



WALTER AUGUSTO FONSÊCA DE CARVALHO

**PROCESSOS DE RENOVAÇÃO OBTIDOS POR
AGREGAÇÃO DE ESTADOS A PARTIR DE UM PROCESSO
MARKOVIANO**

CAMPINAS

2014



UNIVERSIDADE ESTADUAL DE CAMPINAS
Instituto de Matemática, Estatística
e Computação Científica

WALTER AUGUSTO FONSÊCA DE CARVALHO

**PROCESSOS DE RENOVAÇÃO OBTIDOS POR
AGREGAÇÃO DE ESTADOS A PARTIR DE UM PROCESSO
MARKOVIANO**

Tese apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutor em estatística.

Orientadora: Nancy Lopes Garcia

Coorientador: Alessandro Giacomo Grimbert Gallo

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA TESE DEFENDIDA PELO ALUNO WALTER AUGUSTO FONSÊCA DE CARVALHO, E ORIENTADA PELA PROFA. DRA. NANCY LOPES GARCIA.

Assinatura da Orientadora

A handwritten signature in black ink, appearing to read "Nancy Lopes Garcia", written over a horizontal line.

Assinatura do Coorientador

A handwritten signature in black ink, appearing to read "Alessandro Giacomo Grimbert Gallo", written over a horizontal line.

CAMPINAS
2014

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Matemática, Estatística e Computação Científica
Maria Fabiana Bezerra Muller - CRB 8/6162

C253p Carvalho, Walter Augusto Fonsêca, 1964-
Processos de renovação obtidos por agregação de estados a partir de um processo markoviano / Walter Augusto Fonsêca de Carvalho. – Campinas, SP : [s.n.], 2014.

Orientador: Nancy Lopes Garcia.

Coorientador: Alexsandro Giacomo Grimbert Gallo.

Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Processo markoviano oculto. 2. Processos de renovação. 3. Agregabilidade. 4. Método d-bar. I. Garcia, Nancy Lopes, 1964-. II. Gallo, Alexsandro Giacomo Grimbert. III. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. IV. Título.

Informações para Biblioteca Digital

Título em outro idioma: Renewal processes obtained by aggregation of states from a markovian process

Palavras-chave em inglês:

Hidden Markov process

Renewal processes

Lumpability

D-bar method

Área de concentração: Estatística

Titulação: Doutor em Estatística

Banca examinadora:

Nancy Lopes Garcia [Orientador]

Christophe Frédéric Gallesco

Miguel Natalio Abadi

Florencia Graciela Leonardi

Pablo Martin Rodriguez

Data de defesa: 28-02-2014

Programa de Pós-Graduação: Estatística

Tese de Doutorado defendida em 28 de fevereiro de 2014 e aprovada

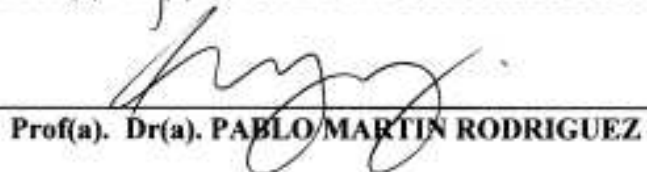
Pela Banca Examinadora composta pelos Profs. Drs.


Prof(a). Dr(a). NANCY LOPES GARCIA


Prof(a). Dr(a). CHRISTOPHE FRÉDÉRIC GALLESKO


Prof(a). Dr(a). MIGUEL NATALIO ABADI


Prof(a). Dr(a). FLORENCIA GRACIELA LEONARDI


Prof(a). Dr(a). PABLO MARTIN RODRIGUEZ

Abstract

This thesis is devoted to the study of binary renewal processes obtained as aggregation of states from Markov processes with finite alphabet.

In the first part, we use a matrix approach to obtain conditions under which the aggregated process belongs to each of the following classes: (1) Markov of finite order, (2) process of infinite order with continuous transition probabilities, (3) Gibbsian process.

The second part deals with the distance \bar{d} between binary renewal processes. We obtain conditions under which this distance can be achieved between these processes.

Keywords: Agregating Transformation, Hidden Markov Process, Renewal Processes, Lumpability, Partition-Distance, \bar{d} -Distance.

Resumo

Esta tese é dedicada ao estudo dos processos de renovação binários obtidos como agregação de estados a partir de processos Markovianos com alfabeto finito.

Na primeira parte, utilizamos uma abordagem matricial para obter condições sob as quais o processo agregado pertence a cada uma das seguintes classes: (1) Markoviano de ordem finita, (2) processo de ordem infinita com probabilidades de transição contínuas, (3) processo Gibbsiano.

A segunda parte trata da distância \bar{d} entre processos de renovação binários. Obtivemos condições sob as quais esta distância pode ser atingida entre tais processos.

Palavras chaves: Transformação Agregante, Processo Markoviano Oculto, Processos de Renovação, Agregabilidade, Distância de Partição, Distância \bar{d} .

Sumário

1	Introdução	1
2	Processos De Renovação Obtidos Por Agregação De Estados - Alfabeto Finito	3
2.1	Notação E Definições	3
2.2	Abordagem Matricial	6
2.3	Agregabilidade De Um Processo Markoviano	7
2.4	Continuidade Das Probabilidades De Transição	8
2.5	Gibbsianidade	19
2.6	Exemplos	21
3	Acoplamentos E Discrepância Entre Processos De Renovação	32
3.1	Acoplamento De Renovação	32
3.2	Discrepância Entre Processos De Renovação	33
3.3	Transformação Agregante E Distância De Partição	35
3.4	Transformação Agregante E Distância \bar{d}	37
3.5	Condição De Ordenação E Acoplamentos De Renovação	40
4	Conclusão	46
	Referências Bibliográficas	48
A	Álgebra Matricial	52
B	Teorema De Perron-Frobênus	55

Capítulo 1

Introdução

A imagem de um processo Markoviano através de uma transformação sobrejetiva é conhecida na literatura sob nomes distintos associados a várias áreas. *Processo de Markov Oculto* (Hidden Markov models) [Juang & Rabiner \(1991\)](#), [Baum & Petrie \(1966\)](#), [Han & Marcus \(2006\)](#), muito utilizado em aplicações como na área de reconhecimento de padrões comportamentais, biológicos e temporais como a fala, a escrita, os gestos e a bioinformática, *Fator da cadeia de Markov* (1-block factor) [Boyle & Petersen \(2011\)](#) em teoria ergódica, *Processos difusos* (fuzzy processes) [Małłoka \(2000\)](#) ou *amalgamados* em física estatística, *Fonte Markoviana* (Markov Source) [Massey \(1978\)](#), [Cénac et al. \(2012\)](#) comumente utilizados na teoria da comunicação, *Cadeia de Markov Agregada* (Lumped/Aggregated Markov Chain) [Kemeny & Snell \(1976\)](#), [Boucherie \(1993\)](#), são modelos baseados em regras estocásticas para redes, *Cópia com Ruído* (Copy With noise) [Lestas et al. \(2008\)](#) em redes genéticas. De forma geral, tais processos aparecem quando um certo fenômeno é observado na presença de um ruído, interferência, que pode levar a confundir, ou não poder se distinguir certos estados. Muitas vezes o interesse é inferir sobre o processo de Markov original a partir do processo observado (imagem), daí, por exemplo, o nome de processo oculto. No entanto, a questão inversa tem atraído bastante interesse na literatura de física estatística, no contexto de Gibbsianidade, de grupos de renormalização e Modelos de Ising da área de mecânica estatística. Supondo que um sistema Markoviano é observado de forma que não se pode distinguir certos estados, que podemos dizer a respeito da distribuição do que vemos?

Nesta tese, propomos estudar a imagem de processos de Markov (com alfabeto finito) por uma transformação sobrejetiva particular chamada agregante. Neste trabalho, uma transformação será dita agregante se ela transforma um elemento do alfabeto original no símbolo 1 e todos os demais no símbolo 0. Desta forma, o processo binário resultante é chamado *processo agregado*. De um modo geral, sabe-se que um processo oculto não é Markoviano de nenhuma ordem, [Harris \(1955\)](#), [Chazottes & Ugalde \(2011\)](#). Nossos processos agregados pertencem à classe dos processos binários de renovação, [Shields \(1973\)](#).

O primeiro problema abordado nesta tese é obter condições, sobre a matriz de transição do

processo original, para que o processo agregado tenha probabilidades condicionais com respeito ao passado, contínuas. Usamos para isso, uma abordagem matricial semelhante à de Darroch e Seneta (1967), em um estudo sobre quasi-estacionariedade. Condições suficientes para continuidade é encontrada e, neste caso, estimativas da taxa de continuidade são obtidas como função dos dois maiores auto-valores de uma submatriz da matriz de transição do processo Markoviano original. Também apresentamos condições sobre os auto-valores garantindo que o processo agregado é Markoviano de ordem k . Harris (1955) provou que, se a cadeia original possui matriz de transição com entradas estritamente positivas (uma condição mais forte que a nossa) então a cadeia imagem possui probabilidade de transição contínua.

Mais recentemente, Chazottes & Ugalde (2003) obtiveram condições para que o processo imagem seja Gibbsiano no sentido de Bowen (2008). Neste sentido, mostramos que sob a mesma condição mencionada no parágrafo acima, o processo agregado é Gibbsiano no sentido de Dobrushin (1968), isto é, que as probabilidades condicionais com respeito ao passado e ao futuro são contínuas. Se um processo é Gibbsiano no sentido de Bowen, então ele é Gibbsiano no sentido de Dobrushin (Fernández *et al.* (2011)). Segue que uma comparação com o resultado de Chazottes e Ugalde não é direta.

Na segunda parte desta tese estudamos acoplamentos, sob o ponto de vista de Fernández *et al.* (2001) e a possibilidade de atingir a distância \bar{d} entre dois processos agregados. Conforme definido por Ornstein (1973) em um estudo sobre teoria ergódica, a distância \bar{d} é o mínimo de discrepância que pode ser atingida entre dois processos acoplados. A questão de poder construir um acoplamento atingindo esta distância foi extensivamente estudada por (Ellis (1976), Ellis (1978), Ellis (1980a), Ellis (1980b)) no caso de processos Markovianos. Mais recentemente, Galves *et al.* (2010) mostraram que o acoplamento maximal atinge esta distância no caso de cadeias ordenadas, ordenação esta definida no Capítulo 4. Nós tomamos vantagem da noção de processo agregado para obter exemplos de processos de ordem infinita para os quais sabemos construir um acoplamento atingindo \bar{d} sem assumir a ordenação de Galves *et al.* (2010).

A tese está organizada da seguinte forma. No Capítulo 2 introduzimos a notação e definições básicas que serão utilizadas em toda tese, apresentamos a principal ferramenta a ser utilizada na tese, que é a construção de um processo de renovação com alfabeto binário, a partir de um processo Markoviano com alfabeto finito, através da transformação agregante. Condições para a Markovianidade do processo agregado são apresentadas, assim como as condições para a continuidade e Gibbsianidade. A medida estacionária do processo de renovação, obtido por agregação de estados a partir de um processo Markoviano, através da Transformação Agregante, pode ser facilmente calculada a partir da medida do processo Markoviano original, no Capítulo 3 esse raciocínio é desenvolvido e o alcance a distância \bar{d} é estudado.

Capítulo 2

Processos De Renovação Obtidos Por Agregação De Estados - Alfabeto Finito

Iniciamos este capítulo com a notação e as definições iniciais, que serão utilizadas ao longo de toda a tese. Em seguida estudamos a construção de um processo de renovação com espaço de estados binário, a partir de um processo Markoviano com espaço de estados finito, através de uma transformação sobrejetiva. E, finalmente, estudamos a propriedade de continuidade das probabilidades de transição do processo obtido por este método, bem como a Gibbsianidade.

2.1 Notação E Definições

Seja um conjunto finito A , que chamaremos de alfabeto, cuja cardinalidade denotaremos por $|A|$. Uma sequência finita qualquer de símbolos no alfabeto A será denominada palavra sobre A . Para todo inteiro n , A^n denota o conjunto das palavras de n símbolos. Dados dois inteiros $-\infty < m \leq n < +\infty$, a sequência de símbolos em A , x_m, x_{m+1}, \dots, x_n será denotada por x_m^n e o seu comprimento por $|x_m^n| = n - m + 1$. Dadas duas palavras $x_1^m = x_1 \dots x_m$ e $y_1^n = y_1 \dots y_n$ sobre A , a cocatenação de x_1^m e y_1^n será denotada por x_1^m, y_1^n e definida por $x_1^m, y_1^n = x_1 \dots x_m y_1 \dots y_n$. Semelhantemente, $x_{-\infty}^n$ denota a sequência $(x_i)_{i \in \mathbb{N}}$, toda história passada até o tempo n , e $A_{-\infty}^n$ o conjunto de tais sequências. As sequências genéricas em A serão denotadas negritadas e sem sub ou superíndices, $\mathbf{x} \in A$.

Todos os processos mencionados aqui são sequências de variáveis aleatórias discretas definidas em um mesmo espaço de probabilidade (Ω, F, \mathbb{P}) .

Utilizaremos letras maiúsculas em negrito, como \mathbf{X} , para representar processos, isto é, $\mathbf{X} = (X_n)_{n \in \mathbb{Z}}$ é um processo estacionário sobre um alfabeto finito A . Apesar de não convencional, utilizaremos a notação $P^{\mathbf{X}}$ para representar as probabilidades de transição do processo \mathbf{X} . Assim, para quaisquer sequências de símbolos $x_{-\infty}^{-1}$ e qualquer símbolo $a \in A$ teremos

$$P^{\mathbf{X}}(a|x_{-\infty}^{-1}) = \mathbb{P}(X_0 = a|X_{-\infty}^{-1} = x_{-\infty}^{-1}).$$

Denotaremos por $\mu_{\mathbf{X}}$ a distribuição estacionária do processo \mathbf{X} , isto é

$$\mu_{\mathbf{X}}(a_1^n) = \mathbb{P}(X_1^n = a_1^n)$$

para todos $a_1^n \in A^n$ e $n \in \mathbb{Z}$.

Apresentamos a definição de processo de renovação que será utilizada em toda a tese.

Definição 2.1 *Um processo estocástico estacionário $(X_n)_{n \in \mathbb{Z}}$ de medida estacionária $\mu_{\mathbf{X}}$ e alfabeto finito A é um processo de renovação, se existe um símbolo $e \in A$, tal que*

$$\mathbb{P}(X_0 = a | X_{-i}^{-1} = a_{-i}^{-1}, X_{-i-1} = e, X_{-r}^{-i-2} = a_{-2}^{-i-r}) = \mathbb{P}(X_0 = a | X_{-i}^{-1} = a_{-i}^{-1}, X_{-i-1} = e), \quad (2.1)$$

para todos $r \in \mathbb{N}$, $i \leq r$, $a_{-i}^{-1} \in A^i$ e $a \in A$.

Um processo de renovação binário, renovando no símbolo 1, possui esta mesma definição, com $e = 1$. A probabilidade de transição $\mathbb{P}(X_0 = 1 | X_{-i}^{-1} = 0_{-i}^{-1}, X_{-i-1} = 1)$ será denotada por $P^{\mathbf{X}}(1|0^i1)$, onde 0^i1 representa i símbolos 0 antes do símbolo 1. Em todas as nossas considerações, suporemos que

$$\lim_{n \rightarrow \infty} P(1|0^n1) = P(1|\underline{0})$$

sempre que houver convergência de $P(1|0^n1)$ quando $n \rightarrow \infty$. Onde $\underline{0}$ representa sequência infinita de símbolos todos iguais a zero.

A não ser nos casos mencionados especificamente, todos os processos de renovação abordados nesta tese são definidos no alfabeto $\{0, 1\}$, isto é, são binários, construídos através da transformação agregante, que é definida a seguir.

Definição 2.2 *Seja o conjunto finito A e $a \in A$. A função sobrejetiva $t : A \rightarrow \{0, 1\}$ definida por*

$$t(x) = \begin{cases} 1, & \text{se } x = a \\ 0, & \text{se } x \neq a \end{cases} \quad (2.2)$$

é chamada de *Transformação Agregante*.

Sem perda de generalidade consideramos sempre $A = \{1, 2, \dots, m\}$.

Com abuso de notação, utilizaremos $t : A^n \rightarrow \{0, 1\}^n$ para denotar a transformação que mapeia a sequência x_1^n na sequência $t(x_1^n)$, $t(\mathbf{X}) = \mathbf{Y} = (y_n)_{n \in \mathbb{N}}$, tal que $y_n = t(x_n)$, $n \in \mathbb{N}$. Como exemplo, considerando o alfabeto $A = \{1, 2, 3\}$ e $a = 1$, a Transformação Agregante (2.2) transforma um processo \mathbf{X} em $t(\mathbf{X})$, isto é, a partir da sequência de símbolos

1 1 2 1 2 3 2 1 1 2 2 3 ... 3 2 3 1 construímos 1 1 0 1 0 0 0 1 1 0 0 0 ... 0 0 0 1.

A generalização para sequências infinitas e sequências de variáveis aleatórias é análoga.

A Proposição 2.1, a seguir, nos fornece o método construtivo de se obter um processo de renovação por agregação de estados, a partir de um processo Markoviano com alfabeto finito.

Proposição 2.1 *Sejam os conjuntos finitos $A = \{1, 2, \dots, m\}$ e $B = \{0, 1\}$. Seja \mathbf{X} um processo Markoviano estacionário com alfabeto A e probabilidades de transição $P^{\mathbf{X}}(a|b) = p_{ba}, \forall a, b \in A$. Considere a Transformação Agregante t definida por (2.2), então $t(\mathbf{X})$ é um processo de renovação estacionário.*

Prova da Proposição 2.1

Vamos mostrar que se o processo \mathbf{X} é estacionário então o processo $\mathbf{Y} = t(\mathbf{X})$ também é. Como \mathbf{X} é estacionário,

$$\mu_{\mathbf{X}}(a_1^n) = \mathbb{P}(X_1^n = a_1^n) = \mathbb{P}(X_{t+1}^{t+n} = a_1^n)$$

para todo $n \in \mathbb{N}, t \in \mathbf{Z}$ e $a_1^n \in A^n$. Portanto, para todo $b_1^n \in B^n$,

$$\mathbb{P}(Y_{t+1}^{t+n} = b_1^n) = \sum_{a_1^n \in t^{-1}(b_1^n)} \mathbb{P}(X_{t+1}^{t+n} = a_1^n) = \sum_{a_1^n \in t^{-1}(b_1^n)} \mathbb{P}(X_1^n = a_1^n) = \mathbb{P}(Y_1^n = b_1^n).$$

Logo, \mathbf{Y} é estacionário.

Provaremos agora que $\mathbf{Y} = t(\mathbf{X})$ é um processo de renovação. Para todo $n \in \mathbb{N}$ e $b_{-r-n-1}^{-n-2} \in A^r$ e pela propriedade de Markov, temos

$$\begin{aligned} P^{\mathbf{Y}}(1|0^n 1 b_1^r) &= \mathbb{P}(Y_0 = 1 | Y_{-n}^{-1} = 0_{-n}^{-1}, Y_{-n-1} = 1, Y_{-r-n-1}^{-n-2} = b_{-r-n-1}^{-n-2}) \\ &= \mathbb{P}(X_0 = 1 | X_{-n}^{-1} \in t^{-1}(0_{-n}^{-1}), X_{-n-1} = 1, X_{-r-n-1}^{-n-2} \in t^{-1}(b_{-r-n-1}^{-n-2})) \\ &= \frac{\mathbb{P}(X_{-r-n-1}^{-n-2} \in t^{-1}(b_{-r-n-1}^{-n-2}), X_{-n-1} = 1, X_{-n}^{-1} \in t^{-1}(0_{-n}^{-1}), X_0 = 1)}{\mathbb{P}(X_{-r-n-1}^{-n-2} \in t^{-1}(b_{-r-n-1}^{-n-2}), X_{-n-1} = 1, X_{-n}^{-1} \in t^{-1}(0_{-n}^{-1}))} \\ &= \frac{\mathbb{P}(X_{-r-n-1}^{-n-2} \in t^{-1}(b_{-r-n-1}^{-n-2}), X_{-n-1} = 1)}{\mathbb{P}(X_{-r-n-1}^{-n-2} \in t^{-1}(b_{-r-n-1}^{-n-2}), X_{-n-1} = 1)} \\ &\quad \cdot \frac{\mathbb{P}(X_0 = 1, X_{-n}^{-1} \in t^{-1}(0_{-n}^{-1}) | X_{-n-1} = 1, X_{-r-n-1}^{-n-2} \in t^{-1}(b_{-r-n-1}^{-n-2}))}{\mathbb{P}(X_{-n}^{-1} \in t^{-1}(0_{-n}^{-1}) | X_{-n-1} = 1, X_{-r-n-1}^{-n-2} \in t^{-1}(b_{-r-n-1}^{-n-2}))} \\ &= \frac{\mathbb{P}(X_0 = 1, X_{-n}^{-1} \in t^{-1}(0_{-n}^{-1}) | X_{-n-1} = 1)}{\mathbb{P}(X_{-n}^{-1} \in t^{-1}(0_{-n}^{-1}) | X_{-n-1} = 1)} \\ &= \mathbb{P}(X_0 = 1 | X_{-n}^{-1} \in t^{-1}(0_{-n}^{-1}), X_{-n-1} = 1) \\ &= P^{\mathbf{Y}}(1|0^n 1). \end{aligned}$$

■

A expressão geral para o cálculo das probabilidades de transição do processo agregado binário será

$$P^{\mathbf{Y}}(1|0^n 1) = \frac{\sum_{x_i \in t^{-1}(0)} p_{1x_i} \cdot \prod_{i=1}^{n-1} (p_{x_i x_{i+1}}) \cdot p_{x_n 1}}{\sum_{x_i \in t^{-1}(0)} p_{1x_i} \cdot \prod_{i=1}^{n-1} (p_{x_i x_{i+1}})}. \quad (2.3)$$

Com as devidas adaptações, a Proposição 2.1 pode ser generalizada para uma transformação agregante onde o conjunto B é não binário, mas, pelo menos um elemento de B é imagem de um

único elemento de A .

2.2 Abordagem Matricial

Seja $(X_n)_{n \in \mathbb{Z}}$ um processo Markoviano com espaço de estados $A = \{1, 2, \dots, m\}$ e probabilidade de transição $P^{\mathbf{X}}$. Utilizaremos a decomposição da matriz $P^{\mathbf{X}}$ como

$$P^{\mathbf{X}} = \begin{bmatrix} p_{11} & V \\ W' & P \end{bmatrix}, \quad (2.4)$$

onde

- $p_{11} = \mathbb{P}(X_0 = 1 | X_{-1} = 1)$.
- V é o vetor-linha formado pelas probabilidades de transição $p_{1i} = \mathbb{P}(X_0 = i | X_{-1} = 1)$, $i \in A \setminus \{1\}$.
- W é o vetor-linha formado pelas probabilidades de transição $p_{i1} = \mathbb{P}(X_0 = 1 | X_{-1} = i)$, $i \in A \setminus \{1\}$, W' é o vetor-coluna obtido pela transposição de W .
- P é a submatriz de $P^{\mathbf{X}}$, de ordem $(m-1) \times (m-1)$ com $p_{ij} = \mathbb{P}(X_0 = j | X_{-1} = i)$, $i, j \in A \setminus \{1\}$.

Defina $\mathbf{1}$ e $\mathbf{0}$ como sendo os vetores linha, formados todos, respectivamente, de uns e de zeros, de comprimento $m-1$. Para evitar os casos triviais, consideraremos sempre em nossas análises $V \neq \mathbf{0}$ e $W \neq \mathbf{0}$. Como $P^{\mathbf{X}}$ é uma matriz estocástica, $0 \leq P_{ij}^{\mathbf{X}} \leq 1$ e $P^{\mathbf{X}} \cdot \mathbf{1}' = \mathbf{1}'$, então a submatriz P é sub-estocástica, isto é, $0 \leq p_{ij} \leq 1$ e $P \cdot \mathbf{1}' \leq \mathbf{1}'$.

Considerando a Transformação Agregante t dada por (2.2), as probabilidades de transição do processo agregado podem ser calculadas através de

$$P^{t(\mathbf{X})}(1|0^{n+1}1) = \begin{cases} p_{11}, & \text{se } n = -1 \\ \frac{V \cdot P^n \cdot W'}{V \cdot P^n \cdot \mathbf{1}'}, & \text{se } n \geq 0 \end{cases}, \quad (2.5)$$

onde $P^0 = I$ é matriz identidade de ordem $m-1$.

As probabilidades de transição do processo agregado dadas por (2.3) são equivalentes à expressão (2.5). Isso pode ser mostrado por indução da forma seguinte.

1. A expressão (2.5) é válida para $n = 0$ e $n = 1$ de modo óbvio. Vamos mostrar usando o produto de matrizes, que é válido para $n = 2$. A expressão simplificada (2.3) pode ser escrita na forma

$$P^{t(\mathbf{X})}(1|0^31) = \frac{\sum_{k=2}^m \left[\sum_{j=2}^m p_{1j} \cdot \left(\sum_{i=2}^m p_{ji} \cdot p_{ik} \right) \right] \cdot p_{k1}}{\sum_{k=2}^m \left[\sum_{j=2}^m p_{1j} \cdot \left(\sum_{i=2}^m p_{ji} \cdot p_{ik} \right) \right]}.$$

Pela definição de produto de matrizes

$$\sum_{i=2}^m p_{ji} \cdot p_{ik} = P^2$$

e

$$P^{t(\mathbf{X})}(1|0^3 1) = \frac{V \cdot P^2 \cdot W'}{V \cdot P^2 \cdot \mathbf{1}'}$$

2. Supondo (2.5) válida para n , do mesmo modo,

$$P^{t(\mathbf{X})}(1|0^{n+1} 1) = \frac{V \cdot P^n \cdot W'}{V \cdot P^n \cdot \mathbf{1}'}$$

Então, para $n + 1$ teremos

$$P^{t(\mathbf{X})}(1|0^{n+2} 1) = \frac{\sum_{i_{n+1}=2}^m \left[\sum_{j=2}^m p_{1j} \cdot \left(\sum_{i_n=2}^m \cdots \sum_{i_2=2}^m \sum_{i_1=2}^m p_{j i_1} \cdot p_{i_1 i_2} \cdots p_{i_n i_{n+1}} \right) \right] \cdot p_{i_{n+1}}}{\sum_{i_{n+1}=2}^m \left[\sum_{j=2}^m p_{1j} \cdot \left(\sum_{i_n=2}^m \cdots \sum_{i_2=2}^m \sum_{i_1=2}^m p_{j i_1} \cdot p_{i_1 i_2} \cdots p_{i_n i_{n+1}} \right) \right]}$$

Pela definição de produto de matrizes, de modo semelhante, teremos

$$P^{t(\mathbf{X})}(1|0^{n+2} 1) = \frac{V \cdot P^{n+1} \cdot W'}{V \cdot P^{n+1} \cdot \mathbf{1}'}$$

Mostrando assim, a equivalência entre (2.3) e (2.5).

2.3 Agregabilidade De Um Processo Markoviano

As condições sob as quais o processo $t(\mathbf{X})$ continua sendo um processo Markoviano foi estudado por vários autores como, Burke & Rosenblatt (1958), Rubino & Sericola (1989), Rubino & Sericola (1991), Ball & Yeo (1993), Buchholz (1994), Ledoux *et al.* (1994), Peng (1996), Gurvits & Ledoux (2005), e outros. Utilizaremos a versão de Kemeny & Snell (1976).

Definição 2.3 *Um processo Markoviano é dito ser **agregável** com respeito à partição $\{A_1, A_2, \dots, A_k\}$ de A , se o processo agregado definido pela transformação sobrejetiva $g(x) = i \cdot \mathbf{1}(x \in A_i)$ é um processo Markoviano.*

Teorema 2.1 *Uma condição necessária e suficiente para um processo Markoviano \mathbf{X} ser agregável com respeito à partição $\{A_1, A_2, \dots, A_k\}$ é que, para todo par de conjuntos A_i, A_j e cada $b \in A_j$*

$$P^{\mathbf{X}}(b|a) = \mathbb{P}(X_0 = b | X_{-1} = a) = c, \text{ para todo } a \in A_i.$$

Esses valores comuns formam a matriz de transição $p_{ij}^{t(\mathbf{X})}$ do processo agregado.

Note que a condição de agregabilidade do Teorema 2.1 no caso binário, pode ser escrita, utilizando a forma matricial, como $W' = c \cdot \mathbf{1}'$, isto implica que $p_{i1} = \mathbb{P}(X_0 = 1 | X_{-1} = i) = c$ é constante

para todo $i \in A \setminus \{1\}$. Assim, se a condição do Teorema 2.1 é satisfeita com respeito à partição $\{\{1\}, \{2, 3, \dots, m\}\}$, o processo agregado $t(\mathbf{X})$ é Markoviano com matriz de transição dada por

$$P^{t(\mathbf{X})} = \begin{bmatrix} p_{11} & 1 - p_{11} \\ c & 1 - c \end{bmatrix}.$$

Observe que a condição de agregabilidade dada no Teorema 2.1 acima, se refere a um processo Markoviano de ordem 1. Outras perguntas que surgem estão apresentadas a seguir.

- (1) A partir de um processo Markoviano de ordem 1, é possível construir um processo Markoviano de qualquer ordem finita $k \neq 1$?
- (2) Quais condições um processo Markoviano e uma partição do alfabeto devem satisfazer para que seja possível produzir um processo Markoviano de ordem finita?
- (3) No caso de um processo não agregável, que conclusões podemos tirar quanto à continuidade das probabilidades de transição do processo agregado? Quais condições um processo não agregável deve satisfazer, para que o processo agregado obtido possua probabilidades de transição contínuas? Neste caso, qual a velocidade de convergência de tais probabilidades condicionais quando $n \rightarrow \infty$? A continuidade será definida na próxima Seção.
- (4) As medidas de probabilidade do processo agregado mantém a propriedade Gibbsiana? Quais são as condições sob as quais a Gibbsianidade é garantida? A Gibbsianidade será definida na Seção 3.4.

As perguntas de (1) a (3) são respondidas na Seção 3.3, a pergunta (4) é abordada na Seção 3.4 deste capítulo.

Observamos, através da abordagem matricial, que para o processo agregado $t(\mathbf{X})$ ser Markoviano de ordem k , temos que verificar a igualdade

$$P^{t(\mathbf{X})}(1|0^{n+1}1) = \frac{V.P^n.W'}{V.P^n.\mathbf{1}'} = \frac{V.P^{k-1}.W'}{V.P^{k-1}.\mathbf{1}'} = P^{t(\mathbf{X})}(1|0^k1), \text{ com } k \leq n, \text{ para todo } n.$$

2.4 Continuidade Das Probabilidades De Transição

Dedicamos esta seção a encontrar as condições sobre $P^{\mathbf{X}}$ para a continuidade das probabilidades de transição do processo agregado $t(\mathbf{X})$. Lembremos, inicialmente, a definição de continuidade para processos estocásticos.

Definição 2.4 A taxa de continuidade $(\beta^{\mathbf{Y}}(n))_{n \in \mathbb{N}}$ do processo estocástico \mathbf{Y} é definida como

$$\beta^{\mathbf{Y}}(n) = \sup_{\mathbf{x}, \mathbf{y}} \{|P^{\mathbf{Y}}(x_0|x_{-\infty}^{-1}) - P^{\mathbf{Y}}(x_0|x_{-n}^{-1}, y_{-\infty}^{-n-1})|\}.$$

Definição 2.5 A probabilidade de transição $P^{\mathbf{Y}}$ é dita ser contínua se

$$\beta^{\mathbf{Y}}(n) \rightarrow 0, \text{ quando } n \rightarrow \infty.$$

Como, em nosso contexto, o alfabeto é $A = \{0, 1\}$, basta considerar $x_0 = 1$, pois $P^{\mathbf{Y}}(1|x_{-\infty}^{-1}) = 1 - P^{\mathbf{Y}}(0|x_{-\infty}^{-1})$. Então

$$\begin{aligned} \beta^{\mathbf{Y}}(n) &= \sup_{\mathbf{x}, \mathbf{y}} \{ |P^{\mathbf{Y}}(x_0|x_{-\infty}^{-1}) - P^{\mathbf{Y}}(x_0|x_{-n}^{-1}, y_{-\infty}^{-n-1})| \} \\ &= \sup_{\mathbf{x}, \mathbf{y}} \{ |\mathbb{P}(Y_0 = x_0 | Y_{-\infty}^{-1} = x_{-\infty}^{-1}) - \mathbb{P}(Y_0 = x_0 | Y_{-n}^{-1} = x_{-n}^{-1}, Y_{-\infty}^{-n-1} = y_{-\infty}^{-n-1})| \} \\ &= \sup_{\mathbf{x}, \mathbf{y}} \{ |\mathbb{P}(Y_0 = 1 | Y_{-\infty}^{-1} = x_{-\infty}^{-1}) - \mathbb{P}(Y_0 = 1 | Y_{-n}^{-1} = x_{-n}^{-1}, Y_{-\infty}^{-n-1} = y_{-\infty}^{-n-1})| \}. \end{aligned}$$

Como o Processo \mathbf{Y} renova em 1, o supremo é atingido quando $x_{-\infty}^{-1} = 0_{-\infty}^{-1} = \underline{0}$, caso contrário, haveria um símbolo 1, esqueceríamos o passado a partir dele e, desta forma, teríamos $\beta^{\mathbf{Y}}(n) = 0$. Assim, para $l \geq n$, $x_{-l}^{-1} = 0_{-l}^{-1}$ e $Y_{-l-1} = 1$, podemos escrever

$$\begin{aligned} \beta^{\mathbf{Y}}(n) &= \sup_{\mathbf{y}} \{ |\mathbb{P}(Y_0 = 1 | Y_{-\infty}^{-1} = 0_{-\infty}^{-1}) - \mathbb{P}(Y_0 = 1 | Y_{-n}^{-1} = 0_{-n}^{-1}, Y_{-n-1} = 1, Y_{-\infty}^{-n-2} = y_{-\infty}^{-n-2})| \} \\ &= \sup_{l \geq n} \{ |\mathbb{P}(Y_0 = 1 | Y_{-\infty}^{-1} = 0_{-\infty}^{-1}) - \mathbb{P}(Y_0 = 1 | Y_{-l}^{-1} = 0_{-l}^{-1}, Y_{-l-1} = 1)| \} \\ &= \sup_{l \geq n} \{ |P^{\mathbf{Y}}(1|0_{-\infty}^{-1}) - P^{\mathbf{Y}}(1|0_{-l}^{-1}1)| \} \\ &= \sup_{l \geq n} \{ |P^{\mathbf{Y}}(1|\underline{0}) - P^{\mathbf{Y}}(1|0^l1)| \}. \end{aligned}$$

Harris (1955), citado por Verbitskiy (2011a) e Verbitskiy (2011b), mostrou que sob a condição de positividade das probabilidades de transição do processo original, processos agregados que são funções de processos Markovianos, como a Transformação Agregante t que estamos estudando, têm probabilidades de transição contínuas.

Proposição 2.2 Para todo $t(\mathbf{X}) = \mathbf{Y} = (Y_n)_{n \in \mathbb{Z}}$, se $P^{\mathbf{X}} = p_{ij} > 0$, qualquer que seja $i, j \in \mathbb{N}$, então

$$\mathbb{P}(Y_0 = y_0 | Y_{-\infty}^{-1} = y_{-\infty}^{-1}) := \lim_{n \rightarrow \infty} \mathbb{P}(Y_0 = y_0 | Y_{-n}^{-1} = y_{-n}^{-1}),$$

para todo $n \in \mathbb{N}$ e $y_{-\infty}^{-1}, y_{-n}^{-1} \in B^{\mathbb{N}}$, B alfabeto de \mathbf{Y} . Existe $0 \leq c < 1$, tal que

$$\beta^{\mathbf{Y}}(n) \leq (c)^n.$$

Outra indagação que naturalmente surge é, a condição $P^{\mathbf{X}} = p_{ij} > 0$ é necessária para a continuidade das probabilidades de transição? que outra condição, mais fraca, garante a continuidade das probabilidades de transição num processo agregado de ordem infinita?

O Caso Primitivo

O Teorema 2.2 dá uma condição suficiente mais fraca que a positividade, para a continuidade das probabilidades de transição do processo agregado $t(\mathbf{X})$. Uma matriz irreduzível aperiódica é aqui chamada de **primitiva**; se ela é periódica então a chamamos de **cíclica**. Maiores detalhes são apresentados no Apêndice B. Utilizaremos a abordagem matricial desenvolvida na Seção 3.1, bem como a Teoria de Perron-Frobenius apresentada no Apêndice B.

Teorema 2.2 *Seja $(X_n)_{n \in \mathbb{N}}$ um processo Markoviano estacionário, com espaço de estados $A = \{1, 2, \dots, m\}$ e matriz das probabilidades de transição decomposta na forma (2.4). Suponha a submatriz P irreduzível.*

Então, o processo agregado $t(\mathbf{X})$ possui probabilidades de transição contínuas se a submatriz P é primitiva.

E ainda

1. *A taxa de continuidade $\beta^{t(\mathbf{X})}(n)$ das probabilidades de transição do processo agregado é da ordem de*

$$n^{d_2-1} \cdot \left(\frac{|\lambda_2|}{\lambda_1} \right)^n,$$

onde λ_1 é o maior auto-valor de P , λ_2 é o segundo maior auto-valor, em módulo, de P e d_2 é sua respectiva multiplicidade algébrica.

2. *O processo agregado $t(\mathbf{X})$ é processo Markoviano de ordem finita $1 < k \leq m-1$ se a submatriz P possui um único auto-valor não-nulo e todos os demais nulos, com k linhas não nulas em sua forma triangularizada.*

Prova do Teorema 2.2

Já vimos que a submatriz P é uma matriz sub-estocástica, logo é não negativa, como requer o Teorema de Perron-Frobenius. Considere $\lambda_1, \lambda_2, \dots, \lambda_r$ todos os auto-valores de P , ordenados de tal forma que

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_r|. \quad (2.6)$$

Se $|\lambda_2| = |\lambda_j|$ para algum $j \geq 3$, então $d_2 \geq d_j$, onde d_j é a multiplicidade algébrica de λ_j . Assim, se P é primitiva, temos um único auto-valor igual ao seu raio espectral, $\lambda_1 > 0$ e

$$\lim_{n \rightarrow \infty} P^{t(\mathbf{X})}(1|0^{n+1}1) = \lim_{n \rightarrow \infty} \frac{V.P^n.W'}{V.P^n.\mathbf{1}'} = \lim_{n \rightarrow \infty} \frac{V.\left(\frac{P}{\lambda_1}\right)^n.W'}{V.\left(\frac{P}{\lambda_1}\right)^n.\mathbf{1}'}$$

Pelo Teorema B.4,

$$\lim_{n \rightarrow \infty} \left(\frac{P}{\lambda_1} \right)^n = G$$

e

$$\lim_{n \rightarrow \infty} P^{t(\mathbf{X})}(1|0^{n+1}1) = \frac{V.G.W'}{V.G.1'} \quad (2.7)$$

com $G = \phi.\psi' > 0$, onde ϕ e ψ são os auto-vetores correspondentes a λ_1 de P e P' , respectivamente. Como $l \geq n + 1$, $\beta^{t(\mathbf{X})}(n) \rightarrow 0$ e temos a continuidade.

Para provar o Ítem 1, utilizaremos a primeira parte do Teorema B.5,

$$\left(\frac{P}{\lambda_1}\right)^n = G + O\left(n^{d_2-1} \cdot \left(\frac{|\lambda_2|}{\lambda_1}\right)^n\right), \quad (2.8)$$

onde $O\left(n^{d_2-1} \cdot \left(\frac{|\lambda_2|}{\lambda_1}\right)^n\right) \rightarrow 0$, quando $n \rightarrow \infty$.

Para mostrar que

$$\beta^{t(\mathbf{X})}(n) = \sup_{l \geq n+1} \{|P^{\mathbf{Y}}(1|0) - P^{\mathbf{Y}}(1|0^l1)|\} = O\left(n^{d_2-1} \cdot \left(\frac{|\lambda_2|}{\lambda_1}\right)^n\right),$$

basta mostrar que

$$\sup_{l \geq n+1} \left\{ \frac{|P^{t(\mathbf{X})}(1|0^l1) - \frac{V.G.W'}{V.G.1'}|}{l^{d_2-1} \cdot \left(\frac{|\lambda_2|}{\lambda_1}\right)^l} \right\} \rightarrow \text{constante.}$$

É suficiente mostrar que o argumento

$$\frac{|P^{t(\mathbf{X})}(1|0^l1) - \frac{V.G.W'}{V.G.1'}|}{l^{d_2-1} \cdot \left(\frac{|\lambda_2|}{\lambda_1}\right)^l} \rightarrow \text{constante}$$

quando $n, l \rightarrow \infty$. Seja $K_l = l^{d_2-1} \cdot \left(\frac{|\lambda_2|}{\lambda_1}\right)^l$, então

$$\begin{aligned} \frac{\left| \frac{V.P^l.W'}{V.P^l.1'} - \frac{V.G.W'}{V.G.1'} \right|}{K_l} &= \frac{\left| \frac{V \cdot \left(\frac{P}{\lambda_1}\right)^l \cdot W'}{V \cdot \left(\frac{P}{\lambda_1}\right)^l \cdot 1'} - \frac{V.G.W'}{V.G.1'} \right|}{K_l} \\ &= \frac{\left| \frac{V.(G+O(K_l)).W'}{V.(G+O(K_l)).1'} - \frac{V.G.W'}{V.G.1'} \right|}{K_l} \\ &= \frac{\left| \frac{V.G.W'+V.O(K_l).W'}{V.G.1'+V.O(K_l).1'} - \frac{V.G.W'}{V.G.1'} \right|}{K_l} \\ &= \left| \frac{\frac{V.G.W'.V.G.1'+V.O(K_l).W'.V.G.1'-V.G.W'.V.G.1'-V.G.W'.V.O(K_l).1'}{(V.G.1')^2+V.G.1'.V.O(K_l).1'}}{K_l} \right| \\ &= \left| \frac{\frac{V.G.1'.V.O(K_l).W'-V.G.W'.V.O(K_l).1'}{K_l}}{(V.G.1')^2 + V.G.1'.V.O(K_l).1'} \right| \rightarrow \text{constante.} \end{aligned}$$

Pois $V.O(K_l).\mathbf{1}' \rightarrow 0$ quando $n \rightarrow \infty$.

Portanto a taxa de continuidade das probabilidades de transição do processo agregado é dada pelos dois maiores auto-valores, em valor absoluto, da submatriz P .

Para provar o Ítem 2, note que no caso não agregável temos $k \neq 1$. E ainda, como P é irredutível primitiva, se $\lambda_2 = 0$, pela ordenação (2.6), P apresenta apenas λ_1 não-nulo e todos os demais nulos. Então, de acordo com a segunda parte do Teorema B.5, é válida a igualdade

$$P^n = \lambda_1^n \cdot \phi \cdot \psi'$$

com $n \geq m - 1$. Então teremos

$$P^{t(\mathbf{X})}(1|0^{n+1}\mathbf{1}) = \frac{V.P^n.W'}{V.P^n.\mathbf{1}'} = \frac{V.\lambda_1^n \cdot \phi \cdot \psi'.W'}{V.\lambda_1^n \cdot \phi \cdot \psi'.\mathbf{1}'} = \frac{V.G.W'}{V.G.\mathbf{1}'}$$

Note que o processo agregado $t(\mathbf{X})$ é Markoviano de ordem $\leq m - 1$.

Note ainda que, segundo Lima (1996) (pág.283), se os auto-valores de P são todos reais, P é *triangularizável*, isto é, existem Q e Δ , tais que

$$P = Q \cdot \Delta \cdot Q',$$

onde Δ é triangular superior e Q é ortogonal. Mais ainda, os auto-valores de Δ são os elementos de sua diagonal principal. Suponha ainda que P tenha k linhas não nulas em Δ . Note que Δ satisfaz as condições do Lema A.2 presente no Apêndice A, então teremos

$$\Delta^n = \lambda_1^{n-k+1} \cdot \Delta^{k-1},$$

assim

$$\begin{aligned} P^{t(\mathbf{X})}(1|0^{n+1}\mathbf{1}) &= \frac{V.P^n.W'}{V.P^n.\mathbf{1}'} \\ &= \frac{V.(Q \cdot \Delta \cdot Q')^n.W'}{V.(Q \cdot \Delta \cdot Q')^n.\mathbf{1}'} \\ &= \frac{V.Q \cdot (\Delta)^n \cdot Q'.W'}{V.Q \cdot (\Delta)^n \cdot Q'.\mathbf{1}'} \\ &= \frac{V.Q \cdot \lambda_1^{n-k+1} \cdot \Delta^{k-1} \cdot Q'.W'}{V.Q \cdot \lambda_1^{n-k+1} \cdot \Delta^{k-1} \cdot Q'.\mathbf{1}'} \\ &= \frac{V.(Q \cdot \Delta \cdot Q')^{k-1}.W'}{V.(Q \cdot \Delta \cdot Q')^{k-1}.\mathbf{1}'} \\ &= \frac{V.P^{k-1}.W'}{V.P^{k-1}.\mathbf{1}'} \\ &= P^{t(\mathbf{X})}(1|0^k\mathbf{1}). \end{aligned}$$

O processo agregado $t(\mathbf{X})$ é Markoviano de ordem $k \leq m - 1$. ■

O Caso Cíclico

O nosso objetivo é buscar uma condição necessária e suficiente para a continuidade das probabilidades de transição do processo agregado $t(\mathbf{X})$, sem impor hipóteses muito restritivas sobre o processo original \mathbf{X} . Abordaremos agora o caso em que a submatriz P é irredutível e cíclica de índice h . Neste contexto utilizaremos a forma de Frobênius de uma matriz cíclica (B.2) apresentada no Apêndice B, Teorema B.6.

Teorema 2.3 *Seja $(X_n)_{n \in \mathbb{N}}$ um processo Markoviano estacionário, com espaço de estados $A = \{1, 2, \dots, m\}$ e matriz das probabilidades de transição decomposta na forma (2.4). Suponha a submatriz P irredutível e cíclica de índice h . O processo agregado $t(\mathbf{X})$ possui probabilidades de transição contínuas se, somente se*

$$\frac{V.G^*.P^r.W'}{V.G^*.P^r.\mathbf{1}'} = C \quad (2.9)$$

com C constante para todo $1 \leq r \leq h - 1$, onde $G^* = \lim_{n \rightarrow \infty} \left(\frac{P^n}{\lambda_1^*} \right)^n$ e λ_1^* é o auto-valor maximal de P^h .

Prova do Teorema 2.3

Note que a matriz sub-estocástica P é cíclica, cujos blocos cíclicos são submatrizes primitivas. Escrevendo P na forma de Frobênius (B.2)

$$P = \begin{bmatrix} \mathbf{0} & B_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & B_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & B_{h-1} \\ B_h & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \end{bmatrix}. \quad (2.10)$$

Suas potências apresentarão as formas

$$P^2 = \begin{bmatrix} \mathbf{0} & \mathbf{0} & B_1.B_2 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & B_2.B_3 & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ B_{h-1}.B_h & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & B_h.B_1 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \end{bmatrix} \quad (2.11)$$

e assim sucessivamente até

$$P^h = \begin{bmatrix} B_1.B_2.\dots.B_h & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & B_2.B_3.\dots.B_h.B_1 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & B_h.B_1.\dots.B_{h-1} \end{bmatrix}. \quad (2.12)$$

Note que P^h é primitiva, então, pelo Teorema B.4

$$\lim_{n \rightarrow \infty} \left(\frac{P^h}{\lambda_1^*} \right)^n = G^*$$

com $G^* = \phi.\psi' > 0$, onde ϕ e ψ são os auto-vetores correspondentes a λ_1^* de P^h e $(P^h)'$, respectivamente. Então, podemos escrever

$$P^{t(\mathbf{X})}(1|0^{n.h+1}1) = \frac{V.P^{n.h}.W'}{V.P^{n.h}.\mathbf{1}'} = \frac{V.\left(\frac{P^h}{\lambda_1^*}\right)^n.W'}{V.\left(\frac{P^h}{\lambda_1^*}\right)^n.\mathbf{1}'}$$

e

$$P^{t(\mathbf{X})}(1|0^{n.h+r+1}1) = \frac{V.P^{n.h+r}.W'}{V.P^{n.h+r}.\mathbf{1}'} = \frac{V.\left(\frac{P^h}{\lambda_1^*}\right)^n.P^r.W'}{V.\left(\frac{P^h}{\lambda_1^*}\right)^n.P^r.\mathbf{1}'},$$

assim

$$\lim_{n \rightarrow \infty} P^{t(\mathbf{X})}(1|0^{n.h+r+1}1) = \lim_{n \rightarrow \infty} \frac{V.\left(\frac{P^h}{\lambda_1^*}\right)^n.P^r.W'}{V.\left(\frac{P^h}{\lambda_1^*}\right)^n.P^r.\mathbf{1}'} = \frac{V.G^*.P^r.W'}{V.G^*.P^r.\mathbf{1}'} = C(r).$$

Logo haverá continuidade das probabilidades de transição do processo agregado $t(\mathbf{X})$ se e somente se, $C(r) = C = \text{constante}$ para todo $r \in \{0, \dots, h-1\}$, pois, da mesma forma que o caso primitivo, $l \geq n.h+r+1$ e $\beta^{t(\mathbf{X})}(n) \rightarrow 0$, garantindo a continuidade das probabilidades de transição. ■

Observe que, se a submatriz P , cíclica de índice h , apresenta os blocos cíclicos idênticos, podemos verificar que a condição do Teorema 2.3 é satisfeita. De fato, escreva P na forma

$$P = \begin{bmatrix} \mathbf{0} & B & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & B & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & B \\ B & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \end{bmatrix}.$$

A sua potência h será

$$P^h = \begin{bmatrix} B^h & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & B^h & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & B^h \end{bmatrix}.$$

Note que P^h é primitiva com auto-valor maximal igual a $(\lambda_1)^h$, onde λ_1 é o auto-valor maximal do bloco B . Então, pelo Teorema B.4

$$\lim_{n \rightarrow \infty} \left(\frac{P}{\lambda_1} \right)^{n.h} = G^*.$$

Os blocos B são matrizes primitivas, quadradas e idênticas, então quando $n \rightarrow \infty$, G^* é formada pelos blocos limites $\lim_{n \rightarrow \infty} \left(\frac{B}{\lambda_1} \right)^{n.h} = H^* = \phi.\psi'$, na diagonal principal e zero nas outras entradas. onde ϕ e ψ são os auto-vetores à direita e à esquerda, respectivamente, de B e B' , correspondente à λ_1 . Operando em blocos, vamos escrever V_i e W'_i , $i \in \{1, 2, \dots, h\}$, como subvetores, respectivamente, de V e W' , com a mesma dimensão de B . Então

$$\begin{aligned} \lim_{n \rightarrow \infty} P^{t(\mathbf{X})}(1|0^{n.h+r+1}1) &= \lim_{n \rightarrow \infty} \frac{V. \left(\frac{P}{\lambda_1} \right)^{n.h} . P^r . W'}{V. \left(\frac{P}{\lambda_1} \right)^{n.h} . P^r . \mathbf{1}'} \\ &= \lim_{n \rightarrow \infty} \frac{\sum_{i=0}^{h-1} V_i. \left(\frac{B}{\lambda_1} \right)^{n.h} . B^r . W'_i}{\sum_{i=0}^{h-1} V_i. \left(\frac{B}{\lambda_1} \right)^{n.h} . B^r} \\ &= \frac{\sum_{i=0}^{h-1} V_i. H^* . B^r . W'_i}{\sum_{i=0}^{h-1} V_i. H^* . B^r} \\ &= \frac{\sum_{i=0}^{h-1} V_i. \phi. \psi' . B^r . W'_i}{\sum_{i=0}^{h-1} V_i. \phi. \psi' . B^r}. \end{aligned}$$

Note que ϕ e ψ são auto-vetores, respectivamente de B e B' , correspondente à λ_1 , então

$$\begin{aligned} \phi. \psi' . B^r &= \phi. \psi' . B. B^{r-1} \\ &= \phi. \lambda_1. \psi' . B^{r-1} \\ &= \lambda_1. H^* . B^{r-1}, \end{aligned}$$

assim, sucessivamente, chegamos a

$$\phi. \psi' . B^r = (\lambda_1)^r . H^*.$$

Substituindo acima, temos então

$$\begin{aligned}
\lim_{n \rightarrow \infty} P^{t(\mathbf{X})}(1|0^{n.h+r+1}1) &= \frac{\sum_{i=0}^{h-1} V_i \cdot (\lambda_1)^r \cdot H^* \cdot W'_i}{\sum_{i=0}^{h-1} V_i \cdot (\lambda_1)^r \cdot H^*} \\
&= \frac{\sum_{i=0}^{h-1} V_i \cdot H^* \cdot W'_i}{\sum_{i=0}^{h-1} V_i \cdot H^*} \\
&= \frac{V \cdot G^* \cdot W'}{V \cdot G^* \cdot \mathbf{1}'} = C.
\end{aligned}$$

como C não depende de r , a Condição (2.9) é satisfeita e as probabilidades de transição do processo agregado são contínuas.

O Exemplo 2.4 da Seção 3.5, é outro exemplo que ilustra o Teorema 2.3 mostrando como a condição do teorema pode ser verificada, num processo Markoviano cuja matriz de transição apresenta submatriz P , cíclica, com soma constante nas colunas.

O Caso Redutível

Estudaremos agora o caso em que a submatriz P é redutível. A forma canônica de uma matriz redutível P é escrita como (vide Apêndice B)

$$P = \begin{bmatrix} T_{11} & T_{12} & T_{13} & \dots & T_{1c} \\ \mathbf{0} & T_{22} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & T_{33} & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & T_{cc} \end{bmatrix}, \quad (2.13)$$

onde as submatrizes $T_{22}, T_{33}, \dots, T_{cc}$ são submatrizes quadradas e irredutíveis.

Definição 2.6 *As submatrizes irredutíveis $T_{ii}, i \in \{2, 3, \dots, c\}$ cujo auto-valor é o máximo*

$$\lambda_{1,ii} = \lambda_{1,max} = \max_i |\lambda_{1,i}|,$$

são chamadas de Blocos Dominantes.

Combinando os resultados obtidos nos casos irredutível primitivo e cíclico da submatriz P obtemos o resultado a seguir.

Teorema 2.4 *Seja $(X_n)_{n \in \mathbb{N}}$ um processo Markoviano não agregável, estacionário, com espaço de estados $A = \{1, 2, \dots, m\}$ e matriz das probabilidades de transição decomposta na forma (2.4). Sem perda de generalidade, assumamos que os estados em $A \setminus \{1\}$ sejam ordenados de forma que a submatriz*

P , redutível, seja escrita na forma canônica (B.3). O processo agregado $t(\mathbf{X})$ possui probabilidades de transição contínuas se

i. Os blocos dominantes $T_{ii} \in \{T_{22}, T_{33}, \dots, T_{cc}\}$ são todos primitivos.

ii. Os blocos dominantes $T_{ii} \in \{T_{22}, T_{33}, \dots, T_{cc}\}$ cíclicos, possuem o mesmo índice h e satisfazem

$$\frac{V_i \cdot G_i \cdot (T_{ii})^r \cdot W_i'}{V_i \cdot G_i \cdot (T_{ii})^r \cdot \mathbf{1}'_i} = C_i \quad (2.14)$$

com C_i constante para todo $1 \leq r \leq h-1$, $G_i = \lim_{n \rightarrow \infty} \left(\frac{(T_{ii})^h}{\lambda_{1,ii}^*} \right)^n$, $\lambda_{1,ii}^*$ é o auto-valor maximal de $(T_{ii})^h$, $i \in \{2, 3, \dots, c\}$ e T_{ii} são blocos irredutíveis dominantes.

Prova do Teorema 2.4

Utilizaremos raciocínio análogo ao caso irredutível primitivo da submatriz P , para provar o Ítem i.

Como as submatrizes $T_{22}, T_{33}, \dots, T_{cc}$ são todas primitivas, então, pelo Teorema B.7 existe o limite

$$\lim_{n \rightarrow \infty} \left(\frac{P}{\lambda_{1,max}} \right)^n = \begin{bmatrix} \mathbf{0} & (I - T_{11})^{-1} \cdot T_{12} \cdot G_2 & (I - T_{11})^{-1} \cdot T_{13} \cdot G_3 & \dots & (I - T_{11})^{-1} \cdot T_{1c} \cdot G_c \\ \mathbf{0} & G_2 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & G_3 & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & G_c \end{bmatrix} = \widehat{G}.$$

Considerando $\lambda_{1,max} = \max_i \{\lambda_{1,ii}\}$ e definindo os blocos V_i , W_i' e $\mathbf{1}'_i$, $i \in \{2, 3, \dots, c\}$, com dimensões compatíveis com G_i . Operando em blocos, teremos

$$G_i = \lim_{n \rightarrow \infty} \left(\frac{T_{ii}}{\lambda_{1,max}} \right)^n = \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^c V_i \cdot \left(\frac{T_{ii}}{\lambda_{1,ii}} \right)^n \cdot \left(\frac{\lambda_{1,ii}}{\lambda_{1,max}} \right)^n \cdot W_i'}{\sum_{i=1}^c V_i \cdot \left(\frac{T_{ii}}{\lambda_{1,ii}} \right)^n \cdot \left(\frac{\lambda_{1,ii}}{\lambda_{1,max}} \right)^n \cdot \mathbf{1}'_i},$$

onde

$$\lim_{n \rightarrow \infty} \left(\frac{\lambda_{1,ii}}{\lambda_{1,max}} \right)^n = \mathbb{I}(\lambda_{1,ii} = \lambda_{1,max}), \quad (2.15)$$

\mathbb{I} representa a função indicadora. Finalmente, então teremos

$$\lim_{n \rightarrow \infty} P^{t(\mathbf{X})}(1|0^{n+1}1) = \lim_{n \rightarrow \infty} \frac{V \cdot P^n \cdot W'}{V \cdot P^n \cdot \mathbf{1}'} = \frac{V \cdot \widehat{G} \cdot W'}{V \cdot \widehat{G} \cdot \mathbf{1}'},$$

onde \widehat{G} é dado acima e

$$\begin{bmatrix} G_2 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & G_3 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & G_c \end{bmatrix} = \begin{bmatrix} \mathbb{I}(\lambda_{1,22} = \lambda_{1,max}) \cdot \phi_{22} \cdot \psi'_{22} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbb{I}(\lambda_{1,cc} = \lambda_{1,max}) \cdot \phi_{cc} \cdot \psi'_{cc} \end{bmatrix}.$$

Para provar o Ítem *ii.*, utilizaremos o mesmo raciocínio utilizado no caso irredutível cíclico da submatriz P . Escrevendo P na forma canônica, chamaremos os blocos dominantes de T_{ii} . Como são cíclicos de índice h , podemos escrevê-los na forma

$$T_{ii} = \begin{bmatrix} \mathbf{0} & B_{1,i} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & B_{2,i} & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & B_{h-1,i} \\ B_{h,i} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \end{bmatrix}.$$

T_{ii}^h é primitiva, então, pelo Teorema B.4

$$\lim_{n \rightarrow \infty} \left(\frac{T_{ii}^h}{\lambda_{1,max}^*} \right)^n = G_i^*$$

com $G_i^* = \phi_i \cdot \psi'_i > 0$, onde ϕ_i e ψ_i são os auto-vetores correspondentes a $\lambda_{1,max}^*$ de T_{ii}^h e $(T_{ii}^h)'$, respectivamente, e $\lambda_{1,max}^*$ é o autovalor maximal de T_{ii}^h .

Sabemos pelo Teorema B.7 que, se o limite $G_i = \lim_{n \rightarrow \infty} \left(\frac{T_{ii}}{\lambda_{1,max}} \right)^{n.h}$ existir, teremos

$$\lim_{n \rightarrow \infty} \left(\frac{P}{\lambda_{1,max}} \right)^{n.h} = \begin{bmatrix} \mathbf{0} & (I - T_{11})^{-1} \cdot T_{12} \cdot G_2 & (I - T_{11})^{-1} \cdot T_{13} \cdot G_3 & \dots & (I - T_{11})^{-1} \cdot T_{1c} \cdot G_c \\ \mathbf{0} & G_2 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & G_3 & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & G_c \end{bmatrix} = \widehat{G},$$

onde

$$G_i = \lim_{n \rightarrow \infty} \left(\frac{T_{ii}}{\lambda_{1,max}} \right)^{n.h} = \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^c V_i \cdot \left(\frac{T_{ii}}{\lambda_{1,ii}} \right)^{n.h} \cdot \left(\frac{\lambda_{1,ii}}{\lambda_{1,max}} \right)^{n.h} \cdot W'_i}{\sum_{i=1}^c V_i \cdot \left(\frac{T_{ii}}{\lambda_{1,ii}} \right)^{n.h} \cdot \left(\frac{\lambda_{1,ii}}{\lambda_{1,max}} \right)^{n.h} \cdot \mathbf{1}'_i}.$$

Então, a condição para a convergência do limite acima é a convergência dos blocos dominantes. De acordo com o caso cíclico das matrizes irredutíveis, a condição necessária e suficiente para esta

convergência é que

$$\frac{V_i \cdot G_i \cdot (T_{ii})^r \cdot W_i'}{V \cdot G_i \cdot (T_{ii})^r \cdot \mathbf{1}'_i} = C_i(r)$$

com $C_i(r)$ constante para todo $1 \leq r \leq h-1$ e T_{ii} , $i \in \{2, 3, \dots, c\}$ são os blocos dominantes. ■

A taxa de continuidade $\beta^{t(\mathbf{X})}(n)$ das probabilidades de transição do processo agregado no caso redutível é da ordem de

$$n^{d_{2,max}-1} \cdot \left(\frac{|\lambda_{2,max}|}{\lambda_{1,max}} \right)^n,$$

onde $\lambda_{2,max}$ é o segundo maior auto-valor do bloco dominante e $d_{2,max}$ é sua respectiva multiplicidade algébrica. A demonstração deste fato é análoga ao caso irredutível.

2.5 Gibbsianidade

Nesta tese consideramos a noção de Gibbsianidade introduzida por Dobruschin (1968), a qual está relacionada às propriedades estudadas no presente capítulo para probabilidades de transição, condicionando no passado e no futuro. Uma caracterização desta Gibbsianidade, que é ao mesmo tempo, simples e adequada para o propósito desta tese, é dada a seguir.

Definição 2.7 *Uma medida de probabilidade não nula, sobre $(\Omega, \mathcal{F}, \mathbb{P})$, associada ao processo $(X_n)_{n \in \mathbb{Z}}$ é uma medida Gibbsiana se, e somente se as sequências*

$$[\mathbb{P}(X_0 = x_0 | X_{-n}^{-1} = x_{-n}^{-1}, X_1^r = x_1^r)]_{n,r \geq 1} \text{ convergem uniformemente quando } n, r \rightarrow \infty$$

para todo $x_0 \in A$, $x_{-n}^{-1} \in A^n$ e $x_1^r \in A^r$.

A principal referência sobre medidas Gibbsianas é Georgii (2011). Uma referência muito ligada ao nosso problema e à nossa abordagem é Fernández *et al.* (2011).

O nosso objetivo nesta seção é verificar as condições que devem ser impostas sobre o processo Markoviano original, para que o processo agregado $t(\mathbf{X}) = \mathbf{Y}$ seja Gibbsiano.

Novamente, suporemos que

$$\lim_{n,r \rightarrow \infty} \mathbb{P}(Y_0 = 1 | Y_{-n-1} = 1, Y_{-n}^{-1} = 0_{-n}^{-1}, Y_1^r = 0_1^r, Y_{r+1} = 1) = P^{\mathbf{Y}}(1 | 0_{-\infty}^{-1} 0_1^{+\infty})$$

sempre que houver convergência.

Para abreviar a notação, vamos escrever

$$P^{\mathbf{Y}}(n, r) = \mathbb{P}(Y_0 = 1 | Y_{-n-1} = 1, Y_{-n}^{-1} = 0_{-n}^{-1}, Y_1^r = 0_1^r, Y_{r+1} = 1).$$

Utilizando a propriedade de Markov, a expressão geral para o cálculo desta medida de probabilidade será

$$\begin{aligned}
P^{\mathbf{Y}}(n, r) &= \mathbb{P}(Y_0 = 1 | Y_{-n-1} = 1, Y_{-n}^{-1} = 0_{-n}^{-1}, Y_1^r = 0^r, Y_{r+1} = 1) \\
&= \mathbb{P}(X_0 = 1 | X_{-n-1} = 1, X_{-n}^{-1} \in t^{-1}(0_{-n}^{-1}), X_1^r \in t^{-1}(0^r), X_{r+1} = 1) \\
&= \frac{\mathbb{P}(X_{-n-1} = 1, X_{-n}^{-1} \in t^{-1}(0_{-n}^{-1}), X_0 = 1, X_1^r \in t^{-1}(0^r), X_{r+1} = 1)}{\mathbb{P}(X_{-n-1} = 1, X_{-n}^{-1} \in t^{-1}(0_{-n}^{-1}), X_1^r \in t^{-1}(0^r), X_{r+1} = 1)} \\
&= \frac{\sum_{i_1, \dots, n \in A \setminus \{1\}} \sum_{j_1, \dots, r \in A \setminus \{1\}} p_{1i_1} \cdot \prod_{u=1}^{n-1} (p_{i_u i_{u+1}}) \cdot p_{i_n 1} \cdot p_{1j_1} \cdot \prod_{v=1}^{r-1} (p_{j_v j_{v+1}}) \cdot p_{j_r 1}}{\sum_{i_1, \dots, n \in A \setminus \{1\}} \sum_{j_1, \dots, r \in A \setminus \{1\}} p_{1i_1} \cdot \prod_{u=1}^{n-1} (p_{i_u i_{u+1}}) \cdot p_{i_n j_1} \cdot \prod_{v=1}^{r-1} (p_{j_v j_{v+1}}) \cdot p_{j_r 1}}.
\end{aligned}$$

Na versão matricial (2.4), esta expressão se torna,

$$\begin{aligned}
&\mathbb{P}(Y_0 = 1 | Y_{-n-2} = 1, Y_{-n-1}^{-1} = 0_{-n-1}^{-1}, Y_1^{r+1} = 0^{r+1}, Y_{r+2} = 1) \\
&= \begin{cases} p_{11}, & \text{se } n, r = -1 \\ \frac{V \cdot P^n \cdot W' \cdot V \cdot P^r \cdot W'}{V \cdot P^{n+r+1} \cdot W'}, & \text{se caso contrário} \end{cases}, \quad (2.16)
\end{aligned}$$

onde $P^0 = I$ e $n, r \in \{0, 1, 2, \dots\}$.

Teorema 2.5 *Seja $(X_n)_{n \in \mathbb{N}}$ um processo Markoviano não agregável, estacionário, com espaço de estados $A = \{1, 2, \dots, m\}$ e matriz das probabilidades de transição decomposta na forma (2.4). Suponha a submatriz P irredutível. O processo agregado $t(\mathbf{X})$ é Gibbsiano se a submatriz P for primitiva. Com essa condição, a taxa de continuidade é da ordem de*

$$n^{d_2-1} \cdot \left(\frac{|\lambda_2|}{\lambda_1} \right)^n,$$

onde λ_1 é o auto-valor maximal de P , λ_2 é o segundo auto-valor maximal de P e d_2 é sua respectiva multiplicidade algébrica.

Prova do Teorema 2.5

A prova é feita de modo semelhante à seção anterior, utilizando o Teorema B.4. O limite $\lim_{n, r \rightarrow \infty} \mathbb{P}(Y_0 = 1 | Y_{-n-2} = 1, Y_{-n-1}^{-1} = 0_{-n-1}^{-1}, Y_1^{r+1} = 0^{r+1}, Y_{r+2} = 1)$ será dado por

$$\begin{aligned}
\lim_{n, r \rightarrow \infty} \frac{V \cdot P^n \cdot W' \cdot V \cdot P^r \cdot W'}{V \cdot P^{n+r+1} \cdot W'} &= \lim_{n, r \rightarrow \infty} \frac{V \cdot \left(\frac{P}{\lambda_1} \right)^n \cdot W' \cdot V \cdot \left(\frac{P}{\lambda_1} \right)^r \cdot W'}{\lambda_1 \cdot V \cdot \left(\frac{P}{\lambda_1} \right)^{n+r+1} \cdot W'} \\
&= \frac{V \cdot G \cdot W' \cdot V \cdot G \cdot W'}{\lambda_1 \cdot V \cdot G \cdot W'} \\
&= \frac{V \cdot G \cdot W'}{\lambda_1}.
\end{aligned}$$

A não nulidade vem do fato de que $G = \phi.\psi' > 0$ e supondo que os vetores $V, W' \neq 0$, para evitar os casos triviais, como já mencionado.

Para obter a taxa de continuidade, basta utilizar a primeira parte do Teorema B.5, onde

$$P^n = \lambda_1^n \cdot G + O(n^{d_2-1} \cdot |\lambda_2|^n).$$

Então a convergência de $\mathbb{P}(Y_0 = 1 | Y_{-n-2} = 1, Y_{-n-1}^{-1} = 0_{-n-1}^{-1}, Y_1^{r+1} = 0^{r+1}, Y_{r+2} = 1)$ possui taxa dada pela convergência de $\left(\frac{P}{\lambda_1}\right)^n$. Assim, de modo análogo ao Teorema 2.2, a velocidade de convergência da medida Gibbsiana é da ordem de

$$n^{d_2-1} \cdot \left(\frac{|\lambda_2|}{\lambda_1}\right)^n.$$

■

Este resultado também pode ser percebido como consequência do Teorema 4.22 de Fernández & Maillard (2004), onde a não nulidade e a taxa exponencial de continuidade das probabilidades de transição, garantem a Gibbsianidade do processo agregado $t(\mathbf{X})$.

2.6 Exemplos

Apresentamos agora alguns exemplos que ilustram os resultados dos teoremas apresentados neste capítulo.

O Exemplo 2.1 é uma aplicação dos Teoremas 2.2 e 2.5, sendo importante por dois motivos

- i. Primeiro, mostra que a condição $P^{\mathbf{X}} = p_{ij} > 0$ da Proposição 2.2 não é necessária para a continuidade das probabilidades de transição do processo agregado.
- ii. Segundo, temos um processo agregado não markoviano de nenhuma ordem e com medida Gibbsiana.

Exemplo 2.1 *Considere o processo Markoviano \mathbf{X} com espaço de estados $A = \{1, 2, 3\}$ e matriz das probabilidades de transição dada por*

$$P^{\mathbf{X}} = \begin{bmatrix} 0.3 & 0.4 & 0.3 \\ 0.2 & 0 & 0.8 \\ 0.5 & 0.4 & 0.1 \end{bmatrix}.$$

Note, claramente, que para a partição $\{1\}, \{2, 3\}$ a condição de agregabilidade não é satisfeita, já que não existe $c \in \mathbb{R}$ tal que

$$W' = \begin{bmatrix} 0.2 \\ 0.5 \end{bmatrix} = c \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ isto é, } W' \neq c \cdot \mathbf{1}'.$$

O processo agregado obtido não é Markoviano de nenhuma ordem, sendo um processo de ordem infinita, isto é, processo de renovação.

Utilizando a abordagem matricial teremos

$$P^{\mathbf{X}} = \begin{bmatrix} p_{11} & V \\ W' & P \end{bmatrix}.$$

Observando $P^{\mathbf{X}}$, verificamos que a submatriz P é irredutível e ainda

$$P^2 = \begin{bmatrix} 0 & 0.8 \\ 0.4 & 0.1 \end{bmatrix}^2 = \begin{bmatrix} 0.32 & 0.08 \\ 0.04 & 0.33 \end{bmatrix} > 0, \quad (2.17)$$

verificamos no Apêndice B, pelo Teorema B.3, que P possui índice de primitividade 2.

Pelo Teorema 2.2

$$\lim_{k \rightarrow \infty} \left(\frac{P}{\lambda_1} \right)^k = G = \phi \cdot \psi'.$$

Os auto-valores de P são $\lambda_1 = 0.6178908$ e $\lambda_2 = -0.5178908$. Assim o raio espectral de P será $\lambda_1 = 0.6178908$. Correspondentemente a este auto-valor, obtemos os auto-vetores $\phi = \begin{bmatrix} 0.7914243 \\ 0.6112672 \end{bmatrix}$

de P e $\psi = \begin{bmatrix} 0.5434316 \\ 0.8394535 \end{bmatrix}$ de P' . Assim podemos calcular G

$$G = \phi \cdot \psi' = \begin{bmatrix} 0.4300849736 & 0.6643638986 \\ 0.3321819125 & 0.5131303905 \end{bmatrix}.$$

Obtemos, finalmente

$$\begin{aligned} \lim_{n \rightarrow \infty} P^{t(\mathbf{X})}(1|0^n 1) &= \frac{V \cdot G \cdot W'}{V \cdot G \cdot 1'} \\ &= \frac{\begin{bmatrix} 0.4 & 0.3 \end{bmatrix} \cdot \begin{bmatrix} 0.4300849736 & 0.6643638986 \\ 0.3321819125 & 0.5131303905 \end{bmatrix} \cdot \begin{bmatrix} 0.2 \\ 0.5 \end{bmatrix}}{\begin{bmatrix} 0.4 & 0.3 \end{bmatrix} \cdot \begin{bmatrix} 0.4300849736 & 0.6643638986 \\ 0.3321819125 & 0.5131303905 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}} \\ &= 0,382109164. \end{aligned}$$

Com respeito à Gibbsianidade, aplicando o Teorema 2.5,

$$\lim_{n,r \rightarrow \infty} \mathbb{P}(Y_0 = 1 | Y_{-n-1} = 1, Y_{-n}^{-1} = 0_{-n}^{-1}, Y_1^r = 0^r, Y_{r+1} = 1) = \frac{V \cdot G \cdot W'}{\lambda_1}$$

$$= \frac{\begin{bmatrix} 0.4 & 0.3 \end{bmatrix} \cdot \begin{bmatrix} 0.4300849736 & 0.6643638986 \\ 0.3321819125 & 0.5131303905 \end{bmatrix} \cdot \begin{bmatrix} 0.2 \\ 0.5 \end{bmatrix}}{0.6178908} = 0,427551358.$$

■

Antes de apresentar o exemplo seguinte, observamos que, pelo Teorema 2.2, um processo Markoviano pode gerar, pela Transformação Agregante (2.2), processos Markovianos de ordem 1 a $k = m - 1$, nos demais casos será de ordem infinita. O Exemplo 2.2 a seguir é uma aplicação do Teorema 2.2, Ítem 2. O Apêndice A será útil neste exemplo, o qual ilustra um processo Markoviano k -agregável.

Exemplo 2.2 *Seja $(X_n)_{n \in \mathbb{N}}$ um processo Markoviano não agregável, estacionário, com espaço de estados $A = \{1, 2, 3, 4, 5\}$ e matriz das probabilidade de transição decomposta na forma (2.4)*

$$P^{\mathbf{X}} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} & p_{15} \\ p_{21} & p_{22} & p_{23} & p_{24} & p_{25} \\ p_{31} & p_{32} & p_{33} & p_{34} & p_{35} \\ p_{41} & p_{42} & p_{43} & p_{44} & p_{45} \\ p_{51} & p_{52} & p_{53} & p_{54} & p_{55} \end{bmatrix} = \begin{bmatrix} p_{11} & V \\ W' & P \end{bmatrix}.$$

Suponha que a submatriz P seja primitiva com $\lambda_1 > \lambda_2 = \lambda_3 = \lambda_4 = 0$. Como os auto-valores de P são todos reais, P é **triangularizável**, e sua forma triangularizada apresenta os auto-valores na diagonal principal. Por (2.5)

$$P^{t(\mathbf{X})}(1|0^{n+1}1) = \begin{cases} p_{11}, & \text{se } n = -1 \\ \frac{V.P^n.W'}{V.P^n.\mathbf{1}'}, & \text{se } n \geq 0. \end{cases}$$

Portanto,

$$\begin{aligned} P^{t(\mathbf{X})}(1|0^{n+1}1) &= \frac{V.P^n.W'}{V.P^n.\mathbf{1}'} \\ &= \frac{V.(C.\Delta.C')^n.W'}{V.(C.\Delta.C')^n.\mathbf{1}'} \\ &= \frac{V.C.\Delta^n.C'.W'}{V.C.\Delta^n.C'.\mathbf{1}'} \end{aligned}$$

Note que Δ satisfaz às condições do Lema A.2. Vamos considerar os casos em que Δ possua uma, duas e três linhas não nulas.

1. Se Δ possui uma linha não nula,

$$P = C.\Delta.C', \text{ onde } \Delta = \begin{bmatrix} \lambda_1 & q_{23} & q_{24} & q_{25} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Pelo Lema A.2, para todo $n < \infty$, $\Delta^{n-1} = \lambda_1^{n-2}.\Delta$, então

$$\begin{aligned} P^{t(\mathbf{X})}(1|0^n 1) &= \frac{V.(C.\Delta^{n-1}.C').W'}{V.(C.\Delta^{n-1}.C').\mathbf{1}'} \\ &= \frac{V.(C.\lambda_1^{n-2}.\Delta.C').W'}{V.(C.\lambda_1^{n-2}.\Delta.C').\mathbf{1}'} \\ &= \frac{V.(C.\Delta.C').W'}{V.(C.\Delta.C').\mathbf{1}'} \\ &= \frac{V.P.W'}{V.P.\mathbf{1}'} \\ &= P^{t(\mathbf{X})}(1|0^2 1). \end{aligned}$$

No caso de $n \rightarrow \infty$

$$\begin{aligned} \lim_{n \rightarrow \infty} P^{t(\mathbf{X})}(1|0^n 1) &= \lim_{n \rightarrow \infty} \frac{V.P^{n-1}.W'}{V.P^{n-1}.\mathbf{1}'} \\ &= \lim_{n \rightarrow \infty} \frac{V.\left(\frac{P}{\lambda_1}\right)^{n-1}.W'}{V.\left(\frac{P}{\lambda_1}\right)^{n-1}.\mathbf{1}'} \\ &= \lim_{n \rightarrow \infty} \frac{V.\frac{C.\Delta^{n-1}.C'.W'}{\lambda_1^{n-1}}}{V.\frac{C.\Delta^{n-1}.C'.\mathbf{1}'}{\lambda_1^{n-1}}} \\ &= \lim_{n \rightarrow \infty} \frac{V.\frac{C.\lambda_1^{n-2}.\Delta.C'.W'}{\lambda_1^{n-1}}}{V.\frac{C.\lambda_1^{n-2}.\Delta.C'.\mathbf{1}'}{\lambda_1^{n-1}}} \\ &= \frac{V.\frac{C.\Delta.C'.W'}{\lambda_1}}{V.\frac{C.\Delta.C'.\mathbf{1}'}{\lambda_1}} \\ &= \frac{V.P.W'}{V.P.\mathbf{1}'} \\ &= P^{t(\mathbf{X})}(1|0^2 1). \end{aligned}$$

Logo $t(\mathbf{X})$ é Markoviano de ordem 2.

2. Se Δ possui duas linhas não nulas,

$$P = C.\Delta.C', \text{ onde } \Delta = \begin{bmatrix} \lambda_1 & q_{23} & q_{24} & q_{25} \\ 0 & 0 & q_{34} & q_{35} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

De modo semelhante, para todo $n < \infty$, $\Delta^{n-1} = \lambda_1^{n-3}.\Delta^2$, então

$$\begin{aligned} P^{t(\mathbf{X})}(1|0^n 1) &= \frac{V.(C.\Delta^{n-1}.C').W'}{V.(C.\Delta^{n-1}.C').\mathbf{1}'} \\ &= \frac{V.(C.\lambda_1^{n-3}.\Delta^2.C').W'}{V.(C.\lambda_1^{n-3}.\Delta^2.C').\mathbf{1}'} \\ &= \frac{V.(C.\Delta^2.C').W'}{V.(C.\Delta^2.C').\mathbf{1}'} \\ &= \frac{V.P^2.W'}{V.P^2.\mathbf{1}'} \\ &= P^{t(\mathbf{X})}(1|0^3 1). \end{aligned}$$

No caso de $n \rightarrow \infty$

$$\begin{aligned} \lim_{n \rightarrow \infty} P^{t(\mathbf{X})}(1|0^n 1) &= \lim_{n \rightarrow \infty} \frac{V.P^{n-1}.W'}{V.P^{n-1}.\mathbf{1}'} \\ &= \lim_{n \rightarrow \infty} \frac{V.\left(\frac{P}{\lambda_1}\right)^{n-1}.W'}{V.\left(\frac{P}{\lambda_1}\right)^{n-1}.\mathbf{1}'} \\ &= \lim_{n \rightarrow \infty} \frac{V.\frac{C.\Delta^{n-1}.C'.W'}{\lambda_1^{n-1}}}{V.\frac{C.\Delta^{n-1}.C'.\mathbf{1}'}{\lambda_1^{n-1}}} \\ &= \lim_{n \rightarrow \infty} \frac{V.\frac{C.\lambda_1^{n-3}.\Delta^2.C'.W'}{\lambda_1^{n-1}}}{V.\frac{C.\lambda_1^{n-3}.\Delta^2.C'.\mathbf{1}'}{\lambda_1^{n-1}}} \\ &= \frac{V.\frac{C.\Delta^2.C'.W'}{\lambda_1^2}}{V.\frac{C.\Delta^2.C'.\mathbf{1}'}{\lambda_1^2}} \\ &= \frac{V.P^2.W'}{V.P^2.\mathbf{1}'} \\ &= P^{t(\mathbf{X})}(1|0^3 1). \end{aligned}$$

Logo $t(\mathbf{X})$ é Markoviano de ordem 3.

3. Se Δ possui três linhas não nulas,

$$P = C.\Delta.C', \text{ onde } \Delta = \begin{bmatrix} \lambda_1 & q_{23} & q_{24} & q_{25} \\ 0 & 0 & q_{34} & q_{35} \\ 0 & 0 & 0 & q_{45} \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Analogamente aos casos anteriores, pelo Lema A.2, $\Delta^{n-1} = \lambda_1^{n-4}.\Delta^3$, então $t(\mathbf{X})$ é Markoviano de ordem 4.

Podemos notar que o número de linhas não nulas na forma triangularizada de P determina, de fato, a ordem do processo markoviano agregado. ■

O exemplo a seguir ilustra como a estrutura de uma matriz cíclica de índice $m - 1$, leva à não continuidade das probabilidades de transição do processo agregado $t(\mathbf{X})$. O Teorema B.6 presente no Apêndice B, apresenta a *Forma de Frobenius* de uma matriz cíclica, a qual será útil neste exemplo.

Exemplo 2.3 Seja $(X_n)_{n \in \mathbb{N}}$ um processo Markoviano não agregável, estacionário, com espaço de estados $A = \{1, 2, \dots, m\}$ e matriz das probabilidade de transição decomposta na forma (2.4)

$$P^{\mathbf{X}} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & \dots & p_{15} \\ p_{21} & p_{22} & p_{23} & \dots & p_{25} \\ p_{31} & p_{32} & p_{33} & \dots & p_{35} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{51} & p_{52} & p_{53} & \dots & p_{55} \end{bmatrix} = \begin{bmatrix} p_{11} & V \\ W' & P \end{bmatrix}.$$

Suponha que a submatriz P seja cíclica de índice $h = m - 1$, isto é, $(m - 1)$ -cíclica. O processo agregado $t(\mathbf{X})$ possui probabilidades de transição descontínuas.

Pelo Teorema B.6 podemos escrever:

$$P = C.\Delta.C', \text{ onde } \Delta = \begin{bmatrix} 0 & p_{23} & 0 & \dots & 0 \\ 0 & 0 & p_{34} & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & p_{m-1 m} \\ p_{m2} & 0 & 0 & \dots & 0 \end{bmatrix}.$$

Aplicando (2.5) para as probabilidades de transição do processo agregado $t(\mathbf{X})$ teremos

$$P^{t(\mathbf{X})}(1|0^{n+1}1) = \frac{V.P^n.W'}{V.P^n.1'} = \frac{V.(C.\Delta.C')^n.W'}{V.(C.\Delta.C')^n.1'} = \frac{V.C.\Delta^n.C'.W'}{V.C.\Delta^n.C'.1'}.$$

Observe também que

$$\Delta^{m-1} = p_{23} \cdot p_{34} \cdot \dots \cdot p_{m2} \cdot \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} = \prod_{i=2}^{m-1} (p_{i,i+1}) \cdot p_{m2} \cdot I,$$

note que $\prod_{i=2}^{m-1} (p_{i,i+1}) \cdot p_{m2}$ é uma constante não nula e I é a matriz identidade. Assim teremos $m-1$ valores alternados para as probabilidades de transição. Vamos mostrar que para todo $k \geq 0$ e $i = \{1, 2, 3, \dots, m-1\}$,

$$P^{t(\mathbf{X})}(1|0^{k(m-1)+i}1) = P(1|0^i1). \quad (2.18)$$

Vamos mostrar por indução em k .

1. (2.18) é válido para $k=0$. De fato,

$$P(1|0^{k \cdot (m-1)+i}1) = P(1|0^i1).$$

2. Suponha válido para k ,

$$P^{t(\mathbf{X})}(1|0^{k \cdot (m-1)+i}1) = P^{t(\mathbf{X})}(1|0^i1).$$

Então, para $k+1$, (2.18) é válido para $i=1$, pois

$$\begin{aligned} P^{t(\mathbf{X})}(1|0^{(k+1) \cdot (m-1)+1}1) &= \frac{V \cdot (C \cdot \Delta \cdot C')^{(k+1)(m-1)+0} \cdot W'}{V \cdot (C \cdot \Delta \cdot C')^{(k+1)(m-1)+0} \cdot \mathbf{1}'} \\ &= \frac{\overbrace{V \cdot (C \cdot \Delta \cdot C')^{m-1} \cdot (C \cdot \Delta \cdot C')^{m-1} \cdot \dots \cdot (C \cdot \Delta \cdot C')^{m-1}}^{k+1 \text{ vezes}} \cdot (C \cdot \Delta \cdot C')^0 \cdot W'}{V \cdot \underbrace{(C \cdot \Delta \cdot C')^{m-1} \cdot (C \cdot \Delta \cdot C')^{m-1} \cdot \dots \cdot (C \cdot \Delta \cdot C')^{m-1}}_{k+1 \text{ vezes}} \cdot (C \cdot \Delta \cdot C')^0 \cdot \mathbf{1}'} \\ &= \frac{V \cdot \left(\prod_{i=2}^{m-1} p_{i,i+1} \cdot p_{m2} \cdot I \right)^{k+1} \cdot I \cdot W'}{V \cdot \left(\prod_{i=2}^{m-1} p_{i,i+1} \cdot p_{m2} \cdot I \right)^{k+1} \cdot I \cdot \mathbf{1}'} \\ &= \frac{V \cdot I \cdot W'}{V \cdot I \cdot \mathbf{1}'} \\ &= P^{t(\mathbf{X})}(1|0^11). \end{aligned}$$

Supondo válido para $i = r$, teremos

$$\begin{aligned}
P^{t(\mathbf{X})}(1|0^{(k+1).(m-1)+r}1) &= \frac{V.(C.\Delta.C')^{(k+1)(m-1)+r-1}.W'}{V.(C.\Delta.C')^{(k+1)(m-1)+r-1}.\mathbf{1}'} \\
&= \frac{V.(C.\Delta.C')^{r-1}.W'}{V.(C.\Delta.C')^{r-1}.\mathbf{1}'} \\
&= P^{t(\mathbf{X})}(1|0^r1).
\end{aligned}$$

Então para $i = r + 1$, teremos

$$\begin{aligned}
P^{t(\mathbf{X})}(1|0^{(k+1).(m-1)+r+1}1) &= \frac{V.(C.\Delta.C')^{(k+1)(m-1)+r}.W'}{V.(C.\Delta.C')^{(k+1)(m-1)+r}.\mathbf{1}'} \\
&= \frac{\overbrace{V.(C.\Delta.C')^{m-1}.(C.\Delta.C')^{m-1}.\dots.(C.\Delta.C')^{m-1}}^{k+1 \text{ vezes}}.(C.\Delta.C')^r.W'}{V.\overbrace{(C.\Delta.C')^{m-1}.(C.\Delta.C')^{m-1}.\dots.(C.\Delta.C')^{m-1}}^{k+1 \text{ vezes}}.(C.\Delta.C')^r.\mathbf{1}'} \\
&= \frac{V.\left(\prod_{i=2}^{m-1} p_{i,i+1}.p_{m2}.I\right)^{k+1}.(C.\Delta.C')^r.W'}{V.\left(\prod_{i=2}^{m-1} p_{i,i+1}.p_{m2}.I\right)^{k+1}.(C.\Delta.C')^r.\mathbf{1}'} \\
&= \frac{V.(C.\Delta.C').(C.\Delta.C')^{r-1}.W'}{V.(C.\Delta.C').(C.\Delta.C')^{r-1}.\mathbf{1}'} \\
&= P^{t(\mathbf{X})}(1|0^{r+1}1).
\end{aligned}$$

Dessa forma, vê-se que há $m - 1$ valores alternados para as probabilidades de transição, logo não há convergência. ■

O Exemplo 2.4 é uma aplicação do Teorema 2.3. Ele ilustra o fato de que, quando há soma constante, k , nas colunas de P , a condição do Teorema 2.3 é satisfeita. Este exemplo é importante por mostrar, novamente, como a condição do Teorema 2.3 é verificada e mostrar que a probabilidade limite, neste caso, é dada por $1 - k$.

Exemplo 2.4 Seja $(X_n)_{n \in \mathbb{N}}$ um processo Markoviano não agregável, estacionário, com alfabeto $A = \{1, 2, \dots, m\}$ e matriz das probabilidades de transição decomposta na forma (2.4). Suponha a submatriz P cíclica de índice h , com soma constante k em suas colunas e cada bloco cíclico B_i com dimensão $\frac{m-1}{h}$. Vamos verificar que, neste caso, a Condição (2.9) é satisfeita.

Escreva P na forma de Frobenius (2.10) apresentada na prova do Teorema 2.3.

Do mesmo modo, P^h tem forma (2.12) e é primitiva com auto-valor maximal igual a $\lambda_1^* = k^h$, onde k é o auto-valor maximal dos blocos B_i , pois, por hipótese, temos soma constante k , nas colunas de P . Então, pelo Teorema B.4

$$\lim_{n \rightarrow \infty} \left(\frac{P}{k} \right)^{n.h} = G^*.$$

Os blocos B_i são matrizes primitivas e quadradas, então quando $n \rightarrow \infty$, G^* é formada pelos blocos limites $H_i^* = \phi_i \cdot \psi'_i$, na diagonal principal e zero nas outras entradas, onde ϕ_i e ψ_i são os auto-vetores à direita e à esquerda, respectivamente, do i -ésimo bloco da diagonal principal de P^h , correspondente à k^h . Operando em blocos, escrevemos V_i e W'_i , $i \in \{1, 2, \dots, h\}$, como subvetores, respectivamente, de V e W' , com a mesma dimensão de B_i . Então a expressão $\frac{V \cdot \left(\frac{P}{k}\right)^{n,h} \cdot P^r \cdot W'}{V \cdot \left(\frac{P}{k}\right)^{n,h} \cdot P^r \cdot \mathbf{1}'}$ será

$$\frac{V_1 \cdot \left(\frac{\prod_{i=1}^h B_i}{k^h}\right)^n \cdot (\prod_{i=1}^r B_i) \cdot W'_{j_1} + V_2 \cdot \left(\frac{\prod_{i=2}^h B_i \cdot B_1}{k^h}\right)^n \cdot (\prod_{i=2}^{r+1} B_i) \cdot W'_{j_2} + \dots + V_h \cdot \left(\frac{B_h \cdot \prod_{i=1}^{h-1} B_i}{k^h}\right)^n \cdot (B_h \cdot \prod_{i=1}^{r-1} B_i) \cdot W'_{j_h}}{V_1 \cdot \left(\frac{\prod_{i=1}^h B_i}{k^h}\right)^n \cdot (\prod_{i=1}^r B_i) + V_2 \cdot \left(\frac{\prod_{i=2}^h B_i \cdot B_1}{k^h}\right)^n \cdot (\prod_{i=2}^{r+1} B_i) + \dots + V_h \cdot \left(\frac{B_h \cdot \prod_{i=1}^{h-1} B_i}{k^h}\right)^n \cdot (B_h \cdot \prod_{i=1}^{r-1} B_i)}$$

Então, como afirmamos anteriormente, $\lim_{n \rightarrow \infty} \left(\frac{\prod_{i=1}^h B_i}{k^h}\right)^n = H_1^*$, $\lim_{n \rightarrow \infty} \left(\frac{\prod_{i=2}^h B_i \cdot B_1}{k^h}\right)^n = H_2^*$, ..., $\lim_{n \rightarrow \infty} \left(\frac{B_h \cdot \prod_{i=1}^{h-1} B_i}{k^h}\right)^n = H_h^*$. Logo

$$\begin{aligned} \lim_{n \rightarrow \infty} P^{t(\mathbf{X})}(1|0^{n,h+r+1}1) &= \frac{V_1 \cdot H_1^* \cdot (\prod_{i=1}^r B_i) \cdot W'_{j_1} + V_2 \cdot H_2^* \cdot (\prod_{i=2}^{r+1} B_i) \cdot W'_{j_2} + \dots + V_h \cdot H_h^* \cdot (B_h \cdot \prod_{i=1}^{r-1} B_i) \cdot W'_{j_h}}{V_1 \cdot H_1^* \cdot (\prod_{i=1}^r B_i) + V_2 \cdot H_2^* \cdot (\prod_{i=2}^{r+1} B_i) + \dots + V_h \cdot H_h^* \cdot (B_h \cdot \prod_{i=1}^{r-1} B_i)} \\ &= \frac{V_1 \cdot \phi_1 \cdot \psi'_1 \cdot (\prod_{i=1}^r B_i) \cdot W'_1 + V_2 \cdot \phi_2 \cdot \psi'_2 \cdot (\prod_{i=2}^{r+1} B_i) \cdot W'_{j_2} + \dots + V_h \cdot \phi_h \cdot \psi'_h \cdot (B_h \cdot \prod_{i=1}^{r-1} B_i) \cdot W'_h}{V_1 \cdot \phi_1 \cdot \psi'_1 \cdot (\prod_{i=1}^r B_i) + V_2 \cdot \phi_2 \cdot \psi'_2 \cdot (\prod_{i=2}^{r+1} B_i) + \dots + V_h \cdot \phi_h \cdot \psi'_h \cdot (B_h \cdot \prod_{i=1}^{r-1} B_i)} \end{aligned}$$

Note que ϕ_i e ψ_i são auto-vetores correspondentes à k^h . Como a soma nas colunas é constante, $\psi_i = \mathbf{1}'_i$ é um auto-vetor de B'_i , para todo $i \in \{1, 2, \dots, h\}$. Suponha, sem perda de generalidade que $i = 1$, assim teremos

$$\begin{aligned} \phi_1 \cdot \mathbf{1}_1 \cdot \prod_{i=1}^r B_i &= \phi_1 \cdot \mathbf{1}_1 \cdot B_1 \cdot \prod_{i=2}^r B_i \\ &= \phi_1 \cdot k \cdot \mathbf{1}_1 \cdot \prod_{i=2}^r B_i \\ &= k \cdot H_1^* \cdot \prod_{i=2}^r B_i, \end{aligned}$$

sucessivamente, chegamos a

$$= k^r \cdot H_1^*.$$

Para $i \neq 1$ o raciocínio é análogo. Substituindo acima, teremos

$$\begin{aligned} \lim_{n \rightarrow \infty} P^{t(\mathbf{X})}(1|0^{n,h+r+1}1) &= \frac{V_1 \cdot k^r \cdot H_1^* \cdot W'_{j_1} + V_2 \cdot k^r \cdot H_2^* \cdot W'_{j_2} + \dots + V_h \cdot k^r \cdot H_h^* \cdot W'_{j_h}}{V_1 \cdot k^r \cdot H_1^* + V_2 \cdot k^r \cdot H_2^* + \dots + V_h \cdot k^r \cdot H_h^*} \\ &= \frac{\sum_{i=1}^h V_i \cdot H_i^* \cdot W'_{j_i}}{\sum_{i=1}^h V_i \cdot H_i^*} \\ &= \frac{V \cdot G^* \cdot W'}{V \cdot G^* \cdot \mathbf{1}'} = C \end{aligned}$$

com C constante para todo $r \in \mathbb{N}$. Assim a Condição 2.9 é satisfeita e as probabilidades de transição do processo agregado são contínuas. ■

O Exemplo 2.5 a seguir ilustra o Teorema 2.4, mostrando que, quando a submatriz P é redutível ainda pode haver continuidade das probabilidades de transição.

Exemplo 2.5 Considere o processo Markoviano \mathbf{X} com espaço de estados $A = \{1, 2, 3, 4, 5, 6, 7\}$ e matriz das probabilidades de transição dada por

$$P^{\mathbf{X}} = \begin{bmatrix} 0.2 & 0.3 & 0.1 & 0.2 & 0.1 & 0 & 0.1 \\ 0.2 & 0.3 & 0.2 & 0.1 & 0 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.2 & 0 & 0.3 & 0.1 & 0.2 \\ 0 & 0 & 0 & 0.4 & 0.6 & 0 & 0 \\ 0.3 & 0 & 0 & 0.2 & 0.5 & 0 & 0 \\ 0.7 & 0 & 0 & 0 & 0 & 0.3 & 0 \\ 0.6 & 0 & 0 & 0 & 0 & 0 & 0.4 \end{bmatrix}.$$

Note que para a partição $\{1\}, \{2, 3, 4, 5, 6, 7\}$ a condição de agregabilidade não é satisfeita, o processo agregado obtido não é Markoviano de qualquer ordem, sendo um processo de ordem infinita.

Utilizando a abordagem matricial teremos

$$P^{\mathbf{X}} = \begin{bmatrix} p_{11} & V \\ W' & P \end{bmatrix}.$$

Observando $P^{\mathbf{X}}$, verificamos que a submatriz P é redutível e está escrita na forma canônica (B.3)

$$P = \begin{bmatrix} 0.3 & 0.2 & 0.1 & 0 & 0.1 & 0.1 \\ 0.1 & 0.2 & 0 & 0.3 & 0.1 & 0.2 \\ 0 & 0 & 0.4 & 0.6 & 0 & 0 \\ 0 & 0 & 0.2 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.4 \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} & T_{13} & T_{14} \\ \mathbf{0} & T_{22} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & T_{33} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & T_{44} \end{bmatrix}.$$

Como as submatrizes T_{22} , T_{33} , T_{44} são todas primitivas. Pelo Teorema 2.4 devemos obter os seus auto-valores, que são, respectivamente, $\lambda_{1,22} = 0.8$ e $\lambda_{1,33} = 0.3$, $\lambda_{1,44} = 0.4$. O bloco dominante é T_{22} e $\lambda_{1,max} = 0.8$. Assim, correspondentemente ao auto-valor maximal obtemos os auto-vetores

$$\phi_{max} = \begin{bmatrix} 0.8320503 \\ 0.5547002 \end{bmatrix} \text{ de } T_{22} \text{ e } \psi_{max} = \begin{bmatrix} 0.4472136 \\ 0.8944272 \end{bmatrix} \text{ de } T'_{22}. \text{ Assim}$$

$$\phi_{max} \cdot \psi'_{max} = \begin{bmatrix} 0.372104 & 0.744208 \\ 0.248069 & 0.496139 \end{bmatrix}$$

e calculando $\lim_{n \rightarrow \infty} \left(\frac{T_{22}}{\lambda_{1,max}} \right)^n$, teremos

$$\lim_{n \rightarrow \infty} \left(\frac{T_{22}}{\lambda_{1,max}} \right)^n = \begin{bmatrix} 0.372104 & 0.744208 & 0 & 0 \\ 0.248069 & 0.496139 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Obtendo $(I - T_{11})^{-1} \cdot T_{12} \cdot \lim_{n \rightarrow \infty} \left(\frac{T_{22}}{\lambda_{1,22}} \right)^n$ teremos

$$\begin{bmatrix} 0.082689741 & 0.165379593 & 0 & 0 \\ 0.103362093 & 0.206724574 & 0 & 0 \end{bmatrix}.$$

Podemos agora calcular \hat{G}

$$\hat{G} = \begin{bmatrix} 0 & 0 & 0.082698 & 0.165380 & 0 & 0 \\ 0 & 0 & 0.103362 & 0.206725 & 0 & 0 \\ 0 & 0 & 0.372104 & 0.744208 & 0 & 0 \\ 0 & 0 & 0.248069 & 0.496139 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Finalmente, pelo Teorema 2.4

$$\lim_{n \rightarrow \infty} P^{t(\mathbf{X})}(1|0^n1) = \frac{V \cdot \hat{G} \cdot W'}{V \cdot \hat{G} \cdot \mathbf{1}'} =$$

$$\begin{bmatrix} 0.3 & 0.1 & 0.2 & 0.1 & 0 & 0.1 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 0.082698 & 0.165380 & 0 & 0 \\ 0 & 0 & 0.103362 & 0.206725 & 0 & 0 \\ 0 & 0 & 0.372104 & 0.744208 & 0 & 0 \\ 0 & 0 & 0.248069 & 0.496139 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0.2 \\ 0.1 \\ 0 \\ 0.3 \\ 0.7 \\ 0.6 \end{bmatrix} = 0.2.$$

$$\begin{bmatrix} 0.3 & 0.1 & 0.2 & 0.1 & 0 & 0.1 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 0.082698 & 0.165380 & 0 & 0 \\ 0 & 0 & 0.103362 & 0.206725 & 0 & 0 \\ 0 & 0 & 0.372104 & 0.744208 & 0 & 0 \\ 0 & 0 & 0.248069 & 0.496139 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

■

Capítulo 3

Acoplamentos E Discrepância Entre Processos De Renovação

Neste capítulo abordamos acoplamentos e a discrepância entre processos estocásticos. Utilizamos a noção de processos agregados para obter resultados relacionados a este tema e construímos exemplos de acoplamentos entre processos binários, que alcançam a mínima discrepância entre os processos acoplados. Todos os processos mencionado neste capítulo são estacionários.

Iniciamos nossa abordagem com algumas definições iniciais que foram extraídas e adaptadas a partir de [Fernández *et al.* \(2001\)](#) e [Ellis \(1980b\)](#).

3.1 Acoplamento De Renovação

Definição 3.1 *Sejam dois processos estocásticos \mathbf{X} e \mathbf{Y} com alfabeto $A = \{0, 1\}$ e distribuições estacionárias $\mu_{\mathbf{X}}$ e $\mu_{\mathbf{Y}}$, respectivamente. Um acoplamento entre \mathbf{X} e \mathbf{Y} , denotado por $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})$, é um processo com alfabeto $A \times A = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$, cujas marginais são $\mu_{\mathbf{X}}$ e $\mu_{\mathbf{Y}}$.*

Acoplar dois processos estocásticos significa construí-los simultaneamente utilizando o mesmo procedimento aleatório. Se \mathbf{X} e \mathbf{Y} são dois processos estocásticos estacionários, com mesmo alfabeto finito A , ao acoplarmos \mathbf{X} e \mathbf{Y} , obtemos um processo \mathbf{Z} com alfabeto $A \times A$, tal que,

$$\sum_{a \in A} \mu_{\mathbf{Z}}(a, b) = \mu_{\mathbf{Y}}(b) \quad (3.1)$$

e

$$\sum_{b \in A} \mu_{\mathbf{Z}}(a, b) = \mu_{\mathbf{X}}(a) \quad (3.2)$$

para todo $a, b \in A$.

Acoplamento Markoviano e acoplamento de renovação são definidos a seguir.

Definição 3.2 *Sejam dois processos estocásticos \mathbf{X} e \mathbf{Y} com alfabeto $A = \{0, 1\}$. O processo $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})$ com alfabeto $A \times A = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$ é um acoplamento Markoviano, se \mathbf{Z} é um processo markoviano e é um acoplamento entre \mathbf{X} e \mathbf{Y} de acordo com a Definição 3.1.*

Definição 3.3 *Sejam dois processos de renovação \mathbf{X} e \mathbf{Y} com alfabeto $A = \{0, 1\}$. O processo $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})$ com alfabeto $A \times A = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$ é um acoplamento de renovação, se é um processo de renovação de acordo com a Definição 2.1, e é um acoplamento entre \mathbf{X} e \mathbf{Y} de acordo com a Definição 3.1.*

3.2 Discrepância Entre Processos De Renovação

Considere dois processos de renovação \mathbf{X} e \mathbf{Y} com medidas estacionárias $\mu_{\mathbf{X}}$ e $\mu_{\mathbf{Y}}$, respectivamente. Seja $C_{\mathbf{X}\mathbf{Y}}$ o conjunto dos acoplamentos entre \mathbf{X} e \mathbf{Y} .

Definição 3.4 *Dados dois processos \mathbf{X} e \mathbf{Y} , seja $C_{\mathbf{X}\mathbf{Y}}$ o conjunto dos acoplamentos entre \mathbf{X} e \mathbf{Y} . Defina a função*

$$d : C_{\mathbf{X}\mathbf{Y}} \rightarrow [0, 1], \text{ dada por } d(\mathbf{Z}) = \sum_{a \neq b} \mu_{\mathbf{Z}}(a, b), \text{ para todo } a, b \in A, \mathbf{Z} \in C_{\mathbf{X}\mathbf{Y}}.$$

Neste caso $d(\mathbf{Z})$ é a medida dos pares de símbolos do processo acoplado cujas coordenadas de \mathbf{X} e \mathbf{Y} são diferentes. $d(\mathbf{Z})$ será chamada de discrepância entre \mathbf{X} e \mathbf{Y} alcançada pelo acoplamento \mathbf{Z} .

Definição 3.5 *A distância \bar{d} entre \mathbf{X} e \mathbf{Y} , denotada por $\bar{d}(\mathbf{X}, \mathbf{Y})$, é*

$$\bar{d}(\mathbf{X}, \mathbf{Y}) = \inf\{d(\mathbf{Z}) : \mathbf{Z} \in C_{\mathbf{X}\mathbf{Y}}\}. \quad (3.3)$$

Definição 3.6 *Sejam \mathbf{X} e \mathbf{Y} processos Markovianos. Considere $M_{\mathbf{X}, \mathbf{Y}}$ o conjunto dos acoplamentos Markovianos entre \mathbf{X} e \mathbf{Y} . A distância Markoviana entre \mathbf{X} e \mathbf{Y} é definida por:*

$$d_M(\mathbf{X}, \mathbf{Y}) = \inf\{d(\mathbf{Z}) : \mathbf{Z} \in M_{\mathbf{X}, \mathbf{Y}}\}. \quad (3.4)$$

Definição 3.7 *A distância de partição entre os processos aleatórios \mathbf{X} e \mathbf{Y} , é definida por*

$$d_p(\mathbf{X}, \mathbf{Y}) = 1 - \sum_{a \in A} \min\{\mu_{\mathbf{X}}(a), \mu_{\mathbf{Y}}(a)\}, \quad (3.5)$$

onde $\mu_{\mathbf{X}}$ e $\mu_{\mathbf{Y}}$ são as medidas estacionárias de \mathbf{X} e \mathbf{Y} , respectivamente.

No caso binário, a distância de partição é dada por

$$\begin{aligned} d_p(\mathbf{X}, \mathbf{Y}) &= 1 - \sum_{a \in A} \min\{\mu_{\mathbf{X}}(a), \mu_{\mathbf{Y}}(a)\} \\ &= 1 - \min\{\mu_{\mathbf{X}}(0), \mu_{\mathbf{Y}}(0)\} - \min\{\mu_{\mathbf{X}}(1), \mu_{\mathbf{Y}}(1)\} \\ &= \max\{\mu_{\mathbf{X}}(1), \mu_{\mathbf{Y}}(1)\} - \min\{\mu_{\mathbf{X}}(1), \mu_{\mathbf{Y}}(1)\} \\ &= |\mu_{\mathbf{X}}(1) - \mu_{\mathbf{Y}}(1)|. \end{aligned}$$

NOTA: Entre dois processos estacionários a distância de partição representa a diferença entre as medidas estacionárias de cada símbolo nos processos. Note que essa medida não tem nenhuma relação com os acoplamentos, leva em consideração apenas as medidas marginais dos dois processos considerados. Se existir um acoplamento que alcança a distância de partição, então também alcança a distância \bar{d} . De fato, suponha $\mu_{\mathbf{X}}(1) \leq \mu_{\mathbf{Y}}(1)$ (o caso $\mu_{\mathbf{X}}(1) \geq \mu_{\mathbf{Y}}(1)$ é análogo), então para qualquer acoplamento $Z \in C_{XY}$, temos

$$\begin{aligned}
d_p(\mathbf{X}, \mathbf{Y}) &= 1 - \sum_{i \in A} \min\{\mu_{\mathbf{X}}(i), \mu_{\mathbf{Y}}(i)\} \\
&= 1 - \min\{\mu_{\mathbf{X}}(1), \mu_{\mathbf{Y}}(1)\} - \sum_{i \in A \setminus \{1\}} \min\{\mu_{\mathbf{X}}(i), \mu_{\mathbf{Y}}(i)\} \\
&= \sum_{i=2}^m \mu_{\mathbf{X}}(i) - \sum_{i=2}^m \min\{\mu_{\mathbf{X}}(i), \mu_{\mathbf{Y}}(i)\} \\
&= \sum_{i=2}^m (\mu_{\mathbf{X}}(i) - \min\{\mu_{\mathbf{X}}(i), \mu_{\mathbf{Y}}(i)\}) \leq \sum_{i=2}^m |\mu_{\mathbf{X}}(i) - \mu_{\mathbf{Y}}(i)| \\
&= \sum_{i=2}^m \left| \sum_{k=1}^m (\mu_{\mathbf{Z}}(i, k) - \mu_{\mathbf{Z}}(k, i)) \right| \\
&= \sum_{i=2}^m \left| \sum_{k \neq i} (\mu_{\mathbf{Z}}(i, k) - \mu_{\mathbf{Z}}(k, i)) \right| \leq \sum_{i=1}^m \sum_{k \neq i} (\mu_{\mathbf{Z}}(i, k) + \mu_{\mathbf{Z}}(k, i)) \\
&= d(\mathbf{Z}).
\end{aligned}$$

Portanto, se um acoplamento alcança a distância de partição, alcança também distância \bar{d} , porém é possível alcançar a distância \bar{d} sem alcançar a distância de partição, como ilustrado por Ellis (1976) (pág.439). Devido à facilidade no cálculo da distância de partição, utilizaremos em muitos casos a comparação com ela em nossas análises.

Teorema 3.1 *Sejam dois processos \mathbf{X} e \mathbf{Y} com alfabeto $\{0, 1\}$. Se o acoplamento \mathbf{Z} entre \mathbf{X} e \mathbf{Y} alcança a distância \bar{d} , então $\mu_{\mathbf{Z}}(1, 1) > 0$ e $\mu_{\mathbf{Z}}(0, 0) > 0$.*

Prova do Teorema 3.1

Suponha que $\mu_{\mathbf{Z}}(1, 1) = 0$. A discrepância entre \mathbf{X} e \mathbf{Y} alcançada por \mathbf{Z} é dada por

$$d(\mathbf{Z}) = \mu_{\mathbf{Z}}(1, 0) + \mu_{\mathbf{Z}}(0, 1).$$

Pela definição de acoplamentos

$$\mu_{\mathbf{X}}(1) = \mu_{\mathbf{Z}}(1, 0) + \mu_{\mathbf{Z}}(1, 1) = \mu_{\mathbf{Z}}(1, 0)$$

e

$$\mu_{\mathbf{Y}}(1) = \mu_{\mathbf{Z}}(0, 1) + \mu_{\mathbf{Z}}(1, 1) = \mu_{\mathbf{Z}}(0, 1).$$

Portanto,

$$d(\mathbf{Z}) = \mu_{\mathbf{X}}(1) + \mu_{\mathbf{Y}}(1).$$

Tome outro acoplamento \mathbf{W} entre \mathbf{X} e \mathbf{Y} com $\mu_{\mathbf{W}}(1, 1) > 0$. A discrepância alcançada por esse acoplamento é dada por

$$d(\mathbf{W}) = \mu_{\mathbf{W}}(1, 0) + \mu_{\mathbf{W}}(0, 1),$$

pela definição de acoplamentos

$$\mu_{\mathbf{X}}(1) = \mu_{\mathbf{W}}(1, 0) + \mu_{\mathbf{W}}(1, 1)$$

e

$$\mu_{\mathbf{Y}}(1) = \mu_{\mathbf{W}}(0, 1) + \mu_{\mathbf{W}}(1, 1).$$

Então

$$d(\mathbf{Z}) = \mu_{\mathbf{X}}(1) + \mu_{\mathbf{Y}}(1) = \mu_{\mathbf{W}}(1, 0) + \mu_{\mathbf{W}}(0, 1) + 2\mu_{\mathbf{W}}(1, 1) > d(\mathbf{W}),$$

pois $\mu_{\mathbf{W}}(1, 1) > 0$.

O argumento para $\mu(0, 0)$ é análogo. ■

3.3 Transformação Agregante E Distância De Partição

Sejam \mathbf{X} e \mathbf{Y} , processos Markovianos com alfabeto finito $A = \{1, 2, 3, \dots, m\}$ e medidas estacionárias $\mu_{\mathbf{X}}$ e $\mu_{\mathbf{Y}}$, respectivamente. Suponha que $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})$ seja um acoplamento entre \mathbf{X} e \mathbf{Y} com alfabeto $A \times A$ e medida dada por $\mu_{\mathbf{Z}}$. Considere uma partição $A = A_1 \cup A_2 \cup \dots \cup A_k$, tal que $A_i \cap A_j = \emptyset$, para algum A_i unitário, com $i \leq k$.

Seja $B = \{1, 2, 3, \dots, k\}$ com $k < m$, defina a **Transformação Agregante** $t : A \rightarrow B$ por

$$t(a) = \sum_{i \in B} i \cdot \mathbb{I}(a \in A_i). \quad (3.6)$$

Defina a **Transformação Agregante Estendida** $T : A \times A \rightarrow B \times B$ por

$$T(a, a') = \sum_{(i, j) \in B^2} (i, j) \cdot \mathbb{I}(a \in A_i) \cdot \mathbb{I}(a' \in A_j). \quad (3.7)$$

Defina σ e σ' por

$$\sigma(\mu_{\mathbf{X}}, \mu_{\mathbf{Y}}) = \sum_i \min \left\{ \sum_{a \in A_i} \mu_{\mathbf{X}}(a), \sum_{a \in A_i} \mu_{\mathbf{Y}}(a) \right\} - \sum_i \sum_{a \in A_i} \min \{ \mu_{\mathbf{X}}(a), \mu_{\mathbf{Y}}(a) \} \quad (3.8)$$

e

$$\sigma'(\mu_{\mathbf{Z}}) = \sum_{a \neq b, a, b \in A} \mu_{\mathbf{Z}}(a, b) - \sum_{i, j} \sum_{a \in A_i, b \in A_j, i \neq j} \mu_{\mathbf{Z}}(a, b). \quad (3.9)$$

Lema 3.1 *Sejam \mathbf{X} e \mathbf{Y} processos Markovianos. Suponha que exista um acoplamento $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})$ que alcança a distância de partição entre \mathbf{X} e \mathbf{Y} . Se $\sigma(\mu_{\mathbf{X}}, \mu_{\mathbf{Y}}) = \sigma'(\mu_{\mathbf{Z}})$, então o acoplamento $T(\mathbf{Z})$ alcança a distância de partição entre $t(\mathbf{X})$ e $t(\mathbf{Y})$.*

Prova do Lema 3.1

A distância alcançada pelo acoplamento \mathbf{Z} será dada por $d(\mathbf{Z}) = \sum_{a \neq b} \mu_{\mathbf{Z}}(a, b)$. Como alcança a distância de partição entre \mathbf{X} e \mathbf{Y} , pela Definição 3.7,

$$\begin{aligned} d(\mathbf{Z}) &= \sum_{a, b \in A, a \neq b} \mu_{\mathbf{Z}}(a, b) \\ &= d_p(\mathbf{X}, \mathbf{Y}) \\ &= 1 - \sum_{a \in A} \min\{\mu_{\mathbf{X}}(a), \mu_{\mathbf{Y}}(a)\} \\ &= 1 - \sum_i \sum_{a \in A_i} \min\{\mu_{\mathbf{X}}(a), \mu_{\mathbf{Y}}(a)\}. \end{aligned}$$

A distância alcançada pelo acoplamento $T(\mathbf{Z})$ será dada por

$$\begin{aligned} T(\mathbf{Z}) &= \sum_{i, j \in B, i \neq j} \mu_{T(\mathbf{Z})}(i, j) \\ &= \sum_{i, j \in B} \sum_{a \in A_i, a' \in A_j, i \neq j} \mu_{\mathbf{Z}}(a, a') \leq \sum_{a \neq a'} \mu_{\mathbf{Z}}(a, a') \\ &= d(\mathbf{Z}). \end{aligned}$$

Vamos obter agora a distância de partição entre $t(\mathbf{X})$ e $t(\mathbf{Y})$,

$$\begin{aligned} d_p(t(\mathbf{X}), t(\mathbf{Y})) &= 1 - \sum_{i \in B} \min\{\mu_{t(\mathbf{X})}(i), \mu_{t(\mathbf{Y})}(i)\} \\ &= 1 - \sum_{i \in B} \min \left\{ \sum_{a \in A_i} \mu_{\mathbf{X}}(a), \sum_{a \in A_i} \mu_{\mathbf{Y}}(a) \right\} \leq 1 - \sum_{i \in B} \sum_{a \in A_i} \min\{\mu_{\mathbf{X}}(a), \mu_{\mathbf{Y}}(a)\} \\ &= d_p(\mathbf{X}, \mathbf{Y}). \end{aligned}$$

Como $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})$ alcança a distância de partição entre \mathbf{X} e \mathbf{Y} e usando a hipótese, $\sigma(\mu_{\mathbf{X}}, \mu_{\mathbf{Y}}) = \sigma'(\mu_{\mathbf{Z}})$, teremos

$$0 = d_p(\mathbf{X}, \mathbf{Y}) - d(\mathbf{Z}) = 1 - \sum_i \sum_{a \in A_i} \min\{\mu_{\mathbf{X}}(a), \mu_{\mathbf{Y}}(a)\} - \sum_{a, b \in A, a \neq b} \mu_{\mathbf{Z}}(a, b)$$

$$\sigma(\mu_{\mathbf{X}}, \mu_{\mathbf{Y}}) - \sigma'(\mu_{\mathbf{Z}}) = \sum_i \min \left\{ \sum_{a \in A_i} \mu_{\mathbf{X}}(a), \sum_{a \in A_i} \mu_{\mathbf{Y}}(a) \right\} - \sum_i \sum_{a \in A_i} \min\{\mu_{\mathbf{X}}(a), \mu_{\mathbf{Y}}(a)\}$$

$$- \left(\sum_{a,b \in A, a \neq b} \mu_{\mathbf{Z}}(a,b) - \sum_{i,j} \sum_{a \in A_i, b \in A_j, i \neq j} \mu_{\mathbf{Z}}(a,b) \right).$$

Rearranjando os termos, teremos

$$1 - \sum_i \min \left\{ \sum_{a \in A_i} \mu_{\mathbf{X}}(a), \sum_{a \in A_i} \mu_{\mathbf{Y}}(a) \right\} = \sum_{i,j} \sum_{a \in A_i, b \in A_j, i \neq j} \mu_{\mathbf{Z}}(a,b)$$

logo $d_p(t(\mathbf{X}), t(\mathbf{Y})) = d(T(\mathbf{Z}))$. ■

Considere a seguinte condição de ordenação sobre a partição de A

Condição 3.1 Para cada subconjunto A_i da partição de A , $\mu_{\mathbf{X}}(a) \leq \mu_{\mathbf{Y}}(a)$, para todo $a \in A_i$.

Teorema 3.2 Sejam as transformações sobrejetivas t e T , definidas em (3.6) e (3.7), respectivamente, cuja partição satisfaz a Condição 3.1. Se o acoplamento $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})$ alcança a distância de partição entre \mathbf{X} e \mathbf{Y} , então o acoplamento $T(\mathbf{Z})$ alcança a distância de partição entre $t(\mathbf{X})$ e $t(\mathbf{Y})$.

Prova do Teorema 3.2

Pela Condição 3.1 para cada $i \in \mathbb{N}$, sabemos que $\mu_{\mathbf{X}}(a) \leq \mu_{\mathbf{Y}}(a)$ para todo $a \in A_i$, consequentemente $\sum_{a \in A_i} \mu_{\mathbf{X}}(a) \leq \sum_{a \in A_i} \mu_{\mathbf{Y}}(a)$ para todo $i \in \mathbb{N}$. Neste caso,

$$\begin{aligned} d_p(t(\mathbf{X}), t(\mathbf{Y})) &= 1 - \sum_{i \in B} \min \left\{ \sum_{a \in A_i} \mu_{\mathbf{X}}(a), \sum_{a \in A_i} \mu_{\mathbf{Y}}(a) \right\} \\ &= 1 - \sum_{A_i} \sum_{a \in A_i} \min\{\mu_{\mathbf{X}}(a), \mu_{\mathbf{Y}}(a)\} \\ &= d_p(\mathbf{X}, \mathbf{Y}). \end{aligned}$$

Como $d(T(\mathbf{Z})) \leq d(\mathbf{Z}) = d_p(\mathbf{X}, \mathbf{Y})$ então o acoplamento $T(\mathbf{Z})$ alcança a distância de partição entre $t(\mathbf{X})$ e $t(\mathbf{Y})$. ■

3.4 Transformação Agregante E Distância \bar{d}

Proposição 3.1 Sejam \mathbf{X} e \mathbf{Y} , processos Markovianos com alfabeto finito $A = \{1, 2, 3, \dots, m\}$. Considere a Transformação Agregante dada por (2.2) e T , Transformação Agregante Estendida definida por (3.7). Se o acoplamento $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})$ alcança a distância \bar{d} entre \mathbf{X} e \mathbf{Y} , e

$$\sum_{a,b \in A \setminus \{1\}, a \neq b} \mu_{\mathbf{Z}}(a,b) \geq \sum_{a,b \in A \setminus \{1\}, a \neq b} \mu_{\delta}(a,b) \quad (3.10)$$

para qualquer $\delta \in C_{\mathbf{X}\mathbf{Y}}$, então o acoplamento $T(\mathbf{Z})$ alcança a distância \bar{d} entre $t(\mathbf{X})$ e $t(\mathbf{Y})$.

Prova da Proposição 3.1

A distância alcançada pelo acoplamento \mathbf{Z} é

$$\begin{aligned} d(\mathbf{Z}) &= \sum_{a,b \in A, a \neq b} \mu_{\mathbf{Z}}(a, b) \\ &= \sum_{a \in A \setminus \{1\}} \mu_{\mathbf{Z}}(a, 1) + \sum_{b \in A \setminus \{1\}} \mu_{\mathbf{Z}}(1, b) + \sum_{a,b \in A \setminus \{1\}, a \neq b} \mu_{\mathbf{Z}}(a, b) = \bar{d}(\mathbf{X}, \mathbf{Y}). \end{aligned}$$

Note que

$$\sum_{a \in A} \mu_{\mathbf{Z}}(a, b) = \mu_{\mathbf{Y}}(b)$$

e

$$\sum_{b \in A} \mu_{\mathbf{Z}}(a, b) = \mu_{\mathbf{X}}(a)$$

para todos $a, b \in A$.

Considerando a Transformação T definida por (3.7), a distância alcançada por este acoplamento será dada por

$$\begin{aligned} d(T(\mathbf{Z})) &= \mu_{T(\mathbf{Z})}(0, 1) + \mu_{T(\mathbf{Z})}(1, 0) \\ &= \sum_{a \in A \setminus \{1\}} \mu_{\mathbf{Z}}(a, 1) + \sum_{b \in A \setminus \{1\}} \mu_{\mathbf{Z}}(1, b). \end{aligned}$$

Suponha, por contradição, que $T(\mathbf{Z})$ não alcança a distância $\bar{d}(t(\mathbf{X}), t(\mathbf{Y}))$, então existe τ , acoplamento entre $t(\mathbf{X})$ e $t(\mathbf{Y})$, tal que

$$d(T(\mathbf{Z})) > d(\tau).$$

Assim

$$\mu_{T(\mathbf{Z})}(0, 1) + \mu_{T(\mathbf{Z})}(1, 0) > \mu_{\tau}(0, 1) + \mu_{\tau}(1, 0).$$

Portanto, existe $\delta \in C_{\mathbf{X}\mathbf{Y}}$ tal que

$$\begin{aligned} &\sum_{a \in A \setminus \{1\}} \mu_{\mathbf{Z}}(a, 1) + \sum_{b \in A \setminus \{1\}} \mu_{\mathbf{Z}}(1, b) + 2 \cdot \mu_{\mathbf{Z}}(1, 1) - 2 \cdot \mu_{\mathbf{Z}}(1, 1) \\ &> \sum_{a \in A \setminus \{1\}} \mu_{\delta}(a, 1) + \sum_{b \in A \setminus \{1\}} \mu_{\delta}(1, b) + 2 \cdot \mu_{\delta}(1, 1) - 2 \cdot \mu_{\delta}(1, 1). \end{aligned}$$

Neste caso,

$$\mu_{\mathbf{X}}(1) + \mu_{\mathbf{Y}}(1) - 2 \cdot \mu_{\mathbf{Z}}(1, 1) > \mu_{\mathbf{X}}(1) + \mu_{\mathbf{Y}}(1) - 2 \cdot \mu_{\delta}(1, 1)$$

e conseqüentemente

$$\mu_{\mathbf{Z}}(1, 1) < \mu_{\delta}(1, 1), \tag{3.11}$$

onde δ é obtido pela imagem inversa de τ pela Transformação Agregante T . Analogamente

$$\mu_{T(\mathbf{Z})}(0, 1) + \mu_{T(\mathbf{Z})}(1, 0) > \mu_\tau(0, 1) + \mu_\tau(1, 0)$$

implica que

$$\begin{aligned} & \mu_{T(\mathbf{Z})}(0, 1) + \mu_{T(\mathbf{Z})}(1, 0) + 2 \cdot \mu_{T(\mathbf{Z})}(0, 0) - 2 \cdot \mu_{T(\mathbf{Z})}(0, 0) \\ & > \mu_\tau(0, 1) + \mu_\tau(1, 0) + 2 \cdot \mu_\tau(0, 0) - 2 \cdot \mu_\tau(0, 0). \end{aligned}$$

Assim,

$$\sum_{a \in A \setminus \{1\}} \mu_{\mathbf{X}}(a) + \sum_{b \in A \setminus \{1\}} \mu_{\mathbf{Y}}(b) - 2 \cdot \mu_{T(\mathbf{Z})}(0, 0) > \sum_{a \in A \setminus \{1\}} \mu_{\mathbf{X}}(a) + \sum_{b \in A \setminus \{1\}} \mu_{\mathbf{Y}}(b) - 2 \cdot \mu_\tau(0, 0)$$

e

$$\mu_{T(\mathbf{Z})}(0, 0) < \mu_\tau(0, 0).$$

Consequentemente teremos

$$\sum_{a \in A \setminus \{1\}} \mu_{\mathbf{Z}}(a, a) + \sum_{a, b \in A \setminus \{1\}, a \neq b} \mu_{\mathbf{Z}}(a, b) < \sum_{a \in A \setminus \{1\}} \mu_\delta(a, a) + \sum_{a, b \in A \setminus \{1\}, a \neq b} \mu_\delta(a, b). \quad (3.12)$$

Somando (3.11) com (3.12), teremos

$$\mu_{\mathbf{Z}}(1, 1) + \sum_{a \in A \setminus \{1\}} \mu_{\mathbf{Z}}(a, a) + \sum_{a, b \in A \setminus \{1\}, a \neq b} \mu_{\mathbf{Z}}(a, b) < \mu_\delta(1, 1) + \sum_{a \in A \setminus \{1\}} \mu_\delta(a, a) + \sum_{a, b \in A \setminus \{1\}, a \neq b} \mu_\delta(a, b),$$

por (3.10), teremos, finalmente

$$\mu_{\mathbf{Z}}(1, 1) + \sum_{a \in A \setminus \{1\}} \mu_{\mathbf{Z}}(a, a) < \mu_\delta(1, 1) + \sum_{a \in A \setminus \{1\}} \mu_\delta(a, a)$$

e $d(\mathbf{Z}) > d(\delta)$. Contradição. ■

Antes de desenvolvermos a seção seguinte, apresentamos a Proposição 3.2 que nos dá a relação existente entre a distância Markoviana e a distância entre processos de renovação agregados, nos moldes construtivos que estamos trabalhando.

Definição 3.8 *Sejam $t(\mathbf{X})$ e $t(\mathbf{Y})$, processos de renovação gerados a partir de \mathbf{X} e \mathbf{Y} , respectivamente, pela Transformação Agregante (2.2). Seja \mathbf{Z} um acoplamento Markoviano entre \mathbf{X} e \mathbf{Y} . Seja $T(\mathbf{Z})$ o processo de renovação gerado a partir de \mathbf{Z} pela Transformação Agregante Estendida (3.7). A distância entre processos de renovação agregados é definida por*

$$d_r(t(\mathbf{X}), t(\mathbf{Y})) = \inf\{d(T(\mathbf{Z})) : \mathbf{Z} \in M(\mathbf{X}, \mathbf{Y})\}.$$

Vamos chamar a distância entre processos de renovação agregados de *Distância de Renovação*, com a notação acima.

Proposição 3.2

$$d_r(t(\mathbf{X}), t(\mathbf{Y})) \leq d_M(\mathbf{X}, \mathbf{Y}),$$

onde d_M é a distância Markoviana definida em 3.6.

Prova da Proposição 3.2

Sejam os processos Markovianos, \mathbf{X} e \mathbf{Y} com alfabeto $A = \{1, 2, \dots, m\}$ e sejam $t(\mathbf{X})$ e $t(\mathbf{Y})$ os processos de renovação com alfabeto $B = \{0, 1\}$ obtidos pela Transformação (2.2).

Considere \mathbf{Z} o acoplamento Markoviano entre \mathbf{X} e \mathbf{Y} que alcança a distância Markoviana ($d_M(\mathbf{X}, \mathbf{Y})$), então

$$\begin{aligned} d_M(\mathbf{X}, \mathbf{Y}) &= \sum_{i,j \in A, i \neq j} \mu_{\mathbf{Z}}(i, j) = \mu_{\mathbf{Z}}(1, 2) + \mu_{\mathbf{Z}}(1, 3) + \dots + \mu_{\mathbf{Z}}(1, m) + \\ &\mu_{\mathbf{Z}}(2, 3) + \mu_{\mathbf{Z}}(2, 4) \dots + \mu_{\mathbf{Z}}(2, m) + \dots + \mu_{\mathbf{Z}}(m-1, m). \end{aligned}$$

Note que os pares de símbolos discrepantes $(i, j), i \neq j$ com $i, j \neq 1$, deixam de ser discrepantes quando aplicamos a Transformação (2.2), pois $t(i) = t(j) = 0$, sempre que $i, j \neq 1$. Por esse argumento,

$$d_M(\mathbf{X}, \mathbf{Y}) \geq \sum_{i,j \in B, i \neq j} \mu_{t(\mathbf{Z})}(i, j) = d_r(t(\mathbf{X}), t(\mathbf{Y})).$$

■

Podemos observar que a distância de renovação só se iguala à distância Markoviana quando tivermos $\mu_{\mathbf{Z}}(i, j) = 0$, sempre que $i \neq j$ com $i, j \neq 1$, caso contrário $d_r(t(\mathbf{X}), t(\mathbf{Y})) < d_M(\mathbf{X}, \mathbf{Y})$. Note então, que é possível que o acoplamento Markoviano não alcance a distância \bar{d} , enquanto que o acoplamento de renovação a alcança, isso é ilustrado pelo Exemplo 3.2.

3.5 Condição De Ordenação E Acoplamentos De Renovação

Galves *et al.* (2010) construíram um acoplamento estacionário alcançando a distância \bar{d} entre pares de processos binários de ordem infinita ordenados no seguinte sentido,

$$x_{-\infty}^{-1} \leq y_{-\infty}^{-1} \Rightarrow P^{\mathbf{X}}(1|x_{-\infty}^{-1}) \leq P^{\mathbf{Y}}(1|y_{-\infty}^{-1}). \quad (3.13)$$

O Exemplo 3.1 ilustra o fato de que esta condição de ordenação, não é necessária para que um acoplamento entre processos de renovação alcance a distância \bar{d} .

Exemplo 3.1 *Sejam dois processos Markovianos \mathbf{U} e \mathbf{V} , com alfabeto $B = \{1, 2, 3\}$ e probabilidades*

de transição dadas, respectivamente, pelas matrizes

$$P^U = \begin{bmatrix} \beta & 1-\beta & 0 \\ \alpha & 0 & 1-\alpha \\ \beta & 1-\beta & 0 \end{bmatrix}$$

e

$$P^V = \begin{bmatrix} \beta & 1-\beta & 0 \\ \alpha' & 0 & 1-\alpha' \\ \beta & 1-\beta & 0 \end{bmatrix}$$

com $\alpha' \leq \beta \leq \alpha$.

Tomemos os processos de renovação $\mathbf{X} = t(\mathbf{U})$ e $\mathbf{Y} = t(\mathbf{V})$ com alfabeto $A = \{0, 1\}$, obtidos pela Transformação Agregante (2.2). Seja o acoplamento Markoviano padrão $\mathbf{W} = (\mathbf{U}, \mathbf{V})$ com alfabeto $B \times B$, cujas probabilidades de transição são dadas pela matriz

$$P^W(U, V) = \begin{bmatrix} \beta & 0 & 0 & 0 & 1-\beta & 0 & 0 & 0 & 0 \\ \alpha' & 0 & \beta-\alpha' & 0 & 0 & 1-\beta & 0 & 0 & 0 \\ \beta & 0 & 0 & 0 & 1-\beta & 0 & 0 & 0 & 0 \\ \beta & \alpha-\beta & 0 & 0 & 0 & 0 & 0 & 1-\alpha & 0 \\ \alpha' & 0 & \alpha-\alpha' & 0 & 0 & 0 & 0 & 0 & 1-\alpha \\ \beta & 0 & 0 & 0 & 1-\beta & 0 & 0 & 0 & 0 \\ \alpha' & 0 & \beta-\alpha' & 0 & 0 & 1-\beta & 0 & 0 & 0 \\ \beta & 0 & 0 & 0 & 1-\beta & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Tomemos agora o processo de renovação $\mathbf{Z} = T(\mathbf{W})$ com alfabeto $A \times A$, obtido pela Transformação Agregante Estendida (3.7). Neste caso, \mathbf{Z} alcança a distância \bar{d} entre \mathbf{X} e \mathbf{Y} .

Este exemplo é uma aplicação do Teorema 3.2.

Observando as matrizes de transição de \mathbf{U} e \mathbf{V} , obtemos as medidas estacionárias

$$(\mu_{\mathbf{U}}(1), \mu_{\mathbf{U}}(2), \mu_{\mathbf{U}}(3)) = \left(\frac{\alpha + \beta - \alpha\beta}{2 - \beta}, \frac{1 - \beta}{2 - \beta}, \frac{(1 - \alpha)(1 - \beta)}{2 - \beta} \right)$$

e

$$(\mu_{\mathbf{V}}(1), \mu_{\mathbf{V}}(2), \mu_{\mathbf{V}}(3)) = \left(\frac{\alpha' + \beta - \alpha'\beta}{2 - \beta}, \frac{1 - \beta}{2 - \beta}, \frac{(1 - \alpha')(1 - \beta)}{2 - \beta} \right).$$

Construímos os processos de renovação através da Transformação **Agregante** (2.2): $\mathbf{X} = t(\mathbf{U})$ e $\mathbf{Y} = t(\mathbf{V})$. As medidas estacionárias dos processos de renovação gerados por esses processos serão, respectivamente

$$(\mu_{\mathbf{X}}(1), \mu_{\mathbf{X}}(0)) = (\mu_{\mathbf{U}}(1), \mu_{\mathbf{U}}(2) + \mu_{\mathbf{U}}(3)) = \left(\frac{\alpha + \beta - \alpha\beta}{2 - \beta}, \frac{(2 - \alpha)(1 - \beta)}{2 - \beta} \right)$$

$$(\mu_{\mathbf{Y}}(1), \mu_{\mathbf{Y}}(0)) = (\mu_{\mathbf{V}}(1), \mu_{\mathbf{V}}(2) + \mu_{\mathbf{V}}(3)) = \left(\frac{\alpha' + \beta - \alpha'\beta}{2 - \beta}, \frac{(2 - \alpha')(1 - \beta)}{2 - \beta} \right).$$

Podemos calcular a distância de partição utilizando a Definição 3.7. É fácil ver que as distâncias de partição, Markoviana e de renovação, são iguais e dadas por

$$d_p(\mathbf{U}, \mathbf{V}) = d_p(\mathbf{X}, \mathbf{Y}) = \frac{(1 - \beta)(\alpha - \alpha')}{2 - \beta}.$$

Vamos verificar que a Transformação **Agregante** (2.2) satisfaz a Condição 3.1. De fato, tome a partição: $A_1 = \{1\}$ e $A_2 = \{2, 3\}$, note que $\mu_{\mathbf{U}}(1) \leq \mu_{\mathbf{V}}(1)$ ou $\mu_{\mathbf{U}}(1) \geq \mu_{\mathbf{V}}(1)$ em A_1 , como $\mu_{\mathbf{U}}(2) = \mu_{\mathbf{V}}(2)$ o mesmo acontece para os elementos de A_2 . Então como o acoplamento Markoviano $\mathbf{W} = (\mathbf{U}, \mathbf{V})$ alcança a distância de partição entre \mathbf{U} e \mathbf{V} , pelo Teorema 3.2 o acoplamento de renovação $\mathbf{Z} = T(\mathbf{W})$ obtido pela Transformação Agregante Estendida (3.7) alcança a distância de partição entre $\mathbf{X} = t(\mathbf{U})$ e $\mathbf{Y} = t(\mathbf{V})$. ■

Pelo Exemplo 3.1, podemos concluir que a Condição De Ordenação (3.13) não é uma condição necessária entre dois processos de renovação, para que exista um acoplamento que alcance a distância \bar{d} . Como no exemplo em questão, temos $\alpha' \leq \beta \leq \alpha$. Se considerarmos a sequência de passados: $x_{-\infty}^{-1} = 0001$ de \mathbf{X} e $y_{-\infty}^{-1} = 0011$ de \mathbf{Y} . Temos que $x_{-\infty}^{-1} \leq y_{-\infty}^{-1}$, mas

$$P^{\mathbf{X}}(1|x_{-\infty}^{-1}) \geq P^{\mathbf{Y}}(1|y_{-\infty}^{-1}).$$

O exemplo seguinte mostra um acoplamento de renovação entre pares de processos de renovação, que alcança a distância \bar{d} , sem, porém, satisfazer às condições do Teorema 3.2

Exemplo 3.2 *Sejam dois processos Markovianos \mathbf{U} e \mathbf{V} , com alfabeto $B = \{1, 2, 3\}$ e probabilidades de transição dadas, respectivamente, pelas matrizes*

$$P^{\mathbf{U}} = \begin{bmatrix} \beta & 1 - \beta & 0 \\ \alpha & 0 & 1 - \alpha \\ \beta & 1 - \beta & 0 \end{bmatrix}$$

e

$$P^{\mathbf{V}} = \begin{bmatrix} \beta' & 1 - \beta' & 0 \\ \alpha & 0 & 1 - \alpha \\ \beta' & 1 - \beta' & 0 \end{bmatrix}$$

com $\beta' \leq \alpha \leq \beta$.

Tomemos os processos de renovação $\mathbf{X} = t(\mathbf{U})$ e $\mathbf{Y} = t(\mathbf{V})$ com alfabeto $A = \{0, 1\}$, obtidos pela Transformação Agregante (2.2). Seja o acoplamento Markoviano padrão $\mathbf{W} = (\mathbf{U}, \mathbf{V})$ com alfabeto

$B \times B$, cujas probabilidades de transição são dadas pela matriz

$$P^W(U, V) = \begin{bmatrix} \beta' & \beta - \beta' & 0 & 0 & 1 - \beta & 0 & 0 & 0 & 0 \\ \alpha & 0 & \beta - \alpha & 0 & 0 & 1 - \beta & 0 & 0 & 0 \\ \beta' & \beta - \beta' & 0 & 0 & 1 - \beta & 0 & 0 & 0 & 0 \\ \beta' & \alpha - \beta' & 0 & 0 & 0 & 0 & 0 & 1 - \alpha & 0 \\ \alpha & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 - \alpha \\ \beta' & \alpha - \beta' & 0 & 0 & 0 & 0 & 0 & 1 - \alpha & 0 \\ \beta' & \beta - \beta' & 0 & 0 & 1 - \beta & 0 & 0 & 0 & 0 \\ \alpha & 0 & \beta - \alpha & 0 & 0 & 1 - \beta & 0 & 0 & 0 \\ \beta' & \beta - \beta' & 0 & 0 & 1 - \beta & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Neste caso, \mathbf{Z} é um acoplamento de renovação entre os processos de renovação \mathbf{X} e \mathbf{Y} , que alcança a distância \bar{d} .

Inicialmente vamos mostrar que \mathbf{Z} é de fato um acoplamento entre os processos \mathbf{X} e \mathbf{Y} . Para isso mostraremos que $\mu_{\mathbf{X}}$ e $\mu_{\mathbf{Y}}$ são marginais de $\mu_{\mathbf{Z}}$ utilizando (3.1) e (3.2).

Calculando as medidas estacionárias de \mathbf{U} e \mathbf{V} obtemos

$$\begin{aligned} (\mu_{\mathbf{U}}(1), \mu_{\mathbf{U}}(2), \mu_{\mathbf{U}}(3)) &= \left(\frac{\alpha + \beta - \alpha \cdot \beta}{2 - \beta}, \frac{1 - \beta}{2 - \beta}, \frac{(1 - \beta)(1 - \alpha)}{2 - \beta} \right) \\ (\mu_{\mathbf{V}}(1), \mu_{\mathbf{V}}(2), \mu_{\mathbf{V}}(3)) &= \left(\frac{\alpha + \beta' - \alpha \cdot \beta'}{2 - \beta'}, \frac{1 - \beta'}{2 - \beta'}, \frac{(1 - \beta')(1 - \alpha)}{2 - \beta'} \right). \end{aligned}$$

Pela Transformação Agregante (2.2), obtemos as medidas de \mathbf{X} e \mathbf{Y}

$$\begin{aligned} (\mu_{\mathbf{X}}(1), \mu_{\mathbf{X}}(0)) &= \left(\frac{\alpha + \beta - \alpha \cdot \beta}{2 - \beta}, \frac{(1 - \beta)(2 - \alpha)}{2 - \beta} \right) \\ (\mu_{\mathbf{Y}}(1), \mu_{\mathbf{Y}}(0)) &= \left(\frac{\alpha + \beta' - \alpha \cdot \beta'}{2 - \beta'}, \frac{(1 - \beta')(2 - \alpha)}{2 - \beta'} \right). \end{aligned}$$

Da mesma forma, calculamos as medidas de probabilidade do acoplamento \mathbf{W}

$$\begin{aligned} &(\mu_W(1, 1), \mu_W(1, 2), \mu_W(1, 3), \mu_W(2, 1), \mu_W(2, 2), \mu_W(2, 3), \mu_W(3, 1), \mu_W(3, 2), \mu_W(3, 3)) \\ &= \left(\frac{\alpha\beta' - \alpha - \beta'}{2 - \beta'}, \frac{\alpha\beta - \alpha\beta' - \beta\beta' - \alpha\beta^2 + \alpha\beta\beta' + \beta^2}{\beta(2 - \beta)(2 - \beta')}, \frac{\beta^2 - \beta\beta' + \alpha\beta' - \alpha\beta}{\beta(2 - \beta)(2 - \beta')}, 0, \right. \\ &\quad \frac{\beta^2\beta' - \beta^2 - 2\beta\beta' + \beta + \beta'}{\beta(2 - \beta)(2 - \beta')}, \frac{(1 - \beta)(\beta - \beta')}{\beta(2 - \beta)(2 - \beta')}, 0, \frac{(\beta - \beta')(1 - \alpha)(1 - \beta)}{(2\beta - \beta^2)(2 - \beta')}, \\ &\quad \left. \frac{\beta - \alpha\beta - \beta^2 + \alpha\beta^2 + \beta' - \alpha\beta' - 2\beta\beta' - 2\alpha\beta\beta' + \beta^2\beta' - \alpha\beta^2\beta'}{\beta(2 - \beta)(2 - \beta')} \right). \end{aligned}$$

Assim, o acoplamento de renovação $\mathbf{Z} = T(\mathbf{W})$ terá medida estacionária

$$(\mu_{\mathbf{Z}}(1, 1), \mu_{\mathbf{Z}}(1, 0), \mu_{\mathbf{Z}}(0, 1), \mu_{\mathbf{Z}}(0, 0)) = \left(\frac{\alpha + \beta' - \alpha\beta'}{2 - \beta'}, \frac{(2 - \alpha)(\beta - \beta')}{(2 - \beta)(2 - \beta')}, 0, \frac{(2 - \alpha)(1 - \beta)}{2 - \beta} \right).$$

Então

$$\begin{aligned} \mu_{\mathbf{X}}(1) &= \mu_{\mathbf{Z}}(1, 1) + \mu_{\mathbf{Z}}(1, 0) \\ &= \frac{\alpha + \beta' - \alpha\beta'}{2 - \beta'} + \frac{(2 - \alpha)(\beta - \beta')}{(2 - \beta)(2 - \beta')} \\ &= \frac{\alpha + \beta - \alpha\beta}{2 - \beta}, \end{aligned}$$

$$\begin{aligned} \mu_{\mathbf{X}}(0) &= \mu_{\mathbf{Z}}(0, 0) + \mu_{\mathbf{Z}}(0, 1) \\ &= \frac{(2 - \alpha)(1 - \beta)}{2 - \beta} + 0 \\ &= \frac{(2 - \alpha)(1 - \beta)}{2 - \beta}. \end{aligned}$$

Da mesma forma temos

$$\begin{aligned} \mu_{\mathbf{Y}}(1) &= \mu_{\mathbf{Z}}(1, 1) + \mu_{\mathbf{Z}}(0, 1) \\ &= \frac{\alpha + \beta' - \alpha\beta'}{2 - \beta'} + 0 \\ &= \frac{\alpha + \beta' - \alpha\beta'}{2 - \beta'}, \end{aligned}$$

$$\begin{aligned} \mu_{\mathbf{Y}}(0) &= \mu_{\mathbf{Z}}(0, 0) + \mu_{\mathbf{Z}}(1, 0) \\ &= \frac{(2 - \alpha)(1 - \beta)}{2 - \beta} + \frac{(2 - \alpha)(\beta - \beta')}{(2 - \beta)(2 - \beta')} \\ &= \frac{(2 - \alpha)(1 - \beta')}{2 - \beta'}. \end{aligned}$$

Portanto \mathbf{Z} é acoplamento entre \mathbf{X} e \mathbf{Y} . \mathbf{Z} é também um processo de renovação pois possui um subconjunto unitário $\{(1, 1)\}$ na partição de $B \times B$.

Vamos verificar agora o alcance à distância \bar{d} . Podemos calcular a distância de partição utilizando a Definição 3.7. É fácil ver que as distâncias de partição são dadas por

$$d_p(\mathbf{U}, \mathbf{V}) = d_p(\mathbf{X}, \mathbf{Y}) = \frac{(2 - \alpha)(\beta - \beta')}{(2 - \beta)(2 - \beta')}.$$

O acoplamento \mathbf{W} tem discrepância

$$d(\mathbf{W}) = \frac{\beta(2 + \alpha - 2\beta) + \beta'(2 - \alpha - 2\beta + \beta^2)}{\beta(2 - \beta)(2 - \beta')},$$

que não alcança a distância de partição.

Utilizando raciocínio análogo, vê-se que acoplamento de renovação alcança a distância de partição e a distância \bar{d} , isto é

$$d(\mathbf{Z}) = \frac{(2 - \alpha)(\beta - \beta')}{(2 - \beta)(2 - \beta')}$$

■

Algumas observações devem ser feitas em relação ao Exemplo 3.2

1. Os processos \mathbf{X} e \mathbf{Y} não satisfazem a Condição De Ordenação (3.13).
2. Note que o acoplamento de renovação \mathbf{Z} alcança a distância de partição e a distancia \bar{d} no caso $\beta' \leq \alpha \leq \beta$. Nos demais casos não é difícil mostrar que a distância de partição não é alcançada.
3. O acoplamento Markoviano \mathbf{W} não alcança a distância \bar{d} , enquanto o acoplamento de renovação \mathbf{Z} a alcança, sendo assim, uma ilustração da proposição 3.2.

Uma observação em relação a ambos os Exemplos, 3.1 e 3.2, é que eles possuem probabilidades de transição descontínuas. Um tema interessante de pesquisa é tratar apenas dos casos contínuos e verificar se a Condição De Ordenação (3.13) é necessária ou se há outra condição melhor.

Capítulo 4

Conclusão

A área de processos estocásticos de ordem infinita, acoplamentos e distância \bar{d} , tem sido objeto de estudo de muitos pesquisadores. As principais contribuições desta tese são

- Utilização de uma abordagem matricial para processos de renovação binários, obtidos a partir de um processo Markoviano através da Transformação Agregante (2.2), a qual permite o cálculo mais simplificado das probabilidades de transição, bem como a percepção das condições para a agregabilidade, k -agregabilidade, continuidade e Gibbsianidade.
- Obtenção de condições necessárias e suficientes para que um processo de renovação binário, obtido por agregação de estados a partir de um processo Markoviano com alfabeto finito, possua probabilidades de transição contínuas. Assim como, obtenção de condições suficientes para a Gibbsianidade do mesmo processo.
- Estudos sobre acoplamentos de renovação que alcançam a distância \bar{d} , utilizando agregação de estados por uma transformação agregante.

Outros temas de pesquisa que esta tese sugere são

- (1) Estudar processos de renovação não binários, obtidos através da Transformação Agregante (2.2). Desenvolver uma abordagem matricial equivalente e buscar condições para que o processo agregado continue Markoviano de ordem finita $k \geq 1$, bem como para a continuidade das probabilidades de transição. Estudar, neste caso, as condições para a Gibbsianidade de medidas do processo de renovação não binário, obtido com este procedimento.
- (2) Determinar sob que condições a distância \bar{d} é alcançada por um acoplamento de renovação, entre dois processos de renovação binários com probabilidades de transição contínua. Esta condição se estende a processos não binários? A continuidade das probabilidades de transição garantiria que a Condição De Ordenação (3.13) fosse necessária e suficiente para o acoplamento alcançar a distância \bar{d} ?
- (3) Buscar uma generalização para processos de renovação obtidos por uma transformação agregante, a partir de um processo Markoviano com alfabeto infinito enumerável.

- (4) Generalizar o estudo realizado nesta tese considerando uma transformação não determinística, mas uma transformação aleatória qualquer.

Referências Bibliográficas

- BALL, F. & YEO, G. F. (1993). Lumpability and marginalisability for continuous-time Markov chains. *J. Appl. Probab.* **30**(3), 518–528. 7
- BAUM, L. E. & PETRIE, T. (1966). Statistical inference for probabilistic functions of finite state Markov chains. *Ann. Math. Statist.* **37**, 1554–1563. 1
- BOUCHERIE, R. J. (1993). Aggregation of Markov chains. *Stochastic Process. Appl.* **45**(1), 95–114. 1
- BOWEN, R. (2008). *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, vol. 470 of *Lecture Notes in Mathematics*. Berlin: Springer-Verlag, revised ed. With a preface by David Ruelle, Edited by Jean-René Chazottes. 2
- BOYLE, M. & PETERSEN, K. (2011). Hidden Markov processes in the context of symbolic dynamics. In: *Entropy of hidden Markov processes and connections to dynamical systems*, vol. 385 of *London Math. Soc. Lecture Note Ser.* Cambridge: Cambridge Univ. Press, pp. 5–71. 1
- BRÉMAUD, P. (1999). *Markov chains*, vol. 31 of *Texts in Applied Mathematics*. New York: Springer-Verlag. Gibbs fields, Monte Carlo simulation, and queues. 55, 56, 58
- BUCHHOLZ, P. (1994). Exact and ordinary lumpability in finite Markov chains. *J. Appl. Probab.* **31**(1), 59–75. 7
- BURKE, C. J. & ROSENBLATT, M. (1958). A Markovian function of a Markov chain. *Ann. Math. Statist.* **29**, 1112–1122. 7
- CÉNAC, P., CHAUVIN, B., PACCAUT, F. & POUYANNE, N. (2012). Context trees, variable length Markov chains and dynamical sources. In: *Séminaire de Probabilités XLIV*, vol. 2046 of *Lecture Notes in Math.* Heidelberg: Springer, pp. 1–39. 1
- CHAZOTTES, J.-R. & UGALDE, E. (2003). Projection of Markov measures may be Gibbsian. *J. Statist. Phys.* **111**(5-6), 1245–1272. 2
- CHAZOTTES, J.-R. & UGALDE, E. (2011). On the preservation of Gibbsianness under symbol amalgamation. In: *Entropy of hidden Markov processes and connections to dynamical systems*,

- vol. 385 of *London Math. Soc. Lecture Note Ser.* Cambridge: Cambridge Univ. Press, pp. 72–97. [1](#)
- DOBRUSCHIN, P. (1968). The description of a random field by means of conditional probabilities and conditions of its regularity. *Journal of Applied Probability* **13**(2), 197–224. [2](#), [19](#)
- ELLIS, M. H. (1976). The \bar{d} -distance between two Markov processes cannot always be attained by a Markov joining. *Israel J. Math.* **24**(3-4), 269–273. [2](#), [34](#)
- ELLIS, M. H. (1978). Distances between two-state Markov processes attainable by Markov joinings. *Trans. Amer. Math. Soc.* **241**, 129–153. [2](#)
- ELLIS, M. H. (1980a). Conditions for attaining \bar{d} by a Markovian joining. *Ann. Probab.* **8**(3), 431–440. [2](#)
- ELLIS, M. H. (1980b). On Kamae’s conjecture concerning the \bar{d} -distance between two-state Markov processes. *Ann. Probab.* **8**(2), 372–376. [2](#), [32](#)
- FERNÁNDEZ, R., FERRARI, P. A. & GALVES, A. (2001). Coupling, renewal and perfect simulation of chains of infinite order. [2](#), [32](#)
- FERNÁNDEZ, R., GALLO, S. & MAILLARD, G. (2011). Regular C -measures are not always Gibbsian. *Electron. Commun. Probab.* **16**, 732–740. [2](#), [19](#)
- FERNÁNDEZ, R. & MAILLARD, G. (2004). Chains with complete connections and one-dimensional Gibbs measures. *Electron. J. Probab.* **9**, no. 6, 145–176 (electronic). [21](#)
- FROBENIUS, G. (1911). Gegenseitige Reduktion algebraischer Körper. *Math. Ann.* **70**(4), 457–458. [56](#)
- GALVES, A., GARCIA, N. L. & PRIEUR, C. (2010). Perfect simulation of a coupling achieving the \bar{d} -distance between ordered pairs of binary chains of infinite order. *J. Stat. Phys.* **141**(4), 669–682. [2](#), [40](#)
- GEORGII, H.-O. (2011). *Gibbs measures and phase transitions*, vol. 9 of *de Gruyter Studies in Mathematics*. Berlin: Walter de Gruyter & Co., second ed. [19](#)
- GURVITS, L. & LEDOUX, J. (2005). Markov property for a function of a Markov chain: a linear algebra approach. *Linear Algebra Appl.* **404**, 85–117. [7](#)
- HAN, G. & MARCUS, B. (2006). Analyticity of entropy rate of hidden Markov chains. *IEEE Trans. Inform. Theory* **52**(12), 5251–5266. [1](#)
- HARRIS, T. E. (1955). On chains of infinite order. *Pacific J. Math.* **5**, 707–724. [1](#), [2](#), [9](#)

- JUANG, B.-H. & RABINER, L. R. (1991). Hidden Markov models for speech recognition. *Technometrics* **33**(3), 251–272. 1
- KEMENY, J. G. & SNELL, J. L. (1976). *Finite Markov chains*. New York: Springer-Verlag. Reprinting of the 1960 original, Undergraduate Texts in Mathematics. 1, 7
- KITCHENS, B. P. (1998). *Symbolic dynamics*. Universitext. Berlin: Springer-Verlag. One-sided, two-sided and countable state Markov shifts. 56
- LEDOUX, J., RUBINO, G. & SERICOLA, B. (1994). Exact aggregation of absorbing Markov processes using the quasi-stationary distribution. *J. Appl. Probab.* **31**(3), 626–634. 7
- LESTAS, I., PAULSSON, J., ROSS, N. E. & VINNICOMBE, G. (2008). Noise in gene regulatory networks [mr2492563]. *IEEE Trans. Automat. Control* (Special Issue on Systems Biology), 189–200. 1
- LIMA, E. L. (1996). *Algebra Linear - 2ed.* Colecao Matematica Universitaria. Rio de Janeiro: IMPA-CNPq. 12, 52
- MASSEY, J. L. (1978). Joint source and channel coding. In: *Communication systems and random process theory (Proc. 2nd NATO Advanced Study Inst., Darlington, 1977)*. Alphen aan den Rijn: Sijthoff & Noordhoff, pp. 279–293. NATO Advanced Study Inst. Ser., Ser. E: Appl. Sci., No. 25. 1
- MATŁOKA, M. (2000). Convex fuzzy processes. *Fuzzy Sets and Systems* **110**(1), 109–114. 1
- MEYER, C. (2000). *Matrix analysis and applied linear algebra*. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM). With 1 CD-ROM (Windows, Macintosh and UNIX) and a solutions manual (iv+171 pp.). 55, 56, 58
- NOUTSOS, D. (2006). On Perron-Frobenius property of matrices having some negative entries. *Linear Algebra Appl.* **412**(2-3), 132–153. 56
- ORNSTEIN, D. S. (1973). An application of ergodic theory to probability theory. *Ann. Probability* **1**(1), 43–65. 2
- PENG, N.-F. (1996). On weak lumpability of a finite Markov chain. *Statist. Probab. Lett.* **27**(4), 313–318. 7
- PERRON, O. (1907). Zur theorie der matrizen. *Math. Ann.* **64**(1), 248–263. 56
- RENCHER, A. C. & SCHAALJE, G. B. (2008). *Linear models in statistics*. Hoboken, NJ: Wiley-Interscience [John Wiley & Sons], second ed. 52

- RUBINO, G. & SERICOLA, B. (1989). On weak lumpability in Markov chains. *J. Appl. Probab.* **26**(3), 446–457. [7](#)
- RUBINO, G. & SERICOLA, B. (1991). A finite characterization of weak lumpable markov processes.part i: The discrete time case. *Stoch. Proc. and Appl.* **38**(2), 195–204. [7](#)
- SENETA, E. (2006). *Non-negative matrices and Markov chains*. Springer Series in Statistics. New York: Springer. Revised reprint of the second (1981) edition [Springer-Verlag, New York; MR0719544]. [55](#), [56](#), [58](#)
- SHIELDS, P. (1973). *The theory of Bernoulli shifts*. The University of Chicago Press, Chicago, Ill.-London. Chicago Lectures in Mathematics. [1](#)
- VERBITSKIY, E. (2011a). On factors of g -measures. *Indag. Math. (N.S.)* **22**(3-4), 315–329. [9](#)
- VERBITSKIY, E. (2011b). Thermodynamics of hidden Markov processes. In: *Entropy of hidden Markov processes and connections to dynamical systems*, vol. 385 of *London Math. Soc. Lecture Note Ser.* Cambridge: Cambridge Univ. Press, pp. 258–272. [9](#)

Apêndice A

Álgebra Matricial

Matrizes é o conjunto de números reais (ou complexos) distribuídos em m linhas e n colunas, com $m, n \in \mathbb{N}$. Algumas poucas propriedades de matrizes serão revistas neste apêndice, sendo extraídas principalmente de [Lima \(1996\)](#) e [Rencher & Schaalje \(2008\)](#). A notação natural de uma matriz é $P = (a_{ij})_{m \times n}$, com $i \in \{1, 2, \dots, m\}$ e $j \in \{1, 2, \dots, n\}$.

As operações de adição de matrizes, multiplicação por escalar e a multiplicação de matrizes podem ser combinadas de várias maneiras, e essas operações satisfazem a várias propriedades que nos são familiares. Gostaríamos de mencionar algumas.

Multiplicação de Matrizes

Para multiplicar a matriz P pela matriz Q de maneira a formar o produto $P.Q$, a dimensão das colunas de P deve ser igual à dimensão das linhas de Q . Portanto, seja P uma matriz $m \times n$ e seja Q uma matriz $n \times p$. Então, a multiplicação de matrizes é definida como

$$P.Q = \sum_{k=1}^n a_{ik}.b_{kj}.$$

Em outras palavras, o (i, j) -ésimo elemento da nova matriz $P.Q$ é obtido pela multiplicação de cada elemento na i -ésima linha de P pelo elemento correspondente na j -ésima coluna de Q e somando esses n produtos. Também podemos multiplicar uma matriz e um vetor, basta tratar o vetor como uma matriz-coluna $n \times 1$, ou como vetor-linha $1 \times n$ e considerar as mesmas condições e definição acima.

Potência de uma Matriz Quadrada de Ordem m

$$\begin{aligned} P^0 &= I_m \\ P^1 &= P \\ P^2 &= P.P \\ P^3 &= P.P.P \\ &\vdots \\ P^k &= P^{k-1}.P = P.P^{k-1}. \end{aligned}$$

Toda matriz quadrada P comuta com qualquer potência n de P , $n \in \mathbb{N}$.

A seguir apresentamos alguns resultados que serão úteis nas provas dos resultados desta tese.

Lema A.1 *Seja $P = (a_{ij})$ uma matriz quadrada de ordem m com termos $0 \leq a_{ij} \leq 1$, tal que P possui uma única coluna (ou linha) não nula, digamos $a_{.k}$, então*

$$P^n = a_{kk}^{n-1} \cdot P$$

para todo $n \in \mathbb{N}$.

Prova do Lema A.1

Façamos a demonstração por indução em n .

1. O Lema A.1 é válido para $n = 0$ e $n = 1$ de modo óbvio. Vamos mostrar que é válido para $n = 2$. De fato, considere, sem perda de generalidade, que apenas a primeira coluna $a_{.1}$ não seja nula, então

$$P = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & 0 & \dots & 0 \end{bmatrix},$$

assim

$$P^2 = \begin{bmatrix} a_{11}^2 & 0 & \dots & 0 \\ a_{11} \cdot a_{21} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{11} \cdot a_{m1} & 0 & \dots & 0 \end{bmatrix} = a_{11} \cdot P.$$

2. Supondo o Lema A.1 válido para n , isto é, $P^n = a_{11}^{n-1} \cdot P$. Então para $n + 1$, teremos

$$P^{n+1} = P^n \cdot P = a_{11}^{n-1} \cdot P \cdot P = a_{11}^{n-1} \cdot P^2 = a_{11}^{n-1} \cdot a_{11} \cdot P = a_{11}^n \cdot P.$$

■

Lema A.2 *Seja $P = (a_{ij})$ uma matriz quadrada de ordem m , com termos $0 \leq a_{ij} \leq 1$. Se P é triangular superior (ou inferior) tal que, para algum $h \in \{1, \dots, m\}$, $a_{hh} \neq 0$ e $a_{ii} = 0$, para todo $i \neq h, i = 1, \dots, m$ então*

$$P^n = a_{hh}^{n-k} \cdot P^k$$

para todo $n \in \mathbb{N}$, onde $k < n$ é o número de linhas não nulas e $n \geq m$.

Prova do Lema A.2

Faremos a demonstração por indução em n .

1. De modo óbvio, o Lema A.2 é válido para $n = 0$ e $n = 1$. Vamos mostrar que é válido para $n = 2$. De fato, como $k < n$, a condição implica que P possui uma única coluna não nula e, pelo Lema A.1, o resultado segue.
2. Suponha o Lema A.2 válido para n , isto é, P é triangular com $a_{ii} = 0$, para todo $i \neq h, i = 1, \dots, m$. Suponha, sem perda de generalidade, que P_m possua as $k \leq m - 1$ primeiras linhas não nulas. Então

$$P^n = a_{hh}^{n-k} \cdot P^k.$$

Portanto para $n + 1$ teremos

$$P^{n+1} = P^n \cdot P = a_{hh}^{n-k} \cdot P^k \cdot P = a_{hh}^{n-k} \cdot P^{k+1},$$

como $k + 1 > k$, novamente temos a hipótese indutiva,

$$= a_{hh}^{n-k} \cdot a_{hh}^{k+1-k} \cdot P^k = a_{hh}^{n+1-k} \cdot P^k,$$

que encerra a prova.

■

Apêndice B

Teorema De Perron-Frobênus

Quando o alfabeto de um processo Markoviano é finito, podemos contar com os resultados de álgebra linear para estudar o seu comportamento assintótico. Realmente, esse comportamento, no tempo n , é totalmente descrito pelo comportamento da n -ésima potência da sua matriz de transição, P . O Teorema de Perron-Frobenius descreve a auto-estrutura de matrizes não negativas, e portanto, descreve tudo que precisamos saber para entender o comportamento assintótico dos processos Markovianos estudados nesta tese. Uma revisão básica da Teoria de Perron-Frobênus apresentada neste apêndice foi extraída de [Brémaud \(1999\)](#), [Meyer \(2000\)](#) e [Seneta \(2006\)](#).

O principal resultado de Perron e Frobenius, é que a convergência para um estado de equilíbrio de um processo Markoviano é geométrica, com uma taxa relativa igual ao módulo do segundo maior auto-valor.

Algumas definições precisam ser introduzidas neste momento.

Definição B.1 *Uma matriz permutação é uma matriz quadrada, binária, que quando multiplicada, simultaneamente, à direita e à esquerda tem o efeito de gerar uma permutação nos elementos de um vetor, ou entre as linhas ou colunas de uma matriz. É formada apenas de zeros e uns, sendo o valor de apenas um elemento por linha e por coluna igual a um.*

Em particular, uma matriz permutação C é uma matriz ortogonal, de modo que $C^{-1} = C'$, isto é, a inversa de C é também uma matriz de permutação. Uma matriz transposição T é um caso bem particular de matriz permutação, que é igual a matriz identidade exceto em duas posições, sendo, então uma matriz simétrica. O efeito de multiplicar uma matriz P por uma matriz transposição a esquerda é trocar a posição de duas linhas da matriz P , e a multiplicação de P pela matriz de transposição a direita troca a posição de duas colunas de P . Uma propriedade importante da matriz permutação é que toda matriz permutação C é o produto de matrizes transposição. Assim, a matriz

$$C.P.C' = T_m \dots T_1.P.T_1 \dots T_m$$

é obtida pela permutação de linhas e colunas de P , de tal forma que nenhum elemento é criado ou perdido.

Definição B.2 Uma matriz $P_{m \times m}$ não-negativa é dita ser **Redutível** se existe uma matriz-permutação C tal que

$$P = C \cdot \Gamma \cdot C', \text{ onde } \Gamma = \begin{bmatrix} P_{11} & P_{12} \\ \mathbf{0} & P_{22} \end{bmatrix},$$

onde $P_{11} \in P_{r \times r}$, $P_{22} \in P_{(n-r) \times (n-r)}$, e $P_{12} \in P_{r \times (n-r)}$.

Definição B.3 Uma matriz $P_{m \times m}$ não-negativa é dita ser **irredutível** se não for redutível.

Uma consequência imediata da definição é que, para que a matriz P , quadrada de ordem m , seja redutível, ela deve ter pelo menos $m - 1$ zeros.

A prova dos teoremas e corolários relacionados à Teoria de Perron-Frobenius, abaixo, são encontrados em [Kitchens \(1998\)](#), [Brémaud \(1999\)](#), [Meyer \(2000\)](#), [Seneta \(2006\)](#) e [Noutsos \(2006\)](#).

Os auto-valores de uma matriz quadrada real P são números complexos e determinam o raio espectral da matriz. Ordenando todos os auto-valores $\lambda_1, \lambda_2, \dots, \lambda_r$ da matriz P de tal forma que

$$|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_r|. \quad (\text{B.1})$$

Definimos o raio espectral da forma a seguir.

Definição B.4 O raio espectral de uma matriz $P_{m \times m}$, denotado por $\rho(P)$ é dado por

$$\rho(P) = |\lambda_1|.$$

Definição B.5 O espectro de uma matriz $P_{m \times m}$ é o conjunto de todos os auto-valores de P .

Definição B.6 O círculo espectral de uma matriz $P_{m \times m}$ é o módulo do auto-valor de maior valor absoluto ($|\lambda_1|$).

Através da teoria de Perron-Frobenius podemos concluir que P^n , com $n \rightarrow \infty$, é controlado pelo auto-valor de maior valor absoluto. O Teorema de Perron-Frobenius nos fornece as propriedades dos auto-valores e de seus respectivos auto-vetores. Apresentamos duas versões deste importante teorema, a primeira versão é para matrizes positivas, trata-se do Teorema de Perron, [Perron \(1907\)](#), apresentada no Teorema [B.1](#), e a segunda para matrizes não negativas, tratando-se do Teorema de Perron-Frobenius, [Frobenius \(1911\)](#), apresentada no Teorema [B.2](#).

Teorema B.1 Seja $P_{m \times m}$ uma matriz positiva. Então

1. P tem pelo menos um auto-valor real positivo, o maior auto-valor positivo é o raio espectral $\rho(P)$.
2. λ_1 é um auto-valor simples de P , isto é, é uma raiz simples do polinômio característico de P , gerando o auto-espaço de P de dimensão 1 (chamado auto-espaço direito). Analogamente teremos o auto-espaço de P' de dimensão 1 (chamado auto-espaço esquerdo).

3. O auto-vetor ϕ correspondente a λ_1 é positivo, isto é, suas coordenadas são estritamente maiores de zero.

$$P.\phi = \lambda_1.\phi, \text{ com } \phi > 0.$$

4. Não existe outro auto-vetor positivo de P , exceto $\phi' = c.\phi$, $c \in \mathbb{N}$.

5. λ_1 é o único auto-valor sobre o círculo espectral de P .

O maior auto-valor de P é chamado de **Auto-valor de Perron-Frobenius**, o qual denotaremos por λ_1 .

Considerando agora a matriz P com, pelo menos uma entrada nula, a proposição B.1, a seguir, mostra que o Teorema B.1 se estende para matrizes não-negativas sacrificando a existência de um auto-valor positivo por um não-negativo.

Proposição B.1 *Seja $P_{m \times m}$ uma matriz não-negativa. Então*

1. P tem pelo menos um auto-valor real não-negativo, o maior auto-valor não-negativo é o raio espectral $\rho(P)$.

2. O auto-vetor ϕ correspondente a λ_1 é positivo: $P.\phi = \lambda_1.\phi$, com $\phi > 0$.

Os resultados acima são os únicos que valem para matrizes quadradas não-negativas, sem hipóteses adicionais. As propriedades de Perron-Frobenius para matrizes não-negativas são apresentados a seguir.

Teorema B.2 *Seja $P_{m \times m}$ uma matriz não-negativa e irredutível. Então*

1. P tem pelo menos um auto-valor real não-negativo, o maior auto-valor não-negativo é o raio espectral - $\lambda_1 = \rho(P)$.

2. λ_1 é um auto-valor simples de P .

3. O auto-vetor ϕ correspondente a λ_1 é positivo: $P.\phi = \lambda_1.\phi$, com $\phi > 0$.

4. Não existe outro auto-vetor positivo de P , exceto um múltiplo de ϕ .

Observe que a única propriedade do Teorema de Perron que a irredutibilidade não é capaz de manter, é a unicidade de λ_1 no círculo espectral. A propriedade de ter ou não ter um único auto-valor sobre o círculo espectral, divide o conjunto das matrizes não negativas irredutíveis em duas classes de matrizes: **Primitivas e cíclicas**.

Definição B.7 *Uma matriz não-negativa e irredutível $P_{m \times m}$ é dita ser Primitiva se ela possui um único auto-valor sobre o seu círculo espectral, isto é, λ_1 é o único auto-valor de módulo igual a $\rho(P)$.*

Definição B.8 Uma matriz não-negativa e irredutível $P_{m \times m}$ é dita ser *Imprimitiva* ou *Cíclica* de índice $h > 1$, se ela possui h auto-valores de módulo igual a $\rho(P)$.

Teorema B.3 Seja $P_{m \times m}$, uma matriz não-negativa e irredutível. Então existe $r \geq 1$ tal que $P^r > 0$ se, e somente se, P é primitiva.

O menor inteiro r , tal que $P^r > 0$, é chamado índice de primitividade.

Corolário B.1 Seja $P_{m \times m}$, uma matriz não-negativa e irredutível com espaço de estados finito, ela é primitiva se, e somente se ela é irredutível e aperiódica.

Para a prova do Corolário B.1, basta verificar que a condição necessária e suficiente para que a matriz P seja aperiódica é a existência de um inteiro k tal que P^k tenha todas as entradas positivas.

O Teorema B.4, seguinte, devido a Frobenius, foi extraído de Brémaud (1999), Meyer (2000) e Seneta (2006) e será fundamental para a prova dos Teoremas apresentados na Seção 2.4 desta tese.

Teorema B.4 Seja $P_{m \times m}$, uma matriz não-negativa e irredutível com $\lambda_1 = \rho(P)$. Então P é primitiva se, e somente se, o limite $\lim_{k \rightarrow \infty} \left(\frac{P}{\lambda_1}\right)^k$ existe, e neste caso

$$\lim_{k \rightarrow \infty} \left(\frac{P}{\lambda_1}\right)^k = G = \phi \cdot \psi' > 0,$$

onde ϕ e ψ são os auto-vetores de Perron de P e P' , respectivamente.

O Teorema B.5, a seguir, extraído de Seneta (2006) e Brémaud (1999), é utilizado na prova da taxa de continuidade das probabilidades de transição dos processos agregados, apresentada na Seção 2.4.

Sejam $\lambda_1, \lambda_2, \dots, \lambda_r$ todos os auto-valores da matriz P ordenados de tal forma que

$$\lambda_1 > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_r|.$$

Se $|\lambda_2| = |\lambda_j|$ para algum $j \geq 3$, então $m_2 \geq m_j$, onde m_j é a multiplicidade algébrica de λ_j .

Teorema B.5 Para $P_{m \times m}$ uma matriz primitiva

1. Se $\lambda_2 \neq 0$ então

$$P^k = \lambda_1^k \cdot \phi \cdot \psi' + O(k^{m_2-1} \cdot |\lambda_2|^k)$$

quando $k \rightarrow \infty$.

2. Se $\lambda_2 = 0$ então para $k \geq n - 1$

$$P^k = \lambda_1^k \cdot \phi \cdot \psi'.$$

Em ambos os casos, ϕ e ψ' são os auto-vetores positivos de Perron-Frobenius, correspondentes ao auto-valor λ_1 garantidos pelos Teoremas B.2 e B.4.

O Teorema B.6 apresenta a forma de Frobenius de uma matriz cíclica e é apresentado a seguir.

Teorema B.6 *Para cada matriz cíclica $P_{m \times m}$ de índice h , existe uma matriz-permutação C , tal que:*

$$P = C.\Delta.C' \text{ onde } \Delta = \begin{bmatrix} \mathbf{0} & P_{12} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & P_{23} & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & P_{h-1,h} \\ P_{h1} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \end{bmatrix}, \quad (\text{B.2})$$

onde blocos nulos sobre a diagonal principal são quadrados e os blocos P_{ij} serão chamados de **blocos cíclicos**.

O Teorema de Perron-Frobênus não é aplicável diretamente a matrizes não negativas redutíveis. Para estudar a sua aplicabilidade a essas matrizes, devemos entender a forma canônica das matrizes redutíveis.

Seja uma Matriz P redutível, logo, existe uma matriz permutação Q , e matrizes quadradas X , Y e Z , tais que

$$Q'.P.Q = \begin{bmatrix} X & Y \\ \mathbf{0} & Z \end{bmatrix},$$

cuja notação usaremos $P = \begin{bmatrix} X & Y \\ \mathbf{0} & Z \end{bmatrix}$.

Se as submatrizes X ou Z forem redutíveis, da mesma forma, existirão matrizes permutações, de tal forma que

$$P = \begin{bmatrix} X & Y \\ \mathbf{0} & Z \end{bmatrix} = \begin{bmatrix} R & S & T \\ \mathbf{0} & U & V \\ \mathbf{0} & \mathbf{0} & W \end{bmatrix}$$

e assim sucessivamente até obter

$$P = \begin{bmatrix} Q_{11} & Q_{12} & \dots & Q_{1r} \\ \mathbf{0} & Q_{22} & \dots & Q_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & Q_{rr} \end{bmatrix},$$

onde Q_{ii} são irredutíveis ou 0. É possível permutar todos os blocos diagonais de entradas não nulas para a parte inferior da matriz, obtendo assim a **forma canônica de uma matriz redutível**,

$$P = \begin{bmatrix} T_{11} & T_{12} & T_{13} & \dots & T_{1c} \\ \mathbf{0} & T_{22} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & T_{33} & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & T_{cc} \end{bmatrix}, \quad (\text{B.3})$$

onde

$$T_{11} = \begin{bmatrix} P_{11} & P_{12} & P_{13} & \dots & P_{1k} \\ \mathbf{0} & P_{22} & P_{23} & \dots & P_{2k} \\ \mathbf{0} & \mathbf{0} & P_{33} & \dots & P_{3k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & P_{kk} \end{bmatrix}$$

é formada pelas submatrizes $P_{11}, P_{22}, \dots, P_{kk}$, quadradas, irredutíveis ou 0 e de tal forma que, se fossem estocásticas, corresponderiam aos estados transientes e $P_{ij}, i < j$. As demais submatrizes $T_{22}, T_{33}, \dots, T_{cc}$, são quadradas, irredutíveis e, da mesma forma, corresponderiam aos estados recorrentes. Todas as matrizes redutíveis nesta tese, são supostas estarem nesta forma canônica.

O Teorema de Perron-Frobenius para matrizes não-negativas redutíveis é apresentado a seguir.

Teorema B.7 *Seja $P_{m \times m}$ uma matriz não-negativa, redutível escrita na forma canônica (B.3), com auto-valor maximal λ_{max} . Então, $\lim_{n \rightarrow \infty} \left(\frac{P}{\lambda_{max}} \right)^n$ existe se, somente se as submatrizes $T_{22}, T_{33}, \dots, T_{cc}$ são cada uma primitiva, e este limite é dado por*

$$\lim_{n \rightarrow \infty} \left(\frac{P}{\lambda_{max}} \right)^n = \begin{bmatrix} \mathbf{0} & (I - T_{11})^{-1} \cdot T_{12} \cdot \left(\frac{\lambda_{1,2}}{\lambda_{max}} \right)^n \cdot \phi_2 \cdot \psi'_2 & \dots & (I - T_{11})^{-1} \cdot T_{1c} \cdot \left(\frac{\lambda_{1,c}}{\lambda_{max}} \right)^n \cdot \phi_c \cdot \psi'_c \\ \mathbf{0} & \left(\frac{\lambda_{1,2}}{\lambda_{max}} \right)^n \cdot \phi_2 \cdot \psi'_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \left(\frac{\lambda_{1,c}}{\lambda_{max}} \right)^n \cdot \phi_c \cdot \psi'_c \end{bmatrix},$$

onde ϕ_j e ψ'_j , $j = \{2, 3, \dots, c\}$ são os auto-vetores de Perron de cada uma das submatrizes de $T_{22}, T_{33}, \dots, T_{cc}$, que possuem auto-valor maximal $\lambda_{1,j}$.