**Reconstrução de cenários em
 Z^d seguindo um passeio
aleatório com ramificação**

Angélica Yohana Pachón Pinzón

DISSERTAÇÃO APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA OBTENÇÃO DO TÍTULO
DE
MESTRE EM ESTATÍSTICA

Área de Concentração: Probabilidade

Orientador: Prof. Dr. Serguei Popov

Apoio financeiro do CNPq

- São Paulo, Fevereiro de 2006 -

Reconstrução de cenários em Z^d seguindo um passeio aleatório com ramificação

Este exemplar corresponde à redação final
da dissertação devidamente corrigida e
defendida por Angélica Yohana Pachón Pinzón
e aprovada pela comissão julgadora.

São Paulo, 16 de Fevereiro de 2006.

Banca Examinadora:

Prof. Dr. Serguei Popov (Presidente) - IME - USP

Prof. Dra. Marina Vachkovskaia - IMECC - UNICAMP

Prof. Dr. Fabio Machado - IME - USP

Dedicatória.

Dedicatória.

Dedicatória

Dedicatória.

*A Camilo, meu irmão,
hoje um milagre que me da
a força para lutar nesses
momentos em que perco o
alento e nada tem sentido.*

Irmãozinho, obrigada por existir.

Agradecimentos

Agradeço em primeiro lugar a Deus por ter-me escutado e dado uma nova vida, a meus pais por ter-me feito a mulher que sou hoje, e a meu orientador Prof. Serguei Popov pela guia, a dedicação e acima de tudo por seu exemplo de alguém que é feliz fazendo probabilidade.

Sumário

No presente trabalho nós estudamos o problema de reconstrução de cenários em d dimensões usando observações registradas em caixas de tamanho k (para k fixo), vistas ao longo de um passeio aleatório com ramificação. Nós provamos que, para quase todos os passeios aleatórios, quase todos os cenários podem ser reconstruídos. A reconstrução trabalha modulo equivalência.

Abstract

In this work we study the problem of scenery reconstruction in d dimensions, using observations registered in boxes of size k (for k fixed), seen throughout a random walk with ramification. We prove that, for almost all the random walks, almost all the scenarios can be reconstructed. The reconstruction will work only up to equivalence.

Índice

Introdução	1
1 História	4
2 Descrição do problema	7
2.1 Definição do cenário	7
2.2 Definição do processo	8
3 Demonstração do teorema	13
3.1 Existência de um aglomerado infinito de pontos bons	14
3.2 Recorrência de passeios aleatórios com ramificação restrito ao aglomerado infinito de pontos bons	19
3.3 Reconstrução do cenário nos pontos bons	21
3.4 Reconstrução do cenário fora do aglomerado de pontos bons .	24

Introdução

Reconstrução de cenários pertence à área geral da teoria da probabilidade que trata de propriedades ergódicas das observações produzidas por um processo estocástico em um meio aleatório, mais também pode ser considerado interessante por se mesmo. Considere um grafo (V, E) com cores em seus vértices dispostos de maneira arbitrária. Esta coração será chamada um cenário sobre (V, E) , então, um passeio aleatório começa se mover sobre (V, E) registrando a cor onde ele se encontrar mais não sua localização, assim o que se observa é uma seqüência de cores vistas pelo passeio em cada vértice. A questão é: é possível dizer algo de como V foi colorido? O problema, foi investigado independentemente por Benjamini, den Hollander e Keane [5] e de ele se originam basicamente três ramos:

1. Podem dois cenários (conhecidos) ser distinguíveis com somente uma seqüência observada de cada um?
2. É possível reconstruir cenários (desconhecidos) com as observações obtidas por um passeio aleatório sobre V ?
3. Têm cenários que não podem ser reconstruídos ou distinguidos pelo registro de cores do passeio aleatório?

Reconstrução de cenários corresponde à segunda questão, se temos apenas as observações vistas ao longo do passeio aleatório podemos reconstruir

a coloração inicial do espaço? A resposta em termos gerais é “não”. Entretanto, sob restrições apropriadas a resposta torna-se “sim” em $\mathbb{Z}$ e $\mathbb{Z}^2$. Vamos explicar estas restrições: primeiro, se dois cenários ξ e ξ^* são equivalentes, (dizemos que dois cenários são equivalentes se um deles é obtido a partir do outro por translações ou reflexões), em geral, não conseguimos distinguir se as observações estão vindo de ξ ou ξ^* . Assim, só é possível conseguir reconstruir ξ módulo equivalência. Segundo, a reconstrução funciona no melhor dos casos *quase certamente*(q.c).

É importante notar, embora, que Lindenstrauss em [10] dá exemplos de cenários que não podem ser reconstruídos, a maioria dos cenários podem ser reconstruídos módulo equivalência.

Os métodos usados para reconstrução de cenários mudam dependendo do número das cores no cenário. (Exceto por [14], todos os trabalhos de reconstrução de cenários assumem que as cores do cenário são i.i.d.)

Kesten observou que os métodos conhecidos em reconstrução de cenários também parecem depender da unidimensionalidade do problema, então ele perguntou se uma reconstrução ainda é possível em duas ou mais dimensões. Löwe e Matzinger em [11] respondem a esta pergunta afirmativamente no caso de duas dimensões sob a condição que o número das cores no cenário seja suficientemente grande. O problema da reconstrução em dimensões maiores torna-se impossível considerando um passeio aleatório simples. Se a dimensão é pelo menos 3 o processo é transiente, logo não vamos ver o cenário em muitos pontos.

Neste trabalho estamos interessados no problema em maiores dimensões mas considerando um passeio aleatório com ramificação, usando observações registradas em caixas ou janelas de tamanho k centradas no vértice onde o passeio se encontrar (para k fixo), vistas ao longo de uma realização do processo. Assim, a pergunta que nos fazemos é a seguinte: é possível reconstruir o cenário q.c. em d dimensões dado somente uma seqüência de janelas

observadas obtidas pelo passeio aleatório com ramificação?

Nós conseguimos responder positivamente esta pergunta se o tamanho das janelas é suficientemente grande. Vamos usar uma idéia nova para construir o algoritmo empregando teoria de percolação em sítios. A reconstrução funciona módulo equivalência.

A organização da dissertação é como segue. O próximo capítulo será dedicado à história da reconstrução de cenários. No Capítulo 2 descrevemos o problema e apresentamos o teorema que estabelece a reconstrução de cenários sobre Z^d usando um passeio aleatório com ramificação e observações registradas em caixas de tamanho k . Finalmente, no Capítulo 3 demonstramos o teorema.

Capítulo 1

História

Reconstrução de cenários está fortemente relacionado ao problema de cenários distinguíveis, o qual surgiu primeiro. O interesse no problema de cenários distinguíveis surgiu quando Keane e den Hollander, e também Benjamini, se perguntaram se cada par de cenários não-equivalentes poderiam ser distinguidos. Lindenstrauss [10] conseguiu provar que existem pares de cenários os quais não podem ser distinguidos. Vamos dar um breve resumo. Sejam ξ^a e ξ^b dois cenários não equivalentes os quais conhecemos. Assuma que o cenário ξ é igual a um dos dois ξ^a ou ξ^b mas não sabemos a qual. Se temos apenas uma realização do processo observado χ do cenário ξ vista ao longo de um passeio aleatório S , podemos determinar quase certamente se ξ é igual a ξ^a ou igual a ξ^b ? Se a resposta fosse positiva, nós dizemos que os cenários ξ^a e ξ^b são distinguíveis.

Benjamini e Kesten [3] mostram que quase todos os pares de cenários são distinguíveis, mesmo no caso de duas dimensões com somente duas cores. Mais recentemente, Howard [4] mostrou que qualquer par de cenários periódicos não equivalentes e cenários periódicos com um simples defeito são distinguíveis.

O problema de distinguir dois cenários os quais diferem em somente um

ponto é chamado detecção de um único defeito num cenário. Kesten [6] conseguiu mostrar que q.c. é possível detectar um único defeito no caso de cenários com quatro cores. Uma pergunta que Kesten fez sobre a detecção de defeitos em cenários permitiu a Matzinger [15, 16, 17] começar com a pesquisa do problema de reconstrução de cenários.

O primeiro resultado positivo no problema da reconstrução de cenários foi encontrado por Matzinger em sua tese de doutorado [15]. O orientador Harry Kesten, notou que o resultado [15] era inapropriado para o problema de reconstrução em duas dimensões, como também quando os passeios aleatórios permitem saltos. Esta observação provocou intensa pesquisa ao respeito. Durante os seguintes três anos no instituto Eurandom em Eindhoven, Jyri Lember, Mathias Löwe, Heinrich Matzinger, Franz Merkl and Silke Rolles gastaram muito tempo e energia reconstruindo cenários. Seguidamente, Löwe e Matzinger [11] provaram que a reconstrução de cenários é também possível em cenários em duas dimensões com muitas cores. Mais tarde, Löwe, Matzinger e Merkl [13] provaram que com suficientes cores em uma dimensão, é possível reconstruir mesmo se o passeio aleatório tem saltos e assim não é um passeio aleatório simples. Em geral, a reconstrução de cenários é mais difícil quando o número de cores diminui (exceto no caso trivial quando há somente uma cor). O caso mais difícil de reconstrução considerando um passeio aleatório com saltos sob cenários com duas cores, foi resolvido por Lember e Matzinger [8]. den Hollander perguntou se seria possível fazer a reconstrução se os saltos fossem não limitados. Lenstra e Matzinger [9] respondem positivamente esta pergunta. Seguindo também a questão de den Hollander, Löwe e Matzinger [12] perguntam pela possibilidade de reconstruir cenários correlacionados.

Finalmente, reconstrução de cenários chega ao Chile onde aparece uma versão do problema de cenários distinguíveis com observações defeituosas. Mais uma vez o cenário ξ é igual a ξ^a ou ξ^b , os quais são conhecidos, en-

tretanto, as observações são agora defeituosas através de um processo de erro $\{\nu_t\}_{t \geq 0}$, o qual assume-se uma sequência de variáveis aleatórias Bernoulli i.i.d. com parâmetro estritamente menor que $1/2$ e independentes de ξ e S . Matzinger e Andrew Hart [1] do Núcleo Milênio de Servet Martinez trabalharam sobre marcas para observações com erros.

No momento Matzinger e Popov [18] trabalham sobre cenários contínuos.

Capítulo 2

Descrição do problema

Este problema trata-se a reconstrução de cenários em $\mathbb{Z}^d$ usando as observações registradas em caixas de tamanho k (para k fixo) vistas ao longo de um passeio aleatório com ramificação.

Começamos a nossa descrição por definir um cenário em $\mathbb{Z}^d$ e depois o que é um passeio aleatório com ramificação.

2.1 Definição do cenário

Um cenário em $\mathbb{Z}^d$ é uma coloração ξ do espaço com C_0 cores, $\{0, 1, \dots, C_0 - 1\}$, localizadas em cada vértice do reticulado $\mathbb{Z}^d$. Essa coloração é obtida de maneira independente e identicamente distribuída. Formalmente, um cenário é uma função $\xi : \mathbb{Z}^d \rightarrow \{0, 1, \dots, C_0 - 1\}$.

Dizemos que dois cenários ξ e ξ^* são equivalentes (o que denotaremos por $\xi \sim \xi^*$), se existe um $a \in \mathbb{Z}^d$ tal que $\xi(x) = \xi^*(x + a)$ para todo ponto x em $\mathbb{Z}^d$. Em outras palavras, ξ e ξ^* são equivalentes se eles podem ser obtidos um do outro por translações com relação a cada eixo.

Consideraremos o cenário com duas cores, $\{0, 1\}$, localizados em cada

vértice de $\mathbb{Z}^d$ com probabilidades $(1 - p)$ e p respectivamente, isto é,

$$\xi : \mathbb{Z}^d \rightarrow \{0, 1\}.$$

2.2 Definição do processo

O processo que consideramos é um passeio aleatório com ramificação que pode ser descrito da seguinte maneira. O processo começa com uma única partícula na origem de $\mathbb{Z}^d$ que forma a geração inicial. No tempo 1 com probabilidade b , a partícula escolhe se dividir em duas formando a geração 1 de tamanho N_1 , $N_1 = 1 + X_1^1$, onde $P(X_1^1 = 0) = 1 - P(X_1^1 = 1) = 1 - b$. Além disso, independentemente cada partícula da primeira geração decide saltar com probabilidade uniforme para qualquer um dos $2d$ vértices vizinhos. O processo continua com as partículas da nova geração, então, no tempo n , cada partícula da geração $n - 1$ decide se dividir em duas com probabilidade b , formando a geração n de tamanho N_n , $N_n = N_{n-1} + X_1^n, \dots, + X_{N_{n-1}}^n$, em que $P(X_i^n = 0) = 1 - P(X_i^n = 1) = 1 - b$, $1 \leq i \leq N_{n-1}$. Imediatamente depois cada uma destas novas partículas decide com probabilidade $\frac{1}{2d}$ a direção do salto.

É importante esclarecer que as partículas de cada geração atuam independentemente uma da outra.

Antes de enunciar o teorema que queremos provar é preciso introduzir mais algumas notações.

Seja $\eta_n(z) :=$ o número de partículas em z no momento n , $z \in \mathbb{Z}^d$ e $n \geq 0$, com $\eta_0(z) = 1\{z = 0\}$. Denotaremos por $\eta_n = (\eta_n(z), z \in \mathbb{Z}^d)$ o passeio aleatório com ramificação em $\mathbb{Z}^d$ com probabilidade de ramificação b começando na origem até o momento n como foi descrito acima, por $\Omega_1 = \{(\eta_n)_{n \in \mathbb{N}}\}$ o espaço de todos os passeios aleatórios com ramificação, por $\Omega_2 = \{0, 1\}^{\mathbb{Z}^d}$ o espaço de todos os possíveis cenários ξ , por $G_n = \{v_1^n, \dots, v_{N_n}^n\}$ o conjunto de vértices (partículas) da n -ésima geração e por $G = \cup_{n \in \mathbb{N}} G_n$

Figura 2.1: Árvore genealógica até $n = 4$

a árvore genealógica obtida pelo processo de ramificação. Um desenho da árvore genealógica pode ser visto na Figura 2.1.

Seja $\Psi : G \rightarrow \mathbb{Z}^d$ uma função tal que $\Psi(v_j^i) = z$, quer dizer, ela dá a posição da partícula v_j^i em $\mathbb{Z}^d$ depois do salto, e $\mathfrak{C}^k(z)$ o cenário contido no hipercubo d -dimensional de tamanho k com centro em z , isto é, o cenário contido na caixa $z + [-k, k]^d$. A matriz d -dimensional $\mathfrak{C}^k(z)$ também será chamada uma configuração de tamanho k centrada em z .

Assim, de acordo com o processo de ramificação, o cenário observado será a árvore com janelas $\chi = \cup_{n \in \mathbb{N}} \chi_n$, em que $\chi_n = [\mathfrak{C}^k(\Psi(v_j^i)), v_j^i \in G_n]$, e definimos $\Omega_3 = \{(\chi_n)_{n \in \mathbb{N}}\}$ o espaço de todos os cenários observados. Um exemplo do processo η_n o qual produz uma árvore com janelas pode ser visto na Figura 2.2 e na Figura 2.3, respectivamente.

Assumimos que $(\eta_n)_{n \in \mathbb{N}}$ e ξ são independentes e distribuídos de acordo às leis P_1 e P_2 .

Formalmente, o que queremos é provar o seguinte teorema.

Figura 2.3: Árvore com janelas obtida do processo na Figura 2.2

Teorema 1 *Existe uma função mensurável $\Lambda : \Omega_3 \rightarrow \Omega_2$ tal que $P(\Lambda(\chi) \sim \xi) = 1$.*

Aqui a medida P se define no espaço $\Omega_1 * \Omega_2$, logo P é a medida produto $P_1 * P_2$.

Assim vamos provar que para quase todos os passeios aleatórios, quase todos os cenários podem ser reconstruídos. A reconstrução se dá módulo equivalência.

Capítulo 3

Demonstração do teorema

Definição 1 *Sejam $\vec{e}_i$ os vetores canônicos d -dimensionais, $i = 1, \dots, d$, e $\mathfrak{C}_{\vec{e}}^k(z)$ para $\vec{e} \in \{\pm\vec{e}_i, i = 1, \dots, d\}$ o cenário contido na caixa $z + [-k, k]^d \cap (z + \vec{e}) + [-k, k]^d$, isto é, a caixa $z + [-k, k]^d$ translada na direção de $\vec{e}$, mas eliminando a última fase nesta direção.*

Dizemos que $\mathfrak{C}^k(z)$ é uma boa configuração se todas as translações $\{\mathfrak{C}_{\vec{e}}^k(z), \vec{e} \in \{\pm\vec{e}_i, i = 1, \dots, d\}\}$ são diferentes, e chamaremos um ponto z de bom se ele é o centro de uma boa configuração.

Lema 1

$$\lim_{k \rightarrow \infty} P_2(\text{configuração de tamanho } k \text{ centrada na origem é boa}) = 1 \quad (3.1)$$

Demonstração: Definimos os eventos:

$$\begin{aligned} A_{mn}^k &= \{\mathfrak{C}_{\vec{e}_m}^k(0) \neq \mathfrak{C}_{\vec{e}_n}^k(0); (\vec{e}_m \neq \vec{e}_n) \in \{\pm\vec{e}_i, i = 1, \dots, d\}, 0 = \Psi(v_1^0), v_1^0 \in G\} \\ A^k &= \cap_{m,n} \{A_{m,n}^k\}, \end{aligned}$$

em palavras podemos escrever estes eventos como,

$$\begin{aligned} A_{mn}^k &:= \{\text{as translações } \mathfrak{C}_{\vec{e}_m}^k(0) \text{ e } \mathfrak{C}_{\vec{e}_n}^k(0) \text{ são diferentes}\}, \text{ e} \\ A^k &:= \{\text{Todas as translações são diferentes}\}. \end{aligned}$$

Para que uma configuração seja boa é preciso que todas as translações sejam diferentes, logo provar (3.1) é equivalente a provar que $\lim_{k \rightarrow \infty} P(A^k) = 1$. Neste sentido calculamos

$$\begin{aligned} P(A^k) &= P(\cap_{m,n} A_{mn}^k) \\ &= 1 - P(\cup_{m,n} (A_{mn}^k)^c) \\ &\geq 1 - \sum_{m,n} P((A_{mn}^k)^c) \end{aligned}$$

e como

$$\begin{aligned} P((A_{mn}^k)^c) &= P(\cap \{c_m = c_n\}) \\ &= \frac{1}{2^{(2k)^d}}, \end{aligned}$$

a intersecção é sob todos os pares de cores c_m e c_n localizados nos mesmos lugares das configurações $\mathfrak{C}_m^k(0)$ e $\mathfrak{C}_n^k(0)$ respectivamente, (ver Figura 3.1). Então

$$P(A^k) \geq 1 - \frac{\binom{2d}{2}}{2^{(2k)^d}},$$

portanto

$$\lim_{k \rightarrow \infty} P(A^k) = 1.$$

□

3.1 Existência de um aglomerado infinito de pontos bons

Lema 2 *Para um k suficientemente grande existe percolação de pontos bons, isto é:*

$$P(\text{Existir um aglomerado infinito de pontos bons sob } k, \text{ para } k \text{ grande}) > 0. \quad (3.2)$$

Figura 3.1: Translações para cima e para direita no caso $d = 2$, $z = (5, 5)$ e $k = 4$. Os quadros com $\{0, 1\}$ são as configuração nas caixas $z + [-k, k]^d$, os de linha pontilhada são as configurações $\mathfrak{C}_1^k(0)$ e $\mathfrak{C}_2^k(0)$ nas caixas $(z + \vec{e}_1) + [-k, k]^d$ e $(z + \vec{e}_2) + [-k, k]^d$, respectivamente. Nas zonas mais claras estão os pares localizados nos mesmos lugares das configurações $\mathfrak{C}_1^k(0)$ e $\mathfrak{C}_2^k(0)$.

Demonstração: Primeiro vejamos que o evento “ x é bom no modelo com caixas de tamanho k' ”, é crescente em k' . Com efeito, supondo que ele é bom sob k^* , então para $k' > k^*$, cada uma das configurações $\mathfrak{C}_\varepsilon^{k^*}(x)$ estão contidas na respectiva configuração $\mathfrak{C}_\varepsilon^{k'}(x)$. Portanto, se todas são diferentes sob k^* também o serão sob k' , e assim, x também é bom sob k' . Isto é

$$\{x \text{ é bom no modelo com } k^*\} \subseteq \{x \text{ é bom no modelo com } k'\},$$

portanto

$$P(x \text{ é bom no modelo com } k^*) \leq P(x \text{ é bom no modelo com } k'),$$

desta maneira podemos escolher k' tal que $P(x \text{ seja bom sob } k') > p_c$, em que p_c é o ponto crítico para percolação por sítios em $\mathbb{Z}^d$.

Agora dividamos o espaço em caixas disjuntas de tamanho k' , assim, construímos um modelo independente de percolação em sítios, que são, as configurações correspondentes a estas caixas.

Como o evento “ x é bom sob k' ” é equivalente a “ x é o centro de uma boa configuração de tamanho k' ”, temos que

$$P(\mathfrak{C}^{k'}(x) \text{ é bom}) = P(x \text{ é bom sob } k') > p_c,$$

logo temos percolação de boas configurações de tamanho k' , isto é, com probabilidade positiva temos um aglomerado infinito de boas configurações de tamanho k' .

Embora, nos queremos percolação de pontos bons, portanto consideremos agora caixas de tamanho $k = 2k' + 1$. Então, para qualquer ponto y de $\mathbb{Z}^d$ o qual também está na caixa $x + [-k', k']^d$, em que $x + [-k', k']^d$ é uma das caixas da partição, temos que todas as translações da caixa $y + [-k, k]^d$ contêm a caixa $x + [-k', k']^d$.

Este fato implica que se a configuração dentro da caixa $x + [-k', k']^d$ é boa, então, todas as translações da configuração dentro da caixa $y + [-k, k]^d$ são diferentes. Logo, a configuração dentro da caixa $y + [-k, k]^d$, ou seja $\mathfrak{C}^k(y)$, é também boa e y um ponto bom ao considerar caixas de tamanho k .

Disto temos que tomando caixas de tamanho k um ponto é bom se ele está dentro de alguma boa configuração de tamanho k' , e como provamos que existe percolação de boas configurações de tamanho k' , então temos conseguido percolação de pontos bons no modelo sob k . Isto termina a demonstração do Lema 2. $\square$

Lema 3 *O passeio aleatório com ramificação η é recorrente em $\mathbb{Z}^d$, isto é, todos os sítios são visitados um número infinito de vezes com probabilidade 1.*

Demonstração: Seja S_n um passeio aleatório simples em $\mathbb{Z}^d$ começando na origem, $p(n, x) = P(S_n = x)$ e $\bar{p} = 2\left(\frac{d}{2n\pi}\right)^{\frac{d}{2}} e^{\frac{-d|x|^2}{2n}}$, com $|x|$ a norma do vetor x . Um resultado do teorema central do limite local (ver [7] Teorema 1)

mostra que

$$|E(n, x)| := |\bar{p}(n, x) - p(n, x)| \leq O(n^{\frac{-(d+2)}{2}}) \leq \frac{c}{n^{\frac{(d+2)}{2}}}, \quad (3.3)$$

em que c é uma constante que depende somente da dimensão d . Usando este fato obtemos uma cota inferior para a probabilidade da partícula estar na origem no tempo n , assim

$$|E(n, x)| = |\bar{p}(n, x) - p(n, x)| \geq \bar{p}(n, x) - p(n, x),$$

então

$$p(n, 0) \geq \bar{p}(n, 0) - |E(n, 0)| \geq 2\left(\frac{d}{2n\pi}\right)^{\frac{d}{2}} - \frac{c}{n^{\frac{(d+2)}{2}}} = \frac{c^*}{n^{\frac{d}{2}}}, \quad (3.4)$$

em que c^* só depende de d .

Agora, calculamos o número esperado de partículas no tempo n do processo η_n , isto é, $E(N_n)$. Para isto usamos que $E(N_n) = E(E(N_n/N_{n-1}))$. Assim, calculamos primeiro

$$\begin{aligned} E(N_n/N_{n-1} = i) &= E(i + X_1^n + \cdots + X_i^n) \\ &= i + bi \\ &= i(1 + b), \end{aligned}$$

logo, $E(N_n/N_{n-1}) = (1 + b)N_{n-1}$. Tomando valor esperado nos ambos os lados desta última igualdade temos que $E(N_n) = (1 + b)E(N_{n-1})$ com o que recursivamente chegamos a

$$E(N_n) = (1 + b)^n. \quad (3.5)$$

De (3.4) e (3.5) temos que

$$E(\# \text{ de partículas de } \eta_n \text{ na origem}) \geq (1 + b)^n \frac{c^*}{n^{\frac{d}{2}}}, \quad (3.6)$$

que é maior que 1 para n suficientemente grande. Vamos chamar o mínimo destes n de t_0 .

Agora vamos construir um processo de Galton-Watson [2] com esperança do número de filhos maior que 1. Este é construído usando (3.6) e portanto pode ser restrito a uma caixa de tamanho t_0 como segue: começamos com uma partícula na origem no tempo 0 e tomamos como primeira geração o número de partículas na origem no tempo t_0 , o que chamamos tamanho da primeira geração e denotamos por Z_1 . Para formar a segunda geração de tamanho Z_2 , de cada uma destas Z_1 partículas, consideramos os descendentes que estão na origem no tempo $2t_0$. Assim por diante, a geração n de tamanho Z_n estará composta pelos descendentes da geração $n - 1$ que estão na origem no tempo nt_0 .

Da teoria de processos de ramificação sabemos que se a média do número de filhos, $E(Z_1)$, é maior que um, como é o caso, o processo não morre com probabilidade positiva, (ver [2] Teorema 1). Isto é,

$$P(\text{sobreviver}) = P(Z_n > 0 \text{ para todo } n) = p^*, \text{ com } p^* \in (0, 1).$$

Além disso, como as partículas no processo η não morrem, o número de partículas em $\mathbb{Z}^d$ vai se incrementando, então é possível encontrar partículas em infinitos lugares diferentes de $\mathbb{Z}^d$, digamos em $z_i, i = 1, \dots, \infty$. Definamos os seguintes eventos:

$$B_i := \{\text{Um processo de Galton-Watson como acima começando em } z_i \text{ sobrevive}\},$$

$i = 1, \dots, \infty$. Estes eventos têm probabilidade p^* e são independentes, pois lembremos que as partículas movem-se de maneira independente umas das outras.

Assim, temos uma sucessão de eventos independentes com $\sum_{i=1}^{\infty} P(B_i) = \infty$, e pelo segundo lema de Borel-Cantelli sabemos que $P(\limsup B_i) = 1$.

Como

$$P(\cup_{i=1}^{\infty} B_i) \geq P(\cap_{n=1}^{\infty} \cup_{i=n}^{\infty} B_i) = P(\limsup B_i) = 1,$$

então, temos que

$$P(\text{pelo menos um dos } B_i) = P(\text{pelo menos um dos processo sobreviver}) = 1.$$

Isto mostra que o processo η visita pelo menos um destes sítios infinitas vezes com probabilidade 1, ou seja, pelo menos um sítio é recorrente. Entendemos como visita do processo η a z no tempo n , $z \in \mathbb{Z}^d$, se $\eta_n(z) \geq 1$. Em palavras, há pelo menos uma partícula no momento n em z .

Por outro lado, sabemos que um passeio aleatório simples em $\mathbb{Z}^d$ é irreduzível, ou seja, para quaisquer par de pontos x, y do espaço, uma partícula que parte de x estará em y com probabilidade positiva num número finito de passos, digamos m . Disto segue que num passeio aleatório com ramificação, uma partícula que parte de x ou qualquer um de seus descendentes chegará a y com probabilidade positiva em m' passos, com $m' \leq m$. Logo, um passeio aleatório com ramificação também é irreduzível.

Este fato implica que se um sítio é recorrente então todos os sítios também o são, ou seja, todos os sítios são visitados um número infinito de vezes com probabilidade 1. Assim, provamos a recorrência de η . $\square$

3.2 Recorrência de passeios aleatórios com ramificação restrito ao aglomerado infinito de pontos bons

Nesta seção a recorrência de passeios aleatórios com ramificação restrito, é um tipo de recorrência a qual acontece com probabilidade positiva. Antes de enunciar o próximo Lema é preciso fazer a seguinte definição:

Definição 2 Dizemos que um passeio aleatório com ramificação η é restrito a um subconjunto ζ de pontos de $\mathbb{Z}^d$, o que denotaremos por $\eta^{r,\zeta}$, se as partículas morrem quando saem do subconjunto.

Lema 4 Com probabilidade positiva, o passeio aleatório com ramificação “restrito” ao aglomerado infinito de sítios bons, visita todos seus pontos um número infinito de vezes.

Demonstração: Como vimos no Lema 2, o aglomerado infinito de sítios bons sobre k contém um aglomerado de boas configurações de tamanho maior ou igual a k' , onde k' é tal que $P(\text{um sítio seja bom sob } k') > p_c$, p_c é o ponto crítico para percolação por sítio, e $k = 2k' + 1$. Nesta prova precisaremos aumentar ainda mais o tamanho da caixa para pelo menos t_0 , em que

$$t_0 = \{\min n; E(\# \text{ de partículas de } \eta_n \text{ na origem}) \geq 1\}.$$

Escolhamos qualquer uma das caixas do tamanho maior ou igual a t_0 que formam o aglomerado de boas configurações, lembremos que uma configuração é a porção de cenário contida numa caixa. Agora começamos no centro desta caixa um processo de Galton-Watson restrito a uma caixa de tamanho t_0 como o que construímos na demonstração do Lema 3, em que o tamanho da n -ésima geração foi denotado por Z_n .

Pela construção, $E(Z_1) > 1$, logo temos que

$$P(\text{sobreviver}) = P(Z_n > 0 \text{ para todo } n) > 0,$$

em outras palavras, a probabilidade do centro desta caixa ter pelo menos uma partícula em cada tempo nt_0 , para $n \geq 1$, é maior que zero, assim, a probabilidade de que o número de visitas deste processo ao centro da caixa seja igual a infinito é maior que zero.

Além disso, como a caixa escolhida está contida no aglomerado de sítios bons, podemos encontrar um caminho dentro do aglomerado que liga o centro

da caixa com qualquer outro ponto dele, ou seja, com probabilidade positiva, uma partícula localizada no centro da caixa ou qualquer um de seus descendentes percorre este caminho sem morrer. Desta maneira, cada vez que o centro da caixa é visitado, com probabilidade positiva qualquer outro sítio do aglomerado também o será.

Portanto, como o centro é visitado infinitas vezes com probabilidade maior que zero, temos que qualquer outro ponto do aglomerado também é visitado infinitas vezes com probabilidade maior que zero. Assim, concluímos a demonstração do Lema 4. $\square$

3.3 Reconstrução do cenário nos pontos bons

A idéia do algoritmo é começar a reconstrução nos pontos do aglomerado infinito de sítios bons, já que nestes conhecemos a direção do salto de uma partícula aí localizada (ver Definição 1).

Lembremos que estes sítios são bons no modelo sob k , $k = 2k' + 1$, em que k' é tal que $P(\text{um sítio seja bom sob } k') > p_c$, e p_c o ponto crítico para percolação por sítio.

Denotemos por $\mathcal{A}_k$ o aglomerado infinito de sítios bons sob k e por $\mathcal{V}_{\mathcal{A}_k}^{(x)}$ o número de visitas a x do processo $\eta^{\tau, \mathcal{A}_k}$, $x \in \mathcal{A}_k$. Isto é, $\mathcal{V}_{\mathcal{A}_k}^{(x)}$ é o número de visitas a x do processo η restrito ao aglomerado $\mathcal{A}_k$ (ver Definição 2).

Lema 5 *Com probabilidade 1 existem infinitas subárvores infinitas de somente configurações boas, as quais contêm em seus vértices todos os sítios bons do aglomerado $\mathcal{A}_k$.*

Demonstração: Do Lema 3 sabemos que todos os sítios em $\mathbb{Z}^d$ são visitados infinitas vezes pelo processo η com probabilidade 1, logo, em particular, isto se verifica para todos os sítios do aglomerado infinito $\mathcal{A}_k$. Tomemos um ponto y_1 da fronteira de $\mathcal{A}_k$, denotemos esta por $\delta(\mathcal{A}_k)$.

Seja h_1 o momento da primeira visita do processo η a y_1 . Consideremos qualquer instante depois de h_1 em que o processo η visita um sítio não bom e observemos este daí para frente, como se começarmos a partir deste ponto um novo processo η . Vamos chamar este novo processo de $\eta^{(2)}$, então, $\eta^{(2)}$ também é recorrente e portanto visitará todos os sítios de $\mathbb{Z}^d$ infinitas vezes com probabilidade 1. Denotemos por h_2 o tempo da primeira visita neste novo processo $\eta^{(2)}$ a um ponto de $\delta(\mathcal{A}_k)$, digamos y_2 . Assim por diante, seja $\eta^{(i)}$ o processo que começa em qualquer instante depois de h_{i-1} no qual o processo η_{i-1} visita um sítio não bom, então, como $\eta^{(i)}$ também é recorrente, ele visitará todos os sítios de $\mathbb{Z}^d$ infinitas vezes com probabilidade 1. Denotamos por h_i o tempo da primeira visita neste novo processo $\eta^{(i)}$ a um ponto de $\delta(\mathcal{A}_k)$, digamos y_i .

Seja E_i o evento que o processo restrito $\eta^{r, \mathcal{A}_k}$ visita todos os pontos do aglomerado $\mathcal{A}_k$ a partir de h_i , formalmente

$$E_i := \left\{ \text{a partir de } h_i, \text{ para cada } x \in \mathcal{A}_k, \mathcal{V}_{\mathcal{A}_k}^{(x)} \geq 1 \right\}.$$

Estes eventos E_i , ($i = 1, \dots, \infty$), assim considerados são independentes e podem ser escritos simplesmente como

$$E_i := \left\{ \text{para cada } x \in \mathcal{A}_k, \mathcal{V}_{\mathcal{A}_k}^{(x)} \geq 1 \right\}.$$

Pelo Lema 4 sabemos

$$P \left(\text{para cada } x \in \mathcal{A}_k, \mathcal{V}_{\mathcal{A}_k}^{(x)} = \infty \right) > 0,$$

e como

$$\left\{ \text{para cada } x \in \mathcal{A}_k, \mathcal{V}_{\mathcal{A}_k}^{(x)} = \infty \right\} \subset \left\{ \text{para cada } x \in \mathcal{A}_k, \mathcal{V}_{\mathcal{A}_k}^{(x)} \geq 1 \right\} = E_i,$$

então

$$0 < P \left(\text{para cada } x \in \mathcal{A}_k, \mathcal{V}_{\mathcal{A}_k}^{(x)} = \infty \right) < P(E_i), \quad (3.7)$$

logo, $P(E_i)$ é positiva, digamos $p' > 0$ para todo $i = 1, \dots, \infty$.

Usando novamente o segundo Lema de Borel-Cantelli temos que

$$P(\limsup E_i) = 1,$$

ou seja,

$$P(E_i \text{ acontecer infinitas vezes}) = 1. \quad (3.8)$$

□

Observação: Note que quando o processo η chega a um ponto da fronteira de $\mathcal{A}_k$, digamos y_1 , e restringimos o processo a $\mathcal{A}_k$, então, visto desde a árvore obtida pelo processo de ramificação o que estamos fazendo é eliminar da subárvore que começa nesse ponto y_1 , aqueles ramos com vértices em sítios que não são bons, ou seja, estamos podando a árvore. Assim, qualquer realização vista na árvore de janelas de cada um dos eventos E_i , são realmente subárvores de somente configurações boas.

Portanto, ao que temos chegado é que com probabilidade 1 existem infinitas subárvores de somente configurações boas as quais contêm todos os vértices correspondentes ao aglomerado infinito de sítios bons. Com qualquer uma delas reconstruímos o cenário contido nos pontos deste aglomerado.

Como não sabemos quando estamos num destes ramos, o algoritmo para a reconstrução do cenário nos pontos bons é construído da seguinte maneira:

Cada vez que chegamos a um ponto bom, mesmo que não seja do aglomerado $\mathcal{A}_k$, (pois isso não sabemos por enquanto), começamos a reconstrução seguindo os ramos na árvore de janelas que contêm só configurações boas. Com probabilidade 1 sabemos que existem infinitas subárvores de só configurações boas, as quais contêm todos os pontos do aglomerado. Então, como estamos seguindo todas, com qualquer uma destas o cenário dentro do aglomerado fica construído.

Usando este fato obtemos o cenário nos pontos do aglomerado infinito $\mathcal{A}_k$. Assim, conseguimos a primeira parte de um algoritmo que prova o Teorema

1, isto é, temos obtido uma função mensurável, que aplicada à subárvore de janelas, nos permite obter um cenário equivalente ao cenário original ξ , mas só nos pontos bons que pertencem ao aglomerado.

3.4 Reconstrução do cenário fora do aglomerado de pontos bons

Seja w um ponto da $\delta(\mathcal{A}_k)$ e definimos por $\rho(w)$ a linha vertical semi-infinita que sai dele em direção de $-\vec{e}_d$ onde $-\vec{e}_d$ é o d -ésimo vetor canônico d -dimensional (em dimensão dois e três esta é a direção para baixo).

Lema 6 *Com probabilidade 1 para todo w em $\delta(\mathcal{A}_k)$, $\rho(w)$ contém infinitos pontos bons.*

Demonstração: Seja $(\gamma_i(w))_{i \geq 0}$ uma sequência de variáveis aleatórias Bernoulli onde $\gamma_i(w)$ pode ser escrita como $\gamma_i(w) = f(T^i(w))$ onde $T^i(w) = w + i(-\vec{e}_d)$, e

$$f(x) = \begin{cases} 1 & \text{se } x \text{ é um ponto bom} \\ 0 & \text{caso contrario} \end{cases}$$

para $i = 0, \dots, \infty$.

Como f é integrável e T é uma transformação ergódica, então pelo teorema ergódico (ver [19] Seção 1.2) temos

$$\lim_{n \rightarrow \infty} \frac{\gamma_0 + \gamma_1 + \gamma_2 + \dots + \gamma_{n-1}}{n} = E(f) \quad \text{q.c.,}$$

mas $E(f) = P(x \text{ é um ponto bom}) > p_c$, ou seja, $E(f)$ é igual a uma constante positiva, logo

$$\gamma_0 + \gamma_1 + \gamma_2 + \dots + \gamma_{n-1} \xrightarrow[n \rightarrow \infty]{} \infty.$$

Assim, provamos que há infinitos pontos bons sob a linha $\rho(w)$. □

O lema anterior nos garante que com probabilidade 1 para todo w em $\delta(\mathcal{A}_k)$ existe um w' também em $\delta(\mathcal{A}_k)$ que está justamente embaixo de w .

Lema 7 *O aglomerado infinito de pontos bons $\mathcal{A}_k$ é “unicamente determinado”, isto é, não existe translação (i.e., função $f(x) = x + \vec{a}$, com $\vec{a}$ um vetor em $\mathbb{Z}^d$) que mande o aglomerado para ele mesmo.*

Demonstração: Para provar o lema anterior basta mostrar que $\mathcal{A}_k + \vec{a} \neq \mathcal{A}_k$ para qualquer $\vec{a}$, com $\vec{a}$ um vetor em $\mathbb{Z}^d$. Aqui não vamos mais pensar no cenário dentro do aglomerado, basta só a forma dele.

Consideremos os seguintes eventos:

$$H_n = \{(n-1)\vec{a} \text{ é bom e pertence ao aglomerado, } n\vec{a} \text{ não é bom.}\}$$

É fácil mostrar que $P(H_0) > 0$ e como a sequência $\{1_{H_n}\}_{n \geq 0}$ é ergódica (ver [19] Seção 1.2), então, pelo teorema ergódico com probabilidade 1 pelo menos um dos eventos H_n ocorre. Isto mostra que $\mathcal{A}_k + \vec{a}$ é diferente de $\mathcal{A}_k$, pois $(n-1)\vec{a}$ pertence ao aglomerado mas $(n-1)\vec{a} + \vec{a} = n\vec{a}$ não.

□

O que o lema anterior nos diz é que não importa de onde começamos ver o aglomerado de pontos bons, o fato é que sempre será possível identificar estes pontos de partida em qualquer um dos aglomerados. Assim todos estes aglomerados podemse “relacionar”, (isto é, por exemplo, é possível identificar o ponto onde começamos ver o primeiro aglomerado no segundo e vice-versa).

Para algumas colorações determinísticas, isso não é possível. Considere-mos, por exemplo, a seguinte coloração:

```

0 0 0 1 1 1 0 0 0 1 1 1 0 0 0 1 1 1 0 0 0 1 1 1
0 0 0 1 1 1 0 0 0 1 1 1 0 0 0 1 1 1 0 0 0 1 1 1
0 0 0 1 1 1 0 0 0 1 1 1 0 0 0 1 1 1 0 0 0 1 1 1
1 1 1 0 0 0 1 1 1 0 0 0 1 1 1 0 0 0 1 1 1 0 0 0
1 1 1 0 0 0 1 1 1 0 0 0 1 1 1 0 0 0 1 1 1 0 0 0
1 1 1 0 0 0 1 1 1 0 0 0 1 1 1 0 0 0 1 1 1 0 0 0

```

Para esta coloração e k grande o suficiente, todos os pontos vão ser bons, mas se consideramos duas subárvores, não é possível identificar os pontos do aglomerado gerado pela primeira subárvore no aglomerado gerado pela segunda.

Denotemos por $g_i(y_i)$ a i -ésima das subárvores de somente configurações boas que começa em y_i (y_i como na demonstração do Lema 5) e por $g = \cup_{i=1} g_i(y_i)$ a floresta composta pelas subárvores $g_i(y_i)$; $i = 1, \dots, \infty$.

Como sobre os pontos bons conhecemos a direção em cada salto, então, no aglomerado gerado por $g_1(y_1)$ temos a localização relativa de cada um de seus pontos com respeito a y_1 . Além disso, como o aglomerado é unicamente determinado, então, podemos identificar y_1 em qualquer um dos aglomerados gerados por $g_i(y_i)$, para $i = 2, \dots, \infty$. Destes fatos, nós sempre saberemos diferenciar quando estamos em x ou em z em qualquer um dos aglomerados, aliás a configuração vista por uma partícula localizada em x seja igual á configuração vista por uma partícula localizada em z , para x e z quaisquer par de pontos em $\mathcal{A}_k$.

Por outro lado, como cada uma das subárvores de g contém todos os pontos bons, então, pelo menos um dos vértices de cada $g_i(y_i)$ está localizado em w , e pelo Lema 7 podemos saber quando estamos nele em qualquer uma das subárvores, para todo w em $\delta(\mathcal{A}_k)$ e todo $i = 1, \dots, \infty$.

Em cada subárvore $g_i(y_i)$ vamos localizar todos os vértices em w , e sobre os ramos que tínhamos eliminado quando restringimos o processo η a $\mathcal{A}_k$, isto é, sobre os ramos que contêm vértices localizados fora do aglomerado de pontos bons, consideremos o conjunto de todos os caminhos que ligam w e w' , com w' o ponto em $\delta(\mathcal{A}_k)$ que está justamente embaixo de w .

Lembremos que o processo η é irredutível, ou seja, todos os pontos são visitados infinitas vezes, então, é evidente que há partículas que percorrem o caminho vertical entre w e w' .

Mais uma vez, como sobre os pontos bons sabemos a direção de cada salto, e o aglomerado de pontos bons é unicamente determinado, então conseguimos conhecer sua forma, assim a distância entre w e w' também fica conhecida.

Seja $d(w, w')$ a distância entre w e w' , $\mathcal{W}_{(w, w')}$ o conjunto de todos os caminhos que ligam estes pontos, e denotemos por $T_i(w, w')$ o tempo para chegar de w a w' no i -ésimo destes caminhos.

Escolhendo dentro de $\mathcal{W}_{(w, w')}$ o caminho para o qual $T_i(w, w') = d(w, w')$, temos encontrado o caminho mais curto entre w e w' , este é o caminho vertical. Quando o processo percorre ele obtemos a porção de cenário ali contido.

Fazendo isto para todos os pontos de $\delta(\mathcal{A}_k)$ reconstruímos o cenário fora do aglomerado de pontos bons.

Assim temos conseguido reconstruir todo o cenário, provando o Teorema 1.

Referências Bibliográficas

- [1] Hart A., Matzinger H. (2003). Markers for error-corrupted observations. Preprint. Submitted.
- [2] Athreya K.B, Ney P.E. (1972). Branching Processes Springer New York.
- [3] Benjamini I., Kesten H. (1996). Distinguishing sceneries by observing the scenery along a random walk path. *J. Anal. Math.* **69** 97–135.
- [4] Howard C. D. (1997). Distinguishing certain random sceneries on $\mathbb{Z}$ via random walks. *Statist. Probab. Lett.* **34** 123–132.
- [5] Keane M., den Hollander W. Th. F. (1986). Ergodic properties of color records. *Phys. A* **138** 183–193.
- [6] Kesten H. (1996). Detecting a single defect in a scenery by observing the scenery al a random walk path. *Itô's Stochastic Calculus and Probability Theory* 171–183. Springer, Tokyo.
- [7] Lawler G. (1991). Intersections of Random Walk Birkhauser Boston.
- [8] Lember J., Matzinger H. (2003). Reconstructing a 2-color scenery by observing it along a recurrent random walk path with bounded jumps. Eurandom. In preparation.

- [9] Lenstra A., Matzinger H. (2001). Reconstructing a 4-color scenery by observing it along a recurrent random walk path with unbounded jumps. *Eurandom*.
- [10] Lindenstrauss E. (1999). Indistinguishable sceneries. *Random Structures Algorithms* **14** 71–86.
- [11] Löwe M., Matzinger H. (2002). Scenery reconstruction in two dimensions with many colors. *Ann. Appl. Probab.* **12** 1322–1347.
- [12] Löwe M., Matzinger H. (2003). Reconstruction of sceneries with correlated colors. *Stochastic Process. Appl.* **105** 175–210.
- [13] Löwe M., Matzinger H., Merkl F. (2004). Reconstructing a multicolor random scenery seen along a random wal path with bounded jumps. *Electron. J. Probab.* **9** 436–507.
- [14] Mattias L., Matzinger H. (2003). Reconstruction of scenerues with correlated colors. *Stochastic Processes and their Applications*, **105**(2) 175–210.
- [15] Matzinger H. (1999). *Reconstructing a 2-color scenery by observing it along a simple random walk path with holding*. Ph.D. thesis, Cornell University.
- [16] Matzinger H. (1999b). Reconstructing a three-color scenery by observing it along a simpl random walk path. *Random Structures Algorithms* **15** 196–207.
- [17] Matzinger H. Reconstructing a 2-color scenery by observing it along a simple random walk path. *Ann. Appl. Probab.* **15** 778–819.
- [18] Popov S., Matzinger H. (2005). Detecting a local perturbation in a continuous scenery In preparation.

- [19] Ulrich K. (1985). Ergodic Theorems Walter de Gruyter Berlin-New York.