## Simulação perfeita via cadeias de Markov

Edvaldo Capobiango Coelho

## DISSERTAÇÃO APRESENTADA AO INSTITUTO DE MATEMÁTICA E ESTATÍSTICA DA UNIVERSIDADE DE SÃO PAULO PARA OBTENÇÃO DO GRAU DE MESTRE EM ESTATÍSTICA Área de Concentração: PROBABILIDADE

Orientadora: Prof<sup>a</sup>. Dra. Elisabeti Kira

O autor recebeu auxílio financeiro da CNPQ. Processo 138843/1999-6

São Paulo, agosto de 2001

## Simulação perfeita via cadeias de Markov

Este exemplar corresponde à redação final da dissertação corrigida e defendida por Edvaldo Capobiango Coelho e aprovada pela comissão julgadora.

São Paulo, 14 de agosto de 2001.

Banca examinadora:

- Prof<sup>a</sup>. Dra. Elisabeti Kira (Presidente)- IME-USP
- Prof. Dr. Pablo Augusto Ferrari IME-USP
- Prof. Dr. Hervé Jean François Guiol IMECC-UNICAMP

Dedico aos meus pais, José Geraldo e Marlene. A meu irmão Iuander e ao seu filho Gabriel. A minha esposa Niara e ao meu querido filho Divaldinho. E ao meu querido tio Henrique (in memorian). Por sempre acreditarem em mim e por tudo que fizeram para que este trabalho ocorresse com sucesso.

7

## Agradecimentos

A orientadora Elisabeti Kira pela dedicada orientação e pela paciência comigo.

Aos professores da banca Pablo e Hervé pela colaboração.

A Niara e ao Divaldinho por não se esquecerem de mim mesmo eu tendo que ficar vários dias longe deles.

Ao grande amigo André pela estadia em sua casa e pelos conselhos que foram muito importantes para mim.

Aos eternos amigos Fábio e Josué pela paciência que sempre tiveram comigo nos estudos e por não ficarem com raiva pelas inúmeras brincadeiras que fiz com eles, mas no fundo sei que não ficavam com raiva é porque eu dava brilhantes passes para eles fazerem gols nas nossas peladas.

Ao Marcelo, Flávio, Mauro por sempre servirem de sparrings no penbolin, mesmo que fracos mas eles sempre faziam o melhor, e por sempre assumirem que sou o melhor jogador de penbolin.

A garota Sandy (Gisela), por sempre pagar as cervejas que apostavamos mesmo sabendo que eu não iria pagar se tivesse perdido.

Se fosse citar todos aqueles que participaram de alguma forma deste trabalho estes agradecimentos ficariam maior do que esta dissertação. Assim meus agradecimentos a todos que de alguma forma foram importantes e em particular ao CNPQ pelo auxílio financeiro.

## Resumo

Neste trabalho apresentamos dois algoritmos de simulação perfeita usando cadeias de Markov e as respectivas provas de que fornecem amostras de acordo com a distribuição desejada. Estes algoritmos são o *Coupling from the past*-CFTP introduzido por Propp e Wilson e o algoritmo de Fill. Para ilustração destes algoritmos aplicamo-os ao modelo de Ising ferromagnético em duas dimensões, usando o Amostrador de Gibbs (*Gibbs Sampler*).

## Abstract

In this work we present two perfect simulation algorithms using Markov chains and the respective demonstrations that they provide samples in accordance with the desired distribution. These algorithms are Coupling from the past-CFTP introduced by Propp and Wilson and the algorithm of Fill. For illustration of these algorithms we have applied them to the ferromagnetic Ising model in two dimensions by using Gibbs Sampler.

# Índice

1	Inti	rodução	1			
<b>2</b>	2 Acoplamento do Passado - CFTP					
	2.1	Acoplamento	4			
	2.2	O algoritmo CFTP	7			
	2.3	Algoritmo CFTP-monótono	12			
	2.4	Extensões do CFTP	19			
3 Algoritmo de Fill						
	3.1	O Algoritmo	20			
	3.2	Extensões do algoritmo de Fill	23			
4	4 Modelo de Ising					
	4.1	Medida de Gibbs	25			
	4.2	Modelo de Ising	27			
	4.3	Amostrador de Gibbs	30			
		4.3.1 Algoritmo Metropolis-Hastings	30			
		4.3.2 Amostrador de Gibbs	31			
5	Sim	ulações	36			

	5.1	Algoritmo CFTP-monótono aplicado ao modelo de Ising					
	5.2 Algoritmo de Fill aplicado ao modelo de Ising						
	5.3	Resultados	45				
Α	Apêndice						
	A.1	Prova da Proposição 2.7	47				
	A.2	Prova da equação (3.3)	51				
	A.3	Prova do resultado (4.19)	52				
Re	Referências Bibliográficas						

## Capítulo 1

## Introdução

Existem vários modelos, tanto na teoria estatística quanto na teoria de probabilidades, que nos levam a uma distribuição de probabilidades sob um espaço de estados E de dimensão imensa ou até infinita. Tais modelos tem origem em física-estatística, inferência Bayesiana, geometria estocástica, etc. Para inferir sob uma determinada população, considerando que a população tenha a distribuição de um destes modelos, necessitamos obter uma ou mais amostras desta população. No entanto não temos uma forma simples ou direta para obter esta amostra, devido a alta dimensão de E.

Nesta última década os métodos de Monte Carlo via Cadeias de Markov (*Markov Chain Monte Carlo* - MCMC) foram muito utilizados, e ainda são, em processamento de imagens, sistemas de dinâmica molecular, modelos hierárquicos etc., como pode ser visto por exemplo em [Gilks et al., 1996] e [Gamerman, 1997]. Estes métodos são baseados na simulação de cadeias de Markov que convergem para a distribuição desejada. Entre os métodos MCMC mais importantes está o algoritmo de Metropolis-Hastings, que foi introduzido por Metropolis em 1953 e aperfeiçoado por Hastings em [Hastings, 1970], sendo que um caso especial deste algoritmo é o Amostrador de Gibbs (*Gibbs Sampler*) introduzido por Geman e Geman em [Geman and Geman, 1984]. Um problema destes métodos é que eles garantem que a cadeia convergiu para a distribuição desejada somente após um número infinito de transições. Como isto é impossível, conseguimos apenas encontrar um número de transições suficientemente grande, tal que a distribuição da cadeia difere da distribuição desejada por no máximo uma quantidade pequena especificada. Logo, não sabemos se esta cadeia já atingiu a distribuição desejada ou quantas transições deveriam ser feitas para que esta cadeia atingisse esta distribuição desejada. Assim, no instante em que interrompemos ou finalizamos uma simulação do MCMC não sabemos se a amostra obtida tem a distribuição desejada.

Apresentamos dois algoritmos de obtenção de uma amostra que esteja de acordo com a distribuição desejada que são Acoplamento do Passado (*Coupling from the past*-CFTP) e o algoritmo de Fill. Chamamos a amostra, obtida por estes métodos, de amostra perfeita ou exata.

O algoritmo CFTP foi proposto por Propp e Wilson em [Propp and Wilson, 1996]. Este algoritmo também é baseado na simulação de cadeias de Markov e irá nos fornecer uma amostra perfeita da distribuição desejada definida em um espaço de estados *E* grande mas finito. Uma desvantagem deste algoritmo é a existência de correlação entre o estado amostrado e o número de transições da cadeia de Markov. Pela existência desta correlação este algoritmo só irá liberar a amostra perfeita se o número de transições previamente determinado for executado. Assim o usuário impaciente que interromper uma simulação, antes do algoritmo ter percorrido todas as transições determinadas, poderá ter um vício na amostra obtida. O algoritmo CFTP e a prova de que ele produz uma amostra com distribuição exata serão apresentados no Capítulo 2. Além disso, veremos que a técnica de acoplamento em cadeias de Markov é fundamental para construção deste algoritmo.

Fill em [Fill, 1998] propôs um algoritmo alternativo que fornece uma amostra perfeita da distribuição desejada. Este algoritmo também trabalha com cadeias de Markov, mas baseia-se no método de rejeição da amostra. Diferentemente do CFTP, neste algoritmo não há correlação entre o estado amostrado e o número de transições da cadeia, eliminando assim o vício que, eventualmente, ocorra quando se interrompe uma simulação. O algoritmo interrompível de Fill e a prova de que este algoritmo fornece uma amostra com distribuição exata são apresentados no Capítulo 3.

Ilustramos estes dois algoritmos aplicando-os ao modelo de Ising. Assim, no

Capítulo 4 definimos campo aleatório, campo aleatório Markoviano e medida de Gibbs. Descrevemos neste capítulo o modelo de Ising ferromagnético em duas dimensões com suas principais características. Apresentamos também o amostrador de Gibbs e mostramos como aplicá-lo ao modelo de Ising.

No Capítulo 5 fazemos a aplicação do algoritmo CFTP e do algoritmo de Fill para o modelo de Ising ferromagnético em duas dimensões usando o amostrador Gibbs. E, apresentamos amostras da distribuição deste modelo sob determinadas condições.

## Capítulo 2

## Acoplamento do Passado - CFTP

Queremos obter uma amostra de uma certa distribuição  $\pi$ , com  $\pi$  definida em um espaço de estados E (cuja dimensão é muito grande mas finita). No entanto uma amostra direta desta distribuição complexa, devido a alta dimensão do espaço de estados E, não é possível. Entre as distribuições complexas, em particular, temos por exemplo o modelo de Ising em física estatística, processos pontuais Markovianos em geometria estocástica, entre outros.

Veremos neste capítulo que o algoritmo proposto por Propp e Wilson em [Propp and Wilson, 1996], conhecido como Acoplamento do Passado (*Coupling from the past*-CFTP), fornece uma amostra perfeita (exata) da distribuição complexa desejada. Este algoritmo também é baseado na construção de cadeias de Markov ergódicas e ao contrário dos métodos MCMC irá produzir uma amostra somente após a cadeia de Markov ter alcançado o equilíbrio. Para a construção deste algoritmo a técnica de acoplamento em cadeias de Markov será fundamental.

### 2.1 Acoplamento

A técnica de acoplamento, em linhas gerais, consiste em construir duas ou mais medidas de probabilidade, variáveis aleatórias, processos estocásticos, etc, sobre um mesmo espaço de probabilidade. Para aplicação da técnica de acoplamento em geral ver, por exemplo, [Lindvall, 1992] e [Thorisson, 2000]. A seguir, introduzimos a técnica de acoplamento em cadeias de Markov, como pode ser visto, por exemplo, em [Galves and Ferrari, 1997], pois esta técnica é fundamental para a construção do algoritmo CFTP apresentado na próxima seção.

Consideramos uma cadeia de Markov X estacionária, com espaço de estados finito E (com |E| = N, onde |E| denota a cardinalidade de E) e matriz de probabilidades de transição P. A evolução de uma cadeia de Markov  $X = \{X_t\}_{t \in \mathbb{Z}}$ , onde t representa o tempo e  $\mathbb{Z}$  o conjunto dos números inteiros, pode ser feita usando uma regra de transição f obtida a partir de P e uma seqüência  $\{U_t\}_{t \in \mathbb{Z}}$  de variáveis aleatórias independentes uniformes em [0,1], sendo que esta regra de transição f determina  $X_t$ como uma função de  $X_{t-1}$  e  $U_t$ , ou seja,

$$X_t = f(X_{t-1}, U_t). (2.1)$$

Esta função  $f : E \times [0,1] \to E$ , que chamamos de regra de transição, pode ser construída de várias formas (ver por exemplo [Galves and Ferrari, 1997] e o exemplo abaixo), no entanto tem que satisfazer

$$\mathbb{P}(f(x,U) = y) = P(x,y), \ \forall x, y \in E,$$
(2.2)

onde  $P(x, y) = \mathbb{P}(X_t = y | X_{t-1} = x).$ 

Chamamos de dispositivo aleatório a regra de transição e a variável aleatória uniforme usadas para evoluir uma cadeia de Markov X.

Exemplificamos a seguir como é possível obter várias funções de transição satisfazendo a equação (2.2).

Exemplo 2.1 Considere uma cadeia de Markov X com espaço de estados  $E = \{0, 1, 2\}$ e com a seguinte matriz de transição

$$P = \left(\begin{array}{rrrr} 0,5 & 0,5 & 0 \\ 0,5 & 0 & 0,5 \\ 0 & 0,5 & 0,5 \end{array}\right).$$

De acordo com esta matriz de transição P podemos definir as seguintes regras de transição que satisfazem a equação (2.2)

$$f(X_{t-1}, U) = \begin{cases} X_t = \max\{X_{t-1} - 1, 0\} & \text{se } U \le 0,5; \\ X_t = \min\{X_{t-1} + 1, 2\} & \text{se } U > 0,5; \end{cases}$$
(2.3)

e

$$f'(X_{t-1}, U) = \begin{cases} X_t = 0 & \text{se } X_{t-1} = 0 \text{ ou } 1 \text{ e } U \le 0, 5; \\ X_t = 1 & \text{se } X_{t-1} = 0 \text{ ou } 2 \text{ e } U > 0, 5; \\ X_t = 2 & \text{se } X_{t-1} = 1 \text{ e } U > 0, 5 \\ & \text{ou se } X_{t-1} = 2 \text{ e } U \le 0, 5. \end{cases}$$
(2.4)

Note que existem outras regras de transição obtidas a partir de P que também satisfazem a equação (2.2).

**Definição 2.2** Dizemos que duas ou mais cadeias de Markov estão acopladas se elas evoluem de acordo com o mesmo dispositivo aleatório, ou seja, dentro de um mesmo espaço de probabilidade.

Assim, para uma cadeia de Markov, com espaço de estados finito, ergódica (irredutível e aperiódica), duas ou mais trajetórias desta cadeia serão ditas acopladas se evoluem de acordo com a mesma regra de transição f (obtida a partir de P) e a mesma seqüência de números aleatórios  $\cdots, U_t, U_{t+1}, \cdots$ , ou seja, evoluem de acordo com um mesmo dispositivo aleatório. Para maiores detalhes sobre uma cadeia de Markov ergódica veja, por exemplo, [Ross, 1994].

Dizemos que as trajetórias de duas cadeias, partindo de dois estados quaisquer de E, coalesceram se suas trajetórias se juntaram após um tempo aleatório T e permanecem juntas após este tempo. Devemos observar que se iniciarmos N trajetórias a partir de cada um dos N estados de E, usando o acoplamento em cadeias de Markov, certamente (pela ergodicidade) obteremos a coalescência para um tempo suficientemente grande. Assim, a partir do instante em que se tem a coalescência, a evolução conjunta da trajetória restante não dependerá mais do estado inicial da cadeia.

Neste trabalho consideramos cadeias acopladas como trajetórias acopladas de uma cadeia.

### 2.2 O algoritmo CFTP

Apresentamos, nesta seção, o algoritmo CFTP, introduzido por Propp e Wilson em [Propp and Wilson, 1996], e a respectiva prova de que a amostra fornecida tem a distribuição exata  $\pi$ . Este algoritmo é baseado na construção de cadeias de Markov.

Para construirmos uma trajetória (evolução) de uma cadeia em geral podemos usar uma composição das regras de transição de X,  $f_t$ 's, denotada por

$$F_{0,t}(x) = f_t \circ \dots \circ f_1(x), \quad t \in \mathbb{N}, \ x \in E;$$

$$(2.5)$$

onde  $F_{0,t}(x)$  é o estado da cadeia no instante t de uma trajetória de X iniciada no estado x no instante 0 e as funções  $f_t$ 's são definidas por

$$f_t(\cdot) = f(\cdot, U_t), \quad t \in \mathbb{N}.$$
(2.6)

Estas regras de transição são independentes pois as  $U_t$ 's são variáveis aleatórias independentes uniformes no intervalo [0,1]. Esta composição de funções, como na equação (2.5), descreve o mapeamento de uma trajetória da cadeia até o instante t iniciada no estado x.

Assim, evoluindo as cadeias (trajetórias) acopladas, com cada trajetória partindo de um dos estados de E no instante 0 e fazendo um número infinito de transições, certamente teremos um único estado amostrado, visto que a cadeia é ergódica. Suponha que a coalescência tenha ocorrido no instante S, então a partir daí teremos a evolução conjunta das trajetórias da cadeia como se fosse apenas uma. O fato é que no instante em que ocorre a coalescência, não podemos dizer que a cadeia está em equilíbrio e não se sabe quanto tempo esta cadeia teria que percorrer para atingir o equilíbrio. Pois, de acordo com alguns exemplos, como o Exemplo 2.1 dos Capítulos 2 e 4 de [Thönnes, 1999b] e o Exemplo 2.6 neste trabalho, vemos que no instante emque ocorre a coalescência a cadeia não está em equilíbrio.

Propp e Wilson em [Propp and Wilson, 1996] reverteram a ordem da regra de transição composta acima (2.5) iniciando todas as cadeias no instante -k, evoluindo estas cadeias acopladas até o instante 0, até termos um único estado amostrado no

instante 0, ou seja, até ocorrer a coalescência antes de 0. Podemos imaginar que a cadeia está vindo de um tempo  $-\infty$ , assim no instante 0 o estado amostrado terá a distribuição estacionária  $\pi$ . Revertendo-se a ordem de (2.5) temos que a composição de funções  $f_k$  será dada por

$$F_{-k,0}(x) = f_0 \circ \cdots \circ f_{-k+1}(x); \quad k \in \mathbb{N}, \ x \in E.$$
 (2.7)

Pela composição acima verifica-se que  $F_{-k,0}(x)$  é o estado da cadeia no instante 0 de uma trajetória de X iniciada no instante -k no estado x.

A cada iteração  $i, i \in \mathbb{N} - \{0\}$ , do CFTP escolhemos  $k = k_i > k_{i-1}$  e uma vez ocorrido a coalescência em uma determinada iteração i, finalizamos a simulação no instante 0 desta iteração. A idéia proposta por eles é que a cada iteração i do CFTP, isto é, para cada valor de -k estaremos fixando a nossa regra de transição  $f_{-k}$ , ou seja, quando -k = -1 estará fixado  $f_0$  (pois escolhemos aleatoriamente e fixamos o valor de  $U_0$ ). Para -k = -2 fixaremos  $f_{-1}$  (escolhemos e fixamos  $U_{-1}$ ) e compomos com o valor de  $f_0$  fixado anteriormente, e assim por diante.

Temos então que evoluindo as trajetórias acopladas da cadeia (ergódica) partindo de todos os estados de E no instante -k, para  $-k \to -\infty$ , a coalescência das trajetórias ocorrerá antes do instante 0. Assim a imagem de  $\lim_{k\to\infty} F_{-k,0}(x), \forall x \in E$ , será única e será uma amostra exata da distribuição  $\pi$ .

Apresentamos a seguir o algoritmo CFTP.

#### PASSOS DO ALGORITMO CFTP

**Passo 1)** Iniciar com  $i := 1 e k_0 = 0$ ;

**Passo 2)** Assumir  $k = k_i > 0$ . Iniciar as trajetórias da cadeia no instante  $-k_i$ , a partir de todos os estados x de E, e evoluir estas trajetórias acopladas até o instante  $-k_{i-1}$  usando a regra de transição composta

$$F_{-k_{i},-k_{i-1}}(x) = f_{-k_{i-1}} \circ \cdots \circ f_{-k_{i}+1}(x), \quad \forall x \in E.$$

**Passo 3)** Evoluir o restante das trajetórias da cadeia do instante  $-k_{i-1}$ , a partir dos

estados  $F_{-k_i,-k_{i-1}}(x)$ ,  $\forall x \in E$ , até o instante 0 usando a regra de transição composta

$$F_{-k_{i-1},0}(F_{-k_{i},-k_{i-1}}(x)) = f_0 \circ \cdots \circ f_{-k_{i-1}+1}(F_{-k_{i},-k_{i-1}}(x))$$

e reusando a seqüência de variáveis aleatórias uniformes geradas na iteração anterior. Passo 4) Verificar se as trajetórias coalesceram até o instante 0. Caso isto tenha ocorrido o estado amostrado no instante 0 será a amostra da distribuição desejada. Passo 5) Senão reiniciamos o procedimento para uma nova iteração, ou seja, para um novo valor de k, fazendo i := i + 1 e  $k_i < k_{i+1}$  e retornando ao Passo 2.

Abaixo apresentamos o resultado e a respectiva prova de que o algoritmo CFTP produz uma amostra com distribuição exata.

**Teorema 2.3** Seja X uma cadeia de Markov ergódica com espaço de estados E finito, matriz de transição P e distribuição estacionária  $\pi$ . O CFTP (Coupling from the Past) produz uma amostra perfeita de acordo com a distribuição  $\pi$ .

Uma prova heurística de que o algoritmo CFTP fornece uma amostra exata segue abaixo.

Na trajetória em que a cadeia está indo do presente para o futuro, como descrita na composição (2.5), temos que quando  $t \to \infty$  (t suficientemente grande) a cadeia atingirá a distribuição estacionária devido a ergodicidade da cadeia. O problema é que temos garantia de que esta cadeia terá convergido para a distribuição desejada somente após um número infinito de transições. Como na prática é impossível percorrer uma cadeia por um número infinito de transições Propp e Wilson propuseram o CFTP descrito.

Se a cadeia é iniciada no passado e evolui até o instante 0, como na composição (2.7), podemos supor que as trajetórias desta cadeia iniciam de todos os estados no instate  $-k = -\infty$ . No entanto como as  $f_{-k}$ 's são independentes, como em (2.6), teremos no instante 0 um único estado amostrado que estará de acordo com a distribuição estacionária, devido a ergodicidade e por estarmos percorrendo a cadeia por um período infinito. Seja t um valor fixo,  $t \in \mathbb{N}$ , tal que  $F_{-t,0}$  tem um único estado amostrado. Este estado tem a distribuição estacionária pois de acordo com a composição de funções abaixo

$$F_{-t,0}(x) = f_0 \circ \cdots \circ f_{-t+1}(x), \ \forall x \in E,$$

$$(2.8)$$

podemos imaginar que a cadeia inicia (em todos os estados) em um instante  $-\infty$  e evolui até o instante -t, ou seja,

$$F_{-\infty,-t}(\cdot) = f_{-t} \circ \cdots \circ f_{-\infty}(\cdot).$$

Pelo fato da cadeia ser ergódica e percorrer um tempo infinito de  $-\infty$  até -t, conseqüentemente esta cadeia estará em equilíbrio no instante -t e em um único estado  $x \in E$ . Para qualquer que seja este estado x, o valor de  $F_{-t,0}(x)$  será o mesmo pois como dito anteriormente  $F_{-t,0}$  tem um único estado amostrado para qualquer estado inicial. Assim deduzimos o estado amostrado de uma simulação de um tempo infinito no instante 0, em tempo finito. Como  $t \in \mathbb{N}$  nos dá um único estado amostrado existirá um tempo de coalescência,  $T \in \mathbb{N}$ , tal que T < t e conseqüentemente  $\mathbb{P}(T < \infty) = 1$  (como veremos na primeira etapa da prova abaixo).

#### Prova do Teorema 2.3:

A prova abaixo segue as linhas de [Thönnes, 1999b] e a prova original pode ser vista em [Propp and Wilson, 1996].

A prova deste teorema consiste de duas etapas:

(I) Mostrar que a imagem de  $F_{-k,0}(\cdot)$  (como na equação (2.7)) quase certamente vai ser única quando  $k \longrightarrow \infty$ , ou seja, partindo de todos os estados no instante -k e evoluindo as trajetórias paralelas até o instante 0, teremos que todas elas estarão num mesmo estado, ou seja, irão coalescer até o instante 0.

(II) Mostrar que esta imagem única de  $\lim_{k\to\infty} F_{-k,0}(\cdot)$  tem distribuição estacionária  $\pi$ .

#### Prova de (I)

Seja  $y \in E$  um estado arbitrário do espaço de estados de X. Como X é ergódica, existe um N(y) > 0 finito tal que, para qualquer  $x \in E$ , teremos

$$P^{N(y)}(x,y) = \mathbb{P}(X_{N(y)} = y | X_0 = x) > 0$$

 $(P^N$  denota a matriz de probabilidades de transição de X em N passos). Então seja $N=\max_{y\in E}\{N(y)\}$ tal que

$$\epsilon = \min\{P^N(x, y) : \text{ para todo } x, y \in E\} > 0.$$

Para todo  $k \in \mathbb{N}^*$ , definimos os eventos

$$C_k = \{\text{imagem de } F_{-kN,-(k-1)N} \notin \text{única}\}$$

assim

$$\mathbb{P}\{C_k\} = \mathbb{P}\{\text{todas as trajetórias coalescerem em } (-kN, -(k-1)N)\} > \epsilon.$$

Note que  $\{C_k\}_{k\in\mathbb{N}}$  são independentes pois cada  $C_k$  é função de  $F_{-kN, -(k-1)N}$  que são independentes (pois  $U_j$ , para  $j \in \mathbb{Z}$ , são independentes e identicamente distribuídas) para diferentes valores de k.

Pelo lema de Borel-Cantelli, como

$$\sum_{k=1}^{\infty} \mathbb{P}(\{C_k\}) = \infty$$

e  $\{C_k\}_{k\in\mathbb{N}}$  são independentes, então

$$\mathbb{P}(\{C_k\} \ i.v.) = \mathbb{P}(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} C_k) = 1,$$

onde i.v. = infinitas vezes. Seja

 $S = \min\{k \in \mathbb{N} : \text{imagem de } F_{-kN,-(k-1)N} \notin \text{única}\} = \min\{k \in \mathbb{N} : \{C_k\} \text{ ocorre}\}$ 

Provaremos que S é finito quase certamente pois, temos que

$$\{S < \infty\} = C_1 \cup C_2 \cup C_3 \cup \dots = \bigcup_{k=1}^{\infty} C_k$$

e então

$$(\{C_k\} \ i.v.) = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} C_k \subseteq \bigcup_{k=1}^{\infty} C_k = \{S < \infty\};$$

conseqüentemente

$$\mathbb{P}\{ S < \infty \} \ge \mathbb{P}(\{C_k\} i.v.) = 1,$$

ou seja,

$$\mathbb{P}\{S < \infty\} = 1.$$

Logo, a coalescência ocorre em um tempo aleatório finito quase certamente (isto é, com probabilidade 1), então:

$$F_{-SN,0} = F_{-(S-1)N,0} \circ F_{-SN,-(S-1)N}, \qquad (2.9)$$

terá imagem única em tempo quase certamente finito.

#### Prova de (II)

Vamos mostrar que a imagem única de  $F_{-SN,0}$  tem a distribuição  $\pi$ .

Temos que o mapa composto para frente é

$$F_{0,k}(x) = f_k \circ \cdots \circ f_1(x), x \in E$$

e  $F_{0,k}(x)$ , para  $k \in \mathbb{N} - \{0\}$ , é o caminho de X iniciado em x no instante 0. Devido a ergodicidade de X temos que a distribuição de  $F_{0,k}(x)$  converge para  $\pi$  quando  $k \longrightarrow \infty$ .

Como  $F_{-k,0}$  é composta de k regras de transição independentes então  $F_{-k,0}(x)$ tem a mesma distribuição de  $F_{0,k}(x)$ , quando k aproximar-se do infinito. Temos que

$$\lim_{k \to \infty} F_{-k,0}(\cdot) = \{ \text{imagem única de } F_{-SN,0} \}$$

pois  $F_{-SN,0}$  terá imagem única em tempo quase certamente finito (observe a equação (2.9)). Assim a imagem única de  $F_{-SN,0}$  terá também a distribuição estacionária  $\pi$  visto que  $\lim_{k\to\infty} F_{-k,0}(\cdot)$  tem distribuição  $\pi$ .

### 2.3 Algoritmo CFTP-monótono

O problema em aplicarmos o algoritmo CFTP apresentado na seção anterior é que, em geral, o espaço de estados E é muito grande, tornando impraticável evoluir as trajetórias da cadeia partindo de todos os possíveis estados de E. Para evitar este problema considere a cadeia com as propriedades de monotonicidade abaixo. Suponha que o espaço de estados E de uma cadeia de Markov admita uma ordem parcial natural, denotada por  $\leq$ .

Definição 2.4 a) Seja P uma matriz de transição de uma cadeia de Markov em um espaço de estados parcialmente ordenado  $(E, \preceq)$ . Dizemos que P, ou a cadeia, é monótona se P preserva a ordem parcial, isto é, se

$$P(x,\cdot) \leq P(y,\cdot)$$

estocasticamente se  $x \leq y$ .

b) Dizemos que uma regra de transição f associada a P é monótona se para  $x \leq y$ , temos que  $f(x,U) \leq f(y,U)$  quase certamente com respeito a U.

Suponha também que E tenha estados máximo  $(\hat{1})$  e mínimo  $(\hat{0})$  tal que  $\hat{0} \leq x \leq \hat{1}$ , para todo  $x \in E$ . Então, devido a monotonicidade, será necessário verificar a coalescência apenas entre as trajetórias partindo dos estados  $\hat{1}$  e  $\hat{0}$  para verificarmos a coalescência completa.

Pelo exemplo abaixo, observamos que para uma mesma matriz de transição P monótona encontramos uma regra de transição que é monótona com respeito a uma ordem parcial e outra que não é monótona.

Exemplo 2.5 Considere a cadeia X com espaço de estados  $E = \{0, 1, 2\}$  do Exemplo 2.1 dotado com a ordem parcial  $\hat{0} = 0 \leq 1 \leq \hat{1} = 2$  e a regra de transição  $f(\cdot, U)$ (definida na equação (2.3)). Note que f é monótona, ou seja, se  $x \leq y, x, y \in E$ então  $f(x, U) \leq f(y, U)$ . Já a regra de transição  $f'(\cdot, U)$  (definida em (2.4)) não é monótona com respeito a ordem parcial  $\hat{0} = 0 \leq 1 \leq \hat{1} = 2$ , pois quando  $1 \leq 2$  e U > 0.5 teremos que  $f'(1, U) = 2 \leq f'(2, U) = 1$ , ou seja, f' não mantém a ordem parcial.

A seguir apresentamos o algoritmo CFTP para o caso em que há monotonicidade da regra de transição. Considere uma cadeia de Markov definida em um espaço de estados E parcialmente ordenado possuindo estados máximo e mínimo, com matriz de transição monótona com respeito a esta ordem parcial e com regra de transição monótona. Assim no algoritmo CFTP-monótono abaixo será necessário verificar a coalescência das trajetórias partindo apenas dos estados máximo e mínimo.

Propp e Wilson argumentaram em [Propp and Wilson, 1996] que, para o algoritmo CFTP-monótono apresentado a seguir, uma escolha próxima da ótima para -ké, a cada iteração  $i \in \mathbb{N} - \{0\}$ ,

$$-k = -k_i = -2^{i-1}$$
.

Mais detalhes com respeito a justificativa da escolha de -k pode ser vista na seção 5.2 de [Propp and Wilson, 1996].

#### PASSOS DO ALGORITMO CFTP-MONÓTONO

**Passo 1**) Iniciar com  $i := 1 e k_0 = 0$ ;

**Passo 2)** Assumir  $k = k_i = 2^{i-1}$ . Iniciar duas trajetórias da cadeia no instante  $-k_i$ , sendo uma partindo do estado máximo  $\hat{1}$  e a outra do estado mínimo  $\hat{0}$ , e evoluir estas trajetórias acopladas até o instante  $-k_{i-1}$  usando a regra de transição composta

$$F_{-k_{i},-k_{i-1}}(\hat{0}) = f_{-k_{i-1}} \circ \cdots \circ f_{-k_{i+1}}(\hat{0}) \quad e$$
$$F_{-k_{i},-k_{i-1}}(\hat{1}) = f_{-k_{i-1}} \circ \cdots \circ f_{-k_{i+1}}(\hat{1}).$$

**Passo 3)** Evoluir o restante das trajetórias da cadeia do instante  $-k_{i-1}$ , a partir dos estados  $F_{-k_i,-k_{i-1}}(\hat{0}) \in F_{-k_i,-k_{i-1}}(\hat{1})$ , acopladas até o instante 0 usando a regra de transição composta

$$F_{-k_{i-1},0}(F_{-k_{i},-k_{i-1}}(\hat{0})) = f_0 \circ \cdots \circ f_{-k_{i-1}+1}(F_{-k_{i},-k_{i-1}}(\hat{0})) \quad e$$
  
$$F_{-k_{i-1},0}(F_{-k_{i},-k_{i-1}}(\hat{1})) = f_0 \circ \cdots \circ f_{-k_{i-1}+1}(F_{-k_{i},-k_{i-1}}(\hat{1}))$$

e reusando a mesma seqüência de variáveis aleatórias uniformes geradas na iteração anterior.

Passo 4) Verificar se as duas trajetórias coalesceram até o instante 0. Caso isto tenha ocorrido o estado amostrado no instante 0 será a amostra da distribuição  $\pi$ .

Passo 5) Senão reiniciamos o procedimento para uma nova iteração, ou seja, para um novo valor de k, fazendo i := i + 1 e retornando ao Passo 2.

No algoritmo CFTP o tempo da trajetória do CFTP, em termos do número de iterações (transições), não é independente do estado amostrado, ou seja, existe uma correlação entre o número de iterações (transições) e o estado amostrado pois a cada nova iteração do CFTP estamos reusando a mesma seqüência de números aleatórios geradas nas iterações anteriores. Entretanto para a liberação do resultado é necessário percorrer um número de iterações fixado. Assim o usuário impaciente que interromper uma simulação, antes do número de iterações previamente fixado ter sido alcançado, poderá provocar um vício na amostra favorecendo assim as amostras (estados) que demorarem menos tempo para serem geradas.

A seguir consideramos novamente a cadeia de Markov do Exemplo 2.1, no qual podemos observar que no instante em que ocorre a coalescência, esta cadeia não terá a distribuição desejada  $\pi$  e que para qualquer número de transições par esta cadeia não terá distribuição  $\pi$ .

**Exemplo 2.6** Considerando a cadeia de Markov do Exemplo 2.1 obtemos com um cálculo simples que a distribuição estacionária da cadeia é  $\pi(0) = \frac{1}{3}$ ,  $\pi(1) = \frac{1}{3}$ ,  $\pi(2) = \frac{1}{3}$ . Consideramos a mesma ordem parcial do Exemplo 2.5, ou seja,  $\hat{0} = 0 \leq 1 \leq \hat{1} = 2$ , e a regra de transição monótona f dada por (2.3).

Descrevemos abaixo as trajetórias paralelas da cadeia partindo dos estados máximo e mínimo para n = 2 (duas) transições, obtendo assim qualquer um dos seguintes resultados.

Caso 1) Para  $U_{-1} \leq 0.5$ ;  $U_0 \leq 0.5$   $x_{-2} = 0 \longrightarrow x_{-1} = 0 \longrightarrow x_0 = 0$   $x_{-2} = 2 \longrightarrow x_{-1} = 1 \longrightarrow x_0 = 0$ Caso 2) Para  $U_{-1} \leq 0.5$ ;  $U_0 > 0.5$   $x_{-2} = 0 \longrightarrow x_{-1} = 0 \longrightarrow x_0 = 1$   $x_{-2} = 2 \longrightarrow x_{-1} = 1 \longrightarrow x_0 = 2$ Caso 3) Para  $U_{-1} > 0.5$ ;  $U_0 \leq 0.5$   $x_{-2} = 0 \longrightarrow x_{-1} = 1 \longrightarrow x_0 = 0$  $x_{-2} = 2 \longrightarrow x_{-1} = 2 \longrightarrow x_0 = 1$  Caso 4) Para  $U_{-1} > 0.5; U_0 > 0.5$  $x_{-2} = 0 \longrightarrow x_{-1} = 1 \longrightarrow x_0 = 2$  $x_{-2} = 2 \longrightarrow x_{-1} = 2 \longrightarrow x_0 = 2$ 

De acordo com estas trajetórias, podemos observar que não é possível ocorrer a coalescência das trajetórias em uma transição da cadeia e que no primeiro instante em que ocorre a coalescência (na segunda transição) a cadeia não tem a distribuição  $\pi$ .

Observamos também que o algoritmo CFTP só irá liberar o resultado se estiver diante do primeiro ou do quarto caso. Então temos que a distribuição da cadeia obtida no instante zero, condicionado que houve coalescência, será  $\pi'(0) = \frac{1}{2}$ ,  $\pi'(1) = 0$ ,  $\pi'(2) = \frac{1}{2}$ , que é diferente da distribuição desejada.

Denote por c (cara) se  $U \leq 0.5$  e  $\bar{c}$  (coroa) se U > 0.5. As trajetórias da cadeia apresentada acima podem ser resumida na Tabela 2.1.

Sequências de números				
aleatórios possíveis	Estado amostrado			
ĒĒ	$x_0 = 0$			
- c c	não coalesce			
$car{c}$	não coalesce			
сс	$x_0 = 2$			

Tabela 2.1: Situações possíveis para duas transições das cadeias paralelas

A Tabela 2.2 fornece os resultados para n = 3 transições. Vemos que a distribuição da cadeia, condicionado que houve a coalescência em até 3 transições é  $\pi'(0) = \frac{1}{3}$ ,  $\pi'(1) = \frac{1}{3}$ ,  $\pi'(2) = \frac{1}{3}$ , que é igual a distribuição desejada.

Seq. de números	Estado	Seq. de números	Estado
aleat. possíveis	amostrado	aleat. possíveis	amostrado
$\overline{c}\overline{c}\overline{c}$	$x_{0} = 0$	cēē	$x_{0} = 0$
$\overline{c}\overline{c}c$	$x_0 = 1$	cēc	não coalesce
$\overline{c}c\overline{c}$	não coalesce	$cc\bar{c}$	$x_0 = 1$
<i>ccc</i>	$x_0 = 2$	ccc	$x_0 = 2$

Tabela 2.2: Situações possíveis para três transições das cadeias paralelas

Considerando n = 4 transições, obtemos a Tabela 2.3 e observamos que a distribuição da cadeia, condicionado que houve a coalescência em até 4 transições é  $\pi'(0) = \frac{5}{14}, \pi'(1) = \frac{4}{14}, \pi'(2) = \frac{5}{14}$ , que é diferente da distribuição desejada.

Estado	Seq. de números	Estado
amostrado	aleat. possíveis	amostrado
$x_{0} = 0$	cccc	$x_{0} = 0$
$x_0 = 1$	cccc	$x_0 = 1$
$x_{0} = 0$	cēcē	não coalesce
$x_{0} = 0$	ccīcī	$x_0 = 0$
$x_0 = 2$	cccc	$x_0 = 1$
não coalesce	ccēc	$x_0 = 2$
$x_0 = 1$	cēcc	$x_0 = 2$
$x_0 = 2$	cccc	$x_0 = 2$
	Estado amostrado $x_0 = 0$ $x_0 = 1$ $x_0 = 0$ $x_0 = 2$ não coalesce $x_0 = 1$ $x_0 = 2$	Estado amostradoSeq. de números aleat. possíveis $x_0 = 0$ $c\bar{c}\bar{c}\bar{c}$ $x_0 = 1$ $c\bar{c}\bar{c}c$ $x_0 = 0$ $c\bar{c}c\bar{c}$ $x_0 = 0$ $c\bar{c}c\bar{c}$ $x_0 = 2$ $ccc\bar{c}$ $x_0 = 1$ $c\bar{c}cc$ $x_0 = 1$ $c\bar{c}cc$ $x_0 = 2$ $cccc$

Tabela 2.3: Situações possíveis para quatro transições das cadeias paralelas

As Tabelas 2.1, 2.2 e 2.3 acima ilustram a evolução das trajetórias paralelas e nos fornecem a amostra, dado que houve a coalescência. Por estas tabelas observamos que a distribuição da cadeia para n = 2 e 4 transições é diferente da distribuição estacionária  $\pi$  e que a distribuição da cadeia para n = 3 transições é igual a distribuição estacionária  $\pi$ . A proposição a seguir fornece a distribuição da cadeia em ntransições, dado que houve coalescência em até n transições. **Proposição 2.7** Seja X uma cadeia de Markov com espaço de estados  $E = \{0, 1, 2\}$ , matriz de transição P igual a

$$\left(\begin{array}{rrrr} 0,5 & 0,5 & 0 \\ 0,5 & 0 & 0,5 \\ 0 & 0,5 & 0,5 \end{array}\right)$$

e com distribuição estacionária  $\pi(0) = \frac{1}{3}$ ,  $\pi(1) = \frac{1}{3}$ ,  $\pi(2) = \frac{1}{3}$ . Considere a regra de transição

$$f(X_{t-1}, U) = \begin{cases} X_t = \max\{X_{t-1} - 1, 0\} & se \ U \le 0, 5; \\ X_t = \min\{X_{t-1} + 1, 2\} & se \ U > 0, 5; \end{cases}$$
(2.10)

obtida a partir de P.

Sejam os eventos

 $C_n = \{ocorrência de coalescência das trajetórias da cadeia em até n transições\} e$  $\overline{C}_n = \{n \tilde{a} o \ ocorrência de \ coalescência das trajetórias da cadeia em até n transições\}.$ Para um número de transições n, n par, temos

$$\mathbb{P}(X_n = 0 \mid C_n) = \frac{\frac{2}{3}(2^{n-1} - 2) + 1}{2^n - 2}, 
\mathbb{P}(X_n = 1 \mid C_n) = \frac{\frac{2}{3}(2^{n-1} - 2)}{2^n - 2}, 
\mathbb{P}(X_n = 2 \mid C_n) = \frac{\frac{2}{3}(2^{n-1} - 2) + 1}{2^n - 2},$$
(2.11)

e para um número de transições n, n ímpar, temos

$$\mathbb{P}(X_n = x \mid C_n) = \frac{1}{3} , \forall x \in E.$$
(2.12)

Podemos observar pela proposição acima que para um número de transições n, n par, teremos uma distribuição da cadeia, condicionado que houve a coalescência, que não está de acordo com a distribuição desejada. Já para um número de transições n, nímpar, temos que a distribuição da cadeia, condicionado que houve a coalescência, está de acordo com a distribuição desejada.

A prova da Proposição 2.7 pode ser vista no Apêndice A.1.

### 2.4 Extensões do CFTP

Nelander e Häggström em [Häggström and Nelander, 1998] extenderam o CFTP para o caso anti-monótono ilustrando com simulações do modelo *hard-core* e do modelo de aglomerados aleatórios (*random cluster model*).

Foss, Tweedie e Corcoran em [Foss et al., 1998] desenvolveram uma variação do CFTP que pode ser usada quando a cadeia não for finita ou não for monótona, e que não requer um acoplamento vertical, ou seja, iniciando de todos os estados.

Murdoch e Green em [Murdoch and Green, 1998] extenderam o CFTP para um espaço de estados contínuo.

Wilson em [Wilson, 2000] propôs uma modificação do CFTP para gerar uma amostra aleatória perfeita da distribuição estacionária de uma cadeia de Markov. Este método é conhecido como *Read-Once* CFTP (chamaremos de leitura única do CFTP). Este método não guarda a seqüência de variáveis aleatórias geradas em uma iteração para ser reusada na próxima iteração, ou seja, a cada nova iteração ele gera uma nova seqüência de variáveis aleatórias independentes, que não depende da iteração anterior. Além disso, funciona como uma espécie de acoplamento indo para o futuro, evoluindo a cadeia por mais uma certa quantidade de tempo (aleatória) após a ocorrência da coalescência. Uma explicação de porque este método produz uma amostra (perfeita) da distribuição estacionária da cadeia de Markov pode ser vista em [Häggström, 2001].

Outras extensões e aplicações do algoritmo CFTP podem ser vistas em http://dimacs.rutgers.edu/~dbwilson/exact/.

## Capítulo 3

## Algoritmo de Fill

Assim como o algoritmo CFTP-monótono, o algoritmo interrompível de Fill assume que as cadeias de Markov sejam ergódicas e monótonas, requerendo que o espaço de estados seja parcialmente ordenado e que a matriz de transição P seja estocasticamente monótona com respeito a esta ordem parcial, para produzir uma amostra perfeita (exata) da distribuição  $\pi$ . Como veremos o algoritmo de Fill é baseado no método de rejeição da amostra e, ao contrário do CFTP, protege contra o vício causado pelo usuário impaciente pois, neste algoritmo, não existe correlação entre o tempo da trajetória (em termos do número de transições) e o estado amostrado. Este algoritmo foi introduzido por [Fill, 1998] e extendido em [Fill et al., 2000] para cadeias monótonas ou não-monótonas e com espaço de estados discreto ou contínuo (não necessariamente finito).

### 3.1 O Algoritmo

Nesta seção apresentamos o algoritmo Fill, descrito na seção 7.2 de [Fill et al., 2000], para o caso em que o espaço de estados é finito.

Seja X uma cadeia de Markov em um espaço de estados discreto finito parcialmente ordenado  $(E, \leq)$ , possuindo um único elemento mínimo  $\hat{0}$  e um único elemento máximo  $\hat{1}$ , e com matriz de transição P estocasticamente monótona (como definido em (2.4)). Além disso, seja

$$\widetilde{P}(x,y) = P(y,x).\pi(y)/\pi(x)$$
(3.1)

a matriz de probabilidades de transição em tempo reverso de X, onde  $\pi$  é a distribuição estacionária de P. Denotamos por  $P^t$  a matriz de probabilidades de transição de X em t-passos e  $\tilde{P}^t$  a matriz de probabilidades de transição em tempo reverso de X em t-passos.

A seguir apresentamos o algoritmo de Fill que, como qualquer procedimento de rejeição, consiste em propor uma amostra de  $\tilde{P}^t(\hat{0}, \cdot)$  como amostra de  $\pi(\cdot)$  a qual será aceita com probabilidade apropriada.

#### PASSOS DO ALGORITMO DE FILL

**Passo 1)** Escolher e fixar um inteiro  $t = t_i$  (qualquer) e  $X_t = \hat{0}$ . **Passo 2)** Percorrer a cadeia em tempo reverso  $\widetilde{P}$  por t passos obtendo  $\widetilde{X}_t = \hat{0}, \ \widetilde{X}_{t-1}, \cdots, \ \widetilde{X}_0 = x \text{ com } \widetilde{X}_{t-1} | \widetilde{x}_t, \cdots, \widetilde{X}_0 | \widetilde{x}_1 \text{ sucessivamente.}$ 

Para gerar esta trajetória, em tempo reverso, usamos a matriz de transição em tempo reverso  $\tilde{P}$ , como na equação (3.1), para obtermos uma regra de transição  $\tilde{f}$  dada por

$$\widetilde{X}_N = \widetilde{f}(\widetilde{X}_{N+1}, V_N)$$

com  $V_N$  sendo uma variável aleatória uniforme em [0,1] e  $N \in \{0,1,2,\cdots,t-1\}$ . **Passo 3)** Gerar  $[U_1 | \tilde{x}_0, \tilde{x}_1], \cdots, [U_t | \tilde{x}_{t-1}, \tilde{x}_t]$ . Sendo que  $U_N$  é uniforme no conjunto  $\{ u : \tilde{x}_N = f(\tilde{x}_{N-1}, u) \}$  para  $N \in \{1, 2, \cdots, t\}$  e f é a regra de transição monótona associada a P.

**Passo 4)** Revertendo-se novamente a direção do tempo (percorrendo a cadeia para frente) gera-se a trajetória  $\vec{Y} = (Y_0, \dots, Y_t) \operatorname{com} Y_0 = \hat{1}$ . Para obter esta trajetória usa-se a regra de transição f e a seqüência de variáveis aleatórias  $U_N$  geradas no passo anterior, ou seja,

$$Y_N = f(Y_{N-1}, U_N).$$

**Passo 5)** Verificar se  $Y_t = \hat{0}$ . Caso isto ocorra então, por construção, esta trajetória

terá coalescido com trajetória  $\vec{X} = (x_0 = x, \dots, x_t = \hat{0})$  que é vista também como uma trajetória de P. Neste caso, o valor  $\tilde{X}_0 = x$  será aceito como uma observação de  $\pi$ .

Passo 6) Senão repetimos o procedimento escolhendo um outro valor  $t = t_{i+1}$  com  $t_{i+1} > t_i$  retornando ao Passo 1.

Este algoritmo considera cada iteração como uma rotina de duas fases: na primeira fase propõe a amostra e na segunda aceita com determinada probabilidade, como uma implementação do método de rejeição da amostra (veja [Ross, 1994]). A idéia por trás do método de rejeição é gerarmos x de acordo com  $\tilde{P}^t(\hat{0}, \cdot)$  e aceitarmos como uma amostra (observação) de  $\pi(\cdot)$  com probabilidade

$$0 \leq C^{-1} \cdot \frac{\pi(x)}{\widetilde{P}^t(\hat{0}, x)} \leq 1, \ x \in E,$$

$$(3.2)$$

onde  $C \geq \frac{\pi(x)}{\tilde{P}^{t}(\hat{0},x)}$ ,  $\forall x \in E$ . Provamos no Apêndice A.2 que

$$C^{-1} \cdot \frac{\pi(x)}{\widetilde{P}^t(\hat{0}, x)} = \frac{P^t(\hat{1}, \hat{0})}{P^t(x, \hat{0})}.$$
(3.3)

Abaixo apresentamos o resultado e a respectiva prova de que o algoritmo Fill produz uma amostra com distribuição exata  $\pi$ .

**Teorema 3.1** Seja X uma cadeia de Markov ergódica com espaço de estados E finito, matriz de transição P e distribuição estacionária  $\pi$ . O Algoritmo de Fill fornece uma amostra exata da distribuição  $\pi$  via método de rejeição da amostra.

#### Prova:

Temos que a amostra fornecida pelo algoritmo de Fill na primeira fase, no Passo 2  $(\tilde{X}_0 = x)$ , é aceita como amostra de  $\pi$ , desde que na segunda fase, no Passo 4, se tenha  $Y_t = \hat{0}$ . Desta forma provamos a seguir que a probabilidade de aceitação desta amostra deverá ser idêntica a fornecida em (3.3) pelo método de rejeição, ou seja,

$$\mathbb{P}(Y_t = \hat{0} \mid X_0 = x, X_t = \hat{0}, Y_0 = \hat{1}) = \frac{P^t(\hat{1}, \hat{0})}{P^t(x, \hat{0})}.$$
(3.4)

Temos que

$$\mathbb{P}(Y_t = \hat{0} \mid X_0 = x, X_t = \hat{0}, Y_0 = \hat{1}) = \frac{\mathbb{P}(Y_t = \hat{0}, X_0 = x, X_t = \hat{0}, Y_0 = \hat{1})}{\mathbb{P}(X_0 = x, X_t = \hat{0}, Y_0 = \hat{1})} \\
= \frac{\mathbb{P}(Y_t = \hat{0}, X_t = \hat{0} \mid X_0 = x, Y_0 = \hat{1})}{\mathbb{P}(X_t = \hat{0} \mid X_0 = x, Y_0 = \hat{1})} \cdot \frac{\mathbb{P}(X_0 = x, Y_0 = \hat{1})}{\mathbb{P}(X_0 = x, Y_0 = \hat{1})} \quad (3.5) \\
= \frac{\mathbb{P}(Y_t = \hat{0} \mid X_0 = x, Y_0 = \hat{1})}{\mathbb{P}(X_t = \hat{0} \mid X_0 = x, Y_0 = \hat{1})} \quad (3.6)$$

$$\frac{\mathbb{P}(Y_t = 0 \mid X_0 = x, Y_0 = 1)}{P^t(x, \hat{0})}$$
(3.6)

$$=\frac{P^{t}(\hat{1},\hat{0})}{P^{t}(x,\hat{0})}$$
(3.7)

Na prova acima para passarmos da equação (3.5) para (3.6), no numerador temos pela monotonicidade que

$$Y_t = \hat{0} \Longrightarrow X_t = \hat{0}$$

e no denominador  $Y_0 = \hat{1}$  (fixo) não é dependente de  $(X_0 = x, \dots, X_t = \hat{0})$ . E passamos da equação (3.6) para (3.7) por termos  $X_0 = x$  independente de  $(Y_0 = \hat{1}, \dots, Y_t = \hat{0})$ .

Logo, aceitamos  $X_0 = x$  como amostra com probabilidade equivalente a probabilidade fornecida pelo método de rejeição da amostra em (3.3), o que implica que  $\mathbb{P}(X_0 = x) = \pi(x)$ .

Temos que o algoritmo de Fill é não viciado, com respeito ao usuário impaciente, pois a cada iteração i na qual não obtemos o resultado não reusamos informações desta iteração i na iteração i + 1.

### 3.2 Extensões do algoritmo de Fill

Murdoch e Rosenthal em [Murdoch and Rosenthal, 1998] fizeram uma extensão do algoritmo de Fill para cadeias não monótonas, o qual é conhecido como método FMR.

Thönnes em [Thönnes, 1999a] mostrou como o algoritmo de Fill pode ser extendido para um contexto com espaço de estados infinito, ilustrando como este algoritmo pode ser combinado com o amostrador de Gibbs para produzir uma amostra exata do processo de mistura de esferas penetráveis e modelos relacionados. Moller e Schladitz em [Moller and Schladitz, 1999] também extenderam o algoritmo de Fill para cadeias com espaço de estados infinito.

Huber em [Huber, 1999] mostrou como usar o FMR para gerar amostras do hard core gas model e do Windom-Rowlinson mixture model contínuo.

Outras extensões e aplicações do algoritmo de Fill podem ser vistas em http://dimacs.rutgers.edu/~dbwilson/exact/ .

### Capítulo 4

## Modelo de Ising

Para ilustrar o algoritmo CFTP-monótono e o algoritmo de Fill apresentados, aplicamos estes algoritmos ao modelo de Ising ferromagnético em duas dimensões no próximo capítulo usando o amostrador de Gibbs (*Gibbs Sampler*). Assim, neste capítulo, mostramos como aplicar o amostrador de Gibbs ao modelo de Ising para simular uma amostra da distribuição deste modelo. Mas primeiro apresentamos o modelo de Ising em duas dimensões com suas principais características. Em seguida descrevemos o amostrador de Gibbs que consiste em construir uma cadeia de Markov ergódica que converge para a distribuição  $\pi$ . A cadeia construída pelo amostrador de Gibbs, aplicada ao modelo de Ising ferromagnético em duas dimensões, é monótona em um espaço de estados E parcialmente ordenado com elementos máximo e mínimo, e possui uma regra de transição monótona com respeito a esta ordem parcial.

### 4.1 Medida de Gibbs

Apresentamos algumas definições e teoremas importantes sobre campo aleatório Markoviano e medida de Gibbs.

Seja  $\Lambda \subseteq \mathbb{Z}^2$  um modelo no reticulado ou sistema de spins. Este reticulado  $\Lambda$  consiste em um conjunto de elementos, onde cada elemento  $j = (j_1, j_2)$  será chamado de sítio. Além disso, seja  $E = A^{\Lambda}$  o espaço de estados ou conjunto de todas as configurações possíveis, onde cada configuração  $x = \{x(j), j \in \Lambda\}$  é uma função

 $x : \Lambda \to A$  que atribui a cada sítio  $j \in \Lambda$  um valor de A, denotado por x(j). A este valor x(j) chamamos de spin do sítio j.

Para  $j, l \in \Lambda$ ,  $l \sim j$  denota que l é vizinho adjacente de j. Assim definimos os vizinhos do sítio j como

$$\rho(j) = \{l : l \sim j\}$$

e

$$\rho = \{\rho(j); j \in \Lambda\},\$$

ou seja,  $\rho$  representa uma coleção de subconjuntos de  $\Lambda$  chamado de sistema de vizinhança.

A seguir apresentamos as definições de campo aleatório e campo aleatório Markoviano.

**Definição 4.1** Um vetor aleatório  $x = (x(j), j \in \Lambda)$  definido em um espaço de probabilidade e assumindo valores em  $E = A^{\Lambda}$  é chamado de campo aleatório.

**Definição 4.2** Uma configuração  $x \in E = A^{\Lambda}$  é dita ser um campo aleatório Markoviano em  $\Lambda$  se para cada  $x \in E$  a distribuição condicional de x(j) dado x(l), para todo  $l \in \Lambda$  e  $l \neq j$ , depende apenas dos valores de x(l) tais que  $l \sim j$ . Em outras palavras, x é um campo aleatório Markoviano se para todo j,  $l \in \Lambda$  e  $l \neq j$  tal que  $\mathbb{P}(x(l), para todo l \neq j) > 0$ , tem-se

$$\mathbb{P}(x(j) = a | x(l), \ \forall \ l \neq j) = \ \mathbb{P}(x(j) = a | x(l), \ \forall \ l \sim j), \ a \in A.$$

Logo, em um campo aleatório Markoviano, a probabilidade do spin de um sítio assumir um valor, condicionado à todo campo aleatório, não depende de todo o campo aleatório, mas somente dos valores dos spins dos sítios que são vizinhos mais próximos deste sítio.

Temos que um subconjunto  $B \subseteq \Lambda$  é chamado de clique se  $l \in \rho(j)$  para todo  $l, j \in B$ , ou seja, se todos os sítios de B são vizinhos entre si.

Definição 4.3 A medida de Gibbs para um sistema de vizinhos  $\rho$  é um campo aleatório da seguinte forma

$$\pi(x) = \frac{1}{Z} \cdot \exp\{-H(x)\} > 0, \ \forall x \in E,$$
(4.1)

sendo  $Z = \sum_{x \in E} \exp\{-H(x)\}$  a função de partição ou constante de normalização e  $H(x) = \sum_{B \subset \Lambda} \phi_B(x)$  a função de energia, sendo a soma dada sobre os subconjuntos B de  $\Lambda$  de potenciais de interação  $\phi_B$ , tais que  $\phi_B = 0$  se B não é um clique.

De acordo com o teorema de Hammersley-Clifford, (ver, por exemplo, [Fismen, 1997]), apresentado a seguir, existe uma equivalência entre a medida de Gibbs e um campo aleatório Markoviano.

**Teorema 4.4** (Hammersley-Clifford) Seja  $\rho$  um sistema de vizinhança. Um campo aleatório é um campo Markoviano com respeito a  $\rho$  se e só se é uma medida de Gibbs com respeito a  $\rho$ .

Para muitos campos aleatórios Markovianos encontramos problemas na obtenção de uma amostra, pois a função de partição  $Z = \sum_{x \in E} \exp\{-H(x)\}$  é uma soma difícil de ser calculada devido a alta dimensão do espaço de estados E destes modelos.

### 4.2 Modelo de Ising

O modelo de Ising foi proposto pelo físico Ernst Ising um estudante de Lens que escreveu sua tese de doutorado neste modelo. Ising tentou explicar, através deste modelo, certos fatos observados empiricamente sob materiais ferromagnéticos. Quando Ising em 1925 publicou um resumo deste resultado, ele fixou que o modelo foi sugerido por Lens em um paper em 1920 no qual dava uma vaga idéia deste modelo. Maiores informações sobre a origem e desenvolvimento do modelo de Ising podem ser vistas em [Kindermann and Snell, 1980] e em referências lá citadas. O modelo de Ising é um caso particular da medida de Gibbs e pelo Teorema 4.4 temos que é um campo aleatório Markoviano, possuindo as seguintes características: \* Considere que temos um reticulado quadrado de sítios  $\Lambda \subseteq \mathbb{Z}^2$ .

\* Cada sítio possui um spin com duas orientações possíveis, para cima ou para baixo, que denotamos por +1 ou -1, respectivamente, ou seja,  $A = \{-1, +1\}$ .

\* Existe uma energia de interação  $\phi_B$  entre dois spins  $x(j) \in x(l)$ , com  $j, l \in B$ , tal que

$$\phi_B = \begin{cases} 0 & \text{se B não \'e clique;} \\ Jx(j)x(l) & \text{se B \'e clique} \end{cases}$$
(4.2)

onde J é uma propriedade do material e o sistema de vizinhança é definido por  $(l \sim j) \iff |l - j| = 1.$ 

 $\star$  Pode existir interação com um campo magnético externo  $\alpha$ .

Portanto, este modelo possui a seguinte função de energia para uma configuração  $x \in E$ , onde  $E = \{-1, +1\}^{\Lambda}$  é o espaço de estados,

$$H(x) = -\beta \left[J \sum_{j \in \Lambda} \sum_{l \sim j} x(j) x(l) + \sum_{j \in \Lambda} \alpha(j) x(j) + \sum_{r \in \mathbb{Z}^2 \setminus \Lambda} \sum_{\substack{j \in \Lambda \\ j \sim r}} J_{jr} x(j) y(r)\right] \quad (4.3)$$

sendo  $j = (j_1, j_2) \neq l = (l_1, l_2), x(j) \in \{-1, +1\}, \beta = \frac{1}{KT}$  sendo K a constante de Boltzmann e T a temperatura absoluta do material,  $\alpha(j)$  é a força do campo magnético externo medida no sítio  $j \in y \in \{-1, +1\}^{\mathbb{Z}^2 \setminus \Lambda}$  (onde  $\mathbb{Z}^2 \setminus \Lambda$  denota todo o reticulado  $\mathbb{Z}^2$  exceto a região de intersecção com o reticulado  $\Lambda$ ). A parcela  $\sum_{r \in \mathbb{Z}^2 \setminus \Lambda} \sum_{\substack{j \in \Lambda \\ j \sim r}} J_{jr}x(j)y(r)$ , na equação (4.3), representa interação dos spins do reticulado  $\Lambda$  com os spins fora do reticulado, e é conhecida como condição de fronteira y para volumes finitos  $\Lambda$ .

Definimos o conjunto fronteira  $(\partial \Lambda)$  de um volume  $\Lambda$  como sendo o conjunto de todos os sítios  $r \notin \Lambda$ , tal que  $r \in \rho(j)$ ,  $j \in \Lambda$ . Uma condição de fronteira é positiva se, para todo  $r \in \partial \Lambda$ , y(r) assumir valor +1, e uma condição de fronteira é negativa se, para todo  $r \in \partial \Lambda$ , y(r) assumir valor -1. Uma condição de fronteira é chamada de livre se  $J_{jr} = 0$  na equação (4.3), ou seja, se não existir fronteira fixada. Existe também uma condição de fronteira que é chamada de periódica e que não entraremos em detalhes, mas o leitor interessado pode ver, por exemplo, em [Kindermann and Snell, 1980].

Pela equação (4.3), no caso em que J > 0, chamado de caso ferromagnético ou caso atrativo, a energia de uma configuração será mínima se todos os spins tiverem a mesma direção, ou seja, se x(j) e x(l) tiverem os mesmos sinais para todo  $j, l \in \Lambda$ tais que  $l \sim j$ . E, neste caso, um determinado spin estará provavelmente alinhado na mesma direção da maioria de seus vizinhos mais próximos e conseqüentemente se existir um campo externo os spins tendem a ficar orientados de acordo com o campo externo. Já no caso em que J < 0, chamado de caso antiferromagnético ou caso repulsivo, a energia será mínima quando os spins tiverem sinais contrários, ou seja, se x(j) e x(l) tiverem sinais diferentes para todo  $j \in \Lambda$  tal que  $l \sim j$ . Consideraremos modelo de Ising para o caso ferromagnético com J = 1 e sem a presença de campo externo, ou seja, na equação (4.3)  $\alpha(j) = 0$ .

O modelo de Ising é um caso particular da medida de Gibbs. Assim este modelo terá a seguinte distribuição, para todo  $\Lambda \subseteq \mathbb{Z}^2$ ,

$$\pi(x) = \frac{1}{Z} \cdot \exp\{-H(x)\} > 0, \ \forall \ x \in E = \{-1, +1\}^{\Lambda},$$
(4.4)

onde H(x) é a função de energia do modelo de Ising dada em (4.3).

Consideremos agora um subconjunto  $\Lambda \subset \mathbb{Z}^2$ , finito centrado na origem, dado por

$$\Lambda = \{ j \in \mathbb{Z}^2 : -L \le j_i \le L, \ i = 1, 2 \}; \ L \in \mathbb{N}.$$
(4.5)

Este subconjunto será chamado de caixa finita. Conseqüentemente por  $\Lambda$  ser finito teremos alguma condição de fronteira.

Definiremos agora a medida de Gibbs em volume infinito.

**Definição 4.5** A medida de probabilidade  $\pi^+$  ( $\pi^-$ ) definida sobre  $\{-1, +1\}^{\mathbb{Z}^2}$  é chamada medida limite de Gibbs do modelo de Ising ferromagnético em duas dimensões sobre  $\mathbb{Z}^2$ , com temperatura inversa  $\beta$  e condição de fronteira positiva (negativa, respectivamente) se existir uma seqüência crescente de caixas, satisfazendo  $\Lambda_n \longrightarrow \mathbb{Z}^2$  quando  $n \longrightarrow \infty$ , tal que

$$\pi^{+} = \lim_{\Lambda_n \nearrow \mathbb{Z}^2} \pi^{+}_{\Lambda_n} \quad (\pi^{-} = \lim_{\Lambda_n \nearrow \mathbb{Z}^2} \pi^{-}_{\Lambda_n} , \quad respectivamente)$$
(4.6)

Teorema 4.6 (Limite Termodinâmico) Para o modelo de Ising ferromagnético de vizinhos mais próximos e com condição de fronteira positiva (negativa, respectivamente) definido sobre  $\mathbb{Z}^2$ , a temperatura inversa  $\beta$ , existe o limite termodinâmico

$$\pi^{+} = \lim_{\Lambda_n \nearrow \mathbb{Z}^2} \pi^{+}_{\Lambda_n} \quad (\pi^{-} = \lim_{\Lambda_n \nearrow \mathbb{Z}^2} \pi^{-}_{\Lambda_n} , \quad respectivamente)$$
(4.7)

A prova deste resultado pode ser vista em [Holley, 1974].

Temos que o modelo de Ising em duas dimensões sem campo magnético externo é importante por ser o modelo mais simples, em mecânica estatística, que exibe transição de fase. O modelo possui um valor crítico para  $\beta$  (ou para temperatura T) denotado por  $\beta_c$  ( $T_c$  respectivamente). A seguir definimos quando ocorre a transição de fase.

**Definição 4.7** Diremos que existe a transição de fase no modelo de Ising ferromagnético sem campo magnético externo em duas dimensões se existe um  $\beta_c$  tal que se  $\beta > \beta_c$  então

$$\pi^+ \neq \pi - \tag{4.8}$$

Sabe-se que o valor crítico de beta, em duas dimensões, é  $\beta_c = \frac{1}{2} \log(1 + \sqrt{2}) \simeq 0.4407$ (ver, por exemplo [Georgii, 1988]).

### 4.3 Amostrador de Gibbs

#### 4.3.1 Algoritmo Metropolis-Hastings

Suponha que queremos obter uma amostra de uma distribuição  $\pi$ . O algoritmo Metropolis-Hastings consiste em construir uma cadeia de Markov com espaço de estados E, matriz de transição de probabilidades P e distribuição estacionária  $\pi$ . Esta distribuição estacionária pode ser obtida resolvendo as equações

$$\pi(x) = \sum_{y \in E} \pi(y) P(y, x) \quad e \quad \sum_{y \in E} \pi(y) = 1.$$
(4.9)

Se a cadeia é reversível com respeito a  $\pi$ , isto é,

$$\pi(x)P(x,y) = \pi(y)P(y,x) \quad \forall \ x,y \in E$$

$$(4.10)$$

então ela satisfaz as equações (4.9).

O algoritmo Metropolis-Hastings considera probabilidades de transição P(x,y)que sejam reversíveis com respeito a  $\pi$ . As probabilidades de transição P(x,y) são dadas por

$$P(x,y) = q(x,y)\sigma_{xy} \tag{4.11}$$

onde os q(x,y)'s são as probabilidades de transição de outra cadeia arbitrária tal que

$$q(x,y) > 0 \Longrightarrow q(y,x) > 0 \quad \forall x, y \in E,$$
(4.12)

e  $\sigma_{xy}$ é a probabilidade de aceitação do estado y proposto.

Portanto, o algoritmo de Metropolis-Hastings consiste de duas etapas.

Etapa 1) A partir de um estado x, propor um estado y com probabilidade q(x, y). Etapa 2) Consiste em aceitar o estado y com probabilidade  $\sigma_{xy}$ , caso contrário a cadeia permanece no estado x.

Hastings em [Hastings, 1970] mostrou que escolhendo a probabilidade de aceitação como sendo

$$\sigma_{xy} = \min\left(1, \frac{\pi(y)q(y, x)}{\pi(x)q(x, y)}\right)$$
(4.13)

a distribuição estacionária  $\pi$  é obtida.

### 4.3.2 Amostrador de Gibbs

O amostrador de Gibbs é um caso especial do Metropolis-Hastings. Este algoritmo utiliza distribuições condicionais para produzir uma amostra de uma distribuição  $\pi$ .

Suponha que se deseje simular uma configuração aleatória x tomando valores em  $\{-1, +1\}^{\Lambda}$  de acordo com a medida de Gibbs  $\pi$  em (4.1). O amostrador de Gibbs consiste em construir uma cadeia de Markov que é irredutível e aperiódica, e se  $\Lambda$  for finito conseqüentemente será ergódica, tal que esta cadeia tenha distribuição estacionária  $\pi$ .

Para obter uma amostra de  $\pi$  usando o amostrador de Gibbs, inicia-se a cadeia com uma configuração  $x \in \{-1, +1\}^{\Lambda}$  e vai-se atualizando os sítios desta configuração, ou seja, percorrendo uma cadeia de Markov. Para atualização destes sítios utiliza-se a distribuição condicional de x(j) dado os outros elementos de x, ou seja,  $\pi(x(j)|x(-j))$ . Estes sítios são escolhidos aleatoriamente, um de cada vez, para serem atualizados e os valores nos sítios  $l \neq j$  não mudam quando estamos atualizando o sítio j. Assim, passamos de um estado x para um estado y, onde  $x \in y$  diferem apenas no sítio j, com probabilidade

$$q(x,y) = \pi(x(j)|x(-j)). \tag{4.14}$$

E no amostrador de Gibbs a probabilidade de aceitação é

$$\sigma_{xy} = \min\left(1, \frac{\pi(y)q(y,x)}{\pi(x)q(x,y)}\right) = 1 .$$
(4.15)

Conseqüentemente, no sítio j, teremos o valor x(j) = +1 com a seguinte probabilidade condicional

$$\pi(x(j) = +1|x(-j)) = \frac{\pi(x(j) = +1, x(-j))}{\pi(x(j) = +1, x(-j)) + \pi(x(j) = -1, x(-j))}$$
(4.16)

Para a atualização dos sítios a cada instante t escolhemos um número  $U_t$  uniforme no intervalo [0,1] e um sítio  $j_t \in \Lambda$ , aleatoriamente (uniformemente) em  $\Lambda$ , para ser atualizado, independentemente para cada t. Então atribuímos o valor +1 para o sítio  $j_t$  se

$$U_t \leq \pi(x(j_t) = +1|x(-j_t)) \tag{4.17}$$

e atribuímos o valor -1 para o sítio  $j_t$  caso contrário.

Veremos agora como aplicar o amostrador de Gibbs no modelo de Ising.

Não temos uma forma direta para obter uma amostra da distribuição do modelo de Ising ferromagnético em duas dimensões (equação (4.4)), devido ao amplo número de configurações possíveis. Por exemplo, se tivermos  $|\Lambda| = 64 \times 64$  a função de partição de (4.4) será uma soma de  $2^{64\times 64}$  termos, que é extremamente difícil de ser calculada. No amostrador de Gibbs não calculamos a função de partição pois este método é baseado em distribuições condicionais. A seguir vemos como aplicar este método ao modelo de Ising.

Aplicando o amostrador de Gibbs ao modelo de Ising ferromagnético, é fácil ver que  $\pi(x(j) = +1|x(-j))$  será maior quanto maior for o número de vizinhos do sítio j assumindo valor +1, pois

$$\pi(x(j) = +1|x(-j)) = \frac{\pi(x(j) = +1, x(-j))}{\pi(x(j) = +1, x(-j)) + \pi(x(j) = -1, x(-j))}$$
(4.18)  
=  $\left( 1 + \exp\left(-2\beta \sum_{\substack{l \in \Lambda \\ l \sim j}} x(l)\right) \right)^{-1}$ , (4.19)

onde  $\pi(x(j) = \cdot, x(-j))$  é a medida de Gibbs com a função de energia do modelo de Ising sem a presença de campo externo, com condição de fronteira livre e J = 1. A probabilidade acima é provada no Apêndice A.3 e observe que ela não depende da função de partição o que facilitará muito os cálculos.

De acordo com a probabilidade (4.18) temos uma matriz de probabilidades de transição P dada por

$$P(x, y = (x; x(j) \leftarrow a)) = \begin{cases} \frac{1}{|\Lambda|} \cdot \pi(x(j) = a | x(-j)) & \text{se apenas o sítio } j \\ & \text{foi atualizado;} \\ 0 & \text{se } y \in x \text{ diferem em} \\ & \text{mais que 1 sítio;} \end{cases}$$
(4.20)

onde  $y = (x; x(j) \leftarrow a)$  denota a configuração x a qual tem spin  $a \in \{-1, +1\}$  no sítio j.

Na prática, associado a esta matriz de transição P, temos uma regra de transição

da seguinte forma:

$$f(x,j,U) = \begin{cases} \{x(j) = +1, x(-j)\}, & \text{se } U \le \left(1 + \exp\left(-2\beta \sum_{l \in \Lambda} x(l)\right)\right)^{-1}; \\ \{x(j) = -1, x(-j)\}, & \text{caso contrário.} \end{cases}$$
(4.21)

A regra de transição (4.21) é da forma  $X_t = f(X_{t-1}, j_t, U_t)$ , onde  $X_t$  difere de  $X_{t-1}$ no máximo no sítio  $j_t$  (escolhido no instante t para ser atualizado). Considerando  $|\Lambda| = N$ , essa regra é construída como  $f : E \times \text{Uniforme}\{1, 2, \dots, N\} \times [0, 1] \longrightarrow E$ .

A seguir mostramos que o modelo de Ising ferromagnético admite uma ordem parcial.

De acordo com a probabilidade (4.19), uma configuração x será considerada menor do que uma configuração y ( $x \leq y$ , onde  $\leq$  representa a ordem parcial) se tivermos

$$x(j) \le y(j), \ \forall j \in \Lambda.$$
 (4.22)

Com a ordem parcial descrita em (4.22), o espaço de estados terá uma configuração máxima ( $\hat{1}$ ) e uma mínima ( $\hat{0}$ ), onde a máxima é aquela com todos os sítios em  $\Lambda$  assumindo valor +1 (spins para cima) e a mínima com todos os sítios assumindo valor -1 (spins para baixo).

Temos também da ordem parcial descrita em (4.22) que, se  $x \leq y$ , então

$$\pi(x(j) = +1 \mid x(-j)) \leq \pi(y(j) = +1 \mid y(-j)), \quad \forall j \in \Lambda.$$
(4.23)

Baseado na ordem parcial (4.22), mostramos a seguir que a regra de transição (4.21) é monótona.

Dada a probabilidade em (4.23), para atualizarmos um sítio j, geramos um número aleatório U uniforme em [0,1] e atribuímos +1 para  $x(j) \in y(j)$  se

$$U \leq \pi(x(j) = +1|x(-j)) \leq \pi(y(j) = +1|y(-j)), \quad \forall j \in \Lambda.$$
(4.24)

Assim, de acordo com (4.24) temos que se a regra de transição  $f = f(\cdot, j, U)$  em (4.21) atribuir o valor +1 para x(j) então também atribuirá um valor +1 para y(j), mas a recíproca nem sempre será verdadeira bastando verificar as probabilidades de

(4.24). Da mesma se forma se f atribui um valor -1 para x(j) e y(j) se U é maior que as probabilidades acima e devemos observar que quando f atribuir um valor -1para y(j) então irá atribuir também um valor -1 para x(j), e novamente a recíproca nem sempre será verdadeira.

Portanto, temos que a regra de transição (4.21) é monótona com respeito a ordem parcial  $\leq$ . Pois se tivermos  $x \leq y$  a atualização destas configurações, feitas por esta regra de transição, irá sempre manter a sua ordem parcial, ou seja,  $f(x, j, U) \leq$ f(y, j, U), para todo U e j fixados.

Mostramos que o modelo de Ising ferromagnético em duas dimensões possui uma ordem parcial  $\leq$  e uma regra de transição monótona com respeito a  $\leq$ , então de acordo com a monotonicidade descrita no Capítulo 2, P é estocasticamente monótona. Vimos também que o espaço de estados E possui estados máximo e mínimo.

Temos então que o modelo de Ising em duas dimensões satisfaz as condições impostas pelos algoritmos CFTP-monótono e de Fill apresentados. Assim, no próximo capítulo, apresentaremos resultados de simulações usando estes algoritmos.

### Capítulo 5

## Simulações

Neste capítulo o objetivo é ilustrar os métodos de simulação perfeita apresentados, isto é, o algoritmo CFTP-monótono e o algoritmo de Fill descritos nos Capítulos 2 e 3, respectivamente. Para esta ilustração fizemos uma aplicação destes métodos ao modelo de Ising em duas dimensões para obtermos uma amostra perfeita da distribuição  $\pi$  deste modelo, tal que  $\pi : E \longrightarrow [0,1]$  onde  $E = \{-1,+1\}^{\Lambda}$ .

Consideramos o modelo de Ising ferromagnético sem a presença de campo externo descrito no Capítulo 4 e com condições de fronteira livre, positiva e negativa em reticulados  $\Lambda$  de tamanho 64×64, 256×256 e 512×512. As simulações foram feitas usando-se o Matlab (versão 6.0) na rede UNIX do IME-USP para as duas técnicas apresentadas.

## 5.1 Algoritmo CFTP-monótono aplicado ao modelo de Ising

No capítulo anterior, verificamos que o espaço de estados E do modelo de Ising ferromagnético em duas dimensões admite uma ordem parcial  $\leq$ , possuindo estados máximo  $\hat{1}$  e mínimo  $\hat{0}$ , tais que  $\hat{0} \leq x \leq \hat{1}$ , para todo  $x \in E$ . De acordo com esta ordem parcial obtemos que a regra de transição dada por (4.21) é monótona. Então será suficiente iniciarmos as trajetórias da cadeia de Markov partindo dos dois estados extremos  $\hat{0}$  (máximo) e  $\hat{1}$  (mínimo), que correspondem às configurações em que todos os sítios assumem valor +1 e -1, respectivamente, e verificarmos a coalescência completa entre essas trajetórias. Pois uma vez ocorrido a coalescência entre as trajetórias iniciadas dos estados  $\hat{0}$  e  $\hat{1}$ , terá ocorrido também a coalescência entre todas as outras trajetórias se a cadeia fosse iniciada em um outro estado qualquer.

Vimos no capítulo anterior que o amostrador de Gibbs constrói cadeias de Markov irredutíveis e aperiódicas (ergódicas) garantindo a existência de uma única distribuição estacionária que, por construção, coincidirá com  $\pi$ .

Para aplicarmos o CFTP-monótono seguimos os mesmos passos do algoritmo apresentado no Capítulo 2 usando o amostrador de Gibbs com a regra de transição (4.21). Como esta regra de transição também usa os sítios para serem feitas as transições, no **Passo 3** do algoritmo, estamos reusando além da mesma seqüência de variáveis aleatórias uniformes ( $\{U_k\}_{k=-k_{i-1}+1}^0$ ) a mesma seqüência de sítios ( $\{j_k\}_{k=-k_{i-1}+1}^0$ ) geradas na iteração anterior.

Nas Figuras 5.1 e 5.2 apresentamos amostras perfeitas geradas de acordo com as suposições acima, em reticulados  $|\Lambda| = 64 \times 64$ ,  $|\Lambda| = 256 \times 256$  com condição de fronteira livre. Os sítios que assumem o valor +1 correspondem aos sítios brancos e os sítios que assumem o valor -1 correspondem aos sítios pretos.

Observe que nessas amostras assumimos diferentes valores de  $\beta$  o que nos proporcionou um número diferente de iterações *i* do algoritmo. Lembramos que cada valor de *i* corresponde a  $k_i = 2^{i-1}$  transições da cadeia.

Como estamos trabalhando com o modelo de Ising ferromagnético podemos observar a influência dos vizinhos de um sítio nas Figuras 5.3, 5.4, 5.5 e 5.6, onde as condições de fronteira são positiva ou negativa. Podemos observar também que quando  $\beta$  é pequeno ( $\beta = 0.1$ , por exemplo) a fronteira infuencia muito pouco no alinhamento dos spins dos sítios do reticulado. Esta influência é evidente quando  $\beta = 0.3$  pois os spins dos sítios da borda de  $\Lambda$  ficam em sua grande maioria orientados com a fronteira. No entanto, podemos observar que a medida que os reticulados vão crescendo, as amostras obtidas com condição de fronteira positiva são cada vez mais parecidas com as que tem condição de fronteira negativa, visto que a influência das fronteiras ocorre basicamente nos sítios da borda do reticulado.

Sabemos que não existe transição de fase para reticulados finitos. Mas fizemos as simulações para  $\beta < \beta_c$ , pois quando  $\beta$  se aproxima de  $\beta_c$  a simulação torna-se extremamente lenta.



Figura 5.1: Amostras obtidas pelo algoritmo CFTP-monótono com  $|\Lambda| = 64 \times 64$  e condição de fronteira livre.



Figura 5.2: Amostras obtidas pelo algoritmo CFTP-monótono com  $|\Lambda| = 256 \times 256$ e condição de fronteira livre.



Figura 5.3: Amostras obtidas pelo algoritmo CFTP-monótono com  $|\Lambda| = 64 \times 64$ ,  $|\Lambda| = 256 \times 256$  e  $|\Lambda| = 512 \times 512$  respectivamente e condição de fronteira positiva.



Figura 5.4: Amostras obtidas pelo algoritmo CFTP-monótono com  $|\Lambda| = 64 \times 64$ ,  $|\Lambda| = 256 \times 256$  e  $|\Lambda| = 512 \times 512$  respectivamente e condição de fronteira negativa.



Figura 5.5: Amostras obtidas pelo algoritmo CFTP-monótono com  $|\Lambda| = 64 \times 64$ ,  $|\Lambda| = 256 \times 256$  e  $|\Lambda| = 512 \times 512$  respectivamente e condição de fronteira positiva.



Figura 5.6: Amostras obtidas pelo algoritmo CFTP-monótono com  $|\Lambda| = 64 \times 64$ ,  $|\Lambda| = 256 \times 256$  e  $|\Lambda| = 512 \times 512$  respectivamente e condição de fronteira negativa.

### 5.2 Algoritmo de Fill aplicado ao modelo de Ising

Para aplicarmos este algoritmo ao modelo de Ising usamos também o amostrador de Gibbs. Então, seguindo os passos do algoritmo apresentado no Capítulo 3 temos: Passo 1) A cada iteração escolhemos um novo valor de  $t = t_i = t_0 \cdot 2^{i-1}$  e  $X_t = \hat{0}$ , onde  $t_0 = |\Lambda|$ .

**Passo 2)** Obter  $\widetilde{X}_t = \hat{0}, \ \widetilde{X}_{t-1}, \cdots, \ \widetilde{X}_0 = x$  usando a regra de transição  $\widetilde{f}$  que é

$$\widetilde{X}_{t-1} = \widetilde{f}(\widetilde{X}_t, j_{t-1}, V_{t-1})$$

onde  $V_t$  é uma variável aleatória uniforme em [0,1] e esta regra de transição é obtida a partir da cadeia em tempo-reverso  $\tilde{P}$ . Como a cadeia de Markov construída pelo amostrador de Gibbs é reversível, temos que  $\tilde{P} = P$ . Logo  $\tilde{f} = f$  e f é dada em (4.21). **Passo 3)** Para gerarmos variáveis aleatórias uniformes em  $[U_1|\tilde{x}_0, \tilde{x}_1], \dots, [U_t|\tilde{x}_{t-1}, \tilde{x}_t]$ , devemos usar a função de transição f de P dada em (4.21), procedendo da seguinte forma:

Suponha que estamos indo de  $\tilde{x}_t$  para  $\tilde{x}_{t-1}$  com o sítio j escolhido no instante t-1para ser atualizado. Vamos supor também que no instante t este sítio possui spin para cima então devemos gerar uma variável aleatória  $U_t$  uniforme no conjunto

$$\left[0, \left(1 + \exp\left(-2\beta \sum_{\substack{l \in \Lambda \\ l \sim j}} x_t(l)\right)\right)^{-1}\right]$$

não dependendo se este sítio vai assumir spin para cima ou para baixo no instante (t-1). Da mesma forma se o sítio possuir spin para baixo devemos gerar uma variável aleatória  $U_t$  uniforme no conjunto

$$\left[ \left( 1 + \exp\left(-2\beta \sum_{\substack{l \in \Lambda \\ l \sim j}} x_t(l)\right) \right)^{-1}, 1 \right]$$

e também não vai depender se este sítio vai assumir spin para baixo ou para cima no instante (t-1).

Passo 4) Percorrendo a cadeia para frente partindo de  $Y_0 = \hat{1}$  usamos a regra de

transição dada em (4.21), ou seja,

$$Y_t = f(Y_{t-1}, j_t, U_t)$$

com a seqüência de uniformes geradas no passo anterior e também a mesma seqüência de sítios do passo anterior.

**Passo 5**) A cada iteração verificamos se  $Y_t = \hat{0}$  caso isso ocorra aceitamos  $\widetilde{X}_0 = x$  como amostra de  $\pi$  e interrompemos o algoritmo, senão retornamos ao **Passo 1** com i := i + 1.

Nas Figuras 5.7, 5.8, 5.9, 5.10, 5.11 e 5.12 apresentamos amostras geradas de acordo com algoritmo acima, com valores iniciais de t iguais a  $t_0 = 64^2 = 2^{12}$ ,  $t_0 = 256^2 = 2^{16}$  e  $t_0 = 512^2 = 2^{18}$ , para reticulados de tamanho  $64 \times 64$ ,  $256 \times 256$ e  $512 \times 512$ , respectivamente. Pois se tomarmos valores de t menores que  $t_0$  para estes reticulados, de acordo com a equação (3.4) do Capítulo 3, a probabilidade de que a amostra obtida seja aceita é zero. Como no CFTP obtemos também uma amostra perfeita para alguns valores de  $\beta$ , com seus respectivos números de iterações e transições.

As amostras obtidas pelo algoritmo de Fill tem o mesmo comportamento das obtidas pelo algoritmo CFTP-monótono, assim os comentários são análogos.



Figura 5.7: Amostras obtidas pelo algoritmo de Fill com  $|\Lambda| = 64 \times 64$  e condição de fronteira livre.



Figura 5.8: Amostras obtidas pelo algoritmo de Fill com  $|\Lambda| = 256 \times 256$  e condição de fronteira livre.



Figura 5.9: Amostras obtidas pelo algoritmo de Fill com  $|\Lambda| = 64 \times 64$ ,  $|\Lambda| = 256 \times 256$ e  $|\Lambda| = 512 \times 512$  respectivamente e condição de fronteira positiva.



Figura 5.10: Amostras obtidas pelo algoritmo de Fill com  $|\Lambda| = 64 \times 64$ ,  $|\Lambda| = 256 \times 256$  e  $|\Lambda| = 512 \times 512$  respectivamente e condição de fronteira negativa.



Figura 5.11: Amostras obtidas pelo algoritmo de Fill com  $|\Lambda| = 64 \times 64$ ,  $|\Lambda| = 256 \times 256$  e  $|\Lambda| = 512 \times 512$  respectivamente e condição de fronteira positiva.



Figura 5.12: Amostras obtidas pelo algoritmo de Fill com  $|\Lambda| = 64 \times 64$ ,  $|\Lambda| = 256 \times 256$  e  $|\Lambda| = 512 \times 512$  respectivamente e condição de fronteira negativa.

### 5.3 Resultados

O tempo de processamento no algoritmo CFTP-monótono foi menor do que no algoritmo de Fill para os mesmos valores de  $\beta \in \Lambda$ . O que já era de se esperar, pois no algoritmo de Fill, no **Passo 2**, quando percorremos a cadeia em tempo reverso a cada novo sítio a ser atualizado temos que gerar uma variável aleatória uniforme como apresentada no **Passo 3**. Estas variáveis serão usadas no **Passo 4** quando percorremos a cadeia para frente, o que torna seu processamento mais lento pois temos que percorrer a cadeia duas vezes para frente e para trás. Por termos que percorrer a cadeia nestes dois sentidos estaremos usando mais memória do computador. Um outro motivo, deste processamento ocorrer mais lentamente, é que no algoritmo CFTP-monótono reusamos a seqüência de uniformes e a seqüência de sítios geradas na iteração anterior. Enquanto que no algoritmo de Fill temos que, a cada iteração, gerar novas seqüências para as escolhas dos sítios e novas seqüências de variáveis aleatórias uniformes.

Apesar do processamento do algoritmo de Fill ser mais lento, o número máximo de transições apresentados na última iteração por ambas as técnicas foram equivalentes (mas não necessariamente isto deveria acontecer), entretanto o número de sítios visitados na última iteração foram os mesmos para ambos os algoritmos e para os mesmos valores de  $\beta$  e  $\Lambda$ . Apesar do tempo necessário para processar uma amostra ser diferente, as amostras apresentadas pelos dois algoritmos para os mesmos valores de  $\beta$  e  $\Lambda$  são muito parecidas. Isto nos permite escolher qualquer uma das técnicas quando formos fazer simulação de cadeia de Markov, pelo menos para o modelo de Ising.

## APÊNDICE A

## Apêndice

### A.1 Prova da Proposição 2.7

Considerando a cadeia de Markov X com espaço de estados E, matriz de transição P e regra de transição f dados pela proposição 2.7. Para provarmos esta proposição temos que provar que para um número de transições n, n par, temos

$$\mathbb{P}(X_n = 0 \mid C_n) = \frac{\frac{2}{3}(2^{n-1} - 2) + 1}{2^n - 2}, 
\mathbb{P}(X_n = 1 \mid C_n) = \frac{\frac{2}{3}(2^{n-1} - 2)}{2^n - 2}, 
\mathbb{P}(X_n = 2 \mid C_n) = \frac{\frac{2}{3}(2^{n-1} - 2) + 1}{2^n - 2},$$
(A.1)

e para um número de transições n, n ímpar, temos

$$\mathbb{P}(X_n = x \mid C_n) = \frac{1}{3}, \forall x \in E.$$
(A.2)

Observando as Tabelas 2.1, 2.2 e 2.3, é possível vermos que para qualquer número de transições sempre teremos 2 situações em que as trajetórias não coalesceram. Isto acontece quando temos seqüências alternadas de números aleatórios para qualquer número de transições, ou seja, quando tivermos seqüências da forma ( $\bar{c}c\bar{c}c\cdots$ ).

Assim

$$\mathbb{P}(\bar{C}_n) = \frac{2}{2^n} , \forall n \in \mathbb{N}, \ n > 1,$$
(A.3)

ou seja,

$$\mathbb{P}(C_n) = \frac{2^n - 2}{2^n} , \forall n \in \mathbb{N}, \ n > 1.$$
(A.4)

Para um número de transições n, n par, temos o seguinte comportamento para as probabilidades conjuntas  $\mathbb{P}(X_n = x, C_n), x \in E$ ,

$$\mathbb{P}(X_n = 0, \ C_n) = \frac{\frac{2}{3}(2^{n-1} - 2) + 1}{2^n},$$
  

$$\mathbb{P}(X_n = 1, \ C_n) = \frac{\frac{2}{3}(2^{n-1} - 2)}{2^n},$$
  

$$\mathbb{P}(X_n = 2, \ C_n) = \frac{\frac{2}{3}(2^{n-1} - 2) + 1}{2^n},$$
  
(A.5)

Já para um número de transições n, n ímpar, temos o seguinte comportamento

$$\mathbb{P}(X_n = x, \ C_n) = \frac{\frac{1}{3}(2^n - 2)}{2^n}, \forall x \in E.$$
(A.6)

Provamos por indução finita que  $\mathbb{P}(X_n = x, C_n), \forall x \in E$ , satisfaz as equações (A.5) e (A.6) para todo n, e conseqüentemente estamos provando as equações (A.1) e (A.2).

Primeiro verificamos que as equações (A.5) e (A.6) são válidas para n = 2 e 3 transições respectivamente,

$$\mathbb{P}(X_2 = 0, C_2) = \frac{\frac{2}{3}(2^{2-1} - 2) + 1}{2^2} = \frac{1}{4},$$
  

$$\mathbb{P}(X_2 = 1, C_2) = \frac{\frac{2}{3}(2^{2-1} - 2)}{2^2} = 0,$$
  

$$\mathbb{P}(X_2 = 2, C_2) = \frac{\frac{2}{3}(2^{2-1} - 2) + 1}{2^2} = \frac{1}{4},$$
  

$$\mathbb{P}(X_3 = x, C_3) = \frac{\frac{1}{3}(2^3 - 2)}{2^3} = \frac{2}{8}, \forall x \in E = \{0, 1, 2\},$$

logo temos que as equações são válidas para n = 2 e 3 transições respectivamente (basta verificar os resultados apresentados nas Tabelas 2.1 e 2.2). Temos,  $\forall n = k + 1$  e para todo  $x \in E$ , que

$$\mathbb{P}(X_{k+1} = x, C_{k+1}) = \mathbb{P}(X_{k+1} = x, C_{k+1} | X_k = 0, C_k) \cdot \mathbb{P}(X_k = 0, C_k) + \mathbb{P}(X_{k+1} = x, C_{k+1} | X_k = 1, C_k) \cdot \mathbb{P}(X_k = 1, C_k) + \mathbb{P}(X_{k+1} = x, C_{k+1} | X_k = 2, C_k) \cdot \mathbb{P}(X_k = 2, C_k) + \mathbb{P}(X_{k+1} = x, C_{k+1} | \bar{C}_k) \cdot \mathbb{P}(\bar{C}_k) .$$
(A.7)

Suponha que a equação (A.6) seja válida para n = k, k ímpar, então

$$\mathbb{P}(X_k = x, C_k) = \frac{\frac{1}{3}(2^k - 2)}{2^k}, \forall x \in E.$$

Vamos provar que é válido para n = k + 1, k + 1 par. De acordo com a equação (A.7) quando x = 0 teremos,

$$\mathbb{P}(X_{k+1} = 0, \ C_{k+1}) = \frac{1}{2} \cdot \frac{\frac{1}{3}(2^k - 2)}{2^k} + \frac{1}{2} \cdot \frac{\frac{1}{3}(2^k - 2)}{2^k} + 0 + \frac{1}{2} \cdot \frac{1}{2^k}$$
(A.8)  
$$= \frac{\frac{2}{3}(2^k - 2) + 1}{2^{k+1}} ,$$

logo, pela suposição n = k, k ímpar, a equação (A.6) é válida para x = 0. Explicamos a seguir o último termo de (A.8), ou seja,

$$\mathbb{P}(X_{k+1} = 0, \ C_{k+1} | \bar{C}_k) \cdot \mathbb{P}(\bar{C}_k) = \frac{1}{2} \cdot \frac{1}{2^k} .$$
(A.9)

Para k transições, k ímpar, os dois casos em que as trajetórias não coalesceram estão associadas às seqüências de números aleatórios  $\underline{ccc} \cdots \underline{cc}$  e  $\underline{ccc} \cdots \underline{cc}$ . Entretanto para que ocorra a coalescência no estado 0 em k + 1 transições, a partir das seqüências apresentadas acima, será necessário que se tenha a seqüência  $\underline{ccc} \cdots \underline{cc}$  em k transições e  $\overline{c}$  na transição k+1, ou seja, é necessário que se tenha a seqüência  $\underline{ccc} \cdots \underline{cc}_{k+1}$ em k+1 transições. Já a partir da seqüência  $\underline{ccc} \cdots \underline{cc}_{k}$  é possível obter a coalescência apenas no estado 2 para k+1 transições. No restante da prova aparecerá novamente a equação (A.9) e a justificativa é análoga a esta. De acordo com a equação (A.7) quando x = 1 teremos,

$$\mathbb{P}(X_{k+1} = 1, \ C_{k+1}) = \frac{1}{2} \cdot \frac{\frac{1}{3}(2^k - 2)}{2^k} + 0 + \frac{1}{2} \cdot \frac{\frac{1}{3}(2^k - 2)}{2^k} + 0$$
$$= \frac{\frac{2}{3}(2^k - 2)}{2^{k+1}},$$

logo, pela suposição n = k, k ímpar, a equação (A.6) é válida quando x = 1. A prova da equação (A.6) quando x = 2 é análoga a que apresentamos quando x = 0.

Suponha agora que  $\mathbb{P}(X_n = x, C_n)$  seja válido para n = k, k par, então

$$\mathbb{P}(X_k = 0, \ C_k) = \frac{\frac{2}{3}(2^{k-1} - 2) + 1}{2^k},$$
$$\mathbb{P}(X_k = 1, \ C_k) = \frac{\frac{2}{3}(2^{k-1} - 2)}{2^k} e$$
$$\mathbb{P}(X_k = 2, \ C_k) = \frac{\frac{2}{3}(2^{k-1} - 2) + 1}{2^k}.$$

Vamos provar que é válido para n = k + 1, k + 1 ímpar. Considerando x = 0 na equação (A.7), teremos

$$\mathbb{P}(X_{k+1} = 0, \ C_{k+1}) = \frac{1}{2} \cdot \frac{\frac{2}{3}(2^{k-1} - 2) + 1}{2^k} + \frac{1}{2} \cdot \frac{\frac{2}{3}(2^{k-1} - 2)}{2^k} + 0 + \frac{1}{2} \cdot \frac{1}{2^k}$$
$$= \frac{\frac{4}{3}(2^{k-1} - 2) + 2}{2^{k+1}}$$
$$= \frac{\frac{2^{k+1} - 8 + 6}{3}}{2^{k+1}}$$
$$= \frac{\frac{1}{3}(2^{k+1} - 2)}{2^{k+1}}.$$

logo, pela suposição n = k, k par, a equação (A.5) é válida quando x = 0.

Considere agora x = 1 na equação (A.7), assim

$$\mathbb{P}(X_{k+1} = 1, \ C_{k+1}) = \frac{1}{2} \cdot \frac{\frac{2}{3}(2^{k-1} - 2) + 1}{2^k} + 0 + \frac{1}{2} \cdot \frac{\frac{2}{3}(2^{k-1} - 2) + 1}{2^k} + 0$$
$$= \frac{\frac{4}{3}(2^{k-1} - 2) + 2}{2^{k+1}}$$
$$= \frac{\frac{1}{3}(2^{k+1} - 2)}{2^{k+1}}.$$

logo, pela suposição n = k, k par, a equação (A.5) é válida quando x = 1. A prova da equação (A.5) quando x = 2 é análoga ao caso x = 0. Portanto, provamos as equações (A.1) e (A.2) e conseqüentemente a Proposição 2.7.

## A.2 Prova da equação (3.3)

No método de rejeição geramos x de acordo com  $\widetilde{P}^t(\hat{0}, \cdot)$  e aceitamos x como amostra de  $\pi(x)$  com probabilidade apropriada. Provamos que esta probabilidade é

$$C^{-1} \cdot \frac{\pi(x)}{\widetilde{P}^{t}(\hat{0}, x)} = \frac{P^{t}(\hat{1}, \hat{0})}{P^{t}(x, \hat{0})}.$$
(A.10)

Temos da definição da matriz de transição para cadeia em tempo reverso que

$$\frac{\pi(x)}{\widetilde{P}^{t}(\hat{0},x)} = \frac{\pi(\hat{0})}{P^{t}(x,\hat{0})}$$
(A.11)

e como  $C \ge \frac{\pi(x)}{\tilde{P}^{t}(\hat{0},x)}$  (pelo método de rejeição),  $\forall x \in E$ , podemos escolher C tal que

$$C = \sup_{x \in E} \frac{\pi(x)}{\widetilde{P}^{t}(\hat{0}, x)} = \sup_{x \in E} \frac{\pi(\hat{0})}{P^{t}(x, \hat{0})} = \frac{\pi(\hat{0})}{\inf_{x \in E} P^{t}(x, \hat{0})} = \frac{\pi(\hat{0})}{P^{t}(\hat{1}, \hat{0})}.$$
 (A.12)

Substituindo C na equação (A.10) aceitamos x com a probabilidade abaixo:

$$C^{-1} \cdot \frac{\pi(x)}{\widetilde{P}^{t}(\hat{0}, x)} = \frac{P^{t}(\hat{1}, \hat{0})}{\pi(\hat{0})} \cdot \frac{\pi(x)}{\widetilde{P}^{t}(\hat{0}, x)} = \frac{P^{t}(\hat{1}, \hat{0})}{\pi(\hat{0})} \cdot \frac{\pi(\hat{0})}{P^{t}(x, \hat{0})} = \frac{P^{t}(\hat{1}, \hat{0})}{P^{t}(x, \hat{0})}.$$
 (A.13)

Em (A.12) acima a segunda igualdade ocorre pela equação (A.11), a terceira igualdade é trivial, já a quarta explicamos abaixo.

Considere a notação  $(\hat{1} \longrightarrow^t \hat{0})$  significando que a cadeia vai do estado  $\hat{1}$  para  $\hat{0}$  em t passos. Então, pela monotonicidade de P, se para algum t

$$(\hat{1} \longrightarrow^t \hat{0})$$
 obviamente teremos que  $(x \longrightarrow^t \hat{0}), \forall x \in E,$ 

logo

 $(\hat{1} \longrightarrow^t \hat{0}) \subset (x \longrightarrow^t \hat{0}), \ \forall x \in E,$ 

assim

$$\mathbb{P}(\hat{1} \longrightarrow^{t} \hat{0}) \leq \mathbb{P}(x \longrightarrow^{t} \hat{0}) , \text{ isto \acute{e}, } P^{t}(\hat{1}, \hat{0}) \leq P^{t}(x, \hat{0}), \quad \forall \ x \in E,$$

ou seja,

$$\inf_{x \in E} P^t(x, \hat{0}) = P^t(\hat{1}, \hat{0}).$$

## A.3 Prova do resultado (4.19)

Usamos aqui a função de energia dada por (4.2) sem a presença de campo externo, possuindo condição de fronteira livre e com J = 1.

$$\begin{aligned} \pi(x_t(j) &= +1 | x_t(-j)) = \\ &= \frac{\pi(x_t(j) = +1, x_t(-j))}{\pi(x_t(j) = +1, x_t(-j)) + \pi(x_t(j) = -1, x_t(-j))} \\ &= \left(1 + \frac{\pi(x_t(j) = -1, x_t(-j))}{\pi(x_t(j) = 1, x_t(-j))}\right)^{-1} \\ &= \left(1 + \frac{\frac{1}{Z} \exp \beta \left(\sum_{\substack{m \in \Lambda \\ m \neq j}} \sum_{\substack{l \in \Lambda \\ m \neq j}} x_t(m) \cdot x_t(l) + \sum_{\substack{l \in \Lambda \\ l \sim j}} x_t(j) \cdot x_t(l)\right)}{\frac{1}{Z} \exp \beta \left(\sum_{\substack{m \in \Lambda \\ m \neq j}} \sum_{\substack{l \in \Lambda \\ m \neq j}} x_t(m) \cdot x_t(l) + \sum_{\substack{l \in \Lambda \\ l \sim j}} x_t(j) \cdot x_t(l)\right)}\right)^{-1} \\ &= \left(1 + \frac{\exp \left(\beta \sum_{\substack{l \in \Lambda \\ l \sim j}} 1 \cdot x_t(l)\right)}{\exp \left(\beta \sum_{\substack{l \in \Lambda \\ l \sim j}} 1 \cdot x_t(l)\right)}\right)^{-1}. \end{aligned}$$

## **Referências Bibliográficas**

- [Fill, 1998] Fill, J. A. (1998). An interruptible algorithm for perfect sampling via Markov chains. Annals of Applied Probability, 8:131-162.
- [Fill et al., 2000] Fill, J. A., Machida, M., Murdoch, D. J., and Rosenthal, J. S. (2000). Extension of Fill's perfect rejection sampling algorithm to general chains (extended abstract). Pages 37-52 in *Monte Carlo Methods* (ed.: N. Madras), *Fields Institute Communications* 26, American Mathematical Society.
- [Fismen, 1997] Fismen, M. (1997). Exact simulation using Markov chains. Master's thesis, Department for Mathematical Sciences, The Norwegian University.
- [Foss et al., 1998] Foss, S. G., Tweedie, R. L., and Corcoran, J. (1998). Simulating the invariant measures of Markov chain using backward coupling at regeneration times. Prob. Eng. Inf. Sci., 12:303-320.
- [Galves and Ferrari, 1997] Galves, A. and Ferrari, P. (1997). Acoplamento e Processos Estocásticos. IMPA, 21<sup>o</sup> Colóquio Brasileiro de Matemática.
- [Gamerman, 1997] Gamerman, D. (1997). Stochastic Simulation for Bayesian Inference. Chapman and Hall, London, First edition.
- [Geman and Geman, 1984] Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. Pattn.* Anal. Mach. Intel., 6:721-741.

- [Georgii, 1988] Georgii, H.-O. (1988). *Gibbs Measures and Phase Transitions*. Walter de Gruyter, New York.
- [Gilks et al., 1996] Gilks, W. R., Richardson, S., and Spiegelhalter, D. J. (1996). Markov Chain Monte Carlo in Practice. Chapman and Hall, First edition.
- [Häggström, 2001] Häggström, O. (2001). Finite Markov chains and algoritmic applications. Notas de aula.
- [Häggström and Nelander, 1998] Häggström, O. and Nelander, K. (1998). Exact sampling from anti-monotone systems. *Statistica Neerlandica*, 52:360-380.
- [Hastings, 1970] Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57:97-109.
- [Holley, 1974] Holley, R. (1974). Recent results on the stochastic Ising model. Journal of Mathematics, 4(3):479-496.
- [Huber, 1999] Huber, M. (1999). Perfect sampling without a lifetime commitment. Preprint.
- [Kindermann and Snell, 1980] Kindermann, R. and Snell, J. L. (1980). Markov random fields and their applications. American Mathematical Society, Providence, RI.
- [Lindvall, 1992] Lindvall, T. (1992). Lectures on the Coupling Method. John Wiley and Sons, New York.
- [Moller and Schladitz, 1999] Moller, J. and Schladitz, K. (1999). Extensions of Fill algorithm for perfect simulation. Journal of the Royal Statistical Society, Series B, 61.
- [Murdoch and Green, 1998] Murdoch, D. J. and Green, P. J. (1998). Exact sampling from a continuous state space. *Scandinavian Journal or Statistics*, 25(3):483-502.
- [Murdoch and Rosenthal, 1998] Murdoch, D. J. and Rosenthal, J. S. (1998). An extension of Fill's exact sampling algorithm to non-monotone chains. Preprint.

- [Propp and Wilson, 1996] Propp, J. G. and Wilson, D. B. (1996). Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures and Algorithms*, 9:223-252.
- [Ross, 1994] Ross, S. M. (1994). A First Course in Probability. Macmillan College Publishing Company, New York, Fourth edition.
- [Thönnes, 1999a] Thönnes, E. (1999a). Perfect simulation of some point processes for the impatient user. Adv. Appl. Prob., 31:69-87.
- [Thönnes, 1999b] Thönnes, E. (1999b). A primer on perfect simulation. Technical report, Department of Mathematical Statistics, Chalmers University of Technology.
- [Thorisson, 2000] Thorisson, H. (2000). Coupling, stationarity, and regeneration (Probability and its Applications). Springer-Velarg, New York.
- [Wilson, 2000] Wilson, D. B. (2000). How to couple from the past using a read-once source of randomeness. *Random Structures and Algorithms*, 16:85-113.