Introdução às Cadeias Estocásticas com Memória de Alcance Variável

Caio Théodore Genovese Huss Oliveira, IMECC
Orientador: Christophe Frédéric Gallesco, IMECC

Palavras-chave: *CEMAV, Redução de Ordem, Esqueleto da Cadeia de Markov* *Financiamento: PIBIC/CNPq*

XXXIII Congresso de Iniciação Científica da Unicamp

Introdução

Cadeias Estocásticas com Memória de Alcance Variável, ou CEMAVs (VLMC em inglês), foram introduzidas em 1983 [3], como um algoritmo de compressão de dados. Diferem de Cadeias de Markov de ordem fixa pela variação do passado relevante (ou memória). O alcance variável da memória resulta em expressivos ganhos de eficiência, já que o número de passados possíveis de uma cadeia de ordem fixa cresce exponencialmente com sua ordem, além de expôr relações entre os símbolos observados e o tamanho do passado relevante. Os contextos (passados relevantes) são representados por uma árvore, onde as folhas representam o símbolo mais longe no passado, como pode ser observado na Figura 2. A *ordem* de uma CEMAV é dada pelo comprimento de seu maior contexto.

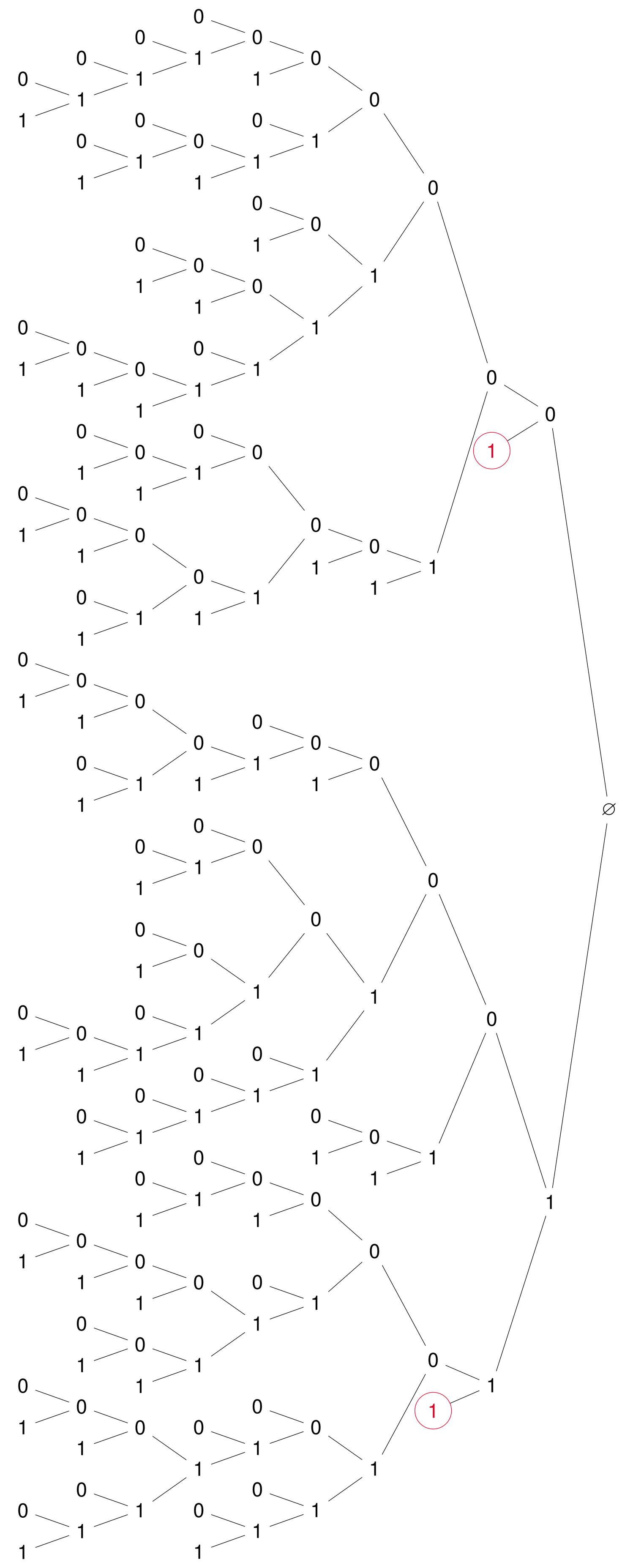

Fig. 2: Árvore de contextos da *Cadeia 1*, contextos que proíbem transições destacados.

Resultados

CEMAVs foram estudadas profundamente, pela literatura e por diversas simulações em R, dentre as quais encontramos a CEMAV da Figura 2, *Cadeia 1*, binária (assumindo apenas os símbolos 0 ou 1) de ordem 10, utilizada para o desenvolvimento da implementação dos algoritmos desenvolvidos neste projeto. Em 2025, uma redução de ordem de Cadeias de Markov foi desenvolvida [1].

O *esqueleto* é um caso específico desta redução, limitando o passado ao comprimento necessário para determinar que transições são proibidas (com probabilidade 0). O *esqueleto* frequentemente possui ordem menor, e retém as propriedades de *irredutibilidade essencial [2]* e periodicidade (para as classes recorrentes) da cadeia original [2]. Isto reduz drasticamente o custo de análise das propriedades mencionadas, que são exploradas pela *matriz de transição do esqueleto*, matriz binária de tamanho $|\mathcal{A}|^d \times |\mathcal{A}|^d$, onde $\mathcal{A}$ é o conjunto de símbolos possíveis (alfabeto), e d a ordem do *esqueleto*, cujas entradas 1 indicam uma transição com probabilidade positiva. O *esqueleto* e sua *matriz de transição* da *Cadeia 1* pode ser observado na Figura 3 e 4, respectivamente, de ordem 3.

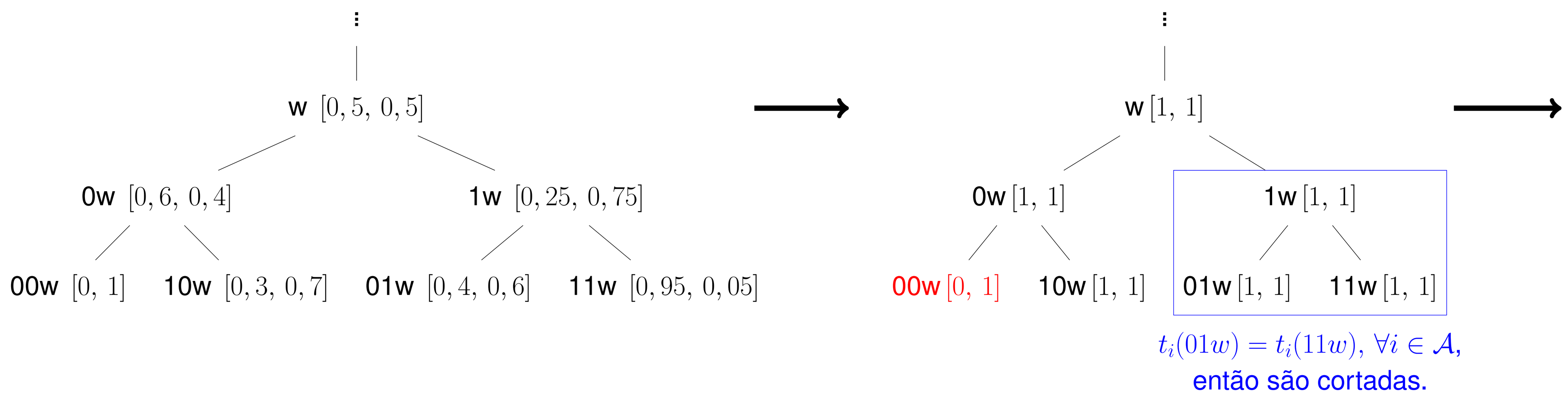

Fig. 1: Algoritmo de extração do esqueleto para uma subseção de árvore binária.

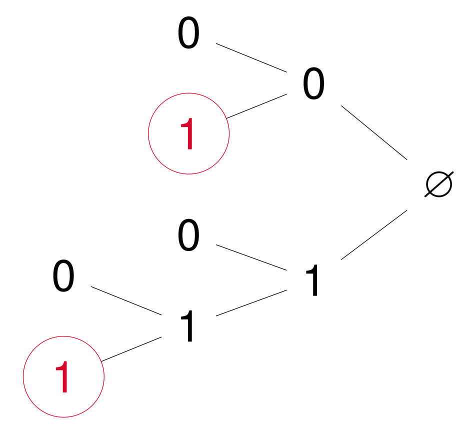

Fig. 3: Esqueleto da *Cadeia 1*, ordem 3, contextos que proíbem transições destacados.

Neste projeto, dois algoritmos foram desenvolvidos para extrair o *esqueleto* de uma Cadeia de Markov, um teórico (extração a partir de uma cadeia) e outro empírico (extração a partir de dados gerados pela cadeia). A Figura 1 descreve o processo: Após obter os vetores de probabilidades associados com cada passado possível (por estimação no caso empírico), os mesmos são transformados em *vetores de transição*, e então *folhas irmãs [2]* são cortadas caso seus vetores de transição sejam idênticos. Um dos principais resultados do projeto é que, sendo w_s^0 o contexto de *esqueleto* associado a uma sequência s, e w_s^m sua estimação pelo algoritmo empírico, após m observações da sequência s, temos a garantia da convergência do algoritmo, pelo teorema:

$$\textbf{Teorema 1: } \lim_{m\to+\infty} \mathbb{P}[w_s^m = w_s^0] = 1.$$

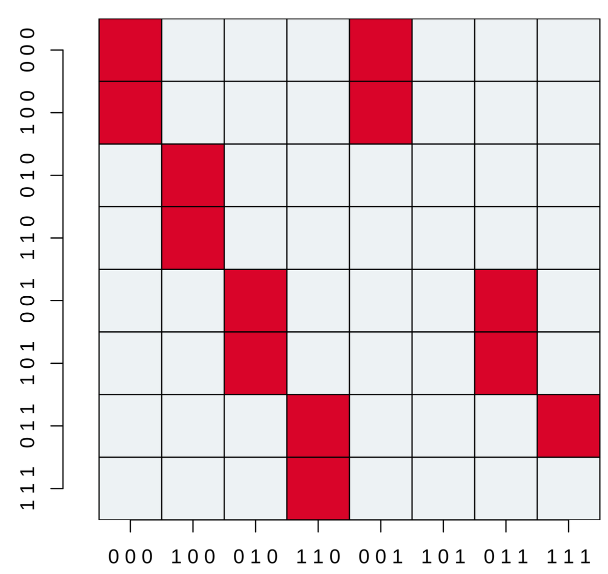

Fig. 4: Matriz de transição do esqueleto da *Cadeia 1*.

Discussão

Abordo agora o ganho de eficiência na análise de uma cadeia utilizando o esqueleto. Seja $\mathcal{A}$ o alfabeto, k a ordem da cadeia original, e d a ordem de seu *esqueleto*. Uma análise da *irredutibilidade essencial*, comparando com um método simples (produto matricial), possui os custos máximos:

- *Tradicional:* $|\mathcal{A}|^{3k}$ produtos de vetor.
- *Esqueleto:* $|\mathcal{A}|^{3d}$(análise) $+ |\mathcal{A}|^{k}$(extração de esqueleto) produtos de vetor e cortes de nodo.

A análise pela extração prévia do *esqueleto* se torna rapidamente mais eficiente com um *esqueleto* menor que a cadeia original, maioria dos casos, e, quando $d << k$, o termo $|\mathcal{A}|^{3d}$ se torna proporcionalmente desprezível, e portanto obtém-se uma redução aproximadamente **cúbica**.

Olhando para a irredutibilidade "estrita", sendo m a ordem da cadeia original, a cadeia é irredutível se, e somente se, seu esqueleto possuir ordem m ou 0 [2], este não é o caso da *Cadeia 1*.

Finalmente, a *Cadeia 1* é *essencialmente irredutível*, e possui uma única probabilidade invariante, resultado obtido com um método aproximadamente **700.000 vezes mais eficiente**.

Referências

[1] Christophe Gallesco, Alexsandro Gallo e Daniel Yasumasa Takahashi. Uniqueness of stationary compatible probability measure for chains of infinite order with forbidden transitions. 2025. arXiv: 2507.16981 [math.PR]. URL: https://arxiv.org/abs/2507.16981. *[2] Christophe Gallesco, Caio Théodore G. Huss Oliveira e Daniel Yasumasa Takahashi. «Reduction algorithm for high order Markov chains». Em desenvolvimento. 2025.* *[3] Jorma Rissanen. «A universal data compression system». Em: Information Theory, IEEE Transactions on 29 (out. de 1983), pp. 656–664. DOI: 10.1109/TIT.1983.1056741.*