3
Cadeias de Markov Geométricas

Neste capítulo introduzimos dois conceitos fundamentais para o nosso trabalho: cadeias de Markov com suas propriedades espectrais e núcleos geométricos. As cadeias de Markov construidas a partir dos núcleos geométricos é o que chamamos de Cadeias de Markov Geométricas.

Aqui construimos os fundamentos de uma parte importante da estrutura de dados associada à malha: adicionamos ao 1-esqueleto da malha - a união do conjunto de vértices com o conjunto de arestas - uma estrutura que é ao mesmo tempo topológica, geométrica e estocástica. A topologia é definida pelo próprio 1-esqueleto da malha e suas estruturas de incidências, e é geométrica e estocástica porque as cadeias de Markov geométricas permitem construir matrizes de transição que levam em conta a proximidade de vértices com pesos nas arestas, que são na realidade probabilidades atribuidas às conexões de vértices.

3.1
Cadeias de Markov

Definição 3.1 (Variável aleatória) *Considere um experimento e seu espaço amostral $\Omega = \{a_1, a_2, \ldots, a_n\}$. Variável aleatória é qualquer função que associa a cada elemento do espaço amostral um valor numérico mensurável (31).*

Definição 3.2 (Cadeia de Markov) *Uma cadeia de Markov é formada por uma coleção de variáveis aleatórias e uma sequência de estados, cuja probabilidade condicional de qualquer estado futuro depende apenas do estado atual (30, 31), ou seja*

$$P(X_{n+1} = x_{n+1} | \, X_k = x_k, 0 \leq k \leq n) = P(X_{n+1} = x_{n+1} | \, X_n = x_n), \quad (3\text{-}1)$$

Definição 3.3 (Cadeia de Markov irredutível) *Uma cadeia de Markov é dita irredutível se para cada par de estados x_i e x_j existe uma probabilidade* *de transição positiva entre eles (31).*

Definição 3.4 (Cadeia de Markov aperiódica) *Uma cadeia de Markov é dita aperiódica se ela é irredutível e se para algum $n \geq 0$ e algum estado x_j,*

$$P(X_n = x_j | X_0 = x_j) > 0 \; e \; P(X_{n+1} = x_j | \, X_0 = x_j) > 0. \tag{3-2}$$

Considere $P_{ij} = p(x_i, x_j)$ a probabilidade de transição para se passar do estado x_i para o x_j, com $0 \leq i \leq n$ e $0 \leq j \leq n$, em uma cadeia de Markov. O relacionamento existente entre as transições garante a seguinte propriedade:

$$\sum_{j=0}^{n} P_{ij} = 1. \tag{3-3}$$

A associação das probabilidades de transição permite agrupá-las em uma matriz, denominada matriz de transição de estados ou matriz de Markov, denotada por P. Indexados os estados, a probabilidade de transição existente entre x_i e x_j pode ser encontrada na i-ésima linha e j-ésima coluna da matriz P (31, 28).

3.2 Núcleos Geométricos

Com o propósito de criar uma cadeia de Markov irredutível e aperiódica é definida uma função núcleo responsável por atribuir pesos às relações que existem entre os estados, que são os vértices, no caso da malha. Essa função realiza a ligação entre a geometria do conjunto e a cadeia de Markov, tornando-a geométrica. Para cumprir tal objetivo ela deve apresentar as seguintes propriedades: positividade e não-negatividade. Além disso, a atribuição dos pesos deve possuir simetria, retirando a influência das direções de propagação sobre o problema (8)

Considere um conjunto X e uma função núcleo k, tal que $k : X \times X \rightarrow R$, que para todo x_i e $x_j \in X$, satisfaz:

$$k(x_i, x_j) = k(x_j, x_i), \tag{3-4}$$

$$k(x_i, x_i) > 0 \;\; e \;\; k(x_i, x_j) \geq 0 \;\; se \;\; x_i \neq x_j. \tag{3-5}$$

As probabilidades da matriz de transição surgem pela normalização dos pesos adquiridos pela função núcleo. Deste modo, $p(x_i, x_j)$ é calculada da forma

$$p(x_i, x_j) = \frac{k(x_i, x_j)}{d(x_i)}, \tag{3-6}$$

onde o fator $d(x_i)$ é um termo normalizador que garante a propriedade mostrada na equação 3-3 e é definido por

$$d(x_i) = \sum_{x_j \in X} k(x_i, x_j). \tag{3-7}$$

A matriz de transição assim construída captura a estrutura geométrica local de um elemento. A interpretação do que representa essa geometria local depende das propriedades utilizadas para a definição da função núcleo sobre o conjunto.

Existe uma distribuição estacionária associada à cadeia de Markov construída, e é dada por

$$\pi(x_i) = \frac{d(x_i)}{\sum_{x_k \in X} d(x_k)}. \tag{3-8}$$

A cadeia é reversível por construção e verifica-se que

$$\pi(x_i)\, p(x_i, x_j) = \pi(x_j)\, p(x_j, x_i). \tag{3-9}$$

A matriz de transição captura a anisotropia do conjunto utilizado para construí-la e se torna dependente da distribuição estacionária associada. Existe um operador simétrico de difusão, denotado por A, ligado à matriz P, cujos elementos, $a(x_i, x_j)$, são calculados utilizando os elementos de P e a distribuição estacionária π, da forma

$$a(x_i, x_j) = \sqrt{\frac{\pi(x_i)}{\pi(x_j)}}\, p(x_i, x_j). \tag{3-10}$$

É importante ressaltar que o operador simétrico A e a matriz P possuem os mesmos autovalores.

3.3 Propriedades Espectrais

O espectro da matriz de transição revela diversas propriedades estruturais do grafo que representa. Autovalores e autovetores são usados para compreender os diversos parâmetros que compõem o complexo relacionamento existente entre os seus vértices (3).

Existe uma decomposição espectral da matriz de transição, tal que

$$p(x_i, x_j) = \sum_{l \geq 0} \lambda_l\, \psi_l(x_i)\, \varphi_l(x_j), \tag{3-11}$$

onde $\psi_l(\cdot)$ e $\varphi_l(\cdot)$ são, respectivamente, as coordenadas da dimensão 1 das autofunções à esquerda e à direita da matriz P. Uma importante característica sobre o comportamento dos autovalores da matriz P é que (25)

$$\lambda_0 = 1 > |\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|. \tag{3-12}$$

Teorema 3.1 (Convergência Espectral) *Os autovalores λ_i de uma matriz de transição P são tais que $|\lambda_i| \leq 1$ e $\lim_{n\to\infty} \lambda_n = 0$ (24).*

A Figura 3.1 apresenta o espectro de uma matriz de transição, com valores positivos na diagonal e não-negativos nas outras posições, de dimensão 400×400, que foi normalizada. O gráfico mostra a tendência de queda dos autovalores e como uma grande quantidade deles se aproxima de zero.

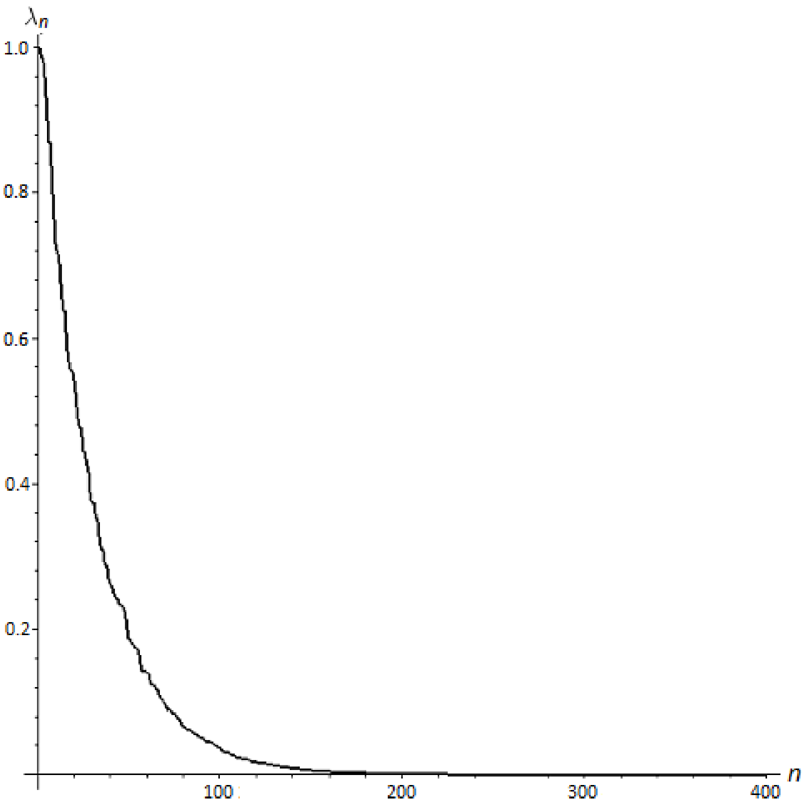

Figura 3.1: Comportamento dos autovalores de uma matriz de transição .

3.4 Algoritmo

O algoritmo 1 sintetiza os passos necessários para a construção da cadeia de Markov associada a um conjunto X. Esta cadeia é descrita como uma matriz de probabilidades de transição, P, que guarda os pesos que a função núcleo atribui aos relacionamentos existentes entre os elementos do conjunto trabalhado (2).

Algoritmo 1 Montagem da matriz de transição associada a um conjunto X.

Entrada: O conjunto X.
Saída: A matriz de transição P.

1. Escolhe-se uma função, $h : R \rightarrow R$, e define-se o núcleo em função dela, tal que $k(x_i, x_j) = h(\|x_i - x_j\|^2)$, com $h(0) > 0$ e $h(t) \geq 0$, $t > 0$.

2. Calcula-se o fator normalizador para cada um dos elementos, $d(x_i) = \sum_{x_j \in X} k(x_i, x_j)$.

3. Calcula-se as probabilidades de transição pela normalização dos pesos, $p(x_i, x_j) = \frac{k(x_i, x_j)}{d(x_i)}$.