

Universidade Federal do Rio Grande do Sul
Instituto de Matemática
Programa de Pós-Graduação em Matemática

Estimativa da Taxa de Convergência de uma Dinâmica de Glauber

por

Vilarbo da Silva Júnior

Prof. Dr. Artur Oscar Lopes

Orientador

Porto Alegre, Junho de 2007

Dissertação submetida por Vilarbo da Silva Júnior¹ como requisito parcial para a obtenção do grau de Mestre em Matemática pelo Programa de Pós-Graduação em Matemática do Instituto de Matemática da Universidade Federal do Rio Grande do Sul.

Professor Orientador: Dr. Artur Oscar Lopes

Banca Examinadora:

Dr. Artur Oscar Lopes

Dr. Alexandre Tavares Baraviera

Dr. Eduardo E. Brietzke

Dr. Leonardo Fernandes Guidi

Data de Defesa: 18 de junho de 2007.

¹Bolsista do Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

Resumo: Este trabalho aborda um estudo, a nível introdutório, sobre *processos de Markov reversíveis*, onde seguimos a exposição de Stroock [24]. Mais especificamente, estudamos cadeias de Markov sobre S (espaço de estados) finito que sejam reversíveis com respeito a uma dada distribuição inicial μ . Analisamos também a validade do limite $\lim_{t \rightarrow \infty} \|\mathcal{P}^t f - \langle f \rangle_\mu\|_{2,\mu} = 0$, nos preocupando, principalmente, com a taxa de tal convergência. Por fim, propomos uma *Dinâmica de Glauber* para o *Estado de Gibbs* $\mu(\beta)$ e damos uma cota superior e inferior para a taxa de convergência λ_β do limite $\lim_{t \rightarrow \infty} \|\mathcal{P}^t f - \langle f \rangle_{\mu(\beta)}\|_{2,\mu(\beta)} = 0$. Concluindo, apresentamos aplicações, em modelos físicos, da teoria desenvolvida ao longo desta dissertação.

Abstract: This work approach a study, in introductory level, about reversible Markov processes, where we follow the Stroock's exposition [24]. More specifically, we study Markov chains about S (space of states) finite that is reversible with respect a initial distribution μ give. Too we analysis the validate of limit $\lim_{t \rightarrow \infty} \|\mathcal{P}^t f - \langle f \rangle_\mu\|_{2,\mu} = 0$, preoccupying, principally, with the rate of convergence. Finish, we propose a *Glauber Dynamics* to *Gibbs States* $\mu(\beta)$ and we give a lower and upper boundary to rate of convergence λ_β of the limit $\lim_{t \rightarrow \infty} \|\mathcal{P}^t f - \langle f \rangle_{\mu(\beta)}\|_{2,\mu(\beta)} = 0$. Concluding, we presents applications, in physical models, of the developed theory in the course these dissertation.

Agradecimentos

Agradeço inicialmente a meu orientador Dr. Artur Oscar Lopes, pelo ótimo convívio durante os encontros nestes dois anos, não somente por aqueles em que discutimos e dirigimos os estudos desta dissertação, como também aqueles em sala de aula nas diversas disciplinas que tive o prazer de tê-lo como ministrante. Agradeço, principalmente, pelas contribuições técnicas e intelectuais que enriqueceram esta dissertação.

Quero agradecer aos demais professores da banca Alexandre Tavares Baraviera, Eduardo Brietzke e Leonardo Guidi, que não somente aceitaram gentilmente ler esta dissertação, como também participaram dos seminários realizados por mim, colaborando com críticas e sugestões, o que foi de fundamental importância.

Durante minha estadia no curso fiz verdadeiros amigos e companheiros. Destaco alguns destes e lhes deixo meu profundo agradecimento.

Professor Eduardo Brietzke que, mesmo sem me conhecer gentilmente recebi-me em sua sala para discutirmos Análise real.

Professor Leonardo Bonorino, pelas ótimas aulas de Análise I e II que foram muito importantes para minha formação básica como matemático.

Adriana Neumann e Cínthya Schneider foram colegas muito especiais com as quais pude trocar experiências e conhecimentos nas diversas disciplinas que cursamos juntos. Meu colega da época da graduação Alexsandro M. Carvalho, por incessantemente me incentivar a nunca desistir por mais difícil que pareça a vida.

Deixo aqui um agradecimento especial a minha namorada, Lidiane de Matos Kerschner que já me acompanha há muitos anos e que é minha razão por estar sempre batalhando. Por fim, agradeço a meus familiares pelo apóio e compreensão.

Sumário

Agradecimentos	1
Introdução	3
1 Elementos de Processos Estocásticos	6
1.1 Introdução	6
1.2 Teoria da Medida	6
1.3 Processos Estocásticos	8
1.3.1 Processos de Markov a Tempo Discreto	9
1.3.2 Processos de Markov a Tempo Contínuo	13
2 Processos de Markov Reversíveis e Forma de Dirichlet-Tempo Discreto	16
2.1 Introdução	16
2.2 Reversibilidade	16
2.2.1 Conceitos Gerais	16
2.3 Convergência em $L^2(\mu)$ - Tempo Discreto	19
2.3.1 O Gap Espectral	23
2.3.2 Mais Sobre a Notação $\mathcal{P}f$	28
2.4 A Forma de Dirichlet	29
2.4.1 Forma de Dirichlet e Desigualdade de Poincaré	29
2.4.2 Estimando β_{\pm}	36
3 Processos de Markov Reversíveis e Forma de Dirichlet-Tempo Contínuo	43
3.1 Introdução	43
3.2 Reversibilidade	43
3.3 Convergência em $L^2(\mu)$ - Tempo Contínuo	46
3.3.1 Estimando λ	49
4 Elementos de Mecânica Estatística	51
4.1 Introdução	51
4.2 Postulados da Mecânica Estatística de Equilíbrio	51
4.2.1 Parâmetros Intensivos da Termodinâmica	53
4.3 Ensemble Canônico	54
4.3.1 Valor Esperado e Variância	56
4.4 A Equação Mestra	56
5 Dinâmica de Glauber	58
5.1 Introdução	58
5.2 Construção da Dinâmica de Glauber	58
5.3 Convergência em $L^2(\mu(\beta))$ da Dinâmica de Glauber	61
6 Aplicações	66
6.1 Introdução	66

6.2 Sistema de Partículas com Dois Níveis de Energia .	66
6.2.1 Modelagem	66
6.2.2 Construção das Probabilidades de Transição	66
6.3 Modelo de Núcleos Magnéticos Localizados	70
6.3.1 Modelagem	70
6.3.2 Construção das Probabilidades de Transição	70
Referências Bibliográficas	75

Introdução

Quando Markov introduziu seu modelo em 1906, não havia se preocupado com as aplicações. De fato, ele o introduzira para mostrar que independência não é necessário para a Lei (fraca) dos Grandes Números. Hoje em dia existem inúmeras aplicações, podemos citar algumas como por exemplo: *fenômenos biológicos*, teoria de populações e genética; *ciências sociais*, a mobilidade social pode ser descrita em termos markovianos; *engenharia elétrica*, processamento de imagens, comunicações de redes; a *física* é, provavelmente, o maior patrono da teoria das cadeias de Markov, modelos de Markov têm sido usados no entendimento qualitativo de fenômenos complexos (o modelo da difusão de Ehrenfest, o modelo de Ising-Peierls de transição de fase).

O presente trabalho apresenta um estudo sobre uma classe especial de Processos de Markov, denominados *Processos de Markov Reversíveis* (veja definição 2.2.1). Seguimos, na maior parte do tempo, a exposição dada por Stroock [24].

Mais especificamente, vamos desenvolver a teoria geral sobre Processos Estocásticos de Markov Reversíveis com respeito a uma dada distribuição inicial de probabilidades μ , sobre S (Espaço de Estados, ver definição 1.3.4) **finito**, com o intento de aplicar tal teoria ao caso particular de uma *Dinâmica de Glauber* (veja definição 5.1.1).

Feito isto, temos como objetivo específico dar uma cota superior e inferior para a taxa (velocidade) de convergência do limite $\lim_{t \rightarrow \infty} \|\mathcal{P}^t f - \langle f \rangle_\mu\|_{2,\mu} = 0$, onde $f \in L^2(\mu) := \{f : S \rightarrow \mathbb{R} : \|f\|_{2,\mu} < \infty\}$ e $\mathcal{P}^t f$ está definida em 2.10. Tal fato está demonstrado no teorema 5.3.1, cujo cerne para demonstração é a importante definição de *Forma de Dirichlet* (ver definição 2.4.4).

Nesta dissertação não temos a ambição de sermos originais nos assuntos aqui desenvolvidos (exceto pelo capítulo 6 que, conforme nossa particular pesquisa bibliográfica, não detectamos nada do gênero), de fato, nossa contribuição neste trabalho reside apenas na organização dos assuntos, escolha adequada de exemplos, no estilo próprio de comentar os resultados e no detalhamento das demonstrações apresentadas (sendo que algumas destas não foram indicadas por Stroock [24]).

A escolha do tema deste trabalho deve-se a dois motivos principais. Primeiramente, procuramos algo que relacionasse Processos de Markov e Mecânica Estatística. Num segundo momento, preocupamo-nos em elaborar um material para que o leitor não familiarizado com processos reversíveis, e tão pouco com a Dinâmica de Glauber, possa aqui encontrar um material introdutório e com alto grau de detalhe nas demonstrações e exemplos de tais assuntos (de fato, tivemos muitas dificuldades na aquisição de referências sobre tais assuntos).

No capítulo 1, apresentamos os rudimentos sobre processos estocásticos com o intento de relembrar ao leitor conceitos e resultados relevantes que serão utilizados incessantemente nos capítulos seguintes, e também fixar as notações empregados doravante, cuja escolha dá-se, simplesmente, pelo particular gosto.

Iniciamos o capítulo 2 com a importante definição de Processos de Markov Reversíveis (ver definição 2.2.1) e também apresentamos uma condição necessária e suficiente para que um processo de Markov seja reversível com respeito a uma dada distribuição inicial μ , a saber, a *Condição do Balanço Detalhado* (ver proposição 2.2.3) (de fato, quando falarmos sobre processos de Markov Reversíveis, devemos pensar na condição 2.4).

Neste capítulo nos concentramos nos processos de Markov a tempo discreto, dando enfoque à teoria geral sobre a convergência de $\mathcal{P}^n f$ (conforme 2.10) ao valor esperado $\langle f \rangle_\mu$ quando $n \rightarrow \infty$, objetivando dar uma estimativa para a taxa de convergência deste limite, apresentamos também a definição de *Forma de Dirichlet* (ver definição 2.4.4). Na subseção 2.4.2 procuramos dar uma estimativa para o *Gap Spectral* (veja proposição 2.3.7) em termos da geometria dos *Grafos de Transição* (que será fundamental para o que segue no capítulo 3).

Grosso modo estenderemos, no capítulo 3, os resultados e definições vistas no capítulo 2 (tempo discreto) ao caso de Processos de Markov Reversíveis a tempo contínuo. O que faremos aqui é, simplesmente, contextualizar e identificar analogias entre os dois casos (tempo discreto e tempo contínuo). Este capítulo apresenta as ferramentas necessárias para a demonstração do principal teorema deste trabalho (teorema 5.3.1).

Apresentamos no capítulo 4, sucintamente, os rudimentos da Mecânica Estatística de Equilíbrio (de um ponto de vista físico, onde seguimos [20] e [22]). Aqui, vamos nos preocupar em descrever o chamado *Ensemble Canônico* (ver definição 4.3.4), pois é nele que se alicerça a construção da *Dinâmica de Glauber*. Damos, na seção 4.3, um esboço da dedução da distribuição de probabilidades $\mu(\beta)$ conhecida como *Estado de Gibbs* (ver proposição 4.3.2). Finalizamos este capítulo construindo, de forma heurística, uma *Equação Mestra* (veja seção 4.4) para $(\mu(\beta))_i$, i.e., uma equação que descreve a dinâmica da evolução temporal da coordenada i de $\mu(\beta)$. Com tal Equação Mestra, damos uma justificativa física para a Condição do Balanço Detalhado 2.4, que caracteriza os processos reversíveis!

Propomos, no capítulo 5, a construção de uma Dinâmica de Glauber para o Estado de Gibbs. Além disso, utilizamos os resultados oriundos dos capítulos 2 e 3, para provar o teorema 5.3.1 que nos dá uma cota superior e inferior para a taxa de convergência, λ_β , com que $\|\mathcal{P}^t f - \langle f \rangle_{\mu(\beta)}\|_{2,\mu(\beta)} \rightarrow 0$ quando $t \rightarrow \infty$, onde \mathcal{P}^t é a Dinâmica de Glauber descrita na seção 5.2.

Finalmente, explicitamos no capítulo 6 uma Dinâmica de Glauber (como proposta na seção 5.2) para dois particulares sistemas físicos, que são, *Sistema de Partículas com Dois Níveis de Energia* e o *Modelo de Núcleos Magnéticos Localizados*.

Interpretamos alguns casos limites de \mathcal{P}^t (matriz de transição de probabilidades a tempo contínuo) para cada sistema em estudo. Por fim, aplicamos alguns resultados dos capítulos anteriores para certas variáveis aleatórias (o hamiltoniano \mathcal{H} no primeiro exemplo e a variável de spin σ no segundo), mostrando com isto que a Dinâmica de Glauber \mathcal{P}^t , agindo sobre estas variáveis aleatórias (no sentido da relação 2.10), converge ao valor esperado das mesmas e, além disso, estimamos a taxa de tal convergência.

1 Elementos de Processos Estocásticos

1.1 Introdução

Neste capítulo veremos alguns resultados dos *processos estocásticos* e demais conteúdos afins que são relevantes para o bom entendimento do que segue nos capítulos seguintes. O leitor que já tenha feito um curso introdutório de processos estocásticos poderá, perfeitamente, omitir a leitura deste capítulo, mas recomendamos sua leitura para fixar as notações. Vale ressaltar que, como mencionado anteriormente, o propósito deste capítulo é apenas resumir informações necessárias para o entendimento dos capítulos restantes desse trabalho. Sendo assim, a maior parte dos resultados serão dados sem indicação de suas provas. Para um aprofundamento e verificação das provas dos teoremas e proposições, indicamos as seguintes referências: [3], [6], [9], [14], [17].

1.2 Teoria da Medida

Começamos simplesmente definindo o importantíssimo conceito de σ -Álgebra.

Definição 1.2.1

Seja Ω um conjunto não-vazio e $\mathcal{A} \subseteq P(\Omega)$ uma família não-vazia de subconjuntos de Ω , onde $P(\Omega)$ denota o conjunto das partes de Ω (i.e., formado por todos subconjuntos de Ω). Dizemos que \mathcal{A} é uma σ -Álgebra sobre Ω se,

1. $\Omega \in \mathcal{A}$;
2. $A \in \mathcal{A} \Rightarrow A^c \equiv \Omega \setminus A \in \mathcal{A}$;
3. $\{A_n\}_{n=1}^{\infty} \subseteq \mathcal{A} \Rightarrow \cup_{n=1}^{\infty} A_n \in \mathcal{A}$.

◇

Observação 1.2.2

Note que, se \mathcal{A} é uma σ -Álgebra vem, pelas propriedades elementares de manipulações com conjuntos, que

1. $A, B \in \mathcal{A} \Rightarrow A \cap B \in \mathcal{A}$;
2. $A, B \in \mathcal{A} \Rightarrow A \setminus B \in \mathcal{A}$;
3. $\{A_n\}_{n=1}^{\infty} \subseteq \mathcal{A} \Rightarrow \cap_{n=1}^{\infty} A_n \in \mathcal{A}$.

◇

Observação 1.2.3

Se a união, na definição 1.2.1, for apenas finita, dizemos então que \mathcal{A} é uma *Álgebra* de subconjuntos de Ω .

◇

Definição 1.2.4

Dado $\Upsilon \subset P(\Omega)$, a σ -*Álgebra Gerada* por Υ é simplesmente a intersecção de todas as σ -Álgebras de Ω contendo Υ . Denotamos por $\sigma(\Upsilon)$ tal intersecção.

◇

Um dos conceitos mais relevantes, utilizado com frequência neste trabalho, é o de *Probabilidade*, que será dado na definição abaixo.

Definição 1.2.5

Uma função de conjunto $\mathbb{P} : \mathcal{A} \rightarrow [0, 1]$, onde \mathcal{A} é dada como na definição 1.2.1, é dita uma *Probabilidade* se,

1. $\mathbb{P}(\Omega) = 1$;
2. \mathbb{P} é σ -aditiva, isto é, dado uma seqüência de conjuntos $A_n \in \mathcal{A}$, dois a dois disjuntos, tal que sua união $A = \sum_{n=1}^{\infty} A_n = \bigsqcup_{n=1}^{\infty} A_n$ também pertença a \mathcal{A} , então $\mathbb{P}(A) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$.

◇

Definição 1.2.6

Fixado uma probabilidade sobre uma σ -Álgebra \mathcal{A} , a *Probabilidade Condicional* de ocorrer o evento $A \in \mathcal{A}$, dado que ocorreu o evento $B \in \mathcal{A}$, é o número

$$\mathbb{P}(A|B) := \frac{\mathbb{P}(A, B)}{\mathbb{P}(B)}, \quad (1.1)$$

quando $\mathbb{P}(B) > 0$, aqui a vírgula entre A e B denota a intersecção entre os eventos.

◇

Teorema 1.2.7 *Regra de Bayes*

(**Versão 1**) Se $\mathbb{P}(A) > 0$, com $A \in \mathcal{A}$, então

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A|B)\mathbb{P}(B)}{\mathbb{P}(A)}. \quad (1.2)$$

(**Versão 2**) Sejam B_1, B_2, \dots elementos da σ -Álgebra \mathcal{A} onde \mathbb{P} está definida, tais que

$$\sum_{i=1}^{\infty} B_i = \Omega,$$

então, $\forall A \in \mathcal{A}$, vale

$$\mathbb{P}(A) = \sum_{i=1}^{\infty} \mathbb{P}(A|B_i)\mathbb{P}(B_i). \quad (1.3)$$

(**Versão 3**) Para qualquer seqüência de eventos $A_1, \dots, A_n \in \mathcal{A}$, temos

$$\mathbb{P}(\cap_{i=1}^n A_i) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1 \cap A_2)\dots\mathbb{P}(A_n|\cap_{i=1}^{n-1} A_i). \quad (1.4)$$

◇

Definição 1.2.8

Um *Espaço de Probabilidade* é um terno $(\Omega, \mathcal{A}, \mathbb{P})$, onde Ω é um conjunto, \mathcal{A} é uma σ -Álgebra de subconjuntos de Ω , e \mathbb{P} é uma probabilidade sobre \mathcal{A} (como na definição 1.2.5). Denomina-se *Espaço Mensurável* o par (Ω, \mathcal{A}) .

◇

Teorema 1.2.9 (Teorema de Extensão de Hahn-Kolmogorov)

Seja Ω um conjunto, \mathcal{A}_0 uma Álgebra de subconjuntos de Ω e $\mathbb{P}_0 : \mathcal{A}_0 \rightarrow [0, 1]$ uma função de conjunto aditiva. Então, se \mathcal{A} é a σ -Álgebra gerada por \mathcal{A}_0 , existe uma só probabilidade $\mathbb{P} : \mathcal{A} \rightarrow [0, 1]$ tal que $\mathbb{P}|_{\mathcal{A}_0} = \mathbb{P}_0$.

◇

Definição 1.2.10

Uma função $X : (\Omega, \mathcal{A}) \rightarrow (S, \mathcal{G})$, onde (Ω, \mathcal{A}) e (S, \mathcal{G}) são espaços mensuráveis, é dita *Mensurável* se $X^{-1}(A) \in \mathcal{A}$ para todo $A \in \mathcal{G}$. A função mensurável X é usualmente chamada de *Variável Aleatória*.

◇

1.3 Processos Estocásticos

Definição 1.3.1

Considere $(\Omega, \mathcal{A}, \mathbb{P})$ um espaço de probabilidade, (S, \mathcal{G}) um espaço mensurável e ainda uma família de variáveis aleatórias X_t , indexadas por um parâmetro $t \in T$, onde $T \subset \mathbb{R}$, e onde cada $X_t : (\Omega, \mathcal{A}) \rightarrow (S, \mathcal{G})$ é mensurável. Dizemos que tal $\{X_t\}_{t \in T}$ é um *Processo Estocástico*. Trataremos quase que na totalidade deste texto S finito ou enumerável, e assim $\mathcal{G} = P(S)$.

◇

Definição 1.3.2

O conjunto $T \neq \emptyset$, contido em \mathbb{R} , é denominado *Espaço de Parâmetros* ou *Índices do Processo*. O conjunto T possui uma ordem e vamos pensar que, para cada $t \in T$, a variável X_t descreve o que acontece com o processo no tempo t . Destacamos dois casos:

Parâmetro Discreto: Quando $T = \mathbb{N}$, ou \mathbb{Z} , ou ainda $\{1, 2, \dots, n\}$.

Parâmetro Contínuo: Quando $T = [a, b]$, ou $T = \mathbb{R}^+$ ou $T = \mathbb{R}$.

◇

Observação 1.3.3

Quanto o processo estocástico em questão for de tempo contínuo usaremos a letra t para os elementos de T , por outro lado, quando o tempo for discreto usaremos a letra n para nos referirmos aos elementos de T .

◇

Definição 1.3.4

O *Espaço de Estados* é o conjunto S , ou seja, o elenco dos possíveis valores de cada variável aleatória X_t .

◇

Definição 1.3.5

Uma matriz $\#S$ por $\#S$

$$\mathcal{P} = (P)_{ij}, \quad (1.5)$$

onde $(i, j) \in S^2$ é dita *Estocástica* (ou *Linha Estocástica*) se, para cada i fixo, vale que $\sum_{j \in S} (P)_{ij} = 1$, e ainda $(P)_{ij} \geq 0$.

◇

Definição 1.3.6

Um vetor (linha) $\mu = ((\mu)_i)_{i \in S}$ é dito um *Vetor de Probabilidade* sobre S se cada entrada $(\mu)_i \geq 0$ e $\sum_{i \in S} (\mu)_i = 1$.

$\mu = ((\mu)_i)_{i \in S}$ é dito *Vetor de Probabilidade Estacionário* (ou uma *Distribuição Estacionária*) para a matriz \mathcal{P} , dada por 1.5, se $\mu\mathcal{P} = \mu$.

◇

1.3.1 Processos de Markov a Tempo Discreto**Definição 1.3.7**

Um *Processo Estocástico de Markov* (com tempo discreto) em um espaço de estados S (finito ou enumerável) é uma família de variáveis aleatórias $\{X_n\}_{n \geq 0}$, tomando valores em S , com a propriedade que, para todo $n \geq 0$ e $(i_0, \dots, i_n, j) \in S^{n+2}$, tenhamos

$$\mathbb{P}(X_{n+1} = j | X_0 = i_0, \dots, X_n = i_n) = \mathbb{P}(X_{n+1} = j | X_n = i_n), \quad (1.6)$$

sempre que $\mathbb{P}(X_0 = i_0, \dots, X_n = i_n) > 0$.

◇

Observação 1.3.8

A notação utilizada na definição 1.3.7, $\mathbb{P}(A|B)$, significa a *Probabilidade Condicional* de ocorrer o evento A dado que ocorreu o evento B , conforme definição 1.2.6.

◇

Definição 1.3.9

Define-se *Probabilidade de Transição* a probabilidade $P^{n,n+1}(i, j) = \mathbb{P}(X_{n+1} = j | X_n = i)$ do processo estar no estado j no tempo $n + 1$, dado que o processo está no estado i no tempo n .

◇

Definição 1.3.10

Seja $\{X_n\}_{n \geq 0}$ um Processo de Markov, dizemos que ele tem *Probabilidade de Transição Estacionária* se, e somente se, existe uma matriz $\mathcal{P} = (P)_{ij}$ do tipo $\#S$ por $\#S$, tal que, $\forall n \geq 0$ e $\forall (i, j) \in S^2$, se tenha

$$\mathbb{P}(X_{n+1} = j | X_n = i) = \mathbb{P}(X_1 = j | X_0 = i) = (P)_{ij}. \quad (1.7)$$

A matriz $\mathcal{P} = (P)_{ij}$, definida em 1.7, é chamada de *Matriz de Transição* sobre S associada ao processo.

◇

Observação 1.3.11

A matriz \mathcal{P} , mencionada na definição 1.3.10, é uma *Matriz Estocástica*, no sentido da definição 1.3.5.

◇

Definição 1.3.12

Se um vetor de probabilidade (definição 1.3.6) $\mu = ((\mu)_i)_{i \in S}$ tem a i -ésima entrada dada por $(\mu)_i = \mathbb{P}(X_0 = i)$, então μ é chamado a *Distribuição Inicial* do Processo Estocástico (algumas vezes denotado por $\mu^0 = ((\mu^0)_i)_{i \in S}$, para enfatizar o tempo inicial zero). Além disso, $\forall n \geq 0$ e $\forall i \in S$, vale que

$$(\mu \mathcal{P}^n)_i = \mathbb{P}(X_n = i), \quad (1.8)$$

onde, por convenção, $\mathcal{P}^0 = I$, $\mathcal{P}^n = \mathcal{P} \mathcal{P}^{n-1} \quad \forall n \geq 1$, e \mathcal{P} é a *matriz de transição*. Uma distribuição inicial $\mu = ((\mu)_i)_{i \in S}$ tal que $\mu \mathcal{P} = \mu$ é dita uma *Distribuição Inicial Estacionária* para a matriz estocástica \mathcal{P} .

◇

Teorema 1.3.13

Seja $\{X_n\}_{n \geq 0}$ um processo estocástico de Markov com probabilidade de transição estacionária. Então, dada a matriz de transição $\mathcal{P} = (P)_{ij}$, $(i, j) \in S^2$ (associada ao processo conforme a definição 1.3.10) e uma distribuição inicial $\mu = ((\mu)_i)_{i \in S}$, o processo fica determinado completamente, no sentido que é possível determinarmos a probabilidade $\mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n)$ (conhecida como *Distribuição Conjunta*). De fato,

$$\mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = (\mu)_{i_0} (P)_{i_0 i_1} (P)_{i_1 i_2} \dots (P)_{i_{n-1} i_n}. \quad (1.9)$$

◇

Definição 1.3.14

Uma *Cadeia de Markov* (com estados em S finito ou enumerável) associada à matriz estocástica \mathcal{P} fixada, tipo $\#S$ por $\#S$, é a classe de todos os processos estocásticos de Markov que podemos obter, a partir de todas as distribuições iniciais $\mu = ((\mu)_i)_{i \in S}$.

◇

Observação 1.3.15

Os elementos da matriz \mathcal{P}^n serão denotados por $(P^n)_{ij}, \forall (i, j) \in S^2$.

◇

Definição 1.3.16

Seja $\mathcal{P} = (P)_{ij}$ uma matriz de transição com $(i, j) \in S^2$. Então,

1. Se $(P^n)_{ij} > 0$, para algum $n \geq 0$, dizemos que j é *Acessível* a i , ou que i *Conduz* a j e escrevemos $i \rightarrow j$.
2. Se $i \rightarrow j$ e $j \rightarrow i$ dizemos que i e j se *Comunicam* e escrevemos $i \leftrightarrow j$. (i.e., $i \leftrightarrow j \Leftrightarrow \exists n_1, n_2 \geq 0$ tal que $(P^{n_1})_{ij}, (P^{n_2})_{ji} > 0$).

◇

Proposição 1.3.17

A relação \leftrightarrow acima, é uma relação de equivalência. Além disso, \leftrightarrow divide S em *Classes de Equivalência*, ou seja, se existirem m classes de equivalências $C_i \subset S$, com $i \in \{1, \dots, m\}$, para a relação de comunicação \leftrightarrow , então $S = \bigcup_{i=1}^m C_i$.

◇

Definição 1.3.18

Se existe somente uma classe de equivalência para a relação de comunicação \leftrightarrow , a cadeia é dita *Irreduzível*. Diremos também que \mathcal{P} é uma *Matriz Irreduzível*.

◇

Teorema 1.3.19 *Teorema de Perron-Frobenius*

Seja \mathbf{A} uma matriz $N \times N$ irreduzível. Então, \mathbf{A} possui um autovalor λ_1 (denominado o autovalor dominante de \mathbf{A}) tal que:

1. Se $\lambda_2, \lambda_3, \dots, \lambda_N$ são os outros autovalores de \mathbf{A} , então podemos ordená-los em relação a λ_1 do seguinte modo

$$\lambda_1 > |\lambda_2| \geq \dots \geq |\lambda_N|;$$

2. $\dim(\mathbf{A} - \lambda_1 \mathbf{I})^{-1}(0) = 1$;

3. O autovetor a esquerda u_1 e o autovetor a direita v_1 , associados a λ_1 , podem ser escolhidos positivos e tais que $\langle u_1, v_1 \rangle_{\mathbb{R}^N} = 1$.

◇

Corolário 1.3.20

Se, além das hipóteses anteriores, tivermos \mathbf{A} (linha) estocástica, então

1. $\lambda_1 = 1$;
2. Existe um único $v = ((v)_1, \dots, (v)_N)$ tal que $v\mathbf{A} = v$ e $\sum_i (v)_i = 1$;
3. Se $x = ((x)_1, \dots, (x)_N)$ satisfaz $\sum_i (x)_i = 1$, então

$$\lim_{n \rightarrow \infty} x\mathbf{A}^n = v.$$

◇

Proposição 1.3.21 *Desigualdade de Schwarz*

Para quaisquer seqüências reais $\{a_n\}_{n \in \mathbb{Z}}$ e $\{b_n\}_{n \in \mathbb{Z}}$ tem-se

$$\sum_{n \in \mathbb{Z}} |a_n b_n| \leq \sqrt{\sum_{n \in \mathbb{Z}} a_n^2} \sqrt{\sum_{n \in \mathbb{Z}} b_n^2}. \quad (1.10)$$

◇

Teorema 1.3.22

Seja \mathcal{P} matriz estocástica irredutível sobre S finito. Então:

1. \mathcal{P} tem um único vetor de probabilidade estacionário μ e as componentes de μ são todas positivas, i.e., μ é o único vetor de probabilidade tal que $\mu\mathcal{P} = \mu$, $\sum_{i \in S} (\mu)_i = 1$ com $(\mu)_i > 0 \forall i \in S$;
2. Se p é qualquer vetor de probabilidade, então a seqüência de vetores $p\mathcal{P}, p\mathcal{P}^2, \dots$ converge à distribuição estacionária μ , i.e.,

$$\lim_{n \rightarrow \infty} p\mathcal{P}^n = \mu;$$

3. As entradas das matrizes $\mathcal{P}, \mathcal{P}^2, \dots$ obtidas a partir de \mathcal{P} convergem para as entradas da matriz \mathcal{M} cujas linhas são todas iguais ao vetor estacionário μ .
Em outras palavras

$$\lim_{n \rightarrow \infty} \mathcal{P}^n = \mathcal{M} = \begin{bmatrix} (\mu)_1 & (\mu)_2 & \dots & (\mu)_N \\ (\mu)_1 & (\mu)_2 & \dots & (\mu)_N \\ \vdots & \vdots & \vdots & \vdots \\ (\mu)_1 & (\mu)_2 & \dots & (\mu)_N \end{bmatrix}.$$

◇

Proposição 1.3.23 *Desigualdade de Jensen*

Seja $\psi : [0, \infty) \rightarrow [0, \infty)$ uma função não-decrescente, contínua e convexa, $\mathbf{g} = (g(1), g(2), \dots, g(N))$ o vetor determinado por $g : S \rightarrow [0, \infty)$ (onde $S = \{1, 2, \dots, N\}$) e $(\alpha_1, \alpha_2, \dots, \alpha_N) \in [0, 1]^N$ com a propriedade que $\sum_{j=1}^N \alpha_j = 1$. Então,

$$\psi \left(\sum_{j=1}^N \alpha_j g(j) \right) \leq \sum_{j=1}^N \alpha_j \psi(g(j)). \quad (1.11)$$

◇

1.3.2 Processos de Markov a Tempo Contínuo

De forma análoga à definição 1.3.7 definimos *Processo de Markov a Tempo Contínuo* como segue abaixo.

Definição 1.3.24

Um *Processo Estocástico de Markov* (com tempo contínuo) em um espaço de estados S (finito ou enumerável) é uma família de variáveis aleatórias $\{X_t\}_{t \geq 0}$, tomando valores em S , com a propriedade que, para todo $t \geq 0$ e $(i_0, \dots, i_n, j) \in S^{n+2}$, tenhamos

$$\mathbb{P}(X_t = j | X_0 = i_0, X_{t_1} = i_1, \dots, X_{t_n} = i_n) = \mathbb{P}(X_t = j | X_{t_n} = i_n), \quad (1.12)$$

sempre que $0 < t_1 < t_2 < \dots < t_n < t$ e $\mathbb{P}(X_0 = i_0, X_{t_1} = i_1, \dots, X_{t_n} = i_n) > 0$.

◇

Pelo que vimos no caso do tempo discreto, na definição 1.3.12 e no teorema 1.3.13, a probabilidade \mathbb{P} sobre $S^{\mathbb{N}}$, associada a um processo Markoviano, fica determinada a partir da matriz estocástica \mathcal{P} e uma distribuição inicial μ sobre S . Em outras palavras, \mathbb{P} fica determinada pela família de matrizes \mathcal{P}^n , que satisfazem a propriedade que $\forall n, m \in \mathbb{N}$, $\mathcal{P}^n \mathcal{P}^m = \mathcal{P}^{n+m}$, e uma distribuição inicial μ sobre S .

Motivados pelo parágrafo anterior, queremos saber se existe uma família de matrizes estocásticas \mathcal{P}^t , $t \in \mathbb{R}^+$, do tipo $\#S$ por $\#S$ tal que

1. para todo $t, s \in \mathbb{R}^+$, vale que $\mathcal{P}^t \mathcal{P}^s = \mathcal{P}^{t+s}$,
2. para cada $t \in \mathbb{R}^+$, a matriz \mathcal{P}^t é estocástica,
3. $\mathcal{P}^0 = I$,

e tal que, a partir de qualquer vetor de probabilidade μ possamos determinar, completamente, a probabilidade \mathbb{P} .

A resposta para nossa dúvida é afirmativa e, todas as questões envolvendo a demonstração destes fatos despenderiam algumas páginas; como este não é nosso objetivo, destacaremos apenas os resultados mais relevantes.

Definição 1.3.25

Seja S um conjunto finito ou enumerável. Uma Q -matriz em S é uma matriz $Q = (Q)_{ij}$, para $(i, j) \in S^2$, satisfazendo as seguintes condições:

1. $0 \leq -(Q)_{ii} < \infty \forall i \in S$,
2. $(Q)_{ij} \geq 0 \forall i \neq j, (i, j) \in S^2$,
3. $\sum_{j \in S} (Q)_{ij} = 0 \forall i \in S$.

A Q -matriz é, muitas vezes, chamada de matriz tipo *Linha Soma Zero*, em jus à propriedade 3.

◇

Vamos agora expor a conexão existente entre a Q -matriz e o *semi-grupo a tempo contínuo* $\mathcal{P}^t, t \in \mathbb{R}^+$, sobre S .

Lema 1.3.26

Sejam Q_1 e Q_2 matrizes do tipo $\#S$ por $\#S$, se $Q_1 Q_2 = Q_2 Q_1$, então $e^{Q_1 + Q_2} = e^{Q_1} e^{Q_2} = e^{Q_2} e^{Q_1}$.

◇

Teorema 1.3.27

Seja Q uma matriz em um conjunto S finito. Defina $\mathcal{P}^t = e^{tQ}$ (onde e^{tQ} denota o limite $\sum_{k=0}^{\infty} t^k Q^k / k!$). Então, \mathcal{P}^t para $t \in \mathbb{R}^+$ tem as seguintes propriedades:

1. $\mathcal{P}^{(t+s)} = \mathcal{P}^t \mathcal{P}^s \forall t, s$ (propriedade de semi-grupo),
2. $\mathcal{P}^0 = I$,
3. \mathcal{P}^t para $t \in \mathbb{R}^+$ é a única solução da *forward equation*

$$\frac{d}{dt} \mathcal{P}^t = \mathcal{P}^t Q; \quad \mathcal{P}^0 = I,$$
4. \mathcal{P}^t para $t \in \mathbb{R}^+$ é a única solução da *backward equation*

$$\frac{d}{dt} \mathcal{P}^t = Q \mathcal{P}^t; \quad \mathcal{P}^0 = I,$$
5. para $k = 0, 1, 2, \dots$ temos
$$\frac{d^k}{dt^k} \mathcal{P}^t |_{t=0} = Q^k.$$

◇

O próximo teorema fecha as propriedades requeridas sobre \mathcal{P}^t .

Teorema 1.3.28

Uma matriz Q é uma Q -matriz se, e somente se, $\mathcal{P}^t = e^{tQ}$ é uma matriz estocástica, para todo $t \geq 0$.

◇

Fixado um semi-grupo a parâmetro contínuo \mathcal{P}^t , e uma distribuição inicial μ , existe uma expressão análoga a dada no teorema 1.3.13, a saber:

$$\mathbb{P}(X_0 = i_0, X_{t_1} = i_1, X_{t_2} = i_2, \dots, X_{t_n} = i_n) = (\mu)_{i_0} (P^{t_1})_{i_0 i_1} (P^{t_2 - t_1})_{i_1 i_2} \dots (P^{t_n - t_{n-1}})_{i_{n-1} i_n}. \quad (1.13)$$

Em resumo, os teoremas 1.3.27 e 1.3.28, juntamente com o parágrafo anterior, nos afirmam que se \mathcal{Q} é uma \mathcal{Q} -matriz, então o semi-grupo a tempo contínuo $\mathcal{P}^t = e^{t\mathcal{Q}}$ é estocástico e, desta forma, a partir de uma distribuição inicial μ sobre S , podemos definir, completamente, $\mathbb{P} \cdot \mathcal{P}^t$ para $t \geq 0$ é, assim, a matriz de transição associada ao processo Markoviano com tempo contínuo que procurávamos.

Em quase sua totalidade, as definições e demais resultados vistos no caso de processos de Markov com tempo discreto se estendem, de forma natural, ao caso de tempo contínuo com as devidas adaptações lógicas.

2 Processos de Markov Reversíveis e Forma de Dirichlet-Tempo Discreto

2.1 Introdução

Desenvolveremos o estudo de uma classe especial de Processos de Markov sobre S finito, a saber, processos de Markov *reversíveis*, i.e., que admitem uma distribuição inicial com respeito a qual eles são *reversíveis*, no sentido que, para qualquer intervalo de tempo, a *distribuição conjunta* (ver teorema 1.3.13) do processo é a mesma quando ele está evoluindo em tempo reverso (passado), assim como quando ele está evoluindo em tempo à frente (futuro).

Neste caso, tanto a cadeia original, quanto a de tempo reverso, são estocasticamente as mesmas, posto que, pelo teorema 1.3.13, a distribuição conjunta de uma cadeia de Markov está inteiramente determinada pela distribuição inicial e a matriz de transição.

Primeiramente, vamos derivar condições necessárias e/ou suficientes para que um Processo de Markov, tanto a tempo discreto quanto a tempo contínuo, seja reversível com respeito a uma dada distribuição inicial μ . Estudaremos detalhadamente a importante definição de *Forma de Dirichlet* e suas conseqüências, que serão fundamentais para o que segue nos capítulos 3 e 5, em particular, estamos focados em desenvolver ferramentas para estimar uma cota inferior para β que aparece na proposição 2.3.7.

2.2 Reversibilidade

2.2.1 Conceitos Gerais

Daremos agora a definição de Processo de Markov Reversível.

Definição 2.2.1

Seja $\{X_t\}_{t \geq 0}$ um Processo de Markov com distribuição inicial μ e matriz de transição \mathcal{P} irredutível. Dizemos que $\{X_t\}_{t \geq 0}$ é um *Processo de Markov Reversível* com respeito a μ se, para qualquer $n \geq 1$ e $(i_0, \dots, i_n) \in S^{n+1}$ vale, no tempo discreto

$$\mathbb{P}(X_0 = i_0, \dots, X_n = i_n) = \mathbb{P}(X_0 = i_n, \dots, X_n = i_0), \quad (2.1)$$

e no tempo contínuo

$$\mathbb{P}(X_0 = i_0, \dots, X_{t_n} = i_n) = \mathbb{P}(X_0 = i_n, \dots, X_{t_n - t_1} = i_1, X_{t_n} = i_0), \quad (2.2)$$

sempre que $0 = t_0 < t_1 < \dots < t_n$.

◇

Definição 2.2.2

De modo geral, dizemos que uma matriz $\mathcal{A} = (A)_{ij}$, $(i, j) \in \mathbb{Z}^2$ satisfaz a *Condição do Balanço Detalhado* para um vetor $v = ((v)_i)_{i \in \mathbb{Z}}$ se, e somente se, vale $\forall (i, j) \in \mathbb{Z}^2$

$$(v)_i(A)_{ij} = (v)_j(A)_{ji}. \quad (2.3)$$

◇

A próxima proposição nos fornece um critério para inferirmos sobre a reversibilidade de uma cadeia de Markov.

Proposição 2.2.3

Sejam $\{X_t\}_{t \geq 0}$ um Processo de Markov (qualquer) com matriz de transição de probabilidades \mathcal{P} irredutível e μ uma distribuição inicial. Então, $\{X_t\}_{t \geq 0}$ é um *Processo de Markov Reversível* com respeito a μ se, e somente se, vale a *condição do balanço detalhado* para \mathcal{P} e μ , i.e. $\Leftrightarrow \forall (i, j) \in S^2$ vale que

$$(\mu)_i(P)_{ij} = (\mu)_j(P)_{ji} \quad (2.4)$$

◇

Prova

Provaremos apenas para o caso de tempo discreto, o caso de tempo contínuo se prova analogamente.

Suponhamos que $\{X_n\}_{n \geq 0}$ seja reversível, então tomando $n = 1$ em 2.1 temos, $\forall (i, j) \in S^2$ que

$$\begin{aligned} (\mu)_i(P)_{ij} &= (\mu)_i \mathbb{P}(X_1 = j | X_0 = i) \stackrel{1}{=} (\mu)_i \frac{\mathbb{P}(X_1 = j, X_0 = i)}{\mathbb{P}(X_0 = i)} \stackrel{2}{=} \\ &= (\mu)_i \frac{\mathbb{P}(X_0 = j, X_1 = i)}{(\mu)_i} = \mathbb{P}(X_1 = i, X_0 = j) \stackrel{3}{=} \\ &= \mathbb{P}(X_0 = j) \mathbb{P}(X_1 = i | X_0 = j) = (\mu)_j(P)_{ji}, \end{aligned}$$

onde a 1 e 3 seguem da definição de probabilidade condicional (ver definição 1.2.6), utilizamos a equação 2.1 para justificar 2. Logo, $\forall (i, j) \in S^2$, vale a condição do balanço detalhado 2.4.

Reciprocamente, suponhamos válida, $\forall (i, j) \in S^2$, 2.4. Utilizaremos nesta parte da demonstração indução em n .

Com efeito, para $n = 1$ temos (ver teorema 1.3.13) que

$$\begin{aligned} \mathbb{P}(X_0 = i_0, X_1 = i_1) &= (\mu)_{i_0}(P)_{i_0 i_1} \stackrel{1}{=} (\mu)_{i_1}(P)_{i_1 i_0} \stackrel{2}{=} \mathbb{P}(X_0 = i_1, X_1 = i_0) = \\ &= \mathbb{P}(X_1 = i_0, X_0 = i_1), \end{aligned}$$

onde em 1 utilizamos a condição 2.4 e em 2 novamente o teorema 1.3.13. Logo, vale a equação 2.1 para $n = 1$. No que consiste, suponhamos válida a equação 2.1 para um certo $n = k$, i.e, que

$$\begin{aligned} \mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_k = i_k) &= (\mu)_{i_0}(P)_{i_0 i_1} \dots (P)_{i_{k-1} i_k} = \mu_{i_k}(P)_{i_k i_{k-1}} \dots (P)_{i_1 i_0} = \\ &= \mathbb{P}(X_0 = i_k, X_1 = i_{k-1}, \dots, X_k = i_0). \end{aligned}$$

Vamos então verificar que vale a equação 2.1 para $n = k + 1$, ou seja, denotando $\mathbb{P}^{k+1} = \mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_{k+1} = i_{k+1})$, temos

$$\begin{aligned}
\mathbb{P}^{k+1} & \stackrel{1}{=} (\mu)_{i_0}(P)_{i_0 i_1} \cdots (P)_{i_{k-1} i_k} (P)_{i_k i_{k+1}} \stackrel{2}{=} \\
& = (\mu)_{i_k}(P)_{i_k i_{k-1}} \cdots (P)_{i_1 i_0} (P)_{i_k i_{k+1}} = \\
& = (\mu)_{i_k}(P)_{i_k i_{k+1}} (P)_{i_k i_{k-1}} \cdots (P)_{i_1 i_0} \stackrel{3}{=} \\
& = (\mu)_{i_{k+1}}(P)_{i_{k+1} i_k} (P)_{i_k i_{k-1}} \cdots (P)_{i_1 i_0} \stackrel{4}{=} \\
& = \mathbb{P}(X_0 = i_{k+1}, X_1 = i_k, \dots, X_{k+1} = i_0),
\end{aligned}$$

onde 1 e 4 seguem do teorema 1.3.13, 2 segue da hipótese de indução acima, 3 vem da condição de balanço detalhado 2.4 aplicada à $(\mu)_{i_k}(P)_{i_k i_{k+1}}$.

Ou seja, vale a equação 2.1 para o sucessor de k e, portanto, pelo princípio da indução, vale $\forall n \in \mathbb{N}$. Logo, o processo é reversível com respeito a μ .

□

Exemplo 2.1

Considere o processo estocástico determinado pela distribuição inicial $\mu = (1/4, 1/2, 1/4)$ e pela matriz estocástica \mathcal{P} dada por

$$\mathcal{P} = \begin{bmatrix} 1/4 & 1/2 & 1/4 \\ 1/4 & 1/2 & 1/4 \\ 1/4 & 1/2 & 1/4 \end{bmatrix},$$

Vamos verificar que o processo de Markov associado é reversível. Com efeito, pela proposição 2.2.3 é suficiente mostrar a validade da relação 2.4, note que

$$(\mu)_1(P)_{12} = \frac{1}{4} \frac{1}{2} = \frac{1}{2} \frac{1}{4} = (\mu)_2(P)_{21};$$

$$(\mu)_1(P)_{13} = \frac{1}{4} \frac{1}{4} = (\mu)_3(P)_{31};$$

e

$$(\mu)_2(P)_{23} = \frac{1}{2} \frac{1}{4} = \frac{1}{4} \frac{1}{2} = (\mu)_3(P)_{32}.$$

Logo, pela proposição 2.2.3, o processo é reversível.

Além disso, é relativamente fácil ver que $\mu\mathcal{P} = \mu$, ou seja, μ é uma distribuição estacionária para \mathcal{P} . Tal fato visto acima é geral para processos reversíveis, como segue no corolário abaixo.

□

Corolário 2.2.4

Se $\{X_t\}_{t \geq 0}$ é um processo de Markov reversível com respeito a μ , então μ é distribuição estacionária para \mathcal{P} , i.e., $\mu\mathcal{P} = \mu$.

◇

Prova

Como $\{X_t\}_{t \geq 0}$ é reversível com respeito a μ vem, pela proposição 2.2.3, que vale a condição do balanço detalhado para μ e \mathcal{P} , $\forall (i, j) \in S^2$. Logo, utilizando a definição de produto matricial, vem

$$(\mu\mathcal{P})_i = \sum_{k \in S} (\mu)_k (P)_{ki} \stackrel{1}{=} \sum_{k \in S} (\mu)_i (P)_{ik} = (\mu)_i \sum_{k \in S} (P)_{ik} \stackrel{2}{=} (\mu)_i,$$

onde 1 segue da condição 2.4 e 2 vem do fato de \mathcal{P} ser (linha) estocástica. Logo, $(\mu\mathcal{P})_i = (\mu)_i \forall i \in S$, donde μ é distribuição estacionária para \mathcal{P} .

□

O seguinte exemplo mostra que a recíproca do corolário 2.2.4 é, em geral, falsa.

Exemplo 2.2

Consideremos uma cadeia de Markov cuja matriz de transição de probabilidades seja dada por

$$\mathcal{P} = \begin{bmatrix} 0 & 2/3 & 1/3 \\ 1/3 & 0 & 2/3 \\ 2/3 & 1/3 & 0 \end{bmatrix},$$

e seja $\mu = (1/3, 1/3, 1/3)$. Claramente μ é uma distribuição estacionária para \mathcal{P} . Note que $(\mu)_1 = (\mu)_2 = (\mu)_3 = 1/3$, então para verificarmos a condição 2.4 basta vermos, por exemplo, que $(P)_{12} = 1/3 \neq 2/3 = (P)_{21}$. Logo, **não** vale a condição do balanço detalhado e, assim, pela proposição 2.2.3, vem que o processo de Markov que tem esta \mathcal{P} como matriz de transição e este μ como distribuição estacionária **não** é reversível.

□

2.3 Convergência em $L^2(\mu)$ - Tempo Discreto

Queremos compreender como atua a dinâmica de uma Cadeia de Markov \mathcal{P} Irredutível, Reversível com respeito a μ , sobre uma *função mensurável* (ver definição 1.2.10) $f : S \rightarrow \mathbb{R}$, onde S é o espaço de estados que estamos considerando finito (digamos que possua N estados). Mais especificamente, estamos interessados em saber o quão rápido os iterados de \mathcal{P} , agindo sobre f (no sentido da equação 2.10), se aproximam do valor esperado de f , ou de um estado de equilíbrio.

Para motivação, sabe-se, pelo Teorema de Perron - Frobenius (ver teorema 1.3.19), que o autovalor de segundo maior módulo de uma matriz estocástica nos dá a taxa de convergência ao equilíbrio (ver exemplo 1.4, página 198 de [4]).

A definição abaixo será utilizada comumente ao longo do restante deste trabalho evidenciando, deste modo, sua importância.

Definição 2.3.1

Seja μ uma distribuição inicial estritamente positiva (i.e., $(\mu)_i > 0 \forall i \in S$) e considere as funções $f, g : S \rightarrow \mathbb{R}$.

Definimos o *produto interno com respeito a μ* entre f e g por

$$\langle f, g \rangle_{2, \mu} := \sum_{i \in S} f(i)g(i)(\mu)_i. \quad (2.5)$$

A *norma de f com respeito a μ* definimos como

$$\|f\|_{2, \mu} := \sqrt{\langle f, f \rangle_{2, \mu}} = \sqrt{\sum_{i \in S} f(i)^2(\mu)_i}. \quad (2.6)$$

Já o *valor esperado de f com respeito a μ* é

$$\langle f \rangle_{\mu} := \langle f, \mathbf{1} \rangle_{2, \mu} = \sum_{i \in S} f(i)(\mu)_i, \quad (2.7)$$

onde $\mathbf{1} : S \rightarrow \mathbb{R}$ é tal que $\mathbf{1}(i) = 1 \forall i \in S$.

Finalmente, a *variância de f com respeito a μ* será dada por

$$Var_{\mu}(f) := \sum_{i \in S} \left(f(i) - \langle f \rangle_{\mu} \right)^2 (\mu)_i = \langle f^2 \rangle_{\mu} - \langle f \rangle_{\mu}^2. \quad (2.8)$$

◇

Não é uma tarefa difícil provar que 2.5 e 2.6 satisfazem, de fato, os axiomas de produto interno e norma respectivamente.

Observação 2.3.2

Para o produto interno definido em 2.5 vale a *Desigualdade de Cauchy-Schwarz*, ou seja, $\forall f, g \in L^2(\mu)$ temos

$$|\langle f, g \rangle_{2, \mu}| \leq \|f\|_{2, \mu} \|g\|_{2, \mu}. \quad (2.9)$$

Para um esboço da demonstração de que $\|\cdot\|_{2, \mu}$ é, de fato, uma norma e da desigualdade 2.9 veja [24].

◇

Observação 2.3.3

Utilizaremos a notação $L^2(\mu) := \{f : S \rightarrow \mathbb{R} : \|f\|_{2, \mu} < \infty\}$, i.e, $f : S \rightarrow \mathbb{R}$ limitadas. $L^2(\mu)$ é um espaço vetorial de dimensão finita com métrica induzida $d_{\mu}(f, g) = \|f - g\|_{2, \mu}$. Este espaço métrico, munido com a métrica $d_{\mu}(f, g)$ é completo.

◇

Doravante, utilizaremos a seguinte *notação* (veja também subseção 2.3.2)

$$\mathcal{P}f(i) := \sum_{j \in S} f(j)(P)_{ij} = (\mathcal{P}\mathbf{f})_i, \quad (2.10)$$

onde \mathbf{f} é o vetor coluna determinado por f (i.e., $\mathbf{f} := (f(1), f(2), \dots, f(N))$).

Teorema 2.3.4

Uma matriz de transição de probabilidades irreduzível \mathcal{P} é reversível com respeito a uma distribuição inicial μ se, e somente se, \mathcal{P} é auto-adjunta em $L^2(\mu)$, i.e.,
 \Leftrightarrow

$$\langle \mathcal{P}f, g \rangle_{2, \mu} = \langle f, \mathcal{P}g \rangle_{2, \mu}, \quad (2.11)$$

$\forall f, g \in L^2(\mu)$.

◇

Prova

Suponhamos que \mathcal{P} seja uma matriz de transição de probabilidades irreduzível e reversível com respeito a μ , então, pela proposição 2.2.3 temos, $\forall (i, j) \in S^2$, que

$$(\mu)_i(P)_{ij} = (\mu)_j(P)_{ji}.$$

Logo, $\forall f, g \in L^2(\mu)$, temos

$$\begin{aligned} \langle \mathcal{P}f, g \rangle_{2, \mu} &= \sum_{i \in S} (\mathcal{P}f(i))g(i)(\mu)_i = \sum_{i \in S} \left\{ \left(\sum_{j \in S} f(j)(P)_{ij} \right) g(i)(\mu)_i \right\} = \\ &= \sum_{i \in S} \sum_{j \in S} (\mu)_i(P)_{ij} f(j)g(i) \stackrel{1}{=} \sum_{i \in S} \sum_{j \in S} (\mu)_j(P)_{ji} f(j)g(i) = \\ &= \sum_{j \in S} \left\{ f(j) \left(\sum_{i \in S} g(i)(P)_{ji} \right) (\mu)_j \right\} = \langle f, \mathcal{P}g \rangle_{2, \mu}, \end{aligned}$$

onde 1 segue da condição do balanço detalhado dada acima.

Donde que, $\langle \mathcal{P}f, g \rangle_{2, \mu} = \langle f, \mathcal{P}g \rangle_{2, \mu} \forall f, g \in L^2(\mu)$ e, daí, \mathcal{P} é auto-adjunta em $L^2(\mu)$.

Reciprocamente, suponhamos que \mathcal{P} seja auto-adjunta em $L^2(\mu)$, ou seja, que valha 2.11 $\forall f, g \in L^2(\mu)$. Defina, para cada $k \in S$, $\delta_k(i) := \delta_{k,i}$, i.e., o Delta de Kronecker tal que $\delta_{k,i} = 0$ e $\delta_{k,k} = 1$. Então, tomando $f(k) = \delta_i(k)$ e $g(k) = \delta_j(k)$ em 2.11 temos, para cada $(i, j) \in S^2$, que

$$\begin{aligned} \langle \mathcal{P}\delta_i, \delta_j \rangle_{2, \mu} &= \sum_{k \in S} (\mathcal{P}\delta_i(k))\delta_j(k)(\mu)_k = \sum_{k \in S} \left(\sum_{l \in S} \delta_i(l)(P)_{kl} \right) \delta_j(k)(\mu)_k = \\ &= \sum_{k \in S} (\mu)_k(P)_{ki}\delta_j(k) = (\mu)_j(P)_{ji}. \end{aligned}$$

De forma análoga

$$\begin{aligned} \langle \delta_i, \mathcal{P}\delta_j \rangle_{2, \mu} &= \sum_{k \in S} \delta_i(k)(\mathcal{P}\delta_j(k))(\mu)_k = \sum_{k \in S} \delta_i(k) \left(\sum_{l \in S} \delta_j(l)(P)_{kl} \right) (\mu)_k = \\ &= \sum_{k \in S} (\mu)_k(P)_{kj}\delta_i(k) = (\mu)_i(P)_{ij}. \end{aligned}$$

Portanto, voltando à equação 2.11, somos levados a concluir, $\forall (i, j) \in S^2$, que $(\mu)_i(P)_{ij} = (\mu)_j(P)_{ji}$ e daí, pela proposição 2.2.3, \mathcal{P} é reversível com respeito a μ .

□

Prosseguindo, devemos notar que, para cada $f \in L^2(\mu)$, $\mathcal{P}f(i)$ está bem definida (pois, sendo S finito a soma definida em 2.10 sempre é um número real!).

Proposição 2.3.5

Seja \mathcal{P} uma matriz de transição de probabilidades irredutível, reversível com respeito a μ . Então, \mathcal{P} é uma *Contração Auto-Adjunta* no Espaço de Hilbert ¹ $L^2(\mu)$.

Mais especificamente vale, $\forall f, g \in L^2(\mu)$, que

$$\langle \mathcal{P}f, g \rangle_{2, \mu} = \langle f, \mathcal{P}g \rangle_{2, \mu}, \quad (2.12)$$

e

$$\|\mathcal{P}f\|_{2, \mu} \leq \|f\|_{2, \mu}. \quad (2.13)$$

◇

Prova

Já foi provado no teorema 2.3.4 que \mathcal{P} é auto-adjunta, i.e., que vale 2.12. Basta mostrar, então, que \mathcal{P} é uma contração em $L^2(\mu)$, ou seja, mostrar a validade da relação 2.13. Com efeito, note primeiramente que $\forall i \in S$

$$\begin{aligned} [\mathcal{P}f(i)]^2 & \stackrel{1}{=} \left[\sum_j f(j)(P)_{ij} \right]^2 \stackrel{2}{=} \left| \sum_j f(j)(P)_{ij} \right|^2 \leq^3 \left[\sum_j |f(j)|(P)_{ij} \right]^2 \leq^4 \\ & \leq \sum_j |f(j)|^2 (P)_{ij} \stackrel{5}{=} \sum_j f(j)^2 (P)_{ij} \stackrel{6}{=} \mathcal{P}(f^2)(i), \end{aligned}$$

onde 1 segue da relação 2.10, 2 e 5 vem do fato que $\forall x \in \mathbb{R} \ |x|^2 = x^2$, 3 vem da desigualdade triangular, 4 vem da proposição 1.3.23 tomando $\psi(x) = x^2$ (que é não-negativa em $[0, \infty)$, contínua e *convexa*) e lembrando que \mathcal{P} é (linha) estocástica (logo, $\sum_{j \in S} (P)_{ij} = 1, \forall i \in S$), já 6 segue novamente pela relação 2.10.

Portanto,

$$\begin{aligned} \|\mathcal{P}f\|_{2, \mu}^2 & = \sum_{k \in S} (\mathcal{P}f(k))^2 (\mu)_k \leq^1 \sum_{k \in S} \mathcal{P}(f^2(k)) (\mu)_k = \sum_{k \in S} \left(\sum_{j \in S} f^2(j)(P)_{kj} \right) (\mu)_k = \\ & = \sum_{j \in S} f^2(j) \left(\sum_{k \in S} (\mu)_k (P)_{kj} \right) \stackrel{2}{=} \sum_{j \in S} f^2(j) (\mu)_j \stackrel{3}{=} \|f\|_{2, \mu}^2, \end{aligned}$$

onde 1 segue do arazoado feito acima, 2 vale pois μ é invariante para \mathcal{P} , 3 vem da definição de norma 2.3.1.

Ou seja,

$$\|\mathcal{P}f\|_{2, \mu} \leq \|f\|_{2, \mu}.$$

$\forall f \in L^2(\mu)$. Mostramos com isto que \mathcal{P} é uma contração, concluindo, deste modo, a prova desta proposição.

□

¹Um espaço de Hilbert é um espaço vetorial equipado com um produto interno que determina uma norma para a qual a métrica associada é completa (veja, [19]).

Exemplo 2.3

Vamos considerar o espaço de estados $S = \{1, 2, 3\}$, o vetor de probabilidade $\mu = (1/3, 1/3, 1/3)$ e as funções $f, g : S \rightarrow \mathbb{R}$ dadas por $f(i) := i^2$ e $g(i) := 2i$. Além disso, seja \mathcal{P} a matriz das probabilidades de transição definida abaixo.

$$\mathcal{P} = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{bmatrix}.$$

É relativamente fácil ver que \mathcal{P} , dada acima, é irredutível e reversível com respeito a μ .

Mostraremos que a conclusão da proposição 2.3.5 é verdadeira neste exemplo particular, i.e., \mathcal{P} é uma contração auto-adjunta em $L^2(\mu)$. Com efeito, primeiramente pela definição de f e g vale que $(f(1), f(2), f(3)) = (1, 4, 9)$ e $(g(1), g(2), g(3)) = (2, 4, 6)$. Além disso, pela relação 2.10 temos que $(\mathcal{P}f(1), \mathcal{P}f(2), \mathcal{P}f(3)) = (13/2, 5, 5/2)$ e, analogamente, $(\mathcal{P}g(1), \mathcal{P}g(2), \mathcal{P}g(3)) = (5, 4, 3)$.

Logo, pela definição de produto interno em $L^2(\mu)$ (definição 2.3.1) e pelas relações do parágrafo anterior.

$$\langle \mathcal{P}f, g \rangle_{2,\mu} = \sum_{i=1}^3 \mathcal{P}f(i)g(i)(\mu)_i = \frac{1}{3} \left[\frac{13}{2} \cdot 2 + 5 \cdot 4 + \frac{5}{2} \cdot 6 \right] = 16,$$

e, de forma análoga

$$\langle f, \mathcal{P}g \rangle_{2,\mu} = \sum_{i=1}^3 f(i)\mathcal{P}g(i)(\mu)_i = \frac{1}{3} [1 \cdot 5 + 4 \cdot 4 + 9 \cdot 3] = 16.$$

Ou seja, $\langle \mathcal{P}f, g \rangle_{2,\mu} = \langle f, \mathcal{P}g \rangle_{2,\mu}$.

Finalmente, pela definição de norma em $L^2(\mu)$ (definição 2.3.1) temos

$$\|f\|_{2,\mu} = \sqrt{\sum_{i=1}^3 f(i)^2(\mu)_i} = \sqrt{\frac{1}{3}(1^2 + 4^2 + 9^2)} = 5, 71.$$

e

$$\|\mathcal{P}f\|_{2,\mu} = \sqrt{\sum_{i=1}^3 \mathcal{P}f(i)^2(\mu)_i} = \sqrt{\frac{1}{3} \left(\left(\frac{13}{2}\right)^2 + 5^2 + \left(\frac{5}{2}\right)^2 \right)} = 4, 95.$$

Donde que, $\|\mathcal{P}f\|_{2,\mu} \leq \|f\|_{2,\mu}$ como afirma a proposição 2.3.5.

□

2.3.1 O Gap Espectral

Voltando à problemática introduzida na seção 2.3 provaremos, com resultados da mesma seção que, $\forall f \in L^2(\mu)$, vale

$$\lim_{n \rightarrow \infty} \|\mathcal{P}^n f - \langle f \rangle_\mu\|_{2,\mu} = 0, \quad (2.14)$$

onde \mathcal{P} é estocástica, reversível com respeito a μ .

Obteremos também uma primeira estimativa para a taxa de tal convergência.

Podemos identificar o espaço das funções $f : S \rightarrow \mathbb{R}$ com o \mathbb{R}^N (lembre que $\mathbf{f} = (f(1), \dots, f(N))$). Deste modo, o produto interno em $L^2(\mu)$ tornar-se-á o produto interno $\sum_{i \in S} (\mu)_i (v)_i (w)_i$ dos vetores $\mathbf{v}, \mathbf{w} \in \mathbb{R}^N$.

Posto isto, a proposição 2.3.5 nos afirma que \mathcal{P} atua como uma contração simétrica em \mathbb{R}^N com este produto interno.

Defina $\mathcal{P}^* := D^{1/2} \mathcal{P} D^{-1/2}$, onde $D = \text{diag}\{(\mu)_1, (\mu)_2, \dots, (\mu)_N\}$ então, pela relação 2.4, \mathcal{P}^* é simétrica com respeito ao produto interno padrão $\langle \mathbf{v}, \mathbf{w} \rangle_{\mathbb{R}^N} = \sum_{i \in S} (v)_i (w)_i$ em \mathbb{R}^N , com $\|\cdot\|_{\mathbb{R}^N}$ sendo a norma associada. Logo, podemos enunciar a seguinte proposição.

Proposição 2.3.6

Como um operador sobre o \mathbb{R}^N , \mathcal{P}^* é uma contração, i.e.,

$$\|\mathcal{P}^* \mathbf{f}\|_{\mathbb{R}^N} \leq \|\mathbf{f}\|_{\mathbb{R}^N}. \quad (2.15)$$

◇

Prova

Utilizando a definição de $\|\cdot\|_{\mathbb{R}^N}$, temos

$$\begin{aligned} \|\mathcal{P}^* \mathbf{f}\|_{\mathbb{R}^N}^2 &:= \langle \mathcal{P}^* \mathbf{f}, \mathcal{P}^* \mathbf{f} \rangle_{\mathbb{R}^N} = \sum_{i \in S} (\mathcal{P}^* \mathbf{f})_i^2 = \sum_{i \in S} (D^{1/2} \mathcal{P} D^{-1/2} \mathbf{f})_i^2 = 1 \\ &= \sum_{i \in S} (\mu)_i (\mathcal{P} D^{-1/2} \mathbf{f})_i^2 = \|\mathcal{P} g\|_{2, \mu}^2 \leq \|g\|_{2, \mu}^2 = \langle g, g \rangle_{2, \mu} = \\ &= \sum_{i \in S} g(i)^2 (\mu)_i = \sum_{i \in S} (D^{-1/2} \mathbf{f})_i^2 (\mu)_i = \sum_{i \in S} (D^{1/2} D^{-1/2} \mathbf{f})_i^2 = \\ &= \sum_{i \in S} (\mathbf{f})_i^2 = \|\mathbf{f}\|_{\mathbb{R}^N}^2, \end{aligned}$$

onde g é a função determinada pelo vetor $D^{-1/2} \mathbf{f}$. 1 e 3 seguem do fato que, sendo D diagonal $\Rightarrow (D^{1/2})_{ii} = (D)_{ii}^{1/2} = \sqrt{(\mu)_i}$, já 2 segue em virtude de 2.13.

□

É bem conhecido da teoria das matrizes simétricas em \mathbb{R}^N ([10], [16]) que \mathcal{P}^* admite autovalores $1 \geq \lambda_1 \geq \dots \geq \lambda_N \geq -1$, com autovetores associados $\{\mathbf{v}_1, \dots, \mathbf{v}_N\}$ que são ortonormais em relação ao produto interno $\langle \cdot, \cdot \rangle_{\mathbb{R}^N}$, i.e., $\langle \mathbf{v}_k, \mathbf{v}_l \rangle_{\mathbb{R}^N} = \delta_{kl}$, onde δ_{ij} é o Delta de Kronecker.

Além disso, \mathcal{P} e \mathcal{P}^* possuem os mesmos autovalores (são conjugadas por $D^{1/2}$, ver exercício 8, página 56 de [7]), como sabemos que \mathcal{P} é estocástica vem, pelo Teorema de Perron-Frobenius (teorema 1.3.19), que $\lambda_1 = 1$.

Sendo \mathbf{v}_1 o autovetor de \mathcal{P}^* associado a $\lambda_1 = 1$ temos, por um lado, que

$$(\mathbf{v}_1)_i = (\mathcal{P}^* \mathbf{v}_1)_i = \sum_{k \in S} (P^*)_{ik} (\mathbf{v}_1)_k. \quad (2.16)$$

Por outro lado, temos que

$$\sqrt{(\mu)_i} = \sum_{k \in S} (P^*)_{ik} \sqrt{(\mu)_k}. \quad (2.17)$$

Comparando 2.16 com 2.17 somos levados a tomar $(\mathbf{v}_1)_i = \sqrt{(\mu)_i}$, para cada $i \in S$. A validade da relação 2.17 será analisada para o caso particular $N = 2$, o caso geral se demonstra de forma análoga. Primeiramente note que,

$$\mathcal{P}^* = D^{1/2} \mathcal{P} D^{-1/2} = \begin{bmatrix} (P)_{11} & (P)_{12} \frac{\sqrt{(\mu)_1}}{\sqrt{(\mu)_2}} \\ (P)_{21} \frac{\sqrt{(\mu)_2}}{\sqrt{(\mu)_1}} & (P)_{22} \end{bmatrix}.$$

Logo, utilizando $(P^*)_{ij} \forall (i, j) \in \{1, 2\}^2$ dado pela expressão acima temos,

$$\begin{aligned} \sum_{k=1}^2 (P^*)_{1k} \sqrt{(\mu)_k} &= (P^*)_{11} \sqrt{(\mu)_1} + (P^*)_{12} \sqrt{(\mu)_2} = \\ &= (P)_{11} \sqrt{(\mu)_1} + (P)_{12} \frac{\sqrt{(\mu)_1}}{\sqrt{(\mu)_2}} \sqrt{(\mu)_2} = \\ &= (P)_{11} \sqrt{(\mu)_1} + (P)_{12} \sqrt{(\mu)_1} = \\ &= ((P)_{11} + (P)_{12}) \sqrt{(\mu)_1} = 1 \sqrt{(\mu)_1}, \end{aligned}$$

onde 1 segue do fato que \mathcal{P} é (linha) estocástica.

Analogamente,

$$\sum_{k=1}^2 (P^*)_{2k} \sqrt{(\mu)_k} = \sqrt{(\mu)_2},$$

justificando a equação 2.17 para o caso $N = 2$ como desejávamos.

Finalmente, definamos $\mathbf{g}_k := D^{-1/2} \mathbf{v}_k$ e seja g_k a associada função em S . Vemos, para cada $i \in S$, (já que $\mathcal{P}^* = D^{1/2} \mathcal{P} D^{-1/2} \Rightarrow D^{-1/2} \mathcal{P}^* = \mathcal{P} D^{-1/2}$) que

$$\mathcal{P} \mathbf{g}_k = \mathcal{P} D^{-1/2} \mathbf{v}_k = D^{-1/2} \mathcal{P}^* \mathbf{v}_k = D^{-1/2} \lambda_k \mathbf{v}_k = \lambda_k D^{-1/2} \mathbf{v}_k = \lambda_k \mathbf{g}_k.$$

Donde pode-se concluir que

$$\mathcal{P} g_k = \lambda_k g_k. \quad (2.18)$$

Também, com a identificação feita acima, $(\mathbf{v}_1)_i = \sqrt{(\mu)_i} \forall i \in S$, temos que $g_1(i) = 1 \forall i \in S$ e, além disso,

$$\begin{aligned} \langle g_l, g_k \rangle_{2, \mu} &= \sum_{i \in S} g_k(i) g_l(i) (\mu)_i = \sum_{i \in S} \frac{(\mathbf{v}_k)_i}{\sqrt{(\mu)_i}} \frac{(\mathbf{v}_l)_i}{\sqrt{(\mu)_i}} (\mu)_i = \sum_{i \in S} \frac{(\mathbf{v}_k)_i (\mathbf{v}_l)_i}{(\mu)_i} (\mu)_i = \\ &= \sum_{i \in S} (\mathbf{v}_k)_i (\mathbf{v}_l)_i = \langle \mathbf{v}_k, \mathbf{v}_l \rangle_{\mathbb{R}^N} = \delta_{kl}. \end{aligned}$$

Resumindo, mostramos que \mathcal{P} , atuando em $L^2(\mu)$, tem autovalores $1 = \lambda_1 \geq \dots \geq \lambda_N \geq -1$ que correspondem às autofunções g_1, \dots, g_N que são ortonormais com respeito ao produto interno $\langle \cdot, \cdot \rangle_{2, \mu}$. Ao mesmo tempo, já que $L^2(\mu)$ tem dimensão N e, pela ortonormalidade, as g_k , para $k \in \{1, \dots, N\}$, são linearmente independentes,

vem que $\zeta = \{g_1, \dots, g_N\}$ é uma base ortonormal em $L^2(\mu)$. Por conseguinte, toda $f \in L^2(\mu)$ pode ser escrita como combinação linear dos elementos da base $\zeta = \{g_1, \dots, g_N\}$, mais explicitamente temos

$$f = \sum_{k \in S} \langle f, g_k \rangle_{2,\mu} g_k. \quad (2.19)$$

A proposição abaixo nos mostra que não apenas, de fato, vale o limite 2.14, como também nos fornece a taxa de tal convergência.

Proposição 2.3.7

Seja \mathcal{P} uma matriz de transição de probabilidades irredutível, reversível com respeito a μ . Então, $\forall f \in L^2(\mu)$ e $n \geq 0$, vale que

$$\|\mathcal{P}^n f - \langle f \rangle_\mu\|_{2,\mu} \leq (1 - \beta)^n \|f - \langle f \rangle_\mu\|_{2,\mu}, \quad (2.20)$$

onde $\beta := (1 - \lambda_2) \wedge (1 + \lambda_N)$ é chamado o *Gap Espectral* entre $\{-1, 1\}$ e $\{\lambda_2, \dots, \lambda_N\}$ (notação $x \wedge y := \min\{x, y\}$).

Em particular, como $\beta \in (0, 1]$, vale o limite 2.14.

◇

Prova

É relativamente fácil ver que de 2.18 vem, por indução em n , que $\mathcal{P}^n g_k = \lambda_k^n g_k \forall k \in S$, com isto, levando-se em conta que \mathcal{P} é um operador linear vem, utilizando a relação 2.19, que

$$\mathcal{P}^n f = \sum_{k \in S} \lambda_k^n \langle f, g_k \rangle_{2,\mu} g_k.$$

Desta última expressão segue que (utilizando a definição de valor esperado 2.3.1)

$$\begin{aligned} \mathcal{P}^n f &= \sum_{k \in S} \lambda_k^n \langle f, g_k \rangle_{2,\mu} g_k = \langle f, g_1 \rangle_{2,\mu} g_1 + \sum_{k \in S \setminus \{1\}} \lambda_k^n \langle f, g_k \rangle_{2,\mu} g_k = \\ &= \langle f, \mathbf{1} \rangle_{2,\mu} \mathbf{1} + \sum_{k \in S \setminus \{1\}} \lambda_k^n \langle f, g_k \rangle_{2,\mu} g_k = \langle f \rangle_\mu + \sum_{k \in S \setminus \{1\}} \lambda_k^n \langle f, g_k \rangle_{2,\mu} g_k. \end{aligned}$$

Donde vem que,

$$\mathcal{P}^n f - \langle f \rangle_\mu = \sum_{k \in S \setminus \{1\}} \lambda_k^n \langle f, g_k \rangle_{2,\mu} g_k,$$

e daí,

$$\begin{aligned} \|\mathcal{P}^n f - \langle f \rangle_\mu\|_{2,\mu}^2 &= \left\langle \sum_{k \in S \setminus \{1\}} \lambda_k^n \langle f, g_k \rangle_{2,\mu} g_k, \sum_{k' \in S \setminus \{1\}} \lambda_{k'}^n \langle f, g_{k'} \rangle_{2,\mu} g_{k'} \right\rangle_{2,\mu} =^0 \\ &= \sum_{k \in S \setminus \{1\}} \lambda_k^{2n} \langle f, g_k \rangle_{2,\mu}^2 \langle g_k, g_k \rangle_{2,\mu} =^1 \sum_{k \in S \setminus \{1\}} \lambda_k^{2n} \langle f, g_k \rangle_{2,\mu}^2 \leq^2 \\ &\leq (1 - \beta)^{2n} \|f - \langle f \rangle_\mu\|_{2,\mu}^2, \end{aligned}$$

onde 0 segue da linearidade do produto interno juntamente com a ortonormalidade da base $\zeta = \{g_1, \dots, g_N\}$, enquanto que 1 segue somente desta última, já 2 vem do fato que $\lambda_k \leq 1 - \beta$, para $k \in S \setminus \{1\}$.

Para completar a justificativa de 2 note que, analogamente ao que fizemos acima vem, pela equação 2.19, que

$$f - \langle f \rangle_\mu = \sum_{k \in S \setminus \{1\}} \langle f, g_k \rangle_{2,\mu} g_k,$$

e daí,

$$\|f - \langle f \rangle_\mu\|_{2,\mu}^2 = \sum_{k \in S \setminus \{1\}} \langle f, g_k \rangle_{2,\mu}^2 \langle g_k, g_k \rangle_{2,\mu} = \sum_{k \in S \setminus \{1\}} \langle f, g_k \rangle_{2,\mu}^2.$$

Portanto,

$$\|\mathcal{P}^n f - \langle f \rangle_\mu\|_{2,\mu} \leq (1 - \beta)^n \|f - \langle f \rangle_\mu\|_{2,\mu}.$$

$\forall f \in L^2(\mu)$ e $\forall n \geq 0$.

Como $\beta \in (0, 1]$ é óbvio que vale o limite 2.14.

□

A relação 2.20 é nosso foco principal e dela vemos que a convergência é polinomial de taxa $r = 1 - \beta$, logo toda e qualquer questão sobre a velocidade de convergência r pode ser traduzida em termos de β .

Exemplo 2.4

Retornando à cadeia de Markov do exemplo 2.3 temos que seus autovalores são $\lambda_1 = 1$ (dominante), $\lambda_2 = \lambda_3 = -1/2$. Como $\beta = (1 - \lambda_2) \wedge (1 + \lambda_3) = (3/2) \wedge (1/2) = 1/2$.

Logo, a taxa de convergência será $r = 1 - \beta = 1/2$.

□

A próxima proposição nos dá uma expressão mais interessante para o β que aparece em 2.20.

Proposição 2.3.8

A relação 2.20 é válida quando

$$\beta = 1 - \sup\{\|\mathcal{P}f - \langle f \rangle_\mu\|_{2,\mu} : f \in L^2(\mu) \text{ com } \|f\|_{2,\mu} = 1\}. \quad (2.21)$$

◇

Prova

Para provar tal proposição, utilizaremos indução em n . Note que, quando β é dado por 2.21 tem-se, para $n = 1$, que

$$\|\mathcal{P}f - \langle f \rangle_\mu\|_{2,\mu} = \|f\|_{2,\mu} \left\| \mathcal{P} \left(\frac{f}{\|f\|_{2,\mu}} \right) - \left\langle \frac{f}{\|f\|_{2,\mu}} \right\rangle_\mu \right\|_{2,\mu} \leq (1 - \beta) \|f\|_{2,\mu},$$

quando $f \in L^2(\mu) \setminus \{0\}$ é qualquer. É claro que $\|\mathcal{P}f - \langle f \rangle_\mu\|_{2,\mu} \leq \|f\|_{2,\mu}$ quando $f = 0$. Além do mais, observe que

$$\begin{aligned} \langle \mathcal{P}^n f \rangle_\mu &= \sum_{i \in S} \left(\sum_{k \in S} f(k) (P^n)_{ik} \right) (\mu)_i = \sum_{k \in S} f(k) \sum_{i \in S} (\mu)_i (P^n)_{ik} = \sum_{k \in S} f(k) (\mu)_k = \langle f \rangle_\mu \quad \forall n \geq 0, \end{aligned}$$

onde 1 é válido pois μ é invariante para \mathcal{P} , já 2 vem da definição de valor esperado 2.3.1.

Logo, supondo válida a relação para um certo n fixo, temos

$$\begin{aligned} \|\mathcal{P}^{n+1} f - \langle f \rangle_\mu\|_{2,\mu} &= \|\mathcal{P}(\mathcal{P}^n f) - \langle \mathcal{P}^n f \rangle_\mu\|_{2,\mu} \leq (1 - \beta) \|\mathcal{P}^n f - \langle \mathcal{P}^n f \rangle_\mu\|_{2,\mu} = \\ &= (1 - \beta) \|\mathcal{P}^n f - \langle f \rangle_\mu\|_{2,\mu} \leq (1 - \beta)^{n+1} \|f\|_{2,\mu}. \end{aligned}$$

Daí, por indução, temos que $\|\mathcal{P}^n f - \langle f \rangle_\mu\|_{2,\mu} \leq (1 - \beta)^n \|f\|_{2,\mu} \quad \forall n \geq 0$. Logo, para $f \in L^2(\mu)$, definamos $\bar{f} := f - \langle f \rangle_\mu$. Temos finalmente, $\forall n \geq 0$, que

$$\begin{aligned} \|\mathcal{P}^n f - \langle f \rangle_\mu\|_{2,\mu} &= \|\mathcal{P}^n(f - \langle f \rangle_\mu)\|_{2,\mu} = \|\mathcal{P}^n \bar{f}\|_{2,\mu} = \|\mathcal{P}^n \bar{f} - 0\|_{2,\mu} = \\ &= \|\mathcal{P}^n \bar{f} - \langle \bar{f} \rangle_\mu\|_{2,\mu} \leq (1 - \beta)^n \|\bar{f}\|_{2,\mu} = (1 - \beta)^n \|f - \langle f \rangle_\mu\|_{2,\mu}. \end{aligned}$$

Assim sendo,

$$\|\mathcal{P}^n f - \langle f \rangle_\mu\|_{2,\mu} \leq (1 - \beta)^n \|f - \langle f \rangle_\mu\|_{2,\mu}.$$

$\forall f \in L^2(\mu)$ e $n \geq 0$, provando, assim, esta proposição. □

2.3.2 Mais Sobre a Notação $\mathcal{P}f$

Nesta subseção vamos procurar esclarecer um pouco mais o significado da notação $\mathcal{P}f$, introduzida em 2.10, assim como dar uma justificativa para o limite $\lim_{n \rightarrow \infty} \|\mathcal{P}^n f - \langle f \rangle_\mu\|_{2,\mu} = 0$.

Lembremos que, pela definição de valor esperado 2.3.1, temos que

$$\langle f \rangle_\mu = \sum_{i \in S} f(i) (\mu)_i = f(1) (\mu)_1 + f(2) (\mu)_2 + \cdots + f(N) (\mu)_N. \quad (2.22)$$

Por outro lado, por 2.10 temos, para cada $n \in \mathbb{N}$ e $i \in S$, que

$$\mathcal{P}^n f(i) = \sum_{j \in S} f(j) (P^n)_{ij} = f(1) (P^n)_{i1} + f(2) (P^n)_{i2} + \cdots + f(N) (P^n)_{iN}. \quad (2.23)$$

Como, para cada $i \in S$, $\sum_{j \in S} (P^n)_{ij} = 1$ (já que \mathcal{P}^n também é estocástica), podemos identificar $\mathcal{P}^n f(i)$ como sendo o valor esperado de $f(i)$ em relação à distribuição de probabilidades $\omega_i := ((P^n)_{ij})_{j \in S}$, para cada $i \in S$. Mais do que isto, como estamos supondo \mathcal{P} irredutível e μ é distribuição estacionária (pois \mathcal{P} é reversível com respeito a μ , ver corolário 2.2.4) temos, pelo teorema 1.3.22, que para cada $(i, j) \in S^2$

$$\lim_{n \rightarrow \infty} (P^n)_{ij} = (\mu)_j.$$

Juntando esta última expressão com a equação 2.23 somos levados a concluir que

$$\lim_{n \rightarrow \infty} \mathcal{P}^n f(i) = f(1)(\mu)_1 + f(2)(\mu)_2 + \cdots + f(N)(\mu)_N = \langle f \rangle_\mu,$$

onde 1 segue da equação 2.22.

Logo, faz sentido perguntar sobre a validade do limite $\lim_{n \rightarrow \infty} \|\mathcal{P}^n f - \langle f \rangle_\mu\|_{2,\mu} = 0$ e concluímos, desta forma, nossa explanação.

2.4 A Forma de Dirichlet

Ao longo desta seção vamos desenvolver ferramentas para que possamos obter, de forma explícita, uma estimativa para o valor ótimo (i.e., os maiores valores) de β em 2.20 (por conseguinte o menor valor de $r = 1 - \beta$) para o caso de processos a tempo discreto.

2.4.1 Forma de Dirichlet e Desigualdade de Poincaré

Nossa primeira tarefa é obter uma outra expressão (mais conveniente) para β , em particular vamos trabalhar um pouco mais a expressão 2.21. Para tal, iniciamos com a seguinte proposição.

Proposição 2.4.1

A expressão 2.21 é equivalente a

$$1 - \beta = \sup\{\|\mathcal{P}f\|_{2,\mu} : f \in L_0^2(\mu) \text{ com } \|f\|_{2,\mu} = 1\}, \quad (2.24)$$

onde $L_0^2(\mu) := \{f \in L^2(\mu) : \langle f \rangle_\mu = 0\}$.

◇

Prova Ora, como $L_0^2(\mu) \subset L^2(\mu)$ vem (de forma abreviada) que

$$\sup\{\|\mathcal{P}f - \langle f \rangle_\mu\|_{2,\mu}, L^2(\mu)\} \geq \sup\{\|\mathcal{P}f - \langle f \rangle_\mu\|_{2,\mu}, L_0^2(\mu)\} = \sup\{\|\mathcal{P}f\|_{2,\mu}, L_0^2(\mu)\}.$$

Para provar a desigualdade oposta, seja $f \in L^2(\mu)$ com $\|f\|_{2,\mu} = 1$, então ou $\|f - \langle f \rangle_\mu\| = 0$ e, neste caso, $\|\mathcal{P}f - \langle f \rangle_\mu\|_{2,\mu} = 0$, ou $1 \geq \|f - \langle f \rangle_\mu\| > 0$ e, neste caso, defina $\bar{f} := \frac{f - \langle f \rangle_\mu}{\|f - \langle f \rangle_\mu\|}$. Note que $\langle \bar{f} \rangle_\mu = 0$, ou seja, $\bar{f} \in L_0^2(\mu)$ e $\|\bar{f}\|_{2,\mu} = 1$, daí

$$\|\mathcal{P}f - \langle f \rangle_\mu\|_{2,\mu} \leq \left\| \mathcal{P} \left(\frac{f - \langle f \rangle_\mu}{\|f - \langle f \rangle_\mu\|} \right) \right\|_{2,\mu} = \|\mathcal{P}\bar{f}\|_{2,\mu}.$$

Logo,

$$\sup\{\|\mathcal{P}f - \langle f \rangle_\mu\|_{2,\mu}, L^2(\mu)\} \leq \sup\{\|\mathcal{P}f\|_{2,\mu}, L_0^2(\mu)\}.$$

Portanto, 2.21 é equivalente a 2.24.

□

Prosseguindo, utilizaremos mais adiante a seguinte proposição.

Proposição 2.4.2

Vale a igualdade

$$\begin{aligned} \sup\{\|\mathcal{P}f\|_{2,\mu} : f \in L_0^2(\mu) \text{ com } \|f\|_{2,\mu} = 1\} = \\ \sup\{|\langle f, \mathcal{P}f \rangle_{2,\mu}| : f \in L_0^2(\mu) \text{ com } \|f\|_{2,\mu} = 1\}. \end{aligned} \quad (2.25)$$

◇

Prova

Para provar a proposição note que, pela desigualdade de Cauchy-Schwarz (equação 2.9) temos (escrevendo de forma abreviada)

$$\sup\{|\langle f, \mathcal{P}f \rangle_{2,\mu}|\} \leq \sup\{\|f\|_{2,\mu}\|\mathcal{P}f\|_{2,\mu}\} = 1 \sup\{\|\mathcal{P}f\|_{2,\mu}\},$$

onde 1 segue do fato de $\|f\|_{2,\mu} = 1$.

Para provar a desigualdade oposta, seja $f \in L_0^2(\mu)$ com $\|f\|_{2,\mu} = 1$ dada, assumamos que $\|\mathcal{P}f\|_{2,\mu} > 0$, e defina

$$g := \frac{\mathcal{P}f}{\|\mathcal{P}f\|_{2,\mu}}.$$

Então, $g \in L_0^2(\mu)$ já que

$$\langle g \rangle_\mu = \left\langle \frac{\mathcal{P}f}{\|\mathcal{P}f\|_{2,\mu}} \right\rangle_\mu = \frac{1}{\|\mathcal{P}f\|_{2,\mu}} \langle \mathcal{P}f \rangle_\mu = 1 \frac{1}{\|\mathcal{P}f\|_{2,\mu}} \langle f \rangle_\mu = 0,$$

onde 1 está justificada no primeiro parágrafo da prova da proposição 2.3.8. Além disso, é claro que $\|g\|_{2,\mu} = 1$.

Conseqüentemente, se σ denota o supremo no lado direito de 2.25 tem-se,

$$\begin{aligned} 4\|\mathcal{P}f\|_{2,\mu} &= 4 \frac{\|\mathcal{P}f\|_{2,\mu}^2}{\|\mathcal{P}f\|_{2,\mu}} = 4 \langle \mathcal{P}f, \mathcal{P}f \rangle_{2,\mu} \frac{1}{\|\mathcal{P}f\|_{2,\mu}} = 4 \left\langle \frac{\mathcal{P}f}{\|\mathcal{P}f\|_{2,\mu}}, \mathcal{P}f \right\rangle_{2,\mu} = \\ &= 4 \langle g, \mathcal{P}f \rangle_{2,\mu} = 1 \langle (f+g), \mathcal{P}(f+g) \rangle_{2,\mu} - \langle (f-g), \mathcal{P}(f-g) \rangle_{2,\mu} \leq 2 \\ &\leq \sigma(\|f+g\|_{2,\mu}^2 - \|f-g\|_{2,\mu}^2) \leq \sigma(\|f+g\|_{2,\mu}^2 + \|f-g\|_{2,\mu}^2) = 3 \\ &= 2\sigma(\|f\|_{2,\mu}^2 + \|g\|_{2,\mu}^2) = 4 \, 4\sigma. \end{aligned}$$

Donde, $\|\mathcal{P}f\|_{2,\mu} \leq \sigma$ e daí,

$$\sup\{\|\mathcal{P}f\|_{2,\mu}\} \leq \sup\{|\langle f, \mathcal{P}f \rangle_{2,\mu}|\}.$$

Concluimos, assim, que vale a igualdade 2.25.

A igualdade 1 nos cálculos acima se justifica da seguinte forma (omitiremos abaixo o subscrito $2, \mu$ de $\langle \cdot, \cdot \rangle_{2,\mu}$)

$$\begin{aligned} \langle (f+g), \mathcal{P}(f+g) \rangle - \langle (f-g), \mathcal{P}(f-g) \rangle &= 1 \langle f, \mathcal{P}f \rangle + \langle f, \mathcal{P}g \rangle + \langle g, \mathcal{P}f \rangle + \\ &+ \langle g, \mathcal{P}g \rangle - \langle f, \mathcal{P}f \rangle + \langle f, \mathcal{P}g \rangle + \\ &+ \langle g, \mathcal{P}f \rangle - \langle g, \mathcal{P}g \rangle = \\ &= 2 \langle f, \mathcal{P}g \rangle + 2 \langle g, \mathcal{P}f \rangle = 2 \\ &= 2 \langle \mathcal{P}f, g \rangle + 2 \langle g, \mathcal{P}f \rangle = 3 \\ &= 2 \langle g, \mathcal{P}f \rangle + 2 \langle g, \mathcal{P}f \rangle = 4 \langle g, \mathcal{P}f \rangle, \end{aligned}$$

onde 1 vem da (bi) linearidade do produto interno, 2 segue pelo teorema 2.3.4, já 3 vem da comutatividade do produto interno. Justificamos, deste modo, a igualdade 1 dos cálculos anteriormente mencionados.

A desigualdade 2 é justificada do seguinte modo. Como observamos anteriormente, σ denota o supremo no lado direito de 2.25, então, (de forma abreviada) $\forall h \in L^2(\mu)$, temos que

$$\sigma = \sup\{|\langle f, \mathcal{P}f \rangle|\} \geq \left\langle \frac{h}{\|h\|_{2,\mu}}, \mathcal{P} \frac{h}{\|h\|_{2,\mu}} \right\rangle_{2,\mu} = \frac{\langle h, \mathcal{P}h \rangle_{2,\mu}}{\|h\|_{2,\mu}^2}.$$

Donde, $\langle h, \mathcal{P}h \rangle_{2,\mu} \leq \sigma \|h\|_{2,\mu}^2$. Concluindo a justificativa da desigualdade 2.

A igualdade 3 nos cálculos acima é devida a *Identidade do Paralelogramo* (ver exercício 10.12 página 133 de [16]). Finalmente, 4 segue de $\|f\|_{2,\mu}^2 = \|g\|_{2,\mu}^2 = 1$.

□

Pela proposição 2.4.2, podemos escrever 2.24 como

$$1 - \beta = \sup\{|\langle f, \mathcal{P}f \rangle_{2,\mu}| : f \in L_0^2(\mu) \text{ com } \|f\|_{2,\mu}^2 = 1\}. \quad (2.26)$$

Daí, segue a seguinte proposição.

Proposição 2.4.3

Suponhamos que vale a igualdade

$$\beta = \inf\{\langle f, (I - \mathcal{P})f \rangle_{2,\mu} : f \in L_0^2(\mu) \text{ com } \|f\|_{2,\mu}^2 = 1\}, \quad (2.27)$$

e que \mathcal{P} seja *não-negativa definida* (i.e., $\langle f, \mathcal{P}f \rangle_{2,\mu} \geq 0 \forall f \in L^2(2, \mu)$). Então, vale 2.26.

◇

Prova

Para provarmos a proposição, devemos observar primeiramente que, $\forall f \in L_0^2(\mu)$,

$$\begin{aligned} \langle f, (I - \mathcal{P})f \rangle_{2,\mu} &= \langle f, If - \mathcal{P}f \rangle_{2,\mu} = \langle f, f \rangle_{2,\mu} - \langle f, \mathcal{P}f \rangle_{2,\mu} = \\ &= \|f\|_{2,\mu}^2 - \langle f, \mathcal{P}f \rangle_{2,\mu} = 1 - \langle f, \mathcal{P}f \rangle_{2,\mu}. \end{aligned}$$

Daí, supondo 2.27 temos (de forma abreviada) que

$$\beta = \inf\{\langle f, (I - \mathcal{P})f \rangle_{2,\mu}\} = \inf\{1 - \langle f, \mathcal{P}f \rangle_{2,\mu}\} = 1 - \sup\{\langle f, \mathcal{P}f \rangle_{2,\mu}\}.$$

Donde vem $1 - \beta = \sup\{|\langle f, \mathcal{P}f \rangle_{2,\mu}|\}$.

□

Definição 2.4.4

O produto interno que aparece em 2.27 é conhecido como a *Forma de Dirichlet* associada à cadeia de Markov \mathcal{P} , reversível com respeito a μ . Utilizaremos a notação $\xi(f, f) := \langle f, (I - \mathcal{P})f \rangle_{2,\mu}$.

◇

Corolário 2.4.5

$$\beta = \inf\{\langle f, (I - \mathcal{P})f \rangle_{2,\mu} : f \in L^2(\mu) \text{ com } Var_\mu(f) := \|f - \langle f \rangle_\mu\|_{2,\mu}^2 = 1\}. \quad (2.28)$$

◇

Prova

Note que

$$(I - \mathcal{P})(f - \langle f \rangle_\mu) = f - \langle f \rangle_\mu - \mathcal{P}f + \langle f \rangle_\mu = f - \mathcal{P}f = (I - \mathcal{P})f.$$

Defina, $\forall f \in L^2(\mu)$,

$$\bar{f} := \frac{f - \langle f \rangle_\mu}{\|f - \langle f \rangle_\mu\|_{2,\mu}}.$$

É fácil ver que $\langle \bar{f} \rangle_\mu = 0$ e $\|\bar{f}\|_{2,\mu} = 1$.Podemos então, aplicar a proposição 2.4.3 a \bar{f} . Daí,

$$\bar{f} \in L_0^2(\mu) \Leftrightarrow f - \langle f \rangle_\mu \in L_0^2(\mu) \Rightarrow f \in L^2(\mu).$$

Além disso, pelo que vimos no início desta prova

$$\begin{aligned} \langle \bar{f}, (I - \mathcal{P})\bar{f} \rangle_{2,\mu} &= \left\langle \frac{f - \langle f \rangle_\mu}{\|f - \langle f \rangle_\mu\|_{2,\mu}}, (I - \mathcal{P}) \frac{f - \langle f \rangle_\mu}{\|f - \langle f \rangle_\mu\|_{2,\mu}} \right\rangle_{2,\mu} = \\ &= \frac{1}{\|f - \langle f \rangle_\mu\|_{2,\mu}^2} \langle f - \langle f \rangle_\mu, (I - \mathcal{P})(f - \langle f \rangle_\mu) \rangle_{2,\mu} =^1 \\ &= \frac{1}{Var_\mu(f)} \langle f - \langle f \rangle_\mu, (I - \mathcal{P})f \rangle_{2,\mu} = \frac{1}{Var_\mu(f)} \langle f, (I - \mathcal{P})f \rangle_{2,\mu} - \\ &- \frac{1}{Var_\mu(f)} \langle \langle f \rangle_\mu, (I - \mathcal{P})f \rangle_{2,\mu} = \frac{1}{Var_\mu(f)} \langle f, (I - \mathcal{P})f \rangle_{2,\mu} - \\ &- \frac{1}{Var_\mu(f)} \langle f \rangle_\mu \langle 1, f \rangle_{2,\mu} + \frac{1}{Var_\mu(f)} \langle f \rangle_\mu \langle 1, \mathcal{P}f \rangle_{2,\mu} = \\ &=^2 \frac{1}{Var_\mu(f)} \langle f, (I - \mathcal{P})f \rangle_{2,\mu}, \end{aligned}$$

onde 1 segue da definição 2.3.1 e 2 segue da mesma definição e do fato que $\langle \mathcal{P}f \rangle_\mu = \langle f \rangle_\mu$, como já vimos anteriormente.

Portanto, (da proposição 2.4.3 aplicada à \bar{f})

$$\beta = \inf\{\langle \bar{f}, (I - \mathcal{P})\bar{f} \rangle_{2,\mu} : \bar{f} \in L_0^2(\mu) \text{ com } \|\bar{f}\|_{2,\mu} = 1\},$$

é equivalente a

$$\beta = \inf\left\{\frac{1}{Var_\mu(f)} \langle f, (I - \mathcal{P})f \rangle_{2,\mu} : f \in L^2(\mu) \text{ com } \|f\|_{2,\mu} = 1\right\},$$

que é, finalmente, equivalente a escrever

$$\beta = \inf\{\langle f, (I - \mathcal{P})f \rangle_{2,\mu} : f \in L^2(\mu) \text{ com } Var_\mu(f) = 1\}.$$

□

Proposição 2.4.6

Para todo processo de Markov reversível com respeito a uma dada distribuição inicial μ vale que

$$\beta = \beta_+ \wedge \beta_-, \quad (2.29)$$

onde

$$\beta_{\pm} = \inf\{\langle f, (I \mp \mathcal{P})f \rangle_{2,\mu} : f \in L^2(\mu) \text{ com } \text{Var}_{\mu}(f) = 1\}. \quad (2.30)$$

◇

Prova

Primeiramente, como vimos na prova da proposição 2.4.3

$$\langle f, (I - \mathcal{P})f \rangle_{2,\mu} = 1 - \langle f, \mathcal{P}f \rangle_{2,\mu}$$

e, analogamente

$$\langle f, (I + \mathcal{P})f \rangle_{2,\mu} = 1 + \langle f, \mathcal{P}f \rangle_{2,\mu}.$$

Note agora que, pelas expressões que definem β_{\pm} tem-se (de forma abreviada)

$$\begin{aligned} \beta_+ \wedge \beta_- &= \inf\{\langle f, (I - \mathcal{P})f \rangle_{2,\mu}\} \wedge \inf\{\langle f, (I + \mathcal{P})f \rangle_{2,\mu}\} =^1 \\ &= \inf\{1 - \langle f, \mathcal{P}f \rangle_{2,\mu}\} \wedge \inf\{1 + \langle f, \mathcal{P}f \rangle_{2,\mu}\}. \end{aligned}$$

Daí, existem duas possibilidades, ou $\langle f, \mathcal{P}f \rangle_{2,\mu} \leq 0$ e, neste caso

$$\langle f, \mathcal{P}f \rangle_{2,\mu} = -|\langle f, \mathcal{P}f \rangle_{2,\mu}|$$

e contas acima tornar-se-ão

$$\begin{aligned} \beta_+ \wedge \beta_- &= \inf\{1 + |\langle f, \mathcal{P}f \rangle_{2,\mu}|\} \wedge \inf\{1 - |\langle f, \mathcal{P}f \rangle_{2,\mu}|\} =^2 \\ &= \inf\{1 - |\langle f, \mathcal{P}f \rangle_{2,\mu}|\} = 1 - \sup\{|\langle f, \mathcal{P}f \rangle_{2,\mu}|\} =^3 \beta, \end{aligned}$$

ou $\langle f, \mathcal{P}f \rangle_{2,\mu} \geq 0$ e, neste caso

$$\langle f, \mathcal{P}f \rangle_{2,\mu} = |\langle f, \mathcal{P}f \rangle_{2,\mu}|$$

e contas acima tornar-se-ão

$$\begin{aligned} \beta_+ \wedge \beta_- &= \inf\{1 - |\langle f, \mathcal{P}f \rangle_{2,\mu}|\} \wedge \inf\{1 + |\langle f, \mathcal{P}f \rangle_{2,\mu}|\} =^4 \\ &= \inf\{1 - |\langle f, \mathcal{P}f \rangle_{2,\mu}|\} = 1 - \sup\{|\langle f, \mathcal{P}f \rangle_{2,\mu}|\} =^5 \beta, \end{aligned}$$

onde a igualdade 1 se deve ao arrazoado feito no início desta prova; 2 e 4 são evidentes pois $1 - |\langle f, \mathcal{P}f \rangle_{2,\mu}| \leq 1 + |\langle f, \mathcal{P}f \rangle_{2,\mu}|$, 3 e 5 seguem da equação 2.26, concluindo assim a prova desta proposição.

□

Observação 2.4.7

Levando em conta as expressões de β_{\pm} podemos observar que, no caso particular em que \mathcal{P} é não-negativa definida ($\langle f, \mathcal{P}f \rangle_{2,\mu} \geq 0 \forall f \in L^2(\mu)$: notação $\mathcal{P} \geq 0$), teremos (de forma abreviada e utilizando os mesmos argumentos empregados na prova da proposição 2.4.6) que

$$\beta_+ = 1 - \sup\{\langle f, \mathcal{P}f \rangle_{2,\mu}\} \leq 1,$$

e

$$\beta_- = 1 + \sup\{\langle f, \mathcal{P}f \rangle_{2,\mu}\} \geq 1.$$

Portanto, vale que $\beta_+ \leq 1 \leq \beta_-$, daí $\beta = \beta_+ \wedge \beta_- = \beta_+$. Tal observação é relevante pois para o caso de processos reversíveis a tempo contínuo mostraremos (ver lema 3.3.1) que o semigrupo estocástico a tempo contínuo goza da propriedade $\mathcal{P}^t \geq 0$. Portanto, tendemos a dar mais ênfase ao parâmetro β_+ .

◇

Lema 2.4.8

Se \mathcal{P} é uma matriz de transição de probabilidades irredutível, reversível com respeito a uma distribuição inicial μ , então o período de \mathcal{P} é 1 ou 2. Além disso, o período é 2 se e só se, existe uma $f \in L^2(\mu) \setminus \{0\}$ tal que $f = -\mathcal{P}f$.

◇

A prova deste lema é do mesmo nível das já apresentadas neste trabalho e pode ser encontrada em [24].

Proposição 2.4.9

Defina β_- como na proposição 2.4.6. Então, $\beta_- = 0$ a não ser que a cadeia seja *aperiódica*.

◇

Prova

Trivialmente $\beta_- \geq 0$. Pelo lema 2.4.8 o período da cadeia só pode ser 1 ou 2, daí se for 2, i.e., se ela for periódica vem que existe uma $f_1 \in L^2(\mu) \setminus \{0\}$ tal que $f_1 = -\mathcal{P}f_1$. Agora, voltando com esta f_1 na definição de β_- teremos que (de forma abreviada)

$$\begin{aligned} \beta_- &= \inf\{\langle f, (I + \mathcal{P})f \rangle_{2,\mu}\} \leq \langle f_1, f_1 + \mathcal{P}f_1 \rangle_{2,\mu} = \\ &= \langle f_1, -\mathcal{P}f_1 + \mathcal{P}f_1 \rangle_{2,\mu} = 0, \end{aligned}$$

$\Rightarrow \beta_- \leq 0$ e, portanto, $\beta_- = 0$ e não ser que a cadeia reversível seja aperiódica.

□

A expressão para β dada por 2.28 no corolário 2.4.5 é bem importante do ponto de vista computacional. De fato, todo nosso trabalho nesta seção vai ao encontro da idéia de expressar β de modo que seja conveniente sua manipulação do ponto de vista algébrico.

Com este intento, obteremos uma expressão um pouco mais explícita para a Forma de Dirichlet $\xi(f, f) := \langle f, (I - \mathcal{P})f \rangle_{2,\mu}$, como irá mostrar o seguinte teorema.

Teorema 2.4.10

A Forma de Dirichlet $\xi(f, f)$ pode ser escrita como

$$\xi(f, f) = \frac{1}{2} \sum_{i \neq j} (\mu)_i (P)_{ij} (f(j) - f(i))^2. \quad (2.31)$$

◇

Prova

Temos, pela definição de produto interno (definição 2.3.1), por um lado que

$$\begin{aligned} \xi(f, f) &= \langle f, (I - \mathcal{P})f \rangle_{2, \mu} = \sum_{i \in S} f(i) (\mu)_i [(I - \mathcal{P})f](i) = \\ &= \sum_{i \in S} f(i) (\mu)_i [f(i) - \mathcal{P}f](i) = \sum_{i \in S} f(i) (\mu)_i \left[f(i) - \sum_{j \in S} f(j) (P)_{ij} \right] =^1 \\ &= \sum_{i \in S} f(i) (\mu)_i \left[f(i) \sum_{j \in S} (P)_{ij} - \sum_{j \in S} f(j) (P)_{ij} \right] = \\ &= \sum_{i \in S} \sum_{j \in S} f(i) (\mu)_i (P)_{ij} (f(i) - f(j)) = \sum_{(i, j) \in S^2} f(i) (\mu)_i (P)_{ij} (f(i) - f(j)), \end{aligned}$$

onde 1 vem do fato que $\sum_j (P)_{ij} = 1$ para cada $i \in S$.

Por outro lado,

$$\begin{aligned} \xi(f, f) &= \sum_{(i, j) \in S^2} f(i) (\mu)_i (P)_{ij} (f(i) - f(j)) =^1 \sum_{(i, j) \in S^2} f(i) (\mu)_j (P)_{ji} (f(i) - f(j)) = \\ &=^2 \sum_{(i, j) \in S^2} f(j) (\mu)_i (P)_{ij} (f(j) - f(i)), \end{aligned}$$

onde 1 segue da condição do balanço detalhado 2.4 e 2 vem da mudança de índice.

Somando-se estas duas expressões encontradas para $\xi(f, f)$ temos

$$\begin{aligned} 2\xi(f, f) &= \sum_{(i, j) \in S^2} f(i) (\mu)_i (P)_{ij} (f(i) - f(j)) + \sum_{(i, j) \in S^2} f(j) (\mu)_i (P)_{ij} (f(j) - f(i)) = \\ &= \sum_{(i, j) \in S^2} (\mu)_i (P)_{ij} [f(i)(f(i) - f(j)) + f(j)(f(j) - f(i))] = \\ &= \sum_{(i, j) \in S^2} (\mu)_i (P)_{ij} [f(i)^2 - f(i)f(j) + f(j)^2 - f(j)f(i)] = \\ &= \sum_{(i, j) \in S^2} (\mu)_i (P)_{ij} [f(i)^2 - 2f(i)f(j) + f(j)^2] = \\ &=^1 \sum_{i \neq j} (\mu)_i (P)_{ij} (f(j) - f(i))^2, \end{aligned}$$

onde 1 segue do fato que o ponto $(i, i) \in S^2$ não contribui para a soma.

Por fim,

$$\xi(f, f) = \frac{1}{2} \sum_{i \neq j} (\mu)_i (P)_{ij} (f(j) - f(i))^2.$$

□

Na maioria dos casos, quando mencionarmos *Forma de Dirichlet*, estaremos pensando na equação 2.31.

De forma análoga ao que fizemos acima é possível mostrar que

$$\langle f, (I + \mathcal{P})f \rangle_{2,\mu} = \tilde{\xi}(f, f) = \frac{1}{2} \sum_{(i,j)} (\mu)_i (P)_{ij} (f(j) + f(i))^2. \quad (2.32)$$

Observação 2.4.11

A forma quadrática $\xi(f, f)$, dada por 2.31, é uma analogia discreta da forma quadrática $\frac{1}{2} \int |\nabla f|^2(x) dx$ introduzida por Dirichlet, justificando assim o nome dado a $\xi(f, f)$.

Prosseguindo com esta analogia, podemos interpretar β como a *Constante de Poincaré* na *Desigualdade de Poincaré*

$$\beta \text{Var}_\mu(f) \leq \xi(f, f), \quad (2.33)$$

$\forall f \in L^2(\mu)$.

◇

Pelo corolário 2.4.5 e teorema 2.4.10 β será expresso, finalmente, como

$$\beta = \inf\{\xi(f, f) : f \in L^2(\mu) \text{ com } \text{Var}_\mu(f) = 1\} \wedge \inf\{\tilde{\xi}(f, f) : f \in L^2(\mu) \text{ com } \text{Var}_\mu(f) = 1\}. \quad (2.34)$$

A desigualdade 2.33 vem do fato de β ser expresso como um ínfimo, donde que $\beta \leq \frac{\xi(f,f)}{\text{Var}_\mu(f)}$.

2.4.2 Estimando β_\pm

Esta parte do texto se destina a encontrar uma estimativa para β_+ em termos da geometria do *Grafo* de transição (ver definição 2.4.13) associado à matriz de transição de probabilidades \mathcal{P} .

Com o objetivo de estimar β_+ utilizando uma desigualdade análoga a 2.33, devemos escrever $\text{Var}_\mu(f)$ de uma forma mais conveniente, como segue no lema abaixo.

Lema 2.4.12

Seja $f \in L^2(\mu)$. Então,

$$\text{Var}_\mu(f) = \frac{1}{2} \sum_{ij} (f(i) - f(j))^2 (\mu)_i (\mu)_j. \quad (2.35)$$

◇

Prova

No que consiste, temos que

$$\begin{aligned}
\frac{1}{2} \sum_{ij} (f(i) - f(j))^2 (\mu)_i (\mu)_j &= \frac{1}{2} \sum_{i \in S} \sum_{j \in S} (f(i)^2 + f(j)^2 - 2f(i)f(j)) (\mu)_i (\mu)_j = \\
&= \frac{1}{2} \sum_{i \in S} \sum_{j \in S} (\mu)_i (\mu)_j f(i)^2 + \frac{1}{2} \sum_{i \in S} \sum_{j \in S} (\mu)_i (\mu)_j f(j)^2 - \\
&- \sum_{i \in S} \sum_{j \in S} (\mu)_i (\mu)_j f(i) f(j) = \frac{1}{2} \sum_{i \in S} (\mu)_i f(i)^2 \sum_{j \in S} (\mu)_j + \\
&+ \frac{1}{2} \sum_{i \in S} f(i)^2 (\mu)_i \sum_{j \in S} (\mu)_j - \sum_{i \in S} (\mu)_i f(i) \sum_{j \in S} (\mu)_j f(j) =^1 \\
&= \frac{1}{2} \langle f^2 \rangle_\mu + \frac{1}{2} \langle f^2 \rangle_\mu - \langle f \rangle_\mu \langle f \rangle_\mu = \\
&= \langle f^2 \rangle_\mu - \langle f \rangle_\mu^2 = \text{Var}_\mu(f),
\end{aligned}$$

onde 1 vem do fato de μ ser vetor de probabilidade e da definição de valor esperado 2.3.1. Concluindo, desta forma, a prova do lema.

□

A conveniência de 2.35 é que ela expressa $\text{Var}_\mu(f)$ em termos da diferença entre valores de f em diferentes pontos de S (lembre que $\xi(f, f)$ também é dado em termos de tais diferenças). Todavia, as diferenças que aparecem em $\xi(f, f)$ são, unicamente, entre os valores de f em pares de pontos (i, j) para os quais $(P)_{ij} > 0$, enquanto que, 2.35 é dada para qualquer $(i, j) \in S^2$.

Antes de prosseguirmos, daremos alguns conceitos cuja definição segue abaixo.

Definição 2.4.13

Grosso modo, um *Grafo de Transição* associado a \mathcal{P} , é um diagrama cujas *Vértices* representam os elementos de S (i.e., os estados) e cujas *Arestas* orientadas $e = (i_1, i_2)$ representam a probabilidade de transição de um estado i_1 para um outro estado i_2 , ou seja, representam $(P)_{i_1 i_2}$, conforme mostra a figura 2.1.

Um *Caminho Acessível* é uma $(n + 1)$ -upla $p(i, j) = (k_0, \dots, k_n) \in S^{n+1}$ tal que, $k_0 = i$ é o *vértice inicial*, $k_n = j$ é o *vértice final* e, além disso, $(P)_{k_{m-1} k_m} > 0$ para cada $1 \leq m \leq n$. Denotaremos por A o conjunto das arestas.

◇

Como vimos na definição 2.4.13, denotaremos por e a aresta $e = (i, j)$, conectando os vértices i e j (para uma melhor compreensão veja exemplo 2.5). Definamos então, a grandeza

$$\rho(e) := (\mu)_k (P)_{kl}, \quad (2.36)$$

quando $e = (k, l)$.

Segue então a seguinte definição.

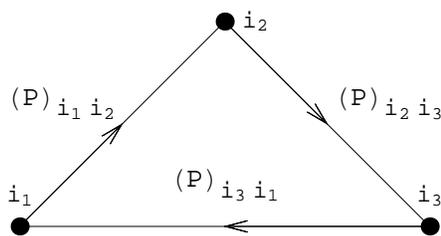


Figura 2.1 - Grafo de transição associado a \mathcal{P} com vértices i_1 , i_2 e i_3 e arestas (i_1, i_2) , (i_2, i_3) e (i_3, i_1) .

Definição 2.4.14

Vamos definir uma escolha de caminhos acessíveis \mathcal{P} como segue. Para cada par ordenado de estados distintos (i, j) selecionamos um, e somente um, caminho acessível $p(i, j)$. Seja, $\mathcal{P} = \{p(i, j) : (i, j) \in S^2 \setminus D\}$, onde $D = \{(i, j) \in S^2 : i = j\}$ uma tal escolha.

Definimos o *Coefficiente de Poincaré* por

$$W(\mathcal{P}) := \sup_{e \in A} \sum_{\{p \in \mathcal{P} \text{ tal que } e \in p\}} \omega(p), \quad (2.37)$$

onde

$$\omega(p) := (\mu)_i (\mu)_j \sum_{e' \in p} \frac{1}{\rho(e')}, \text{ com } p = p(i, j). \quad (2.38)$$

◇

O próximo teorema vai nos dar uma estimativa para $Var_\mu(f)$ em termos de $\xi(f, f)$ e sua prova será baseada no que vimos acima.

Teorema 2.4.15

Seja \mathcal{P} uma matriz de transição irredutível, reversível com respeito a μ em S finito. Então,

$$Var_\mu(f) \leq W(\mathcal{P})\xi(f, f). \quad (2.39)$$

◇

Prova

Para provar este teorema, devemos nos concentrar na definição 2.4.14. Portanto, é necessário escolher, para cada $(i, j) \in S^2 \setminus D$, um único caminho acessível $p(i, j) = (k_0, \dots, k_n) \in S^{n+1}$ com $k_0 = i$ e $k_n = j$. Observando então que

$$(f(i) - f(j))^2 = \left(\sum_{e \in p(i, j)} \Delta_e f \right)^2,$$

onde a soma em e é tomada sobre os seguimentos orientados (k_{m-1}, k_m) para $1 \leq m \leq n$ no caminho $p(i, j)$, e, para $e = (k, l)$, $\Delta_e f := f(l) - f(k)$. Daí, pela

Desigualdade de Schwarz 1.3.21 temos, multiplicando o argumento da soma acima por $(\alpha(e)/\rho(e))^{1/2}(\rho(e)/\alpha(e))^{1/2}$, (note que $\rho(e) > 0$, pois estamos assumindo $(\mu)_i > 0 \forall i \in S$ e $(P)_{kl} > 0$ já que $p(i, j)$ é um caminho acessível, já $\alpha(e)$ é um coeficiente genérico qualaquer) que

$$\begin{aligned} \left(\sum_{e \in p(i, j)} \frac{\alpha(e)^{1/2}}{\rho(e)^{1/2}} \frac{\rho(e)^{1/2}}{\alpha(e)^{1/2}} \Delta_e f \right)^2 &\leq \left[\sqrt{\sum_{e \in p(i, j)} \left(\frac{\alpha(e)^{1/2}}{\rho(e)^{1/2}} \right)^2} \sqrt{\sum_{e \in p(i, j)} \left(\frac{\rho(e)^{1/2}}{\alpha(e)^{1/2}} \Delta_e f \right)^2} \right]^2 = \\ &= \left(\sum_{e \in p(i, j)} \frac{\alpha(e)}{\rho(e)} \right) \left(\sum_{e \in p(i, j)} \frac{\rho(e)}{\alpha(e)} (\Delta_e f)^2 \right). \end{aligned}$$

Portanto, tomando desde já $\alpha(e) \equiv 1$ (por conveniência, já que no caso de maior interesse, a saber, processos a tempo contínuo é esta escolha que faremos) e para qualquer seleção P de caminhos acessíveis temos, pelo lema 2.4.12, que

$$\begin{aligned} Var_\mu(f) &= \frac{1}{2} \sum_{ij} (f(i) - f(j))^2 (\mu)_i (\mu)_j \leq^1 \\ &\leq \frac{1}{2} \sum_{p \in P} (\mu)_i (\mu)_j \left(\sum_{e \in p(i, j)} \frac{1}{\rho(e)} \right) \sum_{e \in p(i, j)} (\Delta_e f)^2 \rho(e) =^2 \\ &= \frac{1}{2} \sum_{p \in P} \omega(p) \sum_{e \in p(i, j)} (\Delta_e f)^2 \rho(e) = \frac{1}{2} \sum_{e \in P} (\Delta_e f)^2 \rho(e) \left(\sum_{p \ni e} \omega(p) \right) \leq^3 \\ &\leq W(P) \xi(f, f), \end{aligned}$$

onde 1 vem do que foi exposto no início desta prova, 2 se deve a equação 2.38, e a soma entre parênteses deve ser feito sobre os $p \in P$ nos quais e aparece, já 3 vem do fato de $W(P)$ ser definido em termos de um supremo (ver equação 2.37) e também não é difícil ver que a primeira soma é, de fato, a Forma de Dirichlet $\xi(f, f)$ dada por 2.31.

Concluimos, com isto, que

$$Var_\mu(f) \leq W(P) \xi(f, f).$$

□

Este teorema (equação 2.39) nos diz que $\frac{\xi(f, f)}{Var_\mu(f)} \geq \frac{1}{W(P)}$. Com isto, somos levados a concluir (pela expressão 2.30 para β_+) que

$$\beta_+ = \inf \left\{ \frac{\xi(f, f)}{Var_\mu(f)} : f \in L^2(\mu) \right\} \geq \inf \left\{ \frac{1}{W(P)} : f \in L^2(\mu) \right\} = \frac{1}{W(P)}.$$

Logo, para qualquer escolha de caminhos acessíveis P vale que.

$$\beta_+ \geq \frac{1}{W(P)}. \quad (2.40)$$

Na maior parte das aplicações de 2.40 escolhemos P conforme a conveniência da particular situação em consideração.

De forma totalmente análoga ao que fizemos acima, considerando agora o caso em que a cadeia é aperiódica, podemos escolher uma seleção de caminhos acessíveis \bar{P} , onde cada caminho acessível é da forma $p(i, j) = (k_0 = i, \dots, k_{2n+1} = j)$, e daí definir

$$\bar{\omega}(p) = (\mu)_i \sum_{e \in p} \frac{1}{\rho(e)}, \quad (2.41)$$

com

$$\bar{W}(\bar{P}) = \sup_{e \in A} \sum_{\{p \in \bar{P} : e \in p\}} \bar{\omega}(p). \quad (2.42)$$

Logo, a estimativa para β_- fica

$$\beta_- \geq \frac{2}{\bar{W}(\bar{P})}. \quad (2.43)$$

Exemplo 2.5

Neste exemplo vamos mostrar como estimar β por meio das relações 2.40 e 2.43 que nos dam uma cota inferior para β (por conseguinte uma cota superior para a taxa de convergência $r = 1 - \beta$ da cadeia).

Consideremos a seguinte cadeia de Markov,

$$P = \begin{bmatrix} 1/2 & 1/4 & 1/4 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{bmatrix}.$$

Tal matriz de transição é irredutível e reversível com respeito a $\mu = (1/2, 1/4, 1/4)$ (como se verifica facilmente).

Seja $S = \{1, 2, 3\}$ o espaço de estados associado. Note que, $\forall i \neq j$ tem-se que $(P)_{ij} > 0$. Daí, para calcular $W(P)$ devemos escolher (conforme definição 2.4.14), para cada $i \neq j$, um único caminho acessível $p(i, j)$ para construir a seleção P de tais caminhos. Observe que são possíveis muitas escolhas para P , das quais destacamos duas:

$$P_1 = \{p_1(1, 2) = (1, 2), p_1(2, 3) = (2, 3), p_1(3, 2) = (3, 2), p_1(2, 1) = (2, 1), \\ p_1(1, 3) = (1, 3), p_1(3, 1) = (3, 1)\}.$$

e

$$P_2 = \{p_2(1, 2) = (1, 2), p_2(2, 3) = (2, 3), p_2(3, 2) = (3, 2), p_2(2, 1) = (2, 1), \\ p_2(1, 3) = (1, 2, 3), p_2(3, 1) = (3, 2, 1)\}.$$

Além disso, o conjunto das arestas é:

$$A = \{e_1 = (1, 2), e_2 = (2, 3), e_3 = (3, 2), e_4 = (2, 1), e_5 = (1, 3), e_6 = (3, 1)\}.$$

Posto isto, iniciaremos calculando $W(P_1)$ dado pela relação 2.37, para tal devemos calcular os $\omega(p_1)$ em 2.38. Com efeito,

$$\omega(p_1(1, 2)) = (\mu)_1(\mu)_2 \sum_{e \in p_1(1,2)} \frac{1}{\rho(e)} = (\mu)_1(\mu)_2 \frac{1}{\rho(e_1)} = \frac{(\mu)_1(\mu)_2}{(\mu)_1(P)_{12}} = 1;$$

$$\omega(p_1(2, 3)) = (\mu)_2(\mu)_3 \sum_{e \in p_1(2,3)} \frac{1}{\rho(e)} = (\mu)_2(\mu)_3 \frac{1}{\rho(e_2)} = \frac{(\mu)_2(\mu)_3}{(\mu)_1(P)_{23}} = 1/2,$$

e de forma totalmente análoga,

$$\omega(p_1(3, 2)) = 1/2, \quad \omega(p_1(2, 1)) = 1, \quad \omega(p_1(1, 3)) = 1, \quad \omega(p_1(3, 1)) = 1.$$

Portanto, por 2.37

$$W(P_1) = \sup_{e \in A} \sum_{p_1 \ni e} \omega(p_1) = 1 \max\{1, 1/2, 1/2, 1, 1, 1\} = 1,$$

onde 1 se deve ao fato que cada caminho é formado por uma única aresta, donde que $\sum_{p_1 \ni e} \omega(p_1) = \omega(p_1)$.

Procederemos de forma análoga para o cálculo de $W(P_2)$. Com efeito,

$$\begin{aligned} \omega(p_2(1, 2)) &= \omega(p_1(1, 2)) = 1, \\ \omega(p_2(2, 3)) &= \omega(p_1(2, 3)) = 1/2, \\ \omega(p_2(3, 2)) &= \omega(p_1(3, 2)) = 1/2, \\ \omega(p_2(2, 1)) &= \omega(p_1(2, 1)) = 1. \end{aligned}$$

Entretanto, nos dois últimos caminhos restantes aparece uma mudança interessante.

$$\begin{aligned} \omega(p_2(1, 3)) &= (\mu)_1(\mu)_3 \sum_{e \in p_2(1,3)} \frac{1}{\rho(e)} = (\mu)_1(\mu)_3 \left(\frac{1}{\rho(e_1)} + \frac{1}{\rho(e_2)} \right) = \\ &= (\mu)_1(\mu)_3 \left(\frac{1}{(\mu)_1(P)_{12}} + \frac{1}{(\mu)_2(P)_{23}} \right) = 2, \end{aligned}$$

e

$$\begin{aligned} \omega(p_2(3, 1)) &= (\mu)_3(\mu)_1 \sum_{e \in p_2(3,1)} \frac{1}{\rho(e)} = (\mu)_3(\mu)_1 \left(\frac{1}{\rho(e_3)} + \frac{1}{\rho(e_4)} \right) = \\ &= (\mu)_3(\mu)_1 \left(\frac{1}{(\mu)_3(P)_{23}} + \frac{1}{(\mu)_2(P)_{21}} \right) = 2. \end{aligned}$$

Observando P_2 e o conjunto de arestas A vemos que a aresta e_1 aparece no primeiro e no quinto caminho, e_2 aparece no segundo e no quinto caminho, e_3 aparece no terceiro e sexto caminho, e_4 no quarto e sexto caminho, e_5 e e_6 apenas nos respectivos caminhos quinto e sexto. Logo,

$$W(P_2) = \max\{1 + 2, (1/2) + 2, (1/2) + 2, 1 + 2, 2, 2\} = \max\{3, 5/2, 5/2, 3, 2, 2\} = 3.$$

Além disso, pela relação 2.41 temos que (note que apenas P_1 é possível pois, para este caso, nossos caminhos devem ser da forma $p(i, j) = (k_0 = i, \dots, k_{2n+1})$ para $n \geq 0$)

$$\bar{\omega}(p_1(1, 2)) = 4, \quad \bar{\omega}(p_1(2, 3)) = 2, \quad \bar{\omega}(p_1(3, 2)) = 2, \quad \bar{\omega}(p_1(2, 1)) = 2,$$

$$\bar{\omega}(p_1(1, 3)) = 4, \quad \bar{\omega}(p_1(3, 1)) = 2.$$

Donde que, pela relação 2.42 vem

$$\overline{W}(P_1) = \max\{2, 4\} = 4.$$

Agora, se escolhermos P_1 como seleção de caminhos acessíveis temos, por 2.40 que $\beta_+ \geq 1$ e por 2.43 que $\beta_- \geq 1/2$, logo por 2.29 segue que $\beta \geq \min\{1, 1/2\} = 1/2$. Por outro lado, se consideramos P_2 como sendo outra escolha de cominhos acessíveis, vem que $\beta_+ \geq 1/3$ e, como antes, $\beta_- \geq 1/2$. Portanto, $\beta \geq \min\{1/2, 1/3\} = 1/3$. Ou seja, em nosso exemplo particular, vamos tomar P_2 como nossa escolha (de modo a minimizar β , i.e., maximizar $r = 1 - \beta$). Daí, por 2.40 temos, finalmente que $\beta \geq 1/3$ (de fato $\beta = 1/2$).

□

Note também que, $W(P)$ dado por 2.37, pode ser escrito como

$$W(P) = \sup_{e \in A} \sum_{\{p \in P : e \in p\}} \sum_{e' \in p} \frac{(\mu(p))_- - (\mu(p))_+}{\rho(e')}, \quad (2.44)$$

onde $(\mu(p))_- = (\mu)_i$ e $(\mu(p))_+ = (\mu)_j$, quando p inicia em i e termina em j .

Em geral, 2.40 não dá informação. De fato, embora a irreduzibilidade de \mathcal{P} garanta a existência de pelo menos um caminho conectando todo par de pontos, quando S é infinito não há garantia que P possa ser escolhido de modo que $W(P) < \infty$. Além disso, mesmo que S seja finito, e então $W(P) < \infty$ para toda escolha de P , somente uma escolha adequada produz uma boa estimativa em 2.40.

3 Processos de Markov Reversíveis e Forma de Dirichlet-Tempo Contínuo

3.1 Introdução

Vamos estender as idéias vistas no capítulo 2 sobre o conceito de reversibilidade e convergência a processos com tempo contínuo.

De fato, a parte mais árdua do trabalho já está feita no capítulo 2, tentaremos apenas adaptar estes resultados conhecidos para nosso caso em consideração (tempo contínuo).

Este capítulo apresentará as ferramentas necessárias para provar o teorema 5.3.1, principal foco deste trabalho.

3.2 Reversibilidade

O conceito de reversibilidade para processos de tempo contínuo está dado na definição 2.2.1 e, como vimos no capítulo 2, a proposição 2.2.3 nos dá uma condição necessária e suficiente para que uma matriz de transição de probabilidades (semi-grupo a um parâmetro contínuo) \mathcal{P}^t , $t \geq 0$, juntamente com uma distribuição inicial μ determinem um processo de Markov reversível, a saber, a *condição do balanço detalhado* 2.4.

Continuando, para o que segue abaixo, é necessária a definição de \mathcal{Q} -matriz, definição 1.3.25.

Apresentaremos a seguir, um procedimento para obter uma matriz estocástica Π , a partir de uma \mathcal{Q} -matriz. Tal matriz estocástica é denominada a *Matriz Jump* associada à \mathcal{Q} -matriz e sua importancia ficará clara a medida que avançarmos.

Definição 3.2.1

Seja \mathcal{Q} uma \mathcal{Q} -matriz sobre S finito, então, a *Matriz Jump* $\Pi = (\pi)_{ij}$ com, $(i, j) \in S^2$, associada à \mathcal{Q} -matriz $\mathcal{Q} = (Q)_{ij}$, com $(i, j) \in S^2$, é dada por:

$$\begin{aligned} (\pi)_{ij} &:= \begin{cases} \frac{(Q)_{ij}}{-(Q)_{ii}} & \text{se } j \neq i \text{ e } -(Q)_{ii} \neq 0 \\ 0 & \text{se } j \neq i \text{ e } (Q)_{ii} = 0, \end{cases} \\ (\pi)_{ii} &:= \begin{cases} 0 & \text{se } -(Q)_{ii} \neq 0 \\ 1 & \text{se } (Q)_{ii} = 0. \end{cases} \end{aligned}$$

◇

Proposição 3.2.2

A *Matriz Jump* $\Pi = (\pi)_{ij}$, definida acima, é uma matriz (linha) estocástica.

◇

Prova

Provaremos apenas o caso $j \neq i$ e $-(Q)_{ij} \neq 0$. Com efeito

$$\sum_{j \neq i} (\pi)_{ij} = \sum_{j \neq i} \frac{(Q)_{ij}}{-(Q)_{ii}} = \frac{1}{-(Q)_{ii}} \sum_{j \neq i} (Q)_{ij} = \frac{1}{-(Q)_{ii}} (-(Q)_{ii}) = 1,$$

onde 1 segue do fato que, sendo $\sum_{j \in S} (Q)_{ij} = 0 \Rightarrow$ que $\sum_{j \neq i} (Q)_{ij} + (Q)_{ii} = 0 \Rightarrow -(Q)_{ii} = \sum_{j \neq i} (Q)_{ij}$. Logo, Π é (linha) estocástica. □

Proposição 3.2.3

Toda \mathcal{Q} -matriz admite uma decomposição da forma $\mathcal{Q} = \mathcal{R}(\Pi - I)$, onde $\mathcal{R} = \text{diag}\{-(Q)_{ii}\} = \text{diag}\{(R)_{ii}\}$ para todo $i \in S$, Π é a Matriz Jump e I é a matriz identidade sobre S . ◇

Prova

Note primeiramente que a matriz $\Pi - I$ é dada por $(\pi)_{ij}$ se $i \neq j$ e por $(\pi)_{ii} - 1$ se $i = j$. Mas, pela definição de Π (definição 3.2.1) temos que

$$\begin{aligned} (\pi)_{ij} &= \begin{cases} \frac{(Q)_{ij}}{-(Q)_{ii}} & \text{se } j \neq i \text{ e } -(Q)_{ii} \neq 0 \\ 0 & \text{se } j \neq i \text{ e } (Q)_{ii} = 0, \end{cases} \\ (\pi)_{ii} - 1 &= \begin{cases} -1 & \text{se } -(Q)_{ii} \neq 0 \\ 0 & \text{se } (Q)_{ii} = 0. \end{cases} \end{aligned}$$

Logo, utilizando a definição de produto matricial vem, $\forall j \neq i$, que

$$\left[\mathcal{R} \left(\Pi - I \right) \right]_{ij} = \sum_{k \in S} (R)_{ik} (\pi)_{kj} = (R)_{ii} (\pi)_{ij} = -(Q)_{ii} \frac{(Q)_{ij}}{-(Q)_{ii}} = (Q)_{ij},$$

onde 1 segue do fato que $j \neq i$ e daí utilizamos a primeira expressão acima, além disso, devemos salientar que \mathcal{R} é diagonal.

Para $j = i$ temos,

$$\left[\mathcal{R} \left(\Pi - I \right) \right]_{ii} = \sum_{k \in S} (R)_{ik} (\pi)_{ki} = (R)_{ii} ((\pi)_{ii} - 1) = -(Q)_{ii} (-1) = (Q)_{ii},$$

aqui 1 se deve ao fato que $j = i$ e, portanto, utilizamos a segunda das expressões acima. Logo, $\mathcal{Q} = \mathcal{R}(\Pi - I)$, concluindo desta forma a prova da proposição. □

Definição 3.2.4

A matriz $\mathcal{R} = \text{diag}\{-(Q)_{ii}\} = \text{diag}\{(R)_{ii}\}$ mencionada na proposição 3.2.3 acima é chamada a *Matriz das Taxas* (\mathcal{R} de Rates) $(R)_{ii} = -(Q)_{ii} \forall i \in S$. Dizemos que as taxas $(R)_{ii}$ são *limitadas* se $\sup_{i \in S} (R)_{ii} < \infty$. ◇

Exemplo 3.1

A \mathcal{Q} -matriz

$$\mathcal{Q} = \begin{bmatrix} -2 & 1 & 1 \\ 1 & -1 & 0 \\ 2 & 1 & -3 \end{bmatrix},$$

tem como matriz jump associada a matriz

$$\mathbf{\Pi} = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1 & 0 & 0 \\ 2/3 & 1/3 & 0 \end{bmatrix},$$

e a matriz das taxas será

$$\mathcal{R} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

É relativamente fácil verificar que $\mathcal{Q} = \mathcal{R}(\mathbf{\Pi} - I)$.

□

Observação 3.2.5

A idéia de *Processos de Markov Explosivos* não será dada aqui, pois é importante apenas no caso em que S é infinito. Para um tratamento análogo ao que faremos, porém sobre S infinito, referenciamos o leitor para [24].

◇

Prosseguindo, vamos assumir que a \mathcal{Q} -matriz \mathcal{Q} em consideração seja irreduzível, no sentido que (ver definição 1.3.18) $(Q^n)_{ij} > 0$ para algum $n \in \mathbb{N}$.

Denotemos, como de costume, $\mathcal{P}^t = e^{t\mathcal{Q}}$ a matriz de transição de probabilidades a parâmetro contínuo gerada por \mathcal{Q} .

Nosso principal objetivo agora é mostrar que, se uma distribuição inicial de probabilidades μ satisfaz a condição do balanço detalhado para a \mathcal{Q} -matriz \mathcal{Q} , então, μ satisfaz a condição do balanço detalhado para o semi-grupo a tempo contínuo \mathcal{P}^t , e daí, o processo a tempo contínuo $\{X_t\}_{t \geq 0}$ é reversível com respeito a μ . Como estamos considerando S finito, as taxas $(R)_{ii}$ são limitadas, posto isto, podemos enunciar o seguinte teorema.

Teorema 3.2.6

Seja \mathcal{Q} uma \mathcal{Q} -matriz irreduzível sobre S finito. Assuma que uma distribuição inicial de probabilidades μ satisfaça a condição do balanço detalhado para \mathcal{Q} , i.e., que vale, $\forall (i, j) \in S^2$,

$$(\mu)_i(Q)_{ij} = (\mu)_j(Q)_{ji}. \quad (3.1)$$

Então vale, $\forall (i, j) \in S^2$ e $t \geq 0$, que

$$(\mu)_i(P^t)_{ij} = (\mu)_j(P^t)_{ji}. \quad (3.2)$$

Ou seja, dado que \mathcal{Q} satisfaz a condição do balanço detalhado 3.1 para μ , \mathcal{P}^t também a satisfaz (ver equação 3.2) e, daí, \mathcal{P}^t é reversível com respeito a μ .

◇

Prova

Mostraremos, primeiramente, que $(\mu)_i(Q^n)_{ij} = (\mu)_j(Q^n)_{ji} \forall (i, j) \in S^2$ e $\forall n \in \mathbb{N}$. Faremos esta parte da prova por indução em n . No que consiste, para $n = 1$ temos

$$(\mu)_i(Q)_{ij} = (\mu)_j(Q)_{ji},$$

que é, justamente, nossa hipótese 3.1, ou seja, a propriedade desejada vale para $n = 1$.

Supondo válida para um certo $n = k$ fixo, i.e.,

$$(\mu)_i(Q^k)_{ij} = (\mu)_j(Q^k)_{ji} \tag{3.3}$$

vamos mostrar que isto implica sua validade para $n = k + 1$. Com efeito, temos

$$\begin{aligned} (\mu)_i(Q^{k+1})_{ij} &=^1 (\mu)_i \sum_{l \in S} (Q^k)_{il} (Q)_{lj} = \sum_{l \in S} (\mu)_i (Q^k)_{il} (Q)_{lj} =^2 \\ &= \sum_{l \in S} (\mu)_l (Q^k)_{li} (Q)_{lj} = \sum_{l \in S} (Q^k)_{li} (\mu)_l (Q)_{lj} =^3 \\ &= \sum_{l \in S} (Q^k)_{li} (\mu)_j (Q)_{jl} = (\mu)_j \sum_{l \in S} (Q)_{jl} (Q^k)_{li} =^4 \\ &= (\mu)_j (Q^{k+1})_{ji}, \end{aligned}$$

onde 2 se deve a hipótese de indução 3.3, 3 se deve a hipótese 3.1, 1 e 4 à definição de produto matricial. Portanto, $(\mu)_i(Q^{k+1})_{ij} = (\mu)_j(Q^{k+1})_{ji} \forall (i, j) \in S^2$. Logo, pelo princípio da indução, temos que $(\mu)_i(Q^n)_{ij} = (\mu)_j(Q^n)_{ji} \forall (i, j) \in S^2$ e $\forall n \in \mathbb{N}$.

Da maneira que definimos \mathcal{P}^t (ver teorema 1.3.27), i.e.,

$$\mathcal{P}^t := e^{tQ} = \sum_{n=0}^{\infty} \frac{t^n Q^n}{n!},$$

temos, $\forall (i, j) \in S^2$ e $\forall t \geq 0$, que

$$\begin{aligned} (\mu)_i(\mathcal{P}^t)_{ij} &= (\mu)_i \sum_{n=0}^{\infty} \frac{t^n (Q^n)_{ij}}{n!} = \sum_{n=0}^{\infty} \frac{t^n (\mu)_i (Q^n)_{ij}}{n!} = \sum_{n=0}^{\infty} \frac{t^n (\mu)_j (Q^n)_{ji}}{n!} = \\ &= (\mu)_j \sum_{n=0}^{\infty} \frac{t^n (Q^n)_{ji}}{n!} = (\mu)_j (\mathcal{P}^t)_{ji}. \end{aligned}$$

Ou seja, vale 3.2 e, então, podemos concluir que \mathcal{P}^t é reversível com respeito a μ pela proposição 2.2.3.

□

3.3 Convergência em $L^2(\mu)$ - Tempo Contínuo

Pelo teorema 3.2.6, vimos que um processo de Markov a tempo contínuo é reversível com respeito a μ se, e somente se, vale $\forall (i, j) \in S^2$,

$$(\mu)_i(Q)_{ij} = (\mu)_j(Q)_{ji}, \tag{3.4}$$

onde \mathcal{Q} é a \mathcal{Q} -matriz (conforme definição 1.3.25).

Portanto, valendo 3.4 tem-se, $\forall (i, j) \in S^2$ e $\forall t > 0$, que

$$(\mu)_i(P^t)_{ij} = (\mu)_j(P^t)_{ji}, \quad (3.5)$$

onde \mathcal{P}^t é o semi-grupo a tempo contínuo gerado pela \mathcal{Q} -matriz.

Com a validade de 3.5, estamos tentados a utilizar as idéias desenvolvidas na seção 2.4 para estimar a taxa com que $\mathcal{P}^t f$ (análogo a 2.10) converge a $\langle f \rangle_\mu$ em $L^2(\mu)$. Ou seja, a taxa com que

$$\lim_{t \rightarrow \infty} \|\mathcal{P}^t f - \langle f \rangle_\mu\|_{2,\mu} = 0. \quad (3.6)$$

Para sermos mais precisos, primeiramente, enunciaremos o seguinte lema.

Lema 3.3.1

Para cada $h > 0$, temos que

$$\langle f, \mathcal{P}^h f \rangle_{2,\mu} = \|\mathcal{P}^{h/2} f\|_{2,\mu}^2 \geq 0. \quad (3.7)$$

◇

Prova

Com efeito, pela equação 2.10 temos, para cada $i \in S$, que

$$\mathcal{P}^{h/2} f(i) = \sum_{k \in S} f(k)(P^{h/2})_{ik}.$$

Logo, pela definição de norma 2.3.1, temos

$$\begin{aligned} \|\mathcal{P}^{h/2} f\|_{2,\mu}^2 &= \sum_{l \in S} (\mathcal{P}^{h/2} f(l))^2 (\mu)_l = \sum_{l \in S} (\mathcal{P}^{h/2} f(l)) (\mathcal{P}^{h/2} f(l)) (\mu)_l = \\ &= \sum_{l \in S} \left[\sum_{k \in S} f(k)(P^{h/2})_{lk} \right] \left[\sum_{j \in S} f(j)(P^{h/2})_{lj} \right] (\mu)_l = \\ &= \sum_{l \in S} \left[\sum_{k \in S} f(k)(P^{h/2})_{lk} \right] \left[\sum_{j \in S} f(j)(\mu)_l (P^{h/2})_{lj} \right] =^1 \\ &= \sum_{l \in S} \left[\sum_{k \in S} f(k)(P^{h/2})_{lk} \right] \left[\sum_{j \in S} f(j)(\mu)_j (P^{h/2})_{jl} \right] = \\ &= \left[\sum_{k \in S} f(k) \right] \left[\sum_{j \in S} f(j)(\mu)_j \sum_{l \in S} (P^{h/2})_{jl} (P^{h/2})_{lk} \right] =^2 \\ &= \left[\sum_{k \in S} f(k) \right] \left[\sum_{j \in S} f(j)(\mu)_j (P^h)_{jk} \right] =^3 \\ &= \left[\sum_{k \in S} f(k) \right] \left[\sum_{j \in S} f(j)(\mu)_k (P^h)_{kj} \right] = \\ &= \sum_{k \in S} f(k) \left(\sum_{j \in S} f(j)(P^h)_{kj} \right) (\mu)_k =^4 \sum_{k \in S} f(k) \mathcal{P}^h f(k) (\mu)_k =^5 \\ &= \langle f, \mathcal{P}^h f \rangle_{2,\mu}, \end{aligned}$$

onde 1 e 3 se devem a condição do balanço detalhado 3.5, já 2 vem da definição de produto matricial e da propriedade de semi-grupo, 4 vem de 2.10 e, finalmente, 5 vem da definição de produto interno 2.3.1, provando assim o lema. \square

Com este lema, temos que $\beta = \beta_+$ e daí temos a seguinte identificação para β , (comparando com aquelas expressões dadas por 2.21, 2.28 e 2.34).

$$\begin{aligned} \beta(h) &= 1 - \sup\{|\langle f, \mathcal{P}^h f \rangle_{2,\mu}| : f \in L_0^2(\mu) \text{ com } \|f\|_{2,\mu} = 1\} = \\ &= \inf\{\langle f, (I - \mathcal{P}^h)f \rangle_{2,\mu} : f \in L^2(\mu) \text{ com } \text{Var}_\mu(f) = 1\} = \\ &= \inf\{\xi_h(f, f) : f \in L^2(\mu) \text{ com } \text{Var}_\mu(f) = 1\}, \end{aligned} \quad (3.8)$$

onde

$$\xi_h(f, f) = \frac{1}{2} \sum_{i \neq j} (\mu)_i (P^h)_{ij} (f(j) - f(i))^2. \quad (3.9)$$

A equação 3.9 é a *Forma de Dirichlet* de \mathcal{P}^t em $L^2(\mu)$. Conseqüentemente, em analogia com a relação 2.20, tem-se, para qualquer $t > 0$ e $n > 0$, que

$$\|\mathcal{P}^t f - \langle f \rangle_\mu\|_{2,\mu} \leq (1 - \beta(t/n))^n \|f - \langle f \rangle_\mu\|_{2,\mu}. \quad (3.10)$$

Para prosseguirmos, devemos relembrar que, como estamos considerando apenas o caso em que S é finito, as taxas \mathcal{R} são limitadas (ver definição 3.2.4).

Posto isto, podemos enunciar a seguinte proposição.

Proposição 3.3.2

Seja $\xi_h(f, f)$ dada como em 3.9 e $f \in L^2(\mu)$. Então,

$$\lim_{h \rightarrow 0} \frac{\xi_h(f, f)}{h} = \xi^{\mathcal{Q}}(f, f), \quad (3.11)$$

onde

$$\xi^{\mathcal{Q}}(f, f) = \frac{1}{2} \sum_{i \neq j} (\mu)_i (Q)_{ij} (f(j) - f(i))^2. \quad (3.12)$$

\diamond

Prova

Com efeito,

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{\xi_h(f, f)}{h} &= \lim_{h \rightarrow 0} \frac{1}{2h} \sum_{i \neq j} (\mu)_i (P^h)_{ij} (f(j) - f(i))^2 = \\ &= \frac{1}{2} \sum_{i \neq j} (\mu)_i \lim_{h \rightarrow 0} \left(\frac{(P^h)_{ij} - 0}{h} \right) (f(j) - f(i))^2 =^1 \\ &= \frac{1}{2} \sum_{i \neq j} (\mu)_i \lim_{h \rightarrow 0} \left(\frac{(P^h)_{ij} - (I)_{ij}}{h} \right) (f(j) - f(i))^2 =^2 \\ &= \frac{1}{2} \sum_{i \neq j} (\mu)_i (Q)_{ij} (f(j) - f(i))^2, \end{aligned}$$

onde 1 vem do fato que a soma está sendo feita apenas para $i \neq j$ e, neste caso, os elementos fora da diagonal da matriz identidade I são nulos, i.e., $(I)_{ij} \equiv 0$, além disso, pelo teorema 1.3.27 temos que $\mathcal{P}^0 = I$. Já 2 segue também pelo teorema 1.3.27, pois o limite que aparece dentro do somatório na igualdade 1 é, de fato, igual a $\frac{d}{dt}(P^t)_{ij}|_{t=0} = (Q)_{ij}$.

□

Da proposição 3.3.2 concluímos que, $\lim_{h \rightarrow 0} \beta(h)/h$ existe e, além do mais, para $\beta(h)$ dado por 3.8, tem-se:

$$\lambda := \lim_{h \rightarrow 0} \frac{\beta(h)}{h} = \inf\{\xi^{\mathcal{Q}}(f, f) : f \in L^2(\mu) \text{ com } \text{Var}_{\mu}(f) = 1\}, \quad (3.13)$$

em analogia com a relação 2.34.

Finalmente, no caso em consideração (S finito e, portanto, taxas \mathcal{R} limitadas), concluímos (ver quação 3.10) que

$$\|\mathcal{P}^t f - \langle f \rangle_{\mu}\|_{2,\mu} \leq e^{-\lambda t} \|f - \langle f \rangle_{\mu}\|_{2,\mu}. \quad (3.14)$$

Para ver isto, lembremos primeiramente que, da *análise na reta*, vale o limite

$$\lim_{s \rightarrow \infty} \left(1 - \frac{a}{s}\right)^s = e^{-a}. \quad (3.15)$$

Logo,

$$\begin{aligned} \lim_{n \rightarrow \infty} (1 - \beta(t/n))^n & \stackrel{1}{=} \lim_{n \rightarrow \infty} \left(1 - \frac{\beta(t/n)}{(t/n)} (t/n)\right)^n = \lim_{n \rightarrow \infty} \left(1 - \frac{t\beta(t/n)}{(t/n)} \frac{1}{n}\right)^n \stackrel{2}{=} \\ & = e^{-\lambda t}, \end{aligned}$$

onde em 1 apenas multiplicamos e dividimos por t/n . Já 2 segue do limite 3.15, além disso, quando $n \rightarrow \infty$ temos $\frac{t}{n} \rightarrow 0$ e portanto, por 3.13 $\beta(\frac{t}{n})/(\frac{t}{n}) \rightarrow \lambda$. Em analogia com a expressão 2.20. Vemos assim que vale o limite 3.6 e, além disso, a convergência é exponencial com taxa λ .

3.3.1 Estimando λ

De forma análoga ao que fizemos na subseção 2.4.2, podemos estimar λ em 3.14, assim como estimamos β em 2.20.

No que consiste, devemos escolher uma seleção de caminhos P consistindo de caminhos $p(i, j)$, para cada $(i, j) \in S^2 \setminus D$, com a propriedade que, se $p(i, j) = (k_0, \dots, k_n)$, com $k_0 = i$ sendo o vértice inicial e $k_n = j$ o vértice final então, $p(i, j)$ é *acessível* no sentido que $(Q)_{k_{m-1}k_m} > 0$ para cada $1 \leq m \leq n$, onde Q é a Q -matriz geradora do semi-grupo a tempo contínuo \mathcal{P}^t .

Logo, da mesmo forma como procedemos em na subseção 2.4.2, temos

$$\lambda \geq \frac{1}{W(P)}, \quad (3.16)$$

onde,

$$W(\mathbf{P}) = \sup_{e \in \mathbf{A}} \sum_{\{p \in \mathbf{P} : e \in p\}} \sum_{e' \in p} \frac{(\mu(p))_- - (\mu(p))_+}{\rho(e')}, \quad (3.17)$$

e onde o supremo é tomado sobre as arestas orientadas $e = (k, l)$ com $(Q)_{kl} > 0$, a primeira soma é sobre $p \in \mathbf{P}$ em que a aresta e aparece, a segunda soma é feita sobre as arestas e' que aparecem no caminho p , também $(\mu(p))_- = (\mu)_i$ e $(\mu(p))_+ = (\mu)_j$, se $p = p(i, j)$ e $\rho(e') = (\mu)_k(Q)_{kl}$ se $e' = (k, l)$ (veja também o exemplo 2.5).

Com isto, completamos a analogia com o caso de processos a tempo discreto.

Temos, portanto, ferramentas suficientes para desenvolver um dos principais capítulos deste trabalho (capítulo 5) e provar o importante teorema lá enunciado (teorema 5.3.1).

4 Elementos de Mecânica Estatística

4.1 Introdução

Neste capítulo iremos apresentar, de forma sucinta, os rudimentos da mecânica estatística de equilíbrio, mais especificamente, nos focaremos na descrição do chamado *Ensemble Canônico*. A abordagem dada aqui será, praticamente, do ponto de vista físico e a omissão da leitura deste capítulo não compromete, em nada, o entendimento dos seguintes.

De fato, a exposição abaixo tem o objetivo de motivar e justificar algumas idéias e conceitos apresentadas nos capítulos 5 e 6, além de dar um ar de completude para este trabalho. Para um tratado completo sobre os tópicos apresentados abaixo indicamos as seguintes referências: [13], [15], [20] e [22].

4.2 Postulados da Mecânica Estatística de Equilíbrio

Antes de introduzir os postulados da termodinâmica de equilíbrio, vamos definir *Sistema Simples* e *Sistema Composto*.

Definição 4.2.1

Sistemas Simples são macroscopicamente homogêneos, isotrópicos, descarregados, quimicamente inertes e suficientemente grandes. Muitas vezes, um Sistema Simples é chamado de *Fluido Puro*.

Um *Sistema Composto* é constituído por um conjunto de Sistemas Simples separados por *paredes* ou *vínculos*.

Vínculos são divisórias ideais que podem ser restritivas a certas variáveis. Os principais tipos de vínculos são:

1. Adiabáticos: São restritivos à troca de energia na forma de calor.
2. Fixos: São restritivos à variação do volume.
3. Impermeáveis: Impedem a passagem de partículas de um ou mais componentes do fluido.

◇

Primeiro Postulado: O *Estado Microscópico* de um fluido puro é completamente caracterizado pela energia interna U , pelo volume V e pelo número de partículas N .

◇

O número de *Estados Microscópicos* de um fluido termodinâmico, com energia E , volume V e número de partículas N , na presença de um conjunto $\{\chi_i\}$ de vínculos internos, é dado por uma função $\Omega = \Omega(E, V, N; \{\chi_i\})$. Por exemplo, se tivermos N_i partículas associadas ao vínculo χ_i , muito freqüentemente na literatura se encontra

$$\Omega(E, V, N; \{\chi_i\}) = \frac{N!}{\prod_i (N_i!)}, \quad (4.1)$$

onde $\sum_i N_i = N$.

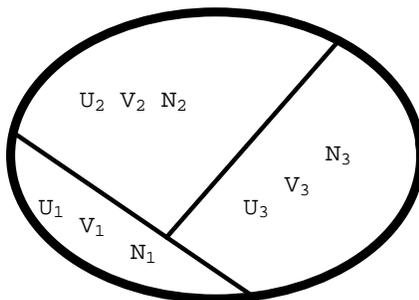


Figura 4.1 - Sistema composto constituído por três fluidos simples separados por vínculos adiabáticos, fixos e impermeáveis.

O problema fundamental da termodinâmica, que será respondido pelos dois postulados seguintes, consiste na determinação do estado final de equilíbrio atingido após a remoção de vínculos internos de um sistema composto. Por exemplo, qual seria o estado final de equilíbrio quando um vínculo do sistema composto, representado pela figura 4.1, se transforma em *Diatérmico* (antônimo de adiabático)?

Segundo Postulado: Existe uma função de todos *Parâmetros Extensivos* de um sistema composto, denominada *Entropia*, $S = S(U_1, V_1, N_1, U_2, V_2, N_2, \dots)$, que está definida para todos os estados de equilíbrio. Na remoção de um vínculo interno, os parâmetros extensivos assumem valores que maximizam a entropia.

◇

Terceiro Postulado: A entropia de um sistema composto é aditiva sobre cada um dos seus componentes. A entropia é um função contínua, diferenciável e monótona crescente.

◇

Temos, portanto,

$$S(U_1, V_1, N_1, U_2, V_2, N_2, \dots) = \sum_j S(U_j, V_j, N_j). \quad (4.2)$$

Além disso, a aditividade da entropia significa que $S = S(U, V, N)$ é uma função homogênea de primeiro grau das suas variáveis, i.e., que

$$S(\alpha U, \alpha V, \alpha N) = \alpha S(U, V, N). \quad (4.3)$$

Finalmente, enunciaremos o principal postulado da mecânica estatística de equilíbrio, o *postulado fundamental da mecânica estatística* ou *postulado das probabilidades iguais a priori*.

Postulado fundamental da mecânica estatística: Em um sistema estatístico fechado, com energia fixa, todos os estados microscópicos acessíveis são igualmente prováveis.

◇

Vamos dar agora a definição da *Entropia Termodinâmica* de Boltzmann.

Definição 4.2.2

A *Entropia Termodinâmica* de um sistema isolado com energia E é dada por

$$S(E) := k_B \log(\Omega(E)), \quad (4.4)$$

onde k_B é a constante de Boltzmann, e $\Omega(E)$ é o número de estados microscópicos acessíveis ao sistema com essa energia (ver, por exemplo, equação 4.1).

◇

4.2.1 Parâmetros Intensivos da Termodinâmica

Na *Representação da Entropia*, a *Equação Fundamental* de um fluido puro é dada pela função $S = S(U, V, N)$. Temos, então, a forma diferencial

$$dS = \left(\frac{\partial S}{\partial U} \right)_{V,N} dU + \left(\frac{\partial S}{\partial V} \right)_{U,N} dV + \left(\frac{\partial S}{\partial N} \right)_{U,V} dN. \quad (4.5)$$

Por outro lado, pela *Lei da Conservação da Energia*, temos

$$dS = \frac{1}{T} dU + \frac{p}{T} dV - \frac{\gamma}{T} dN. \quad (4.6)$$

Donde, comparando 4.5 com 4.6, obtemos as seguintes definições dos *Parâmetros Intensivos* da termodinâmica (*Equações de Estado*):

$$\frac{1}{T} := \left(\frac{\partial S}{\partial U} \right)_{V,N}, \quad \frac{p}{T} := \left(\frac{\partial S}{\partial V} \right)_{U,N} \quad \text{e} \quad -\frac{\gamma}{T} := \left(\frac{\partial S}{\partial N} \right)_{U,V}, \quad (4.7)$$

onde as equações de estado $T = T(U, V, N)$, $p = p(U, V, N)$ e $\gamma = \gamma(U, V, N)$ são, respectivamente a *temperatura*, a *pressão* e o *potencial químico*.

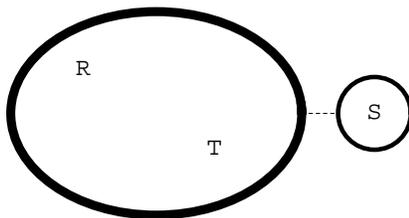


Figura 4.2 - Sistema \mathcal{S} em contato com o reservatório térmico R a temperatura T .

4.3 Ensemble Canônico

Vamos considerar um sistema simples \mathcal{S} em contato com um reservatório térmico com temperatura T por meio de uma parede diatérmica, mas fixa e impermeável. Na figura 4.2, o reservatório R é muito grande em relação ao sistema de interesse \mathcal{S} .

Quando o sistema composto $\mathcal{S} + R$ estiver isolado, com energia total E_0 , valem os postulados da mecânica estatística de equilíbrio (ver seção 4.2). Portanto, a distribuição de probabilidades $\mu = ((\mu)_1, (\mu)_2, \dots)$, com $(\mu)_i$ sendo a probabilidade de encontrar o sistema \mathcal{S} num *particular estado microscópico* i , será dada por

$$(\mu)_i = c \Omega_R(E_0 - E_i), \quad (4.8)$$

onde c é uma constante de normalização, E_i é a energia do sistema no particular estado microscópico i e $\Omega_R(E)$ é o número de estados microscópicos acessíveis ao reservatório R com energia E .

Antes de enunciarmos a próxima proposição, vamos dar a importante definição de *Função Canônica de Partição* (ou soma sobre os estados, do alemão *Zustandsumme*).

Definição 4.3.1

Define-se *Função Canônica de Partição* como a soma

$$Z(\beta) := \sum_j e^{-\beta E_j}, \quad (4.9)$$

onde $\beta = 1/(k_B T)$ e E_j é a função energia do sistema no estado microscópico j .

◇

A função canônica de partição Z está associada a normalização de probabilidade $(\mu)_i$, como veremos na proposição abaixo. (Note que β **Não** é o *Gap Spectral*).

Proposição 4.3.2

Para o caso em consideração, a distribuição de probabilidades μ , onde $(\mu)_i$ está dado como em 4.8, é dada por

$$(\mu(\beta))_i = \frac{e^{-\beta E_i}}{Z(\beta)}, \quad (4.10)$$

onde $Z(\beta)$ é a função canônica de partição dada por 4.9.

◇

Prova

Como o reservatório R é sempre muito grande, a energia E_i deve ser muito menor do que E_0 (ver relação 4.8). Então, tomando o logaritmo em 4.8, podemos fazer a seguinte expansão

$$\begin{aligned} \log((\mu)_i) &= \log(c) + \log(\Omega_R(E_0)) + \left[\frac{\partial \log(\Omega_R(E))}{\partial E} \right]_{E=E_0} (-E_i) + \\ &+ \frac{1}{2} \left[\frac{\partial^2 \log(\Omega_R(E))}{\partial E^2} \right]_{E=E_0} (-E_i)^2 + \dots \end{aligned}$$

Usando a definição de entropia termodinâmica dada pela equação 4.4 e a primeira das três equações de estado em 4.7 temos

$$\frac{\partial \log(\Omega_R(E))}{\partial E} = \frac{1}{k_B T},$$

onde T é a temperatura do reservatório. Também devemos ter

$$\frac{\partial^2 \log(\Omega_R(E))}{\partial E^2} = \frac{1}{k_B} \frac{\partial}{\partial E} \left(\frac{1}{T} \right) \rightarrow 0,$$

no limite de um verdadeiro reservatório térmico, em que a temperatura está praticamente fixa. Portanto, a expansão inicial se reduz à forma

$$\log((\mu)_i) = \tilde{c} - \frac{1}{k_B T} E_i,$$

ou seja,

$$(\mu)_i = \frac{e^{-\beta E_i}}{\sum_j e^{-\beta E_j}},$$

onde estamos utilizando a notação $\beta = 1/(k_B T)$.

Logo, valendo-nos da definição de função canônica de partição (definição 4.3.1) concluímos, finalmente, que

$$(\mu(\beta))_i = \frac{e^{-\beta E_i}}{Z(\beta)}.$$

□

Observação 4.3.3

A distribuição de probabilidades $\mu(\beta)$, onde $(\mu(\beta))_i$ é dada por 4.10, é denominada um *Estado de Gibbs*.

Em certas situações, é necessário dar mais, ou menos importância para um determinado estado microscópico i , basta considerar então um vetor $\nu = ((\nu)_1, (\nu)_2, \dots)$, chamado *Vetor de Pesos* (físicamente corresponde à *degenerescência* do estado de energia mas, na maioria dos casos, consideramos ν uniforme, i.e., $(\nu)_i \equiv 1 \forall i$). Daí, o Estado de Gibbs tornar-se-á:

$$(\mu(\beta))_i = \frac{e^{-\beta E_i} (\nu)_i}{Z(\beta)}. \quad (4.11)$$

◇

Estamos aptos a definir o que, de fato, é o *Ensemble Canônico*.

Definição 4.3.4

O *Ensemble Canônico* é constituído pelo conjunto de estados microscópicos $\{i\}$, associados à distribuição de probabilidades dada pela equação 4.11, acessíveis a um sistema \mathcal{S} , em contato com um reservatório térmico R a temperatura T .

◇

4.3.1 Valor Esperado e Variância

Certamente existe flutuações da energia no ensemble canônico. Utilizando o Estado de Gibbs $(\mu)_i$, dado por 4.10, podemos obter o valor médio, ou valor esperado probabilístico, da energia do sistema \mathcal{S} , é fácil ver que

$$-\frac{\partial}{\partial \beta} \log(Z(\beta)) = \frac{\sum_j E_j e^{-\beta E_j}}{Z(\beta)} = \sum_j E_j \frac{e^{-\beta E_j}}{Z(\beta)} \stackrel{1}{=} \sum_j E_j (\mu)_j \stackrel{2}{=} \langle E_j \rangle_{\mu(\beta)},$$

onde 1 vem da relação 4.10 e 2 vem da definição de valor esperado em relação a $\mu(\beta)$, como na definição 2.3.1.

Logo, o valor esperado da energia pode ser dado, em termos da função canônica de partição $Z(\beta)$, por

$$\langle E_j \rangle_{\mu(\beta)} = -\frac{\partial}{\partial \beta} \log(Z(\beta)). \quad (4.12)$$

Vamos agora obter uma expressão para a variância. Pela definição 2.3.1, a variância de E_j , com respeito ao Estado de Gibbs $\mu(\beta)$, será

$$\begin{aligned} \text{Var}_{\mu(\beta)}(E_j) &= \langle E_j^2 \rangle_{\mu(\beta)} - \langle E_j \rangle_{\mu(\beta)}^2 = \frac{\sum_j E_j^2 e^{-\beta E_j}}{Z(\beta)} - \left[\frac{\sum_j E_j e^{-\beta E_j}}{Z(\beta)} \right]^2 = \\ &= \frac{\partial}{\partial \beta} \left[\frac{1}{Z(\beta)} \frac{\partial Z(\beta)}{\partial \beta} \right] = -\frac{\partial}{\partial \beta} \langle E_j \rangle_{\mu(\beta)} = \frac{\partial^2}{\partial \beta^2} \log(Z(\beta)). \end{aligned}$$

Donde que

$$\text{Var}_{\mu(\beta)}(E_j) = \frac{\partial^2}{\partial \beta^2} \log(Z(\beta)). \quad (4.13)$$

4.4 A Equação Mestra

Vamos introduzir a *Equação Mestra* que governa a evolução temporal dos processos de Markov. Com ela, daremos uma justificativa para a *condição do balanço detalhado* 2.4 vista no capítulo 2.

Seja $(\mu(t))_i$ a probabilidade de encontrar um sistema no estado microscópico i , num determinado instante t . De forma ilustrativa, podemos escrever

$$\frac{\partial}{\partial t} (\mu(t))_i = R_{dentro} - R_{fora}, \quad (4.14)$$

onde a taxa de variação da probabilidade “para dentro” do estado i é dada por

$$R_{dentro} := \sum_j (\mu(t))_j (P^t)_{ji}, \quad (4.15)$$

com $(P^t)_{ji}$ interpretado como sendo a probabilidade, na unidade de tempo, de que o sistema mude do estado j para o estado i . Da mesma maneira, a taxa de variação da probabilidade “para fora” do estado i é escrita como

$$R_{fora} := (\mu(t))_i \sum_j (P^t)_{ij}. \quad (4.16)$$

Temos, portanto, a seguinte definição para a *Equação Mestra*.

Definição 4.4.1

A *Equação Mestra* é definida por

$$\frac{\partial}{\partial t} (\mu(t))_i := \sum_j [(\mu(t))_j (P^t)_{ji} - (\mu(t))_i (P^t)_{ij}]. \quad (4.17)$$

◇

Toda dificuldade reside na obtenção das *probabilidades de transição* $(P^t)_{ij}$. Em problemas de interesse físico (veja capítulo 6) devemos calcular essas probabilidades de transição adotando formas plausíveis, consistentes com os aspectos físicos subjacentes.

Nos *estados estacionários*, a probabilidade $(\mu(t))_i$ não deve ser uma função explícita do tempo, ou seja

$$\frac{\partial}{\partial t} (\mu(t))_i = 0 \quad (4.18)$$

na situação de equilíbrio. Observando a equação 4.17, uma *condição suficiente* para o equilíbrio é dada, $\forall i, j$, por

$$(\mu)_i (P^t)_{ij} = (\mu)_j (P^t)_{ji}, \quad (4.19)$$

que é, justamente, a *Condição do Balanço Detalhado* (ver equação 2.4).

Essa equação de balanço detalhado, como o próprio nome indica, tem um significado intuitivo muito claro: na situação estacionária, devemos ter o mesmo número de transições de i para j ou na direção contrária, de j para i .

Uma das estratégias mais freqüentes nessa área consiste em escolher $(P^t)_{ij}$ a fim de satisfazer a equação do balanço detalhado no equilíbrio, ou seja, tal que

$$\lim_{t \rightarrow \infty} (\mu)_i (P^t)_{ij} = \lim_{t \rightarrow \infty} (\mu)_j (P^t)_{ji}. \quad (4.20)$$

Com essa escolha que, em geral, não é única, temos a certeza de atingir um estado final de equilíbrio para tempos suficientemente grandes.

5 Dinâmica de Glauber

5.1 Introdução

Como vimos no capítulo 4, um *Estado de Gibbs* (ver equação 5.3) é um modelo de equilíbrio, portanto é de interesse nos perguntarmos qual a dinâmica (cadeia de Markov) para que tal estado seja, de fato, um estado de equilíbrio.

Devemos, portanto, procurar um Processo de Markov para o qual o Estado de Gibbs seja uma *distribuição estacionária* (definição 1.3.6). Além disso, por razões físicas, deve ser *reversível* (ver definição 2.2.1 e também seção 4.4). Em outras palavras, devemos olhar os processos de Markov que são reversíveis com respeito ao Estado de Gibbs (pois reversibilidade implica invariância, i.e. $\mu\mathcal{P} = \mu$ ver corolário 2.2.4).

Tais processos foram introduzidos neste contexto por R. J. Glauber (veja também um trabalho original em [11]).

Iniciaremos pois, por construir uma Dinâmica de Glauber para o Estado de Gibbs 5.3.

Além disso, utilizaremos as ferramentas desenvolvidas ao longo do capítulo 3 para dar uma cota superior e inferior para a taxa λ_β (teorema 5.3.1) de convergência do limite

$$\lim_{t \rightarrow \infty} \|\mathcal{P}^t f - \langle f \rangle_{\mu(\beta)}\|_{2, \mu(\beta)} = 0, \quad (5.1)$$

onde $f \in L^2(\mu(\beta))$, \mathcal{P}^t é uma Dinâmica de Glauber e $\mathcal{P}^t f$ é dada como na equação 2.10 do capítulo 2.

Definição 5.1.1

Um processo de Markov que é reversível com respeito a um Estado de Gibbs é chamado de *Dinâmica de Glauber* para este Estado de Gibbs.

◇

Desenvolveremos aqui o estudo sobre Estados de Gibbs e sua Dinâmica de Glauber associada.

5.2 Construção da Dinâmica de Glauber

Lembre que, salvo menção explícita, vamos considerar apenas o caso em que o espaço de estados S é finito.

Como vimos na observação 4.3.3 é, muitas vezes, conveniente considerar, para cada $i \in S$, um *Vetor de Pesos* ν , tal que sua i -ésima coordenada seja $(\nu)_i$. Também observamos que, em muitas aplicações, ν é uniforme, i.e., dá peso 1 para cada $i \in S$, mas, em outras situações, é conveniente não assumir, *a priori*, que ν seja uniforme.

Vamos considerar uma função $\mathcal{H} : S \rightarrow \mathbb{R}^+$, o *Hamiltoniano* do sistema (função energia, ver capítulo 4) com a propriedade que

$$Z(\beta) = \sum_{i \in S} e^{-\beta \mathcal{H}(i)} (\nu)_i < \infty, \quad (5.2)$$

para cada $\beta \in (0, \infty)$. (O que sempre ocorre no caso de S finito).

Como vimos no capítulo 4, fisicamente $\beta = 1/(k_B T)$, onde k_B é a *constante de Boltzmann*, T é a *temperatura absoluta* e a função $Z : (0, \infty) \rightarrow (0, \infty)$ tal que, $\beta \mapsto Z(\beta)$ dada em 5.2, é a *Função Canônica de Partição* (ver definição 4.3.1).

Definição 5.2.1

Para cada $\beta \in (0, \infty)$, o *Estado de Gibbs* $\mu(\beta)$ é, por definição, o vetor de probabilidade dado por

$$(\mu(\beta))_i = \frac{1}{Z(\beta)} e^{-\beta \mathcal{H}(i)} (\nu)_i, \quad (5.3)$$

para $i \in S$.

◇

A prova da proposição 4.3.2 nos dá um esboço da dedução da distribuição 5.3.

Num ponto de vista físico, tudo de interessante está na função de partição 5.2. Por exemplo, como vimos no capítulo 4 subseção 4.3.1, é relativamente fácil obter o valor médio (valor esperado) e a variância da função energia \mathcal{H} , basta tomar o logaritmo de Z e derivar em relação a β , ou seja,

$$\langle \mathcal{H} \rangle_{\mu(\beta)} = -\frac{\partial}{\partial \beta} \log(Z(\beta)) \quad \text{e} \quad \text{Var}_{\mu(\beta)}(\mathcal{H}) = \frac{\partial^2}{\partial \beta^2} \log(Z(\beta)), \quad (5.4)$$

onde o subscrito $\mu(\beta)$ indica que, tanto o valor esperado, quanto a variância, estão sendo calculadas em relação ao Estado de Gibbs $\mu(\beta)$ (veja também definição 2.3.1).

Para descrever a Dinâmica de Glauber, iniciamos com uma matriz \mathcal{A} que tenha entradas não-negativas e, cujas entradas da diagonal, sejam nulas. Além disso, vamos assumir que \mathcal{A} seja irredutível, no sentido que, $\forall (i, j) \in S^2$, se tenha

$$\sup_{n \geq 0} (A^n)_{ij} > 0, \quad (5.5)$$

(ver definição 1.3.18), e que seja reversível com respeito a ν , (proposição 2.2.3), no sentido que, $\forall (i, j) \in S^2$, valha

$$(\nu)_i (A)_{ij} = (\nu)_j (A)_{ji}, \quad (5.6)$$

que é, justamente, a *condição do balanço detalhado*.

Finalmente, devemos notar que

$$\sum_{j \in S} e^{-\beta \mathcal{H}(j)} (A)_{ij} < \infty, \quad (5.7)$$

$\forall i \in S$ e $\beta > 0$.

Continuando, existem várias maneiras de construir uma Dinâmica de Glaber. Adotaremos uma que melhor se adapta a nosso propósito, cuja \mathcal{Q} -matriz seja dada por

$$\begin{aligned} (Q(\beta))_{ij} &= e^{-\beta(\mathcal{H}(j)-\mathcal{H}(i))^+} (A)_{ij}, \quad \text{se } j \neq i \\ (Q(\beta))_{ii} &= -\sum_{j \neq i} (Q(\beta))_{ij}, \end{aligned} \quad (5.8)$$

onde $x^+ := x \vee 0$ é a parte não-negativa do número $x \in \mathbb{R}$.

A próxima proposição nos mostra que o processo Markoviano gerado pela \mathcal{Q} -matriz 5.8 tem as propriedades requeridas.

Proposição 5.2.2

Seja $\{X_t\}_{t \geq 0}$ o processo de Markov com tempo contínuo, cuja matriz de transição de probabilidades \mathcal{P}^t é gerada pela \mathcal{Q} -matriz 5.8. Então, $\{X_t\}_{t \geq 0}$ é um processo de Markov reversível com respeito ao Estado de Gibbs $\mu(\beta)$ dado por 5.3, i.e., $\{X_t\}_{t \geq 0}$ é uma Dinâmica de Glauber.

◇

Prova

Pelo teorema 3.2.6, é suficiente provar que a \mathcal{Q} -matriz 5.8 é reversível com respeito ao Estado de Gibbs $\mu(\beta)$, i.e., que vale, $\forall (i, j) \in S^2$, a condição do balanço detalhado

$$(\mu(\beta))_i (Q(\beta))_{ij} = (\mu(\beta))_j (Q(\beta))_{ji}.$$

Com efeito, por 5.3 e 5.8, vem para $i \neq j$ que

$$\begin{aligned} (\mu(\beta))_i (Q(\beta))_{ij} &= \frac{1}{Z(\beta)} e^{-\beta\mathcal{H}(i)} (\nu)_i e^{-\beta(\mathcal{H}(j)-\mathcal{H}(i))^+} (A)_{ij} = \\ &= \frac{1}{Z(\beta)} e^{-\beta(\mathcal{H}(i)+(\mathcal{H}(j)-\mathcal{H}(i))^+)} (\nu)_i (A)_{ij} = 1 \\ &= \frac{1}{Z(\beta)} e^{-\beta\mathcal{H}(i) \vee \mathcal{H}(j)} (\nu)_j (A)_{ji} = \\ &= \frac{1}{Z(\beta)} e^{-\beta\mathcal{H}(j) \vee \mathcal{H}(i)} (\nu)_j (A)_{ji} = \\ &= \frac{1}{Z(\beta)} e^{-\beta(\mathcal{H}(j)+(\mathcal{H}(i)-\mathcal{H}(j))^+)} (\nu)_j (A)_{ji} = \\ &= \frac{1}{Z(\beta)} e^{-\beta\mathcal{H}(j)} (\nu)_j e^{-\beta(\mathcal{H}(i)-\mathcal{H}(j))^+} (A)_{ji} = (\mu(\beta))_j (Q(\beta))_{ji}, \end{aligned}$$

onde 1 segue da condição 5.6.

Logo, pelo teorema 3.2.6 vem que \mathcal{P}^t é reversível com respeito ao Estado de Gibbs e, portanto, $\{X_t\}_{t \geq 0}$ é uma Dinâmica de Glauber.

□

É claro que esta escolha para a construção da Dinâmica de Glauber não é única, e cada escolha depende da particular situação em consideração. Além disso, assumindo, *a priori*, que S seja finito vem, pela observação 3.2.5, que, seja qual for nossa escolha, para cada $\beta > 0$, $\mathcal{Q}(\beta)$ determina um processo de Markov cujas taxas \mathcal{R} são limitadas. Para uma outra construção da Dinâmica de Glauber referenciamos o leitor para o exercício 5.6.14, página 143 de [24].

5.3 Convergência em $L^2(\mu(\beta))$ da Dinâmica de Glauber

Vamos contextualizar os resultados vistos na seção 3.3, com a Dinâmica de Glauber construída acima.

Definamos, como na equação 3.13 o seguinte:

$$\lambda_\beta := \inf\{\xi_\beta(f, f) : f \in L^2(\mu(\beta)) \text{ com } Var_\beta(f) = 1\}, \quad (5.9)$$

onde $\xi_\beta(f, f)$ é a *Forma de Dirichlet* (ver equação 3.12), dada por

$$\xi_\beta(f, f) = \frac{1}{2} \sum_{j \neq i} (\mu(\beta))_i (\mathcal{Q}(\beta))_{ij} (f(j) - f(i))^2. \quad (5.10)$$

Na equação 5.9 estamos utilizando a notação abreviada $Var_\beta(f)$ para significar $Var_{\mu(\beta)}(f)$, i.e., a variância de f com respeito ao Estado de Gibbs $\mu(\beta)$.

Para prosseguirmos, vamos introduzir as seguintes notações, para um caminho $p = (i_0, \dots, i_n) \in S^{n+1}$ (ver definição 2.4.13 do capítulo 2).

$$Elev(p) := \max_{0 \leq m \leq n} \mathcal{H}(i_m), \quad (5.11)$$

e

$$e(p) := Elev(p) - \mathcal{H}(i_0) - \mathcal{H}(i_n). \quad (5.12)$$

Então, como $\mu(\beta)$ é dado por 5.3 e $\mathcal{Q}(\beta)$ é dada por 5.8 vemos, para $p = (i_0, \dots, i_n)$, (ver equação 3.17) que

$$\begin{aligned} \omega_\beta &:= (\mu(\beta))_{i_0} (\mu(\beta))_{i_n} \sum_{e \in p} \frac{1}{\rho(e)} = \sum_{m=1}^n \frac{(\mu(\beta))_{i_0} (\mu(\beta))_{i_n}}{(\mu(\beta))_{i_{m-1}} (\mathcal{Q}(\beta))_{i_{m-1} i_m}} = \\ &= \frac{1}{Z(\beta)} \sum_{m=1}^n \frac{e^{-\beta \mathcal{H}(i_0)} (\nu)_{i_0} e^{-\beta \mathcal{H}(i_n)} (\nu)_{i_n}}{e^{-\beta \mathcal{H}(i_{m-1})} (\nu)_{i_{m-1}} e^{-\beta (\mathcal{H}(i_m) - \mathcal{H}(i_{m-1}))} (A)_{i_{m-1} i_m}} = \\ &= \frac{1}{Z(\beta)} \sum_{m=1}^n \frac{(\nu)_{i_0} (\nu)_{i_n}}{(\nu)_{i_{m-1}} (A)_{i_{m-1} i_m}} \frac{e^{-\beta (\mathcal{H}(i_0) + \mathcal{H}(i_n))}}{e^{-\beta (\mathcal{H}(i_m) \vee \mathcal{H}(i_{m-1}))}} = \\ &= \frac{1}{Z(\beta)} \sum_{m=1}^n \frac{(\nu)_{i_0} (\nu)_{i_n}}{(\nu)_{i_{m-1}} (A)_{i_{m-1} i_m}} e^{\beta (\mathcal{H}(i_m) \vee \mathcal{H}(i_{m-1}) - \mathcal{H}(i_0) - \mathcal{H}(i_n))}. \end{aligned}$$

Mas, para todo $i_m \in S$ com $m \in \{0, \dots, n\}$ vale, por 5.11 e 5.12 que

$$\mathcal{H}(i_m) \vee \mathcal{H}(i_{m-1}) - \mathcal{H}(i_0) - \mathcal{H}(i_n) \leq Elev(p) - \mathcal{H}(i_0) - \mathcal{H}(i_n) = e(p).$$

Portanto,

$$\omega_\beta \leq \frac{1}{Z(\beta)} \sum_{m=1}^n \frac{(\nu)_{i_0}(\nu)_{i_n}}{(\nu)_{i_{m-1}}(A)_{i_{m-1}i_m}} e^{\beta e(p)} = \frac{1}{Z(\beta)} \omega(p) e^{\beta e(p)}, \quad (5.13)$$

onde

$$\omega(p) := \sum_{m=1}^n \frac{(\nu)_{i_0}(\nu)_{i_n}}{(\nu)_{i_{m-1}}(A)_{i_{m-1}i_m}}. \quad (5.14)$$

Além disso, para qualquer escolha de caminhos P , sabemos que (ver equação 3.17)

$$W_\beta(P) := \sup_{e \in A} \sum_{\{p \in P: e \in p\}} \omega_\beta(p) \leq \frac{1}{Z(\beta)} e^{\beta E(P)} W(P), \quad (5.15)$$

onde

$$W(P) := \sup_{e \in A} \sum_{\{p \in P: e \in p\}} \omega(p) \quad \text{e} \quad E(P) := \sup_{p \in P} e(p). \quad (5.16)$$

Portanto, concluímos (ver equação 3.16) que

$$\lambda_\beta \geq \frac{1}{W_\beta(P)} \geq \frac{1}{\frac{1}{Z(\beta)} e^{\beta E(P)} W(P)} = \frac{Z(\beta) e^{-\beta E(P)}}{W(P)}. \quad (5.17)$$

Como estamos considerando S finito, é claro que 5.17 está bem definida já que, neste caso, $W(P) < \infty$. Ao mesmo tempo, ela mostra que se o interesse está nos β' grandes, então é importante escolher P de modo que $E(P)$ seja o menor possível.

A finitude de S permite-nos obter, para cada $(i, j) \in S^2$, um caminho $p(i, j)$ que minimize $Elev(p)$ dentre todos os *caminhos acessíveis* de i a j e, claramente, qualquer P consistindo de tais caminhos minimiza $E(P)$ (ver equação 5.11). Em qualquer caso, sempre que S for finito e P consistir de caminhos $p(i, j)$ que minimizem $Elev(p)$ sobre todos os caminhos p de i a j , $E(P)$ tem uma interpretação.

Pense S como um conjunto de *sítios* em um mapa e \mathcal{H} como a função que dá a altitude de cada *sítio* em relação ao nível do mar, i.e., $\mathcal{H}(i)$ é a altitude do *sítio* $i \in S$ em relação ao nível do mar.

Sem perda de generalidade, vamos supor que o *sítio* de menor altitude $k_0 \in S$ esteja ao nível do mar, ou seja, que $\mathcal{H}(k_0) = 0$.¹

Quando existe tal k_0 , nós interpretamos (de forma metafórica) $E(P)$ como a maior (diferença de) altitude que uma pessoa caminhando deve atingir, não importa onde ele tenha iniciado, ou que caminho acessível tenha escolhido, para chegar ao nível zero (do mar seguindo a metáfora).

Para ver isto, primeiramente observe que se p e p' são um par de caminhos acessíveis e, se o ponto final de p é o ponto inicial de p' , então o caminho q é acessível e $Elev(q) \leq Elev(p) \vee Elev(p')$, onde q é obtido conectando-se p e p' , i.e., se $p = (i_0, \dots, i_n)$ e $p' = (i'_0, \dots, i'_n)$ então $q = (i_0, \dots, i_n, i'_0, \dots, i'_n)$.

¹caso contrário, i.e., se $\mathcal{H}(k_0) \neq 0$ basta substituir \mathcal{H} por $\mathcal{H} - \mathcal{H}(k_0)$, daí $(\mu)_i^{\mathcal{H} - \mathcal{H}(k_0)} = e^{-\beta(\mathcal{H}(i) - \mathcal{H}(k_0))} (\nu)_i / (e^{\beta \mathcal{H}(k_0)} Z(\beta)) = (\mu)_i^{\mathcal{H}}$ e, da mesma forma, $(Q)_{ij}^{\mathcal{H} - \mathcal{H}(k_0)} = e^{-\beta((\mathcal{H}(j) - \mathcal{H}(k_0)) - (\mathcal{H}(i) - \mathcal{H}(k_0)))} (A)_{ij} = e^{-\beta(\mathcal{H}(j) - \mathcal{H}(i))} (A)_{ij} = (Q)_{ij}^{\mathcal{H}}$. Donde, a desigualdade 5.17 permanece, exatamente, a mesma.

Daí, para qualquer (i, j) , $e(p(i, j)) = e(p(i, k_0)) \vee e(p(j, k_0))$, donde fica claro que $E(P) = \max_{i \in S} e(p(i, k_0))$.

Finalmente, já que, para cada i , $e(p(i, k_0)) = \mathcal{H}(l) - \mathcal{H}(i)$, onde l é o mais alto ponto ao longo do caminho $p(i, k_0)$, concluimos nossa explanação.

Quando S é finito, podemos mostrar que, ao menos para β 's suficientemente grandes, a desigualdade 5.17 nós dá uma boa estimativa para λ_β . Para sermos mais precisos, vejamos o teorema que segue abaixo.

Teorema 5.3.1

Assuma que S seja finito e que $\mathcal{Q}(\beta)$ seja dada por 5.8. Seja $m = \min_{i \in S} \mathcal{H}(i)$ e $S_0 = \{i \in S : \mathcal{H}(i) = m\}$, e seja ϵ o menor valor de $E(P)$ tomado sobre todas seleções de caminhos acessíveis P . Então, $\epsilon \geq -m$ e $\epsilon = -m$ se, e somente se, para cada $(i, j) \in S \times S_0$ existe um caminho acessível p de i a j tal que $Elev(p) = \mathcal{H}(i)$.

Mais geralmente, qualquer que seja o valor de ϵ , existem constantes $0 < c_- \leq c_+ < \infty$, independentes de \mathcal{H} , tais que

$$c_- e^{-\beta(\epsilon+m)} \leq \lambda_\beta \leq c_+ e^{-\beta(\epsilon+m)}, \quad \forall \beta \geq 0.$$

◇

Prova

Primeiramente note que, nem $\mu(\beta)$ nem $\mathcal{Q}(\beta)$ mudam se \mathcal{H} for substituído por $\mathcal{H} - m$, enquanto que ϵ muda para $\epsilon + m$. Com efeito, $\forall i \in S$

$$(\mu(\beta))_i^{\mathcal{H}-m} = \frac{1}{Z(\beta)} e^{-\beta(\mathcal{H}(i)-m)} (\nu)_i = \frac{e^{-\beta\mathcal{H}(i)} (\nu)_i e^{\beta m}}{\sum_{i \in S} e^{-\beta\mathcal{H}(i)} (\nu)_i e^{\beta m}} = \frac{e^{-\beta\mathcal{H}(i)}}{\sum_{i \in S} e^{-\beta\mathcal{H}(i)}} = (\mu(\beta))_i^{\mathcal{H}}.$$

Logo, $(\mu(\beta))_i^{\mathcal{H}-m} = (\mu(\beta))_i^{\mathcal{H}}$, onde estamos supondo estar claro o emprego do sobrescrito na expressão acima.

Também, para $i \neq j$ temos

$$(Q(\beta))_{ij}^{\mathcal{H}-m} = e^{-\beta[(\mathcal{H}(j)-m)-(\mathcal{H}(i)-m)]^+} (A)_{ij} = e^{-\beta(\mathcal{H}(j)-\mathcal{H}(i))^+} (A)_{ij} = (Q(\beta))_{ij}^{\mathcal{H}},$$

e portanto $(Q(\beta))_{ij}^{\mathcal{H}-m} = (Q(\beta))_{ij}^{\mathcal{H}}$.

Finalmente, utilizando as relações 5.11 e 5.12, temos

$$\begin{aligned} \epsilon^{\mathcal{H}-m} &= \min_P E(P)^{\mathcal{H}-m} = \min_P \left(\sup_{p \in P} e(p)^{\mathcal{H}-m} \right) = \\ &= \min_P \left(\sup_{p \in P} [Elev(p)^{\mathcal{H}-m} - (\mathcal{H}(i_0) - m) - (\mathcal{H}(i_n) - m)] \right) = \\ &= \min_P \left(\sup_{p \in P} \left[\max_{0 \leq j \leq n} (\mathcal{H}(j) - m) - \mathcal{H}(i_0) - \mathcal{H}(i_n) + 2m \right] \right) = \\ &= \min_P \left(\sup_{p \in P} \left[\max_{0 \leq j \leq n} (\mathcal{H}(j)) - \mathcal{H}(i_0) - \mathcal{H}(i_n) + m \right] \right) = \\ &= \min_P \left(\sup_{p \in P} \left[\max_{0 \leq j \leq n} (\mathcal{H}(j)) - \mathcal{H}(i_0) - \mathcal{H}(i_n) \right] \right) + m = \\ &= \min_P \left(\sup_{p \in P} [Elev(p)^{\mathcal{H}} - \mathcal{H}(i_0) - \mathcal{H}(i_n)] \right) + m = \\ &= \min_P \left(\sup_{p \in P} e(p)^{\mathcal{H}} \right) + m = \min_P E(P)^{\mathcal{H}} + m = \epsilon^{\mathcal{H}} + m. \end{aligned}$$

O que mostra que $\epsilon^{\mathcal{H}-m} = \epsilon^{\mathcal{H}} + m$.

Posto toda esta argumentação inicial, podemos supor então, sem perda de generalidade, que $m = 0$.

Daí, escolha uma seleção de caminhos acessíveis $P = \{p(i, j) : (i, j) \in S^2\}$ de modo que, para cada $(i, j) \in S^2$, $p(i, j)$ minimize $e(p)$ sobre todos os caminhos acessíveis de i para j . Agora, seja $k_0 \in S_0$ fixo.

Pelo que vimos na metáfora acima,

$$\epsilon = \max_{i \in S} e(p(i, k_0)). \quad (5.18)$$

Em particular, como $e(p(i, k_0)) = Elev(p(i, k_0)) - \mathcal{H}(i) - \mathcal{H}(k_0) = Elev(p(i, k_0)) - \mathcal{H}(i) \geq 0$, pois $k_0 \in S_0$ e $Elev(p(i, k_0)) = \max_{0 \leq j \leq n} \mathcal{H}(j) \forall i \in S$, provando assim que $\epsilon = \max_{i \in S} e(p(i, k_0)) \geq 0$ e, é claro que, $\epsilon = 0$ se, e somente se, $Elev(p(i, k_0)) - \mathcal{H}(i) = 0 \Leftrightarrow Elev(p(i, k_0)) = \mathcal{H}(i) \forall i \in S$.

Para estimar o limite inferior de λ_β , observe que, como estamos supondo $m = 0$, e $Z(\beta) = \sum_{i \in S} e^{-\beta \mathcal{H}(i)} (\nu)_i \geq e^{-\beta \mathcal{H}(k_0)} (\nu)_{k_0} = (\nu)_{k_0} > 0$ vem, por 5.17, que devemos tomar $c_- = \frac{(\nu)_{k_0}}{W(P)}$, que não depende de \mathcal{H} .

Finalmente, para provar o limite superior, escolha $l_0 \in S \setminus \{k_0\}$ de modo que $e(p_0) = \epsilon$, onde $p_0 := p(l_0, k_0)$, e seja Γ o conjunto dos $i \in S$ com a propriedade que ou $i = k_0$ ou $Elev(p(i)) < Elev(p_0)$ para o caminho $p(i) := p(i, k_0) \in P$ de i a k_0 , e defina $f := \chi_\Gamma$, i.e., a função indicador do conjunto Γ , tal que, $f(i) = 1$ se $i \in \Gamma$ e $f(i) = 0$ se $i \in \Gamma^c$.

Então, já que $k_0 \in \Gamma$ e $l_0 \in \Gamma^c$, vem (ver equação 2.8)

$$\begin{aligned} Var_\beta(f) &=^0 \langle f^2 \rangle - \langle f \rangle^2 =^1 \sum_{i \in S} f^2(i) (\mu(\beta))_i - \left[\sum_{i \in S} f(i) (\mu(\beta))_i \right]^2 =^2 \\ &= \sum_{i \in \Gamma} (\mu(\beta))_i - \left[\sum_{i \in \Gamma} (\mu(\beta))_i \right]^2 =^3 \sum_{i \in \Gamma} (\mu(\beta))_i \left[1 - \sum_{i \in \Gamma} (\mu(\beta))_i \right] =^4 \\ &= \left(\sum_{i \in \Gamma} (\mu(\beta))_i \right) \left(\sum_{j \in \Gamma^c} (\mu(\beta))_j \right) \geq (\mu(\beta))_{k_0} (\mu(\beta))_{l_0} = \\ &= \frac{e^{-\beta \mathcal{H}(k_0)} (\nu)_{k_0}}{Z(\beta)} \frac{e^{-\beta \mathcal{H}(l_0)} (\nu)_{l_0}}{Z(\beta)} = \frac{(\nu)_{k_0} (\nu)_{l_0} e^{-\beta(\mathcal{H}(k_0) + \mathcal{H}(l_0))}}{Z(\beta)^2}, \end{aligned}$$

onde 0 e 1 seguem, respectivamente, da definição de variância e valor esperado dada na definição 2.3.1, 2 segue do fato que podemos decompor $S = \Gamma \cup \Gamma^c$ e $f := \chi_\Gamma$, em 3 apenas fatoramos a soma, já em 4 usamos o fato que $1 = \sum_{i \in S} (\mu(\beta))_i = \sum_{i \in \Gamma} (\mu(\beta))_i + \sum_{i \in \Gamma^c} (\mu(\beta))_i$ o que implica $1 - \sum_{i \in S} (\mu(\beta))_{i \in \Gamma} = \sum_{j \in S} (\mu(\beta))_{j \in \Gamma^c}$, onde nesta última passagem fizemos uma mudança de índices ($i \leftrightarrow j$!).

Ao mesmo tempo, pela equação 5.10, temos

$$\begin{aligned} \xi_\beta(f, f) &= \frac{1}{2} \sum_{j \neq i} (\mu(\beta))_i (Q(\beta))_{ij} (f(j) - f(i))^2 = \sum_{(i, j) \in \Gamma \times \Gamma^c} (\mu(\beta))_i (Q(\beta))_{ij} = \\ &= \sum_{(i, j) \in \Gamma \times \Gamma^c} \frac{1}{Z(\beta)} e^{-\beta \mathcal{H}(i)} (\nu)_i e^{-\beta(\mathcal{H}(j) - \mathcal{H}(i))^+} (A)_{ij} = \end{aligned}$$

$$= \frac{1}{Z(\beta)} \sum_{(i,j) \in \Gamma \times \Gamma^c} (\nu)_i (A)_{ij} e^{-\beta \mathcal{H}(i) \vee \mathcal{H}(j)}.$$

Mas, se $i \in \Gamma \setminus \{k_0\}$, $j \in \Gamma^c$ e $(A)_{ij} > 0$ então $\mathcal{H}(j) \geq Elev(p_0)$. Para ver isto, considere o caminho q obtido indo-se em uma etapa de j para i e, então, seguindo $p(i)$ de i para k_0 . Claramente, q é um caminho acessível de j para k_0 e, deste modo, $Elev(q) \geq Elev(p(j)) \geq Elev(p_0)$. Mas, isto significa que

$$Elev(p_0) \leq Elev(p(j)) \leq Elev(q) = Elev(p(j)) \vee \mathcal{H}(j),$$

que, junto com $Elev(p(i)) < Elev(p_0)$ (ver definição de Γ), nos leva a concluir que $\mathcal{H}(j) \geq Elev(p_0)$. Da mesma forma, vemos que $\mathcal{H}(j) \geq Elev(p_0)$ se $j \in \Gamma^c$ e $(A)_{k_0 j} > 0$ já que, neste caso, o caminho $p(j, k_0)$ é acessível e

$$\mathcal{H}(j) = Elev(p(j, k_0)) \geq Elev(p(j)) \geq Elev(p_0).$$

Conseqüentemente, a expressão anterior de $\xi_\beta(f, f)$ pode ser escrita como

$$\xi_\beta(f, f) \leq \frac{1}{Z(\beta)} e^{-\beta Elev(p_0)} \sum_{(i,j) \in \Gamma \times \Gamma^c} (\nu)_i (A)_{ij}.$$

Como, por hipótese, $\epsilon = e(p_0) = Elev(p_0) - \mathcal{H}(k_0) - \mathcal{H}(l_0)$, concluímos que

$$\begin{aligned} \lambda_\beta &= \inf\{\xi_\beta(f, f) : Var_\beta(f) = 1\} \leq \frac{\xi_\beta(f, f)}{Var_\beta(f)} \leq \\ &\leq \frac{1}{Z(\beta)} e^{-\beta Elev(p_0)} \left(\sum_{(i,j) \in \Gamma \times \Gamma^c} (\nu)_i (A)_{ij} \right) \frac{Z(\beta)^2}{(\nu)_{k_0} (\nu)_{l_0} e^{\beta(\mathcal{H}(k_0) + \mathcal{H}(l_0))}} = \\ &= \frac{Z(\beta)}{(\nu)_{k_0} (\nu)_{l_0}} \left(\sum_{(i,j) \in \Gamma \times \Gamma^c} (\nu)_i (A)_{ij} \right) e^{-\beta Elev(p_0)} e^{\beta(Elev(p_0) - \epsilon)} = \\ &= \frac{Z(\beta)}{(\nu)_{k_0} (\nu)_{l_0}} \left(\sum_{(i,j) \in \Gamma \times \Gamma^c} (\nu)_i (A)_{ij} \right) e^{-\beta \epsilon}. \end{aligned}$$

Finalmente, note que

$$\begin{aligned} Z(\beta) &= |Z(\beta)| = \left| \sum_{i \in S} e^{-\beta \mathcal{H}(i)} (\nu)_i \right| \leq \sum_{i \in S} |e^{-\beta \mathcal{H}(i)} (\nu)_i| = \sum_{i \in S} |e^{-\beta \mathcal{H}(i)}| |(\nu)_i| \leq \\ &\leq \sum_{i \in S} |(\nu)_i| = \| \nu \|_\nu, \end{aligned}$$

onde, $\| f \|_\nu := \sum_{i \in S} |f(i)|$ é a *norma variacional* (veja subseção 2.1.2 em [24]).

Logo, devemos tomar

$$c_+ = \frac{\| \nu \|_\nu}{(\nu)_{k_0} (\nu)_{l_0}} \left(\sum_{(i,j) \in \Gamma \times \Gamma^c} (\nu)_i (A)_{ij} \right).$$

Concluindo, deste modo, a demonstração do limite superior e, por conseguinte, a do teorema. \square

6 Aplicações

6.1 Introdução

Neste capítulo vamos construir, explicitamente, a *Dinâmica de Glauber*, da maneira que foi proposta no capítulo 5, para dois sistemas físicos, a saber, *Sistema de Partículas com Dois Níveis de Energia* e o *Modelo de Núcleos Magnéticos Localizados*.

De fato, vamos obter os correspondentes semi-grupos a tempo contínuo \mathcal{P}^t , para $t \geq 0$, que nos fornecem as probabilidades de transição para cada um dos sistemas físicos mencionados, e que sejam reversíveis com respeito ao Estado de Gibbs (ver definição 5.2.1).

Será dada uma interpretação para alguns casos limite de \mathcal{P}^t , em cada sistema.

Na parte final de cada exemplo, mostraremos que a dinâmica de \mathcal{P}^t agindo sobre certas variáveis aleatórias, no sentido da relação 2.10, converge de forma bem rápida (exponencial) ao valor esperado destas variáveis aleatórias e vamos estimar a taxa de tal convergência.

6.2 Sistema de Partículas com Dois Níveis de Energia

6.2.1 Modelagem

Vamos considerar um gás de N partículas semi-clássicas que podem ser encontradas em dois estados, com energia $\varepsilon_1 = 0$ e $\varepsilon_2 = \varepsilon > 0$, respectivamente, numa situação em que a energia total do sistema pode flutuar, mas a temperatura permanece com valor fixo T .

Um estado microscópico do sistema global fica caracterizado pela atribuição do estado de energia de cada partícula. Podemos, então, introduzir uma variável s_j , associada à j -ésima partícula, com $j \in \{1, 2, \dots, N\}$, tal que $s_j = 0$, quando a partícula j tiver energia nula, e $s_j = 1$, quando a partícula j tiver energia $\varepsilon > 0$.

Um estado microscópico do sistema, caracterizado pelo conjunto de valores $\{s_j\}_{j \in \{1, 2, \dots, N\}}$, tem *Hamiltoniano global*

$$\tilde{\mathcal{H}}(\{s_j\}) = \sum_{i=1}^N \varepsilon s_i. \quad (6.1)$$

Concluindo, desta forma, a modelagem do sistema.

6.2.2 Construção das Probabilidades de Transição

Com o intento de obter, explicitamente, o semi-grupo a tempo contínuo \mathcal{P}^t , vamos seguir o roteiro introduzido na seção 5.2.

No que consiste, note que $S = \{0, 1\}$ pois, neste caso, existem apenas dois estados possíveis de energia, a saber, 0 e $\varepsilon > 0$ que, como vimos na modelagem, estão associados aos valores 0 e 1 respectivamente. Além disso, vamos considerar ν uniforme, i.e., $(\nu)_i = 1 \forall i \in S$. Note que, $\tilde{\mathcal{H}} : S^N \rightarrow \mathbb{R}^+$ é tal que $\tilde{\mathcal{H}} = \Sigma \mathcal{H}$, com $\mathcal{H} : S \rightarrow \mathbb{R}^+$ sendo o hamiltoniano de partícula única (dado por 6.3).

Desta forma, a função canônica de partição (global) pode ser escrita na forma

$$\begin{aligned} Z(\beta) &= \sum_{\{s_j\}} e^{-\beta \tilde{\mathcal{H}}(\{s_j\})} = \sum_{s_1 \in \{0,1\}} \sum_{s_2 \in \{0,1\}} \dots \sum_{s_N \in \{0,1\}} e^{-\sum_{i=1}^N \beta \varepsilon s_i} = \\ &= \sum_{s_1 \in \{0,1\}} e^{-\beta \varepsilon s_1} \sum_{s_2 \in \{0,1\}} e^{-\beta \varepsilon s_2} \dots \sum_{s_N \in \{0,1\}} e^{-\beta \varepsilon s_N} = \\ &= \left(\sum_{s \in \{0,1\}} e^{-\beta \varepsilon s} \right)^N = Z_1(\beta)^N, \end{aligned}$$

mostrando que a função canônica de partição do sistema (global) se fatora, donde

$$Z_1(\beta) = \sum_{s \in \{0,1\}} e^{-\beta \varepsilon s} = 1 + e^{-\beta \varepsilon}. \quad (6.2)$$

Ou seja, tudo se passa como se estivéssemos considerando apenas uma única partícula e, de fato, é 6.2 que vamos considerar (todo sistema cuja função canônica de partição se fatora pode ser abordado desta forma!).

Neste ponto de vista, a equação 6.1 tornar-se-á $\mathcal{H} : S \rightarrow \mathbb{R}^+$, dada por

$$\mathcal{H}(s) = \varepsilon s. \quad (6.3)$$

Daí, o Estado de Gibbs (equação 5.3) fica, $\forall s \in \{0, 1\} = S$,

$$(\mu(\beta))_s = \frac{e^{-\beta \varepsilon s}}{Z_1(\beta)} = \frac{e^{-\beta \varepsilon s}}{1 + e^{-\beta \varepsilon}}. \quad (6.4)$$

Avançando nas idéias do capítulo 5, vamos considerar a matriz \mathcal{A} lá mencionada. Como estamos assumindo ν uniforme, temos de tomar \mathcal{A} simétrica para que satisfaça a condição do balanço detalhado 5.6. Explicitamente, considere a seguinte matriz \mathcal{A}

$$\mathcal{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \quad (6.5)$$

Prosseguindo, para construir a Dinâmica de Glauber devemos obter a \mathcal{Q} -matriz $\mathcal{Q}(\beta)$, conforme a equação 5.8. Com efeito, levando-se em conta as relações 6.3 e 6.5, temos

$$\begin{aligned} (\mathcal{Q}(\beta))_{01} &= e^{-\beta(\mathcal{H}(1) - \mathcal{H}(0))^+} (\mathcal{A})_{01} = e^{-\beta(\varepsilon - 0)^+} 1 = e^{-\beta \varepsilon}; \\ (\mathcal{Q}(\beta))_{10} &= e^{-\beta(\mathcal{H}(0) - \mathcal{H}(1))^+} (\mathcal{A})_{10} = e^{-\beta(0 - \varepsilon)^+} 1 = e^{-\beta(-\varepsilon)^+} = 1; \\ (\mathcal{Q}(\beta))_{00} &= -(\mathcal{Q}(\beta))_{01} = -e^{-\beta \varepsilon}; \\ (\mathcal{Q}(\beta))_{11} &= -(\mathcal{Q}(\beta))_{10} = -1. \end{aligned}$$

Portanto,

$$\mathcal{Q}(\beta) = \begin{bmatrix} -e^{-\beta\varepsilon} & e^{-\beta\varepsilon} \\ 1 & -1 \end{bmatrix}. \quad (6.6)$$

O polinômio característico associado à matriz 6.6 é $p(\lambda) = \lambda^2 - \lambda \text{Tr}(\mathcal{Q}(\beta)) + \det(\mathcal{Q}(\beta))$, mas $\text{Tr}(\mathcal{Q}(\beta)) = -(1 + e^{-\beta\varepsilon})$ e $\det(\mathcal{Q}(\beta)) = 0$, portanto

$$p(\lambda) = \lambda^2 + \lambda(1 + e^{-\beta\varepsilon}). \quad (6.7)$$

Daí, temos como autovalores, i.e., solução de $p(\lambda) = 0$,

$$\lambda_1 = 0 \quad \text{e} \quad \lambda_2 = -(1 + e^{-\beta\varepsilon}) = -Z_1(\beta), \quad (6.8)$$

com respectivos autovetores

$$v_1 = (1, 1)^T \quad \text{e} \quad v_2 = (-e^{-\beta\varepsilon}, 1)^T. \quad (6.9)$$

Logo, a matriz $\mathcal{Q}(\beta)$, dada por 6.6, é *conjugada* a uma matriz diagonal D , i.e., existe uma matriz B invertível, (mudança de coordenadas) cujas colunas são v_1 e v_2 respectivamente, e tal que $\mathcal{Q}(\beta) = BDB^{-1}$, onde

$$D = \begin{bmatrix} 0 & 0 \\ 0 & -Z_1(\beta) \end{bmatrix}, \quad (6.10)$$

$$B = \begin{bmatrix} 1 & -e^{-\beta\varepsilon} \\ 1 & 1 \end{bmatrix} \quad \text{e} \quad B^{-1} = \begin{bmatrix} \frac{1}{Z_1(\beta)} & \frac{e^{-\beta\varepsilon}}{Z_1(\beta)} \\ -\frac{1}{Z_1(\beta)} & \frac{1}{Z_1(\beta)} \end{bmatrix}. \quad (6.11)$$

Finalmente, como vimos na subseção 1.3.2, a matriz de transição probabilidades a tempo contínuo é dada por $\mathcal{P}^t = e^{t\mathcal{Q}(\beta)} \forall t \geq 0$, logo, por 6.10 e 6.11 vem, lembrando que $\exp(\text{diag}\{a_1, a_2\}) = \text{diag}\{e^{a_1}, e^{a_2}\}$, que

$$\begin{aligned} \mathcal{P}^t &= e^{t\mathcal{Q}(\beta)} = B e^{tD} B^{-1} = \begin{bmatrix} 1 & -e^{-\beta\varepsilon} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & e^{-tZ_1(\beta)} \end{bmatrix} \begin{bmatrix} \frac{1}{Z_1(\beta)} & \frac{e^{-\beta\varepsilon}}{Z_1(\beta)} \\ -\frac{1}{Z_1(\beta)} & \frac{1}{Z_1(\beta)} \end{bmatrix} = \\ &= \begin{bmatrix} \frac{1}{Z_1(\beta)} + \frac{1}{Z_1(\beta)} e^{-\beta\varepsilon - tZ_1(\beta)} & \frac{e^{-\beta\varepsilon}}{Z_1(\beta)} - \frac{1}{Z_1(\beta)} e^{-\beta\varepsilon - tZ_1(\beta)} \\ \frac{1}{Z_1(\beta)} - \frac{1}{Z_1(\beta)} e^{-tZ_1(\beta)} & \frac{e^{-\beta\varepsilon}}{Z_1(\beta)} + \frac{1}{Z_1(\beta)} e^{-tZ_1(\beta)} \end{bmatrix}. \end{aligned}$$

Donde vem, $\forall t \geq 0$

$$\mathcal{P}^t = \frac{1}{Z_1(\beta)} \begin{bmatrix} 1 + e^{-\beta\varepsilon - tZ_1(\beta)} & e^{-\beta\varepsilon} - e^{-\beta\varepsilon - tZ_1(\beta)} \\ 1 - e^{-tZ_1(\beta)} & e^{-\beta\varepsilon} + e^{-tZ_1(\beta)} \end{bmatrix}. \quad (6.12)$$

Note que o semi-grupo a parâmetro contínuo 6.12 satisfaz a condição do balanço detalhado para o estado de Gibbs 6.4. Com efeito,

$$\begin{aligned} (\mu(\beta))_0 (P^t)_{01} &= \frac{1}{Z_1(\beta)} \left(\frac{e^{-\beta\varepsilon}}{Z_1(\beta)} - \frac{e^{-\beta\varepsilon - tZ_1(\beta)}}{Z_1(\beta)} \right) = \frac{e^{-\beta\varepsilon}}{Z_1(\beta)} \left(\frac{1}{Z_1(\beta)} - \frac{e^{-tZ_1(\beta)}}{Z_1(\beta)} \right) = \\ &= (\mu(\beta))_1 (P^t)_{10}. \end{aligned}$$

Logo, o processo estocástico de Markov a tempo contínuo $\{X_t\}_{t \geq 0}$, onde a variável aleatória X_t indica o estado de energia de cada partícula no instante t , é uma Dinâmica de Glauber (ver definição 5.1.1).

Existem alguns limites interessantes de \mathcal{P}^t , por exemplo, quando $t \rightarrow \infty$, temos

$$\lim_{t \rightarrow \infty} \mathcal{P}^t = \begin{bmatrix} \frac{1}{Z_1(\beta)} & \frac{e^{-\beta\varepsilon}}{Z_1(\beta)} \\ \frac{1}{Z_1(\beta)} & \frac{e^{-\beta\varepsilon}}{Z_1(\beta)} \end{bmatrix} = \begin{bmatrix} (\mu(\beta))_0 & (\mu(\beta))_1 \\ (\mu(\beta))_0 & (\mu(\beta))_1 \end{bmatrix}. \quad (6.13)$$

Ou seja, $\mu(\beta) = ((\mu(\beta))_0, (\mu(\beta))_1)$ é, por construção, o (único) ponto fixo para \mathcal{P}^t . Em outras palavras, $\mu(\beta)$ é o estado de equilíbrio do sistema e daí, $\forall p$ vetor de probabilidade inicial, $\lim_{t \rightarrow \infty} p\mathcal{P}^t = \mu(\beta)$ (veja teorema 1.3.22).

Além disso, note que sempre temos $(\mu(\beta))_0 \geq (\mu(\beta))_1$, ou seja, o nível de energia mais baixa (0) está sempre mais populado de partículas do que o nível mais alto ($\varepsilon > 0$), tal fato particular vai ao encontro do princípio físico geral de que, *grosso modo*, “*todo sistema físico tende a ocupar o estado de menor energia*”. Podemos ver por 6.13 que, no equilíbrio, dado que uma partícula está no estado fundamental (0), é mais provável que lá ela permaneça, do que ela *salte* para o estado *excitado* ($\varepsilon > 0$). Do mesmo modo, dado que uma partícula está no estado excitado ($\varepsilon > 0$), é mais provável que ela vá para o estado fundamental (0), do que ela permaneça neste estado.

Lembrado que $\beta = 1/(k_B T)$, verifica-se a igualdade de $(\mu(\beta))_0 \geq (\mu(\beta))_1$ apenas para $T \rightarrow \infty$ (i.e., $\beta \rightarrow 0$) e, neste caso, $\lim_{\beta \rightarrow 0} (\mu(\beta))_0 = \lim_{\beta \rightarrow 0} \frac{1}{1+e^{-\beta\varepsilon}} = 1/2$ e $\lim_{\beta \rightarrow 0} (\mu(\beta))_1 = \lim_{\beta \rightarrow 0} \frac{e^{-\beta\varepsilon}}{1+e^{-\beta\varepsilon}} = 1/2$, indicando que os dois níveis de energia estão igualmente populados. Para $T \rightarrow 0$ (i.e., $\beta \rightarrow \infty$), $\lim_{\beta \rightarrow \infty} (\mu(\beta))_0 = \lim_{\beta \rightarrow \infty} \frac{1}{1+e^{-\beta\varepsilon}} = 1$ e $\lim_{\beta \rightarrow \infty} (\mu(\beta))_1 = \lim_{\beta \rightarrow \infty} \frac{e^{-\beta\varepsilon}}{1+e^{-\beta\varepsilon}} = 0$, indicando que todas partículas estão no estado fundamental, com energia nula.

Vamos agora nos perguntar sobre a validade do limite

$$\lim_{t \rightarrow \infty} \|\mathcal{P}^t \mathcal{H} - \langle \mathcal{H} \rangle_{\mu(\beta)}\|_{2, \mu(\beta)} = 0, \quad (6.14)$$

onde a notação $\mathcal{P}^t \mathcal{H}$ é a mesma já empregada no capítulo 2 (equação 2.10) que quer dizer

$$\mathcal{P}^t \mathcal{H}(i) := \sum_{j \in S} \mathcal{H}(j) (P^t)_{ij}, \quad (6.15)$$

a norma e o valor esperado apresentados na equação 6.14 são, a norma e o valor esperado em $L^2(\mu(\beta))$ (ver definição 2.3.1), com $\mu(\beta)$ sendo o Estado de Gibbs dado em 6.4, \mathcal{H} é o Hamiltoniano dado por 6.3. Sendo válido o limite 6.14 que estimativa podemos dar a taxa de tal convergência?

Como $\mathcal{H} \in L^2(\mu(\beta))$ vem, pelos argumentos do capítulo 3, que vale o limite 6.14. Mais do que isto, vale que

$$\|\mathcal{P}^t \mathcal{H} - \langle \mathcal{H} \rangle_{\mu}\|_{2, \mu} \leq e^{-\lambda_\beta t} \|\mathcal{H} - \langle \mathcal{H} \rangle_{\mu}\|_{2, \mu},$$

onde a estimativa para a taxa λ_β é dada em 5.17, i.e., $\lambda_\beta \geq Z_1(\beta) e^{-\beta E(P)}/W(P)$.

Para explicitarmos tal estimativa, notemos que a única seleção de caminhos acessíveis P é $P = \{p(0, 1) = (0, 1), p(1, 0) = (1, 0)\}$. Mas $E(P) = \sup_{p \in P} e(p)$. Então, pela definição de $e(p)$ dada em 5.12 temos que

$$e(p(0, 1)) = \varepsilon - \mathcal{H}(0) - \mathcal{H}(1) = \varepsilon - 0 - \varepsilon = 0,$$

e

$$e(p(1, 0)) = \varepsilon - \mathcal{H}(1) - \mathcal{H}(0) = \varepsilon - \varepsilon - 0 = 0.$$

Donde que, $E(\mathbf{P}) = \max\{0, 0\} = 0$.

Vimos na seção 5.3 que, para um caminho $p = (i_0, \dots, i_n)$

$$\omega(p) = \sum_{m=1}^n \frac{(\nu)_{i_0}(\nu)_{i_n}}{(\nu)_{i_{m-1}}(A)_{i_{m-1}i_m}} \stackrel{!}{=} 1,$$

onde 1 se deve ao fato que em nosso caso particular $(\nu)_i = (A)_{ij} \equiv 1 \forall (i, j) \in \{0, 1\}^2$. Donde, vem que (veja equação 5.16) $W(\mathbf{P}) = \max\{1, 1\} = 1$. Daí, retornando à estimativa, vemos que $\lambda_\beta \geq Z_1(\beta) e^{-\beta E(\mathbf{P})} / W(\mathbf{P}) = Z_1(\beta)$ i.e., $\lambda_\beta \geq Z_1(\beta)$.

□

6.3 Modelo de Núcleos Magnéticos Localizados

6.3.1 Modelagem

Os núcleos dos átomos de certos sólidos têm *spin* $\sigma = 1$. De acordo com a teoria quântica, cada núcleo pode ter três estados quânticos de spin, a saber, $\sigma = 1$, $\sigma = 0$ e $\sigma = -1$. Esse número quântico mede a projeção do spin nuclear ao longo do eixo cristalino do sólido. Como a distribuição de carga nuclear não é esfericamente simétrica, a energia do núcleo depende da orientação de seu spin em relação ao *campo elétrico* local.

Assim, um núcleo nos estados $\sigma = \pm 1$ tem energia $D > 0$ e um núcleo no estado $\sigma = 0$ tem energia nula. Considere agora este sistema, com N núcleos magnéticos, a temperatura T . O *Hamiltoniano global* do sistema pode ser escrito como

$$\tilde{\mathcal{H}}(\{\sigma_j\}) = D \sum_{i=1}^N \sigma_i^2, \quad (6.16)$$

onde D é positivo e o estado microscópico do sistema fica caracterizado pelo conjunto de variáveis de spin $\{\sigma_j\}$ com $\sigma_j \in \{1, 0, -1\} = S$, para qualquer *sítio* (localização) $j \in \{1, 2, \dots, N\}$.

6.3.2 Construção das Probabilidades de Transição

Procederemos de forma análoga à subseção 6.2.2. No que consiste, para obter o semi-grupo a tempo contínuo \mathcal{P}^t , vamos considerar $S = \{1, 0, -1\}$ e $\nu = (1, 1, 1)$.

Daí, a função canônica de partição (global) pode ser escrita como

$$\begin{aligned} Z(\beta) &= \sum_{\{\sigma_i\}} e^{-\beta \tilde{\mathcal{H}}(\{\sigma_i\})} = \sum_{\sigma_1 \in \{1, 0, -1\}} \dots \sum_{\sigma_N \in \{1, 0, -1\}} e^{-\beta D \sum_{i=1}^N \sigma_i^2} = \\ &= \sum_{\sigma_1 \in \{1, 0, -1\}} e^{-\beta D \sigma_1^2} \dots \sum_{\sigma_N \in \{1, 0, -1\}} e^{-\beta D \sigma_N^2} = Z_1(\beta)^N, \end{aligned}$$

onde

$$Z_1(\beta) = \sum_{\sigma \in \{1, 0, -1\}} e^{-\beta D \sigma^2} = e^{-\beta D} + 1 + e^{-\beta D} = 1 + 2e^{-\beta D}. \quad (6.17)$$

Ou seja, novamente a função canônica de partição do sistema se fatora e, daí, temos de utilizar apenas 6.17 (isto se deve ao fato que os sítios são *não-interagentes*).

Logo, o Hamiltoniano $H : S \rightarrow \mathbb{R}^+$ de cada sítio será

$$\mathcal{H}(\sigma) = D\sigma^2, \quad (6.18)$$

para $\sigma \in S$.

Portanto, o Estado de Gibbs $\mu(\beta)$ será dado, $\forall \sigma \in S$, por

$$(\mu(\beta))_\sigma = \frac{e^{-\beta D\sigma^2}}{Z_1(\beta)} = \frac{e^{-\beta D\sigma^2}}{1 + 2e^{-\beta D}}. \quad (6.19)$$

Seguindo a ordem que adotamos na subseção 6.2.2, devemos construir a matriz \mathcal{A} satisfazendo a condição do balanço detalhado com ν , como estamos considerando ν uniforme, \mathcal{A} deve ser simétrica, além disso, como S é constituído de tres estados, \mathcal{A} deve ser uma matriz 3×3 . De fato, vamos assumir \mathcal{A} da seguinte forma

$$\mathcal{A} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}. \quad (6.20)$$

Agora, para obter a \mathcal{Q} -matriz $\mathcal{Q}(\beta)$, devemos levar em consideração as expressões 6.18 e 6.20, e também lembrar que $D > 0$. Então,

$$\begin{aligned} (Q(\beta))_{10} &= e^{-\beta(\mathcal{H}(0)-\mathcal{H}(1))^+} (A)_{10} = e^{-\beta(-D)^+} 1 = 1; \\ (Q(\beta))_{1-1} &= e^{-\beta(\mathcal{H}(-1)-\mathcal{H}(1))^+} (A)_{1-1} = e^{-\beta(D-D)^+} 1 = 1; \\ (Q(\beta))_{01} &= e^{-\beta(\mathcal{H}(1)-\mathcal{H}(0))^+} (A)_{01} = e^{-\beta(D)^+} 1 = e^{-\beta D}; \\ (Q(\beta))_{0-1} &= e^{-\beta(\mathcal{H}(-1)-\mathcal{H}(0))^+} (A)_{0-1} = e^{-\beta(D)^+} 1 = e^{-\beta D}; \\ (Q(\beta))_{-11} &= e^{-\beta(\mathcal{H}(1)-\mathcal{H}(-1))^+} (A)_{-11} = e^{-\beta(D-D)^+} 1 = 1; \\ (Q(\beta))_{-10} &= e^{-\beta(\mathcal{H}(0)-\mathcal{H}(-1))^+} (A)_{-10} = e^{-\beta(-D)^+} 1 = 1; \\ (Q(\beta))_{11} &= -(Q(\beta))_{10} - (Q(\beta))_{1-1} = -1 - 1 = -2; \\ (Q(\beta))_{00} &= -(Q(\beta))_{01} - (Q(\beta))_{0-1} = -e^{-\beta D} - e^{-\beta D} = -2e^{-\beta D}; \\ (Q(\beta))_{-1-1} &= -(Q(\beta))_{-11} - (Q(\beta))_{-10} = -1 - 1 = -2. \end{aligned}$$

Daí, a matriz $\mathcal{Q}(\beta)$ será dada, em forma matricial, por

$$\mathcal{Q}(\beta) = \begin{bmatrix} -2 & 1 & 1 \\ e^{-\beta D} & -2e^{-\beta D} & e^{-\beta D} \\ 1 & 1 & -2 \end{bmatrix}. \quad (6.21)$$

O polinômio característico associado à matriz 6.21 é

$$p(\lambda) = -e^{-\beta D} \lambda(3 + \lambda)(2 + e^{\beta D}(1 + \lambda)). \quad (6.22)$$

Os autovalores são as soluções de $p(\lambda) = 0$, i.e., são os valores

$$\lambda_1 = -3, \quad \lambda_2 = 0 \quad \text{e} \quad \lambda_3 = -Z_1(\beta), \quad (6.23)$$

onde $Z_1(\beta)$ é a função canônica de partição dada por 6.17.

Seus respectivos autovetores são

$$v_1 = (-1, 0, 1)^T, \quad v_2 = (1, 1, 1)^T, \quad \text{e } v_3 = (1, -2e^{-\beta D}, 1)^T. \quad (6.24)$$

Desta forma, concluímos que a \mathcal{Q} -matriz $\mathcal{Q}(\beta)$ é conjugada a uma matriz diagonal D , ou seja, existe uma matriz invertível B , tal que, $\mathcal{Q}(\beta) = BDB^{-1}$, onde

$$D = \begin{bmatrix} -3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -Z_1(\beta) \end{bmatrix}; \quad (6.25)$$

$$B = \begin{bmatrix} -1 & 1 & 1 \\ 0 & 1 & -2e^{-\beta D} \\ 1 & 1 & 1 \end{bmatrix}; \quad (6.26)$$

e

$$B^{-1} = \begin{bmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \\ \frac{e^{-\beta D}}{Z_1(\beta)} & \frac{1}{Z_1(\beta)} & \frac{e^{-\beta D}}{2Z_1(\beta)} \\ \frac{1}{2Z_1(\beta)} & -\frac{1}{Z_1(\beta)} & \frac{1}{2Z_1(\beta)} \end{bmatrix}. \quad (6.27)$$

Logo, pelos mesmos argumentos feitos na subseção 6.2.2, temos que \mathcal{P}^t será dada por

$$\mathcal{P}^t = \begin{bmatrix} -1 & 1 & 1 \\ 0 & 1 & -2e^{-\beta D} \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} e^{-3t} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{-Z_1(\beta)t} \end{bmatrix} \begin{bmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \\ \frac{e^{-\beta D}}{Z_1(\beta)} & \frac{1}{Z_1(\beta)} & \frac{e^{-\beta D}}{2Z_1(\beta)} \\ \frac{1}{2Z_1(\beta)} & -\frac{1}{Z_1(\beta)} & \frac{1}{2Z_1(\beta)} \end{bmatrix},$$

ou seja, $\forall t \geq 0$ temos, finalmente que

$$\mathcal{P}^t = \begin{bmatrix} \frac{e^{-3t}}{2} + \frac{e^{-\beta D}}{Z_1(\beta)} + \frac{e^{-Z_1(\beta)t}}{2Z_1(\beta)} & \frac{1}{Z_1(\beta)} - \frac{e^{-Z_1(\beta)t}}{Z_1(\beta)} & -\frac{e^{-3t}}{2} + \frac{e^{-\beta D}}{Z_1(\beta)} + \frac{e^{-Z_1(\beta)t}}{2Z_1(\beta)} \\ \frac{e^{-\beta D}}{Z_1(\beta)} - \frac{e^{-\beta D - Z_1(\beta)t}}{Z_1(\beta)} & \frac{1}{Z_1(\beta)} + \frac{2e^{-\beta D - Z_1(\beta)t}}{Z_1(\beta)} & \frac{e^{-\beta D}}{Z_1(\beta)} - \frac{e^{-\beta D - Z_1(\beta)t}}{Z_1(\beta)} \\ -\frac{e^{-3t}}{2} + \frac{e^{-\beta D}}{Z_1(\beta)} + \frac{e^{-Z_1(\beta)t}}{2Z_1(\beta)} & \frac{1}{Z_1(\beta)} - \frac{e^{-Z_1(\beta)t}}{Z_1(\beta)} & \frac{e^{-3t}}{2} + \frac{e^{-\beta D}}{Z_1(\beta)} + \frac{e^{-Z_1(\beta)t}}{2Z_1(\beta)} \end{bmatrix}. \quad (6.28)$$

Veremos agora que o semi-grupo a tempo contínuo \mathcal{P}^t , para $t \geq 0$, dado por 6.28, satisfaz a condição do balanço detalhado para o Estado de Gibbs 6.19. Com efeito,

$$\begin{aligned} (\mu(\beta))_1(P^t)_{10} &= \frac{e^{-\beta D}}{Z_1(\beta)} \left(\frac{1}{Z_1(\beta)} - \frac{e^{-Z_1(\beta)t}}{Z_1(\beta)} \right) = \frac{1}{Z_1(\beta)} \left(\frac{e^{-\beta D}}{Z_1(\beta)} - \frac{e^{-\beta D - Z_1(\beta)t}}{Z_1(\beta)} \right) = \\ &= (\mu(\beta))_0(P^t)_{01}; \end{aligned}$$

$$(\mu(\beta))_1(P^t)_{1-1} = \frac{e^{-\beta D}}{Z_1(\beta)} \left(-\frac{e^{-3t}}{2} + \frac{e^{-\beta D}}{Z_1(\beta)} + \frac{e^{-Z_1(\beta)t}}{2Z_1(\beta)} \right) = (\mu(\beta))_{-1}(P^t)_{-11};$$

e

$$\begin{aligned} (\mu(\beta))_{-1}(P^t)_{-10} &= \frac{e^{-\beta D}}{Z_1(\beta)} \left(\frac{1}{Z_1(\beta)} - \frac{e^{-Z_1(\beta)t}}{Z_1(\beta)} \right) = \frac{1}{Z_1(\beta)} \left(\frac{e^{-\beta D}}{Z_1(\beta)} - \frac{e^{-\beta D - Z_1(\beta)t}}{Z_1(\beta)} \right) = \\ &= (\mu(\beta))_0(P^t)_{0-1}. \end{aligned}$$

Portanto, o processo estocástico de Markov a tempo contínuo $\{X_t\}_{t \geq 0}$, onde a variável aleatória X_t indica, no instante t , o estado quântico de spin de cada núcleo localizado, é uma Dinâmica de Glauber.

Assim como fizemos na subseção 6.2.2, vamos analisar alguns casos limite, por exemplo, quando $t \rightarrow \infty$ tem-se

$$\lim_{t \rightarrow \infty} \mathcal{P}^t = \begin{bmatrix} \frac{e^{-\beta D}}{Z_1(\beta)} & \frac{1}{Z_1(\beta)} & \frac{e^{-\beta D}}{Z_1(\beta)} \\ \frac{e^{-\beta D}}{Z_1(\beta)} & \frac{1}{Z_1(\beta)} & \frac{e^{-\beta D}}{Z_1(\beta)} \\ \frac{e^{-\beta D}}{Z_1(\beta)} & \frac{1}{Z_1(\beta)} & \frac{e^{-\beta D}}{Z_1(\beta)} \end{bmatrix} = \begin{bmatrix} (\mu(\beta))_1 & (\mu(\beta))_0 & (\mu(\beta))_{-1} \\ (\mu(\beta))_1 & (\mu(\beta))_0 & (\mu(\beta))_{-1} \\ (\mu(\beta))_1 & (\mu(\beta))_0 & (\mu(\beta))_{-1} \end{bmatrix}. \quad (6.29)$$

Assim, concluímos que $\mu(\beta) = ((\mu(\beta))_1, (\mu(\beta))_0, (\mu(\beta))_{-1})$ é o único vetor de probabilidade invariante para \mathcal{P}^t , e daí, seja qual for a distribuição inicial p , temos sempre $\lim_{t \rightarrow \infty} p\mathcal{P}^t = \mu(\beta)$ (conforme teorema 1.3.22).

Além do mais, note que $(\mu(\beta))_0 \geq (\mu(\beta))_1 = (\mu(\beta))_{-1}$, conclui-se daí que, após um amplo intervalo de tempo, a maior parte dos núcleos estão ocupando o estado quântico de spin 0 e, dos núcleos restantes, metade estão ocupando o estado quântico de spin 1 e a outra metade o estado quântico de spin -1 . Novamente, como justificado na subseção 6.2.2, este *número de ocupação* se deve ao fato que o estado quântico de spin 0 é *menos energético* do que os outros dois estados 1 e -1 (que são igualmente energéticos, ver equação 6.18).

Vemos por 6.29 que, no equilíbrio, dado que um núcleo está no estado quântico de spin 1 é mais provável que ele vá para o estado de spin 0, e é igualmente provável que ele permaneça no estado quântico de spin 1, ou que ele vá para o estado quântico de spin -1 . Do mesmo modo, dado que um núcleo está no estado quântico de spin 0 é mais provável que ele permaneça neste estado quântico, e é igualmente provável que ele vá para os estados quânticos de spin 1 e -1 . Finalmente, dado que um núcleo está no estado quântico de spin -1 é mais provável que ele vá para o estado de spin 0, e é igualmente provável que ele permaneça neste estado quântico de spin, ou vá para o estado de spin 1 (Veja figura 6.1).

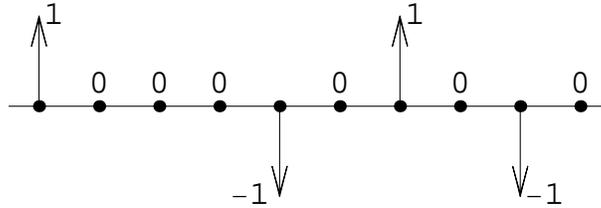


Figura 6.1 - Uma possível configuração para o sistema de N núcleos magnéticos localizados, onde estamos convencionando seta *para cima* quando $\sigma = 1$, *ausência* de seta quando $\sigma = 0$ e seta *para baixo* quando $\sigma = -1$.

Outra comprovação dos fatos mencionados no parágrafo anterior pode ser dada calculando-se o *Valor Esperado de Spin*, i.e., $\langle \sigma \rangle_{\mu(\beta)}$, pela definição 2.3.1 do capítulo 2 temos

$$\langle \sigma \rangle_{\mu(\beta)} = \sum_{\sigma \in S} \sigma (\mu(\beta))_{\sigma} = \frac{e^{-\beta D}}{Z_1(\beta)} + 0 - \frac{e^{-\beta D}}{Z_1(\beta)} = 0; \quad (6.30)$$

ou seja, escolhendo-se um núcleo ao acaso, é mais provável encontrá-lo no estado quântico de spin 0, confirmando as previsões feitas nos parágrafos anteriores.

Prosseguindo a análise dos limites, temos a igualdade em $(\mu(\beta))_0 \geq (\mu(\beta))_1 = (\mu(\beta))_{-1}$ apenas para $T \rightarrow \infty$ (i.e., $\beta \rightarrow 0$) e, neste caso, não é difícil ver que $\lim_{\beta \rightarrow 0} (\mu(\beta))_0 = \lim_{\beta \rightarrow 0} \frac{1}{1+2e^{-\beta D}} = 1/3$, e que $\lim_{\beta \rightarrow 0} (\mu(\beta))_1 = \lim_{\beta \rightarrow 0} (\mu(\beta))_{-1} = \lim_{\beta \rightarrow 0} \frac{e^{-\beta D}}{1+2e^{-\beta D}} = 1/3$. Com isto, somos levados a concluir que, para altas temperaturas, os tres estados quânticos de spin estão igualmente ocupados. Por outro lado, quando $T \rightarrow 0$ (i.e., $\beta \rightarrow \infty$), $\lim_{\beta \rightarrow \infty} (\mu(\beta))_0 = \lim_{\beta \rightarrow \infty} \frac{1}{1+2e^{-\beta D}} = 1$ e, além disso, $\lim_{\beta \rightarrow \infty} (\mu(\beta))_1 = \lim_{\beta \rightarrow \infty} (\mu(\beta))_{-1} = \lim_{\beta \rightarrow \infty} \frac{e^{-\beta D}}{1+2e^{-\beta D}} = 0$, indicando que, para baixas temperaturas, todos os núcleos estão no estado quântico de spin 0, correspondente ao estado de menor energia.

Vamos agora nos perguntar sobre a veracidade do limite (note que, por 6.30 $\langle \sigma \rangle_{\mu(\beta)} = 0$)

$$\lim_{t \rightarrow \infty} \|\mathcal{P}^t \sigma - \langle \sigma \rangle_{\mu(\beta)}\|_{\mu(\beta)} = \lim_{t \rightarrow \infty} \|\mathcal{P}^t \sigma\|_{\mu(\beta)} = 0, \quad (6.31)$$

onde $\mathcal{P}^t \sigma$ é uma notação análoga a relação 6.15. Como $\sigma \in L^2(\mu)$, pelas mesmas observações da subseção 6.2.2 é válido o limite 6.31. O quão rápida é tal convergência? Da mesma forma que fizemos acima (subseção 6.2.2) devemos estimar para a taxa λ_β utilizando 5.17, i.e., $\lambda_\beta \geq Z_1(\beta) e^{-\beta E(\mathbf{P})} / W(\mathbf{P})$.

No que consiste, seja

$$\mathbf{P} = \{p(1,0) = (1,0), p(0,1) = (0,1), p(0,-1) = (0,-1), p(-1,0) = (-1,0), \\ p(1,-1) = (1,-1), p(-1,1) = (-1,1)\},$$

a (boa) seleção de caminhos acessíveis escolhida. Pela definição de $e(p)$ dada em 5.12 temos, levando-se em conta o Hamiltoniano 6.18, que

$$e(p(1,0)) = e(p(0,1)) = e(p(0,-1)) = e(p(-1,0)) = 0,$$

e

$$e(p(1,-1)) = e(p(-1,1)) = -D < 0.$$

Logo, $E(\mathbf{P}) = \max\{0, -D\} = 0$.

Como vimos na seção 5.3, para um caminho $p = (i_0, \dots, i_n)$

$$\omega(p) = \sum_{m=1}^n \frac{(\nu)_{i_0} (\nu)_{i_n}}{(\nu)_{i_{m-1}} (A)_{i_{m-1} i_m}} = 1,$$

onde 1 se deve ao fato que, novamente, em nosso caso particular $(\nu)_i = (A)_{ij} \equiv 1 \forall (i,j) \in S^2$. Donde, vem que (veja equação 5.16) $W(\mathbf{P}) = \max\{1, 1\} = 1$. Daí, retornando à estimativa, vemos que $\lambda_\beta \geq Z_1(\beta) e^{-\beta E(\mathbf{P})} / W(\mathbf{P}) = Z_1(\beta)$ i.e., $\lambda_\beta \geq Z_1(\beta)$.

Para qualquer outra escolha de \mathbf{P} vamos obter $W(\mathbf{P}) > 1$. Como já observamos a seleção de caminhos acessíveis \mathbf{P} deve ser tomada de modo a minimizar $W(\mathbf{P})$.

□

Referências Bibliográficas

- [1] Balakrishnan, A. V. Applied Functional Analysis. Springer-Verlag, 1976.
- [2] Berger, Noam, et. al. Glauber Dynamics on Trees and Hyperbolic Graphs. Preprint available at, <http://front.math.ucdavis.edu/math.PR/0308284>, 2006.
- [3] Billingsley, P. Probability and Measure. John Wiley & Sons, Inc., 1995.
- [4] Brémaud, Pierre. Markov Chains: Gibbs Fields, Mont Carlo Simulation and Queues. Springer-Verlag, 1999.
- [5] Brin, M. Stuck, G. Introduction to Dynamical Systems. Cambridge University Press, 2002.
- [6] Castro Jr, A. Armando de. Curso de Teoria da Medida. Projeto Euclides-Impa, 2004.
- [7] Doering, Claus Ivo. Lopes, Artur Oscar. Equações Diferenciais Ordinárias. Coleção Matemática Universitária-Impa, 2005.
- [8] Ellis, R. S. Entropy, Large Deviations, and Statistical Mechanics. Springer-Verlag, 1985.
- [9] Fernandez, P. Medida e Integração. Projeto Euclides, 1996.
- [10] Gantmacher, F.R. The Theory of Matrices, vol. 2. Chelsea Publishing Company, New York, N.Y., 1959.
- [11] Glauber, Roy. J. Time-Dependent Statistics of the Ising Model, Journal of Mathematical Physics, Vol. 7, 1963.
- [12] Gibbs, J. W. Elementary Principles in Statistical Mechanics. New Haven, Yale University Press, 1902.
- [13] Huang, K. Statistical Mechanics. New York, John Wiley and Sons, 1963.
- [14] Karlin, S., Taylor, H. A first course in stochastic processes. Academic Press, Inc. 1975.
- [15] Knipis, C., Landim, C. Scaling Limits of Interacting Particles Systems. Springer-Verlag, 1999.
- [16] Landau, L. D., Lifshitz, E. M. Physique Statistique. Moscou, edições MIR, 1967.
- [17] Lima, Elon Lages. Álgebra Linear. Coleção Matemática Universitária-Impa, Inc. 2001.

- [18] Norris, J. R. Markov Chains. Cambridge University Press, 1998.
- [19] Pollicott, M. Yuri, M. Dynamical Systems and Ergodic Theory. Cambridge University Press, 1998.
- [20] Reed, Michael, Simin, Barry Methods of Modern Mathematical Physics: Vol 1. Functional Analysis. Academic Press, 1980.
- [21] Reif, F. Fundamentals of Statistical and Thermal Physics. New York, McGraw-Hill Book Company, 1965.
- [22] Royden, H. L. Real Analysis. New York, Macmillan Company, 1963.
- [23] Salinas, Sílvia R. A. Introdução à Física Estatística. edusp, 1999.
- [24] Sinai, Ya. Probability Theory. Springer-Verlag, 1992.
- [25] Stroock, W. Daniel An Introduction to Markov Processes. Springer-Verlag, 2000.
- [26] Taylor, J. C. An Introduction to Measure and Probability. Springer-Verlag, 1997.
- [27] Walters, P. An Introduction to Ergodic Theory. Springer-Verlag, New York, 1982.