left\{ \begin{pmatrix} x & y \\ z & -x \end{pmatrix}; x, y, z \in \mathbb{R}. \right\}$$

Teorema 4.5. O espaço tangente $T_1(G)$ de um subgrupo G de $GL(n, \mathbb{R})$ é uma álgebra de Lie.

Demonstração. Inicialmente, mostremos que $T_1(G)$ é um \mathbb{R} -espaço vetorial. Dadas as matrizes $X, Y \in T_1(G)$, existem os seguintes caminhos diferenciáveis

$$g_X : [0, 1] \rightarrow G \quad \text{e} \quad g_Y : [0, 1] \rightarrow G,$$

satisfazendo

$$\begin{cases} g_X(0) = I_n \\ g'_X(0) = X \end{cases} \quad \text{e} \quad \begin{cases} g_Y(0) = I_n \\ g'_Y(0) = Y \end{cases}.$$

Observe que a função

$$\begin{aligned} h : [0, 1] &\rightarrow G \\ t &\mapsto g_X(t) \cdot g_Y(t) \end{aligned}$$

é diferenciável e $h(0) = I_n$. Tomando a derivada de h no ponto 0, obtemos

$$\begin{aligned} h'(0) &= g'_X(0) \cdot g_Y(0) + g_X(0) \cdot g'_Y(0) \\ &= X + Y, \end{aligned}$$

portanto, $X + Y = h'(0) \in T_1(G)$. Agora, dado um escalar $\lambda \in \mathbb{R}$ e uma matriz $X \in T_1(G)$ qualquer, obtemos, como anteriormente, um caminho diferenciável $g_X : [0, 1] \rightarrow G$, com $g_X(0) = I_n$ e $g'_X(0) \in X$. Tome a função

$$\begin{aligned} \psi : [0, 1] &\rightarrow [0, 1] \\ t &\mapsto \begin{cases} \lambda t \cdot \mathbb{1}\{t \leq 1/\lambda\}, & \text{se } \lambda \neq 0; \\ 0, & \text{se } \lambda = 0. \end{cases} \end{aligned}$$

Vemos que ψ é diferenciável e, portanto, a composta $g_X \circ \psi$ é um caminho diferenciável, com $(g_X \circ \psi)(0) = I_n$. Tomando a derivada de $(g_X \circ \psi)$, obtemos, pela regra da cadeia, que

$$\begin{aligned} (g_X \circ \psi)'(0) &= g'_X(\psi(0)) \cdot \psi'(0) \\ &= g'_X(0) \cdot \lambda \\ &= \lambda X. \end{aligned}$$

Portanto, $T_1(G)$ é um \mathbb{R} -espaço vetorial. Agora, provemos que este espaço vetorial tem estrutura de uma álgebra de Lie. Como sabemos que o conjunto das matrizes $\mathcal{M}_n(\mathbb{R})$ munido do colchete

$$\begin{aligned} [\cdot, \cdot] : \mathcal{M}_n(\mathbb{R}) \times \mathcal{M}_n(\mathbb{R}) &\rightarrow \mathcal{M}_n(\mathbb{R}) \\ (A, B) &\mapsto AB - BA \end{aligned}$$

é uma álgebra de Lie, devemos provar que o colchete $[\cdot, \cdot]$ restrito a $T_1(G)$ tem imagem ainda em $T_1(G)$, i.e., dados $A, B \in T_1(G)$, devemos provar que $[A, B] \in T_1(G)$.

Dadas as matrizes $A, B \in T_1(G)$, existem os caminhos diferenciáveis

$$g_A : [0, 1] \rightarrow G \quad \text{e} \quad g_B : [0, 1] \rightarrow G,$$

satisfazendo

$$\begin{cases} g_A(0) = I_n \\ g'_A(0) = A \end{cases} \quad \text{e} \quad \begin{cases} g_B(0) = I_n \\ g'_B(0) = B \end{cases}.$$

Definimos agora a função auxiliar

$$\begin{aligned} \zeta : [0, 1] \times [0, 1] &\rightarrow G \\ (s, t) &\mapsto g_A(s) \cdot g_B(t) \cdot g_A^{-1}(s). \end{aligned}$$

Tomando a derivada de ζ em relação a coordenada t , obtemos

$$\frac{d}{dt}\zeta(s, t) = g_A(s) \cdot g'_B(t) \cdot g_A^{-1}(s).$$

Logo, para todo $s \in [0, 1]$, vale

$$\zeta(s, 0) = I_n \quad \text{e} \quad \frac{d}{dt}\zeta(s, 0) = g_A(s) \cdot B \cdot g_A^{-1}(s) \in T_1(G).$$

Com isto, definimos o operador $\rho : [0, 1] \rightarrow T_1(G)$, com $\rho(s) := \frac{d}{dt}\zeta(s, 0)$.

Agora, observe que o espaço tangente é um subespaço vetorial de $\mathcal{M}_n(\mathbb{R})$, que por sua vez é um espaço vetorial normado de dimensão finita. Logo, $T_1(G)$ é um subconjunto fechado em $\mathcal{M}_n(\mathbb{R})$. Tomando a derivada do operador ρ , obtemos

$$\frac{d}{ds}\rho(s) = g'_A(s) \cdot B \cdot g_A^{-1}(s) + g_A(s) \cdot B \{-g_A^{-1}(s) \cdot g'_A(s) \cdot g_A^{-1}(s)\}$$

Esta derivada, calculada em $s = 0$, fornece $\rho'(0) = AB - BA$. Porém, a derivada de ρ em relação ao ponto 0 pode ser vista como

$$\begin{aligned} \frac{d}{ds}\rho(0) &= \lim_{h \rightarrow 0} \left\{ \frac{g_A(h) \cdot B \cdot g_A^{-1}(h) - g_A(0) \cdot B \cdot g_A^{-1}(0)}{h} \right\} \\ &= \lim_{h \rightarrow 0} \left\{ \frac{g_A(h) \cdot B \cdot g_A^{-1}(h) - B}{h} \right\}. \end{aligned}$$

Ou seja, para $n \in \mathbb{N}$, a sequência

$$x_n = \frac{g_A(1/n) \cdot B \cdot g_A^{-1}(1/n) - B}{1/n} \in T_1(G)$$

converge para $AB - BA$. Como $T_1(G)$ é um conjunto fechado, obtemos de

$$x_n \xrightarrow{n \rightarrow \infty} AB - BA,$$

que:

$$[A, B] = AB - BA \in T_1(G).$$

Sendo assim, concluímos que $T_1(G)$ é uma álgebra de Lie. □

4.3 BIÁLGEBRAS

Antes de definirmos uma biálgebra precisamos estabelecer o conceito de coálgebras:

Definição 4.8. Dada a terna $(\mathcal{C}, \delta, \varepsilon)$, onde \mathcal{C} representa um \mathbb{K} -espaço vetorial munido dos operadores $\Delta : \mathcal{C} \rightarrow \mathcal{C} \otimes \mathcal{C}$, chamado de coproduto, e $\varepsilon : \mathcal{C} \rightarrow \mathbb{K}$, denotado por counidade. Dizemos que esta terna é uma coálgebra se os diagramas abaixo comutarem:

$$\begin{array}{ccc}
 \mathcal{C} \otimes \mathcal{C} \otimes \mathcal{C} & \xleftarrow{Id \otimes \Delta} & \mathcal{C} \otimes \mathcal{C} \\
 \uparrow \Delta \otimes Id & & \uparrow \Delta \\
 \mathcal{C} \otimes \mathcal{C} & \xleftarrow{\Delta} & \mathcal{C}
 \end{array}$$

Diagrama 1, coassociatividade.

$$\begin{array}{ccccc}
 & & \mathcal{C} \otimes \mathcal{C} & & \\
 \varepsilon \otimes Id \swarrow & & \uparrow \Delta & \searrow Id \otimes \varepsilon & \\
 \mathbb{K} \otimes \mathcal{C} & & & & \mathcal{C} \otimes \mathbb{K} \\
 s_1^{-1} \swarrow & & \mathcal{C} & \searrow s_2^{-1} & \\
 & & & &
 \end{array}$$

Diagrama 2, existência da counidade

Acima, temos que as funções Id, s_1^{-1} e s_2^{-1} são dadas por:

$$\begin{array}{lll}
 Id : \mathcal{C} \rightarrow \mathcal{C} & s_1^{-1} : \mathcal{C} \rightarrow \mathbb{K} \otimes \mathcal{C} & s_2^{-1} : \mathcal{C} \rightarrow \mathcal{C} \otimes \mathbb{K} \\
 x \mapsto x & x \mapsto 1 \otimes x & x \mapsto x \otimes 1.
 \end{array}$$

Uma coálgebra importante em nossos estudos é a *álgebra universal envelopante*. Esta álgebra, de certa forma, “estende” álgebras de Lie para álgebras associativas com unidade, preservando o operador colchete.

Definição 4.9. (*álgebra universal envelopante*) Seja \mathfrak{g} uma álgebra de Lie, $T(\mathfrak{g})$ sua álgebra tensorial e $I(\mathfrak{g})$ o ideal dado por:

$$I(\mathfrak{g}) = \text{span} \{a \otimes b - b \otimes a - [a, b] \in T(\mathfrak{g}); a, b \in \mathfrak{g}\}. \quad (*)$$

Denotamos por $U(\mathfrak{g}) = T(\mathfrak{g})/I(\mathfrak{g})$, o quociente da álgebra tensorial pelo ideal $I(\mathfrak{g})$, como sendo a álgebra universal envelopante de \mathfrak{g} .

Observação. Dado \mathfrak{g} uma álgebra de Lie, $I(\mathfrak{g})$ o ideal dado acima na Definição 4.9 e $U(\mathfrak{g})$ sua álgebra universal envelopante, temos, para todo $a, b \in \mathfrak{g}$, que

$$(a \otimes b - b \otimes a - [a, b]) + I(\mathfrak{g}) = I(\mathfrak{g})$$

pois $a \otimes b - b \otimes a - [a, b] \in I(\mathfrak{g})$. Logo,

$$(a \otimes b - b \otimes a) + I(\mathfrak{g}) = ([a, b]) + I(\mathfrak{g}),$$

ou seja, o operador colchete da álgebra de Lie \mathfrak{g} é dado pelo comutador na álgebra universal envelopante.

Considerando a álgebra de Lie dada pelo espaço das matrizes quadradas $n \times n$ com entrada nos reais munida do comutador. Mostremos que a álgebra universal envelopante $U(\mathfrak{g})$ pode ser vista como uma coálgebra. Para simplificar a notação, denotaremos o espaço $\mathcal{M}_n(\mathbb{R})$ por apenas \mathfrak{g} (i.e., $\mathfrak{g} = \mathcal{M}_n(\mathbb{R})$) e denotaremos por $I(\mathfrak{g})$ o ideal descrito acima em (*).

Dada a transformação linear $\Delta' : \mathfrak{g} \rightarrow U(\mathfrak{g}) \otimes U(\mathfrak{g})$, satisfazendo $x \mapsto x \otimes I_n + I_n \otimes x$, com I_n a matriz identidade, obtemos, para todo $x \in \mathfrak{g}$, que

$$\begin{aligned} ((Id \otimes \Delta') \circ \Delta')(x) &= (Id \otimes \Delta')(x \otimes I_n + I_n \otimes x) \\ &= x \otimes \Delta'(I_n) + I_n \otimes \Delta'(x) \\ &= 2x \otimes I_n \otimes I_n + I_n \otimes x \otimes I_n + I_n \otimes I_n \otimes x. \end{aligned}$$

Note que $[x, I_n] = 0$, pois $[x, I_n] = x \cdot I_n - I_n \cdot x$. Logo, decorre de

$$(x \otimes I_n - I_n \otimes x - [x, I_n]) + I(\mathfrak{g}) = I(\mathfrak{g})$$

que $(x \otimes I_n) + I(\mathfrak{g}) = (I_n \otimes x) + I(\mathfrak{g})$ e, portanto, na álgebra universal envelopante podemos escrever $x \otimes I_n \otimes I_n = I_n \otimes I_n \otimes x$ e $I_n \otimes I_n \otimes x = x \otimes I_n \otimes I_n$. Ou seja,

$$((Id \otimes \Delta') \circ \Delta')(x) = 2I_n \otimes I_n \otimes x + I_n \otimes x \otimes I_n + x \otimes I_n \otimes I_n.$$

Todavia, temos que

$$\begin{aligned} ((\Delta' \otimes Id) \circ \Delta')(x) &= (\Delta' \otimes Id)(x \otimes I_n + I_n \otimes x) \\ &= \Delta'(x) \otimes I_n + \Delta(I_n) \otimes x \\ &= x \otimes I_n \otimes I_n + I_n \otimes x \otimes I_n + 2I_n \otimes I_n \otimes x. \end{aligned}$$

Sendo assim, a transformação linear Δ' satisfaz a condição $(Id \otimes \Delta') \circ \Delta' = (\Delta' \otimes Id) \circ \Delta'$ para todo $x \in \mathfrak{g}$. Agora, pela propriedade universal da álgebra tensorial, Proposição 4.1, existe um homeomorfismo de álgebras $\Delta'' : T(\mathfrak{g}) \rightarrow U(\mathfrak{g}) \otimes U(\mathfrak{g})$ tal que $\Delta'(x) = \Delta''(x)$, para todo $x \in \mathfrak{g}$. Logo, obtemos o operador $\Delta : U(\mathfrak{g}) \rightarrow U(\mathfrak{g}) \otimes U(\mathfrak{g})$ dado por

$$\Delta(x + I(\mathfrak{g})) = \Delta''(x)$$

que satisfaz a restrição $(Id \otimes \Delta) \circ \Delta = (\Delta \otimes Id) \circ \Delta$. Portanto, Δ é um coproduto da álgebra universal envelopante das matrizes quadradas.

Finalmente, para definirmos a counidade observamos que a transformação linear $\varepsilon' : \mathfrak{g} \rightarrow \mathbb{R}$ dada por $x \mapsto 0$, para toda matriz quadrada x , estabelece, devido a propriedade universal da álgebra tensorial, um homeomorfismo de álgebras $\varepsilon : U(\mathfrak{g}) \rightarrow \mathbb{R}$ tal que, para toda matriz x , vale $\varepsilon(x) = \varepsilon'(x)$ e $\varepsilon(1) = 1$. Resta, mostrarmos que a propriedade de counidade é preservada por ε .

Note que, para todo x diferente dos escalares, vale:

$$\begin{aligned} ((Id \otimes \varepsilon) \circ \Delta)(x) &= (Id \otimes \varepsilon)(x \otimes I_n + I_n \otimes x) \\ &= x \otimes 0 + I_n \otimes 0 \\ &\cong 0 \end{aligned}$$

e

$$\begin{aligned} ((\varepsilon \otimes Id) \circ \Delta)(x) &= (\varepsilon \otimes Id)(x \otimes I_n + I_n \otimes x) \\ &= 0 \otimes I_n + 0 \otimes x \\ &\cong 0. \end{aligned}$$

Logo, para todo elemento da álgebra universal das matrizes diferente dos reais vale a igualdade: $(Id \otimes \varepsilon) \circ \Delta = (\varepsilon \otimes Id) \circ \Delta$. De forma semelhante mostramos que esta igualdade também é válida para os reais. Logo, a terna $(U(\mathfrak{g}), \Delta, \varepsilon)$ é uma coálgebra.

Em posse destes resultados podemos definir o conceito de biálgebra.

Definição 4.10. Seja H um espaço vetorial. Dizemos que H é uma biálgebra se o mesmo possuir estrutura de álgebra (H, M, u) , estrutura de coálgebra (H, Δ, ε) e os operadores de comultiplicação e counidade forem homeomorfismos de álgebras.

Decorre desta definição que a álgebra universal envelopante das matrizes quadradas é uma biálgebra, pois a mesma é uma álgebra, uma coálgebra e a comultiplicação e counidade são homeomorfismos de álgebras.

5

MATRIZES ESTOCÁSTICAS

Cadeias de Markov a tempo discreto (X_n) e a tempo contínuo $(X_t, t \geq 0)$ com espaço de estados Ω podem ser caracterizadas por meio de suas *matrizes de transição* $M = (M_{ij})_{i,j \in \Omega}$ e por meio de suas matrizes geradoras $L = (L_{ij})_{i,j \in \Omega}$, respectivamente. Estas matrizes devem satisfazer as seguintes restrições

- (i) $M_{ij} \geq 0$ para todo $i, j \in \Omega$ e $\sum_j M_{ij} = 1$ para todo $i \in \Omega$. Tais matrizes são chamadas de *matrizes de transição* ou *matrizes estocásticas*;
- (ii) $L_{ii} \leq 0$ para todo $i \in \Omega$ e $L_{ij} \geq 0$ para todos elementos $i \neq j$ no espaço de estado. Além do mais, deve valer a igualdade $\sum_j L_{ij} = 0$ para todo $i \in \Omega$. Matrizes que satisfazem estas condições são chamadas de *matrizes geradoras* ou *L-matrizes*.

Neste capítulo iremos estudar uma forma de gerarmos cadeias de Markov a tempo discreto por meio de L-matrizes, e vice-versa. Por exemplo, dada uma L-matriz, que caracteriza uma cadeia de Markov a tempo contínuo, estaremos interessados em obter uma matriz de transição M de um processo de Markov a tempo discreto, como por exemplo a cadeia dada na Figura 5.

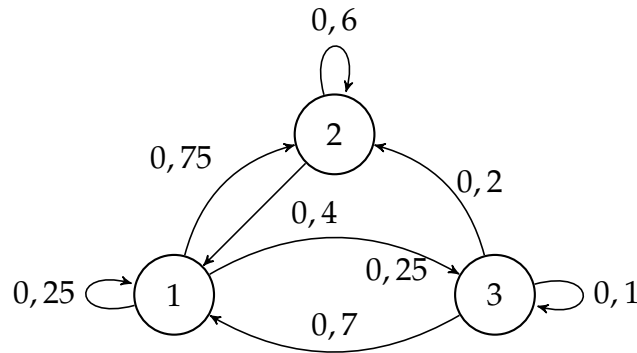


Figura 5: Exemplo de uma cadeia de Markov a tempo discreto com espaço de estados $\Omega = \{1, 2, 3\}$.

Acontece que, ao estudarmos o grupo de matrizes de Lie, dado pelas matrizes estocásticas, percebemos existir uma relação com seu espaço tangente, que contém as

L-matrizes desejadas. Neste capítulo estudaremos tais relações, o qual foram estudadas em [Summer] para o caso de matrizes quadradas de dimensão 2.

Nosso resultado principal foi generalizar o obtido em [Summer] para matrizes estocásticas de dimensão qualquer, o qual provamos que a componente conexa da identidade \mathcal{G}^0 em relação ao grupo \mathcal{G} pode ser decomposta da seguinte forma

$$\mathcal{G}^0 = \bigcup_{t \in \mathbb{R}} e^{Q_0 t} \mathcal{H}, \quad (*)$$

com Q_0 uma L-matriz e \mathcal{H} o subgrupo de \mathcal{G} formado pelas matrizes de determinante um.

Para obtermos esta decomposição provamos o seguinte resultado (cuja demonstração não encontramos na literatura e que pode ser vista no Anexo A), o qual afirmamos que a componente conexa $\mathcal{G}^0 \leq \mathcal{G}$ nada mais é que o subconjunto de \mathcal{G} formado pelas matrizes estocásticas de determinante positivo. Tal afirmação estende o proposto em [Summer], o qual provamos sua validade para grupos estocásticos de matrizes de dimensão qualquer. Além do mais, esta afirmação é fundamental para obtermos a decomposição (*).

Para encerrar este capítulo calculamos uma L-matriz a partir do espaço tangente do grupo das matrizes de Lie referente as matrizes estocásticas 3×3 . Para isto, utilizaremos como ferramenta, o operador coproduto da álgebra universal envelopante, como proposto em [Redig].

5.1 O GRUPO DAS MATRIZES ESTOCÁSTICAS E SEU ESPAÇO TANGENTE

Dado o vetor $\mathbf{1} = (1, 1, \dots, 1)$, cujas coordenadas valem 1, temos que uma matriz $A = (a_{ij})_{n \times n}$ é estocástica se, e somente se, a condição abaixo for satisfeita

$$A \cdot \mathbf{1} = \mathbf{1} \quad \text{e} \quad 0 \leq a_{ij} \leq 1, \quad \forall i, j \in \{1, 2, \dots, n\}.$$

Logo, da definição de matrizes estocásticas, definimos o seguinte conjunto:

$$\mathcal{G} = \{A \in GL(n, \mathbb{R}); A \cdot \mathbf{1} = \mathbf{1}\}.$$

O conjunto \mathcal{G} , definido desta forma, é um subgrupo do grupo linear geral, pois:

- $I_n \cdot \mathbf{1} = \mathbf{1}$, imediato;
- Para todas matrizes $A, B \in \mathcal{G}$, temos: $(AB) \cdot \mathbf{1} = A(B \cdot \mathbf{1}) = A \cdot \mathbf{1} = \mathbf{1}$. Logo, $AB \cdot \mathbf{1} = \mathbf{1}$, e concluímos que $AB \in \mathcal{G}$;
- Dado $A \in \mathcal{G}$, vale

$$A^{-1} \cdot \mathbf{1} = A^{-1} \cdot (A \cdot \mathbf{1}) = (A^{-1} \cdot A) \cdot \mathbf{1} = \mathbf{1},$$

portanto, $A^{-1} \in \mathcal{G}$.

O subgrupo \mathcal{G} será chamado de grupo das matrizes estocásticas.

Mostremos que \mathcal{G} é um subconjunto fechado de $GL(n, \mathbb{R})$, ou seja, \mathcal{G} é um grupo de matrizes de Lie. Para isto, tome a seguinte função

$$\begin{aligned} \phi : GL(n, \mathbb{R}) &\rightarrow \mathbb{R}^n \\ A &\mapsto A \cdot \mathbf{1}. \end{aligned}$$

Dadas $A, B \in GL(n, \mathbb{R})$ matrizes quaisquer, esta função satisfaz a condição de que

$$\|\phi(A) - \phi(B)\| = \|(A - B) \cdot \mathbf{1}\| \leq \sqrt{n} \cdot \|A - B\|,$$

pois $\|\mathbf{1}\| = \sqrt{n}$. Consequentemente, esta desigualdade significa que a função ϕ é uma função lipzchitziana e, portanto, ϕ é uma função contínua. Finalmente, observe que

$$\mathcal{G} = \phi^{-1}(\{\mathbf{1}\}).$$

Como o conjunto $\{\mathbf{1}\}$ é fechado em \mathbb{R}^n e ϕ é uma função contínua, concluímos que \mathcal{G} é um subconjunto fechado do grupo linear geral e, desta forma, \mathcal{G} é um grupo de matrizes de Lie.

Estudemos agora o espaço tangente de \mathcal{G} . Dado um caminho diferenciável

$$A : [0, 1] \rightarrow \mathcal{G}, \text{ com } A(0) = I_n,$$

temos, pela restrição $A(t) \cdot \mathbf{1} = \mathbf{1}$, para todo $t \in [0, 1]$, que

$$A'(t) \cdot \mathbf{1} = \mathbf{0}.$$

Logo, concluímos que $A'(0) \cdot \mathbf{1} = \mathbf{0}$ e, portanto, a seguinte inclusão é verdadeira

$$T_1(\mathcal{G}) \subseteq \{X \in \mathcal{M}_n(\mathbb{R}); X \cdot \mathbf{1} = \mathbf{0}\}.$$

Agora, devido a Propriedade 5, seção 3.3, sobre a exponencial de matrizes, o qual diz que, dado uma matriz A , de autovalor λ e autovetor v , a seguinte igualdade é válida

$$e^{tA}v = e^{t\lambda}v,$$

para t um real qualquer, obtemos, para toda matriz $A \in \{X \in \mathcal{M}_n(\mathbb{R}); X \cdot \mathbf{1} = \mathbf{0}\}$ (o qual admitem o vetor $\mathbf{1}$ como autovetor e 0 como autovalor), a igualdade:

$$e^{tA}\mathbf{1} = e^{t0}\mathbf{1} = \mathbf{1}.$$

Desta forma, para a função $f(t) = e^{tA}$ vale a relação: $f(t) = e^{tA} \in \mathcal{G}$, para todo $t \in [0, 1]$. Consequentemente,

$$f(0) = I_n \quad \text{e} \quad f'(0) = A \in T_1(\mathcal{G}).$$

Sendo assim, concluímos que

$$\{X \in \mathcal{M}_n(\mathbb{R}); X \cdot \mathbf{1} = \mathbf{0}\} \subseteq T_1(\mathcal{G})$$

e, portanto, obtemos a igualdade desejada

$$T_1(\mathcal{G}) = \{X \in \mathcal{M}_n(\mathbb{R}); X \cdot \mathbf{1} = \mathbf{0}\}.$$

Em outras palavras, toda matriz $X \in T_1(\mathcal{G})$ pode ser escrita da seguinte forma

$$X = \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,n} \end{pmatrix}, \text{ com } \sum_{j=1}^n x_{ij} = 0, \forall i \in \{1, 2, \dots, n\}$$

Portanto, da restrição $\sum_{j=1}^n x_{ij} = 0$ obtemos para cada linha i que $x_{in} = -\sum_{j=1}^{n-1} x_{ij}$. Assim, definindo as matrizes elementares E_{ij} como sendo as matrizes com entrada 1 na posição (i, j) e 0 no resto, obtemos

$$X = \sum_{\substack{i \in \{1, \dots, n\} \\ j \in \{1, \dots, n-1\}}} x_{ij} \cdot (E_{ij} - E_{in}).$$

Desta forma, concluímos que a dimensão do espaço tangente será igual a

$$\dim_{\mathbb{R}} T_1(\mathcal{G}) = n(n-1).$$

5.2 RELAÇÃO ENTRE MATRIZES ESTOCÁSTICAS E MATRIZES GERADORAS

Nesta seção mostraremos como a componente conexa da identidade, denotada por \mathcal{G}^0 , do grupo das matrizes estocásticas \mathcal{G} se relaciona com matrizes geradoras de um processo markoviano a tempo contínuo. Para isto, utilizaremos o seguinte resultado:

Proposição 5.1. *O conjunto das matrizes estocásticas de determinante positivo*

$$\{X \in \mathcal{G}; \det(X) > 0\}$$

é conexo por caminhos.

Demonstração. A demonstração desta proposição pode ser vista no Apêndice A. \square

Em posse deste resultado obtém-se imediatamente que

Corolário 5.2. *A componente conexa \mathcal{G}^0 do grupo das matrizes estocásticas é igual ao conjunto das matrizes estocásticas de determinante positivo, i.e.,*

$$\mathcal{G}^0 = \{X \in \mathcal{G}; \det(X) > 0\}$$

Demonstração. Sabemos, vide [Elon1], que a componente conexa do grupo linear geral é igual ao conjunto das matrizes de determinante positivo. Logo,

$$\mathcal{G}^0 \subseteq \{X \in \mathcal{G}; \det(X) > 0\}.$$

Além do mais, decorre da Proposição anterior que o conjunto das matrizes estocásticas de determinante positivo é conexo por caminhos. Portanto, a seguinte inclusão é verdadeira

$$\{X \in \mathcal{G}; \det(X) > 0\} \subseteq \mathcal{G}^0,$$

e encerramos a demonstração. \square

Agora, tomemos o seguinte conjunto

$$\mathcal{H} = \{A \in \mathcal{G}; \det(A) = 1\}.$$

\mathcal{H} será um subgrupo normal, pois

- A identidade está em \mathcal{H} ;

- $A, B \in \mathcal{H}$ implica $\det(AB) = \det(A) \cdot \det(B)$ e, portanto, $AB \in \mathcal{H}$;
- Para $A \in \mathcal{H}$, temos $\det(A^{-1}) = \det(A)^{-1} = 1$, ou seja, $A^{-1} \in \mathcal{H}$;
- E, finalmente, para $A \in \mathcal{H}$ e $B \in \mathcal{G}$, temos

$$\det(BAB^{-1}) = \det(B) \cdot \det(A) \cdot \det(B)^{-1} = 1,$$

$$\text{logo } B\mathcal{H}B^{-1} \subseteq \mathcal{H}.$$

Além do mais, como toda matriz $A \in \mathcal{G}$ de determinante positivo pertence a componente conexa \mathcal{G}^0 (Corolário 5.2), temos que \mathcal{H} é um subgrupo normal de \mathcal{G}^0 .

Em posse destes resultados somos capazes de demonstrar o Teorema:

Teorema 5.3. *Seja \mathcal{G} o grupo das matrizes estocásticas e \mathcal{G}^0 a componente conexa da identidade. Então, existe uma matriz geradora de uma cadeia de Markov a tempo contínuo (equivalentemente, uma L-matriz) Q_0 tal que a seguinte decomposição é válida:*

$$\mathcal{G}^0 = \bigcup_{t \in \mathbb{R}} e^{Q_0 t} \mathcal{H}.$$

Demonstração. Para realizar a demonstração, utilizemos a ideia proposta em [Summer]. Dado o grupo $\mathbb{R}_{>0}$ (dos números reais estritamente positivos) munido do operador multiplicação, e cuja identidade é dada pelo real 1. Existe um homomorfismo entre os grupos \mathcal{G}^0 e $\mathbb{R}_{>0}$, dado por

$$\begin{aligned} \psi : \mathcal{G} &\rightarrow \mathbb{R}_{>0} \\ X &\mapsto \det(X). \end{aligned}$$

Segue do Teorema do Isomorfismo, da teoria dos Grupos Clássicos (vide Teorema 3.1, Seção 3.1), que o quociente de \mathcal{G}^0 pelo núcleo de ψ é isomorfo a imagem de ψ , i.e., $\mathcal{G}^0 / \ker(\psi) \cong \text{Im}(\psi)$. Porém, segue do fato de que $\text{Im}(\psi) = \mathbb{R}_{>0}$ (Corolário 5.2), e do fato de que $\ker(\psi) = \mathcal{H}$, a seguinte relação

$$\mathcal{G}^0 / \mathcal{H} \cong \mathbb{R}_{>0}.$$

Podemos, também, escrever a componente conexa como união de coclasses

$$\mathcal{G}^0 = \bigcup_{Q \in \mathcal{G}^0} Q\mathcal{H}.$$

Finalmente, fixemos uma matriz $Q_0 \in T_1(\mathcal{G})$ com traço diferente de zero. Sabemos que $e^{Q_0} \in \mathcal{G}$ e $\det(e^{Q_0}) = e^{\text{tr}(Q_0)} > 0$, ou seja, $e^{Q_0} \in \mathcal{G}^0$. Ademais, dadas duas matrizes $M, P \in \mathcal{G}^0$, temos que suas coclasses serão iguais quando

$$M\mathcal{H} = P\mathcal{H} \iff \det(M) = \det(P),$$

devido ao homomorfismo ψ . Consequentemente, dada uma matriz $Q \in \mathcal{G}^0$ qualquer, tomando o real $t_0 = \ln(\det(Q))/\text{tr}(Q_0)$, obtemos

$$\det(e^{t_0 Q_0}) = e^{t_0 \text{tr}(Q_0)} = \det(Q),$$

ou seja, $Q\mathcal{H} = e^{(t_0 Q_0)}\mathcal{H}$ e, portanto, a igualdade abaixo é verdadeira

$$\mathcal{G}^0 = \bigcup_{t \in \mathbb{R}} e^{Q_0 t} \mathcal{H}.$$

□

Corolário 5.4. *Para toda matriz M , pertencente ao grupo das matrizes estocásticas, e de determinante positivo, existe um real $t_0 \in \mathbb{R}$, e uma matriz $P \in \mathcal{G}^0$ de determinante igual a um, tal que*

$$M = e^{Q_0 \cdot t_0} e^P,$$

onde Q_0 é uma L -matriz. Ou seja, através de uma matriz geradora de uma cadeia de Markov a tempo contínuo podemos gerar uma matriz de transição de uma cadeia de Markov a tempo discreto.

Demonstração. Segue imediatamente do Teorema acima. □

5.2.1 Aplicação do Teorema 5.3

Para ilustrar o Corolário 5.4, iremos gerar, por meio de simulações computacionais, matrizes de transição a partir de matrizes geradoras L, L_1 e L_2 , utilizando a igualdade $M = e^{L \cdot t_0} e^P$, onde t_0 é um real e P uma matriz de traço zero com linhas que também somam zero.

Exemplo 5.2.1. Sejam as matrizes geradoras

$$L_1 = \frac{1}{3} \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \quad e \quad L_2 = \begin{pmatrix} -0,2244 & 0,0719 & 1,1525 \\ 0,0309 & -0,1419 & 0,1110 \\ 0,1490 & 0,2695 & -0,4186 \end{pmatrix}.$$

Obtemos os seguintes resultados:

(i) Para a matriz geradora L_1 , consideremos o real $t_0 = 4,3329$ e a matriz

$$P = \begin{pmatrix} 0,1070 & 0,3447 & -0,4517 \\ 0,0462 & 0,1351 & -0,1813 \\ 0,9554 & -0,7133 & -0,2421 \end{pmatrix}.$$

Logo, realizando a transformação $M = e^{Lt_0}e^P$, obtemos a matriz de transição

$$M = \begin{pmatrix} 0,4415 & 0,5543 & 0,0042 \\ 0,3771 & 0,4091 & 0,2138 \\ 0,4257 & 0,3147 & 0,2596 \end{pmatrix}.$$

(ii) Agora, para a matriz geradora L_2 , se considerarmos o real $t_0 = 2,5490$ e a matriz

$$P = \begin{pmatrix} 0,0660 & 0,0392 & -0,1053 \\ 0,1219 & 0,0189 & -0,1408 \\ 0,0957 & -0,0107 & -0,0850 \end{pmatrix}$$

obtemos a seguinte matriz de transição

$$M = \begin{pmatrix} 0,6911 & 0,2205 & 0,0884 \\ 0,1928 & 0,7799 & 0,0274 \\ 0,2932 & 0,3957 & 0,3111 \end{pmatrix}.$$

Para encerrarmos nossos exemplos, calculemos a matriz de transição a partir de matrizes geradoras 4×4 .

Exemplo 5.2.2. Seja a matriz geradora

$$L = \begin{pmatrix} -2,0403 & 0,5717 & 0,7233 & 0,7453 \\ 0,9234 & -1,4091 & 0,1230 & 0,3627 \\ 0,5302 & 0,4849 & -1,6914 & 0,6763 \\ 0,9791 & 0,3013 & 0,1422 & -1,4226 \end{pmatrix}.$$

Tomando a constante como $t_0 = 4,5445$ e a matriz P como

$$P = \begin{pmatrix} 0,5348 & 0,6335 & 0,8027 & -1,9710 \\ -0,8745 & 0,1085 & 0,2200 & 0,5461 \\ 0,4856 & -1,3721 & 0,0726 & 0,8139 \\ -0,6206 & 0,7202 & 0,6163 & -0,7159 \end{pmatrix}$$

obtemos a matriz de transição

$$M = \begin{pmatrix} 0,3656 & 0,0374 & 0,5579 & 0,0391 \\ 0,3653 & 0,0377 & 0,5578 & 0,0392 \\ 0,3657 & 0,0373 & 0,5580 & 0,0391 \\ 0,3657 & 0,0373 & 0,5579 & 0,0391 \end{pmatrix}.$$

5.3 CONSTRUÍND O UMA MATRIZ ESTOCÁSTICA A PARTIR DO ESPAÇO TANGENTE

Dado o espaço tangente $T_1(\mathcal{G})$ do grupo das matrizes estocásticas calculado na seção 5.1, estamos interessados em obter uma matriz geradora de um processo de Markov a tempo contínuo através desta álgebra. Isto é possível se realizarmos os seguintes procedimentos:

- (i) Calcularmos o elemento do Casimir da álgebra de Lie $T_1(\mathcal{G})$, i.e., devemos achar um elemento no centro da álgebra $T_1(\mathcal{G})$;
- (ii) Definirmos uma coálgebra em $T_1(\mathcal{G})$ tal que tenhamos uma estrutura de biálgebra;
- (iii) Calcularmos o coproduto do elemento Casimir.

O resultado obtido do coproduto do Casimir nos dará, finalmente, uma matriz geradora. Para isto, trabalharemos com o espaço tangente do grupo das matrizes 3×3 com entradas nos reais:

$$T_1(\mathcal{G}) = \begin{pmatrix} -x_{12} - x_{13} & x_{12} & x_{13} \\ x_{21} & -x_{21} - x_{23} & x_{23} \\ x_{31} & x_{32} & -x_{31} - x_{32} \end{pmatrix}$$

Todavia, ocorre que nossa álgebra de Lie $T_1(\mathcal{G})$ não é semi-simples e, desta forma, não possuímos uma fórmula fechada para calcular o elemento Casimir da mesma. Porém, devido ao fato de $T_1(\mathcal{G})$, também, não ser uma álgebra solúvel, temos garantido pelo Teorema de Levi, descrito abaixo e cuja demonstração pode ser vista em [\[San Martin\]](#), a existência de uma subálgebra semi-simples \mathfrak{s} de $T_1(\mathcal{G})$.

Sendo assim, ao invés de trabalharmos com o espaço tangente $T_1(\mathcal{G})$ iremos trabalhar com sua subálgebra semi-simples \mathfrak{s} , pois a mesma, por ser uma álgebra semi-simples,

terá uma fórmula fechada para o cálculo de seu elemento do Casimir, facilitando nosso trabalho.

Teorema 5.5 (Teorema de Levi). *Seja \mathfrak{g} uma álgebra de Lie de dimensão finita e $\mathfrak{r}(\mathfrak{g})$ o seu radical ($\mathfrak{r}(\mathfrak{g})$ é igual ao maior ideal solúvel da álgebra \mathfrak{g}). Então, existe uma subálgebra semi-simples \mathfrak{s} o qual \mathfrak{g} se decompõe como soma direta da subálgebra \mathfrak{s} e do radical $\mathfrak{r}(\mathfrak{g})$.*

Com o auxílio computacional, utilizando o software [Octave] ¹, obtivemos a seguinte subálgebra semi-simples do espaço tangente $T_1(\mathcal{G})$:

$$\mathfrak{s} = \text{span} \left\{ X_1 := \begin{pmatrix} 0 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, X_2 := \begin{pmatrix} 1 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, X_3 := \begin{pmatrix} 1 & -1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \right\}$$

Observação. A álgebra de Lie \mathfrak{s} é isomórfica a álgebra $sl_2(\mathbb{R})$. O isomorfismo decorre do fato de \mathfrak{s} admitir a base

$$e = X_2, \quad h = -(X_1 + X_3), \quad f = -X_1,$$

cujas constantes estruturais valem

$$[e, f] = h, \quad [h, f] = -2f, \quad [h, e] = 2e$$

que são iguais as constantes estruturais de $sl_2(\mathbb{R})$, ou seja, $\mathfrak{s} \cong sl_2(\mathbb{R})$.

O elemento do Casimir da álgebra semi-simples \mathfrak{s} é dado, então, pela fórmula (vide [San Martin], pg. 95):

$$C = \sum_{i=1}^3 X_i Y_i,$$

o qual Y_i representa a base dual da base (X_1, X_2, X_3) , i.e., dada a forma bilinear não degenerada de Cartan-Killing para a álgebra semi-simples \mathfrak{s} :

$$\begin{aligned} \langle, \rangle : \mathfrak{s} \times \mathfrak{s} &\rightarrow \mathbb{R} \\ (X, Y) &\mapsto \text{Tr}(\text{Ad}(X) \circ \text{Ad}(Y)), \end{aligned}$$

com $\text{Ad}(X)$ representando a matriz da representação adjunta calculada no elemento $X \in \mathfrak{s}$, temos que Y_i é dado pela igualdade $\langle Y_i, X_j \rangle = \delta_{ij}$, com $i, j = 1, 2, 3$ e δ o símbolo

¹ O código do programa utilizado para obter a álgebra semi-simples \mathfrak{s} pode ser visto no Apêndice B

de Kronecker. Realizando os cálculos, obtemos para nossa álgebra \mathfrak{s} o seguinte elemento do Casimir:

$$\begin{aligned} C &= -\{(X_1)^2 + (X_3)^2 - 2(X_1X_2 + X_2X_1) + (X_1X_3 + X_3X_1)\} \\ &= \begin{pmatrix} -3 & 0 & 3 \\ 0 & -3 & 3 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Sabemos, vide Seção 4.3, que a álgebra universal envelopante das matrizes quadradas é uma biálgebra, com coproduto dado por:

$$\Delta(x) = x \otimes I_n + I_n \otimes x,$$

para x uma matriz quadrada qualquer. Como a álgebra \mathfrak{s} está imersa naturalmente na álgebra universal envelopante das matrizes 3×3 , podemos calcular o coproduto do elemento do Casimir C :

$$\Delta(C) = \begin{pmatrix} -6 & 0 & 3 & 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & -6 & 3 & 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & -3 & 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & -6 & 0 & 3 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & -6 & 3 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & -3 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & -3 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -3 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Portanto, como era esperado, a partir do elemento do Casimir C , da álgebra semi-simples \mathfrak{s} , obtivemos uma matriz geradora $L = \Delta(C)$ de um processo markoviano a tempo contínuo.

6

SISTEMAS DE PARTÍCULAS

Neste capítulo descreveremos o conceito de sistema de partículas e analisaremos como podemos gerar um sistema de partículas a partir da matriz de taxas L' obtida na seção 5.3. Nossa abordagem seguirá de perto o exposto em [Biane and Durret] e em [Belitsky and Schütz].

Como resultado obtivemos um sistema de partículas de dois sítios e três estados a partir da matriz L' . Em seguida, estendemos nosso processo estocástico para um sistema de N sítios, utilizando o gerador do sistema de duas partículas dado pela matriz L' e calculamos operadores de autodualidade para estes sistemas de partículas.

Abaixo segue o operador de autodualidade calculado por nós

$$D(\eta, \xi) = \prod_{l=1}^N (2 - \eta_l)(2 - \xi_l),$$

com η, ξ processos pertencentes ao espaço de estado Ω e η_l (resp. ξ_l) denotando o estado do vértice l relativo ao processo η (resp. ξ). Para finalizar, utilizando o trabalho de [Baez and Fong], provamos o Teorema de Noether para a versão estocástica.

Desta forma, neste capítulo estabelecemos uma conexão entre sistema de partículas de N sítios e a álgebra de Lie $sl_2(\mathbb{R})$, álgebra esta obtida na Seção 5.3 e que nos possibilitou o cálculo da matriz L' .

6.1 UMA BREVE INTRODUÇÃO

Dado um conjunto enumerável \mathbf{S} , cujos elementos $x \in \mathbf{S}$ serão denotados como o indivíduo ou a partícula na posição x , \mathbf{F} um conjunto finito, representando os possíveis estados de uma partícula (por exemplo, \mathbf{F} pode representar o fato de uma partícula estar ou não numa posição $x \in \mathbf{S}$), e $\xi : \mathbf{S} \rightarrow \mathbf{F}$ uma função que representa exatamente os estados de cada indivíduo do espaço em estudo. Dizemos, então, que um sistema

de partículas consiste no estudo da evolução temporal dos estados de cada ponto do espaço \mathbf{S} . Em outras palavras, dado o conjunto

$$\mathcal{D} = \{\eta : \mathbf{S} \rightarrow \mathbf{F}\}$$

de todas configurações possíveis de estados para nosso espaço, o estudo do sistema de partículas consiste na análise da função

$$\begin{aligned} \zeta^\zeta : \mathbb{R}_{\geq 0} &\rightarrow \mathcal{D} \\ t &\mapsto \zeta_t, \end{aligned}$$

onde $\zeta^\zeta(0) = \zeta$ e, para qualquer tempo positivo t , $\zeta^\zeta(t)$ representa os estados de cada partícula do nosso sistema. Por comodidade, denotaremos sempre ζ_t^ζ ao invés de $\zeta^\zeta(t)$.

Exemplo 6.1.1. Suponha o conjunto \mathbf{S} dado por

$$\mathbf{S} = \{-5, -4, -3, \dots, 3, 4, 5\}^2$$

e \mathbf{F} como sendo o conjunto $\mathbf{F} = \{0, 1\}$. Podemos interpretar \mathbf{F} como o fato de um indivíduo, situado numa posição $x \in \mathbf{S}$, ter ou não uma doença. (Aqui, 1 representa que o indivíduo está doente e 0 representa o indivíduo sadio.) O sistema de partículas definido pelo espaço \mathbf{S} e estado \mathbf{F} será representado pela função ζ_t^ζ , com

$$\zeta_t^\zeta(x) = \begin{cases} 1 & \text{se o indivíduo na posição } x \text{ e no tempo } t \text{ está doente;} \\ 0 & \text{se o indivíduo na posição } x \text{ e no tempo } t \text{ não está doente.} \end{cases}$$

A Figura 6 fornece um exemplo de uma função ζ_t^ζ , onde os pontos azuis representam indivíduos sadios enquanto que os pontos vermelhos representam pessoas doentes.

Dado um conjunto finito e não nulo $\mathcal{N} \subseteq \mathbf{S}$ (\mathbf{S} um espaço vetorial), dizemos que dois pontos $x, y \in \mathbf{S}$ são vizinhos se a relação $x - y \in \mathcal{N}$ for satisfeita. Para o Exemplo 6.1.1, podemos considerar o conjunto \mathcal{N} como o conjunto

$$\mathcal{N} = \{(1, 0), (0, 1), (-1, 0), (0, -1)\}.$$

Em geral, quando o espaço \mathbf{S} for um espaço vetorial normado, o conjunto \mathcal{N} pode ser definido como o conjunto

$$\mathcal{N}_r = \{x \in \mathbf{S}; \|x\| \leq r\},$$

com $\|\cdot\|$ uma norma.

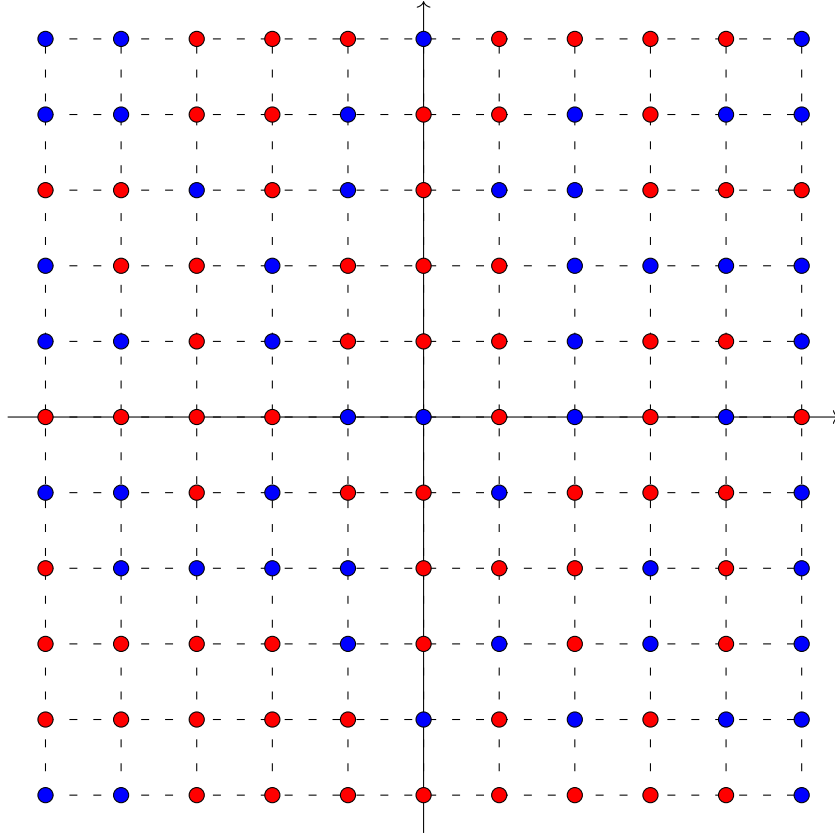


Figura 6: Exemplo de um sistema de partículas que modela a distribuição de uma certa doença no espaço dado por $S = \{-5, -4, \dots, 4, 5\}^2$. Aqui, pontos azuis representam pessoas saudáveis e os pontos vermelhos representam pessoas doentes.

No estudo de sistema de partículas estaremos interessados em estudar o comportamento da função ζ^ζ ao decorrer do tempo, isto é, como o estado das partículas irá mudar com o passar do tempo. O comportamento de ζ_t^ζ será dado pela seguinte regra: dado um ponto $x \in S$ qualquer, x irá para o estado $i \in F$ numa taxa $c_i(x, \zeta_t^\zeta)$, cujo valor depende dos estados dos pontos vizinhos de x , i.e.,

$$c_i(x, \zeta_t^\zeta) = g_i(\zeta_t^\zeta(x + z_0), \zeta_t^\zeta(x + z_1), \dots, \zeta_t^\zeta(x + z_n)),$$

com $g_i(\dots)$ uma função definida nos pontos $z_0, z_1, \dots, z_n \in \mathcal{N}$. Ou seja, a taxa com que um ponto x muda de estado está descrita em função de sua vizinhança. Suponhamos que $\zeta_t^\zeta(x) \neq i$, então a taxa c_i pode ser vista como o limite da razão da probabilidade de transição de estados dividido pelo tempo para ocorrer a transição:

$$\lim_{s \rightarrow 0} \frac{P(\zeta_{t+s}^\zeta(x) = i | \zeta_t^\zeta(x))}{s} = c_i(x, \zeta_t^\zeta).$$

6.2 DINÂMICA DO SISTEMA DE PARTÍCULAS

Para estudarmos a evolução de um sistema de partículas, com configuração inicial ξ , consideraremos as duas regras:

- (i) (*Tempo de saltos entre estados*) Todo ponto $x \in \mathbf{S}$ deverá esperar por instantes de tempo $T^{x,i}$ para “decidir” se irá mudar de seu estado atual para o estado i . Tais instantes de tempo $T^{x,i}$ serão dados por processos de Poisson independentes de taxa 1. Isto quer dizer que, caso no instante s tenhamos $\xi_s^\xi(x) = j$, x deverá esperar por um tempo aleatório $T^{x,i}$, cuja distribuição de probabilidade é dada por uma exponencial de parâmetro 1, para decidir se muda para o estado i ou não. A família de tempos de espera, dada por variáveis aleatórias de distribuição exponencial de taxa 1, é ilustrada pela Figura 7.

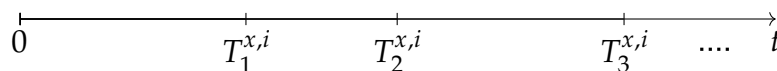


Figura 7: Representação do tempo de espera $T^{x,i}$, dado por um processo de Poisson PPP(1), para um ponto x mudar, ou não, para o estado i .

- (ii) (*Critério para um ponto x mudar de seu estado para um estado i*). Após um tempo de espera $t = T^{x,i}$, o ponto x muda de seu estado para o estado i se, dado um valor U , definido por uma distribuição uniforme $U[0, 1]$, tivermos

$$U < \frac{c_i(x, \xi_{t-}^\xi)}{\sum_{l \neq \xi_{t-}^\xi(x)} c_l(x, \xi_{t-}^\xi)},$$

com c_i as taxas de transição, definidas na seção anterior. Agora, caso a desigualdade acima não seja satisfeita, teremos que x mantém-se no seu estado atual. Um exemplo pode ser visto na Figura 8.

Para finalizarmos a construção de um sistema de partículas, enunciamos o argumento de Harris (1972), cuja ideia baseia-se em responder a seguinte pergunta: “Dado que um ponto $x \in \mathbf{S}$ muda de estado em função de taxas de transição $c_i(x, \xi_{t-}^\xi)$, cujos valores estão em função dos estados $\xi_{t-}^\xi(y)$ dados pelos pontos vizinhos de x . Queremos saber, então, se dado um instante de tempo t qualquer, por quantos pontos x tem sua taxa de

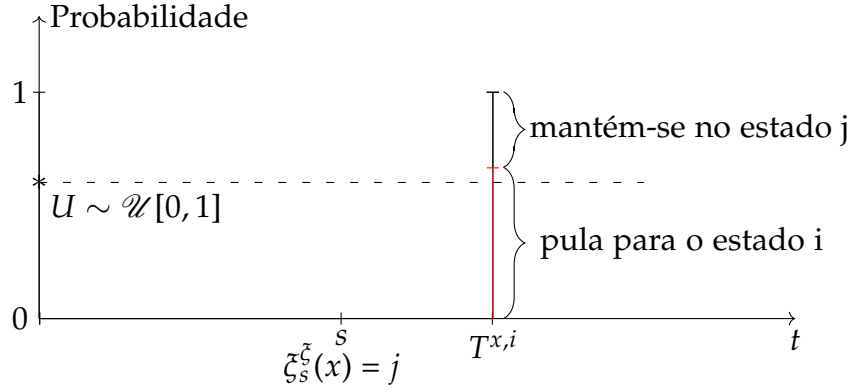


Figura 8: Exemplo de transição, no instante s , do ponto x no estado j para o estado i , no instante de tempo $t = s + T^{x,i}$.

transição influenciada”. Por influenciada, queremos dizer o seguinte: sabemos que a taxa de transição de x para o estado i é dada por

$$c_i(x, \xi_{t-}^{\xi}) = g_i(\xi_{t-}^{\xi}(z_0), \dots, \xi_{t-}^{\xi}(z_n)),$$

com z_0, \dots, z_n vizinhos de x cujos estados influenciam o valor da taxa de transição de x . Portanto, dizemos que z_0, \dots, z_n exercem influencia no valor da taxa de transição de x . Além do mais, temos, por exemplo, que

$$c_i(z_0, \xi_{t-}^{\xi}) = g_i(\xi_{t-}^{\xi}(w_0), \dots, \xi_{t-}^{\xi}(w_n)),$$

com w_0, \dots, w_n vizinhos de z_0 . Como os pontos w_0, \dots, w_n influenciam no valor da taxa de transição de z_0 que, por sua vez, também influencia na taxa de transição de x , obtemos, de forma indireta, que os pontos w_0, \dots, w_n influenciam a taxa de transição do ponto x . Desta forma, podemos ter uma sequência infinita de pontos que exercem influência nas taxas de transições do ponto x , como exemplificado na Figura 9.

Agora, observe que, caso os pontos w_0, \dots, w_n não tenham mudado de estado no intervalo de tempo $(0, t_0)$, então, obtemos que estes pontos já não exercem nenhum tipo de influencia na taxa de transição do ponto x . Seguindo esta intuição, enunciamos o argumento de Harry:

“Para compreendermos a evolução de nosso sistema de partículas devemos estudar o grafo aleatório, dado pelas partículas \mathbf{S} , e cujas arestas sejam dadas pela seguinte regra: fixado um tempo t_0 suficientemente pequeno, duas partículas $x, y \in \mathbf{S}$ estão conectadas por uma aresta se, ambas forem

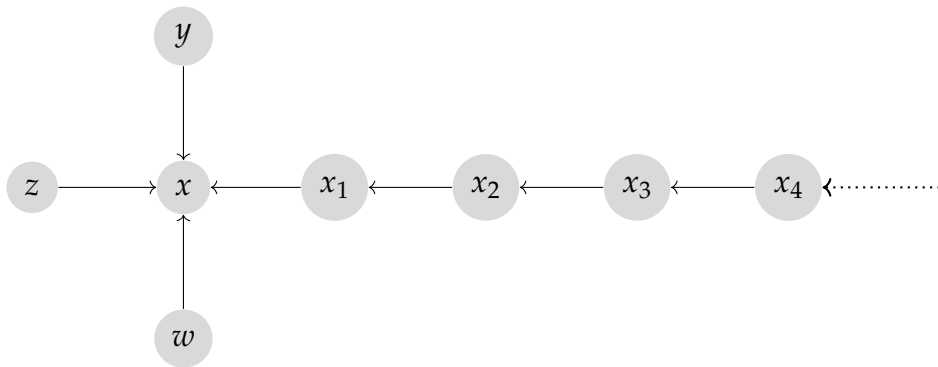


Figura 9: Exemplo de uma sequência infinita de pontos, x_1, x_2, \dots , que influenciam as taxas de transições do ponto x .

vizinhas, e existir um tempo $T^{x,i} < t_0$ em que ocorre transição de estados. Ou seja, caso a partícula x avalie mudar para o estado i , no tempo $T^{x,i} < t_0$, então os pontos vizinhos x e y serão conectados por uma aresta.”

Em posse do grafo aleatório, resultante do argumento de Harry, e pelo resultado (vide Teorema [Biane and Durrett]) o qual nos garante, com probabilidade 1, que as componentes conexas deste grafo serão finitas para um tempo t_0 suficientemente pequeno (vide Figura 10), obtemos uma metodologia para estudarmos a evolução de nosso sistema de partículas, a partir destas componentes conexas. Em outras palavras, como sabemos que para cada partícula x de uma componente conexa qualquer, definida no intervalo $[0, t_0)$, sofre influência de uma quantidade finita de partículas (pertencentes a mesma componente conexa que contém a partícula x), podemos descrever a evolução do estado desta partícula sem problemas. Além do mais, de forma interativa, podemos descrever a evolução do sistema de partículas nos intervalos de tempo $(0, t_0), [t_0, 2t_0), \dots [nt_0, (n+1)t_0), \dots$, e, portanto, temos bem estabelecida a dinâmica de nosso sistema de partículas para qualquer instante de tempo t .

6.3 PROPRIEDADES E O GERADOR DE UM SISTEMA DE PARTÍCULAS

Decorre da forma como estabelecemos a dinâmica de um sistema de partículas que a evolução temporal de seus estados, dados pela função ξ_t , é uma cadeia de Markov

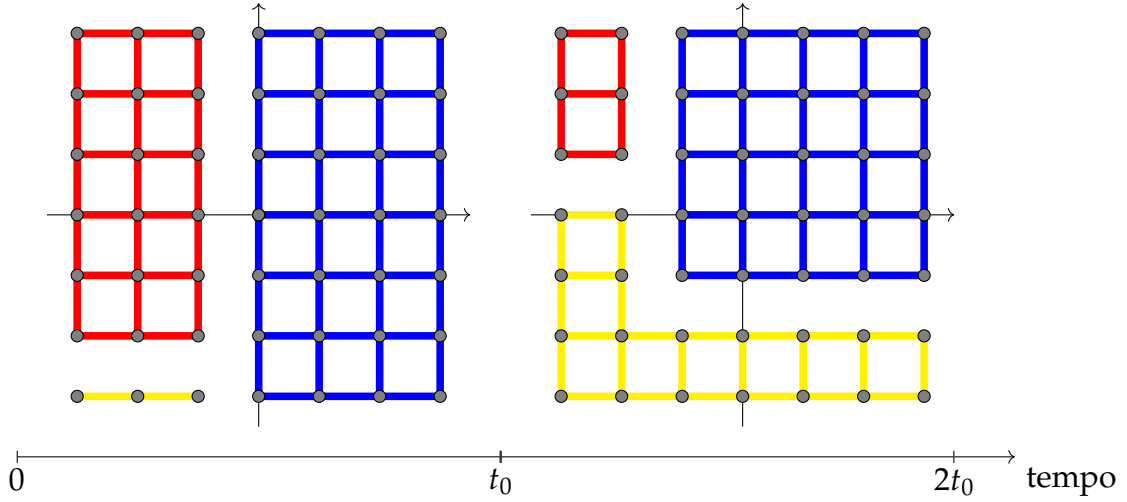


Figura 10: Exemplo de um sistema de partículas nos intervalos de tempo $[0, t_0)$ e $[t_0, 2t_0)$. Para cada intervalo de tempo, destacamos, com cores distintas três componentes conexas do grafo. Também ressaltamos que tais componentes conexas do grafo aleatório podem mudar com o tempo.

a tempo contínuo (a demonstração deste fato, como as demais demonstrações das afirmações feitas nesta seção, podem ser vistas em [Biane and Durrett] pg 119-122), i.e., sejam $s, t \in \mathbf{R}_{\geq 0}$ instantes de tempo quaisquer, então

$$P(\xi_{t+s} = \eta_1 | \xi_s = \eta_0) = P(\xi_t = \eta_1 | \xi_0 = \eta_0),$$

com $\eta_1, \eta_0 : \mathbf{S} \rightarrow \mathbf{F}$ funções representando os estados das partículas.

Seja $\mathcal{D} = \{\eta : \mathbf{S} \rightarrow \mathbf{F}\}$ o conjunto de todas configurações possíveis de estados para cada partícula do sistema e $f : \mathcal{D} \rightarrow \mathbb{R}$ uma função. Dizemos que f é uma função local ou cilíndrica se existir um conjunto finito $H \subseteq \mathbf{S}$ tal que, para $\eta, \xi \in \mathcal{D}$ e para todo $x \in \mathbf{S} \setminus H$, tivermos que $\eta(x) = \xi(x)$ implica $f(\eta) = f(\xi)$ (para uma maior explicação deste tipo de funções vide [Daniela]).

Seja \mathcal{D}_L o conjunto formado por todas as funções locais, $\Psi = \{\phi : \mathcal{D} \rightarrow \mathbb{R}\}$, e P_{ξ} a lei do processo com configuração inicial ξ . A partir destes conjuntos e de um instante de tempo t fixado, definimos o seguinte operador:

$$\begin{aligned} T_t : \mathcal{D}_L &\rightarrow \Psi \\ f &\mapsto T_t(f) : \mathcal{D} \rightarrow \mathbb{R} \\ \xi &\mapsto E_{\xi} T_t(f)(\xi), \end{aligned}$$

onde $E_{\xi}(\cdot)$ é a esperança com relação à medida de probabilidade P_{ξ} . Tal operador T_t satisfaz a condição de ser um semigrupo, i.e., o operador T_t satisfaz (vide [Baez and Fong]):

- (i) T_t é um operador linear, em relação ao tempo t , e preserva distribuições de probabilidades;
- (ii) $T_{s+t} = T_s T_t$, para $s, t \in \mathbb{R}_{\geq 0}$;
- (iii) $T_0 = Id$, com Id o operador identidade;
- (iv) T_t é contínua em relação ao instante de tempo t ;
- (v) Além do mais, para o operador T_t em específico, vale: T_t é um semigrupo de Feller, i.e., se f for uma função contínua em relação a topologia produto do conjunto \mathcal{D} , então T_t é uma função contínua.

Para encerrar esta seção, estabelecemos o gerador do semigrupo. Seja f uma função local. Então o operador \mathcal{L} dado por

$$\mathcal{L}f(\eta) = \sum_{\eta' \in \mathcal{D} \setminus \eta} C(\eta, \eta') (f(\eta') - f(\eta)),$$

com $C(\eta, \eta')$ a taxa de transição da configuração η para a η' , é chamado de gerador do processo markoviano a tempo contínuo de taxas $C(\eta, \eta')$. Em notação matricial, este gerador é dado pela L-matriz

$$\mathcal{L} = (L_{\eta\xi})_{\eta, \xi \in \mathcal{D}} \quad \text{com } L_{\eta\xi} = \begin{cases} C(\eta, \xi) & \text{se } \eta \neq \xi; \\ -\sum_{\eta' \neq \eta} C(\eta, \eta') & \text{c.c.} \end{cases}$$

6.4 ESTABELECENDO UM SISTEMA DE PARTÍCULAS A PARTIR DA MATRIZ DE TAXAS DA SEÇÃO 5.3

Dada a matriz de taxas

$$L' = \begin{pmatrix} -6 & 0 & 3 & 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & -6 & 3 & 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & -3 & 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & -6 & 0 & 3 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & -6 & 3 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & -3 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & -3 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -3 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

obtida na seção 5.3 a partir do coproduto do elemento do Casimir da álgebra universal envelopante $U(\mathfrak{s})$, resultante da álgebra de Lie

$$\mathfrak{s} = \text{span} \left\{ X_1 := \begin{pmatrix} 0 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, X_2 := \begin{pmatrix} 1 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, X_3 := \begin{pmatrix} 1 & -1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \right\},$$

iremos construir um sistema de partículas a partir desta matriz.

Nosso sistema de partículas será da seguinte forma: teremos duas partículas interagindo no sistema as quais podem assumir três estados distintos, ou seja, $S = \{0, 1\}$ e $F = \{0, 1, 2\}$. A partir deste sistema, teremos que as possíveis configurações de estados para cada partícula serão dadas por

$$\mathcal{D} = \{\eta : S \rightarrow F\} = \{(0, 0), (0, 1), (0, 2), (1, 0), (1, 1), (1, 2), (2, 0), (2, 1), (2, 2)\},$$

com $\eta = (0, 2)$, por exemplo, representando a configuração ao qual a partícula 0 está no estado 0 e a partícula 1 está no estado 2. Agora, ordenando o conjunto \mathcal{D} de acordo com a regra

$$i_2(\eta) = 1 + \sum_{k=1}^2 \eta(k)3^{k-1},$$

o qual pode ser vista em [Belitsky and Schütz], obtemos a seguinte ordenação das funções $\eta : S \rightarrow F$:

$$\eta_1 = (0, 0), \eta_2 = (1, 0), \eta_3 = (2, 0), \eta_4 = (0, 1), \eta_5 = (1, 1),$$

$$\eta_6 = (2, 1), \eta_7 = (0, 2), \eta_8 = (1, 2), \eta_9 = (2, 2).$$

Resulta desta ordem a cadeia de Markov a tempo contínuo dada pela matriz geradora:

$$L' = \begin{matrix} & \begin{matrix} (0,0) & (1,0) & (2,0) & (0,1) & (1,1) & (2,1) & (0,2) & (1,2) & (2,2) \end{matrix} \\ \begin{matrix} (0,0) \\ (1,0) \\ (2,0) \\ (0,1) \\ (1,1) \\ (2,1) \\ (0,2) \\ (1,2) \\ (2,2) \end{matrix} & \begin{pmatrix} -6 & 0 & 3 & 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & -6 & 3 & 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & -3 & 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & -6 & 0 & 3 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & -6 & 3 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & -3 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & -3 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -3 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{matrix}.$$

Portanto, pela matriz geradora acima, temos que, dado uma configuração do sistema de partículas qualquer, digamos η_t^η (com o sobrescrito indicando a configuração inicial e t o tempo), teremos, por exemplo, que

$$\lim_{s \rightarrow 0} \frac{P[\eta_{t+s}^\eta = (2,0) | \eta_t^\eta = (0,1)]}{s} = 3,$$

ou seja, a taxa de transição da configuração $(0,1)$ para a configuração $(2,0)$ vale 3.

Além do mais, podemos estender nosso sistema de duas partículas para um sistema de N -partículas. Para isto, utilizaremos a metodologia e a mesma notação apresentada em [Belitsky and Schütz], o qual estabelece que a matriz geradora para o sistema de N -partículas é dada por

$$L = \sum_{k=1}^{N-1} h_{k,k+1},$$

com $h_{k,k+1} = I_3^{\otimes(k-1)} \otimes L' \otimes I_3^{\otimes(N-k-1)}$. (I_3 representa a matriz identidade com $I_3^{\otimes 0} = 1$, 1 a identidade do corpo, e $I_3^{\otimes k} = I_3 \otimes \dots \otimes I_3$ k vezes.) Utilizando a base canônica de \mathbb{R}^3 dada pelos vetores coluna

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

e estabelecendo as matrizes auxiliares

$$\begin{aligned} a^+ &= e_1 \otimes e_2^T, & a^- &= e_2 \otimes e_1^T, & \hat{a} &= e_1 \otimes e_1^T \\ b^+ &= e_3 \otimes e_2^T, & b^- &= e_2 \otimes e_3^T, & \hat{b} &= e_3 \otimes e_3^T \\ c^+ &= e_1 \otimes e_3^T, & c^- &= e_3 \otimes e_1^T, & \hat{v} &= e_2 \otimes e_2^T \end{aligned} \quad (*)$$

que nada mais são que a base canônica do espaço vetorial das matrizes 3×3 , podemos escrever nossa matriz L' em termos destas matrizes:

$$\begin{aligned} L' = & -6(\hat{a} \otimes \hat{a} + \hat{a} \otimes \hat{v} + \hat{v} \otimes \hat{a} + \hat{v} \otimes \hat{v}) + 3(\hat{a} \otimes c^+ + \hat{a} \otimes b^- - \hat{a} \otimes \hat{b} + \hat{v} \otimes c^+ + \hat{v} \otimes b^- - \hat{v} \otimes \hat{b} \\ & + c^+ \otimes \hat{a} + b^- \otimes \hat{a} - \hat{b} \otimes \hat{a} + c^+ \otimes \hat{v} + b^- \otimes \hat{v} - \hat{b} \otimes \hat{v} + c^+ \otimes \hat{b} + b^- \otimes \hat{b} + \hat{b} \otimes c^+ + \hat{b} \otimes b^-). \end{aligned}$$

Agora, seja u, v alguma das matrizes definidas em $(*)$, temos que as matrizes

$$u_k = I_3^{\otimes(k-1)} \otimes u \otimes I_3^{\otimes(N-k)}, \quad v_{k+1} = I_3^{\otimes(k)} \otimes v \otimes I_3^{\otimes(N-k-1)}$$

satisfazem $u_k v_{k+1} = I_3^{\otimes(k-1)} \otimes (u \otimes v) \otimes I_3^{\otimes(N-k-1)}$. Sendo assim, o operador $h_{k,k+1} = I_3^{\otimes(k-1)} \otimes L' \otimes I_3^{\otimes(N-k-1)}$ pode ser escrito da seguinte forma

$$\begin{aligned} h_{k,k+1} = & -6(\hat{a}_k \hat{a}_{k+1} + \hat{a}_k \hat{v}_{k+1} + \hat{v}_k \hat{a}_{k+1} + \hat{v}_k \hat{v}_{k+1}) + 3(\hat{a}_k c_{k+1}^+ + \hat{a}_k b_{k+1}^- - \hat{a}_k \hat{b}_{k+1} + \hat{v}_k c_{k+1}^+ + \hat{v}_k b_{k+1}^- \\ & - \hat{v}_k \hat{b}_{k+1} + c_k^+ \hat{a}_{k+1} + b_k^- \hat{a}_{k+1} - \hat{b}_k \hat{a}_{k+1} + c_k^+ \hat{v}_{k+1} + b_k^- \hat{v}_{k+1} - \hat{b}_k \hat{v}_{k+1} + c_k^+ \hat{b}_{k+1} + b_k^- \hat{b}_{k+1} \\ & + \hat{b}_k c_{k+1}^+ + \hat{b}_k b_{k+1}^-) \end{aligned}$$

e, conseqüentemente, temos o gerador de um sistema de N -partículas

$$L = \sum_{k=1}^{N-1} h_{k,k+1}$$

bem definido.

6.5 AUTODUALIDADE

Nesta seção iremos estabelecer um operador de autodualidade D' para nosso sistema de duas partículas e, a partir do mesmo, definiremos um dual para o sistema de N -partículas cujo gerador foi estabelecido no final da seção anterior. Nossa metodologia será estabelecida de forma similar a construída em [Daniela]. Desta forma, iniciamos nossa discussão definindo o conceito de autodualidade:

Definição 6.1. Sejam as cadeias de Markov a tempo contínuo, e independentes, $(\eta_t)_{t \geq 0}$ e $(\xi_t)_{t \geq 0}$, com espaço de estados Ω , e $D : \Omega \times \Omega \rightarrow \mathbb{R}$ uma função mensurável. A função D é dita ser um operador de autodualidade entre os processos (η_t) e (ξ_t) se

$$E_\eta[D(\eta_t, \xi)] = E_\xi[D(\eta, \xi_t)]$$

para todo $t \geq 0$ e para todo par de processos markovianos $(\eta, \xi) \in \Omega^2$.

Desta definição temos a seguinte caracterização:

Proposição 6.1. *Seja Ω um espaço de estados, \mathcal{L} um gerador de um processo markoviano e $D : \Omega \times \Omega \rightarrow \mathbb{R}$ uma função. Então, se a função D pertence ao domínio do gerador, temos que D é um operador autodual se, e somente se,*

$$\mathcal{L}D(\eta, \cdot)(\xi) = \mathcal{L}D(\cdot, \xi)(\eta),$$

para todo $\eta, \xi \in \Omega$.

Demonstração. Vide [Daniela]. □

Observação. Relembremos que para todo gerador \mathcal{L} de um processo markoviano a tempo contínuo, e com espaço de estados Ω , existe uma matriz associada ao gerador, denotada por $L = (L_{\eta\xi})_{\eta, \xi \in \Omega}$, satisfazendo:

$$L_{\eta\xi} = \begin{cases} C(\eta, \xi) & \text{se } \eta \neq \xi; \\ -\sum_{\eta' \neq \eta} C(\eta', \eta) & \text{c.c.} \end{cases}$$

com $C(\eta, \xi)$ as taxas de transição da cadeia de Markov.

Observação. Note que a igualdade descrita na Proposição 6.1 pode ser reescrita como

$$\sum_{\sigma \in \Omega} \mathcal{L}(\xi, \sigma)D(\eta, \sigma) = \sum_{\sigma \in \Omega} \mathcal{L}(\eta, \sigma)D(\sigma, \xi),$$

para todo $\xi, \eta \in \Omega$. Note, também, que a igualdade acima é a mesma se considerarmos a matriz geradora L , associada ao gerador \mathcal{L} , i.e.,

$$\sum_{\sigma \in \Omega} L_{\xi, \sigma}D(\eta, \sigma) = \sum_{\sigma \in \Omega} L_{\eta, \sigma}D(\sigma, \xi).$$

Em posse desta definição, a Proposição abaixo (vide [Aidan Sudbury and Peter Lloyd]) será útil para estabelecermos operadores de autodualidade.

Proposição 6.2. Dado o espaço de estados finito Ω , a matriz geradora \mathbf{L} de uma cadeia de Markov a tempo contínuo e uma função $D : \Omega \times \Omega \rightarrow \mathbb{R}$. Se considerarmos a matriz $\mathbf{D} = (D_{\eta, \xi})_{\eta, \xi \in \Omega}$ como a matriz dada por $\mathbf{D}_{\eta, \xi} = D(\eta, \xi)$, então, teremos que a função D é um operador de autodualidade se, e somente se, a igualdade

$$\mathbf{L}\mathbf{D} = \mathbf{D}\mathbf{L}^T$$

for válida.

Desta maneira, para definirmos uma autodualidade para nosso sistema de duas partículas, dado pela matriz geradora

$$\mathbf{L}' = \begin{matrix} & \begin{matrix} (0,0) & (1,0) & (2,0) & (0,1) & (1,1) & (2,1) & (0,2) & (1,2) & (2,2) \end{matrix} \\ \begin{matrix} (0,0) \\ (1,0) \\ (2,0) \\ (0,1) \\ (1,1) \\ (2,1) \\ (0,2) \\ (1,2) \\ (2,2) \end{matrix} & \begin{pmatrix} -6 & 0 & 3 & 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & -6 & 3 & 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & -3 & 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & -6 & 0 & 3 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & -6 & 3 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & -3 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & -3 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -3 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{matrix},$$

devemos achar uma matriz \mathbf{D} tal que a igualdade

$$\mathbf{L}'\mathbf{D} = \mathbf{D}\mathbf{L}'^T$$

seja satisfeita. Resolvendo este sistema linear, é possível determinar o seguinte operador autodual para nosso sistema de duas partículas

$$D'((\eta_1, \eta_2), (\xi_1, \xi_2)) = \prod_{l=1}^2 (2 - \eta_l)(2 - \xi_l)$$

com $\eta = (\eta_1, \eta_2), \xi = (\xi_1, \xi_2)$ processos pertencentes ao espaço de estados de duas partículas. Para o operador D' , temos a seguinte representação matricial:

$$D' = \begin{matrix} & \begin{matrix} (0,0) & (1,0) & (2,0) & (0,1) & (1,1) & (2,1) & (0,2) & (1,2) & (2,2) \end{matrix} \\ \begin{matrix} (0,0) \\ (1,0) \\ (2,0) \\ (0,1) \\ (1,1) \\ (2,1) \\ (0,2) \\ (1,2) \\ (2,2) \end{matrix} & \left(\begin{array}{ccccccccc} 16 & 8 & 0 & 8 & 4 & 0 & 0 & 0 & 0 \\ 8 & 4 & 0 & 4 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 8 & 4 & 0 & 4 & 2 & 0 & 0 & 0 & 0 \\ 4 & 2 & 0 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right) \end{matrix}$$

que satisfaz $L'D = DL'^T$.

Agora, sendo que para um sistema de duas partículas, com gerador dado pela matriz L' , somos capazes de definir um operador de autodualidade D' , nosso objetivo, a seguir, será estender este operador para um sistema de N-partículas de gerador definido por

$$L = \sum_{k=1}^{N-1} h_{k,k+1},$$

com $h_{k,k+1} = I_3^{\otimes(k-1)} \otimes L' \otimes I_3^{\otimes(N-k-1)}$ (I_3 a matriz identidade 3×3). Uma forma natural, para o operador de autodualidade de um sistema de N-partículas, é a seguinte função:

$$D(\eta, \xi) = \prod_{l=1}^N (2 - \eta_l)(2 - \xi_l),$$

com $\eta = (\eta_1, \dots, \eta_N)$ e $\xi = (\xi_1, \dots, \xi_N)$ pertencentes ao espaço de estados de um sistema de N-partículas. Abaixo mostraremos que D é uma função de autodualidade.

Proposição 6.3. *Seja o sistema de N-partículas, de espaço de estados Ω , dado pelo gerador*

$$L = \sum_{k=1}^{N-1} h_{k,k+1},$$

com $h_{k,k+1} = \mathbb{1}^{\otimes(k-1)} \otimes L' \otimes \mathbb{1}^{\otimes(N-k-1)}$, onde L' é nossa matriz geradora, já definida no início desta seção, e D' o autodual definido para o sistema de duas partículas. Então, o operador D dado por

$$D(\eta, \xi) = \prod_{i=1}^N (2 - \eta_i)(2 - \xi_i),$$

com $\xi, \eta \in \Omega$, é um operador de autodualidade.

Demonstração. Observação, para não sobrecarregarmos a notação, denotaremos a entrada da matriz $L_{\eta\tilde{\zeta}}$ por apenas $L(\eta, \tilde{\zeta}) = L_{\eta\tilde{\zeta}}$ (respectivamente, fazemos o mesmo para as matrizes $h_{k,k+1}$).

Considerando o gerador \mathcal{L} associado a matriz geradora L , e \mathcal{L}' ao gerador referente a matriz L' , temos, para todos $\eta, \tilde{\zeta} \in \Omega$, a igualdade

$$\begin{aligned}\mathcal{L}D(\eta, \cdot)(\tilde{\zeta}) &= \sum_{\sigma \in \Omega} L(\tilde{\zeta}, \sigma)D(\eta, \sigma) \\ &= \sum_{k=1}^{N-1} \sum_{\sigma \in \Omega} h_{k,k+1}(\tilde{\zeta}, \sigma)D(\eta, \sigma) \\ &= \sum_{k=1}^{N-1} \sum_{\sigma \in \Omega} L'(\tilde{\zeta}^k, \sigma^k) \prod_{i=k}^{k+1} (2 - \eta_i)(2 - \sigma_i) \prod_{i \in \{1,2,\dots,N\} \setminus \{k,k+1\}} (2 - \eta_i)(2 - \tilde{\zeta}_i),\end{aligned}$$

com $\sigma^k = (\sigma_k, \sigma_{k+1})$ e $\tilde{\zeta}^k = (\tilde{\zeta}_1, \tilde{\zeta}_2)$. Chamando de

$$A_k(\eta, \tilde{\zeta}) = \prod_{i \in \{1,2,\dots,N\} \setminus \{k,k+1\}} (2 - \eta_i)(2 - \tilde{\zeta}_i),$$

obtemos

$$\mathcal{L}D(\eta, \cdot)(\tilde{\zeta}) = \sum_{k=1}^{N-1} A_k(\eta, \tilde{\zeta}) \sum_{\sigma \in \Omega} L'(\tilde{\zeta}^k, \sigma^k) D'(\eta^k, \sigma^k).$$

Pela autodualidade de D' , conseguimos a igualdade

$$\mathcal{L}D(\eta, \cdot)(\tilde{\zeta}) = \sum_{k=1}^{N-1} A_k(\eta, \tilde{\zeta}) \mathcal{L}' D'(\cdot, \tilde{\zeta}^k)(\eta^k).$$

Porém,

$$\begin{aligned}\mathcal{L}D(\cdot, \tilde{\zeta})(\eta) &= \sum_{\sigma \in \Omega} L(\eta, \sigma)D(\sigma, \tilde{\zeta}) \\ &= \sum_{k=1}^{N-1} \sum_{\sigma \in \Omega} h_{k,k+1}(\eta, \sigma)D(\sigma, \tilde{\zeta}) \\ &= \sum_{k=1}^{N-1} \sum_{\sigma \in \Omega} L'(\eta^k, \sigma^k) \prod_{i=k}^{k+1} (2 - \sigma_i)(2 - \tilde{\zeta}_i) \prod_{i \in \{1,2,\dots,N\} \setminus \{k,k+1\}} (2 - \eta_i)(2 - \tilde{\zeta}_i) \\ &= \sum_{k=1}^{N-1} A_k(\eta, \tilde{\zeta}) \sum_{\sigma \in \Omega} L'(\eta^k, \sigma^k) D'(\sigma^k, \tilde{\zeta}^k) \\ &= \sum_{k=1}^{N-1} A_k(\eta, \tilde{\zeta}) \mathcal{L}' D'(\cdot, \tilde{\zeta}^k)(\eta^k).\end{aligned}$$

Ou seja, $\mathcal{L}D(\cdot, \tilde{\zeta})(\eta) = \mathcal{L}D(\eta, \cdot)(\tilde{\zeta})$ e, portanto, D é um operador de autodualidade. \square

6.6 UMA VERSÃO DO TEOREMA DE NOETHER PARA O CASO ESTOCÁSTICO

Em mecânica quântica, o Teorema de Noether é utilizado para estabelecer relações entre medidas que se conservam com o tempo e suas simetrias (vide [Baez and Fong]). No entanto, este Teorema também possui uma variante para processos estocásticos, o qual estudaremos nesta seção. (Aqui, toda notação, quanto a ideia para demonstrar o teorema, foram retirados de [Baez and Fong].)

Antes de enunciarmos e demonstrarmos a versão estocástica do Teorema de Noether, devemos estabelecer algumas definições e um pouco de notação. Dado um espaço de probabilidade finito, denotado por $X = \{x_1, x_2, \dots, x_n\}$ e com medida de probabilidade $\phi : X \rightarrow [0, 1]$, dada pelo vetor

$$\phi = [\phi(x_1) \ \phi(x_2) \ \dots \ \phi(x_n)].$$

Dizemos que toda função que mapeia o espaço X nos reais, $O : X \rightarrow \mathbb{R}$, é um observável, com representação matricial dada por

$$O = (O_{ij}), \text{ com } O_{ij} = \delta_{ij} O(x_i)$$

e, realizaremos o abuso de notação, definindo o vetor formado pelos valores $O(x_i)$ pelo mesmo símbolo que representa a matriz do observável, i.e., denotaremos também por O o vetor

$$O = [O(x_1) \ O(x_2) \ \dots \ O(x_n)].$$

Ainda, em relação ao observável $O : X \rightarrow \mathbb{R}$, escreveremos sua média e segundo momento em função do produto linear canônico \langle, \rangle em \mathbb{R}^n :

$$E[O] = \sum_{i=1}^n O(x_i) \phi(x_i) = \langle O, \phi \rangle \quad \text{e} \quad E[O^2] = \sum_{i=1}^n O^2(x_i) \phi(x_i) = \langle O^2, \phi \rangle.$$

Para finalizar, seja $\pi = [\pi_1 \ \pi_2 \ \dots \ \pi_n]^T$ um vetor e $f : \mathbb{R} \rightarrow \mathbb{R}$ uma função qualquer. Denotaremos, em negrito, o vetor $f(\pi)$ como sendo o vetor

$$f(\pi) = [f(\pi_1) \ f(\pi_2) \ \dots \ f(\pi_n)].$$

Definição 6.2. Dada uma L-matriz¹ H , definimos o grafo orientado, chamado de grafo de transição, como o grafo cujos vértices são formados pelo conjunto $X = \{x_1, x_2, \dots, x_n\}$

¹ relembre que uma matriz $H \in \mathbb{R}^{n \times n}$ é chamada de L-matriz quando todas suas linhas somam zero e seus valores fora da diagonal são positivos.

e, dados dois pontos quaisquer $x_i, x_j \in X$, existe uma aresta conectando eles se $H_{ij} \neq 0$ ou $H_{ji} \neq 0$.

Uma componente conexa do grafo de transição será um grafo, cujos vértices são dados por um subconjunto $Y \subseteq X$ tal que, para dois pontos $x_i, x_j \in Y$ quaisquer, deverá existir uma sequência de pontos $x_i = x_{l_1}, x_{l_2}, \dots, x_{l_m}, x_{l_{m+1}} = x_j$, com $x_{l_1}, x_{l_2}, \dots, x_{l_{m+1}} \in Y$, satisfazendo

$$H_{l_p l_{p+1}} \neq 0 \text{ ou } H_{l_{p+1} l_p} \neq 0 \quad \text{para } p = 1, 2, \dots, m.$$

Antes de demonstrarmos a versão estocástica do Teorema de Noether, utilizaremos o seguinte resultado:

Teorema 6.4. *Dado um conjunto finito X de cardinalidade $|X| = n$, uma L -matriz $\mathbf{H} \in \mathbb{R}^{n \times n}$, e um observável $O : X \rightarrow \mathbb{R}$. As seguintes afirmações são equivalentes:*

(i) *O comutador da matriz do observável O pela L -matriz \mathbf{H} é nulo, i.e.,*

$$[O, \mathbf{H}] = O\mathbf{H} - \mathbf{H}O = \mathbf{0};$$

(ii) *Para todo polinômio $f : \mathbb{R} \rightarrow \mathbb{R}$ e para toda medida de probabilidade $\phi : X \rightarrow \mathbb{R}$ satisfazendo $\frac{d}{dt}\phi(t) = \phi(t)\mathbf{H}$, temos*

$$\frac{d}{dt}\langle f(\mathbf{O}), \phi(t) \rangle = 0;$$

(iii) *para toda medida de probabilidade $\phi : X \rightarrow \mathbb{R}$ satisfazendo $\frac{d}{dt}\phi(t) = \phi(t)\mathbf{H}$, temos*

$$\frac{d}{dt}\langle \mathbf{O}, \phi(t) \rangle = \frac{d}{dt}\langle \mathbf{O}^2, \phi(t) \rangle = 0;$$

(iv) *$O(x_i) = O(x_j)$ se x_i, x_j pertencem a mesma componente conexa do grafo construído a partir da L -matriz \mathbf{H} e do conjunto de pontos X .*

Demonstração. Mostremos que (i) \implies (ii). Observe que $[O, \mathbf{H}] = \mathbf{0}$ implica que $O\mathbf{H} = \mathbf{H}O$, ou seja para $i, j = 1, 2, \dots, n$ quaisquer, vale

$$\sum_{k=1}^n O_{ik} H_{kj} = \sum_{l=1}^n H_{il} O_{lj} \implies O_{ii} H_{ij} = H_{ij} O_{jj}.$$

Agora, calculemos a derivada do produto interno

$$\frac{d}{dt}\langle f(\mathbf{O}), \phi(t) \rangle = \langle f(\mathbf{O}), \frac{d}{dt}\phi(t) \rangle = \langle f(\mathbf{O}), \phi(t)\mathbf{H} \rangle.$$

Observe que

$$\phi(t)H = \left[\sum_{k=1}^n \phi_k H_{k1} \quad \dots \quad \sum_{k=1}^n \phi_k H_{kn} \right],$$

com $\phi_k = \phi(x_k)$. Desta forma, o produto interno escrito acima fica igual a

$$\langle f(O), \phi(t)H \rangle = \sum_{l=1}^n \left(f(O_l) \sum_{k=1}^n \phi_k H_{kl} \right),$$

com $O_l = O(x_l)$. Decorre de f ser um polinômio de grau m que $f(O_l) = \sum_{i=0}^m a_i O_l^i$, com a_i os coeficientes do polinômio. Assim, obtemos de $O_i H_{ij} = H_{ij} O_j$ que

$$f(O_l)H_{lk} = \sum_{i=0}^m a_i (O_l^i H_{lk}) = \sum_{i=0}^m a_i (O_k^i H_{lk}) = f(O_k)H_{lk}.$$

Portanto,

$$\langle f(O), \phi(t)H \rangle = \sum_{l=1}^n \left(\sum_{k=1}^n \phi_k f(O_l) H_{kl} \right) = \sum_{l=1}^n \left(\sum_{k=1}^n \phi_k f(O_k) H_{kl} \right) = \sum_{k=1}^n \left(\phi_k f(O_k) \sum_{l=1}^n H_{kl} \right),$$

ou seja, $\langle f(O), \phi(t)H \rangle = 0$, pois $\sum_{l=1}^n H_{kl} = 0$, já que H é uma L-matriz. Desta forma, concluimos que

$$\frac{d}{dt} \langle f(O), \phi(t) \rangle = 0.$$

Para mostrarmos que (ii) \implies (iii) basta tomarmos os polinômios $f_1(x) = x$ e $f(x) = x^2$, caso trivial. Agora, para obtermos que (iii) \implies (iv) notemos que se os pontos x_i, x_j pertencem a mesma componente conexa então existe uma sequência de pontos $x_i = x_{l_1}, x_{l_2}, \dots, x_{l_m}, x_{l_{m+1}} = x_j$ tal que

$$H_{l_p l_{p+1}} \neq 0 \text{ ou } H_{l_{p+1} l_p} \neq 0 \quad \text{para } p = 1, 2, \dots, m.$$

Desta forma, precisamos mostrar apenas que

$$O(x_{l_1}) = O(x_{l_2}), O(x_{l_2}) = O(x_{l_3}), \dots O(x_{l_m}) = O(x_{l_{m+1}}).$$

Para isto, utilizaremos o seguinte truque. Fixado um $p \in \{1, 2, 3, \dots, m\}$ e assumindo, sem perda de generalidade, que $H_{l_p l_{p+1}} \neq 0$, analisemos a somatória abaixo

$$\sum_{k=1}^n (O_k - O_{l_p})^2 H_{l_p k}.$$

A somatória pode ser reescrita como

$$\begin{aligned}\sum_{k=1}^n (O_k - O_{l_p})^2 H_{l_p k} &= \sum_{k=1}^n O_k^2 H_{l_p k} + \sum_{k=1}^n O_{l_p}^2 H_{l_p k} - 2 \sum_{k=1}^n O_k O_{l_p} H_{l_p k} \\ &= \sum_{k=1}^n O_k^2 H_{l_p k} + O_{l_p}^2 \sum_{k=1}^n H_{l_p k} - 2 O_{l_p} \sum_{k=1}^n O_k H_{l_p k}\end{aligned}$$

e, observando que o semigrupo $U_i(t) = e^{Ht} e_i$, com e_1, \dots, e_n a base canônica do \mathbb{R}^n , satisfaz $\frac{d}{dt} U_i(t) = H U_i(t)$, obtemos, pelo item (iii) e pelo fato de H ser uma L-matriz, as seguintes identidades

- $\sum_{k=1}^n H_{l_p k} = 0$;
- $0 = \frac{d}{dt} \langle O, U_i(t) \rangle = \langle O, H U_i(t) \rangle$ para todo instante de tempo t . Tomando $t = 0$ e $i = l_p$ temos $U_{l_p}(0) = I_n e_{l_p}$. Logo

$$0 = \langle O, H e_{l_p} \rangle = \sum_{k=1}^n O_k H_{l_p k};$$

- $0 = \frac{d}{dt} \langle O^2, U_i(t) \rangle = \langle O^2, H U_i(t) \rangle$ para todo instante de tempo t . Tomando $t = 0$ e $i = l_p$, temos, novamente, $U_{l_p}(0) = I_n e_{l_p}$. Logo

$$0 = \langle O^2, H e_{l_p} \rangle = \sum_{k=1}^n O_k^2 H_{l_p k}.$$

Desta forma, concluímos que $\sum_{k=1}^n (O_k - O_{l_p})^2 H_{l_p k} = 0$ e, portanto, como cada parcela desta soma é maior ou igual a zero pois $H_{ij} \geq 0$ para $i \neq j$ e para $i = j$ temos $(O_i - O_j) = 0$, então, concluímos que $(O_{l_{p+1}} - O_{l_p}) = 0$, já que assumimos que $H_{l_p l_{p+1}} \neq 0$. Sendo assim, para cada ponto da sequência $x_i = x_{l_1}, x_{l_2}, \dots, x_{l_m}, x_{l_{m+1}} = x_j$ obtemos

$$O(x_{l_1}) = O(x_{l_2}), O(x_{l_2}) = O(x_{l_3}), \dots O(x_{l_m}) = O(x_{l_{m+1}}).$$

A partir daí concluímos que $O_i = O_j$ se x_i, x_j pertencerem a mesma componente conexa.

Finalmente, para encerrarmos a demonstração do Teorema, devemos mostrar que (iv) \implies (i). Sabendo que $H_{ij} \neq 0$ significa que os pontos x_i, x_j pertencem a mesma componente conexa que, pelo item (iv), significa $O_i = O_j$, obtemos

$$OH = \left(\sum_{m=1}^n O_{km} H_{ml} \right)_{kl}, \text{ com } \sum_{m=1}^n O_{km} H_{ml} = \begin{cases} O_k H_{kl}, & \text{se } H_{kl} \neq 0 \\ 0, & \text{caso contrário.} \end{cases}$$

Pelo item (iv), a somatória acima fica igual a

$$\sum_{m=1}^n O_{km} H_{ml} = \begin{cases} O_l H_{kl}, & \text{se } H_{kl} \neq 0 \\ 0, & \text{caso contrário.} \end{cases}$$

Porém, o produto \mathbf{HO} é dado por

$$\mathbf{HO} = \left(\sum_{m=1}^n H_{km} O_{ml} \right)_{kl}, \text{ com } \sum_{m=1}^n H_{km} O_{ml} = \begin{cases} O_l H_{kl}, & \text{se } H_{kl} \neq 0 \\ 0, & \text{caso contrário.} \end{cases}$$

Ou seja, $[\mathbf{O}, \mathbf{H}] = \mathbf{0}$. □

Em posse deste Teorema enunciamos o Teorema de Noether (cuja demonstração resulta imediatamente do Teorema acima):

Teorema 6.5 (Teorema de Noether, versão estocástica). *Dado um conjunto finito X de cardinalidade $|X| = n$, uma L -matriz $\mathbf{H} \in \mathbb{R}^{n \times n}$, e um observável $O : X \rightarrow \mathbb{R}$. Então, a matriz do observável O comuta com a L -matriz \mathbf{H} , i.e., $[\mathbf{O}, \mathbf{H}]$, se e somente se para toda medida de probabilidade $\phi : X \rightarrow [0, 1]$ satisfazendo $\frac{d}{dt}\phi(t) = \mathbf{H}\phi(t)$ tivermos que o valor esperado e o segundo momento do observável O são invariantes com o tempo, i.e.,*

$$\frac{d}{dt} \langle \mathbf{O}, \phi(t) \rangle = \langle \mathbf{O}^2, \phi(t) \rangle = 0.$$

7

CONCLUSÃO

Inicialmente, neste trabalho, generalizamos (com relação a dimensão das matrizes em estudo), quando possível, os resultados propostos em [Summer] envolvendo os grupos de matrizes de Lie, relativo as matrizes estocásticas, e seus espaços tangentes. Como resultado principal, mostramos que a componente conexa da identidade, denotada por \mathcal{G}^0 admite a seguinte decomposição

$$\mathcal{G}^0 = \bigcup_{t \in \mathbb{R}} e^{Qt} \mathcal{H},$$

para Q uma L-matriz, de dimensão $n \times n$ ($n \geq 2$), e \mathcal{H} um subgrupo normal da componente conexa. Além do mais, mostramos que a componente conexa \mathcal{G}^0 é igual a todas as matrizes estocásticas $n \times n$ de determinante positivo.

Em seguida, utilizando os resultados propostos em [Redig], obtivemos uma matriz geradora L' a partir do espaço tangente $T_1(\mathcal{G})$ do grupo das matrizes de Lie relativo as matrizes estocásticas de dimensão 3×3 . O processo para obtenção da matriz L' baseou-se em, primeiro, calcular uma subálgebra semi-simples \mathfrak{s} de $T_1(\mathcal{G})$, o qual garante uma fórmula fechada para o cálculo de seu elemento do Casimir C . Em posse da matriz C , estabelecemos a álgebra universal envelopante de \mathfrak{s} e, através do coproduto

$$\Delta(x) = I_n \otimes x + x \otimes I_n,$$

com I_n a matriz identidade $n \times n$, calculamos a matriz L' , dada por $L' = \Delta(C)$.

Na parte final do trabalho estabelecemos um sistema de partículas de dois sítios, cujo gerador é definido pela matriz geradora L' , e construímos um operador de autodualidade para o sistema. Em seguida, mostramos ser possível estender o sistema de duas para N partículas e, conseqüentemente, estendemos a função de autodualidade anterior para o nosso novo sistema de N partículas.

APÊNDICE

A. DEMONSTRAÇÃO DA PROPOSIÇÃO 5.1

1 NOTAÇÃO

Seja $GL(n+1, \mathbb{R})$ o grupo linear geral das matrizes $(n+1) \times (n+1)$. Denotemos por $\mathbf{W}^{(1)}, \dots, \mathbf{W}^{(n)} \in GL(n+1, \mathbb{R})$ as matrizes que podem ser escritas da seguinte forma:

$$\mathbf{W}^{(1)} = \begin{pmatrix} 1-a_1 & a_1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ b_1 & 1-b_1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{pmatrix},$$

$$\mathbf{W}^{(2)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1-a_2 & a_2 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & b_2 & 1-b_2 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{pmatrix},$$

$$\mathbf{W}^{(3)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1-a_3 & a_3 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & b_3 & 1-b_3 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{pmatrix},$$

$$\vdots$$

$$\mathbf{W}^{(n)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1-a_n & a_n \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & b_n & 1-b_n \end{pmatrix},$$

com $a_1, \dots, a_n, b_1, \dots, b_n$ valores reais.

Em outras palavras, as matrizes $\mathbf{W}^{(1)}, \dots, \mathbf{W}^{(n)}$ podem ser vistas como as matrizes de blocos (aqui, a matriz $\mathbf{I}_{k \times l}$ denota a matriz identidade, de k linhas e l colunas, enquanto que $\mathbf{0}_{k \times l}$ denota a matriz de zeros, com k linhas e l colunas):

$$\mathbf{W}^{(1)} = \begin{pmatrix} \mathbf{P}_{2 \times 2}^{(1)} & \mathbf{0}_{2 \times (n-1)} \\ \mathbf{0}_{(n-1) \times (n-1)} & \mathbf{I}_{(n-1) \times (n-1)} \end{pmatrix},$$

$$\mathbf{W}^{(i)} = \begin{pmatrix} \mathbf{I}_{(i-1) \times (i-1)} & \mathbf{0}_{(i-1) \times 2} & \mathbf{0}_{(i-1) \times (n-i)} \\ \mathbf{0}_{2 \times (i-1)} & \mathbf{P}_{2 \times 2}^{(i)} & \mathbf{0}_{2 \times (n-i)} \\ \mathbf{0}_{(n-i) \times (i-1)} & \mathbf{0}_{(n-i) \times 2} & \mathbf{I}_{(n-i) \times (n-i)} \end{pmatrix},$$

com $i \in \{2, 3, \dots, n-1\}$ e

$$\mathbf{W}^{(n)} = \begin{pmatrix} \mathbf{I}_{(n-1) \times (n-1)} & \mathbf{0}_{(n-1) \times 2} \\ \mathbf{0}_{2 \times (n-1)} & \mathbf{P}_{2 \times 2}^{(n)} \end{pmatrix},$$

com $\mathbf{P}_{2 \times 2}^{(1)}, \dots, \mathbf{P}_{2 \times 2}^{(n)}$ matrizes estocásticas 2×2 .

Dizemos que uma matriz \mathbf{X} é do tipo $\mathbf{W}^{(i)}$, para $i \in \{1, \dots, n\}$, se \mathbf{X} puder ser escrita como $\mathbf{X} = \mathbf{W}^{(i)}$.

2 DEMONSTRAÇÃO

Seja \mathcal{G}_n o grupo das matrizes estocásticas $n \times n$ dado por

$$\mathcal{G}_n = \{\mathbf{X} \in GL(n, \mathbb{R}); \mathbf{X} \cdot \mathbf{1} = \mathbf{1}\},$$

com $\mathbf{1} = [1, 1, \dots, 1]^T$ o vetor cujas entradas são os inteiros 1. Mostremos a seguinte Proposição:

Proposição. *O conjunto das matrizes estocásticas de determinante positivo*

$$\{\mathbf{X} \in \mathcal{G}_n; \det(\mathbf{X}) > 0\}$$

é conexo por caminhos.

Antes de provarmos a Proposição acima, provemos o seguinte Lema:

Lema. *Seja $\mathbf{Q} \in \mathcal{G}_{(n+1)}$ uma matriz de determinante positivo e $n \geq 2$. Então, existe um sequência finita de matrizes de determinante positivo $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_m$ tal que*

$$\mathbf{Q} \cdot \mathbf{X}_1 \cdot \mathbf{X}_2 \cdots \mathbf{X}_m = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ y_{2,1} & y_{2,2} & \cdots & y_{2,n+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n+1,1} & y_{n+1,2} & \cdots & y_{n+1,n+1} \end{pmatrix}.$$

Além do mais, as matrizes $\mathbf{X}_i, i \in \{1, \dots, m\}$, são do tipo

$$\mathbf{X}_i = \mathbf{W}^{(j)},$$

para algum $j \in \{1, \dots, n\}$.

Demonstração. Tome $\mathbf{Q} \in \mathcal{G}_{(n+1)}$ ($n \geq 2$) a matriz de determinante positivo

$$\mathbf{Q} = \begin{pmatrix} Q_{1,1} & Q_{1,2} & \cdots & Q_{1,n-1} & Q_{1,n} & Q_{1,n+1} \\ Q_{2,1} & Q_{2,2} & \cdots & Q_{2,n-1} & Q_{2,n} & Q_{2,n+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ Q_{n+1,1} & Q_{n+1,n+1} & \cdots & Q_{n+1,n-1} & Q_{n+1,n} & Q_{n+1,n+1} \end{pmatrix}.$$

Nosso primeiro passo é mostrar que existe as matrizes \mathbf{X}_1 e \mathbf{X}_2 tal que

$$\mathbf{Q} \cdot \mathbf{X}_1 \cdot \mathbf{X}_2 = \mathbf{Q}^{(1)},$$

com

$$\mathbf{Q}^{(1)} = \begin{pmatrix} Q_{1,1}^{(1)} & Q_{1,2}^{(1)} & \cdots & Q_{1,n}^{(1)} & 0 \\ Q_{2,1}^{(1)} & Q_{2,2}^{(1)} & \cdots & Q_{2,n+1}^{(1)} & Q_{2,n+1}^{(1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ Q_{n+1,1}^{(1)} & Q_{n+1,2}^{(1)} & \cdots & Q_{n+1,n+1}^{(1)} & Q_{n+1,n+1}^{(1)} \end{pmatrix}.$$

Para mostrarmos esta igualdade devemos observar as seguintes possibilidades:

- (i) $Q_{1,n} + Q_{1,n+1} > 0$.

Neste caso, tome a matriz \mathbf{X}_1 dada por

$$\mathbf{X}_1 = \mathbf{W}^{(n)} = \begin{pmatrix} \mathbf{I}_{(n-1) \times (n-1)} & \mathbf{0}_{(n-1) \times 2} \\ \mathbf{0}_{2 \times (n-1)} & \mathbf{P}_{2 \times 2}^{(n)} \end{pmatrix},$$

com

$$\mathbf{P}_{2 \times 2}^{(n)} = \begin{pmatrix} 1 + Q_{1,n+1} & -Q_{1,n+1} \\ 1 - Q_{1,n} & Q_{1,n} \end{pmatrix}$$

e $\mathbf{X}_2 = \mathbf{I}_{(n+1) \times (n+1)}$. Desta forma,

$$\mathbf{Q} \cdot \mathbf{X}_1 \cdot \mathbf{X}_2 = \mathbf{Q}^{(1)} = \begin{pmatrix} Q_{1,1}^{(1)} & Q_{1,2}^{(1)} & \cdots & Q_{1,n}^{(1)} & 0 \\ Q_{2,1}^{(1)} & Q_{2,2}^{(1)} & \cdots & Q_{2,n+1}^{(1)} & Q_{2,n+1}^{(1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ Q_{n+1,1}^{(1)} & Q_{n+1,2}^{(1)} & \cdots & Q_{n+1,n+1}^{(1)} & Q_{n+1,n+1}^{(1)} \end{pmatrix},$$

com $\det(\mathbf{X}_1) = Q_{1,n} + Q_{1,n+1} > 0$.

- (ii) $Q_{1,n} + Q_{1,n+1} < 0$.

Para esta desigualdade, temos algo similar ao exposto acima no item (i). Tome a matriz \mathbf{X}_1 dada por

$$\mathbf{X}_1 = \mathbf{W}^{(n)} = \begin{pmatrix} \mathbf{I}_{(n-1) \times (n-1)} & \mathbf{0}_{(n-1) \times 2} \\ \mathbf{0}_{2 \times (n-1)} & \mathbf{P}_{2 \times 2}^{(n)} \end{pmatrix},$$

com

$$\mathbf{P}_{2 \times 2}^{(n)} = \begin{pmatrix} 1 - Q_{1,n+1} & Q_{1,n+1} \\ 1 + Q_{1,n} & -Q_{1,n} \end{pmatrix}$$

e $\mathbf{X}_2 = \mathbf{I}_{(n+1) \times (n+1)}$. Desta forma,

$$\mathbf{Q} \cdot \mathbf{X}_1 \cdot \mathbf{X}_2 = \mathbf{Q}^{(1)} = \begin{pmatrix} Q_{1,1}^{(1)} & Q_{1,2}^{(1)} & \cdots & Q_{1,n}^{(1)} & 0 \\ Q_{2,1}^{(1)} & Q_{2,2}^{(1)} & \cdots & Q_{2,n+1}^{(1)} & Q_{2,n+1}^{(1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ Q_{n+1,1}^{(1)} & Q_{n+1,2}^{(1)} & \cdots & Q_{n+1,n+1}^{(1)} & Q_{n+1,n+1}^{(1)} \end{pmatrix},$$

com $\det(\mathbf{X}_1) = -(Q_{1,n} + Q_{1,n+1}) > 0$.

(iii) $Q_{1,n} + Q_{1,n+1} = 0$.

Dada a restrição $Q_{1,n} = -Q_{1,n+1}$, consideramos a matriz \mathbf{X}_1 dada por

$$\mathbf{X}_1 = \mathbf{W}^{(n-1)} = \begin{pmatrix} \mathbf{I}_{(n-2) \times (n-2)} & \mathbf{0}_{(n-2) \times 2} & \mathbf{0}_{(n-2) \times 1} \\ \mathbf{0}_{2 \times (n-2)} & \mathbf{P}_{2 \times 2}^{(n-1)} & \mathbf{0}_{2 \times 1} \\ \mathbf{0}_{1 \times (n-2)} & \mathbf{0}_{1 \times 2} & \mathbf{I}_{1 \times 1} \end{pmatrix},$$

com

$$\mathbf{P}_{2 \times 2}^{(n-1)} = \begin{pmatrix} 1 & 0 \\ 0,5 & 0,5 \end{pmatrix}.$$

Logo, a matriz

$$\tilde{\mathbf{Q}} := \mathbf{Q} \cdot \mathbf{X}_1$$

satisfaz a condição de que $\tilde{\mathbf{Q}}_{1,n} + \tilde{\mathbf{Q}}_{1,n+1} \neq 0$. Portanto, aplicando o item (i) ou (ii) para a matriz $\tilde{\mathbf{Q}}$, obtemos uma matriz \mathbf{X}_2 do tipo $\mathbf{W}^{(n)}$, i.e.,

$$\mathbf{X}_2 = \mathbf{W}^{(n)} = \begin{pmatrix} \mathbf{I}_{(n-1) \times (n-1)} & \mathbf{0}_{(n-1) \times 2} \\ \mathbf{0}_{2 \times (n-1)} & \mathbf{P}_{2 \times 2}^{(n)} \end{pmatrix}$$

tal que

$$\mathbf{Q}^{(1)} = \tilde{\mathbf{Q}} \cdot \mathbf{X}_2 = \mathbf{Q} \cdot \mathbf{X}_1 \cdot \mathbf{X}_2 = \begin{pmatrix} Q_{1,1}^{(1)} & Q_{1,2}^{(1)} & \cdots & Q_{1,n}^{(1)} & 0 \\ Q_{2,1}^{(1)} & Q_{2,2}^{(1)} & \cdots & Q_{2,n}^{(1)} & Q_{2,n+1}^{(1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ Q_{n+1,1}^{(1)} & Q_{n+1,2}^{(1)} & \cdots & Q_{n+1,n}^{(1)} & Q_{n+1,n+1}^{(1)} \end{pmatrix}.$$

Dada a matriz $\mathbf{Q}^{(1)}$ realizamos o mesmo procedimento feito acima. Ou seja, dependendo do valor de $\mathbf{Q}_{1,n}^{(1)} + \mathbf{Q}_{1,n-1}^{(1)}$, obtemos matrizes $\mathbf{X}_3, \mathbf{X}_4$ tal que

$$\mathbf{Q}^{(1)} \cdot \mathbf{X}_3 \cdot \mathbf{X}_4 := \mathbf{Q}^{(2)},$$

com

$$\mathbf{Q}^{(2)} = \begin{pmatrix} Q_{1,1}^{(2)} & Q_{1,2}^{(2)} & \cdots & Q_{1,n-1}^{(2)} & 0 & 0 \\ Q_{2,1}^{(2)} & Q_{2,2}^{(2)} & \cdots & Q_{2,n-1}^{(2)} & Q_{2,n}^{(2)} & Q_{2,n+1}^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ Q_{n+1,1}^{(2)} & Q_{n+1,2}^{(2)} & \cdots & Q_{n+1,n-1}^{(2)} & Q_{n+1,n}^{(2)} & Q_{n+1,n+1}^{(2)} \end{pmatrix}.$$

Para checarmos esta igualdade devemos, novamente, observar as seguintes possibilidades (mesmo o procedimento da demonstração sendo similar ao anterior, iremos checar os três casos possíveis para ilustrar a ideia dos cálculos):

(i) $\mathbf{Q}_{1,n}^{(1)} + \mathbf{Q}_{1,n-1}^{(1)} > 0$.

Tome a matriz

$$\mathbf{X}_3 = \mathbf{W}^{(n-1)} = \begin{pmatrix} \mathbf{I}_{(n-2) \times (n-2)} & \mathbf{0}_{(n-2) \times 2} & \mathbf{0}_{(n-2) \times 1} \\ \mathbf{0}_{2 \times (n-2)} & \mathbf{P}_{2 \times 2}^{(n-1)} & \mathbf{0}_{2 \times 1} \\ \mathbf{0}_{1 \times (n-2)} & \mathbf{0}_{1 \times 2} & \mathbf{I}_{1 \times 1} \end{pmatrix},$$

com

$$\mathbf{P}_{2 \times 2}^{(n-1)} = \begin{pmatrix} 1 + Q_{1,n+1} & -Q_{1,n+1} \\ 1 - Q_{1,n} & Q_{1,n} \end{pmatrix}$$

e $\mathbf{X}_4 = \mathbf{I}_{(n+1) \times (n+1)}$. Logo,

$$\mathbf{Q}^{(1)} \cdot \mathbf{X}_3 \cdot \mathbf{X}_4 = \mathbf{Q}^{(2)} = \begin{pmatrix} Q_{1,1}^{(2)} & Q_{1,2}^{(2)} & \cdots & Q_{1,n-1}^{(2)} & 0 & 0 \\ Q_{2,1}^{(2)} & Q_{2,2}^{(2)} & \cdots & Q_{2,n-1}^{(2)} & Q_{2,n}^{(2)} & Q_{2,n+1}^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ Q_{n+1,1}^{(2)} & Q_{n+1,2}^{(2)} & \cdots & Q_{n+1,n-1}^{(2)} & Q_{n+1,n}^{(2)} & Q_{n+1,n+1}^{(2)} \end{pmatrix},$$

com $\det(\mathbf{X}_3) = Q_{1,n-1}^{(1)} + Q_{1,n} > 0$.

(ii) $\mathbf{Q}_{1,n}^{(1)} + \mathbf{Q}_{1,n-1}^{(1)} < 0$.

Para esta desigualdade tome a matriz \mathbf{X}_3 dada por

$$\mathbf{X}_3 = \mathbf{W}^{(n-1)} = \begin{pmatrix} \mathbf{I}_{(n-2) \times (n-2)} & \mathbf{0}_{(n-2) \times 2} & \mathbf{0}_{(n-2) \times 1} \\ \mathbf{0}_{2 \times (n-2)} & \mathbf{P}_{2 \times 2}^{(n-1)} & \mathbf{0}_{2 \times 1} \\ \mathbf{0}_{1 \times (n-2)} & \mathbf{0}_{1 \times 2} & \mathbf{I}_{1 \times 1} \end{pmatrix},$$

com

$$\mathbf{P}_{2 \times 2}^{(n-1)} = \begin{pmatrix} 1 - Q_{1,n+1} & Q_{1,n+1} \\ 1 + Q_{1,n} & -Q_{1,n} \end{pmatrix}$$

e $\mathbf{X}_4 = \mathbf{I}_{(n+1) \times (n+1)}$. Sendo assim,

$$\mathbf{Q}^{(1)} \cdot \mathbf{X}_3 \cdot \mathbf{X}_4 = \mathbf{Q}^{(2)} = \begin{pmatrix} Q_{1,1}^{(2)} & Q_{1,2}^{(2)} & \cdots & Q_{1,n-1}^{(2)} & 0 & 0 \\ Q_{2,1}^{(2)} & Q_{2,2}^{(2)} & \cdots & Q_{2,n-1}^{(2)} & Q_{2,n}^{(2)} & Q_{2,n+1}^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ Q_{n+1,1}^{(2)} & Q_{n+1,2}^{(2)} & \cdots & Q_{n+1,n-1}^{(2)} & Q_{n+1,n}^{(2)} & Q_{n+1,n+1}^{(2)} \end{pmatrix},$$

com $\det(\mathbf{X}_3) = -(Q_{1,n-1}^{(1)} + Q_{1,n}^{(1)}) > 0$.

(iii) $\mathbf{Q}_{1,n}^{(1)} + \mathbf{Q}_{1,n-1}^{(1)} = 0$.

Dada a restrição $Q_{1,n-1}^{(1)} = -Q_{1,n}^{(1)}$, tomamos a matriz

$$\mathbf{X}_3 = \mathbf{W}^{(n-2)} = \begin{pmatrix} \mathbf{I}_{(n-3) \times (n-3)} & \mathbf{0}_{(n-3) \times 2} & \mathbf{0}_{(n-3) \times 2} \\ \mathbf{0}_{2 \times (n-3)} & \mathbf{P}_{2 \times 2}^{(n-2)} & \mathbf{0}_{2 \times 2} \\ \mathbf{0}_{2 \times (n-3)} & \mathbf{0}_{2 \times 2} & \mathbf{I}_{2 \times 2} \end{pmatrix},$$

com

$$\mathbf{P}_{2 \times 2}^{(n-2)} = \begin{pmatrix} 1 & 0 \\ 0,5 & 0,5 \end{pmatrix}.$$

Portanto, a matriz

$$\tilde{\mathbf{Q}} := \mathbf{Q}^{(1)} \cdot \mathbf{X}_3$$

satisfaz a condição de que $\tilde{\mathbf{Q}}_{1,n-1} + \tilde{\mathbf{Q}}_{1,n} \neq 0$. Consequentemente, aplicando o item (i) ou (ii) para a matriz $\tilde{\mathbf{Q}}$, obtemos uma matriz \mathbf{X}_4 do tipo $\mathbf{W}^{(n-1)}$

$$\mathbf{X}_2 = \mathbf{W}^{(n)} = \begin{pmatrix} \mathbf{I}_{(n-1) \times (n-1)} & \mathbf{0}_{(n-1) \times 2} \\ \mathbf{0}_{2 \times (n-1)} & \mathbf{P}_{2 \times 2}^{(n)} \end{pmatrix}$$

tal que

$$\mathbf{Q}^{(2)} := \tilde{\mathbf{Q}} \cdot \mathbf{X}_4 = \mathbf{Q}^{(1)} \cdot \mathbf{X}_3 \cdot \mathbf{X}_4 = \begin{pmatrix} Q_{1,1}^{(2)} & Q_{1,2}^{(2)} & \cdots & Q_{1,n-1}^{(2)} & 0 & 0 \\ Q_{2,1}^{(2)} & Q_{2,2}^{(2)} & \cdots & Q_{2,n-1}^{(2)} & Q_{2,n}^{(2)} & Q_{2,n+1}^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ Q_{n+1,1}^{(2)} & Q_{n+1,2}^{(2)} & \cdots & Q_{n+1,n-1}^{(2)} & Q_{n+1,n}^{(2)} & Q_{n+1,n+1}^{(2)} \end{pmatrix},$$

Podemos continuar esse processo indutivamente obtendo a sequência de matrizes: $\mathbf{Q}^{(1)}, \mathbf{Q}^{(2)}, \dots, \mathbf{Q}^{(n-1)}$. A matriz $\mathbf{Q}^{(n-1)}$ será da forma:

$$\mathbf{Q}^{(n-1)} = \begin{pmatrix} Q_{1,1}^{(n-1)} & Q_{1,2}^{(n-1)} & 0 & \dots & 0 & 0 & 0 \\ Q_{2,1}^{(n-1)} & Q_{2,2}^{(n-1)} & Q_{2,3}^{(n-1)} & \dots & Q_{2,n-1}^{(n-1)} & Q_{2,n}^{(n-1)} & Q_{2,n+1}^{(n-1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ Q_{n+1,1}^{(n-1)} & Q_{n+1,2}^{(n-1)} & Q_{n+1,3}^{(n-1)} & \dots & Q_{n+1,n-1}^{(n-1)} & Q_{n+1,n}^{(n-1)} & Q_{n+1,n+1}^{(n-1)} \end{pmatrix}.$$

Dado o fato de $\mathbf{Q}^{(n-1)}$ ser uma matriz estocástica temos que $Q_{1,1}^{(n-1)} + Q_{1,2}^{(n-1)} = 1$, ou seja, como já visto no item (i) ou (ii) anteriores, existe uma matriz $\mathbf{X}_{(n-1) \cdot 2+1}$ de determinante positivo e da forma $\mathbf{W}^{(1)}$ tal que

$$\mathbf{Q}^{(n)} := \mathbf{Q}^{(n-1)} \cdot \mathbf{X}_{(n-1) \cdot 2+1} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ Q_{2,1}^{(n)} & Q_{2,2}^{(n)} & Q_{2,3}^{(n)} & \dots & Q_{2,n-1}^{(n)} & Q_{2,n}^{(n)} & Q_{2,n+1}^{(n)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ Q_{n+1,1}^{(n)} & Q_{n+1,2}^{(n)} & Q_{n+1,3}^{(n)} & \dots & Q_{n+1,n-1}^{(n)} & Q_{n+1,n}^{(n)} & Q_{n+1,n+1}^{(n)} \end{pmatrix}.$$

Sendo assim, obtemos uma sequência de matrizes $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{(n-1) \cdot 2+1}$ de determinante positivo e do tipo $\mathbf{W}^{(i)}$, para $i = 1, 2, \dots, (n-1) \cdot 2+1$, satisfazendo

$$\mathbf{Q} \cdot \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{(n-1) \cdot 2+1} = \mathbf{Q}^{(n)} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ Q_{2,1}^{(n)} & Q_{2,2}^{(n)} & Q_{2,3}^{(n)} & \dots & Q_{2,n-1}^{(n)} & Q_{2,n}^{(n)} & Q_{2,n+1}^{(n)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ Q_{n+1,1}^{(n)} & Q_{n+1,2}^{(n)} & Q_{n+1,3}^{(n)} & \dots & Q_{n+1,n-1}^{(n)} & Q_{n+1,n}^{(n)} & Q_{n+1,n+1}^{(n)} \end{pmatrix}.$$

□

Provemos agora a Proposição.

Proposição. O conjunto das matrizes estocásticas de determinante positivo

$$\{\mathbf{X} \in \mathcal{G}_n; \det(\mathbf{X}) > 0\}$$

é conexo por caminhos.

Demonstração. Façamos a prova por indução em relação a n , cujo valor denota a dimensão das matrizes estocásticas $n \times n$.

Para $n = 2$ sabemos que este resultado é verdadeiro (vide [Summer]).

Suponhamos que o resultado valha para n qualquer e analisemos o caso \mathcal{G}_{n+1} . Dada uma matriz $\mathbf{Q} \in \mathcal{G}_{n+1}$ de determinante positivo, sabemos, pelo Lema anterior, que existe uma sequência finita de matrizes de determinante positivo $\mathbf{X}_1, \dots, \mathbf{X}_m$ tal que

$$\mathbf{Q} \cdot \mathbf{X}_m \cdot \mathbf{X}_{m-1} \cdots \mathbf{X}_1 = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ Q'_{2,1} & Q'_{2,2} & \cdots & Q'_{2,n+1} \\ \vdots & \vdots & \ddots & \vdots \\ Q'_{n+1,1} & Q'_{n+1,2} & \cdots & Q'_{n+1,n+1} \end{pmatrix} := \mathbf{Q}'.$$

Dada a matriz \mathbf{Q}' teremos que realizar algumas transformações na matriz tal que tenhamos a seguinte desigualdade:

$$Q'_{i,1} < 1, \forall i = 2, 3, \dots, n+1.$$

Tal propriedade será útil para nós. Para isso, devemos tomar o conjunto

$$\mathcal{I} := \{i \in \{2, 3, \dots, n+1\}; Q'_{i,1} \geq Q'_{j,1} \forall j \in \{2, 3, \dots, n+1\}\}$$

que representa as linhas i tais que $Q'_{i,1}$ é maximal. Agora, definimos por $i_0 \in \mathcal{I}$ como uma linha qualquer cujo valor $Q'_{i,1}$ é maximal e avaliamos as seguintes possibilidades:

(i) $Q'_{i_0,1} > 1$

Nesta situação podemos ter dois casos:

$$\mathcal{I} = \{2, 3, \dots, n+1\} \text{ ou } \{2, 3, \dots, n+1\} \setminus \mathcal{I} \neq \emptyset.$$

Consideremos inicialmente o caso em que $\{2, 3, \dots, n+1\} \setminus \mathcal{I} \neq \emptyset$. Para todo $l \in \{2, 3, \dots, n+1\} \setminus \mathcal{I}$ tomamos a matriz estocástica

$$\mathbf{Y}^{(l)} = \left(Y_{i,j}^{(l)} \right)_{i,j} \text{ tal que } Y_{i,j}^{(l)} = \begin{cases} 1 & \text{se } i = j \text{ e } i \neq l; \\ \alpha & \text{se } i = j = l; \\ 1 - \alpha & \text{se } i = l \text{ e } j = i_0; \\ 0 & \text{caso contrário,} \end{cases}$$

com α satisfazendo: $\alpha > \frac{Q'_{i_0,1}-1}{Q'_{i_0,1}-Q'_{l,1}}$. Desta forma, se tomarmos o produto

$$\mathbf{Y}^{(l)} \cdot \mathbf{Q}'$$

que nada mais é que copiar a matriz \mathbf{Q}' trocando a linha l pela soma $\alpha v_l + (1 - \alpha)v_{i_0}$, com v_l e v_{i_0} vetores representando a linha l e a linha i_0 da matriz \mathbf{Q}' , respectivamente. Observe também que $\det(\mathbf{Y}^{(l)}) > 0$ e que a matriz $\mathbf{Y}^{(l)}$ pode ser escrita pela matriz de blocos:

$$\mathbf{Y}^{(l)} = \begin{pmatrix} \mathbf{I}_{1 \times 1} & \mathbf{0}_{1 \times n} \\ \mathbf{0}_{n \times 1} & \mathbf{A}_{n \times n}^{(l)} \end{pmatrix}$$

com $\mathbf{A}_{n \times n}^{(l)}$ uma matriz estocástica de determinante positivo.

Ao tomarmos a multiplicação $\mathbf{Y}^{(l)} \cdot \mathbf{Q}'$ teremos, na primeira coluna com a linha l , o valor

$$(1 - \alpha)Q'_{i_0,1} + \alpha Q'_{l,1},$$

o qual satisfaz $(1 - \alpha)Q'_{i_0,1} + \alpha Q'_{l,1} < 1$. Sendo assim, definindo por

$$r_1 := |\{2, 3, \dots, n+1\} \setminus \mathcal{I}|,$$

temos uma sequência de matrizes estocásticas de determinante positivo $\mathbf{Y}^{(1)}, \mathbf{Y}^{(2)}, \dots, \mathbf{Y}^{(r_1)}$ tal que a matriz

$$\mathbf{Q}'' := \mathbf{Y}^{(r_1)} \cdot \mathbf{Y}^{(r_1-1)} \dots \mathbf{Y}^{(1)} \cdot \mathbf{Q}' = \begin{pmatrix} 1 & 0 & \dots & 0 \\ Q''_{2,1} & Q''_{2,2} & \dots & Q''_{2,n+1} \\ \vdots & \vdots & \ddots & \vdots \\ Q''_{n+1,1} & Q''_{n+1,2} & \dots & Q''_{n+1,n+1} \end{pmatrix}$$

satisfaz $Q''_{l,1} < 1$, para todo $l \in \{2, 3, \dots, n+1\} \setminus \mathcal{I}$.

Resta agora avaliar as linhas $k \in \mathcal{I}$. Para isto, fixe a linha $i_1 \in \{2, 3, \dots, n+1\} \setminus \mathcal{I}$ da matriz \mathbf{Q}'' , que satisfaz $Q''_{i_1,1} < 1$. O argumento a seguir é o mesmo descrito acima. Para cada $k \in \mathcal{I}$ definimos a matriz

$$\mathbf{Z}^{(k)} = \left(Z_{i,j}^{(k)} \right) \quad \text{tal que} \quad Z_{i,j}^{(k)} = \begin{cases} 1 & \text{se } i = j \text{ e } i \neq k; \\ \alpha & \text{se } i = j = k; \\ 1 - \alpha & \text{se } i = k \text{ e } j = i_1; \\ 0 & \text{caso contrário,} \end{cases}$$

com $0 < \alpha < \frac{1 - Q''_{i_1,1}}{Q''_{k,1} - Q''_{i_1,1}}$. Como anteriormente, $\mathbf{Z}^{(k)}$ será uma matriz estocástica de determinante positivo o qual pode ser escrita como uma matriz de blocos:

$$\mathbf{Z}^{(k)} = \begin{pmatrix} \mathbf{I}_{1 \times 1} & \mathbf{0}_{1 \times n} \\ \mathbf{0}_{n \times 1} & \mathbf{A}_{n \times n}^{(k)} \end{pmatrix}$$

com $\mathbf{A}_{n \times n}^{(k)}$ uma matriz estocástica de determinante positivo.

Finalmente, se multiplicarmos a matriz \mathbf{Q}'' pela matriz $\mathbf{Z}^{(k)}$, i.e.,

$$\mathbf{Z}^{(k)} \cdot \mathbf{Q}'',$$

teremos como entrada na linha k e coluna 1 o valor:

$$(1 - \alpha)Q''_{i_1,1} + \alpha Q''_{k,1} < 1.$$

Desta forma, definindo $r_2 := |\mathcal{I}|$, temos uma sequência de matrizes $\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}, \dots, \mathbf{Z}^{(r_2)}$ tal que, ao multiplicarmos \mathbf{Q}'' por estas matrizes, obtemos

$$\mathbf{Q}''' := \mathbf{Z}^{(r_2)} \cdot \mathbf{Z}^{(r_2-1)} \dots \mathbf{Z}^{(1)} \cdot \mathbf{Q}''$$

que satisfaz

$$Q'''_{i,1} < 1, \forall i \in \{2, 3, \dots, n+1\}.$$

Assim encerramos o caso em que $\{2, 3, \dots, n+1\} \setminus \mathcal{I} \neq \emptyset$. Agora, consideremos quando

$$\mathcal{I} = \{2, 3, \dots, n+1\}.$$

Observe que neste caso basta tomarmos a matriz (estocástica e de determinante positivo)

$$\mathbf{F} = (F_{i,j}) \text{ tal que } F_{i,j} = \begin{cases} 1 & \text{se } i = j \text{ e } i \neq 2; \\ 0,5 & \text{se } i = j = 2; \\ 0,5 & \text{se } i = 2 \text{ e } j = 1; \\ 0 & \text{caso contrário} \end{cases}$$

pois, se a multiplicarmos por \mathbf{Q}' , i.e., se tomarmos

$$\mathbf{F} \cdot \mathbf{Q}',$$

teremos que a entrada da matriz $\mathbf{F} \cdot \mathbf{Q}'$ da linha 2 e coluna 1 será menor que os outros valores das linhas $3, 4, \dots, n+1$ da coluna 1. Logo, ao invés de considerarmos a matriz \mathbf{Q}' , basta realizarmos o mesmo procedimento feito no início do item (i) para a matriz $\mathbf{F} \cdot \mathbf{Q}'$.

(ii) $Q'_{i_0,1} = 1$

Sabemos que deve haver algum $j_0 \in \{2, 3, \dots, n+1\}$ tal que $Q'_{i_0, j_0} \neq 0$, pois, caso contrário, teríamos que a primeira linha e a linha i_0 da matriz \mathbf{Q}' seriam iguais, ou seja, \mathbf{Q}' teria determinante zero e, portanto, a matriz \mathbf{Q} também teria determinante zero, porém, isto é falso por hipótese. Sabemos, também, que $\sum_{j=2}^{n+1} Q'_{i_0, j} = 0$, pois $Q'_{i_0, 1} = 1$ e \mathbf{Q}' é uma matriz estocástica. Sendo assim, deve existir $l_0, l_1 \in \{2, 3, \dots, n+1\}$ tal que $Q'_{i_0, l_0} \neq 0$ e $Q'_{i_0, l_1} \neq 0$. Portanto, existe um elemento na linha i_0 da matriz \mathbf{Q}' que é diferente de zero e que não se posiciona na última coluna. Chamemos de k_0 tal elemento, i.e., $Q'_{i_0, k_0} \neq 0$ e $k_0 < n+1$.

Consideremos a matriz

$$\mathbf{Y} = (Y_{i,j}) \quad \text{tal que } Y_{i,j} = \begin{cases} 1 & \text{se } i = j \text{ e } i \neq k_0; \\ 1 - \alpha & \text{se } i = k_0 \text{ e } j = 1; \\ \alpha & \text{se } i = k_0 \text{ e } j = k_0; \\ 0 & \text{caso contrário,} \end{cases}$$

onde α satisfaz a seguinte condição:

$$0 < \alpha < 1 \text{ se } Q'_{i_0, k_0} > 0, \text{ ou}$$

$$\alpha > 1 \text{ se } Q'_{i_0, k_0} < 0.$$

A matriz definida desta forma é estocástica, tem determinante positivo e pode ser escrita como uma matriz de blocos:

$$\mathbf{Y} = \begin{pmatrix} \mathbf{A}_{n \times n} & \mathbf{0}_{1 \times 1} \\ \mathbf{0}_{1 \times n} & \mathbf{I}_{1 \times 1} \end{pmatrix}$$

o qual $\mathbf{A}_{n \times n}$ é uma matriz estocástica de determinante positivo.

Agora, tome a matriz

$$\mathbf{Q}'' := \mathbf{Q}' \cdot \mathbf{Y}.$$

Tal matriz terá a forma

$$\mathbf{Q}'' = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ Q''_{2,1} & Q''_{2,2} & \cdots & Q''_{2,n+1} \\ \vdots & \vdots & \ddots & \vdots \\ Q''_{n+1,1} & Q''_{n+1,2} & \cdots & Q''_{n+1,n+1} \end{pmatrix}$$

com $Q''_{i_0, 1} > 1$. A partir daqui, o procedimento é igual ao do item (i).

(iii) $Q'_{i_0,1} < 1$

Neste caso não há nada a fazer.

Pelo item (i), (ii) e (iii) acima concluímos que existe uma sequência finita de matrizes estocásticas de determinante positivo $\mathbf{Y}, \mathbf{Y}^{(1)}, \mathbf{Y}^{(2)}, \dots, \mathbf{Y}^{(p)}$ (aqui denoto todas as matrizes encontradas nos itens (i), (ii) e (iii) por \mathbf{Y} para facilitar a notação), que podem ser escritas como

$$\mathbf{Y}^{(l)} = \begin{pmatrix} \mathbf{I}_{1 \times 1} & \mathbf{0}_{1 \times n} \\ \mathbf{0}_{n \times 1} & \mathbf{A}_{n \times n}^{(l)} \end{pmatrix} \text{ ou } \mathbf{Y}, \mathbf{Y}^{(l)} = \begin{pmatrix} \mathbf{A}_{n \times n}^{(l)} & \mathbf{0}_{n \times 1} \\ \mathbf{0}_{1 \times n} & \mathbf{I}_{1 \times 1} \end{pmatrix},$$

com $\mathbf{A}_{n \times n}^{(l)}$ matrizes estocásticas de determinante positivo e $l = 1, 2, \dots, p$. Deste resultado, obtemos a matriz

$$\tilde{\mathbf{Q}} := \mathbf{Y}^{(p)} \cdot \mathbf{Y}^{(p-1)} \dots \mathbf{Y}^{(1)} \cdot \mathbf{Q}' \cdot \mathbf{Y} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ \tilde{Q}_{2,1} & \tilde{Q}_{2,2} & \dots & \tilde{Q}_{2,n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{Q}_{n+1,1} & \tilde{Q}_{n+1,2} & \dots & \tilde{Q}_{n+1,n+1} \end{pmatrix}$$

que satisfaz $\tilde{Q}_{i,1} < 1$, para todo $i = 2, 3, \dots, n+1$, e $\det(\tilde{\mathbf{Q}}) > 0$.

Agora definimos a matriz diagonal

$$\mathbf{P} = (P_{ij}) \text{ com } P_{ij} = \begin{cases} 1 & \text{se } i = j = 1; \\ \sum_{l=2}^{n+1} \tilde{Q}_{il} & \text{se } i = j \text{ e } i > 1; \\ 0 & \text{caso contrário.} \end{cases}$$

Observe que, para $i = 1, 2, \dots, n+1$, $P_{ii} > 0$. Isto é verdade pois, para $i = 1$ temos $P_{11} = 1$ e, para $i > 1$, obtemos da desigualdade $\tilde{Q}_{i,1} < 1$ que $P_{ii} > 0$. Desta forma, a matriz \mathbf{P} é invertível e com determinante positivo. Multipliquemos a matriz $\tilde{\mathbf{Q}}$ por \mathbf{P}^{-1} :

$$\mathbf{P}^{-1} \cdot \tilde{\mathbf{Q}} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ \frac{\tilde{Q}_{2,1}}{\sum_{l=2}^{n+1} \tilde{Q}_{2,l}} & \frac{\tilde{Q}_{2,2}}{\sum_{l=2}^{n+1} \tilde{Q}_{2,l}} & \dots & \frac{\tilde{Q}_{2,n+1}}{\sum_{l=2}^{n+1} \tilde{Q}_{2,l}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\tilde{Q}_{n+1,1}}{\sum_{l=2}^{n+1} \tilde{Q}_{n+1,l}} & \frac{\tilde{Q}_{n+1,2}}{\sum_{l=2}^{n+1} \tilde{Q}_{n+1,l}} & \dots & \frac{\tilde{Q}_{n+1,n+1}}{\sum_{l=2}^{n+1} \tilde{Q}_{n+1,l}} \end{pmatrix} := \mathbf{Z}.$$

Observe que podemos escrever a matriz \mathbf{Z} como a matriz de blocos:

$$\mathbf{Z} = \begin{pmatrix} \mathbf{I}_{1 \times 1} & \mathbf{0}_{1 \times n} \\ \mathbf{Z}_{n \times 1}^{(1)} & \mathbf{Z}_{n \times n}^{(2)} \end{pmatrix},$$

com

$$\mathbf{Z}_{n \times 1}^{(1)} = \begin{pmatrix} \frac{\tilde{Q}_{2,1}}{\sum_{l=2}^{n+1} \tilde{Q}_{2,l}} \\ \vdots \\ \frac{\tilde{Q}_{n+1,1}}{\sum_{l=2}^{n+1} \tilde{Q}_{n+1,l}} \end{pmatrix} \quad \text{e} \quad \mathbf{Z}_{n \times n}^{(2)} = \begin{pmatrix} \frac{\tilde{Q}_{2,2}}{\sum_{l=2}^{n+1} \tilde{Q}_{2,l}} & \dots & \frac{\tilde{Q}_{2,n+1}}{\sum_{l=2}^{n+1} \tilde{Q}_{2,l}} \\ \vdots & \ddots & \vdots \\ \frac{\tilde{Q}_{n+1,2}}{\sum_{l=2}^{n+1} \tilde{Q}_{n+1,l}} & \dots & \frac{\tilde{Q}_{n+1,n+1}}{\sum_{l=2}^{n+1} \tilde{Q}_{n+1,l}} \end{pmatrix},$$

uma matriz estocástica de determinante positivo. Note que, para $i = 2, 3, \dots, n+1$, podemos escrever $\tilde{Q}_{i,1} = 1 - \sum_{l=2}^{n+1} \tilde{Q}_{i,l}$ e, portanto, reescrevemos $\mathbf{Z}_{n \times 1}^{(1)}$ como

$$\mathbf{Z}_{n \times 1}^{(1)} = \begin{pmatrix} \frac{1}{\sum_{l=2}^{n+1} \tilde{Q}_{2,l}} - 1 \\ \vdots \\ \frac{1}{\sum_{l=2}^{n+1} \tilde{Q}_{n+1,l}} - 1 \end{pmatrix}.$$

Tomando as constantes $a_2 = \sum_{l=2}^{n+1} \tilde{Q}_{2,l}$, $a_3 = \sum_{l=2}^{n+1} \tilde{Q}_{3,l}$, \dots , $a_{n+1} = \sum_{l=2}^{n+1} \tilde{Q}_{n+1,l}$ temos

$$\mathbf{Z}_{n \times 1}^{(1)} = \begin{pmatrix} \frac{1}{a_2} - 1 \\ \vdots \\ \frac{1}{a_{n+1}} - 1 \end{pmatrix} \quad \text{e} \quad \mathbf{P} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & a_2 & 0 & \dots & 0 & 0 \\ 0 & 0 & a_3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & a_n & 0 \\ 0 & 0 & 0 & \dots & 0 & a_{n+1} \end{pmatrix}.$$

Após todas estas contas, faremos o caminho reverso agora. Da igualdade $\mathbf{P}^{-1} \cdot \tilde{\mathbf{Q}} = \mathbf{Z}$ temos que

$$\tilde{\mathbf{Q}} = (\mathbf{P} \cdot \mathbf{Z}).$$

Porém, como $\tilde{\mathbf{Q}} = \mathbf{Y}^{(p)} \cdot \mathbf{Y}^{(p-1)} \dots \mathbf{Y}^{(1)} \cdot \mathbf{Q}' \cdot \mathbf{Y}$, obtemos

$$\mathbf{Q}' = [\mathbf{Y}^{(p)} \cdot \mathbf{Y}^{(p-1)} \dots \mathbf{Y}^{(1)}]^{-1} (\mathbf{P} \cdot \mathbf{Z}) \mathbf{Y}^{-1}.$$

Finalmente, como $\mathbf{Q} \cdot \mathbf{X}_m \cdot \mathbf{X}_{m-1} \dots \mathbf{X}_1 = \mathbf{Q}'$, concluímos que

$$\mathbf{Q} = [\mathbf{Y}^{(p)} \cdot \mathbf{Y}^{(p-1)} \dots \mathbf{Y}^{(1)}]^{-1} (\mathbf{P} \cdot \mathbf{Z}) \mathbf{Y}^{-1} [\mathbf{X}_m \cdot \mathbf{X}_{m-1} \dots \mathbf{X}_1]^{-1}$$

Agora relembre que

- As matrizes $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_m$ terão a forma de alguma das matrizes abaixo (vide o Lema inicial deste Apêndice):

$$\mathbf{W}^{(1)} = \begin{pmatrix} \mathbf{P}_{2 \times 2}^{(1)} & \mathbf{0}_{2 \times (n-1)} \\ \mathbf{0}_{(n-1) \times (n-1)} & \mathbf{I}_{(n-1) \times (n-1)} \end{pmatrix} \quad \mathbf{W}^{(n)} = \begin{pmatrix} \mathbf{I}_{(n-1) \times (n-1)} & \mathbf{0}_{(n-1) \times 2} \\ \mathbf{0}_{2 \times (n-1)} & \mathbf{P}_{2 \times 2}^{(n)} \end{pmatrix}$$

$$\mathbf{W}^{(j)} = \begin{pmatrix} \mathbf{I}_{(j-1) \times (j-1)} & \mathbf{0}_{(j-1) \times 2} & \mathbf{0}_{(j-1) \times (n-j)} \\ \mathbf{0}_{2 \times (j-1)} & \mathbf{P}_{2 \times 2}^{(j)} & \mathbf{0}_{2 \times (n-j)} \\ \mathbf{0}_{(n-j) \times (j-1)} & \mathbf{0}_{(n-j) \times 2} & \mathbf{I}_{(n-j) \times (n-j)} \end{pmatrix}$$

com $j \in \{2, 3, \dots, n-1\}$ e $\mathbf{P}_{2 \times 2}^{(1)}, \dots, \mathbf{P}_{2 \times 2}^{(n)}$ matrizes estocásticas 2×2 . Como as matrizes $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_m$ tem determinante positivo obtemos, por hipótese de indução, os caminhos contínuos $\mathbf{X}_i(t)$ satisfazendo

$$\mathbf{X}_i(0) = \mathbf{I}_{n+1} \quad \text{e} \quad \mathbf{X}_i(1) = \mathbf{X}_i,$$

com $i = 1, 2, \dots, m$ e $t \in [0, 1]$. Isto é verdade pois \mathbf{X}_i é uma matriz formada pelo bloco $\mathbf{P}_{2 \times 2}^{(i)}$ que, por hipótese de indução, admite um caminho contínuo o ligando a identidade.

- O mesmo vale para a nossa sequência de matrizes $\mathbf{Y}, \mathbf{Y}^{(1)}, \mathbf{Y}^{(2)}, \dots, \mathbf{Y}^{(p)}$, que podem ser escritas como

$$\mathbf{Y}^{(l)} = \begin{pmatrix} \mathbf{I}_{1 \times 1} & \mathbf{0}_{1 \times n} \\ \mathbf{0}_{n \times 1} & \mathbf{A}_{n \times n}^{(l)} \end{pmatrix} \quad \text{ou} \quad \mathbf{Y}, \mathbf{Y}^{(l)} = \begin{pmatrix} \mathbf{A}_{n \times n}^{(l)} & \mathbf{0}_{n \times 1} \\ \mathbf{0}_{1 \times n} & \mathbf{I}_{1 \times 1} \end{pmatrix},$$

com $\mathbf{A}_{n \times n}^{(l)}$ matrizes estocásticas de determinante positivo. Ou seja, como as matrizes \mathbf{Y} e $\mathbf{Y}^{(i)}$, para $i = 1, 2, \dots, p$, podem ser escritas por matrizes bloco que satisfazem a hipótese de indução, obtemos os caminhos contínuos $\mathbf{Y}^{(i)}(t)$ e $\mathbf{Y}(t)$ satisfazendo:

$$\begin{aligned} \mathbf{Y}^{(i)}(0) &= \mathbf{I}_{n+1}, \quad \mathbf{Y}^{(i)}(1) = \mathbf{Y}^{(i)} \quad \text{e} \\ \mathbf{Y}(0) &= \mathbf{I}_{n+1}, \quad \mathbf{Y}(1) = \mathbf{Y}, \end{aligned}$$

com $t \in [0, 1]$.

- Para encerrar, olhemos para o produto **(PZ)**. Inicialmente, definimos a função contínua

$$\mathbf{P}(t) = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & e^{t \cdot \ln(a_2)} & 0 & \dots & 0 & 0 \\ 0 & 0 & e^{t \cdot \ln(a_3)} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & e^{t \cdot \ln(a_n)} & 0 \\ 0 & 0 & 0 & \dots & 0 & e^{t \cdot \ln(a_{n+1})} \end{pmatrix}$$

com $t \in [0, 1]$. Em seguida, dado que

$$\mathbf{Z} = \begin{pmatrix} \mathbf{I}_{1 \times 1} & \mathbf{0}_{1 \times n} \\ \mathbf{Z}_{n \times 1}^{(1)} & \mathbf{Z}_{n \times n}^{(2)} \end{pmatrix},$$

temos que a matriz estocástica de determinante positivo $\mathbf{Z}_{n \times n}^{(2)}$ admite, por hipótese de indução, o caminho contínuo

$$\mathbf{Z}_{n \times n}^{(2)}(t) \text{ satisfazendo } \mathbf{Z}_{n \times n}^{(2)}(0) = \mathbf{I}_{n+1} \text{ e } \mathbf{Z}_{n \times n}^{(2)}(1) = \mathbf{Z}_{n \times n}^{(2)},$$

com $t \in [0, 1]$. Agora, fixemos o caminho contínuo para $\mathbf{Z}_{n \times 1}^{(1)}$ como sendo

$$\mathbf{Z}_{n \times 1}^{(1)}(t) = \begin{pmatrix} \frac{1}{e^{t \cdot \ln(a_2)}} - 1 \\ \vdots \\ \frac{1}{e^{t \cdot \ln(a_{n+1})}} - 1 \end{pmatrix}.$$

Sendo assim, obtemos o caminho contínuo

$$\mathbf{Z}(t) = \begin{pmatrix} \mathbf{I}_{1 \times 1} & \mathbf{0}_{1 \times n} \\ \mathbf{Z}_{n \times 1}^{(1)}(t) & \mathbf{Z}_{n \times n}^{(2)}(t) \end{pmatrix},$$

com $\mathbf{Z}(0) = \mathbf{I}_{n+1}$ e $\mathbf{Z}(1) = \mathbf{Z}$. Portanto, a função $(\mathbf{PZ})(t) = \mathbf{P}(t)\mathbf{Z}(t)$ é contínua, estocástica, e satisfaz $(\mathbf{PZ})(0) = \mathbf{I}_{n+1}$ e $(\mathbf{PZ})(1) = \mathbf{PZ}$.

Logo, concluímos que a função

$$\mathbf{Q}(t) = [\mathbf{Y}^{(p)}(t) \cdot \mathbf{Y}^{(p-1)}(t) \cdots \mathbf{Y}^{(1)}(t)]^{-1} (\mathbf{PZ})(t) \mathbf{Y}^{-1}(t) [\mathbf{X}_m(t) \cdot \mathbf{X}_{m-1}(t) \cdots \mathbf{X}_1(t)]^{-1}$$

é contínua, tem determinante positivo e é uma matriz estocástica em todos os seus pontos. Além do mais, temos $\mathbf{Q}(0) = \mathbf{I}_{n+1}$ e $\mathbf{Q}(1) = \mathbf{Q}$. Com isto encerramos a demonstração de que o conjunto das matrizes estocásticas de determinante positivo

$$\{\mathbf{X} \in \mathcal{G}_n; \det(\mathbf{X}) > 0\}$$

é conexo por caminhos.

□

B. PROGRAMA PARA CÁLCULO DA ÁLGEBRA SEMI-SIMPLES DA SEÇÃO 5.3

```
# script escrito para rodar no octave versao 4.0.0
# Necessario o pacote "linear-algebra" para rodar
# a funcao cartprod().
# Dada uma base de um ideal soluvel, o programa ira tentar calcular uma
# base para uma algebra semi-simples, tendo em vista o teorema de Levi.
# Para avaliar se uma algebra eh semi-simples inspecionamos, via forca
# bruta, se a matriz da forma de Cartan-Killing eh inversivel.
#
# O programa, assim que achar uma algebra semi-simples, ira imprimir os
# resultados em um arquivo txt.
1;

pkg load all; # carrega o pacote linear-algebra

##### Base do espaco tangente
N = 3;
bases = zeros(N, N, N*(N-1) );
k = 1;
for i = 1:N
    for j = 1:N
        if ( i != j)
            bases(i,i,k) = -1;
            bases(i,j,k) = 1;
            k++;
        endif
    endfor
endfor
#####

##### Funcoes auxiliares
```

```
### Operador colchete
```

```
function ret = colchete(a, b)
```

```
    ret = a*b - b*a;
```

```
endfunction
```

```
### Calcula a matriz da transformacao adjunta
```

```
function transf_matrix = transformacao(base, X_sist_linear, n)
```

```
    transf_matrix = zeros(n*(n-1), n*(n-1), n*(n-1) );
```

```
    for i = 1:(n*(n-1))
```

```
        for j = 1:(n*(n-1))
```

```
            y = ( colchete(base(:, :, i), base(:, :, j) ) )' ;
```

```
            y = y(:);
```

```
            transf_matrix(:, j, i) = X_sist_linear \ y;
```

```
        endfor
```

```
    endfor
```

```
endfunction
```

```
### Calcula a matriz da forma de Killing
```

```
function matriz_killing = forma_killing( index_2, n)
```

```
    matriz_killing = zeros(n, n);
```

```
    for i = 1:n
```

```
        for j = 1:n
```

```
            matriz_killing( i, j) = trace( index_2(:, :, i) * index_2(:, :, j) );
```

```
        endfor
```

```
    endfor
```

```
endfunction
```

```
#####
```

```
Pontos = cartprod (1:6, 1:6)
```

```
Total =6**2;
```

```
### Testando Bases para achar a algebra semi-simples
```

```
LIbases = zeros(Total, 6);
```

```
for i = 1:Total
```

```

    LIbases(i, Pontos(i,1) ) = 1;
    LIbases(i, Pontos(i,2) ) = -1;
endfor

LI = zeros(6,6);
LI(1,:) = [0,1,-1,1,-1,0];
LI(2,:) = [-1,1,0,1,0,-1];

novaBase = zeros(N, N, N*(N-1) );
novaBase(:,:,1) = bases(:,:,2) + bases(:,:,4) - ( bases(:,:,5) + bases(:,:,3) );
novaBase(:,:,2) = bases(:,:,2) + bases(:,:,4) - ( bases(:,:,1) + bases(:,:,6) );

interacao = 1;
for i = unique(1:Total)
    for j = setdiff(1:Total, i)
        for k = setdiff(1:Total, union(i, j) )
            for l = setdiff(1:Total, union( union(i,j), k ) )
                novaBase(:,:,3) = bases(:,:, Pontos(i,1) ) - bases(:,:, Pontos(i,2) ) ;
                novaBase(:,:,4) = bases(:,:, Pontos(j,1) ) - bases(:,:, Pontos(j,2) ) ;
                novaBase(:,:,5) = bases(:,:, Pontos(k,1) ) - bases(:,:, Pontos(k,2) ) ;
                novaBase(:,:,6) = bases(:,:, Pontos(l,1) ) - bases(:,:, Pontos(l,2) ) ;
                interacao++;

                LI(3,:) = LIbases(i,:);
                LI(4,:) = LIbases(j,:);
                LI(5,:) = LIbases(k,:);
                LI(6,:) = LIbases(l,:);

                if ( rank(LI) == 6 )
                    X = zeros(N**2, N*(N-1) );

                    for i = 1:(N*(N-1))
                        aux = novaBase(:,:,i)';
                        X(:, i) = aux(:);
                    endfor
                transformacaoMatriz = transformacao(novaBase, X, N);
            end
        end
    end
end

```

```

transformacaoMatriz2 = zeros(3,3,3);
transformacaoMatriz2(:,:,1) = transformacaoMatriz(4:6,4:6,4);
transformacaoMatriz2(:,:,2) = transformacaoMatriz(4:6,4:6,5);
transformacaoMatriz2(:,:,3) = transformacaoMatriz(4:6,4:6,6);

a1 = round( transformacaoMatriz(1:3,4:6,4) .* 100000) ./100000;
a2 = round( transformacaoMatriz(1:3,4:6,5) .* 100000) ./100000;
a3 = round( transformacaoMatriz(1:3,4:6,6) .* 100000) ./100000;

if ( all( all(a1 == 0 ) ) && all( all(a2 == 0 ) ) &&
    all( all(a3 == 0 ) ) )
    resultado = forma_killing( transformacaoMatriz2, 3 );
    rank(resultado)
    if (rank (resultado ) == 3 )
        filename1 = strcat( num2str(interacao), '.txt');

        myfile = fopen(filename1,'a');
        fprintf (myfile,
            '*****');
        fprintf (myfile, 'i j k l');
        fprintf (myfile, [i j k l] );
        fprintf (myfile, LI );
        fprintf (myfile, transformacaoMatriz );
        fprintf (myfile, ' ' );

        fprintf (myfile, resultado );
        fclose(myfile);
    endif

endif

endif

endfor

endfor

endfor

endfor

```

BIBLIOGRAFIA

- [Elon1] Lima, Elon Lages; *Espaços Métricos*, Projeto Euclides, CNPQ, 2003.
- [Elon2] Lima, Elon Lages; *Elementos de Topologia Geral*, Editora SBM, Rio de Janeiro, 2009.
- [Pedro Fernandez] Fernandez, Pedro J.; *Medida e Integração*, Projeto Euclides, Rio de Janeiro, 2007.
- [Carlos Isnard] Isnard, Carlos; *Introdução à Medida e Integração*, Projeto Euclides, Rio de Janeiro, 2007.
- [Stillwell] Stillwell, John; *Naive Lie Theory*, Springer Verlag, 2008.
- [Summer] Summer, Jeremy G.; *Lie geometry of 2×2 Markov matrices*, 2017.
- [Baez and Fong] Baez, John C. and Fong, Brendan; *A Noether Theorem for Markov Processes*, 2012.
- [San Martin] San Martin, Luiz Antonio Barrera; *Álgebras de Lie*, Editora UNICAMP, 2010.
- [Octave] John W. Eaton, David Bateman, Søren Hauberg, and Rik Wehbring; *GNU Octave version 4.0.0 manual: a high-level interactive language for numerical computations*, 2015.
- [Barry James] James, Barry R.; *Probabilidade: um curso em nível intermediário*, Projeto Euclides, Rio de Janeiro, 2008.
- [John L. Kelley] Kelley, John L.; *General Topology*, Dover edition, Mineola, New York, 2017.
- [Lang] Lang, Serge; *$SL_2(\mathbb{R})$* , Springer-Verlag New York Inc., 1985.
- [Paulo A. Martin] Martin, Paulo A.; *Grupos, Corpos e Teoria de Galois*, Editora Livraria da Física, São Paulo, 2010.

- [Feller] Feller, William; *An introduction to Probability Theory and Its Applications Volume I*, John Wiley & Sons, Inc., 1968.
- [Durrett] Durrett, Richard; *Essentials of Stochastic Processes*, Springer, 2012.
- [Biane and Durrett] Durrett, Richard and Biane, Philippe; *Lectures on Probability theory*, Springer, 1993.
- [Belitsky and Schütz] V. Belitsky and G.M Schütz; *Quantum algebra symmetry of the ASEP with second-class particles*, 2015.
- [Daniela] Cuesta, Daniela Sabrina; *Proceso de inclusión en Z : dualidad, acoplamiento y propagación de equilibrio local*, Tesis de Licenciatura, Universidad de Buenos Aires, 2015.
- [Redig] Redig, Frank; *Duality and exactly solvable models in non-equilibrium*, 2015.
- [Aidan Sudbury and Peter Lloyd] Sudbury, Aidan and Lloyd, Peter; *Quantum Operators in Classical Probability Theory: II. The Concept of Duality in Interacting Particle Systems*, The Annals of Probability, Vol. 23, N^o 4 (oct., 1995), pp 1816-1830.
- [Mac Lane and Birkhoff] Mac Lane, Saunders and Birkhoff, Garret; *Algebra*, Chelsea Publishing Company, New York, Third Edition, 1988.
- [Frank Redig and Federico Sau] Redig, Frank and Sau, Federico; *Stochastic duality and eigenvectors*, 2018.
- [Ken Brown] Brown, Ken A.; *Hopf Algebras*, Lecture Notes.